AULA 5 - MQO Simples: Variância Desconhecida e Teorema de Gauss-Markov

Econometria I

Área quantitativa - IE/UFRJ

Estimando $V[\hat{\beta}|X]$ com σ^2 desconhecida

- A derivação na aula passada nos mostrou um estimador para a variância de $\hat{\beta}$

$$V[\hat{\beta}|X] = \frac{\sigma^2}{\sum(x_i - \overline{x})^2}$$

- No entanto, lembremos que: da hipótese de que $V[u_i|X] = \sigma^2$, onde σ^2 é um parâmetro populacional e desconhecido.
- Portanto, precisamos de uma maneira de encontrar (estimar) o parâmetro σ^2:
  - Caso os erros u_i fossem observados, o que podemos fazer? sendo $V[u|X] = E(u^2|X) - [\underbrace{E(u|X)}_{0}]^2 = E(u^2|X) = \sigma^2$, o que significa também que vale para o caso não condicional: $V[u] = E(u^2) = \sigma^2$, assim uma estimativa inicial seria $\hat{\sigma}^2 = \frac{\sum u_i^2}{n}$, por exemplo?
  - No entanto, não observamos $\sum u_i^2$

Estimando $V[\hat{\beta}|X]$ com σ^2 desconhecida

- Uma alternativa consiste em utilizar os resíduos $\hat{u}_i = \hat{y}_i - y_i$ em lugar de u_i. Teríamos então o seguinte estimador:
  $$\hat{\sigma}^2 = \sum \frac{\hat{u}_i^2}{n}$$
  o qual é computável.
- O problema é que este estimador é viesado. A nossa estratégia a partir de agora será:
  - Mostrar que o estimador acima é viesado.
  - Ao mostrar isso, identificaremos o viés e conseguiremos sugerir um **novo estimador não viesado**.

Estimando $V[\hat{\beta}|X]$ com σ^2 desconhecida

- Primeiro, vamos montar a seguinte equação:
  (i) $y_i = \hat{\alpha} + \hat{\beta} x_i + \hat{u}_i$
  (ii) $y_i = \alpha + \beta x_i + u_i$
  Fazendo (i)=(ii):

$$\hat{u}_i = u_i - (\hat{\alpha} - \alpha) - (\hat{\beta} - \beta) x_i \tag{1}$$

- Ou seja, escrevemos $\hat{u}_i$ em função de u_i populacional (não observado). Agora, tomemos a média da equação acima:

$$\sum \frac{\hat{u}_i}{n} = \sum \frac{u_i}{n} - (\hat{\alpha} - \alpha) - (\hat{\beta} - \beta) \sum \frac{x_i}{n}$$

- Temos então:

$$0 = \overline{u} - (\hat{\alpha} - \alpha) - (\hat{\beta} - \beta)\overline{x} \tag{2}$$

Estimando $V[\hat{\beta}|X]$ com σ^2 desconhecida

- Subtraindo (1)-(2), temos:

$$\hat{u}_i = (u_i - \overline{u}) - (\hat{\beta} - \beta)(x_i - \overline{x})$$

E, elevando os dois lados ao quadrado:

$$\hat{u}_i^2 = (u_i - \overline{u})^2 + (\hat{\beta} - \beta)^2(x_i - \overline{x})^2 - 2(u_i - \overline{u})(\hat{\beta} - \beta)(x_i - \overline{x})$$

Tomando o somatório dos dois lados em $i = 1, ..., n$:

$$\sum \hat{u}_i^2 = \underbrace{\sum (u_i - \overline{u})^2}_{(a)} + \underbrace{\sum (\hat{\beta} - \beta)^2 (x_i - \overline{x})^2}_{(b)}$$
$$- 2 \underbrace{\sum (u_i - \overline{u})(\hat{\beta} - \beta)(x_i - \overline{x})}_{(c) = -2(\hat{\beta} - \beta)\sum (x_i - \overline{x}) u_i}$$

- Agora, basta tomar o valor esperado dos dois lados da expressão acima:

Estimando $V[\hat{\beta}|X]$ com σ^2 desconhecida

- Vamos tomar o valor esperado de cada termo separadamente, a começar pelo primeiro:

$$(a):\ \sum(u_i - \overline{u})^2 = \sum u_i^2 - n\overline{u}^2 \quad \text{ver apêndice A.7 Wooldridge}$$

Fazendo para cada um dos dois termos acima:
$E[\sum u_i^2] = \sum E[u_i^2] = n\sigma^2$
$E[-n\overline{u}^2] = -nE[\overline{u}^2] = -n\,\underbrace{[E\overline{u}^2 - (\underbrace{E(\overline{u})}_{0})^2]}_{\sigma^2/n} = -n\frac{\sigma^2}{n} = -\sigma^2$

- Logo, fechamos o primeiro termo:

$$E[(a)] = E[\sum(u_i - \overline{u})^2] = n\sigma^2 - \sigma^2 = (n-1)\sigma^2$$

Estimando $V[\hat{\beta}|X]$ com σ^2 desconhecida

- Para o segundo termo:

$$(b): \sum(\hat{\beta}-\beta)^2(x_i-\overline{x})^2$$

$$E[\sum(\hat{\beta}-\beta)^2(x_i-\overline{x})^2] = \sum(x_i-\overline{x})^2 \underbrace{E[(\hat{\beta}-\beta)^2]}_{\frac{\sigma^2}{\sum(x_i-\overline{x})^2}} = \sigma^2$$

- Assim: $E[(b)] = \sigma^2$
- Por fim, com relação ao terceiro termo, é possível mostrar também que:

$$\begin{aligned} E[(c)] &= E[-2(\hat{\beta}-\beta)\sum(x_i-\overline{x})u_i] \\ &= -2E[(\hat{\beta}-\beta)^2]\sum(x_i-\overline{x})^2 = -2\sigma^2 \end{aligned}$$

Estimando $V[\hat{\beta}|X]$ com σ^2 desconhecida

- Logo, podemos concluir que o valor esperado da expressão abaixo:

$$\sum \hat{u}_i^2 = \underbrace{\sum (u_i - \overline{u})^2}_{(a)} + \underbrace{\sum (\hat{\beta} - \beta)^2 (x_i - \overline{x})^2}_{(b)}$$
$$- 2 \underbrace{\sum (u_i - \overline{u})(\hat{\beta} - \beta)(x_i - \overline{x})}_{(c) = -2(\hat{\beta} - \beta) \sum (x_i - \overline{x}) u_i}$$

- Pode ser escrito como:

$$E[\sum \hat{u}_i^2] = E[(a)] + E[(b)] + E[(c)] = (n-1)\sigma^2 + \sigma^2 - 2\sigma^2 = (n-2)\sigma^2$$

- Neste caso, notem que o estimador

$$\hat{\sigma}^2 = \sum \frac{\hat{u}_i^2}{n} = \frac{SQR}{n} \rightarrow \text{viesado, já que } \frac{(n-2)}{n}\sigma^2 \neq \sigma^2$$

- Qual seria então um estimador não viesado para σ^2?

Estimando $V[\hat{\beta}|X]$ com σ^2 desconhecida

- Agora ficou fácil ver que o estimador

$$\hat{\sigma}^2 = \sum \frac{\hat{u}_i^2}{n-2} = \frac{SQR}{n-2} \rightarrow \textbf{não viesado}\text{, já que } \frac{(n-2)}{n-2}\sigma^2 = \sigma^2$$

- Agora que já encontramos um estimador não viesado e computável para o parâmetro populacional σ^2, podemos sugerir o seguinte estimador para $V[\hat{\beta}|X]$:

$$V[\hat{\beta}|X] = \frac{\hat{\sigma}^2}{\sum(x_i - \overline{x})^2} = \frac{\frac{SQR}{n-2}}{\sum(x_i - \overline{x})^2}$$

- Já temos então todos os elementos para computar $V[\hat{\beta}|X]$.

Re-interpretando o estimador de MQO

- Antes de enunciar e provar o Teorema de Gauss-Markov, é útil refazer algumas contas de modo diferente de forma a revelar interpretações alternativas sobre o MQO. Defina:

$$\hat{\beta} = \sum k_i y_i, \text{ onde } k_i = \frac{(x_i - \overline{x})}{\sum (x_i - \overline{x})^2}$$

- Aqui vemos claramente uma implicação da hipótese de linearidade no modelo populacional: $\hat{\beta}$ é uma função linear de y_i. Mais precisamente, é uma média ponderada dos y_i, onde os pesos são dados por $k_i = \frac{(x_i - \overline{x})}{\sum (x_i - \overline{x})^2}$
- Vejamos algumas propriedades desses pesos. Vocês podem prova-las como exercício:
  1. $\sum k_i = 0$, pois $\sum x_i - \sum \overline{x} = \sum x_i - n \sum x_i / n = 0$
  2. $\sum k_i^2 = \frac{1}{\sum (x_i - \overline{x})^2}$
  3. $\sum k_i (x_i - \overline{x}) = \sum k_i x_i = 1$

Re-interpretando o estimador de MQO

- Notem também que:
$\sum k_i y_i = \sum k_i(\alpha + \beta x_i + u_i) = \underbrace{\alpha \sum k_i}_{0} + \underbrace{\beta \sum k_i x_i}_{1} + \sum k_i u_i$
$\sum k_i y_i = \beta + \sum k_i u_i$
- Assim, vocês podem tomar o valor esperado e a variância da expressão acima para encontrar as porpriedades estatísticas do MQO novamente. Se valem as hipóteses enunciadas até aqui:
$E[\hat{\beta}|X] = \beta + \underbrace{\sum k_i E[u_i|X]}_{0} = \beta$
$V[\hat{\beta}|X] = E[(\hat{\beta} - \underbrace{E(\hat{\beta})}_{\beta})^2|X)|X] = E[(\hat{\beta} - \beta)^2|X] = E[\sum(k_i u_i)^2|X]$

Re-interpretando o estimador de MQO

Desenvolvendo o somatório

$E[\sum(k_i u_i)^2|X] =$

$E[k_1^2 u_1^2 + k_2^2 u_2^2 + ... + k_n^2 u_n^2 + 2k_1 k_2 u_1 u_2 + ... + 2k_{n-1} k_n u_{n-1} u_n|X]$

Onde:

$E[u_i^2] = \sigma^2$

$E[u_i u_j] = 0 \quad \forall\, i \neq j$

- Logo:

$$V[\hat{\beta}|X] = \sigma^2 \sum k_i^2 = \frac{\sigma^2}{\sum(x_i - \overline{x})^2}$$

Teorema de Gauss-Markov

Theorem (Gauss-Markov)

Sejam as hipóteses clássicas de um modelo de Regressão Linear

1. *Linearidade:* $y = \alpha + \beta x + u$
2. *Amostragem aleatória* (x_i, y_i); $i = 1, ..., n$. *A amostra é suficientemente grande, com variação em x.*
3. *Independência entre x e u,* $E[u] = 0$ *e* $E[u|X] = 0$.
4. *Homocedasticidade:* $Var(u|X) = \sigma^2$

Então os estimadores por MQO são BLUE/MELNV (Best Linear Unbiased Estimator/Melhor Estimador Linear Não Viesado). Ou seja, na classe dos estimadores lineares não viesados, são aqueles com variância mínima (mais eficientes).

Teorema de Gauss-Markov

Prova:
Seja $\hat{\beta}_{MQO} = \sum k_i y_i$, onde k_i é o fator de ponderação do MQO.
Defina um $\hat{\beta}_{alternativo} = \beta^* = \sum w_i y_i$, onde w_i é um fator alternativo a k_i. Logo:

$$E[\beta^*|X] = \sum w_i(\alpha + \beta x_i + u_i) = \alpha \sum w_i + \beta \sum w_i x_i + \sum w_i u_i$$

Além da independência entre x e u, para β^* ser não viesado precisamos $\sum w_i = 0$ e $\sum w_i x_i = 1$.
Já que sob as condições acima ambos os estimadores são não viesados, agora temos que comparar suas variâncias para ver qual é o mais eficiente. Fazendo:

$$V(\beta^*) = Var(\sum w_i y_i) = \sum w_i^2 Var(y_i) = \sigma^2 \sum w_i^2$$

Teorema de Gauss-Markov

Continuação da prova: como visto acima $V(\beta^*) = \sigma^2 \sum w_i^2$
Defina uma $w_i = k_i + d_i$, sendo k_i constante e d_i constante arbitrária.

$$\begin{aligned} V(\beta^*) &= \sigma^2 \sum (k_i + d_i)^2 \\ &= \sigma^2 (\sum k_i^2 + \sum d_i^2 + 2 \sum k_i d_i) \\ &= \sigma^2 \sum k_i^2 + \sigma^2 \sum d_i^2 + \sigma^2 2 \sum k_i d_i \end{aligned}$$

onde $\sigma^2 \sum k_i^2 = V(\hat{\beta})$
$\sum k_i d_i = \sum k_i (w_i - k_i) = \sum k_i w_i - \sum k_i^2 =$
$\sum w_i \frac{(x_i - \overline{x})}{\sum (x_i - \overline{x})^2} - \frac{1}{\sum (x_i - \overline{x})^2} = \frac{\sum w_i x_i - \overline{x} \sum w_i}{\sum (x_i - \overline{x})^2} - \frac{1}{\sum (x_i - \overline{x})^2} = 0$
Assim

$$V(\beta^*) = V(\hat{\beta}) + \sigma^2 \sum d_i^2$$

para ser mínima a variância o $d_i = 0$ o que implica que $w_i = k_i$

Teorema de Gauss-Markov

Continuação da prova:
Concluindo: não existe qualquer outro fator de ponderação melhor que o de MQO (isto é, não existe um estimador que seja simultaneamente não viesado e mais eficiente que o MQO).