On the Incremental Computation of Simplicial Homology
TJHSST Senior Research Project
Computer Systems Lab 2009-2010

Brian Hamrick

April 6, 2010

Abstract

Homology groups are a fundamental algebraic invariant in algebraic topology that allows the discrimination of topological spaces. Methods to compute the homology groups of a simplicial complex are known in general, but they require expensive matrix computations. For specific classes of simplicial complexes, both geometric and incremental methods are known. However, these classes generally preclude the presense of torsion in the homology groups. This paper investigates the possibility of an incremental homology type computation that can account for torsion and thus apply to a larger class of simplicial complexes than previous results.

Keywords: algebraic topology, homology, algorithmic complexity

1 Introduction

Topology is the study of geometric structure on sets. One of the big questions of topology is to determine when are two spaces essentially the same. Algebraic topology is a branch of topology which computes algebraic objects which remain invariant when the underlying structure remains the same. These algebraic invariants include homology groups which encapsulate certain characteristics about the underlying structure of the space. In this paper I consider the homology groups of a certain class of spaces known as simplicial complexes. Algorithms to compute these homology groups are known, but they are generally inefficient. This project will investigate the efficiency of various algorithms to compute these homology groups. Improving the efficiency of such algorithms is applicable in experimental mathematics, where computers can perform computations beyond the reach of human work.

2   Background

2.1   Previous Results

The computation of simplicial homology via reduction of the boundary matrices to Smith normal form is rather classic. For sparse matrices, [1] gives an efficient algorithm to compute the homology groups. This algorithm works in full generality; it will compute the homology of any simplicial complex regardless of factors such as orientability. However, the algorithm presented in [1] is probabilistic and the worst case analysis of the algorithm is unsatisfactory. [4] describes a method by which simplicial complexes embedded in R^3 may be transformed into a homotopically equivalent three-dimensional manifold, for which the homology groups may be computed efficiently via classical results of the structure of such spaces. This method allows deterministic computation of the homology type of a simplicial complex embedded in R^3 in linear time, even faster than the probabilistic methods proposed by [1] for general simplicial complexes. For the remainder of this paper, we investigate extensions of this method to dimensions higher than 3.

2.2   Definitions

In this paper, we deal solely with triangulated spaces, so a general notion of a topological space is unnecessary. A k-simplex is the convex hull of k points in general position. We denote the k-simplex determined by $v_1, v_2, \ldots, v_k$ as $[v_1 v_2 \cdots v_k]$. A k-chain is a formal sum of k-simplices. The orientation of a k-simplex is taken into account in the k-chain such that $[v_1 v_2 \cdots v_k] = -[v_2 v_1 \cdots v_k]$, and similarly for any other transposition of two vertices. The boundary of a k-simplex $[v_1 v_2 \cdots v_k]$ is the $(k-1)$-chain $\sum_{i=1}^{k} (-1)^{i-1} [v_1 \cdots v_{i-1} \hat{v}_i v_{i+1} \cdots v_k]$, where $[v_1 \cdots v_{i-1} \hat{v}_i v_{i+1} \cdots v_k]$ is the $(k-1)$-simplex determined by all the vertices except for v_i. The boundary of a k-chain is the sum of the boundaries of the k-simplices that form the chain.

Definition 2.1. *The cycle groups C_k of a simplicial complex are defined as the set of k-chains with a null boundary. The group operation is addition.*

3   Preliminary Results

3.1   The Abstract Incremental Algorithm

In this section, we will establish an abstract algorithm for incrementally computing the homology type of a simplicial complex. We will work with a given simplicial complex $K = \bigcup_{i=1}^{N} \sigma_i$, where σ_i ranges over all the simplexes of K such that $K_i = \bigcup_{j=1}^{i} \sigma_j$ is a valid simplicial complex for all i. In this incremental algorithm, we will compute the homology type of K_i for all i, starting from $i = 1$. The homology type of K_1 is trivial to compute: $H_n(K_1) = 0$ for all n.

Next, we need to analyze the incremental step. We wish to compute the homology type of $K_i = K_{i-1} \cup \sigma_i$ from the homology type of K_{i-1}. To do this, we will appeal to the following theorem.

Theorem 3.1 (The Mayer-Vietoris Sequence). *Given two subspaces A and B of X whose interiors cover X, the following sequence is exact:*

$$\cdots \to H_{n+1}(X) \overset{\partial_*}{\to} H_n(A \cap B) \overset{(i_*, j_*)}{\to} H_n(A) \oplus H_n(B) \overset{k_* - l_*}{\to} H_n(X) \overset{\partial_*}{\to}$$

$$\overset{\partial_*}{\to} H_{n-1}(A \cap B) \to \cdots \to H_0(A) \oplus H_0(B) \overset{k_* - l_*}{\to} H_0(X) \to 0$$

Given two spaces A and B which cover a space X, the Mayer-Vietoris Sequence relates the homology groups of A, B, $A \cap B$, and X. We'll apply this to our simplicial complex with $A = K_{i-1}$, $B = \sigma_i$, and $X = K_i$. The Mayer-Vietoris Sequence then tells us that the sequence

$$H_n(K_{i-1} \cap \sigma_i) \to H_n(K_{i-1}) \oplus H_n(\sigma_i) \to H_n(K_i) \to H_{n-1}(K_{i-1} \cap \sigma_i) \to H_{n-1}(K_{i-1})$$

is exact for all n.

However, this relation can be simplified significantly. $K_{i-1} \cap \sigma_i = \partial \sigma_i$, the boundary of σ_i, as K_i must be a valid simplicial complex, so the simplices forming $\partial \sigma_i$ must be in K_i, but they are also not σ_i, so they were in K_{i-1} as well while the interior of σ_i was not. Furthermore, σ_i is homotopy equivalent to a single point, so $H_n(\sigma_i) = 0$ for all n. Using these simplifications, we arrive at the following result.

Corollary 3.1. *Given two simplicial complexes K_{i-1} and K_i and a simplex σ_i such that $K_i = K_{i-1} \cup \sigma_i$, the sequence*

$$H_n(\partial \sigma_i) \to H_n(K_{i-1}) \to H_n(K_i) \to H_{n-1}(\partial \sigma_i) \to H_{n-1}(K_{i-1})$$

is exact for all n.

Let us now consider three cases. If σ_i is a k-simplex, we consider $n = k - 1$, $n = k$, and $n \neq k - 1, k$ separately.

First, consider $n \neq k - 1, k$. Then we have $H_n(\partial \sigma_i) = H_{n-1}(\partial \sigma_i) = 0$. Therefore, corollary 3.1 tells us that $0 \to H_n(K_{i-1}) \to H_n(K_i) \to 0$ is exact, and it follows that $H_n(K_{i-1}) \simeq H_n(K_i)$. Since this map from $H_n(K_{i-1}) \to H_n(K_i)$ is induced by the inclusion map $\iota : K_{i-1} \to K_i$, the generating set for $H_n(K_i)$ is the same as the generating set for $H_n(K_{i-1})$.

Second, consider $n = k - 1$. Then we have $H_n(\partial \sigma_i) \simeq \mathbb{Z}$ and $H_{n-1}(\partial \sigma_i) = 0$. Then corollary 3.1 yields that $H_{k-1}(\partial \sigma_i) \overset{\iota_*}{\to} H_{k-1}(K_{i-1}) \to H_{k-1}(K_i) \to 0$ is exact, so $H_{k-1}(K_i) = H_{k-1}(K_{i-1})/\iota_*(H_{k-1}(\partial \sigma_i))$, where ι_* is the map induced by the inclusion $\iota : \partial \sigma_i \hookrightarrow K_{i-1}$.

Finally, consider $n = k$. Then we have $H_n(\partial \sigma_i) = 0$ and $H_{n-1}(\partial \sigma_i) \simeq \mathbb{Z}$. Then corollary 3.1 yields that the sequence $0 \to H_k(K_{i-1}) \to H_k(K_i) \overset{\partial_*}{\to} \mathbb{Z} \overset{\iota_*}{\to} H_{k-1}(K_{i-1})$ is exact. $\partial_*(H_k(K_i))$ is a subgroup of $\mathbb{Z}$, all of which are isomorphic to $\mathbb{Z}$. Let x be an n-cycle that

3

contains σ_i such that the coefficient of σ_i is minimal but positive. Suppose that the coefficient of σ_i in x is a. Then for any k-cycle y, the coefficient of σ_i will be a multiple of a, or else we can subtract an appropriate multiple of x to find a k-cycle with a smaller coefficient of σ_i. Therefore, we can write any k-cycle as $cx + y$, where x is our specific cycle containing σ_i and y is a k-cycle class in K_{i-1}. Then the homoology class of this k-cycle is simply $c[x] + [y]$ where $[x]$ is a new generator and $[y]$ is a homology class in $H_k(K_{i-1})$. All the requisite propeties of the addition operations can be verified to see that $H_k(K_i) \simeq \mathbb{Z} \times H_k(K_{i-1})$ where the new generator is our specific k-cycle x containing σ_i the minimum positive number of times.

Putting these together, we have the following tasks that need to be completed to create a bona fide algorithm for incrementally computing the homology type and generating set of a complex:

- Write the homology class of $\partial \sigma_i$ as a linear combination of the generating set for $H_{k-1}(K_{i-1})$.

- Compute the quotient of $H_{k-1}(K_{i-1})$ by the subgroup generated by the homology class of $\partial \sigma_i$.

- Detect whether a new k-simplex is part of a k-cycle.

- If it exists, compute a k-cycle with the minimal positive coefficient of σ_i.

3.2 The One Dimensional Algorithm

Let us first consider the simple case of a one dimensional finite simplicial complex. The result that we will expect from this algorithm is a number telling us the number of connected components and the number of independent cycles in the complex, as well as a list of those components and independent cycles.

Because our simplex is one dimensional, $H_n(K_i) = 0$ for all $n \geq 2$ and all i. Furthermore, we only need to consider two operations: adding a 0-simplex and adding a 1-simplex.

Assume we are working on the step to compute the homology type and generators for K_i from that of K_{i-1}. Let us first consider the case where σ_i is a 0-simplex. This 0-simplex was not already in our complex, so it is not part of the boundary of any 1-simplex. Therefore, this 0-simplex is independent of everything in $H_0(K_{i-1})$, so we have $H_0(K_i) \simeq H_0(K_{i-1}) \times \mathbb{Z}$, where the new generator is simply the 0-cycle composed of this simplex.

Second, suppose σ_i is a 1-simplex. We know that $\partial \sigma_i$ is of the form $y - x$ for two points x and y. Let us treat σ_i as a path from x to y. There are now two cases:

In the first case, the two bounding points of σ_i are in different connected components (they have distinct homology classes in K_{i-1}). Then we have that $H_0(K_i) \simeq H_0(K_{i-1})/\iota_*(H_0(\partial \sigma_i))$. Therefore, when we mod by the subgroup generated by $[\partial \sigma_i]$, we are simply modding by the relation $[x] \sim [y]$, so the result is that two elements of the generating set, corresponding to $[x]$ and $[y]$, are associated. This simply corresponds to a reduction of the homology group and its generating set. No other elements of the generating set are affected.

In the second case, the two bounding points of σ_i are in the same connected component. Then if we add σ_i to an existing path from y to x, we have a 1-cycle that contains σ_i exactly once, so it is a cycle containing σ_i the minimal positive number of times. We then know that $H_1(K_i) \simeq H_1(K_{i-1}) \times \mathbb{Z}$ where the new generator is our found 1-cycle.

3.3 Implementation Details

Because of the low-dimensionality of this case, the task of writing $\partial \sigma_i$ as a linear combination of the generating set for $H_{k-1}(K_{i-1})$ is rendered moot by the fact that the generators simply correspond to the connected components that the two endpoints lie in. Updating the generator corresponding to each point is a relatively simple task. When we need to merge two generators, we choose either one of them and mark all of the 0-simplices in its connected component as belonging to the other connected component. In this way, each connected component retains a canonical name, so that we may easily find all the 0-simplices in any given component.

Therefore, the only remaining task is that of finding the path from y to x in the last step of the above section. To do this, we will use a floodfill algorithm. Breadth-first search is used so that the shortest path is found and used. Because of the low-dimensionality, these cycles are never modified in the generating set for the first homology group, so we use the shortest path to simplify the output.

Each of these two operations is at most linear in the number of simplices, so this algorithm takes at most quadratic time in the number of simplices.

4 The Main Result

In this section we present an algorithm for incrementally computing the homology type and generating sets for the homology groups of a 2-dimensional simplicial complex where the coboundary of each 1-simplex consists of at most two 2-simplices. We follow the same incremental method based on the Mayer-Vietoris sequence as before, making use of a few key lemmas.

Lemma 4.1. *If a 2-simplex is part of a 2-cycle, then there exists a 2-cycle containing that 2-simplex exactly once.*

Proof. We will call two 2-simplices *neighbors* if they share a bounding 1-simplex. Two 2-simplices σ and τ are *connected* if there is a sequence of 2-simplices $\sigma = \sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_l = \tau$ such that σ_i and σ_{i+1} are neighbors for all i. Let x be a 2-cycle containing a given 2-simplex σ with a coefficient of c. Let y be the set of simplices in x that are connected to σ. y is clearly a 2-cycle. I claim that all of the simplices in y must have the same coefficient up to negation.

Let c denote the coefficient of σ in y. For any simplex τ in y, consider a sequence of 2-simplices $\sigma = \sigma_1, \sigma_2, \ldots, \sigma_l = \tau$. We will induct on i to show that the coefficient of σ_i is the same as that of σ. The base case is clear, so we will only show the inductive step. For

a given i, consider the 1-simplex that is part of the boundary of both σ_i and σ_{i+1}. Call this 1-simplex v_i.

Because y is a 2-cycle, it must have no boundary. However, the only contributions of v_i to the boundary are from σ_i and σ_{i+1} by assumption. Since the contribution of v_i from σ_i is equal to either the coefficient of σ_i or its negation. Similarly, the contribution from σ_{i+1} is equal to either the coefficient of σ_{i+1} or its negation. Since these two contributions must cancel out, the coefficients of σ_i and σ_{i+1} must be either the same or negations of each other. This completes the inductive step, so the coefficient of each σ_i is the same as that of σ, up to negation.

Therefore, if c is the coefficient of σ in y, y/c is also a 2-cycle and contains σ exactly once, so it is the desired cycle. $\qquad\square$

4.1 The Two Dimensional Algorithm

As before, we begin with a sequence of simplicial complexes $K_0, K_1, K_2, \ldots$ such that $K_i = \bigcup_{n=0}^{i} \sigma_n = K_{i-1} \cup \sigma_i$ for all i. For each simplex K_i we will compute the following:

- An independent generating set for each of the homology groups of K_i.

- The torsion coefficients corresponding to each of the above generators.

- A cycle containing σ_i exactly once, if one exists. We will call this cycle S_i.

- The representation of each of the generators for $H_n(K_{i-1})$ in terms of the generators of $H_n(K_i)$.

In this algorithm, we will internally represent the generators as linear combinations of the S_i.

For the incremental step, we will look at three situations.

4.1.1 Adding a 0-simplex

If σ_i is a 0-simplex, then we know that σ_i is a 0-cycle by itself, so $H_n(K_i) \simeq H_n(K_{i+1})$ for $n \geq 1$ and $H_0(K_i) \simeq H_0(K_{i-1}) \times \mathbb{Z}$, where the new generator is $S_i = \sigma_i$.

4.1.2 Adding a 1-simplex

If σ_i is a 1-simplex, then we will proceed almost identically to the one dimensional algorithm. If the two vertices of σ_i are in different components, then we reduce the generating set for H_0 by identifying the two generators corresponding to the two endpoints. If the two vertices of σ_i are in the same component, then we use a breadth-first search to find a cycle containing σ_i exactly once. This cycle becomes S_i and we update the generating set for H_1 by augmenting S_i with a torsion coefficient of 0.

4.1.3 Adding a 2-simplex

If σ_i is a 2-simplex, then the first order of business is to determine whether it is part of a 2-cycle or not. If σ_i is part of a 2-cycle, then we know that $H_2(K_i) \simeq H_2(K_{i-1}) \times \mathbb{Z}$, where the new generator is C_i, the 2-cycle that we found from before. If σ_i is not part of a 2-cycle then we write $\partial\sigma_i$ as a linear combination of the generators for $H_1(K_{i-1})$ and then we have that $H_1(K_i) \simeq H_1(K_{i-1})/\langle[\partial\sigma_i]\rangle$.

4.2 Implementation Details

Definition 4.1. *A* fundamental cycle S_i *is a cycle of any dimension that contains* σ_i *exactly once and for any other simplex* $\sigma_j \in S_i$, *we have* $j < i$.

Theorem 4.1. *The fundamental cycles* $\{S_i\}_{i \leq n}$ *form generating sets for the cycle groups* C_k *of the simplicial complex* K_n.

Proof. We induct on n. The base case, $n = 0$, is trivial as K_0 is empty. Now suppose that the fundamental cycles $\{S_i\}_{i \leq n}$ form generating sets for the cycle groups of K_n. Given a cycle Z in K_{n+1}, consider the coefficient c of σ_{n+1}. If there is a cycle containing σ_{n+1}, then we know that S_{n+1} exists and contains σ_{n+1} exactly once. Then $Z - cS_{n+1}$ is a cycle completely contained in K_n, and thus can be written as a linear combination of $\{S_i\}_{i \leq n}$. By adding cS_{n+1} to this representation, we obtain a representation of Z as a linear combination of $\{S_i\}_{i \leq n+1}$, as desired. This completes our induction. $\square$

Throughout the implementation of this algorithm, all cycles are kept as linear combinations of either simplices, fundamental cycles, or current generators for the homology groups. To translate between these, we keep a list of representations of each of the fundamental cycles S_i in terms of the generators G_i and use the theorem presented above to create representations in terms of the fundamental cycles from representations in terms of simplices.

To determine whether a 2-simplex is part of a cycle, a floodfill is used. Lemma 4.1 tells us that the only possible cycle containing a given 2-simplex is exactly the sum of the 2-simplices that are connected to the given one. Computation of this sum is implemented using a floodfill algorithm and the boundary of the sum determines whether or not the simplex is in fact part of a cycle or not.

If not, the boundary of the new simplex is written as a linear combination of fundamental cycles via the following reduction algorithm. At each step, find the last index such that the corresponding simplex has a nonzero coefficient. Call the index i, so that the corresponding term is $c_i\sigma_i$. Then we add c_iS_i to our representation and subtract c_iS_i from our cycle, so all the remaining terms are only composed of indices strictly smaller than i. After at most n steps, this process will reduce any cycle in K_n to the trivial cycle, at which point we have finished constructing the desired representation.

Finally, when computing the quotient of a homology group by a new relation, an elementary Smith Normal Form reduction algorithm is used so that we have a sequence of row and column operations that reduce the natrix to Smith Normal Form. Row operations simply

compose the relations in different ways, so they do not change the generators. However, when applying a column operation, we are taking two terms in each relation $c_i G_i + c_j G_j$ and replacing them with $c_i' G_i' + c_j' G_j'$. The column operations that we perform are made such that $c_i' = \gcd(c_i, c_j)$.

By Bezout's Identity, we have integers x and y such that $xc_i + yc_j = c_i'$. Letting $\alpha = \frac{c_i}{c_i'}, \beta = \frac{c_j}{c_i'}$, we get that we want to carry out the following transformation: $c_i' = xc_i + yc_j$, $c_j' = -\beta c_i + \alpha c_j$, $G_i' = \alpha G_i + \beta G_j$, $G_j' = -yG_i + xG_j$. When we apply these transformations, we see that $c_i' G_i' + c_j' G_j' = (xc_i + yc_j)(\alpha G_i + \beta G_j) + (-\beta c_i + \alpha c_j)(-yG_i + xG_j) = (xc_i\alpha + yc_j\alpha + y\beta c_i - y\alpha c_j)G_i + (x\beta c_i + y\beta c_j - x\beta c_i + x\alpha c_j)G_j = c_i G_i + c_j G_j$. Therefore, by applying these transformations to the generators and the representations of the fundamental cycles, we obtain new generators and new representations. The integers x and y are constructed via the Extended Euclidean Algorithm.

References

[1] Bruce Randall Donald, David Renpan Chang, "On the complexity of computing the homology type of a triangulation", *Proceedings of the 32nd annual symposium on Foundations of computer science*, p.650-661, September 1991, San Juan, Puerto Rico.

[2] Cecil Jose A. Delfinado, Herbert Edelsbrunner. "An incremental algorithm for Betti numbers of simplicial complexes", *Proceedings of the ninth annual symposium on Computational geometry*, p.232-239, May 18-21, 1993, San Diego, California, United States

[3] Hatcher, Allen. "Algebraic Topology". Cambridge University Press, Cambridge, 2002.

[4] Tamal K. Dey and Sumanta Guha. "Computing Homology Groups of Simplicial Complexes in R^3", *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, pp. 398-407, May 22-24, 1996, Philadelphia, Pennsylvania, United States.

[5] Patrick J. Morandi. "The Smith Normal Form of a Matrix", Unpublished, February 17, 2005.