FLORIAN KÖLBL

Aggregation of AR(2) Processes

DIPLOMARBEIT

zur Erlangung des akademischen Grades Diplomingenieur der Studienrichtung Technische Mathematik

Technische Universität Graz

Betreuer:

Univ.-Prof. István Berkes
Institut für Statistik
Technische Universität Graz

Abstract

We consider the least square estimators of the classical AR(2) process when the underlying variables are aggregated sums of independent random coefficient AR(2) models. We establish the asymptotic of the corresponding statistics and show that in general these estimators are not consistent estimators of the expected values of the autoregressive coefficients when the number of aggregated terms and the sample size tend to infinity. Furthermore the asymptotic behavior of some statistics which can be used to estimate parameters and a central limit theorem for the case that the number of aggregated terms is much larger than the number of observations is given. A method how parameters of a distribution of the random coefficients can be estimated and examples for possible distributions are given.

Keywords: random coefficient AR(2), least square, aggregation, parameter estimation, central limit theorem

Danksagung

Hiermit möchte ich mich bei meinen Eltern, meinem Bruder und meiner Freundin Sandra bedanken, die mich während meines Studiums immer unterstützt und motiviert haben. Spezieller Dank gilt meinen Betreuern István Berkes und Lajos Horváth, die es mir ermöglicht haben diese Diplomarbeit im Zuge eines Auslandsaufenthalts an der University of Utah zu verfassen, mir mit ihrem Wissen und Können tatkräftig zur Seite standen und mir vieles gelernt haben.

Acknowledgment

I want to thank my parents, my brother and my girlfriend Sandra for their support and for motivating me during my study. Special thanks go to my advisors István Berkes and Lajos Horváth who gave me the chance to write this diploma thesis in the course of a stay abroad at the University of Utah, who supported me actively with their knowledge and competence and who taught me a lot.

Florian Kölbl
Graz, 1.April 2006

Contents

1 Introduction **4**

2 Introduction to Time Series **6**
- 2.1 Basic Definitions 6
- 2.2 Important Time Series 9

3 Properties of AR(2) Processes **12**
- 3.1 Representations and Stationarity 12
- 3.2 Least Square Estimators 18

4 Aggregation of AR(2) Processes **20**
- 4.1 Motivation for Aggregation 20
- 4.2 Basic Definitions, Assumptions and Properties 21
- 4.3 Asymptotic Behavior and Properties of the Least Square Estimators 23
- 4.4 Asymptotic Results for some Useful Statistics 29
- 4.5 Differences of Covariances 32

5 Estimators for Distribution Parameters **34**
- 5.1 Ideas to Find Estimators 34
- 5.2 Estimation Method 35
- 5.3 Estimators for Beta-Distributed Coefficients 38
- 5.4 Estimators for Standard Two-Sided Power-Distributed Coefficients 42

Chapter 1

Introduction

This thesis is concerned with the aggregation of second-order autoregressive processes with random coefficients.

Aggregated observations are very common data sets in the studies of macroeconomic time series. Usually to describe such data, classical autoregression (AR) or autoregressive moving average (ARMA) models are used. Granger and Morris (1976) explained how more complicated time series models can arise from an aggregation of simpler models. For example, adding N independent AR(1) models can produce the ARMA(N, $N-1$) model. This was a starting point to consider aggregation of infinitely many simple AR models with random coefficients, which can produce a time series with long memory, see Granger (1980) and the earlier contribution by Robinson (1978). This approach was later generalized by a number of authors, see e.g. Oppenheim and Viano (2004), Zaffaroni (2004) and especially Horváth and Leipus (2005). Horváth and Leipus (2005) studied the behavior of the least square estimator under the aggregation of N independent random coefficient AR(1) models when the number of observations and the number of aggregated models tend to infinity. They have shown that the least square estimator of the classical AR(1) model can not be used as an estimator for the expected value

$\mathbb{E}a$ in the aggregated model and they proposed a consistent estimator for $\mathbb{E}a$. Moreover they gave a central limit theorem for the case when the number of observations is much larger than the number of aggregated terms. Usually the opposite case appears. In practice the number of aggregated terms is much larger than the number of observations. Moreover, in the literature the aggregation of higher order autoregressive processes is rarely discussed. These two facts have motivated the consideration of the aggregation of AR(2) models and the discussion of the case when the number of aggregated terms is much larger than the number of observations. Therefore this thesis is considering the aggregation of second-order autoregessive models, the asymptotic behavior of the least square estimators is discussed, it is shown that in general the least square estimators are no longer consistent estimators for the expected values $\mathbb{E}a_1$ and $\mathbb{E}a_2$ of the random coefficients. Furthermore we give methods how parameters of parametric distributions of the random coefficients can be estimated and we present a central limit theorem for the important case when the number of aggregated terms is much larger than the number of observations.

The structure of this thesis is the following. In Chapter 2 we give a short introduction in time series. In Chapter 3 some important properties of AR(2) processes are discussed. The asymptotic behavior of the least square estimators, the asymptotic behavior of some statistics which can be used to estimate parameters and a central limit theorem for the case that the number of aggregated terms is much larger than the number of observations is given in Chapter 4. A method how parameters of a distribution of the random coefficients can be estimated and examples for possible distributions are given in Chapter 5.

Chapter 2

Introduction to Time Series

Much of economics is concerned with modeling dynamics. There has been an explosion of research in this area in the last twenty years, as "time series econometrics" has practically come to be synonymous with "empirical macroeconomics". This section is an introduction to the basic ideas of time series analysis and stochastic processes. Of particular importance are the concepts of stationarity and autocovariance. Important examples of time series models are presented.

2.1 Basic Definitions

The first step in the analysis of a time series is the selection of a suitable mathematical model (or class of models) for the data. To allow for the possibly unpredictable nature of future observations it is natural to suppose each observation x_t is the observed value of a certain random variable X_t. The time series $\{x_t, t \in T_0\}$ is then a realization of the family of random variables $\{X_t, t \in T\}$. These considerations suggest modeling the data as the realization (or part of a realization) of a stochastic process $\{X_t, t \in T_0\}$ where $T_0 \subseteq T$. To clarify these ideas we need to define precisely what is meant by a stochastic process and its realizations.

Definition 2.1.1 (Stochastic Process) *A stochastic process is a family of random variables* $\{X_t, t \in T\}$ *defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$.

Remark 2.1.1 *In the following* T *will be the set* $\mathbb{Z}$ *where* $\mathbb{Z} = \{0, \pm 1, \pm 2, ...\}$.

Recalling the definition of a random variable we note that for each fixed $t \in T$, X_t is in fact a function $X_t(\cdot)$ on the set Ω. On the other hand, for each fixed $\omega \in \Omega$, $X_.(\omega)$ is a function on T.

Definition 2.1.2 (Realization of a Stochastic Process) *The functions* $\{X_.(\omega), \omega \in \Omega\}$ *on* T *are known as the realizations or sample paths of the process* $\{X_t, t \in T\}$.

In time series analysis we usually want to describe a stochastic process $\{X_t, t \in T_1\}$ with the knowledge of a realization $\{x_t, t \in T_0\}$ where $T_0 = [l, k]$ and $T_1 = (k, \infty)$ where $l < k < \infty$.

When dealing with a finite number of random variables, it is often useful to compute the covariance matrix in order to gain insight into the dependence between them. For a time series $\{X_t, t \in \mathbb{Z}\}$ we need to extend the concept of covariance matrix to deal with infinite collections of random variables. The autocovariance function provides us with the required extension.

Definition 2.1.3 (The Autocovariance Function) *If* $\{X_t, t \in \mathbb{Z}\}$ *is a process such that* $Var(X_t) < \infty$ *for each* $t \in \mathbb{Z}$, *then the autocovariance function* $\varphi_X(r, s)$ *of* $\{X_t, t \in \mathbb{Z}\}$ *is defined by*

$$\varphi_X(r, s) = Cov(X_r, X_s) = \mathbb{E}\left((X_r - \mathbb{E}X_r)(X_s - \mathbb{E}X_s)\right) \text{ with } r, s \in \mathbb{Z}.$$

One of the most important concepts in time series analysis is the concept of stationarity.

Definition 2.1.4 (Stationarity) *The time series* $\{X_t, t \in \mathbb{Z}\}$ *is said to be stationary if*

(i) $\mathbb{E}X_t^2 < \infty$ *for all* $t \in \mathbb{Z}$*,*

(ii) $\mathbb{E}X_t = \mu$ *for all* $t \in \mathbb{Z}$ *and where* μ *is a constant,*

(iii) $\varphi_X(r, s) = \varphi_X(r + t, s + t)$ *for all* $r, s, t \in \mathbb{Z}$*.*

This means that neither the mean μ nor the autocovariances $\varphi_X(r, s)$ of the process $\{X_t, t \in \mathbb{Z}\}$ depend on the date t.

Remark 2.1.2 *Stationary as just defined is frequently referred to in the literature as weak stationary, covariance stationary, stationary in the wide sense or second-order stationary.*

Definition 2.1.5 (Strict Stationarity) *The time series* $\{X_t, t \in \mathbb{Z}\}$ *is said to be strictly stationary if the joint distribution of the vector* $(X_{t_1}, ..., X_{t_k})$ *and* $(X_{t_{1+h}}, ..., X_{t_{k+h}})$ *are the same for all positive integers* k *and for all* $t_1, ..., t_k \in \mathbb{Z}$*.*

Definition 2.1.6 *If* $\{X_t, t \in \mathbb{Z}\}$ *is a stationary process, define*

$$\varphi_k = \varphi_X(r, s)$$

with $k = |r - s|$ *for all* $r, s \in \mathbb{Z}$*.*

Proposition 2.1.1 *If* φ *is the autocovariance function of a stationary process* $\{X_t, t \in \mathbb{Z}\}$*, then*

(i) $\varphi_0 \geq 0$

(ii) $|\varphi_k| \leq \varphi_0 \quad$ *for all* $k \in \{0, 1, 2, ...\}$*.*

Proof: The first property is a statement of the obvious fact that $Var(X_t) \geq 0$. The second is an immediate consequence of the Cauchy-Schwarz inequality.

$$|Cov(X_t, X_{t+h})| \leq (Var(X_t))^{\frac{1}{2}} \cdot (Var(X_{t+h}))^{\frac{1}{2}}$$

□

Definition 2.1.7 (Gaussian Time Series) *The process $\{X_t, t \in \mathbb{Z}\}$ is a Gaussian time series if and only if the distribution function of $(X_{t_1}, ..., X_{t_s})$ are all multivariate normal.*

2.2 Important Time Series

This section introduces some important classes of time series $\{X_t, t \in \mathbb{Z}\}$ defined in terms of linear difference equations. The so defined processes are used in the following sections.

The basic building block for all the processes considered in this work is the so called white noise process.

Definition 2.2.1 (White Noise Process) *A white noise process is a sequence $\{\epsilon_t, t \in \mathbb{Z}\}$ whose elements have zero mean and variance σ^2,*

$$\mathbb{E}\epsilon_t = 0$$

$$\mathbb{E}\epsilon_t^2 = \sigma^2$$

and for which the ϵ´s are uncorrelated

$$\mathbb{E}\epsilon_t\epsilon_s = 0 \quad for\ t \neq s\ .$$

If we replace the last condition with the slightly stronger condition that the ϵ´s are independent, the sequence is called independent white noise process.

Finally, an independent white noise where the ϵ´s are normally distributed

$$\epsilon_t \sim N(0, \sigma^2)$$

is called Gaussian white noise process.

A time series is a collection of observations indexed by the time of each observation. Usually we have collected data beginning with some particular time (say, $t = 0$) and ending at another time (say, $t = n$)

$$(X_0, X_1, ..., X_n).$$

We often imagine that we also have could earlier observations $(..., X_{-2}, X_{-1})$ or later observations $(X_{n+1}, X_{n+2}, ...)$, had the process been observed for more time. The observed sample $(X_0, X_1, ..., X_n)$ could be viewed as a finite segment of a doubly infinite sequence, denoted $\{X_t, t \in \mathbb{Z}\}$.

Typically, a time series $\{X_t, t \in \mathbb{Z}\}$ is identified by describing the *t-th* element.

Definition 2.2.2 (Moving Average Process) *A p-th order moving average process, denoted MA(p) is a stochastic process $\{X_t, t \in \mathbb{Z}\}$ characterized by*

$$X_t = \epsilon_t + a_1\epsilon_{t-1} + ... + a_p\epsilon_{t-p}$$

where $\{\epsilon_t, t \in \mathbb{Z}\}$ is an independent white noise process and $a_1, ..., a_p$ are any real numbers.

Definition 2.2.3 (Autoregressive Process) *A p-th order autoregressive process, denoted AR(p) is a stochastic process $\{X_t, t \in \mathbb{Z}\}$ characterized by*

$$X_t = a_1X_{t-1} + ... + a_pX_{t-p} + \epsilon_t$$

where $\{\epsilon_t, t \in \mathbb{Z}\}$ is an independent white noise process and $a_1, ..., a_p$ are any real numbers.

Remark 2.2.1 *In the further sections, necessary and sufficient conditions for the stationarity of an autoregressive process are given.*

An ARMA process includes both autoregressive and moving average terms.

Definition 2.2.4 (ARMA Process) *An ARMA(p,q) process is a stochastic process* $\{X_t, t \in \mathbb{Z}\}$ *characterized by*

$$X_t = a_1 X_{t-1} + ... + a_p X_{t-p} + \epsilon_t + b_1 \epsilon_{t-1} + ... + b_q \epsilon_{t-q}$$

where $\{\epsilon_t, t \in \mathbb{Z}\}$ *is an independent white noise process and* $a_1, ..., a_p, b_1, ..., b_q$ *could be any real numbers.*

Chapter 3

Properties of AR(2) Processes

The main part of this work deals with AR(2) processes, so it is necessary to discuss some properties of an AR(2) process. For example, conditions for the stationarity are given, useful representations are introduced and the least square estimators of the coefficients a_1 and a_2 are computed.

3.1 Representations and Stationarity

We consider a second order autoregressive process given by

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \epsilon_t \tag{3.1}$$

where $\{\epsilon_t, t \in \mathbb{Z}\}$ is an independent white noise process.

We will show that under certain conditions X_t can be written as the infinite sum

$$X_t = \sum_{i=0}^{\infty} \alpha_i \epsilon_{t-i} \tag{3.2}$$

with appropriate coefficients α_i for $i = 0, 1, \dots$.

That means that an AR(2) process can be written as a MA(∞)-process.

The question is how to find the coefficients α_i as a function of the parameters a_1 and a_2 and under what conditions this representation is valid. The first method to find the α_i´s is to use the following recursion

$$\begin{aligned}\alpha_0 &= 1\\ \alpha_1 &= a_1\\ \alpha_i &= a_1\alpha_{i-1} + a_2\alpha_{i-2} \quad \text{for } i \geq 2.\end{aligned}$$

With this recursion all α_i´s can be computed, if the representation exists. Nevertheless it is very useful to have an explicit representation of all α_i´s as a function of a_1 and a_2.

Theorem 3.1.1 *Let λ_1 and λ_2 be the roots of the so called characteristic equation $\lambda^2 - a_1\lambda - a_2 = 0$. If λ_1 and λ_2 lie inside the unit circle and λ_1 and λ_2 are distinct then X_t can be written as*

$$X_t = \sum_{i=0}^{\infty} \alpha_i \epsilon_{t-i}$$

with

$$\alpha_i = \frac{\lambda_1^{i+1} - \lambda_2^{i+1}}{\lambda_1 - \lambda_2} .$$

Proof: First we write (3.1) in a different way. Define

$$\begin{aligned}\psi_t &= \begin{pmatrix} X_t \\ X_{t-1} \end{pmatrix}\\ F &= \begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix}\\ v_t &= \begin{pmatrix} \epsilon_t \\ 0 \end{pmatrix}.\end{aligned}$$

Then (3.1) can be written as

$$\psi_t \;=\; F\psi_{t-1} + v_t \;.$$

Furthermore we get

$$\begin{aligned} \psi_t &= F\psi_{t-1} + v_t = F(F\psi_{t-2} + v_{t-1}) + v_t = ... \\ &= F^p\psi_{t-p} + F^{p-1}v_{t-p+1} + ... + Fv_{t-1} + v_t \;. \end{aligned}$$

If F^p converges to $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ for $p \to \infty$, then ψ_t is given by

$$\psi_t \;=\; \sum_{i=0}^{\infty} F^i v_{t-i} \;.$$

Let

$$F^i = \begin{pmatrix} f_{11}^{(i)} & f_{12}^{(i)} \\ f_{21}^{(i)} & f_{22}^{(i)} \end{pmatrix},$$

then

$$\begin{pmatrix} X_t \\ X_{t-1} \end{pmatrix} = \sum_{i=0}^{\infty} \begin{pmatrix} f_{11}^{(i)} & f_{12}^{(i)} \\ f_{21}^{(i)} & f_{22}^{(i)} \end{pmatrix} \cdot \begin{pmatrix} \epsilon_{t-i} \\ 0 \end{pmatrix}$$

and therefore

$$X_t \;=\; \sum_{i=0}^{\infty} f_{11}^{(i)} \epsilon_{t-i} \;.$$

Now we look at the eigenvalues of F which are the values of λ which satisfy

$$|F - \lambda I_2| = 0 \tag{3.3}$$

where I_2 the (2×2)-matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $|\cdot|$ denotes the determinant. Then (3.3) gives

$$\begin{vmatrix} a_1 - \lambda & a_2 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - a_1\lambda - a_2 = 0$$

hence

$$\lambda_1 = \frac{a_1 + \sqrt{a_1^2 + 4a_2}}{2}$$
$$\lambda_2 = \frac{a_1 - \sqrt{a_1^2 + 4a_2}}{2} .$$

Under the assumption that the eigenvalues of F are distinct we can assume, without loss of generality, that $\lambda_1 > \lambda_2$. By standard facts in linear algebra there exists a nonsingular matrix T which satisfies

$$F = T\Lambda T^{-1} \quad \text{where} \quad \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Observe that

$$F^2 = T\Lambda T^{-1} T\Lambda T^{-1} = T\Lambda^2 T^{-1},$$

and so F^i can be written as

$$F^i = T\Lambda^i T^{-1} \text{ where } \Lambda = \begin{pmatrix} \lambda_1^i & 0 \\ 0 & \lambda_2^i \end{pmatrix} . \tag{3.4}$$

Hence if λ_1 and λ_2 lie inside the unit circle, F^i converges as $i \to \infty$ to the (2×2) zero matrix.

Now we can use (3.4) to compute $f_{11}^{(i)}$ explicitly

$$F^i = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1^i & 0 \\ 0 & \lambda_2^i \end{pmatrix} \cdot \begin{pmatrix} t^{11} & t^{12} \\ t^{21} & t^{22} \end{pmatrix}$$

that gives us

$$f_{11}^{(i)} = (t_{11}t^{11})\lambda_1^i + (t_{12}t^{21})\lambda_2^i .$$

Define $c_1 = t_{11}t^{11}$ and $c_2 = t_{12}t^{21}$ and observe that $c_1 + c_2$ is nothing else than the (1,1)-element of the matrix $TT^{-1} = I_1$. It follows that $c_1 + c_2 = 1$

what gives with $t_{11}t^{11}\lambda_1 + t_{12}t^{21}\lambda_2 = a_1 = \lambda_1 + \lambda_2$ the unique solution for c_1 and c_2

$$c_1 = \frac{\lambda_1}{\lambda_1 - \lambda_2}$$
$$c_2 = \frac{-\lambda_2}{\lambda_1 - \lambda_2} .$$

This and

$$f_{11}^{(i)} = c_1\lambda_1^i + c_2\lambda_2^i$$

give the following representation of X_t

$$X_t = \sum_{i=0}^{\infty} \left(\frac{\lambda_1^{i+1} - \lambda_2^{i+1}}{\lambda_1 - \lambda_2} \cdot \epsilon_{t-i} \right) . \tag{3.5}$$

□

Remark 3.1.1 *In the following sections the coefficients a_1 and a_2 will be random variables. We will consider continuously distributed coefficients, hence the roots λ_1 and λ_2 of the characteristic equation are also continuously distributed so that the case $\lambda_1 = \lambda_2$ will occur with probability 0.*

Remark 3.1.2 *Note that the representation given in (3.5) is in general not unique, because there is also the possibility to write X_t in terms related to $\{\epsilon_t, \epsilon_{t+1}, \epsilon_{t+2}, ...\}$ i.e.*

$$X_t = \sum_{i=0}^{\infty} \beta_i \epsilon_{t+i} .$$

We are usually looking at past observations and so we are interested in the representation related to the past

$$X_t = \sum_{i=0}^{\infty} \alpha_i \epsilon_{t-i} .$$

If we are only looking at representations where $X_t \in \sigma(\epsilon_i, i \leq t)$ where $\sigma(\epsilon_i, i \leq t)$ is the σ-algebra generated by $\{\epsilon_i, i \leq t\}$, then the representation given in (3.5) is unique.

Proposition 3.1.1 *Let $\{\epsilon_t, t \in \mathbb{Z}\}$ be an independent white noise process. If the roots λ_1, λ_2 of the characteristic equation lie inside the unit circle then $\{X_t, t \in \mathbb{Z}\}$ is a stationary process.*

Proof:

$$\begin{aligned}
\mathbb{E}X_t &= \mathbb{E}\left(\sum_{i=0}^{\infty} \alpha_i \epsilon_{t-i}\right) = \sum_{i=0}^{\infty} \alpha_i \mathbb{E}\epsilon_{t-i} = 0 \\
\mathbb{E}X_t^2 &= \mathbb{E}\left(\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \alpha_i \alpha_j \epsilon_{t-i}\epsilon_{t-j}\right) \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \alpha_i \alpha_j \mathbb{E}\epsilon_{t-i}\epsilon_{t-j} = \sigma^2 \sum_{i=0}^{\infty} \alpha_i^2
\end{aligned}$$

since $\mathbb{E}\epsilon_{t-i} = 0$ and $\mathbb{E}\epsilon_{t-i}\epsilon_{t-j} = \sigma^2$ for $i = j$ and 0 otherwise. Further,

$$\begin{aligned}
\mathbb{E}X_t X_{t+h} &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \alpha_i \alpha_j \mathbb{E}\epsilon_{t-i}\epsilon_{t+h-j} \\
&= \sigma^2 \sum_{i=0}^{\infty} \alpha_i \alpha_{i+h}
\end{aligned}$$

since $\mathbb{E}\epsilon_{t-i}\epsilon_{t+h-j} = \sigma^2$ for $i = j - h$ and 0 otherwise. Note that we have already used that the sums $\sum_{i=0}^{\infty} \alpha_i$, $\sum_{i=0}^{\infty} \alpha_i^2$ and $\sum_{i=0}^{\infty} \alpha_i \alpha_{i+h}$ are finite, this follows from the fact that λ_1 and λ_2 lie inside the unit circle and α_i is given by $\frac{\lambda_1^{i+1} - \lambda_2^{i+1}}{\lambda_1 - \lambda_2}$ and therefore the sums converge.

□

In the sequel, it will be useful to have nice representations for $\sum_{i=0}^{\infty} \alpha_i$, $\sum_{i=0}^{\infty} \alpha_i^2$ and $\sum_{i=0}^{\infty} \alpha_i \alpha_{i+h}$. Observe that

$$\sum_{i=0}^{\infty}\alpha_i = \sum_{i=0}^{\infty}\frac{\lambda_1^{i+1}-\lambda_2^{i+1}}{\lambda_1-\lambda_2} = \frac{1}{\lambda_1-\lambda_2}\sum_{i=1}^{\infty}(\lambda_1^i-\lambda_2^i) \tag{3.6}$$

$$= \frac{1}{\lambda_1-\lambda_2}\sum_{i=0}^{\infty}(\lambda_1^i-\lambda_2^i) = \frac{1}{\lambda_1-\lambda_2}\left(\frac{1}{1-\lambda_1}-\frac{1}{1-\lambda_2}\right)$$

$$= \frac{1}{(1-\lambda_1)(1-\lambda_2)}$$

$$\sum_{i=0}^{\infty}\alpha_i^2 = \frac{1}{(\lambda_1-\lambda_2)^2}\sum_{i=0}^{\infty}(\lambda_1^{i+1}-\lambda_2^{i+1})^2 \tag{3.7}$$

$$= \frac{1}{(\lambda_1-\lambda_2)^2}\sum_{i=0}^{\infty}(\lambda_1^i-\lambda_2^i)^2$$

$$= \frac{1}{(\lambda_1-\lambda_2)^2}\left(\frac{1}{1-\lambda_1^2}+\frac{1}{1-\lambda_2^2}-\frac{2}{1-\lambda_1\lambda_2}\right)$$

$$= \frac{1+\lambda_1\lambda_2}{(1-\lambda_1^2)(1-\lambda_2^2)(1-\lambda_1\lambda_2)}$$

$$\sum_{i=0}^{\infty}\alpha_i\alpha_{i+h} = \frac{1}{(\lambda_1-\lambda_2)^2}\left(\frac{\lambda_1^h}{1-\lambda_1^2}+\frac{\lambda_2^h}{1-\lambda_2^2}-\frac{\lambda_1^h+\lambda_2^h}{1-\lambda_1\lambda_2}\right) \tag{3.8}$$

Note that λ_1 and λ_2 are the roots of the equation $\lambda^2 - a_1\lambda - a_2 = 0$ and therefore $a_1 = \lambda_1 + \lambda_2$ and $a_2 = -\lambda_1\lambda_2$ so that

$$\sum_{i=0}^{\infty}\alpha_i^2 = \frac{1-a_2}{(1+a_2)((1-a_2)^2-a_1^2)}\ . \tag{3.9}$$

3.2 Least Square Estimators

In time series theory we usually have observations of past realizations denoted by $(X_0, ..., X_n)$ of the process $\{X_t, t \in \mathbb{Z}\}$ and if we assume that the process $\{X_t, t \in \mathbb{Z}\}$ is an AR(2) process we are interested in estimators for the parameters a_1 and a_2. This can be done by least square estimation. It can be shown that for an AR(2) process the below presented least square estimators are consistent estimators for the parameters a_1 and a_2. *(see e.g. Brockwell and Davis (1991), p.257-258)*

To find the least square estimators we have to minimize

$$\sum_{t=2}^{n}(X_t - a_1X_{t-1} - a_2X_{t-2})^2 \ .$$

Partial differentiation gives

$$\begin{aligned}
\frac{\partial \sum_{t=2}^{n}(X_t - a_1X_{t-1} - a_2X_{t-2})^2}{\partial a_1} &= -2\sum_{t=2}^{n} X_{t-1}(X_t - a_1X_{t-1} - a_2X_{t-2}) \\
\frac{\partial \sum_{t=2}^{n}(X_t - a_1X_{t-1} - a_2X_{t-2})^2}{\partial a_2} &= -2\sum_{t=2}^{n} X_{t-2}(X_t - a_1X_{t-1} - a_2X_{t-2}) \ .
\end{aligned}$$

Setting the differentials equal to zero gives the estimators we are looking for

$$\begin{aligned}
\hat{a}_1 &= \frac{\sum_{t=1}^{n-1} X_tX_{t-1} \sum_{t=0}^{n-2} X_t^2 - \sum_{t=1}^{n-1} X_tX_{t-1} \sum_{t=2}^{n} X_tX_{t-2}}{\left(\sum_{t=0}^{n-2} X_t^2\right)^2 - \left(\sum_{t=1}^{n-1} X_tX_{t-1}\right)^2} \\
\hat{a}_2 &= \frac{\sum_{t=2}^{n} X_tX_{t-2} \sum_{t=0}^{n-2} X_t^2 - \left(\sum_{t=1}^{n-1} X_tX_{t-1}\right)^2}{\left(\sum_{t=0}^{n-2} X_t^2\right)^2 - \left(\sum_{t=1}^{n-1} X_tX_{t-1}\right)^2} \ .
\end{aligned}$$

Chapter 4

Aggregation of AR(2) Processes

4.1 Motivation for Aggregation

Aggregated observations are very common data sets in the studies of macroeconomic time series. For example there are government reports of sectors of the economy in which only the performance of a sector is reported. The data comes from an aggregation of a huge number of single processes (e.g. companies) so that the number of aggregated terms is very large. The number of observations is usually small (e.g. quarterly reports). Therefore we discuss especially the case when the number of aggregated terms is much larger than the number of observations.

This chapter discusses the aggregation of AR(2) processes, asymptotic results for the least square estimators are given, it is shown that the least square estimators, which are consistent estimators in the nonaggregated case are in general not consistent any more for the aggregated case. Moreover a central limit theorem when the number of aggregated terms is much larger than the number of observations is given.

4.2 Basic Definitions, Assumptions and Properties

We consider the aggregated process $Y_t^{(N)}$ defined by

$$Y_t^{(N)} = \frac{1}{N}\left(X_t^{(1)} + X_t^{(2)} + ... + X_t^{(N)}\right) \tag{4.1}$$

where

$$\left\{X_t^{(i)}, t \in \mathbb{Z}\right\}_{i=1,...,N} \tag{4.2}$$

are independent identically distributed random coefficient AR(2) processes. This means that for any $i = 1, ..., N$

$$X_t^{(i)} = a_1^{(i)} X_{t-1}^{(i)} + a_2^{(i)} X_{t-2}^{(i)} + \epsilon_t^{(i)} \quad , t \in \mathbb{Z} \tag{4.3}$$

where

$$\left\{\epsilon_t^{(i)}, t \in \mathbb{Z}\right\}_{i=1,...,N} \tag{4.4}$$

are independent identically distributed random variables with $\mathbb{E}\epsilon_t = 0$ and $\mathbb{E}\epsilon_t^2 = \sigma^2$ where $0 < \sigma^2 < \infty$,

$$\left\{a_1^{(i)}\right\}_{i=1,...,N} \quad \text{and} \quad \left\{a_2^{(i)}\right\}_{i=1,...,N} \tag{4.5}$$

are independent identically continuously distributed random variables and

$$\left\{\epsilon_t^{(i)}, t \in \mathbb{Z}\right\}_{i=1,...,N}, \quad \left\{a_1^{(i)}\right\}_{i=1,...,N} \quad \text{and} \quad \left\{a_2^{(i)}\right\}_{i=1,...,N} \tag{4.6}$$

are independent of each other. We will further assume

$$P(|a_1| < 2,\ |a_2| < 1,\ 1 - a_1 - a_2 > 0,\ 1 + a_1 - a_2 > 0) = 1 \tag{4.7}$$

where $P(A)$ denotes the probability of the event A. Furthermore

$$\mathbb{E}\frac{1 - a_2}{(1 + a_2)((1 - a_2)^2 - a_1^2)} < \infty \ . \tag{4.8}$$

Proposition 4.2.1 *If (4.1)–(4.8) are satisfied, all roots of the characteristic equation (see Chapter 3) lie inside the unit circle and are distinct with probability 1.*

Proof: The roots of the characteristic equation $\lambda^2 - a_1\lambda - a_2 = 0$ are given by

$$\lambda_1 = \frac{a1 + \sqrt{a_1^2 + 4a_2}}{2}$$

$$\lambda_2 = \frac{a1 - \sqrt{a_1^2 + 4a_2}}{2}$$

The roots are real when $a_1^2 + 4a_2 \geq 0$. In this case we have to show that

$$|a_1 + \sqrt{a_1^2 + 4a_2}| < 2 \quad \text{and} \tag{4.9}$$

$$|a_1 - \sqrt{a_1^2 + 4a_2}| < 2 \tag{4.10}$$

where $\sqrt{a_1^2 + 4a_2} \geq 0$. We know that $\lambda_1 + \lambda_2 = a_1$ and $-\lambda_1\lambda_2 = a_2$. Therefore the following must hold

$$|a_1| < 2 \; , \quad |a_2| < 1 \; .$$

Now to guarantee (4.9) and (4.10) the following must be satisfied

$$a_1 + \sqrt{a_1^2 + 4a_2} < 2 \quad \text{what is equivalent to} \quad 1 + a_1 - a_2 > 0$$

$$a_1 - \sqrt{a_1^2 + 4a_2} > -2 \quad \text{what is equivalent to} \quad 1 - a_1 - a_2 > 0 \; .$$

The roots are complex when $a_1^2 + 4a_2 \geq 0$. In this case we have to show that

$$|a_1 + i \cdot \sqrt{-a_1^2 - 4a_2}| < 2 \tag{4.11}$$

$$|a_1 - i \cdot \sqrt{-a_1^2 - 4a_2}| < 2 \; . \tag{4.12}$$

Now to guarantee (4.11) and (4.12) the following must be satisfied

$$(2 - a_1)^2 > \left(i \cdot \sqrt{-a_1^2 - 4a_2}\right)^2 \quad \text{what is equivalent to} \quad 1 - a_1 - a_2 > 0$$

$$(2 + a_1)^2 > \left(i \cdot \sqrt{-a_1^2 - 4a_2}\right)^2 \quad \text{what is equivalent to} \quad 1 + a_1 - a_2 > 0 \; .$$

□

Remark 4.2.1 *If (4.1)–(4.8) are satisfied, it follows immediately that the system is stationary with probability 1.*

4.3 Asymptotic Behavior and Properties of the Least Square Estimators

We assume that $Y_1^{(N)}, Y_2^{(N)}, ..., Y_n^{(N)}$ are observed from the aggregated process $\left\{Y_t^{(N)}, t \in \mathbb{Z}\right\}$ given in (4.1).

The "least square" estimators are (see Chapter 3)

$$T_{n,N}^{(1)} = \frac{\left(\sum_{t=2}^{n-1} Y_t^{(N)} Y_{t-1}^{(N)}\right)\left(\sum_{t=1}^{n-2}\left(Y_t^{(N)}\right)^2\right) - \left(\sum_{t=2}^{n-1} Y_t^{(N)} Y_{t-1}^{(N)}\right)\left(\sum_{t=3}^{n} Y_t^{(N)} Y_{t-2}^{(N)}\right)}{\left(\sum_{t=1}^{n-2}\left(Y_t^{(N)}\right)^2\right)^2 - \left(\sum_{t=2}^{n-1} Y_t^{(N)} Y_{t-1}^{(N)}\right)^2}$$

$$T_{n,N}^{(2)} = \frac{\left(\sum_{t=3}^{n} Y_t^{(N)} Y_{t-2}^{(N)}\right)\left(\sum_{t=1}^{n-2}\left(Y_t^{(N)}\right)^2\right) - \left(\sum_{t=2}^{n-1} Y_t^{(N)} Y_{t-1}^{(N)}\right)^2}{\left(\sum_{t=1}^{n-2}\left(Y_t^{(N)}\right)^2\right)^2 - \left(\sum_{t=2}^{n-1} Y_t^{(N)} Y_{t-1}^{(N)}\right)^2}.$$

Remark 4.3.1 *If (4.1)–(4.8) are satisfied, then by Proposition 3.1.1, λ_1 and λ_2 are different and lie inside the unit circle. Now from (3.7) and (3.9) we get that*

$$\mathbb{E}\frac{1-a_2}{(1+a_2)((1-a_2)^2-a_1^2)} = \mathbb{E}\frac{1+\lambda_1\lambda_2}{(1-\lambda_1^2)(1-\lambda_2^2)(1-\lambda_1\lambda_2)} > 0$$

By $\xrightarrow{p}$ we denote convergence in probability and by $\xrightarrow{d}$ we denote convergence in distribution.

The following theorems describe the asymptotic behavior of $T_{n,N}^{(1)}$ and $T_{n,N}^{(2)}$.

Theorem 4.3.1 *If (4.1)–(4.8) hold then for any fixed* $n \geq 3$ *we have*

$$T^{(1)}_{n,N} \xrightarrow{d} \tilde{T}^{(1)}_n \quad \textit{and}$$

$$T^{(2)}_{n,N} \xrightarrow{d} \tilde{T}^{(2)}_n \quad \textit{for} \quad N \to \infty$$

where

$$\tilde{T}^{(1)}_n = \frac{\sum_{t=2}^{n-1} \xi_t \xi_{t-1} \sum_{t=1}^{n-2} \xi_t^2 - \sum_{t=2}^{n-1} \xi_t \xi_{t-1} \sum_{t=3}^{n} \xi_t \xi_{t-2}}{\left(\sum_{t=1}^{n-2} \xi_t^2\right)^2 - \left(\sum_{t=2}^{n-1} \xi_t \xi_{t-1}\right)^2}$$

$$\tilde{T}^{(2)}_n = \frac{\sum_{t=3}^{n} \xi_t \xi_{t-2} \sum_{t=1}^{n-2} \xi_t^2 - \left(\sum_{t=2}^{n-1} \xi_t \xi_{t-1}\right)^2}{\left(\sum_{t=1}^{n-2} \xi_t^2\right)^2 - \left(\sum_{t=2}^{n-1} \xi_t \xi_{t-1}\right)^2}$$

and $(\xi_1, \xi_2, ..., \xi_n)$ *is a n-variate normal vector with mean* $\mathbb{E}\xi_t = 0$ *and co-variances*

$$\varphi_k = \mathbb{E}\xi_i \xi_{i+k} = \sigma^2 \mathbb{E}\left(\frac{1}{(\lambda_1 - \lambda_2)^2}\left[\frac{\lambda_1^k}{1-\lambda_1^2} + \frac{\lambda_2^k}{1-\lambda_2^2} - \frac{\lambda_1^k + \lambda_2^k}{1-\lambda_1\lambda_2}\right]\right) .$$

Proof: First we write

$$\sum_{t=1}^{n} \left(Y_t^{(N)}\right)^2 = \frac{1}{N^2} \sum_{t=1}^{n} \left(\sum_{i=1}^{N} X_t^{(i)}\right)\left(\sum_{j=1}^{N} X_t^{(j)}\right)$$

$$\sum_{t=2}^{n} Y_t^{(N)} Y_{t-1}^{(N)} = \frac{1}{N^2} \sum_{t=2}^{n} \left(\sum_{i=1}^{N} X_t^{(i)}\right)\left(\sum_{j=1}^{N} X_{t-1}^{(j)}\right)$$

$$\sum_{t=3}^{n} Y_t^{(N)} Y_{t-2}^{(N)} = \frac{1}{N^2} \sum_{t=3}^{n} \left(\sum_{i=1}^{N} X_t^{(i)}\right)\left(\sum_{j=1}^{N} X_{t-2}^{(j)}\right) .$$

Set

$$Y_N = \left(Y_1^{(N)}, Y_2^{(N)}, ..., Y_n^{(N)}\right) .$$

We want to use the multivariate central limit theorem. First we observe that $X_t^{(1)}, ..., X_t^{(N)}$ are independent identically distributed random variables for all $t \in \mathbb{Z}$. Let $\varphi_k = \mathbb{E}X_i X_{i+k}$. We know that

$$\varphi_0 \;=\; \sigma^2 \cdot \mathbb{E}\frac{1-a_2}{(1+a_2)((1-a_2)^2 - a_1^2)} \ .$$

(4.8) yields that $\varphi_0 < \infty$. Now we show that $|\varphi_k| \le \varphi_0$. The Cauchy-Schwarz inequality gives

$$|\varphi_k| = |Cov(X_t, X_{t+k})| \le (Var(X_t))^{\frac{1}{2}} (Var(X_{t+k}))^{\frac{1}{2}} = \varphi_0 \ ,$$

and therefore the covariance matrix is given by

$$\Sigma = (\sigma_{ij})_{i,j} = (\varphi_{|i-j|})_{i,j} \ .$$

Note that $\mathbb{E}X_t = 0$. Now we can use the multivariate central limit theorem which gives

$$\sqrt{N} \cdot Y_N \xrightarrow{d} (\xi_1, ..., \xi_n) \tag{4.13}$$

where $(\xi_1, ..., \xi_n)$ is a multivariate normal vector with zero mean and covariance matrix

$$\Sigma = (\sigma_{ij})_{i,j} = (\varphi_{|i-j|})_{i,j}$$

where $\varphi_{|i-j|}$ is given by

$$\varphi_k = \sigma^2 \cdot \mathbb{E}\left(\frac{1}{(\lambda_1 - \lambda_2)^2}\left[\frac{\lambda_1^k}{1-\lambda_1^2} + \frac{\lambda_2^k}{1-\lambda_2^2} - \frac{\lambda_1^k + \lambda_2^k}{1-\lambda_1\lambda_2}\right]\right) \ . \tag{4.14}$$

The continuous mapping theorem yields now the result.

□

Remark 4.3.2 *If*

$$\mathbb{E}\frac{1}{(1-\lambda)^3} < \infty \tag{4.15}$$

where $\lambda = max(|\lambda_1|, |\lambda_2|)$ and λ_1, λ_2 are the roots of the characteristic equation then (4.8) is satisfied.

Sometimes it is easier to check this condition instead of (4.8).

Proof:

$$\begin{aligned}
\left|\mathbb{E}\frac{1-a_2}{(1+a_2)((1-a_2)^2-a_1^2)}\right| &= \left|\mathbb{E}\frac{1+\lambda_1\lambda_2}{(1-\lambda_1^2)(1-\lambda_2^2)(1-\lambda_1\lambda_2)}\right| \\
&\le \mathbb{E}\left|\frac{1+\lambda_1\lambda_2}{(1-\lambda_1^2)(1-\lambda_2^2)(1-\lambda_1\lambda_2)}\right| \\
&\le C\cdot\mathbb{E}\frac{1}{(1-|\lambda_1|)(1-|\lambda_2|)(1-|\lambda_1\lambda_2|)} \\
&\le C\cdot\mathbb{E}\frac{1}{(1-\lambda)^3}
\end{aligned}$$

with some constant C.

□

Theorem 4.3.2 *If (4.1)–(4.8) are satisfied and*

$$\mathbb{E}\frac{1}{(1-\lambda)^5} < \infty \tag{4.16}$$

where $\lambda = max(|\lambda_1|, |\lambda_2|)$ *and* λ_1, λ_2 *are the roots of the characteristic equation then*

$$\begin{aligned}
\tilde{T}_n^{(1)} &\xrightarrow{p} \frac{\varphi_0\varphi_1-\varphi_1\varphi_2}{\varphi_0^2-\varphi_1^2}\ , \\
\tilde{T}_n^{(2)} &\xrightarrow{p} \frac{\varphi_0\varphi_2-\varphi_1^2}{\varphi_0^2-\varphi_1^2}
\end{aligned}$$

as $n\to\infty$.

Proof: By the Gaussianity and the assumptions of the theorem we have *(see. Shirayev (1996), p.293)*

$$|E\xi_t\xi_{t-1}\xi_s\xi_{s-1} - E\xi_t\xi_{t-1}E\xi_s\xi_{s-1}| = |E\xi_t\xi_s E\xi_{t-1}\xi_{s-1} + E\xi_t\xi_{s-1}E\xi_s\xi_{t-1}| \le$$

$$\leq \; C \cdot \left[E\left(\frac{1}{(\lambda_1-\lambda_2)^2} \left[\frac{\lambda_1^{|t-s|}}{1-\lambda_1^2} + \frac{\lambda_2^{|t-s|}}{1-\lambda_2^2} - \frac{\lambda_1^{|t-s|}+\lambda_2^{|t-s|}}{1-\lambda_1\lambda_2} \right] \right) \right]^2$$

$$\leq \; C \cdot \left| E\left(\frac{1}{(\lambda_1-\lambda_2)^2} \left[\frac{\lambda_1^{|t-s|}}{1-\lambda_1^2} + \frac{\lambda_2^{|t-s|}}{1-\lambda_2^2} - \frac{\lambda_1^{|t-s|}+\lambda_2^{|t-s|}}{1-\lambda_1\lambda_2} \right] \right) \right|$$

with some constant C. We will show that

$$\frac{1}{n}\sum_{t=1}^{n} \xi_t\xi_{t-1} \;\xrightarrow{p}\; \varphi_1$$

for $n \to \infty$. Now we use the Chebyshev inequality to get

$$P\left(\left| \frac{1}{n}\sum_{t=1}^{n} \xi_t\xi_{t-1} - E\xi_0\xi_1 \right| > \delta \right) \leq \frac{1}{n^2\delta^2} \cdot Var\left(\sum_{t=1}^{n} \xi_t\xi_{t-1} \right)$$

We have to show that

$$Var\left(\sum_{t=1}^{n} \xi_t\xi_{t-1} \right) \;\leq\; C \cdot n^{2-\alpha}$$

with some $\alpha > 0$ and some constant C. We have

$$Var\left(\sum_{t=1}^{n} \xi_t\xi_{t-1} \right) \;=\; \sum_{t=1}^{n}\sum_{s=1}^{n} \left(E\xi_t\xi_{t-1}\xi_s\xi_{s-1} - E\xi_t\xi_{t-1}E\xi_s\xi_{s-1} \right) \leq$$

$$\leq \; C \cdot \sum_{s,t=1}^{n} \left| E\left(\frac{1}{(\lambda_1-\lambda_2)^2} \left[\frac{\lambda_1^{|t-s|}}{1-\lambda_1^2} + \frac{\lambda_2^{|t-s|}}{1-\lambda_2^2} - \frac{\lambda_1^{|t-s|}+\lambda_2^{|t-s|}}{1-\lambda_1\lambda_2} \right] \right) \right|$$

$$\leq \; C \cdot n \cdot \sum_{t=1}^{\infty} \left| E\left(\frac{1}{(\lambda_1-\lambda_2)^2} \left[\frac{\lambda_1^{t}}{1-\lambda_1^2} + \frac{\lambda_2^{t}}{1-\lambda_2^2} - \frac{\lambda_1^{t}+\lambda_2^{t}}{1-\lambda_1\lambda_2} \right] \right) \right|$$

$$= \; C \cdot n \cdot \sum_{t=1}^{\infty} \left| E\left(\frac{1}{(\lambda_1-\lambda_2)^2} \sum_{i=1}^{\infty} \left(\lambda_1^{t+2i} + \lambda_2^{t+2i} - \lambda_1^t(\lambda_1\lambda_2)^i - \lambda_2^t(\lambda_1\lambda_2)^i \right) \right) \right|$$

$$= \; C \cdot n \cdot \sum_{t=1}^{\infty} \left| E\left(\frac{1}{(\lambda_1-\lambda_2)^2} \sum_{i=1}^{\infty} \left(\lambda_1^{t+i}[\lambda_1^i - \lambda_2^i] + \lambda_2^{t+i}[\lambda_2^i - \lambda_1^i] \right) \right) \right|$$

$$= \; C \cdot n \cdot \sum_{t=1}^{\infty} \left| E\left(\frac{1}{(\lambda_1-\lambda_2)^2} \sum_{i=1}^{\infty} \left([\lambda_1^{t+i} - \lambda_2^{t+i}][\lambda_1^i - \lambda_2^i] \right) \right) \right|$$

$$
\begin{aligned}
&\le\ C\cdot n\cdot\sum_{t=1}^{\infty}\left|E\left(\frac{1}{(\lambda_1-\lambda_2)^2}\cdot(\lambda_1-\lambda_2)^2\sum_{i=1}^{\infty}\left([t+i]\cdot|\lambda|^{t+i}\cdot i\cdot|\lambda|^{i}\right)\right)\right| \\
&\le\ C\cdot n\cdot E\left(\sum_{t=1}^{\infty}\sum_{i=1}^{\infty}\left([t+i]\cdot|\lambda|^{t+i}\cdot i\cdot|\lambda|^{i}\right)\right) \\
&\le\ C\cdot n\cdot E\left(\sum_{t=1}^{\infty}t\cdot|\lambda|^{t}\left(\sum_{i=1}^{\infty}i\cdot|\lambda|^{i}+\sum_{i=1}^{\infty}i^2\cdot|\lambda|^{i}\right)\right) \\
&\le\ C\cdot n\cdot E\left(\frac{1}{(1-|\lambda|)^2}\left(\frac{1}{(1-|\lambda|)^2}+\frac{1}{(1-|\lambda|)^3}\right)\right) \\
&\le\ C\cdot n\cdot E\frac{1}{(1-|\lambda|)^5} \\
&\le\ C\cdot n
\end{aligned}
$$

Analogously we get that

$$
\frac{1}{n}\sum_{t=0}^{n}\xi_t^2 \ \xrightarrow{p}\ \varphi_0
$$

$$
\frac{1}{n}\sum_{t=2}^{n}\xi_t\xi_{t-2} \ \xrightarrow{p}\ \varphi_2
$$

and therefore

$$
\frac{\sum_{t=2}^{n}\xi_t\xi_{t-1}\sum_{t=1}^{n}\xi_t^2-\sum_{t=2}^{n}\xi_t\xi_{t-1}\sum_{t=3}^{n}\xi_t\xi_{t-2}}{\left(\sum_{t=1}^{n}\xi_t^2\right)^2-\left(\sum_{t=2}^{n}\xi_t\xi_{t-1}\right)^2} \ \xrightarrow{p}\ \frac{\varphi_0\varphi_1-\varphi_1\varphi_2}{\varphi_0^2-\varphi_1^2}
$$

$$
\frac{\sum_{t=3}^{n}\xi_t\xi_{t-2}\sum_{t=1}^{n}\xi_t^2-\left(\sum_{t=2}^{n}\xi_t\xi_{t-1}\right)^2}{\left(\sum_{t=1}^{n}\xi_t^2\right)^2-\left(\sum_{t=2}^{n}\xi_t\xi_{t-1}\right)^2} \ \xrightarrow{p}\ \frac{\varphi_0\varphi_2-\varphi_1^2}{\varphi_0^2-\varphi_1^2}\ .
$$

□

From Theorem 4.3.1 and 4.3.2 it follows immediately that $T_{n,N}^{(1)}$ and $T_{n,N}^{(2)}$ cannot be used as estimators for $\mathbb{E}a_1$ and $\mathbb{E}a_2$ because usually the limits of $T_{n,N}^{(1)}$ and $T_{n,N}^{(2)}$, namely $\frac{\varphi_0\varphi_1-\varphi_1\varphi_2}{\varphi_0^2-\varphi_1^2}$ and $\frac{\varphi_0\varphi_2-\varphi_1^2}{\varphi_0^2-\varphi_1^2}$ are not equal to $\mathbb{E}a_1$ and $\mathbb{E}a_2$. This means that only in very special cases, the least square estimators will converge to the right limit, for example if the coefficients a_1 and a_2 are non-random constants.

4.4 Asymptotic Results for some Useful Statistics

In this section we introduce some statistics which are used in the next chapter to estimate distribution parameters. We give asymptotic results for these statistics and present a central limit theorem for the important case when the number of aggregated terms is much larger than the number of observations.

We consider the following statistics:

$$S_{n,N}^{(k)} = \frac{N}{n-k}\sum_{t=k}^{n} Y_t^{(N)} Y_{t-k}^{(N)} .$$

Remark 4.4.1 *From the proof of Theorem 4.3.1 it follows that $S_{n,N}^{(k)}$ converges in distribution to $\tilde{S}_n^{(k)}$ for $N \to \infty$, where*

$$\tilde{S}_n^{(k)} = \frac{1}{n-k}\sum_{t=k}^{n} \xi_t \xi_{t-k}$$

and the ξ's have the same properties as in Theorem 4.3.1, provided that all conditions of the theorem are satisfied.

Remark 4.4.2 *From the proof of Theorem 4.3.2 it follows that the quantity $\tilde{S}_n^{(k)} = \frac{1}{n-k}\sum_{t=k}^{n} \xi_t \xi_{t-k}$ converges in probability to $\mathbb{E}\xi_t\xi_{t-k} = \varphi_k$ for $n \to \infty$, provided that all conditions of the theorem are satisfied.*

We are interested in the case when the number of aggregated terms is much larger than the number of observations since this is the case that occurs in practice. The following theorem gives a central limit theorem for the given statistics which can be used for the estimation of distribution parameters and in addition for statistical tests.

Theorem 4.4.1 *If (4.1)–(4.8) and (4.16) are satisfied then*

$$\sqrt{n-k}\left(S_{n,N}^{(k)} - \varphi_k\right) \overset{d}{\to} \mathbb{N}(0, \sigma_k^2)$$

for some $\sigma_k^2 \geq 0$*, for all* $k \geq 0$, $k \in \mathbb{Z}$*, if* $n \to \infty$, $N \to \infty$, $N = f(n)$ *and the function* $f(x) \to \infty$ *sufficiently rapidly.*

To proof this theorem we need the following properties.

Lemma 4.4.1 *If (4.1)–(4.8) and (4.16) are satisfied then*

$$\sqrt{n-k}\left(\tilde{S}_{n}^{(k)} - \varphi_k\right) \overset{d}{\to} \mathbb{N}(0, \sigma_k^2)$$

for some $\sigma_k^2 \geq 0$*, for all* $k \geq 0$, $k \in \mathbb{Z}$*, if* $n \to \infty$.

Proof of Lemma 4.4.1:
This Lemma follows as an application of *(Breuer and Major(1983), Theorem 1)*. We only have to show that

$$\sum_{t=1}^{\infty} \left| E\left(\frac{1}{(\lambda_1 - \lambda_2)^2}\left[\frac{\lambda_1^t}{1-\lambda_1^2} + \frac{\lambda_2^t}{1-\lambda_2^2} - \frac{\lambda_1^t + \lambda_2^t}{1-\lambda_1\lambda_2}\right]\right)\right| < \infty$$

what we have already done in the proof of Theorem 4.3.2.

□

Lemma 4.4.2 *If (4.1)–(4.8) are satisfied then*

$$\sqrt{N}(Y_{t_1}^{(N)}, ..., Y_{t_k}^{(N)}) \overset{d}{\to} (\xi_{t_1}, ..., \xi_{t_k})$$

for any fixed k, $t_1 < t_2 < ... < t_k$*, as* $N \to \infty$.

Note that Lemma 4.8 is a result of (4.13).

Proof of Theorem 4.4.1:
By $\mathcal{L}(\cdot,\cdot)$ we denote the Lévy distance given by

$$\mathcal{L}(A,B) = inf\,[\varepsilon > 0 : F(x) \leq G(x+\varepsilon)+\varepsilon \text{ and } G(x) \leq F(x+\varepsilon)+\varepsilon,\ \forall\ x]$$

where F and G are the distribution functions of the random variables A and B.

We will use the following property of the Lévy distance

$$A_n \xrightarrow{d} A \ \Leftrightarrow\ \mathcal{L}(A_n, A) \to 0$$

where $\{A_n\}_{n\geq 0}$ is a sequence of random variables.

Fix $k \in \mathbb{N}$. From Lemma 4.4.1 we know that

$$\sqrt{n-k}\left(\tilde{S}_n^{(k)} - \varphi_k\right) \ \xrightarrow{d}\ \mathbb{N}(0,\sigma_k^2)$$

which shows that

$$\mathcal{L}\left(\sqrt{n-k}\left(\tilde{S}_n^{(k)} - \varphi_k\right), \mathbb{N}(0,\sigma_k^2)\right) \ \leq \epsilon_n \tag{4.17}$$

for some sequence $\epsilon_n \to 0$. Put

$$a_{n,N} \ = \ \mathcal{L}\left(\sqrt{n-k}\left(S_{n,N}^{(k)} - \varphi_k\right), \mathbb{N}(0,\sigma_k^2)\right) \ .$$

From Lemma 4.4.2 we know that for any fixed $n \in \mathbb{N}$

$$a_{n,N} \ \to \ \mathcal{L}\left(\sqrt{n-k}\left(\tilde{S}_n^{(k)} - \varphi_k\right), \mathbb{N}(0,\sigma_k^2)\right)$$

and thus by (4.17) there exists a number $N_0 = f(n)$ such that

$$a_{n,N} \leq 2\epsilon_n \ \text{ for } N \geq N_0$$

or equivalently

$$\mathcal{L}\left(\sqrt{n-k}\left(S_{n,N}^{(k)}-\varphi_k\right), \mathbb{N}(0,\sigma_k^2)\right) \;\; \leq 2\epsilon_n$$

for $N \geq f(n)$. This proves Theorem 4.4.1.

□

Theorem 4.4.1 gives a central limit theorem for the important case that the number of aggregated terms is much larger than the number of observations. Horváth and Leipus (2005) proposed a consistent estimator for $\mathbb{E}a$ in the AR(1) case. They gave a central limit theorem for the case that the number of observations is much larger than the number of aggregated terms. Usually the opposite case appears. Therefore, for practical use Theorem 4.4.1 is the important direction. We want to point out that as a consequence of the proof of Theorem 4.4.1 similar results will hold for the AR(1) case. Moreover the idea of Theorem 4.4.1 gives also a similar central limit theorem for the aggregation of AR(1) models when the number of aggregated terms is much larger than the number of observations and can be used for the estimator proposed by Horváth and Leipus (2005).

In the upcoming sections we will give an idea how we can use the statistics $S_{n,N}^{(k)}$ to estimate parameters.

4.5 Differences of Covariances

For constructing estimators, it will be useful to look at the differences of the covariances φ_i and ratios of them. In the following we will see that in many cases estimators of the coefficients or the parameters of the assumed distribution of the coefficients a_1 and a_2 can be found by using these differences. Furthermore we want to describe the structure of covariances we will use in the next chapter. Remember that the covariances can either be written in terms of the roots λ_1, λ_2 of the characteristic equation or in terms of the coefficients a_1 and a_2. We will use the expressions in terms of a_1 and a_2. If we have an observed data set, we can estimate the φ_i ´s from the data. The goal is to estimate distribution parameters. Now we investigate the differences of the covariances in terms of a_1 and a_2. We are interested in a representation of the differences $\varphi_i - \varphi_{i+2}$ in terms of the coefficients a_1 and a_2. Therefore we get

$$\varphi_0 - \varphi_2 = \sigma^2 \cdot \mathbb{E}\frac{1}{1+a_2} \tag{4.18}$$

$$\varphi_1 - \varphi_3 = \sigma^2 \cdot \mathbb{E}\frac{a_1}{1+a_2} \tag{4.19}$$

$$\varphi_2 - \varphi_4 = \sigma^2 \cdot \mathbb{E}\frac{a_1^2 + a_2}{1+a_2} \tag{4.20}$$

$$\varphi_3 - \varphi_5 = \sigma^2 \cdot \mathbb{E}\frac{a_1^3 + 2a_1a_2}{1+a_2} \tag{4.21}$$

$$\varphi_4 - \varphi_6 = \sigma^2 \cdot \mathbb{E}\frac{a_1^4 + 3a_1^2a_2 + a_2^2}{1+a_2} \tag{4.22}$$

$$\varphi_5 - \varphi_7 = \sigma^2 \cdot \mathbb{E}\frac{a_1^5 + 4a_1^3a_2 + 3a_1a_2^2}{1+a_2} \tag{4.23}$$

$$\varphi_6 - \varphi_8 = \sigma^2 \cdot \mathbb{E}\frac{a_1^6 + 5a_1^4a_2 + 6a_1^2a_2^2 + a_2^3}{1+a_2}\ . \tag{4.24}$$

Chapter 5

Estimators for Distribution Parameters

For the AR(1) case Horváth and Leipus (2005) proposed an unbiased estimator for the expected value of the random coefficient by using the idea of differences of covariances. In the AR(2) case the situation is more complex and the estimation of the random coefficients a_1 and a_2 by using these differences of covariances fails. But if we assume a parametric distribution for the random coefficients, we can use the results of the previous chapter to estimate these parameters. In this chapter we show how to estimate distribution parameters of the random coefficients a_1 and a_2.

5.1 Ideas to Find Estimators

As mentioned in the last section, we can estimate all covariances φ_i and all differences of them. If we assume a distribution for a_1 and a_2 with k unknown parameters, it is clear from the equations at the end of the last chapter that we can find k equations in terms of the parameters and terms we can estimate. There exist not too many distributions which satisfy all conditions for the asymptotic results. If we are only interested in numerical approximations, instead of explicitly given estimators, this can be done in more cases. Two important examples, namely the cases with independent Beta-distributed and Standard Two-Sided Power-distributed random coefficients are discussed at the end of this chapter.

5.2 Estimation Method

We look at continuously distributed parametric distributions on the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$ with at most two parameters which satisfy all conditions for the asymptotic result in the previous chapter. Then all moments are functions of these parameters. We propose a method to find equations in terms of the unknown parameters and in terms which can be estimated. Therefore we get an equation for each unknown parameter and we can compute estimators of the parameters. In practice this will be done numerically.

The goal is to estimate unknown parameters, denoted by $\alpha_1, \beta_1, \alpha_2, \beta_2$ and σ^2, where α_1 and β_1 are the parameters of the distribution of a_1 and α_2 and β_2 are the parameters of the distribution of a_2 and σ^2 is the variance of the white noise process introduced in the previous chapter.

We can estimate φ_k for all $k \in \mathbb{N}$. Remember that we assume that $\{a_1^{(i)}\}_{i=1,2,...}$ and $\{a_2^{(i)}\}_{i=1,2,...}$ are independent. The estimation will be done in the following steps

Step 1: Estimation of α_1, β_1:

By using the independence of a_1 and a_2 we get

$$\frac{\varphi_1 - \varphi_3}{\varphi_0 - \varphi_2} = \frac{\sigma^2 \cdot \mathbb{E}\frac{a_1}{1+a_2}}{\sigma^2 \cdot \mathbb{E}\frac{1}{1+a_2}} = \mathbb{E}a_1 \tag{5.1}$$

$$\varphi_3 - \varphi_5 = \sigma^2 \cdot \mathbb{E}\frac{a_1^3 + 2a_1 a_2}{1 + a_2}$$

$$(\varphi_3 - \varphi_5) + 2(\varphi_1 - \varphi_3) = \sigma^2 \cdot \left[\mathbb{E}\frac{1}{1+a_2} \cdot \mathbb{E}a_1^3 + 2\mathbb{E}a_1\right]$$

$$\frac{(\varphi_3 - \varphi_5) + 2(\varphi_1 - \varphi_3)}{\varphi_0 - \varphi_2} = \mathbb{E}a_1^3 + \frac{2\mathbb{E}a_1}{\mathbb{E}\frac{1}{1+a_2}} .$$

From (5.1) follows that

$$\frac{(\varphi_3 - \varphi_5) + 2(\varphi_1 - \varphi_3) - (\varphi_0 - \varphi_2) \cdot \mathbb{E}a_1^3}{2(\varphi_1 - \varphi_3)} = \frac{1}{\mathbb{E}\frac{1}{1+a_2}} \tag{5.2}$$

$$\frac{\varphi_0 - \varphi_4}{\varphi_0 - \varphi_2} = \mathbb{E}a_1^2 + \frac{1}{\mathbb{E}\frac{1}{1+a_2}} .$$

Combining the last relation with (5.2) we get

$$\mathbb{E}a_1^2 = \frac{\varphi_0 - \varphi_4}{\varphi_0 - \varphi_2} - \frac{(\varphi_3 - \varphi_5) + 2(\varphi_1 - \varphi_3) - (\varphi_0 - \varphi_2) \cdot \mathbb{E}a_1^3}{2(\varphi_1 - \varphi_3)} . \tag{5.3}$$

By using (5.1), (5.3) we can get estimators for α_1 and β_1.

Observe that these estimates give us estimators of all $\mathbb{E}a_1^k$ where $k \in \mathbb{N}$.

Step 2: Estimation of σ^2: By (4.18) we get

$$\sigma^2 = \frac{\varphi_0 - \varphi_2}{\mathbb{E}\frac{1}{1+a_2}} ,$$

and thus from (5.2) it follows that

$$\sigma^2 = \frac{\varphi_0 - \varphi_2}{2(\varphi_1 - \varphi_3)} \left[-\varphi_5 - \varphi_3 + 2\varphi_1 - (\varphi_0 - \varphi_2) \cdot \mathbb{E}a_1^3\right] , \tag{5.4}$$

what gives an estimator for σ^2.

Step 3: Estimation of α_2, β_2:

$$\begin{aligned}
\frac{\varphi_4-\varphi_6}{\sigma^2} &= \mathbb{E}\frac{a_1^4+3a_1^2a_2+a_2^2}{1+a_2} \\
&= \mathbb{E}a_1^4\cdot\mathbb{E}\frac{1}{1+a_2}+3\mathbb{E}a_1^2\cdot\mathbb{E}\frac{a_2}{1+a_2} \\
&\quad + \mathbb{E}\frac{a_2^2}{1+a_2}
\end{aligned}$$

$$\begin{aligned}
\frac{(\varphi_4-\varphi_6)+3(\varphi_2-\varphi_4)}{\sigma^2} &= \mathbb{E}a_1^4\cdot\mathbb{E}\frac{1}{1+a_2}+3\mathbb{E}a_1^2 \\
&\quad + 3\mathbb{E}\frac{a_2}{1+a_2}+\mathbb{E}\frac{a_2^2}{1+a_2}
\end{aligned}$$

$$\begin{aligned}
\frac{(\varphi_4-\varphi_6)+3(\varphi_2-\varphi_4)+2(\varphi_0-\varphi_2)}{\sigma^2} &= \mathbb{E}a_1^4\cdot\mathbb{E}\frac{1}{1+a_2}+3\mathbb{E}a_1^2 \\
&\quad + 2+\mathbb{E}a_2
\end{aligned}$$

Now we get

$$\begin{aligned}
\mathbb{E}a_2 &= \frac{-\varphi_6-2\varphi_4+\varphi_2+2\varphi_0-\mathbb{E}a_1^4\cdot(\varphi_0-\varphi_2)}{\sigma^2} \\
&\quad - 3\mathbb{E}a_1^2-2\ .
\end{aligned} \tag{5.5}$$

With the same method of adding and subtracting covariance differences and ratios of them we get the following more complicated equation for $\mathbb{E}a_2^2$

$$\begin{aligned}
\mathbb{E}a_2^2 &= \frac{-\varphi_8-5\varphi_6-6\varphi_4+5\varphi_2+7\varphi_0}{\sigma^2} \\
&\quad - \frac{\mathbb{E}a_1^4\cdot(\varphi_0-\varphi_2)-\mathbb{E}a_1^6\cdot(\varphi_0-\varphi_2)}{\sigma^2}-5\mathbb{E}a_1^4 \\
&\quad - 12\mathbb{E}a_1^2-5\mathbb{E}a_1-6\mathbb{E}a_1^2\cdot\mathbb{E}a_2\ .
\end{aligned} \tag{5.6}$$

From (5.5), (5.6) and the estimators of α_1, β_1 and σ^2 we can also get estimators for α_2 and β_2. Numerically this can be done easily.

5.3 Estimators for Beta-Distributed Coefficients

In this section estimators for Beta-distributed coefficients $\{a_1^{(i)}\}_{i=1,2,\ldots}$ and $\{a_2^{(i)}\}_{i=1,2,\ldots}$ are given. The main assumption is the following

$$\{a_1^{(i)}\}_{i=1,2,\ldots} \text{ and } \{a_2^{(i)}\}_{i=1,2,\ldots} \text{ are independent.} \tag{5.7}$$

We assume further that $\{a_1^{(i)}\}_{i=1,2,\ldots}$ and $\{a_2^{(i)}\}_{i=1,2,\ldots}$ are sets of independent identically distributed random variables with the following distributions:

$$a_1^{(i)} \sim Beta(\alpha_1, \beta_1) \quad \text{on} \quad \left[-\frac{1}{2}, \frac{1}{2}\right] ,$$

$$a_2^{(i)} \sim Beta(\alpha_2, \beta_2) \quad \text{on} \quad \left[-\frac{1}{2}, \frac{1}{2}\right] .$$

We want to use the results of Theorem 4.3.1, Theorem 4.3.2 and Theorem 4.4.1. Therefore we have to guarantee that all conditions of the theorems are satisfied. So we have to find conditions on the Beta-distribution so that the theorems can be used.

Remark 5.3.1 *If a_1 and a_2 are distributed as mentioned above we have to guarantee that*

$$\mathbb{E}\frac{1}{(1-\lambda)^5} < \infty \quad \textit{where} \quad \lambda = \max(|\lambda_1|, |\lambda_2|) . \tag{5.8}$$

Let f_x be the density of a $Beta(\alpha_1, \beta_1)$-distributed random variable X on $[0, 1]$ and f_y the density of a $Beta(\alpha_2, \beta_2)$-distributed random variable Y on $[0, 1]$ where f_x and f_y are given by

$$f_x(x) = \frac{1}{B(\alpha_1, \beta_1)} x^{\alpha_1 - 1}(1-x)^{\beta_1 - 1}$$

$$f_y(y) = \frac{1}{B(\alpha_2, \beta_2)} y^{\alpha_2 - 1}(1-y)^{\beta_2 - 1}$$

where $B(\cdot, \cdot)$ is the Beta function, $0 < x < 1$ and $0 < y < 1$.

To guarantee (5.8) we are interested in conditions on the distribution parameters which imply (5.8). The next proposition gives sufficient conditions for the existence of the expected value in (5.8) based on the distribution parameters of a_1, namely α_1 and β_1.

Proposition 5.3.1 *If a_1 and a_2 are distributed as mentioned above and $\alpha_1, \beta_1 > 5$, then (5.8) is satisfied.*

Proof: (5.8) is equivalent to

$$\mathbb{E}\frac{2}{(1-|\lambda_1|)^5} = \mathbb{E}\left(\frac{2}{2-\left|a_1+\sqrt{a_1^2+4a_2}\right|}\right)^5 < \infty$$

and

$$\mathbb{E}\frac{1}{(1-|\lambda_2|)^5} = \mathbb{E}\left(\frac{1}{2-\left|a_1-\sqrt{a_1^2+4a_2}\right|}\right)^5 < \infty \ .$$

Since $|a_1|, |a_2| \leq \frac{1}{2}$, $\left|a_1+\sqrt{a_1^2+4a_2}\right|$ is close to 2 if and only if a_1 is close to $\frac{1}{2}$ and a_2 is close to $\frac{1}{2}$. $\left|a_1-\sqrt{a_1^2+4a_2}\right|$ is close to 2 if and only if a_1 is close to $-\frac{1}{2}$ and a_2 is close to $\frac{1}{2}$. Furthermore

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{2-\left|a_1+\sqrt{a_1^2+4a_2}\right|}\right)^5 &\leq C\cdot\mathbb{E}\left(\frac{1}{2-\left|a_1+\sqrt{\frac{1}{4}+2}\right|}\right)^5 \\
&= C\cdot\mathbb{E}\left(\frac{1}{2-\left|a_1+\frac{3}{2}\right|}\right)^5 \\
&= C\cdot\mathbb{E}\left(\frac{1}{\frac{1}{2}-a_1}\right)^5
\end{aligned}$$

and

$$\mathbb{E}\left(\frac{1}{2-\left|a_1-\sqrt{a_1^2+4a_2}\right|}\right)^5 \leq C\cdot\mathbb{E}\left(\frac{1}{2-\left|a_1-\sqrt{\frac{1}{4}+2}\right|}\right)^5$$
$$= C\cdot\mathbb{E}\left(\frac{1}{2-|a_1-\frac{3}{2}|}\right)^5$$
$$= C\cdot\mathbb{E}\left(\frac{1}{\frac{1}{2}+a_1}\right)^5 .$$

Now we use the transformation $a_1 \leftrightarrow a-\frac{1}{2}$ where a is $Beta(\alpha_1,\beta_1)$-distributed on $[0,1]$ and we get that (5.8) holds if the expected values

$$\mathbb{E}\frac{1}{(1-a)^5} < \infty$$

and

$$\mathbb{E}\frac{1}{a^5} < \infty .$$

exist. The function f_x is integrable for any $\alpha_1 > 0, \beta_1 > 0$. Now the condition $\alpha_1, \beta_1 > 5$ gives the result.

□

If a random variable X is $Beta(\alpha,\beta)$-distributed on $[0,1]$, we can compute the expected value $\mathbb{E}X^k$ for all $k \in \mathbb{N}$ in terms of α and β *(see e.g. Gupta and Nadarajah (2004), p.35-36)* .

$$\mathbb{E}X^k = \frac{\alpha+k-1}{\alpha+\beta+k-1}\cdot\mathbb{E}X^{k-1} \text{ for k } \in \mathbb{N} . \tag{5.9}$$

We will need expressions for $\mathbb{E}a_1$, $\mathbb{E}a_1^2$, $\mathbb{E}a_1^3$, $\mathbb{E}a_1^4$, $\mathbb{E}a_1^6$, $\mathbb{E}a_2$, $\mathbb{E}a_2^2$ in terms of the distribution parameters.

Since $a_1 = a-\frac{1}{2}$, $a_2 = \hat{a}-\frac{1}{2}$ where a, $\hat{a}$ are Beta-distributed on $[0,1]$, we get

$$\begin{aligned}\mathbb{E}a_1^k &= \int_1^0 \left(x - \frac{1}{2}\right)^k f_x dx \\ \mathbb{E}a_2^l &= \int_1^0 \left(y - \frac{1}{2}\right)^l f_y dy \ .\end{aligned}$$

These integrals can be easily solved by expanding the polynomial $(x - \frac{1}{2})^k$ and using (5.9).

$$\mathbb{E}a_1 = \mathbb{E}X - \frac{1}{2} \quad (5.10)$$

$$\mathbb{E}a_1^2 = \mathbb{E}X^2 - \mathbb{E}X + \frac{1}{4} \quad (5.11)$$

$$\mathbb{E}a_1^3 = \mathbb{E}X^3 - \frac{3}{2}\mathbb{E}X^2 + \frac{3}{4}\mathbb{E}X - \frac{1}{8} \quad (5.12)$$

$$\mathbb{E}a_1^4 = \mathbb{E}X^4 - 2\mathbb{E}X^3 + \frac{3}{2}\mathbb{E}X^2 - \frac{1}{2}\mathbb{E}X + \frac{1}{16} \quad (5.13)$$

$$\begin{aligned}\mathbb{E}a_1^6 &= \mathbb{E}X^6 - 3\mathbb{E}X^5 + \frac{15}{4}\mathbb{E}X^4 - \frac{5}{2}\mathbb{E}X^3 \quad (5.14) \\ &\quad + \frac{15}{16}\mathbb{E}X^2 - \frac{3}{16}\mathbb{E}X + \frac{1}{64}\end{aligned}$$

$$\mathbb{E}a_2 = \mathbb{E}X - \frac{1}{2} \quad (5.15)$$

$$\mathbb{E}a_2^2 = \mathbb{E}X^2 - \mathbb{E}X + \frac{1}{4} \quad (5.16)$$

where $\mathbb{E}X^k$ is given by (5.9) and replacing α by α_1, β by β_1 in (5.10) - (5.14) and α by α_2, β by β_2 in (5.15) and (5.16).

Now we have all moments we need in terms of the distribution parameters. By using the method proposed in the previous section we get estimators for the distribution parameters.

5.4 Estimators for Standard Two-Sided Power-Distributed Coefficients

In this section we give an alternative example of a distribution of the coefficients $\{a_1^{(i)}\}_{i=1,2,...}$ and $\{a_2^{(i)}\}_{i=1,2,...}$. This shows that the method presented at the beginning of this chapter can be used also for other parametric distributions on the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$. We assume (5.7) and further that $\{a_1^{(i)}\}_{i=1,2,...}$ and $\{a_2^{(i)}\}_{i=1,2,...}$ are sets of independent identically distributed random variables with the following distributions:

$$a_1^{(i)} \sim STSP(\theta_1, n_1) \quad \text{on} \quad \left[-\frac{1}{2}, \frac{1}{2}\right] \ ,$$

$$a_2^{(i)} \sim STSP(\theta_2, n_2) \quad \text{on} \quad \left[-\frac{1}{2}, \frac{1}{2}\right] \ .$$

where $STSP(\theta, n)$ is the Standard Two-Sided Power Distribution with parameters θ and n.

Again we have to show that (5.8) is satisfied. Let f_x be the density of a $STSP(\theta_1, n_1)$-distributed random variable X on $[0, 1]$ and f_y the density of a $STSP(\theta_2, n_2)$-distributed random variable Y on $[0, 1]$ where f_x and f_y are given by

$$f_x(x|\ \theta_1, n_1) = \begin{cases} n_1 \left(\frac{x}{\theta_1}\right)^{n_1-1} &, \ \text{for } 0 \le x \le \theta_1 \\ \\ n_1 \left(\frac{1-x}{1-\theta_1}\right)^{n_1-1} &, \ \text{for } \theta_1 \le x \le 1 \end{cases}$$

$$f_y(y|\ \theta_2, n_2) = \begin{cases} n_2 \left(\frac{x}{\theta_2}\right)^{n_2-1} &, \ \text{for } 0 \le x \le \theta_2 \\ \\ n_2 \left(\frac{1-x}{1-\theta_2}\right)^{n_2-1} &, \ \text{for } \theta_2 \le x \le 1 \end{cases}$$

where $0 < x < 1$, $0 < y < 1$, $0 \leq \theta_1 \leq 1$, $0 \leq \theta_2 \leq 1$, $n_1 > 0$ and $n_2 > 0$.

Proposition 5.4.1 *Similarly to Proposition 5.3.1 we can show that if* a_1 *and* a_2 *are distributed as mentioned above and* $n_1 > 5$, *then (5.8) is satisfied.*

If a random variable X is $STSP(\theta, n)$-distributed on $[0, 1]$, we can compute the expected value $\mathbb{E}X^k$ for all $k \in \mathbb{N}$ in terms of θ and n *(see e.g. Kotz and van Dorp (2004), p.73)* .

$$\mathbb{E}X^k \quad = \quad \frac{n\theta^{k+1}}{n+k} - \sum_{i=0}^{k} \left(\begin{array}{c} k \\ k-i \end{array} \right) \frac{n(\theta-1)^{i+1}}{n+i} \quad \text{for k } \in \mathbb{N}\ . \qquad (5.17)$$

By using the same transformation as in the previous chapter and by using (5.17) and (5.10) - (5.16) we can apply the proposed method to estimate the parameters θ_1, n_1 and θ_2, n_2.

Bibliography

[1] Breuer P., Major P. (1983) Central Limit Theorems for Non-Linear Functionals of Gaussian Fields. *Journal of Multivariate Analysis,* **13**, 425-441.

[2] Brockwell P.J., Davis R.A. (1991) *Time Series: Theory and Methods, Second Edition.* Springer: New York.

[3] Gupta A.K., Nadarajah S. (2004) *Handbook of Beta Distribution and It Applications.* Marcel Dekker: New York.

[4] Granger C.W.J., Morris M.J. (1976) Time series modelling and interpretation. *Journal of the Royal Statistical Society, Ser A,* **139**, 246-257.

[5] Granger C.W.J. (1980) Long memory relationship and the aggregation of dynamic models. *Journal of Econometrics,* **14**, 227-238.

[6] Hamilton J.D. (1994) *Time Series Analysis.* Princeton: New Jersey.

[7] Horváth L., Leipus R. (2005) Effect of aggregation on estimators in AR(1) sequence. Preprint, University of Utah. 2005.

[8] Kotz S., van Dorp J.R. (2004) *Beyond Beta.* World Scientific: Singapore.

[9] Oppenheim G., Viano M.C. (2004) Aggregation of random parameters Ornstein-Uhlenbeck or AR processes: some convergence results. *Journal of Time Series Analysis,* **25**, 335-350.

[10] Robinson P.M. (1978) Statistical inference for a random coefficient autoregressive model. *Scandinavian Journal of Statistics,* **5**, 163-168.

[11] Shiryaev A.N. (1996) *Probability.* Springer: New York.

[12] Zaffaroni P. (2004) Contemporaneous aggregation of linear dynamic models in large economies. *Journal of Econometrics,* **120**, 75-102.