A nonparametric Bayesian approach to simultaneous subject and cell heterogeneity discovery for single cell RNA-seq data

Xiangyu Luo

RENMIN UNIVERSITY of CHINA
E-mail: xiangyuluo@ruc.edu.cn

Abstract: The advent of the single cell sequencing era opens new avenues for the personalized treatment. The first but important step is discovering the subject heterogeneity at the single cell resolution. We address the two-level-clustering problem of simultaneous subject subgroup discovery (subject level) and cell type detection (cell level) based on the single cell RNA sequencing (scRNA-seq) data from multiple subjects. However, current approaches either cluster cells without considering the subject heterogeneity or group subjects not using the single cell information. We develop a solid nonparametric Bayesian model SCSC (Subject and Cell clustering for Single-Cell data) to achieve subject and cell grouping at the same time without pre-specifying the subject subgroup number or the cell type number. An efficient blocked Gibbs sampler is then proposed for the posterior inference. The simulation study and the real application demonstrate the good performance of our model.