Estimating parameters in the damped exponential model

Nandini Kannan[a], Debasis Kundu[b], *

[a] *Division of Mathematics and Statistics, The University of Texas at San Antonio, San Antonio, TX 78249, USA*
[b] *Department of Mathematics, Indian Institute of Technology Kanpur, P.O. Box I.I.T., Kanpur 208016, India*

Abstract

In this paper we consider the problem of estimation of the frequencies and damping factors of exponential signals in the presence of noise. We propose a non-iterative method based on using forward–backward linear prediction and the notion of extended order modeling. In addition to providing estimators of the unknown parameters in the model, the proposed method can be used to specify initial values in any standard minimization algorithm to obtain the least-squares estimators. For the undamped exponential model, it is well known that any estimator is inconsistent under the usual definition of consistency. We redefine the model so that the sampling interval is finite, and prove the consistency and asymptotic normality of the least-squares estimators under this new assumption. It is observed that the dispersion matrix of the least-squares estimators attains the Cramer–Rao bound. © 2001 Elsevier Science B.V. All rights reserved.

1. Introduction

Consider the model where the complex response at time t, y_t is composed of a sum of M superimposed damped exponential signals and additive noise, i.e.

$$y_t = \sum_{k=1}^{M} \alpha_k \, e^{(-\delta_k + j\omega_k)t} + e_t, \quad \text{for } t = 1, \ldots, n.$$

(1)

This model is referred to as the multiple sinusoid model. Here δ_k's are the damping factors of the signals assumed to be positive, α_k's are the amplitudes

of the signals which are unknown complex valued parameters assumed to be different from zero, and $j = \sqrt{-1}$. The frequencies of the signals ω_k's are distinct and assumed to belong in the interval $[0, 2\pi]$. The error random variables, e_t's are independent and identically distributed (i.i.d.) random variables with mean zero and finite variance. We assume that the number of signals, M, is known.

Given a sample of size n, the problem is to estimate the unknown parameters α's, ω's, δ's and in certain situations, the variance of the error term. Note that for model (1) the parameters α_k's appear as linear parameters while the ω_k's and δ_k's are non-linear parameters. It is evident that the estimation of the non-linear parameters is a much more challenging problem. If we can obtain efficient estimates of the non-linear parameters, the estimation of the linear parameters reduces to the

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

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

standard least-squares problem. This ability to separate the linear and non-linear parameters was introduced by Richards [14] and termed separable regression. In this paper, we will address the problem of efficient estimation of the non-linear parameters: the frequencies and damping factors of the signals.

The multiple sinusoid model has been widely discussed and applied in the literature of Statistical Signal Processing. For example, in electromagnetic pulse (EMP) situations [15,16], the EMP pickup can be characterized by a sum of complex exponentials. The parameters provide a means of coding the various pulse wave forms and the signal approximation thus obtained can be readily employed to analyze responses in various subsystems under the EMP environment. In system identification problems, the characterization of the impulse responses of a linear system by a sum of complex exponentials and then identifying or approximating the complex amplitudes and natural frequencies with high degree of accuracy is of special importance.

There are several references in the literature dealing with the estimation of the parameters for this model. The review articles of Rao [13], Prasad et al. [12] and also the paper by Stoica [17] provide an extensive list of references. The estimation problem is well known to be numerically difficult [2,11], and it has been observed [7] that general purpose algorithms such as the Gauss–Newton, Newton–Raphson or their variants take a long time to converge especially with poor initial values. There has been considerable research [8] in the past 20 years to provide iterative algorithms for obtaining the maximum likelihood estimators. However, all these methods are computationally complex and suffer from dependence on the initial values, and convergence to possibly undesirable local extrema.

In this paper, we propose a non-iterative algorithm for estimation of the frequencies. It is observed that the proposed estimators are not exactly the least-squares estimators, but the mean squared errors (MSEs) of these estimators are quite close to those of the least-squares estimators (LSEs). The basic idea of the proposed method can be traced to Prony. We modify Prony's method using the notion of forward–backward linear prediction and using the singular value decomposition.

Traditionally, in estimating the parameters of the damped exponential model, either the original data (forward) or the conjugate (backward) of the data is used. It has been observed that for the undamped exponential model, if the original data and their conjugates are combined, they produce better estimates in terms of lower mean squared errors. Tufts and Kumaresan observed this phenomenon in a series of papers, (see, for example, [18] and references therein). Kannan [6] observed that the coefficients in Prony's method, when the damping factors are absent, satisfy a conjugate symmetric property. This symmetry property was exploited to obtain more efficient estimates in the undamped exponential model in [6,7]. In this paper we observe that even in the case of the damped exponential model, the original and conjugate data can be utilized to establish the conjugate symmetry property.

We also derive the asymptotic properties of the least-squares estimators and the Cramer–Rao bound of the unknown parameters. It can be established from Wu [19] that any estimators of the unknown parameters of model (1) are inconsistent (see [10]) in the usual sense. We redefine the model using the technique of Kahn et al. [5] and Kundu [10], and obtain the consistency and asymptotic distribution of the least-squares estimators of the redefined model. It is observed that the asymptotic variance of the least-squares estimators coincides with the Cramer–Rao bound.

The rest of the paper is organized as follows. In Section 2, we briefly describe Prony's algorithm and discuss how the original and conjugate data can be used in estimating the parameters of the damped exponential model. We apply the idea of extended order modeling as developed by Tufts and Kumaresan [18] to improve the performance of the modified estimator. Asymptotic properties of the least-squares estimators and the Cramer–Rao bound are discussed in Section 4. The results of a simulation study are provided in Section 5, and final conclusions appear in Section 6.

2. Prony's algorithm and its modifications

Consider the model given by (1). If there is no noise in the data, i.e., $e_t = 0$ for $t = 1, \ldots, n$, then

it can be easily observed that there exist $(M + 1)$ complex valued constants $g_1, \ldots, g_{M+1}$, such that

$$g_1 y_1 + \cdots + g_{M+1} y_{M+1} = 0$$

$$\vdots \qquad\qquad (2)$$

$$g_1 y_{n-M} + \cdots + g_{M+1} y_n = 0.$$

The constants $g_1, \ldots, g_{M+1}$ are the coefficients of the following polynomial equation:

$$g_1 + g_2 z + \cdots + g_{M+1} z^M = 0, \qquad (3)$$

whose roots are $e^{-\delta_1 + j\omega_1}, \ldots, e^{-\delta_M + j\omega_M}$. Here $g_1, \ldots, g_{M+1}$ are unique up to a constant multiple. From (2), it is clear that the conjugate data also satisfies a system of linear equations,

$$\bar{g}_1 \bar{y}_1 + \cdots + \bar{g}_{M+1} \bar{y}_{M+1} = 0$$

$$\vdots \qquad\qquad (4)$$

$$\bar{g}_1 \bar{y}_{n-M} + \cdots + \bar{g}_{M+1} \bar{y}_n = 0.$$

It can also be seen that the polynomial equation

$$\bar{g}_{M+1} + \bar{g}_M z + \cdots + \bar{g}_1 z^M = 0 \qquad (5)$$

has roots $e^{\delta_1 - j\omega_1}, \ldots, e^{\delta_M - j\omega_M}$.

Tufts and Kumaresan [18] used only the conjugate of the data to obtain estimates of the damping factors and the frequencies. We can combine the systems of equations given by (2) and (4) into the forward–backward prediction equations given by

$$\begin{bmatrix} y_1 & \cdots & y_{M+1} & \bar{y}_n & \cdots & \bar{y}_{n-M} \\ & \vdots & & & \vdots & \\ y_{n-M} & \cdots & y_n & \bar{y}_{M+1} & \cdots & \bar{y}_1 \end{bmatrix} \begin{bmatrix} g_1 \\ \vdots \\ g_{M+1} \\ \bar{g}_{M+1} \\ \vdots \\ \bar{g}_1 \end{bmatrix} = \mathbf{0},$$

$$(6)$$

which may be written in a compact form as

$$\begin{bmatrix} \mathbf{Y} : \mathbf{J}\bar{\mathbf{Y}}\mathbf{J} \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ \mathbf{J}\bar{\mathbf{g}} \end{bmatrix} = \mathbf{0}. \qquad (7)$$

Here $\mathbf{Y}$ is an $N - M \times M + 1$ data matrix, '⁻' denotes the conjugate of each element of a vector or

a matrix, and $\mathbf{J}$ is the $M + 1 \times M + 1$ exchange matrix defined as

$$\begin{bmatrix} 0 & \ldots & 1 \\ & \vdots & \\ 1 & \ldots & 0 \end{bmatrix}.$$

It can be easily seen that the rank of $(\mathbf{Y} : \mathbf{J}\bar{\mathbf{Y}}\mathbf{J})$ is $2M$. Consider the following matrix:

$$\begin{bmatrix} \mathbf{Y}^H \\ \mathbf{J}\bar{\mathbf{Y}}^H \mathbf{J} \end{bmatrix} \begin{bmatrix} \mathbf{Y} & \mathbf{J}\bar{\mathbf{Y}}\mathbf{J} \end{bmatrix} = \mathbf{B}^H \mathbf{B} \quad (say), \qquad (8)$$

which is clearly also of rank $2M$. Here H denotes the complex conjugate transpose of a matrix. The null space of this matrix can be generated by the following vectors;

$$\begin{bmatrix} \mathbf{g} \\ \mathbf{0} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{0} \\ \mathbf{J}\bar{\mathbf{g}} \end{bmatrix}. \qquad (9)$$

This representation of the noise space of the matrix in terms of these two vectors is exploited to yield estimates of the parameters of the damped exponential model even in the presence of noise. From the observed data, we construct the matrix $\mathbf{B}^H \mathbf{B}$ and obtain the eigenvector $\mathbf{u}$, corresponding to its smallest eigen value. Since $\mathbf{B}^H \mathbf{B}$ is a centro-symmetric matrix, we know that its eigenvectors must be either conjugate symmetric ($\mathbf{u} = \mathbf{J}\bar{\mathbf{u}}$) or anti-conjugate symmetric ($\mathbf{u} = -\mathbf{J}\bar{\mathbf{u}}$), (see [1]). Because of the centro-symmetric structure of $\mathbf{B}^H \mathbf{B}$ and the representation of its null space, it is always possible to choose the vector $\mathbf{u}$, such that $\hat{g}_1 > 0$, and $\mathbf{u}$ is of the form:

$$\mathbf{u} = [\hat{g}_1, \ldots, \hat{g}_{M+1}, \hat{\bar{g}}_{M+1}, \ldots, \hat{\bar{g}}_1]^T. \qquad (10)$$

Using $\mathbf{u}$, we solve the polynomial equation

$$\hat{\bar{g}}_{M+1} + \hat{\bar{g}}_M z + \cdots + \hat{\bar{g}}_1 z^M = 0 \qquad (11)$$

to obtain roots of the form

$$e^{\hat{\delta}_1 - j\hat{\omega}_1}, \ldots, e^{\hat{\delta}_M - j\hat{\omega}_M},$$

where $\hat{\delta}_1, \ldots, \hat{\delta}_M$ are estimators of the damping factors and $\hat{\omega}_1, \ldots, \hat{\omega}_M$ are estimators of the frequencies.

3. Extended order modeling

In this section, we use the idea of extended order modeling proposed by Tufts and Kumaresan [18] to obtain more efficient estimators of the parameters. Again, assuming that there is no noise in the data, we can show that there exists $b_1, \ldots, b_L$, such that

$$
\begin{bmatrix}
y_1 & \cdots & y_{L+1} & \bar{y}_n & \cdots & \bar{y}_{n-L} \\
& \vdots & & & \vdots & \\
y_{n-L} & \cdots & y_n & \bar{y}_{L+1} & \cdots & \bar{y}_1
\end{bmatrix}
\begin{bmatrix}
1 \\ b_1 \\ \vdots \\ b_L \\ \bar{b}_L \\ \vdots \\ \bar{b}_1 \\ 1
\end{bmatrix} = \mathbf{0}. \quad (12)
$$

It can be easily shown [18] that the polynomial equation

$$
z^L + \bar{b}_1 z^{L-1} + \cdots + \bar{b}_L = 0 \quad (13)
$$

has roots given by

$$
e^{\delta_1 - j\omega_1}, \ldots, e^{\delta_M - j\omega_M},
$$

whenever $M \leqslant L \leqslant N - M$. These zeros are called the signal zeros and have magnitudes all greater than 1 since $\delta_1 > 0, \ldots, \delta_M > 0$. If $L > M$, clearly the equation has $L - M$ additional zeros, usually called extraneous zeros. For $L > M$, (12) has more than one solution. We can use the unique minimum norm solution to (12) i.e., the solution that minimizes $||\mathbf{b}||^2$, where $\mathbf{b} = (\bar{b}_1, \ldots, \bar{b}_L)$. For the minimum norm solution, Kumaresan [9] has shown that the $L - M$ extraneous zeros have magnitudes all less than one, and thus provides a way of separating the signal and extraneous zeros. This idea can be used even when there is noise in the data as follows: Note that (12) can be written in the following form:

$$
\begin{bmatrix}
y_2 & \cdots & y_{L+1} & \bar{y}_n & \cdots & \bar{y}_{n-L+1} \\
& \vdots & & & \vdots & \\
y_{n-L+1} & \cdots & y_n & \bar{y}_{L+1} & \cdots & \bar{y}_2
\end{bmatrix}
\begin{bmatrix}
b_1 \\ \vdots \\ b_L \\ \bar{b}_L \\ \vdots \\ \bar{b}_1
\end{bmatrix}
$$

$$
= \begin{bmatrix}
y_1 + \bar{y}_{n-L} \\ \vdots \\ y_{n-L} + \bar{y}_1
\end{bmatrix} \quad (14)
$$

or

$$
\mathbf{Ab} = -\mathbf{Y} \quad (say).
$$

We need to solve (14) to obtain an estimate of $\mathbf{b}$. Using the ideas of Tufts and Kumaresan [18], we propose an estimate of $\mathbf{b}$ using the singular value decomposition as follows:

$$
\hat{\mathbf{b}} = \sum_{k=1}^{2M} \sigma_k^{-1} [\mathbf{u}_k^{\mathrm{H}} \mathbf{Y}] \mathbf{v}_k, \quad (15)
$$

where $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_{2L}$ are the singular values of $\mathbf{A}$, $\mathbf{u}_k$ and $\mathbf{v}_k$ are the eigenvectors of $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ and $\mathbf{A}\mathbf{A}^{\mathrm{H}}$, respectively. Note that it is always possible to choose $\mathbf{u}_k$ and $\mathbf{v}_k$ such that they are conjugate symmetric. Since $\mathbf{Y}$ is also conjugate symmetric, this guarantees the conjugate symmetric nature of $\hat{\mathbf{b}}$. Therefore it is possible to write

$$
\hat{\mathbf{b}}^{\mathrm{T}} = [\hat{b}_1 \ldots, \hat{b}_L, \hat{\bar{b}}_L, \ldots, \hat{\bar{b}}_1].
$$

We use $\hat{\bar{b}}_1, \ldots, \hat{\bar{b}}_L$ in Eq. (13) to obtain estimates of $(\delta_1, \omega_1), \ldots, (\delta_M, \omega_M)$.

4. Least-squares estimators and Cramer–Rao bound

In this section we redefine the model (1) in the manner of Kahn et al. [5] or Kundu [10] to allow us to establish the standard consistency property of the least-squares estimators for large samples. The new model will coincide with model (1) for finite

samples. The reason for this redefinition is because any estimator of the parameters of this model are inconsistent in the usual sense, i.e. as n tends to infinity.

We redefine the model (1) as follows:

$$y_{ni} = \sum_{k=1}^{M} \alpha_k e^{(-\delta_k + j\omega_k)t_{ni}} + e_{ni}, \qquad (16)$$

where α_k's, δ_k's and ω_k's have the same definitions as before. Here $t_{ni} = a + (b-a)/(n-1)(i-1)$ where, $t_{n1} = a$, the initial sampling point and $t_{nn} = b$, the extreme sampling point. The $\{e_{ni}\}$ for $i = 1, \ldots, n$ and $n = 1, 2 \ldots$ is a double array of random variables, with $\{e_{ni}\}$, $i = 1, \ldots, n$ assumed to be i.i.d. with mean zero. It is important to observe that the models (1) and (16) are equivalent for any finite n, although asymptotically they are different.

Suppose α_{Rk} and α_{Ik} are the real and imaginary parts of α_k and define $\theta_{4(k-1)+1} = \alpha_{Rk}$, $\theta_{4(k-1)+2} = \alpha_{Ik}$, $\theta_{4(k-1)+3} = \delta_k$ and $\theta_{4k} = \omega_k$. Let $\theta = (\theta_1, \ldots, \theta_{4M})$ be the vector of unknown parameters and

$$\mu(\theta, t) = \sum_{k=1}^{M} \alpha_k e^{(-\delta_k + j\omega_k)t} \qquad (17)$$

represent the constant term in the model. In order to establish the large sample properties of the least-squares estimators, we need to make the following assumption.

Assumption 1. $\{e_{ni}\}$ for $i = 1, \ldots, n$ and $n = 1, 2, \ldots$ is a double array of complex valued random variables with mean zero and finite variance $\sigma^2/2$ for both the real part and the imaginary part which are assumed to be independent. For each n, $e_{n1}, \ldots, e_{nn}$ are i.i.d. random variables. We assume $\theta \in \Theta$ where the parameter space Θ is a compact subset of $\mathbf{R}^{4M}$, and the true parameter value θ_0 is an interior point of Θ.

Note that the function

$$R(\theta) = \int_a^b |\mu(\theta, t) - \mu(\theta_0, t)|^2 \, dt$$

has a unique minimum at $\theta = \theta_0$ if at least two of the δ_k's or two of the ω_k's are different.

Theorem 1. *Under Assumption* 1, *the LSEs,* $\hat{\theta}$ *of* θ *in the model* (16) *are strongly consistent and*

$$\sqrt{n}(\hat{\theta} - \theta_0) \to N(\mathbf{0}, \sigma^2(2\mathbf{A})^{-1}).$$

Here ' $\to$ *' means converges in distribution and* $\mathbf{A} = ((a_{ij}))$ *with*

$$a_{ij} = \frac{1}{b-a} \int_a^b \mu_i'(\theta, t) \mu_j'(\theta, t) \, dt \quad and$$

$$\mu_i'(\theta, t) = \frac{\partial}{\partial \theta_i} \mu(\theta, t).$$

Proof. Consider

$$R(\theta) = \frac{1}{n} \sum_{i=1}^{n} |y_{ni} - \mu(\theta, t_{ni})|^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} |\mu(\theta_0, t_{ni}) - \mu(\theta, t_{ni})|^2 + \frac{1}{n} \sum_{i=1}^{n} |e_{ni}|^2$$

$$+ \frac{2}{n} Re \left(\sum_{i=1}^{n} [e_{ni}(\mu(\theta, t_{ni}) - \mu(\theta_0, t_{ni})] \right).$$

It can be shown that

$$\frac{1}{n} \sum_{i=1}^{n} |\mu(\theta_0, t_{ni}) - \mu(\theta, t_{ni})|^2$$

$$\to \int_a^b |\mu(\theta_0, t) - \mu(\theta, t)|^2 \, dt$$

uniformly. The second term on the right-hand side of $R(\theta)$ converges to σ^2 almost surely and the third term converges to zero almost surely. Therefore, using the same kind of arguments as in [4] or [10] the consistency of the least-squares estimator follows.

To prove the asymptotic normality of $\hat{\theta}$, let us write

$$Q(\theta) = nR(\theta) = \sum_{i=1}^{n} |y_{ni} - \mu(\theta, t_{ni})|^2 \qquad (18)$$

and

$$Q'(\theta) = \left(\frac{\partial}{\partial \theta_1} Q(\theta), \ldots, \frac{\partial}{\partial \theta_{4M}} Q(\theta) \right)$$

$$= (Q_1'(\theta), \ldots, Q_{4M}'(\theta)). \qquad (19)$$

Then,

$$\sqrt{n}(Q'(\hat{\boldsymbol{\theta}}) - Q'(\boldsymbol{\theta_0})) = \sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0})Q''(\bar{\boldsymbol{\theta}}), \quad (20)$$

where $Q''(\bar{\boldsymbol{\theta}})$ is a $4M \times 4M$ matrix, such that

$$Q''(\boldsymbol{\theta}) = ((Q''_{ij}(\boldsymbol{\theta}))) = \left(\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}Q(\boldsymbol{\theta})\right)\right) \quad (21)$$

and $\bar{\boldsymbol{\theta}}$ is a point between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta_0}$. Since $Q'(\hat{\boldsymbol{\theta}}) = 0$, (20) can be written as

$$-\sqrt{n}(Q'(\boldsymbol{\theta_0})) = \sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0})Q''(\bar{\boldsymbol{\theta}})$$

or

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta_0}) = -\sqrt{n}(Q'(\boldsymbol{\theta_0}))[Q''(\bar{\boldsymbol{\theta}})]^{-1}$$

$$= -\frac{1}{\sqrt{n}}Q'(\boldsymbol{\theta_0})\left[\frac{1}{n}Q''(\bar{\boldsymbol{\theta}})\right]^{-1}. \quad (22)$$

Observe that

$$\lim_{n\to\infty}\frac{1}{n}Q''(\bar{\boldsymbol{\theta}}) = \lim_{n\to\infty}\frac{1}{n}Q''(\boldsymbol{\theta^0}) \to 2\mathbf{A} \quad (23)$$

and

$$\frac{1}{\sqrt{n}}Q'(\boldsymbol{\theta_0}) \to N(0, 4\sigma^2\mathbf{A}). \quad (24)$$

From (23) and (24), the result follows immediately. Note that $\sigma^2(2\mathbf{A})^{-1}$ is the Cramer–Rao lower bound and the Cramer–Rao lower bound obtained by Kumaresan [9] is the discrete version of it. □

5. Numerical experiments

In this section we present some simulation results to see how the proposed methods behave for different signal-to-noise ratios and for different choices of L. We consider the following model:

$$y_t = 2.0\mathrm{e}^{(-.02+\mathrm{i}2.0)t} + 2.0\mathrm{e}^{(-.01+\mathrm{i}1.5)t} + e_t,$$

$$t = 1,\dots,25. \quad (25)$$

Here e_t's are i.i.d. complex valued normal random variables with mean zero and variance $\frac{1}{2}\sigma^2$ for both the real and imaginary parts. The real and imaginary parts are taken to be independent. We consider different signal to noise ratios (SNR $= 10\log_{10}1/\sigma^2$) varying from 10 to 25 dB. For each SNR, we generate random samples of

size 25 from model (25) and compute the Prony estimators (PEs) as discussed in Section 2. We also compute the extended ordered estimators (EOEs) as discussed in Section 3 for different choices of L, viz. $L = 6, 8, 10, 12, 14, 16, 18, 20, 21$. Finally, we compute the least-squares estimators using the proposed estimators as initial guesses, the backward linear prediction estimators of Tufts and Kumaresan [18] and also the matrix pencil method of Hua and Sarkar [3]. For the backward linear prediction estimators, we choose $L = 3*n/4 \approx 18$ and for the matrix pencil method we take $L = 2*n/3 \approx 17$ as they provide the best results in the respective cases (see [3,18]). The number of simulations is 1000, and we compute the average value of all the estimates and the average mean squared errors (MSEs). Since the results are quite similar for both the δ_i's and ω_i's, we report the results only for δ_1 and ω_1. The results for the PEs and the EOEs for different values of L, and for different SNR are reported in Table 1 and the results for the LSEs, the backward linear prediction estimators (BLPEs) and the matrix pencil methods (MPEs) are reported in Table 2. For comparison purposes, in Table 2 we report the Cramer–Rao lower bound (CRLB) and also the results of the EOEs for the best choice of L. The first figure in each box represents the average estimate of δ_1 and the corresponding MSE is represented within bracket. Similarly, the second figure represents the average estimate of ω_1 and the corresponding MSE is represented within bracket.

From Tables 1 and 2, for all the methods the MSEs decrease as the SNR increases (σ^2 decreases). This illustrates the consistency (an alternative definition of consistency proposed by Kahn et al. [5]) of all the estimators as σ^2 goes to zero. From Table 1, we observe that the PEs defined in Section 2 do not work very well at least for low signal-to-noise ratio. Similarly as L increases, the performances of the EOEs become better in terms of the biases and the mean squared errors. The MSEs and biases are lowest for $L = 20$ and then they start increasing. Therefore, it seems for $L = 4n/5$, the EOEs provide the best results. Now comparing the results in Table 2, it is clear that in most cases, the proposed EOEs for $L = 20$ behave better than the best possible backward linear prediction estimators or the best possible matrix pencil

Table 1
The PEs and the EOEs of δ_1 and ω_1 for different SNRs are reported[a]

	SNR = 10	SNR = 15	SNR = 20	SNR = 25
PE	0.0233(9.999E-4)	0.0282(9.815E-4)	0.0254(9.237E-4)	0.0256(9.127E-4)
	2.0570(1.080E-3)	2.0572(8.412E-4)	2.0164(7.833E-4)	2.0060(1.193E-4)
$L = 6$	0.0251(9.981E-4)	0.0257(9.797E-4)	0.0278(9.156E-4)	0.0160(9.052E-4)
	1.9561(8.901E-4)	1.9620(7.397E-4)	1.9623(6.230E-4)	1.9552(7.093E-5)
$L = 8$	0.0249(9.554E-4)	1.9716(9.333E-4)	0.0266(8.999E-4)	0.0117(3.885E-5)
	1.9516(6.333E-4)	2.0331(5.390E-4)	2.0255(4.705E-4)	1.9579(2.963E-5)
$L = 10$	0.0254(8.385E-4)	0.0296(7.330E-4)	0.0284(5.561E-4)	0.0259(8.374E-5)
	2.0718(2.439E-4)	2.0639(2.413E-4)	2.0546(1.415E-4)	2.0422(3.249E-5)
$L = 12$	0.0239(5.609E-4)	0.0222(5.498E-4)	0.0285(4.414E-4)	0.0218(7.928E-5)
	1.9387(2.231E-4)	1.9204(2.200E-4)	2.0441(9.809E-5)	1.9914(2.505E-5)
$L = 14$	0.0189(2.881E-4)	0.0192(2.136E-4)	0.0182(2.036E-4)	0.0186(2.012E-5)
	1.9584(2.101E-4)	1.8914(2.000E-4)	1.8806(7.567E-5)	1.8766(9.204E-6)
$L = 16$	0.0076(2.023E-4)	0.0087(1.869E-4)	0.0108(1.794E-4)	0.0155(1.058E-5)
	2.0792(1.063E-4)	2.1701(1.001E-4)	2.1284(9.592E-5)	2.0761(8.117E-6)
$L = 18$	0.0201(1.465E-4)	0.0163(2.172E-4)	0.0218(1.567E-4)	0.0205(8.136E-6)
	1.9729(9.280E-5)	1.9305(9.075E-5)	1.9774(8.273E-5)	2.1775(7.334E-6)
$L = 20$	0.0197(8.711E-5)	0.0194(6.345E-5)	0.0179(3.636E-5)	0.0183(1.667E-6)
	1.9785(5.690E-5)	1.9766(6.271E-5)	1.9739(7.876E-5)	1.9726(3.895E-6)
$L = 21$	0.0206(1.534E-4)	0.0169(1.985E-4)	0.0192(1.898E-4)	0.0201(7.098E-6)
	1.6142(9.119E-5)	1.9435(8.986E-5)	1.9609(9.174E-5)	1.6314(6.167E-6)

[a]The sample size is 25. In each box, the figure in the first line represents the average estimate of δ_1 and the figure within the bracket represents the corresponding MSE. Similarly, the results of ω_1 are presented in the second line.

Table 2
The BLPEs, the EOEs, the LSEs, the MPEs and the Cramer–Rao lower bounds of δ_1 and ω_1 for different SNRs are presented[a]

	SNR = 10	SNR = 15	SNR = 20	SNR = 25
CRLB	0.0200(9.096E-6)	0.0200(2.877E-6)	0.0200(9.099E-7)	0.0200(2.877E-7)
	2.0000(6.988E-6)	2.0000(2.209E-6)	2.0000(6.987E-7)	2.0000(2.209E-7)
LSE	0.0196(9.565E-6)	0.0199(2.928E-6)	0.0199(9.971E-7)	0.0200(2.889E-7)
	1.9895(5.036E-6)	2.0000(2.345E-6)	2.0000(7.414E-7)	2.0000(2.344E-7)
MPE	0.0193(2.216E-5)	0.0191(6.254E-5)	0.0186(1.740E-5)	0.0177(4.569E-6)
	1.9749(3.434E-3)	1.9879(1.119E-3)	2.0004(1.559E-5)	2.0006(6.015E-6)
BLPE	0.0182(1.112E-4)	0.0187(6.283E-5)	0.0195(1.756E-5)	0.0198(5.524E-6)
	1.9711(1.574E-2)	1.9984(1.340E-3)	2.0004(1.952E-5)	2.0001(6.142E-6)
EOE	0.0197(8.711E-5)	0.0194(6.345E-5)	0.0179(3.636E-5)	0.0183(1.667E-6)
($L = 20$)	1.9785(5.690E-5)	1.9766(6.271E-5)	1.9739(7.876E-5)	1.9726(3.895E-6)

[a]The sample size is 25. In each box, the figure in the first line represents the average estimate of δ_1 and the figure within the bracket represents the corresponding MSE. Similarly, the results of ω_1 are presented in the second line.

estimators. Although the gain is not that significant at high SNR, for low SNR the EOEs behave much better than the backward linear prediction estimators or the matrix pencil estimators. Matrix pencil estimators behave almost similarly as the backword linear prediction estimators in most of the cases. Sometimes it behaves slightly better than the backward linear prediction estimators. Moreover, the EOEs can be used quite effectively as initial guesses for computing the least-squares estimators and the MSEs of the least-squares estimators as expected are quite close to the Cramer–Rao lower bound as obtained in Section 4.

In terms of computational complexities of the different methods, the matrix pencil method involves the generalized eigenvalues computations of two matrices and the other three estimators, namely the PEs, backward linear prediction estimators and the EOEs, involve computing the singular value decomposition of a complex matrix. For a given n and M, the matrix pencil method involves computing the generalized eigenvalues of two $n/3 \times 2n/3$ matrices, the Prony estimators involve computing the singular value decomposition of a $2M \times 2M$ matrix, whereas backward linear prediction estimators require the singular value decomposition of a $3n/4 \times 3n/4$ matrix and the EOEs involve the computation of a $4n/5 \times 4n/5$ singular value decomposition of a matrix. Therefore for large n, the proposed estimators are computationally more intensive than the backward linear prediction estimators or the matrix pencil estimators, but for small n the difference is not that great.

6. Conclusions

In this paper we consider the estimation of the parameters of a damped exponential model in presence of additive errors. We propose a non-iterative method to estimate the non-linear parameters using both the original data and their conjugates. To the author's knowledge, this use of forward and backward data has never been attempted before for the damped exponential model. We observe that the proposed estimators work quite well and perform better than the backward linear prediction estimators or the matrix pencil estimators in most cases.

The proposed estimators can be used as effective initial guesses for the least-squares estimators. Numerically, it is observed that the performances of the proposed estimators are best for $L = 4n/5$, although at present it is not possible to afford any theoretical justifications for that choice. More work is needed in this direction to determine the optimal order of the model.

References

[1] A. Cantoni, P. Butler, Eigenvalues and eigenvectors of symmetric centrosymmetric matrices, Linear Algebra Appl. 13 (1976) 275–288.

[2] G.H. Golub, V. Pereyra, The differentiation of pseudoinverses and nonlinear least squares problems whose variable separates, SIAM J. Numer. Anal. 10 (2) (1973) 413–432.

[3] Y. Hua, T.K. Sarkar, Matrix pencil method for estimating parameters of exponentially damped/ undamped sinusoids in noise, IEEE Trans. Acoust. Speech Signal Process. 38 (1990) 814–824.

[4] R.I. Jennrich, Asymptotic properties of the non-linear least-squares estimators, Ann. Math. Statist. 40 (1969) 633–643.

[5] M. Kahn, M.S. Mackisack, M.R. Osborne, G. Smyth, On the consistency of Prony's method and related algorithms, J. Comput. Graphical Statist. 1 (4) (1992) 329–349.

[6] N. Kannan, Estimation of direction of arrival in signal processing, Ph.D. Dissertation, The Pennsylvania State University, USA, 1992.

[7] N. Kannan, D. Kundu, On modified EVLP and ML methods for estimating superimposed exponential signals, Signal Process. 39 (1994) 223–233.

[8] S. Kay, Modern Spectral Estimation: Theory and Applications, 2nd Edition, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[9] R. Kumaresan, Estimating the parameters of exponentially damped or undamped sinusoidal signals in noise, Ph.D. Dissertation, The University of Rhode Island, USA, 1982.

[10] D. Kundu, A modified Prony algorithm for sum of damped or undamped exponential signals, Sankhya Ser. A 56 (3) (1994) 525–544.

[11] M.R. Osborne, G. Smyth, A modified Prony algorithm for exponential function fitting, SIAM J. Comput. 16 (1) (1995) 119–138.

[12] S. Prasad, M. Chakraborty, H. Parthasarathy, The role of statistics in signal processing a brief review and some emerging trends, Indian J. Pure Appl. Math. 26 (6) (1995) 547–578.

[13] C.R. Rao, Some results in signal processing, in: S.S. Gupta, J.O. Berger (Eds.), Statistical Decision Theory and Related Topics, IV, Vol. 2, Springer, New York, 1988, pp. 319–332.

[14] F.S.G. Richards, A method of maximum likelihood estimation, J. Roy. Statist. Soc. Ser. B 23 (1961) 469–475.

[15] L.W. Ricketts, J.E. Bridges, J. Milletta, EMP Radiation and Protection Technique, Wiley, New York, 1976.

[16] P. Sircar, Accurate parameter estimation of damped sinusoidal signals samples at non-uniform spacings, Ph.D. Dissertation, Syracuse University, USA, 1987.

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

[18] D.W. Tufts, R. Kumaresan, Estimation of frequencies of multiple sinusoids; making linear prediction perform like maximum likelihood, Proc. IEEE 70 (1982) 975–989.

[19] C.F.J. Wu, Asymptotic theory of nonlinear least squares estimation, Ann. Statist. 9 (1981) 501–513.