AULA 4 - MQO Simples: Propriedades algébricas e Estatísticas

Econometria I

Área quantitativa - IE/UFRJ

Estimação: MQO recapitulando

- Na aula passada aprendemos estimação por MQO. Recapitulando brevemente
- Em 1o lugar, partimos das seguintes hipóteses populacionais:
  - O modelo populacional existe, com linearidade nos parâmetros:

$$Y = \alpha + \beta X + u$$

  - Temos uma amostra aleatória da população para X e Y: $\{(X_i, Y_i), i = 1, 2, ..., n\}$, tamanho n.
  - $E(u) = 0$ e $E(u|X) = 0$.
- Nosso objetivo: estimar os parâmetros populacionais α e β não observados com base em nossa amostra de dados.

Estimação: MQO recapitulando

- Reescrevendo o modelo populacional com base na amostra indexada por i:

$$y_i = \alpha + \beta x_i + u_i$$

- Onde o par (x_i, y_i) é observado para todo i, subscrito para o indivíduo $i = 1, 2, ..., n$; e onde, naturalmente, u_i não é observado.
- Como vimos, o estimador de MQO consiste na solução $\hat{\alpha}$ e $\hat{\beta}$ que resolve o programa de minimização:

$$\min_{\alpha,\beta} S(\alpha, \beta) = \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2$$

Estimação: MQO recapitulando

- Resolvendo, temos um sistema com 2 equações e 2 incógnitas:

$$\frac{\partial S}{\partial \alpha} = 0 \Rightarrow -2\sum_{i=1}^{n}(y_i - \hat{\alpha} - \hat{\beta}x_i) = 0$$

$$\frac{\partial S}{\partial \beta} = 0 \Rightarrow -2\sum_{i=1}^{n}x_i(y_i - \hat{\alpha} - \hat{\beta}x_i) = 0$$

- Onde $\hat{\alpha}$ e $\hat{\beta}$ são solução:

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

$$\hat{\alpha} = \overline{y} - \hat{\beta}\overline{x}$$

Propriedades Algébricas do MQO

- Preciamos agora derivar e interpretar as propriedades de nossa solução, os estimadores de MQO
- Começaremos pelas propriedades **algébricas** do MQO.
- Em 1o lugar, analisemos o vetor de resíduos u. Lembrando, para todo i deve valer:

$$\hat{u}_i = y_i - \hat{\alpha} - \hat{\beta} x_i = y_i - \hat{y}_i$$

- É fácil notar, portanto, que devem valer as seguintes propriedades para os resíduos $\hat{u}_i$:
  1. $\sum_{i=1}^{n} \hat{u}_i = 0$
  2. $\sum_{i=1}^{n} \hat{u}_i x_i = 0$

Propriedades Algébricas do MQO

- Vejamos mais algumas propriedades algébricas. Faça:

$$\hat{u}_i = y_i - \hat{y}_i \Rightarrow \sum_{i=1}^{n} \frac{\hat{u}_i}{n} = \sum_{i=1}^{n} \frac{y_i}{n} - \sum_{i=1}^{n} \frac{\hat{y}_i}{n}$$

  Mas, como vale a propriedade $\sum_{i=1}^{n} \hat{u}_i = 0$, deve valer também:

$$\sum_{i=1}^{n} \frac{y_i}{n} - \sum_{i=1}^{n} \frac{\hat{y}_i}{n} = 0 \Rightarrow \overline{y} = \overline{\hat{y}}$$

  O que significa isso? Note também que:

$$\overline{y} = \hat{\alpha} + \hat{\beta}\overline{x} \Rightarrow \overline{\hat{y}} = \hat{\alpha} + \hat{\beta}\overline{x}$$

- O que significa isso? A reta estimada passa pelos pontos médios (y, x).

Propriedades Algébricas do MQO

- Por fim, vamos mostrar que $cov(\hat{y}_i, \hat{u}_i) = 0$. Mas por que isso é importante? Reescrevendo:

$$cov(\hat{y}_i, \hat{u}_i) = cov(\hat{\alpha} + \hat{\beta}x_i, \hat{u}_i)$$

- E então, por que isso importa? Fazendo mais algumas contas:

$$= \sum_{i=1}^{n} \frac{(\hat{\alpha} + \hat{\beta}x_i)\hat{u}_i}{n} - \sum_{i=1}^{n} \frac{(\hat{\alpha} + \hat{\beta}x_i)}{n} \underbrace{\sum_{i=1}^{n} \frac{\hat{u}_i}{n}}_{=0} =$$

$$\hat{\alpha} \underbrace{\sum_{i=1}^{n} \frac{\hat{u}_i}{n}}_{=0} + \hat{\beta} \underbrace{\sum_{i=1}^{n} \frac{x_i \hat{u}_i}{n}}_{=0} = 0$$

- Por que isso é importante?

Propriedades Algébricas do MQO

- Como $y_i = \hat{y}_i + \hat{u}_i \Rightarrow cov(\hat{y}_i, \hat{u}_i) = 0$, temos então que somos capazes de decompor perfeitamente a variação que existe em y_i em variações que existem em $\hat{y}_i$ (explicadas por x) e em $\hat{u}_i$ (não explicadas por x).
- Vejamos porque isso é útil. Vamos partir da variação em y_i que existe na amostra:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}[\underbrace{(y_i - \hat{y}_i)}_{\hat{u}_i} + (\hat{y}_i - \overline{y})]^2 =$$

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}\hat{u}_i^2 + \underbrace{2\sum_{i=1}^{n}\hat{u}_i(\hat{y}_i - \overline{y}_i)}_{=0} + \sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2$$

Propriedades Algébricas do MQO

- Continuando...como $\overline{\hat{u}} = 0$, podemos escrever:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}(\hat{u}_i - \overline{\hat{u}}_i)^2 + 0 + \sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2$$

- Assim sendo:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}(\hat{u}_i - \overline{\hat{u}}_i)^2 + \sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2$$

- O que significam esses termos?

Propriedades Algébricas do MQO

- A interpretação é a seguinte:

$$\underbrace{\sum_{i=1}^{n}(y_i - \overline{y})^2}_{\substack{\text{SQT:} \\ \text{var. total existente em } y_i}} = \underbrace{\sum_{i=1}^{n}(\hat{u}_i - \overline{\hat{u}}_i)^2}_{\substack{\text{SQR:} \\ \text{var. total resíduos}}} + \underbrace{\sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2}_{\substack{\text{SQE:} \\ \text{var. total em } \hat{y}_i \text{(explicada pelo modelo)}}}$$

- A partir da expressão acima é possível então definir:

$$R^2 = \frac{SQE}{SQT} = 1 - \frac{SQR}{SQT}, \quad R^2 \in (0,1)$$

- O R^2 mede quanto da variação em y é explicada por variações em x.

Propriedades Algébricas do MQO

- Acabamos de passar por algumas propriedades algébricas do MQO.
- Importante: estas propriedades valem sempre, são propriedades algébricas:
  - Basta traçar uma reta por MQO, e esta reta as respeitará necessariamente.
- Agora, essas propriedades não nos dizem nada sobre a qualidade desta reta:
  - A reta estimada é uma boa aproximação da reta populacional?
  - Para responder esta pergunta, precisamos de propriedades estatísticas do MQO

Propriedades Estatísticas do MQO: Introdução

- Novamente... as hipóteses suficientes para a estimação por MQO:
  - O modelo populacional existe, com linearidade nos parâmetros:

$$y = \alpha + \beta x + u$$

  - Temos uma amostra aleatória da população para x e y: $\{(x_i, y_i), i = 1, 2, ..., n\}$, tamanho n suficientemente grande e com variação em x_i.
- Com essas duas hipóteses, fomos capazes de traçar a reta por MQO e derivar todas as propriedades algébricas. Em particular, estimamos o coeficiente de inclinação:

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})y_i}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

Exercício: Mostre que $\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{n}(x_i - \overline{x})y_i$

Propriedades Estatísticas do MQO: Introdução

- Agora, vamos entender melhor quem é $\hat{\beta}$. Vamos usar a hipótese 1 acima ao substituir y_i em $\hat{\beta}$:

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})y_i}{\sum_{i=1}^{n}(x_i - \overline{x})^2} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(\alpha + \beta x_i + u_i)}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

- Notem que, ao usar a hipótese populacional, conseguimos incluir termos populacionais na fórmula de $\hat{\beta}$. Continuando:

$$\hat{\beta} = \frac{\alpha \sum(x_i - \overline{x}) + \beta \sum(x_i - \overline{x})x_i + \sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}$$

Onde, pelas propriedades do somatório:
$\alpha \sum(x_i - \overline{x}) = 0$
$\sum(x_i - \overline{x})x_i = \sum(x_i - \overline{x})^2$

Exercício: Mostre que $\sum(x_i - \overline{x})x_i = \sum(x_i - \overline{x})^2$

Propriedades Estatísticas do MQO: Introdução

- Portanto, conseguimos reescrever $\hat{\beta}$ da seguinte forma:

$$\hat{\beta} = \frac{0 + \beta \sum (x_i - \overline{x})^2 + \sum (x_i - \overline{x}) u_i}{\sum (x_i - \overline{x})^2} = \beta + \frac{\sum (x_i - \overline{x}) u_i}{\sum (x_i - \overline{x})^2}$$

- Essa expressão nos diz que conseguimos decompor o nosso estimador $\hat{\beta}$ em duas partes:
  1. O parâmetro populacional β, o qual queremos identificar.
  2. Uma combinação linear do termo de erro u_i: populacional, estocástico e não observado.
- Pergunta: quando teremos $\hat{\beta} = \beta$? Provavelmente nunca.
  - Por que? A cada nova amostra, a realização do vetor de erros será diferente, e a reta traçada também. Além disso, β populacional nunca será observado.

Propriedades Estatísticas do MQO: Introdução

- Notem que $\hat{\beta}$ é uma função de u. Logo, $\hat{\beta}$ é estocástico (pois $\hat{\beta}$ é um estimador).
- E, por isso, a cada nova amostra, obteremos uma nova reta estimada...
  - Ou seja, novas estimativas para seu intercepto e inclinação... que são função exclusivamente dos valores realizados para y e x amostrais: $\{(y_i, x_i), i = 1, 2, ..., n\}$... sendo, por sua vez, $y|x$ função de u.
- Em última instância, se u é uma variável aleatória $\rightarrow y$ é variável aleatória $\rightarrow \hat{\beta}$ será uma variável aleatória. E, portanto, terá:
  - Um valor esperado e uma variância.
  - Na verdade, uma distribuição de probabilidade completa...

Propriedades Estatísticas do MQO: Valor Esperado

- Dada essa introdução, agora podemos avançar na derivação das propriedades estatísticas do MQO. Vamos partir de:

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

- E derivar expressões para o valor esperado $E[\hat{\beta}|X]$ e a variância $V[\hat{\beta}|X]$. Por que $|X$? Neste curso vamos considerar a distribuição de X dada.
- Começaremos pelo valor esperado.
- Vamos simplesmente passar o valor esperado na expressão para $\hat{\beta}$ acima:

$$E[\hat{\beta}|X] = E\left[\beta + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}|X\right]$$

Propriedades Estatísticas do MQO: Valor Esperado

- Pelas propriedades do valor esperado:

$$E[\hat{\beta}|X] = E\left[\beta + \frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}|X\right] = E[\beta|X] + E\left[\frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}|X\right]$$

Onde:

$E[\beta|X] = \beta$

$E\left[\frac{\sum(x_i-\overline{x})u_i}{\sum(x_i-\overline{x})^2}|X\right] =?$

- Sabemos que u é um termo de erro estocástico populacional, e portanto não observado.

Propriedades Estatísticas do MQO: Valor Esperado

- Por enquanto, podemos apenas escrever:

$$E[\hat{\beta}|X] = \beta + E\left[\frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}|X\right] = \beta + \frac{\sum(x_i - \overline{x})E[u_i|X]}{\sum(x_i - \overline{x})^2}$$

- Portanto, para falarmos mais sobre $E[\hat{\beta}|X]$, teremos que, necessariamente, assumir alguma hipótese sobre o termo estocástico da expressão acima.
- Mais especificamente, sobre a relação entre u e x.

Propriedades Estatísticas do MQO: Valor Esperado

- Tínhamos duas hipóteses suficientes à estimação por MQO. Vamos então adicionar mais uma hipótese ao nosso modelo populacional:
  1. O modelo populacional existe, com linearidade nos parâmetros:

     $$y = \alpha + \beta x + u$$

  2. Temos uma amostra aleatória da população para x e y: $\{(x_i, y_i), i = 1, 2, ..., n\}$, tamanho n suficientemente grande e com variação em x_i.
  3. A variável aleatória u é independente de x (hipótese + forte). Essa hipótese implica em cov(u, x) = 0 (hipótese um pouco + fraca).
- Vamos então assumir independência entre u e x e definir, sem perda de generalidade:

  $$E[u] = 0 \quad \text{e} \quad E[u|X] = 0$$

  .

Propriedades Estatísticas do MQO: Valor Esperado

- Portanto, se válidas as hipóteses anteriores, podemos então dar este último passo:

$$E[\hat{\beta}|X] = \beta + E\left[\frac{\sum(x_i - \overline{x})u_i}{\sum(x_i - \overline{x})^2}|X\right] = \beta + \frac{\sum(x_i - \overline{x})E[u_i|X]}{\sum(x_i - \overline{x})^2} = \beta$$

- Em suma, se válidas as hipóteses, podemos dizer que $\hat{\beta}$ é não viesado.
- Novamente, quando poderemos dizer que $\hat{\beta} = \beta$? Provavelmente, nunca. Mas, se válidas as hipóteses enumeradas até aqui, poderemos dizer que $E[\hat{\beta}|X] = \beta$.
- Por isso, em aplicações, é crucial refletirmos sobre a validade de $u \perp x$. Essa hipótese não é necessária para a estimação por MQO. Mas nos garante qualidade da estimação.

Propriedades Estatísticas do MQO: Variância

- O valor esperado é uma parte da história. Agora vejamos a segunda parte, a variância de $\hat{\beta}|X$.
  - Em econometria estamos preocupados em (i) acertar na média ($E[\hat{\beta}|X] = \beta$; sem viés), (ii) com precisão (variância mínima).
- Então voltemos à expressão inicial para $\hat{\beta}$:

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

- Vamos fazer a variância

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

Propriedades Estatísticas do MQO: Variância

- Para simplificar as contas, vamos denotar $(x_i - \overline{x}) = d_i$. Logo, pelas propriedades da variância:

$$V[\hat{\beta}|X] = \left[\frac{1}{\sum d_i^2}\right]^2 V[\sum d_i u_i|X] =?$$

- Desenvolvendo o último termo:

$$V[\sum d_i u_i|X] = V[d_1 u_1 + d_2 u_2 + ... + d_n u_n|X]$$

- Mas, como os u_i são independentes entre si $(\forall\, 1, 2, ..., n,$ assumimos amostra aleatória) $\rightarrow \quad cov(u_i, u_j) = 0$. Pelas propriedades da variância, temos então que $V(\sum d_i u_i) = \sum V(d_i u_i)$. Portanto:

$$V[\hat{\beta}|X] = \left[\frac{1}{\sum d_i^2}\right]^2 \cdot \underbrace{V(d_1 u_1 + d_2 u_2 + ... + d_n u_n|X)}_{d_1^2 V[u_1|X]+...+d_n^2 V[u_n|X]} =?$$

Propriedades Estatísticas do MQO: Variância

- Mais uma vez, paramos no meio do caminho: quem são $V[u_i|X], \quad \forall\, 1, 2, ..., n$?
- Lembrem-se mais uma vez que u é um termo de erro estocástico populacional, e portanto não observado.
- Portanto, para falarmos mais sobre $V[\hat{\beta}|X]$, teremos que, necessariamente, assumir alguma hipótese sobre $V[u|X]$...

Propriedades Estatísticas do MQO: Variância

- Vamos ent ao adicionar mais uma hipótese ao nosso modelo populacional:
  1. O modelo populacional existe, com linearidade nos parâmetros:

     $$y = \alpha + \beta x + u$$

  2. Temos uma amostra aleatória da população para x e y: $\{(x_i, y_i), i = 1, 2, ..., n\}$, tamanho n suficientemente grande e com variação em x_i.
  3. A variável aleatória u é independente de x (hipótese + forte). Essa hipótese implica em cov(u, x) = 0 (hipótese um pouco + fraca).
  4. O termo de erro é homocedástico: $V[u|X] = \sigma^2$ constante, onde σ^2 é um parâmetro populacional.
- Ou seja, vamos assumir que a variância de u é constante para valores distintos de X. O que isso significa, alguma interpretação econômica?

Propriedades Estatísticas do MQO: Variância

- Continuando as contas, assumindo homocedasticidade:

$$V[\hat{\beta}|X] = \left[\frac{1}{\sum d_i^2}\right]^2 \cdot \underbrace{V(d_1u_1 + d_2u_2 + ... + d_nu_n|X)}_{d_1^2V[u_1|X]+...+d_n^2V[u_n|X]}$$

assumindo que $V[u_i|X] = \sigma^2, \quad \forall\, 1, 2, ..., n.$ Logo

$$V[\hat{\beta}|X] = \left[\frac{1}{\sum d_i^2}\right]^2 \sum d_i^2\sigma^2 = \sigma^2 \left[\frac{1}{\sum d_i^2}\right]^2 \cdot \sum d_i^2 = \frac{\sigma^2}{\sum d_i^2}$$

- Ou melhor, sob homocedasticidade:

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

Propriedades Estatísticas do MQO: Variância

- Vejamos algumas propriedades:
  - Quanto maior $V[u|X] = \sigma^2 \Rightarrow$ Menos precisão, pois $V[\hat{\beta}|X]$ maior.
  - Quanto maior a variação em x em $\sum(x_i - \overline{x})^2 \quad \Rightarrow$ Mais precisão, pois $V[\hat{\beta}|X]$ menor.
- Notem que, pelas propriedades da variância:

$$V[u|X] = V[y - \alpha - \beta x|X] = V[y|X] - \underbrace{V[\alpha|X]}_{=0} - \underbrace{V[\beta x|X]}_{=0}$$

  Logo: $V[y|X] = \sigma^2$. A interpretação da homocedasticidade é essencialmente econômica.
- Vocês já são capazes de estimar $V[\hat{\beta}|X]$? Não, para tanto precisaremos estimar o parâmetro populacional σ^2 que veremos na próxima aula..