Journal of the Korean
Statistical Society
Vol. 24, No. 2, 1995

Testing for a Unit Root in an ARIMA$(p, 1, q)$ Signal Observed with Measurement Error

JongHyup Lee[1] and DongWan Shin[2]

ABSTRACT

An ARIMA signal observed with measurement error is shown to have another ARIMA representation with nonlinear restrictions on parameters. For this model, the restricted Newton-Raphson estimator(RNRE) of the unit root is shown to have the same limiting distribution as the ordinary least squares estimator of the unit root in an AR(1) model tabulated by Dickey and Fuller(1979). The RNRE of parameters of the ARIMA$(p, 1, k)$ process and unit root tests based on the RNRE are developed.

KEYWORDS: Measurement error, Restricted ARIMA model, Newton-Raphson procedure, Restricted maximum likelihood estimation, Unit root tests.

1. INTRODUCTION

Testing for a unit root in ARIMA model has attracted a broad attention and has wide applications. We consider the following model:

$$x_t = \rho x_{t-1} + z_t, \tag{1.1}$$

[1]Department of Statistics, Sungshin Women's University, Seoul, 136–742, Korea.
[2]Department of Statistics, Ewha Womans University, Seoul, 120–750, Korea.

with

$$z_t + \alpha_1 z_{t-1} + \cdots + \alpha_p z_{t-p} = a_t + \beta_1 a_{t-1} + \cdots + \beta_q a_{t-q}, \tag{1.2}$$

for $t = 1, 2, \ldots, n$, where p and q are known nonnegative integers and the a_t are i.i.d. $\mathrm{N}(0, \sigma_{aa})$. It is not always possible to observe a time series x_t directly. Instead of observing x_t one observes

$$y_t = x_t + u_t, \tag{1.3}$$

where u_t ia a measurement error, the u_t are i.i.d. $\mathrm{N}(0, \sigma_{uu})$ and the a_t and the u_t are independent unobservable processes. We assume that

$$A(m) = 1 + \alpha_1 m + \cdots + \alpha_p m^p$$

and

$$B(m) = 1 + \beta_1 m + \cdots + \beta_q m^q$$

have all roots outside the unit circle, that is, $A(m) \neq 0$ and $B(m) \neq 0$ for all m with $|m| \leq 1$.

The x_t defined by (1.1) and (1.2) is an ARIMA(p,1,q) process if $\rho = 1$. We are interested in the estimation for model (1.1)–(1.3) in the case when $\rho = 1$, and testing of the hypothesis that $\rho = 1$. For ARIMA(p,1,q) models there are several methods for estimating the unit root parameter ρ and then testing the hypothesis that $\rho = 1$ based on derived estimators. Said and Dickey(1985) suggested a one-step Gauss-Newton method. Phillips(1987) and Phillips and Perron(1988) proposed alternative methods under very weak conditions on the innovations. Hall(1989) proposed an instrumental variable approach. Pantula and Hall(1991) extended Hall(1989)'s results to ARIMA(p, 1, q) models. Shin and Sarkar(1993) developed unit root tests based on the restricted maximum likelihood estimation of an ARIMA(p, 1, 0) signal disturbed by MA(q) noise with known coefficients, which is motivated by rotational sampling scheme conducted by Bureau of Census in America or Statistics Canada. Shin and Sarkar(1994) considered various aspects about the modified restricted maximum likelihood estimation in rotational survey sampling, limiting distribution, an algorithm for computing the first derivatives of the modified likelihood function.

However, the model studied by Shin and Sarkar(1993,1994) is only applicable under the rotation sampling scheme. Also, most of their estimators are obtained through intensive complicated numerical burden because the observation $\{y_t\}$ is an ARIMA with nonlinear restrictions on the parameters. In this study, we present a simpler and easier-to-compute unit root tests based on Newton-Raphson approximation to the least squares estimator and the modified maximum likelihood estimator of the measurement error model (1.1)–(1.3).

The remainder of the paper is organized as follows. In Section 2, a new ARIMA representation for y_t with parametric restrictions is discussed. In Section 3 the Newton-Raphson estimators of parameters and test statistics for a unit root are developed. In Section 4 large sample properties of the test statistics are studied. Various computational aspects such as explicit expressions of derivatives of the restrictions are discussed in the appendix.

2. A NEW ARIMA REPRESENTATION FOR OBSERVATIONS

In Section 2, we give an ARIMA representation for y_t, in which the new set of parameter and the old set of parameter are nonlinearly related. Also the nonlinear relations are clarified in order to be used in improving the performance of unit root tests.

We can express the process $\{y_t\}$ as an autoregressive integrated moving average. From (1.1)–(1.3), eliminating x_t, we have

$$\sum_{j=0}^{p} \alpha_j (y_{t-j} - \rho y_{t-1-j}) = \sum_{j=0}^{p+1} \alpha_j^* u_{t-j} + \sum_{j=0}^{q} \beta_j a_{t-j}, \tag{2.1}$$

where $\alpha_j^* = \alpha_j - \rho\alpha_{j-1}$, $j = 0, 1, \ldots, p+1$, $\alpha_0 = 1$ and $\alpha_{-1} = \alpha_{p+1} = 0$. Since the autocovariance function of the right hand side of (2.1) is zero when lag is greater than k=max$(p+1, q)$, we can find a k-th order moving average

$e_t + \sum \gamma_j e_{t-j}$ whose autocovariance function is the same as that of the right hand side of (2.1). Hence we can write

$$\sum_{j=0}^{p} \alpha_j (y_{t-j} - \rho y_{t-1-j}) = e_t + \sum_{j=1}^{k} \gamma_j e_{t-j}, \tag{2.2}$$

where $\{e_t\}$ is a sequence of i.i.d. $N(0, \sigma_{ee})$ random variables, $\sigma_{ee} > 0$ and $\gamma_0 = 1$. Using arguments similar to those used in Lemma 1 through Lemma 3 of Pagano(1974), it can be shown that the $\{\gamma_j, j = 1, \ldots, k\}$ can be uniquely chosen so that the roots of the characteristic equation

$$C(m) = 1 + \gamma_1 m + \cdots + \gamma_k m^k = 0 \tag{2.3}$$

lie outside the unit circle and the roots of $C(m)$ are different from the roots of $A(m)$. Equating the autocovariance function of the right hand side of (2.1) to that of the right hand side of (2.2), for $h = 0, 1, \ldots, k$, we get

$$f_h = \sigma_{ee} \sum_{j=0}^{k-h} \gamma_j \gamma_{j+h} - \sigma_{aa} \sum_{j=0}^{q-h} \beta_j \beta_{j+h} - \sigma_{uu} \sum_{j=0}^{p+1-h} \alpha_j^* \alpha_{j+h}^* = 0, \tag{2.4}$$

where $\alpha_0^* = \beta_0 = \gamma_0 = 1$. Thus, the parameters $(\gamma_1, \ldots, \gamma_k, \sigma_{ee})$ are fuctionally related to $(\rho, \alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \sigma_{aa}, \sigma_{uu})$ and the reparameterized model (2.2) is a restricted ARIMA model of order $(p, 1, k)$.

Let $\alpha = (\alpha_1, \ldots, \alpha_p)'$, $\beta = (\beta_1, \ldots, \beta_q)'$, $\gamma = (\gamma_1, \ldots, \gamma_k)'$, $\theta = (\alpha', \gamma')'$, and $\xi = (\rho, \theta', \sigma_{ee})'$. In order to get real restriction on the parameter ξ of the reparametrized model (2.2), we should eliminate the nuisance parameter β and possibly σ_{uu} from the restriction (2.4). We clarify the restriction (2.4).

(I) When σ_{uu} is unknown; When $p \geq q + 1$, the equations $f_h = 0$ impose parametric restrictions on the parameter ξ of the transformed model. For the time being assume $p \geq q + 1$. Since the restriction $f_h = 0$ contains parameter $(\beta', \sigma_{aa}, \sigma_{uu})'$ which is not being estimated in (2.2), we should eliminate $(\beta', \sigma_{aa}, \sigma_{uu})'$ from (2.4) in order to get real restrictions on ξ. Note that there is no (β', σ_{aa}) in f_h for $h = q+1, \ldots, p+1$. Also, (β', σ_{aa})

can be uniquely determined from ξ through $f_h = 0,\ h = 0, 1, \ldots, q$. Therefore, we can eliminate (β', σ_{aa}) from (2.4) without any loss of information by removing $f_h = 0,\ h = 0, 1, \ldots, q$ from the restrictions. Having eliminated (β', σ_{aa}), we still have one more parameter σ_{uu}, which can be eliminated by noting $f_{p+1} = \sigma_{ee}\gamma_{p+1} - \sigma_{uu}\alpha^*_{p+1} = 0$ and hence $\sigma_{uu} = \sigma_{ee}\gamma_{p+1}/\alpha^*_{p+1}$. Therefore, the real restrictions become

$$\boldsymbol{g} = (g_{q+1}, \ldots, g_p)' = 0, \tag{2.5}$$

where

$$g_h = \alpha^*_{p+1} \sum_{j=0}^{p+1-h} \gamma_j\gamma_{j+h} - \gamma_{p+1} \sum_{j=0}^{p+1-h} \alpha^*_j\alpha^*_{j+h},\ h = q+1, \cdots, p.$$

Note that the restriction $\boldsymbol{g} = \mathbf{0}$ does not contain σ_{ee}. The restriction (2.5) gives some information on ξ that can be used in improving the power of unit root tests. However, when $p < q+1$, the number $(q+2)$ of nuisance parameters in $(\beta', \sigma_{aa}, \sigma_{uu})$ is greater than the number $(q+1)$ of equations in (2.4). Hence, (2.4) imposes no real parametric restriction on ξ and does not give any information on ξ to improve the power of unit root tests.

(II) When σ_{uu} is known; When $p \geq q$, the real restrictions of ξ become

$$f_h = \sigma_{ee} \sum_{j=0}^{p+1-h} \gamma_j\gamma_{j+h} - \sigma_{uu} \sum_{j=0}^{p+1-h} \alpha^*_j\alpha^*_{j+h} = 0,\ h = q+1, \ldots, p+1. \tag{2.6}$$

Thus, σ_{ee} cannot be eliminated from the restrictions. When $p < q$, the equation (2.4) imposes no restriction on ξ.

3. TEST STATISTICS FOR A UNIT ROOT

We develop Newton-Raphson(N-R) procedure for estimating the parameters. Our approach is to use all available information to improve the performance of unit root tests, which can be possible by taking full feature of nonlinear structure in the estimation of ARIMA model and taking full advantage of information contained in the nonlinear restriction (2.5) or (2.6). For the case of no restriction as for $p \leq q$ and σ_{uu} unknown or $p \leq q-1$ and σ_{uu} known, the N-R estimation is the unrestricted nonlinear least squares estimation because the maximum likelihood estimator is equivalent to the least squares estimator. Also when $p \geq q+1$ and σ_{uu} is unknown, the estimator is based on least squares estimator, which takes advantage of the nonlinear restriction $\boldsymbol{g} = \mathbf{0}$ in (2.5). However, when $p \geq q$ and σ_{uu} is known, a restricted maximum likelihood estimation is considered because σ_{ee} in (2.6) can not be separated out.

Let $\psi = (\rho, \theta')'$ and $e(t;\psi)$ be the prediction error of the best linear prediction based on $y_{t-1}, \ldots, y_1$. That is,

$$e(t;\psi) = y_t - \mathrm{E}_\psi[y_t \mid y_{t-1}, y_{t-2}, \ldots, y_1].$$

Then, using (2.2), $e_1, e_2, \ldots, e_n$ can be solved in terms of $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)'$ and ψ. The negative log likelihood function of $\boldsymbol{y}$ conditional on y_0, up to a constant, is

$$L_n(\xi) = \sigma_{ee}^{-1} S_n(\psi) + n log_e \sigma_{ee}, \text{ where } S_n(\psi) = \sum_{t=1}^{n} e^2(t;\psi).$$

When the errors are normally distributed, the least squares estimator is equivalent to the maximum likelihood estimator which minimizes $L_n(\xi)$. This is due to the fact that the restriction $\boldsymbol{g} = \mathbf{0}$ does not contain σ_{ee}. Thus, the least squares estimator minimizes the sum of squares $S_n(\psi)$. Since estimation for ARIMA model is essentially nonlinear due to the fact that derivatives of $S_n(\psi)$ are nonlinear in ψ, one usually use a N-R procedure which is a numerical procedure for solving $S_\psi = 0$. Let

$$S_\psi = \partial S_n / \partial \psi = \begin{bmatrix} \partial S_n / \partial \rho \\ \partial S_n / \partial \theta \end{bmatrix} = \begin{bmatrix} S_\rho \\ S_\theta \end{bmatrix},$$

$$S_{\psi\psi} = \partial^2 S_n / \partial \psi \partial \psi' = \begin{bmatrix} \partial^2 S_n / \partial \rho \partial \rho & \partial^2 S_n / \partial \rho \partial \theta' \\ \partial^2 S_n / \partial \theta \partial \rho & \partial^2 S_n / \partial \theta \partial \theta' \end{bmatrix} = \begin{bmatrix} S_{\rho\rho} & S_{\rho\theta} \\ S_{\theta\rho} & S_{\theta\theta} \end{bmatrix}.$$

Then Taylor approximation gives $S_\psi = \tilde{S}_\psi + (\psi - \tilde{\psi})\tilde{S}_{\psi\psi}$, where $\tilde{\psi}$ is an initial estimator, and $\tilde{S}_\psi$ and $\tilde{S}_{\psi\psi}$ are evaluated at $\tilde{\psi}$. The unrestricted N-R procedure for estimating model (2.2) is

$$\hat{\psi} = \tilde{\psi} - \tilde{S}_{\psi\psi}^{-1}\tilde{S}_\psi. \tag{3.1}$$

Since σ_{ee} can be concentrated out from the likelihood function, we have not included derivatives with respect to σ_{ee} in the N-R procedure (3.1).

When $p < q + 1$, unrestricted least squares estimator gives fully efficient estimator. When $p \geq q + 1$, the restriction $\boldsymbol{g} = \mathbf{0}$ of (2.5) gives extra information on ξ and should be incoporated in the estimation of ξ in order to produce efficient estimator. Now we modify this N-R procedure to accomodate the restriction $\boldsymbol{g} = \mathbf{0}$. The maximum likelihood estimator minimizes S_n subject to the restriction $\boldsymbol{g} = \mathbf{0}$ and hence is equivalent to minimizing the Lagrangian function

$$S_n + \boldsymbol{g}'\boldsymbol{\lambda}, \boldsymbol{\lambda} \in \mathrm{R}^{p-q}$$

with respect to ρ, θ and $\boldsymbol{\lambda}$. Differentiating and setting to zero, we get $S_\psi + \mathrm{G}'\boldsymbol{\lambda} = \mathbf{0}$ and $\boldsymbol{g} = \mathbf{0}$, where $\mathrm{G} = \partial\boldsymbol{g}/\partial\psi' = [\partial\boldsymbol{g}/\partial\rho\ \partial\boldsymbol{g}/\partial\theta'] = [\mathrm{G}_\rho\ \mathrm{G}_\theta]$. Combining Taylor expansions $S_\psi = \tilde{S}_\psi + \tilde{S}_{\psi\psi}(\psi - \tilde{\psi})$ and $\boldsymbol{g} = \tilde{\boldsymbol{g}} + \tilde{\mathrm{G}}(\psi - \tilde{\psi})$, we have $\tilde{S}_{\psi\psi}(\psi - \tilde{\psi}) + \tilde{\mathrm{G}}'\boldsymbol{\lambda} = -\tilde{S}_\psi$ and $\tilde{\mathrm{G}}$ are evaluated at $\tilde{\psi}$. Therefore, a restricted N-R procedure which takes advantage of the information in the restriction $\boldsymbol{g} = \mathbf{0}$ is

$$\begin{bmatrix} \hat{\rho} \\ \hat{\theta} \\ \hat{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} \tilde{\rho} \\ \tilde{\theta} \\ 0 \end{bmatrix} - \begin{bmatrix} \tilde{S}_{\rho\rho} & \tilde{S}_{\rho\theta} & \tilde{\mathrm{G}}'_\rho \\ \tilde{S}_{\theta\rho} & \tilde{S}_{\theta\theta} & \tilde{\mathrm{G}}'_\theta \\ \tilde{\mathrm{G}}_\rho & \tilde{\mathrm{G}}_\theta & 0 \end{bmatrix}^{-1} \begin{bmatrix} \tilde{S}_\rho \\ \tilde{S}_\theta \\ \tilde{\boldsymbol{g}} \end{bmatrix}. \tag{3.2}$$

Note that the algorithm of computing the N-R estimator from (3.2) is much simpler than that for computing the modified maximum likelihood estimator given in Shin and Sarkar(1994, § 4). Explicit expressions for G_θ and approximations to $S_\rho, S_\theta, S_{\rho\rho}, S_{\rho\theta}$ and $S_{\theta\theta}$ are given in the appendix.

Having obtained estimator $\hat{\psi}$ by (3.2) we can compute $\hat{\sigma}_{ee} = \sum e^2(t;\hat{\psi})/(n - 2p - 1)$ and $\hat{\sigma}_{uu} = \hat{\sigma}_{ee}\hat{\gamma}_{p+1}/\hat{\alpha}^*_{p+1}$. Also $(\hat{\beta}, \hat{\sigma}_{aa})$ can be computed by applying the Wilson's (1969) algorithm to $f_h = 0\ h = 0, 1, \ldots, q$.

However, when σ_{uu} is known and $p > q-1$, σ_{ee} cannot be eliminated from the restriction. Therefore, σ_{ee} should not be concentrated out from the likelihood function and we should use the maximum likelihood estimator to get efficient estimator which include σ_{ee} in the N-R procedure like

$$\begin{bmatrix} \hat{\rho} \\ \hat{\theta} \\ \hat{\sigma}_{ee} \\ \hat{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} \tilde{\rho} \\ \tilde{\theta} \\ \tilde{\sigma}_{ee} \\ 0 \end{bmatrix} - \begin{bmatrix} \tilde{L}_{\rho\rho} & \tilde{L}_{\rho\theta} & \tilde{L}_{\rho\sigma} & \tilde{\mathrm{F}}'_{\rho} \\ \tilde{L}_{\theta\rho} & \tilde{L}_{\theta\theta} & \tilde{L}_{\theta\sigma} & \tilde{\mathrm{F}}'_{\theta} \\ \tilde{L}_{\sigma\rho} & \tilde{L}_{\sigma\theta} & \tilde{L}_{\sigma\sigma} & \tilde{\mathrm{F}}'_{\sigma} \\ \tilde{\mathrm{F}}_{\rho} & \tilde{\mathrm{F}}_{\theta} & \tilde{\mathrm{F}}_{\sigma} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \tilde{L}_{\rho} \\ \tilde{L}_{\theta} \\ \tilde{L}_{\sigma} \\ \tilde{\boldsymbol{f}} \end{bmatrix}, \tag{3.3}$$

where $L_\xi = \partial L_n/\partial\xi$, $L_{\xi\xi} = \partial^2 L_n/\partial\xi\partial\xi'$, $\boldsymbol{f} = (f_{q+1}, \ldots, f_{p+1})'$ and $\mathrm{F}_\xi = \partial \boldsymbol{f}/\partial\xi'$.

Testing for a unit root hypothesis $H_0 : \rho = 1$ can be performed by $n(\hat{\rho}-1)$ or $\hat{\tau} = (\hat{\rho}-1)/\sqrt{c_{11}\hat{\sigma}_{ee}}$, where $\hat{\rho}$ is computed from (3.1), (3.2), or (3.3), c_{11} is the (1,1) element of the inverted matrix in (3.3).

4. LIMITING DISTRIBUTIONS OF THE TEST STATISTICS

In this section, we show that the limiting distribution of $n(\hat{\rho}-1)$ is the same as that of Dickey and Fuller(1979) for AR(1) model with no measurement error.

Assume that $\tilde{\psi}$ is $n^{-1/2}$-consistent, i.e., $(\tilde{\rho}-1) = O_p(n^{-1/2})$, $(\tilde{\theta}-\theta) = O_p(n^{-1/2})$. We first consider the estimator (3.2) for unknown σ_{uu}. Combining $\tilde{S}_\psi = S_\psi + \tilde{S}_{\psi\psi}(\tilde{\psi}-\psi) + O_p(1)$ and $\tilde{\boldsymbol{g}} = \boldsymbol{g} + \tilde{\mathrm{G}}(\tilde{\psi}-\psi) + O_p(1)$ with (3.2) we get

$$\begin{bmatrix} \hat{\rho}-1 \\ \hat{\theta}-\theta \\ \hat{\boldsymbol{\lambda}} \end{bmatrix} = - \begin{bmatrix} \tilde{S}_{\rho\rho} & \tilde{S}_{\rho\theta} & \tilde{\mathrm{G}}'_{\rho} \\ \tilde{S}_{\theta\rho} & \tilde{S}_{\theta\theta} & \tilde{\mathrm{G}}'_{\theta} \\ \tilde{\mathrm{G}}_{\rho} & \tilde{\mathrm{G}}_{\theta} & 0 \end{bmatrix}^{-1} \begin{bmatrix} S_\rho + O_p(1) \\ S_\theta + O_p(1) \\ O_p(1) \end{bmatrix}.$$

Note that $\tilde{S}_{\rho\theta} = O_p(n)$, $n^{-2}\tilde{S}_{\rho\rho} \xrightarrow{d} \sigma_{ee}\boldsymbol{G}$, $n^{-1}S_\rho \xrightarrow{d} -2^{-1}\sigma_{ee}(\boldsymbol{T}^2-1)$ by Shin and Sarkar(1993, p.202) and $\tilde{\mathrm{G}}_\rho = O_p(1)$. By Brockwell and Davis(1990, p.267),

$n^{-1/2}S_\theta \xrightarrow{d} \mathrm{N}(0, \sigma_{ee}^2 \mathrm{I}(\theta))$ and $n^{-1}\tilde{S}_{\theta\theta} \xrightarrow{p} \sigma_{ee}\mathrm{I}(\theta)$, and $\tilde{\mathrm{G}}_\theta \xrightarrow{p} \mathrm{G}_\theta$, where $\mathrm{I}(\theta)$ is the information matrix of θ in the unrestricted model (2.2). Also,

$$(n^{-2}\tilde{S}_{\rho\rho}, n^{-1}S_\rho, n^{-1}S_\theta) \xrightarrow{d} (\sigma_{ee}\mathrm{G}, -2^{-1}\sigma_{ee}(T^2-1), \mathrm{N}(0, \sigma_{ee}^2\mathrm{I}(\theta))). \qquad (4.1)$$

The joint convergence in distribution in (4.1) is justified by Yap and Reinsel(1995, (6)-(8)) and Chan and Wei(1988, Theorem 2.2). Now we have

$$\begin{bmatrix} n(\hat{\rho}-1) \\ n^{1/2}(\hat{\theta}-\theta) \\ n^{-1/2}\hat{\lambda} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \boldsymbol{G} & 0 & 0 \\ 0 & \mathrm{I}(\theta) & \mathrm{G}_\theta' \\ 0 & \mathrm{G}_\theta & 0 \end{bmatrix}^{-1} \begin{bmatrix} 2^{-1}(\boldsymbol{T}^2-1) \\ -\mathrm{N}\left(\mathbf{0}, \begin{pmatrix} \mathrm{I}(\theta) & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}\right) \end{bmatrix}. \qquad (4.2)$$

Therefore, the limiting distribution of $n(\hat{\rho}-1)$ and $\hat{\tau}$ are

$$2^{-1}\boldsymbol{G}^{-1}(\boldsymbol{T}^2-1) \text{ and } 2^{-1}\boldsymbol{G}^{-1/2}(\boldsymbol{T}^2-1), \qquad (4.3)$$

respectively, where $(\boldsymbol{G}, \boldsymbol{T}) = (\sum_{i=1}^\infty d_i^2\epsilon_i^2, \sum_{i=1}^\infty 2^{1/2}d_i\epsilon_i)$, $d_i = (-1)^{i+1}2[(2k-1)\pi]^{-1}$, ϵ_i's are i.i.d. $\mathrm{N}(0,1)$ random variables. Since the limiting distribution of $n(\hat{\rho}-1)$ is the same as that of Dickey and Fuller(1976), one can use the table of Fuller(1976) to perform unit root test $H_0 : \rho = 1$. From (4.2), we have the following.

Theorem 1. Assume that a time series follows the model (2.2) with the restriction (2.5). We also assume that the e_t's are normal and $\boldsymbol{g}$ is continuously differentiable. Assume that the initial estimator of $\tilde{\theta}$ of θ satisfies $n^{1/2}(\tilde{\theta}-\theta) = O_p(1)$. Let $(\hat{\theta}', \hat{\boldsymbol{\lambda}}')'$ be as defined in (3.2). Then the asymptotic distribution of $[n^{1/2}(\hat{\theta}-\theta)', n^{-1/2}\hat{\boldsymbol{\lambda}}']'$ is

$$\mathrm{N}\left(0, \begin{bmatrix} \mathrm{I}(\theta) & \mathrm{G}_\theta' \\ \mathrm{G}_\theta & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{I}(\theta) & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathrm{I}(\theta) & \mathrm{G}_\theta' \\ \mathrm{G}_\theta & 0 \end{bmatrix}^{-1}\right). \qquad (4.4)$$

Using the formular for inverses of partitioned matrices, one may easily show that $\begin{bmatrix} \mathrm{I}(\theta) & \mathrm{G}_\theta' \\ \mathrm{G}_\theta & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{I}(\theta) & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathrm{I}(\theta) & \mathrm{G}_\theta' \\ \mathrm{G}_\theta & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \mathrm{I}(\theta)^{-1} - \boldsymbol{JKJ}' & \mathbf{0} \\ \mathbf{0} & \boldsymbol{K} \end{bmatrix}$ whose upper-left $(2p+1)\times(2p+1)$ block, say, $\boldsymbol{V}$ is the same as that of $\begin{bmatrix} \mathrm{I}(\theta) & \mathrm{G}_\theta' \\ \mathrm{G}_\theta & 0 \end{bmatrix}^{-1}$, where $\boldsymbol{J} = \mathrm{G}_\theta \mathrm{I}(\theta)^{-1}$ and $\boldsymbol{K} = [\mathrm{G}_\theta \mathrm{I}(\theta)^{-1}\mathrm{G}_\theta']^{-1}$. Note that the asymptotic covariance matrix $\boldsymbol{V}$ has much simpler expression compared with that of Pagano(1974, p.101), Sakai and Arase(1979, equation 38), and Shin(1993, p.633).

We consider the case of known σ_{uu}. Note that $n^{1/2}(n^{-1}S_n - \sigma_{ee}) \xrightarrow{d} \mathrm{N}(0, 2\sigma_{ee}^2)$. Using the above results, the limiting distribution of $n(\hat{\rho} - 1)$ and $\hat{\tau}$ are $2^{-1}\boldsymbol{G}^{-1}(\boldsymbol{T}^2 - 1)$ and $2^{-1}\boldsymbol{G}^{-1/2}(\boldsymbol{T}^2 - 1)$, respectively. Also we have the following.

Corollary 1. Assume that a time series follows the model (2.2) with the restriction (2.6). We also assume that the e_t's are normal and $\boldsymbol{f}$ is continuously differentiable. Assume that the initial estimator of $\tilde{\xi}$ of ξ satisfies $n^{1/2}(\tilde{\xi} - \xi) = O_p(1)$. Let $(\hat{\theta}', \hat{\sigma}_{ee}, \hat{\boldsymbol{\lambda}}')'$ be as defined in (3.3). Then the asymptotic distribution of $[n^{1/2}(\hat{\theta} - \theta)', n^{1/2}(\hat{\sigma}_{ee} - \sigma_{ee}), n^{-1/2}\hat{\boldsymbol{\lambda}}']'$ is $\mathrm{N}(0, \mathrm{V}^{-1}(0)\ \mathrm{V}(1)\ \mathrm{V}^{-1}(0))$, where

$$\mathrm{V}(\delta) = \begin{bmatrix} \mathrm{I}(\theta) & 0 & \mathrm{F}_\theta' \\ 0 & (1+\delta)\sigma_{ee}^{-2} & \mathrm{F}_\sigma' \\ \mathrm{F}_\theta & \mathrm{F}_\sigma & 0 \end{bmatrix}, \tag{4.5}$$

$\mathrm{F}_\theta = \partial \boldsymbol{f}/\partial\theta'$ and $\mathrm{F}_\sigma = \partial \boldsymbol{f}/\partial\sigma_{ee}$.

REFERENCES

(**1**) Brockwell, P.J. and Davis, R.A. (1990). *Time Series - Theory and Methods*. Springer-Verlag, New York.

(**2**) Chan, N.H. and Wei, C.Z. (1988). Limiting Distributions of Least Squares Estimates of Unstable Autoregressive Process. *The Annals of Statistics*, **16**, 367–401.

(**3**) Dickey, D.A. and Fuller, W.A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, **74**, 427–431.

(4) Fuller, W.A. (1976). *Introduction to Statistical Time Series.* John Wiley and Sons, New York.

(5) Hall, A. (1989). Testing for a Unit Root in the Presence of Moving Average Errors. *Biometrika*, **76**, 49–56.

(6) Pagano, M. (1974). Estimation of Models of Autoregressive Signal Plus White Noise. *The Annals of Statistics*, **2**, 99–108.

(7) Pantular, S.G. and Hall, A. (1991). Testing for Unit Roots in Autoregressive Moving Average Models. An Instrumental Variable Approach. *Journal of Econometrics*, **48**, 325–353.

(8) Phillips, P.C.B. (1987). Time Series Regression with a Unit Root. *Econometrica*, **55**, 277–301.

(9) Phillips, P.C.B. and Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika*, **75**, 335–346.

(10) Said, S.E. and Dickey, D.A. (1985). Hypothesis Testing in ARIMA$(p, 1, q)$ Models. *Journal of the American Statistical Association*, **80**, 369–374.

(11) Sakai, H. and Arase, M. (1979). Recursive Parameter Estimation of an Autoregressive Process Disturbed by White Noise. *International Journal of Control*, **30**, 949–966.

(12) Shin, D.W. (1993). Maximum Likelihood Estimation for Autoregressive Process Disturbed by a Moving Average. *Journal of the Time Series Analysis*, **14**, 629–644.

(13) Shin, D. and Sarkar, S. (1993). A Note on Testing for a Unit Root in an ARIMA$(p, 1, 0)$ Signal Observed with MA(q) Noise. *Statistics & Probability Letters*, **18**, 195–203.

(14) Shin, D.W. and Sarkar, S. (1994). Testing for a Unit Root in an ARIMA$(p,1,0)$ Signal Observed with MA(q) Noise. *Communications in Statistics - Theory and Methods*, **23(9)**, 2643–2670.

(15) Wilson, G. (1969). Factorization of the Covariance Generating Function of a Pure Moving Average Process. *SIAM Journal of Numerical Analysis*, **6**, 1–7.

(16) Yap, S.F. and Reinsel, G.C. (1995). Results on Estimation and Testing for a Unit Root in the Nonstationary Autoregressive Moving-Average Model. *Journal of the Time Series Analysis*, **16**, 339–353.

APPENDIX

We discuss how to get simple computations for the estimators in (3.1), (3.2), and (3.3). For that, we need simple expressions for the terms in (3.1)–(3.2).

We first give approximations to S_ρ, S_θ, $S_{\rho\rho}$, $S_{\rho\theta}$, and $S_{\theta\theta}$. We can approximate $e(t;\psi)$ following the approach of Fuller(1987, §8.3). Letting $y_t = e_t = 0$ for $t \le 0$, usually $e(t;\psi)$'s are approximated by

$$e_t(\psi) = \sum_{k=0}^{p} \alpha_k (y_{t-k} - \rho y_{t-1-k}) - \sum_{k=1}^{p+1} \gamma_k e_{t-k}(\psi),$$

for $t = 1, \ldots, n$. Then derivatives of $e(t;\psi)$ are approximated by $W_{\rho,t}(\psi) = \partial e_t(\psi)/\partial\rho$, $W_{\alpha,t}(\psi) = \partial e_t(\psi)/\partial\alpha$, and $W_{\gamma,t}(\psi) = \partial e_t(\psi)/\partial\gamma$, which are computed recursively

$$\begin{aligned}
W_{\rho,t}(\psi) &= -\sum_{l=0}^{p} \alpha_l y_{t-1-l} - \sum_{l=1}^{p+1} \gamma_l W_{\rho,t-l}(\psi), \\
W_{\alpha_j,t}(\psi) &= y_{t-j} - \rho y_{t-1-j} - \sum_{l=1}^{p+1} \gamma_l W_{\alpha_j,t-l}(\psi), j = 1, \ldots, p, \\
W_{\gamma_j,t}(\psi) &= -e_{t-j}(\psi) - \sum_{l=1}^{p+1} \gamma_l W_{\gamma_j,t-l}(\psi), j = 1, \ldots, p+1,
\end{aligned}$$

where $W_{\rho,t}(\psi) = W_{\alpha_j,t}(\psi) = W_{\gamma_j,t}(\psi) = 0$, for $t \leq 0$. Differentiating $2^{-1}S_n(\psi)$, we have $2^{-1}S_\rho = \sum W_{\rho,t}(\psi)e_t(\psi)$ and $2^{-1}S_\theta = \sum W_{\theta,t}(\psi)e_t(\psi)$, where $W_{\theta,t}(\psi) = [W'_{\alpha,t}(\psi)W'_{\gamma,t}(\psi)]'$. In differentiating S_ρ and S_θ, taking terms of $O_p(n)$ and ignoring terms of $O_p(n^{1/2})$, approximations of $S_{\rho\rho}$, $S_{\rho\theta}$, and $S_{\theta\theta}$ are given by $S_{\rho\rho} \cong \sum W^2_{\rho,t}(\psi)$, $S_{\rho\theta} \cong \sum W_{\rho,t}(\psi)W'_{\theta,t}(\psi)$, $S_{\theta\theta} \cong \sum W_{\theta,t}(\psi)W'_{\theta,t}(\psi)$. Also we have approximations of $L_\xi \cong [\sigma_{ee}^{-1}S_\rho, \sigma_{ee}^{-1}S_\theta, -\sigma_{ee}^{-2}(S_n(\psi) - n\sigma_{ee})]'$ and $L_{\xi\xi}$ from the approximations of $S_{\rho\rho}$, $S_{\rho\theta}$, $S_{\theta\theta}$, σ_{ee}.

We next give explicit expressions for G_θ. Direct differentiation gives for $h = 0, 1, \ldots, k$ and $i = 1, 2, \ldots, k$,

$$
\begin{aligned}
\partial g_h/\partial\alpha_i^* &= -\gamma_{p+1}(\alpha_{h+i}^* + \alpha_{-h+i}^*) \\
\partial g_h/\partial\gamma_i &= \alpha_{p+1}^*(\gamma_{h+i} + \gamma_{-h+i}) \\
\partial g_h/\partial\alpha_{p+1}^* &= \sum_{j=0}^{p+1-h} \gamma_j\gamma_{j+h} - \gamma_{p+1}(\alpha_{h+i}^* + \alpha_{-h+i}^*) \\
\partial g_h/\partial\gamma_{p+1} &= \alpha_{p+1}^*(\gamma_{h+i} + \gamma_{-h+i}) - \sum_{j=0}^{p+1-h} \alpha_j^*\alpha_{j+h}^*.
\end{aligned}
$$