Applications of dimension theory to embeddability problems in topological data analysis: the case study of the Gromov-Hausdorff distance

Nicolò Zava

Institute of Science and Technology Austria (ISTA)
nicolo.zava@ist.ac.at

Topological data analysis (TDA) is a recent and fast-growing subject aiming to apply topological techniques to a wide range of applications. Areas where topological data analysis has found applications include neuroscience, image recognition, biology and evolutionary networks. Treating datasets or topologically inspired invariants assigned to them as points of a metric space is a central idea. However, comparing objects in this new metric space is often computationally challenging. Therefore, a technique usually deployed is mapping those objects in a Hilbert space using *vectorisation methods* or *kernels* to allow their implementation in machine learning pipelines. The quest to construct such maps that distort the distances in a controlled way is a crucial research area.

In this talk, we present how dimension theory can fruitfully assist this search. Regarding these maps as examples of certain metric embedding classes is the key-viewpoint shift. As a leading example, we discuss embeddability results for families of metric spaces endowed with the Gromov-Hausdorff distance. In addition to the intrinsic interest of these results given by the importance of this metric in Riemannian geometry and geometric group theory, they also impact TDA since the Gromov-Hausdorff distance was recently proposed as a theoretical framework for shape and dataset comparisons. More precisely, we show the following results:

- the space $\mathcal{GH}_n$ of metric spaces with at most n elements can be coarsely embedded into a Hilbert space;
- their union $\bigcup_{n\in\mathbb{N}}\mathcal{GH}_n$ cannot be coarsely embedded into any Hilbert space;
- the space of metric spaces whose diameter is bounded by r cannot be bi-Lipschitz embedded into any $\mathbb{R}^m$.

If time permits, we conclude by presenting results concerning the embeddability of persistence diagrams, one of the cornerstone notions of TDA, and periodic point sets, a generalisation of lattices used to represent crystals in material science and pharmaceutics.

References

[1] N. Zava, *Coarse and bi-Lipschitz embeddability of subspaces of the Gromov-Hausdorff space into Hilbert spaces*, arXiv:2303.04730.