Let \( \{X_t\} \) be a Gaussian ARMA process with spectral density \( f_g(\lambda) \), where \( \theta \) is an unknown parameter. To estimate \( \theta \) we propose a minimum contrast estimation method which includes the maximum likelihood method and the quasi-maximum likelihood method as special cases. Let \( \hat{\theta}_T \) be the minimum contrast esti-
mator of $\theta$. Then we derive the Edgeworth expansion of the dis-
tribution of $\hat{\theta}_T$ up to third order, and prove its validity. By
this Edgeworth expansion we can see that this minimum contrast
estimator is always second-order asymptotically efficient in the
class of second-order asymptotically median unbiased estimators.
Also the third-order asymptotic comparisons among minimum con-
trast estimators will be discussed.
VALIDITY OF EDGECROTH EXPANSIONS
OF MINIMUM CONTRAST ESTIMATORS
FOR GAUSSIAN ARMA PROCESSES

By Masanobu Taniguchi
University of Pittsburgh
and
Hiroshima University

Center for Multivariate Analysis
University of Pittsburgh

Approved for public release; distribution unlimited.
VALIDITY OF EDGEWORTH EXPANSIONS OF MINIMUM CONTRAST ESTIMATORS FOR GAUSSIAN ARMA PROCESSES

By Masanobu Taniguchi
University of Pittsburgh
and
Hiroshima University

December 1985

Technical Report No. 85-49

Center for Multivariate Analysis
Fifth Floor, Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

This work is supported by Contract F49620-85-C-0008 of the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is approved for public release IAW AFR 190-12. Distribution is unlimited.

THEW J. KENPER
Chief, Technical Information Division
Abstract

Let \( \{X_t\} \) be a Gaussian ARMA process with spectral density \( f_\theta(\lambda) \), where \( \theta \) is an unknown parameter. To estimate \( \theta \) we propose a minimum contrast estimation method which includes the maximum likelihood method and the quasi-maximum likelihood method as special cases. Let \( \hat{\theta}_T \) be the minimum contrast estimator of \( \theta \). Then we derive the Edgeworth expansion of the distribution of \( \hat{\theta}_T \) up to third order, and prove its validity. By this Edgeworth expansion we can see that this minimum contrast estimator is always second-order asymptotically efficient in the class of second-order asymptotically median unbiased estimators. Also the third-order asymptotic comparisons among minimum contrast estimators will be discussed.
1. Introduction

Recently some systematic studies of higher order asymptotic theory for stationary processes have been developed. In many cases such studies have used the formal Edgeworth expansions. Thus it has been required to prove their validities. Götze and Hipp (1983) showed that formal Edgeworth expansions are valid for sums of weakly dependent vectors. Durbin (1980) and Tani-guchi (1984) showed the validity of Edgeworth expansions of statistics derived from observations which are not necessarily independent and identically distributed. However their sufficient conditions for the validity are hard to check even in the fundamental statistics.

In this paper we propose a minimum contrast estimation method which includes the maximum likelihood method and the quasi-maximum likelihood method as special cases. Suppose that \((X_t)\) is a Gaussian ARMA process with spectral density \(f_\theta(\lambda)\), where \(\theta\) is an unknown parameter. Let \(\hat{\theta}_T\) be the minimum contrast estimator of \(\theta\). Then we give the Edgeworth expansion of the distribution of \(\hat{\theta}_T\) up to third order, and prove its validity. That is, as special cases we get the valid Edgeworth expansions for the maximum likelihood estimator and the quasi-maximum likelihood estimator which is defined by the value minimizing \(\int_{-\pi}^{\pi} \{\log f_\theta(\lambda) + I_T(\lambda)/f_\theta(\lambda)\} d\lambda\) with respect to \(\theta\), where \(I_T(\lambda)\) is the periodogram.

In Section 7 we consider the transformed statistic \(\hat{\theta}_m = \hat{\theta}_T + \frac{1}{T} m(\hat{\theta}_T)\), where \(m(\cdot)\) is a smooth function. Then we give the valid Edgeworth expansion for \(\hat{\theta}_m\). By this Edgeworth expansion
we can see that our minimum contrast estimator is always second-order asymptotically efficient in the class of second order asymptotically median unbiased estimators if efficiency is measured by the degree of concentration of the sampling distribution up to second order. Also the third-order asymptotic comparisons among minimum contrast estimators will be given.
2. Minimum Contrast Estimator

We propose a minimum contrast estimator which includes the maximum likelihood estimator and the quasi-maximum likelihood estimator as special cases.

Let $D_d$ and $D^C_{\text{ARMA}}$ be spaces of functions on $[-\pi, \pi]$ defined by

$$D_d = \{ f: f(\lambda) = \sum_{u=-\infty}^{\infty} a(u)e^{iu\lambda}, a(u) = a(-u), \quad \sum_{u=-\infty}^{\infty} (1 + |u|)|a(u)| < d, \text{ for some } d < \infty \},$$

$$D^C_{\text{ARMA}} = \{ f: f(\lambda) = \frac{2}{\sigma^2} \left| \sum_{j=0}^{p} a_j e^{ij\lambda} \right|^{2}, (\sigma^2 > 0), \quad \sum_{j=0}^{p} b_j e^{ij\lambda} \leq c, \text{ for } |z| \leq 1, \quad 0 < c < \overline{c} < \infty \}.$$

We set down the following assumptions.

Assumption 1. \( \{ X_t \} \) is a Gaussian stationary process with the spectral density \( f_{\theta}(\lambda) \in D^C_{\text{ARMA}}, \theta_0 \in C \subset \Theta \subset \mathbb{R}^1, \) and mean 0. 
Here \( \Theta \) is an open set of \( \mathbb{R}^1 \) and \( C \) is a compact subset of \( \Theta \).

Assumption 2. The spectral density \( f_{\theta}(\lambda) \) is continuously five times differentiable with respect to \( \theta \in \Theta \), and the derivatives
\[ \frac{\partial f_\theta}{\partial \theta}, \frac{\partial^2 f_\theta}{\partial \theta^2}, \frac{\partial^3 f_\theta}{\partial \theta^3}, \frac{\partial^4 f_\theta}{\partial \theta^4} \text{ and } \frac{\partial^5 f_\theta}{\partial \theta^5} \text{ belong to } D_d. \]

**Assumption 3.** There exists \( d_1 > 0 \) such that

\[
I(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left( \frac{\partial}{\partial \theta} \log f_\theta(\lambda) \right)^2 d\lambda \geq d_1 > 0, \text{ for all } \theta \text{ in } \Theta.
\]

Suppose that a stretch \( X_T = (X_1, \ldots, X_T)' \) of the series \( \{X_t\} \) is available. Let \( \Sigma_T = \Sigma_T(\Theta_0) \) be the covariance matrix of \( X_T \). The \((m,n)\)th element of \( \Sigma_T \) is given by \( \int_{-\pi}^{\pi} \exp[i(m-n)\lambda] f_\theta(\lambda) d\lambda \).

Let \( A_T(\theta) \) and \( B_T(\theta) \) be \( T \times T \)-Toeplitz matrices associated with harmonic functions \( g_\theta(\lambda) \) and \( h_\theta(\lambda) \), where \( g_\theta \in \mathcal{D}^C_{\text{ARMA}}, \ h_\theta \in \mathcal{D}_d \) (i.e., the \((m,n)\)th element of \( A_T(\theta) \) and \( B_T(\theta) \) are given by \( \int_{-\pi}^{\pi} \exp[i(m-n)\lambda] g_\theta(\lambda) d\lambda \) and \( \int_{-\pi}^{\pi} \exp[i(m-n)\lambda] h_\theta(\lambda) d\lambda \), respectively.

We impose the following assumptions.

**Assumption 4.** The functions \( g_\theta \) and \( h_\theta \) are continuously four times differentiable with respect to \( \theta \in \Theta \), and the derivatives \( \frac{\partial g_\theta}{\partial \theta}, \ldots, \frac{\partial^4 g_\theta}{\partial \theta^4} \), \( \frac{\partial h_\theta}{\partial \theta}, \ldots, \frac{\partial^4 h_\theta}{\partial \theta^4} \), \( \theta \in \Theta \), belong to \( D_d \). Also \( g_\theta \) and \( h_\theta \) satisfy

\[
g_\theta(\lambda)^{-2} h_\theta(\lambda) = \frac{1}{2} f_\theta(\lambda)^{-2} \frac{\partial}{\partial \theta} f_\theta(\lambda). \tag{2.1}
\]

**Assumption 5.** A function \( b_T(\theta) \) is four times continuously differentiable with respect to \( \theta \), and is written as

\[
b_T(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} f_\theta(\lambda)^{-1} \frac{\partial}{\partial \theta} f_\theta(\lambda) d\lambda + O(T^{-1}).
\]

Now consider the following equation;

\[
\frac{1}{T} \sum_{t=1}^{T} A_T(\theta)^{-1} b_T(\theta) A_T(\theta)^{-1} X_t = b_T(\theta), \quad \theta \in \Theta. \tag{2.2}
\]
A minimum contrast estimator $\hat{\theta}_T$ of $\theta_0$ is defined by a value of $\theta$ that satisfies the equation (2.2). This estimator $\hat{\theta}_T$ includes the following cases:

**Example 1.** Put $g_\theta = f_\theta$, $h_\theta = \frac{1}{2} \frac{\partial f_\theta}{\partial \theta}$ and $b_T(\theta) = \frac{1}{2T} \text{tr} \Sigma^{-1} \frac{\partial}{\partial \theta} \Sigma$, then by Theorem 1 in Taniguchi (1983)

$$b_T(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} f_\theta^{-1} \frac{\partial}{\partial \theta} f_\theta d\lambda + O(T^{-1}).$$

The estimator $\hat{\theta}_T$ becomes the maximum likelihood estimator (see Taniguchi (1983) or (1985)).

**Example 2.** Put $g_\theta = \frac{1}{2\pi}$, $h_\theta = \frac{1}{8\pi^2} \frac{\partial f_\theta}{\partial \theta} f_\theta^{-2}$ and

$$b_T(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} f_\theta^{-1} \frac{\partial}{\partial \theta} f_\theta d\lambda.$$ Then (2.2) is written as

$$X_T^T \left( \begin{array}{c} m \\
\vdots \\
m-n \end{array} \right) \int_{-\pi}^{\pi} e^{i(m-n)\lambda} \frac{1}{2\pi} \frac{\partial}{\partial \theta} f_\theta f_\theta^{-2} d\lambda = X_T,$$

$$= \frac{1}{4\pi} \int_{-\pi}^{\pi} f_\theta^{-1} \frac{\partial}{\partial \theta} f_\theta d\lambda.$$

We can see that the equation (2.3) is equivalent to

$$\frac{\partial}{\partial \theta} \int_{-\pi}^{\pi} \left[ \log f_\theta(\lambda) + \frac{I_T(\lambda)}{f_\theta(\lambda)} \right] d\lambda = 0,$$

where $I_T(\lambda) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} X_t e^{i\lambda} \right|^2$. Thus the estimator $\hat{\theta}_T$ becomes the quasi-maximum likelihood estimator (see Dunsmuir and Hannan (1976), Hosoya and Taniguchi (1982)).

At first we present the following basic theorem which is
useful for the higher-order asymptotic theory up to third order in time series analysis.

**Theorem 1.** Assume that Assumptions 1-5 hold. Let $\alpha$ be an arbitrary fixed number such that $0 < \alpha < 3/8$.

(1) There exists a statistic $\hat{\theta}_T$ which solves (2.2) such that for some $d_1 > 0$,

$$P_{\theta_0}^T \left[ |\hat{\theta}_T - \theta_0| < d_1 T^{\alpha - 1/2} \right] = 1 - o(T^{-1}),$$

uniformly for $\theta_0 \in \mathcal{C}$.

(2) For $\{\hat{\theta}_T\}$ satisfying (2.4),

$$\sup_{B \in \mathcal{A}_0} \left| \frac{1}{T} \left[ T(\theta_0) \right]^{1/2} \left( \hat{\theta}_T - \theta_0 \right) \in B \right|$$

$$- \int_B \varnothing(x) p_3^T(x) dx = o(T^{-1}),$$

uniformly for $\theta_0 \in \mathcal{C}$, where $\mathcal{A}_0$ is a class of Borel sets of $\mathbb{R}^1$ satisfying

$$\sup_{B \in \mathcal{A}_0} \int_{(\delta B)^C} \varnothing(x) p_3^T(x) dx = o(\varepsilon).$$

(2.5)

Here $\varnothing(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, and $p_3^T(x) = 1 + \frac{q(x)}{\sqrt{T}} + \frac{r(x)}{T}$ where $q(x)$ and $r(x)$ are polynomials.

In Section 6 we shall give the coefficients of $q(x)$ and $r(x)$ by using the spectral density $f_\theta$. 
3. Stochastic expansion of minimum contrast estimator.

In this section we derive a stochastic expansion of \( \hat{\theta}_T \).

We set down
\[
\ell_T(\theta) = X_T^\top H_T(\theta) X_T - T \beta_T(\theta),
\]
where \( H_T(\theta) = A_T(\theta)^{-1} B_T(\theta) A_T(\theta)^{-1} \).

Let
\[
Z_1(\theta) = \frac{1}{\sqrt{T}} (X_T^\top H_T(\theta) X_T - T \beta_T(\theta)),
\]
\[
Z_2(\theta) = \frac{1}{\sqrt{T}} (X_T^\top \tilde{H}_T(\theta) X_T - \text{tr} \Sigma_T(\theta) \hat{\theta}_T(\theta)),
\]
\[
Z_3(\theta) = \frac{1}{\sqrt{T}} (X_T^\top \tilde{H}_T(\theta) X_T - \text{tr} \Sigma_T(\theta) \hat{\theta}_T(\theta)),
\]
where \( \tilde{H}_T(\theta) = \frac{\partial}{\partial \theta} H_T(\theta) \) and \( \tilde{H}_T(\theta) = \frac{\partial^2}{\partial \theta^2} H_T(\theta) \). Henceforth, for simplicity, we sometimes use \( A, B, H, \Sigma, Z_1, Z_2 \) and \( Z_3 \) instead of \( A_T(\theta), B_T(\theta), H_T(\theta), \Sigma_T(\theta), Z_1(\theta), Z_2(\theta) \) and \( Z_3(\theta) \), respectively.

It is easy to show that
\[
\hat{\theta} = -A^{-1} \hat{\theta} A^{-1} - A^{-1} B A^{-1} \hat{\theta} A^{-1} + A^{-1} B A^{-1},
\]
\[
\tilde{H} = A^{-1} \tilde{H} A^{-1} + A^{-1} B A^{-1} \tilde{H} A^{-1} - A^{-1} \tilde{H} A^{-1} - A^{-1} B A^{-1} \tilde{H} A^{-1} + A^{-1} \tilde{H} A^{-1} + A^{-1} B A^{-1} \tilde{H} A^{-1} + 2A^{-1} \tilde{H} A^{-1} - A^{-1} \tilde{H} A^{-1} - A^{-1} B A^{-1} \tilde{H} A^{-1} + A^{-1} B A^{-1}.
\]
Since the minimum contrast estimator is approximated by simple functions of \( Z_1, Z_2 \) and \( Z_3 \). To give the asymptotic expansion, we must evaluate the asymptotic cumulants (moments) of \( Z_1, Z_2 \) and \( Z_3 \). The following lemma is useful to evaluate them (see Taniguchi (1983)).

**Lemma 1.** Suppose that \( f_1(\lambda), \ldots, f_s(\lambda) \in D_{\text{d}}, g_1(\lambda), \ldots, g_s(\lambda) \in D_{\text{ARMA}} \). We define \( \Gamma_1, \ldots, \Gamma_s, \Lambda_1, \ldots, \Lambda_s \), the TXT-Toeplitz type matrices associated with \( f_1(\lambda), \ldots, f_s(\lambda), g_1(\lambda), \ldots, g_s(\lambda) \), respectively. Then

\[
T^{-1} \text{tr} \Gamma_1^{-1} \Lambda_1^{-1} \Gamma_2^{-1} \Lambda_2^{-1} \ldots \Gamma_s^{-1} \Lambda_s^{-1}
\]

\[
= \frac{1}{2\pi} \int_{-\pi}^{\pi} f_1(\lambda) \ldots f_s(\lambda) g_1(\lambda)^{-1} \ldots g_s(\lambda)^{-1} d\lambda + O(T^{-1}).
\]

We write

\[
E(\theta) = \frac{\mu(\theta)}{\sqrt{T}} + o(T^{-1}). \tag{3.6}
\]

Here \( \mu(\theta) \) will be evaluated explicitly for some cases in Section 7. Using Lemma 1 and (2.1), it is not difficult to show the following lemma.

**Lemma 2.** Under Assumptions 1-5, we have

\[
E(\theta) = \frac{1}{\sqrt{T}} + o(T^{-1}), \tag{3.7}
\]

\[
E(\theta) = \frac{2}{\sqrt{T}} u(\theta) + o(T^{-1}), \tag{3.8}
\]

\[
E(\theta) = \frac{3}{\sqrt{T}} \mu(\theta) + o(T^{-3/2}). \tag{3.9}
\]
\begin{align}
E_\theta[Z_1(\theta)Z_3(\theta)] &= L(\theta) + O(T^{-1}), \quad (3.10) \\
E_\theta[Z_2(\theta)]^2 &= M(\theta) + O(T^{-1}), \quad (3.11) \\
E_\theta[Z_1(\theta)^2Z_2(\theta)] &= \frac{1}{\sqrt{T}} N(\theta) + \frac{2}{\sqrt{T}} J(\theta) \mu(\theta) + O(T^{-3/2}), \quad (3.12) \\
\text{cum}_\theta[Z_1(\theta), Z_1(\theta), Z_1(\theta), Z_1(\theta)] &= \frac{1}{T} H(\theta) + O(T^{-2}), \quad (3.13) \\
E_\theta\left[\frac{1}{T} \frac{\partial}{\partial \theta} \ell_T(\theta)\right] &= - I(\theta) + O(T^{-1}), \quad (3.14) \\
E_\theta\left[\frac{1}{T} \frac{\partial^2}{\partial \theta^2} \ell_T(\theta)\right] &= - 3J(\theta) - K(\theta) + O(T^{-1}), \quad (3.15) \\
E_\theta\left[\frac{1}{T} \frac{\partial^3}{\partial \theta^3} \ell_T(\theta)\right] &= - 4L(\theta) - 3M(\theta) - 6N(\theta) - H(\theta) + O(T^{-1}) \quad (3.16)
\end{align}

where

\begin{align*}
J(\theta) &= - \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\partial f_\theta(\lambda)}{\partial \theta}\right]^3 f_\theta(\lambda)^{-3} d\lambda \\
&\quad + \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[\frac{\partial f_\theta(\lambda)}{\partial \theta}\right]\left[\frac{\partial^2 f_\theta(\lambda)}{\partial \theta^2}\right]^2 f_\theta(\lambda)^{-2} d\lambda, \\
K(\theta) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\partial f_\theta(\lambda)}{\partial \theta}\right]^3 f_\theta(\lambda)^{-3} d\lambda, \\
L(\theta) &= \frac{3}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\partial f_\theta(\lambda)}{\partial \theta}\right]^4 f_\theta(\lambda)^{-4} d\lambda + \frac{3}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\partial f_\theta(\lambda)}{\partial \theta}\right]^2 \left[\frac{\partial^2 f_\theta(\lambda)}{\partial \theta^2}\right] f_\theta(\lambda)^{-3} d\lambda \\
&\quad + \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[\frac{\partial f_\theta(\lambda)}{\partial \theta}\right]\left[\frac{\partial^3 f_\theta(\lambda)}{\partial \theta^3}\right]^2 f_\theta(\lambda)^{-2} d\lambda, \\
M(\theta) &= \frac{1}{T} \int_{-\pi}^{\pi} \left[\frac{\partial f_\theta(\lambda)}{\partial \theta}\right]^4 f_\theta(\lambda)^{-4} d\lambda - \frac{1}{T} \int_{-\pi}^{\pi} \left[\frac{\partial f_\theta(\lambda)}{\partial \theta}\right]^2 \left[\frac{\partial^2 f_\theta(\lambda)}{\partial \theta^2}\right] f_\theta(\lambda)^{-3} d\lambda
\end{align*}
\[ N(\theta) = - \frac{1}{2\pi} \int_{-\pi}^{\pi} \{ \frac{\partial^2}{\partial \theta^2} f_\theta(\lambda) \}^2 f_\theta(\lambda)^4 d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} \{ \frac{\partial}{\partial \theta} f_\theta(\lambda) \}^2 \{ \frac{\partial^2}{\partial \theta^2} f_\theta(\lambda) \} f_\theta(\lambda)^3 d\lambda, \]

\[ H(\theta) = \frac{3}{2\pi} \int_{-\pi}^{\pi} \{ \frac{\partial}{\partial \theta} f_\theta(\lambda) \}^4 f_\theta(\lambda)^4 d\lambda. \]

Put \( E_T = A_1^{-1} \Gamma_1 A_2^{-1} \cdots \Gamma_s A_s^{-1} \), where \( \Gamma_1, \ldots, \Gamma_{s-1}, A_1, \ldots, A_s \) are TxT-Toeplitz type matrices associated with some harmonic functions \( u_\theta^{(1)}(\lambda) \in D_d, \ldots, u_\theta^{(s-1)}(\lambda) \in D_d, v_\theta^{(1)}(\lambda) \in \mathcal{D}_{ARMA}, \ldots, v_\theta^{(s)}(\lambda) \in \mathcal{D}_{ARMA} \), respectively.

**Lemma 3.** Under Assumption 1, for every \( \delta > 0 \), and some \( d_2 > 0 \), we have

\[ P_\delta \left[ \frac{1}{\sqrt{T}} |X_T^{IB}X_T - E(\theta X_T^{IB}X_T) | > d_2 T^\delta \right] = o(T^{-1}), \tag{3.17} \]

uniformly for \( \theta \in \Theta \).

**Proof** Choose an integer \( \eta \geq 1 \) so that \( 2\eta \delta > 1 \). By Tchebychev's inequality, we have

\[ P_\delta \left[ \frac{1}{\sqrt{T}} |X_T^{IB}X_T - E(\theta X_T^{IB}X_T) | > d_2 T^\delta \right] \leq E_\delta [X_T^{IB}X_T - E(\theta X_T^{IB}X_T)]^2 \eta / (d_2 T^\delta)^{2\eta}. \tag{3.18} \]

Since \( E_\delta [X_T^{IB}X_T - E(\theta X_T^{IB}X_T)]^{2\eta} = O(1) \) (see Lemma 4 of Taniguchi (1985)), (3.18) implies (3.17).

**Lemma 4.** Let \( Y_T \) be a random variable which has the stochastic
expansion

\[ Y_T = Y_T^{(3)} + T^{-3/2} \xi_T, \quad (3.19) \]

where the distribution of \( Y_T^{(3)} \) has the following Edgeworth expansion:

\[ P[Y_T^{(3)} \in B] = \int_B \phi(x)p_3^T(x)dx + o(T^{-1}), \quad (3.20) \]

where \( B \) is a Borel set of \( \mathbb{R}^1 \) satisfying (2.6). Also \( \xi_T \) satisfies

\[ P[|\xi_T| > \rho_T T] = o(T^{-1}), \quad (3.21) \]

where \( \rho_T \to 0 \), \( \rho_T T^{1/2} \to \infty \) as \( T \to \infty \). Then

\[ P[Y_T \in B] = \int_B \phi(x)p_3^T(x)dx + o(T^{-1}), \quad (3.22) \]

for \( B \in \mathcal{G}_o \).

The above proof proceeds on a similar way to Chibisov (1972).

**PROOF OF (1) IN THEOREM 1.**

In this proof we develop the discussion by using the argument similar to that of Bhattacharya and Ghosh (1978) and Taniguchi (1985). Consider the equation

\[ 0 = T^{-1} l_T(\theta_o) + T^{-1}(\theta - \theta_o) \frac{\partial}{\partial \theta} l_T(\theta_o) + (2T)^{-1}(\theta - \theta_o)^2 \frac{\partial^2}{\partial \theta^2} l_T(\theta_o) \]

\[ + (6T)^{-1}(\theta - \theta_o)^3 \frac{\partial^3}{\partial \theta^3} l_T(\theta_o) \]

\[ + R_T(\theta), \quad (3.23) \]
where \( R_T(\theta) \) is the usual remainder in the Taylor expansion, for which it holds that

\[
|R_T(\theta)| \leq \frac{1}{24T^2} |\theta - \theta_0|^{\frac{1}{4}} \sup_{|\theta' - \theta| \leq |\theta - \theta_0|} \left| \frac{\partial^4}{\partial \theta^4} \ell_T(\theta') \right|. \tag{3.24}
\]

In view of Lemma 3, we can see that for every \( \alpha > 0 \) there exist positive constants \( d_3 \) and \( d_4 \) such that

\[
P_{\theta_0}^T \left[ |Z_1(\theta_0)| > d_3 T^\alpha \right] = o(T^{-1}), \tag{3.25}
\]

\[
P_{\theta_0}^T \left[ |Z_2(\theta_0)| > d_3 T^\alpha \right] = o(T^{-1}), \tag{3.26}
\]

\[
P_{\theta_0}^T \left[ |Z_3(\theta_0)| > d_3 T^\alpha \right] = o(T^{-1}), \tag{3.27}
\]

\[
P_{\theta_0}^T \left[ \frac{1}{\sqrt{T}} \left| \frac{\partial^3}{\partial \theta^3} \ell_T(\theta_0) - \mathbb{E}_{\theta_0} \frac{\partial^3}{\partial \theta^3} \ell_T(\theta_0) \right| > d_3 T^\alpha \right] = o(T^{-1}), \tag{3.28}
\]

\[
P_{\theta_0}^T \left[ |R_T(\theta)| > |\theta - \theta_0|^{\frac{1}{4}} (d_4 T^\alpha) \right] = o(T^{-1}). \tag{3.29}
\]

Therefore, on a set having \( P_{\theta_0}^T \)-probability at least 1 - \( o(T^{-1}) \), for some constants \( d_5 > 0 \) and \( d_6 > 0 \) we can rewrite (3.23) as

\[
\theta - \theta_0 = (I(\theta_0) + \eta_T)^{-1} [\delta_T + (2T)^{-1/(2+\alpha)} \mathbb{E}_{\theta_0} \frac{\partial^2}{\partial \theta^2} \ell_T(\theta_0)] + (6T)^{-1/(2+\alpha)} \mathbb{E}_{\theta_0} \frac{\partial^3}{\partial \theta^3} \ell_T(\theta_0) + d_5 |\theta - \theta_0|^{\frac{1}{4}} \zeta_T, \tag{3.30}
\]

where \( \eta_T \) and \( \delta_T \) are random variables whose absolute values are less than \( d_5 T^{-1/2+\alpha} \) and \( \zeta_T \) is a random variable whose absolute value is less than \( d_4 T^\alpha \). There exist a sufficiently large \( d_7 > 0 \) and an integer \( T_0 \) such that if \( T > T_0 \) and \( |\theta - \theta_0| \leq d_7 T^{-1/2+\alpha}, (0 < \alpha < 3/8) \),
the right-hand side of (3.30) is less than \(d_T^{-1/2+\alpha}\). Applying
the Brouwer fixed point theorem to the right-hand side of (3.30)
we have proved (2.4).

Now we set down
\[
V_T = \sqrt{T}(\hat{\theta}_T - \theta_0),
\]
\[
I_T(\theta) = - \frac{1}{T}E_\theta [\frac{\partial}{\partial \theta} \varphi_T(\theta)],
\]
and
\[
U_T(\theta) = \frac{Z_1(\theta)}{I_T(\theta)} + \frac{Z_1(\theta)Z_2(\theta)}{I(\theta)^2 \sqrt{T}} - \frac{3J(\theta) + K(\theta)}{2I(\theta)^3 \sqrt{T}} Z_1(\theta)^2
\]
\[
+ \frac{1}{I(\theta)^3} [Z_1(\theta)Z_2(\theta)^2 + \frac{1}{2}Z_1(\theta)^2Z_3(\theta) + \frac{3}{2} (-3J(\theta) - K(\theta))] Z_1(\theta)^2 Z_2(\theta)
\]
\[
+ \frac{[3J(\theta) + K(\theta)]^2}{2I(\theta)} Z_1(\theta)^3 - \frac{4I(\theta) + 3M(\theta) + 6N(\theta) + H(\theta)}{6I(\theta)} Z_1(\theta)^3].
\]

**Lemma 5.** Under Assumptions 1-5, we have the following stochastic
expansion
\[
\sqrt{T}(\hat{\theta}_T - \theta_0) = U_T(\theta_0) + T^{-3/2} \zeta_T,
\]  
(3.31)

where \(\zeta_T\) satisfies \(Pr_\theta \{ |\zeta_T| > \rho_T \sqrt{T} \} = o(T^{-1})\) for some sequence
\(\rho_T \to 0, \rho_T \sqrt{T} \to \infty \) as \(T \to \infty\).

**Proof**

From the equation \(\ell_T(\hat{\theta}_T) = 0\), we have
\[
0 = Z_1(\theta_0) + T^{-1/2} Z_2(\theta_0) V_T - I_T(\theta_0) V_T
\]
\[
+ \frac{1}{2T^{-3/2}} \left[ \frac{\partial^2}{\partial \theta^2} \ell_T(\theta_0) \right] V_T^2 + \frac{1}{6T^{3/2}} \left[ \frac{\partial^3}{\partial \theta^3} \ell_T(\theta_0) \right] V_T^3
\]
\[
+ \frac{1}{24T^{5/2}} \left[ \frac{\partial^4}{\partial \theta^4} \ell_T(\theta^*) \right] V_T^4.
\]  
(3.32)
where \( |\theta^*-\theta_0| \leq |\hat{\theta}_T-\theta_0| \). We rewrite (3.32) as

\[
V_T = \frac{Z_1(\theta_0)}{T} + \frac{Z_2(\theta_0)V_T}{\sqrt{T}I_T(\theta_0)} + \frac{1}{2\sqrt{T}I_T(\theta_0)} \left( \frac{1}{T} \frac{\partial^2}{\partial \theta^2} \ell_T(\theta_0) \right) V_T^2
\]

\[
+ \frac{1}{6I_T(\theta_0)} \left( \frac{1}{T} \frac{\partial^3}{\partial \theta^3} \ell_T(\theta_0) \right) V_T^3 + \frac{1}{24I_T(\theta_0)T\sqrt{T}} \left( \frac{1}{T} \frac{\partial^4}{\partial \theta^4} \ell_T(\theta_0^*) \right) V_T^4. 
\tag{3.33}
\]

Noting (2.4), (3.25)-(3.29) with \( 0 < \alpha < 1/10 \), we can write (3.33) as

\[
V_T = \frac{Z_1}{T} + \frac{a_T(1)}{\sqrt{T}}, \tag{3.34}
\]

where \( P_{\theta_0}^T [ |a_T(1)| > d_8T^{2\alpha} ] = o(T^{-1}) \), for some \( d_8 > 0 \). Substituting (3.34) for the right-hand side of (3.33), and noting (3.15) we have

\[
V_T = \frac{Z_1}{T} + \frac{Z_2}{\sqrt{T}I^2} - \frac{3J+K}{2\sqrt{T}I^3} Z^2 + \frac{a_T(2)}{T}, \tag{3.35}
\]

where \( P_{\theta_0}^T [ |a_T(2)| > d_9T^{3\alpha} ] = o(T^{-1}) \), for some \( d_9 > 0 \). Again substituting (3.35) for the right-hand side of (3.33), and noting (3.16) we have

\[
V_T = U_T(\theta_0) + \zeta_T/T^{3/2}, \tag{3.36}
\]

where \( P_{\theta_0}^T [ |\zeta_T| > d_{10}T^{5\alpha} ] = o(T^{-1}) \), for some \( d_{10} > 0 \). Since \( 0 < \alpha < 1/10 \), we have the desired result.

**REMARK.** By Lemma 4, the Edgeworth expansion for \( \sqrt{T}(\hat{\theta}_T-\theta_0) \) (up to order \( T^{-1} \)) is equal to that for \( U_T(\theta_0) \) on \( B \in \mathfrak{A}_0 \). Thus we have only to derive the Edgeworth expansion for \( U_T(\theta_0) \).

As we saw in the previous section we have to seek the Edgeworth expansion for $U_T(\theta_0)$. To do so we have to derive the Edgeworth expansion for $Z = (Z_1(\theta), Z_2(\theta), Z_3(\theta))'$. Thus, in this section, we give an asymptotic expansion of the characteristic function of $Z$.

Put

$$\tau(\xi) = E_q\{e^{i\xi'Z}\},$$

where $\xi = (t_1, t_2, t_3)'$. Then it is easy to show

$$\tau(\xi) = \det(I(TxT)) - \frac{2i}{\sqrt{T}} \Sigma^{1/2} (t_1 H + t_2 H + t_3 H) \Sigma^{1/2} - \frac{1}{\sqrt{T}} (t_1 Tb_T(\theta) + t_2 trH\Sigma + t_3 trH\Sigma), \quad (4.1)$$

where $I(TxT)$ is the TxT-identity matrix. Let $\rho_j$ be the $j$th latent root of $S = \Sigma^{1/2} (t_1 S_1 + t_2 S_2 + t_3 S_3) \Sigma^{1/2}$ ($\rho_1^2 \geq \ldots \geq \rho_T^2 \geq 0$). Of course each $\rho_j$ is a real number. Then we have

$$\log \tau(\xi) = -\frac{1}{2} \sum_{j=1}^{T} \log(1 - \frac{2i}{\sqrt{T}} \rho_j)$$

$$- \frac{1}{\sqrt{T}} (t_1 Tb_T(\theta) + t_2 trH\Sigma + t_3 trH\Sigma). \quad (4.2)$$

Notice the relation

$$\log(1 - ih) = -ih + \frac{h^2}{2} + \frac{ih^3}{3} - \frac{h^4}{4} - \frac{ih^5}{5}$$

$$+ h^6 \int_0^1 (1 - v)^5 \frac{dv}{(1 - 1vh)^5}, \quad (4.3)$$
where
\[ \left| \int_0^1 (1 - v)^5 \frac{dv}{(1 - ivh)^5} \right| \leq 1 \]

(e.g., Bhattacharya and Rao (1976, p.57)). By (4.3), the relation (4.2) is

\[
\log \tau(t) = -\frac{1}{2} \sum_{j=1}^T \left[ -\frac{2i\rho j}{\sqrt{T}} + \frac{4\rho j^2}{2T} + \frac{8i\rho j^3}{3T\sqrt{T}} \right]
\]

\[
-\frac{16\rho j^4}{4T^2} - \frac{2i\rho j^5}{5T^{5/2}} + \frac{2^6\rho j^6\gamma_1}{T^3}
\]

\[
- \frac{1}{\sqrt{T}}(t_1 T b_{T}(\theta) + t_2 trH \Sigma + t_3 trH \Sigma), \quad (4.4)
\]

where \(|\gamma_1| \leq 1\). Remembering (3.6), we have

\[
\log \tau(t) = \frac{1}{2} \sum_{j=1}^T \left[ -\frac{2i\rho j}{\sqrt{T}} + \frac{4\rho j^2}{2T} + \frac{8i\rho j^3}{3T\sqrt{T}} \right]
\]

\[
+ \frac{1^2}{T} tr S^2 + \frac{4^3}{T^2} tr S^3 + \frac{16^5}{5T^{5/2}} tr S^5 + R_6, \quad (4.5)
\]

where \(|R_6| \leq \frac{25}{T^3} tr S^6\). Using Lemma 1 we can rewrite as

\[
\frac{2^2}{T} tr S^2 = \sum_{j, k=1}^3 \left( A_{jk} + \frac{B_{jk}}{T} + o(T^{-3/2}) \right) (it_j)(it_k), \quad (4.6)
\]

\[
\frac{8^3}{T} tr S^3 = \sum_{j, k, \ell=1}^3 \left( A_{jk\ell} + o(T^{-1}) \right) (it_j)(it_k')(it_\ell), \quad (4.7)
\]
\[
\frac{481^4}{T}\text{tr}S^4 = \frac{3}{T} \sum_{j,k,l,m=1} [A_{jklm} + O(T^{-1})] \\
x (it_j)(it_k)(it_l)(it_m), \quad (4.8)
\]
\[
\frac{3841^5}{T}\text{tr}S^5 = \frac{3}{T} \sum_{j,k,l,m,n=1} [A_{jklmn} + O(T^{-1})] \\
x (it_j)(it_k)(it_l)(it_m)(it_n), \quad (4.9)
\]
\[
\frac{16}{T}\text{tr}S^6 = \frac{3}{T} \sum_{j,k,l,m,n,p=1} [A_{jklmp} + O(T^{-1})] \\
x (it_j)(it_k)(it_l)(it_m)(it_n)(it_p). \quad (4.10)
\]

For examples we can see that \(A_{11} = I(\theta)\), \(A_{12} = J(\theta)\), \(A_{13} = L(\theta)\), \(A_{22} = M(\theta)\), \(A_{111} = K(\theta)\), \(A_{112} = N(\theta)\), \(A_{1111} = H(\theta)\), e.t.c..

Thus (4.5) is written as

\[
\log \tau(t) = it_1 \left[ \frac{\mu(\theta)}{\sqrt{T}} + o(T^{-1}) \right]
+ \frac{1}{2} \sum_{j,k=1}^3 [A_{jk} + B_{jk}/T + o(T^{-1/2})](it_j)(it_k)
+ \frac{1}{6\sqrt{T}} \sum_{j,k,l=1}^3 [A_{jkl} + O(T^{-1})](it_j)(it_k)(it_l)
+ \frac{1}{24T} \sum_{j,l,m=1}^3 [A_{jlm} + O(T^{-1})](it_j)(it_k)(it_l)(it_m)
+ \frac{1}{120T^{3/2}} \sum_{j,k,l,m,n=1}^3 [A_{jklmn} + O(T^{-1})](it_j)(it_k)(it_l)(it_m)(it_n)
+ R_\xi. \quad (4.11)
\]
We set down $\Omega = \{A_{jk}\}$, 3x3-matrix, and $\|\mathbf{z}\| = \sqrt{t_1^2 + t_2^2 + t_3^2}$. If $\Omega$ is singular it is not difficult to show that

$$Z_1(\theta) = c_1(\theta)Z_2(\theta) + d_1(\theta) = c_2(\theta)Z_3(\theta) + d_2(\theta), \text{ a.s.}$$

for some constants $c_1(\theta), d_1(\theta), (i = 1,2), (4.12)$

which implies that the joint distribution of $\mathbf{z}$ is reduced to that of $Z_1$. Thus, without loss of generality, henceforth we consider the case when $\Omega$ is nonsingular.

**Lemma 6.** If we take $T$ sufficiently large, then for a $\delta_1 > 0$ and for all $\mathbf{z}$ satisfying $\|\mathbf{z}\| \leq \delta_1\sqrt{T}$, there exists a positive definite matrix $Q_0$ and polynomial functions $F_1(\cdot)$ and $F_2(\cdot)$ such that

$$|\tau(\mathbf{z}) - A(\mathbf{z};3)|$$

$$= \exp\left(-\frac{1}{2} \mathbf{z}'Q_0\mathbf{z}\right) \times F_1(\|\mathbf{z}\|) \cdot O(T^{-3/2})$$

$$+ \exp\left(-\frac{1}{2} \mathbf{z}'Q_0\mathbf{z}\right) \times F_2(\|\mathbf{z}\|) \cdot O(T^{-3/2}), \quad (4.13)$$

where

$$A(\mathbf{z};3) = \exp\left(-\frac{1}{2} \mathbf{z}'\Omega\mathbf{z}\right) \times \left[1 + \frac{1\mu t_1}{\sqrt{T}}\right]$$

$$+ \frac{1}{6\sqrt{T}} \sum_{j,k,\ell=1}^{3} A_{jk\ell}(it_j)(it_k)(it_\ell)$$

$$+ \frac{1}{2T} \sum_{j,k=1}^{3} B_{jk}(it_j)(it_k) + \frac{u^2(it_1)^2}{2T}$$

$$+ \frac{u(it_1)}{6T} \sum_{j,k,\ell=1}^{3} A_{jk\ell}(it_j)(it_k)(it_\ell)$$
\[ + \frac{1}{24T} \sum_{j,k,l,m=1}^{3} A_{jklm}(it_j)(it_k)(it_l)(it_m) \]
\[ + \frac{1}{72T} \sum_{j,k,l'}^{3} A_{jkl'}A_{jk'l'}(it_j)(it_k)(it_{l'}) (it_{l'}) \]

[PROOF] From (4.11) we have
\[ \tau(t) = \exp[-\frac{1}{2T} \Omega t] \times \exp[i t \left( \frac{\mu}{\sqrt{T}} + o(T^{-1}) \right)] \]
\[ + \frac{1}{2T} \sum_{j,k=1}^{3} B_{jk}(it_j)(it_k) + \frac{1}{6\sqrt{T}} \sum_{j,k,l=1}^{3} A_{jkl}(it_j)(it_k)(it_l) \]
\[ + \frac{1}{24T} \sum_{j,k,l,m=1}^{3} A_{jklm}(it_j)(it_k)(it_l)(it_m) \]
\[ + F_3(\|t\|) o(T^{-3/2}) \] (4.14)

where \( F_3(\cdot) \) is a polynomial function. Applying the relation
\[ |e^z - 1 - z - \frac{z^2}{2}| \leq \frac{|z|^3}{3!} e^{|z|} \] (4.15)
to the second exponential in the right-hand side of (4.14) we have
\[ |\tau(t) - A(t;3)| = \exp[-\frac{1}{2T} \Omega t] \cdot F_1(\|t\|) o(T^{-3/2}) \]
\[ + o(T^{-3/2}) \cdot F_4(\|t\|) \cdot \exp[-\frac{1}{2T} \Omega t] \]
\[ \times \exp[i t \left( \frac{\mu}{\sqrt{T}} + o(T^{-1}) \right)] + \frac{1}{2T} \sum_{j,k} \frac{[B_{jk} + o(T^{-1/2})]}{T} (it_j)(it_k) \]
\[ + \frac{1}{6\sqrt{T}} \sum_{j,k,l} [A_{jkl} + o(T^{-1})] (it_j)(it_k)(it_l) \]
\[ + \frac{1}{24T} \sum_{j,k,l,m} [A_{jklm} + o(T^{-1})] (it_j)(it_k)(it_l)(it_m) \]
\[
\frac{1}{120T^{3/2}} \sum_{j,k,\ell,m,n} \left\{ A_{jk\ell mn} + O(T^{-1}) \right\}(it_j)(it_k)(it_\ell)(it_m)(it_n) + R_6, 
\]

where \( F_4(\cdot) \) is a polynomial function. Let \( \omega > 0 \) be the smallest eigen value of \( \Omega \). Then for sufficiently large \( T \), we can choose \( \delta_1 > 0 \) so that

\[
\frac{\omega}{4} - \frac{1}{4} \sum_{j,k,\ell,m,n} |A_{jk\ell mn}| - \frac{1}{24} \sum_{j,k,\ell,m,n} |A_{jk\ell mn}| - \frac{1}{120} \sum_{j,k,\ell,m,n} |A_{jk\ell mn}| - 2^{5/4} \sum_{j,k,\ell,m,n,p} |A_{jk\ell mnp}| > 0. 
\]

Thus the last exponential term in (4.16) is dominated by

\[
\exp\left[\delta_1 \left(u + o(T^{-1/2})\right)\right]\exp\left[\|\xi\|^2 \left(\frac{\omega}{4} + o(T^{-1})\right)\right],
\]

for \( \|\xi\| \leq \delta_1\sqrt{T} \), (4.18)

which implies the existence of \( Q_0 \) in (4.13).

We also have the following lemma.

**Lemma 7.** Under Assumptions 1-5, for every \( \gamma > 0 \), there exists \( \delta_2 > 0 \) such that

\[
|\tau(\xi)| \leq (1 + 4\delta_2 \gamma^2) \cdot q(T)^{1/4}, \tag{4.19}
\]

for all \( \xi \) satisfying \( \|\xi\| \geq \gamma\sqrt{T} \), where \( q(T) = [\gamma T] \), for some constant \( c \).
[PROOF]

Notice that \( \rho_1^2 = \max \mathbb{E}'S_2^2 \mathbb{E} \), where \( \mathbb{E} = (e_1, \ldots, e_T)' \in \mathbb{R}^T \) and \( \mathbb{E}'\mathbb{E} = 1 \). Also we have

\[
\mathbb{E}'S^2_2 \mathbb{E} = \mathbb{E}'\left(\Sigma^{1/2}(t_1\mathbb{H} + t_2\mathbb{H} + t_3\mathbb{H})\Sigma^{1/2}\right)^2 \mathbb{E}
\]

\[
\leq 2t_1^2\mathbb{E}'\Sigma^{1/2}\mathbb{H}\Sigma^{1/2}\mathbb{E} + 2t_2^2\mathbb{E}'\Sigma^{1/2}\mathbb{H}\Sigma^{1/2}\mathbb{E} + 2t_3^2\mathbb{E}'\Sigma^{1/2}\mathbb{H}\Sigma^{1/2}\mathbb{E}
\]

\[
(4.20)
\]

It is not difficult to show that

\[
\mathbb{E}'\Sigma^{1/2}\mathbb{H}\Sigma^{1/2}\mathbb{E} \leq c_1,
\]

\[
(4.21)
\]

\[
\mathbb{E}'\Sigma^{1/2}\mathbb{H}\Sigma^{1/2}\mathbb{E} \leq c_2,
\]

\[
(4.22)
\]

\[
\mathbb{E}'\Sigma^{1/2}\mathbb{H}\Sigma^{1/2}\mathbb{E} \leq c_3,
\]

\[
(4.23)
\]

where \( c_1, c_2 \) and \( c_3 \) are some positive constants. For exposition we prove (4.21). Since \( f_\delta(\lambda), h_\delta(\lambda) \in D_d \) and \( g_\delta \in D^c_{ARMA} \), we can set

\[
f_1 = \max_{\lambda} f_\delta(\lambda) < \infty,
\]

\[
h_1 = \max_{\lambda} |h_\delta(\lambda)| < \infty,
\]

\[
\delta_1 = \min_{\lambda} \delta_\delta(\lambda) > 0.
\]

Thus, using discussions of Anderson (1971, p.573-4) we have

\[
\mathbb{E}'\Sigma^{1/2}\mathbb{H}\Sigma^{1/2}\mathbb{E} = \mathbb{E}'\Sigma^{1/2}A^{-1}\mathbb{B}A^{-1}\Sigma^{-1}\mathbb{B}A^{-1}\Sigma^{1/2}\mathbb{E}
\]

\[
\leq \mathbb{E}'\Sigma^{1/2}A^{-1}\mathbb{B}A^{-1}\left(\begin{array}{cc}
2f_1 & 0 \\
0 & 0
\end{array}\right)A^{-1}\mathbb{B}A^{-1}\Sigma^{1/2}\mathbb{E}
\]
Thus we have proved (4.21). The proofs of (4.22) and (4.23) are similar. From (4.20) we have

$$\rho_T^2 \leq \cdots \leq \rho_1^2 \leq \|t\|^2 \cdot d_{11}$$

(4.25)

for any $t$, where $d_{11}$ is a positive constant. While by Lemma 1, we get

$$T^{-1} \sum_{j=1}^{T} \rho_j^2 = T^{-1} tr S^2$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( t_1 A(\lambda) + t_2 B(\lambda) + t_3 C(\lambda) \right)^2 d\lambda + \|t\|^2 o(T^{-1})$$

$$= \frac{\|t\|^2}{2\pi} \int_{-\pi}^{\pi} \left( \frac{t_1 A(\lambda) + t_2 B(\lambda) + t_3 C(\lambda)}{\|t\|} \right)^2 d\lambda$$

$$+ \|t\|^2 o(T^{-1}),$$

(4.26)
where \( A(\lambda) = \frac{1}{2} \frac{\delta}{\delta \theta} f_\theta(\lambda) \cdot f_\theta(\lambda)^{-1}, \)

\[
B(\lambda) = \frac{1}{2} \left\{ -2f_\theta(\lambda)^{-2} \left( \frac{\delta}{\delta \theta} f_\theta(\lambda) \right)^2 + f_\theta(\lambda)^{-1} \frac{\delta^2}{\delta \theta^2} f_\theta(\lambda) \right\},
\]

and

\[
C(\lambda) = \left\{ 6f_\theta(\lambda)^{-2} \left( \frac{\delta}{\delta \theta} f_\theta(\lambda) \right)^3 - 6f_\theta(\lambda)^{-2} \frac{\delta}{\delta \theta} f_\theta(\lambda) \cdot \frac{\delta^2}{\delta \theta^2} f_\theta(\lambda) \right\}/2.
\]

Since we are now assuming that \( \Omega \) in (4.11) is nonsingular, the functions \( A(\lambda), B(\lambda) \) and \( C(\lambda) \) are linearly independent in the \( L_2 \)-norm \( (\int_{-\pi}^{\pi} |\cdot|^2 d\lambda) \). So we can show that for sufficiently large \( T \), there exists \( d_{12} > 0 \) such that

\[
T^{-1} \sum_{j=1}^{T} \rho_j^2 \geq \|x\|_{d_{12}}^2, \tag{4.27}
\]

for any \( x \). The relations (4.25) and (4.27) imply that there exist \( \delta_2 > 0 \) and \( q(t) = [cT] \) such that

\[
\rho_1^2 \geq \ldots \geq \rho_q(T) \geq \delta_2 \|x\|^2.
\]

Noting that

\[
|\tau(t)| = \prod_{j=1}^{q(T)} \left( 1 + \frac{4\delta_2}{T} \rho_j^2 \right)^{-1/4}
\leq \prod_{j=1}^{q(T)} \left( 1 + \frac{4\delta_2}{T} \|x\|^2 \right)^{-1/4}
\leq (1 + 4\delta_2^n)^{-q(T)/4}, \quad \text{for } \|x\| \geq n/T,
\]

the proof is completed.
5. Edgeworth expansion for $Z$

In this section we shall give the Edgeworth expansion for $Z$. We set down $B(x,r) = \{ z \in \mathbb{R}^p : \| z - x \| \leq r, x \in \mathbb{R}^p \}$. For a probability measure $P$, we denote the variation norm of $P$ by $\|P\|$. The following lemma is known as a smoothing lemma (see Bhattacharya and Rao (1976, p.97-98 and p.113)).

**Lemma 8.** Let $P$ and $Q$ be probability measures on $\mathbb{R}^p$ and $\mathcal{P}$ the class of all Borel subsets of $\mathbb{R}^p$. Let $\varepsilon$ be a positive number. Then there exists a kernel probability measure $K_\varepsilon$ such that

$$\sup_{B \in \mathcal{P}} |P(B) - Q(B)|$$

$$\leq \frac{2}{3} \|P - Q\| K_\varepsilon + \frac{4}{3} \sup_{B \in \mathcal{P}} Q(\{ (\alpha B)^{2\varepsilon} \})$$

(5.1)

where $K_\varepsilon$ satisfies

$$K_\varepsilon(B(0, r)^c) = O\left( \frac{\varepsilon^3}{r} \right),$$

(5.2)

and the Fourier transform $\hat{K}_\varepsilon$ satisfies

$$\hat{K}_\varepsilon(\xi) = 0 \text{ for } \|\xi\| > 8p^{1/3}/\pi^{1/3} \varepsilon.$$

For $B \in \mathcal{A}^3$, define

$$Q^{(3)}_Z(B) = \mathbb{E} N(Z: \Omega)[1 + \frac{u}{\sqrt{T}} H_1(Z)$$

$$+ \frac{1}{6T} \sum_{j,k,l=1}^3 A_{jk} H_{jk}(Z) + \frac{u^2}{2T} H_{11}(Z)$$
\[ + \frac{1}{2T} \sum_{j,k=1}^{3} B_{jk} H_{jk}(z) + \frac{1}{6T} \sum_{j,k,l=1}^{3} \mu A_{jk\ell} H_{jk\ell}(z) \]

\[ + \frac{1}{24T} \sum_{j,k,\ell}^{3} \sum_{m=1}^{2} A_{jk\ell m} H_{jk\ell m}(z) \]

\[ + \frac{1}{72T} \sum_{j,k,\ell}^{3} \sum_{j',k',\ell'=1}^{3} A_{jk\ell aj'k'\ell'} H_{jk\ell j'k'\ell'}(z), \] \hspace{1cm} (5.3)

where \( z = (z_1, z_2, z_3)' \),

\[ N(z; \Omega) = (2\pi)^{-3/2} |\Omega|^{-1/2} \exp -\frac{1}{2} z' \Omega^{-1} z, \]

\[ H_{j_1 \ldots j_s}(z) = \frac{(-1)^s}{N(z; \Omega)} \frac{\partial^s}{\partial z_{j_1} \ldots \partial z_{j_s}} N(z; \Omega). \]

This measure \( Q_{z}^{(3)}(\cdot) \) corresponds to the characteristic function \( A(z; 3) \) in Lemma 6. Then we have the following lemma.

**Lemma 9.** Suppose that Assumptions 1-5 are satisfied. Then

\[ \sup_{B \in \mathbb{A}_3^3} \left| P_{z}^{T}(z \in B) - Q_{z}^{(3)}(B) \right| = o(T^{-1}) \]

\[ + \frac{4}{3} \sup_{B \in \mathbb{A}_3^3} Q_{z}^{(3)}((\delta B)^{2\varepsilon}), \] \hspace{1cm} (5.4)

uniformly for \( \theta \in \Theta \), where \( \varepsilon = T^{-1-\rho}, 0 < \rho < 1/2. \)

**Proof** Substituting \( P_{z}^{T}(z \in B) \) and \( Q_{z}^{(3)}(B) \) for \( P(z) \) and \( Q(z) \) in Lemma 8, respectively, we get
\[
\sup_{B \in \mathfrak{B}_3} |P_\theta^T(\mathbb{Z} \in B) - Q_\mathbb{Z}^{(3)}(B)| \\
\leq \frac{2}{3} \| (P_\theta^T - Q_\mathbb{Z}^{(3)}) * K_e \|
\]
\[
+ \frac{4}{3} \sup_{B \in \mathfrak{B}_3} Q_\mathbb{Z}^{(3)}((\delta B)^{2e}). \\
(5.5)
\]

Notice that
\[
\|(P_\theta^T - Q_\mathbb{Z}^{(3)}) * K_e\|
\]
\[
= 2 \sup \{|(P_\theta^T - Q_\mathbb{Z}^{(3)}) * K_e(B)|; B \in \mathfrak{B}_3\}
\]
\[
\leq 2 \sup \{|(P_\theta^T - Q_\mathbb{Z}^{(3)}) * K_e(B)|; B \in \mathfrak{B}_3 \text{ and } B \subseteq B(2, r_T)\}
\]
\[
+ 2 \sup \{|(P_\theta^T - Q_\mathbb{Z}^{(3)}) * K_e(B)|; B \in \mathfrak{B}_3 \text{ and } B \subseteq B(2, r_T)^c\},
\]
\[
(5.6)
\]

where \( r_T = T^p, \ 0 < p < 1/6 \). Here we put \( e = T^{-1-p}, \ 0 < p < 1/2 \).

Then, for \( B \subseteq B(2, r_T)^c \) we have
\[
|P_\theta^T - Q_\mathbb{Z}^{(3)}(B)| \leq |P_\theta^T * K_e(B)| + |Q_\mathbb{Z}^{(3)} * K_e(B)|
\]
\[
\leq P_\theta^T(\|\mathbb{Z}\| \geq r_T/2) + K_e(B(2, r_T/2)^c)
\]
\[
+ Q_\mathbb{Z}^{(3)}(B(2, r_T/2)^c)) + K_e(B(2, r_T/2)^c). \\
(5.7)
\]

It is easy to check
\[
Q_\mathbb{Z}^{(3)}(B(2, r_T/2)) = o(T^{-1}). \hspace{1cm} (5.8)
\]
The relations (3.25), (3.26) and (3.27) imply

$$P_T^T(\|z\| \geq r_T/2) = o(T^{-1}).$$  

(5.9)

While (5.2) implies

$$K_e[B(0, r_T/2)^c] = O(T^{-3-3\rho-3\tau}) = o(T^{-1}).$$  

(5.10)

Thus we have

$$| (P_T^T - Q_z(3))^* K_e(B) | = o(T^{-1}),$$

(5.11)

for $B \subseteq B(0, r_T)^c$. Now we have only to evaluate

$$\sup\{ | (P_T^T - Q_z(3))^* K_e(B) | ; B \subseteq B(0, r_T) \}.$$  

By Fourier inversion we have

$$| (P_T^T - Q_z(3))^* K_e(B) |$$

$$\leq (2\pi)^{-3} \frac{\pi^{3/2} r_T^3}{3^{\Gamma(3)/2}} \int | (\hat{P}_T^T - \hat{Q}_z(3))(t) \hat{K}_e(t) | dt.$$  

(5.12)

By Lemma 6 and noting $\hat{Q}_z(3)(t) = A(t; 3)$, the right-hand side of (5.12) is dominated by

$$O(T^{3\tau-3/2}) \int_{\|z\| \leq \delta_1 / T} | \exp\{- \frac{1}{2} z^2 \langle z \rangle \} \times F_1(\|z\|)$$

$$+ \exp\{- z^2 \langle z \rangle \} \times F_2(\|z\|) | \hat{K}_e(t) | dt$$

$$+ O(T^{3\tau}) \int_{\delta_1 / T < \|z\| \leq 3(3)^{4/3} r_T^{1+\rho/3}} | (\hat{P}_T^T - \hat{Q}_z(3))(t) \hat{K}_e(t) | dt.$$  

(5.13)
Evidently the first term of the above is of order $O(T^{-4/3})$. Also by Lemma 7, we have

$$O(T^{3\tau}) \int_{\delta_1 \sqrt{T} \leq \|\xi\| \leq 8(3)^{4/3} T^{1+\rho/3} \leq 1} |(P^T_\theta - Q^{(3)}_\mathbb{Z}) (\xi) \mathcal{K}_\varepsilon (\xi)| \, d\xi$$

$$\leq O(T^{3\tau}) \int_{\delta_1 \sqrt{T} \leq \|\xi\| \leq \delta_4 T^{1+\rho} d_{13} (1 + 4\delta_2 \delta_4^{-1})^{-\sigma(0)/4}} \, d\xi + o(T^{-1}),$$

(5.14)

where $d_{13}$ and $d_{14}$ are appropriate positive constants. The above (5.14) is dominated by

$$O(T^{3\tau+3+3\rho})(1 + 4\delta_2 \delta_4^{-1})^{-\sigma(0)/4} + o(T^{-1}) = o(T^{-1}).$$

(5.15)

Therefore we have proved

$$\sup \{ |(P^T_\theta - Q^{(3)}_\mathbb{Z}) \mathcal{K}_\varepsilon (B)| ; B \subseteq B(Q, r_T) \} = o(T^{-1}),$$

which completes the proof. \(\Box\)
6. Proof for (2) of Theorem 1.

Consider the following transformation

\[
\begin{align*}
W_1(\theta) &= Z_1(\theta) \\
W_2(\theta) &= Z_2(\theta) - J(\theta)I(\theta)^{-1}Z_1(\theta), \\
W_3(\theta) &= Z_3(\theta) - L(\theta)I(\theta)^{-1}Z_1(\theta).
\end{align*}
\] (6.1)

Henceforth, for simplicity we sometimes use \(W_1, W_2\) and \(W_3\), instead of \(W_1(\theta), W_2(\theta)\) and \(W_3(\theta)\), respectively. Evidently (6.1) is a continuous bijective transformation. We denote (6.1) by \(\mathcal{W} = \chi(\mathbb{Z})\), where \(\mathcal{W} = (W_1, W_2, W_3)'\). By Lemma 9, we have

\[
\sup_{B \in \mathcal{A}} \left| P_{\mathcal{E}}^T[Z \in \chi^{-1}(B)] - Q_\mathcal{Z}^{(3)}[\chi^{-1}(B)] \right|
\]

\[
= \frac{4}{3} \sup_{B \in \mathcal{A}} Q_\mathcal{Z}^{(3)}[(\Delta \chi^{-1}(B))^{2\mathcal{E}}] + o(T^{-1}).
\] (6.2)

Here we put \(Q_\mathcal{W}^{(3)}(B) = Q_\mathcal{Z}^{(3)}[\chi^{-1}(B)]\). Then it is not difficult to show

\[
Q_\mathcal{W}^{(3)}(B) = \int_B N(w_1; I)N(w_2, w_3; \Omega_2)[1 +
\sum_{j=1}^{3} \frac{c_j^{(1)}}{\sqrt{T}} H_j(\mathcal{W}) + \frac{l}{6\sqrt{T}} \sum_{j,k,l=1}^{3} c_{jkl}^{(1)} H_{jkl}(\mathcal{W})
\]

\[
+ \frac{1}{2T} \sum_{j,k=1}^{3} (c_j(3) + c_j^{(1)}c_k^{(1)}) H_{jk}(\mathcal{W})
\]

\[
+ \frac{1}{6T} \sum_{j,k,l,m=1}^{3} c_j^{(1)}c_k^{(1)} c_{jklm} H_{jklm}(\mathcal{W})
\].
\[ + \frac{1}{24T} \sum_{j,k,l,m=1} c_{jklm}^{(1)} \]
\[ + \frac{1}{72T} \sum_{j,k,l,j',k',l'=1} c_{jkl}^{(1)} c_{j'k'l'}^{(1)} H_{jkl} j'k'l'(w) \]
\[ = \int_B q_T(w) d\mu, \text{ say,} \] (6.3)

where \( \mu = (w_1, w_2, w_3)' \), \( N(w_1: I) = (2\pi)^{-1/2} I^{-1/2} \exp - \frac{w_1^2}{2I} \)
and \( N(w_2, w_3: \Omega_2) = (2\pi)^{-1} |\Omega_2|^{-1/2} \exp - \frac{1}{2}(w_2, w_3)\Omega_2^{-1}(w_2, w_3) \)
\[ \Omega_2 = \begin{pmatrix} \Omega_{22} & \Omega_{23} \\ \Omega_{32} & \Omega_{33} \end{pmatrix}, \text{ 2x2-matrix.} \]

For examples we can see
\[ c_1^{(1)} = \mu(\theta), \quad c_2^{(1)} = -J(\theta)\mu(\theta)/I(\theta), \quad c_3^{(1)} = -L(\theta)\mu(\theta)/I(\theta), \]
\[ \Omega_{22} = M(\theta) - J(\theta)^2 I(\theta)^{-1}, \quad c_{112}^{(1)} = N(\theta) - J(\theta) K(\theta)/I(\theta), \]
\[ c_{11111}^{(1)} = H(\theta), \text{ e.t.c..} \]

Since \( \chi \) is continuous, we have
\[ \partial \chi^{-1}(B) \subset \chi^{-1}(\partial B), \]
\[ (\partial \chi^{-1}(B))^2 \varepsilon \subset (\chi^{-1}(\partial B))^{2\varepsilon}. \] (6.4)

By the continuity of \( \chi \), there exists \( a > 0 \) such that
\[ (\chi^{-1}(\partial B))^2 \varepsilon \subset \chi^{1(\partial B)a\varepsilon}. \] (6.5)

Thus we have
LEMMA 10. Under Assumptions 1-5

\[
\sup_{B \in \mathcal{B}^3} \left| p_T^*(\mathcal{W} \in B) - q^{(3)}(B) \right|
\]

\[
= \frac{4}{3} \sup_{B \in \mathcal{B}^3} q^{(3)}((\lambda B) \mathcal{E}) + o(T^{-1}),
\]

(6.6)

uniformly for \( \theta \in \Theta \), where \( \mathcal{E} \) is a positive constant and \( \mathcal{E} = T^{-1-\rho} \), \( 0 < \rho < 1/2 \).

Now we rewrite \( U_T(\theta) \) in Lemma 5 as

\[
U_T(\mathcal{W}) = \frac{W_1}{I^T_T} + \frac{W_1 W_2}{I^2 \sqrt{T}} - \frac{(J + K) W_1^2}{2 I^3 \sqrt{T}}
\]

\[
+ \frac{1}{I^3} (W_1 W_2^2 + \frac{5J - 3K}{2I} W_1 W_2^2 + \frac{1}{2} W_1 W_2)
\]

\[
+ \frac{2J^2 + 3KJ + K^2}{2I^2} W_1^3 - \frac{L + 3M + 6N + H}{6I} W_1^3.
\]

(6.7)

Consider the following transformation

\[
\begin{cases}
S_1 = U_T(\mathcal{W}) \\
S_2 = W_2 \\
S_3 = W_3
\end{cases}
\]

(6.8)

We denote (6.8) by \( \mathcal{S} = \psi(\mathcal{W}) \), where \( \mathcal{S} = (S_1, S_2, S_3)' \). For sufficiently large \( T \), we can take a set

\[
M_T = \left\{ \mathcal{W} : |W_i| \leq c_i T^\alpha, 0 < \alpha < 1/6, c_1 > 0, i = 1, 2, 3 \right\}
\]

such that \( \psi \) is a \( C^\infty \)-mapping on \( M_T \).
By (6.6),
\[
\sup_{B \in \mathcal{B}} |P_B^T(W \in \psi^{-1}(B \times \mathbb{R}^2)) - Q^3(W \in \psi^{-1}(B \times \mathbb{R}^2))| = \frac{4}{3} \sup_{B \in \mathcal{B}} Q^3((a \psi^{-1}(B \times \mathbb{R}^2)) \omega^2) + o(T^{-1}).
\] (6.9)

We can see that
\[
Q^3(W \in \psi^{-1}(B \times \mathbb{R}^2)) = \int_{\psi^{-1}(B \times \mathbb{R}^2)} q_T(W) dW
\]
\[
= \int_{M_T \cap \psi^{-1}(B \times \mathbb{R}^2)} q_T(W) dW + o(T^{-1})
\]
\[
= \int_{(B \times \mathbb{R}^2) \cap N_T} q_T(\psi^{-1}(\xi)) |J| d\xi + o(T^{-1}),
\] (6.10)

where \(N_T = \psi(M_T)\) and \(|J|\) is the Jacobian. Since we can solve so that
\[
W_1 = I_T S_1 - \frac{S_1 S_2}{\sqrt{T}} + \frac{(J + K) S_1^2}{2 \sqrt{T}} + \frac{J S_1 S_2}{2T} - \frac{S_1 S_3}{2T}
\]
\[
+ \frac{(-J^2 - JK) S_3^3}{2T} + \frac{L + 3M + 6N + H S_3}{6T} + o(T^{-1}),
\] (6.11)

uniformly on \(M_T\), it is not difficult to show that
\[
q_T(\psi^{-1}(\xi)) |J|
\]
\[
= N(I_T S_1) N(S_2, S_3) \times \left[ I_T + \frac{p_1(\xi)}{\sqrt{T}} + \frac{p_2(\xi)}{T} + o(T^{-1}) \right],
\] (6.12)

uniformly on \(N_T\), where \(p_1(\xi)\) and \(p_2(\xi)\) are polynomials of \(\xi\). Thus
we have
\[
\mathbb{Q}^{(3)}(g^{-1}(B \times \mathbb{R}^2)) = \int_{\{B \times \mathbb{R}^2\} \cap N} N(I_{T}S_{1})N(S_{2}, S_{3}; \mathbb{E})
\]
\[
x \left[ I_{T} + \frac{P_{1}(\mathbb{E})}{\sqrt{T}} + \frac{P_{2}(\mathbb{E})}{T} + o(T^{-1}) \right] d\mathbb{E} + o(T^{-1})
\]
\[
= \int_{B} N(I_{T}S_{1})[\int_{\mathbb{R}^2} \left( I_{T} + \frac{P_{1}(\mathbb{E})}{\sqrt{T}} + \frac{P_{2}(\mathbb{E})}{T} \right) d\mathbb{E}]
\]
\[
x N(S_{2}, S_{3}; \mathbb{E}) d\mathbb{E} dS_{1} + o(T^{-1}). \tag{6.13}
\]

Calculating the square bracket in (6.13), and noting that
\[
\mathbb{Q}^{(3)}((\partial \psi^{-1}(B \times \mathbb{R}^2); \mathbb{E})
\]
\[
\leq \mathbb{Q}^{(3)}(\psi^{-1}((\partial B)^{b} \times \mathbb{R}^2)), \text{ for some } b > 0,
\]
we have
\[
\sup_{B \in \mathcal{B}_{0}} |P_{3}^{T}(\hat{\theta}_{T} - \theta) \in B| - \int_{B} \varphi(x)p_{3}^{T}(x)dx |
\]
\[
= \frac{4}{3} \sup_{B \in \mathcal{B}_{0}} \int_{(\partial B)^{b}} \varphi(x)p_{3}^{T}(x)dx + o(T^{-1}), \tag{6.14}
\]

(remember (6.9)). Here
\[
p_{3}^{T}(x) = 1 + \frac{a_{1} x}{\sqrt{T}} + \frac{\gamma_{1}}{6\sqrt{T}}(x^{3} - 3x)
\]
\[
+ \frac{1}{2}(\frac{P_{2}}{T} + \frac{a_{2}^{2}}{T})(x^{2} - 1) + (\frac{\delta_{1}}{24T} + \frac{\gamma_{1} \gamma_{2}}{6T})(x^{4} - 6x^{2} + 3)
\]
\[
+ \frac{\gamma_{2}}{72T}(x^{6} - 15x^{4} + 45x^{2} - 15), \tag{6.15}
\]
where
\[
\alpha_1 = - \frac{J + K}{2I^{3/2}} + \frac{u}{I^{1/2}}
\]
\[
\gamma_1 = - \frac{3J + 2K}{I^{3/2}}
\]
\[
\beta_2 = \frac{2n}{I} + \frac{\Delta}{I} + \frac{7J^2 + 14JK + 5K^2}{2I^3} - \frac{L + 4N + H}{I^2} - \frac{2\mu(2J + K)}{I^2}
\]
\[
\delta_1 = \frac{12(2J + K)(J + K)}{I^3} - \frac{4L + 12N + 3H}{I^2},
\]

where
\[
\text{Var}(Z_1(\theta)) = I(\theta) + \frac{\Delta(\theta)}{T} + o(T^{-1}), \quad (6.16)
\]
\[
I_T(\theta) = I(\theta) - \frac{n(\theta)}{T} + o(T^{-1}). \quad (6.17)
\]

Remembering (2.6) we have proved (2) of Theorem 1. More explicit forms of (6.15) for the exact maximum likelihood estimators are given in Taniguchi (1985).
7. Third order asymptotic properties of minimum contrast estimators.

Taniguchi (1985) discussed third order asymptotic properties of maximum likelihood estimators in the class of third order asymptotically median unbiased (AMU) estimators, and showed a certain optimality of maximum likelihood estimators. Using the Edgeworth expansions of minimum contrast estimators we can discuss their third order asymptotic properties in this class.

If an estimator \( \hat{\theta}_T \) satisfies the equations

\[
\lim_{T \to \infty} T^{(k-1)/2} \mathbb{P}_\theta \left( \sqrt{T} (\hat{\theta}_T - \theta) \leq 0 \right) - 1/2 = 0, \tag{7.1}
\]

\[
\lim_{T \to \infty} T^{(k-1)/2} \mathbb{P}_\theta \left( \sqrt{T} (\hat{\theta}_T - \theta) > 0 \right) - 1/2 = 0, \tag{7.2}
\]

then \( \hat{\theta}_T \) is called kth-order asymptotically median unbiased (kth-order AMU for short). We denote the set of kth-order AMU estimators by \( A_k \). In general the minimum contrast estimator \( \hat{\theta}_T \) is not third order AMU. To be so a modification of \( \hat{\theta}_T \) is required. The following theorem gives the validity of Edgeworth expansion for modified estimators of \( \hat{\theta}_T \).

**Theorem 2.** Suppose that \( m(\theta) \) is a continuously twice differentiable function. Define

\[
\hat{\theta}_m = \hat{\theta}_T + \frac{1}{2} m(\hat{\theta}_T).
\]

Then

\[
\sup_{F \in \mathcal{F}_0} \left| \mathbb{P}_\theta \left( \sqrt{T} (\hat{\theta}_m - \theta) \in \mathcal{B} \right) - \int_{\mathcal{B}} \mathcal{A}(y) q_m^T(y) dy \right| = o(T^{-1}). \tag{7.3}
\]
uniformly for $\theta \in C$, where
\[
q_{m3}^T(y) = 1 + \frac{1}{\sqrt{T}}[\alpha_1 + \sqrt{T}m(\theta)] y + \frac{\gamma_1}{6\sqrt{T}}(y^3 - 3y)
+ \frac{1}{2T}r_2 + \alpha_1^2 + 2m'(\theta) + \text{Im}(\theta)^2 + 2\alpha_1\sqrt{T}m(\theta)](y^2 - 1)
+ \frac{\delta_1}{T}[\frac{\alpha_1^2}{\gamma_1} + \frac{\gamma_1\sqrt{T}m(\theta)}{6}] (y^4 - 6y^2 + 3)
+ \frac{\gamma_1^2}{72T}(y^6 - 15y^4 + 45y^2 - 15).
\]

[PROOF]

Since $m(\cdot)$ is continuously twice differentiable, we have
\[
\sqrt{T}\{\hat{\theta}_m - \theta\} = \sqrt{T}\{\hat{\theta}_T - \theta\} + \frac{\sqrt{T}}{\sqrt{T}}m(\theta)
+ \sqrt{T}\{\hat{\theta}_T - \theta\}m'(\theta) + T^{-3/2}\{\sqrt{T}\{\hat{\theta}_T - \theta\}\}^2m''(\theta^*)/\sqrt{T},
\]
where $0 \leq \theta^* \leq \hat{\theta}_T$.

By (1) of Theorem 1 we have
\[
P_{\theta}[\{\sqrt{T}\{\hat{\theta}_T - \theta\}\}^2m'(\theta^*)/\sqrt{T} > T^{2\alpha}] = o(T^{-1}),
\]
for $0 < \alpha < 1/4$. Putting $\rho_T = T^{2\alpha-1/2}$ in Lemma 4, we have only to derive the Edgeworth expansion for $a_TU_T + s_T$, where $a_T = [1 + m'(\theta)/T]$, $U_T = \sqrt{T}\{\hat{\theta}_T - \theta\}$ and $s_T = \frac{\sqrt{T}}{T}m(\theta)$. By Theorem 1, we have
\[
\sup_{B \in B_0} \left| P_{\theta}[U_T \in B] - \int_B \varphi(x)p_{T}^*(x)dx \right| = o(T^{-1}).
\]

Lemma 4 implies that
Also we have
\[ \sup_{B \in \mathcal{B}_0} |P_T(\sqrt{T}(\hat{\theta}_m - \theta) \in B) - P_T[a_T U_T + s_T \in B]| = o(T^{-1}). \quad (7.7) \]

Transforming \( y = a_T x + s_T \), it is not difficult to show
\[ \int_{a_T x + s_T \in B} \varphi(x)p_3^T(x)dx = \int_{B} \varphi(y)q_3^T(y)dy + o(T^{-1}). \quad (7.9) \]

The relations (7.7), (7.8) and (7.9) imply our assertion.  \( \square \)

For \( m(\theta) = \frac{K(\theta)}{6I(\theta)^2} - \frac{\mu(\theta)}{I(\theta)} \), we denote \( \hat{\theta}_T^* = \hat{\theta}_m^* \). In this case we have

**COROLLARY 2.**

\[ \sup_{B \in \mathcal{B}_0} |P_T(\sqrt{T}(\hat{\theta}_T^* - \theta) \in B)| \]
\[ - \int_{B} \varphi(x)[1 + \frac{\gamma_1 x}{6\sqrt{T}} + \frac{\gamma_1}{6\sqrt{T}}(x^3 - 3x)] \]
\[ + \frac{1}{2T} \left[ \frac{2n}{I} + \frac{\Delta}{I} - \frac{2\mu I}{I} + \frac{135J^2 + 216JK + 70K^2}{36I^3} \right] \]
\[ - \frac{3L + 9N + 2H}{3I^2} (x^2 - 1) + \frac{1}{T} \left[ \frac{1}{24} + \frac{\gamma_2^2}{36} (x^4 - 6x^2 + 3) \right] \]
\[ + \frac{\gamma_2^2}{72T} (x^6 - 15x^4 + 45x^2 - 15)]dx| = o(T^{-1}). \quad (7.10) \]
REMARK

Of course $\hat{\theta}_T^*$ belongs to $A_2$, and we can see that the asymptotic distribution of $\hat{\theta}_T^*$ (up to second order) coincides with that of the second order efficient estimator (see Taniguchi (1983) or (1985)).

REMARK

It is easy to check that $\hat{\theta}_T^*$ is third order AMU. Also it is $e_d = 2n + \Delta - 2\mu'$ that depends on the minimum contrast estimator.

Let $\hat{\theta}_1^*(i = 1,2)$ be the modified minimum contrast estimators with $\mu_1$, $\Delta_1$, $\eta_1$ and $m_1(i = 1,2)$ in place of $\mu$, $\Delta$, $\eta$ and $m$, respectively. Then we have

COROLLARY 3. For $B = (-a,a)$, $a > 0,$

$$\lim_{T \to \infty} T[P_0^T(\hat{\theta}_1^* - \theta) \in B] - P_0^T[\sqrt{T}(\hat{\theta}_2^* - \theta) \in B]) = \frac{1}{n} a \phi(a)[2(\eta_2 - \eta_1) + \Delta_2 - \Delta_1 + 2(\mu_1 - \mu_2)],$$ (7.11)

Thus if $2\eta_1 + \Delta_1 - 2\mu_1'$ is smaller than $2\eta_2 + \Delta_2 - 2\mu_2'$, then $\hat{\theta}_1^*$ is better than $\hat{\theta}_2^*$ in third order sense.

EXAMPLE 3. Let $\{X_t\}$ be a Gaussian autoregressive process with the spectral density

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi \lambda} \frac{1}{|1 - \theta e^{i\lambda}|^2},$$

where $|\theta| < 1$.

Let $\hat{\theta}_1^*$ be the modified maximum likelihood estimator of $\theta$ defined in Example 1). Also let $\hat{\theta}_2^*$ be the modified quasi-maximum
likelihood estimator of $\theta$ (defined in Example 2). Then we have

$$u_1 = 0, \Delta_1 = \frac{3\theta^2 - 1}{(1 - \theta^2)^2}, \eta_1 = -\Delta_1,$$

$$u_2 = \frac{-\theta}{(1 - \theta^2)}, \Delta_2 = \frac{-\theta^2 - 1}{(1 - \theta^2)^2}, \eta_2 = 0.$$

For this case, the right-hand side of (7.11) is equal to

$$\frac{4}{4} a \mathcal{Q}(a) \frac{\theta^2}{1 - \theta^2} \geq 0,$$

which coincides with the result of Fujikoshi and Ochi (1984). That is, $\hat{\theta}_1^*$ is better than $\hat{\theta}_2^*$.

Acknowledgments

This work was done during the author's stay in the Center for Multivariate Analysis, University of Pittsburgh. The author is grateful to Professor P.R. Krishnaiah for his kindness and encouragement.
REFERENCES


VALIDITY OF EDGEWORTH EXPANSIONS OF MINIMUM CONTRAST ESTIMATORS FOR GAUSSIAN ARMA PROCESSES

Masanobu Taniguchi

Center for Multivariate Analysis
515 Thackeray Hall
Univ. of Pittsburgh, Pittsburgh, PA 15260

Air Force Office of Scientific Research
Department of the Air Force
Bolling Air Force Base DC 20332

Approved for public release; distribution unlimited

Let \( (X_t) \) be a Gaussian ARMA process with spectral density \( f_\theta(\lambda) \), where \( \theta \) is an unknown parameter. To estimate \( \theta \) we propose a minimum contrast estimation method which includes the maximum likelihood method and the quasi-maximum likelihood method as special cases. Let \( \hat{\theta}_m \) be the minimum contrast esti-