Journal of Systems Engineering and Electronics
Vol. 35, No. 3, June 2024, pp.532 – 540

DOA estimation of high-dimensional signals based on Krylov subspace and weighted l_1-norm

YANG Zeqi[1,2,3], LIU Yiheng[1,2,3], ZHANG Hua[1,2,3], MA Shuai[1,2,3], CHANG Kai[4], LIU Ning[4], and LYU Xiaode[1,2,*]

1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;
2. National Key Laboratory of Microwave Imaging Technology, Beijing 100190, China;
3. School of Electronic Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China; 4. Northern Institute of Electronic Equipment, Beijing 100191, China

Abstract: With the extensive application of large-scale array antennas, the increasing number of array elements leads to the increasing dimension of received signals, making it difficult to meet the real-time requirement of direction of arrival (DOA) estimation due to the computational complexity of algorithms. Traditional subspace algorithms require estimation of the covariance matrix, which has high computational complexity and is prone to producing spurious peaks. In order to reduce the computational complexity of DOA estimation algorithms and improve their estimation accuracy under large array elements, this paper proposes a DOA estimation method based on Krylov subspace and weighted l_1-norm. The method uses the multistage Wiener filter (MSWF) iteration to solve the basis of the Krylov subspace as an estimate of the signal subspace, further uses the measurement matrix to reduce the dimensionality of the signal subspace observation, constructs a weighted matrix, and combines the sparse reconstruction to establish a convex optimization function based on the residual sum of squares and weighted l_1-norm to solve the target DOA. Simulation results show that the proposed method has high resolution under large array conditions, effectively suppresses spurious peaks, reduces computational complexity, and has good robustness for low signal to noise ratio (SNR) environment.

Keywords: direction of arrival (DOA), compressed sensing (CS), Krylov subspace, l_1-norm, dimensionality reduction.

DOI: 10.23919/JSEE.2023.000145

1. Introduction

One of the research hotspots of array signal processing is to use array antennas to estimate the angle parameters of illuminators and achieve precise localization of these sources. The precise estimation of direction of arrival (DOA) has been widely studied in fields such as radar positioning, array direction finding, and satellite communication. In recent years, in order to obtain more flexible beam control, higher angular resolution, and more accurate coverage [1], the number of antennas has increased significantly. For example, compared with traditional four-element and eight-element arrays, the number of elements in large-scale phased-array antennas can range from several hundreds to tens of thousands. For large-scale multiple input multiple output (MIMO) radar arrays [2], the data dimension is the product of the transmission and reception dimensions, so its dimension is even higher. With the application of large-scale arrays, the number of array elements is increasing, making the computational complexity of traditional spatial spectrum estimation methods significantly increased, which is proportional to the cube of the number of elements and difficult to process in real-time due to the strong temporal variation of the received signal with an increase in the number of illuminators [3,4]. The "high dimensionality" of high-dimensional signals leads to a high computational complexity of matrix inversion and covariance matrix operations, and spurious peaks are easily generated under low signal to noise ratio (SNR). The spatial sparsity of the signals is of great practical significance for estimating the DOAs of high-dimensional signals.

Development of DOA estimation algorithms has experienced from conventional beamformer (CBF)/Bartlett beamformer [5] to adaptive beamforming methods that

Manuscript received March 27, 2023.
*Corresponding author.
This work was supported by the National Basic Research Program of China.

solve adaptive weights with different optimal criteria, such as Burg's maximum entropy method (MEM) [6] and Capon's minimum variance method [7]. After 1986, Schmidt proposed the multiple signal classification (MUSIC) algorithm [8], which divides array received signals into two mutually orthogonal subspaces and uses the orthogonal characteristics of the two subspaces to determine the DOA. Using the rotation-invariant characteristics of the signal subspace, Roy et al. proposed the estimation of signal parameters via rotational invariance techniques (ESPRIT) algorithm [9]. Compared with the MUSIC algorithm, it does not require spectral peak searching and directly estimates DOA using the least squares method. Super-resolution spectral estimation methods further improve the resolution of the algorithm, breaking the "Rayleigh limit" and developing rapidly in many fields such as radar, communication and sonar. The limitation of these traditional algorithms is that the conventional signal subspace estimation method requires estimation of the covariance matrix of the received signals and performs eigenvalue decomposition or singular value decomposition (SVD) on it, and its computational complexity is proportional to the cube of the number of elements. Therefore, under the condition of a large number of array elements, it usually cannot meet the real-time estimation requirements. The Krylov subspace [10] is a method for dimensionality reduction and solving large sparse matrix, which has been proven to be equivalent to the signal subspace under certain conditions. The multistage Wiener filter (MSWF) method proposed by [11,12] is a method that uses the Krylov subspace to solve DOA, which does not require estimation of the covariance matrix, has low computational complexity and fast convergence speed, but requires prior knowledge of the expected signal and cannot be directly applied to DOA estimation.

With the application of the sparse reconstruction of compressed sensing theory [13−17] in DOA estimation, high-dimensional signals are usually spatial sparsity, and reconstructing signals at a lower sampling rate has great advantages in solving high-dimensional signals processing caused by the increasing number of array elements [18,19]. Since Donoho [20] combined the spatial sparsity of the incident signal to reconstruct the received signal, which achieves super-resolution estimation. Using sparse reconstruction into DOA estimation has been widely studied. Malioutov et al. [21] first combined l_1-norm and SVD to achieve accurate DOA estimation at high SNR and small number of snapshots by solving optimization functions. However, this algorithm has low resolution and inaccurate estimation at low SNR. Cong et al. [22−24] proposed a DOA estimation based on covariance sparse reconstruction, which performs well under conditions of low SNR and small number of snapshots, but has a high complexity. Zuo et al. [25−27] proposed a method based on weighted l_1-SVD algorithm, which is determined by the orthogonality of the subspace and determined a weighted matrix, the DOA of sparse signal is obtained by solving the optimization function. Wang et al. proposed a DOA estimation method based on weighted improved l_0-norm [28], which has good DOA estimation performance under low SNR and few snapshots. These algorithms can effectively suppress the deviation, but the calculation complexity is high under a large number of array elements.

Based on the sparsity of signal in spatial domain, the iterative algorithm can be used to realize signal reconstruction, such as orthogonal matching pursuit (OMP) [29] and approximate message passing (AMP). The greedy algorithm, represented by OMP, solves the optimal solution by iteration, which has low computational complexity, but the accuracy of DOA estimation deteriorates in the case of low SNR. The AMP algorithm is a classical sparse reconstruction method by iterative denoising. The AMP-net [30] algorithm has been applied to deep learning model as a data-driven sparse reconstruction algorithm. Making full use of the sparsity of l_1-norm and the optimization of sampling matrix can effectively improve the reconstruction performance.

In order to solve the problem of high computational complexity and inaccurate estimation of DOA of high-dimensional signals under a large number of array elements, this paper proposes a method based on Krylov subspace and weighted l_1-norm for DOA estimation. Firstly, the received signal model of a uniform linear array (ULA) is established, and Krylov subspace is introduced to iteratively solve the signal subspace, which reduces the computational complexity caused by the matrix SVD in the l_1-SVD algorithm under a large-scale array. The signal subspace is estimated quickly by implementing MSWF iterations based on the basis of the Krylov subspace, and the weighted matrix is determined by the orthogonality of the subspace. Then, a measurement matrix is designed to further reduce the dimensions of the signal subspace and the perception matrix, and a DOA estimation model is established based on the joint weighted l_1-norm of the sparse signal residual sum of squares. The robustness of the algorithm is improved by constructing the objective function using a weighted scheme. Finally, the second-order cone optimization problem is solved to obtain an accurate estimation of the DOA. In this paper, the algorithm avoids the calculation of covariance matrix of received data and reduces the computational complexity of the algorithm through two operations to reduce dimension in the case of large-scale array. And the objective function is constructed by using the weighted l_1-norm, so that the reconstructed signal is equally constrained, the sparsity of the solution is guaranteed, and the estimation accuracy of the algorithm is improved.

2. Algorithm description

Assuming that the array consists of M identical antennas uniformly arranged in a straight line, $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_K]$ represents the DOA of uncorrelated signals incident on the array by K ($M>K$) illuminators, as shown in Fig. 1. Assuming that K is known and the array element spacing is half wavelength, the data model received by the array at time t is given by

$$\boldsymbol{x}(t) = \sum_{k=1}^{K} \boldsymbol{a}_k(\theta_k)\boldsymbol{s}_k(t) + \boldsymbol{n}(t),\ \ t = 1, 2, \cdots, T \tag{1}$$

where

$$\boldsymbol{a}_k(\theta_k) = \left[1, \mathrm{e}^{-\frac{\mathrm{j}2\pi d\sin\theta_k}{\lambda}}, \cdots, \mathrm{e}^{-\frac{\mathrm{j}2\pi(M-1)d\sin\theta_k}{\lambda}}\right]^{\mathrm{T}},$$

d is the array element spacing, λ is the wavelength of the signal, T is the number of snapshots, and θ_k is the DOA of the signal.

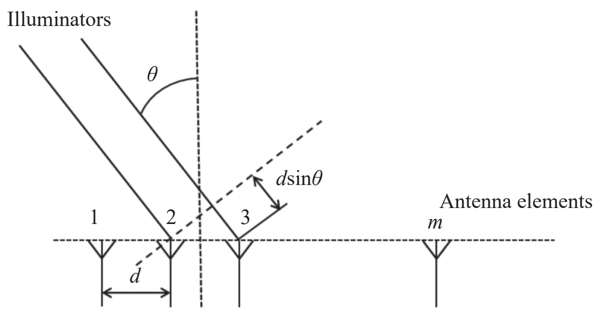


Fig. 1 ULA

Write (1) in vector form as follows:

$$\boldsymbol{X}(t) = \boldsymbol{A}(\theta)\boldsymbol{S}(t) + \boldsymbol{N}(t) \tag{2}$$

where $\boldsymbol{A}(\theta) = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_K)]$ is an array manifold matrix of $M\times K$ dimensions; $\boldsymbol{S}(t) = [\boldsymbol{s}_1(t), \boldsymbol{s}_2(t), \cdots, \boldsymbol{s}_K(t)]^{\mathrm{T}}$ is a signal of $K\times T$ dimensions; $\boldsymbol{N}(t)$ is Gaussian white noise with a mean value of zero and a variance of σ_n^2.

Due to the application of large-scale array, the number of array elements increases significantly, resulting in a high-dimensional received signal. The Krylov subspace algorithm is an iterative algorithm for solving large sparse linear equations. In this paper, The Krylov subspace algorithm is used to reduce the dimension of the received signal to obtain a low-dimensional signal subspace.

Let $\boldsymbol{R}_{x_0} \in \mathbf{C}^{m\times m} \boldsymbol{f} \in \mathbf{C}^{m\times 1}$, then the Krylov subspace of D dimension is defined as

$$\mathcal{K}^D(\boldsymbol{R}_{x_0}, \boldsymbol{f}) = \mathrm{span}\left(\boldsymbol{f}, \boldsymbol{R}_{x_0}\boldsymbol{f}, \cdots, {\boldsymbol{R}_{x_0}}^{D-1}\boldsymbol{f}\right). \tag{3}$$

Assuming that there is any scalar ρ, Krylov subspace has the following properties:

$$\mathcal{K}^D(\boldsymbol{R}_{x_0}, \boldsymbol{f}) = \mathcal{K}^D(\boldsymbol{R}_{x_0} - \rho\boldsymbol{I}_M, \boldsymbol{f}). \tag{4}$$

The covariance matrix $\boldsymbol{R}_{\mathrm{xx}}$ of the data received by the array is

$$\begin{gathered}\boldsymbol{R}_{\mathrm{xx}} = \mathrm{E}\left\{\boldsymbol{X}(t)\boldsymbol{X}^{\mathrm{H}}(t)\right\} = \\ \mathrm{E}\left\{\{\boldsymbol{A}(\theta)\boldsymbol{S}(t) + \boldsymbol{N}(t)\}\{\boldsymbol{A}(\theta)\boldsymbol{S}(t) + \boldsymbol{N}(t)\}^{\mathrm{H}}\right\} = \\ \boldsymbol{A}\mathrm{E}\left\{\boldsymbol{S}(t)\boldsymbol{S}(t)^{\mathrm{H}}\right\}\boldsymbol{A}^{\mathrm{H}} + \mathrm{E}\left\{\boldsymbol{N}(t)\boldsymbol{N}(t)^{\mathrm{H}}\right\} = \\ \boldsymbol{A}(\theta)\boldsymbol{R}_s\boldsymbol{A}(\theta)^{\mathrm{H}} + \sigma_n^2\boldsymbol{I}_M\end{gathered} \tag{5}$$

where $\boldsymbol{R}_s = \mathrm{E}\left\{\boldsymbol{S}(t)\boldsymbol{S}(t)^{\mathrm{H}}\right\}$. $\boldsymbol{I}_M$ is the identity matrix whose dimension is $M\times M$.

The array covariance matrix $\boldsymbol{R}_{\mathrm{xx}}$ satisfies the Krylov subspace property, including

$$\mathcal{K}^D(\boldsymbol{R}_{\mathrm{xx}}, \boldsymbol{f}) = \mathcal{K}^D\left(\boldsymbol{R}_{\mathrm{xx}} - \sigma_n^2\boldsymbol{I}_M, \boldsymbol{f}\right) = \mathcal{K}^D\left(\boldsymbol{A}(\theta)\boldsymbol{R}_s\boldsymbol{A}(\theta)^{\mathrm{H}}, \boldsymbol{f}\right). \tag{6}$$

When the vector $\boldsymbol{f}$ is not orthogonal to the signal subspace, the maximum dimension of $\mathcal{K}^D(\boldsymbol{R}_{\mathrm{xx}}, \boldsymbol{f})$ is $K+1$ dimension, let $\boldsymbol{f}$ be a linear combination of basis of the signal subspace $\boldsymbol{U}_s$. $\boldsymbol{f}$ can be expressed as

$$\boldsymbol{f} = \alpha_1\boldsymbol{\lambda}_1 + \alpha_2\boldsymbol{\lambda}_2 + \cdots + \alpha_k\boldsymbol{\lambda}_k. \tag{7}$$

Since the signal subspace is orthogonal to the noise subspace, $\boldsymbol{f}$ is orthogonal to the noise subspace, and the rank of Krylov subspace is K, i.e.,

$$\mathrm{span}\left(\boldsymbol{f}, \boldsymbol{R}_{\mathrm{xx}}\boldsymbol{f}, \cdots, {\boldsymbol{R}_{\mathrm{xx}}}^{K-1}\boldsymbol{f}\right).$$

Weippert et al. [11] proved that there exists a set of orthogonal basis $\boldsymbol{G} = [g_1, g_2, \cdots, g_K]_{M\times K}$, and a nonsingular matrix $\boldsymbol{F} \in \mathbf{C}^{K\times K}$ in K-dimensional Krylov subspace:

$$\begin{gathered}\boldsymbol{G} = \mathcal{K}^K(\boldsymbol{R}_{\mathrm{xx}}\boldsymbol{f})\boldsymbol{F} = \mathcal{K}^K\left(\boldsymbol{A}(\theta)\boldsymbol{R}_s\boldsymbol{A}(\theta)^{\mathrm{H}}, \boldsymbol{f}\right)\boldsymbol{F} = \\ \left[\boldsymbol{f}, \boldsymbol{U}_s\boldsymbol{V}_s{\boldsymbol{U}_s}^{\mathrm{H}}\boldsymbol{f}, \cdots, \boldsymbol{U}_s{\boldsymbol{V}_s}^{K-1}{\boldsymbol{U}_s}^{\mathrm{H}}\boldsymbol{f}\right]\boldsymbol{F} = \\ \boldsymbol{U}_s\left[{\boldsymbol{U}_s}^{\mathrm{H}}\boldsymbol{f}, \boldsymbol{V}_s{\boldsymbol{U}_s}^{\mathrm{H}}\boldsymbol{f}, \cdots, {\boldsymbol{V}_s}^{K-1}{\boldsymbol{U}_s}^{\mathrm{H}}\boldsymbol{f}\right]\boldsymbol{F} = \boldsymbol{U}_s\boldsymbol{Q}\end{gathered} \tag{8}$$

where

$$\boldsymbol{Q} = \left[{\boldsymbol{U}_s}^{\mathrm{H}}\boldsymbol{f}, \boldsymbol{V}_s{\boldsymbol{U}_s}^{\mathrm{H}}\boldsymbol{f}, \cdots, {\boldsymbol{V}_s}^{K-1}{\boldsymbol{U}_s}^{\mathrm{H}}\boldsymbol{f}\right]\boldsymbol{F}.$$

Therefore, if the linear combination of all input signals can be represented by vector $\boldsymbol{f}$, then

$$\mathcal{K}^K(\boldsymbol{R}_{xx},\boldsymbol{f})=\operatorname{span}(\boldsymbol{U}_s). \tag{9}$$

The Wiener-Hopf equation is solved by forward iteration of K-order Wiener filter: $\boldsymbol{R}_{x_0}\boldsymbol{W}_{x_0}=\boldsymbol{r}_{x_0d_0}$. In this paper, $\boldsymbol{R}_{x_0}$ is the received data, and $\boldsymbol{r}_{x_0d_0}$ is the cross-correlation matrix between the received signal and the expected signal, $\boldsymbol{B}_i$ is a blocking matrix that suppresses signals in the desired direction. Wiener filter is composed of decomposition filter bank and synthesis filter bank based on the minimum mean square error. The input signal is decomposed by orthogonal projection property, and the filter output error is obtained by forward recursion. MSWF does not require covariance matrix estimation and has fast convergence speed, so it has low computational complexity and high real-time performance in the processing of high-dimensional signals. When the column vectors of the dimensionality reduction matrix are orthogonal, the K-order Wiener filter is equivalent to a set of orthogonal basis of the Wiener-Hopf equation in the K-dimensional Krylov subspace $\mathcal{K}^K(\boldsymbol{R},\boldsymbol{r})$.

The structural schematic diagram of the MSWF is shown in Fig. 2.

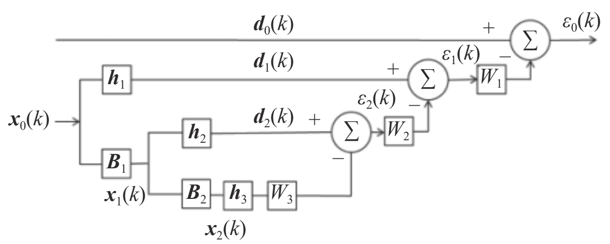


Fig. 2 Two-stage Wiener filtering

In this paper, the blocking matrix is chosen as $\boldsymbol{B}_i=\operatorname{null}\{\boldsymbol{h}_i\}$ and $\boldsymbol{B}_1^{\mathrm{H}}\boldsymbol{h}_1=0$.

Initial value:

$$\begin{cases}\boldsymbol{X}_0=\boldsymbol{X}\\ \boldsymbol{d}_0=\operatorname{mean}(\boldsymbol{X}_0)\end{cases}. \tag{10}$$

Forward recursion:

$$\begin{cases}\boldsymbol{r}_{x_id_i}=\mathrm{E}[\boldsymbol{x}_i(k)\boldsymbol{d}_i^*(k)]\\ \boldsymbol{h}_{i+1}=\dfrac{\boldsymbol{r}_{x_id_i}}{\operatorname{norm}(\boldsymbol{r}_{x_id_i})}\\ \boldsymbol{d}_{i+1}(k)=\boldsymbol{h}_{i+1}^{\mathrm{H}}\boldsymbol{x}_i(k)\\ \boldsymbol{x}_{i+1}(k)=\boldsymbol{x}_i(k)-\boldsymbol{h}_{i+1}\boldsymbol{d}_{i+1}(k)\end{cases}. \tag{11}$$

Orthogonal vector $\boldsymbol{T}_K=[\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_K]$ is obtained by K times forward recursion, which is the signal subspace $\boldsymbol{U}_s$.

Because of the large number of elements, the dimension of the signal subspace $\boldsymbol{U}_s$ is still high. In this paper, the signal subspace obtained is further reduced in dimension, and the signal subspace is sparse in space. Therefore, a Gaussian random measurement matrix $\boldsymbol{\Phi}\in\mathbf{C}^{p\times M}(p<N)$ is designed to compress and observe the vectorized signal subspace, and the following results are obtained:

$$\boldsymbol{U}_{ss}=\boldsymbol{\Phi}\times\boldsymbol{U}_s\in\mathbf{C}^{p\times K}. \tag{12}$$

Construct a sparse representation DOA estimation model: divide the possible incident angle range $[-90,90]$ of the signal into $\{\theta_1,\theta_2,\cdots,\theta_L\}$ at equal intervals, then each direction corresponds to a column of the array manifold matrix, and construct a spatial over-complete dictionary set $\boldsymbol{B}=[a(\theta_1),a(\theta_2),\cdots,a(\theta_L)]$.

Sparse signal $\boldsymbol{S}=[s_1(t),s_2(t),\cdots,s_L(t)]^{\mathrm{T}}$, when $\theta_l=\theta_k$, $\boldsymbol{S}$ is a non-zero value. The DOA estimation model expressed by sparse signals is

$$\boldsymbol{X}=\boldsymbol{BS}+\boldsymbol{N}.$$

Using the spatial sparsity of the incident signal, the signal $\boldsymbol{S}$ is reconstructed by the processed received data $\boldsymbol{X}$ and the over-complete dictionary set $\boldsymbol{B}$. In this paper, the optimization problem is constructed by using the residual sum of squares:

$$\min\sum_{i=1}^{l}|\boldsymbol{X}-\boldsymbol{BS}|^2+\gamma\|\boldsymbol{WS}\|_0 \tag{13}$$

where γ is a regularization factor; $\boldsymbol{W}$ is the weighted matrix; $\|\boldsymbol{S}\|_1$ is penalty function, which is a function with non-zero amplitude value in constraint $\boldsymbol{S}$, thus limiting the sparsity of the model. The residual sum of squares between the actual value and the fitting value is used as the standard to measure the accuracy of reconstruction, and the l_0- norm of the signal is added as the constraint condition of signal sparsity, and the objective function and constraint function are combined to construct an optimization expression to solve it.

l_0-norm is the number of non-zero elements in the vector, and its numerical solution is non-deteministic polynomial (NP)-hard. Dohono et al. [20] proved that l_1-norm will lead to a sparse solution under certain conditions, so l_1-norm is the optimal convex approximation of l_0-norm. Using l_1-norm instead of l_0-norm to solve the optimization problem reduce the complexity of l_0-norm [31−33].

When the residual ϵ obeys normal distribution, the optimal parameter estimation value appears when the maximum likelihood function of the sample reaches the maximum value, then

$$\min\sum_{i=1}^{l}(\boldsymbol{X}-\boldsymbol{BS})^2.$$

That indicates the parameter estimation is optimal when the residual sum of squares is minimum.

In order to make the amplitude of sparse signal equally binding on the reconstructed signal and improve the recovery accuracy of the sparse reconstruction algorithm, weighted constraints are applied to it. The position with the illuminator is multiplied by a smaller weight, and the position without the illuminator is multiplied by a larger weight to enhance the sparseness of the results. Using the recursive noise subspace, the weighted matrix is constructed as

$$\boldsymbol{w}_i = \frac{\boldsymbol{a}(\theta)^{\mathrm{T}} \boldsymbol{R}_n \boldsymbol{a}(\theta)}{\left\| \boldsymbol{a}(\theta)^{\mathrm{T}} \boldsymbol{R}_n \boldsymbol{a}(\theta) \right\|} \tag{14}$$

where $\boldsymbol{R}_n = \boldsymbol{I}_M - \boldsymbol{U}_s \boldsymbol{U}_s^{\mathrm{H}}$.

Substituting it into the signal subspace after compressed observation to obtain the following sparse reconstruction expression:

$$\boldsymbol{U}_{ss} = \boldsymbol{B}_1 \boldsymbol{S} + \boldsymbol{N}. \tag{15}$$

The new sensing matrix is

$$\boldsymbol{B}_1 = \boldsymbol{\Phi}\boldsymbol{\Psi} \in \mathbf{C}^{p \times l}. \tag{16}$$

The objective function of optimization is

$$\min \sum_{i=1}^{l} |\boldsymbol{U}_{ss} - \boldsymbol{B}_1 \boldsymbol{S}|^2 + \gamma \|\boldsymbol{W}\boldsymbol{S}\|_1. \tag{17}$$

Solving this convex optimization problem by second-order cone optimization:

$$\begin{gathered} \min p + \gamma q \\ \text{s.t.} \begin{cases} \displaystyle\sum_{i=1}^{l} |\boldsymbol{U}_{ss} - \boldsymbol{B}_1 \boldsymbol{S}|^2 \leqslant p \\ \gamma \|\boldsymbol{W}\boldsymbol{S}\|_1 \leqslant q \end{cases}. \end{gathered} \tag{18}$$

The above formula can be solved to get the DOA estimation value using software package.

The algorithm flow is presented in Algorithm 1.

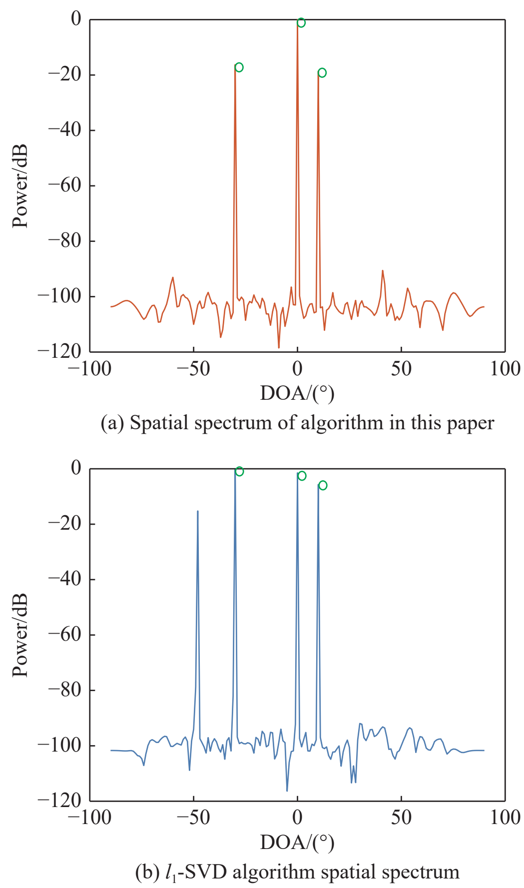
Algorithm 1 The proposed DOA estimation algorithm

1. The array received data $\boldsymbol{X}(t)$ is observed, flow pattern matrix $\boldsymbol{A}(\theta)$
2. Selected MSWF forward recursive initial value: $\boldsymbol{X}_0 = \boldsymbol{X}, \boldsymbol{d}_0 = \mathrm{mean}(\boldsymbol{X}_0)$
3. Carry out K times forward recursion to obtain Krylov subspace $\boldsymbol{T}_K = [h_1, h_2, \cdots, h_K]$, which is equivalent to $\boldsymbol{U}_s$.
4. Determine the weighted matrix: $\boldsymbol{w}_i = \dfrac{\boldsymbol{a}(\theta)^{\mathrm{T}} \boldsymbol{R}_n \boldsymbol{a}(\theta)}{\boldsymbol{a}(\theta)^{\mathrm{T}} \boldsymbol{R}_n \boldsymbol{a}(\theta)}$, $\boldsymbol{R}_n = \boldsymbol{I}_M - \boldsymbol{U}_s \boldsymbol{U}_s^{\mathrm{H}}$
5. Construct the measurement matrix $\boldsymbol{\Phi}$, reduce the dimension of $\boldsymbol{U}_s$ by compressive observation, and get $\boldsymbol{U}_{ss}$ and $\boldsymbol{B}_1$.
6. Solving the weighted l_1-norm convex optimization problem to get the DOA estimation: $\min(\sum_{i=1}^{l} |\boldsymbol{U}_{ss} - \boldsymbol{B}_1 \boldsymbol{S}|^2 + \gamma \|\boldsymbol{W}\boldsymbol{S}\|_1)$

3. Simulation results

Experiment 1 Comparison of the spatial spectral estimation performance of the proposed algorithm and the l_1-SVD algorithm under the large number of array element.

Assuming that the number of ULA elements M=64, the number of sources K=2, the number of snapshots F_s=200, and the illuminator DOA parameters are $\boldsymbol{\theta} = [-30, 0, 10]$ and SNR=−10 dB, the simulation results are shown in Fig. 3.



(a) Spatial spectrum of algorithm in this paper

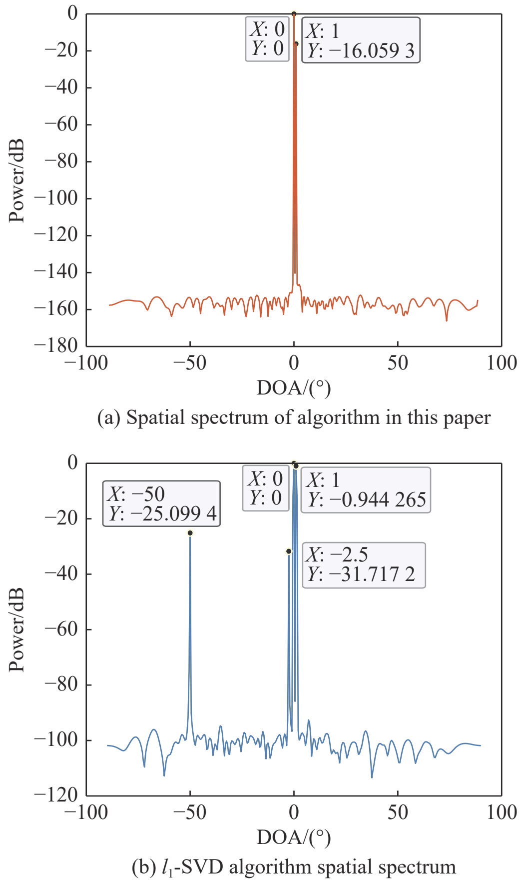
(b) l_1-SVD algorithm spatial spectrum

Fig. 3 Comparison of two algorithms for spatial spectrum estimation

As shown in Fig. 3, under the condition of low SNR, the proposed algorithm improves the accuracy of the algorithm due to the weighted matrix and uses the residual sum of squares as the optimized objective function, thus it can accurately distinguish two similar signals and effectively suppress the formation of spurious peaks. the l_1-SVD algorithm is prone to pseudo-peaks and inaccurate estimation results under low SNR. Therefore, the spatial spectrum estimation performance of the proposed algorithm performs better under certain conditions.

Experiment 2 Comparison of the spatial spectrum estimation performance of the proposed algorithm and the l_1-SVD algorithm under high SNR.

Assuming that the number of ULA elements M=64, the number of sources K=2, the number of snapshots F_s=20, the illuminator DOA parameters are $\boldsymbol{\theta}=[0,1]$ and SNR=10 dB, the simulation results are shown in Fig. 4.



(a) Spatial spectrum of algorithm in this paper

(b) l_1-SVD algorithm spatial spectrum

Fig. 4 Comparison of two algorithms for spatial spectrum estimation

As shown in Fig. 4, the proposed algorithm can distinguish the signals within a beam width at high SNR, while the l_1-SVD algorithm is prone to producing spurious peaks at high SNR and small angular intervals, and the estimation results are inaccurate.

Experiment 3 Comparison of the spatial spectrum estimation performance of the proposed algorithm, the l_1-SVD algorithm, the OMP algorithm and the MUSIC algorithm under low SNR.

Assuming that the number of ULA elements M=64, the number of sources K=5, the number of snapshots F_s=80, the illuminator DOA parameters are $\boldsymbol{\theta}=[-10°,0°,5°,30°,60°]$, and the SNR=−10 dB, the simulation results are shown in Fig. 5.

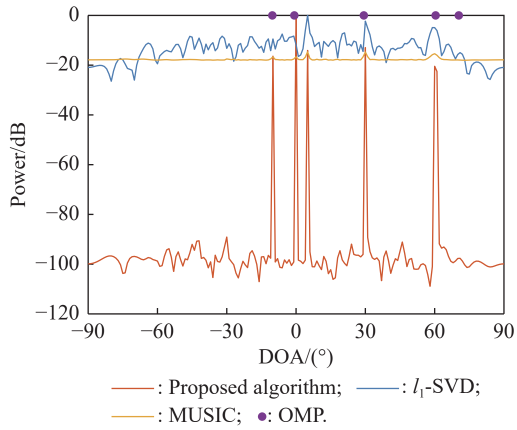


Fig. 5 Comparison of four algorithms for spatial spectrum estimation

It can be seen that the proposed algorithm can correctly distinguish five signals with narrow beam width and good resolution. In contrast, the l_1-SVD algorithm, the OMP algorithm and the MUSIC algorithm cannot correctly distinguish five signals when the number of sources is large and the SNR is low, and the estimation is inaccurate.

Experiment 4 Variation curve of ULA DOA estimation performance with SNR.

DOA estimation is performed in the simulated ULA using 64 array elements for two sources, and 100 Monte Carlo experiments are performed to compare the root mean square error of SNR at −15 dB to 15 dB, using the proposed algorithm, the l_1-SVD algorithm, the OMP algorithm, and the MUSIC algorithm, respectively. The simulation results are shown in Fig. 6.

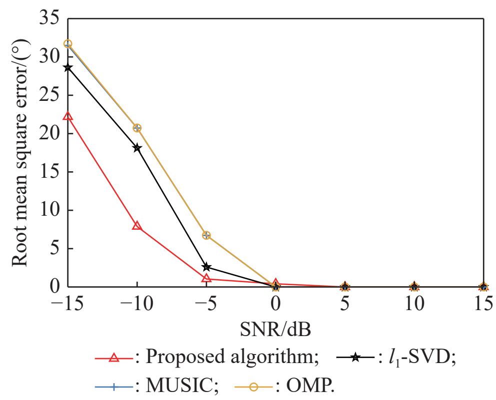


Fig. 6 Comparison of root mean square error of four algorithms

As can be seen from Fig. 6, the proposed algorithm introduces a weighted matrix and uses the residual sum of squares as the optimization objective function, and the root mean square error is smaller than the other three algorithms when the SNR is low, so its estimation accuracy of DOA is higher than that of several other algorithms.

Experiment 5 Comparing the computational complexity of the proposed algorithm, the l_1-SVD algorithm, and the MUSIC algorithm.

Assuming a ULA with M array elements, K sources, N snapshots, and a sampling rate of P for compressed observations, compare the computational complexity of several algorithms, as shown in Table 1. The conventional subspace-like algorithm requires the estimation of covariance matrix and eigenvalue decomposition, and its computational complexity is $O(NM^2+M^3)$. The main computational complexity of the l_1-SVD algorithm is $O(N^3)+O\left((N\times K)^3\right)$ during the solution of the SVD and the second-order cone optimization. The proposed algorithm requires K forward MSWF recursions, and the weighted process is negligible compared with the computational complexity of second-order cone optimization, which is $O(KMN)+O\left((P\times K)^3\right)$. Therefore, in the case of large-scale arrays, the computational complexity of the proposed algorithm is related to the one-time square of the number of array elements, while the l_1-SVD algorithm is proportional to the cube of the number of array elements. The computational complexity of the proposed algorithm is better than that of the l_1-SVD algorithm, and the computational complexity is lower than that of the subspace algorithm in large-scale arrays. For example, when the number of array elements is 64 and the number of snapshots is 80, the computational complexity of the MUSIC algorithm reaches the order of 10^5, the computational complexity of the l_1-SVD algorithm is 10^5, and the computational complexity of the proposed algorithm is 10^3. The computational speed of the proposed algorithm is improved when the number of array elements is larger.

Table 1 Comparison of computational complexity of three algorithms

Algorithm	Algorithm complexity
MUSIC	$O\left(NM^2+M^3\right)$
l_1-SVD	$O\left(N^3\right)+O\left((N\times K)^3\right)$
The proposed algorithm	$O(KMN)+O\left((P\times K)^3\right)$

4. Conclusions

In this paper, the proposed DOA estimation method for high-dimensional signals based on weighted l_1-norm combines the MSWF algorithm with the sparse reconstruction in compressed sensing, obtains the signal subspace without calculating the covariance matrix and the inverse of the matrix, and further reduces the dimensionality through the measurement matrix. Since the weighted matrix is used to constrain the signal amplitude to be closer to the l_1-norm model, the proposed algorithm is more accurate in estimating DOA and can effectively suppress the pseudo-peaks at the same time. The use of the residual sum of squares as the objective function of the optimization problem allows it to be closer to the maximum likelihood estimation of the sample under non-ideal conditions, and performs well under conditions such as large number of array elements and low SNR, with higher resolution, smaller root mean square error and high estimation accuracy. The algorithm proposed in this paper has higher resolution and better robustness compared with the traditional l_1-SVD algorithm, the OMP algorithm and the MUSIC algorithm under the conditions of lower SNR and smaller number of snapshots. Simulations show that the Krylov iterative subspace algorithm is introduced in combination with the sparse reconstruction method to solve the DOA estimation problem of high-dimensional large-scale signals.

As one of the meaningful further consideration, although the proposed algorithm has some advantages comparing with the classical DOA estimation algorithm under low SNR, however, when the number of illuminators exceeds the maximum freedom of the array, the resolution performance of the illuminators deteriorates. In order to solve the problem of multi-dimensional parameter estimation, tensor calculation can be used to expand the algorithm in this paper and establish the tensor model of the array.

With the future wide application of large-scale arrays in radar, communication, astronomy and other fields, the dimensionality reduction method in this paper can be further extended to two-dimensional DOA estimation, thus reducing the computational complexity and improving the accuracy of DOA estimation.

References

[1] ZHANG X F, WANG F, XU D Z. Theory and application of array signal processing. Beijing: National Defense Industry Press, 2010. (in Chinese)

[2] TANG J K, LIU Z, XIE R, et al. Optimal design method for sparse array of MIMO radar. Systems Engineering and Electronics, 2022, 44(12): 3661–3666. (in Chinese)

[3] XIE H. Research on adaptive signal processing method of high-dimensional small sample array. Xi'an: Xidian University, 2015. (in Chinese)

[4] PAN L, ZHAO H F. Direction of arrival estimation for high frequency source in the presence of three-dimensional sensor position errors. Proc. of the IEEE 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, 2020: 489–494.

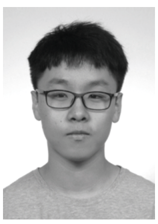
[5] KRIM H. Two decades of array signal processing research. IEEE Signal Processing Magazine, 1994, 13(4): 67–94.

[6] ULRYCH T J, BISHOP T N. Maximum entropy spectral

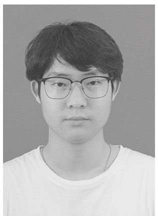
analysis and autoregressive decomposition. Reviews of Geophysics, 1975, 13(1): 183–200.
[7] CAPON J. High-resolution frequency-wavenumber spectrum analysis. Proceedings of the IEEE, 1969, 57(8): 1408–1418.
[8] SCHMIDT R, SCHMIDT R O. Multiple emitter location and signal parameter estimation. IEEE Trans. on Antennas & Propagation, 1986, 34(3): 276–280.
[9] ROY R, KAILATH T. ESPRIT - estimation of signal parameters via rotational invariance techniques. IEEE Trans. on Acoustics Speech and Signal Processing, 1989, 37(7): 984–995.
[10] BONOTTO M, BETTINI P, CENEDESE A. Model-order reduction of large-scale state-space models in fusion machines via Krylov methods. IEEE Trans. on Magnetics, 2017, 53(6): 7204304.
[11] WEIPPERT M E, HIEMSTRA J D, GOLDSTEIN J S, et al. Insights from the relationship between the multistage Wiener filter and the method of conjugate gradients. Proc. of the IEEE Sensor Array & Multichannel Signal Processing Workshop, 2002: 388–392.
[12] GOLDSTEIN J S, REED I S, SCHARF L L. A multistage representation of the Wiener filter based on orthogonal projections. IEEE Trans. on Information Theory, 1998, 44(7): 2943–2959.
[13] CANDES E J, TAO T. Decoding by linear programming. IEEE Trans. on Information Theory, 2005, 51(12): 4203–4215.
[14] CANDES E J, ROMBERG J, TAO T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans. on Information Theory, 2006, 52(2): 489–509.
[15] HAN P, XU H T, CUI W J, et al. A novel DOA estimation method for uncorrelated and coherent signals via compressed sensing in sparse arrays. Electronics Letters, 2021, 57(25): 995–997.
[16] JI W P, HONG W, LEE J H. Direction-of-departure and direction-of-arrival estimation algorithm based on compressive sensing: data fitting. Remote Sensing, 2020, 12(17): 2773–2797.
[17] WANG S X, ZHAO Y, LBRAHIM L, et al. Joint 2D DOA and Doppler frequency estimation for L-shaped array using compressive sensing. Journal of Systems Engineering and Electronics, 2020, 31(1): 32–40.
[18] GKILLAS A, AMPELIOTIS D, BERBERIDIS K. Connections between deep equilibrium and sparse representation models with application to hyperspectral image denoising. IEEE Trans. on Image Processing, 2023, 32: 1513–1528.
[19] CHENG H, LIU Z C, YANG L, et al. Sparse representation and learning in visual recognition: theory and applications. Signal Processing, 2013, 93(6): 1408–1425.
[20] DONOHO D L. Compressed sensing. IEEE Trans. on Information Theory, 2006, 52(4): 1289–1306.
[21] MALIOUTOV D, CETIN M, WILLSKY A S. A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE Trans. on Signal Processing, 2005, 53(8): 3010–3022.
[22] CONG J Y, WANG X P, WAN L T, et al. Neural network-aided sparse convex optimization algorithm for fast DOA estimation. Transactions of the Institute of Measurement and Control, 2022, 44(8): 1649–1655.
[23] XU F J, LIU A F, MO S Q, et al. DOA estimation method using sparse representation with orthogonal projection. Journal of Beijing Institute of Technology, 2021, 30(4): 397–402.
[24] QIU W, BAO C C. DOA estimation method based on the sparsity of array covariance matrix. Journal of National University of Defense Technology, 2020, 42(5): 37–45. (in Chinese)
[25] ZUO M, XIE S G, ZHANG X, et al. DOA estimation based on weighted l_1-norm sparse representation for low SNR scenarios. Sensors. 2021, 21(13): 4614–4627.
[26] XU X, WEI X H, YE Z F. DOA estimation based on sparse signal recovery utilizing weighted l_1-norm penalty. IEEE Signal Processing Letters, 2012, 19(3): 155–158.
[27] HUO H Y. Stable recovery of weighted sparse signals from phaseless measurements via weighted l_1 minimization. Mathematical Methods in the Applied Sciences, 2022, 45(9): 4929–4937.
[28] WANG Y, LI T, XIANG J H. A DOA estimation method based on weighted improved smoothing l_0-norm. Applied Science and Technology, 2022, 49(4): 38–43. (in Chinese)
[29] TROPP J A, GILBERT A C. Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans. on Information Theory, 2007, 53(12): 4655–4666.
[30] ZHANG Z H, LIU Y P, LIU J N, et al. AMP-Net: denoising-based deep unfolding for compressive image sensing. IEEE Trans. on Image Processing, 2020, 30: 1487–1500.
[31] BOYD S, VANDENBERGHE L, FAYBUSOVISH L. Convex Optimization. IEEE Trans. on Automatic Control, 2006, 51(11): 1859–1859.
[32] WANG S X, ZHAO Y, IBRAHIM L, et al. Joint 2D DOA and Doppler frequency estimation for L-shaped array using compressive sensing. Journal of Systems Engineering and Electronics, 2020, 31(1): 28–36.
[33] LIU Q, LI X P. Efficient low-rank matrix factorization based on $l_{1,\epsilon}$-norm for online background subtraction. IEEE Trans. on Circuits and Systems for Video Technology, 2021, 32(7): 4900–4904.

Biographies

YANG Zeqi was born in 2000. She received her B.S. degree from Central South University in 2022. She is working toward her Ph.D. degree in signal and information processing in the University of Chinese Academy of Sciences. Her research interest is radar signal processing.
E-mail: yangzeqi22@mails.ucas.ac.cn



LIU Yiheng was born in 1999. He received his B.S. degree from the University of Chinese Academy of Sciences in 2021. He is working toward his Ph.D. degree in signal and information processing in the University of Chinese Academy of Sciences. His research interest is signal processing of passive radar.
E-mail: liuyiheng17@mails.ucas.ac.cn



ZHANG Hua was born in 1998. He received his B.S. degree from Nanjing University of Science and Technology in 2021. He is working toward his Ph.D. degree in signal and information processing in University of Chinese Academy of Sciences. His research interests are radar systems and signal processing.
E-mail: zhanghua211@mails.ucas.ac.cn

MA Shuai was born in 2000. He received his B.S. degree from Zhengzhou University in 2022. He is working toward his M.E. degree in signal and information processing in the University of Chinese Academy of Sciences. His research interest is radar signal processing.
E-mail: mashuai22@mails.ucas.ac.cn

CHANG Kai was born in 1987. He received his Ph.D. degree in precision test and measurement from Beijing Institute of Technology in 2016. He works in Northern Institute of Electronic Equipment. His research interests include information control of multi-unmanned aerial vehicle (UAV) systems and adaptive control.
E-mail: kerkaichance@gmail.com

LIU Ning was born in 1965. She received her M.S. degree from the University of Electronic Science and Technology of China. She then joined the Northern Institute of Electronic Equipment, Beijing, China. Her research interests include information control, stochastic geometry, and wireless positioning technology.

LYU Xiaode was born in 1969. He received his B.S. and Ph.D. degrees in electrical engineering from Hebei University of Technology, Tianjin, and Xi'an JiaoTong University, Xi'an, in 1991 and 1997, respectively. In 1998, he engaged in postdoctoral research in the Department of Electronic Engineering of Beijing Institute of Technology. In 2000, he joined the Aerospace Information Research Institute (AIR, formerly the Institute of Electronics), Chinese Academy of Sciences (CAS). Since 2005, he has been a Principal Investigator (PI) in AIRCAS. His research interests include radar signal processing, airborne/spaceborne synthetic aperture radar (SAR), and applied electromagnetism.
E-mail: lvxd@aircas.ac.cn