Departamento de Matemática
Faculdade de Ciências e Tecnologia · U C · Faculdade de Economia

Sobre a estimação de processos auto-regressivos de ordem um.

Ana Isabel de Sousa Quintas

Dissertação para a obtenção do Grau de

Mestre em Métodos Quantitativos em Finanças

Júri

Presidente:	Doutor Carlos Manuel Rebelo Tenreiro da Cruz
Co-Orientador:	Doutora Nazaré Mendes Lopes
Co-Orientador:	Doutora Esmeralda Gonçalves
Vogal:	Doutora Cristina Maria Tavares Martins

Data: 27 de Julho de 2012

Resumo

O objetivo da presente tese é realizar uma análise estatística sobre os processos auto-regressivos de ordem 1.

Após uma breve introdução teórica, começamos por estudar o processo X_t quando $t \in \mathbb{N}_0$ e com condição inicial $X_0 = 0$. Depois de provarmos que tal processo é de segunda ordem, analisamos a sua estacionaridade. Em seguida, determinamos o estimador da máxima verosimilhança para os parâmetros do modelo supondo que o ruído associado é normal, centrado e de variância σ^2, concluindo-se também que o estimador obtido para φ coincide com o estimador dos mínimos quadrados.

Posteriormente, estudamos o processo quando $t \in \mathbb{Z}$. Começamos por analisar a estacionaridade fraca, forte e a ergodicidade do processo. Finalmente determinamos o estimador da máxima verosimilhança para os respetivos parâmetros.

Em seguida, estudamos as propriedades assintóticas e a consistência dos estimadores de φ, obtidos para os dois processos.

Finalmente, determinamos a lei assintótica do estimador de φ no caso em que o processo é explosivo, isto é, quando $|\varphi| > 1$ e $X_0 = 0$ e ilustramos os resultados obtidos usando dados simulados e o teste de ajustamento do χ^2.

Palavras Chave: Processo auto-regressivo, Estimador da máxima verosimilhança,

Comportamento assintótico do estimador

Abstract

The aim of this thesis is to perform a statistical analysis on the first-order autoregressive processes.

After a brief introduction, we begin by studying the process X_t when $t \in \mathbb{N}_0$ and with initial condition $X_0 = 0$. Once you prove that this process is of second-order, we analyse its stationarity. Next, we determine the maximum likelihood estimator for the parameters of the model assuming that the associated noise is distributed normally with mean 0 and variance σ^2, concluding that the estimator obtained for φ coincides with the estimator of the least squares.

After, we study the case when $t \in \mathbb{Z}$. We begin by analysing the weak and strong stationary and ergodicity of the process. Finally we determine the maximum likelihood estimator for the respective parameters.

Then we study the asymptotic properties of estimators and consistency of φ, obtained for the two processes.

Finally, we determine the asymptotic law of the estimator of φ in the case where the process is explosive, i.e., when $\mid \varphi \mid > 1$ and $X_0 = 0$ and illustrate the results obtained using simulated data and adjustment test of χ^2.

Keywords: First-order autoregressive processes, Maximum likelihood estimator,

Asymptotic behaviour of the estimator

Agradecimentos

Aos meus pais, que sempre acreditaram no meu empenho e dedicação, pela amizade, motivação e carinho.

A todos os meus amigos pelo companheirismo e dedicação em todos os momentos.

Às minhas orientadoras Doutora Nazaré Mendes Lopes e Doutora Esmeralda Gonçalves pela paciência, dedicação e orientação.

Conteúdo

Capítulo 1

Introdução

O objetivo da presente tese é realizar uma análise estatística sobre os processos-auto-regressivos de ordem 1, AR(1). Em particular, serão obtidos os estimadores de máxima verosimilhança dos parâmetros do modelo e estudada, num caso a precisar, a correspondente lei assintótica.

Na primeira secção do segundo capítulo, após uma breve introdução teórica, começamos por estudar o processo $X = (X_t)$ seguindo um modelo AR(1) quando $t \in \mathbb{N}_0$ e com condição inicial $X_0 = 0$. Depois de provarmos que o processo X é de segunda ordem, analisamos a sua estacionaridade. Em seguida, deduzimos a lei do vetor $(X_1, \cdots, X_T)$, $T \in \mathbb{N}$, de forma a encontrar a função de verosimilhança do processo. Encontrada a função de verosimilhança associada a uma realização da amostra arbitrariamente fixa, $(x_1, \cdots, x_T)$, determinamos o estimador da máxima verosimilhança para o par de parâmetros do modelo em estudo, $\left(\varphi, \sigma^2\right)$, supondo que o ruido ε é tal que $\forall t \in \mathbb{N}, \varepsilon_t \sim N(0, \sigma)$, $\sigma > 0$. Finalmente concluímos que o estimador obtido para φ coincide com o estimador dos mínimos quadrados.

Na segunda secção do segundo capítulo, estudamos o processo $X = (X_t)$ seguindo um modelo AR(1) quando $t \in \mathbb{Z}$. Começamos por estudar a sua estacionaridade no sentido fraco considerando três casos distintos, nomeadamente $\mid \varphi \mid < 1$, $\mid \varphi \mid > 1$ e $\varphi = \pm 1$. Em seguida, analisamos a estacionaridade forte e a ergodicidade do processo. Finalmente, deduzimos a lei do vetor $(X_1, \cdots, X_T)$, $T \in \mathbb{N}$, de forma a encontrar a função de verosimilhança do processo quando $\mid \varphi \mid < 1$, também sob a hipótese de normalidade de ε_t, o que permitirá determinar o estimador da máxima verosimilhança para os parâmetros do modelo.

No terceiro capítulo, começamos por provar que, quando $\mid \varphi \mid < 1$, os estimadores da máxima verosimilhança dos parâmetros φ e σ associados aos processos em estudo na primeira secção e na segunda secção do segundo capítulo são assintóticamente equivalentes (quando $T \to +\infty$) em probabilidade. Finalmente, provamos que tais estimadores são consistentes em probabilidade.

No quarto capítulo, determinamos a lei assintótica do estimador de φ no caso em que o processo AR(1) é explosivo, isto é, quando $|\varphi| > 1$ e $X_0 = 0$. Limitamos o estudo a este caso por se revelar diferente dos casos habitualmente encontrados em que a lei limite é frequentemente gaussiana. Provamos, desse modo, que a distribuição limite do estimador devidamente normalizado é a distribuição de Cauchy standard. Finalmente, ilustramos os resultados obtidos usando dados simulados e o teste de ajustamento do χ^2.

Capítulo 2

Estimação de processos auto-regressivos de ordem 1

Seja $\mathcal{T}$ um conjunto de índices.

Um processo estocástico real $X = (X_t, t \in \mathcal{T})$ admite uma representação auto-regressiva de ordem 1, denotada AR(1), se existem $\varphi \in \mathbb{R} \setminus \{0\}$ e uma família de variáveis aleatórias reais, $\varepsilon = (\varepsilon_t, t \in \mathcal{T})$, centradas, não correlacionadas e de variância $\sigma^2 (\sigma^2 > 0)$, tais que

$$X_t = \varphi X_{t-1} + \varepsilon_t. \qquad \textbf{(1)}$$

Se introduzirmos o operador de atraso L associado a X, tal que $LX_t = X_{t-1}$, e designarmos o polinómio auto-regressivo por $\Phi(L) = 1 - \varphi L$, a relação de recorrência presente na definição anterior escreve-se na forma

$$\Phi(L)X_t = \varepsilon_t.$$

2.1. Modelo com espaço de tempos $\mathcal{T} = \mathbb{N}_0$ e condição inicial $X_0 = 0$

Nesta seção iremos estudar o processo AR(1) definido anteriormente quando $\mathcal{T} = \mathbb{N}_0$ e com a condição inicial $X_0 = 0$.

Assim consideremos

$$\begin{cases} X_t = \varphi X_{t-1} + \varepsilon_t, & t \in \mathbb{N} \\ X_0 = 0 \end{cases} . \qquad \textbf{(A)}$$

Começaremos por provar que o processo X é de segunda ordem e em seguida estudaremos a sua estacionaridade.

Após uma breve introdução teórica sobre o Método da máxima verosimilhança, iremos deduzir a lei do vetor $(X_1, \cdots, X_T), T \in \mathbb{N}$, de forma a encontrar a função de verosimilhança do processo. Encontrada a função de verosimilhança associada a uma realização da amostra arbitrariamente fixa, $(x_1, \cdots, x_T)$, determinaremos o estimador da máxima verosimilhança para o par (φ, σ^2) do modelo em estudo supondo que $\forall t \in \mathbb{N}, \varepsilon_t \sim N(0, \sigma), \sigma > 0$. Posteriormente veremos que o estimador obtido para φ coincide com o estimador dos mínimos quadrados.

2.1.1. Análise da estacionaridade

Dados $\varepsilon = (\varepsilon_t, t \in \mathbb{N}), \varphi$ e fixando $X_0 = 0$, é possível obter todos os valores do processo X. De facto, tem-se

$$\begin{aligned}
X_1 &= \varepsilon_1 \\
X_2 &= \varphi X_1 + \varepsilon_2 = \varphi \varepsilon_1 + \varepsilon_2 \\
X_3 &= \varphi X_2 + \varepsilon_3 = \varphi^2 \varepsilon_1 + \varphi \varepsilon_2 + \varepsilon_3 \\
&\vdots \\
X_t &= \varphi^{t-1} \varepsilon_1 + \varphi^{t-2} \varepsilon_2 + \cdots + \varepsilon_t \\
&= \sum_{j=0}^{t-1} \varphi^j \varepsilon_{t-j}, \quad \forall t \in \mathbb{N}.
\end{aligned}$$

Provemos que o processo $(X_t, t \in \mathbb{N}_0)$ apresentado é de segunda ordem. Temos

$$\begin{aligned}
E(X_t^2) &= E\left(\sum_{j=0}^{t-1} \varphi^j \varepsilon_{t-j} \sum_{k=0}^{t-1} \varphi^k \varepsilon_{t-k} \right) \\
&= \sum_{j=0}^{t-1} \varphi^{2j} E\left(\varepsilon_{t-j}^2\right),
\end{aligned}$$

pois $E(\varepsilon_{t-j} \varepsilon_{t-k}) \neq 0$ se e só se $j = k$.

Assim,

$$E(X_t^2) = \frac{1 - \varphi^{2t}}{1 - \varphi^2} \sigma^2, \text{ se } \mid \varphi \mid \neq 1$$

e

$$E(X_t^2) = t\sigma^2, \text{ se } \mid \varphi \mid = 1.$$

Como $E(X_t^2) < +\infty, \forall t \in \mathbb{N}_0$, concluímos que o processo $X = (X_t, t \in \mathbb{N}_0)$ é de segunda ordem, qualquer que seja $\varphi \in \mathbb{R} \setminus \{0\}$.

Como $E(X_t^2)$ depende de t, podemos concluir de imediato que o processo não é estacionário no sentido fraco ou forte.

Sendo X de segunda ordem, determinemos em seguida, os seus momentos de 1ª e 2ª ordem.

$$\begin{aligned} E(X_t) &= E\left(\sum_{j=0}^{t-1} \varphi^j \varepsilon_{t-j}\right) \\ &= \sum_{j=0}^{t-1} \varphi^j E\left(\varepsilon_{t-j}\right) \\ &= 0, \quad \forall t \in \mathbb{N}_0, \end{aligned}$$

e

$$\begin{aligned} Cov(X_t, X_s) &= E(X_t X_s) - \underbrace{E(X_t)}_{=0}\underbrace{E(X_s)}_{=0} \\ &= E\left(\sum_{j=0}^{t-1} \varphi^j \varepsilon_{t-j} \sum_{k=0}^{s-1} \varphi^k \varepsilon_{s-k}\right) \\ &= \sum_{j=0}^{t-1}\sum_{k=0}^{s-1} \varphi^{j+k} E\left(\varepsilon_{t-j}\varepsilon_{s-k}\right). \end{aligned}$$

Consideremos $t \geq s$.

$$E\left(\varepsilon_{t-j}\varepsilon_{s-k}\right) = \begin{cases} \sigma^2, & t-j = s-k \\ 0, & \text{se não} \end{cases} = \begin{cases} \sigma^2, & j = t-s+k \\ 0, & \text{se não} \end{cases}$$

Então

$$\begin{aligned} Cov(X_t, X_s) &= \sum_{k=0}^{s-1} \varphi^{t-s+k+k} E\left(\varepsilon_{s-k}^2\right) \\ &= \varphi^{t-s}\sigma^2 \sum_{k=0}^{s-1} \varphi^{2k}. \end{aligned}$$

Se $t < s$ obtemos, de modo análogo,

$$\begin{aligned} Cov(X_t, X_s) &= \sum_{j=0}^{t-1} \varphi^{j+s-t+j} E\left(\varepsilon_{t-j}^2\right) \\ &= \varphi^{s-t}\sigma^2 \sum_{j=0}^{t-1} \varphi^{2j}. \end{aligned}$$

Assim $\forall t, s \in \mathbb{N}_0$,

$$Cov(X_t, X_s) = \varphi^{|t-s|}\sigma^2 \frac{1-\varphi^{2\min(t,s)}}{1-\varphi^2}, \text{ se } \mid \varphi \mid \neq 1$$

e

$$Cov(X_t, X_s) = \varphi^{|t-s|} \min(t,s)\sigma^2, \text{ se } \mid \varphi \mid = 1.$$

Em particular, reencontramos

$$V(X_t) = \begin{cases} \sigma^2 \dfrac{1-\varphi^{2t}}{1-\varphi^2}, & \mid \varphi \mid \neq 1 \\ \\ t\sigma^2, & \mid \varphi \mid = 1 \end{cases}.$$

Note-se que $Cov(X_t, X_s)$ não depende apenas da diferença dos índices, $t-s$, o que vem ao encontro da não estacionaridade em sentido fraco de X qualquer que seja $\varphi \in \mathbb{R} \setminus \{0\}$.

Notemos que, quando t tende para $+\infty$ e $\mid \varphi \mid < 1$, $V(X_t)$ e $Cov(X_t, X_{t+h})$ convergem, respetivamente, para $\dfrac{\sigma^2}{1-\varphi^2}$ e $\dfrac{\sigma^2\varphi^{|h|}}{1-\varphi^2}$.

De facto,

$$\lim_{t\to+\infty} V(X_t) = \lim_{t\to+\infty} \frac{1-\varphi^{2t}}{1-\varphi^2}\sigma^2 = \frac{\sigma^2}{1-\varphi^2}$$

e

$$\lim_{t\to+\infty} Cov(X_t, X_{t+h}) = \frac{\sigma^2\varphi^{|h|}}{1-\varphi^2}.$$

Em contrapartida, quando $\mid \varphi \mid > 1$, a variância cresce exponencialmente. Este facto leva a designar tais processos por explosivos.

2.1.2. Estimação dos parâmetros do modelo

Estimadores da máxima verosimilhança

Iniciamos este parágrafo com uma breve revisão sobre o Método da máxima verosimilhança.

Consideremos uma amostra $(X_1, \cdots, X_T)$ de dimensão T, $T \in \mathbb{N}$, associada a um processo estocástico X estacionário e absolutamente contínuo.

Supomos que a lei do vector $(X_1, \cdots, X_T)$ depende de um parâmetro θ, $\theta \in \Theta \subset \mathbb{R}^k$, e admite densidade $f_{(X_1,\cdots,X_T)}(\cdot, \theta)$[1]. Seja $(x_1, \cdots, x_T)$ uma realização arbitrariamente fixa da amostra.

A função de verosimilhança associada à realização $(x_1, \cdots, x_T)$ é a função definida sobre Θ por

$$L\left((x_1, \cdots, x_T); \theta\right) = f_{(X_1,\cdots,X_T)}\left((x_1, \cdots, x_T); \theta\right).$$

Um estimador $\theta_T^* = \theta_T^*(X_1, \cdots, X_T)$ de θ diz-se *estimador da máxima verosimilhança* (e.m.v.) de θ se

$$L(x, \theta_T^*(x)) = \sup_{\theta \in \Theta} L(x, \theta), \quad \forall x = (x_1, \cdots, x_T) \in \mathbb{R}^T.$$

Mediante certas condições de regularidade da função de verosimilhança, podemos utilizar uma metodologia que simplifica bastante a obtenção de um seu maximizante. Tal metodologia é geral e pode resumir-se do seguinte modo:

Seja Θ um aberto de $\mathbb{R}^k$ e suponha-se que para todo o $x \in \mathbb{R}^T, L(x, .)$ é uma função definida em Θ duas vezes derivável. Nestas condições, um estimador da máxima verosimilhança para $\theta = (\theta_1, \cdots, \theta_k)$ é solução do sistema

$$\forall i \in \{1, 2, \cdots, k\}, \frac{\partial}{\partial \theta_i} L(x, \theta) = 0. \qquad \textbf{(1.1)}$$

Suponha-se ainda que

$$\forall \theta \in \Theta, \forall x \in \mathbb{R}^T, L(x, \theta) = L_1(x, \theta) \mathbb{1}_A(x)$$

onde A é um subconjunto de $\mathbb{R}^T$ independente de Θ e L_1 uma função estritamente positiva. Como a função logaritmo é estritamente crescente, um estimador de máxima verosimilhança é então solução do novo sistema

$$\forall i \in \{1, 2, \cdots, k\}, \frac{\partial}{\partial \theta_i} \log L_1(x, \theta) = 0 \qquad \textbf{(1.2)}$$

[1] Limitamo-nos a este caso por ser o que interessa no contexto em estudo.

com x arbitrariamente fixo em A.

Observe-se finalmente que uma solução de **(1.1)** (resp. **(1.2)**) não é necessariamente um maximizante de $L(x,.)$(resp. $L_1(x,.)$). Há assim necessidade de, em qualquer dos casos, analisar as condições usuais de $2^{\underline{a}}$ ordem associadas ao estudo de extremos de funções.

Passemos então à determinação do e.m.v. do parâmetro (φ, σ^2) do modelo em estudo supondo que $\forall t \in \mathbb{N}, \varepsilon_t$ segue a lei Normal centrada de variância σ^2, $N(0,\sigma)$, $\sigma > 0$.

Torna-se então necessário deduzir a lei do vector $(X_1, \cdots, X_T), T \in \mathbb{N}$, de modo a encontrar a função de verosimilhança do processo.

Vamos começar por mostrar que $X_t = \sum\limits_{k=0}^{t-1} \varphi^k \varepsilon_{t-k}$ segue uma lei Normal, $\forall t \in \{1, \cdots, T\}$.

Seja Φ_{X_t} a função característica de X_t. Temos

$$\forall u \in \mathbb{R}, \qquad \Phi_{X_t}(u) = \Phi_{\sum\limits_{k=0}^{t-1} \varphi^k \varepsilon_{t-k}}(u) = \prod_{k=0}^{t-1} \Phi_{\left(\varphi^k \varepsilon_{t-k}\right)}(u)$$

pois $\varepsilon_1, \cdots, \varepsilon_t$ são independentes.

Então,

$$\Phi_{X_t}(u) = \prod_{k=0}^{t-1} \Phi_{\varepsilon_t}\left(\varphi^k u\right) = \prod_{k=0}^{t-1} \exp\left(i\varphi^k u 0\right) \exp\left(-\frac{\sigma^2 \varphi^{2k} u^2}{2}\right)$$

pois se $U \sim N(m,\sigma)$, então $\Phi_U(u) = \exp(imu) \exp\left(-\frac{u^2\sigma^2}{2}\right), \quad u \in \mathbb{R}$.

Obtemos então

$$\Phi_{X_t}(u) = \exp\left(-\frac{u^2}{2}\sigma^2 \sum_{k=0}^{t-1} \varphi^{2k}\right) = \begin{cases} \exp\left(-\frac{u^2}{2}\sigma^2 t\right), & \text{se } \mid \varphi \mid = 1 \\ \\ \exp\left(-\frac{u^2}{2}\sigma^2 \frac{1-\varphi^{2t}}{1-\varphi^2}\right), & \text{se } \mid \varphi \mid \neq 1 \end{cases}.$$

Como a função característica determina a lei, concluímos que, $\forall t \in \mathbb{N}$,

$$X_t \sim N\left(0, \sigma\sqrt{t}\right), \quad \text{se } \mid \varphi \mid = 1$$

e

$$X_t \sim N\left(0, \sigma\sqrt{\frac{1-\varphi^{2t}}{1-\varphi^2}}\right), \quad \text{se } \mid \varphi \mid \neq 1.$$

Para determinar a lei do vector $(X_1, \cdots, X_T)$ vamos atender a que $(X_1, \cdots, X_T) = H(\varepsilon_1, \cdots, \varepsilon_T)$ com $H : \mathbb{R}^T \to \mathbb{R}^T$ tal que

$$\begin{cases} X_1 = \varepsilon_1 \\ X_2 = \varphi X_1 + \varepsilon_2 = \varphi \varepsilon_1 + \varepsilon_2 \\ \vdots \\ X_T = \varphi X_{T-1} + \varepsilon_T \end{cases} \Leftrightarrow \begin{cases} \varepsilon_1 = X_1 \\ \varepsilon_2 = X_2 - \varphi X_1 \\ \vdots \\ \varepsilon_T = X_T - \varphi X_{T-1} \end{cases}$$

pelo que

$$H^{-1}(x_1, \cdots, x_T) = (x_1, x_2 - \varphi x_1, \cdots, x_T - \varphi x_{T-1}) .$$

O Jacobiano da transformação inversa é

$$J = det \left[\frac{\partial H^{-1}(x_1, \cdots, x_T)}{\partial x_j}\right]_{j=1,\cdots,T} = \begin{vmatrix} 1 & -\varphi & 0 & \cdots & 0 \\ 0 & 1 & -\varphi & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & -\varphi \\ 0 & 0 & & & 1 \end{vmatrix} = 1.$$

Pelo Teorema da Mudança de Variável para vectores aleatórios reais com densidade vem

$$\begin{aligned} f_{(X_1,\cdots,X_T)}(x_1, \cdots, x_T) &= f_{(\varepsilon_1,\cdots,\varepsilon_T)}\left(H^{-1}(x_1, \cdots, x_T)\right) \mid J \mid \\ &= f_{(\varepsilon_1,\cdots,\varepsilon_T)}\left(x_1, x_2 - \varphi x_1, \cdots, x_T - \varphi x_{T-1}\right) . \end{aligned}$$

Como $\varepsilon_1, \cdots, \varepsilon_T$ são variáveis aleatórias reais independentes e identicamente distribuídas (v.a.r.i.i.d) e admitem densidades $f_1, f_2, \cdots, f_T$, respetivamente, então $(\varepsilon_1, \cdots, \varepsilon_T)$ é um vector aleatório sobre $\mathbb{R}^T$ absolutamente contínuo de densidade

$$\begin{aligned} f_{(\varepsilon_1,\cdots,\varepsilon_T)}(x_1, \cdots, x_T) &= \prod_{i=1}^{T} f_i(x_i) \\ &= \prod_{i=1}^{T} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2} x_i^2\right), \text{ pois } \quad \varepsilon_t \sim N(0, \sigma) \\ &= \frac{1}{\sigma^T(\sqrt{2\pi})^T} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{T} x_i^2\right). \end{aligned}$$

Então

$$\begin{aligned} f_{(X_1,\cdots,X_T)}(x_1,\cdots,x_T) &= \frac{1}{\sigma^T(\sqrt{2\pi})^T}\exp\left\{-\frac{1}{2\sigma^2}\left(x_1^2+(x_2-\varphi x_1)^2+\cdots+(x_T-\varphi x_{T-1})^2\right)\right\} \\ &= \frac{1}{\sigma^T(\sqrt{2\pi})^T}\exp\left\{-\frac{1}{2\sigma^2}\left(x_1^2+\sum_{i=2}^{T}(x_i-\varphi x_{i-1})^2\right)\right\}. \end{aligned}$$

Assim, a função de verosimilhança do processo associada à realização da amostra $x=(x_1,\cdots,x_T)$ é

$$L((x_1,\cdots,x_T),\theta)=\frac{1}{(\sigma^2)^{T/2}(2\pi)^{T/2}}\exp\left\{-\frac{1}{2\sigma^2}\left(x_1^2+\sum_{i=2}^{T}(x_i-\varphi x_{i-1})^2\right)\right\}$$

onde $\theta=(\varphi,\sigma^2)$.

Como a função de verosimilhança, L, que estamos a considerar coincide com L_1, é estritamente positiva e possui derivadas de 2ª ordem então um possível e.m.v. do par (φ,σ^2) é solução de

$$\begin{cases} \dfrac{\partial \log L\left(x,(\varphi,\sigma^2)\right)}{\partial\varphi}=0 \\ \\ \dfrac{\partial \log L\left(x,(\varphi,\sigma^2)\right)}{\partial\sigma^2}=0 \end{cases} \qquad \textbf{(1.3)}$$

com

$$\log L=-\frac{T}{2}\log(2\pi)-\frac{T}{2}\log(\sigma^2)-\frac{1}{2\sigma^2}\left[x_1^2+\sum_{i=2}^{T}(x_i-\varphi x_{i-1})^2\right].$$

Ora

$$\frac{\partial \log L\left(x,(\varphi,\sigma^2)\right)}{\partial\varphi}=\frac{2}{2\sigma^2}\sum_{i=2}^{T}(x_i-\varphi x_{i-1})\,x_{i-1}=\frac{1}{\sigma^2}\sum_{i=2}^{T}\left(x_ix_{i-1}-\varphi x_{i-1}^2\right)$$

$$\frac{\partial \log L\left(x,(\varphi,\sigma^2)\right)}{\partial\sigma^2}=-\frac{T}{2\sigma^2}+\frac{x_1^2+\sum_{i=2}^{T}(x_i-\varphi x_{i-1})^2}{2\sigma^4}.$$

Assim, o sistema **(1.3)** é equivalente a

$$\begin{cases} \sum_{i=2}^{T}\left(x_ix_{i-1}-\varphi x_{i-1}^2\right)=0 \\ \\ -\dfrac{T}{2\sigma^2}+\dfrac{x_1^2+\sum_{i=2}^{T}(x_i-\varphi x_{i-1})^2}{2\sigma^4}=0 \end{cases} \Leftrightarrow \begin{cases} \varphi\sum_{i=2}^{T}x_{i-1}^2-\sum_{i=2}^{T}x_ix_{i-1}=0 \\ \\ \sigma^2=\dfrac{x_1^2+\sum_{i=2}^{T}(x_i-\varphi x_{i-1})^2}{T} \end{cases}$$

$$\Leftrightarrow \begin{cases} \varphi = \dfrac{\sum\limits_{i=2}^{T} x_i x_{i-1}}{\sum\limits_{i=2}^{T} x_{i-1}^2} \\ \sigma^2 = \dfrac{x_1^2 + \sum\limits_{i=2}^{T} (x_i - \varphi x_{i-1})^2}{T} \end{cases}.$$

Analisemos as condições de 2ª ordem.

$$\begin{cases} \dfrac{\partial^2 \log L\big(x, (\varphi, \sigma^2)\big)}{\partial \varphi^2} = -\dfrac{\sum\limits_{i=2}^{T} x_{i-1}^2}{\sigma^2} \\ \dfrac{\partial^2 \log L\big(x, (\varphi, \sigma^2)\big)}{\partial (\sigma^2)^2} = \dfrac{T}{2\sigma^4} - \dfrac{x_1^2 + \sum\limits_{i=2}^{T} (x_i - \varphi x_{i-1})^2}{\sigma^6} \\ \dfrac{\partial^2 \log L\big(x, (\varphi, \sigma^2)\big)}{\partial \varphi \partial \sigma^2} = \dfrac{\partial^2 \log L\big(x, (\varphi, \sigma^2)\big)}{\partial \sigma^2 \partial \varphi} = -\dfrac{1}{\sigma^4} \sum\limits_{i=2}^{T} (x_i x_{i-1} - \varphi x_{i-1}^2) \end{cases}.$$

Substituindo os valores de φ e de σ^2 pelos provenientes das condições de 1ª ordem, respetivamente $\hat{\varphi}$ e $\hat{\sigma}^2$, obtemos

$$\begin{cases} \dfrac{\partial^2 \log L\big(x, (\hat{\varphi}, \hat{\sigma}^2)\big)}{\partial \varphi^2} = -\dfrac{\sum\limits_{i=2}^{T} x_{i-1}^2}{\hat{\sigma}^2} \\ \dfrac{\partial^2 \log L\big(x, (\hat{\varphi}, \hat{\sigma}^2)\big)}{\partial (\sigma^2)^2} = \dfrac{T}{2\hat{\sigma}^4} - \dfrac{T\hat{\sigma}^2}{\hat{\sigma}^6} = -\dfrac{T}{2\hat{\sigma}^4} \\ \dfrac{\partial^2 \log L\big(x, (\hat{\varphi}, \hat{\sigma}^2)\big)}{\partial \varphi \partial \sigma^2} = \dfrac{\partial^2 \log L\big(x, (\hat{\varphi}, \hat{\sigma}^2)\big)}{\partial \sigma^2 \partial \varphi} = -\dfrac{1}{\hat{\sigma}^4} \left[\sum\limits_{i=2}^{T} x_i x_{i-1} - \dfrac{\sum\limits_{i=2}^{T} x_i x_{i-1}}{\sum\limits_{i=2}^{T} x_{i-1}^2} \sum\limits_{i=2}^{T} x_{i-1}^2 \right] = 0 \end{cases}.$$

Então a matriz relativa as condições de 2ª ordem é

$$\begin{bmatrix} \dfrac{\partial^2 \log L\big(x, (\varphi, \sigma^2)\big)}{\partial \varphi^2} & \dfrac{\partial^2 \log L\big(x, (\varphi, \sigma^2)\big)}{\partial \varphi \partial \sigma^2} \\ \dfrac{\partial^2 \log L\big(x, (\varphi, \sigma^2)\big)}{\partial \sigma^2 \partial \varphi} & \dfrac{\partial^2 \log L\big(x, (\varphi, \sigma^2)\big)}{\partial (\sigma^2)^2} \end{bmatrix}_{[\hat{\varphi}_T, \hat{\sigma}_T^2]} = \begin{bmatrix} -\dfrac{\sum\limits_{i=1}^{T} x_{i-1}^2}{\hat{\sigma}^2} & 0 \\ 0 & -\dfrac{T}{2\hat{\sigma}^4} \end{bmatrix}$$

que é definida negativa, pois os seus valores próprios são negativos.

Assim, concluímos que

$$\begin{cases} \hat{\varphi}_T = \dfrac{\sum\limits_{i=2}^{T} X_i X_{i-1}}{\sum\limits_{i=2}^{T} X_{i-1}^2} \\ \\ \hat{\sigma}_T^2 = \dfrac{X_1^2 + \sum\limits_{i=2}^{T} \left(X_i - \hat{\varphi}_T X_{i-1}\right)^2}{T} \end{cases}$$

é um estimador da máxima verosimilhança para (φ, σ^2).

Como $X_0 = 0$, os estimadores obtidos podem também escrever-se na forma

$$\begin{cases} \hat{\varphi}_T = \dfrac{\sum\limits_{i=1}^{T} X_i X_{i-1}}{\sum\limits_{i=1}^{T} X_{i-1}^2} \\ \\ \hat{\sigma}_T^2 = \dfrac{\sum\limits_{i=1}^{T} \left(X_i - \hat{\varphi}_T X_{i-1}\right)^2}{T} \end{cases}.$$

Note-se que o estimador obtido para φ coincide com o estimador obtido pelo Método dos mínimos quadrados. De facto, de acordo com este método, queremos determinar φ, $\varphi \in \mathbb{R}$, que minimize a função:

$$Q(\varphi) = \sum_{i=1}^{T} \left(x_i - \varphi x_{i-1}\right)^2 = x_1^2 + \sum_{i=2}^{T} \left(x_i - \varphi x_{i-1}\right)^2.$$

Temos

$$\frac{\partial Q(\varphi)}{\partial \varphi} = -2 \sum_{i=2}^{T} \left(x_i - \varphi x_{i-1}\right) x_{i-1}.$$

Assim,

$$\begin{aligned} \frac{\partial Q(\varphi)}{\partial \varphi} = 0 &\Longleftrightarrow \sum_{i=2}^{T} \left(x_i - \varphi x_{i-1}\right) x_{i-1} = 0 \\ &\Longleftrightarrow \sum_{i=2}^{T} x_i x_{i-1} - \varphi \sum_{i=2}^{T} x_{i-1}^2 = 0 \\ &\Longleftrightarrow \varphi = \frac{\sum\limits_{i=2}^{T} x_i x_{i-1}}{\sum\limits_{i=2}^{T} x_{i-1}^2}. \end{aligned}$$

Basta então provar que $\hat{\varphi}$ é um minimizante de $Q(\varphi)$. Ora,

$$\frac{\partial^2 Q(\hat{\varphi})}{\partial \varphi^2} = 2\sum_{i=2}^{T} x_{i-1}^2 > 0$$

portanto $\hat{\varphi}$ é um minimizante de $Q(\varphi)$ e, como tal, concluímos que o estimador dos mínimos quadrados (e.m.q.) de φ é

$$\hat{\varphi}_T = \frac{\sum\limits_{i=2}^{T} X_i X_{i-1}}{\sum\limits_{i=2}^{T} X_{i-1}^2}.$$

2.2. Modelo com o espaço de tempos $\mathcal{T} = \mathbb{Z}$

Nesta seção iremos estudar o processo X seguindo um modelo AR(1) apresentado em **(1)** quando $\mathcal{T} = \mathbb{Z}$.

Assim consideremos

$$X_t = \varphi X_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}. \qquad \textbf{(B)}$$

Começaremos por estudar a sua estacionaridade considerando três casos distintos, nomeadamente $\mid \varphi \mid < 1$, $\mid \varphi \mid > 1$ e $\varphi = \pm 1$.

Posteriormente, iremos deduzir a lei do vetor $(X_1, \cdots, X_T), T \in \mathbb{N}$, de forma a encontrar a função de verosimilhança do processo quando $\mid \varphi \mid < 1$, supondo que $\forall t \in \mathbb{Z}$, $\varepsilon_t \sim N(0, \sigma)$, $\sigma > 0$.
Encontrada a função de verosimilhança associada a uma realização da amostra arbitrariamente fixa, $(x_1, \cdots, x_T)$, determinaremos o estimador da máxima verosimilhança para o par (φ, σ^2) do modelo.

2.2.1. Estacionaridade fraca

Dados $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ e φ, é possível escrever o processo estocástico X na seguinte forma

$$\begin{aligned} X_t &= \varphi X_{t-1} + \varepsilon_t \\ &= \varphi\left(\varphi X_{t-2} + \varepsilon_{t-1}\right) + \varepsilon_t \\ &= \cdots \\ &= \varepsilon_t + \varphi\varepsilon_{t-1} + \cdots + \varphi^n \varepsilon_{t-n} + \varphi^{n+1} X_{t-n-1} \\ &= \sum_{i=0}^{n} \varphi^i \varepsilon_{t-i} + \varphi^{n+1} X_{t-n-1}, \forall n \quad \textbf{(2.2.1)} \end{aligned}$$

Consideremos $\mid \varphi \mid < 1$.

Comecemos por supor que existe uma solução estacionária e verifica **(2.2.1)**. Vejamos que $\sum\limits_{i=0}^{n} \varphi^i \varepsilon_{t-i}$ tende (em L_2) para $\sum\limits_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i}$, quando $n \to +\infty$.

$$\begin{aligned} \left\| \sum_{i=0}^{n} \varphi^i \varepsilon_{t-i} - \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i} \right\| &= \left\| \sum_{i=n+1}^{+\infty} \varphi^i \varepsilon_{t-i} \right\| \\ &\leq \sum_{i=n+1}^{+\infty} \mid \varphi^i \mid \|\varepsilon_{t-i}\| \\ &= \sum_{i=n+1}^{+\infty} \mid \varphi^i \mid \left[E\left(\varepsilon_{t-i}^2\right)\right]^{1/2} \\ &= \sigma \sum_{i=n+1}^{+\infty} \mid \varphi^i \mid \end{aligned}$$

que tende para zero, quando $n \to +\infty$, por ser o resto de ordem $n+1$ de uma série convergente (pois $\sum\limits_{i=0}^{+\infty} \mid \varphi \mid^i$ é uma série geométrica de razão $|\varphi| < 1$).

A parcela $\varphi^{n+1} X_{t-n-1}$ tende (em L^2) para zero, quando $n \to +\infty$. De facto,

$$\begin{aligned} \left\|\varphi^{n+1} X_{t-n-1}\right\| &= \left[E\left(\varphi^{n+1} X_{t-n-1}\right)^2\right]^{1/2} \\ &= \mid \varphi \mid^{n+1} \left[E(X_{t-n-1}^2)\right]^{1/2} \underset{n \to +\infty}{\longrightarrow} 0 \end{aligned}$$

uma vez que sendo X estacionário, se tem $\forall t \in \mathbb{Z}, E(X_t^2) = \text{constante} < +\infty$ e $|\varphi| < 1$.

Assim, um processo estacionário solução do modelo **(B)** é necessariamente da forma

$$X_t = \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i}.$$

Inversamente mostremos agora que se X admite esta representação média móvel infinita ($MA(\infty)$), então X é estacionário.

É claro que X é de 2ª ordem.

Determinemos $E(X_t), t \in \mathbb{Z}$.

Como $\sum_{i=0}^{n} \varphi^i \varepsilon_{t-i}$ tende (em L_2) para $\sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i}$ então,

$$\lim_{n \to +\infty} E\left(\left| \sum_{i=0}^{n} \varphi^i \varepsilon_{t-i} - \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i} \right|^2 \right) = 0.$$

Mas, da desigualdade de Cauchy-Schwartz, [1]

$$\left| E\left(\sum_{i=0}^{n} \varphi^i \varepsilon_{t-i} - \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i} \right) \right| \leq \left[E\left(\left| \sum_{i=0}^{n} \varphi^i \varepsilon_{t-i} - \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i} \right|^2 \right) \right]^{1/2},$$

então

$$\lim_{n \to +\infty} \left| E\left(\sum_{i=0}^{n} \varphi^i \varepsilon_{t-i} - \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i} \right) \right| = 0.$$

Assim,

$$\lim_{n \to +\infty} E\left(\sum_{i=0}^{n} \varphi^i \varepsilon_{t-i} - \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i} \right) = 0 \Leftrightarrow \lim_{n \to +\infty} E\left(\sum_{i=0}^{n} \varphi^i \varepsilon_{t-i} \right) = E\left(\sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i} \right).$$

Mas,

$$\lim_{n \to +\infty} \sum_{i=0}^{n} \varphi^i E\left(\varepsilon_{t-i}\right) = 0$$

e então

$$E\left(\sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i} \right) = 0$$

isto é, $E(X_t) = 0, \forall t$.

Provemos agora que $Cov(X_t, X_{t+h})$ depende apenas de h.

Como o processo X é centrado, $Cov(X_t, X_{t+h}) = E(X_t X_{t+h})$.

Seja $S_{n,t} = \sum_{i=0}^{n} \varphi^i \varepsilon_{t-i}$.

Ora

$$\begin{aligned} \left|E\left(S_{n,t} S_{m,t+h} - X_t X_{t+h}\right)\right| &= \left|E\left[S_{n,t}\left(S_{m,t+h} - X_{t+h}\right) + X_{t+h}\left(S_{n,t} - X_t\right)\right]\right| \\ &\leq \left|E\left[S_{n,t}\left(S_{m,t+h} - X_{t+h}\right)\right]\right| + \left|E\left[X_{t+h}\left(S_{n,t} - X_t\right)\right]\right| \end{aligned}$$

[1] Desigualdade de Cauchy-Schwartz: Tem-se $|< x, y >| \leq \| x \| \| y \|, \forall x, y \in H$ onde H é um espaço vetorial normado com produto interno representado por $< x, y >$.

Então, pela desigualdade de Cauchy-Schwartz,

$$\begin{aligned}|E\left(S_{n,t}S_{m,t+h}-X_tX_{t+h}\right)| \leq & \left[E\left(S_{n,t}\right)^2\right]^{1/2}\left[E\left(S_{m,t+h}-X_{t+h}\right)^2\right]^{1/2} \\ & +\left[E\left(X_{t+h}\right)^2\right]^{1/2}\left[E\left(S_{n,t}-X_t\right)^2\right]^{1/2}.\end{aligned}$$

Como já foi provado,

$$S_{n,t} \underset{n\to+\infty}{\overset{L^2}{\longrightarrow}} X_t \text{ e } S_{m,t+h} \underset{m\to+\infty}{\overset{L^2}{\longrightarrow}} X_{t+h},$$

ou seja

$$\left[E\left(S_{n,t}-X_t\right)^2\right]^{1/2} \underset{n\to+\infty}{\longrightarrow} 0 \text{ e } \left[E\left(S_{m,t+h}-X_{t+h}\right)^2\right]^{1/2} \underset{m\to+\infty}{\longrightarrow} 0$$

e $E\left(S_{n,t}\right)^2<+\infty$ pois X_t é de ordem, então

$$\lim_{\substack{n\to+\infty\\ m\to+\infty}} |E\left(S_{n,t}S_{m,t+h}-X_tX_{t+h}\right)|=0.$$

Assim,

$$\lim_{\substack{n\to+\infty\\ m\to+\infty}} E\left(S_{n,t}S_{m,t+h}-X_tX_{t+h}\right)=0 \Leftrightarrow \lim_{\substack{n\to+\infty\\ m\to+\infty}} E\left(S_{n,t}S_{m,t+h}\right)=E\left(X_tX_{t+h}\right)$$

ou seja,

$$\lim_{\substack{n\to+\infty\\ m\to+\infty}} E\left(\sum_{i=0}^{n}\sum_{j=0}^{m}\varphi^i\varepsilon_{t-i}\varphi^j\varepsilon_{t+h-j}\right)=E\left(X_tX_{t+h}\right).$$

Ora,

$$\lim_{\substack{n\to+\infty\\ m\to+\infty}} E\left(\sum_{i=0}^{n}\sum_{j=0}^{m}\varphi^i\varepsilon_{t-i}\varphi^j\varepsilon_{t+h-j}\right)=\lim_{\substack{n\to+\infty\\ m\to+\infty}}\sum_{i=0}^{n}\sum_{j=0}^{m}\varphi^{i+j}E\left(\varepsilon_{t-i}\varepsilon_{t+h-j}\right).$$

Consideremos $h\geq 0$.

$$E\left(\varepsilon_{t-i}\varepsilon_{t+h-j}\right)=\begin{cases}\sigma^2, & -i=h-j\\ 0, & \text{se não}\end{cases}=\begin{cases}\sigma^2, & j=i+h\\ 0, & \text{se não}\end{cases}.$$

Então

$$\begin{aligned}
\lim_{\substack{n\to+\infty \\ m\to+\infty}} \sum_{i=0}^{n}\sum_{j=0}^{m} \varphi^{i+j} E\left(\varepsilon_{t-i}\varepsilon_{t+h-j}\right) &= \lim_{n\to+\infty} \sum_{i=0}^{n-h} \varphi^{i+i+h} E(\varepsilon_{t-i}^2) \\
&= \sigma^2 \varphi^h \lim_{n\to+\infty} \sum_{i=0}^{n-h} \varphi^{2i+h} \\
&= \sigma^2 \varphi^h \lim_{n\to+\infty} \frac{1-\varphi^{2(n-h)}\varphi^2}{1-\varphi^2} \\
&= \frac{\sigma^2\varphi^h}{1-\varphi^2}.
\end{aligned}$$

Concluímos que

$$E\left(X_t X_{t+h}\right) = \frac{\sigma^2\varphi^h}{1-\varphi^2}.$$

Se $h < 0$, obtemos de modo análogo

$$E\left(\varepsilon_{t-i}\varepsilon_{t+h-j}\right) = \begin{cases} \sigma^2, & -i = h-j \\ 0, & \text{se não} \end{cases} = \begin{cases} \sigma^2, & i = j-h \\ 0, & \text{se não} \end{cases}.$$

e então

$$\begin{aligned}
\lim_{\substack{n\to+\infty \\ m\to+\infty}} \sum_{i=0}^{n}\sum_{j=0}^{m} \varphi^{i+j} E\left(\varepsilon_{t-i}\varepsilon_{t+h-j}\right) &= \sigma^2 \varphi^h \lim_{m\to+\infty} \sum_{j=0}^{m+h} \varphi^{2j-h} \\
&= \sigma^2 \varphi^{-h} \lim_{m\to+\infty} \frac{1-\varphi^{2(m+h)}\varphi^2}{1-\varphi^2} \\
&= \frac{\sigma^2\varphi^{-h}}{1-\varphi^2}.
\end{aligned}$$

Concluímos que

$$E\left(X_t X_{t+h}\right) = \frac{\sigma^2\varphi^{-h}}{1-\varphi^2}.$$

Assim, $\forall t, h \in \mathbb{Z}$

$$\gamma(h) = Cov(X_t, X_{t+h}) = E(X_t X_{t+h}) = \frac{\sigma^2\varphi^{|h|}}{1-\varphi^2}.$$

Vejamos que a solução estacionária obtida corresponde a uma representação canónica, isto é, tal que o ruído branco interveniente, ε_t, é ortogonal ao passado do processo X, $\underline{H}_{t-1}^X$[1].

[1] $\underline{H}_t^X$ é o menor subespaço fechado de L^2 gerado por $X_\tau, \tau \leq t$, e pelas constantes. Assim, $\underline{H}_t^X$ é constituído pelos limites, no sentido em L^2, das combinações lineares afins $\sum_{i=1}^{k} A_i X_{\tau_i} + b, \tau_i \leq t$, com $A_1, A_2, \cdots, A_k$ e b reais. Dizemos que $\underline{H}_t^X$ é o subespaço linear gerado pelo passado do processo no instante t.

Ora,

$$X_t = \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i} \in \underline{H}_t^{\varepsilon}$$

de onde deduzimos que $\underline{H}_t^X \subset \underline{H}_t^{\varepsilon}, \forall t \in \mathbb{Z}$.

Mas ε_t é ortogonal a $\underline{H}_{t-1}^{\varepsilon}$. Então ε_t é ortogonal a $\underline{H}_{t-1}^X$.

Observemos que sempre que temos uma representação canónica de X, ε_t é a inovação no instante t. De facto, a esperança linear[2], E_L, de X_t dado o seu passado é

$$\begin{aligned} E_L\left(X_t/\underline{H}_{t-1}^X\right) &= E_L\left(\varphi X_{t-1} + \varepsilon_t/\,\underline{H}_{t-1}^X\right) \\ &= \varphi X_{t-1} \end{aligned}$$

pois ε_t é ortogonal a $\underline{H}_{t-1}^X$.

Assim $\varepsilon_t = X_t - \varphi X_{t-1} = X_t - E_L\left(X_t/\underline{H}_{t-1}^X\right)$ é a inovação de X no instante t.

Concluímos que se $\mid \varphi \mid < 1$ o modelo admite uma solução estacionária dada por

$$X_t = \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i}.$$

Além disso a representação $X_t = \varphi X_{t-1} + \varepsilon_t$ é canónica.

Estudemos agora o caso em que $\mid \varphi \mid > 1$.

Suponhamos que há uma solução estacionária.

Podemos escrever $X_{t-1} = \frac{1}{\varphi} X_t - \frac{1}{\varphi}\varepsilon_t$, isto é, para todo o $t \in \mathbb{Z}$ temos

$$\begin{aligned} X_t &= \frac{1}{\varphi} X_{t+1} - \frac{1}{\varphi}\varepsilon_{t+1} \\ &= \frac{1}{\varphi}\left(\frac{1}{\varphi} X_{t+2} - \frac{1}{\varphi}\varepsilon_{t+2}\right) - \frac{1}{\varphi}\varepsilon_{t+1} \\ &= \frac{1}{\varphi^2} X_{t+2} - \frac{1}{\varphi^2}\varepsilon_{t+2} - \frac{1}{\varphi}\varepsilon_{t+1} \\ &= \cdots \\ &= -\sum_{i=1}^{n} \frac{1}{\varphi^i}\varepsilon_{t+i} + \frac{1}{\varphi^n} X_{t+n} \end{aligned}$$

[2]A esperança linear é uma esperança condicional particular, pelo que goza do mesmo tipo de propriedades. Para um estudo mais detalhado ver Gouriéroux, Monfort (1989), Vol. 2, (Annexe).

Vejamos que $-\sum_{i=1}^{n} \frac{1}{\varphi^i}\varepsilon_{t+i}$ converge em L^2 para $-\sum_{i=1}^{+\infty} \frac{1}{\varphi^i}\varepsilon_{t+i}$.

$$\begin{aligned}
\left\| -\sum_{i=1}^{n} \frac{1}{\varphi^i}\varepsilon_{t+i} + \sum_{i=1}^{+\infty} \frac{1}{\varphi^i}\varepsilon_{t+i} \right\| &= \left\| \sum_{i=n+1}^{+\infty} \frac{1}{\varphi^i}\varepsilon_{t+i} \right\| \\
&\leq \sum_{i=n+1}^{+\infty} \mid \frac{1}{\varphi^i} \mid \|\varepsilon_{t+i}\| \\
&= \sum_{i=n+1}^{+\infty} \frac{1}{\mid \varphi^i \mid} \left[E(\varepsilon_{t+i})^2\right]^{1/2} \\
&= \sigma \sum_{i=n+1}^{+\infty} \frac{1}{\mid \varphi^i \mid}
\end{aligned}$$

que converge para zero, quando $n \to +\infty$, pois $\mid \frac{1}{\varphi} \mid < 1$.

A parcela $\frac{1}{\varphi^n}X_{t+n}$ tende (em L^2) para zero, quando $n \to +\infty$. De facto,

$$\|\frac{1}{\varphi^n}X_{t+n}\| = \left[E(\frac{1}{\varphi^n}X_{t+n})^2\right]^{1/2} \underset{n\to+\infty}{\longrightarrow} 0$$

uma vez que, sendo X estacionário, $E(X_t^2) = \text{constante} < +\infty$ e $\mid \frac{1}{\varphi} \mid < 1$.

Obtemos então uma única solução estacionária dada por

$$X_t = -\sum_{i=1}^{+\infty} \frac{1}{\varphi^i}\varepsilon_{t+i}.$$

Analisemos se esta representação é canónica, isto é, se o ruído branco interveniente ε_t é ortogonal ao passado $\underline{H}_{t-1}^X$. Ora

$$\begin{aligned}
Cov(\varepsilon_t, X_{t-1}) &= E(\varepsilon_t X_{t-1}) \\
&= E\left[\varepsilon_t\left(-\frac{1}{\varphi}\varepsilon_t - \frac{1}{\varphi^2}\varepsilon_{t+1} - \cdots\right)\right] \\
&= -\frac{1}{\varphi}\sigma^2.
\end{aligned}$$

Como $Cov(\varepsilon_t, X_{t-1}) \neq 0, \varepsilon_t$ não é ortogonal a $\underline{H}_{t-1}^X$ pelo que a representação $X_t = \varphi X_{t-1} + \varepsilon_t$ não é canónica.

Vejamos que o processo X admite no entanto uma representação canónica AR(1) com um processo de inovação diferente de ε.

Comecemos por determinar a densidade espetral de ε, f_ε.[1]

Como ε é um ruído branco de variância σ^2, temos

$$\gamma(h) = \begin{cases} \sigma^2, & h = 0 \\ 0, & h \neq 0 \end{cases}.$$

Então

$$\sum_{h=-\infty}^{+\infty} | \gamma(h) |= \gamma(0) = \sigma^2 < +\infty$$

pelo que $f_\varepsilon(\omega)$ existe e é dada por

$$f_\varepsilon(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \gamma(h) e^{i\omega h} = \frac{\sigma^2}{2\pi}, \omega \in [-\pi, \pi[\,.$$

De $X_t - \varphi X_{t-1} = \varepsilon_t$ deduzimos que a densidade espetral de X verifica [2]

$$f_X(\omega) \mid 1 - \varphi e^{i\omega} \mid^2 = f_\varepsilon(\omega), \quad \omega \in [-\pi, \pi[\qquad \textbf{(2.2.2)}$$

e consequentemente

$$f_X(\omega) = \frac{\sigma^2}{2\pi} \frac{1}{\mid 1 - \varphi e^{i\omega} \mid^2}.$$

Consideremos o processo η definido por

$$\eta_t = X_t - \frac{1}{\varphi} X_{t-1}, \quad t \in \mathbb{Z}.$$

[1] Teorema. Se a função de autocovariância de um processo estacionário é absolutamente somável, isto é, se

$$\sum_{h=-\infty}^{+\infty} | \gamma(h) |< +\infty$$

o processo admite densidade espetral sendo uma versão dada por

$$f(\omega) = \frac{1}{2\pi} \sum_{h=-\infty}^{+\infty} \gamma(h) e^{i\omega h}.$$

[2] Seja $A(L)X_t = Y_t$, tal que $A(L) = \sum_{i=-\infty}^{+\infty} a_i L^i$. Se $(a_i, i \in \mathbb{Z})$ é uma sucessão real absolutamente somável (i.e. $\sum_{i=-\infty}^{+\infty} \mid a_i \mid < +\infty$) e existe f_Y, então a densidade espetral de X existe e é dada por $f_X(\omega) \mid A(e^{i\omega}) \mid^2 = f_Y(\omega), \omega \in [-\pi, \pi[$. (Gonçalves, E., N., Mendes-Lopes, 2008, p.63)

A sua densidade espetral é

$$\begin{aligned} f_\eta(\omega) &= \mid 1 - \frac{1}{\varphi} e^{i\omega} \mid^2 f_X(\omega) \\ &= \frac{\sigma^2}{2\pi} \frac{\mid 1 - \frac{1}{\varphi} e^{i\omega} \mid^2}{\mid 1 - \varphi e^{i\omega} \mid^2} \\ &= \frac{\sigma^2}{2\pi} \frac{1}{\mid \varphi \mid^2} \frac{\mid 1 - \frac{1}{\varphi} e^{i\omega} \mid^2}{\mid \frac{1}{\varphi} - e^{i\omega} \mid^2} \\ &= \frac{\sigma^2}{2\pi} \frac{1}{\mid \varphi \mid^2}, \omega \in [-\pi, \pi[\,. \end{aligned}$$

Então η é um ruído branco de variância $\frac{\sigma^2}{\mid \varphi \mid^2}$, tendo em conta a caraterização de ruído branco através da densidade espetral.

O processo X admite, assim, as duas seguintes representações auto-regressivas de ordem 1

$$\begin{aligned} X_t - \varphi X_{t-1} = \varepsilon_t, \quad V(\varepsilon_t) = \sigma^2 \\ X_t - \frac{1}{\varphi} X_{t-1} = \eta_t, \quad V(\eta_t) = \frac{\sigma^2}{\mid \varphi \mid^2}. \end{aligned}$$

Esta segunda representação é canónica pois $\mid \frac{1}{\varphi} \mid < 1$.

Concluímos, então, que se $|\varphi| > 1$ o modelo admite uma solução estacionária dada por

$$X_t = -\sum_{i=0}^{+\infty} \frac{1}{\varphi^i} \varepsilon_{t+i}.$$

Além disso se considerarmos $\varphi^* = \frac{1}{\varphi}$, $|\varphi^*| < 1$, obtemos uma representação canónica dada por

$$X_t - \varphi^* X_{t-1} = \eta_t,$$

onde $\eta = (\eta_t, t \in \mathbb{Z})$ se interpreta como inovação do processo X.

Analisemos finalmente os casos em que $\varphi = 1$ ou $\varphi = -1$.

Suponhamos que há uma solução X estacionária. Podemos sempre escrever

$$X_t - \varphi^n X_{t-n} = \varepsilon_t + \varphi \varepsilon_{t-1} + \cdots + \varphi^{n-1} \varepsilon_{t-(n-1)}. \quad \textbf{(2.2.3)}$$

Vejamos quando vale a variância de cada um dos membros desta igualdade.

$$\begin{aligned} V\left(X_t - \varphi^n X_{t-n}\right) &= V(X_t) + V(X_{t-n}) - 2\varphi^n cov(X_t, X_{t-n}) \\ &= 2V(X_t) - 2\varphi^n cov(X_t, X_{t-n}) \\ &\leq 2V(X_t) + 2\left|cov(X_t, X_{t-n})\right|, \quad \text{pois } \mid \varphi^n \mid = 1 \\ &\leq 4V(X_t) \end{aligned}$$

pois, pela desigualdade de Schwartz, $|Cov(X,Y)|^2 \leq V(X)V(Y)$.

Assim $X_t - \varphi^n X_{t-n}$ é majorado, em L^2, por $4\gamma(0)$.

Do lado direito de **(2.2.3)**, temos

$$V\left(\varepsilon_t + \varphi\varepsilon_{t-1} + \cdots + \varphi^{n-1}\varepsilon_{t-(n-1)}\right) = V(\varepsilon_t) + V\left(\varphi\varepsilon_{t-1}\right) + \cdots + V\left(\varphi^{n-1}\varepsilon_{t-(n-1)}\right),$$

pois $Cov(\varepsilon_s, \varepsilon_t) = 0, \quad \forall t \neq s$.

Assim

$$\begin{aligned} V\left(\varepsilon_t + \varphi\varepsilon_{t-1} + \cdots + \varphi^{n-1}\varepsilon_{t-(n-1)}\right) &= V(\varepsilon_t) + \varphi^2 V\left(\varepsilon_{t-1}\right) + \cdots + \varphi^{2(n-1)} V\left(\varepsilon_{t-(n-1)}\right) \\ &= n\sigma^2, \quad \text{pois } \varphi = \pm 1 \end{aligned}$$

que tende para $+\infty$ quando $n \to +\infty$.

Encontramos pois uma contradição pelo que, quando $\varphi = \pm 1$, não há solução estacionária.

Pelo exposto limitamo-nos a considerar, no que se segue, apenas o caso em que $|\varphi| < 1$, pois para $\varphi = \pm 1$ o processo não admite solução estacionária e o caso em que $|\varphi| > 1$ pode ser sempre reduzido a este, bastando para isso considerar $\varphi^* = \dfrac{1}{\varphi}$ e um ruído adequado η que é o processo inovação de X.

2.2.2. Estacionaridade forte e Ergocidade

O estudo que apresentamos nos capítulos seguintes está bastante dependente das propriedades de estacionaridade forte e de ergocidade do processo em estudo.

Supomos que $(\varepsilon_t, t \in \mathbb{Z})$ é uma sucessão de v.a.i.i.d.. O caso particular em que vamos desenvolver a estimação do modelo **(B)** está nestas condições pois suporemos que ε um ruído branco Gaussiano.

Vamos então provar que $X_t = \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i}$ é a única solução fortemente estacionária e ergódica de **(B)**.

Relativamente à estacionaridade forte começamos por provar, usando o critério de Cauchy, que a série $\sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i}$ é q.c. convergente.
Ora

$$\begin{aligned}
\lim_{i \to +\infty} \left(\mid \varphi \mid^i \mid \varepsilon_{t-i} \mid \right)^{1/i} &\overset{q.c.}{=} \lim_{i \to +\infty} \mid \varphi \mid \mid \varepsilon_{t-i} \mid^{1/i} \\
&\overset{q.c.}{=} \mid \varphi \mid \lim_{i \to +\infty} e^{\frac{1}{i} \ln \mid \varepsilon_{t-i} \mid}.
\end{aligned}$$

Provemos que $\lim_{i \to +\infty} \frac{1}{i} \ln \mid \varepsilon_{t-i} \mid = 0$, q.c.. O facto de ε_t ser uma sucessão de variáveis aleatórias, identicamente distribuídas implica

$$P\left(\left\{ \lim_{i \to +\infty} \frac{1}{i} \ln \mid \varepsilon_{t-i} \mid = 0 \right\} \right) = P\left(\left\{ \lim_{i \to +\infty} \frac{1}{i} \ln \mid \varepsilon_t \mid = 0 \right\} \right) = 1,$$

pois, $\ln \mid \varepsilon_t \mid$ não depende de i.

Concluímos assim que

$$\lim_{i \to +\infty} \left(\mid \varphi \mid^i \mid \varepsilon_{t-i} \mid \right)^{1/i} = \mid \varphi \mid, \text{ q.c..}$$

Como $|\varphi| < 1$, o critério de Cauchy para séries numéricas de termos positivos permite então afirmar que $X_t = \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i}, t \in \mathbb{Z}$, é q.c. convergente, definindo assim um processo fortemente estacionário uma vez que, nestas condições, é função mensurável de $\varepsilon_t, \varepsilon_{t-1}, \cdots$ e ε é um processo fortemente estacionário. [1]

Para provar a ergodicidade de $X_t = \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i}$ vamos ter em conta dois casos importantes de processos ergódicos, concretamente as sucessões de variáveis aleatórias,

[1] A estacionaridade forte de uma sério temporal $(X_t, t \in \mathbb{Z})$ implica a estacionaridade forte da série temporal $(Z_t, t \in \mathbb{Z})$ com

$$Z_t = f\left(\cdots, X_{t-1}, X_t, X_{t+1}, \cdots\right),$$

para toda a função mensurável f definida em $\mathbb{R}^{\mathbb{Z}}$ e com valores em $\mathbb{R}$ (Azencott, Dacunha-Castelle, 1984, p.30).

independentes e identicamente distribuídas e as funções mensuráveis de processos ergódicos. Com efeito, qualquer sucessão $(Y_t, t \in \mathbb{Z})$ de v.a.i.i.d. constitui um processo ergódico (Azencott, Dacunha-Castelle, 1984, p.32); além disso, se $(Y_t, t \in \mathbb{Z})$ é um processo ergódico, o mesmo acontece ao processo $(Z_t, t \in \mathbb{Z})$, tal que

$$Z_t = f\left(\cdots, Y_{t-2}, Y_{t-1}, Y_t, X_{t+1}, Y_{t+2}, \cdots\right),$$

com f uma função mensurável definida sobre $\mathbb{R}^{\mathbb{Z}}$ (Rosenblatt, 1978, p.33).

Ora, $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ é um processo ergódico porque é uma sucessão de v.a.i.i.d. e, tendo-se

$$X_t \overset{q.c.}{=} \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i}, \quad \forall t \in \mathbb{Z},$$

conclui-se que X é uma função mensurável do processo ergódico ε, pelo que X é também um processo ergódico.

Para este tipo de modelo é válido o seguinte teorema, designado por Teorema Ergódico, ao qual recorreremos algumas vezes no capítulo 3, relativo à análise estatística do modelo AR(1).

Teorema Ergódico: Seja $X = (X_t, t \in \mathbb{Z})$ uma série temporal fortemente estacionária e ergódica. Consideremos a seguinte função de translação τ de $\mathbb{R}^{\mathbb{Z}}$ em $\mathbb{R}^{\mathbb{Z}}$,

$$\tau\left[(x_t, t \in \mathbb{Z})\right] = (y_t, t \in \mathbb{Z}), \qquad \text{com } y_t = x_{t+1}.$$

Se g é uma função de $\mathbb{R}^{\mathbb{Z}}$ em $\mathbb{R}$, P_X-integrável, tem-se

$$\lim_{T \to +\infty} \frac{1}{T}\left[g(x) + g \circ \tau(x) + \cdots + g \circ \tau^T(x)\right] \overset{P_X-q.c.}{=} \int_{\mathbb{R}^{\mathbb{Z}}} g(x) P_X(dx).$$

Suponhamos que g é da forma

$$\forall x = (x_n)_{n \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}, \quad g(x) = f(x_1, \cdots, x_k),$$

com f de $\mathbb{R}^k$ em $\mathbb{R}$ tal que $E\left[f\left(X_1, \cdots, X_k\right)\right]$ existe.

O teorema ergódico permite escrever

$$\lim_{T \to +\infty} \frac{1}{T}\left[f\left(X_1, \cdots, X_k\right) + f\left(X_2, \cdots, X_{k+1}\right) + \cdots + f\left(X_{T+1}, \cdots, X_{T+k}\right)\right]$$

$$\overset{P_X-q.c.}{=} E\left[f\left(X_1, \cdots, X_k\right)\right].$$

2.2.3. Estimação dos parâmetros do modelo

Estimadores da máxima verosimilhança

Consideremos a amostra $(X_1, \cdots, X_T)$ de dimensão $T, T \in \mathbb{N}$, associada ao processo estocástico X verificando o modelo **(B)** sujeito às seguintes condições $\mid \varphi \mid < 1$ e $\varepsilon_t \sim N(0, \sigma)$.

Para determinar o e.m.v. do parâmetro (φ, σ^2) torna-se necessário deduzir a lei de X_t, $t \in \mathbb{Z}$, de modo a encontrar a função de verosimilhança do processo. Como a função característica determina a lei, começaremos por determinar a função característica de X_t. Uma vez que $S_{n,t}$ converge em média quadrática quando $n \to +\infty$ (o que implica que também converge em lei), é possível determinar a função característica de X_t através da função característica de $S_{n,t}$.

Consideremos, então, $S_{n,t} = \sum\limits_{i=0}^{n} \varphi^i \varepsilon_{t-i}$. Vamos agora determinar a sua função característica, $\Phi_{S_{n,t}}$.

$$\forall u \in \mathbb{R}, \qquad \Phi_{S_{n,t}}(u) = \Phi_{\sum\limits_{i=0}^{n} \varphi^i \varepsilon_{t-i}}(u) = \prod_{i=0}^{n} \Phi_{(\varphi^i \varepsilon_{t-i})}(u)$$

pois $\varepsilon_{t-n}, \cdots, \varepsilon_t$ são independentes.

Então,

$$\begin{aligned}
\Phi_{S_{n,t}}(u) &= \prod_{i=0}^{n} \Phi_{\varepsilon_t}\left(\varphi^i u\right) = \prod_{j=0}^{n} \exp\left(i \varphi^j u 0\right) \exp\left(-\frac{\sigma^2 \varphi^{2j} u^2}{2}\right) \\
&= \exp\left(-\frac{u^2}{2} \sigma^2 \sum_{j=0}^{n} \varphi^{2j}\right) \\
&= \exp\left(-\frac{u^2}{2} \sigma^2 \frac{1 - \varphi^{2n}}{1 - \varphi^2}\right), \quad \text{pois } \mid \varphi \mid < 1.
\end{aligned}$$

Daqui concluímos que

$$\forall u \in \mathbb{R}, \lim_{n \to +\infty} \Phi_{S_{n,t}}(u) = \exp\left(-\frac{u^2}{2} \frac{\sigma^2}{1 - \varphi^2}\right) = \Phi_Y(u)$$

com $Y \sim N\left(0, \sqrt{\dfrac{\sigma^2}{1 - \varphi^2}}\right)$.

Como $S_{n,t}$ converge em média quadrática quando $n \to +\infty$ (o que implica que também converge em lei) para X_t e usando o teorema de Paul Lévy[1] podemos concluir que

$$X_t \sim N\left(0, \sqrt{\frac{\sigma^2}{1-\varphi^2}}\right).$$

Para obter a lei do vector $(X_1, \cdots, X_T), T \in \mathbb{N}$, comecemos por provar que $\forall j \in \{1, \cdots, T\}$,

$$F_{X_j|X_1\cdots,X_{j-1}}(x) = F_{X_j|X_{j-1}} = F_{\varepsilon_j}(x - \varphi x_{j-i}), P_{X_{j-1}} - q.c.$$

Seja $x \in \mathbb{R}$. Ora,

$$\begin{aligned}
F_{X_j|X_1\cdots,X_{j-1}}(x) &= P\left(X_j \leq x \mid X_1 = x_1, \cdots, X_{j-1} = x_{j-1}\right), P_{(X_1,\cdots,X_{j-1})} - q.c.\\
&= P\left(\varphi X_{j-1} + \varepsilon_j \leq x \mid X_1 = x_1, \cdots, X_{j-1} = x_{j-1}\right)\\
&= P\left(\varphi x_{j-1} + \varepsilon_j \leq x \mid X_1 = x_1, \cdots, X_{j-1} = x_{j-1}\right)\\
&= P\left(\varepsilon_j \leq x - \varphi x_{j-1}\right), \text{ pois } \varepsilon_j \text{ é independente de } X_1, \cdots, X_{j-1}\\
&= F_{\varepsilon_j}(x - \varphi x_{j-1}).
\end{aligned}$$

Por outro lado

$$\begin{aligned}
F_{X_j|X_{j-1}}(x) &= P\left(X_j \leq x \mid X_{j-1} = x_{j-1}\right), P_{X_{j-1}} - q.c.\\
&= P\left(\varphi X_{j-1} + \varepsilon_j \leq x \mid X_{j-1} = x_{j-1}\right)\\
&= P\left(\varphi x_{j-1} + \varepsilon_j \leq x \mid X_{j-1} = x_{j-1}\right)\\
&= P\left(\varepsilon_j \leq x - \varphi x_{j-1}\right), \text{ pois } \varepsilon_j \text{ é independente de } X_{j-1}\\
&= F_{\varepsilon_j}(x - \varphi x_{j-1}).
\end{aligned}$$

Concluímos então que $F_{X_j|X_1\cdots,X_{j-1}}(x) = F_{X_j|X_{j-1}}(x)$.

Tendo em conta que a densidade de $(X_1, \cdots, X_T), T \in \mathbb{Z}$, é

$$f_{(X_1,\cdots,X_T)}(x_1, \cdots, x_T) = f_{X_1}(x_1) f_{X_2|X_1}(x_2) \cdots f_{X_T|X_1,\cdots,X_{T-1}}(x_T)$$

e que

$$\begin{aligned}
f_{X_1}(x_1) &= \frac{\sqrt{1-\varphi^2}}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(1-\varphi^2)x_1^2\right\}\\
f_{X_2|X_1}(x_2) &\overset{q.c.}{=} f_{X_2|X_1=x_1}(x_2) = f_{\varepsilon_j}(x_2 - \varphi x_1) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x_2 - \varphi x_1)^2\right\}
\end{aligned}$$

$$\vdots$$

[1] Teorema de Paul Lévy: Uma condição necessária e suficiente para que uma sucessão de v.a.r. $(X_n)_{n\in\mathbb{Z}}$ tenda em lei para uma v.a.r. X é que $\forall u \in \mathbb{R}, \lim\limits_{n\to+\infty} \Phi_{X_n}(u) = \Phi_X(u)$.

$$f_{X_T|X_1,\cdots,X_{T-1}}(x_T) = f_{\varepsilon_j}(x_T - \varphi x_{T-1}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x_T - \varphi x_{T-1})^2\right\}$$

obtemos

$$\begin{aligned} f_{(X_1,\cdots,X_T)}(x_1,\cdots,x_T) = & \frac{\sqrt{1-\varphi^2}}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(1-\varphi^2)x_1^2\right\}\frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x_2 - \varphi x_1)^2\right\} \\ & \cdots \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x_T - \varphi x_{T-1})^2\right\} \\ = & \frac{\sqrt{1-\varphi^2}}{\sigma^T(\sqrt{2\pi})^T} \exp\left\{-\frac{1}{2\sigma^2}\left((1-\varphi^2)x_1^2 + \sum_{i=2}^{T}(x_i - \varphi x_{i-1})^2\right)\right\}. \end{aligned}$$

A função de verosimilhança do modelo associada a uma realização da amostra $x = (x_1, \cdots, x_T)$ é então

$$L((x_1,\cdots,x_T),\theta) = \frac{\sqrt{1-\varphi^2}}{\sigma^T(\sqrt{2\pi})^T} \exp\left\{-\frac{1}{2\sigma^2}\left((1-\varphi^2)x_1^2 + \sum_{i=2}^{T}(x_i - \varphi x_{i-1})^2\right)\right\}$$

onde $\theta = (\varphi, \sigma^2)$.

Usando os mesmos argumentos que no modelo anterior, um possível e.m.v. do par (φ, σ^2) é solução de

$$\left\{ \begin{array}{l} \dfrac{\partial \log L\left(x,(\varphi,\sigma^2)\right)}{\partial\varphi} = 0 \\ \\ \dfrac{\partial \log L\left(x,(\varphi,\sigma^2)\right)}{\partial\sigma^2} = 0 \end{array} \right. . \qquad \mathbf{(2.1)}$$

com

$$\begin{aligned} \log L = & \frac{1}{2}\log(1-\varphi^2) - \frac{T}{2}\log(2\pi) - \frac{T}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\left[(1-\varphi^2)x_1^2 + \sum_{i=2}^{T}(x_i - \varphi x_{i-1})^2\right] \\ = & \frac{1}{2}\log(1-\varphi^2) - \frac{T}{2}\log(2\pi) - \frac{T}{2}\log(\sigma^2) - \\ & - \frac{1}{2\sigma^2}\left[x_1^2 + x_T^2 - \varphi^2 x_1^2 + \sum_{i=2}^{T-1} x_i^2 + \varphi^2\sum_{i=2}^{T} x_{i-1}^2 - 2\varphi\sum_{i=2}^{T} x_i x_{i-1}\right] \\ = & \frac{1}{2}\log(1-\varphi^2) - \frac{T}{2}\log(2\pi) - \frac{T}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\left[x_1^2 + x_T^2 + (1+\varphi^2)\sum_{i=2}^{T-1} x_i^2 - 2\varphi\sum_{i=2}^{T} x_i x_{i-1}\right] \end{aligned}$$

Para simplificar a notação, sejam

$$P_0' = x_1^2 + x_T^2$$

$$P_0 = \sum_{i=2}^{T-1} x_i^2$$

$$P_1 = \sum_{i=2}^{T} x_i x_{i-1}.$$

Ora

$$\frac{\partial \log L\left(x,(\varphi,\sigma^2)\right)}{\partial \varphi} = -\frac{\varphi}{1-\varphi^2} + \frac{P_1}{\sigma^2} - \frac{\varphi P_0}{\sigma^2}$$
$$\frac{\partial \log L\left(x,(\varphi,\sigma^2)\right)}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{P_0' + (1+\varphi^2)P_0 - 2\varphi P_1}{2\sigma^4}.$$

Assim, o sistema **(2.1)** é equivalente a

$$\left\{\begin{array}{l} -\dfrac{\varphi}{1-\varphi^2} + \dfrac{P_1}{\sigma^2} - \dfrac{\varphi P_0}{\sigma^2} = 0 \\ \\ -\dfrac{T}{2\sigma^2} + \dfrac{P_0' + (1+\varphi^2)P_0 - 2\varphi P_1}{2\sigma^4} = 0 \end{array}\right. .$$

Então a solução $(\hat{\varphi}, \hat{\sigma}^2)$ do sistema verifica

$$\left\{\begin{array}{ll} -\hat{\sigma}^2\hat{\varphi} + (1-\hat{\varphi}^2)\,(P_1 - \hat{\varphi}P_0) = 0 & \textbf{(2.2)} \\ \\ \hat{\sigma}^2 = \dfrac{P_0' + (1+\hat{\varphi}^2)P_0 - 2\hat{\varphi}P_1}{T} & \textbf{(2.3)} \end{array}\right.$$

Ora a equação **(2.2)** é equivalente a

$$\begin{aligned} &-\hat{\varphi}\frac{P_0' + (1+\hat{\varphi}^2)P_0 - 2\hat{\varphi}P_1}{T} + (1-\hat{\varphi}^2)\,(P_1 - \hat{\varphi}P_0) = 0 \\ &\Leftrightarrow P_1 - \hat{\varphi}P_0 - \hat{\varphi}^2P_1 + \hat{\varphi}^3P_0 - \frac{\hat{\varphi}P_0'}{T} - \frac{\hat{\varphi}P_0}{T} - \frac{\hat{\varphi}^3P_0}{T} + \frac{2\hat{\varphi}^2P_1}{T} = 0 \\ &\Leftrightarrow \hat{\varphi}^3P_0\frac{T-1}{T} - \hat{\varphi}^2P_1\frac{T-2}{T} - \hat{\varphi}\left(\frac{T+1}{T}P_0 + \frac{P_0'}{T}\right) + P_1 = 0. \quad \textbf{(2.4)} \end{aligned}$$

Analisemos a localização das raízes da equação **(2.4)**.

Se $\hat{\varphi} = -1$, o polinómio presente na equação **(2.4)** é igual a

$$\begin{aligned} &-P_0\frac{T-1}{T} - P_1\frac{T-2}{T} + \frac{T+1}{T}P_0 + \frac{P_0'}{T} + P_1 \\ &= 2\frac{P_0}{T} + 2\frac{P_1}{T} + \frac{P_0'}{T} \\ &= \frac{1}{T}\left(2P_0 + 2P_1 + P_0'\right) \\ &= \frac{1}{T}\left(\sum_{i=2}^{T-1} x_i^2 + \sum_{i=2}^{T-1} x_i^2 + 2\sum_{i=2}^{T} x_ix_{i-1} + x_1^2 + x_T^2\right) \\ &= \frac{1}{T}\left(\sum_{i=2}^{T} x_i^2 + \sum_{i=2}^{T} x_{i-1}^2 + 2\sum_{i=2}^{T} x_ix_{i-1}\right) \\ &= \frac{1}{T}\sum_{i=2}^{T}(x_i + x_{i-1})^2 > 0. \end{aligned}$$

Se $\hat{\varphi} = 1$, o polinómio presente na equação **(2.4)** é igual a

$$
\begin{aligned}
& P_0 - \frac{P_0}{T} - P_1 + 2\frac{P_1}{T} - P_0 - \frac{P_0}{T} - \frac{P_0^{'}}{T} + P_1 \\
&= \frac{-2P_0 + 2P_1 - P_0^{'}}{T} \\
&= -\frac{1}{T}\left(\sum_{i=2}^{T-1} x_i^2 + \sum_{i=2}^{T-1} x_i^2 - 2\sum_{i=2}^{T} x_i x_{i-1} + x_1^2 + x_T^2\right) \\
&= -\frac{1}{T}\left(\sum_{i=2}^{T} x_i^2 + \sum_{i=2}^{T} x_{i-1}^2 - 2\sum_{i=2}^{T} x_i x_{i-1}\right) \\
&= -\frac{1}{T}\sum_{i=2}^{T} (x_i - x_{i-1})^2 < 0.
\end{aligned}
$$

Se $\hat{\varphi} = 0$, o polinómio presente na equação **(2.4)** é igual a P_1.

Concluímos que a equação **(2.4)** é positiva se $\hat{\varphi} = -1$, é negativa se $\hat{\varphi} = 1$ e igual a P_1 se $\hat{\varphi} = 0$; assim se $P_1 = 0$ então 0 é uma raiz, se $P_1 > 0$ existe pelo menos uma raiz entre 0 e 1 e, finalmente, se $P_1 < 0$ existe pelo menos uma raiz entre -1 e 0. Assim um estimador para φ obtido pelo m.m.v. será em módulo inferior a 1 pelo que é compatível com as condições de estacionaridade do modelo.

Observa-se que, segundo Anderson, 1971, p. 355, uma boa aproximação de uma solução de **(2.4)** é $\hat{\varphi} = \frac{P_1}{P_0}\left(1 - \frac{1}{T}\right)$.

Quando $T \to +\infty$ a equação **(2.4)** é assintóticamente equivalente em probabilidade a

$$
\begin{aligned}
& \hat{\varphi}^3 P_0 - \hat{\varphi}^2 P_1 - \hat{\varphi} P_0 + P_1 = 0 \\
& \Leftrightarrow \hat{\varphi}^3 - \hat{\varphi}^2 \frac{P_1}{P_0} - \hat{\varphi} + \frac{P_1}{P_0} = 0 \quad \mathbf{(2.5)}
\end{aligned}
$$

Analisemos as raízes da equação **(2.5)**.

Facilmente se verifica que $\hat{\varphi} = \pm 1$ são raízes. Deste modo a equação **(2.5)** é equivalente a

$$(\hat{\varphi} + 1)(\hat{\varphi} - 1)\left(\hat{\varphi} - \frac{P_1}{P_0}\right) = 0.$$

Concluímos então que as raízes da equação são

$$\hat{\varphi} = \pm 1 \quad \text{e} \quad \hat{\varphi} = \frac{P_1}{P_0}.$$

Assim, desprezando os termos de ordem $\frac{1}{T}$ (em probabilidade) obtém-se como estimador de φ, $\hat{\varphi} = \frac{P_1}{P_0}$.

Vejamos agora a equação **(2.3)**. Substituindo pelos valores de P_0', P_0 e P_1, obtemos

$$\hat{\sigma}_T^2 = \frac{x_1^2 + x_T^2 + \left(1+\hat{\varphi}^2\right)\sum\limits_{i=2}^{T-1} x_i^2 - 2\hat{\varphi}\sum\limits_{i=2}^{T} x_i x_{i-1}}{T}$$

$$= \frac{x_1^2 + x_T^2 + \sum\limits_{i=2}^{T-1} x_i^2 + \hat{\varphi}^2\sum\limits_{i=2}^{T-1} x_i^2 - 2\hat{\varphi}\sum\limits_{i=2}^{T} x_i x_{i-1}}{T}$$

$$= \frac{x_1^2 + x_T^2 + \sum\limits_{i=2}^{T-1} x_i^2 + \hat{\varphi}^2\sum\limits_{i=2}^{T-1} x_i^2 - 2\hat{\varphi}\sum\limits_{i=2}^{T} x_i x_{i-1} - \hat{\varphi}^2 x_1^2 + \hat{\varphi}^2 x_1^2}{T}$$

$$= \frac{x_1^2 - \hat{\varphi}^2 x_1^2 + x_T^2 + \sum\limits_{i=2}^{T-1} x_i^2 + \hat{\varphi}^2\sum\limits_{i=2}^{T-1} x_i^2 + \hat{\varphi}^2 x_1^2 - 2\hat{\varphi}\sum\limits_{i=2}^{T} x_i x_{i-1}}{T}$$

$$= \frac{1}{T}\sum_{i=2}^{T}\left(x_i^2 - 2\hat{\varphi}x_{i-1}x_i + \hat{\varphi}^2 x_{i-1}^2\right) + \frac{(1-\hat{\varphi}^2)x_1^2}{T}.$$

O estudo anterior permite concluir que

$$\begin{cases} \hat{\varphi}_T = \dfrac{\sum\limits_{i=2}^{T} X_i X_{i-1}}{\sum\limits_{i=2}^{T} X_{i-1}^2} \\ \hat{\sigma}_T^2 = \dfrac{1}{T}\sum\limits_{i=2}^{T}\left(X_i - \hat{\varphi}_T X_{i-1}\right)^2 + \dfrac{(1-\hat{\varphi}_T^2)X_1^2}{T} \end{cases}$$

é um novo estimador para o par (φ, σ^2) assintóticamente equivalente, em probabilidade, ao e.m.v. do par (φ, σ^2) associado ao modelo **(B)**.

Note-se que este equivalente assintótico do e.m.v. de $\left(\varphi, \sigma^2\right)$ coincide com o obtido para o modelo **(A)** a menos da parcela $\dfrac{(1-\hat{\varphi}_T^2)X_1^2}{T}$ presente em $\hat{\sigma}_T^2$.

Capítulo 3

Comportamento Assintótico dos estimadores

3.1. Equivalência assintótica dos momentos empíricos de segunda ordem para os dois processos

Com o objetivo de estudar o comportamento assintótico dos e.m.v. dos parâmetros φ e σ^2 associados ao modelo **(1)** quando $\mathcal{T} = \mathbb{N}_0$ (obtido em **(C)**) ou $\mathcal{T} = \mathbb{Z}$ (obtido em **(D)**), começamos por provar que, quando $\mid \varphi \mid < 1$, tais estimadores, são assintóticamente equivalentes. Como tais estimadores são função dos momentos empíricos de $2^{\underline{a}}$ ordem, nomeadamente da autocorrelação empírica, iremos estabelecer a equivalência assintótica destes momentos nesses modelos (Anderson, 1970, p.188).

Designemos por X_t os valores do processo correspondentes ao modelo **(B)** ($t \in \mathbb{Z}$) e X_t^* tais valores para o modelo **(A)** ($t \in \mathbb{N}_0$) quando $X_0 = 0$ e $\mid \varphi \mid < 1$.

Vejamos que $\sum\limits_{t=1}^{T} X_t^2 - \sum\limits_{t=1}^{T} (X_t^*)^2$ dividido por $T^\alpha, \forall \alpha > 0$, converge em probabilidade para zero.

Então

$$X_t^* = \sum_{s=0}^{t-1} \varphi^s \varepsilon_{t-s} \qquad \textbf{(3.1)}$$

$$X_t = \sum_{s=0}^{+\infty} \varphi^s \varepsilon_{t-s} \qquad \textbf{(3.2)}$$

De **(3.1)** e **(3.2)** vem

$$\begin{aligned} X_t - X_t^* &= \sum_{s=0}^{+\infty} \varphi^s \varepsilon_{t-s} - \sum_{s=0}^{t-1} \varphi^s \varepsilon_{t-s} \\ &= \sum_{s=t}^{+\infty} \varphi^s \varepsilon_{t-s} \\ &= \varphi^t \sum_{k=0}^{+\infty} \varphi^k \varepsilon_{-k} \\ &= \varphi^t X_0. \end{aligned}$$

Comecemos por estudar $\sum_{t=1}^{T} X_t^2 - \sum_{t=1}^{T} (X_t^*)^2$. Temos

$$\begin{aligned}\sum_{t=1}^{T} X_t^2 - \sum_{t=1}^{T} (X_t^*)^2 &= \sum_{t=1}^{T} (X_t - X_t^*)(X_t + X_t^*)\\ &= \sum_{t=1}^{T} \varphi^t X_0 \left(2X_t^* + \varphi^t X_0\right)\\ &= 2\sum_{t=1}^{T} \varphi^t X_0 X_t^* + X_0^2 \sum_{t=1}^{T} \varphi^{2t}. \qquad \textbf{(3.3)}\end{aligned}$$

Estudemos a primeira parcela de **(3.3)**.

$$\begin{aligned}E\left[\sum_{t=1}^{T} \varphi^t X_0 X_t^*\right]^2 &= E\left[\sum_{t=1}^{T} \varphi^t X_0 X_t^* \sum_{k=1}^{T} \varphi^k X_0 X_k^*\right]\\ &= \sum_{t=1}^{T}\sum_{k=1}^{T} E\left((X_0)^2 X_t^* X_k^*\right)\varphi^{t+k}\\ &= \sum_{t=1}^{T}\sum_{k=1}^{T} E(X_0)^2\, \underbrace{E(X_t^* X_k^*)}_{A}\, \varphi^{t+k},\end{aligned}$$

pois $X_0 = f(\varepsilon_0, \varepsilon_{-1}, \varepsilon_{-2}, \cdots)$ e t e $k \in \{1, \cdots, T\}$.

Analisemos A.

Como $E\left[(X_t^*)^2\right] = \sigma^2 \dfrac{1-\varphi^{2t}}{1-\varphi^2}$, vem

$$|E(X_t^* X_k^*)| \leq \sqrt{E\left[(X_t^*)^2\right]}\sqrt{E\left[(X_k^*)^2\right]} = \sigma^2 \sqrt{\frac{1-\varphi^{2t}}{1-\varphi^2}}\sqrt{\frac{1-\varphi^{2k}}{1-\varphi^2}}.$$

Então, retomando a primeira parcela de **(3.3)**

$$\begin{aligned}E\left[\sum_{t=1}^{T} \varphi^t X_0 X_t^*\right]^2 &\leq \sum_{t=1}^{T}\sum_{k=1}^{T} \varphi^{t+k}\sigma^2 \left(\frac{1}{1-\varphi^2}\right)^2 \sigma^2 \sqrt{1-\varphi^{2t}}\sqrt{1-\varphi^{2k}}\\ &= \sigma^4 \left(\frac{1}{1-\varphi^2}\right)^2 \sum_{t=1}^{T}\sum_{k=1}^{T} \varphi^{t+k}\sqrt{1-\varphi^{2t}}\sqrt{1-\varphi^{2k}}\\ &= \sigma^4 \left(\frac{1}{1-\varphi^2}\right)^2 \left[\sum_{t=1}^{T} \varphi^t \sqrt{1-\varphi^{2t}}\right]^2\\ &\leq \sigma^4 \left(\frac{1}{1-\varphi^2}\right)^2 \left(\sum_{t=1}^{T} |\varphi|^t\right)^2\\ &\leq \frac{\sigma^4}{(1-\varphi^2)^2}\left(\sum_{t=1}^{+\infty} |\varphi|^t\right)^2\\ &= \frac{\sigma^4}{(1-\varphi^2)^2}\frac{\varphi^2}{(1-|\varphi|)^2}.\end{aligned}$$

Ou seja, $E\left[\sum\limits_{t=1}^{T}\varphi^t X_0 X_t^*\right]^2$ é limitado por uma quantidade independente de T. Então

$$E\left[\frac{\left|\sum\limits_{t=1}^{T}\varphi^t X_0 X_t^*\right|}{T^\alpha}\right] \leq \frac{1}{T^\alpha}\sqrt{E\left|\sum_{t=1}^{T}\varphi^t X_0 X_t^*\right|^2} = \frac{1}{T^\alpha}\sqrt{\frac{\sigma^4}{(1-\varphi^2)^2}\frac{\varphi^2}{(1-\mid\varphi\mid)^2}} \underset{T\to+\infty}{\longrightarrow} 0,$$

pois $\alpha > 0$.

Assim concluímos que

$$\frac{1}{T^\alpha}\sum_{t=1}^{T}\varphi^t X_0 X_t^* \underset{T\to+\infty}{\overset{L^1}{\longrightarrow}} 0, \quad \forall\alpha > 0$$

$$\Rightarrow \frac{1}{T^\alpha}\sum_{t=1}^{T}\varphi^t X_0 X_t^* \underset{T\to+\infty}{\overset{P}{\longrightarrow}} 0, \quad \forall\alpha > 0$$

pois se $U_n \overset{L^1}{\longrightarrow} U \Rightarrow U_n \overset{P}{\longrightarrow} U$, quando $n \to +\infty$.

Analisemos a segunda parcela de **(3.3)**

$$\begin{aligned} E\left(X_0^2\sum_{t=1}^{T}\varphi^{2t}\right) &= E\left[X_0^2\right]\sum_{t=1}^{T}\varphi^{2t} \\ &= \sigma^2\frac{1}{1-\varphi^2}\frac{\varphi^2-\varphi^{2T+2}}{1-\varphi^2} \\ &\leq \sigma^2\frac{\varphi^2}{(1-\varphi^2)^2}. \end{aligned}$$

Então

$$\frac{E\left(X_0^2\sum\limits_{t=1}^{T}\varphi^{2t}\right)}{T^\alpha} \underset{T\to+\infty}{\longrightarrow} 0, \qquad \text{pois } \alpha > 0.$$

Concluímos que

$$\frac{1}{T^\alpha}X_0^2\sum_{t=1}^{T}\varphi^{2t} \underset{T\to+\infty}{\overset{L^1}{\longrightarrow}} 0, \quad \forall\alpha > 0$$

$$\Rightarrow \frac{1}{T^\alpha}X_0^2\sum_{t=1}^{T}\varphi^{2t} \underset{T\to+\infty}{\overset{P}{\longrightarrow}} 0, \quad \forall\alpha > 0.$$

Então

$$\frac{1}{T^\alpha}\left[\sum_{t=1}^{T}X_t^2 - \sum_{t=1}^{T}(X_t^*)^2\right] \underset{T\to+\infty}{\overset{P}{\longrightarrow}} 0, \quad \forall\alpha > 0$$

pois se $U_n \overset{P}{\longrightarrow} U, Y_n \overset{P}{\longrightarrow} Y \Rightarrow U_n \pm Y_n \overset{P}{\longrightarrow} U \pm Y$, quando $n \to +\infty$.

Analisemos agora $\sum\limits_{t=1}^{T}X_{t-1}X_t - \sum\limits_{t=1}^{T}X_{t-1}^*X_t^*$ **(3.4)**.

Temos

$$\sum_{t=1}^{T} X_{t-1}X_t - \sum_{t=1}^{T} X_{t-1}^* X_t^* = \sum_{t=1}^{T} X_{t-1}\left(\varphi X_{t-1} + \varepsilon_t\right) - \sum_{t=1}^{T} X_{t-1}^*\left(\varphi X_{t-1}^* + \varepsilon_t\right)$$
$$= \varphi\left[\sum_{t=1}^{T} X_{t-1}^2 - \sum_{t=1}^{T} (X_{t-1}^*)^2\right] + \sum_{t=1}^{T}\left(X_{t-1} - X_{t-1}^*\right)\varepsilon_t. \quad \textbf{(3.5)}$$

Comecemos por analisar a segunda parcela de **(3.5)** do ponto de vista da convergência em média quadrática.

$$E\left[\sum_{t=1}^{T}\left(X_{t-1} - X_{t-1}^*\right)\varepsilon_t\right]^2 = E\left[\sum_{t=1}^{T}\varphi^{t-1}X_0\varepsilon_t\sum_{k=1}^{T}\varphi^{k-1}X_0\varepsilon_k\right]$$

pois $X_j - X_j^* = \varphi^j X_0$.

Então

$$E\left[\sum_{t=1}^{T}\left(X_{t-1} - X_{t-1}^*\right)\varepsilon_t\right]^2 = \sum_{t=1}^{T}\sum_{k=1}^{T}\varphi^{t+k-2}E\left[X_0^2\varepsilon_t\varepsilon_k\right]$$
$$= \sum_{t=1}^{T}\varphi^{2t-2}\sigma^2 E\left[X_0^2\right]$$

pois $X_0 = f\left(\varepsilon_0, \varepsilon_{-1}, \varepsilon_{-2}, \cdots\right)$ e t e $k \in \{1, \cdots, T\}$.

Assim

$$E\left[\sum_{t=1}^{T}\left(X_{t-1} - X_{t-1}^*\right)\varepsilon_t\right]^2 = \sigma^4\frac{\varphi^{-2}}{1-\varphi^2}\frac{\varphi^2 - \varphi^{2T}}{1-\varphi^2}$$
$$\leq \sigma^4\frac{1}{(1-\varphi^2)^2}.$$

Ora,

$$E\left[\frac{\sum_{t=1}^{T}\left(X_{t-1} - X_{t-1}^*\right)\varepsilon_t}{T^\alpha}\right]^2 \leq \frac{1}{T^{2\alpha}}\sigma^4\frac{1}{(1-\varphi^2)^2} \underset{T\to+\infty}{\longrightarrow} 0, \qquad \text{pois } \alpha > 0.$$

Concluímos então que

$$\frac{1}{T^\alpha}\sum_{t=1}^{T}\left(X_{t-1} - X_{t-1}^*\right)\varepsilon_t \underset{T\to+\infty}{\overset{m.q.}{\longrightarrow}} 0, \quad \forall\alpha > 0$$
$$\Rightarrow \frac{1}{T^\alpha}\sum_{t=1}^{T}\left(X_{t-1} - X_{t-1}^*\right)\varepsilon_t \underset{T\to+\infty}{\overset{P}{\longrightarrow}} 0, \quad \forall\alpha > 0.$$

Finalmente, a primeira parcela de **(3.5)** verifica

$$\frac{1}{T^\alpha}\left[\sum_{t=1}^{T}(X_{t-1})^2 - \sum_{t=1}^{T}(X_{t-1}^*)^2\right] = \frac{1}{T^\alpha}\left[\sum_{t=1}^{T}X_t^2 - \sum_{t=1}^{T}(X_t^*)^2 - X_T^2 + (X_T^*)^2 + X_0^2 - (X_0^*)^2\right].$$

Ora, para qualquer $\alpha > 0$

– pela igualdade **(3.3)**, sabemos que $\frac{1}{T^\alpha}\left[\sum_{t=1}^{T} X_t^2 - \sum_{t=1}^{T}(X_t^*)^2\right] \underset{T\to+\infty}{\overset{P}{\longrightarrow}} 0,$

– os últimos 4 termos convergem também em probabilidade para zero.

Podemos então concluir que

$$\frac{1}{T^\alpha}\left[\sum_{t=1}^{T} X_{t-1}^2 - \sum_{t=1}^{T}(X_{t-1}^*)^2\right] \underset{T\to+\infty}{\overset{P}{\longrightarrow}} 0, \quad \forall\alpha > 0.$$

Então

$$\frac{1}{T^\alpha}\left[\sum_{t=1}^{T} X_{t-1}X_t - \sum_{t=1}^{T} X_{t-1}^*X_t^*\right] \underset{T\to+\infty}{\overset{P}{\longrightarrow}} 0, \quad \forall\alpha > 0$$

pois se $U_n \overset{P}{\longrightarrow} U, Y_n \overset{P}{\longrightarrow} Y \Rightarrow U_n \pm Y_n \overset{P}{\longrightarrow} U \pm Y$, quando $n \to +\infty$.

Por fim, provemos que o e.m.v. associado ao modelo **(A)** e o e.m.v. associado ao modelo **(B)** são assintóticamente equivalentes, em probabilidade.

Usando definições evidentes para U_T, V_T, U_T^* e V_T^*, consideremos

$$\hat{\varphi}_T = \frac{\sum_{t=1}^{T} X_{t-1}X_t}{\sum_{t=1}^{T} X_{t-1}^2} = \frac{U_T}{V_T}$$

o estimador associado ao modelo **(B)** e

$$\hat{\varphi}_T^* = \frac{\sum_{t=1}^{T} X_{t-1}^*X_t^*}{\sum_{t=1}^{T} (X_{t-1}^*)^2} = \frac{U_T^*}{V_T^*}$$

o estimador associado ao modelo **(A)**.

Então

$$\begin{aligned}
\hat{\varphi}_T - \hat{\varphi}_T^* = \frac{\sum_{t=1}^{T} X_{t-1}X_t}{\sum_{t=1}^{T} X_{t-1}^2} - \frac{\sum_{t=1}^{T} X_{t-1}^*X_t^*}{\sum_{t=1}^{T} (X_{t-1}^*)^2} &= \frac{U_T}{V_T} - \frac{U_T^*}{V_T^*} \\
&= \frac{U_TV_T^* - V_TU_T^*}{V_TV_T^*} \\
&= \frac{U_TV_T^* - U_TV_T + U_TV_T - U_T^*V_T}{V_TV_T^*} \\
&= \frac{U_T\left(V_T^* - V_T\right) + \left(U_T - U_T^*\right)V_T}{V_TV_T^*} \\
&= \frac{\frac{U_T}{T^2}\left(V_T^* - V_T\right) + \frac{1}{T^2}\left(U_T - U_T^*\right)V_T}{\frac{1}{T}V_T\frac{1}{T}V_T^*}.
\end{aligned}$$

Note-se que

$$\forall \alpha > 0, \quad \frac{1}{T^{\alpha}} \left(V_T^* - V_T \right) \underset{T \to +\infty}{\overset{P}{\longrightarrow}} 0$$

e

$$\forall \alpha > 0, \quad \frac{1}{T^{\alpha}} \left(U_T - U_T^* \right) \underset{T \to +\infty}{\overset{P}{\longrightarrow}} 0.$$

Em particular, para $\alpha = 1$ temos

$$\frac{1}{T} \left(V_T^* - V_T \right) \underset{T \to +\infty}{\overset{P}{\longrightarrow}} 0$$

e

$$\frac{1}{T} \left(U_T - U_T^* \right) \underset{T \to +\infty}{\overset{P}{\longrightarrow}} 0.$$

Além disso

$$\frac{U_T}{T} = \frac{1}{T} \sum_{t=1}^{T} X_t X_{t-1} \underset{T \to +\infty}{\overset{q.c.,P}{\longrightarrow}} E \left(X_t X_{t-1} \right)$$

com $E \left(X_t X_{t-1} \right) = \dfrac{\sigma^2 \varphi}{1 - \varphi^2}$, tendo em conta o Teorema ergódico.

De facto, o processo $(X_t X_{t-1}, t \in \mathbb{Z})$ é fortemente estacionário e ergódico pois

$$\begin{aligned} \forall t, \quad X_t X_{t-1} &= \sum_{i=0}^{+\infty} \sum_{j=0}^{+\infty} \varphi^{i+j} \varepsilon_{t-i} \varepsilon_{t-1-j} \quad \text{(q.c.)} \\ &= f \left(\varepsilon_t, \varepsilon_{t-1}, \cdots \right) \end{aligned}$$

ou seja, $X_t X_{t-1}$ é, para todo o $t \in \mathbb{Z}$, uma função mensurável do processo ergódico $(\varepsilon_t, t \in \mathbb{Z})$.

Do mesmo modo,

$$\frac{1}{T} V_T = \frac{1}{T} \sum_{t=1}^{T} X_{t-1}^2 \underset{T \to +\infty}{\overset{q.c.,P}{\longrightarrow}} \frac{\sigma^2}{1 - \varphi^2}.$$

Como

$$\frac{1}{T} \left(V_T^* - V_T \right) \underset{T \to +\infty}{\overset{P}{\longrightarrow}} 0$$

podemos afirmar que

$$\frac{1}{T} V_T^* = \frac{1}{T} \sum_{t=1}^{T} (X_{t-1}^*)^2 = \frac{1}{T} V_T + o_T(1) \underset{T \to +\infty}{\overset{q.c.,P}{\longrightarrow}} \frac{\sigma^2}{1 - \varphi^2}$$

onde $o_T(1)$ é tal que $\lim\limits_{T \to +\infty} o_T(1) = 0$.

Assim

$$\lim_{T \to \infty} (P) \left(\frac{U_T}{V_T} - \frac{U_T^*}{V_T^*} \right) = \lim_{T \to \infty} (P) \frac{U_T \left(V_T^* - V_T \right) + \left(U_T - U_T^* \right) V_T}{V_T V_T^*} = 0.$$

Podemos então concluir que

$$\hat{\varphi}_T - \hat{\varphi}_T^* \xrightarrow[T\to+\infty]{P} 0,$$

isto é, as propriedades assintóticas em probabilidade do estimador $\hat{\varphi}_T$ são as mesmas quer para o processo X_t quer para o X_t^*.

De facto, a convergência anterior permite escrever, por exemplo,

$$\lim_{T\to\infty}(P)\hat{\varphi}_T\left(1-\frac{\hat{\varphi}_T^*}{\hat{\varphi}_T}\right)=0.$$

Assim se

$$\lim_{T\to\infty}(P)\hat{\varphi}_T=\varphi<+\infty$$

necessariamente

$$\lim_{T\to\infty}(P)\frac{\hat{\varphi}_T^*}{\hat{\varphi}_T}=1,\ \text{isto é,}\ \frac{\hat{\varphi}_T^*}{\hat{\varphi}_T}=1+o_T(1).$$

A mesma conclusão poderia tirar-se se assumíssemos a convergência para φ de $\hat{\varphi}_T^*$.

Analisemos agora a diferença $\hat{\sigma}_T^2-(\hat{\sigma}_T^*)^2$, supondo que $\hat{\varphi}_T \xrightarrow[T\to+\infty]{P} \varphi_T$. Temos

$$\hat{\sigma}_T^2=\frac{1}{T}\sum_{t=1}^{T}\left(X_t^2-2\hat{\varphi}_TX_{t-1}X_t+\hat{\varphi}_T^2X_{t-1}^2\right)+\frac{(1-\hat{\varphi}_T^2)X_1^2}{T}$$

e

$$(\hat{\sigma}_T^*)^2=\frac{1}{T}\sum_{t=1}^{T}\left[(X_t^*)^2-2\hat{\varphi}_T^*X_{t-1}^*X_t^*+(\hat{\varphi}_T^*)^2(X_{t-1}^*)^2\right].$$

Então

$$\begin{aligned}\hat{\sigma}_T^2-(\hat{\sigma}_T^*)^2=&\frac{1}{T}\left[\sum_{t=1}^{T}\left(X_t^2-(X_t^*)^2\right)\right]-\frac{2}{T}\sum_{t=1}^{T}\left[\hat{\varphi}_TX_{t-1}X_t-\hat{\varphi}_T^*X_{t-1}^*X_t^*\right]\\&+\frac{1}{T}\sum_{t=1}^{T}\left[\hat{\varphi}_T^2X_{t-1}^2-(\hat{\varphi}_T^*)^2(X_{t-1}^*)^2\right]+\frac{(1-\hat{\varphi}_T^2)X_1^2}{T}.\qquad\textbf{(3.6)}\end{aligned}$$

Ora

$$\frac{1}{T}\sum_{t=1}^{T}\left(X_t^2-(X_t^*)^2\right)\xrightarrow[T\to+\infty]{P}0\quad\text{e}\quad\frac{(1-\hat{\varphi}_T^2)X_1^2}{T}\xrightarrow[T\to+\infty]{P}0.$$

Analisemos a segunda parcela de **(3.6)**.

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\left[\hat{\varphi}_TX_{t-1}X_t-\hat{\varphi}_T^*X_{t-1}^*X_t^*\right]&=\frac{1}{T}\hat{\varphi}_T\sum_{t=1}^{T}\left[X_{t-1}X_t-(1+o(1))X_{t-1}^*X_t^*\right]\\&=\hat{\varphi}_T\frac{1}{T}\left[\sum_{t=1}^{T}\left(X_{t-1}X_t-X_{t-1}^*X_t^*\right)\right]+\hat{\varphi}_T\frac{o_T(1)}{T}\sum_{t=1}^{T}X_{t-1}^*X_t^*.\end{aligned}$$

Como

$$\hat{\varphi}_T \xrightarrow[T\to+\infty]{P} \varphi_T$$

$$\frac{1}{T}\left[\sum_{t=1}^{T}\left(X_{t-1}X_t - X_{t-1}^* X_t^*\right)\right] \xrightarrow[T\to+\infty]{P} 0$$

$$\frac{1}{T}\sum_{t=1}^{T} X_{t-1}^* X_t^* \xrightarrow[T\to+\infty]{P} \frac{\sigma^2\varphi}{1-\varphi^2}$$

e

$$\lim_{T\to\infty} \frac{o_T(1)}{T} = 0,$$

concluímos que

$$\frac{1}{T}\sum_{t=1}^{T}\left[\hat{\varphi}_T X_{t-1}X_t - \hat{\varphi}_T^* X_{t-1}^* X_t^*\right] \xrightarrow[T\to+\infty]{P} 0.$$

Para a terceira parcela de **(3.6)**, temos

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\left[\hat{\varphi}_T^2 X_{t-1}^2 - (\hat{\varphi}_T^*)^2 (X_{t-1}^*)^2\right] &= \hat{\varphi}_T^2 \frac{1}{T}\sum_{t=1}^{T}\left(X_{t-1}^2 - (1+o_T(1))^2 (X_{t-1}^*)^2\right) \\ &= \hat{\varphi}_T^2 \frac{1}{T}\sum_{t=1}^{T}\left(X_{t-1}^2 - (X_{t-1}^*)^2\right) + \hat{\varphi}_T^2 \frac{o_T(1)}{T}\sum_{t=1}^{T}(X_{t-1}^*)^2,\end{aligned}$$

pelo que concluímos, de forma análoga,

$$\frac{1}{T}\sum_{t=1}^{T}\left[\hat{\varphi}_T^2 X_{t-1}^2 - (\hat{\varphi}_T^*)^2 (X_{t-1}^*)^2\right] \xrightarrow[T\to+\infty]{P} 0.$$

Então a equação **(3.6)** converge em probabilidade para zero quando $T \to +\infty$ e portanto

$$\hat{\sigma}_T^2 - (\hat{\sigma}_T^*)^2 \xrightarrow[T\to+\infty]{P} 0$$

$$\Rightarrow (\hat{\sigma}_T^*)^2 = \hat{\sigma}_T^2 + o_T(1).$$

Provámos, então, que as propriedades assintóticas em probabilidade dos estimadores $\hat{\sigma}_T$ e $\hat{\sigma}_T^*$ são iguais.

3.2. Consistência dos estimadores

Nesta secção iremos estudar a consistência dos estimadores da máxima verosimilhança para os modelos **(A)** e **(B)**. Como na secção anterior provamos que tais estimadores são assintóticamente equivalentes, torna-se mais conveniente trabalhar com o processo estacionário, isto é, quando $\mathcal{T} = \mathbb{Z}$.

Como vimos, os e.m.v. associados ao modelo **(B)** são

$$\begin{cases} \hat{\varphi}_T = \dfrac{\sum\limits_{t=1}^{T} X_t X_{t-1}}{\sum\limits_{t=1}^{T} X_{t-1}^2} \\ \\ \hat{\sigma}_T^2 = \dfrac{\sum\limits_{t=1}^{T} \left(X_t - \hat{\varphi}_T X_{t-1}\right)^2}{T} + \dfrac{(1-\hat{\varphi}_T^2)X_1^2}{T} \end{cases}.$$

O estimador $\hat{\varphi}_T$ pode ser escrito na forma

$$\begin{aligned} \hat{\varphi}_T &= \frac{\sum\limits_{t=1}^{T} \left(\varphi X_{t-1} + \varepsilon_t\right) X_{t-1}}{\sum\limits_{t=1}^{T} X_{t-1}^2} \\ &= \varphi \frac{\sum\limits_{t=1}^{T} X_{t-1}^2}{\sum\limits_{t=1}^{T} X_{t-1}^2} + \frac{\sum\limits_{t=1}^{T} X_{t-1}\varepsilon_t}{\sum\limits_{t=1}^{T} X_{t-1}^2} \\ &= \varphi + \frac{\sum\limits_{t=1}^{T} X_{t-1}\varepsilon_t}{\sum\limits_{t=1}^{T} X_{t-1}^2}. \end{aligned}$$

Então

$$\hat{\varphi}_T - \varphi = \frac{\sum\limits_{t=1}^{T} X_{t-1}\varepsilon_t}{\sum\limits_{t=1}^{T} X_{t-1}^2}.$$

Vejamos o comportamento em probabilidade de $\dfrac{1}{T}\sum\limits_{t=1}^{T} X_{t-1}\varepsilon_t$ quando $T \to +\infty$.

Ora, $(X_{t-1}\varepsilon_t, t \in \mathbb{Z})$ é um processo ergódico pois

$$\begin{aligned} \forall t, \quad X_{t-1}\varepsilon_t &= \sum_{i=0}^{+\infty} \varphi^i \varepsilon_{t-i}\varepsilon_t, \quad q.c. \\ &= f\left(\varepsilon_t, \varepsilon_{t-1}, \cdots\right) \end{aligned}$$

ou seja, $X_{t-1}\varepsilon_t$ é, para todo $t \in \mathbb{Z}$, uma função mensurável do processo ergódico $(\varepsilon_t, t \in \mathbb{Z})$.

Assim, tendo em conta o Teorema ergódico,

$$\frac{1}{T}\sum_{t=1}^{T} X_{t-1}\varepsilon_t \underset{T\to+\infty}{\overset{q.c.,P}{\longrightarrow}} E\left(X_{t-1}\varepsilon_t\right).$$

Ora,

$$E\left(X_{t-1}\varepsilon_t\right) = E\left(X_{t-1}\right)E\left(\varepsilon_t\right) = 0$$

pois $X_{t-1} = f\left(\varepsilon_{t-1}, \varepsilon_{t-2}, \cdots\right)$ e portanto independente de ε_t.

Assim

$$\frac{1}{T}\sum_{t=1}^{T} X_{t-1}\varepsilon_t \xrightarrow[T\to+\infty]{P} 0.$$

Quanto a $\frac{1}{T}\sum_{t=1}^{T} X_{t-1}^2$ já foi provado na secção anterior que

$$\frac{1}{T}\sum_{t=1}^{T} X_{t-1}^2 \xrightarrow[T\to+\infty]{P} \frac{\sigma^2}{1-\varphi^2}.$$

Concluímos então que

$$\hat{\varphi}_T - \varphi \xrightarrow[T\to+\infty]{P} 0$$

$$\Rightarrow \hat{\varphi}_T \xrightarrow[T\to+\infty]{P} \varphi$$

e portanto $\hat{\varphi_T}$ é consistente em probabilidade.

Por fim, provemos que $\hat{\sigma}_T^2$ também é um estimador consistente em probabilidade. Temos

$$\begin{aligned}
\hat{\sigma}_T^2 &= \frac{\sum_{t=1}^{T}\left(X_t - \hat{\varphi}_T X_{t-1}\right)^2}{T} + \frac{(1-\hat{\varphi}_T^2)X_1^2}{T} \\
&= \frac{1}{T}\sum_{t=1}^{T}\left(X_t^2 - 2\hat{\varphi}_T X_{t-1}X_t + \hat{\varphi}_T^2 X_{t-1}^2\right) + \frac{(1-\hat{\varphi}_T^2)X_1^2}{T}.
\end{aligned}$$

Tendo em conta que $\hat{\varphi}_T \xrightarrow[T\to+\infty]{P} \varphi$ e usando, sempre que necessário, o teorema ergódico, vem

$$\frac{(1-\hat{\varphi}_T^2)X_1^2}{T} \xrightarrow[T\to+\infty]{P} 0$$

$$2\hat{\varphi}_T \frac{1}{T}\sum_{t=1}^{T} X_{t-1}X_t \xrightarrow[T\to+\infty]{P} \frac{2\varphi^2}{1-\varphi^2}\sigma^2$$

$$\hat{\varphi}_T^2 \frac{1}{T}\sum_{t=1}^{T} X_{t-1}^2 \xrightarrow[T\to+\infty]{P} \frac{\varphi^2\sigma^2}{1-\varphi^2}.$$

e

$$\frac{1}{T}\sum_{t=1}^{T} X_t^2 \xrightarrow[T\to+\infty]{P} \frac{\sigma^2}{1-\varphi^2},$$

podemos então concluir que

$$\hat{\sigma}_T^2 \xrightarrow[T\to+\infty]{P} \frac{\sigma^2}{1-\varphi^2} - 2\varphi\frac{\sigma^2\varphi}{1-\varphi^2} + \frac{\varphi^2\sigma^2}{1-\varphi^2} = \frac{\sigma^2 - \varphi^2\sigma^2}{1-\varphi^2}$$

$$\Rightarrow \hat{\sigma}_T^2 \underset{T \to +\infty}{\overset{P}{\longrightarrow}} \sigma^2.$$

e portanto $\hat{\sigma}_T^2$ é consistente em probabilidade.

Capítulo 4

Lei limite do estimador do parâmetro do processo AR(1) explosivo

4.1. Estudo da lei limite

Recordemos o modelo **(A)** que foi estudado no capítulo 1.

$$\begin{cases} X_t = \varphi X_{t-1} + \varepsilon_t, \quad t \in \mathbb{N} \\ X_0 = 0 \end{cases} \qquad \textbf{(4.1)}$$

com $\varphi \in \mathbb{R} \setminus \{0\}$ e uma sucessão de variáveis aleatórias $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ centradas, não correlacionadas e de variância σ^2 ($\sigma^2 > 0$).

Neste capítulo vamos dedicar-nos ao estudo da distribuição assintótica do estimador de φ no caso em que $| \varphi |> 1$ e $X_0 = 0$. [1]

Limitamos o estudo a este caso por se revelar diferente dos casos habitualmente encontrados em que a lei limite é frequentemente gaussiana.

Em particular, se $| \varphi |< 1$ a distribuição limite de $\sqrt{T}\,(\hat{\varphi}_T - \varphi)$ é normal com média 0 e variância $1 - \varphi^2$. (Anderson, 1959, p.686)

Finalmente, iremos ilustrar os resultados obtidos usando dados simulados e o teste de ajustamento do χ^2. Provaremos, desse modo, que a distribuição limite do estimador devidamente normalizado é a distribuição de Cauchy standard.

Como vimos o estimador dos mínimos quadrados e da máxima verosimilhança para φ é

$$\hat{\varphi}_T = \frac{\sum_{i=2}^{T} X_i X_{i-1}}{\sum_{i=2}^{T} X_{i-1}^2} \qquad \textbf{(4.2)}.$$

[1]Em Hwang (2007) pode ser encontrada uma generalização deste estudo com $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ seguindo um modelo condicionalmente heteroscedástico.

Usando o teorema seguinte obtêm-se a lei limite do estimador $\hat{\varphi}_T$ convenientemente normalizado.

Teorema: Seja $X = (X_t, t \in \mathbb{N})$ o processo estocástico real definido por **(4.1)** com $|\varphi| > 1$ e $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ uma sucessão de variáveis aleatórias tal que $\varepsilon_t \sim N\left(0, \sigma^2\right)$. Seja $\hat{\varphi}_T$ o estimador dos mínimos quadrados e da máxima verosimilhança para o parâmetro φ. A distribuição limite de $\left(\varphi^2 - 1\right)^{-1} \varphi^T \left(\hat{\varphi}_T - \varphi\right)$ é a distribuição de Cauchy standard.[1]

Prova:

Como já foi provado, na secção **(3.2)**, a diferença entre o e.m.v. e o valor de φ é dado por

$$\hat{\varphi}_T - \varphi = \frac{\sum\limits_{i=1}^{T} X_{i-1}\varepsilon_i}{\sum\limits_{i=1}^{T} X_{i-1}^2} \qquad \textbf{(4.3)}$$

Recorde-se que, como foi provado no capítulo 2, o processo X pode escrever-se na seguinte forma

$$X_t = \sum_{j=0}^{t-1} \varphi^j \varepsilon_{t-j}, \quad \forall t \in \mathbb{N}.$$

Vem que

$$\begin{aligned} X_t &= \sum_{j=0}^{t-1} \varphi^j \varepsilon_{t-j} \\ &= \sum_{i=1}^{t} \varphi^{t-i} \varepsilon_i \\ &= \varphi^t \sum_{i=1}^{t} \varphi^{-i} \varepsilon_i \\ &= \varphi^t Y_t \end{aligned}$$

com $Y_t = \sum\limits_{i=1}^{t} \varphi^{-i} \varepsilon_i$.

[1] Uma lei de probabilidade Q sobre $(\mathbb{R}, \mathcal{B})$ é de Cauchy de parâmetro (α, λ), $\alpha \in \mathbb{R}$, $\lambda > 0$, se é absolutamente contínua de densidade

$$f(x) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (x-\alpha)^2}, \quad x \in \mathbb{R}.$$

Se a v.a.r. X segue a lei de Cauchy de parâmetro (α, λ) escrevemos $X \sim C(\alpha, \lambda)$. A lei diz-se de Cauchy standard se $\alpha = 0$ e $\lambda = 1$. (Gonçalves, E., N. Mendes-Lopes, 2003, p. 304)

Prova-se que

$$Y_t \underset{t\to+\infty}{\overset{q.c.}{\longrightarrow}} Y$$

e consequentemente,

$$\varphi^{-t} X_t \underset{t\to+\infty}{\overset{q.c.}{\longrightarrow}} Y$$

com $Y = \sum\limits_{i=1}^{+\infty} \varphi^{-i}\varepsilon_i$.

Basta ver que a série $\sum\limits_{i=1}^{+\infty} \varphi^{-i}\varepsilon_i$ é q.c. convergente. De facto, pelo critério de Cauchy

$$\lim_{i\to+\infty} \left(|\varphi|^{-i} \mid \varepsilon_i \mid\right)^{1/i} \overset{q.c.}{=} \frac{1}{\mid \varphi \mid} \lim_{i\to+\infty} e^{\frac{1}{i}\log|\varepsilon_0|}, \text{ q.c., P.}$$

pois como ε_i é uma sucessão de variáveis aleatórias identicamente distribuídas

$$P\left(\lim_{i\to+\infty} \frac{1}{i} \log \mid \varepsilon_i \mid = a\right) = P\left(\lim_{i\to+\infty} \frac{1}{i} \log \mid \varepsilon_0 \mid = a\right).$$

Ora

$$P\left(\left\{\lim_{i\to+\infty} \frac{1}{i} \log \mid \varepsilon_0 \mid = 0\right\}\right) = 1.$$

Concluímos assim que

$$\lim_{i\to+\infty} \left(|\varphi|^{-i} \mid \varepsilon_i \mid\right)^{1/i} = \frac{1}{\mid \varphi \mid}, \text{ q.c. .}$$

Como $\frac{1}{\mid \varphi \mid} < 1$, o critério de Cauchy para séries numéricas de termos positivos permite então afirmar que $\sum\limits_{i=1}^{+\infty} \varphi^{-i}\varepsilon_i$ é q.c. convergente.

Analisemos a ordem de convergência de Y_t para Y.[1]

$$\begin{aligned}
|Y_t - Y| &= \left|\sum_{i=1}^{+\infty} \varphi^{-i}\varepsilon_i - \sum_{i=1}^{t} \varphi^{-i}\varepsilon_i\right| \\
&= \left|\sum_{i=t+1}^{+\infty} \varphi^{-i}\varepsilon_i\right| \\
&= O_p\left(\mid \varphi \mid^{-t}\right),
\end{aligned}$$

[1]Diz-se que f é da ordem de g na vizinhança de x_0, e escrevemos $f = O(g)$, se

$$0 < \lim_{x\to x_0} \frac{f(x)}{g(x)} < +\infty.$$

(Gonçalves, E., N. Mendes-Lopes, 2000, p. 187)

Esta definição alarga-se às convergências estocásticas sendo a ordem de convergência considerada no sentido de convergência usada.

de facto,

$$
\begin{aligned}
0 < \frac{\left|\sum\limits_{i=t+1}^{+\infty} \varphi^{-i}\varepsilon_i\right|}{\mid \varphi \mid^{-t}} &\leq \sum_{i=t+1}^{+\infty} \mid \varphi \mid^{-i+t} \mid \varepsilon_i \mid \\
&= \sum_{j=1}^{+\infty} \mid \varphi \mid^{-j} \mid \varepsilon_{t+j} \mid \\
&< +\infty, \text{ q.c., P}
\end{aligned}
$$

Analisemos agora o comportamento assintótico do denominador de **(4.3)** considerando

$$
\begin{aligned}
\varphi^{-2T} \sum_{t=1}^{T} X_{t-1}^2 &= \varphi^{-2T} \sum_{t=1}^{T} \varphi^{2(t-1)} Y_{t-1}^2 \\
&= \varphi^{-2T} \sum_{t=1}^{T} \varphi^{2(t-1)} \left(Y_{t-1} - Y + Y\right)^2 \\
&= \sum_{t=1}^{T} \varphi^{2(t-T)-2} Y^2 + 2 \sum_{t=1}^{T} \varphi^{2(t-T)-2} \left(Y_{t-1} - Y\right) Y + \\
&+ \sum_{t=1}^{T} \varphi^{2(t-T)-2} \left(Y_{t-1} - Y\right)^2 \\
&= \sum_{j=0}^{T-1} \varphi^{-2j-2} Y^2 + 2 \sum_{t=1}^{T} \varphi^{2(t-T)-2} \left(Y_{t-1} - Y\right) Y + \\
&+ \sum_{t=1}^{T} \varphi^{2(t-T)-2} \left(Y_{t-1} - Y\right)^2 . \quad \textbf{(4.4)}
\end{aligned}
$$

Analisemos a $3^{\underline{a}}$ parcela de **(4.4)**.

Ora

$$
\begin{aligned}
Y_{t-1} - Y &= \sum_{j=1}^{t-1} \varphi^{-j}\varepsilon_j - \sum_{j=1}^{+\infty} \varphi^{-j}\varepsilon_j \\
&= - \sum_{j=t}^{+\infty} \varphi^{-j}\varepsilon_j ,
\end{aligned}
$$

então

$$\begin{aligned}
E\left[(Y_{t-1}-Y)^2\right] &= E\left(\sum_{j=t}^{+\infty}\varphi^{-j}\varepsilon_j\right)^2\\
&= E\left(\sum_{j=t}^{+\infty}\sum_{k=t}^{+\infty}\varphi^{-(j+k)}\varepsilon_j\varepsilon_k\right)\\
&= \sum_{j=t}^{+\infty}\varphi^{-2j}E\left(\varepsilon_j^2\right)\\
&= \frac{\varphi^{-2t}}{1-\dfrac{1}{\varphi^2}}\sigma^2\\
&= O\left(|\varphi|^{-2t}\right),
\end{aligned}$$

uma vez que

$$\lim_{t\to+\infty}\sigma^2\frac{\varphi^{-2t}}{1-\dfrac{1}{\varphi^2}}\frac{1}{|\varphi|^{-2t}}=\frac{\sigma^2}{1-\dfrac{1}{\varphi^2}}<+\infty.$$

Assim o 3º termo do 2º membro da igualdade **(4.4)** é tal que

$$\sum_{t=1}^{T}\varphi^{2(t-T)-2}\left(Y_{t-1}-Y\right)^2=O\left(T\mid\varphi\mid^{-2T}\right).$$

De facto,

$$\begin{aligned}
0 &< \lim_{T\to+\infty}\frac{\sum\limits_{t=1}^{T}\varphi^{2(t-T)-2}\left(Y_{t-1}-Y\right)^2}{T\mid\varphi\mid^{-2T}}\\
&= \lim_{T\to+\infty}\varphi^{-2T-2}\frac{1}{T}\sum_{t=1}^{T}\varphi^{2t}\frac{O_p\left(\mid\varphi\mid^{-2t}\right)}{\mid\varphi\mid^{-2T}}\\
&= \lim_{T\to+\infty}\frac{\varphi^{-2}}{T}\sum_{t=1}^{T}\frac{O_p\left(\mid\varphi\mid^{-2t}\right)}{\varphi^{-2t}}\\
&< \lim_{T\to+\infty}\frac{\varphi^{-2}}{T}TM\\
&= \varphi^{-2}M<+\infty,
\end{aligned}$$

pois $\lim\limits_{T\to+\infty}\dfrac{O_p\left(\mid\varphi\mid^{-2t}\right)}{\mid\varphi\mid^{-2t}}<+\infty$, logo a sucessão é limitada a partir de uma certa ordem.

Consideremos agora o 2º termo do 2º membro de **(4.4)**. Ora

$$
\begin{aligned}
V\left(\sum_{t=1}^{T}\varphi^{2(t-T)-2}\left(Y_{t-1}-Y\right)\right) &= V\left(\sum_{t=1}^{T}\varphi^{2(t-T)-2}\sum_{j=t}^{+\infty}\varphi^{-j}\varepsilon_j\right)\\
&= V\left(\sum_{t=1}^{T}\sum_{j=t}^{+\infty}\varphi^{2(t-T)-2}\varphi^{-j}\varepsilon_j\right)\\
&= \sum_{t=1}^{T}\sum_{j=t}^{+\infty}\varphi^{4(t-T)-4}\varphi^{-2j}\sigma^2\\
&= \sigma^2\sum_{t=1}^{T}\varphi^{4(t-T)-4}\sum_{j=t}^{+\infty}\varphi^{-2j}\\
&= \sigma^2\varphi^{-4T-4}\sum_{t=1}^{T}\varphi^{4t}\frac{\varphi^{-2t}}{1-\frac{1}{\varphi^2}}\\
&= \sigma^2\frac{\varphi^{-4T-4}}{1-\varphi^{-2}}\sum_{t=1}^{T}\varphi^{2t}\\
&= \sigma^2\frac{\varphi^{-4T-4}}{1-\varphi^{-2}}\frac{\varphi^2-\varphi^{2T+2}}{1-\varphi^2}\\
&= \sigma^2\frac{\varphi^{-4T-2}-\varphi^{-2T-2}}{\left(1-\varphi^{-2}\right)\left(1-\varphi^2\right)}.
\end{aligned}
$$

Como

$$
\lim_{T\to+\infty}\frac{\sigma^2}{\left(1-\varphi^{-2}\right)\left(1-\varphi^2\right)}\frac{\varphi^{-4T-2}-\varphi^{-2T-2}}{\varphi^{-2T}} = \lim_{T\to+\infty}\frac{\sigma^2}{\left(1-\varphi^{-2}\right)\left(1-\varphi^2\right)}\left(\varphi^{-2T-2}-\varphi^{-2}\right) < +\infty,
$$

o 2º termo do 2º membro é tal que

$$
2Y\sum_{t=1}^{T}\varphi^{2(t-T)-2}\left(Y_{t-1}-Y\right) = 2YO_{L^2}\left(\mid\varphi\mid^{-2T}\right) = 2YO_p\left(\mid\varphi\mid^{-2T}\right).
$$

Assim retomando a expressão em **(4.4)** vem

$$
\begin{aligned}
\varphi^{-2T}\sum_{t=1}^{T}X_{t-1}^2 &= Y^2\varphi^{-2}\frac{1-\varphi^{-2T}}{1-\varphi^{-2}} + 2YO_p\left(\mid\varphi\mid^{-2T}\right) + O_p\left(T\mid\varphi\mid^{-2T}\right)\\
&= Y^2\frac{1}{\varphi^2-1} - \frac{\varphi^2\varphi^{-2T}}{\varphi^2-1} + 2YO_p\left(\mid\varphi\mid^{-2T}\right) + O_p\left(T\mid\varphi\mid^{-2T}\right)\\
&= Y^2\frac{1}{\varphi^2-1} - O_p\left(\mid\varphi\mid^{-T}\right). \quad \mathbf{(4.5)}
\end{aligned}
$$

De facto,

$$
\begin{aligned}
&\lim_{T\to+\infty}\left[\frac{Y^2\varphi^{-2T}}{\left(\varphi^2-1\right)\mid\varphi\mid^{-T}} + 2Y\frac{O_p\left(\mid\varphi\mid^{-2T}\right)}{\mid\varphi\mid^{-T}} + \frac{O_p\left(T\mid\varphi\mid^{-2T}\right)}{\mid\varphi\mid^{-T}}\right]\\
&= \lim_{T\to+\infty}\left[\frac{Y^2}{\varphi^2-1}\mid\varphi\mid^{-T} + 2Y\mid\varphi\mid^{-T}\frac{O_p\left(\mid\varphi\mid^{-2T}\right)}{\mid\varphi\mid^{-2T}} + T\mid\varphi\mid^{-T}\frac{O_p\left(T\mid\varphi\mid^{-2T}\right)}{T\mid\varphi\mid^{-2T}}\right] < +\infty,
\end{aligned}
$$

em probabilidade.

Relativamente ao numerador de **(4.3)** tem-se

$$\begin{aligned}
\varphi^{-T}\sum_{t=1}^{T}X_{t-1}\varepsilon_t &= \varphi^{-T}\sum_{t=1}^{T}\varphi^{(t-1)}Y_{t-1}\varepsilon_t\\
&= \varphi^{-T}\sum_{t=1}^{T}\varphi^{(t-1)}\left(Y_{t-1}-Y+Y\right)\varepsilon_t\\
&= \varphi^{-T}Y\sum_{t=1}^{T}\varphi^{(t-1)}\varepsilon_t+\varphi^{-T}\sum_{t=1}^{T}\left(Y_{t-1}-Y\right)\varepsilon_t\\
&= Y\sum_{t=1}^{T}\varphi^{t-T-1}\varepsilon_t+\varphi^{-T}\sum_{t=1}^{T}\left(Y_{t-1}-Y\right)\varepsilon_t. \qquad \textbf{(4.6)}
\end{aligned}$$

Analisemos o 2º termo do 2º membro de **(4.6)**.

Como

$$\begin{aligned}
E\left(|Y_{t-1}-Y|\mid\varepsilon_t\mid\right) &\leq E\left(\sum_{j=t}^{+\infty}\mid\varphi\mid^{-j}\mid\varepsilon_j\mid\mid\varepsilon_t\mid\right)\\
&= \sum_{j=t+1}^{+\infty}\mid\varphi\mid^{-j}\mid E\left(\mid\varepsilon_j\mid\right)E\left(\mid\varepsilon_t\mid\right)+\mid\varphi\mid^{-t}E\left(\varepsilon_t^2\right)\\
&\leq \sigma^2\sum_{j=t}^{+\infty}\mid\varphi\mid^{-j}\\
&= \sigma^2\frac{\mid\varphi\mid^{-t}}{1-\mid\varphi\mid^{-1}}
\end{aligned}$$

podemos concluir que

$$|Y_{t-1}-Y|\mid\varepsilon_t\mid=O_p\left(\mid\varphi\mid^{-t}\right).$$

Então

$$\begin{aligned}
\lim_{T\to+\infty}\frac{\left|\varphi^{-T}\sum\limits_{t=1}^{T}\left(Y_{t-1}-Y\right)\varepsilon_t\right|}{T\mid\varphi\mid^{-T}} &\leq \lim_{T\to+\infty}\frac{1}{T}\sum_{t=1}^{T}O_p\left(\mid\varphi\mid^{-t}\right)\\
&\leq \lim_{T\to+\infty}\frac{c}{T}\sum_{t=1}^{T}\mid\varphi\mid^{-t}=c<+\infty
\end{aligned}$$

e portanto

$$\varphi^{-T}\sum_{t=1}^{T}\left(Y_{t-1}-Y\right)\varepsilon_t=O_p\left(T\mid\varphi\mid^{-T}\right).$$

Assim

$$\varphi^{-T}\sum_{t=1}^{T}X_{t-1}\varepsilon_t=Y\sum_{t=1}^{T}\varphi^{t-T-1}\varepsilon_t+O_p\left(T\mid\varphi\mid^{-T}\right). \qquad \textbf{(4.7)}$$

Passemos agora ao estudo da lei limite de $\left(\varphi^2-1\right)^{-1}\varphi^T\left(\hat{\varphi}_T-\varphi\right).$

Ora

$$\begin{aligned}
\varphi^T\left(\hat{\varphi}_T-\varphi\right) &= \frac{\varphi^{-T}\sum\limits_{t=1}^{T} X_{t-1}\varepsilon_t}{\varphi^{-2T}\sum\limits_{t=1}^{T} X_{t-1}^2} \\
&= \frac{Y\sum\limits_{t=1}^{T}\varphi^{t-T-1}\varepsilon_t+O_p\left(T\mid\varphi\mid^{-T}\right)}{Y^2\dfrac{1}{\varphi^2-1}+O_p\left(\mid\varphi\mid^{-T}\right)} \qquad \text{usando } \textbf{(4.5)} \text{ e } \textbf{(4.7)} \\
&= \frac{\sum\limits_{t=1}^{T}\varphi^{t-T-1}\varepsilon_t}{Y\dfrac{1}{\varphi^2-1}+\dfrac{1}{Y}O_p\left(\mid\varphi\mid^{-T}\right)}+\frac{O_p\left(T\mid\varphi\mid^{-T}\right)}{Y^2\dfrac{1}{\varphi^2-1}+O_p\left(\mid\varphi\mid^{-T}\right)} \\
&= \frac{\sum\limits_{t=1}^{T}\varphi^{t-T-1}\varepsilon_t}{Y\dfrac{1}{\varphi^2-1}+o_P(\dfrac{1}{T})}+o_P(\frac{1}{T})
\end{aligned}$$

pois

$$\frac{\dfrac{O_p\left(T\mid\varphi\mid^{-T}\right)}{Y^2\dfrac{1}{\varphi^2-1}+O_p\left(\mid\varphi\mid^{-T}\right)}}{\dfrac{1}{T}}=\frac{\dfrac{O_p\left(T\mid\varphi\mid^{-T}\right)}{T\mid\varphi\mid^{-T}}}{\dfrac{Y^2\dfrac{1}{\varphi^2-1}}{T^2\mid\varphi\mid^{-T}}+\dfrac{O_p\left(\mid\varphi\mid^{-T}\right)}{T^2\mid\varphi\mid^{-T}}}\underset{T\to+\infty}{\longrightarrow}0$$

e

$$\frac{\dfrac{1}{Y}O_p\left(\mid\varphi\mid^{-T}\right)}{\dfrac{1}{T}}\underset{T\to+\infty}{\longrightarrow}0$$

então podemos escrever

$$\begin{aligned}
\varphi^T\left(\hat{\varphi}_T-\varphi\right) &= \frac{\sum\limits_{j=1}^{T}\varphi^{-j}\varepsilon_{T-j+1}}{\dfrac{1}{\varphi^2-1}\sum\limits_{i=1}^{+\infty}\varphi^{-i}\varepsilon_i+o_p\left(\dfrac{1}{T}\right)}+o_p\left(\frac{1}{T}\right) \\
&= \frac{\sum\limits_{j<T/2}\varphi^{-j}\varepsilon_{T-j+1}+O_p\left(\mid\varphi\mid^{-T}\right)}{\dfrac{1}{\varphi^2-1}\sum\limits_{i<T/2}\varphi^{-i}\varepsilon_i+\underbrace{O_p\left(\mid\varphi\mid^{-T}\right)+o_p\left(\dfrac{1}{T}\right)}_{O_p(|\varphi|^{-T})}}+o_p\left(\frac{1}{T}\right). \qquad \textbf{(4.8)}
\end{aligned}$$

Finalmente, multiplicando **(4.8)** por $\left(\varphi^2-1\right)^{-1}$, tem-se

$$\left(\varphi^2-1\right)^{-1}\varphi^T\left(\hat{\varphi}_T-\varphi\right)=\frac{\sum\limits_{j<T/2}\varphi^{-j}\varepsilon_{T-j+1}+O_p\left(\mid\varphi\mid^{-T}\right)}{\sum\limits_{i<T/2}\varphi^{-i}\varepsilon_i+O_p\left(\mid\varphi\mid^{-T}\right)}+o_p\left(\frac{1}{T}\right)$$

Assim a lei limite de $\left(\varphi^2-1\right)^{-1}\varphi^T\left(\hat{\varphi}_T-\varphi\right)$ é a lei do quociente de variáveis

$$\frac{\sum\limits_{j<T/2}\varphi^{-j}\varepsilon_{T-j+1}}{\sum\limits_{i<T/2}\varphi^{-i}\varepsilon_i}$$

que são independentes entre si. Além disso, ambas as variáveis são normais centradas e com a mesma variância, logo a lei limite é de Cauchy standard.

4.2. Exemplo de simulação

O teste de ajustamento do χ^2 destina-se a ensaiar a hipótese de adequação da lei de uma população a uma determinada lei de probabilidade.

Suponhamos que dispomos de uma amostra, $(X_1, \cdots, X_n)$, de uma variável aleatória real X de lei Q desconhecida. Pretendemos testar se tal lei coincide com uma dada lei Q_0. Neste modelo estatístico queremos testar

$$\mathcal{H}_0 : Q = Q_0 \quad \text{contra } \mathcal{H}_1 : Q \neq Q_0.$$

Para testar estas hipóteses vamos utilizar a estatística

$$T_n = \sum_{i=1}^{r} \frac{(N_i - np_i)^2}{np_i},$$

sugerida pelo teorema de Karl Pearson[1], cuja lei, sob $\mathcal{H}_0$, é a lei do χ^2_{r-1}.
Para tal, consideremos uma partição do suporte de Q em r classes

$$]a_0, a_1],]a_1, a_2], \cdots,]a_{r-1}, a_r],$$

Em seguida, calculam-se as duas quantidades que serão comparadas na estatística do teste:

- frequências esperadas, np_i, sabendo que

$$p_i = Q\left(]a_{i-1}, a_i]\right) = F_Q(a_i) - F_Q(a_{i-1})$$

- frequências observadas, N_i, isto é, a frequência de X na classe $]a_{i-1}, a_i]$.

Assim a decisão será tomada à custa do teste de região crítica, $\mathcal{C}$, definida pela condição

$$\mathcal{C} = \left\{ (x_1, \cdots, x_n) : T_n = \sum_{i=1}^{r} \frac{(N_i - np_i)^2}{np_i} > c \right\},$$

com c tal que

$$P\left(\mathcal{C} | \mathcal{H}_0\right) = \alpha.$$

Vamos, então, usar o teste de adequação do qui-quadrado ao nível de significância de $\alpha = 0.05$, para ilustrar o resultado estabelecido no teorema anterior.

[1] Teorema de Pearson - Seja $(N_1, \cdots, N_k)$ um vector aleatório sobre $\mathbb{R}^k$ seguindo a lei multinomial de parâmetro $(n, p_1, \cdots, p_k)$. Então a v.a.r.

$$V_{k,n} = \sum_{j=1}^{k} \frac{(N_i - np_j)^2}{np_j}$$

converge em lei, quando $n \to +\infty$, para a lei do χ^2_{k-1}. (Gonçalves, E., N., Mendes-Lopes, 2003, p.256)

Consideremos um processo AR(1) explosivo dado por

$$\begin{cases} X_t = \varphi X_{t-1} + \varepsilon_t, & t \in \mathbb{N} \\ X_0 = 0 \end{cases}$$

com $\mid \varphi \mid > 1$ e uma sucessão de variáveis aleatórias $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ tal que $\varepsilon_t \sim N(0,1)$.

Designando por Y a v.a.r. que representa $\left(\varphi^2 - 1\right)^{-1} \varphi^T \left(\hat{\varphi}_T - \varphi\right)$, queremos ensaiar a hipótese

$$\mathcal{H}_0 : Y \sim C(0,1).$$

Usando o Software EViews 7, com base numa amostra de 100 valores de $\hat{\varphi}_T$, geraram-se 100 valores para Y considerando-se $\varphi = 2$ e T com ordem de grandeza razoável ($T = 50$). [1]

Em seguindo, foi usado o Software SPSS, para fazer a análise estatística descritiva de Y bem como o seu histograma.

Statistics

Y

N	Valid	100
	Missing	0
Mean		1,00833337
Median		,08333350
Std. Deviation		8,434324702
Skewness		3,474
Std. Error of Skewness		,241
Kurtosis		19,368
Std. Error of Kurtosis		,478
Minimum		-17,583330
Maximum		54,833330

Figura 4.1: Análise estatística descritiva da v.a.r. $Y = \frac{1}{3} 2^{50} \left(\hat{\varphi}_T - 2\right)$

Pela análise da Figura 4.1 podemos desde já concluir que a v.a.r. Y não segue uma distribuição normal, uma vez que esta é leptocúrtica. A distribuição de Cauchy não possui momentos, no entanto a sua mediana é igual ao parâmetro de escala α, que como podemos ver se aproxima de 0.

[1] O programa usado está no Anexo A.

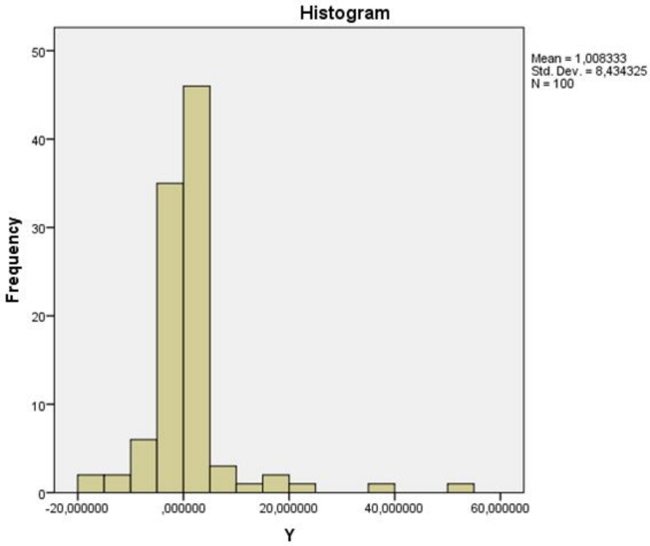

Figura 4.2: Histograma da v.a.r. Y

Após a análise do histograma foram consideradas as seguintes classes :

$$]-\infty, -5],]-5, 0],]0, 5],]5, +\infty[.$$

Em seguido, usamos o SPSS para obter os valores das frequências esperadas e das frequências observadas, que estão reunidos na seguinte tabela.

Classes	$]-\infty, -5]$	$]-5, 0]$	$]0, 5]$	$]5, +\infty[$
Frequências Esperadas (np_i)	6	44	44	6
Frequências Observadas (N_i)	10	40	41	9

Tabela 4.1: Frequências esperadas e frequências observadas

Sendo a amostra observada tal que, para todo $i \in \{1, 2, 3, 4\}$, $np_i > 5$, a utilização do teste do qui-quadrado é legítima.
Da tabela do qui-quadrado obtém-se que $k_{0.05} = 7.8$. Assim a decisão será tomada à custa do teste de região crítica, $\mathcal{C}$, definida pela condição

$$\frac{(N_1 - 6)^2}{6} + \frac{(N_2 - 44)^2}{44} + \frac{(N_3 - 44)^2}{44} + \frac{(N_4 - 6)^2}{6} > 7.8.$$

Para a amostra observada o valor da estatística de teste é 4.735, pelo que não pertence à região critica; somos então levados aceitar a $\mathcal{H}_0$, concluindo-se que $Y \sim C(0, 1)$.

O teste do qui-quadrado foi também feito usando o SPSS, obtendo-se o quadro seguinte

Test Statistics

	freq
Chi-Square	4,735[a]
df	3
Asymp. Sig.	,192

a. 0 cells (0,0%) have expected frequencies less than 5. The minimum expected cell frequency is 6,0.

Figura 4.3: Teste de adequação do qui-quadrado usando o SPSS

Através da análise da Figura 4.3 concluímos que de facto a estatística de teste coincide com a obtida anteriormente, 4.735, pelo que somos levados aceitar a hipótese $\mathcal{H}_0$.

Apêndice A

Programa usado no Capítulo 4

```
scalar nrep=100
vector(100) result
for !i = 1 to nrep
smpl @first @last
series e = @rnorm
smpl @first @first
series x=e
smpl @first+1 @last
series x=2*x(-1)+e
smpl @first @first
series aux=0
series aux1=0
smpl @first+1 @last
series aux=x*x(-1)
series aux1=x(-1)*x(-1)
scalar b =@sum(aux)
scalar h=@sum(aux1)
scalar w=b/h
scalar f = 1/3 * (2^50) * (w - 2)
result(!i)=f
next
```

Bibliografia

[1] Anderson, T.W. (1959) On asymptotic distributions of estimates of parameters of stochastic difference equations. Ann. Math. Statist. 30, 676-687.

[2] Anderson, T.W. (1971) The statistical analysis of time series, Wiley, New York.

[3] Azencott, R., D. Dacunha-Catelle (1984) Séries d'observations irrégulières, modélisation et prévision, Masson, Paris.

[4] Fuller, Wayne A. (1996) Introduction to Statistical Time Series, John Wiley and Sons. Inc., New York.

[5] Gonçalves, E., N., Mendes-Lopes (1998) Probabilidades - Princípios Teóricos, Escolar Editora, Lisboa.

[6] Gonçalves, E., N., Mendes-Lopes (2003) Estatística - Teoria Matemática e Aplicações, Escolar Editora, Lisboa.

[7] Gonçalves, E., N., Mendes-Lopes (2008) Séries temporais. Modelações lineares e não lineares, Sociedade Portuguesa de Estatística.

[8] Gouriéroux, Ch., A. Monfort (1989) Statistique et modèles économétriques, Vols. I, II, Economica, Paris.

[9] Hwang, S.Y., S. Kim, S.D. Lee, I.V. Basawa (2007) Generalized least squares estimation for explosive AR(1) processes with conditionally heteroscedastic errors, Statistics and Probability Letters 77, 1439-1448.

[10] Rosenblatt, M. (1978) Dependence and asymptotic independence for random processes, Studies in Probability Theory, Studies in Mathematics, 18, 24-44, The Mathematical Association of America.