Syntheses Hyvarinen

Nathan de Lara

November 27, 2015

1 Results

ICA Assume that we observe n linear mixtures $x_1, ..., x_n$ of n independent components:

$$x_j = a_{j1}s_1 + a_{j2}s_2 + ... + a_{jn}s_n, \forall j$$

where all x_j, s_j are random variables with zero mean and the s_j are statistically independent. Equivalently:

$$x = As$$

Given the entry x, the goal is to estimate both A and s.

Constraint The vector s must be *nongaussian*. One can prove that the distribution of any orthogonal transformation of the Gaussian (x_1, x_2) has exactly the same distribution as (x_1, x_2), and that x_1 and x_2 are independent. Thus, in the case of Gaussian variables, we can only estimate the ICA model up to an orthogonal transformation. In other words, the matrix A is not identifiable for Gaussian independent components.

Notations $y = w^T x$ and $z = A^T w$ such that: $y = z^T s$

Idea Use the Central Limit Theorem to determine w so that it would equal one of the rows of the inverse of A. Choose w such that it corresponds to a z with only one nonzero coordinate. Consequently, $z^T s$ is equal to one of the wanted components up to a multiplicative sign.

Kurtosis This quantity defined for a zero-mean random variable with unit variance by:

$$kurt(y) = \mathbb{E}(y^4) - 3(\mathbb{E}(y^2))^2$$

is commonly used to measure nongaussianity as is it zero for a Gaussian variable and often nonzero for a nongaussian variable. Besides, for independent variables:

$$kurt(x_1 + x_2) = kurt(x_1) + kurt(x_2)$$
$$kurt(\alpha x_1) = \alpha^4 kurt(x_1)$$

Problem formulation

$$maximize\ \ |kurt(y)|$$
$$st\ \ y = z^T s$$
$$and\ \ Var(y) = 1$$

The drawback is that the kurtosis measure is not robust.

Negentropy Another measure of the nongaussianity of a random variable is its negentropy:

$$J(y) = H(y_{gauss}) - H(y)$$

Where y_{gauss} is the Gaussian random variable with the same covariance matrix than y. The entropy has the property of being maximal for a Gaussian variable among all variables of equal variance. Thus, the negentropy is zero for a Gaussian variable and nonnegative for other variables. Besides, negentropy is invariant for invertible linear transformations. in this form, the problem consists in maximizing the negentropy.

Preprocessing for ICA

- Centering
- Whitening
- Dimension reduction (PCA)
- Band-Pass filtering

2 FastICA

Idea Use Gram-Schmidt-like decorrelation to compute independent w_i.

Algorithm Assuming that the data is whitened, initialize W at random and compute for any norm different of the Frobenius norm:

1. $W = W/\sqrt{||WW^T||}$
2. repeat until convergence: $W = \frac{3}{2}W - \frac{1}{2}WW^TW$

3 Applications

- Separation of Artifacts in MEG Data
- Finding Hidden Factors in Financial Data (data available on the challenges website)
- Reducing Noise in Natural Images