Fourth Moments and Independent Component Analysis

Jari Miettinen[1], **Klaus Nordhausen**[2], **Hannu Oja**[2] **and Sara Taskinen**[1]
[1] University of Jyväskylä, Finland
[2] University of Turku, Finland

In the basic independent component model the observable p-variate random vector $\mathbf{x} = (x_1, \ldots, x_p)'$ is a linear mixture of p-variate latent source vector $\mathbf{z} = (z_1, \ldots, z_p)'$, whose components are non-Gaussian and mutually independent. The model is then written as

$$\mathbf{x} = \mu + \Omega\mathbf{z},$$

where the unknown $p \times p$ mixing matrix Ω is invertible and the p-vector μ is a location parameter. Under these assumptions, Ω (or the unmixing matrix $\Gamma = \Omega^{-1}$) can be estimated based on a random sample $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ from the distribution of $\mathbf{x}$, but only up to scales, signs and permutations of the columns of Ω (or rows of Γ). Hence, we may assume that $\mathbf{z}$ is standardized, that is, $E(\mathbf{z}) = \mathbf{0}$ and $E(\mathbf{z}\mathbf{z}') = I_p$.

Many classical independent component analysis functionals use fourth moments to separate the components. FOBI (Fourt Order Blind Identification) (Cardoso, 1989) is one of the oldest and simplest methods, where the covariance matrix and a matrix based on fourth moments are diagonalized simultaneously. The FOBI estimate is consistent only if the source components have distinct kurtosis values. JADE (Joint Approximate Diagonalization of Eigen matrices) (Cardoso and Souloumiac, 1993) uses approximate joint diagonalization of fourth order cumulant matrices and is able to separate also identically distributed components. On the other hand, computation of $p(p+1)/2$ matrices of size $p \times p$ and joint diagonalization of them makes JADE impractical for very high-dimensional data. Also the original deflation-based and symmetric FastICA estimates (Hyvärinen and Oja, 1997) employ fourth moments, and like JADE, require that at most one of the components has zero kurtosis value.

Asymptotic properties of the FOBI estimate were studied in Ilmonen et al. (2010), and the deflation-based FastICA estimate was considered in Ollila (2010) and Nordhausen et al. (2011). Here we give the corresponding results for JADE and symmetric FastICA. All these estimates are affine equivariant and the asymptotic variances of the kl:th elements of $\hat{\Gamma}\Omega$ depend only on the kth and lth source components, with the exception of FOBI, where the other components have small effect, too. We compare asymptotic variances of the different estimates, when the source components come from the families of the exponential power distributions and the chi-squared distributions.

The comparison reveals some surprising asymptotic equalities between the estimates.

Keywords: Asymptotic properties, FastICA, FOBI, JADE.

References:

Cardoso, J.F. (1989). Source separation using higher order moments. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Glasgow, UK*, 2109–2112

Cardoso, J.F., A. Souloumiac (1993). Blind beamforming for non Gaussian signals. *IEE Proceedings-F*, *140*, 362–370.

Hyvärinen, A. and E. Oja (1997). A fast fixed-point algorithm for independent component analysis. *Neural Computation*, *9*, 1483–1492.

Ilmonen, P., J. Nevalainen, and H. Oja (2010). Characteristics of multivariate distributions and the invariant coordinate system. *Statistics & Probability Letters*, *80*, 1844–1853.

Nordhausen, K., P. Ilmonen, A. Mandal, H. Oja and E. Ollila (2011). Deflation-based FastICA reloaded. in *Proceedings of the 19th European Signal Processing Conference 2011 (EUSIPCO 2011),* Barcelona, 1854–1858.

Ollila, E. (2010). The Deflation-Based FastICA Estimator: Statistical Analysis Revisited. *IEEE Transactions on Signal Processing*, *58*, 1527–1541.