THE UNIVERSITY OF CHICAGO | THE COMMITTEE ON COMPUTATIONAL AND APPLIED MATHEMATICS

Dissertation Defense:

Zhao Lyu

Computational and Applied Mathematics
University of Chicago

"Cross-Validation for Structural Model Selection"

Monday, June 15, 2026, at 8:30 AM
Location: Jones 111, 5747 S. Ellis

ABSTRACT

Cross-validation is widely used in modern machine learning for model selection and hyperparameter tuning. Its appeal comes from its model-agnostic nature: it selects models through held-out predictive performance rather than through problem-specific assumptions. In structure learning and related structural model selection problems, however, model selection has an additional consequence. The selected model may encode a graph, a sparsity pattern, a set of active variables, a representation, an architecture, or another structural object. This presentation studies when such a prediction-based selection principle can be trusted for structural model selection, and what must change when it cannot.

We address this question in three steps. First, in Gaussian graphical model selection, we show that cross-validation can fail to be model-selection consistent even when consistent graph recovery is statistically achievable, revealing a fundamental mismatch between predictive validation loss and exact structural recovery. Second, we show that this negative result does not render validation information useless: by using cross-validation-based quantities as scalable signals rather than treating the validation-loss minimizer as the final model selector, we develop approximate cross-validation methods for consistent and efficient graph recovery in general graphical models. Finally, we extend cross-validation-based selection to nonparametric and deep learning settings, where structural choices are often implicit and exact retraining is infeasible. We develop a new proxy for cross-validation selection and show that it preserves the model-selection behavior of cross-validation under suitable conditions.

Together, these results reposition cross-validation in structural model selection. Cross-validation is not automatically a consistent structural selector, but its validation principle can be diagnosed, redesigned, and computationally extended for classical and modern learning systems.