



THE UNIVERSITY OF
CHICAGO

DEPARTMENT OF STATISTICS

KEVIN LIN

Department of Statistics and Data Science
The Wharton School, University of Pennsylvania

“Quantifying common and distinct information in multi-modal single-cell data via matrix factorization”

TUESDAY, JANUARY 10, 2023, at 2:00 PM
Jones 303, 5747 S. Ellis Avenue

ABSTRACT

Recently, multi-modal single-cell data has been growing in popularity in many areas of biomedical research and provides new opportunities to learn how different modalities coordinate within each cell. Many existing dimension reduction methods for such data estimate a low-dimensional embedding that captures all the axes of variation from either modality. While these current methods are useful, we develop the Tilted-CCA in this talk to perform a fundamentally different task. This method is a novel matrix factorization that estimates low-dimensional embeddings separating the axes of variation shared between both modalities (i.e., "common geometry," capturing the coordination between both modalities) from axes of variation unique to a particular modality (i.e., "distinct geometry"). Methodologically, Tilted-CCA achieves this by combining ideas from Canonical Correlation Analysis (CCA) and density clustering. Our method first uses the nearest-neighbor graphs from each modality to infer the common geometry between both modalities and decomposes the canonical scores from CCA to approximate this geometry. Biologically, Tilted-CCA unveils the cellular dynamics in developmental systems based on the proportion of variation between the common and distinct embeddings. More broadly, Tilted-CCA invites new theoretical questions regarding dimension reduction and can be applied to any domain beyond single-cell genomics.