EXAMINATION OF INDEPENDENCE IN INDEPENDENT COMPONENT ANALYSIS

S. Shimizu and Y. Kano

OSAKA UNIVERSITY, JAPAN
shimizu@hus.osaka-u.ac.jp, kano@hus.osaka-u.ac.jp

The well-known factor analysis model (FA) for an m-dimensional observed vector $\mathbf{x}$ is defined as

$$\mathbf{x} = A\boldsymbol{s} + \boldsymbol{\epsilon} \tag{1}$$

where A is an unknown constant matrix of factor loadings, $\boldsymbol{s}$ is an n-dimensional common factor with $V(\boldsymbol{s}) = I_n$. and $\boldsymbol{\epsilon}$ is a noise vector. If the components of $\boldsymbol{s}$ are nonnormally but independently distributed, this model is said to be the noisy ICA or Independent FA(Attias, 1998). If one further assumes that $\boldsymbol{\epsilon} = \mathbf{0}$, (1) is said to be the noise-free ICA(Comon, 1994).

Since the ICA was proposed in the field of informatics around 1990, many procedures for estimation problem have been developed, some of which are based on maximization of kurtosis or negentropy (e.g., Comon, 1994; Hyvärinen, 1997,1999). On the other hand, Mooijaart(1985) employed the generalized least squares (GLS) approach to estimate parameters in the noisy ICA, that is,

$$T_{AB} = \min \left\| \begin{bmatrix} \boldsymbol{m}_2 \\ \boldsymbol{m}_3 \end{bmatrix} - \begin{bmatrix} \boldsymbol{\sigma}_2(\boldsymbol{\theta}) \\ \boldsymbol{\sigma}_3(\boldsymbol{\theta}) \end{bmatrix} \right\|^2 \tag{2}$$

where $\boldsymbol{m}_2, \boldsymbol{m}_3$ are the vectorized second and third order moments and $E(\boldsymbol{m}_2) = \boldsymbol{\sigma}_2(\boldsymbol{\theta})$, and $E(\boldsymbol{m}_3) = \boldsymbol{\sigma}_3(\boldsymbol{\theta})$. The parameter $\boldsymbol{\theta}$ consists of the model parameters including A.

It is very important to examine the independence of the latent variables $\boldsymbol{s}$ for the case where the independence assumption does not obviously hold. This case often happens in social sciences. If the independence assumption fails, the ICA is nothing but projection pursuit in multivariate analysis. The residual T_{AB} could be used to test the independence assumption while Mooijaart (1985) mentioned nothing about it. Murata (1999) used empirical characteristic functions to test the independence assumption.

The GLS approach by Mooijaart can be generalized. Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be a sample from the ICA models, and let $\boldsymbol{h}(\mathbf{x})$ (and also $\boldsymbol{h}_i(\mathbf{x})$) be an integrable vector-valued function. We consider the following objective function to perform ICA:

$$T_1 = \min \|\bar{\boldsymbol{h}} - E_{\boldsymbol{s}}(\boldsymbol{h}(\mathbf{x}))\|^2, \tag{3}$$

where $\bar{\boldsymbol{h}}$ is the empirical mean vector defined as $\frac{1}{N}\sum_i \boldsymbol{h}(\mathbf{x}_i)$ and $E_{\boldsymbol{s}}(\cdot)$ denotes expectation under the assumption that the components of $\boldsymbol{s}$ be independent of each other. Here $\boldsymbol{h}_{\mathbf{x}}$ is not quadratic. Alternatively we can consider

$$T_2 = \min \|\bar{c}_{h_1,h_2} - Cov(\boldsymbol{h}_1(\mathbf{x}), \boldsymbol{h}_2(\mathbf{x}))\|^2, \tag{4}$$

where $\bar{c}_{h_1,h_2}$ is the empirical covariance matrix. To construct statistically optimum testing procedure, we have to consider $N \times T_1$ or $N \times T_2$ and choose an appropriate weight matrix. Of course, T_1 and T_2 depend on the choice of $\boldsymbol{h}(\mathbf{x})$. One can consider a class of function $\boldsymbol{h}(\mathbf{x})$ and take the supremum of T_1 or T_2 over the function class.

The procedure above can also be considered as an estimation method in ICA and they are related to many existing estimation methods in ICA. Hyvärinen (1999) proposed choosing weight vectors which locates empirical means as far as possible from the expectation under normality. Our approach locates empirical means as close as possible to the expectation under independence assumption. Both are also closely related.

Key words: independence, nonnormality, social sciences

References

Attias, H. (1998). Independent factor analysis. *Neural Computation, 11,* 803-851.

Comon, P. (1994). Independent component analysis, A new concept? *Signal Processing, 36,* 287–314.

Hyvärinen, A. and Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural Computation, 9,*(2) 1483–1492.

Hyvärinen, A. (1999). Fast and robust fixed-point algorithm for independent component analysis. *IEEE Trans. on Neural Networks, 10,*(3) 626–634.

Mooijaart, A. (1985). Factor analysis for non-normal variables. *Psychometrika, 50,* 323–342.

Murata, N. (1999). Properties of empirical characteristic function and its application to test for independence and independent component analysis. *Workshop on Information-Based Induction Sciences.*