ON THE YULE WALKER EQUATIONS FOR THE ALL-POLE COEFFICIENTS

Adnan Al-Smadi
Department of Electronics Engineering
Yarmouk University
Irbid, Jordan
smadi98@yahoo.com

ABSTRACT

In this paper, a method for estimating the coefficients of an all-pole random process based on measurements corrupted by additive white Gaussian noise is described. The system is driven by a zero-mean independent and identically distributed (i.i.d) non-Gaussian sequence. The method proposes the estimation of the all-pole coefficients by extending the Yule Walker Equations to a triple correlation of the contaminated observations of the output sequence. Simulations verify the performance of the proposed method. The proposed method was compared with the Yule-Walker and Burg methods at different levels of signal-to-noise ratio (SNR) on the output.

KEYWORDS

Yule Walker Equations; All-Pole Coefficients; triple correlation; third order cumulants; signal-to-noise ratio.

1 INTRODUCTION

The use of all-pole autoregressive (AR) models has played an important role in the analysis of time series since it was introduced by Yule in 1926 [1]. The AR model is by far the most widely used because it is suitable for representing spectra with narrow peaks and it results in very simple linear equations for the AR parameters [2]. The AR methods tend to adequately describe spectra of data that is “peaky”; i.e., data whose power spectral density (PSD) is large at certain frequencies. The data in many practical applications (such as speech processing) tends to have “peaky spectra’ so that AR models are often useful. In addition, the AR models lead to a system of linear equations which is relatively simple to solve [3].

There are a number of applications where signals are modeled as all-pole AR random processes. These include spectral estimation, speech processing, noise cancellation, biomedical signal processing, vibration analysis, and time series forecasting [4-10]. For example, AR models driven by white noise have been one of the most popular models for representing speech waveforms [11].

Researchers have given considerable attention to the problem of estimating the coefficients of the AR signals from their noisy observations. That is because neglecting the presence of noise yields models that do not characterize properly the processes under consideration [12]. Many papers have been published in the technical literature for AR parameters estimation using a variety of second and higher-order statistics. Yule [1] first introduced the AR model and applied it to Wolfer’s sunspot data. Walker [13, 14] extended Yule’s work. The Yule-Walker (YW) AR method computes the AR coefficients by forming a biased estimate of the signal’s autocorrelation function, and solving the least squares minimization of the forward prediction error. The use of a biased estimate of the autocorrelation function ensures that the autocorrelation matrix is positive definite. Hence, the matrix is invertible and a solution is guaranteed to exist. The Yule-Walker equations can be solved efficiently via Levinson’s algorithm.

In1968, Burg [15] proposed a method for estimating the AR coefficients which can be viewed as an order-recursive least-squares lattice method. It
is based on the minimization of the forward and backward errors in linear predictors, with the constraint that the AR coefficients satisfy the Levinson-Durbin recursion. It avoids calculating the autocorrelation function, and instead estimates the reflection coefficients directly.

This paper presents a new approach for estimating the coefficients of non-Gaussian AR process in additive white Gaussian noise. The linear system is driven by a zero-mean independent and identically distributed (i.i.d) non-Gaussian sequence. The method proposes the estimation of the all-pole coefficients by extending the Yule Walker Equations to a triple correlation of the contaminated observations of the output sequence. Section 2 presents the proposed AR coefficients estimation technique. Numerical examples are given in Section 3. Section 4 presents some concluding remarks.

2 ALL-POLE MODEL

A noisy p-th order all-pole autoregressive (AR) of a time series with current value, \( x(k) \), is expressed as a linear function of previous values plus an error term, \( v(k) \), as

\[
x(k) = -a_1 x(k-1) - a_2 x(k-2) - \cdots - a_p x(k-p) + v(k)
\]

\[
y(k) = x(k) + w(k)
\]

where \( v(k) \) is a zero-mean driving white Gaussian noise sequence (prediction error) with finite variance \( \sigma_v^2 \), \( w(k) \) is a zero-mean white Gaussian process represents observation (or measurement) errors with finite variance \( \sigma_w^2 \), \( x(k) \) denotes the regression of the AR signal (the true AR signal), \( y(k) \) represents the noisy AR signal, \( a(k) \) are the system coefficients. The noise sequence \( v(k) \), and the observation error, \( w(k) \), are assumed to be mutually uncorrelated statically.

The problem under study is, given the all-pole AR order \( p \) and using noisy observations \( y(k) \), to make consistent estimation of the AR model coefficients \( \mathbf{a} = [a_1, a_2, \cdots, a_p]^T \). In practice, the order of the model, \( p \), is unknown that is to be determined from the data. One of the popular information criteria is the minimum description length (MDL) [16]. The MDL criterion is based on selecting the order that minimizes the description length and is defined as

\[
MDL(p) = N \ln \hat{\sigma}^2 + p \ln N
\]

where \( N \) is the length of the output sequence and \( \hat{\sigma}^2 \) is the estimated variance of the linear prediction error of the model in (1).

Let \( \{x(k), k = 0,1, \cdots, N - 1\} \) represent a segment of a stationary ergodic random process. It is assumed that the mean, if nonzero, has been removed. Multiplying both sides of Equation (1) by \( x(k+l) \) and taking the expected value, we obtain a second order autocorrelation function \( \phi_{xx}(l) \):

\[
\phi_{xx}(l) = E[x(k)x(k+l)]
\]

\[
= \sum_{k=-\infty}^{\infty} x(k)x(k-l)
\]

This function leads to the Yule Walker Equations [17] which are represented as a linear combination between \( \phi_{xx}(l) \) and the coefficients \( \{a(k)\} \). The triple autocorrelation (or third order cumulants) of \( x(k) \) are given by [18]

\[
R_{xx}(l, n) = E[x(k)x(k+l)x(k+n)]
\]

The triple correlation (or third order cross-cumulant) between \( y(k) \) and \( x(k) \) are given by [18]

\[
R_{yx}(l, n) = E[y(k)x(k+l)x(k+n)]
\]

Assuming \( y(k) \) is stationary, the third order cumulants of the noisy observed signal \( y(k) \) is

\[
Cum[y(k)] = E[y(k)y(k+l)y(k+n)] = Cum[x(k)+w(k)]
\]

Since \( x(k) \) and \( w(k) \) are independent, the cumulant of \( y(k) \) is the sum of cumulants of \( x(k) \) and \( w(k) \). Therefore,
\[ Cum[y(k)] = Cum[x(k)] + Cum[w(k)] \]
\[ = E[x(k)x(k+l)x(k+n)] + E[w(k)w(k+l)w(k+n)] \]  
(8)

It is assumed in Eq. (2) that \( w(k) \) is Gaussian. Hence, the cumulant of the second term in (8) goes to zero. Therefore, the cumulant of \( y(k) \) and \( x(k) \) are the same; that is

\[ R_{3y}(l,n) = R_{3x}(l,n) \]  
(9)

This paper describes an algorithm for estimating the coefficients of an all-pole AR model. The method proposes the estimation of the all-pole coefficients by extending the Yule Walker Equations to a triple correlation of the contaminated observations of the output sequence. Now, multiplying both sides of Equation (1) by \( x(k+l)x(k+n) \),

\[ x(k)x(k+l)x(k+n) = -a_1x(k-1)x(k+l)x(k+n) \]

\[-a_2x(k-2)x(k+l)x(k+n) \]

\[\cdots -a_p x(k-p)x(k+m)x(k+n)+ \]

\[\nu(k)x(k+m)x(k+n) \]  
(10)

Taking the expectation of (10) results in

\[ E[x(k)x(k+l)x(k+n)] = -a_1 E[x(k-1)x(k+l)x(k+n)] \]

\[-a_2 E[x(k-2)x(k+l)x(k+n)] \]

\[\cdots -a_p E[x(k-p)x(k+m)x(k+n)]+ \]

\[E[\nu(k)x(k+m)x(k+n)] \]  
(11)

Notice that the terms in Equation (11) are third order cumulants and cross-cumulants terms as defined in (5) and (6). Hence

\[ R_{3x}(k,n) = -a_1 R_{3x}(k+1,n+1)-a_2 R_{3x}(k+2,n+2) - \cdots - \]

\[a_p R_{3x}(k+p,n+p) + R_{3y}(k,n) \]  
(12)

This system of equations can be expressed in the matrix form as

\[ \mathbf{R}_3 = -\mathbf{R}_3 \mathbf{a} + \mathbf{v}_{33} \]  
(13)

where \( \mathbf{R}_3 \) is a vector of third order cumulant of the AR signal at lags (0,0), \( \mathbf{a} \) is a vector containing the AR coefficients, \( R_{3x} \) is a third order cumulant matrix, \( \mathbf{v}_{33} \) is a vector of third order cross-cumulant between the prediction error and the AR signal. Ideally, the third order modeling error should be zero (if the cumulants are computed accurately and the AR model order, \( p \), is selected correctly). As a result, Eq. (13) becomes

\[ \mathbf{R}_3 = -\mathbf{R}_3 \mathbf{a} \]  
(14)

Now, we wish to find the vector of the least squares estimator (LSE), \( \hat{\mathbf{a}} \), that minimizes the squared error between \( \mathbf{a} \) and \( R_{3y} \). The LSE is found by minimizing the error criterion

\[ J(\mathbf{a}) = (\mathbf{r}_3 + \mathbf{R}_3 \mathbf{a})^T (\mathbf{r}_3 + \mathbf{R}_3 \mathbf{a}) \]

\[ = \mathbf{r}_3^T \mathbf{r}_3 + \mathbf{r}_3^T \mathbf{R}_3 \mathbf{a} + \mathbf{a}^T \mathbf{R}_3^T \mathbf{r}_3 + \mathbf{a}^T \mathbf{R}_3^T \mathbf{R}_3 \mathbf{a} \]

\[ = \mathbf{r}_3^T \mathbf{r}_3 + 2 \mathbf{r}_3^T \mathbf{R}_3 \mathbf{a} + \mathbf{a}^T \mathbf{R}_3^T \mathbf{r}_3 + \mathbf{a}^T \mathbf{R}_3^T \mathbf{R}_3 \mathbf{a} \]  
(15)

To minimize (15), we take the gradient of \( J(\mathbf{a}) \) with respect to \( \mathbf{a} \).

\[ \frac{\partial J(\mathbf{a})}{\partial \mathbf{a}} = 2 \mathbf{r}_3^T \mathbf{a} + 2 \mathbf{r}_3^T \mathbf{R}_3 \mathbf{a} \]  
(16)

Setting the gradient equal to zero yields the LSE

\[ \hat{\mathbf{a}} = (\mathbf{R}_3^T \mathbf{R}_3)^{-1} \mathbf{R}_3^T \mathbf{r}_3 \]  
(17)

There are several measures taken to reveal the performance of the proposed method. The error of the estimator \( \hat{\mathbf{a}} \) is defined as the difference between the estimate and the true value of the coefficients being estimated. That is

\[ \text{Error}(\hat{\mathbf{a}}) = \hat{\mathbf{a}} - \mathbf{a} \]  
(18)

A method used to test the optimality of the estimator that combines both bias and variance is the mean-square error (MSE). The MSE of an estimator \( \hat{\mathbf{a}} \) is defined as the expected value of the squared errors and is given by [19]

\[ \text{MSE}(\hat{\mathbf{a}}) = E[(\hat{\mathbf{a}} - \mathbf{a})^2] \]  
(19)
To measure the strength of the noise, the signal-to-noise ratio (SNR) was calculated as follows [12].

\[
\text{SNR} = 10 \log_{10} \left( \frac{\sum_{k=1}^{N} (x(k))^2}{\sum_{k=1}^{N} (w(k))^2} \right) \text{ dB} \quad (20)
\]

3 Numerical Example

In this section, the performance and effectiveness of the proposed procedure have been tested using Monte Carlo computer experiments on a number of simulated examples. These examples have been simulated for both noise-free and noisy cases. The systems considered are driven by a zero-mean exponentially distributed, i.i.d random sequences. The computations were performed using MATLAB. The commands *aryule* and *arburg* were used from the Signal Processing Toolbox User’s Guide [3] to estimate the AR coefficients using YW and Burg methods, respectively. A comparison of the performance of the proposed method with several existing methods has been made. In order to guarantee statistical independence, all the results are a mean of 100 Monte Carlo runs, using different seed in each case.

A. Example 1: Consider the AR(3) process that is generated by the difference equation

\[
x(k) = 2x(k-1) - 1.7(x(k-2) + 0.5x(k-3) \quad (21)
\]

This is an AR with three poles. They are located at 0.7205 ± j06127 and 0.5590. The observed time series is \( y(k) = x(k) + w(k) \). The input sequence of length \( N=1500 \) was drawn from a zero-mean non-Gaussian distribution. Then, the input sequence was passed through the filter in equation (21). After that, the output of the filter was corrupted with additive white Gaussian noise at a SNR of 20 dB on the output sequence. The simulation with different seeds was performed 100 times using the proposed technique, YW, and Burg methods. Then, these simulations were averaged to compute the AR coefficients estimate. The performance measures considered are the arithmetic mean and the MSE of the estimates for the parameter vector \( \hat{a} \). Table 1 shows the arithmetic mean results of 100 Monte Carlo simulations for all the compared techniques. Table 2 shows the MSE results for these techniques.

<table>
<thead>
<tr>
<th>TABLE 1. TRUE AND ESTIMATED AR(3) MODEL COEFFICIENTS FOR EXAMPLE 1, ( a(0)=1 ).</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>(a(1))</td>
</tr>
<tr>
<td>(a(2))</td>
</tr>
<tr>
<td>(a(3))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2. THE MSE RESULTS FOR EXAMPLE 1.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burg</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>(a(1))</td>
</tr>
<tr>
<td>(a(2))</td>
</tr>
<tr>
<td>(a(3))</td>
</tr>
</tbody>
</table>

4 Conclusion

This paper presents a method for estimating the coefficients of an all-pole random process based on measurements corrupted by additive white Gaussian noise. The proposed algorithm uses a covariance matrix (R) formed by performing a triple correlation of the contaminated observations of the output sequence. A comparison of the performance of the proposed algorithm with the Yule-Walker and the

Burg methods at different levels of signal-to-noise ratio (SNR) on the output was made. The results presented here at 20 SNR. The presented simulation results demonstrate the effectiveness of the proposed method.

5 REFERENCES