



THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY

Department of Mathematics

SEMINAR ON STATISTICS AND DATA SCIENCE

**Efficient Embedding and
Generative Modeling for Hypergraphs**

By

Prof. Ji ZHU

University of Michigan

Abstract

Data that represent relations and interactions are ubiquitous in science, engineering, business, and medicine. Traditional analytical methods for such data primarily focus on pairwise relations; however, real-world interactions often involve more than two entities and are inherently multi-way. In current practice, these multi-way interactions are typically projected into pairwise relations before analysis, which causes substantial information loss. Directly studying hypergraphs, which naturally encode general multi-way interactions, allows for more effective extraction of information from such relational data. In this talk, I will discuss our development of generative models for hypergraphs. The first part introduces a general latent embedding framework that overcomes key limitations of existing hypergraph modeling methods. We establish identifiability of the latent embedding space and develop a likelihood-based estimator for the latent embeddings. We further derive consistency guarantees and asymptotic distributions for the parameter estimates, enabling efficient inference from an observed hypergraph. Building on these results, the second part of the talk introduces Denoising Diffused Embeddings (DDE), a generative architecture for hypergraphs that produces new hyperlinks not seen in the observed data. DDE connects discrete hyperlinks to a continuous latent space through a conditional hyperlink likelihood model, and then reconstructs that space using a denoising diffusion process. Compared with existing generative models, DDE is computationally efficient to train and sample from, and it also offers interpretability from the likelihood perspective. Our theoretical and empirical studies demonstrate its advantages as a flexible framework for generative modeling of hypergraphs. Together, these results provide new statistical tools for modeling and generating multi-way relational data and highlight the potential of combining probabilistic modeling with modern generative methods.

Short bio: Ji Zhu is Susan A. Murphy Professor of Statistics at the University of Michigan, Ann Arbor. He received his B.Sc. in Physics from Peking University, China in 1996, and his M.Sc. and Ph.D. in Statistics from Stanford University in 2000 and 2003, respectively. His primary research interests include statistical machine learning, statistical network analysis, and their applications to health and natural sciences. He received an NSF CAREER Award in 2008 and was elected as a Fellow of the American Statistical Association (ASA) in 2013 and a Fellow of the Institute of Mathematical Statistics (IMS) in 2015. From 2014 to 2020, he was recognized as a Highly Cited Researcher by Web of Science, which annually identifies leading researchers in the sciences and social sciences worldwide. In 2022, he received the International Chinese Statistical Association Pao-Lu Hsu Award. He served as the Editor-in-Chief of the Annals of Applied Statistics from 2022 to 2024 and was selected to deliver an IMS Medallion Lecture in 2027.

Date : 12 June 2026 (Friday)
Time : 11:00a.m.-12:00noon
Venue : Room 4504 (near Lift 25/26)

All are welcome