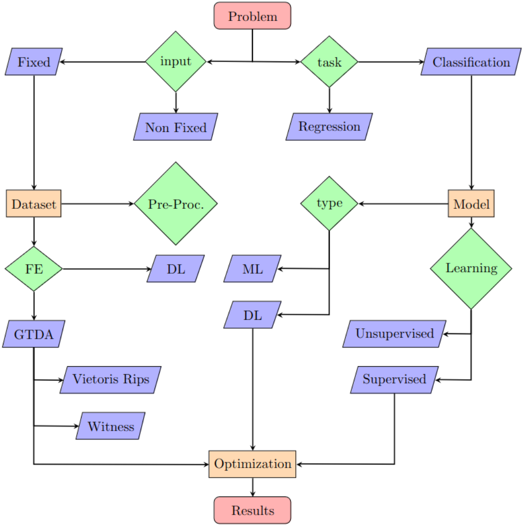
Persistent homology for TS classification

Andrea Chelini

Università di Trento

30 Maggio 2023


Problem
input
task
Fixed
Classification
Non Fixed
Regression
Dataset
Pre-Proc.
type
Model
FE
DL
ML
Learning
DL
GTDA
Unsupervised
Vietoris Rips
Supervised
Witness
Optimization
Results

Problems with naive classification methods of TS

Definition (**Dynamic Time Warping (DTW)**)

Let $S_1 = \{x_i\}_{i \in I}, S_2 = \{x'_j\}_{j \in J}$ be two sequence of elements in $\mathbb{R}^d$ and let $\mathcal{A}(S_1, S_2)$ be the set of all path alignment between I and J. Then the DTW of S_1 and S_2 is defined as

$$DTW_q(S_1, S_2) = \min_{\pi \in \mathcal{A}(S_1, S_2)} \left(\sum_{(i,j) \in \pi} \|x_i - x'_j\|^q_{l^q} \right)^{\frac{1}{q}}$$

It is possible to dynamically program $\mathcal{A}(S_1, S_2)$. Let $n = |S_1|$ and $m = |S_2|$ then

$$D(n, m) = \begin{cases} 0 & m < 0 \vee n < 0 \\ 1 & m = 0 \vee n = 0 \\ \begin{matrix} D(n-1,m)+D(n-1,m-1) \\ +D(n,m-1) \end{matrix} & otherwise \end{cases}$$

From TS to cloud point

Theorem (Takens' embedding theorem)

Let M be a $m-$dimensional manifold, $\phi \in Diff(C^2(M))$ and $y \in C^2(M, \mathbb{R})$ then the map

$$\begin{aligned} \Phi_{(\phi,y)} : M &\to \mathbb{R}^{2m+1} \\ x &\mapsto \Phi_{(\phi,y)}(x) := \left(y(x), y(\phi(x)), \cdots, y\left(\phi^{2m}\right)\right) \end{aligned}$$

is an embedding.

Definition (**Embedding**)

An injective continuous map $f : X \to Y$ between topological spaces X and Y is a topological embedding if f yields a homeomorphism between X and $f(X)$.

Definition (**Diffeomorphism**)

$f : M \to N$ is a diffeomorphism, $f \in Diff(C^0(M))$ if

- *f is a differentiable bijection.*
- *M, N are two manifolds.*
- *f^{-1} is smooth*

More in general if $f, f^{-1} \in C^k$ then $f \in Diff\left(C^k(M)\right)$

Definition (**Manifold**)

$M \subset R^d$ is a d-dimensional manifold if $\exists\{(W_i, h_i)\}_{i\in I}$ such that:

- *$\{W_i\}_{i\in I}$ is cover of M, i.e., $M = \bigcup_i W_i$*
- *$\forall i \in I, h_i : W_i \to U_i \subset R^d$ is a diffeomorphism and the set is called coordinate maps*
- *$\forall j, i \in I, h_j \circ h_i^{-1} : U_i \to U_j$ is smooth*

The inverses, $h_i^{-1} : U_i \to W_i$, are called parametrizations. The collection $\{(W_i, h_i)\}_{i\in I}$ is called atlas of the manifold.

From the previous definitions, it is possible to see that

1. $M = (0, 1)$ is a 1-dimensional manifold.
2. wlog it is possible to consider every time series $y : (0, 1) \to \mathbb{R}$.
3. The function $\phi(t) = t - \tau, \tau \in \mathbb{R}$ is such that $\phi \in Diff(C^\infty(\mathbb{R}))$ so $\phi \in Diff(C^2((0, 1)))$

This implies that

$$\{(y(t), y(t-\tau), y(t-2\tau)) | t \in (2\tau, 1), \tau \in (0, 0.5)\}$$

is an embedding of the time series $y(t)$

Definition (**Simplex**)

A point $x \in \mathbb{R}^d$ *is a convex combination of the points in* S *if it is an affine combination with non-negative coefficients. The convex hull of* S *is the set of all convex combinations of* S*, i.e.,*

$$conv(S) := \left\{ \sum_{i=0}^{k} \lambda_i u_i \,\middle|\, \sum_{i=0}^{k} \lambda_i = 1, \lambda \geq 0 \right\}$$

A $k-$*simplex is the convex hull of* $k+1$ *affinely independent points in* $\mathbb{R}^d$ *and it has dimension* k*.*

Definition (**Simplicial complex**)

A simplicial complex K *is a finite collection of simplexes such that*

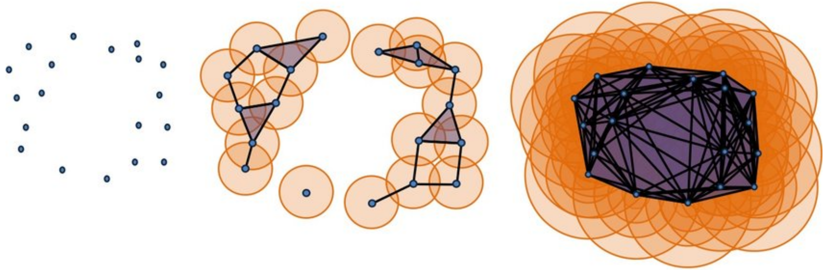
1. *if* $\sigma \in K$ *and* $\tau \leq \sigma$ *then* $\tau \in K$*.*
2. *if* $\sigma_1, \sigma_2 \in K$ *then* $\sigma_1 \cap \sigma_2 \in K$

Complexes

Definition (Vietoris-Rips (VR) complexes)

Let S be a set of points in R^d. The Vietoris-Rips complex of S with radius r is the collection of subsets of S of diameter at most $2r$, i.e.,

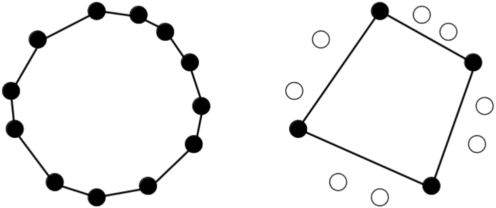
$$VR_r(S) = \{\sigma \subset S : diam(\sigma) \leq 2r\}$$



Definition (**Witness complex**)

Let L, Z be two cloud point such that $L \subset Z$ and $n = |L|, N = |Z|$. Let D be $n \times N$ matrix of non-negative entries regarded as the matrix of distances between the points in L and Z. The (strict) witness complex $W_\infty(D)$ with vertices $\{1, \cdots, n\}$ is defined as

$$\sigma := [a_0, \cdots, a_p] \in W_\infty(D) \iff$$
$$\exists i \in \{1, \cdots, N\} | D(a_j, i) \leq D(b, i), \forall b \in \{1, \cdots, N\} \backslash \{a_1, \cdots, a_n\}$$



Definition (**p-chain**)

Let K be a simplicial complex and A an abelian group. A p-chain is a formal sum of p-dimensional faces of K, i.e.

$$C_p(K) = \left\{ \sum_{\sigma \in F_p(K)} a_\sigma \sigma | a_\sigma, \in A \right\}$$

Definition

Let K be a simplex. A p-cycle $Z_p(K)$ is a chain with empty boundary

$$Z_p(K) := \{\sigma \in Ker(\partial_p(C_p(K)))\}$$

A p-boundary $B_p(K)$ is a chain which is a boundary of a (p+1)-chain

$$B_p(K) := \{\sigma \in C_p(K) | \exists \sigma' \in C_{p+1}(K), \partial_{p+1}\sigma' = \sigma\}$$

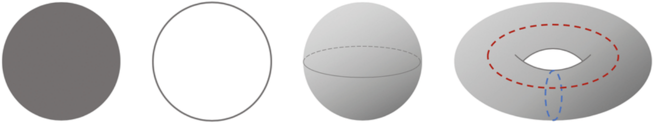
Since $Z_p(K) \supset B_p(K)$ the following definition has meaning.

Definition (p-th homology group)

Let K be a simplicial complex. The p-th homology group is

$$H_p(K) := {}^{Z_p(K)}\!\big/_{B_p(K)}$$

and $\beta_p(K) := \mathrm{rank}\,(H_p(K))$ is the p-th Betti number.



$\beta_0 = 1$	$\beta_0 = 1$	$\beta_0 = 1$	$\beta_0 = 1$
$\beta_1 = 0$	$\beta_1 = 1$	$\beta_1 = 0$	$\beta_1 = 2$
$\beta_2 = 0$	$\beta_2 = 0$	$\beta_2 = 1$	$\beta_2 = 1$

Definition (Filtration)

A filtration $\mathcal{F}$ is an indexed family $\{S_i\}_{i \in I}$ of sub-objects of a given algebraic structure S such that

$$S_i \subseteq S_j \qquad \forall i, j \in I, i \leq j$$

Let K be a simplicial complex and let $f : K \to \mathbb{R}$ be a monotonic function. The induced filtration of K is $\{K_i\}_{i=0}^{n}$ where $K_0 = \phi, K_n = K$ and

$$a_i := \min \left\{ r \,\middle|\, K_i = f^{-1}((-\infty, r])) \right\}$$

By construction it holds that $\forall i \leq j, K_i \subset K_j$ then it is possible to define $f_p^{i,j} : H_p(K_i) \to H_p(K_j)$ where $f_p^{i,j}$ is the induced map.

Definition

The p-th persistent homology groups are the images of the homomorphism induced by inclusion so

$$H_p^{i,j}(K) = immf_p^{i,j}$$

The ranks of these groups are the p-th persistence Betti numbers

$$\beta_p^{i,j} = rank\left(H_p^{i,j}(K)\right)$$

Definition (life-and-death)

Give the p-th homology group $H_p(K_i)$ *and* $\gamma \in H_p(K_i)$ *then*

- γ *is born at* K_i *if* $\gamma \notin H_p^{i-1,i}$
- γ *dies in* K_j *if it merges with an older class as we go from* K_{j-1} *to* K_j.

Moreover, consider a_i, a_j, *previously defined. Then the persistence of* γ *is*

$$pers(\gamma) = a_j - a_i$$

and (a_j, a_i) *is called the persistent pair.*

Features extraction

1. The number of elements in each dimension.
2. Maximum holes lifetime in each dimension.
3. Number of relevant holes which is the fraction of points that have a similar distance.
4. Average lifetime in dimension.
5. Sum of all lifetime (wants to represent the integral on the PD graph)
6. Betti numbers.

Classifier

Definition (**Neural network (NN)**)

A neural network $\mathcal{N}$, is a sequence $(D_0, D_1, \cdots, D_L, W^1, \theta^1, W^2, \theta^2, \cdots, W^L, \theta^L)$ where:

- $L \in \mathbb{N} \setminus \{0\}$ *and it represent the depth of* $\mathbb{N}$.
- $(D^0, \cdots, D^L) \in \mathbb{N}^{L+1}$ *is the layout of the network and* D_0 *is the number of features of the input and* D_L *is the number of the features of the output.*
- $W^l = \left(W^l_{j,k}\right) \in \mathbb{R}^{D_l \times D_{l-1}}$ *matrices whose entries are referred to be the network's weights and* j *is referred to the* $j-th$ *nodes in the* D^l *layer while* k *to the* $k-th$ *node of the* D^{l-1} *layer.*
- $(\theta^l) = \left(\theta^l_j\right) \in \mathbb{R}^{D_l}$ *where* $l \in \{1, \cdots, L\}$ *and it is called the bias.*

Definition (**Forward propagation**)

Given a neural network $\mathcal{N}$ and $\rho = \{\rho_i\}_{i=1}^{L}$ the set of activation function associated at each level of $\mathcal{N}$. The function defined by $\mathcal{N}$ equipped with ρ is defined as

$$\begin{aligned} < \mathcal{N} >^{\rho} \colon \mathbb{R}^{D_0} &\to \mathbb{R}^{D_L} \\ x &\mapsto < \mathbb{N} >^{\rho} (x) = y^{[L]} \end{aligned}$$

where

$$\begin{cases} y^{[0]}(x) = 0 \\ y^{[l]}(x) = \rho^{l}\left(W^{l} y^{[l-1]}(x) + \theta^{l}\right) \quad l \in \{1, \cdots, L\} \end{cases}$$

and this system defines forward propagation.