AULA 8 - MQO em regressão múltipla: Definição, Estimação e Propriedades Algébricas

Econometria I

Área quantitativa - IE/UFRJ

Regressão Múltipla: Definição e Derivação

- A partir de agora vamos alterar o nosso modelo populacional de modo a incorporarmos mais variáveis explicativas em nossa análise.
- Seja, portanto, o modelo:

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + u \Rightarrow \text{linear nos parâmetros}$$

  onde $\alpha + \beta_1 x_1 + \beta_2 x_2$ é termo determinítico, u é termo estocástico independente de X, com $E[u|x_1, x_2] = 0$.
- Ou seja, u não é correlacionado com x_1 e x_2 por hipótese (supondo que o modelo populacional é assim).
- Vamos começar por: interpretação, estimação e propriedades algébricas do MQO múltiplo.

Regressão Múltipla: Definição e Derivação

- Interpretação, preliminares: efeitos de x_1 ou x_2 sobre y

$$\frac{\partial y}{\partial x_1} = \beta_1$$

$$\frac{\partial y}{\partial x_2} = \beta_2$$

- Notem que aqui devemos nos lembrar da interpretação da derivada parcial: o efeito de Δx_1 sobre y é β_1 , ceteris paribus (tudo o mais constante, incluindo x_2)

Estimação por MQO

- Vamos continuar com o modelo populacional

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + u$$

  com $E[u|x_1, x_2] = 0$. Suponha que temos uma amostra aleatória em mãos $\{(y_i, x_{i1}, x_{i2}), i = 1, ..., n\}$.
- Como estimar os parâmetros α, β_1 e β_2 ?
- Os métodos que aprendemos para regressão simples podem ser aplicados aqui também.
- Vejamos o MQO: vamos escolher $\hat{\alpha}$, $\hat{\beta}_1$ e $\hat{\beta}_2$ que minimizem a soma dos resíduos ao quadrado (logo é análogo):

$$\min_{\alpha,\beta_1,\beta_2} S(\alpha, \beta_1, \beta_2) = \sum_{i=1}^{n} (y_i - \alpha - \beta_1 x_{i1} - \beta_2 x_{i2})^2$$

Estimação por MQO

- Agora, no entanto, caímos em um sistema de 3 equações e 3 incógnitas:

$$\frac{\partial S(\alpha, \beta_1, \beta_2)}{\partial \alpha} = 0$$

$$\frac{\partial S(\alpha, \beta_1, \beta_2)}{\partial \beta_1} = 0$$

$$\frac{\partial S(\alpha, \beta_1, \beta_2)}{\partial \beta_2} = 0$$

- Podemos generalizar nosso problema para $k \geq 2$ variáveis independentes. Suponha agora que o modelo populacional é:

$$y = \alpha + \beta_1 x_1 + ... + \beta_k x_k + u$$

onde

$$E[u|x_1, ..., x_k] = 0$$

e temos amostra aleatória para $\{(y_i, x_{1i}, ..., x_{ki}),\ i = 1, ..., n\}$

Estimação por MQO

- Neste caso, temos k variáveis e $k+1$ parâmetros a serem estimados por $\min_{\alpha,\beta_1,...,\beta_k} S(\alpha, \beta_1, ..., \beta_k)$
- Resolvendo um sistema com $k+1$ equações...

$$\frac{\partial S(\cdot)}{\partial \alpha} = -2\sum(y_i - \hat{\alpha} - \hat{\beta}_1 x_{1i} - ... - \hat{\beta}_k x_{ki})$$

$$\frac{\partial S(\cdot)}{\partial \beta_1} = -2\sum x_{1i}(y_i - \hat{\alpha} - \hat{\beta}_1 x_{1i} - ... - \hat{\beta}_k x_{ki})$$

$$\vdots$$

$$\frac{\partial S(\cdot)}{\partial \beta_k} = -2\sum x_{ki}(y_i - \hat{\alpha} - \hat{\beta}_1 x_{1i} - ... - \hat{\beta}_k x_{ki})$$

- Solução do sistema:

$$\hat{\alpha}, \hat{\beta}_1, ..., \hat{\beta}_k \Rightarrow \text{Estimadores de MQO.}$$

Estimação por MQO

- Observações importantes:
  1. Para resolver sistemas lineares $\Rightarrow$ revisão de álgebra linear.
     - Veremos isso com mais detalhes quando revisitarmos a regressão múltipla com notação matricial. Por enquanto, basta a intuição - lembrando que nem sempre um sistema de equações lineares tem solução única.
     - Em particular: não pode existir colinearidade perfeita entre as explicativas. No MQO simples, precisávamos de menos. Agora precisamos impor restrições sobre como essas variáveis $x_1, x_2, ..., x_k$ se relacionam entre si na amostra. Ainda: Na amostra, $n > k + 1$.
  2. Maneiras alternativas de estimar o nosso modelo:
     - Método de momentos: $E[u] = 0$, $E[u\, x_j] = 0$ p/ $j = 1, 2, ..., k$.
     - MV (supondo distribuição conhecida p/ u):

$$\max \alpha, \beta_1, ..., \beta_k \log L(\alpha, \beta_1, ..., \beta_k)$$

Estimação por MQO

- Terminologia análoga á regressão simples:
  - Modelo populacional: $y = \alpha + \beta_1 x_1 + ... + \beta_k x_k + u$
  - Reta de regressão (amostral): $\hat{y} = \hat{\alpha} + \hat{\beta}_1 x_1 + ... + \hat{\beta}_k x_k$ ou, para um dado i: $\hat{y}_i = \hat{\alpha} + \hat{\beta}_1 x_{i1} + ... + \hat{\beta}_k x_{ik}$
  - Resíduos: $\hat{u}_i = y_i - \hat{y}_i$
  - Estimativas: $\hat{\alpha}, \hat{\beta}_1, ..., \hat{\beta}_k$ assumem valores para dada amostra.
- Interpretação dos coeficientes: efeito independente de x_1 em y não importando que valores $x_2, ..., x_k$ assumam
  $\frac{\partial y}{\partial x_1} = \hat{\beta}_1 \Rightarrow$ efeito de x_1 em $\hat{y}$ mantendo-se $x_2, ..., x_k$ fixos ou condicional em $x_2, ..., x_k$ ou controlando por $x_2, ..., x_k$.

Propriedades Algébricas do MQO

- Assim como no caso simples, temos:
  $\hat{u}_i = y_i - \hat{y}_i \Rightarrow \overline{\hat{u}} = \frac{\sum y_i - \hat{y}_i}{n} = 0 \Rightarrow \overline{y}_i = \overline{\hat{y}}_i$
- Os valores estimados de MQO e os resíduos têm algumas propriedades que são extensões imediatas do caso simples:
  1. A média amostral dos resíduos é zero.
  2. $Cov(x_j, u) = 0 \, \forall \, j = 1, ..., k$. Consequentemente, a covarância amostral entre os valores estimados de MQO e os resíduos de MQO é zero.
  3. O ponto $(\overline{y}, \overline{x}_1, \overline{x}_2, ..., \overline{x}_k)$ está sob a reta de regressão MQO: $\overline{y} = \hat{\alpha} + \hat{\beta}_1 \overline{x}_1 + \hat{\beta}_2 \overline{x}_2 + \hat{\beta}_k \overline{x}_k$

  As duas primeiras propriedades são consequências imediatas do conjunto de equações usadas para obter as estimativas de MQO. Vimos que no sistema acima a primeira equação diz que a soma dos resíduos é zero. As equações restantes são da forma $\sum x_{ij} \hat{u}_i = 0$, implicando que cada variável independente tem cov amostral zero com $\hat{u}_i$.
  A propriedade 3 decorre imediatamente da propriedade 1.

Propriedades Algébricas do MQO

- Grau de ajuste: assim como na regressão simples, podemos definir:

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

- A partir da expressão acima é possível então definir:

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

é interpretado como a proporção da variação amostral em y_i que é explicada pela reta de regressão de MQO.

Propriedades Algébricas do MQO

- Um fato importante sobre R^2 é que ele nunca diminui, e geralmente aumenta, quando outra variável independente é adicionada à regressão.
- Esse fato algébrico ocorre por definição, pois a soma dos resíduos quadrados nunca aumenta quando regressores adicionais são acrescentados ao modelo.
- Isso faz com que R^2 seja um instrumento fraco para decidir se uma variável ou diversas variáveis deveriam ser adicionadas ao modelo.
- O fato que deve determinar se uma variável explicativa pertence a um modelo é se a variável explicativa tem, na população um efeito parcial sobre y diferente de zero (veremos mais tarde em inferência estatística).

R^2 ajustado

- Com a inclusão de variáveis explicativas sabemos que R^2 ńunca diminui.
- Para contornarmos esse problema, podemos considerar um R^2 alternativo, chamado R^2 ajustado, denotado por $\bar{R}^2$.
- O termo ajustado significa ajustado pelos graus de liberdade associados á soma de quadrados:

$$\bar{R}^2 = 1 - \frac{SQR/(n-k)}{SQT/(n-1)} = 1 - \frac{SQR(n-1)}{SQT(n-k)}$$

  em que k $=$ número de parâmetros do modelo, incluindo o termo de intercepto.
- Sendo $R^2 = 1 - \frac{SQR}{SQT}$, então podemos escrever

$$\bar{R}^2 = 1 - (1 - R^2)\frac{(n-1)}{(n-k)}$$