SOME CONTRIBUTIONS TO THE STUDY OF BILINEAR TIME SERIES MODELS

by

IHEANYICHUKWU SYLVESTER AKAMANAM

A Thesis Submitted to the University of Sheffield for the Degree of Doctor of Philosophy.

AUGUST 1983

Department of Probability and Statistics

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. M. Bhaskara Rao, for his invaluable assistance during the preparation of this thesis. I am grateful to the Association of Commonwealth Universities for providing me with a grant to undertake this research.

A special acknowledgement should be made to my wife Chinyere for her moral support and for giving up time that belonged to her. Thanks also to Alexandra Crawshaw for her excellent typing.

To my beloved parents

SUMMARY

This thesis is concerned with the study of the properties and applications of bilinear stochastic processes X_t, $t \in Z$ satisfying

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + \sum_{\substack{i=1 \\ i \geq j}}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-i} e_{t-j} + e_t$$

for some sequence e_t, $t \in Z$ of independent identically distributed real random variables with common mean 0 and variance $\sigma^2 < \infty$ and constants $a_1, a_2, \ldots, a_r$, $b_1, b_2, \ldots, b_h$, β_{ij}, $i \leqslant l \leqslant m$, $l \leqslant j \leqslant \ell$, $i \geqslant j$.

The basic properties of stationary time series are outlined in Chapter 1. Some properties of linear time series models, such as autoregressive, moving average and mixed autoregressive-moving average models are also given. Bilinear time series models are introduced and the Subba Rao - Gabr test of linearity of stationary time series is reviewed.

Chapter 2 presents existence theorems for bilinear models. Sufficient conditions for the existence of a stationary process X_t, $t \in Z$ satisfying the bilinear model above are obtained. Ergodicity of the process X_t, $t \in Z$ defined above is also discussed. Stationarity conditions for linear models are derived from those of bilinear models.

Chapter 3 gives a method of obtaining expressions for the mean, variance and covariances of bilinear models. The mean, variance and covariances of linear models are derived from those of bilinear models. It is demonstrated that bilinear models are not necessarily distinguishable from linear models as far as covariance properties are concerned.

A class of bilinear processes that appear to be white noise under second-order analysis are analysed in some detail in Chapter 4. Two methods that use higher order moments in discriminating between a true white noise and a bilinear process with the same covariances are

presented. Also considered is the classical invertibility problem for bilinear processes.

The estimation of the parameters of bilinear models is considered in Chapter 5. A method of order detemination based on covariances and the information criterion of Akaike is given. A rule for forming forecasts for bilinear models is given. Bilinear models are fitted to three real time series and forecasts obtained from the bilinear models are compared with the forecasts obtained from linear models.

CONTENTS

INDEX OF TABLES

INDEX OF FIGURES

CHAPTER 1

PRELIMINARIES

1.1 INTRODUCTION

In this chapter, we compile some basic ideas needed for an understanding of the subsequent chapters. Section 1.2 discusses the statistical properties of a stationary time series. Autocovariance and autocorrelation functions are defined and their properties are given. Spectral density function and bispectral density function are given as two special cases of polyspectra.

Section 1.3 discusses the properties of linear time series models such as autoregressive, moving average, and mixed autoregressive-moving average models. The autocovariance function and the spectral density function of these models are given. Bilinear time series models are introduced in section 1.4. Some results on the simple bilinear process satisfying (1.4.2) are provided. Finally, a brief review of the test for linearity of stationary time series, as was developed by Subba Rao and Gabr [38], is given in section 1.5.

1.2 BASIC CONCEPTS

1.2.1 Stocastic (Random) Processes And Time Series Analysis

A stochastic (or random) process is defined as a family of random variables X_t where t is a parameter running over a suitable index set T and is denoted by X_t, $t \varepsilon T$. If $T = \{... , -1, 0, 1, ...\} = Z$, then X_t, $t \varepsilon T$ is said to be a discrete parameter process. If $T = \{-\infty, \infty\}$, then X_t, $t \varepsilon T$ is called a continuous parameter process. The state space, S of X_t, $t \varepsilon T$ is the space in which the possible values of X_t, $t \varepsilon T$ lie. If $S = (-\infty, \infty)$, then we call X_t, $t \varepsilon T$ a real-valued stochastic process. We will consider only the real-valued discrete parameter process X_t, $t \varepsilon Z$ in what follows.

The term 'time series' is used in statistical literature to mean a collection of observations of a random process made sequentially in time. When considering time series, there is taken to be an underlying real-valued discrete-time stochastic process X_t, $t \in Z$ and the available data x_t, $t = 1, 2, \ldots, N$ is a sample segment from all of the possible series that X_t, $t \in Z$ could have produced. Time series analysis refers to that body of principles and techniques which deal with analysis of the observed data x_t, $t = 1, 2, \ldots, N$. Usually, the data are analysed to try to find a model that approximates the true underlying generating random process X_t, $t \in Z$.

1.2.2 Stationary Processes

Intuitively speaking, a random process X_t, $t \in Z$ is said to be stationary if the statistical properties of the process do not change over time. There are two notions of stationarity, namely complete (strict) stationarity and stationarity up to order m, $m > 0$, $m \in Z$.

A completely (strictly) stationary process is a stochastic process X_t, $t \in Z$ with the property that for any positive integer n and any points $t_1, t_2, \ldots, t_n$ and $h \in Z$, the joint probability distribution of $\{X_{t_1}, X_{t_2}, \ldots, X_{t_n}\}$ is the same as the joint probability distribution of

$$\{X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_n+h}\}.$$

A stochastic process X_t, $t \in Z$ is said to be stationary up to order m, if for any positive integer n and any points $t_1, t_2, \ldots, t_n$ and $h \in Z$, the joint moments up to order m of $\{X_{t_1}, X_{t_2}, \ldots, X_{t_n}\}$ is the same as the joint moments up to order m of $\{X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_n+h}\}$. That is,

$$E\{(X_{t_1})^{k_1}(X_{t_2})^{k_2}\dots(X_{t_n})^{k_n}\}$$

$$= E\{(X_{t_1+h})^{k_1}(X_{t_2+h})^{k_2}\dots(X_{t_n+h})^{k_n}\} \quad (1.2.1)$$

for all possible non-negative integers $k_1, k_2, \dots, k_n$ satisfying $k_1 + k_2 + \dots + k_n \leqslant m$.

If we are given that X_t, $t \in Z$ is stationary up to order 2 then we have,

$$\left.\begin{aligned}
&(a)\quad E(X_t) = \mu \text{, independent of } t\\
&(b)\quad Var(X_t) = E\{(X_t - \mu)^2\}\\
&\qquad\qquad = \sigma_x^2 \text{, independent of } t\\
&(c)\quad Cov(X_t\, X_{t+k}) = E\{(X_t - \mu)(X_{t+k} - \mu)\}\\
&\qquad\qquad = R(k),\ k = 0, \pm1, \pm2, \dots\\
&\qquad\qquad = \text{a function of } k \text{ only}
\end{aligned}\right\} \quad (1.2.2)$$

1.2.3 Autocovariance And Autocorrelation Functions

The function $R(k)$, $k \in Z$ defined in (1.2.2) is known as the autocovariance function of lag k, and

$$\rho_k = R(k)/R(0) \quad (1.2.3)$$

for $k = 0, \pm1, \pm2, \dots$ is known as the autocorrelation function of lag k. The autocovariance and autocorrelation functions possess the following properties

$$\left.\begin{aligned}
&(i)\quad R(0) = Var(X_t) = \sigma_x^2 \Longleftrightarrow \rho_0 = 1\\
&(ii)\quad |R(k)| \leqslant R(0) \Longleftrightarrow |\rho_k| \leqslant 1,\ k \in Z\\
&(iii)\quad R(-k) = R(k) \Longleftrightarrow \rho_{-k} = \rho_k,\ k \in Z
\end{aligned}\right\} \quad (1.2.4)$$

(iv) $R(k)$ and ρ_k are both positive semi-definite in the sense that for any set of time points $t_1, t_2, \dots t_n \in Z$, and all real numbers

$a_1, a_2, \ldots a_n,$

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j R(t_i - t_j) \geq 0 \qquad (1.2.5)$$

and

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \rho_{t_i - t_j} \geq 0 . \qquad (1.2.6)$$

1.2.4 Polyspectra

Let X_t, $t \in Z$ be a random process stationary up to order k. Let

$$M_k(s_1, s_2, \ldots , s_{k-1}) = E\{X_t X_{t+s_1} \ldots X_{t+s_{k-1}}\} \qquad (1.2.7)$$

be the product moment of order k of the random variables $X_t, X_{t+s_1}, \ldots$ $\ldots, X_{t+s_{k-1}}$. Then $M_k(s_1, s_2, \ldots , s_{k-1})$ is the coefficient of $(\theta_1 \theta_2 \ldots \theta_k)$ in the expansion of the moment generating function

$$M(\theta_1, \theta_2, \ldots , \theta_k) = E\{ \exp(\theta_1 X_t + \theta_2 X_{t+s_1} + \ldots + \theta_k X_{t+s_{k-1}})\}. \qquad (1.2.8)$$

Let $C_k(s_1, s_2, \ldots , s_{k-1})$ be the joint cumulant of order k. Then $C_k(s_1, s_2, \ldots , s_{k-1})$ is the coefficient of $(\theta_1 \theta_2 \ldots \theta_k)$ in the expansion of the cumulant generating function

$$k(\theta_1, \theta_2, \ldots , \theta_k) = \log_e\{M(\theta_1, \theta_2, \ldots , \theta_k)\} \qquad (1.2.9)$$

By the stationarity condition, $M_k(s_1, s_2, \ldots , s_{k-1})$ and $C_k(s_1, s_2, \ldots$ $\ldots , s_{k-1})$ do not depend on t.

The Fourier transforms of the k-th order cumulants are called 'polyspectra'.

Definition 1.2.1. The k-th order polyspectrum (or k-th order cumulant spectrum) is defined by

$$f(\omega_1, \omega_2, \ldots , \omega_{k-1})$$
$$= \left(\frac{1}{2\pi}\right)^{k-1} \sum_{s_1=-\infty}^{\infty} \ldots \sum_{s_{k-1}=-\infty}^{\infty} C_k(s_1, s_2, \ldots , s_{k-1})$$

$$\exp\{-i(\omega_1 s_1 + \omega_2 s_2 + \dots + \omega_{k-1} s_{k-1})\} \quad (1.2.10)$$

$$(-\pi \leqslant \omega_i \leqslant \pi,\ i = 1, 2, \dots, k-1)$$

A sufficient condition for the existence of (1.2.10) is that the cumulants $C_k(s_1, s_2, \dots, s_{k-1})$ are absolutely summable, ie

$$\sum_{s_1=-\infty}^{\infty} \dots \sum_{s_{k-1}=-\infty}^{\infty} |C_k(s_1, s_2, \dots, s_{k-1})| < \infty \quad (1.2.11)$$

Polyspectra were introduced by Shiryaev [33]. Brillinger [6] and Brillinger and Rosenblatt [7] have given a comprehensive treatment of the theorectical properties of polyspectra and have discussed also the estimation of polyspectra from sample data. An important property of polyspectra is that all polyspectra of higher order than the second vanish when X_t, $t \in Z$ is a Gaussian process. See Priestley [30, p.872].

The Second Order Polyspectrum

Since

$$C_2(s_1) = \mathrm{Cov}(X_t\, X_{t+s_1})$$
$$= R(s_1), \quad (1.2.12)$$

it follows that the second order polyspectrum is given by

$$f(\omega_1) = f(\omega)$$

$$= \frac{1}{2\pi} \sum_{s=-\infty}^{\infty} R(s) e^{-i\omega s}, \quad -\pi \leqslant \omega \leqslant \pi$$

$$= \frac{1}{2\pi}\left\{R(0) + 2 \sum_{s=1}^{\infty} R(s) \cos \omega s\right\} \quad (1.2.13)$$

The function of $f(\omega)$, $-\pi \leqslant \omega \leqslant \pi$ is called the (power) spectral density function. The spectral density function always has the following properties

(i) $$R(0) = \sigma_x^2 = \int_{-\pi}^{\pi} f(\omega) d\omega \quad (1.2.14)$$

(ii) $f(\omega) \geqslant 0$, $-\pi \leqslant \omega \leqslant \pi$

(iii) for real valued processes

$$f(\omega) = f(-\omega) \text{ , } -\pi \leqslant \omega \leqslant \pi \text{ .}$$

The spectral density function is comprehensively treated in standard books like Priestley [30].

The Third Order Polyspectrum

The third-order cumulant $C_3(s_1, s_2)$ is identical with the third-order moment about the mean, ie

$$C_3(s_1, s_2) = C(s_1, s_2)$$

$$= E\{(X_t - \mu)(X_{t+s_1} - \mu)(X_{t+s_2} - \mu)\}$$

where $\mu = E(X_t)$. The third order polyspectrum may be written as

$$f(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \sum_{s_1=-\infty}^{\infty} \sum_{s_2=-\infty}^{\infty} C(s_1, s_2) e^{-i(\omega_1 s_1 + \omega_2 s_2)}$$

$$(-\pi \leqslant \omega_1, \omega_2 \leqslant \pi) \tag{1.2.15}$$

The function $f(\omega_1, \omega_2)$ is called the bispectral density function (or simply 'Bispectrum'). For a real valued process, the third-order central moments satisfy the following symmetric relations

$$C(s_1, s_2) = C(s_2, s_1) = C(-s_1, s_2 - s_1) = C(s_1 - s_2, -s_2) \tag{1.2.16}$$

while the following symmetric relations hold for $f(\omega_1, \omega_2)$

$$f(\omega_1, \omega_2) = f(\omega_2, \omega_1) = f(\omega_1, -\omega_1 - \omega_2) = f(-\omega_1 - \omega_2, \omega_2)$$

$$= \bar{f}(-\omega_1, -\omega_2) \tag{1.2.17}$$

where $\bar{f}(-\omega_1, -\omega_2)$ is the complex conjugate of $f(-\omega_1, -\omega_2)$.

Bispectra are discussed by Tukey [42] and Akaike [1]. Applications of bispectral analysis are described by Hasselman, Munk and MacDonald [23] and Lii, Rosenblatt and Van Atta [27] amongst others. Some progress has been made in the use of Bispectral techniques to describe non-linear processes. See Brillinger [6], Godfrey [13] and

Subba Rao and Gabr [38].

1.3 LINEAR TIME SERIES MODELS

The theory of linear time series models such as autoregressive, moving average, or mixed autoregressive-moving average models are well developed and excellent accounts of this theory can be found in Box and Jenkins [5] and Chatfield [9] amongst many other standard books. We give below a brief summary.

1.3.1 Purely Random Processes: "White Noise"

The process X_t, $t \in Z$ is called a purely random process if it consists of a sequence of uncorrelated random variables. For such a process to be stationary up to order 2, we require only that

$E(X_t) = \mu$, independent of t

$Var(X_t) = \sigma_x^2$, independent of t

$Cov(X_t, X_{t+s}) = 0$, for all $s \neq 0$

The spectral density funciton of the stationary purely random process is given by

$$f(\omega) = \sigma_x^2/2\pi \ , \ -\pi \leqslant \omega \leqslant \pi \qquad (1.3.1)$$

The purely random process is often called "white noise", particularly in the engineering literature. From now on, we denote a 'purely random process' by e_t, $t \in Z$. The e_t, $t \in Z$ are usually assumed to be normally distributed with mean zero and variance $\sigma^2 < \infty$.

1.3.2 The General Linear Process

A stochastic process X_t, $t \in Z$ is said to be a general linear process if it can be expressed in the form

$$X_t = \sum_{u=0}^{\infty} g_u e_{t-u} \qquad \text{a.e } [P] \quad (1.3.2)$$

for every t in Z where e_t, $t \in Z$ is a purely random process with mean zero and variance $\sigma^2 < \infty$, and g_u, $u \geq 0$ is a given sequence of constants satisfying

$$\sum_{u=0}^{\infty} g_u^2 < \infty$$

The series (1.3.2) converges in the quadratic mean.

If we let $g_u = 0$, $u < 0$, then we may re-write (1.3.2) as

$$X_t = \sum_{u=-\infty}^{\infty} g_u e_{t-u}$$

$$= G(B)e_t, \qquad (1.3.3)$$

where

$$G(B) = \sum_{u=-\infty}^{\infty} g_u B^u, \qquad (1.3.4)$$

and

$$B^u e_t = e_{t-u}.$$

If we assume that the inverse

$$G^{-1}(B) = \pi(B) \qquad (1.3.5)$$

exists, we can write (1.3.3) in the alternative form

$$\pi(B)X_t = e_t \qquad (1.3.6)$$

If $G(B)$, for $|B| \leq 1$, converges for all complex numbers B on or within the unit circle, we say that the model (1.3.2) is stationary. We shall say that the series (1.3.2) is 'invertible' if $\pi(B)$, for $|B| \leq 1$ converges.

The autocovariance function of X_t, $t \in Z$ satisfying (1.3.2) is

$$R(s) = E(X_t X_{t+s})$$

$$= \sigma^2 \sum_{u=-\infty}^{\infty} g_u g_{u+s} \qquad (1.3.7)$$

The spectral density funciton of X_t, $t \in Z$ is given by

$$f(\omega) = \frac{\sigma^2}{2\pi} G(e^{-i\omega})G(e^{i\omega})$$

$$= \frac{\sigma^2}{2\pi} |H(\omega)|^2 , \ -\pi \leqslant \omega \leqslant \pi \qquad (1.3.8)$$

where

$$H(\omega) = G(e^{-i\omega})$$

$$= \sum_{u=-\infty}^{\infty} g_u e^{-i\omega u} \qquad (1.3.9)$$

is known as the 'transfer function'. Also, the third-order central moments of X_t, $t \in Z$ satisfying (1.3.2) is given by

$$C(s_1, s_2) = E(X_t X_{t+s_1} X_{t+s_2})$$

$$= \mu_3 \sum_{u=-\infty}^{\infty} g_u \, g_{u+s_1} \, g_{u+s_2} \qquad (1.3.10)$$

where e_t, $t \in Z$ are independent and

$\mu_3 = E(e_t^3)$.

The bispectral density function is given by

$$f(\omega_1, \omega_2) = \frac{\mu_3}{(2\pi)^3} H(\omega_1)H(\omega_2)H(-\omega_1 - \omega_2) \qquad (1.3.11)$$

$$(-\pi \leqslant \omega_1, \omega_2 \leqslant \pi)$$

1.3.3 Autoregressive Process

A stochastic process X_t, $t \in Z$ is said to be an autoregressive process of order r, denoted by AR(r), if it satisfies the difference equation

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + e_t \qquad \text{a.e } [P] \ (1.3.12)$$

for every t in Z where $a_1, a_2, \ldots , a_r$ are constants and e_t, $t \in Z$ is a purely random process with mean zero and variance $\sigma^2 < \infty$. Equation (1.3.12) may be written in the form

$$\alpha(B)X_t = e_t \tag{1.3.13}$$

where

$$\alpha(B) = 1 - a_1 B - a_2 B^2 - \dots - a_r B^r \tag{1.3.14}$$

For stationarity we require that all the roots of $\alpha(B)$ must lie outside the unit circle.

The autocovariance function of X_t, $t \varepsilon Z$ satisfying (1.3.12) is

$$\left.\begin{aligned} R(s) &= a_1 R(1) + a_2 R(2) + \dots + a_r R(r) + \sigma^2 , \quad s = 0 \\ &= a_1 R(s-1) + a_2 R(s-2) + \dots + a_r R(s-r) , \\ & \qquad s = \pm 1, \pm 2 \dots \end{aligned}\right\} \tag{1.3.15}$$

The second set of equations in (1.3.15) is called the Yule-Walker equations for an AR(r) process. The spectral density funciton of the stationary AR(r) process is given by

$$f(\omega) = \frac{\sigma^2}{2\pi}\left| \alpha^{-1}(e^{-i\omega}) \right|^2 , \ -\pi \leqslant \omega \leqslant \pi \tag{1.3.16}$$

1.3.4 Moving Average Process

A stochastic process X_t, $t \varepsilon Z$ is said to be a moving average process of order h, denoted by MA(h), if it satisfies the difference equation

$$X_t = \sum_{j=1}^{h} b_j e_{t-j} + e_t \qquad \text{a.e } [P] \tag{1.3.17}$$

for every t in Z where $b_1, b_2, \dots , b_h$ are constants, and e_t, $t \varepsilon Z$ is a purely random process with mean zero and variance $\sigma^2 < \infty$.
Equation (1.3.17) may be written in the form

$$X_t = \beta(B)e_t \tag{1.3.18}$$

where

$$\beta(B) = 1 + b_1 B + b_2 B^2 + \dots + b_h B^h \tag{1.3.19}$$

For invertibility of the MA(h) process, we require that all the roots of $\beta(B)$ must lie outside the unit circle.

The autocovariance function of X_t, $t \varepsilon Z$ satisfying (1.3.17) is

$$R(s) = \begin{cases} \sigma^2 \sum_{j=0}^{h} b_j & , s = 0 \\ \sigma^2 \sum_{j=0}^{h-s} b_j b_{j+s} & , s = \pm 1, \pm 2, \ldots, \pm h \\ 0 & , |s| > h \end{cases} \qquad (1.3.20)$$

where $b_0 = 1$. The spectral density function of the invertible MA(h) process is given by

$$f(\omega) = \frac{\sigma^2}{2\pi} \left| \beta(e^{-i\omega}) \right|^2 , \quad -\pi \leqslant \omega \leqslant \pi \qquad (1.3.21)$$

1.3.5 Mixed Autoregressive-Moving Average Process

A stochastic process X_t, $t \varepsilon Z$ is said to be a mixed autoregressive-moving average process of order (r, h), denoted by ARMA(r, h), if it satisfies an equation of the type

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + e_t \qquad \text{a.e } [P] \quad (1.3.22)$$

for every t in Z where $a_1, a_2, \ldots, a_r, \; b_1, b_2, \ldots, b_h$ are constants and e_t, $t \varepsilon Z$ is a purely random process with common mean zero and variance $\sigma^2 < \infty$. Equivalently, we may write (1.3.22) as

$$\alpha(B)X_t = \beta(B)e_t \qquad (1.3.23)$$

where $\alpha(B)$ and $\beta(B)$ are given by (1.3.14) and (1.3.19) respectively. The process X_t, $t \varepsilon Z$ satisfying (1.3.22) is stationary if all the roots of $\alpha(B)$ lie outside the unit circle and invertible if all the roots of $\beta(B)$ lie outside the unit circle.

The autocovariance function of X_t, $t \varepsilon Z$ satisfying (1.3.22) will be like that of an autoregressive series after lag h. That is,

$$R(s) = a_1 R(s-1) + a_2 R(s-2) + \ldots + a_r R(s-r) \; , \; s \geqslant h \quad (1.3.24)$$

The first h autocovariances depend on the moving average parameters $b_1, b_2, \ldots, b_h$ as well as on the autoregressive parameters $a_1, a_2, \ldots, a_r$. Equation (1.3.24) is called the Yule-Walker equations for an ARMA(r, h). The spectral density funciton of the ARMA(r, h) process is given by

$$f(\omega) = \frac{\sigma^2}{2\pi} \frac{|\beta(e^{-i\omega})|^2}{|\alpha\ (e^{-i\omega})|^2}, -\pi \leq \omega \leq \pi \qquad (1.3.25)$$

1.4 BILINEAR TIME SERIES MODELS

Linear time series models are widely used in many fields because these models can be analysed with considerable ease and they provide fairly good approximations for the true underlying generating random process. However, the underlying structure of the series may not be linear and what is more, the series may not be Gaussian. In these situations, second-order properties, such as covariances and spectra, can no longer adequately characterize the properties of the series and one is led then to consider non-linear nodels which can provide a better fit.

A particular class of non-linear models which has been found to be useful in many fields is the bilinear models. Bilinear models have been extensively discussed in the control theory literature. One could check Ruberti, Isidori and d'Alessandro [32] and Bruni, Dupillo and Koch [8] for further details. Until recently the theory of bilinear models dealt with the structural theory of deterministic bilinear differential equations. The study of bilinear models as stochastic models was initiated by Subba Rao [34, 35, 36, 37] and Granger and Andersen [14, 15].

Let e_t, $t \in Z$ be a sequence of independent identically distributed random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let $a_1, a_2, \ldots, a_r$,

$b_1, b_2, \ldots, b_h$ and β_{ij}, $1 \leq i \leq m$, $1 \leq j \leq \ell$ be real constants. The general form of the bilinear model, as defined in [15] is

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-i} e_{t-j} + e_t \quad \text{a.e } [P] \quad (1.4.1)$$

for every t in Z. If X_t, $t \in Z$ satisfies (1.4.1), Granger and Andersen [15] uses the notation that X_t, $t \in Z$ is BARMA(r, h, m, ℓ,), where BARMA is the abbreviation for Bilinear Autoregressive Moving Average Model.

Various simple forms of (1.4.1) are discussed in the literature by the following authors: Granger and Andersen [14, 15], Subba Rao [34, 35 , 36, 37], Tuan Dinh Pham and Lanh Tat Tran [41], Subba Rao and Gabr [39], Bhaskara Rao, Subba Rao and Walker [4], Tong [40], Hannan [22], Quinn [31] and Guegan [18]. The simple bilinear process X_t, $t \in Z$ satisfying

$$X_t = \beta X_{t-1} e_{t-1} + e_t \quad \text{a.e } [P] \quad (1.4.2)$$

for every t in Z for some constant β, where e_t, $t \in Z$ are independent and each e_t is distributed as $N(0, \sigma^2)$, is extensively studied in [15] where it has been shown that the autocorrelation sequence for X_t, $t \in Z$ is

$$\rho_k = \begin{cases} 1 & k = 0 \\ \dfrac{\lambda^2(1 - \lambda^2)}{1 + \lambda^2 + \lambda^4} & k = \pm 1 \\ 0 & \text{elsewhere} \end{cases} \quad (1.4.3)$$

where $\lambda = \sigma\beta$. We have found the expressions given therein for the third and fourth central moments to be incorrect. Under the normality assumption for e_t series, the third and fourth central moments can be shown to be

$$E\left[(X_t - E(X_t))^3\right] = \frac{2\sigma^3\lambda^3}{1 - \lambda^2}\left\{4 + 5\lambda^2\right\}, \lambda^2 < 1 \quad (1.4.4)$$

$b_1, b_2, \ldots, b_h$ and β_{ij}, $1 \leq i \leq m$, $1 \leq j \leq \ell$ be real constants. The general form of the bilinear model, as defined in [15] is

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-i} e_{t-j} + e_t \qquad \text{a.e } [P] \quad (1.4.1)$$

for every t in Z. If X_t, $t \varepsilon Z$ satisfies (1.4.1), Granger and Andersen [15] uses the notation that X_t, $t \varepsilon Z$ is BARMA(r, h, m, ℓ,), where BARMA is the abbreviation for Bilinear Autoregressive Moving Average Model.

Various simple forms of (1.4.1) are discussed in the literature by the following authors: Granger and Andersen [14, 15], Subba Rao [34, 35 , 36, 37], Tuan Dinh Pham and Lanh Tat Tran [41], Subba Rao and Gabr [39], Bhaskara Rao, Subba Rao and Walker [4], Tong [40], Hannan [22], Quinn [31] and Guegan [18]. The simple bilinear process X_t, $t \varepsilon Z$ satisfying

$$X_t = \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P] \quad (1.4.2)$$

for every t in Z for some constant β, where e_t, $t \varepsilon Z$ are independent and each e_t is distributed as $N(0, \sigma^2)$, is extensively studied in [15] where it has been shown that the autocorrelation sequence for X_t, $t \varepsilon Z$ is

$$\rho_k = \begin{cases} 1 & k = 0 \\ \dfrac{\lambda^2(1 - \lambda^2)}{1 + \lambda^2 + \lambda^4} & k = \pm 1 \\ 0 & \text{elsewhere} \end{cases} \qquad (1.4.3)$$

where $\lambda = \sigma\beta$. We have found the expressions given therein for the third and fourth central moments to be incorrect. Under the normality assumption for e_t series, the third and fourth central moments can be shown to be

$$E\left[(X_t - E(X_t))^3\right] = \frac{2\sigma^3\lambda^3}{1 - \lambda^2}\left\{4 + 5\lambda^2\right\}, \lambda^2 < 1 \qquad (1.4.4)$$

and

$$E\left[(X_t - E(X_t))^4\right]$$

$$= \frac{3\sigma^4}{(1-\lambda^2)(1-3\lambda^4)}\left\{1 + 3\lambda^2 + 19\lambda^4 + 31\lambda^6 + 39\lambda^8 + 45\lambda^{10}\right\}, \quad (1.4.5)$$

provided that $3\lambda^4 < 1$. This correction affects column 5 of Table 1 in [15].

An interesting generalisation of (1.4.2) is the process X_t, $t \in Z$ satisfying

$$X_t = a\,X_{t-1} + \beta\,X_{t-1}e_{t-1} + e_t \qquad \text{a.e } [P] \quad (1.4.6)$$

for every t in Z and for some constants a and β where e_t, $t \in Z$ are independent and each e_t is distributed as $N(0, \sigma^2)$. The existence of a stationary and invertible process X_t, $t \in Z$ satisfying (1.4.6) is discussed in [15], [37] and [41]. The existence problem for a process X_t, $t \in Z$ satisfying special cases of the model

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{\substack{i=1\\ i\geq j}}^{m}\sum_{j=1}^{\ell} \beta_{ij} X_{t-i}e_{t-j} + e_t \qquad \text{a.e } [P] \quad (1.4.7)$$

for every t in Z are discussed in [4], [37], [39] and [46].

So far all these studies have failed to include the moving average part $\sum_{j=1}^{h} b_j e_{t-j}$ in (1.4.1). This study, outlined in chapters 2, 3, 4 and 5, is devoted to the study of the full bilinear model (1.4.1). For mathematical convenience, we will confine ourselves to the study of the process X_t, $t \in Z$ satisfying

$$X_t = e_t + \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + \sum_{\substack{i=1\\ i\geq j}}^{m}\sum_{j=1}^{\ell} \beta_{ij} X_{t-i}e_{t-j} \qquad \text{a.e } [P] \quad (1.4.8)$$

for every t in Z.

1.5 TEST FOR LINEARITY OF STATIONARY TIME SERIES

Subba Rao and Gabr [38] have constructed two tests aimed at detecting whether a stationary time series is Gaussian and if the process is non-Gaussian, whether it conforms to a linear model. Hinich [24] has given a modification of their approach that makes use of the large sample properties of the sample bispectrum. We give below a brief summary of the Subba Rao - Gabr approach.

We have indicated in section 1.2.4 that if a process is Gausssian then all its polyspectra of higher order than the second are identically zero. In particular, relation (1.3.11) shows that if $\mu_3 = 0$, then the bispectral density function of the general linear model is identically zero for all frequencies. Of course, if the random variables e_t, $t \in Z$ are Gaussian, then $\mu_3 = 0$ and $f(\omega_1, \omega_2) = 0$, $-\pi \leqslant \omega_1, \omega_2 \leqslant \pi$. Thus, if the process X_t, $t \in Z$ is Gaussian (Gaussianity of e_t, $t \in Z$ imples Gaussianity of X_t, $t \in Z$ for X_t satisfying (1.3.2)), then the bispectral density function $f(\omega_1, \omega_2) = 0$, $-\pi \leqslant \omega_1, \omega_2 \leqslant \pi$.

The test for Gaussianity is carried out by examining the null-hypothesis that the bispectrum is zero at all frequencies. The test statistic has a form similar to Hotelling's T^2 statistic and is constructed from the values of the estimated bispectrum over a grid of frequencies. Using (1.3.8) and (1.3.11) we obtain

$$X_{ij} = \frac{|f(\omega_i, \omega_j)|^2}{f(\omega_i)f(\omega_j)f(\omega_i + \omega_j)}, \quad -\pi \leqslant \omega_i, \omega_j \leqslant \pi$$

$$= \frac{\mu_3^2}{2\pi\sigma^6}$$

$$= \text{constant.}$$

The test for linearity is based on the constancy of the sample values of X_{ij} over a grid of frequencies, and the test statistic again has a form similar to Hotelling's T^2.

The fact that a process X_t, $t \varepsilon Z$ has a zero bispectrum does not necessarily mean that X_t, $t \varepsilon Z$ is a Guassian process for it is possible for a non-Gaussian process to have a zero bispectrum. See Priestley [30, p.877]. Also, the constancy of X_{ij} does not necessarily imply that the process X_t, $t \varepsilon Z$ follows a linear model. As pointed out in [30], it is reasonable to suppose that in most practical situations deviations from Gaussianity or linearity would show up in the form of the bispectrum.

CHAPTER 2

EXISTENCE THEOREMS FOR BILINEAR MODELS

2.1 INTRODUCTION

2.1.1 General Form of Bilinear Models - Vectorial Representation

Let X_t, $t = \dots, -1, 0, 1, \dots$ and e_t, $t = \dots -1, 0, 1, \dots$, be two real stochastic processes defined on some probability space $(\Omega, \mathcal{F}, P)$. Let Z denote the set of all integers. We call e_t, $t \in Z$ the input or unobservable process and X_t, $t \in Z$ the output or observable process. X_t, $t \in Z$ is said to be a Bilinear Model with respect to the input process e_t, $t \in Z$ if

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-i} e_{t-j} + e_t \quad \text{a.e } [P] \quad (2.1.1)$$

for every t in Z, for some constants $a_1, a_2, \dots a_r$, $b_1, b_2, \dots, b_h$ and β_{ij}, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant \ell$.

We usually assume the input process e_t, $t \in Z$ to be independent identically distributed with common mean 0 and variance $\sigma^2 < \infty$.

The first part of (2.1.1) is identifiable as the autoregressive part of the process X_t, $t \in Z$, the second as the moving average part of X_t, $t \in Z$ and the third part is the 'pure' bilinear part of X_t, $t \in Z$. A study of bilinear models subsumes the study of autoregressive models as well as the study of moving average models and mixed autoregressive-moving average models.

One cannot fail to notice that (2.1.1) can also be labelled as the bilinear model for e_t, $t \in Z$ with respect to X_t, $t \in Z$. The labelling is a matter of semantics, and we will always put the observable process on the left of (2.1.1).

For mathematical convenience, we wish to consider the process X_t, $t \in Z$ satisfying

$$X_t = e_t + \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + \sum_{\substack{i=1 \\ i \geq j}}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-i} e_{t-j} \qquad \text{a.e } [P] \quad (2.1.2)$$

for every t in Z.

If X_t, $t \in Z$ satisfies (2.1.2), we use the notation that X_t, $t \in Z$ is BARMA (r, h, m, ℓ), where BARMA is the abbreviation for Bilinear Autoregressive Moving Average Model. The phrase "with respect to e_t, $t \in Z$" in the description of the process X_t, $t \in Z$ is omitted without undue misunderstanding.

The purpose of this chapter is to examine under what conditions a process X_t, $t \in Z$ exists satisfying (2.1.2) for a given sequence e_t, $t \in Z$ of independent identically distributed random variables with common mean 0 and common variance $\sigma^2 < \infty$ and constants $a_1, a_2, \dots, a_r$, $b_1, b_2, \dots, b_h$ and β_{ij}, $1 \leq i \leq m$, $1 \leq j \leq \ell$ where $i \geq j$. This problem has been tackled by Bhaskara Rao, Subba Rao and Walker [4] for the special class of models satisfying

$$X_t = \sum_{j=1}^{p} a_j X_{t-j} + \sum_{j=1}^{q} \beta_j X_{t-j} e_{t-j} + e_t \qquad \text{a.e } [P] \quad (2.1.3)$$

for every t in Z.

After putting this model in vector form, they gave a sufficient condition for the existence of a strictly stationary process satisfying (2.1.3). Earlier, Subba Rao and Gabr [39] gave a sufficient condition for the existence of a second-order stationary process X_t, $t \in Z$, satisfying (2.1.3) with $p = q$. The sufficient conditions in both situations were the same. Subba Rao and Gabr [39] also obtained the same sufficient conditions for the existence of a second-order stationary process X_t, $t \in Z$ satisfying

$$X_t = e_t + \sum_{j=1}^{p} a_j X_{t-j} + \sum_{\substack{i=1 \\ i \geq j}}^{p} \sum_{j=1}^{p} \beta_{ij} X_{t-i} e_{t-j} \qquad \text{a.e } [P] \quad (2.1.4)$$

for every t in Z.

Adapting the method given in Bhaskara Rao, Subba Rao and Walker [4], we give a sufficient condition under which a strictly stationary process X_t, $t \in Z$ exists satisfying (2.1.2). Before that, we would like to put (2.1.2) in vector form.

Theorem 2.1.1. Suppose a process X_t, $t \in Z$ satisfies (2.1.2). Let

$$p = \max\{r, m\} \quad (2.1.5)$$

$$g = \min\{m, \ell\} \quad (2.1.6)$$

$$q = \max\{h, g\} \quad (2.1.7)$$

$$\underset{p\times p}{A} = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_r & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \underbrace{0 \quad 0 \quad \cdots \quad 1}_{(p-r)} & & & & 0 \end{bmatrix} \quad (2.1.8)$$

$$\underset{1\times p}{C^T} = (1, 0, 0, \ldots, 0) \quad (2.1.9)$$

$$\underset{1\times p}{\underline{b}_j^T} = (b_j, 0, 0, \ldots, 0), \; j = 1, 2, \ldots, h \quad (2.1.10)$$

$$= \underline{0}, \text{ for all } j > h \text{ when } h < g$$

where $\underline{0}$ is the null vector in which every entry is zero.

$$\underset{p\times p}{B_j} = \begin{bmatrix} \beta_{jj} & \beta_{j+1,j} & \cdots & \beta_{mj} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \underbrace{0 & 0 & \cdots & 0}_{(p-m+j-1)} \end{bmatrix} \qquad (2.1.11)$$

$(j = 1, 2, \ldots, g)$

$= \underline{0}$, for all $j > g$ when $h > g$

where $\underline{0}$ is the null matrix in which every entry is zero.

$$\underset{1\times p}{\underline{X}_t^T} = (X_t, X_{t-1}, \ldots, X_{t-p+1}), \quad t \in Z \qquad (2.1.12)$$

Then

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j\, e_{t-j} + \sum_{j=1}^{q} B_j\, \underline{X}_{t-j}\, e_{t-j} + Ce_t \quad \text{a.e } [P] \quad (2.1.13)$$

for every t in Z.

Proof By direct verification.

Now, we replace the existence problem as follows. We are given a sequence e_t, $t \in Z$ of independent identically distributed real random variables with common mean 0 and common variance $\sigma^2 < \infty$ and matrices $\underset{p\times p}{A}$, $\underset{p\times p}{B_j}$, $j = 1, 2, \ldots, q$, $\underset{p\times 1}{\underline{b}_j}$, $j = 1, 2, \ldots, q$ and $\underset{p\times 1}{C}$. Under what condition does there exist a vector-valued process $\underline{X}_t$, $t \in Z$ each with p-components satisfying

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j\, e_{t-j} + \sum_{j=1}^{q} B_j\, \underline{X}_{t-j}\, e_{t-j} + Ce_t \quad \text{a.e } [P] \quad (2.1.14)$$

for every t in Z.

We will also discuss some special cases at the end of this chapter.

2.2 PRELIMINARIES

2.2.1 On Kronecker Product of Matrices

In this section, we record some results on kronecker product of matrices. These results are used in the proofs of some of the results in sections 2.3 and 2.4.

For any two matrices $A = (a_{ij})$ and $B = (b_{ij})$ of orders $m \times n$ and $r \times s$ respectively, we denote the kronecker product of A and B by $A \otimes B$ and is defined to be the following matrix of order $mr \times ns$

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & & & \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix} \qquad (2.2.1)$$

We adopt the following notation. For any matrix D, we denote the (i,j)-th element of D by $(D)_{ij}$ or $(D)_{i,j}$, if the elements of D are not indicated specifically before. If C is a column vector, the i-th component of C is denoted by $(C)_i$ or $(C)_{i1}$ when the elements of C are not explicitly indicated before. I_n stands for the identity matrix of order $n \times n$ in which every diagonal entry is equal to unity and every non-diagonal entry is equal to zero. Also, $\underline{0}$ stands for the null matrix or vector as the case may be.

The element $a_{ij}b_{uv}$ which is the $((i-1)r+u, (j-1)s+v)$-th element of the matrix $A \otimes B$ for $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$; $u = 1, 2, \dots, r$; and $v = 1, 2, \dots, s$; is denoted by $(A \otimes B)_{ij;uv}$. In fact,

$$\begin{aligned}(A \otimes B)_{ij,uv} &= a_{ij}b_{uv} \\ &= (A \otimes B)_{(i-1)r+u,(j-1)s+v} \qquad (2.2.2)\end{aligned}$$

We give below some of the properties of kronecker products (for details see Neudecker [29]).

<u>Lemma 2.2.1</u>

(a) For any three matrices A, B and C,

$$(A \otimes B) \otimes C = A \otimes (B \otimes C) \qquad (2.2.3)$$

(b) For any four matrices A, B, C and D, where A and B are of the same order and C and D are of the same order,

$$(A + B) \otimes (C + D) = (A \otimes C) + (A \otimes D) + (B \otimes C) + (B \otimes D) \tag{2.2.4}$$

(c) For any four matrices A, B, C, and D,

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD), \tag{2.2.5}$$

provided the matrices involved are comformable for multiplication.

If $A_1, A_2, \ldots, A_n$, $B_1, B_2, \ldots, B_n$ are 2n matrices, then

$$(A_1 \otimes B_1)(A_2 \otimes B_2) \ldots (A_n \otimes B_n) = (A_1A_2 \ldots A_n) \otimes (B_1B_2 \ldots B_n), \tag{2.2.6}$$

provided the matrices involved are conformable for multiplication.

(d) Let A_{pxp} and B_{qxq} be two square matrices with eigen values $\alpha_1, \alpha_2, \ldots, \alpha_p$ and $\beta_1, \beta_2, \ldots, \beta_q$ respectively. Then the eigen values of $A \otimes B$ are the pq numbers $\alpha_i\beta_j$ where i = 1, 2, ... , p and j = 1, 2, ... , q. See Lancaster [26, p. 260].

2.2.2 Spectral Radius of a Matrix

Let $A = (a_{ij})$ be a square matrix of order n x n with eigen values $\lambda_1, \lambda_2, \ldots, \lambda_n$. We denote the maximum of $|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|$ by $\rho(A)$, and $\rho(A)$ is usually called the spectral radius of A. We give below some of the properties of $\rho(A)$.

Lemma 2.2.2

(a) $\rho(A) \leq ||A||$ for any norm $||\cdot||$ on the linear space of all square matrices of the same order.

(b) There exists a positive constant K such that for any positive integer m, we have

$$|(A^m)_{ij}| \leq K(\rho(A))^m \tag{2.2.7}$$

for all i and j.

(c)
$$\rho(A) \leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|$$

$$\leqslant \max_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} |a_{ij}|$$

$$\leqslant \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|$$

$$\leqslant \left(\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 \right)^{\frac{1}{2}}$$

$$\leqslant n \max_{1 \leqslant i,j \leqslant n} |a_{ij}| \qquad (2.2.8)$$

(d) Let $\underset{p\times p}{A}$ be a square matrix. Then $\rho(A \otimes A) < 1$ if and only if $\rho(A) < 1$.

Proof (a), (b) and (c) are standard fare treated in any good book on matrix algebra. One could check Kato [25, p.36] for (c). We prove (d) now.

Let $\alpha_1, \alpha_2, \ldots, \alpha_p$ be the eigen values of A. Then

$\rho(A) < 1$ if and only if $|\alpha_i| < 1$, for all $i = 1, 2, \ldots, p$,

if and only if $|\alpha_i \alpha_j| < 1$, for all $i,j = 1, 2, \ldots, p$,

if and only if $\rho(A \otimes A) < 1$,

by Lemma 2.2.1 (d).

2.2.3 Convergence of Sequences of Random Vectors

Let $\{\underline{Y}_n, n \geqslant 1\}$ be a sequence of random vectors each of the same order $p \times 1$ defined on some probability space $(\Omega, \mathcal{F}, P)$. We say that $\sum_{n\geqslant 1} \underline{Y}_n$ converges absolutely almost surely [P] if

$$\sum_{n\geqslant 1} |(\underline{Y}_n)_i| < \infty \qquad \text{a.e } [P] \quad (2.2.9)$$

for every $i = 1, 2, \ldots, p$.

We say that $\sum_{n\geqslant 1} \underline{Y}_n$ converges in the mean if there exists a random vector $\underline{Y}$ of p-components such that

$$\lim_{m\to\infty} E \left| \sum_{n=1}^{m} (\underline{Y}_n)_i - (\underline{Y})_i \right| = 0 \qquad (2.2.10)$$

for every $i = 1, 2, \ldots, p$.

We say that $\sum_{n \geq 1} \underline{Y}_n$ converges in the quadratic mean if there exists a random vector $\underline{Y}$ of p-components such that

$$\lim_{m \to \infty} E \left| \sum_{n=1}^{m} (\underline{Y}_n)_i - (\underline{Y})_i \right|^2 = 0 \tag{2.2.11}$$

for every $i = 1, 2, \ldots, p$.

2.3 EXISTENCE THEOREM: THE CASE $q = 1$

In this section, we give a simple sufficient condition for the existence of a strictly stationary vector-valued process $\underline{X}_t$, $t \in Z$ conforming to the bilinear model

$$\underline{X}_t = A\, \underline{X}_{t-1} + \underline{b}e_{t-1} + B\, \underline{X}_{t-1} e_{t-1} + Ce_t \qquad \text{a.e } [P] \tag{2.3.1}$$

for every t in Z for a given sequence e_t, $t \in Z$ of independent identically distributed random variables with common mean 0 and variance $\sigma^2 < \infty$ and matrices $\underset{p\times p}{A}$, $\underset{p\times p}{B}$, $\underset{p\times 1}{\underline{b}}$, $\underset{p\times 1}{C}$.

We treat the case $q = 1$ in (2.1.14) specially because we can describe the process $\underline{X}_t$, $t \in Z$, if it exists, explicitly as an infinite series in e_t's. The case $q > 1$ is more complicated. We will discuss this case in section 2.4.

We begin with a Lemma.

Lemma 2.3.1 Let f be a vector-valued measurable function defined on the infinite-dimensional Euclidian space $R \times R \times R \times \ldots$, ie, for every sequence $(x_1, x_2, \ldots, \ldots)$ of real numbers, $f(x_1, x_2, \ldots)$ is a vector in some p-dimensional Euclidian space R^p. Let e_t, $t \in Z$ be a real strictly stationary process. Then

$$f(e_t, e_{t-1}, e_{t-2}, \ldots), \quad t \in Z$$

is strictly stationary.

Proof Obvious.

The following is the main result of this section.

Theorem 2.3.2 Let e_t, $t \in Z$ be a sequence of independent identically distributed real random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ such that $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let A and B be two matrices of order pxp such that

$$\rho(A \otimes A + \sigma^2 B \otimes B) = \lambda < 1$$

Let $\underline{b}$ and C be two column vectors with components $b_1, b_2, \ldots, b_p$ and $C_1, C_2, \ldots, C_p$ respectively. Then, the series of random vectors

$$\sum_{r\geq 1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})$$

converges absolutely almost surely [P] as well as in the mean for every fixed t in Z. Further, if

$$\underline{X}_t = Ce_t + \underline{b}e_{t-1} + \sum_{r\geq 1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1}), \ t \in Z$$

then $\underline{X}_t$, $t \in Z$ is a strictly stationary process conforming to the bilinear model

$$\underline{X}_t = A\,\underline{X}_{t-1} + \underline{b}e_{t-1} + B\,\underline{X}_{t-1}e_{t-1} + Ce_t$$

for every t in Z.

Conversely, if $\underline{X}_t$, $t \in Z$ is a strictly stationary vector-valued process satisfying

$$\underline{X}_t = A\,\underline{X}_{t-1} + \underline{b}e_{t-1} + B\,\underline{X}_{t-1}\,e_{t-1} + Ce_t \qquad \text{a.e } [P] \quad (2.3.2)$$

for every t in Z for some sequence e_t, $t \in Z$ of independent identically distributed random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$ and for some matrices A, B, $\underline{b}$, and C of orders pxp, pxp, px1, px1 respectively satisfying

$$\rho(A \otimes A + \sigma^2 B \otimes B) = \lambda < 1,$$

then

$$\underline{X}_t = Ce_t + \underline{b}e_{t-1} + \sum_{r\geq 1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1}) \quad \text{a.e } [P]$$

for every t in Z.

Proof. The proof given below is an adaptation of the proof given in section 3 of Bhaskara Rao, Subba Rao and Walker [4] and is carried out in the following steps.

1^0. For almost sure convergence, we show that

$$\sum_{r\geq 1} E\left|\left(\prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})\right)_i\right| < \infty \tag{2.3.3}$$

for every i = 1, 2, ... , p. This would then imply that the series

$$\sum_{r\geq 1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})$$

is absolutely convergent almost surely [P] as well as in the mean. See Chung [10, (xi), p. 42].

2^0. We establish (2.3.3) for i = 1. The general case is clear. First, we note that for every t in Z, $r \geq 1$ and s = 1, 2, ... , p,

$$E|((A + Be_{t-r})(Ce_{t-r} + \underline{b}e_{t-r-1}))_s|$$

$$= E\left|\sum_{j=1}^{p} (A)_{sj}\, C_j\, e_{t-r} + \sum_{j=1}^{p}(A)_{sj}\, b_j\, e_{t-r-1} + \sum_{j=1}^{p} (B)_{sj}\, C_j\, e^2_{t-r} + \sum_{j=1}^{p} (B)_{sj}\, b_j\, e_{t-r}\, e_{t-r-1}\right|$$

$$\leq \sum_{j=1}^{p} |(A)_{sj}||C_j|E|e_{t-r}| + \sum_{j=1}^{p} |(A)_{sj}||b_j|E|e_{t-r-1}| + \sum_{j=1}^{p} |(B)_{sj}||C_j|E(e^2_{t-r}) + \sum_{j=1}^{p} |(B)_{sj}||b_j|E|e_{t-r}\, e_{t-r-1}|$$

$$\leq \left\{\sum_{j=1}^{p} |(A)_{sj}||C_j| + \sum_{j=1}^{p} |(A)_{sj}||b_j|\right\}\sigma + \left\{\sum_{j=1}^{p} |(B)_{sj}||C_j| + \sum_{j=1}^{p} |(B)_{sj}||b_j|\right\}\sigma^2$$

(by Cauchy-Schwartz in equality)

$$\leq K_0$$

where K_0 is a constant which depends only on A, B, $\underline{b}$, C, σ^2 and independent of r and t.

3^0. If $r \geq 2$, we show that

$$E\left|\left(\prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})\right)_1\right|$$

$$\leq K_1 \, p \, \lambda^{\frac{r-1}{2}} \tag{2.3.4}$$

for some constant $K_1 > 0$.

Observe that

$$E\left|\left(\prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})\right)_1\right|$$

$$= E\left|\left(\left(\prod_{j=1}^{r-1} (A + Be_{t-j})\right)\left((A + Be_{t-r})(Ce_{t-r} + \underline{b}e_{t-r-1})\right)\right)_1\right|$$

$$= E\left|\sum_{s=1}^{p} \left(\prod_{j=1}^{r-1} (A + Be_{t-j})\right)_{1s}((A + Be_{t-r})(Ce_{t-r} + \underline{b}e_{t-r-1}))_s\right|$$

$$\leq \sum_{s=1}^{p}\left(E\left|\left(\prod_{j=1}^{r-1} (A + Be_{t-j})\right)_{1s}\right|\right)(E|((A + Be_{t-r})(Ce_{t-r} + \underline{b}e_{t-r-1}))_s|)$$

(In the above derivation, we have used the fact that $\prod_{j=1}^{r-1} (A + Be_{t-j})$ and $(A + Be_{t-r})(Ce_{t-r} + \underline{b}e_{t-r-1})$ are independently distributed.)

$$\leq K_0 \sum_{s=1}^{p} \left\{E\left(\left(\prod_{j=1}^{r-1} (A + Be_{t-j})\right)_{1s}\right)^2\right\}^{\frac{1}{2}}$$

(By 2^0 and Cauchy-Schwartz inequality.)

Now, for any $s = 1, 2, \ldots, p$,

$$\left(\left(\prod_{j=1}^{r-1} (A + Be_{t-j})\right)_{1s}\right)^2$$

$$= \left(\left(\prod_{j=1}^{r-1} (A + Be_{t-j})\right) \otimes \left(\prod_{j=1}^{r-1} (A + Be_{t-j})\right)\right)_{1s;1s}$$

(By 2.2.2)

$$= \left(\prod_{j=1}^{r-1} (A + Be_{t-j}) \otimes (A + Be_{t-j})\right)_{1s;1s}$$

(By Lemma 2.2.1 (c).)

Consequently

$$E\left[\left(\prod_{j=1}^{r-1} (A + Be_{t-j})\right)_{1s}\right]^2$$

$$= \prod_{j=1}^{r-1} (E(A + Be_{t-j}) \otimes (A + Be_{t-j}))_{1s;1s}, \text{ because } e_t\text{'s are independent.}$$

$= ((E[(A + Be_t) \otimes (A + Be_t)])^{r-1})_{1s;1s}$, because e_t's are identically distributed.

$= ((E[A \otimes A + e_t\, A \otimes B + e_t\, B \otimes A + e_t^2\, B \otimes B])^{r-1})_{1s;1s}$, by Lemma 2.2.1 (b).

$= ((A \otimes A + \sigma^2 B \otimes B)^{r-1})_{1s;1s}$

$\leqslant K\, \lambda^{r-1}$, for some constant $K > 0$ by (2.2.7)

Hence,

$$E\left|\left(\prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-1} + \underline{b}e_{t-r-1})\right)_1\right| < K_1\, p\, \lambda^{\frac{r-1}{2}}$$

for a suitable choice of the constant $K_1 > 0$.

4^0. Since $\lambda < 1$, we have

$$\sum_{r\geqslant 1} E\left|\left(\prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})\right)_1\right| < \infty$$

Thus (2.3.3) is established.

5^0. By Lemma 2.3.1, the vector-valued process $\underline{X}_t$, $t \in Z$ defined by

$$\underline{X}_t = Ce_t + \underline{b}e_{t-1} + \sum_{r\geqslant 1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1}),\ t \in Z$$

is strictly stationary. Further, we have

$$\underline{X}_t = Ce_t + \underline{b}e_{t-1} + (A + Be_{t-1})\Big[Ce_{t-1} + \underline{b}e_{t-2} + \sum_{r\geqslant 1} \prod_{j=1}^{r} (A + Be_{t-1-j})(Ce_{t-1-r} + \underline{b}e_{t-2-r})\Big]$$

$= Ce_t + \underline{b}e_{t-1} + (A + Be_{t-1})\, \underline{X}_{t-1}$, $t \in Z$

$= A\, \underline{X}_{t-1} + \underline{b}e_{t-1} + B\, \underline{X}_{t-1}e_{t-1} + Ce_t$, $t \in Z$

6^0. Conversely, if $\underline{X}_t$, $t \in Z$ conforms to the bilinear model (2.3.2) above, we observe that for any $n \geqslant 2$

$$\underline{X}_t = Ce_t + \underline{b}e_{t-1} + \sum_{r=1}^{n-1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})$$

$$+ \prod_{j=1}^{n} (A + Be_{t-j})\, \underline{X}_{t-n} \qquad \text{a.e } [P] \quad (2.3.5)$$

for every t in Z.

As in 3^0, we note that for any $1 \leqslant u, v \leqslant p$,

$$E\left|\left(\prod_{j=1}^{n} (A + Be_{t-j})\right)_{uv}\right|$$

$$\leqslant \left\{E\left[\left(\prod_{j=1}^{n} (A + Be_{t-j})\right)_{uv}\right]^2\right\}^{\frac{1}{2}}$$

$$\leqslant K_2 \lambda^{\frac{n}{2}}, \text{ for some positive constant } K_2 > 0.$$

Since $\lambda < 1$,

$$\lim_{n\to\infty} E \prod_{j=1}^{n} (A + Be_{t-j}) = 0.$$

Since $\underline{X}_t$, $t \in Z$ is a strictly stationary process, $\underline{X}_{t-n}$, $n \geqslant 1$ converges to X_1 in distribution. Consequently,

$\prod_{j=1}^{n} (A + Be_{t-j})\, \underline{X}_{t-n}$, $n \geqslant 2$ converges to 0 in distribution and hence in probability. See Chung [10, Theroem 4.4.6, p.92]. We can find a subsequence of this sequence which converges to 0 a.e [P]. See Chung [10, Theorem 4.2.3, p.73]. Taking Limits along this subsequence in (2.3.5) we obtain

$$\underline{X}_t = Ce_t + \underline{b}e_{t-1} + \sum_{r\geqslant 1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1}) \qquad \text{a.e } [P]$$

for every t in Z. The almost sure convergence of the above series follows from the first part of the theorem.

—— 0 —— 0 —— The Theorem is proved —— 0 —— 0 ——

REMARKS 2.3.3

(1) If we are looking for a real valued strictly stationary process X_t, $t \in Z$ conforming to the bilinear model

$$X_t = a X_{t-1} + be_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P] \quad (2.3.6)$$

for every t in Z for a given sequence e_t, $t \in Z$ of independent identically distributed real random variables with common mean 0 and variance $\sigma^2 < \infty$, a sufficient condition for its existence is given by

$$a^2 + \beta^2 \sigma^2 < 1 \qquad (2.3.7)$$

The result follows at once by taking $p = q = 1$ in Theorem 2.3.2. The coefficient b plays no role at all in the above condition (2.3.7).

If we are merely looking for a real valued strictly stationary process X'_t, $t \in Z$ satisfying

$$X'_t = a X'_{t-1} + \beta X'_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P] \quad (2.3.8)$$

for every t in Z under the above assumptions on the e_t's, the same condition (2.3.7) is sufficient for its existence. The moving average part be_{t-1} of the process has no bearing on the existence of a strictly stationary process confoming to (2.3.6).

The model (2.3.6) without the moving average part was extensively studied by Tuan Dinh Pham and Lanh Tat Tran [41], Subba Rao [37] and Granger and Andersen [15] among others.

(2) The phenomenon described above in (1) also runs true in the general case (2.3.1). The same sufficient condition

$$\rho(A \otimes A + \sigma^2 B \otimes B) < 1$$

works true for the existence of a vector-valued strictly stationary process conforming to (2.3.1) with or without the presence of the moving average part $\underline{b}e_{t-1}$ in (2.3.1).

(3) The above phenomenon is not surprising. In Linear models, the above runs true. We will come to this part in the form of corollaries at the end of this section.

(4) Let e_t, $t \in Z$ be a sequence of independent identically distributed real random variables with common mean 0 and variance $\sigma^2 < \infty$. We are interested in the existence of a real strictly stationary BARMA $(r, 1, \ell, 1)$ model X_t, $t \in Z$, ie, X_t, $t \in Z$ satisfying

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + be_{t-1} + \sum_{j=1}^{\ell} \beta_{j1} X_{t-j} e_{t-1} + e_t \quad \text{a.e } [P] \quad (2.3.9)$$

This model can be put in the vector form as follows.

Let

$$p = \max \{r, \ell\}$$

$$q = 1$$

$$\underset{p\times p}{A} = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_r & \overbrace{0 \; 0 \; \dots \; 0 \; 0}^{(p-r)} \\ 1 & 0 & 0 & \dots & 0 & 0 \; 0 \; \dots \; 0 \; 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \; 0 \; \dots \; 0 \; 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \; \vdots \quad\;\; \vdots \; \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 \; 0 \; \dots \; 1 \; 0 \end{bmatrix}$$

$$\underset{p\times p}{B} = \begin{bmatrix} \beta_{11} & \beta_{21} & \dots & \beta_{\ell 1} & \overbrace{0 \; 0 \; \dots \; 0}^{(p-\ell)} \\ 0 & 0 & \dots & 0 & 0 \; 0 \; \dots \; 0 \\ 0 & 0 & \dots & 0 & 0 \; 0 \; \dots \; 0 \\ \vdots & \vdots & & \vdots & \vdots \; \vdots \quad\; \vdots \\ 0 & 0 & \dots & 0 & 0 \; 0 \; \dots \; 0 \end{bmatrix}$$

$$\underset{1\times p}{\underline{b}^T} = (b, 0, 0, \dots, 0)$$

$$\underset{1\times p}{C^T} = (1, 0, 0, \dots, 0)$$

$$\underset{1\times p}{X^T_t} = (X_t, X_{t-1}, \dots, X_{t-p+1}), \; t \in Z$$

Then,

$$\underline{X}_t = A\,\underline{X}_{t-1} + \underline{b}e_{t-1} + B\,\underline{X}_{t-1}e_{t-1} + Ce_t \qquad \text{a.e } [P]$$

for every t in Z.

A sufficient condition is

$\rho(A \otimes A + \sigma^2 B \otimes B) < 1$

Subba Rao [37] considered the model (2.3.9) without the moving average part be_{t-1} in (2.3.9) and with $r = \ell$.

(5) Let e_t, $t \in Z$ be a sequence of independent identically, distributed real random variables with common mean 0 and variance $\sigma^2 < \infty$. Suppose $a^2 + \beta^2 \sigma^2 = 1$ and $|a| < 1$. If e_1 is not two-valued, then there exists a stationary real valued process X_t, $t \in Z$ such that

$$X_t = a X_{t-1} + be_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P]$$

for every t in Z. To show the existence of the process, first we show that

$$\sum_{r \geq 1} E\left| \prod_{j=1}^{r} (a + \beta e_{t-j})(e_{t-r} + be_{t-r-1}) \right| < \infty$$

from which it follows that the series

$$\sum_{r \geq 1} \prod_{j=1}^{r} (a + \beta e_{t-j})(e_{t-r} + be_{t-r-1})$$

converges absolutely almost surely [P] as well as in the mean.

Note that for $r \geq 2$

$$E\left| \prod_{j=1}^{r} (a + \beta e_{t-j})(e_{t-r} + be_{t-r-1}) \right|$$

$$= E\left| \prod_{j=1}^{r-1} (a + \beta e_{t-j})(a + \beta e_{t-r})(e_{t-r} + be_{t-r-1}) \right|$$

$$= \left\{ \prod_{j=1}^{r-1} E|a + \beta e_{t-j}| \right\} E|(a + \beta e_{t-r})(e_{t-r} + be_{t-r-1})|$$

$$= K\, d^{r-1}$$

where $d = E|a + \beta e_1|$ and

$$K = E|(a + \beta e_{t-r})(e_{t-r} + be_{t-r-1})|$$

$$\leq |a| E|e_{t-r}| + |ab| E|e_{t-r-1}| + |\beta| E(e_{t-r}^2) + |b\beta| E|e_{t-r} e_{t-r-1}|$$

$$\leqslant (|a| + |ab|)\sigma + (|\beta| + |b\beta|)\sigma^2$$

and is independent of r and t.

Now, we claim that $d < 1$. Two cases arise. If e_1 is degenerate, then $e_1 = 0$ a.e [P]. Consequently,

$$\begin{aligned} d &= E|a + e_1| \\ &= E|a| \\ &= |a| \\ &< 1 \end{aligned}$$

If e_1 is not degenerate and not two-valued, then $|a + \beta e_1|$ is not degenerate. So,

$$\begin{aligned} d &= E|a + \beta e_1| \\ &< (E(a + \beta e_1)^2)^{\frac{1}{2}}, \text{ by Cauchy-Schwartz inequality} \\ &= (a^2 + \beta^2\sigma^2)^{\frac{1}{2}} \\ &= 1 \end{aligned}$$

This settles the claim. Consequently,

$$\sum_{r\geqslant 1} E\left|\prod_{j=1}^{r} (a + \beta e_{t-j})(e_{t-r} + be_{t-r-1})\right| < \infty .$$

If we set

$$X_t = e_t + be_{t-1} + \sum_{r\geqslant 1}\prod_{j=1}^{r} (a + \beta e_{t-j})(e_{t-r} + be_{t-r-1}),$$

then X_t, $t \in Z$ is the desired process.

(6) If one has a real strictly stationary process e_t, $t \in Z$ with common mean 0 and variance $\sigma^2 < \infty$, one would like to see whether there exists a strictly stationary vector-valued process $\underline{X}_t$, $t \in Z$ satisfying (2.3.2). The proof given above for Theorem 2.3.2 uses strongly the fact that e_t's are independent. Tuan Dinh Pham and Lanh Tat Tran [41] gave a proof based on the stong law of large numbers of the fact that there exists a real strictly stationary process X_t, $t \in Z$ satisfying

$$X_t = a X_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e [P]}$$

for every $t \in Z$ for a given sequence e_t, $t \in Z$ of independent identically distributed real random variables with common mean 0 and variance $\sigma^2 < \infty$ provided

$$a^2 + \beta^2\sigma^2 < 1$$

We establish a result now generalizing the above result of Tuan Dinh Pham and Lanh Tat Tran [41, Theorem 2.1, p. 618].

<u>Proposition 2.3.4</u> Let e_t, $t \in Z$ be an ergodic process with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Then there exists a strictly stationary process X_t, $t \in Z$ satisfying

$$X_t = a X_{t-1} + be_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P]$$

for every t in Z if $a^2 + \beta^2\sigma^2 < 1$.

<u>Proof</u> The definition of ergodic process is measure theoretic. We will give its definition later. Every ergodic process is pre-supposed to be strictly stationary. A consequence of a process Y_t, $t \in Z$ being ergodic is that, if $E|Y_0| < \infty$,

$$\frac{Y_0 + Y_1 + Y_2 + \dots + Y_{n-1}}{n}, \quad n \geq 1$$

converges a.e [P] to $E(Y_0)$.

To establish this result, we first show that the series

$$\sum_{r \geq 1} \prod_{j=1}^{r} (a + \beta e_{t-j})(e_{t-r} + be_{t-r-1})$$

converges a.e [P] for every t in Z.

Let us look at, for t in Z and $r \geq 1$,

$$P(t, r) = \prod_{j=1}^{r} (a + \beta e_{t-j})$$

Taking logarithms, we obtain

$$\frac{1}{r}\log|P(t,r)| = \frac{1}{r} \sum_{j=1}^{r} \log|a + \beta e_{t-j}|$$

Since e_t, $t \in Z$ is ergodic, $\log|a + \beta e_{t-j}|$, j=1, 2, ... , is ergodic.

Let us check

$$E \log |a + \beta e_0| = \tfrac{1}{2} E \log (a + \beta e_0)^2$$
$$\leq \tfrac{1}{2} \log E (a + \beta e_0)^2$$
$$= \tfrac{1}{2} \log (a^2 + \beta^2\sigma^2) < 0.$$

By the ergodic theorem,

$$\lim_{r\to\infty} \frac{1}{r} \sum_{j=1}^{r} \log |a + \beta e_{t-j}| = E \log |a + \beta e_0| \qquad \text{a.e } [P]$$
$$< 0.$$

The rest of the details follow in exactly the same way as in the proof of theorem 2.1 of [41].

<u>Query</u>. Is X_t, $t \in Z$ defined above ergodic?

The answer to this question is in the affirmative. At this juncture, let us recapitulate what is meant by an ergodic process. Let e_t, $t \in Z$ be a strictly stationary process defined on some probability space $(\Omega, \mathcal{B}, P)$. Ergodicity of e_t, $t \in Z$ can be defined in two ways. Let, for each $t \in Z$,

$$\mathcal{F}_t = \sigma\{e_t, e_{t-1}, e_{t-2}, \ldots,\}$$
= the smallest sub-σ-field of $\mathcal{B}$ with respect to which $e_t, e_{t-1}, e_{t-2}, \ldots$ are measurable.

Note that

$$\ldots \supset \mathcal{F}_3 \supset \mathcal{F}_2 \supset \mathcal{F}_1 \supset \mathcal{F}_0 \supset \mathcal{F}_{-1} \supset \mathcal{F}_{-2} \supset \ldots$$

The tail σ-field of e_t, $t \in Z$ is defined to be the σ-field

$$\mathcal{F} = \bigcap_{t=-\infty}^{\infty} \mathcal{F}_t = \bigcap_{t=0}^{-\infty} \mathcal{F}_t .$$

e_t, $t \in Z$ is said to be ergodic, if its tail σ-field $\mathcal{F}$ is P-trivial, ie, for every A in $\mathcal{F}$, $P(A) = 0$ or 1.

Equivalently, ergodicity can be defined in the following way. For each t in Z, let

$$\mathcal{F}'_t = \sigma\{e_t, e_{t+1}, e_{t+2}, \ldots,\}$$

Note that

$$\cdots \supset \mathcal{F}'_{-2} \supset \mathcal{F}'_{-1} \supset \mathcal{F}'_0 \supset \mathcal{F}'_1 \supset \mathcal{F}'_2 \supset \mathcal{F}'_3 \supset \cdots$$

The tail σ-field is defined to be the σ-field

$$\mathcal{F}' = \bigcap_{t=-\infty}^{\infty} \mathcal{F}'_t$$

If $\mathcal{F}'$ is P-trivial, e_t, $t \in Z$ is ergodic.

Proposition 2.3.5 Let e_t, $t \in Z$ be an ergodic process with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. If $a^2 + \beta^2\sigma^2 < 1$, then there exists an ergodic process X_t, $t \in Z$ satisfying

$$X_t = a X_{t-1} + b e_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P]$$

for every t in Z.

Proof By Proposition 2.3.4, existence and strict stationarity of X_t, $t \in Z$ is assured, we show that X_t, $t \in Z$ is ergodic. Let us find its tail σ-field. For each t in Z, let

$$\mathcal{F}_t = \sigma\{e_t, e_{t-1}, e_{t-2}, \ldots\} \text{ and}$$

$$\mathcal{F}_t^* = \sigma\{X_t, X_{t-1}, X_{t-2}, \ldots\}$$

From the representation of X_t as a function of e_t, e_{t-1}, ... it is clear that X_t is measurable with respect to $\mathcal{F}_t$. Also, X_{t-1}, X_{t-2}, ... are all measurable with respect to $\mathcal{F}_t$. Consequently, $\mathcal{F}_t^* \subset \mathcal{F}_t$ for every t in Z. (Note that, if each e_t is a function of $X_t, X_{t-1}, \ldots$, then $\mathcal{F}_t \subset \mathcal{F}_t^*$.). Therefore,

$$\mathcal{F}^* = \text{Tail } \sigma\text{-field of } X_t,\ t \in Z$$

$$= \bigcap_{t=-\infty}^{\infty} \mathcal{F}_t^*$$

$$\subset \bigcap_{t=-\infty}^{\infty} \mathcal{F}_t = \text{Tail } \sigma\text{-field of } e_t,\ t \in Z$$

$$= \mathcal{F}.$$

Since $\mathcal{F}$ is P-trivail, $\mathcal{F}^*$ is P-trivial. Hence X_t, $t \in Z$ is ergodic.

More generally, we have the following result.

Theorem 2.3.6 Let e_t, $t \in Z$ be a sequence of independent identically distributed random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let A and B be two matrices each of order pxp such that $\rho(A \otimes A + \sigma^2 B \otimes B) < 1$. Then given any $\underset{p\times 1}{C}$ and $\underset{p\times 1}{\underline{b}}$, there exists an ergodic process X_t, $t \in Z$ satisfying

$$\underline{X}_t = A\,\underline{X}_{t-1} + \underline{b}e_{t-1} + B\,\underline{X}_{t-1}e_{t-1} + Ce_t \qquad \text{a.e } [P]$$

for every t in Z.

Proof From the representation of $\underline{X}_t$ given in the proof of Theorem 2.3.2, it is clear that each $\underline{X}_t$ is a function of e_t, e_{t-1}, ... Note that any sequence of independent random variables has P-trivial tail σ-field. This is Kolmogorov's Zero-One Law. The rest of the details are similar to the ones given above.

(7) If the stochastic process to be modelled for a given time series data started only a finite number of steps ago, the same condition stipulated in the above theorem guarantees that the process involved is asymptotically stationary. To be more specific, suppose the p-variate process starts at time t = 0 with the initial random vector being $\underline{X}_0$ and satisfies

$$\underline{X}_t = Ce_t + \underline{b}e_{t-1} + A\,\underline{X}_{t-1} + B\,\underline{X}_{t-1}e_{t-1} \qquad \text{a.e } [P] \quad (2.3.10)$$

for t = 1, 2, 3, ... for some sequence $\{e_0, e_1, e_2, \dots\}$ of independent, identically distributed random variables with common mean $E(e_0) = 0$ and common variance $E(e_0^2) = \sigma^2 < \infty$ and for some constant matrices A, B, $\underline{b}$, and C of orders pxp, pxp, px1 and px1 respectively. Repeated use of (2.3.10) gives

$$\underline{X}_t = Ce_t + \underline{b}e_{t-1} + \sum_{r=1}^{t-1}\prod_{j=1}^{r}(A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})$$

$$+ \prod_{j=1}^{t}(A + Be_{t-j})\underline{X}_0$$

for every $t = 2, 3, 4, \ldots$.

The process $\underline{Y}_t$, $t = 2, 3, 4, \ldots$ defined by

$$\underline{Y}_t = Ce_0 + \underline{b}e_1 + \sum_{r=1}^{t-1} \prod_{j=1}^{r} (A + Be_j)(Ce_r + \underline{b}e_{r+1}) + \prod_{j=1}^{t} (A + Be_j)\, \underline{X}_0$$

has the property that $\underline{X}_t$ and $\underline{Y}_t$ have the same distribution for every $t = 2, 3, 4, \ldots$. This follows from the fact that $\{e_0, e_1, e_2, \ldots\}$ are independently identically distributed. Under the condition

$\rho(A \otimes A + \sigma^2 B \otimes B) < 1$,

$$\sum_{r \geq 1} \prod_{j=1}^{r} (A + Be_j)(Ce_r + \underline{b}e_{r+1})$$

converges absolutely a.e [P] and

$$\left\{ \prod_{j=1}^{t} (A + Be_j)\, \underline{X}_0,\ t \geq 2 \right\}$$

converges to 0 in probability. Consequently, the process $\{\underline{Y}_t, t = 2, 3, 4, \ldots\}$, and hence the process $\{\underline{X}_t, t = 2, 3, 4, \ldots\}$ converges to the random vector

$$Ce_0 + \underline{b}e_1 + \sum_{r \geq 1} \prod_{j=1}^{r} (A + Be_j)(Ce_r + \underline{b}e_{r+1})$$

in distribution. See Chung [10, Theorem 4.4.6, p.92]. One cannot fail to notice that the distribution of the limiting random vector above is the same as that of $\underline{X}_t$, $t \in Z$ whose representation is given in the above theorem.

(8) If $E(e_1^4) < \infty$, then we can show that the series

$$\sum_{r \geq 1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})$$

converges in the quadratic mean. The proof given above for the theorem can easily be adapted to establish this.

Finally, as promised in (3) of Remarks 2.3.3, we obtain results on existence of certain linear processes as corollaries of Theorem 2.3.2.

Corollary 2.3.7 Let e_t, $t \in Z$ be a sequence of independent identically distributed real random variables with a common mean 0 and variance $\sigma^2 < \infty$. Then there exists a strictly stationary process X_t, $t \in Z$ satisfying

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + be_{t-1} + e_t \qquad \text{a.e } [P] \quad (2.3.11)$$

for every t in Z if the roots of the polynomial

$$f(x) = 1 - a_1 x - a_2 x^2 - \dots - a_r x^r \qquad (2.3.12)$$

are in absolute value greater than unity.

Proof The model (2.3.11) can be put in the vector form as follows, let

$$p = r$$

$$q = 1$$

$$\underset{p\times p}{A} = \begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_{p-1} & a_p \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

$$\underset{1\times p}{\underline{b}^T} = (b, 0, 0, \dots, 0)$$

$$\underset{1\times p}{C^T} = (1, 0, 0, \dots, 0)$$

$$\underset{1\times p}{X_t^T} = (X_t, X_{t-1}, \dots, X_{t-p+1}), \ t \in Z$$

Then

$$\underline{X}_t = A\,\underline{X}_{t-1} + \underline{b}e_{t-1} + Ce_t \qquad \text{a.e } [P] \quad (2.3.13)$$

for every t in Z.

A sufficient condition for the existence of a strictly stationary vector-valued process $\underline{X}_t$, $t \in Z$ satisfying (2.3.13) is

$$\rho(A \otimes A) < 1 \qquad (2.3.14)$$

By Lemma 2.2.2 (d), condition (2.3.14) is equivalent to $\rho(A) < 1$, which in turn is equivalent to the condition that the roots of the characteristic polynomial

$$g(x) = x^r - a_1 x^{r-1} - a_2 x^{r-2} - \dots - a_r \qquad (2.3.15)$$

are less than one in modulus.

Since the roots of (2.3.12) are the reciprocals of the roots of (2.3.15), the result follows.

Corollary 2.3.8. Let e_t, $t \in Z$ be a sequence of independent identically distributed real random variables with common mean 0 and variance $\sigma^2 < \infty$. Then there exists a strictly stationary process X_t, $t \in Z$ satisfying

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + e_t \qquad \text{a.e } [P] \quad (2.3.16)$$

for every t in Z if the roots of the polynomial

$$f(x) = 1 - a_1 x - a_2 x^2 - \dots - a_r x^r$$

are in absolute value greater than unity.

Proof. Now (2.3.16) is the same as (2.3.11) without the moving average part be_{t-1} in (2.3.11). The result follows from Corollary 2.3.7 since the condition (2.3.14) holds true for the existence of a strictly stationary process conforming to (2.3.11) with or without the presence of the moving average part be_{t-1} in (2.3.11).

2.4 EXISTENCE THEOREM: GENERAL CASE

Before stating our next theorem on the existence of a vector-valued process $\underline{X}_t$, $t \in Z$ satisfying the general vectorial model (2.1.14), we give a Lemma which on its own may be of independent interest.

Lemma 2.4.1. If a_n, $n \geq 1$ is a sequence of real numbers satisfying

$$|a_n - a_{n-1}| < K \lambda^n \qquad (2.4.1)$$

for every $n \geq 2$ for some positive constant K and $\lambda < 1$, then a_n, $n \geq 1$ is a Cauchy sequence and hence convergent.

Proof. Let b_n, $n \geq 1$ be defined by

$$b_n = \sum_{j=1}^{n} \lambda^j \ , \ 0 < \lambda < 1$$

Then $b_n \to \lambda/(1 - \lambda)$ as $n \to \infty$. This implies that for each positive number ε, we can find a positive number N such that

$$|b_n - b_m| \leqslant {}^{\varepsilon}/_{K} \text{ for all integers } n, m \geqslant N$$

for some positive constant K. It is not difficult to show that if $n, m \geqslant N$ then

$$|a_n - a_m| \leqslant K|b_n - b_m|$$
$$< \varepsilon$$

Therefore a_n, $n \geqslant 1$ satisfying (2.4.1) for some positive constant K and $\lambda < 1$ is a Cauchy sequence of real numbers. Hence a_n, $n \geqslant 1$ is convergent.

REMARK 2.4.2

If $\lim_{n\to\infty} |a_n - a_{n-1}| = 0$, it does not necessarily mean that a_n, $n \geqslant 1$ converges. To see this, consider

$$a_n = 1 + \tfrac{1}{2} + \tfrac{1}{3} + \dots + \tfrac{1}{n};\ n \geqslant 1$$

It is evident that $a_n \to \infty$ as $n \to \infty$ but $|a_n - a_{n-1}| = \frac{1}{n} \to 0$ as $n \to \infty$.

The following is the main result of this section.

Theorem 2.4.3 Let e_t, $t \in Z$ be a sequence of independent identically distributed real random variables defined on some probability space $(\Omega, \mathcal{F}, P)$ such that $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let A, B_1, B_2, ... B_q be q + 1 matrices each of order pxp and

$$\Gamma_1 = A \otimes A + \sigma^2 B_1 \otimes B_1 \qquad (2.4.2)$$

$$\Gamma_j = \sigma^2[B_j \otimes (A^{j-1} B_1 + A^{j-2} B_2 + \dots + AB_{j-1} + B_j)$$
$$+ (A^{j-1} B_1 + A^{j-2} B_2 + \dots + AB_{j-1}) \otimes B_j]$$
$$(j = 2, 3, \dots, q) \qquad (2.4.3)$$

Suppose all the eigen values of the matrix

$$\underset{p^2qxp^2q}{\Gamma} = \begin{pmatrix} \Gamma_1 & \Gamma_2 & \Gamma_3 & \cdots & \Gamma_{q-1} & \Gamma_q \\ I_{p^2} & \underline{0} & \underline{0} & \cdots & \underline{0} & \underline{0} \\ \underline{0} & I_{p^2} & \underline{0} & \cdots & \underline{0} & \underline{0} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \underline{0} & \underline{0} & \underline{0} & \cdots & I_{p^2} & \underline{0} \end{pmatrix} \tag{2.4.4}$$

have moduli less than unity, ie $\rho(\Gamma) = \lambda < 1$. Let C, $\underline{b}_1$, $\underline{b}_2$, ... , $\underline{b}_q$ be $q + 1$ column vectors each of order $p \times 1$. Then there exists a vector-valued strictly stationary process $\underline{X}_t$, $t \in Z$ conforming to the bilinear model

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j e_{t-j} + \sum_{j=1}^{q} B_j\,\underline{X}_{t-j} e_{t-j} + Ce_t \quad \text{a.e } [P] \tag{2.4.5}$$

for every t in Z.

<u>Proof</u> For this general case, it is not easy to provide an infinte series representation (as we did in the case of $q = 1$) for each $\underline{X}_t$, $t \in Z$. As a result we will proceed as in Bhaskara Rao, Subba Rao and Walker [4] and exhibit the process $\underline{X}_t$, $t \in Z$ as an almost sure limit of a sequence $\underline{S}_{n,t}$, $t \in Z$, $n \geq 1$ of strictly stationary processes. The proof is broken down into the following steps.

1°. Let the process $\underline{S}_{n,t}$, $n, t \in Z$ be defined as follows

$$\begin{aligned} \underline{S}_{n,t} &= 0\,, \text{ if } n < 0 \\ &= Ce_t\,, \text{ if } n = 0 \\ &= Ce_t + \sum_{j=1}^{q} \underline{b}_j e_{t-j} + (A + B_1 e_{t-1})\underline{S}_{n-1,t-1} + B_2\,\underline{S}_{n-2,t-2}e_{t-2} + \cdots \\ &\quad + B_q\,\underline{S}_{n-q,t-q}e_{t-q}\,, \text{ if } n > 0 \end{aligned} \tag{2.4.6}$$

for every t in Z.

We show that $\lim_{n\to\infty} \underline{S}_{n,t}$ exists almost surely $[P]$ for every t in Z.

If $\underline{X}_t$ is the almost sure limit of $\underline{S}_{n,t}$, $n \geq 1$ for every t in Z, then it is obvious that the process $\underline{X}_t$, $t \in Z$ conforms to the bilinear model (2.4.5). Using Lemma 2.3.1, it is easy to check that for every fixed n in Z, $\underline{S}_{n,t}$, $t \in Z$ is a strictly stationary process.

2^0. Let $\underline{s}_{n,t} = \underline{S}_{n,t} - \underline{S}_{n-1,t}$, n, $t \in Z$.

We show that

$$E|(\underline{s}_{n,t})_i| \leq k\,\lambda^{\frac{n}{2}}$$

for every $n > 0$ and $i = 1, 2, \ldots, p$, where k is a positive constant. Since $\lambda < 1$, this would then imply that $\underline{S}_{n,t}$, $n \geq 1$ converges almost surely [P] for every t in Z (See Lemma 2.4.1).

3^0. We now settle the question of integrability of $\underline{s}_{n,t}$'s. Fix t in Z. Note that

$$\begin{aligned} \underline{s}_{nt}, &= \underline{S}_{n,t} - \underline{S}_{n-1,t} \\ &= (A + B_1 e_{t-1})\underline{s}_{n-1,t-1} + B_2\,\underline{s}_{n-2,t-2}e_{t-2} + \ldots + B_q\,\underline{s}_{n-q,t-q}e_{t-q} \\ &= Q_n(e_{t-1}, e_{t-2}, \ldots, e_{t-n})\underline{s}_{o,t-n} \\ &= Q_n(e_{t-1}, e_{t-2}, \ldots, e_{t-n})Ce_{t-n} \end{aligned} \qquad (2.4.7)$$

where $Q_n(e_{t-1}, e_{t-2}, \ldots, e_{t-n})$ is a matrix of order pxp and each entry of this matrix is a polynomial in $e_{t-1}, e_{t-2}, \ldots, e_{t-n}$ in which the power index of each e_{t-j} is either 0 or 1. Consequently, every entry in $Q_n(e_{t-1}, e_{t-2}, \ldots e_{t-n})$ and hence in $\underline{s}_{n,t}$ is integrable. It is clear that the distribution of $\underline{s}_{n,t}$ does not depend on t.

4^0. It is convenient to deal with the following processes. Define

$$\begin{aligned} \underline{s}^*_{n,t} &= 0\ , \text{ if } n < 0 \\ &= C\ , \text{ if } n = 0 \\ &= Q_n(e_{t-1}, e_{t-2}, \ldots, e_{t-n})C\ , \text{ if } n > 0 \end{aligned}$$

for every t in Z. Equivalently

$$\underline{s}_{n,t} = \underline{s}^*_{n,t}e_{t-n}\ , \ n,\ t \in Z \qquad (2.4.8)$$

From the remark made on $Q_n(\cdot)$'s in 3^0, it is obvious that every entry in $\underline{s}^*_{n,t}$ is square integrable. Further, it is easy to check that the $\underline{s}^*_{n,t}$'s satisfy the following

$$\underline{s}^*_{n,t} = (A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1} + B_2\underline{s}^*_{n-2,t-2}e_{t-2} + \ldots B_q\,\underline{s}^*_{n-q,t-q}e_{t-q} \qquad (2.4.9)$$

for every n,t in Z. Also, the distribution of $\underline{s}^*_{n,t}$ does not depend on t, since the e_t's are independently identically distributed. Since $\underline{s}_{n,t} = \underline{s}^*_{n,t} e_{t-n}$ for all n and t in Z.

$$\begin{aligned} E|(\underline{s}_{n,t})_i| &= E|(\underline{s}^*_{n,t})_i||e_{t-n}| \\ &\leq (E((\underline{s}^*_{n,t})_i)^2)^{\frac{1}{2}}(E(e_{t-n})^2)^{\frac{1}{2}} \\ &\leq \sigma(E((\underline{s}^*_{n,t})_i)^2)^{\frac{1}{2}} \end{aligned}$$

for every i = 1, 2, ... p. It suffices to obtain an upper bound for $E((\underline{s}_{n,t})_i)^2$ for every i = 1, 2, ... , p and n,t in Z. For this we evaluate

$$E(\underline{s}^*_{n,t} \otimes \underline{s}^*_{n,t}) = M_n \text{ , say.}$$

5^0. In the following, we use (2.4.9) and (2.2.4)

$$\begin{aligned} \underline{s}^*_{n,t} \otimes \underline{s}^*_{n,t} &= [(A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1} + B_2 \underline{s}^*_{n-2,t-2} e_{t-2} + \cdots \\ &\quad + B_q \underline{s}^*_{n-q,t-q} e_{t-q}] \otimes [(A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1} \\ &\quad + B_2 \underline{s}^*_{n-2,t-2} e_{t-2} + \cdots + B_q \underline{s}^*_{n-q,t-q} e_{t-q}] \\ &= \{((A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1}) \otimes ((A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1})\} \\ &\quad + \{((A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1}) \otimes (B_2 \underline{s}^*_{n-2,t-2} e_{t-2}) \\ &\quad + (B_2 \underline{s}^*_{n-2,t-2} e_{t-2}) \otimes ((A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1}) \\ &\quad + (B_2 \underline{s}^*_{n-2,t-2} e_{t-2}) \otimes (B_2 \underline{s}^*_{n-2,t-2} e_{t-2})\} \\ &\quad + \{((A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1}) \otimes (B_3 \underline{s}^*_{n-3,t-3} e_{t-3}) \\ &\quad + (B_3 \underline{s}^*_{n-3,t-3} e_{t-3}) \otimes ((A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1}) \\ &\quad + (B_2 \underline{s}^*_{n-2.t-2} e_{t-2}) \otimes (B_3 \underline{s}^*_{n-3,t-3} e_{t-3}) \\ &\quad + (B_3 \underline{s}^*_{n-3,t-3} e_{t-3}) \otimes (B_2 \underline{s}^*_{n-2,t-2} e_{t-2}) \\ &\quad + (B_3 \underline{s}^*_{n-3,t-3} e_{t-3}) \otimes (B_3 \underline{s}^*_{n-3,t-3} e_{t-3})\} \\ &\quad + \cdots\cdots\cdots\cdots\cdots\cdots \\ &\quad + \{((A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1}) \otimes (B_q \underline{s}^*_{n-q,t-q} e_{t-q}) \end{aligned}$$

$$+ (B_q \underline{s}^*_{n-q,t-q} e_{t-q}) \otimes ((A + B_1 e_{t-1}) \underline{s}^*_{n-1,t-1})$$

$$+ (B_2 \underline{s}^*_{n-2,t-2} e_{t-2}) \otimes (B_q \underline{s}^*_{n-q,t-q} e_{t-q})$$

$$+ (B_q \underline{s}^*_{n-q,t-q} e_{t-q}) \otimes (B_2 \underline{s}^*_{n-2,t-2} e_{t-2})$$

$$+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$+ (B_{q-1} \underline{s}^*_{n-q+1,t-q+1} e_{t-q+1}) \otimes (B_q \underline{s}^*_{n-q,t-q} e_{t-q})$$

$$+ (B_q \underline{s}^*_{n-q} e_{t-q}) \otimes (B_{q-1} \underline{s}^*_{n-q+1,t-q+1} e_{t-q+1})$$

$$+ (B_q \underline{s}^*_{n-q,t-q} e_{t-q}) \otimes (B_q \underline{s}^*_{n-q,t-q} e_{t-q})\} \qquad (2.4.10)$$

We evaluate the expectation of each expression within each set of brackets { } in (2.4.10).

6^0. Consider the expression within the first set of brackets { } in (2.4.10). By (2.2.5) and (2.2.4)

$$((A + B_1 e_{t-1}) \underline{s}^*_{n-1,t-1}) \otimes ((A + B_1 e_{t-1}) \underline{s}^*_{n-1,t-1})$$

$$= ((A + B_1 e_{t-1}) \otimes (A + B_1 e_{t-1}))(\underline{s}^*_{n-1,t-1} \otimes \underline{s}^*_{n-1,t-1})$$

$$= (A \otimes A + e_{t-1} A \otimes B_1 + e_{t-1} B_1 \otimes A + e^2_{t-1} B_1 \otimes B_1)(\underline{s}^*_{n-1,t-1} \otimes \underline{s}^*_{n-1,t-1})$$

Since $\underline{s}^*_{n-1,t-1}$ is a function of $e_{t-2}, e_{t-3}, \ldots, e_{t-n}$; $\underline{s}^*_{n-1,t-1}$ and e_{t-1} are independently distributed. So,

$$E((A + B_1 e_{t-1}) \otimes (A + B_1 e_{t-1}))(\underline{s}^*_{n-1,t-1} \otimes \underline{s}^*_{n-1,t-1})$$

$$= (A \otimes A + \sigma^2 B_1 \otimes B_1) M_{n-1} \qquad (2.4.11)$$

$$= \Gamma_1 M_{n-1}$$

7^0. Consider the following expression in the second set of such brackets. Using (2.2.4), (2.2.5) and (2.4.9) to expand $\underline{s}^*_{n-1,t-1}$, we obtain

$$((A + B_1 e_{t-1}) \underline{s}^*_{n-1,t-1}) \otimes (B_2 \underline{s}^*_{n-2,t-2} e_{t-2})$$

$$= ((A + B_1 e_{t-1})((A + B_1 e_{t-2}) \underline{s}^*_{n-2,t-2} + B_2 \underline{s}^*_{n-3,t-3} e_{t-3}$$

$$+ B_3 \underline{s}^*_{n-4,t-4} e_{t-4} + \ldots + B_q \underline{s}^*_{n-q-1,t-q-1} e_{t-q-1})) \otimes (B_2 \underline{s}^*_{n-2,t-2} e_{t-2})$$

$$= ((A + B_1 e_{t-1})(A + B_1 e_{t-2})\underline{s}^*_{n-2,t-2}) \otimes (B_2\, \underline{s}^*_{n-2,t-2} e_{t-2})$$

$$+ ((A + B_1 e_{t-1}) B_2\, \underline{s}^*_{n-3,t-3} e_{t-3}) \otimes (B_2\, \underline{s}^*_{n-2,t-2} e_{t-2})$$

$$+ \cdots\cdots\cdots\cdots\cdots\cdots$$

$$+ ((A + B_1 e_{t-1}) B_q\, \underline{s}^*_{n-q-1,t-q-1} e_{t-q-1}) \otimes (B_2\, \underline{s}^*_{n-2,t-2} e_{t-2})$$

$$= ((A + B_1 e_{t-1})(A + B_1 e_{t-2}) \otimes B_2 e_{t-2})(\underline{s}^*_{n-2.t-2} \otimes \underline{s}^*_{n-2,t-2})$$

$$+ ((A + B_1 e_{t-1}) \otimes B_2 e_{t-2})(B_2\, \underline{s}^*_{n-3,t-3} e_{t-3} \otimes \underline{s}^*_{n-2,t-2})$$

$$+ \cdots\cdots\cdots\cdots\cdots\cdots$$

$$+ ((A + B_1 e_{t-1}) \otimes B_2 e_{t-2})(B_q\, \underline{s}^*_{n-q-1,t-q-1} e_{t-q-1} \otimes \underline{s}^*_{n-2,t-2})$$

Therefore

$$E((A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1}) \otimes (B_2\, \underline{s}^*_{n-2.t-2} e_{t-2})$$

$$= E((A + B_1 e_{t-1})(A + B_1 e_{t-1}) \otimes B_2\, e_{t-2}) E(\underline{s}^*_{n-2,t-2} \otimes \underline{s}^*_{n-2.t-2})$$

$$+ 0 + 0 + \ldots + 0$$

$$= E((A + B_1 e_{t-1})(Ae_{t-2} + B_1 e^2_{t-2}) \otimes B_2) M_{n-2}$$

$$= ((E(A + B_1 e_{t-1}) E(Ae_{t-2} + B_1 e^2_{t-2})) \otimes B_2) M_{n-2}$$

$$= \sigma^2((AB_1) \otimes B_2) M_{n-2}$$

In a similar fashion, we can show that

$$E((B_2\, \underline{s}^*_{n-2,t-2} e_{t-2}) \otimes ((A + B_1 e_{t-1})\underline{s}^*_{n-1,t-1}))$$

$$= \sigma^2(B_2 \otimes (AB_1)) M_{n-2}, \quad \text{and}$$

$$E((B_2\, \underline{s}^*_{n-2,t-2} e_{t-2}) \otimes (B_2\, \underline{s}^*_{n-2,t-2} e_{t-2}))$$

$$= \sigma^2(B_2 \otimes B_2) M_{n-2}.$$

Consequently, the expected value of the entire expression in the second set of such brackets in (2.4.10) is

$$\sigma^2[B_2 \otimes (AB_1) + (AB_1) \otimes B_2 + B_2 \otimes B_2] M_{n-2}$$

$$= \Gamma_2\, M_{n-2} \tag{2.4.12}$$

8^0. Pursuing ideas similar to those used in 7^0, we can show that the expected value of the entire expression in the third set of such brackets in (2.4.10) is

$$\sigma^2[B_3 \otimes (A^2B_1 + AB_2) + (A^2B_1 + AB_2) \otimes B_3 + B_3 \otimes B_3]M_{n-3}$$

$$= \Gamma_3 M_{n-3} \tag{2.4.13}$$

9^0. The expectations of other expressions in (2.4.10) can be evaluated analogously. Finally, we obtain

$$M_n = E(\underline{s}^*_{n,t} \otimes \underline{s}^*_{n,t})$$

$$= \sum_{j=1}^{q} \Gamma_j M_{n-j} \tag{2.4.14}$$

for all n.

10^0. For each $n \geq 1$, let

$$\underset{p^2q\times 1}{\underline{Y}_n} = \begin{pmatrix} M_n \\ M_{n-1} \\ \vdots \\ M_{n-q+1} \end{pmatrix}$$

Then

$$\underline{Y}_n = \begin{pmatrix} \Gamma_1 & \Gamma_3 & \cdots & \Gamma_{q-1} & \Gamma_q \\ I_{p^2} & \underline{0} & \cdots & \underline{0} & \underline{0} \\ \underline{0} & I_{p^2} & \cdots & \underline{0} & \underline{0} \\ \vdots & \vdots & & \vdots & \vdots \\ \underline{0} & \underline{0} & \cdots & I_{p^2} & \underline{0} \end{pmatrix} \begin{pmatrix} M_{n-1} \\ M_{n-2} \\ \vdots \\ M_{n-q} \end{pmatrix}$$

$$= \Gamma \underline{Y}_{n-1}$$

$$= \Gamma^2 \underline{Y}_{n-2} = \cdots = \Gamma^n \underline{Y}_0 \tag{2.4.15}$$

From this, it follows that $|(\underline{Y}_n)_i| \leqslant K_3\ \lambda^n$ for every i = 1, 2, ... , p^2q and $n \geqslant 1$, where $\lambda = \rho(\Gamma)$ and K_3 is a positive constant. In particular, we have $|(M_n)_i| \leqslant K_3\ \lambda^n$ for every i = 1, 2, ... , p^2 and $n \geqslant 1$. Since

$$M_n = E(\underline{s}^*_{n,t} \otimes \underline{s}^*_{n,t}),$$

we have

$$E((\underline{s}^*_{n,t})_i)^2 \leqslant K_3\ \lambda^n$$

for every i = 1, 2, ... , p and n,t in Z. From this inequality, the inequality stated in 2^0. follows. See also 4^0.

—— 0 —— 0 —— The Theorem is proved —— 0 —— 0 ——

REMARKS 2.4.4

(1) The most important feature that emerges by comparing Theorem 2.4.3 above and the Theorem in section 4 of Bhaskara Rao, Subba Rao and Walker [4] is that the presence of moving average part makes no impact on the existence problem. This is also typical of Linear processes as the following corollary shows.

Corollary 2.4.5 Let e_t, t ε Z be a sequence of independent identically distributed real random variables with common mean 0 and variance $\sigma^2 < \infty$. Then there exists a strictly stationary process X_t, t ε Z satisfying

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{\ell} b_j e_{t-j} + e_t \qquad \text{a.e } [P] \quad (2.4.16)$$

for every t in Z if the roots of the polynomial

$$f(x) = 1 - a_1 x - a_2 x^2 - \dots - a_r x^r$$

are in absolute value greater than unity.

Proof. The model (2.4.16) can be put in the vector form as follows. Let

$$p = r$$

$$q = \ell$$

$$\underset{p\times p}{A} = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{p-1} & a_p \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

$$\underset{1\times p}{\underline{b}_j^T} = (b_j, 0, 0, \ldots, 0), \; j = 1, 2, \ldots, q$$

$$\underset{1\times p}{C^T} = (1, 0, 0, \ldots, 0)$$

$$\underset{1\times p}{\underline{X}_t^T} = (X_t, X_{t-1}, \ldots, X_{t-p+1}), \; t \in Z$$

Then

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j e_{t-j} + Ce_t \qquad \text{a.e } [P] \quad (2.4.17)$$

for every t in Z.

A sufficient condition for the existence of a strictly stationary real vector-valued process $\underline{X}_t$, $t \in Z$ satisfying (2.4.17) is

$$\rho(A \otimes A) < 1$$

The rest of the details follow in exactly the same way as in the proof of corollary 2.3.7 of section 2.3.

(2) Finally, we remark that the process satisfying (2.4.5) in Theorem 2.4.3 is ergodic. This follows from the following observations. For every $n \geq 1$, the process $\underline{S}_{n,t}$, $t \in Z$ defined in step 1^0 in the proof of Theorem 2.4.3 is ergodic. Since

$$\lim_{n\to\infty} \underline{S}_{n,t} = \underline{X}_t \qquad \text{a.e } [P]$$

for every t in Z, $\underline{X}_t$, $t \in Z$ is ergodic.

CHAPTER 3

ON THE MOMENTS OF BILINEAR PROCESSES

3.1 INTRODUCTION

The purpose of this chapter is to develop a method of calculating all joint moments up to order 2 of the vector-valued process $\underline{X}_t$, $t \in Z$ satisfying

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j e_{t-j} + \sum_{j=1}^{q} B_j \underline{X}_{t-j} e_{t-j} + Ce_t \quad \text{a.e } [P] \quad (3.1.1)$$

for every t in Z for some sequence e_t, $t \in Z$ of independent identically distributed real random variables with common mean 0 and variance $\sigma^2 < \infty$ and matrices $\underset{p\times p}{A}$, $\underset{p\times 1}{C}$, $\underset{p\times 1}{\underline{b}_j}$, $j = 1, 2, \ldots, q$ and $\underset{p\times p}{B_j}$, $j = 1, 2, \ldots, q$.

It is tempting to assume that since the bilinear process $\underline{X}_t$, $t \in Z$ satisfying (3.1.1) is strictly stationary, it must be stationary up to any order. This is not necessarily true since a process may be strictly stationary even though none of its moments exist. Such a process will not be useful since the main tools of time series analysis have traditionally been the first and second moments of the series. We show that under the strict stationarity condition

$$\rho(A \otimes A + \sigma^2 B \otimes B) < 1$$

for the case $q = 1$, and

$$\rho(\Gamma) < 1$$

for the general case of $q > 1$, and $E(e_t^4) < \infty$, the vector-valued bilinear process defined by (3.1.1) is second-order stationary. This implies, in particular, the existence of all joint moments up to order 2.

Subba Rao [37, p.248] and Granger and Andersen [15] have discussed some special cases of the general model (3.1.1) and showed that for these special cases the covariance structure is identical with the covariance structure of some suitable linear processes. See also Subba Rao and

Gabr [39].

In this chapter, we show that given any real bilinear model (2.1.2), there exists a linear process such that their covariance structures are identical. I was informed by my supervisor that Tuan Dinh Pham [46] has also arrived at the same conclusion after obtaining a Markovian representation of bilinear processes. But our method here is simple and direct.

We also show that for the general vector-valued bilinear model (3.1.1) there exists a vector-valued linear process such that their covariance structures are identical whenever the matrix A in (3.1.1) is of a specified type. Incidentally, the real bilinear model (2.1.2) can be put in the vector form (3.1.1) with the matirix A being of this special type.

3.2 FIRST AND SECOND ORDER MOMENTS OF BILINEAR PROCESSES

Let A, B_1, B_2, ... , B_q be q + 1 matrices each of order pxp and $\underline{b}_1$, $\underline{b}_2$, ... , $\underline{b}_q$ and C be q + 1 column vectors each of p-components. Let e_t, $t \in Z$ be a sequence of independent identically distributed random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let Γ be the matrix given in Theorem 2.4.3 built on A, B_1, B_2, ... B_q and σ^2. If $\rho(\Gamma) < 1$, by Theorem 2.4.3, there exists a strictly stationary vector-valued process $\underline{X}_t$, $t \in Z$ satisfying

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j e_{t-j} + \sum_{j=1}^{q} B_j\,\underline{X}_{t-j} e_{t-j} + Ce_t \quad \text{a.e } [P] \qquad (3.2.1)$$

for every t in Z.

It is natural to enquire whether $E(\underline{X}_t)$ and Disp ($\underline{X}_t$) exist. Strict stationarity does not guarantee existence of moments. If $E(\underline{X}_t)$ and Disp($\underline{X}_t$) exist, we find a way to calculate these moments. This section is devoted to a study of this problem.

The question of existence of $E(\underline{X}_t)$ and Disp($\underline{X}_t$) can easily be settled in the case q = 1 because in this case $\underline{X}_t$ admits an infinite series representation in terms of $e_t, e_{t-1}, \ldots$.

<u>Theorem 3.2.1</u> Let A and B be two square matrices each of order pxp. Let C and $\underline{b}$ be two column vectors each of p-components. Let e_t, $t \in Z$ be a sequence of independent identically distributed random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let

$$\rho(A \otimes A + \sigma^2 B \otimes B) < 1.$$

Then the following are valid.

(i) $\rho(A) < 1$

(ii) For the bilinear process $\underline{X}_t$, $t \in Z$ satisfying

$$\underline{X}_t = A\,\underline{X}_{t-1} + \underline{b}e_{t-1} + B\,\underline{X}_{t-1}e_{t-1} + Ce_t \qquad \text{a.e } [P]$$

for every $t \in Z$, $E(\underline{X}_t)$ exists.

(iii) $E(\underline{X}_t) = \sigma^2(I_p - A)^{-1}\, B\, C$ (3.2.2)

(iv) If $E(e_t^4) < \infty$, then $\text{Disp}(\underline{X}_t)$ exists.

(v) Further, if $V = E(\underline{X}_t\, \underline{X}_t^T)$, then V satisfies

$$\underset{pxp}{V} = A\,V\,A^T + \sigma^2 B\,V\,B^T + \Delta \qquad (3.2.3)$$

for some constant matrix Δ given by

$$\underset{pxp}{\Delta} = A\,S\,B^T + B\,S\,A^T + \sigma^2(C\,C^T + \underline{b}\,\underline{b}^T + B\,\underline{\mu}\,\underline{b}^T + \underline{b}\,\underline{\mu}^T\,B^T$$
$$+ A\,C\,\underline{b}^T + \underline{b}\,C^T\,A^T + 2\sigma^2 B\,C\,C^T\,B^T) + B\,H\,B^T$$
$$+ k_3(B\,C\,\underline{b}^T + \underline{b}\,C^T\,B^T)$$

and

$$\underset{pxp}{S} = \sigma^2(A\,\underline{\mu}\,C^T + C\,\underline{\mu}^T\,A^T + \sigma^2 B\,C\,C^T + \sigma^2\,C\,C^T\,B^T) + k_3\,C\,C^T$$

$$\underset{pxp}{H} = k_3(A\,\underline{\mu}\,C^T + C\,\underline{\mu}^T A^T + \sigma^2 B\,C\,C^T + \sigma^2\,C\,C^T\,B^T) + k_4\,C\,C^T$$

where k_3 and k_4 are the third and fourth-order cumulants respectively of e_t, $t \in Z$, ie,

$$E(e_t^3) = k_3$$
$$E(e_t^4) = 3\sigma^4 + k_4$$

<u>Proof</u> By Theorem 2.3.2.

$$\underline{X}_t = Ce_t + \underline{b}e_{t-1} + \sum_{r\geq 1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1}) \quad \text{a.e } [P]$$

for every t in Z. The above series converges absolutely almost surely [P] as well as in the mean. Since

$$\sum_{r\geq 1} E\left|\left(\prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})\right)_s\right| < \infty$$

for every s = 1, 2, ..., p, $E(\underline{X}_t)$ exists and is given by

$$E(\underline{X}_t) = 0 + 0 + E \sum_{r\geq 1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})$$

$$= \sum_{r\geq 1} E \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})$$

(See Chung [10, (xi), p.42].)

$$= \sum_{r\geq 1} \left(\prod_{j=1}^{r-1} E(A + Be_{t-j})\right) E((A + Be_{t-r})Ce_{t-r})$$

$$+ \sum_{r\geq 1} \left(\prod_{j=1}^{r-1} E(A + Be_{t-j})\right) E((A + Be_{t-r})\underline{b}e_{t-r-1}),$$

because e_t, $t \in Z$ are independent

$$= \sum_{r\geq 1} \sigma^2 A^{r-1} B\, C + 0$$

because e_t's are identically distributed

$$= \sigma^2 \left(\sum_{r\geq 1} A^{r-1}\right) B C$$

Since $E(\underline{X}_t)$ exists $\sum_{r\geq 1} A^{r-1}$ must be convergent.

This happens if and only if $\rho(A) < 1$, if and only if $(I - A)$ is invertible.

In that case,

$$\sum_{r\geq 1} A^{r-1} = (I_p - A)^{-1}.$$

This proves (i), (ii), and (iii).

If $E(e_t^4) < \infty$, one can show that

$$\sum_{r \geq 1} E \left| \left(\prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1}) \right)_s \right|^2 < \infty$$

for every s = 1, 2, ... , p. Consequently, the series

$$\sum_{r \geq 1} \prod_{j=1}^{r} (A + Be_{t-j})(Ce_{t-r} + \underline{b}e_{t-r-1})$$

is convergent in the quadratic mean. From this, it follows that Disp($\underline{X}_t$) exists. This proves (iv).

We will establish (v) as part of a more general result, namely Theorem 3.2.4.

The above proof gives the following corollary.

Corollary 3.2.2. Let A and B be two matrices each of order pxp. If $\rho(A \otimes A + B \otimes B) < 1$, then $\rho(A \otimes A) < 1$ and hence $\rho(A) < 1$.

See Lemma 2.2.2(d). We are unable to establish the above result directly. Existence Theorem 3.2.1 gives a result on matrix algebra! .

Lemma 3.2.3. Let $F = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ be a partitioned matrix in which A and B are square matrices. If $\rho(F) < 1$, then $\rho(A) < 1$.

Proof. Let A and D be of orders pxp and qxq respectively. Let λ be an eigen value of A. We show that λ is an eigen value of F. Note that

$$|F - \lambda I_{p+q}| = \begin{vmatrix} A - \lambda I_p & B \\ C & D - \lambda I_q \end{vmatrix}$$

$$= |A - \lambda I_p| \, |(D - \lambda I_q) - C(A - \lambda I_p)^{-1} B|$$

(See Morrison [28, p.68].)

$$= 0$$

Since $\rho(F) < 1$, $|\lambda| < 1$. Hence $\rho(A) < 1$.

The following result is the main result of this section.

Theorem 3.2.4. Let A, B_1, B_2, ... , B_q be q + 1 matrices each of order pxp and C, $\underline{b}_1$, $\underline{b}_2$, ... $\underline{b}_q$ be q + 1 vectors each of order px1. Let e_t, $t \in Z$ be a sequence of independent identically distributed random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let Γ be the matrix of Theorem 2.4.3 built on $A, B_1, B_2, \ldots, B_q$ and σ^2. Then the following statements are true.

(i) $\rho(A) < 1$

(ii) For the strictly stationary process $\underline{X}_t$, $t \in Z$ satisfying

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j e_{t-j} + \sum_{j=1}^{q} B_j\, \underline{X}_{t-j} e_{t-j} + Ce_t \quad \text{a.e } [P]$$

for every t in Z, $E(\underline{X}_t)$ exists.

(iii) $$E(\underline{X}_t) = \sigma^2 (I_p - A)^{-1} \sum_{j=1}^{q} B_j\, C\,. \tag{3.2.4}$$

(iv) If $E(e_t^4) < \infty$, then $\text{Disp}(\underline{X}_t)$ exists.

(v) Further, if $V = E(\underline{X}_t\, \underline{X}_t^T)$, then V satisfies

$$\begin{aligned}
\underset{p\times p}{V} &= \{A\,V\,A^T + \sigma^2 B_1\, V\, B_1^T\} + \sigma^2\{(B_2 + A\,B_1)V\,B_2^T + B_2\,V(A\,B_1)^T\} \\
&+ \sigma^2\{(B_3 + A\,B_2 + A^2\,B_1)V\,B_3^T + B_3\,V(A\,B_2 + A^2\,B_1)^T\} \\
&+ \sigma^2\{(B_4 + A\,B_3 + A^2\,B_2 + A^3\,B_1)V\,B_4^T + B_4 V(A\,B_3 + A^2\,B_2 \\
&+ A^3\,B_1)^T\} + \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&+ \sigma^2\{(B_q + A\,B_{q-1} + A^2\,B_{q-2} + \ldots\ldots + B^{q-1}\,B_1)V\,B_q^T \\
&+ B_q\,V(A\,B_{q-1} + A^2\,B_{q-2} + \ldots\ldots + A^{q-1}\,B_1)^T\} + \Delta_1 \\
&+ A\,\Delta_2 + (A\,\Delta_2)^T
\end{aligned} \tag{3.2.5}$$

for some constant matrices Δ_1 and Δ_2, where

$$\begin{aligned}
\underset{p\times p}{\Delta_1} &= \sigma^2\, C\, C^T + \sigma^2 \sum_{j=1}^{q} \underline{b}_j\, \underline{b}_j^T + \sigma^2 \sum_{j=1}^{q} B_j\, \underline{\mu}\, \underline{b}_j^T + \sigma^2 \sum_{j=1}^{q} \underline{b}_j\, \underline{\mu}^T\, B_j^T \\
&+ \sigma^4 \sum_{j=1}^{q} B_j\, C\, C^T\, B_j^T + \sigma^4 \sum_{i=1}^{q} \sum_{j=1}^{q} B_i\, C\, C^T\, B_j^T + \sum_{j=1}^{q} A\, \underline{W}(j-1)\underline{b}_j^T \\
&+ \sum_{j=1}^{q} \underline{b}_j\, \underline{W}^T(j-1)A^T + \sum_{j=1}^{q} B_j\, H\, B_j^T + k_3 \sum_{j=1}^{q} B_j\, C\, \underline{b}_j^T + k_3 \sum_{j=1}^{q} \underline{b}_j\, C^T\, B_j^T
\end{aligned}$$

$$\underset{p\times p}{\Delta_2} = \sum_{j=1}^{q} A^{j-1}\, S\, B_j^T$$

$$+ \sigma^2 \sum_{j=2}^{q} (\underline{b}_{j-1} + A\, \underline{b}_{j-2} + A^2\, \underline{b}_{j-3} + \ldots + A^{j-2}\, \underline{b}_1)\underline{\mu}^T B_j^T$$

$$+ \sigma^4 \sum_{j=2}^{q} \left[(B + B_{j-1}) + A(B + B_{j-2}) + A^2(B + B_{j-3})\right.$$

$$\left. + \ldots\ldots + A^{j-2}\,(B + B_1)\right] C\, C^T\, B_j^T$$

$$+ k_3\, \sigma^2 \sum_{j=2}^{q} (\underline{b}_{j-1} + A\, \underline{b}_{j-2} + A^2\, \underline{b}_{j-3} + \ldots\ldots + A^{j-2}\, \underline{b}_1) C^T\, B_j^T$$

$$+ \sum_{j=2}^{q} (B_{j-1} + A\, B_{j-2} + A^2\, B_{j-3} + \ldots\ldots + A^{j-2}\, B_1) H\, B_j^T ,$$

and

$$\underset{p\times p}{S} = \sigma^2 (A\, \underline{\mu}\, C^T + C\, \underline{\mu}^T A^T + \sigma^2\, B\, C\, C^T + \sigma^2\, C\, C^T\, B^T)$$
$$+ k_3\, C\, C^T$$

$$\underset{p\times p}{H} = k_3 (A\, \underline{\mu}\, C^T + C\, \underline{\mu}^T A^T + \sigma^2\, B\, C\, C^T + \sigma^2\, C\, C^T\, B^T)$$
$$+ k_4\, C\, C^T$$

$$\underset{p\times 1}{\underline{W}(k)} = \sigma^2\, C\ , \text{ if } k = 0$$

$$= \sigma^2\, A^k\, C + \sigma^2 \sum_{j=1}^{k} A^{k-j} (\underline{b}_j + B_j\, \underline{\mu})$$

$$+ k_3 \left\{ \sum_{j=1}^{k} A^{k-j}\, B_j \right\} C, \text{ if } k > 0$$

$$\underset{p\times p}{B} = \sum_{j=1}^{q} B_j$$

where k_3 and k_4 are the third and fourth-order cumulants respectively of e_t, $t \in Z$, ie

$$E(e_t^3) = k_3$$

$$E(e_t^4) = 3\sigma^4 + k_4$$

<u>Proof</u> (i) Since $\rho(\Gamma) < 1$, by Lemma 3.2.3

$$\rho(A \otimes A + \sigma^2\, B_1 \otimes B_1) < 1.$$

By Corollary 3.2.2,

$$\rho(A) < 1$$

(ii) Recall from the proof of Theorem 2.4.3, $\underline{X}_t$, $t \in Z$ is obtained as an almost sure limit of a sequence $\underline{S}_{n,t}$, $t \in Z$, $n \geq 1$ of strictly stationary processes. See steps 1^0. and 2^0. of the proof of Theorem 2.4.3. In fact, we now show that the sequence $\underline{S}_{n,t}$, $t \in Z$, $n \geq 1$ of random vectors indeed converges to $\underline{X}_t$ in the mean for every t in Z.

Recall from step 2^0. of Theorem 2.4.3, that we have proved that

$$E|(\underline{S}_{n,t} - \underline{S}_{n-1,t})_i| \leq k\, \lambda^{\frac{n}{2}}$$

for every $n \geq 1$, for $i = 1, 2, \ldots, p$ and for some constant k, where $\lambda = \rho(\Gamma) < 1$. This means that the sequence $\underline{S}_{n,t}$, $n \geq 1$ is Cauchy in the mean. Consequently, there exists an integrable random vector $\underline{Y}_t$, $t \in Z$ such that $\underline{S}_{n,t}$, $n \geq 1$ converges to $\underline{Y}_t$ in the mean. See Theorem B of Halmos [21, p.107]. This $\underline{Y}_t$ is almost surely equal to our $\underline{X}_t$ above. Thus we have proved that $E(\underline{X}_t)$ exists and equal to $\lim_{n\to\infty} E(\underline{S}_{n,t})$. See Chung [10, Theorem 4.5.4, p.97].

(iii) Since $\underline{X}_t$, $t \in Z$ satisfies

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j e_{t-j} + \sum_{j=1}^{q} B_j\,\underline{X}_{t-j}\, e_{t-j} + Ce_t \quad \text{a.e } [P] \qquad (3.2.6)$$

for every t in Z and $E(\underline{X}_t)$ exists, we can take expectations both sides of the above equality after multiplying on either side by e_t. We obtain

$$E(\underline{X}_t e_t) = \sigma^2\, C$$

Now, we take expectations on both sides of the above equality (3.2.6).

Let

$$\underline{\mu} = E(\underline{X}_t)$$

Then

$$\underline{\mu} = A\,\underline{\mu} + \sigma^2 \sum_{j=1}^{q} B_j\, C$$

This means that,

$$(I_p - A)\underline{\mu} = \sigma^2 \sum_{j=1}^{q} B_j C$$

Since $\rho(A) < 1$.

$$\underline{\mu} = \sigma^2 (I_p - A)^{-1} \left[\sum_{j=1}^{q} B_j C \right] . \qquad (3.2.7)$$

(iv) If $E(e_t^4) < \infty$, one can show that

$$E|(\underline{S}_{n,t} - \underline{S}_{n-1,t})_i|^2 \leqslant k \lambda^n$$

for every $n \geqslant 1$, for $i = 1, 2, \ldots, p$ and for some constant $k > 0$ in steps 2^0. and 3^0. in the proof of Theorem 2.4.3.

As in Lemma 2.4.1, we can show that $\underline{S}_{n,t}$, $n \geqslant 1$ is a Cauchy sequence in the quadratic mean. Consequently, $\underline{S}_{n,t}$, $n \geqslant 1$ indeed converges to $\underline{X}_t$ in the quadratic mean. See Halmos [21, Theorem B, p.107]. Hence $\text{Disp}(\underline{X}_t)$ exists and in fact

$$\text{Disp}(\underline{X}_t) = \lim_{n\to\infty} \text{Disp}(\underline{S}_{n,t})$$

See Chung [10, Theorem 4.5.4, p.97].

Now we proceed to obtain $V = E(\underline{X}_t \underline{X}_t^T)$.

(v) Since $\underline{X}_t$, $t \in Z$ satisfies (3.2.7) and $\text{Disp}(\underline{X}_t)$ exists, then $E(\underline{X}_t \underline{X}_t^T)$ exists.

Since $\underline{X}_t$, $t \in Z$ satisfying (3.2.6) is first-order stationary, we obtain the following

$$E(\underline{X}_{t-i} e_{t-i} e_{t-j})$$

$$\left.\begin{aligned} &= \sigma^2 \underline{\mu} + k_3 C , \text{ if } i = j \\ &= \underline{0} \qquad\qquad\quad , \text{ if } i \neq j \end{aligned}\right\} \qquad (3.2.8)$$

and

$$\underset{p\times 1}{\underline{W}(k)} = E(\underline{X}_t e_{t-k})$$

$$= \sigma^2 C , \text{ if } k = 0$$

$$= \sigma^2 A^k C + \sigma^2 \sum_{j=1}^{k} A^{k-j}(\underline{b}_j + B_j \underline{\mu})$$

$$+ k_3 \left[\sum_{j=1}^{k} A^{k-j} B_j \right] C \text{ , if } k > 0 \qquad (3.2.9)$$

Let $E(e_t^3) = k_3$ and $E(e_t^4) = 3\sigma^4 + k_4 < \infty$, where k_3 and k_4 are the third and fourth-order cumulants respectively of the process e_t, $t \in Z$. Let

$$\underset{p\times p}{V} = E(\underline{X}_t \underline{X}_t^T)$$

$$\underset{p\times p}{S(k)} = E(\underline{X}_t \underline{X}_{t-k}^T e_{t-k}) \text{ , } k \geq 0$$

Note that

$$\underline{X}_t \underline{X}_t^T$$

$$= A \underline{X}_{t-1} \underline{X}_{t-1}^T A^T + \sum_{j=1}^{q} A \underline{X}_{t-1} \underline{X}_{t-j}^T e_{t-j} B_j^T + \sum_{j=1}^{q} B_j e_{t-j} \underline{X}_{t-j} \underline{X}_{t-1}^T A^T$$

$$+ \sum_{j=1}^{q} A \underline{X}_{t-1} e_{t-j} \underline{b}_j^T + \sum_{j=1}^{q} \underline{b}_j e_{t-j} \underline{X}_{t-1}^T A^T + \sum_{i=1}^{q} \sum_{j=1}^{q} B_i \underline{X}_{t-i} \underline{X}_{t-j}^T e_{t-i} e_{t-j} B_j^T$$

$$+ \sum_{i=1}^{q} \sum_{j=1}^{q} \underline{b}_i e_{t-i} e_{t-j} \underline{X}_{t-j}^T B_j^T + \sum_{i=1}^{q} \sum_{j=1}^{q} B_i \underline{X}_{t-i} e_{t-i} e_{t-j} \underline{b}_j^T$$

$$+ \sum_{i=1}^{q} \sum_{j=1}^{q} \underline{b}_i e_{t-i} e_{t-j} \underline{b}_j^T + \sum_{j=1}^{q} B_j \underline{X}_{t-j} e_{t-j} e_t C^T$$

$$+ \sum_{j=1}^{q} C e_t e_{t-j} \underline{X}_{t-j}^T B_j^T + \sum_{j=1}^{q} \underline{b}_j e_{t-i} e_t C^T + \sum_{j=1}^{q} C e_t e_{t-j} \underline{b}_j^T$$

$$+ A \underline{X}_{t-1} e_t C^T + C e_t \underline{X}_{t-1}^T A^T + e_t^2 C C^T \qquad (3.2.10)$$

Since $\underline{X}_t$, $t \in Z$ satisfying (3.2.6) is second-order stationary, we obtain the following

$$E(\underline{X}_{t-i} \underline{X}_{t-j}^T e_{t-i} e_{t-j})$$

$$= \sigma^4 C C^T , \qquad \text{if } i \neq j$$

$$= \sigma^2 V + 2\sigma^4 C C^T + H , \text{ if } i = j \qquad (3.2.11)$$

where

$$H = k_3(A \underline{\mu} C^T + C \underline{\mu}^T A^T + \sigma^2 B C C^T + \sigma^2 C C^T B^T) + k_4 C C^T$$

Making use of (3.2.8), (3.2.9), (3.2.11) and the fact that the random variables e_t, $t \in Z$ are independent and identically distributed, we take expectations on both sides of (3.2.10) to obtain

$$V = A V A^T + \sigma^2 \sum_{j=1}^{q} B_j V B_j^T + \sum_{j=1}^{q} A S(j-1) B_j^T$$

$$+ \sum_{j=1}^{q} B_j S^T(j-1) A^T + \Delta_1 \qquad (3.2.12)$$

where Δ_1 is the matrix of Theorem 3.2.4.

Let us now consider $S(k)$, $k \geq 0$. For $k = 0$, we obtain

$$S(0) = S$$

$$= \sigma^2(A \underline{\mu} C^T + C \underline{\mu}^T A^T + \sigma^2 B C C^T + \sigma^2 C C^T B^T) + k_3 C C^T \qquad (3.2.13)$$

For $k \neq 0$, we obtain

$$S(k) = A S(k-1) + \sigma^2 B_k V + \sigma^4(B + B_k) C C^T + \sigma^2 \underline{b}_k \underline{\mu}^T$$

$$+ k_3 \underline{b}_k C^T + B_k H \qquad (3.2.14)$$

By successive substitution for the quantity $S(k)$ in (3.2.14), we obtain

$$S(k) = A^k S + \sigma^2(B_k + A B_{k-1} + A^2 B_{k-2} + \ldots\ldots + A^{k-1} B_1)V$$

$$+ \sigma^4\Big[(B + B_k) + A(B + B_{k-1}) + A^2(B + B_{k-2}) + \ldots\ldots$$

$$+ A^{k-1}(B + B_1)\Big] C C^T + \sigma^2(\underline{b}_k + A \underline{b}_{k-1} + A^2 \underline{b}_{k-2} + \ldots\ldots$$

$$+ A^{k-1} \underline{b}_1)\underline{\mu}^T + k_3(\underline{b}_k + A \underline{b}_{k-1} + A^2 \underline{b}_{k-2} + \ldots\ldots + A^{k-1} \underline{b}_1) C^T$$

$$+ (B_k + A B_{k-1} + A^2 B_{k-2} + \ldots\ldots + A^{k-1} B_1) H \qquad (3.2.15)$$

Now,

$$\sum_{j=1}^{q} S(j-1)B_j^T = \sigma^2 B_1 V B_2^T + \sigma^2(B_2 + A B_1)V B_3^T$$

$$+ \sigma^2(B_3 + A B_2 + A^2 B_1)V B_4^T + \ldots..$$

$$+ \sigma^2(B_{q-1} + A B_{q-2} + A^2 B_{q-3} + \ldots..$$

$$+ A^{q-2} B_1)V B_q^T + \Delta_2 \tag{3.2.16}$$

where Δ_2 is the constant matrix of Theorem 3.2.4.

If we now substitute (3.2.16) into (3.2.12) we obtain the expression (3.2.5).

This completes the proof.

REMARKS 3.2.5

(1) The equation (3.2.5) is linear in V and can be solved explicitly once all the matrices involved are given explicitly.

(2) If $E(e_t^4) < \infty$, and $a^2 + \sigma^2 \beta^2 < 1$, then there exists a strictly stationary and second-order process X_t, $t \in Z$ satisfying

$$X_t = a X_{t-1} + be_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P]$$

for every t in Z with

$$E(X_t) = \mu = \sigma^2 \beta/(1 - a)$$

and

$$E(X_t^2) = \frac{\sigma^2}{1 - a^2 - \beta^2 \sigma^2}\left[1 + b^2 + 2 a b + 2 \beta \mu(1 + a + b)\right]$$

when e_t, $t \in Z$ is a Gaussian process.

(3) If $E(e_t^4) < \infty$ and $\rho(\Gamma) < 1$, where Γ is the matrix of Theorems 2.1.1 and 2.4.3, then there exists a strictly stationary second-order process X_t, $t \in Z$ satisfying

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + \sum_{\substack{i=1 \\ i \geq j}}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-i} e_{t-j} + e_t \qquad \text{a.e } [P]$$

for every t in Z.

The mean of X_t, $t \varepsilon Z$ is given by

$$\mu = E(X_t) = \frac{\sigma^2 \sum_{j=1}^{q'} \beta_{jj}}{1 - \sum_{j=1}^{r} a_j} \tag{3.2.17}$$

where $q' = \min(m, \ell)$. From (3.2.17) one notices that $E(X_t) = 0$ if and only if

$$\sum_{j=1}^{q'} \beta_{jj} = 0$$

3.3 COVARIANCE STRUCTURES OF BILINEAR AND LINEAR PROCESSES

In this section, we show that for every bilinear process (2.1.2), there exists an ARMA process with identical covariance structures. This result comes as a special case of a corresponding result for vector-valued processes satisfying (3.1.1) with the matrix A having some specified structure.

Theorem 3.3.1. Let e_t, $t \varepsilon Z$ be a sequence of independent identically distributed real random variables with $E(e_t) = 0$, $E(e_t^2) = \sigma^2$ and $E(e_t^4) < \infty$. Let A, B_1, B_2, ... , B_q be q + 1 matrices each of order pxp. Let C, $\underline{b}_1$, $\underline{b}_2$, ... , $\underline{b}_q$ be q + 1 column vectors each of order px1. Let $\rho(\Gamma) < 1$, where Γ is the matrix of Theorem 2.4.3 built on A, B_1, B_2, ... , B_q and σ^2.

For the bilinear strictly stationary second-order vector-valued process $\underline{X}_t$, $t \varepsilon Z$ conforming to the model

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j e_{t-j} + \sum_{j=1}^{q} B_j\,\underline{X}_{t-j} e_{t-j} + Ce_t \quad \text{a.e } [P] \tag{3.3.1}$$

for every t in Z with the matrix A given by

$$\underset{p\times p}{A} = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_{p-1} & a_p \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \qquad (3.3.2)$$

there exists an ARMA process of order (p,q) with autoregressive coefficients $a_1, a_2, \dots, a_p$ and moving average coefficients being functions of $A, B_1, B_2, \dots, B_q, \underline{b}_1, \underline{b}_2, \dots, \underline{b}_q$ and C such that they have identical covariance structures.

<u>Proof</u> From (3.3.1), it is easy to show that

$$E(\underline{X}_{t+k} e_{t+k} \underline{X}_t^T) = \begin{cases} S & \text{if } k = 0 \\ \sigma^2 C \underline{\mu}^T, & \text{if } k \neq 0 \end{cases} \qquad (3.3.3)$$

where $\underline{\mu} = E(\underline{X}_t)$ and S is the pxp matrix given by the expression (3.2.13).

Using (3.3.3), we obtain the following

$$E(\underline{X}_{t+1} \underline{X}_t^T) = A V + \sum_{j=1}^{q} B_j S^T(j-1) + \sum_{j=1}^{q} \underline{b}_j \underline{W}^T(j-1),$$

$$E(\underline{X}_{t+k} \underline{X}_t^T) = A^k V + A^{k-1} \sum_{j=1}^{q} B_j S^T(j-1) + A^{k-2} \sum_{j=2}^{q} B_j S^T(j-2)$$

$$+ \dots + A \sum_{j=k-1}^{q} B_j S^T(j-k+1) + \sum_{j=k}^{q} B_j S^T(j-k)$$

$$+ A^{k-1} \sum_{j=1}^{q} \underline{b}_j \underline{W}^T(j-1) + A^{k-2} \sum_{j=1}^{q} \underline{b}_j \underline{W}^T(j-2)$$

$$+ \dots + A \sum_{j=k-1}^{q} \underline{b}_j \underline{W}^T(j-k+1) + \sum_{j=k}^{q} \underline{b}_j \underline{W}^T(j-k)$$

$$+ \sigma^2 \Big[A^{k-2} B_1 + A^{k-3}(B_1 + B_2) + \dots$$

$$+ A(B_1 + B_2 + \dots + B_{k-2}) + (B_1 + B_2 + \dots$$

$$+ B_{k-1})] C \underline{\mu}^T ,$$

$$\text{if } k = 2, 3, \dots , q \qquad (3.3.4)$$

and

$$E(\underline{X}_{t+k}\, \underline{X}_t^T) = A\, E(\underline{X}_{t+k-1}\, \underline{X}_t^T) + \sigma^2 B\, C\, \underline{\mu}^T ,$$

$$\text{if } k > q \qquad (3.3.5)$$

where $V = E(\underline{X}_t\, \underline{X}_t^T)$, $\underline{W}(k)$, $k \geq 0$ is the px1 matrix given by the expression (3.2.9) and $S(k)$, $k \geq 0$ is the pxp matrix given by the expression (3.2.15).

If we now let

$$C(k) = E\{(\underline{X}_{t+k} - \underline{\mu})(\underline{X}_t - \underline{\mu})^T\}$$

$$= E(\underline{X}_{t+k}\, \underline{X}_t^T) - \underline{\mu}\, \underline{\mu}^T ,$$

and noting that

$$\underline{\mu}\, \underline{\mu}^T = A\, \underline{\mu}\, \underline{\mu}^T + \sigma^2 B\, C\, \underline{\mu}^T ,$$

we obtain from (3.3.5)

$$C(k) = A\, C(k-1) , \; k > q \qquad (3.3.6)$$

Since

$$\underset{p\times p}{C(k)} = \begin{bmatrix} R(k) & R(k+1) & R(k+2) & \dots & R(k+p-1) \\ R(k-1) & R(k) & R(k+1) & \dots & R(k+p-2) \\ R(k-2) & R(k-1) & R(k) & \dots & R(k+p-3) \\ \vdots & \vdots & \vdots & & \vdots \\ R(k-p+1) & R(k-p+2) & R(k-p+3) & \dots & R(k) \end{bmatrix} \qquad (3.3.7)$$

where $R(k) = E\{(X_t - \mu)(X_{t+k} - \mu)\}$,

$\mu = E(X_t)$,

expression (3.3.6) is equvalent to

$$R(k) = a_1 R(k-1) + a_2 R(k-2) + \dots + a_p R(k-p) , \; k > q \qquad (3.3.8)$$

provided the constant matrix A is given by (3.3.2). These equations (3.3.8) are the same as the Yule-Walker equations for an ARMA (p,q) and thus show that the process $\underline{X}_t$, $t \in Z$ conforming to the bilinear model (3.3.1) with the matrix A defined by (3.3.2) has identical covariance structure as some ARMA (p,q) process.

Corollary 3.3.2 Let e_t, $t \in Z$ be a sequence of independent identically distributed real random variables with $E(e_t) = 0$, $E(e_t^2) = \sigma^2$ and $E(e_t^4) < \infty$. Let A, $B_1, B_2, \ldots, B_q$ be $q+1$ matrices of Theorem 2.1.1. Let C, $\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_q$ be $q+1$ column vectors of Theorem 2.1.1. Let $\rho(\Gamma) < 1$, where Γ is the matrix of Theorem 2.4.3 built on A, $B_1, B_2, \ldots, B_q$ and σ^2. For the bilinear strictly stationary second-order process X_t, $t \in Z$ conforming to the model

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + \sum_{\substack{i=1 \\ i \geq j}}^{m} \sum_{j=2}^{\ell} \beta_{ij} X_{t-i} e_{t-j} + e_t \qquad \text{a.e } [P] \quad (3.3.9)$$

for every t in Z, there exists an ARMA (r, max (h, g)), g = min (m, ℓ) with autoregressive coefficients $a_1, a_2, \ldots, a_r$ and moving average coefficients being functions of $a_1, a_2, \ldots, a_r, b_1, b_2, \ldots, b_h$ and β_{ij}, $1 \leq i \leq m$, $1 \leq j \leq \ell$, $i \geq j$ such that they have identical covariance structures.

Proof. Replace the matrix A of Theorem 3.3.1 with the matrix A of Theorem 2.1.1. The result follows from Theorem 3.3.1.

We have said in section 2.1.1 of chapter two that the study of bilinear models subsumes the study of ARMA models. We now obtain the second-order moments and autocovariances of some linear models from those of bilinear models

Corollary 3.3.3. Let e_t, $t \in Z$ be a sequence of independent identically distributed real random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$.

Then there exists a strictly stationary second-order process X_t, $t \in Z$ conforming to the linear model

$$X_t = \sum_{j=1}^{h} b_j e_{t-j} + e_t \qquad \text{a.e } [P] \quad (3.3.10)$$

for every t in Z.

Further more,

$$E(X_t) = 0$$

and the autocovariance function of X_t, $t \in Z$ is given by

$$\begin{aligned} R(k) &= \sigma^2 \sum_{j=0}^{h-k} b_j b_{j+k} &&, \text{ if } k = 0, 1, 2, \dots, h \\ &= 0 &&, \text{ if } k > h \\ &= R(-k) &&, \text{ if } k < 0 \end{aligned}$$

where $R(k) = E(X_t X_{t+k}) = R(-k)$ and $b_o = 1$.

Proof. The model (3.3.10) can be put in the vector form as follows. Let

$$p = q = h$$

$$\underset{p\times p}{A} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \qquad (3.3.11)$$

$$\underset{1\times p}{\underline{b}_j^T} = (b_j, 0, 0, \dots, 0), \; j = 1, 2, \dots, q$$

$$\underset{1\times p}{C^T} = (1, 0, 0, \dots, 0)$$

$$\underset{1\times p}{\underline{X}_t^T} = (X_t, X_{t-1}, \dots, X_{t-p+1}), \; t \in Z$$

Then

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j e_{t-j} + Ce_t \qquad \text{a.e } [P] \quad (3.3.12)$$

for every t in Z.

A sufficient condition for the existence of a strictly stationary second-order vector-valued process $\underline{X}_t$, $t \in Z$ satisfying (3.3.12) would be

$$\rho(A \otimes A) < 1 ,$$

which is $\rho(A) < 1$, by Lemma 2.2.2 (d). For the matrix A given by (3.3.11). $\rho(A) < 1$ is always satisfied because $\rho(A) = 0$. Hence no restrictions on the b_j's are required for a process X_t, $t \in Z$ satisfying (3.3.10) to be stationary.

From Theorem 3.2.4, we obtain the following.

$$\underline{\mu} = E(\underline{X}_t) = 0$$

$$V = E(\underline{X}_t \, \underline{X}_t^T) = \mathrm{Disp}(\underline{X}_t)$$

$$= A\ V\ A^T + \sigma^2\ C\ C^T + \sigma^2 \sum_{j=1}^{q} \underline{b}_j\ \underline{b}_j^T + \sum_{j=1}^{q} A\ \underline{W}(j-1)\underline{b}_j^T$$

$$+ \sum_{j=1}^{q} \underline{b}_j\ \underline{W}^T(j-1)A^T \qquad (3.3.13)$$

where

$$W(k) = \sigma^2\ C \quad , \text{ if } k = 0$$

$$= \sigma^2\ A^k\ C + \sigma^2 \sum_{j=1}^{k} A^{k-j}\ \underline{b}_j \quad , \text{ if } k > 0$$

It is not difficult to check that

$$\sum_{j=1}^{q} \underline{W}(j-1)\underline{b}_j^T = \sigma^2 \begin{pmatrix} \sum_{j=0}^{q-1} b_j\ b_{j+1} & 0 & \cdots & 0 \\ \sum_{j=0}^{q-2} b_j\ b_{j+2} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \sum_{j=0}^{q-q} b_j\ b_{j+q} & 0 & \cdots & 0 \end{pmatrix} \qquad (3.3.14)$$

where $b_0 = 1$.

Using (3.3.14), we obtain from (3.3.13) that

$$\begin{bmatrix} R(0) & R(1) & R(2) & \cdots & R(q-1) \\ R(1) & R(0) & R(1) & \cdots & R(q-2) \\ R(2) & R(1) & R(0) & \cdots & R(q-3) \\ \vdots & \vdots & \vdots & & \vdots \\ R(q-1) & R(q-2) & R(q-3) & \cdots & R(0) \end{bmatrix} = V$$

$$= \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & R(0) & R(1) & \cdots & R(q-2) \\ 0 & R(1) & R(0) & \cdots & R(q-3) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & R(q-2) & R(q-3) & \cdots & R(0) \end{bmatrix}$$

$$+ \sigma^2 \begin{bmatrix} \sum_{j=0}^{q} b_j^2 & \sum_{j=0}^{q-1} b_j\, b_{j+1} & \cdots & \sum_{j=0}^{1} b_j\, b_{j+q-1} \\ \sum_{j=0}^{q-1} b_j\, b_{j+1} & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \sum_{j=0}^{1} b_j\, b_{j+q-1} & 0 & \cdots & 0 \end{bmatrix}$$

This establishes the result for $R(k)$, $k = 1, 2, \ldots, q-1$.

From the proof of Theorem 3.3.1, we obtain

$$C(1) = A\,V + \sum_{j=1}^{q} \underline{b}_j\, \underline{W}^T(j-1) \tag{3.3.15}$$

and

$$C(k) = A\,C(k-1)\,,\ k > q \tag{3.3.16}$$

Expanding (3.3.15) we obtain

$$R(q) = \sigma^2 b_q$$

and from (3.3.16), we obtain

$$R(k) = 0, \; k > q.$$

This completes the proof.

Corollary 3.3.4. Let e_t, $t \in Z$ be a sequence of independent identically distributed real random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let the roots of the polynomial

$$f(x) = 1 - a_1 x - a_2 x^2 - \dots - a_r x^r$$

be greater than unity in absolute value. The strictly stationary second-order process X_t, $t \in Z$ conforming to the model

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + e_t \qquad \text{a.e } [P] \quad (3.3.17)$$

for every t in Z has mean zero and autocovariance function given by

$$R(0) = a_1 R(1) + a_2 R(2) + \dots + a_r R(r) + \sigma^2 \qquad (3.3.18)$$

and

$$R(k) = a_1 R(k-1) + a_2 R(k-2) + \dots + a_r R(k-r), \; k > 0 \qquad (3.3.19)$$

where $R(k) = E(X_t X_{t+k}) = R(-k)$

Proof. By Corollary 2.3.8, the strictly stationary second-order process X_t, $t \in Z$ satisfying (2.2.17) admits the vector representation

$$\underline{X}_t = A \underline{X}_{t-1} + Ce_t \qquad \text{a.e } [P]$$

for every t in Z with the matrix A being of the special type (3.3.2). From Theorems 3.2.1 and 3.2.4, we obtain

$$\underline{\mu} = E(\underline{X}_t) = 0$$

$$C(0) = V$$

$$= E(\underline{X}_t \underline{X}_t^T)$$

$$= A V A^T + \sigma^2 C C^T \tag{3.3.20}$$

$$C(k) = E(\underline{X}_t \underline{X}_{t+k}^T)$$

$$= A^k C(0) , k > 0 \tag{3.3.21}$$

Expressions (3.3.18) and (3.3.19) are easily derived from (3.3.20) and (3.3.21) respectively.

Corollary 3.3.5. Let e_t, $t \in Z$ be a sequence of independent identically distributed real random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let the roots of the polynomial

$$f(x) = 1 - a_1 x - a_2 x^2 - \ldots - a_r x^r$$

be greater than unity in absolute value. The strictly stationary second-order process X_t, $t \in Z$ conforming to the model

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + e_t \qquad \text{a.e } [P] \tag{3.3.22}$$

for every t in Z has mean 0 and autocovariance function given by

$$R(k) = a_1 R(k-1) + a_2 R(k-2) + \ldots + a_r R(k-r) , k > h \tag{3.3.23}$$

The first h autocovariances depend on the moving average parameters $b_1, b_2, \ldots , b_h$, as well as on the autoregressive parameters $a_1, a_2, \ldots \ldots , a_r$.

Proof. By Corollary 2.4.5, the strictly stationary second-order process X_t, $t \in Z$ satisfying (3.3.22) admits the vector representation

$$\underline{X}_t = A \underline{X}_{t-1} + \sum_{j=1}^{h} \underline{b}_j e_{t-j} + Ce_t \qquad \text{a.e } [P]$$

for every t in Z with the matrix A being of the special type (3.3.2). The autocovariance function (3.3.23) and the comments following it follow from Theorems 3.2.4 and 3.3.1.

3.4 AUTOREGRESSIVE AND MOVING AVERAGE STRUCTURES OF BILINEAR PROCESSES

In section 3.3, we have seen that for any bilinear process $\underline{X}_t$, $t \in Z$ satisfying

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} \underline{b}_j e_{t-j} + \sum_{j=1}^{q} B_j\,\underline{X}_{t-j} e_{t-j} + Ce_t \quad \text{a.e } [P] \quad (3.4.1)$$

for every t in Z, under some conditions, there exists an ARMA process with identical covariance structures. Suppose in the bilinear model above, the moving average part is missing, ie,

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{q} B_j\,\underline{X}_{t-j} e_{t-j} + Ce_t \quad \text{a.e } [P] \quad (3.4.2)$$

for every t in Z (certainly, we do have an ARMA process whose covariance structure is identical with that of (3.4.2)). Is there an autoregressive process whose covariance structure is identical with the one of (3.4.2)? This question we are unable to settle generally. However, in some special cases, it does indeed work out to be true.

We look at the case q = 1. For given matrices $\underset{p\times p}{A}$ and $\underset{p\times p}{B}$, let $\Gamma = A \otimes A + \sigma^2 B \otimes B$, where σ^2 is the variance of e_t, $t \in Z$. Assume $\rho(\Gamma) < 1$ and $E(e_t^4) < \infty$. Then for the process $\underline{X}_t$, $t \in Z$ satisfying

$$\underline{X}_t = A\,\underline{X}_{t-1} + B\,\underline{X}_{t-1}\, e_{t-1} + Ce_t \quad \text{a.e } [P] \quad (3.4.3)$$

for every t in Z, let $\underline{\mu} = E(\underline{X}_t)$ and $V = E(\underline{X}_t\,\underline{X}_t^T)$.

From Theorems 3.2.1 and 3.3.1, recall:

$$\underline{\mu} = \sigma^2(I - A)^{-1}\, B\, C$$

$$V = A\,V\,A^T + \sigma^2\, B\,V\,B^T + \Delta$$

$$S = \sigma^2(A\,\underline{\mu}\,C^T + C\,\underline{\mu}^T A^T + \sigma^2\, B\,C\,C^T + \sigma^2\, C\,C^T B^T) + k_3\, C\,C^T$$

$$H = k_3(A\,\underline{\mu}\,C^T + C\,\underline{\mu}^T A^T + \sigma^2\, B\,C\,C^T + \sigma^2\, C\,C^T B^T) + k_4\, C\,C^T$$

$$\Delta = A\,S\,B^T + B\,S\,A^T + \sigma^2\,C\,C^T + 2\sigma^2\,B\,C\,C^T\,B^T + B\,H\,B^T$$

$$C(0) = E(\underline{X}_t\,\underline{X}_t^T) - \underline{\mu}\,\underline{\mu}^T$$

$$= A\,C(0)A^T + \sigma^2\,B\,C(0)B^T + \Delta_3$$

$$C(1) = E(\underline{X}_{t+1}\,\underline{X}_t^T) - \underline{\mu}\,\underline{\mu}^T$$

$$= A\,C(0) + \Delta_4 \qquad (3.4.4)$$

$$C(k) = E(\underline{X}_{t+k}\,\underline{X}_t^T) - \underline{\mu}\,\underline{\mu}^T$$

$$= A\,C(k-1)$$

$$= A^{k-1}\,C(1)\,,\ k = 2,\ 3,\ \dots\,. \qquad (3.4.5)$$

where

$$\Delta_3 = \Delta + A\,\underline{\mu}\,\underline{\mu}^T\,A^T + \sigma^2\,B\,\underline{\mu}\,\underline{\mu}^T\,B^T - \underline{\mu}\,\underline{\mu}^T$$

$$\Delta_4 = B\,S + A\,\underline{\mu}\,\underline{\mu}^T - \underline{\mu}\,\underline{\mu}^T\,. \qquad (3.4.6)$$

From (3.4.5), it is obvious that the process $\underline{X}_t$ $t \in Z$ conforming to the bilinear model (3.4.3) with the matrix A being of the special type (3.3.2) has the same covariance structure as an ARMA (p, 1) process. See also Subba Rao [37, p.248]. Such a process will have the same covariance structure as an autoregressive process of order p when $\Delta_4 = \underline{0}$. The matrix Δ_4 is a null matrix if $B\,C = \underline{0}$ and $k_3 = E(e_t^3) = 0$.

An example of such a process is the BARMA (p, 1, ℓ, 1) process $\underline{X}_t$, $t \in Z$ satisfying

$$X_t = \sum_{j=1}^{r} a_j\,X_{t-j} + (\beta_{21}\,X_{t-2} + \beta_{31}\,X_{t-3} + \dots\dots + \beta_{\ell 1}\,X_{t-\ell})e_{t-1}$$

$$+ e_t \qquad \text{a.e } [P] \quad (3.4.7)$$

where e_t, $t \in Z$ is a sequence of independent identically distributed random variables with $E(e_t) = 0$, $E(e_t^2) = \sigma^2$, $E(e_t^3) = 0$ and $E(e_t^4) < \infty$.

In the vector representation of (3.4.7), we notice that

$$\underset{p\times p}{B} = \begin{bmatrix} 0 & \beta_{21} & \beta_{31} & \cdots & \beta_{\ell 1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$\underset{1\times p}{C^T} = (1, 0, 0, \ldots, 0)$$

$$p = \max\{r, \ell\}$$

Hence B C = 0. From the above discussion, we have $\Delta_4 = \underline{0}$.

Corollary 3.3.2 establishes the fact that for any bilinear process X_t, $t \in Z$ satisfying (2.1.2), under some conditions, there exists an ARMA process with identical covariance structures. Suppose in the bilinear model (2.1.2), the autoregressive part is missing, ie,

$$X_t = \sum_{j=1}^{h} b_j e_{t-j} + \sum_{\substack{i=1 \\ i \geq j}}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-i} e_{t-j} + e_t \quad \text{a.e } [P] \quad (3.4.8)$$

for every t in Z, it is not difficult to show that X_t, $t \in Z$ satisfying (3.4.8) admits the vector representation (3.4.1) if we let

$$p = m$$

$$g = \min(m, \ell)$$

$$q = \max(h, g)$$

$$\underset{p\times p}{A} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

$$\underset{1\times p}{\underline{b}_j^T} = (b_j, 0, 0, \ldots 0), \quad j = 1, 2, \ldots, q$$

$$\underset{1\times p}{C^T} = (1, 0, 0, \ldots, 0)$$

$$\underset{1xp}{\underline{X}_t^T} = (X_t, X_{t-1}, \dots, X_{t-p+1}), \; t \in Z$$

If we now let $\rho(\Gamma) < 1$, where Γ is the matrix of Theorem 2.4.3 built on $A, B_1, B_2, \dots, B_q$ and σ^2 with $E(e_t^4) < \infty$, we obtain from equation (3.3.6) of Theorem 3.3.1 that there exists a moving average process with identical covariance structures.

Repeating the above discussion for the purely bilinear process X_t, $t \in Z$ satisfying

$$X_t = \sum_{\substack{i=1 \\ i \geq j}}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-j} + e_t \qquad \text{a.e } [P] \quad (3.4.9)$$

for every t in Z, we note that, under suitable conditions, there exists a moving average process of order $\min\{m, \ell\}$ with identical covariance structures.

3.5 EXAMPLES WITH NUMERICAL ILLUSTRATIONS

3.5.1 Methods of ARMA Model Identification

Before we give examples to illustrate some of the points discussed in chapter 2 and 3, let us first summarise two methods of ARMA model identification.

(a) BOX-JENKINS METHOD

Basic to the Box and Jenkins [5] method of ARMA model identification is the partial autocorrelation function given by

$$\phi_{kk} = \rho_1 \qquad , \text{ if } k = 1$$

$$= |A(k, 0)|/|B(k, 0)| \; , \text{ if } k > 1 \; ,$$

where B(s,t) is the sxs matrix defined by

$$B(s,t) = \begin{bmatrix} \rho_t & \rho_{t-1} & \cdots & \rho_{t-s+1} \\ \rho_{t+1} & \rho_t & \cdots & \rho_{t-s+2} \\ \vdots & \vdots & & \vdots \\ \rho_{t+s-1} & \rho_{t+s-2} & \cdots & \rho_t \end{bmatrix}$$

$\rho_t = R(k)/R(0)$

and $A(s,t)$ is the matrix composed of the first $s - 1$ columns of $B(s,t)$ with the s-th column given by $\underline{\rho}$ where

$\underline{\rho}^T = (\rho_{t+1}, \rho_{t+s}, \ldots, \rho_{t+s})$.

The Box-Jenkins procedure uses the fact that if X_t, $t \in Z$ actually is ARMA (p, 0), then ϕ_{kk} is non-zero for $k \leqslant p$ and identically zero for $k > p$. Also used in the Box-Jenkins procedure is the fact that if the process is ARMA (0, q), then $\rho_k = 0$, $k > q$. The inspection of sample autocorrelations $\hat{\rho}_k$'s and partial autocorrelations $\hat{\phi}_{kk}$'s should indicate the model or models to be entertained. This is done by comparing the estimated functions with their large-lag standard errors (see Table 3.1), and then seeing where the cut-offs, if any, occur in the $\hat{\rho}_k$'s and $\hat{\phi}_{kk}$'s.

<u>TABLE 3.1</u> SUMMARY OF BOX-JENKINS IDENTIFICATION PROCEDURE

(1)	$\hat{\rho}_k \sim N(0, \frac{1}{n}(1 + 2\sum_{k=1}^{q} \hat{\rho}_k^2))$, $k > q$	$\rightarrow$ ARMA (0, q)
(2)	$\hat{\phi}_{kk} \sim N(0, \frac{1}{n})$, $k > p$	$\rightarrow$ ARMA (p, 0)
(3)	Neither (1) nor (2) holds, then ARMA (p,q) is to be tried for some $p,q > 0$.	
	n is the number of observations used in calculating the $\hat{\rho}_k$'s and $\hat{\phi}_{kk}$'s.	

When p and q are both greater than zero, this procedure would not yield unique values of p and q.

(b) GKM R- AND S-ARRAYS METHOD

Gray, Kelly and McIntire [17] uses R- and S-array elements as the following ratios.

$$R_n(f_m) = H_n(f_m)/H_n(1;\ f_m)$$

$$S_n(f_m) = H_{n+1}(1;\ f_m)/H_n(f_m)$$

where $H_n(f_m)$ is the determinant of the nxn matrix with (i, j)-th element given by $f_{m+i+j-2}$, and $H_{n+1}(1;\ f_m)$ is the determinant of the $(n + 1) \times (n + 1)$ matrix with (1, j)-th element equal to 1 and (i, j)-th element for $i \geq 2$ is given by $f_{m+i+j-3}$. In their work, $f_m = \rho_m$ or $f_m = (-1)^m \rho_m$. The properties of the R- and S-arrays on which the GKM procedure depend are summarized below.

Let X_t, $t \in Z$ be a stationary ARMA (p,q) process satisfying

$$X_t - \sum_{j=1}^{p} \phi_j X_{t-j} = a_t - \sum_{j=1}^{q} \theta_j a_{t-j} \qquad \text{a.e } [P]$$

for every t in Z. Suppose that $S_n(f_m)$ and $R_n(f_m)$ are defined, $p > 0$ and $S_n(f_m) \neq 0$, where $f_m = \rho_m$ or $f_m = (-1)^m \rho_m$, then

(1) for some integer m_0 and some constant $C_1 \neq 0$,

$$S_n(f_m) = C_1,\ m \geq m_0$$

$$S_n(f_{m_0-1}) \neq C_1$$

if and only if $n = p$ and $m_0 = q - p + 1$. Also

$$S_n(f_m) = C_2\ ,\ m \leq m_1$$

$$S_n(f_{m_1+1}) \neq C_2$$

for some integer m_1 and some constant $C_2 \neq 0$, if and only if $n = p$ and $m_1 = - q - p$. Moreover,

$$C_1 = (-1)^p(1 - \sum_{j=1}^{p} \phi_j) \text{ if } f_m = \rho_m \text{ or } f_m = (-1)^m \rho_m$$

and

$C_2 = - C_1/\phi_p$ if $f_m = \rho_m$

$= (-1)^{p+1} C_1/\phi_p$ if $f_m = (-1)^m \rho_m$.

(2) for $k > n$,

$S_k(f_{-k-m}) = \pm \infty$

$S_k(f_{-k+m+1}) = (-1)^{k-n} S_n(f_{-n+m+1})$

if and only if $n = p$ and $m = q$. See Woodward and Gray [45] for proof.

(3) $R_{n+1}((-1)^m \rho_m) = R_{n+1}(\rho_m) = 0$, $m \geq m_0$, $m \leq m_1$,

and

$R_{n+1}(\rho_{q-p}) \neq 0$

$R_{n+1}(\rho_{-q-p+1}) \neq 0$

$R_{n+1}((-1)^{q-p} \rho_{q-p}) \neq 0$

$R_{n+1}((-1)^{-q-p+1} \rho_{-q-p+1}) \neq 0$

if and only if $n = p$, $m_0 = q - p + 1$ and $m_1 = - q - p$.

A process is a stationary ARMA (p,q) process if and only if the associated R- and S-arrays are as in Tables 3.2 and 3.3 respectively

TABLE 3.2 R-ARRAY WHERE X_t, $t \in Z$ IS ARMA (p,q)

$(R_n(f_m) = R_n(m))$

m \ n	1	2	p	p+1
-i	$R_1(-i)$	$R_2(-i)$	$R_p(-i)$	0
⋮	⋮	⋮	⋮	⋮
-q-p-1	$R_1(-q-p-1)$	$R_2(-q-p-1)$	$R_p(-q-p-1)$	0
-q-p	$R_1(-q-p)$	$R_2(-q-p)$	$R_p(-q-p)$	NON-ZERO
⋮	⋮	⋮	⋮	⋮
q-p	$R_1(q-p)$	$R_2(q-p)$	$R_p(q-p)$	NON-ZERO
q-p+1	$R_1(q-p+1)$	$R_2(q-p+1)$	$R_p(q-p+1)$	0
⋮	⋮	⋮	⋮	⋮
j	$R_1(j)$	$R_2(j)$	$R_p(j)$	0

TABLE 3.3 S-ARRAY WHERE X_t, $t \in Z$ IS ARMA (p,q)

$$S_n(f_m) = S_n(m)$$

m \ n	1	2	p	p+1
-i	$S_1(-i)$	$S_2(-i)$	C_2	U*
:	:	:	:	:
-q-p-2	$S_1(-q-p-2)$	$S_2(-q-p-2)$	C_2	U*
-q-p-1	$S_1(-q-p-1)$	$S_2(-q-p-1)$	C_2	± ∞
-q-p	$S_1(-q-p)$	$S_2(-q-p)$	C_2	{2q NON-CONSTANT
:	:	:	{2q NON-CONSTANT	
q-p	$S_1(q-p)$	$S_2(q-p)$		$-C_1$
q-p+1	$S_1(q-p+1)$	$S_2(q-p+1)$	C_1	U*
q-p+2	$S_1(q-p+2)$	$S_2(q-p+2)$	C_1	U*
:	:	:	:	:
j	$S_1(j)$	$S_2(j)$	C_1	U*

U* ≡ undefined

In Table 3.3, column p+1 contains several undefined elements. That is, in the presence of noise the column having the characteristics of column p will be followed by a highly variable column. So p is identified as the first column having the correct constant behaviour followed by a highly variable column.

REMARK 3.5.1 When p = 0, the Box-Jenkins method and the R- and S-arrays procedure all use primarily the autocorrelation function with its property that $\rho_k = 0$, for all $k > q$.

3.5.2 Examples

The following examples illustrate further the work of this chapter and chapter 2.

EXAMPLE 3.1. We consider the BARMA (1, 1, 1, 1,) process X_t, $t \in Z$ satisfying

$$X_t = a X_{t-1} + b e_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P] \quad (3.5.1)$$

for every t in Z where e_t, $t \in Z$ are independent for each e_t is distributed as $N(0, \sigma^2)$. The importance of this example is that it illustrates a bilinear model containing an autoregressive part, a moving average part and a bilinear part.

By Remarks 2.3.3 (1), the strict stationarity condition is

$a^2 + \beta^2 \sigma^2 < 1.$

The mean, second moments and covariances are as follows

$$\mu = E(X_t) = \beta \sigma^2/(1 - a)$$

$$m_2 = E(X_t^2)$$

$$= \frac{\sigma^2}{1 - a^2 - \beta^2 \sigma^2}\left\{1 + b^2 + 2 a b + 2 \beta \mu(1 + a + b)\right\}$$

$$E(X_t X_{t+k})$$

$$= a^{k-1}(a m_2 + b \sigma^2 + (1 - 2a)\mu^2) + \mu^2 , \; k > 0$$

$$R(0) = m_2 - \mu^2$$

$$R(k) = a^{k-1}(a m_2 + b \sigma^2 + (1 - 2a)\mu^2) , \; k > 0$$

Our numerical illustration consists of 500 points generated from the process (3.5.1) with $a = 0.5$, $b = 0.4$, $\beta = 0.3$ and $\sigma^2 = 1$. Using these values of a, b, β and σ^2, we obtain the theoretical mean and covariances to be

$$\mu = 0.60$$

$$R(0) = 3.04$$

$$R(k) = 2.10(0.50)^{k-1} , \; k > 0$$

Figure 3.1 gives a graph of the data; columns 2 and 3 of Table 3.4 give the theoretical and sample autocovariances respectively. We note that

$$\hat{R}(k) = \frac{1}{n} \sum_{t=1}^{n-|k|} (X_t - \bar{X})(X_{t+k} - \bar{X}) , \; k \in Z$$

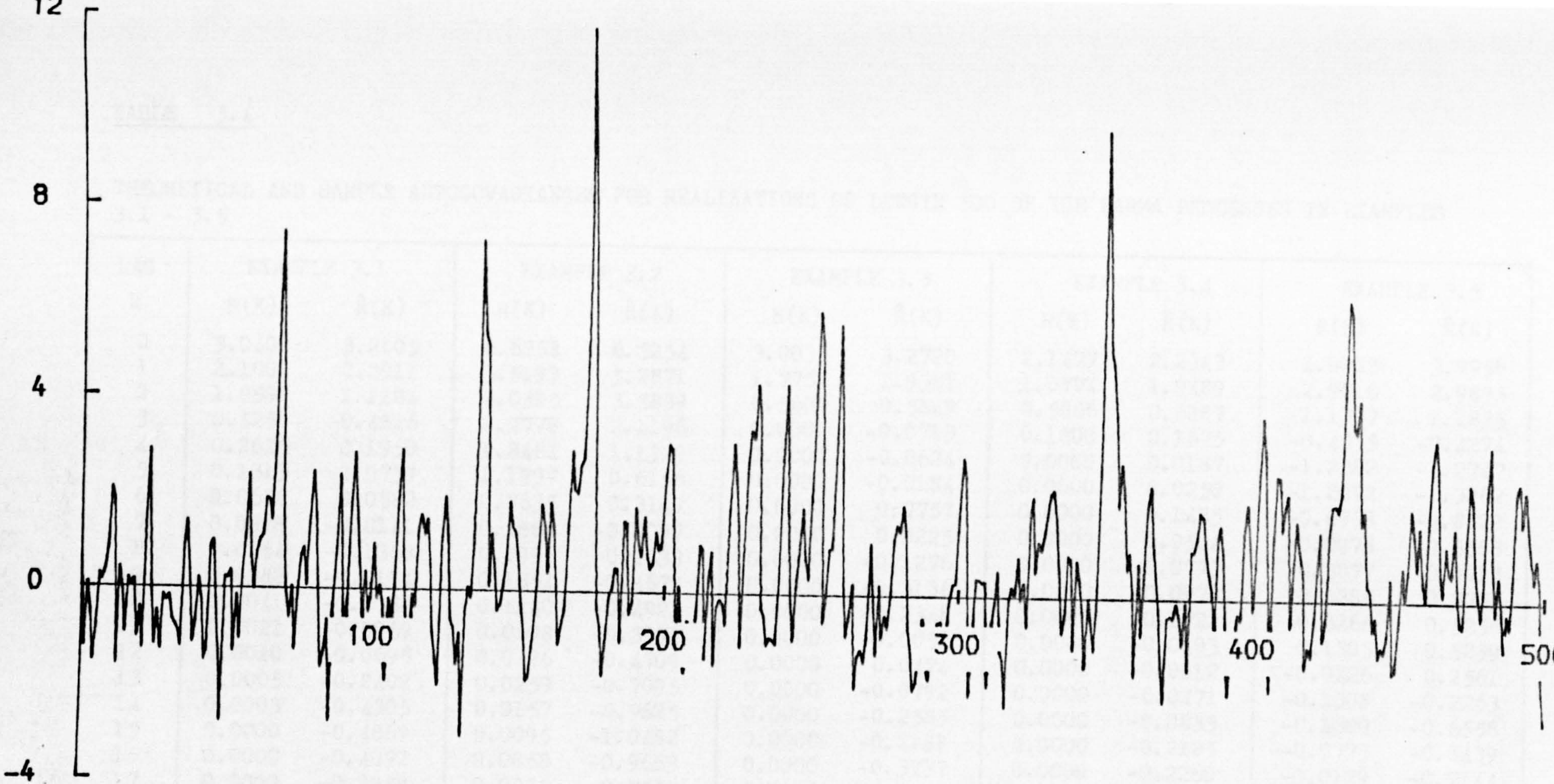


FIG.3.1 500 OBSERVATIONS FROM THE BARMA (1, 1, 1, 1) PROCESS IN EXAMPLE 3.1

TABLE 3.4

THEORETICAL AND SAMPLE AUTOCOVARIANCES FOR REALIZATIONS OF LENGTH 500 OF THE BARMA PROCESSES IN EXAMPLES 3.1 - 3.5

LAG	EXAMPLE 3.1		EXAMPLE 3.2		EXAMPLE 3.3		EXAMPLE 3.4		EXAMPLE 3.5	
k	R(K)	$\hat{R}(K)$	R(K)	$\hat{R}(K)$	R(K)	$\hat{R}(K)$	R(K)	$\hat{R}(K)$	R(K)	$\hat{R}(K)$
0	3.0400	3.2105	6.6358	6.5254	3.0033	3.2720	2.1477	2.2343	4.0003	3.9956
1	2.1000	2.2011	5.5193	5.2871	1.7752	1.9382	1.0724	1.0489	2.9816	2.9835
2	1.0500	1.1104	4.0305	3.5839	0.5600	0.5849	0.5866	0.6287	1.1377	1.1835
3	0.5250	0.4826	2.7778	2.1196	0.0000	-0.0713	0.1800	0.1675	-0.4535	-0.4274
4	0.2625	0.1950	1.8464	1.1323	0.0000	-0.0624	0.0000	0.0187	-1.2382	-1.2747
5	0.1313	0.0739	1.1977	0.6156	0.0000	-0.0184	0.0000	0.0250	-1.2092	-1.3262
6	0.0656	0.0582	0.7635	0.3147	0.0000	0.0757	0.0000	0.1435	-0.6958	-0.8822
7	0.0328	-0.0141	0.4860	-0.0097	0.0000	0.0225	0.0000	0.0544	-0.0973	-0.3588
8	0.0164	-0.1368	0.2996	-0.2633	0.0000	-0.1276	0.0000	-0.0793	0.3077	0.0713
9	0.0082	-0.2184	0.1854	-0.4670	0.0000	-0.3136	0.0000	-0.0978	0.4285	0.3757
10	0.0041	-0.2381	0.1140	-0.4921	0.0000	-0.2348	0.0000	-0.1722	0.3266	0.5258
11	0.0021	-0.0959	0.0698	-0.3971	0.0000	-0.0052	0.0000	-0.0493	0.1305	0.5239
12	0.0010	-0.0695	0.0426	-0.4709	0.0000	0.0474	0.0000	-0.0012	-0.0426	0.2561
13	0.0005	-0.2402	0.0259	-0.7095	0.0000	-0.0792	0.0000	-0.0271	-0.1308	-0.2263
14	0.0003	-0.4305	0.0157	-0.9625	0.0000	-0.2585	0.0000	-0.0833	-0.1309	-0.6558
15	0.0000	-0.4869	0.0095	-1.0482	0.0000	-0.4188	0.0000	-0.2185	-0.0773	-0.8439
16	0.0000	-0.4092	0.0058	-0.9663	0.0000	-0.3737	0.0000	-0.2260	-0.0129	-0.7144
17	0.0000	-0.2958	0.0035	-0.7550	0.0000	-0.2082	0.0000	-0.1766	0.0316	-0.4032
18	0.0000	-0.1102	0.0021	-0.4538	0.0000	0.0140	0.0000	0.0266	0.0459	-0.0120
19	0.0000	-0.0303	0.0012	-0.1959	0.0000	0.0282	0.0000	0.0518	0.0357	0.3142
20	0.0000	-0.0208	0.0007	-0.0146	0.0000	-0.0572	0.0000	-0.0400	0.0149	0.5120
21	0.0000	0.1031	0.0004	0.2061	0.0000	0.0375	0.0000	-0.0067	-0.0039	0.5576
22	0.0000	0.1773	0.0003	0.3503	0.0000	0.0962	0.0000	-0.0362	-0.0133	0.3717
23	0.0000	0.2049	0.0002	0.4312	0.0000	0.1289	0.0000	0.0094	-0.0142	0.0386
24	0.0000	0.2161	0.0001	0.5718	0.0000	0.1603	0.0000	0.1076	-0.0086	-0.2760
25	0.0000	0.2258	0.0001	0.7350	0.0000	0.1405	0.0000	0.1467	-0.0017	-0.3945
$\bar{X}$	0.5988		0.4425		0.7031		0.9347		-0.0239	

$$\hat{\rho}_k = \hat{R}(k)/\hat{R}(0), \quad k \in Z$$

$$\bar{X} = \frac{1}{n} \sum_{t=1}^{n} X_t \quad ,$$

where n is the number of data points. Columns 2 and 3 of Table 3.5 give the sample autocorrelations and partial autocorrelations respectively. While Table 3.6 shows portions of the R- and S-arrays at $f_m = (-1)^m \hat{\rho}_m$. From $S_1(m)$ and $R_2(m)$ columns, it is clear that p = q = 1. The fitted ARMA (1, 1) model is

$$X_t = 0.280 + 0.533\, X_{t-1} + 0.319 a_{t-1} + a_t,$$

where $E(a_t) = 0$ and $Var(a_t) = 1.609$, leading to a 60.9 per cent increase in the error variance.

EXAMPLE 3.2. Let us now consider a bilinear process with the moving average part missing. We consider the BARMA (2, 0, 2, 2) process $X_t, t \in Z$ satisfying

$$X_t = \sum_{j=1}^{2} a_j X_{t-j} + \sum_{\substack{i=1 \\ i \geq j}}^{2} \sum_{j=1}^{2} \beta_{ij} X_{t-i} e_{t-j} + e_t \qquad \text{a.e } [P] \quad (3.5.2)$$

for every t in Z where e_t, $t \in Z$ are independent and each e_t is distributed as $N(0, \sigma^2)$. First we identify the A, B_1, B_2, C and Γ matrices.

$$\underset{2\times 2}{A} = \begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix}$$

$$\underset{2\times 2}{B_1} = \begin{pmatrix} \beta_{11} & \beta_{21} \\ 0 & 0 \end{pmatrix}$$

$$\underset{2\times 2}{B_2} = \begin{pmatrix} \beta_{22} & 0 \\ 0 & 0 \end{pmatrix}$$

$$\underset{1\times 2}{C^T} = (1, 0)$$

TABLE 3.5

SAMPLE AUTOCORRELATIONS AND PARTIAL AUTOCORRELATIONS FOR REALIZATIONS OF LENGTH 500 OF THE BARMA PROCESSES IN EXAMPLES 3.1 - 3.5

LAG	EXAMPLE 3.1		EXAMPLE 3.2		EXAMPLE 3.3		EXAMPLE 3.4		EXAMPLE 3.5	
k	$\hat{\rho}_k$	$\hat{\phi}_{kk}$	$\hat{\rho}_k$	$\hat{\phi}_{kk}$	$\hat{\rho}_k$	$\hat{\phi}_{kk}$	$\hat{\rho}_k$	$\hat{\phi}_{kk}$	$\hat{\rho}_k$	$\hat{\phi}_{kk}$
1	0.686	0.686	0.810	0.810	0.592	0.592	0.469	0.469	0.747	0.747
2	0.346	-0.234	0.549	-0.312	0.179	-0.265	0.281	0.078	0.296	-0.591
3	0.150	0.037	0.325	-0.020	-0.022	0.002	0.075	-0.108	-0.107	-0.062
4	0.061	-0.002	0.174	0.020	-0.019	0.072	0.008	-0.012	-0.319	0.012
5	0.023	-0.002	0.094	0.032	-0.006	-0.050	0.011	0.042	-0.332	-0.020
6	0.018	0.020	0.048	-0.033	0.023	0.052	0.064	0.072	-0.221	-0.023
7	-0.004	-0.048	-0.001	-0.073	0.007	-0.043	0.024	-0.050	-0.090	-0.055
8	-0.043	-0.037	-0.040	0.007	-0.039	-0.047	-0.036	-0.076	0.018	0.042
9	-0.068	-0.017	-0.072	-0.032	-0.096	-0.056	-0.044	0.010	0.094	0.050
10	-0.074	-0.015	-0.075	0.034	-0.072	0.038	-0.077	-0.039	0.132	0.008
11	-0.030	0.066	-0.061	-0.001	-0.002	0.034	-0.022	0.039	0.131	0.016
12	-0.022	-0.070	-0.072	-0.102	0.014	-0.047	-0.001	0.005	0.064	-0.111
13	-0.075	-0.090	-0.109	-0.061	-0.024	-0.030	-0.012	-0.036	-0.057	-0.078
14	-0.134	-0.052	-0.148	-0.027	-0.079	-0.055	-0.037	-0.027	-0.164	-0.001
15	-0.152	-0.018	-0.161	0.017	-0.128	-0.073	-0.098	-0.078	-0.211	-0.043
16	-0.127	0.003	-0.148	-0.015	-0.114	0.007	-0.101	-0.016	-0.179	-0.009
17	-0.092	-0.018	-0.116	0.009	-0.064	-0.010	-0.079	-0.002	-0.101	-0.053
18	-0.034	0.048	-0.070	0.037	-0.004	0.030	0.012	0.068	-0.003	0.042
19	-0.009	-0.030	-0.030	-0.003	0.009	-0.039	0.023	-0.005	0.079	0.011
20	-0.006	0.003	-0.002	0.008	-0.017	-0.009	-0.018	-0.073	0.128	0.022

TABLE 3.6

R- AND S-ARRAYS AT $f_m = (-1)^m \hat{\rho}_m$ FOR A REALIZATION OF LENGTH 500 OF THE BARMA (1, 1, 1, 1) PROCESS IN EXAMPLE 3.1

$R_1(m)$	$R_2(m)$	$R_3(m)$	$R_4(m)$	$R_5(m)$	$R_6(m)$	m	$S_1(m)$	$S_2(m)$	$S_3(m)$	$S_4(m)$	$S_5(m)$	$S_6(m)$
-0.023	-0.003	0.003	-0.030	0.096	-0.232	-5	-3.639	13.050	31.647	132.118	1437.865	2.114
0.061	0.008	0.008	0.014	0.231	-0.005	-4	-3.475	17.906	-89.050	1015.585	-2.158	-2.995
-0.150	-0.033	-0.043	-0.232	0.000	-0.004	-3	-3.301	13.739	-59.084	2.161	27.197	15.050
0.346	0.120	0.241	0.000	0.003	0.001	-2	-2.982	8.880	-2.156	3.682	-12.063	286.392
-0.686	-0.314	-0.009	-0.000	0.004	-0.019	-1	-2.459	2.081	-2.035	-23.180	-14.899	7.167
1.000	0.074	0.001	-0.008	0.011	-0.031	0	-1.686	1.832	-1.200	-10.523	11.523	6.681
-0.685	-0.016	0.000	-0.005	0.023	0.020	1	-1.504	1.721	31.845	10.210	-4.048	2.381
0.346	0.003	0.005	0.058	0.016	0.015	2	-1.435	1.879	-0.869	14.195	-9.025	5.361
-0.150	-0.001	0.006	-0.029	-0.007	0.022	3	-1.404	-7.267	-9.439	6.333	10.195	6.368
0.061	-0.007	0.065	0.007	0.531	0.036	4	-1.379	-0.952	-5.146	-0.131	10.626	5.233
-0.023	-0.010	-0.003	0.007	-0.002	0.035	5	-1.787	4.941	7.865	-10.577	-144.387	2.934

$$\underset{8\times 8}{\Gamma} = \begin{pmatrix} \Gamma_1 & \Gamma_2 \\ I_4 & \underline{0} \end{pmatrix}$$

where

$$\underset{4\times 4}{\Gamma_1} = \begin{pmatrix} a_1^2 + \sigma^2\beta_{11}^2 & a_1a_2 + \sigma^2\beta_{11}\beta_{21} & a_1a_2 + \sigma^2\beta_{11}\beta_{21} & a_2^2 + \sigma^2\beta_{21}^2 \\ a_1 & 0 & a_2 & 0 \\ a_1 & a_2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\underset{4\times 4}{\Gamma_2} = \sigma^2 \begin{pmatrix} \beta_{22}(\beta_{22} + 2a_1\beta_{11}) & a_1\beta_{21}\beta_{22} & a_1\beta_{21}\beta_{22} & 0 \\ \beta_{11}\beta_{22} & \beta_{21}\beta_{22} & 0 & 0 \\ \beta_{11}\beta_{22} & 0 & \beta_{21}\beta_{22} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Let

$$\underset{1\times 2}{\underline{X}_t^T} = (X_t, X_{t-1}), \quad t \in Z$$

then the vector form of (3.5.2) is

$$\underline{X}_t = A\,\underline{X}_{t-1} + \sum_{j=1}^{2} B_j\,\underline{X}_{t-j}\,e_{t-j} + Ce_t \qquad \text{a.e } [P]$$

for every t in Z. The strict stationarity condition $\rho(\Gamma) < 1$ implies that the roots (in modulus) of the equation

$$|\Gamma_2 + y\,\Gamma_1 - y^2 I_4| = 0$$

lie inside the unit circle.

The mean, second order moments and covariances can be evaluated using the methods stated in the proofs of Theorems 3.2.4 and 3.3.1. We state the results.

$$\mu = E(X_t) = \sigma^2(\beta_{11} + \beta_{22})/(1 - a_1 - a_2)$$

Let

$$d_1 = 1 + 2\sigma^2(\beta_{11}^2 + \beta_{22}^2 + \beta_{11}\,\beta_{22})$$

$$+ 2a_1(\mu(2\beta_{11} + \beta_{22} + \beta_{21}) + \beta_{22}(a_1\,\mu + \sigma^2\,\beta_{11}))$$

$$+ 2a_2\mu(\beta_{11} + 2\beta_{22})$$

$$d_2 = \mu(2\beta_{11} + \beta_{22} + \beta_{21}) + \beta_{22}(a_1\,\mu + \sigma^2\,\beta_{11})$$

Then

$$(1 - a_1^2 - a_2^2 - \sigma^2(\beta_{11}^2 + \beta_{21}^2 + \beta_{22}(\beta_{22} + 2a_1\,\beta_{11})))E(X_t^2)$$

$$= 2(a_1\,a_2 + \sigma^2\,\beta_{21}(\beta_{11} + a_1\,\beta_{22}))E(X_t\,X_{t-1}) + \sigma^2\,d_1$$

$$(1 - a_2 - \sigma^2\,\beta_{21}\,\beta_{22})E(X_t\,X_{t-1})$$

$$= (a_1 + \sigma^2\,\beta_{11}\,\beta_{22})E(X_t^2) + \sigma^2\,d_2$$

$$E(X_t\,X_{t-2}) = a_1\,E(X_t\,X_{t-1}) + a_2\,E(X_t^2) + \sigma^2(\beta_{11} + 2\beta_{22})\mu$$

$$E(X_t\,X_{t-k}) = a_1\,E(X_t\,X_{t-k+1}) + a_2\,E(X_t\,X_{t-k+2})$$

$$+ \sigma^2(\beta_{11} + \beta_{22})\mu\ ,\ k > 2$$

and

$$R(k) = a_1\,R(k-1) + a_2\,R(k-2)\ ,\ k > 2.$$

Our numerical illustration consists of 500 points generated from the process (3.5.2) with

$$a_1 = 1.10\ ,$$

$$a_2 = -0.30\ ,$$

$$\beta_{11} = 0.20\ ,$$

$$\beta_{21} = 0.15\ ,$$

$$\beta_{22} = -0.10\ ,$$

$$\sigma^2 = 1\ .$$

Using these values of a_j's and β_{ij}'s, we obtain the following:

$\rho(\Gamma) = 0.663$

$\mu = 0.500$

$R(0) = 6.6358$

$R(1) = 5.5193$

$R(2) = 4.0305$

and

$$R(k) = 1.1\, R(k - 1) - 0.3\, R(k - 2)\ ,\ k > 2.$$

Figure 3.2 gives a graph of the data; columns 4 and 5 of Table 3.4 give the theoretical and sample autocovariances respectively. Columns 4 and 5 of Table 3.5 give the estimated autocorrelations and partial autocorrelations respectively, while Table 3.7 shows portions of the R- and S-arrays at $f_m = (-1)^m\, \hat{\rho}_m$. The $S_1(m)$ column suggests that an ARMA (1, 2) might be an appropriate model, while the $S_2(m)$ column suggests an ARMA (2. 0) model. See also column 5 of Table 3.5. To choose the best ARMA model we employ the information criterion of Akaike (AIC). See Akaike [2] . This is given by

$$AIC = n \log \hat{\sigma}^2 + 2(\text{number of parameters})$$

where n is the effective number of observations used in the estimation process and $\hat{\sigma}^2$ is the sample estimate of σ^2. We give the AIC values for alternative ARMA (p, q), p, q = 1, 2 models fitted to the mean deleted observations, $x_t = X_t - 0.4425$ with n = 498.

FITTED MODEL	AIC VALUE	$\hat{\sigma}^2$
ARMA (1, 1)	357.3	2.0397
ARMA (1, 2)	352.2	2.0148
ARMA (2, 0)	348.6	2.0042
ARMA (2, 1)	350.4	2.0074
ARMA (2, 2)	352.1	2.0101

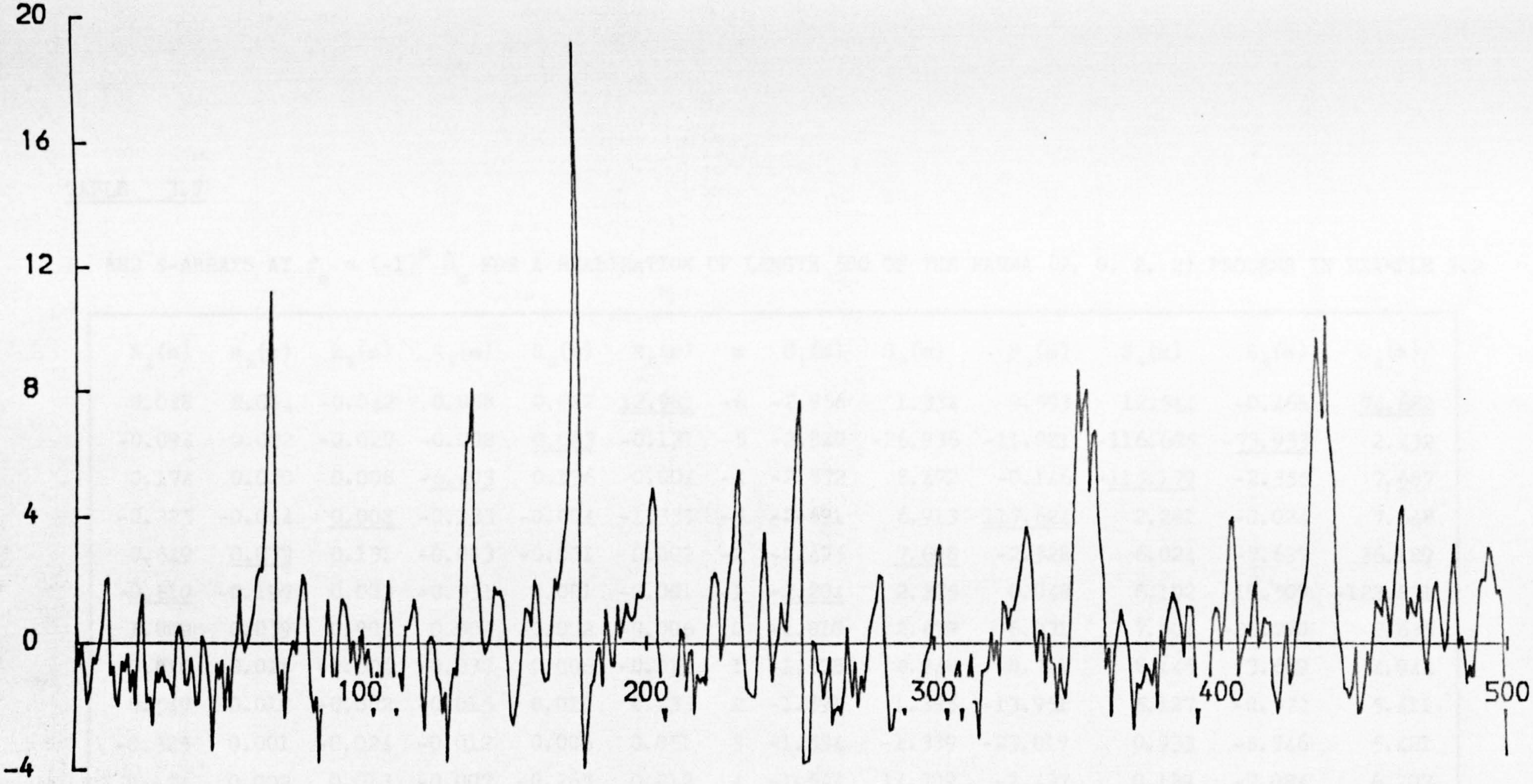


FIG.3.2 500 OBSERVATIONS FROM THE BARMA (2,0,2,2) PROCESS IN EXAMPLE 3.2

TABLE 3.7

R- AND S-ARRAYS AT $f_m = (-1)^m \hat{\rho}_m$ FOR A REALIZATION OF LENGTH 500 OF THE BARMA (2, 0, 2, 2) PROCESS IN EXAMPLE 3.2

$R_1(m)$	$R_2(m)$	$R_3(m)$	$R_4(m)$	$R_5(m)$	$R_6(m)$	m	$S_1(m)$	$S_2(m)$	$S_3(m)$	$S_4(m)$	$S_5(m)$	$S_6(m)$
0.048	0.004	-0.042	0.058	0.082	12.981	-6	-2.956	1.334	9.593	12.544	-0.468	74.682
-0.094	0.002	-0.027	-0.008	0.083	-0.131	-5	-2.840	-26.936	-11.021	-116.685	-73.933	2.432
0.174	0.020	0.008	-6.403	0.136	0.004	-4	-2.872	8.472	-0.146	-115.179	-2.355	7.687
-0.325	-0.044	0.008	-0.133	-0.004	-1.339	-3	-2.691	6.913	117.624	2.282	-0.024	7.648
0.549	0.079	0.131	-0.003	-0.004	0.002	-2	-2.475	7.608	-2.328	6.024	-7.639	36.189
-0.810	-0.190	0.003	-0.331	0.001	-0.001	-1	-2.234	2.375	0.048	6.102	-19.509	-123.831
1.000	0.059	0.003	0.004	-0.002	-0.006	0	-1.810	2.479	-6.031	7.804	-36.790	2.635
-0.810	-0.028	-0.004	-0.037	0.006	-0.011	1	-1.678	2.248	-0.741	8.146	3.669	4.944
0.549	0.012	-0.002	-0.015	0.010	0.133	2	-1.591	1.395	-13.951	5.127	-0.372	5.411
-0.325	0.001	-0.024	-0.012	0.006	0.051	3	-1.534	-1.339	-23.019	0.933	-8.746	5.421
0.174	0.002	0.013	-0.007	-0.267	0.012	4	-1.544	14.709	-2.437	0.189	-7.084	6.207
-0.094	-0.017	0.004	-0.008	0.021	0.005	5	-1.511	3.305	2.685	6.560	-11.038	4.791
0.048	0.042	0.006	0.060	-0.005	0.135	6	-0.969	3.584	-0.769	8.901	0.193	4.501

The minimum AIC value is obtained with an ARMA (2, 0), but it will be seen that the AIC values of some of the models are very close. The fitted ARMA (2, 0) model is

$$X_t = 0.109 + 1.076X_{t-1} - 0.322X_{t-2} + a_t$$

where $E(a_t) = 0$ and $Var(a_t) = 2.0$, leading to a 100 per cent increase in the error variance.

EXAMPLE 3.3. Let us consider a bilinear process with the autoregressive part missing. We consider the BARMA (0, 2, 2, 2) process X_t, $t \varepsilon Z$ satisfying

$$X_t = b_1 e_{t-1} + b_2 e_{t-2} + \beta_{11} X_{t-1} e_{t-1} + \beta_{22} X_{t-2} e_{t-2} + e_t \quad \text{a.e } [P] \quad (3.5.3)$$

for every t in Z where e_t, $t \varepsilon Z$ are independent and each e_t is distributed as $N(0, \sigma^2)$.

The A, B_1, B_2, $\underline{b}_1$, $\underline{b}_2$, C and Γ matrices are identified to be

$$\underset{2x2}{A} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

$$\underset{2x2}{B_1} = \begin{pmatrix} \beta_{11} & 0 \\ 0 & 0 \end{pmatrix}$$

$$\underset{2x2}{B_2} = \begin{pmatrix} \beta_{22} & 0 \\ 0 & 0 \end{pmatrix}$$

$$\underset{1x2}{\underline{b}_1^T} = (b_1, 0)$$

$$\underset{1x2}{\underline{b}_2^T} = (b_2, 0)$$

$$\underset{1x2}{C^T} = (1, 0)$$

$$\underset{4\times4}{\Gamma_1} = \sigma^2\begin{pmatrix} \beta_{11}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\underset{4\times4}{\Gamma_2} = \sigma^2\begin{pmatrix} \beta_{22}^2 & 0 & 0 & 0 \\ \beta_{11}\beta_{22} & 0 & 0 & 0 \\ \beta_{11}\beta_{22} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\underset{8\times8}{\Gamma} = \begin{pmatrix} \Gamma_1 & \Gamma_2 \\ I_4 & \underline{0} \end{pmatrix}$$

The vector form of (3.5.3) is given by (3.1.1) with $p = q = 2$ and

$$\underset{1\times2}{\underline{X}_t^T} = (X_t, X_{t-1}) \text{ , } t \in Z.$$

Γ is the matrix of Theorem 2.4.3 built on A, B_1, B_2 and σ^2. The strict stationarity condition $\rho(\Gamma) < 1$ imples that the roots (in modulus) of the equation

$$y^2 - \sigma^2 \beta_{11}^2 y - \sigma^2 \beta_{22}^2 = 0$$

lie inside the unit circle.

The mean, second-order moments and covariances are given by

$$\mu = E(X_t) = \sigma^2(\beta_{11} + \beta_{22})$$

$$M_2 = E(X_t^2)$$

$$= \frac{\sigma^2}{(1 - \sigma^2(\beta_{11}^2 + \beta_{22}^2))} \Big\{1 + b_1^2 + b_2^2 + 2\sigma^2(\beta_{11}^2 + \beta_{22}^2 + \beta_{11}\beta_{22})$$

$$+ 2\mu(b_1 \beta_{11} + b_2 \beta_{22})\Big\}$$

$$E(X_t X_{t-1}) = \sigma^2 \beta_{11} \beta_{22} M_2 + b_1 + b_2(b_1 + \mu \beta_{11}) + \mu(2\beta_{11} + \beta_{22})$$

$+ \beta_{22}(b_1 \mu + \sigma^2 \beta_{11})$

$$E(X_t X_{t-2}) = \sigma^2(b_2 + \mu \beta_{22}) + \mu^2$$

$$E(X_t X_{t-k}) = \mu^2 , k > 2$$

Thus

$$R(k) = 0 , k > 2$$

Our numerical illustration consists of 500 points generated from the process (3.5.3) with

$b_1 = 0.55$

$b_2 = 0.35$

$\beta_{11} = 0.40$

$\beta_{22} = 0.30$

$\sigma^2 = 1$

Using these values of b_1, b_2, β_{11}, β_{22} and σ^2, we obtain the following

$\rho(\Gamma) = 0.391$

$\mu = 0.700$

$R(0) = 3.0033$

$R(1) = 1.7752$

$R(2) = 0.5600$

$R(k) = 0 , k > 2$

Figure 3.3 gives a graph of the data; columns 6 and 7 of Table 3.4 give the theoretical and sample autocovaraince s respectively. Columns 6 and 7 of Table 3.5 give the sample autocorrelations and partial autocorrelations respectively. Column 6 of Table 3.5 suggests an ARMA (0, 2) model. The fitted ARMA (0, 2) model is

$$X_t = 0.703 + 0.763a_{t-1} + 0.322a_{t-2} + a_t$$

where $E(a_t) = 0$ and $Var(a_t) = 1.962$, leading to a 96 per cent increase in

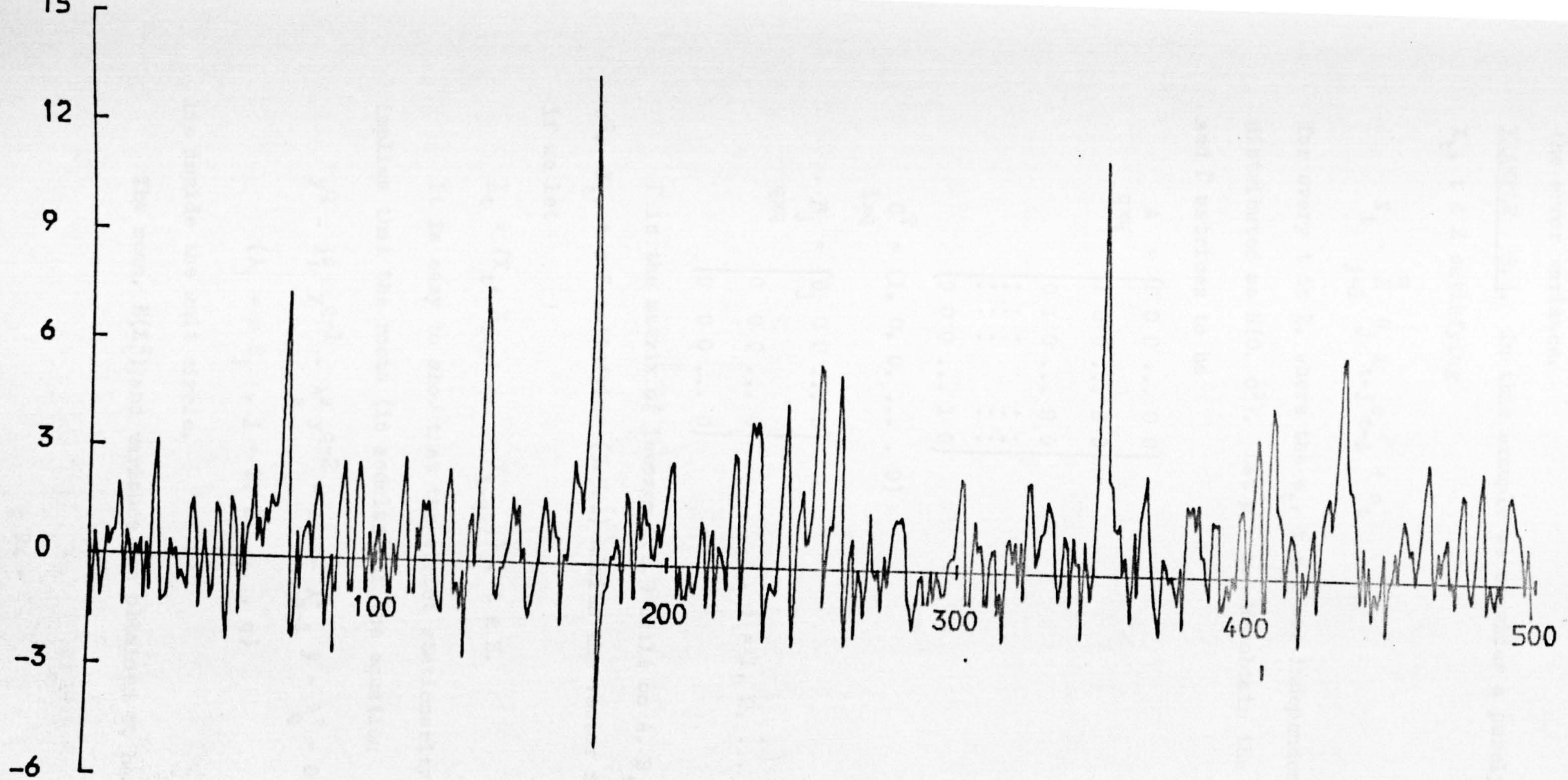


FIG.3.3 500 OBSERVATIONS FROM THE BARMA (0,2,2,2) PROCESS IN EXAMPLE 3.3

the error variance.

EXAMPLE 3.4. In this example we consider a purely bilinear process X_t, $t \in Z$ satisfying

$$X_t = \sum_{j=1}^{q} \theta_j X_{t-j} e_{t-j} + e_t \qquad \text{a.e } [P] \quad (3.5.4)$$

for every t in Z, where the e_t, $t \in Z$ are independent and each e_t is distributed as $N(0, \sigma^2)$. Let p = q. We obtain the A, C, B_1, B_2, ... , B_q and Γ matrices to be

$$\underset{q\times q}{A} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

$$\underset{1\times q}{C^T} = (1, 0, 0, \dots, 0)$$

$$\underset{q\times q}{B_j} = \begin{pmatrix} \theta_j & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad j = 1, 2, \dots, q.$$

Γ is the matrix of Theorem 2.4.3 built on A, B_1, B_2, ... , B_q and σ^2. X_t, $t \in Z$ satisfying (3.5.4) admits the vector representation (3.4.2) if we let

$$\underline{X}_t = (X_t, X_{t-1}, \dots, X_{t-q+1}), \quad t \in Z.$$

It is easy to show that the strict stationarity condition $\rho(\Gamma) < 1$ implies that the roots (in modulus) of the equation

$$y^q - \lambda_1^2 y^{q-1} - \lambda_2^2 y^{q-2} - \dots - \lambda_{q-1}^2 y - \lambda_q^2 = 0$$

$$(\lambda_j = \sigma \theta_j \;, j = 1, 2, \dots, q)$$

lie inside the unit circle.

The mean, $E(X_t^2)$ and variance are obtained to be

$$\mu = E(X_t) = \sigma^2 \sum_{j=1}^{q} \lambda_j$$

$$E(X_t^2) = \frac{\sigma^2}{1 - \sum_{j=1}^{q} \lambda_j^2} \left\{ 1 + 2 \sum_{j=1}^{q} \lambda_j^2 + 2 \sum_{\substack{i=1 \\ i<j}}^{q} \sum_{j=1}^{q} \lambda_i \lambda_j \right\}$$

$$R(0) = \frac{\sigma^2}{1 - \sum_{j=1}^{q} \lambda_j^2} \left\{ 1 + \sum_{j=1}^{q} \lambda_j^2 \left[1 + \left(\sum_{j=1}^{q} \lambda_j \right)^2 \right] \right\}$$

Computation of all second-order moments (and hence covariances) of the bilinear process (3.5.4), in principle, is possible. However, the algebra involved is very tedious. It suffices to observe that

$$E(X_t X_{t-1}) = 2\, \theta_1 \sigma^2 \mu + E(X_t^2 e_t^2) \sum_{i=1}^{q-1} \theta_i \theta_{i+1} + \sigma^4 \sum_{i=2}^{q} \theta_i \left\{ \sum_{\substack{j=1 \\ j \neq i-1}}^{q} \theta_j \right\}$$

where

$$E(X_t^2 e_t^2) = \frac{\sigma^4}{1 - \sum_{j=1}^{q} \lambda_j^2} \left\{ 3 + 2 \sum_{\substack{i=1 \\ i<j}}^{q} \sum_{j=1}^{q} \lambda_i \lambda_j \right\}.$$

Also

$$E(X_t X_{t-q}) = \mu^2 + \sigma^2 \theta_q \mu$$

$$E(X_t X_{t-k}) = \mu^2 \quad , k > q$$

Thus

$$R(k) = 0 \ , \ k > q$$

When $q = 3$, we obtain the following results

$$E(X_t^2) = \frac{\sigma^2}{(1 - \lambda_1^2 - \lambda_2^2 - \lambda_3^2)} \{ 1 + 2\lambda_1^2 + 2\lambda_2^2 + 2\lambda_3^2 + 2\lambda_1\lambda_2$$

$$+ 2\lambda_1\lambda_3 + 2\lambda_2\lambda_3 \}$$

$$E(X_t X_{t-1}) = (\lambda_1 \lambda_2 + \lambda_2 \lambda_3)E(X_t^2) + \sigma^2 \lambda_1(\lambda_1 + \lambda_2 + \lambda_3)$$

$$+ \sigma^2(\lambda_1 \lambda_2 + \lambda_2 \lambda_3) + \mu^2$$

$$E(X_t X_{t-2}) = \lambda_1 \lambda_3 E(X_t^2) + \sigma^2 \lambda_2(\lambda_1 + \lambda_2 + \lambda_3) + \sigma^2 \lambda_1 \lambda_3 + \mu^2$$

$$E(X_t X_{t-3}) = \sigma^2 \lambda_3 (\lambda_1 + \lambda_2 + \lambda_3) + \mu^2$$

$$E(X_t X_{t-k}) = \mu^2 \quad , k > 3$$

Thus

$$R(0) = \frac{\sigma^2}{(1 - \lambda_1^2 - \lambda_2^2 - \lambda_3^2)}\left\{1 + (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)\{1 + (\lambda_1 + \lambda_2 + \lambda_3)^2\}\right\}$$

$$R(1) = \frac{\sigma^2}{(1 - \lambda_1^2 - \lambda_2^2 - \lambda_3^2)}\Big\{\lambda_1(\lambda_1 + \lambda_2 + \lambda_3)(1 - \lambda_1^2 - \lambda_2^2 - \lambda_3^2)$$

$$+ (\lambda_1 \lambda_2 + \lambda_2 \lambda_3)(2 + (\lambda_1 + \lambda_2 + \lambda_3)^2)\Big\}$$

$$R(2) = \frac{\sigma^2}{(1 - \lambda_1^2 - \lambda_2^2 - \lambda_3^2)}\Big\{\lambda_2(\lambda_1 + \lambda_2 + \lambda_3)(1 - \lambda_1^2 - \lambda_2^2 - \lambda_3^2)$$

$$+ \lambda_1 \lambda_3(2 + (\lambda_1 + \lambda_2 + \lambda_3)^2)\Big\}$$

$$R(3) = \sigma^2 \lambda_3(\lambda_1 + \lambda_2 + \lambda_3)$$

$$R(k) = 0 \; , \; k > 3$$

Our numerical illustration consists of 500 points generated from the process (3.5.4) with $q = 3$ and

$$\theta_1 = 0.40$$

$$\theta_2 = 0.30$$

$$\theta_3 = 0.20$$

$$\sigma^2 = 1$$

Using these values of θ_1, θ_2, θ_3 and σ^2, we obtain the following

$\rho(\Gamma) = 0.50$

$\mu = 0.900$

$R(0) = 2.1477$

$R(1) = 1.0724$

$R(2) = 0.5866$

$R(3) = 0.1800$

$R(k) = 0 \quad , k > 3$

Figure 3.4 gives a graph of the data; columns 8 and 9 of Table 3.4 give the theoretical and sample autocovariances respectively. Columns 8 and 9 of Table 3.5 give the sample autocorrelations and partial autocorrelations respectively. Column 8 of Table 3.5 suggests that $p = 0$ and $q = 2$ or 3. However, on the basis of the information criterion of Akaike [2], the ARMA (0, 3) model

$$X_t = 0.935 + 0.438a_{t-1} + 0.339a_{t-2} + 0.119a_{t-3} + a_t$$

with $E(a_t) = 0$ and $Var(a_t) = 1.716$ provided a better fit. This ARMA (0, 3) model leads to 72 per cent increase in the error variance.

<u>EXAMPLE 3.5.</u> Let us consider a bilinear process whose covariance structure is the same as some autoregressive process. We consider the BARMA (2, 0, 3, 1) process X_t, $t \in Z$ satsifying

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + b_1 X_{t-2}e_{t-1} + b_2 X_{t-3}e_{t-1} + e_t \qquad \text{a.e } [P] \quad (3.5.5)$$

for every t in Z where e_t, $t \in Z$ are independent and each e_t is distributed as $N(0, \sigma^2)$.

The A, B, C and Γ matrices are identified to be

$$\underset{3x3}{A} = \begin{pmatrix} a_1 & a_2 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

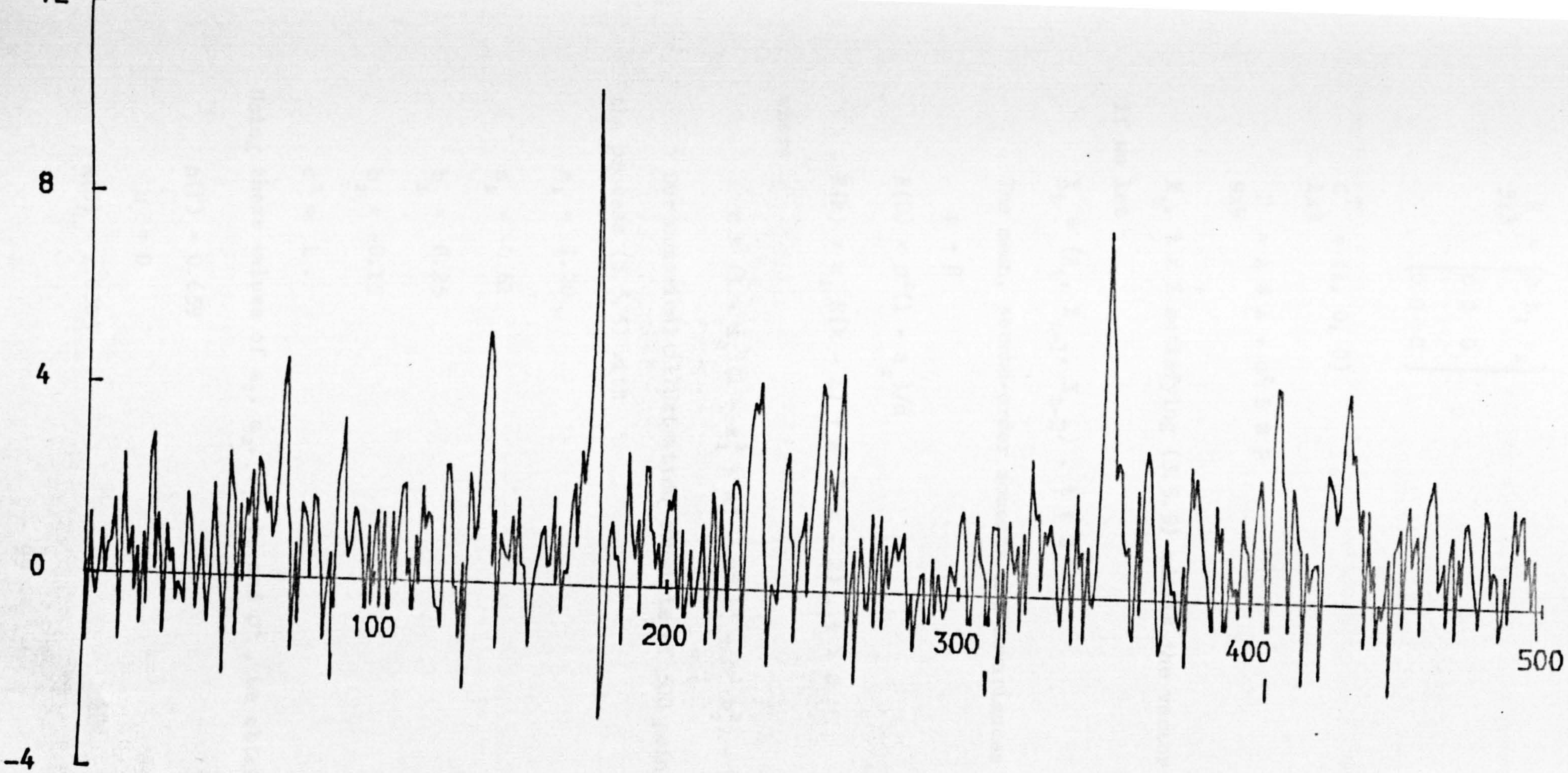


FIG.3.4 500 OBSERVATIONS FROM THE BARMA (0,0,3,3) PROCESS IN EXAMPLE 3.4

$$\underset{3\times 3}{B} = \begin{pmatrix} 0 & b_1 & b_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\underset{1\times 3}{C^T} = (1, 0, 0)$$

$$\underset{9\times 9}{\Gamma} = A \otimes A + \sigma^2 B \otimes B$$

X_t, $t \in Z$ satisfying (3.5.5) admits the vector representation (3.4.3) if we let

$$\underline{X}_t = (X_t, X_{t-1}, X_{t-2}) , \ t \in Z$$

The mean, second-order moments and covariances are then given by

$$\mu = 0$$

$$R(0) = \sigma^2(1 - a_2)/d$$

$$R(k) = a_1 R(k - 1) + a_2 R(k - 2) , \ k > 0$$

where

$$d = (1 - a_2)(1 - a_1^2 - a_2^2 - \sigma^2 b_1^2 - \sigma^2 b_2^2) - 2 a_1(a_1 a_2 + \sigma^2 b_1 b_2)$$

Our numerical illustration consists of 500 points generated from the process (3.5.5) with

$a_1 = 1.20$

$a_2 = -0.61$

$b_1 = 0.25$

$b_2 = -0.15$

$\sigma^2 = 1$.

Using these values of a_1, a_2, b_1, b_2 and σ^2 , we obtain the following

$\rho(\Gamma) = 0.659$

$\mu = 0$

$R(0) = 4.0003$

$$R(k) = 1.20R(k - 1) - 0.61R(k - 2)\ ,\ k > 0$$

Figure 3.5 gives a graph of the data; columns 10 and 11 of Table 3.4 give the theoretical and sample autocovariances respectively. Columns 10 and 11 of Table 3.5 give the estimated autocorrelations and partial autocorrelations respectively, while Table 3.8 shows portions of the R- and S-arrays at $f_m = (-1)^m \hat{\rho}_m$. From $S_2(m)$ and $R_3(m)$ columns, it is clear that p = 2 and q = 0. The ARMA (2. 0) fitted to the data is

$$X_t = 1.206X_{t-1} - 0.605X_{t-2} + a_t$$

with $E(a_t) = 0$ and $Var(a_t) = 1.116$, leading to an 11.6 per cent increase in the error variance.

Summary 3.6

In the above examples, we fitted ARMA models using Box-Jenkins method in some cases and R- and S-arrays method in the remaining cases for known bilinear models. The primary purpose of this study is to examine how the error variance increases with wrong model fitting. The findings are summarized in the following table.

TABLE 3.9 COMPARISON OF ERROR VARIANCES OF LINEAR AND BILINEAR MODELS

True Model	Fitted Model	% Increase In Error Variance
BARMA(1, 1, 1, 1)	ARMA(1, 1)	60.9
BARMA(2. 0. 2. 2)	AR(2)	100
BARMA(0, 2, 2, 2)	MA(2)	96
BARMA(0, 0, 3, 3)	MA(3)	72
BARMA(2, 0, 3, 1)	AR(2)	11.6

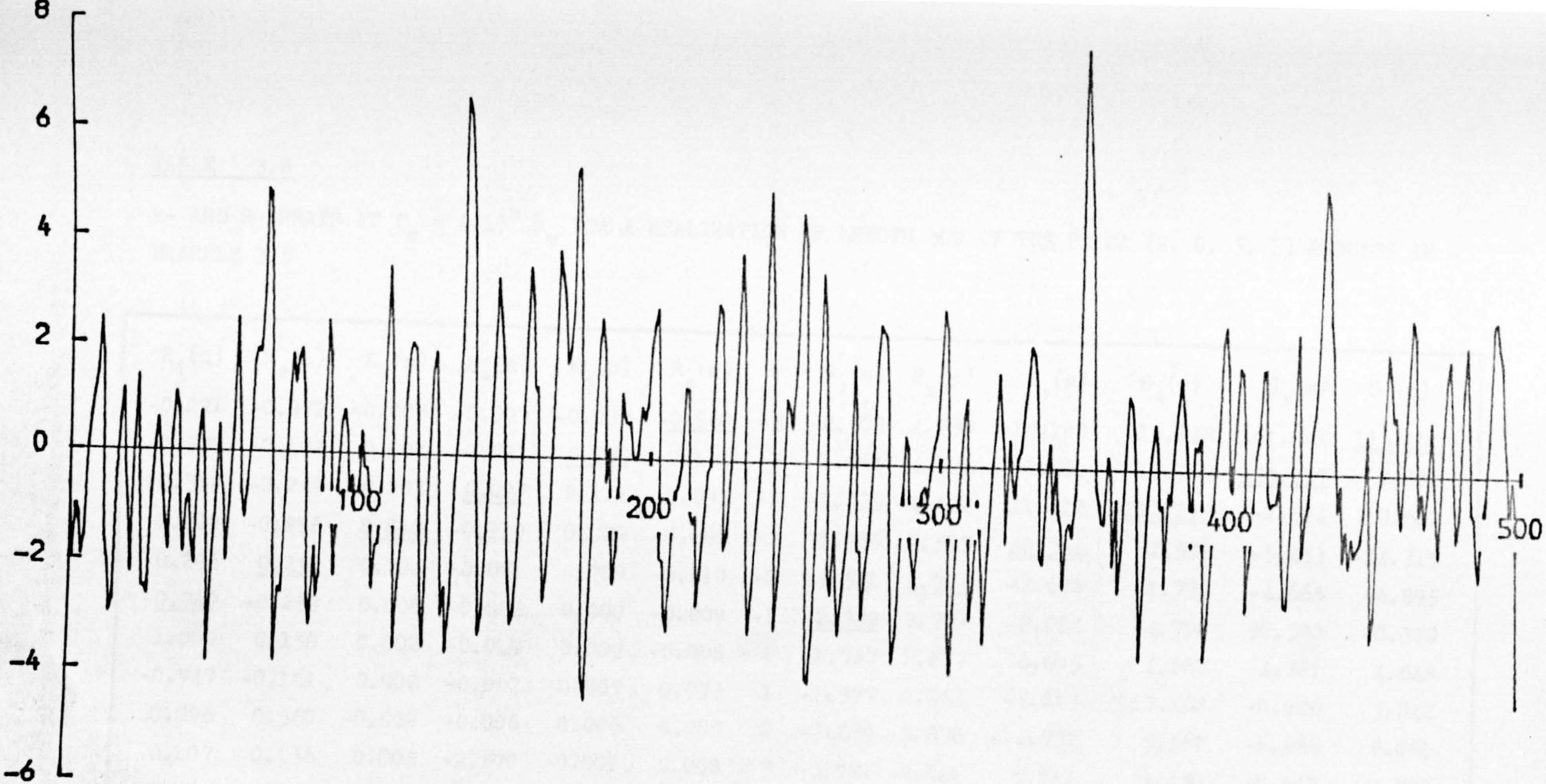


FIG.3.5 500 OBSERVATIONS FROM THE BARMA (2,0,3,1) PROCESS IN EXAMPLE 3.5

TABLE 3.8

R- AND S-ARRAYS AT $f_m = (-1)^m \hat{\rho}_m$ FOR A REALIZATION OF LENGTH 500 OF THE BARMA (2, 0, 3, 1) PROCESS IN EXAMPLE 3.5

$R_1(m)$	$R_2(m)$	$R_3(m)$	$R_4(m)$	$R_5(m)$	$R_6(m)$	m	$S_1(m)$	$S_2(m)$	$S_3(m)$	$S_4(m)$	$S_5(m)$	$S_6(m)$
-0.221	-0.072	-0.003	-0.003	-0.037	-0.060	-6	-2.503	4.738	-10.102	-10.143	161.041	113.102
0.332	0.102	0.004	-0.041	0.273	-0.114	-5	-1.961	4.598	-0.882	-19.362	125.797	2.582
-0.319	-0.249	0.003	0.027	0.111	0.003	-4	-1.335	4.532	11.612	-212.293	-2.624	8.622
0.107	-0.886	0.010	-0.110	0.002	-0.002	-3	1.768	4.517	42.364	2.576	-5.453	-24.315
0.296	0.251	0.104	-0.001	0.009	-0.010	-2	-3.521	4.704	-2.608	-1.736	-4.665	-6.895
-0.747	-0.253	0.006	-0.003	0.000	-0.009	-1	-2.339	2.779	-2.062	4.798	92.582	-0.379
1.000	0.150	0.002	-0.019	0.009	-0.008	0	-1.747	2.897	0.695	4.687	4.721	4.648
-0.747	-0.161	0.002	-0.002	0.007	0.033	1	-1.397	2.864	-5.215	-17.468	-0.926	5.042
0.296	0.560	-0.002	-0.006	0.006	0.007	2	-0.639	2.876	-10.732	5.462	-6.286	8.061
0.107	0.156	0.005	-2.590	-0.001	0.008	3	-3.981	2.641	0.011	5.450	40.647	4.702
-0.319	-0.061	0.005	0.003	-0.011	0.011	4	-2.040	2.601	-5.445	3.635	10.794	-25.530
0.332	0.034	-0.004	-0.002	-0.015	-0.086	5	-1.665	2.994	-2.113	24.094	2.934	-3.636
-0.221	-0.039	-0.002	0.008	-0.013	-0.071	6	-1.407	3.323	-4.672	-23.130	21.778	-66.479

CHAPTER 4

ON PURELY BILINEAR PROCESSES THAT ARE WHITE NOISE AND INVERTIBILITY OF BILINEAR PROCESSES

4.1 INTRODUCTION

In this chapter, we analyse in some detail a class of purely bilinear processes that appear to be white noise under second-order analysis (analysis based just on first-and second-order moments only). In our definition, to be given in section 4.2, such a process is said to be purely bilinear white noise. We are interested in the bilinear white noise process X_t, $t \in Z$ satisfying

$$X_t = e_t + \left(\sum_{j=1}^{m} b_j X_{t-q-j} \right) e_{t-q} \qquad \text{a.e } [P] \quad (4.1.1)$$

for every t in Z for some $q > 0$ and constants $b_1, b_2, \ldots, b_m$ where e_t, $t \in Z$ is a sequence of independent identically distributed random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Purely bilinear white noises are indeed very useful in that they could be used to modify or extend linear models to bilinear models. We will return to this use of purely bilinear white noise in chapter 5.

Also considered in this chapter is the classical invertibility problem for bilinear processes. Some simple invertibility conditions are derived for the second-order stationary process X_t, $t \in Z$ satisfying

$$X_t = a X_{t-1} + b e_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P] \quad (4.1.2)$$

for every t in Z for some real numbers a, b and β where e_t, $t \in Z$ is a sequence of independent identically distributed real random variables with $E(e_t) = 0$, $E(e_t^2) = \sigma^2$ and $E(e_t^4) < \infty$. The classical invertibility problem of the process X_t, $t \in Z$ satisfying (4.1.1), and of the process X_t, $t \in Z$ satisfying (2.3.9) are also studied.

Finally, we consider the problem of distinguishing a purely bilinear white noise from a pure white noise. See definition of pure white noise in section 4.2. We show that the Bispectral density function analysis method of Subba Rao and Gabr [38] works true for X_t, $t \in Z$ satisfying (4.1.1) contrary to an earlier opinion expressed by Granger and Andersen [15, p.43] in relation to the single term model

$$X_t = \beta X_{t-2}e_{t-1} + e_t \qquad \text{a.e } [P] \quad (4.1.3)$$

The Granger and Andersen [15] method of performing second-order analysis on $Y_t = X_t^2$, $t \in Z$ was considered for the more general purely bilinear white noise given by (4.1.1).

4.2 PURELY BILINEAR PROCESSES AND WHITE NOISE

Let e_t, $t \in Z$ be a sequence of random variables with the following properties.

(i) $E(e_t) = 0$ for all $t \in Z$.

(ii) $E(e_t^2) = \sigma^2 < \infty$ for all $t \in Z$

(iii) $R(k) = E(e_t e_{t+k})$

$= 0$ for all $t \in Z$ and $k \in Z$ with $k \neq 0$

Such a process is called White Noise. A simple example of white noise is a sequence of independent identically distributed random variables with common mean 0 and variance $\sigma^2 < \infty$. This type of process is usually assumed to have a normal distribution and is called 'pure white noise'.

In this section, we study some pure bilinear processes which are white noises. We introduce a definition.

Definition 4.2.1. Let e_t, $t \in Z$ be a sequence of independent identically distributed random variables with common mean 0 and variance $\sigma^2 < \infty$. The second-order purely bilinear process X_t, $t \in Z$ satsifying

$$X_t = e_t + \sum_{\substack{i=1 \\ i \geq j}}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-i}e_{t-j} \qquad \text{a.e } [P] \quad (4.2.1)$$

for every t in Z for some constants β_{ij}'s is said to be purely bilinear white noise if

$$E(X_t) = 0 \text{ for all } t \in Z$$

$$R(k) = E(X_t X_{t+k})$$

$$= 0 \text{ for all } t,k \in Z \text{ with } k \neq 0$$

Granger and Andersen [15, p.42] gave the following example of a purely bilinear process whose covariance structure is identical with the covariance structure of some suitable white noise.

$$X_t = e_t + \beta e_{t-k} X_{t-\ell} \qquad \text{a.e } [P] \quad (4.2.2)$$

for every t in Z for some β and $\ell > k$.

We show that for a more general class of purely bilinear processes, the above phenomenon still rings true.

Theorem 4.2.2. Let e_t, $t \in Z$ be a sequence of independent identically distributed random variables with common mean 0 and variance $\sigma^2 < \infty$, ie, e_t, $t \in Z$ is pure white noise. If there is a second-order stationary process X_t, $t \in Z$ satisfying

$$X_t = e_t + \left(\sum_{j=1}^{m} b_j X_{t-q-j} \right) e_{t-q} \qquad \text{a.e } [P] \quad (4.2.3)$$

for some $q > 0$ and $b_1, b_2, \ldots, b_m$ constants, and also satisfying the strict stationarity condition $\rho(\Gamma) < 1$ of Theorem 2.4.3, then X_t, $t \in Z$ is purely bilinear white noise.

Proof. We put (4.2.3) in vector form as follows. Let

$$p = m + 1$$

$$\underset{p\times p}{A} = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix}$$

$$\underset{p\times p}{B} = \begin{pmatrix} 0 & b_1 & b_2 & \cdots & b_m \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & 0 \end{pmatrix}$$

$$\underset{1\times p}{C^T} = (1, 0, 0, \ldots, 0)$$

and

$$\underset{1\times p}{\underline{X}_t^T} = (X_t, X_{t-1}, \ldots, X_{t-p+1}), \ t \in Z$$

Then

$$\underline{X}_t = A\,\underline{X}_{t-1} + B\,\underline{X}_{t-q}\,e_{t-q} + Ce_t \qquad \text{a.e } [P] \quad (4.2.4)$$

for every t in Z. Γ of Theorem 2.4.3 works out to be

$$\Gamma = A \otimes A + \sigma^2 B \otimes B\ , \text{ if } q = 1$$

$$= \begin{pmatrix} A \otimes A & \underline{0} & \underline{0} & \cdots & \underline{0} & \sigma^2 B \otimes B \\ I_p^2 & \underline{0} & \underline{0} & \cdots & \underline{0} & \underline{0} \\ \underline{0} & I_p^2 & \underline{0} & \cdots & \underline{0} & \underline{0} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \underline{0} & \underline{0} & \underline{0} & & I_p^2 & \underline{0} \end{pmatrix}$$

$$\text{if } q > 1$$

Hence, in the framework of the model (2.4.5), $B_1 = B_2 = \ldots = B_{q-1} = \underline{0}$ and $B_q = B$. We easily check that $B\,C = \underline{0}$. From Theorem 2.4.3, we observe that

$$\underline{\mu} = E(\underline{X}_t) = \sigma^2(I_p - A)^{-1}\left[\sum_{j=1}^{q} B_j\right]C$$

$$= \sigma^2(I_p - A)^{-1}\,B\,C$$

$$= 0$$

This can be easily verified directly by working with the model (4.2.3).

Now, let $V = E(\underline{X}_t \underline{X}_t^T)$. This is the matrix of variances and covariances of $X_t, X_{t-1}, \dots, X_{t-m}$. From Theorem 2.4.3, this matrix V satisfies the equation

$$V = A V A^T + \sigma^2 B V B^T + \sigma^2 C C^T \qquad (4.2.5)$$

This equation can be solved easily. We find that

$$\begin{aligned} R_x(k) &= Cov(X_t, X_{t-k}) \\ &= 0 \text{ for } k = 1, 2, \dots, m \\ &= \sigma^2/(1 - \sum_{j=1}^{m} \sigma^2 b_j^2) \text{ for } k = 0 \end{aligned} \qquad (4.2.6)$$

By considering the expressions $E(\underline{X}_{t+k} \underline{X}_t^T)$, $k > 0$ in the proof of Theorem 3.3.1, we have

$$R_x(k) = 0 \text{ for all } k = 1, 2, \dots .$$

This shows that the process X_t, $t \in Z$ is white noise. The above assertion can also be worked out by using directly (4.2.3).

In section 3.4, we have remarked that for any purely bilinear process, there exists a moving average process with identical covariance structrues. There are purely bilinear processes which are not white noise. The following process is an example.

$$X_t = e_t + \sum_{j=1}^{q} \beta_j X_{t-j} e_{t-j} \qquad \text{a.e } [P]$$

for every t in Z. See section 3.5. In view of this, one might wonder whether processes X_t, $t \in Z$ satsifying

$$X_t = e_t + \sum_{\substack{i=1 \\ i>j}}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-i} e_{t-j} \qquad \text{a.e } [P] \quad (4.2.7)$$

for all t in Z would be white noise. Not always. Granger and Andersen [15, p.42] have given the following example

$$X_t = e_t + \beta_1 e_{t-1} X_{t-2} + \beta_2 e_{t-2} X_{t-3} \qquad \text{a.e } [P]$$

for every t in Z.

Granger and Andersen [15, p.42] have given a condition under which (4.2.7) will appear to be white noise. However, this condition is a mere statement which is not substantiated.

We would like to work out the bispectrum of the model considered in Theorem 4.2.2 in section 4.4. Lemma 4.2.3, to be stated below, will be useful in section 4.4. Before that we want to make some comments on vector difference equations of order 1.

Suppose $\underline{Y}_t$, $t = 0, 1, 2, \ldots$ is a sequence of vectors each of order pxl and satisfies the following difference equation

$$\underline{Y}_t = A\,\underline{Y}_{t-1} + \underline{b}$$

for $t = 1, 2, 3, \ldots\ldots$, for some matrices $\underset{p\times p}{A}$ and $\underset{p\times 1}{\underline{b}}$. We can always solve this equation and express $\underline{Y}_t$ as a function of t, $\underline{Y}_0$, A and b. A necessary and sufficient condition for this solution to be independent of t is that $(I_p - A)$ is invertible, or, equivalently, $\rho(A) < 1$.

Lemma 4.2.3. Let e_t, $t \in Z$ be a sequence of independent identically distributed random variables with mean 0 and variance $\sigma^2 < \infty$. Suppose there exists a second-order stationary process X_t, $t \in Z$ satisfying

$$X_t = e_t + \left(\sum_{j=1}^{m} b_j\, X_{t-q-j}\right) e_{t-q} \qquad \text{a.e } [P] \quad (4.2.8)$$

for every t in Z for some constants $b_1, b_2, b_3, \ldots, b_m$ and $q > 0$. Let Γ be the matrix associated with the vector-valued process representation of the above process, ie,

$$\underline{X}_t = A\,\underline{X}_{t-1} + B\,\underline{X}_{t-q} e_{t-q} + e_t \qquad \text{a.e } [P]$$

for every t in Z, where A, B and Γ are as given in Theorem 4.2.2. Assume that the strict stationarity condition $\rho(\Gamma) < 1$ holds. Let

$$\underset{r\times r}{\nabla} = \begin{pmatrix} \overbrace{0\ 0\ 0\ \cdots\ 0}^{q} & \sigma^2 b_1^2 & \sigma^2 b_2^2 & \cdots & \sigma^2 b_{m-1}^2 & \sigma^2 b_m^2 \\ 1\ 0\ 0\ \cdots\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0\ 1\ 0\ \cdots\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0\ 0\ 0\ \cdots\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

where $r = q + m$.

Then $\rho(\nabla) < 1$.

Proof. Since X_t, $t \in Z$ is second-order stationary,

$E(X_t^2)$ = Constant, for all t in Z.

Let us calculate $E(X_t^2)$. Squaring both sides of (4.2.8) we obtain

$$X_t^2 = e_t^2 + 2e_t e_{t-q}\left[\sum_{j=1}^{m} b_j X_{t-q-j}\right] + e_{t-q}^2\left\{\sum_{j=1}^{m} b_j^2 X_{t-q-j}^2 + 2\sum_{\substack{i=1 \\ i<j}}^{m}\sum_{j=1}^{m} b_i b_j X_{t-q-i} X_{t-q-j}\right\} \tag{4.2.9}$$

But

$$E(X_{t-q-i} X_{t-q-j}) = \mathrm{Cov}(X_{t-q-i}\ X_{t-q-j})$$
$$= 0, \text{ from Theorem 4.2.2.}$$

If we now take expectations on both sides of (4.2.9) we obtain

$$E(X_t^2) = \sigma^2 + \sum_{j=1}^{m} \sigma^2 b_j^2 E(X_{t-q-j}^2) \tag{4.2.10}$$

We attempt to put (4.2.10) in the form of a first-order difference equation. Let

$r = q + m$

$$\underset{1\times r}{\underline{W}_t^T} = (E(X_t^2), E(X_{t-1}^2), \ldots, E(X_{t-r+1}^2))$$

$$H^T_{1\times r} = (1, 0, 0, \dots , 0)$$

With this notation, we can write (4.2.10) as the first-order vector difference equation

$$\underline{W}_t = \nabla \underline{W}_{t-1} + \sigma^2 H \qquad (4.2.11)$$

for $t = 1, 2, \dots$.

Because of second-order stationarity of X_t, $t \in Z$, $\underline{W}_t = \underline{W}_{t-1}$ for all t. Consequently, $\rho(\nabla) < 1$. See the remarks preceding Lemma 4.2.3.

4.3 INVERTIBILITY

Suppose X_t, $t \in Z$ and e_t, $t \in Z$ are two stochastic processes satisfying

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + \sum_{i=1}^{m} \sum_{\substack{j=1 \\ i \geq j}}^{\ell} \beta_{ij} X_{t-i} e_{t-j} + e_t \quad \text{a.e } [P]$$

for every t in Z, and for some constants $a_1, a_2, \dots , a_r, b_1, b_2, \dots , b_h$ and β_{ij}, $1 \leq i \leq m$, $1 \leq j \leq \ell$ with $i \geq j$. It is natural to express X_t purely as a function of $e_t, e_{t-1}, e_{t-2}, \dots$, for every t in Z. The results of chapter 2 do attempt to achieve this. Under some conditions on the process e_t, $t \in Z$ and the coefficients $a_1, a_2, \dots , a_r, b_1, b_2, \dots , b_h$ and β_{ij}, $1 \leq i \leq m$, $1 \leq j \leq \ell$, $i \geq j$; X_t, $t \in Z$ is indeed written as a function of $e_t, e_{t-1}, e_{t-2}, \dots$. It is natural to enquire whether one can express e_t purely as a function of $X_t, X_{t-1}, X_{t-2}, \dots$ for every t in Z. This is the classical invertibility problem. We do not know of any nice conditions under which invertibility holds. For some simple models, described in the following theorems, we give some simple invertibility conditions. The result of Theorem 4.3.1, to be stated below, generalizes a result of Tuan Dinh Pham and Lanh Tat Tran [41, p. 622].

Granger and Andersen [16] introduced a notion of invertibility which they claim is relevant to both linear and non-linear time series models. Hallin [19] has studied the relationship between classical invertibility, Granger - Andersen invertibility and what he calls generalized invertibility. According to Hallin [19] the three invertibility concepts are equivalent with respect to Linear ARMA models with constant coefficients. We confine ourselves to the classical concept.

We now give some invertibility conditions.

Theorem 4.3.1. Let e_t, $t \in Z$ be a sequence of independent identically distributed random variables with $E(e_t^4) < \infty$. Let a and β be two real numbers such that

$$a^2 + \beta^2 \sigma^2 < 1.$$

Then the bilinear process X_t, $t \in Z$ satsifying

$$X_t = a X_{t-1} + b e_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P] \quad (4.3.1)$$

for every t in Z for some real number b, is invertible if

$$E \ln|b + \beta X_t| < 0.$$

Proof. Repeated use of (4.3.1) gives

$$e_t = X_t - a X_{t-1} + \sum_{r=1}^{n-1} (-1)^r \prod_{j=1}^{r} (b + \beta X_{t-j})(X_{t-r} - a X_{t-r-1})$$

$$+ (-1)^n \prod_{j=1}^{n} (b + \beta X_{t-j}) e_{t-n} \qquad (4.3.2)$$

for all t in Z and n = 1, 2, 3, ...

From the expression (4.3.2), we can express e_t purely as a function of X_t, X_{t-1}, ... , if we can show

$$\lim_{n\to\infty} \prod_{j=1}^{n} (b + \beta X_{t-j}) e_{t-n} = 0 \qquad \text{a.e } [P]$$

Since e_t, $t \in Z$ are identically distributed, it is enough to show that

$$\lim_{n\to\infty} \prod_{j=1}^{n} (b + \beta X_{t-j}) = 0$$

We observe the following facts.

(1) X_t, $t \in Z$ is ergodic. See Remarks 2.3.3(6)

(2) $\log|b + \beta X_{t-j}|$, $j = 1, 2, 3, \ldots$, is ergodic.

Now, let

$$P(n, t) = \prod_{j=1}^{n} (b + \beta X_{t-j})$$

Taking logarithms, we obtain

$$\frac{1}{n} \log|P(n, t)| = \frac{1}{n} \sum_{j=1}^{n} \log|b + \beta X_{t-j}|$$

By the ergodic theorem,

$$\lim_{n\to\infty} \frac{1}{n} \log|P(n, t)| = E \log|b + \beta X_t| < 0 \qquad \text{a.e } [P]$$

Hence,

$$\lim_{n\to\infty} \prod_{j=1}^{n} (b + \beta X_{t-j}) = 0 \qquad \text{a.e } [P]$$

This completes the proof.

REMARK 4.3.2. The condition

$$E \log|b + \beta X_t| < 0$$

given in Theorem 4.3.1 involves the distribution of X_t, $t \in Z$. We do not yet know the distribution of X_t for a given distribution of e_t. Thus, it is virtually impossible to characterize all values of a, b, β and σ^2 for which the second-order stationary process X_t, $t \in Z$ satisfying (4.3.1) is invertible.

We can obtain a sufficient condition.

$$E \log|b + \beta X_t| = \tfrac{1}{2} E \log (b + \beta X_t)^2$$

$$< \tfrac{1}{2} \log E(b + \beta X_t)^2 \text{ , by Jensen's inequality.}$$

$$< \tfrac{1}{2} \log(b^2 + 2 b \beta E(X_t) + \beta^2 E(X_t^2))$$

Hence a sufficient condition for invertibility is

$$b^2 + 2 b \beta E(X_t) + \beta^2 E(X_t^2) < 1 \qquad (4.3.3)$$

In case the e_t, $t \in Z$ are Gaussian, then

$$\mu = E(X_t) = \sigma^2 \beta/(1 - a)$$

$$E(X_t^2) = \frac{\sigma^2}{1 - a^2 - \sigma^2 \beta^2} \left\{ 1 + b^2 + 2 a b + 2 \beta \mu (1 + a + b) \right\}$$

See Example 3.1 of section 3.5.

When $b = 0$, we obtain from (4.3.3) that a sufficient condition for the invertibility of the process X_t, $t \in Z$ satisfying

$$X_t = a X_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P]$$

is

$$\beta^2 E(X_t^2) < 1.$$

This agrees with the condition obtained by Granger and Andersen [15 , p.74], Subba Rao [37 , p.249] and Tuan Dinh Pham and Lanh Tat Tran [41 , p.622].

The invertibility problem for the more general model (2.3.9) will be discussed in Theorem 4.3.4.

In section 4.2, we initiated the study of the bilinear model X_t, $t \in Z$ satisfying

$$X_t = e_t + \left(\sum_{j=1}^{m} b_j X_{t-q-j} \right) e_{t-q} \qquad \text{a.e } [P]$$

for every t in Z, for some $q > 0$, for some sequence e_t, $t \in Z$ of independent identically distributed real random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$ and constants $b_1, b_2, \dots , b_m$. Next, we study the (classical) invertibility problem of this process.

Theorem 4.3.3. Let e_t, $t \in Z$ be a sequence of independent identically distributed random variables with $E(e_t) = 0$ and $E(e_t^4) < \infty$. Let $\rho(\Gamma) < 1$, where Γ is given in Theorem 4.2.2 for given $b_1, b_2, \ldots, b_m$ and $\sigma^2 = E(e_t^2)$. Then the second-order strictly stationary process X_t, $t \in Z$ satisfying

$$X_t = e_t + \left\{\sum_{j=1}^{m} b_j X_{t-q-j}\right\} e_{t-q} \qquad \text{a.e } [P] \quad (4.3.4)$$

for every t in Z, is invertible if

$$2 \sum_{j=1}^{m} b_j \sigma^2 < 1 \qquad (4.3.5)$$

<u>Proof</u>. Repeated use of (4.3.2) gives

$$e_t = X_t + \sum_{r=1}^{n-1} (-1)^r \prod_{i=1}^{r} \left\{\sum_{j=1}^{m} b_j X_{t-iq-j}\right\} X_{t-rq}$$
$$+ (-1)^n \prod_{i=1}^{n} \left\{\sum_{j=1}^{r} b_j X_{t-iq-j}\right\} e_{t-nq} \qquad (4.3.6)$$

for all t in Z and n = 1. 2, 3, ...

From the expression (4.3.6), we can express e_t purely as a function of $X_t, X_{t-1}, X_{t-2}, \ldots$, if we can show

$$\lim_{n\to\infty} \prod_{i=1}^{n} \left\{\sum_{j=1}^{m} b_j X_{t-iq-j}\right\} e_{t-nq} = 0 \qquad \text{a.e } [P]$$

Since e_t, $t \in Z$ are identically distributed, it is enough to show that

$$\lim_{n\to\infty} \prod_{i=1}^{n} \left\{\sum_{j=1}^{m} b_j X_{t-iq-j}\right\} = 0 \qquad \text{a.e } [P]$$

We observe the following facts.

(1) X_t, $t \in Z$ is ergodic. See Remark 2.4.4 (2)

(2) Fix $t \in Z$. Then

$\sum_{j=1}^{m} b_j X_{t-q-j}$, $\sum_{j=1}^{m} b_j X_{t-2q-j}$, $\sum_{j=1}^{m} b_j X_{t-3q-j}$, ... , is ergodic.

(3) $\text{Var}\left(\sum_{j=1}^{m} b_j X_{t-q-j}\right) = E\left(\sum_{j=1}^{m} b_j X_{t-q-j}\right)^2$ (because mean is 0)

$= \sum_{j=1}^{m} b_j^2 E(X_t^2)$ (because X_t is white noise. See Theorem 4.2.2)

$= \dfrac{\sum_{j=1}^{m} b_j^2 \sigma^2}{1 - \sum_{j=1}^{m} b_j^2 \sigma^2}$ from 4.2.6

(4) Fix t in Z. Then

$$\log\left|\sum_{j=1}^{m} b_j X_{t-q-j}\right|,\ \log\left|\sum_{j=1}^{m} b_j X_{t-2q-j}\right|,\ \log\left|\sum_{j=1}^{m} b_j X_{t-3q-j}\right|, \ldots$$

is ergodic.

(5) $E \log\left|\sum_{j=1}^{m} b_j X_{t-q-j}\right| = \frac{1}{2} E \log\left(\sum_{j=1}^{m} b_j X_{t-q-j}\right)^2$

$\leqslant \frac{1}{2} \log E\left(\sum_{j=1}^{m} b_j X_{t-q-j}\right)^2$

by Jensen's inequality

$\leqslant \frac{1}{2} \log \left(\dfrac{\sum_{j=1}^{m} b_j^2 \sigma^2}{1 - \sum_{j=1}^{m} b_j^2 \sigma^2}\right)$

$< \frac{1}{2} \log 1 = 0$, by (4.3.5)

Now, we look at

$$\frac{1}{n} \sum_{i=1}^{n} \log\left|\sum_{j=1}^{m} b_j X_{t-iq-j}\right|,\ n \geqslant 1$$

By the ergodic theorem, the above sequence has a constant limit almost surely. This limit is < 0. Hence

$$\lim_{n\to\infty} \prod_{i=1}^{n} \left(\sum_{j=1}^{m} b_j X_{t-iq-j}\right) = 0 \qquad \text{a.e } [P]$$

This completes the proof.

Next, we study the classical invertibility problem of the process X_t, $t \in Z$ satisfying

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + be_{t-1} + \left(\sum_{j=1}^{\ell} \beta_{j1} X_{t-j}\right) e_{t-1} + e_t \quad \text{a.e } [P]$$

for every t in Z, for some sequence e_t, $t \in Z$ of independent identically distributed real random variables with $E(e_t) = 0$, $E(e_t^2) = \sigma^2 < \infty$ and constants $a_1, a_2, \ldots, a_r$ $\beta_{11}, \beta_{21}, \ldots, \beta_{\ell 1}$.

<u>Theroem 4.3.4</u>. Let e_t, $t \in Z$ be a sequence of independent identically distributed random varaibles with $E(e_t) = 0$ and $E(e_t^4) < \infty$. Let $\underset{p\times p}{A}$, $\underset{p\times p}{B}$, $\underset{p\times 1}{C}$, $\underset{p\times 1}{b}$ be the matrices of Remarks 2.3.3 (4). Let

$\rho(A \otimes A + \sigma^2 B \otimes B) < 1$

where $\sigma^2 = E(e_t^2)$.

Then the bilinear process X_t, $t \in Z$ satisfying

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + be_{t-1} + \left(\sum_{j=1}^{\ell} \beta_{j1} X_{t-j}\right) e_{t-1} + e_t \quad \text{a.e } [P] \quad (4.3.7)$$

for every t in Z, is invertible if

$$E \log \left| b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-j} \right| = 0$$

<u>Proof</u>. Repeated use of (4.3.7) gives

$$e_t = X_t - \sum_{j=1}^{r} a_j X_{t-j}$$
$$+ \sum_{m=1}^{n-1} (-1)^m \prod_{i=1}^{m} (b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-m+1-j})(X_{t-m} - \sum_{j=1}^{r} a_j X_{t-m-j})$$
$$+ (-1)^n \prod_{m=1}^{n} (b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-m+1-j}) e_{t-n} \quad (4.3.8)$$

for all t in Z and n = 1, 2, 3, ...

From the expression (4.3.8), we can express e_t, $t \varepsilon Z$ purely as a function of $X_t, X_{t-1}, X_{t-2}, \ldots$, if we can show

$$\lim_{n\to\infty} \prod_{m=1}^{n} (b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-m+1-j}) e_{t-n} = 0 \qquad \text{a.e } [P]$$

Since e_t, $t \varepsilon Z$ are identically distributed, it is enough to show that

$$\lim_{n\to\infty} \prod_{m=1}^{n} (b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-m+1-j}) = 0$$

We observe the following facts.

(1) X_t, $t \varepsilon Z$ satisfying (4.3.7) is ergodic. See Theorem 2.3.6.

(2) Fix $t \varepsilon Z$. Then $(b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-j})$, $(b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-1-j})$, $(b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-2-j})$, ... , is ergodic .

(3) Fix $t \varepsilon Z$. Then $\log|b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-j}|$, $\log|b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-1-j}|$, $\log|b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-2-j}|$, ... , is ergodic.

Now let,

$$P(n,t) = \prod_{m=1}^{n} (b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-m+1-j})$$

Taking logarithms, we obtain

$$\frac{1}{n} \log|P(n,t)| = \frac{1}{n} \sum_{m=1}^{n} \log|b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-m+1-j}|$$

By the ergodic theorem

$$\lim_{n\to\infty} \frac{1}{n} \log|P(n,t)| = E \log|b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-j}| \qquad \text{a.e } [P]$$

$$< 0$$

Hence,

$$\lim_{n\to\infty} \prod_{m=1}^{n} (b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-m+1-j}) = 0 \qquad \text{a.e } [P]$$

This completes the proof.

REMARKS 4.3.5

(1) The condition

$$E \log|b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-j}| < 0$$

involves the distribution of X_t, $t \in Z$. We obtain a sufficient condition.

$$E \log|b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-j}| = \tfrac{1}{2} E \log (b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-j})^2$$

$$< \tfrac{1}{2} \log E(b + \sum_{j=1}^{\ell} \beta_{j1} X_{t-j})^2, \text{ by}$$

Jensen's inequality.

$$< \tfrac{1}{2} \log(b^2 + 2 b \mu \sum_{j=1}^{\ell} \beta_{j1}$$

$$+ \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \beta_{i1} \beta_{j1} E(X_{t-i} X_{t-j}))$$

where

$$\mu = E(X_t) = \sigma^2 \beta_{11} / (1 - \sum_{j=1}^{r} a_j).$$

Hence a sufficient condition for invertibility is

$$b^2 + 2 b \mu \sum_{j=1}^{\ell} \beta_{j1} + \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \beta_{i1} \beta_{j1} E(X_{t-i} X_{t-j}) < 1 \qquad (4.3.9)$$

(2) Subba Rao [37] considered the model (4.3.7) without the moving average part be_{t-1} in (4.3.7) and with $r = \ell$. He used the Granger-Andersen invertibility concept to obtain a sufficient condition for invertibility. His condition is

$$C^T B E(\underline{X}_t \underline{X}_t^T) B^T C < 1 \qquad (4.3.10)$$

where $\underset{p\times p}{B}$, $\underset{p\times 1}{C}$ are the matrices of Remarks 2.3.3 (4) and

$$\underset{1\times\ell}{\underline{X}_t^T} = (X_t, X_{t-1}, \dots, X_{t-\ell+1}), \ t \in Z.$$

We can evaluate (4.3.10) to obtain

$$\sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \beta_{i1} \beta_{j1} E(X_{t-i} X_{t-j}) < 1 \qquad (4.3.11)$$

We note that expression (4.3.11) is the same as (4.3.9) with b = 0.

(3) The condition for the invertibility of the moving average process X_t, $t \in Z$ satisfying

$$X_t = be_{t-1} + e_t \qquad \text{a.e } [P] \quad (4.3.12)$$

for every t in Z under the above assumptions on the e_t's, can be deduced from (4.3.9) by putting $\beta_{j1} = 0$ for all $j = 1, 2, \dots, \ell$. The condition for invertibility is

$$|b| < 1 \qquad (4.3.13)$$

4.4 ON THE BISPECTRAL ANALYSIS OF PURELY BILINEAR WHITE NOISE PROCESSES

The sample bispectrum is beginning to play an important role in testing Gaussianity and linearity of stationary time series. See Subba Rao and Gabr [39] and Hinich [24].

Let X_t, $t \in Z$ be a real valued process with finite moments up to the third-order and is stationary up to the third-order. If X_t, $t \in Z$ is pure white noise, then the third-order moments

$$C(k_1, k_2) = E\{(X_t - \mu)(X_{t+k_1} - \mu)(X_{t+k_2} - \mu)\} \qquad (4.4.1)$$

where $\mu = E(X_t)$ will be zero for all values of k_1 and k_2 in Z. Thus, the bispectral density function

$$f(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} C(k_1, k_2) e^{-i(k_1\omega_1 + k_2\omega_2)} \qquad (4.4.2)$$

$$(-\pi \leq \omega_1, \omega_2 \leq \pi)$$

is zero for all frequencies ω_1 and ω_2 when X_t, $t \in Z$ is pure white noise. On the other hand, the third-order moments for some values of k_1 and k_2 and the bispectral density function for many frequencies ω_1 and ω_2 are usually non-zero when X_t, $t \in Z$ is non-linear.

Gabr [12] has obtained an exact expression for the bispectral density function of the process X_t, $t \in Z$ satisfying

$$X_t = a X_{t-1} + \beta X_{t-1} e_{t-1} + e_t \qquad \text{a.e } [P]$$

for every t in Z where e_t, $t \in Z$ is a sequence of independent identically distrubuted normal random variables with $E(e_t) = 0$ and $E(e_t^2) = 1$.

Granger and Andersen [15, p.43] have considered the third-order moments of the process X_t, $t \in Z$ satisfying

$$X_t = b X_{t-2} e_{t-1} + e_t \qquad \text{a.e } [P] \quad (4.4.3)$$

for some t in Z where e_t, $t \in Z$ is a sequence of independent identically distributed normal random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. They come to the conclusion that all the third moments are zero for X_t, $t \in Z$ satisfying (4.4.3) and so are of no use for identification between pure white noise and some bilinear models. Our results below contradicts this claim.

We show below that the third-order moments (and hence the bispectral density function) of the second-order stationary process X_t, $t \in Z$ satisfying (4.2.3) are non-zero for some values of k_1 and $k_2 \in Z$ (and for many frequencies $-\pi \leq \omega_1, \omega_2 \leq \pi$). In view of the symmetric relations (see chapter 1) satisfied by $C(k_1, k_2)$, we restrict attention to the plane $\{0 \leq k_1 < \infty, k_1 \leq k_2 < \infty\}$.

Let e_t, $t \in Z$ be a sequence of independent identically distributed normal random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let $b_1, b_2, \ldots, b_m$ be m real numbers. Let $\rho(\Gamma) < 1$, where Γ is the matrix of Theorem 4.2.2. We proceed to calculate the third moments and the bispectral density function of the process X_t, $t \in Z$ satisfying

$$X_t = e_t + \left\{\sum_{j=1}^{m} b_j X_{t-q-j}\right\} e_{t-q} \qquad \text{a.e } [P] \quad (4.4.4)$$

We assume X_t, $t \in Z$ satsifying (4.4.4) is stationary up to the third-order.

We proceed in the following steps.

1°. The mean, variance and covariances of the second-order stationary process X_t, $t \in Z$ satisfying (4.4.4) are given by

$$\mu = E(X_t) = 0$$

$$E(X_t^2) = \sigma^2/(1 - \sum_{j=1}^{m} \sigma^2 b_j^2)$$

$$\mathrm{Cov}(X_t\ X_{t-k}) = 0\ ,\ k \neq 0$$

2°. We obtain $C(k_1, k_2)$ for all $k_1 \neq k_2$ with k_1 and k_2 lying inside the plane $\{0 \leqslant k_1 < \infty,\ k_1 \leqslant k_2 < \infty\}$. Some rather tedious algebra shows that

$$C(k_1, k_2) = \sigma^2 b_j E(X_t^2)\ ,\ \text{if } k_1 = j \text{ and } k_2 = q + j,\ j = 1, 2, \ldots, m$$

$$= 0 \text{ elsewhere} \qquad (4.4.5)$$

3°. When $k_1 = k_2 = 0$, we obtain from (4.4.4) that

$$X_t^3 = e_t^3 + 3 e_t^2 e_{t-q}\left\{\sum_{j=1}^{m} b_j X_{t-q-j}\right\} + 3 e_t e_{t-q}^2 \left\{\sum_{j=1}^{m} b_j X_{t-q-j}\right\}^2 + e_{t-q}^3\left\{\sum_{j=1}^{m} b_j X_{t-q-j}\right\}^3$$

Since e_t is independent of X_s, $s < t$ and $E(e_t) = 0 = E(e_t^3)$, we obtain

$$C(0, 0) = E(X_t^3) = 0 \qquad (4.4.6)$$

4°. For $k_1 = k_2 = 1$, we have

$$E(X_t\ X_{t+1}^2) = \sum_{j=1}^{m} b_j^2 E(X_t e_{t+1-q}^2\ X_{t+1-q-j}^2)$$

$$+ 2 \sum_{\substack{i=1 \\ i<j}}^{m} \sum_{j=1}^{m} b_i \, b_j \, E(X_t e^2_{t+1-q} \, X_{t+1-q-i} \, X_{t+1-q-j})$$

Let us now consider

$$X_t e^2_{t+1-q} \, X^2_{t+1-q-j} = e_t e^2_{t-q+1} \, X^2_{t+1-q-j}$$

$$+ e_{t-q} e^2_{t+1-q} \, X^2_{t+1-q-j} \left[\sum_{i=1}^{m} b_i \, X_{t-q-i} \right]$$

when $j = 1$,

$$E(X_t e^2_{t+1-q} \, X^2_{t-q}) = \sigma^2 \sum_{i=1}^{m} b_i \, E(X^2_{t-q} e_{t-q} \, X_{t-q-i})$$

$$= 0$$

because $E(X^2_{t-q} e_{t-q} \, X_{t-q-i}) = 0$ for all $i = 1, 2, \dots, m$.

When $j > 1$, it is easy to check that

$$E(X_t e^2_{t+1-q} \, X^2_{t+1-q-j}) = 0$$

Hence

$$E(X_t \, X^2_{t+1}) = 2 \sum_{\substack{i=1 \\ i<j}}^{m} \sum_{j=1}^{m} b_i \, b_j \, E(X_t e^2_{t+1-q} \, X_{t+1-q-i} \, X_{t+1-q-j})$$

$$= 2 \sum_{j=2}^{m} b_1 b_j \, E(X_t e^2_{t+1-q} \, X_{t-q} \, X_{t+1-q-j})$$

(Since $E(X_t e^2_{t+1-q} \, X_{t+1-q-i} \, X_{t+1-q-j}) = 0$ for all $i > 1$ and $j > i$).

We can show that

$$E(X_t e^2_{t+1-q} \, X_{t-q} \, X_{t+1-q-j}) = \sigma^4 \, b_{j-1} \, E(X^2_t) \, , \; j = 2, 3, \dots , m.$$

Thus,

$$C(1, 1) = E(X_t \, X^2_{t+1})$$

$$= 2\sigma^4 \, b_1 \, E(X^2_t) \left[\sum_{j=1}^{m-1} b_j \, b_{j+1} \right] \tag{4.4.7}$$

5^0. For a given second-order stationary process X_t, $t \in Z$ satisfying (4.4.4), all values of $C(k, k)$, $k \geq 1$ can be evaluated and after some initial values, C(k, k) will satisfy a linear difference equation. This linear difference equation is always satisfied for all values of $k \geq 2m + q + 1$. This fact is now demonstrated below.

First we note that

$$X_t X_{t+2m+q+1}^2$$

$$= X_t \left[e_{t+2m+q+1}^2 + 2e_{t+2m+q+1} e_{t+2m+1} \left(\sum_{j=1}^{m} b_j X_{t+2m+1-j} \right) \right.$$

$$\left. + e_{t+2m+1}^2 \left(\sum_{j=1}^{m} b_j^2 X_{t+2m+1-j}^2 \right) + 2 \sum_{\substack{i=1 \\ i<j}}^{m} \sum_{j=1}^{m} b_i b_j X_{t+2m+1-i} X_{t+2m+1-j} \right]$$

Then

$$E(X_t X_{t+2m+q+1}^2) = \sum_{j=1}^{m} \sigma^2 b_j^2 E(X_t X_{t+2m+1-j}^2)$$

$$+ 2\sigma^2 \sum_{\substack{i=1 \\ i<j}}^{m} \sum_{j=1}^{m} b_i b_j E(X_t X_{t+2m+1-i} X_{t+2m+1-j})$$

$$= \sum_{j=1}^{m} \sigma^2 b_j^2 E(X_t X_{t+2m+1-j}^2)$$

(Using 4.4.5)

That is

$$C(2m+q+1, 2m+q+1) = \sum_{j=1}^{m} \sigma^2 b_j^2 C(2m+1-j, 2m+1-j).$$

Generally, we would obtain after some initial values

$$C(k, k) = \sum_{j=1}^{m} \sigma^2 b_j^2 C(k - q - j, k - q - j) \tag{4.4.8}$$

We now have a difference equation (4.4.8), which can be put in the following vector form

$$\begin{bmatrix} C(k, k) \\ C(k-1, k-1) \\ C(k-2, k-2) \\ \vdots \\ C(k-r+1, k-r+1) \end{bmatrix} = \nabla \begin{bmatrix} C(k-1, k-1) \\ C(k-2, k-2) \\ C(k-3, k-3) \\ \vdots \\ C(k-r, k-r) \end{bmatrix} \tag{4.4.9}$$

where $r = q + m$ and ∇ is as given in Lemma 4.2.3. Since $\rho(\nabla) < 1$,

$$\lim_{k\to\infty} C(k, k)$$

exists. We show that this limit a, say, is zero. In (4.4.8), take limit as $k \to \infty$. We have

$$a = \left[\sum_{j=1}^{m} \sigma^2 b_j^2\right] a$$

Since $\sum_{j=1}^{m} \sigma^2 b_j^2 < 1$, the only solution to the above is that $a = 0$.

Hence

$$\lim_{k\to\infty} C(k, k) = 0 . \qquad (4.4.10)$$

6°. Having obtained all the non-zero third-order moments we use the symmetric relations

$$\begin{aligned} C(k_1, k_2) &= C(k_2, k_1) = C(-k_1, k_2-k_1) \\ &= C(k_2-k_1, -k_1) = C(k_1-k_2, -k_2) \\ &= C(-k_2, k_1-k_2) \end{aligned}$$

$$C(k, k) = C(-k, 0) = C(0, -k)$$

to show that the bispectral density function of the third-order stationary process X_t, $t \in Z$ satisfying (4.4.4) is of the form

$$\begin{aligned} f(\omega_1, \omega_2) = \frac{1}{(2\pi)^2}\Big\{ & h_1(\omega_1, \omega_2) + h_2(\omega_1, \omega_2) + \dots \\ & + h_m(\omega_1, \omega_2) + g(\omega_1, \omega_2)\Big\} \end{aligned}$$

$$-\pi \leq \omega_1, \omega_2 \leq \pi \ , \qquad (4.4.11)$$

where

$$\begin{aligned} h_j(\omega_1, \omega_2) = \sigma^2 b_j\, E(X_t^2)\, \{ & e^{-i(j\omega_1+(q+j)\omega_2)} \\ & + e^{-i((q+j)\omega_1+j\omega_2)} + e^{-i(q\omega_1-j\omega_2)} \\ & + e^{i(j\omega_1-q\omega_2)} + e^{i(q\omega_1+(q+j)\omega_2)} \end{aligned}$$

$$+ e^{i((q+j)\omega_1+q\omega_2)}\}$$

$$(-\pi \leq \omega_1, \ \omega_2 \leq \pi) \qquad (4.4.12)$$

and

$$g(\omega_1, \omega_2) = \sum_{k=1}^{\infty} C(k, k)\{e^{-ik(\omega_1+\omega_2)} + e^{ik\omega_1} + e^{ik\omega_2}\}$$

$$-\pi \leq \omega_1, \ \omega_2 \leq \pi \qquad (4.4.13)$$

We now give the third-order moments and bispectral density function of four special cases of (4.4.4).

<u>EXAMPLE 4.4.1</u>. Let us consider the process X_t, $t \in Z$ satisfying

$$X_t = e_t + b_1 e_{t-1} X_{t-2} \qquad \text{a.e } [P]$$

for every t in Z where e_t, $t \in Z$ is a sequence of independent identically distributed normal random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. We assume X_t, $t \in Z$ is stationary up to order 3. Then

$$C(k_1, k_2) = \sigma^2 b_1 E(X_t^2), \text{ if } k_1 = 1 \text{ and } k_2 = 2$$

$$= 0 \text{ elsewhere} \qquad (4.4.14)$$

and

$$f(\omega_1, \omega_2) = \frac{2\sigma^2 b_1 E(X_t^2)}{(2\pi)^2}\left\{\text{Cos}(\omega_1 + 2\omega_2) + \text{Cos}(2\omega_1 + \omega_2) + \text{Cos}(\omega_1 - \omega_2)\right\}$$

$$(-\pi \leq \omega_1, \ \omega_2 \leq \pi) \qquad (4.4.15)$$

where

$$R(0) = E(X_t^2) = \sigma^2/(1 - \sigma^2 b_1^2), \ \sigma^2 b_1^2 < 1.$$

<u>EXAMPLE 4.4.2</u>. Our next example is the process X_t, $t \in Z$ satisfying

$$X_t = b_1 X_{t-3} e_{t-2} + e_t \qquad \text{a.e } [P]$$

for every t in Z where e_t, $t \in Z$ are defined as in Example 4.4.1. Assuming X_t, $t \in Z$ is stationary up to order 3 we obtain

$$C(k_1, k_2) = \sigma^2 b_1 E(X_t^2) \text{ if } k_1 = 1 \text{ and } k_2 = 3$$

$$= 0 \text{ elsewhere} \tag{4.4.16}$$

and

$$f(\omega_1, \omega_2) = \frac{\sigma^2 \, b_1 \, E(X_t^2)}{(2\pi)^2} \left\{ e^{-i(\omega_1+3\omega_2)} + e^{-i(3\omega_1+\omega_2)} + e^{-i(2\omega_1-\omega_2)} \right.$$

$$\left. + \, e^{i(\omega_1-2\omega_2)} + e^{i(2\omega_1+3\omega_2)} + e^{i(3\omega_1+2\omega_2)} \right\} \tag{4.4.17}$$

$$(-\pi \leqslant \omega_1, \ \omega_2 \leq \pi)$$

where

$$E(X_t^2) = \sigma^2 / (1 - \sigma^2 \, b_1^2) \, , \ \sigma^2 \, b_1^2 < 1.$$

<u>EXAMPLE 4.4.3</u>. We now consider a two parameter bilinear white noise. Let X_t, $t \in Z$ be a third-order stationary process satisfying

$$X_t = e_t + e_{t-1}(b_1 \, X_{t-2} + b_2 \, X_{t-3}) \qquad \text{a.e } [P]$$

for every t in Z where e_t, $t \in Z$ are defined as in Example 4.4.1. We give below the third-order moments and the bispectral density function.

For $k_1 \neq k_2$, we obtain

$$C(K_1, k_2) = \sigma \, \lambda_1 \, R(0) \ \text{ if } k_1 = 1 \text{ and } k_2 = 2$$

$$= \sigma \, \lambda_2 \, R(0) \ \text{ if } k_1 = 2 \text{ and } k_3 = 3$$

$$= 0 \text{ elsewhere} \tag{4.4.18}$$

When $k_1 = k_2 = k$, we obtain

$$C(k, k) = \begin{cases} 0 & \text{if } k = 0 \\ 2\sigma \, \lambda_1^2 \, \lambda_2 \, R(0) & \text{if } k = 1 \\ 4\sigma \, \lambda_1^3 \, \lambda_2^2 \, R(0) & \text{if } k = 2 \\ 2\sigma \, \lambda_1^4 \, \lambda_2 \, R(0) & \text{if } k = 3 \\ 2\sigma \, \lambda_1^2 \, \lambda_2(2 \, \lambda_1^3 \, \lambda_2 + \lambda_2^2 + 1)R(0) & \text{if } k = 4 \\ 2\sigma \, \lambda_1 \, \lambda_2(\lambda_1^5 + 2 \, \lambda_1^2 \, \lambda_2^3 + \lambda_2)R(0) & \text{if } k = 5 \end{cases} \tag{4.4.19}$$

and

$$C(k, k) = \lambda_1^2 \, C(k-2, k-2) + \lambda_3^2 \, C(k-3, k-3)$$

$$k = 6, 7, 8, \ldots \tag{4.4.20}$$

where

$$\lambda_j = \sigma b_j , j = 1, 2$$

and

$$R(0) = E(X_t^2) = \sigma^2/(1 - \lambda_1^2 - \lambda_2^2)$$

Note that $\lim_{k\to\infty} C(k, k) = 0$ by (4.4.10)

Also,

$$f(\omega_1, \omega_2) = \frac{2\sigma\lambda_1 R(0)}{(2\pi)^2}\left\{\cos(\omega_1 + 2\omega_2) + \cos(2\omega_1 + \omega_2) + \cos(\omega_1 - \omega_2)\right\}$$

$$+ \frac{\sigma\lambda_2 R(0)}{(2\pi)^2}\left\{e^{-i(2\omega_1+3\omega_2)} + e^{-i(3\omega_1+2\omega_2)} + e^{-i(\omega_1-2\omega_2)}\right.$$

$$\left. + e^{i(2\omega_1-\omega_2)} + e^{i(\omega_1+3\omega_2)} + e^{i(3\omega_1+\omega_2)}\right\}$$

$$+ \sum_{k=1}^{\infty} \frac{C(k, k)}{(2\pi)^2}\left\{e^{-ik(\omega_1+\omega_2)} + e^{ik\omega_1} + e^{ik\omega_2}\right\} \qquad (4.4.21)$$

$$(-\pi \leq \omega_1, \omega_2 \leq \pi)$$

EXAMPLE 4.4.4. Finally, we consider a two parameter bilinear white noise with $q = 2$ and $m = 2$. Let X_t, $t \in Z$ be a third-order stationary process satisfying

$$X_t = e_t + e_{t-2}(b_1 X_{t-3} + b_2 X_{t-4}) \qquad \text{a.e } [P]$$

for every t in Z where e_t, $t \in Z$ are defined as in Example 4.4.1.

For this process we obtained the third-order moments to be

(a) when $k_1 \neq k_2$

$$\begin{aligned} C(k_1, k_2) &= \sigma\lambda_1 R(0) \text{ if } k_1 = 1 \text{ and } k_2 = 3 \\ &= \sigma\lambda_2 R(0) \text{ if } k_1 = 2 \text{ and } k_2 = 4 \\ &= 0 \qquad \text{elsewhere} \end{aligned} \qquad (4.4.22)$$

(b) when $k_1 = k_2 = k$

$$C(k, k) = \begin{cases} 0 & \text{if } k = 0 \\ 2\sigma \lambda_1^2 \lambda_2 R(0) & \text{if } k = 1 \\ 0 & \text{if } k = 2 \\ 4\sigma \lambda_1^3 \lambda_2^2 R(0) & \text{if } k = 3 \\ 2\sigma \lambda_1^4 \lambda_2 R(0) & \text{if } k = 4 \end{cases} \tag{4.4.23}$$

and

$$C(k, k) = \lambda_1^2 C(k - 3, k - 3) + \lambda_2^2 C(k - 4, k - 4)$$

$$k = 5, 6, 7, \ldots \tag{4.4.24}$$

where

$$\lambda_j = \sigma b_j , \ j = 1, 2$$

and

$$R(0) = E(X_t^2) = \sigma^2/(1 - \lambda_1^2 - \lambda_2^2)$$

$$\lim_{k\to\infty} C(k, k) = 0 , \text{ by } (4.4.10)$$

The bispectral density is given by

$$\begin{aligned} f(\omega_1, \omega_2) = {} & \frac{2\sigma \lambda_2 R(0)}{(2\pi)^2}\left\{\mathrm{Cos}(2\omega_1 + 4\omega_2) + \mathrm{Cos}(4\omega_1 + 2\omega_2) + \mathrm{Cos}(2\omega_1 - 2\omega_2)\right\} \\ & + \frac{\sigma \lambda_1 R(0)}{(2\pi)^2}\left\{e^{-i(\omega_1+3\omega_2)} + e^{-i(3\omega_1+\omega_2)} + e^{-i(2\omega_1-\omega_2)}\right. \\ & \left. + e^{i(\omega_1-2\omega_2)} + e^{i(2\omega_1+3\omega_2)} + e^{i(3\omega_1+2\omega_2)}\right\} \\ & + \sum_{k=1}^{\infty} \frac{C(k, k)}{(2\pi)^2}\left\{e^{-ik(\omega_1+\omega_2)} + e^{ik\omega_1} + e^{ik\omega_2}\right\} \end{aligned} \tag{4.4.25}$$

$$(-\pi \leq \omega_1, \ \omega_2 \leq \pi)$$

Let us now consider the implication of applying Subba Rao - Gabr linearity test to a purely bilinear white noise X_t, $t \in Z$ satisfying (4.4.4). By way of introduction, we consider the simple purely bilinear white noise of Example 4.4.1. Since the process in Example 4.4.1 is white noise, its spectral density function is given by

$$f(\omega) = R(0)/2\pi$$

$$= \frac{\sigma^2}{2\pi(1 - \sigma^2 b_1^2)} \qquad (4.4.26)$$

$$= \text{Constant}$$

Using (4.4.15) and (4.4.26) we obtain

$$X_{ij} = \frac{|f(\omega_i, \omega_j)|^2}{f(\omega_i)\, f(\omega_j)\, f(\omega_i + \omega_j)}$$

$$= \frac{2\sigma^2 b_1^2(1 - \sigma^2 b_1^2)}{\pi} \left[\text{Cos}(\omega_i + 2\omega_j) + \text{Cos}(2\omega_i + \omega_j) + \text{Cos}(\omega_i - \omega_j)\right]^2$$

$$(-\pi \leq \omega_i,\ \omega_j \leq \pi) \qquad (4.4.27)$$

Equation (4.4.27) shows that X_{ij} is not a constant for all i and j. This result holds true for the more general bilinear model X_t, $t \in Z$ satisfying (4.2.3). Therefore, the tests, constructed from the bispectral density function, of Subba Rao and Gabr [38] will be of great use for identification between pure white noise and pure bilinear white noise of the type given in (4.2.3).

Finally, we give plots of the bispectral density functions (4.4.15), (4.4.16), (4.4.21) and (4.4.25). Let $b_1 = 0.6$ and $\sigma^2 = 1.0$ in (4.4.15) and (4.4.16). Also let $b_1 = 0.45$, $b_2 = 0.35$ and $\sigma^2 = 1.0$ in (4.4.21) and (4.4.25). Using these values, we present the graphs of the bispectral density functions (4.4.15), (4.4.16), (4.4.21) and (4.4.25) in Figures 4.1, 4.2, 4.3 and 4.4 respectively. In the graphs given, X stands for ω_1, Y stands for ω_2 and

$$Z(x, y) = |f(\omega_1, \omega_2)|\ ,\ -\pi < \omega_1, \omega_2 < \pi\ .$$

The red colour is used to indicate the lower side of the surface while the green colour indicates the upper side.

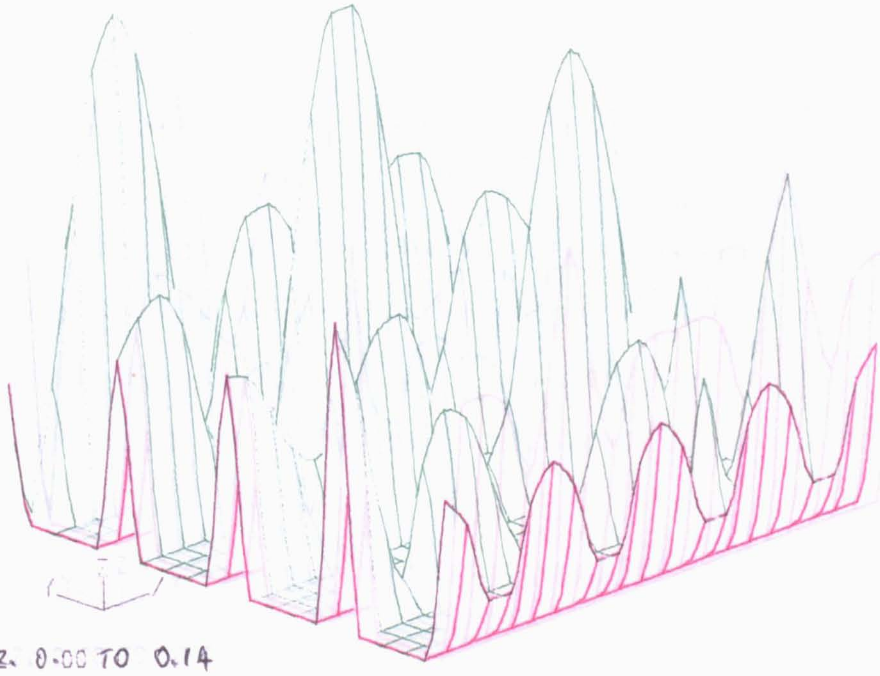


FIG. 4.1 THE MODULUS OF THE BISPECTRUM OF THE PROCESS

$X_t = 0.6X_{t-2}e_{t-1} + e_t;\ e_t \simeq N(0, 1).$

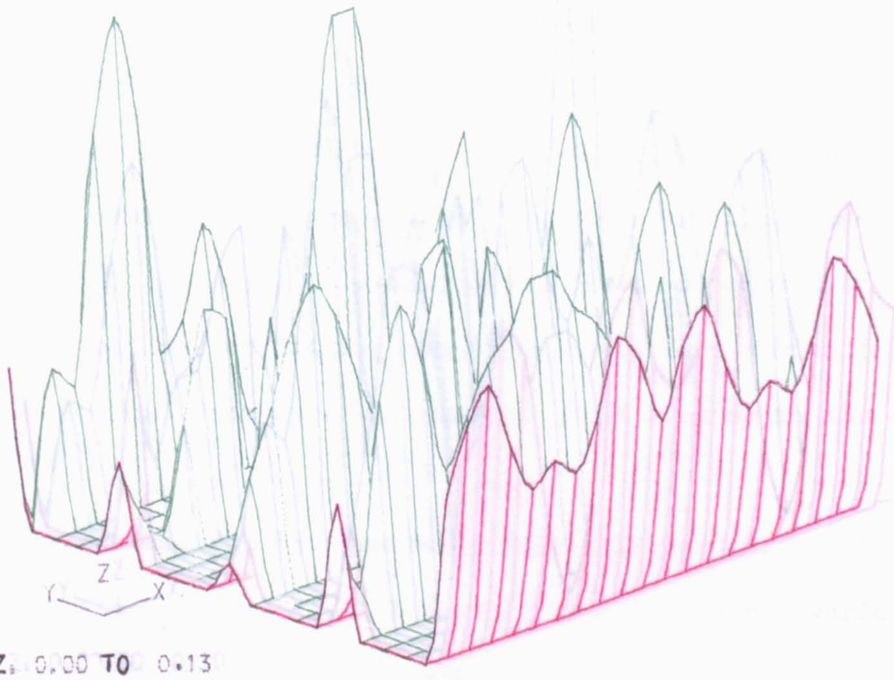


FIG. 4.2 THE MODULUS OF THE BISPECTRUM OF THE PROCESS

$X_t = 0.6X_{t-3}e_{t-2} + e_t;\ e_t \simeq N(0, 1).$

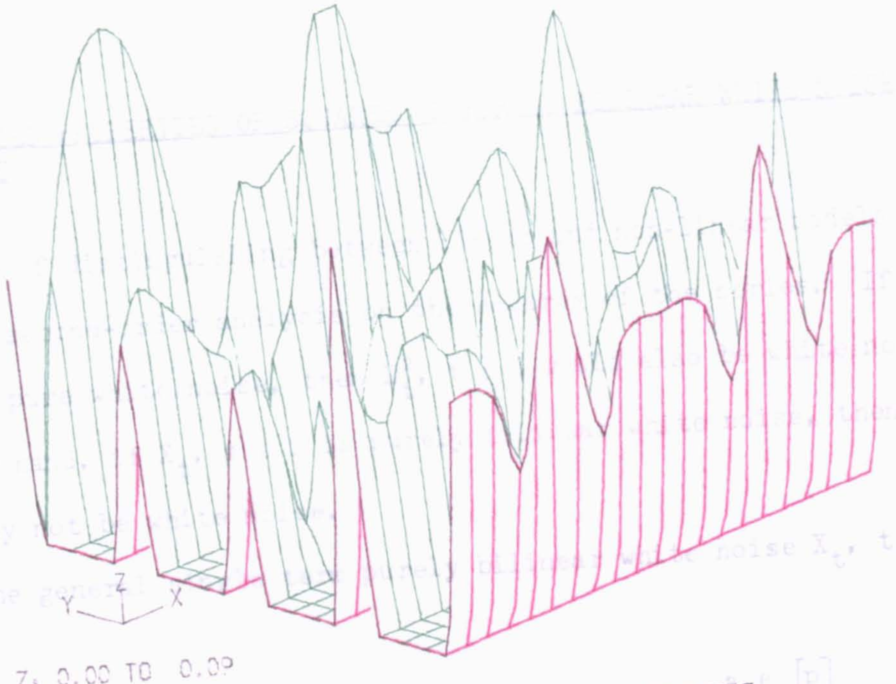


FIG. 4.3 THE MODULUS OF THE BISPECTRUM OF THE PROCESS

$$X_t = (0.45X_{t-2} + 0.35X_{t-3})e_{t-1} + e_t;\ e_t \simeq N(0, 1).$$

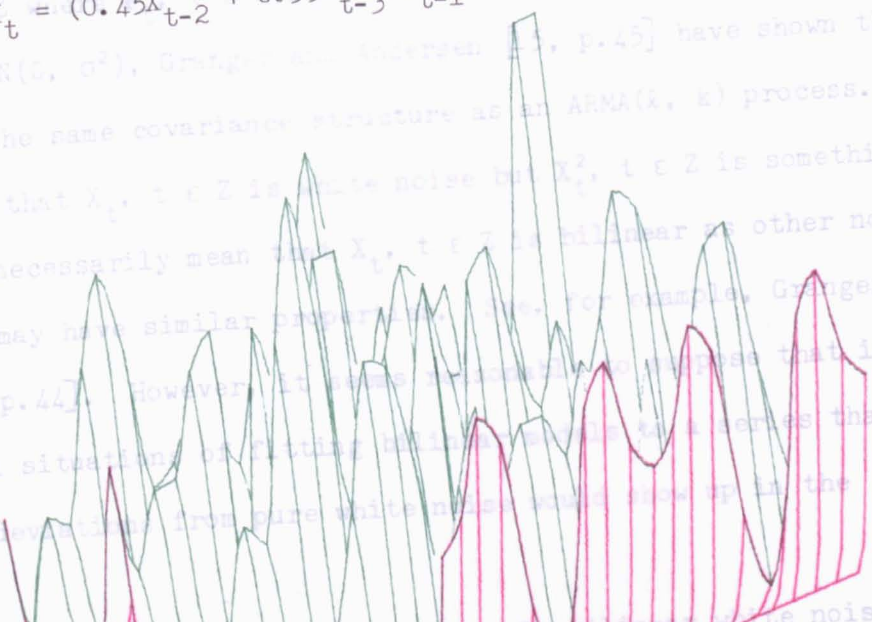


FIG. 4.4 THE MODULUS OF THE BISPECTRUM OF THE PROCESS

$$X_t = (0.45X_{t-3} + 0.35X_{t-4})e_{t-2} + e_t;\ e_t \simeq N(0, 1).$$

4.5 TIME SERIES PROPERTIES OF SQUARES OF PURELY BILINEAR WHITE NOISE PROCESSES

One way of distinguishing between linear and non-linear models is to perform a second-order analysis on the squares of the series. If X_t, $t \in Z$ is pure white noise, then X_t^2, $t \in Z$ would also be white noise. On the other hand, if X_t, $t \in Z$ is purely bilinear white noise, then X_t^2, $t \in Z$ may not be white noise.

For the general single term purely bilinear white noise X_t, $t \in Z$ satisfying

$$X_t = \beta e_{t-k} X_{t-\ell} + e_t \, , \; \ell < k \qquad \text{a.e } [p]$$

for every t in Z where e_t, $t \in Z$ are independent and each e_t, $t \in Z$ is distributed as $N(0, \sigma^2)$, Granger and Andersen [15, p.45] have shown that X_t^2, $t \in Z$ has the same covariance structure as an ARMA(ℓ, k) process.

The fact that X_t, $t \in Z$ is white noise but X_t^2, $t \in Z$ is something else does not necessarily mean that X_t, $t \in Z$ is bilinear as other non-linear series may have similar properties. See, for example, Granger and Andersen [15, p.44]. However, it seems reasonable to suppose that in most practical situations of fitting bilinear models to a series that is white noise, deviations from pure white noise would show up in the X_t^2, $t \in Z$.

We show below that for the general purely bilinear white noise X_t, $t \in Z$ satisfying (4.2.3), X_t^2, $t \in Z$ would have the same covariance structure as some ARMA processes. In what follows, X_t, $t \in Z$ is assumed to be stationary up to order 4 and e_t, $t \in Z$ are assumed to be independent and each e_t is distributed as $N(0, \sigma^2)$.

We use equation (4.2.3) and consider $E(X_t^2 X_{t-k}^2)$. We obtain

$$E(X_t^2 X_{t-k}^2) = \sigma^2 E(X^2) + \sum_{j=1}^{m} \sigma^2 b_j^2 E(X_{t-q-j}^2 X_{t-k}^2)$$

$$+ 2 \sum_{i=1}^{m} \sum_{j=1}^{m} \sigma^2 b_i b_j E(X_{t-q-i} X_{t-q-j} X_{t-k}^2) \quad (4.5.1)$$

For fixed q and m, $E(X_{t-q-i} X_{t-q-j} X_{t-k}^2)$ can be evaluated for all k and a point, say k_0, is reached after which

$$E(X_{t-q-i} X_{t-q-j} X_{t-k}^2) = 0 , \; k > k_0 .$$

Then

$$E(X_t^2 X_{t-k}^2) = \sum_{j=1}^{m} \sigma^2 b_j^2 E(X_{t-q-j}^2 X_{t-k}^2) + \sigma^2 E(X_t^2) , \; k > k_0 \quad (4.5.2)$$

If we now let

$$Y_t = X_t^2$$

$$\mu_y = E(Y_t) = E(X_t^2)$$

$$R_y(k) = Cov(Y_t Y_{t-k}) = E\{(Y_t - \mu_y)(Y_{t-k} - \mu_y)\} ,$$

we obtain from (4.5.2)

$$E(X_t^2 X_{t-k}^2) = \sum_{j=1}^{m} \sigma^2 b_j^2 (R_y(k - q - j) + \mu_y^2) + \mu_y^2 (1 - \sum_{j=1}^{m} \sigma^2 b_j^2)$$

$$\text{(by using } \sigma^2 = (1 - \sum_{j=1}^{m} \sigma^2 b_j^2)\mu_y)$$

$$= \sum_{j=1}^{m} \sigma^2 b_j^2 R_y(k - q - j) + \mu_y^2 , \; k > k_0 .$$

or

$$R_y(k) = \sum_{j=1}^{m} \sigma^2 b_j^2 R_y(k - q - j) , \; k > k_0 \quad (4.5.3)$$

Equation (4.5.2) is the Yule-Walker equation for an ARMA$(q + m, k_0)$.

To illustrate the kind of methods employed to obtain k_0 we consider the following examples.

(1) Let X_t, $t \in Z$ be the process considered in Example 4.4.1, ie, X_t, $t \in Z$ satisfies

$$X_t = e_t + b_1 e_{t-1} X_{t-2} \qquad \text{a.e } [P].$$

Let $\lambda_1 = \sigma b_1$, then (see also Granger and Andersen [15, p.44].)

$$E(X_t^4) = \frac{3\sigma^4(1 + \lambda_1^2)}{(1 - 3\lambda_1^4)(1 - \lambda_1^2)}, \quad 3\lambda_1^4 < 1$$

$$E(X_t^2 X_{t-1}^2) = \frac{\sigma^4(1 + 2\lambda_1^2)}{(1 - \lambda_1^2)^2}$$

$$E(X_t^2 X_{t-k}^2) = \lambda_1^2 E(X_{t-2}^2 X_{t-k}^2) + \sigma^2 E(X_t^2), \quad k_1 > 1$$

Thus,

$$R_y(0) = \frac{2\sigma^4}{(1 - \lambda_1^2)^2(1 - 3\lambda_1^4)} \tag{4.5.4}$$

$$R_y(1) = \frac{2\sigma^4 \lambda_1^2}{(1 - \lambda_1^2)^2}, \tag{4.5.5}$$

and

$$R_y(k) = \lambda_1^2 R_y(k - 2), \quad k \geq 2 \tag{4.5.6}$$

Hence $Y_t = X_t^2$ for the process X_t, $t \in Z$ considered in this example has the same covariance structure as some ARMA(2, 1) process.

(2) Let X_t, $t \in Z$ be the process considered in Example 4.4.3, ie, X_t, $t \in Z$ satisfies

$$X_t = e_t + (b_1 X_{t-2} + b_2 X_{t-3})e_{t-1} \qquad \text{a.e } [P]$$

For this model we obtain

$$E(X_t^4) = 3\lambda_1^4 E(X_{t-2}^4) + 3\lambda_2^4 E(X_{t-3}^4) + 18\lambda_1^2 \lambda_2^2 E(X_{t-2}^2 X_{t-3}^2)$$
$$+ 3\sigma^2 E(X_t^2)\{1 + \lambda_1^2 + \lambda_2^2 + 24\lambda_1^4 \lambda_2^3(3\lambda_1^3 + 2\lambda_2^3 + \lambda_2)\}$$

$$E(X_t^2 X_{t-1}^2) = \lambda_1^2 \lambda_2^2 E(X_{t-3}^4) + \lambda_1^2 E(X_{t-1}^2 X_{t-2}^2) + \lambda_2^4 E(X_{t-3}^2 X_{t-4}^2)$$
$$+ \sigma^2\{1 + 2\lambda_1^2 + 3\lambda_2^2(1 + 12\lambda_1^5 \lambda_2^3) + 4\lambda_1^2 \lambda_2^2(3\lambda_1^4(1 + 2\lambda_2^2)$$
$$+ \lambda_1 \lambda_2^3 + 2\lambda_1 \lambda_2)\} E(X_t^2)$$

$$E(X_t^2 X_{t-2}^2) = \lambda_1^2 E(X_{t-2}^4) + \lambda_2^2 E(X_{t-2}^2 X_{t-3}^2) + \sigma^2 E(X_t^2)(1 + 36\lambda_1^5 \lambda_2^3)$$

$$E(X_t^2\, X_{t-3}^2) = \lambda_1^2\, E(X_{t-2}^2\, X_{t-3}^2) + \lambda_2^2\, E(X_{t-3}^4)$$

$$+ \sigma^2\, E(X_t^2)(1 + 12\lambda_1^4\, \lambda_2^2(1 + 2\lambda_2^2))$$

$$E(X_t^2\, X_{t-4}^2) = \lambda_1^2\, E(X_{t-2}^2\, X_{t-4}^2) + \lambda_2^2\, E(X_{t-3}^2\, X_{t-4}^2) + \sigma^2\, E(X_t^2)(1 + 4\lambda_1^3\, \lambda_2^3)$$

and

$$E(X_t^2\, X_{t-k}^2) = \lambda_1^2\, E(X_{t-2}^2\, X_{t-k}^2) + \lambda_2^2\, E(X_{t-3}^2\, X_{t-k}^2) + \sigma^2\, E(X_t^2)\ ,\ k \geqslant 5$$

where (4.5.7)

$$\lambda_j = \sigma\, b_j\ ,\ j = 1,\ 2\ ,$$

and

$$E(X_t^2) = \sigma^2/(1 - \lambda_1^2 - \lambda_2^2)$$

Thus

$$R_y(k) = \lambda_1^2\, R_y(k - 2) + \lambda_2^2\, R_y(k - 3)\ ,\ k \geqslant 5$$

So, $Y_t = X_t^2$ for the process X_t, $t \in Z$ considered in this example has the same covariance structure as some ARMA(3, 4).

(3) Finally, let X_t, $t \in Z$ be the process considered in Example 4.4.3. Then X_t, $t \in Z$ satisfies

$$X_t = e_t + (b_1\, X_{t-3} + b_2\, X_{t-4})e_{t-2} \qquad \text{a.e } [P].$$

For this model we obtain

$$E(X_t^4) = 3\lambda_1^4\, E(X_{t-3}^4) + 3\lambda_2^4\, E(X_{t-4}^4) + 18\lambda_1^2\, \lambda_2^2\, E(X_{t-3}^2\, X_{t-4}^2)$$

$$+ 3\sigma^2\, E(X_t^2)(1 + \lambda_1^2 + \lambda_2^2).$$

$$E(X_t^2\, X_{t-1}^2) = \lambda_1^2\, \lambda_2^2\, E(X_{t-4}^4) + \lambda_1^2\, E(X_{t-1}^2\, X_{t-3}^2) + \lambda_2^4\, E(X_{t-4}^2\, X_{t-5}^2)$$

$$+ \sigma^2\, E(X_t^2)(1 + \lambda_2^2 + 16\lambda_1^5\, \lambda_2^5)$$

$$E(X_t^2\, X_{t-2}^2) = \lambda_1^2\, E(X_{t-2}^2\, X_{t-3}^2) + \lambda_2^2\, E(X_{t-2}^2\, X_{t-4}^2)$$

$$+ \sigma^2\, E(X_t^2)(1 + 2\lambda_1^2 + 2\lambda_2^2 + 4\lambda_1^3\, \lambda_2^3)$$

$$E(X_t^2\, X_{t-k}^2) = \lambda_1^2\, E(X_{t-3}^2\, X_{t-k}^2) + \lambda_2^2\, E(X_{t-4}^2\, X_{t-k}^2) + \sigma^2\, E(X_t^2)\ ,\ k \geqslant 3$$

(4.5.8)

where $\lambda_j = \sigma b_j$, $j = 1, 2$ and

$$E(X_t^2) = \sigma^2/(1 - \lambda_1^2 - \lambda_2^2) .$$

Thus,

$$R_y(k) = \lambda_1^2 R_y(k - 3) + \lambda_2^2 R_y(k - 4) , \quad k \geq 3 \qquad (4.5.9)$$

So, $Y_t = X_t^2$ for the process X_t, $t \in Z$ considered in this example has the same covariance structure as some ARMA(4, 2).

Figure 4.5 shows a series of 500 terms generated by the bilinear white noise model

$$X_t = (0.45X_{t-2} + 0.35X_{t-3})e_{t-1} + e_t \qquad (4.5.10)$$

where e_t is a normal N(0, 1) white noise. Columns 2 and 3 of Table 4.1 give the sample autocorrelations of the series and the squares of the series respectively. Figure 4.6 shows a series of 500 terms generated by the bilinear white noise model

$$X_t = (0.45X_{t-3} + 0.35X_{t-4})e_{t-2} + e_t \qquad (4.5.11)$$

where e_t is normal N(0, 1) white noise. Columns 4 and 5 of Table 4.1 give the sample autocorrelations of the series and the squares of the series respectively. In all cases the approximate standard error is 0.045. It is seen that in both cases X_t identifies as pure white noise under covariance analysis but X_t^2 certainly does not.

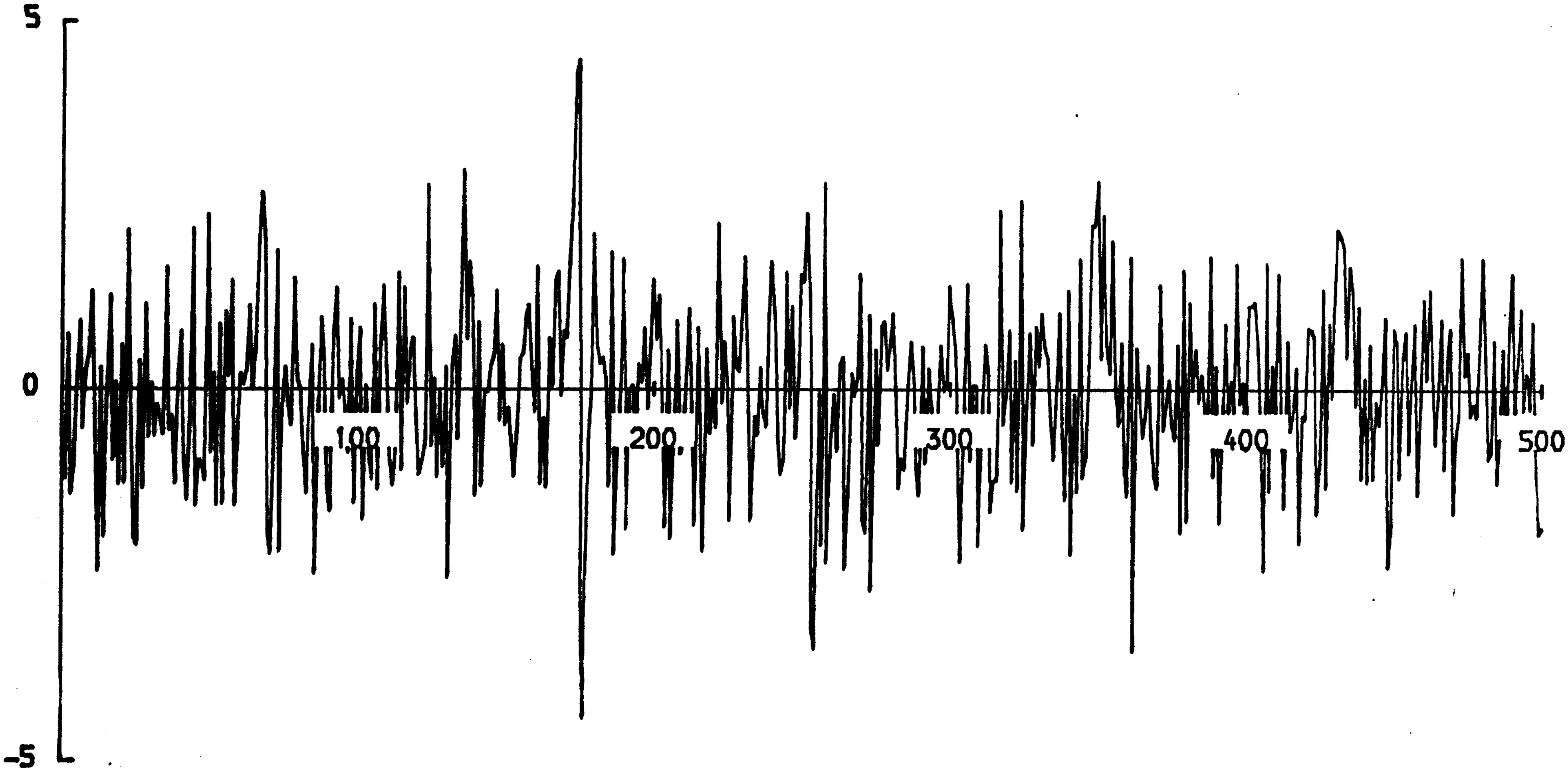


FIG 4.5 500 OBSERVATIONS FROM THE BILINEAR PROCESS $X_t = 0.45X_{t-2}e_{t-1} + 0.35X_{t-3}e_{t-1} + e_t$; $e_t \simeq N(0, 1)$

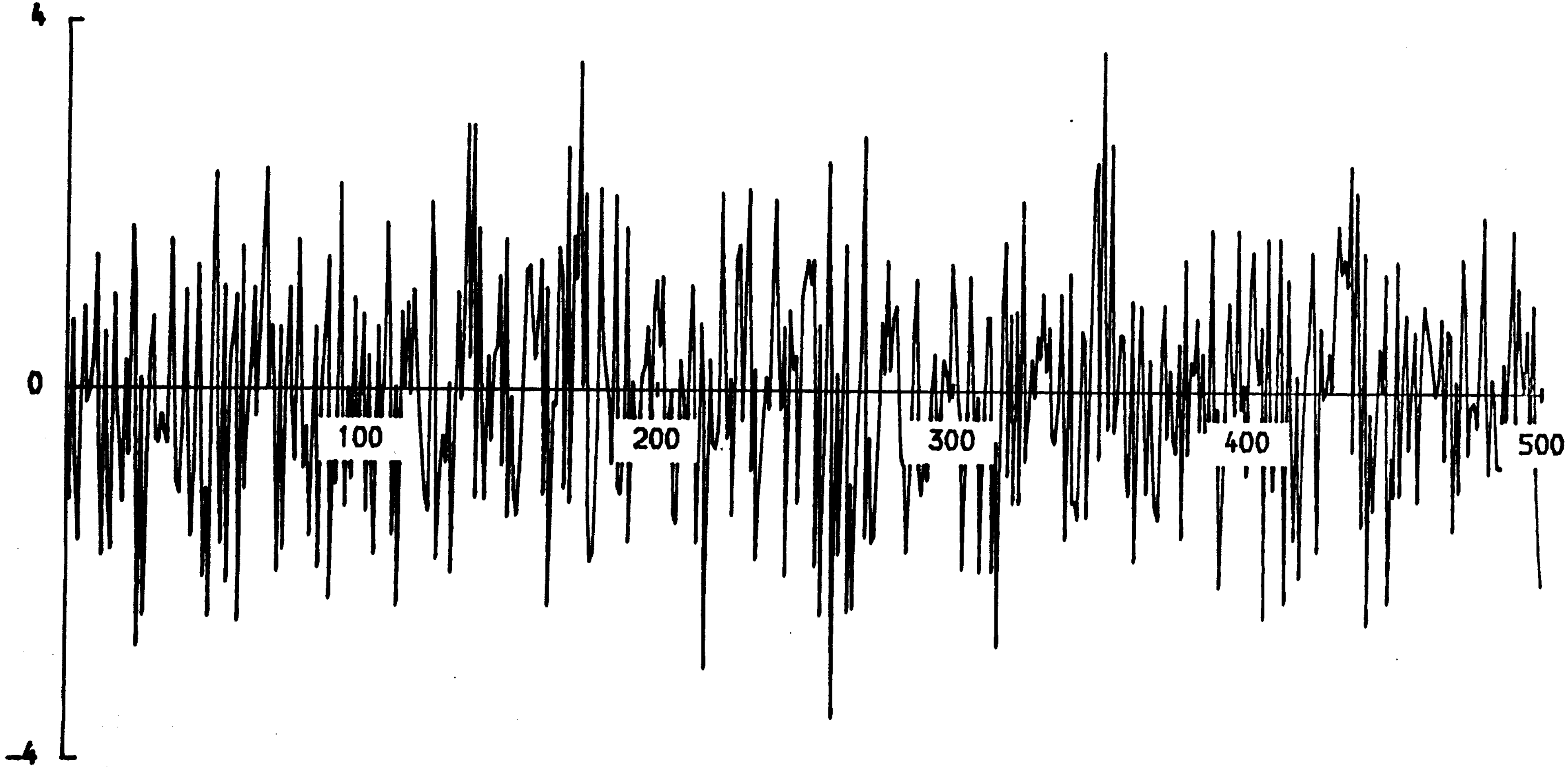


FIG. 4.6 500 OBSERVATIONS FROM THE BILINEAR PROCESS $X_t = 0.45X_{t-3}e_{t-2} + 0.35X_{t-4}e_{t-2} + e_t$, $e_t \simeq N(0, 1)$

TABLE 4.1 SAMPLE AUTOCORRELAITONS FOR REALIZATIONS OF LENGTH 500 OF THE BILINEAR WHITE NOISE MODELS (4.5.10) AND (4.5.11).

LAG	MODEL (4.5.10)		MODEL (4.5.11)	
k	$\hat{\rho}_k(X_t)$	$\hat{\rho}_k(X_t^2)$	$\hat{\rho}_k(X_t)$	$\hat{\rho}_k(X_t^2)$
1	0.007	0.285	-0.037	-0.100
2	0.033	0.234	0.031	0.209
3	-0.049	0.198	0.006	0.037
4	-0.043	0.022	-0.057	0.121
5	-0.006	0.009	-0.047	0.046
6	0.057	-0.051	0.063	0.023
7	-0.025	0.019	0.013	0.032
8	-0.030	-0.077	-0.099	-0.010
9	-0.025	-0.021	0.057	0.022
10	-0.028	0.035	-0.039	0.055
11	0.054	-0.008	-0.001	0.038
12	0.035	0.028	0.050	0.049
13	-0.036	0.044	-0.010	-0.027
14	-0.069	-0.008	-0.026	0.019
15	-0.049	-0.026	-0.081	-0.069
16	-0.008	0.005	-0.015	-0.022
17	-0.025	-0.005	-0.076	-0.042
18	0.007	-0.028	0.056	-0.024
19	0.019	-0.023	-0.005	-0.089
20	-0.047	-0.046	-0.020	-0.043

CHAPTER 5

ON THE FITTING OF BILINEAR MODELS TO TIME SERIES DATA

5.1 INTRODUCTION

Let e_t, $t \in Z$ be a sequence of independent identically distributed real random variables with common mean 0 and variance $\sigma^2 < \infty$. In what follows, we also assume e_t, $t \in Z$ to be normally distributed. Let $x_1, x_2, \ldots, x_n$ be a realisation of the process X_t, $t \in Z$ satisfying

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j} + \sum_{\substack{i=1 \\ i \geq j}}^{m} \sum_{j=1}^{\ell} \beta_{ij} X_{t-i} e_{t-j} + e_t \quad \text{a.e } [P] \quad (5.1.1)$$

for every t in Z, for some constants $a_1, a_2, \ldots, a_r$, $b_1, b_2, \ldots, b_h$ and β_{ij}, $1 \leq i \leq m$, $1 \leq j \leq \ell$, $i \geq j$. Let $\rho(\Gamma) < 1$, where Γ is the matrix of Theorem 2.4.3 built on A, $B_1, B_2, \ldots B_q$ of Theorem 2.4.1 and σ^2. Under the assumption that model (5.1.1) is invertible, this chapter considers the estimation of the parameters of the bilinear time series model given by (5.1.1).

Also considered in this chapter is the problem of selecting r, h, m and ℓ in modeling the BARMA(r, h, m, ℓ) process X_t, $t \in Z$ satisfying (5.1.1). We show how the covariance structure could be utilized to determine r, h, m and ℓ in a BARMA(r, h, m, ℓ) process, despite the fact that the bilinear model (5.1.1) is not necessarily distinguishable from linear ARMA models as far as the covariance properties are concerned.

An important use of time series models is to provide forecasts and sometimes, the performance of a time series model is judged on the basis of its forecasting performance. A rule for forming forecasts for the bilinear model (5.1.1) is given. Finally, we consider the fitting of bilinear time series models to some real time series data. The forecasts obtained from the bilinear models are compared with the forecasts obtained from the best linear ARMA models.

5.2 ESTIMATION OF THE PARAMETERS OF BILINEAR TIME SERIES MODELS

Subba Rao [37] has considered the estimation of the parameters of the bilinear model

$$X_t = a_0 + \sum_{j=1}^{p} a_j X_{t-j} + \sum_{i=1}^{p} \sum_{j=1}^{q} \beta_{ij} X_{t-i} e_{t-j} + e_t \qquad (5.2.1)$$

and Gabr and Subba Rao [11] have considered the estimation of the subset bilinear model

$$X_t = a_0 + \sum_{i=1}^{\ell} a_{k_i} X_{t-k_i} + \sum_{j=1}^{m} b_{r_j s_j} X_{t-r_j} e_{t-s_j} + e_t \qquad (5.2.2)$$

where $k_1, k_2, \ldots, k_\ell$ are subsets of the integers $(1, 2, \ldots, p)$, $1 \leq k_1 \leq k_2 \leq \ldots \leq k_\ell \leq p$; p is the order of the best linear autoregressive model that fits the data, and the pairs of integers

$$(r_1, s_1) \in T_1 = \{(i, j): i = 1, 2, \ldots ; j = 1, 2, \ldots\}$$

$$(r_2, s_2) \in T_2 = T_1 - \{(r_1, s_1)\}$$

$$\vdots$$

$$(r_m, s_m) \in T_m = T_{m-1} - \{(r_{m-1}, s_{m-1})\} , \; m \leq p .$$

The model (5.2.2) is a special case of (5.2.1). The statistical properties, such as stationarity, first and second-order moments and spectral density function, of the above two models in the general form are not yet known.

We wish to restrict attention to the fitting of those bilinear models whose second-order properties we can investigate and interpret. Let $g = \min(m, \ell)$, then the model (5.1.1) splits into two forms depending on whether $m > \ell$ or $m \leq \ell$. When $m > \ell$, model (5.1.1) can be written in the form

$$X_t = e_t + \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j}$$

$$+ (\beta_{11} X_{t-1} + \beta_{21} X_{t-2} + \dots + \beta_{m1} X_{t-m}) e_{t-1}$$

$$+ (\beta_{22} X_{t-2} + \beta_{32} X_{t-3} + \dots + \beta_{m2} X_{t-m}) e_{t-2}$$

$$+ \dots\dots\dots$$

$$+ (\beta_{gg} X_{t-g} + \beta_{g+1,g} X_{t-g-1} + \dots + \beta_{mg} X_{t-m}) e_{t-g}$$

$$(\ell = g) \tag{5.2.3}$$

If $m \leqslant \ell$, then model (5.1.1) can be written in the form

$$X_t = e_t + \sum_{j=1}^{r} a_j X_{t-j} + \sum_{j=1}^{h} b_j e_{t-j}$$

$$+ (\beta_{11} X_{t-1} + \beta_{21} X_{t-2} + \dots + \beta_{m1} X_{t-m}) e_{t-1}$$

$$+ (\beta_{22} X_{t-2} + \beta_{32} X_{t-3} + \dots + \beta_{m2} X_{t-m}) e_{t-2}$$

$$+ \dots\dots\dots$$

$$+ \beta_{mg} X_{t-m} e_{t-g}, \; m = g \tag{5.2.4}$$

Let R be the total number of parameters of the autoregressive, moving average and the pure bilinear parts of the model (5.1.1). From (5.2.3) and (5.2.4), we obtain the value for R to be

$$R = r + h + \tfrac{1}{2}[m(m+1) - (m-\ell)(m-\ell+1)] \text{ , if } m > \ell$$

$$= r + h + \tfrac{1}{2}m(m+1) \text{ , if } m \leqslant \ell \tag{5.2.5}$$

From (5.2.5) one cannot fail to notice that the total number of parameters of the BARMA(r, h, m, ℓ) model (5.1.1) can be excessive.

We now consider the estimation of the parameters of the bilinear time series model satisfying (5.2.3) or (5.2.4) when we have a realization $\{x_1, x_2, \dots x_n\}$ on the time series X_t, $t \in Z$ under the assumption that (5.2.3) and (5.2.4) are invertible. Proceeding as in Subba Rao [37], we can show that maximizing the likelihood function of $\{x_{n_0}, x_{n_0+1}, \dots x_n\}$ is the same as minimizing the function

$$S(\underline{\theta}) = \sum_{n_0}^{n} e_t^2 \tag{5.2.6}$$

with respect to the parameters $\underline{\theta}' = (a_1, a_2, \ldots a_r, b_1, b_2, \ldots, b_h, \beta_{11}, \beta_{21}, \ldots, \beta_{mg})$ where $n_0 = \max\{r, m\} + 1$. Let $\theta_1 = a_1$, $\theta_2 = a_2$, ..., $\theta_r = a_r$, $\theta_{r+1} = b_1$, $\theta_{r+2} = b_2$, ..., $\theta_{r+h} = b_h$, $\theta_{r+h+1} = \beta_{11}$, $\theta_{r+h+2} = \beta_{21}$, ..., $\theta_R = \beta_{mg}$. Then the partial derivatives of $S(\underline{\theta})$ are given by

$$\left.\begin{aligned} \frac{\partial S(\underline{\theta})}{\partial \theta_i} &= 2 \sum_{t=n_0}^{n} e_t \frac{\partial e_t}{\partial \theta_i} ; \\ \frac{\partial^2 S(\underline{\theta})}{\partial \theta_i \partial \theta_j} &= 2 \sum_{t=n_0}^{n} \left(\frac{\partial e_t}{\partial \theta_i}\right)\left(\frac{\partial e_t}{\partial \theta_j}\right) + 2 \sum_{t=n_0}^{n} e_t \frac{\partial^2 e_t}{\partial \theta_i \partial \theta_j} ; \end{aligned}\right\} \tag{5.2.7}$$

where the partial derivatives of e_t satisfy the recursive equations

$$\frac{\partial e_t}{\partial a_i} = - X_{t-i} - \sum_{j=1}^{h} b_j \frac{\partial e_{t-j}}{\partial a_i} - \sum_{s=1}^{g} B_s(t) \frac{\partial e_{t-s}}{\partial a_i} \tag{5.2.8}$$

$(i = 1, 2, \ldots r)$

$$\frac{\partial e_t}{\partial b_k} = - e_{t-k} - \sum_{j=1}^{h} b_j \frac{\partial e_{t-j}}{\partial b_k} - \sum_{s=1}^{g} B_s(t) \frac{\partial e_{t-s}}{\partial b_k} \tag{5.2.9}$$

$(k = 1, 2, \ldots, h)$

$$\frac{\partial e_t}{\partial \beta_{uv}} = - X_{t-u} e_{t-v} - \sum_{j=1}^{h} b_j \frac{\partial e_{t-j}}{\partial \beta_{uv}} - \sum_{s=1}^{g} B_s(t) \frac{\partial e_{t-s}}{\partial \beta_{uv}} \tag{5.2.10}$$

$(v = 1, 2, \ldots, g, \quad u = v, v+1, \ldots, m)$

$$B_s(t) = \sum_{j=s}^{m} \beta_{js} X_{t-j}, \quad s = 1, 2, \ldots, g \tag{5.2.11}$$

In calculating these partial derivatives, we set

$e_t = 0, \quad t = 1, 2, \ldots, n_0 - 1$

and also

$$\frac{\partial e_t}{\partial \theta_i} = 0 \text{ , } t = 1, 2, \ldots , n_0 - 1; \quad i = 1, 2, \ldots , R .$$

In evaluating the second-order partial derivatives we approximate

$$\frac{\partial^2 S(\underline{\theta})}{\partial \theta_i \partial \theta_j} \simeq 2 \sum_{t=n_0}^{n} \left(\frac{\partial e_t}{\partial \theta_i}\right)\left(\frac{\partial e_t}{\partial \theta_j}\right) \tag{5.2.12}$$

as is done in Marquardt algorithm. See also Gabr and Subba Rao [11].

Now let

$$\underset{1\text{x}R}{G^T(\underline{\theta})} = \left[\frac{\partial S(\underline{\theta})}{\partial \theta_1}, \frac{\partial S(\underline{\theta})}{\partial \theta_2}, \ldots , \frac{\partial S(\underline{\theta})}{\partial \theta_R}\right] \tag{5.2.13}$$

and

$$\underset{R\text{x}R}{H(\underline{\theta})} = \left[\frac{\partial^2 S(\underline{\theta})}{\partial \theta_i \partial \theta_j}\right] \tag{5.2.14}$$

Expanding $G(\hat{\underline{\theta}})$ near $\hat{\underline{\theta}} = \underline{\theta}$ in a Taylor series, we obtain

$$0 = G(\underline{\theta}) + H(\underline{\theta})\cdot(\hat{\underline{\theta}} - \underline{\theta})$$

Rewriting this equation, we get

$$(\hat{\underline{\theta}} - \underline{\theta}) = H^{-1}(\underline{\theta})G(\underline{\theta}) \text{ ,}$$

and thus obtain the Newton-Raphson iterative equation

$$\underline{\theta}_{k+1} = \underline{\theta}_k - H^{-1}(\underline{\theta}_k)G(\underline{\theta}_k) \tag{5.2.15}$$

where $\underline{\theta}_k$ is the set of estimates obtained at the k^{th} stage of interation. The estimates obtained by the iterative equations (5.2.15) usually converge, but to obtain a good set of estimates it is necessary that we have good sets of initial values of the parameters. The problem of obtaining the initial values of the parameters is discussed in the next section.

5.3 ORDER DETERMINATION AND INITIAL VALUES

The method of estimating the parameters of the BARMA(r, h, m, ℓ) model (5.1.1) described in section 5.2 is based on the assumption that the order (r, h, m, ℓ) of the model is specified a priori. In practice, the values r, h, m, and ℓ are invariably unknown and suitable values have to be inferred from the data. We give a method based on the observed covariance structure of the data.

We have shown (see Corollary 3.3.2.) that for the bilinear strictly stationary second-order process X_t, $t \in Z$ conforming to the bilinear model (5.1.1) there exists an ARMA $(r, \max(h, g))$, $g = \min(m, \ell)$ with autoregressive coefficients being $a_1, a_2, \ldots, a_r$ and moving average coefficients being functions of $a_1, a_2, \ldots a_r$, $b_1, b_2, \ldots, b_h$ and β_{ij}, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant \ell$, $i \geqslant j$ such that they have identical covariance structures. Thus, given a time series data we can determine r and $q = \max(h, g)$ by using the sample autocovariances of the series and one of the methods of section 3.5. No method is yet available to us for the unique determination of h, m and ℓ. From (5.2.3) and (5.2.4), it is evident that the maximum lag of the input process e_t, $t \in Z$ involved in these difference equations is q.

In view of the above observation, it seems reasonable to consider first the fitting of the best linear ARMA model based on the realisation $\{x_1, x_2, \ldots, x_n\}$. Let the order of this linear ARMA model be (r', q). We then replace r by r', h and ℓ by q to obtain the BARMA(r', q, m, q) model

$$X_t = \sum_{j=1}^{r'} a_j X_{t-j} + \sum_{j=1}^{q} b_j e_{t-j} + \sum_{\substack{i=1 \\ i \geqslant j}}^{m} \sum_{j=1}^{q} \beta_{ij} X_{t-i} e_{t-j} + e_t \tag{5.2.16}$$

The choice of the value of m is made on the basis of the information criterion of Akaike [2] given in section 3.5.2. In choosing the value of m, it is necessary to use the same number of observations over which we wish to fit and compare models for various m values.

To obtain our initial estimates we proceed as follows. When we wish to fit the BARMA(r', q, m, q) model, we choose the coefficients of this model to the corresponding coefficients of the BARMA(r', q, m-1, q) model and set the rest of the g coefficients equal to zero. It may be necessary in some situations to overfit the bilinear part in (5.3.16) to the more elaborate bilinear part

$$\sum_{i=1}^{m} \sum_{\substack{j=1 \\ i \geq j}}^{q+1} \beta_{ij} X_{t-i} e_{t-j} + e_t$$

This kind of overfitting will only be considered when the residuals from the model (5.2.16) do not satisfy the assumptions of normality and independence.

5.4 RESIDUAL ANALYSIS

An important assumption we have made in section 5.1 concerning the errors e_t, $t \in Z$ of the model (5.1.1) is that they are mutually independent and Gaussian. Suppose the correct model is a bilinear model of the form (5.1.1) but an incorrect ARMA model is fitted to the series. Then the residuals from the incorrect ARMA model may be bilinear rather than linear. Similarly, if an incorrect bilinear model is fitted to the series, then the residuals from this incorrect bilinear model will be correlated. One method of differentiating between a pure white noise and a bilinear white noise with the same covariances is to apply covariance analysis to the squares of the series. To check whether the assumptions of pure white noise are satisfied by the errors e_t, $t \in Z$, we examine the serial correlation of the squares of the residuals.

It is important to realize that the residuals are necessarily correlated even if the true errors are independent. In view of this, no elaborate test procedure based on residuals is adopted. We shall only examine the first few values of the serial correlations of both the residuals and the squares of the residuals and see if any are significantly different from zero. We must also bear in mind that if just one value of the serial correlations is significant there would not be enough evidence to reject the model.

In conclusion, we examine the serial correlations of the residuals and the squares of the residuals of the incorrect linear ARMA models fitted to the bilinear models of section 3.5.2. The results are tabulated in Table 5.1. The approximate 95% confidence interval for these values is ±0.09, if true correlation is zero. The correlogram of the residuals themselves do not point to any model inadequacy. The correlogram of the squares of the residuals do suggest that the residuals are not linear in the case of the BARMA processes of Examples 3.1, 3.2, 3.3 and 3.4. The correlogram of the squares of the residuals of the linear AR(2) model of Example 3.5 fails to detect the expected non-linearity in the residuals. In practice, one would not expect to be able to detect model inadequacy or non-linearity of the residuals each time by a mere test of independence or Gaussianity on the residuals. Consequently, this method or any other method of detecting model inadequacy by tests based on residuals can only be used as a general guide.

5.5 FORECASTING

An important use of time series models is to provide forecasts and sometimes the performance of a time series model is judged on the basis of its forecasting performance. Suppose that X_t, $t \in Z$ is a discrete parameter time series and, when at time $t = t_0$, a forecast is required of the future value X_{t_0+k}. Such a forecast has to be based on the past

TABLE 5.1

CORRELOGRAM OF THE RESIDUALS AND SQUARES OF THE RESIDUALS OF THE INCORRECT LINEAR ARMA MODELS FITTED TO THE BARMA PROCESSES IN EXAMPLES 3.1-3.5

LAG	EXAMPLE 3.1		EXAMPLE 3.2		EXAMPLE 3.3		EXAMPLE 3.4		EXAMPLE 3.5	
k	$\hat{\rho}_k(a_t)$	$\hat{\rho}_k(a_t^2)$	$\hat{\rho}_k(a_t)$	$\hat{\rho}_k(a_t^2)$	$\hat{\rho}_k(a_t)$	$\hat{\rho}_k(a_t^2)$	$\hat{\rho}_k(a_t)$	$\hat{\rho}_k(a_t^2)$	$\hat{\rho}_k(a_t)$	$\hat{\rho}_k(a_t^2)$
1	0.009	0.493	-0.006	0.565	-0.003	0.441	-0.001	0.341	-0.029	0.031
2	0.012	0.174	0.022	0.317	0.002	0.243	-0.005	0.089	0.036	-0.016
3	-0.044	0.042	-0.011	0.111	-0.016	0.126	-0.015	0.067	-0.029	-0.077
4	-0.006	-0.005	-0.040	0.053	-0.009	0.021	0.001	0.014	-0.028	0.034
5	-0.028	-0.016	0.007	-0.005	-0.016	-0.007	-0.024	0.013	-0.028	-0.034
6	0.029	-0.028	0.065	-0.043	0.034	-0.018	0.077	-0.010	0.036	-0.047
7	0.014	-0.043	-0.024	-0.035	-0.003	-0.005	0.025	-0.015	-0.013	0.002
8	-0.033	-0.057	0.008	-0.039	0.013	-0.054	-0.047	-0.041	-0.049	-0.089
9	-0.011	-0.029	-0.054	-0.040	-0.080	-0.021	-0.006	-0.031	0.008	0.003
10	-0.086	0.005	-0.060	-0.022	-0.049	0.001	-0.069	0.018	-0.050	0.012
11	0.031	0.036	0.038	-0.012	0.031	0.042	-0.004	0.104	0.032	0.015
12	0.046	0.005	0.020	-0.007	0.014	-0.004	0.026	-0.007	0.063	0.043
13	-0.016	-0.007	-0.013	-0.032	-0.010	0.015	0.005	-0.036	-0.047	0.063
14	-0.068	-0.031	-0.073	-0.033	-0.023	-0.032	0.004	-0.030	-0.059	0.032
15	-0.072	-0.049	-0.045	-0.027	-0.080	-0.028	-0.056	-0.049	-0.090	-0.068
16	-0.021	-0.047	-0.037	-0.015	-0.037	-0.021	-0.068	-0.036	-0.001	-0.033
17	-0.062	-0.054	-0.048	-0.042	-0.054	-0.043	-0.066	-0.023	-0.056	-0.007
18	0.030	-0.037	0.008	-0.033	0.035	-0.041	0.055	-0.011	0.018	0.023
19	0.029	-0.029	0.028	-0.048	0.038	-0.054	0.048	-0.042	0.014	-0.005
20	-0.060	-0.065	-0.047	-0.059	-0.079	-0.063	-0.054	-0.092	-0.031	-0.018

and present of the series, ie X_s, $s \leqslant t_0$. Denote the forecast made at time $t = t_0$ for k-steps ahead by $X_{t_0}(k)$. The forecast error is defined by

$$e_{t_0}(k) = X_{t_0+k} - X_{t_0}(k) \tag{5.5.1}$$

while the k-step forecast error variance or expected square error is defined by

$$\sigma_e^2(k) = E(e_{t_0}^2(k))$$

$$= E[(X_{t_0+k} - X_{t_0}(k))^2] \tag{5.5.2}$$

Then it is well known that $\sigma_e^2(k)$ is minimum if and only if

$$X_{t_0}(k) = E[X_{t_0+k}/X_s \,,\; s \leqslant t_0] \tag{5.5.3}$$

For a bilinear model of the form (5.1.1), formula (5.5.3) can be used to form forecasts, provided the model is invertible. Our rule for forming forecasts is as follows. Write down the equation for X_{t_0+k}; everything on the right-hand side that has already occurred at time t_0 is given its observed value, anything that has yet to occur is replaced by its conditional expectation. Applying this rule to model (5.1.1) we obtain

$$X_{t_0}(k) = \sum_{j=1}^{r} \alpha_j \bar{X}_{t_0}(k-j) + \sum_{j=1}^{h} b_j \bar{e}_{t_0}(k-j)$$

$$+ \sum_{\substack{i=1 \\ i \geqslant j}}^{m} \sum_{j=1}^{\ell} \beta_{ij} E_c[X_{t_0+k-i} e_{t_0+k-j}] \tag{5.5.4}$$

where E_c denotes the conditional expectation given the semi-infinite realisation X_s, $s \leqslant t_0$ and

$$\left.\begin{aligned} \bar{X}_{t_0}(k-j) &= X_{t_0+k-j}\,, \quad \text{for } j \geqslant k \\ &= X_{t_0}(k-j)\,, \text{ for } j < k \end{aligned}\right\} \tag{5.5.5}$$

$$\left.\begin{aligned} \bar{e}_{t_0}(k-j) &= e_{t_0+k-j}\,, \text{ for } j \geqslant k \\ &= 0 \qquad\quad, \text{ for } j < k \end{aligned}\right\} \tag{5.5.6}$$

Since e_t, $t \in Z$ is assumed to be independent of X_s, $s < t$ in (5.1.1), it is not difficult to show that as far as model (5.1.1) is concerned

$$\left.\begin{aligned} E(X_{t-i}\, e_{t-j}) &= \sigma^2 \ , \text{ for } i = j \\ &= 0 \quad , \text{ for } i > j \end{aligned}\right\} \qquad (5.5.7)$$

Thus,

$$E_c[X_{t_0+k-i}\, e_{t_0+k-j}]$$

$$\left.\begin{aligned} &= X_{t_0+k-i}\, e_{t_0+k-j} \ , && \text{for } j \geqslant k \text{ and } i \geqslant j \\ &= \sigma^2\, \beta_{jj} \ , && \text{for } j < k \text{ and } i = j \\ &= 0 \ , && \text{for } j < k \text{ and } i > j \end{aligned}\right\} \qquad (5.5.8)$$

We now use (5.5.4) to write down the forecasting expressions for the BARMA processes of section 3.5.

(a) <u>BARMA Process of Example 3.1</u>

Model:

$$X_t = a\, X_{t-1} + b e_{t-1} + \beta X_{t-1} e_{t-1} + e_t$$

Forecasting Expressions:

$$X_{t_0}(1) = a\, X_{t_0} + b e_{t_0} + \beta\, X_{t_0} e_{t_0}$$

$$X_{t_0}(k) = a\, X_{t_0}(k-1) + \sigma^2\, \beta \ , \ k > 1$$

(b) <u>BARMA Process of Example 3.2</u>

Model:

$$X_t = a_1\, X_{t-1} + a_2\, X_{t-2} + (\beta_{11}\, X_{t-1} + \beta_{21}\, X_{t-2}) e_{t-1} + \beta_{22}\, X_{t-2} e_{t-2} + e_t$$

Forecasting Expressions:

$$X_{t_0}(1) = a_1\, X_{t_0} + a_2\, X_{t_0-1} + \beta_{11}\, X_{t_0} e_{t_0} + \beta_{21}\, X_{t_0-1} e_{t_0}$$

$$+ \beta_{22}\, X_{t_0-1} e_{t_0-1}$$

$$X_{t_0}(2) = a_1\, X_{t_0}(1) + a_2\, X_{t_0} + \beta_{22}\, X_{t_0} e_{t_0} + \sigma^2\, \beta_{11}$$

$$X_{t_0}(k) = a_1\, X_{t_0}(k-1) + a_2\, X_{t_0}(k-2) + \sigma^2(\beta_{11} + \beta_{22}) \ , \ k > 2$$

(c) BARMA Process of Example 3.3

Model:

$$X_t = b_1 e_{t-1} + b_2 e_{t-2} + \beta_{11} X_{t-1} e_{t-1} + \beta_{22} X_{t-2} e_{t-2} + e_t$$

Forecasting Expressions:

$$X_{t_0}(1) = b_1 e_{t_0} + b_2 e_{t_0-1} + \beta_{11} X_{t_0} e_{t_0} + \beta_{22} X_{t_0-1} e_{t_0-1}$$

$$X_{t_0}(2) = b_2 e_{t_0} + \beta_{22} X_{t_0} e_{t_0} + \sigma^2 \beta_{11}$$

$$X_{t_0}(k) = \sigma^2(\beta_{11} + \beta_{22}) , \; k > 2$$

$$= E(X_t)$$

$$= \mu$$

(d) BARMA Process of Example 3.4

Model:

$$X_t = \theta_1 X_{t-1} e_{t-1} + \theta_2 X_{t-2} e_{t-2} + \theta_3 X_{t-3} e_{t-3} + e_t$$

Forecasting Expressions:

$$X_{t_0}(1) = \theta_1 X_{t_0} e_{t_0} + \theta_2 X_{t_0-1} e_{t_0-1} + \theta_3 X_{t_0-2} e_{t_0-2}$$

$$X_{t_0}(2) = \theta_2 X_{t_0} e_{t_0} + \theta_3 X_{t_0-1} e_{t_0-1} + \sigma^2 \theta_1$$

$$X_{t_0}(3) = \theta_3 X_{t_0} e_{t_0} + \sigma^2(\theta_1 + \theta_2)$$

$$X_{t_0}(k) = \sigma^2(\theta_1 + \theta_2 + \theta_3) , \; k > 3$$

$$= E(X_t)$$

$$= \mu$$

(e) BARMA Process of Example 3.5

Model:

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + b_1 X_{t-2} e_{t-1} + b_2 X_{t-3} e_{t-1} + e_t$$

Forecasting Expressions:

$$X_{t_0}(1) = a_1 X_{t_0} + a_2 X_{t_0-1} + b_1 X_{t_0-1} e_{t_0} + b_2 X_{t_0-2} e_{t_0}$$

$$X_{t_0}(2) = a_1 X_{t_0}(1) + a_2 X_{t_0}$$

$$X_{t_0}(k) = a_1 X_{t_0}(k-1) + a_2 X_{t_0}(k-2), \quad k > 2$$

The evaluation of $X_{t_0}(k)$ from the model depends on the unknown parameters. Typically, we substitute the estimates of the parameters obtained by the methods of section 5.2. We estimate σ^2 by

$$\hat{\sigma}^2 = S(\underline{\theta})/(n - n_0 + 1) \tag{5.5.9}$$

The forecasts thus obtained are denoted by $\hat{X}_{t_0}(k)$, $k = 1, 2, 3, \ldots$; and the error by

$$\hat{e}_{t_0}(k) = X_{t_0+k} - \hat{X}_{t_0}(k) \tag{5.5.10}$$

The mean sum of squares of the forecast errors for the period $(t_0 + k, t_0 + k + 1, \ldots, t_0 + k + n)$ is given by

$$\hat{\sigma}^2_{\hat{e}}(k) = \frac{1}{n} \sum_{j=1}^{n} \hat{e}^2_{t_0+j}(k) \tag{5.5.11}$$

The expression $\hat{\sigma}^2_{\hat{e}}(k)$ will be used to measure the superiority of one model over the others.

5.6 NUMERICAL ILLUSTRATIONS

5.6.1 Simulation Studies

(a) BARMA Process of Example 3.1

For the simulated series of Example 3.1 we identified and fitted an ARMA (1, 1) model. Here $r' = q = 1$, and using (5.2.16), we consider the bilinear model

$$X_t = a X_{t-1} + b e_{t-1} + \left(\sum_{j=1}^{m} \beta_{j1} X_{t-j} \right) e_{t-1} + e_t \tag{5.6.1}$$

The AIC value is found to be minimum when $m = 1$ and the estimates obtained are $\hat{a} = 0.5137$, $\hat{b} = 0.3832$, $\hat{\beta} = 0.3057$ and $\hat{\sigma}^2 = 1.0052$. The AIC value is 8.5738.

The true model is the BARMA(1, 1, 1, 1) model

$$X_t = 0.5X_{t-1} + 0.4e_{t-1} + 0.3X_{t-1}e_{t-1} + e_t$$

with $\sigma^2 = 1.0$.

(b) BARMA Process of Example 3.2

The AIC analysis of this series under linear model assumption was less clear cut. The minimum AIC value is attained with an AR(2) model, but as was pointed out in section 3.5 the AIC values for the AR(2), ARMA(1, 2), ARMA(2. 1) and ARMA(2, 2) models are very close. In practice, one would consider all these ARMA models before reaching a decision on the best bilinear model. However, using the ARMA(2, 2) model as the correct linear model, we entertain the bilinear model

$$X_t = \sum_{j=1}^{2} a_j X_{t-j} + \sum_{j=1}^{2} b_j e_{t-j} + \sum_{\substack{i=1 \\ i \geq j}}^{m} \sum_{j=1}^{2} \beta_{ij} X_{t-i} e_{t-j} + e_t \quad (5.6.2)$$

We noted in chapter 3 that the bilinear model (5.1.1) has an ARMA covariance structure with or without the moving average part

$\sum_{j=1}^{h} b_j e_{t-j}$ in (5.1.1). In view of this we also consider model (5.6.2) without the moving average part. When the moving part is missing in (5.6.2) we obtain $q = \min(m, 2) = 2$. Generally, m connot be less than the number of moving average coefficients in the identified linear ARMA model when the moving average part is omitted in (5.1.1).

The AIC value is found to be minimum when m = 2 and without the moving average part in (5.6.2). The estimates obtained are $\hat{a}_1 = 1.0678$, $\hat{a}_2 = -0.2872$, $\hat{\beta}_{11} = 0.2062$, $\hat{\beta}_{21} = 0.1400$, $\hat{\beta}_{22} = -0.1041$ and $\sigma^2 = 0.9916$. The AIC value is 5.8083.

The true model is the BARMA(2, 0, 2, 2) model

$$X_t = 1.10X_{t-1} - 0.30X_{t-2} + 0.20X_{t-1}e_{t-1} + 0.15X_{t-2}e_{t-1} - 0.10X_{t-2}e_{t-2} + e_t$$

with $\sigma^2 = 1.0$.

(c) BARMA Process of Example 3.3

We identified and fitted a linear MA(2) model to this series. So we set $r = 0$ and $h = \ell = 2$, and using (5.2.16) we consider the bilinear model

$$X_t = \sum_{j=1}^{2} b_j e_{t-j} + \sum_{\substack{i=1 \\ i\geq j}}^{m} \sum_{j=1}^{2} \beta_{ij} X_{t-i} e_{t-j} + e_t \qquad (5.6.3)$$

The AIC value is found to be minimum when $m = 2$ and the estimates obtained are $\hat{b}_1 = 0.5381$, $\hat{b}_2 = 0.3492$, $\hat{\beta}_{11} = 0.3843$, $\hat{\beta}_{21} = 0.0172$, $\hat{\beta}_{22} = 0.2879$ and $\hat{\sigma}^2 = 0.9909$. The AIC value is 5.4616.

The estimate of the bilinear coefficient β_{21} is very small when compared with the estimates of the other bilinear coefficients. This suggests a parsimonious model of the form

$$X_t = b_1 e_{t-1} + b_2 e_{t-2} + \beta_{11} X_{t-1} e_{t-1} + \beta_{22} X_{t-2} e_{t-2} + e_t \qquad (5.6.4)$$

On fitting model (5.6.4) to the data we obtain the following estimates: $\hat{b}_1 = 0.5353$, $\hat{b}_2 = 0.3347$, $\hat{\beta}_{11} = 0.4015$, $\hat{\beta}_{22} = 0.3047$ and $\hat{\sigma}^2 = 0,9950$. The AIC value is 5.5038.

The true model is the BARMA(0, 2, 2, 2) model

$$X_t = 0.55 e_{t-1} + 0.35 e_{t-2} + 0.40 X_{t-1} e_{t-1} + 0.30 X_{t-2} e_{t-2} + e_t$$

with $\sigma^2 = 1.0$.

(d) BARMA Process of Example 3.4

We identified and fitted a linear MA(3) model to this series. So we set $r = 0$ and $h = \ell = 3$, and using (5.2.16) we consider the bilinear model

$$X_t = \sum_{j=1}^{3} b_j e_{t-j} + \sum_{\substack{i=1 \\ i\geq j}}^{m} \sum_{j=1}^{3} \beta_{ij} X_{t-i} e_{t-j} + e_t \qquad (5.6.5)$$

The AIC value is found to be minimum when $m = 3$ and without the moving part $\sum_{j=1}^{3} b_j e_{t-j}$ in (5.6.5). The estimates obtained are $\hat{\beta}_{11} = 0.4326$,

$\hat{\beta}_{21} = -0.0160$, $\hat{\beta}_{31} = -0.0200$, $\hat{\beta}_{22} = 0.3479$. $\hat{\beta}_{32} = -0.0251$, $\hat{\beta}_{33} = 0.2247$ and $\hat{\sigma}^2 = 0.9936$. The AIC value is 8.8153. Now some of the bilinear coefficients are very small when compared to other bilinear coefficients. This suggests a parsimonious model of the form

$$X_t = \beta_{11} X_{t-1}e_{t-1} + \beta_{22} X_{t-2}e_{t-2} + \beta_{33} X_{t-3}e_{t-3} + e_t \qquad (5.6.6)$$

The estimates obtained on fitting model (5.6.6) to the data are $\hat{\beta}_{11} = 0.4116$, $\hat{\beta}_{22} = 0.3312$, $\hat{\beta}_{33} = 0.2206$, and $\hat{\sigma}^2 = 0.9996$. The AIC value is 5.8257.

The true model is the BARMA(0, 0, 3, 3) model

$$X_t = 0.40X_{t-1}e_{t-1} + 0.30X_{t-2}e_{t-2} + 0.20X_{t-3}e_{t-3} + e_t$$

with $\sigma^2 = 1.0$.

(e) BARMA Process of Example 3.5

We identified and fitted a linear AR(2) model to this series. As was pointed out in section 3.4, the bilinear model (5.1.1) without the moving average part could have the same covariance structure as some linear autoregressive process. We were unable to give a general condition under which the bilinear model (5.1.1) without the moving average part could have an autoregressive covariance structure. However, we do know that the autoregressive coefficients of the linear AR model are the same as the autoregressive coefficients of the bilinear model. Suppose the correct model for a time series data is the bilinear model

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + \sum_{i=1}^{m} \sum_{\substack{j=1 \\ i\geq j}}^{\ell} \beta_{ij} X_{t-i}e_{t-j} + e_t , \qquad (5.6.7)$$

but the best linear ARMA model that fits the data is the AR(r) model

$$X_t = \sum_{j=1}^{r} a_j X_{t-j} + Z_t. \qquad (5.6.8)$$

Then the errors, Z_t from (5.6.8) are in effect bilinear rather than linear. It follows that we can fit a purely bilinear model of the form

$$Z_t = \sum_{\substack{i=1 \\ i \geq j}}^{m'} \sum_{j=1}^{\ell'} \beta'_{ij} Z_{t-i} e'_{t-j} + e'_t \qquad (5.6.9)$$

to the errors. We can use (5.6.9) to modify (5.6.8) to a bilinear model of the form (5.6.7). This will in general be our modus operandi for fitting bilinear models to time series data that admits an autoregressive model as the best linear ARMA model. Use of residuals to modify a linear model to a bilinear model with identical covariance structure can be applied in all cases of fitting bilinear models to time series data.

The AR(2) model obtained was

$$X_t = 1.2059X_{t-1} - 0.6053X_{t-2} + Z_t \qquad (5.6.10)$$

with $\text{Var}(Z_t) = 1.116$. To the residuals Z_t, we have fitted the bilinear white noise model

$$Z_t = (0.2467Z_{t-2} + 0.1782Z_{t-3})e_{t-1} + e_t \qquad (5.6.11)$$

with $\hat{\sigma}^2 = 1.0280$. On eliminating Z_t between (5.6.10) and (5.6.11), we obtain

$$X_t = 1.2059X_{t-1} - 0.6053X_{t-2} + (0.2467X_{t-2} - 0.1193X_{t-3} - 0.0656X_{t-4} + 0.1079X_{t-5})e_{t-1} + e_t \qquad (5.6.12)$$

which suggests that the bilinear model

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + (\beta_{21} X_{t-2} + \beta_{31} X_{t-3} + \beta_{41} X_{t-4} + \beta_{51} X_{t-5})e_{t-1} + e_t , \qquad (5.6.13)$$

should now be entertained.

On fitting model (5.6.13) to the original series we obtained the following estimates: $\hat{a}_1 = 1.2009$, $\hat{a}_2 = -0.6076$, $\beta_{21} = 0.2377$, $\hat{\beta}_{31} = -0.1115$, $\hat{\beta}_{41} = -0.0715$, $\hat{\beta}_{51} = 0.0415$ and $\hat{\sigma}^2 = 1.0047$. The AIC value is 14.3026. The parsimonious model obtained is the BARMA(2. 0, 3, 1) model

$$X_t = 1.2006X_{t-1} - 0.6059X_{t-2} + 0.2484X_{t-2}e_{t-1} - 0.1648X_{t-3}e_{t-1} + e_t \qquad (5.6.14)$$

with $\hat{\sigma}^2 = 1.0077$, AIC = 11.8149.

The true model is the BARMA(2, 0, 3, 1) model

$$X_t = 1.20X_{t-1} - 0.61X_{t-2} + 0.25X_{t-2}e_{t-1} - 0.15X_{t-3}e_{t-1} + e_t$$

with $\sigma^2 = 1.0$.

5.6.2 Fitting of BARMA Models To Real Data

(a) Ben Nevis Temperatures

First we consider the daily drybulb temperatures (^{0}F) at noon on Ben Nevis for the days 1st., February - 18th., August 1884, giving two hundred observations. The series is referred to in Anderson [3] as series A*. The graph of this data is plotted in Figure 5.1(a) and its sample autocorrelations up to fifty lags are plotted in Figure 5.1(b). The sample autocorrelation function shows that there is a linear trend in the series. In order to remove the trend, Anderson [3, p.112-116] has differenced and obtained the series

$$X_t = (1 - B)Y_t$$

$$Y_t = \text{original series}$$

The series X_t thus obtained seems to be free from trend and the following linear model is identified, estimated and diagnostically checked by Anderson [3, p.116]:

$$X_t = -0.238a_{t-1} - 0.305a_{t-2} + a_t \qquad (5.6.15)$$

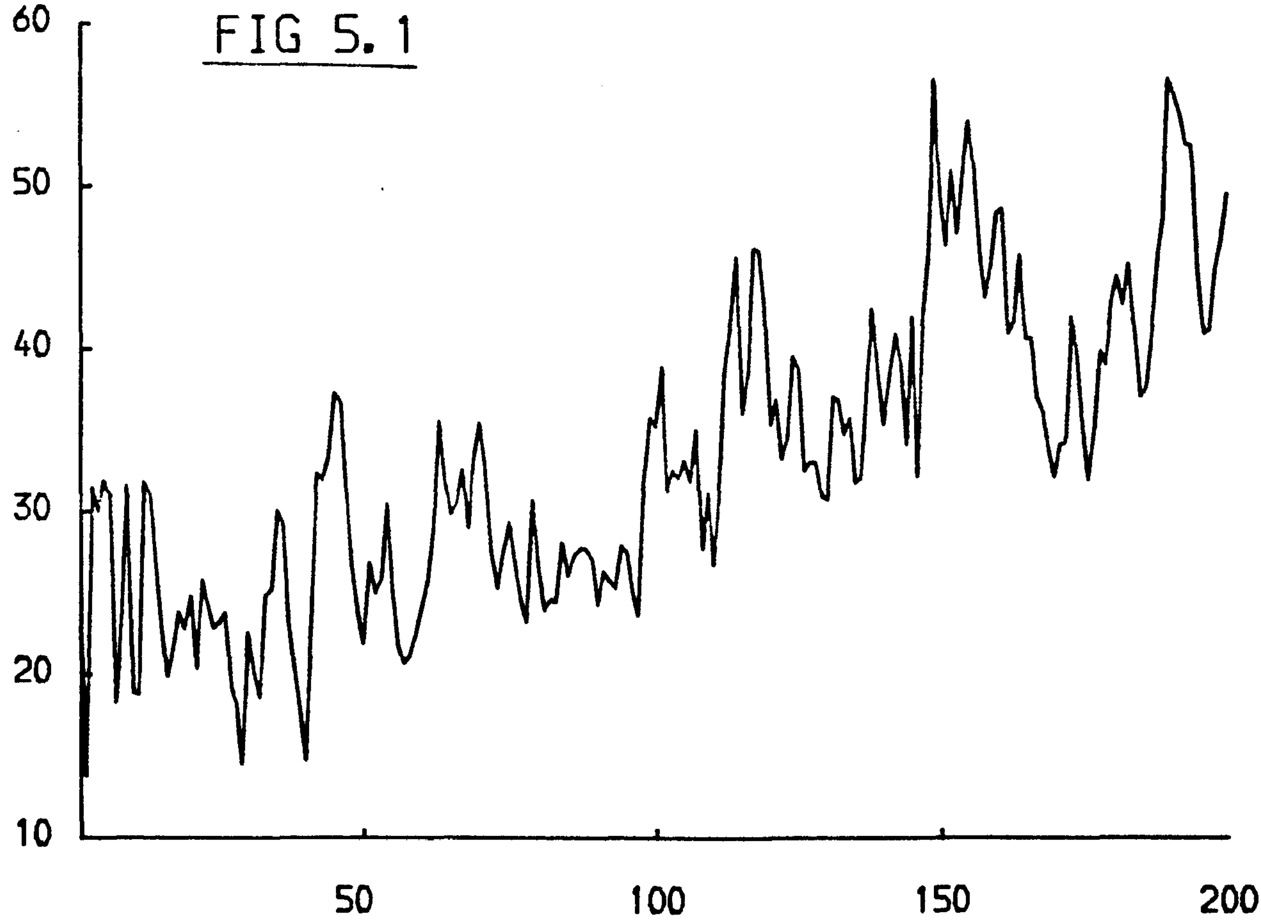
with $\text{Var}(a_t) = 17.91$.

We use the first one hundred and eighty observations for model fitting, and the next twenty observations are used for forecasting purpose. The MA(2) model fitted to the one hundred and eighty observations is the model

$$X_t = -0.2751a_{t-1} - 0.3218a_{t-2} + a_t \qquad (5.6.16)$$

with $\text{Var}(a_t) = 18.0$, where X_t is change in temperature. Since the best linear ARMA model is an MA(2) model, we entertain the model

FIG 5.1



(a) DAILY DRYBULB TEMPERATURES AT NOON ON BEN NEVIS

(b) AUTOCORRELATION FUNCTION OF THE BEN NEVIS TEMPERATURES

$$X_t = \sum_{j=1}^{2} b_j e_{t-j} + \sum_{\substack{i=1 \\ i \geq j}}^{m} \sum_{j=1}^{2} \beta_{ij} X_{t-i} e_{t-j} + e_t \qquad (5.6.17)$$

The AIC value is found to be minimum when m = 2 and the parsimonious BARMA(0, 2, 2, 2) model obtained is

$$X_t = -0.2217e_{t-1} - 0.3060e_{t-2} - 0.0036X_{t-1}e_{t-1} + 0.0175X_{t-2}e_{t-2} + e_t \qquad (5.6.18)$$

where X_t is change in temperature and $\sigma^2 = 16.10$, AIC = 499.92. This leads to 10.6% reduction in error variance.

The first twenty serial correlations of the residuals from the models (5.6.16) and (5.6.18) are given in columns 2 and 3 respectively of Table 5.2. Given in columns 2 and 3 of Table 5.3 are the first twenty serial correlations of the squares of the residuals from the models (5.6.16) and (5.6.18) respectively. The approximate 95% confidence interval for these values is ±0.15, if true correlation is zero. From Table 5.3, it is clear that the values for k = 2, 3, certainly appear significant for the model (5.6.16).

For the next twenty days, both model (5.6.16) and (5.6.18) were used to forecast and the results are given in Table 5.4. The mean sum of squares of the one-step-ahead forecast errors are:

	MA(2) (5.6.16)	BARMA(0, 2, 2, 2) (5.6.18)
First 10 days	21.7	19.4
First 15 days	19.5	17.5
First 20 days	17.5	16.0

The bilinear model (5.6.18) reduces mean square error (MSE) by 10.6% for the first ten days, 10.3% for the first fifteen days and 8.6% for the first twenty days. The bilinear model produces a small but consistent reduction in the mean square error. The performance of the bilinear

TABLE 5.2

CORRELOGRAM OF THE RESIDUALS OF THE ARMA AND BARMA MODELS FITTED TO THE SERIES OF SECTION 5.6.2

LAG k	BEN NEVIS TEMPERATURES		IBM CLOSING STOCK PRICES		SUNSPOT NUMBERS	
	MA(2) (5.6.16) $\hat{\rho}_k(a_t)$	BARMA(0, 2, 2, 2) (5.6.18) $\hat{\rho}_k(e_t)$	MA(1) (5.6.20) $\hat{\rho}(a_t)$	BARMA(0, 1, 1, 1) (5.6.22) $\hat{\rho}(e_t)$	ARMA(8, 1) (5.6.23) $\hat{\rho}(a_t)$	BARMA(8, 1, 5, 2) (5.6.24) $\hat{\rho}(e_t)$
1	0.013	0.007	-0.014	-0.004	0.002	0.039
2	0.068	0.034	-0.077	-0.055	-0.013	0.109
3	-0.048	-0.072	-0.141	-0.117	-0.007	-0.001
4	-0.064	-0.031	0.112	0.110	-0.004	0.081
5	-0.017	-0.042	0.044	0.081	0.012	0.028
6	-0.035	-0.052	-0.044	-0.049	0.033	0.057
7	-0.008	0.076	-0.004	-0.007	0.036	0.101
8	-0.019	-0.041	0.040	0.041	0.119	-0.049
9	0.017	-0.014	-0.069	-0.043	0.107	0.081
10	-0.021	0.011	0.016	-0.017	-0.020	0.031
11	0.000	0.029	0.057	0.086	0.034	0.049
12	-0.006	0.015	-0.002	-0.003	-0.008	0.074
13	-0.045	-0.029	-0.078	-0.087	-0.107	0.027
14	-0.069	-0.065	0.007	0.015	-0.054	0.010
15	-0.144	-0.131	0.069	0.031	0.024	0.100
16	-0.061	-0.030	0.059	0.069	0.003	0.093
17	-0.005	-0.015	-0.030	-0.050	0.173	0.131
18	0.025	0.087	-0.046	-0.061	-0.043	-0.048
19	0.055	0.026	-0.084	-0.067	-0.032	0.001
20	-0.078	-0.047	0.138	0.127	-0.088	0.064

TABLE 5.3

CORRELOGRAM OF THE SQUARES OF THE RESIDUALS OF THE ARMA AND BARMA MODELS FITTED TO THE SERIES OF SECTION 5.6.2

LAG	BEN NEVIS TEMPERATURES		IBM CLOSING STOCK PRICES		SUNSPOT NUMBERS	
k	MA(2) (5.6.16) $\hat{\rho}_k(a_t^2)$	BARMA(0, 2, 2, 2) (5.6.18) $\hat{\rho}_k(e_t^2)$	MA(1) (5.6.20) $\hat{\rho}(a_t^2)$	BARMA(0, 1, 1, 1) (5.6.22) $\hat{\rho}(e_t^2)$	ARMA(8, 1) (5.6.23) $\hat{\rho}(a_t^2)$	BARMA(8, 1, 5, 2) (5.6.24) $\hat{\rho}(e_t^2)$
1	0.090	0.006	0.230	0.130	0.281	0.126
2	0.199	0.123	0.055	0.061	0.154	0.119
3	0.160	0.135	-0.034	-0.003	-0.039	0.003
4	0.093	-0.015	-0.036	-0.015	-0.048	-0.052
5	0.099	0.032	-0.059	-0.075	-0.075	0.032
6	-0.032	-0.044	-0.075	-0.078	-0.070	-0.007
7	0.094	-0.008	-0.049	-0.053	-0.015	-0.047
8	0.006	0.010	-0.050	-0.060	0.037	0.141
9	0.073	0.019	-0.009	-0.044	0.128	0.035
10	-0.044	-0.087	0.021	0.008	0.026	0.027
11	-0.040	-0.004	0.075	0.052	0.020	0.074
12	-0.036	-0.006	0.034	0.019	-0.015	-0.058
13	-0.013	-0.031	0.038	0.105	-0.027	-0.071
14	-0.066	-0.054	0.113	0.081	-0.079	0.014
15	-0.056	-0.031	-0.001	-0.001	0.020	0.014
16	-0.054	-0.020	0.030	0.024	-0.018	0.050
17	-0.111	-0.128	0.005	0.063	-0.054	0.059
18	0.006	0.005	-0.010	0.010	-0.034	-0.080
19	-0.049	-0.066	0.036	0.024	-0.007	0.066
20	-0.028	0.025	0.013	-0.007	0.017	-0.034

TABLE 5.4

FORECASTING THE BEN NEVIS TEMPERATURES FROM MODELS BASED ON 180 OBSERVATIONS

t	X_t	ONE-STEP-AHEAD PREDICTIONS		MORE THAN ONE-STEP-AHEAD PREDICTIONS	
		Predicted X_t MA(2) (5.6.16)	Predicted X_t BARMA(0, 2, 2, 2) (5.6.18)	Predicted X_t MA(2) (5.6.16)	Predicted X_t BARMA(0, 2, 2, 2) (5.6.18)
181	44.4	41.2	41.7	41.2	41.7
182	42.7	41.8	42.7	39.5	40.5
183	45.2	41.4	41.9	39.5	40.7
184	41.0	43.9	44.4	39.5	41.0
185	37.0	40.6	40.9	39.5	41.2
186	37.5	38.9	39.1	39.5	41.4
187	41.0	39.0	39.3	39.5	41.6
188	45.0	40.9	41.1	39.5	41.9
189	47.9	43.2	43.7	39.5	42.1
190	56.5	45.3	46.0	39.5	42.3
191	55.6	51.9	52.8	39.5	42.5
192	54.2	51.0	53.3	39.5	42.8
193	52.5	52.1	53.1	39.5	43.0
194	52.4	51.4	52.3	39.5	43.2
195	44.9	52.0	52.6	39.5	43.4
196	40.9	46.5	46.4	39.5	43.7
197	41.1	44.7	45.5	39.5	43.9
198	44.8	43.9	44.1	39.5	44.1
199	46.7	45.7	45.9	39.5	44.3
200	49.4	46.1	46.4	39.5	44.6

model seems to be better when it is used for forecasting several steps ahead.

(b) IBM Common Stock Closing Prices

For our second example we consider the IBM daily common stock closing prices for one hundred and eighty-five trading days starting 17th., May 1961. The original data which consists of three hundred and sixty-nine observations is series B in Box and Jenkins [5]. Box and Jenkins fitted MA(1) models separately to the first and second halves of the differenced series, as well as to the complete series after differencing. Using the results obtained by fitting the MA(1) models they produced evidence that in later periods the MA(1) model suffers a significant change in parameter value. We confine ourselves to the first half of the series.

The graph of the closing prices for the first one hundred and eighty-five trading days is plotted in Figure 5.2(a) and its sample autocorrelations up to forty-six lags are plotted in Figure 5.2(b). The sample autocorrelation function shows that there is a linear trend in the series. Granger and Andersen [15] have considered the first one hundred and sixty-nine trading days. They have fitted the bilinear model

$$Z_t = 0.02 Z_{t-1} e_{t-1} + e_t \qquad (5.6.19)$$

to the residuals Z_t, obtained from the MA(1) model

$$X_t = 0.26 Z_{t-1} + Z_t , \qquad (5.6.20)$$

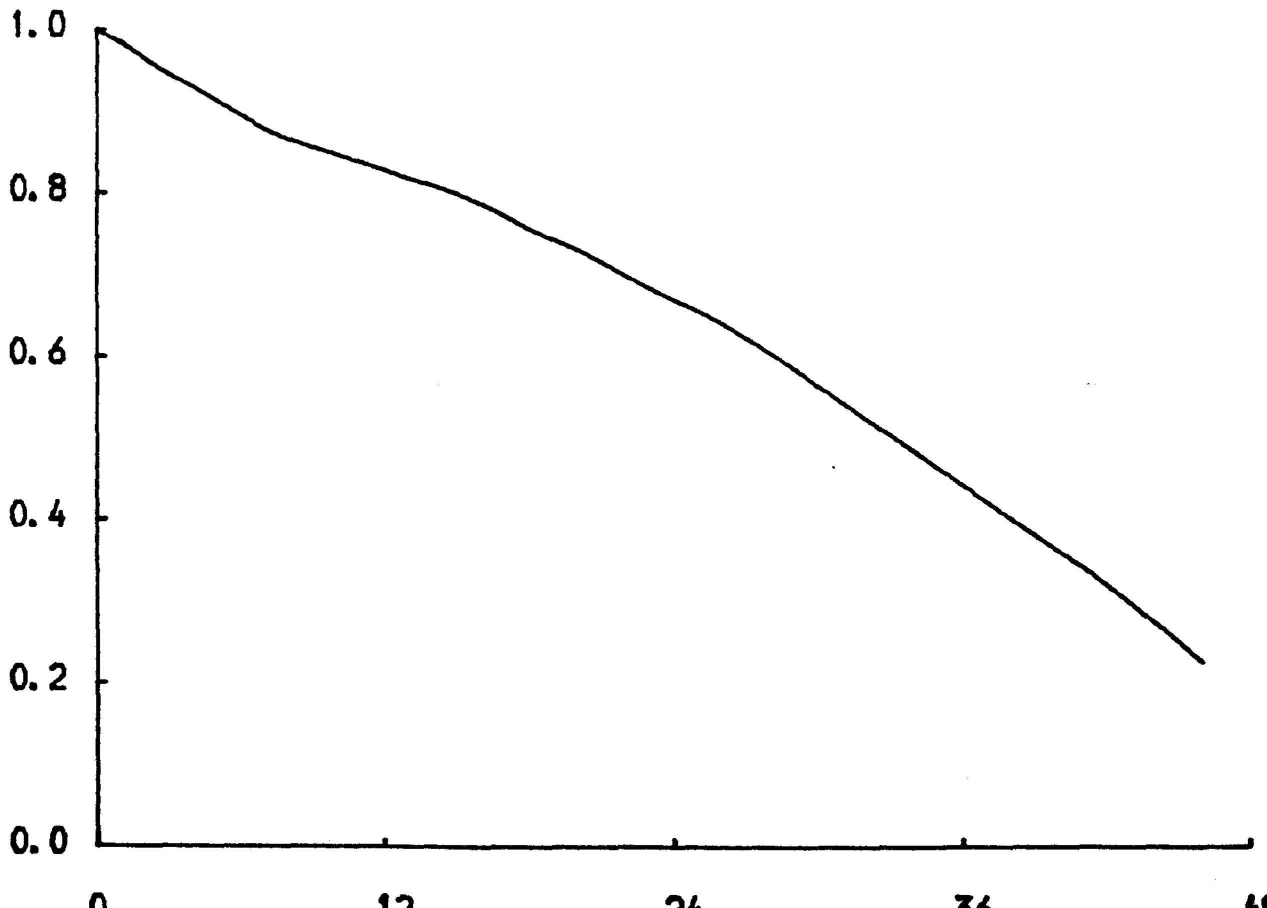
where X_t is change in price. Our interest in this series is to illustrate further the use of residuals to modify a linear model to a bilinear model of the form (5.1.1)

The MA(1) model fitted to the first one hundred and eighty-five observations is the model

$$X_t = 0.2728 a_{t-1} + a_t \qquad (5.6.20)$$

FIG 5.2

(a) IBM DAILY COMMON STOCK CLOSING PRICES



(b) AUTOCORRELATION FUNCTION OF THE IBM CLOSING STOCK PRICES with $\text{Var}(a_t) = 26.49$, where X_t is change in price. To the residuals a_t, we have fitted the bilinear model

$$a_t = 0.0259 a_{t-1} e_{t-1} + e_t \tag{5.6.21}$$

with $\hat{\sigma}^2 = 25.08$. On eliminating a_t between (5.6.20) and (5.6.21) we obtain

$$X_t = 0.2728 e_{t-1} + 0.0259 X_{t-1} e_{t-1} + e_t$$

which suggests that the bilinear model

$$X_t = b e_{t-1} + \beta X_{t-1} e_{t-1} + e_t \tag{5.6.22}$$

should now be entertained. On fitting model (5.6.22) to the first one hundred and eighty-five observations we obtained the following estimates: $\hat{b} = 0.2148$, $\hat{\beta} = 0.0271$ and $\hat{\sigma}^2 = 24.88$. The point to note here is that using our estimation procedure described in sections 5.2 and 5.3, we also arrived at the model (5.6.22). The first twenty serial correlations of the residuals from the models (5.6.20) and (5.6.22) are given in columns 4 and 5 respectively of Table 5.2. Given in columns 4 and 5 of Table 5.3 are the first twenty serial correlations of the squares of the residuals from the models (5.6.20) and (5.6.22) respectively. The approximate 95% confidence interval for these values is ±0.15, if true correlation is zero. From Table 5.3, it is clear that the value for k = 1 certainly appears significant for the model (5.6.20).

We have not calculated forecasts, since the parameter values for the two halves of the series differ significantly.

(c) Wolfer Sunspot Numbers

For our third example we consider the annual sunspot numbers given in Waldmeier [43]. Woodward and Gray [44] have given a list of linear ARMA models fitted in the literature for this yearly data: Granger and Andersen [15] have fitted the bilinear model

$$X_t = 0.202 Z_{t-1} - 0.0222 Z_{t-2} e_{t-1} + e_t$$

to the residuals Z_t, from the AR(2) model

$$X_t = 14.70 + 1.425X_{t-1} - 0.731X_{t-2} + Z_t$$

fitted by Box and Jenkins [5] to the annual sunspot numbers for the period 1770 to 1869. Using the AIC criterion, Subba Rao [37] has fitted a bilinear model of the form (5.2.1) with p = 3 and q = 4 to the annual sunspot numbers for the period 1700 to 1945. Gabr and Subba Rao [11] have also fitted their subset bilinear model (5.2.2) to the annual sunspot numbers for the period 1700 to 1920.

We consider the annual sunspot numbers for the years 1749 to 1924, giving one hundred and seventy-six observations. The series is plotted in Figure 5.3(a) and its sample autocorrelations up to forty-four lags are plotted in Figure 5.3(b). The sample auotcorrelations exhibit an oscillatory behaviour. Woodward and Gray [44] have also considered the one hundred and seventy-six observations and have used the R- and S-arrays of Gray, Kelley and McIntire [17] to identify and fit the following ARMA(8, 1) model

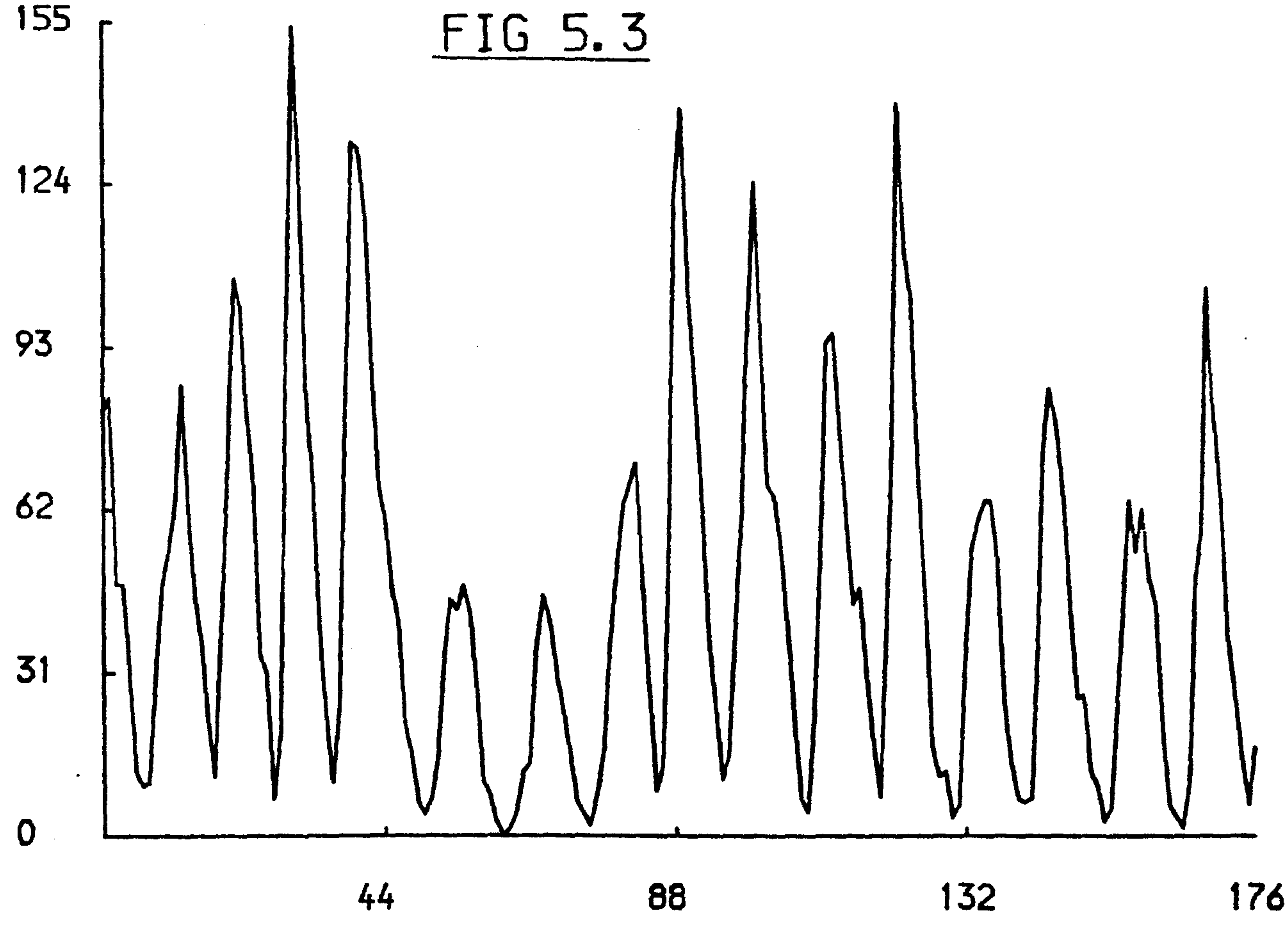
$$(1 - 1.64B + 0.94B^2)(1 - 0.1748B - 0.0309B^2 - 0.0136B^3 - 0.2528B^4 - 0.1429B^5 - 0.1616B^6)X_t = (1 - 0.5972B)a_t \quad (5.6.23)$$

with $Var(a_t) = 215.23$, where $X_t = Y_t - \bar{Y}$ and Y_t denotes the observed sunspot numbers. Model (5.6.23) has roots close to the unit circle. Making use of the observed ARMA(8, 1) covariance structure, we have fitted the following BARMA(8, 1, 5, 2) model to the mean deleted observations

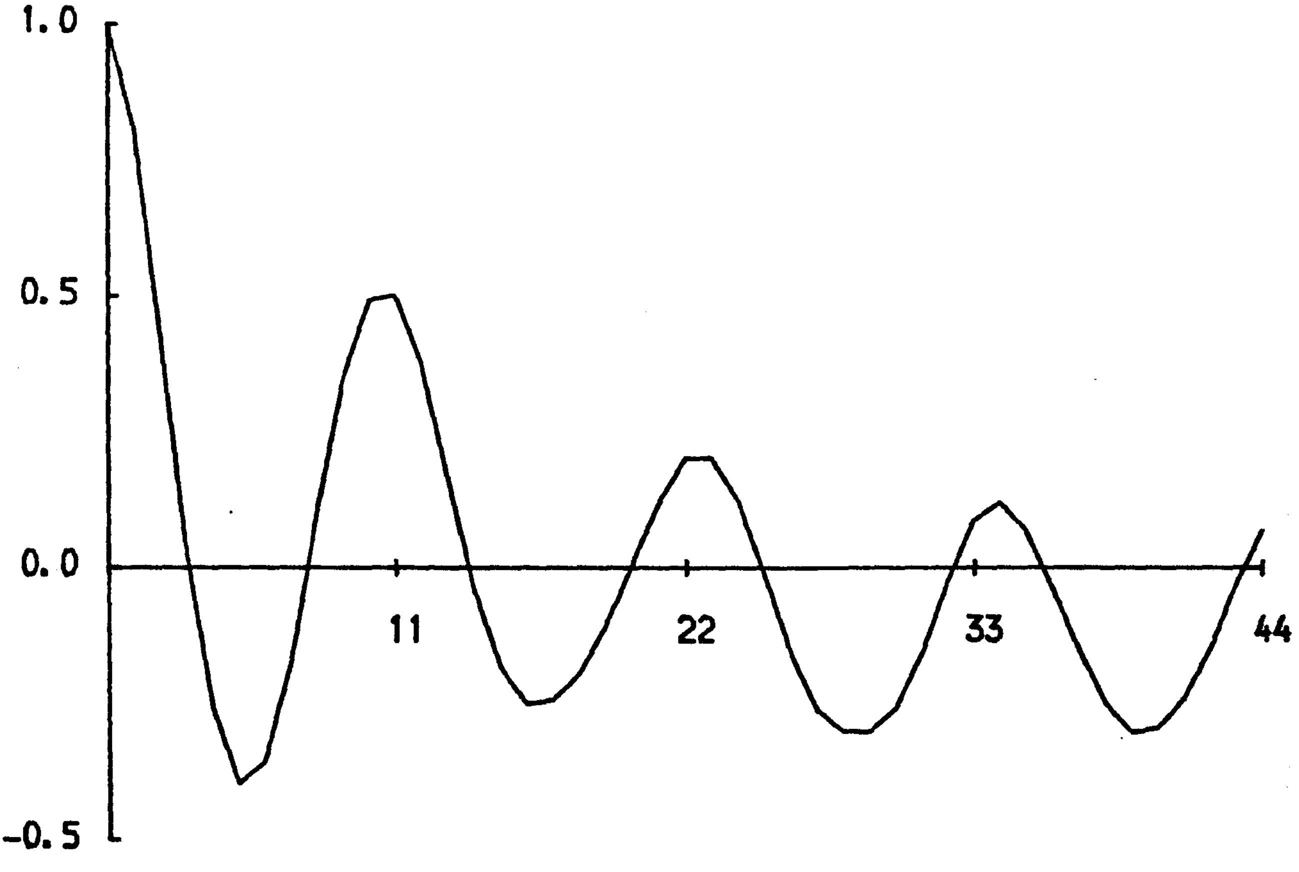
$$\begin{aligned} X_t = {} & 2.0381X_{t-1} - 1.4203X_{t-2} + 0.0704X_{t-3} + 0.2684X_{t-4} - 0.0718X_{t-5} \\ & + 0.0541X_{t-6} - 0.0225X_{t-7} + 0.0112X_{t-8} - 0.8071e_{t-1} \\ & - (0.0304X_{t-2} - 0.0103X_{t-3})e_{t-1} + (0.0322X_{t-3} - 0.0155X_{t-4} \\ & - 0.0038X_{t-5})e_{t-2} + e_t \end{aligned} \quad (5.6.24)$$

with $\hat{\sigma}^2 = 151.43$. We can expand the left hand side of (5.6.23) to obtain

FIG 5.3



(a) SUNSPOT NUMBERS (1749-1924)



(b) AUTOCORRELATION FUNCTION OF THE SUNSPOT NUMBERS

$$X_t = 1.8148X_{t-1} - 1.1958X_{t-2} + 0.1272X_{t-3} + 0.2529X_{t-4} - 0.2589X_{t-5}$$
$$+ 0.1649X_{t-6} - 0.1307X_{t-7} + 0.1519X_{t-8} - 0.5972e_{t-1} + e_t \quad (5.6.25)$$

By comparing (5.6.24) and (5.6.25), one notices some similariites between the autoregressive coefficients of the two models.

The first twenty serial correlations of the residuals from the models (5.6.23) and (5.6.24) are given in columns 6 and 7 respectively of Table 5.2. Given in columns 6 and 7 of Table 5.3 are the first twenty serial correlations of the squares of the residuals from the models (5.6.23) and (5.6.24) respectively. The approximate 95% confidence interval for these values is ±0.15, if true correlation is zero. From Table 5.3, it is clear that the value for k = 1 certainly appears significant for the model (5.6.23).

For the next twenty years, both model (5.6.23) and (5.6.24) were used to forecast and the results are given in Table 5.5. The mean sum of squares of the one-step-ahead forecast errors is 83.69 for the ARMA(8, 1) model (5.6.23) and 49.77 for the BARMA(8, 1, 5, 2) model (5.6.24). Thus, a 40.5% reduction in error mean square results from using the bilinear model over the ARMA model. The bilinear model (5.6.24) also performs better than the linear model (5.6.23) when they are used for forecasting several steps ahead.

TABLE 5.5

FORECASTING THE ANNUAL SUNSPOT NUMBERS FROM MODELS BASED ON THE YEARS 1749 TO 1924

Year	X_t	ONE-STEP-AHEAD PREDICTIONS		MORE THAN ONE-STEP-AHEAD PREDICTIONS	
		Predicted X_t ARMA(8, 1) (5.6.23)	Predicted X_t BARMA(8, 1, 5, 2) (5.6.24)	Predicted X_t ARMA(8, 1) (5.6.23)	Predicted X_t BARMA(8, 1, 5, 2) (5.6.24)
1925	44.3	34.4	39.8	34.4	39.8
1926	63.9	65.2	66.0	53.2	58.6
1927	69.0	75.7	75.0	67.3	69.0
1928	77.8	68.8	68.2	73.3	68.0
1929	64.9	74.6	61.4	69.9	57.7
1930	35.7	51.5	44.0	59.7	42.8
1931	21.2	16.8	16.6	45.1	29.3
1932	11.1	13.9	16.7	30.7	22.0
1933	5.7	10.4	10.7	20.8	22.8
1934	8.7	11.5	12.8	18.4	30.5
1935	36.1	22.2	22.5	23.9	41.7
1936	79.7	57.1	68.3	34.9	52.0
1937	114.4	99.3	100.3	47.8	57.8
1938	109.6	116.0	105.3	58.6	57.6
1939	88.8	88.8	87.6	64.2	52.2
1940	67.8	61.3	61.0	63.3	44.2
1941	47.5	44.6	40.5	56.6	36.6
1942	30.6	26.2	28.4	46.3	32.3
1943	16.3	16.2	20.9	35.8	32.4
1944	9.6	11.7	13.1	28.2	36.3

REFERENCES

[1] Akaike, H. (1966). Note on higher order spectra. Ann. Inst. Statist. Math., 18, 123-126.

[2] Akaike, H. (1977). On entropy maximisation principle. Proc. Symp. on Applications of Statistics (Ed. P. Krishnaiah). Dayton, Ohio.

[3] Anderson, O. D. (1976). Time series analysis and forecasting. The Box-Jenkins approach. Butterworth and Co., London.

[4] Bhaskara Rao, M., Subba Rao, T., and Walker, A. M. (1983). On the existence of some bilinear time series models. To appear in J. of Time Series Analysis.

[5] Box, G. E. P., and Jenkins, G. M. (1970). Time series analysis, forecasting and control. Holden-Day, San Francisco.

[6] Brillinger, D. R. (1965). An introduction to polyspectra. Ann. Math. Statist., 36, 1351-1374.

[7] Brillinger, D. R., and Rosenblatt, M. (1967). Asymptotic theory of estimates of k-th order spectra. In Spectral Analysis of Time Series (Ed. B. Harris). Wiley, New York, 153-188.

[8] Bruni, C., Dupillo, G., and Koch, G. (1974). Bilinear systems: an appealing class of "nearly linear" systems in theory and applications. I.E.E.E. Trans. Auto. Control, AC - 19, 334-338.

[9] Chatfield, C. (1980). The analysis of time sereis: An introduction. Chapman and Hall, London.

[10] Chung, K. L. (1974). A course in probability theory, 2nd ed. Academic Press, London.

[11] Gabr, M. M., and Subba Rao, T. (1981). The estimation and prediction of subset bilinear time series models with applications. J. Time Series Analysis, 2, 3, 155-171.

[12] Gabr, M. M.(1981). Bispectral analysis of non-linear time series and the statistical theory of bilinear time sereis models with applications. Unpublished Ph. D thesis submitted to the University of Manchester Institute of Science and Technology.

[13] Godfrey, M. D. (1965). An exploratory study of the bispectrum of economic time series. Appl. Statist., 14, 48-69.

[14] Granger, C. W. J., and Andersen, A. (1978). Non-linear time series modelling. In Applied Time Series Analysis (Ed. D. F. Findley). Academic Press, New York, 25-38.

[15] Granger, C. W. J., and Andersen, A. (1978). An introduction to bilinear time series models. Vanderhoeck and Reprecht, Gottingen.

[16] Granger, C. W. J., and Andersen, A. (1978). On the invertibility of time sereis models. Stoch. Proc. Appl., 8, 87-92.

[17] Gray, H. L., Kelley, G. D., and McIntire, D. D. (1978). A new approach to ARMA modelling. Commun. Statist., B7, 1-78.

[18] Guegan, D. (1981). Étude d'un modèle non linéaire, le modèle superdiagonal d'ordre 1. C. R. Acad. Sc. Paris, 293, Serie 1, 95-99.

[19] Hallin, M. (1980). Invertibility and generalized invertibility of time series models. J. R. Statist. Soc., B, 42, 210-212.

[20] Hallin, M. (1981). Addendum to 'Invertibility and generalized invertibility of time series models. J. R. Statist. Soc., B, 43, 103.

[21] Halmos, P. R. (1950). Measure theory. D. Van Nostrand Co. London.

[22] Hannan, E. J. (1982). A note on bilinear time series models. Stoch. Proc. Appl., 12, 221-224.

[23] Hasselman, K., Munk, W., and MacDonald, G. (1963). Bispectrum of ocean waves. In Time Series Analysis (Ed. M. Rosenblatt). Wiley, New York, 125-139.

[24] Hinich, M. J. (1982). Testing for Gaussianity and linearity of a stationary time series. J. Time Series Analysis, 3, 3, 169-176.

[25] Kato, T. (1966). Perturbation theory for linear operators. Springer-Verlag, Berlin.

[26] Lancaster, P. (1969). Theory of matrices. Academic Press, London.

[27] Lii, K. S., Rosenblatt, M., and Van Atta, C. (1976). Bispectral measurements in turbulance. J. Fluid. Mech., 77, 45-62.

[28] Morrison, D. F. (1976). Multivariate statistical methods, 2nd ed. McGraw-Hill, New York.

[29] Neudecker, H. (1969). Some theorems on matrix differentiation with special reference to Kronecker matrix products. J. Amer. Stat. Ass., 69, 953-963.

[30] Priestley, M. B. (1981). Spectral analysis and time series, Vol. 1 and 2. Academic Press, London.

[31] Quinn, B. G. (1982). Stationarity and invertibility of simple bilinear models. Stoch. Proc. Appl., 12, 225-230.

[32] Ruberti, A., Isidori, A., and d'Allesandro, P. (1972). Theory of bilinear dynamical systems. Springer-Verlag, Berlin.

[33] Shiryaev, A. N. (1966). Some problems in the spectral theory of higher order moments. J. Theor. Probability Appl., 5, 265-284.

[34] Subba Rao, T., (1978). On the theory of bilinear time series models. Technical Report, No. 87, Department of Mathematics, University of Manchester Institute of Science and Technology.

[35] Subba Rao, T. (1978). On the estimation of parameters of bilienar time series models. Technical Report, No. 79, Department of Mathematics, University of Manchester Institute of Science and Technology.

[36] Subba Rao, T., (1979). On the theory of bilinear time series models, II. Technical Report, No. 121. Department of Mathematics, University of Manchester Institute of Science and Technology.

[37] Subba Rao, T. (1981). On the theory of bilinear time series models. J. Royal Stat. Soc., 43, B, 244-255.

[38] Subba Rao, T., and Gabr, M. M. (1980). A test for linearity of stationary time series. J. Time Series Analysis, 1, 2, 145-158.

[39] Subba Rao, T., and Gabr, M. M. (1981). The properties of bilinear time series models and their usefulness in forecasting - with examples. Paper presented at the Golden Jubilee celebration of the Indian Statistical Institute, Calcutta on "Statistics: applications and new directions". Technical Report, No. 151, Department of Mathematics, University of Manchester Institute of Science and Technology.

[40] Tong, H. (1981). A note on Markov bilinear stochastic process in discrete time. J. Time Series Analysis, 2, 4, 279-284.

[41] Tuan Dinh Pham and Lanh Tat Tran (1981). On the first order bilinear time series model. J. Appl. Prob., 18, 617-627.

[42] Tukey, J. W. (1959). An introduction to the measurement of spectra. In Probability and Statistics (Ed. U. Grenander), Wiley, New York, 300-330.

[43] Waldmeier, M. (1961). The sunspot activity in the years 1610-1960. Schultheses, Zurich.

[44] Woodward, W. A., and Gray, H. L. (1978). New ARMA models for Wolfer's sunspot data. Commun. Statist., B7, 97-115.

[45] Woodward, W. A., and Gray, H. L. (1981). On the relationship between the S-array and the Box-Jenkins method of ARMA model identification. J. Amer. Stat. Ass., 76, 375, 579-587.

[46] Tuan Dinh Pham (1983). Bilinear Markovian Representation and Bilinear Models. Technical Report No. 160, Department of Mathematics, University of Manchester Institute of Science and Technology.