Journal of the Korean
OR/MS Society
Vol. 10, No. 2, Oct. 1985

— Research Paper —

# Joint Distribution of ESACF Arrays within Triangular Zero Boundary

Sung Joo Park*
Tae Joon Jeon*

**Abstract**

An obvious criticism of ESACF approach for model identification by Tsay and Tiao(1984) is that the user may be confused by the elements which are in triangular but marginally larger two standard deviation values. To avoid this drawback, the joint limiting distribution of the vector whose elements are in triangular of ESACF arrays is verified and the statistics to test the nullity of the vector are suggested. We illustrate this approach with three examples.

## 1. Introduction

Suppose $n$ observations are available from the ARMA(p, q) process

$$\Phi(B) Z_t = \theta(B) a_t$$

where $\Phi(B) = U(B)\,\phi(B) = 1 - \Phi_1 B - \dots - \Phi_p B^p$

$$U(B) = 1 - U_1(B) - \dots - U_d B^d$$

$$\phi(B) = 1 - \phi_1(B) - \dots - \phi_{p-d} B^{p-d}$$

$$\theta(B) = 1 - \theta_1(B) - \dots - \theta_q B^q$$

$\{ a_t \}$ ; Gaussian white noise process with mean zero and variance $\sigma_a^2$

We shall require that all the zeros of $U(B)$ are on, those of $\phi(B)$ are outside, and those of $\theta(B)$ are on or outside the unit circle, and also that $\Phi(B)$ and $\theta(B)$ have no common factors. Further, we shall assume that $Z_t$ starts at a finite time point $t_0$ if it is nonstationary.

For tentative model identification of univariate time series, several approaches have been proposed for practical uses. They can be categorized as the post-estimation method and the pre-estimation method. The post-estimation method such as the information criterion (AIC) by Akaike (1974), criterion autoregressive transfer function (CAT) by Parzen(1982), data dependent system (DDS) by Pandit and Wu (1983), select the model after estimation procedure.

* Dept of Management Science, KAIST.

But the pre-estimation method such as sample autocorrelation functon (SACF) and sample partial autocorrelation function (SPACF) approach by Box and Jenkins (1970), the R-and S-array by Gray, et al.(1978), the Pade table approach by Beguin, et al. (1980), and Tucker (1982), suggest the model before estimating the parameters.

Tsay and Tiao (1984) have developed the extended sample autocorrelation function (ESACF) approach which can be categorized as mixed method of post-estimation method and pre-estimation method. Based on the consistent AR estimates produced by iterated regressions, the ESACF is defined and the model is selected by its "triangular cutting off behavior." The ESACF approach is very powerful bacause it can handle nonstationary as well as stationary models in a much more direct manner. But the obvious criticism of ESACF approach is the user may be confused by the elements which are in triangular but marginally larger than two standard deviation value.

The main purpose of this paper is to develop the statistics which represent "composite Triangular zero behavior" and a procedure for identifing the model automatically.

Section 2 investigates the variance and covariance of the ESACF's and developes the limiting distribution of the vector whose elements are within triangular zero boundary of ESACF arrays and the statistics to test the nullity of the vector. Using the statistics in Section 2, the model identification procedure is given in Section 3 with three examples. A summary and discussions are given in Section 4.

## 2. Variance and Covariance of ESACF Array

We first define the autocorrelation $rw(l)$ as

$$rw(l) = r_l(W_t) \tag{2}$$

Then the asymptotic variance of the $r_{j(k)}$'s can be approximately obtained by using Bartelett formula. On the assumption that all the autocorrelations $r_w(l)$ are approximately zero for $l >$ j, Barlett's approximation gives

$$\text{Var}[r_{j(k)}] \doteqdot [1 + 2 \sum_{v=1}^{J-1} [\{r_w(l)\}^2]/(N-k) \tag{3}$$

For ARMA ($p$, $q$) time series, the equation (3) holds when $j-q > k-p \geqq 0$. If $r_{(jk)}$'s are independent, under the null hypothesis:

$$\text{H}_o = \{\rho_{j(k)} = 0,\ j-q > k-p \geqq 0\}$$

the value

$$QI_T(p,\ q) = \sum_{k=p}^{p+K} \sum_{j=q-p+k+1}^{q+K+1} \frac{\{r_{j(k)}\}^2}{\text{Var}[r_{j(k)}]} \tag{4}$$

where $T = \sum_{k=1}^{K} k$

converges in distribution to a Chi-square with degree of freedom $T$. The set $\{QI_T > C_\alpha\}$, where $C_\alpha$ is the $\alpha$-percentage point in $\chi^2(T)$, is an asymtotic critical region at level $\alpha$ for testing $H_o$.

In general, for any nonnegative integer $u$ and $v$, we define the $QI_T(u,\ v)$ value as (4). Suppose that the ARMA process $Z_t$ in (1) is nonstationary, that is, $U(B)=1$. In this case, the values $QI_T(u, v)$, $u< d$, are not converge in distribution to a Chi-square, because asymtotically $r_{v+m(u)} \neq 0$, $m>1$, due to remaining nonstationary part of the AR process.

Using the above result, we can eliminate the possible AR order $u$, $u< d$, by checking $QI_T(u, v)$ values.

The covariances of $r_{j(k)}$ cannot be solved easily because they are covariance of autocorrelations of different series. But we can approximate the covariances of $r_{j(k)}$'s which are in triangular whose vertex is $r_{q+1(p)}$ by following theorems and corollaries.

**Theorem 1.** For ARMA$(p, q)$ time series, let $\{W_t\}$, $\{X_t\}$ and $\{Y_t\}$ be defined respectively by

$$W_t = \Phi_p^{(q)}(B) \quad Z_t = \sum_{m=0}^{q} \theta_m a_{t-m}$$

$$X_t = H_j^{(i)}(B) \quad W_t \qquad (j > i \geqq 0)$$

$$Y_t = H_l^{(k)}(B) \quad W_t \qquad (l > k \geqq 0)$$

where the $a_t$ are independent $(0, \sigma^2)$ random variable with $\mathrm{E}\{a_t^4\} = \eta\sigma^4$.

Then for fixed $h$ and $g$, $i+h \geqq k+g$,

$$\lim_{n\to\infty} (n\text{-}k\text{-}g)\ Cov\ \{\gamma_X(h), \gamma_Y(g)\} = \sum_{t=-q-j}^{q+k-jl} S_{XY}(t)\ S_{XY}(t+jl) \tag{5}$$

where $S_{XY}(t) = H_j^{(i)}(F) H_1^{(k)}(B) \gamma_W(t)$,

$h = q+j$

$g = q+l$

$jl = |j-l|$

**Proof)** See Appendix 1.

**Corollary 1.** Let the assumptions of Theorem 1 hold and $\mathrm{E}\{a_t^6\} = \eta\sigma^6$. Then,

$$Cov\{r_X(h), r_Y(g)\}$$

$$= \{(n\text{-}k\text{-}g)\gamma_X(0)\gamma_Y(0)\}^{-1} \sum_{t=-q-i}^{q+k-jl} S_{XY}(t) S_{XY}(t+j1) + 0(n^{-2}) \tag{6}$$

where $S_{XY}(t) = H_j^{(i)}(F) H_l^{(k)}(B) \gamma_W(t)$

**Proof)** The estimated autocorrelations are bounded and are differentiable functions of the estimated autocovariances on a closed set containing the true parameter vector as an interior point. Furthermore, the derivations bre bounded on that set. Hence, the conditions of Thoerem 5.4.3 (Fuller, 1976) are met with $\{\tilde{\gamma}_X(h), \tilde{\gamma}_Y(g), \tilde{\gamma}_X(0), \tilde{\gamma}_Y(0)\}$ playing the role of $\{X_n\}$.

Since the function $r_X(h)$ and $r_Y(g)$ are bounded, we take $\alpha=1$. Expanding $\mathrm{E}\{r_X(h)-\rho_X(h)]$ $[r_Y(h)-\rho_Y(g)]$ through third order terms and using Theorem 5.4.1 (Fuller, 1976) to establish that the expected value of the third order moment is $0(n^{-2})$, we have

$$Cov\{r_X(h), r_Y(g)\} = [\gamma_X(0)\gamma_Y(0)]^{-1} [Cov\{\tilde{\gamma}_X(h), \tilde{\gamma}_Y(g)\}$$

$$-\rho_X(h)\, Cov\{\tilde{\gamma}_X(0), \tilde{\gamma}_Y(g)\} - \rho_Y(g)\, Cov\{\tilde{\gamma}_X(h), \tilde{\gamma}_Y(0)\}$$

$$+\rho_X(h)\rho_Y(g)\, Cov\{\tilde{\gamma}_X(0), \tilde{\gamma}_Y(0)\}] + 0(n^{-2}).$$

Using Theorem 1, we have the result, because $\rho_X(h)$ and $\rho_Y(g)$ are zero.

we now turn to the large sample properties of the estimated autocovariances and autocorrelations.

**Theorem 2.** For ARMA $(p, q)$ time series, let $\{W_t\}$ and $X_t^{(i,j)}$ be defined respectively by

$$W_t = \Phi_p^{(q)}(B)\ Z_t = \sum_{m=0}^{M} \theta_m a_{t-m}$$

$$X_t^{(i,j)} = H_j^{(i)}(B)\ W_t \qquad (j > i \geqq 0)$$

where the $a_t$ are independent $(0, \sigma^2)$ random variables with fourth moment $\eta\sigma^4$ and sixth moment $\xi$.

And we let $\alpha_1, \alpha_2, \ldots, \alpha_T$ be the elements of finite subset $\{ \tilde{\gamma}_{q+j(p+i)},\ i > i \geqq 0 \}$.

Then the limiting distribution of $n^{1/2}(\alpha_1, \alpha_2, \ldots, \alpha_T)$ is multivariate normal with mean zero and covariance matrix $V$ where the elements of $V$ are defined by Theorem 1.

**Proof**) See Appendix 2.

**Corollary 2.** Let the assumptions of Corolly 1 hold, and $\beta_1, \beta_2, \ldots, \beta_T$ be the elements of finite subset $\{ r_{q+j(p+i)},\ j > i \geqq 0 \}$.

Then the limiting distributing of $n^{1/2}(\beta_1, \beta_2, \ldots, \beta_T)$ is multivariate normal with mean zero and covariance matrix $G$ where the elements of $G$ can be derived by Corollary 1.

**Proof)** Since the $r_{q+j(p+i)}$ are continuous differentiable functions of the $\hat{\gamma}_{q+j(p+1)}$, the result follows immediately from Theorem 5.1.4 (Fuller, 1976), Corrollary 1 and Theorem 2.

Using the above result, under the null hypothesis:

$H_o = \{ \rho_{q+j(p+i)} = 0,\ j > i \geqq 0 \}$

the quadratic form

$$QD_T(p, q) = n(\beta_1, \beta_2, \ldots, \beta_T)'\ G^{-1}(\beta_1, \beta_2, \ldots, \beta_T) \qquad (7)$$

converges in distribution to $\chi^2(T)$. The set $\{ QD_T > C_\alpha \}$, where $C_\alpha$ is the $\alpha$-percentage point in $\chi^2(T)$ is an asymtotic critical region at level $\alpha$ for testing $H_o$.

## 3. Model Identification Procedure

To indentify the ARMA $(p, q)$ model tentatively, the ESACF table and indicator symbol are examined to check the "triangular cutting off behavior". The obvious criticism of the ESACF table and its indicator symbol is that the user may be confused by the elements which are in triangular but marginally larger than two standard deviation values.

To avoid this drawback, the $QI_T$ and $QD_T$ values defined in Section 2 can be used to check the "composite triangular zero behavior" for the model indentification. That is, we search the North-West coordinate $(i, j)$ whose $QI_T(i, j)$ and $QD_T(i, j)$ values are less than $C_\alpha$, where $C_\alpha$ is the $\alpha$-percentage of $\chi^2(T)$, and tentatlvely identify $(i, j)$ as an order of the model.

We illustrate this approach at $T=6$ and $\alpha = 0.05$ with three examples which were also demonstrated by Tsay and Tiao (1984). The elements of ESACF array used to calculate $Q_6(i, j)$ value are shown in Table 1.

**Table I**

The elements of ESACF array for calculating $QI_6(i, j)$ and $QD_6(i, j)$

| AR | MA | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0 | 1 | … | $i$ | $j+1$ | $j+2$ | … |
| 0 | $r_1(0)$ | $r_2(0)$ | … | $r_{j+1}(0)$ | $r_{j+2}(0)$ | $r_{j+3}(0)$ | … |
| 1 | $r_1(1)$ | $r_2(1)$ | … | $r_{j+1}(1)$ | $r_{j+2}(1)$ | $r_{j+3}(1)$ | … |
| ⋮ | ⋮ | ⋮ | | ⋮ | ⋮ | ⋮ | |
| $i$ | $r_1(i)$ | $r_2(i)$ | … | $r_{j+1}(i)$ | $r_{j+2}(i)$ | $r_{j+3}(i)$ | … |
| $i+1$ | $r_1(i+1)$ | $r_2(i+1)$ | … | $r_{j+1}(i+1)$ | $r_{j+2}(i+1)$ | $r_{j+3}(i+1)$ | … |
| $i+2$ | $r_1(i+2)$ | $r_2(i+2)$ | … | $r_{j+1}(i+1)$ | $r_{j+2}(i+2)$ | $r_{j+3}(i+2)$ | … |
| ⋮ | ⋮ | ⋮ | | ⋮ | ⋮ | ⋮ | |

**Example 1.** For the series $C$ in Box and Jenkins (1970), Tsay and Tiao (1984) suggested an AR(2) model by the "triangular cutting off" behavior of the indicator symbols.

Table II gives the $QI_6$ and $QD_6$ arrays and justifies AR(2) model because $QI_6(2, 0)$ and $QD_6(2, 0)$ are less than theoretical $\chi^2(6, 0.05)$ value.

**Table II**

The $QI_6$ and $QD_6$ Arrays of Series $C$

| AR | MA | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| | a. The $QI_6$ array | | | | | | |
| 0 | 391. 2 | 169. 8 | 106. 6 | 74. 5 | 55. 4 | 40. 7 | 33. 5 |
| 1 | 214. 3 | 83. 4 | 44. 7 | 29. 2 | 21. 6 | 15. 6 | 14. 8 |
| 2 | 4. 5* | 9. 4* | 6. 3* | 1. 1* | 1. 3* | 1. 8* | 7. 2* |
| 3 | 79. 7 | 6. 8* | 5. 6* | 2. 8* | . 3* | . 4* | 3. 6* |
| 4 | 50. 2 | 27. 7 | 4. 1* | 3. 2* | . 5* | . 4* | 2. 6* |
| 5 | 94. 3 | 21. 5 | 22. 9 | 2. 1* | . 8* | . 6* | 2. 0* |
| 6 | 31. 2 | 27. 9 | 29. 4 | 14. 2 | . 4* | . 5* | 1. 7* |
| | b. The $QD_6$ array | | | | | | |
| 0 | — | — | — | 8703. 6 | 3408. 5 | 2140. 3 | 2006. 5 |
| 1 | 1121. 2 | — | — | 5072. 5 | 3888. 8 | 2451. 1 | 2038. 2 |
| 2 | 5. 9* | 469. 0 | — | 212. 8 | 91. 0 | 39. 9 | 286. 4 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 3 | — | 74. 5 | 600. 4 | — | 426. 3 | 432. 2 | 2. 3* |
| 4 | — | 93. 7 | — | — | — | . 0* | 79. 7 |
| 5 | — | — | 230. 1 | 134. 2 | — | — | 37. 5 |
| 6 | 3057. 7 | 443. 5 | 1006. 6 | — | — | 461. 5 | 105. 9 |

( * : less than $\chi^2(6, 0.05)$ value
— : greater than 9999.9)

**Example 2.** Tsay and Tiao (1984) generated one hundred observations from the ARMA (4, 1) model and analyze the ESACF table. The $QI_6$ and $QD_6$ arrays are shown in Table III and we can determine ARMA (4, 1) model as suggested by them.

For this example, it is of interest to examine the $QI_6(i, j)$ values when $i$ is less than the order of the nonstationary part of the AR process. For the values $QI_6(i, j)$, $i < 4$, are all greater than $\chi^2(0.05, 6)$, we can eliminate nonstationary factor of the AR process.

**Example 3.** for the series $A$ in Box and Jenkins (1970), ARMA (1, 1) model was suggested by Tiao and Tsay (1984).

The $QI_6$ and $QD_6$ arrays given in Table IV also confirmed ARMA (1, 1) model. In this case, we can see this procedure is not confused by the ESACF elements whose values are marginally larger than the two standard deviation values.

**Table III**

The $QI_6$ and $QD_6$ Arrays of Example 2

| AR | MA | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 0 | 2 | 3 | 4 | 5 | 6 |
| | a. The $QI_6$ Array | | | | | | |
| 0 | 173. 6 | 125. 4 | 77. 8 | 35. 8 | 37. 2 | 53. 0 | 41. 4 |
| 1 | 92. 0 | 88. 8 | 86. 9 | 45. 2 | 27. 2 | 45. 7 | 48. 9 |
| 2 | 93. 2 | 57. 5 | 61. 7 | 45. 3 | 25. 6 | 30. 7 | 33. 6 |
| 3 | 84. 3 | 31. 0 | 18. 1 | 15. 9 | 15. 9 | 16. 8 | 13. 7 |
| 4 | 24. 5 | . 7* | . 3* | . 6* | 1. 4* | 3. 4* | 3. 9* |
| 5 | 18. 0 | 11. 3* | . 3* | . 3* | . 5* | . 9* | . 8* |
| 6 | 31. 1 | 6. 4* | 7. 0* | . 2* | . 5* | . 6* | . 4* |
| | b. The $QD_6$ Array | | | | | | |
| 0 | — | 4013. 2 | 1076. 2 | 1986. 1 | 1490. 2 | 631. 1 | 803. 7 |
| 1 | — | 1520. 5 | 3064. 4 | 1341. 7 | 290. 7 | 1186. 3 | 1200. 7 |
| 2 | — | — | 3. 6* | 1524. 9 | 2883. 0 | 621. 8 | . 1* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 3 | 1571. 6 | 4605. 7 | 1557. 9 | 203. 4 | 524. 1 | 432. 4 | 465. 2 |
| 4 | 1534. 5 | . 4* | — | 28. 1 | 5700. 3 | 45. 8 | 177. 8 |
| 5 | 421. 8 | 122. 5 | — | — | 4. 6* | 115. 1 | 2. 5* |
| 6 | — | 572. 6 | — | — | — | — | 5065. 1 |

( * : less than $\chi^2(6, 0.05)$ value
— : greater than 9999. 9 )

**Table Ⅳ**

The $QI_6$ and $QD_6$ arrays of series A

| AR | MA | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 6 |
| | a. The $QI_6$ array | | | | | | |
| 0 | 109. 0 | 54. 5 | 33. 4 | 26. 8 | 33. 9 | 28. 1 | 21. 4 |
| 1 | 45. 9 | 1. 1* | 1. 5* | 1. 1* | 13. 4 | 9. 2* | 5. 9* |
| 2 | 32. 1 | 15. 3 | . 7* | . 7* | 12. 6 | 9. 5* | 7. 1* |
| 3 | 54. 7 | 2. 5* | 4. 2* | . 5* | 9. 3* | 7. 9* | 6. 9* |
| 4 | 83. 9 | 4. 3* | 11. 4 | 3. 4* | 5. 9* | 6. 1* | 4. 7* |
| 5 | 70. 4 | 58. 9 | 23. 7 | 11. 6* | 12. 6 | 3. 9* | 3. 2* |
| 6 | 51. 9 | 34. 4 | 53. 2 | 20. 5 | 19. 5 | 11. 0* | 2. 7* |
| | b. The $QD_6$ array | | | | | | |
| 0 | 1409. 8 | 9555. 7 | 4725. 7 | 1910. 9 | 1514. 8 | 2739. 9 | 1016. 3 |
| 1 | 8529. 2 | 6. 8* | 5. 5* | 113. 1 | 1880. 1 | 138. 0 | 262. 6 |
| 2 | 4876. 6 | 6463. 6 | 6773. 0 | — | 6. 4* | — | . 4* |
| 3 | — | 36. 0 | 399. 3 | 25. 4 | 25. 2 | 28. 3 | 591. 5 |
| 4 | — | — | . 0* | 204. 5 | 2. 3* | 17. 7 | 247. 2 |
| 5 | — | 645. 5 | 6376. 7 | 44. 8 | 397. 6 | 369. 3 | 2424. 7 |
| 6 | — | — | 2494. 3 | 1550. 9 | . 0* | 3687. 9 | . 2* |

( * : less than $\chi^2(6, 0.05)$ value
— : greater than 9999. 9 )

## 4. Conclusion

In this paper we have investigated the limiting distribution of the vector whose elements are within triangular boundaries of the ESACF array and the statistics, the $QI_T$ and $QD_T$ arrays, to test the nullity of the vector.

The major advantage of model identification procedure using those statistics is that it can avoid the difficulties of manual checking the triangular cutting off behavior of ESACF array, which was the problem of the ESACF approach.

All examples considered were nonseasonal in nature. A study of other seasonal examples is needed in the future.

## Appendix 1. (Proof of Theorem 1)

Let $\theta_m^{j(i)} = H_j^{(i)}(B)\ \theta_m$, then the sequence $\{ \theta_m^{j(i)} \}$ is finite and absolutely summable, because $\theta_m$ is zero for $m > q+i$ and $m < 0$.

From definition,

$$X_t = H_j^{(i)}(B)\ W_t = \sum_{m=-\infty}^{\infty} \theta_m^{j(i)}\ a_{t-m}$$

$$Y_t = H_l^{(k)}(B)\ W_t = \sum_{m=-\infty}^{\infty} \theta_m^{l(k)}\ a_{t-m}$$

and

$$E\ \{ X_t X_{t+h} Y_{t+h+f} Y_{t+h+f+g} \}$$

$$=(\eta-3)\ \sigma^4 \sum_{m=-\infty}^{\infty} \{ \theta_m^{j(i)}\ \theta_{m+h}^{j(i)}\ \theta_{m+h+f}^{l(k)}\ \theta_{m+k+f+g}^{l(k)} \}$$

$$+\sigma^4 \sum_{m=-\infty}^{\infty} \{ \theta_m^{j(i)}\ \theta_{m+h}^{j(i)} \} \sum_{m=-\infty}^{\infty} \{ \theta_{w+h+f}^{l(k)}\ \theta_{m+h+f+g}^{l(k)} \}$$

$$+\sigma^4 \sum_{m=-\infty}^{\infty} \{ \theta_m^{j(i)}\ \theta_{m+h+f}^{l(k)} \{ \theta_{m+h}^{j(i)}\ \theta_{m+h+f+g}^{l(i)} \}$$

$$+\sigma^4 \sum_{m=-\infty}^{\infty} \{ \theta_m^{j(i)}\ \theta_{m+h+f+g}^{l(k)} \} \sum_{m=-\infty}^{\infty} \{ \theta_{m+h}^{j(i)}\ \theta_{m+h+f}^{l(k)} \}$$

$$= (\eta-3)\ \sigma^4 \sum_{m=-\infty}^{\infty} \{ \theta_m^{j(i)}\ \theta_m^{j(i)}\ \theta_{m+h+f+g}^{l(k)} \}$$

$$+\gamma_X(h)\ \gamma_Y(g) + S_{XY}(h+f)\ S_{XY}(f+g) + S_{XY}(h+f+g)\ S_{XY}(f)$$

where $S_{XY}(t) = H_j^{(i)}(F)\ H_l^{(k)}(B)\ \gamma_W(h)$

Thus,

$$E\ \{ \gamma_X(h)\ \gamma_Y(g) \} - \gamma_X(h)\ \gamma_Y(g)$$

$$= \frac{1}{(n-i-h)\ (n-k-g)} \{ \sum_{s=1}^{n-k-g} \sum_{t=1}^{n-i-h} E(X_t X_{t+h} Y_s Y_{s+g}) \} - \gamma_X(h)\ \gamma_Y(g)$$

$$= \frac{(\eta-3)\sigma^4}{(n-i-h)\ (n-k-g)} \sum_{s=1}^{n-k-g} \sum_{t=1}^{n-i-h} \{ \theta_m^{j(i)}\ \theta_{m+h}^{j(i)}\ \theta_{m+s-t}^{l(k)}\ \theta_{m+s-t+g}^{l(k)} \}$$

$$+ \frac{1}{(n-i-h)\ (n-k-g)} \sum_{s=1}^{n-k-g} \sum_{t=1}^{n-i-h} \{ S_{XY}(s-t)\ S_{XY}(s-t-h+g)$$

$$+ S_{XY}(s-t+g)\ S_{XY}(s-t-h) \}$$

Applying Lemma 6.2.1 (Fuller, 1976), we have

$$\lim_{n\to\infty} (n-k-g)\ Cov\ \{ \tilde{\gamma}_X(h),\ \tilde{\gamma}_Y(g) \}$$

$$= (\eta-3)\sigma^4 \sum_{m=-\infty}^{\infty} \sum_{t=-\infty}^{\infty} \{ \theta_m^{j(i)}\ \theta_{m+h}^{j(i)}\ \theta_{m+t}^{l(k)}\ \theta_{m+t+g}^{l(k)} \}$$

$$+ \sum_{t=-\infty}^{\infty} \{ S_{XY}(t)\, S_{XY}(t-h+g) + S_{XY}(t+g)\ S_{XY}(t-h) \}$$

For ARMA$(p,\ q)$ time series, $\gamma_X(h)$, $\gamma_Y(g)$ and $S_{XY}(t)$, $t > q+k, t < -q-i$ are zero. If we let $jl = | j-l |$ , then

$$\sum_{t=-\infty}^{\infty} S_{XY}(t)\, S_{XY}(t-h-q) = \sum_{t=-q-i}^{q+k-jl} S_{XY}(t)\ S_{XY}(t+jl)$$

$$\sum_{t=-\infty}^{\infty} S_{XY}(t+g)\, S_{XY}(t-h) = 0$$

## Appendix 2. (Proof of Theorem 2)

Let $\bar{X}$ is artithmatic mean of $X_t^{(i,j)}$ series, then the estimated covariance is

$$\hat{\gamma}_{q+j(p+j)} = \frac{1}{n-p-i} \sum_{t=1}^{n-p-q-i-j} (X_t^{(i,j)} - \bar{X})\ (X_{t+q+j}^{(i,j)} - \bar{X})$$

$$= \frac{1}{n-p-i} \sum_{t=1}^{n-p-q-i-j} X_t^{(i,j)}\ X_{t+q+j}^{(i,j)} - \frac{\bar{X}}{n-p-i} \sum_{t=1}^{n-p-q-i-j} (X_t^{(i,j)} + X_{t+q+j}^{(i,j)})$$

$$+ \frac{n-p-q-i-j}{n-p-i} \bar{X}^2$$

and the last two terms, when multiplied by $n^{1/2}$ converges in probabiliry to zero. Therefore in investigating the limiting distribution of $n^{1/2}\ \hat{\gamma}_{q+j(p+i)}$ we need only consider the first term. Let

$$S_n = n^{1/2} \sum_{i=0}^{I} \sum_{j=i+1}^{J} \lambda_{ij} \left[\frac{1}{n-p-i} \sum_{t=1}^{n-p-q-i-j} X_t^{(i,j)}\ X_{t+q+j}^{(i,j)} \right]$$

$$= n^{1/2} \sum_{i=0}^{I} \sum_{j=i+1}^{J} \lambda_{ij} \left[\frac{1}{n-p-i} \sum_{t=1}^{n-p-q-i-j} Z_t^{(i,j)} \right]$$

where the $\lambda_{ij}$ are arbitrary real numbers (not all zero) and $Z_t^{(i,j)} = X_t^{(i,j)}\ X_{t+q+j}^{(i,j)}$

Now $Z_t^{(i,j)}$ is an $(2q+i+j)$-dependent covariance stationary time series with mean $\gamma_{q+j(p+i)}$ and covariance function

$E[Z_t^{(i,j)}\ Z_{t+h}^{(k,l)}]$

$$= (\eta-3)\sigma^4 \sum_{m=-\infty}^{\infty} \{ \theta_m^{j(i)}\ \theta_{m+q+j}^{j(i)}\ \theta_{m+h}^{k(l)}\ \theta_{m+h+l}^{k(l)} \}$$

$$+\ \gamma_{q+j(p+i)}\ \gamma_{q+l(p+k)}\ S(h)\, S(h-j+l) + S(h-q-j)\, S(h+q+l)$$

where $S(t) = H_j^{(i)}(B)\, H_1^{(k)}(F)\ \gamma_W(t)$

and it is understood that $\theta_m^{j(i)} = 0$ for $m > q + i$ and $m < 0$.

Thus, the weighted average of the $Z_t^{(i, j)}$ 's

$$Y_t = \sum_{i=1}^{I} \sum_{j=i+1}^{J} \lambda_{ij} Z_t^{(i, j)} = \sum_{j=0}^{I} \sum_{j=i+1}^{J} \lambda_{ij} X_t^{(i, j)} X_{t+q+j}^{(i, j)}$$

is a stationary time series. Furthermore, the time series $Y_t$ is $(2q+I+J)$-dependent and has finite third moment.

Therefore, $S_n$ converges in distribution to a normal random variable. Since $\lambda_{ij}$ were arbitrary, the vector random variable $n^{1/2}(\alpha_1, \alpha_2, \ldots, \alpha_T)'$ converges in distribution to a multivariate normal by Theorem 5.3.3 (Fuller, 1976).

## References


1. Akaike, H., (1974). "A new look at the statistical model identification", *IEEE Tra . AC,* 19, 716-723.
2. Beguin, J. M., Gourieroux, C. & Monfort, A., (1980). "Identification of a mixed autoregressive moving average process : the corner method", in *Time Series* (ed). O. D. Anderson, Amsterdam ; North-Holland, 423-436.
3. Box, G. E. P., Jenkins, G. M., (1970). *Time Series Forecasting and Control,* Holden Day, New York.
4. Fuller, W. A., (1976) *Introducation to Statistical Time Zories,* John Wiley & Sons, New York.
5. Gray, H. L., Kelly, A. G., & McIntire, D. D., (1978) "A new approach to ARMA modelling", *Commun. Statis. B,* 7, 1-77.
6. Pandit, S. M., Wu, S. M., (1983). *Time Series and System Analysis with Applications,* John Wiley and Sons.
7. Parzen, E., (1982). "ARMA Models for time series analysis and forecasting", *J. of Forecasting,* 1, 67-82.
8. Quinn, B. G., (1980). "Limiting behavior of autocorrelation funtion of ARMA process as several roots of characteristic equation approach unit circle", *Commu. Statis. B,* 9, 195-198.
9. Tiao, G. C., & Tsay, R. S., (1983) "Consistency properties of least square estimates of autoregressive paramaters in ARMA models", *Annals of Statistics,* 11, 856-871.
10. Tsay, R. S., & Tiao, G. C., (1984). "Consistent estimates of AR paramaters and ESACF for stationary and nonstationary ARMA models", *J. of American Statistical Association,* 79, 84-96.
11. Tucker, W. T., (1982). "On the Pade table and its relationship to the *R* and *S* arrays and *ARMA* modelling", *Comm. Stat. A,* 11, 1335-79.
12. Woodward, W. A., & Gray, H. L.,(1981). "On the relationship between the S-array and the Box-Jenkins method of ARMA model identification", *J. of American Statistical Association,* 76, 579-587.