For favour of Posting

THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY

Department of Mathematics

SEMINAR ON STATISTICS

HeteroJIVE: Joint Subspace Estimation for Heterogeneous Multi-View Data

By

Dr. Jingyang LI

University of Michigan

Abstract

Many modern datasets consist of multiple related matrices measured on a common set of units, where the goal is to recover the shared low-dimensional subspace. While the Angle-based Joint and Individual Variation Explained (AJIVE) framework provides a solution, it relies on equal-weight aggregation, which can be strictly suboptimal when views exhibit significant statistical heterogeneity (arising from varying SNR and dimensions) and structural heterogeneity (arising from individual components). In this paper, we propose HeteroJIVE, a weighted two-stage spectral algorithm tailored to such heterogeneity. Theoretically, we first revisit the "non-diminishing" error barrier with respect to the number of views K identified in recent literature for the equal-weight case. We demonstrate that this barrier is not universal: under generic geometric conditions, the bias term vanishes and our estimator achieves the $O(K^{-1/2})$ rate without the need for iterative refinement. Extending this to the general-weight case, we establish error bounds that explicitly disentangle the two layers of heterogeneity. Based on this, we derive an oracle-optimal weighting scheme implemented via a data-driven procedure. Extensive simulations corroborate our theoretical findings, and an application to TCGA-BRCA multi-omics data validates the superiority of HeteroJIVE in practice.

Bio : Jingyang Li is a Postdoc at the University of Michigan and will soon join Fudan University as an assistant professor. His research interests lie in high-dimensional statistics, with a focus on matrix and tensor learning, nonconvex optimization, and robust statistics.

Jingyang works on developing scalable algorithms and understanding their theoretical properties for large-scale data problems. He has contributed to methods involving Riemannian optimization for tensor estimation, as well as robust frameworks for high-dimensional regression and low-rank recovery. Currently, he is also exploring questions in online learning and federated learning, seeking to design adaptive algorithms that balance computational trade-offs with privacy constraints.

.

Date : 30 January 2026 (Friday)

Time : 10:00a.m.-11:00a.m.

Venue : Room 2463 (near Lift 25/26)

All are welcome!