Abstract

This paper discusses the asymptotic efficiency of estimators for optimal portfolios when the returns are vector-valued Gaussian stationary processes. Then it is shown that the usual portfolio estimators are not asymptotically efficient if the returns are dependent. Numerical studies for the difference between the asymptotic variance of the portfolio estimators and the Cramer-Rao bound are given. The results clearly illuminate the inefficiency of the usual estimators for vector-valued ARMA(1,2) processes. From this point of view we construct portfolio estimators which are asymptotically efficient.

1. Introduction

In the theory of portfolio analysis, optimal portfolios are determined by the mean $\mu$ and variance $\Sigma$ of the portfolio return. Several authors proposed estimators of the optimal portfolios as the functions of the sample mean $\hat{\mu}$ and the sample variance $\hat{\Sigma}$ for independent returns of assets (e.g. Jobson and Korkie, 1980 and 1989; Lauprete, Samarov and Welsch, 2002). However, empirical studies show that financial return processes are often dependent. From this point of view, Basak, Jagannathan and Sun (2002) showed the consistency of optimal portfolio estimators when the portfolio returns are stationary processes.

In the literature there has been no study on the asymptotic efficiency of estimators for optimal portfolios. Therefore, in this paper, denoting the optimal portfolios by a function $g = g(\mu, \Sigma)$ of $\mu$ and $\Sigma$, we discuss the asymptotic efficiency of estimators $\hat{g} = g(\hat{\mu}, \hat{\Sigma})$ when the returns are vector-valued Gaussian stationary processes.

Section 3 gives the asymptotic distribution of $\hat{g}$. Section 4 addresses the problem of asymptotic efficiency for the class of estimators $\hat{g}$. It is seen that $\hat{g}$ is not asymptotically efficient generally if the returns are dependent. Such examples are provided. For asymptotically inefficient cases, we give some numerical results, which illuminate some interesting feature of them. Our estimators $\hat{g}$ includes many famous portfolio estimators as special cases, and the asymptotic results give a strong warning for use of the usual portfolio estimators when we
observe dependent return processes.

Throughout this paper, $\|A\|_E$ denotes the Euclidean norm of a matrix $A$. If $(X_n)$ is a sequence of random vectors which converges in distribution to a random vector $X$, then we write $X_n \xrightarrow{d} X$. The 'vec' operator transforms a matrix into a vector by stacking the columns, and the 'vech' operator transforms a symmetric matrix into a vector by stacking the elements on and below the main diagonal. For matrices $A$ and $B$, $A \otimes B$ denotes the Kronecker product of $A$ and $B$, whose $(j_1, j_2)$th block is $a_{j_1,j_2}B$.

2. Optimal portfolios

Suppose the existence of a finite number of assets indexed by $i, (i = 1, \ldots, m)$. Let $X(t) = (X_1(t), \ldots, X_m(t))'$ denote the random returns on $m$ assets at time $t$. Write $\mu = E\{X(t)\}$ and $\Sigma = \text{Cov}(X(t))$. Let $\alpha = (\alpha_1, \ldots, \alpha_m)'$ be the vector of portfolio weights. Then the return of portfolio is $X(t)'\alpha$, and the expectation and variance are, respectively, given by $\mu(\alpha) = \mu'\alpha$, $\eta^2(\alpha) = \alpha'\Sigma\alpha$. Optimal portfolio weights have been proposed by various criteria. The followings are the typical ones.

I.

\[
\begin{cases}
\max_{\alpha} \{\mu(\alpha) - a\eta^2(\alpha)\}, \\
\text{subject to } e'\alpha = 1,
\end{cases}
\]  
(2.1)

where $e = (1, \ldots, 1)'$ ($m \times 1$-vector), and $a$ is a given positive number. The solution is given by

\[
\alpha_I = \frac{1}{2a} \left\{ \Sigma^{-1} \mu - \frac{e'\Sigma^{-1}\mu}{e'\Sigma^{-1}e} \right\} + \frac{\Sigma^{-1}e}{e'\Sigma^{-1}e}. 
\]  
(2.2)

If we take $\eta^2(\alpha)$ as the utility function, the criterion is

II.

\[
\begin{cases}
\min_{\alpha} \eta^2(\alpha), \\
\text{subject to } e'\alpha = 1.
\end{cases}
\]  
(2.3)

The solution is given by

\[
\alpha_{II} = \frac{\Sigma^{-1}e}{e'\Sigma^{-1}e}. 
\]  
(2.4)

Let us now suppose that there exists a risk-free asset. We denote by $R_0$ its return, and denote by $\alpha_0$ the amount. The problem to be solved is given by

III.

\[
\begin{cases}
\max_{\alpha_0, \alpha} \{\mu(\alpha) + R_0\alpha_0 - a\eta^2(\alpha)\}, \\
\text{subject to } \sum_{j=0}^{m} \alpha_j = 1.
\end{cases}
\]  
(2.5)
Then the solution for $\alpha$ and $\alpha_0$ are

$$\alpha_{III} = \frac{1}{2a} \Sigma^{-1}(\mu - R_0 e),$$  \hspace{1cm} (2.6)$$

$$\alpha_{0III} = 1 - \frac{1}{2a} e^' \Sigma^{-1}(\mu - R_0 e).$$ \hspace{1cm} (2.7)

Therefore the optimal portfolios can be considered as smooth functions of $\mu$ and $\Sigma$, i.e. we may put

$$g_1(\mu, \Sigma) = \frac{1}{2a} \left\{ \Sigma^{-1} \mu - \frac{\mu^' \Sigma^{-1} \mu}{\mu^' \Sigma^{-1} e} e \right\} + \frac{\Sigma^{-1} e}{\mu^' \Sigma^{-1} e},$$ \hspace{1cm} (2.2)$$

$$g_2(\mu, \Sigma) = \frac{\Sigma^{-1} e}{\mu^' \Sigma^{-1} e},$$ \hspace{1cm} (2.4)$$

$$g_3(\mu, \Sigma) = \frac{1}{2a} \Sigma^{-1}(\mu - R_0 e),$$ \hspace{1cm} (2.6)$$

$$g_4(\mu, \Sigma) = 1 - \frac{1}{2a} e^' \Sigma^{-1}(\mu - R_0 e).$$ \hspace{1cm} (2.7)$$

Unifying the above we consider to estimate a general function $g(\mu, \Sigma)$ of $\mu$ and $\Sigma$. Here it should be noted that the coefficient $\alpha$ satisfies the restriction $e^' \alpha = 1$. Then we have only to estimate the subvector $(\alpha_1, \ldots, \alpha_{m-1})'$. Hence we assume that the function $g(\cdot)$ is $(m-1)$-dimensional, i.e.,

$$g : (\mu, \Sigma) \rightarrow \mathbb{R}^{m-1}.$$ \hspace{1cm} (2.8)

This paper addresses the problem of statistical estimation for $g(\mu, \Sigma)$, which describes various optimal portfolios.

3. **Asymptotic Theory for Fundamental Quantities**

Empirical studies show that financial return processes are often not independent. So it is natural to suppose that the return process concerned is dependent. In this paper we assume that the return process $\{X(t) = (X_1(t), \ldots, X_m(t))'; t \in \mathbb{Z}\}$ is a Gaussian $m$-vector stationary process with mean $\mu = (\mu_1, \ldots, \mu_m)'$ and autocovariance matrix $R(k) = E((X(t) - \mu)(X(t+k) - \mu)').$

**Assumption 1.**

$$\sum_{l=-\infty}^{\infty} \|R(l)\|_E < \infty$$

The spectral density matrix of the process $\{X(t)\}$ exists and is

$$f(\lambda) = \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} R(l) e^{i\lambda l}. \hspace{1cm} (3.1)$$

Note that the autocovariance matrix at lag $k$ is expressed as

$$R(k) = \int_{-\pi}^{\pi} f(\lambda) e^{i\lambda k} d\lambda. \hspace{1cm} (3.2)$$
In what follows, \( k \) is assumed to be nonnegative, since \( R(k) = R(-k)' \).

Let \( X(1), \ldots, X(T) \) be an observed stretch from \( \{X(t)\} \). Then we write
\[
R(0) = \Sigma = \{\sigma_{ij}\}, \quad \theta = (\mu', vech(\Sigma)')',
\]
and
\[
\hat{\mu} = \frac{1}{T} \sum_{t=1}^{T} X(t) \quad (3.3)
\]
\[
\hat{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} (X(t) - \hat{\mu})(X(t) - \hat{\mu})' \quad (3.4)
\]
\[
\hat{\theta} = (\hat{\mu}', vech(\hat{\Sigma})')'. \quad (3.5)
\]

Also we introduce the matrices;
\[
\Omega_1 = 2\pi f(0), \quad (m \times m) - matrix
\]
\[
\Omega_2 = \left\{ 2\pi \int_{-\pi}^{\pi} \{f_{\alpha_1\alpha_3}(\lambda)f_{\alpha_2\alpha_4}(\lambda) + f_{\alpha_1\alpha_4}(\lambda)f_{\alpha_2\alpha_3}(\lambda)\}d\lambda \right\}, \quad (r \times r) - matrix,
\]
where \( r = m(m + 1)/2 \). We write the parameter space of \( \theta \) by \( \Theta \subset \mathbb{R}^{m+r} \).

Then, introduce a function \( g: \Theta \to \mathbb{R}^{m-1} \), which describes various portfolios.

**Assumption 2.** The function \( g(\theta) \) is continuously differentiable.

Then we have the following fundamental result.

**Theorem 1.** Under Assumptions 1 and 2,
\[
\sqrt{T}(g(\hat{\theta}) - g(\theta)) \Rightarrow N \left( 0, \left( \begin{array}{cc} \Omega_1 & 0 \\ 0 & \Omega_2 \end{array} \right) \right), \quad (as \; T \to \infty) \quad (3.6)
\]
where
\[
\left( \frac{\partial g}{\partial \theta} \right) = \left( \begin{array}{c} \frac{\partial g}{\partial \mu'} \quad \frac{\partial g}{\partial vech(\Sigma)'} \end{array} \right) = \left( \begin{array}{c} \frac{\partial g}{\partial \mu_1}, \ldots, \frac{\partial g}{\partial \mu_k}, \frac{\partial g}{\partial \sigma_{11}}, \ldots, \frac{\partial g}{\partial \sigma_{kk}} \end{array} \right).
\]

The proof of Theorem 1 will be given in Section 6. In Section 2 some concrete examples for \( g(\theta) \) were given in (2.2)',(2.4)',(2.6)' and (2.7)'. For these
where $g_k(\theta), k = 1, \ldots, 4$, their derivatives become

$$
\begin{align*}
\frac{\partial g_1}{\partial \mu_i} &= \frac{1}{2a} \left\{ \Sigma^{-1} e_i - \frac{e' \Sigma^{-1} \mu}{e' \Sigma^{-1} e} \right\}, \\
\frac{\partial g_1}{\partial \sigma_{ij}} &= -\frac{1}{2a} \left\{ \Sigma^{-1} E_{ij} \Sigma^{-1} e - \frac{e' \Sigma^{-1} E_{ij} \Sigma^{-1} e}{e' \Sigma^{-1} e} \right\} + \frac{e' \Sigma^{-1} \mu}{e' \Sigma^{-1} e} \left\{ \Sigma^{-1} E_{ij} \Sigma^{-1} e - \frac{e' \Sigma^{-1} E_{ij} \Sigma^{-1} e}{e' \Sigma^{-1} e} \right\}, \\
\frac{\partial g_2}{\partial \mu_i} &= 0, \quad \frac{\partial g_2}{\partial \sigma_{ij}} = -\frac{\Sigma^{-1} E_{ij} \Sigma^{-1} e}{e' \Sigma^{-1} e} + \frac{e' \Sigma^{-1} E_{ij} \Sigma^{-1} e}{(e' \Sigma^{-1} e)^2} \Sigma^{-1} e, \\
\frac{\partial g_3}{\partial \mu_i} &= \frac{1}{2a} \Sigma^{-1} e_i, \quad \frac{\partial g_3}{\partial \sigma_{ij}} = -\frac{1}{2a} \Sigma^{-1} E_{ij} \Sigma^{-1} (\mu - R_0 e), \\
\frac{\partial g_4}{\partial \mu_i} &= -\frac{1}{2a} e' \Sigma^{-1} e_i, \quad \frac{\partial g_4}{\partial \sigma_{ij}} = \frac{1}{2a} e' \Sigma^{-1} E_{ij} \Sigma^{-1} (\mu - R_0 e), \\
\end{align*}
$$

$(i, j = 1, \ldots, m)$

where $e_a = (0, \ldots, 1, 0, \ldots, 0)'$ is an $m \times 1$ vector with 1 at the $a$th position and $E_{ab} = e_a e_b'$.

4. Asymptotic Efficiency of the Estimators of the Optimal Portfolios

Fundamental results concerning the asymptotic efficiency of sample autocovariance matrices of vector Gaussian processes were obtained by Kakizawa(1999). He compared the asymptotic variance (AV) of the sample autocovariance matrices of vector Gaussian processes were obtained by Kakizawa(1999).

Now we return to the setting for $\{X(t)\}$ in Section 3. That is, $\{X(t)\}$ is a zero-mean Gaussian $m$-vector stationary process with spectral density matrix $f(\lambda)$, and satisfies Assumption 1.

Assumption 3.

(i) $f(\lambda)$ is parameterized by $\eta = (\eta_1, \ldots, \eta_q)' \in \mathcal{H} \subset \mathbb{R}^q$ i.e., $f_\eta = f_\eta(\lambda)$.

(ii) For $A^{(j)}(t) \equiv \int_{-\pi}^{\pi} \partial f_\eta(\lambda)/\partial \eta_j d\lambda$, $j = 1, \ldots, q, l \in \mathbb{Z}$, it holds that

$$
\sum_{j=-\infty}^{\infty} \|A^{(j)}(t)\|_E < \infty.
$$

(iii) $q \geq m(m + 1)/2$.

Assumption 4. There exists a positive constant $c$ (independent of $\lambda$) such that $f_\eta(\lambda) - cI_m$ is positive semi-definite, where $I_m$ is the $m \times m$ identity matrix.

Kakizawa(1999) showed that the limit of averaged Fisher information matrix is given by

$$
\mathcal{F}(\eta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \Delta(\lambda)^* [\{f_\eta(\lambda)^{-1}\}' \otimes f_\eta(\lambda)^{-1}] \Delta(\lambda) d\lambda \quad (4.1)
$$

where

$$
\Delta(\lambda) = (\text{vec}\{\partial f_\eta(\lambda)/\partial \eta_1\}, \ldots, \text{vec}\{\partial f_\eta(\lambda)/\partial \eta_q\}) \quad (m^2 \times q) - \text{matrix}.
$$
Assumption 5. The matrix $\mathcal{F}(\eta)$ is positive definite.

In view of the general asymptotic theory,

$$\text{INE} \equiv \det \left\{ \text{Asymptotic variance of } g(\hat{\theta}) \right\} - \mathcal{F}(\eta)^{-1} \geq 0.$$ 

In what follows we, numerically, investigate $\text{INE}$ for various spectral structures.

**Model I** (VMA(1) model). Let the return process be generated by

$$X(t) = \begin{pmatrix} 1 - \eta_1 B & 0 \\ 0 & 1 - \eta_1 B \end{pmatrix} \epsilon(t), \quad \epsilon(t) \sim \text{i.i.d. } \mathcal{N} \left( 0, \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix} \right),$$

where $B$ is the lag operator. For this we plotted the graph of $\text{INE} = \text{INE}(\text{VMA}(1))$ for $\eta_1 = -0.8(0.2)0.8$ in Figure 1. We can see that, as $|\eta_1|$ tends to 1, $\text{INE}$ increases.

**Figure 1** is about here.

**Model II** (VARMA(1,2) model). Let the return process be generated by

$$X(t) = \begin{pmatrix} 1 - \eta_1 B & 0 \\ 0 & 1 - \eta_1 B \end{pmatrix} \begin{pmatrix} 1 - \eta_2 B & 0 \\ 0 & 1 - \eta_2 B \end{pmatrix} \epsilon(t), \quad \epsilon(t) \sim \text{i.i.d. } \mathcal{N} \left( 0, \begin{pmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{pmatrix} \right),$$

(4.2)

For this model we plotted the graph of $\text{INE} = \text{INE}(\text{VARMA}(1,2))$ for $\eta_1 = -0.8(0.2)0.8, \eta_2 = 0.01, \eta_3 = 0.5$ in Figure 2. We can see that if $\eta_1 \searrow -1$, $\text{INE}$ becomes quite large.

**Figure 2** is about here.

**Model III** (VARMA(1,2) model). Let the model be generated by (4.5) with $\eta_1 = 0.01$ and $\eta_2 = 0.5$. For this we plotted the graph of $\text{INE} = \text{INE}(\text{VARMA}(1,2))$ for $\eta_3 = -0.8(0.2)0.8$ in Figure 3. We can see that as $\eta_3 \nearrow 1$, $\text{INE}$ increases.

**Figure 3** is about here.

Summarizing the above we observe that

(i) If $X(t)$ is VMA(1) model, $\text{INE}$ increases as the absolute value of the MA coefficient $\eta_1$ tends to 1.

(ii) If $X(t)$ is VARMA(1,2) model with the MA coefficient $\eta_2 \approx 0$, $\text{INE}$ increases as the AR coefficient $\eta_1$ tends to $-1$. 

6
(iii) If $X(t)$ is VARMA(1,2) model with the AR coefficient $\eta_1 \approx 0$, $INE$ increases as the MA coefficient $\eta_3$ tends to 1.

Although we just examined a few examples of dependent returns, the above studies illuminate inefficiency of the usual portfolio estimators. Therefore it should be noted that the degree of inefficiency becomes quite large if some parameters tend to a boundary value.

5. Construction of efficient estimators

5.1. A quasi-Gaussian maximum likelihood estimator

Let $\{X(t)\}$ be the linear process defined by Sections 3 and 4 with spectral density matrix $f_\eta(\lambda), \eta \in \mathcal{H}$. We assume $\eta = \text{vech}\{R(0)\}$. Denote by $I_x(\lambda)$, the periodogram matrix constructed from a partial realization $\{X(1), \ldots, X(T)\}$;

$$I_x(\lambda) = F_x F_x^*, \text{ with } F_x(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} X(t) e^{it\lambda}, -\pi \leq \lambda \leq \pi. \quad (5.1)$$

In view of Hosoya and Taniguchi(1982) we introduce

$$D(\eta, I_x) = \int_{-\pi}^{\pi} [\log \det f_\eta(\lambda) + \text{tr} f_\eta^{-1}(\lambda) I_x(\lambda)] d\lambda. \quad (5.2)$$

A quasi-Gaussian maximum likelihood estimator $\hat{\eta}$ of $\eta$ is given by

$$\hat{\eta} = \arg \min_{\eta \in \mathcal{H}} D(\eta, I_x). \quad (5.3)$$

Hosoya and Taniguchi(1982) showed that

(i) $p - \lim T \to \infty \hat{\eta} = \eta,$

(ii) $\sqrt{T}(\hat{\eta} - \eta) \overset{d}{\to} N(0, F(\eta)^{-1}).$

Therefore, $\hat{\eta}$ is Gaussian asymptotically efficient, hence $g(\hat{\theta}),$ with $\hat{\theta} = (\hat{\mu}', \hat{\eta}')'$, is asymptotically efficient.

Since the solution of $\partial D(\eta, I_x)/\partial \eta = 0$ is generally nonlinear with respect to $\eta$, we use the Newton-Raphson iteration procedure. A feasible procedure is

$$\hat{\eta}^{(1)} = \text{vech}(\hat{\Sigma})$$

$$\hat{\eta}^{(k)} = \hat{\eta}^{(k-1)} - \left[ \frac{\partial^2 D(\eta, I_x)}{\partial \eta \partial \eta'} \right]^{-1} \frac{\partial D(\eta, I_x)}{\partial \eta} \bigg|_{\eta=\hat{\eta}^{(k-1)}} (k \geq 2) \quad (5.5)$$

From Hosoya and Taniguchi(1982) and Taniguchi and Kakizawa(2000) it is seen that $\hat{\eta}^{(2)}$ is asymptotically efficient. Therefore, $g(\hat{\theta}), \hat{\theta} = (\hat{\mu}', \hat{\eta}^{(2)}')'$, becomes asymptotically efficient. In calculating (5.5), it follows that

$$\frac{\partial D(\eta, I_x)}{\partial \eta_{ij}} = \frac{1}{2\pi} \int_{-\pi}^{\pi} tr \left[ f_\eta(\lambda)^{-1} E_{ij} (I_m - f_\eta(\lambda)^{-1} I_x(\lambda)) \right] d\lambda, \quad (5.6)$$
\[
\frac{\partial^2 D(f_\eta, I_x)}{\partial \eta_{ij} \partial \eta_{kl}} = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \text{tr}[-f_\eta(\lambda)^{-1} E_{kl} f_\eta(\lambda)^{-1} E_{ij} \left(I_m - f_\eta(\lambda)^{-1} I_x(\lambda)\right)] \\
+ f_\eta(\lambda)^{-1} E_{ij} f_\eta(\lambda)^{-1} E_{kl} f_\eta(\lambda)^{-1} I_x(\lambda)] d\lambda,
\]

where \(\eta_{ij}\) is the \((i,j)\)th element of \(R(0)\). To make the step (5.5) feasible, we replace \(f_\eta\) in (5.6) and (5.7) by a nonparametric spectral estimator

\[
\hat{f}_\eta(\lambda) = \int_{-\pi}^{\pi} W_T(\lambda - \mu) I_x(\mu) d\mu.
\]

Here \(W_T(\theta)\) is a class of functions of the form

\[
\hat{f}_\eta(\lambda) = \int_{-\pi}^{\pi} W_T(\lambda - \mu) I_x(\mu) d\mu.
\]

where \(W(\cdot)\) and \(M\) satisfy the following:

(W1) \(W(\cdot)\) is a real, bounded, nonnegative, even function with

\[
\int_{-\infty}^{\infty} W(\theta) d\theta = 1, \int_{-\infty}^{\infty} \theta^2 W(\theta) d\theta = \kappa_2 (0 < \kappa_2 < \infty).
\]

(W2) \(w(x) = \int_{-\infty}^{\infty} W(\theta) e^{i\theta x} d\theta\) satisfies \(|w(x)| \leq \bar{w}(x)\), where \(\bar{w}(x)\) is even, integrable, and monotonically decreasing on \([0, \infty)\).

(M) \(M > 0\) depends on \(T\) in such a way that \(M/T^{1/2} + T^\nu/M \to 0\) as \(T \to \infty\), where \(\nu\) is a nonnegative number.

Then, from Taniguchi and Kakizawa(2000) it follows that

\[
\max_{\lambda \in [-\pi, \pi]} \|\hat{f}(\lambda) - f(\lambda)\|_{L^2} \overset{P}{\to} 0,
\]

and that

\[
\hat{f}(\lambda) = f(\lambda) + O_P\left(\frac{1}{\sqrt{T}B_T}\right),
\]

where \(B_T > 1\) and \(B_T \not\to \infty\), which implies that the convergence rate (consistency order) of \(\hat{f}(\lambda)\) is smaller than \(O(\sqrt{T})\). But if we integrate \(\hat{f}(\lambda)\), the \(\sqrt{T}\)-consistency is recovered, i.e.

\[
\int \Psi(\hat{f}(\lambda)\lambda) d\lambda = \int \Psi(f(\lambda)) d\lambda + O_P\left(\frac{1}{\sqrt{T}}\right),
\]

where \(\Psi(\cdot)\) is a continuous function. Hence,

\[
\frac{\partial D(f_\eta, I_x)}{\partial \eta} \bigg|_{f_\eta=f} = \frac{\partial D(f_\eta, I_x)}{\partial \eta} + O_P\left(\frac{1}{\sqrt{T}}\right),
\]

\[
\frac{\partial^2 D(f_\eta, I_x)}{\partial \eta^2} \bigg|_{f_\eta=f} = \frac{\partial^2 D(f_\eta, I_x)}{\partial \eta^2} + O_P\left(\frac{1}{\sqrt{T}}\right).
\]
It is possible to drop the Gaussian assumption of \( \{X(t)\} \). In fact, in view of Lemma A2.3 of Hosoya and Taniguchi(1982), the statement (6.4) becomes

\[
\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \Omega), \tag{5.15}
\]

where \( \Omega = \begin{pmatrix} \Omega_1 & 0 \\ 0 & \Omega_2 + \Omega_{2N}^G \end{pmatrix} \). Here the explicit form of \( \Omega_{2N}^G \) is given by (6.11) of Hosoya and Taniguchi(1982), and depends on the non-Gaussianity of \( \{X(t)\} \). Hence Theorem 1 becomes

\[
\sqrt{T} \left\{ g(\hat{\theta}) - g(\theta) \right\} \xrightarrow{d} N(0, (\partial g/\partial \theta)' \Omega (\partial g/\partial \theta)'). \tag{5.16}
\]

However, if we discuss the asymptotic efficiency of \( g(\hat{\theta}) \), it is required to develop the asymptotic efficient estimation theory based on the local asymptotic normality (LAN) for non-Gaussian processes. The discussion is much involved, so we make it in a future work.

6. Proof

This section provides the proof of Theorem 1.

Proof of Theorem 1

From Hannan(1970,Theorem 11,p.221) it follows that

\[
\sqrt{T} \{\hat{\mu} - \mu\} \xrightarrow{d} N(0, \Omega_1), \text{ as } T \to \infty. \tag{6.1}
\]

The result

\[
\sqrt{T} \left\{ \text{vech}(\hat{\Sigma}) - \text{vech}(\Sigma) \right\} \xrightarrow{d} N(0, \Omega_2), \text{ as } T \to \infty, \tag{6.2}
\]

follows from Theorem 2.2 of Hosoya and Taniguchi (1982). Since we assume the Gaussianity of \( \{X(t)\} \), it is easily seen that

\[
\lim_{T \to \infty} \text{Cov}[\sqrt{T}(\hat{\mu} - \mu), \sqrt{T}(\text{vech}(\hat{\Sigma}) - \text{vech}(\Sigma))] = 0, \text{ (zero - matrix)} \tag{6.3}
\]

Combining the results by Hannan(1970) and Hosoya and Taniguchi(1982), we can check the joint asymptotic normality of \( \sqrt{T} \{\hat{\mu} - \mu\} \) and \( \sqrt{T} \left\{ \text{vech}(\hat{\Sigma}) - \text{vech}(\Sigma) \right\} \).

Hence,

\[
\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \Omega) \text{ (as } T \to \infty), \tag{6.4}
\]

where \( \Omega = \begin{pmatrix} \Omega_1 & 0 \\ 0 & \Omega_2 \end{pmatrix} \). By use of the \( \delta - \text{method} \) (e.g.,Brockwell and Davis (1991,Proposition 6.4.3)) for (6.4), we observe that

\[
\sqrt{T} \left\{ g(\hat{\theta}) - g(\theta) \right\} \xrightarrow{d} N(0, (\partial g/\partial \theta)' \Omega (\partial g/\partial \theta)'). \text{ (as } T \to \infty). \tag{6.5}
\]

Acknowledgments

The author would like to thank Professor Masanobu Taniguchi Waseda University for his encouragement, guidance, and advice for this paper.
References


Figure 1: Model I: $\eta_1 = -0.8\, (0.2)0.8$

Figure 2: Model II: $\eta_1 = -0.8\,(0.2)0.8$, $\eta_2 = 0.01$, $\eta_3 = 0.5$
Figure 3: Model III: \( \eta_1 = 0.01, \eta_2 = 0.5, \eta_3 = -0.8(0.2)0.8 \)

HIROSHI SHIRAISHI
DEPARTMENT OF MATHEMATICAL SCIENCES
SCHOOL OF SCIENCE AND ENGINEERING
WASEDA UNIVERSITY
3-4-1 OKUBO SHINJUKU-KU
TOKYO 169-8555
JAPAN
E-MAIL: h.shiraishi@moegi.waseda.jp