HKUST ECON Seminar

August 31 (Mon), 3:00 - 4:30 pm

Thin Sets Are Not Equally Thin: Minimax Learning of Submanifold Integrals

Xiaohong Chen, Wayne Gao

Abstract:

Many economic parameters are identified by ``thin sets'' (submanifolds with Lebesgue measure zero) and hence difficult to recover from data in an ambient space. This paper provides a unified theory for estimation and inference of such ``thin-set'' identified functionals. We show that thin sets are \emph{not} equally thin: their intrinsic dimensionality m matters in a precise manner. For a nonparametric regression h_0 with H\"{o}lder smoothness s and d-dimensional covariates in the ambient space, we show that $n^{-\frac{s}{2s+d-m}}$ is the minimax optimal rate of estimating linear and nonlinear (e.g., quadratic, upper contour) integrals of h_0 on an m-dimensional submanifold ($0\leq m < d$), which is the fastest possible attainable rate among all estimators. The minimax lower bound rate result is generalized to estimating submanifold integrals when h_0 is a nonparametric density and a nonparametric instrumental variable function. The asymptotic normality of t statistics is established via sieve Riesz representation, and the corresponding inference is computed using Sobol points.

HKUST ECON Seminar

August 31 (Mon), 3:00 - 4:30 pm

Inference on Irregular Evaluation Functionals under Optimal Treatment Assignments

Xiaohong Chen, Zhenxiao Chen, Wayne Gao

Abstract:

We study estimation and inference for \emph{evaluation functionals}: weighted probabilities of population regions selected by an oracle conditional average treatment effect (CATE) rule, including treatment shares, selected-group average characteristics, and side effects per target or selected unit. When the evaluation weight is nonzero on the zero-CATE boundary, the pathwise derivative is an unbounded surface integral, precluding any regular $\sqrt{n}$ consistent and asymptotically normal estimation. We first establish that $n^{-s/(2s+1)}$ is the minimax rate lower bound among all possible consistent estimators of such functionals under the s H\"{o}lder smoothness assumption on the unknown CATE function. We then show that a simple plug-in linear sieve estimator for such a functional can attain the minimax optimal rate. For confidence set construction based on a sieve student t statistic, a linear-sieve plug-in estimator is pointwise Gaussian at rate $n^{-s/(2s+1)}$ up to logarithmic undersmoothing. The results cover unknown target distributions and weights to be nonparametrically estimated as well. Welfare functional is a regular exception. Simulations and applications to the International Stroke Trial and JTPA training data sets illustrate the effectiveness of our methods.