BILINEAR TIME SERIES MODELS AND ITS APPLICATIONS

Berlin Wu*

吳 柏 林*

摘 要

雙線性時間數列模式及其應用

線性時間數列模式，如ARMA模式已被廣泛地應用在許多學科領域。但是其中之一重要假設爲此時間數列之結構可被一線性模式來描述。此線性假設有時常覺得頗牽強，因而我們考慮是否有更好的模式來做資料之擬似。近年來，屬於非線性模式族之一系的一雙線性模式便引起學者的熱烈討論。本文即針對特定型之雙線性模式，探討其平穩性，可逆性。並做參數估計法則。最後舉例說明有關預測之方法。

Abstract

Linear time series models such as ARMA models have been widely used in many fields. An important assumption is that the structure of the series can be described by a linear model. However, this assumption of linearity is often a dubious one. In some particular situations one may ask if there exist other models which can provide a better fit. A particular class of non-linear models which has received a great deal of attentions is bilinear models. In this paper we investigates some properties of the bilinear model: stationarity and invertibility. Estimation of the parameters are obtained by minimum least squares method. The forecasting of certain bilinear models are also considered.

Keywords: *Time series analysis; ARMA models; Bilinear models; Markovian representation; Stationarity; Invertibility; Forecasting.*

1. Introduction

Linear time series models such as autoregressive moving average (ARMA) models have been widely and successfully used in many fields during the past two

* 作者爲本校應用數學研究所副教授

decades. An important assumption that is made in those models is that the structure of the series can be described by a linear model. However, this assumption of linearity is often a dubious one. In some particular situations one may ask if there exist other models which can provide a better fit. In view of this, several authors (Ozaki, 1978; Tong and Lim, 1980; Subba Rao, 1981; Priestley, 1988) have discussed certain more specific types of non-linear models.

A particular class of non-linear models which has received a great deal of attentions is bilinear models. The interesting feature of a bilinear system is that though it is non-linear, its structuraltheory is analogous to that of linear system. In its most general form a bilinear time series $\{X_t\}$ *with discrete time parameter is defined by*

$$X_t - \sum_{j=1}^{p} \phi_j X_{t-j} = \sum_{j=0}^{q} \theta_j \epsilon_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{k} b_{ij} X_{t-1} \epsilon_{t-j} \qquad (1.1)$$

where $\theta_0 = 1$ and $\{\epsilon_t\}$ is a strict white noise process, i.e. a sequence of independent zero mean and finite variance $\sigma^2_{\epsilon_t}$ random variables. It is apparent that if we set $b_{ij} = 0$ for all i,j, then (1.1) reduces to ARMA models, and thus the bilinear models includes as a special case the standard ARMA models.

In fact, Brockett (1976) has shown that, with suitable choice of the modelparameters, the bilinear model can approximate to an arbitrary degree of accuracy any 'well behaved' Votterra series relationship over a finite time interval. To some extent this parallels the corresponding property of ARMA models, namely that they can approximate to an arbitrary degree of accuracy any general linear relationship between $\{X_t\}$ and $\{\epsilon_t\}$. In this case the bilinear models may be regarded as a natural non-linear extension of the ARMA models.

A simple two-stage procedure to investigate whether or not a bilinear model might be appropriate is to fit an ARMA model and then to consider the autocorrelations of the squared residuals. Maravall (1982) applied this approach to some Spanish monetary data, found evidence of bilinearity, and achieved a modest 10% or so improvement in mean squared forecast error.

Specific results are available for the properties of a number of bilinear models, but the best results are for orders (p,q,r,l). Models of order (1,0,1,1) were considered by Granger and Andersen (1978), of order (p,0,p,1) by Subba Rao (1981, 1984), and of order (p,q,r,2) by Liu and Brockwell (1988).

2. Bilinear time series models

To study some features of bilinear time series, we have generated time series $\{X_t\}$ from the models

(a) $X_t = 0.3X_{t-1} + 0.6X_{t-1}\epsilon_{t-1} + \epsilon_t$.

(b) $X_t = 0.7X_{t-1} + 0.2X_{t-2} + 0.7\ X_{t-1}\epsilon_{t-1} + 0.8X_{t-2}\epsilon_{t-1} + \epsilon_t$.

(c) $X_t = 0.8X_{t-1} - 0.2X_{t-2} + 0.8\ X_{t-1}\epsilon_{t-1} + 0.8X_{t-2}\epsilon_{t-1} + \epsilon_t$.

(d) $X_t = -0.7X_{t-1} + 0.2X_{t-2} - 0.7\ X_{t-1}\epsilon_{t-1} + 0.8X_{t-2}\epsilon_{t-1} + \epsilon_t$.

The series (a), (b), (c) and (d) are ploted in Fig. 2.1, Fig. 2.2, Fig. 2.3 and Fig. 2.4 respectively. In each case the $\{\epsilon_t\}$ are independent $N(0,1)$ random variables, and each realization consists of 150 data points. An examination of the series (a), for which the coefficient of the bilinear term is mediate, has a more or less conventional form. Whereas series (b), with a further AR terms, exhibits a number of 'bursts' of large-amplitude excursions. Series (c) shows that at certain time period, there are high amplitude oscillations. In contrast to the series (a), (b) and (c), the behavior of the series (d) is very remarkable. This type of behavior is a well-known feature of certain types of seismological data, particularly in series relating to earthquakes and underground explosions.

We will denote the general bilinear models by BL(p,q,m,k), the integer p,q,m,k clearly denoting the orders of the various terms in (1.1). It is known that the representation $\{X_t\}$ given by (1.1) is not a Markovian representation. Tuan and Tran (1981) point out that, for the BL($p,0,p,1$) models, namely

$$X_t - \sum_{j=1}^{p} \phi_j X_{t-1} = \epsilon_t + \sum_{i=0}^{p} X_{t-1}\epsilon_{t-1} , \qquad (2.1)$$

a Markovian representation can be derived via the state vector $Z_t = (\Phi + B\epsilon_t)X_t$, where

$$\Phi_{pxp} = \begin{bmatrix} -\phi_1 & -\phi_2 & \cdots & -\phi_p \\ 1 & 0 & \cdots & 0 \\ & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \quad B_{pxp} = \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{pp} \\ 0 & 0 & \cdots & 0 \\ & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

Fig. 2.1 $X_t = 0.3X_{t-1} + 0.6X_{t-1}\epsilon_{t-1} + \epsilon_t$.

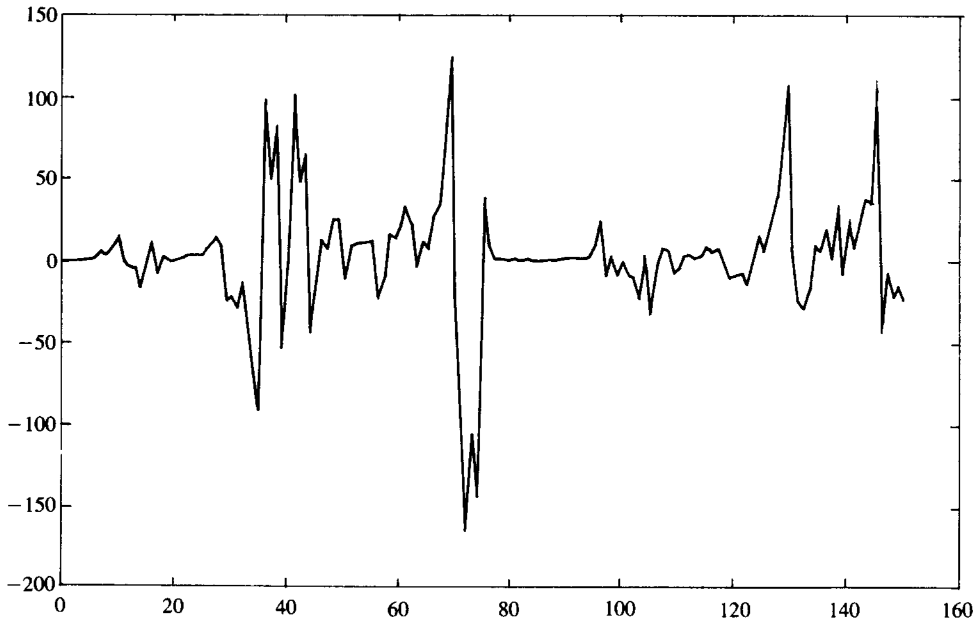


Fig. 2.2 $X_t = 0.7X_{t-1} + 0.2X_{t-2} + 0.7\ X_{t-1}\epsilon_{t-1} + 0.8X_{t-2}\epsilon_{t-1} + \epsilon_t$.

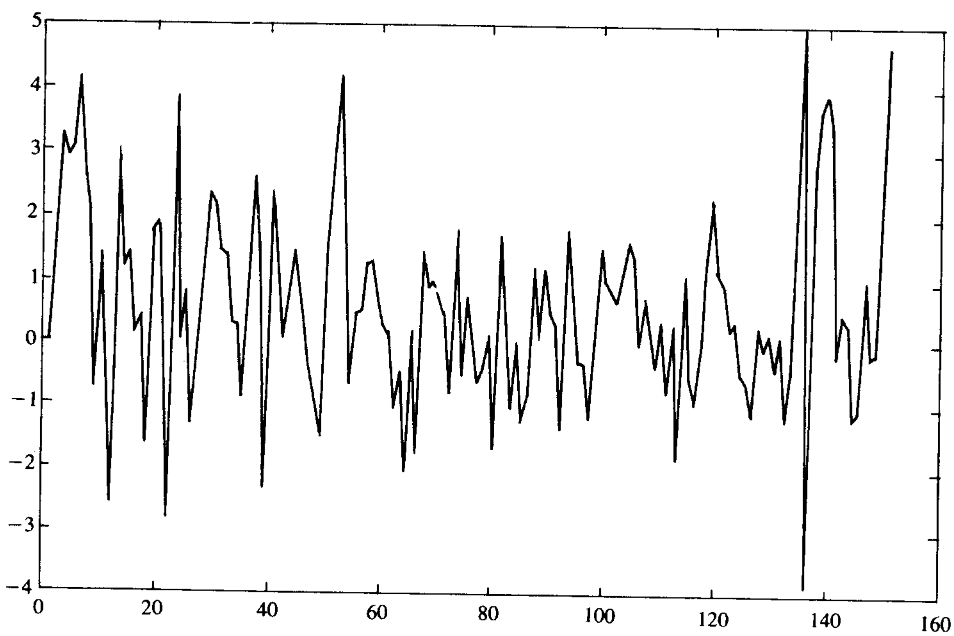


Fig. 2.3 $X_t = 0.8X_{t-1} - 0.2X_{t-2} + 0.8\ X_{t-1}\epsilon_{t-1} + 0.8X_{t-2}\epsilon_{t-1} + \epsilon_t$.

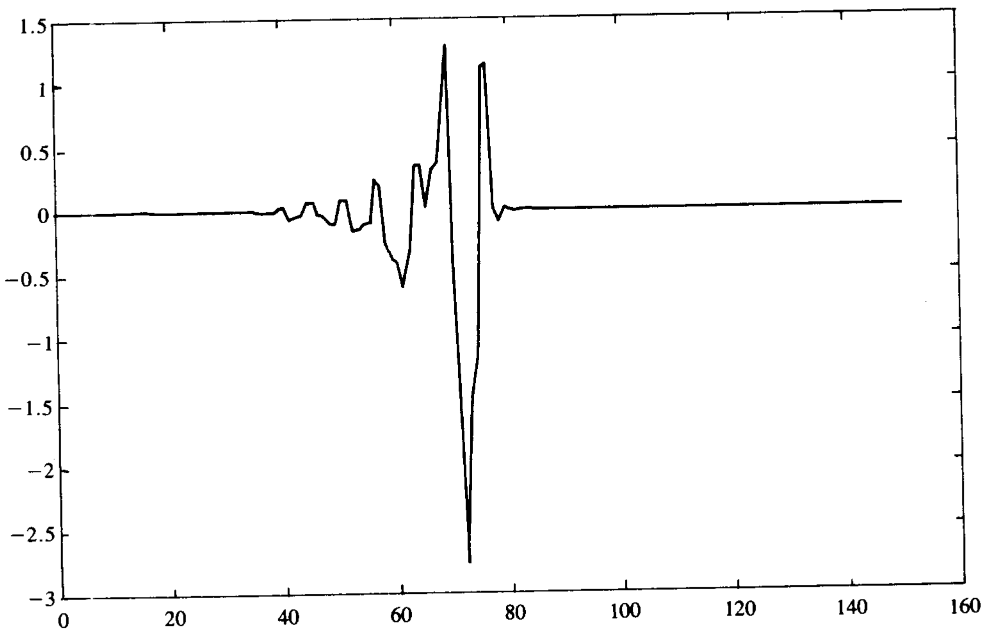
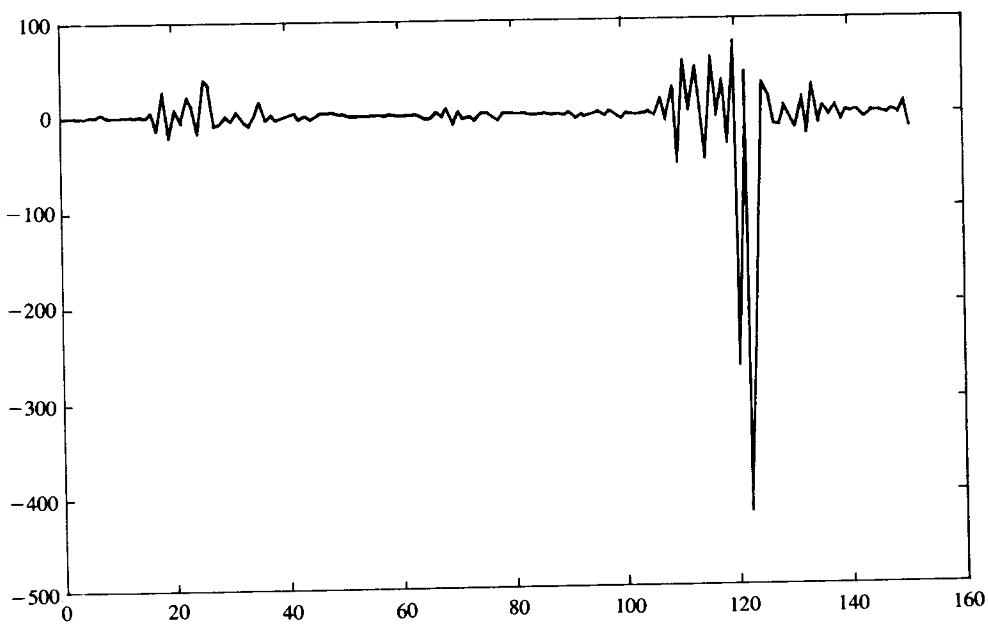


Fig. 2.4 $X_t = -0.7X_{t-1} + 0.2X_{t-2} - 0.7\ X_{t-1}\epsilon_{t-1} + 0.8X_{t-2}\epsilon_{t-1} + \epsilon_t$.



$C_{pxl} = (1,0,\ldots,0)'$ and $\mathbf{X}_t = (X_t, X_{t-1}, \ldots, X_{t-p+1})'$. Then, we can rewrite (2.1) as

$$\begin{cases} \mathbf{Z}_t = (\Phi + \mathbf{B}\epsilon_t)\mathbf{Z}_{t-1} + (\Phi + \mathbf{B}\epsilon_t)\mathbf{C}\epsilon_t \\ \mathbf{X}_t = \mathbf{Z}_{t-1} + \mathbf{C}\epsilon_t \end{cases} \tag{2.2}$$

Tuan (1985) has used this Markovian representation to obtain the moments of $\mathbf{Z}_t$ in the case of first-order models, i.e. when $\mathbf{Z}_t$ is a scalar process.

3. Asymptotic stationarity and covariance

In this section we obtain the conditions for asymptotic stationarity of the time series X_t satisfying the model (2.1); where we assume $\{\epsilon_t\}$ are i.i.d. $N(0,1)$ random variables.

Let $\mathbf{H}_{1xp} = (1,0,\ldots 0)$, from (2.1) we have ,

$$E(X_t) = \mathbf{H}\ E(\mathbf{X}_t)$$

$$\mathrm{Cov}(X_t,X_{t+s}) = \mathbf{H}\{E[(\mathbf{X}_t - E\mathbf{X}_t)\ (\mathbf{X}_{t+s} - E\mathbf{X}_{t+s})']\}\mathbf{H}' \ ,$$

and thus it suffices to consider the second-order properties of the vector process $\mathbf{X}_t$.

Write $\mathbf{u}_t = E(\mathbf{X}_t)$, $\mathbf{V}_t = E(\mathbf{X}_t\mathbf{X}_t')$, $\mathbf{S}_t = E(\mathbf{X}_t\mathbf{X}_t'\epsilon_t^2)$. Taking expectations of both sides of

$$\mathbf{X}_t = \Phi\mathbf{X}_{t-1} + \mathbf{B}\mathbf{X}_{t-1}\epsilon_{t-1} + \mathbf{C}\epsilon_t$$

and using $E(\mathbf{X}_t\epsilon_t) = \mathbf{C}$, we get

$$\begin{aligned} \mathbf{u}_{t+1} &= \Phi\mathbf{u}_t + \mathbf{BC} \\ &= \Phi^t\mathbf{u}_1 + \left[\sum_{i=1}^{t-1} \Phi^i\right]\mathbf{BC}. \end{aligned} \tag{3.1}$$

Iϕ $\mathbf{B} = 0$, and $\mathbf{u}_1 = 0$, then $\mathbf{u}_t = 0$, for all $t \geq 1$, and the process is then stationary to order 1 without any conditions on Φ. Otherwise, let $\rho(\Phi)$ be the spectral radius of a matrix Φ, i.e.

$$\rho(\Phi) = \max_i |\lambda_i\ (\Phi)|, \tag{3.2}$$

where $\{|\lambda_i \rho(\Phi)|\}$ are the i^{th} eigenvalue of Φ and it is known that $|\rho(\Phi)| \leq \|\Phi\|$ where $\|\cdot\|$ is any norm. A sufficient condition for $\lim_{t \to \infty} \left\{ \Phi^t \mathbf{u}_1 + \left[\sum_{i=1}^{t-1} \Phi^i \right] \mathbf{BC} \right\}$ to be finite is the $\rho(\Phi) \leq 1$. When this condition holds the process is asymptotically stationary to order 1, and the limiting mean value $\mathbf{u}$ is then given by

Proposition 3.1.
$$E(\mathbf{X}_t) \to \mathbf{u} = (\mathbf{I}-\Phi)^{-1}\mathbf{BC}. \tag{3.3}$$

As for the second-order moments, Subba Rao (1981) obtained a sufficient condition for stationarity, i.e.

Proposition 3.2. The bilinear the process $\mathbf{X}_t$ defined in (2.2) is asymptotically stationary if

$$\rho[\Phi \oplus \Phi + \mathbf{B} \oplus \mathbf{B}\varphi_\epsilon^2]\ )\ 1. \tag{3.4}$$

The condition (3.4) is a somewhat weaker form of the stationarity condition originally derived by Subba Rao (1981), namely

$$\|\Phi\|^2 + \|\mathbf{B}\|^2 < 1,$$

where $\|\Phi\|^2$ = is the largest eigenvalue of $\Phi\Phi'$, and $\|\mathbf{B}\|$ is similarly defined.

From the model (2.2), we can write

$$\mathbf{X}_t = \Phi\mathbf{X}_{t-1} + \mathbf{B}\mathbf{X}_{t-1}\epsilon_{t-1} + \mathbf{C}\epsilon_t. \tag{3.5}$$

It follows

$$E(\mathbf{X}_{t+1}\mathbf{X}_t') = \Phi' E(\mathbf{X}_t\mathbf{X}_t') + \mathbf{B}E(\mathbf{X}_t\mathbf{X}_t\epsilon_t) \tag{3.6}$$

and for $k > 1$,

$$E(\mathbf{X}_{t+k}\mathbf{X}_t') = \Phi^{k-1} E(\mathbf{X}_{t+1}\mathbf{X}_t') + \left[\sum_{i=1}^{k-2} \Phi^i \mathbf{BC} \right] \mathbf{u}'. \tag{3.7}$$

Let $C(k) = E[(\mathbf{X}_{t+k}-\mathbf{u})(\mathbf{X}_t-\mathbf{u})']$ be the autocorvariance matrix of lag k for $\mathbf{X}_t$. we have

Proposition 3.3.

$$\mathbf{C}(0) = \Phi\mathbf{C}(0)\Phi' + \mathbf{BC}(0)\mathbf{B}' + \xi, \tag{3.8}$$

$$\mathbf{C}(1) = \Phi\mathbf{C}(0) + \eta, \tag{3.9}$$

$$\mathbf{C}(k) = \Phi\mathbf{C}(k-1)\Phi' = \Phi^{k-1}\mathbf{C}(1), \tag{3.10}$$

where

$$\xi = \mathbf{Buu'B'} + \Phi\mathbf{uu'}\Phi' + \Phi\mathbf{SB'} + \mathbf{BS'}\Phi + \mathbf{2BCC'B'} + \mathbf{CC'} - \mathbf{uu'},$$

$$\eta = \Phi\mathbf{uu'} + \mathbf{BS} - \mathbf{uu'}.$$

Example 3.1. Consider the first order BL(1,0,1,1) model given by

$$X_t - \phi X_{t-1} = \epsilon_t + bX_{t-1}\epsilon_{t-1}.$$

Then, by (3.4), the sufficient condition for the asymptotic stationarity of the process $\mathbf{X}_t$ is $\phi^2 + b^2 < 1$, and from Proposition (3.1)

$$EX_t \rightarrow = \frac{b}{1-\phi}$$

$$E(X_t^2) \rightarrow \frac{1+2b^2}{1-\phi^2-b^2} + \frac{4\phi b^2}{(1+\phi)(1-\phi^2-b^2)}$$

$$E(X_{t+k}X_t) \rightarrow \phi E(X_t^2) + \frac{2b^2}{1-\phi}$$

4 Order selection and parameter estimation

A major problem with bilinear series modeling is the problem of model selection. This problem is far more difficult that the ARMA models. In the linear ARMA case, a preliminary identification on the orders of p and q can be done using sample ACF and PACF as in Box-Jenkings procedure. With the presence of bilinear terms in (1.1) the ordinary Box-Jenkings procedure of identification cannot be applied. Therefore, many authors resort to the use of the *Akaike Information Criterion* (AIC), see Akaike (1974), in selecting the right combination of p, q, m and l. The AIC is defined by

$$\text{AIC} = (N-r)\log \hat{\sigma}_\epsilon^2 + 2\times\text{independent number of parameters},$$

where $\hat{\sigma}_\epsilon^2 = \dfrac{1}{N-r} \sum_{t=r-1}^{N} \epsilon_t^2$ and

$N-r$ is the number of observations used for calculating the likelihood function. The model with smallest AIC value will be picked as the adequate model. Note that in using the AIC criterion we are trying to strike a balance between reducing the magnitude of the residual variance and increasing the number of model parameters.

To obtain a good set of estimates it is necessary that we should have a good set of initial values to start the iteration. The algorithm for choosing the order of the bilinear model BL($p,0,m,k$) is described as follows:

(1) Fit an AR(p) model to the data.

(2) Take the AR(p) coefficients obtained from (1) as initial estimates for the autoregregressive part of the BL($p,0,1,1$). Using the Newton-Raphson iteration method to fit this BL($p,0,1,1$) model.

(3) Fit the BL(p,0,1,2) and BL($p,0,2,1$) models using the coefficients obtained from (2) as initial values of the parameters. Of the two models fitted, choose the model which has the small residual variance and AIC values and use its parameters as starting values of fitting a BL($p,0,2,2$) model.

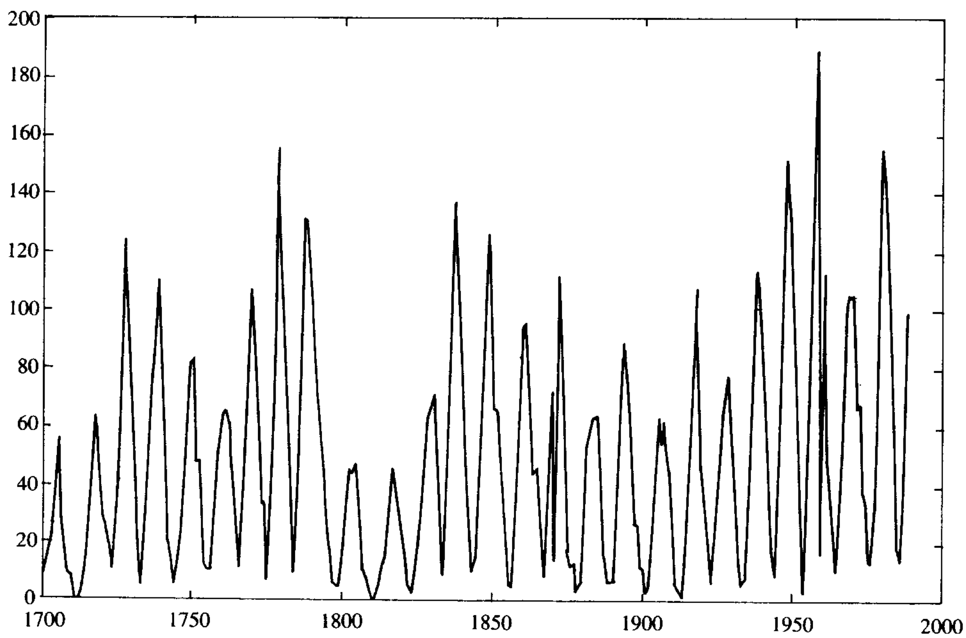
(4) The procedure is continued until the residual variance $\hat{\sigma}_\epsilon^2$ and AIC starts to increase as m, k increase. The final choice of model is then made by comparing the AIC values for each fitted model, and choose that model which had the minimum AIC value.

5. Analysis of the sunspot series

The Wolf sunspot series has attracted the attention of many time series analysts. Scientists believe that the sunspot numbers affect the weather of the earth and hence human activities such agriculture, telecommunications, war, and others. It is generally accepted that the earliest recorded date of a sunspot event was 10 May 28 *BC* during the reign of Emperor *Liu Ao* (*Cheng Di*) of the Western Han Dynasty in China. (see Needham (1959) p. 435). However, data on the Wolf annual sunspot index are available from 1700 onwards. In the year 1843 the sunspot cycle was apparently first noted by the German pharmaceutical chemist and amateur astronomer, Samuel Heinrich Schwabe (1789-1875), after 17 years of painstaking daily observations. For an interesting account of the history of the series see Izenmann (1985).

The data are shown in Fig. 5.1 and it can be seen that the main feature of this series is a cycle of activity varying in duration from about 9 to 14 years (about an average period of 11 years). Another feature of the series is its different gradients of *ascensions* and *descensions*. This suggests that a non-linear model might be appropriate. Subba Rao (1981) fitted various linear and bilinear models and compared their relative fit. The results of their analyses, for 246 observations with mean 43.53, are summarized below. The orders of the models were selected according to the AIC criterion.

Fig. 5.1. Annual sunspot numbers (1700 — 1988)



(I) *AR Models*. The best AR model is AR(9) with coefficients:

$$\hat{\phi}_1 = -1.21,\ \hat{\phi}_2 = 0.47,\ \hat{\phi}_3 = 0.12,\ \hat{\phi}_4 = -0.14,\ \hat{\phi}_5 = 0.16,$$

$$\hat{\phi}_6 = -0.09,\ \hat{\phi}_7 = 0.08,\ \hat{\phi}_8 = -0.09,\ \hat{\phi}_9 = -0.10;$$

$$\hat{\sigma}^2_\epsilon = 194.43,\ \text{AIC} = 1316.44.$$

(II) *Subset AR models*. The best subset AR model includes lags 1,2,9, with coefficients

$$\hat{\phi}_1 = -1.24,\ \hat{\phi}_2 = 0.54,\ \hat{\phi}_9 = -0.15;$$

$$\hat{\sigma}^2_\epsilon = 199.20,\ \text{AIC} = 1310.39.$$

(III) *ARMA models*. The best ARMA model is ARMA(6,6) with AR coefficients

$$\hat{\phi}_1 = 0.52,\ \hat{\phi}_2 = -0.47,\ \hat{\phi}_3 = -0.51,\ \hat{\phi}_4 = -1.09,$$

$$\hat{\phi}_5 = 0.07,\ \hat{\phi}_6 = -0.65,$$

$$\hat{\theta}_1 = 0.71,\ \hat{\theta}_2 = -0.07,\ \hat{\theta}_3 = -1.09,\ \hat{\theta}_4 = -0.08,$$

$$\hat{\theta}_5 = 0.04,\ \hat{\theta}_6 = -0.42,$$

$$\hat{\sigma}^2_\epsilon = 185.27,\ \text{AIC} = 1309.8.$$

(IV) *Bilinear models*. Applying the procedures described at the end of section 4, a BL($p,0,m,k$) of the form (2.1), with the constant c, was chosen with p = 3, m = 3, k = 4. The estimated parameters are:

$$c = 10.91,\ \hat{\phi}_1 = -1.93,\ \hat{\phi}_2 = -0.51,\ \hat{\phi}_3 = -1.09,$$

and the $\hat{b}_{ij}$ values (i = 1, 2, 3, j = 1, 2, 3, 4) are:

$$\hat{b}_{11} = -0.0055,\ \hat{b}_{12} = 0.0032,\ \hat{b}_{13} = -0.0018,\ \hat{b}_{14} = 0.0008,$$

$$\hat{b}_{11} = -0.0057,\ \hat{b}_{12} = 0.0056,\ \hat{b}_{13} = -0.0082,\ \hat{b}_{14} = 0.0058,$$

$$\hat{b}_{11} = -0.0017,\ \hat{b}_{12} = 0.0071,\ \hat{b}_{13} = -0.0110,\ \hat{b}_{14} = 0.0008;$$

$$\hat{\sigma}^2_\epsilon = 143.86,\ \text{AIC} = 1214.58.$$

We have to indicate that a different class of non-linear models, call *threshold models* was introduced by Tong and Lim (1980). They fited a threshold autoregressive model TAR(2;3,11) model to 221 observations on the Wolf sunspot series (from the year 1700 to 1920). The "best" model model is

$$X_t = \begin{cases} 12 + 1.7X_{t-1} - 1.3X_{t-2} + 0.4X_{t-3}^{(1)} + e_t, \text{ if } X_{t-3} \leq 36.6 \\ 7.8 + 0.7X_{t-1} - 0.04X_{t-2} + 0.2X_{t-3} + 0.2X_{t-4} - 0.2X_{t-5} \\ \quad - 0.02X_{t-6} + 0.2X_{t-7} - 0.2X_{t-8} + 0.3X_{t-9} + 0.4X_{t-10} \\ \quad + 0.4X_{t-11} + e_t^{(2)}, \qquad \text{if } X_{t-3} > 36.6 \end{cases}$$

where $\hat{\sigma}^2(e_t^{(1)}) = 254.64$, $\hat{\sigma}^2(e_t^{(2)}) = 66.8$, and *pooled* residual variance $= 153.7$.

6. Forecasting from bilinear models

Given observations on a series up to time t, it is known that minimum mean-square error (MSE) predictor of a future value X_{t+s}, s step ahead of X_t is given by the conditional expectation $\hat{X}_t(s) = E[X_{t+s} | B_t]$, B_t being the σ-algebra generated by X_r, $r \leq t$. For linear models the conditional expectations can be evaluated by a set of recursive equations. While for bilinear models we need to find out those terms $\epsilon_t X_s$. Tong (1990) present the following facts concerning $\epsilon_t X_s$:

(1) For $t > s$:

$$E[\epsilon_t X_s | B_0] = 0, \text{ if } t > 0, \text{ because } \epsilon_t \text{ and } X_s \text{ are independent.}$$

$$E[\epsilon_t X_s | B_0] = \epsilon_t x_s, \text{ if } t \geq 0, \text{ by invertibility.}$$

(2) For $t = s$:

$$E[\epsilon_t X_s | B_0] = \sigma^2_{\epsilon_t}, \text{ if } t > 0, \text{ because } X_t = \epsilon_t + f(\epsilon_{t-1}, \epsilon_{t-2}, \ldots)$$

by stationarity.

$$E[\epsilon_t X_s | B_0] = \epsilon_t x_s, \text{ if } t \leq 0, \text{ by invertibility.}$$

(3) For $t < s$:

We use the bilinear model to express X_s in terms of X_{s-1}, X_{s-2}, ... *and* ϵ_t, ϵ_{t-1}, ... We may repeat this process until the the subscripts of the X terms are < t and then appeal to (1) and (2), the independence of the ϵ term, and marginal distribution of ϵ_0.

Example 6.1. Consider the bilinear model

$$X_t = \phi X_{t-1} + bX_{t-1}\epsilon_{t-1} + \epsilon_t.$$

where $\epsilon_t \overset{iid}{\sim} N(0,\sigma^2)$ and the model is assumed stationary and invertible. Thus, the one step ahead forecast is

$$E[X_1 \mid B_0] = \phi X_0 + bX_0\epsilon_0.$$

And the 2^{nd}, 3^{rd} step ahead forecasts are

$$E[X_2 \mid B_0] = \phi E[X_1 \mid B_0] + bE[X_1\epsilon_1 \mid B_0]$$

$$= \phi^2 X_0 + \phi bX_0\epsilon_0 + b\sigma^2, \text{ using (2) above.}$$

$$E[X_3 \mid B_0] = \phi E[X_2 \mid B_0] + bE[X_2\epsilon_2 \mid B_0]$$

$$= \phi^3 X_0 + \phi^2 bX_0\epsilon_0 + \phi b\sigma^2 + b\sigma^2.$$

Higher-step predictions can be obtained similarly.

REFERENCE

[1] Akaike, H (1974). Anew look at the statistical model identification, *IEEEE Transaction on Automatic control,* AC-19, 716-723.

[2] Brockett, R. W. (1976). Volterra series and geometric control theory. *Automatica,* **12,** 167-176.

[3] Granger, C.W.J. and Andersen, A (1978a). Non-linear time series modeling. *Applied Time series Analysis,* 25-38, (Findley. D. F. ed.) Academic Press, New York.

[4] Granger, C.W.J. and Andersen, A (1978b). An introduction to bilinear time series models. *Vanderhoeck and Reprecht,* Gottingen.

[5] Izenman, A. J. (1985). J. R. Wolf and the Zurich sunspot relative numbers, *The Mathematical Intelligencer,* **7**, No. 1, 27-33.
[6] Liu, J. and Brockwell, P.J. (1988). On the general bilinear time series models. *J. Appl. Prob.,* 25, 553-64.
[7] Maravall, A. (1982). An application of nonlinear time series forecasting, *J. Bus. Econ. Stat.* **1**, 66-74.
[8] Needham, J. (1959). *Science and civilisation in China,* Vol. III. Cambridge University Press.
[9] Ozaki, T. (1978). Non-linear models for non-linear random vibrations. Technical Report no. **92**, Department of Mathematics, University of Manchester Institute of Science and Technology, UK.
[10] Tuan, P. D. and Tran, L. T. (1981). On the first order bilinear time series model. *J. of Appl. Prob.,* 18, 617-627.
[11] Tuan, P. D. (1985). Bilinear Markovian representation and bilinear models. *Stochastic Processes Appl.,* **20**, 295-306.
[12] Priestley, M.B. (1988). *Non-linear and non-stationary time series analysis.* Academic Press, London.
[13] Subba Rao, T. (1981). On the theory of bilinear time series models. *J. Roy. Statist. Soc.* B 43(2), 244-255.
[14] Subba Rao, T. and Gabr, M. M. (1984) *An introduction to Bispectral Analysis and Bilinear Time Series Models.* Lecture Notes in Statistics, Springer-Verlag, London.
[15] Tong, H. and Lim, K.S. (1980). Threshold autoregression, limit cycles and cyclical data. J. Roy. Statist. Soc. Ser. B. **42**, 245-292.
[16] Tong, H. (1990). *Non-linear Time Series.* Oxford University Press.