J. R. Stat. Soc. B (1981),
43, *No.* 2, *pp.* 231–239

Distribution of the Residual Autocorrelations in Multivariate ARMA Time Series Models

By W. K. Li and A. I. McLeod

University of Western Ontario, London, Ontario, Canada N6A 5B9

[Received May 1980. Revised November 1980]

SUMMARY

The large-sample distribution of the multivariate residual autocorrelations in the vector ARMA model is derived. This result is somewhat less complicated for the vector autoregressive model. A new multivariate portmanteau test for checking the adequacy of fitted vector ARMA models is developed. A simulation study shows that a simple modification of the portmanteau test improves its accuracy in small samples.

Keywords: CHECKING MODEL ADEQUACY; MULTIVARIATE RESIDUAL AUTOCORRELATION; MULTIVARIATE TIME SERIES; PORTMANTEAU TEST ; VECTOR AUTOREGRESSION; VECTOR ARMA MODEL

1. Introduction

The recent paper of Tiao and Box (1979) describes a practical approach to vector ARMA modelling of multivariate time series data by following the three stages of statistical model development: identification, estimation and model criticism. At the model criticism stage, Tiao and Box (1979) suggest using the multivariate residual autocorrelations to check the adequacy of a fitted vector ARMA model. In this paper, the large sample distribution of the multivariate residual autocorrelations in vector ARMA models is derived and a new multivariate generalization of the univariate portmanteau test of Box and Jenkins (1970) and Box and Pierce (1970) is given. Hosking (1981) gives an alternative but somewhat less general and less direct derivation of these results using an hypothesis testing technique. Previously, Chitturi (1974) and Hosking (1980) have obtained another multivariate generalization of the portmanteau test but their approaches use a less standard definition of multivariate autocorrelations.

2. Notation and Lemmas

The multivariate zero-mean ARMA model for a k-dimensional time series $\mathbf{Z}_t^{\mathrm{T}} = (Z_{1,t}, \ldots, Z_{k,t})$ can be written

$$\boldsymbol{\phi}(B)\,\mathbf{Z}_t = \boldsymbol{\theta}(B)\,\mathbf{a}_t, \tag{1}$$

where

$$\boldsymbol{\phi}(B) = \mathbf{1}_k - \boldsymbol{\phi}_1 B - \ldots - \boldsymbol{\phi}_p B^p,$$

$$\boldsymbol{\theta}(B) = \mathbf{1}_k - \boldsymbol{\theta}_1 B - \ldots - \boldsymbol{\theta}_q B^q,$$

where

$\mathbf{1}_k$ is the $k \times k$ identity matrix,

$$\boldsymbol{\phi}_l = (\phi_{ij,\,l})_{k \times k},$$

$$\boldsymbol{\theta}_l = (\theta_{ij,\,l})_{k \times k},$$

and B is the backshift operator, $B\mathbf{Z}_t = \mathbf{Z}_{t-1}$ and $B\mathbf{a}_t = \mathbf{a}_{t-1}$.

The innovation series $\mathbf{a}_t$ is a sequence of independent and identically distributed vector random variables with $\langle \mathbf{a}_t \rangle = \mathbf{0}$ and

$$\operatorname{var}(\mathbf{a}_t) = \boldsymbol{\Delta} = (\sigma_{gh})_{k \times k},$$

where $\langle \cdot \rangle$ denotes mathematical expectation and $\boldsymbol{\Delta}$ is positive definite.

The (g,h) element of $\boldsymbol{\Delta}^{-1}$ will be denoted by σ^{gh}. It is assumed that the fourth moment of $a_{i,t}$ exists for $i = 1,2,\ldots,k$ and that the model is stationary and invertible so that for all $|B| \leqslant 1$,

$$\det \boldsymbol{\phi}(B) \neq \mathbf{0} \quad \text{and} \quad \det \boldsymbol{\theta}(B) \neq \mathbf{0}.$$

Consequently, inverses exist and can be written

$$\boldsymbol{\phi}(B)^{-1} = \mathbf{1}_k + \boldsymbol{\phi}'_1 B + \boldsymbol{\phi}'_2 B^2 + \ldots, \tag{2}$$

and

$$\boldsymbol{\theta}(B)^{-1} = \mathbf{1}_k + \boldsymbol{\theta}'_1 B + \boldsymbol{\theta}'_2 B^2 + \ldots.$$

It is further assumed that the model is identifiable so that the large sample information matrix is non-singular. Conditions for identifiability are discussed in Hannan (1969).

Let $\boldsymbol{\beta} = (\operatorname{vec} \boldsymbol{\phi}_1^{\mathrm{T}}, \ldots, \operatorname{vec} \boldsymbol{\phi}_p^{\mathrm{T}}, \operatorname{vec} \boldsymbol{\theta}_1^{\mathrm{T}}, \ldots, \operatorname{vec} \boldsymbol{\theta}_q^{\mathrm{T}})$ where the vec of a $k \times k$ matrix is a vector with the (i,j) matrix element in the $(j-1)k+i$ position.

For any $\dot{\boldsymbol{\beta}}$, and $n \geqslant t \geqslant p+1$, let

$$\dot{\mathbf{a}}_t = \mathbf{Z}_t - \dot{\boldsymbol{\phi}}_1 \mathbf{Z}_{t-1} - \ldots - \dot{\boldsymbol{\phi}}_p \mathbf{Z}_{t-p} + \dot{\boldsymbol{\theta}}_1 \dot{\mathbf{a}}_{t-1} + \ldots + \dot{\boldsymbol{\theta}}_q \dot{\mathbf{a}}_{t-q},$$

where

$$\dot{\mathbf{a}}_t = \mathbf{0} \quad \text{for } t \leqslant p.$$

The corresponding residual autocorrelation matrices are defined by

$$\dot{\mathbf{R}}_l = (\dot{r}_{ij}(l))_{k \times k} \quad (l \geqslant 0), \tag{3}$$

where

$$\dot{r}_{ij}(l) = \dot{c}_{ij}(l) / \sqrt{(\dot{c}_{ii}(0)\, \dot{c}_{jj}(0))} \quad (1 \leqslant i, j \leqslant k),$$

where

$$\dot{c}_{ij}(l) = \frac{1}{n} \sum_{t=l+1}^{n} \dot{a}_{i,t}\, \dot{a}_{j,t-l}. \tag{4}$$

Also, let

$$\dot{\mathbf{r}} = (\operatorname{vec} \dot{\mathbf{R}}_1^{\mathrm{T}}, \operatorname{vec} \dot{\mathbf{R}}_2^{\mathrm{T}}, \ldots, \operatorname{vec} \dot{\mathbf{R}}_m^{\mathrm{T}}).$$

Let $\hat{\boldsymbol{\beta}}$ be an asymptotically efficient estimate of $\boldsymbol{\beta}$ and let $\hat{a}_{i,t}$ and $\hat{\mathbf{R}}_l$ be the corresponding residuals and residual autocorrelations. Similarly $a_{i,t}$ and $\mathbf{R}_l$ are the residuals and residual autocorrelations corresponding to β.

Lemma 1. The distribution of $\hat{\mathbf{r}}$ does not depend on (σ_{ii}), $i = 1, \ldots, k$.

Proof. This follows from the fact that $\dot{r}_{ij}(l)$ is invariant under arbitrary scale transformations on each of the components of the $\mathbf{Z}_t$-series.

From Lemma 1, it can be assumed without loss of generality that $\boldsymbol{\Delta}$ is in correlation form, so that $\sigma_{ii} = 1$, $i = 1, \ldots, k$.

Lemma 2. $\sqrt{n} \cdot \mathbf{r}$ is asymptotically normal with mean $\mathbf{0}$ and covariance matrix

$$\mathbf{Y} = \mathbf{1}_m \otimes (\boldsymbol{\Delta} \otimes \boldsymbol{\Delta}) \tag{5}$$

Proof. Consider the covariance matrix of $\sqrt{n} \operatorname{vec} \mathbf{R}_l^{\mathrm{T}}$, $l > 0$. If this covariance matrix is partitioned into k^2 $k \times k$ blocks, the (i,j) element in the (g,h) block is neglecting quantities of $\mathrm{O}p(1/n)$

$$\begin{aligned} n \cdot \operatorname{cov}(r_{gi}(l), r_{hj}(l)) &= \sum_t \sum_s \langle a_{g,t}\, a_{i,t-l}\, a_{h,s}\, a_{j,s-l} \rangle / n \\ &= \sigma_{ij}\, \sigma_{gh}. \end{aligned} \tag{6}$$

The last line follows from a well-known fourth moment result (Hannan, 1970, p. 23). Similarly, it can be shown that $\sqrt{n}\cdot\operatorname{vec} R_l^{\mathrm{T}}$, and $\sqrt{n}\cdot\operatorname{vec} R_{l'}^{\mathrm{T}}$ are asymptotically uncorrelated when $l\neq l'$. Asymptotic normality is established by the martingale central limit theroem as in McLeod (1978, 1979).

3. THE AUTOREGRESSIVE CASE

For any $\dot{\boldsymbol{\beta}}$ define the auxiliary series $\dot{\mathbf{V}}_t$ by

$$\boldsymbol{\phi}(B)\,\dot{\mathbf{V}}_t = \dot{\mathbf{a}}_t. \tag{7}$$

Then we have

$$-\partial\dot{\mathbf{a}}_t/\partial\dot{\phi}_{ij,l} = \mathbf{D}_{ij}\,\dot{\mathbf{V}}_{t-l}, \tag{8}$$

where $\mathbf{D}_{ij}$ is a $k\times k$ matrix with 1 at position (i,j) and 0 elsewhere.

Given n data points $\mathbf{Z}_1,\mathbf{Z}_2,\ldots,\mathbf{Z}_n$, let $\dot{\boldsymbol{\beta}}$ satisfy the invertibility restriction. Then under assumptions of normality the log-likelihood of $(\dot{\boldsymbol{\beta}},\dot{\boldsymbol{\Delta}})$ is then (Wilson, 1973)

$$L(\dot{\boldsymbol{\beta}},\dot{\boldsymbol{\Delta}}) = -n/2\log|\dot{\boldsymbol{\Delta}}| - \frac{1}{2}\sum_{t=1}^{n}\dot{\mathbf{a}}_t^{\mathrm{T}}\,\dot{\boldsymbol{\Delta}}^{-1}\,\dot{\mathbf{a}}_t. \tag{9}$$

Algorithms for obtaining maximum likelihood estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\Delta}}$ are given by Wilson (1973), Nicholls and Hall (1979) and Tiao and Hillmer (1979). Let

$$S = \sum_{t=1}^{n}\dot{\mathbf{a}}_t^{\mathrm{T}}\,\dot{\boldsymbol{\Delta}}^{-1}\,\dot{\mathbf{a}}_t. \tag{10}$$

Let $\hat{\boldsymbol{\beta}}$ be the least square estimate of $\boldsymbol{\beta}$ by maximizing (9) then $\sqrt{n}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ is asymptotically normal with mean $\mathbf{0}$ and covariance matrix $\mathbf{I}^{-1}$ (Wilson, 1973) where

$$\mathbf{I} = \lim_{n\to\infty}\left\langle\frac{\partial^2 S}{\partial\boldsymbol{\beta}\,\partial\boldsymbol{\beta}^{\mathrm{T}}}\right\rangle\Big/(2n). \tag{11}$$

Lemma 3.

$$\hat{\boldsymbol{\beta}}-\boldsymbol{\beta} = \mathbf{I}^{-1}\mathbf{S}^c + \mathrm{O}p(1/n),$$

where

$$\mathbf{S}^c = -(2n)^{-1}\,\partial S/\partial\boldsymbol{\beta}.$$

Proof. Expanding $\partial S/\partial\boldsymbol{\beta}$ using Taylor's series and evaluating at $\hat{\boldsymbol{\beta}}$ gives

$$\mathbf{0} = \frac{\partial S}{\partial\boldsymbol{\beta}} + \frac{\partial^2 S}{\partial\boldsymbol{\beta}\,\partial\boldsymbol{\beta}^{\mathrm{T}}}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}) + \mathrm{O}p(1). \tag{12}$$

As in McLeod (1978) it can be shown that

$$\frac{1}{2}\frac{\partial^2 S}{\partial\boldsymbol{\beta}\,\partial\boldsymbol{\beta}^{\mathrm{T}}} = n\mathbf{I} + \mathrm{O}p(\sqrt{n}). \tag{13}$$

Lemma 3 now follows by inverting $\partial^2 S/(\partial\boldsymbol{\beta}\,\partial\boldsymbol{\beta}^{\mathrm{T}})$. Note that

$$\begin{aligned} S^c_{ij,l} &= \frac{-1}{(2n)}\partial S/\partial\phi_{ij,l} \\ &= \frac{-1}{n}\sum_t \mathbf{a}_t^{\mathrm{T}}\,\boldsymbol{\Delta}^{-1}\frac{\partial\mathbf{a}_t}{\partial\phi_{ij,l}} \\ &= \frac{1}{n}\sum_t\sum_{f=1}^{k} a_{f,t}\,\sigma^{fi}\,V_{j,t-l}. \end{aligned} \tag{14}$$

Thus, the cross covariance of $\sqrt{n}\cdot(S^c_{ij,l})$ and $\sqrt{n}\cdot c_{gh}(l')$ is given by

$$\begin{aligned} n\cdot\langle S^c_{ij,l}\,c_{gh}(l')\rangle &= \frac{1}{n}\langle(\sum_t V_{j,t-l}\sum_f a_{f,t}\,\sigma^{fi})(\sum_s a_{g,s}\,a_{h,s-l})\rangle \\ &= \frac{1}{n}\sum_{f=1}^{k}\sum_t\sum_s \sigma^{fi}\{\langle V_{j,t-l}\,a_{f,t}\rangle\langle a_{g,s}\,a_{h,s-l'}\rangle+\langle V_{j,t-l}\,a_{g,s}\rangle \\ &\qquad\times\langle a_{f,t}\,a_{h,s-l'}\rangle+\langle V_{j,t-l}\,a_{h,s-l'}\rangle\langle a_{f,t}\,a_{g,s}\rangle\}\quad (l'\geqslant l), \\ &= \frac{1}{n}\sum_t\langle V_{j,t-l}\,a_{h,t-l'}\rangle\quad (g=i \text{ and } l'\geqslant l), \\ &= 0,\quad \text{otherwise.} \end{aligned}\tag{15}$$

The last equality follows from the fact that

$$\mathbf{V}_t=\mathbf{a}_t+\boldsymbol{\phi}'_1\,\mathbf{a}_{t-1}+\boldsymbol{\phi}'_2\,\mathbf{a}_{t-2}+\ldots.$$

Let $\mathbf{X}^{\mathrm{T}}$ be the asymptotic cross covariance matrix of $\mathbf{S}^c$ and $\mathbf{r}$. Then

$$\mathbf{X}^{\mathrm{T}}=\begin{pmatrix}\mathbf{A}_{11} & \mathbf{A}_{12} & \ldots & \mathbf{A}_{1m}\\ \vdots & & & \vdots\\ \mathbf{A}_{p1} & \ldots & & \mathbf{A}_{pm}\end{pmatrix},\tag{16}$$

where

$$\begin{aligned}\mathbf{A}_{ij}&=\begin{pmatrix}\boldsymbol{\phi}'_{j-i}\boldsymbol{\Delta}, & \mathbf{0}, & \mathbf{0}, & \\ \mathbf{0}, & \boldsymbol{\phi}'_{j-i}\boldsymbol{\Delta}, & \mathbf{0}, & \ddots \quad \boldsymbol{\phi}'_{j-i}\boldsymbol{\Delta}\end{pmatrix}\\ &=\mathbf{1}_k\otimes\boldsymbol{\phi}'_{j-i}\boldsymbol{\Delta}\quad (j\geqslant i)\\ &=\mathbf{0},\qquad \text{otherwise.}\end{aligned}$$

Theorem 1. The asymptotic joint distribution of $\sqrt{n}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ and $\sqrt{n}\cdot\mathbf{r}=\sqrt{n}\cdot(\operatorname{vec}\mathbf{R}^{\mathrm{T}}_1,\ldots,\operatorname{vec}\mathbf{R}^{\mathrm{T}}_m)$ is normal with mean $\mathbf{0}$ and covariance matrix

$$\left[\begin{array}{c|c}\mathbf{I}^{-1} & \mathbf{I}^{-1}\mathbf{X}^{\mathrm{T}}\\ \hline \mathbf{X}\mathbf{I}^{-1} & \mathbf{Y}\end{array}\right]_{k^2(p+m)\times k^2(p+m)},\tag{18}$$

where $\mathbf{Y}$ is given by equation (5).

Proof. Asymptotic normality follows from Lemma 3 and the martingale limit theorem (Billingsley, 1961) as in McLeod (1978, 1979) while Lemma 2 and the preceding discussions give us the covariance matrix.

Theorem 2. $\sqrt{n}\cdot\hat{\mathbf{r}}$ is asymptotically normal with mean $\mathbf{0}$ and covariance matrix $(\mathbf{Y}-\mathbf{X}\mathbf{I}^{-1}\mathbf{X}^{\mathrm{T}})$.

Proof. Expanding $\mathbf{R}_l$ using a Taylor's series expansion about $(\boldsymbol{\beta},\boldsymbol{\Delta})$, evaluating at $(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\Delta}})$ and using the fact that $\partial\dot{\mathbf{a}}_t/\partial\dot{\sigma}_{ij}=\mathbf{0}$ gives

$$\hat{\mathbf{r}}=\mathbf{r}-\mathbf{X}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})+\mathrm{O}p(1/n).$$

Theorem 2 follows immediately.

Since $\boldsymbol{\Delta}$ is positive definite and symmetric there exists a non-singular matrix $\mathscr{G}$ such that

$$\mathscr{G}\mathscr{G}^{\mathrm{T}}=\boldsymbol{\Delta}^{-1},\tag{19}$$

$$\mathscr{G}\boldsymbol{\Delta}\mathscr{G}^{\mathrm{T}}=\mathbf{1}_k.$$

Let

$$\mathbf{G} = \mathbf{1}_m \otimes (\mathscr{G} \otimes \mathscr{G}^{\mathrm{T}}) \tag{20}$$

and

$$\tilde{\mathbf{r}} = \mathbf{G}\hat{\mathbf{r}}.$$

By Theorem 2 $\sqrt{n}\cdot\tilde{\mathbf{r}}$ is asymptotically normal with mean $\mathbf{0}$ and covariance matrix $\mathbf{Q}$ where

$$\mathbf{Q} = \mathbf{1}_{mk^2} - \mathbf{G}\mathbf{X}\mathbf{I}^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}. \tag{21}$$

Theorem 3. $\mathbf{X}^{\mathrm{T}}(\mathbf{G}\mathbf{G}^{\mathrm{T}})\mathbf{X} \cong \mathbf{I}$ provided m is large enough so that $\boldsymbol{\phi}'_i \cong \mathbf{0}$ for $i > m$.

Proof. From (14)

$$\begin{aligned}\frac{1}{2}\frac{\partial^2 S}{\partial\phi_{i'j',l'}\,\partial\phi_{ij,l}} &= \sum_t \mathbf{V}_{t-l'}^{\mathrm{T}}\,\mathbf{D}_{i'j'}^{\mathrm{T}}\begin{pmatrix}\sigma^{1i}\\ \vdots \\ \sigma^{ki}\end{pmatrix} V_{j,t-l}\\ &= \sum_t V_{j',t-l'}\,V_{j,t-l}\,\sigma^{i'i}\\ &= \sigma^{i'i}\sum_t V_{j',t-l'}\,V_{j,t-l}.\end{aligned} \tag{22}$$

Taking expectations,

$$\left\langle\frac{\partial^2 S}{\partial\phi_{i'j',l'}\,\partial\phi_{ij,l}}\right\rangle \Big/ (2n) = \sigma^{i'i}\,\gamma_{vv,j'j}(l-l'),$$

where

$$\gamma_{vv,j'j}(l-l') = \langle \mathbf{V}_{t-l'}\,\mathbf{V}_{t-l}^{\mathrm{T}}\rangle_{j'j}.$$

Hence

$$\left\langle\frac{\partial^2 S}{\partial\boldsymbol{\beta}\,\partial\boldsymbol{\beta}^{\mathrm{T}}}\right\rangle \Big/ (2n) = [\boldsymbol{\Delta}^{-1} \otimes \boldsymbol{\gamma}_{vv}(l-l')]. \tag{23}$$

By (20)

$$\mathbf{X}^{\mathrm{T}}(\mathbf{G}\mathbf{G}^{\mathrm{T}}) = \begin{bmatrix}\mathbf{B}_{11} & \cdots & \mathbf{B}_{1m}\\ \vdots & & \vdots\\ \mathbf{B}_{p1} & \cdots & \mathbf{B}_{pm}\end{bmatrix}, \tag{24}$$

where $\mathbf{B}_{ij}$ are $k^2 \times k^2$ matrices of the form

$$\boldsymbol{\Delta}^{-1} \otimes \boldsymbol{\phi}'_{j-i} = \begin{bmatrix}\sigma^{11}\boldsymbol{\phi}'_{j-i} & \cdots & \sigma^{1k}\boldsymbol{\phi}'_{j-i}\\ \vdots & & \vdots\\ \sigma^{1k}\boldsymbol{\phi}'_{j-i} & \cdots & \sigma^{kk}\boldsymbol{\phi}'_{j-i}\end{bmatrix},$$

$(j \geqslant i)$ and $\mathbf{0}$ $(j < i)$. Hence,

$$\mathbf{X}^{\mathrm{T}}(\mathbf{G}\mathbf{G}^{\mathrm{T}})\mathbf{X} = \begin{bmatrix}\mathbf{C}_{11} & \cdots & \mathbf{C}_{1p}\\ \vdots & & \vdots\\ \mathbf{C}_{p1} & & \mathbf{C}_{pp}\end{bmatrix}, \tag{25}$$

where $\mathbf{C}_{l'l}$ are $k^2 \times k^2$ matrices of the form

$$\boldsymbol{\Delta}^{-1} \otimes \sum_{u=l}^{m} \boldsymbol{\phi}'_{u-l'}\,\boldsymbol{\Delta}\boldsymbol{\phi}'^{\mathrm{T}}_{u-l} \quad (\geqslant l'),$$

$$\boldsymbol{\Delta}^{-1} \otimes \sum_{u=l'}^{m} \boldsymbol{\phi}'_{u-l'}\,\boldsymbol{\Delta}\boldsymbol{\phi}'^{\mathrm{T}}_{u-l} \quad (l < l').$$

Now

$$\begin{aligned}\boldsymbol{\gamma}_{vv}(l-l') &= \langle \mathbf{z}_{t-l'}\,\mathbf{z}_{t-l}^{\mathrm{T}}\rangle\\ &= \langle(\boldsymbol{\phi}^{-1}(B)\,\mathbf{a}_{t-l'})(\boldsymbol{\phi}^{-1}(B)\,\mathbf{a}_{t-l})^{\mathrm{T}}\rangle.\end{aligned}$$

For m large enough,

$$\gamma_{vv}(l-l') \cong \sum_{t=0}^{m} \boldsymbol{\phi}'_{t+s}\, \boldsymbol{\Delta}\boldsymbol{\phi}'^{\mathrm{T}}_{t} \quad (\text{if } l \geqslant l' \text{ and } s = l-l'),$$

$$\cong \sum_{t=0}^{m} \boldsymbol{\phi}'_{t}\, \boldsymbol{\Delta}\boldsymbol{\phi}'^{\mathrm{T}}_{t+s} \quad (\text{if } l' > l \text{ and } s = l'-l).$$

Hence, by comparing (23) and (25) for m sufficiently large so that $\boldsymbol{\phi}'_s \cong \mathbf{0}$ for $s > m$, we have $\mathbf{X}^{\mathrm{T}}(\mathbf{G}\mathbf{G}^{\mathrm{T}})\mathbf{X} \cong \mathbf{I}$.

It follows from Theorem 3 that $\mathbf{Q}$ is very nearly idempotent with rank $k^2(m-p)$ provided that m is large enough so that $\boldsymbol{\phi}'_i \cong \mathbf{0}$ for $i > m$. And hence the statistic

$$Q_m = n \cdot \sum_{i=1}^{m} \hat{\mathbf{r}}(i)^{\mathrm{T}}(\hat{\mathbf{R}}_0^{-1} \otimes \hat{\mathbf{R}}_0^{-1})\,\hat{\mathbf{r}}(i), \tag{26}$$

where

$$\hat{\mathbf{r}}(i) = \operatorname{vec} \hat{\mathbf{R}}_i^{\mathrm{T}}$$

is approximately χ^2-distributed with $k^2(m-p)$ DF for suitable m and n. This result provides a new generalization of the portmanteau test of Box and Pierce (1970) for checking model adequacy.

4. THE ARMA CASE

The method of Section 3 directly extends to the ARMA model case. However, the derivation is more involved since the partials $\partial \dot{\mathbf{a}}_t / \partial \dot{\phi}_{\alpha\beta,l}$ and $\partial \dot{\mathbf{a}}_t / \partial \dot{\theta}_{\alpha\beta,l}$ are much more complicated (unlike the corresponding situation in the univariate case). Let

$$\dot{\mathbf{a}}_t = \dot{\boldsymbol{\theta}}^{-1}(B)\dot{\boldsymbol{\phi}}(B)\,\mathbf{Z}_t \tag{27}$$

then

$$\partial \dot{\mathbf{a}}_t / \partial \dot{\theta}_{\alpha\beta,l} = \dot{\boldsymbol{\theta}}^{-1}(B)\,\mathbf{D}_{\alpha\beta}\,\dot{\mathbf{a}}_{t-l} \tag{28}$$

and

$$-\partial \dot{\mathbf{a}}_t / \partial \dot{\phi}_{\alpha\beta,l} = \dot{\boldsymbol{\theta}}^{-1}(B)\,\mathbf{D}_{\alpha\beta}\,\boldsymbol{\phi}^{-1}(B)\,\dot{\boldsymbol{\theta}}(B)\,\dot{\mathbf{a}}_{t-l}. \tag{29}$$

Define

$$\begin{aligned} \boldsymbol{\phi}^{\alpha\beta}(B) &= \dot{\boldsymbol{\theta}}^{-1}(B)\,\mathbf{D}_{\alpha\beta}\,\boldsymbol{\phi}^{-1}(B)\,\dot{\boldsymbol{\theta}}(B), \\ \dot{\boldsymbol{\theta}}^{\alpha\beta}(B) &= \dot{\boldsymbol{\theta}}^{-1}(B)\,\mathbf{D}_{\alpha\beta}. \end{aligned} \tag{30}$$

Let

$$\boldsymbol{\phi}^{\alpha\beta}_s(\boldsymbol{\theta}^{\alpha\beta}_s)$$

be the coefficient of B^s in $\boldsymbol{\phi}^{\alpha\beta}(B)\,(\boldsymbol{\theta}^{\alpha\beta}(B))$.
Let

$$-n \cdot S^{\psi}_{\alpha\beta,l} = \sum_t \mathbf{a}_t^{\mathrm{T}}\,\boldsymbol{\Delta}^{-1}\frac{\partial \mathbf{a}_t}{\partial \psi_{\alpha\beta,l}}, \tag{31}$$

where ψ corresponds to etiher θ or ϕ and let

$$\begin{aligned} -\mathbf{V}^{\alpha\beta}_{s-l} &= \partial \mathbf{a}_s / \partial \phi_{\alpha\beta,l}, \\ \overset{\alpha\beta}{\mathbf{U}}_{s-l}\ & \partial \mathbf{a}_s / \partial \theta_{\alpha\beta,l}. \end{aligned} \tag{32}$$

As in (15) it can be shown that the covariance of $\sqrt{n} \cdot S^{\psi}_{\alpha\beta,l}$ and $\sqrt{n} \cdot c_{gh}(l')$ is equal to

$$\frac{1}{n} \cdot \sum \langle \mathbf{W}^{\alpha\beta}_{t-l}\,\mathbf{a}^{\mathrm{T}}_{t-l'} \rangle_{gh} = (\boldsymbol{\psi}^{\alpha\beta}_s\,\boldsymbol{\Delta})_{gh} \quad (s = l'-l), \tag{33}$$

where W corresponds to V or U whenever

$$\psi = \phi \text{ or } \theta.$$

Let the (g, h) element of $\boldsymbol{\psi}_s^{\alpha\beta}\boldsymbol{\Delta}$ be denoted by

$$\psi_{\alpha\beta,\, gh,\, s}$$

Then

$$\mathbf{X}^{\mathrm{T}} = (X_{uv})$$

$$= \begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1m} \\ \vdots & & \vdots \\ \mathbf{A}_{p+q} & \cdots & \mathbf{A}_{p+qm} \end{bmatrix}_{(p+q)k^2 \times mk^2}, \tag{34}$$

where

$$\mathbf{A}_{ij} = (\psi_{\alpha\beta,\, gh,\, j-i}) \quad (j \geqslant i;\ i \leqslant p, \text{ or } j \geqslant i-p;\ i > p),$$
$$= 0, \qquad \text{otherwise.}$$

In other words,

$$X_{sk^2+ik+g,\, s''k^2+jk+h} = \psi_{ig,\, jh,\, s''-s} \quad (s'' \geqslant s;\ s \leqslant p, \text{ or } s'' \geqslant s-p;\ s > p),$$
$$= 0, \qquad \text{otherwise,}$$

where

$$0 \leqslant s'' \leqslant m-1, \quad 0 \leqslant s \leqslant p+q-1, \quad 1 \leqslant i, j, g, h \leqslant k.$$

Theorem 4. $\sqrt{n}\cdot\hat{\mathbf{r}}$ is asymptotically normal with mean $\mathbf{0}$ and covariance matrix $(\mathbf{Y} - \mathbf{X}\mathbf{I}^{-1}\mathbf{X}^{\mathrm{T}})$ where $\mathbf{X}$ is given by (34) and $\mathbf{I}$ is the information matrix,

$$\mathbf{I} = \lim_{n\to\infty} \langle \partial^2 S / \partial\boldsymbol{\beta}\, \partial\boldsymbol{\beta}^{\mathrm{T}} \rangle / (2n),$$

where

$$S = \sum_{t=1}^{n} \dot{\mathbf{a}}_t^{\mathrm{T}} \boldsymbol{\Delta}^{-1} \dot{\mathbf{a}}_t.$$

Proof. This follows exactly as in Theorem 1 and Theorem 2.

Theorem 5. $\mathbf{X}^{\mathrm{T}}(\mathbf{G}\mathbf{G}^{\mathrm{T}})\mathbf{X} \cong \mathbf{I}$ for m sufficiently large.

Proof. Multiply out $\mathbf{X}^{\mathrm{T}}(\mathbf{G}\mathbf{G}^{\mathrm{T}})\mathbf{X}$ and $\langle \partial^2 S / (\partial\boldsymbol{\beta}\, \partial\boldsymbol{\beta}^{\mathrm{T}}) \rangle$. Again, if m is sufficiently large such that $\boldsymbol{\phi}_s^{\alpha\beta}$ and $\boldsymbol{\theta}_s^{\alpha\beta} \cong \mathbf{0}$ for $s > m$, we see that $\mathbf{X}^{\mathrm{T}}(\mathbf{G}\mathbf{G}^{\mathrm{T}})\mathbf{X} \cong \mathbf{I}$.

It follows from Theorem 5 that the portmanteau test statistic defined in (26) is approximately χ^2-distributed with $k^2(m-p-q)$ *DF* when m and n are large enough.

The extension of the results of this section to the vector seasonal multiplicative ARMA model, discussed by Tiao and Box (1979), is straightforward.

5. MODIFIED PORTMANTEAU TEST

Simulation experiments reported by Davies, Triggs and Newbold (1977) and Ljung and Box (1978) showed that the univariate portmanteau test gives significance levels much lower than that suggested by asymptotic theory even for moderate sample sizes. In this section a modified portmanteau test statistic is suggested and its improvement over the unmodified test statistic is demonstrated by a simulation study.

Our recommended multivariate modified portmanteau statistic is

$$Q_m^* = Q_m + \frac{k^2 m(m+1)}{2n}. \tag{35}$$

Under the null hypothesis of model adequacy Q_m^* is asymptotically χ^2-distributed with $k^2(m-p-q)$ DF. Furthermore, it may be shown, along the lines of Ljung and Box (1978) that

$$\langle Q_m^* \rangle = k^2(m-p-q)$$

which suggests that the portmanteau test using Q_m^* will give a better approximation to the null distribution. This is shown to be the case in the simulation study below. In the univariate case, Ljung and Box (1978, p. 301, Section 4.2) suggested a very similar but not exactly equivalent modification.

One thousand simulations of the first-order bivariate autoregressive model with $n = 200$,

$$\mathbf{\Delta} = \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix},$$

where $\alpha = \pm 0{\cdot}25$, $\pm 0{\cdot}5$, $\pm 0{\cdot}75$ and $\boldsymbol{\phi}_1 = \mathbf{A}, \mathbf{B}, \mathbf{C}$ where

$$\mathbf{A} = \begin{pmatrix} -0{\cdot}2 & 0{\cdot}3 \\ -0{\cdot}6 & 1{\cdot}1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0{\cdot}4 & 0{\cdot}1 \\ -1{\cdot}0 & 0{\cdot}5 \end{pmatrix} \quad \text{and} \quad \mathbf{C} = \begin{pmatrix} -1{\cdot}5 & 1{\cdot}2 \\ -0{\cdot}9 & 0{\cdot}5 \end{pmatrix}$$

were done and the portmanteau statistics defined in (26) and (35) were calculated with $m = 20$.

The 5 per cent empirical significance levels for Q_{20} and Q_{20}^*, shown in Table 1, are defined as the proportion of times that the statistic exceeds the upper 5 per cent point of χ^2_{76}. As expected, the modified portmanteau test provides a significant improvement.

TABLE 1

Empirical significance of portmanteau test at 5 per cent level (in per cent)

	A		**B**		**C**	
α	Q_{20}	Q_{20}^*	Q_{20}	Q_{20}^*	Q_{20}	Q_{20}^*
0·25	32	58	29	57	28	61
−0·25	31	56	28	52	27	55
0·5	33	56	27	56	28	62
−0·5	30	56	26	64	32	66
0·75	33	57	22	57	26	60
−0·75	36	73	26	70	36	74

REFERENCES

BILLINGSLEY, P. (1961). The Lindeberg–Levy theorem for martingales. *Proc. Amer. Math. Soc.*, **12**, 788–792.

BOX, G. E. P. and JENKINS, G. M. (1970). *Time Series Analysis Forecasting and Control.* San Francisco: Holden-Day.

BOX, G. E. P. and PIERCE, D. A. (1970). Distribution of the residual autocorrelations in autoregressive integrated moving average time series models. *J. Amer. Statist. Ass.*, **65**, 1509–1526.

CHITTURI, R. V. (1974). Distribution of residual autocorrelations in multiple autoregressive schemes. *J. Amer. Statist. Ass.*, **69**, 928–934.

DAVIES, N., TRIGGS, C. M. and NEWBOLD, P. (1977). Significance levels of the Box–Pierce portmanteau statistic in finite samples. *Biometrika*, **64**, 517–522.

HANNAN, E. J. (1969). The identification of vector mixed autoregressive-moving average systems. *Biometrika*, **56**, 223–225.

—— (1970). *Multiple Time Series.* New York: Wiley.

HOSKING, J. R. M. (1980). The multivariate portmanteau statistic. *J. Amer. Statist. Ass.*, **75**, 602–607.

—— (1981). Lagrange-multiplier tests of multivariate time-series models. *J. R. Statist. Soc.* B, **43**, 219–230.

LJUNG, G. M. and BOX, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, **65**, 297–303.

MCLEOD, A. I. (1978). On the distribution of residual autocorrelations in Box–Jenkins model. *J. R. Statist. Soc.* B, **40**, 296–302.

—— (1979). Distribution of the residual cross-correlation in univariate ARMA time series models. *J. Amer. Statist. Ass.*, **74**, 849–855.

NICHOLLS, D. F. and HALL, A. D. (1979). The exact likelihood function of multivariate autoregressive-moving average models. *Biometrika*, **66**, 259–64.

TIAO, G. C. and BOX, G. E. P. (1979). An introduction to applied multiple time series analysis. Technical Report, University of Wisconsin.

TIAO, G. C. and HILLMER, S. C. (1979). Likelihood function of stationary multiple autoregressive moving average models. *J. Amer. Statist. Ass.*, **74**, 652–660.

WILSON, G. T. (1973). The estimation of parameters in multivariate time series models. *J. R. Statisti. Soc.* B, **40**, 76–85.