Dimension reduction through local principal curves and manifolds

Jochen Einbeck

Durham University, Durham City, UK jochen.einbeck@durham.ac.uk

Principal curves are today a well-established concept, and the researcher can choose from a variety of concepts for their estimation. Surprisingly, most research has concentrated on the algorithms to fit those curves, but has not proceeded with exploiting their value once they are fitted. In situations where the intrinsic dimensionality of the predictor space of a multivariate regression problem is approximately one-dimensional, it is a straightforward idea to use the principal curve as a compressed representation of the predictor space, hence reducing a multivariate to a univariate regression problem (which then may be dealt with using any nonparametric smoother). However, the implementation of this technique is not trivial, and not all principal curve algorithms are suitable for this task. We will use local principal curves to implement this technique, apply them to astrophysical data from the Galactic Survey mission GAIA, and compare their performance to competing dimension reduction or model selection techniques. The extension to local principal manifolds, and their use as a multivariate, but low-dimensional, nonparametric summary of a high-dimensional predictor space, will also be considered.