Learning cell fate landscapes from spatial transcriptomics using Fused Gromov-Wasserstein

Geert-Jan HUIZING, Institut Pasteur, Université Paris Cité, CNRS UMR 3738, Machine Learning for Integrative Genomics Group - Paris, France

Gabriel PEYRÉ, CNRS and DMA de l'Ecole Normale Supérieure - Paris, France
Laura CANTINI, Institut Pasteur, Université Paris Cité, CNRS UMR 3738, Machine Learning for Integrative Genomics Group - Paris, France

Spatial transcriptomics is an emerging technology that simultaneously measures the gene expression and the spatial location of cells within a tissue. This breakthrough offers new insights into neurodegenerative diseases, cancer, and development.
Recent studies collected spatial transcriptomics data at various time points, providing multimodal snapshots of a population of cells evolving in physical space and gene expression space. There is an urgent need for computational methods to reveal cell fate trajectories and drivers of differentiation from such spatiotemporal atlases.
Wasserstein gradient flows are a promising framework for analyzing cellular population dynamics across time. They provide a causal model of the evolution of cells according to a potential function. Recent works have learned neural potentials from observed gene expression snapshots. However, these methods are ill-equipped to handle the isometries characteristic of physical space.
Here, we propose a novel method to leverage spatial coordinates by employing a Fused Gromov-Wasserstein objective to learn a spatially informed potential. We benchmark our approach using large spatiotemporal atlases encompassing developmental and regenerative processes across three organisms : mice, zebrafish, and axolotls. We demonstrate superior spatial coherence compared to prevailing approaches, which use a Wasserstein objective.

Contact : geert-jan.huizing@pasteur.fr