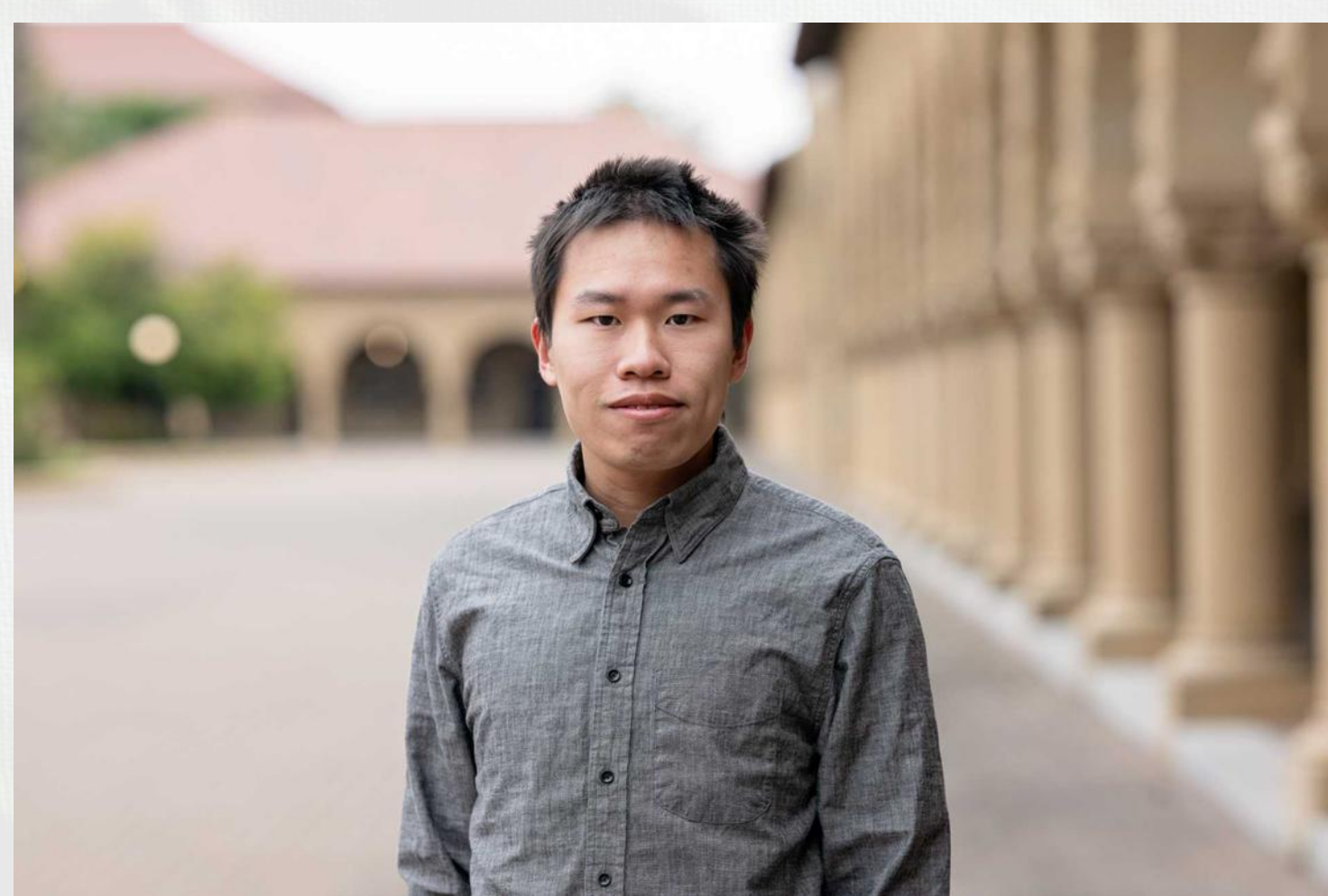




復旦大學

数据科学与深度学习
博士生与青年学者论坛

Automating Causal Inference with Guarantees: From Adaptive Estimation to Formal Verification



Jiyuan Tan

Stanford University

时间: 2026年10月27日10:30-11:30

腾讯会议: 568-614-847

密码: 261027

Abstract: This talk presents two recent works on automation in causal inference: one automates the choice of regularization in estimation, and the other automates theoretical research. In the first part, we study adaptive estimation and inference in linear inverse problems defined by conditional moment restrictions, such as nonparametric instrumental variables. Regularization stabilizes such problems, but the regularization parameter depends on a source parameter that is almost never known. We develop a discrepancy principle that chooses this parameter from empirical losses and prescribed bounds on statistical error, without knowing the source parameter. The framework covers Regularized DeepIV (RDIV) and the Tikhonov Regularized Adversarial Estimator (TRAЕ). Both procedures attain strong- and weak-error bounds that match the rates achievable when the source parameter is known. We also extend the framework to iterative procedures, widening the range of source parameters covered. Finally, we use these adaptive estimators to construct an asymptotically normal doubly robust estimator that adapts to the well-posedness of the primal and dual problems. In the second part, we turn to automating theoretical research, which requires generating candidate results and evaluating them reliably. Models keep getting better at the first, while the second remains hard. A common approach asks one large language model to review what another produced, yet such reviewers are empirically unreliable: they may accept fabricated papers and catch the fabrication at close to chance rates. We present CausalSmith, a framework for automated theoretical research in causal inference built on the Lean proof assistant, where a program checks each proof instead of a referee reading it. CausalSmith rests on Causalean, a foundational Lean library for causal inference holding 8,179 machine-checked definitions and theorems. Around it we build a self-improving agentic pipeline that selects research topics, proposes results, formalizes statements, constructs proofs, and presents the resulting artifacts for human inspection. The pipeline also pairs Lean verification with a statement audit that compares each formal theorem against the informal claim behind it. We evaluate the system on artifacts produced by completed autonomous research runs.

个人简介: Jiyuan Tan is a fourth-year Ph.D. student in Management Science and Engineering at Stanford University, where he is co-advised by Professors Vasilis Syrgkanis and Jose Blanchet. His research focuses on the automation of causal inference with rigorous mathematical guarantees, ranging from statistical estimation to AI agents. His work spans partial identification, adaptive estimation in causal models, and causal reasoning in AI systems. Most recently, he developed CausalSmith, an agentic framework for automated research in causal inference that is formally verified in the Lean proof assistant. His work has appeared at NeurIPS, ICML, and UAI. He received his B.S. in Mathematics from Fudan University.

复旦大学数学国家高层次人才培养中心 —— 数据科学与深度学习博士生与青年学者论坛
复旦大学数学科学学院 上海数学中心