Direction of Arrival Estimation via Multiclass Classification

Hua Bai, *Student Member, IEEE*, Ramakrishna Janaswamy, *Fellow, IEEE*

***Abstract*—In this letter, the angle of impinging sources are estimated via converting it to a classification problem. During the training process, the labels with using k-hot encoding are assigned according to the angle of incoming signal, and the training data is comprised of the received signals from the receiving antenna array in noisy environment. The probability of the angle of test source is predicted with the model trained by the available data and the given labels. Since the number of classes is usually more than 2, the multiclass classification methods are implemented in this letter.**

***Index Terms*—direction of arrival, non-uniform array, k-hot encoding, binary classifier, multiclass classification, complex sources, softmax**

I. Introduction

DIRECTION of arrival (DOA) estimation for linear arrays is a topic of active research. During the past decades, many methods have been proposed and applied to estimate the incoming sources from the standpoint of regression, that the sources are solved by making its associated receiving signal approach to the actual one, then the direction of incoming sources are predicated according to the estimated sources. Since the information regarding the actual sources is limited, the angle grid which includes all possible directions is created with using the regression method, and some specific constraints or strategies such the the sparsity of actual sources in the grid [1] is considered in the estimation procedure in order to improved the estimation accuracy. With the regression method, the estimated sources are figured out at the end, however, in practice, estimating the probability of the direction of incoming sources is still worthwhile. Even though within the sparse Bayesian learning method [2], the variance of estimation result which indicate the uncertainty of the result is calculated along with the estimated mean vector, one parameter which can indicate the probability of incoming sources directly is still necessary for the DoA problem. In this letter, the DoA problem is solved by converting it to a classification problem which has not been studied according to authors' knowledge.

In order to predict the direction of incoming sources, the training data is provided to generate the training model and optimize the parameters of the model. The labels of the training data consist of the angles of incoming sources, since the number of labels, K, is usually larger than 2 for the DoA problem, the DoA problem thus becomes a multiclass classification. For a multiclass classification problem, we can either transform it to a binary classification problem by using K binary classifier (i.e., one vs. rest), or build a linear mapping for the all K classes with the k-hot encoding to map the input vector to K classes. The training data consists of the signal received on the receiver's end with sources incoming from various direction, and the noise is added to the receiver's end to generate the simulation data.

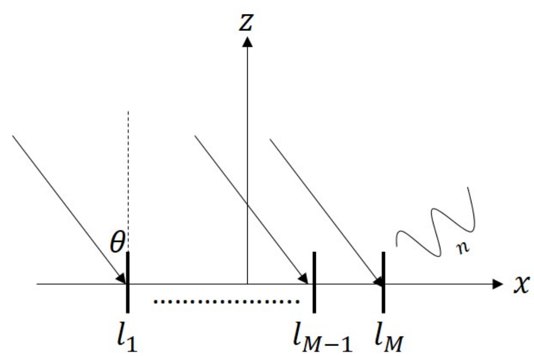

Fig. 1: M-element nonuniform linear array

Once the training model is ready, we can predict the probability of the input for the K directions, then the direction of the source can be estimated via some strategies based on the probabilities instead of the value of estimated sources in the regression way. In this letter, one non-uniform linear array is used as the receiver, since the incoming sources consists of complex signals, the real and imaginary part of the sources are predicted separately.

II. DOA problem formulation

A non-uniform linear array shown in Fig. 1 is comprised of M identical elements with the i-th element located at distance l_i from the origin. The incoming sources which are independent of each other are denoted by $\mathbf{s}_j$, j = 1,2,...,N and the incident angles of the incoming sources are denoted by θ_j, where $\theta_j \in$ [-90°, 90°], j = 1,2,...,N. Independent complex white noise is also present in the system, denoted by n_i; for the far-field sources, the approximate phase variation of received signals for the j-th source at two elements separated by Δl is equal to $k\Delta l \sin\theta_j$ ignoring mutual coupling, $k = \frac{2\pi}{\lambda}$ and λ is the wavelength.

The received signal of the i-th element, y_i, is written as :

$$y_i = \sum_{j=1}^{N} e^{-jkl_i \sin\theta_j} \mathbf{s}_j + n_i, i = 1, 2, ..., M. \tag{1}$$

In matrix form equation (1) is written as :

$$\mathbf{y} = \mathbf{A}\mathbf{s} + \mathbf{n}, \tag{2}$$

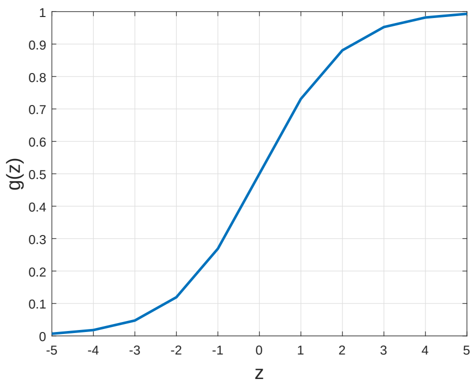

Fig. 2: the logistic function

where $\mathbf{y} = [y_1, y_2 \ldots y_M]^T \in \mathbb{C}^{M\times 1}$, $\mathbf{A} \in \mathbb{C}^{M\times N}$ is the measurement matrix with entries

$$\mathbf{A}(m,n) = e^{-jkl_m \sin\theta_n}, \tag{3}$$

$\boldsymbol{s} = [s_1, s_2 \ldots s_N]^T \in \mathbb{C}^{N\times 1}$ and the noise $\mathbf{n} = [n_1, n_2, \ldots, n_M]^T \in \mathbb{C}^{M\times 1}$ is subject to a circularly symmetric complex normal distribution with zero mean and covariance matrix $\Gamma = \sigma^2\mathbf{I} \in \mathbb{R}^{M\times M}$, with σ^2 *unknown*. The goal is to estimate the direction of sources via the received signal $\mathbf{y}$.

III. The Direction of Complex Source Prediction

A. Binary Classifier

In the binary classification problem, the logistic function which is plotted in Fig. 2 is well known and widely applied:

$$g(z) = \frac{1}{1+e^{-z}}, \tag{4}$$

where z denotes the distance between the input and the classifier, in the 2-D situation with the input being written as $\boldsymbol{x_i} = [x_{i1}, x_{i2}] \in \mathbb{R}^{2\times 1}$, the classier is comprised of a line, and z is defined as :

$$z_i = \alpha_1 x_{i1} + \alpha_2 x_{i2} = \boldsymbol{\alpha^T x_i}, \tag{5}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2] \in \mathbb{R}^{2\times 1}$ denote the two parameters of the model which are *unknwon*. The training data consist of inputs belonging to 2 classes, the label, θ, consists of 0 and 1 is assigned to the specific input according to which class it belongs to, then the probability for the input is defined as :

$$P_i = \begin{cases} P(\theta = 1|\boldsymbol{x_i}, \boldsymbol{\alpha}) = g(\boldsymbol{\alpha^T x_i}) \\ P(\theta = 0|\boldsymbol{x_i}, \boldsymbol{\alpha}) = 1 - g(\boldsymbol{\alpha^T x_i}), \end{cases} \tag{6}$$

for convenience, rewrite (6) as :

$$P(y_i|\boldsymbol{x_i}, \boldsymbol{\alpha}) = (g(\boldsymbol{\alpha^T x_i}))^{\theta_i}(1 - g(\boldsymbol{\alpha^T x_i}))^{1-\theta_i}, \tag{7}$$

assume there are M inputs as the training data which are independent with each other, we have the likelihood of the all inputs :

$$\mathcal{L}(\boldsymbol{x}) = \prod_{i=1}^{M} \left\{(g(\boldsymbol{\alpha^T x_i}))^{\theta_i}(1 - g(\boldsymbol{\alpha^T x_i}))^{1-\theta_i}\right\}, \tag{8}$$

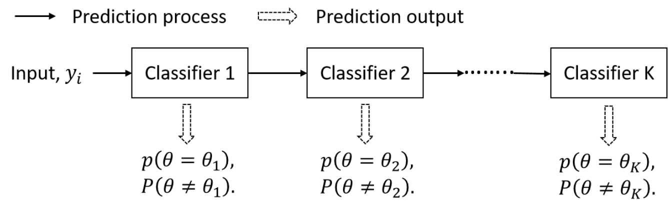

Fig. 3: prediction procedure with K binary classifiers

the log-likelihood is written as :

$$\log\mathcal{L}(\boldsymbol{x}) = \sum_{i=1}^{M}\left\{\theta_i \log(g(\boldsymbol{\alpha^T x_i})) + (1-\theta_i)\log(1 - g(\boldsymbol{\alpha^T x_i}))\right\}, \tag{9}$$

the parameters $\boldsymbol{\alpha}$ can be optimized iteratively by making the log-likelihood approach to its maximum with using the gradient ascent algorithm. When the parameters $\boldsymbol{\alpha}$ is optimized, the probability of test input can be predicated by (6).

B. K Binary Classifier

The labels of the training data in DoA problem are assigned depending on the directions of incoming sources. The number of the labels, K, is usually larger than 2. In order to apply the binary classifier mentioned above to the DoA problem, we can use one binary classifier to predict the probability of sources belong to one direction and outside of this specific direction, therefore K binary classifiers are required in total to predict the probability of the input for the all K directions. During the training procedure, for the i-th classifier, the received signals associated with the sources from $\theta = \theta_i$ are labeled with 1, while the remaining received signals for the sources from other directions are labeled with 0. Since the received signals and incoming sources consist of complex signals, the real and imaginary part are predicted separately for simplicity :

$$\begin{bmatrix}\Re(\mathbf{y})\\ \Im(\mathbf{y})\end{bmatrix} = \begin{bmatrix}\Re(\mathbf{A}) & -\Im(\mathbf{A})\\ \Im(\mathbf{A}) & \Re(\mathbf{A})\end{bmatrix}\begin{bmatrix}\Re(\mathbf{s})\\ \Im(\mathbf{s})\end{bmatrix} + \begin{bmatrix}\Re(\mathbf{n})\\ \Im(\mathbf{n})\end{bmatrix}, \tag{10}$$

The probability of the real part of the received signal is defined as :

$$P_i = \begin{cases} P(\theta|\boldsymbol{x_j}, \boldsymbol{\alpha}) = g(\boldsymbol{\alpha^T x_j}), & \text{if } \theta = \theta_i \\ P(\theta|\boldsymbol{x_j}, \boldsymbol{\alpha}) = 1 - g(\boldsymbol{\alpha^T x_j}), & \text{if } \theta \neq \theta_i, \end{cases} \tag{11}$$

where $\boldsymbol{x_j} \in \mathbb{R}^{M\times 1}$ stand for the real part of the j-th received signal, and $\boldsymbol{\alpha} \in \mathbb{R}^{M\times 1}$ denote the parameters which need to be optimized by the training data. The imaginary part has the similar format as the real part. When the K classifiers are available, the prediction procedure which is illustrated in Fig. 3, we classify the test input with the K classifiers one by one, finally the angle with the highest probability is selected as the direction of incoming source.

C. Softmax Regression

The computational complexity of using the K binary classifiers can be huge especially for a large K, because one sample is used as the training data for K times during the training procedure and the test input needs to be classified K times to obtain the probability. In order to improve the efficiency, the softmax regression can be applied for the DoA problem. Unlike the 2 classes considered in the binary classifier, with softmax regression, the probability of the input belonging to K classes is calculated, the probability for the input is defined as :

$$P_i = \begin{cases} P(\theta = \theta_1 | \boldsymbol{x_i}, \boldsymbol{\alpha}) = \frac{e^{(\boldsymbol{\alpha_1})^T \boldsymbol{x_i}}}{\eta} \\ P(\theta = \theta_2 | \boldsymbol{x_i}, \boldsymbol{\alpha}) = \frac{e^{(\boldsymbol{\alpha_2})^T \boldsymbol{x_i}}}{\eta} \\ \dots \\ P(\theta = \theta_K | \boldsymbol{x_i}, \boldsymbol{\alpha}) = \frac{e^{(\boldsymbol{\alpha_K})^T \boldsymbol{x_i}}}{\eta}, \end{cases} \tag{12}$$

where $\eta = \sum_{l=1}^{K} e^{(\boldsymbol{\alpha_l})^T \boldsymbol{x_i}}$, $\boldsymbol{x_i} \in \mathbb{R}^{M\times 1}$ denotes either the real or imaginary part of the received signal and $\boldsymbol{\alpha} \in \mathbb{R}^{M\times K}$, $\boldsymbol{\alpha}_i \in \mathbb{R}^{M\times 1}, i = 1, 2, ..., K$. When $K = 2$, it can be found (12) is equivalent to (6). The label, $I \in \mathbb{R}^{K\times 1}$, is assigned to the training data with k-hot encoding :

$$I(i) = \begin{cases} 1, \text{if } \theta_i \in \theta_s, \\ 0, \text{if } \theta_i \notin \theta_s, \end{cases} i = 1, 2, ..., K, \tag{13}$$

where θ_s is a set that includes all directions of the sources. Compared with using the integer encoding, one advantage of using k-hot encoding for DoA problem is when there are multiple sources incoming, only the indexes of I corresponding to the directions of sources are set to 1 instead of creating a new label in the integer encoding. With the k-hot encoding, the probability for one input can be rewritten as :

$$P_i = \prod_{l=1}^{K} \left(\frac{e^{(\boldsymbol{\alpha_l})^T \boldsymbol{x_i}}}{\eta} \right)^{I(l)}, \tag{14}$$

assume the M inputs are independent with each other, the log-likelihood becomes :

$$\log \mathcal{L}(\boldsymbol{x}) = \sum_{i=1}^{M} \sum_{l=1}^{K} \left\{ I(l) \log \frac{e^{(\boldsymbol{\alpha_l})^T \boldsymbol{x_i}}}{\eta} \right\}, \tag{15}$$

with (15) the parameters $\boldsymbol{\alpha}$ can be optimized as the similar way in binary classifier. As mentioned above, the complex sources are assumed, the predicted probability is average of predicted probability of real and imaginary part :

$$P = \frac{P_{real} + P_{imag}}{2}. \tag{16}$$

D. Direction of Arrival Estimation

In principle, when the number of incoming sources, N, is known, the angles corresponding to the highest N probability are selected as the directions for the incoming source. However, in practice, N is usually unknown to us, and the number of non-zero predicted probability can be more than the number of actual sources. After the probability of the test input is predicted, in this letter the number of sources, $\widehat{N}$, is estimated via the following strategy :

1) Sort the predicated probability in descending order, $\widehat{P}(\theta)$ denotes the sorted probability, and $\widehat{P}(\theta_i)$ denotes the i-th entry.

2) The highest probability $\widehat{P}(\theta_1)$ is selected.

3) The probability which satisfies the below equation is selected

$$\frac{\widehat{P}_{\theta_i}}{\sum_{n=1}^{i} \widehat{P}_{\theta_i}} > \epsilon, i = 2, 3, 4, ... \tag{17}$$

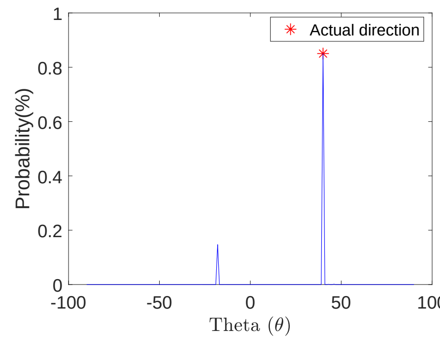

Fig. 4: The predicted probability with N = 1, SNR = 30 dB.

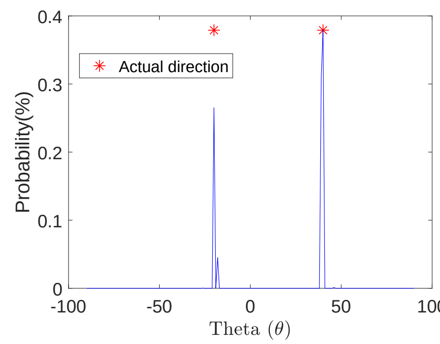

Fig. 5: The predicted probability with N = 2, SNR = 30 dB.

where ϵ is the threshold set by users, and $\widehat{N}$ probabilities can be selected with (17), the corresponding $\widehat{N}$ angles are the estimated directions for the given input.

IV. NUMERICAL RESULTS

In this section, a non-uniform linear array is deployed as the receiver, there are $M = 12$ identical elements in total, and the twelve elements are divided into 4 clusters which are separated by around 100 wavelengths between the neighboring cluster. Each cluster consists of 3 elements, the positions of the 3 elements are randomly settled within the cluster with the range set to 6 wavelengths, the positions are assumed to be fixed. This configuration is referred to *cluster* configuration, which is encountered in real life such as distributed ground communication. The inputs for the training model are the received signals in the noisy environment, the signal to noise ratio (SNR) is defined as

$$\text{SNR} = 10 \log_{10} \frac{\|\mathbf{y_p}\|_\infty^2}{\sigma^2}, \tag{18}$$

where $\mathbf{y_p} = \mathbf{As} \in \mathbb{C}^{M\times 1}$ stands for the signal vector received by the array without noise. The SNR is set to 30 dB to generate the training data, and for each source, 5 samples were generated. The range for the direction of incoming sources is from $-90°$ to $90°$, we used the incoming signals from 2 angles as the source. The size for the entire training data is

REFERENCES

[1] M. Carlin, P. Rocca, G. Oliveri, F. Viani, and A. Massa, “Directions-of-arrival estimation through Bayesian compressive sensing strategies,” *IEEE Trans. on Antennas and Propagation*, pp. 3828–3838, Apr 2013.

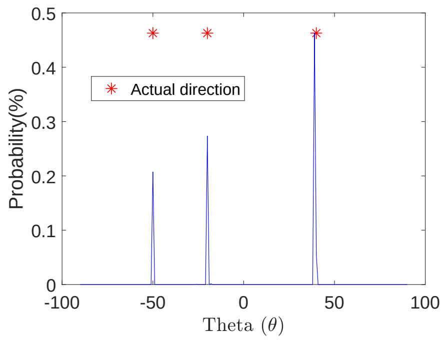

Fig. 6: The predicted probability with N = 3, SNR = 30 dB.

[2] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, Jun 2001.