Machine Learning: Assignment 2: ICA

Vivek Kumar Khetan
UT-EID - vkk287

1. Introduction:

The goal of this study is to learn about the Blind Signal Separation and to use Independent Component Analysis (ICA) to separate a set of source sounds from mixed sound. Blind signal separation is also known as blind source separation. Now, blind source separation is used for images, sounds as well as tensor separation. We will be using ICA for blind source separation for separating mixed sounds and to obtain the original source sounds. We need to aware about the basic assumption before using ICA for blind source separation. Our data should not be Gaussian and we should have enough data to use ICA. If the source sound data is Gaussian, we won't be able to recover our original sound form mixed as there will be a rotational component in the mixing matrix.

One very interesting point is to what point our source signal is recoverable and how good we can recover it. In our experiment, we will observe few very interesting points: -

- We won't be able to recover the source sound in the original scale.
- Sometimes, recovered sound wake will be inverted, i.e., we will recover inverted sound wave. But, it doesn't matter as both will paly the same sound.

2. **METHOD**:

We are given a four second sound clip of five different sources samples at 11025, i.e., a 5 by 44000 matrix. Let's say our source matrix, **U** is a **n** by **t** matrix of **n** source signals of length **t,** where **n** = 5. We will mix these five sound sources using a square inverse matrix A and will obtain a 5 by 44000 matrix X so that, **X = AU** where **A** is an **m** by **n** matrix such that Ai,j is the weight of the jth source signal in the ith mixed signal. Now, our task is to find the optimize value of matrix **W** (separator matrix) so that **Y = WX**. **Y** matrix is our recovered sound sources.

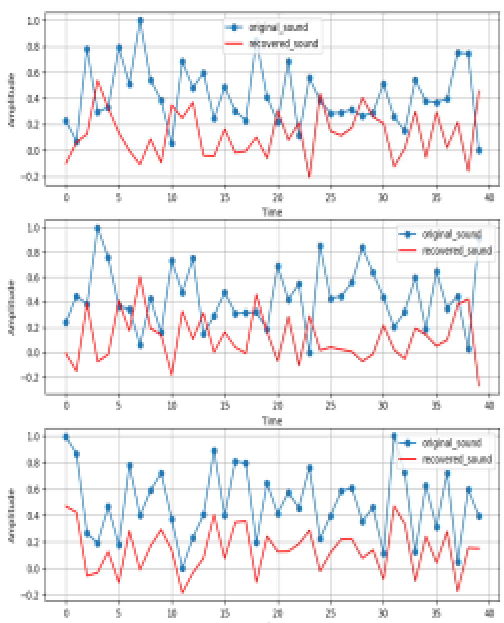
Following are the steps I have taken:

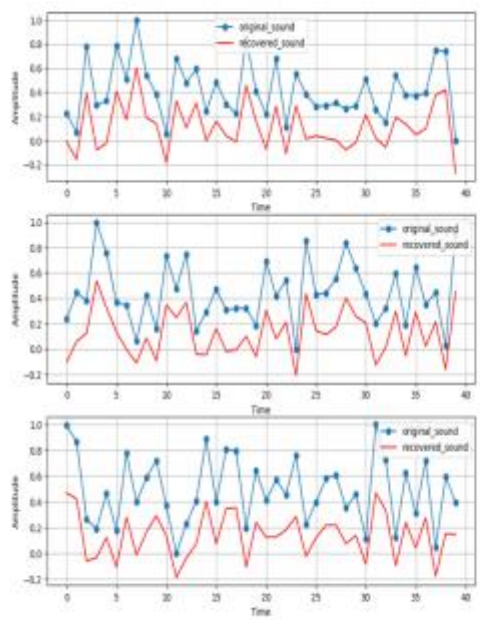
- Assumed **X = AU**.
- Initialized the (**n** by **m**) matrix **W** with small random values.
- Calculated **Y = WX**. **Y** as the current estimate of the source signals.
- Calculated **Z** where $z_{i,j} = g(y_{i,j}) = 1/(1+e^{-y_{i,j}})$ for $i \in [1..n]$ and $j \in [1..t]$ (where **t** is the length of the signals). This helped me traverse the gradient of maximum information separation.
- calculated $\Delta W = \eta(I + (1-2Z)Y^T)W$ where **η** is a small learning rate.
- Updated **W = W + ΔW** and repeat from step 3 until convergence in R_max iterations

3. **Discussion**:

I started with the given test and mixing matrix to decide my algorithm's learning rate as well as no of iteration it was taking to start to converge. It was a hit and trail method. I tried with many different values. I used numpy random generator function to generate my W matrix and then I also used various factor(0<fact<1) to make W smaller. It helped with converge with relatively smaller no of iterations. To confirm that my algorithm has been converging, I printed value of W for each iteration and kept an eye on the computer screen. It was fun finding W getting out of bound sometime and sometime getting forever to converge. Finally, I was able to find my algorithm converging for certain value of learning rate and no of iterations.

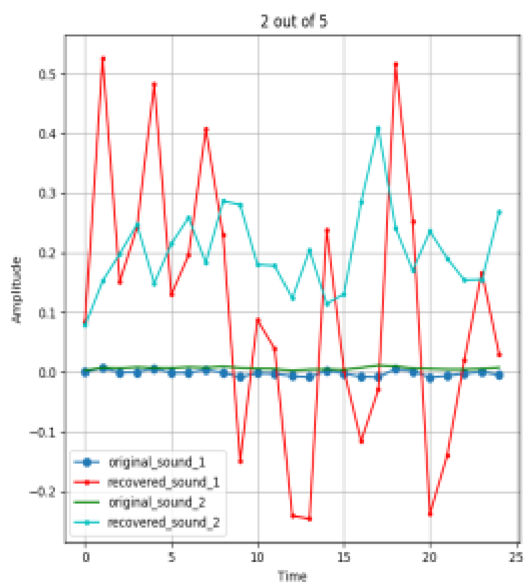
First I used my algorithm on proved test file and was able to retrieve the original source sound.

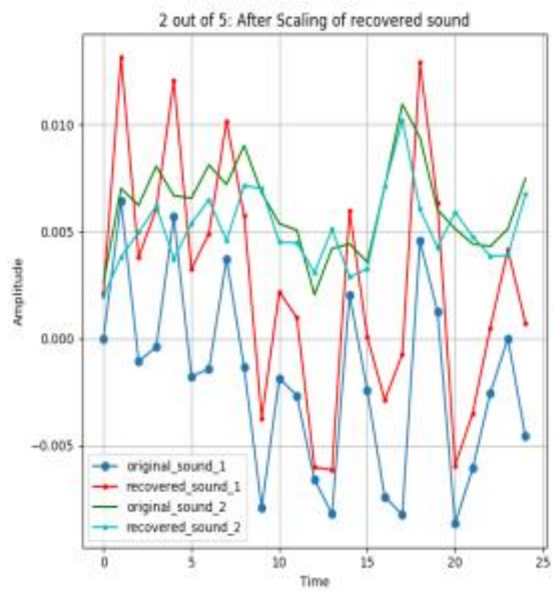




The figure on the left has three different plots of original and recovered source sounds. It is interesting to see how recovered sound source in first and second subplot got exchanges. The plot on the right is the final plot of the provided three source test file. I have to change the position of first and second recovered sound source and it is very evident that how we were able to recover the source sound although it was a little bit scaled down.

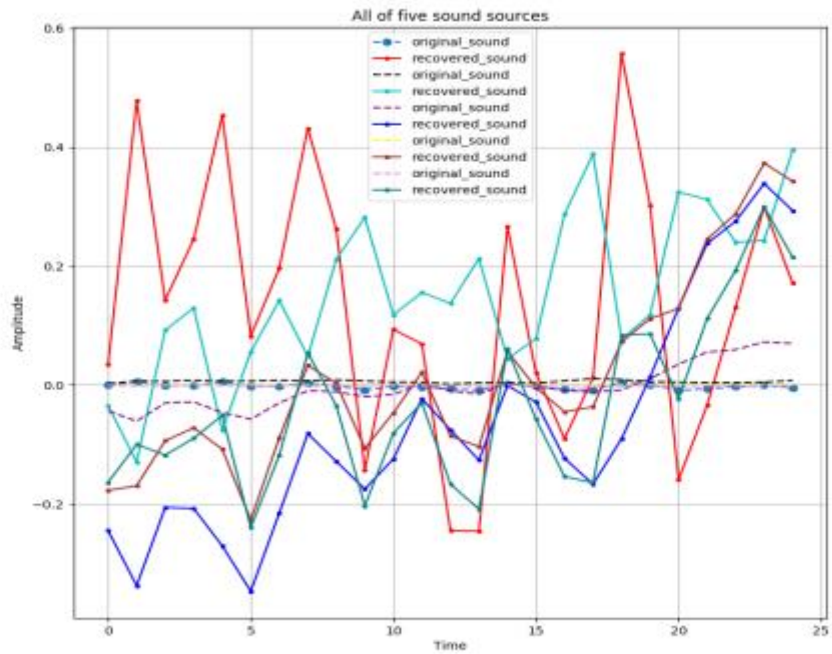
There are many other interesting insights I was able to get while doing my experiment. In the above figure, the recovered sound got scaled down. In one of my experiment recovered source sound was scaled up. I have to scale it down to check how good it was in comparison to original sound.





The above two figures are of the same two sound sources. These two-sound sources belong to the five-provided source matrix. In the figure on the left, at first sight we can't say which one is recovered sound source and which one is original provided sound source. But, after a careful observation one can easily see how the recovered sounds have be scaled up. Therefore, after scaling down (by multiplying with a factor) the recovered sound we can easily see how good our algorithm was in recovering the source sound.

Finally, I used matrix for all the five sound sources, mixed the sounds in random manner and then used my algorithm to recover the original sound for all the five sources. It was quite interesting to see all ten waves in a single plot (five for original and five for recovered sound sources). In the figure below, it is clearly evident how two recovered sounds were scaled down one was flipped and two were scaled up. But if we will play all five in a speaker it will create the same sound, may be of higher or lower intensity.



4. **Conclusion and Future Work:**

With the above results, the usefulness of Independent Component analysis (ICA) is clearly evident. We were able to recover original source sound form mixed noise even without having any initial individual information about the original source sound. Although, we did assume that the source sound is not Gaussian and each sound source is independent of other.

It would be interesting to record sound from my surrounding and then perform this experiment. I would like to record some sound on my own and perform this experiment on that mixed sound. It would be interesting to see; how good I can recover real life sound.