Jinchi Lv, Ph.D.
McAlister Associate Professor
USC
Los Angeles, CA

Olmsted Hall 420
December 4th 2018
3:45-4:45pm
Reception in Olmsted 1331 at 3:15 P.M.

“RANK: LARGE-SCALE INference with GRAPHICAL NONLINEar KNOckoffs”
Abstract

Power and reproducibility are key to enabling refined scientific discoveries in contemporary big data applications with general high-dimensional nonlinear models. In this paper, we provide theoretical foundations on the power and robustness for the model-X knockoffs procedure introduced recently in Candès, Fan, Janson and Lv (2018) in high-dimensional setting when the covariate distribution is characterized by Gaussian graphical model. We establish that under mild regularity conditions, the power of the oracle knockoffs procedure with known covariate distribution in high-dimensional linear models is asymptotically one as sample size goes to infinity. When moving away from the ideal case, we suggest the modified model-X knockoffs method called graphical nonlinear knockoffs (RANK) to accommodate the unknown covariate distribution. We provide theoretical justifications on the robustness of our modified procedure by showing that the false discovery rate (FDR) is asymptotically controlled at the target level and the power is asymptotically one with the estimated covariate distribution. To the best of our knowledge, this is the first formal theoretical result on the power for the knockoffs procedure. Simulation results demonstrate that compared to existing approaches, our method performs competitively in both FDR control and power. A real data set is analyzed to further assess the performance of the suggested knockoffs procedure. This is a joint work with Emre Demirkaya, Yingying Fan and Gaorong Li.

Biography

Jinchi Lv is McAlister Associate Professor in Business Administration in Data Sciences and Operations Department of the Marshall School of Business at the University of Southern California, Associate Professor in Department of Mathematics at USC, and an Associate Fellow of USC Dornsife Institute for New Economic Thinking (INET). He received his Ph.D. in Mathematics from Princeton University in 2007 under the supervision of Jianqing Fan. His research interests include statistical machine learning, deep learning, causal inference, high-dimensional statistics and large-scale inference, big data problems, personalized medicine and choices, scalable Bayesian inference, networks, business and neuroscience applications, and financial econometrics.

His papers have been published in journals in statistics, economics, information theory, biology, and computer science, and one of them was published as a Discussion Paper in Journal of the Royal Statistical Society Series B (2008). He has served as an associate editor of the Annals of Statistics (2013-2018) and Statistica Sinica (2008-2016). He is the recipient of USC Marshall Dean’s Award for Research Impact (2017), Adobe Data Science Research Award (2017), 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), and Zumberge Individual Award from USC’s James H. Zumberge Faculty Research and Innovation Fund (2008).