MARKOV COMPONENT ANALYSIS AND FUNCTIONAL REGRESSION

Bo Markussen[1]
[1] Department of Mathematical Sciences, University of Copenhagen, Denmark

Decomposition of data into underlying components is one of the most used techniques in multivariate statistics and functional data analysis. *Principal Component Analysis (PCA)* is the foremost among the decomposition techniques. For a sample of multivariate functional data $X_n\colon [a,b] \to \mathbb{R}^d$ for $n = 1, \ldots, N$ functional PCA into q components is given by

$$X_n(t) = \sum_{k=1}^{q} S_{nk}\phi_k(t) + \text{error}.$$

The scores $S_{nk} \in \mathbb{R}$ can be interpreted as stochastically independent random variables (Tipping & Bishop, 1999), and the loadings $\phi_k\colon [a,b] \to \mathbb{R}^d$ are deterministic functions. In this talk we propose an alternative decomposition that we call *Markov Component Analysis (MCA)*. This is given by

$$X_n(t) = \sum_{k=1}^{q} Z_{nk}(t) + \text{error}.$$

Here the components $Z_{nk}\colon [a,b] \to \mathbb{R}^d$ are stochastically independent Markov processes. The covariance operators for Gaussian Markov processes can be parametrized by a factorizable structure that extends the concept of loadings from PCA. From a modelling point of view this structure is more versatile, while it retains many of the nice computational properties known from PCA.

In this talk we will compare MCA to PCA, present an algebra that allows for efficient computation within the MCA framework, and apply MCA to a functional regression problem of predicting horse lameness from three dimensional acceleration signals.

Keywords: Functional Data Analysis, Multivariate Analysis, Covariance estimation, Decomposition methods, Sparse representation, Functional Regression.

References:

Tipping, M. E. and Bishop, C. M. (1999). Probabilistic principal component analysis. *J. R. Statist. Soc. B*, *61*, 611–622.

Sørensen, H., Tolver, A., Thomsen, M. H., and Andersen P. H. (2012). Quantification of symmetry for functional data with application to equine lameness classification. *J. Appl. Stat.*, *39*, 337–360.