University of California, Irvine
Statistics Seminar

Statistical Models for Spatial Transcriptomics Data to Study Cell-cell Interactions

Sha Cao
Assistant Professor of Biostatistics & Health Data Sciences
Indiana University

4 p.m., Thursday, February 3, 2022

The recent development of spatial transcriptomics (ST) technologies has provided spatially resolved gene expressions at the resolution of single cells, and are promising to deliver better understanding of the complex interactions within the tissue microenvironment. To computational study the complex interactions, two key questions in the analysis of ST data include: clustering of spatial single cells, and analysis of spatially dependent gene gene associations. To date, most studies used methods that only rely on the expression levels of the genes, while the spatial information has not been used efficiently. To fully utilize the data and to improve the ability to identify biologically meaningful cell clusters and gene gene associations, we have developed 1) a low rank approximation method embedded with spatial smoothness that specifically address the count based nature of the ST data, which will enable efficient low dimensional visualization, as well as clustering of the spatial single cells; and 2) a spatial robust mixture regression model to investigate the relationship between a response variable and a set of explanatory variables over the spatial domain, assuming that the gene gene relationships may exhibit complex spatially dynamic patterns that cannot be captured by constant regression coefficients.