Independent Component Analysis with Neural Networks

Project category: Applications
R. Rojas, SCPD, raulr@stanford.edu, Video: https://youtu.be/pvsSFVYxpmY

Problem and motivation

Independent Component Analysis (ICA) is an interesting problem for diverse applications. Assuming that we have n sensors, and that n mixed sources have been recorded, we would like to separate the individual independent sources from the n recorded mixtures. Independent sources should have minimal correlation. One possible application, is to separate the voices from several people using a microphone array (Fig. 1), or separating signals in an array of EEG sensors coming from different regions of the brain. ICA is sometimes also called blind source separation.

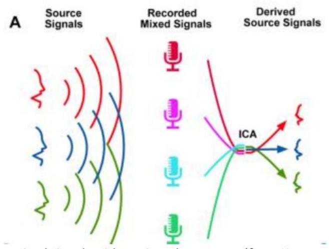


Fig. 1: Recording mixed signals with a microphone array (from Karczweski et. Al. 2013)

Analysis

We call y_i(t) the *m* original (and unknown) signals (at time t). Our task is to recover them from the recorded *m* signals x_i(t). We assume that the mixing process is linear, that is, the vector of signals y_i(t) is multiplied by a matrix of weights M yielding the recorded signals.

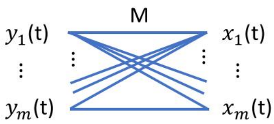


Fig. 2: The original signals are mixed linearly by the matrix of weights M.

In mathematical terms, we have n time series of linear mixtures $x_i(\tau)$, τ being the timeframe. If we multiply the time series by the un-mixing matrix W, we recover an estimation of the independent components $y_i(\tau)$. The mixing matrix M mixes the independent components and provides an estimate of the recordings $\tilde{x}_i(\tau)$. Both M and W are unknown. M is the physical process producing the mixture, W the matrix we need to find in order to implement source separation. We want the sum of squared differences of the recordings $x_i(\tau)$ and the reconstructed recordings to be minimal $\tilde{x}_i(\tau)$.

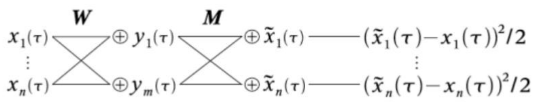


Fig. 3: Unmixing the original signals followed by mixing provides an estimation of the error.

As stated the problem seems ill-posed, since any invertible matrix M and its inverse W would provide a perfect reconstruction of the recorded inputs. Therefore, we introduce constraints to limit the kind of acceptable mixing and unmixing matrices M and W.

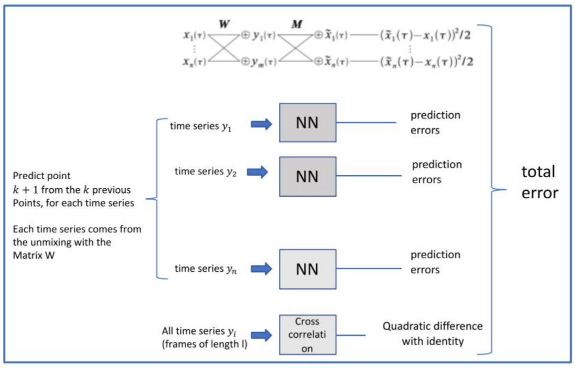
Smoothness constraint

The proposed approach is to take each reconstructed time series $y_i(\tau)$ and predict the next point $k+1$ using the previous k data points. The prediction is done with a neural network for each independent component (we have n neural networks, one for each time series $y_i(\tau)$). The prediction errors are added to the squared differences shown in the diagram. That is, the final network consists of one network for each $y_i(\tau)$ and the unmixing-mixing process shown above (see diagram), together with a computation of the matrix of cross-correlations of the $y_i(\tau)$'s. We initialize W and M randomly and train to minimize the error function. The training process should find the best neural networks for predicting the independent components and, at the same time, the matrices M and W (notice that the diagram represents a composite network). Not every combination of W and its inverse can work, because the predictability of the reconstructed components $y_i(\tau)$ is being maximized at the same time. The intuition is that regulating the capacity of the neural networks we can nudge the system to prefer uncorrelated $y_i(\tau)$'s.

Fig. 4: The composite network for unmixing the recorded signals. Smooth and uncorrelated y's are preferred using the total error function..

Datasets and Evaluation

The network has been tested using self-generated synthetic signals and audio signals. The prediction



networks have been limited to perceptron linear units, which corresponds to Predictive Linear Coding (used in telephony). The current network can separate such synthetic sources very well after a few minutes of computation. The figure below shows the result of an experiment with two audio signals: the first row shows a recording of voice and applause, the middle row the mixed signals, and the last row the reconstructed signals.

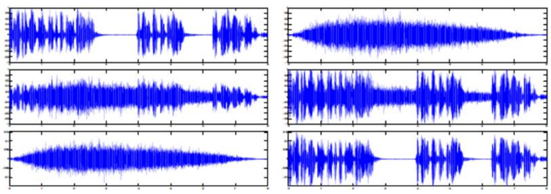


Fig. 4: An experiment with two audio signals. Upper row: original signals. Middle row: mixed signals. Lowest row: reconstructed signals.

References

Aapo Hyvärinen, Erkki Oja und Juha Karhunen, *Independent Component Analysis*, Wiley, 2002.
Konrad Karczweski et al., „Coherent Functional Modules Improve Transcription Factor Target Identification", *PLOS Genetics*, Feb. 2014, V. 10, N. 2