Persistent Homology and TDA Senior Thesis

American University in Bulgaria Departament of Mathematics and Science

Supervisor: Prof. Peter Dalakov

December 11, 2020

Outline

- Chain Complex and Simplicial Homology
 - Δ complex
 - Chain Complex and Homology
 - Computing Homology
 - Maps of Complexes
- Singular Homology
- Čech and Vietoris-Rips Complex
 - Definition
 - Čech Complex Example
 - Vietoris-Rips Complex Example
- Persistent Homology
 - Persistance
 - Computations



Standard Simplex - n-simplex

A n-simplex is denoted by $[v_0, v_1, ..., v_n]$

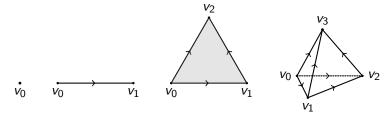


Figure 1: 0-simplex, 1-simpex, 2-simplex, 3-simplex

Δ – complex

Definition (Δ -complex)

A Δ – complex structure on a space X is a collection of maps $\sigma_{\alpha}:\Delta^{n}\to X$, with n depending on the index α , such that:

- The restriction $\sigma_{\alpha} | \mathring{\Delta}^n$ is injective, and each point of X is in the image of exactly one such restriction $\sigma_{\alpha}|\mathring{\Delta}^{n}$.
- **2** Each restriction of σ_{α} to a face of Δ^{n} is one of the maps $\sigma_{\beta}: \Delta^{n-1} \to X$. Here we are identifying the face of Δ^n with Δ^{n-1} by the canonical linear homeomorphism between them that preserves the ordering of the vertices.
- **3** A set $A \subset X$ is open iff $\sigma_{\alpha}^{-1}(A)$ is open in Δ^n for each σ_{α}

Δ – complex – S^1

Consider $X = S^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. We are going to describe explicitly two maps $\sigma_0 : \Delta^0 \to S^1$, $\sigma_1 : \Delta^1 \to S^1$, which equip S^1 with the structure of a Δ -complex.

For the explicit description, keep in mind that $\Delta^0=\{1\}\subseteq\mathbb{R}$ and that

$$\Delta^{1} = \{(t_{0}, t_{1}) | t_{0} + t_{1} = 1\} = \{(t_{0}, 1 - t_{0}), t_{0} \in [0, 1]\}$$

So in particular, any (continuous) map $\Delta^1 \to S^1$ is determined by and determines a (continuous) map $[0;1] \to S^1$.

5 / 42

First, we define $\sigma_0: \Delta^0 \to S^1$ by $\sigma_0(1) = (1,0)$. Next, $\sigma_1: \Delta^1 \to S^1$ is defined by $\sigma_1(t_0, t_1) = (\cos(2\pi t_0), \sin(2\pi t_0))$, which is clearly continuous. The map σ_1 is one-to-one on $\mathring{\Delta}^1$, and so is, trivially, σ_0 on $\mathring{\Delta}^0$. The images of the two maps cover the circle, with $\sigma_1\left(\mathring{\Delta}^1\right)=S^1ackslash\{(1,0)\}$ and $\sigma_0\left(\mathring{\Delta}^0\right) = \{(1,0)\}.$

Finally, we check the compatibility property (2) of σ_1 and σ_0 . The restrictions of σ_1 to the two faces of Δ^1 coincide with σ_0 :

$$egin{aligned} \sigma_1|_{[e_1]} &= \sigma_1(0,1) = (\cos(0),\sin(0)) = (1,0) = \sigma_0(1) \ \sigma_1|_{[e_0]} &= \sigma_1(1,0) = (\cos(2\pi),\sin(2\pi)) = (1,0) = \sigma_0(1) \end{aligned}$$

Property (3) holds as well. There are numerous other Δ -complex structures on S^1 , and we shall discuss some of them later on.

Abstract Simplicial Complex

Definition (Abstract Simplicial Complex)

Given a finite set $\{1, 2, \dots, m\} (= [m])$ an abstract simplicial complex is a collection \mathcal{K} of subsets of [m], such that:

- \bullet $\emptyset \in \mathcal{K}$
- $\{i\} \in \mathcal{K} \text{ (singletons are in } \mathcal{K} \}$
- **3** If $I \in \mathcal{K}$ and $J \subseteq I$, then $\mathcal{J} \in \mathcal{K}$

The elements of [m] are the *vertices*, and the elements of \mathcal{K} are the simplices, where $I \in \mathcal{K}$ is an (|I| - 1)-simplex.

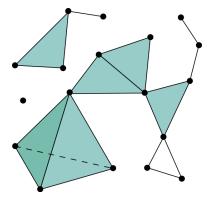
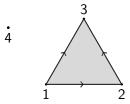


Figure 2: A simplicial 3-complex

Consider the following partially ordered set $V = \{1, 2, 3, 4\}$: The simplicial complex

$$\mathcal{K} = \{I = \{1,2,3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1\}, \{2\}, \{3\}, \{4\}, \emptyset\}$$



Given an abstract simplex K, we can construct its topological realization as

$$|\mathcal{K}| = \bigcup_{\varnothing \neq I \in \mathcal{K}} (\operatorname{Conv}(e_i), i \in I) \subseteq \mathbb{R}^m$$

where $\{e_i\}$ is in the standard basis $e_1, \dots e_m \in \mathbb{R}^m$.

Definition (Chain complex)

Complex of abelian groups.

A chain complex is a sequence of homomorphisms of abelian groups:

$$\cdots \longrightarrow C_{n+1} \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} C_{n-1} \xrightarrow{\partial_{n-1}} \cdots \longrightarrow C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0 = 0} 0$$

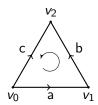
where $\partial_n \partial_{n+1} = 0$ for each n in \mathbb{Z} . The equation $\partial_n \partial_{n+1} = 0$ is equivalent to the inclusion $\text{Im } \partial_{n+1} \subset \ker \partial_n$.

The boundary map $\partial_n : \Delta_n(\mathcal{X}) \to \Delta_{n-1}(\mathcal{X})$ for the would-be chain complex $\Delta_{\bullet}(X)$ is defined as

$$\partial_n(\sigma_\alpha) = \sum_i (-1)^i \sigma_\alpha | [v_0, ..., \widehat{v_i}, ..., v_n].$$

4□ > 4ⓓ > 4 ≧ > 4 ≧ > □ ■ ♥9

Boundary operator



$$\partial[v_0,v_1]=v_1-v_0$$

$$\partial[v_0, v_1, v_2] = [v_0, v_1] + [v_1, v_2] - [v_0, v_2]$$

Homology of a Chain Complex

Definition (Homology Group)

The n-th homology group of the chain complex is defined as the quotient group

$$H_n = \frac{Z_n}{B_n} = \frac{\ker \partial_n}{\operatorname{Im} \partial_{n+1}}$$

Elements of Z_n are called cycles and elements of B_n are called boundaries.

Elements of H_n are cosets of Im ∂_{n+1} , called homology classes. Two cycles representing the same homology class are said to be homologous. This means their difference is a boundary.

Computing Homology of \mathbb{RP}^2 in \mathbb{Z}

One way to calculate the homology groups of a projective plain \mathbb{RP}^2 is by triangulating it into two 2-simplices A and B, upper triangle and lower one respectively.



We can construct the following chain complex which is a sequence of homomorphisms of abelian groups:

$$0 \xrightarrow{\partial_3 = 0} C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0 = 0} 0$$

where $\partial_n \partial_{n+1} = 0$ for each n in \mathbb{Z}

$$\begin{vmatrix} C_0 = \langle v, w \rangle \\ C_1 = \langle a, b, c \rangle \\ C_2 = \langle A, B \rangle \\ C_n = \{0\} \quad \forall n \geqslant 3 \end{vmatrix},$$

$$0 \xrightarrow{\partial_3 = 0} \mathbb{Z}^{\oplus^2} \xrightarrow{\partial_2} \mathbb{Z}^{\oplus^3} \xrightarrow{\partial_1} \mathbb{Z}^{\oplus^2} \xrightarrow{\partial_0 = 0} 0$$

The n-th homology group is defined as $H_n = \ker \partial_n / \operatorname{Im} \partial_n$ The homology groups of the projective plane are:

$$H_n^{\Delta}(\mathbb{RP}^2) \simeq \left\{ egin{array}{ll} \mathbb{Z}, & ext{for } n=0 \ rac{\mathbb{Z}}{2\mathbb{Z}}, & ext{for } n=1 \ 0 & ext{for } n \geqslant 2 \end{array}
ight.$$

Let's compute H_1 :

$$\ker \partial_1 = \langle a - b, c \rangle$$
 since

$$\partial_1(\alpha a + \beta b + \gamma c) = (\alpha + \beta)(w - v) = 0 \implies \alpha = -\beta$$

The general element in C_1 : $(\alpha a + \beta b + \gamma c) = \alpha(a - b) + \gamma c$, so the ker ∂_1 can be generated by the elements a-b and c

Im
$$\partial_2 = \langle -a + b + c, a - b + c \rangle$$
 since

$$\partial_2(\alpha A + \beta B) = \alpha(-a + b + c) + \beta(a - b + c)$$

$$H_1 = rac{\ker \partial_1}{\operatorname{Im} \partial_2} = rac{\langle \mathsf{a} - \mathsf{b}, \mathsf{c}
angle}{\langle -\mathsf{a} + \mathsf{b} + \mathsf{c}, \mathsf{a} - \mathsf{b} + \mathsf{c}
angle}$$

The group $\langle a-b,c\rangle$ can be also generated by the elements

$$m=a-b+c$$
, and c where $a-b=m-c$. So, $H_1=\frac{\langle a-b,c\rangle}{\langle -a+b+c,a-b+c\rangle}=\frac{\langle a-b+c,c\rangle}{\langle a-b+c,-a+b+c\rangle}$

$$H_1 = \frac{\langle a-b,c \rangle}{\langle -a+b+c,a-b+c \rangle} = \frac{\langle a-b+c,c \rangle}{\langle a-b+c,-a+b+c \rangle}$$

If we let t = a - b + c then -a + b + c = -t + 2c then the group

 $\langle t, -t + 2c \rangle$ can be also generated by the elements t and 2c.

In terms of t and c,
$$H_1=rac{\langle t,c
angle}{\langle t,2c
angle}=rac{\langle c
angle}{\langle 2c
angle}\simeqrac{\mathbb{Z}}{2\mathbb{Z}}$$

4□ > 4□ > 4 □ > 4 □ > 4 □ > 4 □ >

Maps of Chain Complexes

Let (C_{\bullet}, ∂) and (D_{\bullet}, δ) be two chain complexes. A map of chain complexes is a morphism f that is a sequence of homomorphisms $(f_n)_{n \in Z}$:

$$(C_{\bullet}, \partial) \qquad C_{\bullet} \dots \longrightarrow C_{n} \xrightarrow{\partial n} C_{n-1} \xrightarrow{\partial n-1} C_{n-2} \xrightarrow{\partial n-2} \dots C_{\bullet}$$

$$\downarrow^{f_{n}} \qquad \downarrow^{f_{n-1}} \qquad \downarrow^{f_{n-2}}$$

$$(D_{\bullet}, \delta) \qquad D_{\bullet} \dots \longrightarrow D_{n} \xrightarrow{\delta n} D_{n-1} \xrightarrow{\delta n-1} D_{n-2} \xrightarrow{\delta n-2} \dots D_{\bullet}$$

$$f_n: C_n \to D_n$$
 s.t, $f_{n-1} \circ \partial n = \delta_n \circ f_n \ \forall n \in \mathbb{Z}$

$$C_n \xrightarrow{\partial n} C_{n-1}$$

$$\downarrow^{f_n} \qquad \downarrow^{f_{n-1}} commutes.$$

$$D_n \xrightarrow{\delta n} D_{n-1}$$

◆ロト ◆個ト ◆差ト ◆差ト を めらぐ

Maps of Chain Complexes induce Maps on Homology

A homomorphism of chain complexes induces a homomorphism on the homology. The induced map can be defined as:

$$H_n(f): H_n(C_{\bullet}) \to H_n(D_{\bullet})$$

 $H_n(f): [x] \mapsto [f_n(x)]$

To prove the claim above it is enough to check that $H_n(f)$ is well-defined. We can prove well-defines by checking if cycles are send to cycles and boundaries to boundaries.

(1) Let us take a cycle $x \in C_n$, so that $x \in \ker(\partial_n)$, $\partial_n(x) = 0$

$$\delta_n \circ f_n(x) = f_{n-1} \circ \partial n(x) = f_{n-1}(0) = 0 \Rightarrow f_n(x) \in \ker \delta_n, f_n(x) \text{ is a cycle}$$

$$\Rightarrow f_n(\ker \partial n) \subseteq \ker \delta_n$$

So, cycles are send to cycles.

◆□▶ ◆□▶ ◆■▶ ◆■ ◆○○○

(2) Let us take a boundary $y \in C_n$, so that $y \in \text{Im } \partial_{n+1} \Rightarrow \exists z \in C_{n+1}$ such that $\partial_{n+1}(z) = v$

$$f_n(y) = f_n(\partial_{n+1}(z)) = \delta_{n+1}(f_{n+1}(z))$$

$$\Rightarrow f_n(y) \in \operatorname{Im} \partial_{n+1} f_n(y) \text{ is a boundary}$$

$$\Rightarrow f_n(\operatorname{Im} \partial_{n+1}) \subseteq \operatorname{Im}(\delta_{n+1})$$

So. boundaries are send to boundaries.

$$H_n(f): H_n(C_{\bullet}) \to H_n(D_{\bullet})$$

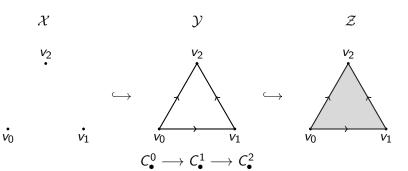
$$H_n(f): \ker \partial_n / \text{Im}(\partial_{n+1}) \to \ker \delta_n / \text{Im}(\delta_{n+1})$$

$$[x] \mapsto [f_n(x)]$$

$$x + \text{Im} \, \partial_{n+1} \mapsto f_n(x) + f_n(\text{Im} \, \delta_{n+1}) = f_n(x) + Im(\delta_{n+1}) = [f_n(x)]$$

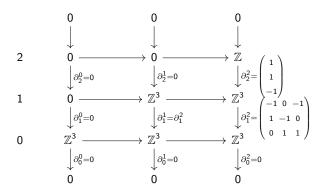
◆□▶ ◆□▶ ◆■▶ ◆■ ◆○○○

Filtered Complex



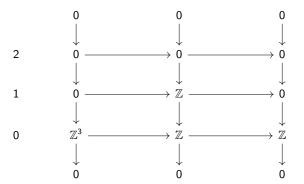
Maps of Complexes

$$C^0_{ullet} \longrightarrow C^1_{ullet} \longrightarrow C^1_{ullet}$$



Maps of Complexes induce maps on Homology

$$H(C^0_{ullet}) \longrightarrow H(C^1_{ullet}) \longrightarrow H(C^2_{ullet})$$



Exact Sequence

Consider a short exact sequence of chain complexes:

$$0 \to A_{\bullet} \xrightarrow{i} B_{\bullet} \xrightarrow{\pi} C_{\bullet} \to 0$$

 A_{ullet} , B_{ullet} , C_{ullet} are chain complexes and i,π are maps between chain complexes where $\ker \pi = \operatorname{Im} i$, $\pi : surjective \ and \ i : injective \ induces long exact sequence of homology:$

$$H_{n+1}(C_{\bullet}) \xrightarrow{\partial n+1} H_n(A_{\bullet}) \xrightarrow{i_*} H_n(B_{\bullet}) \xrightarrow{\pi_*} H_n(C_{\bullet}) \xrightarrow{\delta} H_{n-1}(A_{\bullet}) \rightarrow H_{n-1}(B_{\bullet}) \rightarrow H_{n-1}(C_{\bullet})$$

The Equivalence of Simplicial and Singular Homology

A singular n-simplex in a space X is a map $\sigma: \Delta^n \to X$

 $C_n(X)$ is a free abelian group with generators the set of singluar n-simplexes in X: the continous maps $\sigma:\Delta^n\to X$.

Theorem

The homomorphisms $H_n^{\Delta}(X,A) \to H_n(X,A)$ are isomorphisms for all n and all Δ -complex pairs (X,A).

Mayer-Vietoris sequence

In addition to the long exact sequence of homology groups for a pair (X,A), there is another sort of long exact sequence, known as a Mayer-Vietoris sequence, which is equally powerful but is sometimes more convenient to use. For a pair of subspaces $A,B\subset X$ such that X is the union of the interiors of A and B, this exact sequence has the form

$$\dots \to H_n(A \cap B) \xrightarrow{\phi} H_n(A) \oplus H_n(B) \xrightarrow{\Psi} H_n(X) \xrightarrow{\partial} H_{n-1}(A \cap B) \to \dots \to H_0(X)$$

The Mayer-Vietoris sequence is then the long exact sequence of homology groups associated to the short exact sequence of chain complexes formed by the short exact sequences

$$0 \to C_n(A \cap B) \xrightarrow{\Phi} C_n(A) \oplus C_n(B) \xrightarrow{\Psi} C_n(A+B) \to 0$$

The Nerve Theorem

Consider X is a topological space, and $X = \bigcup_{\alpha} \mathcal{U}_{\alpha}$ where $\mathcal{U} = \{\mathcal{U}_{\alpha}\}_{\alpha \in A}$ are cover indexes.

Theorem (The Nerve Theorem)

If X is a paracompact space, and U is an open cover of X such that the intersection of any finite subfamily of U is either empty or contractible, then the realization of the nerve of U is homotopy equivalent to X.

If \mathcal{X} is a metric space and $r \geq 0$:

Definition (Čech Complex)

Given a collection of points $\{x_{\alpha}\}$ in Euclidean space \mathbb{E}^n , the Čech Complex, \mathcal{C}_{ϵ} , is the abstract simplicial complex whose k-simplices are determined by unordered (k+1)-uple of points $\{x_{\alpha}\}_0^k$ whose closed $\epsilon/2-ball$ neighborhoods have a point of common intersection.

Definition (Vietoris-Rips Complex)

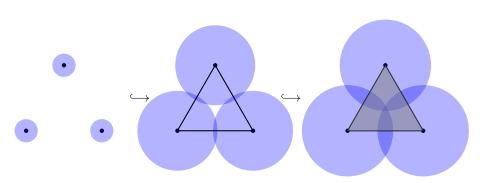
Given a collection of points $\{x_{\alpha}\}$ in Euclidean space \mathbb{E}^n , the Rips Complex, \mathcal{R}_{ϵ} , is the abstract simplicial complex whose k-simplices correspond to unordered (k+1)-uple of points $\{x_{\alpha}\}_0^k$ which are pairwise within distance ϵ .

Relation of Čech and Vietoris-Rips Complex: For each $\epsilon>0$, there is a chain inclusion maps

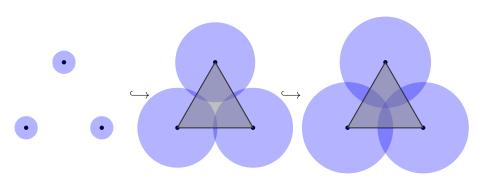
$$\mathcal{R}\hookrightarrow\mathcal{C}_{\epsilon\sqrt{2}}\hookrightarrow\mathcal{R}_{\epsilon\sqrt{2}}$$



Čech Complex Example



Vietoris-Rips Complex Example



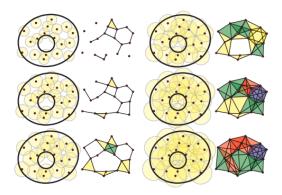


Figure 3: A sequence of Rips Complex from a point clound data set that represent an annulus

Definition

Given a filtered complex, the i-th complex K^i has associated boundary operators ∂^{i_k} , matrices M^i_{ν} , and groups C^i_{ν} , Z^i_{ν} , B^i_{ν} , and H^i_{ν} for all $i, k \ge 0$ The p-persistent k-th homology group of K^i is

$$H_k^{i,p} = Z_k^i / (B_k^{i+p} \cap Z_k^i)$$

Example:
$$p = i = k = 1$$
: $H_1^{1,1} = Z_1^1 / (B_2^1 \cap Z_1^1) \mathbb{Z} / (\{0\} \cap \mathbb{Z}) = \mathbb{Z}$



Definition (Persistence Complex)

A persistence complex \mathcal{C} is a family of chain complexes $\{C_*^i\}_{i>0}$ over R, together with a chain map's $f_i: C_*^i \to C_*^{i+1}$ so that we have the following diagram:

$$C^0_* \xrightarrow{f^0} C^1_* \xrightarrow{f^1} C^2_* \xrightarrow{f^2} \dots$$

Definition (Persistence Module)

A persistence module \mathcal{M} is a family of R-modules, M', together with homomorphism $\varphi^i: M^i \to M^{i+1}$

Suppose we have a persistence module $\mathcal{M}=\{M^i,\varphi^i:M^i\to M^{i+1}\}$ over a ring R, We can equip R[t] with the standard grading and define a graded module over R[t] by

$$\alpha(M) = \underset{i \geq 0}{\oplus} M_i$$

, where the R-module structure is the sum of the structures on the individual components, and where the action of t is given by:

$$t \cdot (m^0, m^1, \cdots) = (0, \varphi^0(m^0), \varphi^1(m^1), \cdots)$$

$$\begin{pmatrix} 0 & & & \\ \varphi^0 & 0 & & \\ \hline & \varphi^1 & 0 & \\ \hline & & \varphi^2 & \end{pmatrix}$$

t simply shift elements of the module up in gradation.

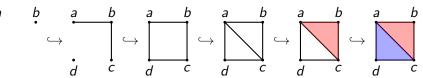
<ロ > ← □

Theorem (Correspondence)

The correspondence α defines an equivalence of categories between the category of persistence modules of finite type over R and the category of finitely generated non-negatively graded modules over R[t].

The Correspondence theorem gives us a simple decomposition when the ground ring is a field F. In this case the graded ring F[t] is a PID and its only graded ideals are homogeneous of form (t_n) , so the structure of the F[t] — module is described in structure theorem:

$$\left(\bigoplus_{i} \Sigma^{\alpha_i} F[t]\right) \oplus \left(\bigoplus_{j} \Sigma^{\gamma_j} F[t] / (t^{n^j})\right). \tag{1}$$



$$C^0 \longrightarrow C^1 \longrightarrow C^2 \longrightarrow C^3 \longrightarrow C^4 \longrightarrow C^4$$

5

In particular, consider the last simplicial complex in the filtration:



The chain complex is $0 \to \mathbb{Z}^2 \xrightarrow{\partial_2} \mathbb{Z}^5 \xrightarrow{\partial_1} \mathbb{Z}^4 \to 0$. The matrices M_2 and M_1 of ∂_2 and ∂_1 without the indicated bases are:

Then ∂_1 (ie M_1) induces a map of $\mathbb{Z}[t]$ - modules $\mathbb{Z}[t]^{\oplus 5} \xrightarrow{M_1} \mathbb{Z}[t]^{\oplus 4}$:

$$\underline{V} = \left(egin{array}{c} p_1(t) \\ p_2(t) \\ p_3(t) \\ p_4(t) \\ p_5(t) \end{array}
ight) \mapsto M_1 \underline{V}_{ullet}$$

Using the grading (filtration) on the simplicial complex we get a chain complex of graded $\mathbb{Z}[t]$ - modules

 $0 \to (t^4) \oplus (t^5) \xrightarrow{\partial_2^1} (t)^{\oplus 2} \oplus (t^2)^{\oplus 2} \oplus (t^3) \xrightarrow{\partial_1^1} (t)^{\oplus 2} \oplus (1)^{\oplus 2} \to 0$ by the procedure defined on. The respective matrices are (without same bases)

$$M_2' = \begin{array}{c} ab \\ ab \\ bc \\ cd \\ ad \\ ac \end{array} \begin{pmatrix} abc & acd \\ t^3 & 0 \\ t^3 & 0 \\ 0 & t^3 \\ 0 & -t^3 \\ -t^3 & t^3 \end{pmatrix}, M_1' = \begin{array}{c} d \\ c \\ b \\ b \\ a \end{array} \begin{pmatrix} ab & bc & cd & ad & ac \\ 0 & 0 & t & t & 0 \\ 0 & 1 & -t & 0 & t^2 \\ t & -t & 0 & 0 & 0 \\ -t & 0 & 0 & -t^2 & -t^2 \\ -t & 0 & 0 & -t^2 & -t^2 \end{array}$$

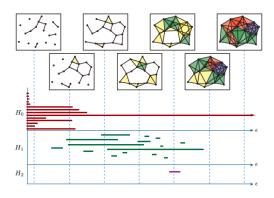


Figure 4: An example of barcode representations of the homology of the sampling of points in an annulus

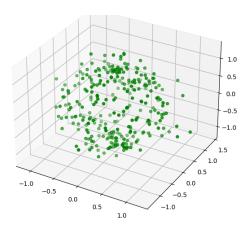


Figure 5: Points on a sphere with noise

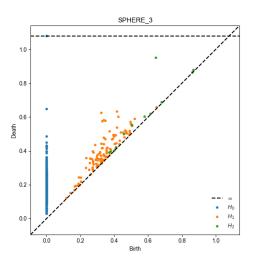


Figure 6: Persistent diagram of homology of sphere (300 points)

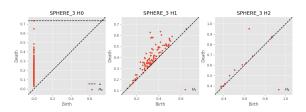


Figure 7: Persistent diagram of homology of sphere, separate (300 points)

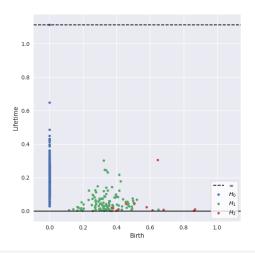


Figure 8: Lifetime persistence

References

- A. Zomorodian and G. Carlsson, "Computing Persistent Homology," Discrete Comput. Geom., 33, (2005), 249274.
- A. Hatcher, Algebraic Topology, Cambridge University Press, (2002).
- V. de Silva and G. Carlsson. "Topological estimation using witness complexes," in SPBG04 Symposium on Point-Based Graphics (2004), 157-166
- Perea, J.A., Harer, J. Sliding Windows and Persistence: An Application of Topological Methods to Signal Analysis. Found Comput Math 15, 799838 (2015).
- H. Kantz and T. Schreiber, Nonlinear Time Series Analysis, Cambridge University Press, 2003.
- Ghrist, Robert. (2008). Barcodes: The persistent topology of data.
 BULLETIN (New Series) OF THE AMERICAN MATHEMATICAL SOCIETY. 45. 10.1090/S0273-0979-07-01191-3.