Questões Adicionais quanto ao Uso do MQO com Dados de Séries Temporais Correlação Serial e Heterocedasticidade em Regressão de Séries Temporais

Profa. R. Ballini

Bibliografia Básica:

Wooldridge (2002) *Introductory Econometric*, Cap. 11 e 12.

Greene (2012). *Econometric Analysis*, Cap. 20.

Processo Estocástico Estacionário:

O processo estocástico $\{x_t : t = 1, 2 \ldots\}$ é **estacionário** se, para todas as coleções de índices temporais $1 \leq t_1 < t_2 < \ldots < t_m$, a distribuição conjunta de $(x_{t_1}, x_{t_2}, \ldots x_{t_m})$ é a mesma que a distribuição conjunta de $(x_{t_1+h}, x_{t_2+h}, \ldots, x_{t_m+h})$ para todos os inteiros $h \geq 1$.

Processo Fracamente Estacionário

Um processo estocástico é dito **fracamente estacionário** se

$$E[y_t] = \mu, \forall t \in T$$

$$E[(y_t - \mu)]^2 = \sigma^2, \text{para } h = 0$$

$$E[(y_t - \mu)(y_{t-h} - \mu)] = \gamma_h, \forall t, h \in T$$

Série Temporal Fracamente Dependente

Um processo estacionário de série temporal $\{y_t : t = 1, \ldots, n\}$ é chamado de fracamente dependente se y_t e y_{t+h} forem "quase independentes" enquanto h aumenta sem limites.

Seqüências com covariância estacionária em que

$$Cov(y_t, y_{t+h}) \to 0$$

conforme $h \to \infty$ são chamadas de assintoticamente não correlacionadas.

Autocovariância e Autocorrelação

Seja uma série temporal $y_1,\ y_2,\ \ldots,\ y_N$, de N observações. A estimativa dos coeficientes de autocovariância é obtida por:

$$c_k = \frac{1}{N-k}\sum_{t=1}^{N-k}(y_t-\overline{y})(y_{t+k}-\overline{y}),\ k=0,\ 1,\ \ldots,\ K \tag{1}$$

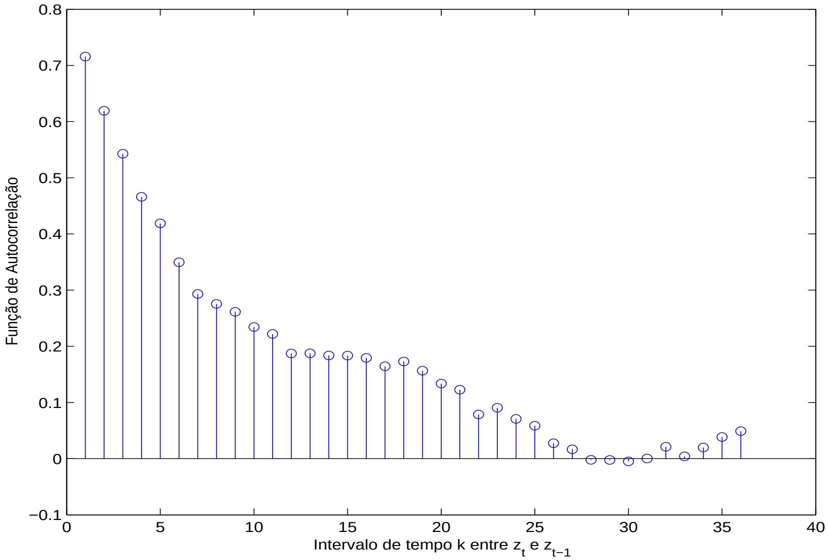
sendo $\overline{y}$ é a média amostral da série temporal.

O coeficiente de autocorrelação é obtido da seguinte forma:

$$\hat{\rho}_k = \frac{c_k}{c_0} \tag{2}$$

em que c_0 é a variância da série y_t.

Exemplo de Série Temporal Fracamente Dependente



- Pressupõe estabilidade ao longo do tempo, no sentido de que β_j não muda ao longo do tempo

- Propriedade de estacionaridade: fundamental para realizarmos inferências estatísticas com base na realização de um processo estocástico.

Teorema do Limite Central (TLC) para dados de Séries Temporais:

Requer estacionaridade e dependência fraca

Séries Temporais Altamente Persistentes (Fortemente Dependentes)

Suponha o seguinte modelo:

$$y_t = \rho_1 y_{t-1} + e_t, \quad t = 1, 2, \ldots$$

$|\rho_1| < 1$ é necessário para que a série seja fracamente dependente. Este modelo é chamado de auto-regressivo de ordem 1 (AR(1)).

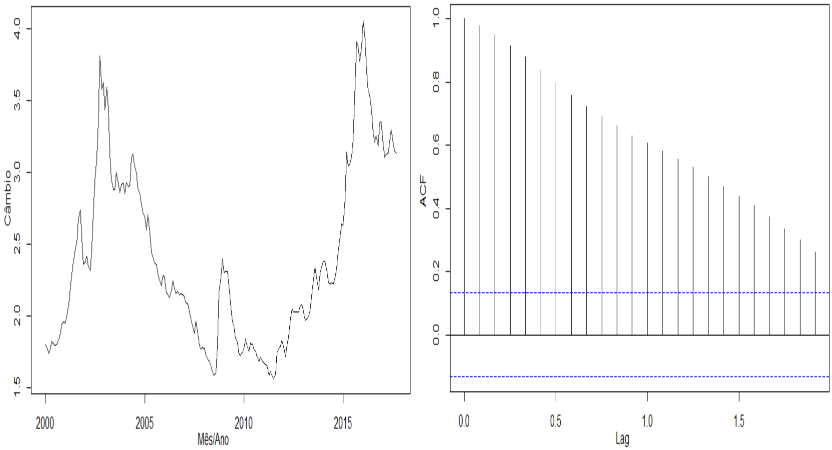
Muitas séries econômicas são caracterizadas com $\rho_1 = 1$:

$$y_t = y_{t-1} + e_t, \quad t = 1, 2, \ldots \tag{3}$$

em que $\{e_t\}$ é i.i.d com $E(e_t) = 0$, $Var(e_t) = \sigma_e^2$; o valor inicial y_0 é independente de e_t para todo $t \geq 1$.

O processo (3) é chamado de **passeio aleatório** (**random walk**).

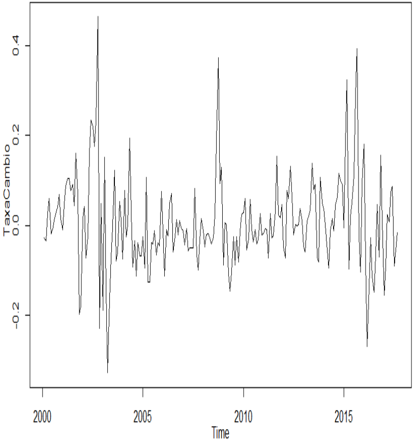
Exemplo: Taxa de Câmbio

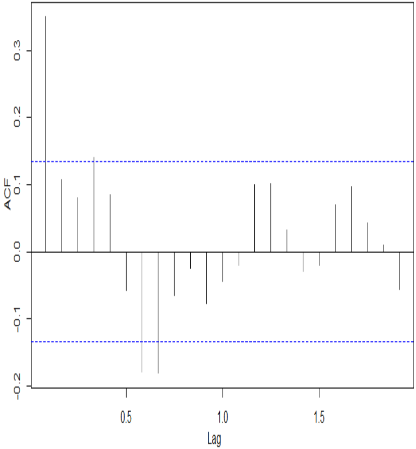


Transformação de Séries Temporais Altamente Persistentes

- Processos fracamente dependentes são chamados de **integrados de ordem zero I(0)**.
- Passeio aleatório: são **I(1)**.
- **Primeira Diferença** do processo é fracamente dependente.

Exemplo: Diferença da Taxa de Câmbio





Decisão sobre uma Série Temporal ser I(1)

Em um processo AR(1), se $|\rho_1| < 1$ então o processo é I(0).

Fazer a correlação amostral entre y_t e y_{t-1}, chamada de **autocorrelação amostral de primeira ordem** de $\{y_t\}$.

Se a série tem tendência: remover a tendência antes de calcular a correlação.

- Outras formas: testes de Raiz Unitária

Analisar Salários e Produtividade

Seja *hrwage* a média de salário por hora na economia dos Estados Unidos, e *outphr* a produção por hora. Estime a elasticidade salário por hora em relação à produção por hora, a partir do modelo:

$$\log(hrwage) = \beta_0 + \beta_1 \log(outpr) + \beta_2 t + u$$

em que t é a tendência temporal. A partir dos dados do arquivo **earns.xlsx** verifique se as séries são estacionárias. Faça a correção necessária para obter séries estacionárias e estime o modelo.

Propriedades do MQO com Erros Serialmente Correlacionados

Inexistência de Viés e Consistência

Seja o modelo de regressão linear simples:

$$y_t = \beta_0 + \beta_1 x_t + u_t \tag{4}$$

Sob as hipóteses ST.1 a ST.3 os estimadores de MQO são não-viesados (desde que os regressores sejam estritamente exógenos).

Relaxando a hipótese de exogeneidade e quando os dados são fracamente dependentes, os estimadores de MQO são consistentes.

Sob a hipótese que

$$u_t = \rho u_{t-1} + e_t$$

em que $|\rho| < 1$ e e_t são variáveis aleatórias não correlacionadas com $E(e_t) = 0$ e $Var(e_t) = \sigma_e^2$.

$$Var(u_t) = \frac{\sigma_e^2}{1-\rho^2} \qquad\qquad E(u_t u_{t+s}) = \rho^s \frac{\sigma_e^2}{1-\rho^2}$$

A variância de $\hat{\beta}_1$ é dada por:

$$Var(\hat{\beta}_1) = \frac{1}{\sum x_t^2}\left(Var(u_t) + \frac{2Var(u_t)}{\sum x_t^2}\sum_{t=1}^{n-1}\sum_{s=1}^{n-t} x_t x_{t+s} E(u_t u_{t+s})\right)$$

Identificação de Correlação Serial

- Método Gráfico

  A análise gráfica dos resíduos $\hat{u}_t$ obtidos por MQO nos fornece uma "pista" sobre a possível presença de autocorrelação nos resíduos, devido a presença de um padrão sistemático.

  Formas de examinar os resíduos:

1. Representá-los graficamente em relação ao tempo.

2. Ou representá-los graficamente $\hat{u}_t$ contra $\hat{u}_{t-1}$.

Teste de Correlação Serial com Regressores Estritamente Exógenos

1) Teste t de Correlação Serial AR(1)

Sob a hipótese de que $E(u_t|x_t) = 0$ e que $u_t = \rho u_{t-1} + e_t$, devemos assumir que:

$$E(e_t|u_{t-1}) = 0 \text{ e } Var(e_t|u_{t-1}) = Var(e_t) = \sigma_e^2.$$

No modelo AR(1), a hipótese nula de que os erros são não correlacionados é:

$$H_0 : \rho = 0$$

Teste de Correlação Serial com Regressores Estritamente Exógenos

Passos para o teste:

1. Estime o modelo de y_t sobre $x_{t1}, x_{t2}, \ldots, x_{tk}$ e obtenha os resíduos do MQO $\hat{u}_t$.
2. Estime a regressão $\hat{u}_t$ sobre $\hat{u}_{t-1}$ e obtenha o coeficiente $\hat{\rho}$ de $\hat{u}_{t-1}$ e sua estatística $t_{\hat{\rho}}$.
3. Use $t_{\hat{\rho}}$ para testar $H_0 : \rho = 0$ contra $H_1 : \rho \neq 0$.

Obs: Ao usar a estatística t usual, a hipótese de homocedasticidade deve ser satisfeita. Caso não seja, usar o teste robusto de heterocedasticidade.

Teste de Correlação Serial com Regressores Estritamente Exógenos

2) Teste d de Durbin-Watson sob as Hipóteses Clássicas de MQO

A estatística para teste d de Durbin-Watson é definida pela seguinte expressão:

$$DW = \frac{\sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n} \hat{u}_t^2} \tag{5}$$

Ou podemos reescrever a equação (5) como:

$$DW \cong 2\,(1 - \hat{\rho}) \tag{6}$$

Definindo:

$$\widehat{\rho} = \frac{\sum \hat{u}_t \hat{u}_{t-1}}{\sum \hat{u}_t^2} \tag{7}$$

como um estimador do coeficiente de autocorrelação de primeira ordem na amostra e como $-1 \leq \widehat{\rho} \leq 1$, temos:

Teste de Correlação Serial (Phillips.xlsx)

Para verificar se existe uma relação, em média, entre inflação e desemprego, podemos testar $H_0 : \beta_1 = 0$ contra $H_1 : \beta_1 < 0$ em:

$$inf_t = \beta_0 + \beta_1 desemp_t + u_t$$

em que *inf* é a taxa de inflação anual e *desemp* é a taxa de desemprego no mesmo período. (Phillips.xlsx)

Teste de Correlação Serial (Phillips.xlsx)

Uma versão linear da curva de Phillips de expectativas aumentadas pode ser escrita como:

$$inf_t - inf_t^e = \beta_1(desemp_t - \mu_0) + e_t$$

em que μ_0 é a taxa natural de desemprego e inf_t^e é a taxa de inflação esperada formada no ano $t-1$. Sob a hipótese de que $inf_t^e = inf_{t-1}$, temos:

$$\Delta inf_t = \beta_0 + \beta_1 desemp_t + e_t$$

Teste em cada uma das equações o termo erro para verificar a existência de correlação serial.

Teste de Correlação Serial AR(1) sem Regressores Estritamente Exógenos

Quando as variáveis explicativas não são estritamente exógenas, os testes t e Durbin Watson não são válidos.

Durbin (1970)[1] propôs uma estatística alternativa para o caso de variáveis explicativas não-estritamente exógenas.

1. Compute a regressão MQO de y_t sobre $x_{t1}, x_{t2}, \ldots, x_{tk}$ e obtenha os resíduos $\hat{u}_t$.

2. Compute a regressão $\hat{u}_t$ sobre $x_{t1}, x_{t2}, \ldots, x_{tk}, \hat{u}_{t-1}$ e obtenha $\hat{\rho}$ de $\hat{u}_{t-1}$ e a estatística t associada;

3. Teste $H_0 : \rho = 0$ contra a hipótese $H_1 : \rho \neq 0$.

[1] Durbin, J. *Testing for Serial Correlation in Least Squares Regressions when Some of the Regressors are Lagged Dependent Variables.* Econometrica, vol. 38, p. 410–421, 1970.

Emprego e Salário Mínimo em Porto Rico

A partir de dados anuais sobre a taxa de emprego, salário mínimo e outras variáveis Castillo-Freeman e Freeman (1992)[2] analisaram os efeitos do salário mínimo norte-americano sobre o emprego em Porto Rico. Uma versão simplifica do modelo é:

$$\log(prepop_t) = \beta_0 + \beta_1 \log(mincov_t) + \beta_2 \log(usgnp_t) + \beta_3 \log(prgnp_t) + \beta_4 t + u_t$$

sendo *prepop* a taxa de emprego em Porto Rico durante o ano t (razão entre as pessoas com trabalho e a população total), *usgnp* o produto nacional bruto real dos USA (em bilhões de dólares) *prgnp* o produto nacional bruto real de Porto Rico (em bilhões de dólares); *mincov* mede a importância do salário mínimo em relação aos salários médios e t representa a inclusão de uma tendência temporal linear no modelo.

[2] Castillo-Freeman, A. J.; Freeman, R. B. *When the Minimum Wage Really Bites: The effect of the U.S.-Level Minimum on Puerto Rico.* In: BORJAS, G.J.; FREEMAN, R.B. (Ed) *Imigration and the Work Force*. Chicago: University of Chicago Press, p. 177-211, 1992.

Teste de Correlação Serial sem Regressores Estritamente Exógenos

Teste de Breusch-Godfrey (BG)

Sob a hipótese nula de ausência de autocorrelação até a defasagem p, o teste é baseado nos resíduos da regressão:

$$y_t = \hat{\beta}\mathbf{x}_t + \hat{u}_t$$

em que $\hat{u}_t$ são os resíduos da regressão estimada de MQO.

Estima-se a regressão $\hat{u}_t$ sobre $x_{t1}, x_{t2}, \ldots, x_{tk}, \hat{u}_{t-1}, \ldots, \hat{u}_{t-q}$.

Obtenha o $R^2_{\hat{u}_t}$ desta regressão e calcule a estatística LM:

$$LM = (n - q)R^2_{\hat{u}_t} \sim \chi^2_q$$

Este teste também exige a hipótese de homocedasticidade.

Pode ser considerado o caso das estatísticas robustas de Heterocedasticidade.

Seja o modelo de regressão na forma matricial:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{u}$$

com

$$E(\mathbf{u}\mathbf{u}') = \sigma^2\mathbf{\Omega}$$

A variância de $\hat{\beta}$ na presença de autocorrelação é dada por:

$$Var(\hat{\beta}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$$

Matriz de Variância e Covariância de Newey e West

Inferência robusta para correlação serial de Newey-West:

1. Estime por MQO $y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_k x_{kt} + u_t$ e obtenha $\hat{\sigma}$, $\hat{u}_t$, "$ep(\hat{\beta}_j)$";
2. Calcule os resíuos $\hat{r}_t$ da regressão de x_{jt} sobre $x_{2t}, \ldots, x_{kt}$, exceto x_{jt};
3. Compute $\hat{a}_t = \hat{r}_t \hat{u}_t, t = 1, \ldots, n$;
4. Escolha um inteiro $g > 0$ e calcule:

$$\hat{v} = \sum_{t=1}^{n} \hat{a}_t^2 + 2 \sum_{h=1}^{g} \left(\frac{1-h}{g+1} \right) \left(\sum_{t=h+1}^{n} \hat{a}_t \hat{a}_{t-h} \right)$$

5. Obtenha o **erro padrão robusto em relação à correlação serial** de $\hat{\beta}_j$ como:

$$ep(\hat{\beta}_1) = \left(\frac{\text{“}ep(\hat{\beta}_j)\text{''}}{\hat{\sigma}} \right) \sqrt{\hat{v}}$$

Matriz de Variância e Covariância de Newey e West

- Erro padrão pode ser utilizado para construir IC e estatísticas t de $\hat{\beta}_j$;
- Útil quando não sabemos se variáveis explicativas são estritamente exógenas;
- Vale também para variáveis explicativas defasadas;
- Maior g, maior o no. de termos incluídos para corrigir a autocorrelação;
- Erro padrão corrigido também é robusto para heterocedasticidade;
- Newey-West sugerem $g = 4 \cdot (n/100)^{2/9}$;
- Para $g = 1$ e $g = 2$, por exemplo:

$$\hat{v} = \sum_{t=1}^{n} \hat{a}_t^2 + \sum_{t=2}^{n} \hat{a}_t \hat{a}_{t-1}$$

$$\hat{v} = \sum_{t=1}^{n} \hat{a}_t^2 + (4/3) \sum_{t=2}^{n} \hat{a}_t \hat{a}_{t-1} + (2/3) \sum_{t=2}^{n} \hat{a}_t \hat{a}_{t-2}$$

Exemplo: Salário Mínimo de Porto-Rico

Obtenha o erro-padrão robusto para o efeito do salário mínimo na equação de desemprego.

Correção da Correlação Serial com Regressores Estritamente Exógenos

Considere o modelo em notação matricial:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{u}$$

em que $E(\mathbf{u}) = 0$ e $E(\mathbf{u}\mathbf{u}') \neq 0$.

Assumindo que as hipóteses ST.1 a ST.4 estão válidas e que os erros seguem um processo AR(1):

$$u_t = \rho u_{t-1} + e_t$$

Sob a hipótese ST.2 $E(u_t|X) = 0$, tem-se

$$Var(u_t) = \frac{\sigma_e^2}{1-\rho^2} \qquad Cov(u_t, u_{t+s}) = \rho^s \frac{\sigma_e^2}{1-\rho^2}$$

A matriz de variância e covariância de **u** é dada por:

$$E(\mathbf{u}\mathbf{u}') = \frac{\sigma_e^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{bmatrix} = \sigma_e^2\mathbf{\Omega}$$

Os estimadores de mínimos quadrados são dados por:

$$\hat{\beta}_{MQG} = (\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{Y}$$

Dado que Ω deve ser estimado, o estimador passa a ser denominado de Mínimos Quadrados Generalizados Factível.

A matriz Ω é função de ρ, devemos estimar consistentemente este parâmetro e, em seguinda, estimar Ω.

Duas formas para estimar ρ:

1. Estimação de Cochrane-Orcutt
2. Estimação de Prais-Winsten

Estimação de Cochrane-Orcutt (CO)

Considere o modelo:

$$y_t = \beta_0 + \beta_1 x_t + u_t \tag{9}$$

com

$$u_t = \rho u_{t-1} + e_t \tag{10}$$

sendo $t = 1, \ldots, T$, $|\rho| < 1$ e $\{e_t\}$ é i.i.d.

O procedimento consiste em:

1. Estimar o modelo (9) por MQO e obter os resíduos $\hat{u}_t$.
2. A partir da série de resíduos, obter a regressão:

$$\hat{u}_t = \hat{\rho}\hat{u}_{t-1} + w_t \tag{11}$$

Estimação de Cochrane-Orcutt (CO)

3. Com o coeficiente $\hat{\rho}$ estimado de (11), obter a regressão de diferença:

$$y_t - \hat{\rho} y_{t-1} = \beta_0^* + \beta_1 (x_t - \hat{\rho} x_{t-1}) + v_t \tag{12}$$

   sendo $\beta_0^* = \beta_0(1 - \hat{\rho})$.

4. Obtenha os resíduos de (12), $\hat{u}^{**}$.
5. A partir dos resíduos $\hat{u}^{**}$ obtenha:

$$\hat{u}_t^{**} = \hat{\hat{\rho}} \hat{u}_{t-1}^{**} + w_t \tag{13}$$

   em que $\hat{\hat{\rho}}$ é o valor estimado na segunda iteração.

6. Este processo deve continuar até que algum critério de convergência definido a priori tenha sido alcançado (por exemplo, se a variação entre ρ estimado em duas iterações consecutivas for menor que $0,001$).

Estimação de Prais-Winsten (PW)

Segue todos os passos do processo de estimação de CO, porém usa a primeira observação computada como:

$$\tilde{y}_1 = (1-\rho^2)^{1/2}\beta_0 + \beta_1\tilde{x}_1 + \tilde{u}_1$$

em que $\tilde{u}_1 = (1-\rho^2)^{1/2}u_1,\ \tilde{y}_1 = (1-\rho^2)^{1/2}y_1$ e $\tilde{x}_1 = (1-\rho^2)^{1/2}x_1$

Exemplo: Estimação de CO e PW

A partir dos dados do modelo relacionados ao emprego e salário em Porto Rico, use a transformação de Cochrane-Orcutt e analise os resultados.
Após use a transformação de Prais-Winsten.

Sugestões de Exercícios e Leitura

1. Leitura Capítulos 10, 11, 12;
2. Leitura Próxima Aula: Capítulo 15;