Estimadores de Mínimos Cuadrados

Regresión Lineal con Dos Regresores

Demostración Matemática

9 de junio de 2025

Contenido

Modelo de Regresión Lineal con Dos Regresores

Modelo Poblacional

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$$

donde:

- Y_i: variable dependiente
- X_{1i}, X_{2i}: variables independientes (regresores)
- $\beta_0, \beta_1, \beta_2$: parámetros poblacionales
- u_i: término de error
- $i = 1, 2, \ldots, n$ observaciones

Supuestos del Modelo

1. **Linealidad**: El modelo es lineal en parámetros
2. **Muestreo aleatorio**: (X_{1i}, X_{2i}, Y_i) es una muestra aleatoria
3. **No colinealidad perfecta**: X_1 y X_2 no son perfectamente colineales
4. **Media condicional cero**: $E[u_i|X_{1i}, X_{2i}] = 0$
5. **Homocedasticidad**: $\text{Var}(u_i|X_{1i}, X_{2i}) = \sigma^2$

Función Objetivo

Suma de Cuadrados de Residuos

Queremos minimizar:

$$SSR(\beta_0, \beta_1, \beta_2) = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2$$

Condiciones de Primer Orden

Derivamos respecto a cada parámetro e igualamos a cero:

$$\frac{\partial SSR}{\partial \beta_0} = 0 \tag{1}$$

$$\frac{\partial SSR}{\partial \beta_1} = 0 \tag{2}$$

$$\frac{\partial SSR}{\partial \beta_2} = 0 \tag{3}$$

Derivación - Paso 1

Primera Derivada respecto a β_0

$$\frac{\partial SSR}{\partial \beta_0} = \frac{\partial}{\partial \beta_0}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2 \tag{4}$$

$$= \sum_{i=1}^{n} 2(Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})(-1) \tag{5}$$

$$= -2\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}) \tag{6}$$

Igualando a cero:

$$\sum_{i=1}^{n} Y_i = n\beta_0 + \beta_1 \sum_{i=1}^{n} X_{1i} + \beta_2 \sum_{i=1}^{n} X_{2i}$$

Derivación - Paso 2

Primera Derivada respecto a β_1

$$\frac{\partial SSR}{\partial \beta_1} = -2\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})X_{1i} \tag{7}$$

Igualando a cero:

$$\sum_{i=1}^{n} X_{1i} Y_i = \beta_0 \sum_{i=1}^{n} X_{1i} + \beta_1 \sum_{i=1}^{n} X_{1i}^2 + \beta_2 \sum_{i=1}^{n} X_{1i} X_{2i}$$

Derivación - Paso 3

Primera Derivada respecto a β_2

$$\frac{\partial SSR}{\partial \beta_2} = -2\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})X_{2i} \tag{8}$$

Igualando a cero:

$$\sum_{i=1}^{n} X_{2i}Y_i = \beta_0 \sum_{i=1}^{n} X_{2i} + \beta_1 \sum_{i=1}^{n} X_{1i}X_{2i} + \beta_2 \sum_{i=1}^{n} X_{2i}^2$$

Ecuaciones Normales

Sistema de 3 ecuaciones con 3 incógnitas

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum X_{1i} + \hat{\beta}_2 \sum X_{2i} = \sum Y_i \tag{9}$$

$$\hat{\beta}_0 \sum X_{1i} + \hat{\beta}_1 \sum X_{1i}^2 + \hat{\beta}_2 \sum X_{1i}X_{2i} = \sum X_{1i}Y_i \tag{10}$$

$$\hat{\beta}_0 \sum X_{2i} + \hat{\beta}_1 \sum X_{1i}X_{2i} + \hat{\beta}_2 \sum X_{2i}^2 = \sum X_{2i}Y_i \tag{11}$$

donde $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ son los estimadores de mínimos cuadrados.

Definamos:

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \qquad \bar{X}_1 = \frac{1}{n}\sum_{i=1}^{n} X_{1i} \qquad \bar{X}_2 = \frac{1}{n}\sum \tag{12}$$

$$S_{11} = \sum_{i=1}^{n}(X_{1i} - \bar{X}_1)^2 \qquad S_{22} = \sum_{i=1}^{n}(X_{2i} - \bar{X}_2)^2 \tag{13}$$

$$S_{12} = \sum_{i=1}^{n}(X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) \tag{14}$$

$$S_{1Y} = \sum_{i=1}^{n}(X_{1i} - \bar{X}_1)(Y_i - \bar{Y}) \qquad S_{2Y} = \sum_{i=1}^{n}(X_{2i} - \bar{X}_2)(Y_i - \bar{Y}) \tag{15}$$

Estimador de $\hat{\beta}_0$

De la primera ecuación normal:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}_1 - \hat{\beta}_2\bar{X}_2$$

Interpretación

El intercepto se ajusta automáticamente para que la línea de regresión pase por el punto de medias $(\bar{X}_1, \bar{X}_2, \bar{Y})$.

Estimadores de las Pendientes

Sustituyendo $\hat{\beta}_0$ en las ecuaciones 2 y 3, y simplificando:

Fórmulas de los Estimadores

$$\hat{\beta}_1 = \frac{S_{1Y}S_{22} - S_{2Y}S_{12}}{S_{11}S_{22} - S_{12}^2} \tag{16}$$

$$\hat{\beta}_2 = \frac{S_{2Y}S_{11} - S_{1Y}S_{12}}{S_{11}S_{22} - S_{12}^2} \tag{17}$$

El denominador $S_{11}S_{22} - S_{12}^2 > 0$ cuando no hay colinealidad perfecta.

Interpretación Geométrica

Caso Sin Correlación ($S_{12} = 0$)

Si X_1 y X_2 no están correlacionadas:

$$\hat{\beta}_1 = \frac{S_{1Y}}{S_{11}} = \frac{\sum(X_{1i} - \bar{X}_1)(Y_i - \bar{Y})}{\sum(X_{1i} - \bar{X}_1)^2} \tag{18}$$

$$\hat{\beta}_2 = \frac{S_{2Y}}{S_{22}} = \frac{\sum(X_{2i} - \bar{X}_2)(Y_i - \bar{Y})}{\sum(X_{2i} - \bar{X}_2)^2} \tag{19}$$

Estos son exactamente los estimadores de regresión simple de Y sobre X_1 y Y sobre X_2 respectivamente.

Caso General con Correlación

Cuando X_1 y X_2 están correlacionadas ($S_{12} \neq 0$):

- $\hat{\beta}_1$ mide el efecto de X_1 sobre Y **manteniendo constante** X_2
- $\hat{\beta}_2$ mide el efecto de X_2 sobre Y **manteniendo constante** X_1
- Los estimadores se ajustan por la correlación entre regresores
- Si $S_{12} > 0$ (correlación positiva), los efectos se "descuentan" mutuamente

Propiedades Estadísticas

Insesgadez

Bajo los supuestos del modelo:

$$E[\hat{\beta}_j] = \beta_j \quad \text{para } j = 0, 1, 2$$

Varianzas

$$\text{Var}(\hat{\beta}_1) = \frac{\sigma^2 S_{22}}{S_{11}S_{22} - S_{12}^2} \tag{20}$$

$$\text{Var}(\hat{\beta}_2) = \frac{\sigma^2 S_{11}}{S_{11}S_{22} - S_{12}^2} \tag{21}$$

$$\text{Var}(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{X}_1^2 S_{22} + \bar{X}_2^2 S_{11} - 2\bar{X}_1\bar{X}_2 S_{12}}{S_{11}S_{22} - S_{12}^2} \right] \tag{22}$$

Teorema de Gauss-Markov

Mejor Estimador Lineal Insesgado (BLUE)

Bajo los supuestos clásicos, los estimadores de mínimos cuadrados son:

- **Lineales**: en Y_i
- **Insesgados**: $E[\hat{\beta}_j] = \beta_j$
- **Eficientes**: tienen la menor varianza entre todos los estimadores lineales insesgados

Normalidad Asintótica

Con muestras grandes:

$$\hat{\beta}_j \sim N\left(\beta_j, \mathrm{Var}(\hat{\beta}_j)\right)$$

Representación Matricial

El modelo se puede escribir como:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

donde:

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & X_{11} & X_{21} \\ 1 & X_{12} & X_{22} \\ \vdots & \vdots & \vdots \\ 1 & X_{1n} & X_{2n} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

Estimador Matricial

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

¿Por qué necesitamos $(\mathbf{X}'\mathbf{X})^{-1}$?

Del sistema de ecuaciones normales:

$$\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{Y}$$

Objetivo: Despejar $\hat{\beta}$

Para aislar $\hat{\beta}$, multiplicamos ambos lados por $(\mathbf{X}'\mathbf{X})^{-1}$:

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

Propiedad de la Matriz Inversa

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} = \mathbf{I}$$

donde $\mathbf{I}$ es la matriz identidad.

Por lo tanto:

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

Cálculo de $(\mathbf{X}'\mathbf{X})^{-1}$

Para el modelo con dos regresores:

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} n & \sum X_{1i} & \sum X_{2i} \\ \sum X_{1i} & \sum X_{1i}^2 & \sum X_{1i}X_{2i} \\ \sum X_{2i} & \sum X_{1i}X_{2i} & \sum X_{2i}^2 \end{pmatrix}$$

$$\mathbf{X}'\mathbf{Y} = \begin{pmatrix} \sum Y_i \\ \sum X_{1i} Y_i \\ \sum X_{2i} Y_i \end{pmatrix}$$

Inversión de la Matriz 3×3

Para una matriz 3×3:

$$\mathbf{A} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

El determinante es:

$$\det(\mathbf{A}) = a(ei - fh) - b(di - fg) + c(dh - eg)$$

La matriz inversa es:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{pmatrix}$$

Aplicación al Caso Específico

Sustituyendo en $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$:

Relación con Fórmulas Anteriores

Al realizar la multiplicación matricial, recuperamos exactamente:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}_1 - \hat{\beta}_2\bar{X}_2 \tag{23}$$

$$\hat{\beta}_1 = \frac{S_{1Y}S_{22} - S_{2Y}S_{12}}{S_{11}S_{22} - S_{12}^2} \tag{24}$$

$$\hat{\beta}_2 = \frac{S_{2Y}S_{11} - S_{1Y}S_{12}}{S_{11}S_{22} - S_{12}^2} \tag{25}$$

Ventaja: La forma matricial se generaliza fácilmente a k regresores.

Conclusiones

- Los estimadores de mínimos cuadrados minimizan la suma de cuadrados de residuos
- Se obtienen resolviendo un sistema de ecuaciones normales
- Cuando los regresores están correlacionados, los efectos se ajustan mutuamente
- Los estimadores son BLUE bajo los supuestos clásicos
- La interpretación es de efectos parciales (ceteris paribus)

Extensión

Este método se generaliza fácilmente a k regresores, manteniendo la misma lógica matemática.

¡Gracias por su atención!

¿Preguntas?