Persistent homology

I. Topological data analysis

Setting: data given as finite set $S \subset \mathbb{R}^n$, n large
$\rightsquigarrow$ distance function on S

Goal: Analyze geometry of S

Why topology?

1. Qualitative information is needed ($\rightsquigarrow$ topology ~~is~~ is "qualitative geometry")
2. Choice of ~~[illegible]~~ metric not theoretically justified
3. Choice of coordinates is not justified

$\rightsquigarrow$ topological results are quite stable under different choices

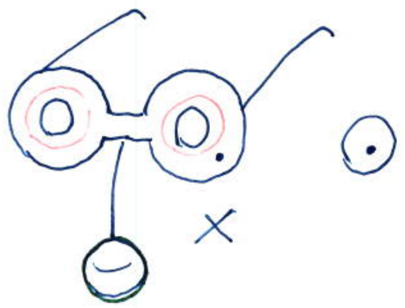
II. Intuition about homology

X top. space, k field, $n \in \mathbb{N}$.

$\rightsquigarrow H_n(X;k)$ k-vector space
"n-th homology with coefficients in k"

$\beta_n = \dim H_n(X;k)$ n-th Betti number

β_n measures "number of n-dim. loops in X"



β_0 = number of components = 2

$\beta_1 = 2$

$\beta_2 = 1$

$\beta_3 = \beta_4 = \ldots = 0$

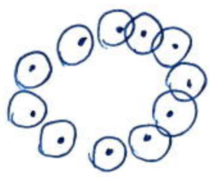
$\mathbb{R}P^2 = S^2/_{x \sim -x}$

$\beta_{1,2}(\mathbb{R}P^2, \mathbb{Z}_2) = 1$

$\beta_{1,2}(\mathbb{R}P^2, \mathbb{Z}_p) = 0$, p prime

III. Idea of persistence

Topology of finite set $S \subset \mathbb{R}^n$ is discrete



$\rightsquigarrow$ replace S by $B_\varepsilon(S) = \bigcup_{p \in S} B_\varepsilon(p)$, $\varepsilon > 0$

Which choice of ε?
- ε small $\rightsquigarrow$ no new information
- ε big $\rightsquigarrow$ gigantic clusters

Solution:

For $\varepsilon \leq \tilde{\varepsilon}$ we have $B_\varepsilon(S) \overset{\iota_{\varepsilon\tilde{\varepsilon}}}{\hookrightarrow} B_{\tilde{\varepsilon}}(S)$

Aside: homology is functorial:

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "g"'] & Y \arrow[dl, "h"] \arrow[d, "1"] \\ Z & Y \end{tikzcd} \rightsquigarrow \begin{tikzcd} H_n(X;k) \arrow[r, "f_*"] \arrow[d] & H_n(Y;k) \arrow[dl] \arrow[d, "1"] \\ H_n(Z;k) & H_n(Y;k) \end{tikzcd}$$

$g = h \circ f \Rightarrow g_* = h_* \circ f_*$

$g \simeq h \Rightarrow g_* = h_*$

$\rightsquigarrow H_n(B_\varepsilon(S);k) \xrightarrow{\iota_{\varepsilon\tilde{\varepsilon}} = (\iota_{\varepsilon\tilde{\varepsilon}})} H_n(B_{\tilde{\varepsilon}}(S);k)$

$\quad !!$

$H_n^\varepsilon(S;k)$

$\Rightarrow H_n^*(S;k)$ is a $\mathbb{R}^+$- persistent k-vector space in the sense of IV

IV. Persistent objects

$(I, \leq)$ partially ordered set, C category (e.g. $C = k\text{-Vect}$)

An I-persistent C-object K is a functor $I \to C$

- a ~~collection~~ family $(K_i)_{i\in I}$ of C-objects
- for $i \leq \tilde{\imath}$ a morphism $\varphi^K_{i\tilde{\imath}} : K_i \to K_{\tilde{\imath}}$ s.t.

  for $i \leq \tilde{\imath} \leq \hat{\imath}$ $\quad \varphi_{\tilde{\imath}\hat{\imath}} \varphi_{i\tilde{\imath}} = \varphi_{i\hat{\imath}}$

Theorem: (Classification of $\mathbb{N}$-pers. vector spaces)

Let $a \in \mathbb{N}$, $b \in \mathbb{N} \cup \{\infty\}$, $a \leq b$,

$k[a,b]$ is the $\mathbb{N}$-pers. k-vect. sp. defined by

$$k[a,b]_n = \begin{cases} k, & a \leq n \leq b \\ 0, & \text{else} \end{cases} \qquad \varphi^{k[a,b]}_{nm} = \begin{cases} 1 & \text{if possible} \\ 0 & \text{else} \end{cases}$$

$$0 \to 0 \to \dots \to k \xrightarrow{\mathbb{1}} k \xrightarrow{\mathbb{1}} \dots \xrightarrow{\mathbb{1}} k \xrightarrow{0} 0 \to 0 \to \dots$$

Let V be some $\mathbb{N}$-pers. k-vect. space

Then there exists a family $B_V : I \to \mathbb{N} \times (\mathbb{N} \cup \{\infty\})$, $i \mapsto (a_i, b_i)$ s.t.

some index set

$$V = \bigoplus_{i \in I} k[a_i, b_i]$$

Furthermore, the cardinality of $B_V^{-1}(a_i, b_i)$ is unique for every $(a_i, b_i) \in \mathbb{N} \times (\mathbb{N} \cup \{\infty\})$.

B_V is called a barcode of V.