L'inférence causale en présence de dépendance et de structure de réseau
« Atelier sur la découverte de la structure causale à haute dimension »
25 au 27 juin 2018

Causal Inference in the Presence of Dependence and Network Structure
"Workshop on Discovery of causal structure in high dimensions"
June 25-27, 2018

Matrix completion methods for causal panel data models

Guido W. Imbens *

imbens@stanford.edu

In this paper we develop new methods for estimating causal effects in settings with panel data, where a subset of units are exposed to a treatment during a subset of periods, and the goal is estimating counterfactual (untreated) outcomes for the treated unit/period combinations. We develop a class of estimators that uses the observed elements of the matrix of control outcomes corresponding to untreated unit/periods to predict the "missing" elements of the matrix, corresponding to treated units/periods. The approach estimates a matrix that well-approximates the original (incomplete) matrix, but has lower complexity according to a matrix norm, where we consider the family of Schatten norms based on the singular values of the matrix. The proposed methods have attractive computational properties. From a technical perspective, we generalize results from the matrix completion literature by allowing the patterns of missing data to have a time series dependency structure. We also present new insights concerning the connections between the interactive fixed effects models and the literatures on program evaluation under unconfoundedness as well as on synthetic control methods. If there are few time periods and many units, our method approximates a regression approach where counterfactual outcomes are estimated through a regression of current outcomes on lagged outcomes for the same unit. In contrast, if there are few units and many periods, our proposed method approximates a synthetic control estimator where counterfactual outcomes are estimated through a regression of the lagged outcomes for the treated unit on lagged outcomes for the control units. The advantage of our proposed method is that it moves seamlessly between these two different approaches, utilizing both cross-sectional and within-unit patterns in the data.

Coauthors, Susan Athey, Mohsen Bayati, Nick Doudchenko, and Khashayar Khosravi.

*Department of Economics, Stanford Graduate School of Business, 655 Knight Way, Stanford, CA 94305-5015, USA