Topological Data Analysis

Topology is the mathematical discipline that deals with the structures of objects that are preserved under continuous deformations. Topological features are the features that do not change when you bend, stretch or massage your objects in any way you want as long as you do not tear them apart, glue them together or stick holes in them.
From a topological viewpoint a standard coffee cup and a doughnut are indistinguishable. In the same way a cube and a ball of the same dimension are topologically identical objects.

Topological Data Analysis is concerned with trying to impute underlying topological features from point clouds of sampled data. These topological features are often described as trying to impute "Betti Numbers" or other "properties of the different Homology Groups" for possible underlying geometrical objects.
Topological data analysis rests on several assumptions, e.g.

- There exist interesting and information bearing underlying structure in the sampled data that is connected with the topology of that underlying structure.
- The data is sufficiently densely sampled to make it possible to robustly estimate and impute the topological features of the underlying structure.
- The data is not too noisy in relation to the topological complexity. Noise can very easily cause topological instability when we try to impute topological features of lower dimensional objects sitting in a high dimensional space.

In Qlucore Omics Explorer you can visually look for indications of geometrical or topological properties in your data by using several of the built in plot types. Notably, with PCA you can inspect a low dimensional approximation (projection) of your data. If you have a strong signal connected with a geometrical object (possible to imbed in low dimensional space) you will often be able to see indications of this in a suitable low dimensional PCA-projection.

Sometimes, topological complexity can be hard to discern in a PCA-plot, but then using ISOMAP you are sometimes able to discover more complicated topological structure. In fact, if there exists a dominant signal in your data that is connected with a low dimensional topological feature, you will often be able to see indications of this using ISOMAP. It is important to note that the reverse statement is false, i.e. if you do see interesting information bearing structure in an ISOMAP plot this does not prove that this structure is per se topological in nature.

In fact, in particular PCA, but also ISOMAP are much broader tools than simply topological and in fact this goes for all tools available in Qlucore Omics Explorer. Our choices of tools in general build on broad applicability and usefulness in data analysis.
It is normally fair to assume that there exist a lot of interesting information in a data set that is not directly linked to one specific model assumption like the topology of an imaginary underlying geometrical object. It is often highly speculative to even assume the existence of any underlying geometrical or topological object in real world data analysis. Using only a

QLUCORE®

topological approach you would be disregarding most of the information in any absolute measurements and almost all of the statistical properties of your underlying variables. So if you do believe that your measurements have any relation to some absolute meaning or that the statistics of your underlying variables are of interest, you should complement any "topological explorations" of your data with several of the broadly applicable methods available in Qlucore Omics Explorer.

Disclaimer

The contents of this document are subject to revision without notice due to continuous progress in methodology, design, and manufacturing.

Qlucore shall have no liability for any error or damages of any kind resulting from the use of this document.

Qlucore Omics Explorer is only intended for research purposes.

Topological Data Analysis