Estimating the Parameters of Multiple Chirp Signals

Ananya Lahiri[1,3] & Debasis Kundu[2] & Amit Mitra[2]
[1] Chennai Mathematical Institute, Siruseri, Chennai, India
[2] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, India
[3] Corresponding author: Ananya Lahiri, e-mail: alahiri@cmi.ac.in, ananya.isi@gmail.com

Abstract

In this paper firstly using a number theoretic result we have simplified the structure of asymptotic dispersion matrix of the least squares estimators (LSEs) of multiple chirp signal model. Secondly, we have proposed a type of sequential estimators, which are strong consistent, for the same model, and provide an alternative way to reduce the high dimensionality problem in optimization for LSEs, in a sequential manner.

Key Words and Phrases: asymptotic distribution; least squares estimators; linear processes; sequential method; strong consistency.

1 Introduction

In this paper we consider the following multiple chirp model;

$$y(n) = \sum_{k=1}^{p} \left\{ A_k^0 \cos(\alpha_k^0 n + \beta_k^0 n^2) + B_k^0 \sin(\alpha_k^0 n + \beta_k^0 n^2) \right\} + X(n);\ n = 1, \cdots, N. \quad (1)$$

Here $y(n)$ is the real valued signal observed at $n = 1, \cdots, N$. For $k = 1, \cdots, p$, A_k^0, B_k^0 are amplitudes, and α_k^0 and β_k^0 are frequency and frequency rate respectively. The additive error $\{X(n)\}$ is a sequence of stationary random variables (r.v.s) with mean zero and finite fourth moment. Detail assumptions will be provided later.

The chirp signal model appears in the signal processing literature, and occurs quite naturally, particularly in physics, sonar, radar and communications. Extensive work on this model, mainly when $p = 1$, has been done by several authors. One is referred to Kundu and Nandi (2008) and the references cited therein for detailed literature. Saha and Kay (2002) introduced the multiple chirp signal model (1), and provided the maximum likelihood estimators of the unknown parameters using importance sampling procedure under the assumptions that $X(n)$'s are independent and identically distributed (*i.i.d.*) normal r.v.s. Kundu and Nandi (2008) proved the strong consistency and asymptotic normality of the LSEs of the model (1)when the

$X(n)$'s are obtained from a linear stationary process. Firstly the structure of the dispersion matrix of the asymptotic distribution of the LSEs is quite complicated. It was observed that the LSE of α_k^0 has the convergence rate $O_p(N^{-3/2})$, whereas the LSE of β_k^0 has the convergence rate $O_p(N^{-5/2})$. But finding the LSEs is a numerically challenging problem, if $p \geq 2$. For the model (1), it involves solving a $2p$-dimensional optimization problem. Therefore our second observation is, for large p, it becomes a highly computer intensive method.

First aim of this paper is to provide a simplified structure of the dispersion matrix of the asymptotic distribution of the LSEs using a number theoretic result of Vinogradov (1954). The second aim is to provide an estimation procedure which is computationally less demanding and produces efficient estimators of the unknown parameters. If p is known, using the idea that the regressors vectors are orthogonal, we provide a step by step sequential estimation procedure for estimating the amplitudes, frequency and frequency rate. It is observed that $2p$-dimensional optimization procedure can be reduced to p sequential 2-dimensional (2-D) optimization problems. Therefore for large p, the proposed sequential method is very effective. If p is not known, and we fit a lower order model, *i.e.* when the assumed number of components is less than the actual number of components then the proposed estimators converge almost surely to true parameter values. If we fit a higher order model, *i.e.* assumed number of components is more than the actual number of components, then the amplitude estimates obtained after p-th step converge to zero almost surely. Due to the complicated nature of the model, we could not establish the asymptotic distribution of the proposed sequential estimators. Based on an unsolved conjecture in number theory it can be shown that the asymptotic distribution of the LSEs and the proposed sequential estimators are same. We perform some simulation experiments and observe that the mean squared errors (MSEs) of the LSEs and sequential estimators are very close to each other. Due to space restriction we are not providing proofs and the simulation results. But we provide the analysis of two real data sets for illustrative purpose. In subsequent sections we provide our findings.

Model Assumptions and Preliminary Results :

Assumption 1: The r.v.s $X(n)$ satisfies the condition $X(n) = \sum_{j=-\infty}^{\infty} a(j)e(n-j)$, where $\{e(n)\}$ is a sequence of *i.i.d.* r.v.s with mean zero,variance σ^2 and finite fourth moment, and $\sum_{j=-\infty}^{\infty} |a(j)| < \infty$.

We use the following notations; The parameter vector as $\theta_k = (A_k, B_k, \alpha_k, \beta_k)$, the true parameter vector as $\theta_k^0 = (A_k^0, B_k^0, \alpha_k^0, \beta_k^0)$, for $k = 1, \cdots, p$, and the parameter space as $\Theta = [-K, K] \times [-K, K] \times [0, \pi] \times [0, \pi], \quad K > 0$.

Assumption 2: It is assumed that θ_k^0 is an interior point of Θ, for $k = 1, \cdots, p$, and α_k^0's are distinct, and similarly β_k^0's are also distinct.

Assumption 3: A_k^0's and B_k^0's satisfy $K^2 > A_1^{0^2} + B_1^{0^2} > \cdots > A_p^{0^2} + B_p^{0^2} > 0$.

We require the following results based on Vinogradov (1954) for further development.

Lemma 1: If (ω_1, ω_2) in $(0, \pi) \times (0, \pi)$, then except for countable number of points the followings are true.

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\cos(\omega_1 n+\omega_2 n^2)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\sin(\omega_1 n+\omega_2 n^2)=0. \tag{2}$$

$$\text{For } t=0,\ 1,\ 2,\quad \lim_{N\to\infty}\frac{1}{N^{t+1}}\sum_{n=1}^{N}n^t\sin(\omega_1 n+\omega_2 n^2)\cos(\omega_1 n+\omega_2 n^2)=0, \tag{3}$$

$$\lim_{N\to\infty}\frac{1}{N^{t+1}}\sum_{n=1}^{N}n^t\cos^2(\omega_1 n+\omega_2 n^2)=\lim_{N\to\infty}\frac{1}{N^{t+1}}\sum_{n=1}^{N}n^t\sin^2(\omega_1 n+\omega_2 n^2)=\frac{1}{2(t+1)}. \tag{4}$$

Lemma 2: If $X(n)$ satisfies Assumption 1 then as $N\to\infty$ and for $s\geq 0$, $i=\sqrt{-1}$,

$$\sup_{\alpha,\beta}\left|\frac{1}{N^{s+1}}\sum_{n=1}^{N}n^s X(n)e^{i(\alpha n+\beta n^2)}\right|\to 0\quad a.s. \tag{5}$$

2 The Two Main Results

Asymptotic Distribution of the LSEs: Readers are referred to Kundu and Nandi (2008) or *www.isid.ac.in/ statmath/eprints/(isid/ms/*2005/08) for the result of the asymptotic distribution of the LSEs where dispersion matrix is quite complicated. Using Lemma 1 and Lemma 2, the following simplified version of the asymptotic distribution of the LSEs can be obtained.

Theorem 1: If the Assumptions 1-3 are satisfied, then the LSEs of the unknown parameters have the following asymptotic $4p$ variate normal distribution

$$((\tilde{\theta_1}-\theta_1^0)D^{-1},\cdots,(\tilde{\theta_p}-\theta_p^0)D^{-1})\xrightarrow{d}\mathcal{N}_{4p}(0,2c\sigma^2\mathbf{\Sigma}(\theta_0)) \tag{6}$$

where $\mathbf{\Sigma}(\theta_0)$ is a $4p\times 4p$ matrix having the following block-diagonal structure as

$$\mathbf{\Sigma}(\theta^0)=\begin{bmatrix}\mathbf{\Sigma}_1 & 0 & \cdots & 0\\ 0 & \mathbf{\Sigma}_2 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & \mathbf{\Sigma}_p\end{bmatrix},\quad c=\sum_{j=-\infty}^{\infty}a(j)^2,$$

$D=diag\left(\frac{1}{\sqrt{N}},\frac{1}{\sqrt{N}},\frac{1}{\sqrt{N^3}},\frac{1}{\sqrt{N^5}}\right)$ is 4×4 diagonal matrix and for $k=1,\cdots,p$,

$$\mathbf{\Sigma}_k=\frac{1}{{A_k^0}^2+{B_k^0}^2}\begin{bmatrix}\frac{1}{2}\left({A_k^0}^2+9{B_k^0}^2\right) & -4A_k^0B_k^0 & -18B_k^0 & 15B_k^0\\ -4A_k^0B_k^0 & \frac{1}{2}\left(9{A_k^0}^2+{B_k^0}^2\right) & 18A_k^0 & -15A_k^0\\ -18B_k^0 & 18A_k^0 & 96 & -90\\ 15B_k^0 & -15A_k^0 & -90 & 90\end{bmatrix}.$$

Sequential Estimation Procedure: Here we propose a sequential procedure to estimate the unknown parameters of the model (1), and prove that they are strongly consistent. Let us use the following notations. The $N\times 2$ matrix $\mathbf{W}(\alpha,\beta)$ is defined as $\mathbf{W}(\alpha,\beta) = \begin{bmatrix} \cos(\alpha+\beta) & \sin(\alpha+\beta) \\ \cos(2\alpha+4\beta) & \sin(2\alpha+4\beta) \\ \vdots & \vdots \\ \cos(N\alpha+N^2\beta) & \sin(N\alpha+N^2\beta) \end{bmatrix}$. At first step of sequential method we minimize $Q_1(A,B,\alpha,\beta)$ with respect to A, B, α, β, where

$$Q_1(A,B,\alpha,\beta) = \left[\mathbf{Y} - \mathbf{W}(\alpha,\beta)\begin{pmatrix} A \\ B \end{pmatrix}\right]^T \left[\mathbf{Y} - \mathbf{W}(\alpha,\beta)\begin{pmatrix} A \\ B \end{pmatrix}\right], \tag{7}$$

and $\mathbf{Y} = [y(1), \cdots, y(N)]^T$ is the data vector. For fixed α and β

$$(\widehat{A}(\alpha,\beta), \widehat{B}(\alpha,\beta)) = [\mathbf{W}(\alpha,\beta)^T\mathbf{W}(\alpha,\beta)]^{-1}\mathbf{W}(\alpha,\beta)^T\mathbf{Y}. \tag{8}$$

minimizes (7). Therefore, using the separable regression technique of Richards (1961), the minimization of $Q_1(A,B,\alpha,\beta)$ can be obtained by minimizing

$$R_1(\alpha,\beta) = Q_1(\widehat{A}(\alpha,\beta), \widehat{B}(\alpha,\beta), \alpha, \beta) = \mathbf{Y}^T(\mathbf{I} - \mathbf{P_1}(\alpha,\beta))\mathbf{Y} \tag{9}$$

with respect to (α,β), where $\mathbf{P_1}(\alpha,\beta) = \mathbf{W}(\alpha,\beta)\left[\mathbf{W}(\alpha,\beta)^T\mathbf{W}(\alpha,\beta)\right]^{-1}\mathbf{W}(\alpha,\beta)^T$ is the projection matrix on the column space of the matrix $\mathbf{W}(\alpha,\beta)$. If $(\widehat{\alpha}_1, \widehat{\beta}_1)$ are the minimizers of $R_1(\alpha,\beta)$, then the estimators of $A_1^0, B_1^0, \alpha_1^0,$ and β_1^0 become $\widehat{A}_1 = \widehat{A}(\widehat{\alpha}_1, \widehat{\beta}_1), \widehat{B}_1 = \widehat{B}(\widehat{\alpha}_1, \widehat{\beta}_1), \widehat{\alpha}_1, \widehat{\beta}_1$ respectively.

Now to compute the estimators of $(A_2^0, B_2^0, \alpha_2^0, \beta_2^0)$, we take out the effect of the first component from the signal, *i.e.*, we consider a new data vector

$$\mathbf{Y}^1 = \mathbf{Y} - \mathbf{W}(\hat{\alpha_1}, \hat{\beta_1})\begin{pmatrix} \widehat{A}_1 \\ \\ \widehat{B}_1 \end{pmatrix} \tag{10}$$

Then using the new data vector $\mathbf{Y}^1$, following the same procedure as before we get $\widehat{A}_2, \widehat{B}_2, \widehat{\alpha}_2, \widehat{\beta}_2$, the estimators of $A_2^0, B_2^0, \alpha_2^0, \beta_2^0$ respectively. Continuing in this manner, at the k-th stage we can obtain estimators of $A_k^0, B_k^0, \alpha_k^0, \beta_k^0$, say $\widehat{A}_k, \widehat{B}_k, \widehat{\alpha}_k, \widehat{\beta}_k$ respectively.

For the consistency results of the proposed estimators we consider two cases separately; (i) for lower order model and (ii) for higher order model. At first step we have the following result.

Theorem 2: If the Assumptions 1-3 are satisfied then $(\widehat{A}_1, \widehat{B}_1, \widehat{\alpha}_1, \widehat{\beta}_1)$ is a strongly consistent estimator of $(A_1^0, B_1^0, \alpha_1^0, \beta_1^0)$.

The estimators obtained at the second step also are strongly consistent and for subsequent steps $3 \leq k \leq p$ along same manner we get consistency. We have the following results.

Theorem 3: If the Assumptions 1-3 are satisfied and $p \geq 2$, then $\widehat{\theta}_2$, the estimator obtained by minimizing $Q_2(A,B,\alpha,\beta)$, where $Q_2(A,B,\alpha,\beta)$ is obtained by replacing $\mathbf{Y}$ with $\mathbf{Y}^1$ in (7), is a strongly consistent estimator of θ_2^0.

Theorem 4: If the Assumptions 1-3 are satisfied and $p \geq k$, then the estimators obtained at the k-th step are strongly consistent.

If the sequential process is continued even after p-th step, then we get:

Theorem 5: If the Assumptions 1-3 are satisfied, and if $\widehat{A}_k, \widehat{B}_k, \widehat{\alpha}_k, \widehat{\beta}_k$ are the estimators obtained at the k-step for $k > p$, then $\widehat{A}_k \to 0$ $a.s.$ and $\widehat{B}_k \to 0$ $a.s.$.

Comments: If the following conjecture, see Montgomery (1990), holds then it can be shown that the proposed sequential estimators have the same asymptotic distribution as that of the ordinary LSEs.

Conjecture: If $\omega_1, \omega_2, \omega_1', \omega_2' \in (0, \pi)$, then except for countable number of points

$$\lim_{N \to \infty} \frac{1}{\sqrt{N} N^t} \sum_{n=1}^{N} n^t \cos(\omega_1 n + \omega_2 n^2) \sin(\omega_1' n + \omega_2' n^2) = 0; \quad t = 0, 1, 2. \tag{11}$$

In addition if $\omega_2 \neq \omega_2'$, then

$$\lim_{N \to \infty} \frac{1}{\sqrt{N} N^t} \sum_{n=1}^{N} n^t \cos(\omega_1 n + \omega_2 n^2) \cos(\omega_1' n + \omega_2' n^2) = 0; \quad t = 0, 1, 2. \tag{12}$$

$$\lim_{N \to \infty} \frac{1}{\sqrt{N} N^t} \sum_{n=1}^{N} n^t \sin(\omega_1 n + \omega_2 n^2) \sin(\omega_1' n + \omega_2' n^2) = 0; \quad t = 0, 1, 2. \tag{13}$$

3 Real Life Data Analysis and Conclusions

Real Life Data Analysis: In this section we perform the analysis of two speech signal data sets; "AHH" and "AWW" vowel sound mainly for illustrative purpose. Both these data sets are obtained from a sound instrument at the Speech Signal Processing laboratory of the Indian Institute of Technology Kanpur. There are 469 data points of "AHH" signal and 512 data points of "AWW" signal and they are sampled at 10 kHz frequency. We have fitted the chirp signal model to these data sets and use the proposed sequential method to compute the unknown parameters. Since the number components of model is not known in this case, we use the Bayesian Information Criterion (BIC) to estimate the number of components. The BIC takes the following form

$$BIC(k) = N \ln(SSE) + \frac{1}{2}(4k + ar_k + 1)\ln(N)$$

in this case, where k is the number of components fitted in the model, ar_k is the number of parameters fitted in the stationary process and N is the data size, SSE is error sum of squares. We choose that model order for which the BIC is minimum. For "AHH" data set, the estimate of p becomes 5, and for "AWW" it is 7. We perform Dickey-Fuller test for checking the stationarity for the residuals, and in both cases the null hypothesis cannot be rejected. We also provide the plots for the fitted and predicted signals and they match quite well in both the cases.

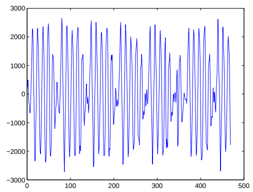

Figure 1: AHH :Original signal

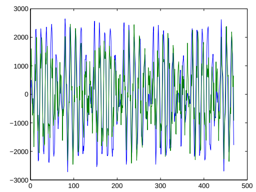

Figure 3: AHH: Fitted (Green) vs Original signal

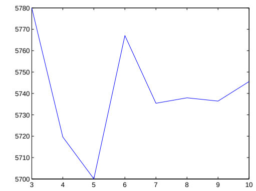

Figure 5: AHH: BIC plot

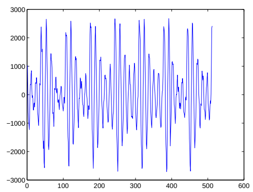

Figure 2: AWW :Original signal

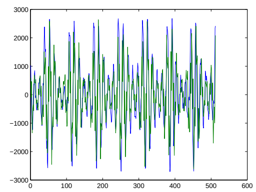

Figure 4: AWW: Fitted (Green) vs Original signal

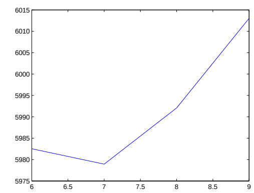

Figure 6: AWW: BIC plot

Conclusions: In this paper using a number theoretic result of Vinogradov (1954), we provide a simplified form of the asymptotic dispersion matrix of the LSEs of multiple chirp signal model and observe that the LSEs of the different chirp components are asymptotically independent. We also provide a sequential estimation procedure of the parameters and prove their strong consistency. The proposed sequential method can be very useful in fitting multiple chirp signal model, particularly when the number of chirp components is large. We analyze two real data sets.

Acknowledgement: For this work first author is financially supported by CSIR, India; second and third authors are supported by DST, India.

References

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

[2] Montgomery H.L. (1990), *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis*, American Mathematical Society, 196.

[3] Richards, F. S. G. (1961) "A method of maximum likelihood estimation", *Jour. Royal Stat. Soc.* B, 23, 469-475.

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

[5] 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.