数理解析研究所講究録
第 777 巻 1992 年 126-138

A NOTE ON MISSPECIFIED ARMA MODEL FITTINGS

Minoru Tanaka(田中 稔)
(Senshu University)
and
Mituaki Huzii(藤井 光昭)
(Tokyo Institute of Technology)

1. Introduction

For a finite dimensional parameter model fitting to a process, the conditional quasi-likelihood estimator can be used to estimate the unknown parameter (Huzii[2]). The estimator is usually given by the local maximal point of the likelihood function in a parameter space. If the fitted model is correct, statistical properties of the estimator have been studied in detail. On the other hand if the model is misspecified, they have not been much studied. Tanaka and Huzii[4] showed that there exist some misspecified MA(1) models whose conditional quasi-likelihood function has several local maximal points in a parameter space when a number of observations is large. To deal with cases of this kind it is worth to investigate whether or not its global maximizer is the best estimator of all the local maximal points.

In this paper, we focus on misspecified ARMA model fittings to a weakly stationary process, and examine a problem of estimating a spectral density of the process by using the fitting models. It will be shown by some concrete examples that the global maximal point of the likelihood function is not always superior to the local one for the estimation of the spectral density of the process when the number of observations is large.

2. ARMA model fittings

Let $\{Z_t; t=0,1,2,\cdots\}$ be a weakly stationary process with $EZ_t=0$. Then $\{Z_t\}$ is said to satisfy an autoregressive-moving average model of order p and q (or ARMA(p,q) model) if $\{Z_t\}$ is satisfies

$$(1+\phi_1 B+\cdots+\phi_p B^p)Z_t = (1-\theta_1 B-\cdots-\theta_q B^q)e_t, \tag{1}$$

where $\{e_t\}$ consists of independently and identically distributed random variables with $Ee_t=0$, $Ee_t^2=\sigma^2$, $\{\phi_i\}$ and $\{\theta_j\}$ are constants which are independent of t, and B is the usual backward shift operator defined by $B^i Z_t = Z_{t-i}$ $(i=0,1,2,\cdots)$. Here we put

$$\phi(B)=1+\phi_1 B+\cdots+\phi_p B^p, \qquad \theta(B)=1-\theta_1 B-\cdots-\theta_q B^q, \tag{2}$$

and assume that equations $\phi(x)=0$ and $\theta(x)=0$ have no common zeros and $\phi(x)\neq 0$ and $\theta(x)\neq 0$ for $|x|\leq 1$. In this case $\{Z_t\}$ has a spectral density

$$f_z(\lambda)= \sigma^2 \left|\frac{\theta(\exp(-2\pi i\lambda))}{\phi(\exp(-2\pi i\lambda))}\right|^2 , \quad -0.5\le\lambda\le 0.5. \qquad (3)$$

Let $\{X_t\}$ be a real-valued weakly stationary linear process with $EX_t=0$, $r_h= EX_tX_{t+h}$ and spectral density $f_x(\lambda)$, such that

$$X_t = \sum_{j=0}^{\infty} g_j e_{t-j}, \qquad (4)$$

where the g_j's are constants being independent of t, satisfy $\sum_{j=0}^{\infty} |g_j|<\infty$ and $g_x(x)= \sum_{j=0}^{\infty} g_j x^j \neq 0$ on $|x|=1$, and $Ee_t^{\,4}<\infty$.

We here consider the fitting of an ARMA(p,q) model to the process $\{X_t\}$. In this case $\{X_t\}$ may not satisfy the model. Let $\theta=(\phi_1,\cdots,\phi_p,\theta_1, \cdots,\theta_q)$ denote the unknown coefficient vector, and let the residual a_t and the sequence $\{F_k(\theta)\}$ be defined by

$$a_t = \theta^{-1}(B)\phi(B)X_t = \{ \sum_{k=0}^{t+p+q-1} F_k(\theta)B^k \}X_t, \qquad (5)$$

where $\{X_t; -(p+q)+1\le t\le 0\}$ should be replaced by suitable initial random variables (Huzii[2]). The notion of $F_k(\theta)$ in (5) can be extended for any k, $0\le k<\infty$. Then, by assuming Gaussian properties, the conditional quasi-maximum likelihood estimator

$\hat{\theta}_T$ based on observations $\{X_{-(p+q)+1}, \cdots, X_{T-1}, X_T\}$ is given by the value which minimizes

$$\hat{\sigma}^2(\theta) = \frac{1}{T} \sum_{t=1}^{T} \left| \{ \sum_{k=0}^{t+p+q-1} F_k(\theta) B^k \} X_t \right|^2 \tag{6}$$

with respect to θ in the parameter space,

$$\Omega_{p,q} = \{\theta \in R^{p+q};\ \phi(z)\theta(z) \neq 0 \text{ for } |z|<1,\ \phi(z)\neq 0 \text{ on } |z|=1$$
$$\text{and } \phi(\cdot),\ \theta(\cdot) \text{ have no common zeros}\}.$$

To obtain the estimator we usually search the parameter space for a local minimal point of $\hat{\sigma}^2(\theta)$. If there exist several local minimal points in the space, then the global one, $\hat{\theta}_T$, will be selected. It is questionable that this selection is appropriate for the estimation in the misspecified model fitting.

By the way, for evaluating the asymptotic properties of the conditional quasi-maximum likelihood estimator $\hat{\theta}_T$, our attention should be turned to a function,

$$S_{p,q}(\theta) = \sum_{h=0}^{\infty} \sum_{k=0}^{\infty} F_h(\theta) F_k(\theta) r_{h-k}, \tag{7}$$

for $\theta \in \Omega_{p,q}$ (see Huzii[2]). From the definition of $F_k(\theta)$ in (5) and using the spectral decomposition of the autocovariance function r_h, we can obtain another representation of (7)

$$S_{p,q}(\theta) = \int_{-0.5}^{0.5} f_x(\lambda) \left| \frac{\phi(\exp(-2\pi i\lambda))}{\theta(\exp(-2\pi i\lambda))} \right|^2 d\lambda \tag{8}$$

(see Åström and Söderström[1] and Kabaila[3]).

When $S_{p,q}(\theta)$ has at least two local minimal points θ_i^* in $\Omega_{p,q}$, there exist estimators $\hat{\theta}_T^i$ which converge in probability to these points as the sample size T tends to infinity (see Huzii[2], Theorem 2). In this case we have a question which one we should use for the estimator of the parameter. To discuss the problem, we shall examine the statistical properties of these estimators $\hat{\theta}_T^i$ by considering the spectral densities of the fitting model which are constructed by the estimators.

Let the estimator of the spectral density of the ARMA(p,q) model whose parameter θ is estimated by the $\hat{\theta}_T$ be defined as

$$\hat{f}_{p,q}(\lambda;\hat{\theta}_T) = \hat{\sigma}^2(\hat{\theta}_T) \left| \frac{\hat{\theta}(\exp(-2\pi i\lambda))}{\hat{\phi}(\exp(-2\pi i\lambda))} \right|^2 . \tag{9}$$

We assume that the estimator $\hat{\theta}_T$ converges in probability to θ^*, which is a local minimal point of $S_{p,q}(\theta)$ in $\Omega_{p,q}$, as T tends to infinity. Then it follows from the definition of $\{X_t\}$ that $\hat{\sigma}^2(\hat{\theta}_T)$ converges in probability to $S_{p,q}(\theta^*)$, and in addition $\hat{f}_{p,q}(\lambda;\hat{\theta}_T)$ converges in probability to

$$f_{p,q}(\lambda;\theta^*) = S_{p,q}(\theta^*) \left| \frac{\theta^*(\exp(-2\pi i\lambda))}{\phi^*(\exp(-2\pi i\lambda))} \right|^2 \tag{10}$$

at each λ. Therefore we have that

$$E[\{\hat{f}_{p,q}(\lambda;\hat{\theta}_T) - f_{p,q}(\lambda;\theta^*)\}^2]$$

and

$$E[\ \hat{f}_{p,q}(\lambda;\hat{\theta}_T) - f_{p,q}(\lambda;\theta^*)]$$

converges to 0 as T tends to infinity, Thus it follows that

$$E[\{\hat{f}_{p,q}(\lambda;\hat{\theta}_T) - f_X(\lambda)\}^2] \tag{11}$$

converges to

$$\{f_{p,q}(\lambda;\theta^*) - f_X(\lambda)\}^2 \tag{12}$$

at each λ, as T tends to infinity. Consequently the integrated mean squared error of $\hat{f}_{p,q}(\lambda;\hat{\theta}_T)$, i.e.

$$\int_{-0.5}^{0.5} E[\{\hat{f}_{p,q}(\lambda;\hat{\theta}_T) - f_X(\lambda)\}^2]d\lambda, \tag{13}$$

converges to

$$G_{p,q}(\theta^*) = \int_{-0.5}^{0.5} \{f_{p,q}(\lambda;\theta^*) - f_x(\lambda)\}^2 d\lambda \tag{14}$$

as T tends to infinity. Thus the goodness of the local minimal points θ^* of $S_{p,q}(\theta)$ can be evaluated by using the function (14). Since it is very hard to derive general results in the ARMA(p,q) model fitting, the following concrete examples of MA(1) and ARMA(1,1) model fittings are considered.

Example 1 (MA(1) model fitting).

In this case $S_{0,1}(\theta)$ is expressed as follows:

$$S_{0,1}(\theta)=\{r_0 + 2\sum_{h=1}^{\infty} r_h \theta^h\}(1-\theta^2)^{-1}, \tag{15}$$

where $r_h=EX_tX_{t+h}$, for $|\theta|<1$. If $\{X_t\}$ is an MA(1) process with true parameter value θ_0 in $\Omega_{0,1}$, then the minimal point of $S_{0,1}(\theta)$ is uniquely determined by θ_0. On the other hand if the model is not correct, $S_{0,1}(\theta)$ may have more than one local minimal points in $\Omega_{0,1}= \{\theta;|\theta|<1\}$ (see Tanaka and Huzii[4]). We consider the following example. Let $\{X_t\}$ be an MA(4) process such that

$$X_t= (1+0.1B-0.7B^2+0.5B^3+0.5B^4)e_t, \quad \sigma^2= 1,$$

where $\{e_t\}$ is defined in (1). Then if an MA(1) model is fitted to this process, it is seen that $S_{0,1}(\theta)$ has two local minimal points, namely θ_1=-0.85 and θ_2= 0.19, such that $S_{0,1}(\theta_1)$=1.87 < $S_{0,1}(\theta_2)$=1.98 (see Figure 1). This implies that the θ_1 is a global miminal point of the function in a parameter space. In this case, our interest is to see whether or not the grobal minimal point θ_1 is superior to the local one θ_2. Calculating the values of $G_{0,1}(\theta_i)$ for $i=1, 2$, we can see that $G_{0,1}(\theta_1)$ > $G_{0,1}(\theta_2)$ (see Figure 1). This shows that the global mnimal point θ_1 is not superior to the local one θ_2 concerning the estimation of the spectral density in terms of $G_{0,1}(\theta)$. Their spectral densities are shown in Figure 2.

Example 2 (ARMA(1,1) model fittings).

Let $\{X_t\}$ be an AR(2) process with parameters ϕ_1, ϕ_2 and σ^2=1. It follows from (8) that, for each $\Theta=(\phi,\theta)\in\Omega_{1,1}$,

$$S_{1,1}(\Theta)= r_0(A\ \theta^2 +B\theta +C)[(1+\phi_2)(1-\theta^2)(1+\phi_1\theta+\phi_2\theta^2)]^{-1}, \tag{16}$$

where $A= -\phi_2[(1+\phi_2)\phi^2-2\phi_1\phi+(1+\phi_2)]$,

$B= -[\phi_1(1-\phi_2)-2(1-\phi_2{}^2)\phi+\phi_1(1-\phi_2)\phi^2]$,

$C= (1+\phi_2)-2\phi_1\phi+(1+\phi_2)\phi^2$,

and r_0= Var(X_t). In addition, from (14), we can obtain

$$G_{1,1}(\theta)=\{(1+4\theta^2+\theta^4)+8\theta(1+\theta^2)\phi+(1+10\theta^2+\theta^4)\phi^2-2\theta^2\phi^4\}(1-\phi^2)^{-3}$$

$$\cdot S_{1,1}(\theta)\}^2 -2S_{1,1}(-\theta,-\phi)S_{1,1}(\phi,\theta) \qquad (17)$$

$$+\int_{-0.5}^{0.5} f_X^2(\lambda)d\lambda,$$

where $f_X(\lambda)$ is the spectral density of the process X_t.

We know that if an MA(1) model is fitted to some AR(2) process, there exists two local minimal points of $S_{0,1}(\theta)$ in the parameter space (Tanaka and Huzii[4]). If we extend the MA(1) model to an ARMA(1,1) model, how the properties $S_{1,1}(\theta)$ has. The following two examples of an ARMA(1,1) model fitting are considered here, and the properties of the function $S_{1,1}(\theta)$ are investigated.

(a) Let $\{X_t\}$ be an AR(2) process with $\phi_1=0$ and $\phi_2=0.7$ such that $X_t+ 0.7X_{t-2} = e_t$. Then $S_{1,1}(\theta)$ has two local minimal points, namely $\theta_1=(0.5, -0.85)$ and $\theta_2=(-0.5, 0.85)$, such that $S_{1,1}(\theta_1)= S_{1,1}(\theta_2)$. In this case we have $G_{1,1}(\theta_1)= G_{1,1}(\theta_2)$.

(b) Let $\{X_t\}$ be an AR(2) process with the parameters $\phi_1=0.1$ and $\phi_2=0.7$. Then $S_{1,1}(\theta)$ has also two local minimal points, i.e., $\theta_1=(0.41, -0.76)$ and $\theta_2=(-0.27, 0.73)$, such that $S_{1,1}(\theta_1) > S_{1,1}(\theta_2)$ (see Figure 3). In addition, it can be seen that

$G_{1,1}(\theta_1) > G_{1,1}(\theta_2)$. This shows that the global minimal point is superior to the local one in terms of the criterion $G_{1,1}(\theta)$.

REFERENCES

[1] Åström,K.J. and Söderström,T.(1974). Uniqueness of the maximum likelihood estimates of the parameters of an ARMA model, IEEE Trans. Automat. Contr. **19** 769-773.

[2] Huzii,M.(1988). Some properties of conditional quasi-likelihood functions for time series model fitting, Journal of Time Series Analysis, **9** 345-352.

[3] Kabaila,P.(1983). Parameter values of ARMA models minimising the one-step-ahead prediction error when the true system is not in the model set, J. Appl. Prob. **20** 405-408.

[4] Tanaka,M. and Huzii,M. Some Properties of Moving-average model fittings, To appear in J. Japan Statist. Soc.

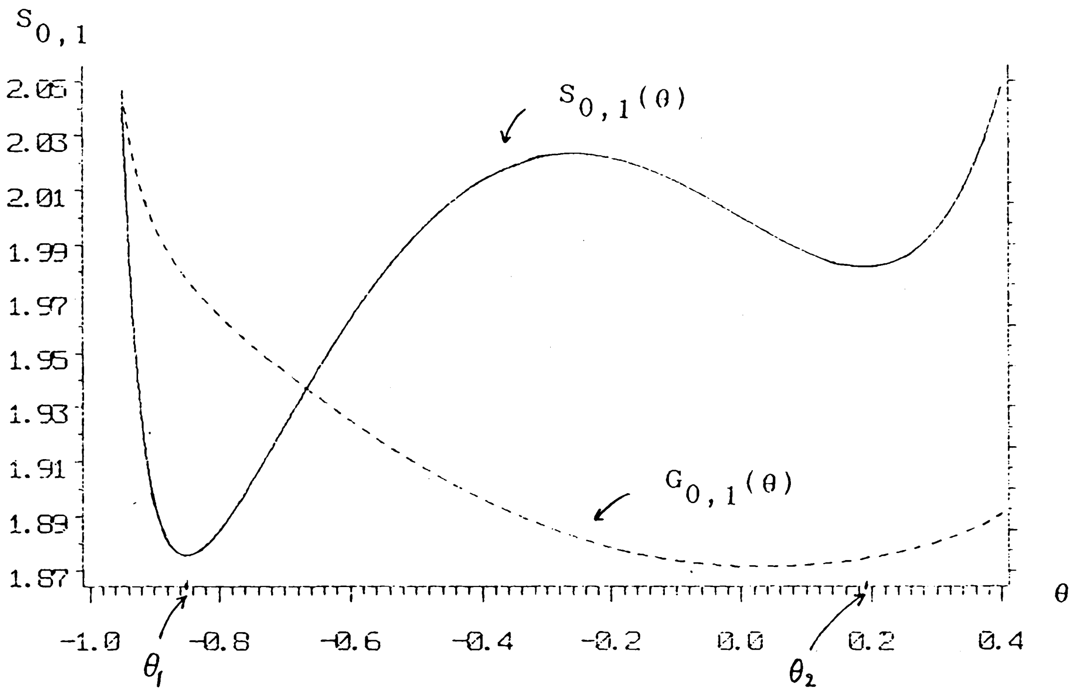

Figure 1. $S_{0,1}(\theta)$ and $G_{0,1}(\theta)$ when MA(1) model is fitted to an MA(4) process (Example 1).

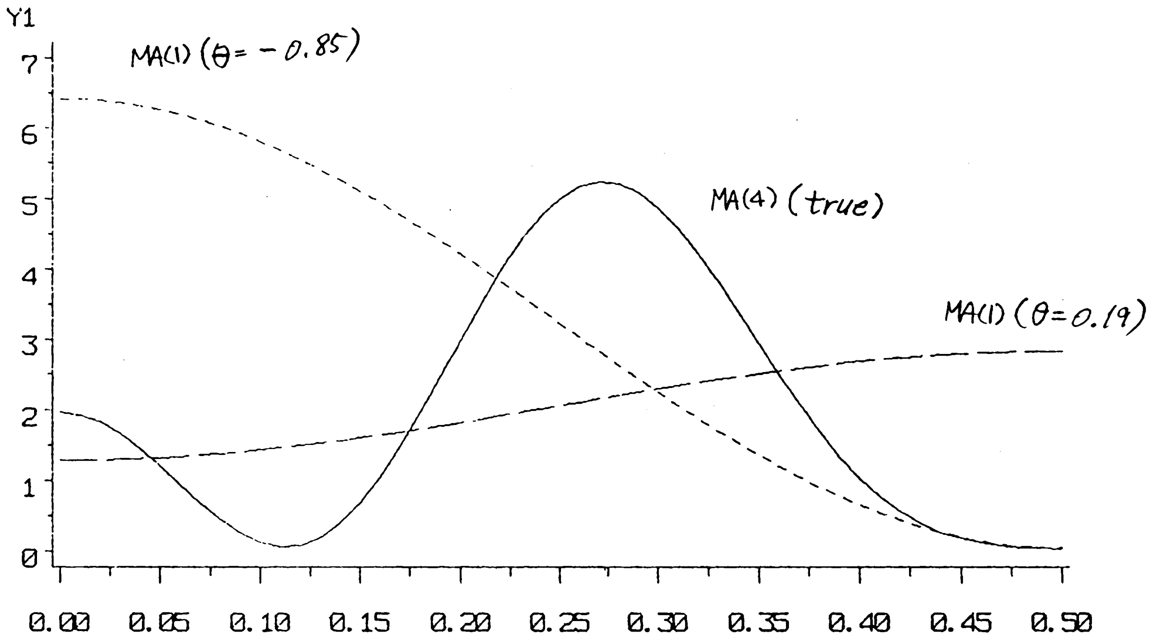

Figure 2. Spectral densities of the MA(1) models and MA(4) process in Example 1.

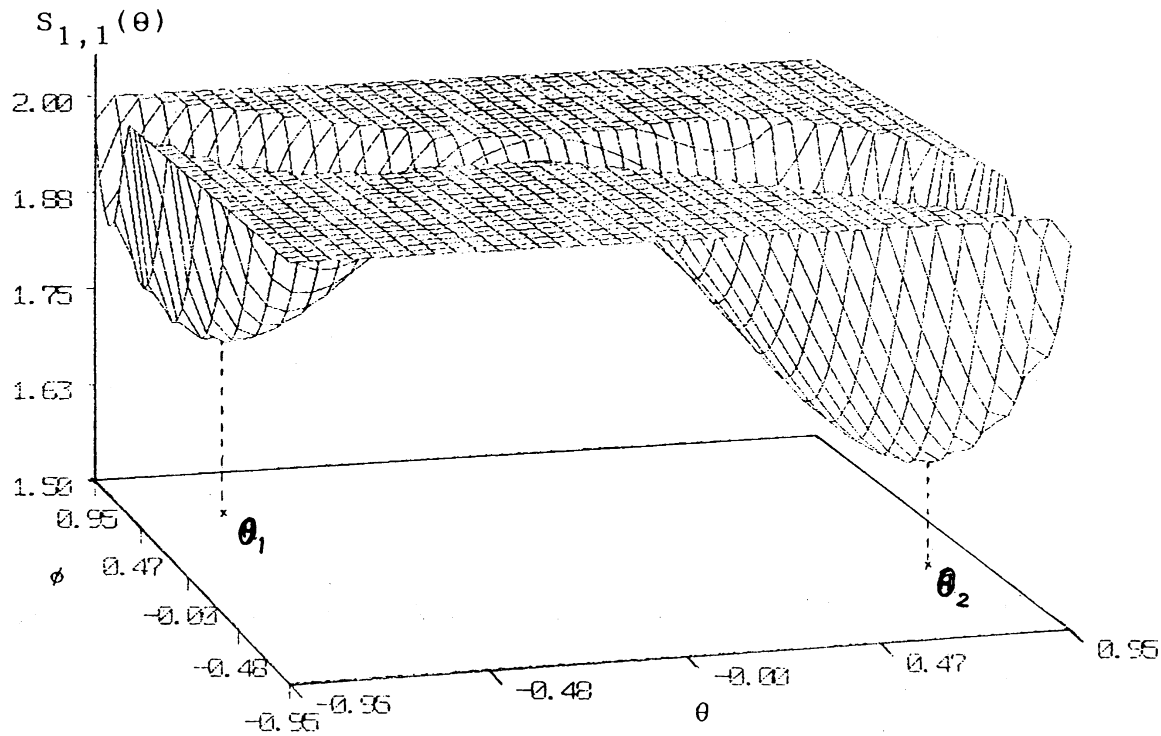

Figure 3. $S_{1,1}(\theta)$ when ARMA(1,1) model is fitted to an AR(2) process (Example 2(b)).