UNIVERSIDAD NACIONAL DE INGENIERIA

Facultad de Ingeniería Económica, Estadistíca y CC.SS.

Curso: Modelos Lineales

Modelos Lineal General

Eduardo Marcos Sánchez

Modelos

- Tenemos el modelo con K variables

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_K X_{Ki} + u_i,\ \ i = 1, \ldots n$$

- Y vimos que usando matrices puede escribirse como

$$Y = X\beta + u$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{K1} \\ 1 & X_{22} & X_{32} & \cdots & X_{K2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{Kn} \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

donde X es $(n \times K)$, Y y u son $(n \times 1)$, y β es $(K \times 1)$.

Definiciones

- Vector de estimadores $\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_K \end{bmatrix}$
- Vector de estimaciones de Y $(n \times 1)$: $\hat{Y} \equiv X\hat{\beta}$
- Vector de residuos o errores de estimación $(n \times 1)$:
  $e \equiv Y - \hat{Y} = Y - X\hat{\beta}$

Definiciones

- $$X'X = \begin{bmatrix} n & \sum X_{2i} & \sum X_{3i} & \cdots & \sum X_{Ki} \\ \sum X_{2i} & \sum X_{2i}^2 & \sum X_{2i}X_{3i} & \cdots & \sum X_{2i}X_{Ki} \\ \sum X_{3i} & \sum X_{2i}X_{3i} & \sum X_{3i}^2 & \cdots & \sum X_{3i}X_{Ki} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum X_{Ki} & \sum X_{2i}X_{Ki} & \sum X_{3i}X_{Ki} & \cdots & \sum X_{Ki}^2 \end{bmatrix}$$

- La existencia de $(X'X)^{-1}$ se garantiza por el supuesto de no multicolinealidad perfecta.
- Recordemos: si no hay multicolinealidad perfecta $\rho(X) = K$, lo que a su vez implica que $\rho(X'X) = K \Rightarrow \exists (X'X)^{-1}$

Supuestos

1. $E(u) = \begin{bmatrix} E(u_1) \\ E(u_2) \\ \vdots \\ E(u_n) \end{bmatrix} = \mathbf{0}$

2. $VarCov(u) = \begin{bmatrix} V[u_1] & Cov[u_1, u_2] & \cdots & Cov[u_1, u_n] \\ Cov[u_2, u_1] & V[u_2] & \cdots & Cov[u_2, u_n] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[u_n, u_1] & Cov[u_n, u_2] & \cdots & V[u_n] \end{bmatrix} = \sigma^2 \mathbf{I_n}$

3. X_{nxK} no estocástica con $\rho(X) = K$

Supuestos

- Insesgadez: $E[\hat{\beta}] = \beta$
- Matriz de varianzas y covarianzas: $VarCov(\hat{\beta}) = \sigma^2 (X'X)^{-1}$

$$= \begin{bmatrix} V[\hat{\beta}_1] & Cov[\hat{\beta}_1, \hat{\beta}_2] & \cdots & Cov[\hat{\beta}_1, \hat{\beta}_K] \\ Cov[\hat{\beta}_1, \hat{\beta}_2] & V[\hat{\beta}_2] & \cdots & Cov[\hat{\beta}_2, \hat{\beta}_K] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[\hat{\beta}_1, \hat{\beta}_K] & Cov[\hat{\beta}_2, \hat{\beta}_K] & \cdots & V[\hat{\beta}_K] \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^2 A_{11} & \sigma^2 A_{12} & \cdots & \sigma^2 A_{1K} \\ \sigma^2 A_{21} & \sigma^2 A_{22} & \cdots & \sigma^2 A_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2 A_{K1} & \sigma^2 A_{K2} & \cdots & \sigma^2 A_{KK} \end{bmatrix}$$

donde A_{kk} es el elemento en la fila k y columna k de la matriz $(X'X)^{-1}$

- El estimador de $VarCov(\hat{\beta})$ es $Va\hat{r}Cov(\hat{\beta}) = S^2 (X'X)^{-1}$ con $S^2 = \frac{e'e}{n-K}$

Supuestos

Definición: matriz semidefinida positiva

Sea H una matriz $m \times m$ y sea c cualquier vector de m constantes. Se dice que H es semidefinida positiva si y sólo si $c'Hc \geq 0 \; \forall c \in \mathbb{R}^m$

- Notar:
  - que $c'Hc$ es (1×1), o sea, un número real.
  - $c'Hc$ es una **forma cuadrática.** Para ver por qué, desarrollar la expresión para $m = 2$.
- **Importante**: cualquier matriz de varianzas y covarianzas es semidefinida positiva.
  - En particular, $VarCov[\hat{\beta}]$ es semidefinida positiva.

Supuestos

- Entonces, para cualquier vector de constantes c de dimensión $K \times 1$ se cumple que $c' VarCov[\hat{\beta}]c \geq 0$.
- La expresión $c' VarCov[\hat{\beta}]c$ es la varianza de la combinación lineal $c'\hat{\beta}$.
- Veamos el ejempo para $K = 2$:
    - $c'\hat{\beta} = c_1\hat{\beta}_1 + c_2\hat{\beta}_2$
    - $V[c'\hat{\beta}] = c' VarCov[\hat{\beta}]c = c_1^2 V[\hat{\beta}_1] + c_2^2 V[\hat{\beta}_2] + 2c_1c_2 Cov[\hat{\beta}_1, \hat{\beta}_2]$
    - La semidefinición positiva garantiza que esta última expresión es ≥ 0.
    - Entonces todos los siguientes casos particulares son ≥ 0:
        - Si $c_1 = 1$ y $c_2 = 0 \Rightarrow V[c'\hat{\beta}] = V[\hat{\beta}_1] \geq 0$
        - Si $c_1 = 0$ y $c_2 = 1 \Rightarrow V[c'\hat{\beta}] = V[\hat{\beta}_2] \geq 0$
        - Si $c_1 = 1$ y $c_2 = 1 \Rightarrow V[c'\hat{\beta}] = V[\hat{\beta}_1] + V[\hat{\beta}_2] + 2Cov[\hat{\beta}_1, \hat{\beta}_2] \geq 0$
        - Si $c_1 = 1$ y $c_2 = -1 \Rightarrow V[c'\hat{\beta}] = V[\hat{\beta}_1] + V[\hat{\beta}_2] - 2Cov[\hat{\beta}_1, \hat{\beta}_2] \geq 0$

Teorema de Gauss Markov

- Bajo todos los supuestos clásicos, el estimador de MCO es el más eficiente de todos los estimadores lineales e insesgados.
- **Mejor Estimador Lineal e Insesgado (MELI).**
- **Además**: sea c un vector de K constantes arbitrarias $c_1, c_2, \ldots c_K$, $c'\hat{\beta}$ es el mejor estimador lineal e insesgado de $c'\beta$.
  - Es decir, la combinación lineal de los estimadores MCO es MELI para estimar la combinación lineal de los parámetros.

MLC: $\quad \epsilon \sim N(\emptyset, \sigma^2 I)$

i) $y = x\beta + \epsilon$

$\bullet)$ $\hat{\beta} \sim N_{p+1}(\beta ; \sigma^2 (x'x)^{-1})$

Dem: $\hat{\beta} = (x'x)^{-1} x'y$

como $y = x\beta + \epsilon$ y $\epsilon \sim N_n(\emptyset, \sigma^2 I_n)$

$\Rightarrow y \sim N_n(x\beta, \sigma^2 I_n)$

y $\Rightarrow \hat{\beta}$ es una Norm. multivar.

$\bullet)$ $\tilde{\sigma}^2 \sim \dfrac{\sigma^2}{n-p-1} \chi^2_{(n-p-1)}$

Dem: $\tilde{\sigma}^2 = \dfrac{\hat{\epsilon}'\hat{\epsilon}}{n-p-1} = \dfrac{\epsilon' M_x \epsilon}{n-p-1} = \epsilon' A \epsilon$

$\Rightarrow A\Sigma = \dfrac{M_x \sigma^2}{n-p-1} = \dfrac{\sigma^2}{n-p-1} M_x$

$A\Sigma$ es multiplo de una matriz M_x Idempotente

y $c = \dfrac{\sigma^2}{n-p-1}$

$\Rightarrow \tilde{\sigma}^2 \sim \dfrac{\sigma^2}{n-p-1} \chi^2_{(n-p-1)}$

• $\dfrac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma^2} \sim \chi^2_{(n-p-1)}$

• $\hat{\beta}_{MCO}$ y $\hat{\sigma}^2_{MCO}$ son indep.

Dem: $\hat{\beta}_{MCO} = (X'X)^{-1}X'Y \quad \ldots (1)$

$\hat{\sigma}^2_{MCO} : \dfrac{\hat{\varepsilon}'\varepsilon}{n-p-1} = \dfrac{y'M_x y}{n-p-1} = \dfrac{1}{n-p-1}(M_x y)'(M_x y)$

$Z = M_x y$

$\Rightarrow \hat{\sigma}^2_{MCO} = \dfrac{1}{n-p-1} \times Z'Z$

$Cov(\hat{\beta}_{MCO}; Z) = Cov\left((X'X)^{-1}X'y \,;\, M_x y\right)$

$= (X'X)^{-1}X'(\sigma^2 I_n) M_x \Rightarrow \sigma^2 (X'X)^{-1} \underbrace{X'M_x}_{0}$

$= 0$

Z y $\hat{\beta}_{MCO}$ son indep.

$\hat{\sigma}^2 = \dfrac{Z'Z}{n-p-1} = \dfrac{\hat{\varepsilon}'\varepsilon}{n-p-1}$ y $\hat{\beta}_{MCO}$

EMV

•) $\hat{\beta}_{MV} = \hat{\beta}_{MCO}$ $\qquad Y \sim N(X\beta ; \sigma^2 I_n)$

•) $\hat{\sigma}^2_{MV} = \dfrac{(Y - X\hat{\beta}_{MV})'(Y - X\hat{\beta}_{MV})}{n}$

Modelo Reg Lineal Multiple

$Z = \begin{bmatrix} Y \\ \hline X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} ; \quad E(Z) = \begin{bmatrix} \mu_Y \\ \hline \mu_X \end{bmatrix} ; \quad V(Z) = \left[\begin{array}{c|c} \sigma^2_Y & \sigma_{10} \\ \hline \sigma_{01} & \Sigma_X \end{array}\right]$

•) $Z \sim N_{p+1}(\mu_Z ; \Sigma_Z)$

•) $\alpha = (\alpha_1, \alpha_2 \dots \alpha_p)' = \underset{\alpha \neq 0}{\arg\max}\, f(Y, \alpha' X)$

$Y_i = \beta_0 + \beta_1 X_{i1} + \beta X_{i2} + \dots + \beta_p X_{ip} + e_i$

$$\bullet)\ y_j = \underbrace{\beta_0 + \beta_1 \bar{x}_1 + \beta_2 \bar{x}_2 + \dots + \beta_p \bar{x}_p}_{\alpha} + \beta_1 (x_{j1} - \bar{x}_1) + \beta_2 (x_{j2} - \bar{x}_2) + \dots + \beta_p (x_{jp} - \bar{x}_p) + e_j$$

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{Y} = \underbrace{\begin{pmatrix} 1 & x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & \dots & x_{1p} - \bar{x}_p \\ 1 & \vdots & \vdots & \dots & \vdots \\ \vdots & & & & \\ 1 & & & & x_{np} - \bar{x}_p \end{pmatrix}}_{X^*} \underbrace{\begin{pmatrix} \alpha \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}}_{\beta^*} + \underbrace{\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}}_{\mathcal{E}}$$

$$Y = X^* \beta^* + \mathcal{E}$$

$$\operatorname{argmax} \{ \rho (y, \alpha' x) \}$$

$$E(Y|x) = \mu_y + \sigma_{10} \Sigma^{-1} (x - \mu_x) = \underbrace{\mu_y - \sigma_{10} \Sigma^{-1} \mu}_{\beta_0} + \underbrace{\sigma_{10} \Sigma^{-1}}_{\alpha' \to \beta_1} x = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

- Sea el modelo Y=XB+e con termino independiente y "β" variables explicativa. Si consideramos un conjunto de variables explicativas alternativo Z=XP, donde X es la matriz de datos original y P es no singular. Pruebe que el ajuste del modelo se mantiene. (suger. Considerar $R^2 = SCReg/SCTotal$

Modelo 1: $y = X\beta + \varepsilon$

$SCReg = y'\left(H_x - \frac{1}{n}J_n\right)y$

$SC_T = y'\left(I_n - \frac{1}{n}J_n\right)y$

$\Rightarrow R^2 = \dfrac{y'\left(H_x - \frac{1}{n}J_n\right)y}{y'\left(I_n - \frac{1}{n}J_n\right)y}$

Modelo 2: $y = Z\beta^* + \varepsilon^*$

$SCReg_{(2)} = y'\left(H_z - \frac{1}{n}J_n\right)y = y'\left(Z(Z'Z)^{-1}Z' - \frac{1}{n}J_n\right)y$

$SCReg_{(2)} = y'\left(H_x - \frac{1}{n}J_n\right)y$

$R^2 = \dfrac{y\left(H_x - \frac{1}{n}J_n\right)y}{y'\left(I_n - \frac{1}{n}J_n\right)y}$

=

$\longrightarrow$ mantiene

- Se tiene un modelo de regresión conformado por p variables explicativas, luego se decide incorporar una variable independiente adicional. Analizar matemáticamente que efecto ocasiona la inclusión de esta nueva variable en el R^2.

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

$$\text{Mod 1}: \; y = X\beta + \varepsilon \Rightarrow SCE = \hat{\varepsilon}'\hat{\varepsilon}$$

$$\text{Mod 2}: \; \underset{n\times 1}{y} = \left(\underset{n\times p}{X} \;\vdots\; \underset{n\times 1}{z}\right)\underset{p+1}{\begin{pmatrix}\beta \\ \alpha\end{pmatrix}} + \varepsilon^*$$

$$\Rightarrow y = X\beta + \alpha z + \varepsilon'$$

$$SCE_n = \underbrace{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*}_{esc} = (y - X\hat{\beta} - \hat{\alpha} z)'(y - X\hat{\beta} - \hat{\alpha} z)$$

$$= \underbrace{(y - X\hat{\beta})'(y - X\hat{\beta})}_{\hat{\varepsilon}} - 2(y - X\hat{\beta})'\hat{\alpha} z + \hat{\alpha}^2 z'z$$

$$\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^* = \hat{\varepsilon}'\hat{\varepsilon} - 2\hat{\alpha}(y - X\hat{\beta})'z + \hat{\alpha}^2 z'z \; \ldots (I)$$

$$\text{Para } z'\hat{\varepsilon}^* = 0 \rightarrow z'(y - X\hat{\beta} - \hat{\alpha} z) = 0$$

$$= z'(y - X\hat{\beta}) = \hat{\alpha} z'z$$

$$\Rightarrow (y - X\hat{\beta})z' = \hat{\alpha} z'z \; \ldots (II)$$

(II) en (I)

$$\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^* = \hat{\varepsilon}'\hat{\varepsilon} - 2\hat{\alpha} z'z + \hat{\alpha}^2 z'z$$

$$\Rightarrow \hat{\varepsilon}^{*\prime}\hat{\varepsilon}^* - \hat{\varepsilon}'\hat{\varepsilon} = -\hat{\alpha}^2 z'z \leq 0$$

$$\underbrace{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*}_{SCE_{(2)}} \leq \hat{\varepsilon}'\hat{\varepsilon}$$

$$SCE_{(2)} \leq SCE_{(1)} \Rightarrow R^2_{(2)} \geq R^2_{(1)}$$