PERGAMON Mathematical and Computer Modelling 36 (2002) 1409–1424

MATHEMATICAL AND COMPUTER MODELLING

www.elsevier.com/locate/mcm

A Note on the Modelling and Analysis of Vector Arma Processes with Nonstationary Innovations

N. SINGH AND V. S. S. YADAVALLI
Department of Industrial and Systems Engineering, University of Pretoria
Pretoria 0002, South Africa

M. S. PEIRIS*
School of Mathematics and Statistics, University of Sydney
Sydney, N.S.W. 2006, Australia

(Received August 2001; revised and accepted April 2002)

Abstract—This paper considers the modelling and analysis of nonstationary vector ARMA processes from the theoretical point of view. Some extensions to the existing work (see, for example [1,2]) incorporating the nonstationarity of the underlying process are given using models with nonstationary innovations. It is shown that this class of models provides a very strong framework for many practical situations. The estimation procedures are discussed. Various prediction results are also given. Some examples are added to illustrate the theory.

Keywords—Hilbert space, Vector processes, Uniformly bounded, Regularity conditions, Nonstationary, White-noise, Prediction, Multivariate, Yule-Walker equations, ARIMA process, Innovations, Estimation.

1. INTRODUCTION

It is well known that the problem of identification of a time series (even if it is a univariate) is not an easy task in practice. Although there are some general guidelines laid down in the literature, no clear deterministic approach to the problem exists. The final decision is rather based on subjective judgment of the analysts. The problem becomes more difficult when the process is multivariate.

In practice, multivariate (or vector) time series arise when we observe k time series $Y_1(t), Y_2(t), \dots, Y_k(t)$; $k \geq 2$ simultaneously at time t. For example, we may be interested in studying the simultaneous variations in current and voltage over a period of time; variation in the volume, temperature, and pressure of a product in a chemical plant. Similarly, in economics, the price

*Author to whom all correspondence should be addressed. Current address: Department of Statistics, The Pennsylvania State University, University Park, PA 16802-2111, U.S.A.

The authors wish to thank two anonymous referees and the editor for their constructive comments and useful suggestions to improve the quality and the readability of the paper. We also thank P. Ainkaran for his help in simulations.

0895-7177/02/$ - see front matter
 Typeset by $\mathcal{AMS}$-TEX
PII: S0895-7177(02)00297-2

of a certain commodity, its production, and demand form a multivariate time series. Seismic records taken at two or more geographical sites form another example of a vector time series. Though each of the phenomena cited above can be studied separately in its own right as a univariate process, however, such a study will not reveal the true and important features such as the interdependency between any two series (the function measuring this interdependency is called the cross-covariance function). For this reason and practical applications points of view, the stationary vector ARMA processes have been studied in the literature (see, for instance [3–10]). Jenkins and Alavi [11] has also reported some interesting results in this class of models. Among others, the theory of nonstationary ARMA processes has been discussed by [2,11–15]. However, there are many practical situations (for example, seismic recordings at two or more sites, variations in current and voltage, and variations in volume, temperature, and pressure, etc.) that the time series analysts observe time series with stationarity in mean but not in the covariance structure. Although the theory of [11–13] can be used in these cases, an alternative simple way of modelling such a situation is to use the vector ARMA models with nonstationary innovations. The main reason for this approach is that the proposed model involves a fewer number of parameters than in [12] or [13]. In their recent paper [2], Singh and Peiris considered the class of multivariate ARMA processes with nonstationary innovations and studied some interesting statistical properties of the underlying process. The univariate version of the problem has been studied by Niemi [16]. In this current paper, the authors further investigate the vector problem and report some extended results. In addition, the estimation and prediction problems are also studied. Some examples are added to support the theoretical investigations and to provide guidelines in practical applications.

With that view in mind, Section 2 of this paper is devoted to review the general theory of vector integrated autoregressive-moving average (VARIMA) process and their properties. Further, we introduce the class of vector ARMA processes with nonstationary covariance structure in the latter part of Section 2. Some basic results are given in Section 3. Section 4 reports the main results for the nonstationary innovations case. Simplified results for some practical models are also given in Section 4. The identification problem is addressed in Section 5 and a graphical approach is suggested for the specification of certain models. A few bivariate examples are discussed. The estimation of parameters is considered in Section 6 and the prediction problem is discussed in Section 7.

2. VARIMA PROCESSES

A vector ARIMA model of orders $(p, d_1, \dots, d_k, q)$ is given by (abbreviated as VARIMA$(p, d_1, \dots, d_k, q)$)

$$\Phi(B) D\left(\nabla^{d_1}, \dots, \nabla^{d_k}\right) \mathbf{Y}(t) = G(B)\epsilon(t), \tag{2.1}$$

where

(i)

$$\begin{aligned} \Phi(B) &= I - \Phi_1 B - \dots - \Phi_p B^p, \quad \text{and} \\ G(B) &= I - G_1 B - \dots - G_q B^q \end{aligned} \tag{2.2}$$

are $k \times k$ matrix operators (called AR and MA operators, respectively) such that $\det[\Phi(z)] \neq 0$ and $\det[G(z)] \neq 0$ for all $|z| \leq 1$;

(ii) $D(\nabla^{d_1}, \dots, \nabla^{d_k})$ is a $k \times k$ diagonal matrix, where ∇^{d_i}, $i = 1, \dots, k$ are the differencing operators, where d_is may or may not be all equal;

(iii) $\mathbf{Y}(t) = (Y_1(t), \dots, Y_k(t))^\top$ and $\boldsymbol{\varepsilon}(t) = (\varepsilon_1(t), \dots, \varepsilon_k(t))^\top$ are k-dimensional vectors, where $\{\boldsymbol{\varepsilon}(t)\}$ is a stationary process such that $E\{\boldsymbol{\varepsilon}(t)\} = \mathbf{0}$, $E\{\boldsymbol{\varepsilon}(t)\boldsymbol{\varepsilon}^\top(s)\} = \Omega\delta_{ts}$ ($\top$ is the transpose operator), $\mathbf{0}$ is the $k \times 1$ column vector of zeros, Ω is a positive definite matrix of order $k \times k$ and δ_{ts} is the Kronecker delta satisfying $\delta_{ts} = 1$ for $t = s$ and 0, otherwise;

(iv) I is the identity matrix of order k.

If sufficient differencing is carried out to ensure the stationarity of $\mathbf{X}(t) = D(\nabla^{d_1}, \dots, \nabla^{d_k})\mathbf{Y}(t)$, equation (2.1) simplifies to

$$\Phi(B)\mathbf{X}(t) = G(B)\varepsilon(t). \tag{2.3}$$

Notice that $\mathbf{X}(t) = (X_1(t), \dots, X_k(t))^\top$ is a stationary vector ARMA(p, q) process. This process has been studied in [3–5,7,11,17].

This classical way of modelling can be done for a particular class of processes satisfying homogeneous nonstationarity (see, for instance [17,18]). However, there are situations in practice that one cannot clearly identify the above homogeneous nonstationarity, however, the process is, in fact, nonstationary. An alternative way of modelling the latter case is to replace the process $\{\epsilon(t)\}$ by a nonstationary process as given below.

Suppose that the vector innovation process $\{\mathbf{e}(t)\}$ is nonstationary in the sense that its covariance structure is time-dependent. Then, obviously, the vector process $\{Y(t)\}$ defined by

$$\Phi(B)\mathbf{Y}(t) = G(B)\mathbf{e}(t) \tag{2.4}$$

is also nonstationary, where $\Phi(B)$ and $G(B)$ are defined in (2.2), the vectors $\mathbf{Y}(t) = \{Y_1(t), \dots, Y_k(t)\}^\top$ and $\mathbf{e}(t) = \{e_1(t), \dots, e_k(t)\}^\top$ are k-dimensional column vectors such that for all u and t

$$E\{\mathbf{Y}(t)\} = 0, \qquad E\mathbf{e}(t)\} = 0, \tag{2.5}$$

$$\mathrm{Cov}\{\mathbf{Y}(t), \mathbf{Y}(u)\} = E\left\{\mathbf{Y}(t)\mathbf{Y}^\top(u)\right\} = \Gamma(t, u), \tag{2.6}$$

$$\mathrm{Cov}\{\mathbf{e}(t), \mathbf{e}(u)\} = E\left\{\mathbf{e}(t)\mathbf{e}^\top(u)\right\} = \Omega(t, u)\delta_{tu}, \tag{2.7}$$

where $\Gamma(t, u)$ and $\Omega(t, u)$ are $k \times k$ positive definite matrices with time-dependent entries and assumed to be uniformly bounded and bounded away from zero for each u and t. For simplicity, we replace $\Omega(t, u)$ by $\Omega(t)$ since $\delta_{tu} = 1$ for $u = t$ and 0, otherwise.

NOTE. A square matrix $M(t)$ of order k is said to be uniformly bounded and bounded away from zero for all t if the maximal characteristic root of $M(t)$ is such that

$$0 < \max_i(\lambda_i(t)) \le c < \infty, \tag{2.8}$$

for all t, where c is a finite constant.

Niemi [16] considered the univariate version of (2.4).

3. SOME BASIC RESULTS

Suppose that $\mathcal{H}(\cdot)$ denotes the Hilbert space of all real-valued univariate random variables with a common mean and finite second-order moments. Let

$$\left\{\boldsymbol{\xi}(t) = (\xi_1(t), \dots, \xi_j(t), \dots, \xi_k(t))^\top;\ t \in \mathbf{Z} = (0, \pm 1, \pm 2, \dots)\right\}$$

be a k-dimensional random process such that each element $\xi_j(t)$, $j = 1, \dots, k$ belongs to $\mathcal{H}(\cdot)$ for all t.

Define a closed linear k-dimensional subspace $\mathcal{H}(\xi, t) = \mathcal{H}(\cdot) \oplus \cdots \oplus \mathcal{H}(\cdot)$, (i.e., the direct sum of k copies of $\mathcal{H}(\cdot)$ spanned by the random variables $\{\xi_j(s);\ \xi_j(s) \in \mathcal{H}(\cdot),\ s \le t,\ j = 1, \dots, k\}$) and, similarly define another closed linear k-dimensional subspace $\mathcal{H}(\xi)$ spanned by all random variables $\{\xi_j(s);\ \xi_j(s) \in \mathcal{H}(\cdot),\ s \in \mathbf{Z},\ j = 1, 2, \dots, k\}$.

Now we state some regularity conditions in parallel with stationarity and invertibility results in the usual case.

Model (2.4) is said to satisfy the AR-regularity condition, if all the zeros of $\det[\Phi(x)]$ lie outside the unit circle. This condition ensures that the matrix expansion

$$\Phi^{-1}(B)\,G(B) = I - \sum_{j=1}^{\infty} \alpha_j B^j \tag{3.1}$$

converges for $|B| < 1$, where I is the $k \times k$ identity matrix. (See, for instance, [3,4,9] for the stationary case.) This implies that the covariance matrix of $Y(t)$ has finite elements. If the process $\{e(t)\}$ is stationary, this condition is called the stationary condition. Furthermore, model (2.4) is said to satisfy the MA regularity condition if all zeros of $\det[G(x)]$ lie outside the unit circle. This is equivalent to the invertibility condition in the usual case. This condition implies that the matrix expansion

$$G^{-1}(B)\,\Phi(B) = I - \sum_{k=1}^{\infty} \beta_k B^k \tag{3.2}$$

converges for $|B| < 1$. This ensures that the forecast weight functions associated with the vector ARMA process $\{Y(t)\}$ would die out for large k.

NOTE. It is easy to verify that (using a similar argument as in the stationary case) under certain conditions on $\Omega(t)$, the solution, $Y(t)$, in (2.4) converges, provided the process is AR regular.

The AR regularity condition ensures the causal representation of (2.4) given by

$$\mathbf{Y}(t) = \alpha(B)\,\mathbf{e}(t) = \sum_{k\geq 0} \alpha_k \mathbf{e}(t-k), \tag{3.3}$$

where $\alpha(B) = \Phi^{-1}(B)G(B)$, $\alpha_0 = I$ and the α_ks are matrices of constants. Similarly, $\det[G(x)] \neq 0$ for all $|x| \leq 1$ ensures the invertible representation

$$\mathbf{e}(t) = \beta(B)\,Y(t) = \sum_{k\geq 0} \beta_k Y(t-k), \tag{3.4}$$

where $\beta(B) = G^{-1}(B)\Phi(B)$, $\beta_0 = I$ and the β_k are matrices of constants (see, for instance [3]).

4. MAIN RESULTS FOR NONSTATIONARY INNOVATIONS

Consider the following lemma for later reference.

LEMMA 4.1. *Let $\{\mathbf{X}(t);\ t \in Z\}$ be a stationary k-dimensional zero mean purely nondeterministic process satisfying*

$$\Phi(B)\,\mathbf{X}(t) = G(B)\,\epsilon(t), \tag{4.1}$$

where

$$E\{\epsilon(t)\} = 0, \qquad E\left\{\epsilon(t)\,\epsilon^{\top}(s)\right\} = V\delta_{ts} \tag{4.2}$$

and V is a $k \times k$ positive definite matrix with constant elements. In this case, it can be seen that

$$\mathcal{H}(\mathbf{X};t) = \mathcal{H}(\epsilon;t). \tag{4.3}$$

(See [12,13,19] for details.)

Let $\{\mathbf{e}(t);\ t \in Z\}$ be a k-dimensional nonstationary stochastic process satisfying

$$E\{\mathbf{e}(t)\} = 0, \qquad E\left\{\mathbf{e}(t)\,\mathbf{e}^{\top}(s)\right\} = \Omega(t)\,\delta_{ts}. \tag{4.4}$$

We now state the following theorem due to [2].

THEOREM 4.1. *For all random processes $\{\boldsymbol{\epsilon}(t)\}$ and $\{\mathbf{e}(t)\}$ satisfying conditions (4.2) and (4.4), there exists a linear transformation S_t such that $S_t : H(e;t) \to H(\epsilon,t)$ with a bounded inverse $S_t^{-1} : H\epsilon(t) \to H(e;t)$.*

PROOF. See [2].

COROLLARY 4.1. *For all $X(t) \in H(\epsilon; t)$ and $Y(t) \in H(\mathbf{e}; t)$ the matrix S_t satisfies $X(t) = S_t Y(t)$ for all $t \in Z$.*

Note that

$$E\{X(t)\} = S_t E\{Y(t)\} \tag{4.5}$$

and

$$E\{X(t) X^\top(t)\} = S_t \{E(Y(t) Y^\top(t))\} S_t^\top = S_t \Gamma_{t,t} S_t^\top, \tag{4.6}$$

where

$$\Gamma_{t,t} = E\{Y(t) Y^\top(t)\}.$$

COROLLARY 4.2. *Let $\{X(t);\ t \in Z\}$ and $\{Y(t);\ t \in Z\}$ be k-dimensional stochastic processes satisfying (2.1) and (2.4). Then*

$$Y(t) = S_t^{-1} X(t) \tag{4.7}$$

is purely nondeterministic. The univariate version of these results are given by [16].

NOTE. The implications of Corollaries 4.1 and 4.2 is that the covariance structures of $X(t)$ and $Y(t)$ defined in (2.4) and (4.1) are similar. It may be noted from the above that $S_t Y(t)$ is a stationary vector process. The covariance structure of $X(t)$ when $X(t)$ is a (stationary) vector

(i) MA(q) process,

(ii) AR(p) process, and

(iii) a mixed ARMA(1, 1) process are given in [11].

In the following, we give the corresponding covariance structures of $Y(t) = S_t^{-1} X(t)$ for some simple cases.

In Theorems 4.2 and 4.4 below, we simplify some of our general results for the popular AR and MA models. We first consider an MA(q) process with nonstationary innovations.

THEOREM 4.2. *Let $\{Y(t)\}$ be a pure k-variate MA(q) process defined by*

$$Y(t) = e(t) - \sum_{j=1}^{q} G_j e(t-j), \tag{4.8}$$

where $e(t) \in H(\cdot)$ is a sequence of vectors satisfying (4.4). Furthermore, for any choice of $G_1, \dots, G_q$, let $Y^*(t)$ be a unique and bounded solution of the difference equation (4.8). Then $Y^*(t)$ has the property

$$E\{Y^*(t) Y^{*\top}(t-h)\} = \Gamma_{t,t-h},$$

where

$$\Gamma_{t,t-h} = \begin{cases} 0, & |h| > q, \\ -G_h \Omega(t-h) + G_{1+h} \Omega(t-h-1) G_1^\top + \cdots + G_q \Omega(t-h) G_{q-h}^\top, & 1 \le h \le q, \\ \Omega(t) + \sum\limits_{j=1}^{q} G_j \Omega(t-j) G_j^\top, & h = 0. \end{cases} \tag{4.9}$$

The following theorem reports the results for an AR(p) process.

THEOREM 4.3. *Let $\{Y(t)\}$ be a pure k-dimensional AR(p) process defined by*

$$\Phi(B) Y(t) = e(t). \tag{4.10}$$

Let $\{Y^*(t)\}$ be a uniquely determined and bounded solution of the difference equation (4.8) satisfying the AR regularity condition (9). Then the covariance function of $Y^*(t)$ is given by

$$\Phi(B) \Gamma_{t,t-j} = \begin{cases} \Omega(t), & j = 0, \\ 0, & j > 0, \end{cases} \tag{4.11}$$

where B operates on the first suffix t in $\Gamma_{t,t-j}$.

Now consider an ARMA(1,1) process.

THEOREM 4.4. *Let $\{Y(t)\}$ be a mixed k-vector ARMA(1,1) process defined by*

$$(I - \Phi_1 B) Y(t) = (I - G_1 B) e(t). \tag{4.12}$$

Suppose that $Y^*(t)$ be the uniquely determined bounded solution of the difference equation (4.8). Then the covariance of $Y^*(t)$ is

$$\Gamma_{t,t-j} = \begin{cases} \Gamma_{t-1,t}\Phi_1^\top + \Omega(t)(\Phi_1 - G_1) - \Omega(t) G_{-1}^\top, & j = 0, \\ \Gamma_{t-1,t-1}\Phi_1^\top - \Omega(t) G_1^\top, & j = 1, \end{cases} \tag{4.13}$$

and

$$\Phi(B)\Gamma_{t,\tau} = 0, \qquad \tau < t - q.$$

REMARK. It may be verified that (4.9), (4.11), and (4.13) are similar in form and structure to equations (3.3), (3.5) and (3.8) of [11] for the corresponding stationary cases.

5. IDENTIFICATION

The problem of identification is consistently one of the foremost problems in time series analysis and is certainly a difficult step in model building. Although, some general guidelines are laid down in the literature, no definite deterministic approach exists. It is rather subjective and heavily relies on judgment. However, the problem of estimation requires some preliminary specification of the model. In our opinion, it is important to have some specification of the model even though it might run the risk of being modified or discarded altogether at a later stage.

5.1. Specification of a Multivariate Stationary ARMA Model

At the start, we make an initial guess of the structure of a multivariate stochastic model based on sample autocorrelation, partial autocorrelation, and cross-correlation functions. This tentatively entertained model can be fitted, estimated, and tested. If the model is found inadequate based on the analysis of residuals, further cycles of identification, estimation, and testing can be carried out until a best fitted model is discovered.

5.2. Specification of a Vector ARMA Process with Nonstationary White-Noise

We have noted that the covariance structure of a nonstationary vector ARMA process $\{Y(t)\}$ defined in (2.4) is similar to that of the underlying stationary process $\{X(t)\}$ defined in (2.3). However, this information alone is not sufficient for the identification of the nonstationary process considered in this paper. As pointed out recently by [20] in the univariate case, there are in fact two problems for identification, namely

(i) the identification of the model, and
(ii) the specification of the error covariance structure.

In the following, we first discuss a method for identifying a model for a nonstationary process with time-dependent innovation process. It may be noted that the covariance structure of a nonstationary (NS) process with time-dependent error covariance (TEC) is similar to that of the underlying stationary process. Hence, a preliminary guess about the type of model for the NS process defined at (2.4) can be made by examining the sample autocorrelation, partial autocorrelation and the cross-correlation (between any two components of the series) functions of the given series $Y(t)$, $t = 1, \dots, n$. However, these correlation functions are of little help in identifying the error covariance structure. For this, we suggest a graphical approach in the following subsection.

5.3. Identification of the Error Covariance Structure

We now address the problem of identification of time-dependent error covariance structure in three relatively simple cases as they can be used in many practical problems. This procedure can be extended easily to more complex situations.

PROPOSITION 5.1. *Let $\{Y(t)\}$ be a k-dimensional time series and let the component variances of the error vector are linear functions of time "t" of the form $\sigma_{it} = a_i t + b_i$, $|a_i| \ll 1$, $i = 1, \ldots, k$, then the plot of the k-dimensional series $Y(t)$ will be a $(k+1)$-dimensional cone diverging out; however, if $\sigma it = 1/(a_i t + b)$, $|a_i| \ll 1$, the cone will be converging out, $i = 1, \ldots, k$. Here a condition like $|a_i| \ll 1$ is necessary to allow the change in variance to be very slow.*

PROPOSITION 5.2. *Let $\{Y(t)\}$ be a k-dimensional time series and let the component variances of error vector be exponential function of "t" of the form $\beta_i \exp(\alpha_i t)$, where $|\alpha_i| \ll 1$, $i = 1, \ldots, k$, then the plot of the series will be $(k+1)$-dimensional bell-shaped diverging (converging) out if $0 < \alpha_i < 1$ $(-1 < \alpha_i < 0)$, $i = 1, \ldots, k$. ($|\alpha_i| \ll 1$ ensures the slow change in $\beta_i \exp(\alpha_i t)$ with time t.)*

PROPOSITION 5.3. *Let $\{Y(t)\}$ be a k-dimensional time series and let the component variances of the error be quadratic in "t" of the form $a_i t^2 + b_i t + c_i$ such that $0 < a_i \ll 1$, $i = 1, \ldots, k$, b_i and c_i are small real numbers, then the plot of the series will be $(k+1)$-dimensional hyperboloid diverging out and converging out if the i^{th} component variance is of the form $(a_i t^2 + b_i t + c_i)^{-1}$.*

The first author examined about 50 bivariate series in each case and found the results conforming to Propositions 5.1–5.3. In Sections 5.4 and 5.5, we discuss two examples of a bivariate time series for illustration.

5.4. When the Component Variances of the White-Noise Process are Linear Functions of Time

Consider the following bivariate AR(1) process given by:

$$\Phi(B)\, X(t) = \epsilon(t), \tag{5.1}$$

where

$$\Phi(B) = \begin{bmatrix} 1 - \phi_{11}B & -\phi_{12}B \\ -\phi_{21}B & 1 - \phi_{22}B \end{bmatrix}.$$

It is assumed that $\epsilon(t) = (\epsilon_1(t), \epsilon_2(t))^\top$ follows the bivariate normal distribution with the covariance matrix

$$V = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}.$$

As in Section 4, let $Y(t) = S_t X(t)$ and $\epsilon(t) = S_t e(t)$. Then it follows that

$$\phi(B)\, Y(t) = e(t), \tag{5.2}$$

such that

$$E[e(t)] = 0, \qquad E\left[e(t)\, e^\top(t)\right] = \Omega(t),$$

where

$$S_t = V^{1/2} \Omega^{-1/2}(t).$$

For simplicity, let the variances of the components of $e(t) = (e_1(t), e_2(t))^\top$ be the linear functions of "t" and let $e_1(t)$ and $e_2(t)$ be uncorrelated such that,

$$\Omega(t) = \begin{pmatrix} at + b & 0 \\ 0 & ct + b \end{pmatrix}, \qquad |a| \ll 1, \quad |c| \ll 1. \tag{5.3}$$

If $V = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, then

$$S_t = \begin{pmatrix} \dfrac{1}{\sqrt{at + b}} & 0 \\ 0 & \dfrac{1}{\sqrt{ct + d}} \end{pmatrix} \tag{5.4}$$

and

$$\begin{pmatrix} Y_1(t) \\ Y_2(t) \end{pmatrix} = \begin{pmatrix} \sqrt{at+b} & 0 \\ 0 & \sqrt{ct+d} \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix}. \tag{5.5}$$

Putting $\phi_{11} = -1$, $\phi_{12} = 2.2$, $\phi_{21} = 1$, and $\phi_{22} = 2$ in (5.1), we have

$$\begin{aligned} X_1(t) &= -X_1(t-1) - 2.2X_2(t-1) + E_1(t), \\ X_2(t) &= 2X_2(t-1) + X_1(t-1) + E_2(t), \end{aligned} \tag{5.6}$$

and for further simplicity, let $a = 0.1$, $b = 0$, $c = 0.15$, and $d = 0$, in (5.5),

$$\begin{aligned} Y_1(t) &= \sqrt{0.1t}X_1(t), \\ Y_2(t) &= \sqrt{0.15t}X_2(t). \end{aligned} \tag{5.7}$$

Since V is the identity matrix of order two, we first generated two standard normal random samples each of size 200. Then the bivariate samples on $X(t)$ and $Y(t)$ were generated using (5.6) and (5.7). They are plotted in Figures 1 and 2, respectively, in the Appendix.

REMARK 5.1. It may be noted from Figures 1 and 2 that $X(t)$ is a stationary series and $Y(t)$ is stationary in mean and nonstationary in variance as expected.

REMARK 5.2. It may also be noted that since the variances of the components of $Y(t)$ are linear functions of time, the three-dimensional plot of $Y(t)$ is cone-shaped. The acfs and pacfs of $Y_1(t)$ and $X_1(t)$ may be looked at for comparison.

INTERPRETATION. It may be reminded that the bivariate series $X(t)$ is stationary and the series $Y(t) = S_t X(t)$ is nonstationary in variance. Now on examining the acfs and pacfs of the components of $X(t)$ and $Y(t)$, of $X_1(t)$, $Y_1(t)$ in Figures 3 and we observe that they are very similar to each other. Similarly, the acfs and pacfs of $X_2(t)$ and $Y_2(t)$ are also very close to each other. Further, the cross-correlations of $X_1(t)$, $X_2(t)$ and those of $Y_1(t)$, $Y_2(t)$ are more or less the same. Therefore, given a bivariate series $Y(t)$ which is nonstationary in variance, the acfs, pacfs of its components and their ccf as such will be misleading since they will suggest that $Y(t)$ is a stationary series in fact it is not. However, they do point out to the specific underlying stationary BARMA process $\{X(t)\}$ which has similar acfs, pacfs and the cross-correlation function. On the other hand, the three-dimensional plots of $Y(t)$ and $X(t)$ are different by their very different nature and further we note that since the variances of the components of $Y(t)$ are linear functions of time, the three-dimensional plot of $Y(t)$ is cone-shaped and that of $X(t)$ is cylindrical. Thus, from the above, we arrive at the following.

PROCEDURE AND CONCLUSION. Given a bivariate time series $Y(1), \dots, Y(n)$, we first obtain the three-dimensional plot the series and then calculate the acfs, pacfs, and ccfs of its components. Now if

(i) the three-dimensional plot of the series is cone-shaped, it suggests that the variances of the components of $Y(t)$ are linear functions of time, and
(ii) the acfs, pacfs, and ccfs are similar in pattern to those of a known stationary bivariate process $X(t)$, then the series $Y(t)$ can be identified as $X(t)$ with its components having linearly time-dependent variances.

In support of the above, we discuss below another example.

5.5. An Example with the Component Variances of the Noise-Process are Exponential Functions of Time

Consider the same bivariate process $\{X(t)\}$ given in (5.2). Let S_t be defined by

$$S_t = \begin{pmatrix} \dfrac{1}{\sqrt{\alpha}\exp(-\beta t/2)} & 0 \\ 0 & \dfrac{1}{\sqrt{\gamma}\exp(-\delta t/2)} \end{pmatrix}. \tag{5.8}$$

Then

$$\begin{pmatrix} Y_1(t) \\ Y_2(t) \end{pmatrix} = \begin{pmatrix} \sqrt{\alpha} \exp\left(\frac{-\beta t}{2}\right) & X_1(t) \\ \sqrt{\gamma} \exp\left(\frac{-\delta t}{2}\right) & X_2(t) \end{pmatrix}. \tag{5.9}$$

A NUMERICAL EXAMPLE. Using Splus, two standard normal random samples were generated each of size 300 and then keeping the same values of ϕ_{11}, ϕ_{12}, ϕ_{21}, and ϕ_{22} as given before, a bivariate sample $(X_1(1),\ X_2(1)), \dots, (X_1(300),\ X_2(300)))$ was generated using (5.6). Then putting $\alpha = \gamma = 0.5$ and $\beta = \delta = 0.03$, we generated two random samples $(Y_1(1),\ Y_2(1), \dots, Y_1(300),\ Y_2(300))$ using (5.9). Samples on $Y_1(t)$, $Y_2(t)$, $t = 1, \dots, 300$ were plotted in Figures 3 and 4, respectively (see Appendix).

REMARK 5.3. It may be observed from Figures 3 and 4 that the component series $Y_1(t)$ and $Y_2(t)$ are nonstationary in variances which are exponential functions of "t". They are almost symmetric on both sides of means and diverging exponentially on both sides of means.

REMARK 5.4. It may be noted from Figures 3 and 4 that the plots of the bivariate series $Y(t)$ with components each having the variance equal to $0.5\exp(0.03t)$, is bell-shaped. Thus, we arrive at the following for a given a bivariate time series $Y(1), \dots, Y(n)$.

5.6. Conclusion

(i) If the plots of the component series are exponentially increasing (decreasing) almost symmetrically on both sides of means and further, if the three-dimensional plot of the bivariate series $Y(t)$ looks bell-shaped, then it can be concluded that the variances of the component series are exponential functions of the type $\alpha\exp(-\beta t)$ for suitable values of α and β.

(ii) Since the acfs, pacfs, and ccfs of the components of $Y(t)$ are very similar to those of $X(t)$, the series $Y(t)$ can be identified (with the underlying series $X(t)$) with error components having variances as time-dependent exponential functions stated above.

Thus, Criteria (i) and (ii) above would lead to the preliminary identification of the given bivariate series $Y(t)$.

6. ESTIMATION OF MODEL PARAMETERS

After arriving at some preliminary specification of the model structure, one should have reasonably good initial estimates of the parameters which may be used later for the maximum likelihood estimators described in a subsequent section.

6.1. Initial Estimates of Parameters

In this section, we discuss methods for calculating initial estimates for pure moving average and pure autoregressive models.

Pure vector moving average model of order q

Following the procedure described in Section 4, if the underlying process is suspected to be an MA(q), $q = 1, 2, \dots$, then the initial estimates of the parameters may be obtained by substituting the sample estimates for the theoretical covariance matrices $\Gamma_{t,t-h}$, $h = 0, 1, 2, \dots$, in (4.9) leading to

$$\hat{\Omega}(t) = C_{t,t} - \sum_{j=1}^{q} \hat{G}_j \hat{\Omega}(t-j) \hat{G}_j^{\top} \tag{6.1}$$

and

$$\hat{\Omega}(t-h)\hat{G}_h^{\top} = C_{t,t-h} + \sum_{j=1}^{q-h} \hat{G}_j \hat{\Omega}(t-h-j)\hat{G}_{j+h}, \tag{6.2}$$

where $C_{t,t-h},\ h = 0,1,2,\dots$ are the sample covariance matrices at time t. Since these matrices are not independent of t, we suggest that they may be estimated using the more homogeneous observations around a fixed time $t = \tau$ by

$$C_{\tau,\tau-h} = \frac{1}{N-1} \sum_{k=-\ell}^{\ell} \left(Y(\tau+h) - \bar{Y}_\tau\right)\left(Y(\tau-h+k) - \bar{Y}_\tau\right)^{\top}, \qquad h = 0,1,2,\dots, \tag{6.3}$$

where $\ell = (N-1)/2$ and N is odd and

$$\bar{Y}_\tau = \sum_{k=-\ell}^{\ell} \frac{Y(\tau+k)}{N}. \tag{6.4}$$

Furthermore, it may be observed from (6.1) and (6.2) that only the matrices $C_{t,t-h}$, $h = 0,1,2,\dots$ for fixed t are known and the matrices $\Omega(t-j)$ and $G_j,\ j = 0,1,2,\dots,q$ are unknown. However, given some initial values of unknowns, the required solutions may be obtained without much difficulty following the procedure illustrated below for a multivariate MA(2) process. First, we consider the case when all covariance matrices are independent of t.

CASE 6.1. When the covariance matrices are independent of t, then for $q = 2$, we have from (6.1) and (6.2) that

$$\Omega = C_0 - G_1 \Omega G_1^{\top} - G_2 \Omega G_2^{\top}, \tag{6.5}$$

$$G_1^{\top} = -\Omega^{-1} C_1 + \Omega^{-1} G_1 \Omega G_2^{\top}, \tag{6.6}$$

$$G_2^{\top} = -\Omega^{-1} C_2. \tag{6.7}$$

Substituting for $G_2^{\top}$ from (6.7) in (6.6) gives

$$G_1^{\top} = -\Omega^{-1} C_1 - \Omega^{-1} G_1 C_2. \tag{6.8}$$

Given the initial value Ω_0 and using the property of three square matrices D, E, and F that

$$\mathrm{Vec}(DEF) = \left(F^{\top} \otimes D\right) \mathrm{Vec}(E) \tag{6.9}$$

and

$$\mathrm{Vec}(DE) = (I \otimes D)\,\mathrm{Vec}(E), \tag{6.10}$$

where $\mathrm{Vec}(A) = (A_{\cdot 1}, A_{\cdot 2}, \dots, A_{\cdot p})^{\top}$, and $A_{\cdot j}$ is the j^{th} column of A and $A \otimes B$ is the Kronecker product of two matrices A and B. Now the solution to (6.8) is given by

$$\mathrm{Vec}\left(G_1^{\top}\right) = \left[I - \left(C_2 \otimes \Omega_0^{-1}\right)\right]^{-1} \left[-\left(I \otimes \Omega_0^{-1}\right) \mathrm{Vec}(C_1)\right]. \tag{6.11}$$

After estimating G_1 from (6.11), and G_2 from (6.7), an estimate of Ω may be obtained from (6.5). Denoting this by Ω_1, the procedure can be repeated until a stable solution is obtained. It may be pointed out that for a convergent solution of $\mathrm{Vec}(G_1)$, a set of sufficient conditions is

$$\begin{aligned} \rho\left[C_2 \otimes \hat{\Omega}^{-1}\right] &< 1, \\ \rho\left[I \otimes \hat{\Omega}^{-1}\right] &< 1, \end{aligned} \tag{6.12}$$

where $\rho[A]$ is the spectral radius of matrix A defined by

$$\rho(A) = \max_i \{|\lambda_i(A)|\},$$

where $\lambda_i(A)$ is the i^{th} eigenvalue of A.

CASE 6.2. If the covariance matrices are not independent of "t", then for any fixed $t = \tau$ and given $\Omega_{\tau-1}$, $\Omega_{\tau-2}$ from some previous records and the sample covariance matrices $C_{\tau\tau}$, $C_{\tau,\tau-1}$, and $C_{\tau,\tau-2}$; G_1, G_2 and $\Omega(\tau)$ may be estimated following the procedure outlined above.

Pure vector autoregressive model of order p

For this process, we discuss two methods, namely

(i) the method of moments using the multivariate Yule-Walker equations, and
(ii) the method of maximum likelihood.

METHOD (i). The multivariate Yule-Walker equations. If the preliminary model to a series $Y(t)$ is found to be a vector AR(p) ($p \geq 1$) following the identification procedure described in Section 4, then the initial estimates of the parameter matrices Φ_is, may be obtained by replacing the theoretical covariance matrices $\Gamma_{t,t-k}$, $k = 0, 1, 2, \ldots$ by their counterpart sample covariances $C_{t,t-k}$. The normal equations are

$$\left(C^{\top}_{\tau-j,\tau-i}\right) \Phi^{\top} = C^{\top}_{\tau,p-i}, \tag{6.13}$$

where

$$\Phi^{\top} = \begin{pmatrix} \Phi_1^{\top} \\ \vdots \\ \Phi_p^{\top} \end{pmatrix}, \qquad C^{\top}_{\tau,p-1} = \begin{pmatrix} C^{\top}_{\tau,\ \tau-1} \\ \vdots \\ C^{\top}_{\tau,\ \tau-p} \end{pmatrix},$$

$i, j = 1, \ldots, p$. The suffix i denotes the i^{th} row and j denotes the j^{th} column. At time τ, the $Y - W$ equations (6.13) give estimates of the coefficient matrices of the VAR(p) model. The estimate of $\Omega(t)$ at $t = \tau$ may then be obtained from

$$\Omega(\tau) = C_{\tau,\tau} - \sum_{h=1}^{p} \hat{\Phi}_h \, C_{p-h,\tau}. \tag{6.14}$$

Estimates obtained from (6.13) and (6.14) may be obtained as preliminary estimates for the maximum likelihood estimation discussed in the succeeding section.

METHOD (ii). Maximum likelihood estimation of model parameters. For simplicity, we consider a k-variate AR(1) process defined by

$$Y(t) = \Phi Y(t-1) + e(t), \tag{6.15}$$

where $\Phi = \mathrm{diag}(\phi_{ii})$, $i = 1, \ldots, k$. The error vector $e(t)$ is assumed to be k-variate with mean 0 and covariance matrix.

$$\Omega(t) = \mathrm{diag}(a_i\, t + b_i), \qquad |a_i| \ll 1, \quad b_i \text{ is a constant} \tag{6.16}$$

and $a_i t + b_i \neq 0$, $i = 1, \ldots, k$.

Given observations $Y(1), Y(2), \ldots, Y(n)$, the ordinary least square estimates are given by

$$\hat{\phi}_{ii} = \frac{\sum_{t=2}^{n} Y_i(t)\, Y_i(t-1)}{\sum_{t=2}^{n} Y_i^2(t-1)}, \qquad i = 1, \ldots, k. \tag{6.17}$$

Assume that the vector residuals follows a k-variate normal distribution with mean 0 and covariance matrix $\Omega(t)$ defined in (6.16). Then conditioning on the initial values of the series $Y(t)$, $t = 1, \ldots, n$ being fixed (including zero) and unknown initial values of residuals $e(t)$ set equal to zero, the log-likelihood function is written as

$$\mathcal{L}\left(\beta, \Omega(t) \mid X(t)\right) \alpha - \frac{n}{2} \ln |\Omega(t)| - \frac{1}{2} \sum_t \left[e^\top(t) \Omega^{-1}(t) e(t)\right], \tag{6.18}$$

where $\beta = ((a_i, b_i, \phi_{11});\ i = 1, \dots, k)$ is the vector parameter and α is the proportionality symbol. Summation in (6.18) starts from the point when the $Y_i(t)$s are available. Replacing $e(t)$ in (6.18) by $Y(t) - \Phi Y(t-1)$, we have

$$\mathcal{L}(\beta, \Omega(t) \mid X(t)) \alpha - \frac{1}{2} n \,|\Omega(t)| - \frac{1}{2} \sum_t \left[Y(t) - \Phi Y(t-1)^\top \Omega^{-1}(t)(Y(t) - \Phi Y(t-1))\right]. \tag{6.19}$$

As in the classical multivariate analysis, it is obvious that maximizing (6.18) or (6.19) by differentiating is equivalent to minimizing $|\Omega(t)|$, that is, to minimizing the volume of the hyperellipsoid of constant probability in the k-dimensional space of $e(t)$. One way to minimizing $|\Omega(t)|$ would be to use the hill-climbing techniques, but as these are not well suited to the present purpose, we adopt the more direct method of differentiating (6.19) first with respect to vector β and then with respect to each parameter. In terms of $\hat{e}(t)$, we then have

$$\beta : \frac{n}{|\Omega(t)|} \frac{\partial |\Omega(t)|}{\partial \beta} + \sum_t e^\top(t) \frac{\partial \Omega^{-1}(t)}{\partial \beta} e^\top(t) = 0, \tag{6.20}$$

$$a_i : n|\Omega(t)|^{-1} t\Omega_i(t) + \sum_t e^\top(t) \left\{-\Omega^{-1}(t) J_{ii} \Omega^{-1}(t)\right\} e(t) = 0, \qquad i = 1, \dots, k, \tag{6.21}$$

where $\Omega_i(t)$ is the cofactor of a_i and J_{ii} is the matrix with the coefficient of a_i, i.e., "t" in the $(i,i)^{\text{th}}$ position and zeros elsewhere.

$$b_i : n|\Omega(t)|^{-1} \Omega_i(t) + \sum_t e^\top(t) \left\{-\Omega^{-1}(t) I_{ii} \Omega^{-1}(t)\right\} e(t) = 0, \qquad i = 1, \dots, k, \tag{6.22}$$

where I_{ii} is the matrix with 1 in the $(i,i)^{\text{th}}$ position and zeros elsewhere.

$$\phi_{ii} : \hat{\phi}_{ii} = \frac{\sum\limits_{t=1}^{n} Y_i(t-1)}{\sum\limits_{j=2}^{n} Y_j^2(t-1)}, \qquad i = 1, \dots, k. \tag{6.23}$$

From (6.17) and (6.23), it may be noted that the least square and maximum likelihood estimators are the same as expected (under the assumption of normality).

However, a better estimation procedure can be developed using the estimation function approach as in [21,22]. This work is in progress and will be presented in a future paper.

7. PREDICTION

The prediction for the process $\{Y(t);\ t \in Z\}$ can be made in the same way as for the stationary process $\{X(t);\ t \in Z\}$ since the spaces $\mathcal{H}(e; t)$ and $\mathcal{H}(\epsilon; t)$ are isomorphic. Let the best linear ℓ-step ahead predictor for $Y(t + \ell)$ be denoted by $\hat{Y}_\ell(t);\ \ell > 0$ and defined by

$$\hat{Y}_\ell(t) = P_{\mathcal{H}(Y;t)} Y(t + \ell), \qquad \ell > 0, \tag{7.1}$$

where $P_{\mathcal{H}(Y;t)}$ denotes the orthogonal projection of $\mathcal{H}(Y)$ onto $\mathcal{H}(Y; t)$. Consider now the following results associated with the prediction problem. In the following theorem, we assume that $\{Y(t)\}$ in (2.4) satisfies both AR and MA regularity conditions.

THEOREM 7.1. *Let the process $\{Y(t);\ t \in Z\}$ be generated by (2.4). Then $\hat{Y}_\ell(t)$ is the uniquely determined bounded solution of the k-variate ARMA$(p, \max(q-\ell, 0))$ model.*

PROOF. From (2.4), we have

$$\Phi(B)\, Y(t+\ell) = G(B)\, e(t+\ell)\,. \tag{7.2}$$

Since $\mathcal{H}(e;t)$ and $\mathcal{H}(\epsilon;t)$ are isomorphic, we have

$$\mathcal{H}(Y;t) = \mathcal{H}(e;t)\,.$$

Considering the orthogonal projection $P_{\mathcal{H}(Y;t)}$ on both sides of (7.2), we get

$$P_{\mathcal{H}(Y;t)}\Phi(B)\, Y(t+\ell) = P_{\mathcal{H}(e;t)} G(B)\, e(t+\ell)\,. \tag{7.3}$$

Since

$$P_{\mathcal{H}(e;t)} G(B)\, e(t+\ell) = \begin{cases} 0, & \text{if } q-\ell < 0, \\ -G_{q-\ell}e(t) - G_{q-\ell+1}e(t-1) - \cdots - G_q e(t-q+\ell)\,, & \text{if } q-\ell \geq 0. \end{cases} \tag{7.4}$$

Thus, from (7.3) and (7.4), it follows that

$$\Phi(B)\, \hat{Y}_\ell(t) = \begin{cases} 0, & \text{if } q-\ell < 0, \\ -G_{q-\ell}e(t) - G_{q-\ell+1}e(t-1) - \cdots - G_q e(t-q+\ell)\,, & \text{if } q-\ell \geq 0. \end{cases} \tag{7.5}$$

Hence, the proof.

THEOREM 7.2. *The $\hat{Y}_\ell(t);\ \ell > 1$ can be recursively obtained from*

$$\sum_{j=0}^{\ell-1} \beta_j \hat{Y}_{t-j}(t) = \sum_{j=0}^{\infty} \beta_{\ell+j} Y(t-j)\,. \tag{7.6}$$

PROOF. From $\sum_{j=0}^{\infty} \beta_j Y(t-j) = e(t)$, we have

$$\sum_{j=0}^{\ell-1} \beta_j Y(t+\ell-j) + \sum_{j=\ell}^{\infty} \beta_j Y(t+\ell-j) = e(t+\ell)\,, \qquad \ell \geq 1.$$

Considering the projection $P_{\mathcal{H}(Y;t)}$, the result follows. (See [23,24] for similar derivations.)

REFERENCES

1. M.S. Peiris, Analysis of multivariate ARMA processes with nonstationary innovations, *Communications Statist. Theory and Methods* **19**, 2847–2852, (1990).
2. N. Singh and M.S. Peiris, A simulation study on vector ARMA processes with nonstationary innovation, *J. Statist. Comput. Simul.* **58**, 37–58, (1997).
3. E.J. Hannan, *Multiple Time Series*, John Wiley, New York, (1970).
4. M.B. Priestly, *Spectral Analysis and Time Series, Volume 2: Multivariate Series, Prediction and Control*, Academic Press, New York, (1989).
5. H. Lutkepohl, *Introduction to Multiple Time Series Analysis*, Second Edition, Springer-Verlag, New York, (1993).
6. S. Mittnik, Multivate time series analysis with state space models, *Computers Math. Applic.* **17** (8/9), 1189–1201, (1989).
7. G.C. Reisnel, *Elements of Multivariate Time Series Analysis*, Springer-Verlag, New York, (1993).
8. T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Second Edition, John Wiley & Sons. New York, (1984).
9. P.J. Brockwell and R.A. Davis, *Time Series: Theory and Methods*, Second Edition, Springer-Verlag, New York, (1991).

10. G.C. Tiao and R.S. Tsay, Model specification in multivariate time series, *Journal of Royal Statist. Soc. B.* **51**, 157–213, (1989).
11. G.M. Jenkins and A.S. Alavi, Some aspects of modelling and forecasting multi-variate time series, *Journal of Time Series Analysis* **2**, 1–47, (1981).
12. M.S. Peiris, On the prediction of multivariate ARMA processes with time dependent covariance structure, *Commun. Stat. Theory Meth.* **17**, 27–37, (1988).
13. M. Hallin, Mixed autoregressive moving average multivariate processes with time dependent coefficients, *J. Multiv. Anal.* **8**, 567–572, (1978).
14. P.J. Brockwell and H. Mitchell, Linear prediction for a class of multivariate stable processes, *Commn. Statist.-Stochastic Models* **14** (1/2), 297–310, (1998).
15. M.S. Peiris and A. Thavaneswaran, Multivariate stable ARMA processes with time dependent coefficients, *Metrika* **54**, 131–138, (2001).
16. H. Niemi, On the effect of non-stationary noise in ARMA models, *Scand. J. Statist.* **10**, 11–17, (1983).
17. G.E.P. Box and G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, California, (1976).
18. M.B. Priestley, *Nonlinear and Nonstationary Time Series Analysis*, Academic Press, New York, (1988).
19. M. Hallin, Nonstationary q-dependent processes and time varying moving-average models: Invertibility properties and the forecasting problem, *Adv. in Appl. Prob.* **18**, 170–210, (1986).
20. N. Singh, Identification of an ARMA process with nonstationary white-noise, Statistics Research Report No. 215, Department of Mathematics, Monash University, Clayton, Vic 3168, Australia, (1992).
21. V.P. Godambe, The foundations of finite sample estimation in stochastic processes, *Biometrika* **72**, 419–428, (1985).
22. A. Thavaneswaran and B. Abraham, Estimation of nonlinear time series models using estimating equations, *J. Time. Ser. Anal.* **9**, 99–108, (1988).
23. M.S. Peiris and A. Thavaneswaran, On the properties of some nonstationary ARMA processes with infinite variance, *Int. J. Modelling Simu.* **21**, 301–304, (2001).
24. K.S. Miller, *Linear Difference Equations*, Benjamin, New York, (1968).

APPENDIX

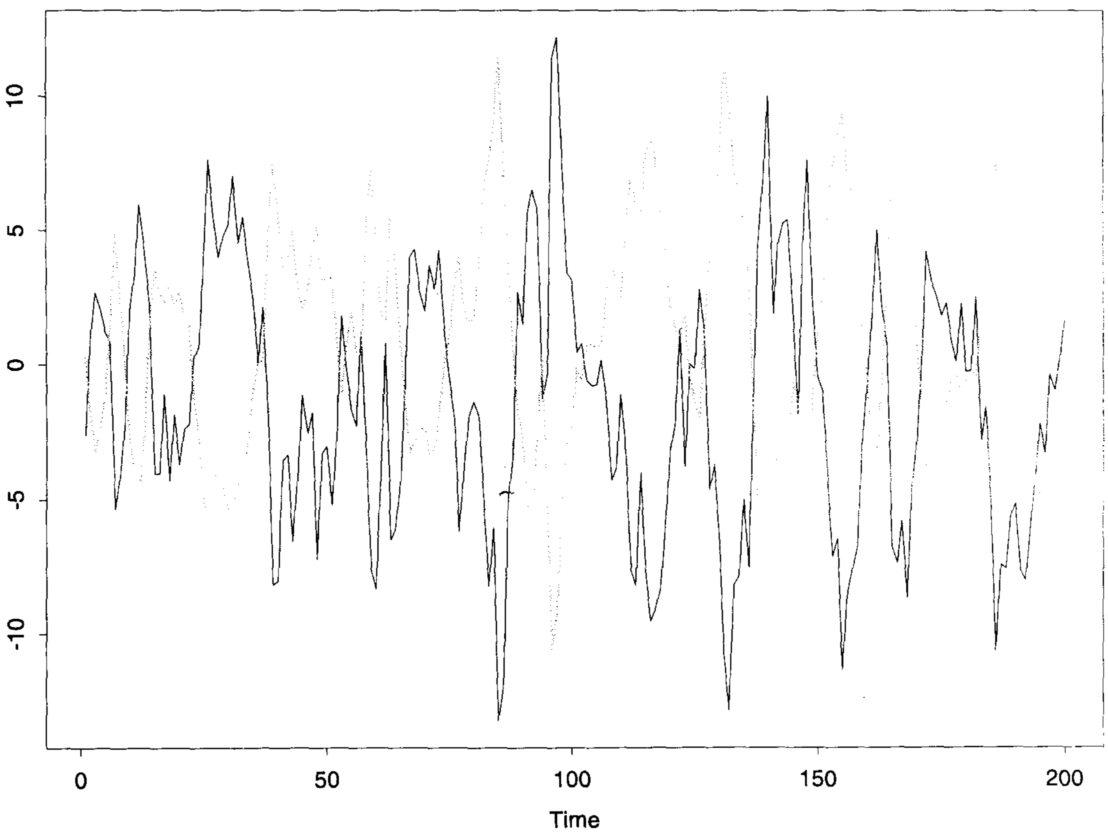

Figure 1. Time series plots of components $X1$ and $X2$ simulation using (5.6).

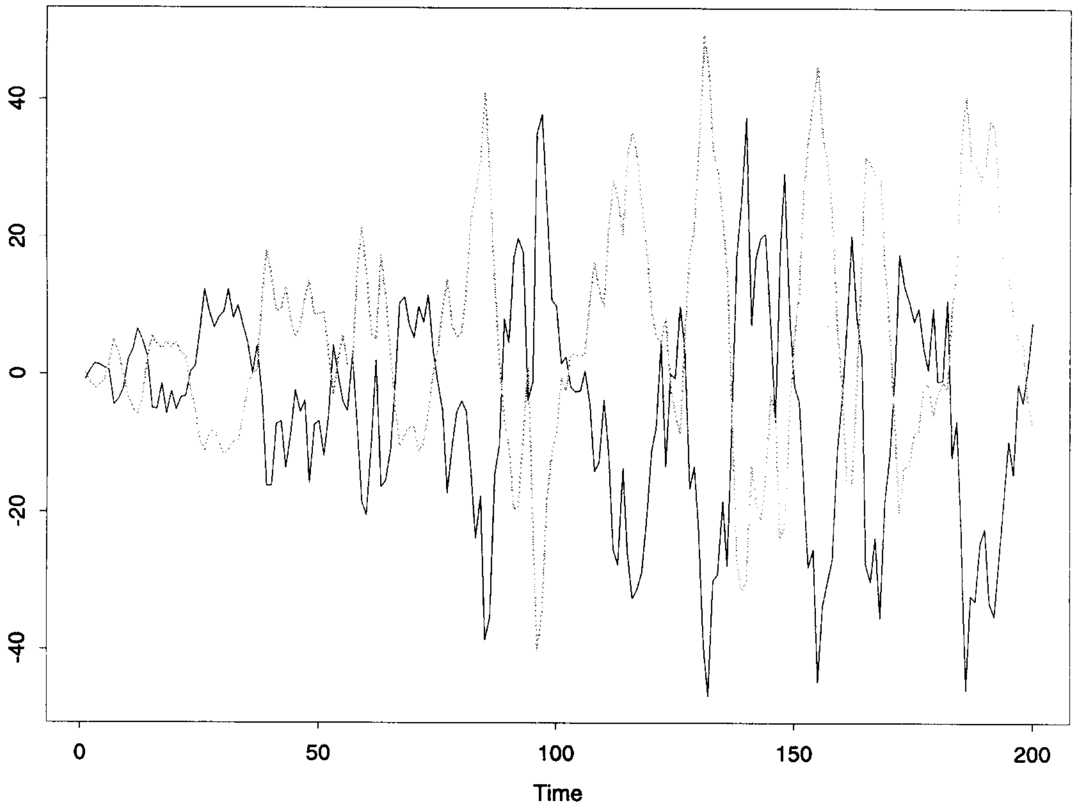

Figure 2. Time series plots of components $Y1$ and $Y2$ simulation using (5.7).

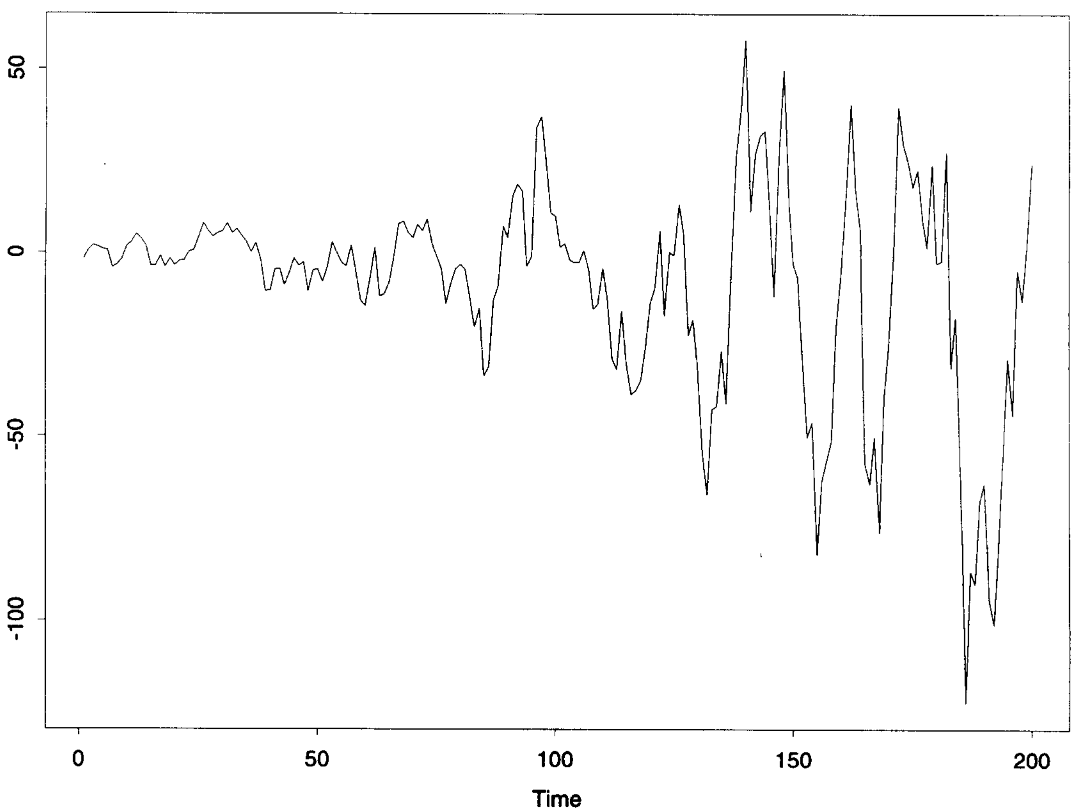

Figure 3. Time series plot of component $Y1$ simulation using (5.9).

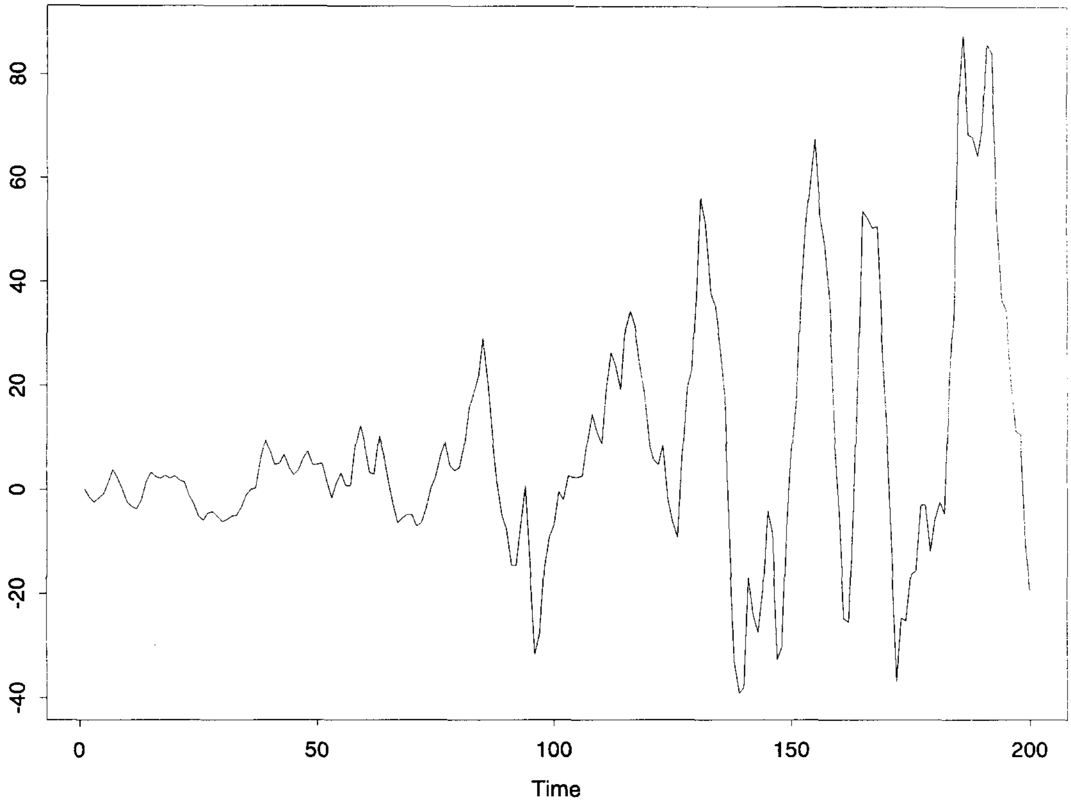

Figure 4. Time series plot of component $Y2$ simulation using (5.9).