HIROSHIMA MATH. J.
21 (1991), 217–251

Higher order asymptotic investigations of weighted estimators for Gaussian ARMA processes

MYINT SWE

(Received January 18, 1990)

Summary

In this paper we investigate higher order asymptotic properties of weighted estimators of Bayes type for a Gaussian ARMA process. First, for a Gaussian ARMA process with a scalar unknown parameter θ we define a quasi-weighted estimator $\hat{\theta}_{qw}$ of Bayes type based on a handy "quasi"-likelihood function. We show that if we modify $\hat{\theta}_{qw}$ to be second-order asymptotically median unbiased (AMU), then it is second-order asymptotically efficient in the class $\mathscr{A}_2$ of second-order asymptotically median unbiased estimators. We also obtain the normalizing transformation of $\hat{\theta}_{qw}$ which vanishes the second order terms of its Edgeworth expansion. Furthermore, we consider the problem of testing $H: \theta = \theta_0$ against $A: \theta \neq \theta_0$. Then higher order local powers are evaluated for a likelihood ratio, Wald and modified Wald' tests based on $\hat{\theta}_{qw}$. Secondly, we define a generalized multiparameter weighted estimator (GMWE) for a Gaussian ARMA process with a multiparameter unknown vector, and discuss its higher order asymptotic efficiency. Thirdly, we extend Akaike's final prediction error (FPE) to the case when the process concerned is a Gaussian ARMA process and evaluate Akaike's FPE up to higher order $O(n^{-2})$. It is shown that the generalized weighted estimator and the maximum likelihood estimator are best up to order n^{-2} in the sense of FPE.

1. Introduction

In the area of time series analysis, Hosoya [17] showed that the maximum likelihood estimator of a spectral parameter is second-order asymptotically efficient in the sense of Rao [27]. Akahira and Takeuchi [3] showed that an appropriately modified maximum likelihood estimator of the coefficient of an autoregressive process of order one is second-order asymptotically efficient in the sense of degree of concertration of the sampling distribution up to second-order. Furthermore, Taniguchi [35] also showed that appropriately modified maximum likelihood and quasi-maximum likelihood estimators for Gaussian ARMA processes are second-order asymptotically efficient in the sense of

Akahira and Takeuchi [3]. Ochi [25] proposed a generalized estimator in the first-order autoregression, which includes the least square estimator as a special case, and gave its third-order Edgeworth expansion. Fujikoshi and Ochi [15] investigated the third-order asymptotic properties of the maximum likelihood estimator and Ochi's generalized estimator.

For independent and identically distributed observations, Takeuchi [32] introduced a natural class $\mathscr{D}$ of estimators, and showed that the maximum likelihood estimator is third-order asymptotically efficient in $\mathscr{D}$. For a Gaussian ARMA process Taniguchi [37] elucidated various third-order asymptotic properties of the maximum likelihood estimator, and showed that it is also third-order asymptotic efficient in $\mathscr{D}$.

For i.i.d. observations, Takeuchi and Akahira [31] and Akahira and Takeuchi [3] showed that the generalized Bayes estimator for symmetric loss function is second-order asymptotically efficient in the class $\mathscr{A}_2$ of second-order asymptotically median unbiased estimators and that it is also third-order asymptotically efficient in the class $\mathscr{D}$. For dependent observations Rao [26] gave a Berry-Essen type of comparison between Bayes estimator and maximum likelihood estimator for a Markov process. For a Gaussian ARMA process with a scalar unknown parameter θ, Myint Swe and Taniguchi [24] investigated various higher order asymptotic properties of a weighted estimator of Bayes type based on the exact likelihood. They obtained the normalizing transformation of the weighted estimator. For the problem of testing $H: \theta = \theta_0$ against $A: \theta \neq \theta_0$, they compared higher order local powers of a likelihood ratio, Wald and modified Wald's test based on it.

However if the sample size n is large, the exact likelihood is intractable in practice because the likelihood function needs the inversion procedure of $n \times n$ covariance matrix. Thus in Part I, a quasi-weighted estimator $\hat{\theta}_{qw}$ of Bayes type based on a handy "quasi-likelihood function is introduced for a Gaussian process with a scalar unknown parameter. In the same way as Myint Swe and Taniguchi [24] we investigate its higher order asymptotic properties, and obtain the normalizing transformation of $\hat{\theta}_{qw}$ which vanishes the second-order terms of its Edgeworth expansion. Furthermore, we consider the problem of testing a simple hypothesis $H: \theta = \theta_0$ against the alternative $A: \theta \neq \theta_0$. Then an attempt is also made to compare higher order local powers of three tests based on $\hat{\theta}_{qw}$.

Myint Swe and Taniguchi [24] developed their discussion when the unknown parameter is scalar. In Part II we extend their results to the case where the unknown parameter is a vector. We define a generalized multiparameter weighted estimator (GMWE) for a Gaussian ARMA process with a multiparameter unknown vector and discuss its higher order efficiency.

The asymptotic mean squared error of estimated predictors is the most fundamental quantity to characterize the best statistical prediction in time series. In many cases finite order autoregressive models have been used for prediction. For the case of one step ahead prediction, Bloomfield [10] derived the asymptotic mean square error for a general mixed ARMA (p, q) model. Bhansali [9] derived the asymptotic mean square error of predicting more than one-step ahead for a general autoregressive model AR(p). Yamamoto [41] gave a manageable expression for the asymptotic mean square error of predicting more than one-step ahead from an estimated autoregressive model up to $O(n^{-1})$, where n is the sample size. Yamamoto [42] generalized the above results to the case where the process concerned is a multivariate autoregressive moving average model. The related works are Baillie [7], [8] and Reinsel [29]. Ray [28] also derived an expression for the asymptotic mean square error in predicting more than one step ahead from a p-variate autoregressive model with random coefficients.

Recently Tanaka and Maekawa [33] considered the prediction in the case of misspecifying the model as AR(1) while the true model is ARMA(1,1). Then they derived the approximate sampling distributions of the prediction error for the case (i) the data used in estimation are independent of the data used in prediction, (ii) the data used in estimation are dependent on the data used in prediction. They evaluated its bias and mean square error up to order $O(n^{-1})$ for the case (i) and (ii). Davies and Newbold [13], Kunitomo and Yamamoto [21] and Lewis and Reinsel [22] also investigated the mean square prediction error with misspecified models. Furthermore, Maekawa [23] gave the asymptotic distribution of h-step ahead prediction error in the AR(p) model up to $O(n^{-1})$ for the case (i) and (ii). He also specified the general formula for the distributions of the predictor errors based on the maximum likelihood, two types of least squared, and the Yule-Walker estimators in the AR(1) model and found that all distributions are the same up to order $O(n^{-1})$ except for the Yule-Walker predictor.

In actual situations, the order of AR model is often unknown. The difficulty is to determine the order of autoregressive model. For autoregressive model fitting, Akaike [4], [5] proposed a simple for final prediction error (FPE) to determine the order of autoregressive model. This criterion is defined as an asymptotically unbiased estimator of the mean squared error of the estimated predictor.

In part III we extend Akaike's definition of FPE to the case when the process concerned is a Gaussian ARMA process and evaluate Akaike's FPE up to order n^{-2}. Furthermore, we show that the generalized weighted estimator and the maximum likelihood estimator are best up to third-order in the sense of FPE.

Part I. Higher order asymptotic properties of quasi-weighted estimators for Gaussian ARMA processes

2. Higher order asymptotic efficiency of quasi-weighted estimators

For a Gaussian ARMA process with a scalar unknown parameter, Myint Swe and Taniguchi [24] investigated various higher order asymptotic properties of a weighted estimator of Bayes type based on the exact likelihood. However, if the sample size n is large, the exact likelihood is intractable in practice because the likelihood function needs the inversion procedure of the $n \times n$ covariance matrix. Thus we introduce a quasi-weighted estimator of Bayes type based on a handy "quasi"-likelihood function.

Let $\{X_t\}$ be a Gaussian ARMA process with spectral density $f_\theta(\lambda)$, where θ is an unknown parameter. In this section, we propose an estimator $\hat{\theta}_{qw}$ of Bayes types based on a quasi-likelihood function. Since our standpoint is different from that of original Bayes idea, we call it a quasi-weighted estimator. First, we derive a stochastic expansion of $\hat{\theta}_{qw}$. Then we show that $\hat{\theta}_{qw}$ is second-order asymptotically efficient in the class $\mathscr{A}_2$ of second-order asymptotically median unbiased estimators and that it belongs to a restricted class $\mathscr{D}$. An approach is also presented that the higher order asymptotic bias can be vanished by choosing the weight function.

We introduce $\mathscr{F}$ and $\mathscr{F}_{\text{ARMA}}$, the spaces of functions on $[-\pi, \pi]$;

$$\mathscr{F} = \{f; f(\lambda) = \textstyle\sum_{u=-\infty}^{\infty} a(u)\exp(-iu\lambda),\ a(u) = a(-u),$$

$$\textstyle\sum_{u=-\infty}^{\infty} (1 + |u|)|a(u)| < \infty\},$$

$$\mathscr{F}_{\text{ARMA}} = \{f; f(\lambda) = \sigma^2(2\pi)^{-1}|\textstyle\sum_{j=0}^{q} a_j e^{ij\lambda}|^2/|\sum_{j=0}^{p} b_j e^{ij\lambda}|^2,$$

$$(\sigma^2 > 0),\ \underline{c} \le |\textstyle\sum_{j=0}^{q} a_j z^j|^2/|\sum_{j=0}^{p} b_j z^j|^2 \le \bar{c},$$

$$\text{for } |z| \le 1,\ 0 < \underline{c} < \bar{c} < \infty\}.$$

We set down the following assumptions.

ASSUMPTION 2.1. The process $\{X_t; t = 0, \pm 1, \cdots\}$ is a Gaussian stationary process with spectral density $f_{\theta_0}(\lambda) \in \mathscr{F}_{\text{ARMA}}$, $\theta_0 \in C \subset \Theta \subset R^1$, and mean 0. Here Θ is an open set of R^1 and C is a compact subset of Θ.

ASSUMPTION 2.2. The spectral density $f_\theta(\lambda)$ is continuously five times differentiable with respect to $\theta \in \Theta$, and the derivatives $\partial f_\theta/\partial\theta$, $\partial^2 f_\theta/\partial\theta^2$, $\partial^3 f_\theta/\partial\theta^3$, $\partial^4 f_\theta/\partial\theta^4$ and $\partial^5 f_\theta/\partial\theta^5$ belong to $\mathscr{F}$.

ASSUMPTION 2.3. If $\theta \neq \theta^*$, then $f_\theta(\lambda) \neq f_\theta^*(\lambda)$ on a set of positive Lebesgue measure.

ASSUMPTION 2.4. There exists $d_1 > 0$ such that

$$I(\theta) = (4\pi)^{-1} \int_{-\pi}^{\pi} \left\{ \frac{\partial}{\partial\theta} \log f_\theta(\lambda) \right\}^2 d\lambda \geq d_1, \qquad \text{for all} \quad \theta \in C.$$

Suppose that a stretch $X_n = (X_1, \cdots, X_n)'$ of the series $\{X_t\}$ is available. In this section we use the following quasi log-likelihood function,

$$\log L_n(\theta) = -(1/2) \textstyle\sum_{j=0}^{n-1} \{\log f_\theta(\lambda_j) + I_n(\lambda_j)/f_\theta(\lambda_j)\},$$

where $I_n(\lambda_j) = (2\pi n)^{-1} |\sum_{t=1}^{n} X_t \exp(it\lambda_j)|^2$, $(\lambda_j = 2\pi j/n)$ is the periodogram. It is known that $\log L_n(\theta)$ is, to within constant terms, an approximation for the exact log-likelihood.

Let $\tilde{Z}_1(\theta) = n^{-1/2} \dfrac{\partial}{\partial\theta} \log L_n(\theta)$,

$$\tilde{Z}_2(\theta) = n^{-1/2} \left\{ \frac{\partial^2}{\partial\theta^2} \log L_n(\theta) - E_\theta\left[\frac{\partial^2}{\partial\theta^2} \log L_n(\theta) \right] \right\}, \text{ and}$$

$$\tilde{Z}_3(\theta) = n^{-1/2} \left\{ \frac{\partial^3}{\partial\theta^3} \log L_n(\theta) - E_\theta\left[\frac{\partial^3}{\partial\theta^3} \log L_n(\theta) \right] \right\}.$$

The asymptotic moments of $\tilde{Z}_1(\theta)$, $\tilde{Z}_2(\theta)$ and $\tilde{Z}_3(\theta)$ are evaluated by Taniguchi [35] as follows.

LEMMA 2.1. *Under Assumptions* 2.1–2.4 *it holds that*

$$E_\theta[\tilde{Z}_1(\theta)] = -B(\theta)/\sqrt{n} + O(n^{-3/2}),$$

$$E_\theta[\tilde{Z}_1(\theta)\tilde{Z}_2(\theta)] = J(\theta) + O(n^{-1}),$$

$$E_\theta[\tilde{Z}_1(\theta)^2] = I(\theta) + O(n^{-1}),$$

$$E_\theta[\tilde{Z}_1(\theta)^3] = K(\theta)/\sqrt{n} - 3B(\theta)I(\theta)/\sqrt{n} + O(n^{-3/2}),$$

$$n^{-1} E_\theta\left[\frac{\partial^2}{\partial\theta^2} \log L_n(\theta) \right] = -I(\theta) + O(n^{-1}).$$

$$n^{-1} E_\theta\left[\frac{\partial^3}{\partial\theta^3} \log L_n(\theta) \right] = -3J(\theta) - K(\theta) + O(n^{-1}),$$

$$Var\left[n^{-1} \frac{\partial^3}{\partial\theta^3} \log L_n(\theta) \right] = O(n^{-1}).$$

Here

$$J(\theta) = -(2\pi)^{-1}\int_{-\pi}^{\pi}\left\{\frac{\partial}{\partial\theta}f_\theta(\lambda)\right\}^3\{f_\theta(\lambda)\}^{-3}\,d\lambda$$

$$+(4\pi)^{-1}\int_{-\pi}^{\pi}\left\{\frac{\partial^2}{\partial\theta^2}f_\theta(\lambda)\right\}\left\{\frac{\partial}{\partial\theta}f_\theta(\lambda)\right\}\{f_\theta(\lambda)\}^{-2}\,d\lambda,$$

$$K(\theta) = (2\pi)^{-1}\int_{-\pi}^{\pi}\left\{\frac{\partial}{\partial\theta}f_\theta(\lambda)\right\}^3\{f_\theta(\lambda)\}^{-3}\,d\lambda$$

$$B(\theta) = (4\pi)^{-1}\int_{-\pi}^{\pi}\left\{\frac{\partial}{\partial\theta}f_\theta(\lambda)\right\}b_\theta(\lambda)\{f_\theta(\lambda)\}^{-2}d\lambda,$$

where

$$b(\theta) = (2\pi)^{-1}\textstyle\sum_{j=-\infty}^{\infty}|j|\gamma(j)e^{ij\lambda} \text{ and } \gamma(j) = E_\theta(X_tX_{t+j}).$$

Occasionally, we shall use the simpler notation $\tilde{Z}_1$, $\tilde{Z}_2$, I, J, K, etc. instead of $\tilde{Z}_1(\theta)$, $\tilde{Z}_2(\theta)$, $I(\theta)$, $J(\theta)$, $K(\theta)$, etc., respectively, when there is no danger of confusion.

We now define a quasi-weighted estimator and investigate its higher order asymptotic efficiency. Let $\xi(\theta)$ be a non-negative weight function of θ. Suppose that $\xi(\theta)$ is continuously two times differentiable with respect to $\theta\in\Theta$. The quasi-weighted estimator of θ is defined by

$$\hat{\theta}_{qw} = \frac{\int_\Theta \theta\, L_n(\theta)\xi(\theta)d\theta}{\int_\Theta L_n(\theta)\xi(\theta)d\theta}.$$

Putting $t = \sqrt{n}(\theta-\theta_0)$, we obtain

$$\sqrt{n}(\hat{\theta}_{qw}-\theta_0) = \frac{\int t\, L_n(\theta_0+t/\sqrt{n})\xi(\theta_0+t/\sqrt{n})\,dt}{\int L_n(\theta_0+t/\sqrt{n})\xi(\theta_0+t/\sqrt{n})\,dt}$$

(2.1)

$$= \frac{\int t\exp[\log L_n(\theta_0+t/\sqrt{n})]\xi(\theta_0+t/\sqrt{n})dt}{\int \exp[\log L_n(\theta_0+t/\sqrt{n})]\xi(\theta_0+t/\sqrt{n})dt}.$$

Expanding $\log L_n(\theta_0 + t/\sqrt{n})$ in a Taylor series at $\theta = \theta_0$, we obtain

$$\log L_n(\theta_0 + t/\sqrt{n}) = \log L_n(\theta_0) + t\frac{\partial}{\partial\theta}\log L_n(\theta_0)/\sqrt{n}$$
$$+ t^2\frac{\partial^2}{\partial\theta^2}\log L_n(\theta_0)/(2n) + t^3 n^{-3/2}\frac{\partial^3}{\partial\theta^3}\log L_n(\theta_0)/6$$
$$+ o_p(n^{-1/2})$$
$$(2.2) \qquad = \log L_n(\theta_0) + t\,\tilde{Z}_1(\theta_0) + t^2\tilde{Z}_2(\theta_0)/(2\sqrt{n})$$
$$+ t^2 E\left[\frac{\partial^2}{\partial\theta^2}\log L_n(\theta_0)\right]/(2n) + t^3 n^{-3/2} E\left[\frac{\partial^3}{\partial\theta^3}\log L_n(\theta_0)\right]/6$$
$$+ o_p(n^{-1/2}).$$

From (2.2) and Lemma 2.1 we have

$$\log L_n(\theta_0 + t/\sqrt{n}) = \log L_n(\theta_0) + t\,\tilde{Z}_1(\theta_0) + t^2\tilde{Z}_2(\theta_0)/(2\sqrt{n})$$
$$(2.3) \qquad - t^2 I(\theta_0)/2 + t^3[-3J(\theta_0) - K(\theta_0)]/(6\sqrt{n})$$
$$+ o_p(n^{-1/2}).$$

Next, expanding $\xi(\theta_0 + t/\sqrt{n})$ in a Taylor series at $\theta = \theta_0$ we obtain

$$t\exp[\log L_n(\theta_0 + t/\sqrt{n})]\xi(\theta_0 + t/\sqrt{n}) = L_n(\theta_0)\exp\{\tilde{Z}_1^2/(2I)\}$$
$$\cdot\exp\{-(I/2)(t - \tilde{Z}_1/I)^2\}[t\xi(\theta_0) + t^3\tilde{Z}_2\xi(\theta_0)/(2\sqrt{n})$$
$$+ t^4\xi(\theta_0)(-3J - K)/(6\sqrt{n}) + t^2\xi^{(1)}(\theta_0)/\sqrt{n} + o_p(n^{-1/2})].$$

By using the moments of normal distribution with mean $\tilde{Z}_1/I$ and variance I^{-1}, we can evaluate

$$(2\pi I^{-1})^{-1/2}\int t\exp[\log L_n(\theta_0 + t/\sqrt{n})]\,\xi(\theta_0 + t/\sqrt{n})\,dt$$
$$(2.4) \qquad = L_n(\theta_0)\exp\{\tilde{Z}_1^2/(2I)\}\,\{A + B/\sqrt{n} + o_p(n^{-1/2})\},$$

where $A = (\tilde{Z}_1\xi)/I$ and

$$B = (\tilde{Z}_2\xi/2)(3I\tilde{Z}_1 + \tilde{Z}_1^3)I^{-3} + [\xi(-3J - K)/6](3I^2 + 6I\tilde{Z}_1^2 + \tilde{Z}_1^4)I^{-4}$$
$$+ \xi^{(1)}(I + \tilde{Z}_1^2)I^{-2}.$$

Similarly it is shown that

$$(2.5)\quad \begin{aligned}(2\pi I^{-1})^{-1/2}\int \exp[\log L_n(\theta_0 + t/\sqrt{n})]\,\xi(\theta_0 + t/\sqrt{n})\,dt \\ = L_n(\theta_0)\exp\{\tilde{Z}_1^2/(2I)\}\,\{\tilde{A} + \tilde{B}/\sqrt{n} + o_p(n^{-1/2})\},\end{aligned}$$

where $\tilde{A} = \xi$ and

$$\begin{aligned}\tilde{B} = (\tilde{Z}_2\xi/2)(I + \tilde{Z}_1^2)I^{-2} + [\xi(-3J - K)/6](3I\tilde{Z}_1 + \tilde{Z}_1^3)I^{-3} \\ + \xi^{(1)}\tilde{Z}_1/I.\end{aligned}$$

The relations (2.1), (2.4) and (2.5) yield

LEMMA 2.2. *Under Assumptions* 2.1–2.4, *the quasi-weighted estimator can be expanded as*

$$\begin{aligned}\sqrt{n}(\hat{\theta}_{qw} - \theta_0) = \tilde{Z}_1/I + \{\tilde{Z}_1\tilde{Z}_2 I^{-2} + (-3J - K)\tilde{Z}_1^2 I^{-3}/2\}/\sqrt{n} \\ + \{(-3J - K)I^{-2}/2 + \xi^{(1)}/(\xi I)\}/\sqrt{n} + o_p(n^{-1/2}).\end{aligned}$$

We shall investigate the higher order asymptotic efficiency of $\hat{\theta}_{qw}$. Let $\mathscr{S}$ and $\mathscr{D}$ be the classes of estimators:

$$\begin{aligned}\mathscr{S} = \{\hat{\theta}_n;\ \sqrt{n}(\hat{\theta}_n - \theta_0) = Z_1/I + Q/\sqrt{n} + o_p(n^{-1/2}), \\ Q = O_p(1),\ E_{\theta_0}(Q) = \mu\}, \\ \mathscr{D} = \{\hat{\theta}_n;\ \hat{\theta}_n \in \mathscr{S},\ E_{\theta_0}(Z_1\tilde{Q}^2) = o(1),\ \tilde{Q} = Q - \mu\},\end{aligned}$$

where $Z_1 = n^{-1/2}(\partial/\partial\theta)\log$ (exact likelihood). Taniguchi [36] showed that the maximum likelihood estimator $\hat{\theta}_{ML}$ and quasi-maximum likelihood estimator $\hat{\theta}_{qML}$ of θ belong to $\mathscr{D}$. It follows from Lemma 2.2 that

$$\sqrt{n}(\hat{\theta}_{qw} - \theta_0) - \sqrt{n}(\hat{\theta}_{qML} - \theta_0) = (\text{constant})/\sqrt{n} + o_p(n^{-1/2}).$$

This implies that $\hat{\theta}_{qw}$ also belongs to $\mathscr{D}$. Thus we establish the following theorem.

THEOREM 2.1. *If we modify the quasi-weighted estimator $\hat{\theta}_{qw}$ to be second-order asymptotically median unbiased* (AMU), *then it is second-order asymptotically efficient in the class $\mathscr{A}_2$ of second-order* AMU *estimators. Also $\hat{\theta}_{qw}$ belongs to the restricted class $\mathscr{D}$.*

From Lemma 2.2, the cumulants of $U_n = \sqrt{nI}(\hat{\theta}_{qw} - \theta_0)$ can be evaluated as follows:

(2.6) $$E(U_n) = \{(-2J - K)I^{-3/2} + (\xi^{(1)}/\xi - B)I^{-1/2}\}/\sqrt{n} + o(n^{-1/2}).$$
$$= C_1/\sqrt{n} + o(n^{-1/2})\,(\text{say}),$$

(2.7) $$Cum\{U_n, U_n\} = 1 + o(n^{-1/2}),$$

(2.8) $$Cum\{U_n, U_n, U_n\} = \{-(3J + 2K)\}I^{-3/2}/\sqrt{n} + o(n^{-1/2}),$$
$$= C_3/\sqrt{n} + o(n^{-1/2})\ (\text{say}),$$
$$Cum^{(J)}\{U_n, \cdots, U_n\} = O(n^{-J/2+1}) \quad \text{for} \quad J \geq 3.$$

Applying the Edgeworth expansion formula (see Taniguchi [37]) to U_n, we get

PROPOSITION 2.1. *Under Assumptions* 2.1–2.4,

$$P^n_{\theta_0}[\sqrt{nI}(\hat{\theta}_{qw} - \theta_0) \leq y] = \Phi(y) - \varphi(y)[C_1/\sqrt{n} + C_3(y^2 - 1)/(6\sqrt{n})] + o(n^{-1/2}),$$

where $\varphi(y) = (2\pi)^{-1/2}\exp(-y^2/2)$ *and* $\Phi(y) = \int_{-\infty}^{y} \varphi(x)\,dx$.

When we modify an estimator to be second-order asymptotically unbiased or second-order AMU, we usually use the adjustment factor. Here we adjust the asymptotic bias by choosing the weight function. As we shall see later we can do this successfully if we know the type of parameter (e.g., θ is an AR part root). In many cases, maximum likelihood estimation requires iterative computational procedure. On the other hand, $\hat{\theta}_{qw}$ has a closed form, and has no need for bias-correction factor if we choose an appropriate weight function. This is the reason why we do not call $\hat{\theta}_{qw}$ the Bayes estimator.

Since the evaluation of $B(\theta)$ for general rational spectral such as

$$f_\theta(\lambda) = \sigma^2(2\pi)^{-1}[\textstyle\prod_{k=1}^{q}(1 - \psi_k e^{i\lambda})(1 - \psi_k e^{-i\lambda})]/[\prod_{k=1}^{p}(1 - \rho_k e^{i\lambda})(1 - \rho_k e^{-i\lambda})]$$

is very complicated, in this part we consider the following ARMA(1,1) spectral density

(2.9) $$f_\theta(\lambda) = \sigma^2(2\pi)^{-1}[(1 - \psi e^{i\lambda})(1 - \psi e^{-i\lambda})]/[(1 - \rho e^{i\lambda})(1 - \rho e^{-i\lambda})],$$

where ψ and ρ are real numbers such that $|\rho| < 1$ and $|\psi| < 1$. For the spectral density of (2.9) the following are evaluated by Taniguchi [35] explicitly.

$$I(\sigma^2) = \sigma^{-4}/2,\ I(\psi) = 1/(1 - \psi^2),\quad I(\rho) = 1/(1 - \rho^2),$$
$$J(\sigma^2) = -\sigma^{-6},\ J(\psi) = 4\psi/(1 - \psi^2)^2,\quad J(\rho) = -2\rho/(1 - \rho^2)^2,$$
$$K(\sigma^2) = \sigma^{-6},\ K(\psi) = -6\psi/(1 - \psi^2)^2,\quad K(\rho) = 6\rho/(1 - \rho^2)^2,$$

$$B(\sigma^2) = -\sigma^{-2}(\rho - \psi)^2(1 - \rho^2)^{-1}(1 - \psi^2)^{-1},$$

$$B(\psi) = \frac{(\psi - \rho)(1 + \psi^2 - 2\psi\rho - \rho^2 + 3\psi^2\rho^2 - 2\psi^3\rho)}{(1 - \psi^2)^2(1 - \psi\rho)(1 - \rho^2)},$$

$$B(\rho) = \frac{(\rho - \psi)(1 - 2\psi\rho + \psi^2)}{(1 - \rho^2)(1 - \rho\psi)(1 - \psi^2)}.$$

Particularly for AR(1) model, $B(\rho) = \rho/(1 - \rho^2)$ and for MA(1) model, $B(\psi) = \psi(1 + \psi^2)/(1 - \psi^2)^2$.

In (2.6), if we choose $\xi(\theta)$ so that $C_1 = 0$ [i.e., $\xi^{(1)}/\xi = (2J + K)/I + B$], then the quasi-weighted estimator becomes second-order asymptotically unbiased. We call the weight function $\xi(\theta)$ chosen by this manner the second-order asymptotically unbiased weight function. For the spectral density of (2.9) the second-order asymptotically unbiased weight functions are given by

$$\xi(\sigma^2) = \sigma^{[-4-2(\rho-\psi)^2(1-\rho^2)^{-1}(1-\psi^2)^{-1}]} \text{ on } (0, \infty) \text{ for } \theta_0 = \sigma^2,$$

$$\xi(\rho) = (1 - \rho^2)^{-3/2} \text{ on } (-1, 1) \text{ for } \theta_0 = \rho \text{ and } \psi = 0,$$

$$\xi(\psi) = (1 - \psi^2)^{-1/2} \exp\{(1 - \psi^2)^{-1}\} \text{ on } (-1, 1) \text{ for } \theta_0 = \psi \text{ and } \rho = 0.$$

In Proposition 2.1, if we choose $\xi(\theta)$ so that $C_1 = C_3/6$ [i.e., $\xi^{(1)}/\xi = (9J + 4K)/(6I) + B$], then the quasi-weighted estimator becomes second-order AMU. We call the weight function $\xi(\theta)$ chosen by this manner the second-order AMU weight function. For the spectral density of (2.9) the second-order AMU weight functions are given by

$$\xi(\sigma^2) = \sigma^{[(-10/3)-2(\rho-\psi)^2(1-\rho^2)^{-1}(1-\psi^2)^{-1}]} \text{ on } (0, \infty) \text{ for } \theta_0 = \sigma^2,$$

$$\xi(\rho) = (1 - \rho^2)^{-1} \text{ on } (-1, 1) \text{ for } \theta_0 = \rho \text{ and } \psi = 0,$$

$$\xi(\psi) = (1 - \psi^2)^{-1/2} \exp\{(1 - \psi^2)^{-1}\} \text{ on } (-1, 1) \text{ for } \theta_0 = \psi \text{ and } \rho = 0.$$

3. Normalizing transformations

For i.i.d. case, Konishi [20] considered a normalizing transformation of statistics based upon the elements of the sample covariance matrix which extinguishes the second-order terms of the Edgeworth expansion. Then he showed that Fisher's Z-transformation gives the normalizing transformation of the correlation coefficient. In the area of time series analysis Taniguchi, Krishnaiah and Chao [40] considered the normalizing transformations of the maximum likelihood estimator and quasi-maximum likelihood estimator for Gaussian ARMA processes. They also showed that Fisher's Z-transformation gives the normalizing transformation for parameters of AR part. In this section we shall seek the normalizing transformation of the estimator $\hat{\theta}_{qw}$, and show that the weight function plays a role of bias adjustment.

For a smooth function $g(\cdot)$ we consider the standardized transformation $\sqrt{nI}\{g(\hat{\theta}_{qw}) - g(\theta_0)\}/g^{(1)}(\theta_0)$ of $\hat{\theta}_{qw}$. Deriving the Edgeworth expansion of standardized transformation, we seek the normalizing transformation which vanishes the second-order terms of the Edgeworth expansion. Suppose that $g(\theta)$ is three times continuously differentiable. We shall derive, in the same way as Taniguchi, Krishnaiah and Chao [40], the Edgeworth expansion of

$$V_n = \sqrt{nI(\theta_0)}\{g(\hat{\theta}_{qw}) - g(\theta_0)\}/g^{(1)}(\theta_0).$$

By using Taylor's expansion of $g(\theta)$ at θ_0, we obtain

$$\begin{aligned}(3.1)\qquad V_n = &\sqrt{nI(\theta_0)}\{\hat{\theta}_{qw} - \theta_0\} + (1/2)\sqrt{nI(\theta_0)}\{\hat{\theta}_{qw} - \theta_0\}^2 g^{(2)}/g^{(1)}(\theta_0)\\ &+ o_p(n^{-1/2}).\end{aligned}$$

Further, from Lemma 2.2 we can write V_n as

$$\begin{aligned}(3.2)\qquad V_n = &P\tilde{Z}_1 + n^{-1/2}\{Q\tilde{Z}_1\tilde{Z}_2 + R\tilde{Z}_1^2 + S + A\tilde{Z}_1^2 g^{(2)}(\theta_0)/g^{(1)}(\theta_0)\}\\ &+ o_p(n^{-1/2}),\end{aligned}$$

where $P = I^{-1/2}$, $Q = I^{-3/2}$, $R = (-3J - K)I^{-5/2}/2$, $A = I^{-3/2}/2$ and $S = (-3J - K)I^{-3/2}/2 + (\xi^{(1)}/\xi)I^{-1/2}$. The asymptotic cumulants of V_n can be evaluated as follows:

$$\begin{aligned}E(V_n) &= \{-B(\theta_0)I^{-1/2} + (-2J - K)I^{-3/2} + (\xi^{(1)}/\xi)I^{-1/2}\\ &\qquad + [g^{(2)}(\theta_0)/g^{(1)}(\theta_0)]I^{-1/2}/2\}/\sqrt{n} + o(n^{-1/2})\\ &= d_1/\sqrt{n} + o(n^{-1/2}), \text{ (say)},\\ Cum\{V_n, V_n\} &= 1 + o(n^{-1/2}),\\ Cum\{V_n, V_n, V_n\} &= \{(-3J - 2K)I^{-3/2} + 3[g^{(2)}(\theta_0)/g^{(1)}(\theta_0)]I^{-1/2}\}/\sqrt{n}\\ &\qquad + o(n^{-1/2}),\\ &= d_3/\sqrt{n} + o(n^{-1/2}), \text{ (say)},\\ Cum\{V_n, \cdots, V_n\} &= O(n^{-J/2+1}), \qquad \text{for } J \geq 3.\end{aligned}$$

Hence we obtain the Edgeworth expansion of V_n in the form

$$(3.3)\qquad P^n_{\theta_0}[V_n \leq x] = \Phi(x) - \varphi(x)\{d_1/\sqrt{n} + d_3(x^2 - 1)/\sqrt{n}\} + o(n^{-1/2}).$$

If we set $d_1 = d_3 = 0$, the second-order terms of (3.3) vanish. This implies the following theorem.

THEOREM 3.1. *If the transformation* $g(\cdot)$ *and the weight function satisfy*

$$g^{(2)}(\theta_0)/g^{(1)}(\theta_0) = (3J + 2K)/(3I) \quad \text{and} \tag{3.4}$$

$$\xi^{(1)}(\theta_0)/\xi(\theta_0) = (9J + 4K)/(6I) + B(\theta_0), \tag{3.5}$$

then $P^n_{\theta_0}\{\sqrt{nI(\theta_0)}[g(\hat{\theta}_{qw}) - g(\theta_0)]/g^{(1)}(\theta_0) \leq x\} = \Phi(x) + o(n^{-1/2})$.

Henceforth the transformation $g(\cdot)$ and the weight function $\xi(\theta)$ satisfying (3.4) and (3.5) are called the normalizing transformation and normalizing weight function, respectively.

EXAMPLE 3.1. The normalizing transformations and the normalizing weight functions for the sepctral density of (2.9) are given as follows:

Normalizing transformations:

(i) If $\theta_0 = \sigma^2$, then $g(\sigma^2) = 3(\sigma^2)^{1/3}$.
(ii) If $\theta_0 = \psi$, then $g(\psi) = \psi$.
(iii) If $\theta_0 = \rho$, then $g(\rho) = \log\{(1 + \rho)/(1 - \rho)\}/2$.

Normalizng weight functions:

(i) If $\theta_0 = \sigma^2$, then $\xi(\sigma^2) = \sigma^{[(-10/3) - 2(\rho - \psi)^2(1 - \rho^2)^{-1}(1 - \psi^2)^{-1}]}$.
(ii) If $\theta_0 = \psi$ and $\rho = 0$, then $\xi(\psi) = (1 - \psi^2)^{-1/2}\exp\{(1 - \psi^2)^{-1}\}$.
(iii) If $\theta_0 = \rho$ and $\psi = 0$, then $\xi(\rho) = (1 - \rho^2)^{-1}$.

REMARK 3.1. In our results, Fisher's $\boldsymbol{Z}$-transformation and the transformation $3(\sigma^2)^{1/3}$ give the normalizing transformations for the AR part parameter ρ and the innovation variance, respectively. For i.i.d. case Konishi [20] showed that Fisher's $\boldsymbol{Z}$-transformation and the transformation $(\cdot)^{1/3}$ give the normalizing transformations for the correlation coefficient and the latent roots of the sample covariance matrix, respectively. It may be noted that the AR part parameter and the correlation coefficients represent a sort of correlation structure, and that the innovation variance and latent roots of the covariance matrix represent a sort of variance structure although our statistical models are essentially different from Konishi's one.

REMARK 3.2. When we consider the normalizing transformation of the maximum likelihood estimator and the quasi-maximum likelihood estimator, we need an adjustment factor C [i.e., $\sqrt{nI}\{g(\hat{\theta}_{ML}) - g(\theta_0) - C/n\}/g^{(1)}(\theta_0)$]. For the normalizing transformation of quasi-weighted estimator $\hat{\theta}_{qw}$, the adjustment factor is not necessary because the normalizing weight function plays a role of the adjustment factor.

4. Higher order investigations for testing problems based on quasi-weighted estimator

For a Gaussian ARMA process with spectral parameter θ, Taniguchi [39] considered the problem of testing a simple hypothesis $H: \theta = \theta_0$ against the alternative $A: \theta \neq \theta_0$, and introduced a class of tests $\mathscr{S}$, which contains the likelihood ratio(LR), Wald(W), modified Wald(MW) and Rao(R), tests. Then he derived the χ^2 type asymptotic expansion of the distribution of a test $T \in \mathscr{S}$ under the sequence of alternatives $A_n: \theta = \theta_0 + \varepsilon/\sqrt{n}$, $\varepsilon > 0$, up to order $n^{-1/2}$, where n is the sample size. He also compared the local powers of these tests on the basis of their asymptotic expansions and showed that there is no uniformly superior test for the local alternatives. Also Myint Swe and Taniguchi [24] developed a similar discussion by use of a weighted estimator of Bayes type based on the exact likelihood.

In this section we shall discuss testing problems based on the quasi-weighted estimator. Consdier the problem of testing a simple hypothesis $H: \theta = \theta_0$ against the alternative $A: \theta \neq \theta_0$ for a Gaussian process with spectral parameter θ. We propose a class of tests $\mathscr{S}_A$, which contains weighted likelihood ratio test (WLR), weighted Wald's test(WW), weighted modified Wald's test (WMW) based on the quasi-weighted estimator. Then we derive the asymptotic expansion of local power $P_{\theta_0+\varepsilon/\sqrt{n}}(T \leq x)$ of $T \in \mathscr{S}_A$ up to order $n^{-1/2}$ and consider the power comparison of WLR, WW and WMW. Then it is shown that none of the above tests is uniformly superior. However if we modify them to be asymptotically unbiased we can show that their local powers are identical.

Consider the transformations given by

$$\tilde{U}_1(\theta) = \tilde{Z}_1(\theta)/\sqrt{I(\theta)}, \tag{4.1}$$

$$\tilde{U}_2(\theta) = [\tilde{Z}_2(\theta) - J(\theta)I(\theta)^{-1}\tilde{Z}_1(\theta)]/\gamma_\theta I(\theta), \tag{4.2}$$

where $\gamma_\theta = [M(\theta)I(\theta) - J(\theta)^2]^{1/2}/I(\theta)^{3/2}$ and

$$M(\theta) = \pi^{-1}\int_{-\pi}^{\pi}\left\{\frac{\partial}{\partial\theta}f_\theta(\lambda)\right\}^4 f_\theta(\lambda)^{-4}d\lambda + (4\pi)^{-1}\int_{-\pi}^{\pi}\left\{\frac{\partial^2}{\partial\theta^2}f_\theta(\lambda)\right\}^2 f_\theta(\lambda)^{-2}d\lambda$$
$$- \pi^{-1}\int_{-\pi}^{\pi}\left\{\frac{\partial}{\partial\theta}f_\theta(\lambda)\right\}^2\left\{\frac{\partial^2}{\partial\theta^2}f_\theta(\lambda)\right\}\{f_\theta(\lambda)\}^{-3}d\lambda.$$

For simplicity, we use $\tilde{U}_1$, $\tilde{U}_2$, $\tilde{Z}_1$, $\tilde{Z}_2$, I, J, K, γ instead of $\tilde{U}_1(\theta)$, $\tilde{U}_2(\theta)$, $\tilde{Z}_1(\theta)$, $\tilde{Z}_2(\theta)$, $I(\theta)$, $J(\theta)$, $K(\theta)$, $\gamma(\theta)$ respectively, if they are evaluated at $\theta = \theta_0 + \varepsilon/\sqrt{n}$. Define the following class of tests;

$$\mathscr{S}_A = \{S;\, S = \{\tilde{U}_1 + I(\theta_0)^{1/2}\varepsilon\}^2 + n^{-1/2}[c_1\tilde{U}_1^3 + c_2\tilde{U}_1^2\tilde{U}_2 + d_1\tilde{U}_1 + \{c_3\tilde{U}_1^2 + c_4\tilde{U}_1\tilde{U}_2 + d_1\sqrt{I}\}\varepsilon + \{c_5\tilde{U}_1 + c_6\tilde{U}_2\}\varepsilon^2 + c_7\varepsilon^3] + o_p(n^{-1/2}), \text{ under } A_n, \text{ where } c_7 = I^{3/2}c_1 - Ic_3 + I^{1/2}c_5\}.$$

This class $\mathscr{S}_A$ is also very natural. It contains the ones defined by Taniguchi [39] and also the three tests defined by the following (i), (ii) and (iii).

(i) Weighted likelihood ratio test $\mathrm{WLR} = 2[l_n(\hat{\theta}_{qw}) - l_n(\theta_0)]$, where $l_n(\theta) = \log L_n(\theta)$. Let $V = \sqrt{n}(\hat{\theta}_{qw} - \theta)$. By expanding WLR in a Taylor series at $\theta = \hat{\theta}_{qw}$, we obtain

$$\mathrm{WLR} = -2(\theta_0 - \hat{\theta}_{qw})\left[\frac{\partial}{\partial\theta}l_n(\theta) + (\hat{\theta}_{qw} - \theta)\frac{\partial^2}{\partial\theta^2}l_n(\theta) + (\hat{\theta}_{qw} - \theta)^2\frac{\partial^3}{\partial\theta^3}l_n(\theta)/2\right] - (\theta_0 - \hat{\theta}_{qw})^2\left[\frac{\partial^2}{\partial\theta^2}l_n(\theta) + (\hat{\theta}_{qw} - \theta)\frac{\partial^3}{\partial\theta^3}l_n(\theta)\right] + (\hat{\theta}_{qw} - \theta_0)^3\frac{\partial^3}{\partial\theta^3}l_n(\theta)/3 + o_p(n^{-1/2}).$$

If we modify $(\hat{\theta}_{qw} - \theta_0) = (\hat{\theta}_{qw} - \theta + \theta - \theta_0)$, then we obtain

$$\mathrm{WLR} = \frac{\partial}{\partial\theta}l_n(\theta)\cdot 2V/\sqrt{n} + \frac{\partial}{\partial\theta}l_n(\theta)\cdot 2\varepsilon/\sqrt{n} + \frac{\partial^2}{\partial\theta^2}l_n(\theta)(V^2 - \varepsilon^2)/n + (1/3)n^{-3/2}\frac{\partial^3}{\partial\theta^3}l_n(\theta)(V^3 + \varepsilon^3) + o_p(n^{-1/2}). \tag{4.3}$$

Substituting

$$V = \tilde{Z}_1/I + n^{-1/2}\{\tilde{Z}_1\tilde{Z}_2 I^{-2} - (3J + K)\tilde{Z}_1^2 I^{-3}/2 - (3J + K)I^{-2}/2 + \xi^{(1)}/(\xi I)\} + o_p(n^{-1/2}),$$

$$n^{-1}\frac{\partial^2}{\partial\theta^2}l_n(\theta) = -I + \tilde{Z}_2/\sqrt{n} + O(n^{-1}),$$

$$n^{-1}\frac{\partial^3}{\partial\theta^3}l_n(\theta) = -3J - K + \tilde{Z}_3/\sqrt{n} + O(n^{-1}),$$

into (4.3), we have

$$\mathrm{WLR} = \{\tilde{U}_1 + \sqrt{I(\theta_0)}\cdot\varepsilon\}^2 + n^{-1/2}\{(-KI^{-3/2}/3)\tilde{U}_1^3 + \gamma\tilde{U}_1^2\tilde{U}_2 + [(J+K)I^{-1/2}\tilde{U}_1 - \gamma I\tilde{U}_2]\varepsilon^2 + [(3J+2K)/3]\varepsilon^3\} + o_p(n^{-1/2}),$$

this implies that WLR belongs to $\mathscr{S}_A$.

Similarly the following results (ii)–(iii) are obtained:

(ii) Weighted Wald's test $\mathrm{WW} = n(\hat{\theta}_{qw} - \theta_0)^2 I(\hat{\theta}_{qw})$ belongs to $\mathscr{S}_A$ with coefficients

$$c_1 = JI^{-3/2},\ c_2 = 2\gamma,\ d_1 = 2(\xi^{(1)}/\xi)I^{-1/2} - (3J+K)I^{-3/2},$$
$$c_3 = (3J+K)/I,\ c_4 = 2\gamma\sqrt{I},\ c_5 = (4J+2K)I^{-1/2},\ c_6 = 0,$$
$$c_7 = (2J+K).$$

(iii) Weighted modified Wald's test $\mathrm{WMW} = n(\hat{\theta}_{qw} - \theta_0)^2 I(\theta_0)$ belongs to $\mathscr{S}_A$ with coefficients

$$c_1 = -(J+K)I^{-3/2},\ c_2 = 2\gamma,\ d_1 = 2(\xi^{(1)}/\xi)I^{-1/2} - (3J+K)I^{-3/2},$$
$$c_3 = -(3J+2K)/I,\ c_4 = 2\gamma\sqrt{I},\ c_5 = -(2J+K)I^{-1/2},\ c_6 = 0.$$
$$c_7 = 0.$$

The cumulants of transformations $\tilde{U}_1$ and $\tilde{U}_2$ can be evaluated as follows:

$$E_\theta(\tilde{U}_1) = n^{-1/2}(-BI^{-1/2}) + O(n^{-3/2}),$$
$$E_\theta(\tilde{U}_2) = n^{-1/2}(JB)/(\gamma I^2) + O(n^{-3/2}),$$
$$Cum(\tilde{U}_1, \tilde{U}_1) = 1 + O(n^{-1}),$$
$$Cum(\tilde{U}_1, \tilde{U}_2) = O(n^{-1}),$$
$$Cum(\tilde{U}_2, \tilde{U}_2) = 1 + O(n^{-1}),$$
$$Cum(\tilde{U}_1, \tilde{U}_1, \tilde{U}_1) = n^{-1/2}(KI^{-3/2}) + O(n^{-3/2}),$$
$$Cum(\tilde{U}_1, \tilde{U}_1, \tilde{U}_2) = n^{-1/2}c_{112}^{(1)} + O(n^{-3/2}),$$
$$Cum(\tilde{U}_1, \tilde{U}_2, \tilde{U}_2) = n^{-1/2}c_{122}^{(1)} + O(n^{-3/2}),$$
$$Cum(\tilde{U}_2, \tilde{U}_2, \tilde{U}_2) = n^{-1/2}c_{222}^{(1)} + O(n^{-3/2}),$$

where $c_{112}^{(1)}$, $c_{122}^{(1)}$ and $c_{222}^{(1)}$ can be expressed by the spectral density (see Taniguchi [35] or [37] for the expressions of I, J, K).

In order to derive an asymptotic expansion of the distribution of $S \in \mathscr{S}_A$ under A_n, by using Edgeworth expansion formula (Taniguchi [37]) we establish the following lemma.

LEMMA 4.1. *Under Assumptions* 2.1–2.4,

$$P^n_{\theta_0}[\tilde{U}_1 < y_1, \tilde{U}_2 < y_2] = \int_{-\infty}^{y_1} \int_{-\infty}^{y_2} f(u_1, u_2) du_1 du_2 + o(n^{-1/2}),$$

where

$$f(u_1, u_2) = \varphi(u_1)\varphi(u_2)[1 + n^{-1/2}(-BI^{-1/2})u_1 + n^{-1/2}\{(JB)/(\gamma I^2)\}u_2$$
$$+ \{1/(6\sqrt{n})\}\{K(\theta_0)I^{-3/2}(u_1^3 - 3u_1) + 3c^{(1)}_{112}(u_1^2 u_2 - u_2)$$
$$+ 3c^{(1)}_{122}(u_1 u_2^2 - u_1) + c^{(1)}_{222}(u_2^3 - 3u_2)\}]$$

and $\phi(u) = (2\pi)^{-1/2}\exp(-u^2/2)$.

By using Lemma 4.1 the characteristic function of $C_S(t)$ of $S \in \mathscr{S}_A$ can be evaluated under A_n. In fact

$$C_S(t) = E_{\theta_0 + \varepsilon/\sqrt{n}}\{e^{itS}\}$$
$$= \iint e^{itS} f(u_1, u_2) du_1 du_2$$
$$= \iint \exp[it\{u_1 + I(\theta_0)^{1/2}\varepsilon\}^2]\cdot[1 + n^{-1/2}(it)\{c_1 u_1^3$$
$$+ c_2 u_1^2 u_2 + d_1 u_1 + (c_3 u_1^2 + c_4 u_1 u_2 + d_1\sqrt{I})\varepsilon$$
$$+ (c_5 u_1 + c_6 u_2)\varepsilon^2 + c_7\varepsilon^3\}]\,\phi(u_1)\,\phi(u_2)\,[1$$
$$+ n^{-1/2}(-BI^{-1/2})u_1 + n^{-1/2}\{(JB)/(\gamma I^2)\}u_2 \tag{4.4}$$
$$+ (n^{-1/2}/6)\{KI^{-3/2}(u_1^3 - 3u_1) + 3c^{(1)}_{112}(u_1^2 u_2 - u_2)$$
$$+ 3c^{(1)}_{122}(u_1 u_2^2 - u_1) + c^{(1)}_{222}(u_2^3 - 3u_2)\}]du_1 du_2$$
$$+ o(n^{-1/2}).$$

Integration of (4.4) with respect to u_2 yields

$$C_S(t) = \exp\left\{\frac{itI(\theta_0)\varepsilon^2}{1 - 2it}\right\}(1 - 2it)^{-1/2}\int (2\pi)^{-1/2}$$
$$\cdot(1 - 2it)^{1/2}\exp\left[-\frac{(1 - 2it)}{2}\left\{u_1 - \frac{2\varepsilon it I(\theta_0)^{1/2}}{1 - 2it}\right\}^2\right]$$

$$\cdot [1 + n^{-1/2}(it)\{c_1 u_1^3 + d_1 u_1 + c_3 u_1^2 \varepsilon + \sqrt{I} d_1 \varepsilon + c_5 u_1 \varepsilon^2 + c_7 \varepsilon^3\} + n^{-1/2}\{-BI^{-1/2}u_1 + KI(\theta_0)^{-3/2}(u_1^3 - 3u_1)/6\}]du_1 + o(n^{-1/2}).$$

By calculating the above integral we get the asymptotic expansion of $C_S(t)$ under A_n.

LEMMA 4.2. *Under Assumptions* 2.1–2.4, *the characteristic function* $C_S(t)$ *of* $S \in \mathscr{S}_A$ *under* $\theta = \theta_0 + \varepsilon/\sqrt{n}$ *has an asymptotic expansion*;

$$C_S(t) = \exp\left\{\frac{itI(\theta_0)\varepsilon^2}{1 - 2it}\right\}(1 - 2it)^{-1/2} \cdot [1 + n^{-1/2}\textstyle\sum_{j=0}^{3} B_j^{(S)}(1 - 2it)^{-j}] + o(n^{-1/2}),$$

where

$$B_0^{(S)} = (1/6)[\{-9I(\theta_0)^{3/2}c_1 + 6I(\theta_0)c_3 - 3I(\theta_0)^{1/2}c_5 - K(\theta_0)\}\varepsilon^3 + \{9I(\theta_0)^{1/2}c_1 - 3c_3 + 3K(\theta_0)/I(\theta_0) - 3d_1\sqrt{I(\theta_0)} + 6B(\theta_0)\}\varepsilon],$$

$$B_1^{(S)} = (1/2)[\{6I(\theta_0)^{3/2}c_1 - 3I(\theta_0)c_3 + I(\theta_0)^{1/2}c_5 + K(\theta_0)\}\varepsilon^3 + \{c_3 - 6I(\theta_0)^{1/2}c_1 - 2K(\theta_0)/I(\theta_0) + d_1\sqrt{I(\theta_0)} - 2B(\theta_0)\}\varepsilon],$$

$$B_2^{(S)} = (1/2)[\{I(\theta_0)c_3 - 4I(\theta_0)^{3/2}c_1 - K(\theta_0)\}\varepsilon^3 + \{3I(\theta_0)^{1/2}c_1 + K(\theta_0)/I(\theta_0)\}\varepsilon],$$

$$B_3^{(S)} = (1/6)[3I(\theta_0)^{3/2}c_1 + K(\theta_0)]\varepsilon^3.$$

This lemma implies

THEOREM 4.1. *Under Assumptions* 2.1–2.4, *the distribution function of* $S \in \mathscr{S}_A$ *for* $\theta = \theta_0 + \varepsilon/\sqrt{n}$ *has the following expansion*

$$P^n_{\theta_0 + \varepsilon/\sqrt{n}}[S \le x] = P[\chi_1^2(\delta) \le x] + n^{-1/2}\textstyle\sum_{j=0}^{3} B_j^{(S)} P[\chi_{1+2j}^2(\delta) \le x] + o(n^{-1/2}),$$

where $\delta^2 = I(\theta_0)\varepsilon^2/2$, *and* $\chi_j^2(\delta)$ *is a noncentral* χ^2 *random variable with* j *degrees of freedom and noncentrality parameter* δ^2.

For the three tests WLR, WW and WMW, we can give more explicit expressions for the coefficients $B_j^{(S)}$ in Theorem 4.1.

EXAMPLE 4.1.

(i) $S = \text{WLR}$ (Weighted likelihood ratio test)

$$B_0^{(\text{WLR})} = -\{3J(\theta_0) + K(\theta_0)\}\varepsilon^3/6 + B(\theta_0)\varepsilon,$$

$$B_1^{(\text{WLR})} = J(\theta_0)\varepsilon^3/2 - B(\theta_0)\varepsilon,\ B_2^{(\text{WLR})} = K(\theta_0)\varepsilon^3/6,\ B_3^{(\text{WLR})} = 0,$$

(ii) $S = \text{WW}$ (Weighted Wald's test)

$$\begin{aligned} B_0^{(\text{WW})} &= (1/6)[-\{3J(\theta_0) + K(\theta_0)\}\varepsilon^3 \\ &\quad - \{6\xi^{(1)}/\xi - 3(3J(\theta_0) + K(\theta_0))/I(\theta_0)\}\varepsilon] + B(\theta_0)\varepsilon, \\ B_1^{(\text{WW})} &= (1/2)[J(\theta_0)\varepsilon^3 - \{3J(\theta_0) + K(\theta_0)\}\varepsilon/I(\theta_0) \\ &\quad + \{2\xi^{(1)}/\xi - (3J(\theta_0) + K(\theta_0))/I(\theta_0)\}\varepsilon] - B(\theta_0)\varepsilon, \\ B_2^{(\text{WW})} &= (1/2)[-J(\theta_0)\varepsilon^3 + \{3J(\theta_0) + K(\theta_0)\}\varepsilon/I(\theta_0)], \\ B_3^{(\text{WW})} &= (1/6)[3J(\theta_0) + K(\theta_0)]\varepsilon^3, \end{aligned}$$

(iii) $S = \text{WMW}$ (Weighted modified Wald's test)

$$\begin{aligned} B_0^{(\text{WMW})} &= (1/6)[\{-3J(\theta_0) - K(\theta_0)\}\varepsilon^3 \\ &\quad - \{6\xi^{(1)}/\xi - 3(3J(\theta_0) + K(\theta_0))/I(\theta_0)\}\varepsilon] + B(\theta_0)\varepsilon, \\ B_1^{(\text{WMW})} &= (1/2)[J(\theta_0)\varepsilon^3 + \{3J(\theta_0) + 2K(\theta_0)\}\varepsilon/I(\theta_0) \\ &\quad + \{2\xi^{(1)}/\xi - (3J(\theta_0) + K(\theta_0))/I(\theta_0)\}\varepsilon - B(\theta_0)\varepsilon, \\ B_2^{(\text{WMW})} &= (1/2)[\{J(\theta_0) + K(\theta_0)\}\varepsilon^3 - \{3J(\theta_0) + 2K(\theta_0)\}\varepsilon/I(\theta_0)], \\ B_3^{(\text{WMW})} &= -(1/6)[3J(\theta_0) + 2K(\theta_0)]\varepsilon^3. \end{aligned}$$

In view of Theorem 4.1 we can investigate the local power properties in the class $\mathscr{S}_A$. By Theorem 4.1 and Example 4.1, it can be easily shown that for $S \in \mathscr{S}_A$,

$$\begin{aligned} &P^n_{\theta_0+\varepsilon/\sqrt{n}}[S > x] - P^n_{\theta_0+\varepsilon/\sqrt{n}}[\text{WLR} > x] \\ &= n^{-1/2}[(1/2)\{P(\chi_7^2(\delta) > x) - P(\chi_5^2(\delta) > x)\}Q_3^{(S)}(\theta_0) \\ &\quad + (1/2)\{P(\chi_5^2(\delta) > x) - P(\chi_3^2(\delta) > x)\}Q_2^{(S)}(\theta_0) \\ &\quad + (1/2)\{P(\chi_3^2(\delta) > x) - P(\chi_1^2(\delta) > x)\}Q_1^{(S)}(\theta_0)] \\ &\quad + o(n^{-1/2}) \end{aligned} \tag{4.5}$$

where $$Q_1^{(S)}(\theta_0) = \{3I(\theta_0)^{3/2}c_1 - 3I(\theta_0)c_3 + I(\theta_0)^{1/2}c_5 - J(\theta_0)\}\varepsilon^3 + \{c_3 - 3I(\theta_0)^{1/2}c_1 - K(\theta_0)/I(\theta_0) + d_1\sqrt{I(\theta_0)}\}\varepsilon,$$

$$Q_2^{(S)}(\theta_0) = \{I(\theta_0)c_3 - 3I(\theta_0)^{3/2}c_1 - K(\theta_0)\}\varepsilon^3 + \{3I(\theta_0)^{1/2}c_1 + K(\theta_0)/I(\theta_0)\}\varepsilon,$$

$$Q_3^{(S)}(\theta_0) = (1/3)\{3I(\theta_0)^{3/2}c_1 + K(\theta_0)\}\varepsilon^3.$$

By using (4.5) and the well known relation

$$P[\chi_{j+2}^2(\delta) > x] - P[\chi_j^2(\delta) > x] = 2p_{j+2}(x;\delta),$$

where $p_j(x;\delta)$ is the probability density function of $\chi_j^2(\delta)$, we establish the following theorem.

THEOREM 4.2. *Under Assumptions* 2.1–2.4,

$$P^n_{\theta_0+\varepsilon/\sqrt{n}}[S > x] - P^n_{\theta_0+\varepsilon/\sqrt{n}}[\mathrm{WLR} > x] = n^{-1/2}[Q_3^{(S)}(\theta_0)p_7(x;\delta) + Q_2^{(S)}(\theta_0)p_5(x;\delta) + Q_1^{(S)}(\theta_0)p_3(x;\delta)] + o(n^{-1/2}), \quad \text{for } S\in\mathscr{S}_A.$$

By using Theorem 4.2 for the spectral of (2.9), we can compare the local power properties among the three tests WLR, WW and WMW. The following local power comparisons are performed for the second-order asymptotically unbiased weight function, i.e., $\xi^{(1)}(\theta_0)/\xi(\theta_0) = \{2J(\theta_0) + K(\theta_0)\}/I(\theta_0) + B(\theta_0)$.

EXAMPLE 4.2. (WW versus WLR under A_n)

$$P^n_{\theta_0+\varepsilon/\sqrt{n}}[\mathrm{WW} > x] - P^n_{\theta_0+\varepsilon/\sqrt{n}}[\mathrm{WLR} > x] = n^{-1/2}(3J + K)[(\varepsilon^3/3)p_7(x;\delta) + (\varepsilon/I)p_5(x;\delta)] + n^{-1/2}[(J + K)/I + 2B]\varepsilon p_3(x;\delta).$$

(1) If $\theta_0 = \sigma^2$, then $3J + K = -2\sigma^{-6} < 0$, $J + K = 0$ and $2B = -2\sigma^{-2}(\rho - \psi)^2(1 - \rho^2)^{-1}(1 - \psi^2)^{-1} < 0$. which implies that WLR is more powerful than WW.

(2) If $\theta_0 = \rho$ and $\psi = 0$, then $3J + K = 0$ and $(J + K)/I + 2B = 6\rho(1 - \rho^2)^{-1}$, which implies that WW is more powerful than WLR if $\rho > 0$ and vice versa.

(3) If $\theta_0 = \psi$ and $\rho = 0$, then $3J + K = 6\psi(1 - \psi^2)^{-2}$ and $(J + K)/I + 2B = 4\psi^3(1 - \psi^2)^{-2}$, which implies that WW is more powerful than WLR if $\psi > 0$ and vice versa.

EXAMPLE 4.3. (WMW versus WLR under A_n)

$$P^n_{\theta_0+\varepsilon/\sqrt{n}}[\mathrm{WMW}] > x] - P^n_{\theta_0+\varepsilon/\sqrt{n}}[\mathrm{WLR} > x]$$
$$= n^{-1/2}(-3J - 2K)[(\varepsilon^3/3)p_7(x;\delta) + (\varepsilon/I)p_5(x;\delta)]$$
$$+ n^{-1/2}[(J+K)/I + 2B]\varepsilon p_3(x;\delta) + o(n^{-1/2}).$$

(1) If $\theta_0 = \sigma^2$, then $(-3J - 2K) = \sigma^{-6} > 0$, $J + K = 0$ and $2B = -2\sigma^{-2}(\rho - \psi)^2(1-\rho^2)^{-1}(1-\psi^2)^{-1} < 0$ which implies that we can not determine which test is more powerful than the other test.

(2) If $\theta_0 = \rho$ and $\psi = 0$, then $(-3J - 2K) = -6\rho(1-\rho^2)^{-2}$ and $(J+K)/I + 2B = 6\rho(1-\rho^2)^{-1}$, which implies that we can not determine which test is more powerful than the other test.

(3) If $\theta_0 = \psi$ and $\rho = 0$, then $(-3J - 2K) = 0$ and $(J+K)/I + 2B = 4\psi^3(1-\psi^2)^{-2}$, which implies that WMW is more powerful than WLR if $\psi > 0$ and vice versa.

These examples show that none of WLR, WW and WMW tests is uniformly superior.

Finally we show that an appropriate modification of $S \in \mathscr{S}_A$ leads to a unified result. First, we note that the coefficients c_1, c_3 and c_5 in the stochastic expansion of the three tests automatically satisfy

(4.6)
$$Ic_3 - 3I^{3/2}c_1 = K,$$
$$I^{1/2}c_5 - Ic_3 = J + K.$$

Henceforth we confine ourselves to a class of tests

$$\mathscr{S}'_A = \{S;\, S \in \mathscr{S}_A \text{ and } c_1,\ c_3 \text{ and } c_5 \text{ satisfy (4.6)}\}.$$

Furthermore, we impose the second-order asymptotic unbiasedness;

(4.7)
$$\frac{\partial}{\partial \varepsilon} P^n_{\theta_0+\varepsilon/\sqrt{n}}[S > x]|_{\varepsilon=0} = o(n^{-1/2}) \text{ for } S \in \mathscr{S}'_A.$$

By Lemma 4.2 and Theorem 4.1, we can see that (4.7) is equivalent to

(4.8)
$$9I^{1/2}c_1 - 3c_3 + 3K/I - 3d_1\sqrt{I} + 6B = 0,$$
$$c_3 - 6I^{1/2}c_1 - 2K/I + d_1\sqrt{I} - 2B = 0,$$
$$3I^{1/2}c_1 + K/I = 0.$$

Consider a class of tests

$$U\mathscr{S}'_A = \{S;\ S \in \mathscr{S}'_A \text{ and satisfies (4.7) and (4.8)}\}.$$

From Example 4.1, it is easy to see $\mathrm{WLR} \in U\mathscr{S}'_A$. From Theorem 4.2, it is not difficult to show

$$P^n_{\theta_0+\varepsilon/\sqrt{n}}[S > x] - P^n_{\theta_0+\varepsilon/\sqrt{n}}[\mathrm{WLR} > x] = o(n^{-1/2})$$

for all $S \in U\mathscr{S}'_A$.

Now we modify $S \in \mathscr{S}'_A$ to be second-order asymptotically unbiased. Put $S^* = m(\hat{\theta}_{qw})S$, where $S \in \mathscr{S}'_A$ and $m(\theta)$ is a continuously twice differentiable function with $m(\theta_0) = 1$. Then we can show that

$$\begin{aligned} S^* &= \{m(\theta) + (\hat{\theta}_{qw} - \theta)m^{(1)}(\theta)\}S + o_p(n^{-1/2}) \\ &= \{1 + n^{-1/2}m^{(1)}(\theta_0)\varepsilon + (\hat{\theta}_{qw} - \theta)m^{(1)}(\theta_0)\}S + o_p(n^{-1/2}) \\ &= \{\tilde{U}_1 + \varepsilon\sqrt{I}\}^2 + n^{-1/2}[(c_1 + h/\sqrt{I})\tilde{U}_1^3 + c_2\tilde{U}_1^2\tilde{U}_2 + d_1\sqrt{I} \\ &\quad + \{(c_3 + 3h)\tilde{U}_1^2 + c_4\tilde{U}_1\tilde{U}_2 + d_1\sqrt{I}\}\varepsilon \\ &\quad + \{(c_5 + 3h\sqrt{I})\tilde{U}_1 + c_6\tilde{U}_2\}\varepsilon^2 + \{c_7 + Ih\}\varepsilon^3] + o_p(n^{-1/2}), \end{aligned}$$

where $h = m^{(1)}(\theta_0)$. Then the second-order asymptotic unbiased condition of S^* leads to

$$\begin{aligned} &9\sqrt{I}(c_1 + h/\sqrt{I}) - 3(c_3 + 3h) + 3K/I - 3d_1\sqrt{I} + 6B = 0, \\ &(c_3 + 3h) - 6\sqrt{I}(c_1 + h/\sqrt{I}) - 2K/I + d_1\sqrt{I} - 2B = 0, \\ &3\sqrt{I}(c_1 + h/\sqrt{I}) + K/I = 0. \end{aligned}$$

Thus we can see that $S^* \in U\mathscr{S}'_A$ if $d_1\sqrt{I} - 2B = 0$, $c_1 + h/\sqrt{I} = -(K/3)I^{-3/2}$ and $c_3 + 3h = 0$. Summarizing the above we have the following unified result.

THEOREM 4.3. *Under Assumptions 2.1–2.4, the local powers of all modified tests* $S^* = m(\hat{\theta}_{qw})S$, $S \in \mathscr{S}'_A$, *with* $m(\theta_0) = 1$, $m^{(1)}(\theta_0) = -\sqrt{I}c_1 - K/(3I)$ *(or* $= -c_3/3$ *equivalently) and* $d_1\sqrt{I} - 2B = 0$ *are identical up to order* $n^{-1/2}$.

This modification procedure can be realized easily. For the AR(1) model we can see $\mathrm{WMW} = n(\hat{\rho}_{qw} - \rho)^2(1 - \rho^2)^{-1}$, and $m^{(1)}(\rho) = 2\rho(1 - \rho^2)^{-1}$. Thus the modified test of WMW is given by $\{1 + (\hat{\rho}_{qw} - \rho)2\rho(1 - \rho^2)^{-1}\}\mathrm{WMW}$ with $\xi(\rho) = (1 - \rho^2)^{-1/2}$.

Part II. Third order asymptotic efficiency of generalized multiparameter weighted estimators

5. Generalized multiparameter weighted estimator (GMWE)

In this section we extend the results of Myint Swe and Taniguchi [24] to the case where the unknown parameter is a vector. That is, we define a generalized multiparameter weighted estimator (GMWE) for a Gaussian ARMA process and discuss its higher order asymptotic efficiency.

Let $\{X_t; t = 0, \pm 1, \pm 2, \cdots\}$ be a Gaussian ARMA process with spectral density $f_{\underset{\sim}{\theta}}(\lambda)$ which depends on an unknown multiparameter $\underset{\sim}{\theta} \in \Theta \subset R^p$. We consider stretch $X_n = (X_1, \cdots, X_n)'$ of the series $\{X_t\}$. Let Σ_n be the covariance matrix of X_n. The likelihood function based on X_n is given by

$$L_n(\underset{\sim}{\theta}) = (2\pi)^{-n/2} |\Sigma_n|^{-1/2} \exp\{ -(1/2) X_n' \Sigma_n^{-1} X_n\}.$$

Now we introduce D_Δ and D_{ARMA}, the spaces of functions on $[-\pi, \pi]$;

$$D_\Delta = \{f; f(\lambda) = \textstyle\sum_{u=-\infty}^{\infty} a(u) \exp(-iu\lambda),\ a(u) = a(-u),$$
$$\textstyle\sum_{u=-\infty}^{\infty} |u|^2 |a(u)| < \infty\},$$

$$D_{\text{ARMA}} = \{f; f(\lambda) = \sigma^2 (2\pi)^{-1} |\textstyle\sum_{j=0}^{q} \alpha_j e^{ij\lambda}|^2 / |\textstyle\sum_{j=0}^{p} \beta_j e^{ij\lambda}|^2,$$

$(\sigma^2 > 0)$, for some positive intergers p and q, where $A(z) = \sum_{j=0}^{q} \alpha_j z^j$ and $B(z) = \sum_{j=0}^{p} \beta_j z^j$ are bounded away from zero for $|z| \le 1\}$.

We set down the following assumptions.

ASSUMPTION 5.1. The process $\{X_t; t = 0, \pm 1, \pm 2, \cdots\}$ is a Gaussian stationary process with the spectral density $f_{\underset{\sim}{\theta}}(\lambda) \in D_{\text{ARMA}}$, $\underset{\sim}{\theta} = (\theta_1, \cdots, \theta_p)' \in \Theta \subset R^p$, and mean 0.

ASSUMPTION 5.2. The spectral density $f_{\underset{\sim}{\theta}}(\lambda)$ is continuously five times differentiable with respect to $\underset{\sim}{\theta}$, and the derivatives $\partial f_{\underset{\sim}{\theta}} / \partial \theta_j$, $\partial^2 f_{\underset{\sim}{\theta}} / \partial \theta_j \partial \theta_k, \cdots, \partial^5 f_{\underset{\sim}{\theta}} / \partial \theta_j \partial \theta_k \partial \theta_l \partial \theta_m \partial \theta_n$ $(j, k, l, m, n = 1, \cdots, p)$ belong to D_Δ.

ASSUMPTION 5.3. If $\underset{\sim}{\theta} \neq \underset{\sim}{\theta}^*$, then $f_{\underset{\sim}{\theta}} \neq f_{\underset{\sim}{\theta}^*}$ on a set of positive Lebesgue measure.

ASSUMPTION 5.4. The matrix

$$I(\underset{\sim}{\theta}) = (4\pi)^{-1} \int_{-\pi}^{\pi} \left\{ \frac{\partial}{\partial \underset{\sim}{\theta}} \log f_{\underset{\sim}{\theta}}(\lambda) \right\} \left\{ \frac{\partial}{\partial \underset{\sim}{\theta}}, \log f_{\underset{\sim}{\theta}}(\lambda) \right\} d\lambda$$

is positive definite for all $\underset{\sim}{\theta} \in \Theta$.

Let

$$(5.1) \qquad Z_\alpha = n^{-1/2}\frac{\partial}{\partial\theta_\alpha}\log L_n(\underset{\sim}{\theta})$$

$$(5.2) \qquad Z_{\alpha\beta} = n^{-1/2}\left\{\frac{\partial^2}{\partial\theta_\alpha\partial\theta_\beta}\log L_n(\underset{\sim}{\theta}) - E\left[\frac{\partial^2}{\partial\theta_\alpha\partial\theta_\beta}\log L_n(\underset{\sim}{\theta})\right]\right\}$$

$$(5.3) \qquad Z_{\alpha\beta\gamma} = n^{-1/2}\left\{\frac{\partial^3}{\partial\theta_\alpha\partial\theta_\beta\partial\theta_\gamma}\log L_n(\underset{\sim}{\theta}) - E\left[\frac{\partial^3}{\partial\theta_\alpha\partial\theta_\beta\partial\theta_\gamma}\log L_n(\underset{\sim}{\theta})\right]\right\}.$$

It is known that their asymptotic moments are given by

$$E_{\underset{\sim}{\theta}}[Z_\alpha Z_\beta] = I_{\alpha\beta} + O(n^{-1}),$$

$$E_{\underset{\sim}{\theta}}[Z_\alpha Z_{\beta\gamma}] = J_{\alpha\beta\gamma} + O(n^{-1}),$$

$$E_{\underset{\sim}{\theta}}[Z_\alpha Z_\beta Z_\gamma] = n^{-1/2}K_{\alpha\beta\gamma} + O(n^{-3/2}),$$

$$n^{-1}E_{\underset{\sim}{\theta}}\left[\frac{\partial^3}{\partial\theta_\alpha\partial\theta_\beta\partial\theta_\gamma}\log L_n(\underset{\sim}{\theta})\right] = -J_{\alpha\beta\gamma} - J_{\beta\gamma\alpha} - J_{\gamma\alpha\beta} - K_{\alpha\beta\gamma} + O(n^{-1}),$$

$$\mathrm{Var}_{\underset{\sim}{\theta}}\left[n^{-1}\frac{\partial^2}{\partial\theta_\alpha\partial\theta_\beta}\log L_n(\underset{\sim}{\theta})\right] = O(n^{-1}),$$

$$\mathrm{Var}_{\underset{\sim}{\theta}}\left[n^{-1}\frac{\partial^3}{\partial\theta_\alpha\partial\theta_\beta\partial\theta_\gamma}\log L_n(\underset{\sim}{\theta})\right] = O(n^{-1}),$$

where $I_{\alpha\beta}$, $J_{\alpha\beta\gamma}$ and $K_{\alpha\beta\gamma}$ are expressed in terms of the spectral density (see Taniguchi [37]).

Let $\underset{\sim}{\theta}_0 = (\theta_{01}, \theta_{02}, \cdots, \theta_{0p})'$ be the true parameter of $\underset{\sim}{\theta}$ and $\xi(\underset{\sim}{\theta})$ be a weight function on Θ. Define

$$p_n(\underset{\sim}{\theta}\,|\,\boldsymbol{x}_n) = \frac{L_n(\underset{\sim}{\theta})\xi(\underset{\sim}{\theta})}{\int_\Theta L_n(\underset{\sim}{\theta})\xi(\underset{\sim}{\theta})d\underset{\sim}{\theta}} \quad \text{and}$$

$$\gamma_n(\boldsymbol{d}\,|\,\boldsymbol{x}_n) = \int_\Theta L[\sqrt{n}(\boldsymbol{d} - \underset{\sim}{\theta})]\, p_n(\underset{\sim}{\theta}\,|\,\boldsymbol{x}_n)d\underset{\sim}{\theta},$$

where L is a loss function. An estimator $\hat{\underset{\sim}{\theta}}$ is called a generalized multiparameter weighted estimator (GMWE) with respect to a loss function L and weight function ξ if

$$(5.4) \qquad \gamma_n(\hat{\underset{\sim}{\theta}}\,|\,\boldsymbol{x}_n) = \inf_{\boldsymbol{d}\in\Theta}\gamma_n(\boldsymbol{d}\,|\,\boldsymbol{x}_n).$$

Here the estimator $\hat{\underset{\sim}{\theta}}$ is exactly a generalized Bayes estimator. However we call it GMWE because our standpoint is different from that of original Bayes idea.

6. Third-order stochastic expansion of GMWE

In this section we shall give a stochastic expansion of the generalized multiparameter weighted estimator defined by (5.4).

First, in $O(n^{-1/2})$ neighbourhood of $\underset{\sim}{\theta}_0$, we have

$$
\begin{aligned}
&p_n(\underset{\sim}{\theta}|\boldsymbol{x}_n)/p_n(\underset{\sim}{\theta}_0|\boldsymbol{x}_n)\\
&\quad = \exp[\log p_n(\underset{\sim}{\theta}|\boldsymbol{x}_n) - \log p_n(\underset{\sim}{\theta}_0|\boldsymbol{x}_n)]\\
&\quad = \exp[\log L_n(\underset{\sim}{\theta}) - \log L_n(\underset{\sim}{\theta}_0) + \log \xi(\underset{\sim}{\theta}) - \log \xi(\underset{\sim}{\theta}_0)]\\
&\quad = \exp[\textstyle\sum_{\alpha=1}^{p} \dfrac{\partial}{\partial\theta_\alpha} \log L_n(\underset{\sim}{\theta}_0)(\theta_\alpha - \theta_{0\alpha})\\
&\qquad + \dfrac{1}{2}\textstyle\sum_{\alpha,\beta=1}^{p} \dfrac{\partial^2}{\partial\theta_\alpha\partial\theta_\beta} \log L_n(\underset{\sim}{\theta}_0)(\theta_\alpha - \theta_{0\alpha})(\theta_\beta - \theta_{0\beta})\\
&\qquad + \dfrac{1}{6}\textstyle\sum_{\alpha,\beta,\gamma=1}^{p} \dfrac{\partial^3}{\partial\theta_\alpha\partial\theta_\beta\partial\theta_\gamma} \log L_n(\underset{\sim}{\theta}^*)(\theta_\alpha - \theta_{0\alpha})(\theta_\beta - \theta_{0\beta})(\theta_\gamma - \theta_{0\gamma})\\
&\qquad + \textstyle\sum_{\alpha=1}^{p} \dfrac{\xi_\alpha^{(1)}(\underset{\sim}{\theta}_0)}{\xi(\underset{\sim}{\theta}_0)}(\theta_\alpha - \theta_{0\alpha}) + o_p(n^{-1/2})],
\end{aligned}
$$

where $\xi_\alpha^{(1)}(\underset{\sim}{\theta}_0) = \dfrac{\partial}{\partial\theta_\alpha}\xi(\underset{\sim}{\theta}_0)$ $(\alpha = 1, \cdots, p)$ and $\underset{\sim}{\theta} < \underset{\sim}{\theta}^* < \underset{\sim}{\theta}_0$ or $\underset{\sim}{\theta} > \underset{\sim}{\theta}^* > \underset{\sim}{\theta}_0$.

Letting $t_\alpha = \sqrt{n}(\theta_\alpha - \theta_{0\alpha})$ $(\alpha = 1, \cdots, p)$ and $\rho_{\alpha\beta\gamma} = J_{\alpha\beta\gamma} + J_{\beta\gamma\alpha} + J_{\gamma\alpha\beta} + K_{\alpha\beta\gamma}$ we obtain

$$
\begin{aligned}
p_n(\underset{\sim}{\theta}|\boldsymbol{x}_n)/p_n(\underset{\sim}{\theta}_0|\boldsymbol{x}_n) = \exp[&\textstyle\sum_{\alpha=1}^{p} Z_\alpha(\underset{\sim}{\theta}_0)t_\alpha\\
&+ (1/2)\textstyle\sum_{\alpha,\beta=1}^{p}\{n^{-1/2}Z_{\alpha\beta}(\underset{\sim}{\theta}_0) - I_{\alpha\beta}(\underset{\sim}{\theta}_0)\}t_\alpha t_\beta\\
&- (6\sqrt{n})^{-1}\textstyle\sum_{\alpha,\beta,\gamma=1}^{p}\rho_{\alpha\beta\gamma}(\underset{\sim}{\theta}_0)t_\alpha t_\beta t_\gamma\\
&+ n^{-1/2}\{\xi(\underset{\sim}{\theta}_0)\}^{-1}\textstyle\sum_{\alpha=1}^{p}\xi_\alpha^{(1)}(\underset{\sim}{\theta}_0)t_\alpha + o_p(n^{-1/2})].
\end{aligned}
$$

Let $I^{\alpha\beta}$ be the (α, β) element of the inverse matrix of the information matrix $\boldsymbol{I}$. By letting $U_\alpha = \sum_{\beta=1}^{p} I^{\alpha\beta}Z_\beta$ we can modify $p_n(\underset{\sim}{\theta}|\boldsymbol{x}_n)/p_n(\underset{\sim}{\theta}_0|\boldsymbol{x}_n)$ as follws:

$$
\begin{aligned}
(6.1)\qquad p_n(\underset{\sim}{\theta}|\boldsymbol{x}_n)/p_n(\underset{\sim}{\theta}_0|\boldsymbol{x}_n) &= \exp\{(1/2)\textstyle\sum_{\alpha,\beta=1}^{p} I_{\alpha\beta}(\underset{\sim}{\theta}_0)U_\alpha U_\beta\}\\
&\quad\cdot[\exp\{-(1/2)\textstyle\sum_{\alpha,\beta=1}^{p} I_{\alpha\beta}(\underset{\sim}{\theta}_0)(t_\alpha - U_\alpha)(t_\beta - U_\beta)\}]
\end{aligned}
$$

$$\cdot\{1 + n^{-1/2}\xi(\theta_0)^{-1}\textstyle\sum_{\alpha=1}^{p}\xi_\alpha^{(1)}(\theta_0)t_\alpha + (2\sqrt{n})^{-1}\sum_{\alpha,\beta=1}^{p}Z_{\alpha\beta}(\theta_0)t_\alpha t_\beta$$

$$- (6\sqrt{n})^{-1}\textstyle\sum_{\alpha,\beta,\gamma=1}^{p}\rho_{\alpha\beta\gamma}(\theta_0)t_\alpha t_\beta t_\gamma + o_p(n^{-1/2})\}$$

$$= \exp\{(1/2)\textstyle\sum_{\alpha,\beta=1}^{p}I_{\alpha\beta}(\theta_0)U_\alpha U_\beta\}q_n(\boldsymbol{t};\theta_0|\boldsymbol{x}_n) \text{ (say)}, \tag{6.1}$$

where $\boldsymbol{t}=(t_1, t_2, \cdots, t_p)'$. Therefore, letting $\hat{\boldsymbol{t}}=\sqrt{n}(\hat{\theta}-\theta_0)$ and recalling $\gamma_n(\hat{\theta}\,|\boldsymbol{x}_n)=\int_\Theta L(\sqrt{n}(\hat{\theta}-\theta)p_n(\theta|\boldsymbol{x}_n)d\theta$, we can rewrite $\gamma_n(\hat{\theta}\,|\boldsymbol{x}_n)$ as

$$\gamma_n(\hat{\boldsymbol{t}}|\boldsymbol{x}_n) = n^{-1/2}p_n(\theta_0|\boldsymbol{x}_n)\exp\{(1/2)\textstyle\sum_{\alpha,\beta=1}^{p}I_{\alpha\beta}(\theta_0)U_\alpha U_\beta\} \cdot \displaystyle\int L(\hat{\boldsymbol{t}}-\boldsymbol{t})q_n(\boldsymbol{t};\theta_0|\boldsymbol{x}_n)d\boldsymbol{t}. \tag{6.2}$$

Furthermore, we assume the following:

ASSUMPTION 6.1. $L(\boldsymbol{u})$ is convex and symmetric about the origin.

ASSUMPTION 6.2. For each $\alpha=1,\cdots,p$, $\int L(-\boldsymbol{u})q_n(\boldsymbol{u}+\boldsymbol{t};\theta_0|\boldsymbol{x}_n)\,d\boldsymbol{u}$ is continuously partially differentiable with respect to t_α under integral sign.

By (6.2) and Assumption 6.1 it is shown that the generalized multiparameter weighted estimator $\hat{\boldsymbol{t}}$ with respect to $L(\cdot)$ and $\xi(\cdot)$ is given as a solution of the equation

$$\frac{\partial}{\partial u_\alpha}\int L(\boldsymbol{u}-\boldsymbol{t})q_n(\boldsymbol{t};\theta_0|\boldsymbol{x}_n)d\boldsymbol{t}=0 \quad (\alpha=1,\cdots,p).$$

By Assumption 6.2 we have

$$\frac{\partial}{\partial u_\alpha}\int L(\boldsymbol{u}-\boldsymbol{t})q_n(\boldsymbol{t};\theta_0|\boldsymbol{x}_n)d\boldsymbol{t}=\int L(-\boldsymbol{u})\left\{\frac{\partial}{\partial t_\alpha}q_n(\boldsymbol{t}+\boldsymbol{u};\theta_0|\boldsymbol{x}_n)\right\}d\boldsymbol{u} \quad (\alpha=1,\cdots,p),$$

which implies that the GMWE $\hat{\boldsymbol{t}}$ is obtained by a solution of the equation

$$\int L(-\boldsymbol{u})\left\{\frac{\partial}{\partial t_\alpha}q_n(\boldsymbol{t}+\boldsymbol{u};\theta_0|\boldsymbol{x}_n)\right\}d\boldsymbol{u}=0 \quad (\alpha=1,\cdots,p). \tag{6.3}$$

Substitution of (6.1) into (6.3) gives

$$0=\int L(-\boldsymbol{u})\exp[-(1/2)\textstyle\sum_{\alpha,\beta=1}^{p}I_{\alpha\beta}(\theta_0)(\hat{t}_\alpha+u_\alpha-U_\alpha)(\hat{t}_\beta+u_\beta-U_\beta)]$$

$$[-\textstyle\sum_{\beta=1}^{p}I_{\alpha\beta}(\theta_0)(\hat{t}_\beta+u_\beta-U_\beta)$$

$$\cdot\{1+n^{-1/2}\xi(\theta_0)^{-1}\textstyle\sum_{\alpha=1}^{p}\xi_\alpha^{(1)}(\theta_0)(\hat{t}_\alpha+u_\alpha)$$

$$+(2\sqrt{n})^{-1}\textstyle\sum_{\alpha,\beta=1}^{p}Z_{\alpha\beta}(\theta_0)(\hat{t}_\alpha+u_\alpha)(\hat{t}_\beta+u_\beta)$$

$$- (6\sqrt{n})^{-1}\textstyle\sum_{\alpha,\beta,\gamma=1}^{p} \rho_{\alpha\beta\gamma}(\underset{\sim}{\theta}_0)(\hat{t}_\alpha + u_\alpha)(\hat{t}_\beta + u_\beta)(\hat{t}_\gamma + u_\gamma)\}$$

$$+ (\sqrt{n}\xi(\underset{\sim}{\theta}_0)^{-1}\xi_\alpha^{(1)}(\underset{\sim}{\theta}_0) + n^{-1/2}\textstyle\sum_{\beta=1}^{p} Z_{\alpha\beta}(\underset{\sim}{\theta}_0)(\hat{t}_\beta + u_\beta)$$

$$- (2\sqrt{n})^{-1}\textstyle\sum_{\beta,\gamma=1}^{p} \rho_{\alpha\beta\gamma}(\underset{\sim}{\theta}_0)(\hat{t}_\beta + u_\beta)(\hat{t}_\gamma + u_\gamma)]\, d\boldsymbol{u}$$

$$+ o_p(n^{-1/2}).$$

We define

$$M = \int L(-\boldsymbol{u})\exp[-(1/2)\textstyle\sum_{\alpha,\beta=1}^{p} I_{\alpha\beta}(\underset{\sim}{\theta}_0)u_\alpha u_\beta]\, d\boldsymbol{u},$$

$$P_{\alpha\beta} = \int L(-\boldsymbol{u})u_\alpha u_\beta \exp[-(1/2)\textstyle\sum_{\alpha,\beta=1}^{p} I_{\alpha\beta}(\underset{\sim}{\theta}_0)u_\alpha u_\beta]\, d\boldsymbol{u}, \quad \boldsymbol{P} = (P_{\alpha\beta}),$$

$$(\alpha, \beta = 1, \cdots, p),$$

$$Q_{\alpha\beta\gamma\delta} = \int L(-\boldsymbol{u})u_\alpha u_\beta u_\gamma u_\delta \exp[-(1/2)\textstyle\sum_{\alpha,\beta=1}^{p} I_{\alpha\beta}(\underset{\sim}{\theta}_0)u_\alpha u_\beta)]\, d\boldsymbol{u},$$

$$(\alpha, \beta, \gamma, \delta = 1, \cdots, p).$$

From Assumption 6.1 it holds that for $\alpha, \beta, \gamma = 1, \cdots, p$,

$$\int L(-\boldsymbol{u})u_\alpha \exp[-(1/2)\textstyle\sum_{\alpha,\beta=1}^{p} I_{\alpha\beta}(\underset{\sim}{\theta}_0)u_\alpha u_\beta]\, d\boldsymbol{u} = 0,$$

$$\int L(-\boldsymbol{u})u_\alpha u_\beta u_\gamma \exp[-(1/2)\textstyle\sum_{\alpha,\beta=1}^{p} I_{\alpha\beta}(\underset{\sim}{\theta}_0)u_\alpha u_\beta\}\, d\boldsymbol{u} = 0.$$

Letting $\hat{u}_\alpha = \hat{t}_\alpha - U_\alpha$, we can see that (6.4) is written as

$$0 = -M\textstyle\sum_{\beta=1}^{p} I_{\alpha\beta}(\underset{\sim}{\theta}_0)\hat{u}_\beta - n^{-1/2}\textstyle\sum_{\beta,\gamma=1}^{p} \dfrac{I_{\alpha\beta}(\underset{\sim}{\theta}_0)\xi_\gamma^{(1)}(\underset{\sim}{\theta}_0)}{\xi(\underset{\sim}{\theta}_0)} P_{\beta\gamma}$$

$$- n^{-1/2}\textstyle\sum_{\beta,\gamma,\delta=1}^{p} I_{\alpha\beta}(\underset{\sim}{\theta}_0)Z_{\gamma\delta}(\underset{\sim}{\theta}_0)U_\gamma P_{\beta\delta}$$

$$+ (6\sqrt{n})^{-1}\textstyle\sum_{\beta,\gamma,\delta,\psi=1}^{p} I_{\alpha\beta}(\underset{\sim}{\theta}_0)\rho_{\gamma\delta\psi}(\underset{\sim}{\theta}_0)(Q_{\alpha\beta\gamma\delta} + 3U_\gamma U_\delta P_{\beta\psi})$$

$$+ n^{-1/2}\dfrac{\xi_\alpha^{(1)}(\underset{\sim}{\theta}_0)}{\xi(\underset{\sim}{\theta}_0)} M + n^{-1/2}\textstyle\sum_{\beta=1}^{p} Z_{\alpha\beta}(\underset{\sim}{\theta}_0)U_\beta M$$

$$- (2\sqrt{n})^{-1}\textstyle\sum_{\beta,\gamma=1}^{p} \rho_{\alpha\beta\gamma}(\underset{\sim}{\theta}_0)(P_{\alpha\beta} + U_\beta U_\gamma M)$$

$$+ \textstyle\sum_{\beta,\gamma,\delta=1}^{p} I_{\gamma\beta}(\underset{\sim}{\theta}_0)I_{\alpha\gamma}(\underset{\sim}{\theta}_0)P_{\delta\gamma}\hat{u}_\beta + o_p(n^{-1/2}).$$

Let $\boldsymbol{E}$ be the identity matrix of order p, an $\underset{\sim}{\xi}^*$, $\boldsymbol{L}$, $\boldsymbol{V}$ and $\boldsymbol{W}$ be column vectors with

$$\xi^* = [\xi_\alpha^{(1)}(\theta_0)/\xi(\theta_0)],$$

$$L = [-(1/6)\sum_{\beta,\gamma,\delta,\psi=1}^p I_{\alpha\beta}(\theta_0)\rho_{\gamma\delta\psi}(\theta_0)Q_{\beta\gamma\delta\psi}$$
$$+ (1/2)\sum_{\beta,\gamma=1}^p \rho_{\alpha\beta\gamma}(\theta_0)P_{\beta\gamma}],$$

$$V = [\sum_{\beta,\gamma=1}^p \rho_{\alpha\beta\gamma}(\theta_0)U_\beta U_\gamma], \quad W = [\sum_{\beta=1}^p U_\beta Z_{\beta\gamma}(\theta_0)].$$

Then it follows that

$$(IPI - MI)\hat{u} = (IP - ME)\xi^*/\sqrt{n} + L/\sqrt{n} - (IP - ME)V/(2\sqrt{n})$$
$$+ (IP - ME)W/\sqrt{n} + o_p(n^{-1/2}).$$

Since $(IPI - MI)$ is positive definite, we obtain a stochastic expansion of the generalized multiparameter weighted estimator,

$$\text{(6.6)} \quad \sqrt{n}(\hat{\theta} - \theta_0) = U + I^{-1}W/\sqrt{n} - I^{-1}V/(2\sqrt{n})$$
$$+ \{I^{-1}\xi^* + (IPI - MI)^{-1}L\}/\sqrt{n} + o_p(n^{-1/2}),$$

where $U = (U_1, \cdots, U_p)'$.

Here we review the classes of estimators S and D (see Taniguchi [37]). Let

$$S = \{\hat{\theta}; \sqrt{n}(\hat{\theta}_n - \theta_0) = U + Q/\sqrt{n} + o_p(n^{-1/2}),$$
$$Q = (Q_1, \cdots, Q_p)' = O_p(1)\}.$$

We assume that $\sqrt{n}(\hat{\theta}_n - \theta_0)$ has the Edgeworth expansion up to order n^{-1} and that

$$E_{\theta_0}\sqrt{n}(\hat{\theta}_n - \theta_0) = \mu/\sqrt{n} + o(n^{-1}), \text{ and}$$
$$E_{\theta_0}Z^{(1)}Z^{(1)'} = I(\theta_0) + \Delta(\theta_0)/n + o(n^{-1}),$$

where $\mu = (\mu_1, \cdots, \mu_p)' = E_{\theta_0}Q$ and $Z^{(1)} = (Z_1, \cdots, Z_p)'$. It is known (Taniguchi [37]) that for $S = (S_1, \cdots, S_p)' = \sqrt{n}[\hat{\theta}_n - E_{\theta_0}(\hat{\theta}_n)] \in S$,

(i) $E_{\theta_0}(S_iS_j) = I^{ij} - \eta_{ij}/n + D_i\mu_j/n + D_j\mu_i/n + Cov(Q_i, Q_j)/n + o(n^{-1})$,

(ii) $E_{\theta_0}(S_iS_jS_k) = \beta_{ijk}/\sqrt{n} + A_{ijk}/(2n) + o(n^{-1})$,

(iii) $Cum(S_i, S_j, S_k, S_m) = \beta_{ijkm} + o(n^{-1})$,

where I^{ij} and η_{ij} are the (i, j) th elements of $I(\theta_0)^{-1}$ and $I(\theta_0)^{-1}\Delta(\theta_0)I(\theta_0)^{-1}$, respectively and $D_i = \sum_{k=1}^p I^{ik}\dfrac{\partial}{\partial\theta_k}$ (differential operator). Here A_{ijk}

$= E_{\underset{\sim}{\theta}_0}(U_i\tilde{Q}_j\tilde{Q}_k) + E_{\underset{\sim}{\theta}_0}(U_j\tilde{Q}_i\tilde{Q}_k) + E_{\underset{\sim}{\theta}_0}(U_k\tilde{Q}_i\tilde{Q}_j)$, and β_{ijk} and β_{ijkm} are expressed in terms of the spectral density.

Now we introduce a class $\boldsymbol{D}(\subset \boldsymbol{S})$ of estimators which satisfy $A_{ijk} = o(1)$ for $i, j, k = 1, \cdots, p$. This class $\boldsymbol{D}$ is a natural one. Taniguchi [37] showed that an estimator $\tilde{\underset{\sim}{\theta}}$ is third-order asymptotically efficient in the clas $\boldsymbol{D}$ if and only if $\tilde{\underset{\sim}{\theta}}$ has the following stochastic expansion;

$$(6.7)\qquad \sqrt{n}(\tilde{\underset{\sim}{\theta}} - \underset{\sim}{\theta}_0) = \boldsymbol{U} + \boldsymbol{I}(\underset{\sim}{\theta}_0)^{-1}\boldsymbol{Z}^{(2)}\boldsymbol{U}/\sqrt{n} + \boldsymbol{I}(\underset{\sim}{\theta}_0)^{-1}R\ldots\cdot \boldsymbol{U}\cdot \boldsymbol{U}/(2\sqrt{n}) + \underset{\sim}{\xi}/\sqrt{n} + o_p(n^{-1/2}),$$

where $\underset{\sim}{\xi}$ is a constant vector, $\boldsymbol{Z}^{(2)} = \{Z_{ij}\}$ and $\boldsymbol{R}\ldots = \{R_{ijk}\}$, $R_{ijk} = -K_{ijk} - J_{ijk} - J_{jki} - J_{kij}$, and $\boldsymbol{R}\ldots\cdot\boldsymbol{U}\cdot\boldsymbol{U}$ is a p-dimensional column vector with i th component $\sum_{j,k} R_{ijk}U_jU_k$.

Comparing with (6.6) and (6.7) we can see that stochastic expansions of $\hat{\underset{\sim}{\theta}}$ and $\tilde{\underset{\sim}{\theta}}$ up to order $n^{-1/2}$ have the same structure. Thus we establish the following theorem.

THEOREM 6.1. *Under Assumptions* 5.1–5.4 *and* 6.1–6.2, *if we modify the generalized multiparameter weighted estimator* $\hat{\underset{\sim}{\theta}}$ *of* $\underset{\sim}{\theta}_0$ *to be third-order asymptotically median unbiased, then it is third-order asymptotically efficient in the class* $\boldsymbol{D}$.

Part III. Higher order evaluation of final prediction error

7. Final prediction error (FPE) for Gaussian ARMA processes

In this section we give the definition of final prediction error for Gaussian ARMA process. First, we review Akaike's definition of "Final Prediction Error (FPE)". Suppose that $\{X(t)\}$ is a stationary autoregressive process generated by the relation

$$(7.1)\qquad X(t) = \textstyle\sum_{m=1}^{M} a(m)X(t-m) + a(0) + \varepsilon(t),$$

where $\varepsilon(t)$ are mutually independently and identically distributed random variables with $E[\varepsilon(t)] = 0$ and $E[\varepsilon^2(t)] = \sigma^2$. Assuming that $\{Y(t)\}$ is generated by the same relation as (7.1) and that $\{Y(t)\}$ is independent of $\{X(t)\}$, Akaike defined an estimated predictor $\hat{Y}(t)$ of $Y(t)$;

$$\hat{Y}(t) = \textstyle\sum_{m=1}^{M} \hat{a}_M(m)Y(t-m) + \hat{a}_M(0),$$

where $\hat{a}_M(m)$, $m = 0, 1, \cdots, M$, are defined as the least squares estimators of $a(m)$ based on the observed stretch $\{X(t); t = -M+1, -M+2, \cdots, n\}$. Then he evaluated the following quantity up to order n^{-1};

$$E_X[E_Y\{Y(t) - \hat{Y}(t)\}^2], \tag{7.2}$$

where E_X and E_Y mean the expectation with respect to $\{X(t)\}$ and $\{Y(t)\}$, respectively. The resulting evaluation is given by

$$(\text{FPE})_M \text{ of } \hat{Y}(t) = (1 + [M + 1]/n)\sigma^2. \tag{7.3}$$

Akaike's FPE criterion, which determines the order of the AR model, is defined as an asymptotically unbiased estimator of (7.3). Therefore it seems important to evaluate the basic quantity (7.2) more accurately.

In Part III we extend the above result to the case when the process concerned is Gaussian ARMA process. Then we evaluate (7.2) up the order n^{-2}., we also show that the generalized weighted estimator and the maximum likelihood estimator are best in the sense of FPE.

Let $\{X_t;\ t = 0, \pm 1, \pm 2, \cdots\}$ be a Gaussian ARMA process with spectral density $f_{\underset{\sim}{\theta}}(\lambda)$ which depends on an unknown vector $\underset{\sim}{\theta} = (\theta_1, \cdots, \theta_p)' \in \Theta \subset R^p$. We consider a stretch $X_n = (X_1, \cdots, X_n)'$ of the series $\{X_t\}$. Let Σ_n be the covariance matrix of X_n. The likelihood function based on X_n is given by

$$L_n(\underset{\sim}{\theta}) = (2\pi)^{-n/2} |\Sigma_n|^{-1/2} \exp\{-(1/2) X_n' \Sigma_n^{-1} X_n\}.$$

In Part III we use the notations and assumptions stated in Section 5. Let $\underset{\sim}{\theta}_0 = (\theta_{01}, \theta_{02}, \cdots, \theta_{0p})'$ be true parameter of $\underset{\sim}{\theta} \in \Theta$.

Now we consider the higher order evaluation of FPE for $\{X_t\}$ satisfying Assumptions 5.1-5.4. Let $\{X_t\}$ be observable. Let $\{Y_t\}$ be the ideal process with the same structure as $\{X_t\}$ and be independent of $\{X_t\}$. Let $\hat{\underset{\sim}{\theta}} = (\hat{\theta}_1, \cdots, \hat{\theta}_p)'$ be a $\sqrt{n}$ consistent estimator of $\underset{\sim}{\theta}$ calculated from $\{X_1, \cdots, X_n\}$, which will be specified later. We construct the predictor $\hat{Y}(t)$ of $Y(t)$ by $f_{\hat{\underset{\sim}{\theta}}}(\lambda)$. Then the mean square error of prediction $\hat{Y}(t)$ can be written in the following form:

$$E_Y\{Y(t) - \hat{Y}(t)\}^2 = [(2\pi)^{-1} \int_{-\pi}^{\pi} \{f_{\underset{\sim}{\theta}_0}(\lambda)/f_{\hat{\underset{\sim}{\theta}}}(\lambda)\} d\lambda$$
$$\cdot \exp\{(2\pi)^{-1} \int_{-\pi}^{\pi} \log [f_{\hat{\underset{\sim}{\theta}}}(\lambda)/f_{\underset{\sim}{\theta}_0}(\lambda)] d\lambda\} - 1]\sigma^2 + \sigma^2$$

(see Grenander and Rosenblatt [16]), p. 261). Henceforth we call $E_X[E_Y\{Y(t) - \hat{Y}(t)\}^2]$ the final prediction error for Gaussian ARMA process.

8. Some best estimators in the sense of FPE

In this section we evaluate the FPE up to higher order $O(n^{-2})$ and discuss an optimality property of GMWE and the maximum likelihood estimator in the sense of FPE. Let

$$F(\hat{\underset{\sim}{\theta}}\,;\underset{\sim}{\theta}_0)=(2\pi)^{-1}\int_{-\pi}^{\pi}\{f_{\underset{\sim}{\theta}_0}(\lambda)/f_{\hat{\underset{\sim}{\theta}}}(\lambda)\}\,d\lambda$$

$$\cdot\exp\{(2\pi)^{-1}\int_{-\pi}^{\pi}\log\,[f_{\hat{\underset{\sim}{\theta}}}(\lambda)/f_{\underset{\sim}{\theta}_0}(\lambda)]\,d\lambda\}-1.$$

For simplicity we sometime use $f_{\underset{\sim}{\theta}_0}, f_{\hat{\underset{\sim}{\theta}}}$ instead of $f_{\underset{\sim}{\theta}_0}(\lambda)$, $f_{\hat{\underset{\sim}{\theta}}}(\lambda)$, respectively.

Expanding $F(\hat{\underset{\sim}{\theta}}\,;\underset{\sim}{\theta}_0)$ in a Taylor series at $\theta=\theta_0$, it holds that

$$\begin{aligned}F(\hat{\underset{\sim}{\theta}}\,;\underset{\sim}{\theta}_0)&=F(\underset{\sim}{\theta}_0\,;\underset{\sim}{\theta}_0)+n^{-1/2}\textstyle\sum_{\alpha=1}^{p}F_\alpha\sqrt{n}(\hat\theta_\alpha-\theta_{0\alpha})\\&+(2n)^{-1}\textstyle\sum_{\alpha,\beta=1}^{p}F_{\alpha\beta}\sqrt{n}(\hat\theta_\alpha-\theta_{0\alpha})\sqrt{n}(\hat\theta_\beta-\theta_{0\beta})\\&+(6n^{3/2})^{-1}\textstyle\sum_{\alpha,\beta,\gamma=1}^{p}F_{\alpha\beta\gamma}\sqrt{n}(\hat\theta_\alpha-\theta_{0\alpha})\sqrt{n}(\hat\theta_\beta-\theta_{0\beta})\sqrt{n}(\hat\theta_\gamma-\theta_{0\gamma})\\&+(24n^2)^{-1}\textstyle\sum_{\alpha,\beta,\gamma,\delta=1}^{p}F_{\alpha\beta\gamma\delta}\sqrt{n}(\hat\theta_\alpha-\theta_{0\alpha})\sqrt{n}(\hat\theta_\beta-\theta_{0\beta})\sqrt{n}(\hat\theta_\gamma-\theta_{0\gamma})\\&\times\sqrt{n}(\hat\theta_\delta-\theta_{0\delta})+o(n^{-2}),\end{aligned}\tag{8.1}$$

where $F_\alpha=\dfrac{\partial}{\partial\hat\theta_\alpha}F(\underset{\sim}{\theta}_0;\underset{\sim}{\theta}_0)$, $F_{\alpha\beta}=\dfrac{\partial^2}{\partial\hat\theta_\alpha\partial\hat\theta_\beta}F(\underset{\sim}{\theta}_0;\underset{\sim}{\theta}_0)$, $F_{\alpha\beta\gamma}=\dfrac{\partial^3}{\partial\hat\theta_\alpha\partial\hat\theta_\beta\partial\hat\theta_\gamma}F(\underset{\sim}{\theta}_0;\underset{\sim}{\theta}_0)$ and $F_{\alpha\beta\gamma\delta}=\dfrac{\partial^4}{\partial\hat\theta_\alpha\partial\hat\theta_\beta\partial\hat\theta_\gamma\partial\hat\theta_\delta}F(\underset{\sim}{\theta}_0;\underset{\sim}{\theta}_0)$. Henceforth we assume that $\underset{\sim}{\theta}_0$ is the innovation free parameter (see Hosoya and Taniguchi [18]). Then the following can be evaluated without difficulty.

$$\begin{aligned}&F(\underset{\sim}{\theta}_0;\underset{\sim}{\theta}_0)=0,\\&F_\alpha=0,\\&F_{\alpha\beta}=(2\pi)^{-1}\int_{-\pi}^{\pi}\left\{\frac{\partial}{\partial\hat\theta_\alpha}\log f_{\hat{\underset{\sim}{\theta}}}\cdot\frac{\partial}{\partial\hat\theta_\beta}\log f_{\hat{\underset{\sim}{\theta}}}\right\}\Big|_{\hat{\underset{\sim}{\theta}}=\underset{\sim}{\theta}_0}d\lambda=2I_{\alpha\beta},\\&F_{\alpha\beta\gamma}=2(J_{\alpha\beta\gamma}+J_{\beta\gamma\alpha}+J_{\gamma\alpha\beta}+K_{\alpha\beta\gamma})\\&F_{\alpha\beta\gamma\delta}=2(2L_{\alpha\beta\gamma\delta}+L_{\gamma\alpha\beta\delta}+L_{\delta\alpha\beta\gamma})+2(M_{\alpha\gamma\beta\delta}+M_{\alpha\delta\beta\gamma}\\&\quad+M_{\gamma\delta\alpha\beta})+(N_{\beta\delta\alpha\gamma}+N_{\alpha\delta\beta\gamma}+N_{\beta\gamma\alpha\delta}+N_{\alpha\beta\gamma\delta}\\&\quad+N_{\alpha\gamma\beta\delta}+N_{\alpha\beta\gamma\delta}+N_{\alpha\gamma\beta\delta}+N_{\beta\gamma\alpha\delta}+N_{\beta\delta\alpha\gamma}+N_{\alpha\delta\beta\gamma}\\&\quad+N_{\beta\delta\alpha\gamma}+N_{\gamma\delta\alpha\beta})+2H_{\alpha\beta\gamma\delta}+4(A_{\alpha\beta}I_{\gamma\delta}+A_{\alpha\gamma}I_{\beta\delta}\\&\quad+A_{\alpha\delta}I_{\beta\gamma}+A_{\beta\gamma}I_{\alpha\delta}+A_{\beta\delta}I_{\alpha\gamma}+A_{\gamma\delta}I_{\alpha\beta})-A_{\alpha\gamma}A_{\beta\delta}\\&\quad-A_{\alpha\beta}A_{\gamma\delta}-A_{\alpha\delta}A_{\beta\gamma}-8I_{\alpha\beta}I_{\gamma\delta}-16I_{\alpha\gamma}I_{\beta\delta}\\&\quad-12I_{\alpha\delta}I_{\beta\gamma},\end{aligned}\tag{8.2}$$

where $A_{ij} = (2\pi)^{-1}\int_{-\pi}^{\pi} \frac{\partial^2}{\partial\hat{\theta}_i\partial\hat{\theta}_j} f_{\underset{\sim}{\theta}}(\lambda)\cdot f_{\underset{\sim}{\theta}}(\lambda)^{-1}d\lambda$. For the explicit forms of I, J, K, L, M, N, H, see Taniguchi [37].

We now assume that the estimator $\hat{\underset{\sim}{\theta}}$ belongs to a natural class $\boldsymbol{D}$ defined in Section 6. Then we can show that

$$E\{\sqrt{n}(\hat{\theta}_\alpha - \theta_{0\alpha})\sqrt{n}(\hat{\theta}_\beta - \theta_{0\beta})\} = I^{\alpha\beta} - \eta_{\alpha\beta}/n + D_\alpha\mu_\beta/n + D_\beta\mu_\alpha/n + \mathrm{Cov}(Q_\alpha, Q_\beta)/n + (\mu_\alpha\mu_\beta)/n + o(n^{-1}),$$

$$(8.3)\quad E\{\sqrt{n}(\hat{\theta}_\alpha - \theta_{0\alpha})\sqrt{n}(\hat{\theta}_\beta - \theta_{0\beta})\sqrt{n}(\hat{\theta}_\gamma - \theta_{0\gamma})\} = \mu_\alpha I^{\beta\gamma}/\sqrt{n} + \mu_\beta I^{\alpha\gamma}/\sqrt{n} + \mu_\gamma I^{\alpha\beta}/\sqrt{n} + \beta_{\alpha\beta\gamma}/\sqrt{n} + o(n^{-1}),$$

$$E\{\sqrt{n}(\hat{\theta}_\alpha - \theta_{0\alpha})\sqrt{n}(\hat{\theta}_\beta - \theta_{0\beta})\sqrt{n}(\hat{\theta}_\gamma - \theta_{0\gamma})\sqrt{n}(\hat{\theta}_\delta - \theta_{0\delta})\} = I^{\alpha\beta}I^{\gamma\delta} + I^{\alpha\gamma}I^{\beta\delta} + I^{\alpha\delta}I^{\beta\gamma} + O(n^{-1}).$$

Using (8.1), (8.2) and (8.3) we obtain

$$E\{F(\hat{\underset{\sim}{\theta}}\,;\underset{\sim}{\theta}_0)\} = p/n + D/n^2 + o(n^{-2}),$$

where
$$D = \textstyle\sum_{\alpha,\beta=1}^{p} I_{\alpha\beta}[-\eta_{\alpha\beta} + D_\alpha\mu_\beta + D_\beta\mu_\alpha + \mu_\alpha\mu_\beta + Cov(Q_\alpha, Q_\beta)] + \sum_{\alpha,\beta,\gamma=1}^{p}(J_{\alpha\beta\gamma} + J_{\beta\gamma\alpha} + J_{\gamma\alpha\beta} + K_{\alpha\beta\gamma})(\mu_\alpha I^{\beta\gamma} + \mu_\beta I^{\alpha\gamma} + \mu_\gamma I^{\alpha\beta} + \beta_{\alpha\beta\gamma}) + \sum_{\alpha,\beta,\gamma,\delta=1}^{p}(1/24)F_{\alpha\beta\gamma\delta}(I^{\alpha\beta}I^{\gamma\delta} + I^{\alpha\gamma}I^{\beta\delta} + I^{\alpha\delta}I^{\beta\gamma}).$$

If we modify $\hat{\underset{\sim}{\theta}} \in \boldsymbol{D}$ to be third-order AMU (say $\hat{\underset{\sim}{\theta}}^*$), then the second-order bias $\underset{\sim}{\mu}$ is specified by β_{iii}. Thus all the other terms in (8.3) can be expressed by the spectral density except for the terms $Cov(Q_\alpha, Q_\beta)$. This implies the undetermined term for $E\{F(\hat{\underset{\sim}{\theta}}\,;\theta_0)\}$ is only $Cov(Q_\alpha, Q_\beta)$. In fact, the term $Cov(Q_\alpha, Q_\beta)$ only depends upon the estimator. Thus we can see that $\hat{\underset{\sim}{\theta}}^*$ minimizes $E\{F(\hat{\underset{\sim}{\theta}};\underset{\sim}{\theta}_0)\}$ up to order n^{-2} if and only if it minimizes the matrix $\{Cov(Q_\alpha, Q_\beta)\}$. This implies the best estimator in the sense of FPE has the stochastic expansion (6.7) (see Proposition 6 of Taniguchi [37]). Summarizing the above we have

THEOREM 8.1. *If we choose the generalized multiparameter weighted estimator or the maximum likelihood estimator of* $\underset{\sim}{\theta}$ *and modify them to be third-order* AMU, *then they minimize* $E\{F(\hat{\underset{\sim}{\theta}}\,;\underset{\sim}{\theta}_0)\}$ (FPE) *up to order* n^{-2} *in* $\boldsymbol{D}$. *That is, the generalized multiparametr weighted estimator and the maximum likelihood estimator are the best estimators in the class* $\boldsymbol{D}$ *in the sense of* FPE.

Note that, Takeuchi and Akahira [3] discussed that the weighted estimator (Bayes estimator) and the maximum likelihood estimator are the best estimators in the class $\boldsymbol{D}$ in the sense of highest probability concentration up to third order.

Since general formular of $E\{F(\hat{\underset{\sim}{\theta}}\,;\underset{\sim}{\theta}_0)\}$ is too complicated, we give its explicit formula for the scalar θ_0. Let $\hat{\theta}$ be the weighted estimator or the maximum likelihood estimator of θ_0. We modify $\hat{\theta}$ to be third-order AMU, and denote it by $\hat{\underset{\sim}{\theta}}{}^*$. Thus the moments of $U_n^* = \sqrt{nI}(\hat{\theta}^* - \theta_0)$ become

$$E(U_n^{*2}) = 1 - \Delta/(In) + (-3L - 9N - 2H)/(3I^2 n) + (135J^2 + 216JK + 70K^2)/(36I^3 n) + o(n^{-1}),$$

(8.4)
$$E(U_n^{*3}) = (-9J - 6K)/(2I^{3/2}\sqrt{n}) + o(n^{-1}),$$

$$E(U_n^{*4}) = 3 + o(n^{-1}),$$

Further,

$$\frac{\partial^2 F(\hat{\theta}^*;\theta_0)}{\partial\hat{\theta}^{*2}}\Big|_{\hat{\theta}^*=\theta_0} = 2I,$$

(8.5)
$$\frac{\partial^3 F(\hat{\theta}^*;\theta_0)}{\partial\hat{\theta}^{*3}}\Big|_{\hat{\theta}^*=\theta_0} = 6J + 2K,$$

$$\frac{\partial^4 F(\hat{\theta}^*;\theta_0)}{\partial\hat{\theta}^{*4}}\Big|_{\hat{\theta}^*=}2H + 8L + 6M + 12N + 3(A - 2I)(6I - A),$$

where $A = (2\pi)^{-1}\int_{-\pi}^{\pi}\left\{\frac{\partial^2}{\partial\theta_0{}^2} f_{\theta_0}(\lambda)\right\} f_{\theta_0}(\lambda)^{-1} d\lambda$. Thus we obtain

(8.6)
$$\begin{aligned} E\{F(\hat{\theta}^*;\theta_0)\} = {} & 1/n - \Delta/(In^2) - (3L + 9N + 2H)/(3I^2 n^2) \\ & + (135J^2 + 216JK + 70K^2)/(36I^3 n^2) \\ & + (-3J - 2K)(3J + K)/(2I^3 n^2) \\ & + [2H + 8L + 6M + 12N + 3(A - 2I)(6I - A)]/(8I^2 n^2) \\ & + o(n^{-2}). \end{aligned}$$

Here we give a more explicit form for (8.6) when $\{X_t\}$ has the following ARMA(1,1) spectral density

$$f_{\theta_0}(\lambda) = (\sigma^2/2\pi)|1 - \psi e^{i\lambda}|^2/|1 - \rho e^{i\lambda}|^2.$$

Then $A(\theta_0)$ for $\theta_0 = \sigma^2$, ρ and ψ are given by

$$A(\sigma^2) = 0,\ A(\rho) = 2(1-\rho^2)^{-1} \quad \text{and} \quad A(\psi) = 2(1-\psi^2)^{-1},$$

respectively. For $\theta_0 = \sigma^2$, ρ and ψ, I, J, K, L, M, N, H and Δ are evaluated by Taniguchi [37]. Finally we obtain the following evaluations.

(i) If $\theta_0 = \sigma^2$, then $E\{F(\hat{\theta}^*; \theta_0)\} = \frac{1}{n} - 35/(18n^2) + o(n^{-2})$.

(ii) If $\theta_0 = \rho$, then

$$E\{F(\hat{\theta}^*; \theta_0)\} = \frac{1}{n} + \frac{1 + 6\rho^2 - 14\rho^3\psi + \psi^2(9\rho^4 - 3\rho^2 + 1)}{n^2(1-\rho^2)(1-\rho\psi)^2} + o(n^{-2}).$$

(iii) If $\theta_0 = \psi$, then

$$E\{F(\hat{\theta}^*; \theta_0)\} = \frac{1}{n} + \frac{10 + 12\psi^2 - \rho(22\psi + 22\psi^3) + \rho^2(14\psi^2 + 9\psi^4 - 1)}{n^2(1-\psi^2)(1-\rho\psi)^2} + o(n^{-2}).$$

REMARK 8.1. It should be noted that for $\theta_0 = \sigma^2$, the term of n^{-2} is independent from σ^2. For $\theta_0 = \rho$ and $\theta_0 = \psi$ we can not neglect the terms of order n^{-2} (i.e., they become large) when the modulus of ρ and ψ tend to 1.

ACKNOWLEDGEMENTS

The author would like to express his hearty thanks to Professor Y. Fujikoshi, Hirohsima University, for his supervision, encouragement and significant suggestions for the outcome of this paper. The author is also deeply grateful to Professor M. Taniguchi, Hiroshima University, for his recommendation to write this paper and for his invaluable guidance, inspiration and encouragement at various stages of this research.

References

[1] M. Akahira, A note on the second order asymptotic efficiency of estimators in an autoregressive process, Rep. Univ. Electro-Comm. 26–1, (Sci. & Tech. Sect.), August (1975), 143–149.

[2] M. Akahira, On the second order asymptotic optimality of estimators in an autoregressive process, Rep. Univ. Electro-Comm. 29–2, (Sci. & Tech. Sect.). February (1979), 213–218.

[3] M. Akahira and K. Takeuchi, Asymptotic efficiency of statistical estimators: Concepts and higher order asymptotic efficiency, Lecture notes in statistics, **7**, Springer-Verlag, 1981.

[4] H. Akaike, Fitting autoregressive models for prediction, Ann. Inst. Statist. Math., **21** (1969), 243–247.

[5] H. Akaike, Statistical predictor identification, Ann. Inst. Statist. Math., **22** (1970), 203–217.

[6] H. Akaike, Autoregressive model fitting for control, Ann. Inst. Statist. Math., **23** (1971), 163–180.

[7] R. T. Baillie, Asymptotic prediction mean squared error for vector autoregressive models, Biometrika, 66 (1979), 675–678.

[8] R. T. Baillie, The asymptotic mean squared error of multistep prediction from the regression model with autoregressive errors, Journal of The American Statistical Association, **74** (1979), 175–184.

[9] R. J. Bhansali, Asymptotic mean-square error of predicting more than one-step ahead using the regression method, Appl. Statist., **23** (1974), 35–42.

[10] P. Bloomfield, On the error of prediction of a time series, Biometrika, **59** (1972), 501–507.

[11] D. R. Brillinger, Asymptotic properties of spectral estimates of second order, Biometrika, **56** (1969), 375–390.

[12] D. R. Brillinger, Time series; Data analysis and theory. Holt, New York, 1975.

[13] N. Davies and P. Newbold, Forecasting with misspecified models, Appl. Statist., **29** (1980), 87–92.

[14] W. A. Fuller, Introduction to statistical time series, New York, John Wiley, 1976.

[15] Y. Fujikoshi and Y. Ochi, Asymptotic properties of the maximum likelihood estimate in the first order autoregressive process, Ann. Inst. Statist. Math., **36** (1984), 119–128.

[16] U. Grenander and M. Rosenblatt, Statistical analysis of stationary time series, New York, Wiley 1957.

[17] Y. Hosoya, Higher-order efficiency in the estimation of linear process, Ann. Statist., **7** (1979), 516–530.

[18] Y. Hosoya and M. Taniguchi, A central limit theorem for stationary process and the parameter estimation of linear processes, Ann. Statist., **10** (1982), 132–153.

[19] S. Konishi, An approximation to the distribution of the sample correlation coefficient, Biometrika, **65** (1978), 654–656.

[20] S. Konishi, Normalizing transformations of some statistics in multivariate analysis, Biometrika, **68** (1981), 647–651.

[21] N. Kunitomo and T. Yamamoto, Properties of predictors in misspecified autoregressive time series models, Journal of American Statistical Association, **80** (1985), 941–950.

[22] R. A. Lewis and G. C. Reinsel, Prediction error of multivariate time series with misspecified models, Journal of Time Series Analysis, **9** (1988), 43–57.

[23] K. Maekawa, Finite sample properties of several predictors from an autoregressive model, Econometric Theory, **3** (1987), 359–370.

[24] Myint Swe and M. Taniguchi, Higher order asymptotic properties of a weighted estimator for Gaussian ARMA processes, to appear in Journal of Time Series Analysis.

[25] Y. Ochi, Asymptotic expansions for the distribution of an estimator in the first-order autoregressive process, Journal of Time Series Analysis, **4** (1983), 57–67.

[26] B. L. S. P. Rao, The equivalence between (modified) Bayes estimator and maximum likelihood estimator for Markov processes, Ann. Inst. Statist. Math., **31** (1979), 499–513.

[27] C. R. Rao, Efficient estimates and optimum inference procedures in large samples, J. Roy. Statist. Soc. B, **24** (1962), 46–72.

[28] D. Ray, Asymptotic mean square prediction error for a multivariate autoregressive model with random coefficients, Journal of Time Series Analysis, **9** (1988), 73–80.

[29] G. Reinsel, Asymptotic properties of prediction error for the multivariate autoregrssive model using estimated parameters, J. Roy. Statist. Soc. B, **42** (1980), 328–333.

[30] K. Takeuchi and M. Akahira, Asymptotic optimality of the generalized Bayes estimator, Rep. Univ. Electro-Comm, 29–1, (Sci. & Tech. Sect.), August (1978), 37–45.

[31] K. Takeuchi and M. Akahira, Asymptotic optimality of the generalized Bayes estimator in multiparameter cases, Ann. Inst. Statist. Math., **31** (1979), 403–415.

[32] K. Takeuchi, Structures of the asymptotic best estimation theory. Surikagaku (Mathematical Science) No. **219** (1981), 5–16 (in Japanese).

[33] K. Tanaka and K. Maekawa, The sampling distributions of the predictor for an autoregressive model under misspecifications, Journal of Econometrics, **25** (1984), 327–351.

[34] M. Taniguchi, On selection of the order of the spectral density model for a stationary process, Ann. Inst. Statist, Math., **32** (1980), 401–419.

[35] M. Taniguchi, On the second order asymptotic efficiency of estimators of Gaussian ARMA processes, Ann. Statist., **11** (1983), 157–169.

[36] M. Taniguchi, Third order efficiency of the maximum likelihood estimator in Gaussian autoregressive moving average processes, In Statistical Theory and Data Analysis, (ed. K. Matsushita), Elesevier Science Publishers B. V. (North-Holland), (1985), 725–743.

[37] M Taniguchi, Third order asymptotic properties of maximum likelihood estimators for Gaussian ARMA processes, J. Multivariate., **18** (1986), 1–31.

[38] M. Taniguchi, Validity of edgeworth expansion of minimum contrast estimators for Gaussian ARMA processes, J. Multivariate Anal., **21** (1987), 1–28.

[39] M. Taniguchi, Asymptotic expansion of the distribution of some tests statistics for Gaussian ARMA processes, J. Multivariate Anal., **27** (1988), 494–511.

[40] M. Taniguchi, P. R. Krishnaiah and R. Chao, Normalizing transformations of some statistics of Gaussian ARMA processes, Ann. Inst. Statist. Math., **41** (1989), 187–197.

[41] T. Yamamoto, Asymptotic mean square prediction error for an autoregressive model with estimated coefficients, Appl. Statist., **25** (1976), 123–127.

[42] T. Yamamoto, Predictions of multivariate autoregressive moving average models, Biometrika, **68** (1981), 485–492.

Department of Mathematics,
Faculty of Science,
Hiroshima University