University of BRISTOL

Overcomplete and Undercomplete ICA

Rachel Wood

24th February 2023

Definitions

As in the lecture notes we use $X \in \mathbb{R}^{M \times T}$ to represent the (centered) observed signals in time and $S \in \mathbb{R}^{N \times T}$ for the source signals at each point in time

Overcomplete ICA:

- More sources than sensors ($N > M$)

Undercomplete ICA:

- Fewer sources than sensors ($M < N$)

Assumptions:

- $X = BS$ for some full rank $B \in \mathbb{R}^{M \times N}$
- $\mathbb{E}[S] = 0$ and Var$(S) = \mathrm{I}_p$
- S is non-Gaussian
- B is time-invariant

Goal: Find $\hat{S}$ and the un-mixing matrix W such that $\hat{S} = WX$ is a reasonable estimate for S

Overcomplete ICA

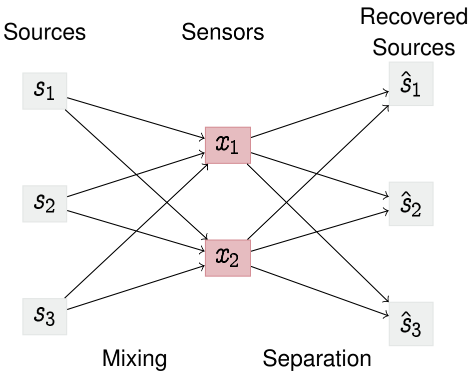

Figure: Example of Overcomplete ICA

- When $M = N$, we often assume the existence of $W = B^{-1}$.
- For $M < N$, reframe the problem as a mixing problem rather than a separation one (i.e estimate B rather than finding W)
- We assume S is sparse (we will expand on this later)

Overcomplete ICA: Maximum a Posteriori

We use a maximum a posteriori approach:

$$\max_{B,S} P(B,S|X) \propto \max_{B,S} P(X|B,S)P(B)P(S)$$

where we assume $P(A)$ to be uniform and model S to be a classical Lagrangian r.v.

$$P(X|B,S) \propto \prod_i \exp\left(-\frac{(x_i-(BS)_i)^2}{2\sigma^2}\right), \qquad P(S) \propto \prod_{i,t} \exp\left(|s_i^t|\right)$$

We also take the logarithm and invert the sign to get

$$\max_{B,S} P(B,S|X) \propto \min_{B,S} \frac{\|BS-X\|^2}{2\sigma^2} + \sum_{i,t} |s_i^t|$$

Overcomplete ICA: Signal Recovery

Since we have assumed $X = BS$ with no noise, we can simplify our optimisation problem to T individual optimisation problems:

$$\text{Minimise} \sum_i |s_i^t|$$

$$\text{Subject to } x^t = Bs^t$$

where B is estimated using external optimisation

$$\hat{B} = \min_B \sum_{i,t} |s_i^t(B)|$$

Overcomplete ICA: Sparsity

It is often not a reasonable assumption for the components of S to be sparse, however we can linearly transform S to obtain a sparse matrix Φ [1]. We say the elements of Φ are atoms of a dictionary and we write

$$s_i = \sum_{t=1}^{T} C_i^t \varphi_t$$

We can then replace S with Φ in our previous steps and recover $\hat{S}$ from the resulting estimate $\hat{\Phi}$:

$$\hat{S} = C\hat{\Phi}$$

Undercomplete ICA

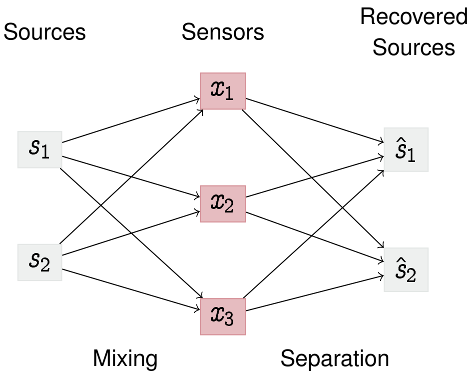

Figure: Example of Undercomplete ICA

When $M > N$ the linear system $X = BS$ is overdetermined, so we can't estimate B as in the lecture notes. We instead use the following steps:

1. Perform PCA to obtain an N-dimensional reduction of the dataset X
2. Use a standard ICA algorithm on the reduced dimensional data

Undercomplete ICA:PCA Step

Goal To find Y such that

$$Y = AX, \text{ where } \mathrm{Var}(Y) = \Lambda \text{ is diagonal}$$

We consider the SVD of B:

$$B = U\Sigma V^T = U \begin{bmatrix} \tilde{\Sigma} \\ 0 \end{bmatrix} V^T$$

where $\Sigma = \mathrm{diag}(\sigma_1, ..., \sigma_N, 0, ..., 0) \in \mathbb{R}^{M\times M}$ and $\tilde{\Sigma} = \mathrm{diag}(\sigma_1, ..., \sigma_N) \in \mathbb{R}^{N\times N}$ for $\sigma_1 > ... > \sigma_N > 0$ and U, V are orthogonal.

Undercomplete ICA:PCA Step

Then the eigen decomposition of the correlation matrix of X is

$$R_X = BR_S B^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma^2 U^T$$

where $\Sigma^2 = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_N^2, 0, \ldots, 0) \in \mathbb{R}^{M\times M}$

If we take

$$A = \begin{bmatrix} A_{1:N} \\ A_{N+1:M} \end{bmatrix} = U^T = U^{-1}, \text{ then } R_Y = AR_X A^T = \Sigma^2$$

and so the elements of Y are uncorrelated.

Undercomplete ICA

Now we can perform ICA on the N-dimensional reduction of the data

$$Y_{1:N} = A_{1:N}X$$

and obtain the estimated signal matrix $\hat{S} = WA_{1:N}X$

References I

- M. Stéphane, “Chapter 1 - sparse representations,” in *A Wavelet Tour of Signal Processing (Third Edition)* (M. Stéphane, ed.), pp. 1–31, Boston: Academic Press, third edition ed., 2009.
- G. Naik and D. Kumar, “An overview of independent component analysis and its applications,” *Informatica*, vol. 35, pp. 63–81, 01 2011.
- P. Bofill and M. Zibulevsky, “Blind separation of more sources than mixtures using sparsity of their short-time fourier transform,” *Proc. ICA2000*, pp. 87–92, 01 2000.
- P. Bofill and M. Zibulevsky, “Underdetermined blind source separation using sparse representations,” *Signal Processing*, vol. 81, no. 11, pp. 2353–2362, 2001.