On Least Absolute Deviation Estimators For One Dimensional Chirp Model

Ananya Lahiri[1] & Debasis Kundu[1,2] & Amit Mitra[1]

Abstract

It is well known that the least absolute deviation (LAD) estimators are more robust than the least squares estimators particularly in presence of heavy tail errors. We consider the LAD estimators of the unknown parameters of one dimensional chirp signal model under independent and identically distributed error structure. The proposed estimators are strongly consistent and it is observed that the asymptotic distribution of the LAD estimators are normally distributed. We perform some simulation studies to verify the asymptotic theory for small sample sizes and the performance are quite satisfactory.

Key Words and Phrases: Chirp signals; least absolute deviation estimators; strong consistency, asymptotic distribution.

[1] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India.

[2] Corresponding author, E-mail: kundu@iitk.ac.in, Phone: 91-512-2597141, Fax: 91-512-2597500.

1 Introduction

Let us consider the following chirp signal model;

$$y(n) = A^0 \cos(\alpha^0 n + \beta^0 n^2) + B^0 \sin(\alpha^0 n + \beta^0 n^2) + X(n); \quad n = 1, \cdots, N. \qquad (1)$$

Here $y(n)$ is the real valued signal observed at $n = 1, \cdots, N$. A^0, B^0 are amplitudes, and α^0 and β^0 are frequency and frequency rate respectively. The additive error $\{X(n)\}$ is a sequence of independent and identically distributed (*i.i.d.*) random variables with mean zero and finite second moment. The explicit assumptions on $X(n)$s will be provided later.

In signal processing literature, chirp signal models are used to detect an object with respect to a fixed receiver. Such models are typically one-dimensional chirp model as described in (1), where the dimension is usually the time. In this model, frequency varies with time in a non-linear fashion like a quadratic function and it is this property that has been exploited for measuring the distance of an object from a fixed receiver. In various areas of science and engineering, for example in sonar, radar and communications systems, such models are used. Oceanography and geology are some other areas where this model has been used quite extensively.

On this model (1) or on its variations, extensive work has been done by several authors, see for example Abatzoglou (1986), Kumaresan and Verma (1987), Djuric and Kay (1990), Gini, Montanari and Verrazani (2000),Saha and Kay (2002), Nandi and Kundu (2004), Kundu and Nandi (2008) and the references cited therein. Nandi and Kundu (2004) first established the consistency and asymptotic normality property of the LSE of the one dimensional (1D) chirp signal model for *i.i.d.* errors. The authors, see Kundu and Nandi (2008), extended the results when $X(n)$'s are obtained from a linear stationary processes. But there is no discussion about any method like least absolute deviation (LAD) estimation which is well known to be more robust than the LSEs, particularly in presence of outliers.

Unfortunately, the model does not satisfy the assumption B5 of Oberhofer (1982) and therefore the strong consistency of the LAD estimators in this case is not immediate. It may be mentioned that even the ordinary sinusoidal model does not satisfy the assumption B5 of Oberhofer (1982), and in that case Kim *et al.* (2000) provided the consistency and asymptotic normality results of the LAD estimators. The main aim of this paper is to provide the consistency and asymptotic normality properties of the LAD estimators of the unknown parameters of model (1).

It is known that the LSE of α^0 has the convergence rate $O_p(N^{-3/2})$, whereas the LSE of β^0 has the convergence rate $O_p(N^{-5/2})$, see Nandi and Kundu (2004). Here $z = O_p(N^{-\delta})$ means zN^{δ} is bounded in probability. In this paper it is observed that the LAD estimators of α^0 and β^0 have the same rates of convergence as the corresponding LSEs. But it is observed that asymptotic efficiency of LAD estimators relative to LSE is $4f(0)^2\sigma^2$, here $f(\cdot)$ is the probability density function (PDF) of the error random variable $X(n)$. Therefore it is clear that LAD estimators are more efficient than LSEs for heavy tailed error distributions. We perform some extensive simulation experiments to study the effectiveness of the LAD estimators for finite samples, and it is observed that the performances of the LAD estimators are quite satisfactory.

The rest of the paper is organized as follows. In Section 2, we mainly provide the model assumptions and methodology. In Section 3 the strong consistency and asymptotic normality of LAD estimators are provided. Numerical results are presented in Section 4, and finally we conclude the paper in Section 5.

2 Model Assumptions and Preliminary Results

2.1 Model Assumptions

We make the following assumptions on the error random variables.

Assumption 1: The error random variable $X(n)$ satisfies the following conditions; $\{X(n)\}$ is a sequence of *i.i.d.* absolute continuous random variables with mean zero, variance σ^2, and has the PDF $f(\cdot)$. It is further assumed that $f(\cdot)$ is symmetric and differentiable in $(0, \epsilon)$ and $(-\epsilon, 0)$ for some $\epsilon > 0$ and $f(0) > 0$.

We use the following notations; $F(\cdot)$ the cumulative distribution function corresponds to $f(\cdot)$. The parameter vector $\theta = (A, B, \alpha, \beta)$, the true parameter vector $\theta^0 = (A^0, B^0, \alpha^0, \beta^0)$, and the parameter space $\Theta = [-M, M] \times [-M, M] \times [0, \pi] \times [0, \pi]$.

Assumption 2: It is assumed that θ^0 is an interior point of Θ.

2.2 Least Absolute Deviation Estimation Procedure

In this section we propose the LAD estimation procedure to estimate the unknown parameters of the model (1). The LAD estimators are obtained by minimizing $Q(\theta)$, with respect to θ, where,

$$Q(\theta) = \sum_{n=1}^{N} \left| y(n) - \left(A\cos(\alpha n + \beta n^2) + B\sin(\alpha n + \beta n^2)\right)\right| \tag{2}$$

we note that

$$Q(A, B, \alpha, \beta) > Q(\widehat{A}(\alpha, \beta), \widehat{B}(\alpha, \beta), \alpha, \beta) > Q(\widehat{A}(\widehat{\alpha}, \widehat{\beta}), \widehat{B}(\widehat{\alpha}, \widehat{\beta}), \widehat{\alpha}, \widehat{\beta})$$

where $\widehat{A}(\alpha, \beta), \widehat{B}(\alpha, \beta)$ are the minimizer of $Q(A, B, \alpha, \beta)$ for known α, β and $\widehat{A}(\widehat{\alpha}, \widehat{\beta})$, $\widehat{B}(\widehat{\alpha}, \widehat{\beta})$ are the minimizer of $Q(A, B, \widehat{\alpha}, \widehat{\beta})$. Now

$$(\widehat{\alpha}, \widehat{\beta}) = \arg\min Q(\widehat{A}(\alpha, \beta), \widehat{B}(\alpha, \beta), \alpha, \beta).$$

So, LAD estimators of θ^0 will be $\widehat{\theta} = (\widehat{A}(\widehat{\alpha}, \widehat{\beta}), \widehat{B}(\widehat{\alpha}, \widehat{\beta}), \widehat{\alpha}, \widehat{\beta}) = (\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{\beta})$.

3 Asymptotic Properties of Least Absolute Deviation Estimators

3.1 Strong Consistency

Now we will provide the consistency results for the proposed estimators.

Theorem 1. *If the Assumptions 1-2 are satisfied then* $\left(\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{\beta}\right)$ *is a strongly consistent estimator of* $(A^0, B^0, \alpha^0, \beta^0)$.

We need the following results to prove Theorem 1.

Lemma 1. *If* (θ_1, θ_2) *in* $(0, \pi) \times (0, \pi)$, $t = 0, 1, 2$ *then except for countable number of points the followings are true.*

(i)

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \cos(\theta_1 n + \theta_2 n^2) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \sin(\theta_1 n + \theta_2 n^2) = 0. \tag{3}$$

(ii)

$$\lim_{N\to\infty} \frac{1}{N^{t+1}} \sum_{n=1}^{N} n^t \cos^2(\theta_1 n + \theta_2 n^2) = \frac{1}{2(t+1)} \tag{4}$$

$$\lim_{N\to\infty} \frac{1}{N^{t+1}} \sum_{n=1}^{N} n^t \sin^2(\theta_1 n + \theta_2 n^2) = \frac{1}{2(t+1)}. \tag{5}$$

$$\lim_{N\to\infty} \frac{1}{N^{t+1}} \sum_{n=1}^{N} n^t \sin(\theta_1 n + \theta_2 n^2) \cos(\theta_1 n + \theta_2 n^2) = 0. \tag{6}$$

Proof: Using the result of Vinogradov (1954) Lemma 1 can be easily established.

Lemma 2. *If,* $D(\theta) = Q(\theta) - Q(\theta^0)$, *then*

$$\frac{1}{N} D(\theta) - \lim_{N\to\infty} E[\frac{1}{N} D(\theta)] \to 0 \ a.s. uniformly \ \forall \ \theta \in \Theta.$$

Proof: Let us denote $W_n(\theta) = |h_n(\theta) + X(n)| - |X(n)|$. Then $\frac{1}{N} D(\theta) = \frac{1}{N} \sum_{n=1}^{N} W_n(\theta)$. We note that $W_n(\theta) = |h_n(\theta) + X(n)| - |X(n)| \leq |h_n(\theta)| \leq 4M$, as the parameter space is compact. Also $W_n(\theta)$s are independent and non identically distributed random variables with $E[W_n(\theta)] < \infty$ and $V[W_n(\theta)] < \infty$. It may be easily seen similarly as in Oberhofer (1982) that these bounds do not depend on n.

Since Θ is a compact set, there exists $\Theta_1, \cdots, \Theta_K$, such that $\Theta = \cup_{i=1}^K \Theta_i$ and on each Θ_i, $\sup_{\theta\in\Theta_i} W_n(\theta) - \inf_{\theta\in\Theta_i} W_n(\theta) < \frac{\epsilon}{4^n}$. *a.s.* Now for $\theta \in \Theta_i$,

$$\begin{aligned}
&\frac{1}{N}D(\theta) - \lim_{N\to\infty} E[\frac{1}{N}D(\theta)] \\
= &\left[\frac{1}{N}\sum_{n=1}^N W_n(\theta) - \frac{1}{N}\sum_{n=1}^N E\sup_{\theta\in\Theta_i} W_n(\theta)\right] + \left[\frac{1}{N}\sum_{n=1}^N E\sup_{\theta\in\Theta_i} W_n(\theta) - \lim_{N\to\infty} E[\frac{1}{N}D(\theta)]\right] \\
= &\mathfrak{A}(\theta) + \mathfrak{B}(\theta)
\end{aligned}$$

$$\begin{aligned}
\text{where,} \qquad \mathfrak{A}(\theta) &= \frac{1}{N}\sum_{n=1}^N W_n(\theta) - \frac{1}{N}\sum_{n=1}^N E\sup_{\theta\in\Theta_i} W_n(\theta) \\
&\le \sup_{\theta\in\Theta_i}[\frac{1}{N}\sum_{n=1}^N W_n(\theta)] - \frac{1}{N}\sum_{n=1}^N E\sup_{\theta\in\Theta_i} W_n(\theta) \\
&\le \frac{1}{N}\sum_{n=1}^N \sup_{\theta\in\Theta_i} W_n(\theta) - \frac{1}{N}\sum_{n=1}^N E\sup_{\theta\in\Theta_i} W_n(\theta).
\end{aligned}$$

Note that $\sup_{\theta\in\Theta_i} W_n(\theta)$'s are independent and non identically distributed random variables with finite mean and variance, and the variance is bounded by a quantity not depending on n. Applying Kolmogorov's strong law of large numbers, choose N_{0i} large enough, so that for $N \ge N_{0i}$, $\mathfrak{A}(\theta) < \frac{\epsilon}{3}$ *a.s.*, uniformly $\forall\ \theta \in \Theta_i$. Now

$$\begin{aligned}
\mathfrak{B}(\theta) &= \frac{1}{N}\sum_{n=1}^N E\sup_{\theta\in\Theta_i} W_n(\theta) - \lim_{N\to\infty} E[\frac{1}{N}D(\theta)] \\
&= \frac{1}{N}\sum_{n=1}^N E\sup_{\theta\in\Theta_i} W_n(\theta) - E\lim_{N\to\infty}[\frac{1}{N}D(\theta)] \qquad \text{using DCT} \\
&= \mathfrak{C}(\theta) + \mathfrak{D}(\theta),
\end{aligned}$$

where

$$\mathfrak{C}(\theta) = \frac{1}{N}\sum_{n=1}^N E\sup_{\theta\in\Theta_i} W_n(\theta) - E\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N \sup_{\theta\in\Theta_i} W_n(\theta),$$

and DCT stands for dominated convergence theorem. We take $U_N(\theta) = \frac{1}{N}\sum_{n=1}^N E\sup_{\theta\in\Theta_i} W_n(\theta)$ and we want to apply DCT to pass the limit inside the expectation and we get N_{*i}

such that $\mathfrak{C}(\theta) < \frac{\epsilon}{3}$. Further, note that

$$\begin{aligned}
\mathfrak{D}(\theta) &= E \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \sup_{\theta\in\Theta_i} W_n(\theta) - E \lim_{N\to\infty} [\frac{1}{N} D(\theta)] \\
&= E \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \sup_{\theta\in\Theta_i} W_n(\theta) - E \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} W_n(\theta) \\
&\leq E \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \sup_{\theta\in\Theta_i} W_n(\theta) - E \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \inf_{\theta\in\Theta_i} W_n(\theta) \\
&\leq \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \frac{\epsilon}{4^n} = 0.
\end{aligned}$$

Combining we get $\frac{1}{N} D(\theta) - \lim_{N\to\infty} E[\frac{1}{N} D(\theta)] \to 0$ $a.s.$uniformly $\forall\ \theta \in \Theta$. ∎

Lemma 3. *The global minimum of* $\lim_{N\to\infty} E[\frac{1}{N} D(\theta)]$ *is attained at* θ^0.

Proof: At θ^0 the value of $\lim_{N\to\infty} E[\frac{1}{N} D(\theta)]$ is zero, and for $\theta \neq \theta^0$, if we can show $\lim_{N\to\infty} E[\frac{1}{N} D(\theta)] > 0$ then we are through. To achieve that we verify the assumptions **B7**, **B8**, **B9** of Lemma 4 by Oberhofer (1982). For convenience we reproduce the assumptions **B7**, **B8**, **B6** as **A1**, **A2**, **A3** respectively, below.

A1: For every closed set Θ_0 not containing θ^0, there exist numbers $c > 0,\ d > 0,\ \mathfrak{N}_0 > 0$ such that for all $\theta \in \Theta_0$ and all $N \geq \mathfrak{N}_0$, $|\{n :\ n \leq \mathfrak{N}_0, |h_n(\theta)| \geq c\}|/N \geq d > 0$.

A2: For every $c > 0$, there exists a real number $\mathfrak{d} > 0$, such that for all n $\min[F_n(c) - 1/2, 1/2 - F_n(-c)] \geq \mathfrak{d} > 0$

A3: There exists $\mathfrak{e} > 0$ and $\mathfrak{N}_0$ such that for all $N \geq \mathfrak{N}_0$,

$$\mathfrak{Q} = \inf_{\Theta_0} \frac{1}{N} \sum_{n=1}^{N} |h_n(\theta)| \min[F_n(c) - 1/2, 1/2 - F_n(-c)] \geq \mathfrak{e} > 0$$

Lemma 4 of Oberhofer (1982) states that **A3** is fulfilled if **A1** and **A2** holds. Note that Lemma 2 of Oberhofer (1982) gives $\frac{1}{N} D(\theta) \geq \mathfrak{Q}$. Then it is enough to show $\lim_{N\to\infty} E[\frac{1}{N} D(\theta)] \geq \lim_{N\to\infty} \mathfrak{Q} > 0$. Now, $\lim_{N\to\infty} \mathfrak{Q} > 0$ condition is same as **A3**. Using Lemma 4 of Oberhofer (1982) instead of **A3** we try to show **A1** and **A2**. If $f(0) > 0$

then **A2** is automatically satisfied. It remains to show that **A1** is satisfied in our case. If there exists $\mathfrak{c} > 0$ such that $\inf_{\Theta_0} \frac{1}{N} \sum_{n=1}^{N} |h_n(\theta)| \geq \mathfrak{c} > 0$ for all $N \geq \mathfrak{N}_0$ then **A1** will be satisfied. Let us consider

$$\Theta_0 = S_c = \{\theta : |\theta - \theta^0| \geq 3c > 0\} \subseteq S_c^A \cup S_c^B \cup S_c^{(\alpha,\beta)} \tag{7}$$

where

$$\begin{aligned} S_c^A &= \{\theta : |A - A^0| \geq c > 0\} \subseteq \{\theta : |A - A^0| \geq c, (\alpha, \beta) = (\alpha^0, \beta^0)\} \\ &\quad \cup \{\theta : |A - A^0| \geq c, (\alpha, \beta) \neq (\alpha^0, \beta^0)\}, \end{aligned}$$

$$\begin{aligned} S_c^B &= \{\theta : |B - B^0| \geq c > 0\} \subseteq \{\theta : |B - B^0| \geq c, (\alpha, \beta) = (\alpha^0, \beta^0)\} \\ &\quad \cup \{\theta : |B - B^0| \geq c, (\alpha, \beta) \neq (\alpha^0, \beta^0)\}, \end{aligned}$$

$$S_c^{(\alpha,\beta)} = \{\theta : |(\alpha, \beta) - (\alpha^0, \beta^0)| \geq c > 0\}$$

Now on the set $\{\theta : |A - A^0| \geq c, (\alpha, \beta) = (\alpha^0, \beta^0)\}$

Case-1, if $B - B^0 = 0$, then

$$h_n(\theta) = (A - A^0) \cos(\alpha^0 n + \beta^0 n^2)$$

$$\begin{aligned} \text{and,} \qquad \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} |h_n(\theta)| &= |A - A^0| \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} |\cos(\alpha^0 n + \beta^0 n^2)| \\ &\geq |A - A^0| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} (\cos(\alpha^0 n + \beta^0 n^2))^2 \\ &= |A - A^0| \frac{1}{2} \geq \frac{c}{2} > 0 \end{aligned}$$

Case-2, if $B - B^0 \neq 0$, then

$$\begin{aligned} h_n(\theta) &= (A - A^0) \cos(\alpha^0 n + \beta^0 n^2) + (B - B^0) \sin(\alpha^0 n + \beta^0 n^2) \\ &= r \cos(\omega) \cos(\alpha^0 n + \beta^0 n^2) + r \sin(\omega) \sin(\alpha^0 n + \beta^0 n^2) \text{ for some } r > 0,\ \omega \\ &= r \cos(\alpha^0 n + \beta^0 n^2 - \omega) \end{aligned}$$

$$\text{So,} \qquad \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} |h_n(\theta)| = |r| \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} |\cos(\alpha^0 n + \beta^0 n^2 - \omega)|$$
$$\geq |r| \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} (\cos(\alpha^0 n + \beta^0 n^2 - \omega))^2$$
$$= |r|\frac{1}{2} > 0$$

On the set $\{\theta : |A - A^0| \geq c, (\alpha, \beta) \neq (\alpha^0, \beta^0)\}$

$$\begin{aligned} h_n(\theta) &= A\cos(\alpha n + \beta n^2) + B\sin(\alpha n + \beta n^2) \\ &- A^0\cos(\alpha^0 n + \beta^0 n^2) - B^0 \sin(\alpha^0 n + \beta^0 n^2) \\ &= r\cos(\alpha n + \beta n^2 - \omega) - r^0 \cos(\alpha^0 n + \beta^0 n^2 - \omega^0) \text{ for some } r, r^0 > 0,\ \omega, \omega^0 \end{aligned}$$

We recall that $|h_n(\theta)| \leq 4M$. Then $\left(\frac{h_n(\theta)}{4M}\right)^2 \leq \left|\frac{h_n(\theta)}{4M}\right| < 1$. We denote $\frac{r}{4M} = R > 0$ and $\frac{r^0}{4M} = R^0 > 0$. Then

$$\begin{aligned} \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} |h_n(\theta)| &= 4M \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} |\frac{h_n(\theta)}{4M}| \\ &\geq 4M \liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \left(\frac{h_n(\theta)}{4M}\right)^2 \\ &= 4M \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} [R\cos(\alpha n + \beta n^2 - \omega) - R^0 \cos(\alpha^0 n + \beta^0 n^2 - \omega^0)]^2 \\ &= 4M\frac{R^2 + R^{02}}{2} > 0 \end{aligned}$$

Similarly on other sets $\liminf_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} |h_n(\theta)| > 0$. So, we get $\lim_{N\to\infty} E[\frac{1}{N} D(\theta)] > 0$ for $\theta \neq \theta^0$. ∎

Proof of Theorem 1:

Now to prove the strong consistency of the LAD estimators, first let us observe that the minimizer of $Q(\theta)$ will be same as the minimizer of $D(\theta) = Q(\theta) - Q(\theta^0)$. So we develop our result based on minimizer of $D(\theta)$ instead of $Q(\theta)$. Note that

$$Q(\theta) = \sum_{n=1}^{N} \left| y(n) - \left(A\cos(\alpha n + \beta n^2) + B\sin(\alpha n + \beta n^2)\right)\right| = \sum_{n=1}^{N} |h_n(\theta) + X(n)| \tag{8}$$

where

$$h_n(\theta) = A^0\cos(\alpha^0 n+\beta^0 n^2)+B^0\sin(\alpha^0 n+\beta^0 n^2)-A\cos(\alpha n+\beta n^2)-B\sin(\alpha n+\beta n^2)$$

and note that $|h_n(\theta)| \le 4M$ for $\theta \in \Theta$ and $Q(\theta^0) = \sum_{n=1}^{N} |X(n)|$. In Lemma 2 we have shown that

$$\frac{1}{N}D(\theta) - \lim_{N\to\infty} E[\frac{1}{N}D(\theta)] \to 0 \ a.s. \qquad \text{uniformly} \ \forall\ \theta \in \Theta.$$

and in Lemma 3 we have shown that θ^0 is the global minimizer of $\lim_{N\to\infty} E\left[\frac{1}{N}D(\theta)\right]$. Therefore, by Lemma 2 of Jennrich (1969) or by Lemma 2.2 of White (1980) we can conclude that minimizer of $D(\theta)$ is a strong consistent estimators of θ^0. ∎

3.2 Asymptotic Normality

Now we want to show that the estimators obtained have the following asymptotic normality result. Let us take $D = diag\left(\frac{1}{\sqrt{N}}, \frac{1}{\sqrt{N}}, \frac{1}{N\sqrt{N}}, \frac{1}{N^2\sqrt{N}}\right)$

Theorem 2. *If the Assumptions 1-2 are satisfied then*

$$(\widehat{\theta} - \theta^0)D^{-1} \xrightarrow{d} N_4\left(0, \frac{1}{f(0)^2}\Sigma\right) \tag{9}$$

$$\Sigma = \frac{1}{A^{0^2}+B^{0^2}}\begin{bmatrix} \frac{1}{2}\left(A^{0^2}+9B^{0^2}\right) & -4A^0B^0 & -18B^0 & 15B^0 \\ -4A^0B^0 & \frac{1}{2}\left(9A^{0^2}+B^{0^2}\right) & 18A^0 & -15A^0 \\ -18B^0 & 18A^0 & 96 & -90 \\ 15B^0 & -15A^0 & -90 & 90 \end{bmatrix}, \tag{10}$$

here, ' $\xrightarrow{d}$ *' means converges in distribution,*

Proof:

We recall that $Q(\theta)$ is not a differentiable function, to find the asymptotic distribution of $\widehat{\theta}$, we want to approximate $Q(\theta)$ by $\widetilde{Q}(\theta)$ with some "nice" property (differentiability). For that purpose we need to approximate $|x|$ by some "nice" function $\rho_N(x)$ near zero, such that $\lim_{N\to\infty} \rho_N(x) = |x|$. Let us consider the interval near zero as $(-\frac{1}{\gamma_N}, \frac{1}{\gamma_N})$ where γ_N is an increasing function of N satisfying $\lim_{N\to\infty} \dfrac{1}{\gamma_N} = 0$. Let us approximate $|x|$ by a polynomial. We want to approximate $|x|$ separately in $(-\frac{1}{\gamma_N}, 0)$ and $(0, \frac{1}{\gamma_N})$. In each of these intervals we observe that the degree of the polynomial has to be at least 3 to make the approximating function twice continuously differentiable. If possible the degree of the polynomial is less than 3, say 2 and it is $\mathcal{P}(x) = \mathcal{A}x^2 + \mathcal{B}x + \mathcal{C}$. Then $\mathcal{P}''(\frac{1}{\gamma_N}) = 2\mathcal{A}$ should match with the second derivative of $|x|$ at boundary point $\frac{1}{\gamma_N}$, which is zero. In that case $\mathcal{A} = 0$ makes polynomial degree 1, if not then there will be a jump discontinuity at $\frac{1}{\gamma_N}$ for the function $\mathcal{P}''(x)$. So, let the approximating polynomial is $\mathcal{P}(x) = \mathcal{A}x^3 + \mathcal{B}x^2 + \mathcal{C}x + \mathcal{D}$ in $(0, \frac{1}{\gamma_N})$. As $|x|$ is symmetric about zero the approximating polynomial in $(-\frac{1}{\gamma_N}, 0)$ will be $\mathcal{P}(x) = -\mathcal{A}x^3 + \mathcal{B}x^2 - \mathcal{C}x + \mathcal{D}$. Now to find the coefficients of the polynomial we match the function value and its derivatives at the joining points. $\mathcal{P}(\frac{1}{\gamma_N}) = |\frac{1}{\gamma_N}|$ gives

$$\frac{\mathcal{A}}{\gamma_N^3} + \frac{\mathcal{B}}{\gamma_N^2} + \frac{\mathcal{C}}{\gamma_N} + \mathcal{D} = \frac{1}{\gamma_N} \tag{11}$$

$\mathcal{P}'(\frac{1}{\gamma_N}) = 1$ gives

$$\frac{3\mathcal{A}}{\gamma_N^2} + \frac{2\mathcal{B}}{\gamma_N} + \mathcal{C} = 1 \tag{12}$$

$\mathcal{P}''(\frac{1}{\gamma_N}) = 0$ gives

$$\frac{6\mathcal{A}}{\gamma_N} + 2\mathcal{B} = 0 \tag{13}$$

and $\mathcal{P}'(0)$ agrees from both parts of the polynomial giving

$$\mathcal{C} = 0. \tag{14}$$

Solving previous four equations we get the suitable cubic spline as

$$\rho_N(x) = \left[-\frac{1}{3}\gamma_N^2 x^3 + \gamma_N x^2 + \frac{1}{3\gamma_N} \right] I_{\left(0 < x \leq \frac{1}{\gamma_N}\right)} + x I_{\left(x > \frac{1}{\gamma_N}\right)}$$

$$\rho_N(-x) = -\rho_N(x)$$

which is symmetric, twice continuously differentiable and γ_N is an increasing function of N satisfying some extra conditions, $N^2 = o(\gamma_N^3)$, $\gamma_N = o(N)$ and $\sum_{N=1}^{\infty} \dfrac{1}{\gamma_N^2} < \infty$, which we will be needing later. After getting the nice function $\rho_N(x)$ we now define

$$\widetilde{Q}(\theta) = \sum_{n=1}^{N} \rho_N(h_n(\theta) + X(n)) \tag{15}$$

and note that $\widetilde{Q}(\theta^0) = \sum_{n=1}^{N} \rho_N(X(n))$. Now we want to prove the following two results (Lemma 4 and Lemma 5) which when combined will give the required asymptotic normality result.

Lemma 4. *If the Assumptions 1-2 are satisfied then* $(\widehat{\theta} - \widetilde{\theta})D^{-1} \xrightarrow{P} 0$ *where* $\xrightarrow{P}$ *means convergence in probability.*

Lemma 5. *If the Assumptions 1-2 are satisfied, then* $\widetilde{\theta}$*, the minimizer of* $\widetilde{Q}(\theta)$ *has the following asymptotic distribution* $(\widetilde{\theta} - \theta^0)D^{-1} \xrightarrow{d} N_4(0, \frac{1}{f(0)^2}\Sigma)$

To prove Lemma 4 and Lemma 5 we need some more lemmas.

Lemma 6. $\sup_{\theta\in\Theta}(\widetilde{Q}(\theta) - Q(\theta)) = o_P(1)$ *and* $\sup_{\theta\in\Theta} \frac{1}{N}|\widetilde{Q}(\theta) - Q(\theta)| \to 0$ *a.s. where* $o_P(1)$ *means converges to zero in probability.*

Proof: To calculate the following quantity $\widetilde{Q}(\theta) - Q(\theta)$. we write explicitly the function $\rho_N(x) - |x|$.

$$\begin{aligned}\rho_N(x) - |x| \quad = \quad & \left[-\frac{1}{3}\gamma_N^2 x^3 + \gamma_N x^2 - x + \frac{1}{3\gamma_N}\right] I_{\left(0<x\leq\frac{1}{\gamma_N}\right)} \\ & + \left[\frac{1}{3}\gamma_N^2 x^3 + \gamma_N x^2 + x + \frac{1}{3\gamma_N}\right] I_{\left(-\frac{1}{\gamma_N}\leq x\leq 0\right)}\end{aligned}$$

and we note that $|\rho_N(x) - |x|| \leq \frac{C_1}{\gamma_N}$. Now

$$\begin{aligned}
P(|\widetilde{Q}(\theta) - Q(\theta)| > \epsilon) &\leq \frac{E|\widetilde{Q}(\theta) - Q(\theta)|}{\epsilon} \\
&\leq \frac{C_1}{\gamma_N}\sum_{n=1}^{N} EI_{\left(0<|h_n(\theta)+X(n)|\leq\frac{1}{\gamma_N}\right)} \\
&= \frac{C_1}{\gamma_N}\sum_{n=1}^{N} P\left(0 < |h_n(\theta) + X(n)| \leq \frac{1}{\gamma_N}\right) \\
&= \frac{C_1}{\gamma_N}\sum_{n=1}^{N} P(-h_n(\theta) - \frac{1}{\gamma_N} \leq X(n) \leq -h_n(\theta) + \frac{1}{\gamma_N}) \\
&= \frac{C_1}{\gamma_N}\sum_{n=1}^{N} F(-h_n(\theta) + \frac{1}{\gamma_N}) - F(-h_n(\theta) - \frac{1}{\gamma_N}) \\
&= C_2\frac{1}{\gamma_N^2}\sum_{n=1}^{N} f(\widetilde{h_n}(\theta)) \text{ using mean value theorem} \\
&\leq C_3\frac{N}{\gamma_N^2} \quad \to 0 \quad \text{as } N \to \infty.
\end{aligned}$$

So, we get $\widetilde{Q}(\theta) - Q(\theta) = o_P(1)$ and hence $\sup_{\theta\in\Theta}(\widetilde{Q}(\theta) - Q(\theta)) = o_P(1)$ as Θ is compact. Also we note that $P(\frac{1}{N}|\widetilde{Q}(\theta) - Q(\theta)| > \epsilon) \leq \frac{C_3}{\gamma_N^2}$ and

$$\sum_{N=1}^{\infty} P(\frac{1}{N}|\widetilde{Q}(\theta) - Q(\theta)| > \epsilon) \leq C_2 \sum_{N=1}^{\infty} \frac{1}{\gamma_N^2} < \infty$$

implies $\frac{1}{N}|\widetilde{Q}(\theta) - Q(\theta)| \to 0$ $a.s.$ and hence $\sup_{\theta\in\Theta} \frac{1}{N}|\widetilde{Q}(\theta) - Q(\theta)| \to 0$ $a.s.$ ■

Lemma 7. *$\widetilde{\theta}$, the minimizer of $\widetilde{Q}(\theta)$ is strong consistent estimator of θ^0.*

Proof: We take $\widetilde{W}_n(\theta) = \rho_N(h_n(\theta) + X(n)) - \rho_N(X(n))$. Then $\widetilde{D}(\theta) = \frac{1}{N}\sum_{n=1}^{N}\widetilde{W}_n(\theta)$ and $\widetilde{D}(\theta) = \widetilde{Q}(\theta) - \widetilde{Q}(\theta^0)$. As before $\widetilde{\theta}$ is also minimizer of $\widetilde{D}(\theta)$. Then we proceed with exactly same technique as that used for proving strong consistency of $\hat{\theta}$ and we finally get $\frac{1}{N}\widetilde{D}(\theta) - \lim_{N\to\infty} E[\frac{1}{N}\widetilde{D}(\theta)] \to 0$ $a.s.$uniformly $\forall\, \theta \in \Theta$. Now at θ^0 the value of $\lim_{N\to\infty} E[\frac{1}{N}\widetilde{D}(\theta)]$ is zero. And for $\theta \neq \theta^0$,

$$\begin{aligned}
& \lim_{N\to\infty} E[\frac{1}{N}\widetilde{D}(\theta)] \\
= & \lim_{N\to\infty} E[\frac{1}{N}\widetilde{D}(\theta) - \frac{1}{N}D(\theta)] + \lim_{N\to\infty} E[\frac{1}{N}D(\theta)] \\
= & \lim_{N\to\infty} E[\frac{1}{N}\widetilde{Q}(\theta) - \frac{1}{N}Q(\theta)] - \lim_{N\to\infty} E[\frac{1}{N}\widetilde{Q}(\theta^0) - \frac{1}{N}Q(\theta^0)] + \lim_{N\to\infty} E[\frac{1}{N}D(\theta)]
\end{aligned}$$

The first two terms converge to zero, using Lemma 6 we get $\lim_{N\to\infty} E[\frac{1}{N}\widetilde{D}(\theta)] > 0$ for $\theta \neq \theta^0$. So, $\widetilde{\theta}$ is strong consistent estimator of θ^0. ∎

Let us denote $\widetilde{Q}'(\theta)$ as the 4×1 first derivative vector and $\widetilde{Q}''(\theta)$ as the 4×4 second derivative matrix of $\widetilde{Q}(\theta)$. To get explicit expressions of $\widetilde{Q}'(\theta)$ and $\widetilde{Q}''(\theta)$, let us write explicitly the functions $\rho_N(x)$, $\rho'_N(x)$ and $\rho''_N(x)$.

$$\begin{aligned}
\rho_N(x) &= \left[-\frac{1}{3}\gamma_N^2 x^3 + \gamma_N x^2 + \frac{1}{3\gamma_N}\right] I_{\left(0<x\leq\frac{1}{\gamma_N}\right)} + xI_{\left(x>\frac{1}{\gamma_N}\right)} \\
&+ \left[\frac{1}{3}\gamma_N^2 x^3 + \gamma_N x^2 + \frac{1}{3\gamma_N}\right] I_{\left(-\frac{1}{\gamma_N}\leq x\leq 0\right)} - xI_{\left(x<-\frac{1}{\gamma_N}\right)}
\end{aligned}$$

$$\begin{aligned}
\rho'_N(x) &= \left[-\gamma_N^2 x^2 + 2\gamma_N x\right] I_{\left(0<x\leq\frac{1}{\gamma_N}\right)} + 1I_{\left(x>\frac{1}{\gamma_N}\right)} \\
&+ \left[\gamma_N^2 x^2 + 2\gamma_N x\right] I_{\left(-\frac{1}{\gamma_N}\leq x\leq 0\right)} - 1I_{\left(x<-\frac{1}{\gamma_N}\right)}
\end{aligned}$$

$$\rho''_N(x) = \left[-2\gamma_N^2 x + 2\gamma_N\right] I_{\left(0<x\leq\frac{1}{\gamma_N}\right)} + \left[2\gamma_N^2 x + 2\gamma_N\right] I_{\left(-\frac{1}{\gamma_N}\leq x\leq 0\right)}$$

Lemma 8. *$D\widetilde{Q}''(\theta^0)D$ converges to $2f(0)\Sigma_1$ which is a positive definite matrix, in probability.*

PROOF: First we note that $Q''(\theta^0)$ depends on N.

STEP-1 Let us calculate the quantity $E\rho''_N(X(n))$. We want to show

$$E[\frac{1}{N}\sum_{n=1}^{N}\rho''_N(X(n))] = 2f(0) + o(1)$$

Also we recall that $X(n)$'s are *i.i.d.* and have symmetric density function f with $f(0) < \infty$. Now f is differentiable in $(0, \frac{1}{\gamma_N})$ and $(-\frac{1}{\gamma_N}, 0)$ for sufficiently large N.

Note that in that case f' is bounded in $(0, \frac{1}{\gamma_N})$, say less than M'.

$$\begin{aligned} E[\rho_N''(X(n))] &= \int_0^{\frac{1}{\gamma_N}} \left[-2\gamma_N^2 x + 2\gamma_N\right] f(x)dx + \int_{-\frac{1}{\gamma_N}}^{0} \left[2\gamma_N^2 x + 2\gamma_N\right] f(x)dx \\ &= 2\int_0^{\frac{1}{\gamma_N}} \left[-2\gamma_N^2 x + 2\gamma_N\right] f(x)dx \\ &= 4\int_0^{\frac{1}{\gamma_N}} \left[-\gamma_N^2 x + \gamma_N\right] f(x)dx \end{aligned}$$

Integration by parts gives this is equal to,

$$\begin{aligned} & 4\left[f(x)\int(-\gamma_N^2 x + \gamma_N)dx - \int f'(x)[-\gamma_N^2\frac{x^2}{2} + \gamma_N x]dx\right]_0^{\frac{1}{\gamma_N}} \\ = \;& 4[\frac{1}{2}f(\frac{1}{\gamma_N})] - R_N = 2f(0) + 2(f(\frac{1}{\gamma_N}) - f(0)) - R_N \end{aligned}$$

where,

$$\begin{aligned} |R_N| &= 4\Big|\left[\int f'(x)[\gamma_N^2\frac{x^2}{2} - \gamma_N x]dx\right]_0^{\frac{1}{\gamma_N}}\Big| \\ &\leq 4M'\Big|\left[\int [\gamma_N^2\frac{x^2}{2} - \gamma_N x]dx\right]_0^{\frac{1}{\gamma_N}}\Big| \\ &= 4M'\Big|\left[\gamma_N^2\frac{x^3}{6} - \gamma_N\frac{x^2}{2}\right]_0^{\frac{1}{\gamma_N}}\Big| \\ &\leq 4M'\Big|\left[\frac{1}{6\gamma_N} + \frac{1}{2\gamma_N}\right]\Big| \to 0 \end{aligned}$$

and $2(f(\frac{1}{\gamma_N}) - f(0)) \to 0$ as $N \to \infty$.

Step-2 Next we want to show

$$V[\frac{1}{N}\sum_{n=1}^{N}\rho_N''(X(n))] = o(1),$$

using Step-1, which is equivalent to

$$E\left[\frac{1}{N}\sum_{n=1}^{N}\rho_N''(X(n))\right]^2 = 4f^2(0) + o(1).$$

For variance calculation let us write the expression for $[\rho_N''(x)]^2$.

$$\begin{aligned} [\rho_N''(x)]^2 &= \left[4\gamma_N^4 x^2 - 8\gamma_N^3 x + 4\gamma_N^2\right] I_{\left(0<x\leq\frac{1}{\gamma_N}\right)} \\ &+ \left[4\gamma_N^4 x^2 + 8\gamma_N^3 x + 4\gamma_N^2\right] I_{\left(-\frac{1}{\gamma_N}\leq x\leq 0\right)} \end{aligned}$$

which is an even function. Then,

$$
\begin{aligned}
& E\left[\rho_N''(X(n))\right]^2 \\
= & \int_0^{\frac{1}{\gamma_N}} \left[4\gamma_N^4 x^2 - 8\gamma_N^3 x + 4\gamma_N^2\right] f(x)dx + \int_{-\frac{1}{\gamma_N}}^0 \left[4\gamma_N^4 x^2 + 8\gamma_N^3 x + 4\gamma_N^2\right] f(x)dx \\
= & \; 2\int_0^{\frac{1}{\gamma_N}} \left[4\gamma_N^4 x^2 - 8\gamma_N^3 x + 4\gamma_N^2\right] f(x)dx \\
\leq & \; 8\int_0^{\frac{1}{\gamma_N}} |\gamma_N^4 x^2 + \gamma_N^2| dx \\
\leq & \; C_0 \gamma_N.
\end{aligned}
$$

Therefore, using the above and independence of $X(n)$ we get

$$E[\frac{1}{N}\sum_{n=1}^{N}\rho_N''(X(n))]^2 = 4f^2(0) + o(1)$$

Step-3: Let us consider the (1,1)-th element of $\lim_{N\to\infty} D\widetilde{Q}''(\theta^0)D$, which is

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\rho_N''(X(n))\cos^2(\alpha^0 n + \beta^0 n^2),$$

Using Step-1 and Step-2, and by applying Chebychev's inequality we get

$$D\widetilde{Q}''(\theta^0)D \overset{P}{=} f(0)\begin{bmatrix} 1 & 0 & \frac{B^0}{2} & \frac{B^0}{3} \\ 0 & 1 & -\frac{A^0}{2} & -\frac{A^0}{3} \\ \frac{B^0}{2} & -\frac{A^0}{2} & \frac{A^{02}+B^{02}}{3} & \frac{A^{02}+B^{02}}{4} \\ \frac{B^0}{3} & -\frac{A^0}{3} & \frac{A^{02}+B^{02}}{4} & \frac{A^{02}+B^{02}}{5} \end{bmatrix} + o_P(1) = 2f(0)\Sigma_1 + o_P(1)$$

Then $D\widetilde{Q}''(\theta^0)D$ converges to $2f(0)\Sigma_1$, which is a positive definite matrix, in probability. ∎

Lemma 9. *If $\breve{\theta}$ is a function of $X(1),\cdots,X(N)$, such that $\breve{\theta}\to\theta^0$ a.s. as $N\to\infty$ then, $D[\widetilde{Q}''(\breve{\theta}) - \widetilde{Q}''(\theta^0)]D \to 0$ a.s.*

Proof: To calculate $D[\widetilde{Q}''(\breve{\theta}) - \widetilde{Q}''(\theta^0)]$ let us consider

$$\frac{1}{N}\sum_{n=1}^{N}[\rho_N''(h_n(\breve{\theta}) + X(n)) - \rho_N''(X(n))].$$

Now note that as $N\to\infty$, $\breve{\theta}\to\theta^0$ $a.s.$ This implies $h_n(\breve{\theta})\to 0$ $a.s.\ \forall n$ as $h_n(\theta)$ is a continuous function of θ which implies for fixed n, $\lim_{k\to\infty} P(\cap_{n=k}^{\infty}|h_n(\breve{\theta})| < \epsilon_1) = 1\ \forall\epsilon_1$.

Now $h_n(\breve{\theta}) + X(n) \rightarrow \; X(n) \;\; a.s. \; \forall n$ and $\rho''_N(.)$ is a continuous function, then given $\epsilon_2 > 0 \; \exists \; \epsilon_1 > 0$ such that for fixed n,

$$|h_n(\breve{\theta})| < \epsilon_1 \quad \Rightarrow \; |\rho''_N(h_n(\breve{\theta}) + X(n)) - \rho''_N(X(n))| < \epsilon_2 \; .$$

which implies, $|\rho''_N(h_n(\breve{\theta}) + X(n)) - \rho''_N(X(n))| \rightarrow \; 0 \; a.s.$ And using this fact we get, $D[\widetilde{Q}''(\breve{\theta}) - \widetilde{Q}''(\theta^0)]D \rightarrow 0 \; a.s.$ ∎

Proof of Lemma 4

Step-1 By definition of $\widehat{\theta}$, $Q(\widetilde{\theta}) - Q(\widehat{\theta}) > 0$. Adding to both sides, $\widetilde{Q}(\widehat{\theta}) - \widetilde{Q}(\widetilde{\theta})$, which is again > 0 by definition of $\widetilde{\theta}$, we get,

$$\widetilde{Q}(\widehat{\theta}) - Q(\widehat{\theta}) + Q(\widetilde{\theta}) - \widetilde{Q}(\widetilde{\theta}) > \widetilde{Q}(\widehat{\theta}) - \widetilde{Q}(\widetilde{\theta}) \tag{16}$$

By Lemma 6 the left hand side of (16) is $o_P(1)$. So is the right hand side. *i.e.* $\widetilde{Q}(\widehat{\theta}) - \widetilde{Q}(\widetilde{\theta}) = o_P(1)$

Step-2 Now by Taylor series expansion of $\widetilde{Q}$ around $\widetilde{\theta}$

$$\widetilde{Q}(\widehat{\theta}) - \widetilde{Q}(\widetilde{\theta}) = (\widehat{\theta} - \widetilde{\theta})\widetilde{Q}'(\widetilde{\theta}) + \frac{1}{2}(\widehat{\theta} - \widetilde{\theta})\widetilde{Q}''(\theta^*)(\widehat{\theta} - \widetilde{\theta})^T.$$

By definition of $\widetilde{\theta}$, $\widetilde{Q}'(\widetilde{\theta}) = 0$, So,

$$\widetilde{Q}(\widehat{\theta}) - \widetilde{Q}(\widetilde{\theta}) = \frac{1}{2}(\widehat{\theta} - \widetilde{\theta})D^{-1}[D\widetilde{Q}''(\theta^*)D]D^{-1}(\widehat{\theta} - \widetilde{\theta})^T \tag{17}$$

where, θ^* is a point on line joining $\widetilde{\theta}$ and $\widehat{\theta}$. We note that $\hat{\theta} \rightarrow \theta^0 \; a.s.$ and $\widetilde{\theta} \rightarrow \theta^0 \; a.s.$ Then $\theta^* \rightarrow \theta^0 \; a.s.$ So, using Lemma 9 $\lim_{N\rightarrow\infty} D\widetilde{Q}''(\theta^*)D$ converges in probability to a positive definite matrix and that implies its minimum eigen value, say λ is strictly positive. By using step-1, the left hand side of (17) is $o_P(1)$. Then

$$\frac{1}{2}(\widehat{\theta} - \widetilde{\theta})D^{-1}D^{-1}(\widehat{\theta} - \widetilde{\theta})^T < \frac{o_P(1)}{\lambda}$$

which implies $(\widehat{\theta} - \widetilde{\theta})D^{-1} \xrightarrow{P} 0$. ∎

Lemma 10. $[\widetilde{Q}'(\theta^0)D] \xrightarrow{d} N(0, \Sigma_1)$.

Proof of Lemma 10

$$\widetilde{Q}'(\theta^0)D = \begin{bmatrix} -\frac{1}{\sqrt{N}}\sum_{n=1}^{N}\cos(\alpha^0 n+\beta^0 n^2)\rho_N'(X(n)) \\ -\frac{1}{\sqrt{N}}\sum_{n=1}^{N}\sin(\alpha^0 n+\beta^0 n^2)\rho_N'(X(n)) \\ \frac{1}{N\sqrt{N}}\sum_{n=1}^{N} n[A^0\sin(\alpha^0 n+\beta^0 n^2)-B^0\cos(\alpha^0 n+\beta^0 n^2)]\rho_N'(X(n)) \\ \frac{1}{N^2\sqrt{N}}\sum_{n=1}^{N} n^2[A^0\sin(\alpha^0 n+\beta^0 n^2)-B^0\cos(\alpha^0 n+\beta^0 n^2)]\rho_N'(X(n)) \end{bmatrix}$$

For investigation about $[\widetilde{Q}'(\theta^0)D]$ let us concentrate on $\rho_N'(X(n))$. We note that $E\rho_N'(X(n)) = 0$ as $\rho_N'(x)$ is an odd function and $X(n)$ has symmetric density f around zero. This gives $E[\widetilde{Q}'(\theta^0)D] = 0$. Now to calculate $V\rho_N'(X(n))$ let us consider the function $[\rho_N'(x)]^2$.

$$\begin{aligned}
[\rho_N'(x)]^2 &= \left[\gamma_N^4x^4 - 4\gamma_N^3x^3 + 4\gamma_N^2x^2\right] I_{\left(0<x\le\frac{1}{\gamma_N}\right)} + 1I_{\left(x>\frac{1}{\gamma_N}\right)} \\
&+ \left[\gamma_N^4x^4 + 4\gamma_N^3x^3 + 4\gamma_N^2x^2\right] I_{\left(-\frac{1}{\gamma_N}\le x\le 0\right)} + 1I_{\left(x<-\frac{1}{\gamma_N}\right)} \\
&= 1 + [\left[\gamma_N^4x^4 - 4\gamma_N^3x^3 + 4\gamma_N^2x^2 - 1\right] I_{\left(0<x\le\frac{1}{\gamma_N}\right)} \\
&\qquad + \left[\gamma_N^4x^4 + 4\gamma_N^3x^3 + 4\gamma_N^2x^2 - 1\right] I_{\left(-\frac{1}{\gamma_N}\le x\le 0\right)}],
\end{aligned}$$

then

$$V[\rho_N'(X(n))] = E[\rho_N'(X(n))]^2 = 1 + R_N^0,$$

where for some constant C^*, $|R_N^0| \le \frac{C^*}{\gamma_N} \to 0$ as $N \to \infty$.

Using above calculated variance, the elements of $[\widetilde{Q}'(\theta^0)D]$ satisfies the conditions of the Central Limit Theorem by Fuller(1996).

To find the asymptotic variance of $\left[\frac{1}{\sqrt{N}}\sum_{n=1}^{N}\cos(\alpha^0 n+\beta^0 n^2)\rho_N'(X(n))\right]$ we need to calculate for $h = 0, \pm 1, \pm 2, \cdots$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N-|h|}\cos(\alpha^0 n+\beta^0 n^2)\cos(\alpha^0(n+h)+\beta^0(n+h)^2).$$

Using Lemma 1, and after some calculations it can be shown that for $h = 0$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N-|h|}\cos(\alpha^0 n+\beta^0 n^2)\cos(\alpha^0(n+h)+\beta^0(n+h)^2) = \frac{1}{2}$$

and it is 0 otherwise. Therefore, using the Central Limit Theorem of linear processes, see Fuller (1996, page 321), the variance turns out to be $\frac{1}{2}$. To find the variance of $\left[\frac{1}{\sqrt{N}}\sum_{n=1}^{N}\sin(\alpha^0 n+\beta^0 n^2)\rho_N'(X(n))\right]$ we need to calculate the above limits where both the cos terms are replaced by sin terms, and we will get similar result using Lemma 1. Now for all h, and for $t=0,1,\cdots$, we also get

$$\lim_{N\to\infty}\frac{1}{N^{t+1}}\sum_{n=1}^{N-|h|} n^t\cos(\alpha^0 n+\beta^0 n^2)\sin(\alpha^0(n+h)+\beta^0(n+h)^2)=0$$

and the variance-covariances of the other terms can be obtained along the same line, using these limits. Finally we get $[\widetilde{Q}'(\theta^0)D]\xrightarrow{d} N(0,\Sigma_1)$. ∎

Proof of Lemma 5: Using multivariate Taylor series expansion we have

$$\widetilde{Q}'(\widetilde{\theta})-\widetilde{Q}'(\theta^0)=(\widetilde{\theta}-\theta^0)\widetilde{Q}''(\bar{\theta}) \tag{18}$$

where $\bar{\theta}$ is a point on line joining $\widetilde{\theta}$ and θ^0. Since, $\widetilde{Q}'(\widetilde{\theta})=0$, (18) can be written as

$$-\widetilde{Q}'(\theta^0)D=(\widetilde{\theta}-\theta^0)D^{-1}[D\widetilde{Q}''(\bar{\theta})D] \tag{19}$$

Note that

$$\left[D\widetilde{Q}''(\bar{\theta})D\right]=D\left[\widetilde{Q}''(\bar{\theta})-\widetilde{Q}''(\theta^0)\right]D+\left[D\widetilde{Q}''(\theta^0)D\right]$$

Using Lemma 9 and Lemma 8 we get $\left[D\widetilde{Q}''(\bar{\theta})D\right]\to\left[\lim_{N\to\infty}D\widetilde{Q}''(\theta^0)D\right]=2f(0)\Sigma_1$ in probability as $N\to\infty$ *i.e.*, $\bar{\theta}\to\theta^0$ *a.s.* (19) gives,

$$(\widetilde{\theta}-\theta^0)D^{-1}=[-\widetilde{Q}'(\theta^0)D][D\widetilde{Q}''(\bar{\theta})D]^{-1} \tag{20}$$

Using Lemma 10 we obtain $[-\widetilde{Q}'(\theta^0)D]\xrightarrow{d}N(0,\Sigma_1)$. So, from (20), combining above two observations we will get the asymptotic distribution of $(\widetilde{\theta}-\theta^0)D^{-1}$. Dividing by $\sqrt{N}$ the expression becomes

$$(\widetilde{\theta}-\theta^0)(\sqrt{N}D)^{-1}=[-\frac{1}{\sqrt{N}}\widetilde{Q}'(\theta^0)D][D\widetilde{Q}''(\bar{\theta})D]^{-1}. \tag{21}$$

Then

$$(\widetilde{\theta}-\theta^0)(\sqrt{N}D)^{-1}\to 0 \text{ in probability,} \tag{22}$$

Theorem 3. *If the Assumptions 1-2 are satisfied, then* $(\widehat{\theta}-\theta^0)(D\sqrt{N})^{-1} \rightarrow 0$ *in probability.*

Proof:

Note that $\Sigma_1^{-1} = 4\Sigma$. As $\lim_{N\rightarrow\infty} D\widetilde{Q}''(\theta^0)D = 2f(0)\Sigma_1$ and $[-\widetilde{Q}'(\theta^0)D] \xrightarrow{d} N(0,\Sigma_1)$ then by (20), $(\widetilde{\theta}-\theta^0)D^{-1} \xrightarrow{d} N_4(0, \frac{1}{4f(0)^2}\Sigma_1^{-1}) = N_4(0, \frac{1}{f(0)^2}\Sigma)$. We note that combining Lemma 5 and Lemma 4 we get $(\widehat{\theta}-\theta^0)D^{-1} \xrightarrow{d} N_4(0, \dfrac{1}{f(0)^2}\Sigma)$. Using Lemma 4 we get $(\widetilde{\theta}-\widehat{\theta})(\sqrt{N}D)^{-1} \rightarrow 0$ in probability. This along with (22) gives $(\widehat{\theta}-\theta^0)(\sqrt{N}D)^{-1} \rightarrow 0$ in probability. ∎

4 Numerical Results and Data Analysis

4.1 Numerical Results:

In this section we perform some simulation experiments to see how the LAD estimators behave for different sample sizes. We consider the following model parameters: $A_1 = 2.0, B_1 = 2.0, \alpha_1 = 1.75, \beta_1 = 1.05$. $X(n)$'s are assumed to be *i.i.d.* Gaussian random variables with mean 0 and variance σ^2. We have taken different sample sizes namely n = 25, 50, 75, 100 and $\sigma^2 = 2$ for our simulation experiments. We compute the average estimates (MEAN), mean squared errors (MSE), variance (VAR) over 1000 replications, and we also provide the asymptotic variances (ASYV) for comparison purposes. We further calculate the asymptotic confidence length (ACON) and coverage probability (CP). To calculate LAD estimators we use the methodology given in Section 3.2. Numerically the minimum value has been obtained by using the *Downhill Simplex Algorithm*, see for example Press *et al.* (1996).

Some of the points are quite clear from the simulation experiments. It is observed that as sample size increases the MSEs, variances and the biases decrease. It verifies the consistency properties of the LAD estimators. The asymptotic variances of the LAD estimators and the MSE's of the different estimators obtained over 1000

Table 1: The results for LADs are reported, when $n = 25$, $\sigma^2 = 2$

PARA	2.00	2.00	1.75	1.05
MEAN	1.883581	1.707926	1.742816	1.042801
MSE	(1.104332)	(1.436890)	(0.037685)	(0.024062)
VAR	(1.090778)	(1.351586)	(0.037633)	(0.024011)
ASYV	(0.628319)	(0.628319)	(0.004825)	(0.724×10^{-5})
ACON	(2.978543)	(3.047005)	(0.270745)	(0.010486)
CP	(0.857000)	(0.870000)	(0.914000)	(0.894000)

Table 2: The results for LADs are reported, when $n = 50$, $\sigma^2 = 2$

PARA	2.00	2.00	1.75	1.05
MEAN	1.964855	1.991527	1.749681	1.050014
MSE	(0.336184)	(0.290056)	(0.000179)	(0.5568×10^{-7})
VAR	(0.334950)	(0.289984)	(0.000179)	(0.5545×10^{-7})
ASYV	(0.314159)	(0.314159)	(0.000603)	(0.2262×10^{-6})
ACON	(2.173244)	(2.149382)	(0.095844)	(0.001856)
CP	(0.911000)	(0.929000)	(0.996000)	(0.995000)

replications are quite close to each other particularly for large sample sizes. So the performances of the LAD estimators are quite satisfactory.

5 Conclusion

In this paper we consider the least absolute deviation estimators for parameters of one dimensional chirp signal. It is observed that the LAD estimators are strongly consistent. Also we found the joint asymptotic normal distribution of the estimators. It is observed that LAD estimators are more efficient than LSE's in presence of additive heavy tailed errors.

Acknowledgement

Part of this work of the first author has been financially supported by the Center for Scientific and Industrial Research (CSIR), and part of this work of the second and

Table 3: The results for LADs are reported, when $n = 75$, $\sigma^2 = 2$

PARA	2.00	2.00	1.75	1.05
MEAN	2.034512	1.999528	1.749682	1.050002
MSE	(0.113840)	(0.120251)	(0.2817×10^{-4})	(0.8081×10^{-8})
VAR	(0.112649)	(0.120251)	(0.2807×10^{-4})	(0.8074×10^{-8})
ASYV	(0.209440)	(0.209440)	(0.1787×10^{-3})	(0.2979×10^{-7})
ACON	(1.776017)	(1.795303)	(0.052159)	(0.000673)
CP	(0.981000)	(0.985000)	(0.995000)	(0.996000)

Table 4: The results for LADs are reported, when $n = 100$, $\sigma^2 = 2$

PARA	2.00	2.00	1.75	1.05
MEAN	2.011717	2.015523	1.750156	1.049995
MSE	(0.078239)	(0.064690)	(0.6823×10^{-5})	(0.1612×10^{-8})
VAR	(0.078102)	(0.064449)	(0.6799×10^{-5})	(0.1616×10^{-8})
ASYV	(0.157080)	(0.157080)	(0.7540×10^{-4})	(0.7069×10^{-8})
ACON	(1.551703)	(1.547225)	(0.033978)	(0.000329)
CP	(0.989000)	(0.995000)	(0.100000)	(0.992000)

third authors have been supported by a grant from the Department of Science and Technology, Government of India. The authors would like to thank the referees for their constructive comments, which had helped to improve the earlier version of the manuscript.

References

[1] Abatzoglou, T. (1986), "Fast maximum likelihood joint estimation of frequency and frequency rate", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, 708 - 714.

[2] Djuric, P.M. and Kay, S.M. (1990), "Parameter estimation of chirp signals", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, 2118 - 2126.

[3] Fuller, W. A. (1996), *Introduction to statistical Time Series, second Edition*, John Wiley and Sons, New York.

[4] Gini, F., Montanari, M. and Verrazzani, L. (2000), "Estimation of chirp signals in compound Gaussian clutter: A cyclostationary approach", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 48, 1029 - 1039.

[5] Jennrich, R.I. (1969), "Asymptotic properties of non-linear least square estimators", *The Annals of Mathematical Statistics*, vol-40, 633-643

[6] Kim, T.S., Kim, H.K. and Choi, H.C. (2000), "Asymptotic properties of LAD estimators of a nonlinear time series regression model", *Journal of Korean Statistical Society*, vol. 29, 2, 187-199.

[7] Kumaresan, R. and Verma, S. (1987), "On estimation the parameters of chirp signals using rank reduction techniques",*Proceedings of 21 st Asilomar Conference*, 555 - 558, Pacific Grove, California.

[8] Kundu, D. and Nandi, S. (2008), "Parameter estimation of chirp signals in presence of stationary noise", *Statistica Sinica*, vol. 18, 187 - 202.

[9] Nandi, S. and Kundu, D. (2004), "Asymptotic properties of the least squares estimators of the parameters of the chirp signals", *Annals of the Institute of Statistical Mathematics*, vol. 56, 529 - 544.

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

[11] Saha, S. and Kay, S. (2002), "Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo importance sampling", *IEEE Transactions on Signal Processing*, vol. 50, 224 - 230.

[12] Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. (1996) *Numerical Recipes in Fortran 90, Second Edition*, Cambridge University Press. Page 402.

[13] Vinogradov, I.M. (1954), "The method of trigonometrical sums in the theory of numbers" , *Interscience*, Translated from Russian. Revised and annotated by K.

F. Roth and Anne Davenport. Reprint of the 1954 translation. Dover Publications, Inc., Mineola, NY, 2004.

[14] White, H. (1980), "Nonlinear regression on cross-section data", *Econometrica*, vol. 48, 721-746.

[15] Weyl, H. (1916), "Ueber die Gleichverteilung von Zahlen mod. Eins" , *Math. Ann.*, vol. 77, 313 - 352.

[16] Wu, C. F. (1981), "Asymptotic theory of non-linear least-squares estimation" , *Annals of Statistics*, vol. 9, No-3, 501 - 513.