Econometria I

O estimador de Mínimos Quadrados Ordinários

Luis A. F. Alvarez

24 de março de 2025

Ambiente

- Pesquisador observa n pares (Y_i, X_i), $i = 1 \ldots, n$, para os quais supõe um modelo linear da forma:

$$Y_i = X_i'\beta + \epsilon_i \ \ldots i = 1, \ldots n \,, \tag{1}$$

onde $\beta \in \mathbb{R}^k$ é um parâmetro desconhecido, e ϵ_i, $i = 1, \ldots, n$ são variáveis aleatórias não observadas.

  - Recorde-se, da última aula, que cabe ao pesquisador postular o modelo e a interpretação dos coeficientes.
  - No caso mais comum, $(Y_i, X_i) \overset{d}{=} (Y, X)$ para $i = 1, \ldots, n$, onde a distribuição de (Y, X) representa a distribuição das variáveis numa população de interesse.
    - Por exemplo, podemos ter que $\{(Y_i, X_i)\}_{i=1}^n$ é uma amostra aleatória de uma população com distribuição $\mathbb{P}_{Y,X}$, para a qual postulamos um modelo linear.
    - Mas também podemos ter que as observações entre pares apresentem dependência entre si, embora com leis $\mathbb{P}_{(Y_i, X_i)}$ comuns a todo i.
  - De modo mais geral, no entanto, pode ser que os (Y_i, X_i) não possuam a mesma distribuição conjunta, mas haja uma relação comum e estável ao longo de i.

Notação matricial

- No que segue, definimos as seguintes matrizes aleatórias:

$$\boldsymbol{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} X_1' \\ X_2' \\ \vdots \\ X_n' \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

- Com base na notação acima introduzida, podemos reescrever (1) em notação matricial como:

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{\epsilon} \qquad (2)$$

Estimador de mínimos quadrados ordinários

- O estimador de mínimos quadrados ordinários de β, denotado por $\hat{b}$, consiste em estimar β minimizando a distância, na norma Euclidiana, entre $\boldsymbol{y}$ e uma combinação linear das colunas de $\boldsymbol{X}$, i.e.

$$\hat{b} \in \text{argmin}_{b \in \mathbb{R}^k} \|\boldsymbol{y} - \boldsymbol{X} b\|_2^2 = \text{argmin}_{b \in \mathbb{R}^k} \frac{1}{n} \sum_{i=1}^{n} (Y_i - X_i' b)^2 \,,$$

  - Em outras palavras, encontramos o coeficiente b que maximiza a contribuição dos X_i à explicação de Y_i, tal qual medida pela média da distância ao quadrado entre os Y_i e $X_i' b$.
- Estimador é obtido como solução a uma otimização de uma função convexa, sem restrições $\implies$ condição de primeira ordem caracteriza o conjunto de soluções.
- Condições de primeira ordem podem ser escritas como:

$$\mathbf{0}_{k \times 1} = \sum_{i=1}^{n} X_i (Y_i - X_i' b) = \boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X} b)$$

Condição de posto e unicidade do mínimo

Sob a condição

Hipótese

A matriz $\boldsymbol{X}$ apresenta posto k.

Temos que $\boldsymbol{X}'\boldsymbol{X}$ é invertível (por quê), de modo que existe uma única solução ao problema de otimização, dada por:

$$\hat{b} = (\boldsymbol{X}'\boldsymbol{X})(\boldsymbol{X}'\boldsymbol{y}) \, .$$

- Condição de posto requer que nenhuma das colunas seja escrita como combinação linear das demais.
  - Se a primeira entrada dos X_i corresponde a um intercepto (i.e. $X_{i,1} = 1$ para todo $i = 1, \ldots, n$), nenhuma das colunas pode ser escrita como função afim das demais
- Observe que, como $\text{rank}(\boldsymbol{X}) \leq \min\{n, k\}$, condição implica que $n \geq k$.

Um caso simples

- Considere, para fixar as ideias, o caso em que $X_i = \begin{bmatrix} 1 & T_i \end{bmatrix}'$.
- Nesse caso, a matriz $\boldsymbol{X}'\boldsymbol{X}$ é dada por:

$$\begin{bmatrix} n & \sum_{i=1}^{n} T_i \\ \sum_{i=1}^{n} T_i & \sum_{i=1}^{n} T_i^2 \end{bmatrix} \tag{3}$$

  de modo que a condição de posto é equivalente a
  $\widehat{V(T)} = \frac{1}{n}\sum_{i=1}^{n} T_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} T_i\right)^2 > 0.$
- Se condição de posto é satisfeita, estimador de MQO é dado por:

$$\hat{b}_2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(T_i - \bar{T})}{\sum_{i=1}^{n}(T_i - \bar{T})^2} = \frac{\widehat{\text{cov}(T, Y)}}{\widehat{V(T)}}$$

MQO: Propriedades Algébricas

Matriz de projeção

- Definimos a matriz de projeção de $\boldsymbol{X}$ como:

$$P = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$$

  - Observe que a matriz de projeção é tal que $X\hat{b} = P\boldsymbol{y}$.
  - $P\boldsymbol{z}$ é a projeção (em termos de minimização da distância Euclidiana) de $\boldsymbol{z} \in \mathbb{R}^n$ no espaço gerado pelas colunas de $\boldsymbol{X}$.
- Matriz de projeção tem as seguintes propriedades:
  - Simétrica.
  - Idempotente ($P^2 = P$).
  - Os autovalores de P são ou 0 ou 1.
  - trace$(P) = k =$ rank(P).

Matriz residualizadora

- A matriz residualizadora (*residual-maker*) de X é dada por:

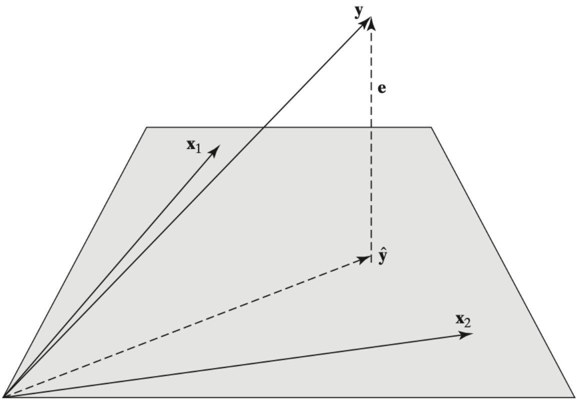
$$M = (I - P)$$

- Matriz residualizadora tem as seguintes propriedades:
  - Simétrica.
  - Idempotente ($M^2 = M$).
  - Os autovalores de M são ou 0 ou 1.
  - $\text{trace}(M) = n - k = \text{rank}(M)$.
  - $\boldsymbol{X}'M = 0$, $M\boldsymbol{X} = 0$, $PM = MP = 0$.
- Matriz residualizadora devolve o erro de projeção $\boldsymbol{z} - P\boldsymbol{z}$.
  - Como $P\boldsymbol{z}$ é o minimizador da distância Euclidiana no espaço gerado pelas colunas de $\boldsymbol{X}$,temos que:

$$(\boldsymbol{Mz}) \cdot (\boldsymbol{Xl}) = \boldsymbol{z}'M\boldsymbol{Xl} = 0 \quad \forall \boldsymbol{z} \in \mathbb{R}^n, \boldsymbol{l} \in \mathbb{R}^k \,.$$

Visualização gráfica

FIGURE 3.2 Projection of **y** into the Column Space of **X.**



Fórmula da inversa particionada

- Suponha que particionemos a matriz $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 & \boldsymbol{X}_2 \end{bmatrix}$ onde $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$ são matrizes de dimensão $n \times k_1$ e $n \times k_2$.
- Nesse caso, a fórmula da inversa particionada nos indica que:

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} \boldsymbol{X_1}'\boldsymbol{X}_1 & \boldsymbol{X_1}'\boldsymbol{X}_2 \\ \boldsymbol{X_2}'\boldsymbol{X}_1 & \boldsymbol{X_2}'\boldsymbol{X}_2 \end{bmatrix}^{-1} =$$

$$\begin{bmatrix} (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1} + (\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{F}\boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1} & -(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2\boldsymbol{F} \\ -\boldsymbol{F}\boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1} & \boldsymbol{F} \end{bmatrix}$$

onde $\boldsymbol{F} = (\boldsymbol{X}_2'\boldsymbol{X}_2 - \boldsymbol{X}_2'\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{X}_2)^{-1} = (\boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}$ onde $\boldsymbol{M}_1$ é a residualizadora de $\boldsymbol{X}_1$.

  - Inversa de $\boldsymbol{X}_2'\boldsymbol{M}_1\boldsymbol{X}_2$ existe pois matriz $\boldsymbol{X}$ tem posto cheio.

Teorema de Frisch-Waugh-Lovell

-
- Da propriedade acima, segue o importante resultado abaixo:

Teorema (Frisch-Waugh-Lovell)

$$\hat{b}_2 = \left(\boldsymbol{X}_2' M_1 \boldsymbol{X}_2\right)^{-1} \left(\boldsymbol{X}_2' M_1 \boldsymbol{y}\right) = \left((M_1\boldsymbol{X}_2)'(M_1\boldsymbol{X}_2)\right)^{-1} \left((M_1\boldsymbol{X}_2)'(M_1\boldsymbol{y})\right)$$

- Resultado acima nos mostra que estimadores de MQO associados a $\boldsymbol{X}_2$ em uma regressão que inclui $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$ são idênticos a:
  - Regredir $\boldsymbol{y}$ em $\boldsymbol{X}_1$, e guardar os resíduos $\boldsymbol{e}_y$.
  - Para cada $j = 1, \ldots k_2$, regredir a j-ésima coluna de $\boldsymbol{X}_2$ em $\boldsymbol{X}_1$, e guardar os resíduos $\boldsymbol{e}_j$.
  - Regredir $\boldsymbol{e}_y$ em $\boldsymbol{e}_1, \ldots \boldsymbol{e}_{k_2}$ e recuperar os coeficientes.