Barcodes and Layers

Maria Walch

July 2022

1 Introduction

Interpretability: given an interval $\mathcal{L} \in \text{Barcode}_n(X)$, how do we udnerstand it in terms of underlying data. Inverse problem: construct geometric strucutres corresponding to each persistent interval in the original input data. Cycle representatives are not uniquely defined.

Definition 1.1 (Homology, Cycle Representatives)**.** The n-th homology group of a simplicial complex K is defined as the quotient:

$$\mathbf{H}_n(K) := \mathbf{Z}_n(K)/\mathbf{B}_n(K). \tag{1}$$

Elements of $\mathbf{H}_n(K)$ are cosets of the form $[\mathbf{z}] = \{\mathbf{z}' \in \mathbf{Z}_(K) | \mathbf{z}' - \mathbf{z} \in \mathbf{B}_n(K)\}$. An element $h \in \mathbf{H}_n(K)$ is called an n-dimensional homology class. We say that a cycle $\mathbf{z} \in \mathbf{Z}_n(K)$ represents h or that $\mathbf{z}$ is a cycle representative of h if $h = [\mathbf{z}]$. We say that $\mathbf{z}$ and $\mathbf{z}'$ are homologous if $[\mathbf{z}] = [\mathbf{z}']$.

Definition 1.2 (Betti Numbers, Cycle Bases)**.** A (dimension-n) homological cycle basis for $\mathbf{H}_n(K)$ is a set of cycles $\mathcal{B} = \{\mathbf{z}^1, ..., \mathbf{z}^m\}$ such that $[\mathbf{z}^i] \neq [\mathbf{z}^j]$ when $i \neq j$ and $\{[\mathbf{z}^1], ..., [\mathbf{z}^m]\}$ is a basis for $\mathbf{H}_n(K)$. Modulo boundaries, every n-cycle can be expressed as a unique linear combination in $\mathcal{B}$.
The rank of the nth homology group

$$\beta_n(K) := \dim(\mathbf{H}_n(K)) = \dim(\mathbf{Z}_n(K)) - \dim(\mathbf{B}_n(K)), \tag{2}$$

is the called the nth Betti number of K and quantifies the number of independent holes in K.

A (dimension-n) persistent homology cycle basis is a subset $\mathcal{B} \subseteq \mathbf{Z}_n(K)$ with the following two properties:

1. Each $\mathbf{z} \in \mathcal{B}$ has a nonempty lifespan interval.
2. For each $i \in \{1, ..., T\}$ the set

   $$\mathcal{B}_{\epsilon_i} := \{\mathbf{z} \in \mathcal{B} : \epsilon_i \in \mathcal{L}(\mathbf{z})\} \tag{3}$$

   is a homological cycle basis for $\mathbf{H}_n(K_{\epsilon_i})$.

Every filtration of simplicial complexes $(K_{\epsilon_i})_{i \in \{1,\dots,T\}}$ admits a persistent homological cycle basis $\mathcal{B}$. Moreover it can be shown that the multiset of lifespan intervals (one for each basis vector), called the dimension-n barcode of $K_\bullet$,

$$\text{Barcode}_n = \{\mathcal{L}(\mathbf{z}) : \mathbf{z} \in \mathcal{B}\} \tag{4}$$

is invariant over all possible choices of persistent homological cycle bases of $\mathcal{B}$.