

R-ESTIMATION FOR ASYMMETRIC INDEPENDENT COMPONENT ANALYSIS

Marc Hallin¹ and Chintan Mehta²

¹ECARES, Université libre de Bruxelles and ORFE, Princeton University

² ORFE, Princeton University

Independent Component Analysis (ICA) recently has attracted much attention in the statistical literature as an appealing alternative to elliptical models. Whereas k -dimensional elliptical densities depend on one single unspecified radial density, however, k -dimensional independent component distributions involve k unspecified component densities. In practice, for given sample size n and dimension k , this makes the statistical analysis much harder. We focus here on the estimation, from an independent sample, of the mixing/demixing matrix of the model. Traditional methods (FOBI, Kernel-ICA, FastICA) mainly originate from the engineering literature. Their consistency requires moment conditions, they are poorly robust, and do not achieve any type of asymptotic efficiency. When based on robust scatter matrices, the two-scatter methods developed by Oja, Sirkiša, and Eriksson (2006) and Nordhausen, Oja, and Ollila (2008) enjoy better robustness features, but their optimality properties remain unclear. The “classical semiparametric” approach by Chen and Bickel (2006), quite on the contrary, achieves semiparametric efficiency, but requires the estimation of the densities of the k unobserved independent components. As a reaction, an efficient (signed-) rank-based approach has been proposed by Ilmonen and Paindaveine (2011) for the case of symmetric component densities. The performance of their estimators is quite good, but they unfortunately fail to be root- n consistent as soon as one of the component densities violates the symmetry assumption. In this paper, using ranks rather than signed ranks, we extend their approach to the asymmetric case and propose a one-step R -estimator for ICA mixing matrices. The finite-sample performances of those estimators are investigated and compared to those of existing methods under moderately large sample sizes. Particularly good performances are obtained from a version involving data-driven scores taking into account the skewness and kurtosis of residuals. Finally, we show, by an empirical exercise, that our methods also may provide excellent results in a context such as image analysis, where the basic assumptions of ICA are quite unlikely to hold.

Keywords: Independent Component Analysis (ICA), local asymptotic normality (LAN), ranks, R -estimation, robustness.

References:

Hallin, M. and Mehta, C. (2013). R-Estimation for Asymmetric Independent Component Analysis. arXiv:1304.3073v2.