Parameter Estimation Techniques for Autoregressive Processes

A Thesis Presented

by

Baret Khamisian

to

The Faculty of Natural and Applied Sciences

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in

Financial Mathematics

Notre Dame University-Louaizé

Zouk Mosbeh, Lebanon

October 9, 2019

Notre Dame University-Louaizé,
Zouk Mosbeh, Lebanon

Department of Mathematics and Statistics

Baret Khamisian

We, the thesis committee for the above candidate for the Master of Science degree, hereby recommend acceptance of this thesis.

John Haddad – Thesis Advisor
Professor, Department of Mathematics

Roger Nakad – First Reader
Professor, Department of Mathematics

Remi Hage – Second Reader
Professor, Department of Mathematics

This thesis is accepted by the Faculty of Natural and Applied Sciences.

Professor George Eid
Dean of the Faculty of Natural and
Applied Sciences

Abstract of the Thesis

Parameter Estimation Techniques for Autoregressive Processes

by

Baret Khamisian

Master of Science

in

Financial Mathematics

Notre Dame University-Louaizé,

Zouk Mosbeh,
Lebanon

2019

The main objective of this work is to find a more straightforward method for estimating the parameters of an equally spaced discrete autoregressive process by using maximum likelihood estimation (MLE) considering it is challenging to obtain the parameters of a nonlinear optimization procedure. The resulting estimated values are tested through simulation and then compared with those obtained using the previous MLE and Yule-Walker estimation. The achieved result yields slightly increased accuracy.

Another problem we tackle is the Yule-Walker estimators for the continuous autoregressive models based on equally spaced discrete-time approximations. Again, these estimators are examined through simulation to demonstrate that the obtained result yields an accurate estimation.

To my family.

Contents

List of Tables

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my mentor, Dr. John Haddad for his patience, motivation, and immense knowledge. His guidance and assistance throughout this program helped me in all times of research and writing of my thesis.

A very special gratitude goes out to the previous chairman Dr. Bassem Ghalayini, who accepted me as a member in his department and provided me with the optimal environment to flourish scientifically and individually.

My sincere acknowledgements also go to both of the members of my thesis committee, Dr. Roger Nakad and Dr. Remi Hage, for their time and constant support.

Finally, I must express my very profound appreciation to my family and friends for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of writing my thesis. This accomplishment would not have been possible without them.

Chapter 0

Introduction

Time is the infinite progress of events. It occurs naturally in many application areas like finance, econometrics, and many more. Some of the numerous components observed over time include stock returns, sales, and price indices. Whether individuals are investing personal money in stocks or developing a forecast to see what change in some variables will affect the future route of others, understanding time is significant to ensure success in business.

Any series of observations measured along time is called time series. We describe a time series as a set of data created by taking a series of measurements in a time sequence. It is valuable in statistics, mathematical finance, and mainly in any domain that requires temporal measurements. To understand the mechanisms of time series a model is to be developed to explain the data in such a way that prediction occurs. Hence understanding time series data is crucial to making better data-informed decisions.

Statistically, we represent time series measurements with a set of random variables $X_1, X_2, \ldots, X_t$ where t is the set of time that the process is mea-

sured. In reality, we can only observe a time series at a finite number of times. However, it is beneficial to allow the number of observations to be infinite. In that case, the underlying sequence of random variables $\{X_t\}$ is called a stochastic process [1]. Yule, who launched the idea of stochasticity in time series, proposed that every time series is the realization of a stochastic process. Therefore to analyze a time series, it is useful to set down a statistical model in the form of a stochastic process. A stochastic process is an underlying process that will be the focus of our theoretical development.

Time series consists of two necessary components - time units and the corresponding value assigned for the given time unit. In a discrete-time series, the values are collected at discrete points of time. On the other hand, a continuous-time series records them continuously through time intervals [2]. For example, the exchange rate between two currencies represents a discrete-time series. However, an earthquake that is recorded continuously through time may depict as a continuous-time series. The observations can be equally spaced, unequally spaced, or have missing data.

The primary purpose of time series modeling is to collect and study past observations and to develop an appropriate model which describes the essential structure of the series. In the 1920s and 1930s, Yule and Walker formulated the concept of some of the standard models in time series; moving average (MA), autoregressive (AR), and autoregressive moving average (ARMA). Moreover, Box and Jenkins originated the building process. These models were then used to generate future observations for the series. In practice, we would certainly not know the values of the parameters of the models. However, we can use the observed data to estimate these unknown quantities. In time series modeling,

the standard parameter estimation methodology includes the Yule-Walker and the maximum likelihood method.

This work is divided into two parts. The central part of it is about the implementation of maximum likelihood estimation through the discrete-time autoregressive process. We first tackle the notion of LU-decomposition for the inverse covariance matrix of the model and then construct a likelihood function to derive the parameters. The latter concerns the parameter estimation of the continuous-time autoregressive process by employing discrete-time approximations to the Yule-Walker equations.

The scope of this work is as follows. In chapter one, we will discuss the properties of stochastic processes, Brownian motion, and the autocorrelation function. Chapter two contains a detailed overview of the linear discrete-time autoregressive process, elaborates the Yule-Walker estimation of parameters regarding the model, and introduces the most crucial part of this work, the maximum likelihood estimation based on Haddad's (1998) inverse covariance matrix. Finally, chapter three describes an approach to the continuous-time autoregressive process concerned with the problem of estimation when taking discrete-time approximations to them.

Chapter 1

Time Series Concepts

In this chapter, we give an overall view of a stochastic process with a discussion of some critical stochastic concepts like stationarity and Brownian motion. We also introduce the notion of autocorrelation, which is crucial in analyzing our time series data. Finally, the last section will include a brief approach to the stochastic differential equations.

1.1 Stochastic Process

A time series could be either modeled using a stochastic process or a deterministic process. A stochastic process can be thought of as evolving in time randomly where a single input leads to different outputs. On the other hand, a deterministic process leads to a single output. However, it is impossible to predict what will occur in the future precisely [3]. Therefore, we model our time series using a stochastic process.

A mathematical expression which describes the probability structure of

the time series is referred to as a stochastic process. A stochastic process, say $\{X_t\}$, is defined as a family of random variables indexed with time t [4]. This process can be depicted as a statistical phenomenon that advances in time. Unquestionably, a continuous-time stochastic process is a stochastic process for which $\{X_t, 0 \leq t < \infty\}$ is an uncountable collection of random variables, as contrasted with a discrete-time process for which $\{X_t, t = 0, 1, 2, \dots\}$ is a finite or countable collection of random variables.

1.2 Stationarity

A crucial class of stochastic processes are those which are stationary. A stochastic process is said to be stationary if all of its statistical properties such as mean, variance, and covariance do not change over time [5]. In other words, the properties of one area of the data are much similar to those of any other area. However, one can distinguish between two different notions of stationarity that have been suggested in time series literature over the years; strictly stationary and weakly stationary.

A stochastic process, $\{X_t\}$, is said to be strictly stationary if its joint probability distribution does not change when shifted in time [2]; i.e., it is the same at time t as at any other time $t + h$.

Note that in practice, it is often useful to define stationarity in a less restricted way than that described above, since strict stationarity is a firm requirement. We, therefore, introduce a much less rigorous property, the notion of weak stationarity, which generally describes the same type of behavior.

A stochastic process, $\{X_t\}$, is said to be weakly stationary if the following

conditions are satisfied:

1. The mean $E[X_t] = \mu$ is independent of time t.

2. $E[X_t^2] < \infty$, which implies the variance $Var[X_t] = \sigma^2$ is finite.

3. $Cov[X_t, X_{t+h}] = E[(X_t - \mu)(X_{t+h} - \mu)] = \gamma_h$, called the autocovariance ACVF at lag h, is independent of time t.

In what follows, the term stationary implies weakly stationary.

Example. [IID noise] The random variables $\{X_t\}$ are IID, i.e., independent and identically distributed, with finite second moment $E[X_t^2] = \sigma^2 < \infty$. Then the first requirement is satisfied, since $E[X_t] = 0$. By the assumed independence,

$$\gamma_h = \begin{cases} \sigma^2 & \text{if } h = 0, \\ 0 & \text{if } h \neq 0, \end{cases}$$

which does not depend on t. Hence the above IID noise is stationary.

Note that mathematical definitions of stationarity will be given in chapters two and three.

1.3 Autocorrelation and Partial Autocorrelation Functions

Unlike regular sampling data, time series takes into consideration the order of occurrence of the observations. Therefore, there is extra information about our sample that we could take advantage of. To accomplish that, we need the

autocorrelation function (ACF) of a stationary process $\{X_t\}$ that measures the correlation between observations separated by various time lags [6].

Suppose we have a stationary stochastic process $\{X_t\}$ with mean μ, variance σ^2, and ACVF γ_h. The ACF is denoted by ρ_h and is given as [7]:

$$\rho_h = \frac{E[(X_t - \mu)(X_{t+h} - \mu)]}{\sqrt{E[(X_t - \mu)^2]E[(X_{t+h} - \mu)^2]}} = \frac{Cov[X_t, X_{t+h}]}{Var[X_t]} = \frac{\gamma_h}{\gamma_0}. \tag{1.3.1}$$

The ACF satisfies:

1. $\rho_0 = 1$
2. $\rho_h = \rho_{-h}$
3. $\mid \rho_h \mid \leq 1$

Investigation in ACF enables us to detect vital conditions in time series data. However, in practice, we do not start with a model but with observed data. This means that the stochastic process governing a time series is unknown and so it is impossible to determine the actual ACF. One of the essential tools we use is the sample ACF of the data which is an estimate of the ACF of $\{X_t\}$ [2]. Sample ACF for both discrete and continuous time series will be introduced in chapters two and three, respectively.

Partial autocorrelation function (PACF) is the correlation between observations at different time lags after removing any linear dependence at shorter lags; i.e., a correlation between X_t and X_{t-h} after removing any linear dependence on $X_1, X_2, \ldots, X_{t-h+1}$. The PACF is denoted by ϕ_{hh} and is defined

as:

$$\phi_{hh} = \frac{\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{h-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{h-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \rho_{h-3} & \dots & \rho_1 & \rho_h \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{h-2} & \rho_{h-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{h-3} & \rho_{h-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \rho_{h-3} & \dots & \rho_1 & 1 \end{bmatrix}}$$

The most crucial role of PACF is identifying the appropriate lags of our autoregressive (AR) model by utilizing the PACF plot [8].

1.4 Brownian Motion

A Brownian motion (or Wiener) is an essential concept in stochastic calculus. It is a fundamental stochastic process which can be exercised in modeling noise, specifically white noise. White noise is a sequence of uncorrelated random variables $\{X_t\}$ each with zero mean and variance σ^2. We will define discrete and continuous white noise in chapters two and three, respectively.

A continuous-time stochastic process $\{W_t, 0 \leq t < \infty\}$ is called a (one-dimensional) Brownian motion if it satisfies the following [9]:

1. $W_0 = 0$.

2. W_t follows a normal distribution with mean 0 and variance $\sigma_W^2 t$; i.e.,

$$W_t \sim N(0, \sigma_W^2 t).$$

3. The process $\{W_t, 0 \leq t < \infty\}$ has stationary increments; i.e., the distribution of $W_{t+s} - W_t$ is independent of t for all $s, t > 0$.

4. The process $\{W_t, 0 \leq t < \infty\}$ has independent increments; i.e., for each $0 \leq t_1 < t_2 < \cdots < t_n$, the variables W_{t_1}, $W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent.

Brownian motion is thus a continuous stochastic process which is normally distributed, and whose variance increases the farther it gets from the origin. It is independent of its remote history, and a change in its location at any increment in time is independent of a change anywhere else over the same time increment.

Note that a continuous-time stochastic process $\{W_t, 0 \leq t < \infty\}$ is said to be a standard Brownian motion if $\sigma_W^2 = 1$, i.e. $W_t \sim N(0, t)$.

1.5 Stochastic Differential Equation

A general form of a stochastic differential equation is defined by

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t \quad \text{for } t \in [0, T], \tag{1.5.1}$$

where dX_t represents the derivative of process $\{X_t\}$, a and b are usually time-varying functions, and W_t is a Brownian process [10]. Then if a solution to (1.5.1) exists, we write

$$X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dWs. \tag{1.5.2}$$

Example. An Ornstein-Uhlenbeck process $\{X_t\}$ is a stochastic differential equation of the form [11]

$$dX_t = -aX_t dt + \sigma dW_t. \tag{1.5.3}$$

Note that the theoretical ACF of this process $\{X_t\}$ is given by $\rho_h = e^{-ah}$ for $h \geq 0$.

Chapter 2

Discrete Autoregressive Process

The determination of a suitable model in time series is critical as it captures the underlying data for the series. Models for time series data can represent different stochastic processes. For instance, one of the most popular used linear stochastic time series models is the autoregressive (AR) model. The AR model depicts how an observation directly relies on p past observations with addition to a white noise term. This form of a time series model is appealing and has been broadly applied to data sets in a wide range of fields.

This chapter deals with the introduction of linear regression, which led to the critical discrete autoregressive (DAR) model. The properties of these models are then explored, including the sample autocorrelation function. Additionally, the Yule-Walker equations are formulated. These will be fundamental in developing estimates of the model parameters. The final section deals with a new approach to the maximum likelihood estimation. Conclusively, the parameter estimates are illustrated through a simulated example. This example is analyzed in detail using R and Maple programs.

2.1 Discrete Autoregressive Process of Order 1

Let $\{X_t, t = 1, 2, \dots\}$ be a discrete-time stochastic process. We say that X_t has a linear regression on X_{t-1} if it satisfies the following equation

$$X_t = \phi X_{t-1} + \upsilon_t,$$

with υ_t being the error term. The fact that X_t has a regression on its past gives rise to the first order discrete autoregressive model, denoted DAR(1), which is written as

$$X_t - \phi X_{t-1} = \upsilon_t, \tag{2.1.1}$$

where ϕ is a constant parameter and $\{\upsilon_t\}$ is a discrete white noise process defined as a sequence of independent and identically distributed normal random variable with mean 0 and variance σ_υ^2. Without loss of generality, σ_υ^2 is assumed to equal one.

To analyze the condition of stationarity for equation (2.1.1), we introduce the backshift/lag operator L that is $LX_t = X_{t-1}$ [12]. Hence, we can rewrite the difference equation (2.1.1) as

$$(1 - \phi L)X_t = \upsilon_t. \tag{2.1.2}$$

For the process $\{X_t\}$ to be stationary, all the roots of the homogeneous equation should lie outside the unit circle; i.e., are larger than one in absolute value.

Therefore the solution of the homogeneous equation $X_t - \phi X_{t-1} = 0$, which can be written as (2.1.2), is $L = \dfrac{1}{\phi}$. Thus $\{X_t\}$ to be stationary, $|\phi| < 1$.

Let us assume $\{X_t, t = 1, 2, \dots\}$ to be a stationary DAR(1) process. The covariance matrix, $\Sigma_1 = \gamma_{(|k-j|)}$ for $j, k = 1, \dots, n$, can be computed by multiplying both sides of equation (2.1.1) by X_{t-h} and taking expectations

$$\gamma_h - \phi\gamma_{h-1} = \begin{cases} 1 & \text{if } h = 0, \\ 0 & \text{if } h > 0, \end{cases}$$

where $\gamma_h = E[X_t X_{t-h}]$ is the ACVF at lag h. Thus by solving the above system of equations, it follows the covariance matrix of a DAR(1) process to be

$$\Sigma_1 = \frac{1}{1-\phi^2} \begin{bmatrix} 1 & \phi & \dots & \phi^{n-1} \\ \phi & 1 & \dots & \phi^{n-2} \\ \vdots & & \vdots & \vdots \\ \phi^{n-1} & \phi^{n-2} & \dots & 1 \end{bmatrix}.$$

Consequently, we can obtain the ACF of a DAR(1) process by utilizing equation (1.3.1).

2.2 Discrete Autoregressive Process of Order P

The discrete-time stochastic process, $\{X_t\}$, that satisfies the difference equation [5]

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = \upsilon_t \quad \text{for t = p, p+1, p+2, } \ldots \tag{2.2.1}$$

is called a zero mean discrete-time autoregressive process of order p , a DAR(p) process, where $\{\upsilon_t\}$ is a Gaussian discrete white noise process; i.e., the process is normally distributed and $\phi_1, \phi_2, \ldots, \phi_p$ are model parameters.

Note that the above equation can also be written as

$$(1 - \phi_1 L - \cdots - \phi_p L^p) X_t = \upsilon_t, \tag{2.2.2}$$

where L is the backshift operator. In fact setting z = 1/L, provides us the characteristic equation

$$z^p - \phi_1 z^{p-1} - \cdots - \phi_p = 0. \tag{2.2.3}$$

Denote by $R = (r_1, \ldots, r_p)'$ the roots (may be real or complex) of the characteristic equation. Thus, as given in Box and Jenkins [8], the process in (2.2.1) can be written as

$$(1 - r_1 L) \ldots (1 - r_p L) X_t = \upsilon_t. \tag{2.2.4}$$

Following the same argument as DAR(1) process, a discrete-time stochastic process $\{X_t\}$, that satisfies (2.2.1), is stationary under the condition that the roots of its homogeneous equation lie outside the unit circle, or each root of the characteristic equation (2.2.3) is less than one in absolute value.

Haddad [13] showed that the inverse covariance matrix of a DAR(p) model is determined in terms of the roots of its characteristic equation

$$\Sigma_p^{-1} = \Omega^{-1}(r_1)\Delta_1^{-1} \dots \Delta_{p-1}^{-1}\Omega^{-1}(r_p), \tag{2.2.5}$$

where $\Omega^{-1}(r_j)$ is the inverse covariance matrix of a DAR(1) model with parameter r_j, and Δ_{p-1}^{-1} is a near identity matrix.

The subsequent step is to find the LDL' decomposition of each inverse covariance matrix $\Omega^{-1}(r_j)$ and rewrite equation (2.2.5) as a function of unit lower triangular matrices

$$\Sigma_p^{-1} = L(r_1)D(r_1)L'(r_1)\Delta_1^{-1} \dots \Delta_{p-1}^{-1}L(r_p)D(r_p)L'(r_p).$$

Following some simple matrix manipulations, we can easily obtain

$$\Sigma_p^{-1} = [L(r_1)L(r_2) \dots L(r_p)][L'(r_1)L'(r_2) \dots L'(r_p)] + \Psi_p,$$

where Ψ_p is a near null matrix.

In what follows,

$$\Sigma_p^{-1} = [L(r_1)L(r_2) \dots L(r_p)][L'(r_1)L'(r_2) \dots L'(r_p)]. \tag{2.2.6}$$

2.3 Sample Autocorrelation Function

While working on practical problems, we do not start with a model, but with observed data $X_1, X_2, \ldots, X_n$. Hence we would certainly not know the values of the model parameters, and consequently, we would not be able to determine the theoretical ACF. Nevertheless, from the observed values, we will estimate the ACVF and the ACF of the underlying process $\{X_t\}$. This is a crucial step to construct an adequate model for the data.

The estimate of the theoretical ACVF γ_h is the sample ACVF $\hat{\gamma}_h$, which is given by [5]

$$\hat{\gamma}_h = \frac{1}{n}\sum_{t=1}^{n-h}(X_t - \bar{X})(X_{t+h} - \bar{X}),$$

where $\bar{X}$ is the estimate of the mean μ and is defined by $\bar{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$. Having estimated the ACVF, we then can easily calculate the ACF at lag h by using (1.3.1), $\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0}$.

2.4 Yule-Walker Equations for Discrete Autoregressive Processes

Throughout this section, we assume we have n observations, $X_1, X_2, \ldots, X_n$, from a stationary DAR(p) process, in which the order of the model, p, is known. We will not discuss the problem of determining p. Our goal in this section is to estimate the parameters $\phi_1, \phi_2, \ldots, \phi_p$ of a stationary DAR(p) process by using the Yule-Walker equations.

Let $\{X_t\}$ be a zero mean stationary DAR(p) process defined by (2.2.1)

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = \upsilon_t \quad \text{for t= p, p+1, p+2, } \ldots.$$

Multiplying both sides by X_{t-h} and taking expectations, we obtain the Yule-Walker equations [14]

$$\gamma_h - \phi_1 \gamma_{h-1} - \cdots - \phi_p \gamma_{h-p} = \begin{cases} 1 & \text{if } h = 0, \\ 0 & \text{if } h > 0, \end{cases} \tag{2.4.1}$$

where $\gamma_h = E[X_t X_{t-h}]$.

The Yule-Walker estimators are then obtained by substituting γ_h in equations (2.4.1) by $\hat{\gamma}_h$, due to the fact the sample ACVF is the estimate of the theoretical ACVF. Thus, the parameter estimates $\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_p$ are acquired by solving the below Yule-Walker estimators

$$\hat{\gamma}_h - \hat{\phi}_1 \hat{\gamma}_{h-1} - \cdots - \hat{\phi}_p \hat{\gamma}_{h-p} = \begin{cases} 1 & \text{if } h = 0, \\ 0 & \text{if } h > 0. \end{cases} \tag{2.4.2}$$

Example. Consider a zero mean stationary DAR(2) process defined by

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = \upsilon_t \quad \text{for t= 2, 3, } \ldots.$$

Let us generate 50 observations in R from a simulated DAR(2) model with parameters $\phi_1 = -0.2$ and $\phi_2 = 0.35$. We utilize R to obtain the Yule-Walker estimators.

Here is the output summarizing the fit. Appendix A.1 contains detailed information regarding the codes in R.

```
Coefficients:
      1         2
-0.1721    0.3164
sigma^2 estimated as  1.285
```

The estimators are $\hat{\phi}_1 = -0.1721$ and $\hat{\phi}_2 = 0.3164$ which provides the fitted DAR(2) model

$$X_t + 0.1721X_{t-1} - 0.3164X_{t-2} = \upsilon_t.$$

The noise variance estimate is $\hat{\sigma}_\nu^2 = 1.285$.

2.5 Maximum Likelihood Estimation

Another method utilized to estimate the parameters of a zero mean DAR(p) model is the maximum likelihood estimator (MLE). This procedure finds the values of the parameters which maximize the probability of obtaining the data that we have observed.

Suppose we have a zero mean stationary DAR(p) process given by (2.2.1). For independent and identically distributed random variables $X_1, X_2, \ldots, X_n$, the likelihood function is defined as the probability density function of $X = (X_1, X_2, \ldots, X_n)'$ under the multivariate Gaussian model [15]

$$L(\phi_1, \phi_2, \ldots, \phi_p; X) = \prod_{i=1}^{n} P(X_i; \phi_1, \phi_2, \ldots, \phi_p) \tag{2.5.1}$$

$$= \frac{1}{(2\pi)^{n/2}|\Sigma_p|^{1/2}} exp\{-\frac{1}{2}X'\Sigma_p^{-1}X\}, \tag{2.5.2}$$

where Σ_p is the covariance matrix of X with the given parameter values, and $|\Sigma_p|$ denotes the determinant of Σ_p.

The goal is then to find the values of the model parameter that maximize the likelihood function. However, rather than maximizing this product (2.5.2), which can be tedious, it is often convenient to work with the natural logarithm of the likelihood function, called the log-likelihood

$$\begin{aligned}\ell(\phi_1, \phi_2, \ldots, \phi_p; X) &= \log L(\phi_1, \phi_2, \ldots, \phi_p; X) \\ &= -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma_p| - \frac{1}{2}\{X'\Sigma_p^{-1}X\}.\end{aligned}$$

We know from linear algebra that $|\Sigma_p| = \dfrac{1}{|\Sigma_p^{-1}|}$, then the log-likelihood function becomes

$$\ell(\phi_1, \phi_2, \ldots, \phi_p; X) = -\frac{n}{2}\log(2\pi) + \frac{1}{2}\log|\Sigma_p^{-1}| - \frac{1}{2}\{X'\Sigma_p^{-1}X\}. \qquad (2.5.3)$$

Now, maximizing $\ell(\phi_1, \phi_2, \ldots, \phi_p; X)$ with respect to ϕ_j; i.e., taking the partial derivative of the log-likelihood function with respect to ϕ_j, and setting to 0, will give us the MLE.

Let us examine a new approach to MLE. Consider again a zero mean stationary DAR(p) process, $\{X_t\}$, with n observations to be defined as (2.2.4)

$$(1 - r_1 L)\ldots(1 - r_p L)X_t = \upsilon_t.$$

To find the maximum likelihood estimates, we maximize the log-likelihood

function (2.5.3)

$$\ell(r_1, r_2, \ldots, r_p; X) = -\frac{n}{2}\log(2\pi) + \frac{1}{2}\log|\Sigma_p^{-1}| - \frac{1}{2}\{X'\Sigma_p^{-1}X\}, \qquad (2.5.4)$$

with respect to r_j where each r_j is the root of the characteristic equation (2.2.3), and $\Sigma_p^{-1} = [L(r_1)L(r_2)\ldots L(r_p)][L'(r_1)L'(r_2)\ldots L'(r_p)]$ is considered as the inverse covariance matrix of $X = (X_1, X_2, \ldots, X_n)'$ presented in (2.2.6). Thus, the estimated roots can be computed by solving the system of equations

$$\frac{\partial\{X'\Sigma_p^{-1}X\}}{\partial r_j} = 0, \qquad (2.5.5)$$

considering $|\Sigma_p^{-1}| = 1$.

Example. We again revisit the simulated DAR(2) process with $\phi_1 = -0.2$ and $\phi_2 = 0.35$ in section (2.4), and use R to refit the model

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} = \upsilon_t,$$

except now using the maximum likelihood estimator (MLE) method. Here is the output from R summarizing the MLE fit. Appendix A.1 includes information regarding the estimation in R.

```
Coefficients:
      1         2
-0.1707    0.3117
sigma^2 estimated as  1.204
```

The estimators are $\hat{\phi}_1 = -0.1707$ and $\hat{\phi}_2 = 0.3117$ which yields the fitted

DAR(2) model

$$X_t + 0.1707X_{t-1} - 0.3117X_{t-2} = \upsilon_t.$$

The noise variance estimate is $\hat{\sigma}_\nu^2 = 1.204$.

However, by using our new strategy, we know that a DAR(2) model can also be represented as

$$(1 - r_1 L)(1 - r_2 L)X_t = \upsilon_t.$$

Expanding the left-hand side of the above equation and comparing it to a DAR(2) model defined by (2.2.1), we can easily obtain $\phi_1 = r_1 + r_2$ and $\phi_2 = -r_1 r_2$.

In order to estimate the coefficients r_1 and r_2, a mathematically based software known as Maple is used to solve the system of equations presented in (2.5.5)

$$\frac{\partial\{X'\Sigma_p^{-1}X\}}{\partial r_j} = 0,$$

where $\Sigma_p^{-1} = [L(r_1)L(r_2)][L'(r_1)L'(r_2)]$ is the inverse covariance matrix of $X = (X_1, X_2, \ldots, X_n)'$. Appendix A.2 includes the complete code regarding the solution of the system of equations in Maple.

```
Coefficients:
     r1        r2
-0.6603    0.4868
```

The calculated estimates are found to be $\hat{\phi}_1 = r_1 + r_2 = -0.1735$ and $\hat{\phi}_2 =$

$-r_1 r_2 = 0.3214$, which produces the fitted DAR(2) model

$$X_t + 0.1735X_{t-1} - 0.3214X_{t-2} = \upsilon_t.$$

2.6 Conclusion

For the simulated zero mean DAR(2) process with parameters $\phi_1 = -0.2$ and $\phi_2 = 0.35$, we represent the estimates from all three methods in the following table.

Table 2.1: Estimated parameters $\hat{\phi}_1$ and $\hat{\phi}_2$ of a DAR(2) process based on three methods.

Method	$\hat{\phi}_1$	$\hat{\phi}_2$
Yule-Walker estimation	-0.1721	0.3164
MLE	-0.1707	0.3117
MLE with $\Sigma_p^{-1} = [L(r_1)L(r_2)][L'(r_1)L'(r_2)]$	-0.1735	0.3214

We analyze Table 2.1 and observe that the three methods are somewhat close. Nevertheless, comparing it to the actual parameter values $\phi_1 = -0.2$ and $\phi_2 = 0.35$ provides us the outcome that using MLE with inverse covariance matrix $\Sigma_p^{-1} = [L(r_1)L(r_2)][L'(r_1)L'(r_2)]$ slightly increases the accuracy of the estimation.

Chapter 3

Continuous Autoregressive Process

Although we defined the discrete autoregressive process in chapter two, many economic and financial problems have continuously observed data. As a result, continuous time series is introduced. Peculiarly, the continuous autoregressive CAR process.

In this chapter, we first discuss the concept of a continuous-time white noise process. We also develop the notion of a CAR process in parallel to the well known discrete-time autoregressive process addressed in chapter two. Further, we introduce the derivative covariance function, which will assist us in establishing the Yule-Walker equations. Finally, estimation of model parameters using the discrete-time approximations of the Yule-Walker equations is considered. In conclusion, the parameter estimates, similar to the previous chapter, are illustrated through a simulated CAR process.

3.1 Continuous-Time White Noise

A continuous-time stochastic process $\{\upsilon_t, 0 \leq t < \infty\}$ is called a continuous-time white noise if it satisfies the following properties:

1. $E[\upsilon_t] = 0$.

2. There exists a constant variance $\sigma_\upsilon^2 \in [0, \infty)$ satisfying

$$Cov[\upsilon_s, \upsilon_t] = \sigma_\upsilon^2 \delta(t - s),$$

where δ represents a Dirac delta function [16].

This particular form does not prove too worrisome since we are going to work with continuous parameter models, ergo we will usually be dealing with integrals of $\{\upsilon_t\}$ rather than $\{\upsilon_t\}$ itself. Note that throughout this chapter, we assume a continuous-time white noise process to be integrable with respect to time.

Suppose that $\{\upsilon_t, 0 \leq t < \infty\}$ is a Gaussian continuous-time white noise process; i.e., $\int_0^t \upsilon_u du$ is normally distributed. If the continuous-time stochastic process $\{W_t, 0 \leq t < \infty\}$ is defined by

$$W_t = \int_0^t \upsilon_u du, \tag{3.1.1}$$

then $\{W_t\}$ is a Brownian motion. Proof of which can be found in Appendix B.

3.2 Continuous Autoregressive Process of Order 1

A continuous-time autoregressive process of order 1, denoted CAR(1), is defined along similar lines as a DAR(1) process, with the most noticeable change being the replacement of the difference equation by a differential equation.

Let $\{X_t, 0 \leq t < \infty\}$ be a continuous-time stochastic process. We say that $\{X_t\}$ is a CAR(1) process if it satisfies the stochastic differential equation

$$X_t^{(1)} + \alpha_0 X_t = \sigma \upsilon_t, \tag{3.2.1}$$

where $X_t^{(1)} = (d/dt)X_t$ is the first derivative of the process $\{X_t\}$, α_0 is the model parameter, and $\{\upsilon_t\}$ is a continuous white noise process.

Following the same argument as used for the discrete model, a continuous-time stochastic process $\{X_t\}$ that satisfies (3.2.1) is said to be stationary if the solution of the homogeneous equation decays to zero as $t \to \infty$. Therefore the solution of the homogeneous equation $X_t^{(1)} + \alpha_0 X_t = 0$, which is of the form $X_t = ce^{-\alpha_0 t}$ where c is an arbitrary constant, decays to zero when $\alpha_0 > 0$.

The solution of (3.2.1) is the Itô integral

$$X_t = X_0 e^{-\alpha_0 t} + \sigma \int_0^t e^{-\alpha_0 (t-u)} dW_u, \tag{3.2.2}$$

where X_0 is an arbitrary constant, and through (3.1.1), $dW_t = \upsilon_t dt$ is the derivative of Brownian motion. Priestley [6] demonstrated that $\gamma_h = \dfrac{\sigma^2}{2\alpha_0} e^{-\alpha_0 h}$ presents the ACVF of the process $\{X_t\}$. Consequently, we can obtain the ACF

by applying equation (1.3.1).

We can see that a CAR(1) model is a representation of the Ornstein-Uhlenbeck (OU) process illustrated in (1.5.3) with ACF $\rho_h = e^{-\alpha_0 h}$, which is congruous with Priestley's result. Furthermore, to estimate the parameter of this OU-process, we discretize it to a DAR(1) process as following.

Let $\{x_t\}$ denote a sample function of a zero mean stationary CAR(1) process that satisfies (3.2.1), and suppose the data consists of observations taken at times $0 = t_0 < t_1 < \cdots < t_n = T$ and are equally spaced; i.e., $\Delta_i = t_{i+1} - t_i = \Delta$ for all i. Formulating $X_t = x_{(t\Delta)}$ into the Itô integral

$$x_t = x_0 e^{-\alpha_0 t} + \sigma \int_0^t e^{-\alpha_0 (t-u)} dw_u,$$

the sampled process can be described by the DAR(1) process $X_t = \phi X_{t-1} + \upsilon_t$ where $X_0 = x_0$, and $\phi = e^{-\alpha_0 \Delta}$. Considering $\phi = e^{-\alpha_0 \Delta}$, the parameter estimate of the above CAR(1) model is defined by

$$\hat{\alpha}_0 = -\frac{1}{\Delta} log(\hat{\phi}), \tag{3.2.3}$$

where $\hat{\phi} = \dfrac{\hat{\gamma}_1}{\hat{\gamma}_0}$ represents the Yule-Walker estimate (2.4.2) of the DAR(1) process.

Example. Consider a zero mean stationary CAR(1) process defined by

$$dX_t = -\alpha_0 X_t dt + \sigma dW_t.$$

Let us simulate n=1000 observations throughout $[0, T = 10]$ of an OU-process, having parameter $\alpha_0 = 0.3$ and volatility $\sigma = 0.1$. The generated observations

are discretized with $\Delta = \dfrac{T - t_0}{n}$ signifying the time step of the simulation. Appendix A.3 carries detailed information regarding the simulation in R.

To acquire the parameter estimate of the above CAR(1) model, we solve (3.2.3) on Maple software to obtain $\hat{\alpha}_0 = 0.3216$. We can see that this estimation method is accurate as it is close to the actual parameter value of $\alpha_0 = 0.3$. Check Appendix A.4 regarding the solution of the above technique in Maple.

Note that to obtain the ACVF at lag 0, we use the ACF of the OU-process and Priestley's computed ACVF, and employ them into equation (1.3.1) to get $\gamma_0 = \dfrac{\sigma^2}{2\alpha_0}$. Hence to get the estimated sigma, solve $\hat{\sigma} = \sqrt{2\hat{\alpha}_0\hat{\gamma}_0}$.

3.3 Continuous Autoregressive Process of Order P

The continuous-time stochastic process, $\{X_t, 0 \leq t < \infty\}$, that satisfies the differential equation

$$\alpha_0 X_t + \cdots + \alpha_{p-1} X_t^{(p-1)} + X_t^{(p)} = \sigma \upsilon_t \tag{3.3.1}$$

is called a zero mean continuous-time autoregressive process of order p, a CAR(p) process, where $\alpha_0, \alpha_1, \ldots, \alpha_{p-1}$ represent model parameters, $X_t^{(j)}$ denotes the jth derivative of X_t, and $\{\upsilon_t\}$ is a Gaussian continuous-time white noise process. However, in the light of (3.1.1), we refer a Gaussian continuous-time white noise process $\{\upsilon_t, 0 \leq t < \infty\}$ as $\{dW_t, 0 \leq t < \infty\}$ where $\{W_t\}$ denotes the Brownian motion. Thus, we introduce a more precise definition of a CAR(p) process.

Suppose that A is a pxp matrix and b is a $px1$ vector, where $p \in \mathbb{N} = \{1, 2, \dots\}$, both defined by

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & \dots & \dots & -\alpha_{p-1} \end{bmatrix} \qquad b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

A CAR(p) process represented by $C_t = [X_t, X_t^{(1)}, \dots, X_t^{(p-1)}]'$, should satisfy the stochastic differential equation

$$dC_t = AC_t dt + \sigma b dW_t, \tag{3.3.2}$$

where $\{W_t\}$ denotes standard Brownian motion.

A vital approach to study stationarity of a continuous-time stochastic process $\{C_t, 0 \leq t < \infty\}$ that satisfies (3.3.2) includes the eigenvalues of A to have negative real parts, which indicates $\lim_{t\to\infty} e^{At} = 0$. The proof of which is found in [17].

Note that Brownian motion is not differentiable [18], indicating dW_t does not exist. In fact, (3.3.2) can be written as an Itô integral equation given as (1.5.2)

$$C_t = C_0 + \int_0^t AC_u du + \sigma \int_0^t b dW_u. \tag{3.3.3}$$

By utilizing equation (3.3.3) and employing some mathematical techniques,

we represent the solution of (3.3.2) to be

$$C_t = e^{At}C_0 + \sigma \int_0^t e^{A(t-u)}bdW_u, \tag{3.3.4}$$

where e^A is the matrix exponential for any square matrix A, and the process $C_0 \sim N\left(0, \sigma^2 \int_0^\infty e^{As}bb'e^{A's}ds\right)$.

3.4 Derivative Covariance Function

For a continuous-time autoregressive process CAR(p), the Yule-Walker equations are written in terms of the derivative covariance function (DCVF) defined below.

Let $\{X_t\}$ be a CAR(p) process. Hyndman, within [19], showed that for $h \geq 0$ and $0 \leq j, k \leq p-1$, there exist a relation between DCVF and ACVF.

$$D_{j,k}(h) = Cov[X_{t+h}^{(j)}, X_t^{(k)}] \tag{3.4.1}$$

$$= (-1)^k \gamma^{(j+k)}(h), \tag{3.4.2}$$

where $\gamma^{(j+k)}(h)$ denotes the $(j+k)^{th}$ derivative of the ACVF. The proof of which can be found in [10]. Additionally, we note that $D_{j,k}(0) = 0$ if $j+k$ is odd. This postulate follows using $D_{j,k}(h) = D_{k,j}(-h)$ when h=0.

Similar to the discrete case, it is impossible to determine the actual DCVF considering we do not start with a model, but with observed data. Accordingly, we need to consider the problem of estimating the DCVF $D_{j,k}(0)$ based on the continuously observed data.

Let $\{x_t\}$ denote the observed data from a stationary CAR(p) process $\{X_t\}$. The estimate of the theoretical DCVF $D_{j,k}(0)$, for $0 \leq j \leq p-1$ and $0 \leq k \leq p$, is the sample DCVF $\hat{D}_{j,k}$ which is given by

$$\hat{D}_{j,k} = \begin{cases} \frac{1}{T}\int_0^T x_t^{(j)} x_t^{(k)} dt & \text{if } k < p, \\ \frac{1}{T}\int_0^T x_t^{(j)} dx_t^{(p-1)} & \text{if } k = p. \end{cases} \tag{3.4.3}$$

3.5 Yule-Walker Equations for Continuous Autoregressive Processes

Our goal in this section is to estimate the parameters $\alpha_0, \alpha_1, \ldots, \alpha_{p-1}$ of a stationary CAR(p) process by using the Yule-Walker equations. To keep the close analogue between the continuous and discrete case, we will not discuss the problem of determining p. For related work in the case of equal and closely spaced data, see Philips [20] and Brockwell [21].

Let $\{X_t\}$ be a zero mean stationary CAR(p) process given by (3.3.1)

$$\alpha_0 X_t + \cdots + \alpha_{p-1} X_t^{(p-1)} + X_t^{(p)} = \sigma \upsilon_t.$$

The Yule-Walker equations for CAR(p) processes are similar to the discrete-time Yule-Walker equations. We simply multiply both sides of the above equation by $X_{t+h}^{(j)}$ and take expectations to obtain

$$\alpha_0 D_{j,0}(h) + \cdots + \alpha_{p-1} D_{j,p-1}(h) + D_{j,p}(h) = 0 \quad 0 \leq j \leq p-1, \tag{3.5.1}$$

where $D_{j,k}(h) = E[X_{t+h}^{(j)} X_t^{(k)}]$ due to (3.4.1).

Consider equation (3.5.1) and let $h = 0$. The Yule-Walker estimators are then obtained by substituting $D_{j,k}$ and $D_{j,p}$ by $\hat{D}_{j,k}$ and $\hat{D}_{j,p}$, respectively, in view of the fact that the sample DCVF is the estimate of the theoretical DCVF

$$\hat{\alpha}_0 \hat{D}_{j,0} + \cdots + \hat{\alpha}_{p-1} \hat{D}_{j,p-1} + \hat{D}_{j,p} = 0 \quad 0 \leq j \leq p-1, \tag{3.5.2}$$

where $\hat{\alpha}_j$ is the Yule-Walker estimator of α_j. Note that to obtain the estimated volatility $\hat{\sigma}$, we substitute $(-1)^k \gamma^{(j+k)}(h)$ for $D_{j,k}(h)$, as a result of (3.4.2), into (3.5.2) and solve the following equation

$$\hat{\alpha}_0 \hat{\gamma}^{(j)}(0) + \cdots + (-1)^{p-1} \hat{\alpha}_{p-1} \hat{\gamma}^{(j+p-1)}(0) + (-1)^p \hat{\gamma}^{(j+p)}(0) = 0, \tag{3.5.3}$$

where $\hat{D}_{p-1,p}(0) = -\dfrac{\hat{\sigma}^2}{2}$ according to Hyndman.

Numerous real-life problems have continuously observed data. However, in practice, measurements are listed at discrete times. Hence to utilize the above Yule-Walker estimators (3.5.2), it is necessary to derive discrete-time approximations to them.

Let $\{x_t\}$ denote a sample function of a zero mean stationary CAR(p) process that satisfies (3.3.1), and suppose we are able to observe it discretely taken at times $0 = t_0 < t_1 < \cdots < t_n = T$ and let $\Delta_i = t_{i+1} - t_i$ for all i. Consider the recorded data to be equally spaced; i.e., $\Delta_i = \Delta = (T - t_0)/n$. For closely-spaced discrete observations, Hyndman introduced the notion of discrete-time approximations to the continuous-time estimators (3.4.3) where

the integrals are replaced by approximating sums

$$\check{D}_{j,k} = \begin{cases} \dfrac{\Delta}{T}\sum_{i=1}^{n} \hat{x}_{t_i}^{(j)}\hat{x}_{t_i}^{(k)} & \text{if } k < p, \\ \dfrac{1}{T}\sum_{i=1}^{n} \hat{x}_{t_i}^{(j)}\{\hat{x}_{t_{i+1}}^{(p-1)} - \hat{x}_{t_i}^{(p-1)}\} & \text{if } k = p, \end{cases} \tag{3.5.4}$$

where $\hat{x}_{t_i}^{(j)}$ is the estimate of $x_{t_i}^{(j)}$ for instance, $\hat{x}_{t_i}^{(1)} = \{x_{t_{i+1}} - x_{t_i}\}/\Delta$.

Thus we define a discrete form of the Yule-Walker estimators (3.5.2) by replacing $\hat{D}_{j,k}$ and $\hat{D}_{j,p}$ by $\check{D}_{j,k}$ and $\check{D}_{j,p}$, respectively.

$$\check{\alpha}_0\check{D}_{j,0} + \cdots + \check{\alpha}_{p-1}\check{D}_{j,p-1} + \check{D}_{j,p} = 0 \quad 0 \le j \le p-1, \tag{3.5.5}$$

where $\check{\alpha}_j$ denotes the estimator which is obtained by solving the above Yule-Walker equation.

Consider, as an illustration, p=1. The estimate for the coefficient α_0 in the CAR(1) model defined by (3.3.1) is (replace $p = 1$ and take $j = 0$ in (3.5.2))

$$\begin{aligned} \hat{\alpha}_0 &= -\frac{\hat{D}_{0,1}}{\hat{D}_{0,0}} \\ &= -\frac{\int_0^T x_t dx_t}{\int_0^T \{x_t\}^2 dt}, \end{aligned}$$

which is congruous with Priestley's result [6]. Nevertheless, observations are recorded at discrete times. Consequently, we use the Yule-Walker equations of the discrete form (3.5.5) and represent the estimate to be

$$\check{\alpha}_0 = -\frac{\sum_{i=1}^{n} \hat{x}_{t_i}\{\hat{x}_{t_{i+1}} - \hat{x}_{t_i}\}}{\Delta\sum_{i=1}^{n}\{\hat{x}_{t_i}\}^2}.$$

Furthermore, Hyndman pointed out that by implementing some simple mathematical manipulations, the above estimate yields to

$$\check{\alpha}_0 = -\frac{1}{\Delta}\big(\frac{\hat{\gamma}_\Delta}{\hat{\gamma}_0} - 1\big), \tag{3.5.6}$$

where $\hat{\gamma}_\Delta = (1/n)\sum_{i=1}^{n-1} x_{t_i} x_{t_{i+1}}$ and $\hat{\gamma}_0$ is the sample ACVF defined in section (2.3).

Example. We revisit the simulated CAR(1) process, which was an OU-process, with $\alpha_0 = 0.3$ and $\sigma = 0.1$ in section (3.2), to refit the model

$$dx_t = -\alpha_0 x_t dt + \sigma dw_t,$$

through estimating the coefficient α_0, except now using Hyndman's Yule-Walker method (3.5.6), including the volatility σ.

Employing equation (3.5.6) in Maple software, the coefficient estimate $\check{\alpha}_0$ is found to be 0.3211. Appendix A.4 covers the code regarding the solution of the CAR(1) parameter estimate. Moreover, solving equation (3.5.3) with p=1 and j=0 provides the volatility estimate $\hat{\sigma}$, which is congruent with Priestley's result. Thus the produced fitted CAR(1) model is defined by

$$dx_t = -0.3211 x_t dt + 0.1301 dw_t.$$

3.6 Conclusion

For the simulated zero mean Ornstein-Uhlenbeck process with parameter $\alpha_0 = 0.3$ and volatility $\sigma = 0.1$, we represent the estimates from both methods in the following table.

Table 3.1: Estimated parameter and volatility of a CAR(1) process based on two methods.

Method	Estimated α_0	Estimated σ
Discretizing to a DAR(1) process	0.3216	0.1302
Hyndman's discrete Yule-Walker estimation	0.3211	0.1301

We analyze Table 3.1 and observe that the two estimates are remarkably close. Nonetheless, comparing it to the actual parameter value $\alpha_0 = 0.3$ and volatility $\sigma = 0.1$ yields us with the outcome that using Hyndman's discrete form of the Yule-Walker estimators slightly improves the accuracy of the estimation.

Conclusion and Future Work

The determination of an adequate autoregressive model to represent an observed stationary time series involves estimation of the unknown parameters; i.e., the coefficients and the white noise variance. In this work, we discussed the problem of different varieties of estimation through discrete and continuous autoregressive models.

A new approach based on the maximum likelihood estimation of the model was introduced. One of the main objectives of our work was to express the inverse covariance matrix with unit lower triangular matrices and to estimate the parameters of our discrete autoregressive model by employing the maximum likelihood estimation technique. A simulation of the process was performed to examine different types of estimation. Results show that our approach obtains a better outcome. It casts a new light on estimation methodology.

Another problem we tackled relies on the parameter estimation of a continuous autoregressive process through Yule-Walker equations. We drive discrete-time approximations to the Yule-Walker estimators as observations are recorded at discrete times. Once more, a simulation of our model was conducted to test different types of estimation, including the method described. Results showed reliable outcomes.

The work presented in this thesis produces a foundation for future research in several fields. For example, further research could investigate the CAR process implemented to the MLE technique, with its inverse covariance matrix formulated as unit lower triangular matrices, similar to chapter two.

Bibliography

[1] Andrew H Jazwinski. *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[2] Peter J Brockwell, Richard A Davis, and Matthew V Calder. *Introduction to time series and forecasting*, volume 2. Springer, 2002.

[3] Keith W Hipel and A Ian McLeod. *Time series modelling of water resources and environmental systems*, volume 45. Elsevier, 1994.

[4] Mark Pinsky and Samuel Karlin. *An introduction to stochastic modeling*. Academic press, 2010.

[5] Chris Chatfield. *The analysis of time series: an introduction*. CRC press, 2016.

[6] Maurice Bertram Priestley. Spectral analysis and time series. 1981.

[7] Douglas C Montgomery, Cheryl L Jennings, and Murat Kulahci. *Introduction to time series analysis and forecasting*. John Wiley & Sons, 2015.

[8] George EP Box and Gwilym M Jenkins. Time series analysis, control, and forecasting. *San Francisco, CA: Holden Day*, 3226, 1976.

[9] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.

[10] Robert B Ash and Melvin F Gardner. *Topics in stochastic processes*, volume 27. Elsevier, 1975.

[11] O Onalan. Financial modelling with ornstein-uhlenbeck processes driven by lévy process. In *Proceedings of the World Congress on Engineering*, volume 2, pages 1–3, 2009.

[12] Helmut Lütkepohl, Markus Krätzig, and Peter CB Phillips. *Applied time series econometrics*. Cambridge University press, 2004.

[13] John N Haddad. A simple method for computing the covariance matrix and its inverse of a stationary autoregressive process. *Communications in Statistics-Simulation and Computation*, 27(3):617–623, 1998.

[14] Ngai Hang Chan. *Time series: Applications to finance with R and S-Plus*, volume 837. John Wiley & Sons, 2011.

[15] Jean-Luc Gauvain and Chin-Hui Lee. Maximum a posteriori estimation for multivariate gaussian mixture observations of markov chains. *IEEE transactions on speech and audio processing*, 2(2):291–298, 1994.

[16] Bruce R Kusse and Erik A Westwig. *Mathematical physics: applied mathematics for scientists and engineers*. John Wiley & Sons, 2010.

[17] Mátyás Arató. *Linear stochastic systems with constant coefficients: a statistical approach*, volume 45. Springer, 1982.

[18] Kiyosi Itô, P Henry Jr, et al. *Diffusion processes and their sample paths*. Springer Science & Business Media, 2012.

[19] Rob J Hyndman. Yule-walker estimates for continuous-time autoregressive models. *Journal of Time Series Analysis*, 14(3):281–296, 1993.

[20] PCB Phillips. The structural estimation of a stochastic differential equation system. *Econometrica: Journal of the Econometric Society*, pages 1021–1041, 1972.

[21] Peter J Brockwell, Richard Davis, and Yu Yang. Continuous-time gaussian autoregression. *Statistica Sinica*, 17:63–80, 2007.

Appendix A

Simulation and Estimation

A.1 Simulating and Estimating Discrete Autoregressive Process of Order Two in R

Let us generate 50 observations for a simulated stationary DAR(2) process $\{X_t\}$ with $\phi_1 = -0.2$ and $\phi_2 = 0.35$. To ensure that we all get the same results, we set the seed to a predetermined value before we generate values for the respective variable, which has been appointed as X.

```
1 set.seed(12345)
2 nu <- rnorm(50)
3 X <- rnorm(50)
4 a <- -0.2
5 b <- 0.35
6 X[2] = a * X[1] + nu[2]
7 for (t in 3:50) X[t] = a * X[t-1] + b * X[t-2] + nu[t]
```

To obtain the Yule-Walker estimates:

```
1 ar.yw <- ar(X, order=2, method="yw", demean=FALSE)
2 ar.yw$ar
```

To obtain the Maximum Likelihood estimates:

```
1 ar.mle <- ar(X, order=2, method="mle", demean=FALSE)
2 ar.mle$ar
```

A.2 Computing MLE of a DAR(2) process by Using Maple

Let us consider the simulated stationary DAR(2) process that we generated in Appendix A.1. In order to compute the MLE of the coefficients r_1 and r_2,

we have to solve the following system of equations

$$\frac{\partial\{X'\Sigma_p^{-1}X\}}{\partial r_j} = 0,$$

where $\Sigma_p^{-1} = [L(r_1)L(r_2)][L'(r_1)L'(r_2)]$. The following are the maple codes required to solve the system of equations.

```
1 with(LinearAlgebra);
2 n := 50
3 Lr1 := Matrix(n, shape = triangular[lower], storage = band[1,
     0], scan = band[1, 0], [[‘$‘(-r1, n-1)], [‘$‘(1, n)]])
4 Lr2 := Matrix(n, shape = triangular[lower], storage = band[1,
     0], scan = band[1, 0], [[‘$‘(-r2, n-1)], [‘$‘(1, n)]])
5 tLr1 := Transpose(Lr1)
6 tLr2 := Transpose(Lr2)
7 interface(rtablesize = infinity);
```

Insert the DAR(2) data, that we simulated by using R, into a $50x1$ column matrix denoted by X.

```
1 f := factor(map(diff,Transpose(X).Lr1.Lr2.tLr1.tLr2.X, r1))
2 g := factor(map(diff,Transpose(X).Lr1.Lr2.tLr1.tLr2.X, r2))
3 solve({f = 0, g = 0}, {r1, r2})
```

A.3 Simulating an Ornstein-Uhlenbeck Process in R

In R, a package named "Sim.DiffProc" provides functions to deal with a wide range of stochastic differential equations including the discrete version of the Ornstein-Uhlenbeck process.

Let us generate 1000 observations for a stationary OU-process $\{X_t\}$ with parameter $\alpha_0 = 0.3$. The following are the R codes required to simulate an OU-process.

```
1 install.packages("Sim.DiffProc")
2 library("Sim.DiffProc")
3 set.seed(1234)
4 X <- OU(N=1000, T=10, t0=0, mu=0.3, sigma=0.1, x0=0)
```

A.4 Computing Estimates of an OU-Process by Using Maple

Let us consider the simulated stationary Ornstein-Uhlenbeck process that we generated in Appendix A.3. In order to compute the estimates on maple, first write down the following codes:

```
1 with(LinearAlgebra);
2 interface(rtablesize = infinity);
3 T := 10
4 n := 1000
5 Delta := T/n
```

Insert the OU data, that we simulated by using R, into a $1000x1$ column matrix denoted by X.

To acquire the estimate $\hat{\alpha}_0 = -\frac{1}{\Delta} log(\frac{\hat{\gamma}_1}{\hat{\gamma}_0})$:

```
1 g := (1/n)*(sum('X[t]'*'X[t+1]', t = 1 .. n-1))
2 y := (1/n)*(sum('X[t]'*'X[t]', t = 1 .. n))
3 (-1/Delta) * log(g/y)
```

To obtain the Yule-Walker estimate given by Hyndman:

```
1 z := (1/n)*(sum('X[t]'*'X[t+1]', t = 1 .. n-1))
2 y := (1/n)*(sum('X[t]'*'X[t]', t = 1 .. n))
3 -(1/Delta)*(z/y-1)
```

Appendix B

Brownian Motion

Suppose that $\{\upsilon_t, 0 \leq t < \infty\}$ is a Gaussian continuous-time white noise process. If the continuous-time stochastic process $\{W_t, 0 \leq t < \infty\}$ is defined by

$$W_t = \int_0^t \upsilon_u du, \tag{B.0.1}$$

then $\{W_t\}$ is a Brownian motion.

Proof. We will show that $\{W_t, 0 \leq t < \infty\}$ satisfies all the properties of Brownian motion that is set out in section (1.4).

1. $W_0 = \int_0^0 \upsilon_u du = 0$.
2. For $t > 0$ and $s > 0$, we have

$$E[W_t] = E\bigg[\int_0^t \upsilon_u du\bigg] = \int_0^t E[\upsilon_u] du = 0,$$

and

$$\begin{aligned} Cov[W_s, W_t] &= Cov\bigg[\int_0^{min\{s,t\}} \upsilon_u du, \int_0^{max\{s,t\}} \upsilon_v dv\bigg] \\ &= \int_0^{min\{s,t\}} \int_0^{max\{s,t\}} \sigma_v^2 \delta(v-u) du dv \\ &= \int_0^{min\{s,t\}} \sigma_v^2 du \\ &= \sigma_v^2 \text{min\{s,t\}}. \end{aligned}$$

In particular, $Var[W_t] = Cov[W_t, W_t] = \sigma_v^2 t$.

3. For $t > 0$ and $s > 0$,

$$E[W_{t+s} - W_t] = E[W_{t+s}] - E[W_t] = 0,$$

and

$$\begin{aligned} Var[W_{t+s} - W_t] &= Var[W_{t+s}] + Var[W_t] \\ &\quad - 2Cov[W_{t+s}, W_t] \\ &= \sigma_v^2(t+s) + \sigma_v^2 t - 2\sigma_v^2 t \\ &= \sigma_v^2 s. \end{aligned}$$

The distribution of $W_{t+s} - W_t$ is independent of t. Hence $\{W_t, 0 \leq t < \infty\}$ has stationary increments.

4. Let t_1, t_2 and t_3 such that $0 \leq t_1 < t_2 < t_3$. We have

$$\begin{aligned} Cov[W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}] &= \sigma_v^2 t_2 - \sigma_v^2 t_2 - \sigma_v^2 t_1 + \sigma_v^2 t_1 \\ &= 0. \end{aligned}$$

Thus $W_{t_2} - W_{t_1}$ and $W_{t_3} - W_{t_2}$ are independent. Therefore the process $\{W_t, 0 \leq t < \infty\}$ has independent increments.

Hence $\{W_t, 0 \leq t < \infty\}$ satisfies all the properties of Brownian motion. $\square$