Applied Topology

Persistent Homology and Sliding Window Embedding

Yi Siheng

faculty of Mathematics
Southern University of Science and Technology

Mar. 14, 2022

Summary

1. Persistent Homology
2. Example: Natural Image
3. Sliding Window Embedding

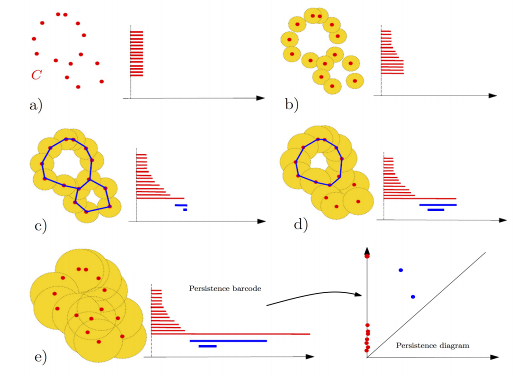
Persistent Homology

What is the Persistent Homology

- Persistent homology is a powerful tool to compute, study and encode efficiently multiscale topological features of nested families of simplicial complexes and topological spaces.

What is the Persistent Homology

- Sampling → Constructing filtration of simplicial complexes → computing homological group of complexes → Product the Barcode and Diagram



Why do we want to use persistent homology

- On some questions, such as when studying protein molecules, we need softer "scales" and "measures".

Figure: Doughnut and Teacup

What information does persistent homology characterize

- Loosely speaking, the persistent homology describes the information of holes in topological spaces, that is, how many holes there are and what kind of holes there are.
- Klein bottle and Torus

$$H_k(\mathbb{K},\mathbb{Z}) = \begin{cases} \mathbb{Z}\oplus\mathbb{Z}/2, & k=0,1 \\ \mathbb{Z}, & k=2 \\ 0, & otherwise \end{cases}$$

$$H_k(\mathbb{T},\mathbb{Z}) = \begin{cases} \mathbb{Z}\oplus\mathbb{Z}, & k=0,1 \\ \mathbb{Z}, & k=2 \\ 0, & otherwise \end{cases}$$

Constructing the Filtration

- Cech Complex
  S is a finite set of points in $\mathbb{R}^d$ and $B_x(r) = x + r\mathbb{B}^d$

Definition

Cech complex of S and r is the nerve of this collection of balls

$$Cech(r) = \{\sigma \subseteq S | \cap_{x\in\sigma} B_x(r) \neq \emptyset\}$$

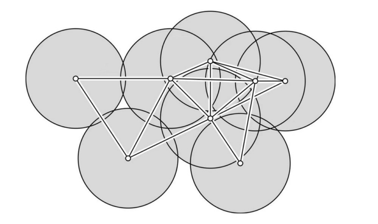


Figure: Cech Complex

Computing Barcode and Diagram

Suppose the filtration

$$\emptyset \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_n = K$$

computing homotopy of K_i

$$0 = H_p(K_0) \to H_p(K_1) \to H_p(K_2) \to \cdots \to H_p(K_n) = H_p(K)$$

What we want is the birth and death of the generators of homology groups

Definition

the p-th persistent homology groups $H_p^{i,j} = im\ f_p^{i,j}$
the p-th persistent Betti numbers $\beta_p^{i,j} = rank(H_p^{i,j})$

$$\mu_p^{i,j} = \beta_p^{i,j} = rank((\beta_p^{i,j-1} - \beta_p^{i,j}) - (\beta_p^{i-1,j-1} - \beta_p^{i-1,j}))$$

Example

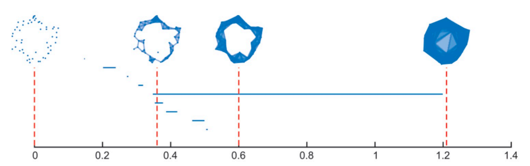


Figure: The Process of Filtration

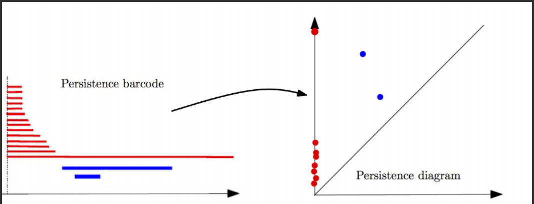
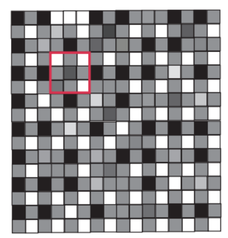


Figure: Barcode and Diagram

Example: Natural Image

"Round about the cauldron go"

- A collection of 4,167 digital photographs of random outdoor scenes was assembled in the late 1990s by van Hateren and van der Schaaf.
- Mumford, Lee, and Pederson construct a data set M. Each such square is a matrix of grey-scale intensities(0-255). The full data set consists of roughly 8,000,000 points in E^9.
- Mean center the data and Normalize the D-norm



"Hover through the fog"

- codensity function $\delta_k(x) = d(x, v_k(x))$, $v_k(x)$ denotes the k-th nearest neighbor of x in X
- $M_0[k, T]$=$\{x \in M_0 | \delta_k(x)$ lies among the $T\%$ lowest values of δ_k in $M_0\}$
- construction of complex: witness complex

"When shall we three meet again?"

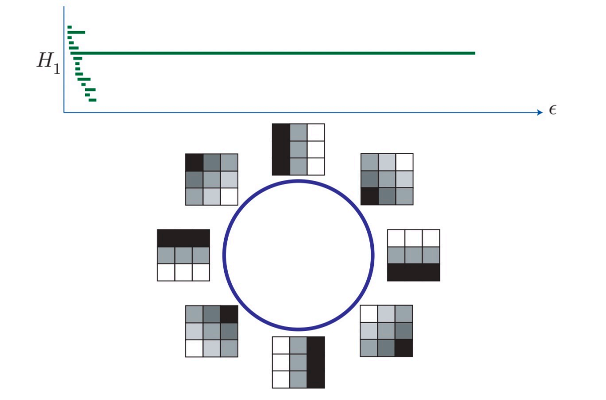


Figure: $M[300, 25]$

“When shall we three meet again?”

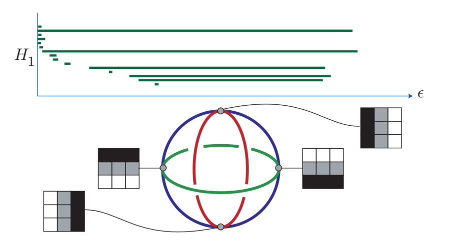
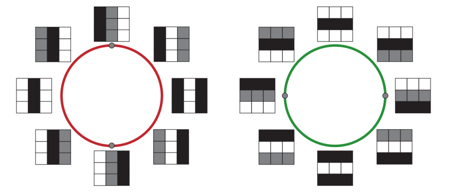


Figure: $M[15, 25]$



"Come like shadows, so depart!"

- When we consider more points

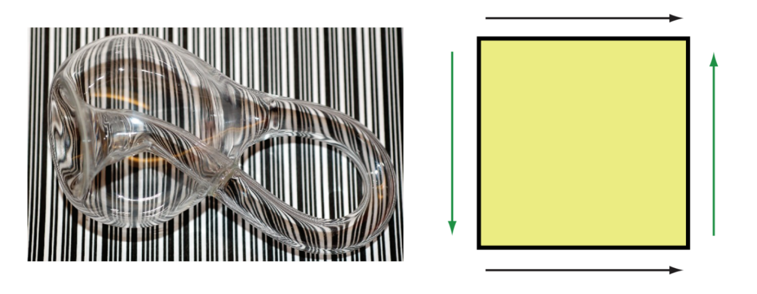
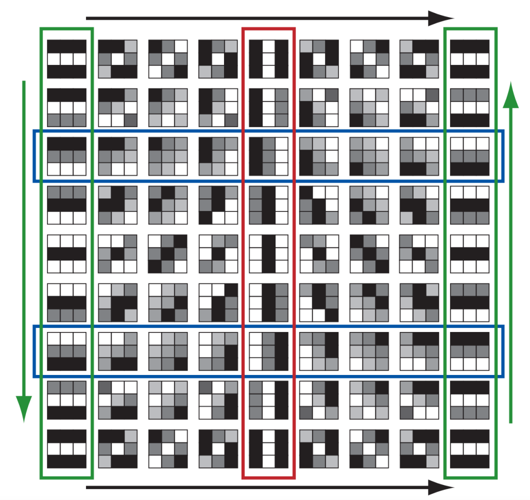


Figure: Klein Bottle

"Come like shadows, so depart!"



Sliding Window Embedding

Attractor

Definition

a global continuous time dynamical system is a pair (M, Φ), where M is a topological space and $\Phi : \mathbb{R} \times M \to M$ is a continuous map so that $\Phi(0, p) = p$, and $\Phi(s, \Phi(t, p)) = \Phi(s + t, p)$ for all $p \in M$ and all $t, s \in \mathbb{R}$.

Definition

an attractor is an subset $A \subset M$, which satisfies three conditions: (1) it is compact, (2) it is an invariant set —— this is ,if $a \in A$ then $\Phi(t, a) \in A$ for all $t \geq 0$ —— and (3) it has an open basin of attraction. In other words, there is an invariant open neighborhood $U \subset M$ of A, so that $\cap_{t\geq 0}\{\Phi(t, p) : p \in U\} = A$.

Strong Attractor

An attractor A is called strange, if (1) there are arbitrarily close points p, p' in a basin of attraction of A for which the distance between $\Phi(t,p)$ and $\Phi(t,p')$ grows exponentially quickly with t, and (2) A has non-integral Hausdorff dimension.

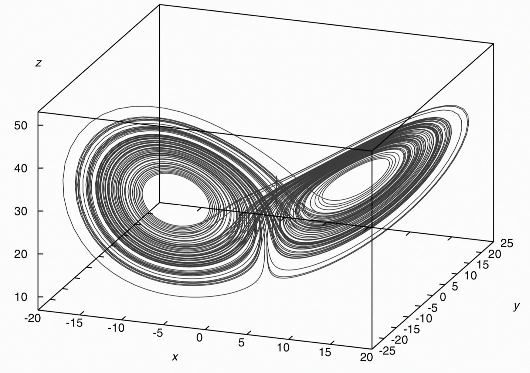
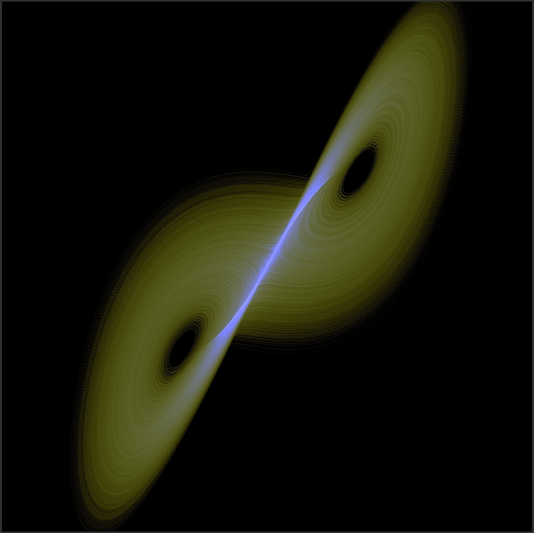


Figure: Lorenz's Butterfly Attractor

Strong Attractor



Takens' Theorem

Observation function is a continuous map $F : M \to \mathbb{R}$. Given an initial state $p \in M$, one obtains the time series

$$\begin{aligned} \phi_p :& \mathbb{R} \to \mathbb{R} \\ & t \mapsto F \circ \Phi(t, p) \end{aligned}$$

Theorem

(Takens' Embedding) Let M be a smooth, compact Riemannian manifold. let τ be a real number; and let $d \geq 2dim(M)$ be an integer. Then, for generic $\Phi \in C^2(\mathbb{R} \times M, M)$ and $F \in C^2(M, \mathbb{R})$, and for $\phi_p(t)$ as defined above, the delay map

$$\begin{aligned} \phi :& M \to \mathbb{R}^{d+1} \\ & p \mapsto (\phi_p(0), \phi_p(\tau), \cdots, \phi_p(d\tau)) \end{aligned}$$

is an embedding.

Sliding Window Embedding

Definition

Let $f : \mathbb{R} \to \mathbb{R}$ be a function, $\tau \geq 0$ a real number, and $d \geq 0$ an integer. The sliding window embedding of f, with parameters d and τ, is the vector-valued function

$$\begin{aligned} SW_{d,\tau} f : &\mathbb{R} \to \mathbb{R}^{d+1} \\ &t \mapsto (f(t), f(t+\tau), f(t+2\tau), \cdots, f(t+d\tau)). \end{aligned}$$

The integer $d+1$ is the dimension, τ is the delay, and the product $d\tau$ is the window size. For $T \subset \mathbb{R}$, the set

$$\mathbb{SW}_{d,\tau} f = \{SW_{d,\tau} f(t) : t \in T\}$$

is the sliding window point cloud associated to the sampling set T.

The Results of Jose A. Perea

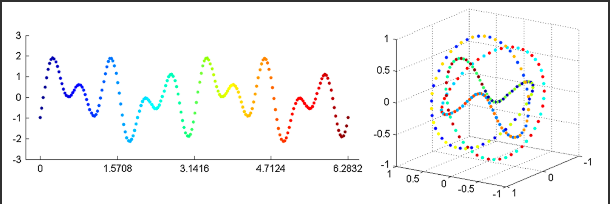
- The behavior of periodic function

Definition

We say that a function f is L-periodic on $[0, 2\pi]$, $L \in \mathbb{N}$, if

$$f(t + \frac{2\pi}{L}) = f(t)$$

for all t.



The Results of Jose A. Perea

- The behavior of quasi-periodic functions

Definition

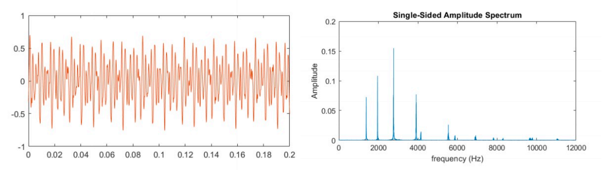
Let $\omega_1, \omega_1, \cdots, \omega_N > 0$ be incommensurate. A function $f : \mathbb{R} \to \mathbb{C}$ is said to be quasi-periodic with frequency vector $\omega = (\omega_1, \omega_1, \cdots, \omega_N)$ if

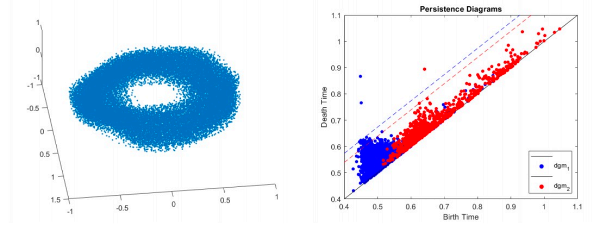
$$f(t) = F(\omega_1 t, \cdots, \omega_N t)$$

for all $t \in \mathbb{R}$ and a continuous function $F : \mathbb{T}^N \to \mathbb{C}$. That is $F \in C(\mathbb{T})$, which we will call a parent function for f.

The Results of Jose A. Perea

- The behavior of quasi-periodic functions





Example: Wheeze detection

The End