



¿Qué es la regresión lineal? Una forma sencilla de entender los mínimos cuadrados

Jesús Villegas Saucillo *

Tecnológico Nacional de México / IT de
Celaya,
Celaya, Guanajuato, México

Mario Alberto Sandoval Hernández

Centro de Bachillerato Tecnológico industrial
y de servicios No. 190,
Boca del Río, Veracruz, México

Fernando Iván Molina Herrera

Tecnológico Nacional de México / IT de
Celaya,
Celaya, Guanajuato, México

Javier Díaz Carmona

Tecnológico Nacional de México / IT de
Celaya,
Celaya, Guanajuato, México

* Autor de correspondencia: jesus.villegas@itcelaya.edu.mx

Resumen: Este artículo tiene como propósito explicar de manera sencilla el método de mínimos cuadrados, una herramienta que permite encontrar una línea que se ajusta lo mejor posible a un conjunto de datos. Se presenta la deducción paso a paso de las fórmulas de regresión lineal, lo que permite entender su origen y utilidad. Además, se incluye un ejemplo práctico que ilustra cómo aplicar el método de una manera sencilla. Todo el contenido está orientado a estudiantes y lectores interesados en comprender cómo se aplica esta técnica de la estadística empleada para analizar tendencias y relaciones entre variables, sin necesidad de conocimientos avanzados.

Palabras clave: Ajuste de datos, matemática educativa, mínimos cuadrados, modelación matemática, regresión lineal.

1. Introducción: lo que debemos saber de inicio

El método de mínimos cuadrados (MMC) ha sido atribuido a matemáticos como Legendre, Gauss y Boscovich (Garzón, 2003). En 1801, tras la observación parcial del planeta Ceres por

el astrónomo G. Piazzi, Carl Gauss utilizó este método para predecir con gran precisión su órbita. A diferencia de otros cálculos, su predicción fue confirmada al redescubrirse el planeta justo en la posición anticipada (Garzón, 2003).

Actualmente, el MMC es una herramienta fundamental tanto en estadística como en análisis numérico. Se emplea en la estimación de parámetros para generar modelos que predicen fenómenos como epidemias o comportamientos físicos complejos (Englezos y Kalogerakis, 2000; Bates y Watts, 1988).

El método de mínimos cuadrados también juega un papel importante en el aprendizaje automático, también conocido como *machine learning*, un área que permite a las computadoras aprender por sí mismas a partir de datos, sin necesidad de ser programadas para cada tarea específica. En vez de seguir instrucciones fijas, estas máquinas encuentran patrones y los usan para hacer predicciones. En el estudio de Kim, Bae y Jang (2022), los investigadores usaron técnicas de aprendizaje automático, como la regresión lineal, para predecir cuánta agua se evapora del suelo y las plantas, lo que se conoce como evapotranspiración. Lo interesante es que no necesitaron todos los datos que normalmente se requieren, como la velocidad del viento o la presión del aire. En su lugar, trabajaron con información sencilla de obtener, como la temperatura, la humedad y la cantidad de sol.

En electrónica, el método MMC tiene aplicaciones importantes en el procesamiento digital de señales, una técnica que permite analizar, modificar o extraer información de señales como audio, video o datos de sensores mediante algoritmos computacionales (Kay, 1993). Además, el MMC se ha empleado para aproximar funciones especiales, que son funciones matemáticas no elementales pero comunes en física e ingeniería y que provienen de la solución de ecuaciones diferenciales, como las funciones, Gamma, error, Bessel, o las funciones elípticas. En particular, Sandoval-Hernández et al. (2018) utilizaron funciones trascendentes como seno, coseno y exponencial para aproximar de manera algebraica las integrales de Fresnel, las cuales aparecen al estudiar fenómenos de difracción, como la dispersión de la luz al pasar por una rendija, una doble rendija o por el borde de un obstáculo. Estos modelos permiten predecir patrones de interferencia luminosa. De manera más reciente, Vázquez-Leal et al. (2025) utilizaron el MMC para ampliar el dominio de validez de métodos asintóticos, los cuales permiten obtener aproximaciones de funciones cuando una variable tiende a valores extremos, es decir, valores muy grandes o valores muy pequeños. Estos métodos son muy útiles cuando no se puede encontrar una solución exacta, pero se

requiere una expresión cercana que describa el comportamiento de una función en ciertas regiones. En su trabajo, los autores aplicaron esta técnica para mejorar la precisión de funciones especiales como la función Gamma y la función elíptica de primera clase.

2. Fundamentos Teóricos: reglas y principios científicos importantes

El método de mínimos cuadrados (MMC) se usa para encontrar una función que se acerque lo mejor posible a un conjunto de datos. Lo que busca este método es que la diferencia entre los datos reales y los que predice la función sea lo más pequeña posible. Para lograrlo, se toman esas diferencias, se elevan al cuadrado y luego se suman. Al hacer esto, se da más importancia a los errores grandes y se obtiene una mejor aproximación.

Imaginemos que tenemos varios puntos con coordenadas $(x_0, y_0), (x_1, y_1), \dots, (x_i, y_i)$, y queremos encontrar una curva que pase cerca de todos ellos. Esa curva se puede escribir como una función $y = f(x_i, a_0, a_1, a_2, \dots, a_j)$, donde los $a_j, (j = 0, 1, 2, \dots, n)$ son números que necesitamos calcular. La idea es ajustar esa función para que los valores que nos da estén lo más cerca posible de los valores reales de y_i . Lo que se hace es sumar las diferencias al cuadrado entre lo que dice el modelo y lo que realmente se midió, y luego encontrar los valores de $a_0, a_1, a_2, \dots, a_j$ de la función S que hacen que esa suma sea la más pequeña posible.

$$S(a_0, a_1, a_2, \dots, a_j) = \sum_{i=1}^n \left((y_i) - f(x_i, a_0, a_1, a_2, \dots, a_j) \right)^2 \quad (1)$$

Esto representa una forma cuadrática que, cuando el modelo es lineal respecto a sus parámetros, garantiza que el error tendrá una única solución mínima. En este tipo de modelo, los parámetros aparecen de manera sencilla, sin multiplicarse entre sí ni elevarse a potencias, lo cual facilita la resolución del sistema. Sin embargo, también existen modelos no lineales, en los que los parámetros intervienen de manera más compleja algebraicamente, por ejemplo, dentro de funciones exponenciales, logarítmicas o como parte de productos entre ellos. Aun en estos casos, el objetivo sigue siendo el mismo: encontrar los valores de los parámetros que hagan que el error sea lo más pequeño posible. Para lograrlo, se utiliza una técnica matemática que consiste en calcular derivadas parciales con respecto a cada parámetro y luego igualarlas a cero. Esto genera un sistema de ecuaciones que permite encontrar los puntos donde el error no cambia, conocidos como

puntos críticos. Estos puntos pueden ser mínimos, máximos o puntos de silla. Un punto de silla es un tipo especial de punto donde la función cambia de dirección: en una dirección puede aumentar y en otra disminuir, como si uno estuviera sentado en una silla de montar, hundido en un sentido y elevado en otro. Sin embargo, en el método de mínimos cuadrados nos interesan principalmente aquellos puntos que representan el menor error posible, es decir, los mínimos.

$$\begin{aligned} \frac{\partial S}{\partial a_0} &= \frac{\partial}{\partial a_0} \left(\sum_{i=1}^n ((y_i) - f(x_i, a_0, a_1, a_2, \dots, a_j))^2 \right) = 0 \\ &\vdots \\ \frac{\partial S}{\partial a_n} &= \frac{\partial}{\partial a_n} \left(\sum_{i=1}^n ((y_i) - f(x_i, a_0, a_1, a_2, \dots, a_j))^2 \right) = 0 \end{aligned} \tag{2}$$

Sin embargo, como la expresión S está construida como una suma de cuadrados, esto garantiza que los puntos que se obtienen al resolver el sistema correspondan a mínimos globales en el caso de modelos lineales, o a mínimos locales en el caso de modelos no lineales, siempre que se cumplan ciertas condiciones de regularidad. El objetivo del método de mínimos cuadrados es, precisamente, encontrar el mejor ajuste posible reduciendo el error; por lo tanto, maximizar S iría en contra de este propósito (Bates y Watts, 1988). Cuando se trabaja con modelos lineales, S es una función estrictamente convexa, lo que asegura la existencia de un único mínimo. En modelos no lineales, aunque esta convexidad no está garantizada, el método sigue enfocado en buscar mínimos utilizando un proceso iterativo que ajusta los parámetros paso a paso. Las condiciones de segundo orden pueden confirmar si se ha encontrado un mínimo, pero en la práctica se utilizan principalmente las condiciones de primer orden, ya que son más eficientes desde el punto de vista computacional (Bates y Watts, 1988). Al resolver el sistema de ecuaciones (2) se obtienen los valores de las constantes de ajuste (a_j), lo cual permite construir una función que minimice el error entre los datos observados y el modelo. Este proceso ayuda a encontrar una curva continua que se ajuste lo mejor posible a los datos. En este enfoque, se pueden probar distintos modelos y elegir aquel que proporcione el mejor ajuste. La Figura 1 ilustra esta metodología: a partir del conjunto de datos (x_i, y_i) , donde $i = 1, 2, 3, \dots$, se determina la función que mejor representa la tendencia de los datos.

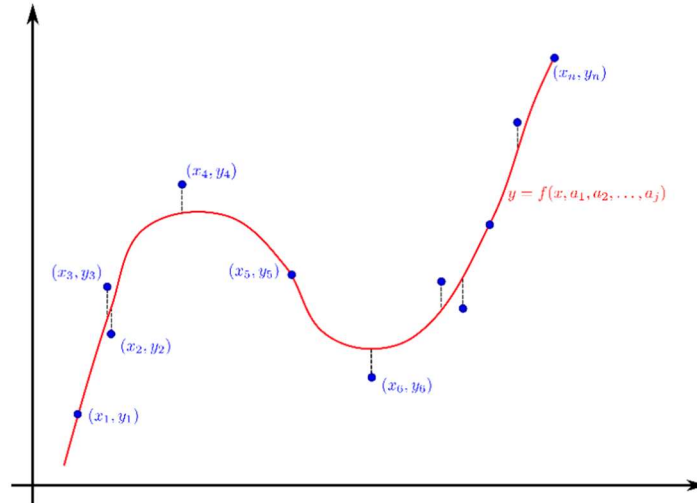


Figura 1 El método de mínimos cuadrados.

Obtenida de: elaboración propia.

3. Desarrollo del Trabajo: *aplicando las reglas y principios científicos*

Para deducir las fórmulas de regresión lineal, se considera el modelo dado por la recta $y = mx + b$. Al sustituir este modelo en la expresión (1) del error cuadrático y derivar con respecto a los parámetros $a_0, a_1, a_2, \dots, a_j$, se obtiene el sistema de ecuaciones necesario para calcular los valores óptimos de estos coeficientes.

$$\begin{aligned} \frac{\partial S(m, b)}{\partial b} &= -2 \left(\sum_{i=1}^n ((y_i) - mx - b) \right) = 0 \\ \frac{\partial S(m, b)}{\partial m} &= -2 \left(\sum_{i=1}^n ((y_i) - mx - b) x \right) = 0 \end{aligned} \quad (3)$$

El sistema de ecuaciones (3) contiene tres términos en cada una de sus ecuaciones. Al aplicar las propiedades de las sumatorias, es posible simplificar el sistema, lo que facilita el cálculo de los coeficientes de la recta de regresión.

$$\begin{aligned} \sum_{i=1}^n y_i - m \sum_{i=1}^n x_i - bn &= 0 \\ \sum_{i=1}^n y_i x_i - m \sum_{i=1}^n x_i^2 - b \sum_{i=1}^n x_i &= 0 \end{aligned} \quad (4)$$

Nótese que se han incluido los subíndices en x e y debido a la presencia de sumatorias, ya que en este paso se sustituirán los datos correspondientes. Al reescribir el sistema (4), se obtiene un sistema de ecuaciones en el que deben determinarse los valores de m y b , que corresponden a la pendiente y al punto de intersección con el eje y , respectivamente.

$$\begin{aligned} m \sum_{i=1}^n x_i + bn &= \sum_{i=1}^n y_i \\ m \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i &= \sum_{i=1}^n x_i y_i \end{aligned} \quad (5)$$

Resolviendo el sistema (5) se obtiene

$$\begin{aligned} m &= \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \\ b &= \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} = \frac{\sum_{i=1}^n y_i - m \sum_{i=1}^n x_i}{n} \end{aligned} \quad (6)$$

Es posible demostrar que la correlación entre dos variables está dada por la fórmula:

$$r = \frac{s_{xy}}{s_x s_y} \quad (7)$$

donde s_{xy} representa la covarianza, que indica cómo varían conjuntamente dos variables: si tienden a aumentar o disminuir al mismo tiempo, la covarianza será positiva, y si una aumenta mientras la otra disminuye, será negativa. Por su parte, s_x y s_y son las desviaciones estándar de x y y respectivamente, las cuales miden cuánto se alejan los datos respecto a su promedio; en otras palabras, indican qué tan dispersos están los valores de cada variable. A continuación, presentamos un caso de aplicación.

En la tabla 1 se tiene el registro de la temperatura y de la longitud de una varilla la cual se sabe que tiene un coeficiente de dilatación que es aproximadamente lineal. Obténgase la longitud de la varilla a una temperatura de 12.5 °C. Determine la temperatura si la longitud es 18.5mm.

La Tabla 1 incluye todas las sumatorias necesarias para realizar los cálculos. Por ejemplo, para obtener la suma de los cuadrados de x (i.e. x^2), se ha elevado al cuadrado cada uno de los valores de x , y el resultado se ha registrado en una columna adicional. Luego, estas sumas se utilizan para sustituir en las fórmulas correspondientes. Al insertar los valores de las

sumatorias obtenidas en la Tabla 1 en la ecuación (5), se procede a resolver el sistema y determinar los parámetros del modelo lineal.

$$m = \frac{5(1855) - 100(87)}{5(2250) - 10000} = 0.46$$

$$b = \frac{87 - 0.46(100)}{5} = 8.2$$
(8)

Tabla 1. Dilatación de una varilla y sumatorias para regresión lineal.

T [°C]	L [mm]	Datos n = 5	Sumatorias			
x	y		$\sum x$	$\sum y$	$\sum x^2$	$\sum xy$
10	13		10	13	100	130
15	15		15	15	225	225
20	17		20	17	400	340
25	20		25	20	625	500
30	22		30	22	900	660
Totales			100	87	2250	1855
		$(\sum x)^2 =$	10000			

Obtenido de: Elaboración propia.

Sustituyendo los valores obtenidos para m y b en el modelo, se tiene la ecuación $\hat{y} = 0.46x + 8.2$, donde el acento circunflejo sobre la variable y indica que se trata de una estimación. Para calcular la longitud de la varilla a una temperatura de 12.5 °C, basta con sustituir este valor en la fórmula, obteniendo así $L = 13.95\text{mm}$. Si en cambio se desea conocer la temperatura correspondiente a una longitud determinada, se debe despejar x de la ecuación, resultando en $\hat{x} = \frac{y-8.2}{0.46}$. Al sustituir el valor de la longitud, se obtiene la temperatura correspondiente en grados Celsius. La Figura 2 muestra el modelo de regresión lineal junto con los datos originales y los valores estimados.

Como era de esperarse, los valores pronosticados se encuentran sobre la línea recta, mientras que los datos reales se sitúan cerca de la recta de regresión. Cuando se dispone de un mayor número de datos, es posible que algunos de ellos coincidan exactamente con la recta, lo cual indica un ajuste preciso del modelo.

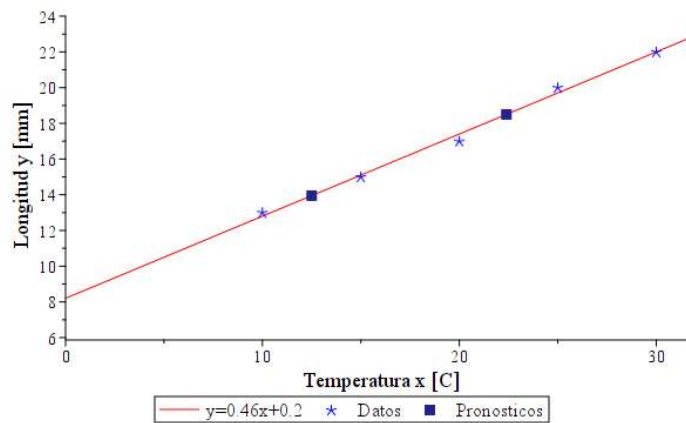


Figura 2. Recta de regresión lineal, datos y valores pronosticados.

Obtenida de: autoría propia

4. Conclusiones: *lo que podemos aprender de este artículo*

En este trabajo se ha resaltado la importancia de MMC como una herramienta fundamental en diversas áreas de la ciencia y la ingeniería. Aunque en la actualidad existen numerosos programas informáticos que permiten calcular fácilmente la regresión lineal, resulta valioso comprender el procedimiento que hay detrás de estos cálculos. Conocer la deducción de las fórmulas permite a los estudiantes desarrollar un pensamiento analítico y fortalecer su comprensión matemática. En este artículo, se presentó de manera clara y accesible la deducción de las expresiones para el caso lineal, así como un ejemplo práctico con cinco datos experimentales. Este enfoque busca facilitar la comprensión del método y destacar su utilidad como puente entre las matemáticas y su aplicación en fenómenos reales.

5. Referencias: *por si quieres seguir conociendo más*

- Bates, D. M. y Watts, D. G. (1988). *Nonlinear regression: iterative estimation and linear approximations*. Wiley.
- Englezos, P., y Kalogerakis, N. (2000). *Applied parameter estimation for chemical engineers*. CRC Press.
- Garzón, G. R. (2003). Los orígenes del método de mínimos cuadrados. *Suma: Revista sobre Enseñanza y Aprendizaje de las Matemáticas*, (43), 31–38.
- Kay, S. M. (1993). *Fundamentals of statistical signal processing: Estimation theory*. Prentice-Hall, Inc.

- Kim, S.-J., Bae, S.-J., & Jang, M.-W. (2022). Linear regression machine learning algorithms for estimating reference evapotranspiration using limited climate data. *Sustainability*, 14(18), 11674. <https://doi.org/10.3390/su141811674>
- Sandoval-Hernández, M., Vázquez-Leal, H., Hernández-Martínez, L., Filobello-Niño, U. A., Jiménez-Fernández, V. M., Herrera-May, A. L., Castañeda-Sheissa, R., Ambrosio-Lázaro, R. C., & Díaz-Arango, G. (2018). Approximation of Fresnel integrals with applications to diffraction problems. *Mathematical Problems in Engineering*, 2018(1). <https://doi.org/10.1155/2018/4031793>
- Vázquez-Leal, H., Sandoval-Hernández, M. A., Filobello-Niño, U. A., Huerta-Chua, J., Aguilar-Velázquez, R., & Domínguez-Chávez, J. A. (2025). Introducing the Leal method for the approximation of integrals with asymptotic behaviour: Special functions. *AppliedMath*, 5(1), 28. <https://doi.org/10.3390/appliedmath5010028>