AULA 13 - MQO em Notação Matricial

Econometria I

Área quantitativa - IE/UFRJ

Modelo de regressão linear com k-variáveis

- Vimos que o modelo populacional pode ser escrito como:

$$y_i = \beta_1 + \beta_2 x_{2i} + ... + \beta_k x_{ki} + u_i$$

onde $i = 1, 2, 3, ..., n$

- A equação acima é uma expressão abreviada para o seguinte conjunto de n equações simultâneas:

$$\begin{aligned} y_1 &= \beta_1 + \beta_2 x_{21} + ... + \beta_k x_{k1} + u_1 \\ y_2 &= \beta_1 + \beta_2 x_{22} + ... + \beta_k x_{k2} + u_2 \\ &\vdots \\ y_n &= \beta_1 + \beta_2 x_{2n} + ... + \beta_k x_{kn} + u_n \end{aligned}$$

Modelo de regressão linear com k-variáveis

- Vamos escrever o sistema acima de uma maneira alternativa, porém mais esclarecedora, como segue.

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{21} & x_{32} & \cdots & x_{k1} \\ 1 & x_{22} & x_{32} & \cdots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & x_{2n} & x_{3n} & \cdots & x_{kn} \end{bmatrix} \times \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

$$\underbrace{\mathbf{y}}_{n\times 1} = \underbrace{\mathbf{X}}_{n\times k} \underbrace{\beta}_{k\times 1} + \underbrace{\mathbf{u}}_{n\times 1}$$

Hipóteses

H1 $E(\mathbf{u}) = \mathbf{0}$, onde $\mathbf{0}$ é um vetor coluna $(n \times 1)$

H2 $E(\mathbf{u}\mathbf{u}') = \sigma^2\mathbf{I}$, onde $\mathbf{I}$ é uma matriz identidade $(n \times n)$

$$E(\mathbf{u}\mathbf{u}') = E\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}$$

onde $\mathbf{u}'$ é a transposta do vetor coluna $\mathbf{u}$. Fazendo a multiplicação:

$$E(\mathbf{u}\mathbf{u}') = E\begin{bmatrix} u_1^2 & u_1u_2 & \cdots & u_1u_n \\ u_2u_1 & u_2^2 & \cdots & u_2u_n \\ \vdots & \vdots & & \vdots \\ u_nu_1 & u_nu_2 & \cdots & u_n^2 \end{bmatrix}$$

$$= \begin{bmatrix} E(u_1^2) & E(u_1u_2) & \cdots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & \cdots & E(u_2u_n) \\ \vdots & \vdots & & \vdots \\ E(u_nu_1) & E(u_nu_2) & \cdots & E(u_n^2) \end{bmatrix}$$

Hipóteses

H2 $E(\mathbf{u}\mathbf{u}') = \sigma^2\mathbf{I}$, onde $\mathbf{I}$ é uma matriz identidade $(n \times n)$
Por causa da hipótese de homocedasticidade e não correlação entre os erros temos

$$
\begin{aligned}
E(\mathbf{u}\mathbf{u}') &= \begin{bmatrix} E(u_1^2) & E(u_1u_2) & \cdots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & \cdots & E(u_2u_n) \\ \vdots & \vdots & & \vdots \\ E(u_nu_1) & E(u_nu_2) & \cdots & E(u_n^2) \end{bmatrix} \\
&= \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} \\
&= \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \sigma^2\mathbf{I}
\end{aligned}
$$

Hipóteses

- H3 A matriz $\mathbf{X}$ $n \times k$ não é estocástica, isto é, ela consiste de um conjunto de número fixados.
- H4 O posto de $\mathbf{X}$ é $p(\mathbf{X}) = k$, onde k é o número de colunas em $\mathbf{X}$ e $k < n$. Isto significa que as colunas da matrix $\mathbf{X}$ são linearmente independentes. Assim sendo não temos o problema de multicolinearidade.
- H5 O vetor $\mathbf{u}$ tem uma distribuição normal multivariada, isto é, $\mathbf{u} \sim (\mathbf{0}, \sigma^2\mathbf{I})$.

Estimação

- Para estimatimar β por MQO, escrevemos o modelo amostral:
$$y_i = \hat{\beta}_1 + \hat{\beta}_2 x_{2i} + ... + \hat{\beta}_k x_{ki} + \hat{u}_i$$
- De forma compacta
$$\mathbf{y} = \mathbf{X}\hat{\beta} + \hat{\mathbf{u}}$$
- Iremos minimizar a soma dos resíduos ao quadrado
$$\sum \hat{u}_i^2 = \sum (y_i - \hat{\beta}_1 - \hat{\beta}_2 x_{2i} - ... - \hat{\beta}_k x_{ki})^2$$
- Na notação matricial temos
$$\hat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\hat{\beta}$$
- Assim
$$\begin{aligned}\hat{\mathbf{u}}'\hat{\mathbf{u}} &= (\mathbf{y} - \mathbf{X}\hat{\beta})'(\mathbf{y} - \mathbf{X}\hat{\beta}) \\ &= \mathbf{y}'\mathbf{y} - \hat{\beta}'\mathbf{X}'\mathbf{y} - \mathbf{y}'\mathbf{X}\hat{\beta} + \hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta} \\ &= \mathbf{y}'\mathbf{y} - 2\hat{\beta}'\mathbf{X}'\mathbf{y} + \hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta}\end{aligned}$$

  usamos o fato de que a transposta de um escalar é um escalar ($y'X\hat{\beta} = \hat{\beta}'X'y$)

Equação Normal

- Então podemos escrever:

$$\min_{\hat{\beta}} S(\hat{\beta}) = \mathbf{y}'\mathbf{y} - 2\hat{\beta}'\mathbf{X}'\mathbf{y} + \hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta}$$

- A **condição necessária** para um mínimo é:

$$\frac{\partial S(\hat{\beta})}{\partial \hat{\beta}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\hat{\beta} = 0$$

- Sendo $\hat{\beta}$ a solução do problema de minimização, obtemos a **equação normal**

$$(\mathbf{X}'\mathbf{X})\hat{\beta} = \mathbf{X}'\mathbf{y}$$

- Se a inversa de $(\mathbf{X}'\mathbf{X})$ existe, então pre-multiplicar ambos lados pela inversa temos:
$(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$

- Sabemos que $(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X}) = I$. Isto nós dá:

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

Matriz varcov de $\hat{\beta}$

- Sabemos que $\hat{\beta} = (\mathbf{X'X})^{-1}\mathbf{X'y}$ e que $\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$ então

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

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

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

- Por definição

$$\begin{aligned} varcov(\hat{\beta}) = {} & E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] \\ & E\{[(X'X)^{-1}X'u][(X'X)^{-1}X'u]'\} \\ & E[(X'X)^{-1}X'uu'X(X'X)^{-1}] \end{aligned}$$

- Sendo X não estocástico

$$\begin{aligned} varcov(\hat{\beta}) &= (X'X)^{-1}X'E[uu']X(X'X)^{-1} \\ &= (X'X)^{-1}X'\sigma^2 IX(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1} \end{aligned}$$