Categorification of Persistent Homology

Michael Rawson
UMD

March 26, 2020

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

Contents

References I

Bubenik, P., de Silva, V., and Scott, J. (2014).
Metrics for generalized persistence modules.
Foundations of Computational Mathematics, 15(6):1501–1531.

Bubenik, P. and Scott, J. A. (2014).
Categorification of persistent homology.
Discrete Computational Geometry, 51(3):600–627.

Cohen-Steiner, D., Edelsbrunner, H., and Harer, J. (2007).
Stability of persistence diagrams.
Discrete Computational Geometry, 37:103–120.

Cohen-steiner, D., Edelsbrunner, H., and Harer, J. (2009a).
Extending persistence using poincare and lefschetz duality.
FOUNDATIONS OF COMPUTATIONAL MATHEMATICS, pages 79–103.

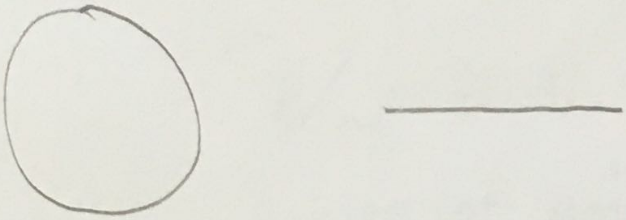
References II

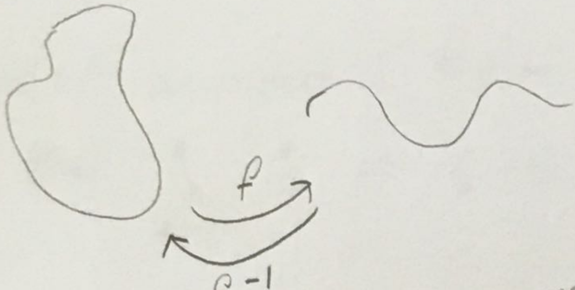
Cohen-steiner, D., Edelsbrunner, H., Harer, J., and Morozov, D. (2009b).
Persistent homology for kernels, images, and cokernels.
In *Proc. Nineteenth Annual ACM -SIAM Symposium on Discrete Algorithms*.

Questions?

Questions?

Topology





f contin.

open balls

$\forall\ U^{\text{open}} \ni f(x_0)\quad \exists\ V^{\text{open}} \ni x_0$ s.t. $f(V) \subset U$

f^{-1} contin

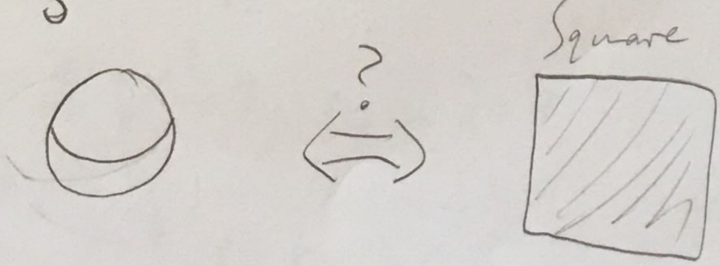
Call f homeomorphism

What is Top/$\sim$?

Classify?

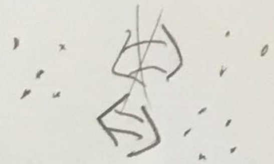
Sphere S^2
(empty)

Square



Homology

In 0 dimension



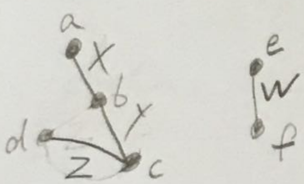
$\mathrm{Top}/_{\sim} \cong \mathbb{N} \cup \{\infty\} \cup \{\infty^{\infty}\} \cup \ldots$

number of points

Connected components: Euler number : $\#V - \#E$

Graph (tree)  $= 6 - 4 = 2$

(with topology)



Map $W \xrightarrow{\partial} e-f$

$X \to a-b$

$Y \to b-c$

$Z \to c-d$

Euler number $= \#V - \#\{a-\cancel{b}+\cancel{b}-\cancel{c}+\cancel{c}-d+e-f\}$

$= \#V - \#\{a-d+e-f\}$

$= 6 - 4 = 2$

Apply Linear Algebra

$V = \mathbb{F}_2^{\#V} = \mathbb{F}_2^{6}$ basis vertices

$E = \mathbb{F}_2^{\#E} = \mathbb{F}_2^{4}$ basis edges

lin map $E \xrightarrow{\partial_1} V \xrightarrow{\partial_0} 0$, $\dim \operatorname{im} \partial_1 = 4$

$$\begin{array}{c c} & \begin{array}{cccc} X & Y & Z & W \end{array} \\ \begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \end{array} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

$\dim \operatorname{im} \partial_0 = 0$, $\dim \frac{\ker \partial_0}{\operatorname{im} \partial_1} = 2$

$\dim \ker \partial_1 = 0$,

$\dim \ker \partial_0 = 6$

for any trianglulated topological space: simplicial Homology

n-chains: C_n

$$\to C_n \xrightarrow{\partial_n} C_{n-1} \xrightarrow{\partial_{n-1}} \cdots \to C_0 \to 0$$

Homology $H_n = \dfrac{\ker \partial_n}{\operatorname{im} \partial_{n+1}}$

Needed: Space
Trianglulation
Field (Ring)
(Recall Gauss Elimination)
(or Smith normal form for rings)

topological spaces X, Y

contin $f : X \to Y$

homomorphism

$f_* : HX \to HY$ (modules)

Can prove $g : Y \to Z$ get $g_* : HY \to HZ$

$(gf)_* = g_* f_*$ Also $f = id \Rightarrow f_* = id_*$

Noticed

H is a nice mapping Top $\to$ Module

and contin maps $\to$ homomorphisms

Define ~~Function~~ Functor btw categories

homomorphism

Functor $H : A \to B$

objects $a \in A \mapsto H(a) \in B$

and $(a \to a') \mapsto H(a \to a') = H(a) \to H(a')$

s.t.

$$\begin{array}{ccccccc} a & \xrightarrow{id} & a & \to & a' & \to & a'' \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ H(a) & \to & H(a) & \to & H(a') & \to & H(a'') \end{array}$$

Needed category have objects and maps (morphisms)

unique identity morphism

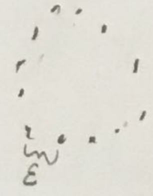
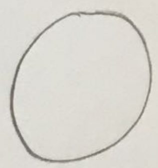
composition that is
$$\begin{tikzcd} a \arrow[r] \arrow[rr, bend right, "\exists"'] & a' \arrow[r] & a'' \end{tikzcd}$$

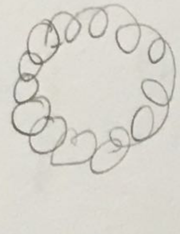
associativity if $a \xrightarrow{f} a' \xrightarrow{g} a'' \xrightarrow{h} a'''$

$$h(gf) = (hg)f$$

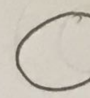
Data Analysis

Samples

$\{x_i\}_{i=1}^{N} = X =$  $\xrightarrow[\text{model}]{\text{Pattern}}$ 

$X_\varepsilon = \{B_\varepsilon(x_i)\}_{i=1}^{N} =$  homology$(X_\varepsilon) = H(S^1)$

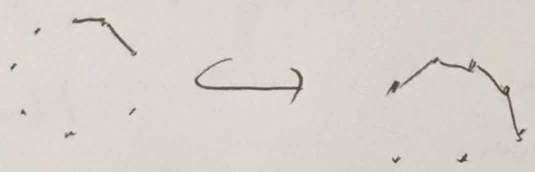
↑ solid

$S^1 =$ 

Find $\varepsilon = ?$

Persistent Homology

$\forall \varepsilon \quad CW(X_\varepsilon) \hookrightarrow CW(X_{\varepsilon'})$ for $\varepsilon \leq \varepsilon'$



$H(X_\varepsilon) \longrightarrow H(X_{\varepsilon'})$

Finite interesting ε

fully connected graph

So $\quad H(X_0) \to H(X_{\varepsilon_1}) \to \cdots \to H(X_{\varepsilon_{max}})$

Persistent Homology

The ε at which connected components merge is only interesting ε's (in H_0).

A set (or multiset) of pairs determines the persistent homology

Called barcode or persistence diagram.

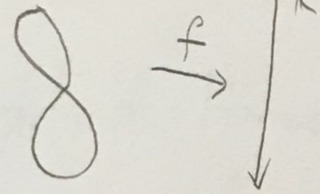
Generalize to any filtration of simplicial complexes.

Extended Persistent Homology

Morse Theory

$f: X \to \mathbb{R}$ in Top

Define $X_\alpha := f^{-1}((-\infty, \alpha])$

X is  by projection

α small large

X_α $\emptyset \hookrightarrow \smile \hookrightarrow \cup \hookrightarrow \cdots \hookrightarrow \delta \hookrightarrow 8$ (f is morse function)

$H\emptyset \to H(\smile) \to H(\cup) \to \cdots$

$X^\alpha := f^{-1}([\alpha, \infty))$

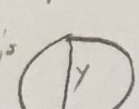
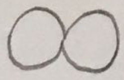
α small large

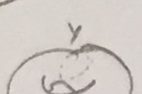
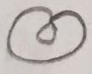
X^α 8 8 $\frown$ $\emptyset$

Extended Persistent Homology extends seq of maps by excision. X, Y topological spaces where $Y \subseteq X$

$H(X,Y) := H(X/Y)$ ($n \geq 1$, X, Y "good" Hatcher)

Collapse Y to a point

if X is  then X/Y is 

or X is  then X/Y is 

Recall n chains $C_n(X)$

let $C_n(X,Y) = \dfrac{C_n(X)}{C_n(Y)}$ (quotient of module)

$$C_n(X,Y) \xrightarrow{\partial_n} C_{n-1}(X,Y) \xrightarrow{\partial_{n-1}} \cdots$$

$$H_n(X,Y) = \frac{\ker \partial_n}{\operatorname{im} \partial_{n+1}}$$

Extended Persistent Homology

α large, $\quad X_\alpha = X / X^\alpha$

$$H(X_\alpha) = H(X, X^\alpha)$$

$$H(\emptyset) \to H(\smile) \to \cdots \to H(8) \to H(8) \to H(8, /) \to$$

$$\to H(8, \wedge) \to H(8, \ell) \to \cdots \to H(8,8) = 0$$

$$\quad\quad\quad\; \| \quad\quad\quad\quad\quad\quad\quad\quad\quad \|$$

$$H(\circ) \quad\quad\quad\quad\quad\quad\quad H(\emptyset)$$

Extended Persistent Homology
on simplicial complex, X (not in $\mathbb{R}^n$)

$f: V \to \mathbb{R}$

$X_\alpha := \iota x\left(f^{-1}((-\infty, \alpha])\right)$

So we get $X_\alpha \hookrightarrow X_{\alpha+\varepsilon}$ and $HX_\alpha \to HX_{\alpha+\varepsilon}$

$X^\alpha := \iota x\left(f^{-1}([\alpha, \infty))\right)$

we get $X/X^\alpha \twoheadrightarrow X/X^{\alpha+\varepsilon}$

and $H(X/X^\alpha) \to H(X/X^{\alpha+\varepsilon})$

Similar to previous def.

$H(\emptyset) \to H(X_\alpha) \to H(X_{\alpha'}) \to \dots \to H(X_{\alpha''}) = H(X, \emptyset) \to$

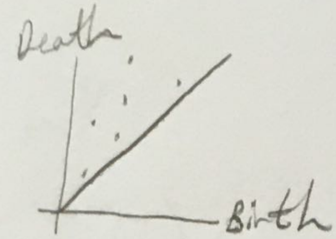
$\to H(X, X_{\alpha'''}) \to \dots \to H(X, X) = 0$

for $\alpha < \alpha' < \alpha'' < \alpha'''$

Bottleneck Distance

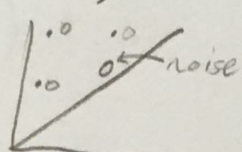
$X =$  in $\mathbb{R}^n$, $\tilde{X} = X + \varepsilon$

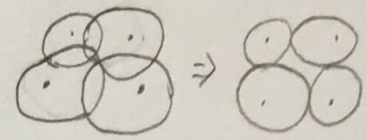
Persistence Diagram



Stability

• is X , o is $\tilde{X}$



ex: 

Metric $d_B(X, Y) = \inf\limits_{\gamma \in \text{Bijection}(\text{Diag}X, \text{Diag}Y)} \sup_{x \in \text{Diag}(X)} \|x - \gamma(x)\|_\infty$

include diagonal for noise in homology

$d_H(X,Y) = \max\{\sup_x \inf_y \|x-y\|_\infty, \sup_y \inf_x \|y-x\|_\infty\}$

$d_H \leq d_B$

Thm [Cohen-Steiner] Let X triangulable space with continuous tame functions $f, g : X \to \mathbb{R}$. Then

$$d_B(D(f), D(g)) \leq \|f-g\|_\infty = \sup_{x \in X} |f(x) - g(x)|$$

Thm $\forall \varepsilon > 0$ with $d_H(X, P) < \varepsilon < hfs\ X/4$ and sufficiently small $\delta > 0$, $\dim H(X^{+\delta}) = \dim P_\varepsilon^{3\varepsilon}$ if either finite

$X^{+\delta} = \{x \in \mathbb{R}^n : d(x, X) \leq \delta\}$

Where homological feature size are critical points of $d^X : \mathbb{R}^n \to \mathbb{R}$ the distance to X

Persistent Homology for Kernels, images, cokernels

Consider X and $Y \subseteq X$

$f: X \to \mathbb{R}$ (tame)

$g: Y \to \mathbb{R}$

$f(y) \leq g(y)$ on Y

$X_a = f^{-1}((-\infty, a])$

$Y_a = g^{-1}((-\infty, a])$

$Y_0 \hookrightarrow X_0$

$HY_0 \to HX_0$

then

$$\begin{tikzcd} HX_0 \arrow[r] & \cdots \arrow[r] & HX_m \\ HY_0 \arrow[u, "j_0"] \arrow[r] & \cdots \arrow[r] & HY_m \arrow[u, "j_m"] \end{tikzcd}$$

commutes [C.-S., E., H.]

We get $\ker j_0 \to \cdots \to \ker j_m$

$\operatorname{im} j_0 \to \cdots \to \operatorname{im} j_m$

$\operatorname{coker} j_0 \to \cdots \to \operatorname{coker} j_m = \dfrac{H(X_m)}{\operatorname{im} j_m}$

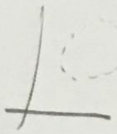
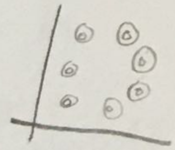
Get Persistence Diagram$(\ker g \to f)$, Dgm$(\operatorname{im} g \to f)$, Dgm$(\operatorname{cok} g \to f)$

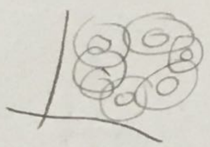
Consider Multiparameter Persistence

Kernel Method

Data $X = \{x_i\}_{i=1}^{N}$

Replace x_i with Gaussian Bump of bandwidth σ

 $\Rightarrow$ σ small  $= X_\sigma$

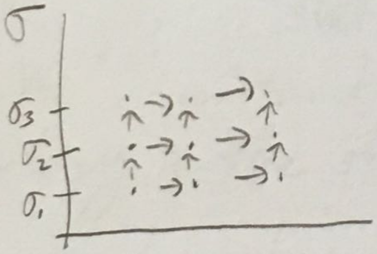
σ large 

height $f_\sigma : X_\sigma \to \mathbb{R}$

$Dgm(f_\sigma)$ varies

If we set $\sigma_1 < \sigma_2 < \ldots < \sigma_m$

looking at $\ker_{im/coker}$ $\sigma_i \to \sigma_{i+1}$ can potentially encapsulate/clarify information



Can look at persistence diagram of any path

Generalize with Category Theory

Thm [Bubenik]: Let B be the set of finite barcodes, d_B the bottleneck distance, and d the interleaving distance. The mapping χ defined by $\chi(\{I_k\}_{k=1}^{n}) = \bigoplus_{k=1}^{n} \chi_{I_k}$ gives an <u>isometric</u> embedding of metric spaces

$$\chi : (B, d_B) \hookrightarrow (Vec^{(\mathbb{R},\leq)}, d).$$

Category $Vec^{(\mathbb{R},\leq)}$ objects are collections of maps $\forall a \in \mathbb{R}$, $a \mapsto A \in Vec$ and $(a \leq b) \mapsto (A \to B)$.

Morphisms are natural transformations η such that

$$\begin{tikzcd}
& a \arrow[dl] \arrow[d] \arrow[dr] & \\
A \arrow[rr, bend right, "\eta(a)"] \arrow[d] & b \arrow[dl] \arrow[dr] & A' \arrow[d] \\
B \arrow[rr, bend right, "\eta(b)"'] & & B'
\end{tikzcd}$$

For $Vec^{(\mathbb{R},\leq)}$ constant except at finite points, define homology in the standard way.

For topological space X and $f : X \to \mathbb{R}$

Let $F \in Top^{(\mathbb{R},\leq)}$ where $F(a) = f^{-1}((-\infty, a])$

Then $H_k F \in Vec^{(\mathbb{R},\leq)}$ some field $\mathbb{F}$

$$HF = \bigoplus_{k=0}^{\infty} H_k F$$

Define functor

$T_\varepsilon : \mathbb{R} \to \mathbb{R}$, $a \mapsto a + \varepsilon$

nat. transf. $\eta_b : Id_{(\mathbb{R},\leq)} \Rightarrow T_b$

Interleaving Distance

D some category

$\varepsilon \geq 0$

$F, G \in D^{(\mathbb{R}, \leq)}$

Def: An ε-interleaving of F and G consists of natural transformations $\phi : F \Rightarrow GT_\varepsilon$ and $\psi : G \Rightarrow FT_\varepsilon$ ie

$$
\begin{tikzcd}
(\mathbb{R}, \leq) \arrow[r, "T_\varepsilon"] \arrow[d, "F"'] \arrow[dr, phantom, "\overset{\phi}{\Rightarrow}"] & (\mathbb{R}, \leq) \arrow[r, "T_\varepsilon"] \arrow[d, "G"] \arrow[dr, phantom, "\overset{\psi}{\Rightarrow}"] & (\mathbb{R}, \leq) \arrow[d, "F"] \\
D \arrow[r, equal] & D \arrow[r, equal] & D
\end{tikzcd}
$$

s.t. $(\psi T_\varepsilon)\phi = F\eta_{2\varepsilon}$ and $(\phi T_\varepsilon)\psi = G\eta_{2\varepsilon}$. $(*)$

Now this implies, $a \leq b$

$$
\begin{tikzcd}
F(a) \arrow[rr] \arrow[dr, "\phi(a)"'] & & F(b) \arrow[dr, "\psi(b)"] & \\
& G(a+\varepsilon) \arrow[rr] & & G(b+\varepsilon)
\end{tikzcd}
,
$$

$$
\begin{tikzcd}
& F(a+\varepsilon) \arrow[rr] & & F(b+\varepsilon) \\
G(a) \arrow[rr] \arrow[ur, "\psi(a)"] & & G(b) \arrow[ur, "\psi(b)"'] &
\end{tikzcd}
$$

And $(*)$ implies
$$
\begin{tikzcd}
F(a) \arrow[rr] \arrow[dr, "\phi(a)"'] & & F(a+2\varepsilon) \\
& G(a+\varepsilon) \arrow[ur, "\psi(a+\varepsilon)"'] &
\end{tikzcd}
$$

and
$$
\begin{tikzcd}
& F(a+\varepsilon) \arrow[dr, "\phi(a+\varepsilon)"] & \\
G(a) \arrow[ur, "\psi(a)"] \arrow[rr] & & G(a+2\varepsilon)
\end{tikzcd}
$$

Def: Interleaving distance

$$d(F,G) := \inf\{\varepsilon \geq 0 : F, G \text{ are } \varepsilon\text{-interleaved}\}.$$

This is an extended pseudometric

Prop: Let $F, G : (\mathbb{R}, \leq) \to D$ and $H : D \to E$. If F and G are ε-interleaved then so are HF and HG. Thus $d(HF, HG) \leq d(F, G)$.

Pf:

$$\begin{tikzcd}
(\mathbb{R},\leq) \arrow[r, "T_\varepsilon"] \arrow[d, "F"'] & (\mathbb{R},\leq) \arrow[r, "T_\varepsilon"] \arrow[d, "G"] & (\mathbb{R},\leq) \arrow[d, "F"] \\
D \arrow[r, equal] \arrow[d, "H"'] & D \arrow[r, equal] \arrow[d, "H"] & D \arrow[d, "H"] \\
E \arrow[r, equal] & E \arrow[r, equal] & E
\end{tikzcd}$$

(by functorality) $\square$.

Thm Stability. $X \in \mathbf{Top}$. $f, g : X \to \mathbb{R}$. $F, G \in \mathbf{Top}^{(\mathbb{R},\leq)}$

$F(a) = f^{-1}((-\infty, a])$. etc.

Then $d(HF, HG) \leq \|f-g\|_\infty$

Pf: $\varepsilon := \|f-g\|_\infty$

$F(a) = f^{-1}(-\infty, a] \subseteq g^{-1}(-\infty, a+\varepsilon] = G(a+\varepsilon)$

$G(a) \subseteq F(a+\varepsilon)$

So F and G are ε-interleaved

So are HF and HG.

Thus $d(HF, HG) \leq \|f-g\|_\infty$. $\square$