CS 6170: Computational Topology, Spring 2019
Lecture 08

Topological Data Analysis for Data Scientists

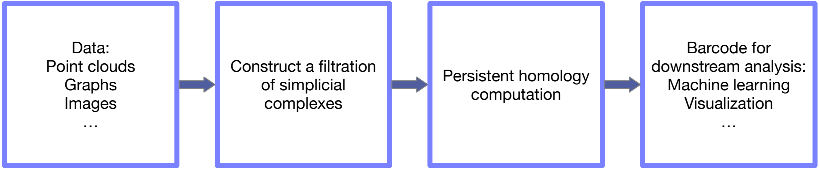
Dr. Bei Wang

School of Computing
Scientific Computing and Imaging Institute (SCI)
University of Utah
www.sci.utah.edu/~beiwang
beiwang@sci.utah.edu

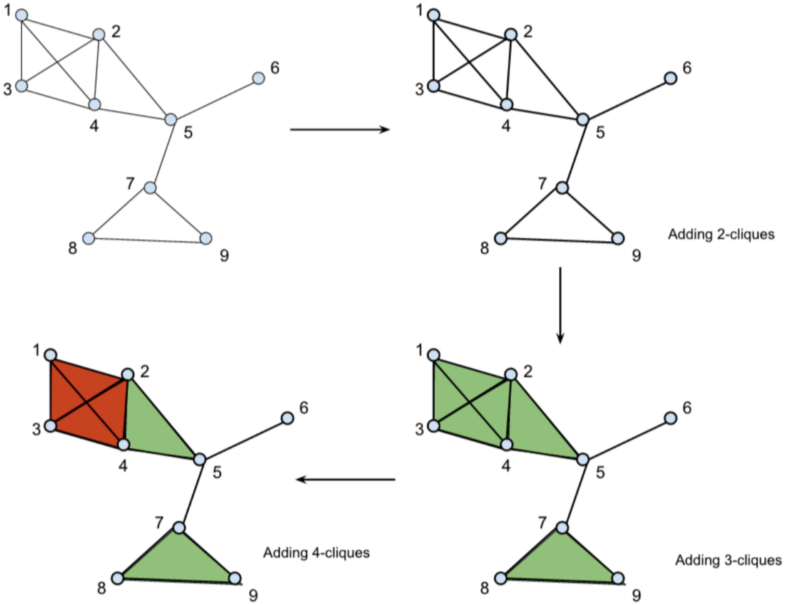
Jan 31, 2019

A TDA Pipeline

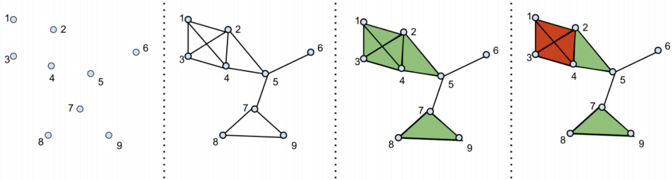
A TDA pipeline with variations



Beyond Rips Complex: Clique Complex of a graph

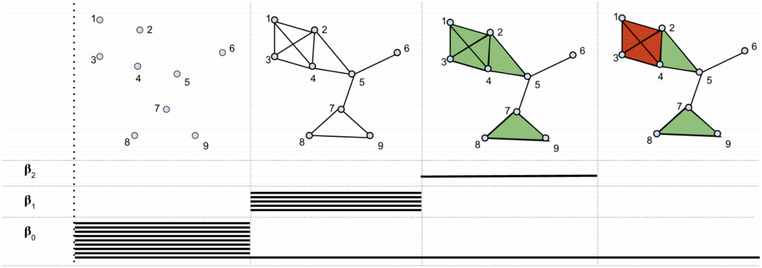


Persistent homology of a graph



β_2	0	0	1	0
β_1	0	5	0	0
β_0	9	1	1	1

Persistent homology of a graph



Persistent Homology: Mathematical Formulations

Book Chapter C.VII, C.VIII

Filtration

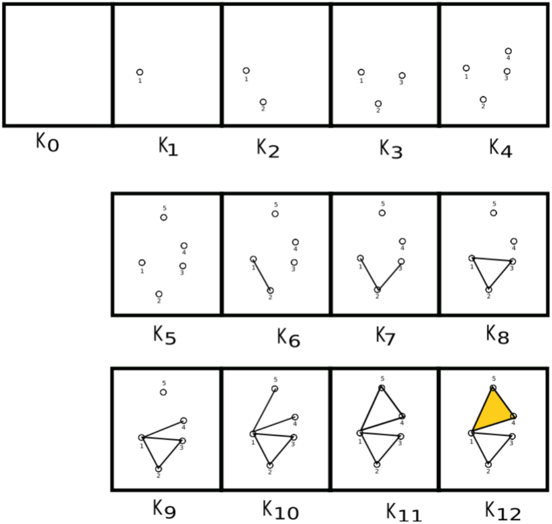
- Consider a function $f : K \to \mathbb{R}$ defined on a s.c. K
- Require f to be *monotonic*: whenever $\sigma < \tau$, $f(\sigma) \leq f(\tau)$.
- Sublevel set: $K(a) = f^{-1}(-\infty, a]$, for $a \in \mathbb{R}$
- Let $< a_1 < a_2 < \cdots < a_n$ be the function values of simplicies in K.
- Set $a_0 = \infty$, $K_i = K(a_i)$.
- A *filtration* of K is a sequence of complexes such that

$$\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K.$$

- Compute homology for each K_i gives rise to a sequence of homology groups connected by homomorphisms,

$$0 = \mathsf{H}_p(K_0) \to \mathsf{H}_p(K_1) \to \cdots \to \mathsf{H}_p(K_n) = \mathsf{H}(K).$$

Filtration: an example



Horak et al. (2009)

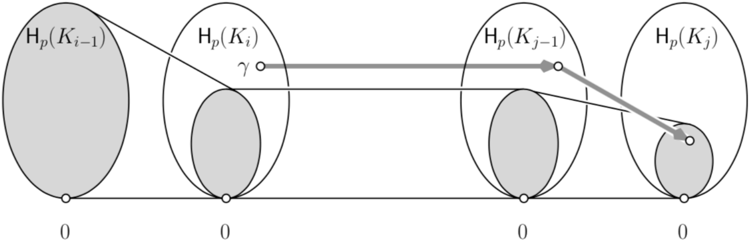
Persistence module

- Recall a *homomorphism* is a map between groups that commutes with the group operation.
- Compute homology for each K_i gives rise to a sequence of homology groups connected by homomorphisms,

$$0 = \mathsf{H}_p(K_0) \to \mathsf{H}_p(K_1) \to \cdots \to \mathsf{H}_p(K_n) = \mathsf{H}(K).$$

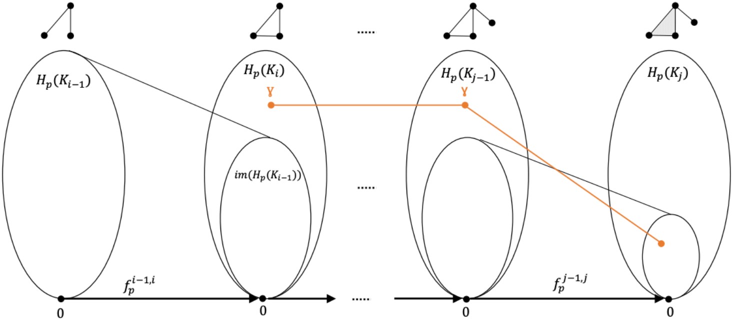
- $f_p^{i,j} : \mathsf{H}_p(K_i) \to \mathsf{H}_p(K_j)$ is a homomorphism (and a linear map) induced by inclusion.
- $f_p^{i,k} = f_p^{j,k} \circ f_p^{i,j}$
- The *p-th persistent homology groups* are the images of the homomorphisms induced by inclusions, $\mathsf{H}_p^{i,j} = \mathrm{im}\ f_p^{i,j}$, for $0 \leq i \leq n$.
- The corresponding *p-th persistent Betti numbers* are the ranks of these groups, $\beta_p^{i,j} = \mathrm{rank}\ \mathsf{H}_p^{i,j}$.

Birth and death of a homology class γ



(Edelsbrunner and Harer, 2010, page 151)

Birth and death of a homology class: an example



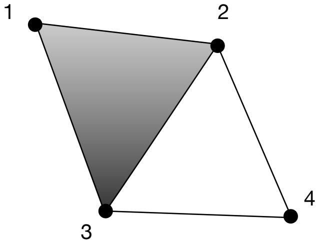
Homology and Computation

Book Chapter B.IV

Review

K: a simplicial complex. p: dimension.

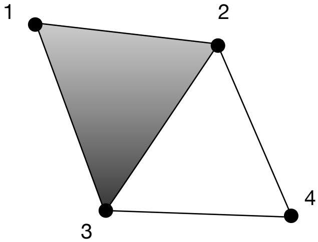
- A *p-chain* is a sum of p-simplices in K. $c \in \mathsf{C}_p$ iff $c = \sum a_i \sigma_i$, $a_i \in \{0, 1\}$.
- A *p-circle* is a p-chain with empty boundary. $c \in \mathsf{Z}_p$ iff $\partial c = 0$.
- A *p-boundary* is a p-chain that is the boundary of a $(p+1)$-chain. $c \in \mathsf{B}_p$ iff $c = \partial d$ for $d \in \mathsf{C}_{p+1}$.



Give examples of a 1-chain, 1-cycle and 1-boundary.

Review

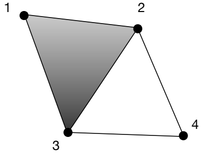
- A 1-boundary is a 1-chain that is the boundary of a 2-chain.



Review

K: a simplicial complex. p: dimension.

- The *p-th homology group* is the p-th cycle group modulo the p-th boundary group,
$$\mathsf{H}_p = \mathsf{Z}_p / \mathsf{B}_p.$$
- The element in H_p is obtained by adding all p-boundaries to a given p-cycle: $c + \mathsf{B}_p$ for $c \in \mathsf{Z}_p$.
- For example, take $c \in \mathsf{Z}_p$, $c'' \in \mathsf{B}_p$, then $c' + \mathsf{B}_p = c + \mathsf{B}_p$ since $c'' + \mathsf{B}_p = \mathsf{B}_p$.
- "Cycles that are not boundaries".



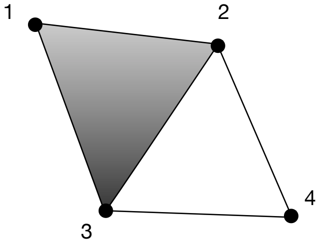
What are the elements in H_1?

Review

- The *p-th Betti number* is the rank of H_p,

$$\beta_p = \text{rank } \mathsf{H}_p.$$

- $\beta_p = \text{rank } \mathsf{Z}_p - \text{rank } \mathsf{B}_p$.



Computing Homology

See whiteboard examples.

References I

Edelsbrunner, H. and Harer, J. (2010). *Computational Topology: An Introduction*. American Mathematical Society, Providence, RI, USA.

Horak, D., Maletić, S., and Rajković, M. (2009). Persistent homology of complex networks. *JSTAT*, page P03034.