Preliminary Test Estimation for Spectra and Its Applications to Financial Hedging Problem

Yusuke Maeyama*
Kenichiro Tamaki†
Masanobu Taniguchi ‡
Department of Applied Mathematics,
School of Fundamental Science and Engineering,
Waseda University, Tokyo, 169-8555, Japan

Summary. For a general non-Gaussian stationary linear process, quasi-maximum likelihood estimation of a subset of the parameters of the spectral density is considered when the complementary subset is suspected to be redundant. A preliminary test quasi-maximum likelihood estimator (q-MLE) of parameters is introduced and, in the light of its mean square error, is compared with the restricted and unrestricted q-MLE.

Key words. ARMA model; generalized error distribution; market efficiency hypothesis; preliminary test; quasi-maximum likelihood estimator; spectral density; stationary linear process.

AMS2000. 62F12; 62M15; 91B28

1 Introduction

In linear regression models, estimation of a subset of parameters is considered when the complementary subset is plausibly redundant. Saleh and Sen [7] considered preliminary test estimator in a non parametric error distribution set-up and obtained restricted, unrestricted and preliminary test estimators using rank estimators which are robust in this situation. Later on, Sen [11], Saleh and Sen [7] and Sen and Saleh [8] treated preliminary test and shrinkage estimation for least squares estimators, maximum likelihood estimators and M-estimators in a linear model with nonparametric independent and identical error distribution. Moreover, Mukherjee [6] discussed shrinkage estimations in linear models with long-memory errors.

*E-mail:maeyama@ruri.waseda.jp
†k-tamaki@toki.waseda.jp
‡taniguchi@waseda.jp

These discussions on shrinkage estimations can be extended to linear time series models. Saleh [9] dealt with a shrinkage estimation for the parameters of autoregressive Gaussian process setting Whittle likelihood ratio as the preliminary test. The asymptotic theory is based on that of Dzhaparidze [2]

In this paper, we consider a general non-Gaussian linear process which includes the usual ARMA processes, and address the problem of estimation for subset of spectral parameters when the complementary subset is suspected to be redundant. A preliminary test quasi-maximum likelihood estimator (q-MLE) and restricted and unrestricted q-MLE's are introduced. When the complementary parameter is contiguous to zero vector, the mean square errors of these estimators are evaluated. Then we observe that the preliminary test estimator is relatively effective among the three ones.

The paper is organized as follows. Section 2 describes a class of vector-valued non-Gaussian linear stationary processes. Section 3 elucidates the asymptotic behaviors of quasi-maximum likelihood estimators and preliminary test quasi-maximum likelihood estimator. Section 4 addresses the problem of comparative performance of the estimators with respect to mean square error under local alternatives, especially when the model is ARMA(1,1). Numerical studies illuminate that the preliminary estimator is relatively effective under a sequence of local alternatives. In Section 5, we mention its applications to financial hedging problem based of preliminary test for efficient market hypothesis.

2 Settings and Problems

Consider the m-vector linear process

$$\mathbf{X}(t) = \sum_{j=0}^{\infty} A_\theta(j)\mathbf{U}(t-j), \quad t \in \mathbf{Z} \tag{2.1}$$

where the $\mathbf{U}(t)$ are i.i.d. m-vector random variables with probability density $p(\mathbf{u}) > 0$ on $\mathbf{R}^m$, and $A_\theta(j) = \{A_{\theta,ab}(j); a, b = 1, \cdots, m\}, j \in \mathbf{Z}$, are $m \times m$ matrices depending on a parameter vector $\theta = (\theta_1, \cdots, \theta_q) \in \Theta \subset \mathbf{R}^q$. Here the coefficients $\{A_\theta(j)\}$ and innovation density $p(\cdot)$ of $\mathbf{U}(t)$ are required to satisfy the following assumption.

(A1) (i) There exists $0 < \rho_A < 1$ so that

$$\| A_\theta(j) \| = \mathcal{O}(\rho_A^j), \quad j \in \mathbf{N}, \tag{2.2}$$

where $\| A_\theta(j) \|$ denotes the sum of absolute values of the entries of $A_\theta(j)$.

(ii) Every $A_\theta(j) = \{A_{\theta,ab}(j)\}$ is continuously two times differentiable with respect to θ, and the derivatives satisfy

$$|\partial_{i_1} \cdots \partial_{i_k} A_{\theta,ab}(j)| = \mathcal{O}(\gamma_A^j), \quad k = 1, 2, \ j \in \mathbf{N} \tag{2.3}$$

for some $0 < \gamma_A < 1$ and for $a, b = 1, \cdots, m$.

(iii) Every $\partial_{i_1} \cdots \partial_{i_k} A_{\theta,ab}(j)$ satisfies the Lipschitz condition for all $i_1, i_2 = 1, \cdots q$ and $j \in \mathbf{N}$.

(iv) $\det \left\{ \sum_{j=0}^{\infty} A_\theta(j) z^j \right\} \neq 0$ for $|z| \leq 1$ and $\left\{ \sum_{j=0}^{\infty} A_\theta(j) z^j \right\}^{-1}$ has the power series expansion

$$\left\{ \sum_{j=0}^{\infty} A_\theta(j) z^j \right\}^{-1} = B_\theta(0) + B_\theta(1) z + B_\theta(2) z^2 + \cdots, \quad |z| \leq 1,$$

where $\| B_\theta(j) \| = \mathcal{O}(\rho_B^j), \ j \in \mathbf{N}$, for some $0 < \rho_B < 1$.

(v) Every $B_\theta(j) = \{B_{\theta,ab}\}$ is continuously two times differentiable with respect to θ, and the derivatives satisfy

$$|\partial_{i_1} \cdots \partial_{i_k} B_{\theta,ab}(j)| = \mathcal{O}(\gamma_B^j), \quad k = 1, 2, \ j \in \mathbf{N} \tag{2.4}$$

for some $0 < \gamma_B < 1$ and for $a, b = 1, \cdots, m$.

(vi) Each $\partial_{i_1} \cdots \partial_{i_k} B_{\theta,ab}(j)$ satisfies the Lipschitz condition for all $i_1, i_2 = 1, \cdots q$ and $j \in \mathbf{N}$.

(A2) (i)

$$\lim_{\|u\| \to \infty} p(\mathbf{u}) = 0, \quad \int \mathbf{u} p(\mathbf{u}) d\mathbf{u} = \mathbf{0},$$

$$\int \mathbf{u}\mathbf{u}' p(\mathbf{u}) d\mathbf{u} = I_m \text{ and } \int \|\mathbf{u}\|^4 p(\mathbf{u}) d\mathbf{u} < \infty,$$

where I_m is the $m \times m$ identity matrix.

(ii) The derivatives Dp and $D^2 p \equiv D(Dp)$ exist on $\mathbf{R}^n$, and every component of $D^2 p$ satisfies the Lipschitz condition.

(iii)

$$\int \| \phi(\mathbf{u}) \|^4 p(\mathbf{u}) d\mathbf{u} < \infty \text{ and } \int D^2 p(\mathrm{u}) = 0,$$

where $\phi(\mathbf{u}) = p(u)^{-1} Dp(u)$.

Then, $\{\mathbf{X}(t)\}$ is a stationary process with spectral density matrix

$$\mathbf{f}_\theta(\lambda) = \frac{1}{2\pi} \left\{ \sum_{j=0}^{\infty} A_\theta(j) e^{ij\lambda} \right\} \left\{ \sum_{j=0}^{\infty} A_\theta(j) e^{ij\lambda} \right\}^* . \tag{2.5}$$

The class of $\{\mathbf{X}(\mathbf{t})\}$ includes that of non-Gaussian vector-valued causal ARMA models, so the class is sufficiently rich.

Let $\mathbf{X}(1), \cdots, \mathbf{X}(n)$ be n observations from the process $\{X(t)\}$. Partition θ as follows:

$$\theta = (\theta_1, \theta_2) = (\theta_{11}, \cdots, \theta_{1q_1}, \theta_{21}, \cdots, \theta_{2q_2}), \quad q_i \geq 1, \ i = 1, 2, \ q_1 + q_2 = q. \tag{2.6}$$

We are now interested in the estimation of θ_1 in the case when θ_2 is close to $\mathbf{0}$. If $\theta_2 = \mathbf{0}$, then we have the restricted model

$$\mathbf{X}(t) = \sum_{j=0}^{\infty} A_{(\theta_1,\mathbf{0})}(j)\mathbf{U}(t-j), \quad t \in \mathbf{Z}. \tag{2.7}$$

3 Preliminary Test Estimator

From (A1) the model (2.1) is representable as

$$\sum_{j=0}^{\infty} B_\theta(j)\mathbf{X}(t-j) = \mathbf{U}(t). \tag{3.1}$$

From (2.1) and (3.1) the likelihood function of $\{\mathbf{U}(s),\ s \le 0,\ \mathbf{X}(1), \cdots, \mathbf{X}(n)\}$ is given by

$$dQ_{n,\theta} = \prod_{t=1}^{n} p\left\{\sum_{j=0}^{t-1} B_\theta(j)\mathbf{X}(t-j) + \sum_{r=0}^{\infty} C_\theta(r,t)\mathbf{U}(-r)\right\} dQ_{\mathbf{u}} \tag{3.2}$$

where $Q_{\mathbf{u}}$ is the probability distribution of $\{\mathbf{U}(s),\ s \le 0\}$, and

$$C_\theta(r,t) = \sum_{r'=0}^{r} B_\theta(r'+t)A_\theta(r-r').$$

Since $\mathbf{U}(s),\ s \le 0$, are unobservable we use the quasi-likelihood

$$L_n(\theta) = \prod_{t=1}^{n} p\left\{\sum_{j=0}^{t-1} B_\theta(j)\mathbf{X}(t-j)\right\} \tag{3.3}$$

for estimation of $\theta = (\theta_1, \theta_2)$. Hence, the quasi-maximum likelihood estimator (QML) θ_n of θ is a solution of the equation

$$\frac{\partial}{\partial\theta}\left[\sum_{t=1}^{n} \log p\left\{\sum_{j=0}^{t-1} B_\theta(j)\mathbf{X}(t-j)\right\}\right] = \mathbf{0} \tag{3.4}$$

with respect to θ. Let

$$\mathcal{F}(p) = \int \phi(\mathbf{u})\phi(\mathbf{u})' p(\mathbf{u}) d\mathbf{u} \quad \text{(Fisher information)},$$

and

$$R(j) = \mathsf{E}[X_t X_{t+j}], \quad j \in \mathbf{Z}.$$

Then there exists a statistic $\tilde{\theta}_n$ that solves (3.4) such that

$$\sqrt{n}\begin{pmatrix} \tilde{\theta}_{1n} - \theta_1 \\ \tilde{\theta}_{2n} - \theta_2 \end{pmatrix} = \Gamma^{-1}\begin{pmatrix} \Delta_{1n} \\ \Delta_{2n} \end{pmatrix} + o_p(1) \tag{3.5}$$

where

$$\begin{pmatrix} \Delta_{1n} \\ \Delta_{2n} \end{pmatrix} \xrightarrow{\mathrm{d}} \mathcal{N}(\mathbf{0}, \Gamma) \tag{3.6}$$

and

$$\Gamma = \left(\mathrm{tr} \left[\mathcal{F}(p) \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \frac{\partial}{\partial \theta_j} B_\theta(i_1) R(i_1 - i_2) \frac{\partial}{\partial \theta_k} B_\theta(i_2) \right] \right)_{j,k=1,\cdots,q}. \tag{3.7}$$

is dependent on the spectral density matrix $\mathbf{f}_\theta(\lambda)$(e.g. Taniguchi and Kakizawa [16]). We call $\tilde{\theta}_n$ the unrestricted QML of θ If $\mathcal{F}(p) = \gamma I_m$ (γ:a positive constant), then Γ is rewritten as

$$\Gamma = \frac{1}{4\pi} \left(\mathrm{tr} \left[\mathcal{F}(p) \int_{-\pi}^{\pi} \mathbf{f}_\theta(\lambda) \frac{\partial}{\partial \theta_j} \{\mathbf{f}_\theta\}^{-1} \mathbf{f}_\theta(\lambda) \frac{\partial}{\partial \theta_k} \{\mathbf{f}_\theta\}^{-1} d\lambda \right] \right)_{j,k=1,\cdots,q}. \tag{3.8}$$

Write the decomposition of Γ as

$$\Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \tag{3.9}$$

where Γ_{ij} is $q_i \times q_j$ matrix for $i, j = 1, 2$. Using this decomposition and

$$\Gamma^{-1} = \begin{pmatrix} \Gamma_{11}^{-1} + \Gamma_{11}^{-1}\Gamma_{12}\Gamma_{22\cdot1}^{-1}\Gamma_{21}\Gamma_{11}^{-1} & -\Gamma_{11}^{-1}\Gamma_{12}\Gamma_{22\cdot1}^{-1} \\ -\Gamma_{22\cdot1}^{-1}\Gamma_{21}\Gamma_{11}^{-1} & \Gamma_{22\cdot1}^{-1} \end{pmatrix}, \tag{3.10}$$

(3.5) can be rewritten as

$$\sqrt{n} \begin{pmatrix} \tilde{\theta}_1 - \theta_1 \\ \tilde{\theta}_2 - \theta_2 \end{pmatrix} = \begin{pmatrix} \Gamma_{11}^{-1} \\ \mathbf{0} \end{pmatrix} \Delta_{1n} + \begin{pmatrix} \Gamma_{11}^{-1}\Gamma_{12}\Gamma_{22\cdot1}^{-1} \\ \Gamma_{22\cdot1}^{-1} \end{pmatrix} \left(\Delta_{2n} - \Gamma_{21}\Gamma_{11}^{-1}\Delta_{1n} \right) + o_p(1) \tag{3.11}$$

where

$$\Gamma_{22\cdot1} = \Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12}. \tag{3.12}$$

In the same way as the above, we get $\hat{\theta}_{1n}$ that solves (3.4) with respect to $(\theta_1, \mathbf{0})$, and show that

$$\sqrt{n} \left(\hat{\theta}_{1n} - \theta_1 \right) = \Gamma_{11}^{-1}\Delta_{1n} + o_p(1). \tag{3.13}$$

We call $\hat{\theta}_{1n}$ the restricted QML of θ_1.

To introduce a compromised estimator of θ_1 under $\theta_2 \approx 0$ we define a test statistic $\mathcal{L}_n$ to test the null-hypothesis $H_0 : \theta_2 = \mathbf{0}$ by

$$\mathcal{L}_n = 2 \log \frac{L_n(\tilde{\theta}_{1n}, \tilde{\theta}_{2n})}{L_n(\hat{\theta}_{1n}, \mathbf{0})}. \tag{3.14}$$

From Taylor's formula (see Fuller [3] or Taniguchi and Amano [15]),

$$\log \frac{L_n(\tilde{\theta}_{1n}, \tilde{\theta}_{2n})}{L_n(\theta_{1n}, \mathbf{0})} = \frac{1}{2} \begin{pmatrix} \Delta_{1n} \\ \Delta_{2n} \end{pmatrix}' \Gamma^{-1} \begin{pmatrix} \Delta_{1n} \\ \Delta_{2n} \end{pmatrix} + o_p(1) \tag{3.15}$$

and

$$\log \frac{L_n(\hat{\theta}_{1n}, \mathbf{0})}{L_n(\theta_{1n}, \mathbf{0})} = \frac{1}{2}\Delta_{1n}'\Gamma_{11}^{-1}\Delta_{1n} + o_p(1) \tag{3.16}$$

under H_0. Hence, from (3.5) and (3.6) it is seen that

$$\mathcal{L}_n = (\Delta_{2n} - \Gamma_{21}\Gamma_{11}^{-1}\Delta_{1n})'\Gamma_{22\cdot 1}^{-1}(\Delta_{2n} - \Gamma_{21}\Gamma_{11}^{-1}\Delta_{1n}) + o_p(1) \xrightarrow{\text{d}} \chi^2_{q_2} \tag{3.17}$$

under H_0 where $\chi^2_{q_2}$ is a chi-square distribution with q_2 degrees of freedom (d.f.). Thus, we choose the α-level critical value $\chi^2_{q_2,\alpha}$ and define the preliminary test quasi-maximum likelihood estimator $\hat{\theta}_{1n}^{PT}$ by

$$\hat{\theta}_{1n}^{PT} = \tilde{\theta}_{1n} - (\tilde{\theta}_{1n} - \hat{\theta}_{1n})\, I\left(\mathcal{L}_n \leq \chi^2_{q_2,\alpha}\right) \tag{3.18}$$

where $I(\cdot)$ is the indicator function.

4 Asymptotic property of the estimators

Since we are interested in estimation θ_1 in the case when $\theta_2 \approx 0$, we consider the asymptotic behavior of $\tilde{\theta}_{1n}$, $\hat{\theta}_{1n}$ and $\hat{\theta}_{1n}^{PT}$ under local alternatives

$$A_n : \theta_2 = \mathbf{h}/\sqrt{n}, \quad \mathbf{h} \in \mathbf{R}^{q_2} \text{ fixed.} \tag{4.1}$$

First, we obtain,

Lemma 4.1. *Under A_n,*

$$\begin{pmatrix} \sqrt{n}\,(\hat{\theta}_{1n} - \theta_1) \\ \sqrt{n}\,(\tilde{\theta}_{1n} - \theta_1) \\ \Delta_{21} - \Gamma_{21}\Gamma_{11}^{-1}\Delta_{1n} \end{pmatrix} \xrightarrow{\text{d}}$$

$$\mathcal{N}\left(\begin{pmatrix} \Gamma_{11}^{-1}\Gamma_{12}\mathbf{h} \\ \mathbf{0} \\ \Gamma_{22\cdot 1}\mathbf{h} \end{pmatrix}, \begin{pmatrix} \Gamma_{11}^{-1} & \Gamma_{11}^{-1} & \mathbf{0} \\ \Gamma_{11}^{-1} & \Gamma_{11\cdot 2}^{-1} & -\Gamma_{11}^{-1}\Gamma_{12} \\ \mathbf{0} & -\Gamma_{21}\Gamma_{11}^{-1} & \Gamma_{22\cdot 1} \end{pmatrix}\right). \tag{4.2}$$

Proof. From LAN theorem (see Hallin, et al.[4] or Taniguchi and Kakizawa [16]), the log-likelihood ratio between H_0 and A_n given by

$$\Lambda_n \equiv \log \frac{L_n(\theta_1, \mathbf{h}/\sqrt{n})}{L_n(\theta_1, \mathbf{0})}$$

has the stochastic expansion

$$\Lambda_n = \mathbf{h}'\Delta_{2n} - \frac{1}{2}\mathbf{h}'\Gamma_{22}\mathbf{h} + o_p(1). \tag{4.3}$$

It follows from (3.11), (3.13), and (4.3) that

$$\begin{pmatrix} \sqrt{n}\,(\hat{\theta}_{1n} - \theta_1) \\ \sqrt{n}\,(\tilde{\theta}_{1n} - \theta_1) \\ \Delta_{21} - \Gamma_{21}\Gamma_{11}^{-1}\Delta_{1n} \\ \Lambda_n \end{pmatrix} \xrightarrow{\text{d}} \mathcal{N}(\mu, \Sigma) \tag{4.4}$$

under H_0 where

$$\mu = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \frac{1}{2}\mathbf{h}'\Gamma_{22}\mathbf{h} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Gamma_{11}^{-1} & \Gamma_{11}^{-1} & \mathbf{0} & \Gamma_{11}^{-1}\Gamma_{12}\mathbf{h} \\ \Gamma_{11}^{-1} & \Gamma_{11\cdot 2}^{-1} & -\Gamma_{11}^{-1}\Gamma_{12} & \mathbf{0} \\ \mathbf{0} & -\Gamma_{21}\Gamma_{11}^{-1} & \Gamma_{22\cdot 1} & \Gamma_{22\cdot 1}\mathbf{h} \\ \mathbf{h}'\Gamma_{21}\Gamma_{11}^{-1} & \mathbf{0} & \mathbf{h}'\Gamma_{22\cdot 1} & \mathbf{h}'\Gamma_{22}\mathbf{h} \end{pmatrix}. \tag{4.5}$$

Hence, by LeCam's third lemma we obtain (4.2). □

Lemma 1 leads to the following conclusions under A_n:

(i) $\sqrt{n}\,(\hat{\theta}_{1n} - \theta)$ and $\mathcal{L}_n$ are asymptotically independent.

(ii)

$$\mathcal{L}_n \xrightarrow{\ \mathrm{d}\ } \chi^2_{q_2}(\Delta) \tag{4.6}$$

where $\Delta = \mathbf{h}'\Gamma_{22\cdot 1}\mathbf{h}$, and $\chi^2_{q_2}(\Delta)$ is a noncentral chi-squared distribution with q_2 d.f. and noncentrality parameter Δ.

(iii)

$$\begin{aligned} \sqrt{n}\,(\tilde{\theta}_{1n} - \theta_1) \ \text{ given } \ \Delta_{2n} - \Gamma_{21}\Gamma_{11}^{-1}\Delta_{1n} &= \Gamma_{22\cdot 1}(\mathbf{z} + \mathbf{h}) \\ &\xrightarrow{\ \mathrm{d}\ } \mathcal{N}\left(-\Gamma_{11}^{-1}\Gamma_{12}\mathbf{z}, \Gamma_{11}^{-1}\right). \end{aligned} \tag{4.7}$$

From the above it is shown that

$$\begin{aligned} &\lim_{n\to\infty} P\left\{\sqrt{n}\,(\hat{\theta}_{1n} - \theta_1) \le \mathbf{x},\ \mathcal{L}_n < \chi^2_{q_2,\alpha} | A_n\right\} \\ &\quad = \Phi_{q_1}(\mathbf{x} - \Gamma_{11}^{-1}\Gamma_{12}\mathbf{h} : \mathbf{0}, \Gamma_{11}^{-1}) H_{q_2}(\chi^2_{q_2,\alpha} : \Delta) \end{aligned} \tag{4.8}$$

and

$$\begin{aligned} &\lim_{n\to\infty} P\left\{\sqrt{n}\,(\tilde{\theta}_{1n} - \theta_1) \le \mathbf{x},\ \mathcal{L}_n \ge \chi^2_{p_2,\alpha} | A_n\right\} \\ &\quad = \int_{(\mathbf{z}+\mathbf{h})'\Gamma_{22\cdot 1}(\mathbf{z}+\mathbf{h}) \ge \chi^2_{q_2,\alpha}} \Phi_{q_1}(\mathbf{x} + \Gamma_{11}^{-1}\Gamma_{12}\mathbf{z} : \mathbf{0}, \Gamma_{11}^{-1})\, d\Phi_{q_2}(\mathbf{z} : \mathbf{0}, \Gamma_{22\cdot 1}^{-1}) \end{aligned} \tag{4.9}$$

where $\Phi_m(\mathbf{x} : \mu, \Sigma)$ is the m-dimensional normal cumulative distribution function (CDF) with mean μ and variance Σ, and $H_m(x : \Delta)$ is the CDF of a noncentral chi-squared distribution with m d.f. and noncentrality parameter Δ. Thus, we get, as the asymptotic CDF of $\sqrt{n}\,(\hat{\theta}^{PT}_{1n} - \theta_1)$

$$\begin{aligned} G(\mathbf{x}) &= \Phi_{q_1}(\mathbf{x} - \Gamma_{11}^{-1}\Gamma_{12}\mathbf{h} : \mathbf{0}, \Gamma_{11}^{-1}) H_{q_2}(\chi^2_{q_2,\alpha} : \Delta) \\ &\quad + \int_{(\mathbf{z}+\mathbf{h})'\Gamma_{22\cdot 1}(\mathbf{z}+\mathbf{h}) \ge \chi^2_{q_2,\alpha}} \Phi_{q_1}(\mathbf{x} + \Gamma_{11}^{-1}\Gamma_{12}\mathbf{z} : \mathbf{0}, \Gamma_{11}^{-1})\, d\Phi_{q_2}(\mathbf{z} : \mathbf{0}, \Gamma_{22\cdot 1}^{-1}). \end{aligned} \tag{4.10}$$

In what follows, denote, by $B(\theta_n)$ and $M(\theta_n)$, the bias and mean square error (MSE) of θ_n under A_n, respectively.

Proposition 4.1. *We have*

(i) $B(\tilde{\theta}_{1n}) = \mathbf{0}$ and $M(\tilde{\theta}_{1n}) = \Gamma_{11}^{-1} + \Gamma_{11}^{-1}\Gamma_{12}\Gamma_{22\cdot 1}^{-1}\Gamma_{21}\Gamma_{11}^{-1}$. (4.11)

(ii) $B(\hat{\theta}_{1n}) = -\Gamma_{11}^{-1}\Gamma_{12}\mathbf{h}$ and $M(\hat{\theta}_{1n}) = \Gamma_{11}^{-1} + \Gamma_{11}^{-1}\Gamma_{12}\mathbf{h}\mathbf{h}'\Gamma_{21}\Gamma_{11}^{-1}$. (4.12)

(iii) $B(\hat{\theta}_{1n}^{PT}) = -\Gamma_{11}^{-1}\Gamma_{12}\mathbf{h}H_{q_2+2}(\chi^2_{q_2,\alpha} : \Delta)$ and (4.13)

$$\begin{aligned} M(\hat{\theta}_{1n}^{PT}) &= \Gamma_{11}^{-1} + \Gamma_{11}^{-1}\Gamma_{12}\Gamma_{22\cdot 1}^{-1}\Gamma_{21}\Gamma_{11}^{-1}\left\{1 - H_{q_2+2}(\chi^2_{q_2,\alpha} : \Delta)\right\} \\ &\quad + \Gamma_{11}^{-1}\Gamma_{12}\mathbf{h}\mathbf{h}'\Gamma_{21}\Gamma_{11}^{-1}\left\{2H_{q_2+2}(\chi^2_{q_2,\alpha} : \Delta) - H_{q_2+4}(\chi^2_{q_2,\alpha} : \Delta)\right\}. \end{aligned} \tag{4.14}$$

Proof. (4.2) implies (4.12) and (4.11). Moreover, (4.13) and (4.14) follow from the results of [Chap. 7 in Saleh [10]]. □

We now evaluate the errors for a special case. Set

$$f_\alpha(\Delta) = 1 - H_3(\chi^2_{1,\alpha} : \Delta) + \left\{2H_3(\chi^2_{1,\alpha} : \Delta) - H_5(\chi^2_{1,\alpha} : \Delta)\right\}\Delta, \tag{4.15}$$

then (4.12) and (4.14) are rewritten as

$$M(\hat{\theta}_{1n}) = \Gamma_{11}^{-1} + \Gamma_{11}^{-1}\Gamma_{12}\Gamma_{22\cdot 1}^{-1}\Gamma_{21}\Gamma_{11}^{-1}\Delta \tag{4.16}$$

and

$$M(\hat{\theta}_{1n}^{PT}) = \Gamma_{11}^{-1} + \Gamma_{11}^{-1}\Gamma_{12}\Gamma_{22\cdot 1}^{-1}\Gamma_{21}\Gamma_{11}^{-1}f_\alpha(\Delta). \tag{4.17}$$

For the comparison between their behaviors, it is important to describe the magnitude of $M(\tilde{\theta}_{1n})$, $M(\hat{\theta}_{1n})$ and $M(\hat{\theta}_{1n}^{PT})$ in terms of that of 1, Δ and $f_\alpha(\Delta)$. Here note that $f_\alpha(\Delta) \to 1$ as $\Delta \to \infty$.

Proposition 4.2. *We obtain*

$$\begin{aligned} &M(\hat{\theta}_{1n}) \le M(\hat{\theta}_{1n}^{PT}) \le M(\tilde{\theta}_{1n}) && \text{if } 0 \le \Delta \le \tilde{\Delta}, && \text{(i)} \\ &M(\hat{\theta}_{1n}) \le M(\tilde{\theta}_{1n}) \le M(\hat{\theta}_{1n}^{PT}) && \text{if } \tilde{\Delta} \le \Delta \le 1, && \text{(ii)} \\ &M(\tilde{\theta}_{1n}) \le M(\hat{\theta}_{1n}) \le M(\hat{\theta}_{1n}^{PT}) && \text{if } 1 \le \Delta \le \hat{\Delta}, && \text{(iii)} \\ &M(\tilde{\theta}_{1n}) \le M(\hat{\theta}_{1n}^{PT}) \le M(\hat{\theta}_{1n}) && \text{if } \hat{\Delta} \le \Delta && \text{(iv)} \end{aligned}$$

where $\Delta = h^2\Gamma_{22\cdot 1}$, *and* $\tilde{\Delta}$ *and* $\hat{\Delta}$ *are defined by values of functions of* α *(or* $\chi^2_{1,\alpha}$*) satisfying* $1 = f_\alpha(\tilde{\Delta})$ *and* $\hat{\Delta} = f_\alpha(\hat{\Delta})$*, respectively.*

Let us examine numerical features of the results. For $\alpha = 0.1,\ 0.05,\ 0.01$ the corresponding $(\tilde{\Delta}, \hat{\Delta})$ are (0.6409735, 1.5044336), (0.6991444, 1.7876485) and (0.8228669, 2.5083560), respectively. Figure 1 plots f_α for α=0.1, 0.05 and 0.01.

Let $\{X(t)\}$ be an ARMA(1,1) process, i.e,

$$X(t) + bX(t-1) = U(t) + aU(t-1) \tag{4.18}$$

where $a, b \in (-1, 1)$, $a \neq b$ and $U(t)$ are i.i.d. random variables. Further, we assume that the distribution of $U(t)$ is a generalized error distribution (GED) with innovation density

$$p(u) = \frac{\eta \exp\left(-\frac{1}{2}|u/\lambda|^{\eta}\right)}{\lambda 2^{(1+1/\eta)}\Gamma(1/\eta)}, \quad 0 < \eta \leq \infty \tag{4.19}$$

where

$$\lambda = \frac{1}{2^{1/\eta}}\sqrt{\frac{\Gamma(1/\eta)}{\Gamma(3/\eta)}}. \tag{4.20}$$

If $\eta = 2$, $\{X(t)\}$ is a Gaussian process. Table 1 shows the values of $\mathcal{F}(p)$ and kurtosis for η =0.75 (0.25) 2.5. From this, we can see that $U(t)$ have heavy tails when $\eta < 2$, and the minimum of Fisher information $\mathcal{F}(p)$ is achieved at $\eta = 2$. Figure 2 plots f_α for $\alpha = 0.05$ and η=1, 1.5 and 2 under $\Gamma_{22\cdot 1}/\mathcal{F}(p)$=1. From Fig.2, it clear that $\hat{\theta}_{1n}^{PT}$ is comparatively effective when the distribution has heavy tail. Moreover from MSE $\propto \mathcal{F}(p)^{-1}$ MSE with $\eta \neq 2$ as a whole, especially $\eta < 2$ is smaller than that with $\eta = 2$.

TABLE 1.
(Skewness and Fisher Information of $U(t)$)

η	0.75	1	1.25	1.5	1.75	2	2.25	2.5
kurtosis	9.6500	6	4.5272	3.7620	3.3026	3	2.7875	2.6312
$\mathcal{F}(p)$	5.7312	2	1.3938	1.0957	1.0181	1	1.0122	1.0421

Let $\theta = (\theta_1, \theta_2) = (a, b - b')$ for given $b' \in (-1, 1)$ (Since a and b' are symmetric in Γ, we consider the case of $(\theta_1, \theta_2) = (b, a - a')$ for given $a' \in (-1, 1)$ similarly). Since the spectral density of $\{X(t)\}$ is

$$f_\theta(\lambda) = \frac{1}{2\pi}\frac{|1 - ae^{i\lambda}|^2}{|1 - be^{i\lambda}|^2}, \tag{4.21}$$

it follows that

$$\Gamma = \mathcal{F}(p)\begin{pmatrix} \frac{1}{1-a^2} & -\frac{1}{1-ab'} \\ -\frac{1}{1-ab'} & \frac{1}{1-b'^2} \end{pmatrix} \tag{4.22}$$

(See Taniguchi [13] or Taniguchi [14]) and

$$\Gamma_{22\cdot 1} = \mathcal{F}(p)\frac{(a - b')^2}{(1 - b'^2)(1 - ab')^2} \tag{4.23}$$

Since we assume the stationarity, it is supposed that a is not close to b'. In Figures 3 and 4 we provide the MSEs for $\alpha = 0.05$, $\eta = 2$, $h = 2$ and b'=0.2. The figures confirm (i) and (iv) in proposition 4.2 numerically. From Fig.3 we observe that PTMLE is better than Unrestricted MLE if $a \approx b'$. While from Fig.4 we mention that PTMLE is better than Restricted MLE otherwise.

It is clear from the numerical studies that $\hat{\theta}_{1n}^{PT}$ is comparatively effective as a whole in comparison with $\tilde{\theta}_{1n}$ and $\hat{\theta}_{1n}$. Thus, we conclude that PTMLE is useful to estimate θ_1 when $\theta_2 \approx 0$.

5 Optimal Portfolio and Market Efficiency

As an application of the results in the previous sections we consider the optimization problem of the portfolio choice of m assets:

$$\text{stock } \mathbf{S} = \{\mathbf{S}(t)\} = \{(S_1(t), \cdots, S_m(t))'\}.$$

Let the return process of $\{\mathbf{S}(t)\}$ be $\mathbf{X}(t)$ defined by (2.1), i.e.,

$$S_j(t) - S_j(t-1) = S_j(t-1)X_j(t), \quad t \in \mathbf{Z}, \ j = 1, \cdots, m.$$

For the optimization problem of their portfolio choice(e.g. Kariya [5] and Shiraishi and Taniguchi [12]), it is important that we estimate the spectral density matrix $\mathbf{f}_\theta(\lambda)$ of $\{\mathbf{X}(t)\}$ accurately.

For example, we consider the mean variance optimal hedging of $\{S_1(t)\}$ by $\{S_2(t), \cdots, S_m(t)\}$, i.e., the problem of choosing the constant μ and $(m-1) \times 1$ vectors $\{\xi(t)\}$ which minimize

$$\mathsf{E}\left[\left\{X_1(t) - \mu - \sum_{j=-\infty}^{\infty} \xi(j)(X_2(t-j), \cdots, X_m(t-j))'\right\}^2\right]. \tag{5.1}$$

Then, using the decomposition of spectral density matrix

$$\mathbf{f}_\theta(\lambda) = \begin{pmatrix} f_{11,\theta}(\lambda) & \mathbf{f}_{1m,\theta}(\lambda) \\ \mathbf{f}_{m1,\theta}(\lambda) & \mathbf{f}_{mm,\theta}(\lambda) \end{pmatrix}, \tag{5.2}$$

the μ and $\xi(j)$ that minimize (5.1) are given by $\mu = 0$ and

$$\xi(j) = \frac{1}{2\pi} \int_0^{2\pi} e^{ij\lambda} \mathbf{f}_{1m,\theta}(\lambda) \mathbf{f}_{mm,\theta}(\lambda)^{-1} d\lambda, \tag{5.3}$$

and the minimum of (5.1) is

$$\int_0^{2\pi} \left\{ f_{11,\theta}(\lambda) - \mathbf{f}_{1m,\theta}(\lambda)\mathbf{f}_{mm,\theta}(\lambda)^{-1}\mathbf{f}_{m1,\theta}(\lambda) \right\} d\lambda \tag{5.4}$$

(See Chapter 8 in Brillinger [1]).

Evidently, if we need a good estimator for $\xi(j)$, we need the one for θ Since market efficiency and stock returns are generally supposed to be uncorrelated. Thus, if we choose a partition of θ for preliminary test such that $A_\theta(0)$ is independent of θ_2, and $A_\theta(j) = 0$ (or $B_\theta(j) = 0$) for $j = 1, 2, \cdots$ when $\theta_2 = \mathbf{0}$, then we can apply the results in Section 4 to this problem. Here, the test statistic $\mathcal{L}_n$ means a test of (weak-form) efficient market hypothesis (EMH).

References

[1] Brillinger, D.R. (1981). *Time Series: Data Analysis and Theory.* expanded-eded. San Francisco: Holden-Day.

[2] Dzhaparidze, K. (1986) *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*. New York: Springer-Verlag.

[3] Fuller, Wayne A. (1976). *Introduction to Statistical Time Series*. Wiley-Interscience.

[4] Hallin, M., Taniguchi, M., Serroukh, A. & Choy, K. (1999). Local Asymptotic Normality for Regression Models with Long-Memory Disturbance. *Ann. of Statistics.* **27**, 6, 2054-2080.

[5] Kariya, T. (1993). *Quantitative Methods for Portfolio Analysis.* Kluwer Academic Publishers, Dordrecht.

[6] Mukherjee, K. (1998). On preliminary test and shrinkage estimation in linear models with long-memory errors *Journal of statistical planning and inference.*

[7] Saleh, A.K.Md. Ehsanes & Sen, P.K. (1978). Non parametric estimation parameter after a preliminary test on regression. *Ann. Statist.* **6**, 154-168.

[8] Saleh, A.K.Md. Ehsanes & Sen, P. K. (1987). Relative performance of Stein-rule and preliminary test estimators in linear models. *Commun. Statist.* **13**, 2725-2747.

[9] Saleh, A. K. Md. Ehsanes (1992). On Shrinkage Estimation of the Parameters of an Autoregressive Gaussian Process. *Theory Prob. Appl.* **37**, 290-300.

[10] Saleh, A. K. Md. Ehsanes (2006). *Theory of Preliminary Test and Stein-Type Estimation with Applications.* Wiley-Interscience.

[11] Sen, P.K. & Saleh, A.K.Md. Ehsanes (1987a). On preliminary test and shrinkage M-estimation in linear models. *Ann. Statist.* **15**, 1580-1592.

[12] Shiraishi, H. & Taniguchi, M. (2008). Statistical Estimation of Optimal Portfolios for non-Gaussian Dependent Returns of Assets. To appear in *J. Forecasting.*

[13] M. Taniguchi (1985). An Asymptotic Expansion for the Distribution of the Likelihood Ratio Criterion for a Gaussian Autoregressive moving Average Process under a Local Alternative. *Econometric Theory* **1**, 73-84.

[14] Taniguchi, M (1991). *Higher Order Asymptotic Theory for Time Series Analysis.* Springer-Verlag.

[15] Taniguchi, M. & Amano, T. (2007). Systematic Approach for Portmanteau Tests in View of Whittle Likelihood Ratio.

[16] Taniguchi,M. &Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series.* New York: Springer-Verlag.

Figure 1.
f_α for α=0.1, 0.05 and 0.01.

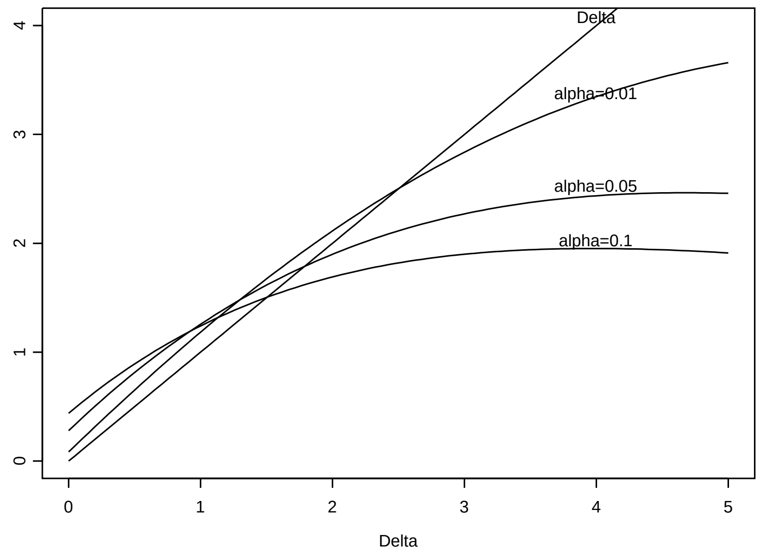

Figure 2.
f_α for α=0.05 and η=1, 1.5 and 2.

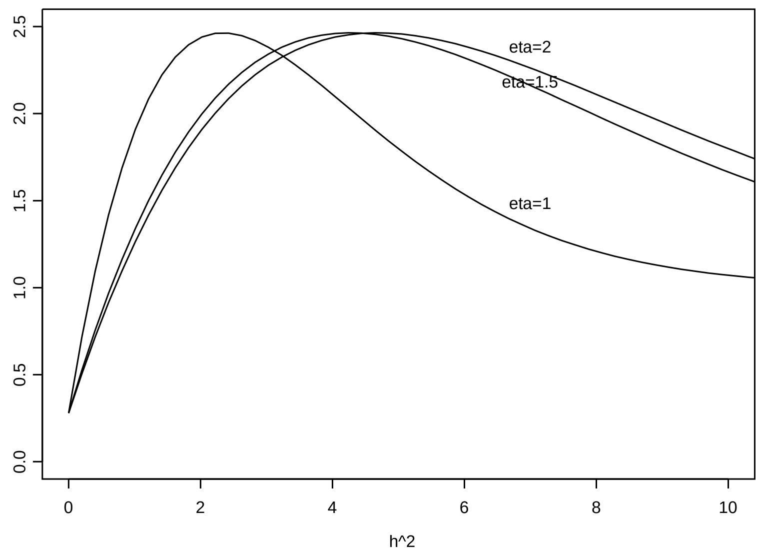

FIGURE 3.
MSEs for α=0.05, η=2, $h = 2$ and b'=0.2

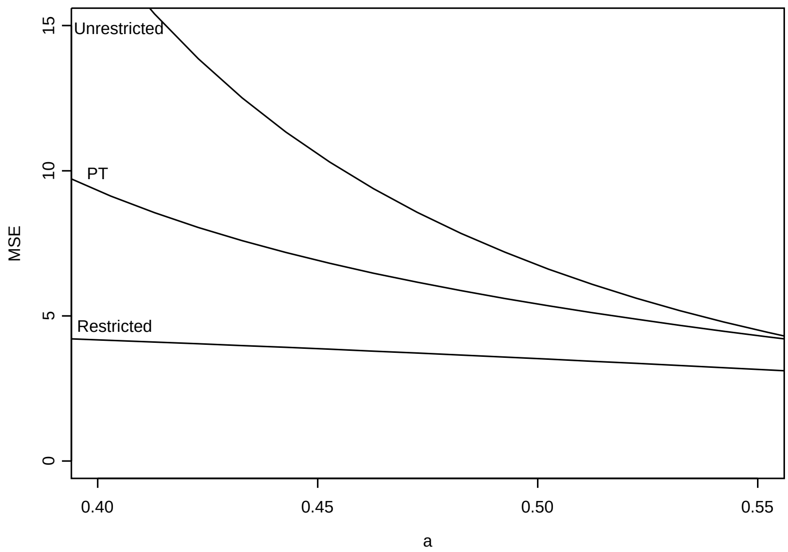

FIGURE 4.
MSEs for α=0.05, η=2, $h = 2$ and b'=0.2