Econometria I

Mínimsos Quadrados Restrito

Victor Oliveira
Núcleo de Economia Internacional e Desenvolvimento Econômico

PPGDE – 2024

Sumário I

Demanda por Moeda

- Da revisão de diversas teorias sobre a demanda por moeda podemos sintetizar a seguinte modelo econométrico:

$$\frac{M^D}{P} = \beta_0 + \beta_1 \boldsymbol{y} + \beta_2 \boldsymbol{r} + \beta_3 \boldsymbol{P}^* + e_i \tag{1}$$

  em que M^D é o valor nominal da quantidade demandada por moeda, P é o nível geral de preços (que é o índice geral de preços com base 1), $\frac{DM}{P}$ é a demanda por saldos reais de moeda, r é a taxa de juros, y é o produto real e P^* é a taxa de inflação.
- Espera-se que $\beta_1 > 0$, $\beta_2 < 0$ e $\beta_3 < 0$.

Restrições

1. $\beta_2 = 0$: demanda por moeda para os clássicos – motivo transações e precaução

$$\frac{M^D}{P} = \beta_0 + \beta_1 \boldsymbol{y} + \beta_3 \boldsymbol{P}^* + e_i \tag{2}$$

2. $\beta_2 \neq 0$: demanda por moeda para os keynesianos – motivo especulação

$$\frac{M^D}{P} = \beta_0 + \beta_1 \boldsymbol{y} + \beta_2 \boldsymbol{r} + \beta_3 \boldsymbol{P}^* + e_i \tag{3}$$

- **Empiricamente podemos visualizar essas duas situações em momento em que a economia está passando por um período de inflação baixa e inflação elevada.**

- No modelo de projeção linear

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + e_i \tag{4}$$
$$\mathbb{E}(\boldsymbol{x}_i e_i) = 0 \tag{5}$$

- Impor uma restrição sobre o vetor de coeficientes $\boldsymbol{\beta}$ é um exercício comum.
- Podemos particionar $\boldsymbol{x}_i' = (\boldsymbol{x}_{1i}', \boldsymbol{x}_{2i}')$ e $\boldsymbol{\beta}' = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')$ e temos uma típica restrição de exclusão da forma $\boldsymbol{\beta}_2 = 0$.
- Neste caso o modelo é:

$$y_i = \boldsymbol{x}_{1i}'\boldsymbol{\beta}_1 + e_i \tag{6}$$
$$\mathbb{E}(\boldsymbol{x}_i e_i) = 0 \tag{7}$$

Matriz de Restrições

- À primeira vista, esta última equação parecer ser a mesma como a do modelo de projeção linear, mas há uma importante diferença: **o termo de erro e_i é não correlacionado com o vetor de regressores $\boldsymbol{x}_i' = (\boldsymbol{x}_{1i}', \boldsymbol{x}_{2i}')$ e não apenas o regressor incluído $\boldsymbol{x}_{1i}'$.**
- De uma forma geral, o conjunto q de restrições lineares toma a forma de

$$\boldsymbol{R}'\boldsymbol{\beta} = \boldsymbol{c} \tag{8}$$

em que $\boldsymbol{R}$ é uma matriz $(k \times q)$, $\text{rank}(\boldsymbol{R}) = q$ e $\boldsymbol{c}$ é um vetor $(q \times 1)$. A matriz $\boldsymbol{R}$ "codifica" as hipóteses a serem testadas. Cada linha corresponde a uma restrição linear sobre o vetor β.

- O pressuposto de que $\boldsymbol{R}$ tem *rank* completo significa que as restrições são linearmente independentes.

- A restrição $\beta_2 = 0$ é um caso especial com

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{I} \end{bmatrix} \tag{9}$$

em que a matriz $\boldsymbol{R}$ "codifica" as hipóteses a serem testadas e $\boldsymbol{c} = 0$.

- Outras restrições poderiam assumir a forma de $\beta_1 + \beta_2 = 1$. Esse é o caso de testarmos retornos de escala constante.
- Poderíamos ter restrições testando a igualdade de coeficiente: $\beta_1 = \beta_2$ ou também uma restrição como $\beta_1 = -\beta_2$.
- Vamos ver alguns exemplos de restrições que podem ser implementadas no contexto de analise de regressão.

- Alguns exemplos de hipóteses que podemos testar:

$$H_0\colon \beta_2 = 0 \tag{10}$$

$$H_0\colon \beta_2 = -1 \tag{11}$$

$$H_0\colon \beta_2 + \beta_3 = 1 \tag{12}$$

$$H_0\colon \beta_2 = \beta_4 \tag{13}$$

$$H_0\colon \beta_2 - \beta_4 = 0 \tag{14}$$

$$H_0\colon \beta_2 = \beta_3 = \beta_4 = \cdots = \beta_k = 0 \tag{15}$$

$$H_0\colon \beta_2 = \beta_3 = 0 \tag{16}$$

- Teríamos as seguintes representações para $\boldsymbol{R}$ e $\boldsymbol{c}$:

$$H_0\colon \beta_2 = 0 \implies \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix} = 0 \tag{17}$$

$$H_0\colon \beta_2 = -1 \implies \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix} = -1 \tag{18}$$

$$H_0\colon \beta_2 + \beta_3 = 1 \implies \begin{bmatrix} 0 & 1 & 1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix} = 1 \tag{19}$$

$$H_0\colon \beta_2 - \beta_3 = 0 \implies \begin{bmatrix} 0 & 1 & -1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix} = 0 \tag{20}$$

$$H_0\colon \beta_2 = \beta_3 = \cdots = \beta_k = 0 \implies \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & . & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{21}$$

- Por que impor uma restrição?
  - Acreditar ou ter a informação de que a restrição é verdadeira
  - Esperamos que melhore a eficiência da estimação
  - Obter estimativas consistentes com variância reduzida em relação ao estimador não restrito
- **Questões**
  1. Como deveríamos estimar o vetor de coeficientes $\boldsymbol{\beta}$ impondo a restrição linear?
  2. Ao impormos essa restrição, qual é a distribuição amostral do estimador resultante?
  3. Como calcular o desvio-padrão?

Mínimos Quadrados Restrito

- Um método intuitivamente atraente para estimar uma projeção linear restrita é minimizar o critério de mínimos quadrados sujeito a restrição $\boldsymbol{R}'\boldsymbol{\beta} = c$. Este estimador é:

$$\widetilde{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{R}'\boldsymbol{\beta}=\boldsymbol{c}} SQE_n(\boldsymbol{\beta}) \tag{22}$$

em que

$$SQE_n(\boldsymbol{\beta}) = \sum_{i=1}^{n}(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2 = \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{y}'\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}. \tag{23}$$

- O estimador $\widetilde{\boldsymbol{\beta}}$ minimiza a soma dos erros ao quadrados sobre todos $\boldsymbol{\beta}$ tal que a restrição (8) é mantida. Chamamos $\widetilde{\boldsymbol{\beta}}$ ou $\widetilde{\boldsymbol{\beta}}_{cls}$ de **estimador de mínimos quadrados restrito**.

- O método utilizado para encontrar a solução para a eq. (23) utiliza a técnica do Multiplicador de Lagrange. O problema (23) é semelhante a minimização do lagrangiano

$$L(\boldsymbol{\beta}, \boldsymbol{\lambda}) = \frac{1}{2} SQE_n(\boldsymbol{\beta}) + \lambda'(\boldsymbol{R'\beta - c}) \tag{24}$$

  sobre $(\boldsymbol{\beta}, \boldsymbol{\lambda})$, em que λ é um vetor $s \times 1$ do multiplicador de Lagrange.

- As CPOs para (24) são:

$$\frac{\partial}{\partial \boldsymbol{\beta}} L(\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\lambda}}) = 0 \quad \Longleftrightarrow \quad -\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{X}'\boldsymbol{X}\widetilde{\boldsymbol{\beta}} + \boldsymbol{R}\widetilde{\boldsymbol{\lambda}} = \mathbf{0} \tag{25}$$

$$\frac{\partial}{\partial \boldsymbol{\lambda}} L(\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\lambda}}) = 0 \quad \Longleftrightarrow \quad \boldsymbol{R}'\widetilde{\boldsymbol{\beta}} - \boldsymbol{c} = \mathbf{0} \tag{26}$$

- Se pré-multiplicarmos a eq. (25) por $\boldsymbol{R}'(\boldsymbol{X}'\boldsymbol{X})^{-1}$, obtemos

$$-\boldsymbol{R}'\widehat{\boldsymbol{\beta}} + \boldsymbol{R}'\widetilde{\boldsymbol{\beta}} + \boldsymbol{R}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}\widetilde{\boldsymbol{\lambda}} = \mathbf{0} \tag{27}$$

em que $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ é o estimador de mínimos quadrados não restrito. Impondo $\boldsymbol{R}'\widetilde{\boldsymbol{\beta}} - \boldsymbol{c} = \mathbf{0}$ e resolvendo para $\widetilde{\boldsymbol{\lambda}}$ encontramos

$$\widetilde{\boldsymbol{\lambda}} = \left[\boldsymbol{R}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}\right]^{-1}\left(\boldsymbol{R}'\widehat{\boldsymbol{\beta}} - \boldsymbol{c}\right) \tag{28}$$

- Como $(\boldsymbol{X}'\boldsymbol{X})^{-1} > 0$ e $\boldsymbol{R}$ possui rank completo, temos que: $\boldsymbol{R}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R} > 0$. Substituindo essa expressão em (25) e resolvendo para $\widetilde{\boldsymbol{\beta}}$ encontramos a solução para o problema de minimização dada pela eq. (22)

$$\widetilde{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}} - (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}\left[\boldsymbol{R}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{R}\right]^{-1}\left(\boldsymbol{R}'\widehat{\boldsymbol{\beta}} - \boldsymbol{c}\right) \tag{29}$$

- A eq. (29) é a fórmula geral para o estimador de mínimos quadrados restrito. Ela ainda pode ser escrita como

$$\widetilde{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{Q}}_{\boldsymbol{xx}}^{-1}\boldsymbol{R}\left[\boldsymbol{R}'\widehat{\boldsymbol{Q}}_{\boldsymbol{xx}}^{-1}\boldsymbol{R}\right]^{-1}\left(\boldsymbol{R}'\widehat{\boldsymbol{\beta}} - \boldsymbol{c}\right) \tag{30}$$

- Dado $\widetilde{\boldsymbol{\beta}}$ os resíduos são:

$$\widetilde{e}_i = y_i - \boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}} \tag{31}$$

- O estimador dos momentos de σ^2 é:

$$\widetilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\widetilde{e}_i^2 \tag{32}$$

Restrição de Exclusão

- A eq. (29) é a fórmula geral para o estimador de mínimos quadrados restrito. Na maioria dos casos pode ser obtido aplicando Mínimos Quadrados para uma equação reparametrizada. Para ver isso, considere o modelo não restrito

$$y_i = \boldsymbol{x}'_{1i}\boldsymbol{\beta}_1 + \boldsymbol{x}'_{2i}\boldsymbol{\beta}_2 + e_i \tag{33}$$

- A restrição de exclusão é $\boldsymbol{\beta}_2 = \mathbf{0}$ e a equação restrita é:

$$y_i = \boldsymbol{x}'_{1i}\boldsymbol{\beta}_1 + e_i \tag{34}$$

- Neste cenário o estimador de mínimos quadrados restrito é o estimador de mínimos quadrados de y_i em x_{1i}. Assim,

$$\widetilde{\boldsymbol{\beta}}_1 = \left(\sum_{i=1}^{n} \boldsymbol{x}_{1i}\boldsymbol{x}'_{1i}\right)^{-1}\left(\sum_{i=1}^{n} \boldsymbol{x}_{1i}\boldsymbol{y}_i\right) \tag{35}$$

- O estimador de mínimos quadrados restrito do vetor inteiro $\boldsymbol{\beta}' = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')$ é:

$$\widetilde{\boldsymbol{\beta}} = \begin{pmatrix} \widetilde{\boldsymbol{\beta}}_1 \\ \mathbf{0} \end{pmatrix} \tag{36}$$

- A eq. (30) e (36) são equivalentes. Para ver isso, seja

$$\widetilde{\boldsymbol{\beta}}_1 = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} \end{pmatrix} \left[\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{Q}}_{\boldsymbol{xx}}^{-1} \begin{pmatrix} \mathbf{0} \\ \boldsymbol{I} \end{pmatrix} \left[\begin{pmatrix} \boldsymbol{I} & \mathbf{0} \end{pmatrix} \widehat{\boldsymbol{Q}}_{\boldsymbol{xx}}^{-1} \begin{pmatrix} \mathbf{0} \\ \boldsymbol{I} \end{pmatrix} \right]^{-1} \begin{pmatrix} \boldsymbol{I} & \mathbf{0} \end{pmatrix} \widehat{\boldsymbol{\beta}} \right] \tag{37}$$

- Usando

$$\begin{aligned} \widehat{\boldsymbol{Q}}_{\boldsymbol{xx}}^{-1} &= \begin{pmatrix} \widehat{\boldsymbol{Q}}_{11} & \widehat{\boldsymbol{Q}}_{12} \\ \widehat{\boldsymbol{Q}}_{21} & \widehat{\boldsymbol{Q}}_{22} \end{pmatrix}^{-1} \stackrel{\text{def}}{=} \begin{pmatrix} \widehat{\boldsymbol{Q}}^{11} & \widehat{\boldsymbol{Q}}^{12} \\ \widehat{\boldsymbol{Q}}^{21} & \widehat{\boldsymbol{Q}}^{22} \end{pmatrix} \\ &= \begin{pmatrix} \widehat{\boldsymbol{Q}}_{11.2}^{-1} & -\widehat{\boldsymbol{Q}}_{11.2}^{-1} \widehat{\boldsymbol{Q}}_{12} \widehat{\boldsymbol{Q}}_{22}^{-1} \\ -\widehat{\boldsymbol{Q}}_{22.1}^{-1} \widehat{\boldsymbol{Q}}_{21} \widehat{\boldsymbol{Q}}_{11}^{-1} & \widehat{\boldsymbol{Q}}_{22.1}^{-1} \end{pmatrix} \end{aligned} \tag{38}$$

- Obtemos

$$
\begin{aligned}
\widetilde{\beta}_1 &= \widehat{\beta}_1 - \widehat{Q}^{12}\left(\widehat{Q}^{22}\right)^{-1}\widehat{\beta}_2 \\
&= \widehat{\beta}_1 + \widehat{Q}_{11.2}^{-1}\widehat{Q}_{12}\widehat{Q}_{22}^{-1}\widehat{Q}_{22.1}\widehat{\beta}_2 \\
&= \widehat{Q}_{11.2}^{-1}(\widehat{Q}_{1y} - \widehat{Q}_{12}\widehat{Q}_{22}^{-1}\widehat{Q}_{2y}) \\
&+ \widehat{Q}_{11.2}^{-1}\widehat{Q}_{12}\widehat{Q}_{22}^{-1}\widehat{Q}_{22.1}\widehat{Q}_{22.1}^{-1}(\widehat{Q}_{2y} - \widehat{Q}_{21}\widehat{Q}_{11}^{-1}\widehat{Q}_{1y}) \\
&= \widehat{Q}_{11.2}^{-1}(\widehat{Q}_{1y} - \widehat{Q}_{12}\widehat{Q}_{22}^{-1}\widehat{Q}_{21}\widehat{Q}_{11}^{-1}\widehat{Q}_{1y}) \\
&= \widehat{Q}_{11.2}^{-1}(\widehat{Q}_{11} - \widehat{Q}_{12}\widehat{Q}_{22}^{-1}\widehat{Q}_{21})\widehat{Q}_{11}^{-1}\widehat{Q}_{1y} \\
&= \widehat{Q}_{11}^{-1}\widehat{Q}_{1y} \qquad (39)
\end{aligned}
$$

que é igual a eq. (36) como desejamos.

- Há duas abordagens para tratar de mínimos quadrados restrito.

Abordagem (1)

1. *Estimam-se os parâmetros.*
2. *Verifica-se se tais estimativas estão muito longe de satisfazer determinadas restrições.*

Abordagem (2)

1. *Impõem-se as restrições lineares diretamente, estimando-se uma **regressão restrita**.*
2. *Estima-se a **regressão não restrita (irrestrita)** usual.*
3. *Comparam-se os resultados das **regressões restrita** e **não restrita**, de modo a verificar se a perda de ajuste ocasionada pela imposição das restrições deve-se a amostragem ou realmente indica que as restrições não são válidas.*

- **Questão**: quais das duas abordagens deve prevalecer?
- As duas abordagens são equivalentes. A opção por uma ou outra dependerá da situação:
  - Para testar uma única restrição ou a "*significância global*" da regressão, a primeira abordagem é preferível.
  - Para testar várias restrições, a segunda abordagem é geralmente preferível.
- A estatística de teste a ser utilizada nas aplicações é:

$$F = \frac{\dfrac{(\widetilde{\boldsymbol{e}}'\widetilde{\boldsymbol{e}} - \widehat{\boldsymbol{e}}'\widehat{\boldsymbol{e}})}{q}}{\dfrac{(\widehat{\boldsymbol{e}}'\widehat{\boldsymbol{e}})}{(n-k)}} \sim F(q, n-k) \tag{40}$$

- Muitas vezes, o processo de identificação-estimação-diagnóstico conduz não a um, mas a uma lista de possíveis modelos.
- Para selecionar entre tais modelos, pode-se usar critérios de informação, que fornecem medidas de ajuste dos modelos que penalizam o aumento do número de regressores.
- Os critérios de informação mais populares são:
  - Critério de Akaike (AIC): $AIC(k) = n\ln(SQR) + 2k$
  - Critério de Shwarz (SIC): $n\ln(SQR) + k\ln(n)$

  em que n é o número de observações, k é o número de parâmetros estimados e **SQR** é a soma dos quadrados dos resíduos.
- Deve-se escolher o modelo com os menores AIC e SIC.

- Considere os seguintes modelos de regressão:

$$\boldsymbol{y} = \boldsymbol{X}_1\beta_1 + \boldsymbol{X}_2\beta_2 + \boldsymbol{e} \tag{41}$$

$$\boldsymbol{y} = \boldsymbol{X}_1\beta_1 + \boldsymbol{e} \tag{42}$$

$$\boldsymbol{y} = \boldsymbol{X}_1\beta_1 + \boldsymbol{X}_3\beta_3 + \boldsymbol{e} \tag{43}$$

- Se $\beta_2 = 0 \Longrightarrow$ Modelo (2)
- Se $\beta_3 = 0 \Longrightarrow$ Modelo (2)
- A seleção dos modelos pode ser realizada por meio de um **teste de hipótese**.
- É o caso de 1 e de 2 a seguir:
  - Modelo (1) e (2) são *nested*
  - Modelo (3) e (2) são *nested*
  - Modelo (1) e (3) não são *nested*

- Considerando a situação dos modelos não *nested*, podemos adotar um modelo híbrido:

$$\boldsymbol{y} = \boldsymbol{X}_1\beta_1 + \boldsymbol{X}_2\beta_2 + \boldsymbol{X}_3\beta_3 + \boldsymbol{e} \tag{44}$$

- O modelo (43) é um caso especial do modelo (44).
- Às vezes não é possível fazer um modelo híbrido dado a forte correlação entre as variáveis explicativas.
- Para comparar os modelos (1) e (2) posso usar o $\overline{\boldsymbol{R}}^2$.
- Para comparar os modelos (1) e (3) posso fazer uso dos critérios de informação.

Questão: Qual dos dois critérios utilizar?

1. Ao comparar modelos, deve-se manter n fixo
2. O SIC leva a modelos mais parcimoniosos
3. O SIC é "assintoticamente consistente"
4. O AIC é "assintoticamente viesado" (sobreparametrização)

Econometria I

Mínimsos Quadrados Restrito

Victor Oliveira
Núcleo de Economia Internacional e Desenvolvimento Econômico

PPGDE – 2024