Using the Mapper method to visualize high-dimensional data should capture the topological properties of the data. Today, there are few or no guarantees that the filtering functions used for the Mapper does this.

We propose:
• Using topological autoencoders (TAEs) to capture local and global topologies, constructing more representative Mapper graphs.
• TAEs capture the shape of the data through a Persistent Homology-regularized loss.
• Comparing with other filtering functions on high-dimensional manifolds to validate mapper graphs.

Our data consists of 11 100-dimensional hyperspheres originally constructed by Moor et al.¹
• 10,000 datapoints.
• Points sampled uniformly at random on surfaces of different spheres.

Results
• TAE reaches optimal separation of spheres into different simplicial complexes (see figure and table).
• TAE superior to other filtration functions in terms of separating the manifolds in this setting.

<table>
<thead>
<tr>
<th>Method</th>
<th>Empirical investigation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Compare 6 different functions for Mapper over different settings.</td>
</tr>
<tr>
<td></td>
<td>Grid search over:</td>
</tr>
<tr>
<td></td>
<td>• Overlap fraction 0.025 – 0.4 (step of 0.025).</td>
</tr>
<tr>
<td></td>
<td>• Number of intervals: 5 – 45 (step of 5).</td>
</tr>
<tr>
<td></td>
<td>Metric</td>
</tr>
<tr>
<td></td>
<td>Average amount of spheres in each node in the Mapper graph.</td>
</tr>
</tbody>
</table>

Mapper pipeline
- Filtering
- DBSCAN
- Partitioning
- Graph

**Data**

**Results**

<table>
<thead>
<tr>
<th>TAE</th>
<th>PCA</th>
<th>t-SNE</th>
<th>SVD</th>
<th>Kernel density</th>
<th>Eccentricity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>1.000</td>
<td>1.217</td>
<td>1.060</td>
<td>1.227</td>
<td>6.609</td>
</tr>
<tr>
<td>Worst</td>
<td>1.028</td>
<td>2.920</td>
<td>1.181</td>
<td>2.920</td>
<td>8.583</td>
</tr>
</tbody>
</table>

Best average number of spheres in each node. Lower is better, 1 is optimal.

**Discussion & Future work**

• Initial results show the potential of TAEs for the Mapper.
• Currently investigating higher dimensional manifolds to confirm hypothesis.
• Creating new validation metrics to validate the graphs in terms of:
  • Manifold separation.
  • Point cloud connectivity.

**Acknowledgements**

This work was partially supported by Wallenberg Autonomous Systems Program (WASP).

---