8307776

Pantula, Sastry Gouripathi

PROPERTIES OF ESTIMATORS OF THE PARAMETERS OF AUTOREGRESSIVE TIME SERIES

Iowa State University PH.D. 1982

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

Properties of estimators of the parameters

of autoregressive time series

by

Sastry Gouripathi Pantula

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1982

TABLE OF CONTENTS

CHAPTER I. INTRODUCTION

Many time series encountered in practice are well-approximated by the representation

$$Y_t = \sum_{i=1}^{r} X_{ti}\beta_i + P_t, \quad t = 1, 2, \ldots \tag{1.1}$$

where the $\{X_{ti}\}$ are fixed sequences and the $\{P_t\}$ is a time series with mean zero. For example, we might have $X_{t1} \equiv 1$, $X_{t2} = t$, $X_{t3} = t^2$. The X_{ti} might also be random functions of time, for example, a stationary time series. If X_{ti} is random, we shall investigate the behavior of the estimators conditional on a particular realization of X_{ti}. Thus, all X_{ti} shall be treated as fixed functions of time. It is assumed that $\{P_t\}$ is independent of $\{X_{ti}\}$. We also consider the case in which $\{P_t\}$ is a seasonal autoregressive process satisfying

$$P_t = \sum_{j=1}^{p} \alpha_j P_{t-jk} + e_t, \tag{1.2}$$

where $\{e_t\}$ is a sequence of uncorrelated $(0,\sigma^2)$ random variables and k is the seasonal period. The values of k that are commonly used are 1, 4, 12 corresponding to yearly, quarterly and monthly observations.

Given a realization $\{Y_t;\ t = 1, 2, \ldots, nk\}$ of nk observations, the least squares procedure is commonly used to estimate the parameters of the seasonal autoregressive process. Under the assumption of normality the method of maximum likelihood is appealing, but is difficult to compute in all but the simplest case of a first-order seasonal autoregressive process with known means.

The asymptotic properties of the least squares estimators depend upon (i) the properties of the $\{e_t\}$ sequence, (ii) the roots of the characteristic equation

$$m^p - \alpha_1 m^{p-1} - \dots - \alpha_p = 0,$$

(iii) the initial conditions $Y_0, Y_{-1}, \dots, Y_{-p+1}$ and (iv) the properties of $\{X_{ti}\}$. In this study, the asymptotic properties of the least squares estimators are examined under a wide variety of assumptions.

In the majority of the situations, the least squares estimators are consistent and asymptotically normal, but are biased in small samples. In econometric work, small sample sizes ranging from 5 to 20 years are frequently encountered. For such samples the bias in the least squares estimators of the autoregressive coefficients is appreciable in magnitude.

Consider the p-th order stationary autoregressive process with period $k = 1$ which satisfies the stochastic difference equation

$$Y_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j Y_{t-j} + e_t \tag{1.3}$$

$$= \mu + \sum_{j=1}^{p} \alpha_j (Y_{t-j} - \mu) + e_t \tag{1.4}$$

where μ is the mean of the time series $\{Y_t\}$ and $\{e_t\}$ is a sequence of independent $(0,\sigma^2)$ random variables. Note that $\alpha_0 = \mu(1-\sum_{j=1}^{p} \alpha_j)^{-1}$. The least squares estimator of $\underset{\sim}{\alpha} = (\alpha_1,\alpha_2, \ldots, \alpha_p)'$ is obtained by regressing $Y_t - \overline{Y}$ on $Y_{t-1} - \overline{Y}$, $Y_{t-2} - \overline{Y}$, ..., $Y_{t-p} - \overline{Y}$, where $\overline{Y} = n^{-1} \sum_{t=1}^{n} Y_t$. In this study, approximate expressions for the bias in the least squares estimator of $\underset{\sim}{\alpha}$ that is due to replacing μ by $\overline{Y}$ are derived. Using the approximate expressions for the bias, modifications of the least squares estimator are proposed. This method of bias correction is extended to the model given in (1.1) for the case $k = 1$. The method is particularly suitable for the case where the X_{ti} are polynomials in time. Estimators are also given for the stationary p-th order seasonal model.

Two Monte Carlo studies examining the small sample properties of various estimators of the parameters of second-order autoregressive processes are considered. A second-order autoregressive process with constant mean, and a second-order autoregressive process with mean function linear in time are considered. Generally speaking, the modified estimators performed better than the least squares estimator.

CHAPTER II. THE ASYMPTOTIC PROPERTIES OF THE LEAST SQUARES ESTIMATOR OF THE PARAMETER OF THE FIRST-ORDER AUTOREGRESSIVE MODEL

Literature Review

A casual inspection of many economic time series leads one to conclude that the observations are not independent. In recent years, autoregressive moving average processes have been proposed for modeling economic data. See Box and Jenkins (1976), Box, Hillmer and Tiao (1976), Fuller (1976), Jenkins and Watts (1968), and Parzen and Pagano (1977). With the advent of the computer, the autoregressive moving average schemes are widely accepted as a reliable method for estimating and predicting the behavior of a real process.

Yule (1927), Walker (1931), and Slutsky (1937) first formulated the concept of autoregressive moving average schemes. In 1938, Wold (1954) obtained a general representation for time series. Since then, a considerable body of literature in the area of time series dealing with the parameter estimation and the order determination of time series models has appeared. More recently, Jenkins and Watts (1968), and Box and Jenkins (1976) extended the autoregressive moving average processes to include seasonal time series.

Most of the results in time series deal with stationary processes. A stochastic process defined on T is said to be (weakly) stationary if its first and second moments exist and

(i) $E\left\{Y_t\right\} = \mu$,

(ii) $E\left\{(Y_t-\mu)(Y_{t+h}-\mu)\right\} = \gamma(h)$, for all t, $t + h$ in T. The autocorrelation function of $\{Y_t\}$ is defined as

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} . \tag{2.1}$$

Much of the early work in time series was concerned with estimating the autocorrelations and in deriving tests of the hypothesis of independence.

Many processes that occur in practice can be well-approximated by the autoregressive process of order p satisfying the stochastic difference equation

$$Y_t = \sum_{i=1}^{r} X_{ti} \beta_i + \sum_{j=1}^{p} \alpha_j (Y_{t-j} - \sum_{i=1}^{r} X_{t-j,i} \beta_i) + e_t \tag{2.2}$$

$$= \sum_{i=1}^{r} \psi_{ti} \beta_i + \sum_{j=1}^{p} \alpha_j Y_{t-j} + e_t \tag{2.3}$$

$t = p + 1, p + 2, \ldots, n$, where the $\{e_t\}$ are uncorrelated $(0, \sigma^2)$ random variables and $Y_1, Y_2, \ldots, Y_p$ are initial conditions. It is assumed $\alpha_p \neq 0$. Let $m_1, m_2, \ldots, m_p$ be the roots of the characteristic equation

$$m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p = 0. \tag{2.4}$$

The parameters of the model and the variance of e_t are to be estimated from an observed sequence $Y_1, Y_2, \ldots, Y_n$. The sampling theory approach to the estimation problem of an autoregressive process has generally been analogous to the treatment of univariate regression model. The ordinary least squares procedure provides the best linear unbiased estimators in the classical linear regression model. The assumptions of the Gauss-Markov theorem are not met in the autoregressive case since lagged values of the dependent variables are not distributed independently of the error term for all lags. Under the assumption of normal errors, the conditional maximum likelihood estimators, conditonal on $Y_1, Y_2, \ldots, Y_p$, of the autoregressive parameters are the least squares estimators. Several other asymptotically equivalent estimators are considered in the next chapter.

Mann and Wald (1943) considered estimation of the parameters of the model (2.3) with the $\{\psi_{ti}\}$ restricted to a constant and the roots of (2.4) less than one in absolute value. Assuming $\{e_t\}$ to be a sequence of normal independent $(0,\sigma^2)$ random variables, they established that the asymptotic distribution of the least squares estimator is normal.

White (1958) obtained the limiting joint moment generating function of the numerator and the denominator of the least squares estimator of α_1 for the case $p = 1$ and no ψ - variables. The moment generating function had three forms, according as the root of the characteristic equation was less than one, equal to one, or greater than one in absolute value.

Anderson (1959) extended Mann and Wald's result to the case where $\{e_t\}$ are assumed to be independent $(0, \sigma^2)$ random variables with bounded $(2+\delta)$-th moments, for some $\delta > 0$. He also studied the case when at least one of the roots of the characteristic equation is greater than one in absolute value.

Rao (1961), Venkataraman (1967, 1968, and 1973), Narasimham (1969), and Stigum (1974) have studied estimation of the model when at least one of the roots of the characteristic equation is greater than one in absolute value.

The limiting behavior of the least squares estimator for a model with fixed ψ - variables and roots of the characteristic equation less than one in absolute value has been investigated by several authors. Among the first to consider the statistical properties of this model were workers at the Cowles Commission; see Anderson and Rubin (1950), Koopmans, Rubin, and Leipnik (1950), and Rubin (1950). Hannan (1965), Amemiya and Fuller (1967), Hatanaka (1974), and Fuller (1976) studied the situation in which there are nonlinear restrictions on the parameters arising from the specification of autocorrelated errors and lagged dependent variables in the equation. Hannan and Heyde (1972), Hannan and Nicholls (1972), Reinsel (1976), Fuller (1976), Anderson and Taylor (1979), Crowder (1980), and Fuller, Hasza, and Goebel (1981) considered estimation of model (2.3) with the roots of (2.4) less than one in absolute value.

Fuller, Hasza, and Goebel assumed that $\{e_t\}$ are independent $(0,\sigma^2)$ random variables with bounded $(2+\delta)$-th moments. Crowder considered the case where $\{e_t\}$ is a sequence of martingale differences. An extension of the results of Fuller, Hasza and Goebel for the stationary case with $\{e_t\}$ a martingale difference sequence and the results of Crowder (1980) are presented in Appendix B.

Dickey (1976), Fuller (1976), and Dickey and Fuller (1979) considered the estimation of equation (2.3) assuming one of the roots of the characteristic equation to be one and permitted the set $\{\psi_{ti}\}$ to include the constant function and time. Hasza (1977) discussed the estimation of equation (2.3) with one of the roots of the characteristic equation greater than one in absolute value. Hasza permitted a set $\{\psi_{ti}\}$ composed of polynomial function of time to enter the equation.

Fuller, Hasza, and Goebel (1981) established the limiting distributions of the least squares estimators of the parameters of (2.2) for cases in which the largest root is less than one, equal to one, and greater than one in absolute value, assuming $\{e_t\}$ to be a sequence of independent $(0,\sigma^2)$ random variables with bounded $(2+\delta)$-th moments. They established that,

(a) if all the roots of the characteristic equation are less than one in absolute value, then the limiting distribution of the least squares estimator is normal under mild regularity conditions,

(b) if one of the roots of the characteristic equation is one and the others are less than one in absolute value, then the limiting distribution depends upon the nature of the set $\{\psi_{ti}\}$ and upon the parameters in the model,

(c) if one of the roots of (2.4) is greater than one in absolute value and the remaining roots are less than one in absolute value, the least squares estimators normalized by the square roots of the sums of squares of the explanatory variables are normal if and only if the e_t are normal independent random variables.

Asymptotic Properties Of The Least Squares Estimator
For The Case $p = 1$ And $\alpha_1 = 1$

In this section, we establish the limiting distribution of the least squares estimator under the assumption that $\{e_t\}$ is a sequence of martingale differences. It is proven that the limiting distribution is the same as that obtained by Dickey (1976) under the assumption that $\{e_t\}$ is a sequence of independent random variables.

Consider the following three models:

$$\text{(i)} \quad Y_t = \rho Y_{t-1} + e_t, \quad t = 1, 2, \ldots \qquad (2.5)$$
$$Y_0 = 0,$$

$$\text{(ii)} \quad Y_t = \mu + \rho Y_{t-1} + e_t, \quad t = 1, 2, \ldots \qquad (2.6)$$
$$Y_0 = 0,$$

and,

$$\text{(iii)} \quad Y_t = \mu + \beta t + \rho Y_{t-1} + e_t, \quad t = 1, 2, \ldots \qquad (2.7)$$
$$Y_0 = 0.$$

We assume n observations $Y_1, Y_2, \ldots, Y_n$ are available. Define the $(n - 1)$ dimensional vectors.

$$\underset{\sim}{1} = (1, 1, \ldots, 1)',$$

$$\underset{\sim}{t} = (1 - \tfrac{n}{2}, 2 - \tfrac{n}{2}, \ldots, n - 1 - \tfrac{n}{2})',$$

$$\underset{\sim}{Y}_t = (Y_2, Y_3, \ldots, Y_n)',$$

and,

$$\underset{\sim}{Y}_{t-1} = (Y_1, Y_2, \ldots, Y_{n-1})'.$$

Let $\underset{\sim}{U}_1 = \underset{\sim}{Y}_{t-1}$, $\underset{\sim}{U}_2 = (\underset{\sim}{1}, \underset{\sim}{Y}_{t-1})$, and $\underset{\sim}{U}_3 = (\underset{\sim}{1}, \underset{\sim}{t}, \underset{\sim}{Y}_{t-1})$.
Define,

$$\hat{\rho} = (\underset{\sim}{U}_1' \underset{\sim}{U}_1)^{-1} \underset{\sim}{U}_1' \underset{\sim}{Y}_t, \tag{2.8}$$

$$\hat{\rho}_\mu = \underset{\sim}{d}_2'(\underset{\sim}{U}_2' \underset{\sim}{U}_2)^{-1} \underset{\sim}{U}_2' \underset{\sim}{Y}_t, \tag{2.9}$$

and,

$$\hat{\rho}_\tau = \underset{\sim}{d}_3'(\underset{\sim}{U}_3' \underset{\sim}{U}_3)^{-1} \underset{\sim}{U}_3' \underset{\sim}{Y}_t \tag{2.10}$$

where $\underset{\sim}{d}_2 = (0,1)'$ and $\underset{\sim}{d}_3 = (0, 0, 1)'$.

The statistics analogous to the regression t statistics for the test of the hypothesis that $\rho = 1$ are

$$\hat{\tau} = (\hat{\rho} - 1)(S_{e1}^2\ c_1)^{-1/2}, \tag{2.11}$$

$$\hat{\tau}_\mu = (\hat{\rho}_\mu - 1)(S_{e2}^2\ c_2)^{-1/2}, \tag{2.12}$$

$$\hat{\tau}_\tau = (\hat{\rho}_\tau - 1)(S_{e3}^2\ c_3)^{-1/2} \tag{2.13}$$

where S^2_{ek} is the appropriate regression residual mean square,

$$S^2_{ek} = (n-k-1)^{-1} [\underset{\sim}{Y}'_t \{\underset{\sim}{I} - \underset{\sim}{U}_k(\underset{\sim}{U}'_k \underset{\sim}{U}_k)^{-1} \underset{\sim}{U}'_k\} \underset{\sim}{Y}_t] \tag{2.14}$$

and c_k is the lower-right element of $(\underset{\sim}{U}'_k \underset{\sim}{U}_k)^{-1}$.

Assume $\{e_t\}$ is a sequence of random variables satisfying,

$$E(e_t | F_{t-1}) = 0 \qquad \text{a.e.}, \tag{2.15}$$

$$E(e_t^2 | F_{t-1}) = \sigma^2 > 0 \qquad \text{a.e.}, \tag{2.16}$$

and,

$$E(e_t^4) < \infty . \tag{2.17}$$

where F_t is the σ-field generated by $(e_1, e_2, \ldots, e_t)$. A sequence $\{e_t\}$ satisfying (2.15), (2.16), and (2.17) is considered. Some of the properties of $\{e_t\}$ are established in the following Lemma.

Lemma 2.1. Assume $\{e_t\}$ is a sequence of random variables satisfying conditions (2.15), (2.16), and (2.17). Then,

$$\mathrm{Cov}(e_t, e_j) = 0 \qquad \text{for } t \neq j,$$

$$\mathrm{Cov}(e_t^2, e_j^2) = 0 \qquad \text{for } t \neq j,$$

and

$$n^{-1} \sum_{t=1}^{n} e_t^2 \to \sigma^2 \quad \text{a.s.} .$$

Proof. We have

$$E(e_t) = E\{E(e_t \mid F_{t-1})\} = 0,$$

and,

$$E(e_t^2) = E\{E(e_t^2 \mid F_{t-1})\} = \sigma^2.$$

Therefore, for $h > 0$

$$\begin{aligned} \mathrm{Cov}(e_t, e_{t+h}) &= E(e_t\, e_{t+h}) \\ &= E\{E(e_t\, e_{t+h} \mid F_t)\} \\ &= E[e_t\, E(e_{t+h} \mid F_t)] \\ &= E[e_t\, E\{E(e_{t+h} \mid F_{t+h-1}) \mid F_t\}] \\ &= 0, \end{aligned}$$

and,

$$\begin{aligned} \mathrm{Cov}(e_t^2, e_{t+h}^2) &= E(e_t^2\, e_{t+h}^2) - \sigma^4 \\ &= E[e_t^2\, E\{E(e_{t+h}^2 \mid F_{t+h-1}) \mid F_t\}] - \sigma^4 \\ &= \sigma^2\, E(e_t^2) - \sigma^4 \\ &= 0. \end{aligned}$$

Since $\{e_t^2\}$ is a sequence of uncorrelated random variables with $E(e_t^4) < \infty$, using Theorem 5.1.2 of Chung (1974, p. 100), we get

$$n^{-1} \sum_{t=1}^{n} e_t^2 \to \sigma^2 \quad \text{a.s.} \qquad \square$$

Following the approach used by Dickey (1976), we obtain the following theorem. This theorem is a representation of the error in the estimator in terms of a transformation of the original variables.

Theorem 2.1. Assume $\{e_t\}$ satisfies (2.15), (2.16), and (2.17). Let

$$Y_t = Y_{t-1} + e_t \qquad t = 1, 2, \ldots$$
$$= 0 \qquad t = 0.$$

and let $\hat{\rho}$, $\hat{\rho}_\mu$, and $\hat{\rho}_\tau$ be defined by (2.8), (2.9), and (2.10). Assume without loss of generality that $\sigma^2 = 1$. Then,

$$n(\hat{\rho}-1) = (2\Gamma_n)^{-1} (T_n^2-1) + o_p(n^{-1/2}),$$

$$n(\hat{\rho}_\mu-1) = (2\Gamma_n-2W_n^2)^{-1} (T_n^2-1-2T_nW_n) + o_p(n^{-1/2}),$$

and

$$n(\hat{\rho}_\tau-1) = [2(\Gamma_n-W_n^2-3V_n^2)]^{-1} [(T_n-2W_n)(T_n-6V_n)-1] + o_p(n^{-1/2}),$$

where

$$\Gamma_n = n^{-2} \sum_{t=2}^{n} Y_{t-1}^2$$
$$= n^{-2} \sum_{t=1}^{n-1} \lambda_{in} Z_{in}^2,$$

$$T_n = n^{-1/2} Y_{n-1}$$
$$= \sum_{i=1}^{n-1} a_{in} Z_{in} + o_p(n^{-1/2}),$$

$$W_n = (n^{-3})^{1/2} \sum_{t=2}^{n} Y_{t-1}$$
$$= \sum_{i=1}^{n-1} b_{in} Z_{in} + o_p(n^{-1/2}),$$

$$V_n = (n^{-5})^{1/2} \sum_{j=1}^{n-1} (n-j)(j-1)\, e_j$$

$$= \sum_{i=1}^{n-1} g_{in}\, Z_{in} + O_p(n^{-1/2}),$$

$$\lambda_{in} = 1/4 \sec^2 \left[\frac{(n-i)\pi}{2n-1} \right],$$

$$\underset{\sim}{Z}_n = (Z_{1n}, Z_{2n}, \ldots, Z_{n-1,n})' = \underset{\sim}{M}_n\, \underset{\sim}{e}_n,$$

$$m_{it}(n) = (i,t)\text{-th element of } \underset{\sim}{M}_n$$

$$= 2(2n-1)^{-1/2} \operatorname{Cos}[(4n-2)^{-1} (2t-1)(2i-1)\pi],$$

$$\underset{\sim}{e}_n = (e_1, e_2, \ldots, e_{n-1})',$$

$$a_{in} = \operatorname{Cov}(T_n, Z_{in}),$$

$$b_{in} = \operatorname{Cov}(W_n, Z_{in}),$$

and,

$$g_{in} = \operatorname{Cov}(V_n, Z_{in}).$$

Proof. See Dickey (1976). For example,

$$n(\hat{\rho}-1) = n\left(\frac{\sum_{t=2}^{n} Y_{t-1} Y_t}{\sum_{t=2}^{n} Y_{t-1}^2} - 1 \right)$$

$$= \frac{n^{-1} \sum_{t=2}^{n} Y_{t-1} e_t}{n^{-2} \sum_{t=2}^{n} Y_{t-1}^2}$$

$$= \frac{n^{-1} \sum_{t=2}^{n} \left(\sum_{j=1}^{t-1} e_j \right) e_t}{\Gamma_n}$$

$$= \frac{(2n)^{-1}\left[\left(\sum_{t=1}^{n} e_t \right)^2 - \sum_{t=1}^{n} e_t^2\right]}{\Gamma_n}$$

$$= \frac{(2n)^{-1}\left[Y_n^2 - \sum_{t=1}^{n} e_t^2 \right]}{\Gamma_n}$$

By Lemma 2.1,

$$n^{-1} \sum_{t=1}^{n} e_t^2 = 1 + 0_p(n^{-1/2}).$$

Also,

$$n^{-1} Y_n^2 = T_{n+1}^2$$

$$= n^{-1} (Y_{n-1} + e_n)^2$$

$$= n^{-1} Y_{n-1}^2 + 0_p(n^{-1/2}).$$

Therefore,

$$n(\hat{\rho}-1) = \frac{\frac{1}{2}(T_n^2-1)}{\Gamma_n} + O_p(n^{-\frac{1}{2}}).$$

Also,

$$\Gamma_n = n^{-2} \sum_{t=2}^{n} Y_{t-1}^2$$

$$= n^{-2} \underset{\sim}{e}_n' \underset{\sim}{A}_n \underset{\sim}{e}_n$$

where,

$$\underset{\sim}{A}_n = \begin{pmatrix} n-1 & n-2 & n-3 & \cdots & 1 \\ n-2 & n-2 & n-3 & \cdots & 1 \\ n-3 & n-3 & n-3 & \cdots & 1 \\ \vdots & & & & \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} = \underset{\sim}{M}_n' \underset{\sim}{\Lambda}_n \underset{\sim}{M}_n,$$

$$\underset{\sim}{\Lambda}_n = \mathrm{diag}(\lambda_{1n}, \lambda_{2n}, \ldots, \lambda_{n-1,n}),$$

λ_{in} are the eigenvalues of $\underset{\sim}{A}_n$ and $\underset{\sim}{M}_n$ consists of the eigenvectors of $\underset{\sim}{A}_n$.

Therefore,

$$\Gamma_n = n^{-2} \sum_{i=1}^{n-1} \lambda_{in} Z_{in}^2. \qquad \square$$

The following results will be used in the derivation of the limiting distribution of $(T_n, W_n, V_n, \Gamma_n)$.

<u>Lemma 2.2.</u> Suppose $\{Z_{in}:\ i = 1,2,\ldots,n-1;\ n = 1,2,\ldots\}$ is a

triangular array of random variables. Suppose $E(Z_{in}) = 0$, $V(Z_{in}) = \sigma^2$, and $\mathrm{Cov}(Z_{in}, Z_{jn}) = 0$ for $i \neq j$. Let $\{w_i\colon\ i = 1,2,\ldots\}$ be a sequence of real numbers and let $\{w_{in}\colon\ i = 1,2,\ldots,n-1;\ n = 1,2,\ldots\}$ be a triangular array of real numbers. If

$$\sum_{i=1}^{\infty} w_i^2 < \infty,$$

$$\lim_{n\to\infty} \sum_{i=1}^{n} w_{in}^2 = \sum_{i=1}^{\infty} w_i^2$$

and,

$$\lim_{n\to\infty} w_{in} = w_i,$$

then,

$$\sum_{i=1}^{n-1} w_{in} Z_{in} = \sum_{i=1}^{n-1} w_i Z_{in} + o_p(1).$$

<u>Proof.</u> Let $\sum_{i=1}^{\infty} w_i^2 = A > 0.$

Define

$$\eta_{in} = \frac{w_{in}}{\{\sum_{t=1}^{n-1} w_{tn}^2\}^{1/2}}.$$

Then, as $n\to\infty$.

$$\eta_{in} \to w_i A^{-1/2} = \eta_i \text{ (say).}$$

We will show that $\sum_{t=1}^{n-1} (\eta_t - \eta_{tn})^2$ converges to zero as n tends to infinity. Note,

$$\sum_{i=1}^{n-1} \eta_{in}^2 = 1,$$

and

$$\sum_{i=1}^{\infty} \eta_i^2 = 1.$$

For $M > 0$,

$$\sum_{t=1}^{M} \eta_{tn}^2 = \sum_{t=1}^{M} \eta_t^2 + 2 \sum_{t=1}^{M} \eta_t(\eta_{tn} - \eta_t) + \sum_{t=1}^{M} (\eta_{tn} - \eta_t)^2.$$

For a given $\varepsilon > 0$, choose M large such that

$$\sum_{t=M+1}^{\infty} \eta_t^2 < \varepsilon.$$

For this choice of M, choose N such that for $n > N > M$

$$\sum_{t=1}^{M} (\eta_t - \eta_{tn})^2 < \frac{\varepsilon^2}{4}.$$

This is possible because $\eta_{tn} \to \eta_t$ as $n \to \infty$. Then, for $n > N$

$$\sum_{t=1}^{M} \eta_t^2 > 1 - \varepsilon,$$

$$\sum_{t=1}^{M} (\eta_t - \eta_{tn})^2 < \tfrac{1}{4}\varepsilon^2,$$

and,

$$2 \left| \sum_{t=1}^{M} (\eta_{tn} - \eta_t)\, \eta_t \right| \leq 2\Big[\sum_{t=1}^{M} \eta_t^2 \sum_{t=1}^{M} (\eta_{tn} - \eta_t)^2 \Big]^{1/2}$$
$$\leq 2\,\varepsilon\,.$$

Therefore, for $n > N$

$$\sum_{t=1}^{M} \eta_{tn}^2 > 1 - 3\,\varepsilon,$$

and,

$$\sum_{t=M+1}^{n-1} \eta_{tn}^2 < 3\,\varepsilon.$$

It follows that, for $n > N$

$$\sum_{t=1}^{n-1} (\eta_t - \eta_{tn})^2 = \sum_{t=1}^{M} (\eta_t - \eta_{tn})^2 + \sum_{t=M+1}^{n-1} (\eta_t - \eta_{tn})^2$$

$$\leq \sum_{t=1}^{M} (\eta_t - \eta_{tn})^2 + 2 \sum_{t=M+1}^{n-1} \eta_t^2 + 2 \sum_{t=M+1}^{n-1} \eta_{tn}^2$$

$$< \tfrac{1}{4}\varepsilon^2 + 2\varepsilon + 6\varepsilon ,$$

and

$$\lim_{n\to\infty} \sum_{t=1}^{n-1} (\eta_t - \eta_{tn})^2 = 0.$$

Because

$$V\left(\sum_{t=1}^{n-1} \eta_{tn} Z_{tn} - \sum_{t=1}^{n-1} \eta_t Z_{tn}\right) = \sum_{t=1}^{n-1} (\eta_t - \eta_{tn})^2 \sigma^2$$

$$\to 0 \quad \text{as} \quad n \to \infty,$$

we have

$$\sum_{t=1}^{n-1} \eta_{tn} Z_{tn} = \sum_{t=1}^{n-1} \eta_t Z_{tn} + o_p(1).$$

Since,

$$A^{-1} \sum_{t=1}^{n-1} w_{tn}^2 \to 1 \quad \text{as} \quad n \to \infty,$$

we get

$$\sum_{i=1}^{n-1} w_{in} Z_{in} = \sum_{i=1}^{n} w_i Z_{in} + o_p(1). \qquad \square$$

Dickey (1976) obtained the following result which is used in deriving the limiting distributions.

Lemma 2.3. Let a_{in}, b_{in}, g_{in} and λ_{in} be as defined in Theorem 2.1. Then,

$$n^{-2}\lambda_{in} - \gamma_i^2 = O(n^{-2}),$$

$$\lim_{n\to\infty} a_{in} = a_i = 2^{1/2}\gamma_i,$$

$$\lim_{n\to\infty} b_{in} = b_i = 2^{1/2}\gamma_i^2,$$

$$\lim_{n\to\infty} g_{in} = g_i = 2^{3/2}\gamma_i^3,$$

$$\sum_{i=1}^{\infty} a_i^2 = 1,$$

$$\sum_{i=1}^{\infty} b_i^2 = \frac{1}{3},$$

and

$$\sum_{i=1}^{\infty} g_i^2 = \frac{1}{30},$$

where

$$\gamma_i = (-1)^{i+1}\sqrt{\gamma_i^2},$$

and

$$\gamma_i^2 = 4[(2i-1)\pi]^{-2}.$$

Proof. See Dickey (1976). □

Using Lemma 2.2 and Lemma 2.3, we obtain

$$T_n = \sum_{i=1}^{n-1} a_i Z_{in} + o_p(1), \tag{2.18}$$

$$W_n = \sum_{i=1}^{n-1} b_i Z_{in} + o_p(1), \tag{2.19}$$

and

$$V_n = \sum_{i=1}^{n-1} g_i Z_{in} + o_p(1). \tag{2.20}$$

Also,

$$\Gamma_n = n^{-2} \sum_{i=1}^{n-1} \lambda_{in} Z_{in}^2$$

$$= \sum_{i=1}^{n-1} \gamma_i^2 Z_{in}^2 + \sum_{i=1}^{n-1} (n^{-2} \lambda_{in} - \gamma_i^2) Z_{in}^2 ,$$

and, because

$$E \left| \sum_{i=1}^{n-1} (n^{-2} \lambda_{in} - \gamma_i^2) Z_{in}^2 \right| < \sum_{i=1}^{n-1} \left| n^{-2}\lambda_{in} - \gamma_i^2 \right|$$

$$\to 0 \qquad \text{as} \quad n \to \infty,$$

we get

$$\Gamma_n = \sum_{i=1}^{n-1} \gamma_i^2 Z_{in}^2 + o_p(1). \tag{2.21}$$

The limiting distribution of $\{Z_{in}\}$ is obtained in the following lemma.

Lemma 2.4. Let $\{e_t\}$ be a sequence of random variables satisfying the conditions (2.15), (2.16), and (2.17). Let $\underset{\sim}{M}_n$ be a (n-1) by (n-1) orthogonal matix with

$$\lim_{n\to\infty} \sup_{1<t<n-1} \left| m_{it}(n) \right| = 0 \qquad \text{for each fixed } i ,$$

where $m_{it}(n)$ is the (i,t)-th element of $\underset{\sim}{M}_n$. Then for a fixed k,

$$\underset{\sim}{Z}_n(k) = (Z_{1n}, Z_{2n}, \ldots, Z_{kn})' \xrightarrow{L} N(\underset{\sim}{0}, \sigma^2 \underset{\sim}{I}_k),$$

where

$$\begin{aligned} Z_n &= (Z_{1n}, Z_{2n}, \ldots, Z_{n-1,n})' \\ &= \underset{\sim}{M}_n \underset{\sim}{e}_n, \end{aligned}$$

and,

$$\underset{\sim}{e}_n = (e_1, e_2, \ldots, e_{n-1})'.$$

Proof. Let $\underset{\sim}{\eta}$ be an arbitrary vector such that $\underset{\sim}{\eta}' \underset{\sim}{\eta} = 1$. Consider

$$W_n = \underset{\sim}{\eta}' \underset{\sim}{Z}_n(k)$$

$$= \sum_{i=1}^{k} \eta_i Z_{in}$$

$$= \sum_{i=1}^{k} \eta_i \sum_{t=1}^{n-1} m_{it}(n) e_t$$

$$= \sum_{t=1}^{n-1} e_t \sum_{i=1}^{k} \eta_i m_{it}(n)$$

$$= \sum_{t=1}^{n-1} X_{tn},$$

where $X_{tn} = e_t \; \Sigma_{i=1}^{k} \eta_i m_{it}(n) = e_t d_{tn}$. We apply Theorem A.8 to obtain the result. Note that

$$E[X_{tn} \mid F_{t-1}] = 0 \quad \text{a.e.},$$

$$E[X_{tn}^2 \mid F_{t-1}] = d_{tn}^2 \sigma^2,$$

and

$$s_{nn}^2 = \sum_{t=1}^{n-1} d_{tn}^2 \sigma^2$$

$$= \sigma^2 \sum_{t=1}^{n-1} \{ \sum_{i=1}^{k} \eta_i m_{it}(n)\}^2$$

$$= \sigma^2 \sum_{i=1}^{k} \eta_i^2$$

$$= \sigma^2,$$

since $\underset{\sim}{M}_n$ is orthogonal. Therefore,

$$V_{nn}^2 = \sum_{t=1}^{n} E[X_{tn}^2 \mid F_{t-1}] \qquad \text{a.s.}$$
$$= s_{nn}^2 \qquad \text{a.s.}$$

It follows that

$$s_{nn}^{-2} V_{nn}^2 \xrightarrow{P} 1,$$

and the first condition of Theorem A.8 is established. To verify the second condition of Theorem A.8, consider

$$s_{nn}^{-2} \sum_{t=1}^{n-1} E[X_{tn}^2 \, I(|X_{tn}| > \varepsilon \, \sigma)]$$
$$= \sigma^{-2} \sum_{t=1}^{n-1} d_{tn}^2 \, E[e_t^2 \, I(|e_t d_{tn}| > \varepsilon \, \sigma)]$$
$$< \sigma^{-2} \sup_{1<t<n-1} E[e_t^2 \, I(|e_t d_{tn}| > \varepsilon \, \sigma)],$$

since $\sum_{t=1}^{n-1} d_{tn}^2 = 1$. Now,

$$\sup_{1<t<n-1} E[e_t^2 \, I(|e_t d_{tn}| > \varepsilon \, \sigma)]$$
$$< \sup_{1<t<n-1} \{E(e_t^4)\}^{1/2} \{P(|e_t d_{tn}| > \varepsilon \, \sigma)\}^{1/2}$$

$$\leq L^{1/2} \sup_{1 \leq t \leq n-1} \{\varepsilon^{-2}\sigma^{-2} E[|e_t|^2 d_{tn}^2]\}^{1/2}$$

$$\leq L^{1/2} \varepsilon^{-1} \sup_{1 \leq t \leq n-1} |d_{tn}|$$

$$\leq L^{1/2} \varepsilon^{-1} \sum_{i=1}^{k} |\eta_i| \sup_{1 \leq t \leq n-1} |m_{it}(n)|$$

$$\longrightarrow 0 \quad \text{as} \quad n \to \infty .$$

Therefore, the second condition of Theorem A.8 is satisfied and

$$\sum_{t=1}^{n-1} X_{tn} \xrightarrow{L} N(0,\sigma^2). \qquad \square$$

Since for $m_{it}(n)$ of Theorem 2.1,

$$| m_{it}(n) | \leq 2(2n-1)^{-1/2}$$

we get

$$\sup_{1 \leq t \leq n-1} |m_{it}(n)| \longrightarrow 0 \quad \text{as} \quad n \to \infty .$$

Therefore, for a fixed k

$$(Z_{1n}, Z_{2n}, \ldots, Z_{kn})' \xrightarrow{L} N(\underset{\sim}{0}, \underset{\sim}{I}_k)$$

where Z_{in} are defined in Theorem 2.1.

Now we obtain the limiting distribution of $(T_n, W_n, V_n, \Gamma_n)$.

<u>Theorem 2.2.</u> Let $\{Z_i\}_{i=1}^{\infty}$ be a sequence of normal independent (0,1) random variables. Let $\underset{\sim}{\eta}_n' = (T_n, W_n, V_n, \Gamma_n)$ where T_n, V_n, W_n, and Γ_n are defined in Theorem 2.1. Let $\underset{\sim}{\eta}' = (T, W, V, \Gamma)$, where

$$T = \sum_{i=1}^{\infty} a_i Z_i,$$

$$W = \sum_{i=1}^{\infty} b_i Z_i,$$

$$V = \sum_{i=1}^{\infty} g_i Z_i,$$

$$\Gamma = \sum_{i=1}^{\infty} \gamma_i^2 Z_i^2,$$

and a_i, b_i, g_i, γ_i are defined in Lemma 2.3. Then,

$$\underset{\sim}{\eta}_n \xrightarrow{L} \underset{\sim}{\eta}.$$

<u>Proof.</u> Using (2.18),

$$T_n = \sum_{i=1}^{n-1} a_i Z_{in} + o_p(1)$$

$$= \sum_{i=1}^{k} a_i Z_{in} + \sum_{i=k+1}^{n-1} a_i Z_{in} + o_p(1).$$

Note

$$V\left(\sum_{i=k+1}^{n-1} a_i Z_{in}\right) = \sum_{i=k+1}^{n-1} a_i^2$$

$$< \sum_{i=k+1}^{\infty} a_i^2,$$

converges to zero uniformly in n. From Lemma 2.3,

$$\sum_{i=1}^{k} a_i Z_{in} \xrightarrow{L} \sum_{i=1}^{k} a_i Z_i \quad \text{as} \quad n \to \infty.$$

Now,

$$\sum_{i=1}^{k} a_i Z_i \xrightarrow{L} T \qquad\qquad \text{as} \quad k \to \infty.$$

Therefore, using Lemma A.1 and Theorem A.7, we get

$$T_n \xrightarrow{L} T$$

Similarly,

$$\eta_n \xrightarrow{L} \eta \ .$$

□

<u>Corollary 2.1.</u> Let $\{Y_t\}$ satisfy (2.5) with $\rho = 1$. Let $\{e_t\}$ be a sequence of random variables satisfying the conditions (2.15), (2.16), and (2.17). Let $\hat{\rho}$, $\hat{\rho}_\mu$, $\hat{\rho}_\tau$ be defined by (2.8), (2.9), and (2.10), respectively and let $\hat{\tau}$, $\hat{\tau}_\mu$, $\hat{\tau}_\tau$ be defined by (2.11), (2.12), and (2.13), respectively. Then,

$$n(\hat{\rho}-1) \xrightarrow{L} \tfrac{1}{2}\,\Gamma^{-1}(T^2-1),$$

$$n(\hat{\rho}_\mu-1) \xrightarrow{L} \tfrac{1}{2}\,(\Gamma-W^2)^{-1}\;[(T^2-1) - 2TW],$$

$$\hat{\tau} \xrightarrow{L} \tfrac{1}{2}\,\Gamma^{-1/2}\,[T^2-1],$$

and

$$\hat{\tau}_\mu \xrightarrow{L} \tfrac{1}{2}\,(\Gamma-W^2)^{-1/2}\;[(T^2-1) - 2TW].$$

Let $\{Y_t\}$ satisfy (2.6) with $\rho = 1$. Then

$$n(\hat{\rho}_\tau-1) \xrightarrow{L} \tfrac{1}{2}\,(\Gamma-W^2-3V^2)^{-1}\;[(T-2W)(T-6V) - 1],$$

and

$$\hat{\tau}_\tau \xrightarrow{L} \tfrac{1}{2}\,(\Gamma-W^2-3V^2)^{-1/2}\;[(T-2W)(T-6V) - 1].$$

Proof. The proof is an immediate consequence of Theorem 2.2 because the denominator quadratic forms in $\hat{\rho}$, $\hat{\rho}_\mu$, $\hat{\rho}_\tau$ are continuous functions of $\underset{\sim}{\eta}$ that have probability 1 of being positive. Under the model (2.5),

$$\begin{aligned} S^2_{e1} &= (n-2)^{-1} \sum_{t=2}^{n} (Y_t - \hat{\rho}\, Y_{t-1})^2 \\ &= (n-2)^{-1} \sum_{t=2}^{n} [e_t - (\hat{\rho}-1)\, Y_{t-1}]^2 \\ &= (n-2)^{-1}\, [\sum_{t=2}^{n} e_t^2 - (\hat{\rho}-1) \sum_{t=2}^{n} Y_{t-1}\, e_t] \end{aligned}$$

$$= (n-2)^{-1} \sum_{t=1}^{n} e_t^2 + 0_p(n^{-1}).$$

$$S_{e2}^2 = (n-3)^{-1} \sum_{t=2}^{n} [Y_t - \hat{\mu} - \hat{\rho}_\mu Y_{t-1}]^2$$

$$= (n-3)^{-1} \sum_{t=2}^{n} [e_t - (\overline{y}_0 - \overline{y}_{-1}) - (\hat{\rho}_\mu - 1)(y_{t-1} - \overline{y}_{-1})]^2$$

$$= (n-3)^{-1} [\sum_{t=2}^{n} e_t^2 - (\hat{\rho}_\mu - 1) \sum_{t=2}^{n} (Y_{t-1} - \overline{y}_{-1}) e_t$$

$$+ (n-1)(\overline{y}_{-1} - \overline{y}_0)^2 + 2 \sum_{t=2}^{n} e_t (\overline{y}_{-1} - \overline{y}_0)]$$

$$= (n-3)^{-1} \sum_{t=2}^{n} e_t^2 + 0_p(n^{-1}),$$

where

$$\hat{\mu} = \overline{y}_0 - \hat{\rho}_\mu \overline{y}_{-1},$$

$$\overline{y}_0 = \frac{1}{n-1} \sum_{t=2}^{n} Y_t,$$

$$\overline{y}_{-1} = \frac{1}{n-1} \sum_{t=2}^{n} Y_{t-1}.$$

Similarly under (2.6),

$$S_{e3}^2 = (n-4)^{-1} \sum_{t=2}^{n} e_t^2 + O_p(n^{-1}) . \qquad \square$$

Note that the limiting distribution of $\hat{\rho}_\mu$ and $\hat{\tau}_\mu$ are obtained under the assumption that the constant term μ is zero. Likewise, the limiting distributions of $\hat{\rho}_\tau$ and $\hat{\tau}_\tau$ are derived under the assumption that the coefficient for time, β, is zero. If $\mu \neq 0$ in (2.6) or $\beta \neq 0$ in (2.7), then the limiting distributions of $\hat{\tau}_\mu$ and $\hat{\tau}_\tau$ are normal.

Extensions of Corollary 2.1 to p-th order case are presented in the Appendix B.

CHAPTER III. AN ADJUSTMENT FOR BIAS IN ESTIMATING THE PARAMETERS OF AN AUTOREGRESSIVE PROCESS DUE TO ESTIMATING CONSTANT MEAN

The methods of maximum likelihood and least squares estimation are commonly used to construct estimators for the parameters of a stationary normal first-order autoregressive process with mean zero. If the mean is unknown, there is no closed analytical form for the maximum likelihood estimator. The complexity of the likelihood equations increases with the order of the process, while the least squares estimation procedure easily extends to higher order processes. The large sample properties of the least squares estimators have received considerable attention. Several authors considered the small sample properties of the least squares estimator for the parameter of the first-order autoregressive process. However, the small sample properties of the least squares estimators have received very little attention in the case of higher order process.

Mariott and Pope (1954), Barnard et al. (1962), Copas (1966), Thornber (1967), Orcutt and Winokur (1969), Salem (1971), Min (1975), Sawa (1978), De Gooijer (1980), Ansley and Newbold (1980), Bora-Senta and Kounias (1980) and Lee (1981) proposed several estimators of α_1 for the first-order autoregressive process. These authors also compared the small sample properties of the various estimators through Monte Carlo studies. See Lee (1981) for details.

Salem (1971) extended the method of Mariott and Pope (1954) to obtain expressions for the approximate biases of the least squares estimators of a second order stationary autoregressive process. Bora-

Senta and Kounias (1980) considered an iterative method of moments procedure as an alternative to the least squares estimation procedure for higher order procedures. Lee (1981) extended Salem's (1971) methods to stationary second-order seasonal autoregressive processes.

Consider the stationary p-th order autoregressive process $\{Y_t\}$ which satisfies the stochastic difference equation

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \dots + \alpha_p Y_{t-p} + e_t \tag{3.1}$$

where $\{e_t\}$ is a sequence of independent $(0,\sigma^2)$ random variables. Assume that the roots of the characteristic equation, $m^p - \alpha_1 m^{p-1} - \dots - \alpha_p = 0$, are less than unity in modulus. Multiplying equation (3.1) by $(Y_{t-h}-\mu)$ for $h > 0$ and taking the expectation of both sides, where $\mu = \alpha_0(1 - \sum_{j=1}^{p} \alpha_j)^{-1}$, one obtains a system of equations relating the autocovariances to the coefficients of the model. The equations corresponding to $h = 1, 2, \dots, p$ are

$$\underset{\sim}{H}\,\underset{\sim}{\alpha} = \underset{\sim}{N} \tag{3.2}$$

where

$$\underset{\sim}{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_p)',$$

$$\underset{\sim}{H} = \begin{pmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(p-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(p-2) \\ \vdots & & & \\ \gamma(p-1) & \gamma(p-2) & \cdots & \gamma(0) \end{pmatrix},$$

and

$$\underset{\sim}{N} = [\gamma(1), \gamma(2), \ldots, \gamma(p)]'.$$

This system of p simultaneous equations is known as the Yule-Walker equations. See Yule (1926), Walker (1931). Levinson (1947) and Durbin (1960) give a recursive procedure for obtaining the Yule-Walker estimates of a p-th order autoregression.

The least squares estimator of $\underset{\sim}{\alpha}$ is given by

$$\underset{\sim}{\alpha}^* = \hat{\underset{\sim}{H}}^{-1} \hat{\underset{\sim}{N}}, \tag{3.3}$$

where

$$\hat{\underset{\sim}{H}} = (n-p)^{-1}$$

$$\begin{bmatrix} \Sigma(Y_{t-1}-\overline{Y}_1)(Y_{t-1}-\overline{Y}_1) & \Sigma(Y_{t-1}-\overline{Y}_1)(Y_{t-2}-\overline{Y}_2) & \cdots & \Sigma(Y_{t-1}-\overline{Y}_1)(Y_{t-p}-\overline{Y}_p) \\ \Sigma(Y_{t-2}-\overline{Y}_2)(Y_{t-1}-\overline{Y}_1) & \Sigma(Y_{t-2}-\overline{Y}_2)(Y_{t-2}-\overline{Y}_2) & \cdots & \Sigma(Y_{t-2}-\overline{Y}_2)(Y_{t-p}-\overline{Y}_p) \\ \vdots & & & \\ \Sigma(Y_{t-p}-\overline{Y}_p)(Y_{t-1}-\overline{Y}_1) & \Sigma(Y_{t-p}-\overline{Y}_p)(Y_{t-2}-\overline{Y}_2) & \cdots & \Sigma(Y_{t-p}-\overline{Y}_p)(Y_{t-p}-\overline{Y}_p) \end{bmatrix},$$

$$\hat{\underset{\sim}{N}} = (n-p)^{-1} \begin{bmatrix} \Sigma(Y_{t-1}-\overline{Y}_1)(Y_t-\overline{Y}_0) \\ \\ \Sigma(Y_{t-2}-\overline{Y}_2)(Y_t-\overline{Y}_0) \\ \vdots \\ \Sigma(Y_{t-p}-\overline{Y}_p)(Y_t-\overline{Y}_0) \end{bmatrix},$$

$$\overline{Y}_i = (n-p)^{-1} \; \Sigma \; Y_{t-i},$$

and all the summations are over $t = p+1, p+2, \ldots, n$.

Burg (1967, 1968) suggested a method of estimating the autoregressive parameters based on the Levinson (1947) - Durbin (1960) procedure used in computing the Yule-Walker estimates. Autoregressions of increasing order are fit in a stepwise fashion. Denote the estimate of the j-th coefficient obtained by fitting an autoregression of order p by $a_{j(p)}$, and the estimate of σ^2 by S_p. The recursion begins with

$$S_0 = n^{-1} \sum_{t=1}^{n} Y_t^2,$$

$$a_{1(1)} = \frac{2 \sum_{t=2}^{n} Y_t Y_{t-1}}{\sum_{t=2}^{n} Y_{t-1}^2 + \sum_{t=2}^{n} Y_t^2} ,$$

$$S_1 = n^{-1} \sum_{t=1}^{n} Y_t^2 (1 - a_{1(1)}^2)$$

when the mean is known and is equal to zero. At the p-th stage, define the residuals from a p-th order autoregression by

$$e_{t(p)} = Y_t - \sum_{j=1}^{p} a_{j(p)} Y_{t-j}, \quad t = p+1, \ldots, n$$

$$= Y_t - \sum_{j=1}^{p-1} (a_{j(p-1)} - a_{p(p)} a_{p-j,(p-1)}) Y_{t-j} - a_{p(p)} Y_{t-p}.$$

Similarly the backward residuals are

$$h_{t(p)} = Y_t - \sum_{j=1}^{p-1} [a_{j(p-1)} - a_{p(p)} a_{p-j,(p-1)}] Y_{t+j}$$

$$- a_{p(p)} Y_{t+p}, \quad t = 1, 2, \ldots, n-p.$$

The coefficient at the p-th stage is chosen to minimize the sum of squares

$$\sum_{t=p+1}^{n} [e_{t(p)}]^2 + \sum_{t=1}^{n-p} [h_{t(p)}]^2,$$

giving

$$a_{p(p)} = \frac{2 \sum_{t=p+1}^{n} e_{t(p-1)} h_{t-p,(p-1)}}{\sum_{t=1}^{n-p} [h_{t(p-1)}]^2 + \sum_{t=p+1}^{n} [e_{t(p-1)}]^2} .$$

The other coefficients are updated by

$$a_{j(p)} = a_{j(p-1)} - a_{p(p)} a_{p-j(p-1)}, \quad j = 1, 2, \ldots, p-1,$$

and the updated estimate of σ^2 is

$$S_p = S_{p-1}(1 - a^2_{p(p)}).$$

When the mean is unknown, $Y_t - \overline{Y}$ is substituted for Y_t and the recursion proceeds as before. See Burg (1975), Ulrych and Bishop (1975), Jones (1978), and Robinson and Silvia (1980).

Box and Jenkins (1976) proposed a method that gives the approximate maximum likelihood estimators in the case of normally distributed errors. The estimators $\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_p$ minimize the sum of squares

$$S(\underset{\sim}{\alpha}) = \sum_{t=-j}^{n} [e_t]^2,$$

where $[e_t] = Y_t - \alpha_0 - \alpha_1 Y_{t-1} - \dots - \alpha_p Y_{t-p}$, $t = p+1, p+2, \dots, n$, and $[e_p], [e_{p-1}], \dots, [e_{-j}]$ are formed from

$$[e_t] = Y_t - \alpha_0 - \alpha_1 Y_{t-1} - \dots - \alpha_p Y_{t-p}, \qquad t = p, p-1, \dots, -j,$$

$$Y_t = \alpha_0 + \alpha_1 Y_{t+1} + \dots + \alpha_p Y_{t+p}, \qquad t < 0.$$

Recursive algorithms such as Marquardt's (1963) algorithm are used to perform the iterations.

The various estimators considered are asymptotically equivalent, but behave differently in small samples. It is well-known that the estimation methods are biased in finite samples although the exact distributions of the estimators are not known. In the case of first-order autoregressive process, a number of methods have been proposed to reduce the bias in the estimate of α_1. See Lee (1981).

Quenouille (1949) suggested a method of removing the bias in the least squares estimators of autoregressive parameters. Assuming the bias is proportional to n^{-1}, the method consists of dividing the series into halves and estimating the autoregressive parameters using the whole series and each half separately. An estimator of $\underset{\sim}{\alpha}$ unbiased to order n^{-1} is obtained as

$$\hat{\hat{\underset{\sim}{\alpha}}} = 2\, \underset{\sim}{\alpha}^* - \tfrac{1}{2} (\hat{\underset{\sim}{\alpha}}' + \hat{\underset{\sim}{\alpha}}'')$$

where $\hat{\underset{\sim}{\alpha}}'$ and $\hat{\underset{\sim}{\alpha}}''$ are the least squares estimator of $\underset{\sim}{\alpha}$ for the first and second halves, respectively.

Salem (1971) obtained the moments of the least squares estimator $\underset{\sim}{\alpha}^*$ up to terms of order n^{-1} for a second-order stationary autoregressive process. The means of α_1^* and α_2^* are

$$E[\alpha_1^*] = \alpha_1 - (n-2)^{-1}(1+\alpha_2) + O(n^{-1}),$$

and

$$E[\alpha_2^*] = \alpha_2 - (n-2)^{-1}(1+\alpha_2) + O(n^{-1}).$$

A linear transformation of α_1^* and α_2^* that is nearly unbiased is constructed based upon the above expressions.

$$\begin{aligned}\hat{\alpha}_2(S) &= [\alpha_2^*(n-2)+1](n-3)^{-1}, && \alpha_2^* \in (-1, 1-2(n-2)^{-1})\\ &= 1, && \alpha_2^* \geq 1-2(n-2)^{-1}\\ &= -1, && \alpha_2^* \leq -1.\end{aligned}$$

and

$$\hat{\alpha}_1(S) = \alpha_1^* + (n-2)^{-1}(1+\hat{\alpha}_2(S)) \ . \qquad (3.4)$$

Bora-Senta and Kounias (1980) recently proposed a method for parameter estimation of an autoregressive model with unknown constant mean. The authors propose an iterative procedure using modified estimators of the autocorrelations

$$\hat{\rho}_h = \frac{\hat{A}}{n\hat{\gamma}_0} + \frac{n}{n-h}\, r_h\left(1 - \frac{\hat{A}}{n\hat{\gamma}_0}\right), \tag{3.5}$$

where

$$C_h = n^{-1} \sum_{t=1}^{n-h} (Y_t - \bar{Y})(Y_{t+h} - \bar{Y}),$$

$$r_h = C_h / C_0,$$

and

$$\frac{\hat{A}}{\hat{\gamma}_0} = \frac{1 - \hat{\alpha}_1\hat{\rho}_1 - \hat{\alpha}_2\hat{\rho}_2 - \dots - \hat{\alpha}_p\hat{\rho}_p}{(1 - \hat{\alpha}_1 - \hat{\alpha}_2 - \dots - \hat{\alpha}_p)^2}. \tag{3.6}$$

The iteration proceeds as follows:

i) As a first approximation $\hat{\rho}_{h,1} = r_h$, $h = 1, 2, \dots, p$.

ii) Using $\hat{\rho}_{h,1}$, $h = 1$ to p, compute the Yule-Walker type estimates $\hat{\alpha}_{1,1}, \hat{\alpha}_{2,1}, \dots, \hat{\alpha}_{p,1}$.

iii) Calculate $\hat{A}/\hat{\gamma}_0$ from (3.6) using the estimates $\hat{\alpha}_{h,1}$, $\hat{\rho}_{h,1}$, $h = 1$ to p.

iv) Obtain second approximations $\hat{\rho}_{h,2}$, $h = 1$ to p, using $\hat{A}/\hat{\gamma}_0$ in (3.5).

v) Check the conditions for stationarity. If violated, take the previous estimates.

vi) If not violated, continue until the sum of squares

$$J_i = \sum_{h=1}^{p} (\hat{\alpha}_{h,i+1} - \hat{\alpha}_{h,i})^2$$

is less than a quantity ε.

Lee (1981) also considered a modified least squares estimator which corrects for the bias in autocovariances. Let

$$\hat{\gamma}(0) = (n-p)^{-1} \sum_{t=p+1}^{n} (Y_t - \overline{Y})^2,$$

and

$$V = \mathrm{Var}(\overline{Y}).$$

Lee suggested the estimator

$$\hat{\hat{\underset{\sim}{\alpha}}} = \hat{\hat{\underset{\sim}{H}}}^{-1} \hat{\hat{\underset{\sim}{N}}},$$

where,

$$\hat{\hat{\underset{\sim}{H}}} = \hat{\underset{\sim}{H}} + \hat{V} \underset{\sim}{J} \underset{\sim}{J}',$$

$$\hat{\hat{\underset{\sim}{N}}} = \hat{\underset{\sim}{N}} + \hat{V} \underset{\sim}{J},$$

$$\underset{\sim}{J} = (1, 1, \ldots, 1)',$$

and $\hat{V}$ is the estimator of V obtained by substituting $\hat{\gamma}(0)$ and $\underset{\sim}{\alpha}^*$ for $\gamma(0)$ and $\underset{\sim}{\alpha}$, respectively.

Fuller and Hasza (1981) established that the least squares estimators for normal autoregressive parameters are integrable. Using Taylor's Theorem, Lee (1981) obtained, for the least squares estimator

$$E(\underset{\sim}{\alpha}^{*}) = \underset{\sim}{\alpha} + E\{-\underset{\sim}{H}^{-1} (\underset{\sim}{A}\underset{\sim}{\alpha}-\underset{\sim}{d}) + \underset{\sim}{H}^{-1} \underset{\sim}{A}\underset{\sim}{H}^{-1} (\underset{\sim}{A}\underset{\sim}{\alpha}-\underset{\sim}{d})\} + O(n^{-2}) \tag{3.7}$$

where $\underset{\sim}{A} = \hat{\underset{\sim}{H}} - \underset{\sim}{H}$ and $\underset{\sim}{d} = \hat{\underset{\sim}{N}} - \underset{\sim}{N}$.

The bias in $\underset{\sim}{\alpha}^{*}$ arises from two sources. The first source of bias is inherent in estimating the product of the inverse of $\underset{\sim}{H}$ and the vector $\underset{\sim}{N}$. The second source of bias results from estimating the mean when the true mean is unknown. The approximate bias in $\underset{\sim}{\alpha}^{*}$ arising from estimating the mean is given by $E\{-\underset{\sim}{H}^{-1}(\underset{\sim}{A}\underset{\sim}{\alpha}-\underset{\sim}{d})\}$ and is evaluated in the following theorem.

<u>Theorem 3.1.</u> Let $\{Y_t\}$ be a stationary time series satisfying the stochastic difference equation

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \dots + \alpha_p Y_{t-p} + e_t, \tag{3.8}$$

where the $\{e_t\}$ are independent normal $(0,\sigma^2)$ random variables and the roots of the characteristic equation, $m^p - \alpha_1 m^{p-1} - \dots - \alpha_p = 0$, are less than unity in modulus. Let the least squares estimator $\underset{\sim}{\alpha}^{*}$ be defined by (3.3). Then,

$$E(\underset{\sim}{A}\underset{\sim}{\alpha}-\underset{\sim}{d}) = \frac{\sigma^2}{(n-p)(1-\sum_{i=1}^{p}\alpha_i)}(1, 1, \ldots, 1)' + O(n^{-2}), \tag{3.9}$$

where $\underset{\sim}{A} = \hat{\underset{\sim}{H}} - \underset{\sim}{H}$ and $\underset{\sim}{d} = \hat{\underset{\sim}{N}} - \underset{\sim}{N}$.

<u>Proof.</u> Let $\gamma(h)$ be the autocovariance function of $\{Y_t\}$. Using Theorem A.9,

$$\mathrm{Var}(\overline{Y}_i) = (n-p)^{-1}\sum_{h=-\infty}^{\infty}\gamma(h) + O(n^{-2}). \tag{3.10}$$

For a stationary p-th order autoregressive process,

$$\sum_{h=-\infty}^{\infty}\gamma(h) = \sigma^2(1-\sum_{i=1}^{p}\alpha_i)^{-2}. \tag{3.11}$$

Using Theorem A.10,

$$\begin{aligned} E[\hat{h}_{ij}] &= E[(n-p)^{-1}\sum_{t=p+1}^{n}(Y_{t-i}-\overline{Y}_i)(Y_{t-j}-\overline{Y}_j)] \\ &= \gamma(j-i) - V(\overline{Y}_0) + O(n^{-2}) \\ &= h_{ij} - \frac{\sigma^2}{(n-p)(1-\sum_{i=1}^{p}\alpha_i)^2} + O(n^{-2}). \end{aligned} \tag{3.12}$$

and

$$E[\hat{N}_i] = E[(n-p)^{-1}\sum_{t=p+1}^{n}(Y_{t-i}-\overline{Y}_i)(Y_t-\overline{Y}_0)]$$

$$= \gamma(i) - V(\bar{Y}_0) + O(n^{-2})$$

$$= N_i - \frac{\sigma^2}{(n-p)(1 - \sum_{i=1}^{p} \alpha_i)^2} + O(n^{-2}) \qquad (3.13)$$

where $\hat{h}_{ij}$ and h_{ij} are (i,j)-th elements of $\hat{\underset{\sim}{H}}$ and $\underset{\sim}{H}$, and $\hat{N}_i$ and N_i are the i-th elements of $\hat{\underset{\sim}{N}}$ and $\underset{\sim}{N}$ respectively. Therefore,

$$E(\underset{\sim}{A}\underset{\sim}{\alpha}-\underset{\sim}{d}) = E[\hat{\underset{\sim}{H}} - \underset{\sim}{H}]\underset{\sim}{\alpha} - E[\hat{\underset{\sim}{N}} - \underset{\sim}{N}]$$

$$= \frac{-\sigma^2}{(n-p)(1 - \sum_{i=1}^{p} \alpha_i)^2} \underset{\sim}{J}\,\underset{\sim}{J}'\underset{\sim}{\alpha}$$

$$+ \frac{\sigma^2}{(n-p)(1 - \sum_{i=1}^{p} \alpha_i)^2} \underset{\sim}{J} + O(n^{-2})$$

$$= \frac{\sigma^2}{(n-p)(1 - \sum_{i=1}^{p} \alpha_i)} \underset{\sim}{J} + O(n^{-2}),$$

where $\underset{\sim}{J} = (1,1,\ldots1)'$, $\underset{\sim}{J}'\,\underset{\sim}{\alpha} = \Sigma\, \alpha_i$, and all the summations are over $i = 1, 2, \ldots, p$. □

Using (3.9) the least squares estimators can be modified to correct for the bias arising from estimating the unknown mean. The following modification is suggested.

(i) Regress Y_t on $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$ with an intercept to obtain the least squares estimator of $(\alpha_0, \alpha_1, \ldots, \alpha_p)$, The least squares estimator of $\underset{\sim}{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_p)'$, is

$$\underset{\sim}{\alpha}^* = \hat{\underset{\sim}{H}}^{-1} \hat{\underset{\sim}{N}},$$

where $\hat{\underset{\sim}{H}}$ and $\hat{\underset{\sim}{N}}$ are given by (3.3).

(ii) Obtain an estimator $\hat{\sigma}^2$ of σ^2. The residual mean square error of the above regression is a consistent estimator of σ^2.

(iii) Construct the new estimator

$$\tilde{\underset{\sim}{\alpha}} = \hat{\underset{\sim}{H}}^{-1} \tilde{\underset{\sim}{N}} \tag{3.14}$$

where

$$\tilde{\underset{\sim}{N}} = \hat{\underset{\sim}{N}} + \frac{\hat{\sigma}^2}{(n-p)(1 - \sum_{i=1}^{p} \alpha_i^*)} (1, 1, \ldots, 1)'.$$

If $p = 1$, the estimator reduces to

$$\tilde{\alpha}_1 = \alpha_1^* + [(n-1)\ \hat{\gamma}(0)\ (1-\alpha_1^*)]^{-1}\ \hat{\sigma}^2$$

$$\doteq \alpha_1^* + (n-1)^{-1}\ (1 + \alpha_1^*).$$

The estimator (3.14) is relatively easy to construct and, hence, is of practical importance. We shall study the estimator and its extensions to the case of alternative mean functions. The Monte Carlo study of Chapter V demonstrates that the mean square error of estimator (3.14) is smaller than that of the least squares estimator for a wide range of parameter values.

Theorem 3.1 also suggests that it is possible to isolate the effect of estimating the mean by transforming model (3.1). For $p \geq 2$, consider the following reparametrization of model (3.1). Let

$$Y_t = \alpha_0 + \delta_1 Y_{t-1} + \sum_{i=2}^{p} \delta_i (Y_{t-i+1} - Y_{t-i}) + e_t \tag{3.15}$$

where

$$\underset{\sim}{\delta} = (\delta_1, \delta_2, \ldots, \delta_p)'$$

$$= \underset{\sim}{C}(\alpha_1, \alpha_2, \ldots, \alpha_p)'$$

for some nonsingular matrix $\underset{\sim}{C}$. For $p = 1$,

$$Y_t = \alpha_0 + \delta_1 Y_{t-1} + e_t$$

and $\delta_1 = \alpha_1$. Note, by (3.15) for $p > 2$,

$$Y_t = \alpha_0 + (\delta_1 + \delta_2)\, Y_{t-1} + (\delta_3 - \delta_2)\, Y_{t-2} + \cdots + (\delta_p - \delta_{p-1})\, Y_{t-p+1} - \delta_p Y_{t-p} + e_t. \tag{3.16}$$

Comparing (3.16) with (3.1), we get

$$\alpha_1 = (\delta_1 + \delta_2)$$
$$\alpha_j = \delta_{j+1} - \delta_j, \qquad j = 2, 3, \ldots, p-1,$$

and

$$\alpha_p = -\delta_p.$$

Therefore,

$$\delta_1 = \sum_{j=1}^{p} \alpha_j. \tag{3.17}$$

Define,

$$P_t = Y_t - \mu,$$

$$\mu = \alpha_0 (1 - \delta_1)^{-1},$$

$$\overline{P} = n^{-1} \sum_{t=1}^{n} P_t,$$

$$Z_t = Y_t - Y_{t-1},$$

$$\overline{Y} = n^{-1} \sum_{t=1}^{n} Y_t,$$

and

$$\begin{aligned} W_t &= Y_t - \overline{Y} \\ &= P_t - \overline{P}. \end{aligned} \tag{3.18}$$

Note that $\{P_t\}$ is a stationary p-th order autoregressive process with mean zero. Also, W_t satisfies

$$\begin{aligned} W_t &= \delta_1 W_{t-1} + \sum_{i=2}^{p} \delta_i Z_{t-i+1} + v_t, && \text{if } p \geq 2 \\ &= \delta_1 W_{t-1} + v_t \qquad\qquad , && \text{if } p = 1 \end{aligned} \tag{3.19}$$

where

$$\begin{aligned} v_t &= e_t - \overline{e} + \frac{\delta_1}{n} P_n - \sum_{i=2}^{p} \frac{\delta_i}{n} P_{n-i+1}, && \text{if } p \geq 2 \\ &= e_t - \overline{e} + \frac{\delta_1}{n} P_n \qquad\qquad , && \text{if } p = 1 \end{aligned} \tag{3.20}$$

and

$$\overline{e} = n^{-1} \sum_{t=1}^{n} e_t.$$

The least squares estimator $\underset{\sim}{\delta}^*$ of $\underset{\sim}{\delta} = (\delta_1, \delta_2, \ldots, \delta_p)'$ is obtained by regressing W_t on $W_{t-1}, Z_{t-1}, \ldots, Z_{t-p+1}$ if $p \geq 2$ and W_t on W_{t-1} if $p = 1$. The error in the fitted equation is

$$\underset{\sim}{\delta}^* - \underset{\sim}{\delta} = \hat{\underset{\sim}{D}}^{-1} \hat{\underset{\sim}{M}} , \tag{3.21}$$

where

$$\hat{\underset{\sim}{D}} = (n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{B}_t' \underset{\sim}{B}_t$$

$$\begin{aligned} \underset{\sim}{B}_t &= (W_{t-1}, Z_{t-1}, \ldots, Z_{t-p+1}) && \text{if } p \geq 2 \\ &= W_{t-1} && \text{if } p = 1 \end{aligned}$$

and

$$\hat{\underset{\sim}{M}} = (n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{B}_t' v_t .$$

The approximate bias in $\underset{\sim}{\delta}^*$ arising from estimating the mean is established in the following theorem.

<u>Theorem 3.2.</u> Let $\{Y_t\}$ be a stationary time series satisfying the stochastic difference equation

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \ldots + \alpha_p Y_{t-p} + e_t ,$$

where the $\{e_t\}$ are independent normal $(0, \sigma^2)$ random variables and the roots of the characteristic equation, $m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p = 0$, are less than unity in modulus. Let $\hat{\underset{\sim}{M}}$ be given by (3.21). Then,

$$\begin{aligned} E(\hat{\underset{\sim}{M}}) &= \left(\frac{-\sigma^2}{(n-p)(1-\delta_1)}, 0, 0, \ldots, 0\right)' + O(n^{-2}), \text{ if } p \geq 2 \\ &= \frac{-\sigma^2}{(n-p)(1-\delta_1)} + O(n^{-2}) \qquad , \text{ if } p = 1 \end{aligned} \tag{3.22}$$

<u>Proof.</u> Let $\bar{P}_1 = n^{-1} \sum_{t=p+1}^{n} P_{t-1}$ and $\bar{Y}_1 = n^{-1} \sum_{t=p+1}^{n} Y_{t-1}$. For $p = 1$, by Theorem 3.1,

$$E(\hat{\underset{\sim}{M}}) = \frac{-\sigma^2}{(n-p)(1-\delta_1)} + O(n^{-2}) .$$

For $p \geq 2$,

$$\begin{aligned} E\left[\sum_{t=p+1}^{n} W_{t-1} v_t\right] &= E\left[\sum_{t=p+1}^{n} (Y_{t-1}-\bar{Y}) v_t\right] \\ &= E\left[\sum_{t=p+1}^{n} (Y_{t-1}-\bar{Y}_1) v_t\right] + E\left[\sum_{t=p+1}^{n} (\bar{Y}_1-\bar{Y}) v_t\right] \\ &= E\left[\sum_{t=p+1}^{n} (P_{t-1}-\bar{P}_1)\{e_t-\bar{e} + \frac{\delta_1}{n} P_n - \sum_{i=2}^{p} \frac{\delta_i}{n} P_{n-i+1}\}\right] \\ &\quad + E\left[(\bar{P}_1-\bar{P}) \sum_{t=p+1}^{n} v_t\right] \end{aligned}$$

$$= E\left[\sum_{t=p+1}^{n} (P_{t-1}-\overline{P}_1)(e_t-\overline{e})\right]$$

$$- E\left[(n^{-1}\sum_{t=1}^{p-1} P_t + n^{-1} P_n)\right.$$

$$\left.\left\{-\sum_{t=1}^{p} e_t + \delta_1 P_n - \sum_{i=2}^{p} \delta_i P_{n-i+1}\right\}\right]$$

$$+ O(n^{-1})$$

$$= - E\left[\sum_{t=p+1}^{n} \overline{P}_1\, e_t\right] + O(n^{-1})$$

$$= - n^{-1} E\left[\sum_{t=p+1}^{n} \sum_{j=p+1}^{n} P_j e_t\right] + O(n^{-1})$$

$$= - n^{-1} E\left[\sum_{t=p+1}^{n} \sum_{j=t}^{n} P_j e_t\right] + O(n^{-1})$$

$$= - \sigma^2 n^{-1} \sum_{t=p+1}^{n} \sum_{j=t}^{n} w_{j-t} + O(n^{-1})$$

$$= - \sigma^2 n^{-1} \sum_{t=p+1}^{n} \sum_{j=0}^{n-t} w_j + O(n^{-1})$$

$$= - \sigma^2 n^{-1} \sum_{j=0}^{n-p-1} w_j (n-j) + O(n^{-1})$$

where $P_t = \sum_{j=0}^{\infty} w_j e_{t-j}$ and w_j satisfy the p-th order difference equations

$$w_j - \alpha_1 w_{j-1} - \cdots - \alpha_p w_{j-p} = 0, \quad j = 1,2,\ldots,$$

with $w_j = 0$ for $j < 0$, and $w_0 = 1$. Since the roots of the characteristic equation are less than one,

$$\sum_{j=0}^{\infty} j \left| w_j \right| < \infty$$

and

$$\sum_{j=n}^{\infty} w_j = O(n^{-1}).$$

Therefore,

$$E[\sum_{t=p+1}^{n} W_{t-1} v_t] = -(\sum_{j=0}^{\infty} w_j)\, \sigma^2 + O(n^{-1})$$

$$= -(1 - \delta_1)^{-1} \sigma^2 + O(n^{-1}).$$

Now for $j = 1, 2, \ldots, p-1$,

$$E[\sum_{t=p+1}^{n} Z_{t-j} v_t] = E[\sum_{t=p+1}^{n} Z_{t-j}\{e_t - \bar{e} + \frac{\delta_1}{n} P_n - \sum_{i=2}^{n} \frac{\delta_i}{n} P_{n-i+1}\}]$$

$$= -E[\sum_{t=p+1}^{n} Z_{t-j}\, \bar{e}] + O(n^{-1})$$

$$= -E[(P_{n-j} - P_{p-j+1})\, \bar{e}] + O(n^{-1})$$

$$= O(n^{-1}),$$

since $\sum_{j=0}^{\infty} | w_j |$ is finite. Therefore, for $p > 2$,

$$E(\hat{\underset{\sim}{M}}) = \left(\frac{-\sigma^2}{(n-p)(1-\delta_1)}, 0, 0, \ldots, 0 \right)' + O(n^{-1}). \qquad \square$$

The approach used in proving Theorem 3.2 is different from that of Theorem 3.1, but the results of Theorem 3.1 and 3.2 are equivalent. We will use the approach of Theorem 3.2 to extend the result for the p-th order autoregressive process with $E(Y_t)$ a polynomial in t.

Theorem 3.2 also makes it possible to establish whether or not the roots of the process associated with the modified estimator are less than one in absolute value. If one of the roots of the characteristic equation is one, then $\delta_1 = 1$. This fact is established in the following lemma.

Lemma 3.1. Consider the polynomial,

$$f(m) = m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p .$$

Then $\sum_{i=1}^{p} \alpha_i = 1$ if and only if $f(1) = 0$.

Proof. The result is immediate because $f(1) = 1 - \sum_{i=1}^{p} \alpha_i$. $\square$

Lemma 3.1 and Theorem 3.2 suggest the following method of correcting for the bias due to estimating the mean

(a) Obtain the least squares estimator $\underset{\sim}{\delta}^*$, for $\underset{\sim}{\delta} = (\delta_1, \delta_2, \ldots, \delta_p)'$, by regressing $W_t = Y_t - \overline{Y}$ on $W_{t-1}, Z_{t-1}, \ldots, Z_{t-p+1}$. Obtain the modified least squares estimator,

$$\begin{aligned} \hat{\underset{\sim}{\delta}} &= \hat{\underset{\sim}{D}}^{-1} \hat{\underset{\sim}{\phi}} &&, \quad \text{if} \quad \delta_1^* < 1 \\ &= \hat{\underset{\sim}{D}}^{-1} [\hat{\underset{\sim}{\phi}} + (D'')^{-1} (1 - \delta_1^*, 0, 0, \ldots, 0)'], && \quad \text{if} \quad \delta_1^* > 1 \end{aligned} \tag{3.23}$$

where

$$\underset{\sim}{\delta}^* = \hat{\underset{\sim}{D}}^{-1} \hat{\underset{\sim}{\phi}},$$

$$\hat{\underset{\sim}{D}} = (n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{B}_t' \underset{\sim}{B}_t,$$

$$\begin{aligned} \underset{\sim}{B}_t &= (W_{t-1}, Z_{t-1}, \ldots, Z_{t-p+1}), && \text{if} \quad p > 2 \\ &= W_{t-1}, && \text{if} \quad p = 1, \end{aligned}$$

$$\hat{\underset{\sim}{\phi}} = (n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{B}_t' W_t,$$

and D'' is the upper left element of $\hat{\underset{\sim}{D}}^{-1}$.

(b) Use the mean square error of the regression in (a) as an estimator $\hat{\sigma}^2$ of σ^2.

(c) Obtain the modified estimates as

$$\tilde{\underset{\sim}{\delta}} = \hat{\underset{\sim}{D}}^{-1} [\hat{\underset{\sim}{\phi}} - \hat{\underset{\sim}{E}}(\hat{\underset{\sim}{M}})] \tag{3.24}$$

where

$$\begin{aligned} \hat{\underset{\sim}{E}}(\hat{M}) &= (\hat{g}_1, 0, 0, \ldots, 0)', \quad \text{if} \quad p > 2 \\ &= \hat{g}_1 \qquad\qquad\qquad\quad , \quad \text{if} \quad p = 1, \end{aligned}$$

and

$$\begin{aligned} \hat{g}_1 &= -(D'')^{-1}(1 - \delta_1^*) \text{ if } \delta_1^* > 1 \text{ or} \\ &\qquad\qquad \text{if } [(n-p)(1-\hat{\delta}_1)]^{-1}\hat{\sigma}^2 D'' > (1 - \hat{\delta}_1), \\ &= -[(n-p)(1 - \hat{\delta}_1)]^{-1} \hat{\sigma}^2 \qquad \text{otherwise.} \end{aligned}$$

(d) The estimate of $\underset{\sim}{\alpha}$ is defined by $\tilde{\underset{\sim}{\alpha}} = \underset{\sim}{C}^{-1} \tilde{\underset{\sim}{\delta}}$ where $\underset{\sim}{C}$ is defined in (3.15).

The estimators defined by (3.23) and (3.24) have the property that $\hat{\delta}_1 < 1$ and $\tilde{\delta}_1 < 1$. Similar modifications can be introduced to guarantee that the smallest root is greater than or equal to negative one by requiring

$$\sum_{i=1}^{p} (-1)^i \alpha_i < 1.$$

The procedure can be extended to check for all roots, but the method outlined in (3.23) and (3.24) should be sufficient for most practical situations.

Let $m_1, m_2, \ldots, m_p$ be the roots of $f(m) = 0$. Assume $m_1 = 1$ and $|m_i| < 1$ for $i = 2, 3, \ldots, p$. Also assume that $\alpha_0 = 0$. Dickey and Fuller (1979) derived the limiting distribution of the t-statistic,

$$\hat{\tau}_\mu = \frac{\delta_1^* - 1}{[\hat{V}(\delta_1^*)]^{1/2}} = \frac{\delta_1^* - 1}{[(n-p)^{-1}\, D''\hat{\sigma}^2]^{1/2}},$$

where D'' is the (1,1)-th element of $\hat{\underset{\sim}{D}}^{-1}$. and $\hat{\underset{\sim}{D}}$ is defined in (3.23). The percentiles of the distribution of the t-statistic are given in Table 8.5.2 of Fuller (1976) for different sample sizes. If $\alpha_0 \neq 0$ then the limiting distribution of $\hat{\tau}_\mu$ is standard normal.

To extend the bias adjustment method to the case in which $\delta_1 \varepsilon [-1, 1]$, several possibilities beyond that of (3.24) exist. One is to use the $\hat{\tau}_\mu$ tables as follows.

(i) Construct

$$\hat{\tau}_\mu = \frac{\delta_1^* - 1}{[\hat{V}(\delta_1^*)]^{1/2}}$$

the regression t-statistic for δ_1. If $\hat{\tau}_\mu$ is greater than the median $\hat{\tau}_{\mu(50)}$, for $\hat{\tau}_\mu$ (about -1.5), set $\tilde{\delta}_1(\tau_\mu) = 1$ by adding $(D")^{-1}(1-\hat{\delta}_1)$ to $(n-p)^{-1} \Sigma_{t=p+1}^{n} W_{t-1} W_t$.

(ii) If $\hat{\tau}_\mu$ < the 0.01 tabular value, $\hat{\tau}_{\mu(01)}$, for $\hat{\tau}_\mu$ (about -3.5), then use the stationary adjustment given in (3.24).

(iii) Make the adjustment continuous in $\hat{\tau}_\mu$ for $\hat{\tau}_{\mu(01)} < \hat{\tau}_\mu < \hat{\tau}_{\mu(50)}$.

One method of obtaining a continuous adjustment is to make the adjustment cubic in $\hat{\tau}_\mu$. The cubic adjustment is selected so that the adjustment does not have a large effect on moderately sized values of δ_1. Let

$$\tilde{\underset{\sim}{\delta}}(\tau_\mu) = \hat{\underset{\sim}{D}}^{-1} (\hat{\underset{\sim}{\phi}} + \hat{\underset{\sim}{f}}) \tag{3.25}$$

where

$$\hat{\underset{\sim}{f}} = (\hat{f}_1, 0, 0, \ldots, 0)', \qquad \text{if } p > 2$$

$$= \hat{f}_1 \qquad , \qquad \text{if } p = 1,$$

$$\hat{f}_1 = [(n-p)(1-\delta_1^*)]^{-1} \hat{\sigma}^2, \qquad \text{if } \hat{\tau}_\mu < \hat{\tau}_{\mu(01)}$$

$$= a + b\,[\hat{\tau}_{\mu(50)} - \hat{\tau}_{\mu(01)}]^{-3}[\hat{\tau}_\mu - \hat{\tau}_{\mu(01)}]^3,$$

$$\text{if } \hat{\tau}_{\mu(01)} < \hat{\tau}_\mu < \hat{\tau}_{\mu(50)}$$

$$= (D'')^{-1}(1 - \delta_1^*) \qquad \text{if } \hat{\tau}_\mu > \hat{\tau}_{\mu(50)},$$

$$a = [(n-p)(1 - \delta_1^*)]^{-1} \hat{\sigma}^2,$$

and

$$b = (D'')^{-1} (1 - \delta_1^*) - a.$$

The above method of adjusting for bias arising from estimating the mean extends immediately to a seasonal p-th order autoregressive process with unknown seasonal means.

Consider the stationary p-th order seasonal process $\{Y_t\}$ which satisfies the stochastic difference equation

$$Y_t = \sum_{i=0}^{k-1} \delta_{it} g_i + \sum_{j=1}^{p} \alpha_j Y_{t-jk} + e_t, \qquad t = 1,2,\ldots,nk, \tag{3.26}$$

where

$$\delta_{it} = 1, \qquad i = (t-1) \bmod k$$
$$= 0, \qquad \text{otherwise},$$

$\{e_t\}$ is a sequence of independent $(0,\sigma^2)$ random variables and g_i are parameters. Observe that this is a pk-th order autoregressive process that is purely seasonal in the sense that it can be written as k independent p-th order processes. This model can also be represented as

$$Y_{ij} = g_i + \sum_{\ell=1}^{p} \alpha_\ell Y_{i,j-\ell} + e_{ij},$$

$$= \beta_i + \sum_{\ell=1}^{p} \alpha_\ell(Y_{i,j-\ell} - \beta_i) + e_{ij}, \qquad i = 0,1,\ldots,k-1;$$
$$j = 1,2,\ldots,n, \tag{3.27}$$

where $g_i = \beta_i(1 - \Sigma_{\ell=1}^{p} \alpha_\ell)$, β_i is the i-th seasonal mean and Y_{ij} is the value for the i-th period and j-th year. It is assumed that the roots of the polynomial equation,

$$m^{kp} - \alpha_1 m^{k(p-1)} - \ldots - \alpha_p = 0, \tag{3.28}$$

lie inside the unit circle.

Consider the following reparametrization of the model (3.26). For $p > 2$, let

$$Y_{ij} = \beta_i + \delta_1(Y_{i,j-1} - \beta_i) + \sum_{\ell=1}^{p-1} \delta_{\ell+1} Z_{i,j-\ell} + e_{ij} \tag{3.29}$$

where $Z_{ij} = Y_{ij} - Y_{i,j-1}$ and $\delta_1, \delta_2, \ldots, \delta_p$ are linear combinations of $\alpha_1, \alpha_2, \ldots, \alpha_p$. Estimate $\underset{\sim}{\delta} = (\delta_1, \delta_2, \ldots, \delta_p)'$ by regressing $W_{ij} = Y_{ij} - \overline{Y}_{i.}$ on $W_{i,j-1}, Z_{i,j-1}, \ldots, Z_{i,j-p+1}$, where $\overline{Y}_{i.} = n^{-1} \Sigma_{j=1}^{n} Y_{ij}$ is the least squares estimator of β_i. If p=1, then $\delta_1 = \alpha_1$ and the least squares estimator of δ_1 is obtained by regressing W_{ij} on $W_{i,j-1}$. Let

$$\underset{\sim}{\delta}^* = \hat{\underset{\sim}{D}}^{-1} \hat{\underset{\sim}{\phi}} \tag{3.30}$$

where

$$\hat{\underset{\sim}{D}} = [k(n-p)]^{-1} \sum_{i=0}^{k-1} \sum_{j=p+1}^{n} \underset{\sim}{B}'_{ij} \underset{\sim}{B}_{ij},$$

$$\hat{\underset{\sim}{\phi}} = [k(n-p)]^{-1} \sum_{i=0}^{k-1} \sum_{j=p+1}^{n} \underset{\sim}{B}'_{ij} W_{ij},$$

and

$$\begin{aligned} \underset{\sim}{B}_{ij} &= (W_{i,j-1}, Z_{i,j-1}, \ldots, Z_{i,j-p+1}), && \text{if } p > 2 \\ &= W_{i,j-1} \qquad\qquad\qquad\qquad\quad , && \text{if } p = 1. \end{aligned}$$

The error in the fitted regression equation is

$$\underset{\sim}{\delta}^* - \underset{\sim}{\delta} = \hat{\underset{\sim}{D}}^{-1} \hat{\underset{\sim}{M}} \tag{3.31}$$

where

$$\hat{\underset{\sim}{M}} = [k(n-p)]^{-1} \sum_{i=0}^{k-1} \sum_{j=p+1}^{n} \underset{\sim}{B}'_{ij} v_{ij},$$

$$v_{ij} = e_{ij} - \bar{e}_{i.} - \frac{\delta_1}{n} P_{in} + \frac{1}{n} \sum_{\ell=2}^{p} \delta_\ell P_{i,n-\ell+1}, \quad \text{if } p > 2$$

$$= e_{ij} - \bar{e}_{i.} - \frac{\delta_1}{n} P_{in}, \quad \text{if } p = 1,$$

$$\bar{e}_{i.} = n^{-1} \sum_{j=1}^{n} e_{ij},$$

and

$$P_{ij} = Y_{ij} - \beta_i.$$

Since the roots of (3.28) are assumed to be less than unity in absolute value, the roots of the polynomial equation, $m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p = 0$, lie inside the unit circle. Therefore, for each fixed i, P_{ij} is a stationary p-th order autoregressive process with mean zero. Using Theorem 3.2, an approximate expression for the mean of $\hat{\underset{\sim}{M}}$ is obtained in the following theorem.

Theorem 3.3. Consider a stationary p-th order seasonal autoregressive process $\{Y_t\}$ given by (3.26). Then,

$$E(\hat{\underset{\sim}{M}}) = \left(\frac{-\sigma^2}{(n-p)(1-\delta_1)}, 0, 0, \ldots, 0\right)' + O(n^{-2}), \quad \text{if } p > 2$$

$$= \frac{-\sigma^2}{(n-p)(1-\delta_1)} + O(n^{-2}), \quad \text{if } p=1,$$

where $\hat{\underset{\sim}{M}}$ is defined following (3.31).

Proof. We have

$$E(\hat{\underset{\sim}{M}}) = [k(n-p)]^{-1} \sum_{i=0}^{k-1} \sum_{j=p+1}^{n} E[\underset{\sim}{B}'_{ij} v_{ij}]$$

$$= k^{-1} \sum_{i=0}^{k-1} E[(n-p)^{-1} \sum_{j=p+1}^{n} \underset{\sim}{B}'_{ij} v_{ij}] \ .$$

For $p \geq 2$,

$$E(\hat{\underset{\sim}{M}}) = k^{-1} \sum_{i=0}^{k-1} [(- \frac{\sigma^2}{(n-p)(1-\delta_1)} , 0, 0, \ldots, 0)'] + O(n^{-2})$$

$$= (- \frac{\sigma^2}{(n-p)(1-\delta_1)} , 0, 0, \ldots, 0)' + O(n^{-2}) \ ,$$

and for $p = 1$,

$$E(\hat{\underset{\sim}{M}}) = k^{-1} \sum_{i=0}^{k-1} [- \frac{\sigma^2}{(n-p)(1-\delta_1)}] + O(n^{-2})$$

$$= - \frac{\sigma^2}{(n-p)(1-\delta_1)} + O(n^{-2}).$$

□

It follows from Theorem 3.3 that a method of correcting for the bias due to estimating the mean is to first construct $\underset{\sim}{\delta}^*$ using (3.30) and then obtain the new estimator

$$\tilde{\underset{\sim}{\delta}} = \hat{\underset{\sim}{D}}^{-1} [\hat{\underset{\sim}{\phi}} - \hat{\underset{\sim}{E}}(\hat{\underset{\sim}{M}})] \tag{3.32}$$

where $\hat{\underset{\sim}{E}}(\hat{\underset{\sim}{M}})$ is defined in (3.24).

To extend the method to the case in which $\delta_1 = 1$, we consider the following procedure.

(i) Obtain $\underset{\sim}{\delta}^*$ and $\hat{\sigma}^2$ as before,

(ii) In the construction of $\underset{\sim}{\delta}^*$, add $\hat{\underset{\sim}{f}}$ to $\hat{\underset{\sim}{\phi}}$ where

$$\begin{aligned} \hat{\underset{\sim}{f}} &= (\hat{f}_1, 0, \ldots, 0)', \quad \text{if } p \geq 2 \\ &= \hat{f}_1 \qquad\qquad\quad , \quad \text{if } p = 1, \end{aligned} \tag{3.33}$$

and,

$$\begin{aligned} \hat{f}_1 &= a \qquad , \quad \text{if } \hat{\tau}_{\mu k} \leq \hat{\tau}_{\mu k(01)} \\ &= a + b\,[\,\hat{\tau}_{\mu k(50)} - \hat{\tau}_{\mu k(01)}]^{-3}[\hat{\tau}_{\mu k} - \hat{\tau}_{\mu k(01)}]^3, \\ &\qquad \text{if } \hat{\tau}_{\mu k(01)} < \hat{\tau}_{\mu k} < \hat{\tau}_{\mu k(50)} \\ &= a + b \qquad , \quad \text{if } \hat{\tau}_{\mu k} \geq \hat{\tau}_{\mu k(50)}, \end{aligned}$$

$$a = [(n-p)(1 - \delta_1^*)]^{-1}\,\hat{\sigma}^2$$

$$b = (D'')^{-1}\,(1 - \delta_1^*) - a$$

$$\hat{\tau}_{\mu k} = \{\hat{V}(\delta_1^*)\}^{-1/2}\,(\delta_1^* - 1),$$

$$\hat{V}(\delta_1^*) = [k(n-p)]^{-1}\,D''\,\hat{\sigma}^2,$$

and $\hat{\tau}_{\mu k(\alpha)}$ is the α - percentile of the t-statistic, $\hat{\tau}_{\mu k}$. Fuller and Hasza have tabulated the percentiles of the statistic $\hat{\tau}_{\mu k}$, for $k = 1$, 4, and 12.

For autoregressive processes that contain seasonal means, but that are not pure seasonal in the sense of (3.26), a slightly different method of adjustment for bias is required.

CHAPTER IV. BIAS ADJUSTMENT FOR THE LEAST SQUARES ESTIMATORS OF A p-TH ORDER AUTOREGRESSIVE PROCESS WITH A NONCONSTANT MEAN

The mean of a stationary p-th order autoregressive process is constant, but the mean of an observed time series is often a function of time, other than the constant function. In many situations, we are able to specify the mean of a time series to be a simple function of time. Mean functions that often appear in practice are low order polynomials in t or trignometric polynomials in t.

Consider the model

$$Y_t = \underset{\sim}{X}_t \underset{\sim}{\beta} + P_t \tag{4.1}$$

where

$$\underset{\sim}{\beta} = (\beta_1, \beta_2, \ldots, \beta_r)',$$

$$\underset{\sim}{X}_t = (X_{t1}, X_{t2}, \ldots, X_{tr}),$$

$$P_t = \sum_{j=1}^{p} \alpha_j P_{t-j} + e_t,$$

and $\{e_t\}$ is a sequence of independent $(0,\sigma^2)$ random variables. It is assumed that the roots of the polynomial equation, $m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p = 0$, are less than unity in modulus. The elements X_{ti} are assumed to be fixed functions of time. Given a sample of n observations, the ordinary least squares estimator of $\underset{\sim}{\beta}$ is given by

$$\hat{\beta} = (X' X)^{-1} X'Y \tag{4.2}$$

where $X = (X_1', X_2', \ldots, X_n')'$ and $Y = (Y_1, Y_2, \ldots, Y_n)$.

The large sample behavior of $\hat{\beta}$ is given in the following theorem and is taken from Fuller (1976).

Theorem 4.1. Let the model (4.1) hold. Assume X_t is fixed and that the roots of the characteristic equation $m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p = 0$, are less than unity in modulus. Assume the e_t are independent $(0,\sigma^2)$ random variables with distributions $F_t(e)$ such that

$$\lim_{\delta\to\infty} \sup_t \int_{|e| > \delta} e^2 \, dF_t(e) = 0. \tag{4.3}$$

Assume that $\{X_{ti}\}$ satisfies

$$\lim_{n\to\infty} \sum_{t=1}^{n} X_{ti}^2 = \infty, \qquad i = 1,2,\ldots,r; \tag{4.4}$$

$$\lim_{n\to\infty} \frac{X_{ni}^2}{\sum_{t=1}^{n} X_{ti}^2} = 0, \qquad i = 1,2,\ldots,r; \tag{4.5}$$

and

$$\lim_{n\to\infty} \frac{\sum_{t=1}^{n-h} X_{ti} X_{t+h,j}}{\{ \sum_{t=1}^{n} X_{ti}^2 \sum_{t=1}^{n} X_{tj}^2 \}^{1/2}} = a_{hij} = a_{-hij}$$

$$h = 0, 1, 2, \ldots, \quad \text{and} \quad i, j = 1, 2, \ldots, r. \tag{4.6}$$

Assume that $\underset{\sim}{X}' \underset{\sim}{X}$ is positive definite for all $n > r$ and that $\underset{\sim}{A}_0$ defined by

$$\lim_{n \to \infty} \underset{\sim}{D}_n^{-1} \underset{\sim}{X}' \underset{\sim}{X} \underset{\sim}{D}_n^{-1} = \underset{\sim}{A}_0 \tag{4.7}$$

is a nonsingular matrix, where the diagonal matrix

$$\underset{\sim}{D}_n = \text{diag}\{(\sum_{t=1}^{n} X_{t1}^2)^{1/2}, (\sum_{t=1}^{n} X_{t2}^2)^{1/2}, \ldots, (\sum_{t=1}^{n} X_{tr}^2)^{1/2}\}. \tag{4.8}$$

Let $\underset{\sim}{B}$ be nonsingular, where the (i,j)-th element of $\underset{\sim}{B}$ is

$$b_{ij} = \sum_{h=-\infty}^{\infty} a_{hij} \, \gamma_P(h) \tag{4.9}$$

and

$$\gamma_P(h) = \text{Cov}(P_t, P_{t+h}).$$

Then

$$\underset{\sim}{D}_n (\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \xrightarrow{L} N(\underset{\sim}{0}, \underset{\sim}{A}_0^{-1} \underset{\sim}{B} \underset{\sim}{A}_0^{-1}).$$

<u>Proof.</u> See Fuller (1976). □

The assumptions of the above theorem are satisfied by polynomial functions of time if they are suitably transformed. See Fuller, Hasza, and Goebel (1981).

The least squares estimator $\underset{\sim}{\alpha}^*$ of $\underset{\sim}{\alpha}$ is obtained by regressing $W_t = Y_t - \underset{\sim}{X}_t \hat{\underset{\sim}{\beta}}$ on $W_{t-1}, W_{t-2}, \ldots, W_{t-p}$. We use an approach similar to that used in Chapter III to adjust for the bias in $\underset{\sim}{\alpha}^*$ that is due to the estimation of $\underset{\sim}{\beta}$.

It is assumed that there exist constants $\{C_m: m = 0, 1, \ldots, q\}$ such that

$$\sum_{m=0}^{q} C_m \underset{\sim}{X}_{t-m} = 0, \tag{4.10}$$

$$\sum_{m=0}^{q} C_m \underset{\sim}{X}_{t+m} = 0, \tag{4.11}$$

$$C_0 = 1,$$

and $r < q$. Note that for a stationary p-th order autoregressive process, with constant mean we have,

$$X_t \equiv 1, \; r = 1, \; q = 1, \quad \text{and} \quad C_1 = -1.$$

Consider a p-th order autoregressive process with

$$\begin{aligned} E\{Y_t\} &= \underset{\sim}{X}_t \underset{\sim}{\beta} \\ &= (1, t, \ldots, t^{r-1}) \underset{\sim}{\beta}. \end{aligned}$$

For this process, $q = r$ and since the r-th difference of $\underset{\sim}{X}_t$ is $\underset{\sim}{0}$, the constants C_m are given by

$$C_m = (-1)^m \binom{r}{m}, \qquad m = 0, 1, \ldots, r.$$

Consider the model (4.1). Define

$$\begin{aligned} Z_t &= \sum_{m=0}^{q} C_m Y_{t-m} \\ &= \sum_{m=0}^{q} C_m P_{t-m}. \end{aligned} \tag{4.12}$$

For $p > q$, consider a reparametrization of the model (4.1) given by

$$\begin{aligned} Y_t = \underset{\sim}{X}_t \underset{\sim}{\beta} &+ \sum_{i=1}^{q} \delta_i (Y_{t-i} - \underset{\sim}{X}_{t-i} \underset{\sim}{\beta}) \\ &+ \sum_{i=q+1}^{p} \delta_i Z_{t-i+q} + e_t \end{aligned} \tag{4.13}$$

where $\underset{\sim}{\delta} = (\delta_1, \delta_2, \ldots, \delta_p)' = \underset{\sim}{C}(\alpha_1, \alpha_2, \ldots, \alpha_p)'$ for some nonsingular matrix $\underset{\sim}{C}$. If $p \leq q$ then take $\underset{\sim}{\delta} = \underset{\sim}{\alpha}$. For $p > q$, the relation between $\underset{\sim}{\alpha}$ and $\underset{\sim}{\delta}$ is derived below. From (4.13),

$$\begin{aligned} P_t &= \sum_{i=1}^{q} \delta_i P_{t-i} + \sum_{i=q+1}^{p} \delta_i Z_{t-i+q} + e_t \\ &= \sum_{i=1}^{q} \delta_i P_{t-i} + \sum_{i=q+1}^{p} \delta_i \sum_{m=0}^{q} C_m P_{t-i+q-m} + e_t \\ &= \sum_{i=1}^{q} \delta_i P_{t-i} + \sum_{i=q+1}^{p} \delta_i \sum_{j=i-q}^{i} C_{j+q-i} P_{t-j} + e_t \end{aligned}$$

$$= \sum_{i=1}^{q} \delta_i P_{t-i} + \sum_{j=1}^{q} P_{t-j} \sum_{i=q+1}^{(j+q)\hat{o}p} C_{j+q-i}\, \delta_i$$

$$+ \sum_{j=q+1}^{p} P_{t-j} \sum_{i=j}^{(j+q)\hat{o}p} C_{j+q-i}\, \delta_i + e_t$$

$$= \sum_{j=1}^{q} (\delta_j + \sum_{i=q+1}^{(j+q)\hat{o}p} C_{j+q-i}\, \delta_i)\, P_{t-j}$$

$$+ \sum_{j=q+1}^{p} P_{t-j} \sum_{i=j}^{(j+q)\hat{o}p} C_{j+q-i}\, \delta_i + e_t$$

$$= \sum_{j=1}^{p} \alpha_j P_{t-j} + e_t,$$

where $j\hat{o}p = \min(j,p)$. Therefore,

$$\alpha_j = \delta_j + \sum_{\ell=q+1}^{(j+q)\hat{o}p} C_{j+q-\ell}\, \delta_\ell, \qquad j < q \tag{4.14}$$

$$= \sum_{\ell=j}^{(j+q)\hat{o}p} C_{j+q-\ell}\, \delta_\ell \quad , \qquad j > q \tag{4.15}$$

For $p > q$

$$\sum_{j=1}^{p} \alpha_j = \sum_{j=1}^{q} \delta_j + \sum_{j=1}^{q} \sum_{\ell=q+1}^{(j+q)\hat{o}p} C_{j+q-\ell}\, \delta_\ell + \sum_{j=q+1}^{p} \sum_{\ell=j}^{(j+q)\hat{o}p} C_{j+q-\ell}\, \delta_\ell$$

$$= \sum_{j=1}^{q} \delta_j + \sum_{j=1}^{p} \delta_j \sum_{m=0}^{q} C_m. \tag{4.16}$$

For $p < q$, $\sum_{j=1}^{p} \alpha_j = \sum_{j=1}^{p} \delta_j$.

Since $\{P_t\}$ is a stationary p-th order autoregressive process,

$$P_t = \sum_{j=0}^{\infty} w_j \, e_{t-j}, \tag{4.17}$$

where $\{w_j\}$ satisfies the p-th order difference equation

$$w_j - \alpha_1 w_{j-1} - \cdots - \alpha_p w_{j-p} = 0, \qquad j = 1,2,\ldots, \tag{4.18}$$

$$w_j = 0 \qquad \text{if } j < 0, \text{ and } w_0 = 1.$$

From (4.13), for $p > q$,

$$\begin{aligned}
W_t &= Y_t - \underset{\sim}{X}_t \hat{\underset{\sim}{\beta}} \\
&= \underset{\sim}{X}_t(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}}) + \sum_{i=1}^{q} \delta_i (Y_{t-i} - \underset{\sim}{X}_{t-i} \underset{\sim}{\beta}) \\
&\qquad + \sum_{i=q+1}^{p} \delta_i \, Z_{t-i+q} + e_t \\
&= \sum_{i=1}^{q} \delta_i \, W_{t-i} + \sum_{i=q+1}^{p} \delta_i \, Z_{t-i+q} + e_t + \underset{\sim}{X}_t(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}}) \\
&\qquad - \sum_{i=1}^{q} \delta_i \, \underset{\sim}{X}_{t-i}(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}}) \\
&= \sum_{i=1}^{q} \delta_i \, W_{t-i} + \sum_{i=q+1}^{p} \delta_i \, Z_{t-i+q} + v_t,
\end{aligned}$$

where

$$v_t = e_t - (\underset{\sim}{X}_t - \sum_{i=1}^{q} \delta_i \underset{\sim}{X}_{t-i})(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta}),$$

$$= e_t - d_t \quad \text{(say)}.$$

For $p < q$,

$$W_t = Y_t - X_t \hat{\underset{\sim}{\beta}}$$

$$= \underset{\sim}{X}_t(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}}) + \sum_{i=1}^{p} \delta_i(Y_{t-i} - \underset{\sim}{X}_{t-i} \underset{\sim}{\beta}) + e_t$$

$$= \sum_{i=1}^{p} \delta_i W_{t-i} + e_t - (\underset{\sim}{X}_t - \sum_{i=1}^{p} \delta_i \underset{\sim}{X}_{t-i})(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta}).$$

Let $\tilde{p} = \min(p,q)$. Then

$$W_t = \sum_{i=1}^{\tilde{p}} \delta_i W_{t-i} + 1_{(p>q)} \cdot \sum_{i=q+1}^{p} \delta_i Z_{t-i+q} + v_t, \tag{4.19}$$

where

$$v_t = e_t - d_t,$$

$$d_t = (\underset{\sim}{X}_t - \sum_{i=1}^{\tilde{p}} \delta_i \underset{\sim}{X}_{t-i})(\hat{\underset{\sim}{\beta}}- \underset{\sim}{\beta}),$$

and

$$1_{(p>q)} = 0, \quad \text{if } p \leq q$$
$$= 1, \quad \text{if } p > q.$$

The ordinary least squares estimator $\underset{\sim}{\delta}^*$ of $\underset{\sim}{\delta}$ is

$$\underset{\sim}{\delta}^* = [(n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{F}_t' \underset{\sim}{F}_t]^{-1} [(n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{F}_t' W_t], \tag{4.20}$$

and the error in the estimator is

$$\underset{\sim}{\delta}^* - \underset{\sim}{\delta} = [(n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{F}_t' \underset{\sim}{F}_t]^{-1} [(n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{F}_t' v_t]$$

where

$$\underset{\sim}{F}_t = (W_{t-1}, W_{t-2}, \ldots, W_{t-q}, Z_{t-1}, \ldots, Z_{t-p+q}), \quad \text{if } p > q$$
$$= (W_{t-1}, W_{t-2}, \ldots, W_{t-p}) \quad , \quad \text{if } p \leq q.$$

The following theorem gives the expected values of the elements of the vector $\Sigma_{t=p+1}^{n} \underset{\sim}{F}_t' v_t$.

<u>Theorem 4.2.</u> Let $\{Y_t\}$ be a stochastic process satisfying

$$Y_t = \underset{\sim}{X}_t \underset{\sim}{\beta} + P_t$$

where $\{P_t\}$ is a stationary p-th order autoregressive process satisfying

$$P_t = \sum_{j=1}^{p} \alpha_j P_{t-j} + e_t .$$

Assume $\{e_t\}$ is a sequence of independent $(0,\sigma^2)$ random variables. Assume that the roots of the polynomial equation,

$$m^p - \alpha_1 m^{p-1} - \dots - \alpha_p = 0,$$

lie inside the unit circle. It is also assumed that there exists $\{C_m: m = 0,1,\dots,q\}$ such that (4.10) and (4.11) hold. Then

$$E[\sum_{t=p+1}^{n} W_{t-i} v_t] = - \sum_{t=p+1}^{n} \underset{\sim}{X}_{t-i} (\underset{\sim}{X}' \underset{\sim}{X})^{-1} \underset{\sim}{X}' \underset{\sim}{w}_t \sigma^2$$

$$- \sum_{j=p+1}^{n} (\underset{\sim}{M}'_j - \sum_{\ell=1}^{\tilde{p}} \delta_\ell \underset{\sim}{M}'_{j-\ell})(\underset{\sim}{\Gamma}_{j-i} - \underset{\sim}{\Gamma}\underset{\sim}{M}_{j-i}), \tag{4.21}$$

and for $p > q$,

$$\sum_{t=p+1}^{n} E[Z_{t-s} v_t] = - \sum_{j=p+1-q}^{p} \sum_{m=p+1-j}^{(n-j)\hat{o}q} C_m (\underset{\sim}{M}'_{j+m} - \sum_{\ell=1}^{q} \delta_\ell \underset{\sim}{M}'_{j-\ell+m}) \underset{\sim}{\Gamma}_{j-s}$$

$$- \sum_{j=n-q+1}^{n} \sum_{m=(p+1-j)v0}^{n-j} C_m (\underset{\sim}{M}'_{j+m} - \sum_{\ell=1}^{q} \delta_\ell \underset{\sim}{M}'_{j-\ell+m}) \underset{\sim}{\Gamma}_{j-s}, \tag{4.22}$$

for $i = 1, 2, \dots, q$; $s = 1, 2, \dots, p-q$, where v_t is given by (4.19), $\underset{\sim}{P} = (P_1, P_2, \dots, P_n)'$, $\underset{\sim}{X}$ is given by (4.2), $pvj = \max(p,j)$,

w_j is given by (4.18),

$$\underset{\sim}{M}_j' = \underset{\sim}{X}_j (\underset{\sim}{X}' \underset{\sim}{X})^{-1} \underset{\sim}{X}',$$

$$\underset{\sim}{\Gamma} = E(\underset{\sim}{P} \underset{\sim}{P}') = (\underset{\sim}{\Gamma}_1, \underset{\sim}{\Gamma}_2, \ldots, \underset{\sim}{\Gamma}_n), \tag{4.23}$$

$$\underset{\sim}{\Gamma}_i = (\gamma_p(i-1), \gamma_p(i-2), \ldots, \gamma_p(i-n))',$$

$$\gamma_p(j) = \mathrm{Cov}(P_t, P_{t-j}),$$

and

$$\underset{\sim}{w}_t = (w_{1-t}, w_{2-t}, \ldots, w_{n-t})'.$$

Proof. For a fixed i ($= 1, 2, \ldots, \tilde{p}$), consider

$$E[\sum_{t=p+1}^{n} W_{t-i} v_t] = \sum_{t=p+1}^{n} E[W_{t-i}(e_t - d_t)]$$

$$= \sum_{t=p+1}^{n} E(W_{t-i} e_t) - \sum_{t=p+1}^{n} E(W_{t-i} d_t)$$

We have

$$E[\sum_{t=p+1}^{n} W_{t-i} e_t] = \sum_{t=p+1}^{n} E[\{P_{t-i} + \underset{\sim}{X}_{t-i}(\underset{\sim}{\beta} - \hat{\underset{\sim}{\beta}})\} e_t]$$

$$= - \sum_{t=p+1}^{n} E[\underset{\sim}{X}_{t-i}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) e_t]$$

since P_{t-i} is independent of e_t. Therefore,

$$E[\sum_{t=p+1}^{n} W_{t-i}\, e_t] = -\sum_{t=p+1}^{n} E[\underset{\sim}{X}_{t-i}(\underset{\sim}{X}'\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}'\,\underset{\sim}{P}\,e_t]$$

$$= -\sum_{t=p+1}^{n} \underset{\sim}{X}_{t-i}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\,\underset{\sim}{X}'\,\underset{\sim}{w}_t\,\sigma^2$$

where

$$\underset{\sim}{w}_t = \sigma^{-2}\,E(\underset{\sim}{P}\,e_t)$$

$$= (w_{1-t}, w_{2-t}, \ldots, w_{n-t})'.$$

Now,

$$E[\sum_{t=p+1}^{n} W_{t-i}\,d_t] = \sum_{t=p+1}^{n} E[W_{t-i}(\underset{\sim}{X}_t - \sum_{\ell=1}^{\tilde{p}} \delta_\ell\,\underset{\sim}{X}_{t-\ell})(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})].$$

Consider for a fixed ℓ,

$$\sum_{t=p+1}^{n} E[\underset{\sim}{W}_{t-i}\,\underset{\sim}{X}_{t-\ell}(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})]$$

$$= \sum_{j=p+1-i}^{n-i} E[W_j\,\underset{\sim}{X}_{j-\ell+i}(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})]$$

$$= \sum_{j=p+1-i}^{n-i} E[(P_j - \underset{\sim}{X}_j(\underset{\sim}{X}'\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}'\,\underset{\sim}{P})\,\underset{\sim}{X}_{j-\ell+i}(\underset{\sim}{X}'\,\underset{\sim}{X})^{-1}\underset{\sim}{X}'\,\underset{\sim}{P}]$$

$$= \sum_{j=p+1-i}^{n-i} \underset{\sim}{X}_{j-\ell+i}(\underset{\sim}{X}'\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}'\,E(\underset{\sim}{P}\,P_j)$$

$$- \sum_{j=p+1-i}^{n-i} \underset{\sim}{X}_{j-\ell+i}(\underset{\sim}{X}'\,\underset{\sim}{X})^{-1}\,\underset{\sim}{X}'\,E(\underset{\sim}{P}\underset{\sim}{P}')\,\underset{\sim}{X}\,(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\,\underset{\sim}{X}_j'$$

$$= \sum_{j=p+1-i}^{n-i} \underset{\sim}{M}'_{j-\ell+i}\, \underset{\sim}{\Gamma}_j - \sum_{j=p+1-i}^{n-i} \underset{\sim}{M}'_{j-\ell+i}\, \underset{\sim}{\Gamma}\, \underset{\sim}{M}_j$$

$$= \sum_{j=p+1}^{n} \underset{\sim}{M}'_{j-\ell}\, (\underset{\sim}{\Gamma}_{j-i} - \underset{\sim}{\Gamma}\, \underset{\sim}{M}_{j-i}).$$

Therefore,

$$E[\sum_{t=p+1}^{n} W_{t-i}\, v_t] = - \sum_{t=p+1}^{n} \underset{\sim}{M}'_{t-i}\, \underset{\sim}{w}_t\, \sigma^2$$

$$- \sum_{t=p+1}^{n} (\underset{\sim}{M}'_t - \sum_{\ell=1}^{\tilde{p}} \delta_\ell\, \underset{\sim}{M}'_{t-\ell})(\underset{\sim}{\Gamma}_{t-i} - \underset{\sim}{\Gamma}\, \underset{\sim}{M}_{t-i}),$$

where $\underset{\sim}{M}'_t$ is the t-th row of $\underset{\sim}{M}$. Let us now consider, for a fixed s, assuming $p > q$,

$$E[\sum_{t=p+1}^{n} Z_{t-s}\, v_t] = E[\sum_{t=p+1}^{n} Z_{t-s}\, e_t] - E[\sum_{t=p+1}^{n} Z_{t-s}\, d_t]$$

$$= - E[\sum_{t=p+1}^{n} Z_{t-s}\{\underset{\sim}{X}_t - \sum_{\ell=1}^{q} \delta_\ell\, \underset{\sim}{X}_{t-\ell}\}(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})]$$

since Z_{t-s} is independent of e_t. Consider for a fixed ℓ,

$$E[\sum_{t=p+1}^{n} Z_{t-s}\, \underset{\sim}{X}_{t-\ell}(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})].$$

Note,

$$\sum_{t=p+1}^{n} Z_{t-s}\, \underset{\sim}{X}_{t-\ell} = \sum_{t=p+1}^{n} \sum_{m=0}^{q} C_m\, P_{t-s-m}\, \underset{\sim}{X}_{t-\ell}$$

$$= \sum_{m=0}^{q} C_m \sum_{t=p+1}^{n} P_{t-s-m} \underset{\sim}{X}_{t-\ell}$$

$$= \sum_{m=0}^{q} C_m \sum_{j=p+1-m}^{n-m} P_{j-s} \underset{\sim}{X}_{j-\ell+m}$$

$$= \sum_{j=p+1-q}^{n} P_{j-s} \sum_{m=(p+1-j)\vee 0}^{(n-j)\wedge q} C_m \underset{\sim}{X}_{j-\ell+m}$$

$$= \sum_{j=p+1-q}^{p} P_{j-s} \sum_{m=p+1-j}^{(n-j)\wedge q} C_m \underset{\sim}{X}_{j-\ell+m}$$

$$+ \sum_{j=n-q+1}^{n} P_{j-s} \sum_{m=(p+1-j)\vee 0}^{n-j} C_m \underset{\sim}{X}_{j-\ell+m}$$

since $\sum_{m=0}^{q} C_m \underset{\sim}{X}_{j+m-\ell} = 0$. Therefore,

$$E[\sum_{t=p+1}^{n} Z_{t-s} \underset{\sim}{X}_{t-\ell}(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})]$$

$$= \sum_{j=p+1-q}^{p} \sum_{m=p+1-j}^{(n-j)\wedge q} C_m \underset{\sim}{X}_{j-\ell+m}(\underset{\sim}{X}'\underset{\sim}{X})^{-1} \underset{\sim}{X}' E(\underset{\sim}{P}\underset{\sim}{P}_{j-s})$$

$$+ \sum_{j=n-q+1}^{n} \sum_{m=(p+1-j)\vee 0}^{n-j} C_m \underset{\sim}{X}_{j-\ell+m}(\underset{\sim}{X}'\underset{\sim}{X})^{-1} \underset{\sim}{X}' E(\underset{\sim}{P}\underset{\sim}{P}_{j-s})$$

$$= \sum_{j=p+1-q}^{p} \sum_{m=p+1-j}^{(n-j)\wedge q} C_m \underset{\sim}{M}'_{j-\ell+m} \underset{\sim}{\Gamma}_{j-s}$$

$$+ \sum_{j=n-q+1}^{n} \sum_{m=(p+1-j)\vee 0}^{n-j} C_m \underset{\sim}{M}'_{j-\ell+m} \underset{\sim}{\Gamma}_{j-s}$$

and

$$E[\sum_{t=p+1}^{n} Z_{t-s} v_t] = -\sum_{j=p+1-q}^{p} \sum_{m=p+1-j}^{(n-j)\hat{o}q} C_m(\underset{\sim}{M}'_{j+m} - \sum_{\ell=1}^{q} \delta_\ell \underset{\sim}{M}'_{j-\ell+m}) \underset{\sim}{\Gamma}_{j-s}$$

$$- \sum_{j=n-q+1}^{n} \sum_{m=(p+1-j)v0}^{n-j} C_m(\underset{\sim}{M}'_{j+m} - \sum_{\ell=1}^{q} \delta_\ell \underset{\sim}{M}'_{j-\ell+m}) \underset{\sim}{\Gamma}_{j-s}.$$

□

The elements of $E[\Sigma_{t=p+1}^{n} \underset{\sim}{F}'_t v_t]$ are expressed as linear combinations of $\underset{\sim}{M}'_j \underset{\sim}{\Gamma}_\ell$ and $\underset{\sim}{M}'_j \underset{\sim}{w}_\ell$. Suppose there exists a finite real number L such that

$$\sup_{1<i<n} |\underset{\sim}{X}_i(\underset{\sim}{X}' \underset{\sim}{X})^{-1} \underset{\sim}{X}'_i| < \frac{L}{n} . \tag{4.24}$$

Under the assumption (4.24), we obtain the order of the elements of $E[\Sigma_{t=p+1}^{n} \underset{\sim}{F}'_t v_t]$ in the following theorem.

<u>Theorem 4.3.</u> Let $\{Y_t\}$ be a stochastic process satisfying the conditions of Theorem 4.2. Assume (4.24) is satisfied. Then

$$E[\sum_{t=p+1}^{n} W_{t-i} v_t] = -\sum_{t=p+1}^{n} \underset{\sim}{M}'_{t-i} \underset{\sim}{w}_t \sigma^2$$

$$- \sum_{j=p+1-i}^{n-i} (\underset{\sim}{M}'_{j+i} - \underset{\sim}{M}'_j) \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j$$

$$+ \sum_{\ell=1}^{\tilde{p}} \delta_\ell \sum_{j=p+1-i}^{n-i} (\underset{\sim}{M}'_{j+i-\ell} - \underset{\sim}{M}'_j) \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j$$

$$+ O(n^{-1})$$

$$= O(1), \quad \text{for } i = 1,2,\ldots,\tilde{p},$$

where $\underset{\sim}{M} = \underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'$ and $(\underset{\sim}{I} - \underset{\sim}{M})_j$ is the j-th column of $(\underset{\sim}{I} - \underset{\sim}{M})$, and for $p > q$,

$$E[\sum_{t=p+1}^{n} Z_{t-s} v_t] = O(n^{-1}), \qquad s = 1, 2, \ldots, p-q.$$

Proof. Note,

$$|M_{ij}| = |\underset{\sim}{X}_i(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}_j'|$$

$$\leq |\underset{\sim}{X}_i(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}_i'|^{1/2}\,|\underset{\sim}{X}_j(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}_j'|^{1/2}$$

$$\leq \frac{L}{n}.$$

Also,

$$\sum_{t=p+1}^{n} \underset{\sim}{M}_{t-i}'\underset{\sim}{w}_t = \sum_{t=p+1}^{n}\sum_{\ell=1}^{n} M_{\ell,t-i}\, w_{\ell-t}$$

$$= \sum_{t=p+1}^{n}\sum_{\ell=t}^{n} M_{\ell,t-i}\, w_{\ell-t}$$

$$= \sum_{t=p+1}^{n}\sum_{\ell=0}^{n-t} M_{t+\ell,t-i}\, w_{\ell}$$

$$= \sum_{\ell=0}^{n-p-1} w_{\ell} \sum_{t=p+1}^{n-\ell} M_{t+\ell,t-i}. \tag{4.25}$$

Therefore,

$$\left| \sum_{t=p+1}^{n} \underset{\sim}{M}'_{t-i}\, \underset{\sim}{w}_t \right| < \sum_{\ell=0}^{n-p-1} \left| w_\ell \right| L \cdot \frac{(n-p-\ell)}{n}$$

$$< \sum_{\ell=0}^{\infty} \left| w_\ell \right|,$$

which is finite. Therefore,

$$\sum_{t=p+1}^{n} \underset{\sim}{M}'_{t-i}\, \underset{\sim}{w}_t = O(1).$$

Now consider

$$\sum_{j=p+1}^{n} \underset{\sim}{M}'_{j-\ell}\, (\underset{\sim}{\Gamma}_{j-i} - \underset{\sim}{\Gamma}\, \underset{\sim}{M}_{j-i})$$

$$= \sum_{j=p+1-i}^{n-i} \underset{\sim}{M}'_{j+i-\ell}\, \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j$$

$$= \sum_{j=p+1-i}^{n-i} \underset{\sim}{M}'_j\, \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j$$

$$+ \sum_{j=p+1-i}^{n-i} (\underset{\sim}{M}'_{j+i-\ell} - \underset{\sim}{M}'_j)\, \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j$$

$$= \sum_{j=1}^{n} \underset{\sim}{M}'_j\, \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j - \sum_{j=1}^{p-i} \underset{\sim}{M}'_j\, \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j$$

$$- \sum_{j=n-i+1}^{n} \underset{\sim}{M}'_j\, \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M}_j)$$

$$+ \sum_{j=p+1-i}^{n-i} (\underset{\sim}{M}'_{j+i-\ell} - \underset{\sim}{M}'_j)\, \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j.$$

Since $(\underset{\sim}{I} - \underset{\sim}{M}) \underset{\sim}{M}' = \underset{\sim}{0}$, we have

$$\sum_{j=1}^{n} (\underset{\sim}{I} - \underset{\sim}{M})_j \underset{\sim}{M}_j' = \underset{\sim}{0}.$$

Therefore,

$$\sum_{j=1}^{n} \underset{\sim}{M}_j' \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j = \mathrm{tr}[\underset{\sim}{\Gamma} \sum_{j=1}^{n} (\underset{\sim}{I} - \underset{\sim}{M})_j \underset{\sim}{M}_j']$$

$$= 0. \tag{4.26}$$

Also,

$$\left|\underset{\sim}{M}_j' \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M}_j)\right| = \left|\sum_{i=1}^{n} M_{ij} \left[\gamma(j-i) - \sum_{s=1}^{n} M_{sj} \gamma(i-s)\right]\right|$$

$$< 2 \frac{(L + L^2)}{n} \sum_{h=0}^{n} \left|\gamma(h)\right|. \; . \tag{4.27}$$

Therefore,

$$\sum_{j=1}^{p-i} \underset{\sim}{M}_j' \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j = O(n^{-1}),$$

and

$$\sum_{j=n+1-i}^{n} \underset{\sim}{M}_j' \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j = O(n^{-1}).$$

From (4.21) it follows that

$$E[\sum_{t=p+1}^{n} W_{t-i} v_t]$$

$$= - \sum_{t=p+1}^{n} \underset{\sim}{M}'_{t-i} \underset{\sim}{w}_t \sigma^2 - \sum_{j=p+1-i}^{n-i} (\underset{\sim}{M}'_{j+i} - \underset{\sim}{M}'_j) \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j$$

$$+ \sum_{\ell=1}^{\tilde{p}} \delta_\ell \sum_{j=p+1-i}^{n-i} (\underset{\sim}{M}'_{j+i-\ell} - \underset{\sim}{M}'_j) \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M})_j$$

$$+ O(n^{-1})$$

$$= O(1).$$

Since,

$$\sup_{j,\ell} \left| \underset{\sim}{M}'_j \underset{\sim}{\Gamma}_\ell \right| < \frac{L}{n} \sum_{h=0}^{\infty} \left| \gamma(h) \right|$$

and since $E[\ \Sigma_{j=p+1}^{n} Z_{t-s} v_t]$ is a linear combination of fixed number of $\underset{\sim}{M}'_j \underset{\sim}{\Gamma}_\ell$, we get

$$E[\sum_{t=p+1}^{n} Z_{t-s} v_t] = O(n^{-1}). \qquad \square$$

For $p > q$, we observe that the bias in the right hand side of the equation associated with $(\hat{\delta}_{q+1}, \hat{\delta}_{q+2}, \ldots, \hat{\delta}_p)$ arising from estimating the mean function is of order n^{-2} and that in $(\hat{\delta}_1, \hat{\delta}_2, \ldots, \hat{\delta}_q)$ is of order n^{-1}. (The order of the bias in the right hand side is the order of the bias for the sample covariance and is the order of the bias for δ_i.)

We will evaluate the bias for the special case where the mean function is a polynomial in t. Consider,

$$E(Y_t) = \underset{\sim}{X}_t \underset{\sim}{\beta} \tag{4.28}$$

where $\underset{\sim}{X}_t = (t^{i(1)}, t^{i(2)}, \ldots, t^{i(r)})$ and $0 \leq i(1) < i(2) < \ldots < i(r)$ are integers. If $i(j) = j-1$, for $j = 1, 2, \ldots, r$, then the mean function of Y_t is a polynomial of degree $r - 1$. For the choice of $\underset{\sim}{X}_t$ in (4.28), the mean function is a polynomial of degree $i(r)$ with some of the coefficients of the polynomial restricted to be zero.

Let $q = i(r) + 1$. Note the q-th difference of $\underset{\sim}{X}_t$ is zero. With $C_m = (-1)^m \binom{q}{m}$, the conditions (4.10) and (4.11) are satisfied. Since

$$\sum_{t=1}^{n} t^{i(j)} = \frac{n^{i(j)+1}}{i(j)+1} + O(n^{i(j)}),$$

we have

$$(\underset{\sim}{X}' \underset{\sim}{X})_{\ell s} = \frac{n^{i(\ell)+i(s)+1}}{i(\ell)+i(s)+1} + O(n^{i(\ell)+i(s)}), \tag{4.29}$$

and

$$(\underset{\sim}{X}' \underset{\sim}{X})^{\ell s} = \frac{C_{\ell s}}{n^{i(\ell)+i(s)+1}} + O\left(\frac{1}{n^{i(\ell)+i(s)+2}} \right),$$

where $(\underset{\sim}{X}'\ \underset{\sim}{X})_{\ell s}$ and $(\underset{\sim}{X}'\underset{\sim}{X})^{\ell s}$ are (ℓ,s)-th elements of $\underset{\sim}{X}'\underset{\sim}{X}$ and $(\underset{\sim}{X}'\ \underset{\sim}{X})^{-1}$, respectively and $C_{\ell s}$ are fixed constants. Because

$$\sum_{s=1}^{r} (\underset{\sim}{X}'\ \underset{\sim}{X})_{\ell s}\ (\underset{\sim}{X}'\ \underset{\sim}{X})^{\ell s} = 1,$$

we have

$$\sum_{s=1}^{r} \frac{C_{\ell s}}{i(\ell)+i(s)+1} = 1 + O(n^{-1}). \tag{4.30}$$

The approximate bias expressions for polynomial trends are evaluated in the following theorem.

<u>Theorem 4.4.</u> Let $\{Y_t\}$ be a stochastic process satisfying the conditions of Theorem 4.2. Assume $\underset{\sim}{X}_t$ is given by (4.28). Then,

$$E[\sum_{t=p+1}^{n} W_{t-i}\ v_t] = -\ r\ \sigma^2\ (\sum_{j=0}^{\infty} w_j) + O(n^{-1}), \quad i = 1,2,\ldots,\breve{p},$$

and for $p > q$,

$$E[\sum_{t=p+1}^{n} Z_{t-s}\ v_t] = O(n^{-1}), \quad s = 1,2,\ldots,p-q.$$

<u>Proof.</u> We will first verify that $\underset{\sim}{X}_t$ satisfies the conditions of Theorem 4.3. We need to compute

$$M_{\ell s} = \underset{\sim}{X}_\ell(\underset{\sim}{X}'\ \underset{\sim}{X})^{-1}\ \underset{\sim}{X}_s'$$

$$= \sum_{j=1}^{r} X_{\ell j} \sum_{m=1}^{r} (\underset{\sim}{X}' \underset{\sim}{X})^{jm} X_{sm}$$

$$= \frac{1}{n} \sum_{j=1}^{r} \left(\frac{\ell}{n} \right)^{i(j)} \sum_{m=1}^{r} C_{jm} \left(\frac{s}{n} \right)^{i(m)} + 0(n^{-2}). \qquad (4.31)$$

Therefore,

$$\sup_{\ell,s} \left| M_{\ell s} \right| < \frac{L}{n}$$

with $L = 1 + \sum_{j=1}^{r} \sum_{m=1}^{r} \left| C_{jm} \right|$. By Theorem 4.3, for $p > q$,

$$E[\sum_{t=p+1}^{n} Z_{t-s} v_t] = 0(n^{-1}).$$

From (4.25),

$$\sum_{t=p+1}^{n} \underset{\sim}{M}'_{t-d} \underset{\sim}{w}_t = \sum_{\ell=0}^{n-p-1} w_\ell \sum_{t=p+1}^{n-\ell} M_{t+\ell,t-d}$$

$$= \sum_{\ell=0}^{n-p-1} w_\ell \sum_{t=p+1}^{n-\ell} \left\{ \sum_{j=1}^{r} \sum_{m=1}^{r} C_{jm} \frac{(t+\ell)^{i(j)}(t-d)^{i(m)}}{n^{i(j)+i(m)+1}} \right\}$$

$$+ 0(n^{-1})$$

$$= \sum_{\ell=0}^{n-p-1} w_\ell \sum_{j=1}^{r} \sum_{n=1}^{r} \frac{C_{jm}}{n} \sum_{t=p+1}^{n-\ell} \frac{(t+\ell)^{i(j)}(t-d)^{i(m)}}{n^{i(j)+i(m)}}$$

$$+ 0(n^{-1})$$

$$= \sum_{\ell=0}^{n-p-1} w_\ell \sum_{j=1}^{r} \sum_{m=1}^{r} \frac{C_{jm}}{n} \sum_{t=p+1}^{n-\ell} \frac{(t+\ell)^{i(j)} t^{i(m)}}{n^{i(j)+i(m)}}$$

$$+ O(n^{-1})$$

$$= \sum_{j=1}^{r} \sum_{m=1}^{r} \frac{C_{jm}}{n} \sum_{\ell=0}^{n-p-1} w_\ell \sum_{t=p+1}^{n-\ell} \frac{(t+\ell)^{i(j)} t^{i(m)}}{n^{i(j)+i(m)}}$$

$$+ O(n^{-1})$$

$$= \sum_{j=1}^{r} \sum_{m=1}^{r} \frac{C_{jm}}{n} \sum_{\ell=0}^{n-p-1} w_\ell \sum_{s=0}^{i(j)} \frac{\binom{i(j)}{s} \ell^{i(j)-s}}{n^{i(j)+i(m)}}$$

$$\sum_{t=p+1}^{n-\ell} t^{s+i(m)} + O(n^{-1})$$

$$= \sum_{j=1}^{r} \sum_{m=1}^{r} \frac{C_{jm}}{n} \sum_{\ell=0}^{n-p-1} \frac{w_\ell}{n^{i(j)+i(m)}}$$

$$\sum_{s=0}^{i(j)} \binom{i(j)}{s} \frac{\ell^{i(j)-s} (n-\ell)^{s+i(m)+1}}{s + i(m) + 1}$$

$$+ O(n^{-1}).$$

For $s \neq i(j)$,

$$\left| \sum_{\ell=0}^{n-p-1} \frac{w_\ell}{n^{i(j)+i(m)+1}} \ell^{i(j)-s} (n-\ell)^{s+i(m)+1} \right|$$

$$< \sum_{\ell=0}^{n-p-1} | w_\ell | \frac{\ell^{i(j)-s}}{n^{i(j)-s}} \left(1 - \frac{\ell}{n} \right)^{s+i(m)+1}$$

$$< \sum_{\ell=0}^{n-p-1} | w_\ell | \left(\frac{\ell}{n} \right)^{i(j)-s}.$$

Since,

$$\sum_{\ell=1}^{\infty} \ell \, | w_\ell | < \infty,$$

we have for $s \neq i(j)$,

$$n^{-1} \sum_{\ell=0}^{n-p-1} \ell \, | w_\ell | \left(\frac{\ell}{n} \right)^{i(j)-s-1} = O(n^{-1}).$$

Therefore,

$$\sum_{t=p+1}^{n} \underset{\sim}{M}'_{t-d} \underset{\sim}{w}_t = \sum_{j=1}^{r} \sum_{m=1}^{r} \frac{C_{jm}}{n} \sum_{\ell=0}^{n-p-1} \frac{w_\ell}{n^{i(j)+i(m)}} \frac{(n-\ell)^{i(j)+i(m)+1}}{i(j)+i(m)+1}$$

$$+ O(n^{-1})$$

$$= \sum_{j=1}^{r} \sum_{m=1}^{r} \frac{C_{jm}}{i(j)+i(m)+1} \sum_{\ell=0}^{n-p-1} w_\ell \left(1 - \frac{\ell}{n}\right)^{i(j)+i(m)+1}$$

$$+ O(n^{-1})$$

$$= \sum_{j=1}^{r} \sum_{m=1}^{r} \frac{C_{jm}}{i(j)+i(m)+1} \sum_{\ell=0}^{n-p-1} w_\ell + O(n^{-1})$$

$$= \left(\sum_{\ell=0}^{\infty} w_\ell \right) \sum_{j=1}^{r} \sum_{m=1}^{r} \frac{C_{jm}}{i(j)+i(m)+1} + O(n^{-1}).$$

Using (4.30) we get,

$$\sum_{t=p+1}^{n} \underset{\sim}{M}'_{t-d}\, \underset{\sim}{w}_t = r \sum_{\ell=0}^{\infty} w_\ell + O(n^{-1}).$$

Now consider, for a fixed h,

$$M_{s,j+h} - M_{s,j} = \frac{1}{n} \sum_{\ell=1}^{r} \sum_{m=1}^{r} \frac{C_{\ell m}\, s^{i(\ell)}[(j+h)^{i(m)} - j^{i(m)}]}{n^{i(m)+i(\ell)}} + O(n^{-2})$$

$$= \frac{1}{n} \sum_{\ell=1}^{r} \sum_{m=1}^{r} C_{\ell m} \left(\frac{s}{n}\right)^{i(\ell)} \sum_{t=0}^{i(m)-1} \binom{i(m)}{t} \left(\frac{j}{n}\right)^{t} \left(\frac{h}{n}\right)^{i(m)-t} + O(n^{-2})$$

$$= O(n^{-2}).$$

Therefore, for a fixed h,

$$\sup_{1<j,s<n} \left| M_{s,j+h} - M_{s,j} \right| < \frac{\tilde{L}}{n^2}$$

for some finite real number $\tilde{L}$. Therefore, for a fixed i and ℓ,

$$\left| \sum_{j=p+1-i}^{n-i} (\underset{\sim}{M}'_{j+i-\ell} - \underset{\sim}{M}'_j)\, \underset{\sim}{\Gamma}(\underset{\sim}{I} - \underset{\sim}{M}_j) \right|$$

$$= \left| \sum_{j=p+1-i}^{n-i} \sum_{s=1}^{n} (M_{s,j+i-\ell} - M_{s,j}) \right.$$

$$\left. [\gamma(j+i-\ell-s) - \sum_{m=1}^{n} M_{m,j}\, \gamma(m-s)] \right|$$

$$< \frac{2(\tilde{L} + \tilde{L}^2)}{n} \sum_{j=0}^{\infty} |\, \gamma(j)\, |.$$

Therefore, by Theorem 4.3.,

$$E[\sum_{t=p+1}^{n} W_{t-i}\, v_t] = -\, r\, \sigma^2 \, (\sum_{j=0}^{\infty} w_j) + O(n^{-1}), \qquad i = 1,2,\ldots,\tilde{p},$$

and for $p > q$,

$$E[\sum_{t=p+1}^{n} Z_{t-s}\, v_t] = O(n^{-1}), \qquad s = 1,2,\ldots,p-q. \qquad \square$$

Note that

$$-\, r\, \sigma^2 \sum_{j=0}^{\infty} w_j = -\, r\, \sigma^2 \, (1 - \sum_{j=1}^{p} \alpha_j)^{-1}.$$

If the trend function is a (q-1)-degree polynomial in t, the q-th difference of $\underset{\sim}{X}_t$ is $\underset{\sim}{0}$. The constants C_m satisfying (4.10) and (4.11) are

$$C_m = (-1)^m \binom{q}{m}, \qquad m = 0,1,\ldots,q.$$

Because $\Sigma_{m=0}^{q} C_m = 0$, by (4.16) we obtain

$$\sum_{j=1}^{p} \alpha_j = \sum_{i=1}^{\tilde{p}} \delta_i .$$

We can isolate the effect of estimating the mean function by transforming the problem. Let

$$\Delta^s W_t = \sum_{j=0}^{s} (-1)^j \binom{s}{j} W_{t-j}$$

denote the s-th difference of $\{W_t\}$. Then,

$$Z_t = \sum_{m=0}^{q} C_m Y_{t-m}$$

$$= \Delta^q Y_t$$

$$= \Delta^q W_t .$$

From (4.19),

$$W_t = \sum_{i=1}^{\tilde{p}} \delta_i W_{t-i} + 1_{(p>q)} \sum_{i=q+1}^{p} \delta_i Z_{t-i+q} + v_t .$$

Consider the following reparametrization of the above model.

$$W_t = \sum_{i=1}^{\tilde{p}} \theta_i \, \Delta^{i-1} W_{t-1} + 1_{(p>q)} \sum_{i=q+1}^{p} \theta_i \, Z_{t-i+q} + v_t. \tag{4.32}$$

The following theorem establishes the relation between $(\theta_1, \theta_2, \ldots, \theta_p)$ and $(\delta_1, \delta_2, \ldots, \delta_p)$.

Theorem 4.5. Let $\underset{\sim}{\delta} = (\delta_1, \delta_2, \ldots, \delta_p)$ and $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$ be as defined in (4.19) and (4.32). Then

$$\delta_i = \sum_{\ell=i-1}^{\tilde{p}-1} \binom{\ell}{i-1} (-1)^{i-1} \theta_{\ell+1}, \quad i = 1,2,\ldots,\tilde{p}$$

and if $p > q$, then $\delta_i = \theta_i$ for $i = q+1, q+2, \ldots, p$. Also,

$$\sum_{i=1}^{\tilde{p}} \delta_i = \theta_1.$$

Proof. From (4.32),

$$\begin{aligned} W_t &= \theta_1 W_{t-1} + \theta_2 \sum_{j=0}^{1} (-1)^j \binom{1}{j} W_{t-j-1} + \ldots + \theta_{\tilde{p}} \sum_{j=0}^{\tilde{p}-1} (-1)^j \binom{\tilde{p}-1}{j} W_{t-j-1} \\ &\quad + 1_{(p>q)} \sum_{i=q+1}^{p} \theta_i \, Z_{t-i-q} + v_t \\ &= W_{t-1}\left\{\sum_{\ell=0}^{\tilde{p}-1} \binom{\ell}{0} (-1)^0 \theta_{\ell+1}\right\} + W_{t-2}\left\{\sum_{\ell=1}^{\tilde{p}-1} \binom{\ell}{1} (-1)^1 \theta_{\ell+1}\right\} \\ &\quad + \ldots + W_{t-\tilde{p}} \binom{\tilde{p}-1}{\tilde{p}-1} (-1)^{\tilde{p}-1} \theta_{\tilde{p}} \\ &\quad + 1_{(p>q)} \sum_{i=q+1}^{p} \theta_i \, Z_{t-i+q} + v_t. \end{aligned} \tag{4.33}$$

Comparing (4.33) with (4.19), we get

$$\delta_i = \sum_{\ell=i-1}^{\tilde{p}-1} \binom{\ell}{i-1} (-1)^{i-1} \theta_{\ell+1}$$

for $i = 1, 2, \ldots, \tilde{p}$ and if $p > q$, then $\delta_i = \theta_i$ for $i = q+1, q+2, \ldots, p$. Now

$$\begin{aligned}
\sum_{i=1}^{\tilde{p}} \delta_i &= \sum_{i=1}^{\tilde{p}} \sum_{\ell=i-1}^{\tilde{p}-1} \binom{\ell}{i-1}(-1)^{i-1} \theta_{\ell+1} \\
&= \sum_{\ell=0}^{\tilde{p}-1} \theta_{\ell+1} \sum_{i=1}^{\ell+1} \binom{\ell}{i-1} (-1)^{i-1} \\
&= \theta_1 + \sum_{\ell=1}^{\tilde{p}-1} \theta_{\ell+1} \sum_{j=0}^{\ell} \binom{\ell}{j} (-1)^j \\
&= \theta_1. \qquad \square
\end{aligned}$$

The simple least squares estimator of $\underset{\sim}{\theta}$ is given by

$$\underset{\sim}{\theta}^* = [(n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{G}_t' \underset{\sim}{G}_t]^{-1} [(n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{G}_t' W_t] \tag{4.34}$$

and the error in the estimator,

$$\underset{\sim}{\theta}^* - \underset{\sim}{\theta} = [(n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{G}_t' \underset{\sim}{G}_t]^{-1} [(n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{G}_t' v_t] \tag{4.35}$$

where

$$\underset{\sim}{G}_t = (W_{t-1}, \Delta W_{t-1}, \ldots, \Delta^{\tilde{p}-1} W_{t-1}, Z_{t-1}, \ldots, Z_{t-p+q}), \quad \text{if} \quad p > q$$

$$= (W_{t-1}, \Delta W_{t-1}, \ldots, \Delta^{\tilde{p}-1} W_{t-1}), \qquad \text{if} \quad p < q.$$

Using Theorem 4.4, approximate expressions for the elements of $E[\ \Sigma_{t=p+1}^{n} \underset{\sim}{G}_t' v_t]$ are obtained in Theorem 4.6.

<u>Theorem 4.6.</u> Let $\{Y_t\}$ be a stochastic process satisfying the conditions of Theorem 4.4. Then

$$E[\sum_{t=p+1}^{n} \underset{\sim}{G}_t' v_t] = (-r\,\sigma^2 (1-\theta_1)^{-1}, 0, 0, \ldots, 0)' + O(n^{-1}).$$

<u>Proof.</u> From Theorems 4.4 and 4.5,

$$E[\sum_{t=p+1}^{n} W_{t-1} v_t] = -r\,\sigma^2 (1-\theta_1)^{-1} + O(n^{-1}),$$

$$E[\sum_{t=p+1}^{n} (\Delta^{\ell} W_{t-1}) v_t] = \Delta^{\ell}\{\sum_{t=p+1}^{n} E[W_{t-1} v_t]\}$$

$$= O(n^{-1}) \qquad \text{for} \quad \ell = 1,2,\ldots,\tilde{p}-1,$$

and if $p > q$,

$$E[\sum_{t=p+1}^{n} Z_{t-s} v_t] = O(n^{-1}), \qquad s = 1,2,\ldots,p-q.$$

□

For the choice of $\underset{\sim}{X}_t$ in (4.28), the effect of estimating the mean function could be isolated by transforming the problem. From Lemma 3.1, we know that $\theta_1 = 1$ if and only if there exists a unit root for the characteristic equation

$$m^p - \alpha_1 m^{p-1} - \dots - \alpha_p = 0.$$

Using the above results, the following method of adjusting for bias in the ordinary least squares estimator is proposed.

Step (i). Regress Y_t on $\underset{\sim}{X}_t = (t^{i(1)}, t^{i(2)}, \dots, t^{i(r)})$ to get the ordinary least squares estimate $\hat{\underset{\sim}{\beta}}$ of $\underset{\sim}{\beta}$. Define

$$W_t = Y_t - \underset{\sim}{X}_t \hat{\underset{\sim}{\beta}},$$

and

$$Z_t = \Delta^q W_t,$$

where $q = i(r) + 1$.

Step (ii). Regress W_t on $\underset{\sim}{G}_t$ to get the ordinary least squares estimate of $\underset{\sim}{\theta}$,

$$\underset{\sim}{\theta}^* = \hat{\underset{\sim}{D}}^{-1} \hat{\underset{\sim}{\phi}} \tag{4.36}$$

where

$$\hat{\underset{\sim}{D}} = (n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{G}' \underset{\sim}{G}_t ,$$

$$\hat{\underset{\sim}{\phi}} = (n-p)^{-1} \sum_{t=p+1}^{n} \underset{\sim}{G}_t' W_t ,$$

and

$$\begin{aligned} \underset{\sim}{G}_t &= (W_{t-1}, \Delta W_{t-1}, \ldots, \Delta^{\tilde{p}-1} W_{t-1}, Z_{t-1}, \ldots, Z_{t-p+q}), && \text{if } p > q \\ &= (W_{t-1}, \Delta W_{t-1}, \ldots, \Delta^{\tilde{p}-1} W_{t-1}), && \text{if } p < q. \end{aligned}$$

Obtain the modified least squares estimator,

$$\begin{aligned} \hat{\underset{\sim}{\theta}} &= \hat{\underset{\sim}{D}}^{-1} \hat{\underset{\sim}{\phi}} &&, \text{ if } \theta_1^* < 1 \\ &= \hat{\underset{\sim}{D}}^{-1} [\hat{\underset{\sim}{\theta}} + (D'')^{-1} (1 - \theta_1^*, 0, 0, \ldots, 0)'], && \text{if } \theta_1^* > 1, \end{aligned} \tag{4.37}$$

where D'' is the upper left element of $\hat{\underset{\sim}{D}}^{-1}$.

<u>Step (iii).</u> Use the mean square error of the regression in Step (ii) as an estimator $\hat{\sigma}^2$ of σ^2.

<u>Step (iv).</u> Obtain the adjusted estimator

$$\tilde{\underset{\sim}{\theta}} = \hat{\underset{\sim}{D}}^{-1} [\hat{\underset{\sim}{\phi}} + \hat{\underset{\sim}{h}}], \tag{4.38}$$

where

$$\hat{\underset{\sim}{h}} = (\hat{h}_1, 0, 0, \ldots, 0)',$$

and

$$\begin{aligned} \hat{h}_1 &= (D'')^{-1}(1 - \theta_1^*) \qquad \text{if} \quad \theta_1^* > 1 \quad \text{or} \\ &\qquad\qquad\qquad\qquad \text{if} \quad [(n-p)(1 - \hat{\theta}_1)]^{-1} r\hat{\sigma}^2 D'' > (1 - \hat{\theta}_1) \\ &= [(n-p)(1 - \hat{\theta}_1)]^{-1} r\sigma^2 \qquad \text{otherwise.} \end{aligned}$$

<u>Step (v).</u> The estimate of $\underset{\sim}{\alpha}$ is obtained using Theorem 4.5, (4.14), and (4.15).

To extend the bias adjustment method to the case in which $\theta_1 \varepsilon[-1, 1]$, several possibilities beyond that of (4.38) exist. The following method which uses the $\hat{\tau}_\tau$ statistic,

$$\hat{\tau}_\tau = [(n-p)^{-1} D'' \hat{\sigma}^2]^{-1/2} (\theta_1^* - 1),$$

is suggested. Let

$$\tilde{\underset{\sim}{\theta}}(\tau_\tau) = \hat{\underset{\sim}{D}}^{-1} (\hat{\underset{\sim}{\phi}} + \hat{\underset{\sim}{f}}) \tag{4.39}$$

where

$$\hat{\underset{\sim}{f}} = (\hat{f}_1, 0, 0, \ldots, 0)',$$

$$\hat{f}_1 = a \qquad , \text{ if } \hat{\tau}_\tau < \hat{\tau}_{\tau(01)}$$

$$= a + b[\hat{\tau}_{\tau(50)} - \hat{\tau}_{\tau(01)}]^{-3}[\hat{\tau}_\tau - \hat{\tau}_{\tau(01)}]^3,$$

$$\text{if } \hat{\tau}_{\tau(01)} < \hat{\tau}_\tau < \hat{\tau}_{\tau(50)}$$

$$= a + b \qquad , \text{ if } \hat{\tau}_\tau > \hat{\tau}_{\tau(50)}$$

$$a = [(n-p)(1 - \theta_1^*)]^{-1} r\sigma^2,$$

and

$$b = (D")^{-1} (1 - \theta_1^*) - a.$$

This method of adjusting for bias arising from estimating the mean function extends immediately to the purely seasonal p-th order autoregressive process, with unknown seasonal mean functions. The procedure parallels the case of unknown seasonal means described in Chapter III and therefore will not be repeated.

CHAPTER V. A MONTE CARLO STUDY

Approximate expressions for the biases of the least squares estimators due to estimating the mean function have been derived. The magnitude of the biases can be substantial for the moderately small samples encountered in practice. It is important to empirically investigate the accuracy of the approximate expressions for the bias in small samples. Modified least squares estimators with corrections for the bias are compared with the least squares estimator in a Monte Carlo study to determine the practical value of adjusting for the bias.

Normal random variables are generated using the GGNML subroutine of the IMSL package. All of the computations are performed using double precision arithmetic. Standard normal error processes are used throughout the study. A second-order autoregressive process with constant mean, and a second-order autoregressive process with mean function linear in time are considered.

Mean Model

The second-order autoregressive model with unknown mean has the form

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + e_t \tag{5.1}$$

$$= \mu + \alpha_1(Y_{t-1} - \mu) + \alpha_2(Y_{t-2} - \mu) + e_t, \tag{5.2}$$

where

$$\alpha_0 = \mu(1 - \alpha_1 - \alpha_2),$$

and $\{e_t\}$ is a sequence of normal (0, 1) variables. The intercept α_0 is set equal to zero in the Monte Carlo study. There is no loss of generality in this choice for the stationary process. For the stationary processes, the initial observations are generated by

$$Y_1 = \{\gamma(0)\}^{1/2} e_1, \tag{5.3}$$

$$Y_2 = \{\gamma(0)\}^{-1/2} \gamma(1) e_1 + [\gamma(0) - \{\gamma(1)\}^2 \{\gamma(0)\}^{-1}]^{1/2} e_2,$$

where

$$\gamma(0) = \frac{1 - \alpha_2}{(1+\alpha_2)(1-\alpha_1-\alpha_2)(1+\alpha_1-\alpha_2)},$$

and

$$\gamma(1) = \frac{\alpha_1}{(1+\alpha_2)(1-\alpha_1-\alpha_2)(1+\alpha_1-\alpha_2)}.$$

For the nonstationary processes, Y_i is set equal to e_i for $i=1,2$. Series lengths of 25 and 50 observations are used in the study. This means that 23 and 48 observations are used in the least squares regressions. For each sample size, 15 values of (α_1, α_2), such that the roots of the associated characteristic equation range from -0.8 to 1.0, are used. The values of (α_1, α_2) and the roots of the corresponding equation that are used in the study are given in Table 5.1.

Table 5.1 The roots of the characteristic equation, the parameter values, and the parameters of the reparametrized model

m_1	m_2	α_1	α_2	δ_1	δ_2
1.0	0.8	1.8	-0.80	1.00	0.80
1.0	0.5	1.5	-0.50	1.00	0.50
1.0	0.0	1.0	0.00	1.00	0.00
1.0	-0.5	0.5	0.50	1.00	-0.50
1.0	-0.8	0.2	0.80	1.00	-0.80
0.8	0.5	1.3	-0.40	0.90	0.40
0.8	0.1	0.9	-0.08	0.82	0.08
0.8	-0.5	0.3	0.40	0.70	-0.40
0.8	-0.8	0.0	0.64	0.64	-0.64
0.5	0.0	0.5	0.00	0.50	0.00
0.5	-0.5	0.0	0.25	0.25	-0.25
0.5	-0.8	-0.3	0.40	0.10	-0.40
0.1	-0.5	-0.4	0.05	0.35	-0.05
0.1	-0.8	-0.7	0.08	-0.62	-0.08
-0.5	-0.8	-1.3	-0.40	-1.70	0.40

For each (α_1, α_2, n) combination, various point estimates are computed using the same set of observations. This is repeated for 1,000 sets of observations. Sample biases and mean square errors for each estimator are obtained by averaging over the 1,000 replications. The numerical results are reported in the following tables.

Four estimators are included in the study. They are,

(i) the modified least squares estimator $\hat{\underset{\sim}{\alpha}}$ defined in (3.23),

(ii) the estimator $\hat{\underset{\sim}{\alpha}}(S)$, based on the work of Salem (1971) defined in (3.4),

(iii) the estimator $\tilde{\underset{\sim}{\alpha}}$, defined in (3.24),

and

(iv) the estimator $\tilde{\underset{\sim}{\alpha}}(\tau_\mu)$ defined in (3.25).

Consider the reparametrization of the model (5.2),

$$Y_t = \mu + \delta_1(Y_{t-1} - \mu) + \delta_2(Y_{t-1} - Y_{t-2}) + e_t \tag{5.4}$$

where

$$\delta_1 = \alpha_1 + \alpha_2,$$

and

$$\delta_2 = -\alpha_2.$$

The least squares estimator $\underset{\sim}{\delta}^* = (\delta^*_1, \delta^*_2)'$ of $\underset{\sim}{\delta} = (\delta_1, \delta_2)'$ is

$$\underset{\sim}{\delta}^* = \hat{\underset{\sim}{G}}^{-1} \underset{\sim}{z}, \tag{5.5}$$

where

$$\hat{\underset{\sim}{G}} = (n-2)^{-1} \begin{bmatrix} \sum_{t=3}^{n} (Y_{t-1} - \overline{Y})(Y_{t-1}-\overline{Y}) & \sum_{t=3}^{n} (Y_{t-1}-\overline{Y})(Y_{t-2}-\overline{Y}) \\ \sum_{t=3}^{n} (Y_{t-1}-\overline{Y})(Y_{t-2}-\overline{Y}) & \sum_{t=3}^{n} (Y_{t-2} -\overline{Y})(Y_{t-2} -\overline{Y}) \end{bmatrix},$$

$$\underset{\sim}{z} = (n-2)^{-1} \begin{bmatrix} \sum_{t=3}^{n} (Y_t -\overline{Y})(Y_{t-1} -\overline{Y}) \\ \sum_{t=3}^{n} (Y_t - \overline{Y})(Y_{t-2} - \overline{Y}) \end{bmatrix},$$

and

$$\overline{Y} = n^{-1} \sum_{t=1}^{n} Y_t .$$

The estimators $\hat{\underset{\sim}{\delta}}$, $\tilde{\underset{\sim}{\delta}}$, and $\tilde{\underset{\sim}{\delta}}(\tau_\mu)$ are constructed using (3.23), (3.24), and (3.25), respectively. The corresponding estimators of $\underset{\sim}{\alpha}$ are obtained using

$$\underset{\sim}{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}. \tag{5.6}$$

Tables 5.2 and 5.3 contain the empirical bias of various estimators of α_1 for n = 25 and 50, respectively. For $m_1 = 1$, the modified least squares estimator $\hat{\alpha}_1$ has the largest absolute bias and the

Table 5.2 Empirical bias of various estimators of α_1 for n = 25

m_1	m_2	$\hat{\alpha}_1$	$\hat{\alpha}_1(S)$	$\tilde{\alpha}_1$	$\tilde{\alpha}_1(\tau_\mu)$
1.0	0.8	-0.223	-0.218	-0.217	-0.207
1.0	0.5	-0.196	-0.174	-0.178	-0.159
1.0	0.0	-0.162	-0.122	-0.127	-0.097
1.0	-0.5	-0.148	-0.089	-0.097	-0.057
1.0	-0.8	-0.133	-0.063	-0.074	-0.025
0.8	0.5	-0.118	-0.091	-0.096	-0.065
0.8	0.1	-0.109	-0.071	-0.067	-0.026
0.8	-0.5	-0.096	-0.041	-0.033	-0.029
0.8	-0.8	-0.086	-0.023	-0.014	0.053
0.5	0.0	-0.066	-0.026	-0.019	0.019
0.5	-0.5	-0.052	-0.003	0.005	0.044
0.5	-0.8	-0.058	-0.003	0.006	0.047
0.1	-0.5	-0.027	0.014	0.021	0.029
0.1	-0.8	-0.011	0.032	0.039	0.044
-0.5	-0.8	0.030	0.056	0.060	0.060

Table 5.3 Empirical bias of various estimators of α_1 for n = 50

m_1	m_2	$\hat{\alpha}_1$	$\hat{\alpha}_1(S)$	$\tilde{\alpha}_1$	$\tilde{\alpha}_1(\tau_\mu)$
1.0	0.8	-0.105	-0.103	-0.101	-0.097
1.0	0.5	-0.098	-0.088	-0.088	-0.080
1.0	0.0	-0.084	-0.064	-0.067	-0.052
1.0	-0.5	-0.076	-0.046	-0.051	-0.031
1.0	-0.8	-0.066	-0.032	-0.038	-0.014
0.8	0.5	-0.046	-0.033	-0.032	-0.021
0.8	0.1	-0.036	-0.017	-0.016	-0.002
0.8	-0.5	-0.052	-0.024	-0.022	-0.002
0.8	-0.8	-0.044	-0.012	-0.010	0.015
0.5	0.0	-0.028	-0.008	-0.006	-0.004
0.5	-0.5	-0.025	-0.000	0.001	0.002
0.5	-0.8	-0.021	0.007	0.009	0.010
0.1	-0.5	-0.009	0.012	0.013	0.013
0.1	-0.8	-0.014	0.008	0.009	0.009
-0.5	-0.8	0.016	0.028	0.029	0.029

modified estimator $\tilde{\alpha}_1(\tau_\mu)$ has the smallest absolute bias. For $m_1 = 1$, the bias of $\hat{\alpha}_1(S)$ is close to that of $\tilde{\alpha}_1$. When both roots are negative, $\hat{\alpha}_1$ has the smallest bias. When at least one of the roots is positive, the least squares estimator $\hat{\alpha}_1$ underestimates α_1. For $m_1 = 0.8$ and m_2 positive, $\tilde{\alpha}_1(\tau_\mu)$ has the smallest absolute bias. Only small differences between the biases of $\hat{\alpha}_1(S)$ and $\tilde{\alpha}_1$ are observed.

Tables 5.4 and 5.5 contain the empirical mean square errors of various estimators of α_1 for $n = 25$ and 50, respectively. For $m_1 = 1$, $\hat{\alpha}_1$ has the largest mean square error and $\tilde{\alpha}_1(\tau_\mu)$ has the smallest mean square error. For $m_1 = 0.8$ and m_2 positive, $\tilde{\alpha}_1(\tau_\mu)$ has the smallest mean square error. There are only small differences between the mean square errors of $\hat{\alpha}_1(S)$ and $\tilde{\alpha}_1$. The ordering of the estimators for α_1 based on the absolute bias and on the mean square error coincide. This is because the variance of the estimators is small compared to the bias for most values of m_1 and m_2. From the point of view of statistical decision theory, Chernoff and Moses (1959, pp. 119-165) conclude that the average risk is the best available criterion for evaluating the relative performances of various estimators. For a uniform weight function, the average risk is the mean of the mean square errors averaged over the values of the parameters considered. For $n = 25$, the average mean square errors of $\hat{\alpha}_1$, $\hat{\alpha}_1(S)$, $\tilde{\alpha}_1$ and $\tilde{\alpha}_1(\tau_\mu)$ are 0.063, 0.057, 0.057 and 0.057, respectively. For $n = 50$, the average mean square errors are 0.023, 0.022, 0.022 and 0.021, respectively.

Tables 5.6 and 5.7 contain the empirical biases of various

Table 5.4 Empirical mean square error multiplied by ten of various estimators of α_1 for n = 25

m_1	m_2	$\hat{\alpha}_1$	$\hat{\alpha}_1(S)$	$\tilde{\alpha}_1$	$\tilde{\alpha}_1(\tau_\mu)$
1.0	0.8	0.951	0.905	0.886	0.828
1.0	0.5	0.883	0.768	0.778	0.696
1.0	0.0	0.820	0.675	0.687	0.598
1.0	-0.5	0.784	0.632	0.629	0.557
1.0	-0.8	0.606	0.472	0.463	0.391
0.8	0.5	0.627	0.542	0.540	0.483
0.8	0.1	0.640	0.546	0.542	0.501
0.8	-0.5	0.572	0.490	0.487	0.509
0.8	-0.8	0.540	0.480	0.481	0.558
0.5	0.0	0.534	0.485	0.483	0.516
0.5	-0.5	0.513	0.488	0.493	0.602
0.5	-0.8	0.504	0.489	0.498	0.657
0.1	-0.5	0.507	0.517	0.525	0.566
0.1	-0.8	0.493	0.528	0.540	0.580
-0.5	-0.8	0.416	0.469	0.477	0.478

Table 5.5 Empirical mean square error multiplied by ten, of various estimators of α_1 for n = 50

m_1	m_2	$\hat{\alpha}_1$	$\hat{\alpha}_1(S)$	$\tilde{\alpha}_1$	$\tilde{\alpha}_1(\tau_\mu)$
1.0	0.8	0.262	0.257	0.245	0.233
1.0	0.5	0.288	0.263	0.263	0.246
1.0	0.0	0.316	0.279	0.282	0.262
1.0	-0.5	0.255	0.214	0.216	0.196
1.0	-0.8	0.204	0.168	0.164	0.146
0.8	0.5	0.212	0.196	0.195	0.185
0.8	0.1	0.238	0.222	0.222	0.220
0.8	-0.5	0.236	0.213	0.214	0.230
0.8	-0.8	0.182	0.165	0.166	0.202
0.5	0.0	0.228	0.218	0.218	0.221
0.5	-0.5	0.229	0.223	0.224	0.228
0.5	-0.8	0.199	0.199	0.200	0.205
0.1	-0.5	0.219	0.223	0.223	0.223
0.1	-0.8	0.214	0.218	0.219	0.219
-0.5	-0.8	0.182	0.194	0.195	0.195

Table 5.6 Empirical bias of various estimators of α_2 for n = 25

m_1	m_2	$\hat{\alpha}_2$	$\hat{\alpha}_2(S)$	$\tilde{\alpha}_2$	$\tilde{\alpha}_2(\tau_\mu)$
1.0	0.8	0.153	0.186	0.169	0.176
1.0	0.5	0.069	0.102	0.092	0.105
1.0	0.0	-0.061	-0.015	-0.024	0.003
1.0	-0.5	-0.173	-0.110	-0.119	-0.081
1.0	-0.8	-0.230	-0.158	-0.169	-0.120
0.8	0.5	-0.002	0.025	0.027	0.051
0.8	0.1	-0.055	-0.016	-0.012	0.028
0.8	-0.5	-0.147	-0.090	-0.084	0.022
0.8	-0.8	-0.192	-0.126	-0.119	-0.051
0.5	0.0	-0.083	-0.041	-0.035	0.004
0.5	-0.5	-0.120	-0.069	-0.063	-0.023
0.5	-0.8	-0.139	-0.082	-0.075	-0.034
0.1	-0.5	-0.088	-0.045	-0.040	-0.032
0.1	-0.8	-0.084	-0.039	-0.034	-0.029
-0.5	-0.8	-0.009	0.018	0.021	0.021

Table 5.7 Empirical bias of various estimators of α_2 for n = 50

m_1	m_2	$\hat{\alpha}_2$	$\hat{\alpha}_2(S)$	$\tilde{\alpha}_2$	$\tilde{\alpha}_2(\tau_\mu)$
1.0	0.8	0.078	0.089	0.083	0.085
1.0	0.5	0.037	0.050	0.046	0.052
1.0	0.0	-0.025	-0.002	-0.007	0.005
1.0	-0.5	-0.081	-0.049	-0.055	-0.036
1.0	-0.8	-0.119	-0.082	-0.089	-0.066
0.8	0.5	-0.006	0.006	0.007	0.019
0.8	0.1	-0.045	-0.026	-0.025	-0.009
0.8	-0.5	-0.076	-0.048	-0.046	-0.025
0.8	-0.8	-0.086	-0.053	-0.051	-0.026
0.5	0.0	-0.045	-0.024	-0.024	-0.021
0.5	-0.5	-0.067	-0.042	-0.041	-0.039
0.5	-0.8	-0.065	-0.037	-0.036	-0.035
0.1	-0.5	-0.045	-0.023	-0.022	-0.022
0.1	-0.8	-0.052	-0.031	-0.030	-0.030
-0.5	-0.8	-0.004	0.009	0.010	0.010

estimators of α_2 for $n = 25$ and 50, respectively. Except for $m_1 = 1$ and m_2 positive, $\hat{\alpha}_2$ underestimates α_2. Also for $m_1 = 1$ and m_2 positive $\hat{\alpha}_2$ has the smallest absolute bias. For $m_1 = 1$ and m_2 nonpositive, $\breve{\alpha}_2(\tau_\mu)$ has the smallest absolute bias. For the parameter values in the stationary region, $\breve{\alpha}_2$ and $\breve{\alpha}_2(\tau_\mu)$ have smaller absolute biases than $\hat{\alpha}_2$ and $\hat{\alpha}_2(S)$, except for $(m_1 = 0.8, m_2 = 0.5)$. For $n = 50$, small differences between the biases of $\hat{\alpha}_2(S)$, $\breve{\alpha}_2$ and $\breve{\alpha}_2(\tau_\mu)$ are found. For the second order process the difference in the bias of $\hat{\alpha}_1$ and $\breve{\alpha}_1$ is the same as the difference in the bias of $\hat{\alpha}_2$ and $\breve{\alpha}_2$. The empirical values of these differences are very similar.

Tables 5.8 and 5.9 contain the empirical mean square error of various estimators of α_2 for $n = 25$ and 50, respectively. For $m_1 = 1$ and m_2 positive, the least squares estimator $\hat{\alpha}_2$ has the smallest mean square error. For the remaining values, $\breve{\alpha}_2(\tau_\mu)$ has the smallest mean square error, except for $(m_1 = 0.8, m_2 = 0.5)$ and $(m_1 = -0.8, m_2 = -0.5)$. There are small differences between the mean square errors of $\hat{\alpha}_2(S)$ and $\breve{\alpha}_2$ for the stationary cases. The estimator $\breve{\alpha}_2$ generally has smaller mean square error than that of $\hat{\alpha}_2(S)$. For $n = 25$, the average mean square errors of $\hat{\alpha}_2$, $\hat{\alpha}_2(S)$, $\breve{\alpha}_2$ and $\breve{\alpha}_2(\tau_\mu)$ are 0.057, 0.055, 0.052, and 0.053, respectively. For $n = 50$, the average mean square errors of $\hat{\alpha}_2$, $\hat{\alpha}_2(S)$, $\breve{\alpha}_2$ and $\breve{\alpha}_2(\tau_\mu)$ are 0.022, 0.022, 0.021, and 0.021, respectively.

Table 5.8 Empirical mean square error multiplied by ten of various estimators of α_2 for n = 25

m_1	m_2	$\hat{\alpha}_2$	$\hat{\alpha}_2(S)$	$\tilde{\alpha}_2$	$\tilde{\alpha}_2(\tau_\mu)$
1.0	0.8	0.642	0.838	0.671	0.682
1.0	0.5	0.467	0.589	0.499	0.523
1.0	0.0	0.470	0.488	0.440	0.432
1.0	-0.5	0.792	0.669	0.642	0.553
1.0	-0.8	0.953	0.715	0.697	0.541
0.8	0.5	0.391	0.436	0.419	0.451
0.8	0.1	0.468	0.482	0.468	0.504
0.8	-0.5	0.615	0.516	0.508	0.506
0.8	-0.8	0.776	0.605	0.592	0.551
0.5	0.0	0.463	0.448	0.454	0.553
0.5	-0.5	0.546	0.486	0.489	0.603
0.5	-0.8	0.660	0.576	0.576	0.712
0.1	-0.5	0.474	0.452	0.452	0.496
0.1	-0.8	0.489	0.472	0.475	0.512
-0.5	-0.8	0.372	0.409	0.412	0.412

Table 5.9 Empirical mean square error multiplied by ten of various estimators of α_2 for n = 50

m_1	m_2	$\hat{\alpha}_2$	$\hat{\alpha}_2(S)$	$\tilde{\alpha}_2$	$\tilde{\alpha}_2(\tau_\mu)$
1.0	0.8	0.204	0.239	0.205	0.205
1.0	0.5	0.192	0.216	0.199	0.204
1.0	0.0	0.220	0.228	0.214	0.214
1.0	-0.5	0.254	0.224	0.220	0.201
1.0	-0.8	0.280	0.214	0.216	0.181
0.8	0.5	0.173	0.181	0.180	0.194
0.8	0.1	0.210	0.205	0.206	0.222
0.8	-0.5	0.255	0.228	0.229	0.252
0.8	-0.8	0.232	0.193	0.193	0.217
0.5	0.0	0.206	0.200	0.200	0.209
0.5	-0.5	0.247	0.228	0.228	0.233
0.5	-0.8	0.239	0.218	0.218	0.222
0.1	-0.5	0.217	0.211	0.211	0.211
0.1	-0.8	0.217	0.207	0.207	0.207
-0.5	-0.8	0.174	0.182	0.182	0.182

Tables 5.10 and 5.11 contain the empirical bias of the various estimators of δ_1 for $n = 25$ and 50, respectively. Except when both roots of the characteristic equation are negative, $\tilde{\delta}_1(\tau_\mu)$ has the smallest absolute bias and $\hat{\delta}_1$ has the largest absolute bias. For $n = 25$, $\hat{\delta}_1$ and $\hat{\delta}_1(S) = \hat{\alpha}_1(S) + \hat{\alpha}_2(S)$, underestimate δ_1 except for ($m_1 = -0.5$, $m_2 = -0.8$). For $n = 50$, all four estimators considered underestimate δ_1, except for ($m_1 = -0.5$, $m_2 = -0.8$). Generally, $\tilde{\delta}_1$ has smaller absolute bias than $\hat{\delta}_1(S)$.

Tables 5.12 and 5.13 contain the empirical mean square errors of various estimators of δ_1 for $n = 25$ and 50, respectively. For $m_1 = 1$, $\tilde{\delta}_1(\tau_\mu)$ has the smallest mean square error and $\hat{\delta}_1$ has the largest mean square error. For $m_1 = 1$, $\tilde{\delta}_1$ has smaller mean square error than $\hat{\delta}_1(S)$. For the remaining values, only small differences between the mean square errors of $\hat{\delta}_1(S)$ and $\tilde{\delta}_1$ are found. Generally, for the stationary values of the parameters, $\hat{\delta}_1(S)$ and $\tilde{\delta}_1$ have smaller mean square error than $\hat{\delta}_1$ and $\tilde{\delta}_1(\tau_\mu)$. For $n = 25$, the average mean square error of $\hat{\delta}_1$, $\hat{\delta}_1(S)$, $\tilde{\delta}_1$ and $\tilde{\delta}_1(\tau_\mu)$ are 0.116, 0.095, 0.094, and 0.098, respectively. For $n = 50$, the average mean square error of $\hat{\delta}_1$, $\hat{\delta}_1(S)$, $\tilde{\delta}_1$, and $\tilde{\delta}_1(\tau_\mu)$ are 0.040, 0.035, 0.035, and 0.035, respectively.

Tables 5.14 and 5.15 contain the frequencies of various types of adjustments made in obtaining $\hat{\delta}_1$, $\tilde{\delta}_1$, and $\tilde{\delta}_1(\tau_\mu)$. The estimator $\hat{\delta}_1$ is obtained by setting $\hat{\delta}_1 = 1$ whenever δ_1^* is greater than one. The estimator $\tilde{\delta}_1(\tau_\mu)$ is set equal to one whenever the "t-statistic", $\hat{\tau}_\mu$ for testing $\delta_1 = 1$, is greater than or equal to $\hat{\tau}_{\mu(50)}$, where

Table 5.10 Empirical bias of various estimators of δ_1 for n = 25

m_1	m_2	$\hat{\delta}_1$	$\hat{\delta}_1(S)$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\mu)$
1.0	0.8	-0.070	-0.032	-0.048	-0.031
1.0	0.5	-0.127	-0.072	-0.086	-0.054
1.0	0.0	-0.223	-0.137	-0.151	-0.094
1.0	-0.5	-0.321	-0.199	-0.216	-0.138
1.0	-0.8	-0.363	-0.221	-0.243	-0.145
0.8	0.5	-0.120	-0.066	-0.064	-0.014
0.8	0.1	-0.164	-0.087	-0.079	0.002
0.8	-0.5	-0.243	-0.132	-0.117	-0.007
0.8	-0.8	-0.278	-0.149	-0.133	0.002
0.5	0.0	-0.149	-0.067	-0.054	-0.023
0.5	-0.5	-0.173	-0.072	-0.058	0.020
0.5	-0.8	-0.197	-0.085	-0.069	0.013
0.1	-0.5	-0.115	-0.030	-0.018	-0.003
0.1	-0.8	-0.095	-0.006	0.005	0.015
-0.5	-0.8	0.021	0.074	0.081	0.082

Table 5.11 Empirical bias of various estimators of δ_1 for n = 50

m_1	m_2	$\hat{\delta}_1$	$\hat{\delta}_1(S)$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\mu)$
1.0	0.8	-0.027	-0.014	-0.018	-0.012
1.0	0.5	-0.061	-0.038	-0.042	-0.028
1.0	0.0	-0.109	-0.066	-0.074	-0.048
1.0	-0.5	-0.157	-0.095	-0.106	-0.067
1.0	-0.8	-0.185	-0.114	-0.127	-0.079
0.8	0.5	-0.052	-0.027	-0.025	-0.002
0.8	0.1	-0.080	-0.043	-0.041	-0.011
0.8	-0.5	-0.128	-0.072	-0.068	-0.027
0.8	-0.8	-0.130	-0.065	-0.061	-0.011
0.5	0.0	-0.072	-0.032	-0.030	-0.025
0.5	-0.5	-0.092	-0.042	-0.039	-0.037
0.5	-0.8	-0.086	-0.030	-0.027	-0.025
0.1	-0.5	-0.054	-0.012	-0.010	-0.010
0.1	-0.8	-0.066	-0.023	-0.021	-0.021
-0.5	-0.8	0.012	0.037	0.038	0.038

Table 5.12 Empirical mean square error multiplied by ten of various estimators of δ_1 for n = 25

m_1	m_2	$\hat{\delta}_1$	$\hat{\delta}_1(S)$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\mu)$
1.0	0.8	0.105	0.083	0.069	0.050
1.0	0.5	0.282	0.195	0.178	0.118
1.0	0.0	0.795	0.520	0.498	0.325
1.0	-0.5	1.709	1.163	1.110	0.788
1.0	-0.8	2.123	1.389	1.335	0.878
0.8	0.5	0.314	0.227	0.216	0.195
0.8	0.1	0.609	0.443	0.426	0.427
0.8	-0.5	1.280	0.920	0.898	0.937
0.8	-0.8	1.800	1.340	1.311	1.378
0.5	0.0	0.796	0.666	0.681	0.950
0.5	-0.5	1.223	1.055	1.066	1.508
0.5	-0.8	1.736	1.543	1.554	2.144
0.1	-0.5	1.369	1.348	1.363	1.532
0.1	-0.8	1.588	1.628	1.652	1.806
-0.5	-0.8	1.468	1.648	1.670	1.682

Table 5.13 Empirical mean square error multiplied by ten of various estimators of δ_1 for n = 50

m_1	m_2	$\hat{\delta}_1$	$\hat{\delta}_1(S)$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\mu)$
1.0	0.8	0.013	0.009	0.009	0.006
1.0	0.5	0.063	0.042	0.041	0.030
1.0	0.0	0.199	0.131	0.130	0.094
1.0	-0.5	0.410	0.268	0.264	0.188
1.0	-0.8	0.573	0.371	0.367	0.258
0.8	0.5	0.076	0.058	0.059	0.071
0.8	0.1	0.187	0.145	0.147	0.177
0.8	-0.5	0.470	0.371	0.372	0.451
0.8	-0.8	0.544	0.433	0.436	0.554
0.5	0.0	0.298	0.265	0.264	0.288
0.5	-0.5	0.542	0.494	0.494	0.512
0.5	-0.8	0.638	0.597	0.598	0.616
0.1	-0.5	0.603	0.598	0.599	0.599
0.1	-0.8	0.719	0.708	0.709	0.709
-0.5	-0.8	0.675	0.716	0.718	0.718

Table 5.14 Number of replications for which various estimators of δ_1 are set equal to unity for n = 25

δ_1	$\hat{\delta}_1$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\mu)$	$\hat{\tau}_\mu < \hat{\tau}_{\mu(01)}$
1.00	127	285	504	13
1.00	85	240	490	10
1.00	61	251	481	6
1.00	51	275	531	8
1.00	46	273	513	10
0.90	14	109	275	25
0.82	4	70	260	17
0.70	3	46	203	19
0.64	2	63	222	16
0.50	0	6	54	93
0.25	0	2	19	123
0.10	0	0	17	164
-0.35	0	0	1	553
-0.62	0	0	1	675
-1.70	0	0	0	991

Table 5.15 Number of replications for which various estimators of δ_1 are set equal to unity for n = 50

δ_1	$\hat{\delta}_1$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\mu)$	$\hat{\tau}_\mu < \hat{\tau}_{\mu(01)}$
1.00	94	244	495	9
1.00	54	235	488	12
1.00	64	244	507	13
1.00	51	252	507	12
1.00	52	263	498	12
0.90	1	17	88	69
0.82	0	8	55	78
0.70	0	4	36	99
0.64	0	5	36	83
0.50	0	0	1	541
0.25	0	0	0	704
0.10	0	0	0	750
-0.35	0	0	0	999
-0.62	0	0	0	1000
-1.70	0	0	0	1000

$\hat{\tau}_{\mu(50)}$ is the median of the approximate distribution of the statistic $\hat{\tau}_{\mu}$. The estimators $\breve{\delta}_1$ and $\breve{\delta}_1(\tau_\mu)$ are the same when $\hat{\tau}_\mu$ is less than or equal to $\hat{\tau}_{\mu(01)}$, where $\hat{\tau}_{\mu(01)}$ is the lower 1-percentile of the statistic $\hat{\tau}_\mu$. For $\delta_1 = 1$, approximately 50% of the times $\breve{\delta}_1(\tau_\mu)$ is set equal to one and approximately 25% of the time $\breve{\delta}_1$ is set equal to one. For δ_1 less than or equal to 0.7, generally, $\hat{\delta}_1$ and $\breve{\delta}_1$ are less than one.

On the basis of this study, one can recommend the use of the estimator $(\breve{\alpha}_1, \breve{\alpha}_2)$. The estimator is relatively easy to construct and is less biased for parameter sets judged common in economics. Also, the average mean square error of $(\breve{\alpha}_1, \breve{\alpha}_2)$ for parameter sets with a positive root is 5 to 10 percent below that of the least squares estimator.

Time Trend Model

The second-order autoregressive model with the mean function linear in time, has the form

$$Y_t = \beta_0' + \beta_1' t + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + e_t \tag{5.7}$$

$$= \beta_0 + \beta_1 t + \alpha_1\{Y_{t-1} - \beta_0 - \beta_1(t-1)\}$$

$$+ \alpha_2\{Y_{t-2} - \beta_0 - \beta_1(t-2)\} + e_t \tag{5.8}$$

where

$$\beta_0' = \beta_0(1 - \alpha_1 - \alpha_2) + \beta_1(\alpha_1 + 2\alpha_2),$$

$$\beta_1' = \beta_1(1 - \alpha_1 - \alpha_2),$$

and $\{e_t\}$ is a sequence of independent normal (0, 1) variables. In this study, the coefficients β_0 and β_1 are set equal to zero. For processes with the roots of the characteristic equation less than unity in absolute value, the initial observations are generated by (5.3). If $\alpha_1 + \alpha_2$ is equal to one, then Y_i is set equal to e_i for $i = 1,2$. Series lengths 25 and 50 observations are considered. For each sample size, the 15 values of (α_1, α_2) given in Table 5.1 are considered.

For each (α_1, α_2, n) combination, various point estimates are computed using the same set of observations. This is repeated for 1,000 sets of observations and the sample biases and mean square errors for each estimator are obtained by averaging over the replications.

Three estimators are included in the study. They are

(i) the modified least squares estimator $\hat{\underset{\sim}{\alpha}}$, defined in (4.37),

(ii) the estimator $\tilde{\underset{\sim}{\alpha}}$, defined in (4.38),

and

(iii) the estimator $\tilde{\underset{\sim}{\alpha}}(\tau_T)$ defined in (4.39).

Consider the reparametrization of the model (5.8),

$$Y_t = \beta_0 + \beta_1 t + \delta_1\{Y_{t-1} - \beta_0 - \beta_1(t-1)\} + \delta_2\{Y_{t-1} - Y_{t-2} - \beta_1\} + e_t \tag{5.9}$$

where

$$\delta_1 = \alpha_1 + \alpha_2 = \theta_1$$

and

$$\delta_2 = -\alpha_2 = \theta_2.$$

The least squares estimator $\underset{\sim}{\delta}^* = (\delta_1^*, \delta_2^*)'$ of $\underset{\sim}{\delta}$ is,

$$\underset{\sim}{\delta}^* = \hat{\underset{\sim}{H}}^{-1} \underset{\sim}{h}, \tag{5.10}$$

where

$$\hat{\underset{\sim}{H}} = (n-2)^{-1} \begin{bmatrix} \sum_{t=3}^{n} W_{t-1}^2 & \sum_{t=3}^{n} (W_{t-1} - W_{t-2}) W_{t-1} \\ \sum_{t=3}^{n} W_{t-1}(W_{t-1} - W_{t-2}) & \sum_{t=3}^{n} (W_{t-1} - W_{t-2})^2 \end{bmatrix}$$

$$W_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 t,$$

$$\underset{\sim}{h} = (n-2)^{-1} \begin{bmatrix} \sum_{t=3}^{n} W_t W_{t-1} \\ \\ \sum_{t=3}^{n} W_t (W_{t-1} - W_{t-2}) \end{bmatrix} ,$$

and

$$\begin{pmatrix} \hat{\beta}_0 \\ \\ \hat{\beta}_1 \end{pmatrix} = \begin{bmatrix} n & \sum_{t=1}^{n} t \\ \\ \sum_{t=1}^{n} t & \sum_{t=1}^{n} t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{n} Y_t \\ \\ \sum_{t=1}^{n} t\, Y_t \end{bmatrix} .$$

The adjusted estimators $\hat{\underset{\sim}{\delta}}$, $\tilde{\underset{\sim}{\delta}}$ and $\tilde{\underset{\sim}{\delta}}(\tau_\tau)$ are obtained using (4.37), (4.38), and (4.39), respectively and the corresponding estimators of $\underset{\sim}{\alpha}$ are obtained using (5.6).

Tables 5.16 and 5.17 contain the empirical bias of various estimators of α_1 for n = 25 and 50, respectively. The adjusted least squares estimator $\hat{\alpha}_1$ has the largest absolute bias, except for ($m_1 = -0.5$, $m_2 = -0.8$) and ($m_1 = 0.1$, $m_2 = -0.8$). For $m_1 = 1$, the estimator $\tilde{\alpha}_1(\tau_\tau)$ has smaller absolute bias than $\tilde{\alpha}_1$. For m_1 less than unity, $\tilde{\alpha}_1$ generally has smaller absolute bias than $\tilde{\alpha}_1(\tau_\tau)$.

Tables 5.18 and 5.19 contain the empirical mean square error of various estimators of α_1. For $m_1 = 1$, the estimator $\tilde{\alpha}_1(\tau_\tau)$ has the smallest mean square error. Also, for $m_1 = 0.8$ and m_2 positive, the estimator has the smallest mean square error. For m_1 less than or equal to 0.1, the modified least squares estimator $\hat{\alpha}_1$ has the smallest

Table 5.16 Empirical bias of various estimators of α_1 for $n = 25$

m_1	m_2	$\hat{\alpha}_1$	$\tilde{\alpha}_1$	$\tilde{\alpha}_1(\tau_\tau)$
1.0	0.8	-0.250	-0.211	-0.196
1.0	0.5	-0.251	-0.192	-0.169
1.0	0.0	-0.280	-0.187	-0.148
1.0	-0.5	-0.278	-0.157	-0.110
1.0	-0.8	-0.282	-0.140	-0.084
0.8	0.5	-0.174	-0.113	-0.087
0.8	0.1	-0.164	-0.081	-0.044
0.8	-0.5	-0.176	-0.055	-0.003
0.8	-0.8	-0.200	-0.059	0.005
0.5	0.0	-0.121	-0.029	0.008
0.5	-0.5	-0.122	-0.010	0.033
0.5	-0.8	-0.111	0.014	0.061
0.1	-0.5	-0.067	0.030	0.044
0.1	-0.8	-0.046	0.055	0.066
-0.5	-0.8	0.122	0.192	0.193

Table 5.17 Empirical bias of various estimators of α_1 for $n = 50$

m_1	m_2	$\hat{\alpha}_1$	$\tilde{\alpha}_1$	$\tilde{\alpha}_1(\tau_\tau)$
1.0	-0.8	-0.109	-0.096	-0.090
1.0	0.5	-0.120	-0.096	-0.085
1.0	0.0	-0.135	-0.094	-0.077
1.0	-0.5	-0.156	-0.096	-0.071
1.0	-0.8	-0.155	-0.085	-0.057
0.8	0.5	-0.072	-0.044	-0.033
0.8	0.1	-0.067	-0.028	-0.010
0.8	-0.5	-0.069	-0.011	0.014
0.8	-0.8	-0.087	-0.018	0.011
0.5	0.0	-0.049	-0.007	-0.000
0.5	-0.5	-0.057	-0.005	-0.000
0.5	-0.8	-0.055	0.004	0.009
0.1	-0.5	-0.032	0.012	0.013
0.1	-0.8	-0.027	0.020	0.020
-0.5	-0.8	0.023	0.050	0.050

Table 5.18 Empirical mean square error multiplied by ten of various estimators of α_1 for n = 25

m_1	m_2	$\hat{\alpha}_1$	$\tilde{\alpha}_1$	$\tilde{\alpha}_1(\tau_\tau)$
1.0	0.8	1.098	0.855	0.767
1.0	0.5	1.175	0.857	0.751
1.0	0.0	1.450	0.964	0.810
1.0	-0.5	1.443	0.897	0.758
1.0	-0.8	1.327	0.734	0.584
0.8	0.5	0.889	0.651	0.580
0.8	0.1	0.908	0.663	0.612
0.8	-0.5	0.891	0.610	0.597
0.8	-0.8	0.886	0.553	0.559
0.5	0.0	0.745	0.591	0.617
0.5	-0.5	0.683	0.558	0.653
0.5	-0.8	0.621	0.558	0.719
0.1	-0.5	0.529	0.539	0.610
0.1	-0.8	0.462	0.540	0.616
-0.5	-0.8	0.566	0.865	0.882

Table 5.19 Empirical mean square error multiplied by ten of various estimators of α_1 for n = 50

m_1	m_2	$\hat{\alpha}_1$	$\tilde{\alpha}_1$	$\tilde{\alpha}_1(\tau_\tau)$
1.0	0.8	0.277	0.234	0.217
1.0	0.5	0.371	0.302	0.277
1.0	0.0	0.463	0.350	0.313
1.0	-0.5	0.476	0.319	0.273
1.0	-0.8	0.431	0.260	0.219
0.8	0.5	0.257	0.213	0.200
0.8	0.1	0.281	0.231	0.226
0.8	-0.5	0.300	0.250	0.267
0.8	-0.8	0.274	0.209	0.236
0.5	0.0	0.274	0.244	0.250
0.5	-0.5	0.256	0.225	0.236
0.5	-0.8	0.245	0.225	0.242
0.1	-0.5	0.236	0.236	0.237
0.1	-0.8	0.228	0.240	0.240
-0.5	-0.8	0.177	0.211	0.211

mean square error. For the remaining values, $\tilde{\alpha}_1$ has the smallest mean square error. For n = 25, the average mean square errors of $\tilde{\alpha}_1$, $\tilde{\alpha}_1$ and $\tilde{\alpha}_1(\tau_\tau)$ are, 0.094, 0.071, and 0.068, respectively. For n = 50, the average mean square errors of $\hat{\alpha}_1$, $\tilde{\alpha}_1$, and $\tilde{\alpha}_1(\tau_\tau)$ are 0.031, 0.025, and 0.025, respectively.

Tables 5.20 and 5.21 contain the empirical bias of various estimators of α_2. For $m_1 = 1$, m_2 positive and for m_1 and m_2 both negative, the modified least squares estimator $\hat{\alpha}_2$ has the smallest absolute bias. For $m_1 = 1$ and m_2 nonpositive, $\tilde{\alpha}_2(\tau_\tau)$ has the smallest absolute bias. Generally, $\tilde{\alpha}_2(\tau_\tau)$ has smaller bias than $\tilde{\alpha}_2$.

Tables 5.22 and 5.23 contain the empirical mean square errors of the various estimators of α_2 for n = 25 and 50. For $m_1 = 1$, m_2 positive and for m_1 and m_2 both negative, $\hat{\alpha}_2$ has the smallest mean square error. For $m_1 = 1$ and m_2 nonpositive, $\tilde{\alpha}_2(\tau_\tau)$ has the smallest mean square error. For m_1 less than or equal to 0.5, $\tilde{\alpha}_2$ has smaller mean square error than $\tilde{\alpha}_2(\tau_\tau)$. For the remaining values, $\tilde{\alpha}_2(\tau_\tau)$ has the smallest mean square error. For n = 25, the average mean square errors of $\hat{\alpha}_2$, $\tilde{\alpha}_2$ and $\tilde{\alpha}_2(\tau_\tau)$ are 0.078, 0.058, and 0.058, respectively. For n = 50, the average mean square errors of $\hat{\alpha}_2$, $\tilde{\alpha}_2$, and $\tilde{\alpha}_2(\tau_\tau)$ are 0.028, 0.023, and 0.022, respectively.

Tables 5.24 and 5.25 contain the empirical bias of various estimators of δ_1 for n = 25 and 50, respectively. Except for ($m_1 = -0.5$, $m_2 = -0.8$), $\tilde{\delta}_1$ and $\tilde{\delta}_1(\tau_\tau)$ have smaller absolute bias than $\hat{\delta}_1$. Generally, $\tilde{\delta}_1(\tau_\tau)$ has the smallest absolute bias.

Table 5.20 Empirical bias of various estimators of α_2 for $n = 25$

m_1	m_2	$\hat{\alpha}_2$	$\tilde{\alpha}_2$	$\tilde{\alpha}_2(\tau_T)$
1.0	0.8	0.082	0.115	0.125
1.0	0.5	-0.015	0.033	0.051
1.0	0.0	-0.158	-0.077	-0.044
1.0	-0.5	-0.316	-0.196	-0.153
1.0	-0.8	-0.390	-0.248	-0.192
0.8	0.5	-0.062	-0.006	0.017
0.8	0.1	-0.149	-0.068	-0.033
0.8	-0.5	-0.251	-0.131	-0.078
0.8	-0.8	-0.283	-0.141	-0.076
0.5	0.0	-0.143	-0.052	-0.014
0.5	-0.5	-0.193	-0.081	-0.038
0.5	-0.8	-0.201	-0.074	-0.027
0.1	-0.5	-0.130	-0.033	-0.018
0.1	-0.8	-0.112	-0.010	0.001
-0.5	-0.8	0.068	0.138	0.139

Table 5.21 Empirical bias of various estimators of α_2 for $n = 50$

m_1	m_2	$\hat{\alpha}_2$	$\tilde{\alpha}_2$	$\tilde{\alpha}_2(\tau_T)$
1.0	0.8	0.046	0.058	0.061
1.0	0.5	-0.000	0.021	0.029
1.0	0.0	-0.075	-0.034	-0.018
1.0	-0.5	-0.155	-0.096	-0.072
1.0	-0.8	-0.206	-0.135	-0.107
0.8	0.5	-0.020	0.007	0.019
0.8	0.1	-0.058	-0.019	-0.002
0.8	-0.5	-0.119	-0.061	-0.037
0.8	-0.8	-0.135	-0.066	-0.037
0.5	0.0	-0.069	-0.026	-0.020
0.5	-0.5	-0.095	-0.042	-0.037
0.5	-0.8	-0.098	-0.038	-0.033
0.1	-0.5	-0.066	-0.020	-0.020
0.1	-0.8	-0.062	-0.015	-0.015
-0.5	-0.8	-0.001	0.026	0.026

Table 5.22 Empirical mean square error multiplied by ten of various estimators of α_2 for n = 25

m_1	m_2	$\hat{\alpha}_2$	$\tilde{\alpha}_2$	$\tilde{\alpha}_2(\tau_\tau)$
1.0	0.8	0.322	0.407	0.443
1.0	0.5	0.277	0.328	0.359
1.0	0.0	0.596	0.469	0.456
1.0	-0.5	1.423	0.890	0.746
1.0	-0.8	1.892	1.046	0.803
0.8	0.5	0.332	0.347	0.374
0.8	0.1	0.547	0.434	0.432
0.8	-0.5	1.033	0.658	0.601
0.8	-0.8	1.220	0.712	0.641
0.5	0.0	0.589	0.499	0.562
0.5	-0.5	0.776	0.563	0.643
0.5	-0.8	0.804	0.552	0.650
0.1	-0.5	0.598	0.537	0.620
0.1	-0.8	0.538	0.508	0.583
-0.5	-0.8	0.477	0.717	0.735

Table 5.23 Empirical mean square error multiplied by ten of various estimators of α_2 for n = 50

m_1	m_2	$\hat{\alpha}_2$	$\tilde{\alpha}_2$	$\tilde{\alpha}_2(\tau_\tau)$
1.0	0.8	0.128	0.141	0.147
1.0	0.5	0.148	0.159	0.168
1.0	0.0	0.239	0.205	0.202
1.0	-0.5	0.429	0.293	0.256
1.0	-0.8	0.597	0.370	0.305
0.8	0.5	0.152	0.162	0.173
0.8	0.1	0.219	0.206	0.218
0.8	-0.5	0.348	0.264	0.271
0.8	-0.8	0.360	0.241	0.248
0.5	0.0	0.241	0.218	0.233
0.5	-0.5	0.290	0.237	0.251
0.5	-0.8	0.295	0.234	0.249
0.1	-0.5	0.257	0.238	0.238
0.1	-0.8	0.246	0.230	0.231
-0.5	-0.8	0.168	0.190	0.190

Table 5.24 Empirical bias in various estimators of δ_1 for n = 25

m_1	m_2	$\hat{\delta}_1$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\tau)$
1.0	0.8	-0.169	-0.096	-0.071
1.0	0.5	-0.266	-0.159	-0.118
1.0	0.0	-0.438	-0.263	-0.192
1.0	-0.5	-0.594	-0.353	-0.263
1.0	-0.8	-0.672	-0.388	-0.276
0.8	0.5	-0.237	-0.119	-0.070
0.8	0.1	-0.313	-0.148	-0.077
0.8	-0.5	-0.427	-0.186	-0.081
0.8	-0.8	-0.483	-0.200	-0.071
0.5	0.0	-0.264	-0.081	-0.006
0.5	-0.5	-0.315	-0.091	-0.005
0.5	-0.8	-0.312	-0.060	0.034
0.1	-0.5	-0.197	-0.003	0.026
0.1	-0.8	-0.158	0.045	0.067
-0.5	-0.8	0.190	0.330	0.332

Table 5.25 Empirical bias of various estimators of δ_1 for n = 50

m_1	m_2	$\hat{\delta}_1$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\tau)$
1.0	0.8	-0.063	-0.038	-0.029
1.0	0.5	-0.120	-0.075	-0.056
1.0	0.0	-0.210	-0.128	-0.095
1.0	-0.5	-0.311	-0.192	-0.143
1.0	-0.8	-0.361	-0.220	-0.164
0.8	0.5	-0.092	-0.037	-0.014
0.8	0.1	-0.126	-0.047	-0.012
0.8	-0.5	-0.188	-0.072	-0.023
0.8	-0.8	-0.222	-0.084	-0.026
0.5	0.0	-0.118	-0.033	-0.020
0.5	-0.5	-0.152	-0.047	-0.037
0.5	-0.8	-0.153	-0.034	-0.024
0.1	-0.5	-0.098	-0.008	-0.007
0.1	-0.8	-0.089	0.005	0.005
-0.5	-0.8	0.022	0.076	0.076

Tables 5.26 and 5.27 contain the empirical mean square error of various estimators of δ_1 for n = 25 and 50. Except for (m_1 = -0.5, m_2 = -0.8), $\hat{\delta}_1$ has the largest mean square error. For m_1 less than or equal to 0.5, $\tilde{\delta}_1$ has smaller mean square error than $\tilde{\delta}_1(\tau_\tau)$. For the remaining values, $\tilde{\delta}_1(\tau_\tau)$ has the smallest mean square error. For n = 25, the average mean square errors of $\hat{\delta}_1$, $\tilde{\delta}_1$, and $\tilde{\delta}_1(\tau_\tau)$ are 0.229, 0.147, and 0.130, respectively. For n = 50, the average mean square errors of $\hat{\delta}_1$, $\tilde{\delta}_1$, and $\tilde{\delta}_1(\tau_\tau)$ are 0.068, 0.047, and 0.046, respectively.

Tables 5.28 and 5.29 contain the frequencies of various types of adjustments made in obtaining $\hat{\delta}_1$, $\tilde{\delta}_1$, and $\tilde{\delta}_1(\tau_\tau)$. The estimator $\hat{\delta}_1$ is obtained by setting $\hat{\delta}_1 = 1$ whenever δ_1^* is greater than one. The estimator $\tilde{\delta}_1(\tau_\tau)$ is set equal to one whenever the "t-statistic", $\hat{\tau}_\tau$ for testing $\delta_1 = 1$, is greater than or equal to $\hat{\tau}_{\tau(50)}$, where $\hat{\tau}_{\tau(50)}$ is the median of the approximate distribution of the statistic $\hat{\tau}_\tau$. The estimators $\tilde{\delta}_1$ and $\tilde{\delta}_1(\tau_\tau)$ are the same when $\hat{\tau}_\tau$ is less than or equal to $\hat{\tau}_{\tau(01)}$, where $\hat{\tau}_{\tau(01)}$ is the lower 1-percentile of the statistic $\hat{\tau}_\tau$. For $\delta_1 = 1$, approximately 50% of the times $\tilde{\delta}_1(\tau_\tau)$ is set equal to one and approximately 15% of the times $\tilde{\delta}_1$ is set equal to one. Except when $m_1 = 1$, $\hat{\delta}_1$ is always less than or equal to one.

On the basis of this study $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ can be recommended over the least squares estimator. The estimator $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ is less biased than

Table 5.26 Empirical mean square error multiplied by ten of various estimators of δ_1 for n = 25

m_1	m_2	$\hat{\delta}_1$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\tau)$
1.0	0.8	0.382	0.181	0.130
1.0	0.5	0.916	0.464	0.354
1.0	0.0	2.371	1.188	0.863
1.0	-0.5	4.446	2.277	1.699
1.0	-0.8	5.513	2.633	1.841
0.8	0.5	0.823	0.415	0.340
0.8	0.1	1.488	0.783	0.676
0.8	-0.5	2.718	1.399	1.255
0.8	-0.8	3.492	1.803	1.670
0.5	0.0	1.404	0.914	1.089
0.5	-0.5	2.048	1.368	1.716
0.5	-0.8	2.302	1.666	2.180
0.1	-0.5	1.655	1.548	1.856
0.1	-0.8	1.692	1.784	2.086
-0.5	-0.8	1.946	3.024	3.092

Table 5.27 Empirical mean square error multiplied by ten of various estimators of δ_1 for n = 50

m_1	m_2	$\hat{\delta}_1$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\tau)$
1.0	0.8	0.055	0.028	0.021
1.0	0.5	0.197	0.107	0.083
1.0	0.0	0.587	0.310	0.234
1.0	-0.5	1.234	0.646	0.481
1.0	-0.8	1.668	0.874	0.660
0.8	0.5	0.151	0.085	0.085
0.8	0.1	0.299	0.175	0.189
0.8	-0.5	0.701	0.434	0.482
0.8	-0.8	0.942	0.571	0.637
0.5	0.0	0.405	0.299	0.341
0.5	-0.5	0.682	0.517	0.566
0.5	-0.8	0.866	0.703	0.767
0.1	-0.5	0.720	0.683	0.685
0.1	-0.8	0.818	0.809	0.811
-0.5	-0.8	0.650	0.762	0.762

Table 5.28 Number of replications for which various estimators of δ_1 are set equal to unity for n = 25

δ_1	$\hat{\delta}_1$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\tau)$	$\hat{\tau}_\tau < \hat{\tau}_{\tau(01)}$
1.00	1	153	495	14
1.00	0	150	499	11
1.00	0	114	487	6
1.00	1	158	511	9
1.00	1	136	528	3
0.90	0	68	376	12
0.82	0	65	340	15
0.70	0	56	357	11
0.64	0	43	353	16
0.50	0	8	136	48
0.25	0	2	66	77
0.10	0	0	70	65
-0.35	0	0	4	335
-0.62	0	0	1	430
-1.70	0	0	0	921

Table 5.29 Number of replications for which various estimators of δ_1 are set equal to unity for n = 50

δ_1	$\hat{\delta}_1$	$\tilde{\delta}_1$	$\tilde{\delta}_1(\tau_\tau)$	$\hat{\tau}_\tau < \hat{\tau}_{\tau(01)}$
1.00	1	146	472	8
1.00	0	130	498	12
1.00	1	148	521	10
1.00	1	136	485	10
1.00	1	140	517	9
0.90	0	28	231	29
0.82	0	18	166	26
0.70	0	13	141	41
0.64	0	4	115	48
0.50	0	0	1	304
0.25	0	0	0	490
0.10	0	0	0	521
-0.35	0	0	0	974
-0.62	0	0	0	992
-1.70	0	0	0	1000

the least squares estimator and has mean square errors about 20 percent below those of the least squares estimator for parameters deemed realistic for economic applications.

CHAPTER VI. SUMMARY AND AN EXAMPLE

The estimation of the parameters of the autoregressive process is investigated. A p-th order autoregressive process defined by

$$Y_t = \underset{\sim}{X}_t \underset{\sim}{\beta} + \sum_{j=1}^{p} \alpha_j (Y_{t-j} - \underset{\sim}{X}_{t-j} \underset{\sim}{\beta}) + e_t, \tag{6.1}$$

where $\{e_t\}$ is a sequence of uncorrelated $(0,\sigma^2)$ random variables, is considered. Given a realization of n observations, the least squares estimators of the parameters are obtained by treating (6.1) as a regression equation. The asymptotic properties of the least squares estimators depend on

(i) the initial conditions $Y_0, Y_{-1}, \ldots, Y_{-p+1}$,

(ii) the roots of the characteristic equation

$$m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p = 0,$$

and

(iii) the properties of the $\{e_t\}$ sequence.

Assuming that the roots of the characteristic equation lie inside the unit circle, the limiting distribution of the least squares estimators are derived in Appendix B. The results are extensions of the results given in Fuller, Hasza and Goebel (1981). The basic difference is that Fuller, Hasza and Goebel assumed that $\{e_t\}$ is a sequence of independent random variables with bounded $(2+\delta)$-th moments, whereas we assume that $\{e_t\}$ is a sequence of martingale differences with constant conditional variance and bounded $(4+\delta)$-th moments.

Dickey (1976) obtained the limiting distribution of $n(\hat{\rho}-1)$, where

$$Y_t = \rho\, Y_{t-1} + e_t ,$$

$$\rho = 1,$$

$$\hat{\rho} = \frac{\sum_{t=2}^{n} Y_t\, Y_{t-1}}{\sum_{t=2}^{n} Y_{t-1}^2} ,$$

and $\{e_t\}$ is a sequence of independent $(0,\sigma^2)$ random variables with finite $(2+\delta)$-th moments. This result is extended in Chapter II to the case where $\{e_t\}$ is a sequence of martingale difference errors with constant conditional variance and bounded fourth moments.

For a p-th order stationary autoregressive process, various estimators are proposed that are asymptotically equivalent. Under a wide variety of assumptions the least squares estimator is consistent. But in small samples the least squares estimator is seriously biased. For the stationary p-th order model, approximate expressions for the bias arising from estimating the unknown mean are derived in Chapter III. Using the approximate expressions for the bias, two adjustments for the least squares estimator are proposed. A Monte Carlo study is conducted to study the small sample behavior of various estimators of a second-order autoregressive process with constant mean. Ordinary least squares, the method suggested by Salem (1971) and the two estimators suggested in Chapter III are compared. The absolute bias and the mean square errors of the estimator (3.14), proposed in Chapter III, are smaller than those of the least squares estimator for a wide range of parameter values. For the parameter values in the stationary region, only small differences in the mean square error are found between the estimator suggested by Salem (1971) and the estimator (3.14).

Also considered is the p-th order autoregressive process with a nonstationary mean function. The model is given by

$$Y_t = \underset{\sim}{X}_t \underset{\sim}{\beta} + \sum_{j=1}^{p} \alpha_j (Y_{t-j} - \underset{\sim}{X}_{t-j} \underset{\sim}{\beta}) + e_t .$$

Assume the roots of the characteristic equation lie inside the unit circle. The ordinary least squares estimator $\hat{\underset{\sim}{\beta}}$ of $\underset{\sim}{\beta}$ is obtained by regressing Y_t on $\underset{\sim}{X}_t$. The least squares estimator $\underset{\sim}{\alpha}^*$ of $\underset{\sim}{\alpha}$ is obtained by regressing $(Y_t - \underset{\sim}{X}_t \hat{\underset{\sim}{\beta}})$ on $(Y_{t-1} - \underset{\sim}{X}_{t-1} \hat{\underset{\sim}{\beta}})$, $(Y_{t-2} - \underset{\sim}{X}_{t-2} \hat{\underset{\sim}{\beta}})$, ..., $(Y_{t-p} - \underset{\sim}{X}_{t-p} \hat{\underset{\sim}{\beta}})$. Approximate expressions for the bias in $\underset{\sim}{\alpha}^*$, arising from estimating $\underset{\sim}{\beta}$, are derived. Particular attention is given to the case,

$$\underset{\sim}{X}_t = (t^{i(1)}, t^{i(2)}, \ldots, t^{i(r)}).$$

where $i(j)$ are integers. Using the approximate expressions for the bias in $\underset{\sim}{\alpha}^*$, arising from estimating $\underset{\sim}{\beta}$, two adjustments are proposed in Chapter IV.

A Monte Carlo study is conducted to study the small sample behavior of various estimators of a second order autoregressive process given by

$$Y_t = \beta_0 + \beta_1 t + \alpha_1\{Y_{t-1} - \beta_0 - \beta_1(t-1)\} + \alpha_2\{Y_{t-2} - \beta_0 - \beta_1(t-2)\} + e_t$$

$$= \beta_0' + \beta_1' t + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + e_t \; , \qquad (6.2)$$

where

$$\beta_0' = \beta_0(1 - \alpha_1 - \alpha_2) + \beta_1(\alpha_1 + 2\alpha_2) \; ,$$

$$\beta_1' = \beta_1(1 - \alpha_1 - \alpha_2) \; ,$$

and $\{e_t\}$ is a sequence of normal independent (0,1) variables. The least squares estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ of β_0 and β_1 are obtained by regressing Y_t on an intercept and time. The least squares estimates α_1^* and α_2^* of α_1 and α_2 are obtained by regressing $(Y_t - \hat{\beta}_0 - \hat{\beta}_1 t)$ on $\{Y_{t-1} - \hat{\beta}_0 - \hat{\beta}_1(t-1)\}$ and $\{Y_{t-2} - \hat{\beta}_0 - \hat{\beta}_1(t-2)\}$. The least squares estimator (4.37) and the two estimators (4.38) and (4.39), suggested in Chapter IV are compared. The Monte Carlo study demonstrates that the mean square errors of the estimators suggested in Chapter IV are smaller than those of the least squares estimator for a wide range of parameter values. Except when both roots of the characteristic equation are negative, the absolute biases in the estimators proposed in Chapter IV are smaller than those of the ordinary least squares estimator. Generally speaking, the adjusted estimators suggested in Chapter IV performed better than the least squares estimator.

An Example

Nelson (1973, p.100) lists 80 observations on seasonally adjusted U.S. gross national product. The data are quarterly data beginning with 1947-1. In modeling the data, we have taken logarithms and have considered autoregressive processes. Hasza (1977) concluded that the observations are generated by a stationary autoregressive process with a time trending mean and is of the form (6.2). Various estimators proposed in Chapter IV are computed.

(i) Regressing Y_t on time, Y_{t-1}, Y_{t-2} and a constant, we get

$$\begin{aligned} Y_t = \underset{(0.1941)}{0.6485} + \underset{(0.0005)}{0.0017t} + \underset{(0.0943)}{1.4475}\, Y_{t-1} - \underset{(0.0931)}{0.5651}\, Y_{t-2}\,, \\ s^2 = 1.4566 \times 10^{-4}\,, \end{aligned} \tag{6.3}$$

where the numbers in parentheses are the standard errors of the coefficients.

(ii) The simple least squares estimators of β_0 and β_1 are obtained by regressing Y_t on time and a constant. The least squares estimate of the mean function of Y_t is

$$\hat{E}\{Y_t\} = 5.4652 + 0.0142t \tag{6.4}$$

Let

$$W_t = Y_t - \hat{E}\{Y_t\} \quad .$$

The least squares estimator of $\underset{\sim}{\alpha} = (\alpha_1, \alpha_2)'$, is obtained by regressing W_t on W_{t-1} and W_{t-2} . The least squares estimate of $\underset{\sim}{\alpha}$ is

$$(\alpha_1^* \ , \ \alpha_2^*) = (1.4530, \ - \ 0.5701) \qquad (6.5)$$

and the least squares estimate of $\underset{\sim}{\delta}$ is

$$(\theta_1^* \ , \ \theta_2^*) = (\delta_1^* \ , \ \delta_2^*) = (0.8829, \ 0.5701)$$

This sequential fitting gives an estimate for σ^2 of

$$\hat{\sigma}^2 = 1.4628 \times 10^{-4} \quad . \qquad (6.6)$$

Combining (6.4) and (6.5) we obtain

$$\begin{aligned} Y_t &= \underset{(0.1941)}{0.6444} + \underset{(0.0005)}{0.0017t} + \underset{(0.0943)}{1.4530}\, Y_{t-1} - \underset{(0.0931)}{0.5701}\, Y_{t-2} \\ &= \underset{(0.1941)}{0.6444} + \underset{(0.0005)}{0.0017t} + \underset{(0.0356)}{0.8829}\, Y_{t-1} + \underset{(0.0931)}{0.5701}\, (Y_{t-1} - Y_{t-2}), \end{aligned}$$

where the standard errors are taken from (6.3).

(iii) Since $\theta_1^* < 1$, the estimator $\tilde{\underset{\sim}{\theta}}$ defined in (4.38) is obtained by adding $\hat{\underset{\sim}{D}}^{-1}\hat{\underset{\sim}{h}}$ to $\underset{\sim}{\theta}^*$ where

$$\hat{\underset{\sim}{D}} = \begin{bmatrix} \sum_{t=3}^{80} W_{t-1}^2 & \sum_{t=3}^{80} W_{t-1}(W_{t-1} - W_{t-2}) \\ \sum_{t=3}^{80} W_{t-1}(W_{t-1} - W_{t-2}) & \sum_{t=3}^{80} (W_{t-1} - W_{t-2})^2 \end{bmatrix}$$

and

$$\hat{\underset{\sim}{h}}' = \{2[(78)\ (1 - \theta_1^*)]^{-1}\ \hat{\sigma}^2,\ 0\} .$$

The estimator

$$\tilde{\underset{\sim}{\theta}}' = (0.9046,\ 0.5611)$$

and

$$\tilde{\underset{\sim}{\alpha}}' = (1.4657,\ -0.5611) .$$

(iv) The estimator $\tilde{\underset{\sim}{\theta}}(\tau_\tau)$ defined in (4.39) is obtained using an adjustment based on the $\hat{\tau}_\tau$ statistic. For this sample,

$$\hat{\tau}_\tau = -\ 3.28 \quad ,$$

and the percentiles $\hat{\tau}_{\tau(01)}$ and $\hat{\tau}_{\tau(50)}$ are approximately -4.10 and -2.435, respectively. Therefore, $\tilde{\theta}(\tau_\tau)$ is obtained by adding $\hat{D}^{-1}\hat{f}$ to θ^*, where

$$\hat{f}' = [a + b\,(1.665)^{-3}\,(\hat{\tau}_\tau + 4.10)^3,\ 0]\ ,$$

$$a = 2[78(1 - \theta_1^*)]^{-1}\,\hat{\sigma}^2\ ,$$

$$b = (D^{11})^{-1}\,(1 - \theta_1^*) - a\ ,$$

and D^{11} is the (1,1)-th element of $\hat{D}^{-1}$. The value of the estimator $\tilde{\theta}(\tau_\tau)$ is

$$\tilde{\theta}(\tau_\tau) = (0.9158,\ 0.5564)'$$

and

$$\tilde{\alpha}(\tau_\tau) = (1.4722,\ -0.5564)'\ .$$

If the parameter α is known then the generalized least squares estimator $(\hat{\beta}_{0G}\ ,\ \hat{\beta}_{1G})$ of $(\beta_0\ ,\ \beta_1)$ is obtained by regressing V_t on A_t and B_t, where

$$V_1 = \{\gamma(0)\}^{-1/2}\,\sigma\,Y_1\ ,$$

$$
\begin{aligned}
V_2 &= d_{22}\, Y_2 - d_{11}\, Y_1 \quad , \\
V_t &= Y_t - \alpha_1\, Y_{t-1} - \alpha_2\, Y_{t-2} \quad , \qquad t = 3,4,\ldots, n, \qquad\qquad (6.7) \\
d_{11} &= \rho(1)\, d_{22} \; , \\
d_{22} &= [\{1 - \rho^2(1)\}\gamma(0)]^{-1/2}\sigma \; , \\
\rho(1) &= \{\gamma(0)\}^{-1}\, \gamma(1) \quad , \\
A_1 &= d_{11} \; , \\
A_2 &= d_{22} - d_{11} \; , \\
A_t &= 1 - \alpha_1 - \alpha_2 \; , \qquad\qquad t = 3,4,\ldots,n \; , \\
B_1 &= d_{11} \; , \\
B_2 &= 2\, d_{22} - d_{11} \; , \\
B_t &= t - \alpha_1(t-1) - \alpha_2(t-2) \; , \qquad t = 3,4,\ldots,n \; ,
\end{aligned}
$$

and, $\gamma(0)$ and $\gamma(1)$ are defined in (5.3). An estimated generalized least squares estimator of β is obtained by regressing $\hat{V}_t$ on $\hat{A}_t$ and $\hat{B}_t$ where $\hat{V}_t$, $\hat{A}_t$ and $\hat{B}_t$ are computed using the estimated values of $\underset{\sim}{\alpha}$ and σ^2 in (6.7).

The generalized least squares estimates of $\underset{\sim}{\beta}_0$ and the standard errors are,

$$\text{(i)} \quad \hat{\beta}_{0G} = \underset{(0.0256)}{5.4752} \; , \quad \hat{\beta}_{1G} = \underset{(0.0005)}{0.0141} \; ,$$

$$\text{(ii)} \quad \hat{\beta}_{0G} = \underset{(0.0257)}{5.4751} \; , \quad \hat{\beta}_{1G} = \underset{(0.0005)}{0.0141} \; ,$$

$$\text{(iii)} \quad \hat{\beta}_{0G} = \underset{(0.0321)}{5.4754} \; , \quad \hat{\beta}_{1G} = \underset{(0.0006)}{0.0141} \; ,$$

and

$$\text{(iv)} \quad \hat{\beta}_{0G} = \underset{(0.0367)}{5.4753} \; , \quad \hat{\beta}_{1G} = \underset{(0.0007)}{0.0141} \; ,$$

respectively.

REFERENCES

Amemiya, T., and W. A. Fuller. 1967. A comparative study of alternative estimators in a distributed lag model. Econometrica 35: 509-529.

Anderson, T. W. 1959. On asymptotic distributions of estimates of parameters in stochastic difference equations. Ann. Math. Statist. 30: 676-687.

Anderson, T. W., and H. Rubin. 1950. The asymptotic properties of estimates of the parameters in a single equation in a complete system of stochastic equations. Ann. Math. Statist. 21: 570-582.

Anderson, T. W., and J. B. Taylor. 1979. Strong consistency of least squares estimates in dynamic models. The Annals of Statistics 7: 484-489.

Ansley, C. F., and P. Newbold. 1980. Finite sample properties of estimators for autoregressive moving average models. J. of Econometrics 13: 159-183.

Barnard, G. A., G. M. Jenkins, and G. B. Winsten. 1962. Likelihood inference and time series. J. R. Statist. Soc., Ser. A, 125: 321-372.

Bora-Senta, E., and S. Kounias. 1980. Parameter estimation and order determination of autoregressive models. Pages 93-108 in O. D. Anderson, ed. In Analyzing Time Series. North Holland, Amsterdam.

Box, G. E. P., S. C. Hillmer, and G. C. Tiao. 1976. Analysis and modeling of seasonal time series. Department of Statistics, Univ. of Wisconsin, Technical Report No. 465.

Box, G. E. P., and G. M. Jenkins. 1976. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.

Burg, J. P. 1967. Maximum entropy spectral analysis. Paper presented at the 37th Annual International Meeting, Soc. of Explor. Geophy., Oklahoma City.

Burg, J. P. 1968. A new analysis technique for time series data. Paper presented at Advanced Study Institute on Signal Processing, NATO, Enschede, Netherlands.

Burg, J. P. 1975. Maximum Entropy Spectral Analysis. Unpublished Ph. D. thesis. Stanford Univ., Stanford, California.

Chernoff, H. and L. Moses. 1959. Elementary Decision Theory. Wiley, New York.

Chung, K. L. 1974. A Course In Probability Theory. Academic Press, New York.

Copas, J. B. 1966. Monte Carlo results for estimation in a stable Markov time series. J. R. Statist. Soc., Ser. B, 28: 110-116.

Crowder, M. J. 1980. On the asymptotic properties of least squares estimators in autoregression. The Annals of Statistics 8: 132-146.

De Gooijer, J. G. 1980. Exact moments of the sample autocorrelations from series generated by general ARIMA processes of order (p,d,q), d = 0 or 1. J. of Econometrics 14: 365-379.

Dickey, D. A. 1976. Hypothesis testing for nonstationary time series. Unpublished Ph. D. thesis. Iowa State University, Ames, Iowa.

Dickey, D. A., and W. A. Fuller. 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Amer. Statist. Assoc. 74: 427-431.

Durbin, J. 1960. Estimation of parameters in time series regression models. J. R. Statist. Soc., Ser. B, 22: 139-153.

Fuller, W. A. 1976. Introduction to Statistical Time Series. Wiley, New York.

Fuller, W. A., and D. P. Hasza. 1981. Properties of predictors for autoregressive time series. J. Amer. Statist. Assoc. 76: 156-161.

Fuller, W. A., D. P. Hasza, and J. J. Goebel. 1981. Estimation of the parameters of stochastic difference equations. The Annals of Statistics 9: 531-543.

Hannan, E. J. 1965. The estimation of relationships involving distributed lags. Econometrica 33: 206-224.

Hannan, E. J., and C. C. Heyde. 1972. On limit theorems for quadratic functions of discrete time series. Ann. Math. Statist. 43: 2058-2066.

Hannan, E. J., and D. F. Nicholls. 1972. The estimation of mixed regression, autoregression, moving average and distributed lag models. Econometrica 40: 529-548.

Hasza, D. P. 1977. Estimation in nonstationary time series. Unpublished Ph. D. thesis. Iowa State University, Ames, Iowa.

Hatanaka, M. 1974. An efficient two-step estimator for the dynamic adjustment model with autoregressive errors. J. Econometrics 2: 199-220.

Jenkins, G. M., and D. G. Watts. 1968. Spectral Analysis and Its Applications. Holden-Day, San Francisco.

Jones, R. H. 1978. Multivariate autoregression estimation using residuals. Pages 139-162 in D. F. Findley, ed. Applied Time Series Analysis. Academic Press, New York.

Koopmans, T. C., H. Rubin, and R. B. Leipnik. 1950. Measuring the equation systems of dynamic economics. Pages 53-237 in T. C. Koopmans, ed. Statistical Inference in Dynamic Economic Models. Wiley, New York.

Lee, H. E. 1981. Estimation of seasonal autoregressive time series. Unpublished Ph. D.thesis. Iowa State University, Ames, Iowa.

Levinson, N. 1947. The Weiner RMS (root mean square) error criterion in filter design and prediction. J. Math. Physics 25: 261-278.

Mann, H. B., and A. Wald. 1943. On the statistical treatment of linear stochastic difference equations. Econometrica 11: 173-220.

Marquardt, D. W. 1963. An algorithm for least squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. 11: 431-441.

Marriott, F. H. C, and J. A. Pope. 1954. Bias in the estimation of autocorrelations. Biometrika 41: 390-402.

Min, A. S. 1975. Study of likelihood functions of ARIMA processes. Unpublished Ph. D. thesis. Univ. of Iowa, Iowa City, Iowa.

Narasimham, G. V. L. 1969. Some properties of estimators occurring in the theory of linear stochastic processes. Pages 375-389 in M. Beckman and H. P. Kunzi, eds. Lecture Notes in Operations Research and Mathematical Economics. Springer, Berlin.

Nelson, C. R. 1973. Applied Time Series Analysis. Holden-Day, San Francisco, California.

Orcutt, G. H., and H. S. Winokur. 1969. First-order autoregression: inference, estimation and prediction. Econometrica 37: 1-14.

Parzen, E., and M. Pagano. 1977. An approach to modeling seasonally stationary time series. Statistical Science Division, SUNY at Buffalo, Technical Report No. 55.

Quenouille, M. H. 1949. Approximate tests of correlation in time series. J. R. Statist. Soc., Ser. B, 11: 63-68.

Quenouille, M. H. 1956. Notes on bias in estimation. Biometrika 43: 353-360.

Rao, M.M. 1961. Consistency and limit distributions of estimators of parameters in explosive stochastic difference equations. Ann. Math. Statist. 32: 195-218.

Reinsel, G. 1976. Maximum likelihood estimation of stochastic linear difference equations with autoregressive moving average errors. Technical Report 112, Dept. Statist. Carnegie-Mellon Univ., Pittsburgh.

Robinson, E. A., and M. T. Silvia. 1980. Digital Foundations of Time Series Analysis: The Box-Jenkins Approach. Holden-Day, San Francisco.

Rubin, H. 1950. Consistency of maximum-likelihood estimates in the explosive case. Pages 356-364 in T. C. Koopmans ed. Statistical Inference in Dynamic Economic Models. Wiley, New York.

Salem, A. S. 1971. Investigation of alternative estimators of the parameters of autoregressive processes. Unpublished M. S. thesis. Iowa State University, Ames, Iowa.

Sawa, T. 1978. The exact moments of the least squares estimator for the autoregressive model. J. of Econometrics 8: 159-172.

Scott, D. J. 1973. Central limit theorems for martingales and for processes with stationary increments using a Skorokhod representation approach. Advances Appl. Probability 5: 119-137.

Slutsky, E. 1937. The summation of random causes as the source of cyclic processes. Econometrica 5: 105-146.

Stigum, B. P. 1974. Asymptotic properties of dynamic parameter estimates (III). J. Multivariate Anal. 4: 351-381.

Thornber, A. 1967. Finite sample Monte Carlo studies, an autoregressive illustration. J. Amer. Statist. Assoc. 62: 801-818.

Ulrych, T. J., and T. N. Bishop. 1975. Maximum entropy spectral analysis and autoregressive decomposition. Rev. of Geophysics and Space Physics 13: 183-200.

Venkataraman, K. N. 1967. A note on the least square estimators of the parameters of a second order linear stochastic difference equation. Calcutta Statist. Assoc. Bull. 16: 15-28.

Venkataraman, K. N. 1968. Some limit theorems on a linear stochastic difference equation with a constant term, and their statistical applications. Sankhya, Series A, 30: 51-74.

Venkataraman, K. N. 1973. Some convergence theorems on a second order linear explosive stochastic difference equation with a constant term. J. Indian Statist. Assoc. 11: 47-69.

Walker, G. T. 1931. On periodicity in series of related terms. Philos. Trans. Roy. Soc. London, Ser. A, 131: 518-532.

White, J. S. 1958. The limiting distribution of the serial correlation coefficient in the explosive case. Ann. Math. Statist. 29: 1188-1197.

Wold, H. O. A. 1954. A study in the Analysis of Stationary Time Series, 2nd ed. Almqvist and Wiksell, Uppsala.

Yule, G. U. 1926. Why do we sometimes get nonsense-correlations between time series? A study in sampling and the nature of time series. J. R. Statist. Soc. 89: 1-64.

Yule, G. U. 1927. On a method of investigating periodicities in distributed series with special reference to Wolfer's sunspot series. Philos. Trans. Roy. Soc. London, Ser. A, 226: 267-298.

ACKNOWLEDGEMENTS

I wish to express my appreciation and gratitude to Professor Wayne A. Fuller who was a constant source of guidance and helpful suggestions during the course of this study. I am grateful to Professor Fuller for sharing his enthusiasm, experience and insight. Without his help I certainly would not have been able to complete this work.

I thank my parents for their constant encouragement and moral support.

A special thanks to Professor Roy D. Hickman and Professor Dean Isaacson for their constant help and for making the days just a little bit brighter.

I cannot fail to acknowledge the contribution of Helen Nelson, who has always been there when I needed her.

Thanks to all my friends, who made the potentially difficult and frustrating years of graduate study into a very nice time which I shall always remember.

It is a pleasure to thank Marylou Nelson for her superb job in typing this manuscript. I thank Jane Stowe for her help in bringing this manuscript into the final form.

APPENDIX A: DEFINITIONS AND PROPERTIES OF STATIONARY TIME SERIES

Some of the properties of stationary time series are presented. We begin with some definitions.

Definition A.1: A time series $\{X_t: t\varepsilon T\}$ is called strictly stationary if the joint distribution of $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ is the same as that of $(X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_n+h})$ for all possible sets of indices $t_1, t_2, \ldots, t_n$ and $t_1 + h, t_2 + h, \ldots, t_n + h$ in the index set T.

Definition A.2: A time series $\{X_t: t\varepsilon T\}$ is called (weakly) stationary if it has finite second moments and

a) the expected value of X_t is a constant for all t,

b) the covariance matrix of $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ is the same as the covariance matrix of $(X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_n+h})$ for all nonempty finite sets of indices $t_1, t_2, \ldots, t_n$ and all h such that $t_1, t_2, \ldots, t_n, t_1 + h, t_2 + h, \ldots, t_n + h$ are contained in the index set.

For a stationary time series $\{X_t\}$ the covarinace of X_{t+h} and X_t depends only on the distance, h, and we may write

$$\mathrm{Cov}(X_t, X_{t+h}) = \gamma(h).$$

The function $\gamma(h)$ is called the autocovariance function of X_t. The autocorrelation function of X_t is defined as

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} .$$

Theorem A.1: The covariance function of a real valued stationary time series is an even function of h. That is, $\gamma(h) = \gamma(-h)$.

Proof. See Fuller (1976, p. 9). □

Definition A.3: The first difference of a sequence $\{y_t\}_{t=0}^{\infty}$ is defined by

$$\Delta y_t = y_t - y_{t-1}, \quad t = 1, 2, \ldots$$

and the n-th difference is defined by

$$\Delta^n y_t = \Delta^{n-1} y_t - \Delta^{n-1} y_{t-1}$$

$$= \sum_{r=0}^{n} (-1)^r \binom{n}{r} y_{t-r}, \quad t = n, n+1, \ldots,$$

where

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

are the binomial coefficients.

Theorem A.2. Let y_t be a polynomial of degree p whose domain is the integers. Then the first difference Δy_t is expressible as a polynomial of degree $(p - 1)$ in t and the $(p + 1)$st difference $\Delta^{p+1} y_t$ is identically zero.

Proof. See Fuller (1976, p. 43). □

A linear difference equation of order p is given by

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + r_t, \quad t = p, p+1, \dots, \tag{A.1}$$

where the $\alpha_1, \alpha_2, \dots, \alpha_p$ are real constants, $\alpha_p \neq 0$, and r_t is a real function of t. The <u>characteristic equation</u> associated with the difference equation (A.1) is

$$m^p - \alpha_1 m^{p-1} - \alpha_2 m^{p-2} - \dots - \alpha_p = 0. \tag{A.2}$$

The solution of a linear difference equation with $r_t \equiv 0$ can be obtained using the roots of the characteristic equation (A.2). The solution is a sum of p terms where:

1. For every real and distinct root, m, a term of the form bm^t is included.
2. For every real root of order s (a root repeated s times), a term of the form

$$(b_1 + b_2 t + \dots + b_s t^{s-1})\, m^t$$

   is included.
3. For each pair of unrepeated complex conjugate roots, a term of the form $\alpha r^t \text{Cos}(t\theta+\beta)$ is included, where $m = r\, e^{i\theta}$.
4. For a pair of complex conjugate roots repeated s times, a term of the form

$$r^t[\alpha_1 \operatorname{Cos}(t\theta+\beta_1) + \alpha_2 t \operatorname{Cos}(t\theta+\beta_2) + \dots + \alpha_s t^{s-1} \operatorname{Cos}(t\theta+\beta_s)]$$

is included.

The solution to the linear difference equation in (A.1) is

$$y_t = \sum_{j=0}^{t-p} w_j \, r_{t-j} + \tau_t,$$

where the w_j are defined by

$$w_0 = 1, \; w_j = 0, \; j < 0,$$

$$w_j - \alpha_1 w_{j-1} - \dots - \alpha_p \, w_{j-p} = 0, \quad j = 1, 2, \dots,$$

τ_t is the first element of the vector $\underset{\sim}{A}^t \, \underset{\sim}{x}_0$, $\underset{\sim}{x}_0 = (y_p, y_{p-1}, \dots, y_1)'$, and,

$$\underset{\sim}{A} = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{p-1} & \alpha_p \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} .$$

<u>Definition A.4.</u> A <u>p-th order autoregressive</u> time series $\{Y_t\}$ is defined by

$$X_t = \sum_{j=1}^{p} \alpha_j \, X_{t-j} + e_t$$

where $\alpha_p \neq 0$, e_t are uncorrelated $(0,\sigma^2)$ random variables, and $X_t = Y_t - E(Y_t)$.

The following two results give an alternative representation for the stationary autoregressive process.

Theorem A.3. Let the roots of the polynomial equation (A.2) be less than one in absolute value, where $\alpha_p \neq 0$, and let the weights $\{w_j\}_{j=0}^{\infty}$ be defined by the solution of the difference equation

$$w_j = \alpha_1 w_{j-1} + \alpha_2 w_{j-2} + \ldots + \alpha_p w_{j-p}, \quad j = 1,2,\ldots$$

subject to the boundary conditions

$$w_0 = 1, \; w_j = 0 \quad \text{for } j < 0.$$

Let $\{e_t\}$ be a sequence of uncorrelated $(0,\sigma^2)$ random variables. Then,

$$\sum_{j=0}^{\infty} | w_j | < \infty,$$

and the mean square limit $X_t = \Sigma_{j=0}^{\infty} w_j e_{t-j}$, is a stationary process.

Moreover, X_t satisfies the stochastic difference equation

$$X_t = \sum_{j=1}^{p} \alpha_j X_{t-j} + e_t$$

for almost every realization of $\{e_t\}$.

Proof. See Fuller (1976, p. 56). □

Theorem A.4. Let X_t be stationary and satisfy

$$X_t = \sum_{j=1}^{p} \alpha_j X_{t-j} + e_t,$$

where the e_t are uncorrelated $(0,\sigma^2)$ random varialbes and the roots of the characteristic polynomial

$$m^p - \alpha_1 m^{p-1} - \dots - \alpha_p = 0,$$

are less than one in absolute value. Then,

$$X_t = \sum_{j=0}^{\infty} w_j e_{t-j}$$

for almost every realization of $\{e_t\}$ where w_j are defined in Theorem A.3. Moreover, the covariance function of X_t satisfies

$$\gamma(h) = \alpha_1 \gamma(h-1) + \alpha_2 \gamma(h-2) + \dots + \alpha_p \gamma(h-p),$$

$$\gamma(h) = \sum_{j=0}^{\infty} w_j w_{j+h} \sigma^2, \qquad h = 1,2,\dots,$$

and,

$$\sum_{h=-\infty}^{\infty} \mid \gamma(h) \mid < \infty .$$

Proof. See Fuller (1976, p. 58). □

Theorem A.5. Let $\{X_t\}$, $\{w_j\}$, and $\{\gamma(h)\}$ be as defined in Theorem A.4. Assume the roots of the polynomial equation (A.2) are less than one in absolute value. Then,

$$\sum_{h=-\infty}^{\infty} \gamma(h) = (\sum_{j=0}^{\infty} w_j)^2 \sigma^2$$

$$= (1 - \sum_{j=1}^{p} \alpha_j)^{-2} \sigma^2 .$$

Proof. By Theorems A.4 and A.5,

$$\sum_{h=-\infty}^{\infty} \mid \gamma(h) \mid < \infty \quad \text{and} \quad \sum_{j=0}^{\infty} \mid w_j \mid < \infty .$$

Therefore,

$$\sum_{h=-\infty}^{\infty} \gamma(h) = 2 \sum_{h=1}^{\infty} \gamma(h) + \gamma(0)$$

$$= 2 \sum_{h=1}^{\infty} \sum_{j=0}^{\infty} w_j w_{j+h} \sigma^2 + \sum_{j=0}^{\infty} w_j^2 \sigma^2$$

$$= \sum_{j=0}^{n} \{w_j^2 + 2 \sum_{h=1}^{\infty} w_j w_{j+h}\} \sigma^2$$

$$= \{ \sum_{j=0}^{\infty} w_j\}^2 \sigma^2.$$

Let $f(m) = m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p$. Since $f(1)$ is not zero,

$1 - \sum_{j=1}^{p} \alpha_j \neq 0$. Note,

$$w_j = \sum_{i=1}^{p} \alpha_i \, w_{j-i}, \qquad i = 1,2,\ldots,$$

with $w_j = 0$ for $j < 0$ and $w_0 = 1$. Therefore,

$$\sum_{j=1}^{\infty} w_j = \sum_{i=1}^{p} \alpha_i \sum_{j=1}^{\infty} w_{j-i}$$

$$= \sum_{i=1}^{p} \alpha_i \sum_{j=0}^{\infty} w_j$$

and,

$$\sum_{j=0}^{\infty} w_j = (1 - \sum_{i=1}^{p} \alpha_i)^{-1}. \qquad \square$$

We now give definitions and results related to order in magnitude and order in probability.

<u>Definition A.5.</u> We say $\{a_n\}$ is of smaller order than $\{g_n\}$ and write $a_n = o(\,|g_n|\,)$ if $\lim_{n\to\infty} g_n^{-1} a_n = 0$. We say $\{a_n\}$ is at most of order $\{g_n\}$ and write $a_n = O(\,|\,g_n\,|\,)$ if there exists a real number M such that $|\,g_n^{-1}\,a_n| \leq M$ for all n.

<u>Definition A.6.</u> Let $\{g_n\}$ be a sequence of positive real numbers and $\{X_n\}$ a sequence of random variables. We say X_n is of smaller order

in probability than g_n and write $X_n = o_p(g_n)$ if $\text{plim } g_n^{-1} X_n = 0$. We say X_n is at most of order in probability g_n and write $X_n = O_p(g_n)$ if, for every $\varepsilon > 0$, there exists a positive real number M_ε such that

$$P[\ |X_n| > M_\varepsilon\ g_n] < \varepsilon$$

for all n.

Definition A.7. If $\{X_n\}$ is a sequence of random variables with distribution functions $\{F_n(x)\}$, then $\{X_n\}$ is said to converge in distribution (or in law) to the random variable X with distribution function $F_X(x)$, and we write $X_n \xrightarrow{L} X$, if $\lim_{n\to\infty} F_n(x) = F_X(x)$ at all x for which $F_X(x)$ is continuous.

Theorem A.6. Let $\{X_n\}$ and X be random variables such that $\text{plim } X_n = X$. If $g(x)$ is a continuous function, then the distribution of $g(X_n)$ converges to the distribution of $g(X)$.

Proof. See Fuller (1976, p. 195). □

Theorem A.7. Let $\{X_n\}$ and $\{Y_n\}$ be two sequences of k-dimensional random variables and let $\{A_n\}$ be a sequence of $k \times k$ random matrices. Suppose there exists a random variable $\underset{\sim}{X}$, a fixed vector $\underset{\sim}{b}$, and a fixed nonsingular matrix $\underset{\sim}{A}$ such that
$\underset{\sim}{X}_n \xrightarrow{L} \underset{\sim}{X}$, $\underset{\sim}{Y}_n \xrightarrow{L} \underset{\sim}{b}$, and $\underset{\sim}{A}_n \xrightarrow{L} \underset{\sim}{A}$. Then,

(i) $\quad \underset{\sim}{X}_n + \underset{\sim}{Y}_n \xrightarrow{L} \underset{\sim}{X} + \underset{\sim}{b}$,

(ii) $\underset{\sim}{X}'_n \, \underset{\sim}{Y}_n \xrightarrow{L} \underset{\sim}{X}' \, \underset{\sim}{b}$,

and,

(iii) $\underset{\sim}{A}_n^{-1} \, \underset{\sim}{X}_n \xrightarrow{L} \underset{\sim}{A}^{-1} \, \underset{\sim}{X}$.

Proof. See Fuller (1976, p. 199). □

We now give the definitions and the results relating to martingale sequences.

Definition A.8. A sequence of random variables and σ - fields $\{X_n, F_n\}$ is called a martingale if we have for each n:

(a) F_n is a sub σ-field of F_{n+1} and
X_n is F_n measurable;

(b) $E \left|X_n\right| < \infty$;

(c) $X_n = E(X_{n+1} \mid F_n)$, a.e.

A sequence $\{e_t, F_t\}$ is called a martingale difference if

(a) $\{F_n\}$ is an increasing sequence of σ-fields and e_n is F_n measurable;

(b) $E \left|e_t\right| < \infty$;

(c) $E(e_t \mid F_{t-1}) = 0$, a.e.

Note if $\{e_t, F_t\}$ is a sequence of martingale differences, then $X_n = \Sigma_{t=1}^n e_t$ with F_n is a martingale.

A version of the martingale central limit theorem is given in the following theorem and is taken from Scott (1973).

Theorem A.8. Let $\{Z_{tn}: 1 \leq t \leq n, \ n = 1,2,\ldots\}$ denote a triangular array of random variables defined on the probability space (Ω, B, P). Let $S_{kn} = \Sigma_{t=1}^{k} Z_{tn}$, for $1 \leq k \leq n$, $n \geq 1$ with $S_{0n} \equiv 0$ for $n \geq 1$. Assume that for $1 \leq k \leq n$

$$E[S_{kn} \mid B_{k-1,n}] \quad \text{a.e.},$$

where $B_{k-1,n}$ denotes the sigma field generated by $S_{1n}, S_{2n}, \ldots, S_{k-1,n}$. Let

$$\delta_{tn}^2 = E[Z_{tn}^2 \mid B_{t-1,n}],$$

$$V_{nn}^2 = \sum_{j=1}^{n} \delta_{jn}^2,$$

$$s_{nn}^2 = E(V_{nn}^2) = E(S_{nn}^2).$$

Assume

(i) $s_{nn}^{-2}\, V_{nn}^2 \xrightarrow{P} 1$

and

(ii) $\lim_{n\to\infty} s_{nn}^{-2} \sum_{j=1}^{n} E[Z_{jn}^2\, I(|Z_{jn}| > \varepsilon\, s_{nn})] = 0,$

for all $\varepsilon > 0$, where $I(A)$ is the indicator function of the set A.

Then,

$$s_{nn}^{-1} S_{nn} \xrightarrow{L} N(0,1).$$

Proof. See Scott (1973). □

We now give some of the limiting properties of the estimator for the autocovariances.

Theorem A.9. A stationary time series with absolutely summable covariance function is ergodic for the mean. Furthermore,

$$\lim_{n\to\infty} n \operatorname{Var}\{\bar{x}_n\} = \sum_{h=-\infty}^{\infty} \gamma(h)$$

where $\bar{x}_n = n^{-1} \sum_{t=1}^{n} X_t$.

Proof. See Fuller (1976, p. 232). □

Theorem A.10. Let the time series $\{X_t\}$ be defined by

$$X_t = \sum_{j=0}^{\infty} w_j \, e_{t-j}$$

where the sequence of $\{w_j\}$ is absolutely summable and the $\{e_t\}$ is a sequence of martingale differences with

$$E(e_t^2 \mid F_{t-1}) = \sigma^2 \quad \text{a.s.} ,$$

$$E(e_t^4) = \eta\sigma^4 ,$$

and F_t is the σ-field generated by $(e_1, e_2, \ldots, e_t)$. Then, for fixed h and q, $(h \geq q \geq 0)$

$$\lim_{n\to\infty} (n-q) \operatorname{Cov}(\tilde{\gamma}(h),\tilde{\gamma}(q)) = (\eta-3)\gamma(h)\gamma(q) + \sum_{p=-\infty}^{\infty} [\gamma(p)\gamma(p-h+q)+\gamma(p+q)\gamma(p-h)],$$

where

$$\tilde{\gamma}(h) = (n-h)^{-1} \sum_{t=1}^{n-h} X_t X_{t+h}.$$

<u>Proof.</u> See Fuller (1976, p. 238). □

<u>Remark.</u> The result given in Fuller (1976) is for $\{e_t\}$ a sequence of independent random variables. But the proof extends immediately for the martingale differences.

<u>Theorem A.11.</u> Given fixed $h \geq q \geq 0$ and a time series X_t satisfying the assumptions of Theorem A.10,

$$E[\hat{\gamma}(h)-\gamma(h)] = -\frac{h}{n}\gamma(h) - \frac{n-h}{n}\operatorname{Var}\{\bar{x}_n\} + O(n^{-2})$$

and

$$\lim_{n\to\infty} \frac{n^2}{(n-h)} \operatorname{Cov}\{\hat{\gamma}(h), \hat{\gamma}(q)\} = \lim_{n\to\infty} (n-q)\operatorname{Cov}\{\tilde{\gamma}(h),\tilde{\gamma}(q)\},$$

where

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-h} (X_t-\bar{x}_n)(X_{t+h}-\bar{x}_n).$$

and

$$\bar{x}_n = n^{-1} \sum_{t=1}^{n} X_t .$$

Proof. See Fuller (1976, p. 239). □

Lemma A.1. Let the random variables Z_n with distribution functions $F_n(z)$ be defined by

$$Z_n = S_{kn} + D_{kn}$$

for $k = 1, 2, \ldots,$ and $n = 1, 2, \ldots$. Let

$$\operatorname*{plim}_{k\to\infty} D_{kn} = 0$$

uniformly in n. Let

$$S_{kn} \xrightarrow{L} \psi_k \quad \text{as} \quad n \to \infty$$

and

$$\psi_k \xrightarrow{L} Z \quad \text{as} \quad k \to \infty.$$

Then,

$$Z_n \xrightarrow{L} Z.$$

Proof. See Fuller (1976, p. 248). □

Theorem A.12. Let the time series X_t be defined by

$$X_t - \mu = \sum_{j=0}^{\infty} w_j e_{t-j}, \quad t = 1,2,\ldots,$$

where $w_0 = 1$, $\Sigma_{j=0}^{\infty} |w_j| < \infty$ and $\{e_t\}$ is a sequence of $(0,\sigma^2)$ random variables with

$$E(e_t \mid F_{t-1}) = 0 \quad \text{a.e.}$$

$$E(e_t^2 \mid F_{t-1}) = \sigma^2 \quad \text{a.e.} ,$$

$$E(e_t^4) < L < \infty .$$

Then,

$$\bar{x}_n \xrightarrow{P} \mu$$

and,

$$\hat{\gamma}(h) = \gamma(h) + O_p(n^{-1/2}), \quad \text{for fixed } h.$$

Proof. Note that if $E(e_t^4) = \eta\sigma^4$ then we have from Theorem A.9 and A.10,

$$\lim_{n\to\infty} n \operatorname{Var}\{\bar{x}_n\} = \left(\sum_{j=0}^{\infty} w_j\right)^2 \sigma^2$$

$$\operatorname{Var}\{\hat{\gamma}(h)\} = O(n^{-1})$$

and,

$$E\{\hat{\gamma}(h)\} = \gamma(h) + O(n^{-1}).$$

The arguments used in the case $E(e_t^4) = \eta\sigma^4$ extend immediately to the case where $E(e_t^4)$ is bounded. Therefore,

$$\operatorname{plim} \bar{x}_n = \mu$$

and,

$$\hat{\gamma}(h) = \gamma(h) + O_p(n^{-1/2}). \qquad \square$$

Consider a stationary p-th order autoregressive process $\{Y_t\}$ satisfying

$$Y_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j Y_{t-j} + e_t, \qquad t = 1,2,\ldots,$$

where $\{e_t\}$ is a sequence of uncorrelated $(0, \sigma^2)$ random variables. The least squares estimator of $(\alpha_0, \alpha_1, \ldots, \alpha_p)$ is obtained by regressing Y_t on $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$ with an intercept. Assume that the roots of the characteristic equation

$$m^p - \sum_{j=1}^{p} \alpha_j m^{p-j} = 0$$

lie inside the unit circle.

Assuming $\{e_t\}$ to be a sequence of normal independent $(0, \sigma^2)$ random variables, Mann and Wald (1943) established that the asymptotic distribution of the least squares estimator is normal.

Anderson (1959) extended Mann and Wald's result to the case where e_t are assumed to be independent $(0, \sigma^2)$ random variables with bounded $(2 + \delta)$ - th moments, for some $\delta > 0$.

Hannan and Heyde (1972) considered the case where e_t are assumed to satisfy

(i) $E[e_t \mid F_{t-1}] = 0$ a.s. ,

(ii) $E[e_t^2 \mid F_{t-1}] = \sigma^2$ a.s. ,

and

(iii) there exists a random variable X, such that

$$P[\,|e_t| > u] \leq c\, P\,[\,|X| > u]$$

for some c real and $E\,|X|^4 < \infty$.

where F_t is the σ-field generated by $(e_1, e_2, \ldots, e_t)$. Under these assumptions they established that the limiting distribution of the least squares estimator is normal.

We consider the case where the condition (iii) on $\{e_t\}$ is replaced by the condition,

(iii') $E(e_t^{4+2\nu}) < L < \infty$

for some $\nu > 0$ and L. The following theorem establishes the limiting distribution of the least squares estimator for this case.

Theorem A.13. Let Y_t satisfy

$$Y_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j Y_{t-j} + e_t, \quad t = 1,2,\ldots,$$

where $Y_0, Y_{-1}, \ldots, Y_{-p+1}$ are initial conditions. Let the roots of

$$m^p - \sum_{j=1}^{p} \alpha_j m^{p-j} = 0$$

be less than one in absolute value. Let $\{e_t\}$ be a sequence of random variables with

$$E\{e_t \mid F_{t-1}\} = 0 \quad \text{a.e.},$$
$$E\{e_t^2 \mid F_{t-1}\} = \sigma^2 \quad \text{a.e.},$$

and

$$E\{e_t^{4+2\nu}\} < L < \infty \text{ for some } \nu > 0 \text{ and real } L.$$

Assume either

(i) $Y_0, Y_{-1}, \ldots, Y_{-p+1}$ are fixed, or

(ii) $Y_0, Y_{-1}, \ldots, Y_{-p+1}$ are random variables independent of $\{e_t\}$ with mean $\alpha_0(1 - \Sigma_{i=1}^{p} \alpha_i)^{-1}$, and the variance and covariances are given by $\gamma(h)$ where $\gamma(h)$ is the autocovariance function of a statinary p-th order autoregressive process with the coefficients $\underset{\sim}{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_p)'$.

Let

$$\hat{\underset{\sim}{\alpha}} = [\sum_{t=1}^{n} \underset{\sim}{X}_t' \underset{\sim}{X}_t]^{-1} \sum_{t=1}^{n} \underset{\sim}{X}_t' Y_t ,$$

where

$$\underset{\sim}{X}_t = (1, Y_{t-1}, \ldots, Y_{t-p}).$$

Then,

$$n^{1/2} (\hat{\underset{\sim}{\alpha}} - \underset{\sim}{\alpha}) \xrightarrow{L} N(\underset{\sim}{0}, \underset{\sim}{\Gamma}^{-1} \sigma^2) ,$$

where

$$\underset{\sim}{\Gamma} = \lim_{n \to \infty} n^{-1} \sum_{t=1}^{n} E[\underset{\sim}{X}_t' \underset{\sim}{X}_t].$$

<u>Proof.</u> We have

$$(\hat{\underset{\sim}{\alpha}} - \underset{\sim}{\alpha}) = (n^{-1} \sum_{t=1}^{n} \underset{\sim}{X}_t' \underset{\sim}{X}_t)^{-1} [n^{-1} \sum_{t=1}^{n} \underset{\sim}{X}_t' e_t].$$

The probability of $\sum_{t=1}^{n} \underset{\sim}{X}_t' \underset{\sim}{X}_t$ being nonsingular converges to one as $n \to \infty$.

Let $\underset{\sim}{\eta}$ be an arbitrary column vector of real numbers such that $\underset{\sim}{\eta}' \underset{\sim}{\eta} \neq 0$. Let

$$n^{-1/2} \underset{\sim}{\eta}' \sum_{t=1}^{n} \underset{\sim}{X}_t' e_t = \sum_{t=1}^{n} Z_{tn} = S_{nn},$$

where

$$Z_{tn} = n^{-1/2} \underset{\sim}{\eta}' \underset{\sim}{X}_t' e_t.$$

Note the σ - field generated by $Z_{1n}, Z_{2n}, \ldots, Z_{t-1,n}$ is

$$B_{t-1,n} = \sigma(e_1, e_2, \ldots, e_{t-1}; Y_0, Y_{-1}, \ldots, Y_{-p+1}).$$

We verify the conditions of Theorem A.8. Since $Y_0, Y_{-1}, \ldots, Y_{-p+1}$ are independent of $\{e_t\}$,

$$E[Z_{tn} \mid B_{t-1,n}] = 0,$$

$$E[Z_{tn}^2 \mid B_{t-1,n}] = n^{-1} \underset{\sim}{\eta}' \underset{\sim}{X}_t' \underset{\sim}{X}_t \underset{\sim}{\eta} \sigma^2,$$

$$V_{nn}^2 = n^{-1} \sum_{t=1}^{n} \underset{\sim}{\eta}' \underset{\sim}{X}_t' \underset{\sim}{X}_t \underset{\sim}{\eta} \sigma^2,$$

$$s_{nn}^2 = n^{-1} \sum_{t=1}^{n} \underset{\sim}{\eta}' E(\underset{\sim}{X}_t' \underset{\sim}{X}_t) \underset{\sim}{\eta} \sigma^2.$$

Under the assumption (ii) $\{Y_t\}$ is a stationary p-th order autoregressive process and by Theorem A.4, it can be represented as $Y_t = \mu + \Sigma_{j=0}^{\infty} w_j e_{t-j}$ with $\Sigma_{j=0}^{\infty} |w_j| < \infty$. Therefore, by Theorem A.12,

$$\operatorname*{plim}_{n\to\infty} n^{-1} \sum_{t=1}^{n} \underset{\sim}{X}_t' \underset{\sim}{X}_t = \underset{\sim}{\Gamma}.$$

Under the assumption (i), the effect of the initial conditions is transient, and

$$\operatorname{plim} n^{-1} \sum_{t=1}^{n} \underset{\sim}{X}_t' \underset{\sim}{X}_t = \underset{\sim}{\Gamma} .$$

Therefore,

$$s_{nn}^{-2} V_{nn}^2 \xrightarrow{P} 1.$$

We now investigate,

$$s_{nn}^{-2} \quad E[\sum_{t=1}^{n} Z_{tn}^2 \, I(|Z_{tn}| > t \, s_{nn})]$$

$$< s_{nn}^{-2-\nu} \varepsilon^{-2} \sum_{t=1}^{n} E \mid Z_{tn} \mid^{2+\nu}$$

Note that

$$E \mid Z_{tn} \mid^{2+\nu} = n^{-1-\frac{\nu}{2}} E[\mid \sum_{j=1}^{p} \eta_j Y_{t-j} e_t + \eta_0 e_t \mid]^{2+\nu}.$$

Consider

$$E\{ | Y_{t-1} e_t |^{2+\nu}\} < [E\{ |Y_{t-1}|^{4+2\nu}\} E\{ |e_t|^{4+2\nu}\}]^{1/2} .$$

Since Y_t is a linear combination of $\{e_t\}$ with absolutely summable weights, $E\{ |Y_t|^{4+2\nu}\} < K$ for some $K < \infty$. Therefore,

$$E\{ |Z_{tn}|^{2+\nu}\} = O(n^{-1-\frac{\nu}{2}}).$$

Therefore, the second condition of Theorem A.8 is satisfied and,

$$s_{nn}^{-1} S_{nn} \xrightarrow{L} N(0, \underset{\sim}{\eta}' \underset{\sim}{\Gamma} \underset{\sim}{\eta}). \qquad \square$$

Some extensions of Theorem A.13 are presented in Appendix B.

Theorem A.14. If $\{e_t\}$ is a sequence of uncorrelated random variables with zero mean and bounded second moments, then

$$n^{-1} \sum_{t=1}^{n} e_t \longrightarrow 0 \quad \text{a.s.}$$

Proof. See Chung (1974, p. 103). $\square$

We establish the order of the difference between two types of least squares estimators of the parameters of the stationary p-th order autoregressive process. Let $\{Y_t\}$ satisfy the stochastic difference equation

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + e_t \ , \tag{A.3}$$

where $\{e_t\}$ is a sequence of independent $N(0,\sigma^2)$ random variables. One form of the least squares estimator $\underset{\sim}{\alpha}^*$ of $\underset{\sim}{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_p)$.

$$\underset{\sim}{\alpha}^* = \hat{\underset{\sim}{H}}^{-1} \hat{\underset{\sim}{N}} \tag{A.4}$$

where $\hat{\underset{\sim}{H}}$ and $\hat{\underset{\sim}{N}}$ are defined in (3.3). This form of the least squares estimator is obtained by regressing Y_t on $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$ with an intercept. Equivalently, $\underset{\sim}{\alpha}^*$ is obtained by regressing $Y_t - \overline{Y}_0$ on $Y_{t-1} - \overline{Y}_1, Y_{t-2} - \overline{Y}_2, \ldots, Y_{t-p} - \overline{Y}_p$, where

$$\overline{Y}_i = (n-p)^{-1} \sum_{t=p+1}^{n} Y_{t-i}, \qquad i = 1,2,\ldots,p \ .$$

An alternative form of the least squares estimator $\underset{\sim}{\alpha}^+$ is obtained by regressing $Y_t - \overline{Y}$ on $Y_{t-1} - \overline{Y}, Y_{t-2} - \overline{Y}, \ldots, Y_{t-p} - \overline{Y}$ where $\overline{Y} = n^{-1} \Sigma_{t=1}^{n} Y_t$. The estimator $\underset{\sim}{\alpha}^+$ is given by

$$\underset{\sim}{\alpha}^+ = (\underset{\sim}{H}^+)^{-1} \underset{\sim}{N}^+ \ , \tag{A.5}$$

where

$$\underset{\sim}{H}^+ = ((h_{ij}^+)) \ ,$$

$$\underset{\sim}{N}^{+} = (N_1^{+}, N_2^{+}, \ldots ; N_p^{+})' ,$$

$$h_{ij}^{+} = (n-p)^{-1} \sum_{t=p+1}^{n} (Y_{t-i} - \overline{Y})(Y_{t-j} - \overline{Y}) ,$$

and

$$N_i^{+} = (n-p)^{-1} \sum_{t=p+1}^{n} (Y_{t-2} - \overline{Y})(Y_t - \overline{Y}) .$$

The following theorem establishes the order in probability of the difference between the estimators $\underset{\sim}{\alpha}^{*}$ and $\underset{\sim}{\alpha}^{+}$.

<u>Theorem A.15.</u> Assume $\{Y_t\}$ is stationary and satisfies (A.3). Let $\underset{\sim}{\alpha}^{*}$ and $\underset{\sim}{\alpha}^{+}$ be as defined in (A.4) and (A.5). Then

$$\underset{\sim}{\alpha}^{*} - \underset{\sim}{\alpha}^{+} = o_p(n^{-2}) .$$

<u>Proof.</u> Let $\hat{h}_{ij}$ and $\hat{N}_i$ denote the (i,j)-th element of $\hat{\underset{\sim}{H}}$ and i-th element of $\hat{\underset{\sim}{N}}$, respectively. Note that

$$\begin{aligned}
\hat{\underset{\sim}{h}}_{ij} &= (n-p)^{-1} \sum_{t=p+1}^{n} (Y_{t-i} - \overline{Y}_i)(Y_{t-j} - \overline{Y}_j) \\
&= (n-p)^{-1} \sum_{t=p+1}^{n} (Y_{t-i} - \overline{Y} + \overline{Y} - \overline{Y}_i)(Y_{t-j} - \overline{Y} + \overline{Y} - \overline{Y}_j) \\
&= (n-p)^{-1} \sum_{t=p+1}^{n} (Y_{t-i} - \overline{Y})(Y_{t-j} - \overline{Y}) \\
&\quad + (n-p)^{-1} \sum_{t=p+1}^{n} (\overline{Y} - \overline{Y}_i)(Y_{t-j} - \overline{Y})
\end{aligned}$$

$$+ (n-p)^{-1} \sum_{t=p+1}^{n} (Y_{t-i} - \overline{Y})(\overline{Y} - \overline{Y}_j)$$

$$+ (\overline{Y}_i - \overline{Y})(\overline{Y}_j - \overline{Y})$$

$$= h^{+}_{ij} - (\overline{Y}_i - \overline{Y})(\overline{Y}_j - \overline{Y}) \ .$$

Now,

$$\begin{aligned}
\overline{Y}_i - \overline{Y} &= (n-p)^{-1} \sum_{t=p+1}^{n} Y_{t-i} - n^{-1} \sum_{t=1}^{n} Y_t \\
&= (n-p)^{-1} \sum_{t=p+1-i}^{n-i} Y_t - n^{-1} \sum_{t=1}^{n} Y_t \\
&= [(n-p)^{-1} - n^{-1}] \sum_{t=p+1-i}^{n-i} Y_t - n^{-1} \sum_{t=1}^{p-i} Y_t - n^{-1} \sum_{t=n-i+1}^{n} Y_t \\
&= p\, n^{-1}(n-p)^{-1} \sum_{t=p+1-i}^{n-i} Y_t - n^{-1} \sum_{t=1}^{p-i} Y_t - n^{-1} \sum_{t=n-i+1}^{n} Y_t \\
&= 0_p(n^{-1}) \ .
\end{aligned}$$

Therefore,

$$\hat{h}_{ij} = h^{+}_{ij} + 0_p(n^{-2}) \ . \qquad \text{(A.6)}$$

Similarly,

$$\hat{N}_i = N_i^+ + o_p(n^{-2}) . \tag{A.7}$$

From Fuller and Hasza (1981), it follows that

$$\hat{\underset{\sim}{H}}^{-1} = O_p(1) ,$$

$$(\underset{\sim}{H}^+)^{-1} = O_p(1) ,$$

and

$$\begin{aligned} \underset{\sim}{\alpha}^* - \underset{\sim}{\alpha}^+ = & - \underset{\sim}{H}^{-1} [(\hat{\underset{\sim}{H}} - \underset{\sim}{H}^+) \underset{\sim}{\alpha} - (\hat{\underset{\sim}{N}} - \underset{\sim}{N}^+)] \\ & + \underset{\sim}{H}^{-1} (\hat{\underset{\sim}{H}} - \underset{\sim}{H}^+) \underset{\sim}{H}^{-1}[(\hat{\underset{\sim}{H}} - \underset{\sim}{H}) \underset{\sim}{\alpha} - (\hat{\underset{\sim}{N}} - \underset{\sim}{N})] \\ & + \underset{\sim}{H}^{-1} (\underset{\sim}{H}^+ - \underset{\sim}{H}) \underset{\sim}{H}^{-1}[(\hat{\underset{\sim}{H}} - \underset{\sim}{H}^+) \underset{\sim}{\alpha} - (\hat{\underset{\sim}{N}} - \underset{\sim}{N}^+)] \\ & + o_p(n^{-2}) \end{aligned}$$

where $\underset{\sim}{H}$ and $\underset{\sim}{N}$ are defined in (3.2). Using (A.6) and (A.7), we get

$$\underset{\sim}{\alpha}^* - \underset{\sim}{\alpha}^+ = o_p(n^{-2}) \qquad \square$$

Fuller and Hasza (1981) established that for n greater than some N_j depending upon j, $E[|\underset{\sim}{\alpha}^+|^{2j}]$ is finite for $j = 1, 2, \ldots$. Using

Theorem 5.4.4 of Fuller (1976, p. 208), it can be shown that

$$E[\underset{\sim}{\alpha}^* - \underset{\sim}{\alpha}^+] = O(n^{-2}) .$$

APPENDIX B: ASYMPTOTIC PROPERTIES OF THE LEAST SQUARES ESTIMATOR

Consider the p-th order autoregressive process with r explanatory variables.

$$Y_t = \underset{\sim}{\psi}_t \underset{\sim}{\beta} + \underset{\sim}{Y}'_{t-1} \underset{\sim}{\alpha} + e_t \tag{B.1}$$

where

$$\underset{\sim}{\beta} = (\beta_1, \beta_2, \ldots, \beta_r)', \quad \underset{\sim}{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_p)',$$
$$\underset{\sim}{Y}_{t-1} = (Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p})', \quad \underset{\sim}{\psi}_t = (\psi_{t1}, \psi_{t2}, \ldots, \psi_{tr});$$

$\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ are parameters to be estimated and $\underset{\sim}{\psi}_t$'s are explanatory variables. Let

$$F_t = \sigma \text{ - field of events generated by } \underset{\sim}{Y}_0, Y_1, \ldots, Y_{t-1}.$$

Assume

$$E(e_t \mid F_{t-1}) = 0 \quad \text{a.s.}$$

Suppose that observations $(Y_{1-p}, \ldots, Y_0, \ldots, Y_n)$ are available and estimate $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ by minimizing the sum of squared residuals. The equations for the least-squares estimators $(\hat{\underset{\sim}{\beta}}, \hat{\underset{\sim}{\alpha}})$ are

$$\Sigma\, \underset{\sim}{J}_t \begin{pmatrix} \hat{\underset{\sim}{\beta}} \\ \hat{\underset{\sim}{\alpha}} \end{pmatrix} = \begin{pmatrix} \Sigma\, \underset{\sim}{\psi}_t'\, Y_t \\ \Sigma\, \underset{\sim}{Y}_{t-1}\, Y_t \end{pmatrix}$$

where

$$\underset{\sim}{J}_t = \begin{bmatrix} \underset{\sim}{\psi}_t'\, \underset{\sim}{\psi}_t & \underset{\sim}{\psi}_t'\, \underset{\sim}{Y}_{t-1}' \\ \underset{\sim}{Y}_{t-1}\, \underset{\sim}{\psi}_t & \underset{\sim}{Y}_{t-1}\, \underset{\sim}{Y}_{t-1}' \end{bmatrix},$$

and all summations are over $t = 1, 2, \ldots, n$ unless otherwise indicated. Let $\underset{\sim}{\mu}_t = E(\underset{\sim}{Y}_t)$, $\sigma_t^2 = V(e_t)$, $\underset{\sim}{V}_n = E[\Sigma\, \underset{\sim}{Y}_{t-1}\, \underset{\sim}{Y}_{t-1}']$, $\underset{\sim}{L}_n = \Sigma[\underset{\sim}{\mu}_{t-1}\, \underset{\sim}{\psi}_t]$, $\underset{\sim}{M}_n = \Sigma[\underset{\sim}{\psi}_t'\, \underset{\sim}{\psi}_t]$, $\underset{\sim}{G}_n = \underset{\sim}{M}_n^{-1}\, \underset{\sim}{L}_n'$ (assuming $\underset{\sim}{M}_n^{-1}$ exists) and $\underset{\sim}{D}_n = E[\Sigma\, \underset{\sim}{J}_t]$. Also, let m_n and d_n be the smallest eigenvalues of $\underset{\sim}{M}_n$ and $\underset{\sim}{E}_n = \underset{\sim}{V}_n - \underset{\sim}{L}_n\, \underset{\sim}{M}_n^{-1}\, L_n'$, respectively. Let z_t denote the vector norm $\left| \underset{\sim}{\psi}_t \right| = (\Sigma_i \{\psi_{ti}\}^2)^{1/2}$ and g_n be the Euclidean matrix norm $\left| G_n \right| = (\Sigma_{ij} \{G_n\}_{ij}^2)^{1/2}$. Let $\sigma_0^2 = \left| V(\underset{\sim}{Y}_0) \right|$ and $\nu_t = \Sigma_{i=0}^t\, \lambda^{2i}\, \sigma_{t-i}^2$ where λ exceeds the largest of the moduli of the roots of the characteristic equation,

$$m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p = 0.$$

The main results of Crowder (1980) are summarized in the next three theorems.

Theorem B.1. The following conditions are sufficient for consistency of the least squares estimators $(\hat{\beta}, \hat{\alpha})$.

(a) $c_n^{-2} \Sigma \lambda^{t-1} (\sum_{r=0}^{n-t} \nu_r \psi_{r+1} \psi'_{r+t}) \longrightarrow 0$ as $n \longrightarrow \infty$

when $c_n = d_n$, d_n/g_n^2 and m_n;

(b) $c_n^{-2} \Sigma z_t^2 \sigma_t^2 \longrightarrow 0$ as $n \to \infty$ when $c_n = d_n$, d_n/g_n and m_n;

(c) $c_n^{-1} \{\nu_n + (\sum_{r=0}^{n} \lambda^r z_{n-r})^2\} \longrightarrow 0$ as $n \to \infty$

when $c_n = d_n$ and d_n/g_n;

(d) either $c_n^{-2} \Sigma \sigma_{1t}^2 \{\nu_{t-1} + (\sum_{r=0}^{t-1} \lambda^r z_{t-1-r})^2\} \longrightarrow 0$

or $c_n^{-2} \Sigma \gamma_t^2 \{\sum_{r=0}^{t-1} \lambda^r (\gamma_{t-1-r} + z_{t-1-r})\}^2 \longrightarrow 0$

for $c_n = d_n$ and d_n/g_n, where $\gamma_t^4 = E(e_t^4)$ and $\sigma_{1t}^2 > E(e_t^2 \mid F_{t-1})$ is a constant bound for the conditional error variance;

(e) $c_n^{-1} \Sigma(e_t^2 - \sigma_t^2) \longrightarrow 0$ in mean square as $n \to \infty$ when $c_n = d_n$ and d_n/g_n.

(f) $\lambda < 1$.

Theorem B.2. Let $X_t = e_t(\underset{\sim}{\psi}_t, \underset{\sim}{Y}'_{t-1})\, \underset{\sim}{x}$ for a given unit vector $\underset{\sim}{x}$, and write $s_n^2 = E[\Sigma\, X_t^2]$, $\underset{\sim}{C}_n = E[\Sigma\, e_t^2\, \underset{\sim}{J}_t]$. Suppose that the conditions of Theorem B.1. hold, and that for all $\underset{\sim}{x}$

$$s_n^{-2}\, \Sigma\, X_t^2 \xrightarrow{P} 1,$$

$$s_n^{-2}\, \Sigma\, X_t^2\, I(\,|X_t| > \varepsilon\, s_n) \xrightarrow{P} 0, \quad \text{for all } \varepsilon > 0.$$

Then the asymptotic distribution of

$$\underset{\sim}{C}_n^{-1/2}\, \underset{\sim}{D}_n \left[\begin{pmatrix} \hat{\underset{\sim}{\beta}} \\ \hat{\underset{\sim}{\alpha}} \end{pmatrix} - \begin{pmatrix} \underset{\sim}{\beta} \\ \underset{\sim}{\alpha} \end{pmatrix}\right] \quad \text{is } N_{r+p}\,(\underset{\sim}{0}, \underset{\sim}{I}) .$$

Theorem B.3. The least squares estimators $(\hat{\underset{\sim}{\beta}}, \hat{\underset{\sim}{\alpha}})$ are consistent if the limits L1 to L4 below hold. If L5 also holds, then their asymptotic distribution is normal as in Theorem B.2.

$$\text{L1:}\quad \underset{\sim}{E}_n^{-1}\, \underset{\sim}{W}_{1n} \xrightarrow{P} 0, \quad \underset{\sim}{G}_n\, \underset{\sim}{E}_n^{-1}\, \underset{\sim}{W}_{1n} \xrightarrow{P} 0,$$

$$\underset{\sim}{E}_n^{-1}\, \underset{\sim}{G}'_n\, \underset{\sim}{W}'_{1n} \xrightarrow{P} 0, \quad \underset{\sim}{G}_n\, \underset{\sim}{E}_n^{-1}\, \underset{\sim}{G}'_n\, \underset{\sim}{E}_n^{-1}\, \underset{\sim}{G}'_n\, \underset{\sim}{W}'_{1n} \xrightarrow{P} 0,$$

$$\underset{\sim}{M}_n^{-1}\, \underset{\sim}{W}'_{1n} \xrightarrow{P} 0;$$

L2: $\underset{\sim}{E}_n^{-1} \underset{\sim}{W}_{2n} \xrightarrow{P} 0$; $\underset{\sim}{G}_n \underset{\sim}{E}_n^{-1} \underset{\sim}{W}_{2n} \xrightarrow{P} 0$

L3: $\underset{\sim}{E}_n^{-1} \underset{\sim}{G}_n' \underset{\sim}{W}_{3n} \xrightarrow{P} 0$, $\underset{\sim}{G}_n \underset{\sim}{E}_n^{-1} \underset{\sim}{W}_{3n} \xrightarrow{P} 0$,

$\underset{\sim}{M}_n^{-1} \underset{\sim}{W}_{3n} \xrightarrow{P} 0$;

L4: $\underset{\sim}{E}_n^{-1} \underset{\sim}{W}_{4n} \xrightarrow{P} 0$, $\underset{\sim}{G}_n \underset{\sim}{E}_n^{-1} \underset{\sim}{W}_{4n} \xrightarrow{P} 0$;

L5: $\underset{\sim}{C}_n^{-1/2} \quad \Sigma\, e_t \begin{pmatrix} \underset{\sim}{\psi}_t' \\ \underset{\sim}{Y}_{t-1} \end{pmatrix} \xrightarrow{L} N_{r+p}(\underset{\sim}{0}, \underset{\sim}{I})$.

where

$$\underset{\sim}{W}_{1n} = \Sigma\, (\underset{\sim}{Y}_{t-1} - \underset{\sim}{G}_n' \underset{\sim}{\psi}_n') \underset{\sim}{\psi}_t,$$

$$\underset{\sim}{W}_{2n} = \Sigma\, (\underset{\sim}{Y}_{t-1} \underset{\sim}{Y}_{t-1}' - E[\underset{\sim}{Y}_{t-1} \underset{\sim}{Y}_{t-1}']),$$

$$\underset{\sim}{W}_{3n} = \Sigma\, \underset{\sim}{\psi}_t' e_t$$

and

$$\underset{\sim}{W}_{4n} = \Sigma\, \underset{\sim}{Y}_{t-1} e_t.$$

Fuller, Hasza, and Goebel (1981) studied the model (B.1) under the assumption that $\{e_t\}$ is a sequence of independent random variables.

They established the limiting distributions of the least squares estimators for the situations where the largest root is less than one, equal to one, and greater than one in absolute value. Their results for the stationary case and the unit root case extend immediately to the situation where e_t are martingale differences. The extensions are given in the following theorems.

Consider the model (B.1). Assume $Y_0, Y_{-1}, \ldots, Y_{-p+1}$ are known and fixed. Assume e_t satisfy,

$$E(e_t \mid F_{t-1}) = 0 \quad \text{a.s.},$$
$$E(e_t^2 \mid F_{t-1}) = \sigma^2 \quad \text{a.s.},$$

and

$$E(e_t^{4+2\nu}) < L < \infty \quad \text{for some } \nu > 0. \tag{B.2}$$

The parameters $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ are fixed unknown constants and ψ_{ti} are fixed functions of time.

The difference equation (B.1) may be solved to obtain

$$Y_t = S_t + u_t,$$

$$u_t = \sum_{j=0}^{t-1} v_j \, e_{t-j},$$

$$S_t = \sum_{j=0}^{p-1} v_{t+j} \, Y_{-j} + \sum_{j=0}^{t-1} v_j \sum_{i=1}^{r} \beta_i \, \psi_{t-j,i}, \tag{B.3}$$

and the v_j are given by

$$v_j - \alpha_1 v_{j-1} - \dots - \alpha_p v_{j-p} = 0$$

with $v_0 = 1$ and $v_j = 0$ for $j < 0$. Set $S_{-t} = Y_{-t}$ for $t = 0, 1, \dots, p-1$. Note that S_t is fixed and u_t is random. Let $m_1, m_2, \dots, m_p$ be the roots of the characteristic equation

$$m^p - \alpha_1 m^{p-1} - \dots - \alpha_p = 0$$

and let m_1 be the root with the largest absolute value. Assume $|m_1| \le 1$. Define

$$\begin{aligned} a &= 1 \quad \text{if} \quad |m_1| = 1, \\ &= 0 \quad \text{if} \quad |m_1| < 1. \end{aligned}$$

If $|m_1| = 1$ then consider the reparametrization,

$$Y_t = \sum_{i=1}^{r} \psi_{ti} \beta_i + \sum_{j=1}^{p-1} \alpha_{p+j}(Y_{t-j} - m_1 Y_{t-j-1}) + \alpha_{p+r} Y_{t-1} + e_t \tag{B.4}$$

where $\alpha_{p+r} = m_1$ and the roots of

$$m^{p-1} - \sum_{j=1}^{p-1} \alpha_{p+j} m^{p-1-j} = 0$$

are $m_2, m_3, \dots, m_p$.

Consider the Gram-Schmidt orthogonalization procedure to reparametrize model (B.1) and the equivalent model (B.4). Given n observations $(n > p + r)$, let

$$x_{t1n} = \psi_{t1}$$

$$x_{tin} = \psi_{ti} - \sum_{j=1}^{i-1} c_{ijn} \psi_{tj}, \qquad i = 2,3,\ldots,r$$

$$x_{tin} = a\, S_{t+r-i} - S_{t+r-i-1} - \sum_{j=1}^{i-1} c_{ijn} x_{tjn},$$

$$i = r+1,\ r+2,\ \ldots,\ r+p-1$$

$$x_{t,p+r,n} = S_{t-1} - \sum_{j=1}^{p+r-1} c_{p+r,jn} x_{tjn}, \tag{B.5}$$

where the c_{ijn} are the multiple regression coefficients obtained by the least squares regression of ψ_{ti} and $a\, S_{t+r-i} - S_{t+r-i-1}$ on x_{tjn}, $j = 1,2,\ldots,i-1$ and $t = 1,2,\ldots,n$. The $c_{p+r,jn}$ are obtained by the least squares regression of S_{t-1} on x_{tjn}, $j = 1,2,\ldots,p+r-1$. It is understood that $c_{ijn} = 0$ if $\Sigma_{t=1}^{n} x_{tjn}^2 = 0$. Define

$$W_{t1n} = a\, Y_{t-1} - Y_{t-2} - \sum_{j=1}^{r} c_{r+1,jn} x_{tjn},$$

$$W_{tin} = a\, Y_{t-i} - Y_{t-i-1} - \sum_{j=1}^{r} c_{r+i,jn}\, x_{tjn} - \sum_{j=1}^{i-1} c_{r+i,r+j,n}\, W_{tjn},$$

$$i = 2,3,\ldots,p-1$$

$$W_{tpn} = Y_{t-1} - \sum_{j=1}^{r} c_{p+r,jn}\, x_{tjn} - \sum_{j=1}^{p-1} c_{p+r,j+r,n}\, W_{tjn}. \quad \text{(B.6)}$$

Let $\underset{\sim}{A}_n$ be the nonsingular transformation matrix defined by (B.5) and (B.6) so that

$$\underset{\sim}{X}'_{tn} = (x_{t1n}, x_{t2n}, \ldots, x_{trn}, W_{t1n}, \ldots, W_{tpn})'$$

$$= \underset{\sim}{A}_n\, (\psi_{t1}, \psi_{t2}, \ldots, \psi_{tr}, Y_{t-1}, \ldots, Y_{t-p})'.$$

Then,

$$W_{tin} = x_{t,r+i,n} + \sum_{j=1}^{p} a_{r+i,r+j,n}\, u_{t-j}, \quad \text{(B.7)}$$

where a_{ijn} is the (i,j)-th element of $\underset{\sim}{A}_n$, and model (B.1) can be written as

$$Y_t = \underset{\sim}{X}_{tn}\, \underset{\sim}{\theta}_n + e_t, \quad \text{(B.8)}$$

where $\underset{\sim}{\theta}'_n = (\beta_1, \beta_2, \ldots, \beta_r, \alpha_1, \alpha_2, \ldots, \alpha_p)\, \underset{\sim}{A}_n^{-1}$. For the stationary case the asymptotic distribution of the least squares estimator is established in the following theorem.

Theorem B.4. Let model (B.1) hold with $|m_1| < 1$. Let $\{e_t\}$ be a sequence of random variables satisfying (B.2). Considering the parametrization in (B.8), define

$$\hat{\theta}_n = \left(\sum_{t=1}^{n} X'_{tn} X_{tn}\right)^{-1} \sum_{t=1}^{n} X'_{tn} Y_t. \tag{B.9}$$

Should the matrix be singular, the inverse is replaced by the Moore-Penrose generalized inverse. Assume

$$\lim_{n\to\infty} \sup_{1<t<n} \left(\sum_{s=1}^{n} x_{sin}^2\right)^{-1} x_{tin}^2 = 0, \qquad i = 1,2,\ldots,r \tag{B.10}$$

and

$$\lim_{n\to\infty} \sup_{1<t<n} \left(n + \sum_{s=1}^{n} x_{sin}^2\right)^{-1} x_{tin}^2 = 0, \qquad i = r+1,\ r+2,\ \ldots,\ r+p.$$

Let D_n be the diagonal matrix whose elements are the square roots of the diagonal elements of $\Sigma_{t=1}^{n} X'_{tn} X_{tn}$ and define

$$G_n = D_n^{-1} \left(\sum_{t=1}^{n} X'_{tn} X_{tn}\right) D_n^{-1}.$$

Let $G_n^{1/2}$ be the symmetric positive definite square root of G_n. Then,

$$\sigma^{-1} G_n^{1/2} D_n(\hat{\theta}_n - \theta_n) \xrightarrow{L} N(0, I) \quad \text{as} \quad n \to \infty.$$

<u>Proof.</u> (The proof is very similar to the one used by Fuller, Hasza, and Goebel. An outline of the proof is given.) The probability that $|G_n| \neq 0$ converges to one as n increases. We have

$$\underset{\sim}{D}_n(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_n) = \underset{\sim}{G}_n^{-1} \underset{\sim}{D}_n^{-1} \sum_{t=1}^{n} \underset{\sim}{X}_{tn} e_t .$$

Consider

$$\sum_{t=1}^{n} W_{tin}^2 = \sum_{t=1}^{n} (x_{t,r+i,n} + \sum_{j=1}^{p} a_{r+i,r+j,n} u_{t-j})^2 .$$

By the definition of u_t and Theorem A.12,

$$n^{-1} \sum_{t=1}^{n} u_t u_{t-j} \xrightarrow{P} \gamma_u(j), \tag{B.11}$$

where $u_t = 0$ for $t < 0$ and $\gamma_u(j)$ is the covariance function of a stationary autoregressive process with characteristic equation $m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p = 0$. It follows that $[n^{-1} \Sigma_{t=1}^{n} W_{tjn}^2]^{-1}$ is $O_p(1)$. Now,

$$(n + \sum_{t=1}^{n} x_{tjn}^2)^{-2} \operatorname{Var}\{\sum_{t=1}^{n} x_{tjn} u_{t-j}\}$$

$$= (n + \sum_{t=1}^{n} x_{tjn}^2)^{-2} \sum_{t=1}^{n} \sum_{s=1}^{n} x_{tjn} x_{sjn} E\{u_{t-j} u_{s-j}\}$$

(B.12)

and the right side of this equation converges to zero because $|E\{u_t u_j\}|$ is bounded by a multiple of $\lambda^{|t-j|}$ for some $0 < \lambda < 1$. Therefore, for $j = 1, 2, \ldots, p$

$$[\sum_{t=1}^{n} E\{W_{tjn}^2\}]^{-1/2} [\sum_{t=1}^{n} W_{tjn}^2]^{1/2} \xrightarrow{P} 1.$$

Let

$$\underset{\sim}{H}_n = [E\{\underset{\sim}{D}_n^2\}]^{-1/2} [\sum_{t=1}^{n} \underset{\sim}{x}'_{tn} \underset{\sim}{x}_{tn} + n \underset{\sim}{F}_n][E\{\underset{\sim}{D}_n^2\}]^{-1/2},$$

$$\underset{\sim}{x}_{tn} = (x_{t1n}, x_{t2n}, \ldots, x_{t,r+p,n}),$$

$$\underset{\sim}{F}_n = \underset{\sim}{A}_n \begin{pmatrix} \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{\Gamma}_n \end{pmatrix} \underset{\sim}{A}'_n,$$

$$\underset{\sim}{\Gamma}_n = E[n^{-1} \sum_{t=1}^{n} (u_{t-1}, u_{t-2}, \ldots, u_{t-p})' (u_{t-1}, u_{t-2}, \ldots, u_{t-p})].$$

Note that $\underset{\sim}{H}_n^{-1}$ is well-defined because $\underset{\sim}{\Gamma}_n$ is positive definite and that $\operatorname{plim}_{n\to\infty} (\underset{\sim}{G}_n^{-1/2} - \underset{\sim}{H}_n^{-1/2}) = \underset{\sim}{0}$. Consider the linear combination

$$\underset{\sim}{\eta}' \underset{\sim}{H}_n^{-1/2} \underset{\sim}{D}_n^{-1} \sum_{t=1}^{n} \underset{\sim}{X}'_{tn} e_t$$

where $\underset{\sim}{\eta}$ is a vector of arbitrary real numbers such that $\underset{\sim}{\eta}' \underset{\sim}{\eta} \neq 0$. Because

$$E\{\sum_{t=1}^{n} W_{tjn} e_t \sum_{s=2}^{n} W_{sin} e_s\} = \sigma^2 \sum_{t=1}^{n} E\{W_{tjn} W_{tin}\},$$

we can write

$$\underset{\sim}{\eta}' \underset{\sim}{H}_n^{-1/2} \underset{\sim}{D}_n^{-1} \sum_{t=1}^{n} \underset{\sim}{X}_{tn}' e_t = \sum_{t=1}^{n} Z_{tn} + o_p(1)$$

$$= S_{nn} + o_p(1),$$

where

$$S_{nn} = \sum_{t=1}^{n} Z_{tn},$$

$$Z_{tn} = (g_{tn} + v_{tn}) e_t,$$

$$g_{tn} = \sum_{i=1}^{r} \eta_i \Big(\sum_{s=1}^{n} x_{sin}^2\Big)^{-1/2} x_{tin}$$

$$+ \sum_{j=1}^{p} \eta_{r+j,n} \Big(\sum_{s=1}^{n} E\{W_{sjn}^2\}\Big)^{-1/2} x_{t,r+j,n},$$

$$v_{tn} = \sum_{i=1}^{p} \eta_{r+i,n} \Big(\sum_{s=1}^{n} E\{W_{sin}^2\}^{-1/2}\Big)\Big(\sum_{j=1}^{p} a_{r+i,r+j,n} u_{t-j}\Big)$$

$$\eta_{r+i,n} = \sum_{j=1}^{p} h_{r+j,r+i,n}^{(-1/2)} \eta_{r+j},$$

and $h_{jin}^{(-1/2)}$ is the (j,i)-th element of $\underset{\sim}{H}_n^{-1/2}$. Observe that $\{g_{tn}: t = 1,2,\ldots,n\}$ is fixed and that the v_{tn} are fixed linear combinations of $\{u_{t-j}: j = 1,2,\ldots,p\}$ for a particular n.

Because $Y_0, Y_{-1}, \ldots, Y_{-p+1}$ are fixed, the sigma field B_{kn}

generated by $(Z_{1n}, Z_{2n}, \ldots, Z_{kn})$ is the sigma field generated by $(e_1, e_2, \ldots, e_k)$. Therefore,

$$E\{Z_{tn}^2 \mid B_{t-1,n}\} = \delta_{tn}^2 = (g_{tn} + v_{tn})^2 \sigma^2.$$

Let

$$V_{nn}^2 = \sum_{t=1}^{n} \delta_{tn}^2$$

and

$$s_{nn}^2 = \sum_{t=1}^{n} E(\delta_{tn}^2) = \sum_{t=1}^{n} [g_{tn}^2 + E(v_{tn}^2)] \sigma^2$$

Using (B.10) and (B.11),

$$s_{nn}^{-2} V_{nn}^2 \xrightarrow{P} 1.$$

To apply the results of Scott (1973), it is sufficient to show that $\{Z_{tn}\}$ satisfies the Lindeberg type condition

$$s_{nn}^{-2} \sum_{t=1}^{n} E\{Z_{tn}^2 I(|Z_{tn}| > \varepsilon s_{nn})\} \longrightarrow 0.$$

Note that

$$s_{nn}^{-2} \sum_{t=1}^{n} E\{Z_{tn}^2 I(|Z_{tn}| > \varepsilon s_{nn})\}$$

$$< s_{nn}^{-2-\nu} \, \varepsilon^{-\nu} \sum_{t=1}^{n} E\{ |Z_{tn}|^{2+\nu} \}$$

By the definitions of $\underset{\sim}{H}_n$ and $\underset{\sim}{H}_n^{-1/2}$, $s_{nn}^2 = (\underset{\sim}{\eta}' \, \underset{\sim}{\eta}) \, \sigma^2$. Now,

$$E\{ |Z_{tn}|^{2+\nu} \} = E\{ |g_{tn} + v_{tn}|^{2+\nu} \, |e_t|^{2+\nu} \}$$

$$< 2^{2+\nu} \, [E\{ |g_{tn}|^{2+\nu} \, |e_t|^{2+\nu} \} + E\{ |v_{tn}|^{2+\nu} \, |e_t|^{2+\nu} \}]$$

$$< 2^{2+\nu} [\, |g_{tn}|^{2+\nu} \, E\{|e_{tn}|^{4+2\nu}\}^{1/2} + \{E|v_{tn}|^{4+2\nu}\}^{1/2} \, \{E|e_t|^{4+2\nu}\}^{1/2} \,]$$

$$< L^{1/2} \, 2^{2+\nu} \, [|g_{tn}|^{2+\nu} + (E\{|v_{tn}|^{4+2\nu}\})^{1/2} \,]$$

Since we can write

$$v_{tn} = \sum_{j=1}^{p} \xi_{tjn} \, u_{t-j},$$

where $\xi_{tjn} = O(n^{-1/2})$ it follows that

$[E\{|v_{tn}|^{4+2\nu}\}]^{1/2} = O(n^{-1-\nu/2})$. Note,

$$s_{nn}^{-2-\nu} \sum_{t=1}^{n} |g_{tn}|^{2+\nu} < \sigma^{-2} \, s_{nn}^{-\nu} \sup_{1<t<n} |g_{tn}|^{\nu}$$

and

$$\sup_{1<t<n} |g_{tn}|^2 < \sup_{1<t<n} (r+p)^2 \left[\sum_{i=1}^{r} \left(\sum_{s=1}^{n} x_{sin}^2 \right)^{-1} x_{tin}^2 \eta_i \right.$$

$$+ \sum_{j=1}^{p} \left(\sum_{s=1}^{n} E\{W_{sjn}^2\} \right)^{-1}$$

$$\left. x_{t,r+j,n}^2 \eta_{r+j,n}^2 \right],$$

which tends to zero by (B.10). It follows that $s_{nn}^{-1} S_{nn}$ converges in distribution to a $N(0,1)$ random variable. The conclusion follows because η was arbitrary. □

For a particular n, the elements of θ_n are fixed linear combination of the parameters α and β. Therefore, for large samples, the above theorem justifies the use of the ordinary regression statistics in making inferential statements regarding the parameters of the model (B.1).

Now we consider the case where $m_1 = 1$. Consider the model (B.1). We consider two cases of practical interest.

(a) $\psi_{t1} \equiv 1$ and

(b) $\psi_{t1} \equiv 1$, $\psi_{t2} = t$.

We introduce an additional modification of the parametrization of (B.8), letting

$$W_{tpn}^{+} = W_{tpn} - n^{-1} \sum_{s=1}^{n} W_{spn} = W_{tpn} - \overline{W}_{.pn}$$

for case (a) and

$$W^{+}_{tpn} = W_{tpn} - \overline{W}_{.pn} - b_{wn}[t - \tfrac{1}{2}(n+1)]$$

for case (b), where $\overline{W}_{.pn}$ is the sample mean of W_{tpn} and b_{wn} is the least squares coefficient obtained by regressing W_{tpn} on $t - \frac{1}{2}(n+1)$. This transformation differs from that used in Theorem (B.4) because the coefficients of x_{t1n} and x_{t2n} defining W^{+}_{tpn} are functions of the random variables $\{u_t\}_{t=1}^{n}$.

Let $\underset{\sim}{A}_{(u)n}$ be the matrix whose first $r + p - 1$ rows are the first $r + p - 1$ rows of $\underset{\sim}{A}_n$ and whose last row $\underset{\sim}{A}_{(u)r+p,.,n}$ is given by the above transformation so that

$$W^{+}_{tpn} = \underset{\sim}{A}_{(u)r+p,.,n}(\psi_{t1}, \psi_{t2}, \ldots, \psi_{tr}, Y_{t-1}, \ldots, Y_{t-p})'.$$

The transformed regression equation is

$$Y_t = \underset{\sim}{X}_{(u)tn}\, \underset{\sim}{\theta}_{(u)tn} + e_t,$$

where

$$\underset{\sim}{X}'_{(u)tn} = \underset{\sim}{A}_{(u)n}(\psi_{t1}, \psi_{t2}, \ldots, \psi_{tn}, Y_{t-1}, \ldots, Y_{t-p})'$$

$$\underset{\sim}{\theta}'_{(u)tn} = (\beta_1, \beta_2, \ldots, \beta_r, \alpha_1, \alpha_2, \ldots, \alpha_p)\, \underset{\sim}{A}^{-1}_{(u)n}.$$

The asymptotic distributions of the least squares estimators are given in the following theorem. Proofs of the theorems parallel the proofs given by Fuller, Hasza, and Goebel with the modifications used in the proof of Theorem B.4 and the results of Chapter II.

Theorem B.5. Let model (B.1) hold with $m_1 = 1$ and $m_2, m_3, \ldots, m_p$ less than one in absolute value, where the m_i are the roots of $m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p = 0$. Let $\{e_t\}$ satisfy the conditions (B.2). Let

$$\hat{\underset{\sim}{\theta}}_{(u)n} = [\sum_{t=1}^{n} \underset{\sim}{X}'_{(u)tn} \underset{\sim}{X}_{(u)tn}]^{-1} [\sum_{t=1}^{n} \underset{\sim}{X}'_{(u)tn} Y_t].$$

Let $\underset{\sim}{D}_{(u)n}$ be the diagonal matrix whose elements are the square roots of the diagonal elements $\sum_{t=1}^{n} \underset{\sim}{X}'_{(u)tn} \underset{\sim}{X}_{(u)tn}$. Let

$$\underset{\sim}{G}_{(u)n} = \underset{\sim}{D}^{-1}_{(u)tn} \sum_{t=1}^{n} \underset{\sim}{X}'_{(u)tn} \underset{\sim}{X}_{(u)tn} \underset{\sim}{D}^{-1}_{(u)n},$$

$$\hat{\underset{\sim}{\delta}}_{(u)n} = \sigma^{-1} \underset{\sim}{G}^{1/2}_{(u)n} \underset{\sim}{D}_{(u)n} (\hat{\underset{\sim}{\theta}}_{(u)n} - \underset{\sim}{\theta}_{(u)n}),$$

where $\underset{\sim}{G}^{1/2}_{(u)n}$ is the positive definite square root of $G_{(u)n}$. Assume (B.10) is satisfied and

$$\lim_{n\to\infty} (n^2 \sum_{t=1}^{n} x^2_{tin})^{-1} \sum_{t=1}^{n} \sum_{h=0}^{n-t} t\, x_{tin}\, x_{t+h,in} = 0 \qquad \text{(B.13)}$$

for $i = 2, 3, \ldots, r$ with case (a) and for $i = 3, 4, \ldots, r$ with case (b). Assume that for $i = r + 1, r + 2, \ldots, r + p$,

$$\lim_{n\to\infty} n^{-2} \sum_{t=1}^{n} x_{tin}^2 = 0,$$

$$\lim_{n\to\infty} n^{-2} (n + \sum_{t=1}^{n} x_{tin}^2)^{-1} \sum_{t=1}^{n} \sum_{h=0}^{n-t} t\, x_{tin}\, x_{t+h,in} = 0.$$

Then the last element of $\hat{\underset{\sim}{\delta}}_n$ converges in distribution to the statistic $\hat{\tau}_\mu$ for case (a) and to $\hat{\tau}_\tau$ for case (b), where $\hat{\tau}_\mu$ and $\hat{\tau}_\tau$ are characterized in Chapter II. The limiting distribution of the remaining $r + p - 1$ elements of $\hat{\underset{\sim}{\delta}}_n$ is normal with zero mean and identity covariance matrix for both cases.

Proof. Following Fuller, Hasza and Goebel (1981), we have

$$u_t = \sum_{i=1}^{t} \sum_{j=0}^{\infty} v_j^* e_i - \sum_{i=1}^{t} \sum_{j=t-i+1}^{\infty} v_j^* e_i$$

$$= A_t + B_t ,$$

where the v_i^* satisfy the homogeneous difference equation

$$v_i^* - \beta_{q+1} v_{i-1}^* - \cdots - \beta_{q+p-1} v_{i-p+1}^* = 0 ,$$

with the initial conditions $v_0 = 1$ and $v_i = 0$, for $i < 0$. It follows that $\Sigma_{j=0}^{\infty} |v_j^*| < \infty$ and $\Sigma_{j=t-i+1}^{\infty} |v_j^*| < M\, \lambda^{t-i+1}$ for some $M < \infty$ and some $0 < \lambda < 1$.

Let A_{tn}^+ and B_{tn}^+ denote the portions of A_t and B_t that are orthogonal to ψ_{t1} under (a) and orthogonal to ψ_{t1} and ψ_{t2} under (b). Following Fuller, Hasza and Goebel (1981), we get that

$$\sum_{t=1}^{n} (B_{tn}^{+})^2 = O_p(n) ,$$

$$\sum_{t=1}^{n} x_{t,p+q,n} B_{tn}^{+} = o_p(n^{3/2}) ,$$

$$\sum_{t=1}^{n} (W_{tpn}^{+})^2 = \sum_{t=1}^{n} [x_{t,p+q,n} + A_{tn}^{+}]^2 + o_p(n^{3/2})$$

and that

$$\mathrm{Var}\{\sum_{t=1}^{n} u_{t-m} x_{tin}\} \le 2\sigma^2 (\sum_{j=0}^{\infty} |v_j^*|)^2 \sum_{t=1}^{n} \sum_{h=0}^{t-1} t\, x_{tin} x_{t+h,in})$$

for $m = 1,2,\ldots, p$. By (B.13), the first q elements of the last row of $G_{(u)n}$ converge in probability to zero because $[\Sigma_{t=1}^{n} (A_{tn}^{+})^2]^{-1} = O_p(n^{-2})$. For $j = 1,2,\ldots, p-1$

$$W_{tjn} = x_{t,q+j,n} + \sum_{i=1}^{j} a_{q+j,q+i,n} \sum_{j=0}^{t-i} v_j^* e_{t-j} .$$

Now we will establish that

$$\mathrm{plim}[\sum_{t=1}^{n} (W_{tpn}^{+})^2 \sum_{t=1}^{n} W_{tjn}^2]^{-1/2} \sum_{t=1}^{n} W_{tpn}^{+} W_{tjn} = 0$$

for $j = 1,2,\ldots, p-1$. Note that

$$\sum_{t=1}^{n} W_{tpn}^{+} W_{tjn} = \sum_{t=1}^{n} x_{t,q+j,n} A_t$$

$$+ \sum_{t=1}^{n} \sum_{i=1}^{j} a_{q+j,q+i,n} \left(\sum_{k=0}^{t-i} v_k^{*} e_{t-k} \right) (x_{t,p+q,n} + A_{tn}^{+})$$

$$+ \sum_{t=1}^{n} W_{tjn} B_{tn}^{+} .$$

Since

$$E\left[\sum_{t=1}^{n} x_{t,q+j,n} A_t\right]^2 < 2\sigma^2 \left(\sum_{j=0}^{\infty} |v_j^{*}|\right)^2 \sum_{t=1}^{n} \sum_{h=0}^{t-1} t \, x_{t,q+j,n} x_{t+h,q+j,n}$$

and

$$E\left[\sum_{t=1}^{n} \left(\sum_{k=0}^{t-i} v_k^{*} e_{t-k}\right) x_{t,p+q,n}\right]^2$$

$$< n \sum_{t=1}^{n} x_{t,p+q,n}^2 \left(\sum_{k=0}^{\infty} v_k^{*2}\right)$$

by (B.13) and (B.10), we get

$$\left[\sum_{t=1}^{n} (W_{tpn}^{+})^2 \sum_{t=1}^{n} W_{tjn}^2\right]^{-1/2} \sum_{t=1}^{n} W_{tpn}^{+} W_{tjn}$$

$$= \left[\sum_{t=1}^{n} (W_{tpn}^{+})^2 \sum_{t=1}^{n} W_{tjn}^2\right]^{-1/2} \left[\sum_{t=1}^{n} \sum_{i=1}^{n} a_{q+j,q+i,n} \left(\sum_{k=0}^{t-i} v_k^{*} e_{t-k}\right) A_{tn}^{+}\right]$$

$$+ o_p(1) .$$

Since

$$n^{-1} \sum_{t=1}^{n} A_t = O_p(n^{1/2}) \ ,$$

$$[\sum_{t=1}^{n} \{t - \tfrac{1}{2}(n+1)\}^2]^{-1} \sum_{t=1}^{n} \{t - \tfrac{1}{2}(n+1)\}A_t = O_p(n^{-1/2}) \ ,$$

$$\sum_{t=1}^{n} \sum_{k=0}^{t-i} v_k^* e_{t-k} = O_p(n^{1/2}) \ ,$$

and

$$\sum_{t=1}^{n} \{t - \tfrac{1}{2}(n+1)\} \sum_{k=0}^{t-i} v_k^* e_{t-i} = O_p(n^{3/2}) \ ,$$

we get

$$[\sum_{t=1}^{n} (W_{tpn}^{+})^2 \sum_{t=1}^{n} W_{tjn}^2]^{-1/2} \sum_{t=1}^{n} W_{tpn}^{+} W_{tjn}$$

$$= [\sum_{t=1}^{n} (W_{tpn}^{+})^2 \sum_{t=1}^{n} W_{tjn}^2]^{-1/2}$$

$$[(\sum_{\ell=0}^{\infty} v_\ell^*) \sum_{i=1}^{j} a_{q+i,q+j,n} \sum_{t=1}^{n} (\sum_{k=0}^{t-i} v_k^* e_{t-k})(\sum_{s=1}^{t} e_s)] + o_p(1) \ .$$

Now consider

$$\sum_{t=1}^{n} (\sum_{k=0}^{t-i} v_k^* e_{t-k})(\sum_{s=1}^{t} e_s)$$

$$= \sum_{t=1}^{n} \sum_{j=i}^{t} v_{t-j}^* e_j^2 + \sum_{t=1}^{n} \sum_{j=i}^{t} v_{t-j}^* \sum_{\substack{k=1 \\ \neq j}}^{t} e_k e_j$$

$$= \sum_{t=1}^{n} \sum_{j=i}^{t} v_{t-j}^{*} e_j A_{tj} + O_p(n) ,$$

where

$$A_{tj} = \sum_{k=1}^{t} e_k - e_j .$$

Now

$$\mathrm{Var}\{ \sum_{t=1}^{n} \sum_{j=i}^{t} v_{t-j}^{*} e_j A_{tj} \}$$

$$< 2 \sum_{t=1}^{n} \sum_{h=0}^{t-1} \mathrm{Cov}(\sum_{j=i}^{t} v_{t-j}^{*} e_j A_{tj} , \sum_{j=i}^{t+h} v_{t+h-j}^{*} e_j A_{t+h,j})$$

$$< 2 \sum_{t=1}^{n} \sum_{h=0}^{t-1} \sum_{j=i}^{t} \sum_{\ell=1}^{t+h} v_{t-j}^{*} v_{t-\ell+h}^{*} \mathrm{Cov}(e_j A_{tj}, e_\ell A_{t+h,\ell})$$

$$< 2 \sum_{t=1}^{n} \sum_{h=0}^{t-1} \sum_{j=i}^{t} \sum_{\ell=1}^{t} v_{t-j}^{*} v_{t-\ell+h}^{*} \sigma^4$$

$$+ 2 \sum_{t=1}^{n} \sum_{h=0}^{t-1} \sum_{j=i}^{t} v_{t-j}^{*} v_{t-j+h}^{*} (t-2) \sigma^4$$

$$= O(n^2) ,$$

where we have used $E\{e_k^2 e_i^2\} = \sigma^4$, $k \neq i$. Therefore,

$$\mathrm{plim}[\sum_{t=1}^{n} (W_{tpn}^{+})^2 \sum_{t=1}^{n} W_{tjn}^2]^{-1/2} [\sum_{t=1}^{n} (W_{tpn}^{+}) W_{tjn}] = 0$$

for $j = 1,2,\ldots, p-1$, and the first $q + p - 1$ coefficients of the last row of $\underset{\sim}{G}_{(u)n}$ converge in probability to zero.

By Theorem B.4, the limiting distribution of the vector composed of the first $q + p - 1$ elements of $\underset{\sim}{D}^{-1}_{u(n)} \Sigma^n_{t=1} \underset{\sim}{X}'_{(u)tn} e_t$ is multivariate normal. Also, the last element of $\underset{\sim}{D}^{-1}_{(u)n} \Sigma^n_{t=1} \underset{\sim}{X}'_{(u)tn} e_t$ is

$$\frac{\Sigma^n_{t=1} W^+_{tpn} e_t}{[\Sigma^n_{t=1} (W^+_{tpn})^2]^{1/2}} = \frac{n^{-1} \Sigma^n_{t=1} u^+_{t-1} e_t}{n^{-2} \Sigma^n_{t=1} (u^+_{t-1})^2} + o_p(1) ,$$

where $u^+_t = A^+_t + B^+_t$. From the results of Chapter II, the limiting distribution of this statistic is that of the $\hat{\tau}_\mu$-statistic for case (a) and that of $\hat{\tau}_\tau$ for case (b). □

Similarly, Theorem 3 of Fuller, Hasza and Goebel can be extended to the case where $\{e_t\}$ is a sequence of martingale differences satisfying the conditions (B.2) and the proof is not included.

Theorem B.6. Let model (B.1) hold with $m_1 = 1$ and $m_2, m_3, \ldots, m_p$ less than one in absolute value. Let $\{e_t\}$ satisfy the conditions in (B.2). Assume (B.10) is satisfied and

$$\lim_{n\to\infty} n^{-2} \sum_{t=1}^{n} x^2_{t,p+r,n} = \infty \tag{B.14}$$

and

$$\lim_{n\to\infty} [\sum_{t=1}^{n} x^2_{tin} \sum_{t=1}^{n} x^2_{t,r+p,n}]^{-1} \sum_{t=1}^{n} \sum_{h=0}^{n-t} t\, x_{tin} x_{t+h,in} = 0 \tag{B.15}$$

for $i = 2, \ldots, r + p - 1$. If $\underset{\sim}{D}_n$, $\hat{\underset{\sim}{\theta}}_n$, $\underset{\sim}{\theta}_n$ and $\underset{\sim}{G}_n^{1/2}$ are as defined in Theorem B.4, then

$$\sigma^{-1} \underset{\sim}{G}_n^{1/2} \underset{\sim}{D}_n(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}_n) \xrightarrow{L} N(\underset{\sim}{0}, \underset{\sim}{I}) \quad \text{as} \quad n \to \infty.$$