AULA 3 - MQO Simples: Estimação

Econometria I

Área quantitativa - IE/UFRJ

Estimação: Preliminares

- Existem distintos métodos de estimação em econometria, por exemplo:
  - Método de Momentos.
  - Método dos Mínimos Quadrados Ordinários, MQO.
  - Método da Maximaverossimilhança (modelos não lineares, necessária hipótese sobre distribuição).
  - Método de Momentos Generalizado (método geral, métodos acima são casos particulares).
- Comecemos com o método de momentos, e logo veremos que o MQO é análogo.
- Vejamos a ideia geral, antes de começarmos com a álgebra...

Estimação: Método de momentos

- Existe uma distribuição conjunta das variáveis (X, Y) na população.
- Usamos então nossas hipóteses populacionais para restringir essa relação, e derivar uma estratégia de estimação:
  - Supondo então que valem $E(u) = 0$ e $E(u|X) = 0$, onde $E(u|X) = 0 \Rightarrow Cov(X, u) = 0$. Reescrevendo essas hipóteses:
    1. $E(u) = 0 \Rightarrow E(Y - \alpha - \beta X) = 0$
    2. $Cov(X, u) = 0 \Rightarrow E(Xu) - E(X)E(u) = 0 \Rightarrow$ $E(X[Y - \alpha - \beta X]) = 0$
  - Para chegar às 2 equações acima, usamos a hipótese sobre a relação linear entre Y e X (ao fazer $u = Y - \alpha - \beta X$) e as hipóteses sobre u.
- Temos então 2 equações e 2 incógnitas (α e β)...

Estimação: Método de momentos

- Agora vamos usar a terceira hipótese, de que temos uma amostra aleatória para (X, Y)...
- Dada uma amostra de dados, podemos reescrever as 2 equações anteriores por equivalência amostral. Ou seja, se α e β são solução para o sistema de equações populacionais, escolhemos as estimativas $\hat{\alpha}$ e $\hat{\beta}$ como solução do sistema

  1. $E(u) = 0 \Rightarrow \sum_{i=1}^{n} \frac{(y_i - \alpha - \beta x_i)}{n} = 0$
  2. $Cov(X, u) = 0 \Rightarrow \sum_{i=1}^{n} \frac{x_i(y_i - \alpha - \beta x_i)}{n} = 0$

- Note então que temos novamente um sistema com 2 equações e 2 incógnitas. Lembrando que observamos (X_i, Y_i), $\forall i = 1, 2, ..., n$.

Estimação: Método de momentos

- Resolvendo... da 1a equação:

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

- Usando a seguinte notação para as médias amostrais $\frac{y_i}{n} = \overline{y}$ e $\frac{x_i}{n} = \overline{x}$, e rearranjando os termos, encontramos

$$\overline{y} - \hat{\alpha} - \hat{\beta}\overline{x} = 0 \Rightarrow \hat{\alpha} = \overline{y} - \hat{\beta}\overline{x}$$

- Substituindo esse resultado no lugar de $\hat{\alpha}$ na 2a equação do sistema

Estimação: Método de momentos

- Passo a passo... primeiro, substituindo:

$$\sum_{i=1}^{n} \frac{x_i[y_i - \alpha - \beta x_i]}{n} = 0 \Rightarrow \sum_{i=1}^{n} x_i[y_i - (\overline{y} - \hat{\beta}\overline{x}) - \hat{\beta}x_i] = 0$$

- Depois, rearranjando de modo a isolar β como função apenas de x e y:

$$\sum_{i=1}^{n} x_i[y_i - (\overline{y} - \hat{\beta}\overline{x}) - \hat{\beta}x_i] = 0 \Rightarrow \sum_{i=1}^{n} x_i(y_i - \overline{y}) = \hat{\beta}\sum_{i=1}^{n} x_i(x_i - \overline{x})$$

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

- E pelas propriedades do somatório...

Estimação: Método de momentos

- Podemos escrever...

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

  ver apêndice A7 e A8 do Wooldridge

- Por fim, basta substituir a expressão acima em $\hat{\alpha} = \overline{y} - \hat{\beta}\overline{x}$ para encontrar

$$\hat{\alpha} = \overline{y} - \frac{Cov(x,y)}{Var(x)}\overline{x}$$

- Pronto, reta estimada.
- Agora, algumas observações operacionais importantes... $\Rightarrow$ Como Var(x) é sempre positiva, o sinal de β será o mesmo de Cov(x, y).

Estimação: Método de momentos

- Algumas observações operacionais importantes... (cont.)
  - Precisamos de variação populacional (e amostral) em X, caso contrário o denominador de β é zero, e não conseguimos identificá-lo.
  - Naturalmente, para conseguir estimar a reta, precisamos também de um tamanho mínimo de amostra: conseguimos traçar uma reta com apenas um ponto no R^2? Com quantos pontos se traça uma reta?
  - Para montar o sistema, partimos das hipóteses populacionais $E(u) = 0$ e $E(u|X) = 0$. No entanto, para estimar uma reta, a partir de uma amostra, não precisamos que essas hipóteses sejam válidas. Precisamos apenas de pelo menos duas observações e $Var(x) > 0$... Para fixar: traçar uma reta é uma questão algébrica. A validade das hipóteses afeta a interpretação dos coeficientes estimados!
  - Isso ficará mais claro no MQO...

Estimação: MQO

- Vejamos agora a ideia por trás da estimação por Mínimos Quadrados Ordinários, MQO.
- Voltemos ao nosso modelo de regressão

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

- Onde o par (x_i, y_i) é observado para todo i da nossa amostra de dados de tamanho n.
- Já vimos que essa amostra pode ser representada por uma nuvem de pontos no R^2 .
- Já vimos também que ao traçarmos uma reta qualquer por essa nuvem, $\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i$ derivaremos um vetor de resíduos de dimensão $n \times 1$. Podemos escrever, para a observação i, $\hat{u}_i = y_i - \hat{y}_i$

Estimação: MQO

- Para cada reta traçada, temos um vetor de resíduos $\hat{u}$ distinto.
- É bastante intuitivo assumir que a soma de todos esses resíduos (desvios dos pontos à reta estimada) represente o nosso grau de incompetência em traçar uma reta que se ajuste bem à nuvem de pontos.
- O método por MQO tem como estratégia encontrar $\hat{\alpha}$ e $\hat{\beta}$ tal que a reta resultante esteja bem ajustada à nuvem de pontos. Isso ocorrerá caso a distância dos y_i observados e dos $\hat{y}_i$ ajustados seja pequena.
- Logo, pelo método por MQO, escolhemos $\hat{\alpha}$ e $\hat{\beta}$ que minimizam a soma dos resíduos ao quadrado (já que desvios positivos e negativos tendem a se anular na soma).

Estimação: MQO

- Formalmente, os estimadores de MQO são a solução para o problema

$$\min_{\alpha,\beta} S(\alpha,\beta) = \sum_{i=1}^{n} (\hat{u}_i)^2$$

- Da regressão populacional: $y_i = \alpha + \beta x_i + u_i$
- Para a amostra: $y_i = \hat{y}_i + \hat{u}_i$ sendo $\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i$

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

- Este é um programa de minimização. Para resolvê-lo basta

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

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

Estimação: MQO

- Note, por fim, que estas condições de 1a ordem (um sistema de 2 equações e 2 incógnitas) são idênticas às equações que montamos anteriormente, quando resolvemos pelo método de momentos.
- Portanto, a solução é a mesma, e os estimadores são análogos. Ao resolver novamente o sistema, encontraremos:

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

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

- Já sabemos então como estimar $\hat{\alpha}$ e $\hat{\beta}$ por MQO para o caso simples.

Estimação: Exemplo

- Vamos treinar um pouco essa álgebra com nosso exemplo de 4 dados (renda e educação).
  - Vamos começar com a mesma distribuição para $X = \{2, 6, 10, 14\}$.
  - Neste mundo, vamos criar o modelo populacional linear $Y = 1000 + 160X + u$
  - E vamos sortear um novo vetor u, por exemplo:

$$u = \{914, -265, 991, -107\}$$

  - Usando o modelo $Y = 1000 + 160X + u$, encontramos:

$$Y = \{2234, 1694, 3592, 3133\}$$

- O problema: não observamos u. Devemos traçar uma reta através dos pontos definidos pelos 4 pares (X, Y) observados.

Estimação: Exemplo

- Nossa amostra, portanto, segue em vermelho na tabela abaixo. Vamos usar as fórmulas para estimar o modelo...

Indivíduo	Educação	Renda	A: Educ - Média(Educ)	B: Renda - Média(Renda)	A*B	A^2
1	2	2234	-6	-429	2575	36
2	6	1694	-2	-969	1938	4
3	10	3592	2	929	1857	4
4	14	3133	6	470	2817	36
Média	8	2663				
Soma					9187	80

- É fácil ver que $\hat{\beta} = \frac{A*B}{A^2} = \frac{9187}{80} = 114,8$. Substituindo na fórmula para $\hat{\alpha} = 2663 - 8 * 114,8 = 1744,6$