On the second order asymptotic optimality of estimators in an autoregressive process

AKAHIRA Masafumi

Reports of the University of Electro-Communications

Volume 29

Number 2

Page range 213-218

Year 1979-02

URL http://hdl.handle.net/2241/119452
On the Second Order Asymptotic Optimality of Estimators in an Autoregressive Process*

Masafumi AKAHIRA**

Abstract

Let \( \{X_t\} \) be defined by \( X_t = \theta X_{t-1} + U_t \) (\( t = 1, 2, \ldots \)), where \( \{U_t\} \) is a sequence of independent identically distributed random variables with mean 0 and variance 1 and \( X_0 \) is a random variable with mean 0 and variance \( \sigma^2 \) and for each \( t \) \( X_0 \) is independent of \( U_t \). We assume that \( |\theta| < 1 \). It is shown that the stochastic expansion of the maximum likelihood estimator (MLE) of \( \theta \) is obtained and the MLE is second order asymptotically efficient and that the least squares estimator is second order asymptotically efficient when \( U_t \) (\( t = 1, 2, \ldots \)) and \( X_0 \) are normally distributed and \( \sigma^2 = 1/(1-\theta^2) \) in the above process. It is noted that the initial condition of \( X_0 \) does not affect the second order asymptotic efficiency of the MLE.

1. Introduction

It is shown by Akahira ([1]) that the bound of the second order asymptotic distributions of second order asymptotically median unbiased (AMU) estimators of \( \theta \) and the estimator of \( \theta \) attaining it, that is, the second order asymptotically efficient estimator of \( \theta \) are obtained in an autoregressive process \( \{X_t\} \) which is defined recursively by \( X_t = \theta X_{t-1} + U_t \) (\( t = 1, 2, \ldots \)), where \( X_0 = 0 \), \( |\theta| < 1 \) and \( \{U_t\} \) is a sequence of independent identically distributed (i.i.d.) random variables with mean 0 and variance \( \sigma^2 \).

In this paper we consider an autoregressive process \( \{X_t\} \) which is defined by \( X_t = \theta X_{t-1} + U_t \) (\( t = 1, 2, \ldots \)), where \( \{U_t\} \) is a sequence of i.i.d. random variables with mean 0 and variance 1 and \( X_0 \) is a random variable with mean 0 and variance \( \sigma^2 \) and for each \( t \) \( X_0 \) is independent of \( U_t \).

We assume that \( |\theta| < 1 \). It is shown that the stochastic expansion of the maximum likelihood estimator (MLE) of \( \theta \) is obtained and the MLE is second order asymptotically efficient and that the least squares estimator is second order asymptotically efficient when \( U_t \) (\( t = 1, 2, \ldots \)) and \( X_0 \) are normally distributed and \( \sigma^2 = 1/(1-\theta^2) \) in the above process. It is noted that the initial condition of \( X_0 \) does not affect the second order asymptotic efficiency of the MLE.

2. Results

Let \( (\mathcal{X}, \mathcal{B}) \) be a sample space and \( \Theta \) a parameter space, which is assumed to be an open set in a Euclidean 1-space \( \mathbb{R}^1 \). We shall denote by \( (\mathcal{X}^T, \mathcal{B}^T) \) the \( T \)-fold direct products of...
(X, B). For each T = 1, 2, ..., the points of $X^{(T)}$ will be denoted by $x_T = (x_1, ..., x_T)$. We consider a sequence of classes of probability measures $\{P_T, \theta : \theta \in \Theta\}$ (T = 1, 2, ...) each defined on $(X^{(T)}, B^{(T)})$ such that for each T = 1, 2, ... and each $\theta \in \Theta$ the following holds:

$$P_T, \theta(B^{(T)}) = P_{T+1}, \theta(B^{(T)} \times X)$$

for all $B^{(T)} \in B^{(T)}$.

An estimator of $\theta$ is defined to be a sequence $\{\hat{\theta}_T\}$ of $B^{(T)}$-measurable functions $\hat{\theta}_T$ on $X^{(T)}$ to $\Theta$ (T = 1, 2, ...). For simplicity we may denote an estimator $\hat{\theta}$ instead of $\{\hat{\theta}_T\}$. For increasing sequence of positive numbers $\{c_T\}$ (c_T tending to infinity) an estimator $\hat{\theta}$ is called consistent with order $\{c_T\}$ (or $\{c_T\}$-consistent for short) if for every $\epsilon > 0$ and every $\delta \in \Theta$ there exist a sufficiently small positive number $\delta$ and a sufficiently large positive number $L$ satisfying the following:

$$\lim_{T \to \infty} \sup_{\theta \in \Theta \cap \{\theta : |\theta - \delta| < \delta\}} \frac{P_T, \theta(c_T |\hat{\theta}_T - \theta|^2 L) < \epsilon}{(2)}.$$ 

A $\{c_T\}$-consistent estimator $\hat{\theta}$ is second order asymptotically median unbiased (or second order AMU) if for any $\delta \in \Theta$, there exists a positive number $\delta$ such that

$$\lim_{T \to \infty} \sup_{\theta \in \Theta \cap \{\theta : |\theta - \delta| < \delta\}} \frac{c_T P_T, \theta(\hat{\theta}_T \leq \theta) - \frac{1}{2}}{0};$$

$$\lim_{T \to \infty} \sup_{\theta \in \Theta \cap \{\theta : |\theta - \delta| < \delta\}} \frac{c_T P_T, \theta(\hat{\theta}_T \geq \theta) - \frac{1}{2}}{0}.$$

We have defined a second order AMU estimator to be second order asymptotically efficient if the second order asymptotic distribution of it attains uniformly the bound of the second order asymptotic distributions of the second order AMU estimators ([3]). In the subsequent discussions we shall deal only with the case when $c_T = \sqrt{T}$.

Let $X_t$ (t = 1, 2, ...) be defined recursively by

$$X_t = \theta X_{t-1} + U_t \quad (t = 1, 2, ...),$$

where $\{U_t\}$ is a sequence of i.i.d. random variables with mean 0 and variance 1 and $X_0$ is a random variable with mean 0 and variance $\sigma^2$ and for each $t$, $X_t$ is independent of $U_t$. We assume that $\theta = (-1, 1)$, i.e., $|\theta| < 1$. It is seen that the MLE of $\theta$ is $\{\sqrt{T}\}$-consistent.

We denote a joint density of $(X_0, X_1, ..., X_T)$ by $L(\theta) = L(\theta : x_0, x_1, ..., x_T)$. Suppose that the joint density satisfies the following:

(i) $\{(x_0, \bar{x}_T) : L(\theta ; x_0, \bar{x}_T) > 0\}$ does not depend on $\theta$;

(ii) For almost all $(x_0, \bar{x}_T)$, $L(\theta ; x_0, \bar{x}_T)$ is three times differentiable in $\theta$;

(iii) For each $\theta \in \Theta$

$$0 < I(\theta) = \frac{1}{T} \mathbb{E} \left[ \left( \frac{\partial}{\partial \theta} \log L(\theta) \right)^2 \right] = -\frac{1}{T} \mathbb{E} \left[ \frac{\partial^2}{\partial \theta^2} \log L(\theta) \right] < \infty;$$

(iv) There exist

$$J(\theta) = \frac{1}{T} \mathbb{E} \left[ \left( \frac{\partial}{\partial \theta} \log L(\theta) \right) \left( \frac{\partial}{\partial \theta} \log L(\theta) \right)^T \right]$$

and

$$K(\theta) = \frac{1}{T} \mathbb{E} \left( \left( \frac{\partial}{\partial \theta} \log L(\theta) \right)^2 \right)$$

and the following holds:

$$\frac{1}{T} \mathbb{E} \left[ \frac{\partial^3}{\partial \theta^3} \log L(\theta) \right] = -3J(\theta) - K(\theta).$$
By the same way as in [3] it is shown that the MLE is second order asymptotically efficient. Let $\hat{\theta}_{ML}$ be the MLE of $\theta$. By Taylor expansion we have

$$0 = \frac{\partial}{\partial \theta} \log L(\hat{\theta}_{ML})$$

$$= \frac{\partial^2}{\partial \theta^2} \log L(\theta) \left( \hat{\theta}_{ML} - \theta \right) + \frac{1}{2} \left( \frac{\partial^3}{\partial \theta^3} \log L(\theta) \right) \left( \hat{\theta}_{ML} - \theta \right)^2,$$

where $|\theta^* - \theta| \leq |\hat{\theta}_{ML} - \theta|$.

Putting

$$T_n = \sqrt{n} (\hat{\theta}_{ML} - \theta);$$
$$Z_1(\theta) = \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} \log L(\theta);$$
$$Z_2(\theta) = \frac{1}{\sqrt{n}} \left[ \frac{\partial^2}{\partial \theta^2} \log L(\theta) - E_\theta \left[ \frac{\partial^2}{\partial \theta^2} \log L(\theta) \right] \right],$$

we obtain

$$Z_1(\theta) = \frac{1}{\sqrt{n}} \frac{\partial}{\partial \theta} \log L(\theta) + \frac{1}{T} \frac{\partial^2}{\partial \theta^2} \log L(\theta) \frac{T_n^2}{n} + \frac{1}{2T} \sqrt{\frac{2}{T} \frac{\partial^3}{\partial \theta^3} \log L(\theta^*)} \frac{T_n^2}{n^2}.$$

Since $(1/T)(\partial^3/\partial \theta^3) \log L(\theta)$ converges in probability to $-3J(\theta) - K(\theta)$ it follows from (2) that

$$0 = Z_1(\theta) - \left[ I(\theta) - \frac{1}{T} Z_1(\theta) \right] \frac{T_n^2}{n} - \frac{1}{2T} \sqrt{\frac{2}{T} \frac{\partial^3}{\partial \theta^3} \log L(\theta^*)} \frac{T_n^2}{n^2}.$$

Hence we have

$$T_n = \frac{Z_1(\theta)}{I(\theta)} \frac{Z_1(\theta)}{T(\theta)} \frac{Z_2(\theta)}{2I(\theta)^{1/2} \sqrt{T}} \frac{3J(\theta) + K(\theta)}{2I(\theta)^{3/2} \sqrt{T}} \frac{T_n^2}{n^2} + o_P \left( \frac{1}{\sqrt{T}} \right).$$

Therefore we have established:

**Theorem 1.** Under the conditions (i)~(iv)

$$\sqrt{n} (\hat{\theta}_{ML} - \theta) = \frac{Z_1(\theta)}{I(\theta)} \frac{Z_1(\theta)}{T(\theta)^{1/2}} \frac{Z_2(\theta)}{2I(\theta)^{3/2} \sqrt{T}} \frac{3J(\theta) + K(\theta)}{2I(\theta)^{3/2} \sqrt{T}} Z_1(\theta)^2 + o_P \left( \frac{1}{\sqrt{T}} \right).$$

We modify $\hat{\theta}_{ML}$ to be second order AMU and denote it by $\hat{\theta}_{ML}^*$. By a similar way as the i.i.d. case ([3]) we have the following:

**Theorem 2.** Under the conditions (i)~(iv) $\hat{\theta}_{ML}^*$ is second order asymptotically efficient.

We further assume that $U_i \ (i=1, 2, \ldots)$ and $X_0$ are normally distributed and $\sigma^2 = 1/(1 - \theta^2)$ in the process (1). Then it is easily seen that the process (1) is stationary. Then the joint density of $(X_0, X_1, \ldots, X_T)$ is given by

$$L(\theta) = L(\theta; x_0, x_1, \ldots, x_T)$$

$$= \frac{1}{(2\pi)^T(1 - \theta^2)^T/2} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T \frac{1}{T} \left( x_t - \theta x_{t-1} \right)^2 - \frac{1}{2} \left( 1 - \theta^2 \right) x_0^2 \right\}.$$

Since

$$\frac{\partial}{\partial \theta} \log L = \sum_{t=1}^T (X_t - \theta X_{t-1}) X_{t-1} + \theta X_t - \theta;$$
$$\frac{\partial^2}{\partial \theta^2} \log L = - \sum_{t=1}^T X_{t-1}^2 + X_t^2 - \frac{1}{2} \frac{1 - \theta^2}{(1 - \theta^2)^2},$$

it follows by the stationarity of $\{X_t\}$ that

$$Z_1(\theta) = \frac{1}{\sqrt{\frac{T}{T}} \sum_{t=1}^T U_t X_{t-1} + \frac{1}{\sqrt{\frac{T}{T}}} \theta X_0^2} - \frac{\theta}{1 - \theta^2} \sqrt{\frac{T}{T}}.$$
By Theorem 1 we have

\[
\sqrt{T}(\theta_{ML} - \theta) = \frac{1}{I(\theta)\sqrt{T}} \left( \sum_{i=1}^{T} U_i X_{i-1} - \frac{\theta}{I(\theta)\sqrt{T}} X_0^2 - \frac{\theta}{1 - \theta^2} I(\theta) \right) + o_p \left( \frac{1}{\sqrt{T}} \right)
\]

\[
\cdot \left[ \left( \frac{1}{\sqrt{T}} \sum_{i=1}^{T} U_i X_{i-1} + \frac{1}{\sqrt{T}} \theta X_0^2 - \frac{\theta}{1 - \theta^2} \sqrt{T} \right) \right]
\]

\[
+ \frac{3J(\theta) + K(\theta)}{2I(\theta)^3} \left( \frac{1}{\sqrt{T}} \sum_{i=1}^{T} U_i X_{i-1} + \frac{1}{\sqrt{T}} \theta X_0^2 - \frac{\theta}{1 - \theta^2} \sqrt{T} \right)^2 + o_p \left( \frac{1}{\sqrt{T}} \right)
\]

\[
= \frac{1}{\sqrt{T}} \sum_{i=1}^{T} U_i X_{i-1} - \frac{(1-\theta^2)X_0^2 - \theta}{\sqrt{T}}
\]

\[
+ \frac{(1-\theta^2)^2}{T^2} \left( \sum_{i=1}^{T} U_i X_{i-1} \right) + \frac{3J(\theta) + K(\theta)}{2I(\theta)^3} \left( \frac{1}{\sqrt{T}} \sum_{i=1}^{T} U_i X_{i-1} \right)^2 + o_p \left( \frac{1}{\sqrt{T}} \right)
\]

Hence we obtain

\[
\sqrt{T}(\theta_{ML} - \theta) = - \frac{2\theta}{\sqrt{T}} + \frac{2\theta(1-\theta^2)}{\sqrt{T}} X_0^2 + \frac{2(1-\theta^2)}{\sqrt{T}} \sum_{i=1}^{T} U_i X_{i-1}
\]

\[
- \frac{(1-\theta^2)^2}{T^2} \left( \sum_{i=1}^{T} X_0^2 \right) \sum_{i=1}^{T} U_i X_{i-1}
\]

\[
+ \frac{(1-\theta^2)^3}{2T^2} \left( 3J(\theta) + K(\theta) \right) \left( \sum_{i=1}^{T} U_i X_{i-1} \right)^2 + o_p \left( \frac{1}{\sqrt{T}} \right)
\]

Since \( E(U_i^3) = 0 \), it is seen that the fifth term of the right-hand side of (2) does not affect the asymptotic distribution of \( \theta_{ML} \) up to order \( T^{-1/2} \) ([1]).

Next it is shown that the least squares estimator

\[
\hat{\theta}_{LS} = \frac{\sum_{i=1}^{T} X_i X_{i-1}}{\sum_{i=1}^{T} X_{i-1}^2}
\]

is also second order asymptotically efficient. It is seen that \( \hat{\theta}_{LS} \) is \( \{ \sqrt{T} \} \)-consistent.

\[
E_\theta \left( \sum_{i=1}^{T} X_i X_{i-1} \right) = \frac{T}{1 - \theta^2} = \mu \text{ (say)};
\]

\[
\sum_{i=1}^{T} X_i X_{i-1} - \theta \sum_{i=1}^{T} X_{i-1}^2 = \sum_{i=1}^{T} (X_i - \theta X_{i-1}) X_{i-1} = \sum_{i=1}^{T} U_i X_{i-1},
\]

it follows that
On the Second Order Asymptotic Optimality of Estimators

\[ \hat{\theta}_{LS} - \theta = \frac{\sum_{t=1}^{T} U_t X_{t-1}}{\sum_{t=1}^{T} X_{t-1}^2} \]

\[ = \frac{\sum_{t=1}^{T} U_t X_{t-1}}{\left( 1 + \frac{\sum_{t=1}^{T} X_{t-1}^2 - \mu}{\mu} \right)} \]

\[ = \frac{1}{\mu} \left( \sum_{t=1}^{T} U_t X_{t-1} \right) \left( 1 - \frac{\sum_{t=1}^{T} X_{t-1}^2 - \mu}{\mu} \right) + o_p \left( \frac{1}{\sqrt{T}} \right) \]

\[ = \frac{1}{\mu} \left( \sum_{t=1}^{T} U_t X_{t-1} \right) + \frac{1}{\mu^2} \left( \sum_{t=1}^{T} U_t X_{t-1} \right)^2 + o_p \left( \frac{1}{\sqrt{T}} \right) \]

\[ = \frac{1}{\mu^2} \left( \sum_{t=1}^{T} U_t X_{t-1} \right)^2 + o_p \left( \frac{1}{\sqrt{T}} \right) \]  

Hence we obtain

\[ \sqrt{T} (\hat{\theta}_{LS} - \theta) = \frac{2(1 - \theta^2)}{\sqrt{T}} \sum_{t=1}^{T} U_t X_{t-1} - \frac{(1 - \theta^2)^2}{T} \left( \sum_{t=1}^{T} X_{t-1} \right) \left( \sum_{t=1}^{T} U_t X_{t-1} \right) + o_p \left( \frac{1}{\sqrt{T}} \right) \]

We modify \( \hat{\theta}_{LS} \) to be second order AMU and denote it by \( \hat{\theta}_{LS}^* \). Since \( E(X_0 U_t X_{t-1}) = 0 \) \((t=1, 2, \ldots, T)\), it follows that the second term with \( X_0 \) of the right-hand side of (2) does not affect the asymptotic distribution of the MLE up to the order \( T^{-1/2} \), that is, the second order asymptotic efficiency of the MLE.

By Theorems 1, 2 and (2) and (3) we have established:

**Theorem 3.** \( \hat{\theta}_{LS}^* \) is second order asymptotically efficient.

Remark: Although it is seen from the above that the second order asymptotic efficiency of the MLE does not depend on the initial condition of \( X_0 \), it is noted that the second term with \( X_0 \) of the right-hand side of (2) may affect the \( m \)-th \((m \geq 3)\) order asymptotic efficiency of the MLE.

**Acknowledgment**

The author wishes to thank Professor K. Takeuchi of Tokyo University for suggestive discussions.

**References**

