A Survey of Persistent Homology

Sean Ippolito

March 29, 2019

Introduction

Motivating Examples

Example 1. The Fundamental Group Functor

- The *functor* π_1 : **Top** $\to$ **Group**
- Sends a topological space X to the group associated with its homotopy classeses $\pi_1(X)$
- sends continuous objects to discrete ones

Example 2. Shape of a Point Cloud

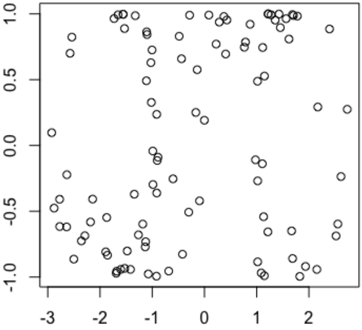


Figure 1:

Example 2. Shape of a Point Cloud

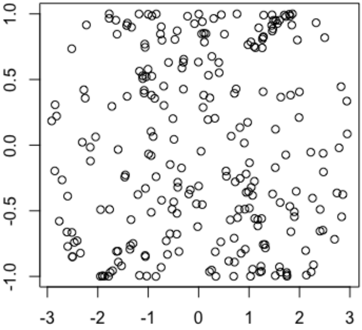
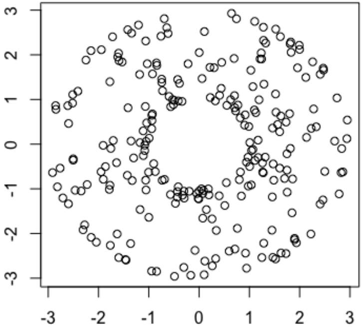


Figure 2:

Example 2. Shape of a Point Cloud

Torus!



Example 3. What do you see?

Figure 4:

The Problem at Hand

- Given a set of points/data can we reconstruct its shape?
- Can we reliably recover the topology of an object?

Preliminaries

Basic Category Theory

Posets & Diagrams

a *poset* is a set P with binary relation $<$ such that the relation is

- irreflexive: $x \not< x$
- anti-symmetric: $x < y$ and $y \not< x$ implies $x = y$
- transitivity: $x < y$ and $y < z \implies x < z$

Example: all totally ordered sets are posets

...think of $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$

Example: let P be a set of subsets of S, like the topology of S

...inclusion $P \hookrightarrow S$ induces a partial order

Posets & Diagrams

define the *oppositeposet* P^{op} by taking the poset $(P, <)$ and defining the opposite poset $>$ as $x > y$ for every $x < y$ in P

define the *product poset* by taking posets P, Q and define a partial order on the product $P \times Q$ where

$$(p, q) \leq (r, s) \iff p \leq r \text{ and } q \leq s$$

$\ldots$ the product poset induces a partial order on P^n for $n \geq 0$

Diagrams

a relation $x < y$ is minimal if it does not factor any further into $x < z < y$

we can visually represent a poset $(P, <)$ as a directed graph with verticies P and edges as the relations $<$

call this the *(Hasse) diagram*

Categories

A category $\mathcal{C}$ consists of

- a collection of objects $obj(\mathcal{C})$
- morphisms between objects $hom(x, y)$ for every $x, y \in obj(\mathcal{C})$
- and a composition rule such that

$$f \in hom(x, y) \text{ and } g \in hom(y, z) \implies g \circ f \in hom(x, z)$$

- identity morphisms: $\forall x \in obj(\mathcal{C})$, $\exists! Id_x \in hom(x, x)$

Categories

Satisfying the properties

- $Id_y \circ f = f = f \circ Id_x$ for any $x \xrightarrow{f} y$
- associative composition $(h \circ g) \circ f = h \circ (g \circ f)$ for

$$x \xrightarrow{f} y \xrightarrow{g} z \xrightarrow{h} w$$

Categories: Examples

Category	Objects	Morphisms
Set	sets	functions
Grp	groups	group homomorphisms
Top	top spaces	continuous functions
$\mathbf{Vect_k}$	vector spaces over field k	linear transformations
Poset	posets	order preserving functions

Functors

for categories $\mathcal{C}$ and $\mathcal{D}$ a functor $F : \mathcal{C} \to \mathcal{D}$ consists of

- an object $F(x) \in \mathcal{D}$ for every $x \in \mathcal{C}$
- a morphism $F(f) \in hom(F(x), F(y))$ for every $f \in hom(x, y)$

this morphism F respects composition and maps identities to identities

Functors: Example

The Fundamental Group Functor

- The *functor* π_1 : **Top** $\rightarrow$ **Group** maps

$f \in hom(X, Y)$ to $\pi_1(f) \in hom(\pi_1(X), \pi_1(Y))$

- for $X, Y \in$ **Top** and $\pi_1(X), \pi_1(Y) \in$ **Grp**

Functors: Even Better Examples

- H_i : $\mathbf{Top} \to \mathbf{Grp}$ for $i \geq 0$ and coeffieceints in $\mathbb{Z}$
- H_i : $\mathbf{Top} \to \mathbf{Vect_k}$ for $i \geq 0$ and coeffieceints in a field K

Simplicial Homology & Betti Numbers

Simplex

an *n-simplex* σ is $n+1$ collection of points $(x_0, ..., x_n)$ such that

- the collection of points is a set of *verticies* $V = (x_0, ..., x_n)$
- every simplex σ induces a total ordering on its verticies le_σ by saying
$$p \leq q \implies x_p \leq_\sigma x_q$$
- a *face* F of the simplex is is a subset of the collection of verticies so that $F \subseteq V$

Simplicial Complexes

an abstract *simplicial complex* is a fininte collection K of simplexes σ such that any face $F \in \sigma$ is also a simplex in K

say the finite collection is indexed by m where

$$K = \cup_i^m \{\sigma_i\}$$

Simplicial Chain

a *simplicial k-chain* C_k is a formal sum of k-simplicies σ_i

$$C_k = \Sigma r_i \sigma_i$$

where $r_i \in R$ ring and σ_i for $0 \leq i \leq m \in \mathbb{N}$

- we can say $C_k =< \sigma_i : 0 \leq i \leq m >$

an *R-module* is the set of all C_k with formal addition over R

Boundaries and Chains

the *boundary* ∂ of a k dimensional simplex σ is the formal sum of all $k-1$ dimensional faces of σ such that

$$\partial_k : C_k \to C_{k-1}$$

the *chain complex* $(C.,\partial.)$ is a sequence of R-modules with boundary maps ∂ such that $\partial_{k-1} \circ \partial_k = 0$ for all k

- for a finite simplicial complex S, the chain complex over a field F induces a vector space

Boundaries and Chains: Continued

- $\partial_{k-1} \circ \partial_k = 0$ in other words... $im\partial_{k+1} \subseteq ker\partial_k$

proof: $\sigma \in im\partial_{k+1}$ implies $\exists\tau$ simplex with $dim(\tau) = k+1$ such that $\partial_{k+1}\tau = \sigma$

$\implies \partial_k(\partial_{k+1}\tau) = \partial_k\sigma = 0 \implies \sigma \in \ker \partial_k \; \square$

- call the $ker\partial_k = Z_k$ the module of cycles
- call the $im\partial_k = B_k$ the module of boundaries
- rewrite previous statement as $B_{k+1} \subseteq Z_k$

Homology

define *k th homology module* of complex S as

$$H_k(S) = ker\partial_k / im\partial_{k+1} = Z_k / B_{k+1}$$

note the k homology of a space is composed of it's torsion and non-torsion coefficients

$$H_k(S) = \bigoplus_i^n \mathbb{Z} + \bigoplus_j^m \mathbb{Z}_{l_j}$$

Homology Examples

Space	H_0	H_1	H_2
Torus T	$\mathbb{Z}$	$\mathbb{Z} \bigoplus \mathbb{Z}$	$\mathbb{Z}$
Projective Plane P	$\mathbb{Z}$	$\mathbb{Z} \mod 2$	0
Sphere S^2	$\mathbb{Z}$	0	$\mathbb{Z}$

Betti Numbers

the k^{th} *betti number* of a space X is defined as $b_k(X) = rankH_k(X)$

Theorem: the euler characteristic $\chi(X) = \sum_i (-1)^i b_i$

example: $\chi(S^2) = 2$ since

- $H_0(S^2) = \mathbb{Z} \implies b_0 = 1$
- $H_1(S^2) = 0 \implies b_1 = 0$
- $H_2(S^2) = \mathbb{Z} \implies b_2 = 1$
- $H_m(S^2) = 0 \implies b_m = 0, \ \forall m \geq 3$

thus $\chi(S^2) = b_0 - b_1 + b_2 - 0 + 0... = 1 - 0 + 1 = 2$

Persistent Homology

Filtrations & Persistence

The Persistent Homology Pipeline

Data > Filtration > Homology > Apply Structure > Analyze Results

1. Take your data and apply a filtration
2. Calculate homology of your filtration to get a persistence module
3. Apply the structure theorem to get a barcode

Filtrations

let J be a poset category $\mathbb{N}, \mathbb{Z}, \mathbb{R}$ or one of their opposite posets

then a J *indexed filtration* is a functor $F : J \to \textbf{Top}$ such that

$$F_r \subset F_s \text{ whenever } r \leq s$$

- an $\mathbb{N}$ indexed filtration

$$F_1 \hookrightarrow F_2 \hookrightarrow F_3 \hookrightarrow F_4 \hookrightarrow \ldots$$

Persistence Modules

a J *indexed persistence module* M is a functor $F : J \to \mathbf{Vect_k}$

- an $\mathbb{N}$ indexed persistence module

$$M_1 \to M_2 \to M_3 \to M_4 \to \ldots$$

an *interval* I in J is a subset such that $a < b < c \in J$ and $a, c \in I$ then $b \in I$

Persistence Modules: Continued

for interval I, the *interval module* K^I to be the persistence module such that

$$K_r^I = \begin{cases} K & \text{if } r \in I \\ 0 & \text{otherwise} \end{cases} \qquad K_{r,s}^I = \begin{cases} id_k & \text{if } r \leq s \in I \\ 0 & \text{otherwise} \end{cases}$$

- **example**: an interval module over $\mathbb{N}$ looks like

$$0 \to 0 \to k \xrightarrow{id_k} k \xrightarrow{id_k} k \to 0 \to 0 \to ...$$

a persistence module M is *pointwise finite dimensional* (pfd) is $dimM < \infty$ for all $r \in R$

Structure Theorem for Persistence Modules

Theorem (Structure of Persistence Modules)

if M is $\mathbb{R}$ or $\mathbb{Z}$ indexed pfd persistence module, then there exists a unique multiset of intervals B_M such that

$$M \cong \bigoplus_{I \in B_M} K^I$$

we call the correspondence B_M the *barcode* of M

note: $\mathbb{Z}$-indexed by Webb (1985), $\mathbb{R}$-indexed by Crawley-Boevy (2012)

note: fininte $\mathbb{Z}$ or $\mathbb{R}$ cases are a variation structure theorem for finitely generated modules over a principal ideal domain

Persistence Modules: Continued

we say M is *essentially discrete* if there is an injection $j : \mathbb{Z} \hookrightarrow \mathbb{R}$ with $lim_{\pm\infty} j(z) = \pm\infty$ such that

$$\forall z \in \mathbb{Z} \text{ and } r \leq s \in [j(z), j(z+1))$$

that $M_{r,s}$ is an isomorphism

- intervals in a barcode of an essentially discrete persistence module take the form $[a, b)$ for $a < b \in \mathbb{R} \cup \{\infty\}$

More Filtrations. . .

a *sublevel-filtration* $S^{\uparrow}(f)$ to be the $\mathbb{R}$ indexed filtration for a topological space X where $f : X \to \mathbb{R}$ where

$$S^{\uparrow}(f)_r = \{p \in X | f(p) \leq r\}$$

example: (Union of balls filtration) let $P \subset \mathbb{R}^n$ be finite set of points, let $d_P : \mathbb{R}^n \to [0, \infty)$ be

$$d_P(x) = \min_{y \in P} \|x - y\|$$

thus $S^{\uparrow}(d_P)_r$ is the union of balls radius r over P

Union of balls filtration

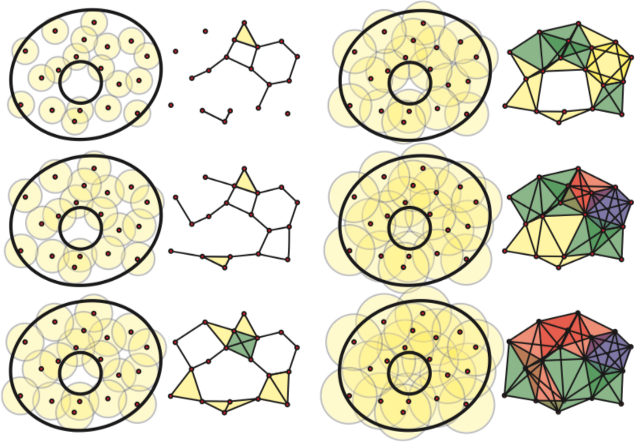


Figure 5:

Making a Simplicial Complex from a Filtration

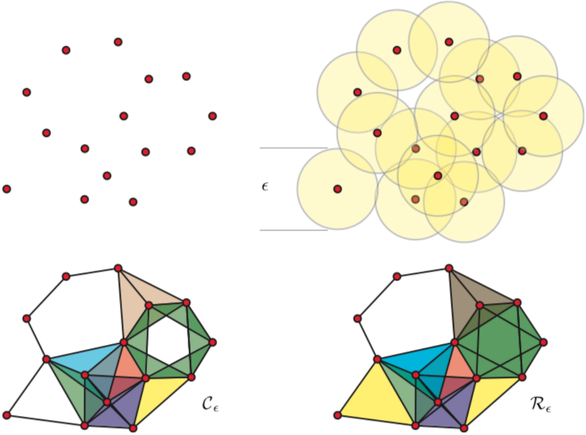


Figure 6:

a Barcode

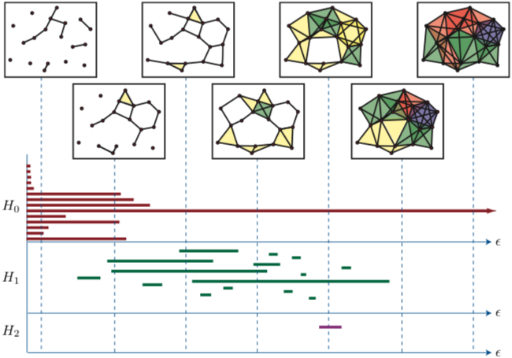


FIGURE 4. [bottom] An example of the barcodes for $H_*(\mathsf{R})$ in the example of Figure 3. [top] The rank of $H_k(\mathcal{R}_{\epsilon_i})$ equals the number of intervals in the barcode for $H_k(\mathsf{R})$ intersecting the (dashed) line $\epsilon = \epsilon_i$.

Figure 7:

More Filtrations...

a *superlevel-filtration* $S^{\downarrow}(f)$ to be the $\mathbb{R}^{op}$ indexed filtration for a topological space X where $f : X \to \mathbb{R}$ where

$$S^{\downarrow}(f)_r = \{p \in X | r \leq f(p)\}$$

example: let T be a Remannian manifold (say $\mathbb{R}^n$ or unit sphere), and $f : T \to \mathbb{R}$ a pdf then $S^{\downarrow}(f)$ tells us about the modes (basins of attraction under gradient flow) of the pdf f and more...

Nerves & Stability

Nerves

let Δ be the finite, non-empty subsets of a set S such that $\sigma \in \Delta$ and $\neq \tau \subset \sigma$, then $\tau \in \Delta$

given the collection of sets $U = \{U^\alpha\}_{\alpha \in S}$ indexed by S, the nerve of U is the simplicial complex

$$N(U) = \{\sigma \subset S | \cap_{\alpha \in \sigma} U^\alpha \neq\}$$

$N(U)$ has the properties

- 0 -simplex $\forall U^\alpha \in U$
- 1 -simplex $\forall \alpha, \beta \in S$ with $U^\alpha \cap U^\beta \neq$
- 2 -simplex $\forall \alpha, \beta, \gamma \in S$ with $U^\alpha \cap U^\beta \cap U^\gamma \neq$
- etc

The Nerve Theorem

Theorem (Nerve Theorem for Open Covers) if U is an open cover of a metrizable space X such that all intersections of finitely many elements in U are contractible, then $X \simeq N(U)$

proof: Hatcher, section 4.G via homotopy theory

proof: Edelsbrunner, Harer crediting Leray

proofs: Borsuk, Weil, and Leray in 1940s-1950s

More Nerves

- nerve theorem can be proven through homology
- nerves can be extended to filtrations

Persistent Nerves

weakly equivalent is defined: if $f : X \to Y$ continuous, then $f_* : \pi_0(X) \to \pi_0(Y)$ is bijective

note: this homotopic notion can be extended to categories, functors, etc.

Theorem (Persistent Nerve Theorem) is U is a cover of a filtration F where for each $r \in \mathbb{R}$, that U_r and F_r satisfy either

1. if U is an open cover of a metrizable space X such that all intersections of finitely many elements in U are contractible
2. is U is a finite, closed, convex cover of $X \subset \mathbb{R}^n$,

then F and $N(U)$ are weakly equivalent

The Persistent Homology Pipeline

1. Take your data and apply a filtration
   - (Union of balls or something more general)
2. Calculate homology of your filtration to get a persistence module
   - linear algebra for computers
3. Apply the structure theorem to get a barcode

if M is $\mathbb{R}$ or $\mathbb{Z}$ indexed pfd persistence module, then there exists a unique multiset of intervals B_M such that

$$M \cong \bigoplus_{I \in B_M} K^I$$

Extensions & Applications

Zig-Zag Persistence

Instead of a a monotone sequence

$$\ldots \to \bullet \to \bullet \to \bullet \to \ldots$$

Have a more general sequence

$$.. \to \bullet \leftarrow \bullet \to \bullet \leftarrow \ldots$$

so we can develop persistence over these modules

- Carlsson, de Silva, and Morozov (2009) "Zigzag Persistent Homology and Real-valued Functions"

Stability of Persistence Modules

Chazal, de Silva, Glisse, and Oudot (2012, 2013) “The structure and stability of persistence modules”

- persistence modules built using measure theory

Botnanand Lesnick (2017) “Algebraic Stability of Zigzag Persistence Modules”

- agebraic stability of the persistent homology of Reeb graphs, persistence modules, and interleavings

Algebra

Adcock, E. Carlsson, G. Carlsson (2013) “The Ring of Algebraic Functions on Persistence Bar Codes”

- the topology of the barcode (collection of intervals) is unusual requiring tools to understand
- identify an algebra of functions on the set of bar codes which is defined in a conceptually coherent way

Multiparameter Persistence

Lesnick (2015) “The Theory of the Interleaving Distance on Multidimensional Persistence Modules”

- the one parameter (thnk r-balls) persistence redily extends to multi-parameter

Miller (20198) “Real Multiparameter Persistent Homology” (Talk)

https://www.youtube.com/watch?v=tBqRbjIWPV0

- extends naturally the idea to the reals
- talks of presentations

Sheaves and Cohomology

Other Ideas

- computer vision
- probability distributions
- dynamical systems
- time series analysis
- behavior in complex systems
- topological data analysis (TDA)
- neuroscience
- Uniform Manifold Approximation and Projection (Leland McInnes, 2018)

Computer Vision & TDA

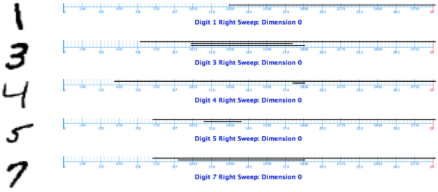


Figure 1: No Loop Digits with Betti 0 barcode, sweep to right

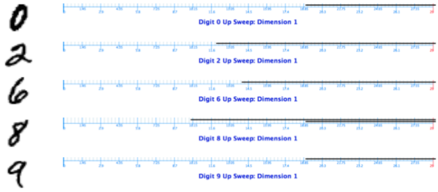


Figure 2: Loop Digits with Betti 1 barcode, sweep to top

Figure 8:

Conclusion

Conclusion

Resources

- Elementary Applied Topology (2014) Robert Ghrist
- Homological Algebra and Data (2010) Rober Ghrist
- THE BASIC THEORY OF PERSISTENT HOMOLOGY (2012) http://math.uchicago.edu/~may/REU2012/REUPapers/WangK.pdf
- Multiparameter Persistence Lecture notes (2019) Michael Lesnick https://www.albany.edu/~ML644186/AMAT_840_Spring_2019/Math840_Notes.pdf
- A User's Guide to Topological Data Analysis (2017) Elizabth Munch