Introduccion a los Modelos de Regresion

Walter Sosa Escudero

(wsosa@udesa.edu.ar)

Universidad de San Andres

Referencias

- Hayashi (2000) Capitulo 1, pp. 3-46.
- Cualquier texto basico de econometria (con matrices!!!)

Introduccion

Modelo lineal:

$$y_i = \beta_1 + \beta_2 x_{2i} + \cdots + \beta_K x_{Ki} + u_i, \qquad i = 1, \ldots, N$$

$$Y = X\beta + u$$

Y, vector de n of observaciones
X, matriz $n \times K$, primera columa con unos.
u, vector de n terminos de error.

- Cuantas variables explicativas hay en este modelo?
- El objetivo es estimar β.
- En los cursos basicos X es tratada como fija. Es extraño en una disciplina no experimental.

Esperanzas condicionales

$$E(Y|X=x) = \int y \;\; f_{Y|X} dy$$

Idea: como cambia la esperanza de Y cuando X cambia. Es una *funcion* que depende de X. Si X es una variable aleatoria, $E(Y|X)$ es una variable aleatoria.

Propiedades:

- $E(X|X) = X$
- $Y = a + bX + U$, entonces $E(Y|X) = a + bX + E(U|X)$.
- *Ley de esperanzas iteradas:* $E(Y) = E\left[E(Y|X)\right]$.

Los Supuestos Clasicos

- Exogeneidad estricta: $E(u_i|X) = 0, \qquad i = 1, 2, \ldots, n.$
  Implica dos cosas:
  - $E(u_i) = 0$. Por LIT, $E(u_i) = E(E(u|X)) = E(0) = 0$.
  - $E(x_{jk} u_i) = 0$ (idem anterior, v/como ejercicio).
- No multicolinealidad: $\rho(X) = K$.
- Varianza esferica
  - *Homocedasticidad:* $V(u_i|X) = E(u_i^2|X) = \sigma^2, \ \forall i.$
  - *No correlacion serial:*
    $Cov(u_i, u_j|X) = E(u_i u_j|X) = 0, \ i \neq j.$

  En terminos matriciales: $V(u|X) = E(uu'|X) = \sigma^2 I.$

Definamos:

$$\hat{Y} \equiv X\hat{\beta}, \qquad e \equiv Y - \hat{Y} \;=\; Y - X\hat{\beta}$$

Estimador minimo-cuadratico:

$$\hat{\beta} = \text{argmin}_b \sum_{i=1}^{n} e_i^2 = \text{argmin}_b \sum_{i=1}^{n} e'e$$

Solucion:

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

Estimation of σ^2: usamos $S^2 = e'e/(n-K)$.

Interpretaciones

$$E(y_i|x_i) = \beta_1 + \beta_2 x_{2i} + \cdots + \beta_{Ki} x_{Ki}$$

- Si las variables x_{ik} son funcionalmente independientes:

$$\beta_k = \frac{\partial E(y_i|x_i)}{\partial x_k}$$

- Si x_{ki} es una variable binaria (1 si i pertenece a una clase, 0 si no)

$$\beta_k = E(y|x_k = 1) - E(y|x_k = 0)$$

- Dos modelos 'no lineales':
  - Cuadratico:
$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + u_i$$
  - Interactivo:
$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i z_i + u_i$$

Como se interpretan los coeficientes en estos casos?

Descomposicion de suma de cuadrados y R^2

Si el modeo tiene una constante:

$$\underbrace{\sum (Y_i - \bar{Y})^2}_{\text{SCT}} = \underbrace{\sum (\hat{Y}_i - \bar{Y})^2}_{\text{SCE}} + \underbrace{\sum e_i^2}_{\text{SCR}}$$

Medida de bondad del ajuste:

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

Propiedades de muestra finita

1. Linealidad: $\hat{\beta}$ es *lineal*: $\exists A_{K\times N}$ tal que $\hat{\beta} = AY$.
Prueba: Tomar $A = (X'X)^{-1}X'$.

2. Insesgadez: $E(\hat{\beta}) = \beta$.

$$\begin{aligned}
\hat{\beta} &= (X'X)^{-1}X'Y \\
&= (X'X)^{-1}X'(X\beta + u) \\
&= \beta + (X'X)^{-1}X'u \\
E(\hat{\beta}) &= \beta + E\left[(X'X)^{-1}X'u\right] \\
&= \beta + E\left[(X'X)^{-1}X'E\left(u|X\right)\right] \\
&\quad \text{(Ley de esperanzas iteradas)} \\
&= \beta \qquad \text{(Dado que } E(u|X) = 0\text{)}
\end{aligned}$$

3. Varianza: $V(\hat{\beta}) = \sigma^2 (X'X)^{-1}$.

$$\begin{aligned} V(\hat{\beta}) &= E[(\beta - E(\hat{\beta}))(\beta - E(\hat{\beta}))'] \\ &= E[(\beta - \beta)(\beta - \beta)'] \quad \text{(porque?)} \\ &= E[(X'X)^{-1} X'uu'X(X'X)^{-1}] \quad \text{(prueba anterior...)} \\ &= \quad \text{(a partir de aca los dejo solos...)} \end{aligned}$$

4. Teorema de Gauss-Markov: Bajo todos los supuestos clasicos, para cualquier estimador lineal e insesgado $\tilde{\beta}$,

$$V(\tilde{\beta}) - V(\hat{\beta}) \geq 0$$

Generalmente se dice que el estimador MCO es el mejor lineal insesgado.

Prueba (X no aleatoria):

Que $\tilde{\beta}$ sea *lineal* implica que existe una matriz no estocastica $A_{K\times n}$ de rango K, tal que $\tilde{\beta} = AY$. Bajo los supuestos clasicos

$$E(\tilde{\beta}) = A\ E(Y) = A(X\beta + E(u)) = AX\beta \quad (1)$$

Que $\tilde{\beta}$ sea *insesgado* implica que

$$E(\tilde{\beta}) = \beta \quad (2)$$

Consecuentemente, que $\tilde{\beta}$ sea *lineal e insesgado* implica que (1) y (2) tienen que cumplirse simultaneamente, lo cual ocurre si y solo si $AX = I$.

Trivialmente:

$$\tilde{\beta} = \hat{\beta} + \tilde{\beta} - \hat{\beta} \equiv \hat{\beta} + \hat{\gamma} \tag{3}$$

con $\hat{\gamma} \equiv \tilde{\beta} - \hat{\beta}$. Notar que:

$$V(\tilde{\beta}) = V(\hat{\beta}) + V(\hat{\gamma}) \tag{4}$$

si $Cov(\hat{\beta}, \hat{\gamma} = 0)$. Esto es lo que probaremos. Notar que

$$\begin{aligned}
\hat{\gamma} &= AY - (X'X)^{-1}X'Y \\
&= (A - (X'X)^{-1}X')Y \\
&= (A - (X'X)^{-1}X')(X\beta + u) \\
&= (A - (X'X)^{-1}X')u \quad \text{(prueba anterior...)}
\end{aligned}$$

ya que $AX = X'A' = I$ por el resultado anterior, de lo que obtenemos

$$E(\hat{\gamma}) = 0$$

Adicionalmente:

$$
\begin{aligned}
Cov(\hat{\beta}, \hat{\gamma}) &= E[(\hat{\beta} - \beta)\hat{\gamma}'] \\
&= E[(X'X)^{-1}X')uu'(A - (X'X)^{-1}X')'] \\
&= \sigma^2[(X'X)^{-1}X'(A' - X(X'X)^{-1})] \\
&= \sigma^2[(X'X)^{-1}X'A' - (X'X)^{-1}X'X(X'X)^{-1})] \\
&= 0
\end{aligned}
$$

donde hemos utilizado que $V(u) = E(uu') = \sigma^2$, y, nuevamente, que $AX = I$.
Entonces, despejando en (4):

$$
V(\tilde{\beta}) - V(\hat{\beta}) = V(\hat{\gamma})
$$

que por definicion es semidefinida positiva, lo cual concluye la prueba. ∎

5. (Insesgadez de S^2): $E(S^2) = \sigma^2$.

Sugiere que podemos estimar $V(\hat{\beta}) = \sigma^2 (X'X)^{-1}$ usando:

$$\hat{V}(\hat{\beta}) = S^2 (X'X)^{-1}$$

Supuesto adicional: (Normalidad) $u|X \sim N(0, \sigma^2 I)$

- Hipotesis individuales: $H_o : \beta_k = \beta_{k0}$ vs. $H_A : \beta_k \neq \beta_{k0}$.
  Resultado: bajo H_0:

$$z_k \equiv \frac{\hat{\beta}_k - \beta_{k0}}{\sqrt{\hat{V}(\hat{\beta}_k)}} \sim t_{n-K}.$$

- Significatividad global: $H_0 : \beta_2 = 0\ ,\ \beta_3 = 0\ ,\ \ldots,\ \beta_K = 0$
  vs. $H_A : \beta_2 \neq 0 \ \vee\ \beta_3 \neq 0 \ \vee\ ,\ldots,\ \vee\ \beta_K \neq 0.$
  Resultado: bajo H_0:

$$F = \frac{ESS/(K-1)}{RSS/(n-K)} \sim F(K-1, n-K)$$