Kernel Testing for Spatially Variable Gene Detection

Lucy Attwood*[1], Franck Picard[1], and Bertrand Michel[2]

[1]Laboratoire de biologie et modélisation de la cellule – Ecole Normale Supérieure de Lyon, Université Claude Bernard Lyon 1, Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique – France

[2]Ecole Centrale De Nantes (ECN) – Ecole Centrale de Nantes – France

Abstract

Spatial transcriptomics allows for the spatial exploration of tissues based on gene expression data. This analysis requires the development of statistical testing procedures to first identify genes with significant spatial variation. Numerous strategies have already been considered and benchmarked based on simulated and experimental data. However the overlap between detected spatially variable genes is small, suggesting that most available variable gene detection methods suffer from high rates of false positives. Comparative studies emphasise the good performance of the method Spark-X, a framework based on kernel methods and the Hilbert-Schmidt Independence Criterion statistic (HSIC). However, Spark-X currently lacks an explicit generative statistical model, requires the p-value to be computed based on permutations, due to an intractable null distribution, and doesn't correct for variation in spatial locations. In this work we provide a generative model of a linear relationship between the signal and the spatial coordinates embedded in distinct reproducible kernel Hilbert spaces (RKHS). This model also allows a visual representations of the genes and spatial coordinates in this functional embedded space as well as the relationship between them. We then propose a new statistical test regarding this relationship, based on a kernelised version of the Hotelling Lawley test, which has a simple and tractable asymptotic distribution. We run our method on benchmark studies that include simulated and biological data to compare with Spark-X and other existing methods.

Keywords: Kernel Methods, Statistical Testing, Spatial Transcriptomics

---

*Speaker