MINIMOS CUADRADOS

Para pequeños conjuntos de datos el ajuste exacto que proporciona un polinomio de interpolación es satisfactorio, pero cuando el conjunto de datos es muy grande lo mas adecuado es una función de aproximación.

El criterio de aproximación o regresión es el que traza una línea o una curva que tiene el mínimo error entre la línea y los datos, describiendo la tendencia de los datos.

El criterio de mínimos cuadrados puede ser aplicado a polinomios de bajo y alto grado.

El polinomio más simple es uno de ***forma lineal***, o sea una línea recta, y se determina mediante la función:

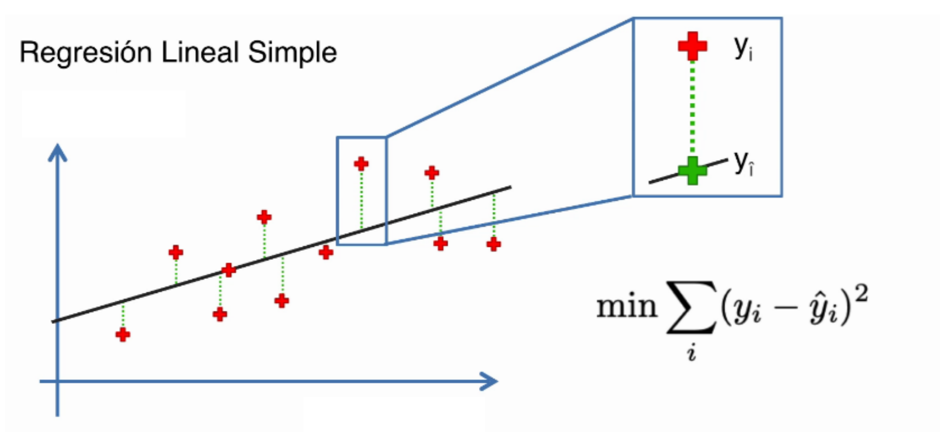
$$y = mx + b \qquad \text{ec. 1}$$

La cual debe ajustarse para N puntos (x_i, y_i), la desviación para cada valor:

$$y_i = mx_i + b \quad \text{para (i = 1,2,…N)}$$

La desviación e_i en cada valor está dada por:

$$e_i = Y_i - y_i$$



El criterio de ajuste es se toma como aquel en que la desviación cuadrática media es mínima, entonces la suma de los cuadrados de las desviaciones define a la función S(b,m):

$$S(b,m) = \sum_{i=1}^{N} (e_i)^2 = \sum_{i=1}^{N} (Y_i - b - mx_i)^2$$

El extremo de una función máximo o mínimo se obtiene cuando las derivadas de b y m sean cero, lo que genera el sistema siguiente:

$$\frac{\partial S}{\partial b} = \sum_{i=1}^{N} 2(Y_i - b - mx_i)(-1) = 0$$

$$\frac{\partial S}{\partial m} = \sum_{i=1}^{N} 2(Y_i - b - mx_i)(-x_i) = 0$$

Dividiendo entre 2, resulta:

$$bN + m\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} Y_i$$

$$b\sum_{i=1}^{N} x_i + m\sum_{i=1}^{N} x_i^2 = \sum_{i=1}^{N} x_i Y_i$$

Las ecuaciones anteriores se llaman ecuaciones normales y se deben resolver para encontrar los valores de m y b.

Otras formas equivalentes de la ecuación 1, son:

$$Y = \beta_0 + \beta_1 x \qquad\qquad y = a_0 + a_1 x$$

EJEMPLO 1: La carga eléctrica neta de proteínas depende fuertemente del pH de la solución acuosa, esto se demuestra en la siguiente tabla que contiene las velocidades de deriva electroforética observadas para un bovino a diferentes valores de pH, encontrar un ajuste lineal mediante mínimos cuadrados.

pH	4.20	4.56	5.20	5.65	6.30	7.00
Velocidad de deriva	0.50	0.18	-0.25	-0.65	-0.90	-1.25

EJEMPLO 2: Coeficiente de temperatura, donde r es la resistencia de una bobina en ohmios, ϴ es la temperatura de la bobina en grados centígrados, ajustar al modelo lineal los siguientes datos:

r	10.421	10.939	11.321	11.799	12.242	12.668
ϴ	10.5	29.49	42.70	60.01	75.51	91.05

$$\theta = a_0 + a_1 r$$

EJEMPLO 3. Determinar la aproximación por mínimos cuadrados de los datos al modelo, (modelo cuadrático):

$$C_p = a + bT + cT^2$$

T	1000	1500	2000	2500	3000
C_P	1.1427	1.2059	1.2522	1.2815	1.2938

$$y = a_0 + a_1 x + a_2 x^2$$

$$(n+1)a_0 + \sum_{k=0}^{n} a_1 + \cdots \sum_{k=0}^{n} x_k^m a_m = \sum_{k=0}^{n} y_k$$

$$\sum_{k=0}^{n} x_k^m a_0 + \sum_{k=0}^{n} x_k^2 a_1 + \cdots \sum_{k=0}^{n} x_k^{m-1} a_m = \sum_{k=0}^{n} x_k y_k$$

$$\sum_{k=0}^{n} x_k^m a_0 + \sum_{k=0}^{n} x_k^{1+l} a_1 + \cdots \sum_{k=0}^{n} x_k^{m+l} a_m = \sum_{k=0}^{n} x_k^l y_k$$

$$\sum_{k=0}^{n} x_k^m a_0 + \sum_{k=0}^{n} x_k^{1+m} a_1 + \cdots \sum_{k=0}^{n} x_k^{2m} a_m = \sum_{k=0}^{n} x_k^m a_m$$