

TIANYI MA

St John's College, Cambridge, CB2 1TP

Email: tm681@cam.ac.uk \diamond mttty2000@gmail.com

Phone: +44 7709906846

EDUCATION

University of Cambridge (St John's College)

PhD in Statistics

Oct. 2023 – Present

Cambridge, UK

- Supervisors: Prof Richard Samworth and Dr Tengyao Wang.

University of Cambridge (St John's College)

MMath and BA in Mathematics

Oct. 2019 – Jul. 2023

Cambridge, UK

- Result: Distinction. Overall mark: 87/100.
- Courses: Topics in Statistical Theory, Modern Statistical Methods, Statistical Learning in Practice, Concentration Inequalities, Information Theory, Topics in Convex Optimisation, Advanced Probability, Probability and Measure, Linear Algebra etc.

RESEARCH EXPERIENCE

Missing Data (Part III Essay)

Supervised by Prof Richard Samworth, University of Cambridge

Mar. 2023 – May. 2023

- In this essay, we considered the problem of estimating covariance and precision matrices, and principal component analysis under the setting of high-dimensional data with missing values.
- We studied two ways of estimating the covariance matrix under the assumption of missing completely at random. We provided finite-sample bounds for the proposed methods under heterogeneous missingness, and showed its minimax optimality under homogeneous missingness.
- We studied the finite-sample behaviour of principal component analysis based on the estimated covariance matrix, and the performance of spectral clustering under heterogeneous missingness.
- We provided finite-sample bounds for the graphical Lasso estimator of the precision matrix based on the estimated covariance matrix under heterogeneous missingness, with extensions to other popular estimators of the precision matrix.
- Finally, we provided simulation results for the problems considered in the essay. Our proposed methods empirically perform well under missing data.

High-Dimensional Nonparametric Two-Sample Testing

Supervised by Dr Tengyao Wang, Associate Professor, LSE

Jul. 2022 – Sep. 2022

- In this project, we aimed to develop a new methodology to detect a sparse change in marginal distributions for high-dimensional two-sample testing (and change-point) problems.
- We reviewed a large amount of literature on two-sample testing and change-point problems for both a change in distribution and a change in the mean vector. I implemented R programmes for these methods, and did extensive simulations to compare their performance under different models.

- We proposed a testing procedure by first standardising the data in a particular way and calculating the coordinate-wise energy statistic, then choosing the sum of k (which may vary with the dimension) largest energy statistics as a test statistic and combine with a permutation test. Simulation studies showed that this method has an improved power compared with other nonparametric tests under a sparse change in marginal distributions.

INTERNSHIP

Data Analyst

Neusoft Hifly

Jul. 2020 – Aug. 2020

Shenyang, China

- Collected and processed medical data from hospitals using MySQL database.
- Feature engineering and data visualisation using Python. Implemented different machine learning algorithms to analyse the relation between vital signs and different clinical diseases.

SKILLS

- Proficient in R and MATLAB.
- Proficient in writing academic papers using LaTeX.
- Familiar with Python.

LANGUAGES

- Chinese: native speaker.
- English: 7.5 in IELTS.