Independent Component Analysis

MVA Project – One page draft

Nathan de Lara Florian Tilquin Vincent Vidal

Abstract

We decided to work on the Independent Component Analysis problem. We intent to implement, apply and compare several algorithms while being certain to understand the link between the likelihood maximisation and the mutual information.

1 Definition of the problem

The general problem can be formalise this way. We suppose to have some random variables $x \in \mathbb{R}^p$ which correspond to a mix of some primitives random sources $s \in \mathbb{R}^n$. The goal is to extract from x every source s_i. To do so, we will suppose here that:
– the sources are independents
– the mix is linear and instantaneous
– at most one source has a gaussian distribution.
So we will write:

$$x = As \quad \text{and} \quad y = Wx, \tag{1}$$

where A is the mixing matrix, W the separation matrix and y the estimation of the sources. The goal is then to find the matrix W that minimise the independence of y. As a measure of independence, we will consider, for theoretical purpose, the mutual information:

$$\mathrm{I}(Y) = \int_{\mathbb{R}^p} P(Y) \log \frac{P(Y)}{\Pi_i P_i(Y_i)} \mathrm{d}Y. \tag{2}$$

For practical purpose, we will consider other contrast function than the mutual information (too hard to compute), which are invariant by permutation and scaling on coordinates and "minimal" for independent coordinates.

2 Algorithms for ICA

2.1 Hérault and Jutten (HJ) algorithm

This method is based on the neural network principle. We write $W = \left(I_n + \widetilde{W}\right)^{-1}$ and for a pair of given function $(f,\, g)$, we adapt the $\widetilde{W}$ as follows:

$$\widetilde{W}_{ij} = f(y_i)g(y_j). \tag{3}$$

2.2 EASI algorithm

For a given contrast function ψ and a cost function J_ψ, we iterate as follows:

$$W_{t+1} = \left(I_n - \lambda_t \nabla J_\psi(y_t)\right) W_t. \tag{4}$$

The EASI algorithm correspond to the functions $\psi(y) = \sum |y_i|^4$ and $J_\psi(y) = \mathbb{E}\big[\, \psi(y) \,\big]$.

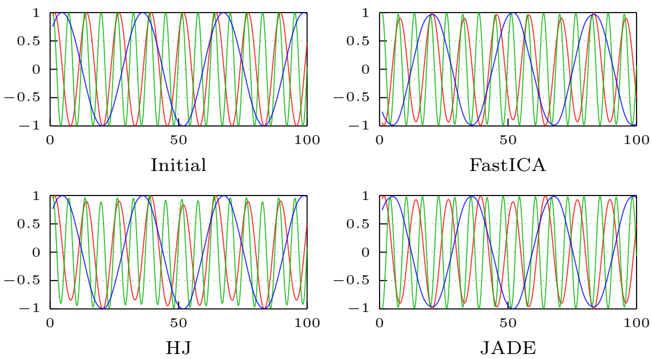

Figure 1: Results of the main algorithm for ICA on simulated data. Each colour correspond to a signal. The "Initial" graph show the initial signals and the other one the results of the algorithms.

2.3 Jade algorithm

Several methods are based on the cumulants. The goal here is to annul all the cross cumulants of order 4. So we want to diagonalize the cumulant tensor which is equivalent to minimise the following contrast function:

$$c\,(x) = \sum_{i,k,l} \left|\mathrm{Cum}\,(x_i, x_i^*, x_k, x_l)\right|^2. \tag{5}$$

2.4 Infomax algorithm

This method is based on the information theory equality $\frac{\partial \mathrm{I}(y,\,x)}{\partial w} = \frac{\partial \mathrm{H}(y)}{\partial w}$, which gives us an update rule for minimising the mutual information:

$$\Delta W = \left[W^\intercal\right]^{-1} + \frac{\partial \ln \Pi_i \left|y_i'\right|}{\partial W}. \tag{6}$$

3 Results

First, we decided to apply the main algorithm to some simple and simulated data. For tree sinusoidal signals, we show the results in figure 1

4 What's next

We intent now to implement our own version of some algorithm and apply it to different type of data such as electroencephalogram, financial data and images.