学 位 請 求 論 文

谷 口 正 信

主　　論　　文

Asymptotic Theory of Estimation of the Spectral Density for a Stationary Process

Masanobu Taniguchi
Hiroshima University

Table of contents

1. Introduction

This dissertation is a compilation of the author's works on the problem of estimating the spectral density of a stationary process which are contained in the papers [23] through [30]. In order to provide a perspective on the whole theory with a wide coverage of topics in a limited space, all lemmas and theorems are stated with only explanation of their implications but without proofs. Readers who are interested in technical details are referred to the original papers.

In the usual estimation theory for independently and identically distributed random variables, asymptotic effects of small perturbations in the underlying model on an estimate have been investigated in detail (e.g., Beran[3]). But, in time series analysis, discussion of this type seem to have been barren, thus we shall develop some discussions of this type for a stationary process.

Let Z be the set of all integers. Suppose that $\{ X(t); t \in Z \}$ is a stationary Gaussian process with zero mean and a spectral density $g(x)$. Then $g(x)$ contains all the information about the process. Assuming that the true spectral density $g(x)$ can be expressed as a parametric form $f_\theta(x)$, many papers have dealt with the problem of estimating the unknown parameter θ. However ' the true structure ' $g(x)$ would be rarely prescribed exactly by any parametric model. Thus a discussion will be needed under the situation that a parametric family of spectral densities does not necessarily contain the true one. The aim of this dissertation is to develop a discussion on fitting the parametric spectral density $f_\theta(x)$ to the true spectral density $g(x)$.

In Section 2, we propose two estimates of θ by fitting $f_\theta(x)$ to $g(x)$, say $\hat{\theta}_1$ and $\hat{\theta}_2$, which minimize two criteria $D_1(f_\theta,g)$ and $D_2(f_\theta,g)$ respectively, which measure the nearness of $f_\theta(x)$ to $g(x)$. Then we investigate some asymptotic behavior of the estimates with respect to efficiency and robustness. We see there a necessity to estimate two types of the integrals $\int_{-\pi}^{\pi} \eta(x)\, g(x)\, dx$ and $\int_{-\pi}^{\pi} \phi(x) \log g(x)\, dx$, where $\eta(x)$ and $\phi(x)$ are continuous functions. If we consider fitting by another criterion, we are required to estimate the integral $\int_{-\pi}^{\pi} \psi(x)\, \Phi(g(x))\, dx$ for some function $\Phi(.)$ and for each continuous function $\psi(x)$. Thus, in Section 3, we shall propose

$$H_n = \int_{-\pi}^{\pi} \psi(x)\, L^{-1}\{\, \Phi(\tfrac{1}{\omega})\, \tfrac{1}{\omega}\,\}_{<I_n(x)>}\, dx,$$

as such an estimate under some assumptions on $\Phi(x)$, where $L^{-1}\{\, F(\omega)\,\}_{<u>}$ denotes the Laplace inverse transform of $F(\omega)$ at the argument u, and $I_n(x)$ is the periodogram of the process.

In Section 4, we propose an estimate of θ to fit $f_\theta(x)$ to $g(x)$, say $\hat{\theta}_n$, by minimizing the criterion

$$D(f_\theta,g) = \int_{-\pi}^{\pi} [\, \Phi(f_\theta(x))^2 - 2\, \Phi(f_\theta(x))\, \Phi(g(x))\,]\, dx.$$

As a parametric family $\{ f_\theta(x) \}$, we shall propose the following type of spectral model

$$f_\theta(x) = \Phi^{-1}\{ \sum_{j=0}^{p} \theta_j\, e_j(x) \},$$

where $\theta = (\theta_1,\ldots,\theta_p)'$, while $\{ e_j(x)\ ;\ j = 1,\ldots,p \}$ is an orthonormal basis of continuous functions in $L^2[-\pi,\pi]$. In general the above model is " non-structural " in the sense that the parameter vector θ is not necessarily explanatory. But it has an advantage that the estimate $\hat{\theta}_n$ can be easily expressed in an explicit form. Further we shall study the asymptotic efficiency of the estimate.

The estimate $\hat{\theta}_1$ introduced in Section 2 is to be called a quasi-Gaussian maximum likelihood estimate. In Section 5, we shall give its asymptotic distribution under the condition that the underlying process is not necessarily Gaussian, and discuss a non-Gaussian robustness.

Up to now we have proceeded under the assumption that the order of the fitted spectral model is known. In Section 6 we assume that the true spectral density $g(x)$ is described by an infinite set of parameters and that the fitted kth order model $f_{\tau(k)}(x)$ tends to $g(x)$ as $k \to \infty$. Then we show that the choice of the order of the model by Akaike's information criterion which is constructed by the Gaussian likelihood is optimal in a certain sense.

In Section 7, we shall point out a kind of similarity between the estimation problem in a time series regression model and the interpolation problem in a stationary process. From a unified view we shall again look at the estimation problem in the time series regression model and propose a parametric method which gives an efficient estimate of the regression coefficient vector involved in the model. Also we shall consider the interpolation problem and the regression problem under the condition that the spectral density of the underlying stationary process is vaguely known (i.e., Huber's ε-contaminated model). Then we can get a minimax robust interpolator and a minimax robust estimate of the regression coefficient.

2. Two estimates of parameters of a Gaussian stationary process

Let F denote the set of all spectral densities with respect to the Lebesgue measure on the real line. For a parametric family of spectral densities $\{ f_\theta ; \theta \in \Theta \}$, $\Theta \subset R^p$, and for $g \in F$, we shall propose the following criterion :

$$D_1(f_\theta, g) = \int_{-\pi}^{\pi} \{ \log f_\theta(x) + g(x)/f_\theta(x) \} dx .$$

Since we can show that for each $x \in [-\pi, \pi]$,

$$\log f_\theta(x) + g(x)/f_\theta(x) \geq \log g(x) + 1 ,$$

and the equality holds if and only if $f_\theta(x) = g(x)$, the criterion can serve as a measure of the nearness of $f_\theta(x)$ to $g(x)$. The functional T_1 on F is defined by the requirement that

(2.1) $$D_1(f_{T_1(g)}, g) = \min_{\theta \in \Theta} D_1(f_\theta, g) \qquad \text{for every } g \in F.$$

Now we shall define a convergence in the set F. If, for a sequence $\{ g_n \}$, $g_n \in F$,

$$\int_{-\pi}^{\pi} \psi(x)\, g_n(x)\, dx \to \int_{-\pi}^{\pi} \psi(x)\, g(x)\, dx \quad \text{as } n \to \infty$$

for every continuous function $\psi(x)$, then we say that $\{ g_n \}$ converges weakly to g and denote by $g_n \overset{W}{\to} g$.

Lemma 2.1.(Taniguchi[24]). Suppose that Θ is a compact subset of R^p, $\theta_1 \neq \theta_2$ implies $f_{\theta_1} \neq f_{\theta_2}$ on a set of positive Lebesgue measure and that $f_\theta(x) > 0$ and also is continuous in θ and x. Then

(i) For every $g \in F$, there exists a value $T_1(g) \in \Theta$ satisfying (2.1).

(ii) Assume $T_1(g)$ is unique for every g ; then if $g_n \overset{W}{\to} g$, $T_1(g_n) \to T_1(g)$ as $n \to \infty$.

(iii) $T_1(f_\theta) = \theta$ uniquely for every $\theta \in \Theta$. ▯

Now we impose further assumptions on f_θ that two p×p matrices of the second partial derivatives $\partial^2 \log f_\theta(x)/\partial\theta\partial\theta'$ and $\partial^2 f_\theta(x)^{-1}/\partial\theta\partial\theta'$ (θ' denoting the transpose of θ) exist and are continuous in θ and x. Then we have the following lemma.

Lemma 2.2.(Taniguchi[24]). Suppose that $T_1(g)$ exists uniquely and lies in $Int(\Theta)$ and that the matrix

$$Q_1 = \int_{-\pi}^{\pi} \left\{ \frac{\partial^2 \log f_\theta(x)}{\partial\theta\,\partial\theta'} + \frac{\partial^2 f_\theta(x)^{-1}}{\partial\theta\,\partial\theta'}\, g(x) \right\}_{\theta=T_1(g)} dx$$

is non-singular. Then for every sequence of spectral densities $\{ g_n \}$ satisfying $g_n \overset{W}{\to} g$, we have

$$T_1(g_n) = T_1(g) + \int_{-\pi}^{\pi} \rho_g(x)\, (\, g_n(x) - g(x)\,)\, dx + a_n \int_{-\pi}^{\pi} \left. \frac{\partial f_\theta(x)^{-1}}{\partial\theta} \right|_{\theta=T_1(g)} (\, g_n(x) - g(x)\,)\, dx,$$

where $\rho_g(x)$ is a p×p matrix defined by

$$\rho_g(x) = -\, Q_1^{-1} \cdot \frac{\partial}{\partial\theta} f_\theta(x)^{-1} \Big|_{\theta=T_1(g)} ,$$

and $\{ a_n \}$ is also a sequence of real p×p matrices which tends to zero as $n \to \infty$. ‖

Remark 2.1. If $g = f_\theta$, then $T_1(g) = \theta$. Thus we have

$$\rho_{f_\theta}(x) = -\left[\int_{-\pi}^{\pi} \frac{\partial \log f_\theta(x)}{\partial\theta} \frac{\partial \log f_\theta(x)}{\partial\theta'}\, dx \right]^{-1} \frac{\partial f_\theta(x)^{-1}}{\partial\theta} . \quad ‖$$

Let $\{ X(t) \; ; \; t \in Z \}$ be a stationary Gaussian process with zero mean and a spectral density g(x). Suppose that a stretch, X(0),...,X(n-1) of the series X(t) is obserbed. Put $I_n(x) = (2\pi n)^{-1} | \sum_{t=0}^{n-1} X(t) \exp(-itx) |^2$,

which is the periodogram of X(t). Define

$$G = \int_{-\pi}^{\pi} \psi(x)\ g(x)\ dx\ , \qquad G_n = \int_{-\pi}^{\pi} \psi(x)\ I_n(x)\ dx\ ,$$

where $\psi(x)$ is a p-vector valued continuous function on $[-\pi,\pi]$ such that $\psi(x) = \psi(-x)$. We shall impose the following.

Assumption 2.1.

$$\sum_{t=-\infty}^{\infty} |t||\ R(t)\ | < \infty\ , \qquad \text{where } R(t) = E[\ X(t+s)X(s)\ ]. \qquad \square$$

Then we have the following lemma.

Lemma 2.3.(Brillinger[5],[6] and Walker[31]). If Assumption 2.1 is satisfied, then G_n is a consistent estimate for G, and the distribution of the vector $\sqrt{n}(\ G_n - G\ )$ tends to the normal distribution

$$N(\ \mathbf{0}_p,\ \ 4\pi \int_{-\pi}^{\pi} \psi(x)\ \psi(x)'\ g(x)^2\ dx\ )$$

as $n \to \infty$, where $\psi(x)'$ stands for the transpose of $\psi(x)$. $\square$

In view of this lemma we can recommend $T_1(I_n)$ as an estimate of $T_1(g)$. Under the assumption that $g = f_\theta$, $D_1(f_\theta,I_n)$ is known to be an ápproximation (neglecting constant term and multiple) for the logarithm of the likelihood of the Gaussian data X(0),...,X(n-1) (c.f., Walker[31], Clevenson[7], Dzhaparidze[11], Hosoya[16], Dunsmuir and Hannan[10] and Dunsmuir[9]). Hereafter we call $T_1(I_n)$ a quasi-Gaussian maximum likelihood estimate under the model f_θ. Then we have the following theorem.

Theorem 2.1.(Taniguchi[24]). Under Assumption 2.1 in addition to the conditions in Lemma 2.2 and as $n \to \infty$, the distribution of the vector $\sqrt{n}(T_1(I_n) - T_1(g))$ tends to the normal distribution

$$N(0_p, 4\pi \int_{-\pi}^{\pi} \rho_g(x) \rho_g(x)' g(x)^2 dx). \qquad \square$$

Remark 2.2. If $g = f_\theta$, the above asymptotic covariance matrix is equal to

$$4\pi[\int_{-\pi}^{\pi} \frac{\partial \log f_\theta(x)}{\partial\theta} \frac{\partial \log f_\theta(x)}{\partial\theta'} dx]^{-1}, \tag{2.2}$$

which is inverse to the limit of Fisher's information matrix(e.g., Dzhaparidze[11] and Clevenson[7]). If the asymptotic variance matrix of an asymptotically unbiassed estimate of θ attains the matrix (2.2) as $T_1(I_n)$ does, we say that the estimate is asymptotically efficient. $\square$

Now we consider another type of approach which proceeds quite similarly. For the parametric family of spectral densities $\{ f_\theta; \theta \in \Theta \}$, $\Theta \subset R^p$, and for $g \in F$, we shall propose the following criterion:

$$D_2(f_\theta,g) = \int_{-\pi}^{\pi} \{ \log^2 f_\theta(x) - 2 \log f_\theta(x) \log g(x) \}dx ,$$

which also measures the nearness of $f_\theta(x)$ to $g(x)$. The functional T_2 on F is defined by the requirement that

$$D_2(f_{T_2(g)},g) = \min_{\theta\in\Theta} D_2(f_\theta,g) \qquad \text{for every } g \in F. \tag{2.3}$$

Here we define a convergence in F. If, for a sequence $\{ g_n \}$, $g_n \in F$,

$$\int_{-\pi}^{\pi} \phi(x) \log g_n(x) dx \to \int_{-\pi}^{\pi} \phi(x) \log g(x) dx \quad \text{as } n \to \infty$$

for every continuous function $\phi(x)$, then we say that $\{ g_n \}$ converges ' log-weakly' to g and denote by $g_n \xrightarrow{L.W} g$.

Lemma 2.4.(Taniguchi[24]). Suppose that Θ is a compact subset of R^p, $\theta_1 \neq \theta_2$ implies $f_{\theta_1} \neq f_{\theta_2}$ on a set of positive Lebesgue measure and that $f_\theta(x) > 0$ and is continuous in θ and x. Then

(i) For every $g \in F$, there exists $T_2(g) \in \Theta$ satisfying (2.3).

(ii) Assume $T_2(g)$ is unique ; then if $g_n \xrightarrow{L.W} g$, $T_2(g_n) \to T_2(g)$ as $n \to \infty$.

(iii) $T_2(f_\theta) = \theta$ uniquely for every $\theta \in \Theta$. ▯

Now we impose further assumptions on $f_\theta(x)$ that $\log f_\theta(x)$ has the derivatives $\partial \log f_\theta(x)/\partial\theta$ and $\partial^2 \log f_\theta(x)/\partial\theta\partial\theta'$ which are continuous in θ and x.

Lemma 2.5.(Taniguchi[24]). Suppose that $T_2(g)$ exists uniquely and lies in $Int(\Theta)$ and that

$$Q_2 = \int_{-\pi}^{\pi} \left\{ \frac{\partial^2 \log f_\theta(x)}{\partial\theta\, \partial\theta'} \log g(x) - \frac{1}{2} \frac{\partial^2 \log^2 f_\theta(x)}{\partial\theta\, \partial\theta'} \right\}_{\theta=T_2(g)} dx$$

is non-singular. Then for every sequence of densities $\{ g_n \}$ satisfying $g_n \xrightarrow{L.W} g$, we have

$$T_2(g_n) = T_2(g) + \int_{-\pi}^{\pi} \sigma_g(x) \left(\log g_n(x) - \log g(x) \right) dx$$

$$+ b_n \int_{-\pi}^{\pi} \frac{\partial \log f_\theta(x)}{\partial\theta} \bigg|_{\theta=T_2(g)} \left(\log g_n(x) - \log g(x) \right) dx,$$

where $\sigma_g(x) = - Q_2^{-1} \dfrac{\partial \log f_\theta(x)}{\partial\theta} \bigg|_{\theta=T_2(g)}$,

and $\{ b_n \}$ is a sequence of p×p matrices which tends to zero as $n \to \infty$. ▯

Remark 2.3. If $g = f_\theta$, then $T_2(f) = \theta$. Thus we have

$$\sigma_{f_\theta}(x) = \left[\int_{-\pi}^{\pi} \frac{\partial \log f_\theta(x)}{\partial \theta} \frac{\partial \log f_\theta(x)}{\partial \theta'} dx \right]^{-1} \frac{\partial \log f_\theta(x)}{\partial \theta} . \qquad \Box$$

Let $\phi(x)$ be a p-dimensional vector consisting of continuous functions on $[-\pi,\pi]$ such that $\phi(x) = \phi(-x)$. Define

$$H = \int_{-\pi}^{\pi} \phi(x) \log g(x)\, dx, \qquad H_n = \int_{-\pi}^{\pi} \phi(x) \log \alpha\, I_n(x)\, dx,$$

where $\alpha = \exp \gamma$, ($\gamma \simeq 0.57721\ldots$, Euler's constant). Comparing Taniguchi [23] and Davis and Jones[8], we have the following lemma.

Lemma 2.6. (Taniguchi[24]). If Assumption 2.1 is satisfied, then H_n is a consistent estimate of H, and the distribution of the vector $\sqrt{n}(\ H_n - H\ )$ tends to the normal distribution

$$N(\ \mathbf{0}_p,\ \frac{2\pi^3}{3} \int_{-\pi}^{\pi} \phi(x)\, \phi(x)'\, dx\ )$$

as $n \to \infty$. $\Box$

From the above three lemmas we obtain the following theorem.

Theorem 2.2. (Taniguchi[24]). Under Assumption 2.1 together with the conditions in Lemma 2.5, the distribution of the vector $\sqrt{n}(\ T_2(\alpha I_n) - T_2(g)\ )$ tends to the normal distribution

$$N(\ \mathbf{0}_p,\ 4\pi \int_{-\pi}^{\pi} \sigma_g(x)\, \sigma_g(x)'\, dx\ )$$

as $n \to \infty$. $\Box$

Remark 2.4. If $g = f_\theta$, the above asymptotic covariance matrix is equal to

$$\frac{2\pi^3}{3} \left[\int_{-\pi}^{\pi} \frac{\partial \log f_\theta(x)}{\partial \theta} \frac{\partial \log f_\theta(x)}{\partial \theta'} dx \right]^{-1},$$

which, combined with Remark 2.2, implies that the estimate $T_1(I_n)$ is better than $T_2(\alpha I_n)$ so long as g belongs to the parametric family $\{ f_\theta ; \theta \in \Theta \}$. ▯

On the contrary, if the true spectral density is not contained in the parametric family $\{ f_\theta ; \theta \in \Theta \}$, we shall show that $T_2(.)$ is optimally insensitive to perturbations of its argument in some sense. To make this assertion precise, consider the set of all functionals U defined on F that have the following two properties for every $\theta \in Int(\Theta)$;

(i) $U(f) = \theta$,

(ii) $$U(g) - U(f_\theta) = \int_{-\pi}^{\pi} \eta(x) (\log g(x) - \log f_\theta(x)) dx + a \int_{-\pi}^{\pi} \psi(x) (\log g(x) - \log f_\theta(x)) dx,$$

where $a \to 0$ as $g \xrightarrow{L.W} f_\theta$ and both $\eta(x)$ and $\psi(x)$ are p-vector-valued continuous functions on $[-\pi,\pi]$. From Lemma 2.5, the functional T_2 has these properties The requirement (i) imposes a further constraint on $\eta(x)$; i.e., for every $\theta \in Int(\Theta)$,

$$\int_{-\pi}^{\pi} \eta(x) \frac{\partial \log f_\theta(x)}{\partial \theta'} dx = I_p, \quad \text{the } p\times p \text{ identity matrix.}$$

Put $\Delta(f_\theta,g) = \| \log f_\theta - \log g \|$, where $\|.\|$ denotes the L^2-norm. The requirement (ii) then can be rewritten as

$$U(g) - \theta = \Delta(f_\theta,g) \int_{-\pi}^{\pi} \eta(x) \delta(x) dx + o(\Delta(f_\theta,g)),$$

where $\delta(x) = \{ \log g(x) - \log f_\theta(x) \}/\Delta(f_\theta,g)$ with $\|\delta\| = 1$. For small and fixed value of $\Delta(f_\theta,g)$, the behavior of $U(g) - \theta$ is determined primarily by the term $\int_{-\pi}^{\pi} \eta(x) \delta(x) dx$. Since $\eta(x)$ is a vector function, we shall investigate

$$L(\eta,\delta) = \left| \int_{-\pi}^{\pi} c'\eta(x) \delta(x) dx \right|$$

for a constant vector c in R^p.

Theorem 2.3.(Taniguchi[24]). Suppose that $\delta \varepsilon L^2$, $\|\delta\| = 1$, $\eta \varepsilon L^2$ and that

$$\int_{-\pi}^{\pi} \eta(x) \frac{\partial \log f_\theta(x)}{\partial \theta'} dx = I_p .$$

Then for every $c \varepsilon R^p$,

$$\max_{\delta} \min_{\eta} L(\eta,\delta) = \min_{\eta} \max_{\delta} L(\eta,\delta) = L(\eta_0,\delta_0), \tag{2.4}$$

where

$$\eta_0 = \left[\int_{-\pi}^{\pi} \frac{\partial \log f_\theta(x)}{\partial \theta} \frac{\partial \log f_\theta(x)}{\partial \theta'} dx \right]^{-1} \frac{\partial \log f_\theta(x)}{\partial \theta} ,$$

$$\delta_0 = \| c' \eta_0 \|^{-1} c' \eta_0 .$$ ▯

Remark 2.5. The above theorem, Lemma 2.5 and Remark 2.3 mean that the functional T_2 ia locally minimax robust at f_θ in the sense of (2.4). That is, for the least favourable δ_0, the functional which minimizes $L(\eta,\delta_0)$ is T_2. But T_1 does not have this type of robustness. ▯

3. Estimation of the integrals of certain functions of spectral density

In view of Lemmas 2.2 and 2.5, we can see that $T_1(g)$ and $T_2(g)$ are continuous functions of $\int_{-\pi}^{\pi} \eta(x)\, g(x)\, dx$ and $\int_{-\pi}^{\pi} \phi(x) \log g(x)\, dx$, respectively, for some continuous functions $\eta(.)$ and $\phi(.)$ on $[-\pi,\pi]$. Thus it is suggested that estimates of $T_1(g)$ and $T_2(g)$ can be obtained by finding estimates of $\int_{-\pi}^{\pi} \eta(x)\, g(x)\, dx$ and $\int_{-\pi}^{\pi} \phi(x) \log g(x)\, dx$, respectively. In the previous section we have proposed $\int_{-\pi}^{\pi} \eta(x)\, I_n(x)\, dx$ and $\int_{-\pi}^{\pi} \phi(x) \log \alpha I_n(x)\, dx$ as such estimates. In Section 4, we shall consider a criterion D as a generalization of the criterion D_2, which requires us to estimate the integral $\int_{-\pi}^{\pi} \psi(x)\, \Phi(g(x))\, dx$ for some function $\Phi(.)$ and for each continuous function $\psi(x)$. Thus we consider to estimate it in this section.

Let $\{ X(t);\ t \in Z \}$ be a stationary Gaussian process with zero mean and a continuous spectral density $g(x)$ such that $0 < g(x) < \infty$ for any x. We shall denote the Laplace transform and the Laplace inverse transform of any function $F(\omega)$ at the argument u by $L\{ F(\omega) \}_{\langle u\rangle}$ and $L^{-1}\{ F(\omega) \}_{\langle u\rangle}$ respectively, i.e.,

$$L\{ F(\omega) \}_{\langle u\rangle} = \int_0^{\infty} e^{-u\omega} F(\omega)\, d\omega,$$

$$L^{-1}\{ F(\omega) \}_{\langle u\rangle} = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{u\omega} F(\omega)\, d\omega,$$

where σ is a real number greater than the abscissa of absolute convergence.

Now we shall impose the following assumptions on $\Phi(\omega)$.

Assumption 3.1. Let $\Phi(\omega)$ be a function on $(0,\infty)$ such that the Laplace inverse transform $L^{-1}\{ \Phi(1/\omega)\, 1/\omega \}_{<u>}$ exists and is continuous in $u \in (0,\infty)$. ▯

Assumption 3.2. For every positive integer m, the Laplace transform

$$L[\; | L^{-1}\{ \Phi(1/\omega)\, 1/\omega \}_{<u>} |^m \;]_{<x>}$$

exists and is continuous in $x \in (0,\infty)$. ▯

Suppose that a stretch, $X(0),\ldots,X(n-1)$ of the series $X(t)$ is observed. Put

$$I_n(x) = (2\pi n)^{-1} \left| \sum_{t=0}^{n-1} X(t) \exp(-itx) \right|^2 .$$

For a function $\Phi(\omega)$ which satisfies Assumptions 3.1 and 3.2, we shall define

$$H = \int_{-\pi}^{\pi} \psi(x)\, \Phi(g(x))\, dx,$$

$$H_n = \int_{-\pi}^{\pi} \psi(x)\, L^{-1}\{ \Phi(1/\omega)\, 1/\omega \}_{<I_n(x)>}\, dx,$$

where $\psi(x)$ is a p-vector valued continuous function on $[-\pi,\pi]$ such that $\psi(x) = \psi(-x)$. Then we have the following theorem.

Theorem 3.1.(Taniguchi[25]). If Assumptions 2.1, 3.1 and 3.2 are satisfied, then H_n is a consistent estimate for H, and the distribution of the vector $\sqrt{n}(\, H_n - H\,)$ tends to the normal distribution

$$N(\, 0_p,\, 4\pi \int_{-\pi}^{\pi} \psi(x)\, \psi(x)'\, \sigma^2(x)\, dx \,)$$

as $n \to \infty$, where

$$\sigma^2(x) = \frac{1}{g(x)}\, L[[\, L^{-1}\{ \Phi(1/\omega)\, 1/\omega \}_{<u>} \,]^2]_{<g(x)^{-1}>} - \Phi(g(x))^2. \;▯$$

Now we shall consider some examples of the function $\Phi(\omega)$ which satisfy

Assumptions 3.1 and 3.2.

Example 3.1. Let us put $\Phi(\omega) = \omega^{\beta}$ ($0 < \beta < \infty$). Then we have

$$L^{-1}\{ \Phi(1/\omega)\ 1/\omega \}_{<u>} = \frac{u^{\beta}}{\Gamma(\beta+1)} ,$$

$$\sigma^2(x) = \{ \frac{\Gamma(2\beta+1)}{\Gamma(\beta+1)^2} - 1 \}\ g(x)^{2\beta}.$$

Thus as a consequence of Theorem 3.1, we obtain the following.

Corollary 3.1. The distribution of the vector

$$\sqrt{n}[\int_{-\pi}^{\pi} \psi(x) \frac{\{ I_n(x) \}^{\beta}}{\Gamma(\beta+1)} dx - \int_{-\pi}^{\pi} \psi(x)\ g(x)^{\beta} dx]$$

tends to the normal distribution

$$N(0_p, 4\pi \{ \frac{\Gamma(2\beta+1)}{\Gamma(\beta+1)^2} - 1 \} \int_{-\pi}^{\pi} \psi(x)\ \psi(x)'\ g(x)^{2\beta} dx)$$

as $n \to \infty$. ▯

Example 3.2. Let us put $\Phi(\omega) = \log \omega$. Then we have

$$L^{-1}\{ \Phi(1/\omega)\ 1/\omega \}_{<u>} = \log \alpha u ,$$

where $\alpha = \exp \gamma$ ($\gamma \simeq 0.57721$, Euler's constant), and

$$\sigma^2(x) = \pi^2/6.$$

Application of Theorem 3.1 to this example yields Lemma 2.6, which also agrees with the results of Hannan and Nicholls[15]. ▯

4. An estimation procedure of parameters of a certain spectral density model

In this section we shall use the same notations and assumptions as in Section 3. Let $\Phi(\omega)$ be a continuous bijective function on $(0,\infty)$ satisfying Assumptions 3.1 and 3.2 and denote by Φ^{-1} the inverse function of Φ. We shall fit a parametric spectral model $f_\theta(x)$ to the true spectral density $g(x)$ by the criterion

$$D(f_\theta, g) = \int_{-\pi}^{\pi} [\ \Phi(f_\theta(x))^2 - 2\ \Phi(f_\theta(x))\ \Phi(g(x))\ ]\ dx$$

and define the functional T on F, the set of all spectral densities, by the requirement that for $\{ f_\theta\ ;\ \theta \in \Theta \}$, $\Theta \subset R^p$,

$$D(f_{T(g)}, g) = \min_{t \in \Theta} D(f_t, g) \quad \text{for every } g \in F.$$

Now we shall define a convergence in F. If , for a sequence $\{ g_n \}$, $g_n \in F$,

$$\int_{-\pi}^{\pi} \phi(x)\ \Phi(g_n(x))\ dx \ \to \int_{-\pi}^{\pi} \phi(x)\ \Phi(g(x))\ dx \quad \text{as } n \to \infty$$

for every continuous function $\phi(x)$, then we denote by $g_n \xrightarrow{\Phi.W} g$. In the sequel in this section we often use the function

$$K(u) = L^{-1}\{\ \Phi(1/\omega)\ 1/\omega\ \}_{<u>}\ .$$

Replacing log(.) in Lemma 2.4 by $\Phi(.)$ we have analogous results for $T(.)$. Suppose that $\Phi(f_\theta(x))$ is twice differentiable with respect to θ, and that the derivatives are continuous in θ and x. Then we have the following lemma.

Lemma 4.1. (Taniguchi[27]). Suppose that T(g) exists uniquely and lies in $Int(\Theta)$ and that

$$G = \int_{-\pi}^{\pi} \{\ \frac{1}{2}\frac{\partial^2 \Phi^2(f_\theta(x))}{\partial\theta\ \partial\theta'} - \frac{\partial^2 \Phi(f_\theta(x))}{\partial\theta\ \partial\theta'}\ \Phi(g(x))\ \}_{\theta=T(g)}\ dx$$

is a non-singular matrix. Then *for* every sequence of spectral densities $\{ g_n \}$ satisfying that $g_n \xrightarrow{\Phi.W} g$, we have

$$T(g_n) - T(g) = \int_{-\pi}^{\pi} \eta_g(x) \{ \Phi(g_n(x)) - \Phi(g(x)) \} dx$$

$$+ a_n \int_{-\pi}^{\pi} \frac{\partial \Phi(f_\theta(x))}{\partial \theta} \Big|_{\theta=T(g)} \{ \Phi(g_n(x)) - \Phi(g(x)) \} dx,$$

where $\eta_g(x) = G^{-1}. \frac{\partial \Phi(f_\theta(x))}{\partial \theta} \Big|_{\theta=T(g)}$,

and $\{ a_n \}$ is a sequence of p×p-matrices which tends to zero as $n \to \infty$. ▯

Remark 4.1. Theorem 3.1 implies that as $n \to \infty$,

$$\Phi^{-1}[K\{ I_n(x)\}] \xrightarrow{\Phi.W} g(x), \text{ in probability.}$$ ▯

Noting Lemma 4.1, Remark 4.1 and Theorem 3.1, we shall recommend $T_n = T[\Phi^{-1}[K\{ I_n(x) \}]]$ as an estimate of $T(g)$.

Theorem 4.1.(Taniguchi[27]). If, in addition to the conditions of Lemma 4.1, Assumptions 2.1, 3.1 and 3.2 are assumed, then the distribution of the vector $\sqrt{n} (T_n - T(g))$ tends to the normal distribution

$$N(0_p, 4\pi \int_{-\pi}^{\pi} \eta_g(x) \eta_g(x)' \sigma^2(x) dx)$$

as $n \to \infty$, where $\sigma^2(x)$ is defined in Theorem 3.1. ▯

Remark 4.2. If $g = f_\theta$, the covariance matrix of the above normal distribution is equal to

$$4\pi[\int_{-\pi}^{\pi} \frac{\partial \Phi(f_\theta)}{\partial \theta} \frac{\partial \Phi(f_\theta)}{\partial \theta'} dx]^{-1}[\int_{-\pi}^{\pi} \frac{\partial \Phi(f_\theta)}{\partial \theta} \frac{\partial \Phi(f_\theta)}{\partial \theta'} \sigma^2(x)dx][\int_{-\pi}^{\pi} \frac{\partial \Phi(f_\theta)}{\partial \theta} \frac{\partial \Phi(f_\theta)}{\partial \theta'} dx]^{-1}.$$ ▯

In this section we shall propose a certain parameterization of $f_\theta(x)$, and then give an explicit estimate of θ which can be expressed in a closed form. As we said in the Introduction we shall fit the following spectral model :

(4.1) $$f_\theta(x) = \Phi^{-1}\{ \sum_{j=1}^{p} \theta_j \, e_j(x) \}.$$

Of course we assume that $\Phi(.)$ and $e_j(x)$ are chosen in advance so that $f_\theta(x) = f_\theta(-x)$, $f_\theta(x) > 0$ for all $x \in [-\pi,\pi]$. It may be noted that in the special case that $\Phi(\omega) = \log \omega$, $e_1(x) = 1/\sqrt{2\pi}$ and $e_j(x) = \cos\{(j-1)x\}/\sqrt{\pi}$ ($j = 2,\ldots,p$), the model (4.1) becomes Bloomfield's [4] exponential model. In general, the model (4.1) is " non-structural" in the sense that the parameter vector θ is not necessarily explanatory. But it has an advantage that the statistic T_n can be expressed in an explicit form as follows.

(4.2) $$T_n = [\int_{-\pi}^{\pi} e_1(x) \, K\{ I_n(x) \}dx,\ldots, \int_{-\pi}^{\pi} e_p(x) \, K\{ I_n(x) \}dx]'.$$

If we cannot compute the above integral easily we shall approximate the integrals by discrete sums.

Hereafter in this section we assume $g = f_\theta$, and discuss the asymptotic efficiency of T_n for a class $\Psi = \{ \Phi(\omega) = \omega^\beta, \ 0 < \beta < \infty \}$, i.e., $f_\theta(x) = \{ \sum_{j=1}^{p} \theta_j \, e_j(x) \}^{1/\beta}$. By Example 3.1 and Remark 4.2, we can see that the asymptotic covariance matrix of T_n is equal to

(4.3) $$4\pi \{ \frac{\Gamma(2\beta+1)}{\Gamma(\beta+1)^2} - 1 \} \int_{-\pi}^{\pi} f_\theta(x)^{2\beta} \, F(x) \, dx ,$$

where $F(x) = e(x)\, e(x)'$, $e(x) = (e_1(x),\ldots,e_p(x))'$. On the other hand, the asymptotic covariance matrix of the efficient estimate given in Remark 2.2 is written in the present case as

(4.4) $$4\pi\beta^2 \left[\int_{-\pi}^{\pi} f_\theta(x)^{-2\beta} F(x)\, dx \right]^{-1}.$$

The following theorem provides a condition for (4.3) to be identical with (4.4).

Theorem 4.2. (Taniguchi[27]). Let $\Phi(\omega) = \omega^\beta$, $0 < \beta < \infty$. Then T_n is asymptotically efficient if and only if $\beta = 1$ and $f_\theta(x)$ is constant on $[-\pi,\pi]$. ▯

This theorem means that the estimate T_n is not recommendable except for a trivial case from the point of view of the efficiency. But it has an advantage that it can be expressed in the explicit form given by (4.2). Thus we can also use this estimate for an initial consistent estimate of θ, which will be useful in getting the efficient estimate discussed in Section 2. Finally we wish to see how inefficient it can be. We introduce

$$\mathrm{eff}(\alpha) = \frac{\text{asymptotic variance of } \alpha'(\text{m.l.e.})}{\text{asymptotic variance of } \alpha'(T(.))},$$

where $\boldsymbol{\alpha}$ is a p-vector with $\boldsymbol{\alpha}'\boldsymbol{\alpha} = 1$.

Theorem 4.3. (Taniguchi[27]). For any p-vector α, eff(α) is bounded below by

$$\beta^2 \left\{ \frac{\Gamma(2\beta+1)}{\Gamma(\beta+1)^2} - 1 \right\}^{-1} (4\, f_{\ell,\theta}^{2\beta}\, f_{u,\theta}^{2\beta})(f_{\ell,\theta}^{2\beta} + f_{u,\theta}^{2\beta})^{-2},$$

where $f_{\ell,\theta} = \min_{x\in[-\pi,\pi]} f_\theta(x)$, $f_{u,\theta} = \max_{x\in[-\pi,\pi]} f_\theta(x)$. ▯

The above lower bound has a property that it tends to zero if $f_{\ell,\theta} \ll f_{u,\theta}$.

5. Asymptotic properties of the quasi-Gaussian maximum likelihood estimate for a non-Gaussian linear process

In Section 2 we have investigated some asymptotic properties of the q-G.M.L.E. (quasi-Gaussian maximum likelihood estimate) for a scalar Gaussian process. In this section we shall investigate some asymptotic properties of the q-G.M.L.E. when the underlying process is not necessarily Gaussian.

Let $\{ \mathbb{X}(t) = (X_1(t),\ldots,X_s(t))' \; ; \; t \in Z \}$ be a vector time series with s components generated by

$$(5.1) \qquad \mathbb{X}(t) = \sum_{j=0}^{\infty} C(j)\, \mathbf{e}(t-j), \qquad C(0) = I_s,$$

where $\{ \mathbf{e}(t) = (e_1(t),\ldots,e_s(t))' \}$ satisfies $E\, \mathbf{e}(t) = \mathbf{0}_s$, $E\, \mathbf{e}(t)\, \mathbf{e}(t)' = K$ and $E\, \mathbf{e}(t)\, \mathbf{e}(u)' = \mathbf{0}_{s \times s}$, $t \neq u$. It will be assumed that

$$\mathrm{tr} \sum_{j=0}^{\infty} C(j)\, K\, C(j)' < \infty ,$$

and that $\{ \mathbf{e}(t) \}$ is strongly stationary. Then the spectral density matrix of $\mathbb{X}(t)$ is

$$g(x) = \frac{1}{2\pi}\, k(x)\, K\, k(x)^*,$$

where $k(x) = \sum_{j=0}^{\infty} C(j)\, e^{ijx}$. Now we set the following conditions (5.2) - (5.5) :

(5.2) The spectral density matrix $g(x)$ is positive definite for all $x \in [-\pi,\pi]$.

(5.3) Every element of $g(x)$ belongs to Λ_α, the Lipshitz class of degree α, $1/2 < \alpha \leq 1$.

(5.4) The process $\{ \mathbb{X}(t) \}$ is uniformly mixing.

$$(5.5) \qquad \sum\sum\sum_{n,p,q=-\infty}^{\infty} | \mathrm{cum}_{abcd}(0,n,p,q) | < \infty,$$

where $\text{cum}_{abcd}(0,n,p,q) = \text{cumulant}\{X_a(t), X_b(t+n), X_c(t+p), X_d(t+q)\}$.

Let $\mathcal{P}$ be the set of all spectral density matrices which satisfy (5.2) and (5.3). For a parametric family of spectral density matrices $\{ f_\theta ; \theta \in \Theta \}$, $\Theta \subset R^p$, we define

$$D(f_t, g) = \int_{-\pi}^{\pi} \{ \log \det f_t(x) + \text{tr}\, f_t(x)^{-1} g(x) \}\, dx$$

and then the functional T by

$$D(f_{T(g)}, g) = \min_{t \in \Theta} D(f_t, g) \qquad \text{for every } g \in \mathcal{P}. \tag{5.6}$$

If for a sequence $\{ g_n \}$, $g_n \in \mathcal{P}$,

$$\int_{-\pi}^{\pi} \text{tr}\, \psi(x)\, g_n(x)\, dx \to \int_{-\pi}^{\pi} \text{tr}\, \psi(x)\, g(x)\, dx \qquad \text{as } n \to \infty$$

for every continuous s×s matrix function $\psi(x)$, then we denote $g_n \overset{W}{\to} g$. Suppose that a stretch, $X(0), \dots, X(n-1)$ of the series $X(t)$ is observed. We define the periodogram matrix by

$$I_n(x) = (2\pi n)^{-1} \{ \sum_{t=0}^{n-1} X(t) \exp(itx) \} \{ \sum_{t=0}^{n-1} X(t) \exp(itx) \}^*.$$

Now we assume that every element of the matrix $f_\theta(x)$ is a twice continuously differentiable function of $\theta \in \Theta$, and that the second derivatives of these elements are continuous in $x \in [-\pi, \pi]$. Then we have the following theorem.

Theorem 5.1. (Taniguchi[30]). Suppose that T(g) exists uniquely and lies in *Int*(Θ) and that

$$M_g = \int_{-\pi}^{\pi} \{ \frac{\partial^2}{\partial\theta\partial\theta'} \text{tr}\, f_\theta(x)^{-1} g(x) + \frac{\partial^2}{\partial\theta\partial\theta'} \log \det f_\theta(x) \}_{\theta = T(g)}\, dx$$

is a non-singular matrix. Assume that $\{ X(t) \}$ defined by (5.1) satisfies the conditions (5.2) - (5.5). Then

(1) $T(I_n) \to T(g)$ a.s., and

(2) the distribution of the vector $\sqrt{n}(\, T(I_n) - T(g)\,)$ tends to the normal distribution

$$N(\, 0_p,\, M_g^{-1}\, V\, M_g^{-1}\,),$$

as $n \to \infty$, where $V = \{ V_{j\ell} \}$, a $p\times p$ matrix, is defined by

$$V_{j\ell} = 4\pi \int_{-\pi}^{\pi} \mathrm{tr}\{\, g(x) \frac{\partial}{\partial\theta_j} f_\theta(x)^{-1}\, g(x) \frac{\partial}{\partial\theta_\ell} f_\theta(x)^{-1} \}_{\theta=T(g)}\, dx$$

$$+ 2\pi \sum\sum\sum\sum_{r,t,u,v=1}^{s} \iint_{-\pi}^{\pi} \frac{\partial}{\partial\theta_j} f_\theta^{(r,t)}(x)^{-1} \frac{\partial}{\partial\theta_\ell} f_\theta^{(u,v)}(y)^{-1} \Big|_{\theta=T(g)}$$

$$\times f_{rtuv}(-x,y,-y)\, dxdy,$$

for which $f_\theta^{(r,t)}(x)^{-1}$ is the (r,t)th element of the matrix $f_\theta(x)^{-1}$, and

$$f_{rtuv}(x,y,z) = \frac{1}{(2\pi)^3} \sum\sum\sum_{n,p,q=-\infty}^{\infty} \mathrm{cum}_{rtuv}(0,n,p,q)\, \exp\{-i(xn+yp+zq)\}. \qquad \square$$

Also we have the following theorem.

Theorem 5.2. (Taniguchi[30]). In addition to the assumptions of Theorem 5.1, furthermore if

(5.7) $$\mathrm{cum}\{\, e_a(n_1),\, e_b(n_2),\, e_c(n_3),\, e_d(n_4)\,\} = \begin{cases} \kappa_{abcd} & \text{if } n_1 = n_2 = n_3 = n_4, \\ 0 & \text{otherwise}, \end{cases}$$

then the asymptotic covariance matrix of the quasi-Gaussian maximum likelihood estimate is written as

$$M_g^{-1}\, U\, M_g^{-1},$$

where $U = \{ U_{j\ell} \}$, a $p\times p$ matrix,

$$U_{j\ell} = 4\pi \int_{-\pi}^{\pi} \mathrm{tr}\{ g(x) \frac{\partial}{\partial\theta_j} f_\theta(x)^{-1} g(x) \frac{\partial}{\partial\theta_\ell} f_\theta(x)^{-1} \}_{\theta=T(g)} dx$$

$$+ \sum_{a,b,c,d=1}^{s}\sum\sum\sum \kappa_{abcd} [\frac{1}{2\pi} \int_{-\pi}^{\pi} k(x)^* \frac{\partial}{\partial\theta_j} f_\theta(x)^{-1} k(x) dx]_{ab}$$

$$\times [\frac{1}{2\pi} \int_{-\pi}^{\pi} k(x)^* \frac{\partial}{\partial\theta_\ell} f_\theta(x)^{-1} k(x) dx]_{cd} \Big|_{\theta=T(g)},$$

$[\ ]_{ab}$ denoting the (a,b)-element of the matrix in the brackets. ‖

Suppose that $f_\theta(x)$ can be expressed as

$$f_\theta(x) = \frac{1}{2\pi} k_1(e^{ix})\ K_1\ k_1(e^{ix})^*,$$

where $k_1(e^{ix}) = \sum_{j=0}^{\infty} C_1(j)\, e^{ijx}$, $C_1(0) = I_s$, and $\det k_1(z)$ is bounded and bounded away from zero for $|z| \le 1$, and K_1 is a positive definite s×s-matrix. If the unknown parameter vector θ specifies only $k_1(.)$, then we say that θ is an innovation free parameter. If θ is innovation free, then the relation

$$(5.8) \qquad \frac{\partial}{\partial\theta} \int_{-\pi}^{\pi} \mathrm{tr}\ f_\theta(x)^{-1} g(x)\, dx \Big|_{\theta=T(g)} = \mathbf{0}_p$$

is satisfied. Under the condition (5.7) we consider to estimate the innovation free parameter.

Remark 5.1. Under the assumptions of Theorem 5.2, in the case of $g(x) = f_\theta(x)$, we can show that if θ is the innovation free parameter, then

$$\int_{-\pi}^{\pi} k(x)^* [\frac{\partial}{\partial\theta_j} f_\theta(x)^{-1}]\, k(x)\, dx = \mathbf{0}_{s\times s}, \quad j = 1,\dots,p.$$

Thus the asymptotic covariance matrix of the q-G.M.L.E. of the innovation free parameters is independent of the fourth cumulant κ_{abcd} (i.e., non-Gaussian robust). However, in the case of $g(x) \ne f_\theta(x)$, it is noted that the relation

$$(5.9) \qquad \int_{-\pi}^{\pi} k(x)^* \{ \frac{\partial}{\partial\theta_j} f_\theta(x)^{-1} \}_{\theta=T(g)} k(x)\, dx = \mathbf{0}_{s\times s},\ j = 1,\ldots,p,$$

is not satisfied generally. For example, suppose that the true spectral density matrix is

$$g(x) = \{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1/2 \\ -1/2 & 1 \end{pmatrix} e^{ix} \} \{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1/2 \\ -1/2 & 1 \end{pmatrix} e^{ix} \}^*,$$

(MA(1)), and the fitted spectral model is

$$f_\theta(x)^{-1} = \{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \theta_1 & \theta_4 \\ \theta_3 & \theta_2 \end{pmatrix} e^{-ix} \} \{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} \theta_1 & \theta_4 \\ \theta_3 & \theta_2 \end{pmatrix} e^{-ix} \}^*,$$

(AR(1)). Then it is not difficult to show

$$\int_{-\pi}^{\pi} k(x)^* \{ \frac{\partial}{\partial\theta_3} f_\theta(x)^{-1} \}_{\theta=T(g)} k(x)dx = \begin{pmatrix} 54\pi/65 & 84\pi/65 \\ 84\pi/65 & -54\pi/65 \end{pmatrix} \neq \mathbf{0}_{2\times 2}.$$

This implies that, in the case of $g(x) \neq f_\theta(x)$, even if (5.7) is satisfied, the q-G.M.L.E. of the innovation free parameters is not always non-Gaussian robust. ▯

If s = 1, i.e., the underlying process is scalar, then we have the following unified result.

Remark 5.2. Assume that (5.7) is satisfied for s = 1. Then q-G.M.L.E. of the innovation free parameter is non-Gaussian robust even if $g(x) \neq f_\theta(x)$. In fact if θ is the innovation free parameter, then (5.8) is satisfied and it implies (5.9) for s = 1. ▯

6. Selection of the order of the spectral density model for a stationary process

Let $\{ X(t) \}$ be a stationary process with zero mean and a specral density $g(x)$, to which we shall fit a kth order parametric spectral model $f_{\tau(k)}(x)$. Without assuming Gaussianity we can obtain an estimate of $\tau(k)$, say $\hat{\tau}(k)$, by maximizing the quasi-Gaussian likelihood of the model. We can then construct the best linear predictor of $X(t)$, which is computed on the basis of the estimated spectral density $f_{\hat{\tau}(k)}(x)$. An asymptotic lower bound of the mean square error of the predictor will be given. The bound is attained if k is selected by using Akaike's information criterion.

Now we shall assume that $\{ X(t) : t \in Z \}$ is a scalar linear process

$$(6.1) \qquad X(t) = \sum_{j=0}^{\infty} a(j;\theta)\, e(t-j), \qquad a(0;\theta) = 1 ,$$

satisfying

$$(6.2) \qquad \sum_{j=0}^{\infty} j^{\beta} \mid a(j;\theta) \mid < \infty \qquad \text{for some } \beta > 1,$$

where $a(j;\theta)$ are known functions of an infinite dimensional parameter vector $\theta = (\theta_1, \theta_2, \ldots)'$, and $e(j)$ are independently and identically distributed with finite cumulant κ_s, $s = 1,\ldots,16$. The spectral density of $X(t)$ is then written as

$$(6.3) \qquad g(x;\theta) = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^{\infty} a(j;\theta) \exp(ijx) \right|^2,$$

where $\sigma^2 = E[\, e(j)^2]$. For notational convenience, we write sometimes simply $g(x)$ in the place of $g(x;\theta)$. We need the following assumptions (6.4) - (6.6) :

(6.4) $g(x;\theta)$ is three times differentiable with respect to each coordinate of $\theta \in \Theta$, and the third order derivative is a continuous function of (x,θ)

$\varepsilon\ [-\pi,\pi]\times\Theta$, where Θ is a compact set in R^{∞}.

(6.5) The associated power series

$$A(z) = 1 + a(1;\theta)\, z + a(2;\theta)\, z^2 + \ldots$$

is not zero for $|z| \le 1$ and for any $\theta \in \Theta$.

(6.6) The true value of θ belongs to the interior of Θ, and it has infinitely many nonzero elements.

Suppose that a stretch, $X(t)$ ($t = 0,\ldots,n-1$) of the time series $X(t)$ is observed. Let $f_{\tau(k)}(x)$ be a spectral model with $(k+1)$-dimensional unknown parameter vector $\tau(k) = (\ \sigma^2(k),\ \theta(k)'\ )'$, $\sigma^2(k) > 0$, $\theta(k) = (\ \theta_1(k),\ldots,\theta_k(k)\ )'$. We assume that $f_{\tau(k)}(x)$ is parametrized so that

$$f_{\tau(k)}(x) = \frac{\sigma^2(k)}{2\pi}\, h_{\theta(k)}(x) = \frac{\sigma^2(k)}{2\pi}\, |\ \dot{h}_{\theta(k)}(e^{ix})\ |^2,$$

where $\dot{h}_{\theta(k)}(0) = 1$. When the spectral model $f_{\tau(k)}(x)$ is applied to $\{\ X(t)\}$, we define the parameter vector $\underline{\tau}(k) = (\ \underline{\sigma}^2(k),\ \underline{\theta}(k)'\ )'$ by

$$(6.7)\qquad \min_{\tau(k)\,\varepsilon\,\Theta_{k+1}} \int_{-\pi}^{\pi} \{\ \log f_{\tau(k)}(x) + g(x)/f_{\tau(k)}(x)\ \}\, dx$$

$$= \int_{-\pi}^{\pi} \{\ \log f_{\underline{\tau}(k)}(x) + g(x)/f_{\underline{\tau}(k)}(x)\ \}\, dx,$$

where Θ_{k+1} is a compact set in R^{k+1}. We can estimate $\underline{\tau}(k)$ by the quasi-Gaussian maximum likelihood estimate $\hat{\tau}(k) = (\ \hat{\sigma}^2(k),\ \hat{\theta}(k)'\ )'$ which is a solution of

$$(6.8)\qquad \min_{\tau(k)\,\varepsilon\,\Theta_{k+1}} \int_{-\pi}^{\pi} \{\ \log f_{\tau(k)}(x) + I_n(x)/f_{\tau(k)}(x)\ \}\, dx$$

$$= \int_{-\pi}^{\pi} \{\ \log f_{\hat{\tau}(k)}(x) + I_n(x)/f_{\hat{\tau}(k)}(x)\ \}\, dx,$$

where $I_n(x) = (2\pi n)^{-1} \mid \sum_{t=0}^{n-1} X(t) \exp(-itx) \mid^2$ is the periodogram of $X(t)$.

Now we postulate the following assumptions (6.9) - (6.15):

(6.9) The number of the parameters, $k+1$, is in the range $1 \leq k \leq K_n$, where $K_n \to \infty$ and $K_n/\sqrt{n} \to 0$ as $n \to \infty$.

(6.10) The spectral density $f_{\tau(k)}(x)$ is three times differentiable with respect to $\tau(k) \in \Theta_{k+1}$. The third order derivative is a continuous function of $(x,\tau(k)) \in [-\pi,\pi]\times\Theta_{k+1}$, and as a function of $x \in [-\pi,\pi]$, the first and the second derivatives satisfy the Lipshitz condition of order 1.

(6.11) $\mid \dot{h}_{\theta(k)}(z) \mid$ is bounded and bounded away from zero for $|z| \leq 1$.

(6.12) For any $1 \leq k \leq K_n$, the $k\times k$ matrix

$$H_k = \int_{-\pi}^{\pi} \frac{\partial^2 h_{\theta(k)}(x)^{-1}}{\partial\theta(k)\,\partial\theta(k)'}\Big|_{\underline{\theta}(k)} g(x)\,dx,$$

is non-singular.

(6.13) The absolute sums of all the elements in each row of the matrices

$$H_k,\ H_k^{-1/2} \quad \text{and} \quad \int_{-\pi}^{\pi} \frac{\partial h_{\theta(k)}(x)^{-1}}{\partial\theta(k)} \frac{\partial h_{\theta(k)}(x)^{-1}}{\partial\theta(k)'}\Big|_{\underline{\theta}(k)} g(x)^2 dx$$

are bounded uniformly with respect to each row as $k \to \infty$.

(6.14) Let

$$W(r,j,m) = \int_{-\pi}^{\pi} \frac{\partial^3 h_{\theta(k)}(x)^{-1}}{\partial\theta_r(k)\partial\theta_j(k)\partial\theta_m(k)}\Big|_{\underline{\theta}(k)} g(x)\,dx\ .$$

Then $\sum_{j,m=1}^{k} \mid W(r,j,m) \mid$ is bounded uniformly in r as $k \to \infty$.

(6.15) For β in (6.2),

$$\mid f_{\underline{\tau}(k)}(x) - g(x) \mid = O(k^{-\beta}) \quad \text{for all } x \in [-\pi,\pi].$$

Example 6.1. Suppose that the true spectral density $g(x;\theta)$ is parametrized so that

$$g(x;\theta) = \frac{\sigma^2}{2\pi} \frac{\left| \sum_{j=0}^{\infty} \mu_{1,j} \exp(ijx) \right|^2}{\left| \sum_{j=0}^{\infty} \mu_{2,j} \exp(ijx) \right|^2} ,$$

$\mu_{1,0} = \mu_{2,0} = 1$, $\theta = (\sigma^2, \mu_{1,1}, \mu_{2,1}, \mu_{1,2}, \mu_{2,2}, \ldots)'$, satisfying $\sum_{j=0}^{\infty} j^{\beta} | \mu_{1,j} | < \infty$, $\sum_{j=0}^{\infty} j^{\beta} | \mu_{2,j} | < \infty$. Also we assume that $\sum_{j=0}^{\infty} \mu_{1,j} z^j$ and $\sum_{j=0}^{\infty} \mu_{2,j} z^j$ are not zero for $|z| \leq 1$. We choose an ARMA(autoregressive moving average) spectral model

$$f_{\tau(k)}(x) = \frac{\sigma^2(k)}{2\pi} \frac{\left| \sum_{j=0}^{p} \theta_{1,j} \exp(ijx) \right|^2}{\left| \sum_{j=0}^{q} \theta_{2,j} \exp(ijx) \right|^2} ,$$

$\theta_{1,0} = \theta_{2,0} = 1$, $\tau(k) = (\sigma^2(k), \theta_{1,1}, \ldots, \theta_{1,p}, \theta_{2,1}, \ldots, \theta_{2,q})'$, $p + q = k$, where $\sum_{j=0}^{p} \theta_{1,j} z^j$ and $\sum_{j=0}^{q} \theta_{2,j} z^j$ are not zero for $|z| \leq 1$. Further assume that $\sum_{j=0}^{p} j^{\beta} | \theta_{1,j} |$ and $\sum_{j=0}^{q} j^{\beta} | \theta_{2,j} |$ are bounded as $p = p(k)$ and $q = q(k)$ tend to infinity so that $p(k)/q(k) \to 1$. Then it is not so hard to show that $f_{\tau(k)}(x)$ satisfies (6.4) - (6.15). ▯

Let $\hat{X}(t)$ be a predictor obtained by fitting the spectral density $f_{\hat{\tau}(k)}(x)$. Then the mean square error is

$$E| X(t) - \hat{X}(t) |^2$$

$$= \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g(x)}{f_{\hat{\tau}(k)}(x)} dx \exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \frac{f_{\hat{\tau}(k)}(x)}{g(x)} dx \right\} - 1 \right]\sigma^2 + \sigma^2$$

(Grenander and Rosenblatt[12]). We can measure the goodness of the estimated spectral model $f_{\hat{\tau}(k)}(x)$ by

(6.16) $D(f_{\hat{\tau}(k)},g)$

$$= \frac{1}{2\pi} \left\{ \int_{-\pi}^{\pi} \frac{g(x)}{f_{\hat{\tau}(k)}(x)} dx \right\} \exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \frac{f_{\hat{\tau}(k)}(x)}{g(x)} dx \right\} - 1.$$

Theorem 6.1.(Taniguchi[28]). Assume (6.4) - (6.6) and (6.9) - (6.14). Then

(6.17) $D(f_{\hat{\tau}(k)},g) = M(f_{\hat{\tau}(k)},g)$ + higher order terms,

where

$$M(f_{\hat{\tau}(k)},g) = D(h_{\underline{\theta}(k)},g) + \frac{1}{4\pi}(\hat{\theta}(k) - \underline{\theta}(k))'H_k(\hat{\theta}(k) - \underline{\theta}(k))$$
$$\times\exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \frac{h_{\underline{\theta}(k)}(x)}{g(x)} dx \right\}$$

and " higher order terms " means stochastically higher order terms as $n \to \infty$ compared with the second term in the right hand side of $M(f_{\hat{\tau}(k)},g)$ uniformly in k. ▯

From Theorem 6.1, neglecting the higher order terms, we can measure the goodness of $f_{\hat{\tau}(k)}(x)$ by $M(f_{\hat{\tau}(k)},g)$. Putting

$$R(n,k) = \frac{k}{n} \exp\left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \frac{f_{\underline{\tau}(k)}(x)}{g(x)} dx \right\} + D(h_{\underline{\theta}(k)},g),$$

we have the following theorem.

Theorem 6.2.(Taniguchi[28]). Assume (6.4) - (6.6) and (6.9) - (6.15).

Then

$$\operatorname*{p-lim}_{n\to\infty} \{ \max_{1 \le k \le K_n} \left| \frac{M(f_{\hat{\tau}(k)}, g)}{R(n,k)} - 1 \right| \} = 0. \qquad ▯$$

We shall define a sequence $\{ k_n^* \}$ by

$$R(n,k_n^*) = \min_{1 \le k \le K_n} R(n,k).$$

Define an asymptotically efficient order selection extending the concept of Shibata[22] in the case of autoregressive model.

Definition 6.1. An order selection $\tilde{k}$ is said to be asymptotically efficient if

$$\operatorname*{p-lim}_{n\to\infty} M(f_{\hat{\tau}(\tilde{k})}, g)/R(n,k_n^*) = 1. \qquad ▯$$

Now let us select the order of $f_{\tau(k)}(x)$ by the criterion AIC (Akaike's information criterion, Akaike[2],[1]), although Gaussianity on which AIC is based is not assumed here. The AIC for $f_{\tau(k)}(x)$ can be written as

$$AIC(k+1) = n \log \hat{\sigma}^2(k) + 2k.$$

Define $Q(n,k)$ by

$$n \exp\{ \frac{1}{n} AIC(k+1) \} = n\, \hat{\sigma}^2(k) \exp \frac{2k}{n}$$

$$= n\, R(n,k) + Q(n,k).$$

Theorem 6.3.(Taniguchi[28]). Assume (6.4) - (6.6) and (6.9) - (6.15). Then

$$\operatorname*{p-lim}_{n\to\infty} \max_{1 \le k \le K_n} | Q(n,k) - Q(n,k_n^*) | / \{ n\, R(n,k) \} = 0. \qquad ▯$$

Let $\hat{k}$ be the order which is selected by minimizing AIC(k+1). By Theorems 6.2 and 6.3, we can show

$$\operatorname*{p-lim}_{n\to\infty} M(f_{\hat{\tau}(\hat{k})}, g)/R(n, k_n^*) = 1.$$

That is, the order selection $\hat{k}$ is asymptotically efficient in the sense of Definition 6.1 although Gaussianity of $\{ X(t) \}$ is not assumed.

Remark 6.1. Shibata[22] showed that a kth order autoregressive model fitting by Akaike's information criterion for a Gaussian linear process with infinite unknown parameters is asymptotically efficient in the sense of Definition 6.1. Needless to say our results can be applied to the fitting of ARMA spectral density model described in Example 6.1. Thus we have extended his results to the case when the process is not necessarily Gaussian and the model is not necessarily autoregressive. ▯

7. Applications for time series regression and interpolation

There is a kind of similarity between the estimation problem in a time series regression model and the interpolation problem in a stationary process. From a unified point of view we shall again look at the estimation problem in the time series regression model, and propose a parametric method which gives an efficient estimate of the regression coefficient vector in the model.

Let $\{ X(t);\ t \in Z \}$ be a non-deterministic stationary process with zero mean and a spectral density $g(x)$. Suppose that values of $X(t)$ are observed for all t except $t \in A_p$, where $A_p = \{ 1,\dots,p \}$. We shall interpolate the unknown values $X(t)$, $t \in A_p$. That is, we seek a p-vector whose components are linear combinations of $X(t)$, $t \in Z - A_p$, and which minimizes the error of interpolation in the mean square sense. Let $M\{\dots\}$ denote the closed linear manifold generated by the elements in the braces with respect to the norm $\|.\|^2 = \int_{-\pi}^{\pi} |.|^2 g(x)dx$. For the present we assume that $g(x)$ belongs to the class

$$\Psi_0 = [\ g\ ;\ g(x)^{-1} \in M\{ e^{-ijx}\ ;\ j \in Z - A_{p-1} - \bar{A}_{p-1} \}\ ],$$

where $\bar{A}_{p-1} = \{ -1,\dots,-p+1 \}$. If $p = 1$, i.e., the usual interpolation problem with one missing time point, we define $A_0 = \bar{A}_0 = \phi$. Mathematically our problem can be described as follows. We seek a response function $\underset{\sim}{h}(x)$ such that

$$(7.1) \qquad \mathrm{tr} \int_{-\pi}^{\pi} [\ \underset{\sim}{e}(x) - \underset{\sim}{h}(x)\ ]\ g(x)\ [\ \underset{\sim}{e}(x) - \underset{\sim}{h}(x)\ ]^*\ dx$$

is minimized. Here $\underset{\sim}{e}(x) = (e^{-ix},\dots, e^{-ipx})'$, and each component of $\underset{\sim}{h}(x)$ belongs to $M\{ e^{-ijx}\ ;\ j \in Z - A_p \}$. For our multiple time points interpolation we have the following theorem.

Theorem 7.1.(Taniguchi[26]). Let { X(t) } be a stationary process with zero mean and a spectral density g(x) which belongs to Ψ_0. Then the response function $\underset{\sim}{h}(x)$ of the optimal interpolating filter and the interpolation error matrix Σ are given by

(7.2) $$\underset{\sim}{h}(x) = [\ I_p - 2\pi\{ \int_{-\pi}^{\pi} (\ F(x)/g(x)\ )\ dx\ \}^{-1}/g(x)\ ]\ \underset{\sim}{e}(x),$$

(7.3) $$\Sigma = 2\pi\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} (\ F(x)/g(x)\ )\ dx\ \}^{-1},$$

where $F(x) = \underset{\sim}{e}(x)\ \underset{\sim}{e}(x)^*$. ▯

In most natural phenomena, the true spectral density is not known a priori. Thus it is of considerable interest to see what happens when an interpolator is computed on the basis of a ' pseudo ' spectral density $f_\theta(x) \in \Psi_0$ instead of the true one, g(x), belonging to the class $\Psi = M\{ e^{-ijx}\ ;\ j \in Z \}$. In this situation we can show the following theorem.

Theorem 7.2.(Taniguchi[26]). Let { X(t) } be a stationary process with zero mean and a spectral density $g(x) \in \Psi$. If the interpolator is computed on the basis of a pseudo spectral density $f_\theta(x) \in \Psi_0$, then the response function $\underset{\sim}{h}_1(x)$ of the pseudo interpolation filter and the pseudo interpolation error matrix Σ_1 are given by

(7.4) $$\underset{\sim}{h}_1(x) = [\ I_p - 2\pi\{ \int_{-\pi}^{\pi} (\ F(x)/f_\theta(x)\ )\ dx\ \}^{-1}/f_\theta(x)\ ]\ \underset{\sim}{e}(x),$$

(7.5) $$\Sigma_1 = \{ \frac{1}{2\pi} \int_{-\pi}^{\pi} (\ F(x)/f_\theta(x)\ )\ dx\ \}^{-1} \{ \int_{-\pi}^{\pi} \frac{g(x)}{f_\theta(x)^2} F(x)\ dx\ \}$$
$$\times \{ \frac{1}{2\pi} \int_{-\pi}^{\pi} (\ F(x)/f_\theta(x)\ )\ dx\ \}^{-1}.$$ ▯

Hereafter we shall look at the results of Grenander and Rosenblatt[12] and Rozanov[20] in relation to the previous interpolation problem. Let

(7.6) $$y(t) = z(t) + u(t)$$

be a process in discrete time, where $z(t) = \alpha_1 z_1(t)+\dots+ \alpha_p z_p(t)$ with known functions $z_1(t),\dots,z_p(t)$ and an unknown parameter $\underset{\sim}{\alpha} = (\alpha_1,\dots,\alpha_p)'$ and $\{ u(t) \}$ is a stationary process with zero mean. Let $\psi_{j,n}(x)$, $j = 1,\dots,p$, be solutions of the equations

$$\int_{-\pi}^{\pi} e^{-ixt}\, \overline{\psi_{j,n}(x)}\, dx = z_j(t), \qquad 0 \le t \le n-1,$$

i.e., $$\psi_{j,n}(x) = \frac{1}{2\pi} \sum_{s=0}^{n-1} z_j(s)\, e^{-isx} .$$

The observed random process is harmonizable, that is, we can express $y(t) = \int_{-\pi}^{\pi} e^{-ixt} z_y(dx)$ with stochastic measure $z_y(dx) = [\, \alpha_1\overline{\psi_{1,n}(x)} +\dots+ \alpha_p\overline{\psi_{p,n}(x)}\,]\, dx + z_u(dx)$, where $z_u(dx)$ is the stochastic spectral measure of the process $u(t)$. We assume that $u(t)$ has the spectral density $g(x)$, which is continuous and positive. Further we assume that the functions $z_1(t),\dots,z_p(t)$ satisfy Grenander's conditions with the regression spectrum matrix $H(x)$. Now we seek a linear unbiased estimate $\hat{\underset{\sim}{\alpha}}$ of $\underset{\sim}{\alpha}$, which can be represented in the following spectral form ;

(7.7) $$\hat{\underset{\sim}{\alpha}} = \int_{-\pi}^{\pi} \underset{\sim}{r}(x)\, z_y(dx),$$

where $\underset{\sim}{r}(x)$ is a p-vector function. We denote $\underset{\sim}{\psi}_n(x) = (\psi_{1,n}(x),\dots,\psi_{p,n}(x))'$ and $D_n = \mathrm{diag}\{ (\sum_{t=0}^{n-1} |z_1(t)|^2)^{1/2},\dots, (\sum_{t=0}^{n-1} |z_p(t)|^2)^{1/2} \}$.

Substituting

$$\underset{\sim}{e}(x) = [\, D_n^{-1} \int_{-\pi}^{\pi} \underset{\sim}{\psi}_n(x)\, \underset{\sim}{\psi}_n(x)^*\, dx\, D_n^{-1}\,]^{-1} D_n^{-1}\, \underset{\sim}{\psi}_n(x)$$

and

$$h(x) = [\, D_n^{-1} \int_{-\pi}^{\pi} \psi_n(x)\, \psi_n(x)^* \, dx \, D_n^{-1} \,]^{-1} D_n^{-1} \eta_n(x)$$

into (7.1), where $\eta_n(x)$ is orthogonal to $\psi_n(x)$ with respect to the L^2-norm, we can see that the best linear unbiased estimate for α is given by

$$(7.8) \quad \hat{\alpha}_1 = [\, \int_{-\pi}^{\pi} g(x)^{-1} \psi_n(x)\, \psi_n(x)^* \, dx \,]^{-1} \int_{-\pi}^{\pi} g(x)^{-1} \psi_n(x)\, z_y(dx).$$

Replacing $F(x)dx$ by $H(dx)$ in (7.3), we have the following classical result.

Theorem 7.3. (Grenander and Rosenblatt[12]). The asymptotic covariance matrix of $D_n(\hat{\alpha}_1 - \alpha)$ is given by

$$2\pi [\, \int_{-\pi}^{\pi} \frac{1}{g(x)} H(dx) \,]^{-1}. \qquad ▯$$

Now we shall proceed to the non-standard situation. Assuming that the process $u(t)$ has a pseudo spectral density $f_\theta(x)$ although the true one is $g(x)$, and replacing $g(x)$ in $\hat{\alpha}_1$ by $f_\theta(x)$ we have the pseudo best linear unbiased estimate of α which is given by

$$(7.9) \quad \hat{\alpha}_2 = [\, \int_{-\pi}^{\pi} f_\theta(x)^{-1} \psi_n(x)\, \psi_n(x)^* \, dx \,]^{-1} \int_{-\pi}^{\pi} f_\theta(x)^{-1} \psi_n(x)\, z_y(dx).$$

We can see that $\hat{\alpha}_2$ is essentially equivalent to the R-estimate of Rozanov [20] (cf. also Rozanov and Kozlov[21] and Kholevo[19]). In relation to Theorem 7.2, replacing $F(x)dx$ by $H(dx)$ in (7.5), we have the following theorem.

Theorem 7.4. (Rozanov[20]). The asymptotic covariance matrix of $\hat{\alpha}_2$ is given by

$$[\, \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{f_\theta(x)} H(dx)]^{-1} [\, \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g(x)}{f_\theta(x)^2} H(dx)] [\, \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{f_\theta(x)} H(dx)]^{-1}. \quad ▯$$

In general the true residual spectral density g(x) is unknown. In this case Hannan[13] and [14] proposed a nonparametric efficient estimate of $\underset{\sim}{\alpha}$ by the following

$$(7.10) \qquad \hat{\underset{\sim}{\alpha}}_3 = \left[\frac{1}{2m} \sum_{j=-m+1}^{m} \hat{g}\left(\frac{\pi j}{m}\right)^{-1} \hat{F}_{\psi\psi}\left(\frac{\pi j}{m}\right) \right]^{-1}$$

$$\times \left[\frac{1}{2m} \sum_{j=-m+1}^{m} \hat{g}\left(\frac{\pi j}{m}\right)^{-1} \hat{\underset{\sim}{f}}_{\psi y}\left(\frac{\pi j}{m}\right) \right],$$

where $\hat{g}(.)$, $\hat{F}_{\psi\psi}(.)$ and $\hat{\underset{\sim}{f}}_{\psi y}(.)$ are nonparametric consistent spectral estimates for g(x), $\underset{\sim}{\psi}_n(.)\underset{\sim}{\psi}_n(.)^*$ and $\underset{\sim}{\psi}_n(x)z_y(dx)/dx$ respectively. Of course these estimates depend on their window designs. It is obvious that (7.10) is an estimated approximation for (7.8). Hannan[13] and [14] showed that (7.10) is asymptotically efficient.

Now we consider a parametric efficient method. Suppose that u(t) in (7.6) is a linear process which satisfies the equation

$$(7.11) \qquad u(t) = \sum_{j=0}^{\infty} a(j;\theta)\, e(t-j) \; , \qquad a(0;\theta) = 1,$$

where $a(j;\theta)$ ($j = 0,1,2,...$) are known functions of an infinite dimensional unknown parameter $\theta = (\theta_1,\theta_2,...)'$, and $e(j)$ ($j = 0,1,2,...$) are independently and identically distributed with $E[e(j)] = 0$, $E[e(j)^2] = \sigma^2 > 0$, and $E[e(j)^4] < \infty$. We impose the previous assumptions (6.4) - (6.6) on the spectral density g(x) of u(t). Since u(t) is not observable we shall fit a k-th order parametric spectral model $f_{\theta(k)}(x)$, $\theta(k) = (\theta_1(k), ..., \theta_k(k))'$, to the observed residual process $\hat{u}(t) = y(t) - \hat{\alpha}_1 z_1(t) - ... \hat{\alpha}_p z_p(t)$, $0 \le t \le n-1$, where $\hat{\alpha}_1,...,\hat{\alpha}_p$ are the least squares estimates for $\alpha_1,...,\alpha_p$. Suppose that Θ_k is a compact set in R^k. Since $\hat{\underset{\sim}{\alpha}}_1$ in

(7.8) is independent of innovation variance of the residual process, we assume the parameter vector $\theta(k)$ does not contain the innovation parameter. When $f_{\theta(k)}(x)$ is applied we shall define a pseudo true value of $\theta(k)$, say $\underline{\theta}(k)$, by

$$\min_{\theta(k) \in \Theta_k} \int_{-\pi}^{\pi} \frac{g(x)}{f_{\theta(k)}(x)} dx = \int_{-\pi}^{\pi} \frac{g(x)}{f_{\underline{\theta}(k)}(x)} dx$$

and estimate $\underline{\theta}(k)$ by $\hat{\theta}(k)$ satisfying

$$\min_{\theta(k) \in \Theta_k} \int_{-\pi}^{\pi} \frac{\hat{I}_n(x)}{f_{\theta(k)}(x)} dx = \int_{-\pi}^{\pi} \frac{\hat{I}_n(x)}{f_{\hat{\theta}(k)}(x)} dx,$$

where $\hat{I}_n(x) = (2\pi n)^{-1} \left| \sum_{t=0}^{n-1} \hat{u}(t) \exp(-itx) \right|^2.$

Suppose that the number of the parameters, $k = k(n)$, is chosen so that $\{ k(n) \}$ is a sequence which tends to infinity and $k(n) = o(\min(\sqrt{n}, K_n))$,

where $$K_n = O[\min_{1 \le j \le p} \{ \sum_{t=0}^{n-1} | z_j(t) |^2 \}^{1/2}].$$

Also we set the previous assumptions (6.10) - (6.15) for $f_{\theta(k)}(x)$. Now we propose an estimate of $\underset{\sim}{\alpha}$ by

$$(7.12) \quad \hat{\underset{\sim}{\alpha}}_4 = \Big[\frac{1}{2n} \sum_{j=-n+1}^{n} f_{\hat{\theta}(k)}(\tfrac{\pi j}{n})^{-1} \underset{\sim}{\psi}_n(\tfrac{\pi j}{n}) \underset{\sim}{\psi}_n(\tfrac{\pi j}{n})^* \Big]^{-1}$$

$$\times \Big[\frac{1}{2n} \sum_{j=-n+1}^{n} f_{\hat{\theta}(k)}(\tfrac{\pi j}{n})^{-1} \underset{\sim}{\psi}_n(\tfrac{\pi j}{n}) \frac{1}{2\pi} \sum_{t=0}^{n-1} y(t) e^{it(\frac{\pi j}{n})} \Big].$$

Then we have the following result.

Theorem 7.5. (Taniguchi[26]). The distribution of $D_n(\hat{\underset{\sim}{\alpha}}_4 - \underset{\sim}{\alpha})$ converges as $n \to \infty$ to a multivariate normal distribution with zero mean vector and covariance matrix

$$[\ \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{g(x)} H(dx)\ ]^{-1},$$

i.e., $\hat{\alpha}_4$ is asymptotically efficient. ∎

Remark 7.1. We can choose the order $k = k(n)$ by Akaike's information criterion which was shown to be optimal in a certain sense in Section 6. ∎

Remark 7.2. In constructing the parametric estimate (7.12), we are free from the problem of choosing window. ∎

Remark 7.3. Of course we can apply this method to the construction of a parametric optimal interpolator. ∎

As the last topic we shall consider the interpolation problem and the regression problem under the condition that the spectral density of a stationary process concerned is vaguely known (i.e., Huber's ε-contaminated model (Huber[18])). In this situation Hosoya[17] considered the problem of linear prediction for a stationary process. Our approach goes in line with the discussion of robust estimation due to Huber[18] and Hosoya[17]. We shall show that we can get a minimax robust interpolator and regression coefficient estimate for the class of spectral densities $\mathcal{S} = \{ g : g(x) = (1 - \varepsilon) f(x) + \varepsilon h(x), h(x) \in \mathcal{D}_0 , 0 < \varepsilon < 1 \}$, where $f(x)$ is a known spectral density and $\mathcal{D}_0$ is a certain class of spectral densities. At first we consider the case $p = 1$. If the pseudo interpolation error (7.5) and the asymptotic covariance matrix for the R-estimate are computed on the basis of a pseudo spectral density $\phi(x)$, then they can be expressed as the following form

$$V(\phi,g) = \{ \int_{-\pi}^{\pi} \frac{1}{\phi(x)} H(dx) \}^{-2} \{ \int_{-\pi}^{\pi} \frac{g(x)}{\phi(x)^2} H(dx) \}.$$

Thus we shall investigate a minimax property of $V(\phi,g)$. Let $\mathcal{D} = \{ g : g(x)$ is continuous and piecewise smooth on $[-\pi,\pi]$, $g(x) = g(-x) > 0$, for all $x \in [-\pi,\pi] \}$. In the first place we assume that $H(x)$ is dominated by the Lebesgue measure such that $H(dx) = \eta(x)\, dx$, $\eta(x) \in \mathcal{D}$. Let $\mathcal{D}_0 = \{ c(x) \in \mathcal{D} ; \int_{-\pi}^{\pi} c(x)/\eta(x)\, dx = 1 \}$. Suppose that the true spectral density $g(x)$ belongs to $\mathcal{S} = \{ g : g(x) = (1 - \varepsilon) f(x) + \varepsilon h(x)$, $h(x) \in \mathcal{D}_0$, $0 < \varepsilon < 1 \}$, where $f(x)$ is a known spectral density and belongs to $\mathcal{D}_0$. Here we adopt the minimax principle as the measure of robustness. That is, we seek an interpolator and a regression coefficient estimate which minimize $\max_{g \in \mathcal{S}} V(\phi,g)$ with respect to $\phi \in \mathcal{D}_0$. Let m be a positive number, and put

$$E_m = \{ x \in [-\pi,\pi] : m \geq (1 - \varepsilon) f(x)/\eta(x) \},$$

$$F_m = \{ x \in [-\pi,\pi] : m < (1 - \varepsilon) f(x)/\eta(x) \}.$$

We set

$$f_m(x) = \begin{cases} (1 - \varepsilon) f(x), & x \in F_m \\ m\,\eta(x), & x \in E_m. \end{cases}$$

Then we have the following theorem.

Theorem 7.6. (Taniguchi[29]). The function $V(\phi,g)$ has a saddlepoint ; there exists ϕ_0 and g_0 such that

$$\max_{g \in \mathcal{S}} V(\phi_0,g) = V(\phi_0,g_0) = \min_{\phi \in \mathcal{D}_0} V(\phi,g_0),$$

where $\phi_0(x) = g_0(x) = f_m(x)$. ▯

Immediately this theorem means that

$$\min_{\phi \in \mathcal{D}_0} \max_{g \in \mathcal{S}} V(\phi, g) = \max_{g \in \mathcal{S}} V(\phi_0, g).$$

Therefore in the case p = 1 we have a minimax robust regression coefficient estimate

$$\hat{\alpha}_n = [\int_{-\pi}^{\pi} | \psi_{1,n}(x) |^2 / f_m(x) \, dx]^{-1} \int_{-\pi}^{\pi} (\psi_{1,n}(x)/f_m(x)) \, z_y(dx).$$

when the residual spectral density $g(x)$ ranges over $\mathcal{S}$. As for the interpolation problem, putting $\eta(x) = 1$ in the construction of $f_m(x)$, we can get a minimax robust interpolation response function

$$h(x) = 1 - 2\pi \{ \int_{-\pi}^{\pi} \frac{1}{f_m(x)} dx \}^{-1} / f_m(x) ,$$

when the spectral density $g(x)$ of the process concerned ranges over $\mathcal{S}$.

Also, in a particular situation, Taniguchi[29] applied the above theorem to the cases that $p > 1$ or $H(x)$ is not dominated.

Acknowledgement

The author wishes to thank Professor M. Okamoto, Osaka University, and Professor Y. Fujikoshi, Hiroshima University, for their encouragements and comments.

References

[1] AKAIKE, H. (1970). Statistical predictor identification. *Ann. Inst. Statist. Math.* 22, 203 - 217.

[2] AKAIKE, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Contr. AC*-19, 716 - 723.

[3] BERAN, R. (1977). Minimum Hellinger distance estimates for parametric models. *Ann. Statist.* 5, 445 - 463.

[4] BLOOMFIELD, P. (1973). An exponential model for the spectrum of a scalar time series. *Biometrika* 60, 217 - 226.

[5] BRILLINGER, D.R. (1969). Asymptotic properties of spectral estimates of second order. *Biometrika* 56, 375 - 390.

[6] BRILLINGER, D.R. (1975). *Time Series : Data Analysis and Theory*. Holt, Rinehart and Winston, NewYork.

[7] CLEVENSON, M.L. (1970). Asymptotically efficient estimates of the parameters of a moving average time series. Ph. D. Thesis, Stanford Univ.

[8] DAVIS, H.T. and JONES, R.H. (1968). Estimation of the innovation variance of a stationary time series. *J. Amer. Statist. Assoc.* 63, 141 - 149.

[9] DUNSMUIR, W. (1979). A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise. *Ann. Statist.* 7, 490 -506.

[10] DUNSMUIR, W. and HANNAN, E.J. (1976). Vector linear time series models. *Adv. Appl. Prob.* 8, 339 - 364.

[11] DZHAPARIDZE, K.O. (1974). A new method for estimating spectral parameters of a stationary regular time series. *Theory Prob. Appl.* 19, 122 - 132.

[12] GRENANDER, U. and ROSENBLATT, M. (1957). *Statistical Analysis of Stationary Time Series.* Wiley, NewYork.

[13] HANNAN, E.J. (1963). Regression for time series. *Time Series Analysis,* Wiley, NewYork, 17 - 37.

[14] HANNAN, E.J. (1970). *Multiple Time Series.* Wiley, NewYork.

[15] HANNAN, E.J.and NICHOLLS, D.F. (1977). The estimation of the prediction error variance. *J. Amer. Statist. Assoc.* 72, 834 - 840.

[16] HOSOYA, Y. (1974). Estimation problems on stationary time series models. Ph. D. Thesis, Yale Univ.

[17] HOSOYA, Y. (1978). Robust linear extrapolations of second-order stationary processes. *Ann. Prob.* 6, 574 - 584.

[18] HUBER, P. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* 35, 73 - 101.

[19] KHOLEVO, A.S. (1969). On estimates of regression coefficients. *Theory Prob. Appl.* 14, 79 - 104.

[20] ROZANOV, Yu.A. (1969). On a new class of statistical estimates. Soviet-Japanese Sympos., Theory of Probability, Novosibirsk, 239 - 252.

[21] ROZANOV, Yu.A. and KOZLOV, M.V. (1969). On asymptotically efficient estimation of regression coefficients. *Dokl. Akad. Nauk* SSSR 10, 1075 - 1078.

[22] SHIBATA, R. (1980). Asymptotically efficient selection of the order of the model for estimating parameters of a linear process. *Ann. Statist.* 8, 147 - 164.

[23] TANIGUCHI, M. (1978). On a generalization of a statistical spectrum analysis. *Math. Japon.* 23, 33 - 44.

[24] TANIGUCHI, M. (1979). On estimation of parameters of Gaussian stationary processes. *J. Appl. Prob.* 16, 575 - 591.

[25] TANIGUCHI, M. (1980). On estimation of the integrals of certain functions of spectral density. *J. Appl. Prob.* 17, 73 - 83.

[26] TANIGUCHI, M. (1980). Regression and interpolation for time series. *Recent Developments in Statistical Inference and Data Analysis :* K. Matusita, ed. 311 - 321. North-Holland Publishing Company.

[27] TANIGUCHI, M. (1980). An estimation procedure of parameters of a certain spectral density. *J. Roy. Statist. Soc. B.*42, to appear.

[28] TANIGUCHI, M. (1980). On selection of the order of the spectral density model for a stationary process. *Ann. Inst. Statist. Math.* 32, to appear.

[29] TANIGUCHI, M. (1980). Robust regression and interpolation for time series. Hiroshima Tech. Rep. No. 15.

[30] TANIGUCHI, M. (1980). Asymptotic properties of quasi-Gaussian maximum likelihood estimates for a linear process under nonstandard conditions. Preprint of Time Series Meeting at the Institute of Statistical Mathematics. Sept. 24 - 26, 1980.

[31] WALKER, A.M. (1964). Asymptotic properties of least-square estimates of parameters of the spectrum of a stationary non-deterministic time-series. *J. Austral. Math. Soc.* 4, 363 - 384.