Signal Processing 142 (2018) 87–95

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Two sparse-based methods for off-grid direction-of-arrival estimation

Xiaohuan Wu[a,*], Wei-Ping Zhu[a,b], Jun Yan[a], Zeyun Zhang[a]

[a] The Key Lab of Broadband Wireless Communication and Sensor Network Technology, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China

[b] Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada

ARTICLE INFO

Article history:
Received 26 April 2017
Revised 15 June 2017
Accepted 7 July 2017
Available online 8 July 2017

Keywords:
Direction-of-arrival (DOA) estimation
Off-grid model
Sparse-based method
Covariance matrix

ABSTRACT

Recently, many sparse-based methods have been proposed for direction-of-arrival (DOA) estimation. However, these methods often suffer from the grid mismatch problem caused by the discretization of the potential angle space. Most of them employ the iterative grid refinement (IGR) method to alleviate this problem. Nevertheless, IGR requires a high computational load and may not comply with the restricted isometry property (RIP) condition in the compressed sensing (CS) theory. This paper aims to overcome the grid mismatch limitation inherent in conventional sparse-based techniques. In particular, we first introduce an off-grid model by incorporating the bias parameter into the signal model, then propose a two-step iterative method named off-grid ℓ_1 Cholesky covariance decomposition (OGL1CCD) to solve the DOA estimation problem. Our method can be accelerated to save computations and the proposed algorithm framework can be extended for any other sparse-based method to improve their estimation accuracy. We then propose another off-grid method named off-grid ℓ_1 covariance matrix reconstruction approach (OGL1CMRA) based on the covariance matrix model. Compared to OGL1CCD, OGL1CMRA is more computationally efficient and accurate, but requires sufficient snapshots and uncorrelated sources. Our proposed methods are superior to many other methods in estimation performance, which is verified by extensive numerical simulations.

1. Introduction

Direction finding, also known as direction-of-arrival (DOA) estimation, is a fundamental problem in array signal processing and has been intensively studied in the past few decades for various applications, e.g., radar [1], sonar and wireless communications [2]. Since the 1970s when Pisarenko found the DOAs can be estimated from the second order statistics of the received signal, a large number of methods have been proposed for DOA estimation, e.g., MUSIC [3], ESPRIT [4] and their variants [5–8]. Although these methods show super resolution ability in some certain scenarios, they often suffer from several limitations. For example, the so-called subspaced-based methods highly depend on the number of sources and the estimation accuracy of the covariance matrix of the array output and are, in general, sensitive to the number of snapshots and signal-to-noise ratio (SNR) as well. Furthermore, the correlation between the impinged signals may cause a rank deficiency problem in the sample covariance matrix, which may lead to a poor estimation performance.

Recently, by exploring the underlying connection between the sparse signal reconstruction (SSR) [9,10] and the DOA estimation, many sparse methods have been proposed by dividing the potential angle space with a predefined set of dense grid points and assuming the fact that the true DOAs exactly lie on the predefined grid [11–20]. With the help of SSR theory, these sparse methods are usually insensitive to the number of snapshots (most of them can correctly locate sources with a single snapshot) and the correlation between the sources. However, the set of grid points usually contains a finite number of atoms while the DOAs of the impinged sources are continuously valued, which leads to infinitely many atoms. The sparse methods then assume that the DOAs of the sources exactly lie on the grids and are also called on-grid method. This assumption, although widely employed by many sparse methods, holds true only when the number of grid points is infinitely large. Consequently, there always exists an unavoidable bias between the true DOA and its nearest grid point, namely, the bias mismatch issue is inherent in the on-grid methods. Using a set of dense grids is able to mitigate this basis mismatch issue but is not an appropriate choice since it not only increases the computational cost but also gives rises to the correlation issue. A high complexity may reduce the practicability of the algorithm while a high correlation may conflict with the restrict isometry property

* Corresponding author.
E-mail address: 2013010101@njupt.edu.cn (X. Wu).

http://dx.doi.org/10.1016/j.sigpro.2017.07.004
0165-1684/

(RIP) in the compressive sensing (CS) theory [21]. Therefore, almost all the sparse methods choose the iterative grid refinement (IGR) procedure to find a balance between the accuracy and the efficiency. The IGR uses an initial set of coarse grid points along with a sparse-based algorithm to obtain a coarse DOA estimate. It then refines the angle space division around the coarse DOA estimate with a new set of finer grid points and recalls the sparse-based algorithm to achieve a better DOA estimate. However, the use of a set of dense grid points in IGR would contradict the RIP condition required by the sparse algorithm. Hence, further improvement of the estimation performance of IGR is impossible.

To address the basis mismatch issue, some so-called off-grid sparse methods have been proposed [22–26]. Compared to the on-grid methods, the discretization is still required in these methods but the DOAs of the sources are not restricted to lie on the grid. Instead, the bias is parameterized into the signal model by using the first order approximation of the manifold matrix. But the new model usually is nonconvex and hence is not easy to solve. For example, the nonconvex model in [22,23] has been solved from the sparse Bayesian learning (SBL) perspective. In [24–26], an alternating procedure is adopted to solve the potential sources and biases. However, a difficult problem in [24] is the parameter tuning while the methods in [25,26] are time-consuming. In conclusion, the main focus of these methods has been to determine the biases. Recently, Tang et al. proposed a gridless method based on the atomic norm theory without requiring grid division [27] and showed that this method can be regarded as the sparse method with an infinite size of the grids, provided that the sources are sufficiently separated. Nevertheless, this precondition prohibits commonly known high resolution of the DOA estimation requirement. We proposed a gridless method named as covariance matrix reconstruction approach (CMRA) by exploiting the Toeplitz structure of the covariance matrix of the array output [28,29]. Since it explores the maximum number of degrees of freedom (DOF), CMRA is able to locate more sources than sensors.

In this paper, we address the DOA estimation problem in an off-grid mode under the sparse framework with an objective of aiming to overcome the grid mismatch limitation inherent in conventional sparse techniques. In particular, we start with a set of coarse grid points and then introduce a parameter to describe the bias between the true DOA and its nearest grid point. The unknown DOAs are determined by its nearest grid point together with the corresponding bias. We propose a two-step iterative technique for joint sparse signal recovery and bias estimation. In the first step, we fix the bias and obtain the peak indices of the recovered sparse signal, while in the second step, we fix the sparse signal and determine the bias by deriving a closed-form solution. To formulate exactly the problem, we propose two methods based on the output of the array and its covariance matrix, named as off-grid ℓ_1 Cholesky covariance decomposition (OGL1CCD) and off-grid ℓ_1 covariance matrix reconstruction approach (OGL1CMRA), respectively. In the first method, the complexity of the problem in step one is reduced by the dimensionality reduction technique. An update rule for the regularization parameter is also proposed. In the second method, we derive the cramer-rao lower bound (CRLB) of the bias to show that the covariance-based model is able to achieve more accurate bias estimates if sufficient snapshots are collected. We also point out that, the dimensionality reduction and the parameter updating are not required in the second method. Furthermore, the proposed algorithm can be extended for any other sparse methods to improve their estimation accuracy.

Notations used in this paper are as follows. $\boldsymbol{A}^*$, $\boldsymbol{A}^T$, $\boldsymbol{A}^H$ and $\boldsymbol{A}^\dagger$ denote the conjugate, transpose, conjugate transpose and pseudo-inverse of matrix $\boldsymbol{A}$, respectively. $\boldsymbol{A}_{(n)}$ denotes the nth row of $\boldsymbol{A}$. vec($\boldsymbol{A}$) denotes the vectorization operator that stacks matrix $\boldsymbol{A}$ column by column. $\boldsymbol{A}\circ\boldsymbol{B}$, $\boldsymbol{A}\otimes\boldsymbol{B}$ and $\boldsymbol{A}\odot\boldsymbol{B}$ are the Hadamard, Kronecker and Khatri-Rao products of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. tr($\bullet$) is the trace operator. $\boldsymbol{I}_N$ denotes the identity matrix of size $N\times N$. $\|\boldsymbol{A}\|_1$ and $\|\boldsymbol{A}\|_F$ denote the ℓ_1-norm and Frobenius norm of $\boldsymbol{A}$, respectively. $\boldsymbol{A}\geq\boldsymbol{0}$ means that matrix $\boldsymbol{A}$ is positive semidefinite. For a vector $\boldsymbol{x}$, $\|\boldsymbol{x}\|_2$ denotes the ℓ_2-norm of $\boldsymbol{x}$. $\boldsymbol{x}\succeq\boldsymbol{0}$ means that every entry of $\boldsymbol{x}$ is non-negative, diag($\boldsymbol{x}$) denotes a diagonal matrix with its diagonal entries being the entries of vector $\boldsymbol{x}$ in turn.

The rest of the paper is organized as follows. Section 2 revisits the on-grid and off-grid models, followed by a CRLB analysis on the bias estimate. Section 3 presents the proposed iterative methods. A new covariance-based off-grid method is proposed in Section 4. Simulations are carried out in Section 5 to demonstrate the performance of our methods. Finally, Section 6 concludes the whole paper.

2. Off-grid signal model

2.1. On-grid signal model

Suppose that K narrowband far-field signals impinge onto an array with M omnidirectional sensors from directions of $\boldsymbol{\theta}=\{\theta_1,\cdots,\theta_K\}$ simultaneously. The array output at time t, which is corrupted by additive circular complex Gaussian white noise, can be expressed as,

$$\boldsymbol{x}(t)=\sum_{k=1}^{K}\boldsymbol{a}(\theta_k)s_k(t)+\boldsymbol{e}(t)=\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(t)+\boldsymbol{e}(t), \tag{1}$$

where $\boldsymbol{x}(t)=[x_1(t),\cdots,x_M(t)]^T$ is the array output, $\boldsymbol{s}(t)=[s_1(t),\cdots,s_K(t)]^T$ is the signal waveform, $\boldsymbol{A}(\boldsymbol{\theta})=[\boldsymbol{a}(\theta_1),\cdots,\boldsymbol{a}(\theta_K)]$ is the array manifold matrix, $\boldsymbol{a}(\theta_k)=[e^{j2\pi f_0\tau_{k,1}},\cdots,e^{j2\pi f_0\tau_{k,M}}]^T$ contains the time-delay of the kth signal received at each sensor relative to the reference sensor, and $\boldsymbol{e}(t)$ is the complex independent white Gaussian noise with zero mean and variance $\sigma\boldsymbol{I}$.[1] When L snapshots are collected, the array output can be given by

$$\boldsymbol{X}=\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{S}+\boldsymbol{E}, \tag{2}$$

where $\boldsymbol{X}=[\boldsymbol{x}(t_1),\cdots,\boldsymbol{x}(t_L)]$, $\boldsymbol{S}=[\boldsymbol{s}(t_1),\cdots,\boldsymbol{s}(t_L)]$ and $\boldsymbol{E}=[\boldsymbol{e}(t_1),\cdots,\boldsymbol{e}(t_L)]$. The goal is to determine the unknown DOAs $\boldsymbol{\theta}$ given the array output $\boldsymbol{X}$.

Uniformly sampling the angle space generates a fixed finite set of N potential angles $\boldsymbol{\vartheta}=\{\vartheta_1,\cdots,\vartheta_N\}$ and a manifold dictionary $\bar{\boldsymbol{A}}=[\boldsymbol{a}(\vartheta_1),\cdots,\boldsymbol{a}(\vartheta_N)]$. Generally speaking, N is much greater than the incident sources number K and the sensors number M. We first assume $\boldsymbol{\theta}\subset\boldsymbol{\vartheta}$, i.e., all interested unknown DOAs exactly lie on the predefined grid. Thus, the array output in (2) can be reformulated as an multiple measurement vectors (MMV) model,

$$\boldsymbol{X}=\bar{\boldsymbol{A}}\bar{\boldsymbol{S}}+\boldsymbol{E}, \tag{3}$$

where $\bar{\boldsymbol{S}}$ is the extension of $\boldsymbol{S}$ from $\boldsymbol{\theta}$ to $\boldsymbol{\vartheta}$ with non-zero entries denoting the true sources locations. Since $N\gg K$, the signal $\bar{\boldsymbol{S}}$ is row-sparse. The problem of DOA estimation is now transformed into an SSR problem and the joint sparsity can be exploited for signal recovery. Usually, following the ℓ_1-norm optimization in a single measurement vector (SMV) problem, the $\ell_{2,1}$-norm, which is defined as $\|\boldsymbol{X}\|_{2,1}=\sum_n\|\boldsymbol{X}_{(n)}\|_2$, is expected to be an appropriate penalty. The problem is then formulated as,

$$\min_{\bar{\boldsymbol{S}}}\ \lambda\left\|\bar{\boldsymbol{S}}\right\|_{2,1}+\frac{1}{2}\left\|\boldsymbol{X}-\bar{\boldsymbol{A}}\bar{\boldsymbol{S}}\right\|_F^2, \tag{4}$$

where λ is the regularization parameter controlling the tradeoff between the sparsity of the solution and the data fitting error.

[1] We have assumed that the noise variances σ_m $(m=1,\cdots,M)$ are equivalent, which is true in most scenarios.