# A note on the distribution of residual autocorrelations in *VARMA(p,q)* models ☆

Huong Nguyen Thu[a,b]

[a] *Department of Business Administration, Technology and Social Sciences*
*Luleå University of Technology, 971 87 Luleå, Sweden*
[b] *Department of Mathematics*
*Foreign Trade University, Hanoi, Vietnam*

**Abstract**

This paper generalizes the distribution of residual autocovariance matrices in *VARMA(p,q)* models obtained previously in Hosking (1980). A new simplified version of the multivariate relation between sample correlation matrix of the errors and its residuals is also established. The modifications are effective tools for identifying and dealing with the curse of dimensionality in multivariate time series.

*Keywords:* Asymptotic distribution; Covariance matrix; Correlation matrix; Multivariate time series; *VARMA(p,q)* models.

## 1. Introduction

Consider a causal and invertible $m-$variate autoregressive moving average *VARMA(p,q)* process

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

where $B$ is backward shift operator $B\mathbf{X}_t = \mathbf{X}_{t-1}$. $\mu$ is the $m \times 1$ mean vector and $\{\varepsilon_t : t \in \mathbf{Z}\}$ is a zero mean white noise sequence $WN(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is a $m \times m$ positive definite matrix. Additionally, $\mathbf{\Phi}(z) = \mathbf{I}_m - \mathbf{\Phi}_1 z - \cdots - \mathbf{\Phi}_p z^p$ and $\mathbf{\Theta}(z) = \mathbf{I}_m + \mathbf{\Theta}_1 z + \cdots + \mathbf{\Theta}_q z^q$ are matrix polynomials, where $\mathbf{I}_m$ is the $m \times m$ identity matrix, and $\mathbf{\Phi}_1, \dots, \mathbf{\Phi}_p, \mathbf{\Theta}_1, \dots, \mathbf{\Theta}_q$ are $m \times m$ real matrices

*Corresponding author
*Email address:* huong.nguyen.thu@ltu.se (Huong Nguyen Thu)

such that the roots of the determinantal equations $|\mathbf{\Phi}(z)| = 0$ and $|\mathbf{\Theta}(z)| = 0$ all lie outside the unit circle. We also assume that both $\mathbf{\Phi}_p$ and $\mathbf{\Theta}_q$ are non-null matrices, and that the identifiability condition of Hannan [3], $r(\mathbf{\Phi}_p, \mathbf{\Theta}_q) = m$, holds.

Let $P = \max(p, q)$ and 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 $\mathbf{\Lambda} = \mathrm{vec}(\mathbf{\Phi}, \mathbf{\Theta})$. The residual vectors $\widehat{\varepsilon}_t$, $t = 1, \ldots, n$, are defined recursively, in the form

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

with the usual conditions $\mathbf{X}_t - \overline{\mathbf{X}}_n \equiv \mathbf{0} \equiv \widehat{\varepsilon}_t$, for $t \leq 0$. In practice, only residual vectors for $t > P = \max(p, q)$ are considered. Define $m \times m$ sample error covariance matrix at lag $k$ with the notation $\mathbf{C}_k = (1/n) \sum_{t=1}^{n-k} \varepsilon_t \varepsilon'_{t+k}$, $0 \leq k \leq n-1$. Similarly, the $m \times m$ $kth$ residual covariance matrix is given by $\widehat{\mathbf{C}}_k = (1/n) \sum_{t>P}^{n-k} \widehat{\varepsilon}_t \widehat{\varepsilon}'_{t+k}$ , $0 \leq k \leq n-(P+1)$. Following Chitturi [2], let $\mathbf{R}_k = \mathbf{C}'_k \mathbf{C}_0^{-1}$ be the $kth$ sample correlation matrix of the errors $\varepsilon_t$. Its residual analogue is given by $\widehat{\mathbf{R}}_k = \widehat{\mathbf{C}}'_k \widehat{\mathbf{C}}_0^{-1}$.

Properties of the residual covariance matrices and correlation matrices and their practical use in multivariate Portmanteau statistic have been considerd by many authors, see, for example, Chitturi [2], Hosking [4], Li and McLeod [6], Li and Hui [5]. The linear relation between the residual and error covariance matrices in Hosking [4] is highly dimensional. For convenience in applications, a new representation is suggested in section 2. To this end, this paper uses the family of statistics proposed in Velilla and Nguyen [8]. The orthogonal projection matrix in Hosking [4] is also generalized. The advantage of the new tool lies in standard properties of the trace operator. The associated correlation matrix of the residuals is useful in identification and diagnostic checking. Section 3 examines the large sample distribution of the transformed multivariate residual autocorrelations. Section 4 contains some final conclusions.

## 2. Distribution of residual autocovariace

This section examines the asymptotic properties of residual autocovariace. We start by considering the $m \times m$ coefficients of the series expansions $\mathbf{\Phi}^{-1}(z)\mathbf{\Theta}(z) = \sum_{j=0}^{\infty} \mathbf{\Omega}_j z^j$ and $\mathbf{\Theta}^{-1}(z) = \sum_{j=0}^{\infty} \mathbf{L}_j z^j$ where $\mathbf{\Omega}_0 = \mathbf{L}_0 = \mathbf{I}_m$. Define also the collection of matrices $\mathbf{G}_k = \sum_{j=0}^{k} (\mathbf{\Sigma}\mathbf{\Omega}'_j \otimes \mathbf{L}_{k-j})$ and $\mathbf{F}_k = \mathbf{\Sigma} \otimes \mathbf{L}_k$, $k \geq 0$. By convention, $\mathbf{G}_k = \mathbf{F}_k = \mathbf{0}$ for $k < 0$, it follows that

$$\widehat{\mathbf{C}}'_k = \mathbf{C}'_k - \sum_{i=1}^{p} \sum_{r=0}^{k-i} \mathbf{L}_{k-i-r} (\widehat{\mathbf{\Phi}}_i - \mathbf{\Phi}_i) \mathbf{\Omega}_r \mathbf{\Sigma} - \sum_{j=1}^{q} \mathbf{L}_{k-j} (\widehat{\mathbf{\Theta}}_j - \mathbf{\Theta}_j) \mathbf{\Sigma} + O_P(\frac{1}{n}) \ . \quad (3)$$

This construction is due to Hosking [4, p.603].

Set the sequence of $m^2M \times m^2(p+q)$ matrices $\mathbf{Z}_M = (\mathbf{X}_M, \mathbf{Y}_M)$, $M \geq 1$, where

$$\mathbf{X}_M = \begin{pmatrix} \mathbf{G}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{G}_1 & \mathbf{G}_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{G}_2 & \mathbf{G}_1 & \mathbf{G}_0 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{G}_{M-1} & \mathbf{G}_{M-2} & \mathbf{G}_{M-3} & \cdots & \mathbf{G}_{M-p} \end{pmatrix} ,$$

and

$$\mathbf{Y}_M = \begin{pmatrix} \mathbf{F}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{F}_1 & \mathbf{F}_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{F}_2 & \mathbf{F}_1 & \mathbf{F}_0 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{F}_{M-1} & \mathbf{F}_{M-2} & \mathbf{F}_{M-3} & \cdots & \mathbf{F}_{M-q} \end{pmatrix} .$$

Put the $Mm^2 \times Mm^2$ block diagonal matrix $\mathbf{W} = \mathrm{diag}(\mathbf{C}_0 \otimes \mathbf{C}_0, \overset{(M)}{\cdots}, \mathbf{C}_0 \otimes \mathbf{C}_0) = \mathbf{I}_M \otimes \mathbf{C}_0 \otimes \mathbf{C}_0$. The residual counterpart is $\widehat{\mathbf{W}} = \mathbf{I}_M \otimes \widehat{\mathbf{\Sigma}} \otimes \widehat{\mathbf{\Sigma}}$. For convenience, consider $Mm^2 \times 1$ random vectors $\widehat{\mathbf{H}}_M = [\mathrm{vec}(\widehat{\mathbf{C}}'_1), \ldots, \mathrm{vec}(\widehat{\mathbf{C}}'_M)]'$ and $\mathbf{H}_M = [\mathrm{vec}(\mathbf{C}'_1), \ldots, \mathrm{vec}(\mathbf{C}'_M)]'$ and $\mathcal{W} = \mathbf{I}_M \otimes \mathbf{\Sigma} \otimes \mathbf{\Sigma}$. After taking vecs in both sides of (3), hence, for each $M \geq 1$,

$$\widehat{\mathbf{H}}_M = \mathbf{H}_M - \mathbf{Z}_M \mathrm{vec}[(\widehat{\mathbf{\Phi}}, \widehat{\mathbf{\Theta}}) - (\mathbf{\Phi}, \mathbf{\Theta})] + O_P(\frac{1}{n}) \ . \quad (4)$$

According to Hosking [4], the below orthogonality condition holds.

$$\mathbf{Z}'_M \mathcal{W}^{-1} \widehat{\mathbf{H}}_M = O_P(1/n) \ . \tag{5}$$

Combining (4) with (5), Hosking [4] concluded that

$$\widehat{\mathbf{W}}^{-1/2} \widehat{\mathbf{H}}_M = (\mathbf{I}_{Mm^2} - \mathbf{P}_H) \mathbf{W}^{-1/2} \mathbf{H}_M + O_P(\frac{1}{n}) \ , \tag{6}$$

where $\mathbf{P}_H = \mathcal{W}^{-1/2} \mathbf{Z}_M (\mathbf{Z}'_M \mathcal{W}^{-1} \mathbf{Z}_M)^{-1} \mathbf{Z}'_M \mathcal{W}^{-1/2}$ is the $Mm^2 \times Mm^2$ orthogonal projection matrix onto the subspace spanned by the columns of $\mathcal{W}^{-1/2} \mathbf{Z}_M$.

In practice, the relation (6) is difficult to deal with as $m \geq 5$ because of the curse of dimensionaltity. As a solution, we define a $Mm^2 \times Mm^2$ matrix $\mathbf{Q}_M = (\mathbf{I}_M \otimes \mathbf{a})(\mathbf{I}_M \otimes \mathbf{a}')$, where $\mathbf{a} = \text{vec}(\mathbf{I}_m)/\sqrt{m}$. Put also $Mm^2 \times Mm^2$ matrix $\mathbf{P}_M = (\mathbf{I}_M \otimes \mathbf{a}') \mathcal{W}^{-1/2} \mathbf{Z}_{\scriptscriptstyle M} (\mathbf{Z}'_M \mathcal{W}^{-1/2} \mathbf{Q}_M \mathcal{W}^{-1/2} \mathbf{Z}_{\scriptscriptstyle M})^{-1} \mathbf{Z}'_M \mathcal{W}^{-1/2} (\mathbf{I}_M \otimes \mathbf{a})$. Note that $\mathbf{P}_M$ is the orthogonal projection matrix onto the subspace spanned by the columns of $(\mathbf{I}_M \otimes \mathbf{a}') \mathcal{W}^{-1/2} \mathbf{Z}_M$.

**Remark 2.1** The matrices $\mathbf{P}_H$ and $\mathbf{P}_M$ are idempotent.

Lemma 2.1 refers to the multivariate linear relation between the residual and error covariance matrices

**Lemma 2.1** Suppose that the error vectors $\{\varepsilon_t\}$ are i.i.d. with $\text{E}[\varepsilon_t] = \mathbf{0}$; $\text{Var}[\varepsilon_t] = \mathbf{\Sigma} > 0$; and finite fourth order moments $\text{E}[\|\varepsilon_t\|^4] < +\infty$. Then, as $n \longrightarrow \infty$,

$$\mathbf{Q}_M \widehat{\mathbf{W}}^{-1/2} \widehat{\mathbf{H}}_M = (\mathbf{I}_{Mm^2} - \mathbf{P}_M) \mathbf{W}^{-1/2} \mathbf{H}_M + O_P(\frac{1}{n}) \ . \tag{7}$$

**Proof 1** We begin by recalling the result in Velilla and Nguyen [8] that

$$\sqrt{n} \mathbf{H}_M \xrightarrow{D} \mathcal{W}^{1/2} [\mathbf{V}_1, \ldots, \mathbf{V}_M]' \ , \quad k \geq 1 \ , \tag{8}$$

where $\mathbf{V}_j$, $j = 1, \ldots, k$ are i.i.d. $\mathbf{N}_{m^2}(\mathbf{0}, \mathbf{I}_{m^2})$. See more details in Velilla and Nguyen [8, Lemma 3.1]. From (8), under the assumptions of Lemma 2.1, we obtain

$$\sqrt{n} \mathbf{Q}_M \mathbf{H}_M \xrightarrow{D} \mathbf{Q}_M \mathcal{W}^{1/2} [\mathbf{V}_1, \ldots, \mathbf{V}_M]' \ , \quad k \geq 1 \ . \tag{9}$$

Notice first that

$$\mathbf{Q}_M \widehat{\mathbf{W}}^{-1/2} \widehat{\mathbf{H}}_M = \mathbf{Q}_M \mathcal{W}^{-1/2} \widehat{\mathbf{H}}_M + \mathbf{Q}_M (\widehat{\mathbf{W}}^{-1/2} - \mathcal{W}^{-1/2}) \widehat{\mathbf{H}}_M \ . \tag{10}$$

As a consequence of (3), (4) and (9), the second summand of (10) is $O_P(1/n)$. We can rewrite (5) as

$$\mathbf{Q}_M \mathbf{Z}'_M \mathcal{W}^{-1} \widehat{\mathbf{H}}_M = O_P(\frac{1}{n}) \ . \tag{11}$$

Taking into account that $\widehat{\mathbf{C}}'_0$ is consistent for the matrix $\boldsymbol{\Sigma}$, hence $\mathbf{W}^{-1/2} - \mathcal{W}^{-1/2} = O_P(1/\sqrt{n})$. Those results combined with (4) and (9) lead to (7) and Lemma 2.1 follows. ■

Next section refers to a simplified multivariate linear relation between the residual autocorrelation matrices and the error ones.

## 3. Distribution of residual autocorrelation

Define $M \times 1$ random vectors $\widehat{\mathbf{T}}_M = [\mathrm{tr}(\widehat{\mathbf{R}}_1), \ldots, \mathrm{tr}(\widehat{\mathbf{R}}_M)]'$ and $\mathbf{T}_M = [\mathrm{tr}(\mathbf{R}_1), \ldots, \mathrm{tr}(\mathbf{R}_M)]'$. Theorem 3.1 provides a solution to the highly dimensional relation (7).

**Theorem 3.1** Under the assumptions of Lemma 2.1, as $n \longrightarrow \infty$,

$$\frac{1}{\sqrt{m}} \widehat{\mathbf{T}}_M = (\mathbf{I}_M - \mathbf{P}_M) \frac{1}{\sqrt{m}} \mathbf{T}_M + O_P(\frac{1}{n}) \ . \tag{12}$$

**Proof 2** We introduce the notion of a linear relationship for dimension-reduction purpose, following Velilla and Nguyen [8],

$$\frac{1}{\sqrt{m}} \mathbf{T}_M = (\mathbf{I}_M \otimes \mathbf{a}') \mathbf{W}^{-1/2} \mathbf{H}_M \ , \quad M \geq 1 \ . \tag{13}$$

Its residual version is given by

$$\frac{1}{\sqrt{m}} \widehat{\mathbf{T}}_M = (\mathbf{I}_M \otimes \mathbf{a}') \widehat{\mathbf{W}}^{-1/2} \widehat{\mathbf{H}}_M \ , \quad M \geq 1 \ . \tag{14}$$

Combining (7), (13) and (14) finishes the proof of (12).■

The objective now is to establish the large sample distribution of the random vector $\widehat{\mathbf{H}}_M$.

**Theorem 3.2** Under the assumptions of Lemma 2.1, as $n \longrightarrow \infty$,

$$\sqrt{n}\mathbf{Q}_M \widehat{\mathbf{W}}^{-1/2} \widehat{\mathbf{H}}_M \overset{D}{\cong} (\mathbf{I}_{Mm^2} - \mathbf{P}_M)\mathbf{N}_{Mm^2}(\mathbf{0}, \mathbf{I}_{Mm^2}) \ . \tag{15}$$

**Proof 3** The proof is completed by using (7) and (9).■

Let us mention an important consequence of the Theorem 3.2. It is a practical tool to construct a goodness of fit process for selecting a proper time series model.

## 4. Conclusions

The main results of this paper were announced in Lemma 2.1, Theorems 3.1 and 3.2. Goodness of fit methods make those properties more attractive in practice especially for large time series datasets.


## 5. Acknowledgment

This work was carried out while the author was a PhD student at Department of Statistics, Universidad Carlos III de Madrid. I am very grateful to Prof. Santiago Velilla for his sharing, encouragement and guidance as my adviser.


## References


[1] , 2007. A goodness-of-fit process for arma models based on a modified residual autocorrelation sequence. Journal of Statistical Planning and Inference 137 (9), 2903 – 2919.

[2] Chitturi, R. V., 1974. Distribution of residual autocorrelations in multiple autoregressive schemes. Journal of the American Statistical Association 69 (348), 928–934.

[3] Hannan, E. J., 1969. The identification of vector mixed autoregressive-moving average systems. Biometrika 56 (1), pp. 223–225.

[4] Hosking, J. R. M., 1980. The multivariate portmanteau statistic. Journal of the American Statistical Association 75 (371), pp. 602–608.

[5] Li, W. K., Hui, Y. V., 1989. Robust multiple time series modelling. Biometrika 76 (2), pp. 309–315.

[6] Li, W. K., McLeod, A. I., 1981. Distribution of the residual autocorrelations in multivariate arma time series models. Journal of the Royal Statistical Society. Series B (Methodological) 43 (2), pp. 231–239.

[7] Velilla, S., 1994. A goodness-of-fit test for autoregressive moving-average models based on the standardized sample spectral distribution of the residuals. Journal of Time Series Analysis 15, 637–647.

[8] Velilla, S., Nguyen, H., 2011. A basic goodness-of-fit process for varma(p,q) models. Statistics and econometrics series 09, Carlos III University of Madrid.