Deterministic Independent Component Analysis (ICA)

Ruitong Huang   András György   Csaba Szepesvári

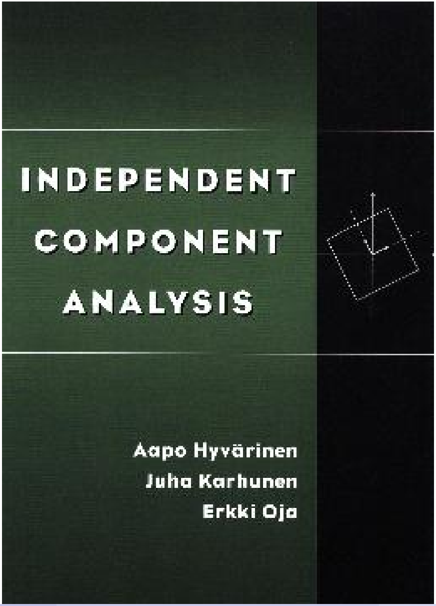
University of Alberta; Imperial College London

December 10, 2015

Outline

1. Introduction
   - What is ICA, really?
2. Deterministic ICA
3. Conclusions

What is Independent Component Analysis (ICA)?



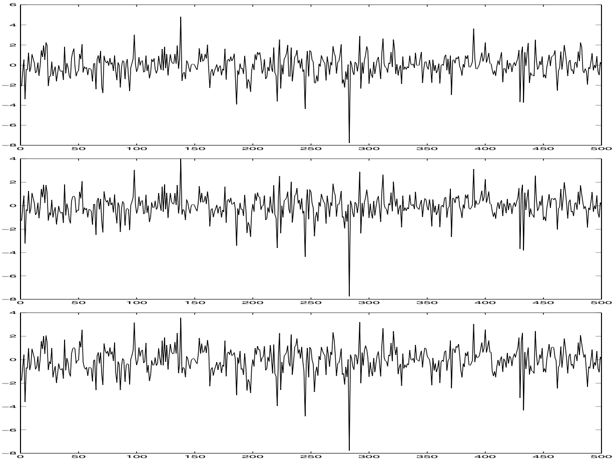


Fig. 1.2 The observed signals that are assumed to be mixtures of some underlying source signals.

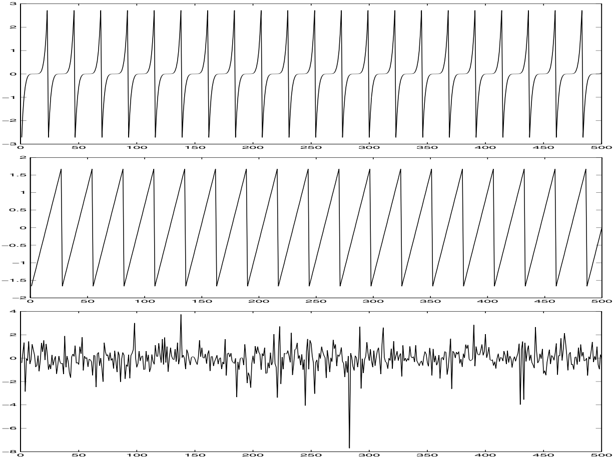


Fig. 1.3 The estimates of the original source signals, estimated using only the observed mixture signals in Fig. 1.2. The original signals were found very accurately.

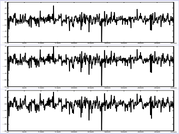


Fig. 1.2 The observed signals that are assumed to be mixtures of some underlying source signals.

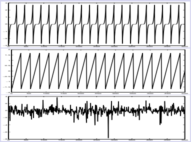


Fig. 1.3 The estimates of the original source signals, estimated using only the observed mixture signals in Fig. 1.2. The original signals were found very accurately.

We were able to estimate the original source signals, using an algorithm that used the information on the independence only. [...] This leads us to the following definition of ICA [...] Given a set of observations of random variables $(x_1(t), x_2(t), \ldots, x_d(t))$, *where t is the time or sample index, assume that they are generated as a linear mixture of independent components:*

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_d(t) \end{pmatrix} = A \begin{pmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_d(t) \end{pmatrix},$$

where A is some unknown matrix.

Good?

Independence?

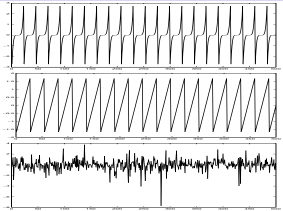


Fig. 1.3 The estimates of the original source signals, estimated using only the observed mixture signals in Fig. 1.2. The original signals were found very accurately.

- Is $s_1(t)$ independent of $s_2(t)$? Sure!
- Any two numbers are independent of each other! All deterministic signal sources are fine then? What if $s_2(t) = 2s_1(t)$?
- Should we be worried about temporal dependencies? No? What if $s_1(t) = s_1(t+1) = \ldots$?
- Can we redefine ICA in a more meaningful way?
- Let's go beyond statistics!

How to go beyond statistical analysis?

1. Perform a deterministic analysis of the algorithm, reducing the problem to perturbation analysis
2. Perform statistical analysis on the size of perturbations when necessary or desired

Let $[T] = \{1, \ldots, T\}$. Sources: $s : [T] \to [-C, C]^d$.
Let $\nu^{(s)}$ be the empirical distribution induced by s; for $B \subset [-C, C]^d$,

$$\nu^{(s)}(B) = \frac{1}{T}|\{t \in [T] : s(\tau) \in B\}|.$$

- Measure of independence: $D_4(\nu^{(s)}, \mu)$, $D_4^{(d,d)}(\nu^{(As,\epsilon)})$;
- Measure of Gaussianness: $\kappa(\nu^{(\epsilon)})$;
- Measure of Zero-Mean: $N(\nu^{(\epsilon)})$, $N(\nu^{(s)})$

Result

There exists a randomized algorithm such that for any $A \in \mathbb{R}^{d\times d}$, and $x, s, \epsilon : [T] \to \mathbb{R}^d$ satisfying $x(t) = As(t) + \epsilon(t)$, the algorithm returns $\hat{A}$ such that:

- The computational complexity is $O(d^3 T)$;
- With high probability,

$$d(\hat{A}, A) \le \inf_{\mu\in\Pi_0} \mathcal{C}(\mu) \min\Big(D_4(\nu^{(s)}, \mu) + \kappa(\nu^{(\epsilon)}) + D_4^{(d,d)}(\nu^{(As,\epsilon)}) + N(\nu^{(\epsilon)}) + N(\nu^{(s)}), \Theta(\mu)\Big),$$

- Here,$\mathcal{C}(\mu)$ and $\Theta(\mu)$ are problem dependent, polynomial in the parameters.

Conclusions

- Independent Component Analysis without probabilities!
- Deterministic analysis: Cleaner, more general, should do it more often! Limits?
- New method: DICA. Universal, strong guarantees.