Geometría y Álgebra Lineal 2

Mauro Polenta Mora

CLASE 15 - 11/06/2025

Mínimos cuadrados

Idea

La idea es resolver un problema de apróximación a partir de datos dados de por ejemplo un experimiento. Los datos que tenemos son pares conformados por: - Una variable independiente. - Una variable de respuesta.

Los mismos se ven de la siguiente forma: - $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$

La idea es modelar algo de la forma $y = \alpha x + \beta$ con $\alpha, \beta \in \mathbb{R}$, de manera que α, β se adapten de la mejor forma posible a los datos que nos fueron dados para el problema.

Resumiendo, buscamos la recta que mejor se adapte (con un criterio que veremos posteriormente) a los datos que recibimos.

Definición (error)

Dada una recta $y = \alpha x + \beta$ definimos el error ε de la siguiente forma:

$$\varepsilon^2 = \sum_{i=1}^{n} \varepsilon_i^2 \text{con } \varepsilon_i = y_i - (\alpha x_i + \beta)$$

Donde: - y_i es el valor observado. - x_i es el valor sobre la recta.

Criterio de ajuste: Calcular α y β que minimizan el error cuadrático.

Enfoque geométrico

Dada la recta $y = \alpha x + \beta$ y las m observaciones (x_i, y_i) para $i = 1, \dots, m$, construimos las siguientes matrices y vectores:

- $Y = (y_1, \dots, y_m)^t \in \mathbb{R}^m$
- $A = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \in \mathcal{M}_{m \times 2}$
- $Z = (\alpha, \beta)^t$

Con esto, podemos tener un vector que contiene todos los errores ε_i:

$$\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{pmatrix} = \begin{pmatrix} y_1 - (\alpha x_1 + \beta) \\ y_2 - (\alpha x_2 + \beta) \\ \vdots \\ y_m - (\alpha x_m + \beta) \end{pmatrix} = Y - AZ$$

Recordemos que lo que queremos es calcular Z que minimiza el error ε^2 que definimos arriba, pero considerando la norma inducida por el producto interno estándar en $\mathbb{R}^m$, tenemos que:

$$\varepsilon^2 = \|(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m)^t\| = \sum_{i=1}^{n} \varepsilon_i^2$$

O en resumen, calcular Z que minimiza la siguiente expresión:

$$\|Y - AZ\|^2$$

Observemos que:

$$AZ = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} + \beta \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \alpha col_1(A) + \beta col_2(A)$$

Por lo que entonces, $AZ \in S$ subespacio de $\mathbb{R}^m$ generado por las columnas de A.

Por lo tanto, dado $Y \in \mathbb{R}^m$ queremos encontrar Z' tal que $\|Y - AZ'\|^2$ sea mínimo. Recordando el concepto de proyección ortogonal, queremos que $AZ' = P_S(Y)$

Lema

Sea $A \in \mathcal{M}_{m\times n}$. Consideremos S subespacio generado por las columnas de A (por lo tanto $S \subset \mathbb{R}^m$). Consideramos $\mathbb{R}^m$ con el producto interno usual. Entonces:

$$S^\perp = \{X \in \mathbb{R}^m : A^t X = \vec{0}\}$$

Demostración

Veamos el siguiente razonamiento

$$\begin{aligned} & X \in S^\perp \iff \\ & X \text{ es ortogonal a todas las columnas de } A \iff \\ & X \text{ es ortogonal a todas las filas de } A^t \iff \\ & A^t X = \vec{0} \end{aligned}$$

El último paso se deriva de que considerando la fila i-ésima de A^t, tenemos que:

$$a_{i1}x_1 + a_{i2}x_2 + ... + a_{im}x_m = \langle col_i(A), X \rangle = 0$$

El resultado se obtiene de que el resultado es el mismo para todas las columnas.

Continuación (enfoque geométrico)

Partimos de la base que $AZ' = P_S(Y)$, podemos reescribir esto como $AZ' = Y - P_{S^\perp}(Y)$. Veamos el siguiente razonamiento:

$$\begin{aligned}
&AZ' = P_S(Y) \\
&\quad \iff \text{(propiedades de la proyección ortogonal)} \\
&AZ' = Y - P_{S^\perp}(Y) \\
&\quad \iff \text{(multiplicando por izquierda por } A^t\text{)} \\
&A^tAZ' = A^tY - A^tP_{S^\perp}(Y) \\
&\quad \iff \text{(por el lema anterior)} \\
&A^tAZ' = A^tY
\end{aligned}$$

Definmos estas ecuaciones como **ecuaciones normales**. Si A tiene rango máximo, se cumple que A^tA es una matriz invertible, y en ese caso podemos obtener directamente la solución como:

$$Z' = (A^tA)^{-1}A^tY$$

Observación: Notemos que:

$$A^tA = \begin{pmatrix} x_1 & x_2 & ... & x_m \\ 1 & 1 & ... & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & m \end{pmatrix}$$

Veamos también lo siguiente:

$$A^tY = \begin{pmatrix} x_1 & x_2 & ... & x_m \\ 1 & 1 & ... & 1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m x_iy_i \\ \sum_{i=1}^m y_i \end{pmatrix}$$

Entonces la solución es la siguiente:

$$Z' = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & m \end{pmatrix}^{-1} \cdot \begin{pmatrix} \sum_{i=1}^m x_iy_i \\ \sum_{i=1}^m y_i \end{pmatrix}$$

Ejemplo

Sean los siguientes pares de datos:

- $(-6,-1),(-2,2),(1,1),(7,6)$

Calculamos las sumatorias que necesitamos para A^tA y A^tY:

- $\sum_{i=1}^4 x_i^2 = 90$
- $\sum_{i=1}^4 x_i = 0$
- $m = 4$
- $\sum_{i=1}^4 x_iy_i = 6-4+1+42 = 45$
- $\sum_{i=1}^4 y_i = 8$

Entonces:

$$Z' = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 90 & 0 \\ 0 & 4 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 45 \\ 8 \end{pmatrix} = \begin{pmatrix} \frac{1}{90} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \cdot \begin{pmatrix} 45 \\ 8 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 2 \end{pmatrix}$$

Entonces, la recta que mejor se ajusta a los datos es $y = \frac{1}{2}x + 2$

Observación: Este método también funciona para polinomios de grado mayor o igual a 2.