

Performance analysis of Group Delay MUSIC estimated through Deep Learning

Naveen Verma (JULY,2020).

ABSTRACT

The spatial spectrum gives the spatial distribution of the source coming from all directions to the sink. Hence from spatial spectrum DoA can be estimated and hence it is also called DoA estimation.

One such algorithm for DoA estimation is Multiple Signal Classification also known as MUSIC. The MUSIC method is unable to resolve the closely spaced sources with minimal number of sensors. Hence a group delay function is computed from the phase spectrum of MUSIC algorithm. This function can be used to resolve the closely spaced sources. We introduce a deep learning framework for such a function. We design deep convolutional neural network which receives the Music Group Delay Spectrum at the output and Covariance Matrix of the DoA at the input, and it learns the complex relationship between them. We can show that it works faster than the original function as the predicted output comes faster.

Index Terms: Deep Learning, DoA estimation, MUSIC, Group-Delay.

INTRODUCTION

Direction of arrival (DoA) estimation is crucial in many fields including Radar, Sonar, communications etc. The method of Multiple signal Classification is used to estimate DoA for this purpose. In Music the magnitude spectrum is used in computing DoA. However, in Group Delay we use negative differential of the unwrapped phase spectrum of MUSIC for DoA estimation.

In this we paper we also discuss why we cannot use MUSIC for closely spaced sources. The examples of closely spaced sources can be taken as meeting room scenarios where the microphones are placed closely. When using the model-based approaches the performance of DoA estimation lies on the accuracy of the input data. To remove this drawback, we use learning-based approach where the mapping between the non linear input output pairs is done by the neural networks. The deep learning is able to learn and predict the complex relationships between input and output in data/signals and can hence achieve better performance. We use the multi-layer neural network architecture to resolve the two sources.

We assume that number of targets are small as the complexity generating of data set increases by increase in targets. We introduce a deep learning framework for Group Delay Music and hence the name DeepGDM. The neural network is fed with Covariance Matrix and output to the network is the Unwrapped Group Delay Music phase spectra. We show that

the proposed approach is faster, data-independent and more accurate than the conventional methods.

DoA Estimation using MUSIC

In Music we use direct relationship between the Direction of Arrival of received signal and the corresponding steering vector and hence we estimate the direction of signals from received signals.

If the M number of signals are detected by N sensors ($N > M$), the MUSIC spectrum is defined as:

$$P_{\text{MUSIC}}(\theta) = \frac{1}{(S^H(\theta) Q_n Q_n^H S(\theta))} \quad (1)$$

Where $S = [s(\theta_1) s(\theta_2) s(\theta_3) s(\theta_4) s(\theta_5) \dots \dots \dots s(\theta_M)]$ is a $N \times M$ matrix of steering vector, Q_n is the $N \times (N-M)$ noise eigenvector matrix. The denominator tends to zero when θ is signal DoA. Therefore, we have M peaks corresponding to M signals. However, when the speech sources are spaced close to each other, MUSIC magnitude spectrum fails to produce results with minimal number of sensors.

In the below figure the MUSIC fails to resolve the DoA placed at 54° and 60° with 5 sensors but it resolves the two sources when number of sensors are increased to 15. The plots are for Magnitude Spectrum.

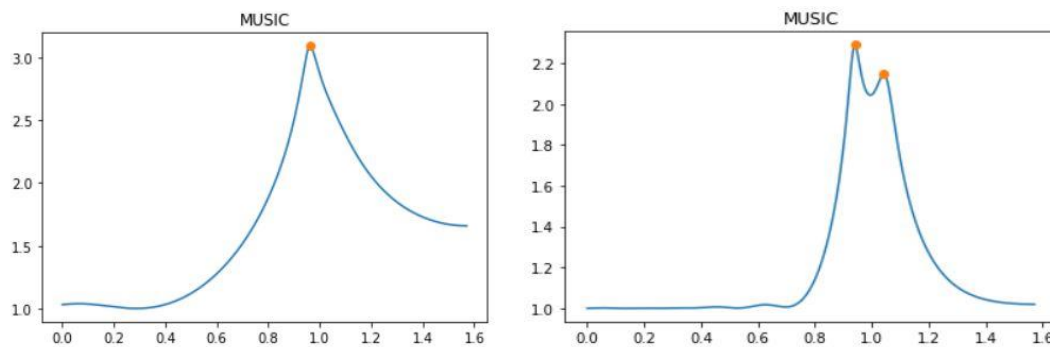


Figure: Music Magnitude Plot for 54° and 60° with 5 sensors(left) , 15 sensors (right).

DoA Estimation using Group Delay of the Music Spectrum:

We propose the use of Group Delay function of the Music phase spectrum for resolving closely spaced sources without the increase in number of sensors. The phase information in the noise eigenvalues of the Music spectrum is used. Let $F_n(\theta)$ denote the phase information.

$$F_n(\theta) = \arg(Q_n^H S(\theta))$$

Where $\arg(\cdot)$ is the instantaneous phase.

However, computing the group delay of the music spectrum can result in sharp spectral changes even when the computed angle is not the DoA. Hence abrupt peaks are formed. Hence to suppress these peaks products of Music spectral magnitude and the group delay of the Music is used so that spurious peaks are withheld and only actual peaks gets amplified.

The Group Delay can be defined as:

$$P_{GDM}(\theta) = \frac{dF'n(\theta)}{d\theta} \times P_{MUSIC}(\theta)$$

This example illustrates the property of Music Group Delay to resolve spatially contagious sources.

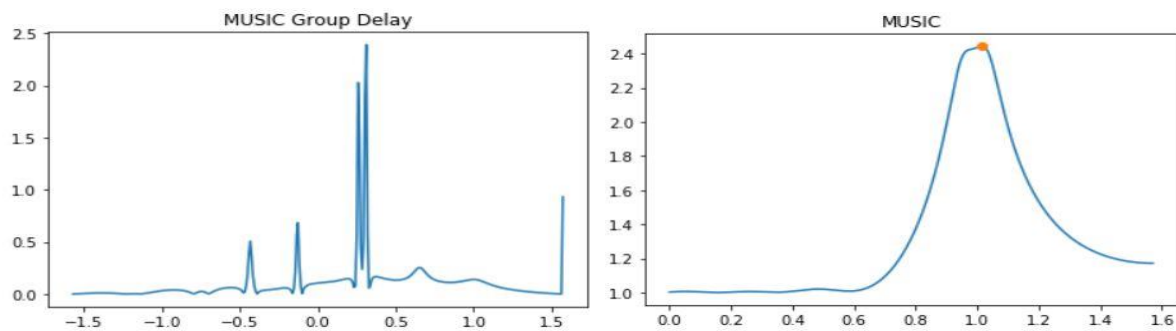


Figure: Group delay(left) and Music magnitude(right) plots for 50° and 55° DoAs.

No. of sensors = 10, snr=20 db.

DoA estimation via Deep Learning

The proposed deep music network is fed with array covariance matrix and yields the Group Delay Music phase spectrum at the output. We first make the inputs and the labels, and then discuss the network architecture and the training.

A. Input Data and Labels

We use 10,000 samples of DoA and corresponding spectra for training. For the DoA of each sample we find out the Group Delay Music Phase spectrum.

As input for the deep learning network we find out the Covariance Matrix from the array steering vector of the DoA. We use the real, imaginary and angle part of the Covariance Matrix \mathbf{R}_y as the input to the convolutional layer of the neural network.

Let \mathbf{X} be $M \times M \times 3$ matrix then first channel of \mathbf{X} for a single sample will be $[[\mathbf{X}];:,1] = \text{Re}\{[\mathbf{R}_y]\}$ for that sample. Similarly $[[\mathbf{X}];:,2] = \text{Im}\{[\mathbf{R}_y]\}$ and $[[\mathbf{X}];:,3] = \text{Angle}\{[\mathbf{R}_y]\}$. Where angle obtained is of imaginary and real parts of \mathbf{R}_y . This is a better method for feature extraction as it helps in input-output mapping.

Using the same covariance matrix, we can find the Group Delay Music phase spectra and use it as label for the entire region of angular spectrum for that sample. Hence, we make input-outputs pairs for 10000 samples.

B. Network Architecture and Training

The deep neural network is composed of 17 layers including input and output layers. The nonlinear mapping function can be represented by $R(M \times M \times 3) \rightarrow R(L)$. Where M =no. of antenna elements and L = No. of angles in the region. The network can be represented by:

$$\text{NN Architecture} = f^{(17)}(f^{(16)}(\dots f^{(1)}(X)))$$

The Fully connected layer is $f^{(14)}$ which maps the input(x) from layer 13 to the output(y) by using the weights W . The output y can be given by the inner product $(W.T \cdot x)$ summed over entire training set N . The $f^{(i)}\{2,5,8,11\}$ represents the convolutional layers with 256 filters of kernel size of 5×5 and 3×3 for the first two and second two layers respectively. The arithmetic operation of a single filter of a convolutional layer can be defined as $Y = \sum \langle W, X \rangle$. Where W is the weight of the convolutional kernel and X is the input to the filter.

$f^{(i)}\{i=3,6,9,12\}$ represents normalization layers and $f^{(i)}\{i=4,7,10,13\}$ represents rectified linear units ReLU layers which can be represented by $\text{ReLU}(x) = \max(0, x)$. $f^{(15)}$ is a dropout layer with dropout=0.5. $f^{(16)}$ is a softmax layer with output $= \frac{\exp(x)}{\sum_i \exp(x)}$. $f^{(17)}$ is a regression layer with size $L \times 1$. L =range of angular spectrum.

The proposed network is trained on Google Colab with GPU accelerator. We used Adam optimizer with the learning rate of 0.01 and $\beta_1=0.9$, $\beta_2=0.999$ with $\epsilon=1e-08$. We take a batch size of 32 samples and train for 100 epochs. The length of Validation dataset is = 20%. The angular range taken is $[-90^\circ, 90^\circ]$ with 360 angles. We take 400 snapshots for every sample in our simulation.

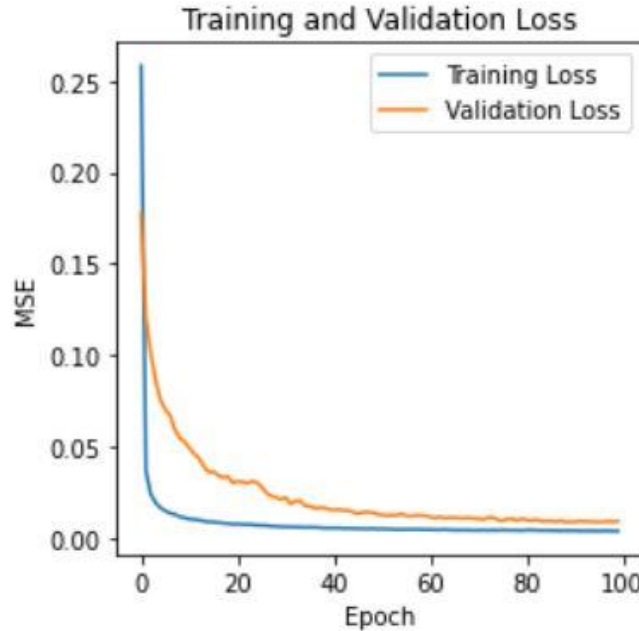


Figure: Training and Validation Loss

C. Predictions and Results.

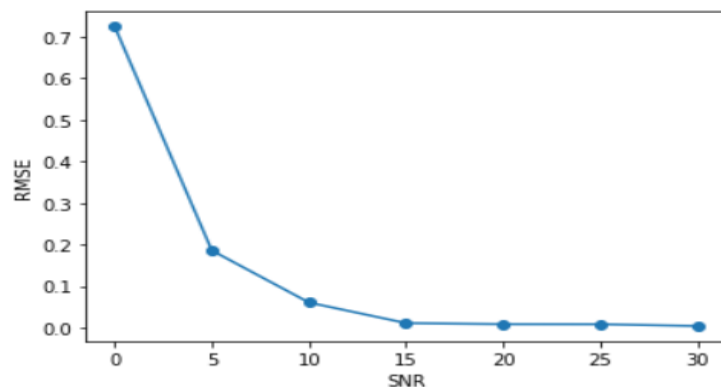
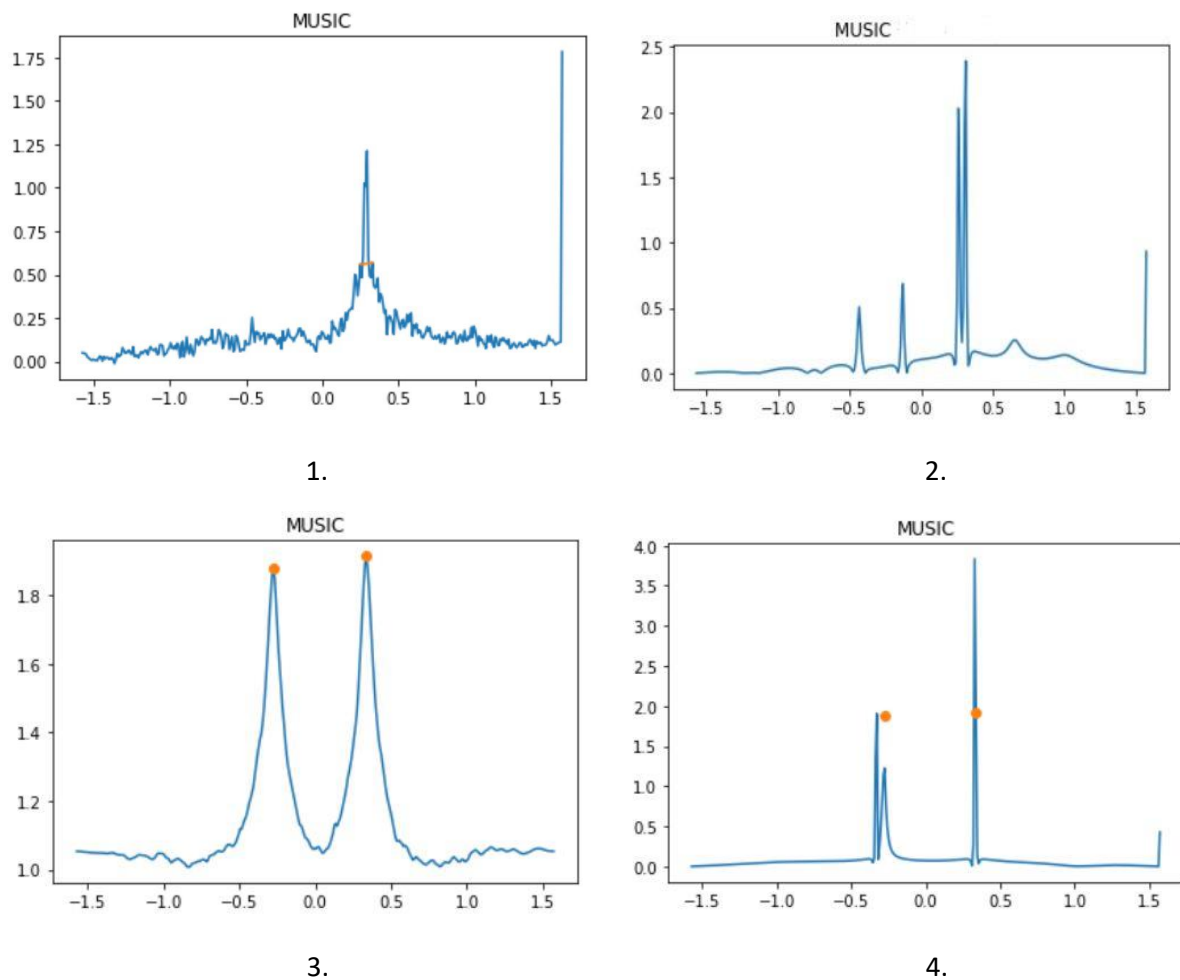


Figure RMSE vs SNR

We see that DeepGDM is able to resolve spatially contagious sources with limited amount of information. The Deep learning framework nearly approximates the actual DoAs and once trained thus becomes independent of input data which might sometimes get corrupted due

to presence of noise. This is the major advantage of using Deep Learning framework for direction estimation.

The proposed approach is also faster as the time taken to calculate predicted output from DeepGDM is 0.0324 seconds and 0.1305 seconds by using Group Delay function. By using convolutional layers there is a better feature extraction of hidden input data which could be attributed for the performance of the model.

The performance deteriorates at low snr due to noisy data as the resolving power of the model increases with increase with snr during training. Once the model is trained it becomes independent of the input data due to good generalisation of neural networks.

Conclusions

We implemented a Deep Group Delay Music phase function which is able to resolve spatially contagious sources with less computational time and provides a good approximation of the group delay function thus making it independent of input data after training. It can work for multiple targets thus increasing its significance.

References

- [1] T. E. Tuncer and B. Friedlander, *Classical and Modern Direction-of Arrival Estimation*. Academic Press, 2009.
- [2] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *Signal Processing Magazine, IEEE*, vol. 13, no. 4, pp. 67–94, Jul 1996.
- [3] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.
- [4] B. Friedlander and A. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Trans. Antennas Propag.*, vol. 39, no. 3, pp. 273–284, 1991.
- [5] Z. Liu, C. Zhang, and P. S. Yu, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7315–7327, Dec 2018.
- [6] L. Wu, Z. Liu, and Z. Huang, "Deep Convolution Network for Direction of Arrival Estimation With Sparse Prior," *IEEE Signal Process. Lett.*, vol. 26, no. 11, pp. 1688–1692, Nov 2019.
- [7] D. Hu, Y. Zhang, L. He, and J. Wu, "Low-Complexity Deep-LearningBased DOA Estimation for Hybrid Massive MIMO Systems with Uniform Circular Arrays," *IEEE Wireless Commun. Lett.*, pp. 1–1, 2019.
- [8] A. M. Elbir, K. V. Mishra, and Y. C. Eldar, "Cognitive radar antenna selection via deep learning," *IET Radar, Sonar & Navigation*, vol. 13, pp. 871–880, 2019.