

Distributional Reduction: Unifying Dimensionality Reduction and Clustering with Gromov-Wasserstein Projection

Hugues VAN ASSEL, ENS - Lyon **Cédric VINCENT-CUAZ**, EPFL - Lausanne

Nicolas COURTY, UBS - Vannes **Rémi FLAMARY**, X - Palaiseau

Pascal FROSSARD, EPFL - Lausanne **Titouan VAYER**, ENS - Lyon

Unsupervised learning aims to capture the underlying structure of potentially large and highdimensional datasets. Traditionally, this involves using dimensionality reduction methods to project data onto lower-dimensional spaces or organizing points into meaningful clusters. In practice, these methods are used sequentially, without guaranteeing that the clustering aligns well with the conducted dimensionality reduction. In this work, we offer a fresh perspective : that of distributions. Leveraging tools from optimal transport, particularly the Gromov-Wasserstein distance, we unify clustering and dimensionality reduction into a single framework called distributional reduction. This allows us to jointly address clustering and dimensionality reduction with a single optimization problem. Through comprehensive experiments, we highlight the versatility of our method on image and genomics datasets. See more in [1].

- [1] H. Van Assel, C. Vincent-Cuaz, N. Courty, R. Flamary, P. Frossard, T. Vayer. *Distributional reduction : Unifying dimensionality reduction and clustering with gromov-wasserstein projection*. arXiv preprint arXiv :2402.02239, 2024.

Contact : hugues.van_assel@ens-lyon.fr