Persistent Relative Homology Using Matrix Factorization

Christian Lentz & Xintan Xia, Macalester College

Advised by Prof. Lori Ziegelmeier & Greg Henselman-Petrusek

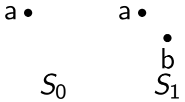
September 2023

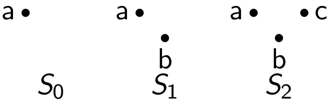
MACALESTER

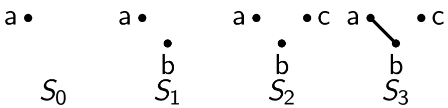
A MOTIVATING EXAMPLE

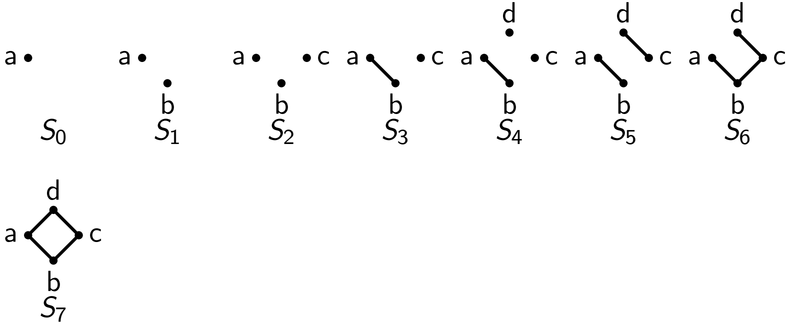
a •

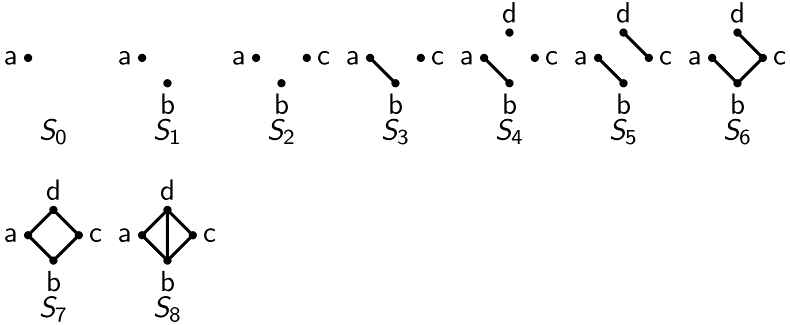
S_0

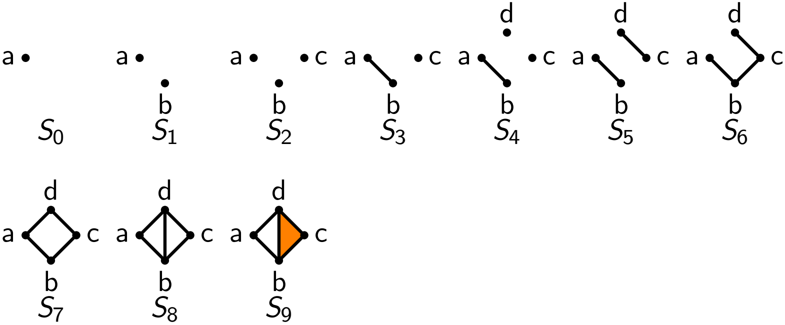

a
a
b
S_0
S_1

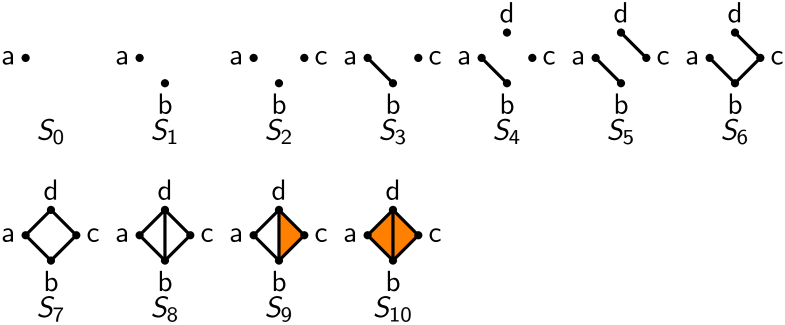

a
a
b
a
c
b
S_0
S_1
S_2

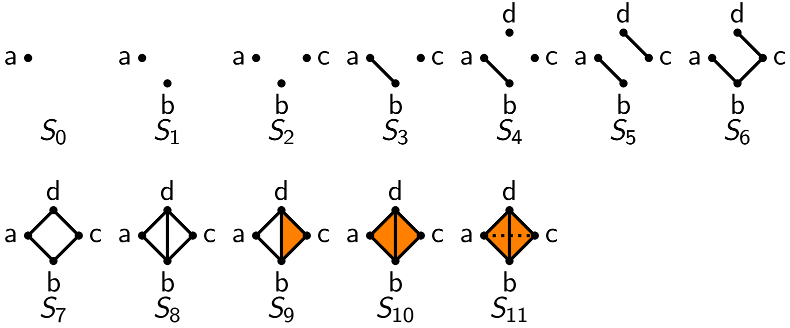

a
a
b
a
c
b
a
c
b
S_0
S_1
S_2
S_3


a
S_0
a
b
S_1
a
c
b
S_2
a
c
b
S_3
d
a
c
b
S_4
d
a
c
b
S_5
d
a
c
b
S_6
d
a
c
b
S_7

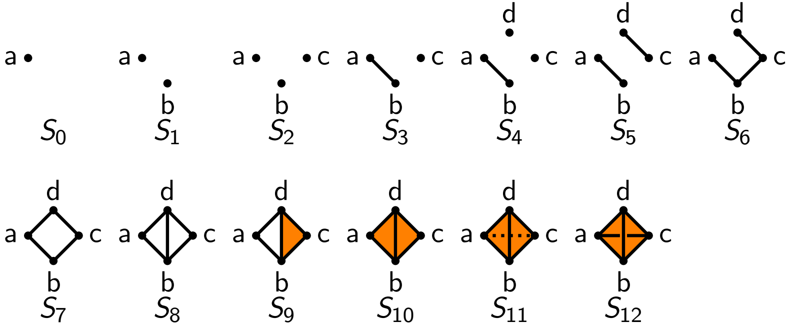

a
S_0
a b
S_1
a b c
S_2
a b c
S_3
a b c d
S_4
a b c d
S_5
a b c d
S_6
a b c d
S_7
a b c d
S_8

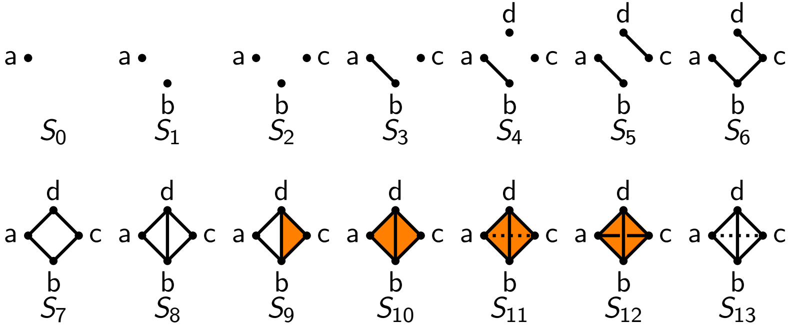

a
b
c
d
S0
S1
S2
S3
S4
S5
S6
S7
S8
S9


a
S_0
a
b
S_1
a
c
b
S_2
a
c
b
S_3
d
a
c
b
S_4
d
a
c
b
S_5
d
a
c
b
S_6
d
a
c
b
S_7
d
a
c
b
S_8
d
a
c
b
S_9
d
a
c
b
S_{10}



a
b
c
d
S_0
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8
S_9
S_{10}
S_{11}


a
b
c
d
S_0
S_1
S_2
S_3
S_4
S_5
S_6
S_7
S_8
S_9
S_{10}
S_{11}
S_{12}


a
a b
a b c
a b c
a b c d
a b c d
a b c d
a b c d
S_0 S_1 S_2 S_3 S_4 S_5 S_6
S_7 S_8 S_9 S_10 S_11 S_12 S_13

- There are a few natural questions at this point:
    - How do we classify holes?
    - Can we take an algorithmic approach?
    - Can we extend this to more complex settings?

BASICS OF HOMOLOGY

Definition

We define the **boundary operator** to be the linear transformation between **chain groups** $\partial_n : C_n(K) \to C_{n-1}(K)$ such that the basis vector $[v_0 v_1 \dots v_n]$ is mapped as follows

$$[v_0 v_1 \dots v_n] \mapsto \sum_{j=0}^{n} (-1)^j [v_0 v_1 \dots \hat{v}_j \dots v_n]$$

where $\hat{v}_j$ is removed.

Example

Definition

A **Chain Complex** is a sequence of vector spaces, in this case chain groups, connected by a sequence of linear transformations such that the composition $\partial_{n-1}\partial_n = 0$.

$$\mathcal{C}_\bullet(K) = \cdots C_{n+1}(K) \xrightarrow{\partial_{n+1}} C_n(K) \xrightarrow{\partial_n} C_{n-1}(K) \xrightarrow{\partial_{n\text{-}1}} \cdots \xrightarrow{\partial_2} C_1(K) \xrightarrow{\partial_1} C_0(K) \xrightarrow{\partial_0} 0.$$

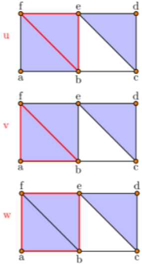
- Since $\partial_{n-1}\partial_n = 0$, then a "boundary has no boundary".
- **Notice**: $Im(\partial_i) \subseteq Ker(\partial_{i-1})$.
- Inclusion of these spaces leads to the idea of basis matching.

Definition

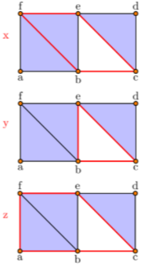
The n^{th} homology group for a simplicial complex K is the quotient vector space,

$$H_n(K) = Ker(\partial_n)/Im(\partial_{n+1}),$$

where $Ker(\partial_n)$ denotes all n-cycles and $Im(\partial_{n+1})$ n-boundaries.



Boundary, and Cycles



Cycles

BASIS MATCHING

- Want a stronger condition than $Im(\partial_{i+1}) \subseteq Ker(\partial_i)$
- Any chain complex which is a **basis matching complex** will satisfy that a set of basis vectors for $Im(\partial_i)$ is a subset of a set of basis vectors for $Ker(\partial_{i-1})$.

Theorem

Every (filtered) chain complex has an associated (filtered) basis matching complex.

- Want an algorithmic process to match the bases of a chain complex

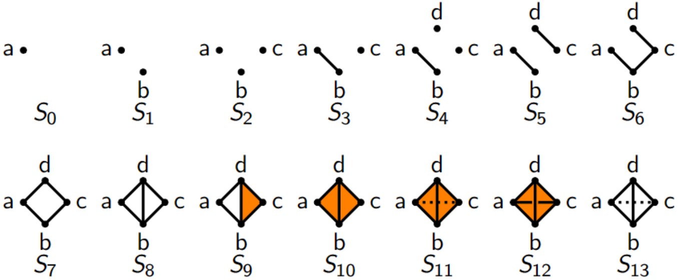
PERSISTENT HOMOLOGY

Definition

A **filtered topological space S** is a nested sequence of sets filtered by some scale parameter

$$S_1 \subseteq S_2 \subseteq \cdots \subseteq S_k = \mathbf{S}$$

and has natural inclusion map $f : S_n \hookrightarrow S_{n+1}$. **Persistent homology** tracks topological features through a filtration.



MATRIX FACTORIZATION

LR DECOMPOSITION

- Starting with $n = 1$, we reduce the boundary matrices of a chain complex, using the following block matrix:

$$A = \begin{pmatrix} \partial_n & I_{c_{n-1}} \\ I_n & 0 \end{pmatrix} \longrightarrow A' = \begin{pmatrix} \partial_n' & L_n \\ R_n & 0 \end{pmatrix}$$

- $I_{C_{n-1}}$ records row operations and I_{C_n} records column operations. We call these L_n and R_n respectively
- Repeat reductions starting with the new block matrix

$$A = \begin{pmatrix} (R_n^{-1})\partial_{n+1} & I_{C_n} \\ I_{C_{n+1}} & 0 \end{pmatrix}$$

- Once all boundary maps have been reduced, we can construct the following products:

$$\hat{\partial}_n = (L_n R_{n-1}^{-1})(\partial_n)(R_n L_{n+1}^{-1}) = \hat{L}_n \partial_n \hat{R}_n$$

- In particular, $\hat{R}_n = (\hat{L}_{n+1})^{-1}$. This is the **matched basis** $\mathcal{B}_n$.

Theorem

The columns of $\mathcal{B}_n$ contain a basis for $Im(\partial_{n+1})$ and $Ker(\partial_n)$.

- We can easily extend this for a method for persistent homology.

U-MATCH DECOMPOSITION

- A tuple of matrices (R, M, D, C) which satisfy the following three conditions:
  - RM = DC
  - M is a **generalized matching matrix**
  - R and C are both upper triangular and invertible
- If we assume that D is square and $D^2 = 0$, then

Theorem

Columns of R contain a basis for $Im(D)$ and columns of C contain a basis for $Ker(D)$. Specific columns depend on pivot locations in M.

- LR Decomposition is a special case of U-Match, however U-Match is more flexible.

PERSISTENT RELATIVE HOMOLOGY

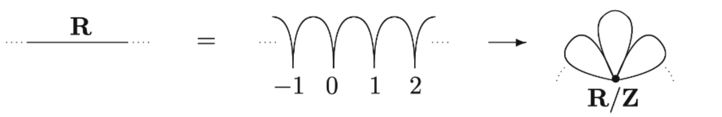
Definition

Let X be a topological space with $A \subseteq X$. We define the **Quotient Space** to be

$$X/A = (X \setminus A) \sqcup *$$

Where $*$ is a single point

Example (The Infinite Bouquet)



- What does the homology of a quotient spacce tell us?

- Suppose simplicial complexes K and K_0 such that $K_0 \subset K$
- **Relative Homology** is the homology of the quotient space K/K_0 ... "how does the homology of a space depend on the homology of some subset of the space?"

Definition

We define a **Relative n-Cycle** to be any n-chain $\alpha \in C_n(K)$ such that $\partial_n(\alpha) \in C_{n-1}(K_0)$.

Definition

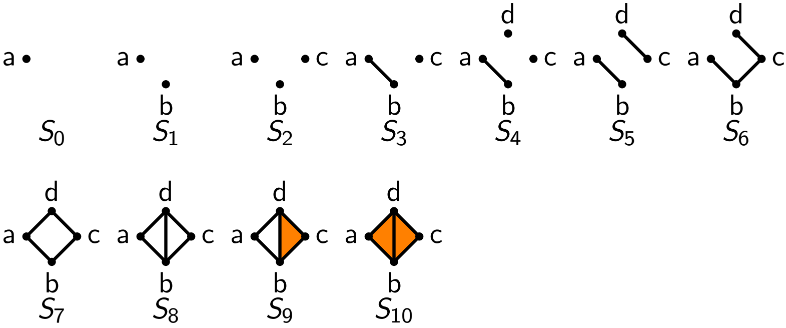
We define a **Relative n-Boundary** to be any relative n-cycle $\alpha = \partial_{n+1}(\beta) + \gamma$ for some $\beta \in C_{n+1}(K)$ and $\gamma \in C_n(K_0)$.

PERSISTENT RELATIVE HOMOLOGY

- Goal: Track the relative homology through a filtration of K. The filtration on K_0 does not necessary follow that of K

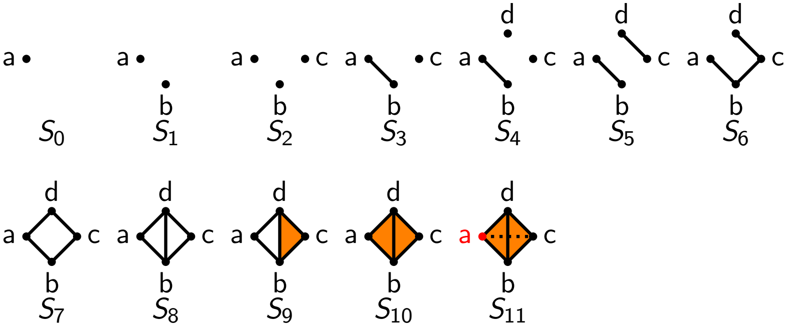
PERSISTENT RELATIVE HOMOLOGY

- Goal: Track the relative homology through a filtration of K. The filtration on K_0 does not necessary follow that of K



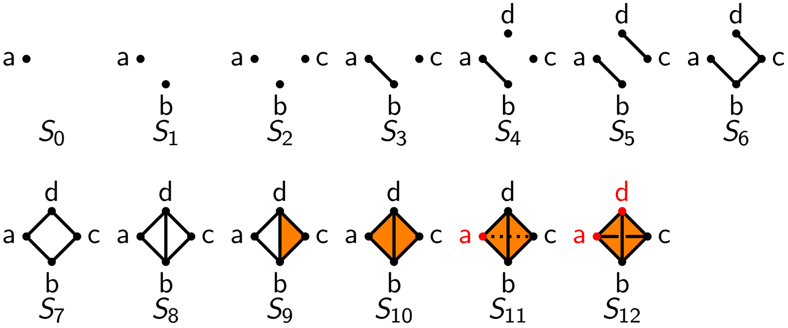
PERSISTENT RELATIVE HOMOLOGY

- Goal: Track the relative homology through a filtration of K. The filtration on K_0 does not necessary follow that of K

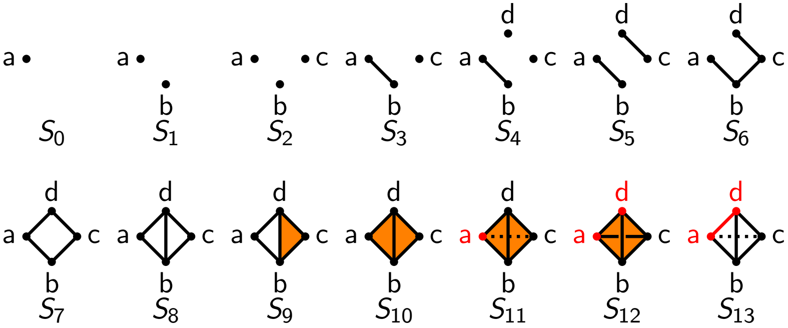


PERSISTENT RELATIVE HOMOLOGY

- Goal: Track the relative homology through a filtration of K. The filtration on K_0 does not necessary follow that of K



- Goal: Track the relative homology through a filtration of K. The filtration on K_0 does not necessary follow that of K



PERSISTENT RELATIVE HOMOLOGY

- Steps to find bases for (persistent) relative homology:
  - Construct the block boundary matrix, D, for K
  - Reorder rows of D to respect filtration on K_0
  - Perform U-Match Decomposition of D to get $RM = DC$

Theorem

Define the set $I = \{c \in COL_{\overline{r_\bullet}}(C) : D(c) \in K_0, D(c) \neq 0\}$. Then $COL_I(C) \cup COL_{\overline{c_\bullet}}(C)$ gives a basis for $\overline{Ker}(D)$.

Theorem

Let I be the set of indices corresponding to the first i columns of R. The subset of the columns of R given by $COL_{r_\bullet}(R) \cup COL_I(R)$ give a basis for $\overline{Im}(D)$.

FUTURE WORK

- This is a preliminary report, and no papers have yet been published.
- We are working to implement this algorithm in the open source project Open Applied Topology (OAT) in the Rust programming language.
- We have conjectured how to “zip” the relative kernel and image into the same basis using U-Match, but have not yet proven this.
- Potential application: knowledge networks and the “science of science” (see the 2:20 talk *Topological Data Analysis of Knowledge Networks*!)

THANK YOU!

Any questions, contact one of us!

- lziegel1@macalester.edu
- clentz@macalester.edu
- xxia@macalester.edu