Computational Statistics & Data Analysis 18 (1994) 525–534
North-Holland

Estimating the parameters of complex-valued exponential signals

Debasis Kundu
Department of Mathematics, Indian Institute of Technology, Kanpur, India

Received August 1992
Revised July 1993

Abstract: Estimating the parameters of complex-valued exponential signals is a very well known problem in signal processing. Osborne (1975) proposed a new iterative algorithm to estimate the unknown parameters when they are real-valued. In this paper the idea has been generalized to the complex parameters case. Theoretical justification has been provided for the convergence of the iterative process.

Keywords: Non-linear regression; Consistent estimator; Asymptotic stability; Least-squares estimation

1. Introduction

In this paper we consider an estimation procedure for the following non-linear regression model. The value of a complex-valued random variable y is observed at n equidistant time points $t_1,\ldots,t_n$. We assume that the observed value $y(t_i)$ at time t_i satisfies

$$y(t_i)=\mu(\alpha_0,\beta_0,t_i)+\epsilon_i,\quad i=1,\ldots,n, \tag{1.1}$$

where $\mu(\alpha,\beta,t)=\sum_{j=1}^{M}\alpha_j e^{\beta_j t}$, $\alpha=(\alpha_1,\ldots,\alpha_M)$ and $\beta=(\beta_1,\ldots\beta_M)$. Here the α_j's and β_j's are unknown complex numbers, the β_j's are assumed to be distinct. ϵ_i is a complex-valued random variable with mean zero for $i=1,\ldots,n$; M is assumed to be known.

This is an important problem in signal processing (see Kay and Marple, 1981). For example, in electromagnetic pulse (EMP) situations (Ricketts, et al., 1976; Sircar, 1987), the EMP pickup can be characterized by a sum of complex exponentials whose parameters are to be determined. The parameters are a means of coding the various pulse wave forms, and the signal approximation thus obtained can be readily employed to analyze responses in various subsys-

Correspondence to: Prof. D. Kundu, Indian Institute of Technology, Kanpur Pin 208016, India.

SSDI 0167-9473(93)E0033-Z

tems under EMP environment. In system identification problems, the characterization of the impulse response of a linear system by a sum of complex exponentials, and then by identifying/approximating the complex amplitudes and natural frequencies with high degree of accuracy, has its special importance in a wide variety of applications. The study of the transient behavior of a system is one of them. The problem is well known to be numerically difficult (Varah, 1985). The general purpose algorithm such as Gauss–Newton, Newton–Rapson or their variants often has great difficulty in converging. Osborne (1975) proposes a new algorithm to estimate the unknown parameters when they are real-valued. He shows by numerical experiments that the proposed algorithm performs much better than any other standard non-linear regression algorithm for the real data sets. The aim of this paper is to generalize the algorithm to the complex parameters case.

Without any distributional assumptions on the error random variables, it may be reasonable to estimate the linear and the non-linear parameters α_j and β_j by minimizing the norm squares error, i.e.,

$$\min_{\alpha,\beta\in\Theta} \frac{1}{n} \sum_{i=1}^{n} |y(t_i) - \mu(\alpha, \beta, t_i)|^2, \tag{1.2}$$

where Θ is a compact subset of C^{2M}. Under the assumption of normality of error components, the least norm squares estimates are the maximum likelihood estimates also. Without a boundedness condition on the t_i's, the maximum likelihood estimates may not be consistent. Wu (1981) and Malinvaud (1970) show that for the first-order decay model

$$y(t) = e^{-\alpha t} + \epsilon_t, \; t = 1, \ldots, n, \tag{1.3}$$

the least squares estimate of α is inconsistent. Osborne and Smyth (1986) also point out the same problem for the model (1.1) when the parameters are real-valued. In this paper we formalize the model in such a way that it satisfies the consistency property of the estimates for large samples as well as coinciding with the model (1.1) for finite samples.

The organization of this paper is as follows. In Section 2, we state the asymptotic assumption of the model clearly. In Section 3, we derive the normal equation. In Section 4, we propose an iterative procedure to obtain the least norm squares estimates. It is observed that the iterative procedure converges very rapidly; some numerical results are presented in Section 5. We give the theoretical justification for convergence of the new algorithm in Section 6.

2. Problem statement and assumptions

It is important to redefine the model (1.1) because of the restriction of t_i. We redefine the model (1.1) as follows:

$$y_{ni} = \mu(\alpha_0, \beta_0, t_{ni}) + \epsilon_{ni}, \quad i = 1, \ldots, n, \tag{2.1}$$

where the α_j's and β_j's are same as defined before, $t_{ni} = i/n$, $\{\epsilon_{ni}\}$, $i = 1, \ldots n$ and $n = 1,2,\ldots$ is a double array of complex-valued random variables. $\{\epsilon_{ni}\}$; $i = 1, \ldots, n$ is assumed to be i.i.d. with mean zero. It is important to observe that the models (1.1) and (2.1) are equivalent for any finite n, although asymptotically they may be different. Let us assume $\theta = (\alpha, \beta)$. We need to make the following assumption.

Assumption 1. $\{\epsilon_{ni}\}$, $i = 1, \ldots, n$ and $n = 1,2,\ldots$ is a double array of complex-valued random variables with mean zero and finite fourth moment for both the real and imaginary part. The real and imaginary part are assumed to be independently distributed. The parameter space Θ is a compact subset of C^{2M} and the true parameter vector $\theta_0 = (\alpha_0, \beta_0)$ is an interior point of Θ.

Observe that the function

$$R(\alpha, \beta) = \int_0^1 |\mu(\alpha, \beta, t) - \mu(\alpha_0, \beta_0, t)|^2 \, dt, \tag{2.2}$$

has a unique minimum at $(\alpha, \beta) = (\alpha_0, \beta_0)$ if β_j's are distinct. Therefore under the Assumption 1, the least norm squares estimates of the model (2.1) are strongly consistent (Kundu, 1991).

3. Normal equation

Prony (1795) observes that for any α and β there exist complex constants $\gamma = (\gamma_1, \ldots, \gamma_{M+1})$ depending only on β such that

$$\sum_{j=1}^{M+1} \gamma_j \mu(\alpha, \beta, t_{n\,k+j}) = 0, \tag{3.1}$$

for $k = 0, 1, \ldots, n - M - 1$. Here $\gamma_1, \ldots, \gamma_{M+1}$ are such that $\bar{e}^{\beta_{1/n}}, \ldots, \bar{e}^{\beta_{M/n}}$ are the roots of the following polynomial equation

$$\gamma_{M+1} z^M + \gamma_M z^{M-1} + \cdots + \gamma_2 z + \gamma_1 = 0. \tag{3.2}$$

Therefore we can conclude that there exists an $n \times (n - M)$ matrix $X(\gamma)$ such that $X^*(\gamma)\mu_{\alpha,\beta} = 0$. Here $\mu_{\alpha,\beta}$ is an $n \times 1$ vector with i-th element $\mu(\alpha, \beta, t_{ni})$ and

$$X^*(\gamma) = \begin{bmatrix} \gamma_1 & \cdot & \cdot & \cdot & \gamma_{M+1} & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \gamma_1 & \cdot & \cdot & \gamma_{M+1} & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 0 & \cdot & \cdot & 0 & \gamma_1 & \cdot & \cdot & \gamma_{M+1} \end{bmatrix}, \tag{3.3}$$

where '$*$' denotes the conjugate transpose of a matrix or of a vector.

Observe that it is possible to separate the linear parameters α_j from the non-linear parameters β_j of the model (2.1). It can be written in the following form

$$Y_n = A(\beta_0)\alpha_0 + \epsilon_n, \tag{3.4}$$

where $\boldsymbol{Y}_n^{\mathrm{T}} = (y_{n1}, \ldots, y_{nn})$, $\boldsymbol{\epsilon}_n^{\mathrm{T}} = (\epsilon_{n1}, \ldots, \epsilon_{nn})$ and $\mathbf{A}(\beta)$ is an $n \times M$ matrix whose (i, j)th element is $\bar{e}^{\beta_j t_{in}}$ for $i = 1, \ldots, n$ and $j = 1, \ldots, M$. Here α_0 and β_0 are the true parameter vectors as it is in (2.1).

The least norm squares estimates of (α, β) can be obtained by minimizing R_n with respect to (α, β) where

$$R_n = \frac{1}{n}(\boldsymbol{Y}_n - \boldsymbol{A}(\beta)\alpha)^*(\boldsymbol{Y} - \boldsymbol{A}(\beta)\alpha). \tag{3.5}$$

Observe that for a fixed value of β, the minimization of R_n with respect to α is a simple linear regression problem. The solution of $\hat{\alpha}(\beta)$ is

$$\hat{\alpha}(\beta) = (\boldsymbol{A}^*(\beta)\boldsymbol{A}(\beta))^{-1}\boldsymbol{A}^*(\beta)\boldsymbol{Y}_n, \tag{3.6}$$

substituting $\hat{\alpha}(\beta)$ in (3.5) we obtain

$$Q_n(\beta) = \frac{1}{n}\boldsymbol{Y}_n^*(\boldsymbol{I} - \boldsymbol{P}_{A(\beta)})\boldsymbol{Y}_n. \tag{3.7}$$

Here $\boldsymbol{P}_{A(\beta)} = \boldsymbol{A}(\beta)\ (\boldsymbol{A}^*(\beta)\ \boldsymbol{A}(\beta)^{-1}\boldsymbol{A}^*(\beta)$ is the projection matrix on the range space spanned by the columns of the matrix $\boldsymbol{A}(\beta)$. The least norm squares estimate $\hat{\beta}$ of β can be obtained by minimizing $Q_n(\beta)$ with respect to β.

We have seen already that for a given β there exists a γ such that $\boldsymbol{X}^*(\gamma)\mu_{\alpha,\beta} = 0$ for all α, therefore $\boldsymbol{X}^*(\gamma)\boldsymbol{A}(\beta) = 0$. This implies $(\boldsymbol{I} - \boldsymbol{P}_{A(\beta)}) = \boldsymbol{P}_{x(\gamma)}$ therefore

$$\min_{\gamma} \frac{1}{n}\boldsymbol{Y}_n^*\boldsymbol{P}_{X(\gamma)}\boldsymbol{Y}_n = \min_{\beta} \frac{1}{n}\boldsymbol{Y}_n^*(\boldsymbol{I} - \boldsymbol{P}_{A(\beta)})\boldsymbol{Y}_n. \tag{3.8}$$

Let us denote

$$\psi_n(\gamma) = \frac{1}{n}\boldsymbol{Y}_n^*\boldsymbol{P}_{X(\gamma)}\boldsymbol{Y}_n \tag{3.9}$$

Observe that $\psi_n(\gamma)$ is invariant under a scalar multiplication, i.e., $\psi_n(\gamma) = \psi_n(c\gamma)$ for any complex constant c. Therefore

$$\min_{\gamma_1, \ldots, \gamma_{M+1}} \psi_n(\gamma) = \min_{\gamma : \gamma_{M+1} = 1} \psi_n(\gamma). \tag{3.10}$$

To minimize $\psi_n(\gamma)$ with respect to γ, we differentiate $\psi_n(\gamma)$ with respect to the real and imaginary parts of γ and equate them to zero. This leads to solving a matrix equation of the following form

$$\frac{1}{n}\boldsymbol{B}(\gamma)\gamma = 0,\ \gamma_{M+1} = 1, \tag{3.11}$$

where $\boldsymbol{B}$ is an $(M+1) \times (M+1)$ Hermitian matrix whose (i, j)th element is given by

$$b_{ij}(\gamma) = \boldsymbol{Y}_n^*\boldsymbol{X}_i(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}_j^{\mathrm{T}}\boldsymbol{Y}_n - \boldsymbol{Y}_n^*\boldsymbol{X}(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}_j^{\mathrm{T}}\boldsymbol{X}_i(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}^*\boldsymbol{Y}_n, \tag{3.12}$$

$i, j = 1, \ldots, M+1$, see Kundu (1993) for details. Here $\boldsymbol{X}_j^{\mathrm{T}}$ for $j = 1, \ldots, M+1$ is an $(n-M) \times n$ matrix whose $(k, k+j-1)$th element is one for $k = 1, \ldots, M+1$ and the rest of the elements are zeroes. Both the matrices $\boldsymbol{X}$ and $\boldsymbol{X}^*$ have been evaluated at γ.

This is a non-linear eigenvalue problem. We discuss in detail the solution procedure of such a problem in Section 4.

4. Modified Prony Algorithm

It is clear that γ satisfying (3.11) should be an eigenvector corresponding to the zero eigenvalue of the matrix $\boldsymbol{B}(\gamma)$. We suggest the following iterative scheme similar to that of Osborne (1975) to solve (3.11):

$$\left(\boldsymbol{B}\left(\gamma^{(k)}\right) - \lambda^{(k+1)}\boldsymbol{I}\right)\gamma^{(k+1)} = 0, \quad \gamma_{M+1}^{(k+1)} = 1. \tag{4.1}$$

Here $\gamma^{(k)}$ denotes the k-th iterate of the above iterative process and $\lambda^{(k+\mathrm{i})}$ is the eigenvalue of $\boldsymbol{B}(\gamma^{(k)})$ which is closest to zero. The iterative procedure should be stopped when $|\lambda^{(k+1)}|$ is small compared to $\|\boldsymbol{B}\|(=\Sigma_i\Sigma_j|b_{ij}|)$. The suggested Modified Prony Algorithm has the following form:

Step 1. Set an initial value $\gamma^{(1)}$ and make the $(M+1)$th component to be 1, i.e.,

$$\gamma^{(1)} = \gamma^{(1)}/\gamma_{M+1}, \quad i = 1.$$

Step 2. Calculate the matrix $\boldsymbol{B}(\gamma^{(k)})$

Step 3. Find the eigenvalue $\lambda^{(k+1)}$ of $\boldsymbol{B}(\gamma^{(k)})$ which is closest to zero and normalize (make the $(M+1)$th component to be 1) the corresponding eigenvector.

Step 4. Test the convergence by testing

$$|\lambda^{(k+1)}| < \epsilon \|\boldsymbol{B}\|.$$

Step 5. If the test in Step 4 fails, $i = i+1$ and go to Step 2.

Once $\hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_{M+1})$ is obtained, construct the polynomial equation

$$\hat{\gamma}_{M+1}z^M + \hat{\gamma}_M z^{M-1} + \cdots + \hat{\gamma}_2 z + \hat{\gamma}_1 = 0, \tag{4.2}$$

obtain the roots of the form $e^{\hat{\beta}_1/n}, \ldots, e^{\hat{\beta}_M/n}$ and take $\hat{\beta}_1, \ldots, \hat{\beta}_M$ (with proper rearrangements) as estimates of $\beta_1, \ldots, \beta_M$.

5. Numerical results

In this section we study the performance of the Modified Prony Algorithm by way of computer simulation. We perform a Monte Carlo simulation study for a model with different sample sizes and with different error variances. All these

Table 5.1

n	σ	Iteration count	$\hat{\beta}_1$	$\hat{\beta}_2$	Smallest eigenvalue
32	0.001	2	−9.10597+i2.22149	−4.12523+i1.14496	0.9891E-07
	0.005	2	−9.09885+i2.25540	−4.11404+i1.16644	0.9210E-07
	0.01	3	−9.16201+i2.20155	−4.14702+i1.13760	0.9988E-07
	0.05	6	−8.15151+i2.29112	−1.37858+i1.28955	0.1321E-07
64	0.001	2	−9.12297+i2.23095	−4.13851+i1.15352	0.9241E-07
	0.005	2	−9.14451+i2.25942	−4.14106+i1.16898	0.1159E-06
	0.01	3	−9.06592+i2.26906	−4.07941+11.19358	0.1086E-06
	0.05	4	−8.99470+i2.35912	−3.77917+i0.60891	0.3787E-06
96	0.001	2	−9.10993+i2.23145	−4.12827+i1.15038	0.9306E-07
	0.005	2	−9.10902+i2.22615	−4.12373+i1.14745	0.1127E-06
	0.01	2	−9.09247+i2.22581	−4.12187+i1.13571	0.1121E-06
	0.05	4	−8.57545+i2.99601	−3.82582+i1.64940	0.3917E-05
128	0.001	2	−9.11449+i2.24339	−4.12821+i1.16141	0.9537E-07
	0.005	2	−9.10954+i2.24384	−4.12996+i1.16314	0.9443E-07
	0.01	2	−9.14165+i2.25532	−4.17314+i1.17229	0.9483E-07
	0.05	3	−9.27804+i1.86285	−3.38798+i1.14538	0.1068E-07

computations are performed on IBM-4381 at the Pennsylvania State University using double precision in Fortran-77.

We present the result in the Table 5.1, when the algorithm (4.1) was applied to the model (2.1) with $M = 2$, $\alpha_1 = (2 + i2)$, $\alpha_2 = (0.5 + i1)$, $\beta_1 = (-9.11542 + i2.23415)$ and $\beta_2 = (-4.13158 + i1.15467)$. The data is generated using different standard deviations viz., $\sigma = 0.001, 0.005, 0.01, 0.05$ and with different sample sizes namely $n = 32, 64, 96, 128$. We consider the case when the real and the imaginary parts of the noise random variables are independently distributed with mean zero and equal variances (i.e., $\sigma_1^2 = \sigma_2^2 = \sigma^2/2$). The data points of the mean vector μ are evaluated at equidistant time points on the closed interval [0,1], i.e., for sample size 32, the distance between two consecutive time points is $\frac{1}{32}$. IMSL random deviate generator is used with DSEED = 12345.DO. For each value of n and σ one hundred and one different data sets are generated. The iterative process is stopped in each case when, $|\lambda^{(k+1)}| < 10^{-15}\sum_{i,j}|b_{ij}(\lambda^{(1)})|$.

The median iteration count and the mean estimates of β_1 and β_2 are reported. The eigenvector corresponding to the smallest eigenvalue of the matrix $\boldsymbol{U}^*\boldsymbol{U}$ where

$$\boldsymbol{U} = \begin{bmatrix} y_1 & \cdots & y_{M+1} \\ \vdots & & \vdots \\ y_{n-M} & \cdots & y_n \end{bmatrix}, \tag{5.1}$$

is used as an initial value in each iteration. It is interesting to observe that if $\sigma = 0$ then the initial estimate of β obtained from $\boldsymbol{U}^*\boldsymbol{U}$ is the true value of β.

So it is expected that when σ is not too large, the initial estimate will be close to the true value. Osborne (1975) also uses this as an initial estimate of β.

The successive iterations show very rapid convergence of the Modified Prony Algorithm. The main trend is that the number of iterations decreases, for a given σ, as n increases. For a fixed σ, the performance of the estimates of β improves as sample size increases. We also observe (not reported here) that as the distance between β increases the performance in iteration count and in estimation also improves for fixed σ and n.

6. Asymptotic stability

In this section we try to give some theoretical support for the convergence of the numerical algorithm (4.1). A complete proof of asymptotic stability is omitted, for details the reader is referred to Kundu (1989).

The iteration of the algorithm can be written as

$$\boldsymbol{F}(\boldsymbol{u}^{(k)}, \boldsymbol{v}^{(k)}) = (\boldsymbol{u}^{(k+1)}, \boldsymbol{v}^{(k+1)}), \tag{6.1}$$

where $\boldsymbol{u}^{(k)}$, and $\boldsymbol{v}^{(k)}$ are the real and imaginary parts of the first M components of $\gamma^{(k)}$, i.e.

$$\gamma^{(k)} = \left(u_1^{(k)} + \mathrm{i}v_1^{(k)}, \ldots, u_M^{(k)}, 1\right) = (\boldsymbol{u}^{(k)}, \boldsymbol{v}^{(k)}). \tag{6.2}$$

Here $\boldsymbol{F}: \mathrm{R}^{2M} \to \mathrm{R}^{2M}$ is defined implicitly through (4.1). Therefore the convergence matrix $\dot{\boldsymbol{F}}$ of the iterative process can be written as

$$\dot{\boldsymbol{F}}(\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}) = \left[\begin{array}{cc} \dfrac{\mathrm{d}u^{(k+1)}}{\mathrm{d}u^{(k)}} & \dfrac{\mathrm{d}v^{(k+1)}}{\mathrm{d}v^{(k)}} \\ \dfrac{\mathrm{d}u^{(k+1)}}{\mathrm{d}v^{(k)}} & \dfrac{\mathrm{d}v^{(k+1)}}{\mathrm{d}v^{(k)}} \end{array}\right]_{(u,v)=(\hat{u},\hat{v})}. \tag{6.3}$$

Here $(\hat{\boldsymbol{u}} + \mathrm{i}\hat{\boldsymbol{v}}, 1) = \hat{\gamma}$ is the solution of (3.11) and all the four submatrices are of the order $M \times M$. The (i, j)th, element of $\mathrm{d}u^{(k+1)}/\mathrm{d}u^{(k)}$ is obtained by taking the partial derivative of $u_i^{(k+1)}$ with respect to $u_j^{(k)}$ and evaluate it at the point $(\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}})$. The sufficient condition, that $\hat{\gamma}$ is a point of attraction of the iterative process (4.1), is the spectral density of $\dot{\boldsymbol{F}}(\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}})$ which is less than one (see Ortega and Rheinboldt, 1970). Writing (4.1) explicitly using the expression of $\boldsymbol{B}$ from (3.12) we obtain

$$\sum_{i=1}^{M+1} \boldsymbol{Y}_n^* \boldsymbol{X}_j (\boldsymbol{X}^*\boldsymbol{X})^{-1} \boldsymbol{X}_i^{\mathrm{T}} \gamma_i^{(k+1)} - \sum_{i-1}^{M+1} \boldsymbol{Y}_n^* \boldsymbol{X} (\boldsymbol{X}^*\boldsymbol{X})^{-1} \boldsymbol{X}_i^{\mathrm{T}} \boldsymbol{X}_j (\boldsymbol{X}^*\boldsymbol{X})^{-1}$$

$$\boldsymbol{X}^*\boldsymbol{Y}_n \gamma_i^{(k+1)} = \lambda^{(k+1)} \gamma_j^{(k+1)}, \qquad j = 1, \ldots, M+1, \quad \gamma_{M+1}^{(k+1)} = 1. \tag{6.4}$$

Here both $\boldsymbol{X}^*$ and $\boldsymbol{X}$ are evaluated at $\gamma^{(k)}$. The elements of the matrix $\dot{\boldsymbol{F}}(\hat{\boldsymbol{u}}, \hat{\boldsymbol{v}})$ can be obtained by differentiating both sides of (6.4) with respect to $u_i^{(k)}$ and

$v_i^{(k)}$ for $i = 1, \ldots, M$ and evaluating at $\hat{\gamma}$. Differentiating both sides of (6.4) with respect to $u_i^{(k)}$ for i $= 1, \ldots, M$ and evaluating at the point $\hat{\gamma}$ yields

$$\boldsymbol{B}(\hat{\gamma})\frac{\mathrm{d}\hat{\gamma}}{\mathrm{d}u_i} + \boldsymbol{C}_i(\hat{\gamma}) = \frac{\mathrm{d}\hat{\lambda}}{\mathrm{d}u_i}\hat{\gamma} + \hat{\lambda}\frac{\mathrm{d}\hat{\gamma}}{\mathrm{d}u_i}. \tag{6.5}$$

Here $\boldsymbol{C}_i(\hat{\gamma})$ is a $(M+1)\times 1$ vector whose j-th. element $c_{ij}(\hat{\gamma})$, for $j = 1, \ldots, (M+1)$ is given by

$$\begin{aligned} c_{ij}(\hat{\gamma}) = & -\boldsymbol{Y}_n^*\boldsymbol{X}_j(\boldsymbol{X}^*\boldsymbol{X})^{-1}(\boldsymbol{X}_i^*\boldsymbol{X}+\boldsymbol{X}^*\boldsymbol{X}_i)(\boldsymbol{X}^*\boldsymbol{X})^{-1} \\ & \times \boldsymbol{X}^*\boldsymbol{Y}_n - \boldsymbol{Y}_n^*\boldsymbol{X}_i(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}^*\boldsymbol{X}_j(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}^*\boldsymbol{Y}_n \\ & + \boldsymbol{Y}_n^*\boldsymbol{X}(\boldsymbol{X}^*\boldsymbol{X})^{-1}(\boldsymbol{X}_i^*\boldsymbol{X}+\boldsymbol{X}^*\boldsymbol{X}_i)(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}^*\boldsymbol{X}_j(\boldsymbol{X}^*\boldsymbol{X})^{-1} \\ & \times \boldsymbol{X}^*\boldsymbol{Y}_n + \boldsymbol{Y}_n^*\boldsymbol{X}(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}^*\boldsymbol{X}_j(\boldsymbol{X}^*\boldsymbol{X})^{-1} \\ & \times (\boldsymbol{X}_i^*\boldsymbol{X}+\boldsymbol{X}^*\boldsymbol{X}_i)(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}^*\boldsymbol{Y}_n \\ & - \boldsymbol{Y}_n^*\boldsymbol{X}(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}^*\boldsymbol{X}_j(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}_i^*\boldsymbol{Y}_n. \end{aligned} \tag{6.6}$$

It is easy to see that

$$\lim_{n\to\alpha}\frac{1}{n}\left\{\frac{\mathrm{d}\hat{\lambda}}{\mathrm{d}u_i}\hat{\gamma} + \hat{\lambda}\frac{\mathrm{d}\hat{\gamma}}{\mathrm{d}u_i}\right\} = 0, \quad i = 1, \ldots, M, \tag{6.7}$$

because all the quantities inside the parentheses are bounded over a compact set.

The main idea of the proof of the asymptotic stability of the iterative process is as follows: $(1/n)\boldsymbol{C}_i(\hat{\gamma})$ converges to a zero vector almost surely for all $i = 1, \ldots, M$ and $(1/n)\boldsymbol{B}(\hat{\gamma})$ converges to a positive semidefinite matrix with null space spanned by γ_0, the true parameter vector, corresponding to β_0. Therefore from (6.5) it follows that there exist a complex sequence $\{a_n\}$ such that

$$\lim_{n\to\alpha}\left(\frac{\mathrm{d}\hat{\gamma}}{\mathrm{d}u_i} - a_n\gamma_0\right) = 0. \tag{6.8}$$

Since $\hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_M, 1)$, therefore $\mathrm{d}\hat{\gamma}_{M+1}/\mathrm{d}u_i = 0$. Again, since the $(M+1)$th, component of γ_0 is 1 $(\neq 0)$ therefore from (6.8) it follows that a_n should converge to zero. The same reasoning holds for $\mathrm{d}\hat{\gamma}_{M+1}/\mathrm{d}u_i$, i $= 1, \ldots, M$. Therefore $\dot{\boldsymbol{F}}$, the convergent matrix, converges to zero matrix almost surely. This implies the asymptotic stability of the proposed algorithm.

We will not prove in detail the above two claims. Interested readers are referred to Kundu (1989), where the details are available. To get some idea we calculate their expectations to see the behaviour for finite n. We show that

$$\begin{aligned} & E\left(\frac{1}{n}\boldsymbol{B}(\gamma_0)\right) = \left(\left(\frac{1}{n}b_{ij}(\gamma_0)\right)\right) = \left(\left(\frac{1}{n}\mu^*_{\alpha_0,\beta_0}\boldsymbol{X}_i(\boldsymbol{X}^*\boldsymbol{X})^{-1}\boldsymbol{X}_j^*\mu_{\alpha_0,\beta_0}\right)\right) \\ & E(\boldsymbol{C}_i(\gamma_0)) = 0, \; i = 1, \ldots, M, \end{aligned} \tag{6.9}$$

here E stands for expectation and μ_{α_0,β_0} is an n-vector as it is defined in Section 3. Both the matrices X and X^* have been evaluated at γ_0, the true parameter value, and the expectation is also taken under the true parameter value. Observe that $E((1/n)B(\gamma_0))$ is a positive semidefinite matrix with null space spanned by γ_0. It can be shown that (see Kundu; 1989):

$$\lim_{n\to\alpha}\frac{1}{n}B(\hat{\gamma}) = \lim_{n\to\alpha}E\left(\frac{1}{n}B(\gamma_0)\right) = V_0$$

and

$$\lim_{n\to\alpha}\frac{1}{n}C_i(\hat{\gamma}) = \lim_{n\to\alpha}E\left(\frac{1}{n}C_i(\gamma_0)\right) = 0, \qquad i = 1,\ldots,M, \tag{6.10}$$

where V_0 is a positive semidefinite matrix with null space spanned by γ_0.

From (3.12) it follows that

$$\begin{aligned} E\big(b_{ij}(\gamma_0)\big) &= \operatorname{trace}\Big\{X_i(X^*X)^{-1}X_j^{\mathrm{T}}\big[(\sigma_1^2+\sigma_2^2)+\mu_{\alpha_0,\beta_0}\mu^*_{\alpha_0,\beta_0}\big]\Big\} \\ &\quad - \operatorname{trace}\Big\{X(X^*X)^{-1}X_j^{\mathrm{T}}X_i(X^*X)^{-1}X\big(\sigma_1^2+\sigma_2^2\big)\Big\} \\ &= \mu^*_{\alpha_0,\beta_0}X_i(X^*X)^{-1}X_j^*\mu_{\alpha_0,\beta_0} \qquad (\text{since } X^*\mu_{\alpha_0,\beta_0}=0). \end{aligned} \tag{6.11}$$

Here σ_1^2 and σ_2^2 denote the variance of the real and imaginary part of the error random variable. $E(C_i(\gamma_0))$ also can be calculated mainly by using the fact $\sum_{j=1}^{M+1}X_j^*\gamma_{0j} = X^*(\gamma_0)$, where γ_{0j} denotes the j-th component of γ_0. From (6.6) it follows that

$$\begin{aligned} &E\big(c_{ij}(\gamma_0)\big) = \\ &\operatorname{trace}\Big\{-X_j(X^*X)^{-1}(X_i^*X+X^*X_i)(X^*X)^{-1}X^* \\ &\quad -X_i(X^*X)^{-1}X^*X_j(X^*X)^{-1}X^* \\ &\quad +X(X^*X)^{-1}(X_i^*X+X^*X_i)(X^*X)^{-1}X^*X_j(X^*X)^{-1}X^* \\ &\quad +X(X^*X)^{-1}X^*X_j(X^*X)^{-1}(X_i^*X+X^*X_i)(X^*X)^{-1}X^* \\ &\quad -X(X^*X)^{-1}X^*X_j(X^*X)^{-1}X_i^*\big(\sigma_1^2+\sigma_2^2\big)\Big\} \\ &= 0. \end{aligned} \tag{6.12}$$

Acknowledgements

The author wishes to thank Professor C.R. Rao for all his guidance in the preparation of this paper. The author also thanks one Associate Editor and the referees for many valuable and constructive suggestions.

References

Kay, S.M. and S.L. Marple, 'Spectrum analysis – A modern perspective'; Proc. IEEE, Vol. 69, **11** (1981) 1380–1419.

Kundu, D., Some results in estimating the parameters of exponential signals in noise, Ph.D. Thesis (Department of Statistics, The Pennsylvania State University, 1989).

Kundu, D., Estimating the parameters of undamped exponential signals, *Technometrics*, **35**, 2, 1993, 215–218.

Kundu, D., Asymptotic properties of the complex-valued nonlinear regression model, *Commun. Statist., Ser. A*, **20** (1991) 3793–3803.

Malinvaud, E., The consistency of non-linear regression, *Ann. Math. Statist.*, 41 (1970) 956–969.

Ortega, J.M. and W.C. Rheinboldt, *Iterative solution of non-linear equation in several variables*, (Academic Press, New York, 1970).

Osborne, M.R., Some special non-linear least squares problem, *SIAM Num. Anal.*, 12, 4 (1975) 572–592.

Osborne, M.R. and G. Smyth, An algorithm for exponential fitting revisited, *J. App. Prob.* (1986) 419–430.

Prony, R., Essay Experimentale et Analytique, *J. École Polytechnique* (Paris) (1795) 24–76.

Ricketts, L.W., J.E. Bridges and J. Miletta, *EMP radiation and protective techniques* (Wiley, New York, 1976).

Sircar, P., Accurate parameter estimation of damped sinusoidal signals samples at non-uniform spacings, Ph.D. Thesis (Syracuse University, 1987).

Varah, J.M., On fitting exponentials by non-linear least squares, *SIAM J. Sc. Stat. Comp.*, 6 (1985) 30–44.

Wu, C.F.J., Asymptotic theory of non-linear least squares estimation, Ann. Stat., 9 (1981) 501–513.