Journal of the Korean Statistical Society (2000), **29**: *2, pp 187–199*

Asymptotic Properties of LAD Esimators of a Nonlinear Time Series Regression Model

Tae Soo Kim, Hae Kyung Kim [1]and Seung Hoe Choi [2]

ABSTRACT

In this paper, we deal with the asymptotic properities of the least absolute deviation estimators in the nonlinear time series regression model. For the sinusodial model which frequently appears in a time series analysis, we study the strong consistency and asymptotic normality of least absolute deviation estimators. And using the derived limiting distributions we show that the least absolute deviation estimators is more efficient than the least squares estimators when the error distribution of the model has heavy tails.

Key Words : Nonlinear Regression; Least Absolute Deviation Estimators; Consistency; Asymptotic Normality; Efficiency.

1. Introduction

Generally, the nonlinear regression model is

$$y_t = f(x_t, \theta_0) + \epsilon_t, \qquad t = 1, 2, \cdots, T$$

where $f(x_t, \theta_0)$ is a real valued nonlinear function defined on R^{p+q}, x_t is a $(1 \times q)$ observed vector, the error term ϵ_t are independent and identically distributed (i.i.d.) with finite variance. The parameter vector θ_0 which is interior point in a compact parameter space Θ is unknown and to be estimated. Jennrich (1969) first rigorously proved the existence of the least squares estimator (LSE) and showed the consistency of the LSE $\check{\theta}_T$ under the following assumption: $F_T(\phi_1, \phi_2)$

[1]Department of Mathematics, Yonsei University, Seoul, 120-749, Korea.
[2]Department of General Studies, Hankuk Aviation University, Koyang, 411-791 Korea.

converges uniformly to a continuous function $F(\phi_1, \phi_2)$ for all $\phi_1 \in \Theta$ and $\phi_2 \in \Theta$ and $F(\phi_1, \phi_2) = 0$ if and only if $\phi_1 = \phi_2$, where

$$F_T(\phi_1, \phi_2) = \frac{1}{T}\sum_{t=1}^{T}\left(f(x_t, \phi_1) - f(x_t, \phi_2)\right)^2.$$

Under some stronger assumptions, asymptotic normality was proved in the same paper. Wu (1981) gave some sufficient conditions under which the LSE converges to θ_0 almost surely, when the growth rate requirement of F_T is replaced by a Lipschitz type condition on the sequence $f(x_t, \theta)$.

The concept of periodicity in time series is of fundermental interest, since it provides a means for formalizing the notions of dependence or correlation between adjacent points. In this paper we think about a sum of sinusoidal components :

$$f(x_t, \theta_o) = \sum_{r=1}^{q}\{A_{ro}\cos(\omega_{ro}t) + B_{ro}\sin(\omega_{ro}t)\},$$

where $\theta_o = (A_{1o}, B_{1o}, \omega_{1o}, \cdots, A_{qo}, B_{qo}, \omega_{qo})$, for $q \geq 1$,A_{ro}, B_{ro}'s are some fixed unknown constants, ω_{ro} is unknown frequency lying between 0 to π $(1 \leq r \leq q)$ and in this case the observed value x_t means t. But in this situation $F_T(\phi_1, \phi_2)$ does not converge uniformly to a continuous function nor it satisfy Wu's Lipschitz type condition, the previous method to gain the LSE is not available. Walker (1971) obtained the asymptotic properties of an approximate LSE. Kundu (1993) and Kundu & Mitra (1996) gave the direct proof of consistency of the LSE and the asymptotic normality results and observed that the approximate LSE and the LSE are asymptotically equal. They found out $P(\breve{\theta}_T) = (P_1(\breve{\theta}_{1T}), P_2(\breve{\theta}_{2T}), \cdots,$ $P_q(\breve{\theta}_{qT}))_{3q\times 1}$ converges in law $N(0,\ \sigma^2\Sigma^{-1})$, where σ^2 is the common variance of errors in the above model and $P_r(\breve{\theta}_{rT}) = (\sqrt{T}(\breve{A}_{rT} - A_{ro}), \sqrt{T}(\breve{B}_{rT} - B_{ro}), \sqrt{T}^3(\breve{\omega}_{rT}$ $- \omega_{ro}))(1 \leq r \leq q)$, and Σ is defined in Theorem 4.2.

In this paper we study the least absolute deviation (LAD) estimators which are defined in (1.2) of the following nonlinear time series model with assumptions A and B,

$$y_t = \sum_{r=1}^{q}\{A_{ro}\cos(\omega_{ro}t) + B_{ro}\sin(\omega_{ro}t)\} + \epsilon_t. \tag{1.1}$$

The LAD estimators of the true parameter $\theta_o = (A_{1o}, B_{1o}, \omega_{1o}, \cdots, A_{qo}, B_{qo}, \omega_{qo})$ denoted by $\hat{\theta}_T = (\hat{A}_{1T}, \hat{B}_{1T}, \hat{\omega}_{1T}, \cdots, \hat{A}_{qT}, \hat{B}_{qT}, \hat{\omega}_{qT})$ is a parameter which minimizes the objective function

$$Q_T(\theta) = \frac{1}{T}\sum_{t=1}^{T}\left| y_t - \sum_{r=1}^{q}\{A_r \cos(\omega_r t) + B_r \sin(\omega_r t)\}\right|, \tag{1.2}$$

where $\theta = (A_1, B_1, \omega_1, \cdots, A_q, B_q, \omega_q)$. On the other hand, Oberhofer (1982) studied the weak consistency about the LAD estimators with the assumptions from B1 to B6 in his paper, but the assumption B5 is equivalent to assumption of Jennrich (1969). So in order to prove the strong consistency of the LAD estimator for this model we must take the different method.

To prove the asymptotic property of $\hat{\theta}_T$, for one harmonic component case, the consistency, asymptotic normality are discussed in section 2 and section 3, respectively. And section 4 consider the case of several harmonic components, the asymptotic relative efficiency (ARE) is stated in section 5.

2. The strong consistency

In this section, for the case of $q = 1$, we will consider the strong consistency of the nonlinear LAD estimators $\hat{\theta}_T = \hat{\theta}_{1T} = (\hat{A}_{1T}, \hat{B}_{1T}, \hat{\omega}_{1T}) = (\hat{A}_T, \hat{B}_T, \hat{\omega}_T)$ for $\theta_o = (A_{1o}, B_{1o}, \omega_{1o}) = (A_o, B_o, \omega_o)$ in a time series with stationary independent residuals model (1.1) with the following assumptions.

Assumption A

The parameter space $\Theta = K \times K \times [0,\ \pi]$, where K is compact subspace of R.

Assumption B

B1: ϵ_t are i.i.d. random variables with the common distribution function G and continuous probability density function $g(x)$ such that G has unique median at zero, and $g(0) > 0$.

B2: $E\{\epsilon_t{}^2\} < \infty$, for all t.

Remark If $\omega_0 = 0$, the model (1.1) is a simple linear model, and if $\omega_0 = \pi$, it is a simple regrssion model with binary regressors. So we may assume from now on that $\omega_0 \in (0, \pi)$.

Since $Q_T(\theta_0)$ defined in (1.2) is independent of $\theta = \theta_1 = (A_1, B_1, \omega_1)$, minimization of $Q_T(\theta)$ is equivalent to minimization of the following new objective function :

$$D_T(\theta) = Q_T(\theta) - Q_T(\theta_0).$$

Theorem 2.1 *Suppose that Assumption A and B are satisfied on the model (1.1). Then the LAD estimators $\hat{\theta}_T$ is strongly consistent for θ_o.*

Proof : Let $D_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} X_t$, where $h_t(\theta) = A_0 \cos\omega_0 t + B_0 \sin\omega_0 t - A\cos\omega t - B\sin\omega t$, and $X_t = |h_t(\theta) + \epsilon_t| - |\epsilon_t|$.

First of all, we consider that the case $h_t(\theta) \geq 0$, then we obtain

$$E(X_t) = 2\int_{-h_t(\theta)}^{0} [h_t(\theta) + \epsilon_t] dG(\epsilon_t).$$

Using the integration by parts and mean value theorem, there exists $-h_t^*(\theta) \in [-h_t(\theta), 0]$ such that

$$E(X_t) = 2h_t(\theta)[G(0) - G(-h_t^*(\theta))].$$

Likewise, for $h_t(\theta) < 0$, we alse take the same results.

$$E(X_t) = 2h_t(\theta)[G(0) - G(-h_t^{**}(\theta))],$$

where for some $-h_t^{**}(\theta) \in [0, -h_t(\theta)]$. Hence we have $E(X_t) < \infty$.

We also get $Var(X_t)$ is bounded. Thus we can apply the Kolmogorov's strong law of large numbers and then we obtain the following result:

$$\left\{ D_T(\theta) - \lim_{T\to\infty} E[D_T(\theta)] \right\} \xrightarrow{a.s.} 0 \ \ uniformly\ for\ all\ \theta\ in\ \Theta. \qquad (2.1)$$

Let $Q(\theta) = \lim_{T\to\infty} E[D_T(\theta)]$. Now, using the next facts $h_t(\theta_0) = 0$ and $G(0) = \frac{1}{2}$, we obtain

$$\frac{\partial Q(\theta)}{\partial \theta}|_{\theta=\theta_0} = (0, 0, 0),$$

$$\begin{aligned} \left|\frac{\partial^2 Q(\theta)}{\partial\theta'\partial\theta}\right|_{\theta=\theta_0} &= \lim_{T\to\infty} T^2 \begin{vmatrix} g(0) & 0 & \frac{1}{2}B_0 g(0) \\ 0 & g(0) & -\frac{1}{2}A_0 g(0) \\ \frac{1}{2}B_0 g(0) & -\frac{1}{2}A_0 g(0) & \frac{1}{3}(A_0^2 + B_0^2) g(0) \end{vmatrix} \\ &= \lim_{T\to\infty} T^2 \cdot \frac{g^3(0)}{12}(A_0^2 + B_0^2) > 0, \end{aligned} \qquad (2.2)$$

and also the leading principal minors are positive. Hence $\frac{\partial^2 Q(\theta)}{\partial\theta'\partial\theta}|_{\theta=\theta_0}$ is a positive-definite and then θ_0 is a local minimum point of $Q(\theta)$.

For all $\theta \neq \theta_0$, we obtain the following

$$\begin{aligned} Q(\theta) &= \lim_{T\to\infty} \frac{2}{T}[\sum_{h_t(\theta)>0} h_t(\theta)[G(0) - G(-h_t^*(\theta))] \\ &\qquad + \sum_{h_t(\theta)<0} h_t(\theta)[G(0) - G(-h_t^{**}(\theta))]] \\ &\geq \lim_{T\to\infty} \frac{2}{T}\sum_{t=1}^{T} |h_t(\theta)| \cdot \min\{G(-h_t^{**}(\theta)) - \frac{1}{2}, \frac{1}{2} - G(-h_t^*(\theta))\}. \end{aligned}$$

It sufficies to show that $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} |h_t(\theta)|^2 > 0$, using the exponent form of the harmonic components, we have that.

Hence we have

$$Q(\theta) \textit{ has the unique minimizer } \theta_0 \textit{ in } \Theta. \tag{2.3}$$

The above results satisfy the conditions of the lemma 2.2 of White(1980), we can state the result as follows : $\hat{\theta}_T$ is a strongly consistent estimator of θ_0. □

3. Asymptotic normality

In present section we consider the asymptotic normality of the proposed estimator $\hat{\theta}_T$ which is one of the most important statistical properties in asymptotic theory. The main idea is to approximate to the function $|x|$ by a smooth function $\rho_T(x)$ such that $\lim_{T\to\infty} \rho_T(x) = |x|$. We take such function

$$\rho_T(x) = \left[-\frac{1}{3}\beta_T^2 x^3 + \beta_T x^2 + \frac{1}{3\beta_T}\right] I_{\{0<x\leq\frac{1}{\beta_T}\}} + xI_{\{\frac{1}{\beta_T}<x\}}$$

and

$$\rho_T(x) = \rho_T(-x)$$

where I_A denotes the indicator function of the event A and β_T is an appropriately chosen increasing function of T such that $\lim_{T\to\infty} \frac{1}{\beta_T} = 0$ with $T^2 = o(\beta_T^3)$ and $\beta_T = o(T)$.

We define

$$Q_T^*(\theta) = \frac{1}{T}\sum_{t=1}^{T} \rho_T(h_t(\theta) + \epsilon_t).$$

For large T, $Q_T^*(\theta)$ is close to $Q_T(\theta)$ defined in (1.2). Since $Q_T^*(\theta)$ is twice continuously differentiable with respect to θ whereas the first derivative of $Q_T(\theta)$ is discontinuous, we use new object function $Q_T^*(\theta)$ instead of $Q_T(\theta)$. And then we prove the asymptotic normality of $\hat{\theta}_T$ by using a Taylor expansion of $\nabla Q_T^*(\theta)$, where $\nabla Q_T^*(\theta)$ is the derivate of $Q_T^*(\theta)$ with respect to θ. Let $\theta_T^* = (A_{1T}^*, B_{1T}^*, \omega_{1T}^*) = (A_T^*, B_T^*, \omega_T^*)$ be the pseudo estimators such that minimize the $Q_T^*(\theta)$.

Lemma 3.1 *For the model (1.1) with assumptions A and B,*

$\left(\sqrt{T}(A_T^* - A_0), \sqrt{T}(B_T^* - B_0), \sqrt{T}^3(\omega_T^* - \omega_0)\right)$ *converges in law* $N\left((0,0,0), \frac{1}{(2g(0))^2}\Sigma^{-1}\right)$, *where*

$$\Sigma = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{4}B_0 \\ 0 & \frac{1}{2} & -\frac{1}{4}A_0 \\ \frac{1}{4}B_0 & -\frac{1}{4}A_0 & \frac{(A_0^2+B_0^2)}{6} \end{pmatrix}. \tag{3.1}$$

Proof : Firstly, we show that the minimum of $Q_T^*(\theta)$ sufficiently close to the minimum of $Q_T(\theta)$. By the definition $Q_T^*(\theta)$ and $Q_T(\theta)$, we obtain

$$\begin{aligned} Q_T^*(\theta) - Q_T(\theta) &= \frac{1}{T}\sum_{t=1}^{T}[(-\frac{1}{3}\beta_T^2 x^3(t;\theta) + \beta_T x^2(t;\theta) + \frac{1}{3\beta_T} - x(t;\theta))I_{\{0<x(t;\theta)\le\frac{1}{\beta_T}\}} \\ &\quad +(\frac{1}{3}\beta_T^2 x^3(t;\theta) + \beta_T x^2(t;\theta) + \frac{1}{3\beta_T} + x(t;\theta))I_{\{-\frac{1}{\beta_T}<x(t;\theta)\le 0\}}] \\ &\stackrel{let}{=} \frac{1}{T}\sum_{t=1}^{T} Y_t, \ \textit{where } x(t;\theta) = h_t(\theta) + \epsilon_t. \end{aligned}$$

By the continuity of $Q_T^*(\theta) - Q_T(\theta)$ and compactness of parameter space, we can choose θ^* such that $T\{Q_T^*(\theta^*) - Q_T(\theta^*)\} = \sup_{\theta\in\Theta} T\{Q_T^*(\theta) - Q_T(\theta)\}$.

Since for all t, $|Y_t| \le \frac{1}{3\beta_T} I_{\{|h_t(\theta)+\epsilon_t|<\frac{1}{\beta_T}\}}$, using mean value theorem and Chebyshev's inequality we have $T\{Q_T^*(\theta) - Q_T(\theta)\} = o_p(1)$. Hence we get that

$$\sup_{\theta\in\Theta} T\{Q_T^*(\theta) - Q_T(\theta)\} = o_p(1). \tag{3.2}$$

Likewise theorem 2.1, by the above (3.2), we easily show that θ_T^* converges almost surely to θ_0 also.

By definition of $\hat{\theta}_T$, we know $Q_T(\hat{\theta}_T) - Q_T(\theta_T^*) \leq 0$, and

$$Q_T^*(\hat{\theta}_T) - Q_T^*(\theta_T^*) \leq Q_T^*(\hat{\theta}_T) - Q_T(\hat{\theta}_T) + Q_T(\theta_T^*) - Q_T^*(\theta_T^*).$$

Hence by (3.2), we obtain that

$$T\{Q_T^*(\hat{\theta}_T) - Q_T^*(\theta_T^*)\} = o_p(1). \tag{3.3}$$

Futhermore, by the Taylor expansion we have the following

$$Q_T^*(\hat{\theta}_T) = Q_T^*(\theta_T^*) + \nabla Q_T^*(\theta_T^*)(\hat{\theta}_T - \theta_T^*) + \frac{1}{2}(\hat{\theta}_T - \theta_T^*)^T \nabla^2 Q_T^*(\bar{\theta}_T)(\hat{\theta}_T - \theta_T^*),$$

where $\bar{\theta}_T = (\bar{A}_{1T}, \bar{B}_{1T}, \bar{\omega}_{1T}) = (\bar{A}_T, \bar{B}_T, \bar{\omega}_T) = \gamma\hat{\theta}_T + (1-\gamma)\theta_T^*$, for some $0 < \gamma < 1$. Let

$$\nabla Q_T^*(\theta) = \frac{\partial Q_T^*(\theta)}{\partial \theta}, \qquad \nabla^2 Q_T^*(\theta) = \frac{\partial^2 Q_T^*(\theta)}{\partial \theta' \partial \theta} = (\alpha_{ij}(T;\theta))_{3\times 3}.$$

Then by the definition of θ_T^*, $\nabla Q_T^*(\theta_T^*) = 0$, hence we obtain the following equation

$$Q_T^*(\hat{\theta}_T) - Q_T^*(\theta_T^*) = \frac{1}{2}(\hat{\theta}_T - \theta_T^*)^T \nabla^2 Q_T^*(\bar{\theta}_T)(\hat{\theta}_T - \theta_T^*). \tag{3.4}$$

Since $\nabla^2 Q_T^*(\bar{\theta}_T)$ is symmetric matrix, by Courant-Fisher minimax characterization, we also have

$$\lambda_1(T)(\hat{\theta}_T - \theta_T^*)^T(\hat{\theta}_T - \theta_T^*) \leq (\hat{\theta}_T - \theta_T^*)^T \nabla^2 Q_T^*(\bar{\theta}_T)(\hat{\theta}_T - \theta_T^*),$$

where $\lambda_1(T)$ is the smallest eigen value of $\nabla^2 Q_T^*(\bar{\theta}_T)$. By (3.4), thus we have

$$T(\hat{\theta}_T - \theta_T^*)^T(\hat{\theta}_T - \theta_T^*) \leq \frac{2T}{\lambda_1(T)}\{Q_T^*(\hat{\theta}_T) - Q_T^*(\theta_T^*)\}.$$

For the proof of $\sqrt{T}(\hat{\theta}_T - \theta_T^*) = o_p(1)$, it remains to show that $\lambda_1(T) > 0$ as $T \to \infty$. Hence it suffices to prove that $\lim_{T\to\infty} \nabla^2 Q_T^*(\bar{\theta}_T)$ is positive definite matrix. And we have the following result by simple operations

$$E\left[\frac{1}{T}\sum_{t=1}^{T}\rho_T''(h_t(\bar{\theta}_T) + \epsilon_t) - 2g(0)\right] = o(1)$$

With the facts $h_t(\bar{\theta}_T) + \epsilon_t \to \epsilon_t$ *a.s.* , $\rho_T''(x)$ is continuous function, and the Assumption B1, we have $\{\rho_T''(h_t(\bar{\theta}_T) + \epsilon_t)\}$ are independent random variables almost surely when T is a sufficiently large number. And so we know that

$Var\left(\frac{1}{T}\sum_{t=1}^{T}\rho_T''(h_t(\bar{\theta}_T)+\epsilon_t)\right) = o(1)$. By the Chebyshev's inequality, we have $\frac{1}{T}\sum_{t=1}^{T}\rho_T''(h_t(\bar{\theta}_T)+\epsilon_t) = 2g(0)+o_p(1)$ and the following results

$$\alpha_{11}(T;\bar{\theta}_T) = \alpha_{22}(T;\bar{\theta}_T) = g(0)+o_p(1), \quad \frac{1}{T}\alpha_{13}(T;\bar{\theta}_T) = \frac{1}{2}B_0 g(0)+o_p(1),$$
$$\frac{1}{T}\alpha_{23}(T;\bar{\theta}_T) = -\frac{1}{2}A_0 g(0)+o_p(1), \quad \frac{1}{T^2}\alpha_{33}(T;\bar{\theta}_T) = \frac{1}{3}(A_0^2+B_0^2)g(0)+o_p(1).$$

Likewise (2.2), we know $\nabla^2 Q_T^*(\bar{\theta}_T)$ is a positive-definite. Therefore we have

$$\sqrt{T}(\hat{\theta}_T-\theta_T^*) = o_p(1). \tag{3.5}$$

Under the same conditions, particularly by the Taylor expansion of $Q_T^*(\theta)$ about only ω, $Q_T^*(\hat{\theta}_T)-Q_T^*(\theta_T^*) = \frac{1}{2}(\hat{\omega}_T-\omega_T^*)^2\frac{\partial^2 Q_T^*(\bar{\theta}_T)}{\partial\omega^2}$, we have

$$\frac{2}{\alpha_{33}(T;\bar{\theta}_T)}T(Q_T^*(\hat{\theta}_T)-Q_T^*(\theta_T^*)) = T^3(\hat{\omega}_T-\omega_T^*)^2.$$

By (3.3), hence we conclude that

$$T^{\frac{3}{2}}(\hat{\omega}_T-\omega_T^*) = o_p(1). \tag{3.6}$$

Since $Q_T^*(\theta)$ is a minimum when $\theta = \theta_T^*$, an application of the mean value theorem gives

$$\begin{aligned}(Q_T^*)_{A_0} &= (Q_T^*)_{\bar{A}\bar{A}}(A_0-A_T^*)+(Q_T^*)_{\bar{A}\bar{B}}(B_0-B_T^*)+(Q_T^*)_{\bar{A}\bar{\omega}}(\omega_0-\omega_T^*),\\ (Q_T^*)_{B_0} &= (Q_T^*)_{\bar{B}\bar{A}}(A_0-A_T^*)+(Q_T^*)_{\bar{B}\bar{B}}(B_0-B_T^*)+(Q_T^*)_{\bar{B}\bar{\omega}}(\omega_0-\omega_T^*),\\ (Q_T^*)_{\omega_o} &= (Q_T^*)_{\bar{\omega}\bar{A}}(A_0-A_T^*)+(Q_T^*)_{\bar{\omega}\bar{B}}(B_0-B_T^*)+(Q_T^*)_{\bar{\omega}\bar{\omega}}(\omega_0-\omega_T^*),\end{aligned} \tag{3.7}$$

where

$$(Q_T^*)_{A_0} = \frac{\partial Q_T^*(A,B,\omega)}{\partial A}|_{(A_0,B_0,\omega_0)},$$
$$(Q_T^*)_{\bar{A}\bar{B}} = \frac{\partial^2 Q_T^*(A,B,\omega)}{\partial A\partial B}|_{(\bar{A}_T,\bar{B}_T,\bar{\omega}_T)}, \ etc,$$

and we use the generic notation $(\bar{A}_T,\bar{B}_T,\bar{\omega}_T)$ for a point on the line joining (A_0,B_0,ω_0) and (A_T^*,B_T^*,ω_T^*), so that $(\bar{A}_T,\bar{B}_T,\bar{\omega}_T) = \gamma(A_0,B_0,\omega_0)+(1-\gamma)(A_T^*,B_T^*,\omega_T^*)$.

The point $(\bar{A}_T, \bar{B}_T, \bar{\omega}_T)$ in (3.7) will in general not be the same, but to distinguish them would complicate the notation, and no ambiguity will arise by not doing so. But (3.7) is replaced by the following

$$\left(\sqrt{T}(Q_T^*)_{A_0}, \sqrt{T}(Q_T^*)_{B_0}, \frac{1}{\sqrt{T}}(Q_T^*)_{\omega_o}\right)$$
$$= -\left(\sqrt{T}(A_T^* - A_0), \sqrt{T}(B_T^* - B_0), \sqrt{T}^3(\omega_T^* - \omega_0)\right) \times W_T^*, \qquad (3.8)$$

where

$$W_T^* = \begin{pmatrix} (Q_T^*)_{\bar{A}\bar{A}} & (Q_T^*)_{\bar{A}\bar{B}} & T^{-1}(Q_T^*)_{\bar{A}\bar{\omega}} \\ (Q_T^*)_{\bar{B}\bar{A}} & (Q_T^*)_{\bar{B}\bar{B}} & T^{-1}(Q_T^*)_{\bar{B}\bar{\omega}} \\ T^{-1}(Q_T^*)_{\bar{\omega}\bar{A}} & T^{-1}(Q_T^*)_{\bar{\omega}\bar{B}} & T^{-2}(Q_T^*)_{\bar{\omega}\bar{\omega}} \end{pmatrix}.$$

Now

$$\sqrt{T}(Q_T^*)_{A_0} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(-\cos\omega_0 t) \times [k_T(\epsilon_t) + l_T(\epsilon_t)],$$

we use the Markov theorem, then we have the following results,

$$\sqrt{T}(Q_T^*)_{A_0} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(-\cos\omega_0 t)k_T(\epsilon_t) + o_p(1), \qquad (3.9)$$

and similarly

$$\sqrt{T}(Q_T^*)_{B_0} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\sin\omega_0 t)k_T(\epsilon_t) + o_p(1), \qquad (3.10)$$

$$\frac{1}{\sqrt{T}}(Q_T^*)_{\omega_o} = \frac{1}{\sqrt{T}^{3/2}}\sum_{t=1}^{T}(A_0 t\sin\omega_0 t - B_0 t\cos\omega_0 t)k_T(\epsilon_t) + o_p(1), \qquad (3.11)$$

where $k_T(\epsilon_t) = I_{\{\epsilon_t \geq \frac{1}{\beta_T}\}} - I_{\{\epsilon_t \leq -\frac{1}{\beta_T}\}}$, and $l_T(\epsilon_t) = l_T^+(\epsilon_t) \cdot I_{\{0<\epsilon_t\leq\frac{1}{\beta_T}\}} + l_T^-(\epsilon_t) \cdot I_{\{-\frac{1}{\beta_T}<\epsilon_t\leq 0\}}$, $l_T^+(\epsilon_t) = -\beta_T^2\epsilon_t^2 + 2\beta_T\epsilon_t$, $l_T^-(\epsilon_t) = \beta_T^2\epsilon_t^2 + 2\beta_T\epsilon_t$.

The sum in (3.9)-(3.11) are of the form $\sum_{t=1}^{T} U_t$, where $E(U_t) = o(1)$, and $B_T^2 = \sum_{t=1}^{T} Var(U_t) = \frac{1}{2} + o(1)$ in (3.9)-(3.10), $B_T^2 = \frac{A_0^2 + B_0^2}{6} + o(1)$ in (3.11). And also we have

$$\lim_{T\to\infty} \frac{1}{B_T^2}\sum_{t=1}^{T} E\left[U_t^2 \cdot I_{\{|U_t|\geq\epsilon B_T\}}\right] = 0.$$

By the Lindberg theorem, we see that $\sqrt{T}(Q_T^*)_{A_0}$, $\sqrt{T}(Q_T^*)_{B_0}$, $\frac{1}{\sqrt{T}}(Q_T^*)_{\omega_o}$ converge in law respectively to $N(0, \frac{1}{2})$, $N(0, \frac{1}{2})$ and $N(0, \frac{A_0^2+B_0^2}{6})$.

For the limiting joint distribution we consider the random variable

$$V_T(\delta_1, \delta_2, \delta_3) = \delta_1\sqrt{T}(Q_T^*)_{A_0} + \delta_2\sqrt{T}(Q_T^*)_{B_0} + \delta_3\sqrt{T}^{-1}(Q_T^*)_{\omega_o},$$

where the δ_i ($i = 1, 2, 3$) are arbitrary real numbers. Now also using the Markov theorem,

$$\begin{aligned} &V_T(\delta_1, \delta_2, \delta_3) \\ &= \frac{1}{T}\sum_{t=1}^{T}\left(\delta_1\sqrt{T}(-\cos\omega_0 t) + \delta_2\sqrt{T}(-\sin\omega_0 t) + \delta_3\frac{1}{\sqrt{T}}C_0(t)\right)k_T(\epsilon_t) + o_p(1), \end{aligned}$$

where $C_0(t) = A_0 t\sin\omega_0 t - B_0 t\cos\omega_0 t$. Likewise the previous case, let $V_T(\delta_1, \delta_2, \delta_3) = \sum_{t=1}^{T} U_t$, we have $E(U_t) = o(1)$ and

$$B_T^2 = \sum_{t=1}^{T} Var(U_t) = \frac{\delta_1^2}{2} + \frac{\delta_2^2}{2} + \frac{A_0^2 + B_0^2}{6}\delta_3^2 + \frac{B_0}{2}\delta_1\delta_3 - \frac{A_0}{2}\delta_2\delta_3 + o(1).$$

Hence by Lindeberg theorem applied to the above sum, we have the fact that $V_T(\delta_1, \delta_2, \delta_3)$ converges in law to a normal distribution with mean zero and variance $\frac{\delta_1^2}{2} + \frac{\delta_2^2}{2} + \frac{A_0^2 + B_0^2}{6}\delta_3^2 + \frac{B_0}{2}\delta_1\delta_3 - \frac{A_0}{2}\delta_2\delta_3$. Consequently, by virtue of the Cramér-Wold device, we see that the joint distribution of $\sqrt{T}(Q_T^*)_{A_0}$, $\sqrt{T}(Q_T^*)_{B_0}$ and $\frac{1}{\sqrt{T}}(Q_T^*)_{\omega_o}$ converge to that $N((0,0,0), \Sigma)$, where $\lim_{T\to\infty} W_T^* = W_0 = 2g(0)\Sigma$, W_T^* is defined in (3.8). So by (3.8) this lemma was proved. □

Theorem 3.2 *For the given model (1.1) with the assumptions A and B, we conclude that* $\left(\sqrt{T}(\hat{A}_T - A_0), \sqrt{T}(\hat{B}_T - B_0), \sqrt{T}^3(\hat{\omega}_T - \omega_0)\right)$ *converges in law* $N\left((0,0,0), \frac{1}{(2g(0))^2}\Sigma^{-1}\right)$.

Proof : Note that

$$\begin{aligned} &\left(\sqrt{T}(\hat{A}_T - A_0), \sqrt{T}(\hat{B}_T - B_0), \sqrt{T}^3(\hat{\omega}_T - \omega_0)\right) \\ &= \left(\sqrt{T}(\hat{A}_T - A_T^*), \sqrt{T}(\hat{B}_T - B_T^*), \sqrt{T}^3(\hat{\omega}_T - \omega_T^*)\right) \\ &\quad + \left(\sqrt{T}(A_T^* - A_0), \sqrt{T}(B_T^* - B_0), \sqrt{T}^3(\omega_T^* - \omega_0)\right). \end{aligned}$$

By the (3.5)-(3.6), we have $\sqrt{T}(\hat{A}_T - A_T^*)$, $\sqrt{T}(\hat{B}_T - B_0^*)$ and $\sqrt{T}^3(\hat{\omega}_T - \omega_T^*)$ converge in probability to zero, and by (3.1), this main theorem was proved.□

4. The case of several harmonic components

Suppose now that the model in (1.1) is generalized to $q > 1$. The function corresponding to (1.2) whose minimization yields estimators $\hat{\theta}_T=(\hat{A}_{1T},\hat{B}_{1T},\hat{\omega}_{1T}, \cdots, \hat{A}_{qT}, \hat{B}_{qT}, \hat{\omega}_{qT})$ becomes (1.2), where $\theta = (A_1, B_1, \omega_1, \cdots, A_q, B_q, \omega_q)$. Also, by Theorem 2.1, for any θ, $D_T(\theta) \to Q(\theta)$ a.s., θ_o is also at least local minimum point of $Q(\theta)$. And we also have for all $\theta \neq \theta_0$ and likewise for the Walker (1971) with the following sufficient condition

$$\lim_{T\to\infty} \min_{1\le r\neq s\le q} (T|\omega_r - \omega_s|) = \infty, \tag{4.1}$$

$Q(\theta) > 0$, i.e. θ_o is the global unique minimizer of $Q(\theta)$. Therefore we we can state the result as a Theorem.

Theorem 4.1 *If $\hat{\theta_T}$ is a LAD estimator of the model (1.1), then it is a strongly consistent estimator of θ_o.*

To prove the asymptotic normality, first of all, we have

$$\nabla^2 Q_T^*(\bar{\theta}_T) = (M(T)_{rs})_{3q\times 3q},$$

where for each $r, s = 1, 2, \cdots, q$, $M(T)_{rs}$ is a 3×3 matrices, and

$$\lim_{T\to\infty} |M(T)_{rs}| = \begin{cases} 0 & for\ r \neq s, \\ \lim\limits_{T\to\infty} T^2 \frac{g^3(0)}{12}(A_{ro}^2 + B_{ro}^2) > 0 & for\ r = s. \end{cases}$$

Hence we know that $\lim\limits_{T\to\infty} \nabla^2 Q_T^*(\bar{\theta}_T)$ is positive definite matrix. This fact indicates for $r = 1, 2, \cdots, q$, $\left(\sqrt{T}(\hat{A}_{rT} - A_{rT}^*), \sqrt{T}(\hat{B}_{rT} - B_{rT}^*), \sqrt{T}^3(\hat{\omega}_{rT} - \omega_{rT}^*)\right) = o_p(1)$, and we can have the fact likewise in (3.8)

$$\begin{aligned} &\left(\sqrt{T}(A_{rT}^* - A_{ro}), \sqrt{T}(B_{rT}^* - B_{ro}), \sqrt{T}^3(\omega_{rT}^* - \omega_{ro})\right) \\ &= -\left(\sqrt{T}(Q_T^*)_{A_{ro}}, \sqrt{T}(Q_T^*)_{B_{ro}}, \frac{1}{\sqrt{T}}(Q_T^*)_{\omega_{ro}}\right) \times (W_{rT}^*)^{-1}, \end{aligned}$$

where W_{rT}^* in (3,8), and $r = 1, 2, \cdots, q$.

Also likewise the theorem 3.2, we obtain the following theorem

Theorem 4.2 *With the same conditions of Theorem 3.2 and the condition (4.1),* $P(\hat{\theta}_T) = (P_1(\hat{\theta}_{1T}), P_2(\hat{\theta}_{2T}), \cdots, P_q(\hat{\theta}_{qT}))$, *where* $P_r(\hat{\theta}_{rT}) = (\sqrt{T}(\hat{A}_{rT} - A_{ro}), \sqrt{T}(\hat{B}_{rT} - B_{ro}), \sqrt{T}^3(\hat{\omega}_{rT} - \omega_{ro}))(1 \le r \le q)$ *converges in law* $N(\mathbb{0}_{3q\times 1}, \frac{1}{(2g(0))^2}\Sigma^{-1})$, *where* $\Sigma = (\Sigma_{rs})_{3q\times 3q}$, *for* $r, s = 1, 2, \cdots, q$,

$$\Sigma_{rs} = \begin{cases} \mathbb{0} & if\ r \neq s, \\ \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{4}B_{ro} \\ 0 & \frac{1}{2} & -\frac{1}{4}A_{ro} \\ \frac{1}{4}B_{ro} & -\frac{1}{4}A_{ro} & \frac{(A_{ro}^2 + B_{ro}^2)}{6} \end{pmatrix} & if\ r = s. \end{cases}$$

5. Conclusions

Since the asymptotic efficiency of LSE $\check{\theta}_T$ relative to LAD estimators $\hat{\theta}_T$ is $\{2g(0)\}^2\sigma^2$, it does impliy that the LAD esitmator is asymptotically more efficient than LSE in the sinusodial model for the heavy tailed error distributions likewise double-exponential and logistic distribution, etc.

ACKNOWLEDGEMENTS

The authors wish to thank two anonymous referees for giving valuable comments which led to an improvement in the presentation of this paper.

REFERENCES

Debasis Kundu (1993), *Asymptotic theory of least squares estimator of a particular nonlinear regression model*, Statistics & Probability Letters 18, 13-17.

Debasis Kundu and Amit Mitra (1996), *Asymptotoic theory of least squares estimator of a nonlinear time series regression model*, COMMUN. STATIST.-THEORY METH., 23(1), 133-141.

Jennrich, R. (1969), *Asymptotic properties of nonlinear least squares estimations*, Ann. Math. Stat. 40, 633-643.

Oberhofer, W. (1982), *The consistency of nonlinear regression minimizing the L_1-norm*, The Annals of Statistics 10, 316-319.

Seung Hoe Choi and Hae Kyung Kim. (1998). *Asymptotic Properties od LAD Estimators in Censored Nonlinear Regression Model*, Journal of the Korean Statistical Society, Vol. 27, No. 1, 101-112.

Walker, A.M. (1971), *On the estimation of a harmonic component in a time series with stationary independent residuals*, Biometrika, 58, 21-36.

White, H. (1980), *Nonlinear regression on cross-section data*, Econometrica 48, 721-746.

Wu, C. F. (1981), *Asymptotic theory of nonlinear least squares estimatiom*, The Annals of Statistics 9, 501-513.