Title:

Feature Space Reduction via Dimension Expansion of High-Dimensional Functional Covariates for Prediction

Abstract:
Due to improvements in technology, the number of high-frequency signals that are collected for prediction of various processes is increasing at a dramatic rate. In many cases, these “big data” covariates have relevant predictive features at different scales or support than the response. In the case where the covariate signal is measured over time, this can be viewed as a mixed frequency model, as it relates features in high-frequency data observed over some period to a response on some other scale. We have found that in many instances one can obtain much more parsimonious representations of such covariates by transforming them to a higher-dimensional space – thereby achieving feature space dimension reduction through an initial dimension expansion. Here we will focus on the case where non-stationary one-dimensional time series covariates are transformed to the time-frequency domain where they can be considered as two-dimensional “images,” and their functional representation provides a mechanism on which we do stochastic variable selection. To account for uncertainty in data, process and parameters, we develop this framework in a Bayesian hierarchical setting. We demonstrate via applications in ecology and finance that we can obtain successful prediction through implicit model averaging. In addition, in many cases we obtain scientifically plausible feature extraction.