**ESTIMAÇÃO DE MODELOS DE REGRESSÃO
EM SÉRIES TEMPORAIS**

Marli Mikael da Costa Neves

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO GRAU DE DOUTOR
EM
ESTATÍSTICA

ORIENTADOR: Prof. Dr. Pedro Alberto Morettin

São Paulo, setembro de 1988

Às minhas meninas

Drica

e

Rê

AGRADECIMENTOS

- Ao Professor Doutor Pedro Alberto Morettin, por orientar este trabalho de forma dedicada, paciente e amiga, como também por toda a contribuição que tem dado à minha formação, desde os primeiros cursos de pós-graduação.
- À Heloísa Nogueira de Lima, pelo carinho e dedicação com que me ajudou na elaboração de programas computacionais.
- À Zara Issa Abud, que de forma muito atenciosa me auxiliou com seus conhecimentos de Cálculo.
- À amiga Silvia Nagib Elian, pelas conversas sempre muito produtivas.
- Ao Professor Luiz K. Hotta pela leitura final deste texto, pelas sugestões e críticas apresentadas.
- À Lourdes Vaz da Silva Netto pelo apoio e pelo primoroso trabalho de datilografia.
- Ao Wladi, pelo apoio e incentivo que seu amor tem me proporcionado.

ABSTRACT

This work deals with the linear regression model

$$y_t = \underline{u}_t' \underline{\beta} + x_t, \quad t = 1, \dots, T,$$

where $\underline{u}_t = (u_{1t}, \dots, u_{kt})'$ and $\underline{\beta} = (\beta_1, \dots, \beta_k)'$. The problem is to estimate the vector of parameters $\underline{\beta}$, from observations $\{y_t, \underline{u}_t\}$, $t = 1, \dots, T$, under certain assumptions on the sequences $\{\underline{u}_t\}$ and $\{x_t\}$, the latter considered to be a stationary random sequence.

The situation of interest is the one in which we consider two-stage Cochrane-Orcutt type estimators, where the order of the autoregression fitted to the first stage residuals is obtained via some information criterion (AIC, BIC, etc). It is shown that the proposed estimator is consistent, asymptotically normal and efficient.

Simulations were performed in order to illustrate the behavior of the estimators, for particular sequences $\underline{u}_t$ and x_t .

RESUMO

Dado o modelo de regressão linear

$$y_t = \underline{u}_t' \underline{\beta} + x_t, \quad t = 1, \dots, T,$$

onde $\underline{u}_t = (u_{1t}, \dots, u_{kt})'$ e $\underline{\beta} = (\beta_1, \dots, \beta_k)'$, este trabalho trata, basicamente, da estimação do parâmetro $\underline{\beta}$, a partir de observações $\{y_t, \underline{u}_t\}$, $t = 1, \dots, T$, sob determinadas suposições a respeito de $\{\underline{u}_t\}$ e $\{x_t\}$, esta última considerada estocástica estacionária.

A situação de maior interesse é aquela em que se considera o estimador de Cochrane-Orcutt em dois estágios, onde a ordem da auto-regressão ajustada aos resíduos de primeiro estágio é obtida via algum critério (AIC, BIC, etc.). Mostra-se a consistência, a normalidade e a eficiência assintóticas do estimador proposto.

Do ponto de vista aplicado foram feitas simulações para ilustrar o comportamento dos estimadores propostos, para algumas se quências particulares $\underline{u}_t$ e x_t .

ÍNDICE

CAPÍTULO 1 - Preliminares	1
1.1- Introdução	1
1.2- Possíveis procedimentos de estimação	3
1.3- A terminologia usada para os estimadores	9
1.4- Objetivos	10
1.5- Suposições	11
 CAPÍTULO 2 - O Estimador de Mínimos Quadrados Ordinários	 14
2.1- Introdução	14
2.2- A consistência do estimador $\hat{\beta}_{MQO}$	15
2.3- A normalidade assintótica de $\hat{\beta}_{MQO}$	16
 CAPÍTULO 3 - O Estimador de Cochrane-Orcutt Generalizado em dois Estágios	 19
3.1- Introdução	19
3.2- O procedimento de estimação	20
3.3- Notações e definições	22
3.4- A convergência das auto-covariâncias	25
3.5- A convergência dos estimadores dos coeficientes auto-regressivos	30
3.6- A consistência de $\hat{\beta}$	36
3.7- A normalidade assintótica de $\hat{\beta}$	44
3.8- O caso vetorial	48
3.9- A eficiência assintótica de $\hat{\beta}$	48

CAPÍTULO 4 - Comparação entre o Estimador de Cochrane-Orcutt e o Estimador de Mínimos Quadrados Ordinários Quando a Variável Independente é Estocástica	52
4.1- Introdução	52
4.2- Variâncias assintóticas dos estimadores $\hat{\beta}_{MQO}$ e $\hat{\beta}_{\phi}$	53
4.3- Comparação dos estimadores	57
 CAPÍTULO 5 - Expansão Assintótica do Erro Quadrático Médio do Estimador de Cochrane-Orcutt em Dois Estágios	 65
5.1- Introdução	65
5.2- Suposições e notações	67
5.3- A expansão de 2ª ordem do erro quadrático médio de $\hat{\beta}_{\phi}$	69
5.4- O caso vetorial	72
5.5- Casos especiais	75
 CAPÍTULO 6 - Simulações	 79
6.1- Introdução	79
6.2- Estimção recursiva e critérios para a determinação de h	79
6.3- As situações onde as simulações foram empregadas	82
6.4- Resultados e Comentários	87
 CAPÍTULO 7 - Discussão e Recomendações	 97
7.1- Discussão	98
7.2- Recomendações	99
7.3- Análise da representação espectral de $\hat{\beta}$	102
7.4- Considerações a respeito do vício e do EQM de $\hat{\beta}$	104
 REFERÊNCIAS	 108

CAPÍTULO 1

PRELIMINARES

1.1 Introdução

Neste trabalho estaremos interessados em vários aspectos do modelo

$$y_t = \underline{u}_t' \underline{\beta} + \underline{x}_t, \quad (1.1.1)$$

onde $\underline{u}_t = (u_{1t}, \dots, u_{kt})'$ e $\underline{\beta} = (\beta_1, \dots, \beta_k)'$, sob várias suposições a respeito das seqüências $\{\underline{u}_t\}$ e $\{\underline{x}_t\}$, $t = 1, 2, \dots$. Em todos os casos, o objetivo básico é estimar o parâmetro $\underline{\beta}$, a partir de dados obtidos nos instantes $t=1, 2, \dots, T$. Os casos de interesse são:

- (i) $\{\underline{u}_t\}$ e $\{\underline{x}_t\}$ são seqüências estocásticas, de modo que $\{\underline{x}_t\}$ seja não observável, $\{\underline{u}_t\}$ e $\{y_t\}$ observáveis;
- (ii) $\{\underline{u}_t\}$ é uma seqüência determinística observável, $\{\underline{x}_t\}$ uma seqüência estocástica não observável e $\{y_t\}$ estocástica observável.

Os dados são $\{y_t, \underline{u}_t\}$, $t=1, \dots, T$ e o modelo (1.1.1) pode ser escrito na forma

$$\underline{y} = \underline{U} \underline{\beta} + \underline{x} \quad (1.1.2)$$

onde $\underline{y} = (y_1, \dots, y_T)'$, $\underline{U} = (\underline{u}_1, \dots, \underline{u}_T)'$ é uma matriz $T \times k$ e

$\underline{x} = (x_1, \dots, x_T)'$. Observe que podemos, evidentemente, escrever (1.1.1) como

$$y_t = \sum_{j=1}^k \beta_j u_{jt} + x_t, \quad t = 1, \dots, T. \quad (1.1.3)$$

Se a sequência $\{x_t\}$ for ruído branco (variáveis aleatórias não correlacionadas com a mesma distribuição) de média zero e variância σ_x^2 , se $\{u_t\}$ for uma sequência fixada e o posto de $\underline{U}$ é k , então o estimador de mínimos quadrados ordinários (MQO) de $\underline{\beta}$,

$$\hat{\underline{\beta}}_{MQO} = (\underline{U}' \underline{U})^{-1} \underline{U}' \underline{y} \quad (1.1.4)$$

é ótimo (BLUE) no sentido habitual. Sob certas condições de regularidade a respeito de $\underline{u}_t$, pode-se mostrar a normalidade assintótica de $\hat{\underline{\beta}}_{MQO}$. Ver Harvey (1981a, pag.46) para detalhes.

Se $\{x_t\}$ não for ruído branco, então procedimentos ótimos envolvem algo que é assintoticamente equivalente a substituir $\underline{U}' \underline{U}$ e $\underline{U}' \underline{y}$, em (1.1.4), por $\underline{U}' \underline{V}^{-1} \underline{U}$ e $\underline{U}' \underline{V}^{-1} \underline{y}$, onde $\underline{V}$ é a matriz $(T \times T)$ de covariâncias de $\underline{x} = (x_1, \dots, x_T)'$.

Iremos supor que $\{x_t\}$ seja um processo estacionário, de modo que $\underline{V}$ é determinada pelas covariâncias

$$\gamma_k = \text{cov} \{x_t, x_{t-k}\}, \quad 0 \leq k < T.$$

Um caso de interesse é aquele em que $\underline{V}$ depende de um vetor de parâmetros $\underline{\theta} = (\theta_1, \dots, \theta_p)$ desconhecidos, situação em que escreveremos $\underline{V}(\underline{\theta})$. Por exemplo, x_t é um processo auto-regressivo de ordem p , AR(p).

De qualquer modo, $\underline{v}$ é desconhecida, e portanto estimadores de mínimos quadrados generalizados (MQG) envolverão dois estágios:

- (A) Estimação dos parâmetros do processo $\{x_t\}$;
- (B) Uso dos resultados de (A) para estimar $\underline{g}$.

Ao implementar (A) será necessário obter um estimador consistente de $\underline{g}$, geralmente através de mínimos quadrados ordinários, como também obter os resíduos $\hat{x}_t$ dessa regressão de primeiro estágio. Assim, os resíduos de mínimos quadrados ordinários, dados por

$$\hat{x}_t = y_t - \underline{u}_t' \hat{\underline{\beta}}_{MQO} \quad (1.1.5)$$

servirão como "substitutos" dos erros x_t , que são desconhecidos.

1.2 Possíveis procedimentos de estimação

Algumas alternativas para implementar os estágios (A) e (B) são mostradas no diagrama abaixo (Pierce, 1980), as quais ilustraremos com maiores detalhes a seguir.

1.2.1 Estimação do processo $\{x_t\}$

Definindo

$$Bx_t = x_{t-1},$$

$$\zeta(B) = 1 + \zeta_1 B + \zeta_2 B^2 + \dots$$

e

$$\Pi(B) = 1 - \Pi_1 B - \Pi_2 B^2 - \dots,$$

sob determinadas condições sobre os pesos Π_j e ζ_j , a representação de médias móveis (MA) e a representação auto-regressiva (AR) de x_t são dadas, respectivamente, por:

$$x_t = \zeta(B) a_t \quad (1.2.1)$$

$$\Pi(B) x_t = a_t, \quad (1.2.2)$$

onde $\{a_t\}$ é o ruído branco de média zero e variância σ_a^2 . Veremos, então, como representar x_t por meio de um modelo que o aproxime bem.

(a) Modelagem ARIMA

É uma forma de representação parcimoniosa para x_t , obtida quando se aproxima $\Pi(B)$ (ou $\zeta(B)$) pela razão de dois polinômios finitos

$$\Pi(B) = \frac{\phi(B)}{\theta(B)}.$$

Se

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

e

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q,$$

obtemos:

$$x_t = \sum_{j=1}^p \phi_j x_{t-j} - \sum_{j=1}^q \theta_j a_{t-j} + a_t \quad (1.2.3)$$

Como estaremos sempre considerando que x_t é estacionário, o modelo dado em (1.2.3) é um auto-regressivo de médias móveis de ordem (p,q) , ARMA(p,q).

Estimar x_t compreende especificar p e q e estimar os parâmetros $\underline{\phi} = (\phi_1, \dots, \phi_p)'$ e $\underline{\theta} = (\theta_1, \dots, \theta_q)'$, o que pode ser feito, por exemplo, usando a metodologia de Box-Jenkins (1976).

(b) Modelagem auto-regressiva

Qualquer processo estacionário pode ser bem aproximado por um modelo auto-regressivo de ordem suficientemente grande. Lembrando que $\pi_j \rightarrow 0$ quando $j \rightarrow \infty$, a equação (1.2.2) pode ser aproximada por

$$x_t = \sum_{j=1}^h \pi_j x_{t-j} + a_t \quad (1.2.4)$$

Aqui também, estimar x_t compreende especificar h e esti-

mar $\underline{\Pi} = (\Pi_1, \dots, \Pi_h)'$. Um estimador consistente de $\underline{\Pi}$ pode ser obtido pela regressão de mínimos quadrados de x_t sobre $x_{t-1}, \dots, x_{t-h}$.

(c) Estimação espectral

Pode-se implementar o estimador de MQG, estimando-se o espectro do processo residual. Um estimador do espectro de $\{x_t\}$ é o periodograma dado por

$$I(\lambda_j) = |\tilde{x}_j|^2, \quad (1.2.5)$$

onde $\tilde{x}_j$ é a transformada de Fourier de $\{x_t\}$, isto é:

$$\tilde{x}_j = \frac{1}{\sqrt{T}} \sum_{t=1}^T x_t \exp \{i t \lambda_j\} \quad (1.2.6)$$

$$\lambda_j = \frac{2\pi j}{T}, \quad j = 0, \dots, T-1.$$

Como o periodograma não é um estimador consistente, na prática usam-se estimadores que são médias ponderadas do periodograma sobre frequências adjacentes. Para detalhes ver Priestley (1981).

Levando em conta as representações AR e ARMA pode-se estimar o espectro de x_t , respectivamente, por

$$\hat{f}_{AR}(\lambda) = \left| \hat{\Pi}(e^{i\lambda}) \right|^{-2} \quad (1.2.7)$$

$$\hat{f}_{ARMA}(\lambda) = \left| \frac{\hat{\Phi}(e^{i\lambda})}{\hat{\Theta}(e^{i\lambda})} \right|^{-2} \quad (1.2.8)$$

1.2.2 Estimação do modelo de regressão

O estágio (A) permite caracterizar o processo x_t pela possibilidade de estimar seu espectro ou sua função de auto-correlação. Vejamos como implementar o estágio B.

(a) Método da "força bruta"

O procedimento consiste em formar $\tilde{V}$, invertê-la e calcular

$$\hat{\beta} = (\tilde{U}' \tilde{V}^{-1} \tilde{U}) \tilde{U}' \tilde{V}^{-1} \tilde{y}.$$

Lembrando que

$$\tilde{V} = \{\gamma_{ij}\} = \{\gamma_{i-j}\}, \quad \gamma_k = \sigma_a^2 \sum_{j=0}^{\infty} \zeta_j \zeta_{j+k}, \quad \zeta_0 = 1$$

onde ζ_j é o coeficiente de B^j em (1.2.1), podemos obter

$$\hat{\zeta}(B) = [\hat{\phi}(B)]^{-1} \hat{\theta}(B)$$

se um modelo ARMA é ajustado, ou

$$\hat{\zeta}(B) = [\hat{\pi}(B)]^{-1}$$

se uma auto-regressão é ajustada.

Como $\zeta_j \rightarrow 0$ se $j \rightarrow \infty$, $\hat{\zeta}(B)$ pode ser truncada num certo ponto Q.

(b) Método de Cochrane-Orcutt

Seja $\hat{\pi} = (\hat{\pi}_1, \dots, \hat{\pi}_h)'$, um estimador consistente de

$\underline{\pi} = (\pi_1, \dots, \pi_h)'$, obtido no estágio A. Se formarmos as seqüências

$$\hat{\eta}_t = y_t - \sum_{i=1}^h \hat{\pi}_i y_{t-i} \quad , \quad (1.2.9)$$

$$\hat{\psi}_{j,t} = u_{j,t} - \sum_{i=1}^h \hat{\pi}_i u_{j,t-i} \quad , \quad j = 1, \dots, k \quad (1.2.10)$$

um estimador factível de MQG de β , é obtido pela regressão de MQO $\hat{\eta}_t$ sobre $\hat{\psi}_{j,t}$, $j = 1, \dots, k$, para $t = h+1, \dots, T$. Tal estimador é de nominado estimador de *Cochrane-Orcutt em dois estágios* (CO2).

O método pode ser usado também no caso de ser ajustado um modelo ARMA a x_t (ver Pierce, 1980).

(c) Regressão no domínio da freqüência

As transformadas de Fourier $\tilde{y}_t$ e $\tilde{u}_{j,t}$ são definidas de modo análogo a $\tilde{x}_j$ em (1.2.6). O modelo (1.1.2) usando as transformadas fica:

$$\tilde{y} = \tilde{U} \tilde{\beta} + \tilde{x} \quad (1.2.11)$$

onde

$$\tilde{y} = W y \quad , \quad \tilde{U} = W U \quad , \quad \tilde{x} = W x \quad ,$$

e W é a matriz, $T \times T$, cujo (j, ℓ) -ésimo elemento é $T^{-1/2} e^{i \ell \lambda_j}$.

Pode-se mostrar (Hannan, 1963) que a matriz de covariâncias de $\tilde{x}$ é

$$E(\tilde{x} \tilde{x}') = W V W' \equiv D \quad (1.2.12)$$

onde

$$\underline{D} = \text{diag} [f_x(\lambda_0), \dots, f_x(\lambda_{T-1})],$$

sendo $f_x(\lambda)$ o espectro de $\{x_t\}$.

O estimador de MQG no domínio da frequência é:

$$\hat{\beta} = (\tilde{U}' \underline{D}^{-1} \tilde{U})^{-1} \tilde{U}' \underline{D}^{-1} \tilde{y} . \quad (1.2.13)$$

A regressão pode também ser implementada no domínio do tempo, o que evita o uso de variáveis complexas e torna viável o uso dos pacotes de regressão usuais. Obtêm-se

$$\hat{\beta} = (\underline{U}^*{}' \underline{U}^*)^{-1} \underline{U}^*{}' \underline{y}^* , \quad (1.2.14)$$

onde

$$\underline{U}^* = \underline{W}' \underline{D}^{-1/2} \underline{W} \underline{U}$$

e

$$\underline{y}^* = \underline{W}' \underline{D}^{-1/2} \underline{W} \underline{y} .$$

É necessário estimar o espectro de $f_x(\lambda_j)$, $j=0, \dots, T-1$. Aqui, uma das três análises da seção 1.2.1 pode ser empregada, levando cada uma a um estimador de espectro.

1.3 A terminologia usada para os estimadores

Seja o modelo (1.1.1) onde x_t pode ser bem representado por um AR(h), isto é, por um modelo como o dado em (1.2.4).

Denominaremos de estimador de *Cochrane-Orcutt Generalizado* (COG), ao estimador obtido pela regressão de MQO de η_t sobre

$\psi_{j,t}$, $j=1,\dots,k$, para $t=h+1,\dots,T$, onde η_t e $\psi_{j,t}$, $j=1,\dots,k$ são obtidas de (1.2.9) e (1.2.10), respectivamente, quando se substitui $\hat{\Pi}_i$ por Π_i , $i=1,\dots,h$. No caso em que $h=1$, o estimador assim obtido será denominado simplesmente de estimador de *Cochrane-Orcutt* (CO).

O estimador de *Cochrane-Orcutt* em dois estágios (CO2) é o definido no item (b) da sub-seção 1.2.2.

1.4 Objetivos

Neste trabalho abordaremos os seguintes tópicos relacionados com o modelo (1.1.1):

(a) Considerando as seqüências $\{u_t\}$ determinística e $\{x_t\}$ estocástica, estudaremos o estimador de MQO de β , sob determinadas suposições para as duas seqüências.

(b) Na situação em que $\{x_t\}$ segue um modelo linear geral, proporemos um estimador de CO em dois estágios para β , onde a ordem h cresce com T . Mais precisamente, suporemos que h é tal que $h \leq H(T)$, onde $H(T)$ cresce com T segundo uma certa ordem, e é obtida via algum critério de "informação" (tais como AIC, BIC, critério de Hannan e Quinn, etc.). Denominaremos tal estimador de estimador de *Cochrane-Orcutt Generalizado em Dois Estágios* (COG2). Estudaremos sua consistência e normalidade assintótica quando $\{u_t\}$ é determinística, e mostraremos que tal estimador é assintoticamente eficiente.

(c) Supondo para $\{x_t\}$ um modelo AR(1), compararemos o

estimador de MQO com o estimador de CO em termos de suas variâncias assintóticas, para $\{u_t\}$ estocástica, seguindo um modelo ARMA particular.

(d) Supondo $\{u_t\}$ determinística e $\{x_t\}$ um processo AR(p), obteremos a expansão assintótica do erro quadrático médio do estimador CO2.

(e) Através de simulações, compararemos o estimador de MQO com o proposto no item (b), quando este último é obtido por cada um dos critérios mencionados. Compararemos também o estimador de MQO com o estimador CO2, na situação convencional de h fixada e grande.

(f) As conclusões e algumas recomendações de problemas a considerar constituem a parte final do trabalho.

1.5 Suposições

Vamos colocar nesta seção as principais suposições que serão utilizadas no decorrer do trabalho. Condições adicionais poderão ser impostas no desenvolvimento dos tópicos mencionados na seção anterior.

Com relação à seqüência $\{x_t\}$, quando admitirmos que ela seja gerada por um processo auto-regressivo de ordem infinita, isto é:

$$\sum_{j=0}^{\infty} a(j) x(t-j) = \epsilon_t, \quad (1.5.1)$$

onde $a(0)=1$, $a(1)$, $a(2)$,... são números reais, suporemos que:

(A1) ϵ_t é uma seqüência de variáveis independentes, normalmente distribuídas, com média zero e variância $\sigma_\epsilon^2 > 0$;

$$(A2) \sum_{j=0}^{\infty} j^{1/2} |a(j)| < \infty;$$

$$(A3) A(Z) = \sum_{j=0}^{\infty} a(j) Z^j \neq 0, \text{ para } |Z| \leq 1.$$

As condições (A2) e (A3) garantem a estacionariedade e a invertibilidade de x_t e também que sua densidade espectral seja limitada e diferente de zero.

Em muitas situações onde a seqüência u_t é determinística, dependendo de T , assumiremos que:

$$(B1) \lim_{T \rightarrow \infty} d_j^2(T) = \infty, \quad d_j^2(T) = \sum_{t=1}^T u_{j,t}^2, \quad j = 1, \dots, k,$$

$$(B2) \lim_{T \rightarrow \infty} \frac{u_{j,T}^2}{d_j^2(T)} = 0, \quad j = 1, \dots, k,$$

$$(B3) \lim_{T \rightarrow \infty} \frac{\sum_{t=1}^{T-|\ell|} u_{j,t} u_{i,t+|\ell|}}{(d_j^2(T))^{1/2} (d_i^2(T))^{1/2}} = r_{i,j}(\ell), \quad \ell = 0, \pm 1, \pm 2, \dots$$

A condição (B1) é necessária para garantir a consistência dos estimadores que estudaremos. Se as condições (B1)-(B3) estão satisfeitas, então a matriz $k \times k$ $R(\ell)$, cujo (i,j) -ésimo elemento é $r_{i,j}(\ell)$, é Hermitiana, não-negativa definida. Assim, ela admite a representação espectral

$$\tilde{R}(\lambda) = \int_{-\Pi}^{\Pi} e^{i\lambda\lambda} \tilde{F}_u(d\lambda), \quad (1.5.2)$$

onde $\tilde{F}_u(\lambda_2) - \tilde{F}_u(\lambda_1)$ é uma matriz Hermitiana não-negativa definida, para todo $-\Pi \leq \lambda_1 \leq \lambda_2 \leq \Pi$. Note que $\tilde{R}(0) = \tilde{F}(\Pi) - \tilde{F}(-\Pi)$ e que se

$$\tilde{D}_T = \text{diag} \left[\left(\sum_{t=1}^T u_{1,t}^2 \right)^{1/2}, \dots, \left(\sum_{t=1}^T u_{k,t}^2 \right)^{1/2} \right] \quad (1.5.3)$$

então

$$\tilde{R}(0) = \lim_{T \rightarrow \infty} \tilde{D}_T^{-1} \tilde{U}' \tilde{U} \tilde{D}_T^{-1}.$$

As condições (B1)-(B3) estão satisfeitas para tendências lineares e polinomiais e para funções do tipo $u_t = \cos \lambda t$, $u_t = \sin \lambda t$ (Hannan 1970).

No caso escalar, isto é, $k=1$, manteremos as notações das condições (B1)-(B3), suprimindo os índices i e j . Assim, $u_{j,T}$ será denotado por u_T , $d_j(T)$ por $d(T)$ e $r_{i,j}(\lambda)$ por $r(\lambda)$.

CAPÍTULO 2

O ESTIMADOR DE MÍNIMOS QUADRADOS ORDINÁRIOS

2.1 Introdução

No modelo clássico de regressão, introduzido no capítulo 1, uma suposição básica é que os erros x_t sejam não-correlacionados. Contudo, em se tratando de regressão com séries temporais, isto não é razoável. Normalmente x_t incorpora um certo número de variáveis omitidas no modelo, sendo muito pouco provável a inexistência de correlação serial. Nesse contexto o estimador $\hat{\beta}_{MQO}$, definido em (1.1.4), não tem propriedades ótimas. Em geral, o estimador de MQO será menos eficiente que o estimador de MQG. A extensão da ineficiência dependerá das estruturas de $\underline{V}$ e $\underline{U}$.

Nosso objetivo nesse capítulo é estudar algumas propriedades do estimador de MQO, supondo que $\{x_t\}$ segue o modelo linear geral, dado em (1.5.1), e que $\{u_t\}$ satisfaz certas condições de regularidade. Essas propriedades nos serão úteis no estudo do estimador em dois estágios que será proposto no capítulo 3.

Os teoremas serão apresentados para $\{u_t\}$ escalar, por facilidade de notação e exposição. As extensões no caso em que $k > 1$, podem ser obtidas sem maiores dificuldades.

2.2 A consistência do estimador $\hat{\beta}_{MQO}$

Teorema 2.2.1. Considere o modelo (1.1.1) com $\{x_t\}$ dado por (1.5.1) e satisfazendo (A1)-(A3) e com $\{u_t\}$ determinística, dependendo de T , satisfazendo (B3) e

$$(B'1) \quad d^2(T) = \sum_{t=1}^T u_t^2 = O(T^\delta), \text{ para algum } \delta > 0.$$

Nessas condições, o estimador $\hat{\beta}_{MQO}$ é consistente.

Prova. Temos que:

$$\text{Var}(\hat{\beta}_{MQO}) = (\underline{u}'\underline{u})^{-1} \underline{u}' \underline{V} \underline{u} (\underline{u}'\underline{u})^{-1},$$

onde

$\underline{u} = (u_1, \dots, u_T)'$ e $\underline{V}$ é a matriz $T \times T$ de covariâncias de $\underline{x} = (x_1, x_2, \dots, x_T)'$, ou seja:

$$\underline{V} = \begin{bmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{T-1} \\ \gamma_1 & \gamma_0 & \dots & \gamma_{T-2} \\ \vdots & & & \vdots \\ \gamma_{T-1} & \gamma_{T-2} & \dots & \gamma_0 \end{bmatrix}$$

com $\gamma_k = E(x_t x_{t-k})$, $k=0, \dots, T-1$.

Como

$$\begin{aligned} \underline{u}' \underline{V} \underline{u} &= \sum_{t=1}^T \sum_{s=1}^T u_t u_s \gamma(t-s) = \\ &= \sum_{k=-T+1}^{T-1} \gamma(k) \sum_{t=1}^{T-|k|} u_t u_{t+|k|} , \end{aligned}$$

então

$$|\underline{u}' \underline{V} \underline{u}| \leq \sum_{k=-T+1}^{T-1} |\gamma(k)| \left| \sum_{t=1}^{T-|k|} u_t u_{t+|k|} \right|$$

Mas, as condições (A1)-(A3) implicam que $\sum_k |\gamma(k)| < \infty$

(Shibata, 1980, p.151). Assim, as condições (B'1) e (B3) implicam que $\underline{u}' \underline{V} \underline{u} = O(T^\delta)$ e portanto $\text{Var}(\hat{\beta}_{MQO}) = O(T^{-\delta})$. Como $\hat{\beta}_{MQO}$ é um estimador não viciado de β , então $\hat{\beta}_{MQO}$ é um estimador consistente de β . Δ

Observação. A condição (B'1) é mais forte do que (B1), ou seja (B'1) $\rightarrow$ (B1).

2.3 A normalidade assintótica de $\hat{\beta}_{MQO}$

Para $\{x_t\}$ ruído branco com média zero e variância σ_x^2 , as condições impostas a $\{u_t\}$ seriam suficientes para garantir a normalidade assintótica de $\hat{\beta}_{MQO}$. O teorema a seguir mostra que a distribuição limite de $\hat{\beta}_{MQO}$ é normal para uma vasta classe de séries temporais. Sua demonstração pode ser vista em Fuller (1976, p.251).

Teorema 2.3.1. Seja $\{x_t, t \in \mathbb{Z}\}$ uma seqüência definida por

$$x_t = \sum_{j=0}^{\infty} c_j \epsilon_{t-j},$$

onde $\sum_{j=0}^{\infty} |c_j| < \infty$, e ϵ_t são variáveis aleatórias independentes, com média zero e variância σ^2 , e com função de distribuição $F_t(\epsilon)$ tal que

$$\lim_{\delta \rightarrow \infty} \sup_{t \in T} \int_{|\epsilon| > \delta} \epsilon^2 dF_t(\epsilon) = 0. \quad (2.3.1)$$

Seja $\{u_t\}$, $t=1,2,\dots$, uma sequência de números reais fixados, satisfazendo as condições (B1)-(B3) e $W = \sum_{\ell=-\infty}^{\infty} r(\ell) \gamma(\ell)$, onde $r(\ell)$ é definida na condição (B3). Então

$$\left[\begin{array}{c} \sum_{t=1}^T u_t^2 \\ \sum_{t=1}^T u_t x_t \end{array} \right]^{-1/2} \xrightarrow{L} N(0, W).$$

Teorema 2.3.2. Sob as condições (A1)-(A3) impostas a x_t e (B1)-(B3), impostas a u_t

$$\left[\begin{array}{c} \sum_{t=1}^T u_t^2 \\ \sum_{t=1}^T u_t x_t \end{array} \right]^{1/2} (\hat{\beta}_{MQO}^{-\beta}) \xrightarrow{L} N(0, W),$$

$$\text{onde: } W = \sum_{\ell=-\infty}^{\infty} r(\ell) \gamma(\ell). \quad (2.3.2)$$

Prova. Temos que:

$$\left[\begin{array}{c} \sum_{t=1}^T u_t^2 \\ \sum_{t=1}^T u_t x_t \end{array} \right]^{1/2} (\hat{\beta}_{MQO}^{-\beta}) = \left[\begin{array}{c} \sum_{t=1}^T u_t^2 \\ \sum_{t=1}^T u_t x_t \end{array} \right]^{-1/2} \sum_{t=1}^T u_t x_t$$

Pelas condições (A1)-(A3),

$$x_t = \sum_{j=0}^{\infty} b(j) \epsilon_{t-j}, \quad \text{onde os } b(j) \text{ satisfazem}$$

$$\sum_{j=0}^{\infty} j^{1/2} |b(j)| < \infty \quad (2.3.3)$$

(ver Hannan e Kavalieres, 1986), e portanto satisfazem também a $\sum_{j=0}^{\infty} |b(j)| < \infty$.

Além disso, ϵ_t são v.a. i.i.d., normais $(0, \sigma_{\epsilon}^2)$, e conseqüentemente têm função de distribuição satisfazendo (2.3.1).

As condições do teorema 2.3.1 estão satisfeitas e portanto o teorema está provado. Δ

CAPÍTULO 3

O ESTIMADOR DE COCHRANE-ORCUTT GENERALIZADO EM DOIS ESTÁGIOS

3.1 Introdução

Estimadores, assintoticamente eficientes, dos parâmetros de um modelo linear de regressão com erros correlacionados, têm sido sugeridos na literatura.

De modo geral, ao se considerar o modelo (1.1.1), admite-se que $\{x_t\}$ seja gerado por um modelo AR(p) ou um modelo ARMA(p,q), onde as ordens p e q são fixadas. Uma suposição mais realista, sem dúvida, é a de que $\{x_t\}$ seja gerado por um processo auto-regressivo de ordem infinita. Nosso objetivo nesse capítulo, é a apresentação de um modelo de estimação sob tal suposição para $\{x_t\}$ e o estudo de algumas propriedades para esse estimador.

Pierce (1971), Harvey (1981a), Fuller (1976) entre outros, investigaram o modelo (1.1.1) sob a suposição de que $\{x_t\}$ seja um ARMA(p,q). Os estimadores de β , estudados por tais autores, dependem dos estimadores do modelo ARMA de ordem fixada, ajustado aos resíduos de mínimos quadrados. Para o nosso procedimento de estimação deveremos nos reportar aos trabalhos daqueles autores que investigaram a estimação de uma série temporal, quando o processo gerador de

tal série é um auto-regressivo de ordem infinita. Nesse sentido têm relevância os trabalhos de Berk (1974), Bhansali (1978), An, Chen e Hannan (1982), Hannan e Kavalieris (1983, 1984a, 1984b, 1986), Shibata (1980), dentre outros.

3.2 O procedimento de estimação

Vamos considerar o modelo (1.1.1) com x_t dado por (1.5.1) e satisfazendo as suposições (A1)-(A3) e com u_t , determinística, es calar, satisfazendo (B1)-(B3).

Dados T pares de observações $\{(y_1, u_1), \dots, (y_T, u_T)\}$, o pro cedimento de estimação que consideraremos segue os seguintes passos:

- (a) Computa-se o estimador de MQO de β , $\hat{\beta}_{MQO}$.
- (b) Consideram-se os resíduos da regressão de MQO,

$$\hat{x}_t = y_t - \hat{\beta}_{MQO} u_t, \quad t=1, \dots, T \quad (3.2.1)$$

e ajusta-se uma auto-regressão de ordem h a $\hat{x}_t$. A ordem h será escolhida de modo a minimizar

$$\log \hat{\sigma}_h^2 + \frac{hC(T)}{T} \quad (3.2.2)$$

para $h \leq H(T)$, onde $H(T)$ cresce com T , e para $\frac{C(T)}{T} \rightarrow 0$. Se $C(T) = 2$ temos o critério AIC (Akaike, 1973), se $C(T) = \log T$ temos o critério BIC (Akaike, 1977; Rissanen, 1978; Schwarz, 1978), e se $C(T) = c \log \log T$, $c > 2$, temos o critério HQ (Hannan e Quinn, 1979).

Os estimadores dos coeficientes da auto-regressão ajustada a $\hat{x}_t$, que serão denotados por $\hat{a}_h(j)$, $j=1, \dots, h$, são estimadores de Yule-Walker, e podem ser obtidos através da recursão de Durbin-Levinson (Durbin, 1960; Morettin, 1984). Na expressão (3.2.2) $\hat{\sigma}_h^2$ é a variância residual para o modelo ajustado de ordem h .

(c) Formam-se as seqüências

$$\hat{\eta}_t = \sum_{j=0}^h \hat{a}_h(j) y_{t-j}, \quad t=h+1, \dots, T \quad (3.2.3)$$

$$\hat{\psi}_t = \sum_{j=0}^h \hat{a}_h(j) u_{t-j}, \quad t=h+1, \dots, T \quad (3.2.4)$$

onde $\hat{a}_h(0)=1$.

(d) O estimador final $\hat{\beta}$ de β é dado pela regressão de $\hat{\eta}_t$ sobre $\hat{\psi}_t$, para $t=h+1, \dots, T$, isto é:

$$\hat{\beta} = \frac{\sum_{t=h+1}^T \hat{\eta}_t \hat{\psi}_t}{\sum_{t=h+1}^T \hat{\psi}_t^2} \quad (3.2.5)$$

O estimador assim obtido é o que denominaremos de estimador de Cochrane-Orcutt Generalizado em dois estágios (COG2).

A justificativa para tal procedimento é que o modelo

$$y_t = \beta u_t + x_t$$

está, por meio dessas transformações, sendo substituído por

$$\hat{\eta}_t = \beta \hat{\psi}_t + \hat{\epsilon}_{t,h}, \quad (3.2.6)$$

onde

$$\hat{\epsilon}_{t,h} = \sum_{j=0}^h \hat{a}_h(j) x_{t-j}. \quad (3.2.7)$$

Podemos ver $\hat{\epsilon}_{t,h}$ como um estimador de ϵ_t , obtido quando se aproxima (1.5.1) por (3.2.7). Como $h=h_T$, à medida que h cresce cada vez mais $\hat{\epsilon}_{t,h}$ se aproxima de ϵ_t (Hannan, 1986, Teorema 2.2), e (3.2.6) se aproxima de um modelo com erros não correlacionados.

3.3 Notações e definições

Seja $\gamma(t)=E(x_s x_{s+t})$ a função de auto-covariâncias de x_t . Pelo fato da correção pela média não alterar os resultados que apresentaremos, definiremos as funções de auto-covariâncias amostrais de x_t e $\hat{x}_t$, respectivamente por:

$$c(t) = \begin{cases} \frac{1}{T} \sum_{s=1}^{T-t} x_s x_{s+t}, & 0 \leq t \leq T-1 \\ 0, & t \geq T \end{cases}$$

$$c(-t) = c(t), \quad t=1,2,\dots$$

$$\hat{c}(t) = \begin{cases} \frac{1}{T} \sum_{s=1}^{T-t} \hat{x}_s \hat{x}_{s+t}, & 0 \leq t \leq T-1 \\ 0 & , \quad t \geq T \end{cases}$$

$$\hat{c}(-t) = \hat{c}(t), \quad t=1,2,\dots$$

Seja

$$A(Z)^{-1} = B(Z) = \sum_{j=0}^{\infty} b(j)Z^j, \quad (3.3.1)$$

então, como verificaram Hannan e Kavalieris (1986), (A2) implica que:

$$\sum_{j=0}^{\infty} j^{1/2} |b(j)| < \infty. \quad (3.3.2)$$

Deste modo, x_t pode ser escrito como

$$x_t = \sum_{j=0}^{\infty} b(j) \epsilon_{t-j}, \quad (3.3.3)$$

onde os $b(j)$ satisfazem (3.3.2).

O melhor preditor de x_t a partir de $x_{t-1}, x_{t-2}, \dots, x_{t-h}$, é dado por

$$E[x_t | x_{t-1}, x_{t-2}, \dots, x_{t-h}] = - \sum_{j=1}^h a_h(j) x_{t-j}, \quad (3.3.4)$$

onde o vetor $\underline{a}_h = [a_h(1), a_h(2), \dots, a_h(h)]'$ é especificado pela equação

$$\underline{a}_h \underline{r}(h) = -[\gamma(1), \gamma(2), \dots, \gamma(h)]', \quad (3.3.5)$$

sendo $\underline{r}(h)$ a matriz, $h \times h$, cujo (j,k) -ésimo elemento é $\gamma(j-k)$, $j,k=1, \dots, h$.

De modo análogo, se $\underline{\tilde{r}}(h)$ é a matriz, $h \times h$, cujo (j,k) -ésimo elemento é $c(j-k)$, $j,k=1, \dots, h$, então o vetor $\underline{\tilde{a}}_h = [\tilde{a}_h(1), \dots, \tilde{a}_h(h)]'$ é definido por meio da equação

$$\underline{\tilde{a}}_h \underline{\tilde{r}}(h) = -[c(1), \dots, c(h)]'. \quad (3.3.6)$$

Finalmente, se $\underline{\hat{r}}(h)$ é a matriz, $h \times h$, cujo (j,k) -ésimo elemento é $\hat{c}(j-k)$, $j,k=1, \dots, h$, então o vetor $\underline{\hat{a}}_h = [\hat{a}_h(1), \dots, \hat{a}_h(h)]'$ é definido por

$$\underline{\hat{a}}_h \underline{\hat{r}}(h) = -[\hat{c}(1), \dots, \hat{c}(h)]'. \quad (3.3.7)$$

Por analogia a (3.2.3), (3.2.4) e (3.2.7), vamos definir

$$\eta_t = \sum_{j=0}^h a_h(j) y_{t-j}, \quad t=h+1, \dots, T \quad (3.3.8)$$

$$\psi_t = \sum_{j=0}^h a_h(j) u_{t-j}, \quad t=h+1, \dots, T \quad (3.3.9)$$

e

$$\epsilon_{t,h} = \sum_{j=0}^h a_h(j) x_{t-j}, \quad t=h+1, \dots, T \quad (3.3.10)$$

onde $a_h(0) = 1$.

Denotaremos por $\|\underline{x}\|_\infty$, a norma uniforme do vetor $\underline{x}$, isto é $\|\underline{x}\|_\infty = \max_j |x_j|$. Para uma matriz $\underline{A}$ teremos que $\|\underline{A}\|_\infty = \max_i \sum_j |a_{ij}|$.

Se $\{y_t\}$, $t=1,2,\dots$, é uma sequência de variáveis aleatórias e $\{g_t\}$, $t=1,2,\dots$, é uma sequência de números reais positivos, escreveremos $y_t = O_p(g_t)$ para denotar que y_t é limitado em probabilidade por g_t , o que significa que para todo $\epsilon > 0$, existe um número positivo M_ϵ , tal que

$$P\{|y_t| \geq M_\epsilon g_t\} \leq \epsilon$$

para todo t . Escreveremos $y_t = o_p(g_t)$ se para todo $\epsilon > 0$

$$\lim_{t \rightarrow \infty} P\{|y_t| > \epsilon g_t\} = 0.$$

Escreveremos $y_t = O(g_t)$ para significar que

$$\limsup_{t \rightarrow \infty} |y_t|/g_t < \infty,$$

quase certamente. Se " $< \infty$ " for substituído por " $= 0$ " na expressão acima, escreveremos $y_t = o(g_t)$.

3.4 A convergência das auto-covariâncias

Teoremas relatando a convergência das auto-covariâncias e auto-correlações amostrais para seus verdadeiros valores, para processos estacionários, foram demonstrados por An, Chen e Hannan (1982) e por Hannan e Kavalieris (1983). As razões de convergência são uniformes até o "lag" $H(T)$, onde $H(T)$ cresce com T . Nosso objetivo nesta seção, é apresentar um teorema que trata da convergência das auto-covariâncias amostrais de $\hat{x}_t$ para seus verdadeiros valores.

O teorema a seguir foi demonstrado por An, Chen e Hannan (1982).

Teorema 3.4.1. Seja $\{x_t\}$ um processo estacionário, ergódico, com média zero e variância finita, tal que:

$$x(t) = \sum_{j=0}^{\infty} k(j) \varepsilon_{t-j}, \quad \sum_{j=0}^{\infty} |k(j)| < \infty, \quad k(0)=1,$$

onde $K(Z) = \sum_{j=0}^{\infty} k(j)Z^j \neq 0$, para $|Z| \leq 1$, e $E(\varepsilon_t \varepsilon_s) = \delta_{st} \sigma^2$. Suponha para

$\{\varepsilon_t\}$ as seguintes condições:

$$E\{\varepsilon_t | \mathfrak{F}_{t-1}\} = 0, \quad (3.4.1)$$

$$E\{\varepsilon_t^2 | \mathfrak{F}_{t-1}\} = \sigma^2, \quad (3.4.2)$$

$$E\{\varepsilon_t^4\} < \infty, \quad (3.4.3)$$

onde $\mathfrak{F}_t$ é a σ -álgebra gerada por ε_s , $s \leq t$. Sob tais condições, se

$$\lim_{T \rightarrow \infty} T^{1/2} \sum_{j=T}^{\infty} k(j) = 0,$$

então

$$\sup_{0 \leq t < \infty} |c(t) - \gamma(t)| = O\{(\log T/T)^{1/2}\}.$$

Com relação às condições impostas a $\{\epsilon_t\}$, pode-se dizer que (3.4.1) garante que o melhor preditor linear é o melhor preditor, e a condição (3.4.2) é uma condição de uniformidade, necessária quando não se tem independência. Quando os ϵ_t são i.i.d., com média zero e variância σ^2 , (3.4.1) e (3.4.2) estão automaticamente satisfeitas.

A prova do teorema a seguir, pode ser vista em Ibragimov e Linnik (1971, Cap. 17).

Teorema 3.4.2. A seqüência Gaussiana $\{x_t\}$ é regular, se e somente se ela puder ser representada na forma

$$x_t = \sum_{j=0}^{\infty} c_j e_{t-j},$$

onde $\sum_{j=0}^{\infty} |c_j|^2 < \infty$, e os e_t são independentes e normalmente distribuídos.

Teorema 3.4.3. Considere o modelo (1.1.1), onde $\{x_t\}$, dado por (1.5.1), satisfaz (A1)-(A3) e $\{u_t\}$ satisfaz (B1)-(B3). Então

$$\sup_{1 \leq t < \infty} |\hat{c}(t) - \gamma(t)| = O_p \{(\log T/T)^{1/2}\}.$$

Prova. Temos

$$|\hat{c}(t) - \gamma(t)| \leq |\hat{c}(t) - c(t)| + |c(t) - \gamma(t)|.$$

As condições (A1)-(A3), garantem que estão satisfeitas as condições do teorema 3.4.2 e portando $\{x_t\}$ é regular e conseqüentemente ergódico. Essas condições garantem também as condições do teorema 3.4.1. Assim,

$$\sup_{0 \leq t < \infty} |c(t) - \gamma(t)| = O\{(\log T/T)^{1/2}\} \quad (3.4.4)$$

Por outro lado, para $t=0,1,\dots, T-1$, temos:

$$\begin{aligned} \hat{c}(t) - c(t) &= \frac{1}{T} \left\{ \sum_{s=1}^{T-t} \hat{x}_s \hat{x}_{s+t} - \sum_{s=1}^{T-t} x_s x_{s+t} \right\} \\ &= \frac{1}{T} \left\{ \sum_{s=1}^{T-t} (y_s - \hat{\beta}_{MQO}) (y_{s+t} - \hat{\beta}_{MQO}) - \sum_{s=1}^{T-t} x_s x_{s+t} \right\} \end{aligned}$$

De (1.1.1) decorre que

$$\begin{aligned} \hat{c}(t) - c(t) &= \frac{1}{T} \left\{ (\beta - \hat{\beta}_{MQO})^2 \sum_{s=1}^{T-t} u_s u_{s+t} + \right. \\ &\quad \left. + (\beta - \hat{\beta}_{MQO}) \left[\sum_{s=1}^{T-t} u_s x_{s+t} - \sum_{s=1}^{T-t} u_{s+t} x_s \right] \right\} \end{aligned}$$

e portanto,

$$\hat{c}(t) - c(t) = \frac{1}{T} (\beta - \hat{\beta}_{MQO})^2 \sum_{s=1}^T u_s^2 \frac{\sum_{s=1}^{T-t} u_s u_{s+t}}{\sum_{s=1}^T u_s^2} +$$

$$+ (\beta - \hat{\beta}_{MQO}) \left(\sum_{s=1}^T u_s^2 \right)^{1/2} \left[\frac{\sum_{s=1}^{T-t} u_s x_{s+t}}{\left(\sum_{s=1}^T u_s^2 \right)^{1/2}} + \frac{\sum_{s=1}^{T-t} u_{s+t} x_s}{\left(\sum_{s=1}^T u_s^2 \right)^{1/2}} \right]$$

Como

$$\lim_{T \rightarrow \infty} \frac{\sum_{s=1}^{T-t} u_s u_{s+t}}{\sum_{s=1}^T u_s^2} = r(t), \quad t=0,1,2,\dots,T-1$$

então

$$\frac{\sum_{s=1}^{T-t} u_s u_{s+t}}{\sum_{s=1}^T u_s^2} = o(1), \quad t=0,1,\dots,T-1.$$

Pelo teorema 2.3.2, temos que:

$$\left(\sum_{s=1}^T u_s^2 \right)^{1/2} (\hat{\beta}_{MQO} - \beta) = o_p(1)$$

$$\left(\sum_{s=1}^T u_s^2 \right) (\hat{\beta}_{MQO} - \beta)^2 = o_p(1),$$

e pelo teorema 2.3.1,

$$\left(\sum_{s=1}^T u_s^2 \right)^{-1/2} \sum_{s=1}^{T-t} u_s x_{s+t} = o_p(1), \quad t=0,1,\dots,T-1$$

Assim,

$$\sup_{0 \leq t < \infty} |\hat{c}(t) - c(t)| = O_p(T^{-1}), \quad (3.4.5)$$

e de (3.4.4) e (3.4.5) segue que:

$$\sup_{0 \leq t < \infty} |\hat{c}(t) - \gamma(t)| = O_p\{(\log T/T)^{1/2}\}. \quad \Delta$$

3.5 A convergência dos estimadores dos coeficientes auto-regressivos

Queremos estabelecer a convergência em probabilidade dos estimadores dos coeficientes auto-regressivos. Para chegarmos ao resultado desejado, necessitamos de dois resultados subsidiários demonstrados por Hannan e Kavalieris (1984a) e um demonstrado por Berk (1974). Os mesmos aparecem enunciados pelos teoremas a seguir.

Teorema 3.5.1. (Hannan e Kavalieris, 1984a). Seja $\{x_t\}$ gerado por (1.5.1), satisfazendo (A2) e (A3) e suponha que ϵ_t é tal que $E(\epsilon_t \epsilon_s) = \delta_{st} \sigma^2$ e adicionalmente satisfaz (3.4.1), (3.4.3) e

$$E\{\epsilon_t^2 | \mathcal{F}_{-\infty}\} = \sigma^2. \quad (3.5.1)$$

Se $H(T)=T$, então

$$\max_{1 \leq j \leq H(T)} \left[T^{-1} \sum_{t=j+1}^T \epsilon_t x_{t-j} \right] = O\{(\log T/T)^{1/2}\} \quad (3.5.2)$$

Teorema 3.5.2. (Hannan e Kavalieris, 1984a). Sob as condições do teorema 3.4.1,

$$\| \tilde{r}(h)^{-1} \|_{\infty} \text{ e } \| \tilde{f}(h)^{-1} \|_{\infty}$$

são uniformemente limitadas em $h \leq H(T) = o\{(T|\log T)^{1/2}\}$.

Teorema 3.5.3. (Berk, 1974). Seja $\{x_t\}$ satisfazendo (1.5.1) com $\sum_{j=0}^{\infty} |a(j)| < \infty$. Defina

$$A_h(Z) = \sum_{j=0}^h a_h(j) Z^j,$$

com $a_h(0)=1$ e $a_h(j)$, $j=1, \dots, h$, definidos por (3.3.5). Então

$$\lim_{h \rightarrow \infty} A_h(Z) = A(Z), \quad |Z| \leq 1.$$

Teorema 3.5.4. Seja $\{y_t\}$ gerado por (1.1.1), com $\{x_t\}$, dado por (1.5.1), satisfazendo (A1)-(A3) e $\{u_t\}$ satisfazendo (B1)-(B3). Então, para $H(T) = o\{(T|\log T)^{1/2}\}$,

$$\max_{1 \leq j \leq h} |\hat{a}_h(j) - a_h(j)| = O_p\{(\log T|T)^{1/2}\},$$

uniformemente em $h \leq H(T)$.

Prova. Temos que

$$\sum_{j=1}^h \{\hat{a}_h(j) - a_h(j)\} \hat{c}(j-k) = \sum_{j=1}^h \hat{a}_h(j) \hat{c}(j-k) - \sum_{j=1}^h a_h(j) \hat{c}(j-k).$$

Fazendo uso de (3.3.5) e (3.3.7), temos:

$$\sum_{j=1}^h \{\hat{a}_h(j) - a_h(j)\} \hat{c}(j-k) = -\hat{c}(k) + \gamma(k) + \sum_{j=1}^h a_h(j) \gamma(j-k)$$

$$- \sum_{j=1}^h a_h(j) \hat{c}(j-k).$$

Como $a(0)=1$, temos:

$$\sum_{j=1}^h \{\hat{a}_h(j) - a_h(j)\} \hat{c}(j-k) = - \sum_{j=0}^h a(j) \hat{c}(j-k) + \sum_{j=1}^h a(j) \hat{c}(j-k)$$

$$- \sum_{j=1}^h a(j) \gamma(j-k) + \sum_{j=0}^h a(j) \gamma(j-k)$$

$$+ \sum_{j=1}^h a_h(j) \gamma(j-k) - \sum_{j=1}^h a_h(j) \hat{c}(j-k).$$

Portanto,

$$\sum_{j=1}^h \{\hat{a}_h(j) - a_h(j)\} \hat{c}(j-k) = - \sum_{j=0}^h a(j) \hat{c}(j-k)$$

$$+ \sum_{j=1}^h \{a(j) - a_h(j)\} \{\hat{c}(j-k) - \gamma(j-k)\}$$

$$+ \sum_{j=0}^h a(j) \gamma(j-k) \quad (3.5.3)$$

Pelo teorema 3.4.3 o 2º termo de (3.5.3) é

$$O_p \{ (\log T/T)^{1/2} \} \sum_{j=1}^h \{a(j) - a_h(j)\},$$

e como pelo teorema 3.5.3

$$\sum_{j=1}^h \{a(j) - a_h(j)\} = o(1), \text{ uniformemente em } h,$$

o 2º termo de (3.5.3) é $o_p\{(\log T/T)^{1/2}\}$, uniformemente em h .

Como

$$\begin{aligned} \gamma(t) &= E[x_s x_{s+t}] = E\left[\sum_{j=0}^{\infty} b(j) \epsilon_{s-j} \sum_{k=0}^{\infty} b(k) \epsilon_{s+t-k}\right] \\ &= \sigma_{\epsilon}^2 \sum_{j=0}^{\infty} b(j)b(t+j), \end{aligned}$$

então

$$|\gamma(t)| \leq \sigma_{\epsilon}^2 \sum_{j=0}^{\infty} |b(j)| \sum_{j=0}^{\infty} |b(j+t)| = \sigma_{\epsilon}^2 \sum_{j=0}^{\infty} |b(j)| \sum_{u=t}^{\infty} |b(u)|,$$

e, por (A2)

$$\lim_{t \rightarrow \infty} t^{1/2} |\gamma(t)| \leq M \lim_{t \rightarrow \infty} t^{1/2} \sum_{u=t}^{\infty} |b(u)| = 0,$$

onde M é uma constante real, o que implica que $t^{1/2} \gamma(t) = o(1)$. Assim, novamente por (A2)

$$\left| \sum_{j=T+1}^{\infty} a(j) \gamma(j-k) \right| = O(T^{-1/2}), \quad (3.5.4)$$

e conseqüentemente para o último termo de (3.5.3) temos:

$$\begin{aligned}
 \sum_{j=0}^h a(j)\gamma(j-k) &= - \sum_{j=h+1}^{\infty} a(j)\gamma(j-k) = \\
 &= - \sum_{j=h+1}^T a(j)\gamma(j-k) - \sum_{j=T+1}^{\infty} a(j)\gamma(j-k) = \\
 &= \sum_{j=h+1}^T a(j)\{\gamma(j-k) - \hat{c}(j-k)\} - \sum_{j=h+1}^T a(j)\hat{c}(j-k) + O(T^{-1/2}) \\
 &= - \sum_{j=h+1}^T a(j)\hat{c}(j-k) + o_p\{(\log T/T)^{1/2}\} - \sum_{j=h+1}^T |a(j)| + O(T^{-1/2}) \\
 &= - \sum_{j=h+1}^T a(j)\hat{c}(j-k) + o_p\{(\log T/T)^{1/2}\}.
 \end{aligned}$$

Voltando a (3.5.3) obtemos:

$$\sum_{j=1}^h \{\hat{a}_h(j) + a_h(j)\} \hat{c}(j-k) = - \sum_{j=0}^T a(j)\hat{c}(j-k) + o_p\{(\log T/T)^{1/2}\}. \quad (3.5.5)$$

Por (3.4.5) o 1º termo de (3.5.5) é

$$- \sum_{j=0}^T a(j)c(j-k) + o_p(T^{-1}), \quad (3.5.6)$$

e o 1º termo de (3.5.6) é

$$\frac{1}{T} \sum_{j=0}^T a(j) \sum_{t=k+1}^T x_{t-j} x_{t-k},$$

como demonstrado por Hannan e Kavalieris (1986).

Portanto,

$$\begin{aligned} \sum_{j=0}^T a(j)c(j-k) &= \frac{1}{T} \sum_{t=k+1}^T \left[\sum_{j=0}^{\infty} a(j)x(t-j) - \sum_{j=t+1}^{\infty} a(j)x(t-j) \right] x(t-k) \\ &= \frac{1}{T} \sum_{t=k+1}^T \epsilon(t)x(t-k) + o(T^{-1/2}), \end{aligned} \quad (3.5.7)$$

pois

$$\sum_{j=T+1}^{\infty} a(j)x(t-j) = o(T^{-1/2}) \quad (\text{Hannan e Kavalieris, 1984a}).$$

Substituindo (3.5.7) em (3.5.5), obtemos:

$$\sum_{j=1}^h \{\hat{a}_h(j) - a_h(j)\} \hat{c}(j-k) = -\frac{1}{T} \sum_{t=k+1}^T \epsilon(t)x_{t-k} + o_p\{(\log T/T)^{1/2}\}$$

e pelo teorema 3.5.1

$$\max_{1 \leq k \leq h} \sum_{j=1}^h \{\hat{a}_h(j) - a_h(j)\} \hat{c}(j-k) = o_p\{(\log T/T)^{1/2}\}, \quad (3.5.8)$$

uniformemente em $h \leq H(T)$, $H(T) = o\{(T|\log T)^{1/2}\}$.

Por outro lado,

$$\begin{aligned} \max_{1 \leq j \leq h} |\hat{a}_h(j) - a_h(j)| &= \|\hat{a}_h - a_h\|_{\infty} = \\ &= \|\hat{f}(h)^{-1} \hat{f}(h) [\hat{a}_h - a_h]\|_{\infty} \end{aligned}$$

$$\leq || \hat{r}(h)^{-1} ||_{\infty} || \hat{r}(h) [\hat{a}_{\sim h} - a_{\sim h}] ||_{\infty}. \quad (3.5.9)$$

Como

$$|| \hat{r}(h) [\hat{a}_{\sim h} - a_{\sim h}] ||_{\infty} = \max_{1 \leq k \leq h} \sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] \hat{c}(j-k),$$

então pelos teoremas 3.5.2 e 3.4.3 e por (3.5.8) e (3.5.9), temos:

$$\max_{1 \leq j \leq h} |\hat{a}_h(j) - a_h(j)| = O_p \{ (\log T/T)^{1/2} \},$$

uniformemente em $h \leq H(T)$. Δ

3.6 A consistência de $\hat{\beta}$

Nesta seção provaremos a consistência do estimador $\hat{\beta}$, dado em (3.2.5). A prova decorre do fato de que a sequência $\{\psi_t\}$ satisfaz as mesmas condições que $\{u_t\}$, como também do teorema 3.5.4.

Teorema 3.6.1. Sob as suposições do teorema 3.5.4, sendo $\hat{\beta}$ definido por (3.2.5) e

$$\hat{\beta}^* = \frac{\sum_{t=h+1}^T \eta_t \psi_t}{\sum_{t=h+1}^T \psi_t^2}, \quad (3.6.1)$$

então

$$\left(\sum_{t=1}^T u_t^2 \right)^{1/2} (\hat{\beta}^* - \hat{\beta}) = o_p(1),$$

uniformemente em $h \leq H(T)$.

Prova. Temos que

$$\hat{\eta}_t = y_t + \sum_{j=1}^h a_h(j) y_{t-j} + \sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] y_{t-j}$$

$$= \eta_t + \sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] y_{t-j},$$

$$\hat{\psi}_t = \psi_t + \sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] u_{t-j},$$

$$\hat{\epsilon}_{t,h} = \epsilon_{t,h} + \sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] x_{t-j}.$$

Assim,

$$\begin{aligned} \left(\sum_{t=1}^T u_t^2 \right)^{-1/2} \sum_{t=h+1}^T \hat{\psi}_t \hat{\epsilon}_{t,h} &= \left(\sum_{t=1}^T u_t^2 \right)^{-1/2} \left\{ \sum_{t=h+1}^T \psi_t \epsilon_{t,h} \right. \\ &\quad + \sum_{t=h+1}^T \psi_t \sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] x_{t-j} \\ &\quad + \sum_{t=h+1}^T \epsilon_{t,h} \sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] u_{t-j} \\ &\quad \left. + \sum_{t=h+1}^T \sum_{j=1}^h \sum_{i=1}^h [\hat{a}_h(j) - a_h(j)] [\hat{a}_h(i) - a_h(i)] u_{t-j} x_{t-i} \right\}. \quad (3.6.2) \end{aligned}$$

O 2º termo do lado direito de (3.6.2) é:

$$\sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] \frac{\sum_{t=h+1}^T \psi_t x_{t-j}}{\left(\sum_{t=1}^T u_t^2 \right)^{1/2}} = \sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] \left\{ \frac{\sum_{t=h+1}^T u_t x_{t-j}}{\left(\sum_{t=1}^T u_t^2 \right)^{1/2}} + \right. \\ \left. + \frac{a_h(1) \sum_{t=h+1}^T u_{t-1} x_{t-j}}{\left(\sum_{t=1}^T u_t^2 \right)^{1/2}} + \dots + \frac{a_h(h) \sum_{t=h+1}^T u_{t-h} x_{t-j}}{\left(\sum_{t=1}^T u_t^2 \right)^{1/2}} \right\} \quad (3.6.3)$$

Como, pelo teorema 2.3.1

$$(\sum u_t^2)^{-1/2} \sum u_{t-k} x_{t-j} = o_p(1),$$

para todo $k, j=1, \dots, h$, conclui-se então, pelo teorema 3.5.4 e pelo fato de $h \leq H(T) = o\{(T \log T)^{1/2}\}$, que (3.6.3) é $o_p(1)$, uniformemente em $h \leq H(T)$. De modo análogo, mostra-se que o 3º e 4º termos de (3.6.2) são $o_p(1)$. Assim, (3.6.2) fica

$$\left(\sum_{t=1}^T u_t^2 \right)^{-1/2} \sum_{t=h+1}^T \hat{\psi}_t \hat{\epsilon}_{t,h} = \left(\sum_{t=1}^T u_t^2 \right)^{-1/2} \sum_{t=h+1}^T \psi_t \epsilon_{t,h} + o_p(1). \quad (3.6.4)$$

Por outro lado,

$$\left(\sum_{t=1}^T u_t^2 \right)^{-1} \sum_{t=h+1}^T \hat{\psi}_t^2 = \left(\sum_{t=1}^T u_t^2 \right)^{-1} \left\{ \sum_{t=1}^T \psi_t^2 + 2 \sum_{t=h+1}^T \psi_t \sum_{j=1}^h |\hat{a}_h(j) - a_h(j)| u_{t-j} \right.$$

$$+ \left\{ \sum_{t=h+1}^T \sum_{j=1}^h \sum_{i=1}^h [\hat{a}_h(j) - a_h(j)][\hat{a}_h(i) - a_h(i)] u_{t-j} u_{t-i} \right\} \quad (3.6.5)$$

O 2º termo do lado direito de (3.6.5) é:

$$2 \sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] \frac{\sum_{t=h+1}^T \psi_t u_{t-j}}{\sum_{t=1}^T u_t^2} = 2 \sum_{j=1}^h [\hat{a}_h(j) - a_h(j)] \left[\frac{\sum_{t=h+1}^T u_{t-j} u_t}{\sum_{t=1}^T u_t^2} + \right. \\ \left. + \frac{a_1(h) \sum_{t=h+1}^T u_{t-1} u_{t-j}}{\sum_{t=1}^T u_t^2} + \dots + \frac{a_h(h) \sum_{t=h+1}^T u_{t-h} u_{t-j}}{\sum_{t=1}^T u_t^2} \right]$$

e por (B3), pelo teorema 3.5.4 e por $h \leq H(T) = o\{(T/\log T)^{1/2}\}$ esse termo é $o_p(1)$, uniformemente em $h \leq H(T)$. De modo análogo, o 3º termo do lado direito de (3.6.5) é $o_p(1)$, uniformemente em $h \leq H(T)$. Assim,

$$\left(\sum_{t=1}^T u_t^2 \right)^{-1} \sum_{t=h+1}^T \hat{\psi}_t^2 = \left(\sum_{t=1}^T u_t^2 \right)^{-1} \sum_{t=h+1}^T \psi_t^2 + o_p(1) \quad (3.6.6)$$

uniformemente em $h \leq H(T)$.

De (3.6.4) e (3.6.6), temos:

$$\frac{\left(\sum_{t=1}^T u_t^2 \right)^{-1/2} \sum_{t=h+1}^T \hat{\psi}_t \hat{\epsilon}_{t,h}}{\left(\sum_{t=1}^T u_t^2 \right)^{-1} \sum_{t=h+1}^T \hat{\psi}_t^2} = \frac{\left(\sum_{t=1}^T u_t^2 \right)^{-1/2} \sum_{t=h+1}^T \psi_t \epsilon_{t,h} + o_p(1)}{\left(\sum_{t=1}^T u_t^2 \right)^{-1} \sum_{t=h+1}^T \psi_t^2 + o_p(1)},$$

e portanto,

$$\left(\sum_{t=1}^T u_t^2 \right)^{1/2} \frac{\sum_{t=h+1}^T \hat{\psi}_t \hat{\epsilon}_{t,h}}{\sum_{t=h+1}^T \hat{\psi}_t^2} = \left(\sum_{t=1}^T u_t^2 \right)^{1/2} \frac{\sum_{t=h+1}^T \psi_t \epsilon_{t,h}}{\sum_{t=h+1}^T \psi_t^2} + o_p(1). \quad (3.6.7)$$

Como o 1º membro de 3.6.7 é $\left(\sum_{t=1}^T u_t^2 \right)^{1/2} (\hat{\beta} - \beta)$ e o 1º termo do 2º membro é $\left(\sum_{t=1}^T u_t^2 \right)^{1/2} (\hat{\beta}^* - \beta)$, temos:

$$\left(\sum_{t=1}^T u_t^2 \right)^{1/2} (\hat{\beta}^* - \hat{\beta}) = o_p(1), \text{ uniformemente em } h \leq H(T). \quad \Delta$$

Teorema 3.6.2. Sob as suposições do teorema (3.5.4), $\hat{\beta}^* \xrightarrow{P} \beta$, quando $T \rightarrow \infty$, para todo $h \leq H(T)$, onde $\hat{\beta}^*$ foi definido no teorema 3.6.1.

Prova. Iremos inicialmente mostrar que $\{\psi_t\}$ satisfaz às condições (B1)-(B3)

(I) Vale (B1)

$$\begin{aligned} \sum_{t=h+1}^T \psi_t^2 &= \sum_{t=h+1}^T \left[u_t + \sum_{j=0}^h a_h(j) u_{t-j} \right]^2 = \sum_{j=0}^h \sum_{t=h+1}^T a_h^2(j) u_{t-j}^2 \\ &\quad + \sum_{\substack{i,j \\ i \neq j}} \sum_{t=h+1}^T a_h(i) a_h(j) u_{t-i} u_{t-j}, \end{aligned}$$

onde $a_h(0)=1$.

A condição (A2) implica que:

$$\sum_{j=0}^h a_h^2(j) = O(1), \text{ uniformemente em } h \leq H(T),$$

e

$$\sum_i \sum_j a_h(i) a_h(j) \leq \sum_i |a_h(i)| \sum_j |a_h(j)| = O(1),$$

uniformemente em $h \leq H(T)$.

Portanto,

$$\sum_{t=h+1}^T \psi_t^2 = O(1) \sum_{t=h+1}^T u_{t-j}^2 + O(1) \sum_{t=h+1}^T u_{t-i} u_{t-j} \quad (3.6.8)$$

e, pelas condições impostas a u_t ,

$$\lim_{T \rightarrow \infty} \sum_{t=h+1}^T \psi_t^2 = \infty, \quad h \leq H(T). \quad (3.6.9)$$

De (3.6.8) decorre também que

$$\frac{\sum_{t=h+1}^T \psi_t^2}{\sum_{t=1}^T u_t^2} = O(1) \quad (3.6.10)$$

(II) Vale (B2)

$$\frac{\psi_T^2}{\sum_{t=h+1}^T \psi_t^2} = \frac{\sum_{t=1}^T u_t^2}{\sum_{t=h+1}^T \psi_t^2} \cdot \frac{\sum_{j=0}^h a_h^2(j) u_{T-j}^2 + \sum_{\substack{i,j \\ i \neq j}} a_h(i) a_h(j) u_{T-i} u_{T-j}}{\sum_{t=1}^T u_t^2}$$

Assim, de (A2), (B2) e (3.6.10) segue que:

$$\lim_{T \rightarrow \infty} \frac{\frac{\sum_{t=h+1}^T \psi_t^2}{T}}{\frac{\sum_{t=h+1}^T \psi_t^2}{T}} = 0.$$

(III) Vale (B3)

$$\frac{\frac{\sum_{t=h+1}^T \psi_t \psi_{t+|\ell|}}{T}}{\frac{\sum_{t=h+1}^T \psi_t^2}{T}} = \frac{\frac{\sum_{t=1}^T u_t^2}{T}}{\frac{\sum_{t=h+1}^T \psi_t^2}{T}} \cdot \frac{\sum_{t=h+1}^T (u_{t+\sum_{h=1}^h a_h(j)} u_{t-j}) (u_{t+|\ell|+\sum_{h=1}^h a_h(j)} u_{t+|\ell|-j})}{\frac{\sum_{t=1}^T u_t^2}{T}}$$

e analogamente, por (A2), (B3) e (3.6.10)

$$\lim_{T \rightarrow \infty} \frac{\frac{\sum_{t=h+1}^T \psi_t \psi_{t+|\ell|}}{T}}{\frac{\sum_{t=h+1}^T \psi_t^2}{T}} = \rho(\ell), \quad \ell=0, \pm 1, \dots \quad (3.6.11)$$

Como

$$\hat{\beta}^{*- \beta} = \frac{\sum_{t=h+1}^T \psi_t^{\epsilon_{t,h}}}{\sum_{t=h+1}^T \psi_t^2}$$

então

$$\left(\sum_{t=h+1}^T \psi_t^2 \right)^{1/2} (\hat{\beta}^{*- \beta}) = \left(\sum_{t=h+1}^T \psi_t^2 \right)^{-1/2} \left[\sum_{t=h+1}^T \psi_t x_{t+} + \sum_{j=1}^h a_h(j) \sum_{t=h+1}^T \psi_t x_{t-j} \right].$$

Como, pelo teorema 2.3.1

$$\left(\sum_{t=h+1}^T \psi_t^2 \right)^{1/2} \sum_{t=h+1}^T \psi_t x_{t-j} = O_p(1), \quad j=1, \dots, h, \quad h \leq H(T),$$

então

$$\left(\sum_{t=h+1}^T \psi_t^2 \right)^{1/2} (\hat{\beta}^* - \beta) = O_p(1), \quad h \leq H(T). \quad \Delta$$

Teorema 3.6.3. Sob as condições do teorema 3.5.4, $\hat{\beta} \xrightarrow{\mathcal{P}} \beta$, quando $T \rightarrow \infty$, uniformemente em $h \leq H(T)$.

Prova. Pelo teorema 3.6.1, temos que:

$$\left(\sum_{t=1}^T u_t^2 \right)^{1/2} (\hat{\beta}^* - \hat{\beta}) = o_p(1), \quad \text{uniformemente em } h \leq H(T),$$

e pelo teorema 3.6.2

$$\left(\sum_{t=h+1}^T \psi_t^2 \right)^{1/2} (\hat{\beta}^* - \beta) = O_p(1), \quad \text{uniformemente em } h \leq H(T).$$

Esses dois resultados juntamente com (3.6.10) levam a

$$\begin{aligned} \left(\sum_{t=1}^T u_t^2 \right)^{1/2} (\hat{\beta} - \beta) &= \left(\sum_{t=1}^T u_t^2 \right)^{1/2} (\hat{\beta} - \hat{\beta}^*) + \left(\sum_{t=1}^T u_t^2 \right)^{1/2} (\hat{\beta}^* - \beta) \\ &= O_p(1), \quad h \leq H(T) \end{aligned}$$

de onde segue que

$$\hat{\beta} \xrightarrow{\mathcal{P}} \beta, \quad \text{uniformemente em } h \leq H(T). \quad \Delta$$

3.7 A normalidade assintótica de $\hat{\beta}$

A normalidade assintótica de $\hat{\beta}$ será provada com base num teorema do limite central demonstrado por Hannan (1973). O teorema de Hannan mostra a distribuição limite de

$$\frac{\sum_{t=1}^T u_t x_t}{\left(\sum_{t=1}^T u_t^2 \right)^{1/2}} \quad (3.7.1)$$

em quatro situações distintas, cada uma impondo um conjunto de condições a respeito de $\{u_t\}$ e $\{x_t\}$, necessárias à obtenção do resultado desejado. O enunciado que apresentaremos abaixo representa apenas a situação que nos interessa.

Teorema 3.7.1. (Hannan, 1973). Seja $x(t)$ uma seqüência estacionária regular, satisfazendo:

$$x(t) = \sum_{j=0}^{\infty} k(j) \varepsilon(t-j), \quad \sum |k(j)| < \infty, \quad k(0) = 1,$$

onde $K(Z) = \sum_{j=0}^{\infty} k(j) Z^j \neq 0$, para $|Z| \leq 1$. Suponha ainda que $E(\varepsilon_t \varepsilon_s) = \delta_{st} \sigma^2$ e que $\{\varepsilon_t\}$ satisfaz (3.4.1). Seja $\{u_t\}$ uma seqüência determinística, satisfazendo (B1)-(B3) e $F_u(\lambda)$ a função definida em (1.5.2). Assumindo que

$$\int_{-\pi}^{\pi} f(\lambda) F_u(d\lambda) > 0 \quad \text{a.s.}, \quad (3.7.2)$$

onde

$$f(\lambda) = \frac{\sigma_\epsilon^2}{2\pi} \left| \sum_{j=0}^{\infty} k(j) e^{ij\lambda} \right|^2,$$

a distribuição de (3.7.1) converge para a normal com média zero e variância a.s.

$$\int_{-\pi}^{\pi} 2\pi f(\lambda) F_u(d\lambda). \quad (3.7.3)$$

Teorema 3.7.2. Sob as condições do teorema 3.5.4 e supondo que (3.7.2) vale, para $f(\lambda)$ substituído por

$$f_h(\lambda) = \frac{\sigma_\epsilon^2}{2\pi} \left| \sum_{k=0}^{\infty} \sum_{j=0}^h a_h(j) b(k-j) e^{ik\lambda} \right|^2,$$

a distribuição assintótica de

$$\left(\sum_{t=h+1}^T \psi_t^2 \right)^{1/2} (\hat{\beta}^* - \beta)$$

é normal, com média zero e variância dada por

$$\int_{-\pi}^{\pi} 2\pi f_h(\lambda) F_\psi(d\lambda), \quad h \leq H(T), \quad (3.7.4)$$

onde $F_\psi(\lambda)$ é a função definida por

$$\rho(\ell) = \lim_{T \rightarrow \infty} \frac{\sum_{t=h+1}^{T-|\ell|} \psi_t \psi_{t+|\ell|}}{\sum_{t=h+1}^T \psi_t^2} = \int_{-\pi}^{\pi} e^{i\ell\lambda} F_\psi(d\lambda) \quad (3.7.5)$$

Prova. Sabemos que:

$$\left(\sum_{t=h+1}^T \psi_t^2 \right)^{1/2} (\hat{\beta}^* - \beta) = \frac{\sum_{t=h+1}^T \psi_t \epsilon_{t,h}}{\left(\sum_{t=h+1}^T \psi_t^2 \right)^{1/2}}$$

onde

$$\epsilon_{t,h} = \sum_{j=0}^h a_h(j) x_{t-j}, \quad a(0)=1.$$

De (3.3.3), segue que:

$$\begin{aligned} \epsilon_{t,h} &= \sum_{j=0}^h a_h(j) \sum_{i=0}^{\infty} b(i) \epsilon_{t-i-j} \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^h a_h(j) b(i) \epsilon_{t-i-j}. \end{aligned}$$

Fazendo $k=i+j$ e

$$c_h(k) = \sum_{j=0}^h a_h(j) b(k-j),$$

temos:

$$\epsilon_{t,h} = \sum_{k=0}^{\infty} c_h(k) \epsilon_{t-k} \quad (3.7.6)$$

onde

$$\sum_{k=0}^{\infty} |c_h(k)| < \infty, \quad h \leq H(T), \quad (3.7.7)$$

por causa da condição (A2). A representação (3.7.6) com (3.7.7), garante que $\{\epsilon_{t,h}\}$ é regular para todo $h \leq H(T)$. Como $\{\psi_t\}$ satisfaz (B1)-(B3), as condições do teorema 3.7.1 estão satisfeitas e o teorema

está provado. Δ

Teorema 3.7.3. Sob as suposições do teorema anterior, a distribuição assintótica de

$$\left(\sum_{t=1}^T u_t^2 \right)^{1/2} (\hat{\beta} - \beta)$$

é normal, com média zero e variância dada por

$$V_h = \int_{-\pi}^{\pi} 2 \pi f_h(\lambda) F^*(d\lambda), \quad h \leq H(T) \quad (3.7.8)$$

onde $F^*(\lambda)$ é tal que

$$v(\ell) = \lim_{T \rightarrow \infty} \frac{\sum_{t=1}^T u_t^2 \cdot \sum_{t=h+1}^{T-|\ell|} \psi_t \psi_{t+|\ell|}}{\left(\sum_{t=h+1}^T \psi_t^2 \right)^2} = \int_{-\pi}^{\pi} e^{i\ell\lambda} F^*(d\lambda) \quad (3.7.9)$$

Prova. É consequência imediata dos teoremas (3.6.1) e (3.7.2). Δ

Note que o limite em 3.7.9 existe, pois de (3.6.10) e (3.6.11), temos:

$$v(\ell) = \lim_{T \rightarrow \infty} \frac{\sum u_t^2}{\sum \psi_t^2} \cdot \frac{\sum_{t=h+1}^T \psi_t \psi_{t+|\ell|}}{\sum_{t=h+1}^T \psi_t^2} = k \cdot \rho(\ell),$$

onde k é uma constante real.

3.8 O caso vetorial

No caso em que consideramos o modelo geral (1.1.1), onde $\underline{u}_t$ é um vetor $k \times 1$, o estimador de COG2 é obtido pelo procedimento descrito na seção 3.2, substituindo-se $\hat{\psi}_t$ por

$$\hat{\psi}_{i,t} = \sum_{j=h+1}^T a_h(j) u_{i,t-j}, \quad i=1, \dots, k.$$

Assim, $\hat{\beta} = \hat{\beta}_{\text{COG2}} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k)'$, onde

$$\hat{\beta}_i = \frac{\sum_{t=h+1}^T \hat{\eta}_t \hat{\psi}_{i,t}}{\sum_{t=h+1}^T \psi_{i,t}^2}, \quad i=1, \dots, k \quad (3.8.1)$$

Pode-se, analogamente ao que foi feito para $k=1$, mostrar que $\hat{\beta}$ é um estimador consistente de β , e que a distribuição limite de $D_T(\hat{\beta} - \beta)$ é normal, onde D_t é dada em (1.5.3).

3.9 A eficiência assintótica de $\hat{\beta}$

De (2.3.2) segue que a variância assintótica de $(\sum_{t=1}^T u_t^2)^{1/2} \times (\hat{\beta}_{\text{MQO}} - \beta)$ é dada por:

$$\begin{aligned} \sum_{\ell=-\infty}^{\infty} r(\ell) \gamma(\ell) &= \int_{-\pi}^{\pi} \left| \sum_{\ell=-\infty}^{\infty} e^{i\ell\lambda} \gamma(\ell) \right| F_u(d\lambda) = \\ &= 2\pi \int_{-\pi}^{\pi} f_x(\lambda) F_u(d\lambda), \end{aligned} \quad (3.9.1)$$

onde $f_x(\lambda)$ é o espectro de x_t .

Por outro lado, a variância assintótica de $(\Sigma u_t^2)^{1/2}(\hat{\beta}_{MQG} - \beta)$ é dada por

$$2\pi \left\{ \int_{-\pi}^{\pi} \frac{1}{f_x(\lambda)} F_u(d\lambda) \right\}^{-1}. \quad (3.9.2)$$

Grenander e Rosemblat (1957), compararam (3.9.1) com (3.9.2), para mostrar sob que condições o estimador de MQO é assintoticamente eficiente.

No caso do estimador $\hat{\beta} = \hat{\beta}_{COG2}$, proposto neste capítulo, obtivemos que a variância assintótica de $(\Sigma u_t^2)^{1/2}(\hat{\beta} - \beta)$ é dada por

$$V_h = \int_{-\pi}^{\pi} 2\pi f_h(\lambda) F^*(d\lambda). \quad (3.9.3)$$

Vejamos o que ocorre com V_h quando $h \rightarrow \infty$.

Sabemos que:

$$\begin{aligned} f_h(\lambda) &= \frac{\sigma^2}{2\pi} \left| \sum_{k=0}^{\infty} \sum_{j=0}^h a_h(j) b(k-j) e^{ik\lambda} \right|^2 \\ &= \frac{\sigma^2}{2\pi} \left| \sum_{k=0}^{\infty} b(k-j) e^{i(k-j)\lambda} \sum_{j=0}^h a_h(j) e^{ij\lambda} \right|^2. \end{aligned}$$

Como

$$\begin{aligned} \lim_{h \rightarrow \infty} A_h(e^{i\lambda}) &= \lim_{h \rightarrow \infty} \sum_{j=0}^h a_h(j) e^{ij\lambda} \\ &= \sum_{j=0}^{\infty} a(j) e^{ij\lambda} = A(e^{i\lambda}), \end{aligned}$$

e

$$A^{-1}(e^{i\lambda}) = B(e^{i\lambda}),$$

então

$$\lim_{h \rightarrow \infty} f_h(\lambda) = \frac{\sigma^2}{2\pi}.$$

Por outro lado temos:

$$\begin{aligned} \lim_{h \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{\sum_{t=h+1}^T \psi_t^2}{\sum_{t=1}^T u_t^2} &= \lim_{h \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{\sum_{t=h+1}^T \sum_{j=0}^h \sum_{k=0}^h a_h(j) a_h(k) u_{t-j} u_{t-k}}{\sum_{t=1}^T u_t^2} \\ &= \lim_{h \rightarrow \infty} \sum_{j=0}^h \sum_{k=0}^h a_h(j) a_h(k) \lim_{T \rightarrow \infty} \frac{\sum_{t=h+1}^T u_{t-k} u_{t-j}}{\sum_{t=1}^T u_t^2} \\ &= \lim_{h \rightarrow \infty} \sum_{j=0}^h \sum_{k=0}^h a_h(j) a_h(k) r(j-k) \\ &= \lim_{h \rightarrow \infty} \sum_{j=0}^h \sum_{k=0}^h a_h(j) a_h(k) \int_{-\pi}^{\pi} e^{i(j-k)\lambda} F_u(d\lambda) \\ &= \lim_{h \rightarrow \infty} \int_{-\pi}^{\pi} A_h(e^{i\lambda}) A_h^*(e^{i\lambda}) F_u(d\lambda), \end{aligned}$$

onde $A_h^*(e^{i\lambda})$ é o complexo conjugado de $A_h(e^{i\lambda})$.

Portanto,

$$\lim_{h \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{\sum_{t=h+1}^T \psi_t^2}{\sum_{t=1}^T u_t^2} = \int_{-\pi}^{\pi} A(e^{i\lambda}) A^*(e^{i\lambda}) F_u(d\lambda). \quad (3.9.4)$$

Lembrando que

$$f_x(\lambda) = \frac{\sigma^2}{2\pi} \left[A(e^{i\lambda}) A^*(e^{i\lambda}) \right]^{-1},$$

de (3.9.4) segue que:

$$\lim_{h \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{\sum_{t=1}^T \psi_t^2}{\sum_{t=1}^T u_t^2} = \int_{-\Pi}^{\Pi} \frac{\sigma^2}{2\Pi} f_X^{-1}(\lambda) F_U(d\lambda). \quad (3.9.5)$$

De (3.7.9) e (3.9.5) temos:

$$\lim_{h \rightarrow \infty} v(0) = \lim_{h \rightarrow \infty} \int_{-\Pi}^{\Pi} F^*(d\lambda) = \frac{2\Pi}{\sigma^2} \left[\int_{-\Pi}^{\Pi} \frac{1}{f_X(\lambda)} F_U(d\lambda) \right]^{-1}$$

Assim,

$$\begin{aligned} \lim_{h \rightarrow \infty} V_h &= \frac{\sigma^2}{2\Pi} \cdot \frac{2\Pi}{\sigma^2} \cdot 2\Pi \left[\int_{-\Pi}^{\Pi} \frac{1}{f_X(\lambda)} F_U(d\lambda) \right]^{-1} \\ &= 2\Pi \left[\int_{-\Pi}^{\Pi} \frac{1}{f_X(\lambda)} F_U(d\lambda) \right]^{-1}, \end{aligned}$$

o que prova a eficiência assintótica do estimador proposto.

CAPÍTULO 4

COMPARAÇÃO ENTRE O ESTIMADOR DE COCHRANE-ORCUTT E O ESTIMADOR DE MÍNIMOS QUADRADOS ORDINÁRIOS QUANDO A VARIÁVEL INDEPENDENTE É ESTOCÁSTICA

4.1 Introdução

Vamos considerar aqui o modelo

$$y_t = \beta u_t + x_t, \quad t = 1, \dots, T \quad (4.1.1)$$

onde $\{x_t\}$ é um processo AR(1), isto é,

$$x_t + \phi x_{t-1} = \epsilon_t, \quad |\phi| < 1. \quad (4.1.2)$$

Considerando as transformações

$$\eta_t = y_t + \phi y_{t-1}, \quad t = 2, \dots, T \quad (4.1.3)$$

$$\psi_t = u_t + \phi u_{t-1}, \quad t = 2, \dots, T, \quad (4.1.4)$$

já vimos que o estimador de Cochrane-Orcutt de β é dado por:

$$\hat{\beta}_\phi = \frac{\sum_{t=2}^T \eta_t \psi_t}{\sum_{t=2}^T \psi_t^2}. \quad (4.1.5)$$

Nosso objetivo agora é comparar o estimador $\hat{\beta}_\phi$ com o estimador $\hat{\beta}_{MQO}$ em termos de suas variâncias assintóticas, supondo que:

- (i) $\{\epsilon_t\}$ são variáveis aleatórias i.i.d $N(0, \sigma_\epsilon^2)$
- (ii) $\{u_t\}$ é sequência estocástica, estacionária, normal com média zero
- (iii) $\{u_t\}$ e $\{x_t\}$ são independentes.

4.2 Variâncias assintóticas dos estimadores $\hat{\beta}_{MQO}$ e $\hat{\beta}_\phi$

Se $\hat{\beta}$ é um estimador qualquer de β , sabemos que sua variância assintótica, denotada por $A Var \hat{\beta}$, é dada por:

$$A Var \hat{\beta} = T^{-1} \lim_{T \rightarrow \infty} E[\sqrt{T} (\hat{\beta} - \beta)]^2 \quad (4.2.1)$$

Denotando por $\gamma_u(t)$ e $\gamma_x(t)$ as funções de autocovariância de u_t e x_t e por $\rho_u(t)$ e $\rho_x(t)$ as respectivas funções de autocorrelação, veremos a seguir como fica (4.2.1) para $\hat{\beta}_{MQO}$ e para $\hat{\beta}_\phi$.

Para $\hat{\beta}_{MQO}$ temos:

$$\hat{\beta}_{MQO} - \beta = \frac{\sum_{t=1}^T u_t x_t}{\sum_{t=1}^T u_t^2} \quad (4.2.2)$$

Assim,

$$E[\sqrt{T} (\hat{\beta}_{MQO} - \beta)]^2 = E \left[\frac{\frac{1}{\sqrt{T}} \sum_{t=1}^T u_t x_t}{\frac{1}{T} \sum_{t=1}^T u_t^2} \right]^2$$

Como

$$\frac{\sum u_t^2}{T} \xrightarrow{p} \gamma_u(0) = E(u_t^2) ,$$

pelo teorema de Isserlis,

$$E \left[\frac{1}{\sqrt{T}} \sum u_t x_t \right]^2 = \frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T E(u_t x_t u_s x_s)$$

$$= \frac{1}{T} \left[\sum_{t=1}^T \sum_{s=1}^T E(u_t u_s) E(x_t x_s) \right] +$$

$$+ \frac{1}{T} \left[\sum_{t=1}^T \sum_{s=1}^T E(u_t x_s) E(x_t u_s) \right] ,$$

e pela independência entre u_t e x_t

$$E \left[\frac{1}{\sqrt{T}} \sum u_t x_t \right]^2 = \frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T \gamma_u(t-s) \gamma_x(t-s) .$$

Então,

$$\lim_{T \rightarrow \infty} [\sqrt{T} (\hat{\beta}_{MQO} - \beta)]^2 = \frac{1}{T} \frac{\sum_{t=1}^T \sum_{s=1}^T \gamma_u(t-s) \gamma_x(t-s)}{\gamma_u^2(0)} .$$

Fazendo $r = \frac{\gamma_x(0)}{\gamma_u(0)}$, temos:

$$\frac{T \text{ A Var } \hat{\beta}_{MQO}}{r} = \frac{1}{T} \sum_{t=1}^T \sum_{s=1}^T \rho_u(t-s) \rho_x(t-s) \quad (4.2.3)$$

Para $\hat{\beta}_\phi$ temos:

$$\hat{\beta}_\phi - \beta = \frac{\sum_{t=2}^T \psi_t \epsilon_t}{\sum_{t=2}^T \psi_t^2} . \quad (4.2.4)$$

Assim,

$$E [\sqrt{T} (\hat{\beta}_\phi - \beta)]^2 = E \left[\frac{1/\sqrt{T} \left(\sum_{t=2}^T (x_t + \phi x_{t-1})(u_t + \phi u_{t-1}) \right)}{1/T \sum_{t=2}^T (u_t + \phi u_{t-1})^2} \right]^2 .$$

Como

$$\frac{\sum u_t^2}{T} + \frac{2\phi \sum u_t u_{t-1}}{T} + \frac{\phi^2 \sum u_{t-1}^2}{T} \xrightarrow{p} (1+\phi^2)\gamma_u(0) + 2\phi\gamma_u(1),$$

e pelo teorema de Isserlis

$$\begin{aligned} E \left[\frac{1}{\sqrt{T}} \sum_{t=2}^T \psi_t \epsilon_t \right]^2 &= \frac{1}{T} \sum_{t=2}^T \sum_{s=2}^T E(\psi_t \psi_s) E(\epsilon_t \epsilon_s) + \\ &+ \frac{1}{T} \sum_{t=2}^T \sum_{s=2}^T E(\psi_s \epsilon_s) E(\psi_s \epsilon_t) \\ &= \frac{\sigma_\epsilon^2}{T} \sum_{t=2}^T E(u_t + \phi u_{t-1})^2 \\ &= \frac{\sigma_\epsilon^2}{T} (T-1) [(1+\phi^2)\gamma_u(0) + 2\phi\gamma_u(1)], \end{aligned}$$

então,

$$\begin{aligned} \lim_{T \rightarrow \infty} E[\sqrt{T}(\hat{\beta}_{\phi} - \beta)]^2 &= \frac{\sigma_{\epsilon}^2}{(1+\phi)^2 \gamma_u(0) + 2\phi \gamma_u(1)} \\ &= \frac{(1+\phi^2) \gamma_x(0) + 2\phi \gamma_x(1)}{(1+\phi^2) \gamma_u(0) + 2\phi \gamma_u(1)} . \end{aligned}$$

Logo,

$$\frac{T \text{ A Var } \hat{\beta}_{\phi}}{r} = \frac{(1+\phi^2) + 2\phi \rho_x(1)}{1+\phi^2 + 2\phi \rho_u(1)} ,$$

e pelo fato de $\rho_x(1) = -\phi$, temos:

$$\frac{T \text{ A Var } \hat{\beta}_{\phi}}{r} = \frac{1 - \phi^2}{1+\phi^2 + 2\phi \rho_u(1)} . \quad (4.2.5)$$

4.3 Comparação dos estimadores

Seja

$$M = \frac{T \text{ A Var } \hat{\beta}_{\phi}}{r} - \frac{T \text{ A Var } \hat{\beta}_{MQO}}{r} .$$

Nas situações em que $\{u_t\}$ é ruído branco, ou um MA(1), ou um AR(1), ou um processo ARMA(1,1), veremos que $M \leq 0$, a menos de termos de $O(T^{-1})$.

4.3.1 $\{u_t\}$ é ruído branco

Neste caso $\rho_u(\ell) = 0$, para todo $\ell \neq 0$. Portanto, de (4.2.3) e (4.2.5) segue que

$$\frac{T \text{ A Var } \hat{\beta}_{MQO}}{r} = 1$$

e

$$\frac{T \text{ A Var } \hat{\beta}_{\phi}}{r} = \frac{1 - \phi^2}{1 + \phi^2}.$$

Assim,

$$M = \frac{-2\phi^2}{1 + \phi^2} \leq 0.$$

4.3.2 $\{u_t\}$ é MA(1)

A expressão (4.2.3) é equivalente a

$$\frac{T \text{ A Var } \hat{\beta}_{MQO}}{r} = \frac{1}{T} \sum_{\ell=-T+1}^{T-1} (T - |\ell|) \rho_u(\ell) \rho_x(\ell), \quad (4.3.1)$$

e como, neste caso, $\rho_u(1) \neq 0$ e $\rho_u(\ell) = 0$, $\forall \ell \geq 2$, temos:

$$\begin{aligned} \frac{T \text{ A Var } \hat{\beta}_{MQO}}{r} &= 1 + \frac{2}{T} (T-1) \rho_u(1) \rho_x(1) \\ &= 1 - 2\phi \rho_u(1) + O(T^{-1}). \end{aligned}$$

Portanto,

$$\begin{aligned}
 M &= \frac{1 - \phi^2 - [1 - 2\phi\rho_u(1)][1 + \phi^2 + 2\phi\rho_u(1)]}{1 + \phi^2 + 2\phi\rho_u(1)} + O(T^{-1}) \\
 &= \frac{2\phi^2[2\rho_u^2(1) + \phi\rho_u(1) - 1]}{1 + \phi^2 + 2\phi\rho_u(1)} + O(T^{-1}) \quad (4.3.2)
 \end{aligned}$$

Como $u_t = a_t - \theta a_{t-1}$, com $|\theta| < 1$, então $\rho_u(1) = \frac{-\theta}{1+\theta^2}$ e $|\rho_u(1)| < \frac{1}{2}$. Assim, o denominador de (4.3.2) que é

$$f(\phi) = \phi^2 + 2\phi\rho_u(1) + 1$$

é sempre positivo, pois $4\rho_u^2(1) - 4 < 0$ se $|\rho_u(1)| < \frac{1}{2}$. O numerador de (4.3.2) dado por $2\phi^2g(\phi)$, onde

$$g(\phi) = \phi\rho_u(1) + [2\rho_u^2(1) - 1],$$

é sempre negativo para $|\rho_u(1)| < \frac{1}{2}$. Suponha, por absurdo, que $g(\phi) \geq 0$. Se $0 < \rho_u(1) < \frac{1}{2}$, teremos $\phi \geq \frac{1 - 2\rho_u^2(1)}{\rho_u(1)}$, o que implica $\phi \geq 1$. Se $-\frac{1}{2} < \rho_u(1) < 0$, teremos $\phi \leq \frac{1 - 2\rho_u^2(1)}{\rho_u(1)}$, o que implica que $\phi \leq -1$.

Portanto $M \leq 0$, a menos de termos de ordem $\frac{1}{T}$.

4.3.3 $\{u_t\}$ é AR(1)

Suponha que u_t seja dado por

$$u_t + \phi_1 u_{t-1} = a_t, \quad |\phi_1| < 1,$$

onde a_t é o ruído branco $(0, \sigma_a^2)$.

De (4.3.1) segue que

$$\begin{aligned} \frac{T A \text{ Var } \hat{\beta}_{MQO}}{r} &= \frac{1}{T} [T\rho_u(0)\rho_x(0) + 2T\rho_u(1)\rho_x(1) - 2\rho_u(1)\rho_x(1) + \\ &+ 2T\rho_u(2)\rho_x(2) - 4\rho_u(2)\rho_x(2) + \dots \\ &+ 2T\rho_u(T-1)\rho_x(T-1) - 2(T-1)\rho_u(T-1)\rho_x(T-1) \\ &= \rho_u(0)\rho_x(0) + 2\rho_u(1)\rho_x(1) + \dots + 2\rho_u(T-1)\rho_x(T-1) + O(T^{-1})]. \end{aligned}$$

Mas, $\rho_u(j) = (-\phi_1)^j$ e $\rho_x(j) = (-\phi)^j$ e portanto

$$\frac{T A \text{ Var } \hat{\beta}_{MQO}}{r} = 1 + \frac{2\phi\phi_1}{1-\phi\phi_1} + O(T^{-1}) = \frac{1+\phi\phi_1}{1-\phi\phi_1} + O(T^{-1}).$$

De (4.2.5) segue que:

$$\frac{T A \text{ Var } \hat{\beta}_\phi}{r} = \frac{1-\phi^2}{1+\phi^2-2\phi\phi_1}.$$

Portanto,

$$\begin{aligned}
 M &= \frac{(1-\phi^2)(1-\phi\phi_1) - (1+\phi\phi_1)(1+\phi^2-2\phi\phi_1)}{(1+\phi^2-2\phi\phi_1)(1-\phi\phi_1)} + O(T^{-1}) \\
 &= \frac{-2\phi^2(1-\phi_1^2)}{(1+\phi^2-2\phi\phi_1)(1-\phi\phi_1)} + O(T^{-1}) \quad (4.3.3)
 \end{aligned}$$

O denominador de (4.3.3) é sempre positivo, pois

$$\phi^2 - 2\phi\phi_1 + 1 > 0, \quad \text{se } |\phi_1| < 1$$

e

$$1 - \phi\phi_1 > 0, \quad \text{se } |\phi\phi_1| < 1.$$

Como $(1-\phi_1^2) > 0$ se $|\phi_1| < 1$, então o numerador de (4.3.3) é sempre ≤ 0 .

Portanto $M \leq 0$, a menos de termos de $O(T^{-1})$.

4.3.4 $\{u_t\}$ é um ARMA(1,1)

Suponha que u_t seja dado por

$$u_t + \phi_1 u_{t-1} = a_t - \theta_1 a_{t-1},$$

com $|\phi_1| < 1$ e $|\theta_1| < 1$.

De (4.3.1) segue que:

$$\frac{T \text{ A Var } \hat{\beta}_{MQO}}{r} = \frac{1}{T} \sum_{\ell=-T+1}^{T-1} (T-|\ell|)(-\phi)^\ell \rho_u(\ell),$$

e como $\rho_u(\ell) = -\phi_1 \rho_u(\ell-1)$, $\ell > 1$, temos:

$$\begin{aligned} \frac{T \text{ A Var } \hat{\beta}_{MQO}}{r} &= 1 - 2\phi \rho_u(1) - 2\phi^2 \phi_1 \rho_u(1) - \dots - 2\phi^{T-1} \phi_1^{T-2} \rho_u(1) \\ &\quad + O(T^{-1}) \\ &= \frac{1-\phi\phi_1-2\phi\rho_u(1)}{1-\phi\phi_1} + O(T^{-1}). \end{aligned}$$

Portanto,

$$\begin{aligned} M &= \frac{(1-\phi^2)(1-\phi\phi_1) - [1-\phi\phi_1-2\phi\rho_u(1)][1+\phi^2+2\phi\rho_u(1)]}{(1-\phi\phi_1)[1+\phi^2+2\phi\rho_u(1)]} + O(T^{-1}) \\ &= \frac{-2\phi^2+2\phi^3\phi_1+2\phi^3\rho_u(1)+2\phi^2\phi_1\rho_u(1)+4\phi^2\rho_u^2(1)}{(1-\phi\phi_1)[1+\phi^2+2\phi\rho_u(1)]} + O(T^{-1}) \end{aligned} \quad (4.3.4)$$

O denominador de (4.3.4) é sempre positivo pois $|\phi_1| < 1$ e $|\phi| < 1$. O numerador é:

$$-2\phi^2[1 - \phi\phi_1 - \phi\rho_u(1) - \phi_1\rho_u(1) - 2\rho_u^2(1)]. \quad (4.3.5)$$

Suponha, por absurdo, que o termo entre colchetes de (4.3.5) seja negativo. Teremos:

$$-\phi[\phi_1 + \rho_u(1)] < -1 + \phi_1 \rho_u(1) + 2\rho_u^2(1) \quad . \quad (4.3.6)$$

Para (4.3.6) podemos distinguir dois casos:

$$(I) \quad \phi_1 + \rho_u(1) > 0 \quad \text{e} \quad (II) \quad \phi_1 + \rho_u(1) < 0.$$

No caso (I), como $|\phi| < 1$, de (4.3.6) segue que:

$$1 > \phi > \frac{1 - \phi_1 \rho_u(1) - 2\rho_u^2(1)}{\phi_1 + \rho_u(1)} \quad ,$$

o que implica que:

$$-2\rho_u^2(1) - \rho_u(1)(1 + \phi_1) + (1 - \phi_1) < 0 \quad .$$

Assim, se

$$h_1(\rho_u(1)) = -2\rho_u^2(1) - \rho_u(1)(1 + \phi_1) + (1 - \phi_1),$$

as raízes de h_1 são $\frac{1 - \phi_1}{2}$ e -1 , o que implica que $\rho_u(1) < -1$ ou $\rho_u(1) > \frac{1 - \phi_1}{2}$. Claramente $\rho_u(1) < -1$ é absurdo e

$$\rho_u(1) = \frac{-(1 + \phi_1 \theta_1)(\phi_1 + \theta_1)}{1 + \theta_1^2 + 2\phi_1 \theta_1} > \frac{1 - \phi_1}{2} \quad ,$$

implica que

$$-\theta_1^2(1+\phi_1) - 2\theta_1(1+\phi_1) - (1+\phi_1) > 0.$$

Como $1 + \phi_1 > 0$, temos:

$$\theta_1^2 + 2\theta_1 + 1 < 0$$

o que também é absurdo.

No caso (II) chegamos, analogamente, a um absurdo.

Conclui-se então que $M \leq 0$, a menos de termos de $O(T^{-1})$.

CAPÍTULO 5

EXPANSÃO ASSINTÓTICA DO ERRO QUADRÁTICO MÉDIO DO ESTIMADOR DE COCHRANE-ORCUTT EM DOIS ESTÁGIOS

5.1 Introdução

Consideraremos agora a situação em que no modelo

$$y_t = \beta u_t + x_t \quad t = 1, 2, \dots \quad (5.1.1)$$

$\{x_t\}$ é gerado por um AR(p), isto é:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t, \quad (5.1.2)$$

onde $\{\varepsilon_t\}$ é uma seqüência de ruído branco, de média zero e variância σ_ε^2 .

Sendo $\hat{\phi} = (\hat{\phi}_1, \dots, \hat{\phi}_p)'$ um estimador consistente de $\phi = (\phi_1, \dots, \phi_p)'$, o estimador de CO2 é definido por:

$$\hat{\beta}_{\hat{\phi}} = \frac{\sum_{t=p+1}^T \hat{\eta}_t \hat{\psi}_t}{\sum_{t=p+1}^T \hat{\psi}_t^2}, \quad (5.1.3)$$

onde

$$\hat{\eta}_t = y_t - \hat{\phi}_1 y_{t-1} - \dots - \hat{\phi}_p y_{t-p}, \quad (5.1.4)$$

e

$$\hat{\psi}_t = u_t - \hat{\phi}_1 u_{t-1} - \dots - \hat{\phi}_p u_{t-p} \quad (5.1.5)$$

Sob as condições usuais impostas a $\{u_t\}$, os estimadores de MQG e CO2 têm a mesma distribuição assintótica. Além disso, o estimador de máxima verossimilhança de β (fixados $y_1, y_2, \dots, y_p$) é idêntico ao estimador CO2 (Harvey, 1981, cap. 6).

Nosso objetivo neste capítulo, é derivar a expansão de 2ª ordem do erro quadrático médio de $\hat{\beta}_{\hat{\phi}}$. Como foi ressaltado por Toyooka (1982), isto pode ser útil na construção de intervalos de confiança mais acurados, baseados no termo de 2ª ordem.

Seja

$$\hat{\epsilon}_t = x_t - \hat{\phi}_1 x_{t-1} - \dots - \hat{\phi}_p x_{t-p} \quad (5.1.6)$$

e sejam $\hat{\psi} = (\hat{\psi}_{p+1}, \dots, \hat{\psi}_T)'$, $\hat{\eta} = (\hat{\eta}_{p+1}, \dots, \hat{\eta}_T)'$, $\hat{\epsilon} = (\hat{\epsilon}_{p+1}, \dots, \hat{\epsilon}_T)'$. Considere as seqüências η_t , ψ_t e ϵ_t obtidas de (5.1.4), (5.1.5) e (5.1.6), respectivamente, quando se substitui $\hat{\phi}$ por ϕ , e sejam η , ψ e ϵ os vetores correspondentes. Seja

$$\hat{\beta}_{\phi} = \frac{\sum_{t=p+1}^T \eta_t \psi_t}{\sum_{t=p+1}^T \psi_t^2} \quad (5.1.7)$$

Note que $\hat{\beta}_{\hat{\phi}}$ e $\hat{\beta}_{\phi}$ podem ser reescritos como

$$\hat{\beta}_{\hat{\phi}} = (\hat{\psi}' \hat{\psi})^{-1} \hat{\psi}' \hat{\eta} \quad (5.1.8)$$

e

$$\hat{\beta}_{\phi} = (\psi' \psi)^{-1} \psi' \eta \quad (5.1.9)$$

respectivamente, e assim:

$$\hat{\beta}_{\hat{\phi}} - \beta = (\hat{\psi}'\hat{\psi})^{-1} \hat{\psi}'\hat{\varepsilon} \quad (5.1.10)$$

e

$$\hat{\beta}_{\phi} - \beta = (\psi'\psi)^{-1} \psi'\varepsilon \quad (5.1.11)$$

5.2 Suposições e Notações

No que segue faremos as seguintes suposições:

Condição 1. O estimador $\hat{\phi}$ é independente de $\underline{y} = (y_1, \dots, y_T)'$.

Condição 2. O $\lim_{T \rightarrow \infty} \frac{\sum_{t=p+1}^T u_{t-i} u_{t-k}}{T} = \gamma_u(i-k)$ existe, para todo i, k .

Condição 3. O $\lim_{T \rightarrow \infty} T \cdot E\{(\hat{\phi}_i - \phi_i)(\hat{\phi}_k - \phi_k)\} = v_{ik}$, existe para todo $i, k = 1, \dots, p$.

Sob as condições acima, temos as seguintes consequências:

$$\begin{aligned} (I) \quad \lim_{T \rightarrow \infty} T^{-1} \psi' \psi &= \gamma_u(0) \sum_{i=0}^p \phi_i^2 - 2 \sum_{i=1}^p \phi_i \gamma_u(i) + 2\gamma_u(1) \sum_{i=1}^{p-1} \phi_i \phi_{i+1} + \\ &+ \dots + 2\gamma_u(j) \sum_{i=1}^{p-j} \phi_i \phi_{i+j} + \dots + 2\gamma_u(p-1) \phi_i \phi_p = \\ &= \omega(\phi), \end{aligned}$$

com $\phi_0 = 1$.

Notamos que:

$$\begin{aligned} \lim_{T \rightarrow \infty} [T^{-1} \frac{\partial}{\partial \phi_k} \psi' \psi] &= \lim_{T \rightarrow \infty} T^{-1} [-2 \sum_{t=p+1}^T u_{t-k} (u_t - \phi_1 u_{t-1} - \dots - \phi_p u_{t-p})] \\ &= \frac{\partial \omega(\underline{\phi})}{\partial \phi_k} \end{aligned} \quad (5.2.1)$$

e

$$\begin{aligned} \lim_{T \rightarrow \infty} T^{-1} [\frac{\partial^2 \psi' \psi}{\partial \phi_j \partial \phi_k}] &= \lim_{T \rightarrow \infty} T^{-1} [2 \sum_{t=p+1}^T u_{t-j} u_{t-k}] = 2 \gamma_u(k-j) \\ &= \frac{\partial^2 \omega(\underline{\phi})}{\partial \phi_j \partial \phi_k} . \end{aligned} \quad (5.2.2)$$

$$(II) \quad E[(\psi'_{\varepsilon})^2] = \sigma_{\varepsilon}^2 \psi' \psi, \quad \text{e portanto}$$

$$\lim_{T \rightarrow \infty} T^{-1} E[(\psi'_{\varepsilon})^2] = \sigma_{\varepsilon}^2 \omega(\underline{\phi}) .$$

No que segue usaremos as seguintes notações:

$$(III) \quad \lim_{T \rightarrow \infty} T^{-1} . E \{ [\frac{\partial}{\partial \phi_i} (\psi'_{\varepsilon})]^2 \} = G_i(\underline{\phi}), \quad i = 1, \dots, p. \quad (5.2.3)$$

$$(IV) \quad \lim_{T \rightarrow \infty} T^{-1} . E \{ [\frac{\partial (\psi'_{\varepsilon})}{\partial \phi_i} \cdot \frac{\partial (\psi'_{\varepsilon})}{\partial \phi_k}] \} = Z_{ik}(\underline{\phi}), \quad i, k = 1, \dots, p. \quad (5.2.4)$$

$$(V) \quad \lim_{T \rightarrow \infty} T^{-1} . E [\frac{\partial (\psi'_{\varepsilon})}{\partial \phi_i} (\varepsilon \psi')'] = P_i(\underline{\phi}), \quad i = 1, \dots, p. \quad (5.2.5)$$

$$(VI) \quad \lim_{T \rightarrow \infty} T^{-1} . E [\frac{\partial^2 (\psi'_{\varepsilon})}{\partial \phi_i \partial \phi_k} (\varepsilon \psi')'] = Q_{ik}(\underline{\phi}), \quad i, k = 1, \dots, p. \quad (5.2.6)$$

5.3 A expansão de 2ª ordem do erro quadrático médio de $\hat{\beta}_{\hat{\phi}}$

Teorema 3.1. Sob as condições 1-3, quando $T \rightarrow \infty$, temos:

$$\begin{aligned} E\{(\hat{\beta}_{\hat{\phi}} - \beta)^2\} &= T^{-1} \omega^{-1}(\underline{\phi}) \sigma_{\epsilon}^2 + T^{-2} \left\{ \sum_{i=1}^p \sum_{k=1}^p v_{ik} \omega^{-2}(\underline{\phi}) [Z_{ik}(\underline{\phi}) + \right. \\ &\quad + Q_{ik}(\underline{\phi}) - 2\sigma_{\epsilon}^2 \gamma_{ik} + 3\omega^{-1}(\underline{\phi}) \sigma_{\epsilon}^2 \frac{\partial \omega(\underline{\phi})}{\partial \phi_i} \cdot \frac{\partial \omega(\underline{\phi})}{\partial \phi_k} - \\ &\quad \left. - 4\omega^{-1}(\underline{\phi}) \sigma_{\epsilon}^2 \frac{\partial \omega(\underline{\phi})}{\partial \phi_i} P_k(\underline{\phi})] \right\} + o(T^{-2}) . \end{aligned}$$

Prova. Temos:

$$\begin{aligned} E\{(\hat{\beta}_{\hat{\phi}} - \beta)^2\} &= E\{(\hat{\beta}_{\hat{\phi}} - \hat{\beta}_{\phi})^2\} + E\{(\hat{\beta}_{\phi} - \beta)^2\} + \\ &\quad + 2 E\{(\hat{\beta}_{\hat{\phi}} - \hat{\beta}_{\phi})(\hat{\beta}_{\phi} - \beta)\} , \end{aligned}$$

e

$$\begin{aligned} \hat{\beta}_{\hat{\phi}} - \beta_{\phi} &= \{[\hat{\psi}'\hat{\psi}]^{-1} - [\psi'\psi]^{-1}\} \hat{\psi}'\hat{\epsilon} + [\hat{\psi}'\hat{\psi}]^{-1} [\hat{\psi}'\hat{\epsilon} - \psi'\epsilon] \\ &= A + B . \end{aligned}$$

Expandindo A em série de Taylor obtemos:

$$\sum_{k=1}^p (\hat{\phi}_k - \phi_k) \frac{\partial}{\partial \phi_k} (\psi'\psi)^{-1} \psi'\epsilon + \frac{1}{2} \sum_i \sum_k (\hat{\phi}_i - \phi_i)(\hat{\phi}_k - \phi_k) \frac{\partial^2 (\psi'\psi)^{-1}}{\partial \phi_i \partial \phi_k} \psi'\epsilon +$$

+ resto.

(5.3.1)

Expandindo B temos:

$$\begin{aligned} & \sum_{k=1}^P (\hat{\phi}_k - \phi_k) (\psi' \psi)^{-1} \frac{\partial}{\partial \phi_k} (\psi' \underline{\varepsilon}) + \sum_k \sum_i (\hat{\phi}_i - \phi_i) (\hat{\phi}_k - \phi_k) \frac{\partial (\psi' \psi)^{-1}}{\partial \phi_i} \frac{\partial (\psi' \underline{\varepsilon})}{\partial \phi_k} + \\ & + \frac{1}{2} \sum_{i=1}^P \sum_{k=1}^P (\hat{\phi}_i - \phi_i) (\hat{\phi}_k - \phi_k) \frac{\partial^2 (\psi' \underline{\varepsilon})}{\partial \phi_i \partial \phi_k} (\psi' \psi)^{-1} + \text{resto.} \end{aligned} \quad (5.3.2)$$

De (5.2.1) e (5.3.1), obtemos:

$$\lim_{T \rightarrow \infty} E(A^2) = T^{-2} \omega^{-3}(\underline{\psi}) \sigma_\varepsilon^2 \sum_i \sum_k v_{ik} \frac{\partial \omega(\underline{\psi})}{\partial \phi_i} \frac{\partial \omega(\underline{\psi})}{\partial \phi_k}. \quad (5.3.3)$$

De (5.3.2), obtemos:

$$\lim_{T \rightarrow \infty} E(B^2) = T^{-2} \omega^{-2}(\underline{\psi}) \sum_i \sum_k v_{ik} Z_{ik}(\underline{\psi}). \quad (5.3.4)$$

e de (5.3.1) e (5.3.2)

$$\lim_{T \rightarrow \infty} E(AB) = -T^{-2} \omega^{-3}(\underline{\psi}) \sum_i \sum_k v_{ik} \frac{\partial \omega(\underline{\psi})}{\partial \phi_i} P_k(\underline{\psi}). \quad (5.3.5)$$

De (5.3.3), (5.3.4) e (5.3.5) segue que:

$$\begin{aligned} \lim_{T \rightarrow \infty} E\{(\hat{\beta}_{\hat{\phi}} - \beta_{\phi})^2\} &= T^{-2} \omega^{-2}(\underline{\psi}) \sum_i \sum_k v_{ik} Z_{ik}(\underline{\psi}) + \\ &+ T^{-2} \omega^{-3}(\underline{\psi}) \sigma_\varepsilon^2 \sum_i \sum_k v_{ik} \frac{\partial \omega(\underline{\psi})}{\partial \phi_i} \frac{\partial \omega(\underline{\psi})}{\partial \phi_k} \end{aligned}$$

$$- 2T^{-2} \omega^{-3}(\underline{\phi}) \sum_i \sum_k v_{ik} \frac{\partial \omega(\underline{\phi})}{\partial \phi_i} P_k(\underline{\phi}) . \quad (5.3.6)$$

De (5.1.11) temos:

$$\lim_{T \rightarrow \infty} E(\hat{\beta}_{\phi} - \beta)^2 = T^{-1} \omega^{-1}(\underline{\phi}) \sigma_{\epsilon}^2 . \quad (5.3.7)$$

Se denotarmos $\hat{\beta}_{\phi} - \beta = C$, então:

$$E\{(\hat{\beta}_{\phi} - \beta)(\hat{\beta}_{\phi} - \beta)\} = E(AC) + E(BC) .$$

De (5.1.11), (5.3.1), (5.2.1) e (5.2.2), temos:

$$\begin{aligned} \lim_{T \rightarrow \infty} E(AC) = & -T^{-2} \omega^{-2}(\underline{\phi}) \sigma_{\epsilon}^2 \sum_i \sum_k v_{ik} \gamma_u^{(k-i)} \\ & + T^{-2} \omega^{-3}(\underline{\phi}) \sigma_{\epsilon}^2 \sum_i \sum_k v_{ik} \frac{\partial \omega(\underline{\phi})}{\partial \phi_i} \frac{\partial \omega(\underline{\phi})}{\partial \phi_k} . \end{aligned} \quad (5.3.8)$$

Também,

$$\begin{aligned} \lim_{T \rightarrow \infty} E(BC) = & -T^{-2} \omega^{-3}(\underline{\phi}) \sum_i \sum_k v_{ik} \frac{\partial \omega(\underline{\phi})}{\partial \phi_i} P_k(\underline{\phi}) \\ & + \frac{1}{2} T^{-2} \omega^{-2}(\underline{\phi}) \sum_i \sum_k v_{ik} Q_{ik}(\underline{\phi}) . \end{aligned} \quad (5.3.9)$$

De (5.3.8) e (5.3.9), temos:

$$\begin{aligned}
 \lim_{T \rightarrow \infty} 2E[(\hat{\beta}_{\phi} - \beta_{\phi})(\hat{\beta}_{\phi} - \beta_{\phi})] &= T^{-2} \omega^{-2}(\phi) \sum_i \sum_k v_{ik} Q_{ik}(\phi) \\
 &- 2T^{-2} \omega^{-3}(\phi) \sum_i \sum_k v_{ik} \frac{\partial \omega(\phi)}{\partial \phi_i} P_k(\phi) - 2T^{-2} \omega^{-2}(\phi) \sigma_{\epsilon}^2 \sum_i \sum_k v_{ik} \gamma_u^{(k-i)} + \\
 &+ 2T^{-2} \omega^{-3}(\phi) \sum_i \sum_k v_{ik} \frac{\partial \omega(\phi)}{\partial \phi_i} \frac{\partial \omega(\phi)}{\partial \phi_k} .
 \end{aligned} \tag{5.3.10}$$

De (5.3.6), (5.3.7) e (5.3.10), obtemos o resultado desejado. Δ

5.4 O caso vetorial

Seja

$$y_t = \underline{u}_t' \underline{\beta} + x_t, \quad t = 1, \dots, T, \tag{5.4.1}$$

onde $\underline{\beta} = (\beta_1, \dots, \beta_q)'$ é um vetor de parâmetros desconhecidos, $\underline{u}_t = (u_{t,1}, \dots, u_{t,q})'$ é um vetor de constantes conhecidas e x_t é dado por (5.1.2).

Podemos escrever (5.4.1) como:

$$\underline{y} = \underline{U} \underline{\beta} + \underline{x}, \tag{5.4.2}$$

onde $\underline{y} = (y_1, \dots, y_T)'$, $\underline{x} = (x_1, \dots, x_T)'$ e $\underline{U}$ é a matriz $T \times q$

$$\underline{U} = \begin{bmatrix} u_{1,1} & \dots & u_{1,q} \\ \vdots & & \vdots \\ u_{T,1} & \dots & u_{T,q} \end{bmatrix}$$

Seja $\hat{\phi} = (\hat{\phi}_1, \dots, \hat{\phi}_p)'$ um estimador consistente de $\phi = (\phi_1, \dots, \phi_p)'$.

Faremos as seguintes suposições:

- (1) $\hat{\phi}$ é independente das observações.
- (2) O posto da matriz $\underline{U}'\underline{U}$, $\rho(\underline{U}'\underline{U})$, é q , $T \geq p+q$.
- (3) O limite

$$\lim_{T \rightarrow \infty} \frac{\sum u_{t-i,r} u_{t-k,s}}{T} = \gamma_{r,s}(i-k)$$

existe, para todo i, k e $r, s = 1, \dots, q$.

- (4) O limite

$$\lim_{T \rightarrow \infty} T \cdot E\{(\hat{\phi}_i - \phi_i)(\hat{\phi}_k - \phi_k)\} = v_{ik}$$

existe, para todo $i, k = 1, \dots, p$.

Vamos formar as seqüências:

$$\eta_t = y_t - \sum_{i=1}^p \phi_i y_{t-i}, \quad \underline{\eta} = (\eta_{p+1}, \dots, \eta_T)',$$

$$\epsilon_t = x_t - \sum_{i=1}^p \phi_i x_{t-i}, \quad \underline{\epsilon} = (\epsilon_{p+1}, \dots, \epsilon_T)'$$

$$\underline{\psi}_t' = \underline{u}_t' - \sum_{i=1}^p \phi_i \underline{u}_{t-i}',$$

$$\underline{\psi} = \begin{bmatrix} \underline{\psi}_{p+1}' \\ \vdots \\ \underline{\psi}_T' \end{bmatrix} = \begin{bmatrix} \psi_{1,p+1} & \dots & \psi_{q,p+1} \\ \vdots & & \vdots \\ \psi_{1,T} & \dots & \psi_{q,T} \end{bmatrix},$$

$$\text{com } \psi_{r,t} = u_{t,r} - \sum_{i=1}^P \phi_i u_{t-i,r} .$$

Sejam $\hat{\eta}_t$, $\hat{\varepsilon}_t$ e $\hat{\psi}'_t$ as seqüências correspondentes, com os ϕ_i substituídos por $\hat{\phi}_i$.

Supondo $\rho(\underline{\psi}'\underline{\psi}) = \rho(\hat{\underline{\psi}}'\hat{\underline{\psi}}) = q$, para $T-p \geq q$, sejam os estimadores:

$$\hat{\beta}_{\hat{\phi}} = (\hat{\underline{\psi}}'\hat{\underline{\psi}})^{-1} \hat{\underline{\psi}}'\hat{\underline{\eta}} \quad (5.4.3)$$

e

$$\hat{\beta}_{\phi} = (\underline{\psi}'\underline{\psi})^{-1} \underline{\psi}'\underline{\eta} . \quad (5.4.4)$$

Também,

$$\hat{\beta}_{\hat{\phi}} - \beta = (\hat{\underline{\psi}}'\hat{\underline{\psi}})^{-1} \hat{\underline{\psi}}'\hat{\underline{\varepsilon}} , \quad (5.4.5)$$

$$\hat{\beta}_{\phi} - \beta = (\underline{\psi}'\underline{\psi})^{-1} \underline{\psi}'\underline{\varepsilon} . \quad (5.4.6)$$

Usaremos as seguintes notações:

$$(I) \quad \lim_{T \rightarrow \infty} T^{-1} \underline{\psi}'\underline{\psi} = W(\phi) ,$$

onde $W(\phi)$ é matriz, $q \times q$, cujo (r,s) -ésimo elemento é $\omega_{rs}(\phi)$

$$(II) \quad E[(\underline{\psi}'\underline{\varepsilon})^2] = \underline{\psi}'E(\underline{\varepsilon}\underline{\varepsilon}')\underline{\psi} = \sigma_{\varepsilon}^2 \underline{\psi}'\underline{\psi}, \text{ e portanto}$$

$$\lim_{T \rightarrow \infty} T^{-1} E[(\underline{\psi}'\underline{\varepsilon})^2] = \sigma_{\varepsilon}^2 W(\phi) .$$

$$(III) \quad \lim_{T \rightarrow \infty} T^{-1} E\left\{\left[\frac{\partial}{\partial \phi_i} (\underline{\psi}'\underline{\varepsilon})\right]^2\right\} = G_i(\phi), \text{ que é matriz } q \times q.$$

$$(IV) \quad \lim_{T \rightarrow \infty} T^{-1} E\left[\frac{\partial(\underline{\psi}'\underline{\varepsilon})}{\partial \phi_i}, \frac{\partial(\underline{\psi}'\underline{\varepsilon})}{\partial \phi_k}\right] = Z_{ik}(\phi), \text{ que é matriz } q \times q.$$

$$(V) \quad \lim_{T \rightarrow \infty} T^{-1} E \left[\frac{\partial(\psi' \underline{\varepsilon})}{\partial \phi_i} (\underline{\varepsilon}' \psi)' \right] = P_i(\phi), \text{ que é matriz } q \times q.$$

$$(VI) \quad \lim_{T \rightarrow \infty} T^{-1} E \left[\frac{\partial^2(\psi' \underline{\varepsilon})}{\partial \phi_i \partial \phi_k} (\underline{\varepsilon}' \psi)' \right] = Q_{ik}(\phi), \text{ que é matriz } q \times q.$$

Temos então o seguinte resultado:

Teorema 5.4.1. Sob as suposições (1)-(4), temos:

$$\begin{aligned} E[(\hat{\beta}_{\hat{\phi}} - \beta)(\hat{\beta}_{\hat{\phi}} - \beta)'] &= T^{-1} \tilde{W}^{-1}(\phi) \sigma_{\varepsilon}^2 + T^{-2} \left\{ \sum_{i=1}^P \sum_{k=1}^P v_{ik} \tilde{W}^{-2}(\phi) [Z_{ik}(\phi) \right. \\ &\quad + Q_{ik}(\phi) - \sigma_{\varepsilon}^2 \left(\frac{\partial^2}{\partial \phi_i \partial \phi_k} W(\phi) \right) + 3\tilde{W}^{-1}(\phi) \sigma_{\varepsilon}^2 \frac{W(\phi)}{\partial \phi_i} \frac{W(\phi)}{\partial \phi_k} \\ &\quad \left. - 4\tilde{W}^{-1}(\phi) \frac{\partial W(\phi)}{\partial \phi_i} P_k(\phi) \right\} + o(T^{-2}) . \end{aligned}$$

5.5 Casos especiais

Considere o modelo (5.1.1) com $\{x_t\}$ dado por (5.1.2) para $p=1$.

Nesse caso temos:

$$(I') \quad \lim_{T \rightarrow \infty} T^{-1} \psi' \psi = (1 + \phi^2) \gamma_u(0) - 2\phi \gamma_u(1) = \omega(\phi) .$$

$$(II') \quad \lim_{T \rightarrow \infty} T^{-1} E(\psi' \underline{\varepsilon})^2 = \sigma_{\varepsilon}^2 \omega(\phi) .$$

$$\begin{aligned} (III') \quad E\left\{\left[\frac{d}{d\phi} (\psi' \underline{\varepsilon})\right]^2\right\} &= \sigma_{\varepsilon}^2 \underline{u}' \underline{u} + \psi' \underline{V} \psi + 2\phi^2 \underline{u}' \underline{V} \underline{u} \\ &\quad - 2\phi \underline{u}' \underline{V} \underline{u} - 2\phi \underline{u}' \underline{M} \underline{u}_+ + 2\underline{u}' \underline{M} \underline{u}_+ , \end{aligned}$$

onde $\underline{u} = (u_1, \dots, u_{T-1})'$, $\underline{u}_+ = (u_2, \dots, u_T)'$, $\underline{V} = E(\underline{x} \underline{x}')$,
 $\underline{x} = (x_1, \dots, x_{T-1})'$, $\underline{M} = E(\underline{x}_+ \underline{x}_+')$, com $\underline{x}_+ = (x_2, \dots, x_T)'$.

Seja

$$\lim_{T \rightarrow \infty} T^{-1} E\left\{\left[\frac{d}{d\phi} (\underline{\psi}' \underline{\varepsilon})^2\right]\right\} = G(\phi) \quad .$$

$$(V') \quad E\left\{\left[\frac{d}{d\phi} (\underline{\psi}' \underline{\varepsilon})\right] \underline{\varepsilon}' \underline{\psi}\right\} = -\sigma_\varepsilon^2 \underline{u}' \underline{u}_+ + \phi \sigma_\varepsilon^2 \underline{u}' \underline{u} - \underline{\psi}' \underline{M} \underline{\psi} + \phi \underline{\psi}' \underline{V} \underline{\psi} \quad .$$

Seja

$$\lim_{T \rightarrow \infty} T^{-1} E\left\{\left[\frac{d}{d\phi} (\underline{\psi}' \underline{\varepsilon})\right] \underline{\varepsilon} \underline{\psi}'\right\} = P(\phi)$$

$$(VI') \quad E\left\{\left[\frac{d^2}{d\phi^2} (\underline{\psi}' \underline{\varepsilon})\right] \underline{\varepsilon}' \underline{\psi}\right\} = 2\phi^2 \underline{u}' \underline{V} \underline{u} - 2\phi \underline{u}' \underline{V} \underline{u}_+ - 2\phi \underline{u}' \underline{M}' \underline{u} + \\ + 2\underline{u}' \underline{M}' \underline{u}_+ \quad .$$

Seja

$$\lim_{T \rightarrow \infty} T^{-1} E\left\{\left[\frac{d^2}{d\phi^2} (\underline{\psi}' \underline{\varepsilon})\right] (\underline{\psi}' \underline{\varepsilon})'\right\} = Q(\phi) \quad .$$

Seja $\hat{\phi}$ o estimador de máxima verossimilhança de ϕ . Assu-
 mindo que x_t é observável, então:

$$\sqrt{T}(\hat{\phi} - \phi) \rightarrow N(0, \sigma_\varepsilon^2 / \gamma_X(0)) \quad (\text{Mann \& Wald, 1943})$$

onde $\sigma_X^2 = \gamma_X(0) = \sigma_\varepsilon^2 / 1 - \phi^2$.

Assim, sob as condições 1 e 2, temos que:

$$E(\hat{\beta}_{\phi} - \beta)^2 = T^{-1} \omega(\phi)^{-1} \sigma_{\epsilon}^2 + \frac{T^{-2} \sigma_{\epsilon}^2}{\gamma_X(0) \omega^2(\phi)} \left\{ \frac{3\sigma_{\epsilon}^2}{\omega(\phi)} \left[\frac{d}{d\phi} \omega(\phi) \right]^2 \right. \\ \left. - \frac{4}{\omega(\phi)} P(\phi) \left[\frac{d}{d\phi} \omega(\phi) \right] + Q(\phi) + G(\phi) - \sigma_{\epsilon}^2 \frac{d^2}{d\phi^2} \omega(\phi) \right\} + o(T^{-2}) \quad (5.5.1)$$

Suponha agora a situação de um modelo de posição com $\{x_t\}$ gerado por um AR(1). Neste caso $\underline{u} = (1, \dots, 1)'$. É fácil verificar que:

$$(i) \quad \omega(\phi) = (1-\phi)^2, \quad \psi = (1-\phi, \dots, 1-\phi)';$$

$$(ii) \quad \lim_{T \rightarrow \infty} T^{-1} \underline{\psi}' \underline{V} \underline{\psi} = \sigma_{\epsilon}^2;$$

$$(iii) \quad \lim_{T \rightarrow \infty} T^{-1} 2\phi^2 \underline{u}' \underline{V} \underline{u} = 2\phi^2 \sigma_{\epsilon}^2 / (1-\phi)^2;$$

$$(iv) \quad \lim_{T \rightarrow \infty} T^{-1} 2\phi \underline{u}' \underline{V} \underline{u}_+ = 2\phi \sigma_{\epsilon}^2 / (1-\phi)^2;$$

$$(v) \quad \lim_{T \rightarrow \infty} T^{-1} 2 \underline{u}' \underline{M} \underline{u}_+ = 2\sigma_{\epsilon}^2 / (1-\phi)^2.$$

Assim,

$G(\phi) = 4\sigma_{\epsilon}^2$, $Q(\phi) = 2\sigma_{\epsilon}^2$, $P(\phi) = 2\sigma_{\epsilon}^2(\phi-1)$, e usando (5.5.1) obtemos:

$$E(\hat{\beta}_{\phi} - \beta)^2 = T^{-1} \sigma_{\epsilon}^2 / (1-\phi)^2 + o(T^{-2}) \quad (5.5.2)$$

Toyooka (1982) fez a expansão de 2ª ordem do erro quadrático médio do estimador $\hat{\beta}_{MQG}$, sob determinadas condições a respeito de $\{u_t\}$. O resultado obtido por Toyooka, para o modelo de posição com $\{x_t\}$ seguindo um AR(1), também é dado por (5.5.2).

CAPÍTULO 6

SIMULAÇÕES

6.1 Introdução

Apresentaremos aqui os resultados das simulações aplicadas aos estimadores propostos. O objetivo será comparar os seus desempenhos.

No caso dos estimadores de Cochrane-Orcutt em dois estágios, aos resíduos de mínimos quadrados ordinários iremos ajustar uma auto-regressão de ordem h , que pode ser fixada (estimador CO2) ou obtida por determinados critérios (estimador COG2). Especialmente nesta última situação será útil estimar os coeficientes da auto-regressão de forma recursiva.

Falaremos inicialmente da estimação recursiva dos coeficientes e dos critérios para a determinação de h . Apresentaremos a seguir as situações onde as simulações foram empregadas, e finalmente os resultados e comentários a respeito dos mesmos.

6.2 Estimação recursiva dos coeficientes e critérios para determinação de h .

Como vimos nas seções 3.2 e 3.3 capítulo 3, ajustamos aos resíduos

$$\hat{x}_t = y_t - \hat{\beta}_{MQO} u_t, \quad (6.2.1)$$

uma auto-regressão de ordem $h \leq H(T)$, onde $H(T)$ cresce com T . Os coeficientes da auto-regressão ajustada a $\hat{x}_t$ são obtidos através das equações de Yule-Walker dadas por:

$$\hat{\tilde{a}}_h \hat{\tilde{r}}(h) = - [\hat{c}(1), \dots, \hat{c}(h)]' \quad (6.2.2)$$

onde:

$$\hat{\tilde{a}}_h = [\hat{a}_h(1), \dots, \hat{a}_h(h)]'$$

e

$$\left\{ \begin{array}{l} \hat{c}(t) = \frac{\sum_{s=1}^{T-t} \{[\hat{x}(s) - \hat{x}(.)] [\hat{x}(s+t) - \hat{x}(.)]\}}{T}, \quad t = 0, 1, \dots, T-1 \\ \hat{c}(t) = 0, \quad t \geq T \end{array} \right.$$

com

$$\hat{x}(.) = \frac{\sum_{t=1}^T \hat{x}(t)}{T},$$

e onde $\hat{\tilde{r}}(h)$ é a matriz, $(h \times h)$, cujo (j, k) -ésimo elemento é $\hat{c}(j-k)$, $j, k = 1, \dots, h$.

A equação (6.2.2) pode ser resolvida recursivamente. Fazendo:

$$\Delta_h = \hat{c}(h) - \sum_{j=1}^{h-1} \hat{a}_{h-1}(j) \hat{c}(h-j)$$

$$P_h = \hat{c}(0) - \sum_{j=1}^{h-1} \hat{a}_{h-1}(j) c(j) ,$$

$$\hat{a}_0(j) = 0 , \quad j = 1, 2, \dots, h ,$$

através da recursão de Durbin - Levinson, obtemos:

$$\hat{a}(h) = \frac{\Delta_h}{P_h} , \quad (6.2.3)$$

$$\hat{a}_h(j) = \hat{a}_{h-1}(j) - \hat{a}_h(h) \hat{a}_{h-1}(h-j) , \quad j = 1, \dots, h-1. \quad (6.2.4)$$

A variância residual para o modelo ajustado de ordem h é dada por:

$$\hat{\sigma}_h^2 = P_{h+1} \quad (6.2.5)$$

A estimação recursiva dada por (6.2.3), (6.2.4) e (6.2.5), tem grande importância na metodologia proposta, pelo fato de h ser obtida pela minimização de um critério do tipo:

$$AIC_{C(T)}(h) = \log \hat{\sigma}_h^2 + \frac{h C(T)}{T} , \quad h \leq H(T) , \quad (6.2.6)$$

onde

$$\frac{C(T)}{T} \rightarrow 0.$$

Utilizamos para $C(T)$, em nossas aplicações, três diferentes funções, a saber:

$C(T) = 2$, correspondendo ao AIC (an Information Criterion), proposto por Akaike (1973a, 1973b, 1974);

$C(T) = \log T$, correspondendo ao BIC (Bayesian Information Criterion), proposto por Akaike (1977) e Schwarz (1978);

$C(T) = c \log \log T$ para $c > 2$, correspondendo ao HQ, critério proposto por Hannan e Quinn (1979). Em nossas aplicações consideramos para c o valor 2,2.

Denotaremos os estimadores, de 2º estágio, obtidos via AIC, BIC e HQ, respectivamente por $\hat{\beta}_{AIC}$, $\hat{\beta}_{BIC}$ e $\hat{\beta}_{HQ}$.

Em nossas simulações consideramos também h fixado e grande. O estimador de 2º estágio, assim obtido será denotado por $\hat{\beta}_{hfix}$.

Alguns autores fizeram avaliações teóricas a respeito dos critérios definidos acima. Assim, Hannan (1980) mostrou que, sob determinadas condições, o BIC e o HQ são consistentes para estimar a ordem de um modelo ARMA(p, q) (os critérios foram redefinidos para essa situação). Shibata (1980) mostrou a "eficiência assintótica" do AIC para um modelo auto-regressivo de ordem infinita.

6.3 As situações onde as simulações foram empregadas

O método de estimação foi aplicado ao modelo

$$y_t = 2u_t + x_t,$$

onde, para u_t , nas situações (I) até (IV), descritas abaixo, foram consideradas as seguintes seqüências determinísticas:

(a) $u_t = t$,

(b) $u_t = \cos \lambda t$, $\lambda = \frac{2\pi}{4}$,

$$(c) \quad u_t = \frac{t^2}{10} ,$$

e na situação (V), consideramos apenas a sequência em (a).

Para x_t considerou-se modelos MA(1) e/ou ARMA(1,1). Os parâmetros de tais modelos foram definidos segundo os objetivos que tínhamos em mente e que descreveremos abaixo.

(I) Para avaliar o comportamento de $\hat{\beta}_{AIC}$, $\hat{\beta}_{BIC}$ e $\hat{\beta}_{HQ}$ em relação a $\hat{\beta}_{MQO}$ quando se modifica o tamanho da amostra, consideramos para T os valores 50, 100 e 200 e para x_t os modelos

$$x_t = a_t - 0,5 a_{t-1}$$

e

$$x_t = 0,5 x_{t-1} + a_t + 0,5 a_{t-1}$$

onde:

na situação (a) tomamos $a_t \sim N(0, 10)$,

na situação (b) tomamos $a_t \sim N(0, 1)$,

na situação (c) tomamos $a_t \sim N(0, 100)$.

Essas diferenças na variabilidade do ruído, σ_a^2 justificam-se devido à diferença de variabilidade das seqüências consideradas para u_t .

Na seção 6.4, tabelas 6.4.1, 6.4.2 e 6.4.3 apresentamos os resultados obtidos.

(II) Pretendíamos avaliar o comportamento de $\hat{\beta}_{AIC}$, $\hat{\beta}_{BIC}$ e $\hat{\beta}_{HQ}$ em relação a $\hat{\beta}_{MQO}$, quando se modifica a razão "sinal/ruído". Normalmente, se u_t é estacionária estocástica, essa razão é dada pelo quociente entre os desvios padrões de u_t e x_t , ou seja por:

$$r = \sqrt{\frac{\gamma_u(0)}{\gamma_x(0)}}$$

Na nossa situação, avaliamos $\gamma_u(0)$ por meio de

$$\gamma_u(0) = \frac{\sum_{t=1}^T (u_t - \bar{u})^2}{T}, \text{ onde } \bar{u} = \frac{\sum_{t=1}^T u_t}{T},$$

quando $u_t = t$ ou $u_t = t^2/10$. Consideramos $\sqrt{\gamma_u(0)}$ como sendo o valor da amplitude da onda cosseno, que quando $u_t = \cos \lambda t$ é igual a 1 (Priestley, 1981, pag. 643).

Consideramos $T = 100$ e para r os valores 2 e $\frac{1}{2}$.

(IIIi) Para fazer $r = 2$ quando

$$x_t = a_t - 0,5 a_{t-1},$$

tomamos

$$a_t \sim N(0,166) \text{ na situação (a).}$$

$$a_t \sim N(0; 0,20) \text{ na situação (b).}$$

$$a_t \sim N(0,18057) \text{ na situação (c).}$$

e quando

$$x_t = 0,5 x_{t-1} + a_t + 0,5 a_{t-1} ,$$

tomamos

$$a_t \sim N(0,89) \text{ na situação (a) ,}$$

$$a_t \sim N(0; 0,11) \text{ na situação (b) ,}$$

$$a_t \sim N(0,9676) \text{ na situação (c).}$$

(IIii) Para fazer $r = \frac{1}{2}$ quando

$$x_t = a_t - 0,5 a_{t-1} ,$$

tomamos

$$a_t \sim N(0,2658) \text{ na situação (a) ,}$$

$$a_t \sim N(0; 3,2) \text{ na situação (b) ,}$$

$$a_t \sim N(0,288906) \text{ na situação (c) ;}$$

e quando

$$x_t = 0,5 x_{t-1} + a_t + 0,5 a_{t-1} ,$$

tomamos

$$a_t \sim N(0,1424) \text{ na situação (a) ,}$$

$$a_t \sim N(0; 1,71) \text{ na situação (b) ,}$$

$$a_t \sim N(0,154814) \text{ na situação (c).}$$

Os resultados obtidos estão apresentados nas tabelas 6.4.4 e 6.4.5.

- (III) Consideramos a situação em que $x_t = a_t - 0,8 a_{t-1}$ com $a_t \sim N(0, 2035)$, $a_t \sim N(0; 2,44)$ e $a_t \sim N(0, 221194)$, respectivamente para as situações (a), (b) e (c); e $x_t = 0,8 x_{t-1} + a_t + 0,8 a_{t-1}$, com $a_t \sim N(0, 410)$, $a_t \sim N(0; 0,49)$ e $a_t \sim N(0, 44617)$, respectivamente para as situações (a), (b) e (c).

O objetivo é avaliar o que ocorre com estimadores $\hat{\beta}_{AIC}$, $\hat{\beta}_{BIC}$ e $\hat{\beta}_{HQ}$, quando os parâmetros dos modelos MA(1) e ARMA(1,1), considerados para x_t , estão mais próximos da fronteira da região de admissibilidade. Essa situação pode ser comparada com a situação (IIii), pois tomamos aqui $T = 100$, e os valores de σ_a^2 considerados tornam $r = \frac{1}{2}$.

Os resultados estão apresentados na tabela 6.4.6.

- (IV) Com o objetivo de comparar $\hat{\beta}_{hfix}$ com $\hat{\beta}_{MQO}$, como também $\hat{\beta}_{AIC}$, $\hat{\beta}_{BIC}$ e $\hat{\beta}_{HQ}$, consideramos $T = 100$, $h = 9$, e para x_t os modelos da situação (I).

Os resultados encontram-se na tabela 6.4.7.

- (V) Finalmente, na situação em que $u_t = t$, consideramos para x_t modelos MA(1) onde $\sigma_a^2 = 10$, e onde o parâmetro de médias móveis, θ , varia de -0,9 até 0,9, com incrementos de 0,2.

Os resultados estão apresentados na tabela 6.4.8.

6.4 Resultados e Comentários

Para todas as situações apresentadas na seção 6.3, foram considerados os resultados de 50 réplicas e avaliou-se o viés amostral de cada estimador e as razões entre a variância amostral de $\hat{\beta}_{MQO}$ e a variância amostral do estimador de 2º estágio. As comparações foram feitas sempre em termos da razão entre as variâncias.

A função $H(T)$ foi considerada como:

$$H(T) = \begin{cases} 8, & \text{se } T = 50 \\ 9, & \text{se } T = 100 \\ 10, & \text{se } T = 200 \end{cases}$$

Sempre que x_t for gerado por um ARMA(1.1) tomamos $x_0 = 0$.

A tabela 6.4.1 apresenta os resultados da situação (I) para $T = 50$. Notamos que o estimado de MQO, foi superior aos estimadores obtidos via AIC, BIC e HQ quando a variável exógena é a função periódica $\cos \lambda t$, e na situação em que $u_t = t$ e x_t é o modelo MA. Nas demais situações os estimadores propostos apresentaram melhor desempenho, sendo este mais significativo para a tendência quadrática.

A tabela 6.4.2 condensa os resultados da situação (I) para $T = 100$. Notamos que os estimadores propostos foram ligeiramente superiores ao estimador de MQO, exceção feita ao caso em que $u_t = \cos \lambda t$ e x_t é gerado pelo MA(1).

A tabela 6.4.3 apresenta os resultados da situação (I) para $T = 200$. Observamos que os estimadores propostos foram superiores a $\hat{\beta}_{MQO}$ em todas as situações.

Tabela 6.4.1 - Viés de $\hat{\beta}_{MQO}$ e dos estimadores obtidos no 2º estágio. Razão entre $\text{var } \hat{\beta}_{MQO}$ e a variância dos estimadores obtidos no 2º estágio. $T = 50$.

Modelo → Viés Razão ↓	$u_t = t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$
Viés de $\hat{\beta}_{MQO}$	- 0,0008	- 0,0019	0,0665	- 0,0019	- 0,0008	- 0,0027
Viés de $\hat{\beta}_{AIC}$	- 0,0006	- 0,0025	0,0517	0,010	- 0,0005	- 0,0005
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{AIC}}$	0,931	1,023	0,992	0,878	1,119	1,117
Viés de $\hat{\beta}_{BIC}$	- 0,0005	- 0,0028	0,0640	- 0,0034	- 0,0006	- 0,0004
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{BIC}}$	0,953	1,004	1,003	0,873	1,083	1,134
Viés de $\hat{\beta}_{HQ}$	- 0,0005	- 0,0026	0,0647	0,0011	- 0,0006	- 0,0001
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{HQ}}$	0,953	1,006	0,993	0,868	1,093	1,123

Tabela 6.4.2 - Viés de $\hat{\beta}_{MQO}$ e dos estimadores obtidos no 2º estágio. Razão entre $\text{var } \hat{\beta}_{MQO}$ e a variância dos estimadores obtidos no 2º estágio. $T = 100$.

Modelo + Viés Razão +	$u_t = t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$
Viés de $\hat{\beta}_{MQO}$	0,0003	0	- 0,0068	- 0,0256	0,0001	0,0008
Viés de $\hat{\beta}_{AIC}$	0,0003	0	- 0,0083	- 0,0252	0	0,0011
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{AIC}}$	1,050	1,019	0,974	1,001	1,056	1,070
Viés de $\hat{\beta}_{BIC}$	0,0003	0	- 0,0069	- 0,0251	0	0,0011
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{BIC}}$	1,040	1,002	0,992	1,031	1,050	1,046
Viés de $\hat{\beta}_{HQ}$	0,0003	0	- 0,0073	- 0,0255	0	0,0011
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{HQ}}$	1,046	1,011	0,997	1,021	1,066	1,046

Tabela 6.4.3 - Viés de $\hat{\beta}_{MQO}$ e dos estimadores obtidos no 2º estágio. Razão entre $\text{var } \hat{\beta}_{MQO}$ e a variância dos estimadores obtidos no 2º estágio. $T = 200$.

Modelo + Viés Razão +	$u_t = t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$
Viés de $\hat{\beta}_{MQO}$	0,0001	0	- 0,0049	- 0,0126	0	0
Viés de $\hat{\beta}_{AIC}$	0,0001	0	- 0,0048	- 0,0120	0	0
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{AIC}}$	1,130	1,042	1,010	1,008	1,104	1,022
Viés de $\hat{\beta}_{BIC}$	0,0001	0	- 0,0053	- 0,0121	0	0
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{BIC}}$	1,120	1,030	1,027	1,012	1,052	1,030
Viés de $\hat{\beta}_{HQ}$	0,0001	0	- 0,0048	- 0,0120	0	0
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{HQ}}$	1,136	1,066	1,023	1,015	1,065	1,030

Observando conjuntamente as três primeiras tabelas, notamos que há uma melhora progressiva dos estimadores de 2º estágio à medida que T cresce, na situação em que u_t é a tendência linear. Tais estimadores não apresentam bom desempenho, relativamente ao estimador de MQO, quando $T = 50$ e $T = 100$, se u_t é a função periódica. Se u_t é a função quadrática, os resultados refletem uma superioridade dos mesmos, relativamente a $\hat{\beta}_{MQO}$, que parece não melhorar quando T cresce. O viés amostral de cada um dos estimadores de 2º estágio nem sempre é menor do que o viés de $\hat{\beta}_{MQO}$. Tanto para este último estimador como para os anteriores, ele é sempre maior para a função cosseno do que para as demais funções. A medida que T cresce, notamos uma diminuição do viés para todas as situações consideradas.

As tabelas 6.4.4 e 6.4.5, apresentam os resultados das situações (IIi) e (IIii), respectivamente. Exceção feita ao $\hat{\beta}_{BIC}$, quando x_t é o modelo MA(1) e $u_t = \cos \lambda t$ na situação (IIi), todos os estimadores apresentaram um melhor desempenho do que $\hat{\beta}_{MQO}$, nas duas situações. Comparando as duas tabelas entre si, notamos que ao se diminuir a razão sinal/ruído aumenta-se o viés dos estimadores em geral, mas diminui-se a variância dos estimadores propostos relativamente à variância de $\hat{\beta}_{MQO}$, exceção feita ao caso em que $u_t = \cos \lambda t$.

Tabela 6.4.4 – Viés de $\hat{\beta}_{MQO}$ e dos estimadores obtidos no 2º estágio. Razão entre $\text{var } \hat{\beta}_{MQO}$ e a variância dos estimadores obtidos no 2º estágio. T = 100, r = 2.

Modelo + Viés Razão +	$u_t = t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$
Viés de $\hat{\beta}_{MQO}$	- 0,0009	- 0,0009	- 0,0089	0,0089	0,0002	0,0004
Viés de $\hat{\beta}_{AIC}$	- 0,0005	- 0,0010	- 0,0068	0,0082	0,0006	0,0006
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{AIC}}$	1,057	1,035	1,005	1,032	1,036	1,026
Viés de $\hat{\beta}_{BIC}$	- 0,0005	- 0,0009	- 0,0075	0,0087	0,0021	0,0006
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{BIC}}$	1,052	1,064	0,984	1,073	1,038	1,010
Viés de $\hat{\beta}_{HQ}$	- 0,0005	- 0,0009	- 0,0065	0,0090	0,0016	0,0006
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{HQ}}$	1,052	1,064	1,012	1,076	1,034	1,008

Tabela 6.4.5 – Viés de $\hat{\beta}_{MQO}$ e dos estimadores obtidos no 2º estágio. Razão entre $\text{var. } \hat{\beta}_{MQO}$ e a variância dos estimadores obtidos no 2º estágio. T = 100, r = 1/2.

Modelo + Viés Razão +	$u_t = t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$
Viés de $\hat{\beta}_{MQO}$	- 0,0024	0,0115	- 0,0538	0,0211	0,0156	0,0247
Viés de $\hat{\beta}_{AIC}$	- 0,0011	0,0004	- 0,0578	0,0257	0,0118	0,0233
- Var $\hat{\beta}_{MQO}$	-	-	-	-	-	-
Var $\hat{\beta}_{AIC}$	1,130	1,067	1,026	1,027	1,082	1,022
Viés de $\hat{\beta}_{BIC}$	- 0,0015	0,0018	- 0,0574	0,0307	- 0,0121	0,0233
- Var $\hat{\beta}_{MQO}$	-	-	-	-	-	-
Var $\hat{\beta}_{BIC}$	1,088	1,059	1,055	1,030	1,110	1,033
Viés de $\hat{\beta}_{HQ}$	- 0,0012	0,0017	- 0,0560	0,0292	0,0120	0,0244
- Var $\hat{\beta}_{MQO}$	-	-	-	-	-	-
Var $\hat{\beta}_{HQ}$	1,138	1,081	1,044	1,035	1,094	1,022

Os resultados da situação (III) encontram-se na tabela 6.4.6. Em todos os casos considerados o desempenho dos estimadores $\hat{\beta}_{AIC}$, $\hat{\beta}_{BIC}$ e $\hat{\beta}_{HQ}$ foi melhor do que o de $\hat{\beta}_{MQO}$. Esses resultados quando comparados com os da situação (IIii), revelaram uma melhora dos estimadores de 2º estágio, quando os coeficientes do modelo considerado para x_t estão mais próximos da fronteira da região de admissibilidade, exceção feita à situação em que $u_t = \cos \lambda t$.

Tabela 6.4.6 - Viés de $\hat{\beta}_{MQO}$ e dos estimadores obtidos no 2º estágio. Razão entre var $\hat{\beta}_{MQO}$ e a variância dos estimadores obtidos no 2º estágio. $T = 100$, $r = 1/2$.

Modelo+ Viés Razão +	$u_t = t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$
Viés de $\hat{\beta}_{MQO}$	0,0027	0,0010	- 0,0071	- 0,0097	0,0001	0,0649
Viés de $\hat{\beta}_{AIC}$	0,0005	0,0008	0,0076	- 0,0088	- 0,0011	0,0597
Var $\hat{\beta}_{MQO}$	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -
Var $\hat{\beta}_{AIC}$	1,593	1,358	1,007	1,015	1,316	1,060
Viés de $\hat{\beta}_{BIC}$	0,0008	0,0009	0,0041	- 0,0069	- 0,0006	0,0664
Var $\hat{\beta}_{MQO}$	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -
Var $\hat{\beta}_{BIC}$	1,535	1,225	1,027	1,030	1,149	1,051
Viés de $\hat{\beta}_{HQ}$	0,0006	0,0008	0,0036	- 0,0097	- 0,0007	0,0626
Var $\hat{\beta}_{MQO}$	- - - -	- - - -	- - - -	- - - -	- - - -	- - - -
Var $\hat{\beta}_{HQ}$	1,623	1,282	1,011	1,020	1,178	1,078

A tabela 6.4.7 mostra que o estimador de 2º estágio obtido para h fixado nem sempre foi superior a $\hat{\beta}_{MQO}$. Na situação em que $u_t = \cos \lambda t$ ele mostrou-se inferior, bem como quando u_t satisfaz (c) e x_t é gerado por um processo ARMA(1.1). Comparando essa tabela com a tabela 6.4.2, notamos também que $\hat{\beta}_{hfix}$ mostrou-se inferior a $\hat{\beta}_{AIC}$, $\hat{\beta}_{BIC}$ e $\hat{\beta}_{HQ}$.

Tabela 6.4.7 - Viés de $\hat{\beta}_{MQO}$ e $\hat{\beta}_{hfix}$. Razão entre $\text{var } \hat{\beta}_{MQO}$ e $\text{var } \hat{\beta}_{hfix}$. $T = 100$, $h = 9$.

Modelo + Viés Razão	$u_t = t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = a_t - 0,5a_{t-1}$	$u_t = \cos \lambda t$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = a_t - 0,5a_{t-1}$	$u_t = t^2/10$ $x_t = 0,5x_{t-1} + a_t + 0,5a_{t-1}$
Viés de $\hat{\beta}_{MQO}$	0,0003	0	- 0,0068	- 0,0256	0,0001	0,0008
Viés de $\hat{\beta}_{hfix}$	0	0,0003	0,0300	- 0,0124	0	0
$\frac{\text{var } \hat{\beta}_{MQO}}{\text{var } \hat{\beta}_{hfix}}$	1,018	1,012	0,969	0,854	1,045	0,9712

Através da tabela 6.4.8 podemos ter uma idéia do que ocorre com o desempenho dos estimadores de 2º estágio, à medida que θ varia. Observamos que, relativamente a $\hat{\beta}_{MQO}$, os estimadores $\hat{\beta}_{AIC}$, $\hat{\beta}_{BIC}$ e $\hat{\beta}_{HQ}$ mostraram-se:

- . significativamente melhores para θ próximo de 1 ($\theta = 0,9$) e ligeiramente inferiores para θ próximo de -1 ($\theta = -0,9$)
- . de modo geral ligeiramente inferiores para θ entre -0,3 e +0,3;
- . superiores para $\theta = |0,6|$ e $\theta = |0,7|$;
- . melhores para valores positivos do que para valores negativos de θ

Tabela 6.4.8 – Viés de $\hat{\beta}_{MQO}$ e dos estimadores obtidos no 2º estágio. Razão entre $\text{Var } \hat{\beta}_{MQO}$ e a variância dos estimadores obtidos no 2º estágio. $T = 100$, $u_t = t$, $x_t = a_t - \theta a_{t-1}$

θ Viés Razão	-0,9	-0,7	-0,5	-0,3	-0,1	0,1	0,3	0,5	0,7	0,9
Viés de $\hat{\beta}_{MQO}$	0,0017	0,0001	0,0021	0,0013	0,0007	-0,0002	-0,0004	0,0003	-0,0001	0
Viés de $\hat{\beta}_{AIC}$	0,0018	0,0001	0,0021	0,0013	0,0007	-0,0002	-0,0004	0,0003	-0,0002	0
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{AIC}}$	- - - 0,9738	- - - 1,036	- - - 1,028	- - - 0,9910	- - - 1,011	- - - 0,9886	- - - 0,9886	- - - 1,050	- - - 1,2951	- - - 2,1878
Viés de $\hat{\beta}_{BIC}$	0,0017	0,0001	0,0021	0,0013	0,0007	-0,0002	-0,0004	0,0003	-0,0002	0
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{BIC}}$	- - - 0,9943	- - - 1,029	- - - 1,025	- - - 0,9929	- - - 0,9962	- - - 0,9970	- - - 0,9966	- - - 1,040	- - - 1,2212	- - - 1,9897
Viés de $\hat{\beta}_{\phi}$	0,0018	0,0001	0,0021	0,0013	0,0007	-0,0002	-0,0004	0,0003	-0,0002	0
$\frac{\text{Var } \hat{\beta}_{MQO}}{\text{Var } \hat{\beta}_{\phi}}$	- - - 0,9828	- - - 1,032	- - - 1,018	- - - 0,9835	- - - 1,002	- - - 0,9958	- - - 0,9905	- - - 1,046	- - - 1,2226	- - - 1,9891

Observando todas as situações consideradas, com o objetivo de comparar $\hat{\beta}_{AIC}$, $\hat{\beta}_{BIC}$ e $\hat{\beta}_{HQ}$ entre si, notamos que não existe evidência de superioridade de qualquer um deles sobre os outros dois, em determinadas situações específicas.

CAPÍTULO 7

DISCUSSÃO E RECOMENDAÇÕES

7.1 Discussão

Neste trabalho apresentamos estimadores para um modelo linear de regressão com erros correlacionados, sob algumas suposições a respeito desses erros e das variáveis independentes.

Do ponto de vista aplicado, consideramos o modelo de regressão linear dado em (1.1.1), com uma só variável exógena a qual supusemos fixada. Comparamos o estimador de MQO com o estimador de COG2 para os três critérios propostos para a determinação de h , e também com o estimador de CO2 (h fixado e grande).

As conclusões que tiramos dos resultados experimentais para as particulares seqüências usadas para gerar $\{u_t\}$ e $\{x_t\}$, e para os particulares tamanhos de amostra utilizados, podem ser sintetizadas como segue:

(a) o estimador de COG2 apresenta melhor desempenho do que o estimador de MQO para amostras médias e grandes ($T=100$ e $T=200$), quando $\{u_t\}$ é a tendência linear ou quadrática; com $\{u_t\}$ assim gerado, a situação em que a variância de $\{x_t\}$ é grande favorece o uso do estimador COG2;

(b) o estimador CO2 apresenta pior desempenho do que o estimador de MQO na situação em que a variável exógena é periódica ($u_t = \cos \lambda t$), e não se mostra significativamente superior quando a mesma é a tendência linear;

(c) para o estimador COG2, não fica evidente a superioridade da escolha de h por qualquer um dos três critérios utilizados.

7.2. Recomendações

Observamos, através da literatura, que pequena experiência tem sido ganha, até hoje, quanto ao mérito de cada um dos procedimentos de estimação em dois estágios descritos na seção 1.2 do Capítulo 1. Sendo assim, algumas recomendações podem ser feitas.

Com relação ao enfoque não paramétrico, isto é, regressão no domínio da frequência, Harvey (1978, 1981b) aponta algumas razões para o seu uso. Uma, é permitir a aplicação da técnica conhecida como "band spectrum regression", que consiste em omitir observações de certas frequências. Isto tem interesse nas aplicações pois permite lidar com sazonalidade e erros nas variáveis independentes. Outra razão seria dada pelo fato de que se os erros são correlacionados no domínio do tempo, eles serão aproximadamente não correlacionados no domínio da frequência, e as vantagens decorrem do fato de que as variâncias dos erros transformados, estão bastante relacionados com a densidade espectral dos erros originais. Também, se o processo gerador do termo de erro no modelo de regressão não é bem especificado, existe uma considerável perda de eficiência. Neste sen

tido o enfoque espectral pode ser sugerido, pois ele é não paramétrico e portanto não se fia numa particular especificação do processo em questão.

Hannan (1963), desenvolveu uma teoria assintótica para estimadores obtidos no domínio da frequência. Embora a metodologia proposta conduza a estimadores assintoticamente eficientes, existem poucos trabalhos experimentais que a utilizam. As razões para isso devem-se provavelmente ao fato dela ser computacionalmente mais trabalhosa, e por presumivelmente exigir amostras grandes. Engle e Gardner (1976) apresentaram os resultados de simulações, para uma variedade de situações econométricas típicas, as quais indicaram que os estimadores espectrais têm um bom desempenho para amostras médias ($T = 100$), quando o processo gerador do erro é altamente dependente, e até para amostras pequenas quando o mesmo é o ruído branco. Esses resultados desmistificam em parte a exigência de amostras grandes.

Outro ponto importante, que merece recomendação, é a comparação dos diversos procedimentos. Assim, se se ajusta uma auto-regressão aos resíduos de primeiro estágio, quando aplicar o estimador de Cochrane-Orcutt e quando usar as transformadas de Fourier no segundo estágio? Ou, dada a decisão de usar métodos no domínio da frequência em que situações deveremos optar por estimador espectral auto-regressivo e em que situações deveremos optar por periodograma suavizado? Para responder a essas e outras questões, o que se sugere é que todas as opções sejam consideradas e, sempre que possível, mais do que uma possa ser usada sob as mesmas suposições a respeito do modelo. Simulações podem ser feitas para comparar os diversos procedimentos, e determinar quão robusto é o estimador de MQG às várias implementações amostrais finitas.

Do ponto de vista teórico, existem trabalhos que tratam de regressão em séries temporais, e que mostram a eficiência e a normalidade assintótica de estimadores de MQG. Assim, Wickens(1969), prova a equivalência assintótica de dois estimadores de MQG para o caso em que os erros seguem um modelo auto-regressivo. Amemiya (1973) mostra a normalidade e eficiência assintóticas, do estimador de MQG que utiliza auto-covariâncias residuais estimadas, num modelo de regressão cujos erros seguem um processo ARMA. Este, para fim de estimação da matriz de covariâncias é aproximado por uma auto-regressão de ordem N. O autor mostra, então, o problema da contraposição entre N e T (tamanho da amostra). Deixa como um problema em aberto, a obtenção de uma função de N e T, $\phi(N,T)$, que aproxime a distribuição do estimador proposto da distribuição do melhor estimador linear não viciado, e a minimização de tal função com respeito a N, para um dado T.

Com relação ao nosso trabalho, uma recomendação para um possível desenvolvimento teórico é a de se avaliar o vício e o erro quadrativo médio (EQM), do estimador proposto, $\hat{\beta}$, para T fixado e h escolhido segundo (6.2.6). Também, sob esse aspecto, talvez se possa determinar que função C(T) é mais recomendada para particulares escolhas de $\{u_t\}$. A análise da representação espectral de $\hat{\beta}$ pode ser útil nesse sentido. Na seção seguinte tecemos alguns comentários a respeito disso, sob a suposição de que $\{u_t\}$ é estocástica. Na última seção mostramos alguns desenvolvimentos que podem ser úteis para a obtenção do viés e do EQM de $\hat{\beta}$.

7.3 Análise da representação espectral de $\hat{\beta}$

Sabemos que o estimador $\hat{\beta}$, baseado em T observações (y_t, u_t) é dado por:

$$\hat{\beta} = \frac{\sum_{t=h+1}^T \hat{\eta}_t \hat{\psi}_t}{\sum_{t=h+1}^T \hat{\psi}_t^2} \quad (7.3.1)$$

Temos que:

$$\begin{aligned} \frac{1}{T} \sum_{t=h+1}^T \hat{\eta}_t \hat{\psi}_t &= \sum_{j=0}^h \sum_{k=0}^h \hat{a}_h(j) \hat{a}_h(k) \cdot \frac{1}{T} \sum_{t=h+1}^T y(t-j) u(t-k) \\ &\approx \sum_{j=0}^h \sum_{k=0}^h \hat{a}_h(j) \hat{a}_h(k) \frac{1}{T} \sum_{t=h+1}^T y(t) u(t+j-k), \end{aligned}$$

onde Σ' é estendida a todos os índices t tais que $1 \leq t, t+j-k \leq T$.

Assim,

$$\frac{1}{T} \sum_{t=h+1}^T \hat{\eta}_t \hat{\psi}_t \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} I_{yu}(w) \left| \sum_{j=0}^h \hat{a}_h(j) e^{-ijw} \right|^2 dw, \quad (7.3.2)$$

onde I_{yu} é o periodograma cruzado

$$I_{yu}(w) = \frac{1}{T} \sum_{t=1}^T y_t e^{-iwt} \sum_{s=1}^T u_s e^{iws}. \quad (7.3.3)$$

De modo análogo, vemos que

$$\frac{1}{T} \sum_{t=h+1}^T \hat{\psi}_t^2 \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} I_{uu}(w) \left| \sum_{j=0}^h \hat{a}_h(j) e^{-ijw} \right|^2 dw, \quad (7.3.4)$$

onde $I_{uu}(w)$ é o periodograma de u_t , $t=1, \dots, T$,

$$I_{uu}(w) = \frac{1}{T} \left| \sum_{t=1}^T u_t e^{-iwt} \right|^2. \quad (7.3.5)$$

Segue de (7.3.1) que $\hat{\beta}$ pode ser escrito como

$$\hat{\beta} \approx \frac{\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_{yu}(w)}{I_{uu}(w)} \left\{ I_{uu}(w) \left| \sum_{j=0}^h \hat{a}_h(j) e^{-ijw} \right|^2 \right\} dw}{\frac{1}{2\pi} \int_{-\pi}^{\pi} I_{uu}(w) \left| \sum_{j=0}^h \hat{a}_h(j) e^{-ijw} \right|^2 dw} \quad (7.3.6)$$

Mas, o denominador de (7.3.6) é uma constante, e portanto:

$$\hat{\beta} = \text{cte} \int_{-\pi}^{\pi} \frac{I_{yu}(w)}{I_{uu}(w)} \left\{ I_{uu}(w) \left| \sum_{j=0}^h \hat{a}_h(j) e^{-ijw} \right|^2 \right\} dw. \quad (7.3.7)$$

Como $\frac{I_{yu}(w)}{I_{uu}(w)}$ é uma estimativa de β , podemos ver (7.3.7) como uma média ponderada de estimativas de β , obtidas para cada frequência. A ponderação é dada por:

$$g(w) = I_{uu}(w) \cdot \left| \sum_{j=0}^h \hat{a}_h(j) e^{-ijw} \right|^2, \quad (7.3.8)$$

que, na realidade, é um estimador de $\frac{f_u(w)}{f_x(w)}$, onde $f_u(w)$ é o espectro de u_t e $f_x(w)$ é o espectro de x_t , dado por

$$f_x(w) = \frac{\sigma_\epsilon^2}{2\pi} \cdot \frac{1}{\left| \sum_{j=0}^{\infty} a(j) e^{-ijw} \right|^2}.$$

Notando que a ponderação ótima é:

$$I_{uu}(w) \left| \sum_{j=0}^{\infty} a(j) e^{-ijw} \right|^2,$$

observamos que:

- (a) Se $f_u(w)$ for bastante concentrado sobre uma faixa estreita de frequências, então uma estimativa muito precisa de $f_x(w)$ não é necessária. Segue que h pode crescer vagarosamente com T , o que implica que $C(T)$ deverá crescer rapidamente com T .
- (b) Se $f_u(w)$ for muito variável, então precisamos de uma boa estimativa de $f_x(w)$ e, para tanto, necessitamos de boas estimativas de $a(j)$. Nesse caso $C(T)$ deve crescer devagar com T ou mesmo ser constante.

7.4 Considerações a respeito do vício e do EQM de $\hat{\beta}$

De (7.3.3) e (7.3.5) segue que:

$$\begin{aligned} \frac{I_{yu}(w)}{I_{uu}(w)} &= \frac{\sum_{t=1}^T y_t e^{-itw}}{\sum_{t=1}^T u_t e^{-itw}} = \beta \cdot \frac{\sum_{t=1}^T u_t e^{-itw}}{\sum_{t=1}^T u_t e^{-itw}} + \frac{\sum_{t=1}^T x_t e^{-itw}}{\sum_{t=1}^T u_t e^{-itw}} = \\ &= \beta + \frac{d_x(w)}{d_u(w)}, \end{aligned} \quad (7.4.1)$$

onde $d_x(w)$ e $d_u(w)$ são as transformadas de Fourier finitas de $x(t)$ e $u(t)$, $t=1, \dots, T$, respectivamente.

De (7.3.6) e (7.4.1) obtemos:

$$\hat{\beta} = \beta + \frac{\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{d_x(w)}{d_u(w)} g(w) dw}{\frac{1}{2\pi} \int_{-\pi}^{\pi} g(w) dw}. \quad (7.4.2)$$

Chamando

$$G(\theta) = \frac{g(\theta)}{\int_{-\Pi}^{\Pi} g(w) dw}, \quad (7.4.3)$$

vemos que $\int_{-\Pi}^{\Pi} G(\theta) d\theta = 1$ e que (7.4.2) pode ser escrita como

$$\hat{\beta} = \beta + J, \quad (7.4.4)$$

onde

$$J = \int_{-\Pi}^{\Pi} \frac{d_x(w)}{d_u(w)} G(w) dw \quad (7.4.5)$$

Como o vício e o EQM de $\hat{\beta}$ são dados respectivamente por $E(J)$ e $E(J^2)$, achamos conveniente o desenvolvimento da integral em (7.4.5), pois ela será útil para a obtenção dos mesmos.

Temos:

$$\begin{aligned} \int_{-\Pi}^{\Pi} \frac{\sum_{t=1}^T x_t e^{-itw}}{\sum_{t=1}^T u_t e^{-itw}} g(w) dw &= \frac{1}{T} \int_{-\Pi}^{\Pi} \frac{\sum_{t=1}^T x_t e^{-itw}}{\sum_{t=1}^T u_t e^{-itw}} \left| \sum_{j=0}^h a_h(j) e^{-ijw} \right|^2 \\ &= \frac{1}{T} \int_{-\Pi}^{\Pi} \left[\sum_{t, t'=1}^T x_t u_{t'} \cos w(t-t') \right] \left| \sum_{j=0}^h a_h(j) e^{-ijw} \right|^2 dw \quad (7.4.6) \end{aligned}$$

Desenvolvendo o integrando de (7.4.6) obtemos:

$$\frac{1}{T} \int_{-\Pi}^{\Pi} \sum_{\substack{1 \leq t, t' \leq T \\ 0 \leq j, j' \leq h}} x_t u_{t'} a_h(j) a_h(j') \cos w(t'-t) \cos w(j'-j) dw \quad (7.4.7)$$

Podemos considerar os seguintes casos:

Caso 1: $t=t'$ e $j \neq j'$ ($j-j' \in \mathbb{Z}$)

$$\int_{-\Pi}^{\Pi} \cos w(j-j') dw = \frac{\sin w(j'-j)}{j'-j} \Big|_{-\Pi}^{\Pi} = 0$$

Caso 2: $t \neq t'$ e $j = j'$

$$\int_{-\Pi}^{\Pi} \cos w(t'-t) dw = 0;$$

Caso 3: $t = t'$ e $j = j'$

$$\int_{-\Pi}^{\Pi} 1 dw = w \Big|_{-\Pi}^{\Pi} = 2\Pi;$$

Caso 4: $t = t'$, $j = j'$, $\alpha = t' - t$, $\beta = j' - j$ e $\alpha \neq \beta$, $\alpha \neq -\beta$

$$\begin{aligned} \int_{-\Pi}^{\Pi} \cos \alpha w \cos \beta w dw &= \int_0^{\Pi} [\cos(\alpha+\beta)w + \cos(\alpha-\beta)w] dw \\ &= \frac{\sin(\alpha+\beta)w}{\alpha+\beta} + \frac{\sin(\alpha-\beta)w}{\alpha-\beta} \Big|_0^{\Pi} = 0; \end{aligned}$$

Caso 5: $t \neq t'$, $j \neq j'$, $\alpha = t' - t$, $\beta = j' - j$, $|\alpha| = |\beta|$

$$\int_{-\Pi}^{\Pi} \cos^2 w(t'-t) dw = 2 \int_0^{\Pi} \cos^2 \alpha w dw = \Pi.$$

Conclui-se então que:

$$\int_{-\Pi}^{\Pi} \frac{d_x(w)g(w)dw}{d_u(w)} = \frac{2\Pi}{T} \sum_{\substack{1 \leq t \leq T \\ 0 \leq j < h}} x_t u_t a_h^2(j) + \frac{\Pi}{T} \sum_{\substack{1 \leq t, t' \leq T \\ 0 \leq j, j' \leq h \\ t-t'=j-j' \\ t \neq t' \text{ e } j \neq j'}} x_t u_t a_h(j) a_h(j'). \quad (7.4.8)$$

Por outro lado,

$$\int_{-\Pi}^{\Pi} g(w) dw = \frac{1}{T} \left| \sum_{t=1}^T u_t e^{-i t w} \right|^2 \left| \sum_{j=0}^h a_h(j) e^{-i j w} \right|^2$$

$$= \frac{1}{T} \int_{-\Pi}^{\Pi} \sum_{\substack{1 \leq t, t' \leq T \\ 0 \leq j, j' \leq h}} u_t u_{t', a_h(j)} a_h(j') \cos w(t-t') \cos w(j-j').$$

Portanto,

$$\int_{-\Pi}^{\Pi} g(w) dw = \frac{2\Pi}{T} \sum_{\substack{1 \leq t \leq T \\ 0 \leq j \leq h}} u_t^2 a_h(j) + \frac{\Pi}{T} \sum_{\substack{1 \leq t, t' \leq T \\ 0 \leq j, j' \leq h \\ t \neq t', j \neq j' \\ t-t'=j-j'}} u_t u_{t', a_h(j)} a_h(j'). \quad (7.4.9)$$

Como J é o quociente entre (7.4.8) e (7.4.9), obtemos:

$$J = \frac{2 \sum_{\substack{1 \leq t \leq T \\ 0 \leq j \leq h}} x_t u_t a_h^2(j) + \sum_{\substack{1 \leq t, t' \leq T \\ 0 \leq j, j' \leq h \\ t \neq t', j \neq j', t-t'=j-j'}} x_t u_{t', a_h(j)} a_h(j')}{2 \sum_{\substack{1 \leq t \leq T \\ 0 \leq j \leq h}} u_t^2 a_h^2(j) + \sum_{\substack{1 \leq t, t' \leq T \\ 0 \leq j, j' \leq h \\ t \neq t', j \neq j', t-t'=j-j'}} u_t u_{t', a_h(j)} a_h(j')} . \quad (7.4.10)$$

A expressão para J, dada em (7.4.10), sugere que a obtenção do vício e o EQM de $\hat{\beta}$, por essa via, deve ser bastante complicada, se não impossível. Uma alternativa seria determinar o limite em probabilidade do denominador de J e de J^2 , e procurar encontrar as esperanças dos respectivos numeradores, nos moldes do que fizemos no capítulo 4.

REFERÊNCIAS

- Akaike, H. (1969). Fitting autoregressive models for prediction. Ann. Inst. Statist. Math. 21, 243-247.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Second International Symposium on Information Theory. Ed. B.N. Petrov and F. Csáki, 267-281. Budapest: Akadémia Kiado.
- Akaike, H. (1977). On entropy maximization principle. In Applications of Statistics. Ed. P.R. Krishnaiah, 27-41. North Holland. Amsterdam.
- Amemiya, T. (1973). Generalized least squares with an estimated autocovariance matrix. Econometrica 41, 723-732.
- An, H.Z., Chen, Z.G. and Hannan, E.J. (1982). Autocorrelation, autoregression and autoregressive approximation. Ann. Statist. 10, 926-936.
- Berk, K. (1974). Consistent autoregressive spectral estimates. Ann. Statist. 2, 489-502.
- Bhansali, R.J. (1978). Linear prediction by autoregressive model fitting in the time domain. Ann. Statist. 6, 224-231.
- Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis, Forecasting and Control, revised edition, Holden-Day, San Francisco.

- Cochrane, D. and Orcutt, G.H. (1949). Application of least squares regression to relationships containing autocorrelated error terms. J. Amer. Statist. Assoc. 44, 32-61.
- Durbin, J. (1960). The fitting of time series models. Rev. Inst. Int. Stat. 28, 233-243.
- Engle, R.F. and Gardner, R. (1976). Some finite sample properties of spectral estimators of a linear regression, Econometrica 44, 149-166.
- Fuller, W.A. (1976). Introduction to Statistical Time Series, John Wiley, New York.
- Grenander, U. and Roseblatt, M. (1957). Statistical Analysis of Stationary Time Series, John Wiley, New York.
- Hannan, E.J. (1963). Regression for time series. (Proc. Symp. on) Time Series Analysis. Ed. M. Roseblatt, John Wiley, New York, 14-37.
- Hannan, E.J. (1970). Multiple Time Series, John Wiley, New York.
- Hannan, E.J. (1973). Central limit theorems for time series regression. Z. Wahrscheinlichkeitstheorie verw Geb. 26, 157-170.
- Hannan, E.J. (1980). The estimation of the order of an ARMA process. Ann. Statist. 8, 1071-1081.
- Hannan, E.J. and Kavalieris, L. (1983). The convergence of autocorrelations and autoregressions. Austral. J. Statist. 25, 287-297.

- Hannan, E.J. and Kavalieris, L. (1984a) Multivariate linear time series models. Adv. Appl. Prob. 16, 492-561.
- Hannan, E.J. and Kavalieris, L. (1984b) A method of autoregressive - moving average estimation. Biometrika 71, 273-280.
- Hannan, E.J. and Kavalieris, L. (1986). Regression, autoregression models. J. Time Series Anal. 7, 27-49.
- Hannan, E.J. and Quinn, B.G. (1979). The determination of the order of an autoregression. J. Roy Statist. Soc. B. 41, 190-195.
- Harvey, A.C. (1978). Linear regression on the frequency domain. International Economic Review 19, 507-512.
- Harvey, A.C. (1981a). The Econometric Analysis of Time Series. Philip Allan Publ.
- Harvey, A.C. (1981b). Time Series Models. Philip. Allan Publ.
- Ibragimov, I.A. and Linnik, Yu. V. (1971). Independent and Stationary Sequences of Random Variables. Wolters-Noordhoff, Groningen.
- Morettin, P.A. (1984). The Levinson algorithm and its applications in time series analysis. Intern. Statist. Rev. 52, 83-92.
- Pierce, D.A. (1971). Least squares estimation in the regression model with ARMA errors. Biometrika 58, 299-312.
- Pierce, D.A. (1980). Multivariate Time Series. A short course at "V ELAM", Argentina.

Priestley, M.B. (1981). Spectral Analysis and Time Series. Academic Press.

Rissanen, J. (1978). Modelling by shortest data description. Automatica 14, 465-471.

Schwarz, G. (1978). Estimating the dimension of a model. Ann. Statist. 6, 461-464.

Shibata, R. (1980). Asymptotically efficient selection of the order of the models for estimating parameters of a linear process. Ann. Statist. 8, 147-164.

Toyooka, T. (1982). Second-order expansion of mean squared error matrix of generalized least squares estimator with estimated parameters. Biometrika 69, 269-273.

Wickens, M.R. (1969). The consistency and efficiency of generalized least squares in simultaneous equation systems with autocorrelated errors. Econometrica 37, 651-659.