Some properties of continuous-time autoregressive moving average time series models

Alet Roux

Submitted in partial fulfilment of the requirements of the degree

Baccalaureus Commercii (Honores)

in

Mathematical Statistics

in the

Faculty of Economic and Management Sciences
University of Pretoria
Pretoria

under the supervision of

Dr. H. Boraine

August 2002

Aan dié persone wie ek dank verskuldig is, spreek ek my hartlike dank uit. Behalwe diverse sensors en kwaadstekers sal die ander seker in elk geval nie hierdie woorde lees nie.

Breyten Breytenbach

I greatly admire my supervisor, Dr. H. Boraine, for her tremendous courage. I am also indebted to my colleagues at the University of Pretoria for their continuing support.

Macta nova virtute, [puella], sic itur ad astra.

Publius Virgil

I hereby declare that the essay submitted hereby in partial fulfilment of the requirements of the degree Baccalaureus Commercii (Honores) at the University of Pretoria contains only my own independent work, and has never before been submitted for any degree at any university.

Alet Roux
Pretoria, August 28, 2002

Contents

Introduction

> The continuous-discrete time series problem is concerned with fitting continuous time series models to discrete time series data. There are a number of reasons for wanting to do this. Firstly, to provide forecasts or interpolates between observations; secondly, to provide estimates of signal derivatives; thirdly, to deal with unequally spaced data; fourthly, as an alternative to fitting splines; fifthly, as part of a continuous time control design scheme. *Solo (1984, p. 326)*

The purpose of this study is to investigate some qualitative and quantitative properties of the class of continuous-time autoregressive moving average (CARMA) time series models, which are defined as follows. If $\{\varepsilon(t)\}_{t\in[0,\infty)}$ is a continuous-time white noise process, and $\alpha_1, \alpha_2, \ldots, \alpha_p$ and $\beta_1, \beta_2, \ldots, \beta_q$ are real constants, the process $\{Y(t)\}_{t\in[0,\infty)}$ satisfying (at least formally) the stochastic differential equation

$$\alpha_p Y(t) + \alpha_{p-1} Y'(t) + \cdots + \alpha_1 Y^{(p-1)}(t) + Y^{(p)}(t) = \varepsilon(t) + \beta_1 \varepsilon'(t) + \beta_2 \varepsilon''(t) + \cdots + \beta_q \varepsilon^{(q)}(t),$$

is called a continuous-time autoregressive moving average process of order (p, q) (or, in short, a CARMA (p, q) process).

Due to their linear specification, CARMA models result in tractable likelihoods for observed discrete-time data. Hence this method is rather popular for the analysis of irregularly sampled time series data, as is often encountered in financial applications and control problems (cf. Robinson, 1977; Jones, 1981; Khabie-Zeitoune, 1982; Pandit and Wu, 1983; Jones, 1984; Masry, 1984; Jones, 1985; Harvey, 1989; Bergstrom, 1990; Tong, 1990; Hyndman, 1993; Jones, 1993; Belcher, Hampton and Tunnicliffe Wilson, 1994; Brockwell, 2000; Tsai and Chan, 2000). Attention has also been paid (by Robinson (1980) and Masry (1997)) to the case where the observation times themselves form stationary point processes. Bartlett (1955), Dzhaparidze (1971) and Priestley (1981) have also presented some results on the estimation of the model from a continuous record.

The application of CARMA models to time series with regularly sampled data, a field conventionally considered the exclusive domain of discrete-time methodology, has also received considerable attention (cf. Bartlett, 1946; Durbin, 1961; Telser, 1967; Hannan, 1970;

Phillips, 1972; Pandit and Wu, 1975; Jones, 1981; Bergstrom, 1984; Solo, 1984). In addition, the embedding of discrete-time autoregressive moving average (DARMA) models into their continuous-time counterparts has also been the subject of a number of papers; the intertextual discourse between Chan and Tong (1987), Liu (1988), He and Wang (1989) and Brockwell (1995) serves as illustration in this regard.

In general, the properties of CARMA processes are well-documented, and have been known for a considerable period of time (cf. Doob, 1953). From a statistical point of view, much insight is to be gained from the study of these properties, without having to delve too deep into serious measure theory.

In Chapter 1, we consider two often-used continuous-time stochastic processes, namely continuous-time white noise and Brownian motion. The most prominent result in this chapter is the additivity property of Brownian motion, expressed as its relationship with Gaussian continuous-time white noise. The presentation of the material covered in this chapter is deliberately non-measure theoretical, with the emphasis being on the development of ideas and concepts for use in the remainder of this work.

Chapter 2 is concerned with the general linear diffusion process, which is defined as the solution of the stochastic differential equation

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

with $\sigma \in [0,\infty)$ being a constant, $\mathbf{A}$ a $p \times p$ matrix, $\mathbf{e}$ a p-dimensional vector and $\{\varepsilon(t)\}_{t\in[0,\infty)}$ a Gaussian continuous-time white noise process, under suitable initial conditions. Section 2.1 constitutes a brief introduction to the Itô integral, which is one of the canonical forms of the stochastic integral (the other being the Stratonovich integral). Some elementary properties of this integral are also presented. In Section 2.2, the interpretation of the above stochastic differential equation is given in terms of Itô calculus. We prove that its solution exists, and also obtain explicit formulae for the first and second order moments of this solution. The chapter concludes with some necessary and sufficient conditions for the weak stationarity of the general linear diffusion process.

The main results of this study are presented in Chapter 3. In Section 3.1, the CARMA process is introduced, and its first and second-order moments and stationarity properties are discussed. Having established that the continuous-time autoregressive (CAR) process is a special CARMA process, Section 3.2 is devoted to establishing the properties of CAR processes. Likewise, the continuous-time moving average (CMA) process forms the focus of Section 3.3.

No attention is paid to the application of these models to actual time series. Several (very involved) methods for estimation of parameters and prediction of observations exist, e.g. the use of Kalman filters (cf. Kalman and Bucy, 1961) and Yule-Walker equations (cf.

Hyndman, 1993). The amount of work necessary for a proper investigation of these methods is of great enough magnitude to warrant an essay of this kind in itself.

Chapter 1

Foundation processes

In this chapter, we study two often-used continuous-time stochastic processes which are inextricably linked, namely, continuous-time white noise and Brownian motion. In Section 1.1, we develop the notion of a continuous-time white noise process in analogy to the well-known discrete-time white noise process. In Section 1.2, we first define Brownian motion in the conventional way, and then investigate its specific relationship with Gaussian continuous-time white noise.

1.1 Continuous-time white noise

We first consider the well-known discrete-time white noise process (cf. Cryer, 1986, pp. 15-16).

Definition 1.1.1 (Discrete-time white noise) *A discrete-time stochastic process $\{\varepsilon_n\}_{n\in\mathbb{N}_0}$ taking values in $\mathbb{R}$ is called a discrete-time white noise process if it has the following properties:*

1. *For $n \in \mathbb{N}_0$,*

$$\mathrm{E}\left(\varepsilon_n\right) = 0. \tag{1.1.1}$$

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

$$\operatorname{cov}\left(\varepsilon_m, \varepsilon_n\right) = \delta_{m-n}\sigma_\varepsilon^2 \textit{ for all } m \in \mathbb{N}_0 \textit{ and } n \in \mathbb{N}_0, \tag{1.1.2}$$

   where the Kronecker delta function $\delta : \mathbb{R} \to \{0, 1\}$ is defined as

$$\delta_m := \begin{cases} 1 & \textit{if } m = 0 \\ 0 & \textit{if } m \neq 0 \end{cases}. \tag{1.1.3}$$

It seems reasonable to expect that a continuous-time white noise process should have similar first and second order moments to a discrete-time white noise process. However, there is one startling difference concerning its variance. We illustrate this by contradiction, essentially following the same line of reasoning as Priestley (1981, Section 3.7.1). Suppose that $\{\varepsilon(t)\}_{t\in[0,\infty)}$ is a continuous-time real-valued stochastic process with expected value

$$\mathrm{E}\left(\varepsilon\left(t\right)\right)=0 \text{ for } t\in[0,\infty) \tag{1.1.4}$$

and, for a constant $\sigma_\varepsilon^2\in[0,\infty)$, covariance

$$\operatorname{cov}\left(\varepsilon\left(s\right),\varepsilon\left(t\right)\right)=\delta_{t-s}\sigma_\varepsilon^2 \text{ for all } s\in[0,\infty) \text{ and } t\in[0,\infty)\,. \tag{1.1.5}$$

Remark 1.1.2 *We assume, in addition, that $\{\varepsilon(t)\}_{t\in[0,\infty)}$ is integrable with respect to time as well as with respect to its probability space. The purpose of this assumption is to satisfy the premises of Fubini's Theorem (cf. Adams and Guillemin, 1996, Section 2.3), which guarantees that integrals with respect to probability spaces (e.g. expected values, covariances) can be interchanged with integrals with respect to time.*

Let us construct a random variable U as some linear combination of $\{\varepsilon(t)\}_{t\in[0,\infty)}$, i.e. for numbers $a\in[0,\infty)$ and $b\in(a,\infty)$ and a square integrable deterministic function $g:[a,b]\to\mathbb{R}$, U is defined as

$$U:=\int_a^b g(t)\varepsilon\left(t\right)\,dt. \tag{1.1.6}$$

The first and second order moments of U are

$$\begin{aligned}
\mathrm{E}\left(U\right)&=\mathrm{E}\left(\int_a^b g(t)\varepsilon\left(t\right)\,dt\right) && \text{cf. (1.1.6)}\\
&=\int_a^b g(t)\mathrm{E}\left(\varepsilon\left(t\right)\right)\,dt && \text{Remark 1.1.2}\\
&=0 && \text{cf. (1.1.4)}
\end{aligned} \tag{1.1.7}$$

and

$$\begin{aligned}
\operatorname{var}\left(U\right)&=\operatorname{cov}\left(\int_a^b g(s)\varepsilon\left(s\right)\,ds,\int_a^b g(t)\varepsilon\left(t\right)\,dt\right) && \text{cf. (1.1.6)}\\
&=\int_a^b\int_a^b g(s)g(t)\operatorname{cov}\left(\varepsilon\left(s\right),\varepsilon\left(t\right)\right)\,ds\,dt && \text{Remark 1.1.2}\\
&=\int_a^b\int_a^b g(s)g(t)\delta_{t-s}\sigma_\varepsilon^2\,ds\,dt && \text{cf. (1.1.5)}\\
&=0. && \text{cf. (1.1.3)}
\end{aligned} \tag{1.1.8}$$

Equations (1.1.7) and (1.1.8) imply that any random variable U of the form (1.1.6) takes the constant value zero almost surely. Hence $\{\varepsilon(t)\}_{t\in[0,\infty)}$ is degenerate in the sense that it is of little use in generating non-trivial stochastic processes.

The only way to obviate this fundamental difficulty is to require the second order moments of a continuous-time white noise process to satisfy

$$\operatorname{cov}(\varepsilon(s), \varepsilon(t)) = \delta(t-s)\,\sigma_\varepsilon^2 \text{ for all } s \in \mathbb{R} \text{ and } t \in \mathbb{R}, \tag{1.1.9}$$

where δ is the Dirac mass function (cf. Pandit and Wu, 1983, Section 6.2) satisfying

$$\int_{\mathbb{R}} g(t)\delta(t)\ dt = g(0) \text{ for all integrable functions } g : \mathbb{R} \to \mathbb{R} \tag{1.1.10}$$

$$\delta(t) = 0 \text{ for } t \neq 0. \tag{1.1.11}$$

Remark 1.1.3 *The Dirac mass is a so-called singular distribution (cf. Sauer, 2001, p. 3), and therefore a member of the class of generalized functions on $\mathbb{R}$. In fact, it is possible (cf. Sauer, 2001, p. 6) to construct the Dirac mass as the pointwise limit of a sequence of functions that are all infinitely differentiable.*

Remark 1.1.4 *Some authors (e.g. Pandit and Wu, 1983; Jones, 1985) prefer to write* (1.1.10) *and* (1.1.11) *(albeit formally) as*

$$\int_{\mathbb{R}} \delta(t)\, dt = 1 \textit{ and } \delta(t) = \begin{cases} \infty & \textit{if } t = 0 \\ 0 & \textit{if } t \neq 0 \end{cases}. \tag{1.1.12}$$

In view of (1.1.12)*, the variance of a continuous-time white noise process can intuitively be regarded as infinite (cf. Chatfield, 1996, pp. 43-44).*

Indeed, if $\{\varepsilon(t)\}_{t\in[0,\infty)}$ satisfies (1.1.9), then U in (1.1.6) has the property

$$\begin{aligned}
\operatorname{var}(U) &= \operatorname{cov}\left(\int_a^b g(s)\varepsilon(s)\ ds, \int_a^b g(t)\varepsilon(t)\ dt\right) && \text{cf. (1.1.6)} \\
&= \int_a^b \int_a^b g(s)g(t)\operatorname{cov}(\varepsilon(s), \varepsilon(t))\ ds\,dt && \text{Remark 1.1.2} \\
&= \int_a^b \int_a^b g(s)g(t)\delta(t-s)\,\sigma_\varepsilon^2\, ds\,dt && \text{cf. (1.1.9)} \\
&= \sigma_\varepsilon^2 \int_a^b (g(t))^2\ dt. && \text{cf. (1.1.10), (1.1.11)} \qquad (1.1.13)
\end{aligned}$$

Clearly, $\text{var}(U) \geq 0$, and $\text{var}(U) = 0$ if and only if g is almost surely zero on the interval $[a, b]$.

It turns out that there does not exist any "reasonable" stochastic process satisfying (1.1.4) and (1.1.9) (cf. Øksendal, 2000). Nevertheless, it is possible to represent $\{\varepsilon(t)\}_{t\in[0,\infty)}$ as a generalized stochastic process called a continuous-time white noise process.

Definition 1.1.5 (Continuous-time white noise) *A continuous-time white noise process* $\{\varepsilon(t)\}_{t\in[0,\infty)}$ *is defined to be the real-valued continuous-time stochastic process satisfying* (1.1.4) *and* (1.1.9).

Remark 1.1.6 *Kallianpur and Karandikar (1988, Chapter 1) follow a more rigorous approach to the definition of continuous-time white noise. Having shown that a continuous-time white noise process cannot be measurable (Remark 2, p. 24), neither can its sample paths be continuous (Remark 1, p. 24), they conclude that the best definition of a continuous-time white noise process is "that stationary process which has constant spectral density" (p. 26). They also show (p. 25–26) that this definition of a continuous-time white noise process leads to* (1.1.4) *and* (1.1.9)*. It is therefore consistent with Definition 1.1.5.*

Remark 1.1.7 *In the remainder of this chapter, we assume a continuous-time white noise process to be integrable with respect to time as well as with respect to its probability space (see Remark 1.1.2). We also need to assume that, at the initial time* ($t = 0$)*, the sample path of such a continuous-time white noise process has a finite value almost surely.*

1.2 Brownian motion

We now turn our attention to a stochastic process which is a prominent member of the class of Wiener processes. Priestley (1981, p. 167) describes Brownian motion as being the

> . . . well known phenomenon in physics which describes the random movement of microscopic particles suspended in a liquid or gas. It was first observed experimentally by Robert Brown in 1827, and is due to the impact on the particles of randomly moving molecules of the liquid or gas. . . . There is a substantial literature on the subject in the context of statistical mechanics—starting with the pioneering work of Einstein. . .

In this work, we will use one-dimensional real-valued Brownian motion, which is defined as follows.

Definition 1.2.1 (Brownian motion) *A real-valued continuous-time stochastic process* $\{W(t)\}_{t\in[0,\infty)}$ *is called Brownian motion if it satisfies the following:*

1. *We have*

$$W(0) = 0 \text{ almost surely.} \tag{1.2.1}$$

2. *There exists a constant* $\sigma_W^2 \in [0, \infty)$*, called the variance of* $\{W(t)\}_{t\in[0,\infty)}$*, such that*

$$W(t) \sim N\big(0, \sigma_W^2 t\big) \text{ for all } t \in (0, \infty), \tag{1.2.2}$$

   i.e. $W(t)$ *has a normal distribution with mean* 0 *and variance* $\sigma_W^2 t$*.*

3. *The process* $\{W(t)\}_{t\in[0,\infty)}$ *has stationary increments, i.e. for all* $s > 0$ *and* $t > 0$*, the distribution of* $W(t+s) - W(t)$ *is independent of* t*.*

4. *The process* $\{W(t)\}_{t\in[0,\infty)}$ *has independent increments, i.e.* $W(t_2) - W(t_1)$ *and* $W(t_3) - W(t_2)$ *are independent for any real numbers* t_1, t_2 *and* t_3 *satisfying*

$$0 \le t_1 \le t_2 \le t_3. \tag{1.2.3}$$

The existence and uniqueness of Brownian motion has been proven conclusively (see, for instance, Karatzas and Shreve (1988, Section 2.1-2.2) and Rogers and Williams (1994*a*, Theorem I.6.1)). In addition, it is related to the class of Gaussian continuous-time white noise processes in a very specific way.

We first define a Gaussian continuous-time stochastic process.

Definition 1.2.2 (Gaussian continuous-time stochastic process) *A continuous-time stochastic process* $\{X(t)\}_{t\in[0,\infty)}$ *is called Gaussian if* $\int_0^t X(u)\,du$ *is normally distributed for all* $t \in (0,\infty)$.

The next result is is an adaptation of a discussion by Priestley (1981, pp. 161-163).

Theorem 1.2.3 *Suppose that* $\{\varepsilon(t)\}_{t\in[0,\infty)}$ *is a Gaussian continuous-time white noise process. If the continuous-time stochastic process* $\{W(t)\}_{t\in[0,\infty)}$ *is defined by*

$$W(t) := \int_0^t \varepsilon(u)\ du \text{ for } t \in [0,\infty), \tag{1.2.4}$$

then $\{W(t)\}_{t\in[0,\infty)}$ *is a Brownian motion.*

Proof. We show that $\{W(t)\}_{t\in[0,\infty)}$ satisfies all the properties of Brownian motion set out in Definition 1.2.1:

1. Since $\int_0^0 \varepsilon(u)\ du = 0$ almost surely (see Remark 1.1.7), Property 1 is satisfied.

2. For $t \in (0,\infty)$, we have

$$\begin{aligned}
\mathrm{E}(W(t)) &= \mathrm{E}\left(\int_0^t \varepsilon(u)\ du\right) && \text{cf. (1.2.4)}\\
&= \int_0^t \mathrm{E}(\varepsilon(u))\ du && \text{Remark 1.1.7}\\
&= 0. && \text{cf. (1.1.4)} \qquad (1.2.5)
\end{aligned}$$

Moreover, there exists a constant $\sigma_\varepsilon^2 \in [0,\infty)$ such that, for $s \in (0,\infty)$ and $t \in (0,\infty)$,

$$\begin{aligned}
\mathrm{cov}(W(s), W(t)) &= \mathrm{cov}\left(\int_0^{\min\{s,t\}} \varepsilon(u)\ du, \int_0^{\max\{s,t\}} \varepsilon(v)\ dv\right) && \text{cf. (1.2.4)}\\
&= \int_0^{\min\{s,t\}} \int_0^{\max\{s,t\}} \mathrm{cov}(\varepsilon(u), \varepsilon(v))\ du\,dv && \text{Remark 1.1.7}\\
&= \int_0^{\min\{s,t\}} \int_0^{\max\{s,t\}} \delta(u-v)\sigma_\varepsilon^2\,du\,dv && \text{cf. (1.1.9)}\\
&= \int_0^{\min\{s,t\}} \sigma_\varepsilon^2\,dv && \text{cf. (1.1.10)}\\
&= \sigma_\varepsilon^2 \min\{s,t\}. && (1.2.6)
\end{aligned}$$

In particular,

$$\mathrm{var}\left(W\left(t\right)\right)=\sigma_\varepsilon^2 t \text{ for } t\in[0,\infty)\,. \tag{1.2.7}$$

In the light of the fact that $\{\varepsilon(t)\}_{t\in[0,\infty)}$ is Gaussian, equations (1.2.5) and (1.2.7) imply that Property 2 is satisfied, with the variance σ_W^2 in Definition 1.2.1 being equal to σ_ε^2.

3. For all $s\in(0,\infty)$ and $t\in(0,\infty)$, $W\left(t+s\right)-W\left(t\right)$ is normally distributed, since it is the difference between the normal variates $W\left(t+s\right)$ and $W\left(t\right)$ (cf. Bain and Engelhardt, 1991, Example 6.4.7). The expected value and variance of $W\left(t+s\right)-W\left(t\right)$ are given, respectively, by

$$\begin{aligned}\mathrm{E}\left(W\left(t+s\right)-W\left(t\right)\right)&=\mathrm{E}\left(W\left(t+s\right)\right)-\mathrm{E}\left(W\left(t\right)\right)\\&=0 \qquad\qquad \text{cf. (1.2.5)}\end{aligned} \tag{1.2.8}$$

and

$$\begin{aligned}\mathrm{var}\left(W\left(t+s\right)-W\left(t\right)\right)&=\mathrm{var}\left(W\left(t+s\right)\right)-2\mathrm{cov}\left(W\left(t+s\right),W\left(t\right)\right)+\mathrm{var}\left(W\left(t\right)\right)\\&=\sigma_\varepsilon^2(t+s)-2\sigma_\varepsilon^2 t+\sigma_\varepsilon^2 t \qquad \text{cf. (1.2.6), (1.2.7)}\\&=\sigma_\varepsilon^2 s.\end{aligned} \tag{1.2.9}$$

Equations (1.2.8) and (1.2.9) permit us to conclude that the distribution of $W\left(t+s\right)-W\left(t\right)$ is independent of t. Hence $\{W(t)\}_{t\in[0,\infty)}$ has stationary increments.

4. For t_1, t_2 and t_3 satisfying (1.2.3), we have

$$\begin{aligned}&\mathrm{cov}\left(W\left(t_2\right)-W\left(t_1\right),W\left(t_3\right)-W\left(t_2\right)\right)\\&=\mathrm{cov}\left(W\left(t_1\right),W\left(t_2\right)\right)-\mathrm{cov}\left(W\left(t_1\right),W\left(t_3\right)\right)-\mathrm{var}\left(W\left(t_2\right)\right)+\mathrm{cov}\left(W\left(t_2\right),W\left(t_3\right)\right)\\&=\sigma_\varepsilon^2 t_1-\sigma_\varepsilon^2 t_1-\sigma_\varepsilon^2 t_2+\sigma_\varepsilon^2 t_2 \qquad \text{cf. (1.2.6), (1.2.7)}\\&=0.\end{aligned} \tag{1.2.10}$$

Thus $W\left(t_2\right)-W\left(t_1\right)$ and $W\left(t_3\right)-W\left(t_2\right)$ are uncorrelated. Since both $W\left(t_2\right)-W\left(t_1\right)$ and $W\left(t_3\right)-W\left(t_2\right)$ are normal random variables, it follows that they are independent (cf. Bain and Engelhardt, 1991, Section 5.4).

■

Chapter 2

A general multivariate stochastic process

This chapter is devoted to the study of the general class of continuous-time stochastic processes to which CARMA processes belong. In Section 2.1, we provide some background information on the Itô integral, while we develop notation and conventions for use in the remainder of this work. In Section 2.2, we turn our attention to the general linear multivariate diffusion process, and investigate its qualitative and quantitative properties.

2.1 The Itô integral

In this section, we briefly define and describe the properties of the Itô integral, which is the most frequently used form of the stochastic integral. We also use the definition and properties of the univariate Itô integral to define and characterize its multivariate counterpart.

Definition 2.1.1 (Itô integral) *For any real-valued continuous-time stochastic process $\{Z(t)\}_{t\in[0,\infty)}$ with continuous sample paths (i.e. Z is a continuous function of time), the Itô integral, if it exists, is defined as the mean square limit*

$$\int_0^\infty Z(u)\,dW(u) := \lim_{n\to\infty,\triangle t\to 0} \sum_{k=0}^{n-1} Z(k\triangle t)\left(W\left((k+1)\triangle t\right) - W\left(k\triangle t\right)\right). \tag{2.1.1}$$

For a given upper limit $t \in [0,\infty)$, the Itô integral on $[0,t]$ is defined as the mean square limit

$$\int_0^t Z(u)\,dW(u) := \lim_{n\to\infty,\triangle t\to 0, n\triangle t = t} \sum_{k=0}^{n-1} Z(k\triangle t)\left(W\left((k+1)\triangle t\right) - W\left(k\triangle t\right)\right). \tag{2.1.2}$$

For a continuous deterministic function $Z : [0,t] \to \mathbb{R}$, we can compare the form of the Itô integral (2.1.2) with the well-known Riemann integral

$$\int_0^t Z(u)\,du := \lim_{n\to\infty,\triangle t\to 0, n\triangle t = t} \sum_{k=0}^{n-1} Z(k\triangle t)\left((k+1)\triangle t - k\triangle t\right). \tag{2.1.3}$$

If $W : [0,t] \to \mathbb{R}$ is simply assumed to be a differentiable deterministic function, then we can also compare (2.1.2) with the Riemann-Stieltjes integral

$$\int_0^t Z(u)\,dW(u) := \lim_{n\to\infty,\triangle t\to 0, n\triangle t = t} \sum_{k=0}^{n-1} Z(k\triangle t)\left(W\left((k+1)\triangle t\right) - W\left(k\triangle t\right)\right). \tag{2.1.4}$$

The main difference between the Itô integral (2.1.2) and the integrals (2.1.3) and (2.1.4) is that, in (2.1.3) and (2.1.4), convergence takes place in $\mathbb{R}$, while the mean square limit in (2.1.2) is again a random variable.

We now have an important result concerning the existence and uniqueness of the Itô integral, as well as its properties.

Theorem 2.1.2 *If a real-valued continuous-time stochastic process $\{Z(t)\}_{t\in[0,\infty)}$ with continuous sample paths is square integrable on an interval $[0,t]$, i.e.*

$$\mathrm{E}\left(\int_0^t (Z(u))^2\,du\right) < \infty, \tag{2.1.5}$$

then its Itô integral on $[0,t]$ exists and is unique. In addition, the Itô integral on $[0,t]$ has the following properties:

1. *It is linear in its argument, i.e. for $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$ and stochastic processes $\{Y(t)\}_{t\in[0,\infty)}$ and $\{Z(t)\}_{t\in[0,\infty)}$ satisfying (2.1.5), we have*

$$\int_0^t \left(\alpha Y(u) + \beta Z(u)\right)\,dW(u) = \alpha\int_0^t Y(u)\,dW(u) + \beta\int_0^t Z(u)\,dW(u). \tag{2.1.6}$$

2. *It has the Itô isometry property, i.e. for $\{Y(t)\}_{t\in[0,\infty)}$ and $\{Z(t)\}_{t\in[0,\infty)}$ satisfying (2.1.5), we have*

$$\mathrm{E}\left(\left(\int_0^t Y(u)\,dW(u)\right)\left(\int_0^t Z(u)\,dW(u)\right)\right) = \mathrm{E}\left(\int_0^t Y(u)Z(u)\,du\right). \tag{2.1.7}$$

3. *It has the martingale property, i.e. for $\{Z(t)\}_{t\in[0,\infty)}$ satisfying (2.1.5), we have*

$$\mathrm{E}\left(\int_0^t Z(u)\,dW(u) \,\middle|\, W(s)\right) = \int_0^s Z(u)\,dW(u) \ \text{ for } s\in[0,t). \tag{2.1.8}$$

Proof. This result is a consequence of the fact that every stochastic process with continuous sample paths can be expressed as the mean square limit of a sequence of step processes, i.e. stochastic processes taking constant values on a finite number of subsets of $[0,\infty)$. The complete proof is given by Brzeżniak and Zastawniak (1999, Section 7.1, Theorem 7.3) ■

The Itô integral is also forced to be additive in its limits by the following definition.

Definition 2.1.3 *We define the Itô integral over* $[s,t]$*, where* $s \in [0,\infty]$ *and* $t \in [0,\infty]$*, of a stochastic process* $\{Z(t)\}_{t\in[0,\infty)}$ *satisfying* (2.1.5) *as*

$$\int_s^t Z(u)\,dW(u) := \int_0^t Z(u)\,dW(u) - \int_0^s Z(u)\,dW(u)\,. \tag{2.1.9}$$

We now turn to the definition of Itô integrals of multivariate stochastic processes.

Definition 2.1.4 (Multivariate Itô integral (vector)) *For* $k = 1,2,\ldots,p$*, suppose that* $\{X_k(t)\}_{t\in[0,\infty)}$ *are real-valued continuous-time stochastic processes with continuous sample paths satisfying* (2.1.5)*. If the* p*-variate stochastic process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *is defined by*

$$\mathbf{X}(t) := \begin{bmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_p(t) \end{bmatrix} \quad \text{for } t \in [0,\infty)\,, \tag{2.1.10}$$

then we define its Itô integral on $[0,t]$ *as*

$$\int_0^t \mathbf{X}(u)\; dW(u) := \begin{bmatrix} \int_0^t X_1(u)\,dW(u) \\ \int_0^t X_2(u)\,dW(u) \\ \vdots \\ \int_0^t X_p(u)\,dW(u) \end{bmatrix}. \tag{2.1.11}$$

The Itô integral of matrix-valued stochastic processes is defined in exactly the same fashion.

Definition 2.1.5 (Multivariate Itô integral (matrix)) *Suppose that* $\{X_{kl}(t)\}_{t\in[0,\infty)}$*, for* $k = 1,2,\ldots,p$ *and* $l = 1,2,\ldots,q$*, are real-valued continuous-time stochastic processes with continuous sample paths satisfying* (2.1.5)*. If the* pq*-variate stochastic process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *is defined by*

$$\mathbf{X}(t) := \begin{bmatrix} X_{11}(t) & X_{12}(t) & \cdots & X_{1q}(t) \\ X_{21}(t) & X_{22}(t) & \cdots & X_{2q}(t) \\ \vdots & \vdots & \ddots & \vdots \\ X_{p1}(t) & X_{p2}(t) & \cdots & X_{pq}(t) \end{bmatrix} \quad \text{for } t \in [0,\infty)\,, \tag{2.1.12}$$

then we define its Itô integral on $[0,t]$ *as*

$$\int_0^t \mathbf{X}(u)\ dW(u) := \begin{bmatrix} \int_0^t X_{11}(u)\,dW(u) & \int_0^t X_{12}(u)\,dW(u) & \cdots & \int_0^t X_{1q}(u)\,dW(u) \\ \int_0^t X_{21}(u)\,dW(u) & \int_0^t X_{22}(u)\,dW(u) & \cdots & \int_0^t X_{2q}(u)\,dW(u) \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^t X_{p1}(u)\,dW(u) & \int_0^t X_{p2}(u)\,dW(u) & \cdots & \int_0^t X_{pq}(u)\,dW(u) \end{bmatrix}. \tag{2.1.13}$$

We define the Riemann integrals of deterministic vector- and matrix-valued functions in obvious analogy.

The following result regarding multivariate Itô integrals, which we will use frequently, follows immediately from Theorem 2.1.2.

Theorem 2.1.6 *The Itô integrals* (2.1.11) *and* (2.1.13) *exist and are both unique. The Itô integral on* $[0,t]$ *of a p-variate stochastic process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *such as in* (2.1.10) *or a pq-variate stochastic process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *such as in* (2.1.12) *has the following properties:*

1. *It is linear.*

2. *It has the Itô isometry property, in the sense that*

$$\mathrm{E}\left(\left(\int_0^t \mathbf{X}(u)\ dW(u)\right)\left(\int_0^t \mathbf{X}(u)\ dW(u)\right)^{\mathrm{T}}\right) = \mathrm{E}\left(\int_0^t \mathbf{X}(u)\,(\mathbf{X}(u))^{\mathrm{T}}\,du\right). \tag{2.1.14}$$

3. *It has the martingale property, i.e.*

$$\mathrm{E}\left(\int_0^t \mathbf{X}(u)\ dW(u)\,\middle|\, W(s)\right) = \int_0^s \mathbf{X}(u)\ dW(u)\ \text{ for } s \in [0,t). \tag{2.1.15}$$

Once again (in similar fashion to Definition 2.1.3), the multivariate Itô integral can be forced to be additive in its limits.

Definition 2.1.7 *We define the Itô integral over* $[s,t]$*, where* $s \in [0,\infty]$ *and* $t \in [0,\infty]$*, of a p-variate stochastic process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *such as in* (2.1.10)*, or a* p^2*-variate stochastic process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *such as in* (2.1.12)*, as*

$$\int_s^t \mathbf{X}(u)\ dW(u) := \int_0^t \mathbf{X}(u)\ dW(u) - \int_0^s \mathbf{X}(u)\ dW(u). \tag{2.1.16}$$

2.2 The general multivariate linear diffusion process

We now turn our attention to the p-variate continuous-time stochastic process $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ satisfying (in the notation of Section 1.2)

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t)\ dt + \sigma\mathbf{e}\,\varepsilon(t)\ \text{ for } t \in [0,\infty)\,, \tag{2.2.1}$$

where $\sigma \in [0,\infty)$ is a constant, $\mathbf{A}$ is a $p\times p$ matrix, $\mathbf{e}$ is a p-dimensional vector and $\{\varepsilon(t)\}_{t\in[0,\infty)}$ is a Gaussian continuous-time white noise process. Moreover, we assume that

$$\mathbf{X}(0)\ \text{ is normally distributed and uncorrelated with } W(t)\ \text{ for all } t \in (0,\infty)\,. \tag{2.2.2}$$

In the light of Theorem 1.2.3, we refer in the sequel to a Gaussian continuous-time white noise process $\{\varepsilon(t)\}_{t\in[0,\infty)}$ as $\{dW(t)\}_{t\in[0,\infty)}$, where $\{W(t)\}_{t\in[0,\infty)}$ is the associated Brownian motion with variance σ_W^2, as defined in (1.2.4). We also restrict our attention to standard Brownian motion, which is the special case of Brownian motion having $\sigma_W^2 = 1$ in (1.2.2).

Employing conventional notation in (2.2.1), the process under consideration is the p-variate stochastic process $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ satisfying the linear stochastic differential equation

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t)\ dt + \sigma\mathbf{e}\,dW(t)\ \text{ for } t \in [0,\infty)\,. \tag{2.2.3}$$

Remark 2.2.1 *In the terminology of Rogers and Williams (1994*b*, Chapter V), the stochastic process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *satisfying* (2.2.3) *is called an* $(\mathbf{A},\sigma\mathbf{e})$ *diffusion.*

Some clarification is necessary as to the exact interpretation of (2.2.3), which, considered on its own, is actually meaningless. In principle this is due to the term $dW(t)$ on the right-hand side: as the sample paths of Brownian motion have infinite variation (cf. Rogers and Williams, 1994*b*, Lemma IV.2.16), Brownian motion is not differentiable (cf. Brzeżniak and Zastawniak, 1999, Theorem 6.6), and therefore $dW(t)$ does not exist. In fact, (2.2.3) is considered to be a formal representation of the Itô integral equation

$$\mathbf{X}(t) = \mathbf{X}(0) + \int_0^t \mathbf{A}\mathbf{X}(u)\ du + \sigma\int_0^t \mathbf{e}\,dW(u)\ \text{ for } t \in [0,\infty)\,. \tag{2.2.4}$$

We can now proceed to find the solution of (2.2.3) and (2.2.4).

Remark 2.2.2 *A number of results in this section depend on the Itô formula, an important result in stochastic calculus which we will not consider here. Statements and proofs of the univariate version of this formula can be found in the books by Brzeżniak and Zastawniak (1999, Theorems 7.5–7.6) (simplified version) and Rogers and Williams (1994*b*, Theorem IV.18.4). The multivariate version of the Itô formula is considered in the works by Øksendal (1989, p. 33), Kallianpur (1980, Section 4.5) and Rogers and Williams (1994*b*, Theorem IV.18.8).*

Theorem 2.2.3 *The solution* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *of the stochastic differential equation* (2.2.3) *exists and satisfies*

$$\mathbf{X}(t) = e^{\mathbf{A}t}\mathbf{X}(0) + \sigma \int_0^t e^{\mathbf{A}(t-u)}\mathbf{e}\, dW(u) \ \text{ for } t \in [0,\infty), \tag{2.2.5}$$

where, for any square matrix $\mathbf{P}$, *the matrix exponential function is defined by*

$$e^{\mathbf{P}} := \sum_{k\in\mathbb{N}_0} \frac{1}{k!}\mathbf{P}^k \ \text{ with } \mathbf{P}^0 := \mathbf{I}. \tag{2.2.6}$$

Proof. If we define the stochastic process $\{\mathbf{Y}(t)\}_{t\in[0,\infty)}$ by

$$\mathbf{Y}(t) := e^{-\mathbf{A}t}\mathbf{X}(t) \ \text{ for } t \in [0,\infty), \tag{2.2.7}$$

then the Itô formula permits us to conclude that

$$\mathbf{Y}(t) = \mathbf{X}(0) + \sigma \int_0^t e^{-\mathbf{A}u}\mathbf{e}\, dW(u) \ \text{ for } t \in [0,\infty). \tag{2.2.8}$$

Thus (2.2.5) holds. ■

The following result follows directly from Theorem 2.2.3 (cf. Brockwell, 2000, p. 4).

Corollary 2.2.4 *The solution of the stochastic differential equation* (2.2.3) *satisfies*

$$\mathbf{X}(t) = e^{\mathbf{A}(t-s)}\mathbf{X}(s) + \sigma \int_s^t e^{\mathbf{A}(t-u)}\mathbf{e}\, dW(u) \ \text{ for } t \in [0,\infty) \ \text{ and } s \in [0,t). \tag{2.2.9}$$

We now present some properties of the solution of (2.2.3).

Theorem 2.2.5 *The solution of* (2.2.3) *has the Markov property (cf. Rogers and Williams, 1994*a*, Definition III.1.1), i.e.*

$$\mathrm{P}\left(\{\mathbf{X}(t+s)\in\mathbf{E}\}|\{\mathbf{X}(u)\}_{u\in[0,t]}\right) = \mathrm{P}\left(\{\mathbf{X}(t+s)\in\mathbf{E}\}|\,\mathbf{X}(t)\right) \ \text{ for } \mathbf{E}\subseteq\mathbb{R}^p, \quad s\in(0,\infty) \ \text{ and } t\in(0,\infty). \tag{2.2.10}$$

Proof. In short, we have

$$\begin{aligned}
&\mathrm{P}\left(\left\{\mathbf{X}(t+s)\in\mathbf{E}\right\}\middle|\left\{\mathbf{X}(u)\right\}_{u\in[0,t]}\right)\\
&=\mathrm{P}\left(\left\{e^{\mathbf{A}}\mathbf{X}(t)+\sigma\int_t^{t+s}e^{\mathbf{A}(t-u)}\mathbf{e}\,dW(u)\in\mathbf{E}\right\}\middle|\,\mathbf{X}(t),\left\{\mathbf{X}(u)\right\}_{u\in[0,t)}\right) && \text{cf. (2.2.9)}\\
&=\mathrm{P}\left(\left\{e^{\mathbf{A}s}\mathbf{X}(t)+\sigma\int_t^{t+s}e^{\mathbf{A}(t-u)}\mathbf{e}\,dW(u)\in\mathbf{E}\right\}\middle|\,\mathbf{X}(t),\left\{W(u)\right\}_{u\in[0,t)}\right) && \text{cf. (2.2.5)}\\
&=\mathrm{P}\left(\left\{e^{\mathbf{A}s}\mathbf{X}(t)+\sigma\int_t^{t+s}e^{\mathbf{A}(t-u)}\mathbf{e}\,dW(u)\in\mathbf{E}\right\}\middle|\,\mathbf{X}(t)\right) && \text{cf. (2.1.2)}\\
&=\mathrm{P}\left(\left\{\mathbf{X}(t+s)\in\mathbf{E}\right\}\middle|\,\mathbf{X}(t)\right). && \text{cf. (2.2.9)} \quad (2.2.11)
\end{aligned}$$

■

The next result provides explicit formulae for the first and second order moments of $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$.

Theorem 2.2.6 *The process $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ in (2.2.3) has expected value function*

$$\mathrm{E}(\mathbf{X}(t))=e^{\mathbf{A}t}\mathrm{E}(\mathbf{X}(0)) \quad \textit{for } t\in[0,\infty) \tag{2.2.12}$$

and covariance function

$$\operatorname{cov}(\mathbf{X}(s),\mathbf{X}(t))=e^{\mathbf{A}s}\operatorname{var}(\mathbf{X}(0))\,e^{\mathbf{A}^{\mathrm{T}}t}+\sigma^2\int_0^{\min\{s,t\}}e^{\mathbf{A}(s-u)}\mathbf{e}\mathbf{e}^{\mathrm{T}}e^{\mathbf{A}^{\mathrm{T}}(t-u)}\,du$$
$$\textit{for } s\in[0,\infty) \textit{ and } t\in[0,\infty). \tag{2.2.13}$$

Proof. Applying the Itô formula yet again, this time to the process $\left\{e^{-\mathbf{A}t}\mathbf{e}W(t)\right\}_{t\in[0,\infty)}$, we obtain

$$\mathbf{e}W(t)=\int_0^t\mathbf{A}e^{\mathbf{A}(t-u)}\mathbf{e}W(u)\,du+\int_0^t e^{\mathbf{A}(t-u)}\mathbf{e}\,dW(u). \tag{2.2.14}$$

For $t\in[0,\infty)$,

$$\begin{aligned}
\mathrm{E}(\mathbf{X}(t))&=\mathrm{E}\left(e^{\mathbf{A}t}\mathbf{X}(0)+\sigma\int_0^t e^{\mathbf{A}(t-u)}\mathbf{e}\,dW(u)\right) && \text{cf. (2.2.5)}\\
&=e^{\mathbf{A}t}\mathrm{E}(\mathbf{X}(0))+\mathrm{E}\left(\int_0^t e^{\mathbf{A}(t-u)}\mathbf{e}\,dW(u)\right)\\
&=e^{\mathbf{A}t}\mathrm{E}(\mathbf{X}(0))+\mathrm{E}\left(\mathbf{e}W(t)-\int_0^t\mathbf{A}e^{\mathbf{A}(t-u)}\mathbf{e}W(u)\,du\right) && \text{cf. (2.2.14)}
\end{aligned}$$

$$= e^{\mathbf{A}t}\mathrm{E}\left(\mathbf{X}\left(0\right)\right) + \mathbf{e}\mathrm{E}\left(W\left(t\right)\right) - \int\limits_0^t \mathbf{A}e^{\mathbf{A}(t-u)}\mathbf{e}\mathrm{E}\left(W\left(u\right)\right)\,du$$

$$= e^{\mathbf{A}t}\mathrm{E}\left(\mathbf{X}\left(0\right)\right). \tag{2.2.15}$$

Thus (2.2.12) holds.

For any $s \in [0, \infty)$ and $t \in [0, \infty)$, we have

$$\begin{aligned}
&\operatorname{cov}\left(\mathbf{X}\left(s\right), \mathbf{X}\left(t\right)\right) \\
&= \mathrm{E}\left(\mathbf{X}\left(s\right)\left(\mathbf{X}\left(t\right)\right)^{\mathrm{T}}\right) - \mathrm{E}\left(\mathbf{X}\left(s\right)\right)\left(\mathrm{E}\left(\mathbf{X}\left(t\right)\right)\right)^{\mathrm{T}} \\
&= \mathrm{E}\left(\left(e^{\mathbf{A}s}\mathbf{X}\left(0\right) + \sigma\int\limits_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\,dW\left(u\right)\right)\left(e^{\mathbf{A}t}\mathbf{X}\left(0\right) + \sigma\int\limits_0^t e^{\mathbf{A}(t-u)}\mathbf{e}\,dW\left(u\right)\right)^{\mathrm{T}}\right) \\
&\quad - e^{\mathbf{A}s}\mathrm{E}\left(\mathbf{X}\left(0\right)\right)\left(\mathrm{E}\left(\mathbf{X}\left(0\right)\right)\right)^{\mathrm{T}}e^{\mathbf{A}^{\mathrm{T}}t} \qquad \text{cf. (2.2.5), (2.2.12)} \\
&= e^{\mathbf{A}s}\operatorname{var}\left(\mathbf{X}\left(0\right)\right)e^{\mathbf{A}^{\mathrm{T}}t} + \sigma e^{\mathbf{A}s}\mathrm{E}\left(\mathbf{X}\left(0\right)\right)\mathrm{E}\left(\left(\int\limits_0^t e^{\mathbf{A}(t-u)}\mathbf{e}\,dW\left(u\right)\right)^{\mathrm{T}}\right) \\
&\quad + \sigma\mathrm{E}\left(\int\limits_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\,dW\left(u\right)\right)\mathrm{E}\left(\left(\mathbf{X}\left(0\right)\right)^{\mathrm{T}}\right)e^{\mathbf{A}^{\mathrm{T}}t} \\
&\quad + \sigma^2\mathrm{E}\left(\int\limits_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\,dW\left(u\right)\left(\int\limits_0^t e^{\mathbf{A}(t-u)}\mathbf{e}\,dW\left(u\right)\right)^{\mathrm{T}}\right) \qquad \text{cf. (2.2.2)} \\
&= e^{\mathbf{A}s}\operatorname{var}\left(\mathbf{X}\left(0\right)\right)e^{\mathbf{A}^{\mathrm{T}}t} + \sigma^2\mathrm{E}\left(\int\limits_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\,dW\left(u\right)\left(\int\limits_0^t e^{\mathbf{A}(t-u)}\mathbf{e}\,dW\left(u\right)\right)^{\mathrm{T}}\right) \\
&\qquad \text{cf. (2.2.15)} \\
&= e^{\mathbf{A}s}\operatorname{var}\left(\mathbf{X}\left(0\right)\right)e^{\mathbf{A}^{\mathrm{T}}t} + \sigma^2\mathrm{E}\left(\int\limits_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\,dW\left(u\right)\left(\int\limits_0^s e^{\mathbf{A}(t-u)}\mathbf{e}\,dW\left(u\right)\right)^{\mathrm{T}}\right) \\
&\quad + \sigma^2\mathrm{E}\left(\int\limits_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\,dW\left(u\right)\left(\int\limits_s^t e^{\mathbf{A}(t-u)}\mathbf{e}\,dW\left(u\right)\right)^{\mathrm{T}}\right) \\
&= e^{\mathbf{A}s}\operatorname{var}\left(\mathbf{X}\left(0\right)\right)e^{\mathbf{A}^{\mathrm{T}}t} + \sigma^2\mathrm{E}\left(\int\limits_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\,dW\left(u\right)\left(\int\limits_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\,dW\left(u\right)\right)^{\mathrm{T}}\right)e^{\mathbf{A}^{\mathrm{T}}(t-s)} \\
&\quad + \sigma^2\mathrm{E}\left(\int\limits_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\,dW\left(u\right)\right)\mathrm{E}\left(\left(\int\limits_s^t e^{\mathbf{A}(t-u)}\mathbf{e}\,dW\left(u\right)\right)^{\mathrm{T}}\right) \qquad \text{cf. (2.1.2)}
\end{aligned}$$

$$= e^{\mathbf{A}s}\mathrm{var}\left(\mathbf{X}\left(0\right)\right)e^{\mathbf{A}^{\mathrm{T}}t} + \sigma^2\mathrm{E}\left(\int_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\mathbf{e}^{\mathrm{T}}e^{\mathbf{A}^{\mathrm{T}}(t-u)}\,du\right) \qquad \text{cf. (2.1.14), (2.2.15)}$$

$$= e^{\mathbf{A}s}\mathrm{var}\left(\mathbf{X}\left(0\right)\right)e^{\mathbf{A}^{\mathrm{T}}t} + \sigma^2\int_0^s e^{\mathbf{A}(s-u)}\mathbf{e}\mathbf{e}^{\mathrm{T}}e^{\mathbf{A}^{\mathrm{T}}(t-u)}\,du.$$

Thus (2.2.13) holds. ■

With the first and second order moments of $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ in hand, we now turn to the stationarity properties of $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$. In practice, strict stationarity is seldom achieved, or even desired; we restrict ourselves to the investigation of weak stationarity.

Definition 2.2.7 (Weak stationarity) *A multivariate stochastic process $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ is called weakly stationary, or second order stationary, if its first and second order moments are independent of time, i.e. its mean is constant and, for $t \in [0,\infty)$ and $s \in [0,\infty)$, $\mathrm{cov}\left(\mathbf{X}\left(s\right), \mathbf{X}\left(t\right)\right)$ depends on time only through $t - s$.*

Asymptotic weak stationarityis weaker than the form of stationarity described in Definition 2.2.7. As the name implies, asymptotically weakly stationary processes can be said to become (approximately) weakly stationary after the passage of enough time.

Definition 2.2.8 (Asymptotic weak stationarity) *A multivariate stochastic process $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ is called asymptotically weakly stationary, or asymptotically second order stationary, if the limits with respect to time of its first and second order moments are independent of time, i.e. both $\lim_{t\to\infty}\mathrm{E}\left(\mathbf{X}\left(t\right)\right)$ and $\lim_{t\to\infty}\mathrm{cov}\left(\mathbf{X}\left(t+h\right), \mathbf{X}\left(t\right)\right)$ exist (the latter for all $h \in \mathbb{R}$).*

In order to study the stationarity properties of $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ in (2.2.3), we need the following auxiliary result.

Lemma 2.2.9 *The eigenvalues of a $p \times p$ matrix $\mathbf{A}$ have negative real parts if and only if*

$$\lim_{t\to\infty} e^{\mathbf{A}t} = \mathbf{0}. \tag{2.2.16}$$

Proof. Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_p$ represent the (not necessarily distinct) eigenvalues of $\mathbf{A}$, and $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_p$ represent the corresponding eigenvectors. If

$$\mathbf{R} := \left[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \cdots \quad \mathbf{r}_p\right], \tag{2.2.17}$$

then $\mathbf{R}^{-1}$ exists, since the eigenvectors of $\mathbf{A}$ are linearly independent. If

$$\mathbf{E} := \mathrm{diag}\left\{\lambda_1, \lambda_2, \ldots, \lambda_p\right\}, \tag{2.2.18}$$

then

$$\begin{aligned} \mathbf{AR} &= \mathbf{RE} \\ \mathbf{A} &= \mathbf{RER}^{-1}. \end{aligned} \tag{2.2.19}$$

Therefore

$$\begin{aligned} e^{\mathbf{A}t} &= e^{\mathbf{RER}^{-1}t} && \text{cf. (2.2.19)} \\ &= \sum_{k\in\mathbb{N}_0} \frac{t^k}{k!} \left(\mathbf{RER}^{-1}\right)^k && \text{cf. (2.2.6)} \\ &= \mathbf{R} \left(\sum_{k\in\mathbb{N}_0} \frac{t^k}{k!} \mathbf{E}^k \right) \mathbf{R}^{-1} \\ &= \mathbf{R} \left(\operatorname{diag} \left\{ \sum_{k\in\mathbb{N}_0} \frac{(\lambda_1 t)^k}{k!}, \sum_{k\in\mathbb{N}_0} \frac{(\lambda_2 t)^k}{k!}, \ldots, \sum_{k\in\mathbb{N}_0} \frac{(\lambda_t t)^k}{k!} \right\} \right) \mathbf{R}^{-1} && \text{cf. (2.2.18)} \\ &= \mathbf{R} \left(\operatorname{diag} \left\{ e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots, e^{\lambda_p t} \right\} \right) \mathbf{R}^{-1}. \end{aligned} \tag{2.2.20}$$

For $k = 1, 2, \ldots, p$, we have (cf. Saff and Snider, 1993, p. 75)

$$\begin{aligned} \lim_{t\to\infty} \left| e^{\lambda_k t} \right| &= \lim_{t\to\infty} e^{\operatorname{Re}(\lambda_k)t} \left| \cos\left(\operatorname{Im}(\lambda_k)\, t\right) + i \sin\left(\operatorname{Im}(\lambda_k)\, t\right) \right| \\ &\le \lim_{t\to\infty} e^{\operatorname{Re}(\lambda_k)t}. \end{aligned} \tag{2.2.21}$$

Thus

$$\lim_{t\to\infty} e^{\lambda_k t} = 0 \text{ if and only if } \operatorname{Re}(\lambda_k) < 0. \tag{2.2.22}$$

In the light of (2.2.20), the relation (2.2.22) implies that $\lim_{t\to\infty} e^{\mathbf{A}t} = 0$ if and only if $\lambda_1, \lambda_2, \ldots, \lambda_p$ all have negative real parts. ■

With this result in hand, we can now obtain a necessary and sufficient condition for $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ in (2.2.3) to be asymptotically weakly stationary.

Theorem 2.2.10 *The process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *in* (2.2.3) *is asymptotically weakly stationary if and only if all the eigenvalues of the matrix* $\mathbf{A}$ *in* (2.2.3) *have negative real parts.*

Proof. In this proof, we essentially follow the same approach as Arató (1982, pp. 118-119). Suppose, on the one hand, that all the eigenvalues of $\mathbf{A}$ have negative real parts. Equation (2.2.16) then follows from Lemma 2.2.9. This implies that the integral

$$\mathbf{\Sigma} := \sigma^2 \int_0^\infty e^{\mathbf{A}u} \mathbf{e}\mathbf{e}^{\mathrm{T}} e^{\mathbf{A}^{\mathrm{T}}u}\, du \tag{2.2.23}$$

exists. For $h \in [0, \infty)$, we have

$$\begin{aligned}
\lim_{t\to\infty} \operatorname{cov}\left(\mathbf{X}(t+h), \mathbf{X}(t)\right) &= \lim_{t\to\infty} \left(e^{\mathbf{A}(t+h)} \operatorname{var}\left(\mathbf{X}(0)\right) e^{\mathbf{A}^{\mathrm{T}} t} + \sigma^2 \int_0^t e^{\mathbf{A}(t+h-u)} \mathbf{e}\mathbf{e}^{\mathrm{T}} e^{\mathbf{A}^{\mathrm{T}}(t-u)}\, du \right) && \text{cf. (2.2.13)} \\
&= e^{\mathbf{A}h} \left(\lim_{t\to\infty} e^{\mathbf{A}t} \right) \operatorname{var}\left(\mathbf{X}(0)\right) \left(\lim_{t\to\infty} e^{\mathbf{A}^{\mathrm{T}} t} \right) \\
&\quad + \sigma^2 e^{\mathbf{A}h} \lim_{t\to\infty} \int_0^t e^{\mathbf{A}(t-u)} \mathbf{e}\mathbf{e}^{\mathrm{T}} e^{\mathbf{A}^{\mathrm{T}}(t-u)}\, du \\
&= e^{\mathbf{A}h} \sigma^2 \int_0^\infty e^{\mathbf{A}u} \mathbf{e}\mathbf{e}^{\mathrm{T}} e^{\mathbf{A}^{\mathrm{T}} u}\, du && \text{cf. (2.2.16)} \\
&= e^{\mathbf{A}h} \mathbf{\Sigma}. && \text{cf. (2.2.23)}
\end{aligned} \tag{2.2.24}$$

In addition, for $h \in (-\infty, 0)$,

$$\begin{aligned}
\lim_{t\to\infty} \operatorname{cov}\left(\mathbf{X}(t+h), \mathbf{X}(t)\right) &= \lim_{t\to\infty} \operatorname{cov}\left(\mathbf{X}(t-|h|), \mathbf{X}(t)\right) \\
&= \left(\lim_{t\to\infty} \operatorname{cov}\left(\mathbf{X}(t), \mathbf{X}(t-|h|)\right) \right)^{\mathrm{T}} \\
&= \left(\lim_{t\to\infty} \operatorname{cov}\left(\mathbf{X}((t-|h|)+|h|), \mathbf{X}(t-|h|)\right) \right)^{\mathrm{T}} \\
&= \left(e^{\mathbf{A}|h|} \mathbf{\Sigma} \right)^{\mathrm{T}} && \text{cf. (2.2.23)} \\
&= \mathbf{\Sigma} e^{-\mathbf{A}^{\mathrm{T}} h}.
\end{aligned} \tag{2.2.25}$$

Moreover,

$$\begin{aligned}
\lim_{t\to\infty} \mathrm{E}\left(\mathbf{X}(t)\right) &= \lim_{t\to\infty} \left(e^{\mathbf{A}t} \mathrm{E}\left(\mathbf{X}(0)\right) \right) && \text{cf. (2.2.12)} \\
&= \left(\lim_{t\to\infty} e^{\mathbf{A}t} \right) \mathrm{E}\left(\mathbf{X}(0)\right) \\
&= \mathbf{0}. && \text{cf. (2.2.16)}
\end{aligned} \tag{2.2.26}$$

Hence $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ is asymptotically weakly stationary.

On the other hand, if one or more of the eigenvalues of $\mathbf{A}$ have non-negative real parts, then $\mathbf{\Sigma}$ in (2.2.23) does not exist (cf. Lemma 2.2.9). Then, in similar fashion to (2.2.24), it follows that $\lim_{t\to\infty} \operatorname{var}\left(\mathbf{X}(t)\right)$ is divergent. Therefore $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ cannot be asymptotically weakly stationary. ■

In the (asymptotically) weakly stationary case, the covariance function simplifies significantly. It is customary to rather express it in terms of the lag, i.e. as

$$\mathbf{\Gamma}(t, h) := \operatorname{cov}\left(\mathbf{X}(t+h), \mathbf{X}(t)\right) \text{ for } t \in [0, \infty) \text{ and } h \in \mathbb{R} \tag{2.2.27}$$

It follows from (2.2.24) and (2.2.25) (cf. Khabie-Zeitoune, 1982, p. 16) that

$$\lim_{t\to\infty} \mathbf{\Gamma}(t,h) = \begin{cases} e^{\mathbf{A}h}\mathbf{\Sigma} & \text{if } h \in [0,\infty) \\ \mathbf{\Sigma}e^{-\mathbf{A}^{\mathrm{T}}h} & \text{if } h \in (-\infty,0) \end{cases}. \tag{2.2.28}$$

The next result follows immediately from Theorem 2.2.10 (cf. Brockwell, 2000, pp. 4-5).

Corollary 2.2.11 *Suppose that, in* (2.2.3)*, the matrix* $\mathbf{A}$ *is chosen such that all its eigenvalues have negative real parts. If* $\mathbf{\Sigma}$ *is defined by* (2.2.23) *and* $\mathbf{X}(0) \sim N(\mathbf{0},\mathbf{\Sigma})$*, then* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *is a stationary zero-mean Gaussian process with covariance function*

$$\mathbf{\Gamma}(h) = \begin{cases} e^{\mathbf{A}h}\mathbf{\Sigma} & \textit{if } h \in [0,\infty) \\ \mathbf{\Sigma}e^{-\mathbf{A}^{\mathrm{T}}h} & \textit{if } h \in (-\infty,0) \end{cases}. \tag{2.2.29}$$

Chapter 3

Moment and stationarity properties of continuous-time autoregressive moving average processes

The main results of this study are presented in this chapter. In Section 3.1, we introduce the most general case, namely, the continuous-time autoregressive moving average (CARMA) time series model. We discuss its first and second-order moments and stationarity properties. Having established that the continuous-time autoregressive (CAR) process is a special CARMA process, Section 3.2 is devoted to establishing the properties of CAR processes. Likewise, the continuous-time moving average (CMA) process forms the focus of Section 3.3.

3.1 Continuous-time autoregressive moving average processes

A discrete-time autoregressive moving average process is defined as follows (cf. Jones (1981, p. 165), Cryer (1986, p. 70) and Brockwell (1995, p. 457)).

Definition 3.1.1 (DARMA (p, q) process) *Suppose that $\{\varepsilon_n\}_{n\in\mathbb{N}_0}$ is a discrete-time white noise process, and $\phi_1, \phi_2, \ldots, \phi_p$ and $\theta_1, \theta_2, \ldots, \theta_q$ (where $q < p$) are real numbers. The discrete-time stochastic process $\{Y_n\}_{n\in\mathbb{N}_0}$ satisfying the difference equation*

$$Y_n - \phi_1 Y_{n-1} - \phi_2 Y_{n-2} - \cdots - \phi_p Y_{n-p} = \varepsilon_n - \theta_1 \varepsilon_{n-1} - \theta_2 \varepsilon_{n-2} - \cdots - \theta_q \varepsilon_{n-q} \quad \text{for } n = p, p+1, p+2, \ldots \tag{3.1.1}$$

is called a discrete-time autoregressive moving average process of order (p, q) (or, in short, a DARMA (p, q) process).

Continuous-time autoregressive moving average (CARMA) processes are defined along similar lines, with the most prominent change being the replacement of the difference equation by a differential equation. The definition that is most often encountered in statistical literature (cf. Robinson (1980, p. 263), Jones (1981, p. 171), Priestley (1981, p. 176), Brockwell (1995, p. 452) and Brockwell (2000, p. 3)) reads as follows. If $\{\varepsilon(t)\}_{t\in[0,\infty)}$ is a Gaussian continuous-time white noise process, and $\alpha_1, \alpha_2, \ldots, \alpha_p$ and $\beta_1, \beta_2, \ldots, \beta_q$ (where $q < p$) are real numbers, then the CARMA (p, q) process $\{Y(t)\}_{t\in[0,\infty)}$ is defined to be a solution of the stochastic differential equation

$$\alpha_p Y(t) + \alpha_{p-1} Y'(t) + \cdots + \alpha_1 Y^{(p-1)}(t) + Y^{(p)}(t) = \varepsilon(t) + \beta_1 \varepsilon'(t) + \cdots + \beta_q \varepsilon^{(q)}(t) \quad \text{for } t \in (0, \infty)\,, \tag{3.1.2}$$

under suitable initial conditions.

In view of the non-realizability of $\{\varepsilon(t)\}_{t\in[0,\infty)}$, the above definition should be interpreted in a formal way. We now present the rigorous definition of a continuous-time autoregressive moving average process (cf. Brockwell (1995, p. 452), Brockwell (2000, pp. 3-4)).

Definition 3.1.2 (CARMA (p, q) process) *For $p \in \mathbb{N}$ and $q \in \mathbb{N}_0$ satisfying $p > q$, let the real numbers $\alpha_1, \alpha_2, \ldots, \alpha_p$ and $\beta_1, \beta_2, \ldots, \beta_q$ be given. Suppose that $\mathbf{A}$ is a $p \times p$ matrix defined as*

$$\mathbf{A} := \begin{cases} -\alpha_1 & \text{if } p = 1 \\ \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -\alpha_p & -\alpha_{p-1} & -\alpha_{p-2} & \cdots & -\alpha_2 & -\alpha_1 \end{bmatrix} & \text{otherwise} \end{cases}, \tag{3.1.3}$$

and let $\mathbf{b}$ and $\mathbf{e}$ be p-dimensional vectors defined, respectively, by

$$\mathbf{b} := \begin{bmatrix} 1 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_q \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and } \mathbf{e} := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{3.1.4}$$

Suppose that there exists a continuous-time real-valued stochastic process $\{X(t)\}_{t\in[0,\infty)}$ *such that the p-variate stochastic process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *defined by*

$$\mathbf{X}(t) := \begin{bmatrix} X(t) \\ X'(t) \\ \vdots \\ X^{(p-1)}(t) \end{bmatrix} \text{ for } t \in [0,\infty) \tag{3.1.5}$$

satisfies the stochastic differential equation

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t)\ dt + \sigma\mathbf{e}\,dW(t) \ \text{ for } t \in [0,\infty), \tag{3.1.6}$$

where $\sigma \in [0,\infty)$ *is given and* $\{W(t)\}_{t\in[0,\infty)}$ *is standard Brownian motion, under the initial condition*

$$\mathbf{X}(0) \text{ is normally distributed and uncorrelated with } W(t) \text{ for all } t \in (0,\infty). \tag{3.1.7}$$

The continuous-time stochastic process $\{Y(t)\}_{t\in[0,\infty)}$ *defined by*

$$Y(t) = \mathbf{b}^{\mathrm{T}}\mathbf{X}(t) \ \text{ for } t \in [0,\infty), \tag{3.1.8}$$

is called a continuous-time autoregressive moving average process of order (p,q) *(in short, a CARMA* (p,q) *process).*

Remark 3.1.3 *In the terminology of the state-space representation approach to linear filtering (cf. Kalman and Bucy, 1961), equation* (3.1.6) *is called the state equation and* (3.1.8) *is called the observation equation. In fact, given the form of* $\mathbf{A}$, $\mathbf{b}$ *and* $\mathbf{e}$ *in* (3.1.3)-(3.1.4), *equations* (3.1.6) *and* (3.1.8) *are called the state and observation equations of the second canonical form of a linear filter (cf. Wiberg, 1971).*

Remark 3.1.4 *The restriction on p and q in Definition 3.1.2 is necessary and sufficient for the CARMA* (p,q) *process* $\{Y(t)\}_{t\in[0,\infty)}$ *to be completely non-deterministic (cf. Doob (1953, p. 579), Priestley (1981, Section 10.1)).*

Equation (3.1.2) should be regarded as a formal representation of (3.1.6) (together with (3.1.8)), which is, in turn (as in Section 2.1), a formal representation of the Itô integral equation

$$\mathbf{X}(t) = \mathbf{X}(0) + \int_0^t \mathbf{A}\mathbf{X}(u)\ du + \sigma\int_0^t \mathbf{e}\,dW(u) \ \text{ for } t \in [0,\infty). \tag{3.1.9}$$

Comparing (3.1.7) with (2.2.2) and (3.1.6) with (2.2.3), it follows that the results of Section 2.2 directly apply to the stochastic process $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$, and, by (3.1.8), to $\{Y(t)\}_{t\in[0,\infty)}$.

According to Theorem 2.2.3, the process $\{Y(t)\}_{t\in[0,\infty)}$ exists and is given by

$$Y(t) = \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}t}\mathbf{X}(0) + \sigma \mathbf{b}^{\mathrm{T}} \int_0^t e^{\mathbf{A}(t-u)}\mathbf{e}\, dW(u) \ \text{ for } t \in [0,\infty). \tag{3.1.10}$$

Theorem 2.2.6 provides us with the exact form of the first and second order moments of $\{Y(t)\}_{t\in[0,\infty)}$. For $t \in [0,\infty)$, the expected value function is given by

$$\begin{aligned} \mathrm{E}(Y(t)) &= \mathbf{b}^{\mathrm{T}}\mathrm{E}(\mathbf{X}(t)) && \text{cf. (3.1.8)} \\ &= \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}t}\mathrm{E}(\mathbf{X}(0)). && \text{cf. (2.2.12)} \end{aligned} \tag{3.1.11}$$

For $s \in [0,\infty)$ and $t \in [0,\infty)$, the covariance function is given by

$$\begin{aligned} \mathrm{cov}(Y(s), Y(t)) &= \mathbf{b}^{\mathrm{T}}\mathrm{cov}(\mathbf{X}(s), \mathbf{X}(t))\,\mathbf{b} && \text{cf. (3.1.8)} \\ &= \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}s}\mathrm{var}(\mathbf{X}(0))\, e^{\mathbf{A}^{\mathrm{T}}t}\mathbf{b} + \sigma^2 \int_0^{\min\{s,t\}} \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}(s-u)}\mathbf{e}\mathbf{e}^{\mathrm{T}} e^{\mathbf{A}^{\mathrm{T}}(t-u)}\mathbf{b}\, du. \\ & && \text{cf. (2.2.12)} \end{aligned} \tag{3.1.12}$$

In order to determine the conditions under which $\{Y(t)\}_{t\in[0,\infty)}$ is asymptotically weakly stationary, we need to find the eigenvalues of $\mathbf{A}$. Easy calculation shows that the p eigenvalues of $\mathbf{A}$ are exactly the roots of the so-called autoregressive polynomial $\alpha(z)$, defined as

$$\alpha(z) := z^p + \alpha_1 z^{p-1} + \cdots + \alpha_p \text{ for } z \in \mathbb{C}. \tag{3.1.13}$$

With Theorem 2.2.10 in hand, the asymptotic weak stationarity of $\{Y(t)\}_{t\in[0,\infty)}$ translates to all the roots of $\alpha(z)$ having negative real parts. If this is assumed, then the asymptotic expected value of $\{Y(t)\}_{t\in[0,\infty)}$ is given by

$$\begin{aligned} \lim_{t\to\infty} \mathrm{E}(Y(t)) &= \lim_{t\to\infty} \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}t}\mathrm{E}(\mathbf{X}(0)) && \text{cf. (3.1.11)} \\ &= 0. && \text{cf. (2.2.26)} \end{aligned} \tag{3.1.14}$$

Moreover, if we define

$$\gamma(t,h) := \mathrm{cov}(Y(t+h), Y(t)) \ \text{ for } t \in [0,\infty) \ \text{ and } h \in \mathbb{R} \tag{3.1.15}$$

and

$$\mathbf{\Sigma} := \sigma^2 \int_0^\infty e^{\mathbf{A}u}\mathbf{e}\mathbf{e}^{\mathrm{T}} e^{\mathbf{A}^{\mathrm{T}}u}\, du, \tag{3.1.16}$$

then

$$\begin{aligned}
\lim_{t\to\infty}\gamma(t,h) &= \mathbf{b}^{\mathrm{T}}\left(\lim_{t\to\infty}\operatorname{cov}\left(\mathbf{X}\left(t+h\right),\mathbf{X}\left(t\right)\right)\right)\mathbf{b} && \text{cf. (3.1.8)}\\
&= \mathbf{b}^{\mathrm{T}}\left(\lim_{t\to\infty}\mathbf{\Gamma}(t,h)\right)\mathbf{b} && \text{cf. (2.2.27)}\\
&= \begin{cases}\mathbf{b}^{\mathrm{T}}e^{\mathbf{A}h}\mathbf{\Sigma}\mathbf{b} & \text{if } h\in[0,\infty)\\ \mathbf{b}^{\mathrm{T}}\mathbf{\Sigma}e^{-\mathbf{A}^{\mathrm{T}}h}\mathbf{b} & \text{if } h\in(-\infty,0)\end{cases} && \text{cf. (2.2.28)}\\
&= \mathbf{b}^{\mathrm{T}}e^{\mathbf{A}|h|}\mathbf{\Sigma}\mathbf{b}, && (3.1.17)
\end{aligned}$$

because of the fact that any real number is equal to its own transpose.

Now, suppose that all the roots of the autoregressive polynomial $\alpha(z)$ in (3.1.13) have negative real parts, and that $\mathbf{X}\left(0\right)\sim \mathrm{N}\left(\mathbf{0},\mathbf{\Sigma}\right)$. Equation (3.1.8) implies that $Y(0)\sim\mathrm{N}\left(0,\mathbf{b}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{b}\right)$. From Corollary 2.2.11, it follows that $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$, and therefore $\{Y(t)\}_{t\in[0,\infty)}$, is a stationary zero-mean Gaussian process. The covariance function of $\{Y(t)\}_{t\in[0,\infty)}$ is then given (for $h\in\mathbb{R}$) by

$$\begin{aligned}
\gamma(h) &= \mathbf{b}^{\mathrm{T}}\mathbf{\Gamma}(h)\mathbf{b} && \text{cf. (3.1.8)}\\
&= \begin{cases}\mathbf{b}^{\mathrm{T}}e^{\mathbf{A}h}\mathbf{\Sigma}\mathbf{b} & \text{if } h\in[0,\infty)\\ \mathbf{b}^{\mathrm{T}}\mathbf{\Sigma}e^{-\mathbf{A}^{\mathrm{T}}h}\mathbf{b} & \text{if } h\in(-\infty,0)\end{cases} && \text{cf. (2.2.29)}\\
&= \mathbf{b}^{\mathrm{T}}e^{\mathbf{A}|h|}\mathbf{\Sigma}\mathbf{b} && (3.1.18)
\end{aligned}$$

(cf. Brockwell, 2000, p. 5), once again for the reason that any real number is its own transpose.

Remark 3.1.5 *In the special case where the autoregressive polynomial $\alpha(z)$ in* (3.1.13) *has p distinct zeroes, then, according to Brockwell (2000, p. 5), the covariance function γ has a very simple form. Indeed, if the so-called moving average polynomial is defined as*

$$\beta(z) := 1+\beta_1 z+\beta_2 z^2+\cdots+\beta_q z^q \text{ for } z\in\mathbb{C}, \tag{3.1.19}$$

then

$$\gamma(h) = \sum_{\{\lambda\in\mathbb{C}\,|\,\alpha(\lambda)=0\}}\frac{e^{\lambda|h|}\beta(\lambda)\beta(-\lambda)}{\alpha'(\lambda)\alpha(-\lambda)} \text{ for } h\in\mathbb{R}. \tag{3.1.20}$$

3.2 Continuous-time autoregressive processes

Discrete-time autoregressive processes are defined as follows (cf. Cryer (1986, p. 59–60), Brockwell (1995, p. 455)).

Definition 3.2.1 (DAR(p) process) *Suppose that $\{\varepsilon_n\}_{n\in\mathbb{N}_0}$ is a discrete-time white noise process, and $\phi_1, \phi_2, \dots, \phi_p$ are real numbers. The discrete-time stochastic process $\{Y_n\}_{n\in\mathbb{N}_0}$ satisfying the difference equation*

$$Y_n - \phi_1 Y_{n-1} - \phi_2 Y_{n-2} - \cdots - \phi_p Y_{n-p} = \varepsilon_n \text{ for } n = p, p+1, p+2, \dots \tag{3.2.1}$$

is called a discrete-time autoregressive process of order p (or a DAR(p) process).

As with CARMA processes, continuous-time autoregressive (CAR) processes are defined by replacing the difference equation by a differential equation. The usual definition then reads as follows (cf. Jones (1981, p. 170), Priestley (1981, p. 174), Jones (1985, p. 653), Hyndman (1993, p. 281)): If $\{\varepsilon(t)\}_{t\in[0,\infty)}$ is a Gaussian continuous-time white noise process, and $\alpha_1, \alpha_2, \dots, \alpha_p$ are real numbers, then the CAR(p) process $\{Y(t)\}_{t\in[0,\infty)}$ is defined to be a solution of the stochastic differential equation

$$\alpha_p Y(t) + \alpha_{p-1} Y'(t) + \cdots + \alpha_1 Y^{(p-1)}(t) + Y^{(p)}(t) = \varepsilon(t) \text{ for } t \in (0, \infty) \tag{3.2.2}$$

coupled with suitable initial conditions.

Once again, the non-realizability of $\{\varepsilon(t)\}_{t\in[0,\infty)}$ poses problems in the validity of the above definition. We now present a rigorous definition of a continuous-time autoregressive process (cf. Jones (1981, p. 654), Jones (1985, pp. 170–171) and Hyndman (1993, p. 282)).

Definition 3.2.2 (CAR(p) process) *For $p \in \mathbb{N}$, let the real numbers $\alpha_1, \alpha_2, \dots, \alpha_p$ be given. Suppose that $\mathbf{e}$ is a p-dimensional vector and $\mathbf{A}$ is a $p \times p$ matrix defined, respectively, by*

$$\mathbf{e} := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \text{ and } \mathbf{A} := \begin{cases} -\alpha_1 & \textit{if } p = 1 \\ \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -\alpha_p & -\alpha_{p-1} & -\alpha_{p-2} & \cdots & -\alpha_2 & -\alpha_1 \end{bmatrix} & \textit{otherwise} \end{cases}. \tag{3.2.3}$$

A continuous-time real-valued stochastic process $\{X(t)\}_{t\in[0,\infty)}$ *is called a continuous-time autoregressive process of order* p *(or a CAR*(p) *process) if the* p*-variate stochastic process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *defined by*

$$\mathbf{X}(t) := \begin{bmatrix} X(t) \\ X'(t) \\ \vdots \\ X^{(p-1)}(t) \end{bmatrix} \quad \text{for } t \in [0,\infty) \tag{3.2.4}$$

satisfies the stochastic differential equation

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t)\ dt + \sigma\mathbf{e}\, dW(t) \ \text{ for } t \in [0,\infty), \tag{3.2.5}$$

where $\sigma \in [0,\infty)$ *is given and* $\{W(t)\}_{t\in[0,\infty)}$ *is standard Brownian motion, under the initial condition*

$$\mathbf{X}(0) \ \text{ is normally distributed and uncorrelated with } W(t) \ \text{ for all } t \in (0,\infty). \tag{3.2.6}$$

Comparing Definition 3.2.2 with Definition 3.1.2, the specific relationship between the CARMA(p,q) process and the CAR(p) process, which is (in view of Theorem 2.2.5) a completely non-deterministic p-dimensional Markov process, is immediately clear. The fact that a CAR(p) process is exactly the same as a CARMA$(p,0)$ process, is due to Doob (1944, Theorems 3.9 and 4.9). Remark 3.1.4 should also be interpreted in the light of this well-known result, which is valid in both the discrete-time and continuous-time cases. It reads as follows.

Theorem 3.2.3 *A one-dimensional process is a DARMA*(p,q)*—or CARMA*(p,q)*—process with* $p > q$ *if and only if it is a component process of a completely non-deterministic discrete-time—or continuous-time—*p*-dimensional Markov process.*

An equivalent definition of a CAR(p) process $\{Y(t)\}_{t\in[0,\infty)}$ would therefore be that it satisfies the state equation (3.2.5) and observation equation

$$Y(t) = \mathbf{b}^{\mathrm{T}}\mathbf{X}(t) \ \text{ for } t \in [0,\infty), \tag{3.2.7}$$

where $\mathbf{b}$ is a p-dimensional vector defined by

$$\mathbf{b} := \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{3.2.8}$$

For the purpose of consistency of notation, we will denote a CAR (p) process by $\{Y(t)\}_{t\in[0,\infty)}$, although it is exactly the same stochastic process as $\{X(t)\}_{t\in[0,\infty)}$ in Definition 3.2.2.

As usual, (3.2.2) is regarded as a formal representation of (3.2.5) (coupled with (3.2.7)), which is, in turn, a formal representation of the now familiar Itô integral equation

$$\mathbf{X}(t) = \mathbf{X}(0) + \int_0^t \mathbf{A}\mathbf{X}(u)\, du + \sigma \int_0^t \mathbf{e}\, dW(u) \ \text{ for } t \in [0,\infty). \tag{3.2.9}$$

Using the properties of CARMA processes established in Section 3.1, the qualitative and quantitative properties of the CAR (p) process $\{Y(t)\}_{t\in[0,\infty)}$ are easily determined. According to Theorem 2.2.3 (cf. (3.1.10)), $\{Y(t)\}_{t\in[0,\infty)}$ exists and satisfies

$$Y(t) = \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}t}\mathbf{X}(0) + \sigma \mathbf{b}^{\mathrm{T}} \int_0^t e^{\mathbf{A}(t-u)}\mathbf{e}\, dW(u) \ \text{ for } t \in [0,\infty). \tag{3.2.10}$$

Once again, Theorem 2.2.6 provides us with the exact form of the first and second order moments of $\{Y(t)\}_{t\in[0,\infty)}$. According to (3.1.11), the expected value function of $\{Y(t)\}_{t\in[0,\infty)}$ is given by

$$\mathrm{E}(Y(t)) = \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}t}\mathrm{E}(\mathbf{X}(0)) \ \text{ for } t \in [0,\infty). \tag{3.2.11}$$

Using (3.1.12), the covariance function of $\{Y(t)\}_{t\in[0,\infty)}$ is of the form

$$\operatorname{cov}(Y(s), Y(t)) = \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}s}\operatorname{var}(\mathbf{X}(0))\, e^{\mathbf{A}^{\mathrm{T}}t}\mathbf{b} + \sigma^2 \int_0^{\min\{s,t\}} \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}(s-u)}\mathbf{e}\mathbf{e}^{\mathrm{T}} e^{\mathbf{A}^{\mathrm{T}}(t-u)}\mathbf{b}\, du$$
$$\text{for } s \in [0,\infty) \text{ and } t \in [0,\infty). \tag{3.2.12}$$

As is the case with CARMA processes, Theorem 2.2.10 allows the asymptotic weak stationarity of $\{Y(t)\}_{t\in[0,\infty)}$ to be translated to all the roots of the autoregressive polynomial

$$\alpha(z) := z^p + \alpha_1 z^{p-1} + \cdots + \alpha_p \text{ for } z \in \mathbb{C} \tag{3.2.13}$$

having negative real parts. If this is indeed the case, then the asymptotic expected value of $\{Y(t)\}_{t\in[0,\infty)}$ is given (according to (3.1.14)) by

$$\lim_{t\to\infty} \mathrm{E}(Y(t)) = 0. \tag{3.2.14}$$

Moreover, if we define γ by (3.1.15) and $\boldsymbol{\Sigma}$ by (3.1.16), then, as in (3.1.17), the asymptotic covariance of $\{Y(t)\}_{t\in[0,\infty)}$ is equal to

$$\lim_{t\to\infty} \gamma(t,h) = \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}|h|}\boldsymbol{\Sigma}\mathbf{b}. \tag{3.2.15}$$

Now, suppose that all the roots of $\alpha(z)$ in (3.2.13) have negative real parts. If we assume that $\mathbf{X}(0) \sim \mathrm{N}(\mathbf{0}, \mathbf{\Sigma})$, then (3.2.7) implies that $Y(0) \sim \mathrm{N}\left(0, \mathbf{b}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{b}\right)$. From Corollary 2.2.11, it follows that $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$, and therefore $\{Y(t)\}_{t\in[0,\infty)}$, is a stationary zero-mean Gaussian process. The covariance function of $\{Y(t)\}_{t\in[0,\infty)}$ is then given by (3.1.18) as

$$\gamma(h) = \mathbf{b}^{\mathrm{T}} e^{\mathbf{A}|h|}\mathbf{\Sigma}\mathbf{b}. \tag{3.2.16}$$

Remark 3.2.4 *In the special case where the autoregressive polynomial $\alpha(z)$ in* (3.2.13) *has p distinct zeroes, then, according to Brockwell (2000, p. 5), the covariance function γ is of the very simple form*

$$\gamma(h) = \sum_{\{\lambda\in\mathbb{C}\,|\,\alpha(\lambda)=0\}} \frac{e^{\lambda|h|}}{\alpha'(\lambda)\alpha(-\lambda)} \textit{ for } h \in \mathbb{R}. \tag{3.2.17}$$

3.3 Continuous-time moving average processes

A discrete-time moving average process is defined as follows (cf. Cryer (1986, p. 54)).

Definition 3.3.1 (DMA (q) process) *Suppose that $\{\varepsilon_n\}_{n\in\mathbb{N}_0}$ is a discrete-time white noise process, and $\theta_1, \theta_2, \ldots, \theta_q$ are real numbers. The discrete-time stochastic process $\{Y_n\}_{n\in\mathbb{N}_0}$ satisfying the difference equation*

$$Y_n = \varepsilon_n - \theta_1\varepsilon_{n-1} - \theta_2\varepsilon_{n-2} - \cdots - \theta_q\varepsilon_{n-q} \text{ for } n = q, q+1, q+2, \ldots \tag{3.3.1}$$

is called a discrete-time moving average process of order q (or, in short, a DMA (q) process).

Continuous-time moving average (CMA) processes are defined similarly, by replacement of the difference equation by a differential equation. The most frequently used definition (cf. Priestley, 1981, p. 176) reads as follows. If $\{\varepsilon(t)\}_{t\in[0,\infty)}$ is a Gaussian continuous-time white noise process, and $\beta_1, \beta_2, \ldots, \beta_q$ are real numbers, then the CMA (q) process $\{Y(t)\}_{t\in[0,\infty)}$ is defined to be a solution of the differential equation

$$Y^{(q+1)}(t) = \varepsilon(t) + \beta_1\varepsilon'(t) + \beta_2\varepsilon''(t) + \cdots + \beta_q\varepsilon^{(q)}(t) \text{ for } t \in (0, \infty), \tag{3.3.2}$$

under suitable initial conditions.

Remark 3.3.2 *Unlike the models encountered before, there is a striking difference between the left-hand sides of* (3.3.1) *and* (3.3.2)*. The object of* (3.3.2) *is not $Y(t)$, as would be expected, but its $(q+1)^{th}$ derivative, while its counterpart in* (3.3.1) *is Y_n, as usual. This is mainly due to Theorem 3.2.3 and the necessity to write the CMA (q) model in terms of a suitable CARMA model.*

In view of the non-realizability of $\{\varepsilon(t)\}_{t\in[0,\infty)}$, the above definition should be interpreted in a formal way. We now present the rigorous definition of a continuous-time moving average process.

Definition 3.3.3 (CMA (q) process) *For $q \in \mathbb{N}$, let the real numbers $\beta_1, \beta_2, \ldots, \beta_q$ be given. Suppose that $\mathbf{A}$ is a $(q+1)\times(q+1)$ matrix defined as*

$$\mathbf{A} := \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \tag{3.3.3}$$

and let $\mathbf{b}$ *and* $\mathbf{e}$ *be* $(q+1)$*-dimensional vectors defined, respectively, by*

$$\mathbf{b} := \begin{bmatrix} 1 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_q \end{bmatrix} \text{ and } \mathbf{e} := \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{3.3.4}$$

Suppose that there exists a continuous-time real-valued stochastic process $\{X(t)\}_{t\in[0,\infty)}$ *such that the* $(q+1)$*-variate stochastic process* $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$ *defined by*

$$\mathbf{X}(t) := \begin{bmatrix} X(t) \\ X'(t) \\ \vdots \\ X^{(q)}(t) \end{bmatrix} \text{ for } t \in [0,\infty) \tag{3.3.5}$$

satisfies the stochastic differential equation

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t)\ dt + \sigma\mathbf{e}\, dW(t) \ \text{ for } t \in [0,\infty) \tag{3.3.6}$$

(called the state equation), where $\sigma \in [0,\infty)$ *is given and* $\{W(t)\}_{t\in[0,\infty)}$ *is standard Brownian motion, under the initial condition*

$$\mathbf{X}(0) \ \text{ is normally distributed and uncorrelated with } W(t) \text{ for all } t \in (0,\infty). \tag{3.3.7}$$

The continuous-time stochastic process $\{Y(t)\}_{t\in[0,\infty)}$ *defined by the observation equation*

$$Y(t) = \mathbf{b}^{\mathrm{T}}\mathbf{X}(t) \ \text{ for } t \in [0,\infty) \tag{3.3.8}$$

is called a continuous-time moving average process of order q *(in short, a CMA*(q) *process).*

Once again, (3.3.2) should be regarded as a formal representation of (3.3.6) (together with (3.3.8)), which is, as usual, a formal representation of

$$\mathbf{X}(t) = \mathbf{X}(0) + \int_0^t \mathbf{A}\mathbf{X}(u)\ du + \sigma \int_0^t \mathbf{e}\, dW(u) \ \text{ for } t \in [0,\infty). \tag{3.3.9}$$

Comparing Definition 3.3.3 with Definition 3.1.2, it is clear that a CMA(q) process is simply a CARMA$(q+1,q)$ process with autoregressive parameters all equal to zero. In addition, the specific form of $\mathbf{A}$ in (3.3.3) permits, for $t \in \mathbb{R}$, the computational simplification

$$
\begin{aligned}
e^{\mathbf{A}t} &= \sum_{k\in\mathbb{N}_0} \frac{t^k}{k!}\mathbf{A}^k \qquad \text{cf. (2.2.6)} \\
&= \sum_{k=0}^{q+1} \frac{t^k}{k!}\mathbf{A}^k \\
&= \begin{bmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{q-1}}{(q-1)!} & \frac{t^q}{q!} \\ 0 & 1 & t & \cdots & \frac{t^{q-2}}{(q-2)!} & \frac{t^{q-1}}{(q-1)!} \\ 0 & 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & t & \frac{t^2}{2} \\ 0 & 0 & \cdots & 0 & 1 & t \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}. \qquad (3.3.10)
\end{aligned}
$$

Using the properties of CARMA processes established in Section 3.1, the qualitative and quantitative properties of the CMA (q) process $\{Y(t)\}_{t\in[0,\infty)}$ are easily determined. Theorem 2.2.3 guarantees that the process $\{Y(t)\}_{t\in[0,\infty)}$ exists and satisfies (when combining (2.2.5) with (3.3.10))

$$
\begin{aligned}
Y(t) = &\begin{bmatrix} 1 & \beta_1 + t & \beta_2 + \beta_1 t + \frac{t^2}{2} & \cdots & \beta_q + \beta_{q-1}t + \cdots + \frac{t^q}{q!} \end{bmatrix} \mathbf{X}(0) \\
&+ \sigma \int_0^t \left(\beta_q + \beta_{q-1}(t-u) + \cdots + \frac{(t-u)^q}{q!} \right) dW(u) \ \text{ for } t \in [0,\infty). \qquad (3.3.11)
\end{aligned}
$$

The exact form of the first and second order moments of $\{Y(t)\}_{t\in[0,\infty)}$ are provided by Theorem 2.2.6. Combining (3.1.11) with (3.3.10), the expected value function is given by

$$
\begin{aligned}
\mathrm{E}(Y(t)) = &\begin{bmatrix} 1 & \beta_1 + t & \beta_2 + \beta_1 t + \frac{t^2}{2} & \cdots & \beta_q + \beta_{q-1}t + \cdots + \frac{t^q}{q!} \end{bmatrix} \mathrm{E}(\mathbf{X}(0)) \\
&\text{for } t \in [0,\infty). \qquad (3.3.12)
\end{aligned}
$$

From (3.1.11) it follows that, for $s \in [0,\infty)$ and $t \in [0,\infty)$, the covariance function is given by

$$
\begin{aligned}
&\operatorname{cov}(Y(s), Y(t)) \\
&= \begin{bmatrix} 1 & \beta_1 + s & \beta_2 + \beta_1 s + \frac{s^2}{2} & \cdots & \beta_q + \beta_{q-1}s + \cdots + \frac{s^q}{q!} \end{bmatrix} \operatorname{var}(\mathbf{X}(0)) \begin{bmatrix} 1 \\ \beta_1 + t \\ \beta_2 + \beta_1 t + \frac{t^2}{2} \\ \vdots \\ \beta_q + \beta_{q-1}t + \cdots + \frac{t^q}{q!} \end{bmatrix}
\end{aligned}
$$

$$+\,\sigma^2 \int_0^{\min\{s,t\}} \left(\beta_q + \beta_{q-1}(s-u) + \cdots + \frac{(s-u)^q}{q!}\right)\left(\beta_q + \beta_{q-1}(t-u) + \cdots + \frac{(t-u)^q}{q!}\right) du. \tag{3.3.13}$$

The specific form of $\mathbf{A}$ in (3.3.3) allows us to calculate the $q+1$ eigenvalues of $\mathbf{A}$ directly: they are all equal to zero. This is consistent with the fact that, in the case of the CMA (q) process, the autoregressive polynomial $\alpha(z)$ (defined in (3.2.13)) reduces to

$$\alpha(z) = z^{q+1} \text{ for } z \in \mathbb{C}. \tag{3.3.14}$$

Unfortunately, in stark contrast to the discrete-time case, Theorem 2.2.10 asserts that $\{Y(t)\}_{t\in[0,\infty)}$ is not asymptotically weakly stationary. Even without referring to Theorem 2.2.10, this disappointing fact is evident from the second term on the right-hand side of (3.3.13); due to its polynomial nature, a limiting covariance matrix such as $\mathbf{\Sigma}$ in (3.1.16) cannot exist.

Now, suppose that $\mathbf{X}(0) \sim \mathrm{N}(\mathbf{0}, \mathbf{\Sigma})$ for some given $(q+1)\times(q+1)$ matrix $\mathbf{\Sigma}$. In the usual fashion, (3.3.8) then implies that $Y(0) \sim \mathrm{N}\left(0, \mathbf{b}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{b}\right)$. In this case, due to the fact that all the eigenvalues of $\mathbf{A}$ are zero, Corollary 2.2.11 does not apply. As consolation, however, it can at least be said that $\{\mathbf{X}(t)\}_{t\in[0,\infty)}$, and therefore $\{Y(t)\}_{t\in[0,\infty)}$ (from (3.3.12)), is a zero-mean Gaussian process.

Bibliography

Adams, M. and Guillemin, V. (1996), *Measure theory and probability*, Birkhäuser, Boston.

Arató, M. (1982), *Linear stochastic systems with constant coefficients: A statistical approach*, number 45 *in* 'Lecture notes in control and information sciences', Springer, Berlin.

Bain, L. J. and Engelhardt, M. (1991), *Introduction to probability and mathematical statistics*, The Duxbury advanced series in statistics and decision sciences, 2nd edn, Duxbury Press, Belmont.

Bartlett, M. S. (1946), 'On the theoretical specification of sampling properties of autocorrelated time series', *Journal of the Royal Statistical Society, Series B* **8**, 27–41.

Bartlett, M. S. (1955), *An introduction to stochastic processes with special reference to methods and applications*, Cambridge University Press, London.

Belcher, J., Hampton, J. S. and Tunnicliffe Wilson, G. (1994), 'Parameterization of continuous time autoregressive models for irregularly sampled time series data', *Journal of the Royal Statistical Society, Series B* **56**, 141–155.

Bergstrom, A. R. (1984), *Continuous time stochastic models and issues of aggregation over time*, Vol. 2, Elsevier, Amsterdam, pp. 1146–1212.

Bergstrom, A. R. (1990), *Continuous time econometric modeling*, Oxford University Press, New York.

Brockwell, P. J. (1995), 'A note on the embedding of discrete-time ARMA processes', *Journal of Time Series Analysis* **16**(5), 451–460.

Brockwell, P. J. (2000), Heavy-tailed and non-linear continuous-time ARMA models for financial time series, *in* W.-S. Chan, W. Li and H. Tong, eds, 'Proceedings of the Hong Kong international workshop on statistics in finance: An interface, 4-8 July 1999', University of Hong Kong: Centre of Financial Time Series, Imperial College Press, London, pp. 3–23.

Brzeżniak, Z. and Zastawniak, T. (1999), *Basic stochastic processes*, Springer undergraduate mathematics series, Springer, London.

Chan, K. S. and Tong, H. (1987), 'A note on embedding a discrete parameter ARMA model', *Journal of Time Series Analysis* **8**(3), 277–281.

Chatfield, C. (1996), *The analysis of time series: An introduction*, 5th edn, Chapman & Hall, London.

Cryer, J. D. (1986), *Time series analysis*, Duxbury Press, Boston.

Doob, J. L. (1944), 'The elementary gaussian process', *Annals of Mathematical Statistics* **25**, 229–282.

Doob, J. L. (1953), *Stochastic processes*, Wiley, New York.

Durbin, J. (1961), 'Efficient fitting of linear models for continuous stationary time series from discrete data', *Bulletin of the International Statistical Institute* **38**, 273–281.

Dzhaparidze, K. O. (1971), 'On methods for obtaining asymptotically efficient spectral parameter estimates for a stationary Gaussian process with rational spectral density', *Theory of Probability and its Applications* **16**, 550–554.

Hannan, E. J. (1970), *Multiple time series*, Wiley, New York.

Harvey, A. (1989), *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, New York.

He, S. W. and Wang, J. G. (1989), 'On embedding a discrete-parameter ARMA model in a continuous-parameter ARMA model', *Journal of Time Series Analysis* **10**(4), 315–323.

Hyndman, R. J. (1993), 'Yule-Walker estimates for continuous-time autoregressive models', *Journal of Time Series Analysis* **14**(3), 281–296.

Jones, R. H. (1981), Fitting a continuous time autoregression to discrete data, *in* D. F. Findley, ed., 'Applied time series analysis II: Proceedings of the second time series symposium held in Tulsa, Oklahoma, 3-5 March 1978', pp. 651–682.

Jones, R. H. (1984), Fitting multivariate models to unequally spaced data, *in* E. Parzen, ed., 'Time series analysis of irregularly observed data. Proceedings of a symposium held at Texas A & M University, College Station, Texas, 10-13 February 1983', number 25 *in* 'Lecture notes in statistics', Springer-Verlag, Berlin-New York, pp. 158–188.

Jones, R. H. (1985), Time series analysis with unequally spaced data, *in* E. J. Hannan, P. R. Krishnaiah and M. M. Rao, eds, 'Handbook of statistics', Vol. 5: Time series in the time domain, Elsevier Science, Amsterdam, pp. 157–177.

Jones, R. H. (1993), *Longitudinal data with serial correction: A state-space approach*, Chapman & Hall, London.

Kallianpur, G. (1980), *Stochastic filtering theory*, number 13 *in* 'Applications of mathematics', Springer-Verlag, New York.

Kallianpur, G. and Karandikar, R. L. (1988), *White noise theory of prediction, filtering and smoothing*, Vol. 3 of *Statistics monographs*, Gordon and Breach Science Publishers, Amsterdam.

Kalman, R. E. and Bucy, B. S. (1961), 'New results in linear filtering and prediction theory', *Transactions of the American Society of Mechanical Engineering, Series D (Journal of Basic Engineering)* **83**, 95–108.

Karatzas, I. and Shreve, S. E. (1988), *Brownian motion and stochastic calculus*, Graduate Texts in Mathematics, Springer-Verlag, New York.

Khabie-Zeitoune, E. (1982), Prediction in continuous time, *in* O. D. Anderson, ed., 'Time Series Analysis: Theory and Practice 2', North-Holland, Amsterdam, pp. 7–24.

Liu, Y. (1988), 'An exploration whether the sample sequence of the stationary markov-2 process possess the no after effect property', *Chinese Journal of Applied Probability and Statistics* **4**, 141–147. (In Chinese).

Masry, E. (1984), Spectral and probability density estimation from irregularly observed data, *in* E. Parzen, ed., 'Time series analysis of irregularly observed data. Proceedings of a symposium held at Texas A & M University, College Station, Texas, 10-13 February 1983', number 25 *in* 'Lecture notes in statistics', Springer-Verlag, Berlin-New York.

Masry, E. (1997), Spectral and multivariate probability density estimations of continuous-time stationary processes from randomly sampled data, *in* T. Subba Rao, M. B. Priestley and O. Lessi, eds, 'Applications of time series analysis in astronomy and meteorology', Chapman & Hall, London, chapter 6.

Øksendal, B. (1989), *Stochastic differential equation: an introduction with applications*, 2nd edn, Springer, Berlin.

Øksendal, B. (2000), *Stochastic differential equation: an introduction with applications*, 5th edn, Springer, Berlin.

Pandit, S. M. and Wu, S. M. (1975), 'Unique estimates of the parameters of a continuous stationary process', *Biometrika* **62**, 497–501.

Pandit, S. M. and Wu, S. M. (1983), *Time series and system analysis with applications*, Wiley, New York.

Phillips, P. C. B. (1972), 'The structural estimation of a stochastic differential equation system', *Econometrica* **40**, 1021–1041.

Priestley, M. B. (1981), *Spectral analysis and time series*, Probability and mathematical statistics, Academic Press, London.

Robinson, P. M. (1977), 'Estimation of a time series model from unequally spaced data', *Stochastic Processes and their Applications* **6**, 9–24.

Robinson, P. M. (1980), Continuous model fitting from discrete data, *in* D. R. Brillinger and G. C. Tiao, eds, 'Reports on directions in time series: Proceedings of the IMS Special Topics meeting on time series analysis, Iowa State University, 1-3 May 1978', Institute of Mathematical Statistics, Hayward, California, pp. 263–278.

Rogers, L. C. G. and Williams, D. (1994*a*), *Diffusions, Markov processes and martingales*, Vol. 1: Foundations, 2nd edn, Cambridge University Press.

Rogers, L. C. G. and Williams, D. (1994*b*), *Diffusions, Markov processes and martingales*, Vol. 2: Diffusions, Markov processes and martingales, 2nd edn, Cambridge University Press.

Saff, E. B. and Snider, A. (1993), *Fundamentals of complex analysis for mathematics, science and engineering*, 2nd edn, Prentice Hall, New York.

Sauer, N. (2001), *Partial differential equations: The next step*, University of Pretoria.

Solo, V. (1984), Some aspects of continuous-discrete time series modelling, *in* E. Parzen, ed., 'Time series analysis of irregularly observed data. Proceedings of a symposium held at Texas A & M University, College Station, Texas, 10-13 February 1983', number 25 *in* 'Lecture notes in statistics', Springer-Verlag, Berlin-New York, pp. 325–345.

Telser, L. G. (1967), 'Discrete samples and moving sums in stationary stochastic processes', *Journal of the American Statistical Association* **62**, 484–499.

Tong, H. (1990), *Non-linear time series: A dynamical system approach*, Oxford University Press, Oxford.

Tsai, H. and Chan, K. S. (2000), Comparison of two discretization methods for estimating continuous-time autoregressive models, *in* W.-S. Chan, W. Li and H. Tong, eds, 'Proceedings of the Hong Kong international workshop on statistics in finance: An interface, 4-8 July 1999', University of Hong Kong: Centre of Financial Time Series, Imperial College Press, London, pp. 68–85.

Wiberg, D. M. (1971), *Theory and problems of state space and linear systems*, Schaum's Outline Series, McGraw-Hill, New York.