PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITE ABOU-BEKR BELKAID - TLEMCEN

FACULTY OF SCIENCES

DEPARTMENT OF MATHEMATICS

MASTER THESIS

SPECIALITY: PROBABILITY AND STATISTICS

PRESENTED BY:

OWUSU ERIC ADJEI

COMPARATIVE STUDY BETWEEN DETERMINISTIC AND RANDOM COEFFICIENT AUTOREGRESSIVE MODELS

September 2022

Mrs. Bensmain Nawel	University of Tlemcen	President
Mrs. Bensaber Fatna	University of Tlemcen	Supervisor
Mrs. Khetteb Zahira	University of Tlemcen	Examiner

Statistics and Random Modeling Laboratory (SRML)

BP 119, 13000 Tlemcen - Algeria

Contents

Acknowledgments **3**

Dedication **4**

Introduction **5**

1 Generalities on Real Stochastic Processes **8**

1.1 Definitions and existence 8

1.1.1 Process Law 10

1.1.2 Gaussian Process 10

1.2 Stationarity 10

1.3 Autocovariance Function and Autocorrelation Function 13

1.4 Linear Process 15

1.5 Moving Average Processes 17

2 Real Autoregressive Process **18**

2.1 Definitions 18

2.2 Infinite Moving Average expression 19

2.3 Stationarity 19

2.4 Yule-Walker Equations 24

2.5 Estimation of an AR(p) Parameters 25

2.5.1 Least Squares Method 25

2.5.2 Maximum Likelihood Method 27

2.6 Simulation and Application 27

2.6.1 A Strong White Noise Simulation . . . 27
2.6.2 An Order One Autoregressive Process Simulation . . . 28
2.6.3 An Order Two Autoregressive Process Simulation . . . 29
2.6.4 An Upper Order Autoregressive Process Simulation . . . 30
2.7 Prediction in an $AR(p)$ Model . . . 31

3 Random Coefficient Autoregressive Process 33
3.1 Preliminary Results . . . 33
3.2 Definition and Stationarity . . . 34
3.3 Markovian Expression of a $RCA(p)$ . . . 35
3.3.1 Conditions for Stationarity . . . 35
3.4 Parameters Estimation . . . 41
3.5 Simulation of Random Coefficient Autoregressive . . . 43
3.5.1 An Order One Autoregressive Process Simulation . . . 43
3.5.2 An Order Two Autoregressive Process Simulation . . . 45
3.6 Prediction in $RCA(p)$ Model . . . 47

4 Comparison and Application 48
4.0.1 Application to Simulated Data . . . 50
4.0.2 Application to Realistic Data . . . 53

Conclusion 56

A Appendix 60

Acknowledgments

This thesis is the result of a protracted process during which numerous people have assisted, encouraged, and supported me, especially Intra-Africa Mobility Scheme, I would want to thank them at this time. I want to start by honoring Professor NEGADI Latifa the Coordinator for the Intra-Africa Mobility Scheme. She owes a lot to this project. I would like to express my profound gratitude to Dr. ALLAM Abdelaziz, Head of the Mathematics Department and Professor at the University of Tlemcen, who despite his heavy academic and teaching workload and allowed me the great pleasure of overseeing the work on this thesis. I would like to express my sincere gratitude to Mad. BENSABER Fatna, Professor at the University of Tlemcen, for agreeing to participate in and instruct me throughout this research period. Also to Mrs. BENSMAIN Nawel, Professor at the University of Tlemcen, who kindly accepted to examine this thesis, I thank her very much for her time. Lastly, I would like to express to Mrs. KHETTAB Zahira, Professor at the University of Tlemcen, my respectful thanks for the interest she has shown in this thesis work by agreeing to participate in the jury. I want to thank all of my friends and colleagues who supported me in various ways, especially Miss Diana Nyamekye Nyarko for her advice and support throughout

Dedication

This thesis is dedicated to every member of my family.

Introduction

In statistics, a chronological series, or time series, designs the modeling of a sequence random and sequentially observed events, usually on a time scale. In other words, time series is a sequence of data points indexed in time order i.e. at different points in time. These data points typically consist of successive measurements made from the same source over a time interval and are used to track change over time. Therefore, a time series can be defined as a collection of observations x_t where the index t represents a unit of the time (a year, a day, an hour...). Time series can be found everywhere, since time is a constituent of everything that is observable. They numerous applications across various fields. One may put this in context through the example of electrical activity in the brain, rainfall measurements, stock prices, number of sunspots, annual retail sales, Heartbeats per minute monthly air passengers (Figure:1)[1] and so on.

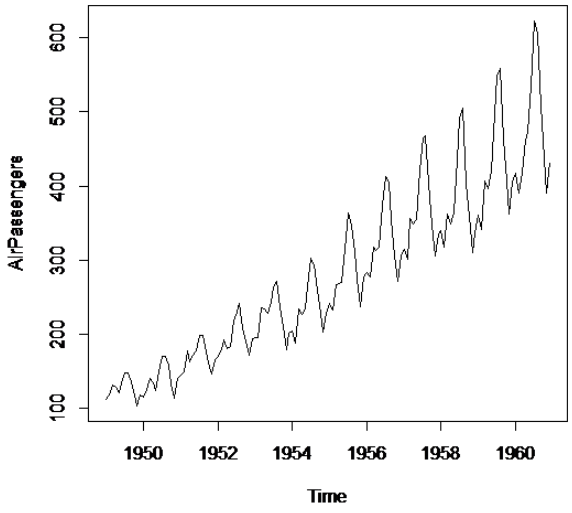


Figure 1: Monthly Airline Passenger Numbers 1949-1960

[1]Data are available in the software R and the plot can be obtained by the command ts.plot(AirPassengers)

The objective of studying time series is to forecast the series evolution. The concept behind the forecasts or prediction is to use previous data points to calculate the future points. For that a non-exhaustive list of mathematical models have been developed such as regression in the case of dependant observations. The mathematical model of a time series is called a stochastic process. It is a mathematical description of a distribution of time series and some time series are not but a realisation of stochastic processes (in term of simulation for instance). From a historical point of view, it is in astronomy that the first time series appear voluntarily for analytical purposes [6]. Systematic observation of the sky dates back to antiquity. For example, the Romans already knew that the year lasts approximately 365 and a quarter days(Julian calendar). According to Kendall (1973), the oldest known graph of a time series (at less in the Western world) is found in a manuscript of the tenth (or eleventh) century and illustrates a commentary on Cicero's Dream of Scipio (De Republica, 6, 14) taken from Saturnalia of Macrobius (395). The graph represents the inclination of the orbits of seven planets as a function of time

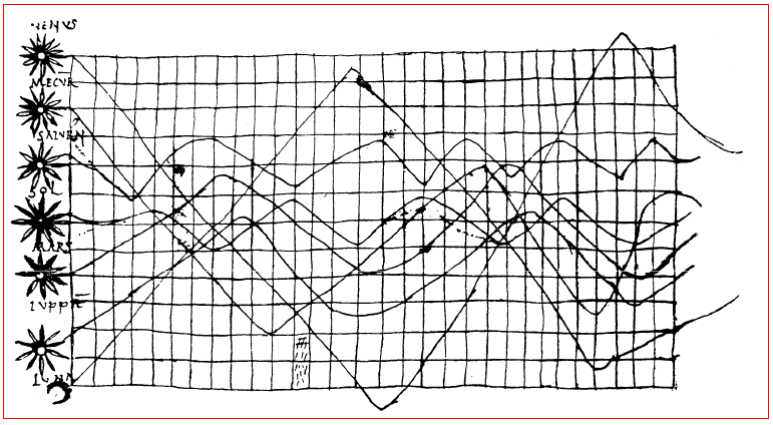


Figure 2: A tenth century chart Source: Funkhauser (1936) and Tufte (1983, p. 28)

However, this graph appears to be an isolated event. Time series charts have reappeared scientific written only during the 18th century (Lambert, Playfair)[6].

From a mathematical point of view, the theory of stochastic processes was settled around 1950. Since then, stochastic processes have become a common tool for mathematicians, physicists, engineers, and the field of application of this theory ranges from the modeling of stock pricing, to a rational option pricing theory, to differential geometry.

However the theoretical developments in time series analysis started early with stochastic processes. The first actual application of autoregressive models (to whom we are interested in

this thesis) to data can be brought back to the work of G. U Yule and J. Walker in the 1920s and 1930s. Autoregressive processes AR are a representative and classical model in time series analysis and it has broad applications in nature, science, and economics. AR can describe certain time-varying processes with a linear relationship in univariate time series.

In statistics, an autoregressive (AR) model is a statistical time series linear model which can be used to predict future values based on previous observations. AR models use regression techniques and rely on autocorrelation in order to make accurate predictions. The autoregressive model $AR(p)$ said of order p determines the value of a process at an arbitrary time step t using a linear combination of the p-last values and superimposed by a so called white noise and assumed to be a sequence of uncorrelated with each other in time and identically distributed, with an expected value of zero and finite variance. The weights of the linear combination are the model parameters. They are considered to be constant. However, when dealing with economic data modelling, Kendall (1953) [10] has recognised that it is more reasonable to generalize the constant coefficient in models to ones changing through time as the economy changes. Attention has been restricted to non-linear autoregressive models with random coefficients. For instance, Garbade (1977) [8] has dealt with regression models estimation where the coefficients are assumed to follow a simple random walk.

So an obvious variation of autoregressive processes are the random coefficient autoregressive RCA models with whom we are in fact concerned in this work. This late is organised as follows: In chapter 1, we give preliminary results on general stochastic process. Then, we introduce real autoregressive processes with deterministic coefficients in chapter 2. Followed by Chapter 3 which is devoted to the random coefficients autoregressive processes. In both chapter 2 and 3, we give conditions of stationarity of the process and we deal with its parameters estimation parameters estimation as well as future predictions. Results are illustrated by simulations. In chapter 4, we compare the two processes by applied the two models to simulated an realistic data.

Chapter 1

Generalities on Real Stochastic Processes

1.1 Definitions and existence

Stochastic processes are used to describe random phenomena which depend on time. Before introducing them let enunciate (without proof) the following result:

Theorem 1.1 (Kolmogorov Theorem)**.** *[11]*
Let T a set of parameters. We denote by $\mathbb{R}^T$ the set of applications defined on T with values in $\mathbb{R}$ and let Σ be the set of finite parts of T. Let $\sigma \in \Sigma$, we denote by π_σ the projection on $\prod_{t\in\sigma}\mathbb{R}$ of $\mathbb{R}^T$, we have

$$\pi_\sigma : \mathbb{R}^T \to \mathbb{R}^\sigma \quad \text{and} \quad \pi_\sigma(\mathbb{R}^T) = \mathbb{R}^\sigma$$

We equip $\mathbb{R}^T$ respectively $\mathbb{R}^\sigma$ by the σ-algebra S^T generated by $(\pi_\sigma, \sigma \in \Sigma)$ and the Borel σ-algebra B_σ respectively. Let $(\mu_\sigma)_{\sigma\in\Sigma}$ a family of probabilities such that:

$$\diamond \quad \mu_\sigma \quad \text{is a probability on} \quad (\mathbb{R}^\sigma, \mathrm{B}_\sigma) \qquad \forall\sigma \in \Sigma$$

$$\diamond\diamond \text{ If } \sigma < \tau, (\sigma,\tau) \in \Sigma^2 \text{ and } \psi_{\tau\sigma} \text{ denote the projection of } \mathbb{R}^\tau \text{ on } \mathbb{R}^\sigma \text{ we have } \mu_\sigma = \mu_\tau \circ \psi_{\tau\sigma}^{-1}$$

Then, It exists a unique probability P on $(\mathbb{R}^T, \mathrm{S}^T)$ such that

$$\mu_\sigma = \mu_T \circ \psi_{T\sigma}^{-1}\mu_\sigma = \mu_T \circ \pi_\sigma^{-1} \qquad \forall\sigma \in \Sigma$$

Definition 1.1.

We define a stochastic process by a triplet $((\Omega, \mathrm{A}, P), (X_t, t \in T), (E, \xi))$ *such that*

a) (Ω, A, P) *is a probability space called the basic space,*

b) (E, ξ) *is a measurable space called the space of states,*

c) $(X_t, t \in T)$ *is a family of random variables defined on* (Ω, A, P) *with values in* (E, ξ)

d) and T *is an arbitrary set of parameters (usually chosen to be set of times).*

♠ The stochastic process will be denoted simply by $(X_t, t \in T)$ where $(X_t)_{t \in T}$ are such that:

$$\begin{array}{rcccl} X_t & : & (\Omega, A, P) & \to & (E, \xi) \\ & & \omega & \mapsto & X_t(\omega) = X(t, \omega) \end{array}$$

For a fixed ω, the application $t \to X_t(\omega) = X(t, \omega)$ is said to be the path.

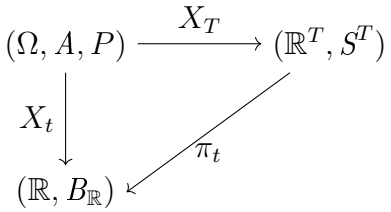
♠ If $T = \mathbb{N}$ or $\mathbb{Z}$, the process X_T is said to be a **discrete-time process** and if T is an interval in $\mathbb{R}$, it is said to be a **continuous-time process**

♠ If $(E, \xi) = (\mathbb{R}, B_{\mathbb{R}})$ where $B_{\mathbb{R}}$ is the Borel σ-algebra of $\mathbb{R}$, X_T is said to be a real process.

♠ Using precedent notations with those of Kolmogorov theorem, $(X_t, t \in T)$ can be considered as a random variable X_T with values in $(\mathbb{R}^T, S^T)$:

$$\begin{array}{rcccl} X_T & : & (\Omega, A, P) & \to & (\mathbb{R}^T, S^T) \\ & & \omega & \mapsto & X_T(\omega) = (X(t, \omega), t \in T) \end{array}$$

This can be schematised by



1.1.1 Process Law

The distribution law $L(X_T)$ of the process is a probability P_{X_T} on $(\mathbb{R}^T, S^T)$ defined by

$$\forall \mathsf{s} \in S^T \qquad P_{X_T}(\mathsf{s}) = P(X_T^{-1}(\mathsf{s}))$$

this law is completely defined by the finite dimensional laws (those of the vectors $(X_{t_1}, \cdots, X_{t_n})$ where $t_1, \cdots, t_n$ are separate elements of T).

1.1.2 Gaussian Process

A real process X_T is said to be gaussian if all its finite dimensional laws are gaussian i.e. if each finite linear combination of $X_t, t \in T$ is a real gaussian random variable.
The law of such a process is completely determined by the given of the mean function and the covariance function:

$$m(t) = E(X_t), \qquad t \in T$$
$$c(s,t) = Cov(X_s, X_t), \qquad s,t \in T$$

Example 1.1.

• *White Noise: a simple stochastic process which is a collection of uncorrelated real random variables* $(\varepsilon_t, t \in T)$ *with zero mean* $E(\varepsilon_t) = 0$ *and finite variance* $Var(\varepsilon_t) = \sigma_\varepsilon^2$

• *Strong White Noise: a particular useful white noise is strong if the* ε_t *are independent and identically distributed normal random variables. It is a gaussian process*

• *White noise is of great interest because the stochastic behavior of almost time series could be explained in terms of the white noise especially those modeled by an autoregressive process (see chapter 2).*

1.2 Stationarity

Stationary processes are those for whom the finite dimensional laws do not change with time. More precisely:

Definition 1.2 (**strong stationarity**).

A process X_T is said to be strictly stationary if for each finite part $(t_1, \cdots, t_n)$ of T and for all $h > 0$ we have:

$$L(X_{t_1+h}, \cdots, X_{t_n+h}) = L(X_{t_1}, \cdots, X_{t_n})$$

where $L(Y)$ is the law of the variable Y.

Example 1.2. *The white noise is the simplest example of stationary processes*

Before given a less rigid definition for stationarity let define second order process.

Definition 1.3.

A real process X_T is said to be a second ordered process if

$$\sup_{t \in T} E(X_t^2) < +\infty$$

Definition 1.4 (**weak stationarity**).

Let X_T to be a second ordered process. It is said to be weakly stationary if

$$E(X_t) = m \qquad \forall t \in T$$

$$Cov(X_s, X_t) = Cov(X_{s+h}, X_{t+h}) \qquad \forall s, t, h$$

Remark 1.1. *It is obvious that for a real gaussian process the strictly stationarity coincide with the weak stationarity.*

Example 1.3.

The process of second order $(X_t, t \in \mathbb{Z})$ defined by

$$X_t = A\cos\theta t + B\sin\theta t, \qquad \theta \in [-\pi, +\pi]$$

where X_t are centered reduced independent random variables. We have

$$E(X_t) = 0 \qquad \forall t \in T$$

$$Cov(X_s, X_t) = \cos\theta h$$

So the process is weakly stationary

Remark 1.2. *It is obvious that a strictly second order process is weakly stationary. However, the reciprocal is not true. Indeed, let the discrete process* $(X_t, t \in \mathbb{N})$ *of independent variables that obey to the normal law* $N(1,1)$ *if* t *is even and to the exponential law* $\xi(1)$ *if* t *is odd. It is clear that* $L(X_{2t}) \neq L(X_{2t+1})$*, so the process is not strictly stationary. But it is easy to see that* $Cov(X_t, X_t) = 1$ *and that* $Cov(X_{t+h}, X_t) = 0, \quad h \neq 0$*, so the covariance does not depend on* h*. That is to say that the process is weakly stationary.*

Remark 1.3.

- ♠ *For a weakly stationary process* $Cov(X_h, X_0) = Cov(X_{t+h}, X_t) =: \gamma_X(h)$
- ♠ $Cov(X_s, X_t) = f(s-t)$

Remark 1.4.

It is possible to construct a weakly stationary process from an arbitrary one. Indeed, let $(Z_t)_{t\in T}$ *an arbitrary process of independent centred random variables with variance* $Var(Z_t) = \sigma^2_{Z_t}$*. We put*

$$X_t = Z_t + \theta Z_{t-1},$$

hence

$$E(X_t) = 0 \text{ and } E|X_t^2| < +\infty$$

$$Cov(X_{t+h}, X_t) = \begin{cases} (1+\theta^2)\sigma^2_{Z_t} & \text{if} \quad h = 0 \\ \theta^2\sigma^2_{Z_t} & \text{if} \quad |h| = 1 \\ 0 & \text{if} \quad |h| \neq 1 \end{cases}$$

Proposition 1.1. *[11] If a process* $(X_t, t \in T)$ *is weakly stationary and gaussian, then it is strictly stationary.*

1.3 Autocovariance Function and Autocorrelation Function

Let $(X_t, t \in (Z))$ a weakly stationary process

Definition 1.5.

The autocovariance function of the process $(X_t, t \in (Z))$ is defined by

$$R(h) = Cov(X_{t+h}, X_t)$$

Definition 1.6.

The autocorrelation function of the process $(X_t, t \in \mathbb{Z})$ is defined by

$$\rho(h) = \frac{R(h)}{R(0)}$$

Proposition 1.2.

1°/ $R(0) = \sigma_{X_0}$.

2°/ $|R(h)| \leq R(0)$ and $R(h) = R(-h)$.

3°/ R is of positive type; that is to say $\forall a_1, \cdots, a_m \in \mathbb{R}$ and $\forall t_1, \cdots, t_m \in \mathbb{Z}$;

$$\sum_{s=1}^{m} \sum_{t=1}^{m} a_s a_t R(s-t) \geq 0$$

. 4°/ $\rho(0) = 1$; $\rho(h) \leq 1$; ρ is of positive type

Proof.

let consider the process $(X_t, t \in (Z))$ to be centred; $E(X_t) = 0$

1°/ $R(0) = Cov(X_t, X_t) = \sigma_{X_0}$.

2°/Using the definition of the covariance and the Holder inequality, we get

$$\begin{aligned} R(h) &= Cov(X_{t+h}, X_t) \\ &= E(X_{t+h}X_t) - E(X_{t+h})E(X_t) \\ &= E(X_{t+h}X_t) \\ &\leq E(X_{t+h}^2)^{\frac{1}{2}} E(X_t^2)^{\frac{1}{2}} \\ &= Cov(X_{t+h}, X_{t+h})^{\frac{1}{2}} Cov(X_t, X_t)^{\frac{1}{2}} \\ &= R(0)^{\frac{1}{2}} R(0)^{\frac{1}{2}} = R(0) \end{aligned}$$

Which proof the first part of 2°. Now to prove the second one we need to introduce the variable changing $s = t + h$ in the definition of $R(h)$:

$$R(h) = E(X_{t+h}X_t) = E(X_s X_{s_h}) = E(X_s X_{s+(_h)}) = R(-h)$$

3°/ Let $a_1, \cdots, a_m \in \mathbb{R}$ and $t_1, \cdots, t_m \in \mathbb{Z}$, we have

$$\begin{aligned} \sum_{i=1}^{m}\sum_{j=1}^{m} a_{t_i}a_{t_j}R(t_i - t_j) &= \sum_{i=1}^{m}\sum_{j=1}^{m} a_{t_i}a_{t_j}Cov(X_{t_i}, X_{t_j}) \\ &= \sum_{i=1}^{m} a_{t_i}^2 Cov(X_{t_i}, X_{t_i}) + \sum_{i=1}^{m}\sum_{\substack{j=1 \\ j\neq i}}^{m} a_{t_i}a_{t_j}Cov(X_{t_i}, X_{t_j}) \\ &\geq \sum_{i=1}^{m} a_{t_i}^2 Cov(X_{t_i}, X_{t_i}) \\ &= \sum_{i=1}^{m} a_{t_i}^2 Var(X_{t_i}) \\ &= \sum_{i=1}^{m} Var(a_{t_i}X_{t_i}) \\ &= Var(\sum_{i=1}^{m} a_{t_i}X_i) \\ &\geq 0 \end{aligned}$$

Now, for $m = 2$ we have:

$$\begin{aligned} Var(a_1X_{t_1} + a_2X_{t_2}) &\Rightarrow E(a_1X_{t_1}^2 + a_2X_{t_2}^2) \\ &= \sum_{s=1}^{n}\sum_{k=t}^{m} E(X_sX_t) \\ &= E(\sum_{s=1}^{n}\sum_{k=t}^{m} X_sX_t) \\ &= Var(\sum_{i=1}^{m} X_{t_i}) \\ &\geq 0 \end{aligned}$$

$4^\circ/$ $\rho(0) = \dfrac{R(0)}{R(0)} = 1; \quad \rho(h) = \dfrac{R(h)}{R(0)} \leq \dfrac{R(0)}{R(0)} = 1;$ ρ is of positive type since $R(h)$ is

□

1.4 Linear Process

Definition 1.7.

A Process $X_{\mathbb{Z}}$ is said to be linear if it can be written in the form

$$X_t = \mu + \sum_{j=-\infty}^{+\infty} a_j\epsilon_{t-j} \quad \forall t \in T$$

where μ, a_j are real parameters such that $\sum_{j=-\infty}^{+\infty} |a_j| < \infty$ and ε_T is a weak white noise,

Proposition 1.3.

A linear process is strictly stationary

Proof. Let $(X_{\mathbb{Z}})$ be a linear process. Notice that it is a zero mean process

$$E(X_t) = \mu + \sum_{j=-\infty}^{+\infty} a_jE(\varepsilon_{t-j}) = \mu$$

and that

$$Var(X_t) = \sum_{j=-\infty}^{+\infty} a_j^2Var(\varepsilon_{t-j}) = \sigma_\varepsilon^2 \sum_{j=-\infty}^{+\infty} a_j^2 < \infty$$

Let now calculate the autocovariance function

$$\begin{aligned} Cov(X_{t+h}, X_t) &= Cov(\mu + \sum_{j=-\infty}^{+\infty} a_j \epsilon_{t+h-j}, \mu + \sum_{j=-\infty}^{+\infty} a_j \varepsilon_{t-j}) \\ &= \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} a_j a_k E(\varepsilon_{t+h-j} \varepsilon_{t-k}) \end{aligned}$$

however,

$$E(X_{t+h-j}, X_{t-k}) = \begin{cases} \sigma_\varepsilon^2 & \text{if } t+h-j = t-l \\ 0 & \text{if not} \end{cases}$$

Hence we find that

$$\begin{aligned} Cov(X_{t+h}, X_t) &= Cov(\mu + \sum_{j=-\infty}^{+\infty} a_j \varepsilon_{t+h-j}, \mu + \sum_{j=-\infty}^{+\infty} a_j \varepsilon_{t-j}) \\ &= \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} a_j a_k E(\varepsilon_{t+h-j} \varepsilon_{t-k}) \\ &= \sigma_\varepsilon^2 \sum_{j=-\infty}^{+\infty} a_{k+h} a_k \\ &= R(h) \end{aligned}$$

□

Remark 1.5.

$$R(0) = \sigma_X^2 = \sigma_\varepsilon^2 \sum_{j=-\infty}^{+\infty} a_j^2 \text{ and } \rho(h) = \frac{R(h)}{R(0)} = \frac{\sum_{j=-\infty}^{+\infty} a_{j+h} a_j}{\sum_{j=-\infty}^{+\infty} a_j^2}$$

Example 1.4.

Let the process $X_t = \varepsilon_t - \theta \varepsilon_{t-1}$ *with* $(\varepsilon_t)_{t \in (Z)}$ *is a white noise. It is a linear one. It is easy to see that*

$$a_j = \begin{cases} 0 & \text{if } j < 0 \text{ and } j \leq 2 \\ 1 & \text{if } j = 0 \\ -\theta & \text{if } j = 1 \end{cases},$$

$R(0) = \sigma_\varepsilon^2(1+\theta^2)$, $\quad R(1) = \sigma_\varepsilon^2 \theta$, $R(h) = 0$ *for* $h > 1$. *and the autocorrelation function is given*

by

$$\rho(h) = \begin{cases} 1 & \text{if } h = 0 \\ \frac{\theta}{1+\theta^2} & \text{if } |h| = 1 \\ 0 & \text{if } |h| > 1 \end{cases}.$$

1.5 Moving Average Processes

A pth-order moving average process, or $MA(p)$ process, is defined by the equation

$$X_t = a_1\varepsilon_t + a_2\varepsilon_{t-1} + \cdots + a_{t-q}\varepsilon_{t-p},$$

where $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a white-noise process. A moving-average process is clearly stationary since any two elements X_t and X_s represent the same function of identically distributed vectors $(\varepsilon_t, \varepsilon_{t-1}, \cdots, \varepsilon_{t-p})$ and $(\varepsilon_s, \varepsilon_{s-1}, \cdots, \varepsilon_{s-p})$. The autocovariance function of a $MA(p)$ process such that

$$X_t = \varepsilon_t - a_2\varepsilon_{t-1} - \cdots - a_{t-q}\varepsilon_{t-p},$$

is given by

$$R(h) = \sigma_\varepsilon^2(1 + a_1^2 + a_2^2 + \cdots + a_p^2), |h| < p$$

Example 1.5. *The first-order moving-average process $X_t = \varepsilon_t + \theta\varepsilon_{t-1}$ has the following autocovariance function:*

$$R(h) = \begin{cases} \sigma_\varepsilon^2(1+\theta^2) & \text{if } h = 0, \\ -\sigma_\varepsilon^2\theta & \text{if } h = 1, \\ 0 & \text{if } h > 1. \end{cases}$$

Chapter 2

Real Autoregressive Process

2.1 Definitions

Let $X_T = (X_t, t \in \mathbb{Z})$ to be a second order centred real process $[E(X_t^2 < +\infty, \forall t \in \mathbb{Z})]$. Let recall the definition of a white noise.

Definition 2.1. *We say that a process $(\varepsilon_t, t \in \mathbb{Z})$ is a weak white noise if*

1°/ $E(\varepsilon_t) = 0$

2°/ $E(\varepsilon_t \varepsilon_s) = \delta_{st}\sigma_\varepsilon^2$ *where* $\delta_{st} = \begin{cases} 1 & \text{if } s = t \\ 0 & \text{if } s \neq t \end{cases}$

The process $(\varepsilon_t, t \in \mathbb{Z})$ is a strong white noise if the variables (ε_t) are i.i.d (independent identically distributed)

Definition 2.2. *A process $(X_t, t \in \mathbb{Z})$ is called to be an autoregressive process of order p, and then denotes AR(p) if there exists $a_1, \cdots, a_p \in \mathbb{R}$ with $a_p \neq 0$ and a weak white noise $(\varepsilon_t, t \in \mathbb{Z})$ such that*

$$X_t = a_1 X_{t-1} + \cdots + a_p X_{t-p} + \varepsilon_t, \qquad t \in \mathbb{Z}$$

This relation can be written as:

$$\Lambda(L) X_t = \varepsilon_t \qquad \text{where} \qquad \Lambda(L) = 1 + a_1 L + \cdots + a_p L^p.$$

2.2 Infinite Moving Average expression

The definition above of autoregressive process does not implies necessarily that this equation has a unique stationary solution. Consequently, such a problem must be related to the question of the invertibility of the polynomial $\Lambda(L)$.

- If $\Lambda(L)$ has its roots of module different than 1, we can inverse $\Lambda(L)$ and the equation possesses a unique solution expressed as an infinite moving average:

$$X_t = \Lambda(L)^{-1}(\varepsilon_t) = \sum_{j=-\infty}^{+\infty} b_j \varepsilon_{t-j}, \qquad with \sum_{j=-\infty}^{+\infty} |b_j| < +\infty$$

  Here we recognise that the value of X_t can depends on past present and future values of ε_t.

- In the case when the roots of $\Lambda(L)$ are greater than 1 and then $\Lambda(L)^{-1}$ have only positive powers of L, the value of X_t depends only on past values of ε_t.

$$X_t = \Lambda(L)^{-1}(\varepsilon_t) = \sum_{j=0}^{+\infty} b_j \varepsilon_{t-j}, \qquad with \sum_{j=0}^{+\infty} |b_j| < +\infty \text{ and } b_0 = 1$$

We see that X_t does not linearly depend but on $\varepsilon_{t-1}, \varepsilon_{t-2}, \cdots$ and since $(\varepsilon_t, t \in \mathbb{Z})$ is a white noise, X_t is then non correlated to ε_t.

2.3 Stationarity

Let define the polynomial associated with the autoregressive process of order p, $X_t, t \in \mathbb{Z}$ by:

$$\pi(z) = z^k - \sum_{i=0}^{k} a_i z^{k_i}$$

Theorem 2.1. *A necessary and sufficient condition for the existence of a weakly stationary autoregressive process of order p; AR(p) is that the roots of the polynomial associated must be of module less than one.*

Proof. First, let's prove the theorem for an $AR(1)$

$$X_t = aX_{t-1} + \varepsilon_t.$$

Suppose that $|a| < 1$, we have

$$\begin{aligned} X_t &= \varepsilon_t + aX_{t-1} \\ &= \varepsilon_t + a(\varepsilon_{t-1} + aX_{t-2}) \\ &= \varepsilon_t + a\varepsilon_{t-1} + a^2 X_{t-2} \\ &= \cdots \\ &= \varepsilon_t + a\varepsilon_{t-1} + a^2\varepsilon_{t-2} + \cdots + a^s\varepsilon_{t-s} + a^{s+1}X_{t-(s+1)}. \end{aligned}$$

That is to say that

$$X_t = a^{s+1}X_{t-(s+1)} + \sum_{i=0}^{s} a^i\varepsilon_{t-i}.$$

Hence

$$E(X_t - \sum_{i=0}^{s} a^i\varepsilon_{t-i})^2 = a^{2(s+1)}E(X_{t-(s+1)}^2) \xrightarrow[s\to+\infty]{} 0.$$

So

$$X_t = \sum_{i=0}^{\infty} a^i\varepsilon_{t-i}$$

and then

$$E(X_t) = \sum_{i=0}^{\infty} a^i E(\varepsilon_{t-i}) = 0.$$

In addition

$$\begin{aligned} E(X_tX_{t+h}) &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a^i a^j E(\varepsilon_{t-i}\varepsilon_{t+h-j}) \\ &= \sigma_\varepsilon^2 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a^i a^j \qquad \text{does not depend on t} \end{aligned}$$

which prove the weakly stationarity of (X_t).

Now suppose that $X_t = aX_{t-1} + \varepsilon_t$ is weakly stationary. At the order $h-1$, we have

$$X_t = \sum_{i=0}^{h-1} a^i \varepsilon_{t-i} + a^h X_{t-h}.$$

And

$$R(h) = E(X_t X_{t+h}) = \sum_{i=0}^{h-1} a^i E(\varepsilon_{t-i} X_{t-h}) + a^h E(X_{t-h}^2).$$

So

$$R(h) = a^h R(0), \text{ and } \rho(h) = a^h,$$

which means that $|a| \leq 1$. However, $|a| \neq 1$. Indeed, If we suppose the contrary, for instance $a = 1$, we get

$$\begin{aligned} E(X_t X_{t+h})^2 &= E(\sum_{i=0}^{h-1} \varepsilon_{t-i})^2 \\ &= \sum_{i=0}^{h-1} \sum_{j=0}^{h-1} E(\varepsilon_{t-i}\varepsilon_{t-j}) \\ &= \sum_{i=0}^{h-1} E(\varepsilon_{t-i})^2 \\ &= h\sigma_\varepsilon^2. \end{aligned}$$

On the other hand, we have

$$\begin{aligned} E(X_t - X_{t+h})^2 &= E(X_t^2) + E(X_{t+h}^2) - 2E(X_t X_{t+h}) \\ &= 2R^2(0) - 2R(h) \\ &= \sum_{i=0}^{h-1} E(\varepsilon_{t-i})^2 \\ &= h\sigma_\varepsilon^2. \end{aligned}$$

This gives

$$R(h) = R(0) - \frac{\sigma_\varepsilon^2}{2} h \xrightarrow[h \to +\infty]{} -\infty$$

which is absurd since it led to $|R(h)| \leq R(0)$ with $\sigma_\varepsilon^2 \geq 0$.

Secondly, Let show the result for an autoregressive process of order two $AR(2)$. Let $(X_t, t \in \mathbb{Z})$ such that $X_t = a_1 X_{t-1} + a_2 X_{t-2} + \varepsilon_t, \quad \in \mathbb{Z}, \quad a_2 \neq 0$. It associated polynomial is given by

$\pi(z) = z^2 - a_1 z - a_2$.

We set

$$Y_t = \begin{pmatrix} X_t \\ X_{t-1} \end{pmatrix}, Y_{t-1} = \begin{pmatrix} X_{t-1} \\ X_{t-2} \end{pmatrix} \text{ and } e_t = \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix}.$$

We get

$$Y_t = \begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix} Y_{t-1} + \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix},$$

or $Y_t = AY_{t-1} + e_t$ which is called the markovian representation of $(X_t, t \in \mathbb{Z})$. We obtain then an autoregressive process of order one in $\mathbb{R}^2$,

Now, if we suppose that $(X_t, t \in \mathbb{Z})$ is weakly stationary. We have

$$Y_t Y_t' = (AY_{t-1} + e_t)(AY_{t-1} + e_t)' = (AY_{t-1} + e_t)(Y_{t-1}' A' + e_t').$$

Hence

$$\Gamma_Y = A\Gamma_Y A' + \Gamma_e$$

where Γ_Y and Γ_e are covariance matrixes. We state that the roots of $\pi(z)$ are the eigenvalues of the matrix A.

So, let v to be the eigenvector of A' associated to the eigenvalue λ. We get

$$A'v = \lambda v \Leftrightarrow v'A = \lambda v'.$$

Then, we can write

$$\begin{aligned} v'\Gamma_e v &= v'\Gamma_Y v - v'A\Gamma_Y Av \\ &= v'\Gamma_Y v - \lambda^2 v'\Gamma_Y v \\ &= (1-\lambda^2) v'\Gamma_Y v. \end{aligned}$$

Notice that $v'\Gamma_e v > 0$. Indeed,

$$e_t e_t' = \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix} \begin{pmatrix} \varepsilon_t & 0 \end{pmatrix} = \begin{pmatrix} \varepsilon_t^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

So

$$E(e_t e_t') = \begin{pmatrix} \sigma_t^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

And then

$$v'\Gamma_e v = \begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} \sigma_t^2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \sigma_\varepsilon^2 v_1 \\ 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \sigma_\varepsilon^2 v_1^2 > 0.$$

Consequently, $1 - \lambda^2 > 0 \Rightarrow |\lambda|^2 < 1 \Rightarrow |\lambda| < 1$.

Suppose, now that the eigenvalues of A are of module less than one. We have

$$\begin{aligned}
Y_t &= AY_{t-1} + e_t \\
&= A(AY_{t-2} + e_{t-1}) + e_t \\
&= A^2 Y_{t-2} + Ae_{t-1} + e_t \\
&= A^2(AY_{t-3} + e_{t-2}) + Ae_{t-1} + e_t \\
&= A^3 Y_{t-3} + A^2 e_{t-2}) + Ae_{t-1} + e_t \\
&= ... \\
&= A^{s+1} Y_{t-(s+1)} + A^s e_{t-s}) + \cdots + Ae_{t-1} + e_t \\
&= A^{s+1} Y_{t-(s+1)} + \sum_{j=0}^{s} A^j e_{t-j}.
\end{aligned}$$

Then, we get

$$\begin{aligned}
E[(Y_t - \sum_{j=0}^{s} A^j Y_{t-j})(Y_t - \sum_{j=0}^{s} A^j Y_{t-j})'] &= E(A^{s+1} Y_{t-(s+1)} Y_{t-(s+1)}'[A^{s+1}]') \\
&= A^{s+1} E(Y_{t-(s+1)} Y_{t-(s+1)}')[A^{s+1}]' \\
&= A^{s+1} \Gamma_Y [A^{s+1}]'.
\end{aligned}$$

Since $|\lambda| < 1$, we get $A^{s+1}\Gamma_Y[A^{s+1}]' \xrightarrow[s \to +\infty]{} 0$. And we obtain then

$$Y_t = \sum_{j=0}^{\infty} A^j \varepsilon_{t-j}, \qquad t \in \mathbb{Z}.$$

Hence

$$X_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}, \qquad t \in \mathbb{Z}$$

with

$$E(X_t) = \sum_{j=0}^{\infty} b_j E(\varepsilon_{t-j}) = 0, \qquad t \in \mathbb{Z}$$

and

$$E(X_t X_{t+h}) = \sum_{i=0} \sum_{j=0}^{\infty} b_i b_{j+h} E(\varepsilon_{t-i}\varepsilon_{t+h-j}) = \sum_{j=0}^{\infty} b_j b_{j+h}\sigma_\varepsilon^2 \qquad \text{does not depend on } t.$$

Which proves the weak stationarity of $(X_t, t \in \mathbb{Z})$. Finally and for a general autoregressive process of order p, we use the same techniques as in the order two case. □

2.4 Yule-Walker Equations

Let $(X_t, t \in \mathbb{Z})$ be a real weakly stationary autoregressive process of order p, $AR(p)$ and let R and ρ be its autocovariance and autocorrelation functions.

In the proof above, we have seen that X_t can be written as

$$X_t = \varepsilon_t + \sum_{j=1}^{p} c_j \varepsilon_{t-j}, \qquad t \in \mathbb{Z}$$

which leads to

$$E(X_t\varepsilon_t) = E(\varepsilon_t^2) + \sum_{j=1}^{p} c_j E(\varepsilon_t \varepsilon_{t-j}) = \sigma_\varepsilon^2$$

since

$$E(\varepsilon_t^2) = \sigma_\varepsilon^2 \quad \text{and } E(\varepsilon_t\varepsilon_{t-j}) = 0.$$

Hence

$$E(X_t X_t) = \sum_{j=1}^{p} c_j E(X_t X_{t-j}) + E(X_t\varepsilon_t).$$

That is to say

$$R(0) = \sum_{j=1}^{p} c_j R(j) + \sigma_\varepsilon^2 \tag{2.1}$$

We have also

$$E(X_{t+h}X_t) = \sum_{j=1}^{p} c_j E(X_{t+h}X_{t-j}) + E(X_{t+h}\varepsilon_t).$$

That is to say

$$R(h) = \sum_{j=1}^{p} c_j R(h-j). \tag{2.2}$$

The equations (2.1) and (2.2) are known as Yule-Walker equations.

The following result is derived directly

Proposition 2.1. *The autocorrelation of an AR(p) is given by*

$$\rho(h) = \sum_{j=1}^{p} c_j \rho(h-j).$$

2.5 Estimation of an AR(p) Parameters

Let $X_t = a_1X_{t-1} + \cdots + a_pX_{t-p} + \varepsilon_t$ be a real autoregressive process of a known order p, $AR(p)$, where the coefficients $a_1, \cdots, a_p$ and σ_ε are unknown.

2.5.1 Least Squares Method

Since the elements ε_t of the white noise are not correlated to the random variables $(X_{t-j})_{j=1,\cdots,p}$, the ordinary least-square method can be applied to estimate the coefficients $a_1, \cdots, a_p$.

Let $X_1, X_2, \cdots, X_N$ be N observations of the process (X_t) with $N > 2p$.

We suppose that $X_0 = X_{-1} = \cdots = X_{-(p-1)} = 0$.

Least square estimators are obtained by the regression of X_t on its past $X_{t-1}, \cdots, X_{t-p}$; that is to say

$$E(X_t/X_{t-1}, \cdots, X_{t-p}) = \sum_{i=1}^{p} a_i X_{t-i}$$

So in order to find the least square estimator, one must minimise the quantity

$$Q(a_1, \cdots, a_p) = \sum_{i=1}^{N} [X_t^{(i)} - (a_1X_{t-1}^{(i)} + \cdots + a_pX_{t-p}^{(i)})]^2.$$

We write

$$\frac{\gamma}{\gamma a_j}Q(a_1,\cdots,a_p)=\sum_{i=1}^{N}-2[X_{t-j}^{(i)}(X_t^{(i)}-a_1X_{t-1}^{(i)}-\cdots-a_pX_{t-p}^{(i)})=0,\qquad j=1,\cdots,p.$$

Thus, we obtain the following system

$$\sum_{i=1}^{N}(a_1X_{t-1}^{(i)}X_{t-j}^{(i)}+\cdots+a_pX_{t-p}^{(i)}X_{t-j}^{(i)})=\sum_{i=1}^{N}X_t^{(i)}X_{t-j}^{(i)},\qquad j=1,\cdots,p$$

whose matrix form is given by

$$\begin{pmatrix}\sum_{i=1}^{N}X_{t-1}^{(i)}X_{t-1}^{(i)} & \sum_{i=1}^{N}X_{t-2}^{(i)}X_{t-1}^{(i)} & \cdots & \sum_{i=1}^{N}X_{t-p}^{(i)}X_{t-1}^{(i)}\\ \sum_{i=1}^{N}X_{t-1}^{(i)}X_{t-2}^{(i)} & \sum_{i=1}^{N}X_{t-2}^{(i)}X_{t-2}^{(i)} & \cdots & \sum_{i=1}^{N}X_{t-p}^{(i)}X_{t-2}^{(i)}\\ \vdots & \vdots & \ddots & \vdots\\ \sum_{i=1}^{N}X_{t-1}^{(i)}X_{t-p}^{(i)} & \sum_{i=1}^{N}X_{t-2}^{(i)}X_{t-p}^{(i)} & \cdots & \sum_{i=1}^{N}X_{t-p}^{(i)}X_{t-p}^{(i)}\end{pmatrix}\begin{pmatrix}\hat{a_1}\\ \hat{a_2}\\ \vdots\\ \hat{a_p}\end{pmatrix}=\begin{pmatrix}\sum_{i=1}^{N}X_t^{(i)}X_{t-1}^{(i)}\\ \sum_{i=1}^{N}X_t^{(i)}X_{t-2}^{(i)}\\ \vdots\\ \sum_{i=1}^{N}X_t^{(i)}X_{t-p}^{(i)}\end{pmatrix}$$

and here we recognise the Yule-Walker equations

$$(\hat{R}(i,j))\begin{pmatrix}\hat{a_1}\\ \hat{a_2}\\ \vdots\\ \hat{a_p}\end{pmatrix}=(\hat{R}(0,j));$$

it sufficient to see that since R(h) (respectively R(0)) is not but the expected value of X_tX_{t+h} (respectively of X_tX_t) its estimator must be

$$\frac{1}{N}\sum_{t=1}^{N}X_{t-i}X_{t-j}=:\hat{R}(i,j)\text{ and }\frac{1}{N}\sum_{t=1}^{N}X_tX_{t-j}=:\hat{R}(0,j)\text{ (respectively)}$$

Now since $E(\varepsilon_t)=0$, we have $\sigma_\varepsilon^2=E(\varepsilon_t^2)$ and then

$$\sigma_\varepsilon^2=E[(X_t-a_1X_{t-1}-\cdots-a_pX_{t-p})^2].$$

Consequently

$$\hat{\sigma_\varepsilon^2}=\frac{1}{N}\sum_{i=1}^{N}(X_t^{(i)}-\hat{a_1}X_{t-1}^{(i)}-\cdots-\hat{a_p}X_{t-p}^{(i)})^2$$

2.5.2 Maximum Likelihood Method

Suppose that $\varepsilon_t, t \in \mathbb{Z}$ are centered gaussian i.i.d. random variables with variance σ_ε^2 and let $X_1, X_2, \cdots, X_N$ be N observations of the process (X_t) where $N > 2p$.

We suppose that $X_0 = X_{-1} = \cdots = X_{-(p-1)} = 0$.

Prestly (1981, p. 374) shows that for $N > 2p$, the likelihood function can be approched by [11]

$$L(a_1, a_2, \cdots, a_p) = -(N-p)\log(\sigma_\varepsilon^2\sqrt{2\pi}) - \frac{1}{\sqrt{2\sigma_\varepsilon^2}}\sum_{i=1}^{N}(X_t - a_1X_{t-1} - \cdots - a_pX_{t-p})^2.$$

We recognise that to maximise $L(a_1, a_2, \cdots, a_p)$ is equivalent to minimise $Q(a_1, \cdots, a_p)$. So the maximum likelihood estimator is identical to the least square estimator.

2.6 Simulation and Application

2.6.1 A Strong White Noise Simulation

As it was defined above a strong white noise process is a sequence of centered independent normal random variables. Hence, knowing that a random vector is normally distributed if and only if its components are normal random variables and in order to simulate this kind of process, it is sufficient to create a vector whose components are values derived from observations of a normal random variable.

The function in Software R that permit to simulate a normal random variable N is

$$\mathbf{rnorm}(\mathbf{n}, \mathbf{mean}, \mathbf{sd})$$

where **n** designs the number of observations, **mean** is the mean of N and **sd** is its standard deviation.

In the figure below, we have simulated a real white noise with $\sigma^2 = 0.01$

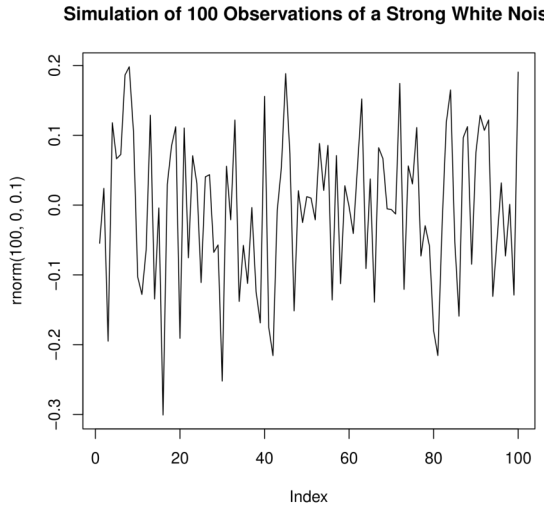


Figure 2.1

2.6.2 An Order One Autoregressive Process Simulation

In order to simulate an autoregressive process of order one, we have created a function on software R called **AR1**(**N**, **a**, **sigma2**) whose arguments are respectively the number of observations, the coefficient of autocorrelation and the variance of the white noise. The function is writing as follows:

```
AR1<-function( N, a,sigma2)
{
E = rnorm(N+2,0,sqrt(sigma2))
X = numeric(N)
X[1] = E[1]
for (i in 2:N) X[i] = E[i] + a*X[i-1]
plot(X[ 1:N ], type="l", main="Example of a One Order Autoregressive",ylab="")
}
```

Below is plotted a one order autoregressive process with an autocorrelation coefficient of 0.12 and $\sigma_\varepsilon^2 = 0.1$. The number of observations being 200 (see figure 2.2).

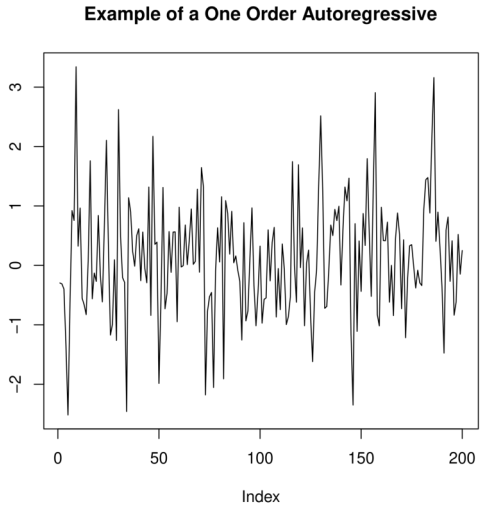


Figure 2.2

2.6.3 An Order Two Autoregressive Process Simulation

In the same way of creating the function that simulate a one order autoregressive process, we give below a function which simulate a two order autoregressive process. The function was named **AR2(N,a1,a2,sigma2)** whose arguments are respectively the number of observations, the two coefficients of autocorrelation and the variance of the white noise. The function is writing as follows:

```
AR2<-function(N=100, a1=0.18828,a2=0.05861,sigma2=1)
{
E = rnorm(N+2,,sqrt(sigma2))
X = numeric(N)
X[1] = E[3] + a1*E[2] + a2*E[1]
X[2] = E[4] + a1*X[1] + a2*E[2]
for (i in 3:N) X[i] = E[i] + a1*X[i-1]+ a2*X[i-2]
plot(X[ 1:N ], type="l", main="Example of a Two Order Autoregressive",xlab="",ylab="")
}
```

Below is plotted an autoregressive process of order two with autocorrelation coefficients a1=0.12 and a2=0.18 and $\sigma_\varepsilon^2 = 1$. The number of observations being 200 (see figure 2.3).

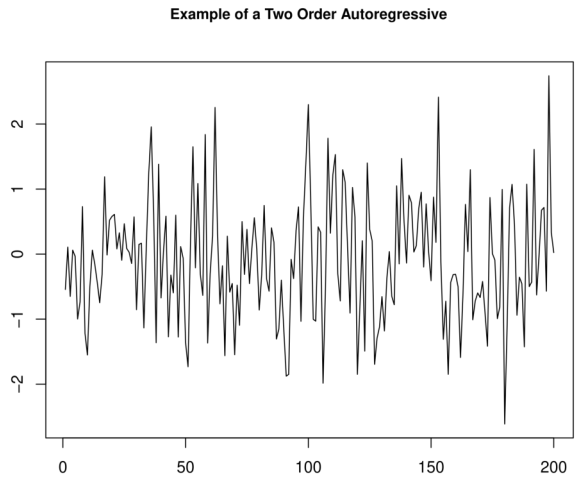


Figure 2.3

2.6.4 An Upper Order Autoregressive Process Simulation

Suppose that we want to simulate N observations of a real autoregressive process of order p. N must be at least equal to $2p$. Indeed, the p first observations will be linearly combined using the p to create the $p+1$th observation X_{p+1}. Coefficient are chosen from p realisations of a uniform random variable on the interval $[-1, 1]$ to assure the stationarity of the process. and then it X_{p+1}, using the same coefficients will be linearly combined will with its $p-1$ antecedents to create the following observation. And so on, we repeat the same algorithm until the N observation. Notice that the p first observations are simulated using the truncated moving average expression. We have written a function that illustrate this algorithm in the software R as follows:

```
ARp<-function( N=100, p=3,a=c(0.12,-0.11,0.18),sigma2=1)
{
E = rnorm(N+2,,sqrt(sigma2))
a=runif(p,-1,1)
ma<-matrix(0,p,p)
ma[,1]<-a[1]*E[1:p]
for (j in 2:p)
ma[,j]<-a[j]*c(rep(0,j-1),E[1:(p-j+1)])
X <-matrix(0,N)
```

for (i in 1:p)

X[i]<-sum(ma[i,])

for (i in (p+1):N)

X[i] =sum(a[1:p]*X[(i-p):(i-1)]) +E[i]

plot(X[1:N],main="Example of an Autoregressive of Order p", type="l", xlab="",ylab="")

return(X)

}

Below is plotted an autoregressive process of order 3 with autocorrelation coefficients a1=0.12, a2=-0.11 and a3=0.18 and $\sigma_\varepsilon^2 = 1$. The number of observations being 200 (see figure 2.4).

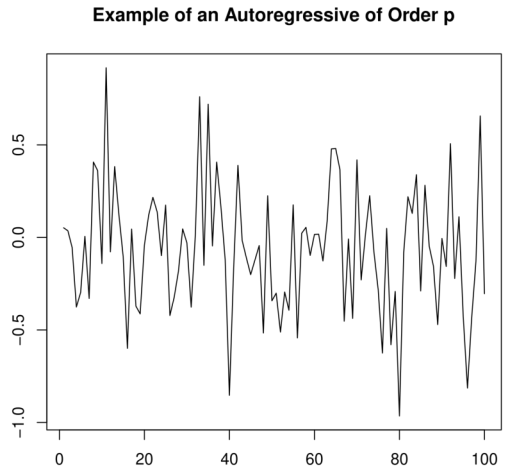


Figure 2.4

2.7 Prediction in an $AR(p)$ Model

An autoregressive process evolves as a linear regression equation in which the current value helps predict the next value. Recall that the model is written as follows

$$X_t = a_1 X_{t-1} + \cdots + a_p X_{t-p} + \varepsilon_t$$

Then the optimal prediction at the time $T+1$ viewing the time T is

$$X_{T+1}^* = E(X_{T+1}/X_T, X_{T-1}, \cdots)$$

So

$$X^*_{T+1} = a_1 X_T + \cdots + a_p X_{T-p}$$

Similary, we get

$$X_{T+h} = a_1 X_{T+h-1} + \cdots + a_p X_{T+h-p} + \varepsilon_{T+h}$$

and then $X^*_{T+h} = E(X_{T+h}/X_T, X_{T-1}, \cdots)$ is given by

$$X^*_{T+h} = \begin{cases} a_{1,T} X^*_{T+h-1} + \cdots + a_{h-1,T} X^*_{T+1} + a_h X_T + \cdots + a_p X_{T+h-p} & \text{if } h \leq p \\ a_{1,T} X^*_{T+h-1} + \cdots + a_{p,T} X^*_{T+h-p} & \text{if } h > p \end{cases}$$

Chapter 3

Random Coefficient Autoregressive Process

As linear models, autoregressive processes have large applications in time series data modeling in several areas of the physical, biological, and behavioral sciences. However, there are fields where they have become inappropriate and then a great interest has been given to the nonlinear models for such applications. Since the late 1970s, there has been an ever-increasing interest in the autoregressive model with random coefficients (RCA) because of its ability to represent a multitude of time series traits. Early contributions in this respect are due to ([3],[4]), Andel [1] who introduced the model for $p = 1$ and highlighted the different sampling of the random coefficients of the RCA model of order p introduced by Nicholls and Quinn (1982) [12] from that of Liu and Tiao (1980).

Before defining what is a random coefficients autoregressive process let recall some intersect notions for the rest of the chapter.

3.1 Preliminary Results

Definition 3.1. *[14]*

Given two matrices A and B of dimensions $m_A \times n_A$ and $m_B \times n_B$ respectively, one may define the Kronecker product $A \otimes B$ of B with A as the $m_A m_B \times n_A n_B$ matrix whose $(i, j)th$ block is the $m_B \times n_B$ matrix $a_{ij}B$, where a_{ij} is the $(i, j)th$ element of A.

Definition 3.2.

Let A be an $m_A \times n_A$ matrix. Then the $m_A n_A$-component vector $vecA$ is obtained from A by stacking the columns of A, one on top of the other, in order, from left to right. $vecA$ has as its elements the elements of A.

The results contained in the following theorem hold for any matrix products which are defined.

Theorem 3.1. *1. $vec(ABC) = (C'\otimes))'A)vecB$.*

2. $tr(AB) = (vec(B'))'vecA = (vecB)'vec(A')$.

3. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

4. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}, \quad (A \otimes B)\prime = A' \otimes B'$

Definition 3.3. *Let A be an $n \times n$ symmetric matrix. The $n(n+1)/2$-component vector $vechA$ (the "vector-half" of A) is obtained from A by stacking those parts of the columns of A, on and below the main diagonal, one on top of the other in order from left to right.*

For symmetric matrices A, it is possible to obtain by linear transformations the vector $vecA$ from the vector $vechA$, and vice versa, which is shown in the following theorem.

Theorem 3.2. *There exists constant $(n(n+1)/2) \times n^2$ matrices K_n and H_n for which*

$$vechA = H_n vecA$$

and

$$vecA = K'vechA$$

for any $n \times n$ symmetric matrix A and $H_n K_n' = I_{n(n+1)/2}$

3.2 Definition and Stationarity

Definition 3.4.

Let $(\varepsilon_t, t \in \mathbb{Z})$ is a white noise process with mean zero and variance σ_ε^2 and $b_1, b_2, ..., b_p$ be given

numbers. A process $X_t, t \in \mathbb{Z})$ is said to be of random coefficient autoregressive of order p often shortened to $RCA(p)$ if it satisfies the following formula

$$X(t) = (b_1 X_{t-1} + ... + b_p X_{t-p}) + (a_1(t)X_{t-1} + \cdots + a_p(t)X_{t-p}) + \varepsilon_t \tag{3.1}$$

where $(\boldsymbol{a}_t)_{t\in\mathbb{Z}} = (a_1(t), \cdots, a_p(t))_{t\in\mathbb{Z}}$ is a sequence of independent centred random p-vectors that are independents from ε_t.

Example 3.1. *$RCA(1)$ Let X be a process defined by $X_t = (0.12 + \xi)X_{t-1} + \varepsilon_t$ where ξ is a uniform $[0, 1]$ distributed random variable and ε_t are i.i.d standard normal random variables.*

3.3 Markovian Expression of a $RCA(p)$

the relation in (3.1) can be expressed in a matrix way. Indeed, if we put

$$Y_t = \begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \end{pmatrix}, \epsilon = \begin{pmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}, B = \begin{pmatrix} b_1 & \cdots & b_p \\ 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \text{ and } A_t = \begin{pmatrix} a_1(t) & \cdots & a_p(t) \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix},$$

then we get

$$Y_t = (B + A_t)Y_{t-1} + \epsilon_t. \tag{3.2}$$

3.3.1 Conditions for Stationarity

In this section, we will investigate the conditions under which a random coefficient autoregressive process is weakly stationary. For this purpose we state the following results from [1]

Lemma 3.3. *A time series$(X_1, \cdots, X_T)$ is stationary if and only if*

$$Var(X_1, \cdots, X_p) = Var(X_2, \cdots, X_{p+1}) \tag{3.3}$$

Proof. Since $(X_1, \cdots, X_p)$ and $Var(X_2, \cdots, X_{p+1})$ have the same law (due to stationarity), it is obvious that $Var(X_1, \cdots, X_p) = Var(X_2, \cdots, X_{p+1})$, . Now, Suppose that $Var(X_1, \cdots, X_p) = Var(X_2, \cdots, X_{p+1})$. This implies that for $k = p, p+1, \cdots, T-1$, we have

$$Var(X_1, \cdots, X_k) = Var(X_2, \cdots, X_{k+1}) \tag{3.4}$$

Indeed, let $p < k < T-1$ and let proof that for $j = 1, 2, \cdots, k$ we have

$$Cov(X_{k+1}, X_{j+1}) = Var(X_k, X_j) \tag{3.5}$$

For $t = k+1$, we multiply the expression (3.1) by X_{j+1}, then by applying the expectation to its two sides we get

$$Cov(X_{k+1}, X_{j+1}) = \sum_{i=1}^{p} b_i Cov(X_{k+1-i}, X_{j+1}) + E(\varepsilon_{k+1} X_{j+1}).$$

And for $t = k1$, we multiply the expression (3.1) by X_j, and by applying again the expectation to its two sides we get

$$Cov(X_k, X_j) = \sum_{i=1}^{p} b_i Cov(X_{k-i}, X_j) + E(\varepsilon_k X_j)$$

Notice that due to the fact that ε_t are centered and independent of ulterior X_t, we have $E(\varepsilon_{k+1} X_{j+1}) = E(\varepsilon_k X_j) = 0$.
In addition,for $i = 1, 2, \cdots, p$ and due to (3.4), we have $Cov(X_{k+1-i}, X_{j+1}) = Cov(X_{k-i}, X_j)$
So (3.5) is proved for $j < k$. Let now $j = k$. In the same way we get

$$E(X_{k+1}^2) = \sum_{i=1}^{p} E(c_i(k+1) X_{k+1} X_{k+1-i}) + E(\varepsilon_{k+1} X_{k+1}) \tag{3.6}$$

$$E(X_k^2) = \sum_{i=1}^{p} E(c_i(k) X_k X_{k-i}) + E(\varepsilon_k X_k) \tag{3.7}$$

where $c_i(t) = b_i + a_i(t)$.
For $t = p+1, \cdots, T$ and from (3.1), we have $E(X_t \varepsilon_t) = \sigma_\varepsilon^2$

So, for $t = p+1, \cdots, T$,$s = 1, 2, \cdots, t-1$ we have

$$\begin{aligned} E(c_i(t)X_tX_{t-s}) &= \sum_{j=1}^{p} E(c_i(t)c_j(t))E(X_{t-j}X_{t-s}) \\ &= \sum_{j=1}^{p} E(\delta_{ij} + b_ib_j)Cov(X_{t-j}, X_{t-s}). \end{aligned}$$

where δ_{ij} are the elements of the covariance matrix of the random vector $\mathbf{c}(t) = (c_1(t), \cdots, c_p(t))$ Hence, for $i = 1, 2, \cdots, p$ we get

$$E(c_i(k+1)X_{k+1}X_{k+1-i}) = \sum_{j=1}^{p} E(\delta_{ij} + b_ib_j)Cov(X_{k+1-j}, X_{k+1-i}).$$

and

$$E(c_i(k)X_kX_{k-i}) = \sum_{j=1}^{p} E(\delta_{ij} + b_ib_j)Cov(X_{k-j}, X_{k-i}).$$

Now, for $i = 1, 2, \cdots, p$ and $j = 1, 2, \cdots, p$ we have already supposed that

$$Cov(X_{k+1-j}, X_{k+1-i}) = Cov(X_{k-j}, X_{k-i}).$$

So we can see that (3.6) and (3.7) implies that $E(X_{k+1}^2) = E(X_k^2)$. here we recognise (3.5) for $j = k$. Hence, the formula (3.4) is proved and then by putting $k = T-1$ we see that the time series $(X_1, \cdots, X_T)$ is stationary. □

Let $V = Var(X_1, \cdots, X_p)'$, $W = Var(a_1(t), \cdots, a_p(t))'$ and J a $p \times p$-matrix with $J_{11} = 1$ and $J_{ij} = 0$ for $i \neq 1, j \neq 1$.

Lemma 3.4. *The formula (3.3) is verified if and only if V satisfies*

$$V = BVB' + (\sigma_\varepsilon^2 + Tr(WV))J \tag{3.8}$$

Proof. let the $p \times p$-matrix C_t defined by $C_t = B + A_t$ for $t = p+1, \cdots, T$. According to (3.2), we have

$$Y_{p+1} = C_{p+1}Y_p + \epsilon_{p+1}.$$

Applying variance to the two sides, we obtain

$$Var(X_1, \cdots, X_p)' = BVB' + \sum_{i=1}^{p}\sum_{j=1}^{p} \delta_{p+1-i,p+1-j} c_{ji} J + \sigma_\varepsilon^2 J.$$

It is sufficient now to see that

$$\sum_{i=1}^{p}\sum_{j=1}^{p} \delta_{p+1-i,p+1-j} c_{ji} = Tr(WV)$$

□

Lemma 3.5. *If the following condition is satisfied*

$$z^p - b_1 z^{p-1} - b_2 z^{p-2} - \cdots_- b_p \neq 0, for |z| \geq 1,$$

then It exists a unique solution V_0 to the equation

$$V = BVB' + \sigma_\varepsilon^2 J \tag{3.9}$$

given explicitly by

$$V_0 = \sum_{i=0}^{\infty} B^i (\sigma_\varepsilon^2 J) B'^i.$$

Moreover, this solution is a positive definite matrix.

Proof. Suppose that (3.3) is satisfied then the eigenvalues of the matrix B are inside the unit circle (notice that $z^p - b_1 z^{p-1} - b_2 z^{p-2} - \cdots_- b_p$ is not but the characteristic polynomial of the matrix B.

Now, let's introduce the operator *vec* in the formula (3.9). We obtain

$$vec(V_0) = B^{\otimes 2} vec(V_0) + vec(\sigma_\varepsilon^2 J)$$

where $B^{\otimes 2}$ is a $p^2 \times p^2$-matrix with eigenvalues $\lambda_i \lambda_j$. λ_i and λ_j are the eigenvalues of the matrix B satisfying $|\lambda_i| < 1$ for all i.

So, $B^{\otimes 2}$ eigenvalues are inside the unit circle. Then $I_{p2} - B^{\otimes 2}$ and hence the equation (3.9) got

a solution given by

$$vec(V_0) = (I_{p2} - B^{\otimes 2})^{-1} vec(\sigma_\varepsilon^2 J).$$

by expanding the matrix, we get

$$\begin{aligned} vec(V_0) &= (I_{p2} - B^{\otimes 2})^{-1} vec(\sigma_\varepsilon^2 J) \\ &= (I_{p2}^{\otimes 2} + B^{\otimes 2} + B^{\otimes 4} + \cdots) vec(\sigma_\varepsilon^2 J) \\ &= I_{p2}^{\otimes 2} vec(\sigma_\varepsilon^2 J) + B^{\otimes 2} vec(\sigma_\varepsilon^2 J) + B^{\otimes 4} vec(\sigma_\varepsilon^2 J) + \cdots \\ &= vec(I_p \sigma_\varepsilon^2 J I_p') + vec(B^2 \sigma_\varepsilon^2 J B'^2) + vec(B \sigma_\varepsilon^2 J B'^2) + \cdots \\ &= \sum_{i=0}^{\infty} vec(B^i \sigma_\varepsilon^2 J B'^i) \end{aligned}$$

Consequently

$$V_0 = \sum_{i=0}^{\infty} B^i (\sigma_\varepsilon^2 J) B'^i$$

□

Lemma 3.6. *Suppose that*

$$z^p - b_1 z^{p-1} - b_2 z^{p-2} - \cdots_- b_p \neq 0, for |z| \geq 1.$$

If the following condition is satisfied

$$1 - \sigma_\varepsilon^{-2} Tr(W V_0) > 0,$$

Then (3.8) has a unique solution given by

$$V = (1 - \sigma_\varepsilon^{-2} Tr(W V_0))^{-1} V_0 \tag{3.10}$$

Proof. According to the lemma (3.5), a solution of the equation (3.8) must satisfy the following equality

$$V = (1 + \sigma_\varepsilon^{-2} Tr(WV)) V_0 \tag{3.11}$$

which has the form

$$V = \alpha V_0 \tag{3.12}$$

where α is a real number. By combining (3.11) and (3.12), we get

$$\alpha(1 - \sigma_\varepsilon^{-2} Tr(WV_0)) = 1 \tag{3.13}$$

If $1 - \sigma_\varepsilon^{-2} Tr(WV_0) = 0$, there will be no solution. Else

$$V = (1 - \sigma_\varepsilon^{-2} Tr(WV_0))^{-1} V_0 \tag{3.14}$$

is a solution of (3.8) whose uniqueness is assured by (3.12) and (3.13).
However, the matrix given in (3.14) is a positive definite one if $1 - \sigma_\varepsilon^{-2} Tr(WV_0) > 0$ □

Lemmas 3.3 to 3.6 allow us to prove the following result

Theorem 3.7. *The p-order time series $X_t, \cdots, X_N$ is stationary if and only if all the roots of the equation $z^p - b_1 z^{p-1} - b_2 z^{p-2} - \cdots_- b_p \neq 0$ are inside the unit circle and the matrix $Var(X_t, \cdots, X_p)'$ is given by the formula (3.10)*

Let A_t be the σ-algebra generated by $\{\mathbf{a}_s, \varepsilon_s), s \leq t\}$

Corollary 3.1. *It exists a unique stationary A_t-mesurable solution for (3.1) if and only if B has all its eigenvalues in the unit circle and $\Gamma\Lambda < 1$ where $\Gamma = E(\boldsymbol{a}_t \otimes \boldsymbol{a}_t')$ and Λ is the last column of $(I - B \otimes B)^{-1}$*

Notice that from corollary (3.1), it follows that if $\Sigma = E(\mathbf{a}_t' \mathbf{a}_t)$, then it is easy to see that

$$vec(\sum) = E(\mathbf{a}_t' \otimes \mathbf{a}_t') = (E(\mathbf{a}_t \otimes \mathbf{a}_t))' = \Gamma'$$

3.4 Parameters Estimation

Knowing that the matrix $\sum$ is symmetric, it follows that it is sufficient to estimate only the $\upsilon = vech(\sum)$.

First of all, we are going to estimate $b_1, \cdots, b_p$. The model (3.1 can be written as

$$X_t = B'Y_{t-1} + u_t \tag{3.15}$$

where $B = (b_1, \cdots, b_p)$ and $u_t = \mathbf{a}_t Y_{t-1} + \varepsilon_t$.

We have

$$E(u_t/A_t) = E(\mathbf{a}_t)Y_{t-1}) + E(\varepsilon_t) = 0$$

and

$$\begin{aligned}
E(u_t^2/A_{t-1}) &= = E((\mathbf{a}_t Y_{t-1})^2/A_{t-1}) + 2E(\mathbf{a}_t Y_{t-1}\varepsilon_t/A_{t-1}) + E(\varepsilon_t^2) \\
&= E(Y'_{t-1}\mathbf{a}'_t\mathbf{a}_t Y_{t-1}/A_{t-1}) + E(\varepsilon_t)E(\mathbf{a}_t Y_{t-1}/A_{t-1}) + \sigma_\varepsilon^2 \\
&= Y'_{t-1}E(\mathbf{a}'_t\mathbf{a}_t/A_{t-1})Y_{t-1} + \sigma_\varepsilon^2 \\
&= Y'_{t-1}\Sigma Y_{t-1} + \sigma_\varepsilon^2 \\
&= (Y'_{t-1} \otimes Y'_{t-1})vec(\Sigma) + \sigma_\varepsilon^2 \\
&= (vec(Y_{t-1}Y'_{t-1}))'K'_p vech(\Sigma) + \sigma_\varepsilon^2
\end{aligned}$$

So

$$E(u_t^2/A_{t-1}) = Z'_t\upsilon + \sigma_\varepsilon^2 = \upsilon' Z_t + \sigma_\varepsilon^2 \tag{3.16}$$

where $Z_t = K_p vech(Y_{t-1}Y'_{t-1})$.

Observing $X_0, \cdots, X_N$, we give the least square estimator $\hat{B}_N$ of $B = (b_1, \cdots, b_p)'$, that min-

imise $\sum_{i=1}^{N} u_i^2$ with respect to B in relation (3.15)by:

$$\hat{B}_N = \frac{\sum_{i=1}^{N} Y_{t-1} X_t}{\sum_{i=1}^{N} Y_{t-1} Y_{t-1}'} \tag{3.17}$$

and according to the relation (3.15), we get for $i = 1, \cdots, N$

$$\hat{u_i} = X_{i-1} - \hat{B}_N Y_{i-1}$$

Now, to estimate δ and σ_ε^2 we must minimise the quantity $\sum_{i=1}^{N} (u_i^2 - \sigma_\varepsilon^2 - Z_t'\delta)^2$ with respect to δ and σ_ε^2 and then obtain

$$\hat{\delta}_N = \frac{\sum_{i=1}^{N} (Z_i - \overline{Z})(Z_i - \overline{Z})'}{\sum_{i=1}^{N} \hat{u}_i^2 (Z_i - \overline{Z})} \tag{3.18}$$

and

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N} \sum_{i=1}^{N} u_i^2 - \hat{\delta}_N \overline{Z} \tag{3.19}$$

It is shown in [11] that $\hat{B}_N$ converge almost sure to B and if $E(X_t^4 < \infty$,

$$\sqrt{N}(\hat{B}_N - B) \Rightarrow \mathcal{N}(0, \Xi)$$

with $\Xi = \sigma_\varepsilon^2 \Delta + \Delta E(X_{t-1} X_{t-1}' \delta' Z_t) \Delta$ where $\Delta = [E(X_{t-1} X_{t-1}')]^{-1}$

Also it shows that both $(\hat{\delta} - \delta)$ and $(\hat{\sigma}_\varepsilon^2 - \sigma_\varepsilon^2)$ converge to zero almost sure and if $E(X_t^4 < \infty$, $\sqrt{N}(\hat{\delta} - \delta)$ and $\sqrt{N}(\hat{\sigma}_\varepsilon^2 - \sigma_\varepsilon^2)$ converge in probability to zero

3.5 Simulation of Random Coefficient Autoregressive

3.5.1 An Order One Autoregressive Process Simulation

A random coefficient autoregressive of order one $RCA(1)$ process $(X_t, t \in \mathbb{Z})$ is said to be of p often shortened to if it satisfies the following formula

$$X(t) = bX_{t-1} + a_t X_{t-1} + \varepsilon_t \tag{3.20}$$

where $(a_t)_{t\in\mathbb{Z}}$ is a sequence of independent identically distributed centred random variables with variance σ_a^2 and which are independents from the white noise (ε_t) of variance σ_ε^2 and b is a given number.

In order to simulate a random autoregressive process of order one, let see that It can be written also

$$X(t) = (b + a_t)X_{t-1} + \varepsilon_t = c_t X_{t-1} + \varepsilon_t \tag{3.21}$$

where the random variable c_t is of mean b and variance σ_a^2. So, we have created a function on software R called **RCA1**$(\mathbf{N}, \mathbf{b}, \mathbf{sigma2}, \mathbf{sigma2a})$ whose arguments are respectively the number of observations N, the given constant b, the variance of the white noise and variance of a. The function is writing as follows:

```
RCA1<-function( N, b=0.8,sigma2=0.1,sigma2a=0.1)
{
E = rnorm(N,,sqrt(sigma2))
c = rnorm(N,b,sqrt(sigma2a)) # c=runif(,Nb-1,b+1)
X = numeric(N)
X[1] = E[1]
x[2]<-(c[2])*x[1]+e[2]
for (i in 3:N) X[i] = E[i] -c[i]*X[i-1]
plot(X[ 1:N ], type="l", main="Example of a Random Coefficient One Order Autoregressive",ylab="")
```

}

Below is plotted a one order random coefficient autoregressive process with c centered gaussian random variables respectively $[b-1, b+1]$-uniform random variable (see figure 3.1 respectively 3.2). Other choices of random variables with mean b are possible.

Notice that to ensure the stationarity of the simulated process one may choose b and σ_a^2 such that $b^2 + \sigma_a^2 < 1$.

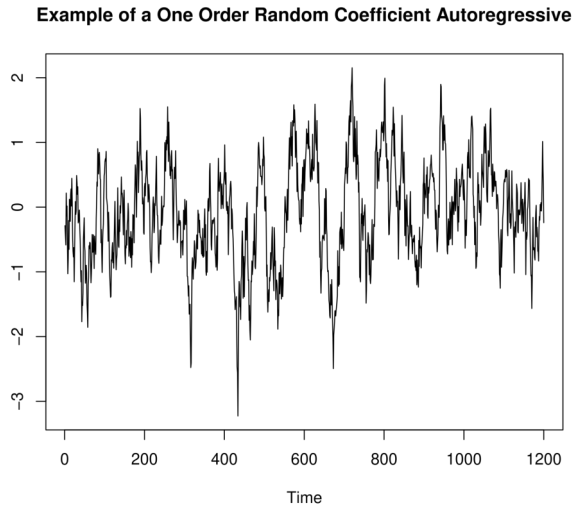


Figure 3.1

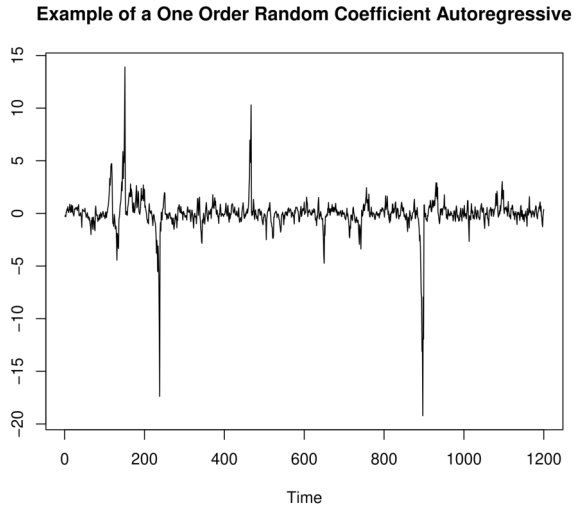


Figure 3.2

3.5.2 An Order Two Autoregressive Process Simulation

In the light of the principle of simulation given in [11] that generate a random coefficients autoregressive processes, where the two sequences $a(t)$ and ε_t are gaussian, we follow the current steps:

1. Calculate the eigenvalues $\lambda_i; i = 1, \cdots, p$ of the matrix B which must be less than one in module.

2. Calculate the parameters $b_i; i = 1, \cdots, p$ of

$$\prod_{i=1}^{p}(1 - \lambda_i z) = 1 - \sum_{i=1}^{p} b_i z^i$$

   and then calculate the square matrix M such as $vec(M)$ is the last collone of $(I - M \otimes M)^{-1}$.

3. Calculate $tr(\Gamma^* M)$ for a positive defined matrix Γ^*.

4. Calculate the diagonal positive lower triangular matrix L such that $LL' = \Gamma$

5. Generate a vector $(v_1, v_2, \cdots, v_{p+1})'$ of random numbers normally distributed.
   Taking $e_t = \sigma_\varepsilon v_1$ where $\sigma_\varepsilon^2 = E(\epsilon_t^2$ and $V_t' = L(v_1, v_2, \cdots, v_{p+1})'$ to insure that e_t and $\mathbf{a}(t)$ are theoretically independents with zero means and that $E(\mathbf{a}(t)'\mathbf{a}(t)) = LL' = \Gamma$.

6. Calculate $X_t = \sum_{i=1}^{p}(b_i + a_i(t))X_{t-i}$ where $X_t = 0$ for $t \leq 0$

7. Repeat step 5 and 6 for $N + k$ times where N is the length of the desired sample and ignore the k first produced values to ensure the stability of X_t.

We give below a function which simulates a two order random coefficient autoregressive process. The function was named **RCA2(N,lamnda1,lambda2,sigma2)** whose arguments are respectively the number of observations, the two eigenvalues of B and the variance of the white noise. The function is writing as follows:

RCA2<-function(N=2500,lambda1=0.2,lambda2=0.3,sigma2=0.1)

{

b1<-lambda1+lambda2

b2<-lambda1*lambda2

```
M<-matrix(c(0,b2,1,b1),nrow=2)
D<-matrix(c(1,0,0,(-b2)^2,0,1,(-b2),(-b1*b2),0,(-b2),1,(-b2*b1),-1,-b1,-b1,1-b1^2),nrow=4)
S<-solve(D)
V<-matrix(S[,4],nrow=2)
J=matrix(c(1,0,0,1),nrow=2)
T<-Y %*% V
c<-sum(diag(T))
r=0.9
K<-Y*(r/c)
L<-chol(K)
X<-rep(0,N)
v<-rnorm(3)
X[1]<-v[1]*sqrt(sigma2)
v<-rnorm(3)
A<-L%*% v[2:3]
X[2]<-(b1+A[1])*X[1]+sqrt(sigma2)*v[1]
for(i in 3:N)
{
v<-rnorm(3)
A=L%*% v[2:3]
X[i]=(b1+A[1])*X[i-1]+(b2+A[2])*X[i-2]+sqrt(sigma2)*v[1]
}
plot(X,type="l",main="Example of a Random Coefficient Autoregressive Process
of Order Two")
}
```

Below is plotted a random coefficient autoregressive process of order two (see figure 3.3).

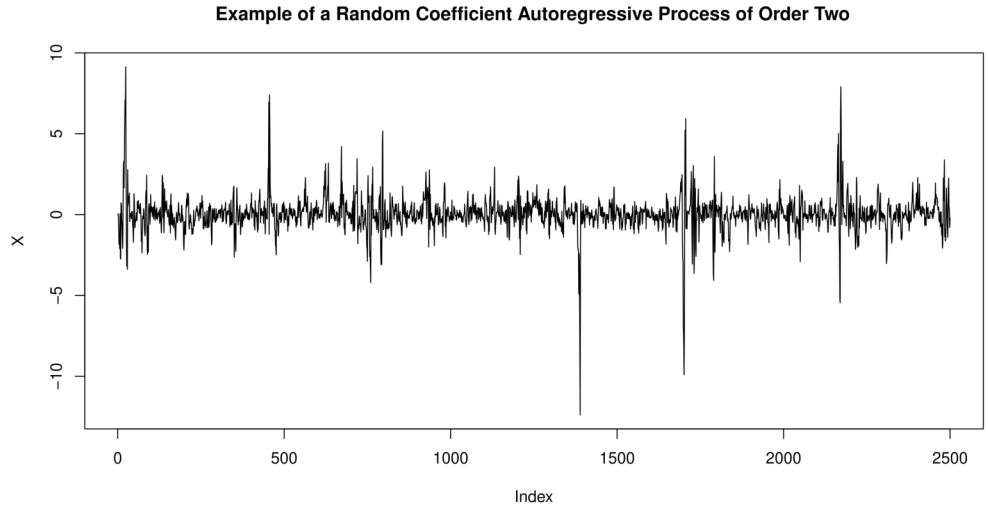


Figure 3.3

3.6 Prediction in $RCA(p)$ Model

According to its definition (3.1) it seems that the best predictor in the sense of least squares for a random coefficient autoregressive process (X_t) knowing $X_{t-1}, X_{t-2}, \cdots$ is $\hat{X}_t = E(X_t/A_{t-1}) = B'Y_{t-1}$. However, this predictor is linear and cannot match with the nonlinear nature of (3.1) which does not allow to the process (X_t) to be normal or even to be near. Since $E(X_t^2/A_{t-1}) = (B'Y_{t-1})^2 + \sigma_\varepsilon^2 + \delta' Z_t$, the natural predictor of (X_t) knowing $X_{t-1}, X_{t-2}, \cdots$ is

$$X_t^* = sgn(B'Y_{t-1})((B'Y_{t-1})^2 + \sigma_\varepsilon^2 + \delta' Z_t)^{\frac{1}{2}}$$

where $sgn(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}$

Chapter 4

Comparison and Application

According to the simulations above one can see that there is a certain similarity between a RCA model and an AR one. Indeed, a RCA model can be seen as an AR model to whom is added a random perturbation to its coefficients. To illustrate such similarity we have create a function that simulate and plot an autoregressive process of order one (in red) and an random coefficient autoregressive process of order one (in blue) where the same autocorrelation coefficient in the first was perturbed by a normal (respectively a uniform) random variable (see figure 4.1 and figure 4.2 respectively).

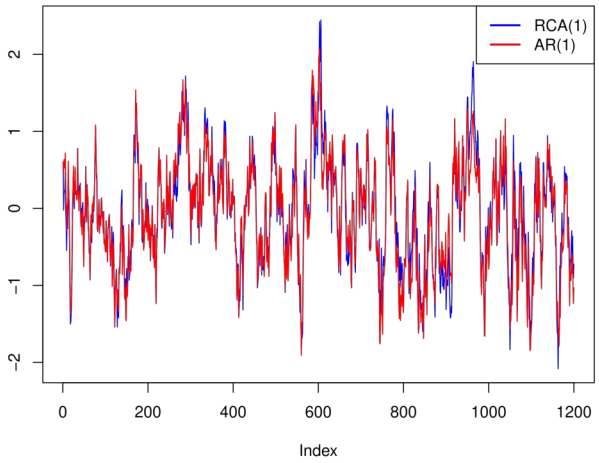


Figure 4.1

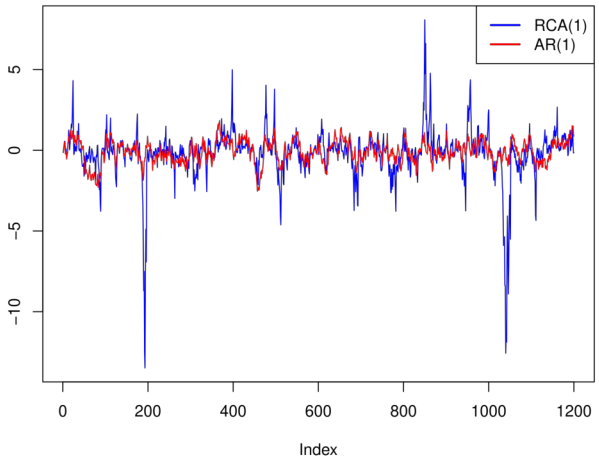


Figure 4.2

The function that gives the above plot is given by:

```
RCA1vsAR1<-function(N=1200,b=0.9,sigma2=0.1,sigma2a=0.1)
{
e<-rnorm(N,0,sqrt(sigma2))
X = numeric(N)
X[1] = e[1]
for (i in 2:N) X[i] = e[i] + b*X[i-1]
x<-rep(0,N)
a<-rnorm(1200,0,sigma2a) #a<-runif(-1,1)
x[1]<-e[1]
x[2]<-(a[2])*x[1]+e[2]
for( i in 3:1200)
x[i]<-(b+a[i])*x[i-1]+e[i]
plot(x,, main="",ylab="",col="blue",type="l")
lines(X, col="red")
legend(x = "topright",legend = c("RCA(1)", "AR(1)"),lty = c(1, 1),
col=c("blue","red"),lwd = 2)
}
```

From Figure 4.1, it can be seen that there is no significant fluctuation throughout the

series of the $AR(1)$ process. The process seems to have no systematic change in the mean and variance.

On the other hand, the maximal behavior of $ARC(1)$ in Figure 4.2, (3.3) is significantly different from $AR(1)$. It is clear that the additive random perturbation caused some amplitude jumps in the series.

4.0.1 Application to Simulated Data

First of all, we simulate N observations of an order one random coefficient autoregressive process with parameters $b = 0.9$, $\sigma_\varepsilon^2 = 1$ and $\sigma_a^2 = 0.25$:

RCA1(N=600,b=0.9,sigma2=1,sigma2a=0.1)

Recall that these parameters estimators are respectively given by [memoire master p28,29]:

$$\hat{b} = \frac{\sum_{i=1}^{N} X_i X_{i-1}}{\sum_{i=1}^{N} X_{i-1}^2},$$

$$\hat{\sigma_\varepsilon}^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \hat{b}X_{i-1})^2 - \hat{\sigma_a}^2 \frac{1}{N}\sum_{i=1}^{N} X_{i-1}^2$$

and

$$\hat{\sigma_a}^2 = \frac{\sum_{i=1}^{N} X_{i-1}^2 (X_i - \hat{b}X_{i-1})^2 - \sum_{i=1}^{N} \hat{\sigma_\varepsilon}^2 X_{i-1}^2}{\sum_{i=1}^{N} X_{i-1}^4}$$

Then we write a function that has as output in the software R the estimations of those parameters using the least square method:

RCA1esimation<-function(x)

{

n<-length(x)

MCE<-rep(0,n)

hatb<-sum(x[2:n]*x[1:(n-1)])/sum((x[1:(n-1)])^2)

MCE<-x[2:n]-hatb*x[1:(n-1)] hatsigma2a<-sum((MCE^2)*((x[1:(n-1)])^2-mean((x[1:(n-1)])^2)))/sum(((x[1:(n-1)])^2-mean((x[1:(n-1)])^4)

hatsigma2<-mean(MCE^2)-hatsigma2a*mean((x[1:(n-1)])^2)

return(list("hatb"=hatb,"hatsigma2a"=hatsigma2a,"hatsigma2"=hatsigma2))

}

we get for instance

RCA1estimation(RCA1(600,0.3,1,0.25))

$ hatb # $\hat{b}$

0.3000723

$ hatsigma2a # $\hat{\sigma_\varepsilon}^2$

0.3600764

$ hatsigma2 # $\hat{\sigma_a}^2$

0.822556

Now, we give a function that takes the m first observations, models them with both random coefficient autoregressive process of order one $RCA(1)$ and ordinary autoregressive process of order one $AR(1)$ and calculates their respective quadratic errors when predicting the last $N-M$ observations using the both $RCA(1)$ and $AR(1)$ models/ The function is given by:

RCA1vsAR1Prediction<-function(x,n=600,m=200)

{

n<-length(x)

if (n<m)

stop("n must be greater than m")

else

MCE<-rep(0,m)

hatb<-sum(x[2:m]*x[1:(m-1)])/sum((x[1:(m-1)])$\hat{2}$)

MCE<-x[2:m]-hatb*x[1:(m-1)]

hatsigma2a<-sum((MCE$\hat{2}$)*((x[1:(m-1)])$\hat{2}$-mean((x[1:(m-1)])$\hat{2}$)))/sum(((x[1:(m-1)])$\hat{2}$-mean((x[1:(m-1)])$\hat{2}$))$\hat{2}$)

hatsigma2<-mean(MCE$\hat{2}$)-hatsigma2a*mean((x[1:(m-1)])$\hat{2}$)

hatx1<-rep(0,n)

```
for(i in 1:(n-m-1))
hatx1[i]=x[i]
for(i in (n-m):n)
hatx1[i]=sign(hatb*x[i-1])*sqrt((hatb*x[i-1])^2+hatsigma2+hatsigma2a*(x[i-1]^2))
QR1<-sqrt(sum((hatx1-x)^2))
e<-rnorm(n,0,(1/m)*sum(MCE^2))
hatx2<-rep(0,n)
for(i in 1:(n-m-1))
hatx2[i]=x[i]
for(i in (n-m):n)
hatx2[i]=hatb*x[i-1]+e[i]
QR2<-sqrt(sum((hatx2-x)^2))
return(list("RCA1-quadratic-error"=QR1,"AR1-quadratic-error"=QR2))
}
RCA1vsARPrediction(RCA1())
```

We get

```
$ RCA1-quadratic-error
5.97764
$ AR1-quadratic-error
5.980647
```

We remark that the quadratic error corresponding to the $RCA(1)$ model is less than the quadratic error corresponding to the $AR(1)$ model. Within the same function we can plot the predicted future values.

Notice that the model is better predicted by the random coefficient autoregressive model than the ordinary one (see figure 4.3 below):

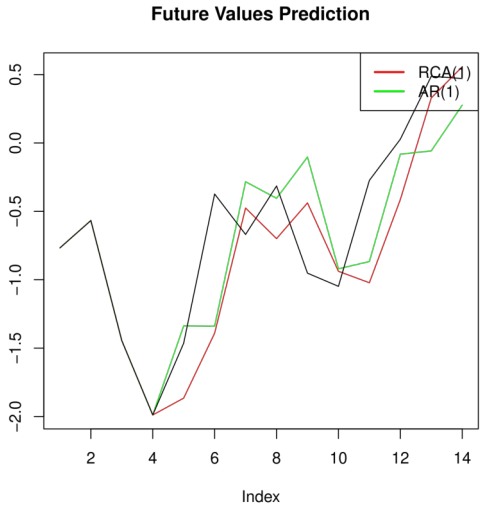


Figure 4.3

4.0.2 Application to Realistic Data

The Southern Oscillation Index (SOI) is a measure of the intensity or strength of the Walker Circulation. It is one of the key atmospheric indices for gauging the strength of El Niño and La Niña events and their potential impacts on the Australian region. Where El Niño and La Niña events are a natural part of the global climate system. They occur when the Pacific Ocean and the atmosphere above it change from their neutral ('normal') state for several seasons. El Niño events are associated with a warming of the central and eastern tropical Pacific, while La Niña events are the reverse, with a sustained cooling of these same areas (see $http://www.bom.gov.au/climate/enso/history/ln-2010-12/SOI-what.shtml$).

The Southern Oscillation Index is defined as the standardized difference between barometric readings at Darwin, Australia and Tahiti. We have chosen the index through the month of January from 1951 to 2022. So, we have $n = 72$ observations. We used $m = 50$ observations to be modeled consecutively by an $AR(1)$ and a $RCA(1)$ and then make predictions about the reminded 22 observation for both models and compare their quadratic errors. Data was registered in an excel file named **soijanvier** of extension ".csv" separator ";" (data are given in appendix A). The file was imported to the console R using the command:

soi<-read.csv2("soijanvier.csv")

and the object **soi** was attached:

attach(soi)

and target data were structured in R as follows:

data<-as.double(data).

Then we applied the function **RCA1vsAR1Prediction** to the argument **data**.

We get

$ RCA1-quadratic-error

18.16607

$ AR1-quadratic-error

28.93774

We remark that the quadratic error corresponding to the RCA(1) model is less than the quadratic error corresponding to the AR(1) model.

Remark 4.1. *For both simulated data and realistic one the quadratic error when modeling and predicting using a random coefficients autoregressive model is less than the quadratic error when modeling and predicting using a deterministic coefficients autoregressive model and then it is more convenient to predict the future using RCA model rather than AR ones (see figure 4.4 and figure 4.5).*

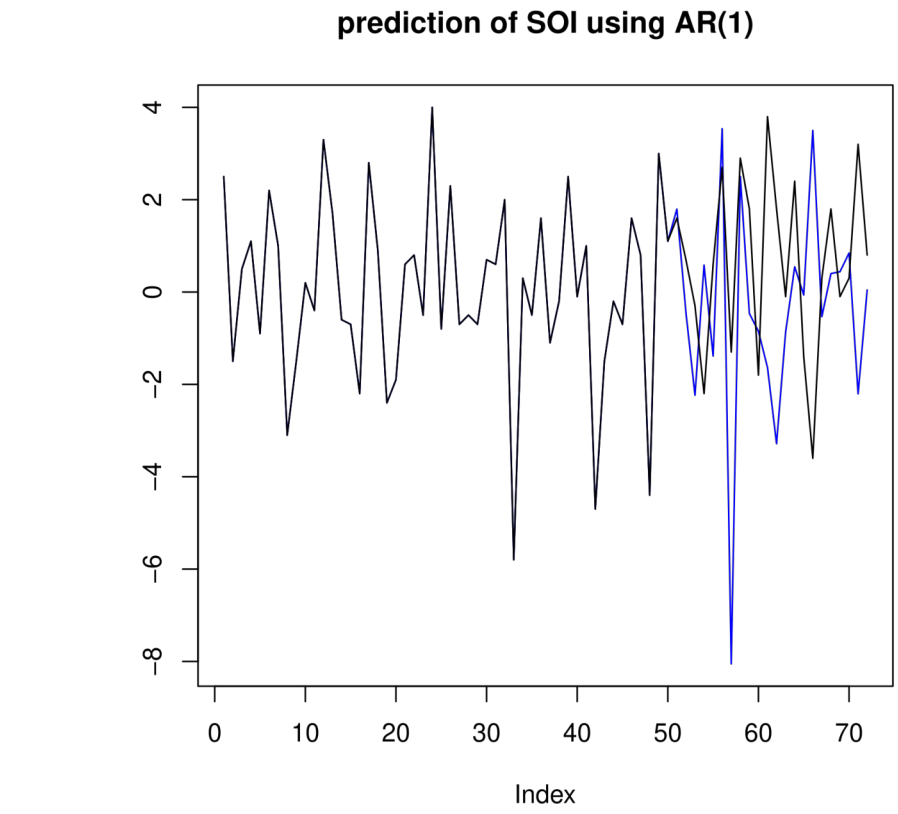


Figure 4.4

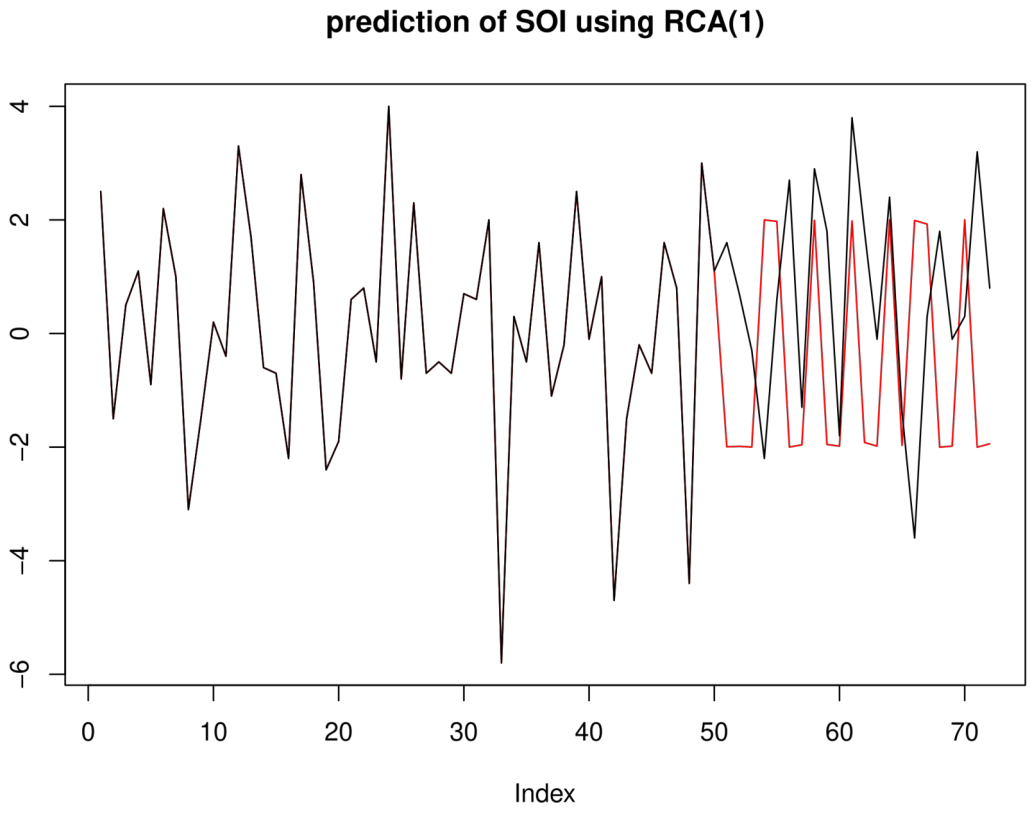


Figure 4.5

Conclusion and Perspectives

The autoregressive model AR is a tool used in time series analysis to describe and model time series data. Its main structure is a linear equation using the previous values to compute the next time step; i.e., the short time relationship is the core component of the autoregressive model. Therefore, short-term effects can be modeled in an easy way, but the global structure of the model is not obvious. In addition, classical autoregressive model is not always suitable for modeling all temporal data, especially in certain fields such as climatology and biology, where there are disturbances random, as well as in finance where the presence of heteroscedasticity and then the necessity of the emergence of other forms of models such as random coefficients autoregressive models RCA.

In this work, we presented some probabilistic properties of classical autoregressive process as well as those with random coefficient. According to the proposed model and the conditions imposed on the coefficients of the model, we have made some statistical inference in these models by applying the ordinary least squares method and then obtain good properties of such estimators. We find that the root mean square error of the modeling by the ordinary autoregressive model is superior to that by the random coefficients autoregressive model for both simulated and realistic data so that we consider that modeling data using random coefficients autoregressive process is much better than a deterministic coefficients ones.

In last decay, several scientific researches have been elaborated about random coefficients autoregressive processes to model a large number of time series data. Yet, this class of modeling is now the object of in-depth research by many researchers where we the subject can be approached from many aspects, including statistical inference by finding other more efficient statistical methods such as the use of non-parametric estimation methods. We can mention also that so far, we have assumed that the order of the autoregressive model is known. However,

in practice, this quantity is generally unknown, and we then seek to estimate its value from the observations, which is commonly referred to as the model selection problem. Moreover, a generalization of these processes to the Functional case, where the coefficient of the representation is a random operator is also a very interesting research topic.

Bibliography

[1] Anděl, J. (1976). Autoregressive series with random parameters. Mathematische Operationsforschung und Statistik, 7(5), 735-741.

[2] Aue, A., Horváth, L., & Steinebach, J. (2006). Estimation in random coefficient autoregressive models. Journal of Time Series Analysis, 27(1), 61-76.

[3] Conlisk, J. (1974). Stability in a random coefficient model. International Economic Review, 529-533.

[4] Conlisk, J. (1976). A further note on stability in a random coefficient model. International Economic Review, 759-764.

[5] Diaz, J. (1990). Bayesian forecasting for AR (1) models with normal coefficients. Communications in statistics-theory and methods, 19(6), 2229-2246.

[6] Dufour, J-M. (2003). Histoire de l'analyse des séries chronologiques. www.researchgate.net.

[7] Fink, T., Kreiss, J.P., (2013). Bootstrap for Random Coefficient Autoregressive Models. Journal of Time Series Analysis. 34(6), 646-667.

[8] Garbade, K. (1977). Two methods for examining the stability of regression coefficients. Journal of the American Statistical Association, 72(357), 54-63.

[9] Hwang, S. Y., & Basawa, I. V. (1993). Parameter estimation in a regression model with random coefficient autoregressive errors. Journal of statistical planning and inference. 36(1), 57-67.

[10] Kendall, M. G., & Hill, A. B. (1953). The analysis of economic time-series-part i: Prices.

Journal of the Royal Statistical Society. Series A (General), 116(1), 11-34.

[11] Mourid, T. (1982). Sur l'identification d'un processus autorégressif (Doctoral dissertation).

[12] Nicholls, D.F. and Quinn, B.G. (1980). The estimation of random coefficient autoregressive models I. Journal of time series analysis. 1(1), 37-46.

[13] Nicholls, D.F. and Quinn, B.G (1981). The estimation of random coefficient autoregressive models II. Journal of time series analysis. 2, 185-203.

[14] Nicholls, D.F, and Quinn, B.G. (1982). Random Coefficient Autoregressive Models. An Introduction. In Lecture notes in statistics. Springer.

[15] Nicholls, D. F. (1986). The Box-Jenkins approach to random coefficient autoregressive modelling. Journal of Applied Probability, 23(A), 231-240.

[16] Quinn, B.G. (1980). Fixed and random coefficient Time series. Unpublished Ph.D. thesis, Austalian National University. *https* : *//openresearch* − *repository.anu.edu.au/bitstream*/1885/138725/2/*b*12091650$_Q uinn_B arry_G .pdf$

[17] Proïa, F., & Soltane, M. (2021). Comments on the presence of serial correlation in the random coefficients of an autoregressive process. Statistics & Probability Letters, 170, 108-988.

[18] Regis, M., Serra, P., & van den Heuvel, E. R. (2022). Random autoregressive models: A structured overview. Econometric Reviews, 41(2), 207-230.

[19] Schick, A. (1996). $\sqrt{n}$-Consistent Estimation in a Random Coefficient Autoregressive Model. Austral. J. Statist.. 38(2), 155-160.

[20] Wang, D. (2003) : Frequentist and bayesian analysis of random coefficient autoregressive model

[21] Wang, D., & Ghosh, S. K. (2004). Bayesian analysis of random coefficient autoregressive models. North Carolina State University. Dept. of Statistics.

Appendix A

Appendix

Year;;data

1951;;2,5;

1952;;-1,5;

1953;;0,5;

1954;;1,1;

1955;;-0,9;

1956;;2,2;

1957;;1;

1958;;-3,1;

1959;;-1,5;

1960;;0,2;

1961;;-0,4;

1962;;3,3;

1963;;1,7;

1964;;-0,6;

1965;;-0,7;

1966;;-2,2;

1967;;2,8;

1968;;0,9;

1969;;-2,4;

1970;;-1,9;

1971;;0,6;

1972;;0,8;

1973;;-0,5;

1974;;4;

1975;;-0,8;

1976;;2,3;

1977;;-0,7;

1978;;-0,5;

1979;;-0,7;

1980;;0,7;

1981;;0,6;

1982;;2;

1983;;-5,8;

1984;;0,3;

1985;;-0,5;

1986;;1,6;

1987;;-1,1;

1988;;-0,2;

1989;;2,5;

1990;;-0,1;

1991;;1;

1992;;-4,7;

1993;;-1,5;

1994;;-0,2;

1995;;-0,7;

1996;;1,6;

1997;;0,8;

1998;;-4,4;

1999;;3;

2000;;1,1;

2001;;1,6;

2002;;0,7;

2003;;-0,3;

2004;;-2,2;

2005;;0,6;

2006;;2,7;

2007;;-1,3;

2008;;2,9;

2009;;1,8;

2010;;-1,8;

2011;;3,8;

2012;;1,8;

2013;;-0,1;

2014;;2,4;

2015;;-1,4;

2016;;-3,6;

2017;;0,3;

2018;;1,8;

2019;;-0,1;

2020;;0,3;

2021;;3,2;

2022;;0,8;

ABSTRACT

Abstract

In this thesis, we provide a comparative study where an autoregressive process with a random coefficient of order p, RCAR(p) is compared to an autoregressive process with a deterministic coefficient of order p, AR(p) knowing that RCAR models are obtained by introducing random coefficients to an AR(1). For that and after checking conditions of stationarity for both of them, we have conducted corresponding simulations which afterward and when applied to real data allow us to make a comparison between them in terms of prediction.

Résumé

Dans cette thèse nous proposons une étude comparative où un processus autorégressif à coefficient aléatoire d'ordre p, RCAR(p) est comparé à un processus autorégressif à coefficient déterministe d'ordre p, AR(p) sachant que les modèles RCAR sont obtenus en introduisant des coefficients aléatoires à ceux AR. Pour cela et après vérification des conditions de stationnarité pour l'un et l'autre, nous avons procédé à des simulations correspondantes qui par la suite et lors de l'application à des données réelles nous permettent de faire une comparaison entre elles en termes de prévision.

ملخص

نقدم في هذه الأطروحة دراسة مقارنتية حيث تتم مقارنة بين عملية الانحدار الذاتي ذات معامل عشوائي بعملية الانحدار الذاتي ذات المعامل القطعي مع العلم أنه يتم الحصول على النموذج الأول عن طريق إدخال معاملات عشوائية إلى النموذج الثاني ومن أجل استبيان هذه المقارنة وبعد التحقق من شروط الثبات لكليهما ، قمنا بإجراء عمليات محاكاة بالنسبة للنوعين مما وعند تطبيقها على بيانات واقعية تسمح لنا بإجراء مقارنة بينها من حيث التنبؤ.