

Causal Mediation Analysis with a Sparse Single-Index

Model

Yi Zhao

Department of Biostatistics and Health Data Science
Indiana University School of Medicine, U.S.A.

Abstract

We study causal mediation analysis in the presence of multiple mediators and pre-treatment covariates using a semiparametric single-index outcome model. The proposed model assumes that multiple mediators jointly affect the outcome through one underlying mediator index, represented as a weighted combination of individual mediators with an unknown smooth relationship to the outcome. This formulation accommodates high-dimensional mediators while preserving interpretability, because the index weights quantify the relative contributions of individual mediators to the indirect pathway. To identify the most relevant mediators, we incorporate a sparsity-inducing penalty that enables automatic mediator selection. Within the potential-outcomes framework, we define natural direct and indirect effects and present identification assumptions that extend standard mediation assumptions to the covariate-adjusted setting. We further develop two multiply robust estimators: one that retains the flexibility of unrestricted nonparametric efficiency theory, and another that exploits the single-index structure to improve efficiency and stability when the number of mediators is large. We establish consistency, asymptotic normality, and valid inference for the proposed index and causal effect estimators, as well as selection consistency for identifying active mediators under suitable tuning.