Generalized Significance in Scale Space: The GS3 Package

Daniel V. Samarov

In traditional approaches to multivariate nonparametric regression the focus is on the estimation of a single optimal bandwidth matrix. An alternative approach which has gained increased attention in the statistics community is based on scale-space theory from computer vision, which takes into consideration a family of smooths. The intuition behind the latter approach is that different levels of smoothness reveal different, potentially significant features about the data. While in the multivariate case much work has been done in determining feature significance (such as local extrema), little work has been done on how to select a sensible family of smooths. In this talk we present the GS3 R package which provides an implementation of a novel approach for finding such a family in the multivariate nonparametric regression setting based on the rodeo algorithm. This methodology, which we will refer to as Generalized Significance in Scale Space (or GS3) is motivated by and applied to the estimation of aerosol extinction in the atmosphere and hyperspectral medical imaging.

References

Chaudhuri, P., Marron, J.S., Scale-space view of curve estimation, Annals of Statistics 28 (2000) 408-428.

Duong, T., Cowling, A., Koch, I., Wand, M.P., Feature significance for multivariate kernel density estimation, Computational Statistics and Data Analysis 52 (2008) 4225-4242.

Godtliebsen, F., Marron, J.S., Chaudhuri, P., Statistical Significance of features in digital images, Image and Vision Computing 22 (2004) 1093-1104.