The Annals of Statistics
1981, Vol. 9, No. 2, 300–306

AN OPTIMAL AUTOREGRESSIVE SPECTRAL ESTIMATE

BY RITEI SHIBATA

Tokyo Institute of Technology

An asymptotic lower bound is obtained for the integrated relative squared error of autoregressive spectral estimate when the order of autoregression is selected. The bound is attained in the limit by the same selection as has been proposed for prediction.

1. Introduction. In a recent paper (Shibata, 1980), the author, assuming data come from an infinite order autoregressive process, has proposed an asymptotically efficient selection of the order of an autoregressive model for estimating parameters of the process. The proposed selection attains a lower bound in the limit for the mean squared error of the estimated predictor. The purpose of the present paper is to apply the result to an autoregressive spectral estimate $\hat{f}_k(\lambda)$ of a spectral density $f(\lambda)$ obtained by a kth order autoregressive model fitting.

Consider a weakly stationary process $\{x_t, t = \cdots, -1, 0, 1, \cdots\}$ with mean 0, which satisfies the equation

$$\sum_{l=0}^{\infty} a_l x_{t-l} = e_t \tag{1.1}$$

where $a_0 = 1$, $\sum_l a_l^2 < \infty$ and $\{e_t\}$ is a sequence of i.i.d. random variables. We suppose that the process $\{x_t\}$ has the spectral density

$$f(\lambda) = \sigma^2 / |A(e^{2\pi i\lambda})|^2, \tag{1.2}$$

where $A(e^{2\pi i\lambda})$ is a boundary function of $A(z) = 1 + a_1 z + a_2 z^2 + \cdots$.

Given observations $x_1, x_2, \cdots, x_n$ from the process $\{x_t\}$, we have the least squares estimate $\hat{a}(k)' = (\hat{a}_1(k), \hat{a}_2(k), \cdots, \hat{a}_k(k))$ of the autoregressive parameters by fitting a kth order autoregressive model

$$\sum_{l=0}^{k} \alpha_l x_{t-l} = \epsilon_t.$$

Let $\{K_n\}$ be a sequence of positive integers and $N = n - K_n$. The estimate $\hat{a}(k)$ is a solution of the following linear equation with $K_n < n$ initial conditions

$$\hat{R}(k)\hat{a}(k) = -\hat{r}(k),$$

where

$$\hat{R}(k) = (\hat{r}_{lm}, 1 \leqq l, m \leqq k), \quad \hat{r}(k)' = (\hat{r}_{10}, \hat{r}_{20}, \cdots, \hat{r}_{k0}),$$
$$\hat{r}_{lm} = \sum_{t=K_n}^{n-1} x_{t+1-l} x_{t+1-m} / N$$

and $k \leqq K_n$.

By using this estimate and an estimate of σ^2, namely

$$\hat{\sigma}_k^2 = \sum_{t=K_n}^{n-1} (x_{t+1} + \hat{a}_1(k) x_t + \cdots + \hat{a}_k(k) x_{t+1-k})^2 / N,$$

we obtain an autoregressive spectral estimate

$$\hat{f}_k(\lambda) = \hat{\sigma}_k^2 / |\hat{A}_k(e^{2\pi i\lambda})|^2,$$

Received November 1978; revised December, 1979.
AMS 1970 *subject classifications.* Primary 62M15; secondary 62M10.
Key words and phrases. Spectral estimate, autoregression, model selection.

where

$$\hat{A}_k(z) = 1 + \hat{a}_1(k)z + \hat{a}_2(k)z^2 + \cdots + \hat{a}_k(k)z^k.$$

Following Akaike (1969, 1970), as the loss function of a spectral estimate $\hat{f}_k(\lambda)$ we will adopt the integrated relative squared error

$$\int \left(\frac{\hat{f}_k(\lambda) - f(\lambda)}{f(\lambda)}\right)^2 d\lambda, \tag{1.3}$$

where the integrals throughout are taken from $-\frac{1}{2}$ to $\frac{1}{2}$.

If k is selected from the given range $1 \leqq k \leqq K_n$, an asymptotic lower bound for this error is given by the following.

Define the norm

$$\|\alpha\|_C = (\textstyle\sum_{l,m} \alpha_l \alpha_m C_{lm})^{1/2}$$

for any vector α, where $C = (C_{lm})$ is a finite or infinite dimensional positive definite matrix. Let $a(k)' = (a_1(k), a_2(k), \cdots, a_k(k), 0, \cdots)$ be the projection of $a' = (a_1, a_2, \cdots)$ on the space

$$V(k) = \{\alpha;\ \alpha' = (\alpha_1, \alpha_2, \cdots, \alpha_k, 0, \cdots)\}$$

with the norm $\|\cdot\|_R$, where $R = (r_{lm}, 1 \leqq l, m < \infty)$ and $r_{lm} = E(x_{t-l}x_{t-m})$. Define a function of k

$$L_n(k) = \|a - a(k)\|_R^2 + k\sigma^2/N$$
$$= \sigma_k^2 - \sigma^2 + k\sigma^2/N,$$

where

$$\sigma_k^2 = E(x_{t+1} + a_1(k)x_t + \cdots + a_k(k)x_{t+1-k})^2.$$

Then let k_n^* be the k which minimizes $L_n(k)$ in $1 \leqq k \leqq K_n$. In Theorem 2.2, we obtain the lower bound $2L_n(k_n^*)/\sigma^2$ for the integrated relative squared error (1.3). It is also shown that this bound is attained by the selection which minimizes the statistic

$$S_n(k) = (N + 2k)\,\hat{\sigma}_k^2.$$

2. An optimal autoregressive spectral estimate.

We shall make use of the following assumptions on the process $\{x_t\}$ and the sequence $\{K_n\}$, which are the same as in Shibata (1980).

ASSUMPTIONS.

(A1) $\{x_t\}$ is a stationary Gaussian process of the form (1.1) and $|a| = \sum_{l=0}^{\infty} |a_l| < \infty$.

(A2) $A(z)$ is nonzero for $|z| \leqq 1$.

(A3) The order k is selected from a given range $1 \leqq k \leqq K_n$, where $K_n \to \infty$ and $K_n/n^{1/2} \to 0$ as $n \to \infty$.

(A4) $\{x_t\}$ is not degenerate to a finite order autoregressive process.

To obtain the lower bound, we need the following lemmas on asymptotics of the estimates $\hat{\sigma}^2(k)$ and $\hat{a}(k)$. In Theorem 2.1 it is shown that the integrated relative squared error is asymptotically equivalent to $L_n(k)$ if k diverges to infinity as $n \to \infty$. From this, the main Theorem 2.2 is easily derived.

LEMMA 2.1. *Assume* (A1) *to* (A3). *Then* $\hat{\sigma}_k^2 - \sigma_k^2$ *converges to zero in probability uniformly in* $1 \leqq k \leqq K_n$, *and for any divergent sequence* $\{k_n\}$, $(\hat{\sigma}_k^2 - \sigma^2)/(L_n(k))^{1/2}$ *converges to zero uniformly in* $k_n \leqq k \leqq K_n$.

PROOF. Putting

$$s_k^2 = \sum_{t=K_n}^{n-1} (x_{t+1} + a_1(k)x_t + \cdots + a_k(k)x_{t+1-k})^2/N$$

we have

$$|\hat{\sigma}_k^2 - \sigma_k^2| \leqq \|\hat{a}(k) - a(k)\|_{\hat{R}(k)} + |s_k^2 - \sigma_k^2|. \tag{2.1}$$

By Proposition 3.2 and Lemma 3.3 of Shibata (1980), we see the first term of the right-hand side of (2.1) is written as $((k/N)\sigma^2 + L_n(k)o_p(1))^{1/2}$ uniformly in k, which then converges to zero in probability from Assumption (A3) and the boundedness of $L_n(k)$. The second term is rewritten as

$$s_k^2 - \sigma_k^2 = (\hat{r}_{00} - r_0) + 2a(k)'(\hat{r}(k) - r(k)) + a(k)'(\hat{R}(k) - R(k))a(k),$$

where $R(k) = (r_{l-m}, 1 \leqq l, m \leqq k)$ and $r(k)' = (r_1, \cdots, r_k)$. By Lemma 4.2 of Shibata (1980), the fourth moment of each term of the right-hand side is bounded by some constant times $1/N^2$ uniformly in $1 \leqq k \leqq K_n$, so that the first assertion is established. In $k_n \leqq k \leqq K_n$, $L_n(k)$ uniformly converges to zero as $n \to \infty$, so that

$$(\|\hat{a}(k) - a(k)\|_{\hat{R}(k)} + \sigma_k^2 - \sigma^2)/(L_n(k))^{1/2}$$

converges to zero in probability uniformly in $k_n \leqq k \leqq K_n$. On the other hand

$$\sum_{k=k_n}^{K_n} 1/(NL_n(k))^2 \leqq \sum_{k=k_n}^{K_n} 1/(k\sigma^2)^2,$$

so that the sum from k equals k_n to K_n of the fourth moments of

$$(s_k^2 - \sigma_k^2)/(L_n(k))^{1/2}$$

converges to zero. This proves the second assertion.

LEMMA 2.2 *Assume* (A1) *to* (A3). *Then*

$$|\hat{a}(k) - a(k)|$$

converges to zero in probability uniformly in $1 \leqq k \leqq K_n$.

PROOF. By using the Euclidean norm $\|\cdot\|$, we have

$$\begin{aligned} |\hat{a}(k) - a(k)| &\leqq k^{1/2} \|\hat{a}(k) - a(k)\| \\ &\leqq k^{1/2} \|R(k)^{-1}\|^{1/2} \|\hat{a}(k) - a(k)\|_{R(k)}, \end{aligned}$$

where $R(k)^{-1}$ is the inverse matrix of $R(k)$, whose operator norm $\|R(k)^{-1}\|$ is bounded (see Berk (1974)). Furthermore

$$k\|\hat{a}(k) - a(k)\|_{R(k)}^2 = \frac{k^2}{N}\left\{\left(\frac{N}{k}\|\hat{a}(k) - a(k)\|_{R(k)}^2 - \sigma^2\right) + \sigma^2\right\}. \tag{2.2}$$

Applying Lemmas 3.3 and 3.4 of Shibata (1980), we see that the right-hand side of (2.2) uniformly converges to zero as $n \to \infty$ in probability. Thus the desired result is obtained.

If we define

$$\tilde{r}_{l,k} = \int e^{2\pi i l\lambda} \hat{f}_k(\lambda)\,d\lambda,$$

then the behavior of the estimated covariance matrix

$$\tilde{R}_k = (\tilde{r}_{l-m,k}, 1 \leqq l, m < \infty)$$

is given in the following lemma.

LEMMA 2.3. *Assume* (A1) *to* (A3). *Then for any divergent sequence* $\{k_n\}$, $\|\tilde{R}_k - R\|$ *converges to zero in probability uniformly in* $k_n \leqq k \leqq K_n$, *where* $\|\cdot\|$ *denotes the operator norm of the matrix.*

PROOF. From the definition of $\tilde{R}_k$, omitting the arguments of $A(e^{2\pi i\lambda})$ and $\hat{A}_k(e^{2\pi i\lambda})$ for simplicity, we have

$$\|\tilde{R}_k - R\| \leqq 2 \max_\lambda |\hat{f}_k(\lambda) - f(\lambda)|$$
$$\leqq 2 \max_\lambda \frac{||A|^2 - |\hat{A}_k|^2|}{|\hat{A}_k|^2 |A|^2} \hat{\sigma}_k^2 + 2 \frac{1}{|A|^2} |\hat{\sigma}_k^2 - \sigma^2|.$$

Now

$$||\hat{A}_k|^2 - |A|^2| \leqq \{|\hat{a}(k) - a(k)| + |a(k) - a|\}^2. \tag{2.3}$$

By Lemma 4 of Berk (1974) and Lemma 2.2, the right-hand side of (2.3) converges to zero as $n \to \infty$ uniformly in $k_n \leqq k \leqq K_n$. The result follows from Lemma 2.1.

THEOREM 2.1. *Assume* (A1) *to* (A3). *Then for any divergent sequence* $\{k_n\}$,

$$\int \left(\frac{\hat{f}_k(\lambda) - f(\lambda)}{f(\lambda)}\right)^2 d\lambda \Big/ L_n(k)$$

converges to $2/\sigma^2$ *as* $n \to \infty$ *in probability uniformly in* $k_n \leqq k \leqq K_n$.

PROOF. By Lemma 2.3 and an application of Proposition 3.2 of Shibata (1980),

$$\|a - \hat{a}(k)\|^2_{\tilde{R}_k}/L_n(k)$$

converges to 1 in probability uniformly in $1 \leqq k \leqq K_n$, where $\hat{a}(k)$ is considered as an infinite dimensional vector with undefined entries zero. On the other hand, $\hat{f}_k(\lambda)$ converges to $f(\lambda)$ in probability as $n \to \infty$ uniformly in $k_n \leqq k \leqq K_n$ and $-\frac{1}{2} \leqq \lambda < \frac{1}{2}$, so that $\hat{F}_k = \max_\lambda \hat{f}_k(\lambda)$ is stochastically bounded uniformly in $k_n \leqq k \leqq K_n$. Putting $f^*(\lambda) = \hat{f}_k(\lambda)$ in the following lemma, completes the proof.

LEMMA 2.4. *Consider another spectral density*

$$f^*(\lambda) = \sigma^{*2}/|A^*(e^{2\pi i\lambda})|^2$$

which induces the covariance matrix

$$R^* = (r^*_{l-m}, 1 \leqq l, m < \infty).$$

Assume that $|a^*| < \infty$ *and* $A^*(z)$ *is nonzero in* $|z| \leqq 1$. *If* $|a - a^*| < \infty$ *and both* σ^2 *and* σ^{*2} *are positive, then*

$$\left| \int \left(\frac{f(\lambda) - f^*(\lambda)}{f(\lambda)}\right)^2 d\lambda - 2\frac{\|a - a^*\|^2_{R^*}}{\sigma^{*2}} \right|$$
$$\leqq \left(\frac{\Delta\sigma^2}{\sigma^2}\right)^2 + \frac{\|a - a^*\|^2_{R^*}}{\sigma^4}\left(|a - a^*|^2 F^* + 4|a - a^*| (F^*)^{12}\sigma^* + 2|\Delta\sigma^2| \left(2 + \frac{\sigma^2}{\sigma^{*2}}\right)\right).$$

Here $F^* = \max_\lambda f^*(\lambda)$, $\Delta\sigma^2 = \sigma^2 - \sigma^{*2}$, *and* $a^{*\prime} = (a_1^*, a_2^*, \cdots)$ *is the vector of the coefficients of* A^*.

The proof of this lemma is placed in the Appendix. To see the behavior of the loss function (1.3) when $1 \leqq k \leqq k_n$, define

$$A_k(z) = 1 + a_1(k)z + \cdots + a_k(k)z^k \quad \text{and} \quad f_k(\lambda) = \sigma_k^2/|A_k(e^{2\pi i\lambda})|^2.$$

PROPOSITION 2.1. *Assume* (A1) *to* (A3). *If* $\{k_n\}$ *is a divergent sequence of integers such that*

$$\frac{N}{k}\int\left(\frac{f(\lambda)-f_k(\lambda)}{f_k(\lambda)}\right)^2 d\lambda$$

diverges to infinity as $n\to\infty$ *uniformly in* $1\leqq k\leqq k_n$, *then*

$$\int\left(\frac{f(\lambda)-\hat{f}_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2 d\lambda\bigg/\int\left(\frac{f(\lambda)-f_k(\lambda)}{f_k(\lambda)}\right)^2 d\lambda$$

converges to 1 *in probability uniformly in* $1\leqq k\leqq k_n$.

PROOF. By Schwartz inequality we have

$$\begin{aligned}\left|\int\left(\frac{f(\lambda)-\hat{f}_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2 d\lambda-\int\left(\frac{f(\lambda)-f_k(\lambda)}{f_k(\lambda)}\right)^2 d\lambda\right|\\ \leqq\left|\int\left(\frac{f(\lambda)-f_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2 d\lambda-\int\left(\frac{f(\lambda)-f_k(\lambda)}{f_k(\lambda)}\right)^2 d\lambda\right|\\ +\left|\int\left(\frac{f_k(\lambda)-\hat{f}_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2 d\lambda\right|\\ +2\left(\int\left(\frac{f_k(\lambda)-\hat{f}_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2 d\lambda\int\left(\frac{f(\lambda)-f_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2 d\lambda\right)^{1/2}\end{aligned}\tag{2.5}$$

From the uniform convergence of $\hat{f}_k(\lambda)$ to $f_k(\lambda)$, it suffices to show that

$$\int\left(\frac{f_k(\lambda)-\hat{f}_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2 d\lambda\bigg/\int\left(\frac{f(\lambda)-f_k(\lambda)}{f_k(\lambda)}\right)^2 d\lambda$$

converges to zero in probability uniformly in $1\leqq k\leqq k_n$. As was shown by Whittle (1963), if $\sigma_k^2\neq 0$, then $A_k(z)\neq 0$ in $|z|\leqq 1$. Thus for any $\epsilon>0$,

$$\lim_{n\to\infty}P(|\hat{A}_k(z)|>\epsilon,\ |z|\leqq 1)=1.$$

We can then apply Lemma 2.4 to $\hat{f}_k(\lambda)$ and $f_k(\lambda)$. From the choice of $\{k_n\}$ and Lemma 3.4 of Shibata (1980), we have the desired result.

LEMMA 2.5. *Assume* (A1) *to* (A3). *Then*

$$\int\left(\frac{f(\lambda)-\hat{f}_k(\lambda)}{f(\lambda)}\right)^2 d\lambda\bigg/\int\left(\frac{f(\lambda)-\hat{f}_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2 d\lambda$$

is stochastically bounded away from zero uniformly in $1\leqq k\leqq K_n$.

PROOF. From the inequality

$$\left(\frac{f(\lambda)-\hat{f}_k(\lambda)}{f(\lambda)}\right)^2\geqq\left(\frac{f(\lambda)-\hat{f}_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2\bigg/\left(1+\left(\frac{f(\lambda)-\hat{f}_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2\right)$$

we may show that

$$\frac{f(\lambda)-\hat{f}_k(\lambda)}{\hat{f}_k(\lambda)}$$

is stochastically bounded uniformly in $1 \leqq k \leqq K_n$ and $-\frac{1}{2} \leqq \lambda < \frac{1}{2}$. Here

$$\begin{aligned}\left(\frac{f(\lambda)-\hat{f}_k(\lambda)}{\hat{f}_k(\lambda)}\right)^2 &\leqq 2\left(1+\left(\frac{f(\lambda)}{\hat{f}_k(\lambda)}\right)^2\right)\\ &\leqq 2(1+\{\max_\lambda f(\lambda)\ \max_\lambda(1/\hat{f}_k(\lambda))\}^2)\\ &\leqq 2(1+|r|^2|\hat{a}(k)|^4/\hat{\sigma}_k^4).\end{aligned} \tag{2.6}$$

As was shown by Berk (1974), $|a(k)|$ is bounded in k, and the vector r of serial correlations has a finite absolute norm $|\ r\ |$, so that, by Lemma 2.2 the right-hand side of (2.6) is stochastically bounded uniformly in $1 \leqq k \leqq K_n$. Thus the proof is complete.

Let k_n^* be the k which attains the minimum of $L_n(k)$ in $1 \leqq k \leqq K_n$. The assumption (A4) then indicates that k_n^* diverges to infinity as $n \to \infty$, at the same time, $L_n(k_n^*)$ goes to zero. The following main theorem shows that the selection $\hat{k}$ which minimizes the statistic

$$S_n(k) = (N + 2k)\hat{\sigma}_k^2$$

is an optimal one.

THEOREM 2.2. *Assume* (A1) *to* (A4). *Then for any order selection* $\tilde{k}$ *which is a random variable possibly depending on the observations* $x_1, x_2, \cdots x_n$, *and for any* $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(\int\left(\frac{\hat{f}_{\tilde{k}}(\lambda)-f(\lambda)}{f(\lambda)}\right)^2 d\lambda \Big/ L_n(k_n^*) \geqq \frac{2}{\sigma^2} - \epsilon\right) = 1$$

and $\hat{k}$ *attains the lower bound in the limit, that is,*

$$\int\left(\frac{\hat{f}_{\hat{k}}(\lambda)-f(\lambda)}{f(\lambda)}\right)^2 d\lambda \Big/ L_n(k_n^*)$$

converges to $2/\sigma^2$ *in probability as* $n \to \infty$.

PROOF. We can choose a divergent sequence of integers $\{k_n^{**}\}$ such that

$$\frac{N}{k}\int\left(\frac{f(\lambda)-f_k(\lambda)}{f_k(\lambda)}\right)^2 d\lambda$$

and

$$\int\left(\frac{f(\lambda)-f_k(\lambda)}{f_k(\lambda)}\right)^2 d\lambda \Big/ L_n(k_n^*)$$

diverge to infinity as $n \to \infty$ uniformly in $1 \leqq k \leqq k_n^{**}$. The first part of the theorem follows from Theorem 2.1 if $\tilde{k} > k_n^{**}$. Otherwise it follows from Proposition 2.1 and Lemma 2.5. Theorem 2.1 together with the proof of Theorem 4.1 of Shibata (1980) implies the last part, since $\hat{k}$ diverges to infinity as $n \to \infty$ in probability.

As was shown by Shibata (1980), Akaike's AIC method (Akaike, 1973) is asymptotically equivalent to our method. Therefore its optimality was also established in the case of autoregressive spectral estimate.

Acknowledgment. The author wishes to thank an associate editor and a referee for their helpful comments.

APPENDIX

PROOF OF LEMMA 2.4. Putting $\Delta|A|^2 = A^*\overline{(A - A^*)} + \overline{A^*}(A - A^*)$, we have

$$\frac{f(\lambda)-f^*(\lambda)}{f(\lambda)} = \frac{\Delta\sigma^2}{\sigma^2} - \frac{\Delta|A|^2}{|A^*|^2} - \frac{|A-A^*|^2}{|A^*|^2}\frac{\sigma^{*2}}{\sigma^2} + \frac{\Delta|A|^2}{|A^*|^2}\frac{\Delta\sigma^2}{\sigma^2}.$$

By Wiener's theorem, the Taylor expansion

$$1/A^*(z) = B^*(z) = 1 + b_1^* z + b_2^* z^2 + \cdots$$

is absolutely convergent on the unit circle, that is, $|b^*| < \infty$, where $b^{*\prime} = (b_1^*, b_2^*, \cdots)$. Therefore

$$\int \left(\frac{A - A^*}{A^*}\right) d\lambda = \int \sum_{l=1}^{\infty} (a_l - a_l^*) \sum_{m=0}^{\infty} b_m^* e^{2\pi i(l+m)\lambda}\, d\lambda$$

$$= \sum_{l=1}^{\infty} \sum_{m=0}^{\infty} (a_l - a_l^*) b_m^* \int e^{2\pi i(l+m)\lambda}\, d\lambda = 0.$$

Thus

$$\int \frac{\Delta |A|^2}{|A^*|^2}\, d\lambda = 2 \int \mathrm{Re}\left(\frac{A - A^*}{A^*}\right) d\lambda = 0.$$

By the same way,

$$\int \left(\frac{A - A^*}{A^*}\right)^2 d\lambda = 0.$$

On the other hand

$$\int \left(\frac{\Delta |A|^2}{|A^*|^2}\right)^2 d\lambda = 2 \int \frac{|A - A^*|^2}{|A^*|^2}\, d\lambda + \mathrm{Re} \int \left(\frac{A - A^*}{A^*}\right)^2 d\lambda$$

$$= (2/\sigma^{*2}) \int |A - A^*|^2 f^*(\lambda)\, d\lambda$$

$$= (2/\sigma^{*2}) \|a - a^*\|_{R^*}^2.$$

Combining these results, we have

$$\int \left(\frac{f(\lambda) - f^*(\lambda)}{f(\lambda)}\right)^2 d\lambda = \left(\frac{\Delta\sigma^2}{\sigma^2}\right)^2 + 2\frac{\|a - a^*\|_{R^*}^2}{\sigma^{*2}}$$

$$+ \frac{\sigma^{*4}}{\sigma^4} \int \left(\frac{|A - A^*|^2}{|A^*|^2} + \frac{2\Delta |A|^2}{|A^*|^2}\right) \frac{|A - A^*|^2}{|A^*|^2}\, d\lambda$$

$$+ \frac{2\|a - a^*\|_{R^*}^2}{\sigma^{*2}} \left\{\left(\frac{\Delta\sigma^2}{\sigma^2}\right)^2 - 2\frac{\Delta\sigma^2}{\sigma^2}\right\} - 2\frac{\Delta\sigma^2}{\sigma^4} \|a - a^*\|_{R^*}^2.$$

It is enough to note that

$$\left|\frac{\Delta |A|^2}{|A^*|^2}\right| = 2\left|\mathrm{Re}\left(\frac{A - A^*}{A^*}\right)\right| \leqq 2|a - a^*|(F^*)^{1/2}/\sigma^*.$$

REFERENCES

[1] Akaike, H. (1969). Power spectrum estimation through autoregressive model fitting. *Ann. Inst. Statist. Math.* **21** 407–419.

[2] Akaike, H. (1970). A fundamental relation between predictor identification and power spectrum estimation. *Ann. Inst. Statist. Math.* **22** 219–223.

[3] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second Internat. Symp. Information Theory.* (B. N. Petrov and F. Csáki, eds.) 267–281. Akadémia Kiado, Budapest.

[4] Akutowicz, E. J. (1957). On an explicit formula in linear least squares prediction. *Math. Scand.* **5** 261–266.

[5] Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.* **2** 489–502.

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

[7] Whittle, P. (1963). On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix. *Biometrika* **50** 129–134.

Department of Mathematics
Tokyo Institute of Technology
Oh-Okayama Meguro, Tokyo, Japan