Jiajun Song

website: jiajunsong629.github.io
email: jiajun.song@alumni.duke.edu
code: github.com/JiajunSong629

Education

Duke University — Durham, NC
M.S. in Statistical Science — 2019–2021

Peking University — Beijing
B.S. in Applied Mathematics — 2014–2018

Research Interests

- High-dimensional statistics, semiparametric statistics, finite-sample property
- Mechanistic interpretability and low-rank structures of Transformer
- Causal inference on complex structured data

Publications

[1] J. Song and Y. Zhong, "Uncovering hidden geometry in transformers via disentangling position and context", *arXiv preprint arXiv:2310.04861*, 2023.

Academic Projects

- Analyzing the Effect of Early Adversities in the Life Expectancy of Wild Baboons — 2021
  Survival Analysis, Causal Inference — Advisor: Fan Li
- Assessing the Racial Disparities in NYC Policing — 2021
  Multilevel Poisson Regression, Spatial Statistics — Advisor: David B. Dunson
- Investigating Mortality Rates from Cardiovascular Pediatric Surgery in STS Public Reporting — 2020
  Bayesian Joint Modeling, Bayesian Calibration
- Python Package: Stochastic Gradient HMC — 2020
  Hamiltonian Monte Carlo, Package Development — Advisor: Cliburn Chan

Research & Work Experience

National Key Laboratory of Artificial General Intelligence — Beijing
Research Engineer — Aug 2022 – Present

- AI For Science: IMO-level Geometry Problem Generation and Solving
- *Neural-Symbolic, Monte Carlo Tree Search*

Warner Music Group — New York, NY
Data Scientist — Jun 2021 – Aug 2022

- Stream Forecast
- *Markov-Switching Autoregressive Model, Bayesian online changepoint detection*

– Playlist Effect Causal Analysis

– *Bayesian structural time series, Synthetic Difference in Difference*

David's Lab, Duke University Durham, NC
Research Statistician Jan 2020 – Jun 2021

– Data-Intensive Medical Device Design

– *Breakpoint detection, Neural Net, Fourier Analysis*

Teaching

- **Teaching Assistant** at Duke University Spring 2021
  Introduction to Mathematical Statistics (STA661), Scott C. Schmidler
- **Teaching Assistant** at Duke University Fall 2020
  Theory of Statistical Inference (STA532), Janson Xu
- **Teaching Assistant** at Duke University Fall 2019
  Introduction to Mathematical Statistics (STA661), Meimei Liu

Skills

- **Programming**: C, Python, R, Stan, SQL, PyTorch
- **Technologies**: Git, Docker, CI/CD, Parallel Computing, Cloud Computing
- **Language**: Mandarin (Native), English (Professional)
- **Statistical Analysis**: Causal Inference, Time Series, Hierarchical/Multilevel Model

Relevant Courses

- **Undergraduate level**: Linear Algebra (97/100), Calculus I & II (95/100), Abstract Algebra (95/100), Statistical Learning (94/100), Probability Theory (97/100), Real Analysis, Measure Theory
- **Graduate level**: Causal Inference (A+), Introduction to Deep Learning (A+), Statistical Programming (A), Bayesian Methods and Modern Statistics (A), Programming, Data Structures, and Algorithms in C++ (A)

Awards

- Selected in Applied Mathematics Program for Elite Students, Peking University 2015–2018
- 1st prize in National High School Mathematics Olympic 2014
- 2nd prize in National High School Mathematics Olympic 2013