The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering and Mathematics (EPSTEM), 2025

Volume 34, Pages 1-7

ICBASET 2025: International Conference on Basic Sciences, Engineering and Technology

A Novel Technique for Estimation the Coefficients of ARMA Model Using Higher Order Moments

Adnan Al-Smadi
Zarqa University
Yarmouk University

Abstract: Time series data is observed in many different areas such as communication systems, signal processing, climate data and earthquake data. Statistical modeling and analysis of time series data includes transformation of the data into stationary times series and fit the time series model to the transformed data. Autoregressive Moving Average (ARMA) model is one of the most often used to fit the time series data. The proper estimation of the coefficients in the time series model is one of the important steps of modeling. In this study, a novel technique for estimating the coefficients of non-Gaussian ARMA model using higher order moments of the observed data. The proposed ARMA coefficients estimator is based on building a special matrix with entries of higher order moments of the observed output only. The observed output data may be corrupted with additive white Gaussian noise. Simulation results promise that the proposed method achieves performance comparable to existing well-known methods even when the available output signal is heavily corrupted with additive white Gaussian noise.

Keywords: Time series data, Coefficients estimation, ARMA model, Higher order moments

Introduction

Time series data is observed in many different areas such as communication systems, signal processing, climate data and earthquake data. Statistical modeling and analysis of time series data includes transformation the data into stationary times series and fit the times series model to the transformed data. Autoregressive Moving Average (ARMA) model is one of the most often used to fit the time series data (Al-Smadi & Wilkes, 2002; Wang et al. 2017). ARMA model is important and extremely useful in modeling predicting future values of a time series. The proper estimation of the coefficients in time series model is one of the important steps of modeling. The literature has a lot of papers that deals with the estimating of the coefficients of a general ARMA model using second order and third order moments (or cumulants) (Giannakis & Mendel, 1989; Swami & Mendel,1990; Wang et al. 2017).

The second order statistics work fine if the analyzed signal has Gaussian probability density function since all of its properties determined by the first and second order moments (Al-Smadi & Smadi, 2003). However, there are many real-life situations where the signal is non-Gaussian. Even though Gaussian random process still plays a significant role when processing data, non-Gaussian random processes and higher order moments are of increasing interest. Giannakis and Mendel (1989) proposed a Residual Time Series (RTS) procedure for the identification of linear time invariant (LTI) nonminimum phase systems using second and third order moments when only output data are available. They assumed that the order is given in modeling an autoregressive moving-average process. The basic idea of the algorithms in literature is to estimate the AR coefficient. Then, a residual MA time series is formed. Finally, the MA coefficients are estimated.

In system identification, the use of the Cholesky decomposition in the prediction and estimation of ARMA time series was used by Ansley (1976). He applied the Cholesky decomposition to the problem of estimating the

- Selection and peer-review under responsibility of the Organizing Committee of the Conference

likelihood function of the MA and the ARMA models. Parzen and Pagano (1979) used Cholesky decomposition for obtaining predictors. Lee et. al. (2017) proposed linear models with a covariance matrix that is modeled using the coefficients of an ARMA Cholesky decomposition (ARMACD). These models allow for non-stationary processes.

In this paper, a novel technique for estimating the coefficients of non-Gaussian ARMA model using higher order moments of the observed data. The proposed ARMA coefficients estimator is based on building a special matrix with entries of higher order moments of the observed output time series data only. The observed output time series data may be corrupted with additive white Gaussian noise (Vinothkumar & Manoj, 2024). Simulation results promise that the proposed method is based on the Cholesky decomposition and achieves good performance comparable to existing well-known methods even when the available output signal is heavily corrupted with additive white Gaussian noise. Section 2 presents the problem formulation.

Problem Formulation

A general model for Autoregressive Moving Average (ARMA) model can be represented as follows:

$$A(z^{-1})x(n) = C(z^{-1})w(n) \tag{1}$$

where $x(n)$ is the observed noiseless output data. The input signal $w(n)$ is zero-mean, white noise, and non-Gaussian random process. The z^{-1} is the backward shift operator, that is,

$$z^{-1}x(n) = x(n\text{-}1) \tag{2}$$

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \ldots a_p z^{-p} \tag{3}$$

$$C(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \ldots b_q z^{-q} \tag{4}$$

In this system, the a_i and b_i are the coefficients of the ARMA model, while p and q are the orders of the denominator and numerator polynomials, respectively. In this study, it is assumed that the orders p and q are known. The output signal $x(n)$ is observed in additive white Gaussian noise $e(n)$ as follows:

$$y(n) = x(n) + e(n) \tag{5}$$

The relationship of (1) can be rewritten in matrix form as follows, assuming the length of the sequence is N.

$$[\,Y_p W_q\,]\,\underline{\theta}_{pq} = \underline{e} \tag{6}$$

Or

$$R_{pq}\,\underline{\theta}_{pq} = \underline{e} \tag{7}$$

where R_{pq} is a composite data matrix such that

$$R_{pq} = [\,Y_p W_q\,] \tag{8}$$

$$Y_p = \begin{bmatrix} y(0) & 0 & \cdots & 0 \\ y(1) & y(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ y(N-1) & y(N-2) & \cdots & y(N-p) \end{bmatrix} \tag{9}$$

$$W_q = \begin{bmatrix} w(0) & 0 & \cdots & 0 \\ w(1) & w(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ w(N-1) & w(N-2) & \cdots & w(N-q) \end{bmatrix} \tag{10}$$

$$\underline{\theta} = [1\ a_1\ \ldots\ a_p\ \ -b_0 \ldots\ -b_q]^T \tag{11}$$

$\underline{\theta}$ is the coefficients vector, and $\underline{e}$ is an $N \times 1$ observation and/or modeling error vector.

Proposed Algorithm

Let $y(k)$ be a zero-mean k^{th}-order stationary random process, then the k^{th}-order cumulant of this process is defined as the joint k^{th}-order cumulant of the random variables $y(k)$, $y(k+t_1)$, $y(k+t_2)$, $\cdots$,$y(k+t_{N-1})$. That is (Mendel, 1991)

$$c_k^y(t_1, t_2, \cdots, t_{N-1}) = Cum(y(k), y(k+t_1), \cdots, y(k+t_{N-1})) \tag{12}$$

which depends only on the time difference t_1, t_2, t_{N-1}, because of the stationarity assumption. Now, if $y(k)$ is stationary random process and its moments up to order N exist, then the following relationships between moment and cumulant sequences of $y(k)$ exist (Nikias & Mendel, 1993). The first order cumulant (mean value):

$$c_1^y = m_1^y = E\{y(k)\} \tag{13}$$

The second order cumulant (covariance sequence):

$$c_2^y(t_1) = m_2^y(t_1) - (m_1^y)^2 \tag{14}$$

where $m_2^y(t_1)$ is the autocorrelation.

The third order cumulant:

$$c_3^y(t_1, t_2) = m_3^y(t_1, t_2) - m_1^y[m_2^y(t_1) + m_2^y(t_2) + m_2^y(t_1 - t_2)] + 2(m_1^y)^3 \tag{15}$$

If the process $y(k)$ is zero-mean, $m_1^y = 0$, then the second order and third order cumulants are identical to the second and third order moments. Hence, from Equation (12),

$$Cum\{y(n)\} = C_y = E[y(n)y(n+m)y(n+k)] \tag{16}$$

where the operator $E[.]$ represents the average value. The third order cross-cumulant between the signal $y(k)$ and the signal $w(k)$ is given by

$$C_{yww}(m,k) = E[y(n)w(n+m)w(n+k)] \tag{17}$$

Multiplying both sides of (5) by $y(n+m)y(n+k)$, we obtain

$$y(n)\, y(n+m)y(n+k) = x(n)\, y(n+m)y(n+k) + e(n)y(n+m)y(n+k) \tag{18}$$

Substituting Equation (5) into the first part of the right side of Equation (18),

$$y(n)\, y(n+m)y(n+k) = x(n)[x(n+m) + e(n+m)][x(n+k) + e(n+k)] + e(n)y(n+m)y(n+k) \tag{19}$$

Simplifying (19) and taking the expected value will result in

$E[y(n)\, y(n+m)y(n+k)] = E[x(n)\, x(n+m)x(n+k)] + E[x(n)\, x(n+m)e(n+k)] + E[x(n)\, e(n+m)x(n+k)] + E[x(n)\, e(n+m)e(n+k)] + E[e(n)\, y(n+m)y(n+k)]$ (20)

Since the additive Gaussian noise, $e(n)$, is independent of both signals $x(n)$ and $y(n)$, then the second, third, fourth, and fifth terms on the right-hand side of Equation (20) go to zeros. Hence, the cumulants of the corrupted sequence $y(n)$ and the noiseless sequence $x(n)$ are theoretically are equal; i.e.,

$$R_y(m,k) = R_x(m,k) \tag{21}$$

Now, multiplying both sides of Equation (1) by $x(n+m)x(n+k)$ and taking expected value gives

$$R_x(m,k) = -a_1\, R_x(m+1,k+1) -...- a_p\, R_x(m+p,k+p) + b_0\, R_{wxx}(m,k) +... + b_q\, R_{wxx}(m+q,k+q) \tag{22}$$

By arranging Equation (22) for several values of m and k, the system in (22) can be as follows.

$$\underline{r} = -(R_X)\, \underline{a} + (R_{WXX})\, \underline{b} \tag{23}$$

where $\underline{r}$ is a vector containing third order cumulants at $m = k = 0$, R_X is a matrix containing the third order cumulants of the output data and R_{WXX} is a matrix containing the third order cross cumulants of the input and output signals. Equation (23) can be written as

$$\underline{r} = R_{pq}\, \underline{\theta} \tag{24}$$

where R_{pq} is a composite data matrix

$$R_{pq} = [R_X\ R_{WXX}] \tag{25}$$

The vector $\underline{\theta}$ contains the ARMA coefficients

$$\underline{\theta} = [-\underline{a} \quad \underline{b}]^T \tag{26}$$

Multiplying both sides of (24) by $(R_{pq})^T$

$$(R_{pq})^T R_{pq}\, \underline{\theta} = (R_{pq})^T\, \underline{r} \tag{27}$$

Let

$$G = (R_{pq})^T R_{pq} \tag{28}$$

Then

$$G\, \underline{\theta} = (R_{pq})^T\, \underline{r} \tag{29}$$

Now, let the right hand side of Equation (29) be

$$\underline{\alpha} = (R_{pq})^T\, \underline{r} \tag{30}$$

Hence, Equation (29) becomes

$$G\underline{\theta} = \underline{\alpha} \tag{31}$$

The matrix G can be decomposed Cholesky decomposition (Higham, 2009).

$$G = LL^T \tag{32}$$

Now, substitute (32) into (31) yields

$$LL^T \underline{\theta} = \underline{\alpha} \tag{33}$$

Now, let

$$\underline{d} = L^T \underline{\theta} \tag{34}$$

We first solve the following using forward substitution,

$$L\underline{d} = \underline{\alpha} \tag{35}$$

The, we solve the second part for $\underline{\theta}$ using back substitution

$$L^T \underline{\theta} = \underline{d} \tag{36}$$

The vector $\underline{\theta}$ contains the desired parameters.

Results and Discussion

The proposed technique for estimating the coefficients of non-Gaussian ARMA model using higher order moments of the observed data has been examined for several cases. A comparison of the performance of the Cholesky decomposition-based (CHDB) algorithm with the Residual Time Series (RTS) algorithm has been made for different SNRs on the output signal. The command *armarts* commands were used from was used from the *Higher-Order Spectral Analysis Toolbox User's Guide* (Swami et al. 1998) to estimate the ARMA coefficients using the RTS method. All the results were taken as the mathematical average of 100 Monte Carlo runs. The computations were performed in MATLAB.

Example 1. The data was generated according to the model

$$x(n) + 0.3x(n-1) + 0.25x(n-2) = w(n) + 0.95\, w(n\text{-}1) + 0.65w(n\text{-}2) \tag{37}$$

This is an ARMA (2,2) which has two poles and two zero. The poles are located at $0.5\, e^{\pm j107}$. The zeros are located at $0.81\, e^{\pm j126}$. The input time series data was drawn from a zero-mean non-Gaussian distribution. The exponential distribution was used. The next step was to the input time series through the system in (37). Then, the output of the system was corrupted with additive white Gaussian noise at signal to noise ratio of 20 dB on the output sequence. To estimate the ARMA coefficients, the composite higher order statistics (or cumulants) matrix R_{pq} in Equation (25) must be built. The matrix R_X consists of third order cumulants of the observed output time series data.

The matrix R_{WXX} consists of third order cross-cumulants of the unseen input time series and the observed output time series data. To estimate the input time series data, the method in Al-Smadi (2003) was used. Simulation with noise realizations based on different seed values was performed 100 times. The ARMA coefficients were estimated using the RTS and the proposed Cholesky decomposition-based (CHDB) algorithms at SNR of 20 dB on the output time series sequence. The average results of 100 Monte Carlo simulations for the RTS and the proposed CHDB algorithms at SNR of 20 dB on the output sequence are displayed in Table 1.

Table 1. True and estimated ARMA (2,2) model coefficients in Example 1

	True	CHDB Method	RTS Method
a(1)	0.3	0.2646	0.1865
a(2)	0.25	0.2511	0.2377
b(1)	0.95	0.9152	0.8517
b(2)	0.65	0.6288	0.5677

Conclusion

This paper presented a technique to estimate the coefficients of a general ARMA process. The proposed method uses the Cholesky decomposition of a special matrix with entries of higher order cumulants (HOC) of the available output data. The available output data may be contaminated by additive white Gaussian noise of unknown power spectral density. The simulation results prove the effectiveness of the proposed technique compared to the RTS, a well-known method in higher order spectral analysis.

Scientific Ethics Declaration

* The author declares that the scientific ethical and legal responsibility of this article published in EPSTEM journal belongs to the author.

Conflict of Interest

* The author declares no conflict of interest.

Funding

* This research is funded (fully/partially) by Zarqa University-Jordan

Acknowledgements or Notes

* This article was presented as an oral presentation at the International Conference on Basic Sciences, Engineering and Technology (www.icbaset.net) held in Trabzon/Türkiye on May 01-04, 2025.

References

Al-Smadi, A. (2003). Cumulant-based inverse filters for blind deconvolution. *International Journal of General Systems*, *32*(5), 503-515.

Al-Smadi, A., & Smadi, M. (2003). Study of the reliability of a binary symmetric channel under non-Gaussian disturbances. *International Journal of Communication Systems*,*16*(10), 865-973.

Al-Smadi, A., & Wilkes, D.M. (2002). Robust and accurate ARX and ARMA model order estimation of non-Gaussian processes. *IEEE Transactions on Signal Processing, 50*(3), 759-763.

Ansley, C. (1976). An algorithm for the exact likelihood of a mixed autoregressive-moving average process. *Biometrika*, *66*, 59-65.

Giannakis, G.B., & Mendel, J.M. (1989). Identification of nonminimum phase systems using higher order statistics. *IEEE Transactions on Acoustics, Speech, Signal Processing*, *37*(3), 360-377.

Higham, N.J. (2009). Cholesky factorization. *WIRE's Computational Statistics*, *1*(2), 251-254.

Lee, K., Baek, C., & Danial, M. (2017). ARMA Cholesky factor models for the covariance matrix of linear models. *Computational Statistics and Data Analysis*, *115*(11), 267-280.

Mendel, M. (1991). Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, 79, 278-305.

Nikias, C.L., & Mendel, J.M. (1993). Signal processing with higher order spectra," *IEEE Signal Processing Magazine,* 10(3), 10-37.

Parzen, E., & Pagsano, M. (1979). An approach to modeling seasonally stationary time series. *Journal of Econometrics*, *9*(1), 137-153.

Swami, A, Mendel, J.M., & Nikias, C. (1998). *Higher order spectral an analysis Toolbox-user's guide*. Natrick, MA: The MathWorks, Inc.

Swami, A., & Mendel, J. (1990). ARMA parameters estimation using only output cumulants. *IEEE Transactions on Signal Processing*, *38*(7), 1257-1265.

Vinothkumar, G., & Manoj -Kumar, D. (2024). A novel adaptive ANC algorithm for removal of background noise in speech applications. *The International Arab Journal of Information Technology (IAJIT)*, *21*(4), 589 - 600.

Wang, K., Wu, Y., & Gao, Y., & Li, Y. (2017). New methods to estimate the observed noise variance for an ARMA model. *Measurement*, *99*, 164-170.

Author(s) Information

Adnan Al-Smadi
Zarqa University
Department of Computer Science
Zarqa-Jordan
Contact e-mail: *a.alsmadi@zu.edu.jo*
Yarmouk University
Department of Electronics Engineering
Irbid-Jordan

To cite this article:

Al-Smadi, A. (2025). A novel technique for estimation the coefficients of ARMA model using higher order moments. *The Eurasia Proceedings of Science, Technology, Engineering and Mathematics (EPSTEM), 34,* 1-7.