Topological Data Analysis

14 November 2019

3 Persistence in homology

3.1 Betti numbers

For a finite ordered abstract simplicial complex K, the *Betti numbers* of K are defined as

$$\beta_n(K) = \operatorname{rank} H_n(K;\mathbb{Z}) = \dim_{\mathbb{Q}} H_n(K;\mathbb{Q})$$

for $n \geq 0$. More generally, for any field $\mathbb{F}$, the nth Betti number of K with coefficients in $\mathbb{F}$ is defined as

$$\beta_n(K;\mathbb{F}) = \dim_{\mathbb{F}} H_n(K;\mathbb{F}).$$

If we denote by D_n the matrix of the boundary operator $\partial_n \colon C_n(K;\mathbb{F}) \to C_{n-1}(K;\mathbb{F})$ in any bases, then, for all n,

$$\beta_n(K;\mathbb{F}) = \dim_{\mathbb{F}} \operatorname{Ker}(\partial_n) - \dim_{\mathbb{F}} \operatorname{Im}(\partial_{n+1}) = F_n - \operatorname{rank} D_n - \operatorname{rank} D_{n+1},$$

where F_n denotes the number of n-faces of K.

Induction shows that, if K has dimension N, then

$$\sum_{n=0}^{N} (-1)^n \, \beta_n(K) = \sum_{n=0}^{N} (-1)^n \, F_n. \tag{3.1}$$

The integer given by (3.1) is called the *Euler characteristic* of K.

3.2 Persistence

A *filtration* of an abstract simplicial complex K is a finite nested sequence of subcomplexes of K that ends with K:

$$K_0 \subseteq K_1 \subseteq \cdots \subseteq K_{m-1} \subseteq K_m = K. \tag{3.2}$$

Our main instances will be the sequence of distinct Vietoris–Rips complexes or Čech complexes of a point cloud X, in which case $K_0 = X$ and $|K| = \Delta^N$ if X has cardinality $N+1$.

Fix any field $\mathbb{F}$ (by default, we will use $\mathbb{F} = \mathbb{Q}$). Given a filtration (3.2) of a finite ordered complex, for all $i, j \in \{0, \ldots, m\}$ with $i \leq j$ and each $n \geq 0$, the inclusion $K_i \hookrightarrow K_j$ induces an $\mathbb{F}$-linear map

$$\varphi_n^{i,j} \colon H_n(K_i;\mathbb{F}) \longrightarrow H_n(K_j;\mathbb{F}).$$

A homology class $\alpha \in H_n(K_j;\mathbb{F})$ is said to be *born* at K_j if it does not belong to the image of $\varphi_n^{i,j}$ for any $i < j$, and a class $\alpha \in H_n(K_i;\mathbb{F})$ *dies* at K_j for $j > i$ if

$\varphi_n^{i,j}(\alpha) = 0$ but $\varphi_n^{i,j-1}(\alpha) \neq 0$. If α is born at K_i and dies at K_j, then the *persistence* of α is defined to be $j - i$.

The image of $\varphi_n^{i,j}$ is an $\mathbb{F}$-subspace of $H_n(K_j; \mathbb{F})$ which is called a *persistent homology group* and denoted by $H_n^{i,j}(K; \mathbb{F})$. It contains those homology classes that are born at or before K_i and survive at least until K_j. The classes that survive until K are called *essential*.

We denote

$$\beta_n^{i,j}(K; \mathbb{F}) = \dim_{\mathbb{F}} H_n^{i,j}(K; \mathbb{F})$$

and call them *persistent Betti numbers* with respect to the filtration (3.2) with coefficients in $\mathbb{F}$. We usually write $\beta_n^{i,j}(K)$ instead of $\beta_n^{i,j}(K; \mathbb{Q})$.

3.3 Barcodes for filtered complexes

Suppose given a filtration of a finite ordered abstract simplicial complex K,

$$K_0 \subseteq K_1 \subseteq \cdots \subseteq K_{m-1} \subseteq K_m = K.$$

The persistence of homology classes can be depicted by means of a *barcode*, which is a collection of horizontal line segments in a plane coordinate system whose x-axis contains $\{0, \ldots, m\}$ and whose y-axis marks the levels of an ordered sequence of homology generators for H_0, H_1, H_2, etc. Homology will be meant with coefficients in $\mathbb{Z}$ (or equivalently $\mathbb{Q}$ if a field is wanted), unless otherwise specified. If a homology class α is born at K_i and dies at K_j, then a segment from i to j will be drawn. We shall use the convention that longer segments are drawn below shorter ones, and those starting later appear above those starting earlier.

Short exercise

(1) Draw a barcode for the persistent homology with $\mathbb{Z}$ coefficients for the Vietoris–Rips filtration of the following point cloud in $\mathbb{R}^2$:

$$X = \{(0,0),\ (1,2),\ (2,1),\ (4,3),\ (4,-3)\}.$$