IPAD: Stable Interpretable Forecasting with Knockoffs Inference

Yoshimasa Uematsu

Tohoku University
E-mail: uematsu0911@gmail.com

Abstract: Interpretability and stability are two important features that are desired in many contemporary big data applications arising in statistics, economics, and finance. While the former is enjoyed to some extent by many existing forecasting approaches, the latter in the sense of controlling the fraction of wrongly discovered features which can enhance greatly the interpretability is still largely underdeveloped. To this end, in this article, we exploit the general framework of model-X knockoffs introduced recently in Candès, Fan, Janson and Lv [(2018), "Panning for Gold: 'model X' Knockoffs for High Dimensional Controlled Variable Selection," Journal of the Royal Statistical Society , Series B, 80, 551–577], which is nonconventional for reproducible large-scale inference in that the framework is completely free of the use of p-values for significance testing, and suggest a new method of intertwined probabilistic factors decoupling (IPAD) for stable interpretable forecasting with knockoffs inference in high-dimensional models. The recipe of the method is constructing the knockoff variables by assuming a latent factor model that is exploited widely in economics and finance for the association structure of covariates. Our method and work are distinct from the existing literature in which we estimate the covariate distribution from data instead of assuming that it is known when constructing the knockoff variables, our procedure does not require any sample splitting, we provide theoretical justifications on the asymptotic false discovery rate control, and the theory for the power analysis is also established. Several simulation examples and the real data analysis further demonstrate that the newly suggested method has appealing finite-sample performance with desired interpretability and stability compared to some popularly used forecasting methods.