PERSISTENT HOMOLOGY AND NEUROSCIENCE

A reading project based on the sources cited in “References”

Kindness Chen[1], Yejin Lee[1,2], Parker Sherry[1,3]

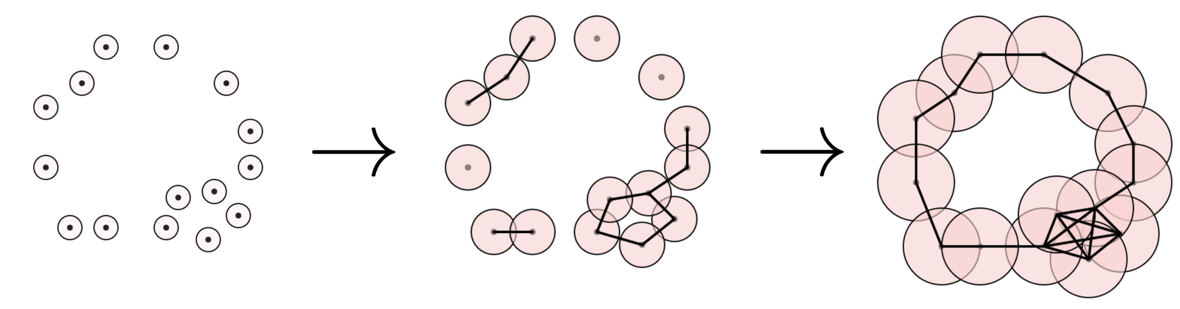
[1] Department of Mathematics and Statistics, McGill University [2] Department of Physiology, McGill University [3] Department of Physics, McGill University

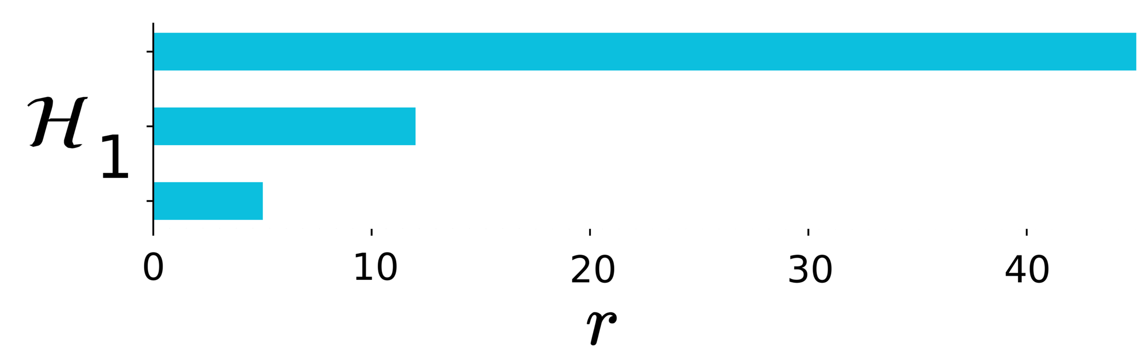
1 OVERVIEW

We aim to identify prevalent shapes formed by collections of points in $\mathbb{R}^n$ by applying concepts from persistent homology. Such points are called vertices, and we identify the shapes by identifying the persistent holes in a progression of Vietoris-Rips simplicial complexes. Finally, we see the applications of this technique onto topics in neuroscience.

Keywords:

- **Simplicial complex**
- **k^{th} homology group and k-dimensional holes**
- **Vietoris-Rips simplicial complex**
- **Functional connectivity**



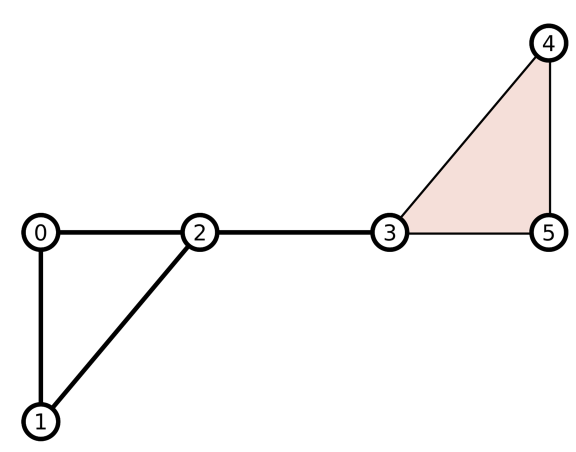


The Vietoris Rips Complex provides an elegant way of avoiding the construction of arbitrary simplices. By letting r increase over a given interval, it is possible to construct a sequence of "nested" complexes such that $VR(X, r_i) \subseteq VR(X, r_{i+1})$. This allows the features of the connected data set to "evolve over time", with the implicit assumption that **features which *persist* over a longer interval are topologically nontrivial**, possibly hinting at some important feature of the data.

2.5 The k^{th} *barcode* of a Vietoris Rips Complex $VR(X,r)$ is a graphical representation of the persistence of k-dimensional holes in X over a range of values of r. The horizontal axis represents the value of r, while the number of bars present at a point along the axis represents how many k-dimensional holes exist at a given r value.

2 DEFINITIONS

2.1 For a finite set of vertices $V = \{v_0, v_1, \dots, v_n\}$ a collection of k+1 vertices is called a *k-simplex*. A subset X of the power set $P(V)$ that is closed under subsets is called an *abstract simplicial complex.*



Simplicial complex with six vertices, seven 1-simplices and one 2-simplex

2.2 Let C_k be the vector space of formal linear combinations of k-simplices over a field F. The k^{th} *boundary map* $D_k : C_k \to C_{k-1}$ is a linear map which sends k-simplices s in X to a finite formal sum of k-1 simplices via

$$D_k(s) = \sum_{i=0}^{k} (-1)^i \left\{ s_0, \dots, \widehat{s_i}, \dots, s_k \right\} = \{s_1, s_2, \dots, s_k\} - \{s_0, s_2, \dots, s_k\} + \dots$$

The kernel of D_k, $ker(D_k)$, consists of linear combinations of k-simplices such that the resulting formal sum in C_{k-1} is zero. Such an object roughly corresponds to a k-cycle. The image of D_k, $Im(D_k)$, is the set of all formal sums of k-1 simplicies that are the result of D_k acting on formal sums of k simplicies, and correspond to the k-1 “boundaries” of X.

2.3 The k^{th} *homology group* of a simplicial complex X is given by the quotient space

$$H_k(X) = \frac{ker(D_k)}{Im(D_{k+1})}$$

The construction of $H_k(X)$ ensures that k-1 cycles which do not enclose k-dimensional features remain in $H_k(X)$, while k-1 cycles generated by such features are removed from the space. Because of this, **the dimension of $H_k(X)$ corresponds to the number of k-dimensional holes in** X, and each basis element of $H_k(X)$ corresponds to the boundaries that outline the holes.

In the transition from a collection of abstract vertices to a real-world data set, two issues arise. The first is that **discrete data points are not connected graphically**. Their distribution may suggest that they adhere to some underlying trend, but there does not normally exist some straight-line path between two neighbouring data points such that all intermediate values lie on said path. This means that the decision to connect any two such data points is somewhat arbitrary, and that any approach to data analysis which involves simplicial homology must account for this fact. The second is that discrete data points do not arrange themselves into neat geometrical figures, as there is always **a degree of randomness or uncertainty in the distribution of any data set**. This means that our algorithm must also be robust to noise.

2.4 A *Vietoris-Rips Complex* $VR(X,r)$ defined on a metric space (M,d) is the simplicial complex generated by a set of points X in M along with a distance r such that a set of k points are joined in a k-1 simplex if the distance between any two points in the set is at most $2r$. Visually, we often represent r as a radius extending from every point in X.

3 COMPUTATION

For the Vietoris-Rips Complex $VR(X,r)$ of a data set X in M with a distance of r, let $M(D_K)$ be the matrix representation of D_K with respect to the basis of C_K and C_{K-1}. By construction, the columns of $M(D_K)$ are the images of the basis vectors of C_K under D_K with respect to the basis of C_{K-1}.

We expect $M(D_K)$ to be either diagonalizable or reducible to reduced row-echelon form using elementary matrix operations. Once this is done, we will have obtained an equivalent matrix whose row rank is the dimension of $Im(D_K)$, and whose nullity is the dimension of $Ker(D_K)$.

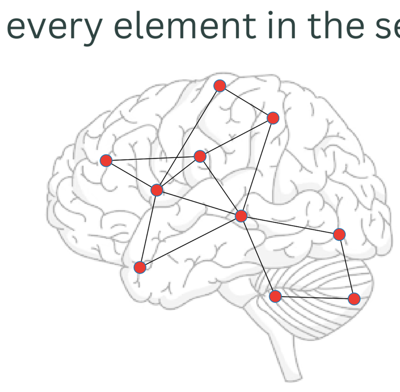
For a given r value, we compute the dimension of the quotient space $H_K(K)$:

$$dim(H_k(K)) = dim(Ker(D_K)) - dim(Im(D_{k-1}))$$

and find the basis of $H_K(K)$ which corresponds to the k-dimensional holes, and append the kth barcode by either beginning a new bar (for a new k-dimensional hole) or extend the length of a pre-existing bar (for a k-dimensional hole that already exists). The placement of the bars along the vertical axis is arbitrary.

This process is repeated for increasing r until the value at which all elements in X are in an $n-1$ simplicial complex, i.e. the point at which every element in the set is connected to every other element.

4 APPLICATIONS



Functional Connectivity of Brain

- Functional connectivity is an approach describing the **brain network across separate brain regions by co-activation patterns and strength**, as opposed to structural connectivity (SC) which anatomically explains the neural network by the physical neuronal pathways of electrical signals.
- Persistent homology-based functional connectivity is used for investigating an overall information transfer in brain and cognitive abilities, offering advantages over traditional graph theory-based measures.
- In this case, a simplicial complex is **a network of nodes (brain regions) and edges (their functional connections).**
- Data are extracted from resting state fMRI scan in time series measuring the O_2 consumption level of neurons, thus they are an indirect and instrumental measure of neuronal activity.
- The advantages include its **robustness to noise & weak connectivity** and that this method takes **integrative information processing** of brain into account.
- The research findings suggest a negative correlation between the dispersion of information and cognitive abilities.

REFERENCES

[1] Adams H., Bush J., Mirth J. (2020). Topological Data Analysis. In Carter N., Data Science for Mathematicians (p.441-473).

[2] Sizemore, A. E., Phillips-Cremins, J. E., Ghrist, R., & Bassett, D. S. (2019). The importance of the whole: Topological data analysis for the network neuroscientist. Network neuroscience, 3(3), 656–673. https://doi.org/10.1162/netn_a_00073

[3] Ryu, H., Habeck, C., Stern, Y., & Lee, S. (2023). Persistent homology-based functional connectivity and its association with cognitive ability: Life-span study. Human Brain Mapping, 44(9), 3669–3683. https://doi.org/10.1002/hbm.26304

ACKNOWLEDGEMENTS

We would like to express our gratitude to Julien Cheng, our project supervisor under the Directed Reading Program offered by the Department of Mathematics and Statistics at McGill University. His guidance and support have been integral in making this project possible.

Special thanks to Professor Seonju Lee of Columbia University for her insights into the biostatistical applications of persistent homology.