The Annals of Statistics
1983, Vol. 11, No. 1, 157–169

ON THE SECOND ORDER ASYMPTOTIC EFFICIENCY OF ESTIMATORS OF GAUSSIAN ARMA PROCESSES

BY MASANOBU TANIGUCHI

Hiroshima University, Japan

In this paper we investigate an optimal property of maximum likelihood and quasi-maximum likelihood estimators of Gaussian autoregressive moving average processes by the second order approximation of the sampling distribution. It is shown that appropriate modifications of these estimators for Gaussian ARMA processes are second order asymptotically efficient if efficiency is measured by the degree of concentration of the sampling distribution up to second order. This concept of efficiency was introduced by Akahira and Takeuchi (1981).

1. Introduction. Let $\{X_t; t = 0, \pm 1, \pm 2, \cdots,\}$ be a Gaussian autoregressive moving average (ARMA) process with spectral density $f_\theta(\lambda)$ which depends on an unknown parameter $\theta \in R^1$. In this paper we show that appropriately modified maximum likelihood and quasi-maximum likelihood estimators (MLE and q-MLE) of θ are second order asymptotically efficient in the sense of Akahira and Takeuchi (1981).

Although there has been much discussion of the "efficiency" of estimators of θ, the term "efficiency" has often been used only in the sense that an estimator has the same limiting distribution as that of Gaussian maximum likelihood estimator. Hosoya (1979), Akahira and Takeuchi (1981), and Takeuchi (1981) deal with higher-order efficiencies for time series analysis. Hosoya (1979) showed that the maximum likelihood estimator of a spectral parameter is second order asymptotically efficient in the sense of Rao (1962). Akahira and Takeuchi (1981) showed that an appropriately modified maximum likelihood estimator of the coefficient of an autoregressive process of order 1 is second order asymptotically efficient in the sense of degree of concentration of the sampling distribution up to second order. This concept of efficiency was introduced by Akahira and Takeuchi (1981), and these results are reviewed in Section 2. Takeuchi (1981) gives a brief guide to higher-order efficiency in time series.

2. Second order efficiency. We consider the approach of Akahira and Takeuchi (1981) whose argument proceeds as follows. Let $\mathbf{X}_T = (X_1, \cdots, X_T)'$ denote a sequence of random variables forming a stochastic process, and possessing the probability measure $P_\theta^T(\cdot)$, where $\theta \in \Theta$, a subset of the real line. If an estimator $\hat{\theta}_T$ satisfies the equation

$$\lim_{T\to\infty} \sqrt{T}\, | P_\theta^T\{\sqrt{T}(\hat{\theta}_T - \theta) \le 0\} - ½ | = 0, \tag{2.1}$$

then $\hat{\theta}_T$ is called a second order asymptotically median unbiased (or second order AMU for short). For this $\hat{\theta}_T$, the asymptotic distribution functions $F_\theta^+(x) + \frac{1}{\sqrt{T}} G_\theta^+(x)$ and $F_\theta^-(x) + \frac{1}{\sqrt{T}} G_\theta^-(x)$ are defined to be the second order asymptotic distributions of $\sqrt{T}\ (\hat{\theta}_T - \theta)$ if

Received February 1982; revised September 1982.

AMS 1980 *subject classifications.* Primary 62F12, 62M15; secondary 62M10, 62E20.

Key words and phrases. Gaussian autoregressive moving average processes, spectral density, periodogram, Toeplitz matrix, maximum likelihood estimator, quasi-maximum likelihood estimator, second order asymptotic efficiency, Gram-Charlier expansion, residue theorem.

$$\lim_{T\to\infty}\sqrt{T}\left| P_\theta^T\{\sqrt{T}(\hat\theta_T-\theta)\le x\} - F_\theta^+(x) - \frac{1}{\sqrt{T}}G_\theta^+(x)\right| = 0 \tag{2.2}$$

for all $x \ge 0$,

$$\lim_{T\to\infty}\sqrt{T}\left| P_\theta^T\{\sqrt{T}(\hat\theta_T-\theta)\le x\} - F_\theta^-(x) - \frac{1}{\sqrt{T}}G_\theta^-(x)\right| = 0 \tag{2.3}$$

for all $x < 0$.

For $\theta_0 \in \Theta$, consider the problem of testing hypothesis $H^+:\theta = \theta_0 + \dfrac{x}{\sqrt{T}}$ $(x>0)$ against alternative $K:\theta=\theta_0$. We define $\beta_{\theta_0}^+(x)$ and $\gamma_{\theta_0}^+(x)$ as follows:

$$\sup_{\{A_T\}\in\Phi_x}\limsup_{T\to\infty}\sqrt{T}\left\{P_{\theta_0}^T(A_T) - \beta_{\theta_0}^+(x) - \frac{1}{\sqrt{T}}\gamma_{\theta_0}^+(x)\right\} = 0, \tag{2.4}$$

where Φ_x is the class of sets $A_T = \{\sqrt{T}(\tilde\theta_T - \theta)\le x\}$, with $\tilde\theta_T$ second order AMU. Then we have for $x>0$

$$P_{\theta_0+x/\sqrt{T}}^T(A_T) = P_{\theta_0+x/\sqrt{T}}^T\{\sqrt{T}(\tilde\theta_T - \theta_0 - x/\sqrt{T})\le 0\} = ½ + o(1/\sqrt{T}).$$

By (2.2) and (2.4) we have

$$\limsup_{T\to\infty}\sqrt{T}\left\{F_{\theta_0}^+(x) + \frac{1}{\sqrt{T}}G_{\theta_0}^+(x) - \beta_{\theta_0}^+(x) - \frac{1}{\sqrt{T}}\gamma_{\theta_0}^+(x)\right\} \le 0 \quad \text{for all} \quad x>0.$$

Also consider the problem of the testing hypothesis $H^-:\theta = \theta_0 + x/\sqrt{T}$ $(x<0)$ against alternative $K:\theta=\theta_0$. Then we define $\beta_{\theta_0}^-(x)$ and $\gamma_{\theta_0}^-(x)$ as follows:

$$\inf_{\{A_T\}\in\Phi_x}\liminf_{T\to\infty}\sqrt{T}\left\{P_{\theta_0}^T(A_T) - \beta_{\theta_0}^-(x) - \frac{1}{\sqrt{T}}\gamma_{\theta_0}^-(x)\right\} = 0. \tag{2.5}$$

In the same way as for the case $x>0$, by (2.3) and (2.5) we have

$$\liminf_{T\to\infty}\sqrt{T}\left\{F_{\theta_0}^-(x) + \frac{1}{\sqrt{T}}G_{\theta_0}^-(x) - \beta_{\theta_0}^-(x) - \frac{1}{\sqrt{T}}\gamma_{\theta_0}^-(x)\right\} \ge 0 \quad \text{for all} \quad x<0.$$

Thus we make the following definition.

DEFINITION 1 (Akahira and Takeuchi, 1981). A second order AMU $\{\hat\theta_T\}$ is called second order asymptotically efficient if for each $\theta\in\Theta$

$$\lim_{T\to\infty}P_\theta^T\{\sqrt{T}(\hat\theta_T-\theta)\le x\} = \begin{cases} \beta_\theta^+(x) + \dfrac{1}{\sqrt{T}}\gamma_\theta^+(x) + o(1/\sqrt{T}) & \text{for all} \quad x\ge 0 \\ \beta_\theta^-(x) + \dfrac{1}{\sqrt{T}}\gamma_\theta^-(x) + o(1/\sqrt{T}) & \text{for all} \quad x< 0. \end{cases}$$

The above definition means that if $\hat\theta_T^*$ is second order asymptotically efficient, then we have, for any second order AMU estimator $\hat\theta_T$,

$$\liminf_{T\to\infty}\sqrt{T}[P_\theta^T\{-a<\sqrt{T}(\hat\theta_T^*-\theta)<b\} - P_\theta^T\{-a<\sqrt{T}(\hat\theta_T-\theta)<b\}]\ge 0,$$

for all $\theta\in\Theta$ and $a>0$, $b>0$. That is, this second order asymptotic efficiency implies the highest probability concentration around the true value with respect to the second order asymptotic distribution. We can regard the bound distribution $\beta_{\theta_0}^+(x) + \gamma_{\theta_0}^+(x)/\sqrt{T}$ as an approximation of the power function of the testing hypothesis $H^+:\theta=\theta_0 + x/\sqrt{T}$ $(x>0)$ against alternative $K:\theta=\theta_0$ at significance level $½ + o(1/\sqrt{T})$. By the fundamental lemma of Neyman and Pearson this bound distribution can be given by deriving the asymptotic expansion of the likelihood ratio test which tests the null hypothesis $H:\theta = \theta_0 + x/\sqrt{T}(x > 0)$ against the alternative $K:\theta = \theta_0$ at significance level $\frac{1}{2} + o(1/\sqrt{T})$. If we consider the problem of testing hypothesis $H^- : \theta = \theta_0 + x/\sqrt{T}$ $(x < 0)$ against the alternative $K:\theta = \theta_0$, we can give $\beta_{\theta_0}^-(x) + \gamma_{\theta_0}^-(x)/\sqrt{T}$ similarly.

3. Basic theorem. In this section we present a basic theorem which enables us to evaluate the asymptotic moments of the maximum likelihood estimator.

We introduce $\mathscr{D}_\Delta$ and $\mathscr{D}_{\text{ARMA}}$, spaces of functions on $[-\pi, \pi]$ defined by

$$\mathscr{D}_\Delta = \{f : f(\lambda) = \textstyle\sum_{u=-\infty}^{\infty} a(u)\exp(-iu\lambda),\ a(u) = a(-u), \sum_{u=-\infty}^{\infty} |u|\,|a(u)| < \infty\},$$

$$\mathscr{D}_{\text{ARMA}} = \left\{ f : f(\lambda) = \frac{\sigma^2}{2\pi} \frac{|\sum_{j=0}^{q} \alpha_j e^{ij\lambda}|^2}{|\sum_{j=0}^{p} \beta_j e^{ij\lambda}|^2},\ (\sigma^2 > 0) \right\}.$$

In this latter expression p and q are positive integers, and $A(z) = \sum_{j=0}^{q} \alpha_j z^j$ and $B(z) = \sum_{j=0}^{p} \beta_j z^j$ are both bounded away from zero for $|z| \le 1$. Noting Theorem 3.8.3 in Brillinger (1975), we have the following proposition.

PROPOSITION 1. (i) *If* $f_1, f_2 \in \mathscr{D}_\Delta$, *then* $f_1 \cdot f_2 \in \mathscr{D}_\Delta$.
(ii) *If* $f \in \mathscr{D}_{\text{ARMA}}$, *then* $f^{-1} \in \mathscr{D}_{\text{ARMA}}$.
(iii) *If* $f \in \mathscr{D}_{\text{ARMA}}$, *then* $f \in \mathscr{D}_\Delta$.

For the subsequent discussions we introduce the following theorem.

THEOREM 1. *Suppose that* $f_1(\lambda), \cdots, f_s(\lambda) \in \mathscr{D}_\Delta, g_1(\lambda), \cdots, g_s(\lambda) \in \mathscr{D}_{\text{ARMA}}$. *We define* $\Gamma_1, \cdots, \Gamma_s, \Lambda_1, \cdots, \Lambda_s$, *the* $T \times T$ *Toeplitz type matrices, by*

$$\Gamma_j = \left(\int_{-\pi}^{\pi} e^{i(m-n)\lambda} f_j(\lambda)\, d\lambda \right),$$

$$\Lambda_j = \left(\int_{-\pi}^{\pi} e^{i(m-n)\lambda} g_j(\lambda)\, d\lambda \right),$$

$m, n, = 1, \cdots, T, j = 1, \cdots, s$. *If* $\phi^{(T)}(k)$, $k = 1, \cdots, T$, *are the eigenvalues of* $\Gamma_1\Lambda_1^{-1}\Gamma_2\Lambda_2^{-1} \cdots \Gamma_s\Lambda_s^{-1}$, *then*

$$\frac{1}{T}\sum_{k=1}^{T} \phi^{(T)}(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_1(\lambda) \cdots f_s(\lambda) g_1(\lambda)^{-1} \cdots g_s(\lambda)^{-1}\, d\lambda + O(T^{-1}).$$

PROOF. First, we show that each Λ_j is nonsingular. Since $g_j \in \mathscr{D}_{\text{ARMA}}$, there exist F_1, F_2 such that $0 < F_1 < g_j(\lambda) < F_2 < \infty$. If $\rho_1 \le \cdots \le \rho_T$ are the eigenvalues of Λ_j, we have $2\pi F_1 \le \rho_1 \le \cdots \le \rho_T \le 2\pi F_2$ (Grenander and Szegö, 1958, page 64), which implies the nonsingularity of Λ_j.

Second we show that

$$T^{-1}\text{tr}\{M_T(\psi_1) \cdots M_T(\psi_\ell) - M_T(\psi_1 \cdots \psi_\ell)\} = O(T^{-1}), \tag{3.1}$$

where $\psi_1, \cdots, \psi_\ell \in \mathscr{D}_\Delta$, and $M_T(\psi_j)$ is the $T \times T$-Toeplitz type matrix,

$$M_T(\psi_j) = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(r-t)\lambda}\psi_j(\lambda)\, d\lambda \right), \quad r, t = 1, \cdots, T,\ j = 1, \cdots, \ell.$$

Denote $m_{rt}(\psi_j)$ for the (r, t)-th element of $M_T(\psi_j)$. Since $\psi_j \in \mathscr{D}_\Delta$, it follows that

$$\psi_j(\lambda) = \textstyle\sum_{u=-\infty}^{\infty} \gamma_j(u) e^{-iu\lambda},$$

where

$$\sum_{u=-\infty}^{\infty} | u | | \gamma_j(u)| < \infty, j = 1, \cdots, \ell.$$

Let $S_T = \frac{1}{T}\operatorname{tr}\{M_T(\psi_1) \cdots M_T(\psi_\ell)\}$ and $L_T = \frac{1}{T}\operatorname{tr}\{M_T(\psi_1, \cdots \psi_\ell)\}$.

We have

$$\begin{aligned} S_T &= \frac{1}{T}\sum_{1\le n_1,\cdots,n_\ell\le T} m_{n_1n_2}(\psi_1)m_{n_2n_3}(\psi_2) \cdots m_{n_\ell n_1}(\psi_\ell) \\ &= \frac{1}{T}\sum_{1\le n_1,\cdots,n_\ell\le T} \frac{1}{2\pi}\int_{-\pi}^{\pi} \psi_1(\lambda)e^{i(n_1-n_2)\lambda}\, d\lambda \cdots \frac{1}{2\pi}\int_{-\pi}^{\pi} \psi_\ell(\lambda)e^{i(n_\ell-n_1)\lambda}\, d\lambda \\ &= \frac{1}{T}\sum_{1\le n_1,\cdots,n_\ell\le T} \gamma_1(n_1-n_2) \cdots \gamma_\ell(n_\ell-n_1) \\ &= \sum_{-T+1\le j_1,\cdots,j_{\ell-1}\le T-1} \gamma_1(j_1) \cdots \gamma_{\ell-1}(j_{\ell-1})\gamma_\ell(-j_1-\cdots-j_{\ell-1}) \times \frac{T-K(j_1,\cdots,j_{\ell-1})}{T}, \end{aligned}$$

where $K(j_1, \cdots, j_{\ell-1})$ is chosen suitably and satisfies

$$| K(j_1, \cdots, j_{\ell-1})| \le | j_1 | + \cdots + | j_{\ell-1} |.$$

On the other hand

$$\begin{aligned} L_T &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \psi_1(\lambda) \cdots \psi_\ell(\lambda)\, d\lambda \\ &= \sum_{-\infty\le j_1,\cdots,j_{\ell-1}\le\infty} \gamma_1(j_1) \cdots \gamma_{\ell-1}(j_{\ell-1})\gamma_\ell(-j_1-\cdots-j_{\ell-1}). \end{aligned}$$

Thus we have

$$\begin{aligned} | S_T - L_T | &\le \sum{}^* | \gamma_1(j_1) \cdots \gamma_\ell(-j_1-\cdots-j_{\ell-1})| \\ &\quad + \sum_{|j_1|,\cdots,|j_{\ell-1}|\le T-1} | \gamma_1(j_1) \cdots \gamma_\ell(-j_1-\cdots-j_{\ell-1})| \frac{| j_1 | + \cdots + | j_{\ell-1} |}{T} \end{aligned} \tag{3.2}$$

where $\sum^* = \sum_{|j_1|,\cdots,|j_{\ell-1}|<\infty} - \sum_{|j_1|,\cdots,|j_{\ell-1}|\le T-1}$.

The first term of the right hand side of (3.2) is bounded by

$$\begin{aligned} &\frac{1}{T}\sum_{k=1}^{\ell-1}\sum_{j_1=-\infty}^{\infty} \cdots \sum_{j_{k-1}=-\infty}^{\infty}\sum_{|j_k|\ge T}\sum_{j_{k+1}=-\infty}^{\infty} \cdots \sum_{j_{\ell-1}=-\infty}^{\infty} | j_k | | \gamma_1(j_1)| \\ &\qquad \times \cdots \times | \gamma_\ell(-j_1-\cdots-j_{\ell-1})| \\ &\le \frac{1}{T}\sum_{k=1}^{\ell-1}\sum_{j_1=-\infty}^{\infty} \cdots \sum_{j_{k-1}=-\infty}^{\infty}\sum_{|j_k|\ge T}^{\infty}\sum_{j_{k+1}=-\infty}^{\infty} \cdots \sum_{j_{\ell-1}=-\infty}^{\infty} | j_k | | \gamma_1(j_1)| \cdots | \gamma_{\ell-1}(j_{\ell-1})| \\ &\qquad \times \sum_{j=-\infty}^{\infty} | \gamma_\ell(j)| = o(T^{-1}). \end{aligned}$$

The second term of the right hand side of (3.2) is bounded by

$$\frac{1}{T}\sum_{-\infty\le j_1,\cdots,j_{\ell-1}\le\infty} (| j_1 | + \cdots + | j_{\ell-1} |)| \gamma_1(j_1)| \cdots | \gamma_{\ell-1}(j_{\ell-1})| \times \sum_{j=-\infty}^{\infty} | \gamma_\ell(j)| = O(T^{-1}).$$

Thus we have completed the proof of (3.1).

In the third step, we show that

$$\frac{1}{T}\operatorname{tr}\{M_T(f_1)M_T(g_1)^{-1} \cdots M_T(f_{s-1})M_T(g_{s-1})^{-1}M_T(f_s)(M_T(g_s)^{-1} - M_T(g_s^{-1}))\} = O(T^{-1}). \tag{3.3}$$

Put $M = M_T(f_1)M_T(g_1)^{-1} \cdots M_T(f_{s-1})M_T(g_{s-1})^{-1}M_T(f_s)\{M_T(g_s)^{-1} - M_T(g_s^{-1})\}$. Then we have

$$\frac{1}{T}\operatorname{tr} M = \frac{1}{2T}\operatorname{tr}(M + M') \le \frac{1}{2T}\| M + M' \| \operatorname{rank}(M + M')$$
$$\le \frac{1}{2T}(\| M \| + \| M' \|)(\operatorname{rank} M + \operatorname{rank} M') \le \frac{2}{T}\| M \| \operatorname{rank} M,$$

where $\|M\|$ = the square root of the largest eigenvalue of MM'. (If M is symmetric, $\|M\|$ = the largest eigenvalue of M.) Here we have

$$\| M \| \le \| M_T(f_1)\| \, \| M_T(g_1)^{-1} \| \cdots \| M_T(f_s)\| \, \| M_T(g_s)^{-1} - M_T(g_s^{-1})\|$$
$$\le \| M_T(f_1)\| \, \| M_T(g_1)^{-1} \| \cdots \| M_T(f_s)\| \{\| M_T(g_s)^{-1} \| + \| M_T(g_s^{-1})\|\}.$$

Since there exist F_j and K_j such that $|f_j(\lambda)| \le F_j < \infty$ and $0 < K_j \le g_j(\lambda)$, then

$$\| M_T(f_j)\| \le F_j, \ \| M_T(g_s)^{-1} \| \le 1/K_j \quad \text{and} \quad \| M_T(g_s^{-1})\| \le 1/K_j$$

(Grenander and Szegö, 1958, page 64). Thus $\|M\|$ is bounded. Now

$$\operatorname{rank} M \le \operatorname{rank}\{M_T(g_s)^{-1} - M_T(g_s^{-1})\} = \min\{2 \max(p, q), T\}$$

(Shaman, 1976) and this implies (3.3). Repeated use of (3.3) shows that

$$\frac{1}{T}\operatorname{tr}\{M_T(f_1)M_T(g_1)^{-1} \cdots M_T(f_s)M_T(g_s)^{-1} - M_T(f_1)M_T(g_1^{-1}) \cdots M_T(f_s)M_T(g_s^{-1})\} = O(T^{-1}).$$

By (3.1) we have

$$\frac{1}{T}\operatorname{tr}\{M_T(f_1)M_T(g_1)^{-1} \cdots M_T(f_s)M_T(g_s)^{-1} - M_T(f_1 g_1^{-1} \cdots f_s g_s^{-1})\} = O(T^{-1}),$$

which completes the proof. □

4. Second order asymptotic efficiency of the maximum likelihood estimator in an ARMA model. In this section we shall show that if we appropriately modify the Gaussian maximum likelihood estimator in an ARMA model, then it is second order asymptotically efficient in the sense of Definition 1. In the first place we shall give the bound distributions $\beta_\theta^+(x) + (1/\sqrt{T})\,\gamma_\theta^+(x)$ and $\beta_\theta^-(x) + (1/\sqrt{T})\,\gamma_\theta^-(x)$ defined in the previous section. Using the fundamental lemma of Neyman and Pearson these are given by the likelihood ratio test which tests the null hypothesis $H: \theta = \theta_0 + x/\sqrt{T}$ against the alternative $K: \theta = \theta_0$.

We now set down the following assumptions.

ASSUMPTION 1. *X_t is a Gaussian stationary process with the spectral density $f_\theta(\lambda) \in \mathscr{D}_{\mathrm{ARMA}}$, $\theta \in R^1$, and mean 0.*

ASSUMPTION 2. *The spectral density $f_\theta(\lambda)$ is continuously three times differentiable with respect to θ, and the derivatives $\partial f_\theta/\partial\theta$, $\partial^2 f_\theta/\partial\theta^2$ and $\partial^3 f_\theta/\partial\theta^3$ belong to $\mathscr{D}_\Delta$.*

ASSUMPTION 3. *If $\theta_1 \ne \theta_2$, then $f_{\theta_1} \ne f_{\theta_2}$ on a set of positive Lebesgue measure.*

ASSUMPTION 4.

$$I(\theta) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left\{\frac{\partial}{\partial\theta}\log f_\theta(\lambda)\right\}^2 d\lambda > 0.$$

Suppose that a stretch $\mathbf{X}_T = (X_1, \cdots, X_T)'$ of the series $\{X_t\}$ is available. Let Σ_T be the covariance matrix of $\mathbf{X}_T$. The (m, n)-element of Σ_T is given by $\int_{-\pi}^{\pi} \exp\{i(m-n)\lambda\} f_\theta(\lambda)\, d\lambda$. The likelihood function based on $\mathbf{X}_T$ is given by

$$L(\theta) = (2\pi)^{-T/2} \mid \Sigma_T \mid^{-1/2} \exp(-\tfrac{1}{2}\, \mathbf{X}_T' \Sigma_T^{-1} \mathbf{X}_T).$$

Consider the problem of testing the hypothesis $H: \theta = \theta_0 + x/\sqrt{T}$ $(x > 0)$ against the alternative $K: \theta = \theta_0$. Let $LR = \log\{L(\theta_0)/L(\theta_1)\}$, where $\theta_1 = \theta_0 + x/\sqrt{T}$. If $\theta = \theta_0$, then we have

$$LR = -\frac{x}{\sqrt{T}}\left\{\frac{\partial}{\partial\theta}\log L(\theta)\right\}_{\theta_0} - \frac{x^2}{2T}\left\{\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right\}_{\theta_0} - \frac{x^3}{6T\sqrt{T}}\left\{\frac{\partial^3}{\partial\theta^3}\log L(\theta)\right\}_{\theta_0} + \text{lower order terms.} \tag{4.1}$$

Now

$$\frac{\partial \log L(\theta)}{\partial\theta} = \frac{1}{2}\mathbf{X}_T'\Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\mathbf{X}_T - \frac{1}{2}\operatorname{tr}(\Sigma_T^{-1}\dot{\Sigma}_T),$$

$$\frac{\partial^2 \log L(\theta)}{\partial\theta^2} = -\mathbf{X}_T'\Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\mathbf{X}_T + \frac{1}{2}\mathbf{X}_T'\Sigma_T^{-1}\ddot{\Sigma}_T\Sigma_T^{-1}\mathbf{X}_T - \frac{1}{2}\operatorname{tr}(\Sigma_T^{-1}\ddot{\Sigma}_T - \Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\dot{\Sigma}_T),$$

and

$$\begin{aligned}\frac{\partial^3 \log L(\theta)}{\partial\theta^3} = {} & 3\mathbf{X}_T'\Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\mathbf{X}_T \\ & - \frac{3}{2}\mathbf{X}_T'\Sigma_T^{-1}\ddot{\Sigma}_T\Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\mathbf{X}_T \\ & - \frac{3}{2}\mathbf{X}_T'\Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\ddot{\Sigma}_T\Sigma_T^{-1}\mathbf{X}_T + \frac{1}{2}\mathbf{X}_T'\Sigma_T^{-1}\dddot{\Sigma}_T\Sigma_T^{-1}\mathbf{X}_T \\ & - \frac{1}{2}\operatorname{tr}\{\Sigma_T^{-1}\dddot{\Sigma}_T - 3\,\Sigma_T^{-1}\dot{\Sigma}_T\Sigma_T^{-1}\ddot{\Sigma}_T + 2(\Sigma_T^{-1}\dot{\Sigma}_T)^3\},\end{aligned}$$

where $\dot{\Sigma}_T$, $\ddot{\Sigma}_T$ and $\dddot{\Sigma}_T$ are the $T \times T$ Toeplitz type matrices whose (m, n)th elements are given by

$$\int_{-\pi}^{\pi} e^{i(m-n)\lambda}\frac{\partial}{\partial\theta} f_\theta(\lambda)\, d\lambda, \quad \int_{-\pi}^{\pi} e^{i(m-n)\lambda}\frac{\partial^2}{\partial\theta^2} f_\theta(\lambda)\, d\lambda \quad \text{and} \quad \int_{-\pi}^{\pi} e^{i(m-n)\lambda}\frac{\partial^3}{\partial\theta^3} f_\theta(\lambda)\, d\lambda,$$

respectively. Hereafter the detailed calculations, which are omitted because of pressure on space, can be obtained from the author. Using Theorem 1 we can show that

$$E_{\theta_0}(LR) = \frac{x^2}{2} I(\theta_0) + \frac{x^3}{6\sqrt{T}}\{3J(\theta_0) + K(\theta_0)\} + O(T^{-1}), \tag{4.2}$$

$$\operatorname{Var}_{\theta_0}(LR) = x^2 I(\theta_0) + \frac{x^3}{\sqrt{T}} J(\theta_0) + O(T^{-1}), \tag{4.3}$$

$$E_{\theta_0}\{LR - E_{\theta_0}(LR)\}^3 = -\frac{x^3}{\sqrt{T}} K(\theta_0) + O(T^{-1}), \tag{4.4}$$

$$E_{\theta_1}(LR) = -\frac{x^2}{2} I(\theta_0) - \frac{x^3}{6\sqrt{T}}\{3J(\theta_0) + 2K(\theta_0)\} + O(T^{-1}), \tag{4.5}$$

$$\operatorname{Var}_{\theta_1}(LR) = x^2 I(\theta_0) + \frac{x^3}{\sqrt{T}}\{J(\theta_0) + K(\theta_0)\} + O(T^{-1}), \tag{4.6}$$

$$E_{\theta_1}\{LR - E_{\theta_1}(LR)\}^3 = -\frac{x^3}{\sqrt{T}} K(\theta_0) + O(T^{-1}), \tag{4.7}$$

where

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

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

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

In general if a random variable Y_T satisfies

$$E_\theta(Y_T) = \mu + \frac{c_1}{\sqrt{T}} + O(T^{-1}), \tag{4.8}$$

$$\mathrm{Var}_\theta(Y_T) = v^2 + \frac{c_2}{\sqrt{T}} + O(T^{-1}), \tag{4.9}$$

$$E_\theta\{Y_T - E_\theta(Y_T)\}^3 = \frac{c_3}{\sqrt{T}} + O(T^{-1}), \tag{4.10}$$

then we have the following Gram-Charlier expansion

$$\begin{aligned} P_\theta^T(Y_T \le a) = \Phi\left(\frac{a-\mu}{v}\right) - \phi\left(\frac{a-\mu}{v}\right)\Bigg[\frac{c_1}{v\sqrt{T}} + \frac{c_2}{2v^2\sqrt{T}}\left(\frac{a-\mu}{v}\right) \\ + \frac{c_3}{6v^3\sqrt{T}}\left\{\left(\frac{a-\mu}{v}\right)^2 - 1\right\}\Bigg] + O(T^{-1}), \end{aligned} \tag{4.11}$$

where $\Phi(x) = \int_{-\infty}^{x} \phi(u)\,du$ and $\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$. If we choose

$$a = -\frac{x^2 I(\theta_0)}{2} - \frac{x^3}{6\sqrt{T}}\{3J(\theta_0) + 2K(\theta_0)\} + \frac{x}{6\sqrt{T}}\frac{K(\theta_0)}{I(\theta_0)} + O(T^{-1}),$$

then, using (4.11), we have

$$P_{\theta_1}^T(LR \ge a) = ½ + O(T^{-1}).$$

Now putting $W_T = -\{LR - a - x^2 I(\theta_0)\}$, we can show that

$$\begin{aligned} &P_{\theta_0}^T\{W_T \le x^2 I(\theta_0)\} \\ &= \Phi(x\sqrt{I(\theta_0)}) + \frac{x^2}{6\sqrt{I(\theta_0)}\sqrt{T}}\{3J(\theta_0) + 2K(\theta_0)\}\phi(x\sqrt{I(\theta_0)}) + O(T^{-1}). \end{aligned} \tag{4.12}$$

If $\{\hat\theta_T\}$ is second order AMU, then remembering (2.4) and the fundamental lemma of Neyman and Pearson, we have

THEOREM 2.

$$\begin{aligned} &\limsup_{T\to\infty} \sqrt{T}\Bigg(P_{\theta_0}^T\{\sqrt{T}(\hat\theta_T - \theta_0) \le x\} \\ &- \Phi(x\sqrt{I(\theta_0)}) - \phi(x\sqrt{I(\theta_0)})\left[\frac{x^2}{6\sqrt{I(\theta_0)T}}\{3J(\theta_0) + 2K(\theta_0)\}\right]\Bigg) \le 0, \quad \text{for } x \ge 0. \end{aligned} \tag{4.13}$$

For $x < 0$, we have

(4.14)
$$\liminf_{T\to\infty} \sqrt{T}\Bigg(P^T_{\theta_0}\{\sqrt{T}(\hat\theta_T - \theta_0) \le x\} - \Phi(x\sqrt{I(\theta_0)}) - \phi(x\sqrt{I(\theta_0)})\left[\frac{x^2}{6\sqrt{I(\theta_0)T}}\{3J(\theta_0) + 2K(\theta_0)\}\right]\Bigg) \ge 0.$$

REMARK 1. In the special case of

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

i.e., an autoregressive model of order 1, the above bound distribution becomes

$$\Phi(x/\sqrt{1-\theta_0^2}) + \phi(x/\sqrt{1-\theta_0^2})\{\theta_0 x^2(1-\theta_0^2)^{-3/2}\}/\sqrt{T},$$

which coincides with the result of Akahira and Takeuchi (1981, pages 134–135) under the Gaussian assumption.

Putting

$$U_T = \sqrt{T}(\hat\theta_{ML} - \theta),$$

$$Z_1(\theta) = \frac{1}{\sqrt{T}}\frac{\partial}{\partial\theta}\log L(\theta) \quad \text{and}$$

$$Z_2(\theta) = \frac{1}{\sqrt{T}}\left[\frac{\partial^2}{\partial\theta^2}\log L(\theta) - E_\theta\left\{\frac{\partial^2}{\partial\theta^2}\log L(\theta)\right\}\right],$$

we can show the following.

THEOREM 3. *Under Assumptions* 1, 2, 3 *and* 4,

$$U_T = \frac{Z_1(\theta)}{I(\theta)} + \frac{Z_1(\theta)Z_2(\theta)}{I(\theta)^2\sqrt{T}} - \frac{3J(\theta)+K(\theta)}{2I(\theta)^3\sqrt{T}}Z_1(\theta)^2 + o_p(1/\sqrt{T}).$$

In the same way as the previous calculations, we can show

$$E_\theta U_T = \frac{-J(\theta) - K(\theta)}{2I(\theta)^2\sqrt{T}} + o(1/\sqrt{T}), \tag{4.15}$$

$$\mathrm{Var}_\theta(U_T) = I(\theta)^{-1} + O(T^{-1}), \tag{4.16}$$

$$E_\theta\{U_T - E_\theta(U_T)\}^3 = -\frac{3J(\theta)+2K(\theta)}{I(\theta)^3\sqrt{T}} + O(T^{-1}). \tag{4.17}$$

Using (4.11), if we put

$$\hat\theta^*_{ML} = \hat\theta_{ML} + \frac{K(\hat\theta_{ML})}{6TI(\hat\theta_{ML})^2} = \hat\theta_{ML} + \frac{K(\theta)}{6TI(\theta)^2} + o_p(T^{-1}),$$

then we obtain

$$P^T_\theta\{\sqrt{T}(\hat\theta^*_{ML} - \theta) \le 0\} = ½ + o(1/\sqrt{T}), \tag{4.18}$$

($\hat\theta^*_{ML}$ is a second order AMU), and

(4.19)
$$P^T_\theta\{\sqrt{T}(\hat\theta^*_{ML} - \theta) \le x\} = \Phi(x\sqrt{I(\theta)}) + \frac{x^2}{6\sqrt{TI(\theta)}}\{3J(\theta) + 2K(\theta)\}\phi(x\sqrt{I(\theta)}) + O(T^{-1}).$$

Remembering Theorem 2, we can see that (4.19) coincides with the bound distribution. Thus we have

THEOREM 4. *The modified maximum likelihood estimator* $\hat{\theta}^*_{ML}$ *is second order asymptotically efficient.*

5. Second order asymptotic efficiency of a quasi-maximum likelihood estimator. In the previous section we showed that an appropriately modified $\hat{\theta}_{ML}$ is second order asymptotically efficient. However if T is large the exact theory is intractable in practice because the likelihood function $L(\theta)$ needs the inversion procedure of the $T \times T$ matrix Σ_T. Thus we often use handy 'quasi' likelihoods as approximations. In this section we shall investigate an optimal property of a quasi-maximum likelihood estimator $\hat{\theta}_{qML}$ of θ, which maximizes the quasi-likelihood

$$\ell_T(\theta) = -\tfrac{1}{2} \textstyle\sum_{j=0}^{T-1} \{\log f_\theta(\lambda_j) + I_T(\lambda_j)/f_\theta(\lambda_j)\},$$

with respect to θ, where $\lambda_j = 2\pi j/T$, and

$$I_T(\lambda_j) = \frac{1}{2\pi T} |\textstyle\sum_{t=1}^T X_t e^{-it\lambda_j}|^2.$$

Then we shall show that an appropriately modified $\hat{\theta}_{qML}$ is second order asymptotically efficient in the sense of Akahira and Takeuchi (1981).

We set

$$\tilde{Z}_1(\theta) = \frac{1}{\sqrt{T}} \frac{\partial \ell_T(\theta)}{\partial \theta}, \tag{5.1}$$

$$\tilde{Z}_2(\theta) = \frac{1}{\sqrt{T}} \left[\frac{\partial^2}{\partial \theta^2} \ell_T(\theta) - E_\theta\left\{ \frac{\partial^2}{\partial \theta^2} \ell_T(\theta) \right\} \right], \tag{5.2}$$

$$B(\theta) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left\{ \frac{\partial}{\partial \theta} f_\theta(\lambda) \right\} b_\theta(\lambda) \{f_\theta(\lambda)\}^{-2} \, d\lambda, \tag{5.3}$$

$$b_\theta(\lambda) = \frac{1}{2\pi} \textstyle\sum_{n=-\infty}^{\infty} |n| \gamma(n) e^{in\lambda}, \quad \text{where } \lambda(n) = E_\theta X_t X_{t+n}. \tag{5.4}$$

Putting $V_T = \sqrt{T}(\hat{\theta}_{qML} - \theta)$, we can show

THEOREM 5. *Under Assumptions* 1, 2, 3 *and* 4,

$$V_T = \frac{\tilde{Z}_1(\theta)}{I(\theta)} + \frac{\tilde{Z}_1(\theta)\tilde{Z}_2(\theta)}{\{I(\theta)\}^2\sqrt{T}} - \frac{3J(\theta) + K(\theta)}{2\{I(\theta)\}^3\sqrt{T}} \tilde{Z}_1(\theta)^2 + o_p\left(\frac{1}{\sqrt{T}}\right).$$

Using fundamental properties of the periodogram (see Brillinger, 1969, 1975) it is not difficult to show that

$$E_\theta V_T = -\frac{B(\theta)}{I(\theta)\sqrt{T}} - \frac{J(\theta) + K(\theta)}{2\{I(\theta)\}^2\sqrt{T}} + o\left(\frac{1}{\sqrt{T}}\right), \tag{5.5}$$

$$\mathrm{Var}_\theta(V_T) = \{I(\theta)\}^{-1} + o\left(\frac{1}{\sqrt{T}}\right), \tag{5.6}$$

$$E_\theta\{V_T - E_\theta(V_T)\}^3 = -\frac{3J(\theta) + 2K(\theta)}{\{I(\theta)\}^3\sqrt{T}} + o\left(\frac{1}{\sqrt{T}}\right). \tag{5.7}$$

Using (4.11), if we put

$$\hat{\theta}^*_{qML} = \hat{\theta}_{qML} + \frac{B(\hat{\theta}_{qML})}{TI(\hat{\theta}_{qML})} + \frac{K(\hat{\theta}_{qML})}{6T\{I(\hat{\theta}_{qML})^2\}} = \hat{\theta}_{qML} + \frac{B(\theta)}{TI(\theta)} + \frac{K(\theta)}{6TI\{I(\theta)\}^2} + o_p\left(\frac{1}{T}\right), \tag{5.8}$$

then we obtain

$$P^T_\theta\{\sqrt{T}(\hat{\theta}^*_{qML} - \theta) \leq 0\} = ½ + o(1/\sqrt{T}),$$

i.e., $\hat{\theta}_{qML}$ is second order AMU, and

$$P^T_\theta\{\sqrt{T}(\hat{\theta}^*_{qML} - \theta) \leq x\} = \Phi(x\sqrt{I(\theta)}) + \frac{x^2}{6\sqrt{TI(\theta)}}\{3J(\theta) + 2K(\theta)\}\phi(x\sqrt{I(\theta)}) + o\left(\frac{1}{\sqrt{T}}\right),$$

which coincides with the bound distribution of Theorem 2. Thus we have

THEOREM 6. *The modified quasi-maximum likelihood estimator $\hat{\theta}^*_{qML}$ is second order asymptotically efficient in the sense of Akahira and Takeuchi.*

6. Calculations of $I(\theta)$, $J(\theta)$, $K(\theta)$ and $B(\theta)$. In this section we shall calculate $I(\theta)$, $J(\theta)$, $K(\theta)$ and $B(\theta)$ for various rational spectra. This enables us to present the second order AMU for these spectra. Of course we can evaluate the asymptotic bias for various estimators.

CASE 1. Consider the ARMA(p, q) spectral density

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi}\frac{|\sum_{j=0}^q \alpha_j e^{ij\lambda}|^2}{|\sum_{j=0}^p \beta_j e^{ij\lambda}|^2}.$$

Suppose that σ^2 is unknown (i.e., $\theta = \sigma^2$), and that $\alpha_0, \cdots, \alpha_q, \beta_0, \cdots, \beta_p$ are known. Then it is easy to show

$$I(\sigma^2) = \frac{1}{2\sigma^4}, \qquad K(\sigma^2) = \frac{1}{\sigma^6}, \qquad J(\sigma^2) = -\frac{1}{\sigma^6}.$$

Let $\hat{\sigma}^2_{ML}$ be the exact maximum likelihood estimator of σ^2. Then we can see that

$$\hat{\sigma}^{2*}_{ML} = \hat{\sigma}^2_{ML} + \frac{K(\hat{\sigma}^2_{ML})}{6T\{I(\hat{\sigma}^2_{ML})\}^2} = \left(1 + \frac{2}{3T}\right)\hat{\sigma}^2_{ML}$$

is second order median unbiased and efficient. Remembering (4.15) we have

$$E_\theta\hat{\sigma}^2_{ML} = \sigma^2 + o(T^{-1}).$$

CASE 2. Consider the following ARMA(p, q) spectral density

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi}\frac{\prod_{k=1}^q (1 - \psi_k e^{i\lambda})(1 - \psi_k e^{-i\lambda})}{\prod_{k=1}^p (1 - \rho_k e^{i\lambda})(1 - \rho_k e^{-i\lambda})}, \tag{6.1}$$

where $\psi_1, \cdots, \psi_q, \rho_1, \cdots, \rho_p$ are real numbers such that $|\psi_j| < 1, j = 1, \cdots, q, |\rho_j| < 1, j = 1, \cdots, p$. Suppose that ψ_m is an unknown parameter (i.e., $\theta = \psi_m$), and that $\rho_1, \cdots, \rho_p, \psi_1, \cdots, \psi_{m-1}, \psi_{m+1}, \cdots, \psi_q$ are known parameters. Then it is easy to show

$$\frac{\partial f_\theta(\lambda)}{\partial\theta}\{f_\theta(\lambda)\}^{-1} = \frac{-z^2 + 2\psi_m z - 1}{(1 - \psi_m z)(z - \psi_m)},$$

where $z = e^{i\lambda}$. In this case

$$I(\psi_m) = \frac{1}{4\pi i}\int_{|z|=1} \frac{(-z^2 + 2\psi_m z - 1)^2}{(1 - \psi_m z)^2(z - \psi_m)^2 z}\, dz, \tag{6.2}$$

$$K(\psi_m) = \frac{1}{2\pi i}\int_{|z|=1} \frac{(-z^2 + 2\psi_m z - 1)^3}{(1 - \psi_m z)^3(z - \psi_m)^2 z}\, dz. \tag{6.3}$$

To evaluate the above integrals we present the residue theorem (e.g., Hille, 1959).

THEOREM 7. *Suppose that $F(z)$ is holomorphic inside and on a "scroc" C, save for a finite number of isolated singularities, $a_1, \cdots, a_r$, none of which lie on C. Then*

$$\int_C F(z)\, dz = 2\pi i \textstyle\sum_{j=1}^{r} \mathrm{Res}(j),$$

where $\mathrm{Res}(j)$ *is the residue of $F(z)$ at a_j. Also if a_j is a pole of order s, then the required residue is given by*

$$\mathrm{Res}(j) = \frac{1}{(s-1)!}\left\{\frac{d^{s-1}}{dz^{s-1}}(z - a_j)^s F(z)\right\}_{z=a_j}.$$

Using this theorem we have $I(\psi_m) = 1/(1 - \psi_m^2)$ and $K(\psi_m) = -6\psi_m/(1 - \psi_m^2)^2$. Let $\hat{\psi}_{m,ML}$ be the exact maximum likelihood estimator of ψ_m. Then we can see that

$$\hat{\psi}^*_{m,ML} = \left(1 - \frac{1}{T}\right)\hat{\psi}_{m,ML}$$

is second order median unbiased and efficient. Similarly we can show

$$J(\psi_m) = 4\psi_m/(1 - \psi_m^2)^2, \quad \text{and} \quad E_\theta\hat{\psi}_{m,ML} = \psi_m + \psi_m/T + o(T^{-1}).$$

CASE 3. We also deal with the rational spectral density (6.1). We assume that ρ_m is an unknown parameter (i.e., $\theta = \rho_m$), and that $\psi_1, \cdots, \psi_q, \rho_1, \cdots, \rho_{m-1}, \rho_{m+1}, \cdots, \rho_p$ are known parameters. Then we have

$$I(\rho_m) = \frac{1}{1 - \rho_m^2}, \qquad K(\rho_m) = \frac{6\rho_m}{(1 - \rho_m^2)^2}, \qquad J(\rho_m) = \frac{-2\rho_m}{(1 - \rho_m^2)^2}.$$

Let $\hat{\rho}_{m,ML}$ be the exact maximum likelihood estimator of ρ_m. Then we can see that

$$\hat{\rho}^*_{m,ML} = \left(1 + \frac{1}{T}\right)\hat{\rho}_{m,ML}$$

is second order median unbiased and efficient. Also we have

$$E_\theta\hat{\rho}_{m,ML} = \rho_m - 2\rho_m/T + o(T^{-1}).$$

Henceforth we shall consider the quasi-maximum likelihood estimation. Since the evaluation of $B(\theta)$ for general rational spectral density such as (6.1) is very complicated, we shall confine ourselves to ARMA(1, 1) spectra. Hereafter we shall consider the following ARMA(1, 1) spectral density.

$$f_\theta(\lambda) = \frac{\sigma^2}{2\pi}\frac{|1 - \psi e^{i\lambda}|^2}{|1 - \rho e^{i\lambda}|^2}, \tag{6.4}$$

where $|\psi| < 1$, $|\rho| < 1$, $\psi \neq \rho$. Then we can show that

$$\gamma(n) = \frac{\sigma^2(1 - \psi\rho)(\rho - \psi)}{(1 - \rho^2)}\rho^{n-1}, \qquad n \geq 1,$$

$$b_\theta(\lambda) = \frac{\sigma^2}{2\pi}\frac{(1 - \psi\rho)(\rho - \psi)}{(1 - \rho^2)}\frac{z\{(z^2 + 1) - 4\rho z + \rho^2(z^2 + 1)\}}{(1 - \rho z)^2(z - \rho)^2}.$$

CASE 4. Suppose that σ^2 is an unknown parameter (i.e., $\theta = \sigma^2$), and that ψ and ρ are known parameters. We can show

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

Let $\hat{\sigma}^2_{qML}$ be the quasi-maximum likelihood estimator of σ^2. Then

$$\hat{\sigma}^{2*}_{qML} = \hat{\sigma}^2_{qML} + \frac{2}{3T}\hat{\sigma}^2_{qML} - \frac{2}{T}\frac{(\rho-\psi)^2}{(1-\rho^2)(1-\psi^2)}\hat{\sigma}^2_{qML} \tag{6.5}$$

is second order asymptotically median unbiased and efficient. By (5.5), we have

$$E_\theta \hat{\sigma}^2_{qML} = \sigma^2 + \frac{2}{T}\frac{(\rho-\psi)^2\sigma^2}{(1-\rho^2)(1-\psi^2)} + o(T^{-1}).$$

CASE 5. In the model (6.4), suppose that ρ is an unknown parameter (i.e., $\theta = \rho$), and that σ^2 and ψ are known parameters. Then it is not difficult to show

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

Let $\hat{\rho}_{qML}$ be the quasi-maximum likelihood estimator of ρ. Then

$$\hat{\rho}^*_{qML} = \hat{\rho}_{qML} + \frac{1}{T}\frac{(\hat{\rho}_{qML}-\psi)(1-2\hat{\rho}_{qML}\cdot\psi+\psi^2)}{(1-\hat{\rho}_{qML}\cdot\psi)(1-\psi^2)} + \frac{1}{T}\hat{\rho}_{qML} \tag{6.6}$$

is second order asymptotically median unbiased and efficient. By (5.5) we have

$$E_\theta \hat{\rho}_{qML} = \rho - \frac{(\rho-\psi)(1-2\rho\psi+\psi^2)}{T(1-\rho\psi)(1-\psi^2)} - \frac{2\rho}{T} + o(T^{-1}). \tag{6.7}$$

Consider the case $\psi = 0$ (i.e., our model is an autoregressive model of order 1), then (6.6) and (6.7) become

$$\hat{\rho}^*_{qML} = \left(1 + \frac{2}{T}\right)\hat{\rho}_{qML}, \quad E_\theta\hat{\rho}_{qML} = \rho - \frac{3\rho}{T} + o(T^{-1}),$$

respectively. By the way, in the case of $\psi = 0$, we can see that $\hat{\rho}_{qML}$ is asymptotically equivalent to the Yule-Walker estimator;

$$\left(\textstyle\sum_{t=1}^{T-1} X_t X_{t+1}\right)/\left(\sum_{t=1}^{T} X_t^2\right),$$

neglecting the terms of order $O_p(\rho^T)$, which do not disturb our asymptotic theory.

CASE 6. In the model (6.4), suppose that ψ is an unknown parameter (i.e., $\theta = \psi$), and that σ^2 and ρ are unknown parameters. It is not so hard to show

$$B(\psi) = \frac{(\psi-\rho)(1+\psi^2-2\psi\rho-\rho^2+3\psi^2\rho^2-2\psi^3\rho)}{(1-\psi^2)(1-\psi\rho)(1-\rho^2)}. \tag{6.8}$$

Let $\hat{\psi}_{qML}$ be the quasi-maximum likelihood estimator of ψ. Then

$$\hat{\psi}^*_{qML} = \hat{\psi}_{qML} - \frac{1}{T}\hat{\psi}_{qML} + \frac{1}{T}\frac{(\hat{\psi}_{qML}-\rho)(1+\hat{\psi}^2_{qML}-2\hat{\psi}_{qML}\cdot\rho-\rho^2+3\hat{\psi}^2_{qML}\cdot\rho^2-2\hat{\psi}^3_{qML}\cdot\rho)}{(1-\hat{\psi}^2_{qML})(1-\hat{\psi}_{qML}\cdot\rho)(1-\rho^2)} \tag{6.9}$$

is second order median unbiased and efficient. Consider the case $\rho = 0$ (i.e., our model is a moving average model of order 1), then (6.8) and (6.9) become

$$B(\psi) = \frac{(1+\psi^2)\psi}{(1-\psi^2)^2}, \quad \hat{\psi}^*_{qML} = \hat{\psi}_{qML} + \frac{1}{T}\frac{2\hat{\psi}^3_{qML}}{(1-\hat{\psi}^2_{qML})},$$

respectively. Also, in the case of $\rho = 0$, we have

$$E_\theta \hat{\psi}_{qML} = \psi - \frac{1}{T} \frac{2\psi^3}{(1-\psi^2)} + o(T^{-1}).$$

REFERENCES

AKAHIRA, M. and TAKEUCHI, K. (1981). *Asymptotic Efficiency of Statistical Estimators: Concepts and Higher Order Asymptotic Efficiency.* Springer Lecture Notes in Statistics, No. 7. Springer-Verlag, New York.

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

BRILLINGER, D. R. (1975). *Time Series; Data Analysis and Theory.* Holt, New York.

GRENANDER, U. and SZEGÖ, G. (1958). *Toeplitz Forms and Their Applications.* Berkeley, Univ. of California Press.

HANNAN, E. J. (1970). *Multiple Time Series.* Wiley, New York.

HILLE, E. (1959). *Analytic Function Theory, Vol.* I. Ginn, Lexington, Massachusetts.

HOSOYA, Y. (1979). High-order efficiency in the estimation of linear processes. *Ann. Statist.* **7** 516–530.

HOSOYA, Y. (1980). Conditionality and maximum-likelihood estimation. *Recent Developments in Statistical Inference and Data Analysis*; K. Matusita, editor, 117–126. North-Holland, Amsterdam.

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

SHAMAN, P. (1976). Approximations for stationary covariance matrices and their inverses with application to ARMA models. *Ann. Statist.* **4** 292–301.

TAKEUCHI, K. (1981). High-order efficiency in estimation of stationary time series. *Proc. Symp. Res. Inst. Math. Sci.*, Kyoto Univ., **418** 41–47.

DEPARTMENT OF APPLIED MATHEMATICS
FACULTY OF ENGINEERING
HIROSHIMA UNVERSITY
SHITAMI, SAIJOCHO
HIGASHIHIROSHIMA, 724 JAPAN.