Frequency Estimation of Irregularly Sampled Data

Using a Sparsity Constrained Weighted Least-Squares Approach

Adel Zahedi
School of Electrical Engineering
Iran University of Science and Technology
Tehran, Iran
azahedi@elec.iust.ac.ir

Mohammad Hossein Kahaei
School of Electrical Engineering
Iran University of Science and Technology
Tehran, Iran
kahaei@iust.ac.ir

Abstract—**In this paper, a new method for frequency estimation of irregularly sampled data is proposed. In comparison with the previous sparsity-based methods where the sparsity constraint is applied to a least-squares fitting problem, the proposed method is based on a sparsity constrained weighted least-squares problem. The resulting problem is solved in an iterative manner, allowing the usage of the solution obtained at each iteration to determine the weights of the least-squares fitting term at the next iteration. Such an appropriate weighting of the least-squares fitting term enhances the performance of the proposed method. Simulation results verify that the proposed method can detect the spectral peaks using a very short data record. Compared to the previous one, the proposed method is less probable to miss the actual spectral peaks and exhibit spurious peaks.**

Keywords-basis pursuit; sparse representation; overcomplete dictionaries; irregular sampling; spectrum estimation

I. INTRODUCTION

Spectral analysis of irregularly sampled data is ubiquitous in science and technology. The range of applications for this branch of signal processing includes, among others, the estimation of Doppler frequency in radars [1], astronomical data analysis [2], the analysis of the heart rate variability in medical sciences [3] and turbulent velocity analysis in measurement systems [4].

Let us denote by $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \end{bmatrix}^T$ the data-set taken at N sampling time instants $\mathbf{t} = \begin{bmatrix} t_1 & t_2 & \cdots & t_N \end{bmatrix}^T$. The frequency grid on which the spectrum is to be computed is denoted by

$$\mathbf{f} = \begin{bmatrix} f_{-M} & \cdots & f_{-1} & f_0 & f_1 & \cdots & f_M \end{bmatrix}^T \quad (1)$$

with

$$f_m = \frac{m}{M} f_{\max}, \ m = -M, \ldots, M \quad (2)$$

for a predefined maximum frequency $f_{\max}$.

Also let us define an overcomplete dictionary of complex exponentials which is a $N \times K$ matrix defined as:

$$\mathbf{W} = \frac{1}{\sqrt{N}} \exp\left(j2\pi \mathbf{t}\mathbf{f}^T \right) \quad (3)$$

where $K = 2M + 1$, $K > N$, and $\exp(\mathbf{A})$ for matrix $\mathbf{A}$ is defined as a matrix of the same size as $\mathbf{A}$ with elements computed as the exponential of the corresponding elements in $\mathbf{A}$. From (3) it can be inferred that the mth column of $\mathbf{W}$ is a complex exponential function of frequency f_m sampled at the sampling instants $\mathbf{t}$.

When the data set is known to be sinusoidal (which is the case for most applications), one of the most powerful techniques of spectrum estimation is to represent the data-set $\mathbf{y}$ as a linear combination of the columns of $\mathbf{W}$; i.e. to obtain a vector $\hat{\mathbf{s}}$ so that $\mathbf{y} = \mathbf{W}\hat{\mathbf{s}}$. If so, the nonzero elements of $\hat{\mathbf{s}}$ demonstrate the amplitude of the spectral components at corresponding frequencies in $\mathbf{f}$. Redundancy of $\mathbf{W}$ gives rise to infinitely many solutions $\hat{\mathbf{s}}$ which opens the way to enforcing some more constraints to the solution. The simplest of them are the least-squares fitting methods; e.g. the Lomb-Scargle periodogram [5]. More sophisticated methods penalize the least-squares fitting term with a sparsity-enforcing term [6-8].

In [2, 9-10], the sparsity constraint is formulated as the ℓ_1 norm of the solution. This method is widely known in the literature as Basis Pursuit [11]. These methods are capable to estimate the spectral lines from very limited data records and allow super-resolution frequency estimation [9-10]. In [9], the dictionary $\mathbf{W}$ is defined as a collection of cisoids instead of complex exponentials. It is shown in [2, 10] that sparse approximation in the complex exponential dictionary outperforms the method of [9]. In this paper, we zoom in on the method of [10] which is the most recent amongst the above methods.

The sparse approximation of $\mathbf{y}$ in $\mathbf{W}$ is the solution to the following optimization problem [12]:

$$\hat{\mathbf{s}} = \underset{\mathbf{s} \in \mathbb{C}^K}{\operatorname{argmin}} \left(\left\| \mathbf{y} - \mathbf{W}\mathbf{s} \right\|^2 + \lambda \left\| \mathbf{s} \right\|_1 \right) \quad (4)$$

where λ is called the regularization parameter and $\left\| \mathbf{s} \right\|_1$ is the norm of $\mathbf{s}$ defined as:

$$\|\mathbf{s}\|_1 = \sum_{i=1}^{K} |s_i| \qquad (5)$$

The cost function in (4) is the sum of two terms. The first one enforces the least squares fitting of the solution to the data set, and the second one guarantees the sparsity of the solution. This problem can be solved in polynomial time using a variety of methods [10]. It has been shown in [10] that if $\hat{\mathbf{s}}$ is a solution to (5) and $\|\hat{\mathbf{s}}\|_0 < N/2$ (where $\|\hat{\mathbf{s}}\|_0$ counts the number of nonzero elements in $\hat{\mathbf{s}}$), then it is the unique sparsest solution.

Roughly speaking, the minimizer of (4) is a sparse vector $\hat{\mathbf{s}} = \begin{bmatrix} \hat{s}_{-M} & \cdots & \hat{s}_{-1} & \hat{s}_0 & \hat{s}_1 & \cdots & \hat{s}_M \end{bmatrix}^T$ so that $\mathbf{W}\hat{\mathbf{s}}$ fits to the data-set in the least-squares sense. $|\hat{s}_m|$ is thus an estimation of the amplitude of the signal component at frequency f_m. Problem (4) is known in the literature as Basis Pursuit De-Noising (BPDN) [10, 11].

In this paper, instead of the sparsity constrained least-squares problem of (4), another cost function is proposed which can be interpreted as a sparsity constrained weighted least-squares problem, and is solved iteratively. In order to refine the fitting to the data-set, in each iteration the obtained temporary solution is utilized to weight the least-squares term in the next iteration. Since the proposed sparsity-based method solves a weighted least-squares problem in an iterative manner, it falls into the class of methods named Iterative Adaptive Approach (IAA) [13-15].

This paper is organized as follows: In Section II, the proposed method is described in detail. In Section III, the proposed method is compared to the BPDN method of [10] formulated in (4) using simulation results. The paper is concluded in Section IV.

II. PROPOSED METHOD

A. Problem Formulation

Problem (4) can be rewritten in the following component-wise form:

$$\hat{s}_m = \underset{s_m \in \mathbb{C}}{\arg\min} \left(\|\mathbf{y} - \mathbf{W}\mathbf{s}\|^2 + \lambda \|\mathbf{s}\|_1 \right), \quad m = -M, \ldots, M \quad (6)$$

The Iterative Coordinate Descent (ICD) method described in [10] can be used to solve this scalar problem considering the following equivalence:

$$\hat{s}_m = \underset{s_m \in \mathbb{C}}{\arg\min} \left(\|\mathbf{y} - \mathbf{W}\mathbf{s}\|^2 + \lambda \|\mathbf{s}\|_1 \right) \quad \Leftrightarrow \quad \hat{s}_m = \phi_\lambda \left(\mathbf{w}_m^H \mathbf{e}_m \right) \quad (7)$$

where $\mathbf{w}_m$ is the mth column of $\mathbf{W}$, $\mathbf{e}_m$ is defined as:

$$\mathbf{e}_m = \mathbf{y} - \mathbf{W}\mathbf{s} + s_m \mathbf{w}_m \qquad (8)$$

and for any complex number $x = |x| e^{j\theta}$ the complex soft shrinkage function $\phi_\lambda(x)$ is defined as [10]:

$$\phi_\lambda(x) = \begin{cases} \left(|x| - \lambda \right) e^{j\theta}, & \text{if } |x| > \lambda \\ 0, & \text{otherwise} \end{cases} \qquad (9)$$

It is necessary to mention again that the cost function in (6) is sum of two terms: a least-squares data-fitting term $\|\mathbf{y} - \mathbf{W}\mathbf{s}\|^2$ and a sparsity enforcing ℓ_1 norm term $\|\mathbf{s}\|_1$. The value of the regularization parameter λ determines the relative emphasis on each one of these two terms. A large λ should be interpreted as a tendency to obtain a very sparse solution, while with a small λ the sparsity of the solution is sacrificed in favor of its fidelity to the data-set. The regularization parameter therefore makes a trade-off between the two important parameters of sparsity and fidelity to the data-set. The level of this trade-off is improved by increase of the number of data samples N. When N is large enough, the solution to (6) can be both sparse and fitted to the data set. The question is that if N is limited, is it possible to somehow improve the trade-off level? One possible answer is to be "parsimonious" in spending the recourses on minimization of the cost function by "wise" fitting to the data-set; i.e., by using a weighted least-squares fitting to the data-set instead of the simple least-squares.

Let us temporarily assume that for a specific $m \left(-M \leq m \leq M \right)$, the solution $\hat{s}_r$ to the optimization problem is available for all $r \neq m$ $\left(-M \leq r \leq M \right)$. Inspired by [13], we define the covariance matrix for the mth component of $\hat{\mathbf{s}}$ as:

$$\mathbf{Q}'_m = \sum_{r=-M, r \neq m}^{M} |\hat{s}_r|^2 \, \mathbf{a}_r \mathbf{a}_r^H \qquad (10)$$

where $\mathbf{a}_r$ is defined as:

$$\mathbf{a}_r = \exp \left(j 2\pi f_r \mathbf{t} \right) \qquad (11)$$

For any nonzero vector $\mathbf{u} \in \mathbb{C}^N$ we have:

$$\mathbf{u}^H \mathbf{Q}'_m \mathbf{u} = \sum_{r=-M, r \neq m}^{M} |\hat{s}_r|^2 \, \mathbf{u}^H \mathbf{a}_r \mathbf{a}_r^H \mathbf{u} = \sum_{r=-M, r \neq m}^{M} |\hat{s}_r|^2 \left| \mathbf{a}_r^H \mathbf{u} \right|^2 \geq 0.$$

Thus $\mathbf{Q}'_m$ is nonnegative definite. Notice that according to (10), $\mathbf{Q}'_m$ is the sum of $2M$ matrices $\mathbf{a}_r \mathbf{a}_r^H$ each of which of rank 1, and weighted by $|\hat{s}_r|^2$. Since the number of nonzero components of $\hat{\mathbf{s}}$ is $\|\hat{\mathbf{s}}\|_0$, $\mathbf{Q}'_m$ is the sum of $\|\hat{\mathbf{s}}\|_0$ or $\left(\|\hat{\mathbf{s}}\|_0 - 1 \right)$ rank 1 matrices (depending on the two different possibilities $\hat{s}_m = 0$ and $\hat{s}_m \neq 0$). The matrix thus has the following rank:

$$rnk \left(\mathbf{Q}'_m \right) = \begin{cases} N, & \text{if } s'_m \neq 0 \text{ and } \|\hat{\mathbf{s}}\|_0 - 1 \geq N \\ \|\hat{\mathbf{s}}\|_0 - 1, & \text{if } s'_m \neq 0 \text{ and } \|\hat{\mathbf{s}}\|_0 - 1 < N \\ N, & \text{if } s'_m = 0 \text{ and } \|\hat{\mathbf{s}}\|_0 \geq N \\ \|\hat{\mathbf{s}}\|_0, & \text{if } s'_m = 0 \text{ and } \|\hat{\mathbf{s}}\|_0 < N \end{cases} \qquad (12)$$

Therefore, it may or may not be of full-rank. Specially, since $\hat{\mathbf{s}}$ is expected to be sparse, the 2nd and 4th cases of the

above relation are likely to happen. The following is a revised definition of the covariance matrix which has full-rank:

$$\mathbf{Q}_m = \sum_{r=-M, r\neq m}^{M} \left| \hat{s}_r \right|^2 \mathbf{a}_r \mathbf{a}_r^H + \varepsilon \mathbf{I}_N = \mathbf{Q}'_m + \varepsilon \mathbf{I}_N \qquad (13)$$

where ε is a small constant and $\mathbf{I}_N$ is the $N \times N$ identity matrix. Since $\mathbf{Q}_m$ is nonnegative definite and full-rank, it is Positive Definite (PD) and of course invertible. The optimization problem (6) is then replaced by the following ℓ_1 regularized weighted least-squares problem:

$$\hat{s}_m = \underset{s_m \in \mathbb{C}}{\operatorname{argmin}} \left\{ (\mathbf{y} - \mathbf{W}\mathbf{s})^H \mathbf{Q}_m^{-1} (\mathbf{y} - \mathbf{W}\mathbf{s}) + \lambda \|\mathbf{s}\|_1 \right\},$$
$$m = -M, \dots, M \qquad (14)$$

The problem (14) obviously cannot be solved directly, since one has to know its solution to compute $\mathbf{Q}_m$ using (13).

B. How to Solve It?

As was shown in Part II.A, the matrix $\mathbf{Q}_m$ is PD, which means that $\mathbf{Q}_m^{-1}$ is also PD. Therefore, there is a unique Cholesky factor $\mathbf{R}_m$ for $\mathbf{Q}_m^{-1}$ so that [16]:

$$\mathbf{Q}_m^{-1} = \mathbf{R}_m^H \mathbf{R}_m \qquad (15)$$

Substituting (15) in (14) yields:

$$\hat{s}_m = \underset{s_m \in \mathbb{C}}{\operatorname{argmin}} \left(\left\| \mathbf{y}'_m - \mathbf{W}'_m \mathbf{s} \right\|^2 + \lambda \|\mathbf{s}\|_1 \right), \quad m = -M, \dots, M \qquad (16)$$

where $\mathbf{y}'_m$ and $\mathbf{W}'_m$ are defined as:

$$\mathbf{y}'_m = \mathbf{R}_m \mathbf{y} \qquad (17)$$
$$\mathbf{W}'_m = \mathbf{R}_m \mathbf{W} \qquad (18)$$

Problem (16) can be solved using the ICD method (7–9) assuming that $\hat{s}_r$ is known for all $r \neq m$, so that $\mathbf{Q}_m$, $\mathbf{R}_m$ and therefore $\mathbf{y}'_m$ and $\mathbf{W}'_m$ can be computed. To cope with the problem of unavailability of $\hat{s}_r$, an iterative scheme can be used: starting from an initial guess $\hat{\mathbf{s}}^{(0)} = \begin{bmatrix} \hat{s}_{-M}^{(0)} & \cdots & \hat{s}_M^{(0)} \end{bmatrix}^T$, at the ith iteration, $\mathbf{Q}_m^{(i)}$ can be computed using $\hat{s}_{-M}^{(i)}, \dots, \hat{s}_{m-1}^{(i)}, \hat{s}_{m+1}^{(i-1)}, \dots \hat{s}_M^{(i-1)}$. The proposed algorithm is summarized in Table I.

It is noteworthy that an excellent initial guess to start the iterations is to use the solution of (4).

C. Computational Complexity

We compute the complexity order of one iteration of the proposed algorithm summarized in Table I. Here is the order of computations for the 5 steps of the algorithm:

1. In Step 1, computation of $\mathbf{a}_r \mathbf{a}_r^H$ should be neglected, since it is performed only for one time at the beginning and will

TABLE I.   PROPOSED ALGORITHM

INPUT:

data-set $\mathbf{y}$, dictionary $\mathbf{W}$, regularization parameter λ, parameter ε, initial guess $\hat{\mathbf{s}}^{(0)} = \begin{bmatrix} \hat{s}_{-M}^{(0)} & \cdots & \hat{s}_M^{(0)} \end{bmatrix}^T$, number of iterations I.

ITERATIONS:

For $i = 1, 2, \dots, I$ do the following:

 For $m = -M, \dots, -1, 0, 1, \dots, M$ perform the following steps:

  1. Compute $\mathbf{Q}_m^{(i)}$ from the following relation:

$$\mathbf{Q}_m = \sum_{r=-M}^{m-1} \left| \hat{s}_r^{(i)} \right|^2 \mathbf{a}_r \mathbf{a}_r^H + \sum_{r=m+1}^{M} \left| \hat{s}_r^{(i-1)} \right|^2 \mathbf{a}_r \mathbf{a}_r^H + \varepsilon \mathbf{I}_N$$

  2. Obtain $\mathbf{R}_m^{(i)}$ from the Cholesky decomposition of $\left\{ \mathbf{Q}_m^{(i)} \right\}^{-1}$.

  3. Compute $\mathbf{y}'_m{}^{(i)} = \mathbf{R}_m^{(i)} \mathbf{y}$ and $\mathbf{W}'_m{}^{(i)} = \mathbf{R}_m^{(i)} \mathbf{W}$.

  4. Compute:

$$\mathbf{e}_m^{(i)} = \mathbf{y}'_m{}^{(i)} - \sum_{r=-M}^{m-1} \hat{s}_r^{(i)} \mathbf{w}'_{m,r}{}^{(i)} - \sum_{r=m+1}^{M} \hat{s}_r^{(i-1)} \mathbf{w}'_{m,r}{}^{(i)}$$

  where $\mathbf{w}'_{m,r}{}^{(i)}$ is the rth column of $\mathbf{W}'_m{}^{(i)}$.

  5. Compute $\hat{s}_m^{(i)}$ using the ICD method as:

$$\hat{s}_m^{(i)} = \phi_\lambda \left(\left\{ \mathbf{w}'_{m,m}{}^{(i)} \right\}^H \mathbf{e}_m^{(i)} \right)$$

---

not vary during the whole process. Also, we define the following matrix:

$$\mathbf{Q}^{(i)} = \sum_{r=-M}^{M} \left| \hat{s}_r \right|^2 \mathbf{a}_r \mathbf{a}_r^H \qquad (19)$$

Computing $\mathbf{Q}^{(i)}$ for only one time, all the matrices $\mathbf{Q}_m^{(i)}$ for $m = -M, \dots, M$ can be computed as:

$$\mathbf{Q}_m^{(i)} = \mathbf{Q}^{(i)} - \left| \hat{s}_m \right|^2 \mathbf{a}_m \mathbf{a}_m^H \qquad (20)$$

Therefore, Step 1 includes one single computation of $\mathbf{Q}^{(i)}$ of order $N^2 M$ and computation of $\mathbf{Q}_m^{(i)}$ from (20) for $m = -M, \dots, M$ each one of order N^2. The overall process will be of order $N^2 M$.

2. This step includes the computation of the inverse of $\mathbf{Q}_m^{(i)}$ and then the Cholesky decomposition (each of order N^3) for $m = -M, \dots, M$. The overall computations will be of order $N^3 M$.

3. The complexity of this step is dominated by the computation of the matrix multiplication $\mathbf{W}'_m{}^{(i)} = \mathbf{R}_m^{(i)} \mathbf{W}$ which is of order $N^2 M$ for each m. The overall cost is thus $N^2 M^2$.

4. Computational cost of Steps 4 and 5 is obviously negligible compared to Steps 2 and 3.

Therefore, the dominant steps are Step 2 of order N^3M and Step 3 of order N^2M^2. Since the dictionary is redundant, M is much larger than N, and therefore the overall computational complexity of the proposed method is of order N^2M^2.

Complexity of the ICD method to solve the BPDN problem (4) is of order NM which implies that the proposed method offers an improved performance in price of higher complexity.

III. SIMULATION RESULTS

In this section, the proposed method is compared with the BPDN method of [10] formulated in (4). To do so, we consider a signal with the following form:

$$y(t) = \sum_{k=1}^{L} A_k \cos\left(2\pi f_k t + \varphi_k\right) + \nu(t) \qquad (21)$$

where φ_k are *i.i.d.* random variables in interval $[0,2\pi)$ with uniform distribution and $\nu(t)$ is the additive noise. A data-set containing N samples is then formed by sampling this signal according to an irregular sampling pattern. The two methods are applied to the resulting data-set several times to be able to compare the frequency estimation methods accurately.

The simulated signal consists of three sinusoidal terms with frequencies 1, 1.1 and 3 Hz with amplitudes of 3, 4 and 1, respectively. The additive noise is assumed to be zero-mean white Gaussian with standard deviation of $\sigma = 0.2$. The sampling pattern shown in Figure 1 follows the Poisson distribution with mean $\Lambda = 1\ \text{s}^{-1}$; i.e. the difference between two adjacent sampling instants follows an exponential distribution with mean 1 s. Number of data samples is assumed to be $N = 24$. The frequency grid **f** is limited to $\pm f_{max} = \pm 10\ \text{Hz}$, therefore covers the interval of -10 to 10 Hz with step-size of 0.05 Hz. Only $I = 5$ iterations of the proposed algorithm were performed.
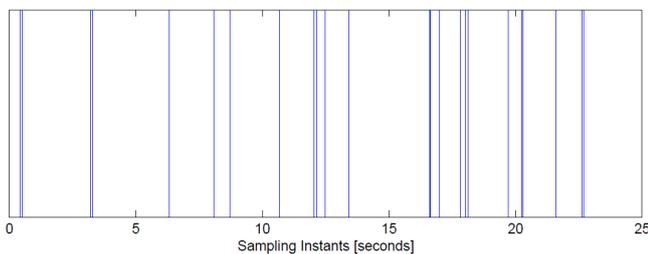
Fig. 1.   Sampling pattern.

The value of the regularization parameter in (4) is selected as:

$$\lambda = \rho\lambda_{max} \qquad (22)$$

where λ_{max} is defined as:

$$\lambda_{max} = \frac{1}{2}\max\left|\mathbf{W}^H\mathbf{y}\right| \qquad (23)$$

and ρ is a constant parameter [10]. For $\rho \geq 1$ the solution to (4) is identically zero [10], therefore ρ is assumed to be less than 1.

In practice, $\rho = 0.05$ seems to be a good choice [10]. In simulations, we use $\rho = 0.05$ for both the proposed method and BPDN method. The value of ε is assumed to be 0.001.

The result of 20 independent simulations altogether has been shown in Figure 2. As can be seen from Figure 2(a), the spectrum estimated by BPDN is full of spurious peaks. The amplitudes of many of these spurious peaks are higher than that of the actual peak at 3 Hz. Figure 2(b) shows that spurious peaks obtained by the proposed method are very rare with very small amplitudes. Moreover, the proposed method is less likely to miss the small actual peak at 3 Hz. To illustrate this, the area nearby the frequency 3 Hz in Figure 2 has been zoomed in and shown in Figure 3. Accurate counting of the number of circles (detected peaks) at 3 Hz shows that the number of those peaks with nonzero amplitude in Figure 3(a) is 15, while in Figure 3(b) is 20. This means that BPDN has missed this peak in 5 experiments, while the proposed method has never missed it.

Figure 2(b) shows that the amplitude of the spectral peaks suggested by the proposed method is not exact. However, this should not be a serious problem. After the detection of the spectral peaks, their amplitude can be corrected by lest-squares fitting to the data-set **y**. The result of fitting the 3 spectral peaks detected by the proposed method to **y** is shown in Figure 4. The actual values of the amplitudes are also shown in this figure with cross signs. As seen, the estimate values are very close to the actual ones.
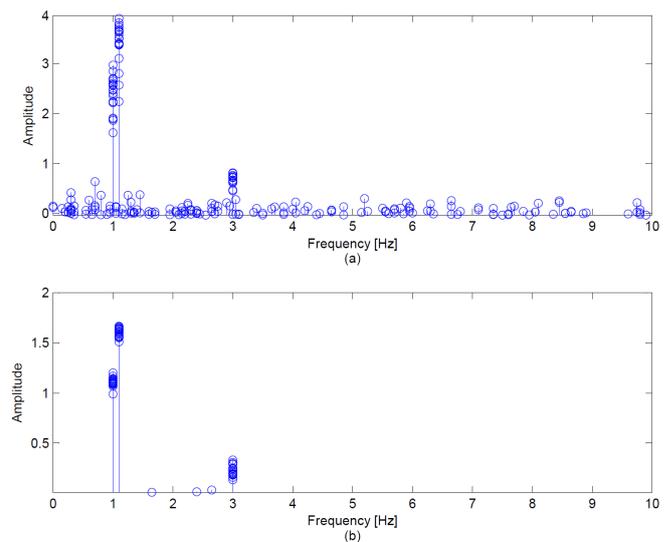
Fig. 2.   Result of 20 independent simulations, (a) BPDN, (b) proposed method.
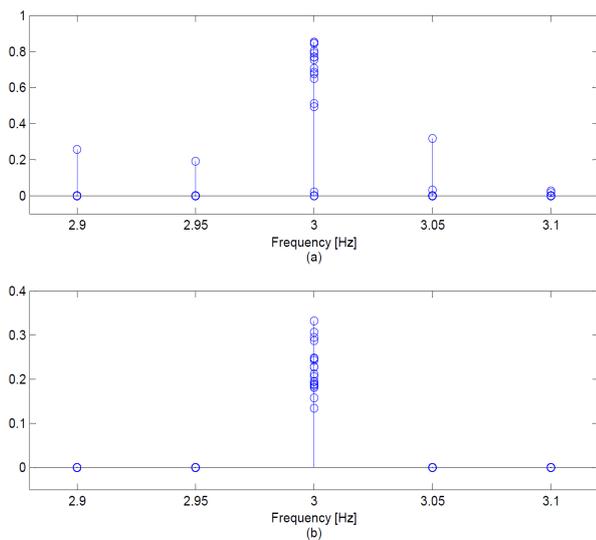
Fig. 3.   Zooming in the nearby of frequency 3 Hz in Figure 2, (a) BPDN, (b) proposed method.
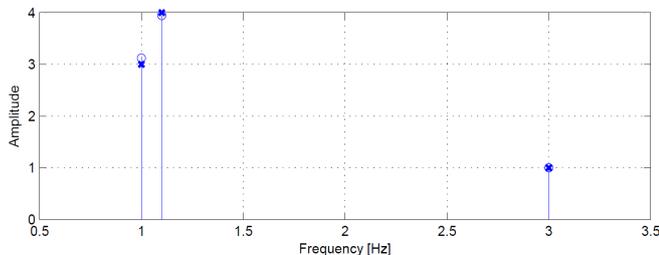
Fig. 4.   Estimate amplitude of the spectral peaks detected by the proposed method after least-squares fitting to the data-set. The actual values are shown by cross signs.

IV.   CONCLUSION

An iterative sparsity-based method was proposed for the estimation of spectral lines of nonuniformly-sampled signals. The proposed method enforces the sparsity constraint to a weighted least-squares problem. It was shown via computer simulations that it is capable of estimating the spectral lines of the signal provided that a very limited number of data samples taken with arbitrary sampling pattern is available. The proposed method seems to be sensitive to the choice of a parameter named ε. Therefore, the selection of an appropriate value for this parameter remains an open problem.

REFERENCES

[1]   R. Pribic, "Radar irregular sampling", IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '04, Vol. 3, pp. 933–936, 2004

[2]   S. Bourguignon, H. Carfantan, T. Böhm, "SparSpec: a new method for fitting multiple sinusoids with irregularly sampled data", Astronomy and Astrophysics, Vol. 462, pp. 379–387, 2007.

[3]   J. Mateo, P. Laguna, "Improved heart rate variability signal analysis from the beat occurrence times according to the IPFM model", IEEE Transactions on Biomedical Engineering, Vol. 47, No. 8, pp. 985–996, 2000.

[4]   L. H. Benedict, H. Nobach, C. Tropea, "Estimation of turbulent velocity spectra from laser Doppler data", Measurement Science and Technology, Vol. 11, No. 8, pp. 1089–1104, 2000.

[5]   N. R. Lomb, "Least-squares frequency analysis of unequally spaced data", Astrophysics and Space Science, Vol. 39, No. 1, pp. 447–462, 1976.

[6]   I. F. Gorodnitsky, B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm", IEEE Transactions on Signal Processing, Vol. 45, No. 3, pp. 600–616, 1997.

[7]   M. D. Sacchi, T. J. Ulrych, C. J. Walker, "Interpolation and extrapolation using a high-resolution discrete Fourier transform", IEEE Transactions on Signal Processing, Vol. 46, No. 1, pp. 31–38, 1998.

[8]   P. Ciuciu, J. Idier J. –F. Giovannelli, "Regularized estimation of mixed spectra using a circular Gibbs-Markov model", IEEE Transactions on Signal Processing, Vol. 49, No. 10, pp. 2202–2213, 2001.

[9]   S. Chen, D. Donoho "Application of basis pursuit in spectrum estimation", IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 3, pp. 1865–1868, 1998.

[10]   S. Bourguignon, H. Carfantan, J. Idier, "A sparsity-based method for the estimation of spectral lines from irregularly sampled data", IEEE Journal of Selected Topics in Signal Processing, Vol. 1, No. 4, pp. 575–585, 2007.

[11]   S. S. Chen, D. L. Donoho, M. Saunders, "Atomic decomposition by basis pursuit", SIAM Journal on Scientific Computing, Vol. 20, No. 1, pp. 33–61, 1999.

[12]   D. L. Donoho, M. Elad, V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise", IEEE Transactions on Information Theory, Vol. 52, No. 1, pp. 6–18, 2006.

[13]   P. Stoica, Jian Li, H. He, "Spectral analysis of nonuniformly sampled data: a new approach versus the periodogram", IEEE Transactions on Signal Processing, Vol. 57, No. 3, pp. 843–858, 2009.

[14]   W. Roberts, P. Stoica, Jian Li, T. Yardibi, F. A. Sadjadi, "Iterative adaptive approaches to MIMO radar imaging", IEEE Journal of Selected Topics in Signal Processing: Special Issue on MIMO Radar and Its Applications, Vol. 4, No. 1, pp. 5–20, 2010.

[15]   T. Yardibi, Jian Li, P. Stoica, Ming Xue, A. B. Baggeroer, "Source localization and sensing: a nonparametric iterative adaptive approach based on weighted least squares", IEEE Transactions on Aerospace and Electronic Systems, Vol. 46, No. 1, pp. 425–443, 2010.

[16]   D. Watkins, Fundamentals of matrix comutations, John Wiley and Sons, 2002.