

scRMD: Imputation for single cell RNA-seq data via robust matrix decomposition

Ruibin Xi

Peking University

E-mail: ruibinxi@math.pku.edu.cn

Abstract: Single cell RNA-sequencing (scRNA-seq) technology is developing at a fast pace and provides a higher resolution for single cell gene profiling, which enables us to better understand biological process at the single cell level. Nevertheless, scRNA-seq data suffers from a significant issue of down sampling and is thus biased by excessive zeros. While some of them are essential zeros indicating no expression, others are missing values known as dropouts due to insufficient amount of mRNA transcripts. To identify dropout zeros and replace them with underlying true expression levels, we develop scRMD, a statistical method for imputation in scRNA-seq data by borrowing information across genes and cells. It is shown that scRMD is able to produce both accurate and robust imputed results that lead to better downstream statistical analyses including detection of differential expression and clustering analysis.