

Data Science Seminar

Hosted by the Department of Mathematics and Statistics

- Date: Tuesday, November 1, 2022
- Time: 12:00pm - 1:00pm
- Room: Via Zoom
- Speaker: Dr. Jinchi Lv (University of Southern California)
- Title: High-Dimensional Knockoffs Inference for Time Series Data.

Abstract

The recently introduced framework of model-X knockoffs provides a flexible tool for exact finite-sample false discovery rate (FDR) control in variable selection in arbitrary dimensions without assuming any dependence structure of the response on covariates. It also completely bypasses the use of conventional p-values, making it especially appealing in high-dimensional nonlinear models. Existing works have focused on the setting of independent and identically distributed observations. Yet time series data is prevalent in practical applications in various fields such as economics and social sciences. This motivates the study of model-X knockoffs inference for time series data. In this paper, we make some initial attempt to establish the theoretical and methodological foundation for the model-X knockoffs inference for time series data. We suggest the method of time series knockoffs inference (TSKI) by exploiting the idea of subsampling to alleviate the difficulty caused by the serial dependence. We establish sufficient conditions under which the original model-X knockoffs inference combined with subsampling still achieves the asymptotic FDR control. Our technical analysis reveals the exact effect of serial dependence on the FDR control. To alleviate the practical concern on the power loss because of reduced sample size cause by subsampling, we exploit the idea of knockoffs with copies and multiple knockoffs. Under fairly general time series model settings, we show that the FDR remains to be controlled asymptotically. To theoretically justify the power of TSKI, we further suggest the new knockoff statistic, the backward elimination ranking (BE) statistic, and show that it enjoys both the sure screening property and controlled FDR in the linear time series model setting. The theoretical results and appealing finite-sample performance of the suggested TSKI method coupled with the BE are illustrated with several simulation examples and an economic inflation forecasting application. This is a joint work with Chien-Ming Chi, Yingying Fan and Ching-Kang Ing.

Biography of the speaker: Dr. Lv is Kenneth King Stonier Chair in Business Administration and Professor in Data Sciences and Operations Department of the Marshall School of Business at the University of Southern California, and Professor in Department of Mathematics at USC. He received his Ph.D. in Mathematics from Princeton University in 2007. He was McAlister Associate Professor in Business Administration at USC from 2016-2019. His research interests include statistics, machine learning, data science, business applications, and artificial intelligence and blockchain. His papers have been published in journals in statistics, economics, business, computer science, information theory, neuroscience, and biology. He is the recipient of the International Congress of Chinese Mathematicians 45-Minute Invited Lecture (2023), NSF Emerging Frontiers (EF) Grant (2022), Fellow of American Statistical Association (2020), NSF Grant (2020), Kenneth King Stonier Chair in Business Administration (2019), Fellow of Institute of Mathematical Statistics (2019), Member of USC University Committee on Appointments, Promotions, and Tenure (UCAPT,

2019-present), USC Marshall Dean's Award for Research Impact (2017), Adobe Data Science Research Award (2017), McAlister Associate Professor in Business Administration (2016), Simons Foundation Grant (2016), the Royal Statistical Society Guy Medal in Bronze (2015), NSF Faculty Early Career Development (CAREER) Award (2010), USC Marshall Dean's Award for Research Excellence (2009), Journal of the Royal Statistical Society Series B Discussion Paper (2008), NSF Grant (2008), and Zumberge Individual Award from USC's James H. Zumberge Faculty Research and Innovation Fund (2008). He has served as an associate editor of Journal of the American Statistical Association (2023-present), Journal of Business & Economic Statistics (2018-present), The Annals of Statistics (2013-2018), and Statistica Sinica (2008-2016).

From:

<https://www2.math.binghamton.edu/> - **Binghamton University Department of Mathematics and Statistics**

Permanent link:

<https://www2.math.binghamton.edu/p/seminars/datasci/110122>



Last update: **2022/10/26 20:50**