

es razonable usar el criterio de seleccionar aquella que minimiza la suma de los cuadrados de las diferencias entre los valores y de los puntos y el valor y correspondiente a la recta. Observe que como la distancia entre (x_1, y_1) y $(x_1, b + mx_1)$ es $y_1 - (b + mx_1)$, el problema (para los n datos) puede establecerse como sigue:

Problema de mínimos cuadrados en el caso de una recta

Encuentre números m y b tales que la suma

$$[y_1 - (b + mx_1)]^2 + [y_2 - (b + mx_2)]^2 + \cdots + [y_n - (b + mx_n)]^2 \quad (6.2.1)$$

sea mínima. Para estos valores de m y b , la recta $y = b + mx$ se llama **aproximación por la recta de mínimos cuadrados a los datos** $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

Una vez definido el problema se busca un método para encontrar la aproximación de mínimos cuadrados. Lo más sencillo es escribir todo en forma matricial. Si los puntos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ están todos sobre la recta $y = b + mx$ (es decir, si son colineales), entonces se tiene

$$\begin{aligned} y_1 &= b + mx_1 \\ y_2 &= b + mx_2 \\ \vdots & \quad \quad \quad \vdots \\ y_n &= b + mx_n \end{aligned}$$

o

$$\mathbf{y} = A\mathbf{u} \quad (6.2.2)$$

donde

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \quad \text{y} \quad \mathbf{u} = \begin{pmatrix} b \\ m \end{pmatrix} \quad (6.2.3)$$

Si los puntos no son colineales, entonces $\mathbf{y} - A\mathbf{u} \neq \mathbf{0}$ y el problema se convierte en

Forma vectorial del problema de mínimos cuadrados

Encuentre un vector \mathbf{u} tal que la forma euclídeana

$$|\mathbf{y} - A\mathbf{u}| \quad (6.2.4)$$

sea mínima.

Observe que en \mathbb{R}^2 , $|(x, y)| = \sqrt{x^2 + y^2}$, en \mathbb{R}^3 , $|(x, y, z)| = \sqrt{x^2 + y^2 + z^2}$, etc. Entonces, minimizar (6.2.4) es equivalente a minimizar la suma de cuadrados en (6.2.1).

Encontrar el vector \mathbf{u} que minimiza no es tan difícil como parece. Como A es una matriz de $n \times 2$ y \mathbf{u} es una matriz de 2×1 , el vector $A\mathbf{u}$ es un vector en \mathbb{R}^n que pertenece a la imagen de A . La imagen de A es un subespacio de \mathbb{R}^n cuya dimensión es a lo más dos (ya que cuando mucho dos columnas de A son linealmente independientes). Así, por el teorema de aproximación de la norma en \mathbb{R}^n (teorema 6.1.8), (6.2.4) es un mínimo cuadrado

$$A\mathbf{u} = \text{proy}_H \mathbf{y}$$