Econometria I

Exame 1 (Gabarito)

26/10/2023

1. a. Sabemos que $\mathbf{b}_1 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$.

$$\begin{aligned} E[\mathbf{b}_1|\mathbf{X},\mathbf{Z}] =& E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}|\mathbf{X},\mathbf{Z}] \\ =& (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E[\mathbf{X}\beta_2 + \mathbf{Z}\gamma + \varepsilon_2|\mathbf{X},\mathbf{Z}] \\ =& \beta_2 + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\gamma \end{aligned}$$

caso o pressuposto $E[\varepsilon_2|\mathbf{X},\mathbf{Z}] = 0$ seja satisfeito. Para encontrarmos a variância:

$$\begin{aligned} Var[\mathbf{b}_1|\mathbf{X},\mathbf{Z}] =& Var[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}|\mathbf{X},\mathbf{Z}] \\ =& (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Var[\mathbf{X}\beta_2 + \mathbf{Z}\gamma + \varepsilon_2|\mathbf{X},\mathbf{Z}]\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ =& \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \end{aligned}$$

caso o pressuposto $Var[\varepsilon_2|\mathbf{X},\mathbf{Z}] = \sigma^2\mathbf{I}$ seja satisfeito.

b. Aplicando o teorema FWL, temos que $\mathbf{b}_2 = (\mathbf{X}'\mathbf{M_z}\mathbf{X})^{-1}\mathbf{X}'\mathbf{M_z}\mathbf{y}$. Portanto,

$$\begin{aligned} E[\mathbf{b}_2|\mathbf{X},\mathbf{Z}] =& E(\mathbf{X}'\mathbf{M_z}\mathbf{X})^{-1}\mathbf{X}'\mathbf{M_z}\mathbf{y}|\mathbf{X},\mathbf{Z}] \\ =& (\mathbf{X}'\mathbf{M_z}\mathbf{X})^{-1}\mathbf{X}'\mathbf{M_z}E[\mathbf{X}\beta_1 + \varepsilon_1|\mathbf{X},\mathbf{Z}] \\ =& \beta_1 \end{aligned}$$

caso o pressuposto $E[\varepsilon_1|\mathbf{X}] = 0$ seja satisfeito. Encontrando a variância:

$$\begin{aligned} Var[\mathbf{b}_2|\mathbf{X},\mathbf{Z}] =& Var[(\mathbf{X}'\mathbf{M_z}\mathbf{X})^{-1}\mathbf{X}'\mathbf{M_z}\mathbf{y}|\mathbf{X},\mathbf{Z}] \\ =& (\mathbf{X}'\mathbf{M_z}\mathbf{X})^{-1}\mathbf{X}'\mathbf{M_z}Var[\mathbf{X}\beta_1 + \varepsilon_1|\mathbf{X},\mathbf{Z}]\mathbf{M_z}\mathbf{X}(\mathbf{X}'\mathbf{M_z}\mathbf{X})^{\mathbf{-1}} \\ =& \sigma^2(\mathbf{X}'\mathbf{M_z}\mathbf{X})^{-1} \end{aligned}$$

caso o pressuposto $Var[\varepsilon_1|\mathbf{X},\mathbf{Z}] = \sigma^2\mathbf{I}$ seja satisfeito.

c. Caso o modelo 1 represente o verdadeiro processo gerador de dados, $E[\mathbf{b}_1|\mathbf{X},\mathbf{Z}] = \beta_1$ e $Var[\mathbf{b}_1|\mathbf{X},\mathbf{Z}] = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. Portanto, $E[\mathbf{b}_1|\mathbf{X},\mathbf{Z}] = E[\mathbf{b}_2|\mathbf{X},\mathbf{Z}] = \beta_1$. Em relação às variâncias, como $\mathbf{M_z}$ é uma matrix de projeção ortogonal, temos que $\mathbf{X}'\mathbf{M_z}\mathbf{X} \leq \mathbf{X}'\mathbf{X}$. Então, $\sigma^2(\mathbf{X}'\mathbf{M_z}\mathbf{X})^{-1} \geq \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \Rightarrow Var[\mathbf{b}_2|\mathbf{X},\mathbf{Z}] \geq Var[\mathbf{b}_1|\mathbf{X},\mathbf{Z}]$

d. Caso o modelo 1 represente o verdadeiro processo gerador de dados, $E[\mathbf{b}_2|\mathbf{X},\mathbf{Z}] = \beta_2$, enquanto que $E[\mathbf{b}_1|\mathbf{X},\mathbf{Z}] = \beta_2 + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\gamma$, e $Var[\mathbf{b}_2|\mathbf{X},\mathbf{Z}] = \sigma^2(\mathbf{X}'\mathbf{M_z}\mathbf{X})^{-1}$. Portanto, também temos neste caso que $Var[\mathbf{b}_2|\mathbf{X},\mathbf{Z}] \geq Var[\mathbf{b}_1|\mathbf{X},\mathbf{Z}]$.

e. Se o modelo 1 for o verdadeiro, ambos $\mathbf{b}_1$ e $\mathbf{b}_2$ são não-viesados, porém $\mathbf{b}_1$ é eficiente, pois possui menor variância. Caso o modelo 2 seja o verdadeiro, $\mathbf{b}_1$ é viesado enquanto $\mathbf{b}_2$ não. Porém $\mathbf{b}_1$ possui menor variância. Temos, portanto, um trade-off viés-variância ao comparar $\mathbf{b}_1$ e $\mathbf{b}_2$.

2. a. De acordo com as definições de $z_{1,i}$, $z_{2,i}$ e $z_{3,i}$ em termos de $x_{1,i}$, $x_{2,i}$ e $x_{3,i}$, temos que

$$\mathbf{Z} = \mathbf{X}\mathbf{A}$$

$$\begin{bmatrix}\mathbf{z}_1 & \mathbf{z}_2 & \mathbf{z}_3\end{bmatrix} = \begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3\end{bmatrix}\begin{bmatrix}1 & 0 & 2\\ -2 & 1 & -3\\ 0 & 4 & 5\end{bmatrix}$$

Note que $\mathbf{A}$ é invertível:

$$\det\ \mathbf{A} = \det\begin{pmatrix}1 & -3\\ 4 & 5\end{pmatrix} + 0 + 2\det\begin{pmatrix}-2 & 1\\ 0 & 4\end{pmatrix} = 5 + 12 + 2(-8) = 1 \neq 0$$

b. Denote $\mathbf{b}$ o estimador de MQO do vetor de parâmetros β e $\hat{\alpha}$ o estimador de MQO do vetor α. Portanto, temos que $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ e $\hat{\alpha} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}$. Usando a relação entre $\mathbf{X}$ e $\mathbf{Z}$ encontrada acima, temos que $\mathbf{X} = \mathbf{Z}\mathbf{A}^{-1}$. Portanto,

$$\begin{aligned}\mathbf{b} =& [(\mathbf{Z}\mathbf{A}^{-1})'\mathbf{Z}\mathbf{A}^{-1}]^{-1}(\mathbf{Z}\mathbf{A}^{-1})'\mathbf{y}\\ =& [(\mathbf{A}')^{-1}\mathbf{Z}'\mathbf{Z}\mathbf{A}^{-1}]^{-1}(\mathbf{A}')^{-1}\mathbf{Z}'\mathbf{y}\\ =& \mathbf{A}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{A}'(\mathbf{A}')^{-1}\mathbf{Z}'\mathbf{y}\\ =& \mathbf{A}\hat{\alpha}\end{aligned}$$

Portanto, $b_1 = \begin{pmatrix}1 & 0 & 2\end{pmatrix}\hat{\alpha} = \hat{\alpha}_1 + 2\hat{\alpha}_3$.

c. Os valores previstos da regressão de $\mathbf{y}$ em $\mathbf{X}$ são dados por

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\mathbf{b}} = \mathbf{X}\mathbf{A}\hat{\alpha} = \mathbf{Z}\hat{\alpha}$$

o que corresponde aos valores previstos da regressão de $\mathbf{y}$ em $\mathbf{Z}$. Como os valores previstos são os mesmos, então os resíduos também são. (Outra forma de chegar este resultado é mostrar que $\mathbf{P}_X = \mathbf{P}_Z$ e $\mathbf{M}_X = \mathbf{M}_Z$). Ambos os modelos são linearmente relacionados através da matriz $\mathbf{A}$. Isto significa que os dois modelos contem o mesmo conjunto de informações e que eles são equivalentes. Um modelo representa uma reformulação do outro.

3. Dada a distribuição assintótica do estimador de MQO $\mathbf{b}$, podemos definir um teste assintótico baseado na estatística de teste Wald.

a. O estimador de MQO $\mathbf{b}$ do vetor de parâmetros β é tal que

$$\sqrt{n}(\mathbf{b} - \beta) \xrightarrow{d} N[0, \sigma^2\mathbf{Q}^{-1}]$$

onde $\mathbf{Q}$ é tal que plim $\frac{(\mathbf{X}'\mathbf{X})}{n} = \mathbf{Q}$. Defina o vetor de restrições

$$g(\beta) = \begin{pmatrix}\beta_1^2 - \beta_2^2 - 5\\ \beta_2 + \beta_3 - 1\end{pmatrix}$$

Portanto, queremos testar

$$H_0 : g(\beta) = \begin{pmatrix}0\\0\end{pmatrix} \text{ vs } H_1 : g(\beta) \neq \begin{pmatrix}0\\0\end{pmatrix}$$

Ademais, defina $\mathbf{C}(\mathbf{b}) = \frac{\partial g(\mathbf{b})}{\partial \mathbf{b}'}$. De acordo com o teorema de Slutsky, plim $g(\mathbf{b}) = g(\beta)$ e, aplicando o método Delta, temos que

$$\sqrt{n}(g(\mathbf{b}) - g(\beta)) \xrightarrow{d} N[\mathbf{0}, \mathbf{\Gamma}[Asy.Var(\mathbf{b})]\mathbf{\Gamma}']$$

onde $\mathbf{\Gamma}$ é tal que plim $\mathbf{C}(\mathbf{b}) = \frac{\partial g(\beta)}{\partial \beta'} = \mathbf{\Gamma}$. A estatística Wald é dada por

$$W = ng(\mathbf{b})'[\hat{\mathbf{\Gamma}}[Asy.\hat{V}ar(\mathbf{b})]\hat{\mathbf{\Gamma}}']^{-1}g(\mathbf{b})$$

onde

$$\hat{\mathbf{\Gamma}} = \frac{\partial g(\mathbf{b})}{\partial \mathbf{b}'} = \begin{pmatrix} 2b_1 & -2b_2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

b. A distribuição limite da estatística de teste sob H_0 é encontrada a seguir:

$$\sqrt{n}g(\mathbf{b}) \xrightarrow{d} N[\mathbf{0}, \mathbf{\Gamma}[Asy.Var(\mathbf{b})]\mathbf{\Gamma}']$$

$$\Rightarrow \sqrt{n}[\mathbf{\Gamma}[Asy.Var(\mathbf{b})]\mathbf{\Gamma}']^{-1/2}g(\mathbf{b}) \xrightarrow{d} N[\mathbf{0}, \mathbf{I}]$$

$$\Rightarrow ng(\mathbf{b})'[\mathbf{\Gamma}[Asy.Var(\mathbf{b})]\mathbf{\Gamma}']^{-1}g(\mathbf{b}) \xrightarrow{d} \chi^2(2)$$

Regra de decisão: Rejeitar H_0 se $W > \chi^2_{1-\alpha}(2)$.

4. a. Precisamos mostrar que plim $\mathbf{b} = \beta$. O estimador de MQO é dado por

$$\mathbf{b} = \beta + \left(\frac{\mathbf{X}'\mathbf{X}}{n}\right)^{-1}\left(\frac{\mathbf{X}'\varepsilon}{n}\right)$$

Assuma plim $\frac{(\mathbf{X}'\mathbf{X})}{n} = \mathbf{Q}$ é uma matriz positiva definida. Portanto,

$$\text{plim } \mathbf{b} = \beta + \mathbf{Q}^{-1}\text{plim } \left(\frac{\mathbf{X}'\varepsilon}{n}\right)$$

Note que podemos escrever $\frac{\mathbf{X}'\varepsilon}{n} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\varepsilon_i$. Como a média amostral é um estimador consistente da média populacional, plim $\left(\frac{\mathbf{X}'\varepsilon}{n}\right) = E(\mathbf{X}'\varepsilon) = 0$ dada a hipótese de exogeneidade $E(\varepsilon|\mathbf{X}) = 0$. Portanto, encontramos plim $\mathbf{b} = \beta$.

b. Escrevemos

$$\mathbf{b} - \beta = \left(\frac{\mathbf{X}'\mathbf{X}}{n}\right)^{-1}\left(\frac{\mathbf{X}'\varepsilon}{n}\right)$$

Multiplicando ambos os lados por $\sqrt{n}$, temos

$$\sqrt{n}(\mathbf{b} - \beta) = \left(\frac{\mathbf{X}'\mathbf{X}}{n}\right)^{-1}\left(\frac{1}{\sqrt{n}}\right)\mathbf{X}'\varepsilon$$

Sabemos, portanto, que a distribuição limite de $\sqrt{n}(\mathbf{b} - \beta)$ é a mesma de $\mathbf{Q}^{-1}\frac{1}{\sqrt{n}}\mathbf{X}'\varepsilon$. Vamos estabelecer agora a distribuição limite de $\frac{1}{\sqrt{n}}\mathbf{X}'\varepsilon$:

$$\frac{1}{\sqrt{n}}\mathbf{X}'\varepsilon = \frac{1}{\sqrt{n}}\sum \mathbf{x}_i\varepsilon_i = \frac{1}{\sqrt{n}}\sum\left(\mathbf{x}_i\varepsilon_i - E(\mathbf{x}_i\varepsilon_i)\right)$$

Defina $\mathbf{w}_i = \mathbf{x}_i\varepsilon_i$, onde $\bar{\mathbf{w}} = \frac{1}{n}\sum \mathbf{x}_i\varepsilon_i$, e multiplique a expressão acima por n/n:

$$\frac{1}{\sqrt{n}}\mathbf{X}'\varepsilon = \sqrt{n}\frac{1}{n}\sum(\mathbf{w}_i - E(\mathbf{w}_i)) = \sqrt{n}[\bar{\mathbf{w}} - E(\bar{\mathbf{w}})]$$

O vetor $\bar{\mathbf{w}}$ é a média de n vetores aleatórios $i.i.d$ com média $\mathbf{0}$ e variáncia $\sigma^2\mathbf{Q}$, assumindo que os erros do modelo de regressão são homoscedásticos e não-correlacionados entre si. Aplicando o teorema do limite central de Lindeberg-Levy, como $\{\mathbf{x}_i\varepsilon_i\}_i$, $i = 1, \cdots, n$, são vetores independentes, cada um distribuído com média $\mathbf{0}$ e variância $\sigma^2\mathbf{Q} < \infty$,

$$\frac{1}{\sqrt{n}}\mathbf{X}'\varepsilon \xrightarrow{d} N[\mathbf{0}, \sigma^2\mathbf{Q}]$$

Multiplicando por $\mathbf{Q}^{-1}$, temos

$$\sqrt{n}(\mathbf{b} - \beta) \xrightarrow{d} N[\mathbf{0}, \sigma^2\mathbf{Q}^{-1}]$$

Logo,

$$\mathbf{b} \overset{a}{\sim} N\left[\beta, \frac{\sigma^2}{n}\mathbf{Q}^{-1}\right]$$

5. A resposta deveria abranger tanto a interpretação dos parâmetros (efeitos marginais) quanto a significância estatística. Os coeficientes são conjuntamente estatisticamente significantes a 5% ($\chi^2(2) = 6.87$). Individualmente, o coeficiente de x_2 é estatisticamente significante a 5%, mas o de x_3 não é. Note que Efeito parcial$_j = g'(1 + 2x_2 + 3x_3)\beta_j$, onde $g'() < 0$ então o sinal do efeito parcial é o inverso do sinal de β_j. Segue-se que $E[y|x_2, x_3]$ diminui quando x_2 aumenta e quando x_3 aumenta.