A Penalized Optimization Approach for Clustering Enhancement through Persistent Homology and Random Dynamic Graph Construction

Serena Grazia De Benedictis, Ang Andersen, Nicoletta Del Buono, Flavia Esposito, Laura Selicato

Abstract

We address a clustering problem in an n-dimensional space by formulating a penalized optimization method that balances two sources of information: a k-means clustering approach on the dataset and a graph-cut method applied to a randomly activated subset of points from a constructed graph. The graph is constructed by randomly selecting a subset of points from the dataset and connecting them based on the Euclidean distance, in order to emphasize the similarity and nearness between points. This dynamic activation process ensures variability in the graph topology and allows the clustering method to adapt to different local structures within the data.

Computationally, the problem is solved by improving the gradient descent algorithm with a computational geometry technique known as Persistent Homology [1-2]. This technique, applied to the evolving graph structure, allows for the identification of the most persistent connected components in the graph in a topologically robust manner by varying a threshold on Euclidean distances between points. The interaction between the clustering process and the dynamically generated graphs provides a richer representation of the intrinsic topology of the data.

In this work, we propose two different extended algorithms that exploit different geometric constructions on the graph: the Vietoris-Rips complex [3] and the random Vietoris-Rips complex [4]. Preliminary results on the Fashion-MNIST dataset show that the proposed method achieves good performance in terms of both computation time and convergence. In particular, it outperforms the gradient descent enhancement that relies solely on the evolution of the graph with random point activation.

References

[1] Dey, Tamal Krishna and Wang, Yusu *Computational Topology for Data Analysis*, Cambridge University Press, 2022. doi:10.1017/9781009099950.

[2] Schenck, Hal *Algebraic Foundations for Applied Topology and Data Analysis* , Springer International Publishing, 2022. doi:10.1007/978-3-031-06664-1.

[3] Hausmann, Jean-Claude *On the Vietoris-Rips complexes and a cohomology theory for metric spaces.* In Prospects in topology (Princeton, NJ, 1994), volume 138 of Ann. of Math. Stud., pages 175–188. Princeton Univ. Press, Princeton, NJ, 1995. doi: 10.1515/9781400882588-013

[4] Kahle, Matthew *Random Geometric Complexes.* Discrete Computational Geometry, 45(3), 553–573, 2011. doi: 10.1007/s00454-010-9319-3