ICA with symmetrized M-estimators of scatter

S. Sirkiä[1] and H. Oja[2]

[1] Department of Mathematics and Statistics, P.O.Box 35, FIN-40014 University of Jyväskylä, Finland
[2] Tampere School of Public Health, FIN-33014 University of Tampere, Finland

Abstract

In this talk we introduce a family of so called symmetrized M-estimators of scatter which are defined to be M-estimators calculated from pairwise differences of the observed data. Such an operation is well known in the univariate case and it is sometimes used in order to make the distribution symmetric, with location at 0. In the multivariate case all marginal distributions become symmetric with location at 0. It does not, however, make the whole distribution for example elliptically symmetric.

Using pairwise differences is still useful for many reasons. Estimators of scatter in the elliptically symmetric family usually require the location either be known or estimated simultaneously. With the estimators at hand, this is simplified as the location of the pairwise differences is always the origin. Also, the efficiency increases, which is intuitively expected due to the "increase" of sample size. We give the asymptotic and small sample efficiencies using certain weigth functions.

Third and maybe the most interesting point is that a symmetrized M-estimator of scatter matrix has the so called independence property: it is diagonal when the components of the original random vector are independent, which is not true for ordinary M-estimators. This property can be used in independent component analysis, or ICA. Briefly, ICA concerns the problem of finding a random vector with independent components when only an unknown linear transformation is observed. Symmetrized M-estimators of scatter produce a family of ICA-estimators. This is also outlined in the talk, together with some simulations.