

Statistical test of structured continuous trees based on discordance matrix

Lin Wan

*Academy of Mathematics and Systems Science, Chinese Academy of Sciences
E-mail: lwan@amss.ac.cn*

Abstract: Cell fate determination is a continuous process in which one cell type diversifies to other cell types following a hierarchical path. Advancements in single-cell technologies provide the opportunity to reveal the continuum of cell progression which forms a structured continuous tree. Computational algorithms, which are usually based on a priori assumptions on the hidden structures, have previously been proposed as a means of recovering pseudo-trajectory along cell differentiation process. However, there still lack of statistical framework on the assessments of intrinsic structure embedded in high-dimensional gene expression profile. We propose an adaptive statistical framework, termed structured continuous tree (SCTree), to test the intrinsic structure of a high-dimensional single-cell dataset. SCTree test is developed on the basis of the tools governing metric geometry and random matrix theory. We show that the SCTree test is most powerful when the signal-to-noise ratio exceeds a moderate value. We also demonstrate that SCTree is able to robustly detect linear, single and multiple branching events with simulated datasets and real scRNA-seq datasets.