Persistence Stability for Metric Thickenings

Michael Moy

March 29, 2021
Applied Topology

- Premise: data can have shape

- The field of applied topology uses tools from topology to study the shape of datasets.
What is the topology of a dataset in \( \mathbb{R}^n \)?

We begin by associating a more interesting topological space to the data points.
A simplicial complex is formed from simplices.
Simplicial Complexes

An abstract simplicial complex records the vertices that form each simplex. The geometric realization has a topology.

\[
\{ \{1\}, \{2\}, \{3\}, \{4\}, \{1, 2\}, \{1, 3\}, \\
\{2, 3\}, \{2, 4\}, \{3, 4\}, \{1, 2, 3\} \}
\]

Starting with a dataset, we can build a simplicial complex using the data points as vertices.
A continuous function between simplicial complexes can be defined by specifying a function on vertices that takes simplices to simplices. This is called a *simplicial map*. 
Given a vertex set that is a metric space, we can form the Vietoris–Rips complex

\[ VR(X; r) = \left\{ \{x_1, \ldots, x_n\} \subseteq X \mid \text{diam} (\{x_1, \ldots, x_n\}) \leq r \right\} \]
The Vietoris–Rips complex grows as the parameter $r$ grows: if $r_1 \leq r_2$, then

$$\text{VR}(X; r_1) \subseteq \text{VR}(X; r_2)$$

Other simplicial complexes are common as well, including Čech complexes.
Consider all parameters at once:

$$VR(X; \_\;) = \{VR(X; r) \mid r \in \mathbb{R}\}$$

This is called a filtration, and comes with inclusion maps $VR(X; r_1) \hookrightarrow VR(X; r_2)$ for all pairs $r_1 \leq r_2$. 
Persistent homology is based on homology.

- $H_n(X)$ is a vector space, with dimension equal to the number of $n$-dimensional holes in $X$.
- We’ll fix $n$ and write $H(X)$.
- $H$ is a functor: given a map $f : X \rightarrow Y$, we have a map $H(f) : H(X) \rightarrow H(Y)$. 
Applying $H$ to $\text{VR}(X; \_)$ gives a *persistence module* $H(\text{VR}(X; \_))$. This consists of vector spaces and linear maps.

Persistent homology records the birth and death times of nonzero elements.
Persistent Homology

\( VR(X; \_\_\_\_) \)

\( H_1 \)
Persistence Diagrams and Barcodes

VR(X; _)

VR(X; 10)  VR(X; 12)  VR(X; 15)  VR(X; 19)

$H_1$

Birth

Death
Stability of Persistent Homology

How do small changes to a dataset affect the persistence diagram?

Dataset 1
Dataset 2

$H_1$ diagrams
Gromov–Hausdorff Distance and Bottleneck Distance

\[ d_{GH}(X, Y) = \frac{1}{2} \inf_{C} \sup \{ |d_X(x, x') - d_Y(y, y')| : (x, y), (x', y') \in C \} \]

\[ d_b(D_1, D_2) = \inf \{ \varepsilon : \text{there exists an } \varepsilon\text{-matching between } D_1 \text{ and } D_2 \} \]
Interleavings

An isomorphism:

\[ H(\text{VR}(X; _)) \]

\[ H(\text{VR}(Y; _)) \]

An \( \varepsilon \)-interleaving:

\[ H(\text{VR}(X; _)) \]

\[ H(\text{VR}(Y; _)) \]
Theorem (Chazal, de Silva, Glisse, and Oudot)

If $\mathbb{U}$ and $\mathbb{V}$ are $q$-tame persistence modules that are $\varepsilon$-interleaved, then $d_b(dgm(\mathbb{U}), dgm(\mathbb{V})) \leq \varepsilon$. 
Lemma (Chazal, de Silva, and Oudot)

Let $X$ and $Y$ be metric spaces. For any $\varepsilon > 2d_{GH}(X, Y)$, the persistence modules $H(VR(X; \_))$ and $H(VR(Y; \_))$ are $\varepsilon$-interleaved.

The interleaving comes from maps on simplicial complexes that commute up to homotopy.

As an example, consider:
Interleavings for Vietoris–Rips Complexes
Theorem (Chazal, de Silva, and Oudot)

Let $X$ and $Y$ be totally bounded metric spaces. Then

$$d_b\left(dgm\left(H(VR(X;\_))\right), dgm\left(H(VR(Y;\_))\right)\right) \leq 2d_{GH}(X, Y)$$
VR(\(X; r\)) can be formed for any metric space \(X\), including those with infinitely many points.

Stability motivates the study of these complexes.

Difficulties arise: the inclusion \(X \hookrightarrow \text{VR}(X; r)\) is not always continuous.
Metric Thickenings

Defined by Adamaszek, Adams, and Frick – an alternate approach for infinite metric spaces.

\[ \text{VR}^m(\mathcal{X}; r) = \left\{ \sum_{i=1}^{n} \lambda_i \delta_{x_i} \mid \lambda_i \geq 0 \text{ for all } i, \sum_{i=1}^{n} \lambda_i = 1, \{x_1, \ldots, x_n\} \in \text{VR}(\mathcal{X}; r) \right\} \]

The \( \delta_{x_i} \) are Dirac delta measures. The space is equipped with the Wasserstein metric.
Metric Thickenings

- Metric thickenings are essentially simplicial complexes built on metric spaces, but given different topologies.
- The topologies agree with simplicial complexes in the finite case, but may be different in the infinite case.
- The inclusion $X \hookrightarrow \text{VR}_m(X; r)$ is continuous (for all $r \geq 0$).
- We get similar filtrations: $\text{VR}_m(X; \_)$
- No exact analog of simplicial maps.
A Potentially Different Topology

Simplicial Complexes

Metric Thickenings
The technique for Vietoris–Rips complexes cannot be used.

We will still construct interleavings starting with maps on spaces.

We begin with a metric space $X$ and a finite $\varepsilon$-sample $F \subseteq X$.

We will define maps between $\operatorname{VR}^m(X;\_)$ and $\operatorname{VR}^m(F;\_)$. 
Given a map $\varphi_r : X \to \text{VR}^m(F; r + \varepsilon)$, we get a continuous induced map $\tilde{\varphi}_r : \text{VR}^m(X; r) \to \text{VR}^m(F; r + \varepsilon)$ defined by $\tilde{\varphi}_r(\sum_i \lambda_i \delta_{x_i}) = \sum_i \lambda_i \varphi_r(x_i)$. 
We choose $\varphi_r$ carefully for arbitrary $X$ and $\varepsilon$-sample $F$. We will require

- $\varphi_r$ is continuous
- $\supp(\varphi_r(x)) \subseteq B_{\varepsilon}(x)$
- For all $f \in F$, $\varphi_r(f) = \delta_f$

Define $\varphi_r(x) = \sum_{j=1}^{n} m_j(x)\delta_{f_j}$ with appropriate $m_j$
The induced maps $\tilde{\varphi}_r : \text{VR}^m(X; r) \to \text{VR}^m(F; r + \varepsilon)$ can be shown to commute with inclusion maps up to homotopy. Applying $H$ gives an interleaving.
Persistence Stability for Metric Thickenings

We’ve compared $X$ to a finite sample $F$. We find

$$H(\text{VR}^m(X; \_))$$

is interleaved with

$$H(\text{VR}^m(F; \_))$$

is isomorphic to

$$H(\text{VR}(F; \_))$$

is interleaved with

$$H(\text{VR}(X; \_))$$
 Persistence Stability for Metric Thickenings

- So $H(VR^m(X;\_))$ is interleaved with $H(VR(X;\_))$, with the interleaving depending on the finite sample.

- If $X$ is totally bounded, the sample can be made arbitrarily fine, so the persistence modules are $\varepsilon$-interleaved for any $\varepsilon > 0$. 
Theorem

If $X$ is a totally bounded metric space, then $H(\text{VR}^m(X;\_))$ and $H(\text{VR}(X;\_))$ have identical persistence diagrams.

This implies

Theorem

If $X$ and $Y$ are totally bounded metric spaces, then

$$d_b\left(dgm\left(H(\text{VR}^m(X;\_))\right), dgm\left(H(\text{VR}^m(Y;\_))\right)\right) \leq 2d_{GH}(X, Y)$$

Similar results hold for both intrinsic and ambient Čech complexes.
Vietoris–Rips complexes and metric thickenings carry the same persistent homology information.

Metric thickenings provide an alternate approach for infinite spaces.

If the persistent homology of either $\text{VR}(X;\_)$ or $\text{VR}^m(X;\_)$ can be found, then the other is known.

These results motivate further work on the homotopy types of metric thickenings.


References II

Robert Ghrist.  
Barcodes: The persistent topology of data.  

Francis C Motta, Rachel Neville, Patrick D Shipman, Daniel A Pearson, and R Mark Bradley.  
Measures of order for nearly hexagonal lattices.  

Paul Bendich, James S Marron, Ezra Miller, Alex Pieloch, and Sean Skwerer.  
Persistent homology analysis of brain artery trees.  