Estimation for the autoregressive moving average model with a unit root

Dongwan Shin
Iowa State University
INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Estimation for the autoregressive moving average model with a unit root

Shin, Dongwan, Ph.D.
Iowa State University, 1990
Estimation for the autoregressive moving average model with a unit root

by

Dongwan Shin

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

1990
# TABLE OF CONTENTS

LIST OF ASSUMPTIONS, DEFINITIONS, LEMMAS, AND THEOREMS

## 1. INTRODUCTION

1.1. Model Description 1

1.2. Literature Review
   1.2.1. Literature on the Consistency in a Stationary Model 3
   1.2.2. Literature on the Unit Root Test Problem 4

1.3. Assumptions and Definitions of Estimators 12

1.4. Summary of Main Results 16

## 2. PRELIMINARY RESULTS

2.1. Literature on the Properties of $\Gamma_n = \Gamma_n(\theta)$ 18

2.2. Properties of $\Gamma_n$ 22

2.3. The Matrix Norm 31

## 3. CONSISTENCY

3.1. A Sufficient Condition for Strong Consistency 45

3.2. Order with Probability One 49

3.3. Strong Consistency in a Stationary Model 52

3.4. Orders of Quadratic Forms in $Z$ 62

3.5. Limiting Behavior of $Y_1' \Gamma_n^{-1} Z$ and $Y_1' Y_1$ 78

3.6. Strong Consistency 91
4. LIMITING DISTRIBUTION
   4.1. Derivatives of $Q_n$ 101
   4.2. Limiting Distribution of Estimators 113
   4.3. Limiting distribution of Regression $t$-statistics 118

5. MODEL WITH AN INTERCEPT
   5.1. Model Description 125
   5.2. Weak Consistency 129
   5.3. Derivatives of $Q_n$ 150
   5.4. Limiting Distribution of Estimators 153
   5.5. Limiting Distribution of Regression $t$-statistics 160

6. SIMULATION STUDY
   6.1. Unit Root Test in a Model without Intercept 164
   6.2. Unit Root Test in a Model with Intercept 166
   6.3. A Test for the Moving Average Parameter 168
   6.4. Computational Aspect 176

7. APPENDIX A PRELIMINARY PROBABILITY THEORY 179
8. APPENDIX B  TECHNICAL DETAILS  
8.1. Appendix for Chapter 2  
8.2. Appendix for Chapter 3  
8.3. Appendix for Chapter 4  
8.4. Appendix for Chapter 5  

9. BIBLIOGRAPHY  

10. ACKNOWLEDGEMENTS
<p>| Assumption 1.1 | Assumptions about a nonstationary model with unit root | 12 |
| Assumption 1.2 | Assumptions about a stationary and invertible autoregressive moving average process | 13 |
| Definition 1.3 | Definition of the least squares estimator | 13 |
| Definition 1.4 | Definition of the maximum likelihood estimator | 15 |
| Definition 1.5 | Definition of the ordinary least squares estimator | 16 |
| Definition 2.1 | Definition of exponential decline | 23 |
| Definition 2.2 | Definition of polynomial | 24 |
| Theorem 2.3 | Solution to the difference equation | 24 |
| Theorem 2.4 | Uniformly exponentially declining property of ( v_j ) satisfying the difference equation | 26 |
| Corollary 2.5 | Uniformly exponentially declining properties of the derivatives of ( f_j ) satisfying | 28 |
| Theorem 2.6 | Uniformly exponentially declining properties of ( d_j ), ( m_{jk} ), ( v_j ) and the derivatives of them | 30 |
| Definition 2.7 | Definition of matrix norm | 31 |
| Lemma 2.8 | Uniform boundedness of the matrix norm of ((A^{-1} + M_n' M_n)^{-1}) | 32 |
| Theorem 2.9 | Gershgorin's theorem | 34 |</p>
<table>
<thead>
<tr>
<th>Theorem/Know/Def</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.10</td>
<td>Uniform boundedness of matrix norms of $\Gamma_n$, $D_n^\prime D_n$, and $\Gamma_n^{-1}$</td>
</tr>
<tr>
<td>2.11</td>
<td>Uniform boundedness of $\det \Gamma_n$</td>
</tr>
<tr>
<td></td>
<td>$\det \Gamma_n &gt; 1$, $\lambda \ln(\det \Gamma_n)^{1/n} \to 1$ as $n \to \infty$.</td>
</tr>
<tr>
<td>2.12</td>
<td>Uniformly exponentially declining property of $\int_0^1 e^{i\omega t} A(e^{i\omega}) d\omega$</td>
</tr>
<tr>
<td>2.13</td>
<td>Uniform exponentially declining property of derivatives of covariance function</td>
</tr>
<tr>
<td>2.14</td>
<td>Uniform boundedness of the matrix norms of the derivatives of covariance matrix and its derivative</td>
</tr>
<tr>
<td>3.1</td>
<td>Wu's sufficient condition of strong consistency</td>
</tr>
<tr>
<td>3.2</td>
<td>A sufficient condition of almost sure convergence of a martingale</td>
</tr>
<tr>
<td>3.3</td>
<td>$n^{-1}</td>
</tr>
<tr>
<td>3.4</td>
<td>Definition of order with probability one</td>
</tr>
<tr>
<td>3.5</td>
<td>$\sup_{\theta}</td>
</tr>
<tr>
<td>3.6</td>
<td>Verification of Wu's condition for the stationary and invertible autoregressive moving average process</td>
</tr>
<tr>
<td>3.7</td>
<td>Convergence of $n^{-1}Z_1 \Gamma_n^{-1}(\theta_n)Z$ and $n^{-1}Z_1 D_n^\prime(\theta_n)D_n(\theta_n)Z$, when $\theta_n$ is consistent</td>
</tr>
<tr>
<td>3.8</td>
<td>Strong consistency of parameter estimates in a stationary and invertible autoregressive moving average process</td>
</tr>
<tr>
<td>3.9</td>
<td>$\sum_{i=1}^{n} \sum_{j=1}^{i-1} \lambda^{i-j} e_i e_j = o_{wp}(n^{1/2+\epsilon}), \lambda \in (0,1), \epsilon &gt; 0$</td>
</tr>
</tbody>
</table>
Lemma 3.10  Orders of quadratic forms in $Z$

Lemma 3.11  Orders of $\sup_{\theta} |Y_1' \Gamma^{-1}_n Z|$ and $\sup_{\theta} |Y_1' D_n Z|$

Lemma 3.12  $\liminf Y_1' Y_1 n^{-2} \log \log n > 0$ a.s. and

$$Y_1' Y_1 - c_0^2 \frac{n-1}{n} \sum_{t=1}^{n} W_t^2 = O_p(n^{3/2})$$

Corollary 3.13  Limiting distribution of $Y_1' Y_1 / n^2$

Lemma 3.14  Verification of Wu's condition in the nonstationary model (1.1)

Theorem 3.15  Strong consistency of parameter estimates in the nonstationary model (1.1)

Lemma 3.16  Order of $(\hat{\rho} - 1)$, $(\rho - 1)$, and $(\hat{\rho} - 1)$ is $o_p(n^{-1/2})$

Lemma 4.1  Orders of the third order derivatives of $Q_n$

Lemma 4.2  Orders of $Q_{p\theta}^0$ and $Q_{p\theta}^*$

Lemma 4.3  Limiting distribution of $(n^{-2} Q_{pp}, n^{-1} Q_{p\theta}, n^{-1/2} Q_{\theta\theta})$

and probability limits of $n^{-1} Q_{\theta\theta}$ and $n^{-3/2} Q_{p\theta}$

Theorem 4.4  Limiting distribution of estimators in the nonstationary model

Theorem 4.5  Limiting distribution of regression $t$-statistics in the stationary model

Corollary 4.6  Limiting distribution of $V(Y; \theta, \rho)$
Assumption 5.1 Assumptions about a model with intercept and a unit root 127
Definition 5.2 Definition of the least squares estimator in the model with intercept 127
Definition 5.3 Definition of the maximum likelihood estimator in the model with intercept 127
Definition 5.4 Definition of the ordinary least squares estimator in the model with intercept 127
Lemma 5.5 Wu's sufficient condition for weak consistency 128
Theorem 5.6 Approximations of $Z'\Gamma_n^{-1}(\theta)Z$, $Y_1'\Gamma_n^{-1}(\theta)Y_1$, $1'\Gamma_n^{-1}(\theta)1$, $Y_1'\Gamma_n^{-1}(\theta)Z$, $1'\Gamma_n^{-1}(\theta)Y_1$, $1'\Gamma_n^{-1}(\theta)Z$ 129
Lemma 5.7 Verification of Wu's condition in the model with intercept 131
Theorem 5.8 Weak consistency of parameter estimates 139
Lemma 5.9 Order of $\hat{\rho} - 1$, $\hat{\rho} - 1$, $\hat{\rho} - 1$ is $o_p(n^{-1/2})$ 147
Lemma 5.10 Orders of the third order derivatives of $Q_n$ 148
Lemma 5.11 Orders of $Q^O_{\rho\theta}$, $Q^O_{\mu\theta}$ 153
Lemma 5.12 Limiting distribution of

\[
(n^{-2}Q^O_{\rho\rho}, n^{-3/2}Q^O_{\rho\mu}, n^{-1}Q^O_{\rho}, n^{1/2}Q^O_{\mu}, n^{-1/2}Q^O_{\theta})
\]

and probability limits of

\[
n^{-1}Q^O_{\theta\theta} \text{ and } n^{-3/2}Q^O_{\rho\theta}
\]

Theorem 5.13 Limiting distribution of parameter estimates in the model with intercept 155
Theorem 5.14 Limiting distribution of regression t-statistics in the model with intercept 162
Theorem A.1  The strong law of large numbers  179
Theorem A.2  The central limit theorem  179
Theorem A.3  The law of the iterated logarithm  179
Theorem A.4  If \( \sum_{n=1}^{\infty} |X_n| < \infty \) then \( \sum_{n=1}^{\infty} X_n \) converges a.s.  180
Example A.5  If \( \sum_{n=1}^{\infty} |a_n| < \infty \) then \( \sum_{n=1}^{\infty} a_n z_n \) converges a.s.  180
Definition A.6  Definition of martingale  180
Theorem A.7  Martingale convergence theorem  181
Theorem A.8  Donsker's invariance principle  181
Theorem A.9  The continuous mapping theorem  182
Example A.10  Limiting distribution of \( \left( n^{-2} \sum_{i=1}^{n} W_{i-1}^2, n^{-1} \sum_{i=1}^{n} W_{i-1} e_i \right) \)  182
Theorem A.11  Strassen's law of the iterated logarithm  182
Corollary A.12  Strassen's law of the iterated logarithm, a different presentation  184
Example A.13  Law of the iterated logarithm for \( \sum_{i=1}^{n} W_{i-1} e_i, \sum_{i=1}^{n} W_{i-1}, \sum_{i=1}^{n} W_{i-1}^2 \)  184
Definition A.14  Definition of the strong \( \alpha \)-mixing  185
Lemma A.15  Kronecker's lemma  186
1. INTRODUCTION

1.1. Model Description

We consider the following time series model.

\[ y_t = \rho y_{t-1} + z_t, \]

\[ z_t + \alpha_1 z_{t-1} + \cdots + \alpha_p z_{t-p} = e_t + \beta_1 e_{t-1} + \cdots + \beta_q e_{t-q}, \quad t = 1, 2, \ldots, n, \]

where \( y_t, t = 0, 1, \ldots, n \), are the observations, \( \{z_t\} \) is an autoregressive moving average process of order \((p, q)\) for some nonnegative integers \( p \) and \( q \), and \( \{e_t\} \) is a sequence of iid \((0, \sigma^2)\) random variables. We shall be interested in estimation of the parameter

\[ (\rho, \alpha', \beta', \sigma) = (\rho, \alpha_1, \alpha_2, \ldots, \alpha_p, \beta_1, \beta_2, \ldots, \beta_q, \sigma) \]

given a sample of \( n+1 \) observations on \( y_t \). We let

\[ A(m) = m^p + \alpha_1 m^{p-1} + \cdots + \alpha_{p-1} m + \alpha_p \]

and

\[ B(m) = m^q + \beta_1 m^{q-1} + \cdots + \beta_{q-1} m + \beta_q \]

be the characteristic equations associated with \( \{z_t\} \). We assume stationarity for process \( \{z_t\} \). Here we define stationarity of a sequence of random variables. A
sequence of random variable \( \{x_t\} \) is said to be covariance stationary if for every positive integers \( t \) and \( h \)

\[
E(x_t) = E(x_1)
\]

and

\[
\text{Cov}(x_t, x_{t+h}) = \text{Cov}(x_1, x_{1+h}).
\]

Throughout this book we use the terminology 'stationary' for 'covariance stationary'. The process \( \{z_t\} \) defined in the second line of (1.1) is stationary if all the roots of the characteristic equation \( A(m) = 0 \) in (1.2) lie inside the unit circle. Next we define invertibility of the process \( \{z_t\} \). The process \( \{z_t\} \) in the second line of (1.1) is said to be invertible if all the roots of the characteristic equation \( B(m) = 0 \) in (1.3) lie inside the unit circle. If \( \{z_t\} \) is invertible there is an absolutely summable sequence \( \{d_j\} \) such that

\[
e_t = \sum_{j=0}^{\infty} d_j z_{t-j} \quad \text{a.s.}
\]

We investigate the limiting behavior of estimators under the stationarity and invertibility condition on \( \{z_t\} \), an identifiability condition of \( \theta \), and the assumption that \( \rho = 1 \). When \( \rho = 1 \), the process \( \{y_t\} \) is nonstationary.
1.2. Literature Review

1.2.1. Literature on the Consistency in a Stationary Model

A number of results on the consistency and limiting distribution of estimators of the parameters of the stationary and invertible autoregressive moving average process

\[ z_t + \alpha_1 z_{t-1} + \cdots + \alpha_p z_{t-p} = e_t + \beta_1 e_{t-1} + \cdots + \beta_q e_{t-q} \]  

(1.4)
can be found in the literature. In model (1.4), \( \{z_t\}_{t=1}^n \) is observation. Hannan (1973) considered model (1.4) under the assumptions that \( \{e_t\} \) is an uncorrelated \((0, \sigma^2)\) sequence and \( \{z_t\} \) is stationary and invertible. He obtained the strong consistency and the limiting distribution of the least squares estimator and of the maximum likelihood estimator of \( \theta' = (\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q) \). His proof is based on a frequency domain approximation of where \( Z = (z_1, \ldots, z_n)' \), and \( \Gamma_n(\theta) = \text{Var}\{\{z_1, \ldots, z_n\}'\} \). The results of Hannan are well described in the book by Brockwell and Davis (1987, pp. 365–386). Dunsmuir and Hannan (1976), using a method similar to that of Hannan (1973), generalized Hannan's result to the multivariate autoregressive moving average process. Rissanen and Caines (1979) considered a stationary and invertible multivariate autoregressive moving average process with multivariate normal innovation \( \{e_t\} \). They gave a Kalman filter type recurrence equation for the least squares prediction errors \( \{\tilde{e}_t\} \) used to approximate \( \{e_t\} \). They approximated \( Z' \Gamma_n^{-1}(\theta)Z \) using \( \{\tilde{e}_t\} \). With this approximation, they showed the strong consistency of the maximum likelihood estimator.
1.2.2. Literature on the Unit Root Test Problem

In recent years, testing for a unit root in an autoregressive process has attracted much interest. Dickey and Fuller (1979) considered the autoregressive model

\[ y_t = \rho y_{t-1} + e_t, \quad t = 1, 2, \ldots, n, \quad (1.5) \]

where \( y_0 = 0 \) and \( \{e_t\} \) is an iid \((0, \sigma^2)\) sequence. They derived the limiting distribution of the ordinary least squares estimator

\[ \hat{\rho} = \frac{n}{\sum_{t=1}^{n} y_t y_{t-1}} \cdot \frac{\sum_{t=1}^{n} y_t^2}{\sum_{t=1}^{n} y_{t-1}^2} \quad (1.6) \]

for the model when the true value of \( \rho \) is one. The limiting distribution of \( \hat{\rho} \) is

\[ n(\hat{\rho} - 1) \Rightarrow 2^{-1} \Gamma^{-1}(T^2 - 1), \quad (1.7) \]

where \( \Rightarrow \) denote convergence in distribution,

\[ \Gamma = \sum_{i=1}^{\infty} \gamma_i^2 \zeta_i, \quad T = \sum_{i=1}^{\infty} 2^{1/2} \gamma_i \zeta_i, \quad \gamma_i = 2(-1)^{i+1}/\{(2i-1)\pi\}, \]

and

\( \{\zeta_i\} \) is an iid \(N(0,1)\) sequence.

Utilizing expression (1.7) for the limiting distribution, Dickey and Fuller prepared
a set of tables of the percentiles of the distribution by Monte Carlo simulation. One version of the table can be found in Fuller (1976, p. 371). White (1958), at the end of his paper, mentioned briefly a different form of the limiting distribution of \( n(\hat{\rho} - 1) \) given by

\[
2^{-1}\{W^2(1) - 1\} / \int_0^1 W^2(\tau) \, d\tau,
\]

(1.8)

where \( W(\cdot) \) is the standard Brownian motion on \([0,1]\). In fact we can show directly that the two limiting distributions are same by expanding the Brownian motion by the reproducing kernel method

\[
W(\tau) = \sum_{i=1}^{\infty} \frac{2^{3/2}}{(2i-1)!} \sin\{(i - 1/2)\pi\tau\} \zeta_i, \quad \tau \in [0,1],
\]

where \( \{\zeta_i\} \) is a sequence of iid \( N(0,1) \) random variables. (For details see Chan and Wei (1988, p. 382)).

Fuller (1976, pp. 373 — 381) considered a \( p \)-th order autoregressive model

\[
y_t + \sum_{j=1}^{p} \alpha_j y_{t-j} = \epsilon_t, \quad t = p+1, p+2, \ldots,
\]

where \( p - 1 \) roots of the characteristic equation \( A(m) = 0 \) are inside the unit circle and the remaining root is one. Fuller rewrote his model

\[
y_t = \psi_1 y_{t-1} + \sum_{j=2}^{p} \psi_j (y_{t-j+1} - y_{t-j}) + \epsilon_t, \quad t = p+1, p+2, \ldots
\]
He regresses \((y_t - y_{t-1})\) on \(y_{t-1}, (y_{t-1} - y_{t-2}), \ldots, (y_{t-p+1} - y_{t-p})\) to get \((\psi_1 - 1), \psi_2, \ldots, \psi_p\), the least squares estimator of \(\psi_1, \psi_2, \ldots, \psi_p\). He showed that, for \(c = (1 + \alpha_1 + \cdots + \alpha_p)^{-1} > 0\), \(nc(\psi_1 - 1)\) has the limiting distribution (1.7).

Dickey and Fuller (1981) also considered the model

\[ y_t = \mu + \delta(t - 1 - n/2) + \rho y_{t-1} + e_t, \quad t = 1, 2, \ldots, n. \]

They derived the limiting distribution of the likelihood ratio statistics for testing the hypotheses \((\mu, \delta, \rho) = (0, 0, 1), (\mu, \delta, \rho) = (\mu, 0, 1)\). They also prepared a set of tables of percentiles for the distributions of the test statistics.

Dickey, Hasza, and Fuller (1984) considered a seasonal autoregressive time series model

\[ y_t = \alpha_d y_{t-d} + e_t, \quad t = 1, 2, \ldots, n, \]

where \(d\) is a fixed positive integer, \(y_{-d+1}, \ldots, y_0\) are fixed values, and \(\{e_t\}\) is a sequence of an iid(0, \(\sigma^2\)) random variables, and \(\alpha_d = 1\). They investigated the limiting distribution of

\[ \hat{\alpha}_d = \left( \sum_{t=1}^{n} y_{t-d}^2 \right)^{-1} \sum_{t=1}^{n} y_{t-d} y_t \]

and showed that

\[ n(\hat{\alpha}_d - 1) \Rightarrow 2^{-1}d \sum_{j=1}^{d} (\Gamma_j)^{-1} \sum_{j=1}^{d} (T_j^2 - 1), \]

where
\[ \Gamma_j = \sum_{i=1}^{\infty} \gamma_1^i \zeta_{ij}^2, \quad T_j = \sum_{i=1}^{\infty} 2^{1/2} \gamma_i \zeta_{ij}, \quad \gamma_1 = 2(-1)^{i+1}/\{2i-1\}, \]

and \( \{\zeta_{ij}\} \) is an iid \( N(0,1) \) double array. They gave a set of tables of percentiles of the limiting distributions of \( \hat{\alpha}_2, \hat{\alpha}_4, \hat{\alpha}_{12} \).

Said and Dickey (1984) suggested an estimation procedure for the parameters in the model

\[ y_t = \beta y_{t-1} + z_t, \]

\[ z_t + \alpha_1 z_{t-1} + \cdots + \alpha_p z_{t-p} = e_t + \beta_1 e_{t-1} + \cdots + \beta_q e_{t-q}. \]

They assumed the usual stationarity and invertibility condition on the process \( \{z_t\} \) and assumed \( \{e_t\} \) be an iid(0, \( \sigma^2 \) ) sequence. They considered \( \rho = 1 \) as a null hypothesis. They approximated the process by a k-th order autoregression and obtained an estimator \( \hat{\rho} - 1 \) of \( \rho - 1 \) as the first coefficient in the regression of \( y_t - y_{t-1} \) on \( y_{t-1}, y_{t-2}, \ldots y_{t-k} \). They assumed \( n^{-1/3}k \to 0 \) as \( n \to \infty \) and that \( ck > n^{1/r} \) for some \( c > 0, r > 0 \). They showed that the limiting distribution of their estimator is the same as that of Dickey and Fuller (1979) up to a scalar multiple. That is,

\[ (1 + \alpha_1 + \cdots + \alpha_p)^{-1}(1 + \beta_1 + \cdots + \beta_p) n(\hat{\rho} - 1) \Rightarrow 2^{-1}i^{-1}(T^2 - 1). \]

Said and Dickey (1985) studied a finite step Gauss–Newton procedure of obtaining an estimator of the parameters in the model
when \( \rho = 1 \) and \( \{z_t\} \) is stationary and invertible. The limiting distribution of \( n(\hat{\rho} - 1) \) is same as that of Dickey and Fuller. For the initialization of the Gauss--Newton procedure, they used one for the initial value of \( \rho \) and a method of moment estimator for the initial value of \( (\beta_1, \ldots, \beta_q) \). Because of the initial values, the power function of the suggested test is not monotone as \( \rho \) moves away from one.

Phillips (1987) generalized the Dickey and Fuller (1979) estimator (1.6) to a more general innovation case. The model is

\[
y_t = \rho y_{t-1} + u_t, \quad t = 1, 2, \ldots, n, (1.9)
\]

where \( \rho = 1 \) and \( \{u_t\} \) is a strong alpha--mixing process satisfying

(a) \( \mathbb{E}(u_t) = 0, \) for all \( t, \)

(b) \( \sup_t \mathbb{E}|u_t|^\delta < \infty \) for some \( \delta > 2, \)

(c) \( \lim_{n \to \infty} \mathbb{E}[n^{-1} \sum_{t=1}^n y_t^2] \) exists and is positive.
(d) The mixing coefficients \( \{ \alpha(m) \} \) satisfy \( \sum_{m=1}^{\infty} \alpha(m)^{1-2/\delta} < \infty \).

For the definition of strong alpha—mixing see Definition A.14 in Appendix A. The limiting distribution of the ordinary least squares estimator \( \hat{\rho} \) of (1.6) for model (1.9) is closely related to the limiting distribution of Dickey and Fuller (1979). Phillips showed

\[
 n(\hat{\rho} - 1) \Rightarrow 2^{-1} \Gamma^{-1}(T^2 - \sigma_u^2 / \kappa^2),
\]

where

\[
 \sigma_u^2 = \lim_{n \to \infty} n^{-1} \sum_{t=1}^{n} \text{E}(u_t^2)
\]

and

\[
 \kappa^2 = \lim_{n \to \infty} n^{-1} \text{E}(u_1 + \cdots + u_n)^2.
\]

Phillips suggested an estimator \((s_u^2, k^2)\) of \((\sigma_u^2, \kappa^2)\) and showed that

\[
 n(\hat{\rho} - 1) - 2^{-1}(k^2 - s_u^2)/(n^{-1} \sum_{t=1}^{n} y_t^2) \Rightarrow 2^{-1} \Gamma^{-1}(T^2 - 1).
\]

However, simulations showed that the test statistics converged to the limit very slowly and that the test has poor power.

Phillips and Perron (1988) generalized Dickey and Fuller (1981) procedures to the strong mixing innovation \( \{ u_t \} \) case described in (a) — (d) above. Their statistics are a simple adjustment of those of Dickey and Fuller (1981).
Hasza and Fuller (1979) discussed the second order autoregressive model

\[ y_t = \eta_1 y_{t-1} + \eta_2 y_{t-2} + e_t, \quad t = 1, 2, \ldots, n, \quad y_0 = y_1 = 0, \]

where \( \{e_t\} \) is an iid \((0, \sigma^2)\) sequence. The corresponding characteristic equation is assumed to have two unit roots. They reparametrized the model

\[ y_t = \alpha y_{t-1} + \beta(y_{t-1} - y_{t-2}) + e_t \]

and obtained the limiting distribution of the least squares estimator of \((\alpha, \beta)\). Let \( (\hat{\alpha}, \hat{\beta}) \) be the regression coefficients in the regression of \( y_t \) on \( y_{t-1} \) and \( (y_{t-1} - y_{t-2}) \). Then

\[ [n^2(\hat{\alpha} - 1), n(\hat{\beta} - 1)] \Rightarrow H^{-1}h, \]

where \( H \) is a 2 x 2 matrix and \( h \) is a 2-dimensional vector whose elements are linear functions of

\[ \sum_{i=1}^{\infty} \gamma_i^1 \zeta_i^1, \sum_{i=1}^{\infty} \gamma_i^2 \zeta_i^2, \sum_{i=1}^{\infty} \gamma_i^3 \zeta_i^3, \sum_{i=1}^{\infty} \gamma_i^4 \zeta_i^4 \]

for an iid \( N(0, 1) \) sequence \( \{\zeta_i\} \).

Chan and Wei (1988) generalized the results of Dickey and Fuller (1979) to the model

\[ y_t + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} = e_t, \quad t = 1, 2, \ldots, n. \]

They assumed that \( \{e_t\} \) is a martingale difference sequence with respect to an
increasing sequence of $\sigma$–field $\{F_t\}$ such that

$$E\{e_t^2 | F_{t-1}\} = 1 \text{ a.s.,}$$

$$\sup_t E\{|e_t|^{2+\delta} | F_{t-1}\} < \infty \text{ a.s. for some } \delta > 0.$$  

The roots of the characteristic equation $m^p + \alpha_1 m^{p-1} + \cdots + \alpha_p = 0$ are assumed to be on or inside the unit circle. The number of roots on the unit circle can be greater than two. Chan and Wei derived the limiting distribution of the least squares estimator of $\alpha = (\alpha_1, \ldots, \alpha_p)'$. Their representation of the limiting distribution is a functional of standard Wiener process.

Fountis and Dickey (1989) studied the unit root problem for multivariate autoregressive time series. They considered the $k$–variate autoregressive process $\{Y_t\}$

$$Y_t = AY_{t-1} + e_t, \ t = 1,2,\ldots,n,$$

where $\{e_t\}$ is a $k$–variate iid$(0,\Sigma)$ sequence. They showed that $\hat{\lambda}_n$, the largest absolute value of eigenvalue of

$$\hat{A} = \frac{1}{n} \sum_{t=1}^{n} Y_t Y'_{t-1}( \sum_{t=1}^{n} Y_{t-1} Y'_t)^{-1}$$

has the same limiting distribution as that of $\hat{\rho}$ of (1.6) as given by Dickey and Fuller (1979). They assumed that $A$ has one eigenvalue of one and that the rest of
the eigenvalues are less than one in magnitude. They also assumed that the eigenvalues are less than one in magnitude. They also assumed that $\Sigma_{i=0}^{k-1} A^i \Sigma (A^i)'$ is of full rank.

1.3. Assumptions and Definitions of Estimators

We consider model defined in (1.1) and (1.4). We define

$$\theta = (\alpha', \beta')' = (\alpha_1, \alpha_2, ..., \alpha_p, \beta_1, \beta_2, ..., \beta_q)',$$

$$Y = (y_1, y_2, ..., y_n)', \quad Y_1 = (y_0, y_1, ..., y_{n-1})', \quad Z = (z_1, z_2, ..., z_n)',$$

$$e = (e_1, e_2, ..., e_n)', \quad \Gamma_n(\theta) = \text{Var}(Z) = E(ZZ').$$

The formal assumptions about model (1.1) are given in Assumption 1.1.

Assumption 1.1. In model (1.1), $\{e_i\}$ is an iid $(0, \sigma^2)$ sequence. The observations are $y_0, y_1, ..., y_n$. We denote by $(\theta^0, \rho^0, \sigma^0)$ the true value of $(\theta, \rho, \sigma)$. The true $\rho^0$ is assumed to be 1. Also $\sigma^0$ is assumed to be positive. We assume that the parameter space is such that for all $\theta$, the equation $A(m) = 0$ and $B(m) = 0$ have roots with absolute value not greater than $1 - \eta$ for some $\eta > 0$ independent of $\theta$. Also for any root $m_a$ of $A(m) = 0$ and any root $m_b$ of $B(m) = 0$ we assume $|m_a - m_b| \geq 1 - \eta$. Denote the set of all those $\theta$ satisfying the above conditions by $\Theta$. 


The condition \(|m_a - m_b| \geq 1 - \eta\) is a kind of identifiability condition for \(\theta\). Let \(m_1, ..., m_p\) be the roots of \(A(m) = 0\) and \(m_1, ..., m_q\) be the roots of \(B(m) = 0\). Since the map \((m_1, ..., m_p, m_1, ..., m_q) \rightarrow \theta\) is continuous (see Brockwell and Davis 1987, p. 366, remark 3), the set \(\Theta\) in Assumption 1.1 is compact.

In addition to considering the estimation problem for model (1.1) we also show the strong consistency of several types of estimators for the stationary and invertible process given in (1.4). Assumptions about model (1.4) are given in assumption 1.2.

Assumption 1.2. In model (1.4), \(\{e_t\}\) is an iid \((0, \sigma^2)\) sequence. The observations are \(x_1, ..., x_n\). We denote by \((\theta^0, \sigma^0)\) the true value of \((\theta, \sigma)\). We assume that \(\sigma^0\) is positive. We assume that the parameter space is such that for all \(\theta\), the equations \(A(m) = 0\) and \(B(m) = 0\) have roots with absolute value not greater than \(1 - \eta\) for some \(\eta > 0\) independent of \(\theta\). Also for any root \(m_a\) of \(A(m) = 0\) and any root \(m_b\) of \(B(m) = 0\) we assume \(|m_a - m_b| \geq 1 - \eta\).

The set of all those \(\theta\) satisfying the above conditions is the set \(\Theta\) given in Assumption 1.1. We now define several types of estimators for model (1.1).

Definition 1.3. The least squares estimator \((\bar{\theta}, \bar{\rho})\) is defined to be the \((\theta, \rho)\) which minimizes

\[
Q_n(\theta, \rho) = (Y - \rho Y_1)' \Gamma_n^{-1} \theta (Y - \rho Y_1)
\]

\[
= Z' \Gamma_n^{-1} (\theta) Z + 2(1 - \rho) Y_1' \Gamma_n^{-1} (\theta) Z + (1 - \rho)^2 Y_1' \Gamma_n^{-1} (\theta) Y_1
\]
over \((\theta, \rho) \in \Theta \times \mathbb{R}\). The least squares estimator \(\hat{\sigma}^2\) of \(\sigma^2\) is

\[
\hat{\sigma}^2 = n^{-1} Q_n(\theta, \rho). \tag{1.13}
\]

Under the assumption of normal innovations \(\{e_t\}\), the conditional likelihood function of \(Y\) given \(y_0\) is

\[
(2\pi \sigma^2)^{-n/2}[\text{det}\Gamma_n(\theta)]^{-1/2} \exp[-Q_n(\theta, \rho)/2\sigma^2]. \tag{1.14}
\]

Maximizing (1.14) is equivalent to minimizing the negative of the logarithm of (1.14). Therefore, it is equivalent to minimizing

\[
L_n(\theta, \rho, \sigma) = Q_n(\theta, \rho)/\sigma^2 + \log \text{det} \Gamma_n(\theta) + 2n \log \sigma. \tag{1.15}
\]

Note that \(\frac{\partial L_n(\theta, \rho, \sigma)}{\partial \sigma} = 0\) gives

\[
\sigma^2 = n^{-1} Q_n(\theta, \rho) = \sigma_1^2 \quad \text{(say)}.
\]

Therefore, for every \((\theta, \rho) \in \Theta \times \mathbb{R}\), we have

\[
L_n(\theta, \rho, \sigma) \geq L_n(\theta, \rho, \sigma_1) = n - n \log n + n \log \{[\text{det} \Gamma_n(\theta)]^{1/n} Q_n(\theta, \rho)}\).
\]

It follows that \(L_n(\theta, \rho, \sigma)\) can be minimized by minimizing
Definition 1.4. The maximum likelihood estimator \((\hat{\theta}, \hat{\rho})\) is the \((\theta, \rho)\) which minimizes (1.16) over \(\theta \times \mathbb{R}\) and the maximum likelihood estimator \(\hat{\sigma}^2\) of \(\sigma^2\) is

\[
\hat{\sigma}^2 = n^{-1}Q_n(\hat{\theta}, \hat{\rho}).
\]

Note that the maximum likelihood estimator is, in fact, conditional on \(y_0\). For convenience, we use the terminology 'maximum likelihood estimator' instead of using the terminology 'conditional maximum likelihood estimator'.

Finally we define the ordinary least squares estimator \((\tilde{\theta}, \tilde{\rho})\). Since in (1.1) the process \(\{z_t\}\) is invertible, we can find a sequence \(\{d_j(\theta)\}\) such that

\[
e_t = \sum_{j=0}^{\infty} d_j(\theta) z_{t-j} \quad \text{for all } t = \ldots, -1, 0, 1, \ldots \quad (1.17)
\]

Given \((y_0, y_1, \ldots, y_n)\), let

\[
e_t(y; \theta, \rho) = \sum_{j=0}^{t-1} d_j(\theta) z_{t-j} = \sum_{j=0}^{t-1} d_j(\theta) (y_{t-j} - \rho y_{t-1-j}), \quad t=1,2,\ldots,n. \quad (1.18)
\]

The \(e_t(Y; \theta, \rho)\) are obtained from (1.17) by truncating the series at \(t-1\).
Definition 1.5. The ordinary least squares estimator $(\hat{\theta}, \hat{\rho})$ is the $(\theta, \rho)$ which minimizes, over $\theta \times \mathbb{R}$,

$$S_n(\theta, \rho) = \sum_{t=1}^{n} e_t^2(Y_t; \theta, \rho).$$

(1.19)

The ordinary least squares estimator $\hat{\sigma}^2$ of $\sigma^2$ is

$$\hat{\sigma}^2 = n^{-1}S_n(\theta, \rho).$$

1.4. Summary of Main Results

In Chapter 2, we derive uniform boundedness properties for the matrix norms of $\Gamma_n(\theta)$ and $\Gamma_n^{-1}(\theta)$, and for the matrix norms of the partial derivatives of $\Gamma_n(\theta)$ and $\Gamma_n^{-1}(\theta)$. In Chapter 3, we establish the strong consistency of the least squares estimator, the maximum likelihood estimator, and the ordinary least squares estimator under Assumption 1.1. We also obtain a new proof of the strong consistency of the three estimators for the stationary model (1.4) under Assumption 1.2. In Chapter 4, the limiting distributions of the three estimators are established. One of the most interesting results is that the limiting distribution of the least squares estimator $\tilde{\rho}$, the maximum likelihood estimator $\hat{\rho}$, and the ordinary least squares estimator $\rho$ are same as that of Dickey and Fuller (1979). Therefore our result can be used for the unit root test $H_0 : \rho = 1$ using their table. The test will be a good candidate test for a unit root even under unknown $p, q$. The case of
unknown \((p,q)\) is discussed by Said and Dickey (1984), and Phillips and Perron (1985). In Chapter 5, model (1.1) is extended to the model with intercept

\[
y_t = \mu + \rho y_{t-1} + z_t,
\]

\[
z_t + \alpha_1 z_{t-1} + \cdots + \alpha_p z_{t-p} = e_t + \beta_1 e_{t-1} + \cdots + \beta_q e_{t-q}, \quad t = 1, 2, \ldots, n.
\]

The weak consistency of the least squares estimator, the maximum likelihood estimator, and the ordinary least squares estimator is established for \((p, \mu) = (1, 0)\). Also the limiting distribution of the estimators of \(\rho, \mu, \alpha,\) and \(\beta\) is derived. In Chapter 6, a Monte Carlo simulation is conducted.
2. PRELIMINARY RESULTS

2.1. Literature on the Properties of $\Gamma_n = \Gamma_n(\theta)$

The functions $S_n(\theta, \rho), Q_n(\theta, \rho), I_n(\theta, \rho, \sigma)$ which define the ordinary least squares estimator, the least squares estimator, and the maximum likelihood estimator contain the terms $\Gamma_n^{-1}$ and $\det(\Gamma_n)$, where $\Gamma_n$ is the $n \times n$ covariance matrix of $Z$ defined in (1.10). There has been a lot of effort to evaluate and characterize $\Gamma_n^{-1}$ and $\det \Gamma_n$.

Shaman (1969) obtained an exact expression for $\Gamma_n^{-1} = \Gamma_n^{-1}(\beta_1)$ for $p=0$ and $q=1$. He showed

$$[\Gamma_n^{-1}(\beta_1)]_{ij} = \frac{(-\beta_1)^{j-i}(1 - \beta_1^{2i})[1 - \beta_1^{2(n-j+1)}]}{\sigma^2(1 - \beta_1^{2})[1 - \beta_1^{2(n+1)}]}, \quad j \geq i.$$  

He also gave two more different expressions for $\Gamma_n^{-1}(\beta_1)$.

Shaman (1973) found an exact expression of $\Gamma_n^{-1}(\theta)$ for $p = 0$ and $q = 2$. The expression is complicated and is not given here. He also suggested techniques which can be used to construct the inverse of the covariance matrix of the process with $p = 0$ and general $q$.

Murthy (1974) proposed a different way of obtaining $\Gamma_n^{-1}$ for $p = 0$ and $q = 1$. First he obtained the inverse of an approximate matrix $\hat{\Gamma}_n^{-1}(\beta_1)$ whose $(1,1)$ and $(n,n)$ elements are 1 instead of $1 + \beta_1^2$ and the remaining elements are the same as
that of $\Gamma_n(\beta_1)$. The inverse of $\hat{\Gamma}_n^{-1}(\beta_1)$ is given by

$$[\hat{\Gamma}_n^{-1}(\beta_1)]_{ij} = (-\beta_1)^{|i-j|/(1-\beta_1^2)}.$$ 

Next Murthy adjusted the inverse to obtain the exact inverse of $\Gamma_n^{-1}(\beta_1)$.

Galbraith and Galbraith (1974) gave a general formula for $\Gamma_n^{-1}$. Let

$$Z^0 = (z_1, z_2, \ldots, z_p, e_1, e_2, \ldots, e_q)'.$$

They considered the following transformation of $Z$ and $Z_0$,

$$\begin{pmatrix} Z^0 \\ e \end{pmatrix} = \begin{pmatrix} 0 \\ D_n \end{pmatrix} Z + \begin{pmatrix} I \\ M_n \end{pmatrix} Z^0 = LZ + KZ^0 $$

(2.1)

where $Z = (z_1, z_2, \ldots, z_n)', e = (e_1, e_2, \ldots, e_n), D_n$ is an $n \times n$ matrix, $M_n$ is an $n \times (p+q)$ matrix, and $0$ is $(p+q)$ dimensional column vector of zeros. From this transformation, they derived the marginal probability density of $Z$,

$$P(Z|\theta, \sigma) = (2\pi\sigma^2)^{-1/2n}[\det C]^{1/2}[\det(K'CK)]^{-1/2}$$

$$\times \exp\left\{ -\frac{1}{2\sigma^2} [LZ - K(K'CK)^{-1}(K'CL)]'C[LZ - K(K'CK)^{-1}(K'CL)] \right\},$$
where \( C = \text{diag}(A, I_n) \), \( \Lambda = \sigma^{-2} \text{Var}(Z^0) \). Comparing this equation with (1.14), where \( Y - \rho Y_1 \) is replaced by \( Z \) in \( Q(\theta, \rho) \), they derived

\[
\Gamma_n^{-1} = D_n' D_n - D_n' M_n (\Lambda^{-1} + M_n' M_n)^{-1} M_n' D_n, \tag{2.2}
\]

\[
\det \Gamma_n = \det(A) \cdot \det(\Lambda^{-1} + M_n' M_n) = \det(I_{p+q} + B' M_n' M_n B),
\]

where \( B \) is a \((p+q) \times (p+q)\) matrix such that \( \Lambda = B'B \). They also showed that the matrix \( D_n \) is lower triangular with \((i,j)\) element \( d_{|i-j|} \) satisfying

\[
d_j + \beta_1 d_{j-1} + \cdots + \beta_q d_{j-q} = \alpha_j, j = 0,1,\ldots,
\]

where \( d_j = 0 \) for \( j < 0 \), \( d_0 = \alpha_0 = 1 \), and \( \alpha_j = 0 \) for \( j > p \). They showed that the elements of the columns of \( M_n \) satisfy the difference equation (2.3) with different starting values. Writing \( m_{ij} \) for the \((i,j)\) element of \( M_n \), we have

\[
m_{ij} + \beta_1 m_{i-1,j} + \cdots + \beta_q m_{i-q,j} = 0, i = q, q+1,\ldots, j = 1,2,\ldots, p+q. \tag{2.4}
\]

Newbold (1974) presented almost the same results as that given by Galbraith and Galbraith (1974).

For later use, we give expressions for \( D_n' D_n \) and \( D_n' M_n \) in terms of \( d_j \)'s and \( m_{jk} \)'s. We have

\[
[D_n' D_n]_{ij} = \sum_{k=0}^{n-\max(i,j)} a_k d_{|i-j|+k}
\]

and
See 8.1.1 in Appendix B. Note that from (1.18), we have

\[ S_n(\theta, \rho) = \sum_{t=1}^{n} e_t^2(Y; \theta, \rho) = (Y - \rho Y_1)' D_n D_n' (Y - \rho Y_1). \]

Shaman (1975) discussed an approximation to the inverse of the covariance matrices of moving average and autoregressive processes. He observed that the inverse \( \Gamma_{AR}^{-1} \) of the covariance matrix of an autoregressive process of order \( q \) is identical to the covariance matrix \( \Gamma_{MA} \) of a moving average process of order \( q \) with the same coefficients up to a \( \sigma^2 \) multiplication, except for the \( q \times q \) submatrices in the upper left and lower right corners. Using this observation he showed that the matrix \( \sigma^2 \Gamma_{MA} \Gamma_{AR}^{-1} - I \) has \( 2q \) positive eigenvalues for a stationary autoregressive process (or for an invertible moving average process).

Anderson (1976a), observing the similarity between \( \Gamma_{MA} \) and \( \Gamma_{AR}^{-1} \), and following the approach due to Murthy (1974), proposed a method for getting \( \Gamma_{AR}^{-1} \) from an adjustment of \( \Gamma_{MA} \). Anderson (1976b), following Anderson (1976a), gave a method of deriving \( \Gamma_n^{-1} \) for general \( p \) and \( q \).

Eltinge (1990) discussed the inversion of the covariance matrix of a multivariate autoregressive moving average process. He gave multivariate versions of (2.2) and (2.3). He also discussed the exponentially declining properties of the
off-diagonal elements of the inverse of the covariance matrix.

2.2. Properties of $\Gamma_n$

In this section, we investigate the properties of $\Gamma_n^{-1}$ and $\det \Gamma_n$. Under a mild regularity condition, uniform boundedness of matrix norms of $\Gamma_n$, $\Gamma_n^{-1}$, derivatives of $\Gamma_n$, and derivatives of $\Gamma_n^{-1}$ are established. Also the uniform boundedness of $\det \Gamma_n^{-1}$ and $\det \Gamma_n'$ is given. In the treatment of $\Gamma_n^{-1}$ and $\det \Gamma_n'$, we heavily use the results (2.2) and (2.3) of Galbraith and Galbraith (1974).

We begin with a representation of $\Gamma_n$. In (1.1), since $\{z_t\}$ is stationary and invertible, we can find sequences $\{d_j\}$ and $\{v_j\}$ such that

\[ e_t = \sum_{j=0}^{\infty} d_j z_{t-j} \]
and
\[ z_t = \sum_{j=0}^{\infty} v_j e_{t-j}. \]  

(2.6)

The sequences $\{d_j\}$ and $\{v_j\}$ satisfy the difference equations

\[ v_j + \alpha_1 v_{j-1} + \cdots + \alpha_p v_{j-p} = 0 \]
and
\[ d_j + \beta_1 d_{j-1} + \cdots + \beta_{j-q} d_{j-q} = 0, \quad j \geq \max (p, q) \]
with some starting values of \( v_j \) and \( d_j, j = 0, 1, \ldots, \max(p, q) - 1 \). Those starting values are polynomials in \( \theta = (\alpha', \beta')' \). See Fuller (1976, pp. 68 - 70). For the definition of polynomials of multiple variables see Definition 2.2 below. Note that the \((i, j)\) element \( r(|i-j|) \) of \( \Gamma_n \) is

\[
r(|i-j|) = E(z_i z_j) = \sum_{k=0}^{\infty} v_k v_{|i-j|+k} \sigma^2.
\]  

(2.7)

Note that the quantities \( d_j, v_j, r(j), D_n, M_n, \) and \( \Lambda \) are functions of \( \theta \). Henceforth, for the convenience of notation, we omit writing the \( \theta \) dependency for those terms and other \( \theta \)-dependent terms that will appear later. The sequences \( \{d_j\}, \{v_j\}, \{r(j)\} \) and many sequences that will appear later decline exponentially to 0 as \( j \) increases to \( \infty \). To be more precise we give the definition of exponential decline.

**Definition 2.1.** A sequence of real functions \( \{a_j(\theta)\} \) of \( \theta \in \theta \) is said to be declining exponentially in \( j \) uniformly in \( \theta \in \theta \) if there exist finite positive real numbers \( M \) and \( \lambda \in (0, 1) \) such that, for all \( j = 0, 1, \ldots \),

\[
\sup_{\theta \in \theta} |a_j(\theta)| \leq M \lambda^j.
\]

Also a sequence of real numbers \( \{b_j\} \) is said to be declining exponentially in \( j \) if there exist finite positive real numbers \( M \) and \( \lambda \in (0, 1) \) such that for all \( j = 0, 1, \ldots \)

\[
|b_j| \leq M \lambda^j.
\]

The terms \( M \) and \( \lambda \) are called the coefficients of exponential decline.
Now we define the polynomial of multiple variables.

**Definition 2.2.** Let $m$ and $k$ be nonnegative integers and let $S$ be the space of all $m \times m$ real matrices. A function $f : S \rightarrow \mathbb{R}$ is said to be a polynomial in $x = (x_1, \ldots, x_k) \in S^k$ if there are real numbers $\{a_{i_1, \ldots, i_k}\}, i_1, \ldots, i_k = 0, \ldots, k$ such that

$$f(x) = \sum_{i_1=0}^{k} \cdots \sum_{i_k=0}^{k} a_{i_1, i_2, \ldots, i_k} x_1^{i_1} x_2^{i_2} \cdots x_k^{i_k}.$$ 

For fixed $j$, $d_j$ and $v_j$ are polynomials in $\theta$. Also partial derivatives of $|A(m)|^2 = A(m)A(m)$ and $|B(m)|^2 = B(m)B(m)$ are polynomials in $\theta$, where $\overline{m}$ is the complex conjugate of $m$. The fixed order derivatives of $\Gamma_n^{-1}(\theta)$, as will be given in (2.18), are polynomials in $(\Gamma_n(\theta), \Gamma_n^{-1}(\theta))$.

For the exponentially declining properties of $d_j$ and $v_j$ in (2.6), we need Theorem 2.3 and Theorem 2.4 on difference equations.

**Theorem 2.3.** Consider the difference equation

$$w_j + a_1w_{j-1} + \cdots + a_{r}w_{j-r} = 0, \; j = 0, 1, \ldots,$$

where $r$ is a fixed nonnegative integer, $w_j = 0$ for $j < 0$, and $w_0 = 1$. Assume the roots $m_1, m_2, \ldots, m_r$ of the equation
have absolute value less than one. Then the unique solution \( w_j \) to the difference equation is given by

\[
w_j = \frac{m_1^{k_1} m_2^{k_2} \cdots m_r^{k_r}}{k_1 + \cdots + k_r = j}
\]

where the summation is understood to be taken over all possible nonnegative integer combinations of \((k_1, k_2, \ldots, k_r)\) with \(k_1 + \cdots + k_r = j\).

Proof. Let

\[
b_j = \begin{cases} 
0 & \text{for } j \neq 0 \\
1 & \text{for } j = 0.
\end{cases}
\]

Let \( B \) be the back shift operator, that is, \( Bx_t = x_{t-1} \). Then

\[
(1 - m_1 B) \cdots (1 - m_r B)w_j = b_j, \quad j = \ldots, -1, 0, 1, \ldots
\]

Therefore,

\[
w_j = [(1 - m_1 B) \cdots (1 - m_r B)]^{-1} b_j
\]

\[
= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_r=0}^{\infty} m_1^{k_1} m_2^{k_2} \cdots m_r^{k_r} b_{j-k_1-k_2-\cdots-k_r}
\]

\[
= \frac{m_1^{k_1} m_2^{k_2} \cdots m_r^{k_r}}{k_1 + \cdots + k_r = j}
\]

\( \square \)
The result of Theorem 2.3 holds regardless of the magnitude of \( m_1, m_2, \ldots, m_r \). However, the proof for the general case is more complicated and, hence, is not given here.

**Theorem 2.4.** Let \( \theta \) be defined in Assumption 1.1. Let \( r \) be a fixed nonnegative integer and let \( \theta \in \Theta \). Consider the difference equation

\[
v_j + a_1(\theta)v_{j-1} + \cdots + a_r(\theta)v_{j-r} = 0, \quad j = r, r+1, \ldots \tag{2.8}
\]

with starting values

\[
v_0 = c_0(\theta), \quad v_1 = c_1(\theta), \quad \ldots, \quad v_{r-1} = c_{r-1}(\theta).
\]

Assume that for all \( \theta \in \Theta \), the roots \( m_1, \ldots, m_r \) of the characteristic equation

\[
m^r + a_1(\theta)m^{r-1} + \cdots + a_{r-1}(\theta)m + a_r(\theta) = 0
\]

have absolute value less than \( 1 - \epsilon \) for some \( \epsilon > 0 \) independent of \( \theta \). Also assume that there is \( C < \omega \) such that \( \sup_{\theta} |c_i(\theta)| \leq C \) for all \( i = 1, \ldots, r \). Then \( v_j \) declines exponentially in \( j \) uniformly in \( \theta \in \Theta \).

**Proof.** For notational convenience we omit writing \( \theta \) in the expressions for the \( c_i \)'s and \( a_i \)'s. Let \( w_j \) be defined in Theorem 2.3 and let
We show that \( v_j \) given in (2.10) satisfy the difference equation (2.8) with initial condition (2.9). For \( 0 \leq j \leq r - 1 \), we know that \( v_j \) in (2.10) satisfy

\[
v_j = \sum_{i=0}^{r} c_i \left( \sum_{k=i}^{r} a_{k-i} w_{j-k} \right)
\]

\[
= \sum_{i=0}^{r} c_i \left( w_{j-i} + a_1 w_{j-i-1} + \cdots + a_{r-i} w_{j-r} \right) = c_j.
\]

Therefore \( v_j \) in (2.10) satisfy the initial condition (2.9). For \( j \geq r \), note that \( v_j \) in (2.10) satisfy

\[
\sum_{s=0}^{r} a_s v_{j-s} = \sum_{s=0}^{r} a_s \sum_{k=0}^{r} \sum_{i=0}^{k} a_{k-i} c_i w_{j-k} = \sum_{k=0}^{r} \sum_{i=0}^{k} a_{k-i} c_i \left( \sum_{s=0}^{r} a_s w_{j-s-k} \right) = 0.
\]

Therefore \( v_j \) in (2.10) satisfy the difference equation (2.8). By the uniqueness of the solution of difference equation (2.8) with initial condition (2.9), \( v_j \) in (2.10) is the unique solution to (2.8). By Theorem 2.3 we know that, for some \( M_1 < \infty \) and \( \lambda \in (1-\epsilon, 1) \),

\[
\sup_{\theta} |w_j| = \sup_{\theta} \frac{m_1^{k_1} m_2^{k_2} \cdots m_r^{k_r}}{k_1 + \cdots + k_r} \leq \frac{(1-\epsilon)^j \leq (j+1)^{r}(1-\epsilon)^j < M_1 \lambda^j}.
\]
Also by the assumption that \( \sup_\theta |c_i(\theta)| \leq C \) for all \( i = 1,...,r \), we can find \( M_2 < \infty \) such that

\[
\sup_\theta |c_k + a_1 c_{k-1} + \cdots + a_k c_0| < M_2, \text{ for every } k = 0,1,...,r.
\]

Therefore,

\[
\sup_\theta |v_j| \leq M_1 M_2 \lambda^j, \quad j = 0,1,\ldots
\]

Corollary 2.5. Let \( \theta \) be given in Assumption 1.1. Consider the difference equation

\[
f_j + \beta_1 f_{j-1} + \cdots + \beta_q f_{j-q} = 0, \quad j = q,q+1,\ldots
\]

where \( f_j = 0 \) for \( j < 0 \). Suppose the starting values \( f_0, f_1, \ldots, f_{q-1} \) are bounded functions of \( \theta \). We assume that the roots \( m_1, m_2, \ldots, m_r \) of the characteristic equation have absolute value less than \( 1 - \epsilon \) for some \( \epsilon > 0 \) independent of \( \theta = (\alpha', \beta') \). Also assume that the starting values of \( f_0, f_1, \ldots, f_{q-1} \) have bounded derivatives with respective to \( \theta \). Let

\[
f_{ij} = \frac{\partial f_j}{\partial \theta_i}, \quad i = 1,2,\ldots, p+q, \quad j = 0,1,\ldots
\]

Then, for all \( i = 1,2,\ldots,p+q \), \( f_{ij} \) declines exponentially in \( j \) uniformly in \( \theta \in \Theta \).

Proof. First let \( i \in \{1,2,\ldots,p\} \). Then noting that \((\theta_1, \theta_2, \ldots, \theta_p) = (\alpha_1, \alpha_2, \ldots, \alpha_p)\), we have

\[
f_{i,j} + \beta_1 f_{i,j-1} + \cdots + \beta_q f_{i,j-q} = 0, \quad j = q,q+1,\ldots
\]
From Theorem 2.4 the result follows for $1 \leq i \leq p$. Next, fix $i \in \{p+1, \ldots, p+q\}$.

Noting that $(\theta_{p+1}, \theta_{p+2}, \ldots, \theta_{p+q}) = (\beta_1, \beta_2, \ldots, \beta_q)$, we have

$$f_{i,j} + \beta_1 f_{i,j-1} + \cdots + \beta_q f_{i,j-q} = -f_{j-i+p}, \quad j = q, q+1, \ldots$$

Let

$$f_j^0 = 0 \quad \text{for } j < 0,$$

$$f_j^0 = (f_{i,j} + \beta_1 f_{i,j-1} + \cdots + \beta_q f_{i,j-q}) \quad \text{for } j = 0, 1, \ldots, q-1,$$

$$= -f_{j-i+p} \quad \text{for } j = q, q+1, \ldots,$$

where $f_{ij} = 0$ for $j < 0$. Then

$$f_{i,j} + \beta_1 f_{i,j-1} + \cdots + \beta_q f_{i,j-q} = f_j^0, \quad j = \ldots, -1, 0, 1, \ldots$$

Therefore, by the method used in the proof of Theorem 2.3,

$$f_{ij} = \sum_{k_1=0}^{\infty} \cdots \sum_{k_q=0}^{\infty} m_1^{k_1} m_2^{k_2} \cdots m_q^{k_q} f_j^0$$

$$= \sum_K m_1^{k_1} m_2^{k_2} \cdots m_q^{k_q} f_j^0$$

where

$$K = \{(k_1, \ldots, k_q) \mid k_a = 0, 1, \ldots, s = 1, 2, \ldots, q, k_1 + \cdots + k_q \leq j\}.$$
Note that \( f_{ij} \) is exponentially declining in \( j \) uniformly in \( \theta \in \Theta \). Therefore for some \( \lambda \in (0,1), M_1, M_2 < \infty, \) and \( \lambda^1 \in (1,1), \) we have

\[
|f_{ij}| \leq |\sum_{k} m_1^{k_1} m_2^{k_2} \cdots m_q^{k_q} f_{j-k_1-\cdots-k_q}^{(0)}| \\
\leq \sum_{k} \lambda_1^{k_1+\cdots+k_q} M_1^{1-k_1-\cdots-k_q} \leq M_1(j+1)^q \lambda_1^j \leq M_2 \lambda_1^j.
\]

Therefore, \( f_{ij} \) declines exponentially in \( j \) uniformly in \( \theta \in \Theta \).

In Theorem 2.6, we establish a characterization for the column vectors of \( D_n \) and \( M_n \) of (2.1). We also show that elements of the column vectors of \( D_n \) and \( M_n \) are uniformly exponentially declining.

**Theorem 2.6.** Let \( \Theta \) be the parameter space defined in Assumption 1.1. The matrix \( D_n \) of equation (2.1) is lower triangular with \( (i,j) \) element \( d_{i-j} \) satisfying

\[
d_j + \beta_1 d_{j-1} + \cdots + \beta_q d_{j-q} = \alpha_j, \quad j = 0, 1, \ldots, \quad (2.11)
\]

where \( d_j = 0 \) for \( j < 0, \) \( d_0 = \alpha_0 = 1, \) and \( \alpha_j = 0 \) for \( j > p. \) The \( (j,k) \) element, \( m_{jk}, \) of the matrix \( M_n \) satisfies

\[
m_{jk} + \beta_1 m_{j-1,k} + \cdots + \beta_q m_{j-q,k} = 0, \quad j = q, q+1, \ldots, \quad k = 1, 2, \ldots, p+q. \quad (2.12)
\]

The elements \( d_j \) in (2.11), \( m_{j1}, \ldots, m_{jp+q} \) in (2.12), \( v_j \) in (2.6), and the derivatives of them are exponentially declining uniformly-in \( \theta \in \Theta. \) Also \( r(j), \) \( j \geq 0, \) is
exponentially declining in $j$ uniformly in $\theta \in \Theta$, where $r(j)$ is the element of $\Gamma_n$ defined in (2.7).

Proof. Equation (2.11) and (2.12) are established by Galbraith and Galbraith (1974). The conclusions about $d_j, m_j, \ldots, m_{j+p+q}$ are consequences of Theorem 2.4 and Corollary 2.5. For the uniformly exponentially declining property of $r(j)$, note that by (2.7),

$$
\sup_{\theta} |r(j)| \leq \sum_{k=0}^{\infty} M^2 \lambda^{2k+j} \sigma^2 \leq M^2 (1-\lambda^2)^{-1} \sigma^2 \lambda^j,
$$

where $M$ and $\lambda \in (0,1)$ are the coefficients of exponential decline of $\{v_j\}$. \hfill \Box

2.3. The Matrix Norm

In deriving our results properties of the matrix norm are used frequently. We define the matrix norm in Definition 2.7.

**Definition 2.7.** For an $m \times n$ matrix $A$, the matrix norm $\|A\|$ of $A$ is defined by

$$
\|A\| = \sup\{|Ax|; \|x\| = 1, x \in \mathbb{R}^n\},
$$

where $|\cdot|$ is the Euclidean norm.

We summarize several properties of the matrix norm $\|\cdot\|$. Let $A$ be an $m \times n$ matrix, let $B$ be an $m \times n$ matrix, and let $C$ be an $n \times k$ matrix. Then
(i) \( \|A + B\| \leq \|A\| + \|B\| \),

(ii) \( \|AC\| \leq \|A\| \|C\| \),

(iii) \( |Ax| \leq \|A\| |x| \), for all \( x \in \mathbb{R}^n \),

(iv) \( \|A\|^2 \leq \text{tr}(AA') \).

Let \( A \) be an \( m \times m \) nonnegative definite symmetric matrix. Then

(v) \( \|A\| \) is the maximum eigenvalue of \( A \).

In Lemma 2.8, we show the uniform boundedness of \( \|(A^{-1} + M_n'M_n)^{-1}\| \).

**Lemma 2.8.** Let \( \theta \) be given in Assumption 1.1. Let \( M_n \) be given by the representation \( e = D_n Z + M_n Z^0 \) in (2.1) and \( \Lambda = \sigma^2 \text{Var}(Z^0) \). Then

\[
\sup_{n, \theta} \|(A^{-1} + M_n'M_n)^{-1}\| < \infty.
\]

**Proof.** First we show \( \sup_n \|A\| < \infty \). Noting that

\[
\sigma^2 A = \text{Var}(Z^0) = \text{Var}\{(z_{1-p}, \ldots, z_0, e_{1-q}, e_{2-q}, \ldots, e_0)'\},
\]

we have

\[
\text{tr}(A) = q + p\sigma^{-2}\text{Var}(Z_0) = q + p \sum_{j=0}^{\infty} v_j^2.
\]
Let \( M < \omega \) and \( \lambda \in (0,1) \) be the coefficients of exponential decline of \( \{v_j\} \). Since \( \Lambda \) is nonnegative definite, \( \|\Lambda\| \leq \text{sum of eigenvalues of } \Lambda = \text{tr}(\Lambda) \). Therefore,

\[
\sup_{\theta} \|\Lambda\| \leq \sup_{\theta} \text{tr}(\Lambda) \leq q + p \sum_{j=0}^{\infty} M \lambda^{2j} < \omega.
\]

Let \( \nu_0 \) be the smallest eigenvalue of \( \Lambda^{-1} \). Then \( \nu_0 = \|\Lambda\|^{-1} \). Since any real symmetric matrix is diagonalizable, we can find an orthonormal matrix \( P \) such that

\[
P(\Lambda^{-1} + M_n'M_n)P' = \text{diag}(\nu_1, \ldots, \nu_{p+q}),
\]

where \( \nu_1 \leq \nu_2 \leq \ldots \leq \nu_{p+q} \) are the eigenvalues of \( \Lambda^{-1} + M_n'M_n \). Let \( x = (1,0,0,\ldots,0)' \) be the unit vector in \( \mathbb{R}^{p+q} \). By the nonnegative definiteness of \( M_n'M_n \),

\[
\nu_1 = x'P(\Lambda^{-1} + M_n'M_n)P'x \geq x'P\Lambda^{-1}P'x
\]

\[
\geq \nu_0 x'PP'x = \nu_0 \quad \text{(see Rao 1973, p. 74)}.
\]

Therefore,

\[
\sup_{n, \theta} \| (\Lambda^{-1} + M_n'M_n)^{-1} \| = \sup_{n, \theta} \nu_1^{-1} \leq \sup_{n, \theta} \nu_0^{-1} = \sup_{\theta} \|\Lambda\| < \omega. \quad \Box
\]

Notice that, in the notation \( \sup_{n, \theta} \) and \( \sup_{n, \theta} \) used in the proof of Lemma 2.8, the range of \( \theta \) is \( \theta \) which is defined in Assumption 1.1. We use this convention throughout.

The following Gershgorin's theorem tells us that all the eigenvalues of a
square matrix are located in some disk centered at a diagonal element of a
particular row with radius not greater than sum of all absolute values of elements
of the row elements. See Kreyszig (1978, p. 313).

Theorem 2.9. (Gershgorin) Let $A = [a_{jk}]$ be an $n \times n$ square matrix. Let $\nu$ be any
eigenvalue of $A$. Then for some $j \in \{1, 2, \ldots, n\},$

$$|\nu - a_{jj}| \leq \sum_{k=1}^{n} |a_{jk}|.$$  

Using Gershgorin's theorem, we can show the uniform boundedness of matrix
norms of $\Gamma_n$, $D_n^r D_n$, and $\Gamma_n^{-1}$.

Theorem 2.10. Let $\theta$ be given in Assumption 1.1. Let $\Gamma_n = \Gamma_n(\theta)$ be the covariance
matrix of $Z$ given in (1.10) and $D_n$ be given by (2.1). Then

$$\sup_{n, \theta} \|\Gamma_n\| < \infty, \quad \sup_{n, \theta} \|D_n^r D_n\| < \infty, \quad \sup_{n, \theta} \|\Gamma_n^{-1}\| < \infty.$$  

Proof. Let $\nu$ be an eigenvalue of $\Gamma_n$. Let $M < \infty$ and $\lambda \in (0,1)$ be the maximums of
the coefficients of exponential decline of $\{v_j\}$ and the coefficients of uniform
exponential decline of $\{d_j\}$ and $\{m_j\}, \ldots, \{m_{j, p+q}\}$. Then, by Gershgorin's
theorem,

$$\sup_{n, \theta} |\nu| \leq 2 \sum_{h=0}^{\infty} \sup_{h=0}^{\infty} |r(h)| = 2 \sum_{h=0}^{\infty} \sum_{j=0}^{\infty} \sup_{\theta} v_j v_{j+h} \sigma^2$$  

\[ \sum_{h=0}^{\infty} \sum_{j=0}^{\infty} \lambda^{j+h} \sigma^2 = M \sigma^2 (1-\lambda)^{-1}(1-\lambda^2)^{-1} < \infty. \]

Since \( \| \Gamma_n \| \) is the largest eigenvalue of \( \Gamma_n \), we have \( \sup \| \Gamma_n \| < \infty. \)

To show that \( \sup \| D_n' D_n \| < \infty \), let \( \nu' \) be an eigenvalue of \( D_n' D_n \). Then by Gershgorin and (2.15), we have

\[ \sup \| \nu' \| \leq 2 \sum_{h=0}^{\infty} \sum_{j=0}^{\infty} \sup \| d_{ij} \| \sigma^2 < \infty. \]

Since \( \| D_n' D_n \| \) is the largest eigenvalue of \( D_n' D_n \), we have \( \sup \| D_n' D_n \| < \infty. \)

Finally, to show that \( \sup \| \Gamma_n^{-1} \| < \infty \), observe that

\[ \sup \| D_n' M_n' \| \leq \sup \| D_n' M_n \| \leq n \sum_{i=1}^{n} \sum_{j=1}^{n} \sup \| d_{ij} \| \leq \sum_{i=1}^{n} \sum_{j=1}^{n} (M\lambda^i)^2 \leq M^2 (p+q)(1-\lambda^2)^{-1} < \infty. \]

Now we use (2.2) and Lemma 2.8 to get

\[ \sup \| \Gamma_n^{-1} \| \leq \sup \| D_n' D_n \| + \sup \| D_n' M_n \| \leq \sup \| (I_n + M_n M_n)^{-1} \| < \infty. \]

In the following theorem, some properties of \( \det \Gamma_n \) are given.
Theorem 2.11. Under the assumptions of Theorem 2.10, we have

(i) \( \sup_{n, \theta} \det \Gamma_n < \alpha, \)

(ii) \( \det \Gamma_n > 1 \) for all \( n = 1, 2, \ldots, \) and all \( \theta \in \theta, \)

(iii) \( \lim_{n \to \infty} [\det \Gamma_n]^{1/n} = 1 \) for all \( \theta \in \theta. \)

Proof. Proofs of ii) and iii) are found in Brockwell and Davis (1987, p. 374). Hence we give only a proof of i). Let \( b_{ik} \) be the \((i, k)\) element of \( B \) given in (2.2), \( i = 1, 2, \ldots, p+q, \) \( k = 1, 2, \ldots, p+q. \) We show that the \( b_{ik} \)'s are bounded in \( \theta. \) Let \( M < \infty \) and \( \lambda \in (0, 1) \) be given in the proof of Lemma 2.8. Then we have

\[
\sum_{i=1}^{p+q} \sum_{k=1}^{p+q} b_{ik}^2 = \text{tr}(B'B) = \text{tr}(\Lambda) \leq \{q + pM(1 - \lambda^2)^{-1}\}.
\]

Therefore,

\[
\sup_{\theta} |b_{ik}| \leq \{q + pM(1 - \lambda^2)^{-1}\}^{1/2} = b_0, \text{ say.}
\]

Let \( M_1 \) and \( \lambda_1 \in (0, 1) \) be the maximums of the coefficients of uniform exponential decline of \( \{m_{i1}\}, \{m_{i2}\}, \ldots, \{m_{i,p+q}\}, \) then

\[
\text{tr}(B'M_nM_nB) = \sum_{j=1}^{p+q} \sum_{i=1}^{n} \left( \sum_{k=1}^{p+q} m_{ik} b_{kj} \right)^2 \leq \sum_{j=1}^{p+q} \sum_{i=1}^{n} \left( \sum_{k=1}^{p+q} M_1 \lambda_1 b_{0j} \right)^2.
\]
\[ \leq M^2 b_0^2 (p + q)^3 (1 - \lambda^2)^{-1}, \text{ for every } n = 1, 2, \ldots, \theta \in \Theta. \]

Therefore the eigenvalues of \((I + B'M_n'M_nB)\) are bounded by \(1 + M^2 b_0^2 (p + q)^3 (1 - \lambda^2)^{-1}\) uniformly in \(n\) and \(\theta\). Since \(\det \Gamma_n\) is the product of \((p + q)\) eigenvalues of \((I + B'M_n'M_nB)\), it is bounded by \([1 + M^2 b_0^2 (p + q)^3 (1 - \lambda^2)^{-1}]^{p+q}\) uniformly in \(n\) and \(\theta \in \Theta\). Hence \(\sup_{n, \theta} \det \Gamma_n < \infty\).

Next we consider uniformly exponentially declining property of \(r(h) = \text{E}(z_1 z_{1+h})\) and that of fixed order partial derivatives of \(r(h)\). Denote by \(g(\omega) = g(\omega; \theta)\) the spectral density of the process \(\{z_t\}\). Then

\[
g(\omega) = \frac{c^2 |B(z)|^2}{2\pi |A(z)|^2}, \quad (2.13)
\]

where

\[ z = e^{i\omega}, \quad |A(z)|^2 = A(z)A(\bar{z}), \quad |B(z)|^2 = B(z)B(\bar{z}), \quad \bar{z} = e^{-i\omega}, \]

and \(A(\cdot)\) and \(B(\cdot)\) are given in (1.2) and (1.3). The covariance function \(r(h)\) of \(\{z_t\}\) is given by

\[ r(h) = \text{E}(z_1 z_{1+h}) = \int_{-\pi}^{\pi} e^{ih\omega} g(\omega) d\omega. \]

Letting

\[ g_{\theta_i}(\omega) = \frac{\partial g(\omega)}{\partial \theta_i}, \quad g_{\theta_i \theta_j}(\omega) = \frac{\partial^2 g(\omega)}{\partial \theta_i \partial \theta_j}, \quad g_{\theta_i \theta_j \theta_k}(\omega) = \frac{\partial^3 g(\omega)}{\partial \theta_i \partial \theta_j \partial \theta_k}, \]

\[ r_{\theta_i}(h) = \frac{\partial r(h)}{\partial \theta_i}, \quad r_{\theta_i \theta_j}(h) = \frac{\partial^2 r(h)}{\partial \theta_i \partial \theta_j}, \quad r_{\theta_i \theta_j \theta_k}(h) = \frac{\partial^3 r(h)}{\partial \theta_i \partial \theta_j \partial \theta_k}, \]
we have

\[ r_{\theta_1}(h) = \int_{-\pi}^{\pi} e^{ih\omega} g_{\theta_1}(\omega) d\omega, \quad (2.14) \]

\[ r_{\theta_i\theta_j}(h) = \int_{-\pi}^{\pi} e^{ih\omega} g_{\theta_i\theta_j}(\omega) d\omega, \tag{2.15} \]

\[ r_{\theta_i\theta_j\theta_k}(h) = \int_{-\pi}^{\pi} e^{ih\omega} g_{\theta_i\theta_j\theta_k}(\omega) d\omega, \quad i, j, k = 1, 2, \ldots, p+q. \tag{2.16} \]

For a justification of the interchange of integration and differentiation see 8.1.2 in Appendix B.

In order to show uniformly exponentially declining property of \( r(h) \) and that of fixed order partial derivatives of \( r(h) \), observe that, by (2.14) – (2.16), \( r(h) \) and the partial derivatives of \( r(h) \) are of the following form

\[ \sigma^2 \sum_{k=-s_1}^{s_1} a_k(\theta) e^{ik\omega} / |A(e^{i\omega})|^{2s_2} \]

for some nonnegative integers \( s_1, s_2 \) and polynomials \( a_k(\theta) \), \( k = -s_1, \ldots, s_1 \) of \( \theta \). In Lemma 2.12, we show uniformly exponentially declining property of

\[ |\int_{-\pi}^{\pi} e^{ih\omega} |A(e^{i\omega})|^{-2s} d\omega| \]

for all nonnegative integers \( s \) and \( h \).
Lemma 2.12. Let $A(\cdot)$ be given by (1.2). Let Assumption 1.1 hold. Then for any fixed nonnegative integer $s$ and integer $h$

$$\sup_{\theta}\left|\int_{-\pi}^{\pi} e^{ih\omega} |A(e^{i\omega})|^{-2s} d\omega\right|$$

is declining exponentially in $h$.

Proof. Fix $\theta \in \mathbb{R}$. Let $m_1 = r_1 e^{i\xi_1}, \ldots, m_p = r_p e^{i\xi_p}$ be the roots of $A(m) = 0$ in (1.2). By Assumption 1.1 we have $|r_j| \leq 1 - \eta$, $\eta > 0$, for all $j = 1, 2, \ldots, p$. Noting that

$$(z - m_j)^{-1} = z^{-1} (1 - m_j z^{-1})^{-1} = z^{-1} \sum_{k=0}^{\infty} (m_j z^{-1})^k,$$

we have

$$(z - m_j)^{-s} = z^{-s} \sum_{k_1=0}^{\infty} \cdots \sum_{k_s=0}^{\infty} (m_j z^{-1})^{k_1} \cdots (m_j z^{-1})^{k_s}.$$

By the dominated convergence theorem, together with $zz = 1$, and $m_j = r_j e^{i\xi_j}$, we have

$$\int_{-\pi}^{\pi} e^{ih\omega} |A(e^{i\omega})|^{-2s} d\omega = \int_{-\pi}^{\pi} e^{ih\omega} \prod_{j=1}^{p} [(z - m_j)^{-s} (z - \overline{m}_j)^{-s}] d\omega$$

$$= \sum_{k_1=0}^{\infty} \cdots \sum_{k_s=0}^{\infty} \cdots \sum_{k_{2p-1}=0}^{\infty} \sum_{k_{2p}=0}^{\infty} r_1^{k_1 + \cdots + k_1 + k_{2p-1} + k_{2p}} r_2^{k_2 + \cdots + k_2 + k_{2p-1} + k_{2p}} \cdots r_p^{k_p + \cdots + k_p + k_{2p-1} + k_{2p}}.$$
\[
\int_{-\pi}^{\pi} \exp\{\sum_{-k_{1,1} \cdots -k_{1,s} + k_{2,1} + \cdots + k_{2,s}} + \sum_{-k_{3,1} \cdots -k_{3,s} + k_{4,1} + \cdots + k_{4,s}} + \cdots + \sum_{-k_{2p-1,1} \cdots -k_{2p-1,s} + k_{2p,1} + \cdots + k_{2p,s}} \}\cdot \omega + i\omega
\]

\[
+ \sum_{k_{1,1} + \cdots + k_{1,s} - k_{2,1} - \cdots - k_{2,s}} \cdot i\xi_1 + \cdots
\]

\[
+ \sum_{k_{2p-1,1} + \cdots + k_{2p-1,s} - k_{2p,1} - \cdots - k_{2p,s}} i\xi_p \cdot d\omega
\]

\[
= \sum_{K} \sum_{k_{1,1} + \cdots + k_{1,s} + k_{2,1} + \cdots + k_{2,s}} \cdot \sum_{k_{3,1} + \cdots + k_{4,s}} \cdot \sum_{k_{2p-1,1} + \cdots + k_{2p,s}}
\]

where

\[
K = \{(k_{1,1}, \ldots, k_{2p,s}) \in \mathbb{N}^+)^{2ps} ; (-k_{1,1} - \cdots - k_{1,s} + k_{2,1} + \cdots + k_{2,s})
\]

\[
+ \cdots + (-k_{2p-1,1} - \cdots - k_{2p-1,s} + k_{2p,1} + \cdots + k_{2p,s}) + h = 0 \}.
\]

The above simplification is justified because for \((k_{1,1}, \ldots, k_{2p,s}) \notin K\), the corresponding integral is 0. Therefore, with \(\eta\) defined in Assumption 1.1,

\[
|\int_{-\pi}^{\pi} e^{ih\omega} | A(e^{i\omega})^{-2g} d\omega| \leq \sum_{K} (1 - \eta)^{k_{1,1} + \cdots + k_{2p,s}} = S, \text{ say.}
\]

Let

\[
t_1 = k_{1,1} + \cdots + k_{1,s} + k_{3,1} + \cdots + k_{3,s} + \cdots + k_{2p-1,1} + \cdots + k_{2p-1,s},
\]

\[
t_2 = k_{2,1} + \cdots + k_{2,s} + k_{4,1} + \cdots + k_{4,s} + \cdots + k_{2p,1} + \cdots + k_{2p,s} + h.
\]
Because \( t_1 = t_2 \) on \( K \), we have

\[
S = \sum_{t_2=h}^{\infty} \sum_{k_1,1=0}^{t_2-k_2,2=0} \cdots \sum_{k_2,p,s=0}^{t_2-k_2,1=0 t_1-k_1,r=0 \cdots -k_2,p-1,s-1}
\]

Now pick \( \lambda \in (1-\eta,1) \) and let \( M' = \max(j+1)^{2p} \left( (1-\eta) / \lambda \right)^{2j} \), then

\[
S \leq \sum_{t_2=h}^{\infty} (j+1)^{2p} \left[ (1-\eta) / \lambda \right]^{2j} \lambda^{2j-h} \leq \sum_{j=h}^{\infty} (j+1)^{2p} (1-\eta)^{2j-h}.
\]

Letting \( M = M'(1-\lambda^2) \), we have

\[
\left| \int_{-\pi}^{\pi} e^{ih\omega} |A(e^{i\omega})|^{-2s} \, d\omega \right| \leq M' \lambda^h. \quad (2.17)
\]

Since the right hand side of (2.17) does not depend on \( \theta \in \theta \), we have the desired declining property.

In Theorem 2.13, uniformly exponentially declining property of \( r_{\theta_1}(h) \), \( r_{\theta_1,\theta_2}(h) \), and \( r_{\theta_1,\theta_2,\theta_3}(h) \) is established.
Theorem 2.13. Under the assumptions of Theorem 2.10, \( r_{\theta_1}(h) \), \( r_{\theta_1 \theta_j}(h) \), and 
\( r_{\theta_1 \theta_j \theta_k}(h) \) are exponentially declining in \( h \) uniformly in \( \theta \in \Theta \).

Proof. Noting that all the partial derivatives of \(|A(e^{i\omega})|^2\) and \(|B(e^{i\omega})|^2\) are 
polynomials of \( \theta \), we know that all functions \( g_{\theta_1}(\omega), g_{\theta_1 \theta_j}(\omega), g_{\theta_1 \theta_j \theta_k}(\omega) \) are of the 
following form 

\[
\sigma^2 \sum_{k=-s_1}^{s_1} a_k(\theta)e^{ik\omega}/|A(e^{i\omega})|^{2s_2},
\]

where \( s_1 \) and \( s_2 \) are nonnegative integers and \( a_k(\theta), k = -s_1, \ldots, s_1 \) are polynomials of \( \theta \). Now the result follows from the observation that \( a_k(\theta), k = -s_1, \ldots, s_1 \) are bounded 
in \( \theta \) and by applying Lemma 2.12.

Finally we discuss uniform boundedness of matrix norms of partial 
derivatives of \( \Gamma_n \) and \( \Gamma_n^{-1} \). Let 

\[
\Gamma_{\theta_1} \equiv \frac{\partial \Gamma_n}{\partial \theta_1}, \quad \Gamma_{\theta_1 \theta_j} \equiv \frac{\partial^2 \Gamma_n}{\partial \theta_1 \partial \theta_j}, \quad \Gamma_{\theta_1 \theta_j \theta_k} \equiv \frac{\partial^3 \Gamma_n}{\partial \theta_1 \partial \theta_j \partial \theta_k}, \\
G_{\theta_1} \equiv \frac{\partial \Gamma_n^{-1}}{\partial \theta_1}, \quad G_{\theta_1 \theta_j} \equiv \frac{\partial^2 \Gamma_n^{-1}}{\partial \theta_1 \partial \theta_j}, \quad G_{\theta_1 \theta_j \theta_k} \equiv \frac{\partial^3 \Gamma_n^{-1}}{\partial \theta_1 \partial \theta_j \partial \theta_k}.
\]

Note that
\[ G_{\theta_i} = -\Gamma_n^{-1} \Gamma_{\theta_i} \Gamma_n^{-1}, \quad G_{\theta_i \theta_j} = 2\Gamma_n^{-1} \Gamma_{\theta_i} \Gamma_n^{-1} \Gamma_{\theta_j} \Gamma_n^{-1} - \Gamma_n^{-1} \Gamma_{\theta_i \theta_j} \Gamma_n^{-1}, \]  

(2.18)

and \( G_{\theta_i \theta_j \theta_k} \) is a polynomial of \( \Gamma_n^{-1}, \Gamma_{\theta_i}, \Gamma_{\theta_j}, \Gamma_{\theta_k}, \Gamma_{\theta_i \theta_j}, \Gamma_{\theta_i \theta_k}, \Gamma_{\theta_j \theta_k}, \Gamma_{\theta_i \theta_j \theta_k} \) of order 7. We do not give the expression for \( G_{\theta_i \theta_j \theta_k} \) because it is messy and the exact form is not needed.

**Theorem 2.14.** Under assumptions of Theorem 2.10,

i) The matrix norms of \( \Gamma_{\theta_i}, \Gamma_{\theta_i \theta_j}, \Gamma_{\theta_i \theta_j \theta_k}, \Gamma_{\theta_i \theta_j \theta_k}, \Gamma_{\theta_i \theta_j \theta_k} \), \( i, j, k = 1, 2, \ldots, p+q \) are bounded by some \( M_1 < \infty \) uniformly in \( n = 1, 2, \ldots, \theta \in \Theta \).

ii) The matrix norms of \( G_{\theta_i}, G_{\theta_i \theta_j}, G_{\theta_i \theta_j \theta_k}, G_{\theta_i \theta_j \theta_k}, G_{\theta_i \theta_j \theta_k} \), \( i, j, k = 1, 2, \ldots, p+q \) are bounded by some \( M_1 < \infty \) uniformly in \( n = 1, 2, \ldots, \theta \in \Theta \).

**Proof of i).** The proof is same as the proof of Theorem 2.10 except Theorem 2.13 is used instead of \( \sup_\theta |r(\theta)| \leq \sigma^2 M_1 \lambda^5 (1 - \lambda^2)^{-1} \). We give only a proof of boundedness of matrix norm of \( \Gamma_{\theta_i} \). Fix \( i \). By Theorem 2.13, \( r_{\theta_i}(h) \) is uniformly exponentially declining in \( h \). Therefore by the same argument used in the proof of Theorem 2.10, for any eigenvalue \( \nu \) of \( \Gamma_{\theta_i} \),

\[ \sup_{n, \theta} |\nu| < \infty. \]

Since \( \| \Gamma_{\theta_i} \| \) is the largest eigenvalue of \( \Gamma_{\theta_i} \), we have
\[ \sup_{n, \theta} \| G_{\theta, i} \| < \infty. \]

Proof of ii). This follows from (2.18), Theorem 2.10, and Theorem 2.14. We give only a proof of boundedness of matrix norm of \( G_{\theta, i} \). Fix \( i \). By (2.18)

\[ \sup_{n, \theta} \| G_{\theta, i} \| = \sup_{n, \theta} \| \Gamma_n^{-1} \Gamma_{\theta, i} \Gamma_n^{-1} \| \leq \sup_{n, \theta} \| \Gamma_{\theta, i} \| \sup_{n, \theta} \| \Gamma_n^{-1} \|^2. \]

By Theorem 2.14–i),

\[ \sup_{n, \theta} \| \Gamma_{\theta, i} \| < \infty. \]

By Theorem 2.10,

\[ \sup_{n, \theta} \| \Gamma_n^{-1} \| < \infty. \]

Therefore,

\[ \sup_{n, \theta} \| G_{\theta, i} \| < \infty. \] \qed
3. CONSISTENCY

In this chapter we shall show that the least squares estimator \((\hat{\theta}, \hat{\rho}, \hat{\sigma})\), the ordinary least squares estimator \((\bar{\theta}, \bar{\rho}, \bar{\sigma})\), and the maximum likelihood estimator \((\hat{\theta}, \hat{\rho}, \hat{\sigma})\) of \((\theta, \rho, \sigma)\) in model (1.1) are all strongly consistent. We also give a new proof of strong consistency of the least squares estimator, the ordinary least squares estimator, and the maximum likelihood estimator of the vector of parameters for the stationary model (1.4).

3.1. A Sufficient Condition for Strong Consistency

Wu (1981) gave a sufficient condition for the strong consistency of an estimator that is obtained by a minimization of a function of data and parameter.

Lemma 3.1. (Wu) Let \(X_n\) be an \(n\)-dimensional random vector whose distribution is indexed by some parameter \(\xi \in \mathbb{R}^k\) for fixed \(k\). Let \(H_n(\xi)\) be a function of \(X_n\) and \(\xi\). Assume \(\hat{\xi}_n\) is a minimizing value of \(H_n(\xi)\). Suppose for any \(\delta > 0\),

\[
\liminf_{n \to \infty} \inf_{\delta > 0} \left| \xi - \xi^0 \right| > 0 \quad \text{a.s.} \tag{3.1}
\]

Then

\[
\hat{\xi}_n \to \xi^0 \quad \text{a.s.}
\]

Proof. If \(\hat{\xi}_n \to \xi^0 \quad \text{a.s.}\) is not true, then there exists a \(\delta > 0\) such that \(P(\omega: \limsup_{n \to \infty} \left| \hat{\xi}_n(\omega) - \xi^0 \right| > \delta) > 0\). From the definition of \(\hat{\xi}_n\), this implies
contradicting (3.1).

Let \( \xi = (\theta, \rho) \) and let \( 1_A \) be the indicator function of a set \( A \). Also let

\[
H_n(\xi) = Q_n(\theta, \rho)1_{\theta} x R, \quad H_{2n}(\xi) = S_n(\theta, \rho)1_{\theta} x R, \quad \text{and} \quad H_{3n}(\xi) = [\text{det}(\theta)]^{1/n} Q_n(\theta, \rho)1_{\theta} x R. \]

By verifying condition (3.1) of Lemma 3.1 for \( H_{1n}(\xi) \), \( H_{2n}(\xi) \), \( H_{3n}(\xi) \) we can establish the strong consistency of estimators of parameters in model (1.1). Similarly we can establish the strong consistency of estimators in model (1.4). To establish the results we need several lemmas. Lemma 3.2 is a strong law of large numbers for a submartingale and can be found in Stout (1974, p. 153). For the definition of submartingale, see Definition A.6 in Appendix A.

**Lemma 3.2. (Chow)** Let \( (T_n, G_n \ n \geq 1) \) be a nonnegative submartingale with \( 0 < a_n \) increasing and \( G_{n-1} \) measurable for \( n \geq 1 \) \( (G_0 = \{\phi, \Omega\}) \). Suppose for some \( \alpha \geq 1 \),

\[
E(T_k^{\alpha}/a_k^{\alpha}) < \infty \quad \text{and} \quad \sum_{k=2}^{\infty} E[(T_k^{\alpha} - T_{k-1}^{\alpha})/a_k^{\alpha}] < \infty. \tag{3.2}
\]

Then

\[
\lim_{n \to \infty} T_n/a_n = 0 \quad \text{a.s.}
\]
In equation (3.12) in the proof of Lemma 3.6, it will be shown that $Z^T \Gamma_n^{-1} Z$ and $Z' D_n^\circ D_n Z$ can be approximated by $c(\theta) \sum_{i=1}^{n} e_i^2$ for some function satisfying $c(\theta^0) = 1$. In equation (4.24) and (4.25) in the proof of Lemma 4.3, it will be shown that $Y_i^T \Gamma_{n-1}^{-1} Z$ and $Y_i D_n^O D_n^O Z$ can be approximated by $\sum_{i=1}^{n-1} \sum_{i_2=1}^{i_1-1} e_{i_1} e_{i_2}$. In the proof of those approximations, Lemma 3.3 plays an important role.

Lemma 3.3. Let $\lambda \in (0,1)$ and let $\{e_n\}$ be an iid $(0, \sigma^2)$ sequence. Then

$$n^{-1}[|\lambda|e_1 e_2 + \cdots + e_{n-1} e_n| + \lambda^2|e_1 e_3 + \cdots + e_{n-2} e_n| + \cdots + \lambda^{n-1}|e_1 e_n|] \to 0 \ a.s.$$ 

Proof. Let $T_n = \lambda|e_1 e_2 + \cdots + e_{n-1} e_n| + \cdots + \lambda^{n-1}|e_1 e_n|$, $n \geq 2$, $T_1 = 0$. Let $G_n = \sigma(e_1, \ldots, e_n)$ be the $\sigma$-algebra generated by $e_1, \ldots, e_n$. From the conditional Jensen's inequality, we can see that $(T_n, G_n, n \geq 1)$ is a nonnegative submartingale (see Appendix 8.2.1). Now by the Cauchy–Schwartz inequality (see Appendix 8.2.2),

$$E(T_{n+1}^2 - T_n^2) \leq \sigma^4(\lambda^2 + \lambda^4 + \cdots + \lambda^{2n})$$

$$+ 2\lambda^3\{E[|(e_1 e_2 + \cdots + e_{n-1} e_n)(e_{n-1} e_{n+1})|$$

$$+ |(e_1 e_3 + \cdots + e_{n-2} e_n)(e_n e_{n+1})| + |e_n e_{n+1} e_{n-1} e_{n+1}|]$$

$$+ \lambda E[|(e_1 e_2 + \cdots + e_{n-1} e_n)(e_{n-2} e_{n+1})|]$$
+ \left| (e_1 e_4 + \cdots + e_{n-3} e_n)(e_n e_{n+1}) \right| + \left| e_n e_{n+1} e_{n-2} e_{n+1} \right|
+ \cdots

+ \lambda^{n-2} E\left[ \left| (e_1 e_2 + \cdots + e_{n-1} e_n)(e_1 e_{n+1}) \right| + \left| 0 \right| + \left| e_n e_{n+1} e_1 e_{n+1} \right| \right]

+ 2\lambda^5 \left\{ E\left[ \left| (e_1 e_3 + \cdots + e_{n-2} e_n)(e_{n-2} e_{n+1}) \right| \right]

+ \left| (e_1 e_4 + \cdots + e_{n-3} e_n)(e_{n-1} e_{n+1}) \right| + \left| e_{n-1} e_{n+1} e_{n-2} e_{n+1} \right|
+ \cdots

+ \lambda^{n-3} E\left[ \left| (e_1 e_3 + \cdots + e_{n-2} e_n)(e_1 e_{n+1}) \right| + \left| 0 \right| + \left| e_{n-1} e_{n+1} e_1 e_{n+1} \right| \right]\right\}

+ \cdots \cdots \cdots

+ 2\lambda^{2(n-1)+1} \left\{ E\left[ \left| e_1 e_n e_{n+1} \right| + 0 + \left| e_2 e_{n+1} e_{n+1} \right| \right] \right\}

\leq e^4 (1 - \lambda^2)^{-1} + 2e^4 (1 - \lambda)^{-1} \left[ 3\lambda^3 (n-1)^{1/2} + 3\lambda^5 (n-2)^{1/2} + \cdots + 3\lambda^{2n+1} \right].

Hence,

\sum_{n=1}^{\infty} E\left[ \left( \frac{T_{n+1}^2 - T_n^2}{n^2} \right) \right]

\leq e^4 (1 - \lambda^2)^{-1} \sum_{n=1}^{\infty} n^{-2} + 6e^4 (1 - \lambda)^{-1} \lambda \sum_{n=1}^{\infty} \sum_{i=1}^{n-1} \lambda^{2i(n-i)} 1/2 / n^2 < \infty

because
Thus by Lemma 3.2, with $a_n = n$ and $\alpha = 2$, the result follows. 

3.2. Order with Probability One

For notational convenience, we define order with probability one: $O_{wp}$ and $o_{wp}$.

**Definition 3.4.** Let $\{U_n\}$, $\{W_n\}$ be sequences of random variables such that $W_n > 0$ with probability one. We use the notation $U_n = O_{wp}(W_n)$ to denote those sequences of random variables such that $\limsup_{n \to \infty} \frac{|U_n/W_n|}{W_n} < \infty$ a.s. Likewise $U_n = o_{wp}(W_n)$ means that $\lim_{n \to \infty} \frac{U_n}{W_n} = 0$ a.s.

It is obvious that $O_{wp}(\cdot)$ and $o_{wp}(\cdot)$ have the following properties. Let $\{U_n\}$ and $\{W_n\}$ be sequences of positive random variables. Let $\{A_n\}$ and $\{B_n\}$ be sequences of random variables such that $A_n = O_{wp}(U_n)$ and $B_n = O_{wp}(W_n)$. Also let $\{C_n\}$ and $\{D_n\}$ be sequences of random variables such that $C_n = o_{wp}(U_n)$ and $D_n = o_{wp}(W_n)$. Then,

$$A_n + B_n = O_{wp}(\max(U_n, W_n)),$$

$$A_n B_n = O_{wp}(U_n W_n).$$
\[
C_n + D_n = o_{wp}(\max(U_n, W_n)),
\]
and
\[
C_n D_n = o_{wp}(U_n W_n). \tag{3.4}
\]

In Lemma 3.5, \(\sup_{\theta} |Y_1'D_n'M_n|\) and \(\sup_{\theta} |Z'D_n'M_n|\) are shown to be \(O_{wp}(1)\).

This together with Lemma 2.8 means that terms like

\[
n^{-1}Z'D_n'M_n(A^{-1} + M_n'M_n)^{-1}M_n'D_nZ
\]

and

\[
n^{-1}Y_1'D_n'M_n(A^{-1} + M_n'M_n)^{-1}M_n'D_nZ
\]

become negligible as \(n\) increases. Those terms are remainder terms when we approximate \(n^{-1}Z'D_nD_nZ\) and \(n^{-1}Y_1'D_nD_nZ\).

Lemma 3.5. Let \(D_n\) and \(M_n\) be given by (2.1). Then under Assumption 1.1,

\[
\sup_{\theta} |Y_1'D_n'M_n| = O_{wp}(1),
\]

and

\[
\sup_{\theta} |Z'D_n'M_n| = O_{wp}(1).
\]

Proof. By (2.5) we have

\[
|Y_1'D_n'M_n|^2 = \sum_{j=1}^{p+q} \left( \sum_{i=1}^n \sum_{k=0}^{n-i} d_{i+k,j} \right)^2.
\]
Observe that by Theorem 2.6 we have, for some $M < \omega$ and $\lambda \in (0,1)$,

$$\sup_{\theta} | \sum_{k=0}^{n-i} d_k m_{i+k,j} | \leq M^2 \sum_{k=0}^{\infty} \lambda^k \lambda^{i+k} = M^2(1-\lambda^2)^{-1}\lambda^i = M_1 \lambda^i,$$

say.

Therefore,

$$\sup_{\theta} | Y_i^2 D'M_n |^2 \leq \sum_{j=1}^{p+q} \sum_{i=1}^{n} | y_{i-1}^i | M_1^i | \lambda \lambda^i^2 = O_{wp}(1)$$

because

$$\sum_{i=1}^{n} | y_{i-1}^i | \lambda \lambda^i \leq \sum_{i=1}^{n} \lambda \lambda^i \left( \sum_{k=1}^{\infty} | z_k | + | y_0 | \right)$$

$$= \sum_{k=1}^{n-1} \lambda \lambda^i \left( \sum_{i=1}^{n} | z_k | + \sum_{i=1}^{n} \lambda \lambda^i | y_0 | \right)$$

$$\leq (1-\lambda)^{-1} \sum_{k=1}^{n-1} \lambda \lambda^{k+1} | z_k | + | y_0 | \right) = O_{wp}(1). \quad (3.5)$$

See Example A.5 in Appendix A. Also

$$\sup_{\theta} | Z'D_n^{-1} M_n |^2 \leq (p+q)M_1^2 \left( \sum_{i=1}^{n} | z_i | \lambda \lambda^i \right)^2 = O_{wp}(1). \quad \square$$
3.3. Strong Consistency in a Stationary Model

Wu's condition (3.1) in Lemma 3.1 for the least squares estimator and the ordinary least squares estimator in the stationary and invertible autoregressive moving average model (1.4) is verified in Lemma 3.6.

Lemma 3.6. Consider model (1.4). Let $D_n$ be given by (2.1) and let $Z$, $\Gamma_n$ be given by (1.10). Given $\delta > 0$, let $\theta_\delta = \{ \theta \in \Theta : | \theta - \theta' | \geq \delta \}$. Then, under Assumption 1.2, for any $\delta > 0$,

i) $\liminf_{n \to \infty} \inf_{\theta_\delta} \{ n^{-1}Z'(D_nD_n - D_0'D_0)Z \} > 0$ a.s.,

ii) $\liminf_{n \to \infty} \inf_{\theta_\delta} \{ n^{-1}Z'(\Gamma_n^{-1} - \Gamma_0^{-1})Z \} > 0$ a.s.,

where $D_0 = D_n(\theta')$ and $\Gamma_0 = \Gamma_n(\theta')$.

Proof of i). Proof is established by showing that $n^{-1}Z'D_nD_nZ$ can be approximated by $c(\theta) n^{-1} \sum_{i=1}^{n} e_i^2$ for some continuous function $c(\theta)$ such that $c(\theta') = 1$ and $c(\theta) > c(\theta')$ for $\theta \neq \theta'$. By (2.5) after some algebra (see Appendix 8.2.3), we have

$$Z'D_nD_nZ = \sum_{i_1 \leq n} \sum_{i_2 \leq n} a_{n,i_1,i_2} e_{i_1} e_{i_2}$$

where
\[ a_{n,i_1,i_2} = \sum_{i_1=\max(1,i_1)}^{n} \sum_{i=\max(1,i_1)}^{n} \sum_{s=0}^{n-\max(i,j)} \left( v_{i-i_1} v_{j-i_2} d_s d_{i-j} + s \right) \]

Notice that \( \{v_j^\circ\} \) is defined at \( \theta = \theta^\circ \) and \( \{d_j\} \) is defined at general \( \theta \), that is,

\[ v_j^\circ = v_j(\theta^\circ) \text{ and } d_j = d_j(\theta). \]

We can write

\[ Z' \mathbf{D}_n^\circ Z = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{n,i_1,i_2} e_{i_1} e_{i_2} + R_n, \]

where

\[ R_n = \sum_{i_1=1}^{n} \sum_{i_2=0}^{\infty} a_{n,i_1,i_2} e_{i_1} e_{i_2} + \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} a_{n,i_1,i_2} e_{i_1} e_{i_2} \]

\[ = R_{1n} + R_{2n} + R_{3n}, \text{ say.} \]

By Theorem 2.4 and Theorem 2.6 we can find \( M_1 < \infty \) and \( \lambda \in (0,1) \) (see Appendix 8.2.4) such that

\[ \sup_{\theta} |a_{n,i_1,i_2}| \leq M_1 \lambda^{-|i_1-i_2|} \]

for \( i_1 \geq 1 \) or \( i_2 \geq 1 \)

\[ \leq M_1 \lambda^{-|i_1-i_2|} \]

for \( i_1 \leq 0 \) and \( i_2 \leq 0 \)

From this, we conclude that \( \sup_{\theta} |R_{1n}|, \sup_{\theta} |R_{2n}|, \text{ and } \sup_{\theta} |R_{3n}| \) are all \( O_{\mathbb{W}_p}(1) \).

For example, see
For some $M_2 < \infty$ and $\lambda_1 \in (\lambda, 1)$ we have (see Appendix 8.2.5)

$$\sup_{\theta} |b_{n, i_1, i_2}| \leq M_2 (n - i_1)(n - i_2) \lambda_1^{2n - i_1 - i_2}.$$

Since

$$\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} (n - i_1)(n - i_2) \lambda_1^{2n - i_1 - i_2} |e_{i_1} e_{i_2}| = \left[ \sum_{i=1}^{n} (n - i) \lambda_1^{n-i} |e_i| \right]^2$$
we have

\[ \sup_\theta \left| \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{i_1-i_2} e_{i_1} e_{i_2} \right| = O_{wp}(1). \]

Note that for some \( M_3, M_4 < \infty \) (see Appendix 8.2.6) and \( \lambda_1 \in (\lambda, 1), \)

\[ \sup_\theta |a_{i_1-i_2}| \leq M_3 \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \lambda^{i_1+j-i_2+|i-j|+2s} \leq M_4 \lambda_1^{1-i_2}. \]

Therefore by Lemma 3.3,

\[ \sup_\theta \left| \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{i_1-i_2} e_{i_1} e_{i_2} \right| \]

\[ \leq 2M_4[\lambda_1 |e_1 e_2 + \cdots + e_{n-1} e_n| + \lambda_1^2 |e_1 e_3 + \cdots + e_{n-1} e_n| + \cdots + \lambda_1^{n-1} |e_1 e_n|] \]

= \( O_{wp}(n) \). Hence,

\[ \sup_\theta \left| \sum_{i=1}^{n} e_i^2 \right| = o_{wp}(n). \]

Now we need another expression for \( a_0 \). Note that
\[ a_0 = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{s=0}^{\infty} v_i^0 v_j^0 d_s d_{i-j} + s. \]

Consider a stationary process \( \{\hat{z}_t(\theta)\} \) defined below

\[ \hat{z}_t(\theta) = \sum_{j=0}^{\infty} d_j \hat{\epsilon}_{t-j}, \text{ where } \{\hat{\epsilon}_t\} \text{ is an iid (0,1) sequence.} \]

Then \( e_t = \sum_{j=0}^{\infty} v_j \hat{z}_{t-j}(\theta). \) (Notice that \( d_j = d_j(\theta) \) and \( v_j = v_j(\theta) \) depend on \( \theta \).

Define

\[ \hat{e}_t(\theta_1; \theta_2) = \sum_{j=0}^{\infty} v_j(\theta_2) \hat{z}_{t-j}(\theta_1), \theta_1, \theta_2 \in \Theta. \]  \hfill (3.9)

Now note that

\[ \text{Cov}(\hat{z}_i(\theta), \hat{z}_j(\theta)) = \sum_{s=0}^{\infty} d_s d_{i-j} + s \]  \hfill (3.10)

and

\[ a_0 = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} v_i^0 v_j^0 \text{Cov}(\hat{z}_i(\theta), \hat{z}_j(\theta)) = \text{Var}(\sum_{i=0}^{\infty} v_i^0 \hat{z}_i(\theta)) = \text{Var}(\hat{\epsilon}_0(\theta; \Theta)). \]  \hfill (3.11)

In (3.11) interchange of the limit operations is justified by the dominated convergence Theorem. Therefore we have

\[ \sup_{\Theta} |Z' D_n' D_n Z - \text{Var}(\hat{\epsilon}_0(\theta; \Theta))| \sum_{i=1}^{n} \epsilon_i^2 | = o_{\text{wp}}(n). \]  \hfill (3.12)

Thus,
Observe that \( \text{Var}(\hat{e}_0(\theta; \theta^2)) \) is continuous function of \( \theta \) (see Appendix 8.2.8), \( \theta \) is compact, and \( \text{Var}(\hat{e}_0(\theta; \theta^2)) > \text{Var}(\hat{e}_0(\theta^2; \theta^2)) \) for \( \theta \neq \theta^2, \theta \in \Theta \). From this we can say (see Appendix 8.2.9)

\[
\inf_{\delta} \{ \text{Var}(\hat{e}_0(\theta; \theta^2)) - \text{Var}(\hat{e}_0(\theta^2; \theta^2)) \} > 0. \tag{3.14}
\]

However by the strong law of large numbers, \( n^{-1} \sum_{i=1}^{n} e_i^2 \to (\sigma^0)^2 > 0 \) a.s. Therefore the term (3.13) converges almost surely to something positive and the result follows.

**Proof of ii).** Observe that

\[
n^{-1}Z'(\Gamma_n^{-1} - \Gamma_0^{-1})Z \\
\geq n^{-1}Z'(D_n D_n - D_0 D_0)Z - n^{-1} \sup_{\theta} |Z'D_n M_n| \sup_{n, \theta} \| (A^{-1} + M_n^T M_n)^{-1} \|.
\]

Therefore, Lemma 3.6-ii) follows from Lemma 3.6-i), Lemma 2.8, and Lemma 3.5.\)

To establish the consistency of estimators of \( \sigma^2 \) in model (1.1) and in model (1.4) we need the following Lemma.
Lemma 3.7. Assume the condition of Lemma 3.6.

i) If \( \theta_n \rightarrow \theta^0 \) a.s. then \( n^{-1} Z' \Gamma_n^{-1}(\theta_n) Z \rightarrow (\sigma^0)^2 \) a.s.

ii) If \( \theta_n \rightarrow \theta^0 \) a.s. then \( n^{-1} Z' D_n(\theta_n) D_n(\theta_n) Z \rightarrow (\sigma^0)^2 \) a.s.

iii) If \( \theta_n \rightarrow \theta^0 \) in probability then i) and ii) are true with 'a.s.' replaced by 'in probability'.

Proof. Let \( \hat{\epsilon}_l(\theta_1; \theta_2) \), \( \theta_1, \theta_2 \in \theta \) be defined by (3.9). Then by (3.10) -- (3.12),

\[
\begin{align*}
    n^{-1} Z' \Gamma_n^{-1}(\theta_n) Z &= \text{Var}(\hat{\epsilon}_o(\theta_o; \theta^0)) \ n^{-1} \sum_{i=1}^{n} e_i^2 + o_{wp}(1) \\
    &= \left\{ \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{s=0}^{\infty} \nu_i^o \nu_j^o d_{s}(\theta_n) d_{|i-j|+s}(\theta_n) \right\} n^{-1} \sum_{i=1}^{n} e_i^2 + o_{wp}(1).
\end{align*}
\]

Letting \( n \rightarrow \infty \), by the strong law of large numbers we have

\[
\lim_{n \rightarrow \infty} n^{-1} Z' \Gamma_n^{-1}(\theta_n) Z = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{s=0}^{\infty} \nu_i^o \nu_j^o d_{s}(\theta^0) d_{|i-j|+s}(\theta^0) (\sigma^0)^2 \quad \text{a.s.}
\]

\[
= \text{Var}(\hat{\epsilon}_o(\theta^0; \theta^0))(\sigma^0)^2 = (\sigma^0)^2.
\]

In the above equation, interchange of limit operations at several stages is justified by the dominated convergence Theorem, by the fact that for some \( M < \infty \) and \( \lambda \in (0,1) \),
and by the continuity of \( d_s(\theta) \) in \( \theta \) (see Appendix 8.2.10). This establishes Lemma 3.7-i).

Lemma 3.7-ii) follows from Lemma 3.7-i) and the observation

\[
\sup_{\theta} | Z' D_n' D_n Z - Z' \Gamma_n^{-1} Z | = o_{wp}(1).
\]

See Lemma 3.5 and Lemma 2.8.

Proof of iii) is similar to the proof of i) and ii).

---

We now give a new proof of the strong consistency of the least squares estimator, the ordinary least squares estimator, and the maximum likelihood estimator for the parameters of model (1.4). There are two different proofs of strong consistency in the literature. One is due to Hannan (1973) and the other is due to Rissanen and Caines (1979).

Hannan used a frequency domain approximation \( \sum_{j=1}^{n} I_n(\omega_j)/g(\omega_j, \theta) \) of

\[
Z' \Gamma_n^{-1} Z,
\]

where \( I_n(\cdot) \) is the periodogram of \( \{z_1, z_2, ..., z_n\} \), that is,

\[
I_n(\omega_j) = 2n^{-1} \{ (\sum_{t=1}^{n} z_t \cos \omega_j)^2 + (\sum_{t=1}^{n} z_t \sin \omega_j)^2 \}, \quad \omega_j = 2\pi j/n,
\]

and

\[
\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{s=0}^{\infty} \sup_{\theta} \left| v_i^0 v_j^0 d_s(\theta) d_{i-j} + s(\theta) \right| \leq M \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{s=0}^{\infty} \lambda^{i+j+s|i-j|} + 2s < \infty,
\]
and \( g(\omega; \theta) \) is the spectral density of the process \( \{z_t\} \) given in (2.13). Hannan obtained a uniform strong convergence of \( n^{-1} \sum_{j=1}^{n} I_n(\omega_j)/g(\omega_j; \theta) \). From this uniform strong convergence, Hannan deduced the strong convergence of the least squares estimator and the maximum likelihood estimator.

Rissanen and Caines obtained Kalman filter predictor \( \{e_t^k(Z; \theta)\} \) of \( \{e^k\} \) and approximated \( Z' \Gamma_n^{-1} Z \) by \( \sum_{t=1}^{n} e_t^2(Z; \theta)/\sigma_t^2(\theta) \), where \( \sigma_t^2(\theta) = \text{var}(e_t(Z; \theta)) \). They calculated \( -2n^{-1} \text{log likelihood} \) by

\[
-2n^{-1} \text{log likelihood} = \sum_{t=1}^{n} \{2 \log \text{det} \sigma_t(\theta) + e_t^2(Z; \theta)/\sigma_t^2(\theta)\}.
\]

By investigating properties of \( e_t(Z; \theta) \) they proved

\[
n^{-1} L_n(\theta, \rho, \sigma) \rightarrow 2 \log \text{det} \sigma_0(\theta) + \text{E} e_0^2(Z, \theta)/\sigma_0^2(\theta) \text{ a.s.}
\]

uniformly in \( \theta \in \Theta \), where \( e_0(Z; \theta) = \sum_{j=0}^{\infty} d_j(\theta) z_j \) and \( \sigma_0^2(\theta) = \text{var}(e_0(Z; \theta)) \). From the uniform almost sure convergence of \( n^{-1} L_n(\theta, \rho, \sigma) \) they deduced the strong consistency of maximum likelihood estimator of \( \theta \).

However our approach is direct in the sense that we express \( Q_n(\theta, \rho), S_n(\theta, \rho), L_n(\theta, \rho, \sigma) \) in terms of \( \{z_t\} \) and expand them using \( z_t = \sum_{j=0}^{\infty} v_j e_{t-j} \) and \( e_t = \sum_{j=0}^{\infty} d_j z_{t-j} \).

In our proof, uniform boundedness of the matrix norms of \( \Gamma_n, \Gamma_n^{-1} \), derivatives of \( \Gamma_n \) and \( \Gamma_n^{-1} \) and the uniformly exponential declining properties of \( \{v_j\} \) and \( \{d_j\} \)
Theorem 3.8. Let model (1.4) and Assumption 1.2 hold. Let \( \Gamma_n = \Gamma_n(\theta) \) be the covariance matrix of \( Z = (z_1, ..., z_n)' \) and let \( D_n \) be given by (2.1). Let \( \theta, \bar{\theta}, \hat{\theta} \) be the values minimizing \( Z'\Gamma_n^{-1}Z, \ Z'D_n'D_nZ, \) and \( [\det \Gamma_n]^{1/n}Z'\Gamma_n^{-1}Z \) respectively. Define \( \hat{\bar{\sigma}}^2, \bar{\sigma}^2, \) and \( \hat{\sigma}^2 \) by \( n^{-1}Z'\Gamma_n^{-1}(\theta)Z, \ n^{-1}Z'D_n'(\bar{\theta})D_n(\bar{\theta})Z, \) and \( n^{-1}Z'\Gamma_n^{-1}(\hat{\theta})Z \) respectively. Then \( (\theta, \hat{\sigma}^2), (\bar{\theta}, \bar{\sigma}^2), \) and \( (\hat{\theta}, \hat{\sigma}^2) \) converge almost surely to \( (\theta_0, (\sigma_0)^2) \).

Proof. The strong consistency of \( \theta, \bar{\theta}, \hat{\sigma}^2, \) and \( \hat{\sigma}^2 \) follows from Lemma 3.1, Lemma 3.6 and Lemma 3.7. To show the consistency of \( \hat{\theta} \), let \( \delta > 0 \) be given. By Theorem 2.10, we can find \( \nu_0 > 0, \) independent of \( \theta, \) such that the eigenvalues \( \nu_1 \leq \nu_2 \leq ... \leq \nu_n \) of \( \Gamma_n \) are all greater than or equal to \( \nu_0 \) for all \( n = 1, 2, ... \) and \( \theta \in \Theta. \) Therefore

\[
[\det \Gamma_n]^{1/n} = [\nu_1 \nu_2 \cdots \nu_n]^{1/n} \geq \nu_0 > 0 \text{ for all } n = 1, 2, ..., \text{ and } \theta \in \Theta.
\]

Consequently, with the notation \( \Gamma_0^0 = \Gamma_n(\hat{\theta}) \) and \( \Theta_\delta = \{ \theta \in \Theta : |\theta - \theta_0| \geq \delta \}, \)

\[
\liminf_{n \to \infty} \inf_{\Theta_\delta} n^{-1}\{ [\det \Gamma_n]^{1/n}Z'\Gamma_n^{-1}Z - [\det \Gamma_0^0]^{1/n}Z'\Gamma_0^{0-1}Z \}
\geq \nu_0 \liminf_{n \to \infty} \inf_{\Theta_\delta} n^{-1}\{ Z'\Gamma_n^{-1}Z - Z'\Gamma_0^{0-1}Z \}
\]

\[
+ \liminf_{n \to \infty} n^{-1/2}\{ [\det \Gamma_n]^{1/n} - [\det \Gamma_0^0]^{1/n} \} \quad \text{a.s.,} \quad (3.17)
\]

where we have used Lemma 3.6, Theorem 2.11, and Lemma 3.7. Hence by Lemma
3.1, \( \hat{\theta} \rightarrow \theta \) a.s. The consistency of \( \hat{\sigma}^2 \) follows from the consistency of \( \hat{\theta} \) and Lemma 3.7.

3.4. Orders of Quadratic Forms in \( \mathbb{Z} \)

We present Lemma 3.9 and Lemma 3.10 for use in checking condition (3.1) in Lemma 3.1 for the parameters of model (1.1). Lemma 3.9 is a preliminary step for Lemma 3.10. Lemma 3.10 shows that if a sequence of random variable \( \{A_j\} \) and a sequence of real number \( \{a_j\} \) are exponentially declining in \( j \) then the orders of

\[
\sum_{j=1}^{n} \sum_{k=1}^{n} |A_{k-j}z_j|, \quad \sum_{j=1}^{n} \sum_{k=1}^{n} |A_{n-j}z_j|, \quad \sum_{j=1}^{n} \sum_{k=1}^{n} a_{n-j}z_j
\]

are the same as those of

\[
\sum_{j=1}^{n} \sum_{k=1}^{n} |A_{k-j}e_k|, \quad \sum_{j=1}^{n} \sum_{k=1}^{n} |A_{n-j}e_k|, \quad \sum_{j=1}^{n} \sum_{k=1}^{n} a_{n-j}e_k
\]

where \( W_k = e_1 + \cdots + e_k, \) \( k = 1, 2, \ldots \), and \( W_0 = 0 \).
Lemma 3.9. Let \( \{e_n\} \) be an iid \((0, \sigma^2)\) sequence and let \( \lambda \in (0,1), \epsilon > 0 \). Then

\[
\sum_{i=1}^{n} \sum_{j=1}^{i-1} \lambda^{i-j} e_i e_j = o_{wp}(n^{1/2} + \epsilon).
\]

Proof. Define

\[
U_i = (\sum_{j=1}^{i-1} \lambda^{i-j} e_j)e_i \quad \text{for } i \geq 1
\]

\[
= 0 \quad \text{for } i = 0.
\]

Then \( E(U_i | e_{i-1}, \ldots, e_1) = 0 \) for every \( i = 1, 2, \ldots \).

Therefore \( V_n = \sum_{i=1}^{n} i^{-1/2} - \epsilon U_i \) is a martingale. Note that

\[
[E|V_n|^2] \leq E(V_n^2) = \sum_{i=1}^{n} i^{-1-2\epsilon} E(U_i^2) = \sum_{i=1}^{n} i^{-1-2\epsilon} \sum_{j=1}^{i-1} \lambda^{2(i-j)} E(e_i^2 e_j^2)\]

\[
\leq \sum_{i=1}^{n} i^{-1-2\epsilon}(1-\lambda^2)(\sigma_0^4) < \infty \quad \text{for every } n = 1, 2, \ldots.
\] (3.18)

Hence by the martingale convergence theorem given in A.7 in Appendix A, \( V_n \) converges a.s. Consequently, by Kronecker's lemma given in A.15 in Appendix A,

\[
n^{-1/2} - \epsilon \sum_{i=1}^{n} U_i = n^{-1/2} - \epsilon \sum_{i=1}^{n} \sum_{j=1}^{i-1} \lambda^{i-j} e_i e_j \to 0 \quad \text{a.s.}
\]
Lemma 3.10. Let \( \{A_j\} \) be a sequence of random variables satisfying \( |A_j| \leq M \lambda^j \) for some \( M < \infty \) and \( \lambda \in (0,1) \). Let \( \{y_t\}, \{z_t\} \) be defined by model (1.1) and assume that \( \{z_t\} \) is stationary and invertible. Then,

i) \[
\sum_{j=1}^{n} \sum_{k=1}^{n} |A_{j-k}| z_k z_j = O_{wp}(n)
\]
and

ii) \[
\sum_{j=1}^{n} \sum_{k=1}^{n} |A_{n-j}| z_k z_j = O_{wp}(n).
\]

If \( \{a_j\} \) is a sequence of real numbers satisfying \( |a_j| \leq M \lambda^j \) then,

iii) \[
\sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-j} z_k z_j = O_p(n^{1/2}),
\]

\[
\sum_{j=1}^{n} \sum_{k=j}^{n} a_{n-k} z_k z_j = O_p(n^{1/2}),
\]

iv) \[
\sum_{j=1}^{n} \sum_{k=1}^{j} a_{j-k} y_{k-1} z_j = O_p(n),
\]

\[
\sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{k-j} y_{k-1} z_j = O_p(n),
\]

v) \[
\sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-j} |y_{k-1} z_j| = O_p(n^{3/2}),
\]

\[
\sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} |y_{k-1} z_j| = O_p(n^{3/2}).
\]
Proof. First we outline the proof. In the proof of i), using (2.6), we express
\[ \prod_{j=1}^{n} \sum_{k=j}^{n} |A_{k-j} z_{j} z_{k}| \] in terms of \{ \{e_{i_1} e_{i_2}\}\} and show that the coefficient of \{e_{i_1} e_{i_2}\} is exponentially declining in |\text{sign}(i_{1})| i_{1} - |\text{sign}(i_{2})| i_{2}|, where \text{sign}(i) = 1 if i \geq 0 and -1 if i < 0. For i_{1} > 0 and i_{2} > 0 apply Lemma 3.9 and the strong law of large numbers to get the order of the corresponding terms. For i_{1} \leq 0 or i_{2} \leq 0 apply the result of example A.5 given in Appendix A that states

\[ \text{if } \sum_{n=1}^{\infty} |a_{n}| < \infty \text{ then } \sum_{n=1}^{\infty} a_{n} e_{n} \text{ converges a.s.} \]

In the proof of iii) we express \[ \prod_{j=1}^{n} \sum_{k=j}^{n} a_{n-j} z_{j} z_{k} \] in terms of \{e_{i_1} e_{i_2}\}. The coefficient of \{e_{i_1} e_{i_2}\} has exponentially declining property as is given in (3.21). The resultant equation contains four terms for \[ [i_1 \leq i_2, i_1 \leq n], [i_2 \leq 0 < i_1 < n], [i_1 \leq 0 < i_2 < n], [i_1 < 0 < i_2 < 0] \] respectively. Next we show that the expectations of the three term for \[ [i_2 \leq 0 < i_1 < n], [i_1 \leq 0 < i_2 < n], [i_1 \leq 0 \text{ and } i_2 \leq 0] \] are all 0. Also the variances of the three terms are all O(n). The case \[ [1 \leq i_1, i_2 \leq n] \] is partitioned into three cases \[ [1 \leq i_2 < i_1 \leq n], [1 \leq i_1 < i_2 \leq n], [1 \leq i_1 = i_2 \leq n] \]. The variances of the terms for \[ [1 \leq i_2 < i_1], [1 \leq i_1 < i_2 \leq n] \] are O(1) and O(n) respectively and the expectations for the terms are all 0. The term for \[ [1 \leq i_1 = i_2 \leq n] \] is O_{p}(1). In the proof of iv) we express \[ \prod_{j=1}^{n} \sum_{k=1}^{j} a_{j-k} y_{k-1} z_{j} \] in terms of \{e_{i_1} e_{i_2}\} and \{y_{i_1} z_{j}\}. The dominating term is the term for \[ [1 \leq i_1, i_2 \leq n] \] and the coefficient for \{e_{i_1} e_{i_2}\} is bounded. From the boundedness of the coefficient of \{e_{i_1} e_{i_2}\} we deduce \[ \sum_{j=1}^{n} \sum_{k=1}^{j} a_{j-k} y_{k-1} z_{j} = O_{p}(n) \]. In
the proof of $v$) we show that the dominating part in

$$
\sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-j} |y_{k-1} z_{j}| \quad \text{is} \quad \sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-j} |v_{j-i_1} e_{i_1} e_{i_2}| = O_p(n^{3/2}).
$$

which is $O_p(n^{3/2})$. Also we show that the dominating part in

$$
\sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} |y_{k-1} z_{j}| \quad \text{is} \quad \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-j} |v_{j-i_1} e_{i_1} e_{i_2}| = O_p(n^{3/2}).
$$

which is $O_p(n^{3/2})$. We now give the details of proof.

Proof of $i$). First we consider $\sum_{j=1}^{n} \sum_{k=j}^{n} A_{k-j} z_{k} z_{j}$. Using $z_{j} = \sum_{i \leq j} v_{j-i} e_{i}$, we have

(see Appendix 8.2.11),

$$
\sum_{j=1}^{n} \sum_{k=j}^{n} A_{k-j} z_{k} z_{j} = \sum_{k=\max(i_1, i_2)}^{n} \sum_{j=i_2}^{k} A_{k-j} v_{k-i_1} v_{j-i_2} e_{i_1} e_{i_2}
$$

$$
+ \sum_{k=i_2}^{n} \sum_{j=i_2}^{k} A_{k-j} v_{k-i_1} v_{j-i_2} e_{i_1} e_{i_2}
$$

$$
+ \sum_{k=1}^{n} \sum_{j=1}^{k} A_{k-j} v_{k-i_1} v_{j-i_2} e_{i_1} e_{i_2}
$$

$$
+ \sum_{k=1}^{n} \sum_{j=1}^{k} A_{k-j} v_{k-i_1} v_{j-i_2} e_{i_1} e_{i_2}.
$$

(3.19)
We can find $\lambda_1 \in (\lambda, 1)$ and $M_1, M_2, M_3, M_4 < \infty$ (see Appendix 8.2.12) such that

$$\sum_{k=\max(i_1,i_2)}^n \sum_{j=i_2}^k |A_{k-j}v^0_{k-i_1}v^0_{j-i_2}| \leq M_1 \lambda_1^{i_1-i_2} \text{ for } i_1, i_2 \geq 1,$$

$$\sum_{k=i_2}^n \sum_{j=i_2}^k |A_{k-j}v^0_{k-i_1}v^0_{j-i_2}| \leq M_2 \lambda_1^{i_2-i_1} \text{ for } i_1 \leq 0 < i_2,$$

$$\sum_{k=1}^n \sum_{j=1}^k |A_{k-j}v^0_{k-i_1}v^0_{j-i_2}| \leq M_3 \lambda_1^{i_1-i_2} \text{ for } i_2 \leq 0 < i_1,$$

and

$$\sum_{k=1}^n \sum_{j=1}^k |A_{k-j}v^0_{k-i_1}v^0_{j-i_2}| \leq M_4 \lambda_1^{-i_1-i_2} \text{ for } i_1, i_2 \leq 0. \quad (3.20)$$

Therefore,

$$\sum_{j=1}^n \sum_{k=j}^n |A_{k-j}v_{k-j}||$$

$$\leq M_1 \sum_{i_1=1}^n \sum_{i_2=1}^n \lambda_1^{i_1-i_2} |e_{i_1} e_{i_2}| + M_2 \sum_{i_1 \leq 0} \sum_{i_2=1}^n \lambda_1^{i_2-i_1} |e_{i_1} e_{i_2}|$$

$$+ M_3 \sum_{i_1=1}^n \sum_{i_2 \leq 0} \lambda_1^{i_1-i_2} |e_{i_1} e_{i_2}| + M_4 \sum_{i_1 \leq 0} \sum_{i_2 \leq 0} \lambda_1^{-i_1-i_2} |e_{i_1} e_{i_2}|.$$

Observe that, by Lemma 3.9 and by the strong law of large numbers,

$$\sum_{i_1=1}^n \sum_{i_2=1}^n \lambda_1^{i_1-i_2} |e_{i_1} e_{i_2}| = 2 \sum_{i_1=1}^{i_1-1} \sum_{i_2=1}^{i_2-1} \lambda_1^{i_1-i_2} |e_{i_1} e_{i_2}| + \sum_{i_1=1}^n \sum_{i_2=1}^n \lambda_1^{i_1-i_2} |e_{i_1} e_{i_2}| + \sum_{i_1=1}^n \sum_{i_2=1}^n \lambda_1^{i_1-i_2} |e_{i_1} e_{i_2}|.$$
\[
= 2 \sum_{i_1=1}^{n} \sum_{i_2=1}^{i_1-1} \lambda_1^{i_1-i_2} \left[ |e_1| - E|e_1| \right] \left[ |e_2| - E|e_2| \right] \\
+ 2E|e_1| \sum_{i_1=1}^{n} \sum_{i_2=1}^{i_1-1} \lambda_1^{i_1-i_2} \left[ |e_1| + |e_2| \right] \\
- 2(E|e_1|)^2 \sum_{i_1=1}^{n} \sum_{i_2=1}^{i_1-1} \lambda_1^{i_1-i_2} + \sum_{i=1}^{n} e_i^2
\]

\[
= O_{wp}(n) + O_{wp}(n) - O(n) + O_{wp}(n) = O_{wp}(n),
\]

\[
\sum_{i_1 \leq 0}^{n} \sum_{i_2=1}^{n} \lambda_1^{i_2-i_1} |e_1 e_2| = (\sum_{i_1 \leq 0}^{n} \lambda_1^{-i_1} |e_1|)(\sum_{i_2=1}^{n} \lambda_1^{i_2} |e_2|) = O_{wp}(1),
\]

\[
\sum_{i_1=1}^{n} \sum_{i_2 \leq 0}^{n} \lambda_1^{i_2-i_1} |e_1 e_2| = O_{wp}(1) \quad \text{by symmetry},
\]

and

\[
\sum_{i_1 \leq 0}^{n} \sum_{i_2 \leq 0}^{n} \lambda_1^{-i_2-i_1} |e_1 e_2| = O_{wp}(1).
\]

Therefore,

\[
\sum_{j=1}^{n} \sum_{k=j}^{n} |A_{k-j} z_k z_j| = O_{wp}(n).
\]

Hence,

\[
\sum_{j=1}^{n} \sum_{k=1}^{n} |A_{k-j} z_k z_j| \leq 2 \sum_{j=1}^{n} \sum_{k=j}^{n} |A_{k-j} z_k z_j| = O_{wp}(n).
\]

Proof of ii). Next observe that
Proof of iii). Observe that \[ \sum_{j=1}^{n} \sum_{k=1}^{n} |a_{n-j}^{k} z_{j}| = \sum_{j=1}^{n} \sum_{k=1}^{n} |\lambda^{n-j} z_{j}| = \sum_{k=1}^{n} |z_{k}| = O_{wp}(n). \]

(3.19) if \( A_{k-j} \) is replaced by \( a_{n-k} \). Therefore for, some \( M'_{1}, M'_{2}, M'_{3}, M'_{4} < \infty \) and \( \lambda \in (\lambda, 1) \) (see Appendix 8.2.13),

\[
\begin{align*}
|a_{n,i_{1},i_{2}}| &= \left| \sum_{k=\text{max}(i_{1},i_{2})}^{n} \sum_{j=i_{2}}^{k} a_{n-k}^{i_{1}} v_{i_{2}}^{j-i_{2}} \right| \leq M'_{1} \lambda^{n-i_{1}} \quad \text{for } 1 \leq i_{1}, i_{2} \leq n, \\
&\quad \text{say } \quad \text{for } i_{1} \leq 0 < i_{2} \leq n, \quad \text{for } i_{2} \leq 0 < i_{1} \leq n, \\
&\quad \text{say } \quad \text{for } i_{1} \leq 0 < i_{2} \leq n, \quad \text{for } i_{2} \leq 0 < i_{1} \leq n, \\
&\quad \text{say } \quad \text{for } i_{1} \leq 0 < i_{2} \leq n, \quad \text{for } i_{2} \leq 0 < i_{1} \leq n.
\end{align*}
\]

From (3.21) we have (see Appendix 8.2.14)

\[
\begin{align*}
\text{Var} \left( \sum_{i_{1}=1}^{n} \sum_{i_{2}=1}^{n} a_{n,i_{1},i_{2}} e_{i_{1}} e_{i_{2}} \right) &= O(n), \\
\text{Var} \left( \sum_{i_{1}=1}^{n} \sum_{i_{2}=1}^{n} a_{n,i_{1},i_{2}} e_{i_{1}} e_{i_{2}} \right) &= O(n).
\end{align*}
\]
Therefore, we have (see Appendix 8.2.14)

\[
\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{n,i_1,i_2} e_{i_1} e_{i_2} + \sum_{i_1 \leq 0}^{n} \sum_{i_2=1}^{n} a_{n,i_1,i_2} e_{i_1} e_{i_2} + \sum_{i_1 \leq 0}^{n} \sum_{i_2 \leq 0}^{n} a_{n,i_1,i_2} e_{i_1} e_{i_2} = O_p(n^{1/2}).
\]

The remaining thing is to show that

\[
\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{n,i_1,i_2} e_{i_1} e_{i_2} = O_p(n^{1/2}).
\]

However, this follows from the observations:

\[
\text{Var} \left( \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{n,i_1,i_2} e_{i_1} e_{i_2} \right) \leq (M_n')^2 \sigma^4 \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \lambda_1^{2(n-i_1)} = O(n),
\]

\[
\text{Var} \left( \sum_{i_1=1}^{n} \sum_{i_2=i_1+1}^{n} a_{n,i_1,i_2} e_{i_1} e_{i_2} \right) \leq (M_n')^2 \sigma^4 \sum_{i_1=1}^{n} \sum_{i_2=i_1+1}^{n} \lambda_1^{2(n-i_1)} = O(1),
\]

and

\[
\sum_{i=1}^{n} a_{n,i} e_{i}^2 \leq M_n' \sum_{i=1}^{n} \lambda_1^{n-i} e_{i}^2 = O_p(1). \tag{3.22}
\]

Therefore,

\[
\sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-j} z_k z_j = O_p(n^{1/2}).
\]

Also

\[
\sum_{j=1}^{n} \sum_{k=j}^{n} a_{n-j} z_k z_j = \sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-k} z_k z_j = O_p(n^{1/2}).
\]
Proof of iv). Now we show \( \sigma \sum_{j=1}^{n} a_{j-k} y_{k-1} z_{j} = O_{P}(n) \). By using \( z_{j} = \sum_{i \leq j} e_{i} \)

and \( y_{k-1} = \sum_{j=1}^{k-1} z_{j} + y_{0} \), we have (see Appendix 8.2.15)

\[
\sum_{j=1}^{n} \sum_{k=1}^{n} a_{j-k} y_{k-1} z_{j} = \sum_{i_{1} \leq n} \sum_{i_{2} \leq n-1} b_{n,i_{1},i_{2}} e_{i_{1}} e_{i_{2}} + \sum_{j=1}^{n} \sum_{k=1}^{n} a_{j-k} y_{0} z_{j}
\]

where

\[
b_{n,i_{1},i_{2}} = \sum_{j=\max(1,i_{1})}^{\min(n,i_{1})} \sum_{k=\max(1,i_{2}+1)}^{\min(n,i_{2})} a_{j-k} y_{0} z_{j}
\]

We can show that there are \( \lambda \in (0,1) \) and \( M_{1}^{''}, M_{2}^{''}, M_{3}^{''}, M_{4}^{''} < \infty \) such that (see Appendix 8.2.16)

\[
|b_{n,i_{1},i_{2}}| \leq M_{1}^{''} \quad \text{for } 1 \leq i_{1}, i_{2} \leq n,
\]

\[
\leq M_{2}^{''} \lambda^{-i_{2}} \quad \text{for } i_{2} \leq 0 < i_{1} \leq n,
\]

\[
\leq M_{3}^{''} \lambda^{-i_{1}} \quad \text{for } i_{1} \leq 0 < i_{2} \leq n,
\]

\[
\leq M_{4}^{''} \lambda^{-i_{1}+i_{2}} \quad \text{for } i_{1} \leq 0, i_{2} \leq 0.
\]

Observe that

\[
\text{Var}(\sum_{i_{1}=1}^{n} \sum_{i_{2}=1}^{n} b_{n,i_{1},i_{2}} e_{i_{1}} e_{i_{2}}) \leq M_{1}^{''} 2^{4} n^{2},
\]

\( i_{1} \neq i_{2} \)
\[
\left| \sum_{i=1}^{n} b_{n,i,i} e_i^2 \right| \leq M_1^n \sum_{i=1}^{n} e_i^2 = \mathcal{O}_p(n),
\]

\[
\text{Var}\left( \sum_{i=1}^{n} \sum_{i_2=1}^{n} b_{n,i,i_2} e_{i_1} e_{i_2} \right) \leq M_2^n \sigma^4 \sum_{i=1}^{n} \lambda^{-i_2} = \mathcal{O}(n),
\]

and

\[
\text{Var}\left( \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} b_{n,i_1,i_2} e_{i_1} e_{i_2} \right) = \mathcal{O}(n), \text{ by symmetry},
\]

and

\[
\left| \sum_{i_1=1}^{n} \sum_{i_2=0}^{n} b_{n,i_1,i_2} e_{i_1} e_{i_2} \right| \leq M_4^n \sum_{i_1=0}^{n} \sum_{i_2=0}^{n} \lambda^{-i_1-i_2} |e_{i_1} e_{i_2}| = \mathcal{O}_p(1). \tag{3.24}
\]

Hence,

\[
\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} b_{n,i_1,i_2} e_{i_1} e_{i_2} = \mathcal{O}_p(n),
\]

\[
\sum_{i_1=1}^{n} \sum_{i_2=0}^{n} b_{n,i_1,i_2} e_{i_1} e_{i_2} = \mathcal{O}_p(n^{1/2}),
\]

and

\[
\sum_{i_1=0}^{n} \sum_{i_2=1}^{n} b_{n,i_1,i_2} e_{i_1} e_{i_2} = \mathcal{O}_p(n^{1/2}),
\]

and

\[
\sum_{i_1=0}^{n} \sum_{i_2=0}^{n} b_{n,i_1,i_2} e_{i_1} e_{i_2} = \mathcal{O}_p(1). \tag{3.25}
\]

Also

\[
\left| \sum_{j=1}^{n} \sum_{k=1}^{j} a_{j-k} y_0 z_j \right| \leq |y_0| \sum_{j=1}^{n} \sum_{k=1}^{j} \lambda^{j-k} |z_j| = \mathcal{O}(n).
\]

Therefore, from (3.23) and (3.25), we have

\[
\sum_{j=1}^{n} \sum_{k=1}^{j} a_{j-k} y_{k-1} z_j = \mathcal{O}(n).
\]
Next we show that
\[ \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{k-j} y_{k-1} z_j = O_p(n). \]

Some algebra yields (see Appendix 8.2.17)
\[
\sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{k-j} y_{k-1} z_j = \sum_{i_1 \leq n} \sum_{i_2 \leq n-1} c_{n,i_1,i_2} e_{i_1} e_{i_2} + \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{k-j} y_0 z_j
\]

where
\[ c_{n,i_1,i_2} = \sum_{j=\max(1,i_1)}^{n-1} \sum_{k=\max(j+1,i_2+1)}^{k-1} \sum_{s=\max(1,i_2)}^{a_{k-j} y_0 z_j}. \]

We can find \( M_1', M_2', M_3', M_4' < \infty \) and \( \lambda_1 \in (\lambda, 1) \) such that (see Appendix 8.2.18)
\[ |c_{n,i_1,i_2}| \leq M_1', M_2', \lambda^{-i_2}, M_3', \lambda_1^{-i_1}, M_4', \lambda^{-i_1-i_2} \]

for each case of
\[ [1 \leq i_1, i_2 \leq n-1], [i_2 \leq 0 < i_1 \leq n-1], [i_1 \leq 0 < i_2 \leq n-1], [i_1 \leq 0, i_2 \leq 0] \]

respectively.

By the same argument used to show \( \sum_{j=1}^{n} \sum_{k=1}^{j} a_{j-k} y_{k-1} z_j = O_p(n) \), we have
\[ \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{k-j} y_{k-1} z_j = O_p(n). \]
The text reads:

Proof of $v$). Note that

\[
\sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{k-j} y_{k-1} z_{j} = O_p(n).
\]

It is easy to show that the second term in the right hand side of (3.27) is $O_p(n)$.

Now note that

\[
\sum_{s=1}^{k-1} \sum_{i=1}^{k-1} z_{j} c_{0} - \sum_{s=k}^{\infty} v_{s-i} e_{i} + z_{j} \sum_{i=1}^{k-1} \sum_{s=1}^{\infty} v_{s-i} e_{i}
\]

\[
= c_{0} z_{j} \sum_{i=1}^{k-1} e_{i} + z_{j} \sum_{i=1}^{k-1} \sum_{s=k}^{\infty} v_{s-i} e_{i}
\]

\[
+ z_{j} \sum_{i=1}^{k-1} \sum_{s=1}^{\infty} v_{s-i} e_{i}
\]

(3.28)
We know

\[ \sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-j} | \sum_{i_1 < k} v_{j-i_1} e_{i_1} \sum_{i_2=1}^{k-1} e_{i_2} | = O_p(n), \]

and

\[ \sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-j} | z_j \sum_{i_1=k}^{\infty} \sum_{i_2=1}^{k-1} v_{s-i_1} e_{i_2} | = O_p(n^{3/2}), \]

and

\[ \sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-j} | z_j \sum_{i_1=k}^{\infty} \sum_{r=1}^{k-1} v_{s-r} e_{r} | = O_p(n). \]  (3.29)

See Appendix 8.2.19. Therefore, if we show

\[ S_n = \sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-j} | \sum_{i_1=k}^{j} v_{j-i_1} e_{i_1} \sum_{i_2=1}^{k-1} e_{i_2} | = O_p(n^{3/2}) \]  (3.30)

we are done. Observe that, by the Holder inequality,

\[ E[ \sum_{i_1=k}^{j} v_{j-i_1} e_{i_1} \sum_{i_2=1}^{k-1} e_{i_2} ] \leq [E[ \sum_{i_1=k}^{j} v_{j-i_1} e_{i_1} ]^2 E[ \sum_{i_2=1}^{k-1} e_{i_2} ]^{1/2} ] \leq [\sigma \sum_{j=0}^{\infty} v_{j}^{o2} k^{1/2}]. \]

Hence, with \( M = (\sigma^4 \sum_{j=0}^{\infty} v_{j}^{o2})^{1/2} \) and by the Cauchy–Schwartz inequality,

\[ E[S_n] \leq M \sum_{j=1}^{n} \lambda^{n-j} k^{1/2} \leq M \sum_{j=1}^{n} \lambda^{n-j} (\sum_{k=1}^{j} k^{1/2}) \leq M \sum_{j=1}^{n} \lambda^{n-j}^{3/2} = O(n^{3/2}). \]

Therefore \( S_n = O_p(n^{3/2}) \) and hence

\[ \sum_{j=1}^{n} \sum_{k=1}^{j} a_{n-j} | y_{k-1} z_j | = O_p(n^{3/2}). \]
Next we show
\[ \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} |y_{k-1}z_j| = O_p(n^{3/2}). \quad (3.31) \]

Note that
\[ \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} |y_{k-1}z_j| \leq \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} \sum_{s=1}^{k-1} \sum_{i=1}^{s} \sum_{i=1}^{s} z_{s,i} + y_0 \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} |z_j|. \quad (3.32) \]

It is easy to show that the second term of the right hand side of (3.32) is \( O_p(n) \).

Now note that by (3.28),
\[ \sum_{s=1}^{k-1} z_{s,j} = c_0 (\sum_{i=1}^{s} v_{i}^{0} e_{i} + z_j (\sum_{i=1}^{k-1} \sum_{s=k}^{k-1} v_{s-i}^{0} e_{i} + \sum_{i=1}^{k-1} \sum_{s=1}^{k-1} v_{s-i}^{0} e_{i})). \]

We know
\[ \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} \sum_{i=1}^{k-1} \left| \sum_{i=1}^{s} v_{i}^{0} e_{i} \right| = O_p(n^{3/2}), \]
\[ \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} \sum_{i=1}^{k-1} \left| \sum_{s=1}^{k-1} v_{s-i}^{0} e_{i} \right| = O_p(n), \]
and
\[ \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} \sum_{i=1}^{k-1} \left| \sum_{s=1}^{k-1} v_{s-i}^{0} e_{i} \right| = O_p(1). \quad (3.34) \]
See Appendix 8.2.20. Therefore, if we show
\[
S'_n = \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} \| \sum_{j-i_1 \leq 1}^{i_1} \sum_{i_2 \geq 1}^{i_2} k-1 \| = O_p(n^{3/2})
\] (3.35)
we are done. Observe that by the Holder inequality,
\[
E \left| \sum_{i_1=1}^{j} v_{j-i_1}^{i_1} e_{i_1} \right|^{k-1} \leq \left[ E \left( \sum_{j-i_1 \leq 1}^{i_1} v_{j-i_1}^{i_1} e_{i_1} \right)^2 \right]^{1/2} \left[ E \left( \sum_{i_2 \geq 1}^{i_2} e_{i_2} \right)^2 \right]^{1/2}
\leq \left( \sigma^4 \sum_{j=0}^{\infty} v_{j}^{2k} \right)^{1/2}.
\]
Hence with \( M = \left( \sigma^4 \sum_{j=0}^{\infty} v_{j}^{2k} \right)^{1/2} \) and by the Cauchy–Schwartz inequality,
\[
E[S'_n] \leq M \sum_{j=1}^{n} \sum_{k=j+1}^{n} \lambda^{n-k} \lambda^{1/2}
\leq M \sum_{j=1}^{n} \left( \sum_{k=j+1}^{n} \lambda^{2(n-k)} \right)^{1/2} \left( \sum_{k=j+1}^{n} \right)^{1/2}
\leq M(1-\lambda^2)^{-1/2} \left( \sum_{j=1}^{n} (n-j)(n+j+1)/2 \right)^{1/2} = O(n^{3/2}).
\] (3.36)
Therefore \( S'_n = O_p(n^{3/2}) \) and hence \( \sum_{j=1}^{n} \sum_{k=j+1}^{n} a_{n-k} \| y_{k-i} z_j \| = O_p(n^{3/2}) \).
3.5. Limiting Behavior of \( Y_{1 n}^{-1} Z \) and \( Y_1 Y_1 \)

In Lemma 3.11, we establish the orders of \( \sup_{\theta} |Y_1' \Gamma_n^{-1} Z| \) and \( \sup_{\theta} |Y_1' D_n' D_n Z| \) both with probability one and in probability. The orders of \( \sup_{\theta} |Y_1' \Gamma_n^{-1} Z| \) and \( \sup_{\theta} |Y_1' D_n' D_n Z| \) are shown to be same as the order of \( \sum_{j=1}^{n} W_{j-1} e_j \) where \( W_j = e_1 + \cdots + e_j, j = 1, 2, \cdots \). Applying Strassen's law of the iterated logarithm for Brownian motion (see Example A.13 in Appendix A), we obtain

\[
\sum_{j=1}^{n} W_{j-1} e_j = O_{wp}(\log \log n), \tag{3.37}
\]

which gives the order with probability one. Also, since

\[
\mathbb{E}(\sum_{j=1}^{n} W_{j-1} e_j) = 0
\]

and

\[
\text{Var}(\sum_{j=1}^{n} W_{j-1} e_j) = \sigma^4 n(n-1)/2
\]

we have

\[
\sum_{j=1}^{n} W_{j-1} e_j = O_p(n) \tag{3.38}
\]

which gives the order in probability.
Lemma 3.11. Let model (1.1) and Assumption 1.1 hold. Then

\[ \sup_\theta |Y_1' \Gamma_n^{-1} D_n Z| = O_{\text{wp}}(\text{nloglog}n), \]

\[ \sup_\theta |Y_1' \Gamma_n^{-1} Z| = O_p(n), \]

\[ \sup_\theta |Y_1' D_n' D_n Z| = O_{\text{wp}}(\text{nloglog}n), \]

and

\[ \sup_\theta |Y_1' D_n' D_n Z| = O_p(n). \]

Proof. First we show \( \sup_\theta |Y_1' D_n' D_n Z| = O_{\text{wp}}(\text{nloglog}n) \). From (2.5) after a considerable algebra (see Appendix 8.2.21),

\[
Y_1' D_n' D_n Z = d^2 \sum_{j=1}^{n} \sum_{k=1}^{j-1} z_k z_j + \sum_{j=1}^{n} \sum_{k=1}^{j-1} b_{n,k,j} z_k z_j + \sum_{j=1}^{n} \sum_{k=1}^{n} c_{n,k,j} z_k z_j + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=0}^{n-\max(i,j)} z_j d_s |i-j| + s' \tag{3.39}
\]

where

\[ d = (\sum_{k=0}^{\infty} d_k), \]

\[ b_{n,k,j} = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} (-d_n-j+1+k_1 d_n-j+1+k_1+k_2-d_{k_1} d_{j-k+k_1+k_2}^--) \]
\[ + d_{n-j+1+k_1} d_{n-k+1+k_1+k_2} - d_{n-j+k_1} d_{n-j+1+k_1+k_2} \]
\[-d_{k_1} d_{n-j+1+k_2} + d_{n-j+k_1} d_{n-j+1+k_2} \]

and

\[ c_{n,k,j} = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} (d_{k_1} d_{k+1-j+k_1+k_2} - d_{n-k+k_1} d_{n+1-j+k_1+k_2} \]
\[-d_{k_1} d_{n-j+1+k_2} + d_{n-k+k_1} d_{n-j+1+k_2} \] (3.40)

Now note that by Theorem 2.6, we can find \( M_1, M_2, M_3, M_4, M_5 < \infty \) and \( \lambda \in (0,1) \) such that

\[ \sup_{\theta} |d^2| \leq M_1 < \infty, \]
\[ \sup_{\theta} |b_{n,k,j}| \leq M_2 \lambda^{j-k} + M_3 \lambda^{n-j}, \]

and

\[ \sup_{\theta} |c_{n,k,j}| \leq M_4 \lambda^{k-j} + M_5 \lambda^{n-k}. \] (3.41)

Also note that (see Appendix 8.2.22)

\[ \sup_{\theta} \left| \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=0}^{n-\max(i,j)} \sum_{s=0}^{n} \sum_{s=0}^{n} z_j d_j d_{|i-j|+s} \right| = O_{wp}(n). \]

Therefore, by Lemma 3.10,

\[ \sup_{\theta} |Y_1 D_n D_n Z| \leq M_1 \left| \sum_{j=1}^{n} \sum_{k=1}^{n} z_k z_j \right| + O_{wp}(n). \]
If we show that \( \sum_{j=1}^{n} \sum_{k=1}^{j-1} z_k z_j = O_{wp}(n \log \log n) \), we are done. After some algebra given in Appendix 8.2.23,

\[
\sum_{j=1}^{n} \sum_{k=1}^{j-1} z_k z_j = \sum_{j=1}^{n} \sum_{i_1=1}^{n-1} \sum_{i_2=1}^{j-1} \sum_{j=\max(i_1,i_2+1)}^{n} \sum_{s=\max(s-i_1-i_2)}^{n} v_{j-i_1}^{i_1} v_{s-i_2}^{i_2} e_{i_1} e_{i_2}
\]

\[
+ \sum_{i_1-1}^{n} \sum_{i_2=1}^{j-1} \sum_{j=\max(i_1,i_2+1)}^{n} \sum_{s=\max(s-i_1-i_2)}^{n} v_{j-i_1}^{i_1} v_{s-i_2}^{i_2} e_{i_1} e_{i_2}
\]

\[
+ \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{j=\max(i_1,i_2+1)}^{n} \sum_{s=\max(s-i_1-i_2)}^{n} v_{j-i_1}^{i_1} v_{s-i_2}^{i_2} e_{i_1} e_{i_2}
\]

\[
+ \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{j=\max(i_1,i_2+1)}^{n} \sum_{s=\max(s-i_1-i_2)}^{n} v_{j-i_1}^{i_1} v_{s-i_2}^{i_2} e_{i_1} e_{i_2} \tag{3.42}
\]

\[
= S_1 + S_2 + S_3 + S_4, \text{ say.}
\]

After some more algebra (see Appendix 8.2.24),

\[
S_1 = (\sum_{j=0}^{\infty} v_{j}^{2}) \sum_{i_1=1}^{n} \sum_{i_2=1}^{i_1-1} e_{i_1} e_{i_2} + \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{n,i_1,i_2} e_{i_1} e_{i_2} + \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} c_{i_1,i_2} e_{i_1} e_{i_2} \tag{3.43}
\]

where

\[
a_{n,i_1,i_2} = -\sum_{j=n+1}^{\infty} \sum_{s=\max(s-i_1-i_2)}^{\infty} v_{j-i_1}^{i_1} v_{s-i_2}^{i_2}
\]

and
From the fact that $|a_{n,i_1,i_2}| \leq M_6 \lambda^{-i_1}$ and $|c_{i_1,i_2}| \leq M_7 \lambda^{-i_2}$ for some $M_6, M_7 < \infty$ (see Appendix 8.2.25) and by Lemma 3.10, we have

$$
\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{n,i_1,i_2} e_{i_1} e_{i_2} = O_{wp}(n)
$$

and

$$
\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} c_{i_1,i_2} e_{i_1} e_{i_2} = O_{wp}(n). \quad (3.44)
$$

By the Strassen's law of the iterated logarithm, we can show (see Example A.13 in Appendix A)

$$
\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} e_{i_1} e_{i_2} = O_{wp}(n \log \log n). \quad (3.45)
$$

Consequently, $S_1 = O_{wp}(n \log \log n)$. Now observe that, for some $M_8, M_9, M_{10} < \infty$,

$$
| \sum_{j=1}^{n} \sum_{s=1}^{j-1} v^o_{j-i_1} v^o_{s-i_2} |, \quad | \sum_{j=1}^{n} \sum_{s=1}^{j-1} v^o_{j-i_1} v^o_{s-i_2} |, \quad | \sum_{j=1}^{n} \sum_{s=1}^{j-1} v^o_{j-i_1} v^o_{s-i_2} |
$$

are less than $M_8 \lambda^{-i_2-i_1}, M_9 \lambda^{-i_2}, M_{10} \lambda^{-i_2-i_1}$ respectively (see Appendix 8.2.26).

Therefore,
\[ S_2 = O_{wp}(1), \quad S_3 = O_{wp}(n), \quad S_4 = O_{wp}(1). \]

For details see Appendix 8.2.27. Hence,

\[
\sum_{j=1}^{n} \sum_{k=1}^{j-1} z_k z_j = \sum_{j=0}^{\infty} v_j^2 \sum_{i_1=1}^{n} \sum_{i_2=1}^{i_1-1} e_{i_1} e_{i_2} + O_{wp}(n)
\]

\[ = O_{wp}(n\log\log n) + O_{wp}(n) = O_{wp}(n\log\log n). \quad (3.46) \]

Next we show \[ \sup_{\theta} |Y_1^{-1} \Gamma_n^{-1} Z| = O_{wp}(n\log\log n). \]

From (2.2) and Lemma 3.5,

\[
\sup_{\theta} |Y_1^{-1} \Gamma_n^{-1} Z| \leq \sup_{\theta} |Y_1' D_n' D_n Z| + \sup_{\theta} |Y_1' D_n' M_n| \sup_{n, \theta} \|(\Lambda_n^{-1} + M_n' M_n)^{-1}\| \sup_{\theta} |M_n' D_n Z|
\]

\[ = O_{wp}(n\log\log n) + O_{wp}(1) = O_{wp}(n\log\log n). \quad (3.47) \]

Finally, following similar arguments with \[ O_{wp}(n\log\log n), \quad O_{wp}(n), \quad \text{and} \quad O_{wp}(1) \]
replaced by \[ O_p(n), \quad O_p(n), \quad \text{and} \quad O_p(1) \] respectively, we have

\[ \sup_{\theta} |Y_1^{-1} \Gamma_n^{-1} Z| = O_p(n) \]

and

\[ \sup_{\theta} |Y_1' D_n' D_n Z| = O_p(n). \]
We now consider the order of $Y_1 Y_1$. Fuller (1976) showed that, for $q = 0$,

$$\frac{1}{n} \sum_{t=1}^{n} y_t^2 - c_0^2 \sum_{t=1}^{n} W_t^2 = O_p(n^{3/2}). \quad (3.48)$$

Since $\sum_{t=1}^{n} W_t^2/n^2$ converges in distribution to $\int_0^1 W^2(r)dr$ with $W(\cdot)$ the standard Brownian motion on $[0,1]$, $\sum_{t=1}^{n} W_t^2$ is $O_p(n^2)$. However in order to verify the condition (3.1) for strong consistency we need the limiting behavior of $\liminf$ of $\sum_{t=1}^{n} W_t^2$. In Lemma 3.12, we show that

$$\frac{1}{n} \sum_{t=1}^{n} y_t^2 - c_0^2 \sum_{t=1}^{n} W_t^2 = O_{wp}(n^{3/2}(\log n)^{1/2}), \quad (3.49)$$

where $c_0 = \sum_{j=0}^{\infty} \delta_j$. By one version of the law of the iterated logarithm due to Donsker and Varadhan (1977), we also know (see also Lai and Wei (1982, p. 364))

$$\liminf_{n \to \infty} \sum_{t=1}^{n} W_t^2 n - 2 \log \log n = \sigma^2/4 \text{ a.s.} \quad (3.50)$$

Combining (3.49) and (3.50),

$$\liminf_{n \to \infty} Y_1 Y_1 n^{-2} \log \log n = c_0^2 (\sigma^2)^{1/2}/4.$$

In Appendix 8.2.33, we show that $c_0^2$ is positive. Also by the Strassen's law of the
iterated logarithm (see Example A.13 in Appendix A),

\[ \limsup_{n \to \infty} \sum_{t=1}^{n} W_t^2 / n \log \log n = 4 \sigma^2 / \pi^2 \text{ a.s.} \]

Therefore,

\[ \limsup_{n \to \infty} \sum_{i=1}^{n} Y_i / n \log \log n = 4c_0^2 \sigma_0^2 / \pi^2. \quad (3.51) \]

**Lemma 3.12.** We consider model (1.1) with Assumption 1.1. We have,

i) \( \liminf_{n \to \infty} \sum_{i=1}^{n} Y_i^2 / n \log \log n = 1/4 \cdot c_0^2 (\sigma_0^2)^2 > 0 \text{ a.s.} \)

ii) \( Y_1^2 Y_1^{-2} \sum_{t=1}^{n} W_t^2 = O_p(n^{3/2}) \),

where \( c_0 = \sum_{j=0}^{\infty} v_j^0 \) and \( W_t = \sum_{j=1}^{t} e_j \).

**Proof of i.** The proof is an adaptation of Fuller's (1976) proof of

\[ \sum_{t=1}^{n} y_t^2 - c_0^2 \sum_{t=1}^{n} W_t^2 = O_p(n^{3/2}) \]

to the proof of

\[ \sum_{t=1}^{n} y_t^2 - c_0^2 \sum_{t=1}^{n} W_t^2 = O_{wp}(n^{3/2}(\log n)^{1/2}). \]

See Fuller (1976, pp. 374–377). We can write
\[ y_t = \sum_{j=1}^{t} \sum_{i=0}^{\infty} v_i^0 e_{j-i} + y_0 \]

\[ = \sum_{j=1}^{t} \sum_{i=0}^{\infty} v_i^0 e_{j-i} - \sum_{j=1}^{t} \sum_{i=0}^{\infty} v_i^0 e_{j-1-j+1} + \sum_{j=0}^{\infty} e_{-j} \sum_{i=j+1}^{j+t} v_i^0 + y_0 \]

\[ = c_0 W_t + U_t + R_t, \quad (3.52) \]

where \( U_t = \sum_{j=1}^{t} g_j e_{t-j+1} \), \( g_j = \sum_{i=j}^{\infty} v_i^0 \), \( R_t = \sum_{j=0}^{\infty} e_{-j} g_{t,j} + y_0 \), and \( g_{t,j} = \sum_{i=j+1}^{j+t} v_i^0 \).

We know that, for some \( M_1 \leq \alpha, \lambda \in (0,1) \), and for all \( t = 0, 1, 2, \ldots \),

\[ |v_i^0| \leq M_1 \lambda^i, \ i=0,1,\ldots, \ |g_j| \leq M_1 \lambda^j, \ \text{and} \ |g_{t,j}| \leq M_1 \lambda^j \ \ j=0,1,\ldots . \]

We have

\[ \sum_{t=1}^{n} y_t^2 - c_0^2 \sum_{t=1}^{n} W_t^2 \]

\[ = 2c_0 \sum_{t=1}^{n} W_t U_t + \sum_{t=1}^{n} U_t^2 + 2c_0 \sum_{t=1}^{n} W_t R_t + 2 \sum_{t=1}^{n} U_t R_t + \sum_{t=1}^{n} R_t^2. \quad (3.53) \]

We will show that (3.53) is \( O_{wp}(n^{3/2}(\log \log n)^{1/2}) \). After some algebra (see Appendix 8.2.28),

\[ \sum_{t=1}^{n} W_t U_t = \sum_{j=1}^{n} (\sum_{t=j}^{j+1} g_{t,j+1} e_j^2 + \sum_{j_1=1}^{j} \sum_{j_2=1}^{j_1-1} (\sum_{t=j_1}^{j_2} g_{t,j_2+1}) e_{j_1} e_{j_2} \quad (3.54) \]
Observe that, for some $M_2, M_3 < \infty$,

$$|S_1| \leq M_2 \sum_{j=1}^{n} \left( \sum_{t=j}^{\infty} \lambda^{t-j+1} e_j^2 \right) \leq M_3 \sum_{j=1}^{n} e_j^2 = O_{wp}(n),$$

and

$$|S_3| \leq M_2 \sum_{j_1=1}^{n} \sum_{j_2=j_1+1}^{\infty} |e_{j_1} e_{j_2}| = O_{wp}(n).$$

Therefore,

$$\sum_{t=1}^{n} W_t U_t = O_{wp}(n \log \log n).$$

(3.55)
We know (see Appendix 8.2.29),

\[ | \sum_{t=\max(j_1,j_2)}^{n} \varepsilon_t \varepsilon_{t+1} \varepsilon_{t+2} | \leq M_2^2 \lambda^{|j_1-j_2|}. \]

Therefore, by Lemma 3.10,

\[ \sum_{t=1}^{n} \sum_{j_1=1}^{\infty} \sum_{j_2=1}^{\infty} \left( \sum_{t=\max(j_1,j_2)}^{n} \varepsilon_t \varepsilon_{t+1} \varepsilon_{t+2} \right) \varepsilon_{j_1} \varepsilon_{j_2} = O_{wp}(n). \]  

(3.56)

We know, for some \( M_4 \) and \( M_4' < \infty \) (see Appendix 8.2.30),

\[ | \sum_{t=1}^{n} g_{t,j} | \leq M_4 n \lambda^j. \]

and

\[ | \sum_{t=1}^{n} g_{t,j_1} g_{t,j_2} | \leq M_4' n \lambda_i^j. \]

Therefore,

\[ \sum_{t=1}^{n} R_t^2 = \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \left( \sum_{t=1}^{n} g_{t,j_1} g_{t,j_2} \right) e_{j_1} e_{j_2} + 2y_0 \sum_{j=0}^{\infty} \sum_{t=1}^{n} g_{t,j} e_{-j} + n y_0^2 = O_{wp}(n). \]

(3.57)

We know, for some \( M_5 < \infty \) (see Appendix 8.2.31),

\[ | \sum_{t=j_1}^{n} g_{t-j_1+1} g_{t-j_2} | \leq M_5 \lambda_i^j. \]

Therefore (see Appendix 8.2.31),

\[ \sum_{t=1}^{n} U_t R_t = \sum_{j_1=1}^{\infty} \sum_{j_2=0}^{\infty} \left( \sum_{t=j_1}^{n} g_{t-j_1+1} g_{t-j_2} \right) e_{j_1} e_{j_2} + y_0 \sum_{t=1}^{n} U_t = O_{wp}(n). \]  

(3.58)
Finally, we have (see Appendix 8.2.32)

\[ \sum_{t=1}^{n} W_t R_t = -\left( \sum_{t=1}^{n} W_t \right) \left( \sum_{j=0}^{\infty} g_{j+1} e^{-j} \right) \]

\[ + \sum_{j_1=1}^{\infty} \sum_{j_2=0}^{\infty} \left( \sum_{t=j_1}^{n} g_{j_2+t+1} e_{j_1} e^{-j_2} \right) + y_0 \sum_{t=1}^{n} W_t. \]

By Strassen's law of the iterated logarithm, we know that (see Example A.13 in Appendix A)

\[ \sum_{t=1}^{n} W_t = O_{wp}(n^{3/2} (\log \log n)^{1/2}). \]  
(3.59)

For some \( M_6 < \alpha \),

\[ |\sum_{t=j_1}^{n} g_{j_2+t+1}| \leq M_6^{j_1+j_2}. \]

Therefore,

\[ \sum_{j_1=1}^{\infty} \sum_{j_2=0}^{\infty} \left( \sum_{t=j_1}^{n} g_{j_2+t+1} e_{j_1} e^{-j_2} \right) = O_{wp}(1). \]

Thus,

\[ \sum_{t=1}^{n} W_t R_t = O_{wp}(n^{3/2} (\log \log n)^{1/2}). \]  
(3.60)

Hence, from (3.53) — (3.60), we have

\[ \sum_{t=1}^{n} y_t^2 - c_0^2 \sum_{t=1}^{n} W_t^2 = O_p(n^{3/2} (\log \log n)^{1/2}). \]  
(3.61)
Because the process \( \{z_t\} \) is invertible and stationary, we have (see Appendix 8.2.33)

\[
c_0 = \sum_{j=0}^{\infty} v_j^o \neq 0.
\]

Consequently, by (3.50) and (3.61),

\[
\liminf_{n \to \infty} n^{-2} \log n \sum_{t=1}^{n} y_t^2 = c_0^2 \liminf_{n \to \infty} n^{-2} \log n \sum_{t=1}^{n} W_t^2 = 4^{-1} c_0^2 \sigma^2 > 0 \text{ a.s.}
\]

Proof of ii) can be obtained by following the same argument as i) if we replace \( O_{wp} \) by \( O_p \) and \( \log \log n \) by one.

Since

\[
\sum_{t=1}^{n} W_t^2/n^2 \Rightarrow (\sigma^0)^2 \int_0^{\infty} W^2(r)dr,
\]

from (3.49), the limiting distribution of \( Y_1^2Y_1/n^2 \) follows.

**Corollary 3.13.** Consider model (1.1). Under Assumption 1.1,

\[
Y_1^2Y_1/n^2 \Rightarrow \sigma^0 c_0^2 \int_0^{\infty} W^2(r)dr,
\]

where \( W(\cdot) \) is the standard Brownian motion on \([0,1]\), \( c_0 = \sum_{j=0}^{\infty} v_j^o \), and \( v_j^o \) are defined in (2.6) and evaluated at \( \theta = \theta^0 \).
3.6. Strong Consistency

In Lemma 3.14, we use Lemma 3.11 and Lemma 3.12 to verify the condition (3.1) for the strong consistency of estimators for the parameters in model (1.1).

**Lemma 3.14.** Consider model (1.1). Let \( Q_n(\theta, \rho) \) and \( S_n(\theta, \rho) \) be defined by (1.12) and (1.19). Suppose Assumption 1.1 holds. Let

\[
\theta_\delta = \{ (\theta, \rho) \in \Theta \times \mathbb{R} : |(\theta', \rho) - (\theta^0', \rho^0)| \geq \delta \}.
\]

Then for any \( \delta > 0 \),

i) \( \liminf_{n \to \infty} n^{-1} \inf_{\theta_\delta} \{ Q_n(\theta, \rho) - Q_n(\theta^0, \rho^0) \} > 0 \) a.s.

and

ii) \( \liminf_{n \to \infty} n^{-1} \inf_{\theta_\delta} \{ S_n(\theta, \rho) - S_n(\theta^0, \rho^0) \} > 0 \) a.s.

**Proof of i).** The function \( Q_n(\theta, \rho) - Q_n(\theta^0, \rho^0) \) increases in the \( \rho \) direction at a faster rate than in the directions associated with \( \Theta \). Therefore it is necessary to recognize this fact in the proof. We do this by partitioning the parameter space. Given \( n = 1, 2, \ldots \), partition the set \( \Theta_\delta \) into two sets \( A_n \) and \( B_n \), where

\[
A_n = \{ (\theta, \rho) \in \Theta \times \mathbb{R} : |(\theta', \rho) - (\theta^0', \rho^0)| \geq \delta, |\rho - \rho^0| < n^{-1/3} \}.
\]
and
\[ B_n = \{(\theta, \rho) \in \theta \times \mathbb{R} : |(\theta', \rho) - (\theta^0, \rho^0)| \geq \delta, |\rho - \rho^0| \geq n^{-1/3}\}. \]

We show result i) by showing that
\[
\liminf_{n \to \infty} n^{-1} \inf_{A_n} \{Q_n(\theta, \rho) - Q_n(\theta^0, \rho^0)\} > 0 \text{ a.s.}
\]
and
\[
\liminf_{n \to \infty} n^{-1} \inf_{B_n} \{Q_n(\theta, \rho) - Q_n(\theta^0, \rho^0)\} > 0 \text{ a.s.}
\]

First consider \( A_n = \inf_{A_n} \{Q_n(\theta, \rho) - Q_n(\theta^0, \rho^0)\} \). Note that if \(|(\theta', \rho) - (\theta^0, \rho^0)| \geq \delta\) and \(|\rho - \rho^0| < n^{-1/3}\) then there is an \( n_0 \) such that, for every \( n > n_0 \) and \((\theta, \rho) \in A_n\), \(|\theta - \theta^0| > \delta/2\). Therefore, for \( n > n_0 \),
\[
a_n = \inf_{A_n} \{Z'(\Gamma_n^{-1} - \Gamma_n^{-1})Z + 2(1-\rho)Y_n^{1}Y_n^{-1}Z + (1-\rho)^2Y_n^{1}Y_n^{-1}Y_n^{1}\}
\]
\[
\geq \inf_{|\theta - \theta^0| \geq \delta/2} \{Z'(\Gamma_n^{-1} - \Gamma_n^{-1})Z\} - 2n^{-1/3} \sup_{\theta} |Y_n^{1}Y_n^{-1}Z| - O_{wp}(n^{2/3/\log\log n}), \tag{3.62}
\]
where we have used nonnegativity of \((1-\rho)^2Y_n^{1}Y_n^{-1}Y_n^{1}\) and Lemma 3.11. Therefore, by Lemma 3.6,
\[ \liminf_{n \to \infty} n^{-1} a_n \geq \liminf_{n \to \infty} \inf \n \to \infty \left[ a_{n+1} \right] + O(W_p(n^{-1/3} \log \log n) > 0 \]
> 0 \ a.s. \quad (3.63) \]

Next we consider \( b_n = \inf_{B_n} \{ Q_n(\theta, \rho) - Q_n(\theta^0, \rho^0) \} \). By Theorem 2.10, for \( \nu_0 = \inf_{\theta, n} \),
\[ ||\Gamma_n^{-1}(\theta)|| = \sup_{\theta, n} ||\Gamma_n(\theta)|| \] > 0, all eigenvalues of \( \Gamma_n^{-1} \) are greater than \( \nu_0 \)
uniformly in \( n \) and \( \theta \in \theta \). Thus,
\[ Y' I_n Y > \nu_0 Y' Y \quad (3.64) \]

See Rao (1973, p. 74). Consequently,
\[ b_n = \inf_{B_n} \{ Z'(\Gamma_n^{-1} - \Gamma_n^{0-1})Z + 2(1 - \rho)Y_1\Gamma_n^{-1}Z + (1 - \rho)^2 Y_1\Gamma_n^{-1}Y_1 \}
\]
\[ \geq \inf_{B_n} \{ (1 - \rho)^2 \nu_0 Y_1 Y_1 - 2|1 - \rho| \sup_{\theta} |Y_1\Gamma_n^{-1}Z| - \sup_{\theta} |Z'(\Gamma_n^{-1} - \Gamma_n^{0-1})Z| \}
\]
\[ \geq \inf_{B_n} \{ (1 - \rho)^2 \nu_0 Y_1 Y_1 - 2|1 - \rho| \sup_{\theta} |Y_1\Gamma_n^{-1}Z| - \sup_{\theta} |Z'(\Gamma_n^{-1} - \Gamma_n^{0-1})Z| \}
\]
\[ \geq \inf_{B_n} \{ (1 - \rho)^2 \nu_0 Y_1 Y_1 - 2|1 - \rho| \sup_{\theta} |Y_1\Gamma_n^{-1}Z| - \sup_{\theta} |Z'(\Gamma_n^{-1} - \Gamma_n^{0-1})Z| \}
\]
\[ \geq \inf_{B_n} \{ (1 - \rho)^2 \nu_0 Y_1 Y_1 - 2|1 - \rho| \sup_{\theta} |Y_1\Gamma_n^{-1}Z| - \sup_{\theta} |Z'(\Gamma_n^{-1} - \Gamma_n^{0-1})Z| \}
\]
\[ \geq n^{-2/3} \nu_0 Y_1 Y_1 \{ 1 + o_{wp}(1) \} - \sup_{\theta} |Z'(\Gamma_n^{-1} - \Gamma_n^{0-1})Z| \]

because $|Y_1\Gamma_n^{-1}Z| = O_{wp}(n\log\log n)$ and $(Y_1'Y_1)^{-1} = O_{wp}(n^{-2}\log\log n)$. Observe that, by Theorem 2.10,

$$\sup_{\theta} |Z'(\Gamma_n^{-1} - \Gamma_n^{(0-1)})Z| \leq 2|Z|\sup_{\theta} \|\Gamma_n^{-1}\| = O_{wp}(n).$$

Therefore, by Lemma 3.12,

$$\liminf_{n \to \infty} n^{-1}b_n \geq \liminf_{n \to \infty} n^{-1}(n^{-2/3}Y_1'Y_1\{1 + o_{wp}(1)\}) + O_{wp}(1)$$

$$= \liminf_{n \to \infty} n^{-1/3} \log\log n (n^{-2}\log\log n)\{1 + o_{wp}(1)\} + O_{wp}(1) = \infty \text{ a.s.} \quad (3.66)$$

Consequently,

$$\liminf_{n \to \infty} \min_{\theta_0} n^{-1}\{Q_n(\theta, \rho) - Q_n(\theta_0, \rho_0)\} = \liminf_{n \to \infty} n^{-1}\min(a_n, b_n) > 0 \text{ a.s.}$$

Proof of ii). By (2.2), Lemma 2.8, and Lemma 3.5,

$$\sup_{\theta, \rho} |S_n(\theta, \rho) - Q_n(\theta, \rho)| = O_{wp}(1).$$

Result ii) then follows from i). \qed

We now give the main result of this chapter, the strong consistency of estimators $(\tilde{\theta}, \tilde{\rho}, \tilde{\sigma})$, $(\overline{\theta}, \overline{\rho}, \overline{\sigma})$, and $(\hat{\theta}, \hat{\rho}, \hat{\sigma})$ for model (1.1).
Theorem 3.15. Consider model (1.1). Suppose Assumption 1.1 holds. Let \((\hat{\theta}, \hat{\rho}), \bar{\theta}, \bar{\rho})\), and \((\hat{\theta}, \hat{\rho})\) be the values minimizing

\[ Q_n(\theta, \rho) = (Y - \rho Y_1)' \Gamma_n^{-1} (Y - \rho Y_1), \]
\[ S_n(\theta, \rho) = (Y - \rho Y_1)' D_n D_n' (Y - \rho Y_1), \]
and
\[ [\det \Gamma_n]^{1/n} (Y - \rho Y_1)' \Gamma_n^{-1} (Y - \rho Y_1), \]
respectively. Define \(\bar{\sigma}^2, \bar{\sigma}^2\), and \(\hat{\sigma}^2\) by \(n^{-1} Q_n(\hat{\theta}, \hat{\rho}), n^{-1} S_n(\hat{\theta}, \hat{\rho})\), and \(n^{-1} Q_n(\hat{\theta}, \hat{\rho})\) respectively. Then

\((\bar{\theta}, \bar{\rho}, \bar{\sigma}), (\bar{\theta}, \bar{\rho}, \bar{\sigma}),\) and \((\hat{\theta}, \hat{\rho}, \hat{\sigma})\) converge to \((\theta^0, \rho^0, \sigma^0)\) a.s.

Proof. The strong consistency of \((\bar{\theta}, \bar{\rho})\) and \((\bar{\theta}, \bar{\rho})\) follows from Lemma 3.1 and Lemma 3.14. The proof of strong consistency for \((\hat{\theta}, \hat{\rho})\) is the same as the proof of strong consistency of \(\hat{\theta}\) in Theorem 3.8 except we use \(Q_n(\theta, \rho)\) in place of \(Z' \Gamma_n^{-1} Z\) and use Lemma 3.14 instead of Lemma 3.6. For the strong consistency of \(\bar{\sigma}^2\), first observe that, by Lemma 3.7,

\[ \bar{\sigma}^2 = n^{-1} Q_n(\hat{\theta}, \hat{\rho}) \leq n^{-1} Q_n(\theta^0, \rho^0) = n^{-1} Z' \Gamma_n^{0-1} Z \overset{\text{a.s.}}{\longrightarrow} (\sigma^0)^2 \]
On the other hand, with \( \nu_0 = \inf_{\theta} \| \Gamma_n^{-1}(\theta) \| > 0 \), (see Theorem 2.10),

\[
n^{-1} Q_n(\hat{\theta}, \hat{\rho}) \geq n^{-1} \{ Z \Gamma_n^{-1}(\hat{\theta}) Z + 2(1-\hat{\rho}) Y_1^\top Y_1 \Gamma_n^{-1}(\hat{\theta}) Z + (1-\hat{\rho})^2 \nu_0 Y_1^\top Y_1 \} \\
\geq n^{-1} Z \Gamma_n^{-1}(\hat{\theta}) Z + n^{-1} \inf_{\theta} (1-\rho)^2 \nu_0 Y_1^\top Y_1 - 2 |1-\rho| \sup_{\theta} |Y_1^\top \Gamma_n^{-1} Z| \\
\geq n^{-1} Z \Gamma_n^{-1}(\hat{\theta}) Z - n^{-1} \sup_{\theta} |Y_1^\top \Gamma_n^{-1} Z|^2 / (\nu_0 Y_1^\top Y_1). \quad (3.67)
\]

Observe that, by Lemma 3.11 and Lemma 3.12,

\[
l\limsup_{n \to \infty} n^{-1} \left[ \sup_{\theta} |Y_1^\top \Gamma_n^{-1} Z|^2 / (\nu_0 Y_1^\top Y_1) \right] \\
\leq \liminf_{n \to \infty} Y_1^\top Y_1 \frac{n^{-2} \log \log n}{n} \to 0 \quad \text{a.s.} \quad (3.68)
\]

Also, by Lemma 3.7,

\[
\lim_{n \to \infty} n^{-1} Z \Gamma_n^{-1}(\hat{\theta}) Z \to (\sigma^0)^2 \quad \text{a.s.} \quad (3.69)
\]

Therefore,

\[
\hat{\sigma}^2 = n^{-1} Q_n(\hat{\theta}, \hat{\rho}) \to (\sigma^0)^2 \quad \text{a.s.} \quad (3.70)
\]
The proof of strong consistency of \( \hat{\sigma}^2 \) is similar. To show the consistency of \( \hat{\sigma}^2 \) observe that

\[
n^{-1}Q_n(\hat{\theta}, \hat{\rho}) \leq n^{-1}Q_n(\hat{\theta}, \hat{\rho})[\text{det} \Gamma_n(\hat{\theta})]^{1/n} \leq n^{-1}Q_n(\theta^0, \rho^0)[\text{det} \Gamma_n(\theta^0)]^{1/n} \longrightarrow (\sigma^0)^2 \text{ a.s.} \quad (3.72)
\]

by Theorem 2.11 and Lemma 3.7.

In Lemma 3.16, we improve order of \((\hat{\rho} - 1), (\tilde{\rho} - 1), \) and \((\hat{\rho} - 1)\) up to \(o_p(n^{-1/2})\). At least \(o_p(n^{-1/2})\) is necessary for establishing the limiting distribution of \((\hat{\rho} - 1), (\tilde{\rho} - 1), \) and \((\hat{\rho} - 1)\) by the Taylor expansion of derivatives of \(Q_n(\theta, \rho),\) \(S_n(\theta, \rho),\) and \(L_n(\theta, \rho, \sigma)\).
Lemma 3.16. Let $\rho, \hat{\rho}, \hat{\rho}$ be defined in Theorem 3.15. Suppose the assumptions of Theorem 3.15 hold. Then

i) $n^{1/2}(\rho - 1) = o_p(1),$

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

iii) $n^{1/2}(\hat{\rho} - 1) = o_p(1).$

Proof of i). We have, with $\nu_0 = \inf_{\theta,n} \| \Gamma_n^{-1}(\theta) \| > 0,$ (see Theorem 2.10),

$$n^{-1}Q_n(\theta, \rho) \geq n^{-1}\left\{ Z' \Gamma_n^{-1}(\theta)Z + 2(1-\rho)Y'_1 \Gamma_n^{-1}(\theta)Z + (1-\rho)^2 \nu_0 Y'_1 Y_1 \right\}. \quad (3.73)$$

By Theorem 3.15 and Lemma 3.7

$$n^{-1}|(1-\rho)Y'_1 \Gamma_n^{-1}(\theta)Z| \leq |1-\rho| n^{-1} \sup_\theta |Y'_1 \Gamma_n^{-1}(\theta)Z| = o_p(1),$$

$$n^{-1}Z' \Gamma_n^{-1}(\theta)Z \longrightarrow (\sigma^2)^2 \quad \text{a.s.,}$$

and

$$n^{-1}Q_n(\theta, \rho) \longrightarrow (\sigma^2)^2 \quad \text{a.s.}$$

Therefore,

$$n^{-1}(1-\rho)^2 \nu_0 Y'_1 Y_1 \longrightarrow 0 \text{ in probability.}$$
Therefore, \( n(1 - \rho)^2 = o_p(1) \) (see Appendix 8.2.34).

**Proof of ii) and iii).** Observe that, as in the proof of Theorem 3.15, we can show

\[
\frac{1}{n} S_n(\theta, \rho) \text{ and } \frac{1}{n} Q_n(\theta, \rho) \rightarrow (\sigma^0)^2 \text{ a.s.}
\]

The remaining steps are the same as those in the proof of i). \( \square \)
4. LIMITING DISTRIBUTION

In this chapter, we derive the limiting distribution of the least squares estimator \((\tilde{\theta}, \tilde{\rho})\), the ordinary least squares estimator \((\hat{\theta}, \hat{\rho})\), and the maximum likelihood estimator \((\hat{\theta}, \hat{\rho})\) for model 1.1. The derivation of the limiting distribution of all the estimators is based on the limiting distribution of the least squares estimator \((\tilde{\theta}, \tilde{\rho})\). Recall that the least squares estimator \((\tilde{\theta}, \tilde{\rho})\) minimizes

\[
Q_n(\theta, \rho) = (Y - \rho Y_1)' \Gamma_n^{-1}(Y - \rho Y_1).
\]

We derive Taylor expansions of the derivatives \(Q_\rho \tilde{\theta}, Q_\rho \tilde{\theta}, Q_{\rho\rho} \tilde{\theta}\) at the true value \((\theta^o, \rho^o)\) up to second order. The limiting distribution of \((\theta, \rho)\) is obtained by solving the expanded equation for \((\rho - 1, \theta - \theta^o)\). It will be shown that

\[
\begin{bmatrix}
\frac{n}{\rho - 1} \\
\frac{n^1/2}{(\theta - \theta^o)}
\end{bmatrix} = - \begin{bmatrix}
\frac{n}{\rho - 1} Q_{\rho}^o + o_p(1) & o_p(1) \\
o_p(1) & n^{-1} Q_{\theta\theta}^o + o_p(1)
\end{bmatrix}^{-1} \begin{bmatrix}
n^{-1} Q_{\rho}^o \\
n^{-1/2} Q_{\theta}^o
\end{bmatrix},
\]

where \((Q_\rho^o, Q_{\theta}^o)\) is the vector of first partial derivatives of \(Q_n(\theta, \rho)\) and \(Q_{\rho\rho}^o, Q_{\theta\theta}^o\) are matrices of second partial derivatives of \(Q_n(\theta, \rho)\) evaluated at \((\theta^o, \rho^o)\). The partial derivatives will be defined below. Therefore by equation (4.2), the limiting distribution of \((\rho, \theta)\) is established by considering the limiting distribution of \((Q_\rho^o, Q_{\theta}^o, Q_{\rho\rho}^o, Q_{\theta\theta}^o)\).
4.1. Derivatives of $Q_n$

We define the partial derivatives of $Q_n = Q_n(\theta, \rho)$ by letting

$$Q_\rho = \frac{\partial Q_n(\theta, \rho)}{\partial \rho} = -2Y_1^1 \Gamma_n^{-1}(Y - \rho Y_1),$$

$$Q_{\rho\rho} = \frac{\partial^2 Q_n(\theta, \rho)}{\partial \rho^2} = 2Y_1^1 \Gamma_n^{-1} Y_1,$$

$$Q_{\rho\rho\rho} = \frac{\partial^3 Q_n(\theta, \rho)}{\partial \rho^3} = 0,$$

$$Q_{\rho\theta_1} = \frac{\partial^2 Q_n(\theta, \rho)}{\partial \rho \partial \theta_1} = -2Y_1^1 G_{\theta_1}(Y - \rho Y_1),$$

$$Q_{\rho\rho\theta_1} = \frac{\partial^3 Q_n(\theta, \rho)}{\partial \rho^2 \partial \theta_1} = 2Y_1^1 G_{\theta_1} Y_1,$$

$$Q_{\rho\theta_1\theta_j} = \frac{\partial^3 Q_n(\theta, \rho)}{\partial \rho \partial \theta_1 \partial \theta_j} = -2Y_1^1 G_{\theta_1 \theta_j}(Y - \rho Y_1),$$

$$Q_{\theta_1} = \frac{\partial Q_n(\theta, \rho)}{\partial \theta_1} = (Y - \rho Y_1)' G_{\theta_1}(Y - \rho Y_1),$$
\[
Q_{\theta \theta} = \frac{\partial^2 Q_n(\theta, \rho)}{\partial \theta \partial \theta} = (Y - \rho Y_1)' G_{\theta \theta}(Y - \rho Y_1), \\
Q_{\theta \theta \theta} = \frac{\partial^3 Q_n(\theta, \rho)}{\partial \theta \partial \theta \partial \theta} = (Y - \rho Y_1)' G_{\theta \theta \theta}(Y - \rho Y_1),
\] (4.3)

for \( i, j, k = 1, 2, \ldots, r \), where \( G_{\theta \theta}, G_{\theta \theta \theta}, G_{\theta \theta \theta \theta} \) are the first, second, and third partial derivatives of \( \Gamma_n^{-1} \) and \( r = p + q \). Let

\[
Q_{\theta} = (Q_{\theta_1}, \ldots, Q_{\theta_r})', \ Q_{\rho \theta} = (Q_{\rho \theta_1}, \ldots, Q_{\rho \theta_r})', \ Q_{\rho \rho \theta} = (Q_{\rho \rho \theta_1}, \ldots, Q_{\rho \rho \theta_r})'.
\] (4.4)

Also let \( Q_{\theta \theta} \) and \( Q_{\rho \theta \theta} \) be the \( r \times r \) matrices with \((i,j)\) elements \( Q_{\theta_i, \theta_j} \) and \( Q_{\rho \theta_i, \theta_j} \), respectively. For each \( i = 1, 2, \ldots, r \), let \( Q_{\theta_i, \theta_0} \) be the \( r \times r \) matrix with \((j,k)\) element \( Q_{\theta_i, \theta_j, \theta_k} \). Also let

\[
Q^\circ = Q_n(\theta^\circ, \rho^\circ), \ Q^* = Q_n(\theta^*, \rho^*), \\
\bar{Q} = Q_n(\bar{\theta}, \bar{\rho}), \ \hat{Q} = Q_n(\hat{\theta}, \hat{\rho}), \ \tilde{Q} = Q_n(\tilde{\theta}, \tilde{\rho}),
\] (4.5)

where \((\theta^*, \rho^*)\) is a vector to be defined later. We use similar notation for \( Q_{\theta}, Q_{\rho}, Q_{\rho \theta}, Q_{\theta \theta}, Q_{\rho \theta \theta}, Q_{\theta \theta \theta} \) evaluated at the different vectors \((\theta^\circ, \rho^\circ), (\theta^*, \rho^*), (\bar{\theta}, \bar{\rho}), (\hat{\theta}, \hat{\rho}), (\tilde{\theta}, \tilde{\rho})\). For example, \( Q^\circ_{\theta} \) is the vector of first order partial derivatives of \( Q_n \) evaluated at \((\theta, \rho) = (\theta^\circ, \rho^\circ)\).
In Lemma 4.1, we establish the orders in probability of the third order derivatives of $Q_n(\theta, \rho)$.

**Lemma 4.1.** Let model (1.1) and Assumption 1.1 hold. Then for $(\theta^*, \rho^*) \in \Theta \times \mathbb{R}$,

i) $Q_{\rho \theta}^* = O_p(n^2),$

ii) $Q_{\rho \theta \theta}^* = O_p(n^{3/2}) + (\rho^* - 1) O_p(n^2),$

iii) $Q_{\theta \theta \theta}^* = O_p(n) + (\rho^* - 1) O_p(n^{3/2}) + (\rho^* - 1)^2 O_p(n^2),$

where $Q_{\rho \theta}^*, Q_{\rho \theta \theta}^*, Q_{\theta \theta \theta}^*$ are defined in (4.5) and thereafter.

iv) If $\rho^* - 1 = O_p(n^{-1/2})$, then

$$Q_{\rho \theta}^* = O_p(n^2), Q_{\rho \theta \theta}^* = O_p(n^{3/2}), \text{ and } Q_{\theta \theta \theta}^* = O_p(n).$$

**Proof.** By (4.3) and (4.5),

$$Q_{\rho \theta \theta}^* = 2Y_1 G_{\theta}^* Y_1,$$

$$Q_{\rho \theta \theta}^* = -2Y_1 G_{\theta \theta}^* (Y - \rho Y_1),$$
By Theorem 2.14,
\[
\sup_{n, \theta} \|G_{\theta_1 i}^1 \| < \omega, \sup_{n, \theta} \|G_{\theta_1 j}^1 \| < \omega, \text{ and } \sup_{n, \theta} \|G_{\theta_1 j}^k \| < \omega.
\]
Therefore,
\[
|Q_{\rho \theta_1 \theta_1}^*| \leq 2|Y_1|^2 \sup_{n, \theta} \|G_{\theta_1}^1 \| = O_p(n^2),
\]
\[
|Q_{\rho \theta_1 \theta_1}^*| \leq 2|Y_1| |Z| \sup_{n, \theta} \|G_{\theta_1}^1 \| + 2|1-\rho|^* |Y_1|^2 \sup_{n, \theta} \|G_{\theta_1}^1 \|
\]
\[
= O_p(n^{3/2}) + (1-\rho^*) O_p(n^2),
\]
and
\[
|Q_{\theta_1 \theta_1 \theta_1}^*| \leq \|[Z]^2 + 2|1-\rho|^* |Y_1||Z| + (1-\rho^*)^2 |Y_1|^2 \sup_{n, \theta} \|G_{\theta_1}^1 \|^k
\]
\[
= O_p(n) + (1-\rho^*) O_p(n^{3/2}) + (1-\rho^*)^2 O_p(n^2). \quad \square
\]

In Lemma 4.2, we establish the orders of \(Q_{\rho \theta}^0\) and \(Q_{\rho \theta}^*\).

**Lemma 4.2.** Let model (1.1) and Assumption 1.1 hold. Let \(Q_{\rho \theta}^0\) and \(Q_{\rho \theta}^*\) be given in (4.5) and after (4.5). Then

i) \(Q_{\rho \theta}^0 = O_p(n)\)

and

ii) if \(\rho^* - 1 = o_p(n^{-1/2})\), then \(Q_{\rho \theta}^* = O_p(n^{3/2})\).

**Proof of i).** Fix \(i\). We first show that
By (2.5) the $j_1$-element of $Y_1D_1^{D^O}D_1^{D^O}D_1^{D^O}$ and the $i_2$-element of $ZD_2^{D^O}D_1^{D^O}$ are

\[
\sum_{i_1=1}^{n} y_{i_1-1} \sum_{k_1=1}^{n} d_1^{D^O} d_{k_1}^O |i_1-j_1| + k_1
\]

and

\[
\sum_{j_2=1}^{n} z_{j_2} \sum_{j_2=1}^{n} d_2^{D^O} d_{j_2}^O |i_2-j_2| + k_2.
\]

Therefore,

\[
Y_1D_1^{D^O}D_1^{D^O}D_1^{D^O}Z
\]

\[
= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} r_{i_1-j_1; \theta^O} \sum_{i_1=1}^{n} y_{i_1-1} \sum_{k_1=1}^{n} d_1^{D^O} d_{k_1}^O |i_1-j_1| + k_1
\]

\[
= \sum_{i_2=1}^{n} \sum_{j_2=1}^{n} \sum_{k_2=1}^{n} \sum_{i_1=1}^{n} \sum_{j_1=1}^{n} \sum_{k_1=1}^{n} n_{\max(i_1,j_1)} n_{\max(i_2,j_2)} \sum_{i_1=1}^{n} y_{i_1-1} \sum_{k_1=1}^{n} d_1^{D^O} d_{k_1}^O |i_1-j_1| + k_1
\]

\[
r_{i_2-j_2; \theta^O} d_2^{D^O} d_{j_2}^O |i_2-j_2| + k_2
\]

\[
= \sum_{i_1=1}^{n} \sum_{j_1=1}^{n} a_{i_1,j_1} y_{i_1-1} z_{j_1-1}.
\]
\[ a_{n,i,j_2} = \sum_{i_2=1}^{n} \sum_{k_2=0}^{n} \sum_{j_1=1}^{n} \sum_{k_1=1}^{n} \max(i_2,j_2) \max(i_1,j_1) \]

\[ r_{\theta}^{(i_2-j_1, \theta^0)} d_{k_1}^{(i_1-j_2)} d_{k_2}^{(i_2-j_2)} + k_1 d_{k_2}^{(i_j-j_2)} + k_2 \]  \hspace{1cm} (4.9)

Note that, for some \( M_1, M_2, M_3 < \infty \) and \( \lambda, \lambda_1 \in (0,1) \), we have (see Appendix 8.3.1),

\[ |a_{n,i,j_2}| \leq M_1 \sum_{i_2=1}^{n} \sum_{k_2=0}^{n} \sum_{j_1=1}^{n} \sum_{k_1=1}^{n} \lambda^{i_2-j_1} + k_1 + |i_1-j_1| + k_1 + k_2 + |i_2-j_2| + k_2 \]

\[ \leq M_2 \sum_{i_2=1}^{n} \sum_{j_1=1}^{n} \lambda^{i_2-j_1} + |i_1-j_1| + |i_2-j_2| \leq M_2^{\lambda_1} |i_1-j_2|. \]  \hspace{1cm} (4.10)

Now, from Lemma 3.10-iv), we conclude that (4.8) is \( O_p(n) \) and hence we have (4.6). Therefore, by Lemma 3.5 and Lemma 2.8 result i) follows. See (4.35).

Proof of ii). Fix \( i \). By Theorem 2.14, Lemma 3.16, and (4.3) we have

\[ |Q_{\mathbf{1}}^{*}| = | \mathbf{-2Y}_1^{*} G_{\theta_1}^{*} Z - 2(1-\rho^{*}) Y_1^{*} G_{\theta_1}^{*} Y_1 | \]

\[ \leq 2\{ |\mathbf{Y}_1| |Z| + |1-\rho^{*}| |\mathbf{Y}_1| |^2 \} \sup_{n, \theta} \| G_{\theta_1}^{*} \| \]

\[ = O_p(n^{3/2}) + O_p(n^{-1/2}) O_p(n^{2}) = O_p(n^{3/2}). \]  \hspace{1cm} (4.11)

\( \square \)
In Lemma 4.3, we establish the limiting distribution of \((n^{-2}Q_{\rho\rho}, n^{-1}Q_{\rho}, n^{-1/2}Q_{\theta\theta})\) and the probability limits of \(n^{-1}Q_{\theta\theta}, n^{-3/2}Q_{\rho\theta}\). These limiting results for the derivatives evaluated at the true values form the foundation for the limiting distribution of the estimator. The limiting distribution of \(n^{-2}Q_{\rho\rho}\) is 
\[2(\sigma^2)^2 \int_0^1 W^2(r)dr\] and the limiting distribution of \(n^{-1}Q_{\rho}\) is 
\[-(\sigma^2)^2 \{W^2(1) - 1\}.

**Lemma 4.3.** Let model (1.1) and Assumption 1.1 hold. Let \(Q_{\rho\rho}^0, Q_{\rho}^0, Q_{\theta\theta}^0, Q_{\rho\theta}^0\) be given by (4.5). Then,

i) \(n^{-2}Q_{\rho\rho}^0, n^{-1}Q_{\rho}^0 \rightarrow 2(\sigma^2)^2 \int_0^1 W^2(r)dr, -2(\sigma^2)^2 \{W^2(1) - 1\},\)

ii) \(n^{-1}Q_{\theta\theta}^0 \rightarrow V(\theta^0)\) in probability,

\[n^{-1/2}Q_{\theta}^0 \rightarrow N(r, V(\theta^0)),\]

\[n^{-3/2}Q_{\rho\theta}^0 \rightarrow 0\] in probability,

where \(V(\theta)\) is an \(r \times r\) matrix with \((i,j)\) element

\[V(\theta)_{ij} = (4\pi)^{-1} \int_{-\pi}^{\pi} \frac{\partial \log g(\omega; \theta)}{\partial \theta_i} \frac{\partial \log g(\omega; \theta)}{\partial \theta_j} d\omega,\]  

(4.12)

\(g(\omega; \theta)\) is the spectral density of the process \(\{z_t\}\) given in (2.13), and \(\rightarrow\) denotes convergence in distribution.
iii) The two limiting distributions of \((n^{-2}Q_{pp}, n^{-1}Q_{p})\) and \(n^{-1/2}Q_{0}\) in i) and ii) are independent.

**Proof of i).** First we show that,

\[
Y_1 D_n^{D_n} D_{n}^{O} = \sum_{j=1}^{n} W_j - e_j + O_p(n^{1/2}),
\]

where \(W_j = e_1 + \cdots + e_j\) and \(W_0 = 0\). Combining (3.39) and (3.40) and applying Lemma 3.10-iii), we get

\[
Y_1 D_n^{D_n} D_{n}^{O} = \left( \sum_{k=0}^{\infty} d_{k}^{o} \right)^2 \sum_{j=1}^{n} \sum_{k=1}^{j-1} z_{kj} - \sum_{j=1}^{n} \sum_{k=1}^{j-1} z_{kj} j k_{1} = 0 \ k_{2} = 0 \ k_{1} = 0 \ k_{2} = 0 \ d_{k_{1} j-k+k_{1}+k_{2}}^{o} \ d_{k_{1} j-k+k_{1}+k_{2}}^{o} \ + \sum_{j=1}^{n} \sum_{k=1}^{j} z_{kj} d_{k_{1} j-k+k_{1}+k_{2}}^{o} \ d_{k_{1} j-k+k_{1}+k_{2}}^{o} + O_p(n^{1/2}).
\]

(4.13)

After some algebra (see Appendix 8.3.2, 8.3.3, and 8.3.4 for (4.14), (4.15), and (4.16) respectively) we can express (4.13) in terms of \(\{e_i\}\). Let \(c_{0} = \left( \sum_{i=1}^{\infty} v_{i}^{o} \right)\). Then

\[
\sum_{j=1}^{n} \sum_{k=1}^{j} z_{kj} = c_{0} \sum_{i=1}^{\infty} v_{i}^{o} + 2^{-1} \sum_{k=0}^{\infty} \sum_{i=1}^{\infty} v_{i}^{o} + O_p(n^{1/2}),
\]

(4.14)

\[
\sum_{j=1}^{n} \sum_{k=1}^{j} z_{kj} d_{k_{1} j-k+k_{1}+k_{2}}^{o} \ d_{k_{1} j-k+k_{1}+k_{2}}^{o} \ = \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} v_{k}^{o} \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} d_{k_{1} j-k+k_{1}+k_{2}}^{o} \ d_{k_{1} j-k+k_{1}+k_{2}}^{o} + O_p(n^{1/2}),
\]

(4.15)
Replacing terms in (4.13) by the terms in (4.14), (4.15), and (4.16), we have

\[ Y_n' D_n' D_n Z = \left( \sum_{k=0}^{\infty} d_k^0 \sum_{i_1=1}^{i_2=1} e_{i_1} e_{i_2} + 2^{-1}(c_0^2 - \sum_{k=0}^{\infty} v_k^{02}) \sum_{i=1}^{n} e_{i}^2 \right) 
+ \left\{ \sum_{k=0}^{\infty} v_k^{02} \sum_{j=0}^{k-1} d_j d_{j-1+k_1+k_2} \right\} \sum_{i=1}^{n} e_{i}^2 + O_p(n^{1/2}). \]
Now observe that the coefficient of $\sum_{i=1}^{n} e_i^2$ is zero because of (4.17) – (4.20) below.

We have

$$\sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1}^0 d_{k_2}^0 = \sum_{k_1=0}^{\infty} d_{k_1}^0 \sum_{k_2=0}^{\infty} d_{k_2}^0 = \sum_{k_1=0}^{\infty} d_{k_1}^0 \sum_{k_2=0}^{\infty} d_{k_2}^0$$

(4.17)

Also

$$c_0 \sum_{k=0}^{\infty} d_k^0 = (\sum_{k=0}^{\infty} v_k^0) (\sum_{k=0}^{\infty} d_k^0) = \sum_{k=0}^{\infty} \sum_{i=0}^{k_1} v_i^0 d_{k-i}^0 = 1,$$

(4.18)

because $\sum_{i=0}^{\infty} v_i^0 d_{k-i}^0 = 1$ for $k = 0$, and is equal to 0 for $k \neq 0$. Also

$$\sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1}^0 d_{k_2}^0 = 2^{-1}\left(\sum_{k=0}^{\infty} d_k^0\right)^2 - \sum_{k=0}^{\infty} d_k^0 d_k^2).$$

(4.19)

Finally,

$$\sum_{j=1}^{k_1} \sum_{k=0}^{\infty} v_k^0 v_j^0 \sum_{k_1=0}^{\infty} d_{k_1}^0 d_{j-k+k_1}^0 = 2^{-1}[1 - \sum_{k=0}^{\infty} v_k^0 \sum_{k=0}^{\infty} d_k^2].$$

(4.20)

For the verification of (4.20), consider the stationary time series

$$\hat{z}_t = \sum_{j=0}^{\infty} d_j^0 \hat{e}_{t-j}.$$
where \( \{\epsilon_t\} \) is an iid \((0,1)\) sequence. Then

\[
\epsilon_t = \sum_{j=0}^{\infty} \epsilon_j^0 z_{t-j}
\]

and

\[
\sum_{k_1=0}^{\infty} \epsilon_{k_1}^0 \epsilon_{j-k+k_1} = \text{cov}(\hat{z}_j, \hat{z}_{k_1}).
\]

Therefore,

\[
\sum_{j=0}^{\infty} \sum_{k=0}^{j-1} \epsilon_j^0 \epsilon_k^0 \sum_{j=0}^{\infty} \sum_{k_1=0}^{j-k+k_1} \epsilon_{k_1}^0 \epsilon_{j-k+k_1} = \sum_{j=1}^{\infty} \sum_{k=0}^{j-1} \epsilon_j^0 \epsilon_k^0 \text{cov}(\hat{z}_j, \hat{z}_k).
\]

\[
= 2^{-1} \left[ \sum_{j=0}^{\infty} \epsilon_j^0 \sum_{k=0}^{j-1} \epsilon_k^0 \text{cov}(\hat{z}_j, \hat{z}_k) \right] - \sum_{k=0}^{\infty} \epsilon_k^0 \text{Var}(\hat{z}_k)
\]

\[
= 2^{-1} \left[ \text{Var}(\sum_{j=0}^{\infty} \epsilon_j^0 \hat{z}_j) - \sum_{k=0}^{\infty} \epsilon_k^0 \sum_{k=0}^{\infty} \epsilon_k^0 \text{d}_k^2 \right] = 2^{-1} \left[ 1 - \sum_{k=0}^{\infty} \epsilon_k^0 \sum_{k=0}^{\infty} \text{d}_k^2 \right].
\]

Therefore,

\[
\sum_{i_1=1}^{n-1} \sum_{i_2=1}^{i_1-1} e_{i_1} e_{i_2} + \mathcal{O}_p(n^{1/2}).
\]

The result (4.24) together with (2.2), Lemma 2.8, and Lemma 3.5 gives us

\[
\sum_{i_1=1}^{n-1} \sum_{i_2=1}^{i_1-1} e_{i_1} e_{i_2} + \mathcal{O}_p(n^{1/2}) = \sum_{j=1}^{n} W_{j-1} e_j + \mathcal{O}_p(n^{1/2}).
\]
Next we have, with \( d^0 = \left( \sum_{s=0}^{\infty} d_s^0 \right)^2 \),

\[
Y_1 D_n^o Y_1 = d^{02} \sum_{j=1}^{n} \sum_{k=1}^{j-1} z_k y_{j-1} + O_p(n^{3/2}) \quad \text{(see Appendix 8.3.5)} \tag{4.26}
\]

\[
= c_0^2 d^{02} \sum_{j=1}^{n} W_{j-1}^2 + O_p(n^{3/2}) + O_p(n^{3/2}) \quad \text{(Lemma 3.12–ii)}
\]

\[
= \sum_{j=1}^{n} W_{j-1}^2 + O_p(n^{3/2}).
\]

Therefore,

\[
Y_1 \Gamma_n^{o-1} Y_1 = \sum_{j=1}^{n} W_{j-1}^2 + O_p(n^{3/2}).
\]

Consequently, by (4.3),

\[
(n^{-2} Q_{rr}, n^{-1} Q_{r}) = 2(n^{-2} Y_1 \Gamma_n^{o-1} Y_1, -n^{-1} Y_1 \Gamma_n^{o-1} Z).
\]

\[
= 2(n^{-2} \sum_{j=1}^{n} W_{j-1}^2, -n^{-1} \sum_{j=1}^{n} W_{j-1} e_j) + o_p(1). \tag{4.27}
\]

From (4.27) the result follows. See Example A.9 in Appendix A.


Proof of iii). The independence of the two limiting distributions is a consequence of Theorem 2.2 of Chan and Wei (1988). \( \square \)
4.2. Limiting Distribution of Estimators

Now we give the main result of this section. The normalized estimators $n(\rho - 1)$, $n(\lambda - 1)$, and $n(\rho - 1)$ have the same limiting distribution. This limiting distribution is shown to be the same as that of Dickey and Fuller (1979) given in (1.7). The limiting distributions of $n^{1/2}(\theta - \theta^0)$, $n^{1/2}(\phi - \phi^0)$, $n^{1/2}(\hat{\theta} - \theta^0)$ are the same and the limiting distribution is a multivariate normal distribution. Moreover the limiting distribution of $n(\rho - 1)$ is independent of the limiting distribution of $n^{1/2}(\theta - \theta^0)$.

Theorem 4.4. Let model (1.1) and Assumption 1.1 hold. Then,

$$n(\rho - 1) \Rightarrow 2^{-1}[W^2(1) - 1] / \int_0^1 W^2(r)\,dr$$

and

$$n^{1/2}(\lambda - \theta^0) \Rightarrow N(0, V^{-1}(\theta^0)),$$

where $V(\theta^0)$ is defined in (4.12). The limiting distribution of $n(\rho - 1)$ is independent of the limiting distribution of $n^{1/2}(\theta - \theta^0)$. The limiting distributions of $n(\rho - 1)$ and $n(\rho - 1)$ are the same as that of $n(\rho - 1)$. The limiting distributions of $n^{1/2}(\theta - \theta^0)$ and $(\theta - \theta^0)$ are the same as that of $n^{1/2}(\theta - \theta^0)$.

Proof. The second order Taylor expansion of $(\hat{Q}_{\rho}, \hat{Q}_{\theta})$ at $(\theta^0, \rho^0)$ is
\[ 0 = \left( \begin{array}{c} \tilde{Q}_p \\ \tilde{Q}_\theta \end{array} \right) = \left( \begin{array}{c} Q^0_p \\ Q^0_\theta \end{array} \right) + \left( \begin{array}{cc} Q^0_{pp} & Q^0_{p\theta} \\ Q^0_{p\theta} & Q^0_{\theta\theta} \end{array} \right) \left( \begin{array}{c} \bar{p} - 1 \\ \bar{\theta} - \tilde{\theta} \end{array} \right) \]

\[ + \frac{1}{2} \left( \begin{array}{ccc} (\bar{\theta} - \tilde{\theta})' & Q^*_{p\theta} & Q^*_{p\theta}(\bar{\rho} - 1) + (\bar{\theta} - \tilde{\theta})'Q^*_{p\theta} \\ Q^*_{p\theta}(\bar{\rho} - 1) + Q^*_{p\theta}(\bar{\theta} - \tilde{\theta}) & Q^*_{p\theta\theta}(\bar{\rho} - 1) + (\bar{\theta} - \tilde{\theta})Q^*_{p\theta\theta} \end{array} \right) \left( \begin{array}{c} \bar{\rho} - 1 \\ \bar{\theta} - \tilde{\theta} \end{array} \right) \]  

where the vector \((\bar{\theta}^*, \tilde{\theta}^*)\) which defines \(Q^*_{p\theta}, Q^*_{p\theta\theta}\), and \(Q^*_{p\theta\theta}\) is between \((\bar{\theta}^*, \bar{\rho}^*)\) and \((\tilde{\theta}, \tilde{\rho})\), and

\[ (\bar{\theta} - \tilde{\theta})Q^*_{p\theta\theta} = \begin{bmatrix} (\bar{\theta} - \tilde{\theta})'Q^*_{\bar{\theta}} \\ \vdots \\ (\bar{\theta} - \tilde{\theta})'Q^*_{\bar{\theta}} \end{bmatrix} \]

is an \(r \times r\) matrix. From Lemma 4.1-i), \(Q^*_{p\theta\theta} = O_p(n^2)\). From Theorem 3.15, \((\bar{\theta} - \tilde{\theta})' = o_p(1)\). Therefore,

\[ (\bar{\theta} - \tilde{\theta})'Q^*_{p\theta\theta} = o_p(n^2), \]

From Lemma 3.16, \(\bar{\rho} - 1 = o_p(n^{-1/2})\). From Lemma 4.1-iv), \(Q^*_{p\theta\theta} = O_p(n^2), Q^*_{p\theta\theta} = O_p(n^{3/2}), \) and \(Q^*_{p\theta\theta} = O_p(n)\). From Theorem 3.15, \((\bar{\theta} - \tilde{\theta})' = o_p(1)\). Therefore
\[ Q^{*}_{\rho \theta}(\rho - 1) + (\hat{\theta} - \theta^0)'Q^{*}_{\rho \theta} = o_p(n^{3/2}), \]

and

\[ Q^{*}_{\rho \theta}(\rho - 1) + (\hat{\theta} - \theta^0)'Q^{*}_{\rho \theta} = o_p(n). \] (4.30)

Also, from Lemma 4.2,

\[ Q^{0}_{\rho \theta} = o_p(n). \]

Therefore,

\[
\begin{bmatrix}
  n(\hat{\rho} - 1) \\
  n^{1/2}(\hat{\theta} - \theta^0)
\end{bmatrix} = \begin{bmatrix}
  n Q^{0}_{\rho \rho} + o_p(1) & o_p(1) \\
  o_p(1) & n Q^{0}_{\theta \theta} + o_p(1)
\end{bmatrix}^{-1} \begin{bmatrix}
  n^{-1} Q^{0}_{\rho} \\
  n^{-1/2} Q^{0}_{\theta}
\end{bmatrix}. \] (4.31)

The result (4.31), together with Lemma 4.3, gives the limiting distribution of \( n(\hat{\rho} - 1) \) and \( n^{1/2}(\hat{\theta} - \theta^0) \). See Example A.10 in Appendix A.

Next we consider the limiting distribution of \( n(\hat{\rho} - 1) \) and \( n^{1/2}(\hat{\theta} - \theta^0) \).

From (1.15), differentiating \( L_n(\theta, \rho, \sigma) \) with respect to \( (\rho, \theta) \), evaluating the derivatives at \( (\hat{\theta}, \hat{\rho}, \hat{\sigma}) \), and setting the derivatives equal to zero gives,

\[ 0 = \frac{\partial L_n(\hat{\theta}, \hat{\rho}, \hat{\sigma})}{\partial \rho} = Q^{*}_{\rho} \hat{\sigma}^{-2}, \]

\[ 0 = \frac{\partial L_n(\hat{\theta}, \hat{\rho}, \hat{\sigma})}{\partial \theta} = Q^{*}_{\theta} \hat{\sigma}^{-2} + \frac{\partial}{\partial \theta} \log \det \Gamma_n(\hat{\theta}) \]
\[ = \hat{Q}_\theta \hat{\sigma}^{-2} + \frac{\partial}{\partial \theta} \log \det \Gamma_n(\theta^\star) + \left\{ \frac{\partial^2}{\partial \theta^2} \log \det \Gamma_n(\theta^\star) \right\}(\hat{\theta} - \theta^\star), \] (4.32)

where \( \theta^\star \) is between \( \theta^0 \) and \( \hat{\theta} \). In the above equation, utilizing the second order Taylor expansion of \((\hat{Q}_\rho, \hat{Q}_\theta)\) at \((\theta^0, \rho^0)\), we have,

\[
\begin{pmatrix}
    n(\hat{\rho} - 1) \\
    n^{1/2}(\hat{\theta} - \theta^0)
\end{pmatrix} = -\begin{pmatrix}
    n^{-2}Q_{\rho\rho}^0 + o_p(1) & o_p(1) \\
    o_p(1) & n^{-1}Q_{\theta\theta}^0 + n^{-1} \frac{\partial^2}{\partial \theta^2} \log \det \Gamma_n(\theta^\star) + o_p(1)
\end{pmatrix}^{-1} \times \begin{pmatrix}
    n^{-1}Q_{\rho}^0 \\
    n^{-1/2}Q_{\theta}^0 + n^{-1/2} \frac{\partial^2}{\partial \theta^2} \log \det \Gamma_n(\theta^\star)
\end{pmatrix}. \] (4.33)

By the result in Brockwell and Davis (1987, p. 383),

\[ n^{-1} \frac{\partial^2}{\partial \theta^2} \log \det \Gamma_n(\theta^\star) = o_p(1) \] (4.34)

and

\[ n^{-1/2} \frac{\partial}{\partial \theta} \log \det \Gamma_n(\theta^\star) = o_p(1). \]

Hence by Lemma 4.3, we have the desired limiting distribution of \( n(\hat{\rho} - 1) \) and \( n^{1/2}(\hat{\theta} - \theta^0) \).

Finally, we derive the limiting distribution of \( n(\rho - 1) \) and \( n^{1/2}(\theta - \theta^0) \). By
Theorem 2.6 and Theorem 2.13, the derivatives of $A$, $D_n$, and $M_n$ have the same uniformly exponentially decaying properties as $A$, $D_n$, and $M_n$, respectively. Therefore, by the same argument used in the proof of Lemma 3.5,

$$\sup_{\theta} \left\| \frac{\partial}{\partial \theta_i}(D_n'M_n) \right\| = O_p(1)$$

and

$$\sup_{\theta} \left\| Z_i \frac{\partial}{\partial \theta_i}(D_n'M_n) \right\| = O_p(1) \quad \text{for all } i = 1, 2, \ldots, r.$$ 

Also, by the same argument used in the proof of Lemma 2.8,

$$\sup_{\theta} \left\| \frac{\partial}{\partial \theta_i} \left( A^{-1} + M_n'M_n \right)^{-1} \right\| < \omega, \ i = 1, 2, \ldots, r.$$ 

Hence,

$$\sup_{\theta} \left| Y_i \frac{\partial}{\partial \theta_i} \{D_n'M_n(A^{-1} + M_n'M_n)^{-1} M_n'D_n\} \right| = O_p(1)$$

and

$$\sup_{\theta} \left| Y_i \frac{\partial}{\partial \theta_i} \{D_n'M_n(A^{-1} + M_n'M_n)^{-1} M_n'D_n\} \right| = O_p(1), \ i = 1, \ldots, r. \quad (4.35)$$

Also, by Lemma 2.8 and Lemma 3.5,

$$\sup_{\theta} \left\| \frac{\partial}{\partial \rho} (Y - \rho Y_1)' \{D_n'M_n(A^{-1} + M_n'M_n)^{-1} M_n'D_n\} (Y - \rho Y_1) \right\| = O_p(1).$$

Consequently, the derivatives $S_\theta$ and $S_\rho$ of $S(\theta, \rho)$ satisfy
\[ 0 = \langle S_{\rho}, S_\theta \rangle = \langle Q_{\rho}, Q_\theta \rangle + O_p(1). \]  

(4.36)

The result (4.36), together with the argument applied to get the limiting distribution of \([n(\tilde{\rho} - 1), n^{1/2}(\bar{\theta} - \theta^0)]\), yields the desired limiting distribution of \([n(\tilde{\rho} - 1), n^{1/2}(\bar{\theta} - \theta^0)]\). Independence of the two limiting distributions \(2^{-1}\{W^2(1) - 1\} / \int_0^1 W^2(r)dr \) and \(N(0, V^{-1}(\theta^0))\) is a consequence of Lemma 4.3-iii). \(\square\)

The limiting distribution of \([n(\tilde{\rho} - 1), n^{1/2}(\bar{\theta} - \theta^0)]\) is completely analogous to that given by Fuller (1976, pp. 373 – 381) for the p-th order autoregressive process. The limiting distribution can be divided into two independent parts. The limiting distribution of \(n(\tilde{\rho} - 1)\) is the same as that of Dickey and Fuller (1979) as is given in (1.17) and independent of the limiting distribution of \(n^{1/2}(\bar{\theta} - \theta^0)\) which is multivariate normal.

4.3. Limiting Distribution of the Regression t-statistics

In this section we define the regression t-statistics and investigate their limiting distributions. To define the test statistics we first define the covariance matrix of the limiting distribution of the estimators. Consider \(e_t(Y; \theta, \rho)\) defined in (1.18). Let

\[ W_{\rho t}(Y; \theta, \rho) = \frac{\partial}{\partial \rho} e_t(Y; \theta, \rho). \]

and
\[ W_{\theta_1}(Y;\theta, \rho) = \frac{\partial}{\partial \theta_1} c_t(Y;\theta, \rho), \quad i = 1, 2, ..., p+q. \] (4.37)

Let \( W_{\theta}(Y;\theta, \rho) \) be the \((p+q)\)-dimensional column vector with \( i \)-th element \( W^{(i)}(Y;\theta, \rho) \). Also define

\[
V(Y;\theta, \rho) = \begin{pmatrix}
\sum_{t=1}^{n} W_{\rho t}^2(Y;\theta, \rho) & \sum_{t=1}^{n} W_{\rho t}(Y;\theta, \rho) W_{\hat{\theta}}(Y;\theta, \rho) \\
\sum_{t=1}^{n} W_{\rho t}(Y;\theta, \rho) W_{\hat{\theta}}(Y;\theta, \rho) & \sum_{t=1}^{n} W_{\hat{\theta}}(Y;\theta, \rho) W_{\hat{\theta}}(Y;\theta, \rho)
\end{pmatrix}
\] (4.38)

The covariance matrix is estimated by replacing \((\theta, \rho)\) of (4.38) with its estimator.

The regression t-statistics are defined as

\[
\bar{\tau} = (\bar{\rho} - 1)(c_L^{-1/2}),
\]

\[
\hat{\tau} = (\hat{\rho} - 1)(c_O^{-1/2}),
\]

\[
\ddot{\tau} = (\ddot{\rho} - 1)(c_M^{-1/2}),
\] (4.39)

where \( c_L, c_O, c_M \) are the upper left element of \( V^{-1}(Y;\theta, \rho) \), \( V^{-1}(Y;\hat{\theta}, \rho) \), and \( V^{-1}(Y;\ddot{\theta}, \rho) \), respectively.

Next we define the t-statistics of \( \theta_1 \). Thinking of the testing \( H_0 : \theta_1 = \theta_1^0 \), we define

\[
\overline{t}_1 = (\overline{\theta}_1 - \theta_1^0)(d_{\overline{\theta}_1}^{-1/2}),
\]
\[ t_i = (\hat{\theta}_i - \theta_i^0)(d_{O,i} \sigma^2)^{-1/2}, \]

\[ \hat{t}_i = (\hat{\theta}_i - \theta_i^0)(d_{M,i} \sigma^2)^{-1/2}, \]

where \( d_{L,i}, d_{O,i}, \) and \( d_{M,i} \) are the \((i+1, i+1)\) element of matrix \( V^{-1}(Y; \hat{\theta}, \hat{\rho}) \), \( V^{-1}(Y; \theta, \rho) \), and \( V^{-1}(Y; \hat{\theta}, \hat{\rho}) \) respectively.

In Theorem 4.5, we establish the limiting distribution of \( \tilde{\tau}, \tilde{t}_i, \tilde{t}_j, \tilde{t}_k, \tilde{t}_l, \tilde{t}_m \), \( i = 1, \ldots, p+q \).

**Theorem 4.5.** Let model (1.1) and Assumption 1.1 hold. Then

\[ \tilde{\tau} \Rightarrow 2^{-1}\{W^2(1) - 1\}/[\int_0^1 W^2(r)dr]^{1/2} \]  \( (4.40) \)

and \( \tilde{t}_i \Rightarrow N(0,1), i = 1, \ldots, p+q \). The limiting distribution of \( \tilde{\tau} \) and \( \{\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_{p+q}\} \) are independent. The limiting distribution of \( (\tilde{\tau}, \tilde{t}_i, i = 1, 2, \ldots, p+q) \) and \( (\tilde{\tau}, \tilde{t}_i, i = 1, \ldots, p+q) \) is the same as that of \( (\tilde{\tau}, \tilde{t}_i, i = 1, 2, \ldots, p+q) \).

**Proof.** First we derive the limiting distribution of \( n^{-2} \sum_{t=1}^n W_{\hat{\rho}_t}(Y; \hat{\theta}, \hat{\rho}) \). Note that

\[ n^{-2} \sum_{t=1}^n W_{\hat{\rho}_t}(Y; \hat{\theta}, \hat{\rho}) = 2^{-1} n^{-2} \sum_{t=1}^n \frac{\partial^2}{\partial \rho} \sum_{t=1}^n e_t^2(Y; \hat{\theta}, \hat{\rho}) = 2^{-1} n^{-2} Q_{\rho \rho}(\hat{\theta}, \hat{\rho}) + o_p(1) \]

by the similar argument used in getting (4.36). By Taylor expansion of \( Q_{\rho \rho}(\hat{\theta}, \hat{\rho}) \) at
\((\theta^0, \rho^0)\) and Lemma 4.1,

\[ n^{-2} Q_{\rho \rho} (\hat{\theta}, \hat{\rho}) = n^{-2} Q_{\rho \rho} (\theta^0, \rho^0) + o_p(1). \]

Therefore, by Lemma 4.3-4),

\[ n^{-2} \sum_{t=1}^{n} W_t^2 (Y; \tilde{\theta}, \tilde{\rho}) \Rightarrow (\sigma^0)^2 \int_0^1 W^2 (r) dr. \quad (4.41) \]

Next we show

\[ n^{-3/2} \sum_{t=1}^{n} W_t (Y; \tilde{\theta}, \tilde{\rho}) W_{\theta t} (Y; \tilde{\theta}, \tilde{\rho}) = o_p(1). \quad (4.42) \]

Note that

\[ \sum_{t=1}^{n} W_t (Y; \tilde{\theta}, \tilde{\rho}) W_{\theta t} (Y; \tilde{\theta}, \tilde{\rho}) = 2^{-1} \sum_{t=1}^{n} \sigma_t^2 (Y; \tilde{\theta}, \tilde{\rho}) - \sum_{t=1}^{n} W_{\rho \theta t} (Y; \tilde{\theta}, \tilde{\rho}) e_t (Y; \tilde{\theta}, \tilde{\rho}), \]

where \( W_{\rho \theta t} (Y; \theta, \rho) \) is the \((p+q)\) dimensional column vector with \( i \)-th element

\[ \frac{\partial^2}{\partial \rho \partial \theta_i} e_t (Y; \theta, \rho), \ t = 1, 2, \ldots. \] Observe that, by Lemma 4.1 and Lemma 4.2,

\[ n^{-3/2} \frac{\partial^2}{\partial \rho \partial \theta_i} \sum_{t=1}^{n} e_t^2 (Y; \tilde{\theta}, \tilde{\rho}) = n^{-3/2} Q_{\rho \theta} (\tilde{\theta}, \tilde{\rho}) + o_p(1) \]

\[ = n^{-3/2} [Q_{\rho \theta}^0 + (\tilde{\rho} - 1) Q_{\rho \rho \theta}^* + (\tilde{\theta} - \theta^0) Q_{\rho \theta \theta}^*] + o_p(1) = o_p(1), \]

where \((\rho^*, \theta^*)\) which defines \(Q_{\rho \rho \theta}^*\) and \(Q_{\rho \theta \theta}^*\) is between \((\rho^0, \theta^0)\) and \((\tilde{\rho}, \tilde{\theta})\). Now, for each \(i = 1, 2, \ldots, (p+q),\)
\[
\sum_{t=1}^{n} W_{\theta_1}(Y; \tilde{\theta}, \tilde{\rho}) e_t(Y; \tilde{\theta}, \tilde{\rho}) = -Y_1 \{ \frac{\partial}{\partial \theta_1} D_n(\tilde{\theta}) \} D_n(\tilde{\theta}) (Y - \tilde{\rho} Y_1)
\]

\[
= -Y_1 \{ \frac{\partial}{\partial \theta_1} D_n(\tilde{\theta}) \} D_n(\tilde{\theta}) Z - (1 - \tilde{\rho}) Y_1 \{ \frac{\partial}{\partial \theta_1} D_n(\tilde{\theta}) \} D_n(\tilde{\theta}) Y_1 = O_p(n)
\]

(4.43)

because, by Corollary 2.5, \( \frac{\partial}{\partial \theta_1} D_n(\tilde{\theta}) \) has the same uniformly exponentially declining property as \( D_n \). Therefore (4.42) holds. Finally, we show

\[
n^{-1} \sum_{t=1}^{n} W_{\theta_1}(Y; \tilde{\theta}, \tilde{\rho}) W_{\theta_1}(Y; \tilde{\theta}, \tilde{\rho}) - P \rightarrow V(\theta^0),
\]

(4.44)

where \( V(\theta^0) \) is defined in (4.12). Note that

\[
\sum_{t=1}^{n} W_{\theta_1}(Y; \tilde{\theta}, \tilde{\rho}) W_{\theta_1}(Y; \tilde{\theta}, \tilde{\rho}) = 2^{-1} \frac{\partial^2}{\partial \theta^2} \sum_{t=1}^{n} e_t^2(Y; \tilde{\theta}, \tilde{\rho}) - \sum_{t=1}^{n} W_{\theta_1}(Y; \tilde{\theta}, \tilde{\rho}) e_t(Y; \tilde{\theta}, \tilde{\rho}),
\]

(4.45)

where \( W_{\theta_1}(Y; \tilde{\theta}, \tilde{\rho}) \) is the \((p+q) \times (p+q)\) matrix with \((i,j)\) element \( \frac{\partial^2}{\partial \theta_i \partial \theta_j} e_t(Y; \tilde{\theta}, \tilde{\rho}) \).

Observe that by Lemma 4.1 and 4.3

\[
\frac{\partial^2}{\partial \theta^2} \sum_{t=1}^{n} e_t^2(Y; \tilde{\theta}, \tilde{\rho}) = n^{-1} Q_{\theta_1}(\tilde{\theta}, \tilde{\rho}) + o_p(1) \rightarrow P \rightarrow V(\theta^0).
\]

Also it can be shown that
\[ \sum_{t=1}^{n} W_{\theta \theta}(Y; \tilde{\theta}, \tilde{\rho}) e_t(Y; \tilde{\theta}, \tilde{\rho}) = o_p(n). \] (4.46)

See Appendix 8.3.6. Therefore (4.44) holds. Now from (4.41), (4.42), and (4.44),

\[ \text{diag}(n^{-1}, n^{-1/2}) V(Y; \tilde{\theta}, \tilde{\rho}) \text{diag}(n^{-1}, n^{-1/2}) \rightarrow \text{diag}\left\{ (\sigma^0)^2 \int_0^1 W^2(r) \, dr, V(\theta^0) \right\}. \] (4.47)

Combining Lemma 4.3 and (4.47), we conclude (4.40).

The limiting distribution of \( \tilde{t}_i, i = 1, \ldots, p+q \) is obtained from Theorem 4.4 and (4.47). The independence of the limiting distributions of \( \tilde{\tau} \) and \( \{\tilde{t}_i, i = 1, \ldots, p+q\} \) is a consequence of Theorem 4.4.

The limiting distribution of \( (\tilde{\tau}, \tilde{t}_i, i = 1, \ldots, p+q) \) and \( (\hat{\tau}, \hat{t}_i, i = 1, \ldots, p+q) \) can be obtained similarly.

\[
\text{In Corollary 4.6, we state the limiting behavior of } V(Y; \theta^*, \rho^*) \text{ for } (\theta^*, \rho^*) = (\tilde{\theta}, \tilde{\rho}), (\hat{\theta}, \hat{\rho}), \text{ and } (\hat{\theta}, \hat{\rho}).
\]

Corollary 4.6. Let model (1.1) and Assumption 1.1 hold. Let \( V(Y; \theta, \rho) \) be defined in (4.39). Then, for all \( (\theta^*, \rho^*) = (\tilde{\theta}, \tilde{\rho}), (\hat{\theta}, \hat{\rho}), \text{ and } (\hat{\theta}, \hat{\rho}) \), we have

\[ \text{diag}(n^{-1}, n^{-1/2}) V(Y; \theta^*, \rho^*) \text{diag}(n^{-1}, n^{-1/2}) \rightarrow \text{diag}\left\{ (\sigma^0)^2 \int_0^1 W^2(r) \, dr, V(\theta^0) \right\}, \]
where $V(\theta^*)$ is defined in (4.12).

Proof. The limiting distribution of $V(Y; \tilde{\theta}, \rho)$ is given in (4.47). The limiting distribution of $V(Y; \tilde{\theta}, \rho)$ and $V(Y; \hat{\theta}, \rho)$ can be obtained analogously. \qed
5. MODEL WITH AN INTERCEPT

5.1. Model Description

In this chapter, we consider the model

\[ y_t = \mu + \rho y_{t-1} + z_t, \]  
\[ z_t + \alpha_1 z_{t-1} + \cdots + \alpha_p z_{t-p} = \epsilon_t + \beta_1 e_{t-1} + \cdots + \beta_q e_{t-q}, \quad t = 1,2,...,n. \]

As in Chapter 1, let

\[ A(m) = m^p + \alpha_1 m^{p-1} + \cdots + \alpha_p \]  
\[ B(m) = m^q + \beta_1 m^{q-1} + \cdots + \beta_q \]

be the characteristic equations associated with \( \{z_t\} \) in model (5.1). We investigate the limiting behavior of estimators under the stationarity and invertibility condition on \( \{z_t\} \), an identifiability condition of \( \theta \), and the assumption that \( (\rho, \mu) = (1,0) \).

When \( \rho = 1 \), the process \( \{y_t\} \) is nonstationary. See Section 1.1 for the definition of stationarity and invertibility. Dickey and Fuller (1979) considered the autoregressive model

\[ y_t = \mu + \rho y_{t-1} + e_t, \quad t = 1,2,...,n, \]

where \( y_0 = 0 \) and the \( \{e_t\} \) is an iid \((0,\sigma^2)\) sequence. When \( \rho = 1 \), they derived the limiting distribution of the ordinary least squares estimator.
\[ \hat{\rho} = \frac{\sum_{t=1}^{n} (y_t - \bar{y}(0))(y_{t-1} - \bar{y}(-1))}{\sum_{t=1}^{n} (y_{t-1} - \bar{y}(-1))^2}, \]

where

\[ \bar{y}(0) = \frac{n}{i=1} y_i/n, \quad \bar{y}(-1) = \frac{n}{i=1} y_{i-1}/n. \]

The limiting distribution of \( \hat{\rho} \) is

\[ n(\hat{\rho} - 1) \Rightarrow 2^{-1}(\Gamma - H^2)^{-1}(T^2 - 1 - 2TH), \]

where \( \Rightarrow \) denote the convergence in distribution,

\[ \Gamma = \sum_{i=1}^{\infty} \frac{\gamma_i^2 \zeta_i^2}{\gamma_i^2}, \quad T = \sum_{i=1}^{\infty} 2^{1/2} \gamma_i \zeta_i, \quad H = \sum_{i=1}^{\infty} 2^{1/2} \gamma_i \zeta_i, \quad \gamma_i = 2(-1)^{i+1}/(2i-1)i, \]

and \( \{\zeta_i\} \) is an iid \( \text{N}(0,1) \) sequence.

Utilizing expression (5.6) for the limiting distribution, they prepared a set of tables of the percentiles of the distribution by Monte Carlo simulation. One version of the table can be found in Fuller (1976, p. 371).

The formal assumptions about model (5.1) are given in Assumption 5.1.
Assumption 5.1. In model (5.1), \( \{e_i\} \) is an iid \((0, \sigma^2)\) sequence. The observations are \( y_0, y_1, \ldots, y_n \). We denote by \((\theta^0, \rho^0, \sigma^0, \mu^0)\) the true value of \((\theta, \rho, \sigma, \mu)\). The true \( \rho^0 \) is assumed to be 1 and true \( \mu^0 \) is assumed to be 0. Also \( \sigma^0 \) is assumed to be positive. We assume that the parameter space is such that for all \( \theta \), the equations \( A(m) = 0 \) and \( B(m) = 0 \) have roots with absolute value not greater than \( 1 - \eta \) for some \( \eta > 0 \) independent of \( \theta \). Also for any root \( m_a \) of \( A(m) = 0 \) and any root \( m_b \) of \( B(m) = 0 \) we assume \( |m_a - m_b| \geq 1 - \eta \). Denote the set of all those \( \theta \) satisfying the above conditions by \( \Theta \).

We now define several types of estimators for model (5.1). Let \( Y, Y_1, Z, \theta, \Gamma_n(\theta), \) and \( e \) be defined by (1.10) and 1 is the \( n \)-dimensional column vector of all 1's.

Definition 5.2. The least squares estimator \((\hat{\theta}, \hat{\rho}, \hat{\mu})\) is defined to be the \((\theta, \rho, \mu)\) which minimizes

\[
Q_n(\theta, \rho, \mu) = (Y - \rho Y_1 - \mu 1)' \Gamma_n^{-1}(\theta)(Y - \rho Y_1 - \mu 1) \tag{5.7}
\]

\[
= Z' \Gamma_n^{-1}(\theta)Z + 2(1 - \rho)Y_1' \Gamma_n^{-1}(\theta)Z + (1 - \rho)^2 Y_1' \Gamma_n^{-1}(\theta)Y_1 + \mu^2 1' \Gamma_n^{-1}(\theta)1 - 2\mu(1 - \rho)1' \Gamma_n^{-1}(\theta)Y_1 - 2\mu 1' \Gamma_n^{-1}(\theta)Z
\]

over \( \theta \times \mathbb{R}^2 \). The least squares estimator \( \hat{\sigma}^2 \) of \( \sigma^2 \) is
Definition 5.3. The maximum likelihood estimator \((\hat{\theta}, \hat{\rho}, \hat{\mu})\) is the \((\theta, \rho, \mu)\) which minimizes

\[
\left[ \det \{ T_n(\theta) \} \right]^{1/n} Q_n(\theta, \rho, \mu)
\]

over \(\theta \times \mathbb{R}^2\). The maximum likelihood estimator \(\hat{\sigma}^2\) of \(\sigma^2\) is

\[
\hat{\sigma}^2 = n^{-1} Q_n(\hat{\theta}, \hat{\rho}, \hat{\mu}).
\]

Note that the maximum likelihood estimator is, in fact, conditional on \(y_0\).

For convenience, we use the terminology 'maximum likelihood estimator' instead of the terminology 'conditional maximum likelihood estimator'.

Finally, we define the ordinary least squares estimator \((\hat{\theta}, \hat{\rho}, \hat{\mu}, \hat{\sigma})\). Since in (5.1) the process \(\{z_t\}\) is invertible, we can find a sequence \(\{d_j(\theta)\}\) such that

\[
e_t = \sum_{j=0}^{\infty} d_j(\theta) z_{t-j} \quad \text{for all } t = \ldots, -1, 0, 1, \ldots
\]

(5.9)

Given \((y_0, y_1, \ldots, y_n)\), let

\[
e_t(Y; \theta, \rho, \mu) = \sum_{j=0}^{t-1} d_j(\theta) z_{t-j} = \sum_{j=0}^{t-1} d_j(\theta) (y_{t-j} - \rho y_{t-1-j} - \mu), \quad t=1,2,\ldots,n.
\]

(5.10)
The $e_t(Y; \theta, \rho, \mu)$ are obtained from (5.9) by truncating the series at $t-1$.

**Definition 5.4.** The ordinary least squares estimator $(\theta, \rho, \mu)$ is the $(\theta, \rho, \mu)$ which minimizes, over $\theta \times \mathbb{R}^2$,

$$S_n(\theta, \rho, \mu) = \sum_{t=1}^{n} e_t^2(Y; \theta, \rho, \mu). \quad (5.11)$$

The ordinary least squares of estimator $\sigma^2$ of $\sigma^2$ is

$$\hat{\sigma}^2 = n^{-1}S_n(\theta, \rho, \mu).$$

### 5.2. Weak Consistency

First we establish weak consistency of the three estimators by checking a sufficient condition for weak consistency given by Wu (1981).

**Lemma 5.5. (Wu)** Let $X_n$ be an $n$-dimensional random vector whose distribution is indexed by some parameter $\xi \in \mathbb{R}^k$ for fixed $k$. Let $H_n(\xi)$ be a function of $X_n$ and $\xi$. Assume $\hat{\xi}_n$ is a minimizing value of $H_n(\xi)$. Suppose for any $\delta > 0$,

$$\lim_{n \to \infty} \mathbb{P}[\inf_{|\xi - \xi^0| > \delta} (H_n(\xi) - H_n(\xi^0)) > 0] = 1. \quad (5.12)$$

Then

$$\hat{\xi}_n \xrightarrow{P} \xi^0$$

in probability.
Proof. If \( \xi_n \rightarrow \xi^0 \) in probability is not true, there is a \( \delta > 0 \) such that

\[
P[|\hat{\xi}_n - \xi^0| > \delta] \text{ does not converge to 0.}
\]

Therefore, there is a subsequence \( \{n_k\} \) and \( \epsilon > 0 \) such that

\[
P[|\hat{\xi}_{n_k} - \xi^0| > \delta] > \epsilon \text{ for all } k = 1, 2, \ldots.
\]

Since \( |\hat{\xi}_{n_k} - \xi^0| > \delta \) implies \( [\inf_{\delta} |\xi - \xi^0| > \delta (H_{n_k}(\xi) - H_{n_k}(\xi^0)) \leq 0] \), we have

\[
P[\inf_{\delta} |\xi - \xi^0| > \delta (H_{n_k}(\xi) - H_{n_k}(\xi^0)) \leq 0] \geq P[|\hat{\xi}_{n_k} - \xi^0| > \delta] > \epsilon \text{ for } k = 1, 2, \ldots.
\]

Therefore,

\[
\limsup_{n \to \infty} P[\inf_{\delta} |\xi - \xi^0| > \delta (H_{n}(\xi) - H_{n}(\xi^0)) \leq 0] > 0
\]

i.e.

\[
\liminf_{n \to \infty} P[\inf_{\delta} |\xi - \xi^0| > \delta (H_{n}(\xi) - H_{n}(\xi^0)) > 0] < 1,
\]

contradicting (5.12).

We establish weak consistency of the three estimators defined in Definition 5.2 — Definition 5.4 by checking (5.12). First we summarize approximations of

\[
Z^n_1(\theta)Z, \ Y^n_1(\theta)Y_1, \ 1^n_1(\theta)1, \ Y^n_1(\theta)Z, \ 1^n_1(\theta)Y_1, \ 1^n_1(\theta)Z.
\]
Theorem 5.6. Consider model (5.1). Under Assumption 5.1, the followings are true

i) \( \sup_{\theta} |Z' \Gamma_n^{-1}(\theta) Z - b(\theta) \sum_{j=1}^{n} e_j^2| = o_p(n), \)

ii) \( \sup_{\theta} |Y_1' \Gamma_n^{-1}(\theta) Y_1 - c_0 d^2(\theta) \sum_{j=1}^{n} W_{j-1}| = O_p(n^{3/2}), \)

iii) \( \sup_{\theta} |1' \Gamma_n^{-1}(\theta) 1 - nd^2(\theta)| = O_p(1), \)

iv) \( \sup_{\theta} |Y_1' \Gamma_n^{-1}(\theta) Z| = O_p(n), \)

v) \( \sup_{\theta} |1' \Gamma_n^{-1}(\theta) Y_1 - c_0 d^2(\theta) \sum_{j=1}^{n} W_{j-1}| = O_p(n), \)

vi) \( \sup_{\theta} |1' \Gamma_n^{-1}(\theta) Z - c_0 d^2(\theta) \sum_{j=1}^{n} e_j| = O_p(1), \)

where

\[
b(\theta) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} v_i^O v_j^O \sum_{s=0}^{\infty} d_s |i-j| + s^i \]

\[
c_0 = \sum_{j=0}^{\infty} v_j^O, \quad d(\theta) = \sum_{j=0}^{\infty} d_j^j \]

\[
W_j = e_1 + \ldots + e_j, \quad j = 1, 2, \ldots, W_0 = 0, \]

for \( \{v_j^O\} \) and \( \{d_j\} \) defined in (2.6). Note that \( v_j^O, j = 0, 1, \ldots \) are the values of \( v_j \) in
(2.6) evaluated at $\theta = \theta^0$.

vii) The result i) — vi) are also true if $D_n^\prime D_n$ is used in place of $\Gamma_n^{-1}(\theta)$.

Proof. Since, if we show i) — vi) for $D_n^\prime D_n$ in place of $\Gamma_n(\theta)$ results i) — vi) follow from Lemma (3.5) and Lemma (2.8), we give only a proof for $D_n^\prime D_n$.

Proof of i). This is a consequence of (3.12).

Proof of ii). Following the argument applied to (3.39),

\[ Y_1D_n^\prime D_n Y_1 = d^2 \sum_{j=1}^{n} \sum_{k=1}^{j-1} z_{kj} + \sum_{j=1}^{n} \sum_{k=1}^{j-1} b_{n,k,j} z_{kj} + \sum_{j=1}^{n} \sum_{k=1}^{j-1} c_{n,k,j} z_{kj} \]

where $d$, $b_{n,k,j}$, and $c_{n,k,j}$ are defined in (3.40). Recall that $b_{n,k,j}$ and $c_{n,k,j}$ are uniformly exponentially declining as are shown in (3.41). Observe that, by Lemma 3.10—v),

\[ \sum_{j=1}^{n} \sum_{k=1}^{j-1} |z_{kj}| = O_p(n^{3/2}) \]  

and

\[ \sum_{j=1}^{n} \sum_{k=1}^{j-1} |z_{kj}| = O_p(n^{3/2}). \]

Also
\[ y_0 \sum_{i=1}^{n} \sum_{j=1}^{n} s_{i-j} d_{i-j} + s = O_p(n^{3/2}). \]  

(5.15)

Now let

\[ a_j = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1} d_{j+k_1+k_2}. \]

If we show

\[ \sup_{\theta} \left| - \sum_{j=1}^{n} \sum_{k=1}^{j-1} a_j z_k y_{j-1} + \sum_{j=1}^{n} \sum_{k=j}^{n} a_{k+1-j} z_k y_{j-1} \right| = O_p(n^{3/2}), \]  

(5.16)

we can conclude, from (3.41) and (5.13) -- (5.16) that

\[ \sup_{\theta} \left| Y_1' D_n' D_n Y_1 - d^2 \sum_{j=1}^{n} \sum_{k=1}^{j-1} z_k y_{j-1} \right| = O_p(n). \]

(5.17)

Now let's consider (5.16). We have

\[ - \sum_{j=1}^{n} \sum_{k=1}^{j-1} a_j z_k y_{j-1} + \sum_{j=1}^{n} \sum_{k=j}^{n} a_{k+1-j} z_k y_{j-1} \]

\[ = - \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} a_j z_k (\sum_{s=1}^{j-1} z_s + y_0) + \sum_{k=1}^{n} \sum_{j=1}^{k+1} a_{k+1-j} z_k (\sum_{s=1}^{j-1} z_s + y_0). \]

We can easily show (see Appendix 8.4.1)

\[ \sup_{\theta} \left| \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} a_j z_{k} y_{0} \right| = O_p(n) \]

(5.18)

and
\[
\sup_{\theta} \left| \sum_{k=1}^{n} \sum_{j=1}^{k} a_{k+1-j} z_{k} y_{0} \right| = O_{p}(n).
\]

(5.19)

Also note that

\[
S_{n} = \sum_{k=1}^{n-1} \sum_{j=1}^{k-1} a_{j-k} z_{k} z_{s} + \sum_{k=1}^{n} \sum_{j=1}^{k-1} a_{k+1-j} z_{k} z_{s}
\]

\[
- \sum_{k=1}^{n-1} \sum_{j=1}^{k} a_{j-k} z_{k} z_{s} - \sum_{k=1}^{n} \sum_{j=1}^{k} a_{k+1-j} z_{k} z_{s}
\]

\[
+ \sum_{k=1}^{n} \sum_{s=1}^{k-1} a_{k+1-j} z_{k} z_{s}.
\]

(5.20)

Here, observe that

\[
- \sum_{k=1}^{n-1} \sum_{s=1}^{k-1} a_{j-k} z_{k} z_{s} + \sum_{k=1}^{n} \sum_{s=1}^{k} a_{k+1-j} z_{k} z_{s}
\]

\[
\sum_{k=1}^{n-1} \sum_{j=1}^{s} a_{j-k} + \sum_{j=1}^{s} a_{j} - \sum_{j=1}^{s} a_{j-k}
\]

\[
+ \sum_{s=1}^{n-1} k
\]

\[
\sum_{s=1}^{n-1} \sum_{j=1}^{s} a_{k+1-j} z_{k} z_{s}
\]

\[
= \sum_{k=1}^{n-1} \sum_{s=1}^{k-1} a_{n,k,s} z_{k} z_{s} + R_{n},
\]

where
\[ a_{n,k,s} = \sum_{j=n-k+1}^{\infty} a_j - \sum_{j=k-s+1}^{\infty} a_j \]

and

\[ R_n = \sum_{s=1}^{n-1} \sum_{j=s+1}^{k} a_{k+1-j} z_n z_s. \] (5.21)

Observe that, for some \( M < \infty \) and \( \lambda \in (0,1) \),

\[ \sup_{\theta} |a_{n,k,s}| \leq M(\lambda^{n-k} + \lambda^{k-s}). \] (5.22)

Therefore, by Lemma 3.10-i) and Lemma 3.10-ii),

\[ \sup_{\theta} \sum_{k=1}^{n-1} |\sum_{j=1}^{k-1} a_{n,k,s} z_k z_s| = O_p(n). \] (5.23)

Also note that

\[ \sup_{\theta} |R_n| = \sup_{\theta} \sum_{s=1}^{n-1} |\sum_{j=s+1}^{k} a_{k+1-j} z_n z_s| \leq M \sum_{s=1}^{n-1} \sum_{j=s+1}^{k} \lambda^{k+1-j} |z_n z_s| \]

\[ \leq M(1-\lambda)^{-1} |z_n| \sum_{s=1}^{n-1} |z_s| = O_p(n). \]

Therefore,

\[ S_n = O_p(n). \] (5.24)

From (5.18), (5.19), and (5.24), we conclude (5.16) and hence (5.17). Finally, observe that, by Lemma 3.12-ii) and Appendix 8.4.2,
\[ \sum_{j=1}^{n} \sum_{k=1}^{j-1} z_{k} y_{j-1}^{2} = \sum_{j=1}^{n} y_{j-1}^{2} - y_{0} \sum_{j=1}^{n} y_{j-1} = x_{0}^{2} \sum_{j=1}^{n} W_{j-1}^{2} + O_{p}(n^{3/2}). \] (5.25)

Therefore from (5.17), (5.25), and boundedness of \( d^{2}(\theta) \), we conclude ii).

Proof of iii). Note that
\[ 1' D_{n} = (\sum_{j=0}^{0} d_{j}, \sum_{j=0}^{1} d_{j}, ..., \sum_{j=0}^{n-1} d_{j}). \] (5.26)

Therefore,
\[ 1' D_{n} D_{n} = \sum_{k=0}^{n-1} \sum_{j=0}^{k} (\sum_{j=0}^{n} d_{j})^{2} = \sum_{k=0}^{n-1} \sum_{j=0}^{k} (\sum_{j=0}^{n} d_{j} - \sum_{j=k+1}^{n} d_{j})^{2} \]
\[ = \sum_{k=0}^{n-1} \sum_{j=0}^{n} (\sum_{j=0}^{k} d_{j})^{2} - 2 \sum_{k=0}^{n-1} \sum_{j=0}^{k} (\sum_{j=0}^{n} d_{j}) (\sum_{j=k+1}^{n} d_{j}) + \sum_{k=0}^{n-1} \sum_{j=k+1}^{n} (\sum_{j=0}^{n} d_{j})^{2} \]
\[ = S_{1} - 2S_{2} + S_{3}, \] say. (5.27)

Now note that
\[ \sup_{\theta} |S_{1} - nd^{2}| = \sup_{\theta} |n(d - \sum_{j=n+1}^{n} d_{j})^{2} - nd^{2}| = O(1), \]
\[ \sup_{\theta} |S_{2}| = O(1), \]
and
\[ \sup_{\theta} |S_{3}| = O(1). \]

Therefore,
\[ \sup_{\theta} |1' D_{n} D_{n} - nd^{2}| = O(1). \]

Proof of iv). This is proved in Lemma 3.11.
Proof of $v)$. Note that
\[
Y_1'D_n = \left( \sum_{j=0}^{n-1} d_jy_j, \sum_{j=0}^{1} (\sum_{j=0}^{n-1-1} d_jy_j), \ldots, \sum_{j=0}^{n-1} d_n - 1-j_jy_j \right).
\] (5.28)

Therefore, combining (5.26) and (5.28),
\[
1'D_n D_n Y_1 = \sum_{k=0}^{n-1} \left( \sum_{j=0}^{k} d_j \right) \left( \sum_{j=0}^{k} d_{k-j}y_j \right) = \sum_{k_1, k_2=0}^{k} d_{k_1} d_{k_2} y_{k_1}.
\] (5.29)

Note that
\[
\sum_{k_1, k_2=0}^{n-1} d_{k_1} d_{k_2} = \sum_{k_1, k_2=0}^{n-1} \sum_{k_1, k_2=k+1}^{n-1} d_{k_1} d_{k_2}
\]
\[
= \sum_{k_1, k_2=0}^{n-1} d_{k_1} d_{k_2} - \sum_{k_1, k_2=0}^{n-1} \sum_{k_1, k_2=k+1}^{n-1} d_{k_1} d_{k_2}
\]
\[
= d^2 - a_{n, j_1} \quad \text{say.}
\]

Now, letting $M$ and $\lambda$ be the coefficients of uniformly exponential decline of $\{d_j\}$,
\[
\sup_\theta |a_{n, j_1}| \leq M^2 \left[ \sum_{k=n}^{\infty} \sum_{j_2=0}^{\infty} \lambda^{k-j_1+j_2} + \sum_{k_1, k_2=k+1}^{\infty} \lambda^{k-j_1+j_2} \right]
\]
\[
\leq M^2 (1-\lambda)^{-1} \left[ \sum_{k=n}^{\infty} \lambda^{k-j_1} + \lambda^{j_1} (1-\lambda^2)^{-1} \right]
\]
\[
\leq M^2 (1-\lambda)^{-1} [(1-\lambda)^{-1} \lambda^{n-j_1} + \lambda^{j_1} (1-\lambda^2)^{-1}].
\] (5.30)
From (5.30), we conclude that
\[
\sup_{\theta} \left| \sum_{j_1=1}^{n-1} a_{n,j_1} y_{j_1} \right| = O_p(n). \tag{5.31}
\]

See Appendix 8.4.3 for details. Therefore,
\[
\sup_{\theta} \left| 1'D_n' D_n' Y_1 - d^2 \sum_{j=0}^{n-1} y_j \right| = O_p(n). \tag{5.32}
\]

Now it is easy to show (see Appendix 8.4.2)
\[
\sum_{j=1}^{n-1} y_j = c_0 \sum_{j=1}^{n-1} W_{j-1} + O_p(n). \tag{5.33}
\]

Therefore, from (5.32) and (5.33), result v) follows.

Proof of vi). By the same argument applied to get (5.32) and by using \( \sum_{j_1=1}^{n-1} a_{n,j_1} z_{j_1} = O_p(1) \) instead of using (5.31), we get
\[
\sup_{\theta} \left| 1'D_n' D_n' Z - d^2 \sum_{j=0}^{n-1} z_j \right| = O_p(1). \tag{5.34}
\]

We have
\[
\sum_{j=1}^{n} z_j = c_0 \sum_{j=1}^{n} e_j + O_p(1). \tag{5.35}
\]
See (8.16) in 8.3.2. Therefore the result follows from (5.34), (5.35), and the boundedness of \( d = d(\theta) \) in \( \theta \in \Theta \).

In Lemma 5.7, utilizing Theorem 5.6, we verify condition (5.12) in Lemma 5.5 for \( Q_n(\theta, \rho, \mu) \) and \( S_n(\theta, \rho, \mu) \) defined in Definition 5.2 and Definition 5.4. Lemma 5.7, combined with Lemma 5.5, gives us the weak consistency of the least squares estimator and the ordinary least squares estimator of the parameters in model (5.1) under Assumption 5.1.

**Lemma 5.7.** Consider model (5.1). Suppose that Assumption 5.1 holds. Let \( Q_n(\theta, \rho, \mu) \) and \( S_n(\theta, \rho, \mu) \) be defined in Definition 5.2 and Definition 5.4. Given \( \delta > 0 \), let

\[
\Theta_\delta = \{ (\theta, \rho, \mu) \in \Theta \times \mathbb{R}^2 : |(\theta - \theta^0, \rho - \rho^0, \mu - \mu^0)| \geq \delta \}.
\]

Then, for any \( \delta > 0 \),

i) \( \lim_{n \to \infty} P[\inf_{\Theta_\delta} \{ Q_n(\theta, \rho, \mu) - Q_n(\theta^0, \rho^0, \mu^0) \} > 0] = 1 \)

and

ii) \( \lim_{n \to \infty} P[\inf_{\Theta_\delta} \{ S_n(\theta, \rho, \mu) - S_n(\theta^0, \rho^0, \mu^0) \} > 0] = 1 \).

**Proof.** We give only a proof of i) because once one of i) or ii) established, the other follows directly from Lemma 2.8 and Lemma 3.5. We can write

\[
Q_n(\theta, \rho, \mu) - Q_n(\theta^0, \rho^0, \mu^0)
\]

\[
= Z' \{ \Gamma_n^{-1}(\theta) - \Gamma_n^{-1}(\theta^0) \} Z + (1 - \rho)^2 Y_1' \Gamma_n^{-1}(\theta) Y_1 - 2 \mu(1 - \rho) Y_1' \Gamma_n^{-1}(\theta) Y_1
\]
\[ + \mu^2 \Gamma_n^{-1}(\theta) \frac{1}{2} + 2(1-\rho)Y_1 \Gamma_n^{-1}(\theta)Z - 2\mu^1 \Gamma_n^{-1}(\theta)Z \]

\[ = \{b(\theta) - b(\theta')\} \sum_{j=1}^{n} e_j^2 + (1-\rho)^2 \{c_0^2 d^2 \sum_{j=1}^{n} W_{j-1}^2 + R_{1n}\} - 2\mu(1-\rho)c_0 d^2 (\sum_{j=1}^{n} W_{j-1} + R_{3n}) \]

\[ + n\mu^2 d^2 + (1-\rho)R_{2n} + R_{4n} \]

where

\[ b(\theta) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} v_i^0 v_j^0 \sum_{s=0}^{\infty} d_s |i-j| + s' \]

\[ d = d(\theta) = \sum_{j=0}^{\infty} d_j(\theta), c_0 = \sum_{j=0}^{\infty} v_j^0, \]

\[ R_{1n} = Y_1 \Gamma_n^{-1}(\theta) Y_1 - c_0^2 d^2 \sum_{j=1}^{n} W_{j-1}^2, \]

\[ R_{2n} = 2Y_1 \Gamma_n^{-1}(\theta)Z, \]

\[ R_{3n} = 1 \Gamma_n^{-1}(\theta) Y_1 - c_0 d^2 \sum_{j=1}^{n} W_{j-1}, \]

\[ R_{4n} = \text{sum of all remaining terms.} \]  

(5.37)

Note that terms in \( R_{4n} \) correspond to terms i), iii), vi) in Theorem 5.6 without \( \sup |\cdot| \). Observe that

\[ \inf_{\theta} \{ Q_n(\theta, \rho, \mu) - Q_n(\theta^0, \rho^0, \mu^0) \} \]
\[ \inf_{\theta} \{ b(\theta) - b(\theta^0) \} \sum_{j=1}^{n} \epsilon_j^2 + \inf_{\theta} \{ (1-\rho)^2 \sum_{j=1}^{n} W_j^2 - r_{1n} \} \]
\[ - 2 \mu(1-\rho)c_0 v^2 \left( \left| \sum_{j=1}^{n} W_j - r_{3n} \right| \right) + n \mu^2 v^2 - |1-\rho|r_{2n} - r_{4n}, \] (5.38)

where \[ r_{1n} = \sup_{\theta} |R_{1n}|, r_{2n} = \sup_{\theta} |R_{2n}|, r_{3n} = \sup_{\theta} |R_{3n}|, r_{4n} = \sup_{\theta} |R_{4n}|, \]

and

\[ v = \inf_{\theta} |d(\theta)|. \]

By Theorem 5.6,

\[ r_{1n} = O_p(n^{3/2}), r_{2n} = O_p(n), r_{3n} = O_p(n), r_{4n} = o_p(n). \] (5.39)

Also, by an argument similar to that used in Appendix 8.2.33,

\[ d = \left| (1 - m_1^*) \cdots (1 - m_q^*) \right|^{-1} \left| (1 - m_1) \cdots (1 - m_p) \right| \]

where \((m_1, \ldots, m_p)\) and \((m_1^*, \ldots, m_q^*)\) are the roots of characteristic equations (5.2) and (5.3). Therefore, with \(\eta\) defined in Assumption 1.1,

\[ v \geq \eta^0 > 0. \]

Now let

\[ \theta_{1\delta} = \theta_\delta \cap \{ |\theta - \theta^0| \geq \delta/3 \}, \]
\[ \theta_{2\delta} = \theta_\delta \cap \{ |\rho - \rho^0| \geq \delta/3 \}, \]

and
\[ \theta_{3\delta} = \theta_\delta \cap \{ |\mu - \mu^0| \geq \delta/3 \} \]

Then
\[ \theta_\delta \subset \theta_{1\delta} \cup \theta_{2\delta} \cup \theta_{3\delta} \] (5.40)

Also let
\[ C_n(\theta) = (b(\theta) - b(\theta^0)) \sum_{j=1}^{n} e_j^2 \]

and
\[ D_n(\rho, \mu) = (1 - \rho)^2 \{ c_0^2 \sum_{j=1}^{n} W_{j-1}^2 - r_{1n} \} - 2|\mu(1 - \rho)c_0| \sum_{j=1}^{n} W_{j-1} r_{3n} + n\mu^2 \rho^2. \]

Since \( b(\theta) > b(\theta^0) \) for all \( \theta \in \Theta \), we have
\[ n^{-1} \inf_{\theta_\delta} C_n(\theta) \xrightarrow{P} \inf_{\theta_\delta} (b(\theta) - b(\theta^0))(\sigma_0^2) = 0. \] (5.41)

We have
\[ P[n^{-1} \inf_{\theta_\delta} \{ D_n(\rho, \mu) - |1 - \rho| r_{2n} - r_{4n} \} \geq 0] \to 1. \] (5.42)

See Appendix 8.4.4. Since \( b(\theta) > b(\theta^0) \) for \( \theta \neq \theta^0 \)(see Appendix 8.2.9)
\[ \lim_{n \to \infty} P[n^{-1} \inf_{\theta_\delta} C_n(\theta) > 0] = 1. \] (5.43)

Observe that
\[ D_n(\rho, \mu) - |1 - \rho| r_{2n} - r_{4n} \]
\[ \geq (1 - \rho)^2 c_0^2 \sum_{j=1}^{n} (W_{j-1} - W)^2 - (1 - \rho)^2 r_{1n} - |1 - \rho| r_{2n} - r_{4n}. \]
Since \( n^{-2} \sum_{j=1}^{n} (W_j - W)^2 \) converges in distribution to \( \int_{0}^{1} \{W(t) - \int_{0}^{1} W(\tau) d\tau\}^2 d\tau \)

which is positive a.e. we get, by (5.39),

\[
\lim_{n \to \infty} P[n^{-1} \inf_{\theta_{2\delta}} \{D_{n}(\rho, \mu) - |1-\rho| r_{2n} - r_{4n} \} > 0] = 1. \tag{5.44}
\]

Next we show

\[
\lim_{n \to \infty} P[n^{-1} \inf_{\theta_{3\delta}} \{D_{n}(\rho, \mu) - |1-\rho| r_{2n} - r_{4n} \} > 0] = 1. \tag{5.45}
\]

Let \( \epsilon = 1/10 \) and define

\[
A_{n} = \theta_{3\delta} \cap \{|\rho - 1| < n^{-1/4+\epsilon}\}
\]

and

\[
B_{n} = \theta_{3\delta} \cap \{|\rho - 1| \geq n^{-1/4+\epsilon}\}.
\]

Also let

\[
a_{n} = \inf_{A_{n}} \{D_{n}(\rho, \mu) - |1-\rho| r_{2n} - r_{4n}\}
\]

and

\[
b_{n} = \inf_{B_{n}} \{D_{n}(\rho, \mu) - |1-\rho| r_{2n} - r_{4n}\}. \tag{5.46}
\]

If we show

\[
\lim_{n \to \infty} P[n^{-1} a_{n} > 0] = 1 \quad \text{and} \quad \lim_{n \to \infty} P[n^{-1} b_{n} > 0] = 1 \tag{5.47}
\]

then we can say (5.45). Due to (5.39) we have
\[
\sup_{A_n} \left|1 - \rho \right| r_{2n} - r_{4n} = o_p(n) \tag{5.48}
\]

and there is \( n_0 \) such that \(|\mu| \geq \delta/4, \) for all \( n > n_0 \) and \( (\theta, \rho, \mu) \in A_n. \) \tag{5.49}

Therefore, by (5.48),

\[
\inf_{A_n} \{D_n(\rho, \mu) - |1 - \rho| r_{2n} - r_{4n}\} = \inf_{A_n} D_n(\rho, \mu) + o_p(n). \tag{5.50}
\]

Since \( \left( \sum_{j=1}^{n} W_j^2 \right)^{-1} r_{1n} = o_p(1) \) and \( \left( \sum_{j=1}^{n} W_{j-1} \right)^{-1} r_{3n} = o_p(1), \) the terms \( r_{1n} \) and \( r_{3n} \) have no effect on the limiting behavior of \( D_n(\rho, \mu). \) Hence we can set \( r_{1n} = r_{3n} = 0 \) for the simplicity of algebra. Setting \( r_{1n} = r_{3n} = 0, \) we have

\[
V^{-2} D_n(\rho, \mu) = [(1 - \rho)^2 c_0^2 \sum_{j=1}^{n} W_j^2 j-1 - 2|\mu(1 - \rho) c_0 \sum_{j=1}^{n} W_{j-1}| + n \mu^2]
\]

\[
= \sum_{j=1}^{n} W_j^2 j-1 [(1 - \rho) c_0 - |\mu \sum_{j=1}^{n} W_{j-1}|/(\sum_{j=1}^{n} W_{j-1})]^2
\]

\[
+ \mu^2 \left[ n - (\sum_{j=1}^{n} W_{j-1})^2 / (\sum_{j=1}^{n} W_{j-1}) \right] \]

\[
\geq \sum_{j=1}^{n} W_j^2 j-1 [(1 - \rho) c_0 - |\mu \sum_{j=1}^{n} W_{j-1}|/(\sum_{j=1}^{n} W_{j-1})]^2
\]
\[
+ n\mu^2 \left( \sum_{j=1}^{n} W_{j-1} - \overline{W} \right)^2 / \sum_{j=1}^{n} W_{j-1}^2. \tag{5.51}
\]

Therefore, combining (5.49) and (5.51), for \( n > n_0 \)

\[
n^{-1}a_n \geq (\delta/4)^2 \nu^2 \left( \sum_{j=1}^{n} W_{j-1} - \overline{W} \right)^2 / \sum_{j=1}^{n} W_{j-1}^2 + o_p(1) \tag{5.52}
\]

\[
\Rightarrow (\delta/4)^2 \nu^2 \left[ \int_0^1 W(r) - \int_0^1 W(t)dt \right]^2 dr / \int_0^1 W^2(r)dr. \tag{5.53}
\]

Since the distribution in (5.53) is positive a.e.,

\[
l \lim \ P[n^{-1}a_n > 0] = 1. \tag{5.54}
\]

Next, on \( B_n \), setting \( r_{1n} = r_{3n} = 0 \),

\[
D_n(\rho,\mu) = |1-\rho|r_{2n} - r_{4n}
\]

\[
n^{-1/2+2\epsilon} c_0 \nu^2 \sum_{j=1}^{n} W_{j-1}^2 - 2n^{-1/4+\epsilon}|\mu c_0| \nu^2 \sum_{j=1}^{n} W_{j-1}^2 - \{(1+\rho)_{2n} + r_{4n}\}\]

\[
n^{-1/2+2\epsilon} c_0 \nu^2 \sum_{j=1}^{n} W_{j-1}^2 + n\mu^2 \nu^2 - 2|\mu c_0| \sum_{j=1}^{n} W_{j-1}^2 n^{-1/4+\epsilon} + o_p(n^{-1/2+2\epsilon+2})
\]

\[
n^{-1/2+2\epsilon} c_0 \nu^2 \sum_{j=1}^{n} (W_{j-1} - \overline{W})^2 + o_p(n^{3/2+2\epsilon}). \tag{5.55}
\]
Since
\[ n^{-2} \sum_{j=1}^{n} (W_{j-1} - \bar{W})^2 \to (\sigma^o)^2 \int_{0}^{1} \{W(r) - \int_{0}^{1} W(t) \, dt\}^2 \, dr \quad (5.56) \]
we have
\[ P[n^{-1} b_n > 0] = P[n^{-3/2} - 2 \epsilon b_n > 0] \to P[(\sigma^o)^2 \int_{0}^{1} W^2(r) \, dr > 0] = 1. \quad (5.57) \]

Therefore, combining (5.38)-(5.45),
\[ \lim_{n \to \infty} P[\inf_{\theta} \{Q_n(\theta, \rho, \mu) - Q_n(\theta, \rho^o, \mu^o)\} > 0] \]
\[ \geq \lim_{n \to \infty} P[\inf_{\theta} \{C_n(\theta) + D_n(\rho, \mu) - |1-\rho| r_{2n} - r_{4n}\} > 0] \quad \text{by (5.38)} \]
\[ \geq \lim_{n \to \infty} P[\inf_{\theta_1 \delta \cup \theta_2 \delta \cup \theta_3 \delta} n^{-1} \{C_n(\theta) + D_n(\rho, \mu) - |1-\rho| r_{2n} - r_{4n}\} > 0] \quad \text{by (5.40)} \]
\[ \geq \lim_{n \to \infty} P[n^{-1} \inf_{\theta_1 \delta \cup \theta_2 \delta} \inf_{\theta_3 \delta} C_n(\theta), \inf_{\theta_2 \delta} \{D_n(\rho, \mu) - |1-\rho| r_{2n} - r_{4n}\} > 0] \]
\[ \quad \text{by (5.41), (5.42)} \]
\[ \inf_{\theta_3 \delta} \{D_n(\rho, \mu) - |1-\rho| r_{2n} - r_{4n}\} > 0 \]
\[ = 1 \quad \text{by (5.43), (5.44), and (5.45)}. \]

In Theorem 5.8, we show the weak consistency of the three estimators of (\theta,
\( \rho, \mu, \sigma \) in model (5.1). The weak consistency of estimators of \( (\theta, \rho, \mu) \) is a consequence of Lemma 5.5 and Lemma 5.7.

**Theorem 5.8.** Consider model (5.1). Suppose Assumption 5.1 holds. Then \( (\hat{\theta}, \hat{\rho}, \hat{\mu}, \hat{\sigma}), (\tilde{\theta}, \tilde{\rho}, \tilde{\mu}, \tilde{\sigma}), \) and \( (\hat{\theta}, \hat{\rho}, \hat{\mu}, \hat{\sigma}) \) converge to \( (\theta^0, \rho^0, \mu^0, \sigma^0) \) in probability.

**Proof.** The weak consistency of \( (\hat{\theta}, \hat{\rho}, \hat{\mu}), (\tilde{\theta}, \tilde{\rho}, \tilde{\mu}) \) is a consequence of Lemma 5.5 and Lemma 5.7. We give a proof of weak consistency of \( \hat{\sigma} \). The proof of weak consistency of \( \tilde{\sigma} \) is similar. The weak consistency of \( \hat{\sigma} \) follows from the observations

\[
\hat{\sigma}^2 = n^{-1}Q_n(\hat{\theta}, \hat{\rho}, \hat{\mu}) \leq n^{-1}Q_n(\theta^0, \rho^0, \mu^0) \xrightarrow{p} (\sigma^0)^2
\]  
(5.58)

and

\[
\tilde{\sigma}^2 = n^{-1}Q_n(\tilde{\theta}, \tilde{\rho}, \tilde{\mu}) = n^{-1}[Z' \Gamma_n^{-1}(\tilde{\theta})Z + (1-\tilde{\rho})^2 Y_1' \Gamma_n^{-1}(\tilde{\theta})Y_1
\]
\[ -2\tilde{\mu}(1-\tilde{\rho})1' \Gamma_n^{-1}(\tilde{\theta})Y_1 + \tilde{\mu}^2 1' \Gamma_n^{-1}(\tilde{\theta})1 + 2(1-\tilde{\rho})Y_1' \Gamma_n^{-1}(\tilde{\theta})Z - 2\tilde{\mu}1' \Gamma_n^{-1}(\tilde{\theta})Z]
\]  
(5.59)

\[
\geq n^{-1}\inf_{\rho} \left[ (1-\rho)^2 \nu_1 Y_1 Y_1 + 2|1-\rho| \sup_{\theta} |Y_1' \Gamma_n^{-1}(\theta)Z - |\tilde{\mu}1' \Gamma_n^{-1}(\theta)Y_1| \right] + o_p(1)
\]

\[
\geq n^{-1}\inf_{\rho} \left[ \mu_1 Y_1' \Gamma_n^{-1}(\theta)Z - n^{-1}\sup_{\theta} |Y_1' \Gamma_n^{-1}(\theta)Z - \tilde{\mu}1' \Gamma_n^{-1}(\theta)Y_1| \right]^{1/2} \xrightarrow{p} (\sigma^0)^2,
\]
where \( \nu_0 = \inf_{\theta} \| \mathbf{r}_n^{-1}(\theta) \| > 0. \) See Theorem 2.10. The weak convergence of \( n^{-1} \mathbf{r}_n^{-1}(\theta) \mathbf{Z} \) to \( (\sigma^2) \) is a consequence of Lemma 3.7–iii). A proof of weak consistency of \((\hat{\theta}, \hat{\rho}, \hat{\mu})\) can be obtained by the same argument given in the proof of Theorem 3.8. A proof of weak consistency of \( \hat{\sigma} \) can be obtained by the same argument given in the proof of Theorem 3.15.

In Lemma 5.9, we improve the order of \((\bar{\rho} - 1), (\rho - 1)\) and \((\hat{\rho} - 1)\) to \( o_p(n^{-1/2}) \). An order of \( o_p(n^{-1/2}) \) for \((\bar{\rho} - 1), (\rho - 1)\) and \((\hat{\rho} - 1)\) is necessary for the expansion used in establishing the limiting distribution of \((\bar{\rho} - 1), (\rho - 1)\) and \((\hat{\rho} - 1)\).

**Lemma 5.9.** Consider model (5.1). Let Assumption 5.2 hold. Then

1. \( n^{1/2}(\bar{\rho} - 1) = o_p(1), \)
2. \( n^{1/2}(\rho - 1) = o_p(1), \)
3. \( n^{1/2}(\hat{\rho} - 1) = o_p(1). \)

**Proof.** Since the proofs of ii) and iii) are similar to the proof of i), we give only a proof of i). From (5.59) and the consistency of \((\bar{\rho}, \bar{\mu}, \bar{\sigma})\), and using \( n^{-1} \mathbf{Z}' \mathbf{r}_n^{-1}(\bar{\theta}) \mathbf{Z} \rightarrow^p \mathbf{Z}' \mathbf{r}_n^{-1}(\bar{\theta}) \mathbf{Y}_1 = o_p(n), \) and \( 1' \mathbf{r}_n^{-1}(\bar{\theta}) \mathbf{Z} = o_p(n^{1/2}), \) we have
Therefore, from Theorem 5.6,
\[
S_n = n^{-1}[(1-\rho)^2 Y_1 \Gamma_n^{-1}(\delta) Y_1 - 2\mu(1-\rho)1' \Gamma_n^{-1}(\delta) Y_1 + \mu^2 1' \Gamma_n^{-1}(\delta) 1] \longrightarrow 0.
\]

where \( R_n = d^{-2}(\delta) Y_1 \Gamma_n^{-1}(\delta) Y_1 - c_0^2 \sum_{j=1}^{n} W_{j-1}^2 \). Therefore,
\[
S_n = \{\mu - n^{-1}(1-\rho)c_0 \sum_{j=1}^{n} W_{j-1}\}^2 + n^{-1}(1-\rho)^2 c_0^2 \{\sum(W_{j-1} - \overline{W})^2 + c_0^2 R_n\} \longrightarrow 0. \tag{5.60}
\]

Consequently,
\[
n^{-1}(1-\rho)^2 c_0^2 \{\sum(W_{j-1} - \overline{W})^2 + c_0^2 R_n\} \longrightarrow 0. \tag{5.61}
\]

By Theorem 5.6, \( R_n = O_p(n^{3/2}) \) and hence
\[
n^{-2}\{\sum(W_{j-1} - \overline{W})^2 + c_0^2 R_n\} \Rightarrow (\sigma^2)^2 \int_{0}^{1} \{W(r) - \int_{0}^{r} W(t) dt\}^2 dr. \tag{5.62}
\]

Therefore,
\[
n^{-1}(1-\rho)^2 = o_p(1)
\]

by the argument used in Appendix 8.2.34. \( \square \)
5.3. Derivatives of $Q_n$

We define the partial derivatives of $Q_n = Q_n(\theta, \rho, \mu)$. We use the same notation for derivatives that was used in Chapter 4 because no confusion will result.

The partial derivatives of

$$Q_n(\theta, \rho, \mu) = Z \Gamma_n^{-1}(\theta)Z + (1-\rho)^2 Y_1 \Gamma_n^{-1}(\theta)Y_1 - 2\mu (1-\rho) 1' \Gamma_n^{-1}(\theta)Y_1$$

$$+ \mu^2 1' \Gamma_n^{-1}(\theta)1 + 2(1-\rho) Y_1' \Gamma_n^{-1}(\theta)Z - 2\mu 1' \Gamma_n^{-1}(\theta)Z$$

are

$$Q_\rho = \frac{\partial Q}{\partial \rho} = -2(1-\rho) Y_1' \Gamma_n^{-1}(\theta)Y_1 + 2\mu 1' \Gamma_n^{-1}(\theta)Y_1 - 2Y_1' \Gamma_n^{-1}(\theta)Z,$$

$$Q_\mu = \frac{\partial Q}{\partial \mu} = -2(1-\rho) Y_1' \Gamma_n^{-1}(\theta)Y_1 + 2\mu 1' \Gamma_n^{-1}(\theta)1 - 2Y_1' \Gamma_n^{-1}(\theta)Z,$$

$$Q_{\theta_i} = \frac{\partial Q}{\partial \theta_i} = (Y - \rho Y_1 - \mu 1)' G_{\theta_i} (Y - \rho Y_1 - \mu 1),$$

$$Q_{\rho \rho} = \frac{\partial^2 Q}{\partial \rho^2} = 2Y_1' \Gamma_n^{-1}(\theta)Y_1,$$

$$Q_{\mu \mu} = \frac{\partial^2 Q}{\partial \mu^2} = 2Y_1' \Gamma_n^{-1}(\theta)1,$$

$$Q_{\rho \mu} = \frac{\partial^2 Q}{\partial \rho \partial \mu} = 2Y_1' \Gamma_n^{-1}(\theta)Y_1.$$
\[
Q_{\rho\theta_1} = \frac{\partial^2 Q}{\partial\rho \partial \theta_1} = -2(1-\rho)Y_1'G_{\theta_1}Y_1 + 2\mu_1'G_{\theta_1}Y_1 - 2Y_1'G_{\theta_1}Z,
\]

\[
Q_{\mu\theta_1} = \frac{\partial^2 Q}{\partial\mu \partial \theta_1} = -2(1-\rho)Y_1'G_{\theta_1}Y_1 + 2\mu_1'G_{\theta_1}Y_1 - 2Y_1'G_{\theta_1}Z,
\]

\[
Q_{\theta_1,\theta_j} = \frac{\partial^2 Q}{\partial \theta_1 \partial \theta_j} = (Y - \rho Y_1 - \mu_1)'G_{\theta_1,\theta_j}(Y - \rho Y_1 - \mu_1),
\]

\[
Q_{\rho\rho\rho} = Q_{\rho\rho\mu} = Q_{\rho\mu\mu} = Q_{\mu\mu\mu} = 0,
\]

\[
Q_{\rho\rho\theta_1} = \frac{\partial^3 Q}{\partial\rho^2 \partial \theta_1} = 2Y_1'G_{\theta_1}Y_1,
\]

\[
Q_{\mu\mu\theta_1} = \frac{\partial^3 Q}{\partial\mu^2 \partial \theta_1} = 2\cdot 1'G_{\theta_1}Y_1,
\]

\[
Q_{\rho\rho\rho\theta_1} = \frac{\partial^3 Q}{\partial\rho\partial\mu \partial \theta_1} = 2\cdot 1'G_{\theta_1}Y_1,
\]

\[
Q_{\rho\theta_1,\theta_j} = \frac{\partial^3 Q}{\partial\rho \partial\theta_1 \partial \theta_j} = -2(1-\rho)Y_1'G_{\theta_1,\theta_j}Y_1 + 2\mu_1'G_{\theta_1,\theta_j}Y_1 - 2Y_1'G_{\theta_1,\theta_j}Z,
\]
\[ \frac{\partial^2 Q}{\partial \theta_i \partial \theta_j} = -2(1-\rho) \frac{\partial G_{1,0,1}}{\partial \theta_i} Y_1 + 2\mu \frac{\partial G_{1,0,1}}{\partial \theta_j} Y_1 - 2\cdot 1' \frac{\partial G_{1,0,1}}{\partial \theta_j} Z, \]

and
\[ \frac{\partial^2 Q}{\partial \theta_i \partial \theta_j \partial \theta_k} = (Y - \rho Y_1 - \mu 1)' G_{1,0,1} (Y - \rho Y_1 - \mu 1), \quad (5.63) \]

for \( i, j, k = 1, 2, \ldots, r \), where \( G_{1,0,1}, G_{1,0,1}, G_{1,0,1} \) are the partial derivatives of \( \Gamma_n^{-1}(\theta) \) and \( r = p + q \). As in Chapter 4, let
\[
Q_{\theta} = (Q_{\theta}, Q_{\rho}, Q_{\mu}), \quad Q_{\rho \theta} = (Q_{\rho \theta 1}, \ldots, Q_{\rho \theta r}), \quad Q_{\mu \theta} = (Q_{\mu \theta 1}, \ldots, Q_{\mu \theta r}),
\]
\[
Q_{\rho \rho \theta} = (Q_{\rho \rho \theta 1}, \ldots, Q_{\rho \rho \theta r}), \quad Q_{\rho \mu \theta} = (Q_{\rho \mu \theta 1}, \ldots, Q_{\rho \mu \theta r}), \quad Q_{\mu \rho \theta} = (Q_{\mu \rho \theta 1}, \ldots, Q_{\mu \rho \theta r}).
\]

Also let \( Q_{\theta \theta} \), \( Q_{\rho \rho} \), \( Q_{\mu \mu} \) be the \( r \times r \) matrices with \((i,j)\) elements \( Q_{\theta_i \theta_j}, Q_{\rho_i \rho_j}, Q_{\mu_i \mu_j} \), respectively. For each \( i = 1, 2, \ldots, r \), let \( Q_{\theta_i \theta} \) be the \( r \times r \) matrix with \((j,k)\) element \( Q_{\theta_i \theta_j \theta_k} \). Let \( Q_{\theta \theta \theta} = (Q_{\theta_1 \theta \theta}, \ldots, Q_{\theta_r \theta \theta}) \). Also let
\[
Q^0 = Q_n(\theta^0, \rho^0, \mu^0), \quad Q^* = Q_n(\theta^*, \rho^*, \mu^*),
\]
\[
\bar{Q} = Q_n(\bar{\theta}, \bar{\rho}, \bar{\mu}), \quad \hat{Q} = Q_n(\hat{\theta}, \hat{\rho}, \hat{\mu}), \quad \check{Q} = Q_n(\check{\theta}, \check{\rho}, \check{\mu}), \quad (5.65)
\]

where \((\theta^*, \rho^*, \mu^*)\) is a vector to be defined. We use similar notation for \( Q_{\rho}, Q_{\rho \rho}, Q_{\rho \mu}, Q_{\mu \rho}, Q_{\rho \mu \rho}, Q_{\rho \mu \mu}, Q_{\mu \rho \rho}, Q_{\mu \rho \mu}, Q_{\mu \mu \rho}, Q_{\mu \mu \mu}, Q_{\rho \rho \rho}, Q_{\rho \rho \mu}, Q_{\rho \mu \rho}, Q_{\rho \mu \mu}, Q_{\mu \rho \rho}, Q_{\mu \rho \mu}, Q_{\mu \mu \rho}, Q_{\mu \mu \mu} \).
evaluated at the different vectors. For example $Q^0_{\theta}$ is the vector of first order partial
derivatives of $Q_n$ evaluated at $(\theta, \rho, \mu) = (\theta^0, \rho^0, \mu^0)$.

5.4. Limiting Distribution of Estimators

In Lemma 5.10, we establish the orders in probability of the third order
derivatives of $Q_n(\theta, \rho, \mu)$.

**Lemma 5.10.** Consider model 5.1. Suppose Assumption 5.2 holds. If $(\rho^* - 1) = o_p(n^{-1/2})$, $\mu^* = o_p(1)$, and $(\theta^* - \theta^0) = o_p(1)$, then

$$Q_{\rho \rho \theta}^* = o_p(n^2), \quad Q_{\rho \mu \theta}^* = o_p(n^{3/2}), \quad Q_{\mu \mu \theta}^* = o_p(n),$$

$$Q_{\rho \theta \theta}^* = o_p(n^{3/2}), \quad Q_{\mu \theta \theta}^* = o_p(n), \quad Q_{\theta \theta \theta}^* = o_p(n). \quad (5.66)$$

**Proof.** This is a consequence of Theorem 5.6, (5.63)–(5.65), and Theorem 2.14. □

By Lemma 5.9, $(\tilde{\rho} - 1), (\tilde{\rho} - 1)$, and $(\tilde{\rho} - 1)$ are $o_p(n^{-1/2})$. Therefore, any
$(\rho^*, \mu^*, \theta^*)$ between $(\rho^0, \mu^0, \theta^0)$ and $(\tilde{\rho}, \tilde{\mu}, \tilde{\theta})$, between $(\rho^0, \mu^0, \theta^0)$ and $(\rho, \mu, \theta)$ or
between $(\rho^0, \mu^0, \theta^0)$ and $(\hat{\rho}, \hat{\mu}, \hat{\theta})$ satisfy (5.66).

In Lemma 5.11, we establish the orders of $Q^0_{\rho \theta}$ and $Q^0_{\mu \theta}$. 
Lemma 5.11. Consider model 5.1. Suppose Assumption 5.2 holds. Then

\begin{enumerate}
\item \( Q^o_{\theta} = O_p(n) \),
\item \( Q^o_{\mu \theta} = O_p(n^{1/2}) \).
\end{enumerate}

Proof of i). This is shown in Lemma 4.2 — i).

Proof of ii). From (5.63), we have

\[ Q^o_{\mu \theta_1} = -2 \cdot 1' G^o_{\theta_1} Z. \]

Therefore it suffices to show

\[ S_n = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{n,i_1,i_2,j_1,j_2} = O_p(n^{1/2}). \]

By the same argument applied to (4.8),

\[ S_n = \sum_{i_1=1}^{n-1} \sum_{j_2=1}^{n} a_{n,i_1,j_2} z_{i_2}. \]

where \( a_{n,i_1,j_2} \) is defined in (4.9). From (4.10), \( a_{n,i_1,i_2} \) is exponentially declining in \(|i_1-j_2|\). Therefore \( S_n = O_p(n^{1/2}) \). See Appendix 8.4.5 for details. \( \square \)
In Lemma 5.12, we establish the limiting distribution of \( (n^{-2}Q_{\rho\rho}^O, n^{-3/2}Q_{\rho\mu}^O, n^{-1}Q_{\rho}^O, n^{-1/2}Q_{\mu}^O, n^{-1/2}Q_{\theta}^O) \) and probability limits of \( n^{-1}Q_{\theta\theta}^O \) and \( n^{-1}Q_{\rho\mu}^O \). In Theorem 5.13, by the second order Taylor expansion of \( \bar{Q}_{\rho}, \bar{Q}_{\mu}, \bar{Q}_{\theta} \) we derive an expression of \( [n(\bar{\theta} - 1), n^{1/2}(\bar{\mu}, n^{1/2}(\bar{\theta} - \theta^2)] \) which is essentially function of \( (n^{-2}Q_{\rho\rho}^O, n^{-3/2}Q_{\rho\mu}^O, n^{-1}Q_{\rho}^O, n^{-1/2}Q_{\mu}^O, n^{-1/2}Q_{\theta}^O, n^{-1}Q_{\theta\theta}^O, n^{-1}Q_{\rho\mu}^O) \).

**Lemma 5.12.** Consider model 5.1. Suppose Assumption 5.2 holds. Then

i) \( 2^{-1}[n^{-2}Q_{\rho\rho}^O, n^{-3/2}Q_{\rho\mu}^O, n^{-1}Q_{\rho}^O, n^{-1/2}Q_{\mu}^O] \)

\[ \Rightarrow \left[ \int_0^1 W^2(r)dr, d_0 \right] \left[ \int_0^1 W(r)dr, -2^{-1}(W^2(1) - 1), -d_0 W(1) \right] \]

\[ 2^{-1}n^{-1}Q_{\rho\mu}^O \xrightarrow{d} d_0^2 \]

ii) \( n^{-1/2}Q_{\theta}^O \Rightarrow N_T(0, V(\theta)) \),

\[ n^{-1}Q_{\theta\theta}^O \xrightarrow{p} V(\theta) \]

where \( V(\theta) \) is defined in Lemma 4.12 and \( d_0 = \sum_{j=0}^\infty d_j(\theta^j) \).

**Proof of i).** From Theorem 5.6,
\[ 2^{-1}(n^{-1}Q_{\rho}^{0}, n^{-1/2}Q_{\mu}^{0}, n^{-2}Q_{\rho \rho}^{0}, n^{-3/2}Q_{\rho \mu}^{0}, n^{-1}Q_{\mu \mu}^{0})' \]

\[ = \left( -n^{-1} \sum_{j=1}^{n} W_{j-1}e^{j} - n^{-1/2} d_{0}W_{n}, n^{-2} \sum_{j=1}^{n} W_{j-1}^{2}, n^{-3/2} \sum_{j=1}^{n} W_{j-1}, d_{0}^{2} \right) ' + o_{p}(1). \]

Therefore, by Donsker's invariance principle (Theorem A.8 in Appendix A), and continuous mapping Theorem (Theorem A.9 in Appendix A) the result i) holds.

Proof of ii). This is done in Lemma 4.3–ii). □

We now give the main result of this section. The limiting distribution of

\[ n(\tilde{\rho} - 1), n(\bar{\rho} - 1), \text{ and } n(\hat{\rho} - 1) \]

is shown to be the same as that of Dickey and Fuller (1979). The limiting distribution of

\[ n^{1/2}(\bar{\theta} - \theta), n^{1/2}(\tilde{\theta} - \theta), n^{1/2}(\hat{\theta} - \theta) \]

is a multivariate normal distribution. Moreover the two limiting distributions are independent.
Theorem 5.13. Let model (5.1) and Assumption 5.1 hold. Then

\[
(n(\rho -1), n^{1/2} \mu)' =>
\]

\[
2^{-1} \left[ W^2(1) - W(1) \right] \left[ \int_0^1 W(r) \, dr \right] / \left[ \int_0^1 \{W(r) - \int_0^1 W(t) \, dt\}^2 \, dr \right]
\]

\[
c_0 \left[ -2^{-1} \left( W^2(1) - 1 \right) \right] \left[ \int_0^1 W(r) \, dr + W(1) \int_0^1 W^2(1) \, dr \right] / \left[ \int_0^1 \{W(r) - \int_0^1 W(t) \, dt\}^2 \, dr \right]
\]

and

\[
n^{1/2}(\hat{\theta} - \theta^0) => N_x(0, V^{-1}(\theta^0))
\]

where \( d_0 = \sum_{j=0}^{\infty} d_j(\theta^0) \) and \( V^{-1}(\theta^0) \) is defined in (4.12). Moreover the two limiting distributions are independent. The limiting distribution of

\[
(n(\rho -1), n^{1/2} \mu, n^{1/2}(\hat{\theta} - \theta^0))
\]

and

\[
(n(\rho -1), n^{1/2} \mu, n^{1/2}(\hat{\theta} - \theta^0))
\]

is the same as the limiting distribution of

\[
(n(\rho -1), n^{1/2} \mu, n^{1/2}(\hat{\theta} - \theta^0)).
\]

Proof. The second order Tylor expansion of \((\tilde{Q}_{\rho}, \tilde{Q}_{\mu}, \tilde{Q}_{\theta})\) at \((\theta^0, \rho^0, \mu^0)\) is
where \((\theta^*, \rho^*, \mu^*)\) which defines \(Q_{\rho \theta}^*\), \(Q_{\rho \mu}^*\), \(Q_{\mu \theta}^*\), \(Q_{\rho 0 0}^*\), \(Q_{\mu 0 0}^*\), \(Q_{\theta 0 0}^*\) is between \((\theta^0, \rho^0, \mu^0)\) and \((\bar{\theta}, \bar{\rho}, \bar{\mu})\), and \((\bar{\theta} - \theta^0)\bar{Q}_{\theta 0 0}^*\) is defined in (4.29). From Theorem 5.8, Lemma 5.9, Lemma 5.10, and Lemma 5.11,

\[
(\bar{\theta} - \theta^0)'Q_{\rho \rho \theta}^* = o_p(n^2),
\]

\[
(\bar{\theta} - \theta^0)'Q_{\rho \mu \theta}^* = o_p(n^{3/2}),
\]

\[
Q_{\rho \rho \theta}^*(\bar{\rho} - 1) + Q_{\rho \mu \theta}^*\bar{\mu} + (\bar{\theta} - \theta^0)'Q_{\rho \theta \theta}^* = o_p(n^{3/2}),
\]
$$(\tilde{\theta} - \theta^0)^* Q_{\mu\mu \theta} = o_p(n),$$

$${Q}_{\mu\mu \theta}(\rho - 1) + Q_{\mu\mu \theta}^* \tilde{\mu} + (\tilde{\theta} - \theta^0)^* Q_{\mu\mu \theta} = o_p(n),$$

and

$$Q_{\mu\mu \theta}(\rho - 1) + Q_{\mu\mu \theta}^* \tilde{\mu} + (\tilde{\theta} - \theta^0)^* Q_{\mu\mu \theta} = o_p(n).$$

(5.69)

Also, from Lemma 5.11

$$Q_{\rho \theta}^0 = O_p(n)$$

and

$$Q_{\mu \theta}^0 = O_p(n^{1/2}).$$

Therefore,

$$\begin{pmatrix}
\frac{n(\tilde{\rho} - 1)}{n^{1/2}} \\
\frac{n^{1/2}}{\tilde{\theta} - \theta^0}
\end{pmatrix}
= \begin{pmatrix}
\frac{n^{-2} Q_{\rho \rho}}{o_p(1)} + \frac{n^{-3/2} Q_{\rho}^0}{o_p(1)} + \frac{p}{o_p(1)}^{1/2} - 1
\frac{n^{-2} Q_{\rho \mu}}{o_p(1)} + \frac{n^{-1} Q_{\rho \theta}}{o_p(1)} + \frac{p}{o_p(1)}
\end{pmatrix}
\begin{pmatrix}
\frac{n^{-1} Q_{\rho}}{o_p(1)}
\frac{n^{-1} Q_{\mu}}{o_p(1)}
\end{pmatrix}$$

(5.70)

The result (5.70), together with Lemma 5.12, gives the limiting distribution of $n(\tilde{\rho} - 1), n^{1/2} \tilde{\mu}$ and $n^{1/2}(\tilde{\theta} - \theta^0)$. The limiting distribution of $n(\tilde{\rho} - 1), n^{1/2} \tilde{\mu}, n^{1/2}(\tilde{\theta} - \theta^0)$ and $n^{1/2}(\tilde{\theta} - 1), n^{1/2} \tilde{\mu}, n^{1/2}(\tilde{\theta} - \theta^0)$ can be obtained by the same method used in the proof of Theorem 4.4. The independence of the two limiting distributions is a consequence of Theorem 2.2 of Chan and Wei (1988).
5.5. Limiting Distribution of Regression t–statistics

In this section, we define the regression t–statistics and investigate their limiting distributions. To define the test statistics we first define the covariance matrix of the limiting distribution of the estimators. Consider $e_t(Y; \theta, \rho, \mu)$ defined in (5.10). Let

$$W_{\rho t}(Y; \theta, \rho, \mu) = \frac{\partial}{\partial \rho} e_t(Y; \theta, \rho, \mu),$$

$$W_{\mu t}(Y; \theta, \rho, \mu) = \frac{\partial}{\partial \mu} e_t(Y; \theta, \rho, \mu),$$

and

$$W_{\theta i t}(Y; \theta, \rho, \mu) = \frac{\partial}{\partial \theta_i} e_t(Y; \theta, \rho, \mu), i = 1, 2, \ldots, p+q. \quad (5.71)$$

Let $W_{\theta t}(Y; \theta, \rho, \mu)$ be the (p+q)–dimensional column vector with i–th element $W_{\theta i t}(Y; \theta, \rho, \mu)$. Also define

$$V(Y; \theta, \rho, \mu) =$$

$$\begin{pmatrix}
\sum_{t=1}^{n} W_{\rho t}^2(Y; \theta, \rho, \mu) & \sum_{t=1}^{n} W_{\rho t}(Y; \theta, \rho, \mu) W_{\mu t}(Y; \theta, \rho, \mu) & \sum_{t=1}^{n} W_{\rho t}(Y; \theta, \rho, \mu) W_{\theta t}(Y; \theta, \rho, \mu) \\
\sum_{t=1}^{n} W_{\mu t}^2(Y; \theta, \rho, \mu) & \sum_{t=1}^{n} W_{\mu t}(Y; \theta, \rho, \mu) W_{\theta t}(Y; \theta, \rho, \mu) \\
\sum_{t=1}^{n} W_{\theta t}^2(Y; \theta, \rho, \mu) & \sum_{t=1}^{n} W_{\theta t}(Y; \theta, \rho, \mu) W_{\theta t}(Y; \theta, \rho, \mu)
\end{pmatrix}
$$

(5.72)
The covariance matrix is estimated by replacing \((\theta, \rho, \mu)\) of (5.72) with its estimator. The regression t-statistics are defined as

\[
\tilde{\tau} = (\rho - 1)(c_L \sigma^2)^{-1/2},
\]

\[
\bar{\tau} = (\rho - 1)(c_O \sigma^2)^{-1/2},
\]

\[
\hat{\tau} = (\rho - 1)(c_M \sigma^2)^{-1/2},
\]

where \(c_L, c_O, c_M\) are be the upper left elements of \(V^{-1}(Y;\tilde{\theta},\tilde{\rho},\tilde{\mu}), V^{-1}(Y;\theta,\rho,\mu),\) and \(V^{-1}(Y;\tilde{\theta},\tilde{\rho},\tilde{\mu})\) respectively, and \(V^{-1}(Y;\theta,\rho,\mu)\) is defined in (5.72).

Next we define t-statistics of \(\theta_i\). Thinking the test \(H_0 : \theta_i = \theta_i^0\), we define

\[
\tilde{t}_i = (\theta_i - \theta_i^0)(d_{L,i} \sigma^2)^{-1/2},
\]

\[
\bar{t}_i = (\theta_i - \theta_i^0)(d_{O,i} \sigma^2)^{-1/2},
\]

\[
\hat{t}_i = (\theta_i - \theta_i^0)(d_{M,i} \sigma^2)^{-1/2}, \quad i = 1, ..., p+q,
\]

where \(d_{L,i}, d_{O,i}, d_{M,i}\) are the \((i+2, i+2)\) elements of matrices \(V^{-1}(Y;\tilde{\theta},\tilde{\rho},\tilde{\mu}), V^{-1}(Y;\theta,\rho,\mu),\) and \(V^{-1}(Y;\tilde{\theta},\tilde{\rho},\tilde{\mu})\), respectively.
In Theorem 5.14, we establish the limiting distribution of $\tilde{\tau}$, $\tilde{\tau}$, $\tau$, $\tilde{t}_i$, $\tilde{t}_i$, and $\hat{t}_i$, $i = 1, \ldots, p+q$.

Theorem 5.14. Let model (5.1) and Assumption 5.1 hold. Then

$$\tilde{\tau} \Rightarrow 2^{-1}\{W^2(1) - 1 - W(1)\int_0^1 W(r)dr\}/[\int_0^1 W(r) - \int_0^1 W(t)dt]^2]^{1/2} \quad (5.75)$$

and $\tilde{t}_i \Rightarrow N(0,1)$, $i = 1, \ldots, p+q$. The limiting distributions of $\tilde{\tau}$ and $\{\tilde{t}_i, i = 1, \ldots, p+q\}$ are independent. The limiting distribution of $(\tilde{\tau}, \tilde{t}_i, i = 1, \ldots, p+q)$ and $(\tilde{\tau}, \hat{t}_i, i = 1, \ldots, p+q)$ is the same as that of $(\tilde{\tau}, \hat{t}_i, i = 1, \ldots, p+q)$.

Proof. Similar to the proof of Theorem 4.5. \qed
6. SIMULATION STUDY

A Monte Carlo experiment was run to investigate the empirical size of the test of the null hypothesis $H_0 : \rho = 1$. The test was based on the three estimators: the least squares estimator, the ordinary least squares estimator, and the maximum likelihood estimator. Also the empirical sizes of the unit root test based on the regression statistics and the empirical sizes of the test of a hypothesis about the moving average parameter are investigated. Data are generated from the model

\[ y_t = y_{t-1} + e_t + \beta e_{t-1}, \quad t = 1, 2, \ldots, n, \quad (6.1) \]

where $y_0 = 0$, $\{e_t\}_{t=0}^n$ is an iid $N(0,1)$ sequence. We considered $n = 100, 500$ and $\beta = -0.8, -0.5, 0, 0.5, 0.8$. For each of the $2 \times 5$ combinations of $(n, \beta)$, one thousand samples of $\{e_t\}_{t=0}^n$ were generated by RNOA in IMSL package. We used 14367 as a seed. The same $(n+1)$ values of $e_t$'s are used for all $\beta = -0.8, -0.5, 0, 0.5, 0.8$.

\[ Y_t = \rho y_{t-1} + e_t + \beta e_{t-1}, \quad (6.2) \]

and

\[ y_t = \mu + \rho \delta y_{t-1} + e_t + \beta e_{t-1}, \quad (6.3) \]

where $(\rho^*, \beta^*)$ is the least squares estimator, the ordinary least squares estimator, or the maximum likelihood estimator in the model without intercept and $(\rho^{**}, \mu^{**}, \beta^{**})$ is the least squares estimator, the ordinary least squares estimator, or the maximum likelihood estimator in the model with intercept. Minimizations are
performed by BCONG in IMSL package. In the minimization steps, we restricted \( \rho \) to \([-2, 2]\), \( \beta \) to \([-0.999, 0.999]\), and \( \mu \) to \([-2, 2]\). The initial values for the minimization were set to the true values.

6.1. Unit Root Test in a Model without Intercept

We consider the test of the null hypothesis that the data are generated by model (6.1) against the alternative hypothesis that the data are generated by the stationary model

\[
y_t = \rho y_{t-1} + e_t + \beta e_{t-1}, \quad t = 1, 2, ..., n.
\]

For each combination of \((n, \beta)\), we fitted model (6.2) one thousand times. The test is based on the estimator \((\rho^*, \beta^*)\) in the fitted model (6.2). For each fitted model the unit root test \( H_0: \rho = 1 \) against \( H_a: \rho < 1 \) with size \( \alpha = 0.05 \) and \( \alpha = 0.01 \) is done. The critical values of the test can be obtained from Fuller (1976, p.371 and p. 373). The test statistics are \( n(\rho^* - 1) \) and the corresponding t–statistics. The t–statistic is \( (\rho^* - 1) \) divided by estimated standard error of \( \rho^* \) and is defined in (4.39). It should be mentioned that, in the calculation of (4.39), \( \sigma^2 = Q_n(\tilde{\beta}, \rho)/(n-2) \), \( \sigma^2 = S_n(\tilde{\beta}, \rho)/(n-2) \), and \( \tilde{\sigma}^2 = Q_n(\hat{\beta}, \rho)/(n-2) \) are used instead of \( Q_n(\tilde{\beta}, \rho)/n \), \( S_n(\tilde{\beta}, \rho)/n \), and \( Q_n(\hat{\beta}, \rho)/n \) respectively.

Table 6.1 and Table 6.2 contain the fractions of estimated values of \( n(\rho^* - 1) \) less than the \( \alpha = 0.05 \) and \( \alpha = 0.01 \) critical values. The critical values obtained from Fuller (1976) are \(-7.9\), \(-8.0\), \(-13.3\), and \(-13.7\) for \((n, \alpha) = (100, 0.05)\), \((500,\)
0.05), (100, 0.01), and (500, 0.01) respectively. We can see that for size \( \alpha = 0.05 \) the empirical sizes of the tests agree well with the theoretical values 0.05 for values of \( \beta = -0.5, 0, 0.5, \) and 0.8. However, for \( \alpha = 0.01 \) and \( n = 100 \), the empirical sizes depend on the values of \( \beta \). When \( \beta = 0 \), the empirical sizes (0.012, 0.012, 0.012) agree with the theoretical value of 0.01 quite well. As \( \beta \) moves away from 0, the empirical sizes deviate from 0.01 moderately. When \( n = 500 \), the empirical sizes are close to the theoretical value of 0.01 for all values of \( \beta \). In the tables we do not see any difference among the empirical sizes of the three estimators, except for \( \beta = -0.8 \) and \( n = 100 \). For \( \beta = -0.8 \) and \( n = 100 \), it seems that the least squares estimator has better size than the ordinary least squares estimator and the maximum likelihood estimator. In general, we can say that \( n(\rho^* - 1) \) converges in distribution to its limit quite well in the sense of size.

Table 6.3 and Table 6.4 contain the fractions of the regression t-statistics in the fitted model (6.2) less than the critical values for size \( \alpha = 0.05 \) and \( \alpha = 0.01 \). The critical values of the regression t-statistics are \(-1.95, -1.95, -2.60, \) and \(-2.58 \) for \((n, \alpha) = (100, 0.05), (500, 0.05), (100, 0.01), \) and \((500, 0.01) \) respectively. For \( n = 100 \), we observe that the t-statistics for \( \beta = -0.8 \) have better better sizes than the t-statistics for other \( \beta \)'s. For \( n = 100 \) and \( \alpha = 0.05 \), and for \( \beta = 0.8 \) the sizes of the t-statistics differ from 0.05. For \( n = 500 \), the t-statistics for \( \beta = 0 \) have better sizes than the t-statistics for other \( \beta \)'s. For \( n = 500 \) and \( \alpha = 0.05 \), the sizes for \( \beta = -0.8 \) and \( \beta = 0.8 \) seems to be far from 0.05. For \( n = 500 \) and \( \alpha = 0.01 \), the t-statistics seem to have empirical sizes close to 0.01. Except for \( n = 100 \) and \( \beta = -0.8 \), the three t-statistics have similar empirical sizes. It seems that the least
squares estimator has better size than the ordinary least squares estimator and the
maximum likelihood estimator for \( n = 100 \) and \( \beta = -0.8 \). Comparing Table 6.3
with Table 6.1 and Table 6.4 with Table 6.2, we see that \( \rho^* \) has better empirical
sizes than that of the \( t \)-statistics except for \( n = 100 \) and \( \beta = -0.8 \).

We conclude that tests based on \( n(\rho^* - 1) \) have quite good empirical sizes for
all value of \( n \) and \( \beta \) considered. The test based on the regression \( t \)-statistics seems
to have reasonable empirical sizes but does not have as good empirical sizes as \( n(\rho^* - 1) \), especially when \( \beta \neq 0 \) and \( n = 500 \). Also we can say that the least squares
estimator offers better size than the ordinary least squares estimator and the
maximum likelihood estimator for \( \beta \) close to \(-1\) and \( n \) small.

6.2. Unit Root Test in a Model with Intercept

We consider the test of the null hypothesis that the data are generated by
model (6.5) with \( \rho = 1 \) against the alternative hypothesis that the data are
generated by model (6.5) with \( \rho < 1 \), where model (6.5) is

\[
y_t = \mu + \rho y_{t-1} + e_t + \beta e_{t-1}, \quad t = 1, 2, \ldots, n.
\]  

(6.5)

Schwert(1989) performed a comprehensive Monte Carlo study for model
(6.5). He considered the procedures proposed by Dickey and Fuller (1979), Phillips
(1987), Said and Dickey (1984, 1985). He calculated empirical sizes of tests of \( H_0: \rho = 1 \) with \( \alpha = 0.01 \) and \( \alpha = 0.05 \).

For each of the \( 2 \times 5 \) combinations of \( (n, \beta) \), we fitted model (6.3) 1000
times. The test is based on the estimator \( (\rho^{**}, \mu^{**}, \beta^{**}) \) in the fitted model (6.3).
For each fitted model the unit root test $H_0 : \rho = 1$ against $H_a : \rho < 1$ with size $\alpha = 0.05$ and $\alpha = 0.01$ is done. The test statistics are $n(\rho^{**} - 1)$ and the corresponding t-statistics. The t-statistics are $(\rho^{**} - 1)$ divided by the estimated standard error of $\rho^{**}$ and are defined in (5.73). It should be mentioned that, in the calculation of (5.73), $\bar{\sigma}^2 = Q_n(\tilde{\beta},\tilde{\rho},\tilde{\mu})/(n-3)$, $\bar{\sigma}^2 = S_n(\beta,\rho,\mu)/(n-3)$, and $\hat{\sigma}^2 = Q_n(\hat{\beta},\hat{\rho},\hat{\mu})/(n-3)$ are used instead of $Q_n(\bar{\beta},\bar{\rho},\bar{\mu})/n$, $S_n(\bar{\beta},\bar{\rho},\bar{\mu})/n$, and $Q_n(\bar{\beta},\bar{\rho},\bar{\mu})/n$ respectively.

Table 6.5 and Table 6.6 contain the fractions of estimated values of $n(\rho^{**} - 1)$ less than the critical values for sizes $\alpha = 0.05$ and $\alpha = 0.01$. The critical values obtained from Fuller (1976) are $-13.7$, $-14.0$, $-19.8$, and $-20.5$ for $(n, \alpha) = (100, 0.05)$, $(500, 0.05)$, $(100, 0.01)$, and $(500, 0.01)$ respectively. First consider the case of $n = 100$. When $\beta = -0.8$, the empirical sizes of the test based on $n(\rho^{**} - 1)$ differ from the theoretical values. The least squares estimator shows less deviation from the theoretical size than does the ordinary least squares estimator and the maximum likelihood estimator. As $\beta$ increases to $-0.5$, the empirical sizes improve to $(0.078, 0.084, 0.088)$ and $(0.028, 0.029, 0.028)$. For $\beta = 0, 0.5, 0.8$, the empirical sizes are reasonable. For $n = 500$, the empirical sizes for $\beta = -0.5, 0, 0.5$, and $0.05$ are reasonable. For $\beta = -0.8$, the empirical sizes are much better at $n = 500$ than at $n = 100$.

Table 6.7 and Table 6.8 contain the fractions of the regression t-statistics less than the critical values for sizes $\alpha = 0.05$ and $\alpha = 0.01$. The critical values obtained from Fuller (1976) are $-2.89$, $-2.87$, $-3.51$, and $-3.44$ for $(n, \alpha) = (100, 0.05)$, $(500, 0.05)$, $(100, 0.01)$, and $(500, 0.01)$ respectively. When $n = 100$, the regression t-statistics have poor empirical sizes both for $\alpha = 0.05$ and $\alpha = 0.01$. 
When \( n = 500 \) and \( \beta = 0, 0.5, 0.8 \), the regression t-statistics have empirical sizes close to the theoretical values both for \( \alpha = 0.05 \) and \( \alpha = 0.01 \). When \( n = 500 \) and \( \beta = -0.8, -0.5 \), the empirical sizes are a little bit away from the theoretical values.

We conclude that the tests based on \( n(\rho^{**} - 1) \) have reasonable sizes when \( n \) is as large as 500. When \( n = 100 \), the \( n(\rho^{**} - 1) \) tests also have reasonable sizes except for \( \beta = -0.8 \). It seems that \( \rho^{**} \) has better size than the regression t-statistic. There is no difference in the sizes of the three type estimators except for \( \beta = -0.8 \) and \( n = 100 \). When \( \beta = -0.8 \) and \( n = 100 \), the least squares estimator has better size than the ordinary least squares estimator and the maximum likelihood estimator.

### 6.3. A Test for the Moving Average Parameter

In this section we investigate the empirical sizes of the one sided test \( H_0 : \beta = \beta^0 \) against \( H_a : \beta < \beta^0 \). Table 6.9 – Table 6.12 contain the fractions of the t-statistics for \( \beta \) less than the critical values for size \( \alpha = 0.05 \) and \( \alpha = 0.01 \). In the fitted model (6.2) and (6.3), t-values for \( \beta^* \) and \( \beta^{**} \) defined in the equation between (4.39) and (4.40) and equation (5.74) with modified \( \sigma^2 \), \( \sigma^2 \), and \( \sigma^2 \) as in section 6.1 and section 6.2 are calculated. The critical values for tests of size \( \alpha = 0.05 \) and \( \alpha = 0.01 \) are obtained from the standard normal distribution and are \(-1.645 \) and \(-2.326 \) respectively. The results in Table 6.9 and Table 6.10 for the model without intercept and the results in Table 6.11 and Table 6.12 for the model with intercept show similar patterns. Therefore statements from this point on are relevant to both set of tables. For \( \beta = 0 \), the empirical sizes are close to the
corresponding theoretical sizes. For $\beta$ not equal to 0, the empirical sizes differ considerably from the theoretical values. As $n$ increases from 100 to 500, the gap between empirical sizes and theoretical sizes decreases. It is interesting to observe that the ordinary least squares estimator has better empirical sizes than the least squares estimator and the maximum likelihood estimator, especially for $n = 100$ and $\beta = -0.8$, $-0.5$ and $n = 500$ and $\beta = -0.8$. Recall that the least squares tests for $\rho$ has better empirical size than the other two estimators for $\beta = -0.8$ and $n$ small.
Table 6.1. Fraction of $n(p^* - 1)$ in fitted model (6.2) without intercept less than the 0.05 critical value (1000 samples)

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 500</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
</tr>
<tr>
<td>-0.8</td>
<td>0.056</td>
<td>0.068</td>
<td>0.081</td>
<td>0.044</td>
<td>0.043</td>
<td>0.046</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.049</td>
<td>0.056</td>
<td>0.057</td>
<td>0.050</td>
<td>0.050</td>
<td>0.050</td>
</tr>
<tr>
<td>0.0</td>
<td>0.049</td>
<td>0.049</td>
<td>0.049</td>
<td>0.055</td>
<td>0.055</td>
<td>0.054</td>
</tr>
<tr>
<td>0.5</td>
<td>0.045</td>
<td>0.045</td>
<td>0.045</td>
<td>0.056</td>
<td>0.056</td>
<td>0.057</td>
</tr>
<tr>
<td>0.8</td>
<td>0.045</td>
<td>0.043</td>
<td>0.042</td>
<td>0.057</td>
<td>0.057</td>
<td>0.057</td>
</tr>
</tbody>
</table>

Table 6.2. Fraction of $n(p^* - 1)$ in fitted model (6.2) without intercept less than the 0.01 critical value (1000 samples)

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 500</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
</tr>
<tr>
<td>-0.8</td>
<td>0.028</td>
<td>0.037</td>
<td>0.036</td>
<td>0.012</td>
<td>0.012</td>
<td>0.012</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.017</td>
<td>0.017</td>
<td>0.017</td>
<td>0.011</td>
<td>0.012</td>
<td>0.012</td>
</tr>
<tr>
<td>0.0</td>
<td>0.012</td>
<td>0.012</td>
<td>0.012</td>
<td>0.011</td>
<td>0.011</td>
<td>0.012</td>
</tr>
<tr>
<td>0.5</td>
<td>0.007</td>
<td>0.007</td>
<td>0.007</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td>0.8</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.012</td>
<td>0.012</td>
<td>0.012</td>
</tr>
</tbody>
</table>
Table 6.3. Fraction of regression $t-$statistics in fitted model (6.2) without intercept less than the 0.05 critical value (1000 samples)

<table>
<thead>
<tr>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 500</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
<td>LSE</td>
</tr>
<tr>
<td>$\beta = -0.8$</td>
<td>0.051</td>
<td>0.053</td>
<td>0.060</td>
<td>0.038</td>
</tr>
<tr>
<td>$\beta = -0.5$</td>
<td>0.041</td>
<td>0.043</td>
<td>0.043</td>
<td>0.043</td>
</tr>
<tr>
<td>$\beta = +0.0$</td>
<td>0.040</td>
<td>0.040</td>
<td>0.040</td>
<td>0.053</td>
</tr>
<tr>
<td>$\beta = +0.5$</td>
<td>0.043</td>
<td>0.041</td>
<td>0.043</td>
<td>0.059</td>
</tr>
<tr>
<td>$\beta = +0.8$</td>
<td>0.038</td>
<td>0.037</td>
<td>0.038</td>
<td>0.061</td>
</tr>
</tbody>
</table>

Table 6.4. Fraction of regression $t-$statistics in fitted model (6.2) without intercept less than the 0.01 critical value (1000 samples)

<table>
<thead>
<tr>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 500</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
<td>LSE</td>
</tr>
<tr>
<td>$\beta = -0.8$</td>
<td>0.012</td>
<td>0.021</td>
<td>0.023</td>
<td>0.009</td>
</tr>
<tr>
<td>$\beta = -0.5$</td>
<td>0.009</td>
<td>0.010</td>
<td>0.010</td>
<td>0.011</td>
</tr>
<tr>
<td>$\beta = +0.0$</td>
<td>0.008</td>
<td>0.008</td>
<td>0.008</td>
<td>0.011</td>
</tr>
<tr>
<td>$\beta = +0.5$</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.010</td>
</tr>
<tr>
<td>$\beta = +0.8$</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.011</td>
</tr>
</tbody>
</table>
Table 6.5. Fraction of $n(\rho^{**} - 1)$ in fitted model (6.3) with intercept less than the 0.05 critical value (1000 samples)

<table>
<thead>
<tr>
<th></th>
<th>n = 100</th>
<th></th>
<th>n = 500</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
</tr>
<tr>
<td>$\beta = -0.8$</td>
<td>0.154</td>
<td>0.217</td>
<td>0.206</td>
<td>0.064</td>
<td>0.076</td>
<td>0.069</td>
</tr>
<tr>
<td>$\beta = -0.5$</td>
<td>0.078</td>
<td>0.084</td>
<td>0.088</td>
<td>0.052</td>
<td>0.054</td>
<td>0.052</td>
</tr>
<tr>
<td>$\beta = +0.0$</td>
<td>0.056</td>
<td>0.057</td>
<td>0.057</td>
<td>0.048</td>
<td>0.048</td>
<td>0.048</td>
</tr>
<tr>
<td>$\beta = +0.5$</td>
<td>0.043</td>
<td>0.045</td>
<td>0.039</td>
<td>0.047</td>
<td>0.048</td>
<td>0.047</td>
</tr>
<tr>
<td>$\beta = +0.8$</td>
<td>0.045</td>
<td>0.043</td>
<td>0.043</td>
<td>0.047</td>
<td>0.042</td>
<td>0.046</td>
</tr>
</tbody>
</table>

Table 6.6. Fraction of $n(\rho^{**} - 1)$ in fitted model (6.3) with intercept less than the 0.01 critical value (1000 samples)

<table>
<thead>
<tr>
<th></th>
<th>n = 100</th>
<th></th>
<th>n = 500</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
</tr>
<tr>
<td>$\beta = -0.8$</td>
<td>0.080</td>
<td>0.143</td>
<td>0.122</td>
<td>0.017</td>
<td>0.021</td>
<td>0.017</td>
</tr>
<tr>
<td>$\beta = -0.5$</td>
<td>0.028</td>
<td>0.029</td>
<td>0.028</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td>$\beta = +0.0$</td>
<td>0.013</td>
<td>0.013</td>
<td>0.012</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>$\beta = +0.5$</td>
<td>0.008</td>
<td>0.008</td>
<td>0.008</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>$\beta = +0.8$</td>
<td>0.008</td>
<td>0.006</td>
<td>0.007</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
</tr>
</tbody>
</table>
Table 6.7  Fraction of regression $t$-statistics for $\beta$ in fitted model (6.3) with intercept less than the 0.05 critical value (1000 samples)

<table>
<thead>
<tr>
<th></th>
<th>n = 100</th>
<th></th>
<th>n = 500</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
<td>LSE</td>
</tr>
<tr>
<td>$\beta = -0.8$</td>
<td>0.053</td>
<td>0.046</td>
<td>0.078</td>
<td>0.019</td>
</tr>
<tr>
<td>$\beta = -0.5$</td>
<td>0.010</td>
<td>0.009</td>
<td>0.011</td>
<td>0.039</td>
</tr>
<tr>
<td>$\beta = +0.0$</td>
<td>0.025</td>
<td>0.025</td>
<td>0.024</td>
<td>0.051</td>
</tr>
<tr>
<td>$\beta = +0.5$</td>
<td>0.039</td>
<td>0.036</td>
<td>0.038</td>
<td>0.051</td>
</tr>
<tr>
<td>$\beta = +0.8$</td>
<td>0.057</td>
<td>0.054</td>
<td>0.055</td>
<td>0.054</td>
</tr>
</tbody>
</table>

Table 6.8. Fraction of regression $t$-statistics for $\beta$ in fitted model (6.3) with intercept less than the 0.01 critical value (1000 samples)

<table>
<thead>
<tr>
<th></th>
<th>n = 100</th>
<th></th>
<th>n = 500</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
<td>LSE</td>
</tr>
<tr>
<td>$\beta = -0.8$</td>
<td>0.033</td>
<td>0.036</td>
<td>0.052</td>
<td>0.006</td>
</tr>
<tr>
<td>$\beta = -0.5$</td>
<td>0.002</td>
<td>0.002</td>
<td>0.001</td>
<td>0.008</td>
</tr>
<tr>
<td>$\beta = +0.0$</td>
<td>0.001</td>
<td>0.000</td>
<td>0.001</td>
<td>0.009</td>
</tr>
<tr>
<td>$\beta = +0.5$</td>
<td>0.006</td>
<td>0.005</td>
<td>0.006</td>
<td>0.008</td>
</tr>
<tr>
<td>$\beta = +0.8$</td>
<td>0.013</td>
<td>0.013</td>
<td>0.012</td>
<td>0.010</td>
</tr>
</tbody>
</table>
Table 6.9. Fraction of t-statistics for $\beta$ in fitted model (6.2) without intercept of $\beta$
less than the 0.05 critical value (1000 samples)

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 500</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
</tr>
<tr>
<td>$\beta = -0.8$</td>
<td>0.221</td>
<td>0.112</td>
<td>0.167</td>
<td>0.091</td>
<td>0.063</td>
<td>0.076</td>
</tr>
<tr>
<td>$\beta = -0.5$</td>
<td>0.094</td>
<td>0.065</td>
<td>0.073</td>
<td>0.069</td>
<td>0.062</td>
<td>0.066</td>
</tr>
<tr>
<td>$\beta = +0.0$</td>
<td>0.061</td>
<td>0.057</td>
<td>0.058</td>
<td>0.057</td>
<td>0.057</td>
<td>0.057</td>
</tr>
<tr>
<td>$\beta = +0.5$</td>
<td>0.036</td>
<td>0.039</td>
<td>0.036</td>
<td>0.041</td>
<td>0.044</td>
<td>0.043</td>
</tr>
<tr>
<td>$\beta = +0.8$</td>
<td>0.018</td>
<td>0.033</td>
<td>0.024</td>
<td>0.022</td>
<td>0.030</td>
<td>0.025</td>
</tr>
</tbody>
</table>

Table 6.10. Fraction of t-statistics for $\beta$ in fitted model (6.2) without intercept
less than the 0.01 critical value (1000 samples)

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 500</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
<td>LSE</td>
<td>OLS</td>
<td>MLE</td>
</tr>
<tr>
<td>$\beta = -0.8$</td>
<td>0.189</td>
<td>0.070</td>
<td>0.134</td>
<td>0.037</td>
<td>0.022</td>
<td>0.030</td>
</tr>
<tr>
<td>$\beta = -0.5$</td>
<td>0.041</td>
<td>0.026</td>
<td>0.032</td>
<td>0.021</td>
<td>0.022</td>
<td>0.018</td>
</tr>
<tr>
<td>$\beta = +0.0$</td>
<td>0.009</td>
<td>0.009</td>
<td>0.008</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
</tr>
<tr>
<td>$\beta = +0.5$</td>
<td>0.006</td>
<td>0.006</td>
<td>0.007</td>
<td>0.012</td>
<td>0.013</td>
<td>0.010</td>
</tr>
<tr>
<td>$\beta = +0.8$</td>
<td>0.003</td>
<td>0.007</td>
<td>0.003</td>
<td>0.002</td>
<td>0.004</td>
<td>0.002</td>
</tr>
</tbody>
</table>
Table 6.11. Fraction of t-statistics for $\beta$ in fitted model (6.3) with intercept less than the 0.05 critical value

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$n = 100$</th>
<th>$n = 500$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
</tr>
<tr>
<td>$\beta = -0.8$</td>
<td>0.251</td>
<td>0.159</td>
</tr>
<tr>
<td>$\beta = -0.5$</td>
<td>0.085</td>
<td>0.059</td>
</tr>
<tr>
<td>$\beta = +0.0$</td>
<td>0.052</td>
<td>0.050</td>
</tr>
<tr>
<td>$\beta = +0.5$</td>
<td>0.030</td>
<td>0.031</td>
</tr>
<tr>
<td>$\beta = +0.8$</td>
<td>0.015</td>
<td>0.027</td>
</tr>
</tbody>
</table>

Table 6.12. Fraction of t-statistics for $\beta$ in fitted model (6.3) with intercept less than the 0.01 critical value (1000 samples)

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$n = 100$</th>
<th>$n = 500$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LSE</td>
<td>OLS</td>
</tr>
<tr>
<td>$\beta = -0.8$</td>
<td>0.223</td>
<td>0.139</td>
</tr>
<tr>
<td>$\beta = -0.5$</td>
<td>0.037</td>
<td>0.025</td>
</tr>
<tr>
<td>$\beta = +0.0$</td>
<td>0.008</td>
<td>0.008</td>
</tr>
<tr>
<td>$\beta = +0.5$</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td>$\beta = +0.8$</td>
<td>0.002</td>
<td>0.006</td>
</tr>
</tbody>
</table>
6.4. Computational Aspect

In this section, we describe a computational aspect obtaining $Q_n(\beta, \rho)$, $S_n(\beta, \rho)$, $Q_n(\beta, \rho, \mu)$, $S_n(\beta, \rho, \mu)$ and derivatives of $Q_n(\beta, \rho)$, $S_n(\beta, \rho)$, $Q_n(\beta, \rho, \mu)$, $S_n(\beta, \rho, \mu)$ for the simulation model (6.4) and model (6.5).

First consider model (6.4). Observe that $e_t(Y; \beta, \rho)$ in (1.18) can be obtained recursively. The recursion is

$$e_t(Y; \beta, \rho) = y_t - \rho y_{t-1} - \beta e_{t-1}(Y; \beta, \rho), \quad t = 1, \ldots, n, \tag{6.6}$$

where $e_0(Y; \beta, \rho) = 0$. Therefore

$$S_n(\beta, \rho) = \sum_{t=1}^{n} e_t^2(Y; \beta, \rho)$$

can be obtained from (6.6). For $Q_n(\beta, \rho)$ we need to calculate

$$R_n(\beta, \rho) = (Y - \rho Y_1)'D_n^{-1}(-A^{-1} + M_n' M_n)^{-1}M_n' D_n(Y - \rho Y_1). \tag{6.7}$$

For model (6.4) and (6.5) we have, by (2.1),

$$M_n = (-\beta, (-\beta)^2, \ldots, (-\beta)^n)'$$
and $A = 1$.

Therefore

$$(A^{-1} + M_n' M_n) = (1 - \beta^{2n+2})/(1 - \beta^2) \tag{6.8}$$
and
\[ R_n(\beta, \rho) = \left\{ \sum_{t=1}^{n} (-\beta)^t e_t(Y; \beta, \rho) \right\}^2 (1 - \beta^2)/(1 - \beta^{2n+2}). \] (6.8)

Therefore we can obtain
\[ Q_n(\beta, \rho) = S_n(\beta, \rho) - R_n(\beta, \rho). \]

The determinant of \( \Gamma_n(\beta) \) is
\[ (1 - \beta^{2n+2})/(1 - \beta^2). \]

For the derivatives of \( Q_n(\beta, \rho), S_n(\beta, \rho), \) and \( V_n(\beta, \rho) \) in (4.39), we need to calculate the derivatives \( W_{\rho t}(Y; \beta, \rho) \) and \( W_{\beta t}(Y; \beta, \rho) \) of \( e_t(Y; \beta, \rho) \). From (6.6), \( W_{\rho t}(Y; \beta, \rho) \) and \( W_{\beta t}(Y; \beta, \rho) \) can be obtained recursively
\[ W_{\rho t}(Y; \beta, \rho) = -y_{t-1} - \beta W_{\rho, t-1}(Y; \beta, \rho) \]
and
\[ W_{\beta t}(Y; \beta, \rho) = -e_{t-1} - \beta W_{\beta, t-1}(Y; \beta, \rho), \] (6.9)

where \( W_{\rho 0}(Y; \beta, \rho) = W_{\beta 0}(Y; \beta, \rho) = 0 \). From (6.9) we can calculate derivatives of \( S_n(\beta, \rho) \) and \( V(Y; \beta, \rho) \). The derivatives of \( S_n(\beta, \rho) \) are
\[ \frac{\partial S_n(\beta, \rho)}{\partial \rho} = 2 \sum_{t=1}^{n} W_{\rho t}(Y; \beta, \rho)e_t(Y; \beta, \rho) \]
and
\[
\frac{\partial S_n(\beta, \rho)}{\partial \beta} = 2 \sum_{t=1}^{n} W_{\beta t}(Y; \beta, \rho)e_t(Y; \beta, \rho).
\] (6.10)

Derivatives of term (6.8) can be obtained by calculating derivatives of \(U_t(Y; \beta, \rho) = (-\beta)^t e_t(Y; \beta, \rho)\) recursively.

\[
\frac{\partial U_t(Y; \rho, \beta)}{\partial \rho} = (-\beta)^t W_{\rho t}(Y; \beta, \rho)
\]

and

\[
\frac{\partial U_t(Y; \rho, \beta)}{\partial \beta} = -t(-\beta)^{t-1} e_t(Y; \beta, \rho) + (-\beta)^t W_{\rho t}(Y; \beta, \rho).
\] (6.11)

From (6.11) and derivatives (6.10) of \(S_n(\beta, \rho)\) we can calculate the derivatives

\[
\frac{\partial Q_n(\beta, \rho)}{\partial \rho} = \frac{\partial S_n(\beta, \rho)}{\partial \rho} - \frac{\partial R_n(\beta, \rho)}{\partial \rho}
\]

and

\[
\frac{\partial Q_n(\beta, \rho)}{\partial \beta} = \frac{\partial S_n(\beta, \rho)}{\partial \beta} - \frac{\partial R_n(\beta, \rho)}{\partial \beta}
\]

of \(Q_n(\beta, \rho)\).
7. APPENDIX A

PRELIMINARY PROBABILITY THEORY

We summarize several probability laws which are used in this research. Let \( \{e_n\} \) be a sequence of iid (0,1) random variables. Define

\[
W_i = e_1 + e_2 + \cdots + e_i, \quad i = 1, 2, \ldots \quad W_0 = 0.
\]  

(7.1)

In Theorem A.1 — A.3, the well known strong law of large numbers, central limit theorem, and law of the iterated logarithm are given.

Theorem A.1. (The strong law of large number) Let \( \{W_n\} \) be defined in (7.1), then

\[
l \lim_{n \to \infty} W_n/n \to 1 \quad \text{a.s.}
\]

Theorem A.2. (The central limit theorem) Let \( \{W_n\} \) be defined in (7.1), then

\[
n^{1/2} W_n \Rightarrow N(0,1).
\]

Theorem A.3. (The law of the iterated logarithm) Let \( \{W_n\} \) be defined in (7.1), then

\[
l \limsup_{n \to \infty} (2n \log \log n)^{-1/2} W_n = 1 \quad \text{a.s.,}
\]
and

\[ \liminf (2n \log \log n)^{-1/2} W_n = -1 \text{ a.s.} \quad \text{as} \quad n \to \infty \]

The following theorem can be found in Chung (1968, p. 122, exercise 7).

**Theorem A.4.** For arbitrary sequence of random variable \( \{X_n\} \), if

\[ \sum_{n=1}^{\infty} \mathbb{E}|X_n| < \infty \]

then \( \sum_{n=1}^{\infty} X_n \) converges a.s.

**Example A.5.** Let the process \( \{z_t\} \) be defined by the model (1.1) with assumption 1.1. Let \( \{a_n\} \) be a absolutely summable sequence, that is \( \sum_{n=1}^{\infty} |a_n| < \infty \). Then

\[ \sum_{n=1}^{\infty} a_n z_n \] converges a.s.

**Proof.** This follows from the stationarity of the process \( \{z_t\} \) and Theorem A.4. \( \square \)

Now we give the martingale convergence theorem (see Ash 1972, p. 292).

**Definition A.6.** Let \( (\Omega, G, P) \) be a probability space, \( \{X_n\} \) be a sequence of integrable random variables on \( (\Omega, G, P) \), and \( G_1 \subset G_2 \subset \ldots \) an increasing sequence of sub \( \sigma \)-fields of \( G \), \( X_n \) is assumed \( G_n \) - measurable. The sequence \( \{X_n, G_n\} \) is said to be a martingale iff for all \( n = 1, 2, \ldots \), \( \mathbb{E}(X_{n+1} \mid G_n) = X_n \) a.e., a submartingale iff \( \mathbb{E}(X_{n+1} \mid G_n) \geq X_n \) a.e., a supermartingale iff \( \mathbb{E}(X_{n+1} \mid G_n) \leq X_n \) a.e.
Theorem A.7. ((Sub)Martingale convergence theorem) Let \( \{X_n, G_n, n = 1, 2, \ldots\} \) be a submartingale. If \( \sup_n E|X_n| < \infty \), there is an integrable random variable \( X \) such that \( X_n \to X \) a.s.

For an example, see Lemma 3.9.

Next we give the Donsker's invariance principle. Let \( C[0,1] \) be the space of all continuous functions on \( [0,1] \) with uniform metric \( |\cdot| \) defined by

\[
|x| = \sup_{1 \leq t \leq 1} |x(t)|.
\]  

(7.2)

Let \( C \) be the \( \sigma \) - algebra generated by the open sets in \( C[0,1] \). Define

\[
W_n(r) = n^{-1/2} \left( W_{i-1} + (nr - (i-1))e_i \right), \text{ if } r \in \left( (i-1)/n, i/n \right], i = 1, 2, \ldots, n
\]

(7.3)

where \( [nt] \) is the smallest integer not greater than \( nt \). Then \( W_n(\cdot) \) is the connected lines at points \( (0,0), (1/n, W_1/n^{1/2}), (2/n, W_2/n^{1/2}), \ldots, (n/n, W_n/n^{1/2}) \). Hence \( W_n(\cdot) \in C[0,1] \). The Donsker's theorem tells us that the limiting distribution of \( W_n(\cdot) \) is the standard Brownian motion on \( [0,1] \). See Billingsley (1968, p. 68).

Theorem A.8. Let \( W_n(\cdot) \) be defined in (7.3). Then we have

\[
W_n(\cdot) \Rightarrow W(\cdot)
\]

where \( W(\cdot) \) is the standard Brownian motion on \([0,1]\).
Now we state the continuous mapping theorem. See Billingsley (1968, p. 31).

**Theorem A.9.** Let $S$ and $S'$ be two metric spaces. Assume $(S, S', P)$ and $(S', S', P')$ be probability spaces and $\{X_n\}$ and $X$ be random elements defined of $(S, S, P)$. Suppose $X_n \Rightarrow X$. Let $\Psi$ be a continuous mapping of $(S, S)$ into $(S', S')$. Then

$$\Psi(X_n) \Rightarrow \Psi(X).$$

We give one application of Donsker's invariance principle and continuous mapping theorem. Consider the probability space $(C[0,1], C, P)$, where $P$ is the Wiener measure. The Wiener measure is the probability measure induced by the standard Brownian motion on $[0,1]$.

**Example A.10.** Let $\{W_i\}$ be defined in (7.1). Then we have

$$(n^{-2} \sum_{i=1}^{n} W_{i-1}^2, n^{-1} \sum_{i=1}^{n} W_{i-1}^2) \Rightarrow \left(\int_0^1 W^2(r)dr, 2^{-1}(W^2(1) - 1)\right). \quad (7.4)$$

**Proof.** Observe that

$$\int_0^1 W^2_n(r)dr = \sum_{i=1}^{n} \int_{(i-1)/n}^{i/n} n^{-1}(W_{i-1} + (nr - (i-1))e_1)^2 dr$$

$$= n^{-2} \sum_{i=1}^{n} W_{i-1}^2 + 2n^{-1} \sum_{i=1}^{n} W_{i-1}^2 \int_{(i-1)/n}^{i/n} (nr-i+1) dr.$$
\[ + n^{-1} \sum_{i=1}^{n} e_i^2 \int_{(i-1)/n}^{i/n} (nr-i+1)^2 \, dr \]

\[ = n^{-2} \sum_{i=1}^{n} W_{i-1}^2 + n^{-2} \sum_{i=1}^{n} W_{i-1} e_i + 3^{-1} n^{-2} \sum_{i=1}^{n} e_i^2 \]

and

\[ 2^{-1} (X_n^2(1) - 1) = \sum_{i=1}^{n} W_{i-1} e_i - 2^{-1} (1 - \sum_{i=1}^{n} e_i^2) \]

Hence observing that \( \sum_{i=1}^{n} W_{i-1} e_i = O_p(n), \sum_{i=1}^{n} e_i^2 = 1 + o_p(1) \), we have

\[ n^{-2} \sum_{i=1}^{n} W_{i-1}^2 = \int_0^1 X_n^2(r) \, dr + O_p(n^{-1}) \]

\[ n^{-1} \sum_{i=1}^{n} W_{i-1} e_i = 2^{-1} (X_n^2(1) - 1) + o_p(1) \]

Now consider the function \( \Psi : C[0,1] \to \mathbb{R}^2 \) defined by

\[ \Psi(X) = \left( \int_0^1 X_n^2(r) \, dr, 2^{-1} (X_n^2(1) - 1) \right) \quad X \in C[0,1] \]

We show that \( \Psi \) is continuous with respect to the supremum metric defined in (7.2).

For this note that if \( X \to 0 \), that is, \( \sup_{0 \leq t \leq 1} |X(t)| \to 0 \) then
\[ |\Psi(X) - \Psi(0)| \leq \left( \sup_{0 \leq t \leq 1} X^4(t) + X^4(1) \right)^{1/2} \rightarrow 0 \]

Therefore by the Donsker's invariance principle (Theorem A.8) and continuous mapping theorem (Theorem A.9) the result follows. \( \square \)

Now we turn to the Strassen’s law of the iterated logarithm for Brownian motion. Let

\[ U_n(r) = (2\log\log n)^{-1/2} W_{[nr]}, \quad 0 \leq r \leq 1. \]  \hspace{1cm} (7.5)

Let

\[ K = \{ x \in C[0,1] : x \text{ is absolutely continuous}, x(0) = 0, \text{ and } \int_0^1 \left( \frac{dx}{dr} \right)^2 \, dr \leq 1 \} \]

(7.6)

The following Theorem A.11 and Corollary A.11 can be found in Stout (1974 p. 284).

**Theorem A.11.** (Strassen) Let \( U_n(\cdot) \) be defined in (7.5). Then with probability one, \( \{ U_n(\cdot), n \geq 3 \} \) has the set \( K \) in (7.6) as the set of limit points.

**Corollary A.12.** Let \( \Psi : C[0,1] \rightarrow \mathbb{R} \) be continuous, then

\[ \limsup_{n \rightarrow \infty} \Psi(U_n) = \sup_{\phi \in K} \Psi(\phi) \quad \text{a.s.} \]
Example A.13. Let \( \{W_n\} \) be defined in (7.1), then

\begin{enumerate}
\item \( \limsup_{n \to \infty} \sum_{i=1}^{n} W_{i-1} e_i \sqrt{\frac{1}{2n \log \log n}} = \frac{1}{2} \),
\item \( \limsup_{n \to \infty} \sum_{i=1}^{n} W_{i-1} \sqrt{\frac{1}{2n^3 \log \log n}} \) = \(3^{-1/2}\),
\item \( \limsup_{n \to \infty} \sum_{i=1}^{n} W_{i-1}^2 \sqrt{\frac{1}{2n \log \log n}} = \frac{4}{\pi^2} \).
\end{enumerate}

Proof of i).

\[ \limsup_{n \to \infty} \sum_{i=1}^{n} W_{i-1} e_i \sqrt{\frac{1}{2n \log \log n}} = \limsup_{n \to \infty} 2^{-1}(W_n^2 - \sum_{i=1}^{n} e_i^2) \sqrt{\frac{1}{2n \log \log n}} \]

\[ = \limsup_{n \to \infty} 2^{-1}(U_n^2(1) - O_{wp}((\log n)^{-1})) \]

\[ = 2^{-1} \sup_{\phi \in K} (\int_0^1 \phi'(r) dr)^2 \leq 2^{-1} \sup_{\phi \in K} (\int_0^1 (\phi'(r))^2 dr) \leq 1/2 \] (Holder inequality).

Note that for \( \phi(r) = r \) the equality holds. Therefore the result follows.

Proof of ii).

\[ \limsup_{n \to \infty} \sum_{i=1}^{n} W_{i-1} \sqrt{\frac{1}{2n^3 \log \log n}} = \limsup_{n \to \infty} \int_0^1 U_n(r) dr \]

\[ = \left( \int_0^1 \int_r^1 dt \right)^{1/2} \] (See Stout 1974, p. 294)
\[ \int_0^1 (1-r)^2 \, dr = 3^{-1/2}. \]

Proof of iii). This is done by Strassen (1964).

Now we define the strong \( \alpha \)-mixing process (see Stout 1974, p. 212).

**Definition A.14.** Let \( \{X_i, i \geq 1\} \) be a sequence of random variables. Define \( B_{ij} = \sigma(X_k, i \leq k \leq j) \), the \( \sigma \)-algebra generated by \( X_k, i \leq k \leq j, 1 \leq i \leq j \leq \omega \). Then \( \{X_i, i \geq 1\} \) is said to be strong \( \alpha \)-mixing if there exists a function \( \alpha \) for which \( \alpha(m) \to 0 \) as \( m \to \omega \) and \( A \in B_{1n}, B \in B_{m+n, \omega} \) implies

\[ |P(A \cap B) - P(A)P(B)| \leq \alpha(m) \text{ for all } m \geq 1 \text{ and } n \geq 1. \]

In Lemma A.15, Kronecker's lemma is given (see Chung 1974, p. 123).

**Lemma A.15. (Kronecker lemma)** Let \( \{x_i\} \) be a sequence of real numbers, \( \{a_i\} \) be a sequence of real numbers \( > 0 \) and \( \uparrow \omega \), then

\[ \sum_{n=1}^{\omega} x_n/a_n < \omega \text{ implies } a_n^{-1} \sum_{j=1}^{n} x_j \to 0. \]
8. APPENDIX B

TECHNICAL DETAILS

8.1. Appendix for Chapter 2

8.1.1. Justification of (2.5).

Let \( i > j \). The \( i \)-th row, the \( j \)-th row of \( D_n^* \), and the \( j \)-th row of \( M_n \) are

\[
(0, 0, ..., 0, d_0, d_1, ..., d_{j-i}, ..., d_{n-i})
\]

\[
(0, 0, ..., 0, 0, 0, 0, 0, ..., d_{n-j})
\]

\[
(m_{1j}, ..., m_{i-1,j}, ..., m_{j-1,j}, ..., m_{n,j})
\]

Therefore we get (2.5).

8.1.2. Justification of (2.14) — (2.16).

Let \( A(\cdot) \) be defined by (1.2). Let \( s_1, s_2, \{a_k(\theta), k = -s_1, ..., s_1\} \) be given in the proof of Theorem 2.13. Then as is given in the proof of Theorem 2.13, \( g_{\theta_1}(\omega), g_{\theta_1,\theta_2} (\omega), g_{\theta_1,\theta_2,\theta_3} (\omega) \) are of the following form

\[
s_2 \sum_{k = -s_1}^{s_1} a_k(\theta)e^{ik\omega}/|A(e^{i\omega})|^{2s_2} \tag{8.1}
\]
Note that
\[ |A(e^{i\omega})| = |\prod_{j=1}^{p} (e^{i\omega} - m_j)| \]
\[ \geq \prod_{j=1}^{p} (1 - |m_j|) \]
\[ \geq \prod_{j=1}^{p} (1 - 1 + \varepsilon) = \varepsilon^p. \]

Therefore if we let \( M \) be the bound of \( \{a_k(\theta); k = -s_1, ..., s_1, \theta \in \theta\} \), then the expression (8.1) is bounded by
\[ \sigma^2 M (2s_1 + 1) / \varepsilon^{2p_2}. \]

Fix \( i \). Now we give a justification of
\[ r_{\theta_i} (h) = \int_{\Pi} e^{ih\omega} g_{\theta_i}(\omega)d\omega, h = ..., -1, 0, 1, ... \] (8.2)

Let \( \theta(i) = (\theta_1, ..., \theta_{i-1}, \theta, \theta_{i+1}, ..., \theta_R) \). Note that for fixed integer \( h \),

\[ r_{\theta_i} (h; \theta(i)) \]
\[ = \lim_{\theta \to \theta(i)} \frac{r(h; \theta) - r(h; \theta(i))}{\theta_i - \theta_i} \]
\[
\lim_{\vartheta_i \to \vartheta_i^o} \int_{-\pi}^{\pi} e^{ih\omega} \left\{ \frac{g(\omega; \vartheta) - g(\omega; \vartheta^o(i))}{\vartheta - \vartheta_i^o} \right\} d\omega.
\]

Now by the mean value theorem,

\[
\frac{g(\omega; \vartheta) - g(\omega; \vartheta^o(i))}{\vartheta - \vartheta_i^o} = g_{\vartheta_1}^o(\vartheta, \vartheta^*(i)),
\]

where \( \vartheta^*(i) \) is between \( \vartheta \) and \( \vartheta^o(i) \). Therefore for all \( \vartheta_i \),

\[
\left| \frac{g(\omega; \vartheta) - g(\omega; \vartheta^o(i))}{\vartheta - \vartheta_i^o} \right| = \left| g_{\vartheta_1}^o(\vartheta, \vartheta^*(i)) \right|
\]

\[
\leq (2\pi)^{-1} \sigma M(2s_1 + 1) / \epsilon^{2p^2}.
\]

Hence we can apply the dominated convergence theorem to make the limit operation in (8.3) go into the integral sign and get result (8.2).

Next fix \( j \). Starting with \( r_{\vartheta_1}(h) \) and \( g_{\vartheta_1}^o(\omega) \) instead of \( r(h; \vartheta) \) and \( g(\omega; \vartheta) \) and letting \( \vartheta_j \to \vartheta_j^o \) and following the same line used in getting (8.2) we can justify (2.15).

Similarly we have (2.16). \( \square \)
8.2. Appendix for Chapter 3

8.2.1. \((T_n, G_n, n \geq 1)\) is a submartingale.

Clearly \(E|T_n| < \infty\) for all \(n\). Next by the conditional Jensen's inequality,

\[
E(T_{n+1}|G_n)
\]

\[
= E[\lambda|e_1e_2 + \cdots + e_{n-1}e_n + e_ne_{n+1}| + \lambda^2|e_1e_3 + \cdots + e_{n-2}e_n + e_{n-1}e_{n+1}|
\]

\[
+ \cdots + \lambda^{n-1}|e_1e_n + e_{n-1}e_{n+1}| + \lambda^n|e_1e_{n+1}| |e_1...e_n|
\]

\[
\geq \lambda|E\{(e_1e_2 + \cdots + e_{n-1}e_n + e_ne_{n+1})| e_1...e_n\}|
\]

\[
+ \lambda^2|E\{(e_1e_3 + \cdots + e_{n-2}e_n + e_{n-1}e_{n+1})| e_1...e_n\}|
\]

\[
+ \cdots + \lambda^{n-1}|E\{(e_1e_n + e_{n+1})| e_1...e_n\}| = T_n.
\]

8.2.2. From

\[
T_{n+1} = \lambda|e_1e_2 + \cdots + e_{n-1}e_n + e_ne_{n+1}| + \lambda^2|e_1e_3 + \cdots + e_{n-2}e_n + e_{n-1}e_{n+1}|
\]

\[
+ \cdots + \lambda^{n-1}|e_1e_n + e_{n+1}| + \lambda^n|e_1e_{n+1}|
\]

we have

\[
E(T_{n+1}^2 - T_n^2) =
\]
\[ E[\lambda^2(e_1e_2 + \cdots + e_{n-1}e_n + e_ne_{n+1})^2 + \cdots + \lambda^{2(n-1)}(e_1e_n + e_ne_{n+1})^2 + \lambda^{2n}(e_1e_{n+1})^2] + \]

\[ - \lambda^2(e_1e_2 + \cdots + e_{n-1}e_n)^2 - \cdots - \lambda^{2(n-1)}(e_1e_n)^2 ] + \]

\[ 2E[\lambda^3\{e_1e_2 + \cdots + e_{n-1}e_n + e_ne_{n+1} \mid e_1e_3 + \cdots + e_{n-1}e_{n+1} \} - e_1e_2 + \cdots + e_{n-1}e_n \mid e_1e_3 + \cdots + e_{n-2}e_n] \]

\[ + \lambda^4\{e_1e_2 + \cdots + e_{n-1}e_n + e_ne_{n+1} \mid e_1e_4 + \cdots + e_{n-2}e_{n+1} \} - e_1e_2 + \cdots + e_{n-1}e_n \mid e_1e_4 + \cdots + e_{n-3}e_n \} \]

\[ \ldots \ldots \]

\[ + \lambda^{n+1}\{e_1e_2 + \cdots + e_{n-1}e_n + e_ne_{n+1} \mid e_1e_{n+1} \} - e_1e_2 + \cdots + e_{n-1}e_n \mid \{0\}] + \]

\[ 2E[\lambda^5\{e_1e_3 + \cdots + e_{n-2}e_n + e_{n-1}e_{n+1} \mid e_1e_4 + \cdots + e_{n-2}e_{n+1} \} - e_1e_3 + \cdots + e_{n-2}e_n \mid e_1e_4 + \cdots + e_{n-3}e_n \} \]

\[ \ldots \ldots \]

\[ + \lambda^{n+2}\{e_1e_3 + \cdots + e_{n-2}e_n + e_{n-1}e_{n+1} \mid e_1e_{n+1} \} - e_1e_3 + \cdots + e_{n-2}e_n \mid \{0\}] \]

\[ + \ldots \ldots + \]

\[ 2E[\lambda^{2n-1}\{e_1e_n + e_ne_{n+1} \mid e_1e_{n+1} \} - 0 \} ] \]
the expression in (3.3).

For the justification of the inequality look, for example,

\[
E[|e_1e_2 + \cdots + e_{n-1}e_n + e_{n}e_{n+1}| - |e_1e_2 + \cdots + e_{n-1}e_n|]
\]

\[
\leq E[|e_1e_2 + \cdots + e_{n-1}e_n| |e_{n+1}| + |e_1e_3 + \cdots + e_{n-2}e_n|]
\]

by Cauchy–Schwartz inequality.

8.2.3.

\[
Z'D'_nD_nZ = \sum_{i=1}^{n} \sum_{j=1}^{n} n_{ij} \sum_{s=0}^{\max(i,j)} d_s d_{|i-j|+s}
\]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=0}^{\max(i,j)} \sum_{i_1 \leq i \leq i_2} \sum_{j_1 \leq j \leq j_2} \prod_{i_1 \leq i \leq i_2} d_{i-i_1} d_{j-j_2} e_i e_j
\]

8.2.4. Let \( M < \infty \) and \( \lambda \in (0,1) \) be the maximums of the coefficients of uniformly exponential decline of the sequence \( \{d_j\} \) and coefficients of exponential decline of the sequence \( \{v_j^0\} \).

1) For \( 1 < i_1 < i_2 \leq n \);
\[
\sup_{\theta} |a_{n,i_1,i_2}| = \sup_{\theta} \left| \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=0}^{n-\max(i,j)} v_{i-i_1} v_{j-i_2} d_{s} d_{i-j+s} \right|
\]

\[
\leq M^{4} \sum_{i=i_1}^{n} \sum_{j=i_2}^{n} \sum_{s=0}^{\infty} \lambda^{i-i_1+j-i_2+s+|i-j|+s} (1-\lambda^2)^{-1}
\]

\[
\leq M^{4} \sum_{i=i_1}^{n} \sum_{j=i_2}^{n} \lambda^{i-i_1+j-i_2+i-j} (1-\lambda^2)^{-1}
\]

\[
\leq M^{4} \sum_{i=i_1}^{i_2} \sum_{j=i_2}^{n} \lambda^{i-i_1+j-i_2+j-i} + \sum_{i=i_2+1}^{n} \sum_{j=i_2}^{n} \lambda^{i-i_1+j-i_2+i-j}
\]

\[
\leq M^{4} \sum_{i=i_2}^{i_2+1} \sum_{j=i_2}^{n} \lambda^{i-i_1+j-i_2+j-i} (1-\lambda^2)^{-1}
\]

\[
\leq M^{4} [(i_2-i_1+1)(1-\lambda^2)^{-2} \lambda^{i_2-i_1} + \sum_{i=i_2+1}^{n} (i_2-i_1+1) \lambda^{2i-i_2-i_1}]
\]

\[
+ (1-\lambda^2)^{-2} \lambda^{i_2-i_1} (1-\lambda^2)^{-1} \leq M_1 \lambda^{i_2-i_1}
\]

where \(\lambda_1\) is a number in \((\lambda, 1)\) and

\[
M_1 = M^{4} [(1-\lambda^2)^{-2} \sup_{i} (i+1)(\lambda/\lambda_1)^i + \sum_{i=0}^{\infty} (i+1)\lambda^{2i} + (1-\lambda^2)^{-2}(1-\lambda^2)^{-1}]
\]

(8.5)

2) For \(1 \leq i_2 \leq i_1 \leq n\), it follows from 1) and symmetry.

3) For \(i_1 < 1 \leq i_2 \leq n\);

\[
\sup_{\theta} |a_{n,i_1,i_2}| = \sup_{\theta} \left| \sum_{i=1}^{n} \sum_{j=i_2}^{n} \sum_{s=0}^{n-\max(i,j)} v_{i-i_1} v_{j-i_2} d_{s} d_{i-j+s} \right|
\]
The remainings are the same as 1) after (8.4).

4) For $i_2 < 1 \leq n$, it follows from 3) and symmetry.

5) For $i_1 < 1$ and $i_2 < 1$;

$$\sup_{\theta} |a_{n,i_1,i_2}| = \sup_{\theta} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=0}^{n-\max(i,j)} \lambda^{i-i_1+j-i_2+s+|i-j|+s}$$

$$\leq M^4 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{s=0}^{\infty} \lambda^{i-i_1+j-i_2+s+|i-j|+s}$$

$$\leq M^4 (1-\lambda)^{-2(1-\lambda^2)-1} \lambda^{-i_1-i_2}.$$

8.2.5.

$$b_{n,i_1,i_2} = a_{n,i_1,i_2} - a_{1-i_2}$$

$$\equiv \sum_{i=i_1}^{n} \sum_{j=i_2}^{n} \sum_{s=0}^{n-\max(i,j)} \lambda^{i-i_1+j-i_2+s+|i-j|+s}$$

$$- \sum_{i=i_1}^{\infty} \sum_{j=i_2}^{\infty} \sum_{s=0}^{\infty} \lambda^{i-i_1+j-i_2+s+|i-j|+s}$$
\[= M^4(1-\lambda^2)^{-1} \sum_{i=n+1}^\infty \sum_{j=i_2}^\infty \lambda^{i-i_1+j-i_2+i-j} + \sum_{i=n+1}^\infty \sum_{j=i+1}^\infty \lambda^{i-i_1+j-i_2+j-i}\]

\[= M^4(1-\lambda^2)^{-1} \sum_{i=n+1}^\infty (i-i_2+1)\lambda^{2i-i_1-i_2} + (1-\lambda^2)^{-1} \sum_{i=n+1}^\infty \lambda^{2i-i_1-i_2}\]

\[\leq M'_2 \lambda_1^{2n-i_1-i_2}, \quad \text{for some } M'_2 < \infty \text{ and } \lambda_1 \in (\lambda, 1).\]

See 8.2.4 for a way of choosing \(M'_2\) and \(\lambda_1\). Also

\[\sup_{\theta} |r_2| \leq M'_2 \lambda_1^{2n-i_1-i_2} \quad \text{by symmetry.}\]

Also

\[\sup_{\theta} |r_3| \leq M^4 \sum_{i=i_1}^n \sum_{j=i_2}^n \sum_{s>n-max(i,j)} |\lambda^{i-i_1+j-i_2+s+i-j}| + s\]

\[\leq M^4(1-\lambda^2)^{-1} \sum_{i=i_1}^n \sum_{j=i_2}^n |\lambda^{i-i_1+j-i_2+i-j}| + 2n-2\max(i,j)\]

\[= M^4(1-\lambda^2)^{-1} \sum_{i=i_1}^n \sum_{j=i_2}^n \lambda^{2n-i_1-i_2} \leq 4M^4(1-\lambda^2)^{-1}(n-i_1)(n-i_2) \lambda^{2n-i_1-i_2}.\]

Also

\[\sup_{\theta} |r_4| \leq M^4 \sum_{i=n+1}^\infty \sum_{j=n+1}^\infty \sum_{s=0}^\infty |\lambda^{i-i_1+j-i_2+i-j}| + s\]

\[= M^4(1-\lambda^2)^{-1} \sum_{i=n+1}^\infty \sum_{j=n+1}^\infty |\lambda^{i-i_1+j-i_2+i-j}|\]
Also
\[ \sup_{\theta} |r_5| \leq M^4 \sum_{i=n+1}^{\infty} \sum_{j=1}^{\infty} \sum_{s=0}^{\infty} \lambda^{i-i_1+j-i_2+s} |i-j| + s \]
\[ \leq M^4 (1-\lambda^2)^{-1} \sum_{i=n+1}^{\infty} \sum_{j=i_2}^{\infty} \lambda^{2i-i_1-i_2} = M^4 (1-\lambda^2)^{-2} (n-i_2+1) \lambda^{2n-i_1-i_2} \]
\[ \leq 2M^4 (1-\lambda^2)^{-2} (n-i_2) \lambda^{2n-i_1-i_2} . \]
Also
\[ \sup_{\theta} |r_6| \leq 2M^4 (1-\lambda^2)^{-2} (n-i_2) \lambda^{2n-i_1-i_2} \] by symmetry.

Finally
\[ \sup_{\theta} |r_7| \leq M^4 (1-\lambda^2)^{-1} (1-\lambda)^{-2} \lambda^2 \lambda^{2n-i_1-i_2} \] by the same method for \( r_4 \).

Therefore
\[ \sup_{\theta} |b_{n,i_1,i_2}| \leq \sup_{\theta} |r_1| + \sup_{\theta} |r_2| + \sup_{\theta} |r_3| + \sup_{\theta} |r_4| + \sup_{\theta} |r_5| \]
\[ + \sup_{\theta} |r_6| + \sup_{\theta} |r_7| \leq M_2 (n-i_1)(n-i_2) \lambda^{2n-i_1-i_2}, \]

for some \( \lambda_1 \in (\lambda, 1) \) and \( M_2 < \infty \).

8.2.6. Let \( i_1 > i_2 \); for some \( M_3' < \infty \) and \( \lambda_1 \in (\lambda, 1) \), we have
\[ \sum_{i=i_1}^{\infty} \sum_{j=i_2}^{\infty} \lambda^{i-i_1+j-i_2+|i-j|} = \sum_{i=i_1}^{\infty} \sum_{j=i_2}^{\infty} \lambda^{i-i_1+j-i_2+i-j} + \sum_{i=i_1}^{\infty} \sum_{j=i_2}^{\infty} \lambda^{i-i_1+j-i_2+j-i} \]

\[ = \sum_{i=i_1}^{\infty} (i-i_2+1) \lambda^{2i-i_1-i_2} + \sum_{i=i_1}^{\infty} \lambda^{2i-i_1-i_2}(1-\lambda^2)^{-1}\lambda^2 \]

\[ \leq M_3 \sum_{i=i_1}^{\infty} \lambda_1^{2i-i_1-i_2} + \sum_{i=i_1}^{\infty} \lambda^{2i-i_1-i_2}(1-\lambda^2)^{-1} \leq M_3 [(1-\lambda_1^2)^{-1} + (1-\lambda_1^2)^{-2}] \lambda_1^{i-i_2}. \]

8.2.7. Continuity of \( d_j(\theta) \).
From the difference equation (2.11), \( d_j(\theta) \), \( j=0,1,\ldots \), can be obtained recursively. The recursion yields a polynomial of \( \theta \), that is, \( d_j(\theta) \) is a polynomial in \( \theta \) with degree \( j \). Therefore \( d_j(\theta) \) is continuous in \( \theta \).

8.2.8. Continuity of \( \text{Var}(e(\theta; \theta_0)) \)
Note that, by the dominated convergence theorem,

\[ \text{Var}(e_0(\theta, \theta_0)) = E[\lim_{n \to \infty} \sum_{i=1}^{n} \sum_{j=1}^{n} v_i^0 v_j^0 \hat{z}_i(\theta) \hat{z}_j(\theta)] \]

\[ = \lim_{n \to \infty} E[\sum_{i=1}^{n} \sum_{j=1}^{n} v_i^0 v_j^0 \hat{z}_i(\theta) \hat{z}_j(\theta)] \]

\[ = \lim_{n \to \infty} \sum_{i=1}^{n} \sum_{j=1}^{n} v_i^0 v_j^0 \sum_{s=0}^{\infty} d_s d_{i-j+s} = \lim_{n \to \infty} f_n(\theta), \text{ say.} \]

Observe that
Thus $f_n(\cdot)$ is continuous in $\theta$ because it is a uniform limit of continuous function.

Therefore $\text{Var}(\hat{e}_0(\theta, \varphi))$, which is again a uniform limit of continuous function $f_n(\cdot)$, is continuous.

8.2.9. $\inf_{\theta, \delta} \{ \text{Var}(\hat{e}_0(\theta; \varphi)) - \text{Var}(\hat{e}_0(\varphi_0; \varphi)) \} > 0$ for all $\delta > 0$.

Assume not. There is a sequence $\{ \theta_n \}$ and $\delta > 0$ such that $|\theta_n - \varphi| \geq \delta$ for all $n$ and

$$\text{Var}(\hat{e}_0(\theta_n; \varphi)) - \text{Var}(\hat{e}_0(\varphi_0; \varphi)) \to 0 \quad \text{as} \quad n \to \infty.$$ 

Since the set $\{ \theta \in \Theta ; |\theta -- \varphi| \geq \delta \}$ is compact we can find a limit point $\theta^*$ of $\{ \theta_n \}$ in that set. Hence there is a subsequence $\{n_k\}$ such that $\theta^* = \lim_{k \to \infty} \theta_{n_k}$. Therefore by the dominated convergence theorem,

$$\text{Var}(\hat{e}_0(\theta^*; \varphi)) = \lim_{k \to \infty} \text{Var}(\hat{e}_0(\theta_{n_k}; \varphi)) = \text{Var}(\hat{e}_0(\varphi; \varphi)).$$

Hence we have $\theta^* = \varphi$. However this contradicts
\[ |\theta - \theta^0| = \lim_{k \to \infty} |\theta_n^k - \theta^0| \geq \delta. \]

8.2.10. For the continuity of \( d_8(\theta) \) see Appendix 8.2.7.

In 8.2.11 – 8.2.20, without loss of generality we can assume that \( M \) and \( \lambda \) are the coefficients of exponential decline of the sequence \( \{v_j^O\} \). Let \( \lambda_1 \in (\lambda, 1) \).

8.2.11.

\[
\sum_{j=1}^{n} \sum_{k=j}^{n} A_{k-j} z_k z_j = \sum_{j=1}^{n} \sum_{k=j}^{n} A_{k-j} \sum_{i_1 \leq k} \sum_{i_2 \leq j} v_{k-i_1}^O v_{j-i_2}^O e_{i_1} e_{i_2}
\]

\[
= \sum_{i_1 \leq n} \sum_{i_2 \leq n} \sum_{k=\max(1, i_1, i_2)}^{n} \sum_{j=\max(1, i_2)}^{k} A_{k-j} v_{k-i_1}^O v_{j-i_2}^O e_{i_1} e_{i_2}
\]

\[
= \text{right side of (3.19)}
\]

8.2.12.

1) \[
\sum_{k=\max(1, i_1, i_2)}^{n} \sum_{j=1}^{i_2} |A_{k-j} v_{k-i_1}^O v_{j-i_2}^O| \leq M^3 \sum_{k=\max(1, i_1, i_2)}^{n} \sum_{j=1}^{i_2} \lambda^{k-j+i_1-j-i_2}
\]

\[
\leq M^3 \sum_{k=\max(1, i_1, i_2)}^{n} (k-i_2+1) \lambda^{2k-i_1-i_2} \leq M_1 \sum_{k=\max(1, i_1, i_2)}^{n} \lambda^{2k-i_1-i_2}
\]
(for some $M_1 < \infty$ by the same method for (8.2.4)).

\[ M_1 \lambda_1^{2 \max(i_1, i_2) - i_1 - i_2} = M_1 \lambda_1^{i_1 - i_2}, \quad \text{where } M_1 = M_1(1-\lambda^2)^{-1}. \]

2) \[ \sum_{k=i_2}^{n} \sum_{j=i_2}^{k} |A_{k-j} v_{k-i_2}^O v_{j-i_2}^O| \leq M^3 \sum_{k=i_2}^{n} \sum_{j=i_2}^{k} \lambda^{2k-i_1-i_2} \leq M_1 \sum_{k=i_2}^{n} \lambda_1^{2k-i_1-i_2} \]

\[ = M_1 \lambda_1^{i_2-i_1}. \]

3) \[ \sum_{k=i_1}^{n} \sum_{j=1}^{k} |A_{k-j} v_{k-i_2}^O v_{j-i_2}^O| \leq M_1 \sum_{k=i_1}^{n} \sum_{j=1}^{k} \lambda^{2k-i_1-i_2} \leq M_1 \sum_{k=i_1}^{n} \sum_{j=1}^{k} \lambda^{(k-i_2+1) \lambda^{2k-i_1-i_2}} \]

\[ \leq M_1 \sum_{k=i_1}^{n} \lambda_1^{2k-i_1-i_2} = M_1(1-\lambda^2)^{-1} \lambda_1^{i_1-i_2}. \]

4) \[ \sum_{k=1}^{n} \sum_{j=1}^{k} |A_{k-j} v_{k-i_2}^O v_{j-i_2}^O| \leq M_1 \sum_{k=1}^{n} \sum_{j=1}^{k} \lambda^{2k-i_1-i_2} \]

\[ \leq M_1 \sum_{k=1}^{n} \lambda_1^{2k-i_1-i_2} \leq M_1(1-\lambda^2)^{-1} \lambda_1^{i_1-i_2}. \]

8.2.13.

1) For $1 \leq i_1, i_2 \leq n$; Note that for some $N_1 < \infty$,

\[ |a_{n,i_1,i_2}| \leq M^3 \sum_{k=\max(i_1, i_2)}^{n} \sum_{j=i_2}^{k} \lambda^{n+j-i_1-i_2} \leq M^3(1-\lambda)^{-1} \sum_{k=\max(i_1, i_2)}^{n} \lambda^{n-i_1} \]

\[ \sum_{j=i_2}^{k} \lambda^{n+j-i_1-i_2} \leq M^3(1-\lambda)^{-1} \sum_{k=\max(i_1, i_2)}^{n} \lambda^{n-i_1} \]
\[ \leq M^3(1-\lambda)^{-1}(n-\max(i_1,i_2)+1)\lambda^{n-i_1} \leq 2M^3(1-\lambda)^{-1}(n-i_1)\lambda^{n-i_1} \leq N_1\lambda^{n-i_1}. \]

2) For \( i_1 \leq 0 < i_2 \leq n \); Note that for some \( N_2, M'_2 < \infty \),

\[ |a_{n,i_1,i_2}| \leq M^3 \Omega \sum_{k=i_2} \sum_{j=i_2} \lambda^{n+j-i_1-i_2} \leq M^3(1-\lambda)^{-1} \sum_{k=i_2} \lambda^{n-i_1} \]

\[ \leq M^3(1-\lambda)^{-1}(n-i_2+1)\lambda^{n-i_1} \leq 2M^3(1-\lambda)^{-1}(n-i_1)\lambda^{n-i_1} \leq M'_2\lambda^{n-i_1}. \]

3) For \( i_2 \leq 0 < i_1 \leq n \); Note that for some \( N_3, M'_3 < \infty \),

\[ |a_{n,i_1,i_2}| \leq M^3 \sum_{k=i_1} \sum_{j=1} \lambda^{n+j-i_1-i_2} \leq M^3(1-\lambda)^{-1} \sum_{k=i_1} \lambda^{n-i_1-i_2} \]

\[ \leq M^3(1-\lambda)^{-1}(n-i_1+1)\lambda^{n-i_1-i_2} \leq 2M^3(1-\lambda)^{-1}(n-i_1-i_2)\lambda^{n-i_1-i_2} \]

\[ \leq M'_3\lambda^{n-i_1-i_2} \leq M'_3\lambda^{n-i_2}. \]

4) For \( i_1 < 1 \) and \( i_2 < 1 \); Note that for some \( N_4, M'_4 < \infty \),

\[ |a_{n,i_1,i_2}| \leq M^3 \sum_{k=1} \sum_{j=1} \lambda^{n+j-i_1-i_2} \leq M^3(1-\lambda)^{-1} \sum_{k=1} \lambda^{n-i_1-i_2} \]

\[ \leq M^3(1-\lambda)^{-1}n\lambda^{n-i_1-i_2} \leq M^3(1-\lambda)^{-1}(n-i_1-i_2)\lambda^{n-i_1-i_2} \leq M'_4\lambda^{n-i_1-i_2}. \]
8.2.14. Let

\[ S_1 = \sum_{i_1=1}^{n} \sum_{i_2 \leq 0} a_{n,i_1,i_2} e_{i_1} e_{i_2}, \quad S_2 = \sum_{i_1 \leq 0}^{n} \sum_{i_2=1} \sum_{i_2 \leq 0} a_{n,i_1,i_2} e_{i_1} e_{i_2}, \quad S_3 = \sum_{i_1 \leq 0}^{n} \sum_{i_2 \leq 0} a_{n,i_1,i_2} e_{i_1} e_{i_2} \]

Clearly \( E(S_1) = E(S_2) = 0 \). Note that,

\[ \text{Var}(S_1) = \sum_{i_1=1}^{n} \sum_{i_2 \leq 0} a_{n,i_1,i_2}^2 \sigma^4 \leq M_3^2 \sum_{i_1=1}^{n} \sum_{i_2 \leq 0} \lambda_2^{2(n-i_2)} \sigma^4 = O(n) \]

and

\[ \text{Var}(S_2) = O(n) \quad \text{by symmetry}. \]

Therefore \( S_1 = O_p(n^{1/2}) \) and \( S_2 = O_p(n^{1/2}) \). Next observe that,

\[ |S_3| \leq M_4^2 \sum_{i_1 \leq 0}^{n} \sum_{i_2 \leq 0} \lambda_1^{n-i_1-i_2} |e_{i_1} e_{i_2}| = M_4^2 \lambda_1^n (\sum_{i_1 \leq 0}^{n} \lambda_1^{-i} |e_i|)^2 = O_p(1). \]

Therefore \( S_1 + S_2 + S_3 = O_p(n^{1/2}) \).

8.2.15.

\[ \sum_{j=1}^{n} \sum_{k=1}^{j} a_{j-k}(y_{k-1}-y_0) \sum_{j=1}^{n} \sum_{k=1}^{j} v_{j-k} \sum_{s=1}^{i_1} \sum_{j-i_1}^{n} v_{s-i_2} e_{i_1} e_{i_2} \]

\[ = \sum_{i_1 \leq n}^{n} \sum_{i_2 \leq n-1}^{i_1} \sum_{j=\max(1,i_1)}^{\max(1,i_2+1)} \sum_{s=\max(1,i_2)}^{k-1} a_{j-k} v_{j-i_1} v_{s-i_2} e_{i_1} e_{i_2} \]

\[ = \sum_{i_1 \leq n}^{n} \sum_{i_2 \leq n-1}^{i_1} b_{n,i_1,i_2} e_{i_1} e_{i_2}. \]
1) For $1 \leq i_1, i_2 \leq n$; Note that for some $M''_1 < \omega$, 

$$|b_{n,i_1,i_2}| \leq M^3 \sum_{j=i_1}^{n} \sum_{k=i_2+1}^{j} \sum_{s=i_2}^{k-1} \lambda^{j-k+j-i_1+s-i_2} \leq M^3 \sum_{j=i_1}^{n} \sum_{k=i_2+1}^{j} \lambda^{2j-k-i_1} (1-\lambda)^{-1} \leq M^3 \sum_{j=i_1}^{n} \lambda^{j-i_1} (1-\lambda)^{-2} \leq M''_1.$$ 

2) For $i_2 \leq 0 < i_1 \leq n$; Note that for some $M''_2 < \omega$, 

$$|b_{n,i_1,i_2}| \leq M^3 \sum_{j=i_1}^{n} \sum_{k=1}^{j} \sum_{s=1}^{k-1} \lambda^{j-k+j-i_1+s-i_2} \leq M^3 \sum_{j=i_1}^{n} \lambda^{2j-k-i_1-i_2} (1-\lambda)^{-1} \leq M^3 \sum_{j=i_1}^{n} \lambda^{j-i_1-i_2} (1-\lambda)^{-2} \leq M''_2 \lambda^{-i_2}.$$ 

3) For $i_1 \leq 0 < i_2 \leq n$; Note that for some $M''_3$ 

$$|b_{n,i_1,i_2}| \leq M^3 \sum_{j=1}^{n} \sum_{k=i_2+1}^{j} \sum_{s=i_2}^{k-1} \lambda^{j-k+j-i_1+s-i_2} \leq M^3 \sum_{j=1}^{n} \sum_{k=i_2+1}^{j} \lambda^{2j-k-i_1} (1-\lambda)^{-1} \leq M^3 \sum_{j=1}^{n} \lambda^{j-i_1} (1-\lambda)^{-2} \leq M''_3 \lambda^{-i_1}.$$ 

4) For $i_1 \leq 0, i_2 \leq 0$; Note that for some $M''_4 < \omega$, 


Now note that for some $M_1''' < \infty$,

i) for $i_1 \leq i_2$;

$$|c_{n,i_1,i_2}| \leq M^3 \left( \sum_{j=i_1}^{i_2} \sum_{k=i_2+1}^{n} \lambda^{k-i_1} + \sum_{j=i_2+1}^{n} \sum_{k=j+1}^{n} \lambda^{k-i_1} \right)(1-\lambda)^{-1}$$

$$\leq M^3 \left[ (i_2-i_1+1)\lambda^{i_2-i_1} (1-\lambda)^{-1} + \lambda^{i_2-i_1} (1-\lambda)^{-2} \right](1-\lambda)^{-1} \leq M_1''' ,$$

ii) for $i_1 \geq i_2$;

$$|c_{n,i_1,i_2}| \leq M^3 \sum_{j=i_1}^{n-1} \sum_{k=j+1}^{n} \lambda^{k-i_1} (1-\lambda)^{-1} \leq M_1'''.$$

2) For $i_2 \leq 0 < i_1 < n-1$; Note that for some $M_2''' < \infty$,

$$|c_{n,i_1,i_2}| \leq M^3 \sum_{j=i_1}^{n-1} \sum_{k=j+1}^{n} \lambda^{k-s-i_1-i_2} \leq M^3 \sum_{j=i_1}^{n-1} \sum_{k=j+1}^{n} \lambda^{k-i_1-i_2} (1-\lambda)^{-1}$$

$$\leq M^3 \sum_{j=i_1}^{n-1} \lambda^{j-i_1-i_2} (1-\lambda)^{-2} \leq M_2''' \lambda^{-i_2}$$

3) For $i_1 \leq 0 < i_2 \leq n-1$; Note that for $M_3''' < \infty$ and $\lambda_1 \in (\lambda, 1)$,

$$|c_{n,i_1,i_2}| \leq M^3 \sum_{j=1}^{n-1} \sum_{k=\max(j+1,i_2+1)}^{n} \lambda^{k+s-i_1-i_2}$$
\[ \leq M^3 \sum_{j=1}^{n-1} \sum_{k=\max(j+1,i_2+1)}^{n} \lambda^{k-i_1} (1-\lambda)^{-1} \]

\[ \leq M^3 \sum_{j=1}^{i_2} \sum_{k=\max(j+1,i_2+1)}^{n-1} \sum_{j=i_2+1}^{n} \lambda^{k-i_1} (1-\lambda)^{-1} \]

\[ \leq M^3 (i_2 \lambda^{i_2-i_1} + \lambda^{i_2-i_1}) (1-\lambda)^{-3} \leq M_3'''' \lambda^{i_2-i_1} \leq M_3'''' \lambda^{-i_1} \]

4) For \( i_1 \leq 0 \) and \( i_2 \leq 0 \); Note that for some \( M_4''' < \omega \),

\[ |c_{n,i_1,i_2}| \leq M^3 \sum_{j=1}^{n-1} \sum_{k=\max(j+1,i_2+1)}^{n} \lambda^{k-s-i_1-i_2} \leq M^3 \sum_{j=1}^{n-1} \sum_{k=\max(j+1,i_2+1)}^{n} \lambda^{k-i_1-i_2} (1-\lambda)^{-1} \]

\[ \leq M^3 \sum_{j=1}^{n-1} \lambda^{j-i_1-i_2} (1-\lambda)^{-2} \leq M_4''' \lambda^{-i_1-i_2} \]

8.2.19.

\[ \sum_{j=1}^{n} \sum_{k=1}^{i_1<\lambda} a_{n-j} \sum_{i_1<\lambda} e_i_1 \sum_{i_2=1}^{i_2<\lambda} e_{i_2} | \leq M^2 \sum_{j=1}^{n} \sum_{k=1}^{i_1<\lambda} \lambda^{n-j} \sum_{i_1<\lambda} e_i_1 \sum_{i_2=1}^{i_2<\lambda} e_{i_2} | \]

\[ \leq M^2 \sum_{j=1}^{n} \sum_{k=1}^{i_1<\lambda} \lambda^{n-j} \sum_{i_1<\lambda} e_i_1 \sum_{i_2=1}^{i_2<\lambda} e_{i_2} | \]

\[ \leq M^2 (\sigma^o)^2 \sum_{j=1}^{n} \sum_{k=1}^{i_1<\lambda} \lambda^{n-j} \sum_{i_1<\lambda} e_i_1 \sum_{i_2=1}^{i_2<\lambda} e_{i_2} | \]

\[ \leq M^2 (\sigma^o)^2 (1-\lambda)^{-1} \sum_{j=1}^{n} \lambda^{n-j} = O(n). \]

For some \( M_1 < \omega \),

\[ \sum_{j=1}^{n} \sum_{k=1}^{i_1<\lambda} a_{n-j} \sum_{i_1<\lambda} e_i_1 \sum_{i_2=1}^{i_2<\lambda} e_{i_2} | \]
\[
\sum_{j=1}^{n} \sum_{k=1}^{n} \lambda^{n-j} [E(z_j)^2]^{1/2} [M(1-\lambda^2)e_1]^2 \leq M_1^2 \sum_{j=1}^{n} \sum_{k=1}^{n} \lambda^{n-j} k^{1/2} \\
\leq M_1 \sum_{j=1}^{n} \sum_{k=1}^{n} \lambda^{n-j} (j^k k)^{1/2} \leq M_1 \sum_{j=1}^{n} \sum_{k=1}^{n} \lambda^{n-j} j^{3/2} = O(n^{3/2}).
\]

For some \( M_2 < \infty \),

\[
\sum_{j=1}^{n} \sum_{k=1}^{n} a_n^{-j} \sum_{i=0}^{k-1} \sum_{s=1}^{k-1} \lambda^{s-i} e_i \leq M_2 \sum_{j=1}^{n} \sum_{k=1}^{n} \lambda^{n-j} k^{1/2} \leq M_2 \sum_{j=1}^{n} \sum_{k=1}^{n} j\lambda^{n-j} = O(n).
\]

8.2.20. For some \( M_3 < \infty \),

\[
\sum_{j=1}^{n} \sum_{k=1}^{n} a_{n-k} \sum_{i_1=0}^{k-1} \sum_{i_2=1}^{k-1} \lambda^{n-k} [E(\sum_{j=i_1}^{k-1} e_i)^2 E(\sum_{i_2=1}^{k-1} e_i)^2]^{1/2} \\
\leq M \sum_{j=1}^{n} \sum_{k=1}^{n} \lambda^{n-k} [E(\sum_{j=i_1}^{k-1} e_i)^2 E(\sum_{i_2=1}^{k-1} e_i)^2]^{1/2} \\
\leq M_3 \sum_{j=1}^{n} \sum_{k=1}^{n} \lambda^{n-k} k^{1/2} = M_3 \sum_{k=1}^{n} \lambda^{n-k} k^{1/2} = O(n^{3/2}).
\]
\[ s + f - 1 \cdot s + f = 1 \quad 0 = s \quad s + f - 1 \cdot s + f = 1 \]

\[ s - u + f - u \quad 1 - u \quad u \]

\[ 1 \leq x + 1 \leq 1 + f - u \quad p + 1 \leq 1 + f - u \quad p + 1 \leq 1 + f - u \]

\[ (1 - x^{f} + x^{g} + x^{k}) \quad 0 = s \quad s + f - 1 \cdot s + f = s \]

\[ s + f - 1 \cdot s + f = 1 \quad 0 = s \quad 1 + f = 1 \quad 0 \]

\[ (f^{*} + x^{f} + x^{k}) \quad 0 = s \quad 1 + f = 1 \quad 0 \]

\[ (f^{*} + x^{f} + x^{k}) \quad 0 = s \quad 1 + f = 1 \quad 0 \]

\[ s + f - 1 \cdot s + f = 1 \quad 0 = s \quad s + f - 1 \cdot s + f = 1 \]

\[ s - u + f - u \quad 1 - u \quad u \]

\[ 1 \leq x + 1 \leq 1 + f - u \quad p + 1 \leq 1 + f - u \quad p + 1 \leq 1 + f - u \]

\[ (1 - x^{f} + x^{g} + x^{k}) \quad 0 = s \quad s + f - 1 \cdot s + f = s \]

\[ s + f - 1 \cdot s + f = 1 \quad 0 = s \quad 1 + f = 1 \quad 0 \]

\[ (f^{*} + x^{f} + x^{k}) \quad 0 = s \quad 1 + f = 1 \quad 0 \]

\[ s + f - 1 \cdot s + f = 1 \quad 0 = s \quad s + f - 1 \cdot s + f = 1 \]

\[ s - u + f - u \quad 1 - u \quad u \]

\[ 1 \leq x + 1 \leq 1 + f - u \quad p + 1 \leq 1 + f - u \quad p + 1 \leq 1 + f - u \]

\[ (1 - x^{f} + x^{g} + x^{k}) \quad 0 = s \quad s + f - 1 \cdot s + f = s \]

\[ s + f - 1 \cdot s + f = 1 \quad 0 = s \quad 1 + f = 1 \quad 0 \]

\[ (f^{*} + x^{f} + x^{k}) \quad 0 = s \quad 1 + f = 1 \quad 0 \]

\[ s + f - 1 \cdot s + f = 1 \quad 0 = s \quad s + f - 1 \cdot s + f = 1 \]

\[ s - u + f - u \quad 1 - u \quad u \]

\[ 1 \leq x + 1 \leq 1 + f - u \quad p + 1 \leq 1 + f - u \quad p + 1 \leq 1 + f - u \]

\[ (1 - x^{f} + x^{g} + x^{k}) \quad 0 = s \quad s + f - 1 \cdot s + f = s \]
\[ t_{g,18} \left. \left( f^{-1} p + \ldots + f^{-2} + f^{-1} + s + f^{-1} + \right. \right) p \left. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right
\[
- \sum_{s_1=0}^{\infty} \sum_{s_2=n-j+1}^{\infty} d_{s_1} d_{s_2} + \sum_{s_1=0}^{\infty} \sum_{s_2=n-k}^{\infty} d_{s_1} d_{s_2}
= \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1} d_{k_1+1-j+k_1+k_2} - \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{n-k+k_1} d_{n+1-j+k_1+k_2}
\]

and

\[
\sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1} d_{k_1+k_2} + \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1} d_{k_1+k_1+k_2} = (\sum_{k=0}^{\infty} d_k)^2. \tag{8.10}
\]

From (8.6) - (8.10), we get (3.39).

8.2.22.

\[
\sup_{\theta} \left| \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=0}^{\max(i,j)} z_j d_s d_{i-j} + s \right| \leq M^2 \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda^{|i-j|} |z_j| (1-\lambda^2)^{-1}
\]

\[
\leq M^2 \sum_{j=1}^{i} \lambda^{|i-j|} |z_j| + \sum_{i=j+1}^{n} \lambda^{|i-j|} |z_j| (1-\lambda^2)^{-1}
\]

\[
\leq 2M^2(1-\lambda)^{-1} \sum_{j=1}^{n} |z_j| (1-\lambda^2)^{-1} = O_{wp}(n).
\]

8.2.23. It is similar to the justification of (3.19) in Lemma 3.10 if we replace \(A_{k-j}\) by 1.
From (6.11) and (8.12) the result \( \delta_{ij} \) follows.

\[
\begin{align*}
\delta_{ij} &= \epsilon_{ij} + \epsilon_{ij} \\
\delta_{ij} &= \epsilon_{ij} + \epsilon_{ij} \\
\delta_{ij} &= \max(\epsilon_{ij} + \epsilon_{ij}) \\
\end{align*}
\]

For \( i > j \), note that:

\[
\begin{align*}
\epsilon_{ij} &= \epsilon_{ij} \\
\epsilon_{ij} &= \epsilon_{ij} \\
\epsilon_{ij} &= \epsilon_{ij} \\
\epsilon_{ij} &= \epsilon_{ij} \\
\epsilon_{ij} &= \epsilon_{ij} \\
\epsilon_{ij} &= \epsilon_{ij} \\
\end{align*}
\]

For \( i > j \), note that.
In 8.2.25 and 8.2.26, let $M$ and $\lambda$ be the coefficients of the exponential decline of the sequence $\{v_j^0\}$.

8.2.25.

$$|a_{n,i_1,i_2}| \leq M^2 \sum_{j=n+1}^{\infty} \lambda^{j-i_1+s-i_2} \leq M^2 \sum_{j=n+1}^{\infty} \lambda^{j-i_1}(1-\lambda)^{-1} \leq M^2 \lambda^{-1}(1-\lambda)^{-2}.$$

For $1 \leq i_2 < i_1 \leq n$;

$$|c_{1,i_1}| \leq M^2 \sum_{j=i_1}^{\infty} \lambda^{j-i_1+ s-i_2} \leq M^2 \sum_{j=i_1}^{\infty} \lambda^{j-i_1}(1-\lambda)^{-1} \leq M^2 \lambda^{-1}(1-\lambda)^{-2}.$$

8.2.26.

$$\left| \sum_{j=i_2+1}^{n} \sum_{s=i_2}^{\infty} v^0_{j-i_1} v^0_{s-i_2} \right| \leq M^2 \sum_{j=i_2+1}^{\infty} \lambda^{j-i_1+s-i_2} \leq M^2 \sum_{j=i_2+1}^{\infty} \lambda^{j-i_1}(1-\lambda)^{-1} \leq M^2 \lambda^{-1}(1-\lambda)^{-2}.$$
The expression given in (3.28) is:

\[
(1)_{\text{dm}}^0 = z \left( |I^0|_{\text{w}} |Y|^0 \right)_{\text{w}} = \left| I^0 \right|_{\text{w}} |Y|^0_{\text{w}} = \left| I^0 \right|_{\text{w}} |Y|^0_{\text{w}} \leq |S|
\]

and

\[
(1)_{\text{dm}}^0 = (|I^0|_{\text{w}} |Y|^0 \rightarrow 0^5_{\text{w}} |X|^0_{\text{w}} |Y|^0_{\text{w}} = \left| I^0 \right|_{\text{w}} |Y|^0_{\text{w}} \leq |S|
\]

\[
(1)_{\text{dm}}^0 = (|I^0|_{\text{w}} |Y|^0 |X|^0_{\text{w}} |Y|^0_{\text{w}} = \left| I^0 \right|_{\text{w}} |Y|^0_{\text{w}} \leq |S|
\]

For some w, we have \( \omega > 0 \).
\begin{align*}
\left\{ 0 \lambda + \left( \frac{t}{1 + t} + f \right) \left( \frac{t}{0} \right) \right\}_{\frac{u}{3}} &= \left( 0 \lambda + \left( \frac{t}{1 + t} + f \right) \left( \frac{t}{0} \right) \right)_{\frac{u}{3}} = \frac{u}{3} \frac{u}{3} \frac{u}{3} \frac{u}{3} \\
\frac{t}{1 + t} + f &= \frac{t}{1 + t} + f
\end{align*}
Let \( \{z_t\} \) be stationary and invertible then \( c_0 = \sum_{j=0}^{\infty} v_j^0 \neq 0 \)

Let \( m_1, \ldots, m_p \) be the roots of \( A(m) = 0 \) in (1.2) and \( m_1^*, \ldots, m_q^* \) be the roots of \( B(m) = 0 \) in (1.47). Then \( |m_i| < 1 \) for \( i = 1 \ldots p \) and \( |m_i^*| < 1 \) for \( i = 1 \ldots q \). Let \( B \) be the back shift operator, that is \( Bz_t = z_{t-1} \). Then

\[
(1 + \alpha_1 B + \alpha_2 B^2 + \cdots + \alpha_p B^p)z_t = (1 + \beta_1 B + \beta_2 B^2 + \cdots + \beta_p B^p)e_t
\]
or

\[
(1 - m_1 B)(1 - m_2 B) \cdots (1 - m_p B)z_t = (1 - m_1^* B)(1 - m_2^* B) \cdots (1 - m_q^* B)e_t.
\]

Therefore

\[
z_t = [(1-m_1 B)(1-m_2 B) \cdots (1-m_p B)]^{-1}[(1-m_1^* B)(1-m_2^* B) \cdots (1-m_q^* B)]e_t
\]

\[
= \sum_{j=0}^{\infty} v_j^0 e_{t-j} = \sum_{j=0}^{\infty} v_j^0 B^j e_t.
\]

Hence

\[
\sum_{j=0}^{\infty} v_j^0 B^j = [(1-m_1 B)(1-m_2 B) \cdots (1-m_p B)]^{-1}[(1-m_1^* B)(1-m_2^* B) \cdots (1-m_q^* B)]
\]

Letting \( B = 1 \), we have
218

\[ c_0 = \sum_{j=0}^{\infty} v_j^0 = [(1 - m_1)(1 - m_2) \cdots (1 - m_p)]^{-1}(1 - m_1^*)(1 - m_2^*) \cdots (1 - m_q^*) \neq 0. \]

8.2.34. Assume that \( n(1 - \rho)^2 \) does not converge to 0 in probability. There is a subsequence \( \{n_k\} \) and \( \epsilon > 0, \delta > 0 \) such that

\[ P(n_k(1 - \rho_{n_k})^2 > \epsilon) > \delta \text{ for every } k = 1, 2, \ldots. \]

Since, by Corollary 3.13 \( Y_1' Y_1 / n_k^2 \) converges in distribution to \( c_0^2 \sigma^2 \int_0^1 W^2(r)dr \),

\[ c_0 = \sum_{j=0}^{\infty} v_j^0 \text{ which is absolutely continuous and positive a.s. we can find a positive integer } K \text{ and } \epsilon_1 > 0 \text{ such that} \]

\[ P(Y_1' Y_1 / n_k^2 \leq \epsilon_1) \leq \delta/2 \text{ for every } k > K. \]

Then, for \( k > K \),

\[ P(n_k^{-1}(1 - \rho_{n_k})^2 Y_1' Y_1 > \epsilon_1) \geq P(n_k(1 - \rho_{n_k})^2 > \epsilon, Y_1' Y_1 / n_k^2 > \epsilon_1) \]

\[ = 1 - P(n_k(1 - \rho_{n_k})^2 \leq \epsilon, \text{ or } Y_1' Y_1 / n_k^2 \leq \epsilon_1) \]

\[ \geq 1 - P(n_k(1 - \rho_{n_k})^2 \leq \epsilon) - P(Y_1' Y_1 / n_k^2 \leq \epsilon_1) \]

\[ = P(n_k(1 - \rho_{n_k})^2 > \epsilon) - P(Y_1' Y_1 / n_k^2 \leq \epsilon_1) > \delta - \delta/2 = \delta/2, \]

contradicting \( n_k^{-1}(1 - \rho_{n_k})^2 Y_1' Y_1 \rightarrow 0 \) in probability. \( \Box \)
8.3. Appendix for Chapter 4

8.3.1. Verification of (4.10)

Assume \( i_1 \geq i_2 \). Note that

\[
\sum_{i_2=1}^{n} \sum_{k_2=0}^{\infty} \sum_{j_1=1}^{\infty} \sum_{k_1=0}^{\infty} \lambda^{k_1+k_2} |i_1-j_1| + k_1+k_2+|i_2-j_2| + k_2+|i_2-j_2| = (1-\lambda^2)^{-2} \left( \sum_{i_2=1}^{n} \sum_{j_1=1}^{i_2} \lambda^{|i_1-j_1|} + |i_2-j_2| + |i_2-j_1| \right)
\]

\[
\leq (1-\lambda^2)^{-2} \sum_{i_2=1}^{n} \sum_{j_1=1}^{i_2} \lambda^{|i_1-j_1|} + |i_2-j_2| + |i_2-j_1|
\]

\[
= (1-\lambda^2)^{-2} \sum_{i_2=1}^{n} \sum_{j_1=1}^{i_2} \lambda^{|i_1-j_1+|i_2-j_2|+i_2-j_1|}
\]

\[
+ \sum_{i_2=1}^{n} \sum_{j_1=i_2+1}^{i_1} \lambda^{|i_1-j_1+|i_2-j_2|-i_2+j_1|} + \sum_{i_2=1}^{n} \sum_{j_1=i_2+1}^{i_1} \lambda^{|i_1-i_2|+|i_2-j_2|-i_2+j_1|}
\]

\[
= (1-\lambda^2)^{-2} \sum_{i_2=1}^{n} \lambda^{|i_2-j_2|+i_1-i_2(1-\lambda^2)^{-1}}
\]

\[
+ \sum_{i_2=1}^{n} (i_1-i_2) \lambda^{i_1-i_2+|i_2-j_2|} + \sum_{i_2=1}^{n} \lambda^{i_2-j_2+|i_1-i_2|} \left( 1-\lambda^2 \right)^{-1}
\]

\[
\leq M_2 \sum_{i_2=1}^{n} \lambda_2^{|i_1-i_2+|i_2-j_2|} \text{ for some } M_2 < \infty \text{ and } \lambda_2 \in (0,1).
\]

Therefore, in general, (8.13) is less than or equal to
for some \( M_2 < \infty \). By the same argument (8.14) is less than or equal to

\[
M_3 \lambda_1 \left| i_r - j_2 \right|
\]

for some \( M_3 < \infty \) and \( \lambda_1 \in (\lambda_2, 1) \).

### 8.3.2. Verification of (4.14)

We have

\[
\sum_{j=1}^{n} \sum_{k=1}^{j-1} z_k z_j = 2^{-1} \left[ (\sum_{k=1}^{n} z_k)^2 - \sum_{k=1}^{n} z_k^2 \right]
\]

(8.15)

Here note that

\[
\sum_{k=1}^{n} z_k = \sum_{k=1}^{n} \sum_{i \leq k} v_{k-i} e_i = \sum_{j=1}^{n} \sum_{k=1}^{n} v_{k-j} e_j + O_p(1) = c_0 \sum_{i=1}^{n} e_i + O_p(1).
\]

(8.16)

Therefore

\[
\left( \sum_{k=1}^{n} z_k \right)^2 = c_0^2 \left( \sum_{i=1}^{n} e_i \right)^2 + O_p(n^{1/2}) = c_0^2 \sum_{i=1}^{n} e_i^2 + 2c_0^2 \sum_{i=1}^{n} \sum_{i_1=1}^{i-1} e_{i_1} e_{i_2} + O_p(n^{1/2})
\]

(8.17)

and
\[
\sum_{k=1}^{n} z_k^2 = \sum_{k=1}^{n} \sum_{i_j \leq k \leq i_{j+1}} v_{k-i_1}^O v_{k-i_2}^O e_{i_1} e_{i_2}
\]

\[
\leq \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{k=\max(i_1, i_2)}^{n} v_{k-i_1}^O v_{k-i_2}^O e_{i_1} e_{i_2} + O_p(1)
\]

\[
= 2S + \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} v_{k-i_1}^O v_{k-i_2}^O e_{i_1} e_{i_2} + O_p(1)
\]

Now we know that \(S = O_p(n^{1/2})\) because \(E(S) = 0,\)

\[
\text{Var}(S) \leq \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{k=\max(i_1, i_2)}^{n} \lambda^{2k-i_1-i_2} 2\sigma^4 \leq \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} (M^2 \sum_{k=1}^{i_1-1} \lambda^{2k-i_1-i_2} 2\sigma^4)
\]

\[
\leq M^4 \sigma^4 \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \lambda^{2(i_1-i_2)(1-\lambda^2)-2} \leq M^4 \sigma^4 \sum_{i_1=1}^{n} (1-\lambda^2)^{-4} = O(n),
\]

where \(M\) and \(\lambda\) are the coefficients of exponential decline of sequence \(\{v_k^O\}\). Also

\[
\sum_{i=1}^{n} \sum_{k=i}^{n} v_{k-i}^O e_{i}^2 = \sum_{i=1}^{n} \sum_{k=i}^{n} \omega v_{k-i}^O e_{i}^2 + O_p(1) = \sum_{k=0}^{\omega} v_k \sum_{i=1}^{n} \sum_{k=i}^{n} e_{i}^2 + O_p(1).
\]

Therefore,
Substituting the terms (\( \sum_{i=1}^{n} z_{k}^{2} \)) and (\( \sum_{k=1}^{n} z_{k}^{2} \)) in (8.15) by the terms (8.17) and (8.18), we have

\[
\sum_{j=1}^{n} \sum_{k=1}^{j-1} z_{k} z_{j} = 2^{-1} \left[ (2c_{2}^{2} \sum_{i_{1}=1}^{i_{1}} \sum_{i_{2}=1}^{i_{2}} e_{i_{1}} e_{i_{2}} + c_{2}^{2} \sum_{i=1}^{n} \sum_{i=1}^{n} e_{i}^{2} ) - \sum_{k=0}^{n} \sum_{i=1}^{n} \sum_{i=1}^{n} e_{i}^{2} \right] + O_p(n^{1/2}).
\]

8.3.3. Verification of (4.15)

\[
\sum_{j=1}^{n} \sum_{k=1}^{j-1} z_{k} z_{j} = 2^{-1} \left[ (2c_{2}^{2} \sum_{i_{1}=1}^{i_{1}} \sum_{i_{2}=1}^{i_{2}} e_{i_{1}} e_{i_{2}} + c_{2}^{2} \sum_{i=1}^{n} \sum_{i=1}^{n} e_{i}^{2} ) - \sum_{k=0}^{n} \sum_{i=1}^{n} \sum_{i=1}^{n} e_{i}^{2} \right] + O_p(n^{1/2}).
\]
\[ (\varphi_S) \text{Var} \]

Also,

\[ z_I = (\varphi^{+1} + z - f) (z - f) 0 = z I = f I = 1 I = f I = 1 \]

Now observe that

\[ (1)^d O + \varphi^{+1} z 0 = z I = f I = 1 I = f I = 1 \]

\[ (1)^d O + \varphi^{+1} (z - f) 0 = z I = f I = 1 I = f I = 1 \]

\[ \Delta \text{Sec. } (1)^d O + z_S + I_S = \]

\[ (1)^d O + \varphi^{+1} (z - f) 0 = z I = f I = 1 I = f I = 1 \]
Therefore, \((8.21)\) combined with \((8.5)\) yields

\[ (u)O = (u)O + (u)O + (u)O = \]

\[
\begin{aligned}
&= I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} + s_{\varphi} Y \frac{3}{u} + \frac{I}{u} I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} \\
&= I^{-1} + I^{-1} = I + I = I + I
\end{aligned}
\]

\[
\begin{aligned}
&= I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} + s_{\varphi} Y \frac{3}{u} + \frac{I}{u} I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} \\
&= I^{-1} + I^{-1} = I + I = I + I
\end{aligned}
\]

\[
\begin{aligned}
&= I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} + s_{\varphi} Y \frac{3}{u} + \frac{I}{u} I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} \\
&= I^{-1} + I^{-1} = I + I = I + I
\end{aligned}
\]

\[
\begin{aligned}
&= I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} + s_{\varphi} Y \frac{3}{u} + \frac{I}{u} I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} \\
&= I^{-1} + I^{-1} = I + I = I + I
\end{aligned}
\]

\[
\begin{aligned}
&= I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} + s_{\varphi} Y \frac{3}{u} + \frac{I}{u} I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} \\
&= I^{-1} + I^{-1} = I + I = I + I
\end{aligned}
\]

Because

\[
\begin{aligned}
&= I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} + s_{\varphi} Y \frac{3}{u} + \frac{I}{u} I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} \\
&= I^{-1} + I^{-1} = I + I = I + I
\end{aligned}
\]

\[
\begin{aligned}
&= I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} + s_{\varphi} Y \frac{3}{u} + \frac{I}{u} I^{-1}(\varphi Y - 1) I_{-}(\varphi Y - 1) \frac{I}{u} \\
&= I^{-1} + I^{-1} = I + I = I + I
\end{aligned}
\]
\[ S_2 = O_p(n^{1/2}) \] (8.22)

Hence from (8.19), (8.20), and (8.22) we have the desired result (4.15).

8.3.4. Verification of (4.16)

\[
\sum_{j=1}^{n} \sum_{k=j}^{n} z_{k} z_{j} \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1}^O d_{k_2}^O
\]

\[ = \sum_{k=1}^{n} \sum_{j=1}^{k-1} z_{k} z_{j} \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1}^O d_{k_2}^O \]

\[ = \sum_{k=1}^{n} \sum_{j=1}^{k-1} z_{k} z_{j} \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1}^O d_{k_2}^O + \sum_{k=1}^{n} z_{k}^2 \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1}^O d_{k_2}^O \]

\[ = S_3 + S_4, \quad \text{say.} \] (8.23)

By the same argument applied to (4.15), we have

\[ S_3 = \sum_{j=1}^{n} \sum_{k=0}^{j-1} v_{k} v_{j} \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1}^O d_{k_2}^O + \sum_{i=1}^{n} e_{i}^2 + O_p(n^{1/2}) \] (8.24)

Now, from (8.18),

\[ S_4 = \sum_{k=0}^{\infty} v_{k}^2 \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} d_{k_1}^O d_{k_2}^O + \sum_{i=1}^{n} e_{i}^2 + O_p(n^{1/2}). \] (8.25)

Combining (8.23), (8.24), and (8.25), we have the desired result (4.16).
8.3.5. Verification of (4.26)

By the same argument applied to (3.39),
\[
Y_1^\prime D_n^0 D_n^0 Y_1 = (d_0)^2 \sum_{j=1}^{n} \sum_{k=1}^{j-1} z_k y_{j-1} + \sum_{j=1}^{n} \sum_{k=1}^{j-1} b_{n,k,j} z_k y_{j-1}
\]
\[
+ \sum_{j=1}^{n} \sum_{k=j}^{n} c_{n,k,j} z_k y_{j-1} + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{s=0}^{n-\max(i,j)} y_{j-1} d_s |i-j| + s
\]

where \(b_{n,k,j}\) and \(c_{n,k,j}\) are defined in (3.40) and evaluated at \(\theta = \theta^0\). By (3.41) and Lemma 3.10-iv), and Lemma 3.10-v), we get (4.26).

8.3.6. Verification of (4.46).

First we show that for all \(k, h = 1, ..., (p+q)\),
\[
Z^\prime \left( \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^0 \right)^\prime D_n^0 Z = O_p(n^{1/2}).
\]

Fix \(k\) and \(h\) in \(\{1, 2, ..., p+q\}\). Let \(g_j(\theta) = \frac{\partial^2}{\partial \theta_k \partial \theta_h} d_j(\theta)\) and \(g_j^0 = g_j(\theta^0)\), \(j = 0, 1, \ldots\). Then \(g_0 = 0\) because \(d_0 = 1\). Since \(g_0 = 0\), the \(t\)-th element of \(\left( \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^0 \right)^\prime D_n^0 Z\)

is \(\sum_{j=1}^{t-1} g_{t-j}^0 z_j\). Also the \(t\)-th element of \(D_n^0 Z\) is \(\sum_{j=1}^{t} d_{t-j}^0 z_j\). Also observing
we have

\[ Z' \left( \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^O \right) D_n^O Z = \sum_{t=1}^{n} \left( \sum_{j=1}^{t-1} e_t e_{t-j} \right) - \sum_{t=1}^{n} \left( \sum_{j=1}^{t-1} e_t e_{t-j} \right) \]

\[ = S_1 - S_2, \text{ say.} \]

It is easy to show

\[ \text{E}(S_1) = 0 \text{ and } \text{Var}(S_1) = O(n). \]

Therefore \( S_1 = O_p(n^{1/2}) \). Also it is easy to show \( S_2 = O_p(n^{1/2}) \). Hence we conclude (8.26). Next if we show

\[ || \left\{ \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^O \right\} D_n^O (\tilde{\theta}) || - \left( \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^O \right)^O D_n^O || = o_p(1) \]

(8.27)

we can say (4.46)

\[ Z' \left\{ \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^O \right\} D_n^O (\tilde{\theta}) Z = o_p(n) \]

because

\[ |Z' \left\{ \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^O \right\} D_n^O (\tilde{\theta}) Z| \]

\[ \leq |Z|^2 || \left\{ \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^O \right\} D_n^O - \left( \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^O \right)^O D_n^O || + |Z' \left\{ \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^O \right\} D_n^O Z| \]

\[ = o_p(n) o_p(1) + O_p(n^{1/2}) = o_p(n). \]
Now we show (8.27). Let \( g_{ij} = g_{ij}(\theta) \) be the (i,j) element of \( \left( \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n \right)' D_n \) and 
\( g_{\theta,ij} \) be the first order partial derivative of \( g_{ij} \) with respect to \( \theta \). Then 

\[
\{g_{ij}(\tilde{\theta}) - g_{ij}(\theta^*)\} = g_{\theta,ij}(\theta^*)(\tilde{\theta} - \theta^*),
\]

for \( \theta^* \) between \( \tilde{\theta} \) and \( \theta^* \). Since, by Corollary 2.5, the derivatives of \( D_n \) enjoy the same uniformly exponentially declining property as \( D_n \), the derivative of 

\( \left( \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n \right)' D_n \) has uniformly exponentially declining off diagonal element.

Therefore, for some \( M < \infty \) and \( \lambda \in (0,1) \), 

\[
|g_{ij}(\tilde{\theta}) - g_{ij}(\theta^*)| = M\lambda |^{i-j} |\tilde{\theta} - \theta^*|.
\]

Let \( \nu \) be any eigenvalue of \( \left\{ \left( \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n \right)' D_n \right\} \) (Theorem 2.9). Now by 

Gershgorin's theorem (Theorem 2.9), for some \( j \), 

\[
|\nu| \leq 2^{\sum_{i=1}^{\infty}} |g_{ij}(\tilde{\theta}) - g_{ij}(\theta^*)| \leq 2M(1-\lambda)^{-1} |\tilde{\theta} - \theta^*|.
\]

Hence 

\[
\|\left\{ \left( \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n \right)' D_n \right\} \| - \left( \frac{\partial^2}{\partial \theta_k \partial \theta_h} D_n^0 \right)' D_n^0 \| \leq 2M(1-\lambda)^{-1} |\tilde{\theta} - \theta^*| = o_p(1)
\]

establishing (8.27).
8.4. Appendix for Chapter 5

8.4.1. Verification of (5.18) and (5.19)

We get (5.18) from the following observation

\[ \sup_{\theta} \left| \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} a_{j-k} z_k \right| \leq M \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \lambda^{j-k} |z_k| = O(n). \]

We get (5.19) from the following observation

\[ \sup_{\theta} \left| \sum_{k=1}^{n-1} \sum_{j=1}^{k} a_{k+1-j} z_k \right| \leq M \sum_{k=1}^{n-1} \sum_{j=1}^{k} \lambda^{k+1-j} |z_k| = O(n). \]

8.4.2. Verification of \( y_j = c_0 \sum_{j=1}^{n-1} W_{j-1} + O_p(n) \)

\[ \sum_{j=1}^{n-1} y_j = \sum_{j=1}^{n-1} (y_0 + \sum_{s=1}^{j} z_s) = (n-1)y_0 + \sum_{j=1}^{n-1} \sum_{s=1}^{j} v_{s-1}^o e_i \]

\[ = (n-1)y_0 + \sum_{i \leq n-1} \sum_{j=i}^{n-1} \sum_{s=1}^{j} v_{s-1}^o e_i \]

\[ = (n-1)y_0 + \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} \sum_{s=i}^{\max(1,i)} \sum_{s=1}^{\max(1,i)} v_{s-1}^o e_i \]

\[ = (n-1)y_0 + \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} \sum_{s=i}^{\max(1,i)} v_{s-1}^o e_i \]

\[ = (n-1)y_0 + \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} \sum_{s=i}^{\max(1,i)} v_{s-1}^o e_i \]

\[ = (n-1)y_0 + \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} v_{s-i}^o e_i + \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} v_{s-i}^o e_i \]

\[ = (n-1)y_0 + \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} v_{s-i}^o e_i + \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} v_{s-i}^o e_i \]
\[ n-1 \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} e_i = c_0 \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} e_i = c_0 \sum_{j=1}^{n} W_{j-1}, \]  

(8.29)

Note that

\[ E|S_2| \leq \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} |v_{s-i}^{0}| E|e_1| \leq M E|e_1| \sum_{i=1}^{n-1} \sum_{j=i}^{n-1} \lambda^{s-i} (1-\lambda)^{-1} = O(n), \]  

(8.30)

and

\[ E|S_3| \leq M \sum_{i=0}^{n-1} \sum_{j=1}^{n-1} \sum_{s=1}^{n-1} \lambda^{s-i} E|e_1| = O(n). \]  

(8.31)

Therefore, from (8.28) — (8.31), we get (5.33).

8.4.3. Verification of (5.31)

Note that

\[ E \sum_{j=1}^{n-1} \lambda^{n-j} y_j \leq E \sum_{j=1}^{n-1} \lambda^{n-j} (|y_0| + \sum_{s=1}^{j} |z_s|) \]

\[ \leq \sum_{j=1}^{n-1} \lambda^{n-j} E|y_0| + E|Z_1| \sum_{j=1}^{n-1} j \lambda^{n-j} = O(n). \]  

(8.32)

Also by (3.5)

\[ \sum_{j=1}^{n-1} \lambda^{j} y_j = O_p(1). \]  

(8.33)
\[ I = \left[ 0 \leq |u_{\theta} - u_{\theta}| = (\nu \cdot \nu) \right]_{\nu} \]

Therefore,

\[ (\nu / I)^{d_{0}} = (\nu / I)^{d_{0}} + (\nu (M - I-M) \frac{I}{I-I})^{d_{0}} / (u_{\theta} - \frac{1}{I}) \]

\[ (\nu / I)^{d_{0}} + (\nu (M - I-M) \frac{I}{I-I})^{d_{0}} / (u_{\theta} - \frac{1}{I}) \]

\[ \varepsilon \left( (M - I-M) \frac{I}{I-I} \nu \nu - |u_{\theta}| \nu \right) = \varepsilon (M - I-M) \frac{I}{I-I} \nu \nu + \]

\[ (u_{\theta} - |I-M \frac{I}{I-I} \nu \nu = u_{\theta} - u_{\theta}) \nu \nu = \varepsilon (M - I-M) \frac{I}{I-I} \nu \nu + \]

\[ \left( \nu_{\theta} - |I-M \frac{I}{I-I} \nu \nu = \varepsilon (M - I-M) \frac{I}{I-I} \nu \nu + \right) \]

\[ \left( \nu_{\theta} - |I-M \frac{I}{I-I} \nu \nu = \varepsilon (M - I-M) \frac{I}{I-I} \nu \nu + \right) \]

\[ \text{Note that} \]

\[ (6.30) \text{, (6.33), and (6.33) we get (6.31)}. \]

\[ 231 \]
Without loss of generality we can assume that \( M \) and \( \lambda \) are the coefficients of exponential decline of \( \{v_j^o\} \). Therefore,

\[
| \sum_{i=1}^{n} \sum_{j_2=k}^{n-1} a_{n,i,j_2} v_{j_2-k} e_k | \leq M^2 \sum_{j_2=k}^{n} \lambda^{i_1-j_2+j_2-k} \leq M_1 \text{ for some } M_1 < \infty.
\]

Therefore

\[
E(S_1) = 0
\]

and

\[
\text{Var}(S_1) \leq M_1^2 \sum_{k=1}^{n-1} E(e_k^2) = O(n).
\]

Hence

\[
S_1 = o_p(n^{1/2}). \quad (8.35)
\]
Also

\[ E|S_2| \leq M^2 \sum \sum_{i_1=1}^{n-1} \lambda^{i_1} E|e_1| \]

\[ \leq M^2 (1-\lambda)^{-1} \sum_{j_2=1}^{n} \sum_{i_1=1}^{n-1} \lambda^{i_1} E|e_1| \]

\[ \leq M^2 (1-\lambda)^{-1} \sum_{j_2=1}^{n} \left\{ \sum_{i_1=1}^{n} \lambda^{j_2} E|e_1| \right\} \]

\[ = M^2 (1-\lambda)^{-1} \left\{ \sum_{j_2=1}^{n} j_2 \lambda^{j_2} + (1-\lambda)^{-1} \sum_{j_2=1}^{n} \lambda^{j_2} \right\} E|e_1| = O(1). \]

Therefore

\[ S_2 = O_p(1). \quad (8.36) \]

Hence, from (8.34) – (8.36), we conclude

\[ S = O_p(n^{1/2}). \]
9. BIBLIOGRAPHY


10. ACKNOWLEDGEMENTS

I wish to express my appreciation to Doctor Wayne A. Fuller for his expert and patient direction of my Ph. D. study. His insightful suggestions and direction have been very helpful to the preparation of this dissertation.

I also wish to express my thanks to my mother, wife, and brothers.