PROJECT 1

FASTICA ALGORTITHM

Paul Pineau
Gdansk University of Technology
Department of Electronic, Telecommunication and Informatics

April 23, 2020

Contents

INTRODUCTION

This paper will explain how to implement the FastICA algorithm. This is an efficient and popular algorithm for independent component analysis invented by Aapo Hyvärinen at Helsinki University of Technology The aim of the algorithm is to unmixed mixed signal. For example is there is two people speaking at the same time in a same room with two microphone, the FastICA algorithm can separate what those two people said in two files. This algorithm is useful in many fields like the medical one.

I FIRST STEP : DATA PROCESSING

The first step of the algorithm is to accomplish some data processing

Center the Data

Let X = $x_{(i,j)} \in \mathbb{R}^{N \times M}$ the matrix of input data where M is the number of mixed signals and N the numbers of independant source signals.

We must center the data ie. :
$\forall (i,j) \in [0,..,N] \times [0,...,M], x_{(i,j)} = x_{(i,j)} - \frac{1}{M}\sum_k x_{(i,k)}$

The mean value of the signal is now 0.

Whiten the data

All the data must be whiten, ie their covariance matrix must be the identity matrix. By the way we are assured that all the data are uncorrelated (by definition of whiten data).
Let x be our signal with 0 as mean value and E his covariance matrix.

$\exists P, N \in \mathbb{R}^{n,n}, E = PDP^{-1}$

We have $\tilde{x}$ the white signal with $\tilde{x} = PD^{-\frac{1}{2}}P^T x$

II SECOND STEP : ONE SOURCE EXTRACTION ALGORITHM

The aim of this part is to unmixed the signal by maximize the non gaussianity of $w^T X$ where w is a column vector and X = $x_{(i,j)} \in \mathbb{R}^{N \times M}$ the matrix of the input data previously whiten

a) Initialisation of w

First of all we must initialize w.
Let w = $[w_{i,j}]$ randomly choosen with $||w|| = 1$

b) One source extraction iteration

We compute the next w_i with this formula:

$$w_i = \sum_t (z(t)[w_{i-1}^T z(t)]^T) - 3w_{i-1}$$

where $z(t) = [z_i(t)]^T$ is the vector of whitened mixture signals.

Then w_i is normalized

$$w_i = \frac{w_i}{||w_i||}$$

III THIRD STEP : VERIFICATION

We have a sequence $w_i, i \in \mathbb{N}$ which converges. We compute the dot-product between $w_i and w_{i-1}$, if it is close to one then our sequence is near the convergence point so we stop the algorithm. If not, we go back to step II)b), 'One source extraction iteration'.
In other words:

Let $\epsilon \in \mathbb{R}$, if $|w_i^T w_{i-1} - 1| > \epsilon$ then go back to step II)b)

else continue the algorithm

IV STEP 4: UNMIXED DATA RECOVERY

Once all the iteration of the One source extraction algorithm are done we must recover the unmixed data. We must compute v, orthogonal vector to w.
In other words

$$\exists v \in \mathbb{R}^N, w^T v = 0 \text{ and } ||v|| = 1$$

Let Y the matrix with the unmixed signals.

$$Y = \begin{bmatrix} w^T \\ v^T \end{bmatrix} \begin{bmatrix} z_1(t) \\ \dots \\ z_N(t) \end{bmatrix}$$

Matlab Implementation

I implemented this algorithm on Matlab for n mixtures.

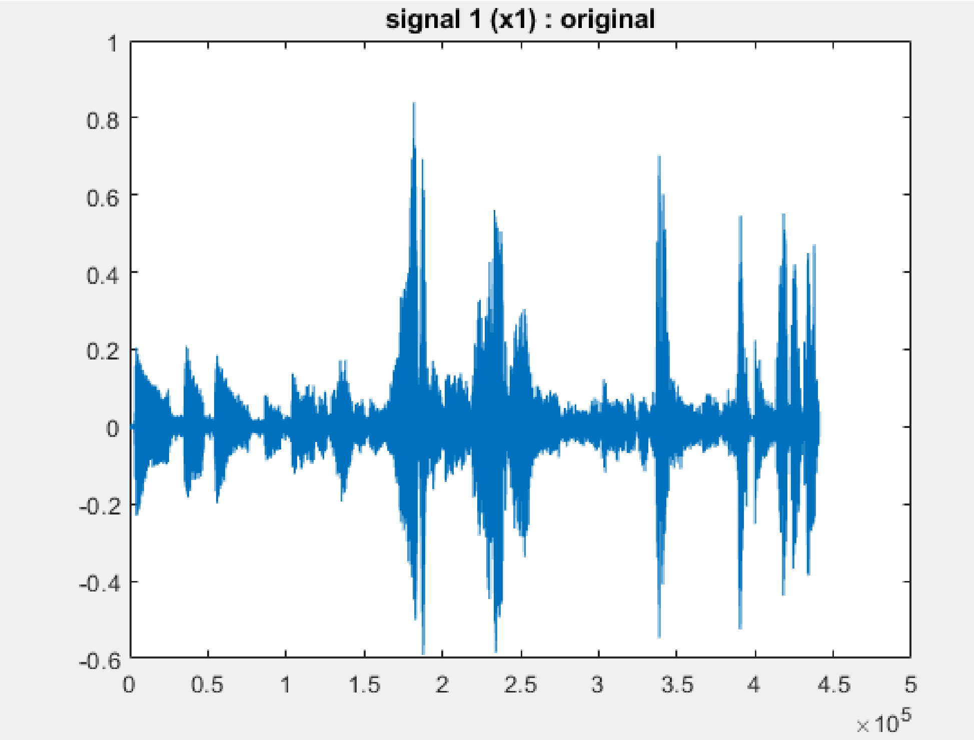

Figure 1: Here is a original signal

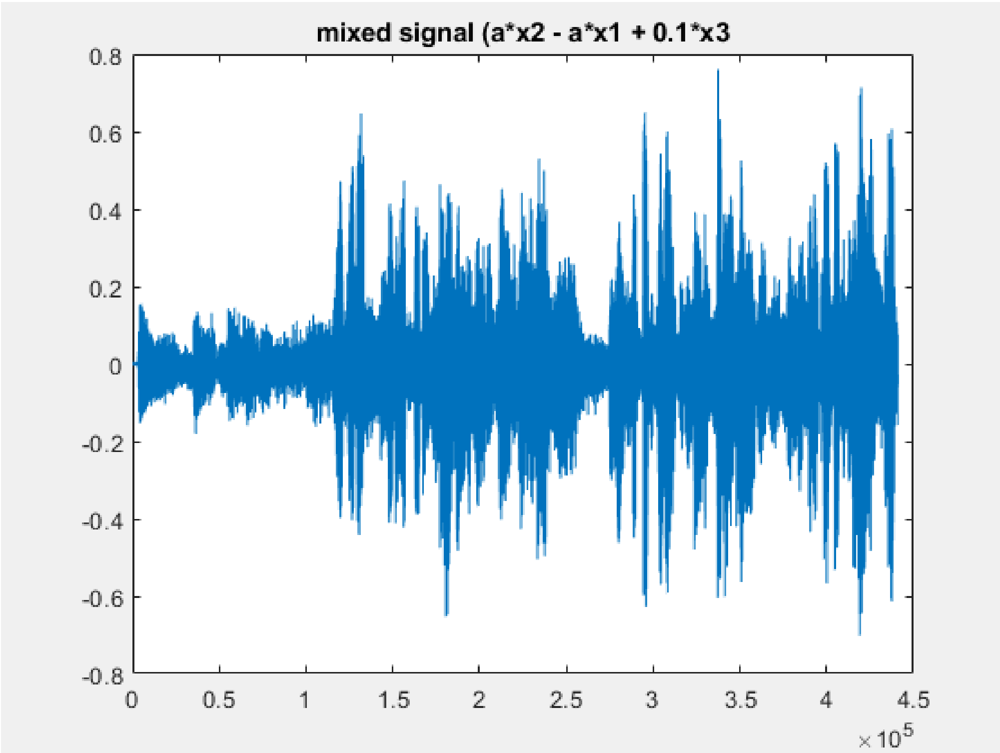

Figure 2: Picture of one mixture

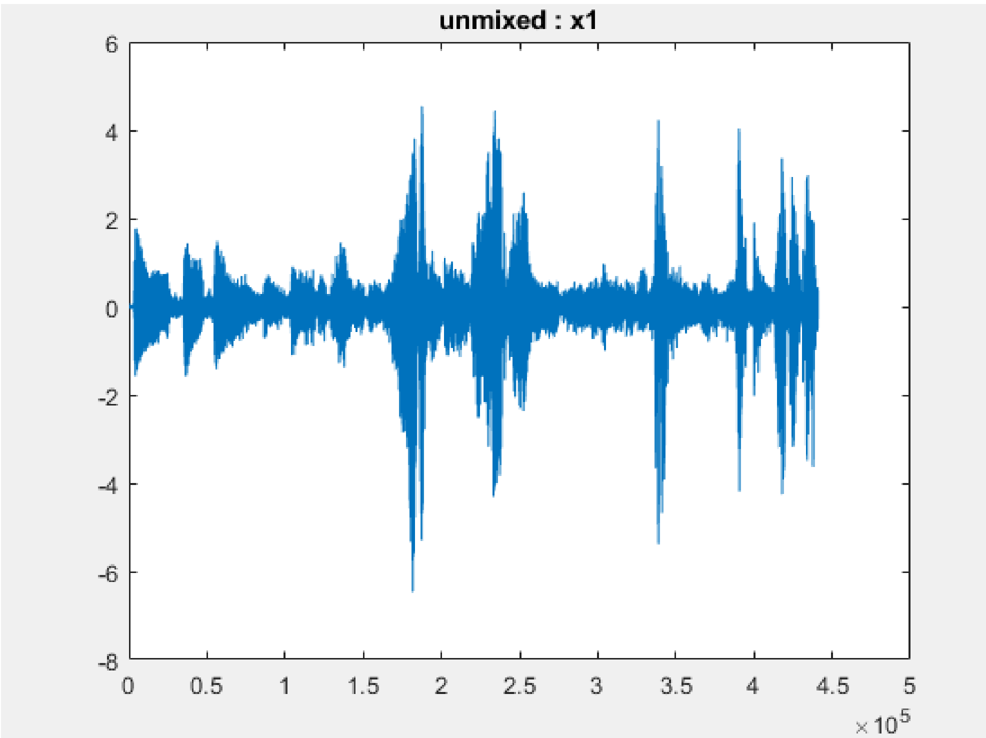

Figure 3: Unmixed signal