ICA

Independent Component Analysis

Zakariás Mátyás

Contents

- Definitions
- Introduction
- History
- Algorithms
- Code
- Uses of ICA

Definitions

- ICA
- Mixture
- Separation
- Signals …typical signals
- Multivariate statistics
- Statistical independence

Definitions

- What is it?

- **Independent component analysis (ICA)** is a method for separating a multivariate signal into subcomponents, supposing the mutual statistical independence of the non-Gaussian source signals. It is a case of blind source separation or blind signal separation.

Definitions

- Mixture

- The data mixture can be defined as the mix of one or more independent components which require separation
- A mixture model is a model in which the independent variables are measured as fractions of a total.
- K-number of components
- a_k – mixture proportion of k
- $h(x|\lambda_k)$ – probability distribution

$$p_X(x) = \sum_{k=1}^{K} a_k h(x|\lambda_k)$$

Definitions

- Multivariate statistics

- **Multivariate statistics** or **multivariate statistical analysis** in statistics describes a collection of procedures: observation and analysis of more than one statistical variable at a time.

- Analysis: regression analysis (linear formula – how variables behave when others change)

What is this ?

Definitions

Why here?

- PCA – principal component analysis (small set of synthetic variables

Why here?

- LDA – linear discriminant analysis (linear predictor from 2 sets of data for new observations)
- Logistic regression, MANOVA, artificial neural networks, multidimensional scale

Why here?

Definitions

- Statistical independence

- In probability theory, to say that two events are **independent** means that the occurrence of one event makes it neither more nor less probable that the other occurs.

$$P_y(y/x) = \frac{P_{x,y}(x,y)}{P_x(x)} = P_y(y) \Leftrightarrow P_{x,y}(x,y) = P_x(X)P_y(y)$$

Definitions

- Separation

- **Blind signal separation**, also known as **blind source separation (BSS)**, is the separation of a set of signals from a set of mixed signals. It is done without the aid of information (or with very little information) about the nature of the signals.

Introd... ...on

ICA statistically illustrated.

- Uniform distributions:

$$p(s_i) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{if } |s_i| \leq \sqrt{3} \\ 0 & \text{otherwise} \end{cases}$$

- Mixing matrix:

$$\mathbf{A}_0 = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$$

Gaussian variables are forbidden, because their joint density shows a completely symmetric density. It does not contain any information on the directions of the columns of the mixing matrix A. This is why A cannot be estimated.

What this means?

Introduction

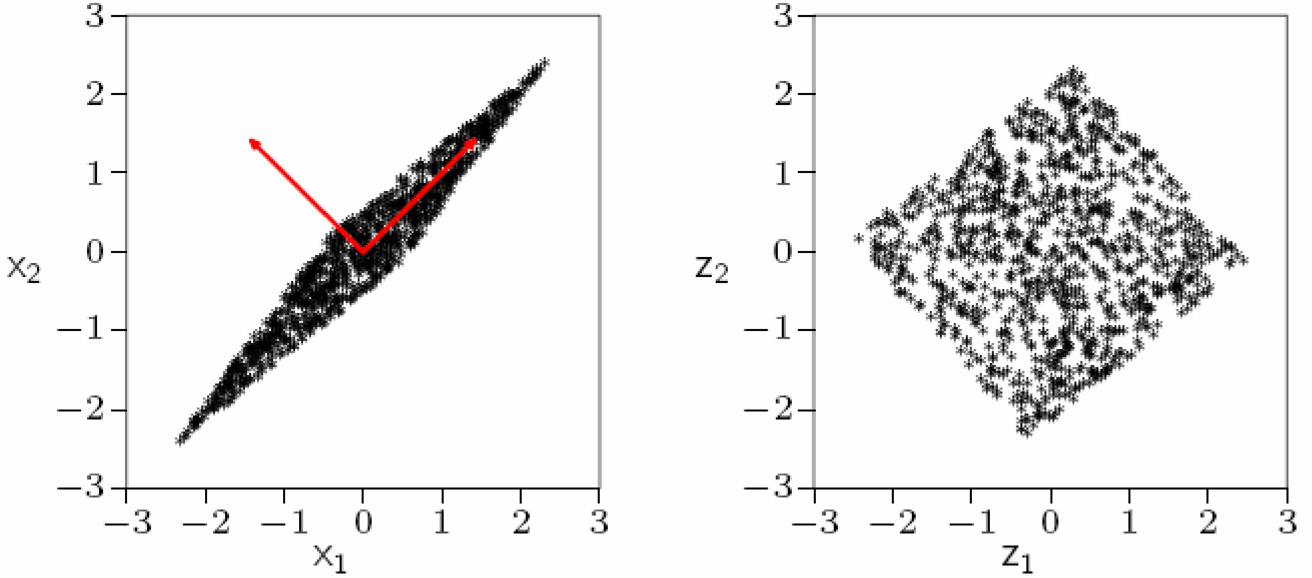
- ICA preprocessing

- Before using any of the ICA algorithms it is useful to do some data preprocessing for simplifying and reducing the complexity of the problem (data):

  1. Centering
  2. Whitening
  3. Other preprocessing steps depending on the application itself (for ex.: dimension reduction)

Introduction

- Whitening:
  - Remove linear dependencies
  - Normalize projection variance



- $\mathbf{z} = \underbrace{\mathbf{D}^{-1/2}\mathbf{E}^T}_{=\mathbf{V}}\mathbf{x}$, $\mathrm{E}\left\{\mathbf{z}\mathbf{z}^T\right\} = \mathbf{I}$

History

- Source separation is a well studied, old problem in electrical engineering too.
- There are many mixed signal processing algorithms.
- It is not easy to use BSS on mixed signals, without knowing any information, that helps us to create a good separating algorithm.

History

- ICA framework was introduced by Jeanny Herault and Christian Jutten in 1986.
- Stated by Pierre Comon in 1994

Infomax algorithm

- 1995 – Tony Bell and Terry Sejnowski created the infomax ICA algorithm, which had a principle introduced by Ralph Linkser in 1992

History

- 1997 – Shun-ichi Amari -> infomax algorithm improvement by natural gradient (Jean-Francois Cardoso)
- Original infomax algorithm was suitable for super-Gaussian sources
- Non-Gaussian signal version developed by Te-Wonn-Lee and Mark Girolami

Algorithms

- ICA algorithms

- FastICA – Aapo Hyvarinen, Erkki Oja, using the cost function: kurtosis

kurtosis - In probability theory and statistics, **kurtosis** is a measure of the "peakedness" of the probability distribution of a real-valued random variable. We measure with it the nongaussianity.

Kurtosis of y: $$\mathrm{kurt}(y) = E\{y^4\} - 3(E\{y^2\})^2$$

Algorithms

- ICA algorithms(2)

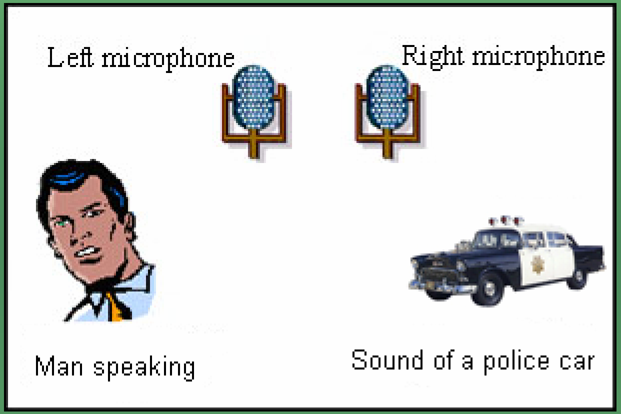
- **Kernel ICA** Contributed by Francis Bach

Implements ICA algorithm for linear independent component analysis (ICA). The Kernel ICA algorithm is based on the minimization of a contrast function based on kernel ideas.

- The well known cocktail-party problem (simplified: only two voices)

- Imagine you're at a cocktail party. For you it is no problem to follow the discussion of your neighbors, even if there are lots of other sound sources in the room: other discussions in English and in other languages, different kinds of music, etc.. You might even hear a siren from the passing-by police car.
- It is not known exactly how humans are able to separate the different sound sources. ICA is able to do it, if there are at least as many microphones or 'ears' in the room as there are different simultaneous sound sources.

cocktail-party problem



The microphones give us two recorded time signals. We denote them with $x=(x_1(t), x_2(t))$. x_1 and x_2 are the amplitudes and t is the time index. We denote the independent signals by $s=(s1(t),s2(t))$; *A - mixing matrix (2x2)*

$$x_1(t) = a_{11}s_1 + a_{12}s_2$$
$$x_2(t) = a_{21}s_1 + a_{22}s_2$$

a_{11}, a_{12}, a_{21}, and a_{22} are some parameters that depend on the distances of the microphones from the speakers. It would be very good if we could estimate the two original speech signals *$s_1(t)$ and $s_2(t)$, using only the* recorded signals *$x_1(t)$ and $x_2(t)$. We need to estimate the a_{ij}.*, but it is enough to assume that *$s_1(t)$* and *$s_2(t)$, at each time instant t, are statistically independent. The main task is to transform the data (x); s=Ax to independent components, measured by function: F(s1,s2)*

Steps

2 vectors containing the points of original sources

Mixing matrix

Mixed signals (begin)

Weight matrix

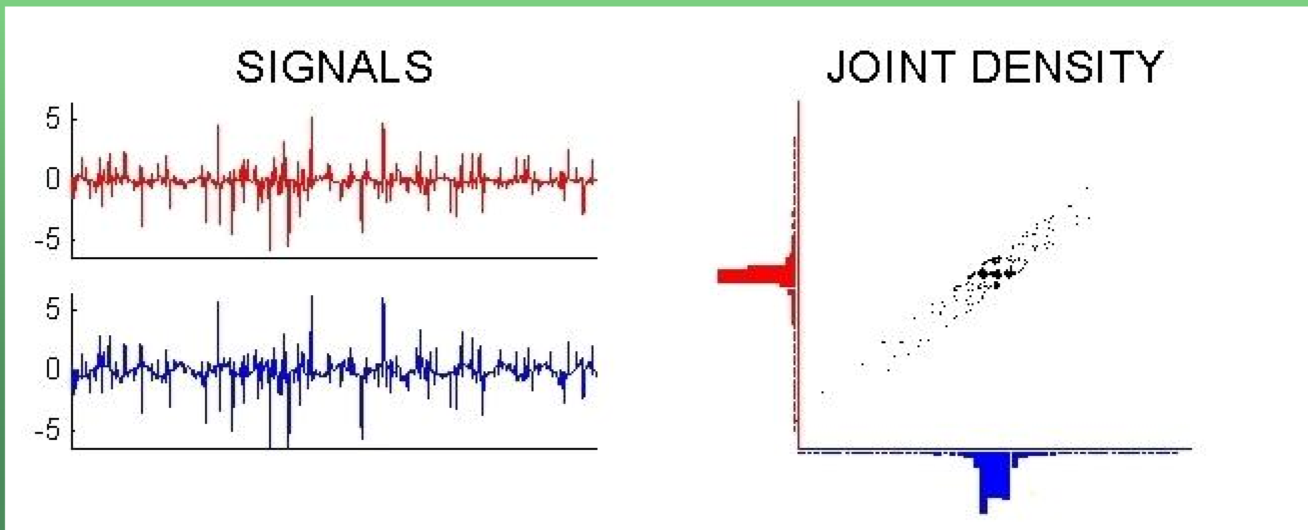
Estimation

$$\begin{bmatrix} \bar{S}_1 \\ \bar{S}_2 \end{bmatrix} \rightarrow \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \rightarrow \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \end{bmatrix}$$

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} * \begin{bmatrix} \bar{S}_1 \\ \bar{S}_2 \end{bmatrix} and \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \end{bmatrix} = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} * \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}$$

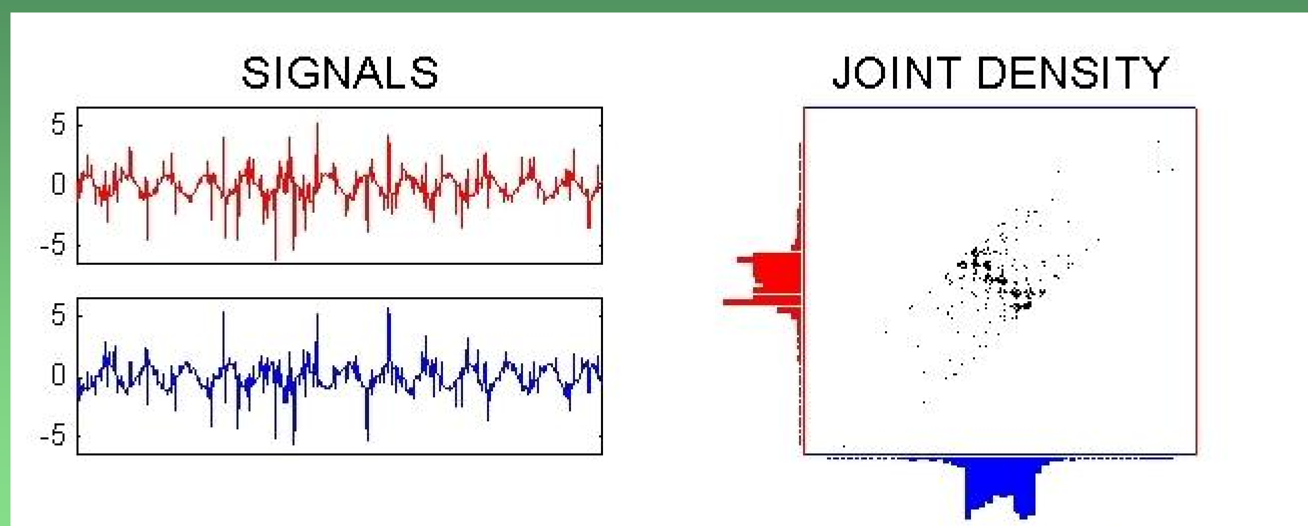
Steps
FastICA

the joint density of two independent variables is just the product of their marginal densities

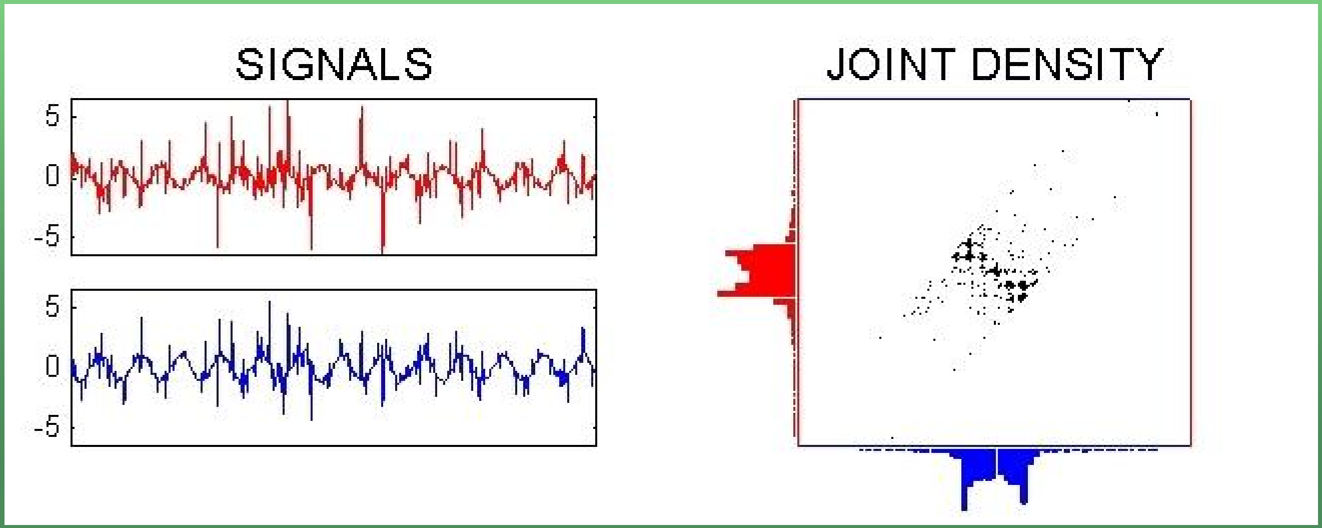


Original data

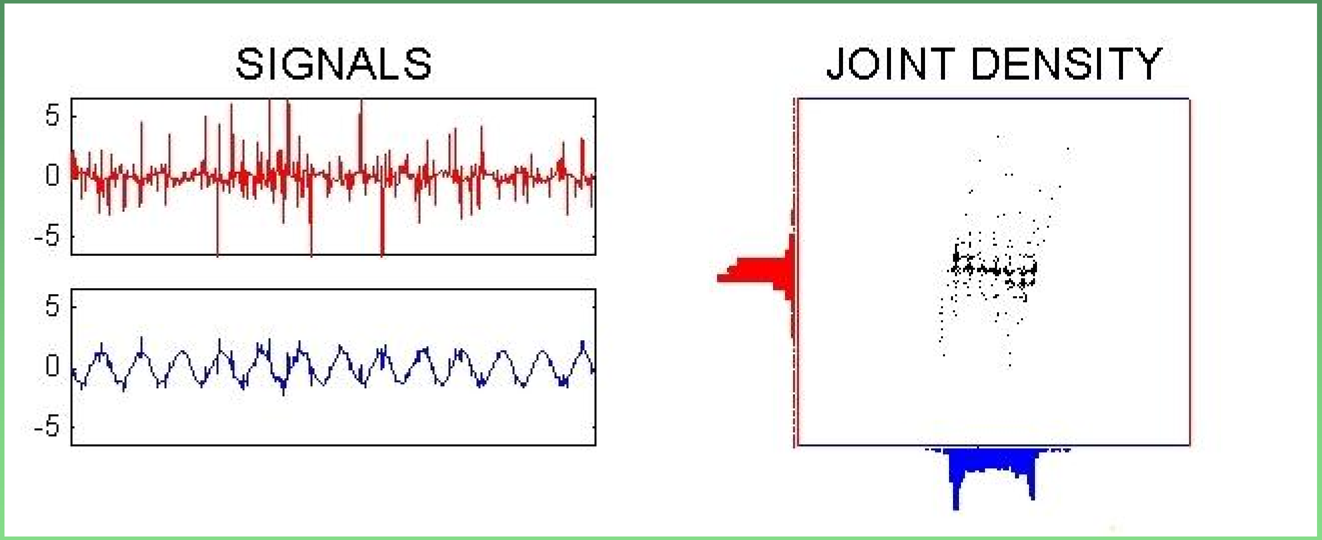
Preprocessing: Whitening->



Steps

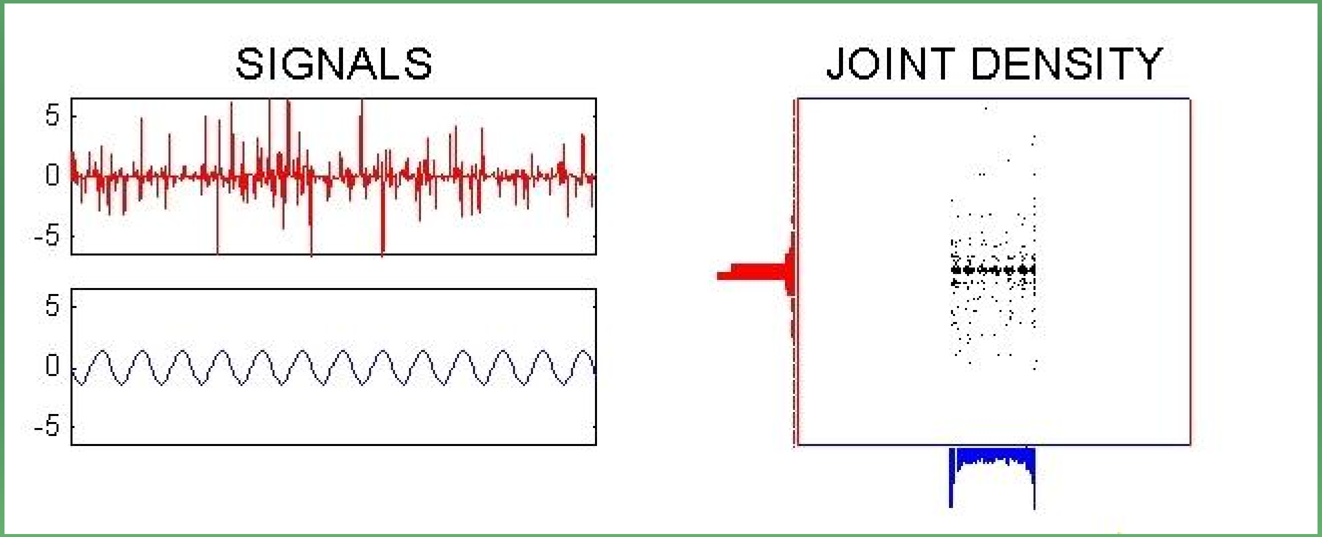


FastICA algorithm, <-first step (rotating begins)



Step 3 (rotating -> continues)

Steps



The last step of the FastICA algorithm (rotating ends)

Matlab Code

- Explain what the PROCEDURES MEAN
- Explain the algorithm on the SOUND MIXTURES.
- 6-7 slides

Usages of ICA

- Separation of Artifacts in MEG (magneto-encephalography) data
- **Finding Hidden Factors in Financial Data**
- Reducing Noise in Natural Images
- **Telecommunications (**CDMA [Code-Division Multiple Access] mobile communications**)**

Sources

- Internet>
  - Wikipedia
- Johan Bylund, Blind signal separation
- A. Hyvärinen, J. Karhunen, E. Oja – Independent Component analysis
- Other useful ICA .pdf files