Universidad Autónoma de Coahuila

FCFM

MÉTODOS NUMÉRICOS

MÍNIMOS CUADRADOS

Profesora: Maria Guadalupe Godina Cubillo

Alumno: Oscar Joel Castro Contreras

6 de junio de 2022

Resumen

En este reporte explico como se aplica el método de mínimos cuadrados en un conjunto de datos para obtener una recta que se ajuste a estos datos.

Palabras clave: Minimos Cuadrados, Recta.

Introducción

Es un procedimiento de análisis numérico en la que, dados un conjunto de puntos, se intenta determinar la función continua que mejor se aproxime a los datos, proporcionando una demostración visual de la relación entre los puntos de estos. En su forma más simple, busca minimizar la suma de cuadrados de las diferencias ordenadas entre los puntos generados por la función y los correspondientes datos. [1]
El método de mínimos cuadrados consiste es una técnica de optimización cuyo objetivo consiste en la obtención de la función que mejor se ajuste, en el sentido de un error cuadrático mínimo a los datos observados de las variables objeto de estudio. [2]
La creación del método de mínimos cuadrados generalmente se le acredita al matemático alemán Carl Friedrich Gauss, quien lo planteó en 1794 pero no lo publicó sino hasta 1809. El matemático francés Andrien-Marie Legendre fue el primero en publicarlo en 1805, este lo desarrolló de forma independiente. Este método se utiliza comúnmente para analizar una serie de datos que se obtengan de algún estudio, con el fin de expresar su comportamiento de manera lineal y así minimizar los errores de la data tomada. [1]

Metodología

El enfoque de mínimos cuadrados para este problema implica determinar la mejor línea de aproximación cuando el error relacionado es la suma de los cuadrados de las diferencias entre los valores y en la línea de aproximación y los valores y proporcionados. Por lo tanto, deben encontrarse las constantes a_0 y a_1 que minimizan el error de mínimos cuadrados:

$$E_2(a_0, a_1) = \sum_{i=1}^{10} [y_i - (a_1 x_i + a_0)]^2$$

El método de mínimos cuadrados es el procedimiento más conveniente para determinar mejores aproximaciones lineales, pero también hay consideraciones teóricas importantes que lo favorecen. En general, mientras el enfoque en los mínimos y máximos asigna demasiado peso a un bit de datos con un gran error, el método de desviación absoluta no da suficiente peso a un punto que está fuera de la línea con la aproximación. El enfoque de mínimos cuadrados asigna considerablemente más peso en un punto que está fuera de la línea que al resto de los datos, pero no permitirá que el punto domine por completo la aproximación. Una razón adicional para considerar el enfoque de mínimos cuadrados implica el estudio de la distribución estadística del error.
El problema general de ajustar la mejor línea de mínimos cuadrados para una recopilación de datos $\{(xi, yi)\}_{i=1}^{m}$ implica minimizar el error total.

$$E = E_2(a_0, a_1) = \sum_{i=1}^{m} [y_i - (a_1 x_i + a_0)]^2$$

respecto a los parámetros a_0 y a_1. Para que se presente un mínimo, necesitamos que la derivada parcial del error respecto a estos parámetros sea cero:

$$\frac{\partial E}{\partial a_0} = 0 \qquad \frac{\partial E}{\partial a_1} = 0,$$

es decir

$$0 = \frac{\partial}{\partial a_0} \sum_{i=1}^{m} [y_i - (a_1 x_i + a_0)]^2$$

y

$$0 = \frac{\partial}{\partial a_1} \sum_{i=1}^{m} [y_i - (a_1 x_i + a_0)]^2.$$

Estas ecuaciones se simplifican en las ecuaciones normales:

$$a_0 m + a_1 \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} y_i$$

y

$$a_0 \sum_{i=1}^{m} x_i + a_1 \sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} x_i y_i.$$

La solución para este sistema de ecuaciones es:

$$a_0 = \frac{\sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} x_i y_i \sum_{i=1}^{m} x_i}{m\left(\sum_{i=1}^{m} x_i^2\right) - \left(\sum_{i=1}^{m} x_i\right)^2}$$

y

$$a_1 = \frac{m \sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i}{m\left(\sum_{i=1}^{m} x_i^2\right) - \left(\sum_{i=1}^{m} x_i\right)^2}$$

con estos parámetros podemos encontrar la recta que se ajuste a los datos de la forma:

$$a_1 x + a_0$$

[3]

Resultado

Por ejemplo tenemos esta tabla de datos (x_i, y_i) de la cual podemos obtener un aproximación a una recta usando lo antes explicado:

x_i	1	2	3	4	5	6	7	8	9	10
y_i	1.3	3.5	4.2	5.0	7.0	8.8	10.1	12.5	13	15.6

Tabla 1: Datos (x_i, y_i) para aproximar a una recta .[3]

Para ayudarnos de otra tabla como la tabla 2 para calcular primero de x_iy_i y x_i^2, después sumamos las columnas para obtener las sumatoria:

	x_i	y_i	x_iy_i	x_i^2
	1	1.3	1.3	1
	2	3.5	7.0	4
	3	4.2	12.6	9
	4	5.0	20.0	16
	5	7.0	35.0	25
	6	8.8	52.8	36
	7	10.1	70.7	49
	8	12.5	100.0	64
	9	13.0	117.0	81
	10	15.6	156.0	100
$\sum =$	55	81.0	572.4	385

Tabla 2: Obtención de x_iy_i y x_i^2 y su sumatoria.[3]

Ahora solo no que da obtener los parámetros para encontrar la recta, con las formulas antes mostradas:

$$a_0 = \frac{385(81) - 55(572.4)}{10(385) - (55)^2} = -0.360$$

y

$$a_1 = \frac{10(572.4) - 55(81)}{10(385) - (55)^2} = 1.538$$

Por lo que la ecuación de la recta es:

$$1.538x - 0.360$$

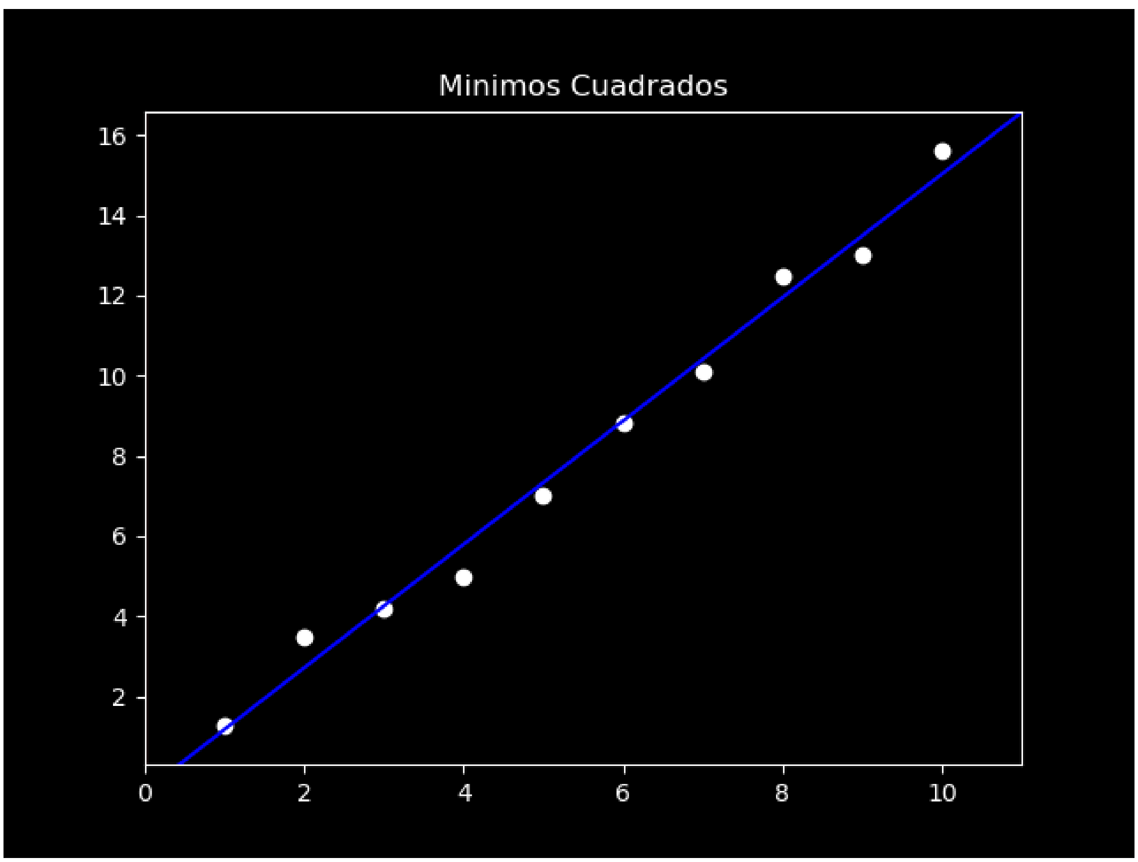
Ahora para comprobar esto, mi programa sigue el mismo procedimiento mostrado por lo que podemos calcular si meto los datos de la tabla 1 mi programa me da lo que se puede ver en la figura 1.

Lista de datos	
x	y
1.0	1.3
2.0	3.5
3.0	4.2
4.0	5.0
5.0	7.0
6.0	8.8
7.0	10.1
8.0	12.5
9.0	13.0
10.0	15.6

Recta de ajuste
1.5381818181818192*x - 0.3600000000000044

Figura 1: Función de la recta que arroja mi programa.

Mi programa también tiene la opción de graficar por lo que si le pido la grafica a mi programa me arroja lo grafica 1.



Grafica 1: Grafica de los datos y la recta que arroja mi programa.

Observación

Como ya vimos este método es muy sencillo de aplicar, a cualquier conjunto de datos, en lo que yo lo estuve probando con distritos datos no encontré ninguna restricción que impidiera crear la recta. Veo que puede ser muy útil para ver la tendencia que tiene los datos, si tienden a bajar o a subir, esta recta se puede usar mucho en estadística, ya que lo que hace es, al sumar el error de los puntos calcula la recta por la que ese error sumado es el más pequeño posible.

Conclusión

El método toma el error de los datos dados y con ellos obtiene los parámetros para poder crear una recta que cruce por la parte donde al sumar el error de cada dato, este sea el más pequeño. Podemos verlo como una recta que esta siendo jalada por el error de cada uno de los datos y cuando la recta queda en equilibrio es cuando al sumar el error de todos los datos es el mas pequeño y hay es donde se coloca la recta. Esta recta es la que se crea con el metodo de minimos cuadrados.

Referencias

[1] P. (2020, 21 abril). Mínimos cuadrados. MiProfe.com. Recuperado 30 de noviembre de 2021, de Pagina web de [1]

[2] Mínimos cuadrados. (s. f.). wolterskluwer.es. Recuperado 30 de noviembre de 2021, de Pagina web de [2]

[3] Burden, A. M., & Faires, D. J. (2017). Análisis numérico (10a. ed.). (10a ed.). Cengage Learning.