Shared Independent Component Analysis for Multi-Subject Neuroimaging

Hugo Richard, Pierre Ablin, Alexandre Gramfort, Bertrand Thirion, Aapo Hyvarinen

NeurIPS @ Paris 2021

Github: @hugorichard

Twitter: @hugorichard

https://hugorichard.github.io

Independent component analysis (noise-free)

ICA model (Jutten, 1991)

- Independent *sources*: $\mathbf{s} \in \mathbb{R}^p$

$p(\mathbf{s}) = p(s_1) \cdots p(s_p)$

- *Sensors*: $\mathbf{x} \in \mathbb{R}^p$

$$\mathbf{x} = A\mathbf{s}$$

where A is the *Mixing matrix*.

GroupICA

Consider m views $\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{R}^p$ such that

$$\forall i \in \{1, \ldots, m\}, \;\; \mathbf{x}_i = A_i\mathbf{s} + \mathbf{n}_i$$

State of the art

The most popular

- ConcatICA [Calhoun, 2001]
- CanICA [Varoquaux, 2010]

Some other related work

- IVA [Lee, 2008]
- Unified approach [Guo, 2008]
- SRM [Chen, 2015]
- MultiViewICA [Richard, 2020]

Our contribution: Shared ICA (ShICA)

ShICA model

$$\mathbf{x}_i = A_i(\mathbf{s} + \mathbf{n}_i), i = 1, \dots, m$$

- $\mathbf{n}_i \sim \mathcal{N}(0, \Sigma_i)$ where Σ_i is diagonal positive.
- $\mathbf{s}$ are independent components some of which may be Gaussian

ShICA-J:

- A fast algorithm based on joint diagonalization and multiset CCA

ShICA-ML

- A maximum likelihood algorithm for ShICA

Our contribution: Shared ICA (ShICA)

ShICA model

Informal theorem (Identifiability): ShICA is identifiable if and only if Gaussian components are "sufficiently diverse".

ShICA-J

1. (Theorem) Apply MultisetCCA framed as an eigenproblem to data following ShICA: the first p eigenvectors yield the unmixing matrices.
2. This fails in practice but joint diagonalization solves the issue.

ShICA-ML

$$\mathbf{x}_i = A_i(\mathbf{s} + \mathbf{n}_i)$$

where $\mathbf{n}_i \sim \mathcal{N}(0, \Sigma_i)$, Σ_i diagonal and $s_j \sim \frac{1}{2}\sum_{\alpha \in \{\frac{1}{2}, \frac{3}{2}\}} \mathcal{N}(0, \alpha)$.
Solved by EM: closed form and efficient E-step.

Experiments

Experiments on synthetic data

- ShICA-ML can separate both Gaussian and non-Gaussian sources
- ShICA-J provide a good computation time / performance ratio

Experiments on real data

- fMRI data: ShICA-ML better reconstructs the data of left-out subjects from data of other subjects.
- MEG data: ShICA-ML yields an improved localization and identification of common sources.

Our paper: https://openreview.net/pdf?id=24-ZY0UZVKD
Our code: https://github.com/hugorichard/ShICA