DISSERTATION

ESTIMATION FOR LÉVY-DRIVEN CARMA PROCESSES

Submitted by

Yu Yang

Department of Statistics

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2008

COLORADO STATE UNIVERSITY

September 24, 2007

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY YU YANG ENTITLED ESTIMATION FOR LÉVY-DRIVEN CARMA PROCESSES BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

P. J. Brockwell

Peter J. Brockwell (Adviser)

Richard A. Davis

Richard A. Davis (Co-adviser)

F. Jay Breidt (Committee Member)

Louis L. Scharf (Committee Member)

F. Jay Breidt (Department Head)

ABSTRACT OF DISSERTATION

ESTIMATION FOR LÉVY-DRIVEN CARMA PROCESSES

This thesis explores parameter estimation for Lévy-driven continuous-time autoregressive moving average (CARMA) processes, using uniformly and closely spaced discrete-time observations.

Specifically, we focus on developing estimation techniques and asymptotic properties of the estimators for three particular families of Lévy-driven CARMA processes. Estimation for the first family, Gaussian autoregressive processes, was developed by deriving exact conditional maximum likelihood estimators of the parameters under the assumption that the process is observed continuously. The resulting estimates are expressed in terms of stochastic integrals which are then approximated using the available closely-spaced discrete-time observations. We apply the results to both linear and non-linear autoregressive processes. For the second family, non-negative Lévy-driven Ornestein-Uhlenbeck processes, we take advantage of the non-negativity of the increments of the driving Lévy process to derive a highly efficient estimation procedure for the autoregressive coefficient when observations are available at uniformly spaced times. Asymptotic properties of the estimator are also studied and a procedure for obtaining estimates of the increments of the driving Lévy process is developed. These estimated increments are important for identifying the nature of the driving Lévy process and for estimating its parameters. For the third family, non-negative Lévy-driven CARMA processes, we estimate the coefficients by maximizing the Gaussian likelihood of the observations and discuss the asymptotic properties of the estimators. We again show how to estimate the increments of the background driving Lévy process and hence to estimate the parameters of the Lévy process itself. We assess the performance of our estimation procedures by simulations and use them to fit models to real data sets in order to determine how the theory applies in practice.

Yu Yang
Department of Statistics
Colorado State University
Fort Collins, Colorado 80523
Summer 2008

ACKNOWLEDGEMENTS

I would like to express my deepest and most sincere gratitude to Professor Peter Brockwell, my advisor, and Professor Richard Davis, my co-advisor. With their enthusiasm, their inspiration, and their great efforts to explain things clearly and simply, they helped to make statistics fun for me. It was a great pleasure for me to conduct this thesis under their supervision. I feel very lucky to have them as my mentors. Knowing them is actually one of the best things that ever happened to me in my life.

I gratefully appreciate the precious opportunity I got to learn from many professors, while attending Colorado State University. I feel especially grateful for having met Professor Hari Iyer, Professor Jan Hannig and Professor Jennifer Hoeting. They always take the time to answer questions and to offer encouragement. My time at CSU wouldn't have been this enjoyable without them.

This thesis could not be finished without many help from my colleagues and friends. I wish to extend my warmest thanks to all of them, especially to Megan Higgs, Derek Sonderegger and Sunghoon Chung for their kindness and support.

I want to give my special thanks to my husband Zhi whose patient love enabled me to complete this work. And I would also like to thank my parents, Chengfu Yang and Dianling Diao. They bore me, raised me, supported me and love me. To them, I dedicate this thesis.

The financial support of the Department of Statistics and the National Science Foundation (DMS 0308109) are gratefully acknowledged.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

1.1 Motivation

Today, in the world's economies and financial markets, one major challenge is to form realistic models to represent the economy and those markets. In response to this challenge, researchers have been doing a great deal of work in econometrics using continuous-time models during the past few decades, partly motivated by the very successful use of continuous-time models in option pricing following the seminal work of Black, Scholes and Merton.

An obvious argument that favors the use of continuous-time models over traditional discrete-time models is that most economic processes are inherently continuous in time. More convenience and flexibility in handling irregularly spaced data also make continuous-time models better candidates. Besides, due to rapidly developing technologies, data are more frequently available at higher frequency, for example, the currently available tick-by-tick transaction data. It is then more natural to deal with fast sampled data using continuous-time models rather than discrete-time models.

However, a major problem faced in continuous-time modeling is that we do not have a continuous-time record of observations. Thus, parameter estimation of continuous-time models using available discrete-time data has become an important subject, which is also the motivating factor behind much of the thesis.

1.2 Lévy Processes

Among all the continuous-time models, we will focus on continuous-time autoregressive moving average (CARMA) processes driven by second-order Lévy processes. Before proceeding any further, we first record a few essential facts concerning Lévy processes, named after the great French mathematician Paul Lévy.

1.2.1 Definition

Suppose we are given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le \infty}, P)$, where $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$ and $(\mathcal{F}_t)$ is right-continuous.

Definition 1.2.1 (Lévy Process). $\{L(t), t \ge 0\}$ is an $(\mathcal{F}_t)$-adapted Lévy process if $L(t) \in \mathcal{F}_t$ for all $t \ge 0$ and

(1) $L(0) = 0$ a.s.

(2) $L(t)$ has independent increments, i.e., $L(t) - L(s)$ is independent of $\mathcal{F}_s$, for any $0 \le s < t < \infty$.

(3) $L(t)$ has stationary increments, i.e., $L(t+s) - L(s)$ has the same distribution as $L(t)$, for any $s, t > 0$.

(4) $L(t)$ is stochastically continuous, i.e., for all $\epsilon > 0$ and all $t \ge 0$,

$$\lim_{s \to t} P\left(|L(t) - L(s)| > \epsilon\right) = 0.$$

Every Lévy process has a unique modification which is càdlàg (right continuous with left limits) and which is also a Lévy process. We shall therefore assume that our Lévy process has these properties. For further properties of Lévy processes, see the books of Protter (2004), Applebaum (2004) and Sato (1999). The characteristic function of $L(t)$, $\phi_t(\theta) := E\left(\exp(i\theta L(t))\right)$, has the form

$$\phi_t(\theta) = \exp(t\xi(\theta)),\ \theta \in \mathbb{R},$$

where $\xi(\theta)$ is often called *characteristic exponent* and satisfies the following Lévy-Khinchin formula,

$$\xi(\theta) = i\theta m - \frac{1}{2}\theta^2\eta^2 + \int_{\mathbb{R}_0} \left(e^{i\theta x} - 1 - i\theta x I_{(-1,1)}(x)\right)\nu(dx),$$

for some $m \in \mathbb{R}$, $\eta \geq 0$, and measure ν on the Borel subsets of $\mathbb{R}_0 = \mathbb{R}\setminus\{0\}$. The measure ν is called the Lévy measure of the process L and has the property,

$$\int_{\mathbb{R}_0} \min(|u|^2, 1)\nu(du) < \infty.$$

From the Lévy-Khinchin formula, we see that, in general, a Lévy process can be decomposed into three parts: a constant drift part, a Brownian motion part, and a pure jump part. If A is a Borel subset of $\{x : |x| > \epsilon\}$ for some $\epsilon > 0$, then the number of jumps with sizes in A, occurring in any time interval of length $t > 0$, has the Poisson distribution with mean $t\nu(A)$. If ν is a finite measure, i.e. $\nu(\mathbb{R}_0) = \int_{\mathbb{R}_0}\nu(dx) < \infty$, then almost all paths of L have a finite number of jumps on every compact interval and the process is said to have *finite activity*. Otherwise, if $\nu(\mathbb{R}_0) = \infty$, then an infinite number of jumps occur in any interval of positive length with probability one and the process is said to have *infinite activity*. As we shall see later in Section 1.2.2, Poisson processes and compound Poisson processes have finite activity, while gamma processes and inverse Gaussian processes have infinite activity.

1.2.2 Examples of Lévy Processes

The triplet (m, η, ν) in the characteristic exponent is called the *Lévy triplet*, which describes each Lévy process completely. A wealth of distributions for $L(t)$ is attainable by suitable choices of Lévy triplets.

In the following we list a few frequently-used Lévy processes.

Brownian Motion. If ν is the zero measure, then $\{L(t)\}$ is a Brownian motion with $E(L(t)) = mt$ and $\mathrm{Var}(L(t)) = \eta^2 t$. A sample path of a Brownian motion with $m = -1$ and $\eta = 1$ is shown in Figure 1.1.

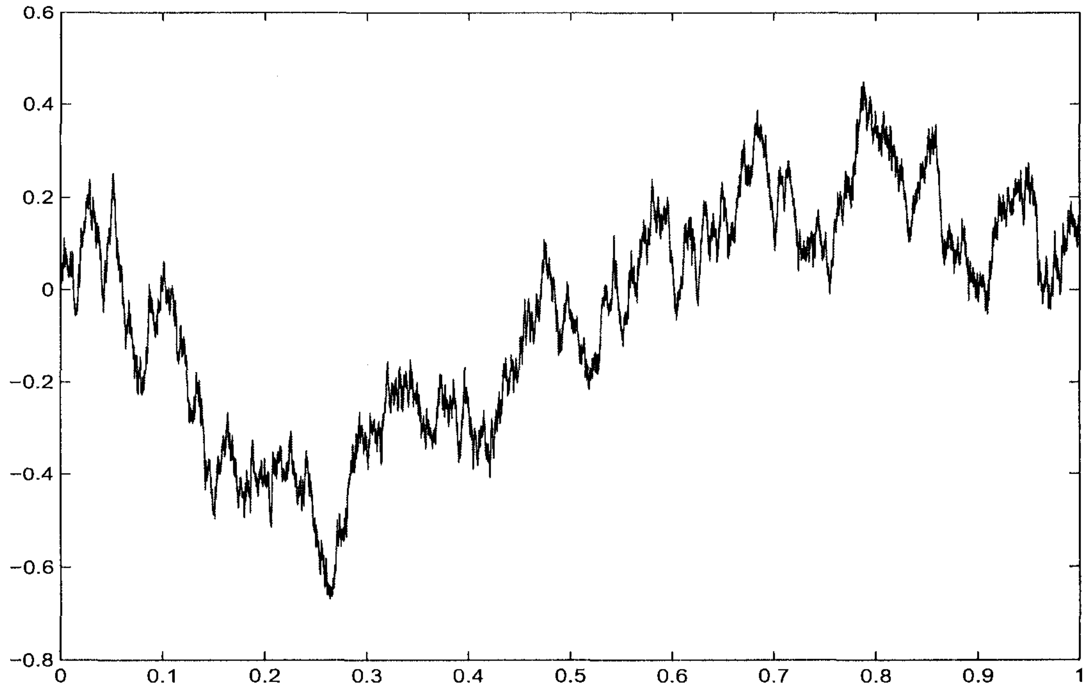


Figure 1.1: A sample path of a Brownian motion.

Poisson Process. A Poisson process with intensity parameter $\lambda > 0$ starts at 0, with $L(s)$ following a Poisson(λs) distribution. The Lévy triplet of a Poisson process is given by $(0, 0, \lambda\delta(1))$, where $\delta(1)$ denotes the Dirac measure with total mass 1 concentrated at the point 1.

Compound Poisson Process. The compound Poisson process is a Lévy process with triplet $(m, 0, \lambda F)$ where $\lambda \in (0, \infty)$ is the mean rate of occurrence of jumps per unit time, F is the distribution of the jump-sizes and $m = \int_{|x|<1} x\lambda dF(x)$. Figure 1.2 shows a sample path of a compound Poisson process with $\lambda = 1$ and jump-sizes following standard normal distribution.

In the special case, when L is a non-decreasing Lévy process (also called a subordinator), its Lévy exponent $\xi(\theta)$ can be written in the form

$$\xi(\theta) = i\theta m^* + \int_{\mathbb{R}_0} \left(e^{i\theta x} - 1\right) \nu(dx), \tag{1.1}$$

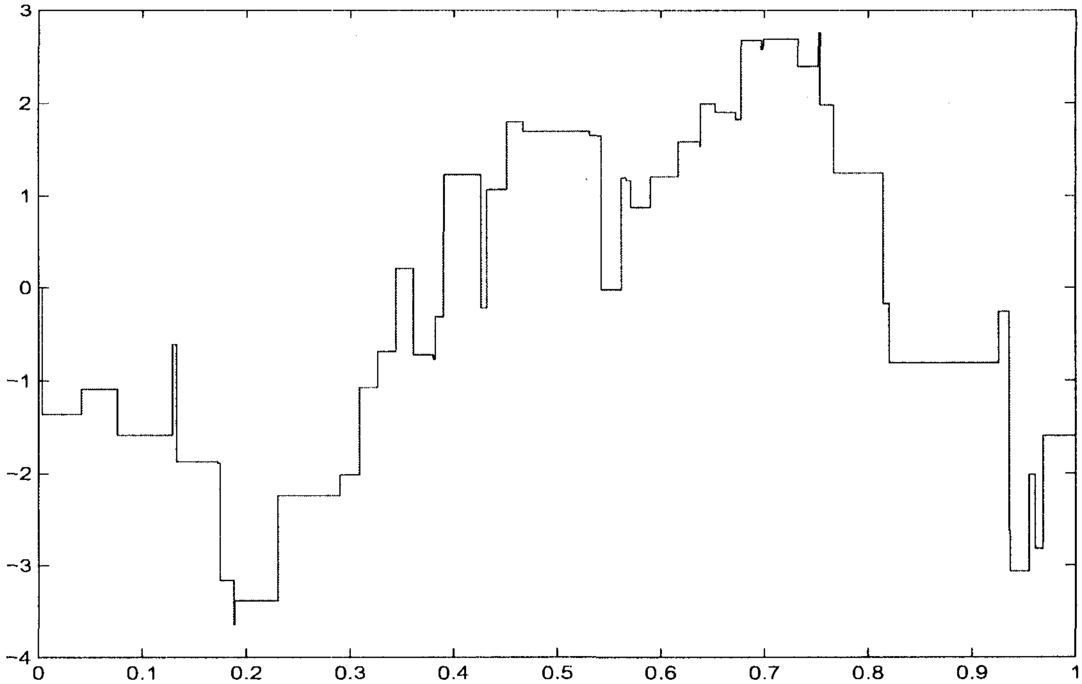


Figure 1.2: A sample path of a compound Poisson process.

where $m^* = m - \int_{|x|<1} x d\nu(x)$ and we shall refer to m^* as the drift. Non-decreasing Lévy processes are widely used for financial modeling. See, for example, Barndorff-Nielsen and Shephard (2001), and Schoutens (2003). Three frequently-used subordinators are defined below.

Gamma Process. If $m^* = 0$ and $\nu(du) = \alpha u^{-1} e^{-\beta u} I_{\{u>0\}} du$, $L(t)$ is a gamma process. $L(s)$ has a Gamma$(\alpha s, \beta)$ distribution. Since $\nu(\mathbb{R}_0) = \int_0^\infty \alpha u^{-1} e^{-\beta u} du = \infty$, gamma processes have an infinite number of jumps on every interval with positive length. Figure 1.3 shows a sample path of a gamma process with $\alpha = 6$ and $\beta = \sqrt{6}$.

Inverse Gaussian Process. If $m^* = 0$ and $\nu(du) = (2\pi)^{(-1/2)} \delta u^{(-3/2)} \exp(-\gamma^2 u/2) I_{\{u>0\}} du$ with $\delta > 0$ and $\gamma \geq 0$, we obtain an inverse Gaussian process $L(t)$. Like the gamma process, the inverse Gaussian process has infinite activity. A sample path of $L(t)$ with $\delta = 3$ and $\gamma = 2$ is shown in Figure 1.4.

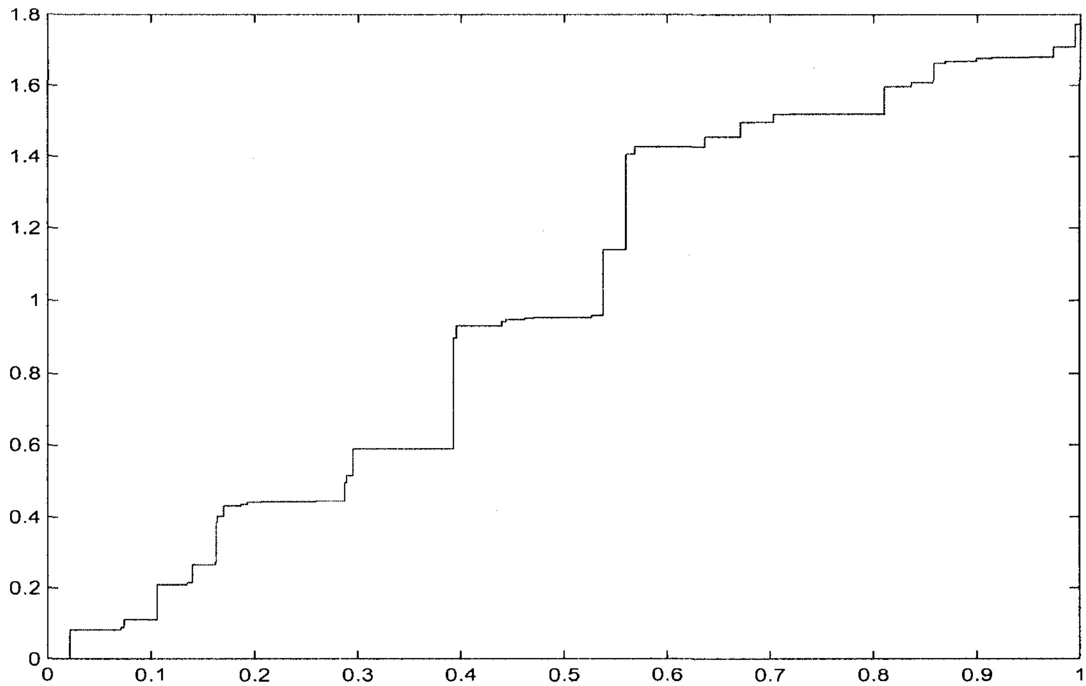


Figure 1.3: A sample path of a gamma process.

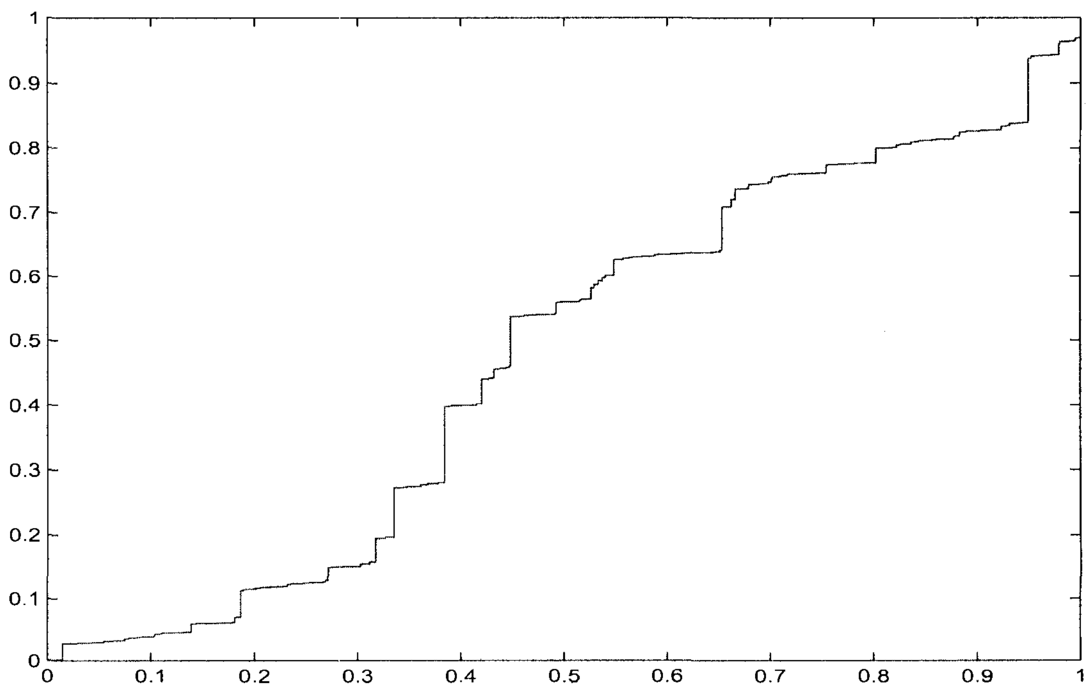


Figure 1.4: A sample path of an inverse Gaussian process.

Stable Subordinator. When $m^* = 0$ and $\nu(du) = \alpha\beta u^{-1-\beta}du$, where $\alpha > 0$ and $0 < \beta < 1$, we obtain the stable subordinator with exponent β and scale parameter $\alpha^{1/\beta}$. In this case $L(t)$ has the positive stable distribution with Laplace transform

$$E\left[\exp\left(-\lambda L(t)\right)\right] = \exp\left[-t\alpha\lambda^{\beta}/\Gamma(1-\beta)\right],\ \lambda > 0.$$

1.2.3 Second-order Lévy Processes

For the second-order Lévy processes, $E(L(1))^2 < \infty$ and there exist real constants μ and σ such that

$$E(L(t)) = \mu t \text{ and } \mathrm{Var}(L(t)) = \sigma^2 t, \text{ for } t \geq 0.$$

To avoid problems of parameter identifiability in the CARMA process defined in the next section, we assume throughout that L is scaled so that $\mathrm{Var}(L(1)) = 1$. Then $\mathrm{Var}(L(t)) = t$ for $t \geq 0$ and we shall refer to the process L as a *standardized second-order Lévy process.* Throughout Chapter 3 and Chapter 4, we shall be concerned with CARMA processes driven by standardized second-order Lévy processes.

1.3 Lévy-driven CARMA Processes

1.3.1 Definition and Properties

Definition 1.3.1 (Lévy-driven CARMA Process). A second-order Lévy-driven continuous-time ARMA(p,q) process is defined (see Brockwell, 2001b) via the state-space representation of the formal stochastic differential equation

$$a(D)Y(t) = \sigma b(D)DL(t),\ t \geq 0, \tag{1.2}$$

where σ is a strictly positive scale parameter, D denotes differentiation with respect to t, $\{L(t), t \geq 0\}$ is a second-order Lévy process,

$$a(z) := z^p + a_1 z^{p-1} + \cdots + a_p\,,$$

$$b(z) := b_0 + b_1 z + \cdots + b_{p-1} z^{p-1},$$

and the coefficients b_j satisfy $b_q = 1$ and $b_j = 0$ for $q < j < p$. The behavior of the process is determined by the process L and the coefficients $\{a_j, 1 \leq j \leq p; b_j, 0 \leq j < q; \sigma\}$. To avoid trivial and easily eliminated complications, we shall assume that $a(z)$ and $b(z)$ have no common factors. The state-space representation includes the *observation* and *state* equations,

$$Y(t) = \sigma \mathbf{b}'\mathbf{X}(t), \tag{1.3}$$

and

$$d\mathbf{X}(t) - A\mathbf{X}(t)dt = \mathbf{e}\, dL(t), \tag{1.4}$$

where the superscript $'$ denotes taking transpose,

$$\mathbf{X}(t) = \begin{bmatrix} X_0(t) \\ X_1(t) \\ \vdots \\ X_{p-2}(t) \\ X_{p-1}(t) \end{bmatrix}, \mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{bmatrix},$$

where $X_j(t)$ is the j^{th} mean-square and pathwise derivative $D^j X_0(t)$, $j = 0, \ldots, p-1$, and

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \cdots & -a_1 \end{bmatrix}.$$

In the special case when $\{L(t)\}$ is a standard Brownian motion, (1.4) is an Itô equation with solution $\{\mathbf{X}(t), t \geq 0\}$ satisfying

$$\mathbf{X}(t) = e^{At}\mathbf{X}(0) + \int_0^t e^{A(t-u)}\mathbf{e}\, dL(u), \tag{1.5}$$

where the integral is defined under the framework of Itô integrals as the L^2 limit of approximating Riemann-Stieltjes sums. For any second-order driving Lévy process, $\{L(t)\}$, the integral can be defined in the same way. If, in addition, $\{L(t)\}$ is a non-decreasing Lévy process (and hence has bounded variation on compact intervals) as is the case in Chapter 3 and Chapter 4, $\{\mathbf{X}(t)\}$ can be defined as a Riemann-Stieltjes integral by (1.5) for each sample path, which really makes things a lot easier. From (1.5), we can also write

$$\mathbf{X}(t) = e^{A(t-s)}\mathbf{X}(s) + \int_s^t e^{A(t-u)}\mathbf{e}\, dL(u), \text{ for all } t > s \geq 0, \tag{1.6}$$

showing, by independence of the increments of $\{L(t)\}$, that $\{\mathbf{X}(t)\}$ is Markov. The following proposition gives necessary and sufficient conditions for stationarity of $\{\mathbf{X}(t)\}$. For a proof, see Brockwell and Marquardt (2005).

Proposition 1.3.1 *If $\{\mathbf{X}(0)\}$ is independent of $\{L(t), t \geq 0\}$ and $E(L(1)^2) < \infty$, then $\{\mathbf{X}(t)\}$ is strictly stationary if and only if the eigenvalues of matrix A all have strictly negative real parts and $\{\mathbf{X}(t)\}$ has the distribution of $\int_0^\infty e^{-Au}\mathbf{e}\, dL(u)$.*

Remark 1.3.1 It is easy to check that the eigenvalues of matrix A, which we shall denote by $\lambda_1, \ldots, \lambda_p$, are the same as the zeroes of the autoregressive polynomial $a(z)$. The corresponding right eigenvectors are

$$\left[\begin{array}{ccccc} 1 & \lambda_j & \lambda_j^2 & \cdots & \lambda_j^{p-1} \end{array}\right]', \; j = 1, \ldots, p.$$

Remark 1.3.2 If we introduce a second Lévy process $\{M(t), 0 \leq t < \infty\}$, independent of L and with the same distribution, the stationary CARMA process defined over non-negative t can be extended so that it is a stationary process over all real t. Define the following extension of L:

$$L^*(t) = L(t)I_{[0,\infty)}(t) - M(-t-)I_{(-\infty,0]}(t), \quad -\infty < t < \infty.$$

Then, provided the eigenvalues of A all have negative real parts, i.e.,

$$\mathcal{R}e(\lambda_j) < 0, \; j = 1, \ldots, p, \tag{1.7}$$

the process $\{\mathbf{X}(t)\}$ defined by

$$\mathbf{X}(t) = \int_{-\infty}^t e^{A(t-u)}\mathbf{e}\, dL^*(u), \tag{1.8}$$

is a strictly stationary process satisfying equation (1.6) (with L replaced by L^*) for all $t > s$ and $s \in (-\infty, \infty)$ with the corresponding CARMA process given by

$$Y(t) = \sigma \mathbf{b}'\mathbf{X}(t) = \int_{-\infty}^{\infty} g(t-u)dL^*(u), \tag{1.9}$$

where the function $g(t) = \sigma \mathbf{b}' e^{At} \mathbf{e} I_{[0,\infty)}(t)$, is referred to as the **kernel of the CARMA process** $\{Y(t)\}$. Henceforth, we restrict our attention to stationary CARMA processes satisfying (1.7) and refer to L^* as the *background driving Lévy process* and denote it by L for simplicity. From (1.9) we easily find that the mean, $EY(t)$, and autocovariance function, $\gamma(h) := \text{Cov}(Y(t+h), Y(t))$, are given by

$$EY(t) = \sigma b_0 \mu / a_p \,,$$

where $\mu = EL(1)$ and

$$\gamma(h) = \sigma^2 \mathbf{b}' e^{A|h|} \Sigma \mathbf{b},$$

where

$$\Sigma = \int_0^{\infty} e^{Ay} \mathbf{e}\mathbf{e}' e^{A'y} dy.$$

It can also be shown that the spectral density of the process Y is

$$f(\omega) = \frac{\sigma^2}{2\pi} \left| \frac{b(i\omega)}{a(i\omega)} \right|^2,$$

which is clearly a rational function of the frequency ω, $-\infty < \omega < \infty$.

Remark 1.3.3 When the zeroes $\lambda_1, \cdots, \lambda_p$ of $a(z)$ are distinct and satisfy the causality condition (1.7), Brockwell (2001a) showed that the expression for the kernel g takes a simple form,

$$g(h) = \sigma \sum_{r=1}^{p} \frac{b(\lambda_r)}{a'(\lambda_r)} e^{\lambda_r h} I_{[0,\infty)}(h), \tag{1.10}$$

and the autocovariance function of Y can be written as

$$\gamma(h) = \sigma^2 \sum_{j=1}^{p} \frac{b(\lambda_j) b(-\lambda_j)}{a'(\lambda_j) a(-\lambda_j)} e^{\lambda_j |h|},$$

where $a'(z)$ denotes the derivative of $a(z)$. When the autoregressive roots are distinct, we also obtain a very useful presentation of the CARMA(p,q) process Y from (1.10). Defining

$$\alpha_r = \sigma \frac{b(\lambda_r)}{a'(\lambda_r)}, \; r = 1, \ldots, p, \tag{1.11}$$

we can write

$$Y(t) = \sum_{r=1}^{p} Y_r(t), \tag{1.12}$$

where

$$Y_r(t) = \int_{-\infty}^{t} \alpha_r e^{\lambda_r(t-u)} dL(u). \tag{1.13}$$

This expression shows that the component processes Y_r satisfy the simple equations,

$$Y_r(t) = Y_r(s)e^{\lambda_r(t-s)} + \int_s^t \alpha_r e^{\lambda_r(t-u)} dL(u), \; t \geq s, \; r = 1, \ldots, p. \tag{1.14}$$

Taking $s = 0$ and using Lemma 2.1 of Eberlein and Raible (1999), we can also write

$$Y_r(t) = Y_r(0)e^{\lambda_r t} + \alpha_r L(t) + \int_0^t \alpha_r \lambda_r e^{\lambda_r(t-u)} L(u) du, \; t \geq 0, \tag{1.15}$$

where the last integral is a Riemann integral and the equality holds for all finite $t \geq 0$ with probability 1. Defining

$$\mathbf{Y}(t) := [Y_1(t), \cdots, Y_p(t)]', \tag{1.16}$$

we obtain from (1.8), (1.10) and (1.13),

$$\mathbf{Y}(t) = \sigma B R^{-1} \mathbf{X}(t), \tag{1.17}$$

where $B = \mathrm{diag}[b(\lambda_i)]_{i=1}^{p}$ and $R = [\lambda_j^{i-1}]_{i,j=1}^{p}$. The initial values $Y_r(0)$ in (1.15) can therefore be obtained from those of the components of the state vector $\mathbf{X}(0)$. The process $\mathbf{Y}$ provides us with an alternative **canonical state representation** of $Y(t)$, $t \geq 0$, namely

$$Y(t) = [1, \cdots, 1]\mathbf{Y}(t), \tag{1.18}$$

where $\mathbf{Y}$ is the solution of

$$d\mathbf{Y}(t) = \mathrm{diag}[\lambda_i]_{i=1}^p \mathbf{Y}(t)dt + \sigma BR^{-1}\mathbf{e}\, dL, \tag{1.19}$$

with $\mathbf{Y}(0) = \sigma BR^{-1}\mathbf{X}(0)$.

Example 1.3.1 (The CARMA(2,1) Process). In this case, $b(z) = b_0 + z$, $a(z) = (z - \lambda_1)(z - \lambda_2)$ and λ_1 and λ_2 satisfy causality condition (1.7). Assuming $\lambda_1 \neq \lambda_2$, we have from (1.10) that

$$g(h) = (\alpha_1 e^{\lambda_1 h} + \alpha_2 e^{\lambda_2 h}) I_{[0,\infty)}(h),$$

where $\alpha_r = \sigma(b_0 + \lambda_r)/(\lambda_r - \lambda_{3-r})$, $r = 1, 2$. From (1.17) the canonical state vector is

$$\mathbf{Y}(t) = \begin{bmatrix} Y_1(t) \\ Y_2(t) \end{bmatrix} = \frac{\sigma}{\lambda_1 - \lambda_2} \begin{bmatrix} \lambda_2(b_0 + \lambda_1) & -(b_0 + \lambda_1) \\ -\lambda_1(b_0 + \lambda_2) & b_0 + \lambda_2 \end{bmatrix} \mathbf{X}(t),$$

and the canonical representation of Y is, based on (1.12) and (1.13),

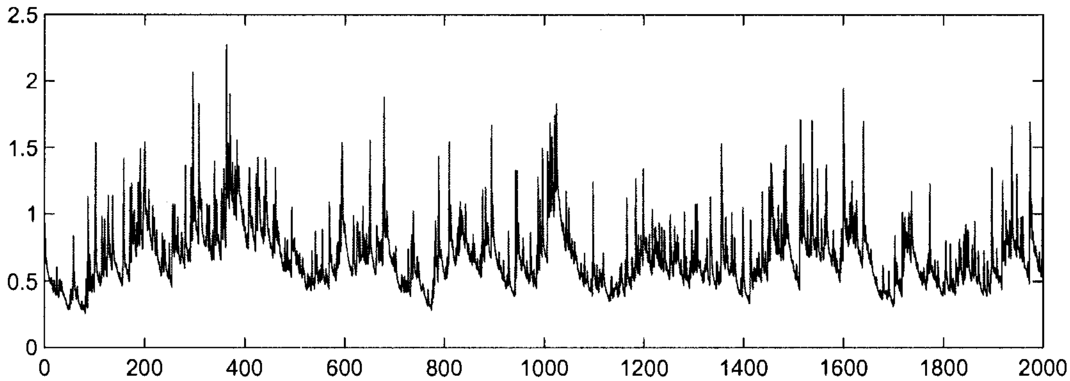
$$Y(t) = Y_1(t) + Y_2(t),$$

where

$$Y_r(t) = \int_{-\infty}^{t} \alpha_r e^{\lambda_r(t-u)} dL(u), \; r = 1, 2.$$

For illustrative purposes, we show one realization of a gamma-driven CARMA(2,1) process on time interval $[0, 2000]$ as well as its sample autocorrelation function (ACF) in Figure 1.5. The parameters we use are $a_1 = 1.3619$, $a_2 = 0.0444$, $b_0 = 0.2062$, $\sigma = 0.2888$ and L is the standardized gamma process with $E(L(t)) = 0.5t$.

Remark 1.3.4 Based on (1.9), we can easily conclude that together with non-negativity of the kernel g and the non-decreasing property of the driving Lévy process L, the process Y will be non-negative as is necessary if Y is used to represent volatility. For general CARMA processes, Tsai and Chan (2005) showed that the kernel is non-negative if and only if the ratio $b(\cdot)/a(\cdot)$ is completely monotone. In



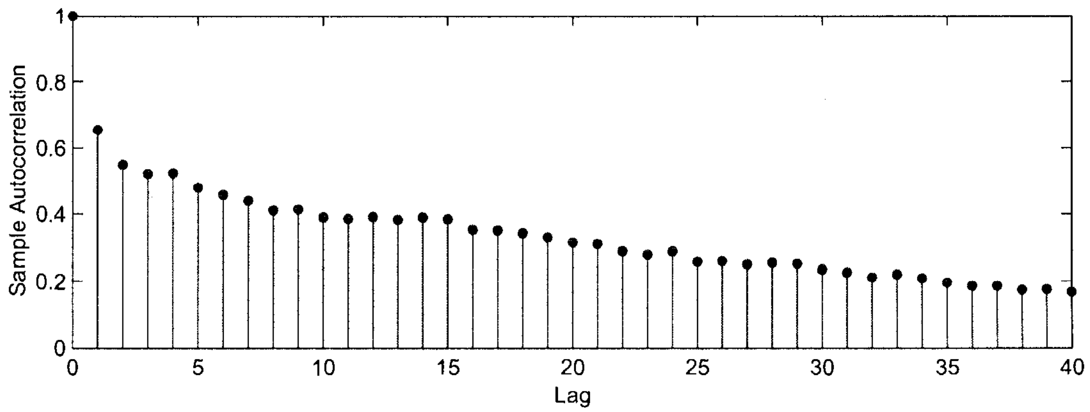


Figure 1.5: Realization and ACF of a gamma-driven CARMA(2,1) process.

particular, for a stationary CARMA(2,1) process defined by (1.2) (we shall discuss this process further in later chapters), a necessary and sufficient condition for the kernel to be non-negative is that the two roots λ_1 and λ_2 of $a(z) = 0$ are real and $b_0 \geq \min(|\lambda_i|)$.

1.3.2 Sampled Process

When we observe a Lévy-driven CARMA process at uniformly spaced times $0, h, 2h, \ldots$, the sampled process we obtain is actually a discrete-time autoregressive moving average (ARMA) process. There are certainly a lot of connections between continuous-time ARMA processes and discrete-time ARMA processes, some of which we shall utilize later in Chapter 4. Thus, it is necessary that we give a brief introduction to the discrete-time ARMA processes. For more details regarding the sampled process, see Brockwell (2001b).

Definition 1.3.2 (ARMA Process). A zero-mean discrete-time ARMA(p, q) process $\{Y_n\}$ with autoregressive coefficients $\phi_1, \dots, \phi_p$, moving average coefficients $\theta_1, \dots, \theta_q$, and white noise variance σ_d^2, is defined to be a (weakly) stationary solution of the p^{th} order linear difference equations,

$$\phi(B)Y_n = \theta(B)Z_n, \; n = 0, \pm 1, \pm 2, \dots, \tag{1.20}$$

where B is the backward shift operator ($BY_n = Y_{n-1}$ and $BZ_n = Z_{n-1}$ for all n), $\{Z_n\}$ is a sequence of uncorrelated random variables with mean zero and variance σ_d^2 (abbreviated to $\{Z_n\} \sim \text{WN}(0, \sigma_d^2)$) and

$$\begin{aligned}\phi(z) &:= 1 - \phi_1 z - \cdots - \phi_p z^p, \\ \theta(z) &:= 1 + \theta_1 z + \cdots + \theta_q z^q,\end{aligned}$$

with $\theta_q \neq 0$ and $\phi_q \neq 0$. We define $\phi(z) := 1$ if $p = 0$ and $\theta(z) := 1$ if $q = 0$. We shall assume that the polynomials $\phi(z)$ and $\theta(z)$ have no common zeroes and that $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ is non-zero for all complex z such that $|z| \leq 1$. This last condition guarantees the existence of a unique stationary solution of (1.20) which is also causal, i.e., is expressible in the form $Y_n = \sum_{j=0}^{\infty} \psi_j Z_{n-j}$ for some absolutely summable sequence $\{\psi_j\}$. The process $\{Y_n\}$ is said to be an ARMA(p, q) process with mean μ if $\{Y_n - \mu\}$ is an ARMA(p, q) process. A more restrictive definition of ARMA process imposes the further requirement that the random variables Z_n be independent and identically distributed, in which case we write $\{Z_n\} \sim \text{IID}(0, \sigma_d^2)$. The process $\{Y_n\}$ is then strictly (as well as weakly) stationary and we shall refer to $\{Y_n\}$ as a *strict ARMA process.* If we impose the further constraint that each Z_n is Gaussian, then we write $\{Z_n\} \sim \text{IIDN}(0, \sigma_d^2)$ and refer to $\{Y_n\}$ as a *Gaussian ARMA process.* There are many structural similarities between ARMA and CARMA processes. For example, there is a corresponding *canonical representation* analogous to that in Remark 1.3.3 of Section 1.3.1. It takes the form (cf. (1.12) and (1.13)),

$$Y_n = \sum_{r=1}^{p} Y_{r,n}, \tag{1.21}$$

and

$$Y_{r,n} = \sum_{k=-\infty}^{n} \beta_r \xi_r^{n-k} Z_k,\ r = 1, \ldots, p, \tag{1.22}$$

where ξ_r^{-1}, $r = 1, \ldots, p$ are the (distinct) zeroes of $\phi(z)$, and

$$\beta_r = -\xi_r \frac{\theta(\xi_r^{-1})}{\phi'(\xi_r^{-1})},\ r = 1, \ldots, p.$$

From (1.22) we also obtain the relations (cf. (1.14)),

$$Y_{r,n} = \xi_r Y_{r,n-1} + \beta_r Z_k,\ n = 0, \pm 1, \ldots;\ r = 1, \ldots, p. \tag{1.23}$$

Remark 1.3.5 If Y is a Gaussian CARMA process defined as in Definition 1.3.1 with standard Brownian motions as the driving process, then it is well-known (see e.g. Doob (1944), Phillips (1959), Brockwell (1995)) that the sampled process $\{Y_n^{(h)} := Y(nh), n = 0, 1, 2, \ldots\}$ with fixed $h > 0$ is a (strict) Gaussian ARMA(r, s) process with $0 \leq s < r \leq p$.

Sampling a general Lévy-driven CARMA(p, q) process with autoregressive roots $\lambda_1, \ldots, \lambda_p$ such that $e^{\lambda_1}, \ldots, e^{\lambda_p}$ are distinct is most clearly illustrated by inspection of (1.14) with $t = nh$ and $s = (n-1)h$. These equations for the discrete-time process $\{Y_n^{(h)}\}$ in the canonical representation of $Y(t)$ match the corresponding equations for the canonical components of a discrete-time ARMA process if we set

$$\xi_r = e^{\lambda_r h} \tag{1.24}$$

and

$$\beta_r Z_n = \int_{(n-1)h}^{nh} \alpha_r e^{\lambda_r (nh-u)} dL(u),\ r = 1, \ldots, p. \tag{1.25}$$

If Y is the CAR(1) process, equations (1.24) and (1.25) show that the sampled process is the strict AR(1) process satisfying

$$Y(nh) = e^{\lambda h} Y((n-1)h) + Z_n, \; n = 0, \pm 1, \ldots, \tag{1.26}$$

where

$$Z_n = \sigma \int_{(n-1)h}^{nh} e^{\lambda(nh-u)} dL(u). \tag{1.27}$$

The noise sequence $\{Z_n\}$ is iid

If the same argument is applied when $p > 1$, the determination of β_r and an iid sequence $\{Z_n\}$ to satisfy the relations (1.25) breaks down since Z_n must be the same for all r. However if we sample each component in the canonical representation of the CARMA process we obtain a strict AR(1) process by the same argument as given above for the CAR(1) process.

Remark 1.3.6 If L is Gaussian, then, as already pointed out, the sampled process is a Gaussian ARMA(r, s) process with $0 \leq s < r$. If L is non-Gaussian, the sampled process will have the same spectral density and autocovariance function as the sampled Gaussian-driven CARMA process with the same parameters. So from a second-order point of view, the two sampled process will be the same.

1.4 Objectives

Our main goal is to make inferences for Lévy-driven CARMA processes based on closely-spaced data. Besides estimating parameters of the CARMA process $\{Y(t)\}$, we would also like to identify the increments of the Lévy process and suggest an appropriate parametric model for the driving Lévy process $L(t)$. The idea we shall employ is to use results for continuously observed processes to assist in making inferences based on closely-space discrete observations.

In particular, three special families of Lévy-driven CARMA models will be discussed in this thesis. They are: (1) CARMA$(p, 0)$ driven by Gaussian processes (i.e., **continuous-time Gaussian autoregressions** or **Gaussian CAR(p)**); (2) CARMA$(1,0)$ driven by non-decreasing Lévy processes (i.e. **non-negative Lévy-driven Ornstein-Uhlenbeck (or CAR(1)) processes**); and (3) general **CARMA(p,q) ($1 \leq q < p$) driven by non-decreasing Lévy processes**.

1.4.1 Background

The problem of fitting *continuous-time Gaussian autoregressions* (linear and non-linear) to closely and regularly spaced data has been of interest for many years. For the linear case Jones (1981) and Bergstrom (1985) used state-space representations to compute exact maximum likelihood estimators and Phillips (1959) did so by fitting an appropriate discrete-time ARMA process to the data. In this thesis, we take a different point of view. We use exact conditional maximum likelihood estimators for the continuously-observed process to derive approximate maximum likelihood estimators based on the closely-spaced discrete observations. We do this for both linear and non-linear Gaussian autoregressions and obtain very clean and elegant results.

In financial econometrics, a stationary non-negative *Lévy-driven Ornstein-Uhlenbeck (or CAR(1)) process* was introduced by Barndorff-Nielsen and Shephard (2001) as a model for stochastic volatility. Their model for log asset price G had the form

$$dG(t) = (\mu + \beta V(t))dt + \sqrt{V(t)}dw(t), \tag{1.28}$$

where $w(t)$ is a standard Brownian motion. The volatility process V is an independent stationary non-negative Ornstein-Uhlenbeck process driven by non-decreasing Lévy process L,

$$V(t) = \sigma \int_{-\infty}^{t} e^{-\lambda(t-u)} dL(u), \ \lambda > 0. \tag{1.29}$$

Such a model allows for a wide variety of marginal distributions for volatility, depending on the driving Lévy process L, and allows also for the presence of jumps.

For such processes we take advantage of the non-negativity of the increments of the driving Lévy process to derive a highly efficient estimation procedure for the parameters when observations are available at uniformly spaced times. We also reconstruct the background driving Lévy process from a continuously observed realization of the process and use this result to estimate the increments of the Lévy process itself when the observation spacing is small. Asymptotic properties of the coefficient estimators are also studied.

According to Barndorff-Nielsen and Shephard's model (1.28) and (1.29), the autocorrelation function of the process V is of the form $\rho(h) = \exp(-\lambda|h|)$. This is quite restrictive for modeling purposes. However, if we replace the process V by a *CARMA process* driven by a non-negative Lévy process, we can obtain a much wider class of not-necessarily monotone autocorrelation functions for the volatility. With very flexible parameters, this class of processes introduces a huge range of possible autocorrelation functions and marginal distributions. For example, a non-negative Lévy-driven CARMA(2,1) process was used by Todorov and Tauchen (2006) and Todorov (2006) to represent stochastic volatility and applied in the latter paper to model the German Deutsche Mark/US Dollar exchange rate. For Todorov's analysis it was only the moments of the CARMA process which were relevant, so that the determination of the type of underlying Lévy process and its parameters was not considered. As to the estimation for a particular Lévy-driven CARMA(2,1) process, we propose using the maximum Gaussian likelihood estimation techniques, separating the estimated CARMA(2,1) model into two CAR(1)s according to the canonical state representation, and then using the CAR(1) which corresponds to the close-to-zero autoregressive root to finally recover the background driving Lévy process. This estimation procedure also applies to more general CARMA processes.

1.4.2 Empirical Data Description

For the empirical application of our estimation procedures for CARMA processes driven by non-decreasing Lévy processes, we shall use the daily returns on the German Deutsche Mark/US Dollar (DM/$) exchange rate series. This data was kindly provided by Viktor Todorov. It covers the period from December 1, 1986 through June 30, 1999. Missing data, weekends, fixed holiday and similar calendar effects were removed with a total of 3045 days left. Figure 1.6 shows the daily returns and realized volatility of this DM/$ exchange rate data.

This daily data originally comes from a larger data set, which is sampled at a higher frequency of 288 times per day (or 5-minute returns). The original data was explained in detail by Andersen et al. (2001). According to Andersen et al., the 288-times-per-day frequency is high enough so that our daily realized volatilities are largely free of measurement error, yet low enough so that microstructure biases (such as non-synchronous trading) are not a major concern. They constructed the daily realized volatilities by summing 288 squared 5-minute returns.

We will explore this daily return data further in Chapter 3 and Chapter 4.

1.5 Overview

The remainder of the thesis focuses on developing estimation techniques and asymptotic properties of the estimators for those three particular families of Lévy-driven CARMA processes discussed in Section 1.4. In Chapter 2, we consider estimation for Gaussian CAR(p) processes, both linear and non-linear. In order to find exact conditional maximum likelihood estimators of the parameters under the assumption that the process is observed continuously, we derive the exact conditional probability density of the $(p-1)^{st}$ derivative of an autoregression of order p with respect to Wiener measure. The resulting estimates are expressed in terms of stochastic integrals which are then approximated using the available discrete-time

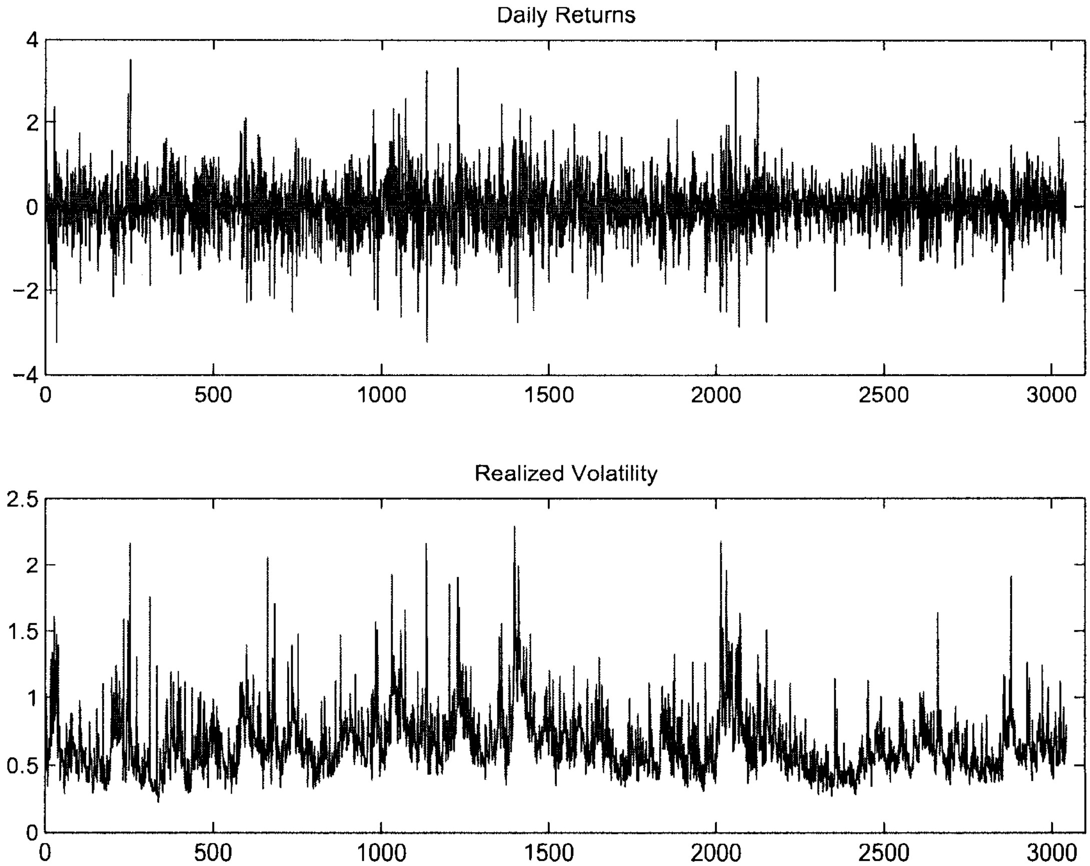


Figure 1.6: Daily returns and realized volatilities for DM/$ exchange rate data.

observations. we apply the results to (linear) CAR(p) processes, deriving explicit expressions for the maximum likelihood estimators of the coefficients and illustrating the performance of the approximations. Since the exact conditional probability density we derived does not require the CAR(p) to be linear, we also apply our results to non-linear autoregressions. In the non-linear examples considered we restrict attention to the continuous-time threshold autoregressive (CTAR) processes, which are continuous-time analogues of the discrete-time threshold models of Tong (1983).

In Chapter 3, we discuss the problem of estimating the parameters of a non-negative Lévy-driven Ornstein-Uhlenbeck process and the parameters of the background driving Lévy process, based on observations made at uniformly and closely spaced times. The idea is to obtain a highly efficient estimator of the CAR(1) coefficient by estimating the corresponding coefficient of the sampled AR(1) process using the estimator of Davis and McCormick (1989) for non-negative discrete-time AR(1) processes. This estimator is then used to estimate the corresponding realization of the driving Lévy process using a generalization of an argument due to Tuan (1977). The exact recovery of the driving Lévy process requires continuous observation of the Ornstein-Uhlenbeck process and deconvolution. The integral expressions determining the driving Lévy process are therefore replaced by approximating sums using the available discrete-time observations.

Chapter 4 is concerned with inference for a non-negative causal CARMA(p, q) process, driven by a non-decreasing Lévy process. It is assumed that observations of the CARMA process at uniformly and closely spaced times are available. The goal is to estimate both the coefficients of the CARMA process and the distribution of the increments of the driving Lévy process. Estimation of the coefficients of the CARMA process is carried out by maximizing the Gaussian likelihood of the observations. As in Chapter 3, the estimation of the distribution of the increments of the driving Lévy process is achieved using a generalization of the argument due to Tuan (1977). The performance of the procedure is illustrated with two simulated examples. We also apply our techniques to the German Deutsche Mark/US Dollar exchange rate series for empirical modeling purposes.

In the last chapter, we summarize the results and indicate future directions for research.

Chapter 2

CONTINUOUS-TIME GAUSSIAN AUTOREGRESSION

2.1 Introduction

In Section 2.2 we define the continuous-time AR(p) (abbreviated to CAR(p)) process driven by Gaussian white noise and briefly indicate the relation between the CAR(p) process $\{Y(t), t \geq 0\}$ and the sampled process $\{Y_n^{(h)} := Y(nh), n = 0, 1, 2, \ldots\}$. The process $\{Y_n^{(h)}\}$ is a discrete-time ARMA process, a result employed by Phillips (1959) to obtain maximum likelihood estimates of the parameters of the continuous-time process based on observations of $\{Y_n^{(h)}, 0 \leq nh \leq T\}$. From the state-space representation of the CAR(p) process it is also possible to express the likelihood of observations of $\{Y_n^{(h)}\}$ directly in terms of the parameters of the CAR(p) process and thereby to compute maximum likelihood estimates of the parameters as in Jones (1981) and Bergstrom (1985). For a CAR(2) process we use the asymptotic distribution of the maximum likelihood estimators of the coefficients of the ARMA process $\{Y_n^{(h)}\}$ to derive the asymptotic distribution, as first $T \rightarrow \infty$ and then $h \rightarrow 0$, of the estimators of the coefficients of the underlying CAR process.

In Section 2.3 we derive the probability density with respect to Wiener measure of the $(p-1)^{st}$ derivative of the (not-necessarily linear) autoregression of order p. This forms the basis for the inference illustrated in Sections 2.4, 2.5 and 2.6. In the non-linear examples considered we restrict attention to continuous-time threshold autoregressive (CTAR) processes, which are continuous-time analogues of the discrete-time threshold models of Tong (1983).

In Section 2.4 we apply the results to (linear) CAR(p) processes, deriving explicit expressions for the maximum likelihood estimators of the coefficients and illustrating the performance of the approximations when the results are applied to a discretely observed CAR(2) process. In Section 2.5 we consider applications to CTAR(1) and CTAR(2) processes with known threshold and in Section 2.6 we show how the technique can be adapted to include estimation of the threshold itself. The technique is also applied to the analysis of the Canadian lynx trappings, 1821 - 1934.

2.2 Gaussian CAR(p) and Corresponding Sampled Process

We begin with the definition of the (linear) Gaussian CAR(p) process.

Definition 2.2.1 (Gaussian CAR(p)). A continuous-time Gaussian autoregressive process of order $p > 0$ is defined symbolically to be a stationary solution of the stochastic differential equation,

$$a(D)Y(t) = \sigma DW(t), \tag{2.1}$$

where $a(D) = D^p + a_1 D^{p-1} + \cdots + a_p$, the operator D denotes differentiation with respect to t and $\{W(t), t \geq 0\}$ is a standard Brownian motion. Since $DW(t)$ does not exist as a random function, we give meaning to equation (2.1) by rewriting it in state-space form,

$$Y(t) = [\sigma, 0, \dots, 0]\, \mathbf{X}(t), \tag{2.2}$$

where the state vector $\mathbf{X}(t) = [X_0(t), \dots, X_{p-1}(t)]'$ satisfies the Itô equation,

$$d\mathbf{X}(t) = A\mathbf{X}(t)dt + \mathbf{e}\, dW(t), \tag{2.3}$$

with

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \cdots & -a_1 \end{bmatrix} \quad \text{and} \quad \mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

As defined in Section 1.3, $X_j(t)$ denotes the j^{th} mean-square and pathwise derivative $D^j X_0(t), j = 0, \ldots, p-1$.

We are concerned in this chapter with inference for the autoregressive coefficients, $a_1, \ldots, a_p$, based on observations of the process Y at times $0, h, 2h, \ldots, h[T/h]$, where h is small and $[x]$ denotes the integer part of x.

One approach to this problem, due to Phillips (1959), is to estimate the coefficients of the discrete-time ARMA process $\{Y_n^{(h)} := Y(nh), n = 0, 1, 2, \ldots\}$ and from these estimates to obtain estimates of the coefficients $a_1, \ldots, a_p$ in equation (2.1). The sampled process $\{Y_n^{(h)}\}$ is a stationary solution of the Gaussian ARMA(p', q') equations,

$$\phi(B) Y_n^{(h)} = \theta(B) Z_n, \quad \{Z_n\} \sim \mathrm{WN}(0, \delta^2(h)), \tag{2.4}$$

where $\phi(B)$ and $\theta(B)$ are polynomials in the backward shift operator B of orders p' and q' respectively, where $p' \leq p$ and $q' < p'$. (For more details see, e.g., Brockwell (1995).)

An alternative approach is to use equations (2.2) and (2.3) to express the likelihood of observations of $\{Y_n^{(h)}\}$ directly in terms of the parameters of the CAR(p) process and then to compute numerically the maximum likelihood estimates of the parameters as in Jones (1981) and Bergstrom (1985).

In this thesis, we take a different point of view by assuming initially that the process Y is observed continuously on the interval $[0, T]$. Under this assumption it is possible to calculate exact (conditional on $\mathbf{X}(0)$) maximum likelihood estimators of $a_1, \ldots, a_p$. To deal with the fact that observations are made only at times $0, h, 2h, \ldots$, we approximate the exact solution based on continuous observations using the available discrete-time observations. This approach has the advantage that for very closely-spaced observations it performs well and is extremely simple to implement.

This idea can be extended to non-linear (in particular threshold) continuous-time autoregressions. We illustrate this in Sections 2.4, 2.5 and 2.6. The assumption of uniform spacing, which we make in all our examples, can also be relaxed providing the maximum spacing between observations is small.

Before considering this alternative approach, we first examine the method of Phillips as applied to CAR(2) processes. This method has the advantage of requiring only the fitting of a discrete-time ARMA process to the discretely observed data and the subsequent transformation of the estimated coefficients to continuous-time equivalents. We derive the asymptotic distribution of these estimators as first $T \to \infty$ and then $h \to 0$.

Example 2.2.1 For the CAR(2) process defined by

$$(D^2 + a_1 D + a_2)Y(t) = \sigma DW(t),$$

the sampled process $\{Y_n^{(h)} := Y(nh), n = 0, 1, \ldots\}$ satisfies

$$Y_n^{(h)} - \phi_1^{(h)} Y_{n-1}^{(h)} - \phi_2^{(h)} Y_{n-2}^{(h)} = Z_n + \theta^{(h)} Z_{n-1}, \quad \{Z_t\} \sim \mathrm{WN}(0, \delta^2(h)).$$

For fixed h, as $T \to \infty$, the maximum likelihood estimator of $\boldsymbol{\beta} = \left[\phi_1^{(h)}, \phi_2^{(h)}, \theta^{(h)}\right]'$ based on observations $Y_1^{(h)}, \ldots, Y_{[T/h]}^{(h)}$ satisfies (see Brockwell and Davis (1991), p.258)

$$\sqrt{T/h}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \Rightarrow N(\mathbf{0}, M(\boldsymbol{\beta})), \tag{2.5}$$

where

$$M(\boldsymbol{\beta}) = \delta^2 \begin{bmatrix} E\mathbf{U}_t\mathbf{U}_t' & E\mathbf{V}_t\mathbf{U}_t' \\ E\mathbf{U}_t\mathbf{V}_t' & E\mathbf{V}_t\mathbf{V}_t' \end{bmatrix}^{-1}, \tag{2.6}$$

and the random vectors $\mathbf{U}_t$ and $\mathbf{V}_t$ are defined as $\mathbf{U}_t = [U_t, \cdots, U_{t+1-p}]'$ and $\mathbf{V}_t = [V_t, \cdots, V_{t+1-q}]'$, where $\{U_t\}$ and $\{V_t\}$ are stationary solutions of the autoregressive equations,

$$\phi(B)U_t = Z_t \qquad \text{and} \qquad \theta(B)V_t = Z_t. \tag{2.7}$$

In order to determine the asymptotic behavior as $T \to \infty$ of the maximum likelihood estimators $\left[\hat{\phi}_1(h), \hat{\phi}_2(h)\right]$, we consider the top left 2×2 submatrix M_2 of the matrix M. For small h we find that M_2 has the representation,

$$\begin{aligned} M_2 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} &(2a_1h + \frac{2}{\sqrt{3}}(2-\sqrt{3})a_1^2h^2 + \frac{4}{3}(2-\sqrt{3})a_1^3h^3) \\ &+ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} a_1a_2h^3 + O(h^4) \quad \text{as } h \to 0. \end{aligned} \tag{2.8}$$

The mapping from (ϕ_1, ϕ_2) to (a_1, a_2) is as follows:

$$\begin{aligned} a_1 &= -\log(-\phi_2)/h, \\ a_2 &= \frac{1}{h^2}\log\left(\frac{\phi_1}{2} + \sqrt{\frac{\phi_1^2}{4} + \phi_2}\right)\log\left(\frac{\phi_1}{2} - \sqrt{\frac{\phi_1^2}{4} + \phi_2}\right). \end{aligned}$$

The matrix

$$C = \begin{bmatrix} \frac{\partial a_1}{\partial \phi_1} & \frac{\partial a_1}{\partial \phi_2} \\ \frac{\partial a_2}{\partial \phi_1} & \frac{\partial a_2}{\partial \phi_2} \end{bmatrix}$$

therefore has the asymptotic expansion

$$C = \begin{bmatrix} 0 & \frac{1}{h}\left(1 + a_1h + \frac{a_1^2}{2}h^2 + \cdots\right) \\ -\frac{1}{h^2}\left(1 + \frac{a_1}{2}h + \frac{a_1^2+2a_2}{12}h^2 + \cdots\right) & -\frac{1}{h^2}\left(1 + \frac{a_1}{2}h + \frac{a_1^2-4a_2}{12}h^2 + \cdots\right) \end{bmatrix}. \tag{2.9}$$

From (2.8) and (2.9) we find that

$$CM_2C' = \frac{1}{h}\begin{bmatrix} 2a_1 & 0 \\ 0 & 2a_1a_2 \end{bmatrix}(1 + o(1)) \ \text{ as } h \to 0, \tag{2.10}$$

and hence, from (2.5) that the maximum likelihood estimator $\hat{\mathbf{a}}$ of $\mathbf{a} = [a_1, a_2]'$ based on observations of Y at times $0, h, 2h, \ldots, h[T/h]$, satisfies

$$\sqrt{T}(\hat{\mathbf{a}} - \mathbf{a}) \Rightarrow N(\mathbf{0}, V), \ \text{ as } T \to \infty,$$

where

$$V = \begin{bmatrix} 2a_1 & 0 \\ 0 & 2a_1a_2 \end{bmatrix}(1 + o(1)) \ \text{ as } h \to 0. \tag{2.11}$$

Remark 2.2.1 Since the moving average coefficient $\theta^{(h)}$ of the sampled process is also a function of the parameters a_1 and a_2, and hence of $\phi_1^{(h)}$ and $\phi_2^{(h)}$, the question arises as to whether the discrete-time likelihood maximization should be carried out subject to the constraint imposed by the functional relationship between $\phi_1^{(h)}, \phi_2^{(h)}$ and $\theta^{(h)}$. However, as we shall see, the unconstrained estimation which we have considered in the preceding example leads to an asymptotic distribution of the estimators which, as $h \to 0$, converges to that of the maximum likelihood estimators based on the process observed continuously on the interval $[0, T]$. This indicates, at least asymptotically, that there is no gain in using the more complicated constrained maximization of the likelihood, so that widely available standard ARMA fitting techniques can be used.

Remark 2.2.2 As the spacing h converges to zero, the autoregressive roots $\exp(-\lambda_j h)$ converge to 1, leading to numerical difficulties in carrying out the discrete-time maximization. For this reason we consider next an approach which uses exact results for the continuously observed process to develop approximate maximum likelihood estimators for closely-spaced discrete-time observations. The same approach can be used not only for linear continuous-time autoregressions, but also for non-linear autoregressions such as continuous-time analogues of the threshold models of Tong (1983).

2.3 Inference for Continuously Observed Autoregressions

We now consider a more general form of (2.1), i.e.

$$(D^p + a_1 D^{p-1} + \cdots + a_p)Y(t) = \sigma(DW(t) + c), \tag{2.12}$$

in which we allow the parameters $a_1, \ldots, a_p$ and c to be bounded measurable functions of $Y(t)$ and assume that σ is constant.

The equation (2.12) has a state-space representation analogous to (2.2) and (2.3), namely

$$Y(t) = \sigma X_0(t), \tag{2.13}$$

where

$$\begin{aligned} dX_0 &= X_1(t)dt, \\ dX_1 &= X_2(t)dt, \\ &\vdots \\ dX_{p-2} &= X_{p-1}(t)dt, \\ dX_{p-1} &= [-a_p X_0(t) - \cdots - a_1 X_{p-1}(t) + c]dt + dW(t), \end{aligned} \tag{2.14}$$

and we have abbreviated $a_i(Y(t))$ and $c(Y(t))$ to a_i and c respectively.

Assuming that $\mathbf{X}(0) = \mathbf{x}$, we can write $\mathbf{X}(t)$ according to (2.14) in terms of $\{X_{p-1}(s), 0 \le s \le t\}$ using the relations, $X_{p-2}(t) = x_{p-2} + \int_0^t X_{p-1}(s)ds$, ..., $X_0(t) = x_0 + \int_0^t X_1(s)ds$. The resulting functional relationship will be denoted by

$$\mathbf{X}(t) = \mathbf{F}(X_{p-1}, t). \tag{2.15}$$

Substituting from (2.15) into the last equation in (2.14), we see that it can be written in the form,

$$dX_{p-1} = G(X_{p-1}, t)dt + dW(t), \tag{2.16}$$

where $G(X_{p-1}, t)$, like $\mathbf{F}(X_{p-1}, t)$ depends on $\{X_{p-1}(s), 0 \le s \le t\}$.

Theorem 2.3.1 *Equation (2.14) with initial condition* $\mathbf{X}(0) = \mathbf{x} = [x_0, x_1, \cdots, x_{p-1}]'$ *has a unique (in law) weak solution* $\mathbf{X} = (\mathbf{X}(t), 0 \le t \le T)$. *The probability density of the random function* $X_{p-1} = (X_{p-1}(t), 0 \le t \le T)$ *conditioning on* $\mathbf{X}(0)$ *with respect to Wiener measure can be written as*

$$M(X_{p-1}, T) = \exp\left[-\frac{1}{2}\int_0^T G^2(X_{p-1}, s)ds + \int_0^T G(X_{p-1}, s)dW(s)\right],$$

where $G(X_{p-1}, s)$ *is defined as in (2.16).*

Proof: Let B be a standard Brownian motion (with $B(0) = x_{p-1}$) defined on the probability space $(C[0,T], \mathcal{B}[0,T], P_{x_{p-1}})$ and, for $t \leq T$, let $\mathcal{F}_t = \sigma\{B(s), s \leq t\} \vee \mathcal{N}$, where $\mathcal{N}$ is the sigma-algebra of $P_{x_{p-1}}$-null sets of $\mathcal{B}[0,T]$. The equations

$$\begin{aligned} dZ_0 &= Z_1 dt, \\ dZ_1 &= Z_2 dt, \\ &\vdots \\ dZ_{p-2} &= Z_{p-1} dt, \\ dZ_{p-1} &= dB(t), \end{aligned} \tag{2.17}$$

with $\mathbf{Z}(0) = \mathbf{x} = [x_0, x_1, \cdots, x_{p-1}]'$, clearly have the unique strong solution, $\mathbf{Z}(t) = \mathbf{F}(B,t)$, where $\mathbf{F}$ is defined as in (2.15). Let G be the functional appearing in (2.16) and suppose that $\hat{W}$ is the Itô integral defined by $\hat{W}(0) = x_{p-1}$ and

$$d\hat{W}(t) = -G(B,t)dt + dB(t) = -G(Z_{p-1},t)dt + dZ_{p-1}(t). \tag{2.18}$$

For each T, we now define a new measure $\hat{P}_{\mathbf{x}}$ on $\mathcal{F}_T$ by

$$d\hat{P}_{\mathbf{x}} = M(B,T)dP_{x_{p-1}}, \tag{2.19}$$

where

$$M(B,T) = \exp\left[-\frac{1}{2}\int_0^T G^2(B,s)ds + \int_0^T G(B,s)dW(s)\right]. \tag{2.20}$$

Then by the Cameron-Martin-Girsanov formula (see e.g. Øksendal (1998), p.152), $\{\hat{W}(t),\ 0 \leq t \leq T\}$ is a standard Brownian motion under $\hat{P}_{\mathbf{x}}$. Hence we see from (2.18) that the equations (2.16) and (2.3) with initial condition $\mathbf{X}(0) = \mathbf{x}$ have, for $t \in [0,T]$, the weak solutions $(Z_{p-1}(t), \hat{W}(t))$ and $(\mathbf{Z}(t), \hat{W}(t))$ respectively. Moreover, by Proposition 5.3.10 of Karatzas and Shreve (1991), the weak solution is unique in law, and by Theorem 10.2.2 of Stroock and Varadhan (1979) it is non-explosive.

If f is a bounded measurable functional on $C[0,T]$,

$$\begin{aligned}\hat{E}_{\mathbf{x}} f(Z_{p-1}) &= E_{x_{p-1}}(M(B,T)f(B)) \\ &= \int f(\xi)M(\xi,T)dP_{x_{p-1}}(\xi).\end{aligned}$$

In other words, $M(\xi,T)$ is the density at $\xi \in C[0,T]$, conditional on $\mathbf{X}(0) = \mathbf{x}$, of the distribution of X_{p-1} with respect to the Wiener measure $P_{x_{p-1}}$ and, if we observed $X_{p-1} = \xi$, we could compute conditional maximum likelihood estimators of the unknown parameters by maximizing $M(\xi,T)$. □

Remark 2.3.1 For parameterized functions a_i and c, this allows the possibility of maximization of the likelihood, conditional on $\mathbf{X}(0) = \mathbf{x}$, of $\{X_{p-1}(t), 0 \le t \le T\}$. Of course a complete set of observations of $\{X_{p-1}(t), 0 \le t \le T\}$ is not generally available unless X_0 is observed continuously. Nevertheless the parameter values which maximize the likelihood of $\{X_{p-1}(t), 0 \le t \le T\}$ can be expressed in terms of observations of $\{Y(t), 0 \le t \le T\}$ as described in subsequent sections. If Y is observed at discrete times, the stochastic integrals appearing in the solution for continuously observed autoregressions will be approximated by corresponding approximating sums. Other methods for dealing with the problem of estimation for continuous-time autoregressions based on discrete-time observations are considered by Stramer and Roberts (2004) and by Tsai and Chan (1999, 2000).

2.4 Estimation for CAR(p) Processes

For the CAR(p) process defined by (2.1), if $\{\mathbf{x}(s) = [x_0(s), x_1(s), \cdots, x_{p-1}(s)]', 0 \le s \le T\}$ denotes the realized state process on the interval $[0,T]$, we have, in the notation of Theorem 2.3.1,

$$-2\log M(x_{p-1}, s) = \int_0^T G^2 ds - 2\int_0^T G dx_{p-1}(s), \tag{2.21}$$

where

$$G = -a_1 x_{p-1}(s) - a_2 x_{p-2}(s) - \cdots - a_p x_0(s). \tag{2.22}$$

Differentiating $\log M$ partially with respect to $a_1, \ldots, a_p$ and setting the derivatives equal to zero gives the maximum likelihood estimators, conditional on $\mathbf{X}(0) = \mathbf{x}(0)$,

$$\begin{bmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_p \end{bmatrix} = - \begin{bmatrix} \int_0^T x_{p-1}^2 ds & \cdots & \int_0^T x_{p-1} x_0 ds \\ \vdots & \ddots & \vdots \\ \int_0^T x_{p-1} x_0 ds & \cdots & \int_0^T x_0^2 ds \end{bmatrix}^{-1} \begin{bmatrix} \int_0^T x_{p-1} dx_{p-1} \\ \vdots \\ \int_0^T x_0 dx_{p-1} \end{bmatrix}. \tag{2.23}$$

Note that this expression for the maximum likelihood estimators is unchanged if x is replaced throughout by y, where y_0 denotes the observed CAR(p) process and y_j denotes its j^{th} derivative.

Differentiating $\log M$ twice with respect to the parameters $a_1, \ldots, a_p$, taking expected values and assuming that the zeroes of the autoregressive polynomial a all have negative real parts, we find that

$$-E \frac{\partial^2 \log M}{\partial \mathbf{a}^2} \sim T\Sigma \quad \text{as} \quad T \to \infty, \tag{2.24}$$

where Σ is the covariance matrix of the limit distribution as $T \to \infty$ of the random vector $[X_{p-1}(t), X_{p-2}(t), \ldots, X_0(t)]'$. It is known (see Arató (1982)) that

$$\Sigma^{-1} = 2 \left[m_{ij} \right]_{i,j=1}^{p}, \tag{2.25}$$

where $m_{ij} = m_{ji}$ and for $j \geq i$,

$$m_{ij} = \begin{cases} 0, & \text{if } j - i \text{ is odd,} \\ \sum_{k=0}^{\infty} (-1)^k a_{i-1-k} a_{j+k}, & \text{otherwise,} \end{cases}$$

where $a_0 := 1$ and $a_j := 0$ if $j > p$ or $j < 0$, and that the estimators given by (2.23) satisfy

$$\sqrt{T}(\hat{\mathbf{a}} - \mathbf{a}) \Rightarrow N(\mathbf{0}, \Sigma^{-1}), \tag{2.26}$$

where Σ^{-1} is given by (2.25). The asymptotic result (2.26) also holds for the Yule-Walker estimates of $\mathbf{a}$ as found by Hyndman (1993).

In the case $p = 1$, $\Sigma^{-1} = 2a_1$ and when $p = 2$, Σ^{-1} is the same as the leading term in the expansion of the covariance matrix V in (2.11).

In order to derive approximate maximum likelihood estimators for closely-spaced observations of the CAR(p) process defined by (2.1) we shall use the result (2.23) with the stochastic integrals replaced by approximating sums. Thus if observations are made at times $0, h, 2h, \ldots$, we replace, for example,

$$\int_0^T x_1(s)^2 ds \quad \text{by} \quad \frac{1}{h} \sum_{i=0}^{[T/h]-1} (x((i+1)h) - x(ih))^2,$$

$$\int_0^T x_1(s) dx_1(s) \quad \text{by} \quad \frac{1}{h^2} \sum_{i=0}^{[T/h]-3} (x((i+1)h) - x(ih))$$
$$\times (x((i+3)h) - 2x((i+2)h) + x((i+1)h)),$$

taking care, as in the latter example, to preserve the non-anticipating property of the integrand in the corresponding approximating sum.

Example 2.4.1 For the CAR(2) process defined by

$$(D^2 + a_1 D + a_2) Y(t) = \sigma DW(t),$$

Table 2.1 shows the result of using approximating sums for the estimators defined by (2.23) in order to estimate the coefficients a_1 and a_2. The coefficients were estimated based on 1000 replicates on time interval $[0, T]$ of the linear CAR(2) process with $a_1 = 1.8$ and $a_2 = 0.5$.

As expected, the variances of the estimators are reduced by a factor of approximately 5 as T increases from 100 to 500 with h fixed. As h increases with T fixed, the variances actually decrease while the bias has a tendency to increase. This leads to mean squared errors which are quite close for $h = .001$ and $h = .01$. The asymptotic covariance matrix Σ^{-1} in (2.26), based on continuously observed data, is diagonal with entries 3.6 and 1.8. For $h = .001$ and $h = .01$, the variances $3.6/T$ and $1.8/T$ agree well with the corresponding entries in the table.

Table 2.1: Estimation results for Gaussian CAR(2) (linear).

h		T=100 Sample mean of estimators	T=100 Estimated variance of estimators	T=500 Sample mean of estimators	T=500 Estimated variance of estimators
0.001	a_1	1.8120	0.03585	1.7979	0.006730
	a_2	0.5405	0.02318	0.5048	0.003860
0.01	a_1	1.7864	0.03404	1.7727	0.006484
	a_2	0.5362	0.02282	0.5007	0.003799
0.1	a_1	1.5567	0.02447	1.5465	0.004781
	a_2	0.4915	0.01902	0.4588	0.003217

2.5 Estimation for CTAR(p) Processes

Definition 2.5.1 (Continuous-time Threshold CAR(p)). We define the continuous-time threshold CAR(p) (abbreviated to CTAR(p)) process with a single threshold at r exactly as in (2.13) and (2.14), except that we allow the parameters $a_1, \ldots, a_p$ and c to depend on $Y(t)$ in such a way that

$$a_i(Y(t)) = a_i^{(J)},\ i = 1, \ldots, p; \quad c(Y(t)) = c^{(J)},$$

where $J = 1$ or 2 according as $Y(t) < r$ or $Y(t) \geq r$.

Thus we obtain a continuous-time analogue of the threshold models of Tong (1983), except that Tong's model has multiple thresholds. Continuous-time threshold models have been used by a number of authors (e.g. Tong and Yeung (1991), Brockwell and Williams (1997)) for the modeling of financial and other time series. In this thesis, we shall consider only the case of a single threshold since the problems associated with more than one are quite analogous.

The density derived in Theorem 2.3.1 is not restricted to linear continuous-time autoregressions as considered in the previous section. It applies also to non-linear autoregressions and in particular to CTAR models as defined in Definition

2.5.1. In the following we illustrate the application of the continuous-time maximum likelihood estimators and corresponding approximating sums to the estimation of coefficients in CTAR(1) and CTAR(2) models.

Example 2.5.1 Consider the CTAR(1) process defined by

$$DY(t) + a_1^{(1)}Y(t) = \sigma DW(t), \quad \text{if } Y(t) < 0,$$

$$DY(t) + a_1^{(2)}Y(t) = \sigma DW(t), \quad \text{if } Y(t) \geq 0,$$

with $\sigma > 0$ and $a_1^{(1)} \neq a_1^{(2)}$. We can write

$$Y(t) = \sigma \; X(t),$$

where

$$dX(t) + a(X(t))X(t)dt = dW(t),$$

and $a(x) = a_1^{(1)}$ if $x < 0$ and $a(x) = a_1^{(2)}$ if $x \geq 0$. Proceeding as in Section 2.4, $-2\log M$ is as in (2.21) with

$$G = -a_1^{(1)}x(s)I_{x(s)<0} - a_1^{(2)}x(s)I_{x(s)\geq 0}. \tag{2.27}$$

Maximizing $\log M$ as in Section 2.4, we find that

$$\hat{a}_1^{(1)} = -\frac{\int_0^T I_{x(s)<0}x(s)dx(s)}{\int_0^T I_{x(s)<0}x(s)^2 ds}$$

and

$$\hat{a}_1^{(2)} = -\frac{\int_0^T I_{x(s)\geq 0}x(s)dx(s)}{\int_0^T I_{x(s)\geq 0}x(s)^2 ds},$$

where, as in Section 2.4, x can be replaced by y in these expressions. For observations at times $0, h, 2h, \ldots$, with h small the integrals in these expressions were replaced by corresponding approximating sums and the resulting estimates are shown in Table 2.2. This time, those estimated coefficients were obtained based on 1000 replicates on $[0, T]$ of the threshold AR(1) with threshold $r = 0$, $a_1^{(1)} = 6$, $a_1^{(2)} = 1.5$.

Table 2.2: Estimation results for CTAR(1) (non-linear, known threshold).

h		T=100		T=500	
		Sample mean of estimators	Estimated variance of estimators	Sample mean of estimators	Estimated variance of estimators
0.001	$a_1^{(1)}$	6.0450	0.41207	5.9965	0.07185
	$a_1^{(2)}$	1.5240	0.04824	1.4986	0.00891
0.01	$a_1^{(1)}$	5.8978	0.39427	5.8472	0.06785
	$a_1^{(2)}$	1.5135	0.04771	1.4875	0.00883
0.1	$a_1^{(1)}$	4.7556	0.27969	4.7085	0.04506
	$a_1^{(2)}$	1.3891	0.03840	1.3682	0.00711

Again we see that as T increases from 100 to 500, the variances of the estimators are reduced by a factor of approximately 5. As h increases with T fixed, the variances decrease while the bias tends to increase, the net effect being (as expected) an increase in mean squared error with increasing h.

Example 2.5.2 Consider the CTAR(2) process defined by

$$D^2Y(t) + a_1^{(1)}DY(t) + a_2^{(1)}Y(t) = \sigma DW(t), \quad \text{if } Y(t) < 0,$$

$$D^2Y(t) + a_1^{(2)}DY(t) + a_2^{(2)}Y(t) = \sigma DW(t), \quad \text{if } Y(t) \geq 0,$$

with $a_1^{(1)} \neq a_1^{(2)}$ or $a_2^{(1)} \neq a_2^{(2)}$, and $\sigma > 0$. We can write

$$Y(t) = [\sigma, 0]\,\mathbf{X}(t),$$

where

$$d\mathbf{X}(t) = A\mathbf{X}(t)dt + \mathbf{e}\; dW(t),$$

and $A = A^{(1)}$ if $x < 0$ and $A = A^{(2)}$ if $x \geq 0$, where

$$A^{(1)} = \begin{bmatrix} 0 & 1 \\ -a_2^{(1)} & -a_1^{(1)} \end{bmatrix}, \; A^{(2)} = \begin{bmatrix} 0 & 1 \\ -a_2^{(2)} & -a_1^{(2)} \end{bmatrix}, \; \mathbf{e} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Proceeding as in Section 2.4, $-2\log M$ is as in (2.21) with

$$G = \left(-a_1^{(1)}x_1(s) - a_2^{(1)}x(s)\right) I_{x(s)<0} + \left(-a_1^{(2)}x_1(s) - a_2^{(2)}x(s)\right) I_{x(s)\geq 0}. \qquad (2.28)$$

Maximizing $\log M$ as in Section 2.4, we find that

$$\begin{bmatrix} \hat{a}_1^{(1)} \\ \hat{a}_2^{(1)} \end{bmatrix} = - \begin{bmatrix} \int_0^T I_{x(s)<0} x_1^2(s)ds & \int_0^T I_{x(s)<0} x_1(s)x_0(s)ds \\ \int_0^T I_{x(s)<0} x_1(s)x_0(s)ds & \int_0^T I_{x(s)<0} x_0^2(s)ds \end{bmatrix}^{-1} \begin{bmatrix} \int_0^T I_{x(s)<0} x_1(s)dx_1(s) \\ \int_0^T I_{x(s)<0} x_0(s)dx_1(s) \end{bmatrix},$$

while $[\hat{a}_1^{(2)}, \hat{a}_2^{(2)}]'$ satisfies the same equation with $I_{x(s)<0}$ replaced throughout by $I_{x(s)\geq 0}$.

As in Section 2.4, x can be replaced by y in these expressions. For observations at times $0, h, 2h, \ldots$, with h small, the integrals in these expressions were replaced by corresponding approximating sums and the resulting estimates are shown in Table 2.3. 1000 replicates on $[0, T]$ of the threshold AR(2) with threshold $r = 0$, $a_1^{(1)} = 1.5$, $a_2^{(1)} = 0.4$, $a_1^{(2)} = 4.6$, $a_2^{(2)} = 2$, were used to estimate those coefficients.

Table 2.3: Estimation results for CTAR(2) (non-linear, known threshold).

h		T=100		T=500	
		Sample mean of estimators	Estimated variance of estimators	Sample mean of estimators	Estimated variance of estimators
0.001	$a_1^{(1)}$	1.5187	0.05441	1.5071	0.01128
	$a_2^{(1)}$	0.4763	0.04119	0.4163	0.00480
	$a_1^{(2)}$	4.6084	0.21224	4.5755	0.03995
	$a_2^{(2)}$	2.3186	0.72069	2.0456	0.08881
0.01	$a_1^{(1)}$	1.5262	0.05234	1.5163	0.01095
	$a_2^{(1)}$	0.4729	0.04056	0.4135	0.00473
	$a_1^{(2)}$	4.3819	0.19823	4.3480	0.03746
	$a_2^{(2)}$	2.2697	0.68915	2.0025	0.08509
0.1	$a_1^{(1)}$	1.5091	0.04177	1.4928	0.00805
	$a_2^{(1)}$	0.4402	0.03489	0.3851	0.00411
	$a_1^{(2)}$	2.7053	0.11312	2.7014	0.01874
	$a_2^{(2)}$	1.7654	0.41380	1.5599	0.05221

The pattern of results is more complicated in this case. As T is increased from 100 to 500 with h fixed, the sample variances all decrease, but in a less regular fashion than in Tables 2.1 and 2.2. As h increases with T fixed, the variances also decrease. The mean squared errors for $h = .001$ and $h = .01$ are again quite close.

2.6 Estimation When the Threshold is Unknown

In the previous section we considered the estimation of the autoregressive coefficients only, under the assumption that the threshold r is known. In this section we consider the corresponding problem when the threshold also is to be estimated. The idea is the same, that is to maximize the (conditional) likelihood of the continuously-observed process, using the closely-spaced discrete observations to approximate what would be the exact maximum likelihood estimators if the continuously-observed data were available. We illustrate first with a CTAR(1) process. The goal is to use observations $\{y(kh), k = 1, 2, \ldots; 0 < kh \le T\}$, with h small, to estimate the parameters $a_1^{(1)}, a_1^{(2)}, c_1^{(1)}, c_1^{(2)}, \sigma$ and r in the following model.

$$\begin{aligned} D\left(Y(t) - r\right) + a_1^{(1)}\left(Y(t) - r\right) + c_1^{(1)} &= \sigma DW_t, \qquad \text{if } Y(t) < r, \\ D\left(Y(t) - r\right) + a_1^{(2)}\left(Y(t) - r\right) + c_1^{(2)} &= \sigma DW_t, \qquad \text{if } Y(t) \ge r. \end{aligned} \tag{2.29}$$

The process $Y^* = Y - r$, satisfies the threshold autoregressive equations,

$$\begin{aligned} DY^*(t) + a_1^{(1)} Y^*(t) + c_1^{(1)} &= \sigma DW_t, \qquad \text{if } Y^*(t) < 0, \\ DY^*(t) + a_1^{(2)} Y^*(t) + c_1^{(2)} &= \sigma DW_t, \qquad \text{if } Y^*(t) \ge 0, \end{aligned}$$

with state-space representation,

$$Y^*(t) = \sigma X(t),$$

where

$$dX(t) = G(X, t)dt + dW(t),$$

as in equation (2.16), and

$$G(x, s) = -\left(a_1^{(1)} x(s) + \frac{c_1^{(1)}}{\sigma}\right) I_{x(s)<0} - \left(a_1^{(2)} x(s) + \frac{c_1^{(2)}}{\sigma}\right) I_{x(s)\ge 0}.$$

Substituting for G in the expression (2.21), we obtain

$$-2\log M(x(s), s) = \int_0^T G^2 ds - 2\int_0^T G dx(s)$$

$$
\begin{aligned}
&= \int_0^T \left(a_1^{(1)} x(s) + \frac{c_1^{(1)}}{\sigma} \right)^2 I_{x(s)<0} ds + \int_0^T \left(a_1^{(2)} x(s) + \frac{c_1^{(2)}}{\sigma} \right)^2 I_{x(s)\geq 0} ds \\
&\quad +2 \int_0^T \left(a_1^{(1)} x(s) + \frac{c_1^{(1)}}{\sigma} \right) I_{x(s)<0} dx(s) + 2 \int_0^T \left(a_1^{(2)} x(s) + \frac{c_1^{(2)}}{\sigma} \right) I_{x(s)\geq 0} dx(s) \\
&= \frac{1}{\sigma^2} \left[\int_0^T \left(a_1^{(1)} y^* + c_1^{(1)} \right)^2 I_{y^*<0} ds + \int_0^T \left(a_1^{(2)} y^* + c_1^{(2)} \right)^2 I_{y^*\geq 0} ds \right. \\
&\quad \left. +2 \int_0^T \left(a_1^{(1)} y^* + c_1^{(1)} \right) I_{y^*<0} dy^* + 2 \int_0^T \left(a_1^{(2)} y^* + c_1^{(2)} \right) I_{y^*\geq 0} dy^* \right].
\end{aligned}
$$

Minimizing $-2 \log M(x(s), s)$ with respect to $a_1^{(1)}$, $a_1^{(2)}$, $c_1^{(1)}$, and $c_1^{(2)}$ with σ fixed gives,

$$
\begin{aligned}
&\hat{a}_1^{(1)}(r) \left[\int_0^T y^{*2} I_{y^*<0} ds \int_0^T I_{y^*<0} ds - \left(\int_0^T y^* I_{y^*<0} ds \right)^2 \right] \\
&\quad = - \left[\int_0^T y^* I_{y^*<0} dy^* \int_0^T I_{y^*<0} ds - \int_0^T I_{y^*<0} dy^* \int_0^T y^* I_{y^*<0} ds \right], \\
&\hat{c}_1^{(1)}(r) \left[\int_0^T y^{*2} I_{y^*<0} ds \int_0^T I_{y^*<0} ds - \left(\int_0^T y^* I_{y^*<0} ds \right)^2 \right] \\
&\quad = - \left[\int_0^T I_{y^*<0} dy^* \int_0^T y^{*2} I_{y^*<0} ds - \int_0^T y^* I_{y^*<0} dy^* \int_0^T y^* I_{y^*<0} ds \right],
\end{aligned}
\tag{2.30}
$$

with analogous expressions for $\hat{a}_1^{(2)}$ and $\hat{c}_1^{(2)}$. An important feature of these equations is that they involve only the values of $y^* = y - r$ and not σ.

For any fixed value of r and observations y, we can therefore compute the maximum likelihood estimators $\hat{a}_1^{(1)}(r)$, $\hat{a}_1^{(2)}(r)$, $\hat{c}_1^{(1)}(r)$ and $\hat{c}_1^{(2)}(r)$ and the corresponding minimum value, $m(r)$, of $-2\sigma^2 \log M$. The maximum likelihood estimator $\hat{r}$ of r is the value which minimizes $m(r)$ (this minimizing value also being independent of σ). The maximum likelihood estimators of $a_1^{(1)}, a_1^{(2)}, c_1^{(1)}$ and $c_1^{(2)}$ are the values obtained from (2.30) with $r = \hat{r}$. Since the observed data are the discrete observations $\{y(h), y(2h), y(3h), \ldots\}$, the calculations just described are all carried out with the integrals in (2.30) replaced by approximating sums as described in Section 2.4.

If the data y are observed continuously, the quadratic variation of y on the interval $[0, T]$ is *exactly* equal to $\sigma^2 T$. The discrete approximation to σ based on

$\{y(h), y(2h), \ldots\}$ is

$$\hat{\sigma} = \sqrt{\sum_{k=1}^{[T/h]-1} (y((k+1)h) - y(kh))^2/T}. \tag{2.31}$$

Example 2.6.1 Table 2.4 shows the results obtained when the foregoing estimation procedure is applied to a CTAR(1) process defined by (2.29) with $a_1^{(1)} = 6$, $c_1^{(1)} = .5$, $a_1^{(2)} = 1.5$, $c_1^{(2)} = .4$, $\sigma = 1$ and $r = 10$. In this table, estimations were made based on 1000 replicates of the process on $[0, T]$.

Table 2.4: Estimation results for CTAR(1) (non-linear, unknown threshold).

h		T=100 Sample mean	T=100 Sample variance	T=500 Sample mean	T=500 Sample variance	T=1000 Sample mean	T=1000 Sample variance
0.001	$a_1^{(1)}$	5.9179	1.5707	5.9758	0.1950	5.9835	0.0904
	$c_1^{(1)}$	0.3787	0.7780	0.3448	0.1561	0.3832	0.0753
	$a_1^{(2)}$	1.7149	0.4105	1.5370	0.0511	1.5178	0.0224
	$c_1^{(2)}$	0.2891	0.1476	0.3415	0.0273	0.3601	0.0133
	σ	0.9996	5.00×10^{-6}	0.9991	4.78×10^{-7}	0.9991	4.84×10^{-7}
	r	9.9963	0.0244	9.9769	0.0041	9.9818	0.0020
0.01	$a_1^{(1)}$	5.7201	1.3175	5.7507	0.1834	5.7614	0.0910
	$c_1^{(1)}$	0.4271	0.7524	0.3235	0.1699	0.3535	0.0705
	$a_1^{(2)}$	1.7373	0.4248	1.5567	0.0538	1.5360	0.0239
	$c_1^{(2)}$	0.2877	0.1598	0.3227	0.0357	0.3407	0.0162
	σ	0.9914	4.82×10^{-5}	0.9913	4.55×10^{-6}	0.9907	4.91×10^{-6}
	r	10.011	0.0278	9.9807	0.0058	9.984	0.0024
0.1	$a_1^{(1)}$	4.1166	0.7861	4.1087	0.1587	4.1115	0.0720
	$c_1^{(1)}$	0.3953	0.5638	0.2834	0.2287	0.2708	0.0944
	$a_1^{(2)}$	1.7308	0.5109	1.5924	0.0805	1.5851	0.0324
	$c_1^{(2)}$	0.2636	0.2391	0.2658	0.1003	0.2666	0.0472
	σ	0.9191	5.00×10^{-4}	0.9208	4.60×10^{-5}	0.9160	4.79×10^{-5}
	r	10.074	0.0425	10.038	0.0191	10.030	0.0086

The pattern of results is again rather complicated. As expected however there is a clear reduction in sample variance of the estimators as T is increased with h fixed. For $T = 1000$ the mean squared errors of the estimators all increase as h

increases, with the mean squared errors when $h = .001$ and $h = .01$ being rather close and substantially better than those when $h = .1$.

Example 2.6.2 Although the procedure described above is primarily intended for use in the modeling of very closely-spaced data, in this example we illustrate its performance when applied to the natural logarithms of the annual Canadian lynx trappings, 1821 - 1934 (see e.g. Brockwell and Davis (1991), p.559). Linear and threshold autoregressions of order two were fitted to this series by Tong and Yeung (1991) and a linear CAR(2) model using a continuous-time version of the Yule-Walker equations by Hyndman (1993).

The threshold AR(2) model fitted by Tong and Yeung (1991) to this series was

$$\begin{aligned} D^2Y(t) \ + \ a_1^{(1)}DY(t) + \ a_2^{(1)}Y(t) &= \sigma_1 DW(t), \qquad \text{if } Y(t) < r, \\ D^2Y(t) \ + \ a_1^{(2)}DY(t) + \ a_2^{(2)}Y(t) &= \sigma_2 DW(t), \qquad \text{if } Y(t) \geq r, \end{aligned} \tag{2.32}$$

with

$$\begin{aligned} a_1^{(1)} = 0.354, \quad a_1^{(2)} = 0.521, \quad \sigma_1 = 0.707, \\ a_1^{(2)} = 1.877, \quad a_2^{(2)} = 0.247, \quad \sigma_2 = 0.870, \end{aligned} \tag{2.33}$$

and threshold $r = 0.857$.

An argument exactly parallel to that for the CTAR(1) process at the beginning of this section permits the estimation of the coefficients and threshold of a CTAR(2) model of this form with $\sigma_1 = \sigma_2 = \sigma$, $h = 1$ and with time measured in years. It leads to the coefficient estimates,

$$\begin{aligned} a_1^{(1)} = 0.3163, \quad a_1^{(2)} = 0.1932, \quad \sigma_1 = 1.150, \\ a_1^{(2)} = 1.2215, \quad a_2^{(2)} = 0.9471, \quad \sigma_2 = 1.150, \end{aligned} \tag{2.34}$$

with estimated threshold $r = 0.478$. (Because of the large spacing of the observations in this case it is difficult to obtain a good approximation to the quadratic variation of the derivative of the process. The coefficient σ was therefore estimated by a simple one-dimensional maximization of the Gaussian likelihood (GL) of the original discrete observations (computed as described by Brockwell (2001a)), with the estimated coefficients fixed at the values specified above.)

In terms of the Gaussian likelihood of the original data, the latter model (with $-2\log(GL) = 220.15$) is considerably better than the Tong and Yeung model (for which $-2\log(GL) = 244.41$). Using our model as an initial approximation for maximizing the Gaussian likelihood of the original data, we obtain the following more general model, which has higher Gaussian likelihood than both of the preceding models ($-2\log(GL) = 161.06$),

$$\begin{aligned} D^2Y(t) \ &+\ 1.1810DY(t) + 0.308Y(t) - 0.345 = 1.050DW(t), \quad \text{if } Y(t) < -0.522, \\ D^2Y(t) \ &+\ 0.0715DY(t) + 0.452Y(t) + 0.500 = 0.645DW(t), \quad \text{if } Y(t) \geq -0.522. \end{aligned} \tag{2.35}$$

Simulations of the model (2.32) with parameters as in (2.33) and (2.34) and of the model (2.35) are shown together with the logged and mean-corrected lynx data in Figure 2.1. Inside Figure 2.1, Figures (a) and (b) show simulations of the CTAR model (2.32) for the logged and mean-corrected lynx data when the parameters are given by (2.33) and (2.34) respectively. Figure (c) shows a simulation (with the same driving noise as in Figures (a) and (b)) of the model (2.35). Figure (d) show the logged and mean-corrected lynx series itself. As expected, the resemblance between the sample paths and the data appears to improve with increasing Gaussian likelihood.

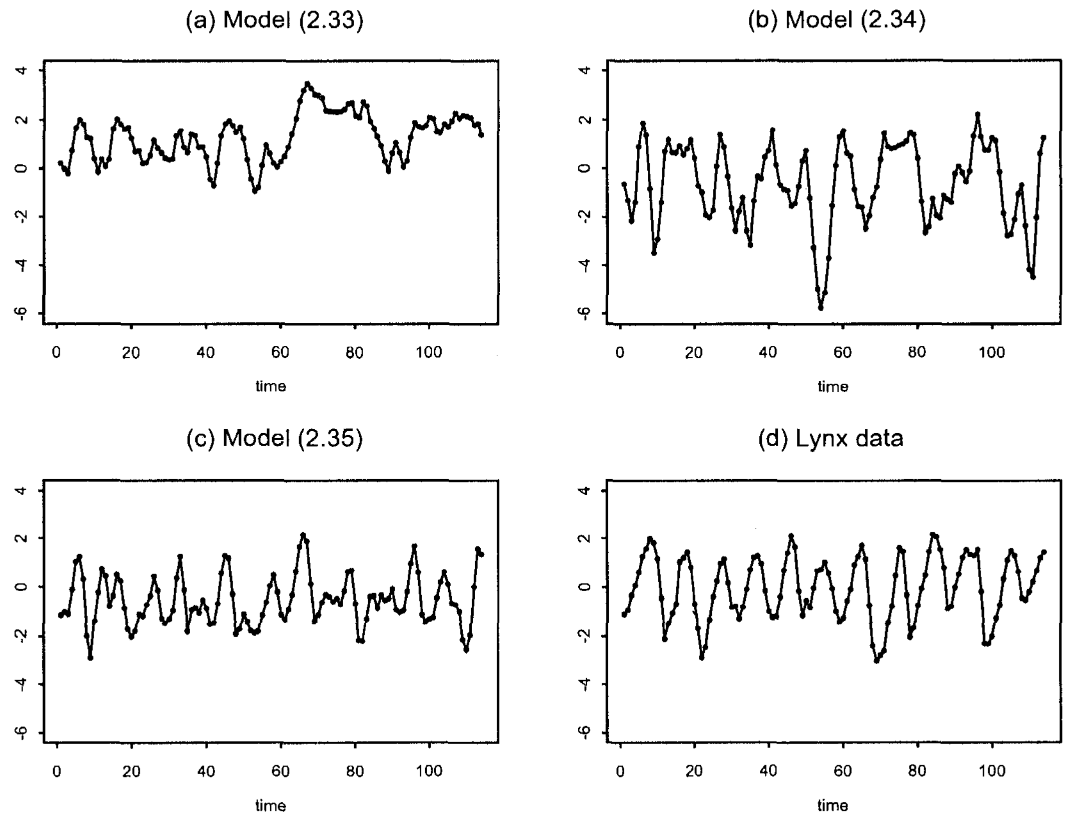


Figure 2.1: Canadian lynx trapping data analysis using CTAR(2) model.

Chapter 3

NON-NEGATIVE LÉVY-DRIVEN ORNSTEIN-UHLENBECK PROCESSES

3.1 Introduction

In Section 3.2, we define the stationary Lévy-driven Ornstein-Uhlenbeck (or CAR(1)) process, $\{Y(t), t \geq 0\}$. In Section 3.3, we characterize the sampled AR(1) process, $\{Y_n^{(h)} = Y(nh), n = 0, 1, 2, \ldots\}$, and the distribution of its driving white noise sequence in terms of the parameters of the underlying CAR(1) process and its driving Lévy process. The autoregressive coefficient of the sampled process is then estimated with very high efficiency using the method of Davis and McCormick (1989). From the relation between the sampled and continuous-time processes we then obtain corresponding parameter estimates for the CAR(1) process. The idea of using the sampled process to estimate the parameters of the underlying continuous-time process was first used by Phillips (1959), but in our case the non-decreasing property of the driving Lévy process and the non-negativity of the corresponding discrete-time increments permits a very large efficiency gain. In Section 3.4, we show how to recover the driving Lévy process under the assumption that the process is observed continuously and then approximate the results using closely-spaced discrete observations. In Section 3.5, we derive the asymptotic distribution of the coefficient estimator when the driving Lévy process is a gamma process and illustrate with a simulated example the performance of the estimators of both the CAR(1) parameters and the driving Lévy process. When the continuously observed process is available, the autoregression coefficient can be identified with probability 1. This is discussed in Section 3.7.

3.2 Stationary Lévy-driven Ornstein-Uhlenbeck Processes

Throughout this chapter, we shall be concerned with CAR(1) (or stationary Ornstein-Uhlenbeck) processes driven by standardized second-order non-decreasing Lévy processes, which are defined in details in Section 1.2.

An interesting fact about any non-decreasing Lévy process is that its Laplace transform $\tilde{f}_{L(t)}(s) := E(\exp(-sL(t)))$ has the form

$$\tilde{f}_{L(t)}(s) = \exp(-t\Phi(s)), \ \Re(s) \geq 0,$$

where

$$\Phi(s) = m + \int_{(0,\infty)} (1 - e^{-sx})\nu(dx),$$

with the drift term $m \geq 0$ and the Lévy measure ν on the Borel subsets of $(0, \infty)$ satisfying

$$\int_{(0,\infty)} \frac{u}{1+u}\nu(du) < \infty.$$

Then, based on Definition 1.3.1, the Lévy-driven CAR(1) process is defined as follows.

Definition 3.2.1 (Lévy-driven CAR(1) process). A CAR(1) process driven by the Lévy process $\{L(t), t \geq 0\}$ is defined to be a strictly stationary solution of the stochastic differential equation,

$$dY(t) + aY(t)dt = \sigma dL(t). \tag{3.1}$$

In the special case when $\{L(t)\}$ is a Brownian motion, (3.1) is interpreted as an Itô equation with solution $\{Y(t), t \geq 0\}$ satisfying

$$Y(t) = e^{-at}Y(0) + \sigma \int_0^t e^{-a(t-u)}dL(u), \tag{3.2}$$

where the integral is defined as the L^2 limit of approximating Riemann-Stieltjes sums. For any second-order driving Lévy process, $\{L(t)\}$, the process $\{Y(t)\}$ can be defined in the same way, and if $\{L(t)\}$ is non-decreasing (and hence of bounded variation on compact intervals) $\{Y(t)\}$ can also be defined pathwise as a Riemann-Stieltjes integral by (3.2). We can also write

$$Y(t) = e^{-a(t-s)}Y(s) + \sigma \int_s^t e^{-a(t-u)} dL(u), \text{ for all } t > s \geq 0, \tag{3.3}$$

showing, by independence of the increments of $\{L(t)\}$, that $\{Y(t)\}$ is Markov. The following proposition implied by Proposition 1.3.1 gives necessary and sufficient conditions for stationarity of $\{Y(t)\}$.

Proposition 3.2.1 *If $Y(0)$ is independent of $\{L(t), t \geq 0\}$ and $E(L(1)^2) < \infty$, then $Y(t)$ is strictly stationary if and only if $a > 0$ and $Y(0)$ has the distribution of $\sigma \int_0^\infty e^{-au} dL(u)$.*

Remark 3.2.1 By introducing a second Lévy process $\{M(t), 0 \leq t < \infty\}$, independent of L and with the same distribution as shown in Remark 1.3.2, we can extend $\{Y(t), t \geq 0\}$ to a process with index set $(-\infty, \infty)$. Then, provided $a > 0$, the process $\{Y(t)\}$ defined by

$$Y(t) = \sigma \int_{-\infty}^t e^{-a(t-u)} dL^*(u), \tag{3.4}$$

is a strictly stationary process satisfying equation (3.3) (with L replaced by L^*) for all $t > s$ and $s \in (-\infty, \infty)$. Henceforth, we refer to L^* as the *background driving Lévy process* (BDLP) and denote it by L for simplicity.

Remark 3.2.2 From (3.4) we have the relation

$$Y(t) = e^{-a(t-s)}Y(s) + \sigma \int_s^t e^{-a(t-u)} dL(u), \quad t \geq s > -\infty. \tag{3.5}$$

Taking $s = 0$ and using Lemma 2.1 of Eberlein and Raible (1999), we find that

$$Y(t) = e^{-at}Y(0) + \sigma L(t) - a\sigma \int_0^t e^{-a(t-u)} L(u)du, \; t \geq 0, \tag{3.6}$$

where the last integral is a Riemann integral and the equality holds for all finite $t \geq 0$ with probability 1.

3.3 Parameter Estimation via the Sampled Process

Setting $t = nh$ and $s = (n-1)h$ in equation (3.5), we see at once that for any $h > 0$, the sampled process $\{Y_n^{(h)}, n = 0, 1, 2, \ldots\}$ is the discrete-time AR(1) process satisfying

$$Y_n^{(h)} = \phi Y_{n-1}^{(h)} + Z_n, \; n = 0, 1, 2, \ldots, \tag{3.7}$$

where

$$\phi = e^{-ah}, \tag{3.8}$$

and

$$Z_n = \sigma \int_{(n-1)h}^{nh} e^{-a(nh-u)} dL(u). \tag{3.9}$$

The noise sequence $\{Z_n\}$ is iid and positive since L has stationary, independent and positive increments.

If the process $\{Y(t), 0 \leq t \leq T\}$ is observed at times $0, h, 2h, \ldots, Nh$, where $N = [T/h]$, i.e., N is the integer part of T/h, then, since the innovations Z_n of the process $\{Y_n^{(h)}\}$ are non-negative and $0 < \phi < 1$, we can use the highly efficient Davis-McCormick estimator of ϕ, namely

$$\hat{\phi}_N^{(h)} = \min_{1 \leq n \leq N} \frac{Y_n^{(h)}}{Y_{n-1}^{(h)}}. \tag{3.10}$$

To obtain the asymptotic distribution of $\hat{\phi}_N^{(h)}$ as $N \to \infty$ with h fixed, we need to suppose that the distribution function F of Z_n satisfies $F(0) = 0$ and that F is regularly varying at zero with exponent α, i.e., that there exists $\alpha > 0$ such that

$$\lim_{t \downarrow 0} \frac{F(tx)}{F(t)} = x^{\alpha} \quad \text{for all } x > 0.$$

(These conditions are satisfied by the gamma-driven CAR(1) process as we shall show in Section 3.5.) Under these conditions on F, the results of Davis and McCormick (1989) imply that $\hat{\phi}_N^{(h)} \to \phi$ *a.s.* as $N \to \infty$ with h fixed and that

$$\lim_{N\to\infty} P\left[k_N^{-1}(\hat{\phi}_N^{(h)} - \phi)c_\alpha \leq x\right] = G_\alpha(x), \tag{3.11}$$

where $k_N = F^{-1}(N^{-1})$, $c_\alpha = (EY_1^{(h)\alpha})^{1/\alpha}$ and G_α is the Weibull distribution function,

$$G_\alpha(x) = \begin{cases} 1 - \exp\{-x^\alpha\}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases} \tag{3.12}$$

From the observations $\{Y_n^{(h)}, n = 0, 1 \ldots, N\}$ we thus obtain the estimator $\hat{\phi}_N^{(h)}$ and, from (3.8), the corresponding estimator,

$$\hat{a}_N^{(h)} = -h^{-1} \log \hat{\phi}_N^{(h)} \tag{3.13}$$

of the CAR(1) coefficient a. Provided the distribution function F of the noise terms Z_n in the discrete-time sampled process satisfies the conditions indicated above, we can also determine the asymptotic distributions of this estimator. In particular, using a Taylor series approximation, we find that

$$\lim_{N\to\infty} P\left[(-h)e^{-ah}c_\alpha k_N^{-1}\left(\hat{a}_N^{(h)} - a\right) \leq x\right] = G_\alpha(x), \tag{3.14}$$

where G_α is given in (3.12). Since $\text{var}(Y^{(h)}) = \sigma^2/(2a)$, we use the estimator,

$$\hat{\sigma}_N^2 = \frac{2\hat{a}_N^{(h)}}{N} \sum_{i=0}^{N} (Y_i^{(h)} - \overline{Y}_N^{(h)})^2 \tag{3.15}$$

to estimate σ^2.

3.4 Estimating the Lévy Increments

So far, we have made no assumptions about the driving Lévy process except for non-negativity and existence of $EL(1)^2$. In order to suggest an appropriate parametric model for L and to estimate the parameters, it is important to recover an approximation to L from the observed data. If the CAR(1) process is continuously observed on $[0, T]$, then the argument of Tuan (1977) can be used to recover $\{L(t), 0 \leq t \leq T\}$. His L^2-based spectral argument which he applied to Gaussian processes, also applies to the Lévy-driven CAR(1) process to give

$$L(t) = \sigma^{-1}\left[Y(t) - Y(0) + a\int_0^t Y(s)ds\right]. \tag{3.16}$$

A direct justification of this result can be obtained by *defining* L as in (3.16) and then showing that $Y(0)e^{-at} + \sigma\int_0^t e^{-a(t-u)}dL(u) = Y(t)$. Thus, using Remark 3.2.2 of Section 3.2, we have

$$\begin{aligned}
&Y(0)e^{-at} + \sigma\int_0^t e^{-a(t-u)}dL(u)\\
=&Y(0)e^{-at} + \sigma L(t) - a\sigma\int_0^t e^{-a(t-u)}L(u)du\\
=&Y(0)e^{-at} + \left[Y(t) - Y(0) + a\int_0^t Y(s)ds\right] - a\int_0^t e^{-a(t-u)}\left[Y(u) - Y(0) + a\int_0^u Y(s)ds\right]du\\
=&Y(t) + a\int_0^t Y(s)ds - a\int_0^t e^{-a(t-u)}Y(u)du - a^2\int_0^t Y(s)\int_s^t e^{-a(t-u)}du\,ds\\
=&Y(t)
\end{aligned}$$

as required.

From (3.16), the increment of the driving Lévy process on the interval $((n-1)h, nh]$ is given by

$$\triangle\, L_n^{(h)} := L(nh) - L((n-1)h) = \sigma^{-1}\left[Y(nh) - Y((n-1)h) + a\int_{(n-1)h}^{nh} Y(u)du\right].$$

Replacing the CAR(1) parameters by their estimators and the integral by a trapezoidal approximation, we obtain the estimated increments,

$$\triangle\, \hat{L}_n^{(h)} = \hat{\sigma}_N^{-1}\left[Y_n^{(h)} - Y_{n-1}^{(h)} + \hat{a}_N^{(h)}h(Y_n^{(h)} + Y_{n-1}^{(h)})/2\right]. \tag{3.17}$$

3.5 Gamma-driven CAR(1) Process

In this section, we illustrate the preceding estimating procedure in the case when L is a standardized gamma process. Thus $L(t)$ has the gamma density $f_{L(t)}$ with exponent γt, scale-parameter $\gamma^{-1/2}$, mean $\gamma^{1/2}t$ and variance t. The Laplace transform of $L(t)$ is

$$\tilde{f}_{L(t)}(s) := E\exp(-sL(t)) = \exp\{-t\Phi(s)\}, \ \Re(s) \geq 0, \tag{3.18}$$

where $\Phi(s) = \gamma\log(1+\beta s)$, $\beta = \gamma^{-1/2}$ and $\gamma > 0$.

Based on the h-spaced observations $\{Y_n^{(h)}, n = 0, 1, \ldots, N\}$, we estimate the discrete-time autoregression coefficient ϕ and the CAR(1) parameters a and σ^2 using (3.10), (3.13) and (3.15) respectively. We then estimate the Lévy increments as in (3.17) and use them to estimate the parameter γ of the standardized gamma process L. To obtain the asymptotic distributions of $\phi_N^{(h)}$ and $\hat{a}_N^{(h)}$ as $N \to \infty$ with h fixed, we first show that the distribution function F of Z_n in (3.7) is regularly varying at zero with exponent γh and then determine the coefficients $k_N = F^{-1}(N^{-1})$ and $c_\alpha = (EY_1^{(h)\alpha})^{1/\alpha}$ in (3.11). To do so, we use the Laplace transform (3.18) to investigate the behavior of the density of $Z_1 = \sigma\int_0^h e^{-a(h-t)}dL(t) = \sigma\int_0^h e^{-at}dL(t)$ near zero.

Define $W_h := Z_1/\sigma = \int_0^h e^{-at}dL(t)$. The Laplace transform of W_h is

$$\begin{aligned}\tilde{f}_{W_h}(s) &= \exp\left[-\int_0^h \Phi(se^{-at})dt\right] \\ &= \exp\left[-\int_0^h \gamma\log(1+\beta se^{-at})dt\right].\end{aligned} \tag{3.19}$$

The exponent in (3.19) has the power series expansion,

$$\begin{aligned}-\int_0^h \gamma\log(1+\beta se^{-at})dt &= -\gamma\int_0^h \log\left[s\beta e^{-at}(1+\frac{1}{\beta se^{-at}})\right]dt \\ \approx \log(s\beta)^{-\gamma h} + \frac{1}{2}\gamma ah^2 &+ \gamma\left[\frac{1}{\beta sa}(1-e^{ah}) - \frac{1}{4\beta^2 s^2 a}(1-e^{2ah}) + \cdots\right],\end{aligned}$$

as $s \to \infty$. Hence $\tilde{f}_{W_h}(s)$ has the corresponding expansion,

$$\tilde{f}_{W_h}(s) \approx \frac{\beta^{-\gamma h}}{s^{\gamma h}} e^{\frac{1}{2}\gamma a h^2} + \frac{C_1}{s^{\gamma h+1}} + \frac{C_2}{s^{\gamma h+2}} + \cdots,$$

where $C_1, C_2, \ldots$ are constants depending on γ, β, h and a. Since $\tilde{f}_{Z_1}(s) = \tilde{f}_{W_h}(\sigma s)$,

$$s^{\gamma h} \tilde{f}_{Z_1}(s) \to (\sigma\beta)^{-\gamma h} e^{\frac{1}{2}\gamma a h^2}, \text{ as } s \to \infty.$$

By Theorem 30.2 of Doetsch (1974), the density f_{Z_1} of Z_1 has the expansion, in a neighborhood of zero,

$$f_{Z_1}(x) = \frac{(\sigma\beta)^{-\gamma h} x^{\gamma h - 1}}{\Gamma(\gamma h)} e^{\frac{1}{2}\gamma a h^2} + \frac{(\sigma x)^{\gamma h} C_1}{\sigma \Gamma(\gamma h + 1)} + \frac{(\sigma x)^{\gamma h+1} C_2}{\sigma \Gamma(\gamma h + 2)} + \cdots.$$

So

$$\frac{f_{Z_1}(x)}{x^{\gamma h - 1}} \to (\sigma\beta)^{-\gamma h} e^{\frac{1}{2}\gamma a h^2} / \Gamma(\gamma h), \text{ as } x \to 0,$$

and

$$F_{Z_1}(x) \sim x^{\gamma h} (\sigma\beta)^{-\gamma h} e^{\frac{1}{2}\gamma a h^2} / \Gamma(\gamma h + 1), \text{ as } x \to 0. \tag{3.20}$$

Thus the distribution F of Z_n is regularly varying at zero with exponent γh.

From the definition of k_N in (3.11), we have $\frac{1}{N} = \int_0^{k_N} F_{Z_1}(du)$. This equation, together with (3.20), gives

$$k_N^{-1} \sim (\sigma\beta)^{-1} \left[\Gamma(\gamma h + 1)\right]^{-1/(\gamma h)} e^{\frac{1}{2}ah} N^{1/(\gamma h)}, \text{ as } N \to \infty. \tag{3.21}$$

In order to calculate $c_{\gamma h}$, we need to find $E[Y_n^{(h)}]^{\gamma h}$, where $Y_n^{(h)} = \sum_{j=0}^{\infty} \phi^j Z_{n-j}$. The Laplace transform of $Y_n^{(h)}$ is

$$\tilde{f}_{Y_n^{(h)}}(s) = E\, e^{-sY_n^{(h)}} = \prod_{j=0}^{\infty} E\, e^{-s\phi^j Z_{n-j}}.$$

So

$$\begin{aligned}
\log \tilde{f}_{Y_n^{(h)}}(s) &= \sum_{j=0}^{\infty} \log \tilde{f}_{Z_1}(s\phi^j) \\
&= \sum_{j=0}^{\infty} \log \tilde{f}_{W_h}(s\sigma\phi^j) \\
&= -\gamma \sum_{j=0}^{\infty} \int_0^h \log(1 + \beta s \sigma \phi^j e^{-ay}) dy,
\end{aligned}$$

and hence

$$\begin{aligned}\tilde{f}_{Y_n^{(h)}}(s) &= \exp\left[\frac{\gamma}{a}\sum_{j=0}^{\infty}\left[\text{dilog}(1+\beta s\sigma\phi^j) - \text{dilog}(1+\beta s\sigma\phi^j e^{-ah})\right]\right] \\ &= \exp\left(\frac{\gamma}{a}\text{dilog}(1+\beta s\sigma)\right),\end{aligned}$$

where dilog is the dilogarithm function, $\text{dilog}(x) = \int_1^x \log(u)/(1-u)du$. Using Theorem 2.1 of Brockwell and Brown (1978), we get

$$\begin{aligned}E[Y_n^{(h)}]^{\gamma h} =& \frac{1}{\Gamma(1-\gamma h)}\int_0^{\infty} s^{-\gamma h}\left|D\tilde{f}_{Y_n^{(h)}}(s)\right| ds \\ =& \frac{\gamma}{a\Gamma(1-\gamma h)}\int_0^{\infty} s^{-\gamma h-1}\exp\left(\frac{\gamma}{a}\text{dilog}(1+\beta s\sigma)\right)\log(1+\beta s\sigma)ds, \quad (3.22)\end{aligned}$$

where Df denotes the derivative of f. Then $c_{\gamma h} = \left[E[Y_n^{(h)}]^{\gamma h}\right]^{1/(\gamma h)}$ can be numerically evaluated from (3.22) for fixed h.

Theorem 3.5.1 *For a sequence of observations $\{Y_n^{(h)}, n = 0, 1, \ldots, N\}$ from a gamma-driven CAR(1) process, we have $\hat{a}_N^{(h)} \to a$ a.s. and*

$$\lim_{N\to\infty} P\left[(-h)e^{-ah}k_N^{-1}(\hat{a}_N^{(h)} - a)c_\alpha \le x\right] = G_\alpha(x),$$

where G_α is as in (3.12), $\alpha = \gamma h$, $\hat{a}_N^{(h)}$ is defined in (3.13), k_N^{-1} is given in (3.21), and c_α is evaluated through (3.22).

Proof: At the beginning of Section 3.3, we have shown that $Y_n^{(h)}$ is a stationary discrete-time AR(1) with autoregression coefficient $\phi \in (0,1)$ and iid noise sequence $\{Z_n\}$. According to (3.20), the distribution function F of Z_n is regularly varying at zero with exponent $\alpha = \gamma h$ and satisfies the condition $F(0) = 0$. Since $0 \le Z_n \le \sigma(L(nh) - L((n-1)h)$, $\int u^\xi F(du) < \infty$ for all $\xi > 0$. By Corollary 2.4 of Davis and McCormick (1989), we have $\hat{\phi}_N^{(h)} \to \phi$ a.s., which implies $\hat{a}_N^{(h)} \to a$ a.s.. From the same corollary, we also conclude that

$$\lim_{N\to\infty} P\left[k_N^{-1}(\hat{\phi}_N^{(h)} - \phi)c_{\gamma h} \le x\right] = G_{\gamma h}(x),$$

where $\hat{\phi}_N^{(h)}$ and k_N^{-1} are given in (3.10) and (3.21) respectively, and $c_{\gamma h}$ is evaluated through (3.22). Using a Taylor series expansion, we find from this result that

$$\lim_{N\to\infty} P\left[(-h)e^{-ah}k_N^{-1}(\hat{a}_N^{(h)} - a)c_{\gamma h} \le x\right] = G_{\gamma h}(x).$$

□

Theorem 3.5.1 gives the limiting distribution of $\phi_N^{(h)}$ for fixed h as $N \to \infty$. It is of interest also to consider the behavior of the estimator as h also goes to zero. For any non-negative random variable Y with density function $f(u)$, we have

$$\begin{aligned} [EY^s]^{1/s} &= \left[\int_0^\infty u^s f(u)du\right]^{1/s} = \left[1 + s\int_0^\infty u^{s-1} f(u)du\right]^{1/s} \\ &\to \exp\left(\int_0^\infty u^{-1} f(u)du\right) = \exp\left(EY^{-1}\right) \quad \text{as} \quad s \to 0, \end{aligned}$$

as long as EY^{-1} is finite. Applying this result to $Y_n^{(h)}$ and using Theorem 2.1 of Brockwell and Brown (1978), we obtain

$$\begin{aligned} \lim_{h\to 0} c_{\gamma h} &= \exp\left(E(Y_n^{(h)})^{-1}\right) = \exp\left(\int_0^\infty \tilde{f}_{Y_n^{(h)}}(s)ds\right) \\ &= \exp\left(\int_0^\infty e^{\frac{\gamma}{a}\mathrm{dilog}(1+\beta s\sigma)}ds\right). \end{aligned} \tag{3.23}$$

The behavior of k_N^{-1}, defined in (3.21), is more complicated. Using L'Hospital's Rule, we have

$$\lim_{s\to 0} -\frac{\log\Gamma(s+1)}{s} = -\lim_{s\to 0}\frac{\Gamma'(s+1)}{\Gamma(s+1)} = -\Gamma'(1) = \gamma_E,$$

where γ_E is the Euler-Mascheroni constant, with numerical value of $0.5772\cdots$. Hence

$$\lim_{h\to 0}\left[\Gamma(\gamma h + 1)\right]^{-1/(\gamma h)} = e^{\gamma_E}$$

and

$$k_N^{-1} \sim (\sigma\beta)^{-1} e^{\gamma_E} N^{1/(\gamma h)} \quad \text{as} \quad N \to \infty \text{ and } h \to 0. \tag{3.24}$$

When h is small, k_N^{-1} and $c_{\gamma h}$ can be well approximated by (3.23) and (3.24). Since the rate of convergence in Theorem 3.5.1, as indicated by k_N^{-1}, increases as h decreases and since the limiting distribution $G_{\gamma h}$ becomes degenerate as $h \to 0$, this suggests the possibility of super-convergence of $\hat{a}_N^{(h)}$ to a as $N \to \infty$ and $h \to 0$. In fact, in Section 3.7, we show that for any fixed $T > 0$, $\hat{a}_{T/h}^{(h)} \to a$ a.s. as $h \to 0$.

3.6 Examples

In this section, we will illustrate the estimation procedure with both a simulated example and the German Deutsche Mark/US Dollars (DM/$) exchange rate series data as described in Section 1.4.2.

Example 3.6.1 Consider the gamma-driven CAR(1) process defined by,

$$DY(t) + 0.6Y(t) = DL(t), \qquad t \in [0, 5000], \tag{3.25}$$

was simulated at times $0, 0.001, 0.002, \ldots, 5000$, using an Euler approximation. The parameter γ of the standardized gamma process was 2. The process was then sampled at intervals $h = 0.01$, $h = 0.1$ and $h = 1$ by selecting every 10^{th}, 100^{th} and 1000^{th} value respectively. We generated 100 such realizations of the process and applied the above estimation procedure to generate 100 independent estimates, for each h, of the parameters a and σ. The sample means and standard deviations of these estimators are shown in Table 3.1, which illustrates the remarkable accuracy of the estimators.

To estimate the parameter γ of the driving standardized gamma process, the following procedure was used. For each h and each realization, the estimated CAR(1) parameters were used in (3.17) to generate the estimated increments $\Delta L_n^{(h)}$, $n = 1, \ldots, 5000/h$. These were then added in blocks of length $1/h$ to obtain 5000 independent estimated increments of L in one time unit. Figure 3.2 shows the histogram (bars) of the increments for one realization with $h = .01$, together

Table 3.1: Estimation results for gamma-driven Ornstein-Uhlenbeck process.

Spacing	Parameter	Gamma increments	
		Sample mean of estimators	Sample std deviation of estimators
h=1	a	0.59269	0.00381
	σ	0.99796	0.01587
h=0.1	a	0.59999	0.00000
	σ	1.00011	0.01281
h=0.01	a	0.60000	0.00000
	σ	0.99990	0.01175

with the true probability density (solid line) of the increments per unit time of the driving gamma process L. Even if we did not know that the background driving Lévy process is a gamma process, the histogram strongly suggests that this is the case. For each h and for each realization of the process, the sample mean $\hat{\gamma}$ of the estimated increments per unit time was then used to estimate the parameter γ of the driving standardized Lévy process, giving a set of 100 independent estimates of γ for each h. The sample means and standard deviations of these estimators are shown in Table 3.2.

Table 3.2: Estimation results for driving Lévy process.

Spacing	Parameter	Sample mean of estimators	Sample std deviation of estimators
$h = 1$	γ	1.99598	0.05416
$h = 0.1$	γ	2.00529	0.03226
$h = 0.01$	γ	2.00547	0.02762

Example 3.6.2 Taking the spacing $h = 1$, we apply our estimation procedure to the realized volatilities constructed from German Deutsche Mark/US Dollar exchange rate series data. For details on this data set, see Section 1.4.2.

The fitted CAR(1) model was

$$DY(t) + 1.3189\,Y(t) = 0.3567\,DL(t), \qquad t \in [0, 3045].$$

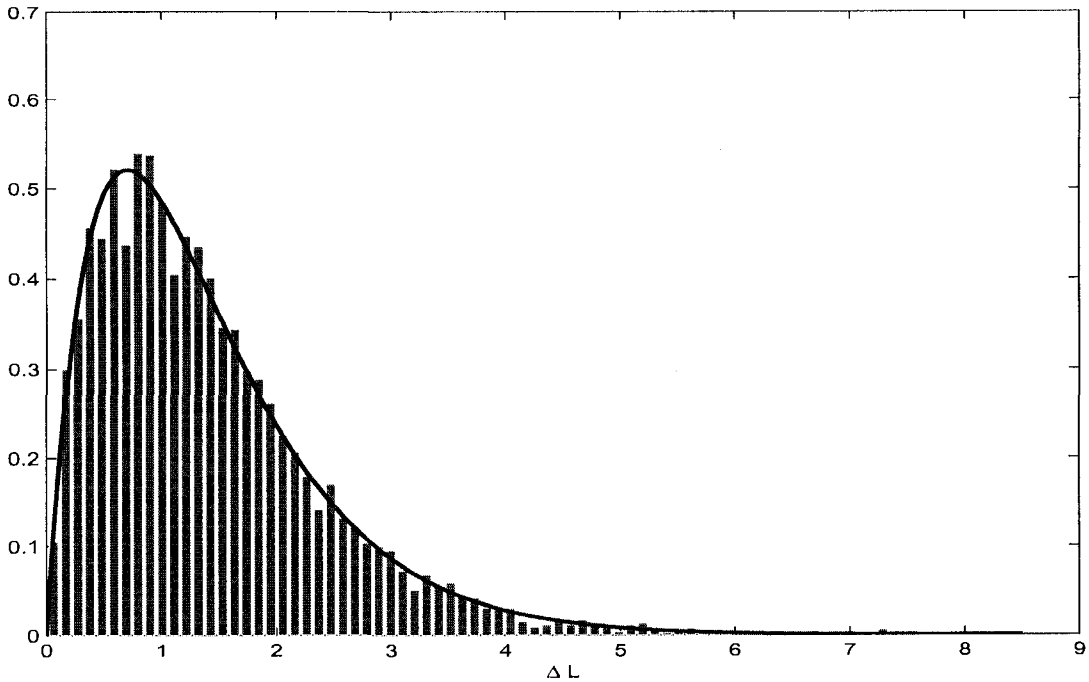


Figure 3.1: True probability density vs. histogram of the estimated Lévy increments.

Then Lévy increments, which turned to be all positive, were recovered as discussed in Section 3.4. Using those increments, we fitted three different Lévy processes by method of moments. The three processes we tried were: gamma process, reciprocal gamma process and inverse gaussian process. Figure 3.2 shows the histogram (bars) of estimated increments and the fitted probability densities (solid lines) for those three Lévy processes. As we shall see, the inverse Gaussian process fits the increments best. The fitted marginal probability density for the inverse Gaussian process is

$$f_{L(t)}(x) = \frac{3.8389t}{\sqrt{2\pi}} e^{5.9518t} x^{-1.5} \exp\{-7.3686t^2 x^{-1} - 1.2019x\} I_{\{x>0\}}(x).$$

3.7 Estimation for Continuously Observed Process

It is interesting to note that from a continuously observed realization on $[0, T]$ of a CAR(1) process driven by a non-decreasing Lévy process with drift $m = 0$, the

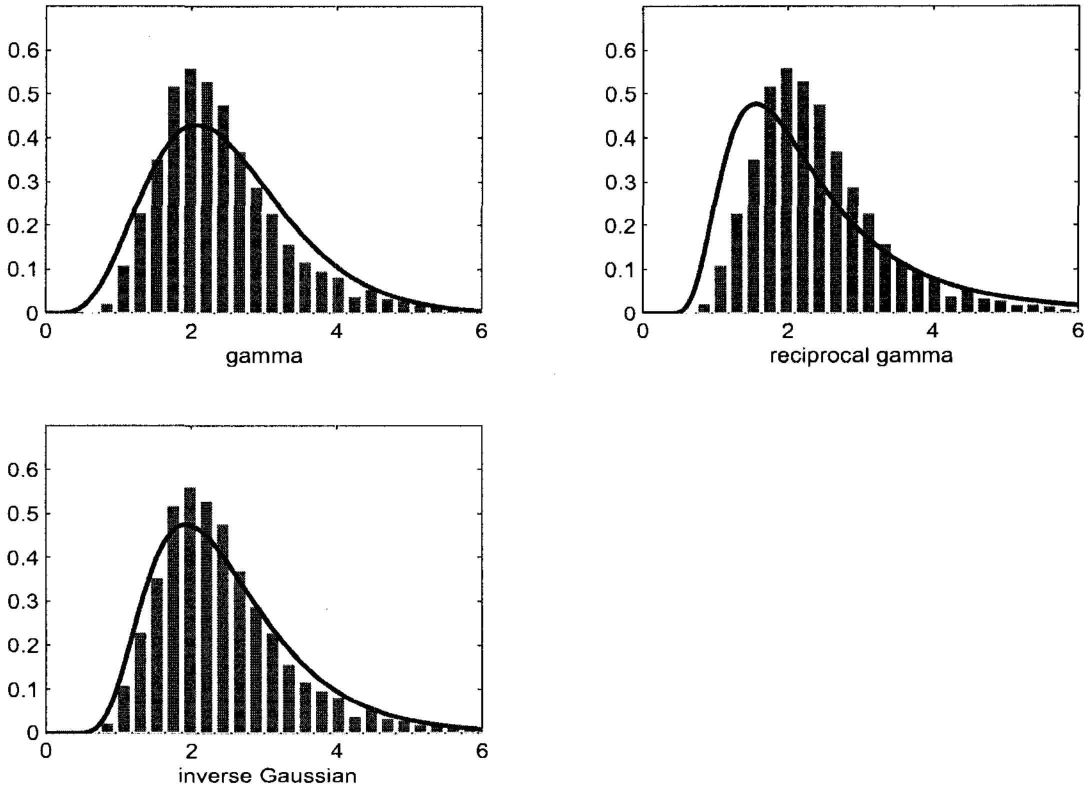


Figure 3.2: Fitted pdf vs. histogram of the estimated Lévy increments (DM/$ data).

value of a can be identified exactly with probability 1. This contrasts strongly with the case of a Gaussian CAR(1) process. The result is a corollary of the following theorem.

Theorem 3.7.1 *If the CAR(1) process $\{Y(t), t \geq 0\}$ defined by (3.1) is driven by a non-decreasing Lévy process L with drift m and Lévy measure ν, then for each fixed t,*

$$\frac{Y(t+h)-Y(t)}{h} + aY(t) \;\rightarrow\; m \text{ a.s. as } h \downarrow 0.$$

Proof: From (3.6) we find that

$$\begin{aligned} Y(t+h) - Y(t) =& Y(0)(e^{-a(t+h)} - e^{-at}) + \sigma(L(t+h) - L(t)) \\ & - a\sigma \int_0^t e^{-a(t-u)}(e^{-ah} - 1)L(u)du - a\sigma \int_t^{t+h} e^{-a(t+h-u)}L(u)du. \end{aligned}$$

Dividing each side by h, letting $h \downarrow 0$, and using the fact (Štatland (1965)) that

$$\lim_{h \downarrow 0}(L(t+h) - L(t))/h = m \quad ,$$

we see that

$$\frac{Y(t+h) - Y(t)}{h} \longrightarrow m - aY(0)e^{-at} + a^2\sigma \int_0^t e^{-a(t-u)}L(u)du - a\sigma L(t) = m - aY(t). \quad \square$$

Corollary 3.7.1 *If $m = 0$ in Theorem 3.7.1 (this is the case if the point zero belongs to the closure of the support of $L(1)$), then for each fixed t, with probability 1,*

$$a = \lim_{h \downarrow 0} \frac{\log Y(t) - \log Y(t+h)}{h}. \tag{3.26}$$

For each fixed $T > 0$, a is also expressible, with probability 1, as

$$a = \sup_{0 \le s < t \le T} \frac{\log Y(s) - \log Y(t)}{t - s}. \tag{3.27}$$

Proof: By setting $L(t) = 0$ for all t in the defining equation (3.1) we obtain the inequality, for all s and t such that $0 \le s < t \le T$,

$$\log Y(s) - \log Y(t) < a(t-s),$$

from which it follows that

$$a \ge \sup_{0 \le s < t \le T} \frac{\log Y(s) - \log Y(t)}{t - s}. \tag{3.28}$$

From Theorem 3.7.1 with $m = 0$ we find that

$$\frac{Y(t) - Y(t+h)}{hY(t)} \longrightarrow a \text{ as } h \downarrow 0.$$

From the inequalities (3.28) and $1 - x \le -\log x$ for $0 < x \le 1$, we obtain the inequalities,

$$\frac{Y(t) - Y(t+h)}{hY(t)} \le \frac{\log Y(t) - \log Y(t+h)}{h} \le a,$$

and letting $h \downarrow 0$ gives (3.26). But this implies that

$$a \leq \sup_{0 \leq s < t \leq T} \frac{\log Y(s) - \log Y(t)}{t - s},$$

which, with (3.28), gives (3.27). □

Remark 3.7.1 If observations are available only at times $\{nh : n = 0, 1, 2, \ldots, [T/h]\}$, and if the driving Lévy process has zero drift, Corollary 3.7.1 suggests the estimator,

$$\hat{a}_T^{(h)} = \sup_{0 \leq n < [T/h]} \frac{\log Y(nh) - \log Y((n+1)h)}{h}.$$

This estimator is precisely the same as the estimator (3.13). Its remarkable accuracy has already been illustrated in Table 3.1. The analogous estimator, based on closely but irregularly spaced observations at times $t_1, t_2, \ldots, t_N$ such that $0 \leq t_1 < t_2 < \cdots < t_N \leq T$, is

$$\hat{a}_T = \sup_n \frac{\log Y(t_n) - \log Y(t_{n+1})}{t_{n+1} - t_n}.$$

By Corollary 3.7.1, both estimators converge almost surely to a as the maximum spacing between successive observations converges to zero.

Chapter 4

NON-NEGATIVE LÉVY-DRIVEN CARMA PROCESSES

4.1 Introduction

In this chapter, we extend the results from previous chapter (see also Brockwell et al. (2007)) to make inference on the general stationary Lévy-driven CARMA process defined in Section 1.3.1, whose sampled process was discussed in Section 1.3.2. In Section 4.2, we describe how estimators of the parameters are obtained by maximizing the Gaussian likelihood of the observations and discuss the asymptotic properties of the estimators. In Section 4.3, we show how the increments of the driving Lévy process are determined for a given specified set of parameters. In Section 4.4, we apply the analysis to simulated CARMA(2,1) processes driven by inverse Gaussian and gamma Lévy processes. For the simulations, we have chosen the CARMA coefficients to be the maximum Gaussian likelihood estimates of the coefficients based on Todorov's realized volatility series. In Section 4.5, we apply the technique to the realized volatility series itself in order to determine which driving Lévy process is most compatible with the data.

4.2 Maximum Gaussian Likelihood Estimation via the Sampled Process

Observations are assumed to be available at the closely and uniformly spaced times $0, h, 2h, \ldots, Nh$, where N denotes the integer part of T/h. Our inference will therefore be based on observations of the sampled process, $\{Y_n := Y(nh), n = 0, 1, 2, \ldots, [T/h]\}$.

If Y is a *Gaussian* CARMA process, i.e. if L in Definition 1.3.1 is a Brownian motion, then it is well-known (see e.g. Doob (1944), Phillips (1959), Brockwell (1995)) that the sampled process $\{Y_n\}$, for any fixed $h > 0$, is a discrete-time ARMA(r, s) process with $0 \leq s < r \leq p$, driven by independent and identically distributed Gaussian white noise.

The probabilistic structure of the sampled Lévy-driven CARMA(p, q) process with distinct autoregressive roots is most clearly illustrated by inspection of (1.12) and (1.14) with $t = nh$ and $s = (n-1)h$. These equations show that the sampled process $\{Y_n\}$ is the sum of the sampled component processes $\{Y_{r,n}\}$ where the r^{th} sampled component process is a discrete-time autoregression of order 1 with coefficient $e^{\lambda_r h}$. Thus $\{Y_{r,n}\}$ satisfies the AR(1) equation

$$Y_{r,n} = e^{\lambda_r h} Y_{r,n-1} + Z_{r,n}, \; n = 0, \pm 1, \ldots, \tag{4.1}$$

with iid noise,

$$Z_{r,n} = \alpha_r \int_{(n-1)h}^{nh} e^{\lambda_r (nh-u)} dL(u). \tag{4.2}$$

If B denotes the backward shift operator, and we apply the operator $\prod_{i \neq r}(1 - e^{\lambda_i h} B)$ to each side of (4.1), then sum over $r = 1, \ldots, p$, we find that

$$\prod_{i=1}^{p} (1 - e^{\lambda_i h} B) Y_n = U_n, \tag{4.3}$$

where the sequence $\{U_n\}$ has the autocovariance structure of a moving average process of order less than p. Hence (see Brockwell and Davis (1991), p.89), equation (4.3) can be written as

$$\prod_{i=1}^{p} (1 - e^{\lambda_i h} B) Y_n = W_n + \theta_1 W_{n-1} + \cdots + \theta_{p-1} W_{n-p+1}, \tag{4.4}$$

where $\{W_n\}$ is a white noise sequence, *but not necessarily iid* (except when $p = 1$). ARMA processes with non-iid driving noise are called *weak*, as opposed to *strong*, ARMA processes.

Remark 4.2.1 As $h \to 0$, it is easy to check that, to within an error which is $o_p(h^2)$, W_n can be replaced on the right-hand side of (4.4) by $\sum_{i=1}^{p} \alpha_i(L(nh) - L((n-1)h))$, and the moving average polynomial, $1 + \theta_1 B + \cdots + \theta_{p-1}B^{p-1}$, by $\sum_{i=1}^{p} \alpha_i \prod_{j\neq i}(1 - e^{\lambda_j h}B)/\sum_{i=1}^{p} \alpha_i$. Since the increments $L(nh) - L((n-1)h)$ are independent, the behavior of the sampled process resembles that of a strong ARMA process for small h.

Remark 4.2.2 Our goal, in estimating the coefficients of the sampled process by maximum Gaussian likelihood, is to obtain almost surely consistent estimators of the coefficients of the sampled process and to use them to find corresponding maximum Gaussian likelihood estimates of $\{a_1, \ldots, a_p; b_1, \ldots, b_q\}$. From these, as will be described in Section 4.3, we estimate the increments of the driving Lévy process. Almost sure consistency of the maximum Gaussian likelihood estimators holds for the (weak) ARMA process $\{Y_n\}$, but in order to establish the standard asymptotic normal distribution of the estimators as $T \to \infty$ with h fixed, stronger conditions on the driving discrete-time white noise are required, e.g. the driving noise should be iid or a martingale difference sequence with constant conditional variance (see Brockwell and Davis (1991), Theorem 10.8.2 and Hannan (1973), Theorem 3). However for small h, Remark 4.2.1 suggests that the standard asymptotic distribution of the estimators under the iid assumption (Brockwell and Davis (1991), p. 386-387) should be approximately valid. This is borne out by the simulations described in Section 4.4 (see Remark 4.4.1 below). In the following example we derive, for the sampled CARMA(2,1) process with assumed iid noise, the asymptotic distribution of the maximum Gaussian likelihood estimators of $\{a_1, a_2, b_0\}$ as $T \to \infty$ with h fixed , and show that the limit of this asymptotic distribution as $h \to 0$ is the same as the asymptotic distribution as $T \to \infty$ of the maximum Gaussian likelihood estimators of a Gaussian CARMA(2,1) process observed continuously on the interval $[0, T]$.

Example 4.2.1 For the CARMA(2,1) process defined in Example 1.3.1, the sampled process $\{Y_n := Y(nh),\ n = 0, 1, \dots\}$ satisfies

$$Y_n - \phi_1^{(h)} Y_{n-1} - \phi_2^{(h)} Y_{n-2} = Z_n + \theta^{(h)} Z_{n-1}, \quad \{Z_n\} \sim \mathrm{WN}(0, \delta^2(h)).$$

For fixed h, as $T \to \infty$, the maximum likelihood estimator of $\boldsymbol{\beta} = \left[\phi_1^{(h)}, \phi_2^{(h)}, \theta^{(h)}\right]'$ based on observations $Y_1, \dots, Y_{[T/h]}$ is strongly consistent and, if we were to make the assumption that the sequence $\{Z_n\}$ is iid (see Remark 4.2.1 above), then (see Brockwell and Davis (1991), p.258)

$$\sqrt{T/h}\,(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \Rightarrow N(\mathbf{0}, M(\boldsymbol{\beta})), \tag{4.5}$$

where

$$M(\boldsymbol{\beta}) = \delta^2 \begin{bmatrix} E\mathbf{U}_t\mathbf{U}_t' & E\mathbf{V}_t\mathbf{U}_t' \\ E\mathbf{U}_t\mathbf{V}_t' & E\mathbf{V}_t\mathbf{V}_t' \end{bmatrix}^{-1}, \tag{4.6}$$

and the random vectors $\mathbf{U}_t$ and $\mathbf{V}_t$ are defined as $\mathbf{U}_t = [U_t, \cdots, U_{t+1-p}]'$ and $\mathbf{V}_t = [V_t, \cdots, V_{t+1-q}]'$, where $\{U_t\}$ and $\{V_t\}$ are stationary solutions of the autoregressive equations,

$$\phi(B)U_t = Z_t \qquad \text{and} \qquad \theta(B)V_t = Z_t. \tag{4.7}$$

In order to study the asymptotic behavior as $T \to \infty$ of the maximum likelihood estimators $\left[\hat{\phi}_1^{(h)}, \hat{\phi}_2^{(h)}, \hat{\theta}^{(h)}\right]'$, when h is small, we consider the expansion

$$\begin{aligned} M = {} & \frac{2a_1}{\mathcal{A}^2} \begin{bmatrix} \mathcal{B} & -\mathcal{B} & -2b_0^2\mathcal{C} \\ -\mathcal{B} & \mathcal{B} & 2b_0^2\mathcal{C} \\ -2b_0^2\mathcal{C} & 2b_0^2\mathcal{C} & b_0\mathcal{C}^2/a_1 \end{bmatrix} h - \frac{2a_1^2}{\mathcal{A}^2} \begin{bmatrix} \mathcal{D}^2 & -\mathcal{B} & -b_0\mathcal{C}\mathcal{D}/a_1 \\ -\mathcal{B} & \mathcal{B} + 4a_2b_0^2 & b_0\mathcal{C}^2/a_1 \\ -b_0\mathcal{C}\mathcal{D}/a_1 & b_0\mathcal{C}^2/a_1 & b_0^2\mathcal{C}^2/a_1^2 \end{bmatrix} h^2 \\ & + \frac{a_1}{3\mathcal{A}^2} \begin{bmatrix} a_1\mathcal{E} & -\mathcal{G} & -\mathcal{C}\mathcal{H}/2 \\ -\mathcal{G} & a_1(\mathcal{E} + 48a_1a_2b_0^2) & \mathcal{C}(\mathcal{H} + 12a_1a_2b_0 + 12a_2b_0^2)/2 \\ -\mathcal{C}\mathcal{H}/2 & \mathcal{C}(\mathcal{H} + 12a_1a_2b_0 + 12a_2b_0^2)/2 & -\mathcal{C}^2\mathcal{I}/(4a_1b_0) \end{bmatrix} h^3 \\ & + O(h^4) \quad \text{as } h \to 0, \end{aligned} \tag{4.8}$$

where

$$
\begin{aligned}
\mathcal{A} =& -a_2 + a_1 b_0 - b_0^2, \\
\mathcal{B} =& a_2^2 - 2a_1 a_2 b_0 + a_1^2 b_0^2 + 2a_2 b_0^2 + 2a_1 b_0^3 + b_0^4, \\
\mathcal{C} =& a_2 + a_1 b_0 + b_0^2, \\
\mathcal{D} =& -a_2 + a_1 b_0 + b_0^2, \\
\mathcal{E} =& 4a_1 a_2^2 + 9a_2^2 b_0 - 8a_1^2 a_2 b_0 + 4a_1^3 b_0^2 - 14a_1 a_2 b_0^2 + 9a_1^2 b_0^3 - 6a_2 b_0^3 + 6a_1 b_0^4 + b_0^5, \\
\mathcal{G} =& 4a_1^2 a_2^2 - 3a_2^3 + 3a_1 a_2^2 b_0 - 8a_1^3 a_2 b_0 + 4a_1^4 b_0^2 - 6a_2^2 b_0^2 + 7a_1^2 a_2 b_0^2 + 9a_1^3 b_0^3 \\
& + a_1 b_0^5 + 6a_1^2 b_0^4 - 3a_2 b_0^4, \\
\mathcal{H} =& 2a_2^2 - 3a_1 a_2 b_0 - 8a_2 b_0^2 + 5a_1^2 b_0^2 + 9a_1 b_0^3 + 6b_0^4, \\
\mathcal{I} =& -a_2^2 - 4a_1 a_2 b_0 + a_1^2 b_0^2 - 2a_2 b_0^2 - 4a_1 b_0^3 - 17b_0^4.
\end{aligned}
$$

Based on the mapping from (ϕ_1, ϕ_2, θ) to (a_1, a_2, b_0), the components $\{C_{ij},\ i,j = 1,2,3\}$ of the matrix

$$
C = \begin{bmatrix} \frac{\partial a_1}{\partial \phi_1} & \frac{\partial a_1}{\partial \phi_2} & \frac{\partial a_1}{\partial \theta} \\ \frac{\partial a_2}{\partial \phi_1} & \frac{\partial a_2}{\partial \phi_2} & \frac{\partial a_2}{\partial \theta} \\ \frac{\partial b_0}{\partial \phi_1} & \frac{\partial b_0}{\partial \phi_2} & \frac{\partial b_0}{\partial \theta} \end{bmatrix} \tag{4.9}
$$

have the asymptotic expansions

$$
\begin{aligned}
C_{11} &= 0, \\
C_{12} = C_{21} &= \frac{1}{h} + a_1 + \frac{a_1^2}{2} h + \cdots, \\
C_{13} = C_{31} &= \frac{a_2 + b_0^2}{12 b_0} + \frac{(a_2 + b_0^2) a_1}{24 b_0} h + \frac{(a_2 + b_0^2)(a_1^2 + 4b_0^2)}{72 b_0} h^2 + \cdots, \\
C_{22} &= -\frac{1}{h^2} - \frac{a_1}{2h} - \frac{a_1^2 - 4a_2}{12} + \cdots, \\
C_{23} = C_{32} &= \frac{a_2 + b_0^2}{12 b_0} + \frac{a_1(3b_0^2 + a_2)}{24 b_0} h + \frac{7b_0^2 a_1^2 + b_0^2 a_2 + a_2 a_1^2 - 3a_2^2 + 4b_0^4}{72 b_0} h^2 + \cdots, \\
C_{33} &= \frac{1}{h} + b_0 + \left(-\frac{a_2}{12} + \frac{3}{8} b_0^2 + \frac{a_2^2}{24 b_0^2} + \frac{a_1^2}{24} \right) h + \left(\frac{b_0^3}{12} + \frac{a_2^2}{12 b_0}\right) h^2 + \cdots.
\end{aligned}
$$

From (4.8) and (4.9) we find that

$$CMC' = \frac{2}{\mathcal{A}^2 h}\begin{bmatrix} a_1\mathcal{B} & 4a_1^2a_2b_0^2 & 2a_1b_0^2\mathcal{C} \\ 4a_1^2a_2b_0^2 & a_1a_2(\mathcal{B}+4a_1a_2b_0-4a_1b_0^3) & 2a_1a_2b_0\mathcal{C} \\ 2a_1b_0^2\mathcal{C} & 2a_1a_2b_0\mathcal{C} & b_0\mathcal{C}^2 \end{bmatrix}(1+o(1)), \tag{4.10}$$

as $h \to 0$, and hence, from (4.5) that the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta} = [a_1, a_2, b_0]'$ based on observations of Y at times $0, h, 2h, \ldots, h[T/h]$, satisfies

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \Rightarrow N(\mathbf{0}, V), \text{ as } T \to \infty, \tag{4.11}$$

where $V = hCMC'$.

Remark 4.2.3 The leading term in the expansion of the covariance matrix V coincides with the asymptotic covariance matrix obtained by Tuan (1977) for the maximum Gaussian likelihood estimator of $\boldsymbol{\theta} := [a_1, a_2, b_0]'$ of the corresponding Gaussian CARMA process observed continuously on the interval $[0, T]$, namely J^{-1}, where

$$J_{jk} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left\{\frac{\partial}{\partial\theta_j}\frac{a(i\omega)}{b(i\omega)}\right\}\left\{\frac{\partial}{\partial\theta_k}\frac{a(-i\omega)}{b(-i\omega)}\right\}\left|\frac{b(i\omega)}{a(i\omega)}\right|^2 d\omega, \quad \text{for } j, k = 1, 2, 3.$$

4.3 Recovering the Background Driving Lévy Process

In order to suggest an appropriate parametric model for the driving Lévy process, it is necessary to recover the realization of L from a given realization of Y on $[0, T]$ for given or estimated values of $\{a_j,\ 1 \le j \le p;\ b_j,\ 0 \le j < q;\ \sigma\}$. In general this requires knowledge of the initial state vector $\mathbf{X}(0)$, but if this is available or if we are willing to assume a plausible value for $\mathbf{X}(0)$, then an argument due to Tuan (1977) can be used to recover L. In this section, we shall assume that the polynomial a and polynomial b have all their zeroes in the left half-plane. Since the covariance structure of our Lévy-driven process is exactly the same (except for slight notational changes) as that of Tuan's Gaussian CARMA process and since his result holds for Gaussian CARMA processes with arbitrary mean, his L^2-based spectral argument can be applied directly to the Lévy-driven CARMA process to give, for $t \geq 0$,

$$L(t) = \sigma^{-1}\left[Y^{(p-q-1)}(t) - Y^{(p-q-1)}(0)\right] - \int_0^t \left[\sum_{j=1}^{q} b_{q-j}X^{(p-j)}(u) - \sum_{j=1}^{p} a_j X^{(p-j)}(u)\right] du, \tag{4.12}$$

where $X^{(0)}, \ldots, X^{(p-1)}$ are the components of the state process $\mathbf{X}$ defined in (1.8). The Lévy process increment on the interval $(s,t]$ is therefore

$$\begin{aligned}\Delta L := L(t) - L(s) =& \sigma^{-1}\left[Y^{(p-q-1)}(t) - Y^{(p-q-1)}(s)\right] \\ &- \int_s^t \left[\sum_{j=1}^{q} b_{q-j}X^{(p-j)}(u) - \sum_{j=1}^{p} a_j X^{(p-j)}(u)\right] du, \ t > s \geq 0.\end{aligned} \tag{4.13}$$

These increments are easily expressed in terms of Y and $\mathbf{X}(0)$ by noting that (1.3) characterizes $X^{(0)}$ as a CARMA$(q,0)$ process driven by the process $\{\sigma^{-1}\int_0^t Y(u)du\}$. Making use of this observation, introducing the $q \times 1$ state vector $\mathbf{X}_q(t) := [X^{(0)}(t), \ldots, X^{(q-1)}(t)]'$ and proceeding exactly as we did in solving the CARMA equations in Section 1.3.1, we find that for $q \geq 1$,

$$\mathbf{X}_q(t) = \mathbf{X}_q(0)e^{Bt} + \sigma^{-1}\int_0^t e^{B(t-u)}\mathbf{e}_q Y(u)du, \tag{4.14}$$

where

$$B = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -b_0 & -b_1 & -b_2 & \cdots & -b_{q-1} \end{bmatrix} \text{ and } \mathbf{e}_q = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

while for $q = 0$,

$$X^{(0)}(t) = \sigma^{-1}Y(t). \tag{4.15}$$

The remaining derivatives of $X^{(0)}$ up to order $p-1$ can be determined from (4.14) and (4.15), completing the determination of the state vector $\mathbf{X}(t)$. Then the driving Lévy process is found from (4.12).

In the CAR(1) case, $L(t)$ takes the extremely simple form,

$$L(t) = \sigma^{-1}\left[Y(t) - Y(0) - \lambda \int_0^t Y(s)ds\right], \tag{4.16}$$

where λ is the autoregressive root, $\lambda = -a_1$.

In the case when the autoregressive roots are distinct, we can take advantage of the simple expression (4.16) and the canonical representation of Y as follows. We use the transformation (1.17) to recover the canonical state process $\mathbf{Y}$ defined by (1.13) and (1.16) from $\mathbf{X}$, and then apply (4.16) to each of the component processes Y_r to obtain p (equivalent) representations of $L(t)$, namely

$$L(t) = \alpha_r^{-1}\left[Y_r(t) - Y_r(0) - \lambda_r \int_0^t Y_r(u)du\right], \quad r = 1, \ldots, p. \tag{4.17}$$

If Y is observed continuously on $[0, T]$ we can use any one of these p very simple equations to recover L from the realization of Y, the value of $\mathbf{X}(0)$ and the parameters of the CARMA process. Of course for calculations it is simplest to choose (if possible) a value r in (4.17) for which λ_r is real.

Example 4.3.1 (The CARMA(2,1) Process). For a specified parameter set $\{a_1, a_2, b_0, \sigma\}$ and assumed initial value $X^{(0)}(0)$, the state vector $\mathbf{X}(t) = [X^{(0)}(t), X^{(1)}(t)]'$ at time $t > 0$ can be constructed from the relations,

$$X^{(0)}(t) = X^{(0)}(0)e^{-b_0 t} + \sigma^{-1}\int_0^t e^{-b_0(t-u)}Y(u)du,$$

and

$$X^{(1)}(t) = -b_0 X^{(0)}(t) + \sigma^{-1}Y(t).$$

The canonical state vector $\mathbf{Y}(t)$ is then given in Example 1.3.1 and the increments of L obtained from either of the two equations (4.17).

We have assumed throughout this section that Y is observed continuously. If Y is observed at closely-spaced discrete times, the integrals must be replaced by approximating sums. This is done in the illustrative examples of the following sections.

4.4 Applications to Simulated Series

In this section, we illustrate the estimation procedure with simulated CARMA(2,1) processes driven by inverse Gaussian and gamma processes. The coefficients of the simulated models were chosen to coincide with the maximum Gaussian likelihood estimates for the realized volatility series of Todorov (2006) which is discussed in more detail in Section 4.5. The corresponding defining differential equation is

$$(D^2+1.36233D+0.04445)Y(t) = 0.28886\,(0.20603+D)DL(t),\ t \in [0, 2000]. \quad (4.18)$$

In Example 4.4.1 we consider this process driven by a standardized inverse Gaussian process and in Example 4.4.2 we suppose that the driving process is a standardized gamma process. In each case we choose the parameter, $\mu := EL(1)$, of the driving process to be 0.50015. (This particular value was chosen so that the mean of the process Y is equal to the sample mean of Todorov's realized volatility series.) Models are fitted to replications of the simulated series and the fitted models are compared with those generating the data.

Example 4.4.1 (The inverse-Gaussian-driven CARMA(2,1) Process). The standardized inverse Gaussian process, with $EL(t) = \mu t$ (and $\mathrm{Var}(L(t)) = t$) is the Lévy process with marginal probability densities

$$f_{L(t)}(x) = \frac{\mu^{3/2} t}{\sqrt{2\pi x^3}} \exp\left\{-\frac{\mu(x-\mu t)^2}{2x}\right\} I_{\{x>0\}}(x). \quad (4.19)$$

A CARMA(2,1) process driven by a standardized inverse Gaussian process is therefore determined by 5 parameters, a_1, a_2, b_0, σ and μ.

In our simulation study, 100 realizations of the CARMA(2,1) process (4.18) driven by the standardized inverse Gaussian process with parameter $\mu = 0.50015$ were generated at times 0, 0.01, 0.02, ..., 2000, using an Euler approximation. Each realization was then sampled at intervals $h = 0.1$ and $h = 1$ by selecting every 10^{th} and every 100^{th} value respectively. For each realization we computed maximum Gaussian likelihood estimators of the parameters a_1, a_2, b_0 and σ. The parameter μ was estimated by equating the mean of the fitted model to the sample mean of the simulated series. The sample means and standard deviations of these estimators are shown in Table 4.1.

Table 4.1: Estimation results for inverse-Gaussian-driven CARMA(2,1) process.

Spacing	Parameter	Sample mean of estimators	Sample std deviation of estimators
h=1	a_1	1.38802	0.13093
	a_2	0.04676	0.01588
	b_0	0.21173	0.03362
	σ	0.28521	0.02194
	μ	0.54724	0.10276
h=0.1	a_1	1.38202	0.06132
	a_2	0.04733	0.01232
	b_0	0.20982	0.02611
	σ	0.28709	0.01670
	μ	0.55254	0.10375

Remark 4.4.1 For the spacing $h = 0.1$, the sample covariance matrix of the estimates of $[a_1, a_2, b_0]'$ is

$$10^{(-4)} \times \begin{bmatrix} 37.6068 & 4.4291 & 11.9388 \\ 4.4291 & 1.5176 & 2.3772 \\ 11.9388 & 2.3772 & 6.8173 \end{bmatrix}.$$

Evaluating the asymptotic covariance matrix $V/2000$ (see equation (4.11)) using the parameters of the simulated process defined by (4.18), we obtain the corresponding covariance matrix,

$$10^{(-4)} \times \begin{bmatrix} 30.9129 & 3.7302 & 11.3213 \\ 3.7302 & 1.4103 & 2.4425 \\ 11.3213 & 2.4425 & 7.4132 \end{bmatrix},$$

showing that the asymptotic covariance V/T of the maximum likelihood estimators for a continuously observed Gaussian process on [0,T] provides a good approximation in this case to the covariance of our estimators when $T = 2000$ and $h = 0.1$. (See Remark 4.2.2.)

To illustrate the estimation of the Lévy increments $L(n)-L(n-1)$, we simulated the same process on the interval [0,5000], sampled it at integer times, estimated the parameters by maximum Gaussian likelihood, and, with these parameter values, estimated the Lévy increments as described in Section 4.3, using the CAR(1) process corresponding to the autoregressive root with smaller absolute value. A kernel density (dash-dot line) estimate of the probability density of the estimated increments is shown in Figure 4.1, together with the inverse Gaussian density (solid line) of the increments per unit time of the Lévy process used to generate the series. A small fraction (.035) of the increments were estimated to be small and negative. These were set to zero. The fit appears to be very good.

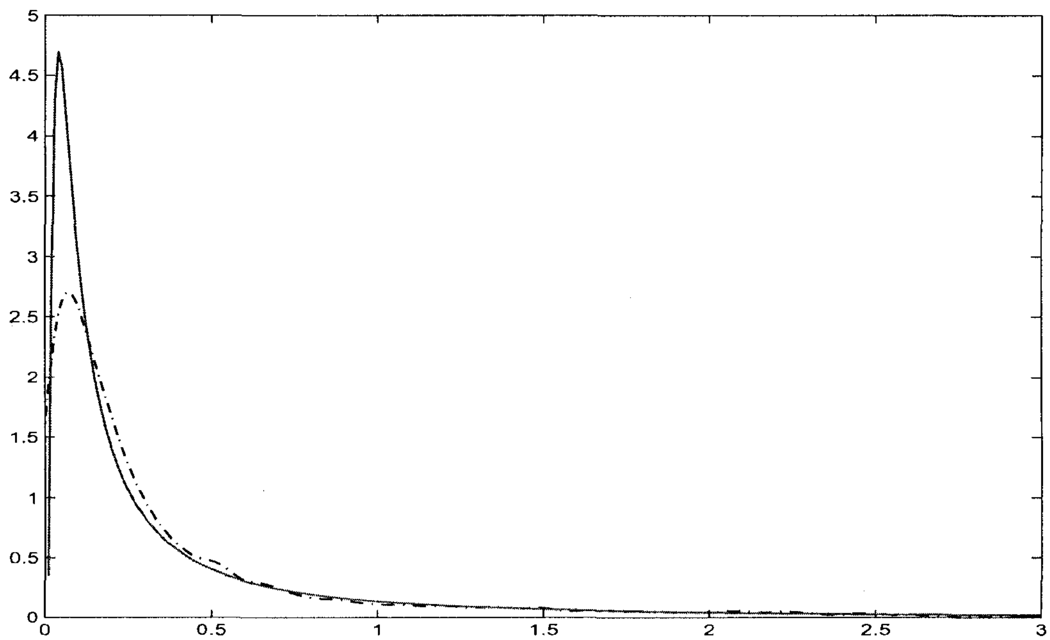


Figure 4.1: Probability density vs. kernel density for inverse Gaussian increments.

Example 4.4.2 (The gamma-driven CARMA(2,1) Process). The standardized gamma process, with $EL(t) = \mu t$ (and $\text{Var}(L(t)) = t$) is the Lévy process with marginal probability densities

$$f_{L(t)}(x) = \frac{x^{\mu^2 t-1}}{\Gamma(\mu^2 t)} \mu^{\mu^2 t} \exp(-\mu x) I_{\{x>0\}}(x). \tag{4.20}$$

The CARMA(2,1) process driven by a standardized gamma process is, like the inverse-Gaussian-driven process, determined by the 5 parameters, a_1, a_2, b_0, σ and μ.

As in the previous example, we generated 100 realizations of the process at times $0, 0.01, 0.02, \ldots, 2000$, using an Euler approximation, and then sampled each realization at intervals $h = 0.1$ and $h = 1$ by selecting every 10^{th} and every 100^{th} value respectively.

For each resulting series we computed maximum Gaussian likelihood estimators of the parameters a_1, a_2, b_0 and σ and estimated μ by equating the mean of the fitted model to the sample mean of the simulated series. The sample means and standard deviations of these estimators are shown in Table 4.2.

Table 4.2: Estimation results for gamma-driven CARMA(2,1) process.

Spacing	Parameter	Sample mean of estimators	Sample std deviation of estimators
h=1	a_1	1.38396	0.10057
	a_2	0.04902	0.01603
	b_0	0.21157	0.02932
	σ	0.29057	0.02187
	μ	0.52306	0.11560
h=0.1	a_1	1.37186	0.06300
	a_2	0.04884	0.01390
	b_0	0.21001	0.02740
	σ	0.28924	0.01683
	μ	0.53017	0.10189

The Lévy process increments were estimated as described in Example 4.4.1, and in Figure 4.2 a comparison is made between the true distribution (solid line) of the increments of the driving process and a kernel density (dash-dot line) based on the estimated Lévy increments. Again the small negative estimates were set to zero before computing the kernel density estimate and again the fit between the generating and empirical densities is good.

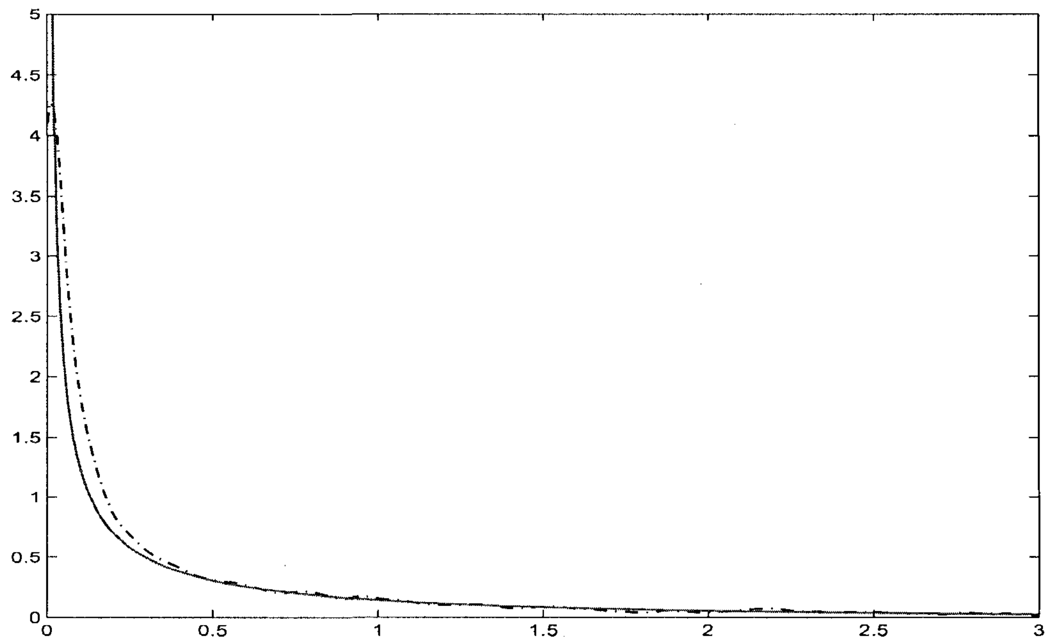


Figure 4.2: Probability density vs. kernel density for gamma increments.

4.5 Analysis of Todorov's Realized Volatility Series

In this section, we will illustrate the estimation procedure with the German Deutsche Mark/US Dollars (DM/\$) exchange rate series data as described in Section 1.4.2.

The CARMA(2,1) model fitted by maximizing the Gaussian likelihood is

$$(D^2+1.36233D+0.04445)Y(t) = 0.28886\ (0.20603+D)DL(t),\ t \in [0, 3045]. \quad (4.21)$$

In Figure 4.3, we show the excellent fit of the autocorrelation function (solid line) of Y to the empirical autocorrelations (vertical bars) up to a lag of 80 days. Our next objective is to estimate the Lévy increments in order to identify an appropriate driving Lévy process to complete the specification of the model (4.21).

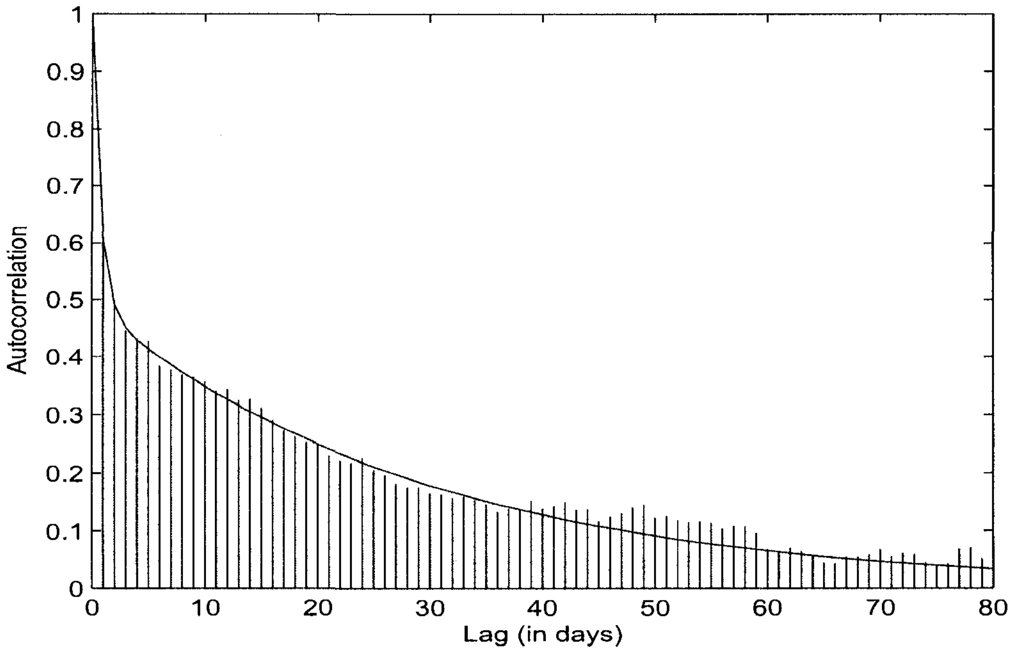


Figure 4.3: Empirical vs. fitted model ACF of the DM/$ realized volatility data.

The estimates of the Lévy increments $L(n) - L(n-1)$ were obtained by constructing a realization of the component CAR(1) process corresponding to the estimated autoregressive root $\hat{\lambda}_1 = -0.03345$ and proceeding as described in Example 4.3.1 to estimate the Lévy increments from it. After setting the estimated small negative increments to zero, a kernel estimate (dash-dot lines) of the probability density of the increments was computed and is shown in Figure 4.4 together with probability densities (solid lines) of four potential densities of $L(1)$ for standardized Lévy processes having the same mean (0.6187) as that of the kernel density estimate. The four potential standardized driving Lévy processes L we tested are inverse Gaussian process with marginal probability density (4.19), gamma process with marginal probability density (4.20), reciprocal gamma process with marginal probability density

$$f_{L(t)}(x) = \frac{(\mu^3 t^2 + \mu t)^{\mu^2 t + 2}}{\Gamma(\mu^2 t + 2)} x^{-\mu^2 t - 3} \exp\left\{-(\mu^3 t^2 + \mu t)x^{-1}\right\} I_{\{x>0\}}(x),$$

and positive hyperbolic process with marginal probability density

$$f_{L(t)}(x) = \frac{\gamma^2}{2} \exp\left\{-\frac{1}{2}(\delta^2 x^{-1} + \gamma^2 x)\right\} I_{\{x>0\}}(x),$$

where $\delta > 0$ and $\gamma \geq 0$ can be evaluated through

$$\mu t = \frac{2}{\gamma^2} - \left(\gamma^2 t + \gamma^2 \mu^2 - 4\mu\right)\left[\ln \gamma + \frac{1}{2}\ln\left(\gamma^2 t + \gamma^2 \mu^2 - 4\mu\right)\right],$$
$$\delta^2 = \gamma^2 t + \gamma^2 \mu^2 - 4\mu.$$

Among the four Lévy processes, the best fitting appears be the standardized gamma process with $\mu = 0.6187$. This suggests the model (4.18) for the data with L a standardized gamma process with $\mu = 0.6187$.

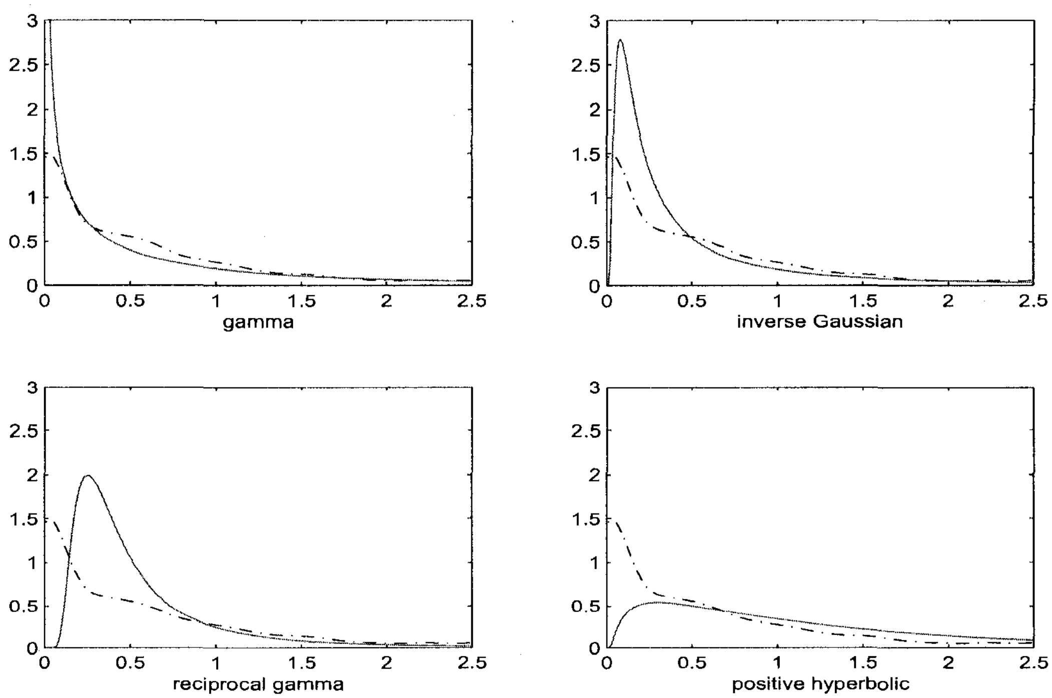


Figure 4.4: Fitted pdf vs. kernel density estimate of Lévy increments (DM/$ data).

Chapter 5

CONCLUSION

In this thesis we have developed inferential techniques for continuous-time Lévy-driven autoregressive moving average (CARMA) processes using uniformly and closely spaced discrete-time observations. In the non-Gaussian case we have gone beyond previous approaches by investigating the nature and parameters of the driving Lévy process itself. Many interesting and important questions remain. Some of these are outlined below.

5.1 Future Work

If we relax the assumption of closely-spaced observations of the CARMA process driven by a Lévy process L, it is possible to write down the moving average component of the sampled (ARMA) process explicitly in terms of integrals with respect to L. If the spacing h is small then it is, to order $o_p(h^2)$, a moving average driven by iid noise. It would be interesting to use the exact representation of this moving average in terms of L to obtain estimation procedures which are good for any value of h and to establish the asymptotic distribution of the maximum Gaussian likelihood estimators as $T \to \infty$ for arbitrary fixed h.

Although, at least in principle, the techniques of Chapter 4 apply to Lévy-driven CARMA(p, q) processes for general $p > q$, our numerical examples have focussed on the CARMA(2,1) case, partly because this model, with one small and one larger autoregressive root, has been found to be effective in modeling certain financial time series. Further numerical studies, applying the techniques to higher-order processes, are needed.

Best linear prediction of a Lévy-driven CARMA process based on discrete observations is of course the same as for a Gaussian process since the predictors and their mean squared errors depend only on the autocovariance structure of the process. It would be interesting however to determine minimum mean squared error predictors and their mean squared errors, taking into account the non-Gaussian nature of the driving process.

In Chapter 3, the asymptotic distribution of the Davis-McCormick estimator of the Ornstein-Uhlenbeck coefficient was derived under the assumption of slow variation of the density of the increments of the sampled process near zero, an assumption satisfied by the gamma-driven process. Although the estimator was also found in simulation studies to perform extremely well for processes not satisfying this assumption, a theoretical investigation of the large-sample behavior of the estimator in such cases would be of great interest. Generalization of the Davis-McCormick estimator to higher order CAR or CARMA would also be of interest.

In the Barndorff-Nielsen and Shephard model for log asset price, (1.27) and (1.28), the volatility is not an observed quantity. For this reason a great deal of effort has gone into the estimation of volatility. Estimated volatility sequences are referred to as realized volatility. Our estimation of the driving Lévy process L thus depends on the reliability of the realized volatility. It is natural therefore to ask if the driving Lévy process can be estimated more directly in terms of the observed log asset prices without the intermediate construction of the realized volatility.

5.2 Summary

This thesis deals with estimation of the parameters of Lévy-driven CARMA processes using available discrete-time observations. Three particular families of Lévy-driven CARMA processes were discussed:

Gaussian Autoregression. From the Radon-Nikodym derivative with respect to Wiener measure of the distribution of the $(p-1)^{\text{th}}$ derivative of a continuous-time linear or non-linear autoregression, observed on the interval $[0, T]$, we have shown how to compute maximum likelihood parameter estimators, conditional on the initial state vector. For closely-spaced discrete observations, the integrals appearing in the estimators are replaced by approximating sums.

The examples illustrate the accuracy of the approximations in special cases. If the observations are not uniformly spaced but the maximum spacing is small, appropriately modified approximating sums can be used in order to approximate the exact solution for the continuously observed process.

Non-negative Lévy-driven Ornstein-Uhlenbeck (or CAR(1)). We developed a highly efficient method, based on observations at times 0, h, $2h$, $\ldots$, Nh, for estimating the parameters of a stationary Ornstein-Uhlenbeck process $\{Y(t)\}$ driven by a non-decreasing Lévy process $\{L(t)\}$. If h is small, we used a discrete approximation to the exact integral representation of $L(t)$ in terms of $\{Y(s), s \leq t\}$ to estimate the increments of the driving Lévy process, and hence to estimate the parameters of the Lévy process. Under specified conditions on the driving Lévy process we obtained the asymptotic distribution of the estimator of the CAR(1) coefficient as $N \to \infty$ with h fixed.

The accuracy of the procedure was illustrated with a simulated example of a gamma-driven process. We also showed for a pure-jump Lévy-driven CAR(1) process, that the coefficient a is determined almost surely by a continuously observed realization of Y on any interval $[0, T]$. This distinguishes Y sharply from the corresponding Gaussian process.

Non-negative Lévy-driven CARMA(p, q). We developed a general procedure to estimate the parameters of a stationary CARMA process $\{Y(t)\}$ driven by a non-decreasing Lévy process $\{L(t)\}$, and to estimate the increments of the background driving Lévy process, based on closely-spaced observations at times $0, h, 2h, \ldots, Nh$. The results extend those obtained for the Lévy-driven stationary Ornstein-Uhlenbeck process. The idea is to use maximum Gaussian likelihood to obtain strongly consistent estimators of the CARMA coefficients and then, with these coefficients, to construct a realization of the canonical CAR(1) components of the CARMA process from the observed data. From these we construct the corresponding realization of the driving Lévy increments.

For the CARMA(2,1) model fitted by maximum Gaussian likelihood to a realized volatility series for the German Deutsche Mark/US Dollar daily exchange rate, simulation studies support the validity of the discrete approximations made to results which are exact for continuously observed processes. When the technique is applied to the data itself, we are able to compare the performance of a variety of possible standardized Lévy models for the driving noise. It appears from our results that, of the four models considered (gamma, inverse Gaussian, reciprocal gamma, and positive hyperbolic), the gamma process is the most compatible with the Lévy increments estimated from the data.

Bibliography

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001). The distribution of realized exchange rate volatility. *J. Amer. Statist. Assoc.*, 96(453):42–55.

Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus*, volume 93 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge.

Arató, M. (1982). *Linear Stochastic Systems with Constant Coefficients: A Statistical Approach*, volume 45 of *Lecture Notes in Control and Information Sciences.* Springer-Verlag, Berlin.

Barndorff-Nielsen, O. E. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B*, 63(2):167–241.

Bergstrom, A. (1985). The estimation of parameters in non-stationary higher-order continuous-time dynamic models. *Econometric Theory*, 1:369–385.

Brockwell, P., Davis, R., and Yang, Y. (2007). Estimation for non-negative Lévy-driven Ornstein-Uhlenbeck processes. Submitted to *J. App. Prob.*

Brockwell, P. J. (1995). A note on the embedding of discrete-time ARMA processes. *J. Time Ser. Anal.*, 16(5):451–460.

Brockwell, P. J. (2001a). Continuous-time ARMA processes. In Shanbhag, D. N. and Rao, C. R., editors, *Stochastic processes: theory and methods*, volume 19 of *Handbook of Statist.*, pages 240–276. North-Holland, Amsterdam.

Brockwell, P. J. (2001b). Lévy-driven CARMA processes. *Ann. Inst. Statist. Math.*, 53(1):113–124.

Brockwell, P. J. and Brown, B. M. (1978). Expansions for the positive stable laws. *Z. Wahrsch. verw. Gebiete*, 45(3):213–224.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods.* Springer Series in Statistics. Springer-Verlag, New York, second edition.

Brockwell, P. J. and Marquardt, T. (2005). Lévy-driven and fractionally integrated ARMA processes with continuous time parameter. *Statist. Sinica*, 15(2):477–494.

Brockwell, P. J. and Williams, R. J. (1997). On the existence and application of continuous-time threshold autoregressions of order two. *Adv. Appl. Prob.*, 29(1):205–227.

Davis, R. A. and McCormick, W. P. (1989). Estimation for first-order autoregressive processes with positive or bounded innovations. *Stochastic Process. Appl.*, 31(2):237–250.

Doetsch, G. (1974). *Introduction to the Theory and Application of the Laplace Transformation.* Springer-Verlag, New York.

Doob, J. L. (1944). The elementary Gaussian processes. *Ann. Math. Statistics*, 15:229–282.

Eberlein, E. and Raible, S. (1999). Term structure models driven by general Lévy processes. *Math. Finance*, 9(1):31–53.

Hannan, E. J. (1973). The asymptotic theory of linear time-series models. *J. App. Prob.*, 10:130–145, corrections, ibid. 10 (1973), 913.

Hyndman, R. J. (1993). Yule-Walker estimates for continuous-time autoregressive models. *J. Time Ser. Anal.*, 14(3):281–296.

Jones, R. (1981). Fitting a continuous time autoregression to discrete data. In Findley., D., editor, *Applied Time Series Analysis*, volume II, pages 651–682. Academic Press, New York.

Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition.

Øksendal, B. (1998). *Stochastic Differential Equations: An Introduction with Applications*. Universitext. Springer-Verlag, Berlin, fifth edition.

Phillips, A. W. (1959). The estimation of parameters in systems of stochastic differential equations. *Biometrika*, 46(1/2):67–76.

Protter, P. E. (2004). *Stochastic Integration and Differential Equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition.

Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge.

Schoutens, W. (2003). *Lévy Processes in Finance: Pricing Financial Derivatives*. Wiley.

Štatland, E. S. (1965). On local properties of processes with independent increments. *Theory of Prob. and Applications*, 10:317–322.

Stramer, O. and Roberts, G. (2004). On Bayesian analysis of non-linear continuous-time autoregression models. Technical Report 327, Department of Statistics and Actuarial Science, The University of Iowa.

Stroock, D. W. and Varadhan, S. R. S. (1979). *Multidimensional Diffusion Processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, Berlin.

Todorov, V. (2006). Econometric analysis of jump-driven stochastic volatility models. Submitted to *J. Econometrics.*

Todorov, V. and Tauchen, G. (2006). Simulation methods for Lévy-driven CARMA stochastic volatility models. *J. Bus. Econ. Statist.*, 24:455 – 469.

Tong, H. (1983). *Threshold Models in Nonlinear Time Series Analysis*, volume 21 of *Lecture Notes in Statistics.* Springer-Verlag, New York.

Tong, H. and Yeung, I. (1991). Threshold autoregressive modelling in continuous time. *Statist. Sinica*, 1(2):411–430.

Tsai, H. and Chan, K. (1999). A new EM method for estimating continuous-time autoregressive models. Technical Report 285, Department of Statistics and Actuarial Science, The University of Iowa.

Tsai, H. and Chan, K. S. (2000). Comparison of two discretization methods for estimating continuous-time autoregressive models. In Chan, W.-S., Li, W. K., and Tong, H., editors, *Statistics and Finance: An Interface*, pages 68–85. Imp. Coll. Press, London.

Tsai, H. and Chan, K. S. (2005). A note on non-negative continuous time processes. *J. R. Stat. Soc. Ser. B*, 67(4):589–597.

Tuan, P. D. (1977). Estimation of parameters of a continuous time Gaussian stationary process with rational spectral density. *Biometrika*, 64(2):385–399.