Symmetry and Topological Data Analysis (TDA)

Algebraic Topology based tools to encode geometrical information about data have seen an unprecedented development and growth during the last decade. Combining statistics with TDA outcomes, subsampling techniques, and multiparameter persistence are one of the directions that have seen a substantial progress. These directions are well represented at KTH. Since TDA invariants are based on geometry, there is hope and growing evidence for their effectiveness also to encode symmetry. For example convolutional neural networks might be seen as capturing translational symmetry. Recent work of G.Carlsson explains how to generalise convolutional networks to capture symmetry described by more complicated spaces of transformations. How to encode symmetry by using TDA based representations of the data (or invariants) is an important direction that is however not represented at KTH.

Presence of noise is a prohibitive factor to use standard ways of encoding symmetry by actions of groups. One way to overcome this difficulty is to enrich the input by injecting symmetry and providing additional measurements obtained from the original measurements by their various modifications. For example assuming that a given data should satisfy rotational symmetry, we can extend the measurements representing the data set by pre composing them with various operations encoding rotational symmetries. This new data set leads to new nontrivial persistent homologies. These new persistent homology invariants composed with various kernels (such as the stable rank developed at KTH) provide novel ways of vectorising data. Features constructed in this way not only contain the standard information but also encode the injected assumption about the symmetry. It should not be surprising that in certain situations machine learning algorithms applied to these new features perform much better than using standard vectorisation methods.

The aim of this project is two fold. First, develop mathematical foundation of how one might inject and use symmetry into data possibly building on the suggestion above and study how this process effects persistent homology. Second, develop an implementation of the new TDA symmetry enriched invariants.

We believe that Brummer & Partners MathDataLab can provide an outstanding environment to bring this important research direction to KTH and the TDA group.

Wojciech Chachólski