Sparse Functional Boxplots for Multivariate Curves

Zhuo Qu[1]; Marc G. Genton[1]

[1] Statistics Program, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia

Abstract:
This paper introduces the sparse functional boxplot and the intensity sparse functional boxplot as practical exploratory tools that make visualization possible for both complete and sparse functional data. These visualization tools can be used either in the univariate or multivariate functional setting. The sparse functional boxplot, which is based on the functional boxplot, depicts sparseness characteristics in the envelope of the 50% central region, the median curve, and the outliers. The proportion of missingness at each time index within the central region is colored in gray. The intensity sparse functional boxplot displays the relative intensity of sparse points in the central region, revealing where data are more or less sparse. The two-stage functional boxplot, a derivation from the functional boxplot to better detect outliers, is also extended to its sparse form. Several depth proposals for sparse multivariate functional data are evaluated and outlier detection is tested in simulations under various data settings and sparseness scenarios. The practical applications of the sparse functional boxplot and intensity sparse functional boxplot are illustrated with two public health datasets.

Keywords:
Depth; Multivariate functional data; Outlier detection; Sparse functional data; Visualization