Linear Dimension Reduction

Hannu Oja

University of Turku, Finland

Abstract

In this talk, unsupervised and supervised linear dimension reduction methods based on various uses of scatter matrices are addressed. In our approach, dimension reduction is done by preprocessing the data, that is, by transforming the data to an invariant coordinate system (ICS). The unsupervised ICS functionals in Tyler et al. (2009) are based on the use of two unsupervised scatter matrix functionals while the supervised functionals such as the sliced inverse regression (SIR) utilize both unsupervised and supervised scatter matrices. See Liski et al. (2013) and Ilmonen et al. (2012). The use of several scatter functionals is discussed as well.

In the so called Blind Source Separation (BSS) model, it is assumed that the observed p-variate observations are all linear combinations of p latent variables. This may be seen also as a dimension reduction problem if only few of the latent variables are relevant in data analysis. The aim is then to find an estimate for an unmixing matrix, which transforms the observed variables back to the latent variables. The most popular BSS approach is independent component analysis (ICA), and the so called AMUSE and SOBI solutions are used to find latent stationary time series in the analysis of multivariate time series. Unmixing matrix estimates are often found as ICS or SICS solutions. See e.g. Miettinen et al. (2014).

Typically, dimension reduction procedures proposed in the literature are highly non-robust as the first step in the algorithms is to standardize the observations using the regular covariance matrix. In this talk, we also briefly discuss the possibility to robustify these procedures. The talk is based on joint work with several coauthors.

Keywords

Blind Source Separation; Independent component analysis; Invariant coordinate system; Sliced inverse regression; Time series

References

Tyler, D., Critchley, F., Dümbgen, L. & Oja, H. (2009). Invariant coordinate selection. *Journal of Royal Statistical Society B*, **71**, 549–592.

Ilmonen, P., Serfling, R. & Oja, H. (2012). Invariant coordinate selection (ICS) functionals. *International Statistical Review*, **80**, 93–110.

Liski, E., Nordhausen, K., and Oja, H. (2013). Supervised invariant coordinate selection. *Statistics*.

Miettinen, J., Nordhausen, K., Oja, H. & Taskinen, S. (2014). Deflation-based separation of uncorrelated stationary time series. *Journal of Multivariate Analysis*, **123**, 214–227