Digital Signal Processing 48 (2016) 148–162

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

An adaptive and computationally efficient algorithm for parameters estimation of superimposed exponential signals with observations missing randomly

Jiawen Bian [a], Jing Xing [b,c], Jianfeng Liu [a,c], Zhiming Li [a,c], Hongwei Li [a,c,*]

[a] School of Mathematics and Physics, China University of Geosciences, Wuhan, 430074, China
[b] Institute of Statistics, Hubei University of Economics, Wuhan, 430205, China
[c] Hubei Subsurface Multi-scale Imaging Key Laboratory, China University of Geosciences, Wuhan, 430074, China

A R T I C L E   I N F O

Article history:
Available online 26 September 2015

Keywords:
Superimposed exponential signals
Least squares estimators
Missing observations
Multiplicative and additive noise
Asymptotically unbiased and consistent estimators

A B S T R A C T

In this paper, we consider the parameters estimation of a model of superimposed exponential signals in multiplicative and additive noise when some observations are missing randomly. The least squares estimators (LSEs) and asymptotic Cramer–Rao low bound (ACRLB) for the considered model are studied and the asymptotic distributions of the LSEs for parameters of frequencies, phases and amplitudes of the considered model are also derived and obtained. An adaptive and computationally efficient iterative algorithm is proposed to estimate the frequencies of the considered model. It can be seen that the iterative algorithm works quite well in terms of biases and mean squared errors and the refined estimators by three iterations are observed to be asymptotically unbiased and consistent. The statistics for iteration are designed to change adaptively according to different missing distributions of time points so as to keep the estimators of frequencies to be asymptotically unbiased. Moreover, the proposed estimators attain the same convergence rate and asymptotic distribution as those of LSEs which are used to obtain the confident intervals and coverage probabilities of the frequencies for finite sample. Since the iterative algorithm needs only three iterations to work, it saves much computation time. So the proposed estimators are LSEs equivalent while avoid the heavy computation cost of LSEs. Finally, several simulation experiments are performed to verify the effectiveness of the proposed algorithm. To examine the robustness of the proposed algorithm, we also test the algorithm on the dual tone multi-frequency (DTMF) signal with observations missing in block and symmetric α-stable (SaS) noise condition, as well as on sinusoidal frequency modulated signals.

1. Introduction

We consider the following model of superimposed exponential signals in multiplicative and additive noise

$$y(t) = x(t)s(t),$$
$$x(t) = \sum_{k=1}^{l} \xi_k(t) e^{i(\omega_k t + \varphi_k)} + \varepsilon(t), \quad t = 1, 2, \cdots, N \tag{1}$$

where $i = \sqrt{-1}$, the ω_k's and φ_k's are unknown frequencies and phases lying strictly between 0 and 2π and they are distinct. Multiplicative noise $\{\xi_k(t)\}$ is a sequence of independent identically distributed (i.i.d.) real random variables with mean $\mu_k \neq 0$ and finite variance $\sigma_k^2 (k = 1, 2, \cdots, l)$. Additive noise $\{\varepsilon(t)\}$ is a sequence of i.i.d. complex random variables with zero mean and finite variance $\sigma_0^2/2$ for both the real and imaginary parts which are assumed to be independent. The multiplicative and additive noise are mutually independent. $\{s(t)\}$ is a sequence of i.i.d. random variable with 0–1 distribution where $P\{s(t) = 0\} = p$ denotes the missing probability at each time point. The number of the signal 'l', the missing rate 'p' and the missing time point for each missing observation are assumed to be known in advance. It is known that the frequencies are nonlinear and also the most important parameters of the considered model. The other parameters can be estimated by a linear regression process once the frequencies are estimated correctly [1]. In this paper, we will first study the LSEs of the parameters of frequencies, amplitudes and phases for the considered model (1). Then we will mainly focus on the

* Corresponding author at: School of Mathematics and Physics, China University of Geosciences, Wuhan, 430074, China.
E-mail address: hwli@cug.edu.cn (H. Li).

http://dx.doi.org/10.1016/j.dsp.2015.09.014
1051-2004/

estimation of frequencies ω_k's, given a sample of size N, namely $\{y(t)\}_{t\in I_N}$ where I_N is the set of observations available.

This is an important problem in statistical signal processing and time series analysis. In the last twenty years a lot of iterative and non-iterative procedures were devised for uniformly sampled complete data sequences [1–3]. However, in many practical applications, the measured data may be incomplete due to, for example, sensor failures, outliers and the data compression needs, etc. In astronomical, meteorological, or satellite based applications, weather or other conditions may disturb sample taking schemes (e.g., measurements are only available during nighttime for astronomical applications), which will result in missing or gapped data [4]. Missing data problems are also encountered in modern radar systems which have multiple duties including searching, tracking and the automatic classification of targets. Switching in and out of these modes leads to incomplete phase history data for synthetic aperture radar (SAR) imaging [5]. For foliage penetrating radar systems, certain radar operating frequency bands are under strong electromagnetic or radio frequency (RF) interference so that the corresponding observations must be discarded resulting in missing data [6].

Several methods are devised to obtain the spectra estimation of the observed signals, which can also be used to estimate the frequencies for the periodic signals. These methods can be classified into two categories. The first one is based on interpolation for the missing observations and the second one does not rely on interpolation. Quite a few techniques are proposed without interpolation. The Lomb periodogram is developed for irregularly sampled (unevenly spaced) data as well as data with missing observations [7,8]. The CLEAN algorithm [9] is used to estimate the spectrum by deconvolving the missing data discrete Fourier transform (DFT) spectrum into the true signal spectrum and the Fourier transform of the windowing function via an iterative approach. The multi-taper methods [10] compute spectral estimates by assuming certain quadratic functions of the available data. The coefficients in the corresponding quadratic functions are optimized according to certain criteria. These methods have nearly the same resolution as the DFT. To achieve high resolution, several parametric algorithms, e.g., those based on autoregressive (AR) or autoregressive moving-average (ARMA) models, were used to handle the missing data problem [11–13]. However, these techniques are all based on the noise free model and are sensitive to the noise. For the non-interpolation methods, the missing observations are reconstructed and the spectra density is estimated from the reconstructed uninterrupted signal. A simple reconstruction uses linear, cubic, or spline interpolation of the missing data points. The best reconstruction method, however, was derived from the EM algorithm for missing data [14,15] where the missing data are reconstructed with a dynamic model. The parameters of the model are updated from the reconstructed uninterrupted signal, and the reconstruction is repeated with the new parameters until convergence. Several nonparametric and adaptive filtering based techniques have also been developed for the missing data problem. In [4] and [16], an extension of the amplitude and phase estimation (APES) method was used to the case of gapped data which iteratively interpolates the missing data and estimates the spectrum. In [17], the APES and Capon spectral estimators were applied to periodically gapped data to obtain the PG-APES and PG-CAPON algorithms. However, these adaptive filtering based methods can only deal with missing data occurring in gaps and do not work well for the more general problem of missing data occurring randomly. Although the methods based on spectral estimation can also be used to estimate the frequencies of the model, they are not devised specifically for parameters estimation and some of them are even biased such as periodogram estimation and multitaper spectrum estimation [18]. Moreover, since the parameters estimation by spectral estimation are obtained by searching the peak of the spectra, the precision and convergence rate of the algorithm are limited. Few works are focused on the parameters estimation of superimposed exponential signals or harmonic signals with missing observations, especially when the observations are missing randomly. An information theoretic criterion and eigenvariation technique were used in [19] to estimate the number and the frequency of the signal simultaneously and the strong convergence of the estimator was obtained. [20] used linear prediction and weighted least squares techniques to estimate the frequencies of the signals with gapped data. However, the linear prediction techniques used in [19] and [20] are both based on noise free model and are sensitive to noise and the methods proposed in [20] can only be used for gapped data and cannot be used for the condition of randomly missing observations.

Recently, an iterative algorithm was proposed in [21] to estimate the frequencies of superimposed exponential signals in additive noise. [22] and [23] generalized the iterative procedure for zero-mean multiplicative noise condition and two-dimensional condition respectively. The iterative algorithm proposed in [21–23] can converge within fixed number of iterations and the corresponding estimators attain the same convergence rate as that of LSEs, thus is LSEs equivalent and computationally efficient. But nowhere, at least not known to the authors, are the LSEs and the iterative procedure for the LSEs of the frequencies for a superimposed exponential model with multiplicative and additive noise and observations missing randomly considered. It will be very appealing if the iterative procedure can be used for the missing observations condition and is LSEs equivalent at the same time. However, the iterative statistics and the corresponding iterative coefficients in [21–23] are based on continuous time series and cannot be used for missing observations condition directly. Moreover, the LSEs for the randomly missing observations condition of model (1) cannot be established directly on the LSEs for complete data of multiplicative and additive noise condition [24]. In this paper, firstly, we study the LSEs as well as the corresponding asymptotic distribution of frequencies, phases and amplitudes for the considered model (1), then an adaptively iterative procedure is proposed to estimate the frequencies of model (1) under the inspiration of the works of [21–23]. The coefficient in the iterative term is devised to change according to the variation of the time point distribution of the missing observations. It can be proved that the precision and the convergence rate of the estimators will be improved after each iteration and the estimators can attain the same convergence rate as that of LSEs after only three iterations while have the same asymptotic distribution as that of LSEs under the same condition. The iterative procedure uses a correction term based on $A_N(j)$ and $B_N(j)$ to be defined in Section 4, which are functions of the data vector and the j-th available frequency estimator. It is observed that if the initial estimator is accurate up to the order $O_p(N^{-1})$ (here $O_p(N^{-\delta})$ means $O_p(N^{-\delta})N^{\delta}$ is bounded in probability[1]), then the iterative procedure produces fully efficient frequency estimator which has convergence rate of $O_p(N^{-3/2})$. The Lomb periodogram maximizers over Fourier frequencies are taken as the initial estimators of frequencies for the iterative algorithm. It is known that the Lomb periodogram maximizers over Fourier frequencies do not generally provide estimators up to the order $O_p(N^{-1})$ [25]. To overcome this problem, we use the varying sample size technique as in [21], i.e., we do not use the fixed sample size available for estimation at each step. At the first step, we use a fraction of it and at the last step we use the whole data set by gradually increasing the effective sample size.

[1] Here bounded in probability means that $\exists$ constant $C > 0$, s.t. $\lim_{N\to\infty} P\{|O_p(N^{-\delta})N^{\delta}| < C\} = 1$.