Lecture 1

Course logistics

Course textbook: Oudot (see link)

Overview of Topological Data Analysis (TDA)

Motivation:

understand structure in data. Exploratory data analysis

- clusters
- loops, holes

} Topology: deformation invariance

"1-dimensional hole" or loop

& hollow void in a sphere

= 2-dimensional hole.

create invariant features for ML

- Molecular database
- texture databases
  - materials design / discovery
- genetics
- incorporation into deep learning
- computational geometry
- robotics / sensor networks.
- ...

Quick history of TDA -

Topology (Euler, Poincaré, pre-1940)

Algebraic topology (1940s - present)

idea of functor - turns a statement/property of topological spaces/maps to a statement about vector spaces/maps

Robbins '00 - early work on homology + sampling

Edelsbrunner, Letscher, Z. 02, persistent homology

Carlsson, Z. 05

Mapper alg. '07

Zigzag homology '09 - '10

---

last 10 years: - applications

and continuing - computational acceleration

- integration w/ ML/Statistics

- generalization/formalization

Simplicial complex: X

X_0: Vertex set (0-simplices) $\Big\}$ Graph (undirected)

X_1: Edge set (1-simplices)

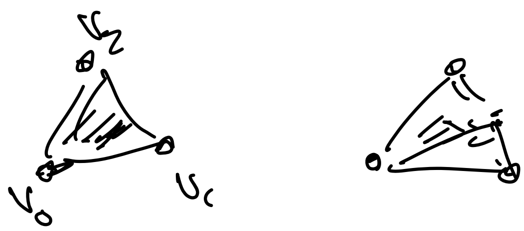
$X_0 = \{(i),\ i = 1 \dots n\}$

$X_1 \subseteq \{(i,j)\ \ i,j = 1 \dots n,\ i<j\}$

X_2: Triangles

X_k: k-simplices:

k-simplex: $(v_0 \dots v_k)$ $\ v_0 \dots v_k \in \{1 \dots n\}$



Chain complex

$$0 \xleftarrow{\partial_0} C_0 \xleftarrow{\partial_1} C_1 \xleftarrow{\partial_2} C_2 \qquad \{C_k(X), \partial_k\} = C_*$$

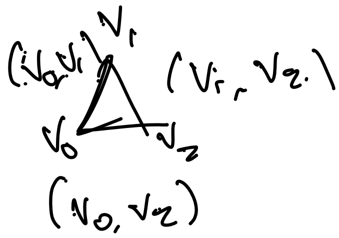
C_k: Vector space w/ basis element for each k-simplex in X $(v_0 \dots v_k)$ is a basis element

∂_k: linear maps from $C_k \to C_{k-1}$

map basis elt. for simplex to lin. comb. of basis elts in simplex boundary.

$\partial(v_0 \dots v_n) := \bigcup (v_0 \dots \hat{v}_i \dots v_k)$

$\uparrow$ removed i^{th} vertex



$$\partial_k : (v_0 \dots v_k) \mapsto \sum_{i=0}^{k} (-1)^i (v_0 \dots \hat{v}_i \dots v_k)$$

$$\partial_k \circ \partial_{k+1} = 0$$

∂_1: incidence matrix of graph.

Homology: $\underline{\partial_k} \circ \underline{\partial_{k+1}} = 0$: $\underline{H_k} = \ker \partial_k / \operatorname{img} \partial_{k+1}$

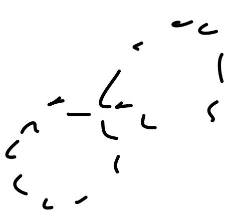
$\dim H_k = \beta_k$ $\leftarrow$ k^{th} betti number

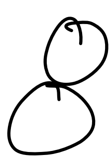
<u>counts</u> k-dimensional topological features.

$\dim H_0 =$ # connected components

$\dim H_1 =$ # loops / 1-dimensional holes.

$\dim H_k =$ # k-dimensional voids.





$\beta_0 = 1$

$\beta_1 = 2$

$\beta_2 = 0$

persistent homology: track how homology changes over a filtration.

$$X(0.2) \subseteq X(0.5) \subseteq X(1.0) \subseteq X(2.0)$$

filtration $X(r) \subseteq X(s)$ $\quad r \leq s$

$$X(r) \hookrightarrow X(s)$$

$$\Downarrow$$

$$H_k(X(r)) \xrightarrow{F_k} H_k(X(s))$$

rank

$$H_k(X_r)$$
$$\parallel$$
$$V(r) \longrightarrow V(s) \longrightarrow V(t) \longrightarrow$$

persistent homology computes

persistence barcode. $\{(b_i, d_i)\}_{i \in I}$