AMAT 584 Lecture 14 2/21/20

<u>Today</u>: Filtrations in TDA
Euler Characteristic Curves

In the last 3 lectures, we discussed three constructions of simplicial complexes from data:
- Čech complex
- Rips complex
- Delaunay / Alpha complex.

All depend on a radius parameter r.

Often choosing a single value of r is problematic:
- Topology of the simplicial complex many be unstable w.r.t. r.
- Different values of r may capture different topological features of the data.

- Even when there is a good choice of r, we may not know it a priori.

<u>Prevailing wisdom</u>: As in hierarchical clustering, one should not consider a single value of r, but look at all values at once.

Recall the following definition from TDA I:

<u>Def</u>: A <u>filtration</u> (indexed by $[0,\infty)$) is a collection of topological spaces

$F = \{F_r\}_{r\in[0,\infty)}$ such that $F_r \subset F_s$ whenever $r<s$.

<u>This definition admits many variants:</u>

- Simplicial complexes instead of topological spaces
- Filtrations indexed by $\mathbb{N}$, $\mathbb{Z}$, or $\mathbb{R}$ instead of $[0,\infty)$.

Allowing r to vary in the various constructions we have seen gives us a filtration:

Let $X \subset \mathbb{R}^n$ be finite.

Filtrations constructed from X:

- Union-of-balls filtration $U(X) = \{U(X,r)\}_{r\in[0,\infty)}$ we saw this in TDA I
- Čech filtration $\check{C}ech(X) = \{\check{C}ech(X,r)\}_{r\in[0,\infty)}$
- Delaunay filtration $Del(X) = \{Del(X,r)\}_{r\in[0,\infty)}$

And for X any finite metric space, we have the

- Vietoris-Rips Filtration $VR(X) = \{VR(X,r)\}_{r\in[0,\infty)}$.

Persistent Homology analyzes data by first constructing a filtration of the data, and then analyzing the topology of the filtration.

In practical computations, the most common choices are the Vietoris-Rips and Delaunay filtrations.

Note that even though $[0,\infty)$ is infinite, since X is finite, the above filtrations contain only finitely many different simplicial complexes, so these are computable.

Euler characteristic curves

- One simple topological topological signature of a filtration.
- Used in some applications, e.g. recent work on glioblastoma, a a brain cancer.

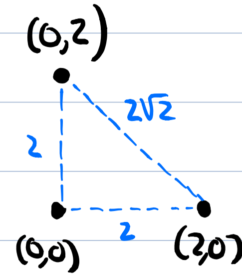
For $F = \{F_r\}_{r \in [0,\infty)}$ a simplicial filtration, let

$\chi^F : [0,\infty) \to \mathbb{Z}$ be the function given by

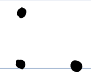
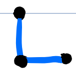
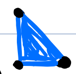
$$\chi^F(r) = \underbrace{\chi(F_r)}_{\text{Euler characteristic of } F_r.}$$

We call this the Euler Characteristic Curve (E.C.C.) of F.

Example: $X = \{(0,0), (2,0), (0,2)\}$

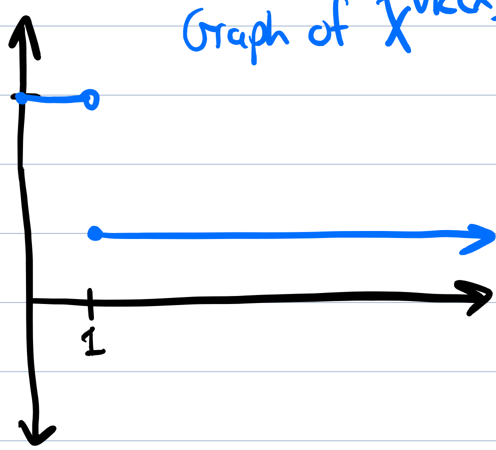


$\mathrm{VR}(X) =$

-  for $r \in [0,1)$
-  for $r \in [1, \sqrt{2})$
-  for $r \geq \sqrt{2}$

So
$$\chi^{\mathrm{VR}(X)} = \begin{cases} 3 & \text{for } r \in [0,1) \\ 1 & \text{for } r \in [1,\sqrt{2}) \\ 1 & \text{for } r \in [\sqrt{2}, \infty) \end{cases} = \begin{cases} 3 & \text{for } r \in [0,1) \\ 1 & \text{for } r \geq 1. \end{cases}$$

Graph of $\chi^{\mathrm{VR}(X)}$



Can use standard tools/ideas in stats for analyzing E.E.C.'s.

Disadvantages of the E.C.C. of data:

- Not stable (at least theoretically)
- Not readily interpreted
- Rather coarse invariant.

Using homology, we can get invariants which are

- Stable (in a sense)
- More easily interpreted
- Capture more information about the topology of the data.