Abstract simplicial complex

Simplicial complex

A *simplicial complex* $\mathcal{C}$ is a set of *finite* sets such that if $\sigma \in \mathcal{C}$ and $\tau \subseteq \sigma$ then $\tau \in \mathcal{C}$. For every $\tau \subseteq \sigma \in \mathcal{C}$, the set τ is a *face* of σ, whereas σ is a *coface* of τ.

k-simplex

$\sigma \in \mathcal{C}$ with $|\sigma| = k + 1$ is called a *k*-simplex.

Orientation

An *orientation* of *k*-simplices is an equivalence class of orderings where two simplices are considered equal if the permutation has a sign of 1.

Geometric realization

- Realize a k-simplex as the *convex hull* of $k+1$ affinely independent points in some $\mathbb{R}^d$, with $d \geq k$.
- Need to ensure that the simplicess only intersect along *shared faces*.
- Geometric intuition:

  0-simplices vertices
  1-simplices edges
  2-simplices triangles
  3-simplices tetrahedra

Not interested in that.

Chain group

kth chain group

The *kth chain group* C_k of $\mathcal{C}$ is the *free abelian group* on the set of oriented k-simplices. The group contains all *abstract combinations* of oriented k-simplices with coefficients from either a *field* or a *principal ideal domain*.

$c \in C_k$ is a *k-chain*, i.e.

$$c = \sum_i \lambda_i[\sigma_i],$$

with $\lambda_i \in \mathbb{Z}$, for example, and $\sigma_i \in \mathcal{C}$.

Boundary operator

The kth boundary operator $\partial_k : C_k \to C_{k-1}$ is a homomorphism whose action on a chain c is defined on a simplex $\sigma = [v_0, v_1, \ldots, v_k]$ by

$$\partial_k \sigma = \sum_i (-1)^i \, [v_0, v_1, \ldots, \hat{v}_i, \ldots, v_k],$$

where $\hat{v}_i$ signifies that the ith vertex is removed from the chain.

Chain complex and subgroups

The boundary operators connect the chain groups of different dimensions. This forms a *chain complex*, i.e.

$$\cdots \to C_{k+1} \to C_k \to C_{k-1} \to \dots$$

Subgroups of C_k

We have the *cycle group* $Z_k = \ker \partial_k$ (mnemonic: "Zykel") and the *boundary group* $B_k = \operatorname{im} \partial_{k+1}$. Since $\partial_k \partial_{k+1} = 0$, the subgroups are nested:

$$B_k \subseteq Z_k \subseteq C_k$$

Homology

kth homology group

$$H_k = Z_k / B_k$$

This is well-defined because the subgroups are nested. The elements of the kth homology group are classes of *homologous cycles*. If the coefficients are taken from a field $\mathbb{F}$ then H_k becomes a *vector space*.

Betti numbers

$$\beta_k = \operatorname{rank} H_k$$

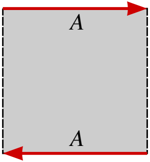
- β_0 is the number of *connected components*
- β_1 is the number of 2-dimensional holes (circles)
- β_2 is the number of 3-dimensional holes (voids)
- ...

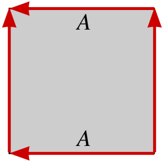
Homology is useful

- Invariants of topological spaces
- "Homology googles" to distinguish different spaces from one another

Classical example

	Möbius strip	Torus
H_0	$\mathbb{Z}$	$\mathbb{Z}$
H_1	$\mathbb{Z}$	$\mathbb{Z} \times \mathbb{Z}$
H_2	0	$\mathbb{Z}$





Input data

Assumption

The input data is given as a high-dimensional *point cloud*. There is some kind of *metric*, i.e. Euclidean distance.

Goal

Identify “interesting” topological structures in the data—especially relevant for time series data.

How to obtain a simplicial complex?

- Use points in point cloud as vertics of a graph
- Determine edges by *proximity*, i.e. take all vertices situated within a distance of ϵ
- This yields the *neighbourhood graph* N_ϵ
- Expand the graph afterwards

Chěch complex

Topologically faithful but very hard to compute. Relies on *precise* distances.

Vietoris-Rips complex

Less expensive calculation but possibly different homotopy type, i.e. we may not "see" what we want to see.

How to choose ϵ?

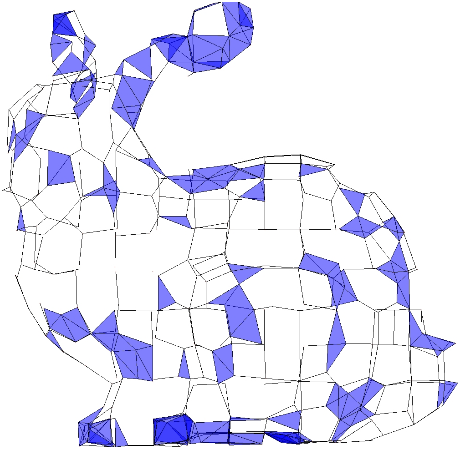


Figure: $\epsilon = 0.013$

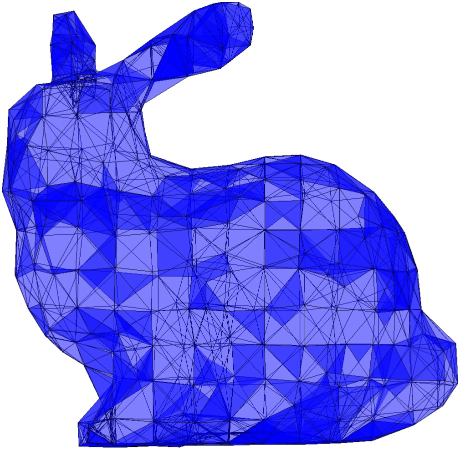


Figure: $\epsilon = 0.019$

Persistent homology

- Need to distinguish between "essential" and "non-essential" holes
- Question of "optimal" values for ϵ is a mistake

Idea

Do *all* computations for a *large range* of parameter values for ϵ. Features that *persist* over the course of varying the parameter are likely to be "real" topological features.

Visualization

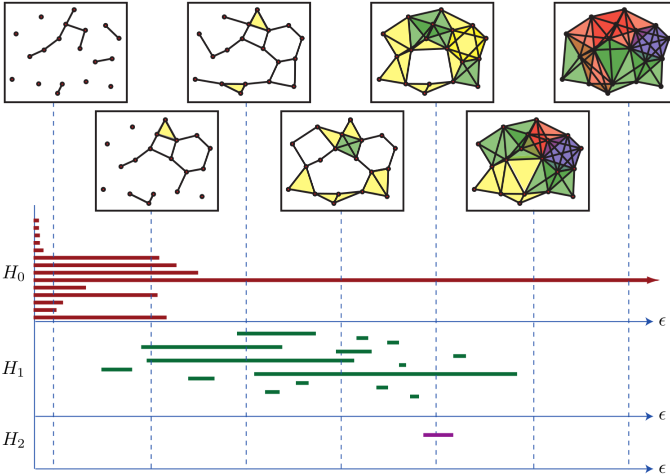
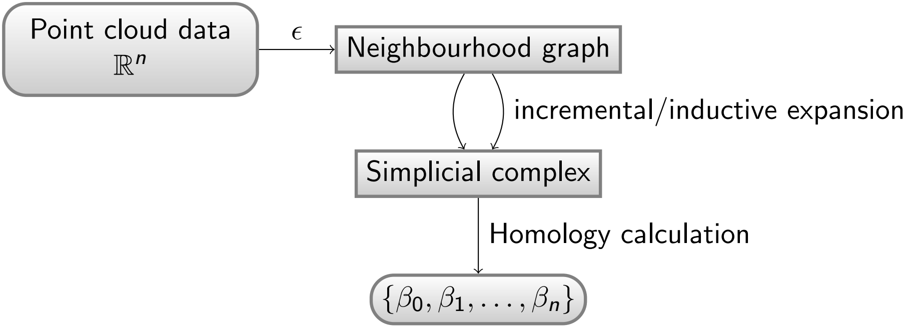


Figure: Default "barcode" visualization taken from [1].

Workflow (so far)



Current status of my work

- Literature survey; we need to know the state of the art
- Implemented algorithms for *constructing* the Vietoris-Rips complex [2]
- Started working on implementation of *persistent homology calculation* [3]

Problems

- Complexes are *very large*
- Calculations are slow
- Not many applications out there (this may be a good thing)

Roadmap

- Even more literature survey
- Examination of some data sets—how can we profit from these methods?
- Try *approximations* to topology (sometimes we know the topology of the underlying space)
- Rather vague: Use *domain knowledge*

Possible applications

- Time-series data
- Clustered data
- ?

- Robert Ghrist.
  Barcodes: The persistent topology of data.
  Bulletin of the American Mathematical Society, 45:61–75, 2008.
- Afra Zomorodian.
  Fast construction of the Vietoris-Rips complex.
  Computers & Graphics, 34(3):263–271, 2010.
- Afra Zomorodian and Gunnar Carlsson.
  Computing persistent homology.
  Discrete and Computational Geometry, 33(2):249–274, 2005.