Proyección Ortogonal sobre un Subespacio

A Calabia, Universidad de Alcala. (DOI:10.5281/zenodo.13957689)

October 20, 2024

Introducción

La proyección ortogonal de un vector sobre un subespacio es una técnica fundamental en álgebra lineal y tiene aplicaciones en diversas áreas, como la regresión lineal y el análisis de datos.

Teoría de la Proyección Ortogonal

Definición de Proyección Ortogonal

La proyección ortogonal de un vector $\mathbf{b}$ sobre un subespacio S es el vector en S que está más cerca de $\mathbf{b}$. Matemáticamente, si $\mathbf{b}$ es un vector en R^n y S es un subespacio de R^n, la proyección ortogonal de $\mathbf{b}$ sobre S se denota como $P_S(\mathbf{b})$.

Derivación de la Fórmula de Proyección Ortogonal

Queremos encontrar un vector $\mathbf{p}$ en el subespacio S tal que la distancia entre $\mathbf{b}$ y $\mathbf{p}$ sea mínima. Es decir, queremos minimizar $\|\mathbf{b} - \mathbf{p}\|$. Dado que $\mathbf{p}$ está en el subespacio generado por las columnas de A, podemos escribir $\mathbf{p}$ como $A\mathbf{c}$ para algún vector $\mathbf{c}$. Por lo tanto, queremos minimizar $\|\mathbf{b} - A\mathbf{c}\|$.

El vector de residuos $\mathbf{r} = \mathbf{b} - \mathbf{p} = \mathbf{b} - A\mathbf{c}$ es la diferencia entre el vector original y su proyección. Este vector de residuos es ortogonal a cada vector en S. Entonces, la proyección ortogonal $\mathbf{p}$ debe ser tal que el vector de residuos $\mathbf{b} - \mathbf{p}$ sea ortogonal a cada columna de A. Esto se puede expresar como:

$$A^T(\mathbf{b} - A\mathbf{c}) = 0$$

Al expandir y reorganizar, obtenemos:

$$A^T\mathbf{b} - A^T A\mathbf{c} = 0 \implies A^T A\mathbf{c} = A^T\mathbf{b}$$

Resolviendo para $\mathbf{c}$, obtenemos:

$$\mathbf{c} = (A^T A)^{-1} A^T\mathbf{b}$$

Finalmente, la proyección ortogonal de $\mathbf{b}$ sobre S es:

$$P_S(\mathbf{b}) = A\mathbf{c} = A(A^T A)^{-1} A^T\mathbf{b}$$

Ejemplo

Supongamos que queremos proyectar el vector $\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ sobre el subespacio generado por las columnas de la matriz $A = \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$.

1. Calcular $A^T A$:

$$A^T A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

2. Calcular $A^T \mathbf{b}$:

$$A^T \mathbf{b} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

3. Calcular $(A^T A)^{-1}$:

$$(A^T A)^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix}$$

4. Calcular la Proyección:

$$P_S(\mathbf{b}) = A(A^T A)^{-1} A^T \mathbf{b} = \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1/2 \\ 3 \end{pmatrix}$$

Por lo tanto, la proyección ortogonal de $\mathbf{b}$ sobre el subespacio generado por las columnas de A es $\begin{pmatrix} -1/2 \\ 1/2 \\ 3 \end{pmatrix}$.

Ejemplo

Supongamos que tenemos los siguientes puntos de datos: $(1, 2)$, $(2, 3)$, $(3, 5)$, $(4, 4)$, $(5, 6)$. Calcular la recta de regresion $y = n + mx$. Podemos escribir estos puntos en forma matricial como $\mathbf{b} = A\mathbf{c}$ donde:

$$\mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ 5 \\ 4 \\ 6 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} n \\ m \end{pmatrix}$$

Entonces:

$$\mathbf{c} = (A^T A)^{-1} A^T \mathbf{b} = \begin{pmatrix} 5 & 15 \\ 15 & 55 \end{pmatrix}^{-1} \begin{pmatrix} 20 \\ 70 \end{pmatrix} = \begin{pmatrix} 1.3 \\ 0.9 \end{pmatrix}$$

Esto nos da la línea de mejor ajuste:

$$y = 1.3 + 0.9x$$

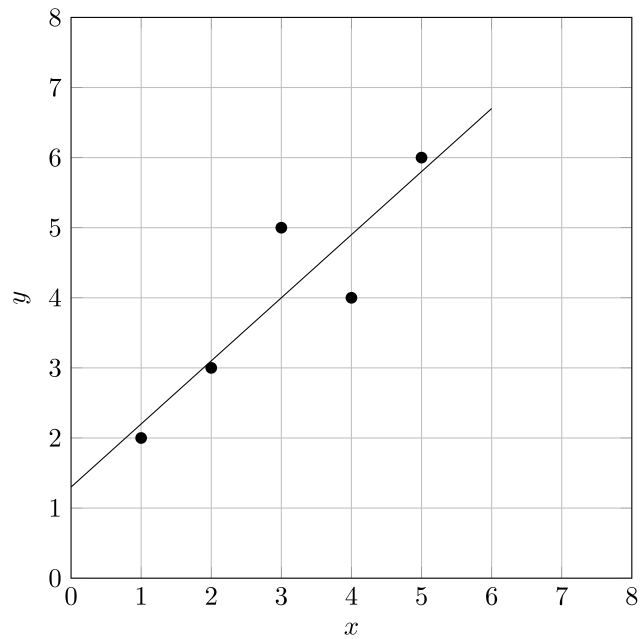

Figure 1: Gráfico de los puntos de datos y la línea de mejor ajuste