Journal of Statistical Planning and Inference 197 (2018) 126–140

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

A goodness-of-fit test for VARMA(p, q) models

Santiago Velilla [a,*], Huong Nguyen Thu [b]

[a] Departamento de Estadística, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe, Madrid, Spain
[b] Division of Mathematical Statistics, Department of Mathematics, Stockholm University, SE - 106 91 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 5 May 2017
Received in revised form 25 October 2017
Accepted 3 January 2018
Available online 17 January 2018

MSC:
62M10

Keywords:
Adjusted residual traces
Brownian bridge
Residual correlation matrices

ABSTRACT

A goodness-of-fit approach for multivariate VARMA(p, q) models is presented. The idea is to consider a stochastic process based on a modified residual correlation matrix sequence, that is shown to converge to the Brownian bridge. Standard criteria based on this new random function, as for instance the Kolmogorov–Smirnov and Cramér–von Mises statistics, will have then a pivotal null asymptotic distribution. The properties of these two methods are investigated by simulation. As compared with the traditional methods in this area, their size does not depend critically on the choice of any lag parameter value, and they have better power properties.

1. Introduction

Consider a causal and invertible m-variate autoregressive moving average VARMA(p, q) process $\{\mathbf{X}_t\}$ of the form

$$\mathbf{\Phi}(B)(\mathbf{X}_t - \boldsymbol{\mu}) = \mathbf{\Theta}(B)\boldsymbol{\varepsilon}_t\,, \tag{1}$$

where $B\mathbf{X}_t = \mathbf{X}_{t-1}$ is the backward shift operator, and $\mathbf{\Phi}(B) = \mathbf{I}_m - \mathbf{\Phi}_1 B - \cdots - \mathbf{\Phi}_p B^p$ and $\mathbf{\Theta}(B) = \mathbf{I}_m + \mathbf{\Theta}_1 B + \cdots + \mathbf{\Theta}_q B^q$ are two polynomials that depend on $\mathbf{I}_m$, the identity of order m, and a collection $\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_p; \mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_q$ of $m \times m$ real matrices that are such that the roots of the determinantal equations $|\mathbf{\Phi}(z)| = 0$ and $|\mathbf{\Theta}(z)| = 0$ are different, and they all lie outside the unit circle. It will be also assumed that $\mathbf{\Phi}_p \neq \mathbf{0} \neq \mathbf{\Theta}_q$, and that the identifiability condition of Hannan (1969), $\mathrm{r}(\mathbf{\Phi}_p, \mathbf{\Theta}_q) = m$, holds. In expression (1), $\boldsymbol{\mu} = \mathsf{E}(\mathbf{X}_t)$ is the mean vector of the process, and $\{\boldsymbol{\varepsilon}_t\}$ a zero mean white noise sequence $\mathrm{WN}(\mathbf{0}, \ \mathbf{\Sigma})$ with covariance matrix $\mathbf{\Sigma} > 0$. In what follows, it will be convenient to define the $m \times mp$ matrix $\mathbf{\Phi} = (\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_p)$, the $m \times mq$ matrix $\mathbf{\Theta} = (\mathbf{\Theta}_1, \ldots, \mathbf{\Theta}_q)$, and the $m^2(p+q) \times 1$ vector of parameters $\mathbf{\Lambda} = \mathrm{vec}(\mathbf{\Phi}, \mathbf{\Theta})$.

Given n observations $\mathbf{X}_1, \ldots, \mathbf{X}_n$ from model (1), the mean vector $\boldsymbol{\mu}$ can be estimated by the average $\overline{\mathbf{X}}_n = n^{-1}\sum_{t=1}^{n}\mathbf{X}_t$. The remaining parameters $(\mathbf{\Phi}, \ \mathbf{\Theta}, \ \mathbf{\Sigma})$ can be estimated, following Lütkepohl (2005, sec. 12.2), by maximizing the Gaussian likelihood function. Once that the corresponding maximum likelihood (ML) estimates $\widehat{\mathbf{\Lambda}} = \mathrm{vec}(\widehat{\mathbf{\Phi}}; \widehat{\mathbf{\Theta}})$ of $\mathbf{\Lambda} = \mathrm{vec}(\mathbf{\Phi}; \mathbf{\Theta})$ have been determined, the $m \times 1$ residual vectors are obtained recursively in the form

$$\widehat{\boldsymbol{\varepsilon}}_t = (\mathbf{X}_t - \overline{\mathbf{X}}_n) - \sum_{i=1}^{p} \widehat{\mathbf{\Phi}}_i(\mathbf{X}_{t-i} - \overline{\mathbf{X}}_n) - \sum_{j=1}^{q} \widehat{\mathbf{\Theta}}_j \widehat{\boldsymbol{\varepsilon}}_{t-j}\,, t = 1, \ldots, n\,, \tag{2}$$

with the conditions $\mathbf{X}_t - \overline{\mathbf{X}}_n \equiv \mathbf{0} \equiv \widehat{\boldsymbol{\varepsilon}}_t$, $t \leq 0$. In practice, only residual vectors $\widehat{\boldsymbol{\varepsilon}}_t$ for $t > P = \max(p, q)$ are considered.

* Corresponding author.
E-mail addresses: santiago.velilla@uc3m.es (S. Velilla), thu@math.su.se (H.N. Thu).

https://doi.org/10.1016/j.jspi.2018.01.002
0378-3758/

This paper presents a new technique for testing the adequacy of a VARMA(p, q) model of the form (1). The motivating idea is to consider a suitable transformation of the sequence of $m \times m$ residual correlation matrices introduced by Chitturi (1974), $\widehat{\mathbf{R}}_k = \widehat{\mathbf{C}}_k' \widehat{\mathbf{C}}_0^{-1}$, $1 \le k \le n-(P+1)$, where

$$\widehat{\mathbf{C}}_k = \frac{1}{n} \sum_{t>P}^{n-k} \widehat{\boldsymbol{\varepsilon}}_t \widehat{\boldsymbol{\varepsilon}}_{t+k}' \,, 0 \le k \le n-(P+1). \tag{3}$$

A goodness-of-fit process is subsequently constructed, that is shown to converge to the Brownian bridge. The associated test statistics, constructed as goodness-of-fit functionals, will have then standard limit distributions that are free of unknown parameters. As an illustration, the properties of the Kolmogorov–Smirnov and Cramér–von Mises statistics are analyzed in some detailed by simulation. As compared with other standard methods, their size does not depend on the selection of any residual lag, and they are more powerful. Results obtained here generalize to a multivariate setting the approach of Ubierna and Velilla (2007) for univariate ARMA(p, q) models.

The rest of this paper is organized as follows. Section 2 establishes notation, and offers additional background and motivation. Sections 3 and 4 contain the main results. Section 5 presents examples of application and comparisons. Section 6 gives some concluding remarks. A final Appendix collects the proofs of all mathematical results.

2. Background and motivation

This section reviews briefly the usual goodness-of-fit methods for univariate and multivariate time series, and motivates the results of this paper.

2.1. Methods for univariate time series

Box and Pierce (1970) consider the residual autocorrelations $\{\widehat{r}_k\}$ of a univariate ARMA(p, q) model $\phi(B)X_t = \theta(B)\varepsilon_t$,

$$\widehat{r}_k = \frac{\sum_{t>P}^{n-k} \widehat{\varepsilon}_t \widehat{\varepsilon}_{t+k}}{\sum_{t>P}^{n} \widehat{\varepsilon}_t^2} \,, 1 \le k \le n-(P+1)\,, \tag{4}$$

where the $\{\widehat{\varepsilon}_t\}$ are the univariate versions of the quantities of (2). Assuming that $M = O(\sqrt{n})$, they establish the representation

$$\widehat{\mathbf{r}}_M = [\mathbf{I}_M - \mathbf{A}_M(\mathbf{A}_M'\mathbf{A}_M)^{-1}\mathbf{A}_M']\,\mathbf{r}_M + O_P(\frac{1}{n})\,, \tag{5}$$

where $\widehat{\mathbf{r}}_M = (\widehat{r}_1, \ldots, \widehat{r}_M)'$, $\mathbf{r}_M = (r_1, \ldots, r_M)'$ is a $M \times 1$ vector that depends on the error counterparts of the statistics of (4), and

$$\mathbf{A}_M = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ a_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{M-1} & a_{M-2} & \cdots & a_{M-(p+q)} \end{pmatrix} \tag{6}$$

is a $M \times (p+q)$ matrix in which the $\{a_r\}$ are the coefficients of the series $a(z) = [\phi(z)\theta(z)]^{-1} = [\theta(z)\phi(z)]^{-1} = \sum_{r=0}^{\infty} a_r z^r$, where $a_0 = 1$. Using (5), it is easy to see that the well-known portmanteau or Box and Pierce statistic,

$$Q_{\mathrm{BP}} = n \sum_{k=1}^{M} \widehat{r}_k^2 \,, \tag{7}$$

is approximately distributed under the null, for n large enough, as a $\chi^2_{M-(p+q)}$.

Modifications of Q_{BP} are introduced by Ljung and Box (1978) and Li and McLeod (1981), in order to improve finite sample size and power properties. Monti (1994) proposes a statistic similar to (7), but based on the residual partial autocorrelations. Peña and Rodríguez (2002) consider a test based on the determinant of the Toeplitz matrix of residual autocorrelations of order M. Li (2004, Chap. 2) gives a thorough review of methods. Francq et al. (2005) study goodness-of-fit beyond the independence assumption for the errors $\{\varepsilon_t\}$. Fisher and Gallaguer (2012) suggest a Weighted portmanteau statistic, whose null asymptotic distribution is a linear combination of M independent χ^2_1 random variables. The latter coincides with that of Peña and Rodríguez (2002). See also Gallaguer and Fisher (2014).

2.2. Multivariate time series

Chitturi (1974) generalizes the univariate residual correlations $\widehat{r}_k$ of (4) by considering the $m \times m$ matrices $\widehat{\mathbf{R}}_k = \widehat{\mathbf{C}}_k' \widehat{\mathbf{C}}_0^{-1}$, $1 \le k \le n-(P+1)$, where the $\widehat{\mathbf{C}}_k$ are as defined in expression (3), $0 \le k \le n-(P+1)$. The matrix $\widehat{\boldsymbol{\Sigma}} = \widehat{\mathbf{C}}_0$ is the ML estimate of the white noise covariance $\boldsymbol{\Sigma}$. Other possibilities for extending the $\{\widehat{r}_k\}$ are reviewed by Reinsel (1997, Chap. 5) and Li (2004, Chap. 3).