Topological Data Analysis and Persistence Theory

NSF/CBMS Conference

Valdosta State University

August 8-12, 2022

Peter Bubenik, University of Florida

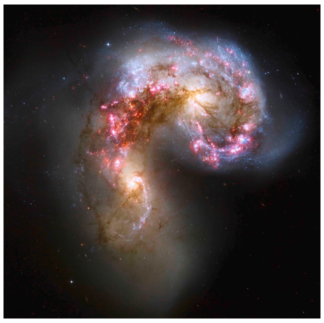
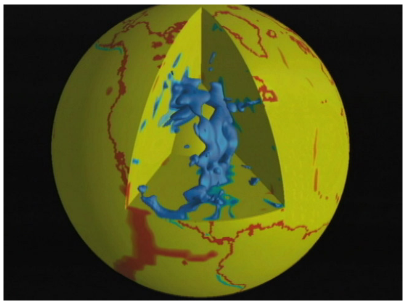
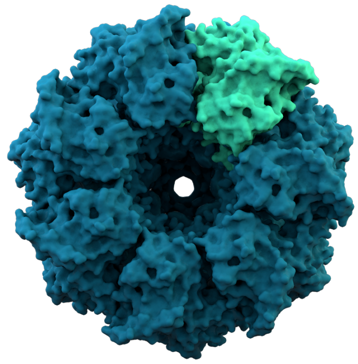
Lecture 1: Motivation and Basic Constructions

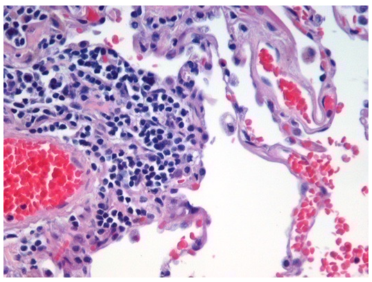
Outline:
1. Background
2. Mathematical Objects
3. Categories of Mathematical Objects
4. Homology and Persistent Homology

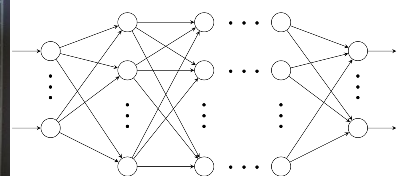
Please interrupt me!!!

1. Background

1.1 Data and Shape



1.2 Topological Data Analysis Pipeline

$$\text{Data} \longrightarrow \text{Geometry} \xrightarrow{\text{Homology}} \text{Algebra} \longrightarrow \text{Summary} \longrightarrow \text{Analysis}$$

Preprocessing
Encoding

Math

Computation

Statistics
Machine
Learning

Art

Science

2. Mathematical Objects

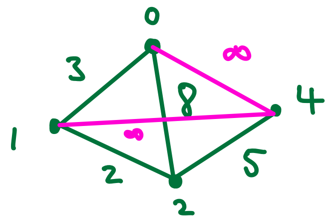
2.1 Discrete Mathematical Objects

$x_1, \dots, x_n \in \mathbb{R}^d$

~~finite~~ metric space

$d(x_i, x_j)$

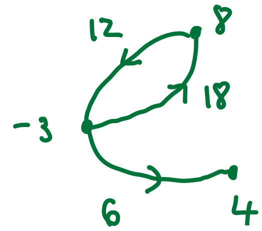
~~finite~~ weighted simple graph



$d(x_i, x_j)$

symmetric

~~finite~~ weighted simple directed graph

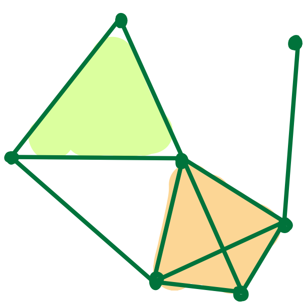


$d(x_i, x_j)$

$d : X \times X \to (-\infty, \infty]$

more general

Simplicial Complex



Geometric

(X, E)

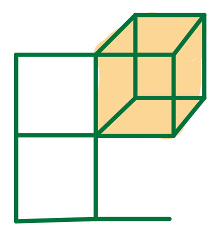
$E \subset 2^X$

$\sigma \in E, \gamma \subset \sigma, \gamma \neq \emptyset$

$\Rightarrow \gamma \in E$

Abstract

Cubical Complex



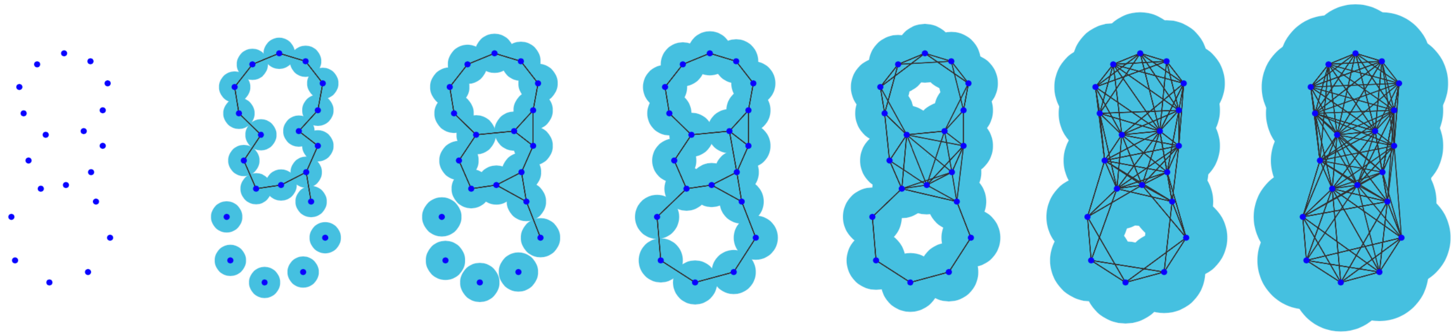
Weighted Simplicial Complex

Weighted Cubical Complex

= Grayscale Digital Image

2.1 Continuous Mathematical Objects   Topological Spaces

2.2 Constructions



Union of Balls $\underbrace{x_1, \dots, x_n}_{X} \in \mathbb{R}^d \longrightarrow \bigcup_i B_r(x_i)$ $r \geq 0$

Čech complex (X, E)

$$E = \{ \underset{\neq \phi}{\sigma} \subset X \mid \bigcap_{x_i \in \sigma} B_r(x_i) \neq \phi \}$$

Vietoris-Rips complex (X, E)

$$E = \{ \underset{\neq \phi}{\sigma} \subset X \mid \forall x_i, x_j \in \sigma \quad d(x_i, x_j) \leq 2r \}$$

Delaunay complex   Use $B_r(x_i) \cap V(x_i)$

$$V(x_i) = \{ x \in X \mid d(x, x_i) \leq d(x, x_j) \ \forall j \}$$

Filtered spaces and filtered complexes   Let r increase above

3. Categories of Mathematical Objects

3.1 Categories

Met objects: metric spaces (X, d)

morphisms: nonexpansive maps 1-Lipschitz

Gph objects: simple graphs (X, E)

morphisms: graph homomorphisms $f : X \to Y$ s.t. $(x,y) \in E$

$\Rightarrow (fx, fy) \in E$ or $fx = fy$

DiGph objects: simple directed graphs

morphisms: graph homomorphisms

Top objects: topological spaces

morphisms: continuous maps

Vect objects: vector spaces over some fixed field $\mathbb{Z}/2\mathbb{Z}$

morphisms: linear maps

$\mathbb{R}$ objects: $\mathbb{R}$ Similarly subsets of $\mathbb{R}$

morphisms: $x \le y$ e.g. $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Z}_{\ge 0}$, $\underline{n} = \{0, 1, \dots, n\}$

P (poset) objects: P

morphisms: $x \le y$

C_{mono} objects: objects in a category C Top_{mono} top spaces &

morphisms: monomorphisms 1-1 cont maps

RTop object: (X, f), $X \in Top$ $f : X \to \mathbb{R}$

morphism:

$$\begin{tikzcd} X \arrow[rr, "h"] \arrow[dr, "f"'] & & Y \arrow[dl, "g"] \\ & \mathbb{R} & \end{tikzcd}$$

$gh \le f$

3.2 Functors

$\underline{n} \to \mathrm{Vect}$ $\quad V_0 \xrightarrow{f_1} V_1 \xrightarrow{f_2} V_2 \xrightarrow{f_3} \cdots \xrightarrow{f_n} V_n$ persistence modules

$\mathbb{R} \to \mathrm{Vect}$ $\quad a \le b \quad V_a \to V_b$

$\underline{n} \longrightarrow C_{mono}$ $\quad A_0 \xrightarrow{f_1} A_1 \xrightarrow{f_2} A_2 \xrightarrow{f_3} \cdots \xrightarrow{f_n} A_n$ $\quad F_{\underline{n}} C$

$\mathbb{R} \longrightarrow C_{mono}$ $\quad F_{\mathbb{R}} C$

Examples $\quad F_{\underline{n}} \mathrm{Simp}$, $\quad F_{\mathbb{R}} \mathrm{Top}$ $\quad$ filtered simpl. complexes / top spaces

$\mathrm{Sub} : \mathbb{R}\mathrm{Top} \longrightarrow F_{\mathbb{R}} \mathrm{Top}$

$X \xrightarrow{f} \mathbb{R} \qquad X_a = \{ x \in X \mid f(x) \le a \}$ $\quad a \le b \Rightarrow X_a \subset X_b$

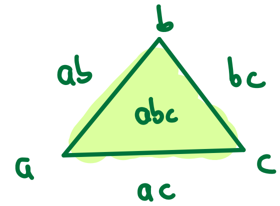
$\check{C} : \mathrm{Met} \longrightarrow F_{\mathbb{R}} \mathrm{Simp}$ $\quad$ Can replace Met with

$VR : \mathrm{Met} \longrightarrow F_{\mathbb{R}} \mathrm{Simp}$ $\quad$ wDiGph

4. Homology and Persistent Homology

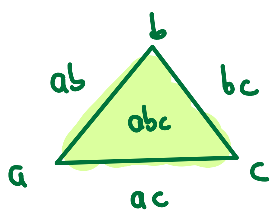
4.1 Homology

Boundary matrix of a finite simplicial complex



Cycles: C_0 = vector space with basis given by 0-cells   vertices
C_1 = vector space with basis given by 1-cells   edges
C_2 = vector space with basis given by 2-cells   triangles

Use $\mathbb{Z}/2\mathbb{Z}$ coefficients   i.e. 0 or 1



C_0 = vector space with basis a, b, c
C_1 = vector space with basis ab, ac, bc
C_2 = vector space with basis abc

Boundary maps $$\partial_j(a_0 a_1 \cdots a_j) = \sum_{i=0}^{j} a_0 \cdots \hat{a}_i \cdots a_j$$

$\partial_2(abc) = bc + ac + ab$
$\partial_1(ab) = b + a$ etc
$\partial_1(a) = 0$

In matrix notation:

$$\partial_3 = 0 \qquad \partial_2 = \begin{array}{c} \\ ab \\ ac \\ bc \end{array}\begin{array}{c} abc \\ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{array} \qquad \partial_1 = \begin{array}{c} \\ a \\ b \\ c \end{array}\begin{array}{c} \begin{array}{ccc} ab & ac & bc \end{array} \\ \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \end{array} \qquad \partial_0 = 0$$

Crucial Fact : $\partial_1 \partial_0 = 0, \quad \partial_2 \partial_1 = 0, \quad \partial_3 \partial_2 = 0$

Chain complex: $0 \to C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \to 0$

Homology vector spaces:

For $j = 0, 1, 2$, $H_j = \ker \partial_j / \operatorname{im} \partial_{j+1}$

Betti numbers: $\beta_j = \dim H_j$

$= \dim \ker \partial_j - \dim \operatorname{im} \partial_{j+1}$

$$\partial_2 \sim \begin{array}{cl} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{array}{l} \text{Row 1} \\ \text{Row 2 + Row 1} \\ \text{Row 3 + Row 1} \end{array} \end{array} \qquad \partial_1 \sim \begin{array}{cl} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} & \begin{array}{l} \text{Row 1} \\ \text{Row 2 + Row 3} \\ \text{Row 3 + Row 1 + Row 2} \end{array} \end{array}$$

$\dim \operatorname{im} \partial_3 = 0$ $\quad$ $\dim \operatorname{im} \partial_2 = 1$ $\quad$ $\dim \operatorname{im} \partial_1 = 2$ $\quad$ $\dim \operatorname{im} \partial_0 = 0$

$\dim \ker \partial_3 = 0$ $\quad$ $\dim \ker \partial_2 = 0$ $\quad$ $\dim \ker \partial_1 = 1$ $\quad$ $\dim \ker \partial_0 = 3$

$\beta_2 = 0 - 0 = 0$ $\quad$ $\beta_1 = 1 - 1 = 0$ $\quad$ $\beta_0 = 3 - 2 =$

H_0 has basis with representative cycle a (or b or c or $a+b$ or $a+c$ or $b+c$ or $a+b+c$)

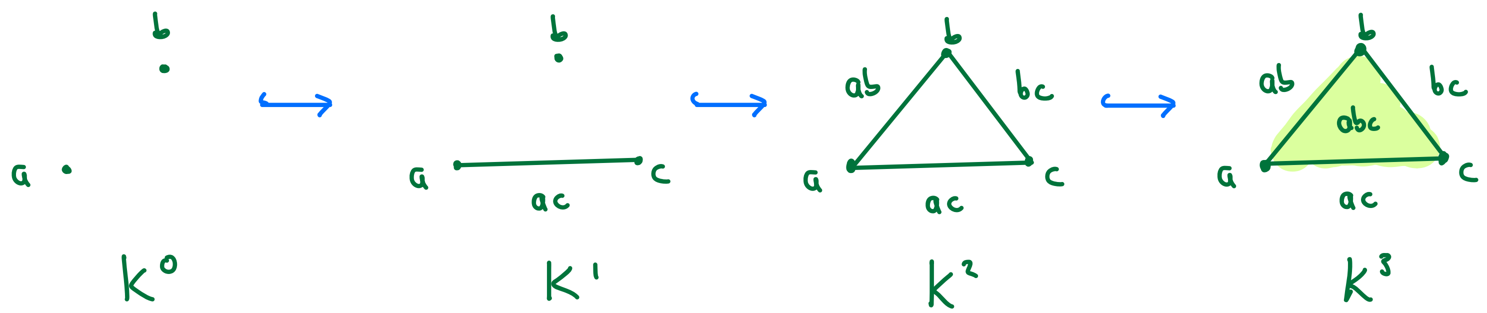
Remark: we don't need a simplicial complex

Any "cell" complex giving a chain complex ($\partial_{j+1} \partial_j = 0 \ \forall j$) will work.

e.g. cubical complex

4.2 Persistent Homology

Filtered simplicial complex :



Apply H_j. Get persistence modules:

$$\begin{array}{ccccccc}
H_0K^0 & \longrightarrow & H_0K^1 & \longrightarrow & H_0K^2 & \longrightarrow & H_0K^3 \\
(\mathbb{Z}/2)^2 & & (\mathbb{Z}/2)^2 & & \mathbb{Z}/2 & & \mathbb{Z}/2 \\
H_1K^0 & \longrightarrow & H_1K^1 & \longrightarrow & H_1K^2 & \longrightarrow & H_1K^3 \\
0 & & 0 & & \mathbb{Z}/2 & & 0
\end{array}$$

Persistent homology vector spaces : $PH_j^{i,\ell} = \operatorname{im}(H_jK^i \to H_jK^\ell)$

$$PH_0^{0,1} = (\mathbb{Z}/2)^2 \qquad PH_0^{0,2} = \mathbb{Z}/2$$

$$\begin{aligned}
PH_j^{i,\ell} &= \operatorname{im}(H_jK^i \to H_jK^\ell) \\
&= \operatorname{im}\left(\ker \partial_j^i / \operatorname{im} \partial_j^{i+1} \longrightarrow \ker \partial_j^\ell / \operatorname{im} \partial_j^{\ell+1}\right) \\
&= \ker \partial_j^i \big/ \left(\operatorname{im} \partial_j^{\ell+1} \cap \operatorname{im} \partial_j^i\right)
\end{aligned}$$

Persistent Betti Numbers : $\beta_j^{i,\ell} = \dim PH_j^{i,\ell}$