E7.22

PARAMETER ESTIMATION FOR A BILINEAR TIME SERIES MODEL*

V. John Mathews and Todd K. Moon

Department of Electrical Engineering
University of Utah
Salt Lake City, Utah 84112

ABSTRACT

This paper presents a direct approach to the estimation of the parameters associated with a bilinear time series model. The approach depends critically on the expressions for certain higher-order statistics of the signals that satisfy the bilinear model. These expressions are linear in most of the parameters of the model. The parameters are then estimated from an overdetermined set of equations. Results of an experiment that employs our technique and demonstrates its good properties are also included in the paper.

I. INTRODUCTION

A large class of parametric spectrum estimation techniques is based on the assumption that the underlying model is linear. While linear models provide adequate performance in many situations, there are several instances where they can give misleading results. Consequently, there has been a lot of activity in the recent past on fitting signals with nonlinear models.

One particular class of nonlinear models that has attracted the attention of several researchers is the bilinear model [2, 4-8]. In the most general form, a discrete-time, random signal X(n) that can be fitted with a bilinear model satisfies the following difference equation

$$X(n) = \sum_{i=1}^{p} a_i X(n-i) + \sum_{i=1}^{s}\sum_{j=1}^{q} b_{i,j} X(n-i)e(n-j) + \sum_{j=1}^{r} c_j e(n-j) \qquad (1)$$

where e(n) belongs to a stationary and zero-mean white process. In statistical literature, the above model is often denoted as the BL(p, r, s, q) model [8]. This model is particularly attractive since it shares several features with linear systems. In particular, if $b_{i,j} = 0$ for all i and j, we get the familiar ARMA model.

In spite of the simplicity of the bilinear system model, there is a large class of nonlinear systems that can be adequately modeled as bilinear systems. For example, it has been shown under fairly mild conditions that a bilinear system model with finite number of coefficients can be used to approximate any Volterra system with arbitrary precision [1]. Consequently, bilinear system models have found a variety of applications including those in control systems, population models, biological systems, economic models, etc. An overview of continuous-time bilinear models and their applications can be found in [2, 6].

In this paper, we present an algorithm for estimating the parameters associated with a bilinear time series represented by a BL(p, 0, p, 1) model

$$X(n) = \sum_{i=1}^{p} a_i X(n-i) + \sum_{i=1}^{p} b_i X(n-i)e(n-1) + e(n) . \qquad (2)$$

Note that the model reduces to the autoregressive (AR) system model if all b_i's are zero, and therefore we can consider the above model as a nonlinear extension to the linear autoregressive models. Note also that there is no loss of generality in setting the ranges of the indices of a_i and b_i to be the same.

The rest of the paper is organized as follows: Section II first develops a set of recursive equations for the autocovariance function and a set of third order cumulants of a bilinear time series. This section then describes how an overdetermined set of these equations are used to find least-squares estimates of the parameters of the bilinear model. Results of a simulation experiment are presented in Section III. The concluding remarks are made in Section IV.

II. PROBLEM STATEMENT, PAST WORK, AND THE NEW PARAMETER ESTIMATION TECHNIQUE

Given a stationary signal that is generated according to (2) when e(n) is zero-mean and Gaussian, the problem is to estimate the parameters a_i, i = 1, 2, ..., p; b_i, i = 1, 2, ..., p; and σ_e^2, the variance of the driving noise e(n) from a single time-limited realization of the process.

Several researchers have considered this problem in the past. Subba Rao [8] studied the general problem of parameter estimation of a BL(p, r, s, q) model by formulating it as a nonlinear optimization problem and estimated the parameters using the Newton-Raphson search procedure. Most of the approaches for direct estimation of the parameters from the measured statistics of the signal deal only with very low orders and as few as one to three parameters [4, 5, 7]. The difficulty arises because the task of developing expressions for the different moments of the bilinear time series becomes very cumbersome as the model order and the number of parameters increase.

Our approach to the estimation of the parameters of the bilinear time series model depends critically on the expressions for the autocovariance function and third-order correlation function of the time series. Let μ denote the mean value of X(n) and let

$$C(n) = X(n) - \mu . \qquad (3)$$

The following results are derived in the appendix.

* This work was supported in part by NSF Grant No. MIP 8922146 and a University of Utah Faculty Fellow Award.

CH2977-7/91/0000-3513 $1.00 © 1991 IEEE

$$\mu = \left(1 - \sum_{i=1}^{p} a_i\right)^{-1} b_1 \sigma_e^2 \tag{4}$$

$$r_{cc}(m) = E\{C(n)C(n-m)\}$$

$$= \begin{cases} \sum_{i=1}^{p} a_i r_{cc}(m-i) & ; m > 1 \\ \sum_{i=1}^{p} a_i r_{cc}(m-i) + \sigma_e^2 \sum_{i=1}^{p} \sum_{j=1}^{p} b_i b_j r_{cc}(i-j) & \\ \quad + \mu \left(\sum_{i=1}^{p} b_i\right) \sigma_e^2 \left\{a_1 + \mu \sum_{i=1}^{p} b_i\right\} + b_1^2 \sigma_e^4 + \sigma_e^2 & ; m = 0 \\ \sum_{i=1}^{p} a_i r_{cc}(m-i) + \mu \left(\sum_{i=1}^{p} b_i\right) \sigma_e^2 & ; m = 1 \end{cases} \tag{5}$$

$$\Gamma_{ccc}(1,m) = E\{C(n)C(n-1)C(n-m)\}$$

$$= \sum_{i=1}^{p} a_i \Gamma_{ccc}(i-1, m-1) + \sigma_e^2 \sum_{i=1}^{p} b_i r_{cc}(m-i) \quad ; m \geq 2. \tag{6}$$

In the above expressions $E\{\bullet\}$ denotes the statistical expectation of $\{\bullet\}$. Several observations can be made now. Note that the covariance function of $C(n)$ for lags larger than 1 is exactly the same as the covariance function of an autoregressive moving average (ARMA) process with p AR parameters $a_1, a_2, ..., a_p$ and one MA parameter. This implies that a bilinear time series cannot be distinguished from an ARMA time series from the knowledge of its covariance function alone. Therefore, one must resort to measurement of higher-order statistics of the signal to identify the parameters that describe the nonlinear nature of the time series. While this result is fairly well known, we believe that the expressions for the third-order cumulants derived here are new.

The second observation that we can make from the above expressions is that the parameters $a_1, a_2, ..., a_p$ and $b_1\sigma_e^2, b_2\sigma_e^2, ..., b_p\sigma_e^2$ appear in a linear fashion in the expressions. Consequently, we can solve for these parameters very easily. Furthermore, given the above parameters the expression for $r_{cc}(0)$ becomes a quadratic function of σ_e^2 and therefore solving for σ_e^2 is also straightforward.

In our approach, we estimate $r_{cc}(m)$ for $m = 0, 1, 2, ..., L_1$, $\Gamma_{ccc}(1, m)$ for $m = 2, 3, ..., L_2$, and $\Gamma_{ccc}(i, m)$ for $i = 0, 1, ..., p - 1$ and $m = 1, 2, ..., L_2 - 1$ from the data. L_1 and L_2 are usually chosen to be much larger than the minimum numbers necessary to solve for the unknown parameters. We can then form an overdetermined set of linear equations in the unknown parameters $a_1, a_2, ..., a_p$ and $\sigma_e^2 b_1, \sigma_e^2 b_2, ..., \sigma_e^2 b_p$ and obtain a least-squares fit for these parameters. We used a singular value decomposition approach described in [3] for estimating the parameters. σ_e^2 is calculated by substituting the estimated values of a_i's and $\sigma_e^2 b_i$'s in the expression for $r_{cc}(0)$, solving the resultant quadratic equations and retaining the positive root.

III. EXPERIMENTAL RESULTS

The time series considered was generated using (2) and parameters as shown in Table 1. The first 50,000 samples of each realization were discarded to ensure that the time series exhibits the steady-state properties and the estimates were made using 2000 samples. Table 1 tabulates the mean and variance of the estimates over 100 independent runs. The estimation procedure employed 10 equations for the autocovariance function and 10 equations for the third-order cumulants.

We can see from the results that the algorithm estimates the unknown parameters fairly accurately. The estimation noise seems to be larger for the b_i coefficients which correspond to the nonlinear terms in the model.

TABLE 1. MEAN AND VARIANCES OF THE PARAMETER ESTIMATES

Parameter	True Value	Mean	Variance
a_1	0.2	0.202	1.87×10^{-3}
a_2	0.1	0.098	4.70×10^{-4}
a_3	0.5	0.498	5.46×10^{-4}
b_1	0.1	0.119	6.54×10^{-2}
b_2	0.2	0.190	6.97×10^{-2}
b_3	0.3	0.294	1.57×10^{-2}
σ_e^2	0.1	0.097	6.12×10^{-5}

IV. CONCLUDING REMARKS

This paper presented a new algorithm for estimating the parameters of a bilinear time series. The method was based on a set of expressions that characterized the behavior of the covariance function and third-order cumulants of the time series. These expressions were linear in most of the parameters and consequently the development of the technique was relatively straightforward once the expressions were derived. The properties of the algorithm are being analytically studied at present. Model-order selection was not addressed in the paper. Our experiments have indicated that the singular values of the correlation and cumulant matrices do give a relatively accurate indication of the number of parameters involved.

APPENDIX

Taking the statistical expectation of both sides of (2) and recognizing that the signal is stationary gives

$$\mu = \left(\sum_{i=1}^{p} a_i\right)\mu + E\{b_1 X(n-1)e(n-1)\} \tag{A1}$$

since $E\{X(n-i)e(n)\} = 0$ for $i > 0$. Expanding $X(n)$ using (2) and taking the expectation of each term separately, we can show that

$$E\{X(n)e(n)\} = \sigma_e^2 \ . \tag{A2}$$

Substituting this in (A1) and solving for μ results in (3).

It is relatively straightforward to see that $C(n)$ satisfies the following difference equation

$$C(n) = \sum_{i=1}^{p} a_i C(n-i) + \sum_{i=1}^{p} b_i C(n-i)e(n-1) + e(n)$$

$$+ \mu \left(\sum_{i=1}^{p} b_i \right) e(n-1) - b_1 \sigma_e^2 . \qquad \text{(A3)}$$

$$r_{cc}(m) = E\Bigg\{ \underbrace{\sum_{i=1}^{p} a_i C(n-i)C(n-m)}_{①} + \underbrace{\sum_{i=1}^{p} b_i C(n-i)C(n-m)e(n-1)}_{②}$$

$$+ \underbrace{C(n-m)e(n)}_{③} + \underbrace{\mu \left(\sum_{i=1}^{p} b_i \right) e(n-1)C(n-m)}_{④} - \underbrace{b_1 \sigma_e^2 C(n-m)}_{⑤} \Bigg\} . \qquad \text{(A4)}$$

It is very easy to see that the following are true.

$$E\{①\} = \sum_{i=1}^{p} a_i r_{cc}(m-i) \qquad \text{(A5)}$$

$$E\{③\} = \begin{cases} \sigma_e^2 \; m = 0 \\ 0 \, ; \; m > 0 \end{cases} \qquad \text{(A6)}$$

$$E\{④\} = \begin{cases} \mu \left(\sum_{i=1}^{p} b_i \right) \sigma_e^2 \; ; \; m = 1 \\ 0 \qquad ; \; m > 1 \end{cases} \qquad \text{(A7)}$$

and

$$E\{⑤\} = 0 . \qquad \text{(A8)}$$

We now need to obtain an expression for $E\{C(n-\ell)\,C(n-m)\,e(n)\}$ to complete the calculation of $r_{cc}(m)$ for $m \geq 1$. Note that if ℓ and m are greater than zero,

$$E\{C(n-\ell)C(n-m)e(n)\} = 0 . \qquad \text{(A9)}$$

If $\ell = 0$ and $m > 0$, expanding C(n) as in (2), and taking the expected value of each term of the product of the expansion with C(n-m)e(n) will show that

$$E\{C(n)C(n-m)e(n)\} = 0 \quad ; m > 0 . \qquad \text{(A10)}$$

If $\ell = m = 0$,

$$E\{C^2(n)e(n)\} = E\Bigg\{ \Bigg(\sum_{i=1}^{p} a_i C(n-i) + \sum_{i=1}^{p} b_i C(n-i)e(n-1)$$

$$+ e(n) + \mu \left(\sum_{i=1}^{p} b_i \right) e(n-1) - b_1 \sigma_e^2 \Bigg)^2 e(n) \Bigg\} . \qquad \text{(A11)}$$

It is straightforward to show that the only nonzero expectations that appear on the right-hand side of (A11) are $2E\{b_1 C(n-1)\; e(n-1)e^2(n)\}$ and $-2b_1\sigma_e^2 E\{e^2(n)\}$ and that the expectations cancel each other, implying that

$$E\{C(n-\ell)C(n-m)e(n)\} = 0 \, ; \quad \ell, m \geq 0 . \qquad \text{(A12)}$$

Substituting (A5)-(A12) in (A4) gives the desried expression for $r_{cc}(m)$ for $m \geq 1$.

To calculate $r_{cc}(0)$, we simply square both sides of (A2) and take the statistical expectations. This yields

$$r_{cc}(0) = E\Bigg\{ \underbrace{\sum_{i=1}^{p} \sum_{j=1}^{p} a_i a_j C(n-i)C(n-j)}_{①}$$

$$+ \underbrace{\sum_{i=1}^{p} \sum_{j=1}^{p} b_i b_j C(n-i)C(n-j)e^2(n-1)}_{②} + \underbrace{e^2(n)}_{③} + \underbrace{\mu^2 \left(\sum_{i=1}^{p} b_i \right)^2 e^2(n-1)}_{④}$$

$$+ \underbrace{b_1^2 \sigma_e^4}_{⑤} + \underbrace{2 \sum_{i=1}^{p} \sum_{j=1}^{p} a_i b_j C(n-i)C(n-j)e(n-1)}_{⑥} + \underbrace{2 \sum_{i=1}^{p} a_i C(n-i)e(n)}_{⑦}$$

$$+ \underbrace{2 \sum_{i=1}^{p} a_i C(n-i)\, \mu \left(\sum_{\ell=1}^{p} b_\ell \right) e(n-1)}_{⑧} - \underbrace{2 \sum_{i=1}^{p} a_i C(n-i) b_1 \sigma_e^2}_{⑨}$$

$$+ \underbrace{2 \sum_{i=1}^{p} b_i C(n-i)e(n-1)e(n)}_{⑩} + \underbrace{2 \sum_{i=1}^{p} b_i C(n-i)e^2(n-1)\, \mu \left(\sum_{\ell=1}^{p} b_\ell \right)}_{⑪}$$

$$- \underbrace{2\, b_1 \sigma_e^2 \sum_{i=1}^{p} b_i C(n-i)e(n-1)}_{⑫} + \underbrace{2\, e(n) \left(\sum_{i=1}^{p} b_i \right) \mu \, e(n-1)}_{⑬}$$

$$- \underbrace{2\, b_1 \sigma_e^2 \, e(n)}_{⑭} - \underbrace{2\, b_1 \sigma_e^2 \, \mu \left(\sum_{i=1}^{p} b_i \right) e(n-1)}_{⑮} \Bigg\} . \qquad \text{(A13)}$$

Direct calculations will show that the statistical expectations of terms ⑥, ⑦, ⑨, ⑩, ⑬, ⑭, and ⑮ are zero. Also

$$E\{⑧\} = 2\mu \left(\sum_{i=1}^{p} b_i \right) a_1 \sigma_e^2 \qquad \text{(A14)}$$

and

$$E\{⑫\} = -2\, b_1^2 \sigma_e^4 . \qquad \text{(A15)}$$

An approach similar to that in [8] will show that

$$E\{C(n-\ell)C(n-m)e^2(n)\} = \begin{cases} r_{cc}(m-\ell)\sigma_e^2 \, ; \text{ if } m > 0 \text{ or } \ell > 0 \\ r_{cc}(0)\, \sigma_e^2 + 2\sigma_e^4 \, ; \; m = \ell = 0 \end{cases} \qquad \text{(A16)}$$

which implies that

$$E\{\textcircled{2}\} = \sigma_e^2 \sum_{i=1}^{p} \sum_{j=1}^{p} b_i b_j r_{cc}(i-j) + 2\, b_1^2 \sigma_e^4 . \qquad \text{(A17)}$$

Finally, one can also show that

$$E\{C(n)e^2(n)\} = 0 . \qquad \text{(A18)}$$

which implies that the expected value of term ① is also zero. Substituting the above results in (A13) and simplifying using the results for $r_{cc}(m)$; $m \geq 1$ will give the desired expression for $r_{cc}(0)$.

Expressions for $\Gamma_{ccc}(1, m)$ can be found in a similar fashion.

$$E\{C(n)C(n-1)C(n-m)\} = E\Bigg\{ \underbrace{\sum_{i=1}^{p} a_i C(n-i)C(n-1)C(n-m)}_{\textcircled{1}}$$

$$+ \underbrace{\sum_{i=1}^{p} b_i C(n-i)C(n-1)C(n-m)e(n-1)}_{\textcircled{2}} + \underbrace{C(n-1)C(n-m)e(n)}_{\textcircled{3}}$$

$$+ \mu \underbrace{\left(\sum_{i=1}^{p} b_i\right) C(n-1)C(n-m)e(n-1)}_{\textcircled{4}} - \underbrace{b_1 \sigma_e^2 C(n-1)C(n-m)}_{\textcircled{5}} \Bigg\} . \qquad \text{(A19)}$$

From (A12) it follows that the mean values of terms ③ and ④ are zero for $m \geq 1$. Expanding $C(n-1)$ as before, it is easy to show that

$$E\{C(n-i)C(n-1)C(n-m)e(n-1)\} = \sigma_e^2 r_{cc}(m-i) \; ; \; i, m > 1 . \qquad \text{(A20)}$$

For $i = 1$ and $m > 1$, we can expand $C^2(n-1)$ using (2), multiply each term in the expansion with $C(n-m)e(n-1)$, and take the expectation. Since $e(n-1)$ is correlated with only those terms involving $e(n-1)$ in the expansion of $C^2(n-1)$, and since $m \geq 1$ the result simplifies to

$$E\{C^2(n-1)C(n-m)e(n-1)\} = E\Bigg\{ \underbrace{2 \sum_{\ell=1}^{p} a\; C(n-\ell-1)C(n-m)e^2(n-1)}_{\textcircled{1}}$$

$$+ \underbrace{2 \sum_{\ell=1}^{p} b\; C(n-\ell-1)C(n-m)e(n-2)e^2(n-1)}_{\textcircled{2}}$$

$$+ \underbrace{2\mu \left(\sum_{\ell=1}^{p} b_\ell\right) e(n-2)e^2(n-1)C(n-m)}_{\textcircled{3}} \Bigg\} . \qquad \text{(A21)}$$

Now,

$$E\{\textcircled{3}\} = \begin{cases} 0 & ; m \geq 3 \\ 2\mu \left(\sum_{i=1}^{p} b_i\right) \sigma_e^4 & ; m = 2 . \end{cases} \qquad \text{(A22)}$$

From (A12),

$$E\{\textcircled{2}\} = 0 \;\; ; m \geq 2 . \qquad \text{(A23)}$$

Similarly,

$$E\{\textcircled{1}\} = 2\, \sigma_e^2 \sum_{\ell=1}^{p} a_i r_{cc}(m-\ell-1) \;\; ; m \geq 2 \qquad \text{(A24)}$$

$$= \begin{cases} 2\, \sigma_e^2 r_{cc}(m-1) & ; m > 2 \\ 2\, \sigma_e^2 \left[r_{cc}(m-1) - \mu \left(\sum_{i=1}^{p} b_i\right) \sigma_e^2 \right] & ; m = 2 . \end{cases} \qquad \text{(A25)}$$

Substituting all the relevant results in (A19) gives (6).

REFERENCES

[1] R. W. Brockett, "Volterra series and geometric control theory," *Automatica,* Vol. 12, pp. 167-176, 1976.

[2] C. Bruni, G. DiPillo, and G. Koch, "Bilinear Systems: An Appealing Class of 'Nearly Linear Systems' in Theory and Applications," *IEEE Trans. Automat. Control,* Vol. AC-19, pp. 334-348: 1974.

[3] J. A. Cadzow, "Spectral Estimation: An Overdetermined Rational Model Equation Approach," *Proc. IEEE,* Vol. 70, No. 9, pp. 907-939: Sept, 1982.

[4] C. W. J. Granger and A. P. Andersen, *Introduction to Bilinear Time Series Models*, Göttingen, Vandenhoeck and Ruprecht: 1978.

[5] W. K. Kim, L. Billard, and I. V. Basawa, "Estimation for the First-Order Diagonal Bilinear Time Series Model," *Journal of Time Series Analysis*, Vol. 11, No. 3, pp. 215-229: 1990.

[6] R. R. Mohler and W. J. Kolodziej, "An Overview of Bilinear System Theory and Applications," *IEEE Trans. Systems, Man and Cybernetics*, Vol. SMC-10, pp. 683-688: Oct, 1980.

[7] T. D. Pham and L. T. Tran, "On the First-Order Bilinear Time Series Model," *J. Applied Probability*, Vol. 18, pp. 617-627: 1981.

[8] T. Subba Rao, "On the Theory of Bilinear Time Series Models," *J. Roy. Statistical Soc. (B)*, Vol. 43, No. 2, pp. 244-255: 1981.