Journal of Statistical Planning and Inference 110 (2003) 23–34

journal of statistical planning and inference

www.elsevier.com/locate/jspi

An efficient algorithm for estimating the parameters of superimposed exponential signals

Z.D. Bai[a], C.R. Rao[b], M. Chow[b], D. Kundu[c,*]

[a] Department of Statistics and Applied Probability, National University of Singapore, 3 Science Drive 2, Singapore 117543, Singapore

[b] Department of Statistics, The Pennsylvania State University, University Park, PA 16802, USA

[c] Department of Mathematics, Indian Institute of Technology, P.O. Box I.I.T. Kanpur, Pin. 208016, India

Received 23 August 2000; received in revised form 19 April 2001; accepted 22 August 2001

Abstract

An efficient computational algorithm is proposed for estimating the parameters of undamped exponential signals, when the parameters are complex valued. Such data arise in several areas of applications including telecommunications, radio location of objects, seismic signal processing and computer assisted medical diagnostics. It is observed that the proposed estimators are consistent and the dispersion matrix of these estimators is asymptotically the same as that of the least squares estimators. Moreover, the asymptotic variances of the proposed estimators attain the Cramer–Rao lower bounds, when the errors are Gaussian.

MSC: 62F10; 60K40

Keywords: Cramer–Rao lower bound; Equivariance linear prediction; Forward and backward linear prediction; Maximum likelihood estimators; Superimposed exponential signals

1. Introduction

We consider the following model:

$$Y(t) = \sum_{k=1}^{M} \alpha_k^0 e^{i\omega_k^0 t} + e(t). \tag{1.1}$$

Here, α_k^0's are unknown complex amplitudes and none of them is identically equal to zero, $i = \sqrt{-1}$. The ω_k^0's are unknown frequencies lying strictly between 0 and 2π and they are distinct. The error random variables $e(t)$'s are independent and identically distributed (i.i.d.) complex valued random variables, with mean 0 and equal variance $\sigma^2/2$ for both the real and imaginary parts. The real and imaginary parts of $e(t)$ are

* Corresponding author. Tel.: +91-512-597-636; fax: +91-512-590-007.

E-mail address: kundu@iitk.ac.in (D. Kundu).

PII: S0378-3758(01)00287-7

assumed to be independent. 'M', the number of signals, is assumed to be known in advance. The problem is to estimate α's, ω's and σ^2 given a sample of size n. Note that although estimation of α's and ω's are both important, estimation of the linear parameters α's is much easier compared to the estimation of the non-linear parameters ω's. In this paper we mainly consider an efficient estimation procedure of the frequencies.

This is a very important and well discussed problem in statistical signal processing. Such data arise in several areas of applications, including telecommunications, radio location of objects, seismic signal processing and computer assisted medical diagnostics. It has received considerable attention during the past 20 years, see for example, the paper of Stoica (1993) for an extensive list of references and also the review articles of Kay and Marple (1981), Rao (1988) and Prasad et al. (1995). This problem is well known to be numerically difficult (Varah, 1985). It is observed (Osborne and Smyth, 1995; Kundu, 1993) that the general purpose algorithms such as Gauss–Newton, Newton–Raphson or their variants take a long time to converge to the least squares estimators even from a good starting value. A considerable amount of research (Kay, 1984; Bresler and Macovski, 1986; Kumaresan et al., 1986) has been done to obtain the least squares estimators (lse's) efficiently. However, all these methods have the drawbacks of high computational complexities and dependence on the initial value chosen.

Among the non-iterative methods, the well known methods are the modified forward backward linear prediction (MFBLP) of Tufts and Kumaresan (1982), equivariance linear prediction (EVLP) of Bai et al. (1986), multiple signal classifications (MUSIC) of Schmidt (1981), eigen space rotation invariance technique (ESPRIT) of Roy (1987) and noise space decomposition (NSD) of Kundu and Mitra (1995). It is well known (Kay, 1988) that MFBLP works better than MUSIC or ESPRIT in accuracy and SNR threshold. However, it is observed (Rao, 1988) that although MFBLP works very well for small samples and at reasonably small SNR, it may not be consistent. Although EVLP provides consistent estimates of the frequencies which are asymptotically normal and have a convergence rate of $O_p(n^{-1/2})$, (here $O_p(.)$ denotes bounded in probability) the convergence rate does not achieve the best possible one. On the other hand, the NSD method provides consistent estimates of the frequencies and its small sample performance is also quite good but the rate of convergence is not easy to obtain theoretically.

In this paper we produce efficient estimates of the frequencies with the best possible convergence rate of $O_p(n^{-3/2})$, within a fixed number of iterations. Therefore, our proposed estimates and the lse's become asymptotically equivalent. It may be mentioned that none of the iterative processes available today guarantees the convergence within a fixed number of iteration.

The rest of the paper is organized as follows. In Section 2, we give a brief description of the Prony's estimates and the asymptotic properties of the lse's. The proposed algorithm is provided in Section 3. Numerical results are presented in Section 4 and the paper concludes in Section 5. All the necessary theoretical results are provided in the appendix.

2. Prony's algorithm and least squares estimators

Prony's (1795) algorithm plays an important role in fitting sum of exponentials to equispaced data. Several articles on numerical analysis are available today regarding Prony's algorithm. A good account of it can be obtained in Barrodale and Olesky (1981).

The basic idea of Prony can be stated as follows: Let $Y(1),\dots,Y(n)$ be n data points satisfying model (1.1) and the data are noiseless, i.e. $\mathrm{Var}(\sigma^2)=0$. Then Prony observed that there exists $(M+1)$ constants, $g_1,\dots,g_{M+1}$ such that

$$\begin{aligned} &g_1 Y(1)+\cdots+g_{M+1}Y(M+1)=0\\ &\vdots\\ &g_1 Y(n-M)+\cdots+g_{M+1}Y(n)=0. \end{aligned} \tag{2.1}$$

Here $\mathbf{g}=(g_1,\dots,g_{M+1})$ is unique up to a constant multiplication and therefore, without loss of generality we can take $|\mathbf{g}|^2=\sum_{i=1}^{M+1} g_i^2=1$. It can be easily seen that $\mathrm{e}^{-\mathrm{i}\omega_1^0},\dots,\mathrm{e}^{-\mathrm{i}\omega_M^0}$ are the roots of the following polynomial equation:

$$B(z)=g_1+g_2 z+\cdots+g_{M+1}z^{M+1}=0. \tag{2.2}$$

Therefore, if the data are noiseless, for each $(\omega_1,\dots,\omega_M)$, there is a unique $\mathbf{g}=(g_1,\dots,g_{M+1})$, such that $|\mathbf{g}|^2=1$ and $g_1>0$. Observe that (2.1) can be written in the matrix form as follows:

$$\mathbf{G}\mathbf{g}=\mathbf{0}, \tag{2.3}$$

where $\mathbf{G}$ is a $(n-M)\times(M+1)$ data matrix. Since $\mathbf{g}$ is unique, it implies that $\mathbf{G}$ is of rank M, if $n-M>M+1$. Note that $\mathbf{g}$ is an eigenvector corresponding to the zero eigenvalue of $\mathbf{G}^{\mathrm{H}}\mathbf{G}$, here 'H' denotes the conjugate transpose of a matrix or of a vector.

This idea was used even when the data are noisy in Bai et al. (1986). Consider the following data matrix:

$$\mathbf{Z}=\frac{1}{n-M}\mathbf{G}^{\mathrm{H}}\mathbf{G}. \tag{2.4}$$

Obtain the eigenvector corresponding to the minimum eigenvalue of $\mathbf{Z}$. It gives an estimator $\tilde{g}=(\tilde{g}_1,\dots,\tilde{g}_{M+1})$ of $\mathbf{g}$. Construct a polynomial equation of the form

$$\tilde{g}_1+\tilde{g}_2 z+\cdots+\tilde{g}_{M+1}z^M=0, \tag{2.5}$$

and obtain the solutions of the form

$$\tilde{\rho}_1\mathrm{e}^{-\mathrm{i}\tilde{\omega}_1},\dots,\tilde{\rho}_M\mathrm{e}^{-\mathrm{i}\tilde{\omega}_M}. \tag{2.6}$$

Now take $\tilde{\omega}_1,\dots,\tilde{\omega}_M$ as estimators of $\omega_1,\dots,\omega_M$. These estimators are known as the EVLP estimators of the frequencies. It is shown in Bai et al. (1986) that $\tilde{\omega}_1,\dots,\tilde{\omega}_M$ are consistent estimators of $\omega_1,\dots,\omega_M$ with the convergence rate $\mathrm{O}_\mathrm{p}(n^{-1/2})$.

Now, we describe briefly the lse's and their asymptotic properties. Note that the most intuitive estimators will be the lse's obtained by minimizing the residual sums of squares

$$R_n(\alpha,\omega)=\sum_{t=1}^{n}\left|Y(t)-\sum_{k=1}^{M}\alpha_k e^{i\omega_k t}\right|^2. \tag{2.7}$$

Note that $R_n(\alpha,\omega)$ can also be written as follows:

$$R_n(\alpha,\omega)=[\mathbf{Y}-\mathbf{A}(\omega)\alpha]^{\mathrm{H}}[\mathbf{Y}-\mathbf{A}(\omega)\alpha], \tag{2.8}$$

where $\mathbf{Y}=(Y(1)=\dots,Y(n))^{\mathrm{T}}$, $\alpha=(\alpha_1,\dots,\alpha_M)^{\mathrm{T}}$, $\omega=(\omega_1,\dots,\omega_M)^{\mathrm{T}}$ and $\mathbf{A}(\omega)$ is a $n\times M$ matrix of the form

$$\mathbf{A}(\omega)=\begin{bmatrix} e^{i\omega_1} & \dots & e^{i\omega_M} \\ \vdots & \vdots & \vdots \\ e^{in\omega_1} & \dots & e^{in\omega_M} \end{bmatrix}. \tag{2.9}$$

From (2.9), it is clear how α can be separated from ω. Now observe that for a fixed ω, the lse of α can be obtained as

$$\hat{\alpha}(\omega)=[\mathbf{A}(\omega)^{\mathrm{H}}\mathbf{A}(\omega)]^{-1}\mathbf{A}(\omega)^{\mathrm{H}}\mathbf{Y}. \tag{2.10}$$

Now, if we replace the estimator of α in (2.8), we obtain

$$Q_n(\omega)=R_n(\hat{\alpha}(\omega),\omega)=\mathbf{Y}^{\mathrm{H}}[\mathbf{I}-\mathbf{P_A}]\mathbf{Y}, \tag{2.11}$$

where

$$\mathbf{P_A}=\mathbf{A}(\omega)[\mathbf{A}(\omega)^{\mathrm{H}}\mathbf{A}(\omega)]^{-1}\mathbf{A}(\omega)^{\mathrm{H}}$$

is the projection matrix on the space spanned by the columns of $\mathbf{A}(\omega)$. Therefore, the lse's of (α,ω) obtained by minimizing $R_n(\alpha,\omega)$ with respect to (α,ω) are the same as obtaining the lse of ω first by minimizing $Q_n(\omega)$ with respect to ω and then use that estimator of ω in (2.10) to obtain the lse of α. Most of the special purpose algorithms namely the methods proposed by Bresler and Macovski (1988), Kumaresan et al. (1986) and Kundu (1993) minimize (2.11), which naturally save computational times.

It is known (Rao and Zhao, 1993; Kundu and Mitra, 1999) that the lse's of the frequencies have the following limiting distribution:

$$n^{3/2}(\hat{\omega}-\omega)\to(\mathbf{0},6\sigma^2(\mathbf{D}^{\mathrm{H}}\mathbf{D})^{-1}), \tag{2.12}$$

where $\mathbf{D}$ is a $M\times M$ diagonal matrix as follows:

$$\mathbf{D}=\mathrm{diag}\{\alpha_1^0,\dots,\alpha_M^0\}.$$

Therefore, it is clear that although the EVLP method provides consistent estimators of the frequencies, the rates of convergence are much slower than the lse's. On the other hand, although the lse's have convergence rates $O_p(n^{-3/2})$, it is not known theoretically how many iterations any iterative procedure may need to obtain the lse's from a given starting value.

3. Proposed algorithm

Let $\tilde{\omega}_j$ (not necessarily EVLP estimator) be a consistent estimator of ω_j, for $j = 1,\dots,M$ and compute $\hat{\omega}_j$ (not necessarily lse) for $j = 1,\dots,M$ as follows:

$$\hat{\omega}_j = \tilde{\omega}_j + \frac{12}{n^2}\mathrm{Im}\left[\frac{C_n}{D_n}\right], \tag{3.1}$$

where

$$C_n = \sum_{t=1}^{n} Y(t)(t - n/2)\mathrm{e}^{-\mathrm{i}\tilde{\omega}_j t}, \qquad D_n = \sum_{t=1}^{n} Y(t)\mathrm{e}^{-\mathrm{i}\tilde{\omega}_j t}$$

and Im(.) means the imaginary part of a complex number. Then we have the following result.

Theorem 1. *If $\tilde{\omega}_j - \omega_j = \mathrm{O_p}(n^{-1-\delta})$ for $\delta \in (0, 1/2]$ for $j = 1,\dots,M$, then*

(1) $\hat{\omega}_j - \omega_j = \mathrm{O_p}(n^{-1-2\delta})$ *if* $\delta \leqslant \frac{1}{4}$,
(2) $n^{3/2}(\hat{\omega} - \omega) \to N_M(\mathbf{0}, 6\sigma^2(\mathbf{D}^{\mathrm{H}}\mathbf{D})^{-1}$ *if* $\delta > \frac{1}{4}$.

Proof. See the appendix.

We start with a consistent estimate of ω_j and improve upon it step by step by a recursive algorithm. The mth stage estimate $\hat{\omega}_j^{(m)}$ is computed from the $(m-1)$th stage estimate $\hat{\omega}_j^{(m-1)}$, by the formula

$$\hat{\omega}_j^{(m)} = \hat{\omega}_j^{(m-1)} + \frac{12}{n_m^2}\mathrm{Im}\left[\frac{C_{n_m}}{D_{n_m}}\right], \tag{3.2}$$

where C_{n_m} and D_{n_m} can be obtained from C_n and D_n by replacing n and $\tilde{\omega}_j$ with n_m and $\hat{\omega}_j^{(m-1)}$, respectively. We apply formula (3.2) repeatedly choosing n_m suitably at each stage as follows.

Step 1: with $m = 1$, choose $n_1 = n^{0.40}$ and $\hat{\omega}_j^{(0)} = \tilde{\omega}_j$, the EVLP estimator. Note that

$$\tilde{\omega}_j - \omega_j = \mathrm{O_p}(n^{-1/2}) = \mathrm{O_p}(n_1^{-1-1/4}). \tag{3.3}$$

Then substituting $n_1 = n^{0.40}$, $\hat{\omega}_j^{(0)} = \tilde{\omega}_j$ in (3.2) and applying Theorem 1, we obtain

$$\hat{\omega}_j^{(1)} - \omega_j = \mathrm{O_p}(n_1^{-1-1/2}) = \mathrm{O_p}(n^{-0.60}). \tag{3.4}$$

Step 2: with $m = 2$, choose $n_2 = n^{.48}$ and compute $\hat{\omega}_j^{(2)}$ from $\hat{\omega}_j^{(1)}$ using (3.2). Since $\hat{\omega}_j^{(1)} - \omega_j = \mathrm{O_p}(n^{-0.60}) = \mathrm{O_p}(n_2^{-1-1/4})$ and using Theorem 1, we obtain

$$\hat{\omega}_j^{(2)} - \omega_j = \mathrm{O_p}(n_2^{-1-1/2}) = \mathrm{O_p}(n^{-0.72}). \tag{3.5}$$

Choosing $n_3,\dots,n_6$ as given below and applying the main theorem in the same way as above, we have

Step 3: $n_3 = n^{0.57}$, yielding $\hat{\omega}_j^{(3)} - \omega_j = \mathrm{O_p}(n^{-0.87})$.
Step 4: $n_4 = n^{0.70}$, yielding $\hat{\omega}_j^{(4)} - \omega_j = \mathrm{O_p}(n^{-1.04})$.
Step 5: $n_5 = n^{0.83}$, yielding $\hat{\omega}_j^{(5)} - \omega_j = \mathrm{O_p}(n^{-1.25})$.
Step 6: $n_6 = n^{0.92}$, yielding $\hat{\omega}_j^{(6)} - \omega_j = \mathrm{O_p}(n^{-1.58})$.

Finally take $n_7 = n$ and compute $\hat{\omega}_j^{(7)}$ from $\hat{\omega}_j^{(6)}$. Now applying Theorem 1, we have

$$n^{3/2}(\hat{\omega}^{(7)} - \omega) \to N_M(\mathbf{0}, 6\sigma^2(\mathbf{D}^{\mathrm{H}}\mathbf{D})^{-1}). \tag{3.6}$$

Once we obtain ω's, the amplitudes α's can be obtained using (2.10).

Comparing (3.6) and (2.12), it is clear that $\hat{\omega}^{(7)}$ has the same asymptotic distribution as the lse's and therefore it is a fully asymptotic efficient estimate of ω.

Comments. Note that the exponents we have used above are not unique. There are several other ways they can be chosen so that the iterative process will converge in seven steps. For example, another set of choices can be $n_1 = n^{0.40}$, $n_2 = n^{0.48}$, $n_3 = n^{0.57}$, $n_4 = n^{0.67}$, $n_5 = n^{0.80}$, $n_6 = n^{0.90}$ and $n_7 = n$. It is not possible to choose a set of exponents to make the iterative process converge in less than seven steps, but definitely there are several sets of exponents for which they may take more than seven steps.

4. Numerical experiments and discussions

In this section, we present some numerical results to compare how the lse's and the proposed estimators behave for different σ^2 when the errors are normally distributed. We also compare the performances of the lse's and the proposed method when the errors are not necessarily normally distributed. All these computations are performed in FORTRAN and using the IMSL random deviate generator. We consider the following model:

$$Y(t) = 2.5\mathrm{e}^{5.5it} + 3.0\mathrm{e}^{3.5it} + e(t). \tag{4.1}$$

Here $e(t)$'s are i.i.d. complex valued random variables with mean zero and variance $\sigma^2/2$ for both the real and imaginary parts. The real and the imaginary parts are taken to be independent and normally distributed. We take $n = 25, 50$ and $\sigma^2 = 1, 2, 3$. We also consider the case when the real and the imaginary parts of $e(t)$'s are independent and double exponential random variables with mean zero and finite variance $\sigma^2/2$ for both the real and imaginary parts. We consider the above medium tailed double exponential error to study the robustness property of the proposed method at least for small sample sizes. In this case also we consider $n = 25, 50$ and $\sigma^2 = 1, 2, 3$.

Ten thousand different data sets were generated for each σ^2. For each data set we estimated $\omega = (5.5, 3.5)$ by the least squares method and also by the proposed method. We report the average of these estimates and also the mean squared errors (mse's) over 10,000 replications. We also report the Cramer–Rao bounds for comparison purposes. All the results are reported in Tables 1–4. Tables 1–2 represent the results when the errors are normally distributed and for the non-normal errors the results are presented in Tables 3–4. In each table the first line represents the true parameter value and the corresponding Cramer–Rao bounds are reported immediately below. The third line represents the average value of the lse's and the fourth line represents the corresponding mse's. Similarly the fifth line and the seventh line represent the average value of the

Table 1
The average estimates and their mse's of the different methods when the errors are normally distributed and the sample size is 25

	$\sigma^2=1$		$\sigma^2=2$		$\sigma^2=3$	
Parameters	5.50000	3.50000	5.50000	3.50000	5.50000	3.50000
CRLB	6.14E−5	4.27E−5	1.23E−4	8.53E−5	1.84E−4	1.28E−4
Lse	5.50010	3.50012	5.49986	3.49987	5.49975	3.50019
Mse	1.34E−4	7.88E−5	2.11E−4	1.17E−4	1.05E−3	1.58E−4
Proposed	5.50613	3.49123	5.50612	3.49127	5.50609	3.49156
Mse	1.38E−4	7.93E−5	1.99E−4	1.22E−4	1.20E−3	1.65E−4
EVLP	5.50022	3.49122	5.50046	3.49968	5.50079	3.49962
Mse	4.03E−4	7.93E−5	1.24E−3	6.51E−4	2.56E−3	1.28E−3

Table 2
The average estimates and their mse's of the different methods when the errors are normally distributed and the sample size is 50

	$\sigma^2=1$		$\sigma^2=2$		$\sigma^2=3$	
Parameters	5.50000	3.50000	5.50000	3.50000	5.50000	3.50000
CRLB	7.68E−6	5.33E−6	1.54E−5	1.07E−5	2.30E−5	1.60E−5
Lse	5.50008	3.50004	5.49999	3.50003	5.49986	3.49937
Mse	1.35E−5	7.91E−6	2.26E−5	1.47E−5	2.99E−5	2.10E−5
Proposed	5.50168	3.49751	5.50167	3.49751	5.50165	3.49750
Mse	1.42E−5	8.16E−6	2.24E−5	1.35E−5	3.06E−5	1.88E−5
EVLP	5.50013	3.50004	5.50024	3.50005	5.50040	3.50006
MSE	1.41E−4	7.63E−5	4.68E−4	2.42E−4	9.92E−4	4.99E−4

Table 3
The average estimates and their mse's of the different methods when the errors are double exponentially distributed and the sample size is 25

	$\sigma^2=1$		$\sigma^2=2$		$\sigma^2=3$	
Parameters	5.50000	3.50000	5.50000	3.50000	5.50000	3.50000
CRLB	6.14E−5	4.27E−5	1.23E−4	8.53E−5	1.84E−4	1.28E−4
Lse	5.50101	3.50008	5.48176	3.49675	5.49987	3.50018
Mse	1.36E−4	7.89E−5	3.19E−4	1.19E−4	1.84E−3	2.19E−4
Proposed	5.50621	3.49131	5.50603	3.49131	5.50597	3.49281
Mse	1.39E−4	8.24E−5	6.18E−4	1.26E−4	5.65E−3	1.15E−3
EVLP	5.50019	3.50003	5.50039	3.49966	5.50051	3.49935
Mse	4.01E−4	2.35E−4	1.27E−3	6.74E−4	2.50E−3	1.32E−3

Table 4
The average estimates and their mse's of the different methods when the errors are double exponentially distributed and the sample size is 50

	$\sigma^2=1$		$\sigma^2=2$		$\sigma^2=3$	
Parameters	5.50000	3.50000	5.50000	3.50000	5.50000	3.50000
CRLB	7.68E−6	5.33E−6	1.54E−5	1.07E−5	2.30E−5	1.60E−5
Lse	5.50008	3.50007	5.49999	3.50008	5.49992	3.49976
Mse	1.36E−5	7.98E−6	2.24E−5	1.51E−5	2.89E−5	1.93E−5
Proposed	5.50167	3.49752	5.50172	3.49760	5.50170	3.49756
Mse	1.44E−5	8.21E−6	2.21E−5	1.38E−5	2.96E−5	1.95E−5
EVLP	5.49995	3.49999	5.49996	3.50018	5.50040	3.50006
Mse	1.41E−4	7.62E−5	4.61E−4	2.43E−4	9.92E−4	4.99E−4

proposed estimators and the EVLP estimators, respectively. The corresponding mse's are reported immediately below in each line.

The numerical results suggest some interesting features. It is clear that for all the methods the biases decrease as the variance decreases. Therefore, all the methods provide asymptotically unbiased estimators of the frequencies. Also the mse's decrease for all the methods as the variance decreases. It is well known (Kundu, 1993) that the asymptotic variances of the lse's reach the Cramer–Rao bounds for undamped exponential model. Here we observe that the mse's of the proposed estimators are quite close to the mse's of the corresponding lse's even when the sample size is small and the error variance is high. Comparing the proposed estimators and the EVLP estimators it is observed that the mse's of the proposed estimators are significantly lower than the corresponding mse's of the EVLP estimators. From the results of Tables 3–4 it is clear that the proposed method works quite well even when the errors are not normally distributed. The proposed method may be robust when the errors are from a medium tailed distribution.

Now to observe how the proposed method works for a particular sample, we generate a particular data set from the model (4.1) with $n=100$, $\sigma^2=2.0$. The real and imaginary parts of the data are presented in Figs. 1 and 2, respectively. We use different procedures to obtain estimates of the different parameters. Using the EVLP method, we obtain the estimates of α_1^0, ω_1^0, α_2^0 and ω_2^0 as 2.05444, 5.49117, 2.67965 and 3.48793, respectively. The corresponding residual sums of squares is 6.47608. Similarly, using the proposed method we obtain the estimates as 2.61968, 5.49982, 3.02893 and 3.49940, respectively. In this case the residual sums of squares is 1.84809. The lse's become 2.61974, 5.49993, 3.02850 and 3.49989, respectively, and the corresponding residual sums of squares is 1.84749. We present the real and imaginary parts of the predicted signal obtained by different methods in Figs. 1 and 2, respectively. From the predicted signals and from the residual sums of squares it is quite clear that the proposed method works much better than the EVLP method and its performance is quite comparable to the least squares estimators.

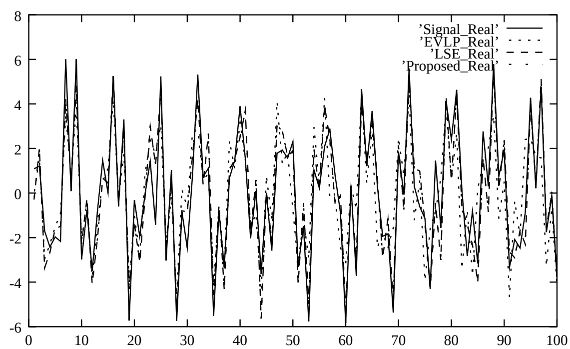

Fig. 1. Real parts of the original signal and the estimated signals are plotted along the Y-axis and the X-axis denotes time t.

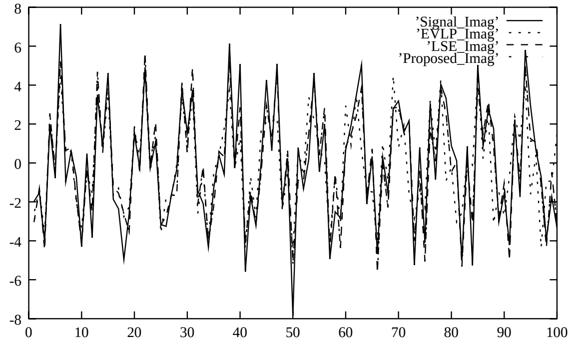

Fig. 2. Imaginary parts of the original signal and the estimated signals are plotted along the Y-axis and the X-axis denotes time t.

Now comparing the numerical computations involved in different methods, it is observed that for the computation of the lse's, one needs to compute the eigenvalues and eigenvectors of an $M \times M$ matrix at each iteration. On the other hand in our proposed method, one needs to compute the eigenvalues and eigenvectors of an $M \times M$ matrix only to obtain the initial guesses and after that we can obtain the final estimators by simple refinement. Therefore, computationally the proposed methods are less intensive compared to the lse's.

5. Conclusions

In this paper, we consider the estimation of the parameters of the undamped exponential model in presence of additive noise when the number of signals (M) is known in advance. If the number of signals is unknown, we need to use some information theoretic criterion or cross validation type techniques to estimate the number of signals (see Kundu, 1992; Kundu and Mitra, 2000). We mainly consider the estimation of the frequencies. It is well known that once the frequencies are estimated efficiently then the estimation of the amplitudes becomes a simple linear regression problem (see Rao, 1988). It is observed that the proposed method works quite well and the mse's of the frequencies obtained by the proposed methods are almost equal to the mse's of the corresponding frequencies obtained by the lse's. Since the proposed method is less computationally intensive and it guarantees the convergence within a fixed number of iteration from the given starting value (EVLP estimators) it can be used instead of the lse's for online implementation.

Acknowledgements

The authors would like to thank an Associate Editor for some very constructive suggestions.

Appendix A. Proof of Theorem 1

First let us consider the term D_n.

$$D_n = \sum_{t=1}^{n} Y(t)\mathrm{e}^{-\mathrm{i}\tilde{\omega}_j t} = \sum_{m=1}^{M} \alpha_m \sum_{t=1}^{n} \mathrm{e}^{\mathrm{i}(\omega_m - \tilde{\omega}_j)t} + \sum_{t=1}^{n} e(t)\mathrm{e}^{-\mathrm{i}\tilde{\omega}_j t}$$

$$= \sum_{m=1}^{M} \alpha_m J_m(n) + R(n) \quad \text{(say)}. \tag{A.1}$$

It can be easily observed that

$$J_m(n) = \mathrm{O_p}(1) \quad \text{if } m \neq j$$

$$= n + i(\omega_j - \tilde{\omega}_j) \sum_{t=1}^{n} t\mathrm{e}^{\mathrm{i}(\omega_j - \omega_j^*)t} \quad \text{if } m = j,$$

where ω_j^* is a point between ω_j and $\tilde{\omega}_j$. Therefore

$$J_m(n) = \mathrm{O_p}(1) \quad \text{if } m \neq j$$

$$= n + \mathrm{O_p}(n^{1-\delta}) \quad \text{if } m = j.$$

Choose L large enough such that $L\delta > 1$. Therefore, using a Taylor series approximation up to the Lth order term

$$R(n) = \sum_{t=1}^{n} e(t)\mathrm{e}^{-\mathrm{i}\tilde{\omega}_j t}$$

$$= \sum_{t=1}^{n} e(t)\mathrm{e}^{-\mathrm{i}\omega_j t} + \sum_{k=1}^{L-1} \frac{(-\mathrm{i}(\tilde{\omega}_j - \omega_j))^k}{k!} \sum_{t=1}^{n} e(t)t^k\mathrm{e}^{-\mathrm{i}\omega_j t}$$

$$+ \frac{\theta(n(\tilde{\omega}_j - \omega_j))^L}{L!} \sum_{t=1}^{n} |e(t)|, \tag{A.2}$$

where $|\theta| \leqslant 1$ and $L\delta > 1$. From (A.2), computing the order of the terms on the right-hand side, we have

$$R(n) = \mathrm{O_p}(n^{1/2}) + \sum_{k=1}^{L-1} \frac{\mathrm{O_p}(n^{-(1+\delta)k})}{k!} \mathrm{O_p}(n^{k+1/2}) + \mathrm{O_p}(1) = \mathrm{O_p}(n^{1/2}). \tag{A.3}$$

Expressions (A.1)–(A.3) imply that

$$D_n = \sum_{t=1}^{n} Y(t)\mathrm{e}^{-\mathrm{i}\tilde{\omega}_j t} = \alpha_j n(1 + \mathrm{O_p}(n^{-\delta})). \tag{A.4}$$

Similarly, it can be shown

$$C_n = \sum_{t=1}^{n} Y(t)\left(t - \frac{n}{2}\right)\mathrm{e}^{-\mathrm{i}\tilde{\omega}_j t}$$

$$= \sum_{t=1}^{n} e(t)\left(t - \frac{n}{2}\right)\mathrm{e}^{-\mathrm{i}\omega_j t} - i\alpha_j\left(\frac{n^3}{12}(1 + \mathrm{O_p}(n^{-\delta}))\right)(\tilde{\omega}_j - \omega_j). \tag{A.5}$$

Now consider (3.1),

$$\hat{\omega}_j = \tilde{\omega}_j + \frac{12}{n^2}\,\mathrm{Im}\left[\frac{C_n}{D_n}\right]$$

$$= \tilde{\omega}_j + \frac{12}{n^2}\,\mathrm{Im}\left[\frac{\sum_{t=1}^{n} e(t)(t-\frac{n}{2})\mathrm{e}^{-\mathrm{i}\omega_j t} - i\alpha_j(n^3/12(1+\mathrm{O_p}(n^{-\delta}))(\tilde{\omega}_j-\omega_j))}{\alpha_j n(1+\mathrm{O_p}(n^{-\delta}))}\right]$$

$$= \omega_j + \mathrm{O_p}(n^{-\delta})(\tilde{\omega}_j-\omega_j) + \frac{12}{n^3}\,\mathrm{Im}\left[\frac{1}{\alpha_j}\sum_{t=1}^{n} e(t)\left(t-\frac{n}{2}\right)\mathrm{e}^{-\mathrm{i}\omega_j t}\right]. \qquad \text{(A.6)}$$

Now note that

$$\mathrm{Im}\left[\left(\frac{n^3}{12}\right)^{-1/2}\sum_{t=1}^{n} e(t)\left(t-\frac{n}{2}\right)\mathrm{e}^{-\mathrm{i}\omega_j t}\right] \to N_M\left(\mathbf{0}, \frac{\sigma^2}{2}\mathbf{I}_M\right), \qquad \text{(A.7)}$$

and for $\omega_j \neq \omega_{j'}$

$$\mathrm{cov}\left\{\mathrm{Im}\left[\left(\frac{n^3}{12}\right)^{-1/2}\sum_{t=1}^{n} e(t)\left(t-\frac{n}{2}\right)\mathrm{e}^{-\mathrm{i}\omega_j t}\right],\right.$$

$$\left.\mathrm{Im}\left[\left(\frac{n^3}{12}\right)^{-1/2}\sum_{t=1}^{n} e(t)\left(t-\frac{n}{2}\right)\mathrm{e}^{-\mathrm{i}\omega_{j'} t}\right]\right\} \to 0. \qquad \text{(A.8)}$$

Therefore if $\tilde{\omega}_j - \omega_j = \mathrm{O_p}(n^{-1-\delta})$ and $\delta \leqslant \frac{1}{4}$, then from (A.6) and (A.7), it is clear that $\hat{\omega}_j - \omega_j = \mathrm{O_p}(n^{-1-2\delta})$. Similarly if $\delta > 1/4$ using (A.6), (A.7) and (A.8), part (2) of Theorem 1 follows immediately.

References

Bai, Z.D., Krishnaiah, P.R., Zhao, L.C., 1986. On the simultaneous estimation of the number of signals and frequencies under a model with multiple sinusoids. Technical report, 86-37, Center for Multivariate Analysis, University of Pittsburgh.

Barrodale, I., Olesky, D.D., 1981. Exponential approximation using Prony's method. In: Baker, C.T.H., Phillips, C. (Eds.), The Numerical Solution of Non-Linear Problems, pp. 258–269.

Bresler, Y., Macovski, A., 1986. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. IEEE Trans. Acoust. Speech Signal Process. ASSP-34 34, 1081–1089.

Kay, S.M., 1984. Accurate frequency estimation at low signal to noise ratio. IEEE Trans. Acoust. Speech Signal Process. ASSP-32 32 (3), 540–547.

Kay, S.M., 1988. Modern Spectral Estimation; Theory and Applications. Prentice-Hall, Englewood Cliffs, NJ.

Kay, S., Marple, S.L., 1981. Spectral analysis: a modern perspective. Proc. IEEE 69, 1380–1419.

Kumaresan, R., Scharff, L.L., Shaw, A.K., 1986. An algorithm for pole-zero modeling and spectral analysis. IEEE Trans. Acoust. Speech Signal Process. ASSP-34 34, 637–640.

Kundu, D., 1992. Estimating the number of signals using information theoretic criterion. J. Statist. Comput. Simulation 44, 117–131.

Kundu, D., 1993. Estimating the parameters of undamped exponential signals. Technometrics 35, 215–218.

Kundu, D., Mitra, A., 1995. Consistent method of estimating superimposed exponential signals. Scand. J. Statist. 22 (1), 73–82.

Kundu, D., Mitra, A., 1999. On asymptotic properties and confidence intervals for exponential signals. Signal Process. 72 (2), 129–139.

Kundu, D., Mitra, A., 2000. Detecting the number of signals for an undamped exponential model using cross validation technique. Signal Process. 80 (3), 525–534.

Osborne, M.R., Smyth, G.K., 1995. A modified Prony algorithm for fitting sum of exponential functions. SIAM J. Sci. Statist. Comp. 16 (1), 119–138.

Prasad, S., Chakroborty, M., Parthasarathy, H., 1995. The role of statistics in signal processing — A brief review and some emerging trends. Ind. J. Pure Appl. Math. 26 (6), 547–578.

Prony, R., 1795. Essai experimentale et analytique. J. Ecole Polytechnique (Paris) 1, 24–76.

Rao, C.R., 1988. Some recent results in signal detection. In: Gupta, S.S., Berger, J.O. (Eds.), Statistical Decision Theory and Related Topics, pp. 319–332.

Rao, C.R., Zhao, L.C., 1993. Asymptotic behavior of the maximum likelihood estimators of superimposed exponential signals. IEEE Trans. Signal Process. 41, 1461–1463.

Roy, R.H., 1987. ESPRIT-Estimation of signal parameter via rotational invariance techniques. Ph.D. Thesis, Stanford University, CA, USA.

Schmidt, R.O., 1981. A signal subspace approach to multiple emitter location and spectral estimation. Ph.D. Thesis, Stanford University, CA, USA.

Stoica, P., 1993. List of references on spectral line analysis. Signal Process. 31 (3), 329–340.

Tufts, D.W., Kumaresan, R., 1982. Estimation of frequencies of multiple sinusoids: making linear prediction perform like maximum likelihood. Proc. IEEE 70, 975–989.

Varah, J.M., 1985. On fitting exponentials by non-linear least squares. SIAM J. Sci. Statist. Comput. 6, 30–44.