Persistence Homology

William Cruse

October 2020

Abstract

Persistence Homology is an algebraic method used to measure topological invariants between different manifolds. By converting a topological manifold into a simplicial complex, abelian groups can be formed out the span of the cycles that make the simplicial complex. These groups can then be ordered by the dimension of the cycles that make them up. By applying a homomorphism, the image and kernal of these groups can be determined and used to create a quotiant group called the Homology Group whose order describes the amount of holes present in a manifold for some dimension. By creating a subset filtration of manifolds, the Persistence Homology Group can be defined to reference the change in the amount of holes between two manifolds.

1 Introduction

Topology is often considered and abstraction of geometry. In topology, shapes or spaces do not have metrics, lengths, or angles. The only things that separate two topological manifolds is the dimension and the amount of holes. Computationally, these holes can be counted by converting a manifold into a collection of simple n-dimensional polygons called a simplicial complex. By creating abelian groups based on the span of the k-dimensional cycles that make up the complex, a boundary homomorphism can be used to examine the image, or boundaries, and the kernal, or cycles, of these abelian groups which are then used to create the quotient Homology Group of cycles modulo boundaries whose order gives information about the amount of k-dimensional holes on some manifold. By creating a subset filtration of a changing manifold or simplicial complex and a positive integer to reference the different manifolds, the Homology Group can be altered to create the Persistence Homology Group of cycles of some i^{th} manifold modulo the intersection of boundaries of a larger manifold in the filtration and the cycles of the i^{th} manifold. The order of the Persistence Homology Group gives information about the amount of holes added as the manifold changes.

Topology is a relatively new subject in mathematics. The concepts of point-set topology have been used as the basis for calculus since the 1700s, but topology on its own was not studied as a subject until the late 1800s. The need to study Topology came from the idea that certain geometric objects can have properties in the most abstract or general settings. The most famous problem in topology is Leonhard Euler's *Seven Bridges of Könisburg* problem in which Euler solved that given four islands and seven bridges, it is impossible to travel to each island by crossing each bridge only once. This gave way to his famous polyhedron formula, Vertices - Edges + Faces = 2. Topology would not be a major subject of mathematics until 1895 with Henri Poincaré's groundbreaking paper, *Analysis Situs* which used algebraic structures to distinguish the differences between topological spaces. Poincaré's paper defined the homotopic function and simplicial Homology, setting the groundwork for the field of algebraic topology.

For many years after *Analysis Situs* was published, homology remained a very abstract subject. The only use it had was to use algebraic structures to count holes in order to distinguish different topological spaces or manifolds. Homology would stay a subject of pure mathematics for over 100 years until the dawn of computers. Computer scientists became interested in writing algorithms to compute the number of holes in a given object. This lead to defining persistence homology as a slight modification of homology to compare the amount of holes between a changing manifold.

By defining Persistence Homology, a new subject of topology was created named Computational Topology. The subject defined multible ways to observe topological invariants of data sets based on the ideas of persistence homology. This field of mathematics is still very new, but researchers are constantly finding applications. For example, researchers are currently using the persistence homology to improve facial recognition software by examining the changes of faces based on different angles.

2 Simplicial Complexes

One of the most important concepts discussed in Poincaré's, *Analysis Situs* is the homotopic function.

Definition. *Let X and Y be two topological manifolds. A **Homotopy** is a function $H : X \times [0, 1] \to Y$ that continuously transforms one topological space into another.*

Example. *Let X be the topological manifold in Figure 1 [6] and Y be the topological manifold in Figure 2 [6]. By using a homotopic function, X can be continuously morphed into Y.*

Note that the use of the real number interval $[0, 1]$ is a part of the function as a continuous reference point. This allows for another homotopic function to be applied as an inverse of sorts. When a homotopic function exists between two manifolds, those manifolds are considered homotopic equivalent. Due note that the holes of both manifolds are still preserved in the morphing process. Homotopic functions will be used to morph a manifold into a simplicial complex.

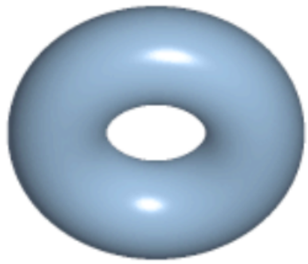

Figure 1: X

Figure 2: Y

In order to make counting cycles easier, a simplex is used rather than a more complicated polygon. Homology can be done with more complicated polygon, but to avoid confusion, this paper will only discuss the use of simplexes.

Definition. *An **simplex** is the simplest euclidean polygon for some n-dimension*

Example. *Figure 3 [7] below details the 0, 1, 2, and 3-simplexes, respectively*

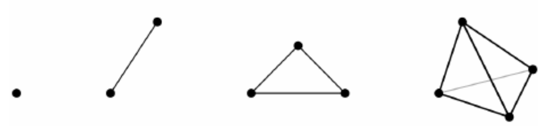

Figure 3: Simplexes

Definition. *A **simplicial complex** is a collection of simplexes.*

Example. *Figure 4 [8] below is a simplicial complex*

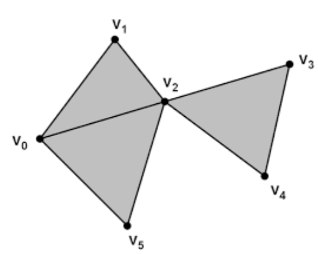

Figure 4: Simplicial Complex

3 Homology

Imagine traversing a simplecial complex like it was an oriented graph. Each simplex of any n-dimension can be considered a cycle. It is important to note that any simplicial complex can be defined by the simplexes that make it up of any smaller n-dimension. This fact is used to create an abelian group.

Definition. *Let K_k be a simplicial complex of dimension k and $\sigma_{i,j} = (x_0...x_k)$ be some i^{th} cycle in K_k indexed by j. Then span of $k-1$ cycles form an abelian group*

$$C_i(K) = span\{\sigma_{i,1}, \sigma_{i,2}, \sigma_{i,3}...\}$$

This allows for simplicial complexes to be defined as groups. The additive inverse of a cycle simply reverses the cycle, for example $-(x_1x_2) = (x_2x_1)$ By examining the $k-1$ cycles of a k-dimensional simplicial complex, we can describe the complex in terms of it's boundaries. For example take the 2-simplex, a simple triangle, and notice that the 1-simplexes that form it are the triangle's boundary. Now that simplicial complexes can be represented as groups, a homomorphism can be applied from an i^{th} dimensional group to an $(i-1)^{th}$ dimensional group.

Definition. *The **boundary homomorphism** is a group homomorphism, $\partial_i : C_i(K) \to C_{i-1}(K)$, that satifies the following conditions*

1. $\partial_i(\sigma_{i,j}) = \sum\limits_{d=1}^{k} (-1)^d (x_1...x_{d-1}x_{d+1}...x_k)$
2. $\partial_{i-1}(\partial_i(\sigma_{i,j})) = 0$

Condition two is important as it intuitively states that the boundary of a boundary does not exist. It will be useful when examining the image later.

Definition. *The **chain complex** is a collection abelian groups linked together by the boundary homomorphism*

$$... \xrightarrow{\partial_{i+1}} C_i(K) \xrightarrow{\partial_i} C_{i-1}(K) \xrightarrow{\partial_{i-1}} ... \xrightarrow{\partial_2} C_1(K) \xrightarrow{\partial_1} C_0(K) \xrightarrow{\partial_0} 0$$

As mentioned previously, homology is about counting the holes by examining a quotient group made of the cycles modulo boundaries. Imagine traversing a cycle on a simplicial complex. Starting at one node, the cycle may be traversed until the node is reached again, ending the cycle. If the boundary homomorphism takes a cycle returns one that starts and ends on the same spot, the cycle belongs to the kernal.

Definition. *The **cycle group**, Z_i, is defined as $\ker(\partial_i)$ and the **boundary group**, B_{i+1}, is defined as $Im(\partial_{i+1})$*

Theorem. *$B_{i+1} = Im(\partial_{i+1})$ is a normal subgroup of $Z_i = \ker(\partial_i)$*

Proof. Since the kernal of an group homomorphism is a subgroup and the groups are abelian, it suffices to show that B_{i+1} is a subset of Z_i. Let $\sigma \in B_{i+1}$. Then

$$\begin{aligned} \sigma &= \partial_{i+1}(\tau) \\ \partial_i(\partial_{i+1}(\tau)) &= 0 \\ \partial_i(\sigma) &= 0 \end{aligned}$$

Therefore $\sigma \in Z_i = \ker(\partial_i)$ and B_{i+1} is a normal subgroup of Z_i $\square$

This theorem states that every boundary is a cycle itself. Using a quotient group, one can examine how many cycles fail to bound a simplicial complex.

Definition. *The **homology group** is the quotiont group of Z_i modulo B_{i+1}*

$$H_i = Z_i / B_{i+1}$$

Definition. *The i^{th}**-Betti number**, β_i, is the number of i^{th}-deminsional holes of a simplicial complex such that $\beta_i = |H_i| = |Z_i| - |B_{i+1}|$.*

4 Persistence Homology

Computing persistence homology only requires a simple modification to the original homology group definition.

Definition. *Let K be a simplicial complex. A **filtration** is a partition of the complex such that $K^1 \subseteq K^2 \subseteq ... \subseteq K$ where each K^m is also complex.*

The filtration allows each complex to be referenced as it changes. The persistence homology group takes advantage of this by examining which cycles of some K^m complex do not become boundaries of a K^{m+p} complex. Intuitively, as a complex increases in size it is subject to change properties. A cycle will either become a boundary or not. This allows for an easy way to compute the differences as a complex changes.

Definition. *The **persistence homology group** is the quotient group of Z_i^m modulo $\{B_i^{m+p} \cap Z_i^m\}$*

$$H_i^{m,p} = Z_i^m / \{B_i^{m+p} \cap Z_i^m\}$$

5 Conclusion

5.1 Application

Beyond all of the algebra, the field is very new to the world of applied mathematics. A lot of the ideas presented can be translated into various programming languages including C++, Python, MATLAB, R, and more. Different methods for measuring topological invariants are being used and developed such as the Vietoris-Rips complex and the Ĉech complex. There are many researchers using these ideas to study more abstract geometric properties such as the effects of gerrymandering or melting ice in the Arctic. On the pure math side, researchers are also attempting to find uses for 3-dimensional holes and beyond. Persistence Homology is still a very new subject having only become famous around 2014. There is a lot of potential for research in statistics, data analysis, computational geometry, and algebraic topology.

5.2 Reflection

Topology is the subject that got me interested in mathematics. When I was about 12 years old I found a video online about turning a sphere inside out. The video showed a grainy animation of how the process is done without ripping, tearing, or squeezing the sphere. My interests in mathematics wouldn't take off until I took linear algebra. I found the most abstract things amazing in mathematics. At the time I was bouncing between being a computer science major and a math major. When Dr. Dan Yasaki taught us vector spaces, I chose my major right then and there. A year later I would study and do research with Dr. Yu-Min Chung in the field of topological data analysis where I first learned about homology. This subject is the perfect balance between every interest I have had a student of mathematics. It is the blend of topology, linear/abstract algebra, and computer science.

Bibliography

1. Poincaré, Henri. Analysis Situs. Paris, France: Gauthier-Villars, 1895.

2. Euler, Leonhard. "Solutio problematis ad geometriam situs pertinentis." Commentarii Academiae Scientiarum Imperialis Petropolitanae 8 (1736): 128–140.

3. Croom, Fred H. Basic Concepts of Algebraic Topology. New York, Heidelberg, Berlin: Springer, 1978.

4. Edelsbrunner, Herbert, and John L. Harer. Computational Topology: an Introduction. Providence, RI: American Mathematical Society, 2010.

5. Wang, Kairui Glen. "The basic theory of persistent homology." URL http://math.uchicago.edu/ may/REU2012/REUPapers/WangK. pdf (2012).

6. Vieira, Lucas. Wikipedia File: Mug and torus morph.gif, 2007. https://en.wikipedia.org/wiki/File:Mug and Torus morph.gif.

7. Kavi, Lord Clifford. Symmetry Breaking in Graphs and Matroids. Ghana. 2015

8. Friedman, Greg. "An elementary illustrated introduction to simplicial sets." arXiv preprint arXiv:0809.4221 (2008).