Detecting Non-linearity Using Squares of Time Series Data

C.O. Omekara
Department of Mathematical and Physical Sciences,
Michael Okpara University of Agriculture,
Umudike, P.M.B. 7267, Umuahia, Abia State, Nigeria

Abstract: The aim of this study is to discuss the properties of squares of a pure diagonal bilinear (PDBL) time series model and how these properties can be used to distinguish between a linear (ARMA) model and a non-linear (bilinear) model. We showed that for a pure diagonal bilinear process, the square of the series have the same covariance structure as an ARMA process. Simulated data was used to illustrate the results obtained in this study.

Key words: ARMA model, bilinear time series, detecting, non-linearity

INTRODUCTION

According to Granger and Andersen (1978) bilinear models are formed by adding a bilinear form to the autoregressive/moving average (ARMA) models leading to

\[ X_t + \sum_{j=1}^{p} a_j X_{t-j} = \varepsilon_t + \sum_{j=1}^{q} C_j \varepsilon_{t-j} + \sum_{j=1}^{n} b_j X_{t-j} \varepsilon_{t-j} \tag{1} \]

where \( \{ \varepsilon_t \} \) is a sequence of i.i.d random variables with zero mean and finite variance \( \sigma^2 \) and \( \varepsilon_t \) is independent of \( X_s \) s < t. The formal difference between a bilinear time series model and an ARMA model is the bilinear term \( eX \).

Bilinear models were first studied in the context of non-linear control systems, but their application as time series model were investigated principally by Granger and Andersen (1978) and Subba Rao (1981). Following Subba Rao (1981) we represent (1) as BL (p, q, m, k) where BL is abbreviation for bilinear. Subba Rao et al. (1984) also gives a comprehensive account of this class of models. Sessay and Subba Rao (1988, 1991), Akamanan et al. (1986), Gabr (1988), Subba Rao and Silva (1993) and many other authors have examined various simple forms of (1) in the context of stationarity, invertibility and estimation.

The motivation for using data values to detect non-linearity is provided by a result inherent in the work of Granger and Newbold (1976). They showed that for a series \( \{X_t\} \) which is normal (and therefore linear)

\[ \rho_k (X^t) = \{ \rho_k (X_j) \}^t \]

where \( \rho_k (\cdot) \) denotes the lag k autocorrelation. Any departures from this result presumably would indicate a degree of non-linearity, a fact pointed out by Granger and Andersen (1978).

Granger and Andersen (1978), have also shown that for single term bilinear time series \( \{X_t\} \) satisfying

\[ X_t = bX_{t-1} \varepsilon_{t-1} + \varepsilon_t \]
\( \{X_i^2\} \) has the same covariance structure as an ARMA \((1,k)\) process.

We show below that for the pure diagonal bilinear process \( \{X_i^2\} \)
would have the same covariance structure as an ARMA \((p,p)\) process.

**Properties of Squares of PDBL Model**

Now consider the pure diagonal bilinear model satisfying

\[
X_t = \sum_{j=1}^{p} b_j X_{t-j} + \epsilon_t
\]

where \( \{\epsilon_t\} \) is a sequence of i.i.d random variables with zero mean and constant variance \( \sigma^2 \).

Let \( W_t = X_t^2 \)

\[
X_t^2 = \sum_{j=1}^{p} b_j X_{t-j}^2 \mu^2 + 2 \sum_{j=1}^{p} b_j \sum_{k=1}^{j} b_{j-k} X_k \mu^2 + \sum_{j=1}^{p} b_j X_j \mu^2 + \epsilon_t^2
\]

(3)

We are going to consider three cases namely: \( k < p \), \( k = p \) and \( k > p \) where \( k \) is the lag of the autocovariance coefficient.

**Case 1: \( k < p \)**

It can be shown that

\[
E(X_t^2 X_{t+1}^2) = b_k^2 E(X_t^4) + \sum_{j=1}^{p-1} b_j^2 E(X_{t-j}^2 X_{t-j}^2) + \sum_{j=1}^{p-1} b_{j-1}^2 E(X_{t-j+1}^2 X_{t-j+1}^2)
\]

\[
+ 2 \sum_{i=1}^{p-k} b_i b_{k+i-1} E(X_{t+i}^2 X_{t+i}^2) + 2 \sum_{i=1}^{p-k} b_i b_{k+i-1} E(X_{t+i}^2 X_{t+i}^2) + \sigma^2 E(X_t^2)
\]

(4)

But, the autocovariance function of a stationary process \( \{X_t\} \) is given by

\[
R(k) = E(X_{t-k}^2 X_{t-k-1}^2) = E(X_t^2 X_{t-1}^2) - \mu^2.\text{ Therefore,}
\]

\[
R_{\epsilon}(k) = E(W_t W_{t+k}) - \mu_{\epsilon}^2
\]

where \( R_{\epsilon}(k) \) is the autocovariance function of \( W_t = X_t^2 \) at lag \( k \) and \( \mu_{\epsilon} = E(X_i^2) \). Therefore,

\[
R_{\epsilon}(k) = E(X_t^4) - \mu_{\epsilon}^2
\]

(5)

**Case 2: \( k = p \)**

It can easily be shown that:

\[
E(X_t^4) = \sigma^4 E(X_t^4) + 12 \sigma^4 E(X_t^4)
\]

and
\[ \text{E}(X_i^2 e_t) = 3\sigma^2 \text{E}(X_i^2) \]

Therefore,

\[ \text{E}(X_i^2; X_{i+p}) = b_j^2 \sigma^4 \text{E}(X_i^2) + 12b_j^4 \sigma^4 \text{E}(X_i^2) + \sum_{j=1}^{p-1} b_j^2 \sigma^4 \text{E}(X_i^2; X_{i+p-j}) + 2 \sum_{j=1}^{p-1} b_j^2 \sigma^4 \text{E}(X_i^2) + 6 \sum_{j=1}^{p-1} b_j b_j^4 \]

\[ \text{E}(X_i^2) + 2 \sum_{j=1}^{p-1} \sum_{j=1}^{p-1} b_j b_j^4 \sigma^4 \text{E}(X_i^2) + \sigma^4 \text{E}(X_i^2) \] (6)

Substituting for \( \text{E}(X_i^2; X_{i+p}) \) in (4) and simplifying we obtain

\[ R_w(k) = \sum_{j=1}^{p-1} b_j^2 \sigma^4 R_w(p-j) + 26 \left( 2 \sum_{j=1}^{p-1} b_j b_j^4 + 5b_j^4 \right) U_w \] (7)

Observe that this is a Yule-Walker type difference equation.

**Case 3: k>p**

In this case, it can easily be shown that

\[ \text{E}(X_i^2; X_{i+p}) = \sum_{j=1}^{p} b_j^2 \sigma^4 \text{E}(X_i^2; X_{i+p-j}) + 2 \sigma^4 \text{E}(X_i^2) \sum_{j=1}^{p} b_j^2 + 2 \sum_{j=1}^{p} \sum_{j=1}^{p} b_j b_j^4 \text{E}(X_i^2) + \sigma^4 \text{E}(X_i^2) \] (8)

Substituting for \( \text{E}(X_i^2; X_{i+p}) \) in (4) and simplifying, we obtain

\[ R_w(k) = \sum_{j=1}^{p} b_j^2 \sigma^4 R_w(k-j), k \geq p + 1 \] (9)

This is the Yule-Walker equation for an ARMA (P,P) model. Present study on squares of \( X_i \) satisfying (1) leads to the following theorem needed for identification purposes.

**Theorem 1**

Let \( \{e_t\} \) be a sequence of independent and identically distributed random variables with \( \text{E}(e_t) = 0 \)

\[ \text{E}(e_t^2) = \sigma^2 < \infty. \]

Suppose there exists a stationary and invertible process \( \{X_i\} \) satisfying

\[ X_i = \sum_{j=1}^{p} b_j X_{i-j} + e_i, \]

for some constants \( b_1, b_2, \ldots, b_p \), \( p > 0 \). Then \( X_i^2 \) will be an ARMA (P,P) model.

**Comparison with a linear model**

Here it is shown that if \( \{X_i\} \) is MA (P), then \( \{X_i^2\} \) is also MA (P). We proceed as follows:-

For the MA (P) model
Then
\[ X_i = \sum_{j=1}^{p} b_j e_{i,j} + c_i \]

It is easy to show that the following are true:
\[ E(\chi_i^2 e_i^j) = \sigma^2 \left( 3 + \sum_{j=1}^{p} b_j^2 \right) \]
\[ E(\chi_i^2) = \sigma^2 \left( 1 + \sum_{j=1}^{p} b_j^2 \right) \]
and
\[ \chi_i^2 \chi_{i,k} = \sum_{j=1}^{p} b_j^2 \chi_i^2 e_{i,k}^j + 2 \sum_{i=1}^{p} b_j b_k \chi_i^2 \chi_{i,k}^j e_{i,k} + 2 \sum_{i=1}^{p} b_j \chi_i^2 e_{i,k}^j e_{i,k}^k + X_i^2 e_{i,k}^j e_{i,k}^k \]

We proceed to treat the autocovariance as follows

**CASE 1: k < p**

For \( k < p \), it can be shown that
\[ \chi_i^2 \chi_{i,k} = b_i^2 \chi_i^2 + \sum_{j=1}^{p} b_j^2 \chi_i^2 + 2 \sum_{k=1}^{p} b_k^2 \chi_i^2 e_{i,k} + 2 \sum_{k=1}^{p} b_k \chi_i^2 e_{i,k} + X_i^2 e_{i,k} \]

Taking expectation, we have
\[ E(\chi_i^2 \chi_{i,k}) = b_i^2 E(\chi_i^2) + \sigma^2 E(\chi_i^2) \left( 1 + \sum_{j=1}^{p} b_j^2 \right) \]

Substituting in Eq. 5, we have that
\[ R_{i,k}^2 = b_i^2 E(\chi_i^2) + \sigma^2 E(\chi_i^2) \left( 1 + \sum_{j=1}^{p} b_j^2 \right) - \left[ \sigma^2 (1 + \sum_{j=1}^{p} b_j^2) \right]^2 \]

**CASE 2: k > p**

We recall that
\[ (X_i^2 X_{1+k}) = \sum_{j=1}^{p-1} b_j X_i^2 e_{i+k-j}^2 + 2 \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} b_j b_{i+j} X_i^2 e_{i+k-j} e_{i+k-j} + 2 \sum_{j=1}^{p} b_j X_i^2 e_{i+k-j} e_{i+k-j} + X_i^2 e_{i+k}^2 \]

and

\[
E(X_i^2 X_{1+k}) = \sum_{i=1}^{p} b_i^2 E(X_i^2) + 2 \sum_{i=1}^{p-1} \sum_{j=1}^{p-1} b_j b_{i+j} E(X_i^2 e_{i+k-j} e_{i+k-j}) + 2 \sum_{j=1}^{p} b_j E(X_i^2 e_{i+k-j} e_{i+k-j}) + \sigma^2 E(X_i^2)
\]

\[
= \sigma^2 \sum_{j=1}^{p} b_i^2 E(X_i^2) + \sigma^2 E(X_i^2) = \sigma^2 E(X_i^2) \left( 1 + \sum_{j=1}^{p} b_i^2 \right)
\]

Thus

\[
R_w(k) = \sigma^2 E(X_i^2) \left( 1 + \sum_{j=1}^{p} b_i^2 \right) - \left( \sigma_i \left( 1 + \sum_{j=1}^{p} b_j^2 \right) \right)^2 = 0
\]

Hence \( X_i^2 \) is also an MA (P).

**SIMULATION RESULTS**

Here we present some simulation to illustrate the results obtained in this study. In what follows, the random variable \( \{e_i\} \) are mutually independent and identically distributed as \( N(0, \sigma^2) \). The processes considered are:

\[
X_i = 0.7X_{i-1} + e_i \quad (10)
\]

\[
Y_i = 0.7 + e_i + 0.146e_{i-1} \quad (11)
\]

The simulation and estimation were done using MINITAB. For purposes of illustration, we have without loss of generality taken \( \sigma^2 = 1 \) for (10) and (11). We generated for each process 200 observations \( X_{i_1}, X_{i_2}, \ldots, X_{i_{200}} \). The autocorrelation for \( X_{i_1}, Y_i \) and were estimated.

The estimator

\[
r_k = R(k) / R(0), \quad k = 1, 2, 3, \ldots
\]

was used to estimate the autocorrelation, where

\[
R(k) = \frac{1}{N-K} \sum_{i=1}^{K} (X_i - \overline{X})(X_{i+k} - \overline{X}), \quad K = 0, 1, 2, \ldots
\]

is the estimate of the autocovariance \( R(k) \) and

\[
\overline{X} = \frac{1}{N} \sum_{i=1}^{N} X_i
\]
Table 1: Showing Estimated Autocorrelation for $X_k$, $X_k^2$, $y_i$ and $y_i^2$

<table>
<thead>
<tr>
<th>Lag $k$</th>
<th>$X_k$</th>
<th>$X_k^2$</th>
<th>$y_i$</th>
<th>$y_i^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.19</td>
<td>0.50</td>
<td>0.51</td>
<td>0.37</td>
</tr>
<tr>
<td>2</td>
<td>0.17</td>
<td>0.28</td>
<td>0.11</td>
<td>0.07</td>
</tr>
<tr>
<td>3</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.13</td>
<td>-0.00</td>
</tr>
<tr>
<td>4</td>
<td>0.01</td>
<td>-0.04</td>
<td>0.03</td>
<td>-0.07</td>
</tr>
<tr>
<td>5</td>
<td>0.03</td>
<td>-0.05</td>
<td>0.05</td>
<td>-0.04</td>
</tr>
<tr>
<td>6</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.21</td>
<td>0.13</td>
</tr>
<tr>
<td>7</td>
<td>0.15</td>
<td>-0.02</td>
<td>0.19</td>
<td>0.02</td>
</tr>
<tr>
<td>8</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.08</td>
<td>-0.01</td>
</tr>
<tr>
<td>9</td>
<td>0.02</td>
<td>-0.06</td>
<td>0.05</td>
<td>-0.04</td>
</tr>
<tr>
<td>10</td>
<td>0.03</td>
<td>-0.06</td>
<td>0.03</td>
<td>-0.08</td>
</tr>
</tbody>
</table>

is the estimate of the mean. As these estimators have been discussed in detail by Chatfield (1980) they have just been stated here. The parameters have been carefully chosen to ensure the invertibility and stationarity of the processes. Table 1 gives the estimated autocorrelation of the models (10) and (11) and their squares.

$X_k$ is seen to identify as an MA (1) under covariance analysis and at least as ARIMA (1,1) as the theory predicted. Both $y_i$ and $y_i^2$ would identify as no more than MA (1). Therefore, looking at the square of a series is a useful way of distinguishing between a linear and a bilinear model having the same covariance analysis properties.

**DISCUSSION AND CONCLUSION**

One way of distinguishing between linear and non-linear models is to perform a second-order analysis on the squares of the series. Some authors have shown that for a series $X_k$ which is normal (and therefore linear)

$$\sigma_k(X_i^2) = [\sigma_k(X_i)]^2$$  \hspace{1cm} (12)

where $\sigma_k(\cdot)$ denotes the lag k autocorrelation. Any departures from this result presumably would indicate a degree of non-linearity, a fact pointed out by Granger and Andersen (1978).

We have, however, shown in this paper that this result (12) does not hold for the pure diagonal bilinear model. We have shown that the covariance structure of the square of a moving sequence time series is the same as the covariance structure of the original series. And this result can be used to distinguish between a pure diagonal and a linear model.

**REFERENCES**


