

中国科学院数学与系统科学研究院

应用数学研究所

华罗庚应用数学青年论坛

报告题目: Representation learning for single-cell multi-omics and spatial-omics data

报告人: 林志翔 副教授 (香港中文大学)

时 间: 2024年7月18日(星期四) 下午 15:00--16:00

地 点: 数学院思源楼 625 会议室

摘 要: In this talk, I will share two of our recent works. First, I will talk about JSNMF, which is a versatile toolkit for the integrative analysis of transcriptomic and epigenomic data profiled from the same cell. The core of JSNMF is an unsupervised method based on jointly semi-orthogonal nonnegative matrix factorization, where it assumes different latent variables for the two molecular modalities, and integrates the information of transcriptomic and epigenomic data with consensus graph fusion, which better tackles the distinct characteristics and levels of noise across different molecular modalities in single-cell multi-omics data. We applied JSNMF to single-cell multi-omics datasets from different tissues and different technologies. The results demonstrate the superior performance of JSNMF in clustering and data visualization of the cells. JSNMF also allows joint analysis of multiple single-cell multi-omics experiments and single-cell multi-omics data with more than two modalities profiled on the same cell. JSNMF also provides rich biological insight on the markers, cell-type-specific region-gene associations and the functions of the identified cell subpopulation.

Second, I will talk about SGCAST, an efficient auto-encoder framework to identify spatial domains. SGCAST adopts a symmetric graph convolutional auto-encoder to learn aggregated latent embeddings via integrating the gene expression similarity and the proximity of the spatial spots. This framework in SGCAST enables a mini-batch training strategy, which makes SGCAST memory-efficient and scalable to high-resolution spatial transcriptomic data with a large number of spots. SGCAST improves the overall accuracy of spatial domain identification on benchmarking data. We also validated the performance of SGCAST on spatial transcriptomic datasets at various scales across multiple platforms. Our study illustrates the superior capacity of SGCAST on analyzing spatial transcriptomic data.

个人简介: Zhixiang Lin is currently Associate Professor in the Department of Statistics at the Chinese University of Hong Kong (CUHK). He received his B.S. from Tsinghua University in 2010 and Ph.D. from Yale University in 2015. During his graduate study at Yale, he was co-advised by Prof. Hongyu Zhao in the Department of Biostatistics and Prof. Matthew State in the Department of Psychiatry. Dr. Lin worked with Prof. Wing Hung Wong at Stanford University as a postdoctoral researcher from 2015 to 2018. His lab develops statistical methods and computational tools for addressing significant scientific questions, especially those related to biomedical research, large-scale genomic data, health and clinical data. The most recent application area include single cell genomics and neuroscience.