Luofeng Liao

liaoluofeng96@gmail.com
5035 S East End Ave, Chicago, IL
(+86) 13420829344

EDUCATION

Ph.D. candidate, Columbia University, Industrial Engineering and Operations Research — 2021 – present

M.S., Statistics, University of Chicago, US — 2019 – 2020
- GPA: 3.90/4.0.
- Concentration: theoretical statistics, statistical learning theory and optimization

B.S., Computer Science (Data Science Track), Fudan University, Shanghai, China — 2015 – 2019
- GPA: 3.69/4.0. rank 2/41.
  - 2015 - 17: School of Computer Science (rank 5/108)
  - 2017 - 19: School of Data Science
- Concentration: algorithm designs, database system, software development

Undergraduate Exchange Program, University of Melbourne, Australia — February – July 2018

RESEARCH INTERESTS

Machine learning for causal inference and econometrics; theoretical reinforcement learning; federated learning.

PUBLICATIONS

Provably Efficient Neural Estimation of Structural Equation Model: An Adversarial Approach — (First author, NeurIPS 2020)
Collaborators: You-Lin Chen, Zhaoran Wang, Mladden Kolar, Bo Dai — March – June 2020

- Motivation: non-parametric estimation techniques in causal inference and econometrics lack scalability; neural nets (NNs) enjoy great representation power but are not well-versed in social sciences.
- Propose a primal-dual formulation for conditional moment problems to facilitate first-order methods. Present theories for (i) global convergence to optima, and (ii) computational and sample complexity, through the lens neural-tangent kernel theory.
- The first scalable NN-based estimation procedure with guarantees for econometrics and causal inference.

Provably Efficient Instrument-Aided Causal Reinforcement Learning with Linear Function Approximation — (Co-first, under review)
Collaborators: Zuyue Fu, Zhaoran Wang, Mladen Kolar — June – November 2020

- Motivation: application of reinforcement learning in social sciences (e.g., education, health care) with access to only observational data requires modeling confounders.
- Propose a model-based IV-adjusted value iteration algorithm to recover the optimal policy using observational data coming from a confounded Markov decision process.
- The first scalable IV-aided policy learning algorithm with guarantees on (i) sample complexity and (ii) convergence to optimal policy, in the presence of confounders.

Local AdaGrad-Type Algorithm for Stochastic Convex-Concave Minimax Problems — (First author, under review)
Collaborators: Li Shen, Jia Duan, Mladen Kolar, Dacheng Tao — March – June 2021

- Motivation: due to privacy/computational concerns, communication-efficient distributed solutions to adversarial training are desirable; GAN are sensitive to the learning rate scheme employed.
- Propose a local-update and periodic-averaging algorithm for distributed minimax problems with adaptive learning rates. Present convergence theories for smooth/nonsmooth minimax problems.
- One of the first provably efficient federated learning algorithm for minimax problems without the need of learning rate tuning.

APPLIED RESEARCH

Leader, Spatio-temporal Modeling of Environmental Data with Additive Models June – September 2018
Advisor: Prof Ruibin Xi, Peking University (summer research project)

- Motivation: Study the effects of PM2.5 on cerebrovascular diseases in Beijing.
- Conduct data cleaning and analysis on an up-to-date cerebrovascular disease database (~500,000 records) using R; transform address texts data into GIS and perform spatial-temporal visualization.
- Propose and implement a new distributed lag model to capture (i) temporal lag effects of PM2.5 and (ii) interaction with atmospheric variables; develop an interactive website for the model; verify robustness with different basis function.

STATISTICAL SOFTWARE

Sole Developer, Modern Multivariate Analysis by Penalization (*sponsored project in Google Summer of Code 2019, Statistical Computing in R*)
Advisors: Prof. Genevera Allen, Dr. Michael Weylandt (Rice University) May – August 2019

- A performant C++ library for multivariate analysis models (PCA, CCA, LDA) that allows data-driven modeling, such as incorporating sparsity/smoothness.
- Implement a penalized singular value decomposition algorithm with different options of penalty; build R interfaces and interactive visualization platforms for data dimension reduction; provide thorough benchmark results.

Sole Developer, GPU-accelerated Bayesian Regression (*ranked 2/30 in student poster session of 2018 Fudan Science and Innovation Forum)*
Advisor: Prof. Quoqi Qian (University of Melbourne) March – July 2018

- A collection of Bayesian regression models which allow for variable selection and robustness in estimates; suitable for ultra-large datasets due to GPU acceleration.
- Develop a parallel Gibbs sampler using CUDA packages; experiment with advanced GPU features such as shared memory, streaming, and asynchronous and overlapping data transfer, which made the result 100x faster than its CPU version.

PROFESSIONAL SERVICES

Reviewer, AISTATS2020, review papers on minimax games and local stochastic gradient algorithms November 2020
Consultant, Consulting Projects in Dept. of Statistics, University of Chicago 2020
- Advise on testing significance of gene expression in mice under medical interventions.

Teaching Assistant, Fudan University, Numerical Algorithms, with Prof Ke Wei 2019
Presenter, HELIOS reading group in Dept. of Statistics, U. of Chicago, present the NeurIPS2020 paper 2019

PRIZES AND AWARDS

Runner's up, student poster session of 2018 Fudan Science and Innovation Forum 2018
Fudan Xiyuan Undergraduate Research Scholarship (800$) For academic excellence and research potential 2018
SCSK Corporation Scholarship (700$) For academic excellence 2017
China National Student Scholarship, Second price (200$) For academic excellence 2015, 2016

SKILLS

Languages: Mandarin and Cantonese (Native), English (Proficient, TOFEL 114, Speaking 27)
Computing: R, Python, MATLAB, SQL, C++, Spark, Linux shell, Latex