Math 412 Topological Data Analysis 08/29/2023

"Basically this room is not actually controlled by Big Trinity"

"Ben-ditch"

Paul Bendich – call him "Paul"

Research prof in math

Runs Datat

Runs Geometry Data Analytics – gov't contracting company

TDA formed at Stanford & Duke

Mixture of Algebraic Topology & Data Science

Applied Math = adapting new math, not necessarily apply old math

Course = 2/3 whiteboard, 1/3 slides

Topology doesn't make sense in the real world: ◯ is more topologically similar to —— than ◯

How we combat this is via point sampling & persistent homology

Sidenote: linalg is important is because it is the last reasonably computable math area

With TDA you can do "Applied Math"

Ask prof ab slide w/ Ezra Miller after class

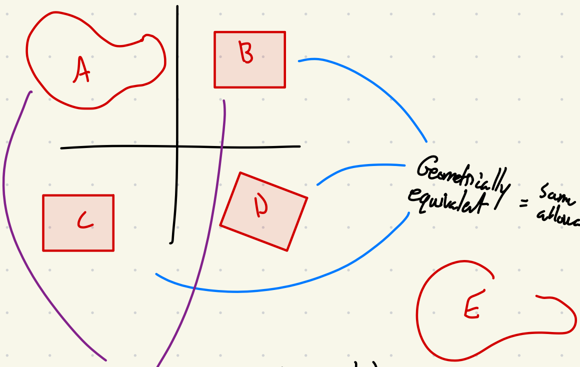
TDA and Deep Learning are becoming friends

- ↳ you can use TDA to improve loss functions – persistent homology helps w/ image segmentation
- ↳ Also helps w/ dimension reduction

Now let's put some rigor to this.

Topology vs. Geometry

"What do we mean by things being the same?"



B, C, and D are geometrically equivalent
A, B, C, and D are homeomorphic

Geometrically equivalent = Same modulo some allowable transformations

Proving positive "is", we can construct.
Proving "impossible," we consider all possibilities

Topologically equivalent (homeomorphic)
= two shapes can be turned into each other without ripping or glueing

How does undecidability work with AI?

Can we think about homeomorphisms algorithmically?

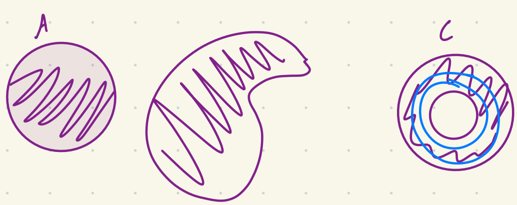
Some inputs cannot return outputs
"

As it turns out, Markov (father of the chain guy, ~1950s) found that the homeomorphy problem is undecidable

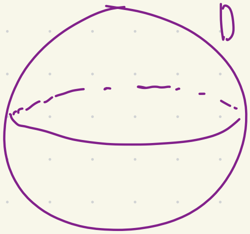
Homology (Noether, ~1920):

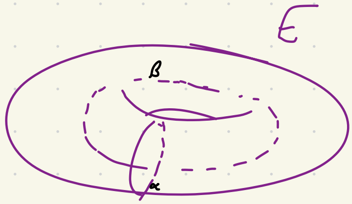
0th-degree homology of a shape
= Number of connected components in a shape
↳ is topologically invariant

1st degree homology
= Number of linearly independent non-bounding holes/loops in the shape



"Any time you can turn something geometric into something algebraic, you are doing deep, deep magic"

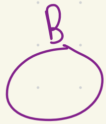




D & E are not homeomorphic but they both have the same number of voids — a notion of having an inside

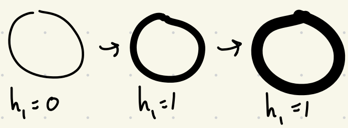
2 non-boundary loops : any loop is either
≅ 0, α, or β

Persistent Homology
(2000s) (1920s)

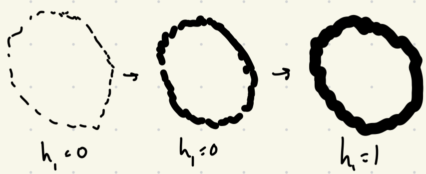


A & C homeomorphic
but this doesn't really make much sense because B & C are much more similar than A & C.

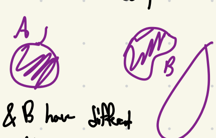
With persistent homology, we look at homology dynamically over time as we make shapes fuzzy



This works well with point clouds



Recall: Connected Components



A & B have different # of connected components
⇒ cannot be homeomorphic

How can we turn this from hand-wavey to algorithmic and computational?

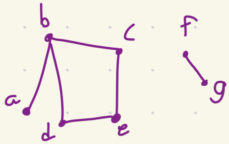
The answer: simplicial complexes/graphs

Def (simple, undirected) A graph ("network") G consists of a set of vertices V and a set of edges E which are unordered pairs of elements of V

$$G = (V, E)$$

Example:

This is a drawing of a graph not a graph itself.



$V = \{a, b, c, d, e, f, g\}$

$E = \{(a,b), (b,c), (c,e), \ldots\}$

Suppose $x, y \in V$ are two vertices in a graph G.

A path in G from x to y is a sequence of vertices

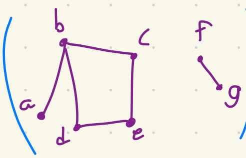
$$x = v_0, v_1, v_2, v_3, \ldots, v_n = y \quad \text{s.t.} \quad \forall i, \ \{v_i, v_{i+1}\} \in E$$

Def G is connected if $\forall x, y \in V$, $\exists$ a path in G from x to y

Def A connected component of G is a subgraph H of G which

1) is connected as its own graph
2) if we add one more vertex to H, it is not connected

Def $\beta_0(G) = \#$ of connected components in a graph G
(0th degree homology)

β_0 () $= 2$

Motivation for filtered graphs:

This point cloud could have 1, 2, or 3 "connected components"

$\beta_0 = 3$, 2 more persistent than
$\beta_0 = 7$

→ we can use the persistent homology thing from above (overlapping disks)

So start with $r = 0$ and watch what happens with the connected components

The length of the # of connected components persists determines significance

Let G be a graph. A filtration of G is a nested sequence of subgraphs of G.

$$\emptyset = G_0 \subseteq G_1 \subseteq G_2 \subseteq \dots \subseteq G_{m-1} \subseteq G_m = G$$

Linear Algebra Review +

Def A vector space is a set V along with two operations: addition and scaling, satisfying:

$\forall v, w \in V,\ v + w \in V$

$v + w = w + v$

$x \in V,\ v + (w + x) = (v + w) + x$

abelian group $\exists 0 \in V$ s.t. $v + 0 = V$.

$\forall v \in V,\ \exists -v \in V$ s.t. $v + (-v) = 0$

$\forall v \in V,\ \forall$ scalars $r,\ rv \in V$

$1v = v$

Examples:
- $(\mathbb{R}^2, +)$, $(\mathbb{R}^3, +)$, $(\mathbb{R}^n, +)$
- Span of functions

Def The Binary Scalar field $\mathbb{Z}/2\mathbb{Z}$ is the set $\{0, 1\}$ with the addition rule $1+1=0$.

$0 + 1 = 1$, $0 + 0 = 0$

Def Suppose V and W are 2 vector spaces. A linear transformation $T : V \to W$ is a function from V to W which satisfies:

$\forall v, v' \in V,\quad T(v + v') = T(v) + T(v')$

$\forall v \in V,\ \forall$ scalars $r,\quad T(rv) = rT(v)$

Category of Vector Space consists of objects called Vector spaces and morphisms which are linear transformations

---

Def Let V be a vector space. Let W be a subset of V. W is a subspace of V if:

- $\forall w, w' \in W,\ w + w' \in W$
- $\forall w \in W,\ \forall$ scalars $r,\ rw \in W$

Examples Space is $\mathbb{R}^3$, Subspaces are any plane or line through the origin, the origin itself, and $\mathbb{R}^3$ itself.

Let $T: V \to V'$ be a linear transformation.

Def: The kernel of T is $\ker(T) = \{v \in V \mid T(v) = 0\}$ ⇔ Nullspace of matrix describing T

Claim [HW]**:** $\ker(T)$ is a subspace of V

Def: The image of T is $\mathrm{Im}(T) = \{v' \in V' \mid \exists v \in V, v' = T(v)\} = \{T(v) \mid v \in V\}$

Claim: $\mathrm{Im}(T)$ is a subspace of V'.

Quotient Spaces

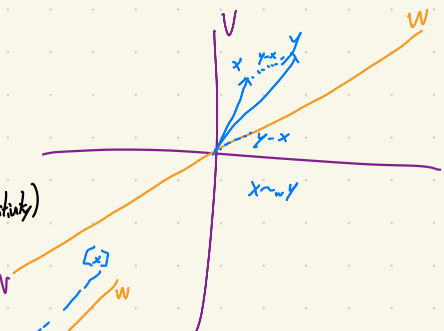
"It's a pretty fundamental concept but it's not that deep"

"x equals y mod W"
⇕

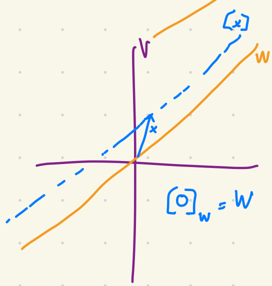
Let V be a vector space, W be a subspace. Suppose $x, y \in V$. We say $x \sim_W y$ if $x - y \in W$

Claim [HW]**:** (Criterea for an equivalence relation)

- $\forall x \in V$, $x \sim_W x$ (reflexivity)
- $\forall x, y \in V$, if $x \sim_W y$ then $y \sim_W x$ (symmetry)
- $\forall x, y, z \in V$, if $x \sim_W y$ and $y \sim_W z$ then $x \sim_W z$ (transitivity)



Def Suppose $x \in V$ and W is a subspace,
The equivalence class of x (mod W) is
$[x]_W = \{y \in V \mid y \sim_W x\}$



Claim [HW]**:** If $x, x' \in W$ then either
$[x]_W = [x']_W$ or $[x]_W \cap [x']_W = \emptyset$

Def: Let V/W be the set of equivalence classes mod W. We give it a vector space structure as follows:

if $[x], [y] \in V/W$, define $[x] + [y] := [x+y]$

also, if r is a scalar, $r[x] = [rx]$

Claim if $x \sim_W x'$ and $y \sim_W y'$ then $x+y \sim_W x'+y'$

Proof: Since $x \sim_W x'$, $\exists w \in W$ s.t. $x - x' = w$.
$y \sim_W y'$, $\exists z \in W$ s.t. $y - y' = z$.

Now we compute
$$(x+y) - (x'+y') = (x-x') + (y-y') = w + z \in W$$
because W is a subspace.

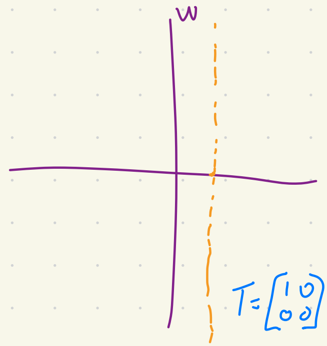
"It's an english paragraph"

So $x+y \sim_W x'+y'$ $\boxed{}$

Note: we should prove the same for scaling $(rx \sim_W rx')$

Claim [HW]: $\dim(V/W) = \dim(V) - \dim(W)$

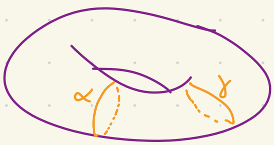
Claim [HW]: If $T: V \to V'$ is a linear transformation
Then $V/\ker T \cong \operatorname{Im} T$ (first isomorphism theorem)



$T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

$\operatorname{Im} = x\text{-axis}$
$\ker = y\text{-axis}$

Motivation:

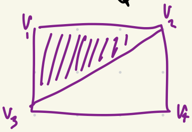


$\alpha \sim \gamma$

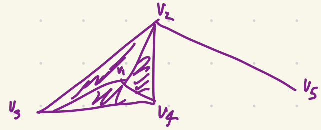
Def: Let $V = \{v_1, v_2, \ldots, v_n\}$ be a finite set called "vertices."
An (abstract) simplicial complex on V is a set K of subsets of V satisfying

- every element in V is in K
- if $\tau \in K$ and $\sigma \subseteq \tau$ then $\sigma \in K$.

drawing of an abstract simplicial complex



"If I have the triangle, I must have the three edges that make up the triangle"



Vocab: if $\sigma \in K$, σ is a simplex $(\dim(\phi) = -1)$
The dimension of σ or $\dim(\sigma) = |\sigma| - 1$
If $\dim(\sigma) = k$, σ is a k-simplex

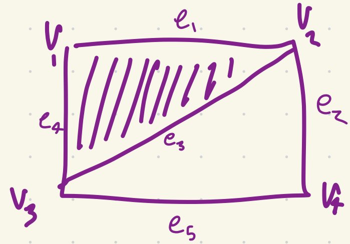
0-simplex = vertex
1-simplex = edge
2-simplex = triangle
3-simplex = tetrahedron

$\dim(K) = \max_{\sigma \in K} \dim(\sigma)$

If $\sigma \in K$ and $\tau \subseteq \sigma$ then τ is a face of σ and σ is a coface of τ.

If τ is a face of σ and $\dim(\tau) = \dim(\sigma) - 1$, then τ is a boundary face of σ.

<u>Def</u> Let K be an abstract simplicial complex.
For each nonnegative integer i, define
$C_i(K)$ to be the vector space (over the binary field)
with basis corresponding to the i-simplicies in K.
We call $C_i(K)$ the i-dimensional chain group of K

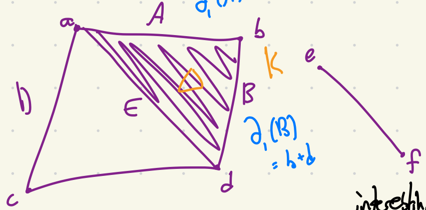


Let K be an abstract simplicial complex.

Recall for each $i = 0, 1, 2, \ldots,$
We define $C_i(K)$ to be the binary vector space
with basis the i-simplicies in K

"the ith chain group of K"

$\partial_1(A) = a+b$



$\partial_1(B) = b+d$

$\partial_2(\Delta) = A+B+E$

interestingly enough,
$\partial_1(\partial_2(\Delta)) = 0$
Which generalizes

$$\cdots\ C_3(K) \xrightarrow{\partial_3} C_2(K) \xrightarrow{\partial_2} C_1(K) \xrightarrow{\partial_1} C_0(K) \longrightarrow 0$$

Define $Z_i(K) := \ker \partial_i$, the group of i-cycles of K.

<u>Boundary maps</u>

"Boundary 1"

Define $\partial_1 : C_1(K) \to C_0(K)$. Suppose α is an edge with verticies x and y.
Then define $\partial_1(\alpha) := x+y$.
Now, extend by linearity. e.x. $\partial_1(A+B) = \partial_1(A) + \partial_1(B)$
$= (a+b) + (b+d)$
$= a+d$

What does the kernel of ∂_1 look like?

$$\partial_1(A+B+E) = \partial_1(A) + \partial_1(B) + \partial_1(E)$$
$$= (a+b) + (b+d) + (d+a)$$
$$= 0$$

$Z_1(K)$:

0

$\alpha = A+B+E$
$\gamma = D+E+C$
$\delta = A+B+C+D = \alpha + \gamma$

but even though we have 4 loops, we only have 1 hole...

So far, we've gotten: $C_2(K) \xrightarrow{\partial_2} C_1(K)$. Let's look at $\partial_2 : C_2(K) \to C_1(K)$.

Given a triangle Δ with edges E_1, E_2, E_3
Define $\partial_2(\Delta) = E_1 + E_2 + E_3$.
Extend by linearity.

Def For each i, the image of $\partial_{i+1} : C_{i+1} \to C_i$ is denoted $B_i(K)$ and is the group of <u>i-boundaries in K</u>

In the above case, $B_1(K) = \{0, \alpha = A+B+E\} \subset Z_1(K)$

Now we can do fun quotient stuff with our lil subspace $B_1(K)$

Claim [HW]: $\forall i$, $\partial_i \circ \partial_{i+1} = 0$.

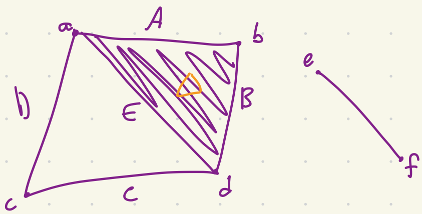
$$\cdots \to C_{i+1} \xrightarrow{\partial_{i+1}} C_i \xrightarrow{\partial_i} C_{i-1} \to \cdots$$

going down twice gives 0 every time.

As a consequence, $\forall i$ B_i is a subspace of Z_i.
image of ∂_{i+1} — Kernel of ∂_i

Def (the big one)
The i-th <u>homology group</u> of K is $H_i(K) := Z_i(K) / B_i(K) = \ker \partial_i / \operatorname{Im} \partial_{i+1}$
it's dimension is $\beta_i(K)$, the ith Betti number of K.

Recall: K



$\dim H_1 = 1$
basis: $\{[\gamma]\}$
$\beta_1(K) = 1$

In H_1, consider $[\alpha] = [0]$ (though in Z_1, $\alpha \neq 0$)
we also have $[\gamma] = [\delta]$ because $\gamma + \delta = \alpha \in B_1$

homology class of γ

note $\partial_0 : C_0 \to 0$ is the zero transformation

$$C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0$$

By definition, $Z_0 = C_1$ and $B_0 = \operatorname{Im} \partial_1$.
Then $H_0 = Z_0 / B_0$.

Elements of C_0 are just combinations of vertices.

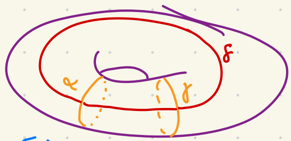
$a + b = \partial_1(A)$ so $a + b \in B_0$
hence $[a] = [b] \in H_0$

$b + c = \partial_1(B + C)$ so $[a] = [b] = [c] \in H_0$

note $a + e \notin B_0$ so $[a] \neq [e] \notin H_0$

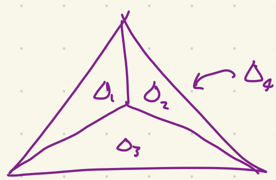
Note $\beta_0(K) = \dim H_0(K) = 2$ with one possible choice of basis being $\{[a], [e]\}$

Return to shapes:

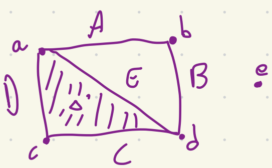


$[\alpha] = [\gamma]$ but $[\delta] \neq [\alpha]$

$$C_3 \xrightarrow{\partial_3} C_2 \xrightarrow{\partial_2} C_1$$



$\partial_2(\Delta_1 + \Delta_2 + \Delta_3 + \Delta_4) = 0$



$H_1(K) = \operatorname{Ker} \partial_1 / \operatorname{Im} \partial_2$

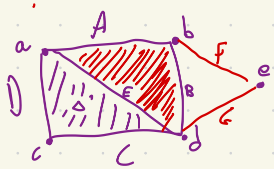
Kernel of ∂_1: $\alpha = A + B + G$
$\gamma = C + D + E$
$\delta = A + B + C + D$
0

$Z_1(K)$ has dimension 2

$$H_1(K) = Z_1(K) / B_1(K)$$

$B_1(K) = \{\partial_2(\Delta)\}$
$= \{C + D + E = \gamma\}$

Since $\alpha + \delta = \gamma \in B_1(K)$,
$[\alpha] = [\delta]$



Consider $\cdots \to C_3(K) \xrightarrow{\partial_3} C_2(K) \xrightarrow{\partial_2} C_1(K) \xrightarrow{\partial_1} C_0(K) \xrightarrow{\partial_0} 0$

In general, define $\partial_i : C_i(K) \to C_{i-1}(K)$. Let $\sigma = \langle v_0 v_1 \dots v_i \rangle$ be a typical i-simplex in K.

We define $\partial_i(\sigma) := \sum_{j=0}^{i} \langle v_0 \cdots v_{j-1} \hat{v}_j v_{j+1} \cdots v_i \rangle$ remove this

If $i=2$ $\partial(\langle v_0 v_1 v_2 \rangle) = \langle \hat{v}_0 v_1 v_2 \rangle + \langle v_0 \hat{v}_1 v_2 \rangle + \langle v_0 v_1 \hat{v}_2 \rangle$

: edge b/w v_1 & v_2

Claim [HW]: $\forall i$, $\partial_i \circ \partial_{i+1} = 0$.

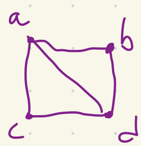
That is, for any $(i+1)$-simplex τ,

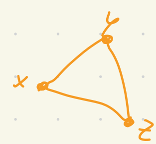
$\partial_i(\partial_{i+1}(\tau)) = 0$.

$C_{i+1}(K) \xrightarrow{\partial_{i+1}} C_i(K) \xrightarrow{\partial_i} C_{i-1}(K)$

Corollary $\forall i$, $B_i(K) \subseteq Z_i(K)$

Simplicial maps:





Suppose K and L are 2 simplicial complexes.
Suppose $f: \mathrm{Vert}(K) \to \mathrm{Vert}(L)$ is a set map.
We say that f is a simplicial map if whenever $\{v_0, v_1, \dots, v_i\}$ form a simplex in K, $\{f(v_0), f(v_1), \dots, f(v_i)\}$ form a simplex in L.

$\{a, b, c, d\} \to \{x, y, z\}$

$f(a) = y$ $f(c) = x$ $f(d) = z$
$f(b) = y$

Induced Maps:

Suppose $f: K \to L$ is a simplicial map.
For each i, f induces a linear transformation $(f_\#)_i : C_i(K) \to C_i(L)$
via if σ is an i-simplex in K,

$$f_\#(\sigma) = \begin{cases} f(\sigma) & \text{if } \dim(f(\sigma)) = \dim(\sigma) \\ 0 & \text{if } \dim(f(\sigma)) < \dim(\sigma) \end{cases}$$

Ex from above $f_\#(\langle a\, b \rangle) = \langle x y \rangle$
$f(\langle abd \rangle) = \langle yz \rangle$

Suppose $f: K \to L$ is a simplicial map.

$$\tau \xrightarrow{\partial_{i+1}} \sigma \qquad \partial^K_i$$

$$\begin{tikzcd}
\cdots \arrow[r] & C_{i+1}(K) \arrow[r] \arrow[d, "(f_\#)_{i+1}"] & C_i(K) \arrow[r] \arrow[d, "(f_\#)_i"] & C_{i-1}(K) \arrow[r] \arrow[d, "(f_\#)_{i-1}"] & \cdots \\
\cdots \arrow[r] & C_{i+1}(L) \arrow[r, "\partial_{i+1}"] & C_i(L) \arrow[r, "\partial^L_i"] & C_{i-1}(L) \arrow[r] & \cdots
\end{tikzcd}$$

$$(f_\#)_{i+1} \circ \partial^K_i \qquad \partial^L_i \circ (f_\#)_i$$

$$f_\#(\tau) \to f_\#(\sigma)$$

[HW]

Lemma

$\forall i$, diagram commutes

Corollary: $f_\# (Z_i(K)) \subseteq Z_i(L)$

Corollary: $f_\#(B_i(K)) \subseteq B_i(L)$

As a consequence, a simplicial map $f: K \to L$ induces linear transformations on homology groups

$$(f_*)_i : H_i(K) \to H_i(L) \quad \forall i$$

$$H_i(K) = Z_i(K)/B_i(K), \qquad H_i(L) = Z_i(L)/B_i(L)$$

defined by

$$f_*([\alpha]) = [f_\#(\alpha)]$$

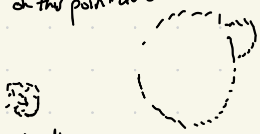
Linear Algebra

Suppose V and V' are vector spaces and $T: V \to V'$ is a linear transformation.
Suppose W is a subspace of V and W' is a subspace of V'.
And suppose $T(W) \subseteq W'$.

Then T induces a well-defined linear transformation $T_*: V/W \to V'/W'$ defined by $T_*([v]_W) = [T(v)]_{W'}$

Persistent Homology

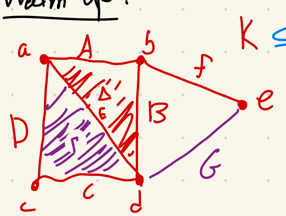
How can we do topology on this point cloud?



Literally, $\beta_0 = 49$, $\beta_1 = 0$
but usually $\beta_0 = 2$, $\beta_1 = 2$

To fix this, let each point be a disk of radius ε; increase ε and watch the data disks overlap.

Warm up:



$K \subseteq L$

Consider the inclusion $i: K \to L$ of K into L $[a] \mapsto [a]$

Consider $i_*: H_0(K) \to H_0(L)$

$i_*: H_1(K) \to H_1(L)$

$[A+B+C+D] \to [A+f+G+C+D]$

$\mapsto 0$

$[A+B+C+D] \neq [0] \in H_1(K)$

but

$i_*([A+B+C+D]) = [0] \in H_1(L)$

Simplest case of Persistent Homology:

Def. A simplex stream is a sequence of simplicies

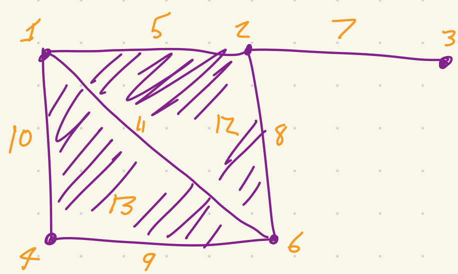
$$\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_N$$

Satisfying if σ_i is a face of σ_j then $i \leq j$.

Def. For each i, define $K^i = \{\sigma_1, \dots, \sigma_i\}$

Note $\forall i$, K^i is a simplicial complex.
Also note if $i < j$, then K^i includes into K^j

Example:



Consider the sequence of inclusions:

$$\emptyset = K^0 \subseteq K^1 \subseteq K^2 \subseteq \dots \subseteq K^{N-1} \subseteq K^N$$

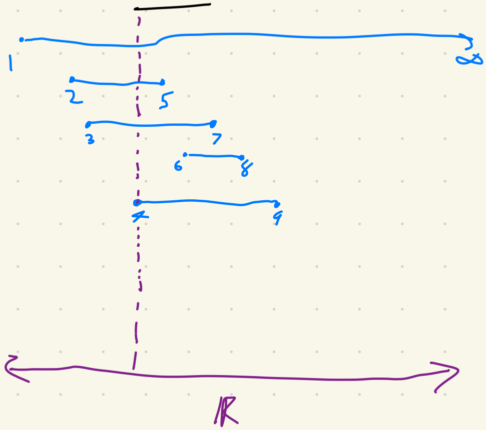
This induces a sequence of linear transformations:

fix p.

$$0 \to H_p(K^1) \to H_p(K^2) \to \dots \to H_p(K^{N-1}) \to H_p(K^N)$$

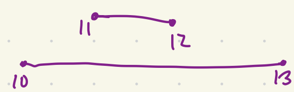
Barcode:

0-dim



$\beta_0 = 4$ @ $t = 4$

1-dim

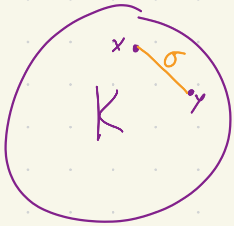


Let's look at how to compute persistent homology: The Persistence Algorithm

Suppose I have a simplicial complex K and I add to it a single simplex σ of dimension p. Call the new simplex ρ.

Then either β_{p-1} goes down by 1 XOR β_p goes up by 1

"σ is a negative p-simplex"



"σ is a positive p-simplex"

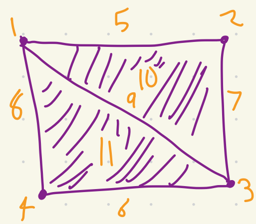
Now for the algorithm: (Edelsbrunner, Letscher, Zomorodian 1999
Zomorodian + Carlsson 2002)

Suppose K has N simplices which stream in as $\sigma_1, \sigma_2, \dots, \sigma_N$
We create an $N \times N$ binary matrix D where $D_{ij} = 1$ only if σ_i is a boundary face of σ_j

$low(5)=2$ $\quad low(8)=4$ $\quad low(11)=8$

$$
\begin{array}{c|ccccccc}
 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\
\hline
1 & 1 & & & 1 & 1 & & \\
2 & 1 & & 1 & & & & \\
3 & & 1 & 1 & & 1 & & \\
4 & & 1 & & 1 & & & \\
5 & & & & & & 1 & \\
6 & & & & & & & 1 \\
7 & & & & & & 1 & \\
8 & & & & & & & 1 \\
9 & & & & & & 1 & 1
\end{array}
$$

not reduced



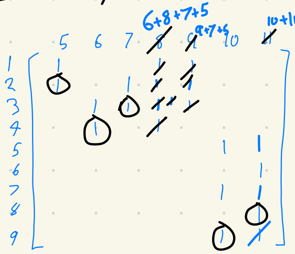
Leave off vertex columns and triangle rows.

Def Let M be any $N \times N$ binary matrix.
Define $low_m : \{1, 2, \dots, N\} \to \{0, 1, 2, \dots, N\}$ by

$$
low_m(j) = \begin{cases} 0 & \text{if column of } m \text{ empty} \\ i & \text{if lowest (downmost) 1 in column } j \text{ of } m \text{ lies in row } i \end{cases}
$$

Def M is reduced if the lowest 1 in each column is in a unique row. That is, if $low_M(j) = low_M(j') \neq 0$ then $j = j'$.

Algorithm: Loosely, we reduce D via left-to-right column additions

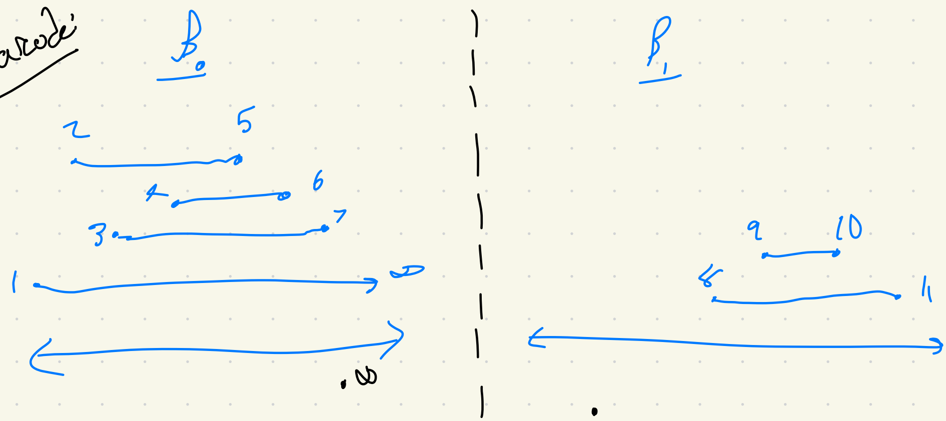


β_0	β_1
$(2, 5)$	$(9, 10)$
$(4, 6)$	$(10, 11)$
$(3, 7)$	
$(1, \infty)$	

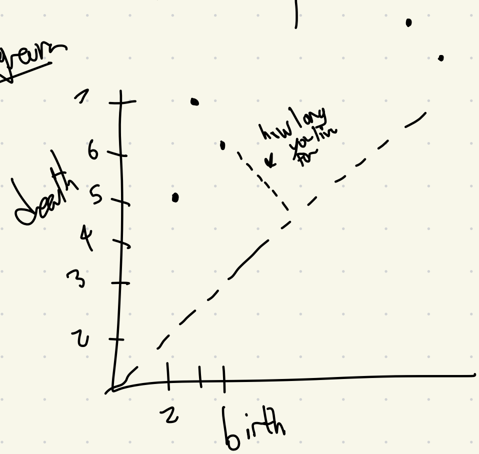
Note: this algorithm has worst case complexity $O(N^3)$

Visual representation of algorithm output

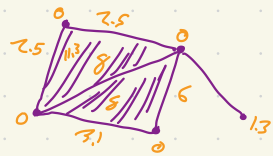
Barcode:



Persistence diagram



Suppose K is a simplicial complex and suppose $f: K \to \mathbb{R}$ is a function



Def. f is a filter function if whenever σ is a face of τ, we have

$$f(\sigma) \leq f(\tau).$$

Def Let $r \in \mathbb{R}$. Define

sublevel set
threshold set

$$K_r = \{\sigma \in K \mid f(\sigma) \leq r\}$$

Claim $\forall r$ K_r is a subcomplex of K

Note: $\forall r < s$, $K_r \subseteq K_s$

Now suppose K has N simplicies and suppose the distinct values taken by f are $r_1 \leq r_2 \leq \cdots \leq r_m$.
$\forall i$, define $K_i := K_{r_i}$ (note $m \leq N$)

Consider the complex K:

$$\emptyset \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_m = K$$

for any p, we get:

$$0 \to H_p(K_1) \to H_p(K_2) \to \cdots \to H_p(K_m) = H_p(K)$$

For any $i<j$, let $f_p^{i,j}: H_p(K_i) \to H_p(K_j)$ be the map induced by including K_i into K_j

Def suppose $[\alpha] \in H_p(K_i)$. $[\alpha]$ is born at r_i if $[\alpha] \notin \mathrm{Im}\, f_p^{i-1,i}$

Def suppose $[\alpha] \in H_p(K_i)$ is born at r_i.
We say $[\alpha]$ dies at r_j if

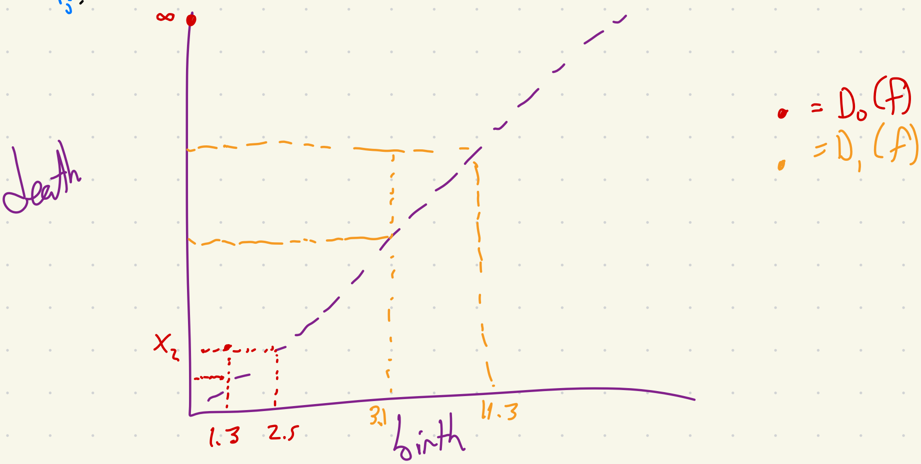
- $f_p^{i,j}[\alpha] \in \mathrm{Im}\, f_p^{i-1,j}$
- $f_p^{i,j-1}[\alpha] \notin \mathrm{Im}\, f_p^{i-1,j-1}$

if $i=j-1$ this is vacuously true.

$$H_p(K_{i-1}) \to H_p(K_i) \to \cdots \to H_p(K_{j-1}) \to H_p(K_j)$$

$$[\alpha] \longmapsto f_p^{i,j}([\alpha])$$

Def: The p-dimensional persistence diagram of f, $\mathrm{Dgm}_p(f)$ $[D_p(f)]$ contains a dot (r_i, r_j) for each p-dimensional homology class born at r_i and dead at r_j.



Def Let A be any set. A metric/distance ρ on A is a function $\rho: A \times A \to \mathbb{R}$ satisfying:

- $\forall a, b \in A,\ \rho(a,b) \geq 0$ (non-negativity)
- $\forall a, b \in A,\ \rho(a,b) = 0 \Leftrightarrow a = b$ (discernability)
- $\forall a, b \in A\ \ \rho(a,b) = \rho(b,a)$ (symmetry)
- $\forall a, b, c \in A,\ \rho(a,c) \leq \rho(a,b) + \rho(b,c)$ (triangle inequality)

(A, ρ) is a metric space

Examples: consider $A = \mathbb{R}^n$

1) $\rho(\vec{x}, \vec{y}) = \left[\sum_{i=1}^{n} (x_i - y_i)^2 \right]^{1/2}$ (l_2, Euclidean)

2) $\rho(\vec{x},\vec{y}) = \sum_{i=1}^{n} |x_i - y_i|$ (ℓ_1, "taxicab metric")

3) $\rho(\vec{x},\vec{y}) = \left[\sum_{i=1}^{n} |x_i - y_i|^p\right]^{1/p}$ for any $p>0$ (ℓ_p)

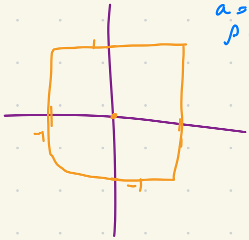
4) $\rho(\vec{x},\vec{y}) = \max_i |x_i - y_i|$ (ℓ_∞)

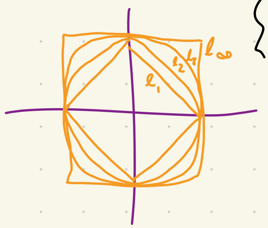
5) $\rho(\vec{x},\vec{y}) = \# i$ s.t. $x_i \neq y_i$ (ℓ_0)

<u>Def</u> Let (A,ρ) be a metric space. Fix $a \in A$ and a radius $R>0$. The closed ball of radius R centered at a is

$$B_\rho(a,R) := \{x \in A \mid \rho(a,x) \le R\}$$

$a = \vec{0}$, $R=1$
$\rho = \ell_\infty$

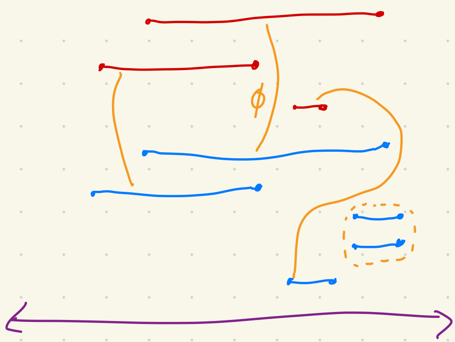


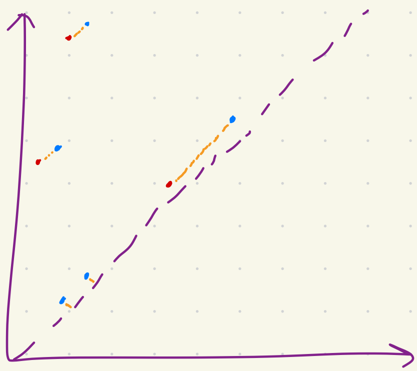


$\left\{\begin{matrix}\text{curse}\\ \text{blessing}\end{matrix}\right\}$ of dimensionality

↳ point distribution of ℓ_p in $\mathbb{R}^n$

Let's look at metrics applied to persistence diagrams!





Suppose D and D' are two different diagrams. How far apart are they?

Let D, D' be two barcodes. A matching between D and D' is a bijection ϕ between a subset E of D and $\phi(E)$.

Given ϕ, fixing $p>0$, the cost of ϕ is $[C_p(\phi)]^p = \sum_{\substack{\text{bars not} \\ \text{in } E}} [\text{length}(\text{bars})]^p + \sum_{\substack{\text{bars not} \\ \text{in } \phi(E)}} [\text{length}(\text{bars})]^p$
$+ \sum_{(a,b)\in E} |a-\phi(a)|^p + |b-\phi(b)|^p$

Finally, the pth Wasserstein distance between D and D' is

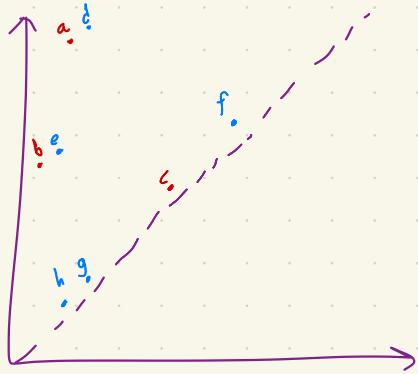
$$W_p(D,D') := \inf_{\substack{\text{matches} \\ \phi: D\to D'}} C_p(\phi)$$

How do we choose p? Well... basically it's arbitrary and up to us.

Claim [HW]: for any $p>0$, W_p is a metric on the space of persistence diagrams

Hint: composing bijections gives you a bijection... but it may not be the best one.

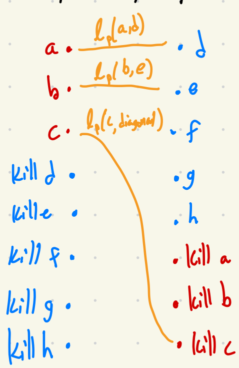
$$W_p(D,D') = \inf_{\phi: D\to D'} C_p(\phi)$$



Turns out, this is reasonably efficiently computable

Create bipartite graph with edge weights the cost of moving a point to where we match it

$\exists$ some algorithm to find best bijection in $O(n^3)$



Digression: doing statistics in a metric space.

Suppose (A,ρ) is a metric space.
Suppose we have a finite set $X=\{x_1,\dots,x_n\}\subseteq A$.
What's the mean of X?

Let's define $\tilde{F}_X : A \to \mathbb{R}$ via $\tilde{F}_X(a) = \sum_{i=1}^{n} [\rho(x_i, a)]^2$

↳ The fréchet function of X.

Def: The fréchet mean of X is

$$\mu_X := \underset{a \in A}{\operatorname{argmin}} \tilde{F}_X(a).$$

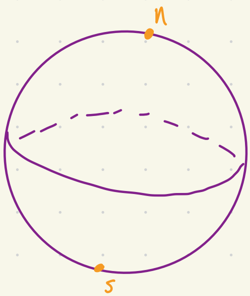
argument which achieves minimum

$$\min_{x \in [0, 2\pi]} (\sin(x)) = -1$$

$$\underset{x \in [0, 2\pi]}{\operatorname{argmin}} (\sin(x)) = 3\pi/2$$

$$\sin(3\pi/2) = -1$$

Claim: Consider the metric space $(\mathbb{R}, \ell_2)$. The fréchet mean of a finite set is the mean you know (and love).



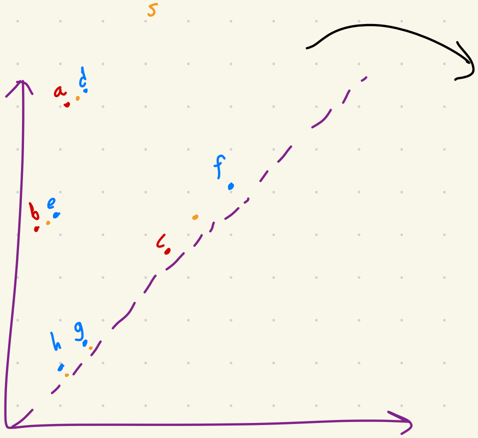
$X = \{n, s\}$

$\mu_X = \{\text{equator}\}$

yikes...

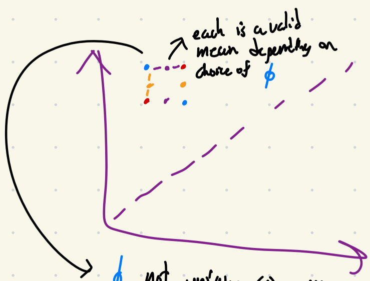
$\mathbb{R}^2 - \{0\}$, $X = \{-5, 5\}$

μ_X DNE



μ_X is a diagram!

• is that μ_X, as it turns out



ϕ not unique so mean not unique

$\Rightarrow$ Space of persistence diagrams basically has infinite curvature :)

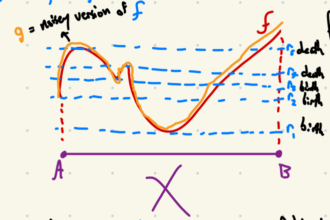
Stability of persistent homology

Prototypical example: changing a set of numbers a little bit changes the mean only a little bit

$(4, 4.5, 5.5, 6) \mapsto 5$

$(4.01, 4.501, 5.501, 6.01) \mapsto \sim 5.01$

Suppose X some topological space.
Suppose $f: X \to \mathbb{R}$ a function.



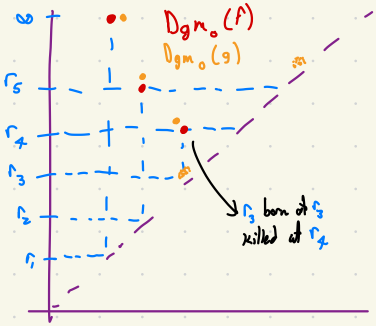
For each threshold value $r \in \mathbb{R}$, consider

$X_r := f^{-1}((-\infty, r]) = \{x \in X \mid f(x) \leq r\}$

note if $r < s$, then $X_r \subseteq X_s$.

$H_p(X_r) \to H_p(X_s)$.

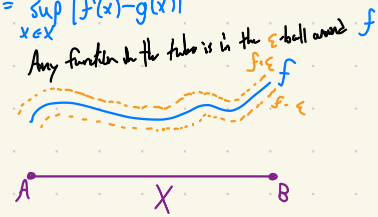
Let's look at the 0-dimensional homology of X under the filtration induced by X_r.



Suppose f and g are two real-valued function on X.
The l_∞-distance between f and g is

$l_\infty(f, g) := \sup_{x \in X} |f(x) - g(x)|$

$\|f - g\|_\infty$



Theorem (Stability theorem): [Haven, Edelsbrunner, Cohen-Steiner 2007]

If f and g are [reasonable] real-valued functions on X

Then $\forall p$, $W_\infty(Dgm_p(f), Dgm_p(g)) \leq l_\infty(f, g)$.

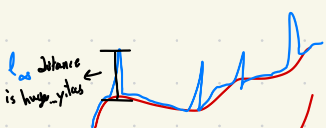
bottleneck distance

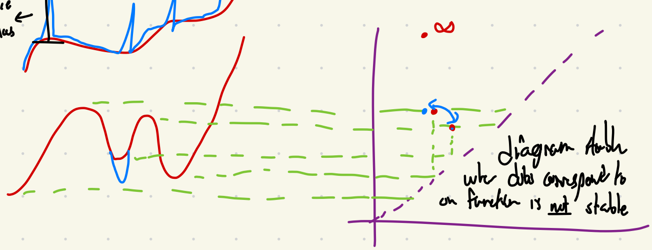
$\inf(\sup(\underline{why}))$ "min-max"

Nice thing: small change in RHS leads to small change in RHS

$f \longmapsto Dgm_p(f)$ respects the metric structure

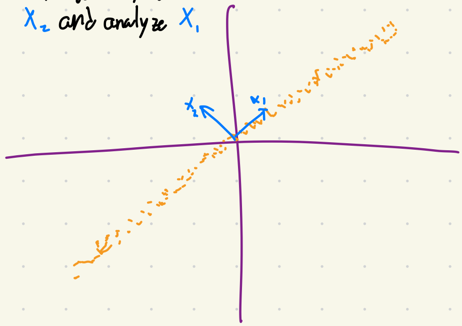
Not so nice things: outliers



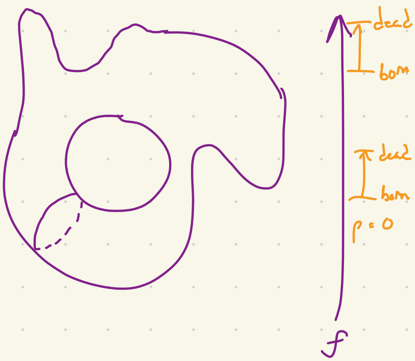


Principle Components Analysis (PCA)

Change basis, get rid of x_2 and analyze x_1



Recall if X is a topological space and $f: X \to \mathbb{R}$ is a function. We have $\mathrm{Dgm}_p(f)$ for each homological dimension p.



Question: how do we compute this?

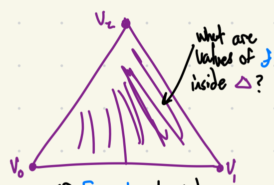
Procedure [Lower-star filtration]:

- Triangulate X with a simplicial complex K.
- To each vertex of K, just associate it's f-value. $\hat{f}(v) = f(v)$.
- To each higher order simplex σ in K, set $\hat{f}(\sigma) = \max_{v \in \sigma} f(v)$.
- Now consider the filtration of K via sublevel sets of $\hat{f}$.
- Compute, for each p, $\mathrm{Dgm}_p(\hat{f})$.

Claim [HW]: $\hat{f}$ is a filter function on K.

Questions:

1) What is relation between $\mathrm{Dgm}_p(\hat{f})$ and $\mathrm{Dgm}_p(f)$?
2) How efficient is this algorithm?



We know (from stability theorem) that $W_\infty(\mathrm{Dgm}_p(f), \mathrm{Dgm}_p(\hat{f})) \leq \|f - \hat{f}\|_\infty$

If f is L-lipschitz, then our bound is nicely controllable

To bound RHS, assume f is L-Lipschitz, assume you have many many simplices in K and that they are super small.

"many many cubed is very many many"

Point Cloud Triangulations:

Let $X = \{x_1, x_2, \ldots, x_N\} \subseteq \mathbb{R}^D$ (or really any metric space).
be a "point cloud".

Define $d_X : \mathbb{R}^D \to \mathbb{R}$ via $d_X(y) := \min_{x \in X} \|y - x\|$.

Abuse (of notation): for any ε, let $X_\varepsilon := \{y \in \mathbb{R}^D \mid d_X(y) \leq \varepsilon\}$.
Note for any ε, $X_\varepsilon = \bigcup_{x \in X} B_\varepsilon(x)$ (closed ε-balls).
And for $\varepsilon < \varepsilon'$, $X_\varepsilon \subseteq X_{\varepsilon'}$.
For every p, we get $\mathrm{Dgm}_p(X) := \mathrm{Dgm}_p(d_X)$

Question: how do we compute this?

Čech Complex: how can we feed the persistence monster

For any $\varepsilon > 0$, define $\check{C}_\varepsilon(X)$ to be a simplicial complex with X as its vertices and where $\sigma \in \check{C}_\varepsilon(X)$ if

$$\bigcap_{x \in \sigma} B_\varepsilon(x) \neq \emptyset$$

Note:

- For any ε, $\check{C}_\varepsilon(X)$ is indeed a simplicial complex because if τ is a face of σ, then

$$\bigcap_{x \in \tau} B_\varepsilon(x) \supseteq \bigcap_{x \in \sigma} B_\varepsilon(x).$$

- For any $\varepsilon < \varepsilon'$, $\check{C}_\varepsilon(X)$ is a subcomplex of $\check{C}_{\varepsilon'}(X)$

Thus we can compute persistence diagrams $\mathrm{Dgm}_p(\check{C}(X))$ associated with this filtration.

Nice fact: $\mathrm{Dgm}_p(\check{C}(X)) = \mathrm{Dgm}_p(X)$ $\forall p$. Magic ingredient: Nerve Lemma.

Not fun facts:

- $\check{C}(X)$ has a lot of simplices!
  In fact, it has 2^N simplices.

  But if we want $Dgm_p(X)$,
  we only need consider $(p+1)$-skeleton
  of $\check{C}(X)$. this is size $\binom{N}{p+1} = O(N^{p+1})$

  or we can just bound ε.

- figuring out when a simplex appears in $\check{C}(X)$ ... is not that fun. $O(N^{\lfloor D/2 \rfloor})$

  complexity dependent on ambient space is bad.

Housekeeping: no class thursday!

Recall: Suppose $X = \{x_1, \dots, x_N\} \subseteq \mathbb{R}^D$ is a point cloud.
We are interested in $Dgm_p(X)$ associated to the filtration of $\mathbb{R}^D$ given by the growing balls around points in X.

Last time, we defined the Čech Complex $\check{C}(X)$.
Where for any $\alpha \in \mathbb{R}$, $\sigma \in \check{C}_\alpha(X)$ if

$$\bigcap_{x \in \sigma} B_\alpha(x) \neq \phi.$$

Victoris-Rips Complex: again given $X = \{x_1, \dots, x_N\} \subseteq \mathbb{R}^D$ (or any metric space)
For any $\alpha \in \mathbb{R}$, define $R_\alpha(X)$ to be the simplicial complex with vertices X
and $\sigma \in R_\alpha(X)$ iff $\forall x_i, x_j \in \sigma$, $B_\alpha(x_i) \cap B_\alpha(x_j) \neq \phi$.

Problem: $R_\alpha(X)$ doesn't give same Betti #'s as X_α

However: $\forall \alpha$, $\check{C}_\alpha(X) \subseteq R_\alpha(X) \subseteq \check{C}_{2\alpha}(X)$ [HW]

Because $\forall i, j$

$$\bigcap_{x \in \sigma} B_\alpha(x) \subseteq B_\alpha(x_i) \cap B_\alpha(x_j)$$

In fact in $\mathbb{R}$, $R_\alpha(X) \subseteq C_{\alpha\sqrt{\frac{2D}{D+1}}}(X)$ [Ghrist, deSilva, Zob]

Rips alternate definition: $R_\alpha(X)$ is the maximal simplicial complex which shares the same 1-skeleton of $\check{C}_\alpha(X)$.

2 negatives, 2 positives:
- Rips is technically wrong
- Rips has a ton of simplices

But
- Easier to check stuff (can just check edges)

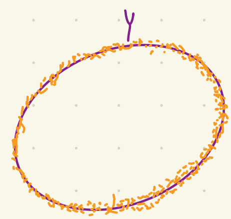
interleaving → gives nice stuff about birth and death

We care about the homology of top row

$$\begin{array}{ccccccccc} & & \check{C}_\alpha(X) & & \check{C}_{2\alpha}(X) & & \check{C}_{4\alpha}(X) & & \\ & \rotatebox{45}{$\subseteq$} & & \rotatebox{-45}{$\subseteq$} & & \rotatebox{-45}{$\subseteq$} & & \rotatebox{-45}{$\subseteq$} & \cdots \\ R_{\alpha/2}(X) & & & & R_\alpha(X) & & R_{2\alpha}(X) & & \end{array}$$

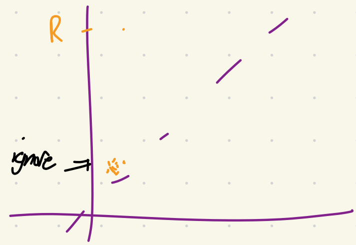
but we can much more easily compute bottom row

Homology inference:



Suppose Y is a topological space embedded in $\mathbb{R}^D$.
Suppose $X=\{X_1,\dots,X_N\}$ is a point sampled from/near Y.

Problem: Can we use X to infer homological things about Y?



Def: Suppose Y is a topological space embedded in $\mathbb{R}^D$. Consider $d_Y:\mathbb{R}^2\to\mathbb{R}$ defined by $d_Y(z):=\inf_{y\in Y}\|z-y\|$.
for any $\alpha\in\mathbb{R}$, define $Y_\alpha:=\{z\in\mathbb{R}^D \mid d_Y(z)\le\alpha\}$

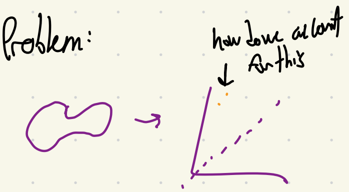
Def: If X and Y are 2 topological spaces embedded in $\mathbb{R}^D$.
The Hausdorff distance between X and Y is

$$d_H(Y,X):=\inf\{\varepsilon \mid Y\subseteq X_\varepsilon \text{ and } X\subseteq Y_\varepsilon\}$$

"is X a good point sample of Y?" controls density & bans outliers

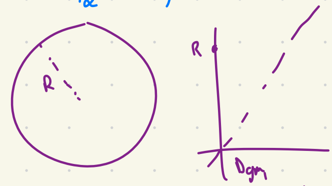
$Y\subseteq X_\varepsilon$
$X\not\subseteq Y_\varepsilon$

Problem:

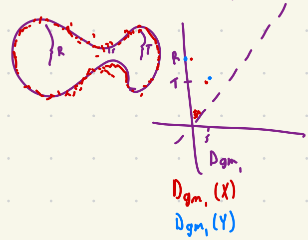


If Y is a space and X is a point-cloud, and $d_H(X,Y)\le\varepsilon$
Then we say "X is an ε-sample of Y"

Note for any space Y embedded in $\mathbb{R}^D$ we can filter $\mathbb{R}^D$ via sublevel sets Y_α of d_Y.
So for any p, we get persistence diagrams $\mathrm{Dgm}_p(Y) := \mathrm{Dgm}_p(d_Y)$.



Q: How do I read $\beta_p(Y)$ from $\mathrm{Dgm}_p(Y)$?
A: # of dots on y-axis.



Claim:[HW] If Y and X are shapes embedded in $\mathbb{R}^D$, then
$$d_H(Y,X) = \|d_Y - d_X\| = \sup_{z\in\mathbb{R}^D} |d_Y(z) - d_X(z)|$$
therefore for all p, $W_\infty(\mathrm{Dgm}_p(Y), \mathrm{Dgm}_p(X)) \leq d_H(Y,X)$ (by stability theorem).

(recall)
Problem: Given a topological space Y and a point cloud X sampled from/near Y, can we infer homology of Y from X

Two major players: first, "how good is the sample?"

Recall: If X and Y are 2 topological spaces embedded in $\mathbb{R}^D$.
The Hausdorff distance between X and Y is
$$d_H(Y,X) := \inf\{\varepsilon \mid \underset{\uparrow}{Y \subseteq X_\varepsilon} \text{ and } \underset{\uparrow}{X \subseteq Y_\varepsilon}\}$$
"is X a good point sample of Y?" controls density & bounds outliers

HW: $d_H(X,Y) = \|d_X - d_Y\|_\infty$

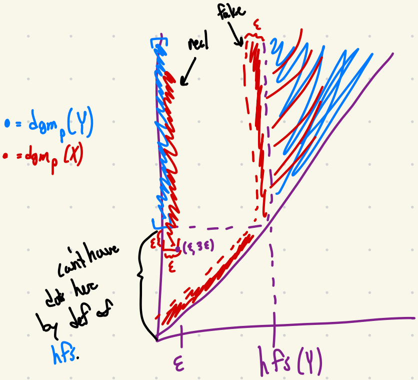
Now by stability $W_\infty(\mathrm{Dgm}_p(Y), \mathrm{Dgm}_p(X)) \leq d_H(Y,X)$

Second major player: resolution
Let's define the homological feature size of Y, denoted $\mathrm{hfs}(Y)$ to be the first nonzero birth and/or death value in the filtration defined by d_Y.

Theorem: (Homology inference theorem) [Cohen-Steiner, Edelsbrunner, Harer, 2007]

Suppose X is a point sample from near Y with $d_H(X,Y) = \varepsilon < \frac{\mathrm{hfs}(Y)}{4}$. Then for all p, $\beta_p(Y)$ is equal to the # of dots born before ε and dead after 3ε.

Proof: HW



Today (10/26): Project

Dates: Project Choice – Due via email in 2 weeks (11/9)

↳ email in advance to discuss before finalizing

↳ no annotated bibliography

Presentation → final exam block; zoom

Paper → due by final exam block

Paper itself: 2 options

① research
- doesn't have to be publishable
- graded on effort and creativity

② expository lecture
- don't expect this to be easy!
- held to high standards

Details on option 1
- find a dataset → make sure you're allowed to use this data!
- do some tda
- find some patterns

Page limit: 6 pages (using template)

Not graded on correct hypothesis!

Option 2
- Choose something we're sure paul didn't cover
- Teach a good class in 20 minutes.

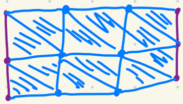
Python package: giotto-tda

↓ calls

ripser

Relative Homology (on the simplicial homology)

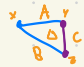
Suppose K is a simplicial complex and L is a subcomplex



for each $p \in \{0, 1, \ldots\}$ define the pth relative chain group to be $C_p(K,L) := C_p(K) / C_p(L)$

Observe that if γ is a p-chain in $C_p(L)$, then $\partial_p(\gamma)$ is a $(p-1)$-chain in $C_{p-1}(L)$.

$$\begin{array}{ccccccc} \cdots \to & C_{p+1}(K) & \xrightarrow{\partial_{p+1}} & C_p(K) & \xrightarrow{\partial_p} & C_{p-1}(K) & \to \cdots \\ & \cup\!| & & \cup\!| & & \cup\!| & \\ & C_{p+1}(L) & & C_p(L) & \to & C_{p-1}(L) & \end{array}$$



Consider $\gamma = A + B$

Then $\gamma = \Delta \in C_2(K,L)$

Thus we can define a map on quotient spaces $\partial_p : C_p(K,L) \to C_{p-1}(K,L)$

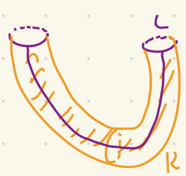
The group of relative p-cycles is $Z_p(K,L)$ is the kernel of this map

The group of relative p-boundaries is $B_p(K,L)$ is the image of $\partial_{p+1} : C_{p+1}(K,L) \to C_p(K,L)$.

Finally, $H_p(K,L) := Z_p(K,L) / B_p(K,L)$ is the relative pth homology group of K mod L.

$\beta_1(K,L) = 1$

\# of connected components in K not "tainted" by L $= \beta_0(K,L) = 0$ $\quad C_1(K,L) \xrightarrow{\partial_1} C_0(K,L) \xrightarrow{\partial_0} 0$



$\beta_0(K) = 1$ — $\beta_0(K,L) = 0$

$\beta_1(K) = 1$ — $\beta_1(K,L) = 1$

$\beta_2(K) = 0$ — $\beta_2(K,L) = 1$

result of manifold modded out by boundary

$K/L =$ pinched torus

$\beta_0(K,L) = 0$

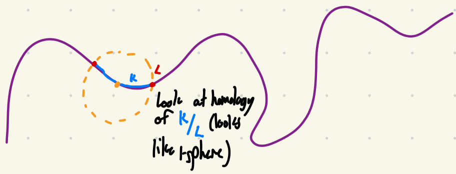
$\beta_1(K,L) = 0$

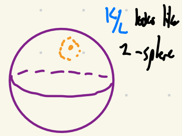
$\beta_2(K,L) = 1$

Local Homology

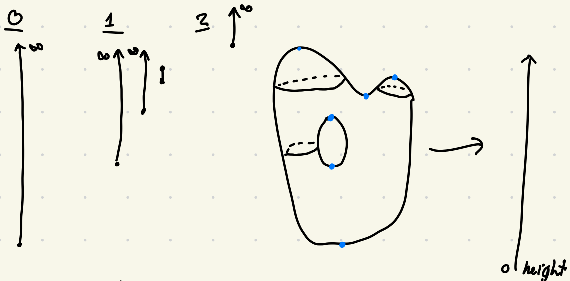
Technically dimension 3 but locally dimension 2. technically dim 2, locally dim 1

local dimension = # of parameters needed (ish)



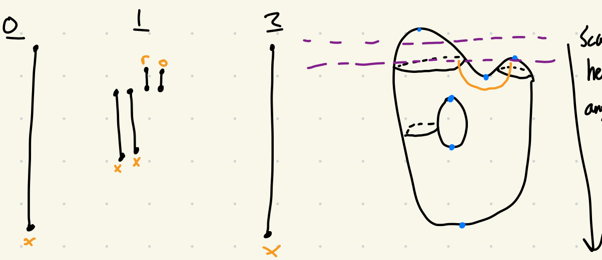


Extended Persistence



Ascend via sublevel sets

Descend via modding out by superlevel sets



Scan down over height, mod out by anything above

Poincaré duality

Manifold of dim d: $\beta_i = \beta_{d-i}$

Elevation Function

Suppose X is a smooth, compact, orientable d-dimensional manifold embedded in $\mathbb{R}^{d+1}$

Given a unit vector $u \in S^d$, consider the filtration function f_u which measures height
"unit sphere in $d+1$ dimensions

in direction u. Formally, $f_u : X \to \mathbb{R}$ via $f_u(x) = x \cdot u$.

Note that every point $x \in X$ is a critical point for precisely 2 functions f_u, where u is either the inward or outward pointing normal vector to X at x.

Fact: x is paired with the same point $p(x)$ for both these directions!

We can define a function on X, called elevation $E : X \to \mathbb{R}$, where

$$E(x) = |f_u(x) - f_u(p(x))|$$

where u is either the inward or outward facing normal vector to X at x.

Persistent Homology Transform (PHT; Turner, Mukerjee, Boyer)

If X is d-dimensional manifold inside $\mathbb{R}^{d+1}$

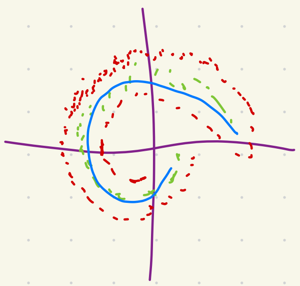
Consider $PHT_X : S^d \to \{\text{all persistence diagrams}\}$
$u \mapsto$ Extended PD of u

Theorem: if $PHT_X = PHT_Y$, then $X = Y$. ...but this is still uncountable...

"How many directions determine a shape?" (Turner, Mukerjee, Curry)

TDA for deep learning

Motivation: Draw me a line that separates these two point sets.

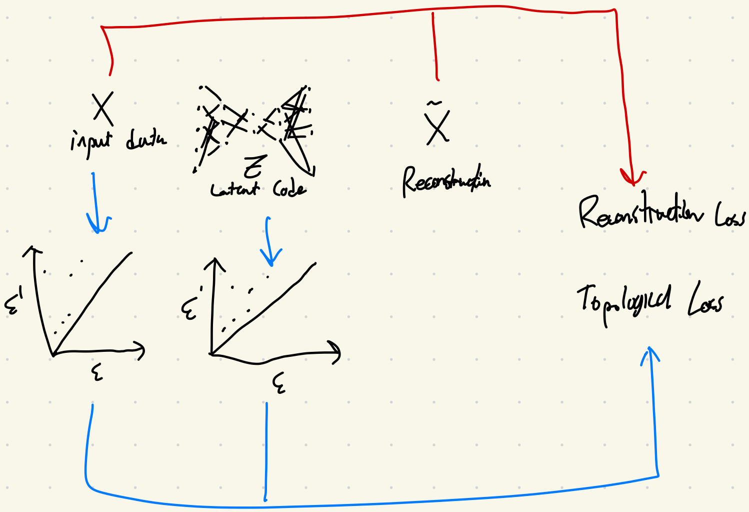


Make topological loss function – restrict connected components instead of smoothness

we want $g \circ f \approx id$

$$\begin{tikzcd} \mathbb{R}^D \arrow[rr, dashed] \arrow[dr, "f"'] & & \mathbb{R}^D \\ & \mathbb{R}^d \arrow[ur, "g"'] & \end{tikzcd}$$

bottleneck $\quad d \ll D$



Turns out we can compute gradients for topological functions based on persistence homology