Sufficient Dimension Reduction in Supervised Forecasting with Many Predictors

Pedro Isaac Chavez Lopez

November 26, 2023

Abstract

We propose a supervised forecasting method that estimates the conditional mean of a time series target variable non-linearly under the presence of a large set of highly correlated covariates. The target variable follows a multi-index model that depends explicitly on the latent factors from the set of predictors. The indices and the relevant unobservable common factors are estimated consistently using the Three Pass Regression Filter (3PRF) of Kelly and Pruitt (2015), given that the number of indices and the number of relevant factors for each index are specified. Our approach considers two stages to reduce the dimension of the covariates: firstly, through latent factors, and secondly, through sufficient dimension reduction, i.e., the projected indices. Our method extends the sufficient dimension reduction to high-dimensional regimes by reducing the cross-sectional information through factor models in a supervised setting. We first generate artificial targets following the Sliced Inverse Regression (SIR). Next, we implement the 3PRF for each target and the set of predictors. Our approach correctly estimates the projection indices of the underlying factors even in the presence of a non-parametric forecasting function and it only estimates the relevant factors of the target. The resulting forecasts are consistent for the nonlinear sufficient infeasible best forecast when both the time dimension and cross section dimension become large. We provide asymptotic results and show the good finite sample performance through simulations. Finally, we confirm the forecasting performance relative to alternatives with empirical applications using US macroeconomic data.

***Keywords**: Supervised forecasting, factor model, sliced inverse regression, learning indices.*