JIRSS (2022)
Vol. 21, No. 02, pp 1-20
DOI:10.22034/JIRSS.2022.706993

ARMA Autocorrelation Analysis: Parameter Estimation and Goodness of Fit Test

Ahmad Reza Soltani [1]

[1] Department of Statistics and Operations Research, Faculty of Science, Kuwait University.

Received: 20/12/2022, Accepted: 15/08/2023, Published online: 14/01/2024

Abstract. The celebrated Ljung-Box residual analysis is a widely used method in time series for the parameter estimation and the goodness of fit test for the ARMA time series models. The question is whether the autocorrelation function of the fitted ARMA(p, q) model for an observed time series, at different lags, in the Ljung-Box estimation method, is close to the correlogram of observed series. The answer indeed is not affirmative. In this article, firstly, we present a new procedure in solving the Yule-Walker equations for *the exact computation* of the autocorrelation functions of ARMA(p, q) models. Secondly, we provided a goodness of fit procedure using the limiting distribution of the the sample correlation function. Thirdly, we establish a new parameters estimation method based on examining the model autocorrelation function against the series autocorrelation coefficients. The effectiveness of the procedure is brought into sight using simulated data.

Keywords. Time Series ARMA(p, q) Models, Yule-Walker Equations, Autocorrelation Function, Parameters Estimation, Goodness of Fit, Model Diagnosing.

MSC: 62M10, 60G10, 60G12.

1 Introduction

In this article, we address the classical topic in time series, namely, parameters estimations for the ARMA(p, q) processes. There are of course various methods in time series for this purpose. The common and most applied method is the residual analysis and the goodness for the fitted model by Ljung-Box (1978) chi-square test. Then estimated parameters, the fits, the residuals and the forecasts will come easily afterwards. The

Corresponding Author: Ahmad Reza Soltani (ar.soltani@ku.edu.kw)

question is whether the fitted series and the observed series exhibit close or nearly close autocorrelation functions, ACFs. The answer, in general, may not be positive. This easily can be seen: For $t = 1, 2$, let $Y_1 = X_1 + Z_1$ and $Y_2 = X_2 + Z_2$, where Z_2 is independent of X_1, X_2 and Z_1 and means are zero. Then $cov(Y_1, Y_2) = cov(X_1, X_2) + cov(Z_1, X_2)$. For time series data, $cov(Z_1, X_2) \neq 0$. Therefore $cov(Y_1, Y_2)$ and $cov(X_1, X_2)$ can be different. Indeed this is the case whenever the current noise and future series values are significantly correlated. Clearly, the autocorrelation function, as well as the autocorrelation coefficients, are essential tools in time series analysis, either in the time domain or in the spectral domain. The main aim in this article is to present a new ARMA(p, q) model building method. This article is indeed three folded. Firstly, we develop a new theoretical and computational method for the exact evaluation of the autocorrelation function of an ARMA(p, q) model, in solving the corresponding "Yule-Walker equations". To the best of our knowledge, the common procedures cited in the well known classical texts in time series, either relay on the roots of certain polynomials, Chatfiled (1975), Brockwell and Davis (1991), Pourahmadi (2001), Shumway and Stoffer (2011), or on the moving average representation, using the moving average coefficients. Neither of these methods, in practice, is handy nor gives the exact values of the autocorrelation function. Secondly, we use the true value of the autocorrelation function of the proposed ARMA(p, q) model to identify the limiting distribution of the correlation coefficients. Then, we provide a goodness of fit for the ARMA(p, q) models, using the asymptotic distribution of the sample ACF. Thirdly, we furnish a new ARMA(p, q) model building procedure by examining the autocorrelation function of the proposed model against the series autocorrelation coefficients. The methodology of our procedure, indeed, is very different from the residual analysis. The fits for the proposed parameters estimation procedure exhibit the closest ACF to the corresponding one of the observed series.

Interestingly, in addition, in our goodness of fit, the limiting distribution of the test statistic at each lag is the standard normal distribution. Hence in contrast to the existing diagnosing methods in time series, at each lag, the p-value of the test is provided by the *standard normal distribution*, as in the goodness of fit for the underlying population distribution given in Soltani (2019). The work of Khmaladze (1998), Khmaladze and Koul (2009) and Koul (2002) give insights to the goodness of fit tests.

This article is organized as follows. In Section 2, we provide preliminaries and establish the foundation of our ACF computation method for $AR(p)$ processes, that also will be used in the ACF computation for the ARMA(p, q) given in Section 3. Section 4 is devoted to the our model parameters estimation procedure. All the programs and numerical derivations are furnished using the Wolfram Mathematica (2020) Version 12.1.

2 AR(p) Autocorrelation Computation

We recall that a discrete time AR(p) time series model assumes the formulation

$$X_t - \mu = \alpha_1(X_{t-1} - \mu) + \alpha_2(X_{t-2} - \mu) + ... + \alpha_p(X_{t-p} - \mu) + Z_t, \tag{2.1}$$

where μ is the mean and $\alpha_1, ..., \alpha_p$ are the AR parameters. It is assumed that all the roots of the model characteristic polynomial $\Psi(z) = 1 - \alpha_1 z - \alpha_2 z^2 - ... - \alpha_p z^p$ lie outside the unit disc $\{|z| \leq 1\}$.

The "Yule-Walker equations" is a very familiar term in classical time series. For a discrete time AR(p) process, it is indeed the system of equations involving the model autocorrelation function (ACF), $\rho(\cdot)$, and the model parameters:

$$\rho(k) = \alpha_1 \rho(k-1) + \cdots + \alpha_p \rho(k-p), \quad \text{for all } \ k > 0, \ \ \rho(-k) = \rho(k). \tag{2.2}$$

As it is reported in well known texts in time series, such as Chatfiled (1975), Brockwell and Davis (1991) and Shumway and Stoffer (2011), the classical technique concerning solving difference equations (2.2) reads $\rho(\cdot)$ as $\rho(k) = A_1 \pi_1^k + \cdots + A_p \pi_p^k, \ \ k \geq 0$, where $\pi_1, \cdots, \pi_p$ are the roots of the auxiliary equation $y^p - \alpha_1 y^{p-1} - \cdots - \alpha_{p-1} y - \alpha_p = 0$, and the coefficients $A_1, \cdots, A_p$ can be derived due to the restrictions that are imposed by the first $p-1$ Yule-Walker equations and $\rho(0) = 1$. Although theoretically the function $\rho(\cdot)$ is provided, but tedious manipulations are required to get it in the closed form for a given model. In general, for $p \geq 3$, there in no formulation for the exact derivation of roots of a polynomial of order p. Since the roots of a polynomial of order p are computed approximately, consequently, the autocorrelation function is approximated, and even advanced computational softwares do not give exact values for the autocorrelation function. The second approach, given in Brockwell and Davis (1991), is based on the moving average representation, using the moving average coefficients. Plausibly, simulated value for large series length gives *estimates* for the autocorrelation.

In time series, mostly, the importance of Yule-Walker equations is realized from the point of the ARMA(p, q) parameters estimation. In the Matrix form, the first $p-1$ equations are written as,

$$\boldsymbol{\rho}_{(p-1)\times 1} = \mathbf{A}_{(p-1)\times p} \boldsymbol{\alpha}_{p\times 1}, \tag{2.3}$$

where

$$\mathbf{A} = \begin{pmatrix} \rho(0) & \rho(1) & \cdots & \rho(p-1) \\ \rho(1) & \rho(0) & \cdots & \rho(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(p-1) & \rho(p-2) & \cdots & \rho(1) \end{pmatrix},$$

$\boldsymbol{\alpha}_{p\times 1} = \left(\alpha_1, \cdots, \alpha_p\right)'$ and $\boldsymbol{\rho}_{(p-1)\times 1} = (\rho(1), \cdots, \rho(p-1))'$. Equation (2.3) is frequently applied for the parameter estimation of the autoregressive coefficients, where the autocorrelation function values $\rho(k)$ are replaced by the corresponding autocorrelation coefficients $r_k, \ k = 1, ..., p-1$.

In this work we bring into sight another application of the AR(p) and ARMA(p, q) Yule-Walker equations, and provide a fairly simple and practical method for the precise computation of the autocorrelation ($\rho(k)$) and autocovariance ($\gamma(k)$) functions, for, $k \geq 0$.

Our approach for the exact computation of the ACF $\rho(\cdot)$, is to rewrite the first $p-1$ equations in (2.2) in a suitable matrix form in which the vector of the first $p-1$ autocorrelation function values is expressed in terms of the autoregressive coefficients. The details are provided in the following lemma.

Lemma 2.1. For $p > 1$, the first $p-1$ Yule-Walker equations in (2.2) can be written in the matrix form as

$$\boldsymbol{\alpha}_{(p-1)\times 1} = \boldsymbol{B}_{(p-1)\times(p-1)}\boldsymbol{\rho}_{(p-1)\times 1}, \tag{2.4}$$

where

$$\boldsymbol{\alpha}_{(p-1)\times 1} = (\alpha_1, \cdots, \alpha_{p-1})', \qquad \boldsymbol{\rho}_{(p-1)\times 1} = (\rho(1), ..., \rho(p-1)), \tag{2.5}$$

and the matrix $\boldsymbol{B}_{(p-1)\times(p-1)} = \{B_{jk}\}_{j,k=1,...,p-1}$ is specified as follows.
For $j < L$:

$$B_{jk} = \begin{cases} -(\alpha_{j-k} + \alpha_{j+k}) & k = 1, 2, \cdots, j-1 \\ 1 - \alpha_{2j} & k = j \\ -\alpha_{j+k} & k = j+1, \cdots, p-j \\ 0 & k = p-j+1, \cdots, p-1, \end{cases} \tag{2.6}$$

for $j = L$:

$$B_{Lk} = \begin{cases} -(\alpha_{L-k} + \alpha_{L+k}) & k = 1, 2, \cdots, L-1 \\ 1 - \alpha_{2L} & k = L \\ 0 & k = L+1, \cdots, p-1, \end{cases} \tag{2.7}$$

for $j > L$:

$$B_{jk} = \begin{cases} -(\alpha_{j+k} + \alpha_{j-k}) & k = 1, 2, \cdots, p-j \\ -\alpha_{j-k} & k = p-j+1, \cdots, j-1 \\ 1 & k = j \\ 0 & k = j+1, \cdots, p-1, \end{cases} \tag{2.8}$$

where $L = p/2$ for p even, and $L = (p-1)/2$ for p odd.

Proof. The term $\rho(0)\alpha_k$ appears in the kth equation, $k = 1, ..., p-1$. Since $\rho(0) = 1$, the first $p-1$ equations are rewritten by keeping α_k in one side of each equation, and moving the other terms to the other side. This rewriting provides the system of equations (2.4). □

In particular, for $p = 2$ and $p = 3$, respectively,

$$\mathbf{B}_{1\times 1} = (1 - \alpha_2), \qquad \mathbf{B}_{2\times 2} = \begin{pmatrix} 1 - \alpha_2 & 0 \\ -(\alpha_1 + \alpha_3) & 1 \end{pmatrix}. \tag{2.9}$$

For $p = 4$ and $p = 5$, respectively,

$$\mathbf{B}_{3\times 3} = \begin{pmatrix} 1 - \alpha_2 & -\alpha_3 & -\alpha_4 \\ -(\alpha_1 + \alpha_3) & 1 - \alpha_4 & 0 \\ -(\alpha_2 + \alpha_4) & -\alpha_1 & 1 \end{pmatrix}, \tag{2.10}$$

$$\mathbf{B}_{4\times 4} = \begin{pmatrix} 1-\alpha_2 & -\alpha_3 & -\alpha_4 & -\alpha_5 \\ -\alpha_1-\alpha_3 & 1-\alpha_4 & 0 & 0 \\ -\alpha_2-\alpha_4 & -\alpha_1-\alpha_5 & 1 & 0 \\ -\alpha_3-\alpha_5 & -\alpha_2 & -\alpha_1 & 1 \end{pmatrix}. \tag{2.11}$$

Example 2.1. In this example, we let $\alpha(k)$ stands for α_k as well.
a. For $p = 2$,

$$\begin{aligned}
\rho(1) &= \frac{\alpha(1)}{1-\alpha(2)}, \\
\rho(2) &= \frac{\alpha(1)^2}{1-\alpha(2)} + \alpha(2), \\
\rho(3) &= \frac{\alpha(1)\alpha(2)}{1-\alpha(2)} + \alpha(1)\left(\frac{\alpha(1)^2}{1-\alpha(2)} + \alpha(2)\right).
\end{aligned}$$

b. For $p = 3$,

$$\begin{aligned}
\rho(1) &= \frac{\alpha(1)}{1-\alpha(2)}, \\
\rho(2) &= \alpha(2) + \frac{\alpha(1)(\alpha(1)+\alpha(3))}{1-\alpha(2)}, \\
\rho(3) &= \alpha(1)\left(\alpha(2) + \frac{\alpha(1)(\alpha(1)+\alpha(3))}{1-\alpha(2)}\right) + \frac{\alpha(1)\alpha(2)}{1-\alpha(2)} + \alpha(3), \\
\rho(4) &= \alpha(1)\left(\alpha(1)\left(\alpha(2) + \frac{\alpha(1)(\alpha(1)+\alpha(3))}{1-\alpha(2)}\right) + \frac{\alpha(1)\alpha(2)}{1-\alpha(2)} + \alpha(3)\right) \\
&+ \left(\alpha(2) + \frac{\alpha(1)(\alpha(1)+\alpha(3))}{1-\alpha(2)}\right)\alpha(2) + \frac{\alpha(1)\alpha(3)}{1-\alpha(2)}.
\end{aligned}$$

c. For $p = 4$,

$$\begin{aligned}
\rho(1) &= \frac{A1}{B}, \\
A1 &= \alpha(1)(1-\alpha(4)) + \alpha(2)(\alpha(3)+\alpha(1)\alpha(4)) + \alpha(3)\left(\alpha(4)-\alpha(4)^2\right), \\
B &= \alpha(4)^3 + \alpha(2)\alpha(4)^2 - \alpha(4)^2 - \alpha(1)^2\alpha(4) - \alpha(1)\alpha(3)\alpha(4) \\
&- \alpha(4) - \alpha(3)^2 - \alpha(2) - \alpha(1)\alpha(3) + 1, \\
\rho(2) &= \frac{A2}{B}, \\
A2 &= \alpha(1)(\alpha(1)+\alpha(3)) + \alpha(3)(\alpha(1)\alpha(4)+\alpha(3)\alpha(4)) \\
&+ \alpha(2)\left(-\alpha(4)^2 - \alpha(2)\alpha(4) - \alpha(2) + 1\right), \\
\rho(3) &= \frac{A3}{B}, \\
A3 &= \alpha(3)\left(-\alpha(3)^2 - \alpha(1)\alpha(3) - \alpha(2) + \alpha(2)\alpha(4) - \alpha(4) + 1\right) \\
&+ \alpha(2)(-\alpha(2)\alpha(1) + \alpha(1) + \alpha(2)\alpha(3) + \alpha(3)\alpha(4)) \\
&+ \alpha(1)\left(\alpha(1)^2 + \alpha(3)\alpha(1) - \alpha(4)^2 + \alpha(2) - \alpha(2)\alpha(4) + \alpha(4)\right).
\end{aligned}$$

For $p > 1$, Lemma 2.1 can be effectively applied to derive the values of the autocorrelation function $\rho(k),\ k = 1, ..., p-1$, analytically and numerically:

$$\boldsymbol{\rho}_{(p-1)\times 1} = \boldsymbol{B}^{-1}_{(p-1)\times(p-1)}\boldsymbol{\alpha}_{(p-1)\times 1}. \tag{2.12}$$

For $p = 1$, $\rho(k) = \alpha^k,\ k \geq 0$. Then, the Yule-Walker equations recursively provide the values for $\rho(k),\ k \geq p$.

For a given value for σ_Z^2, and the ACF $\rho(k),\ k = 0, ..., p$, the variance of the AR(p) model X_t, σ_X^2, is given by

$$\sigma_X^2 = \frac{\sigma_Z^2}{1 - 2\sum_{k=1}^{p} \alpha_k \rho(k) + \sum_{i=1}^{p}\sum_{j=1}^{p} \alpha_i \alpha_j \rho(|i-j|)}. \tag{2.13}$$

3 ARMA(p,q) Processes

We let $\{X_t\}$ denote a mean zero discrete time ARMA(p,q) process,

$$X_t = \alpha_1 X_{t-1} + ... + \alpha_p X_{t-p} + Z_t + \beta_1 Z_{t-1} + ... + \beta_q Z_{t-q}, \quad t \in \mathbb{Z}; \tag{3.1}$$

where $\beta_1, \cdots, \beta_q$ denote the moving average (MA) coefficients. It is also assumed that the roots of the polynomial $\Phi(z) = 1 + \beta_1 z + \beta_2 z^2 + ... + \beta_q z^q$ are not in the unit disc $\{|z| \leq 1\}$.

The corresponding Yule-Walker equations can be written as

$$\rho(k) = \alpha_1 \rho(k-1) + \cdots + \alpha_p \rho(k-p) + \beta_k \eta(0) + \beta_{k+1}\eta(1) + \cdots + \beta_q \eta(q-k), \tag{3.2}$$

$k \geq 1$, where

$$\eta(j) = \begin{cases} E(X_t Z_{t-j})/\sigma_X^2 & j = 1, 2, \cdots \\ \sigma_Z^2/\sigma_X^2 & j = 0 \\ 0 & j = -1, -2, \cdots, \end{cases} \tag{3.3}$$

and satisfy

$$\eta(k) = \alpha_1 \eta(k-1) + \cdots + \alpha_p \eta(k-p) + \beta_k \eta(0), \quad \text{for all } k \geq 1, \tag{3.4}$$

where $\beta_k = 0,\ k > q$.

Lemma 3.1. Let

$$\eta^{(1)}(k) = \frac{\eta(k)}{\eta(0)} = \frac{1}{\sigma_Z^2} E\left[X_t Z_{t-k}\right], \qquad k \geq 0. \tag{3.5}$$

Also let

$$\boldsymbol{\eta}^{(1)}_{(p-1)\times 1} = \left(\eta^{(1)}(1), ..., \eta^{(1)}(p-1)\right)', \quad \text{and} \quad \boldsymbol{\beta}_{(p-1)\times 1} = \left(\beta_1, ..., \beta_{p-1}\right)'. \tag{3.6}$$

Then for $p > 1$,

$$\mathbf{C}_{(p-1)\times(p-1)}\boldsymbol{\eta}^{(1)} = \boldsymbol{\alpha}_{(p-1)\times 1} + \boldsymbol{\beta}_{(p-1)\times 1}, \tag{3.7}$$

where

$$\mathbf{C}_{(p-1)\times(p-1)} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -\alpha_1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ -\alpha_2 & -\alpha_1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ -\alpha_{p-2} & -\alpha_{p-3} & -\alpha_{p-4} & \alpha_{p-5} & \cdots & -\alpha_1 & 1 \end{pmatrix}; \tag{3.8}$$

Indeed

$$\mathbf{C}_{(p-1)\times(p-1)} = [C_{j,k}], \quad C_{j,k} = \begin{cases} -\alpha_{j-k}, & k < j \\ 1, & k = j \\ 0 & k > j, \end{cases} \tag{3.9}$$

Proof. The proof follows from the fact that $\eta^{(1)}(0) = 1$, and the autoregressive structure for $\eta^{(k)}(0),\ k = 1, ..., p-1$. □

By using Lemma 2.1 and Lemma 3.1, the values of $\eta^{(1)}(k),\ k \geq 1$ can be calculated. For calculating $\rho(k)$, it is to compute $\eta(0)$ first. We proceed by deriving $\eta(0)$. Let us first introduce the following notations.

$$\boldsymbol{\beta}^{(k)}_{(q-k+1)\times 1} = \left(\beta_k, ..., \beta_q\right)', \quad \beta_k = 0,\ k > q, \tag{3.10}$$

$$\boldsymbol{\eta}^{(1)}_{(q-k+1)\times 1} = \left(\eta^{(1)}(0), ..., \eta^{(1)}(q-k)\right)', \quad \eta^{(1)}(j) = 0,\ j < 0, \tag{3.11}$$

$$\boldsymbol{\delta}_{(p-1)\times 1} = (\delta_1, ..., \delta_{p-1})', \tag{3.12}$$

$$\delta_k = \boldsymbol{\eta}^{(1)\prime}_{(q-k+1)\times 1} \cdot \boldsymbol{\beta}^{(k)}_{(q-k+1)\times 1},\ k = 1, ..., \min\{p-1, q\}, \tag{3.13}$$

$$\delta_k = 0, \quad k > \min\{p-1, q\},\ k \neq p, \tag{3.14}$$

$$\delta_p = \boldsymbol{\eta}^{(1)\prime}_{(q-p+1)\times 1} \cdot \boldsymbol{\beta}^{(p)}_{(q-p+1)\times 1},\ \ p \leq q; \quad \delta_p = 0,\ \ p > q, \tag{3.15}$$

$$\boldsymbol{\tau}_{(p-1)\times 1} = \left(\tau_1, ..., \tau_{p-1}\right)', \quad \tau_k = \sum_{i=1}^{p-k} \alpha_i\alpha_{i+k} - \alpha_k,\ \ k = 1, ..., p-1. \tag{3.16}$$

Lemma 3.2. For $p > 1,\ q \geq 1$,

a.

$$\mathbf{B}_{(p-1)\times(p-1)}\boldsymbol{\rho}_{(p-1)\times 1} = \boldsymbol{\alpha}_{(p-1)\times 1} + \eta(0)\boldsymbol{\delta}_{(p-1)\times 1}, \tag{3.17}$$

b.

$$\rho(p) = \alpha_p + \boldsymbol{\alpha}^{**\prime}_{(p-1)\times 1}\boldsymbol{\rho}_{(p-1)\times 1} + \eta(0)\delta_p, \tag{3.18}$$

where $\boldsymbol{\alpha}^{**}_{(p-1)\times 1} = (\alpha_{p-1}, \cdots, \alpha_1)'$.

c.

$$(1 + \sum_{j=1}^{p} \alpha_j^2) + 2\sum_{k=1}^{p-1}\{\sum_{i=1}^{p-k} \alpha_i\alpha_{i+k} - \alpha_k\}\rho(k) - 2\alpha_p\rho(p) = (1 + \sum_{j=1}^{q} \beta_j^2)\eta(0). \tag{3.19}$$

Proof. a. Since $\rho(0) = 1$, the same reasoning as in Lemma 2.1, provides (3.17). Part **b** follows from (3.2) with $k = p$ and Part **c** follows from the fact that the expressions on both sides in

$$X_t - \alpha_1 X_{t-1} - \ldots - \alpha_p X_{t-p} = Z_t + \beta_1 Z_{t-1} + \ldots + \beta_q Z_{t-q},$$

possess the same variance. □

Next, we replace $\rho(p)$ in (3.19) by its value in (3.18), then we solve the resulting equation for $\eta(0)$. Indeed

$$\begin{aligned} & (1 + \sum_{j=1}^{p} \alpha_j^2) + 2\boldsymbol{\tau}'_{(p-1)\times 1} \cdot \boldsymbol{\rho}_{(p-1)\times 1} - 2\alpha_p \left(\alpha_p + \boldsymbol{\alpha}^{**\prime}_{(p-1)\times 1}\boldsymbol{\rho}_{(p-1)\times 1} + \eta(0)\delta_p\right) \\ = \; & (1 + \sum_{j=1}^{q} \beta_j^2)\eta(0). \end{aligned}$$

Therefore,

$$\begin{aligned} & (1 + \sum_{j=1}^{p} \alpha_j^2) + 2\boldsymbol{\tau}'_{(p-1)\times 1} \cdot \boldsymbol{\rho}_{(p-1)\times 1} - 2\alpha_p\boldsymbol{\alpha}^{**\prime}_{(p-1)\times 1} \cdot \boldsymbol{\rho}_{(p-1)\times 1} - 2\alpha_p^2 \\ = \; & 2\alpha_p\eta(0)\delta_p + (1 + \sum_{j=1}^{q} \beta_j^2)\eta(0). \end{aligned}$$

Giving that,

$$\begin{aligned} & (1 + \sum_{j=1}^{p} \alpha_j^2) + \left\{2\boldsymbol{\tau}'_{(p-1)\times 1} - 2\alpha_p\boldsymbol{\alpha}^{**\prime}_{(p-1)\times 1}\right\} \cdot \boldsymbol{\rho}_{(p-1)\times 1} - 2\alpha_p^2 \\ = \; & 2\alpha_p\eta(0)\delta_p + (1 + \sum_{j=1}^{q} \beta_j^2)\eta(0). \end{aligned}$$

Thus,

$$\begin{aligned}
&(1+\sum_{j=1}^{p}\alpha_j^2) \\
&\quad +\left\{2\boldsymbol{\tau}'_{(p-1)\times 1}-2\alpha_p\boldsymbol{\alpha}^{**\prime}_{(p-1)\times 1}\right\}\cdot\left\{\mathbf{B}^{-1}_{(p-1)\times(p-1)}\boldsymbol{\alpha}_{(p-1)\times 1}+\eta(0)\mathbf{B}^{-1}_{(p-1)\times(p-1)}\boldsymbol{\delta}_{(p-1)\times 1}\right\} \\
&\quad -2\alpha_p^2 \\
=\;& 2\alpha_p\eta(0)\delta_p+(1+\sum_{j=1}^{q}\beta_j^2)\eta(0).
\end{aligned}$$

Consequently,

$$\begin{aligned}
&(1+\sum_{j=1}^{p}\alpha_j^2)+\left\{2\boldsymbol{\tau}'_{(p-1)\times 1}-2\alpha_p\boldsymbol{\alpha}^{**\prime}_{(p-1)\times 1}\right\}\cdot\left\{\mathbf{B}^{-1}_{(p-1)\times(p-1)}\boldsymbol{\alpha}_{(p-1)\times 1}\right\}-2\alpha_p^2 \\
=\;& -\left\{2\boldsymbol{\tau}'_{(p-1)\times 1}-2\alpha_p\boldsymbol{\alpha}^{**\prime}_{(p-1)\times 1}\right\}\cdot\left\{\mathbf{B}^{-1}_{(p-1)\times(p-1)}\boldsymbol{\delta}_{(p-1)\times 1}\right\}\eta(0) \\
&\quad +2\alpha_p\eta(0)\delta_p+(1+\sum_{j=1}^{q}\beta_j^2)\eta(0),
\end{aligned}$$

giving the formula in the following lemma for the $\eta(0)$.

Lemma 3.3. For $p>1,\ q\geq 1$, the quantity $\eta(0)$ is given by

$$\eta(0)=\frac{C}{D}, \tag{3.20}$$

where

$$C=(1+\sum_{j=1}^{p}\alpha_j^2)+\left\{2\boldsymbol{\tau}'_{(p-1)\times 1}-2\alpha_p\boldsymbol{\alpha}^{**\prime}_{(p-1)\times 1}\right\}\cdot\left\{\mathbf{B}^{-1}_{(p-1)\times(p-1)}\boldsymbol{\alpha}_{(p-1)\times 1}\right\}-2\alpha_p^2, \tag{3.21}$$

$$D=(1+\sum_{j=1}^{q}\beta_j^2)+2\alpha_p\delta_p-\left\{2\boldsymbol{\tau}'_{(p-1)\times 1}-2\alpha_p\boldsymbol{\alpha}^{**\prime}_{(p-1)\times 1}\right\}\cdot\left\{\mathbf{B}^{-1}_{(p-1)\times(p-1)}\boldsymbol{\delta}_{(p-1)\times 1}\right\}. \tag{3.22}$$

We record that $\sigma_X^2=\frac{\sigma_Z^2}{\eta(0)}$.

For $p=1$,

$$\eta^{(1)}(1)=\alpha+\beta_1,\quad \eta(0)=\frac{1-\alpha^2}{2\alpha\delta_1+(1+\sum_{i=1}^{q}\beta_i^2)}. \tag{3.23}$$

3.1 Autocovariance and Autocorrelation Functions Formulations and Computation Procedure

The *formulation-computational algorithm* for computing $\rho(k),\ k \geq 1$ goes as follows.

(I). In spirit of Lemma 3.1, (3.7) is applied to deduce $\eta^{(1)}(k),\ k \geq 1$.

(II). Lemma 3.3, (3.20), together with the derivations in (I) are applied to compute $\eta(0)$.

(III). Lemma 3.1 together with the value of $\eta(0)$ are applied to compute $\rho(k),\ k = 1, ..., p-1$:
For $p > 1$,

$$\boldsymbol{\rho}_{(p-1)\times 1} = \mathbf{B}^{-1}_{(p-1)\times(p-1)}\boldsymbol{\alpha}_{(p-1)\times 1} + \eta(0)\mathbf{B}^{-1}_{(p-1)\times(p-1)}\boldsymbol{\delta}_{(p-1)\times 1}, \tag{3.24}$$

For $p = 1$,

$$\rho(1) = \alpha + \eta(0)\delta_1. \tag{3.25}$$

(IV). Lemma 3.1 together with the derivations in (I), (II) and (III) are applied to compute $\rho(k),\ k \geq p$:

$$\rho(k) = \alpha_1\rho(k-1) + \cdots + \alpha_p\rho(k-p) + \eta(0)\{\beta_k + \beta_{k+1}\eta^{(1)}(1) + \cdots + \beta_q\eta^{(1)}(q-k)\}. \tag{3.26}$$

Numerical implementations are given in Section 5.

4 Parameter Estimation and Goodness of Fit via ACF

As it was discussed in Section 1, the autocorrelation function, ACF, of the fitted ARMA(p, q) model, provided by the residuals, is not necessarily close to the sample ACF of the observed series, more numerical demonstrations will be given in Section 5. In this section, we provide a new ARMA(p, q) parameters estimation procedure based the observed series ACF, SACF in brief, and ACF of the model.

We let $\{x_t,\ t = 1, \cdots, N\}$ to be an observed time series. The aim is to fit an ARMA(p, q) model to the series x_t. It is assumed that x_t exhibits no trend and no seasonality. The derivation of the exact values for the ACF of an ARMA(p, q) model, presented in Sections 2 & 3, enables us to pave the way for the implementation of the procedure. We let $\hat{\rho}(k)$ stands for the SACF of the series x_t and $\rho(k)[\alpha_1, \cdots, \alpha_p;\ \beta_1, \cdots, \beta_q]$ stands for the ACF of an $ARMA(p, q)[\alpha_1, \cdots, \alpha_p;\ \beta_1, \cdots, \beta_q]$ model, at lag k. An SACFvsACF ARMA(p, q) parameters estimates $[\hat{\alpha}_1, \cdots, \hat{\alpha}_p;\ \hat{\beta}_1, \cdots, \hat{\beta}_q]$ are those that minimized the *SACF-ACF Square Deviation*:

$$L[\alpha_1, \cdots, \alpha_p;\ \beta_1, \cdots, \beta_q] = \sum_{k=1}^{K} \left|\hat{\rho}(k) - \rho(k)[\alpha_1, \cdots, \alpha_p;\ \beta_1, \cdots, \beta_q]\right|^2. \tag{4.1}$$

In time series, since the estimation procedures deal with implicit functions of several variables, the estimation procedures are put into action by detecting the unknown model parameters, Chatfield (1999). Nevertheless, advances in computation technology has facilitated the technicalities. In Section 5, we provide numerical demonstrations on the effectiveness of the SACFvsACF estimation procedure. According to the derivations in Section 3, the model ACF $\rho(k)$, $k \geq 1$ are indeed functions of the model parameters $\alpha_1, \cdots, \alpha_p;\ \beta_1, \cdots, \beta_q$, through complicated but computing formulations. Therefore the criteria to minimize (4.1) indeed gives nonlinear least squares parameter estimates. The corresponding model is

$$\hat{\rho}(k) = \rho(k)[\alpha_1, \cdots, \alpha_p;\ \beta_1, \cdots, \beta_q] + \epsilon_k, \quad k = 1, ...K. \tag{4.2}$$

Thus the corresponding estimators possess the properties of the least square estimators. The consistency of the nonlinear is established in Chien-Fu Wu (1981) and Richardson and Bhattacharyya (1986), among others. Applying these results to (4.2) by itself is an interesting research topic.

Goodness of fit. Let us present one more application for the knowledge of the true (exact) values of the autocorrelation function, which is the goodness of fit for a proposed model. The methodology is based on the limiting distribution of $(\hat{\rho}_1, ..., \hat{\rho}_h)$, the sample autocorrelation function.

The commonly applied Box-Pierce (1970) and Ljung-Box (1978) goodness of fit test relies on whether the residual series is purely random, by examining the autocorrelation coefficients of the residual.

Our goodness of fit test for the ARMA(p, q) models relies on examining the autocorrelation coefficients of the observed series to the autocorrelation function values of the proposed model.

According to Theorem 7.2.2 in Brockwell and Davis (1991), for a stationary process

$$X_t - \mu = \sum_{j=-\infty}^{+\infty} \psi_j Z_{t-j}, \qquad \{Z_t\} \sim \mathbf{IID}(0, \sigma_Z^2), \tag{4.3}$$

for which either

$$\sum_{j=-\infty}^{+\infty} |\psi_j| < \infty \quad \text{and} \quad \sum_{j=-\infty}^{+\infty} |j|\psi_j^2 < \infty, \tag{4.4}$$

or

$$\sum_{j=-\infty}^{+\infty} |\psi_j| < \infty \quad \text{and} \quad EZ^4 < \infty, \tag{4.5}$$

is satisfied, then it follows that for each $h \in \{1, 2, ...\}$,

$$\hat{\boldsymbol{\rho}}_{h\times 1} \quad \text{is} \quad \mathbf{AN}\left(\boldsymbol{\rho}_{h\times 1}, n^{-1}\mathbf{W}_{h\times h}\right), \tag{4.6}$$

where $\hat{\boldsymbol{\rho}}_{h\times 1} = (\hat{\rho}(1), ..., \hat{\rho}(h))'$ and the element of the covariance matrix $\mathbf{W}_{h\times h} = [w_{ij}]$, are given by

$$w_{ij} = \sum_{k=1}^{+\infty} c_{ki}c_{kj}, \qquad c_{ki} = \rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k), \quad i, j = 1, ..., h. \tag{4.7}$$

As it is reported in Brockwell and Davis (1991), both conditions (4.4) and (4.5) are satisfied for the ARMA(p,q) models. For more on the sample ACF, see Basrak, Davis and Mikosch (1999), Pourahmadi (2003), Tsay and Pourahmadi (2017).

To introduce our diagnosing procedure, we let

$$\mathbf{1}_{h\times 1} = (1, ..., 1)'_{h\times 1}, \quad \text{and} \quad Y_h = \sum_{i=1}^{h} \hat{\rho}. \tag{4.8}$$

Then for large n,

$$Y_h \sim \mathbf{AN}\left(\sum_{i=1}^{h} \rho(i), n^{-1}\mathbf{1}'_{h\times 1}\mathbf{W}_{h\times h}\mathbf{1}_{h\times 1}\right). \tag{4.9}$$

The null hypothesis is stated as

$$H_0: \quad \mathfrak{X} \sim ARMA(p,q)[\alpha_1, ..., \alpha_p;\ \beta_1, ..., \beta_q], \tag{4.10}$$

where $\mathfrak{X} = \{x_1, ..., x_n\}$ is the observed time series. The test statistics is

$$z(h) = \frac{Y_h - \sum_{k=1}^{h} \rho(k)}{\sigma_{Y_h}}, \tag{4.11}$$

where $\rho(\cdot)$ is the autocorrelation function for the $ARMA(p,q)[\alpha_1, ..., \alpha_p;\ \beta_1, ..., \beta_q]$. The observed time series $\mathfrak{X}$ gives the statistic Y_h. The model information is used in the algorithm, given in Section 3, to compute the exact values for $\sum_{k=1}^{h} \rho(k)$ and σ_{Y_h}.

For a given lag h, $P(|z(h)| < z_{\nu/2}) = 1 - \nu$ for large n. The plot of $z(h)$ in terms h supports the null hypothesis (4.5), whenever $|z(h)| < z_{\nu/2}$, for $100(1-\nu)\%$ of times.

The numerical implementations of the procedures established in Sections 2-4 are given in the next section.

5 Numerical Implementations

In this section we provide numerical implementations of the procedures established in the previous sections. We provide examples accordingly. The programming and numerical derivations are done using Wolfram Mathematica 12.1 (2020).

Example 5.1. In this example, we illustrate the deviation between the corresponding ACF values, the exact values, presented in Section 3 (Algorithm), and the values provided by the software which uses the well known ACF computation technique. Interestingly, as it is evident in Figure 1, even with 1000 iterations, at a specific lag, the deviation is visible.

Model a:
$\{X_t\} \sim ARMA(5,6)\,[0.4,\ -1.3,\ 0.5,\ -0.6,\ 0.2\ ;\ \ -1.7,\ 0.5,\ 0.5,\ -0.3,\ 0.04,\ 0.002]$,
$\mu = 0, \sigma_Z = 1$.
The correlation function $\rho(k)$ is computed and depicted in Figure 1 using the Algorithm: *black tiny discs* plots; using the mean of values provided by the Mathematica, *red tiny squares; light black in black-white display* plots.

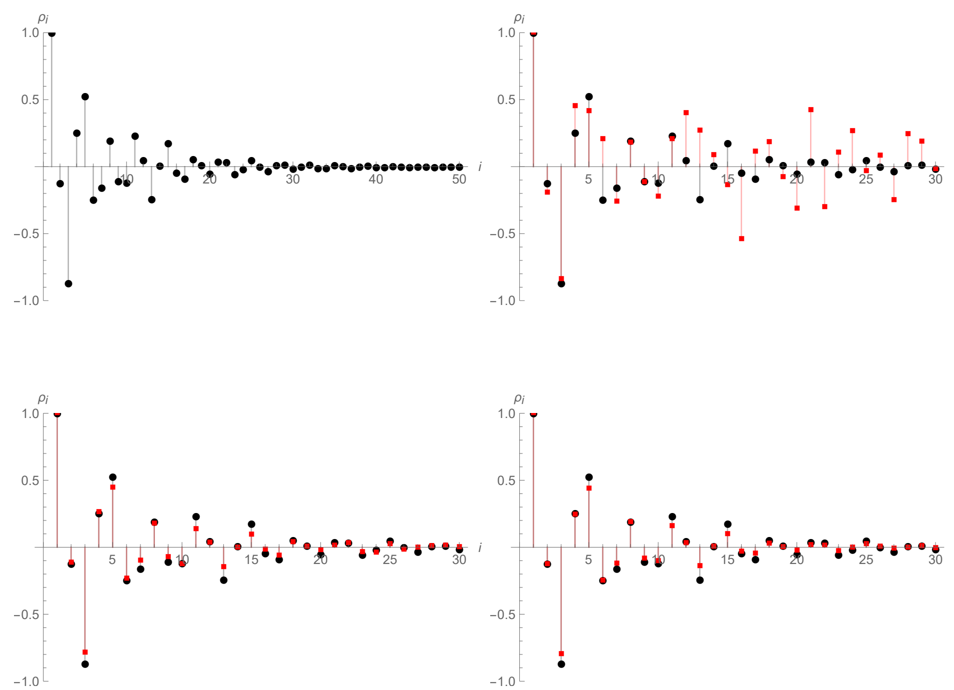

Figure 1: Example 5.1, **Model a**; Plots: $\rho(k)$ in k by the Algorithm (black discs) and the software (red squares); [Top Left: By the Algorithm]; [Top Right: By the Algorithm, and by the software, No. of iterations = 1]; [Bottom Left: By the Algorithm, and by the software, No. of iterations = 100]; [Bottom Right:By the Algorithm, and by the software, No. of iterations = 1000].

Example 5.2. This example concerns the goodness of fit skim developed in Section 4. For the **ARMA**(p,q) **model a** in Example 5.1, $\mathfrak{X}$ is a simulated series of length $n = 160$. Then for a lag h, the test statistic $z(h)$ is computed. This procedure is repeated, r number of times. Then for every lag h, *the percentage of times that* $z(h)$ *falls in the interval* $(-1.96\ ,\ +1.96)$, the coverage index $I(h)$, is measured. The plots of $I(h)$ in h, for $r = 1000,\ 10000$ are depicted in Figure 2. As it is evident in Figure 2, *the coverage index* $I(h)$ indeed surpass, the expected level 95%. This supports that the statistics $z(h)$ behavior is consistent and reliable and it supports the proposed ARMA(p,q) model as well.

We also record that for $\sigma_Z = 1$, the variance of X_t, $\sigma_X^2 = 18.3133$.

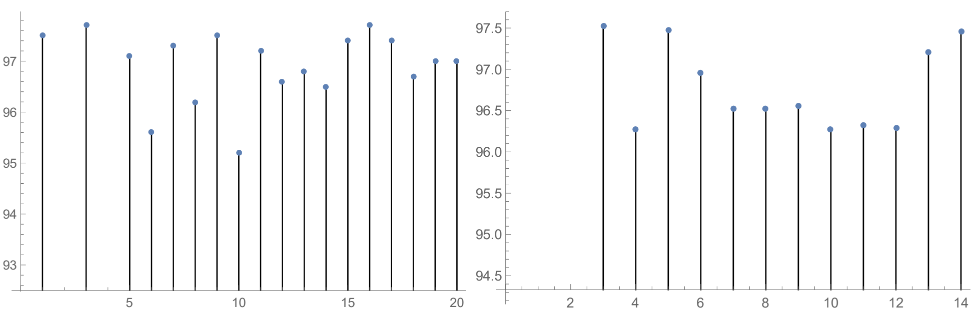

Figure 2: Example 5.2, Coverage Index, $I(h)$ Plot in h; Left: $r = 1000$ iterations; Right $r = 10000$, iterations.

Example 5.3. False Model The procedure in Example 5.2 is done for testing H_0 : $\mathfrak{X} \sim$ *ARMA Model a*. But the series $\mathfrak{X}$ is generated from:

Model b: $ARMA(4,4)\,[0.4,\ 0,\ 0.5,\ -0.6,\ 0.2;\ 0,\ 0.5,\ 0.5,\ -0.3, 0.04,\ 0.002]$, $\sigma_Z = 1$.

The corresponding plots for the coverage index $I(h)$ in h are depicted in Figure 3. The coverage index for the interval $(-1.96\ ,\ +1.96)$ for each h is below 5%, except for lags $h = 10$ and $h = 15$, which is about 10% and 8%, respectively. Thus the statistic $\{z(h),\ h\}$ does not support the Model b.

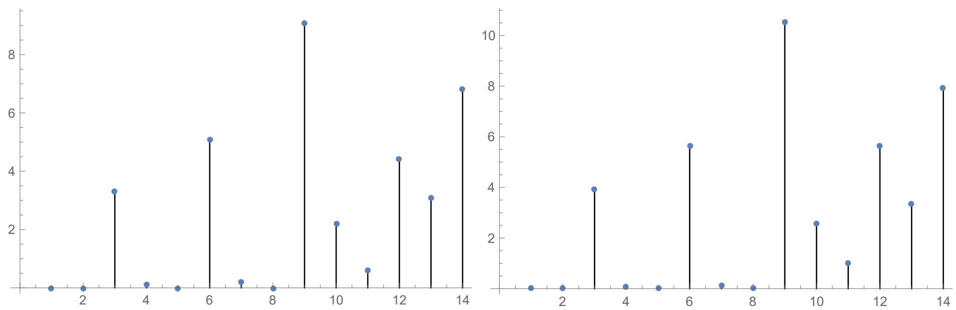

Figure 3: Example 5.3, Coverage Index, $I(h)$ Plot in h; Left: $r = 1000$ iterations; Right $r = 10000$, iterations.

Example 5.4. This example concerns our model building method developed in Section 4. The ACF for the model fitted based onthe Ljung-Box, SACFvsACF model building methods and the actual model are plotted, Figure 4.

Model c: ARMA(1, 1)[{−0.8}; {0.6}, 4], $n = 180$; $\alpha = -0.8$, $\beta = 0.6$, $\sigma_Z^2 = 4$

It is reported in Charfield (1991), and also for $q = 1$, $\delta_1 = \beta$, it follows from (3.23), (3.25) and (3.26) that $\rho(k) = \alpha^{k-1}\frac{(1+\alpha\beta)(\alpha+\beta)}{1+\beta^2+2\alpha\beta}$, $k = 1, 2, ...$

Number of iterations is setto be 100and the reported parameter estimates are the mean values in 100 repeats.

SACFvsACF output data: $\{AD = 0.95951, \quad \hat{\alpha} = -0.736427, \quad \hat{\beta} = 0.519823\}\}$

Ljung-Box Time Series Model Fit $ARMA(1, 1): \quad \tilde{\alpha} = -0.321925, \quad \tilde{\beta} = 0.0568402$

Ljung-Box p-values: {0.569039, 0.431928, 0.262392, 0.249834, 0.228022, 0.239146}

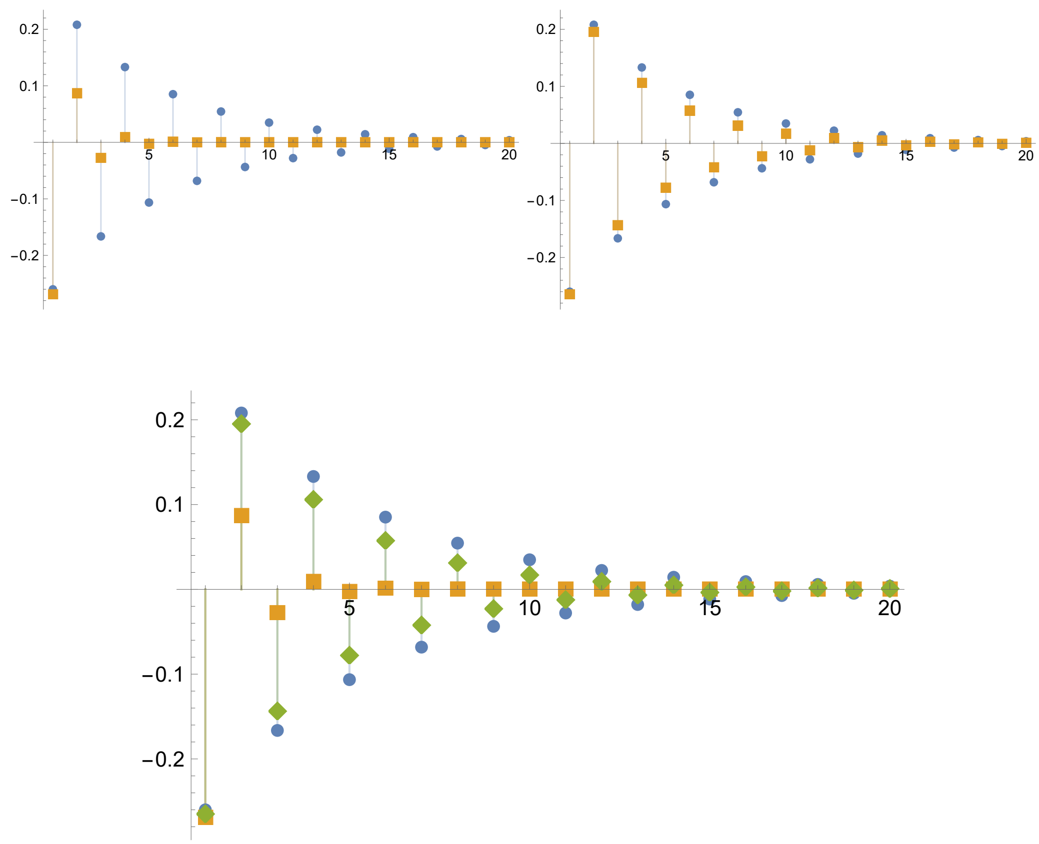

Figure 4: ACF Plots; Top Left: Data ACF in circles (blue color), Ljung-Box Fitted Model ACF in squares (orange color); Top Right: Data ACF in circles (blue color), SACFvsACF Fitted Model ACF in squares (orange color); Bottom: Data ACF in circles (blue color), Ljung-Box Fitted Model ACF in squares (orange color), SACFvsACF Fitted Model ACF in dimonds (green color).

Example 5.5. This example also concerns our model building method developed in Section 4. The ACF for the model fitted using the Ljung-Box, SACFvsACF model building methods and the actual model are plotted in Figure 5.

Model d: ARMA$(2,2)[\{0.4,-0.8\},\ \{0.6,0.8\},4]$, $n = 120$; $\alpha_1 = 0.4$, $\alpha_2 = -0.8$, $\beta_1 = 0.6$, $\beta_2 = 0.8$, $\sigma_Z^2 = 4$.

Number of iterations: 20.

The sum of absolute deviation between the ACF of the simulated data from the model d, and the ACF of ARMA(2,2)$[\{\alpha_1,\ \alpha_2\},\ \{\beta_1,\beta_2\}]$ at lags $j = 1,...,15$ is minimized over the stationary region for the ARMA(2,2) processes. The reported parameter estimates are the mean values in 20 repeats.

SACFvsACF output data: $\{AD = 0.596334,\ \ \hat{\alpha}_1 = 0.450708,\ \hat{\alpha}_2 = -0.791971\ \ \hat{\beta}_1 = 0.607394,\ \hat{\beta}_1 = 0.81144\}\}$

Ljung-Box Time Series Model Fit $ARMA(2,2):\ \ \tilde{\alpha}_1 = 0.484478,\ \ \tilde{\alpha}_2 = -0.739369\ \ \tilde{\beta}_1 = 0.655853,\ \ \tilde{\beta}_2 = 0.619243$

Ljung-Box p-values: $\{0.0774189, 0.103001, 0.132726, 0.176277, 0.192082\}$

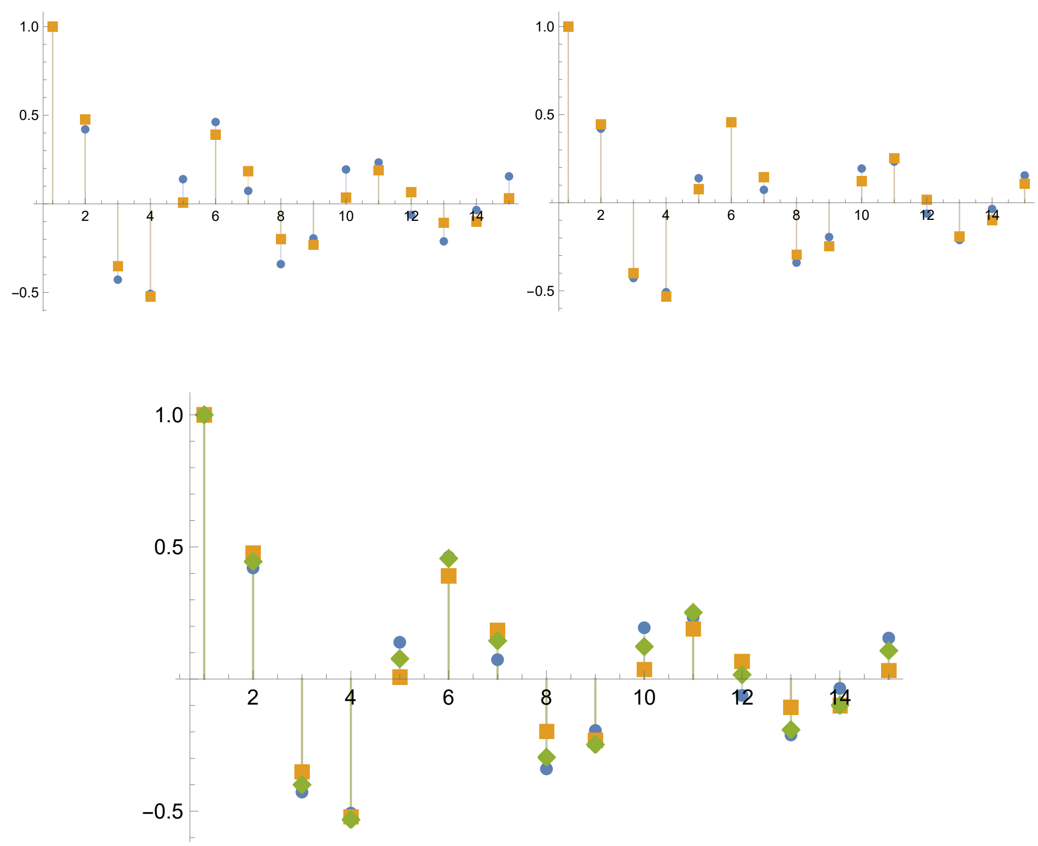

Figure 5: ACF Plots; Top Left: Data ACF in circles (blue color), Ljung-Box Fitted Model ACF in squares (orange color); Top Right: Data ACF in circles (blue color), SACFvsACF Fitted Model ACF in squares (orange color); Bottom: Data ACF in circles (blue color), Ljung-Box Fitted Model ACF in squares (orange color), SACFvsACF Fitted Model ACF in dimonds (green color).

6 Discussion.

In this work, (i): we developed a new practical method for the exact computation of the autocorrelation and the autocovariance functions of a stationary and causal discrete time univariate $ARMA(p,q)$ process. (ii): We presented a goodness of fit procedure for testing whether a time series follows an ARMA(p,q) model, using the exact values for the autocorrelation of the model. (iii): We provided new ARMA(p,q) parameters estimation method by minimizing the squared deviation between the sample and model ACFs. Investigating the statistical properties of the corresponding least squares parameters estimators, not covered here, is an interesting topic for future work. Also we did put light on the fact that the Ljung-Box ARMA(p,q) fitted model ACF might be far away from the series ACF. Also as mentioned by a reviewer, the Durbin Watson (DW) statistic for testing autocorrelation in the residuals from a statistical model or regression analysis, can be revisited by our proposed method developed in this article.

Acknowledgments. The author is indebted to the reviewers for providing constructive comments.

References

Basrak, B., R. A. Davis, and T. Mikosch (1999). The Sample ACF of a Simple Bilinear Process, Stoch. Process. Appl. 83 1-14.

Box, G. E. P. and D. A. Pierce, (1970). Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. J. American Statist. Asso. 65 (332) 1509–1526.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Second edition, Springer-Verlag, Berlin.

Chatfield, Chris (1999). *Introduction to Analysis of Time Series*. Fifth Edition, CRC Press, Boca Raton, USA; first edition (1975), Chapman & Hall, London.

Chien-Fu Wu (1981). Asymptotic theory of nonlinear least squares estimation. Ann. Statist., 9, 3, 501-513.

Khmaladze E. (1998). Goodness of fit tests for "Chimeric" alternatives. Statist. Neerlandica, 52, 1, 90-111.

Khmaladze E, and H. L. Koul (2009). Goodness-of-fit problem for errors in nonparametric regression: distribution free approach. Ann. Statist., 37, 6, 3165-3185.

Koul, H. L. (2002). *Weighted Empirical Processes in Dynamic Nonlinear Models*, 2nd Eds. Springer Lecture Notes in Statistics, 166, Berlin.

Ljung, G. M. and G. E. P. Box (1978). On a Measure of a Lack of Fit in Time Series Models. Biometrika. 65 (2) 297–303.

Pourahmadi, M. (2001). Foundations of Time Series Analysis and Prediction Theory. John Wiley, New York.

Pourahmadi, M. (2003). High-Dimensional Covariance Estimation. John Wiley, New York.

Richardson, G. D. and Bhattacharyya, B. B. (1986). Consistent estimators in nonlinear regression for a noncompact parameter spaces. Ann. Statist., 14, 4, 1591-1596.

Shumway, R. H. and Stoffer, D. S. (2017). *Time Series Analysis and Its Applications*. Fourth Edition, Springer-Verlag, New York, Berlin.

Soltani, A. R. (2019). A pivot function and its limiting distribution: applications in goodness of fit and testing hypothesis. Statistics, 53, 6, 1386-1395.

Mathematica, Version 12.1, (2020). Wolfram Research, Inc. Champaign, IL.

Tsay, R. S. and M. Pourahmadi (2017). Modeling structured correlation matrices. Biometrika, 104, 237-242.