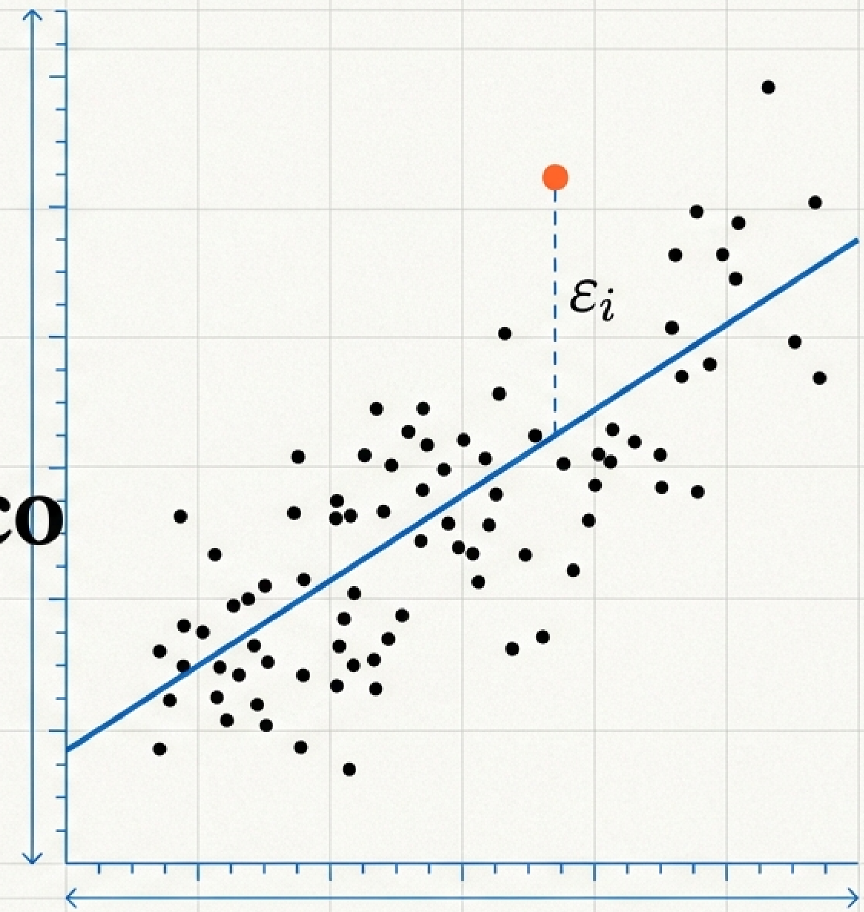
ESTADÍSTICA COMPUTACIONAL AVANZADA

Editorial Mathematics

Modelos de Regresión Simple: Estimación y Diagnóstico

De la intuición geométrica a la formulación matricial.



El Objetivo de la Predicción

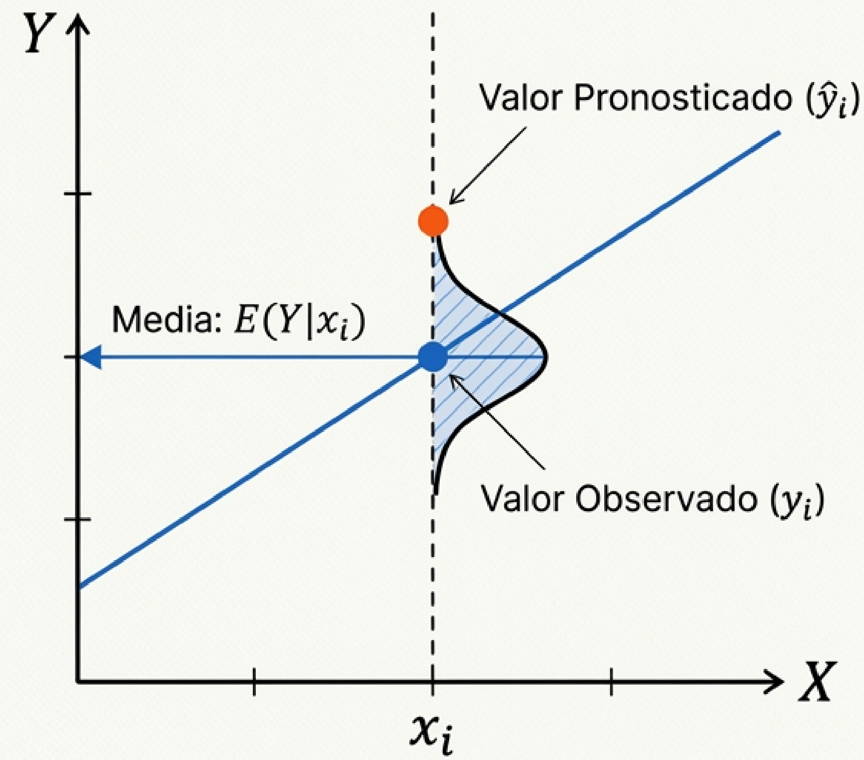
Buscamos un estimador puntual de la media de Y para un x fijo.

La ecuación de predicción:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

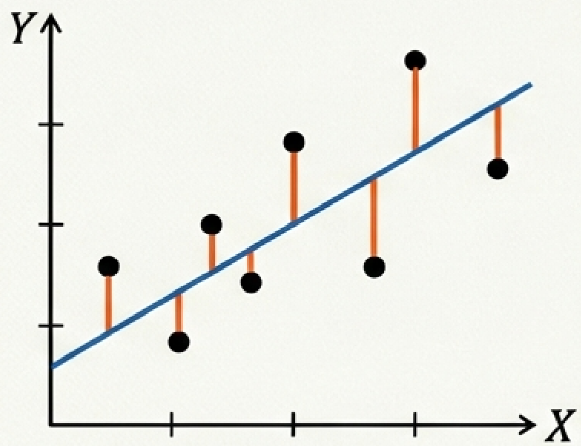
Forma centrada:

$$\hat{y}_i = \bar{y} + \hat{\beta}_1 (x_i - \bar{x})$$



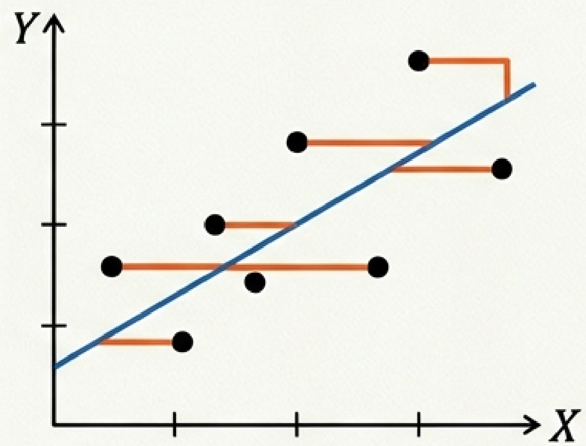
Geometría del Error: ¿Qué distancia minimizamos?

Distancia Vertical
(Regresión Clásica)



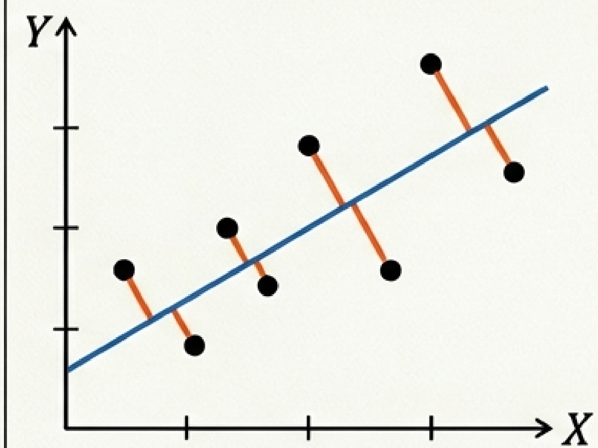
Minimiza error en Y.

Distancia Horizontal
(Regresión Inversa)



Minimiza error en X.

Distancia Perpendicular
(Ortogonal)



Componentes Principales.

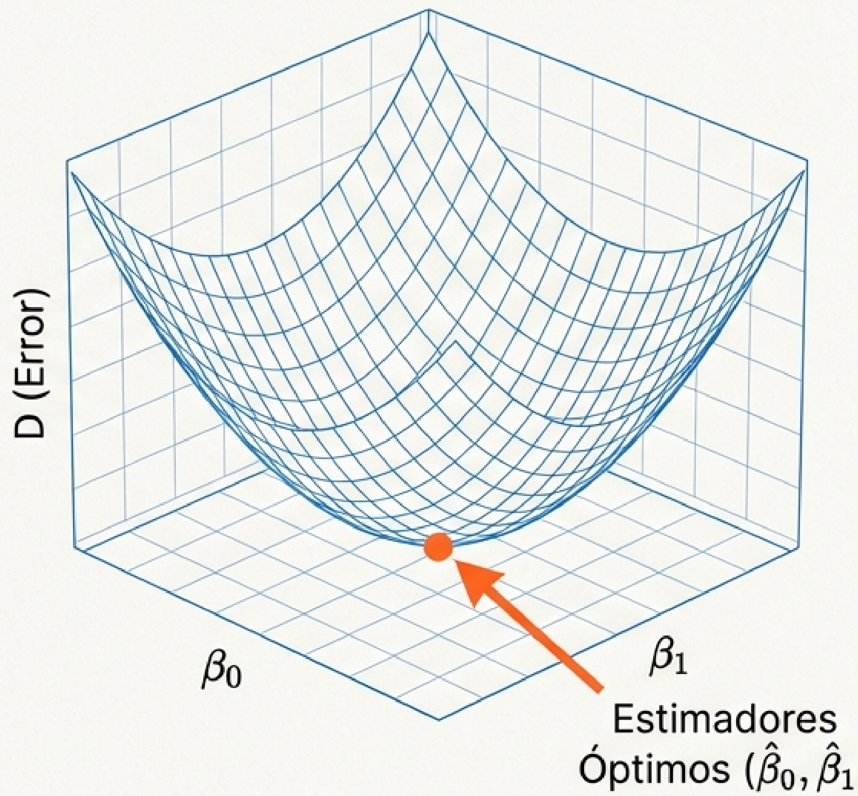
Para estimar la media de Y dado x, el método de Mínimos Cuadrados Ordinarios (MCO) utiliza exclusivamente la **Distancia Vertical**.

El Método de Mínimos Cuadrados Ordinarios (MCO)

Función a minimizar (D):

$$D(\beta_0, \beta_1) = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

Elevamos al cuadrado para penalizar las grandes desviaciones y evitar que los errores positivos y negativos se cancelen.



La Derivación de los Estimadores

1. Derivadas Parciales = 0

$$\frac{\partial D}{\partial \beta_0} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\frac{\partial D}{\partial \beta_1} = -2\sum_{i=1}^{n}x_i(y_i - \beta_0 - \beta_1 x_i) = 0$$

La matriz Hessiana es definida positiva, asegurando un mínimo global.

2. Ecuaciones Normales

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum x_i = \sum y_i$$

$$\hat{\beta}_0 \sum x_i + \hat{\beta}_1 \sum x_i^2 = \sum x_i y_i$$

3. Solución Escalar

$$\boxed{\hat{\beta}_1 = \frac{\mathrm{Cov}(x, y)}{\mathrm{Var}(x)}}$$

$$\boxed{\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}}$$

NotebookLM

Traducción al Lenguaje Matricial

Anatomy of the Equation

$$Y = X\beta + \epsilon$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$n \times 1$
(Respuesta)

Intercepto →

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$$

$n \times 2$
(Matriz de Diseño)

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

2×1
(Parámetros)

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

$n \times 1$
(Errores)

La Solución Elegante

Minimizar la suma de cuadrados en forma matricial:

$$D(\beta) = (Y - X\beta)'(Y - X\beta)$$

Derivando respecto a β:

$$-2X'Y + 2X'X\beta = 0$$

Ecuaciones Normales:

$$X'X\beta = X'Y$$

$$\boxed{\hat{\beta} = (X'X)^{-1}X'Y}$$

Condición: La matriz (***X′X***) debe ser no singular (invertible).

Coeficiente de Determinación (R^2)

Variabilidad Total (SCT)

$$SCT = SCM + SCE$$

Variabilidad Explicada (SCM)

Ruido / Error (SCE)

$$R^2 = \frac{SCM}{SCT} = 1 - \frac{SCE}{SCT}$$

Proporción de la variabilidad de Y explicada por el modelo lineal. $0 \leq R^2 \leq 1$.

Editorial Mathematics

R^2 Ajustado vs. Estándar

R^2 Estándar

El problema: Nunca disminuye al agregar variables, aunque sean irrelevantes.

R^2 Ajustado

La solución: Penaliza la complejidad (número de parámetros p).

$$R^2_{adj} = 1 - \frac{SCE/(n-p)}{SCT/(n-1)}$$

Estatura Hijo-Padre
R^2 = **0.9415**

Estatura Hijo-Padre
R^2_{adj} = **0.9332**

El ajuste corrige los grados de libertad. Si el modelo es parsimonioso, ambos valores deben ser cercanos.

Análisis de los Residuos: La Huella del Modelo

Residuo: $e_i = y_i - \hat{y}_i$

Suma Cero

$$\rightarrow \sum e_i = 0$$

Ortogonalidad

$$\rightarrow \sum e_i x_i = 0$$

Key Metric (Varianza Residual)

$$\hat{\sigma}_R^2 = \frac{\sum e_i^2}{n - 2}$$

Grados de libertad (perdemos **2** al estimar β_0 y β_1).