Topological Data Analysis

PROFESSOR MIKAEL VEJDEMO-JOHANSSON (MVJ)

Topological Data Analysis is a new research field that draws on algebraic topology in pure mathematics to construct applied algorithms, statistical tools and data science approaches.

At CUNY, MVJ is currently the only faculty member primarily interested and active in the TDA community.

Get in touch: mvejdemojohansson@gc.cuny.edu

WEEKLY SEMINAR

The Data Science and Applied Topology Seminar meets Fridays at 11.45am in room 3209. More details: http://cunygc.appliedtopology.nyc

PRIMARY RESEARCH INTERESTS

Persistent Homology
From a point cloud, we can construct a simplicial complex (approx. mesh) that captures the underlying geometry. Topological invariants of this complex help us understand the data.

Persistent Cohomology
Starting the same way as above, Persistent Cohomology can produce new coordinate functions taking values on the circle. These help us understand periodicity in time series as well as directional data.

Mapper
The Mapper algorithm produces a simplicial complex model from arbitrary data, with facilities to mark aspects of the data as important for distinguishing groups.

Software
We maintain the JavaPlex library. It is the only major library available for TDA for Java and Matlab. We contribute to implementations of Mapper.

Multiple Hypotheses
Applications to genetics call for very many simultaneous instances of statistical testing with topological tools. We are just now developing multiple hypothesis correction methods in topology

BUT ALSO

Color Naming Semantics
Systems for naming colors in different languages seem to follow specific patterns. We analyze the World Color Survey at a finer grain using machine learning and data science tools.

Mathematical Art
Using 3d printers and laser cutters we produce and exhibit mathematical art depicting concepts and objects in topology and geometry.

[your area of interest]
We are keen on new areas to work in – Bayesian Statistics; Spatial Statistics; Applied Category Theory… If you want to work in the interface of pure and applied mathematics – see if we can find a project idea!
SOME CURRENT PROJECTS

Statistics in Topological Data Analysis

Hypothesis testing with persistent homology can be done using repeated simulations following a theoretical null model for the data source. If the observed topology is much more extreme than in the null model, the test can be seen as indicating topological structure in the data.

To introduce multiple hypothesis correction we are forced to go back to basics and investigate the behavior of the null model itself, proving that

Oscillator Clustering with Path Distances

Coupled simple oscillators can synchronize – the behavior can be analyzed by representing the oscillators as paths on $S^1 \times \mathbb{R}$. Clustering these paths to determine which oscillator subgroups have synchronized can be approached using homology to calculate distances between paths on the circle.

Pursuit Strategies and the Topology of Evasion

Given a pattern for $N$ agents moving in a region and pursuing an evader, topology can determine whether or not an evasion path exists. From this determination, we can evaluate strategies for pursuit and build new heuristics to direct groups searching for an entity.

Applications of Circular Coordinates

We, and other research groups, have been applying circular coordinates to time series data – of motion capture time series, on financial data – and on directional data for HAM radio fox hunting.

We are looking for more interesting areas of application. Both for new time series questions we have not considered, but also for completely different areas of application.