Sparse Recovery Assisted DOA Estimation Utilizing Sparse Bayesian Learning

Min Huang and Lei Huang

College of Information Engineering, Shenzhen University

Objectives and Contributions

This paper proposes a sparse recovery assisted direction-of-arrival (SR-DOA) estimator.

· The DOA estimation is formulated as a sparse nonnegative least squares problem.

· The SR-DOA method is able to suppress the noise but at the expense of a few degrees-of-freedom, and mitigate the sampling errors by exploiting its asymptotic distribution.

· The spare Bayesian learning with nonnegative Laplace prior is utilized to yield the DOA estimation.

· Numerical results show that the proposed SR-DOA algorithm outperforms the exsiting methods in terms of the estimation accuracy.

Problem formulation

Consider K uncorrelated narrowband far-field signals, $s_k(t)$, $k=1,2,\cdots,K$, impinging on a linear sparse array which consists of M omnidirectional sensors located at $[0,\ d_1,\ \cdots, d_{M-1}]$, where d_m represents the distance between the $(m+1)$-th sensor and the first sensor. Then, the array output vector $\mathbf{x}(t)$ of T snapshots can be expressed as

$$\mathbf{x}(t) = \mathbf{A}\boldsymbol{s}(t) + \boldsymbol{n}(t),\ t=1,2,\cdots,T \quad (1)$$

where $\boldsymbol{s}(t)=[s_1(t)], s_2(t),\cdots, s_K(t)]^T$ and $\boldsymbol{n}(t)$ denote the source signal and additive Gaussian noise, respectively, $\boldsymbol{A}$ consists of K steering vectors. Note that the DOA of the k-th source signal is distributed in the range of $(-90°, 90)$. Thus, by invoking all the possible DOAs, $\mathbf{x}(t)$ in (1) can be written in a high-resolution and sparse representation as

$$\mathbf{x}(t) = \bar{\mathbf{A}}\bar{\boldsymbol{s}}(t) + \boldsymbol{n}(t),\ t=1,2,\cdots,T \quad (2)$$

where $\boldsymbol{A} = [\boldsymbol{a}(\bar{\theta}_1), \boldsymbol{a}(\bar{\theta}_2),\cdots,\boldsymbol{a}(\bar{\theta}_{\bar{K}})]$ and the set of $\bar{\boldsymbol{\theta}} = \{\bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_{\bar{K}}\}$ gives a sampling grid of all possible DOAs, while $\bar{\boldsymbol{s}}(t) = [\bar{s}_1(t)], \bar{s}_2(t),\cdots,\bar{s}_K(t)]^T$ with $\bar{s}_k(t)$ being the possible source signal. In general, we have $\bar{K} \gg K$. Therefore, $\bar{\boldsymbol{s}}(t)$ is a sparse vector, whose i-th row is nonzero and equals to the corresponding row of $\boldsymbol{s}(t)$ in (1). Consequently, the problem of DOA estimation based on (1) is equivalent to identifying the positions of the nonzero rows of $\mathbf{x}(t)$ in (2).

sparse nonnegative least squares (S-NNLS) modeling

To begin with, the sample covariance matrix of $\mathbf{x}(t)$ of (2) can be derived as

$$\hat{\boldsymbol{R}} = \bar{\mathbf{A}}\boldsymbol{R}_s\bar{\mathbf{A}}^H + \boldsymbol{R}_n + \boldsymbol{E} \quad (3)$$

where $\boldsymbol{R}_s = \mathbb{E}[\bar{\boldsymbol{s}}(t)\bar{\boldsymbol{s}}(t)^H] = \text{diag}\{\sigma_1^2,\cdots,\sigma_{\bar{K}}^2\}$ with $\sigma_k^2 = \mathbb{E}[\bar{s}_k(t)\bar{s}_k(t)^H]$ being the power received from the k-th source singal, $\boldsymbol{R}_n = \text{diag}\{\sigma^2,\cdots,\sigma^2\}$ with σ^2 being the variance of noise, while $\boldsymbol{E}$ reflects the error between the covariance matrix of $\mathbf{x}(t)$ given in (2), which is $\bar{\mathbf{A}}\boldsymbol{R}_s\bar{\mathbf{A}}^H + \boldsymbol{R}_n$, and its sample covariance matrix $\hat{\boldsymbol{R}}$ of (3). Let us vectorize (3), yielding an M^2-length vector, which is

$$\boldsymbol{y} \triangleq \text{vec}\{\hat{\boldsymbol{R}}\} = \boldsymbol{V}\boldsymbol{\varsigma} + \boldsymbol{\rho} + \boldsymbol{\xi} \quad (4)$$

where $\boldsymbol{V} \triangleq \bar{\boldsymbol{A}}^* \odot \bar{\boldsymbol{A}}$, $\boldsymbol{\varsigma} \triangleq [\sigma_1^2,\cdots \sigma_{\bar{K}}^2]^T$, $\boldsymbol{\rho} \triangleq \text{vec}(\boldsymbol{R}_n) = [\sigma^2\boldsymbol{e}_1^T,\cdots,\sigma^2\boldsymbol{e}_M^T]^T$ and $\boldsymbol{\xi} \triangleq \text{vec}(\boldsymbol{E})$. Here, $(\cdot)^*$, $\odot$ and $\boldsymbol{e}_i$ denote, respectively, the complex conjugate, Khatri-Rao product, and the i-th column of the identity matrix $\boldsymbol{I}_Q$. Based on (4), our DOA estimation problem is converted to a problem of identifying the locations of nonzero elements in $\boldsymbol{\varsigma}$.

Then, we convert (4) into its real form, which can be expressed as

$$\hat{\boldsymbol{y}} = \hat{\boldsymbol{V}}\boldsymbol{\varsigma} + \hat{\boldsymbol{\rho}} + \hat{\boldsymbol{\xi}} \quad (5)$$

where $\hat{\boldsymbol{y}} = [\Re\{\boldsymbol{y}\}^T, \Im\{\boldsymbol{y}\}^T]^T$, $\hat{\boldsymbol{V}} = [\Re\{\boldsymbol{V}\}^T, \Im\{\boldsymbol{V}\}^T]^T$, $\hat{\boldsymbol{\rho}} = [\boldsymbol{\rho}^T, \boldsymbol{0}^T]^T$ and $\hat{\boldsymbol{\xi}} = [\Re\{\boldsymbol{\xi}\}^T, \Im\{\boldsymbol{\xi}\}^T]^T$. Here,$\boldsymbol{0}$ is an $M^2 \times 1$ zero vector.

Subsequently, the cancellation of the noise resultant components in (5) can be implemented by premultiplying a selection matrix $\boldsymbol{J}$ satisfying $\boldsymbol{J}\hat{\boldsymbol{\rho}} = \boldsymbol{0}$ on $\hat{\boldsymbol{y}}$, yielding

$$\boldsymbol{u} \triangleq \boldsymbol{J}\hat{\boldsymbol{y}} = \boldsymbol{J}\hat{\boldsymbol{V}}\boldsymbol{\varsigma} + \boldsymbol{J}\hat{\boldsymbol{\xi}}. \quad (6)$$

Note that, according to the structure of $\boldsymbol{e}_i$, $\boldsymbol{J}$ is constructed from the identity matrix $\boldsymbol{I}_{2M^2}$ by removing its $\{0\times M+1, 1\times M+2, \cdots, (M-1)\times M+M\}$ rows.

Finally, we may whiten $\boldsymbol{J}\hat{\boldsymbol{\xi}}$ through multiplying $\boldsymbol{u}$ of (6) by $\boldsymbol{G}^{-\frac{1}{2}}$, yielding an S-NNLS model, i.e.,

$$\hat{\boldsymbol{u}} \triangleq \boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{u} = \boldsymbol{\Psi}\boldsymbol{\varsigma} + \boldsymbol{\nu} \quad (7)$$

where $\boldsymbol{\Psi} \triangleq \boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{J}\hat{\boldsymbol{V}}$ and $\boldsymbol{\nu} \sim \mathcal{N}(0, \boldsymbol{I}_{2M^2-M})$ is now a white Gaussian noise vector.

Sparse Bayesian learning with nonnegative Laplace prior

For the model (7), we have the Gaussian likelihood function as

$$p(\hat{\boldsymbol{u}}|\boldsymbol{\varsigma}) \sim \mathcal{N}(\boldsymbol{\Psi}\boldsymbol{\varsigma}, \boldsymbol{I}_{2M^2-M}). \quad (8)$$

In addition, the prior for $\boldsymbol{\varsigma}$ can be considered as a nonnegative Laplace distribution, which is

$$p(\boldsymbol{\varsigma}|\lambda) = \int p(\boldsymbol{\varsigma}|\boldsymbol{\gamma})p(\boldsymbol{\gamma}|\lambda)d\boldsymbol{\gamma} = \sqrt{\lambda}^{\bar{K}} e^{-\sqrt{\lambda}\sum_{k=1}^{\bar{K}}\varsigma_k} \quad (9)$$

where $p(\boldsymbol{\varsigma}|\boldsymbol{\gamma}) = \prod_{k=1}^{\bar{K}} \mathcal{N}_+(\boldsymbol{\varsigma}_k|0,\boldsymbol{\gamma}_k)$ with $\mathcal{N}_+(\boldsymbol{\varsigma}_k|0,\boldsymbol{\gamma}_k) = 2\mathcal{N}(\boldsymbol{\varsigma}_k|0,\boldsymbol{\gamma}_k)$, while $p(\boldsymbol{\gamma}|\lambda) = \prod_{k=1}^{\bar{K}} p(\boldsymbol{\gamma}_k|\lambda)$ with $p(\boldsymbol{\gamma}_k|\lambda) = \frac{\lambda}{2}e^{-\frac{\lambda\gamma}{2}}$, the hyperprior of λ in (9) is assumed to follow Gamma distribution, i.e.,

$$p(\lambda) = \frac{b^a}{\Gamma(a)}\lambda^{a-1}e^{-b\lambda} \quad (10)$$

Based on the Bayes rule, we can estimate $\boldsymbol{\varsigma}$ by maximizing its posterior density, namely,

$$\hat{\boldsymbol{\varsigma}} \propto \arg\max_{\boldsymbol{\varsigma}} \mathcal{N}_+(\boldsymbol{\varsigma}|\boldsymbol{\mu},\boldsymbol{\Sigma}) = \boldsymbol{\mu} \quad (11)$$

where $\boldsymbol{\mu} = \boldsymbol{\Sigma}\boldsymbol{\Psi}^T\hat{\boldsymbol{u}}$ and $\boldsymbol{\Sigma} = \left(\boldsymbol{\Psi}^T\boldsymbol{\Psi} + \boldsymbol{\Lambda}^{-1}\right)^{-1}$ with $\boldsymbol{\Lambda} = \text{diag}\{\boldsymbol{\gamma}\}$. From (11), we readily find that $\hat{\boldsymbol{\varsigma}}$ is a function of $\boldsymbol{\gamma}$. Hence, once $\boldsymbol{\gamma}$ is estimated, the Maximum-A-Posteriori (MAP) estimate of $\hat{\boldsymbol{\varsigma}}$ can be determined by (11).

The $\boldsymbol{\gamma}$ and its associated hyperparameter λ can be estimated by maximizing their posterior density, namely

$$\begin{aligned} \hat{\boldsymbol{\gamma}}, \hat{\lambda} &= \arg\max_{\boldsymbol{\gamma},\lambda} \mathbb{E}[\text{log}p(\boldsymbol{\varsigma},\boldsymbol{\gamma},\lambda|\hat{\boldsymbol{u}})] \\ &\propto \arg\max_{\boldsymbol{\gamma},\lambda} \mathbb{E}[\text{log}p(\boldsymbol{\varsigma},\boldsymbol{\gamma},\lambda,\hat{\boldsymbol{u}})] \\ &\propto \arg\max_{\boldsymbol{\gamma},\lambda} \mathbb{E}[p(\hat{\boldsymbol{u}}|\boldsymbol{\varsigma})p(\boldsymbol{\varsigma}|\boldsymbol{\gamma})p(\boldsymbol{\gamma}|\lambda)p(\lambda)]. \end{aligned} \quad (12)$$

Hence, when λ is given, $\hat{\boldsymbol{\gamma}}$ can be computed by

$$\hat{\gamma}_k = -\frac{1}{2\lambda} + \sqrt{\frac{1}{4\lambda^2} + \frac{w_k}{\lambda}} \quad (13)$$

where w_k is the second-order moment of $\boldsymbol{\varsigma}_k$. Similarly, when $\boldsymbol{\gamma}$ is given, λ can be computed by

$$\hat{\lambda} = \frac{\bar{K}-1+c}{\sum_{k=1}^{\bar{K}} \boldsymbol{\gamma}_k/2 + c}. \quad (14)$$

From (13) and (14), it is easy to see that $\hat{\boldsymbol{\gamma}}$ and λ are the functions of $\{\hat{\boldsymbol{\varsigma}}, \lambda\}$ and $\boldsymbol{\gamma}$, respectively. Recalling that $\hat{\boldsymbol{\varsigma}}$ is a function of $\boldsymbol{\gamma}$, $\hat{\boldsymbol{\varsigma}}$ can be determined in an iterative way.

Simulation results

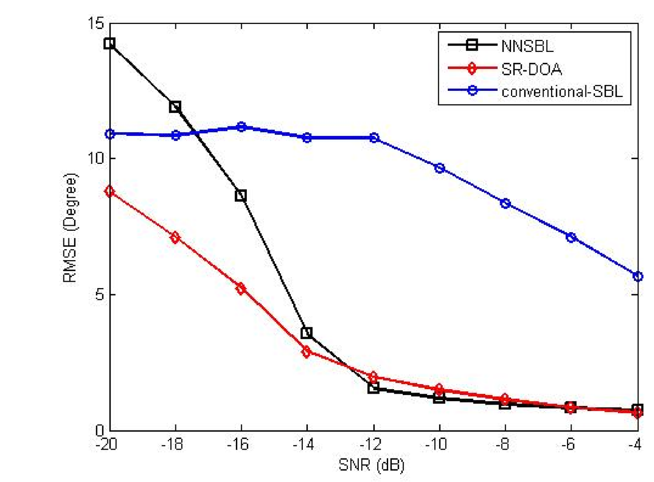

Figure 1: RMSE versus SNR performance for different DOA estimators at T=200.

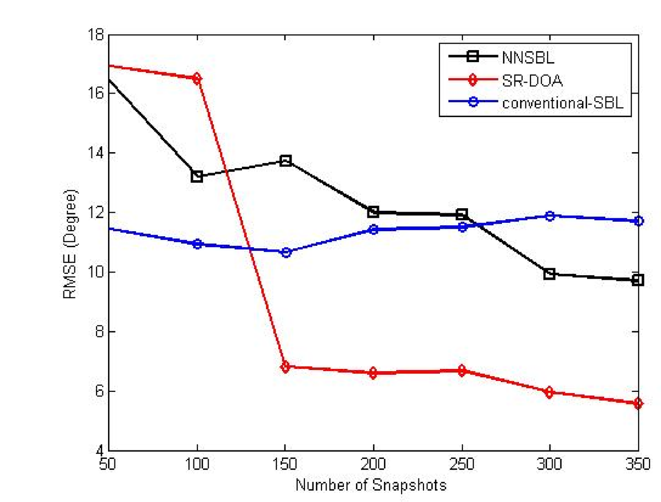

Figure 2: MSE versus number of snapshots for different DOA estimators at $\text{SNR} = -18$ dB.

· From Fig. 1, We can explicitly observe that our proposed SR-DOA method outperforms the other estimators when the SNR is less than -12.5 dB.

· In addition, when the SNR is larger than -12.5 dB, the estimation performance of our proposed SR-DOA algorithm is slightly worse than that of the NNSBL algorithm.

· From Fig. 2, as the number of snapshots increases, the RMSE performance of our proposed SR-DOA and NNSBL algorithms improves, while the performance of the conventional SBL algorithm is almost unchanged.

· Furthermore, from Fig. 2, we can find that the RMSE of the proposed SR-DOA algorithm is relatively low, as long as the number of snapshots is not less than 150.

Contact Information

- Email: mhuang@szu.edu.cn;
- lhuang@szu.edu.cn

PLACEHOLDER PLACEHOLDER