Homological Perspective on Data

Luciano Melodia

Overview

Part I: Topological spaces

Part II: Simplicial complexes

Part III: Homology groups

Part IV: Talks

Part I:

Topological Spaces

Topology saves lives

The case of Mike Hughes

Taken from Süddeutsche Zeitung:
Der Mann, der beweisen will, dass die Erde eine Scheibe ist.

Topological spaces

Aim of algebraic topology

*One of the main ideas of **algebraic topology** is to consider two spaces to be equivalent if they have **'the same' shape** in a sense that is much **broader than homeomorphism**.*

Allen Hatcher, Algebraic Topology, Chap. 0

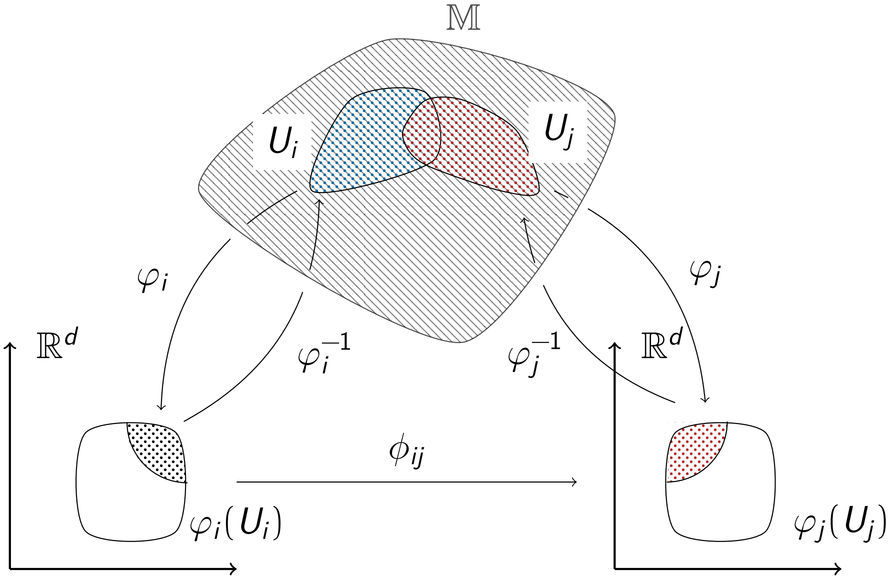
Topological spaces
More structure

A **topological space** is a pair $(X, \mathcal{O})$, consisting of a set X and another set $\mathcal{O}$ of subsets from X, such that:

- An arbitrary union of sets from $\mathcal{O}$ is in $\mathcal{O}$.
- An arbitrary intersection of two sets from $\mathcal{O}$ is in $\mathcal{O}$.
- $\emptyset$ and X are in $\mathcal{O}$.

Topological spaces
Manifolds

A manifold $\mathbb{M}$ of dimension d embedded in some $\mathbb{R}^n$, with $d \ll n$ is a space where every point $p \in \mathbb{M}$ has a neighbourhood that is locally homeomorphic to $\mathbb{R}^d$.



$\mathbb{M}$
U_i
U_j
φ_i
φ_j
φ_i^{-1}
φ_j^{-1}
$\mathbb{R}^d$
$\mathbb{R}^d$
ϕ_{ij}
$\varphi_i(U_i)$
$\varphi_j(U_j)$

Topological spaces
Homeomorphisms

Let X and Y be two topological spaces. A projection $f : X \rightarrow Y$ is called **continuous** if the preimage of the open sets is also open.

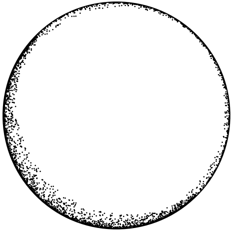
Topological spaces
Homeomorphisms

Let X and Y be two topological spaces. A projection $f : X \longrightarrow Y$ is called **continuous** if the preimage of the open sets is also open.

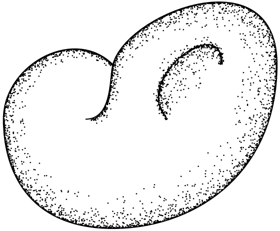
A bijective projection $f : X \longrightarrow Y$ is called **homeomorphism** if f and f^{-1} are both continuous, thus $U \subset X$ is open if and only if $f(U) \subset Y$ is also open.

Topological spaces

Examples of manifolds



A 2-sphere.



A 2-manifold with genus 1.

Part II:
Simplicial complexes

Simplicial complexes

Spanning the 1-skeleton

Data is an unstructured point cloud in some embedding space $\mathbb{R}^n$.

How to capture the shape of this point cloud?

Simplicial complexes

Spanning the 1-skeleton

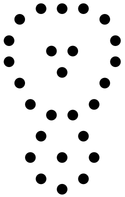
Data is an unstructured point cloud in some embedding space $\mathbb{R}^n$.

How to capture the shape of this point cloud?

Which object do we use for this?

Simplicial complexes

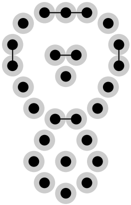
Spanning the 1-skeleton



Imagine your data being organized in the above shape.
How could we get information about the underlying manifold?

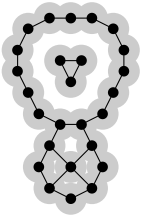
Simplicial complexes

Spanning the 1-skeleton



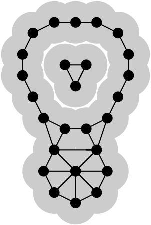
We inflate each point like a balloon and track the radius setting $r = 0.2$. Can you see the blurry shape?

Simplicial complexes
Spanning the 1-skeleton



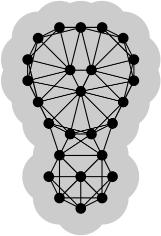
Every time a ball touches another one, we connect them with an edge.
We have $r = 0.4$ and one can already see the sought shape.

Simplicial complexes
Spanning the 1-skeleton



We ommit constructing other simplices than 1-simplices,
to make the visualization more intuitive.
The result is called 1-skeleton.

Simplicial complexes
Spanning the 1-skeleton



The 1-skeleton.

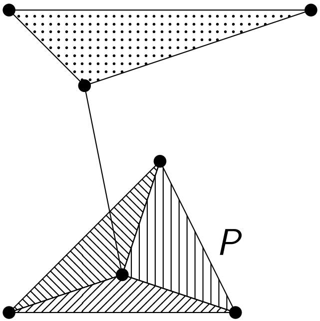
Simplicial complexes
Definition

Given a set $X = \{x_0, \cdots, x_k\} \subset \mathbb{R}^d$ of $k+1$ points that do not lie on a hyperplane with dimension less than d, the k-dimensional simplex σ spanned by X is the set of convex combinations, such that

$$\sum_{i=0}^{k} \lambda_i x_i \quad \text{with} \quad \sum_{i=0}^{k} \lambda_i = 1 \quad \text{and} \quad \lambda_i \geq 0. \tag{1}$$

Simplicial complexes

Example of a Vietoris-Rips complex

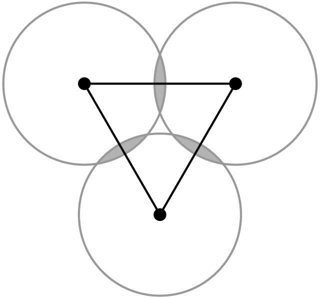


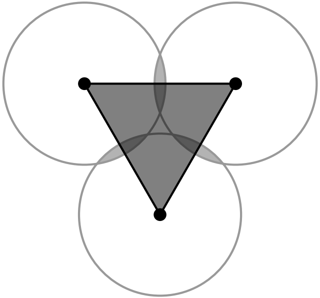
A simplicial complex composed of seven 0-faces, ten 1-faces, five 2-faces and one 3-face.

Simplicial complexes may differ

Čech vs. Vietoris-Rips complex

$C_t(P)$ $R_t(P)$





Taken from Tam Thanh Truong: Persistent Homology via Quotient Spaces.

Part III:
Homology groups

What are we seeking for?
Filtrations of spaces

We believe that different data comes from different spaces. How can we detect this and sort the data accordingly?

We don't want to distinguish data only by their ***holes****, therefore we introduce a* ***magnifying glass*** *which allows us to capture the* ***structure of a topological space*** *in different granularity.*

Chain groups

Chains and boundaries

The k**th chain group** of a **simplicial complex** K is $\langle C_k(K), +\rangle$, the **free abelian group** on the oriented k-simplices, where $[\sigma] = -[\tau]$ if $\sigma = \tau$ and σ and τ have different orientation.

An element of $C_k(K)$ is called k-chain, $\sum_i \lambda_i[\sigma_i], \lambda_i \in \mathbb{Z}, \sigma_i \in K$.

Chain groups
Boundary homomorphism

Let K be a simplicial complex and $\sigma \in K$, $\sigma = [v_0, v_1, \cdots, v_k]$. The **boundary homomorphism** $\delta_k : C_k(K) \to C_{k-1}(K)$ is

$$\delta_k \sigma = \sum_i (-1)^i [v_0, v_1, \cdots, \hat{v}_i, \cdots, v_n]. \tag{2}$$

Cycle group and boundary group
Cycles

The *k***th cycle group** Z_k is defined as

$$Z_k = \ker \delta_k \tag{3}$$
$$= \{c \in C_k \mid \delta_k c = \emptyset\}. \tag{4}$$

An element of this group is called *k*-**cycle**.

Cycle group and boundary group

Boundaries

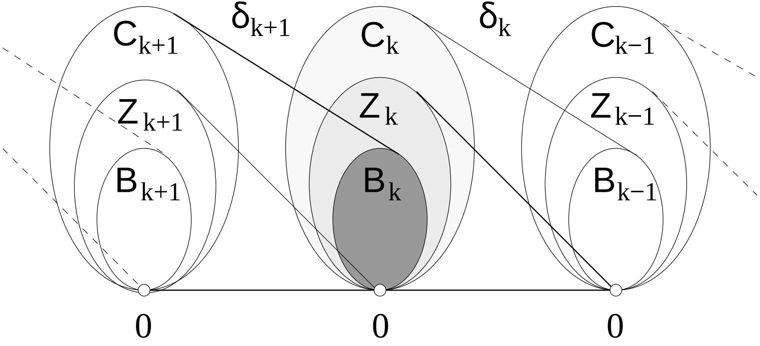
The k**th boundary group** B_k is defined as

$$\begin{aligned} B_k &= \text{im } \delta_{k+1} && (5) \\ &= \{c \in C_k \mid \exists d \in C_{k+1} : c = \delta_{k+1} d\}. && (6) \end{aligned}$$

An element of this group is called k-**boundary**.

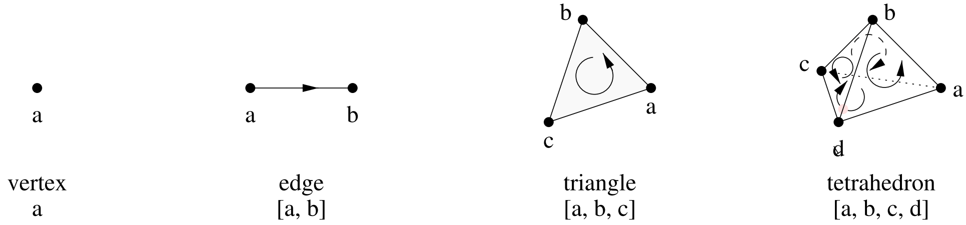
Nested sequences

Following Zomorodian, Edelsbrunner and others



Taken from Afra Zomorodian: Chap. 6, Homology.

Boundaries of simplices
Computation of boundaries



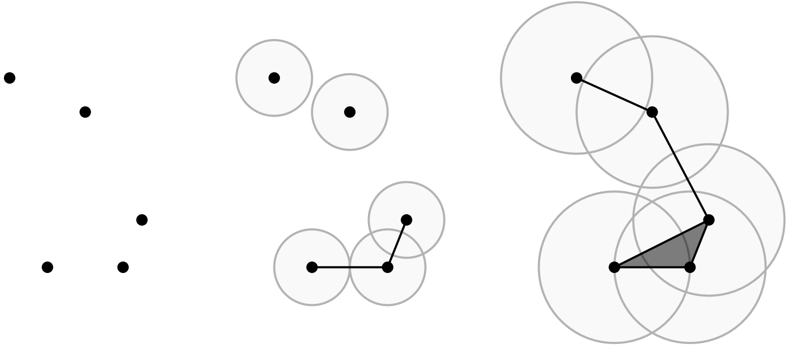
Taken from Afra Zomorodian: Chap. 6, Homology.

$$\begin{aligned} \text{edge:}\quad & \delta_1[a,b] = b - a, \\ \text{triangle:}\quad & \delta_2[a,b,c] = [b,c] - [a,c] + [a,b] = [b,c] + [c,a] + [a,b], \\ \text{tetrahedron:}\quad & \delta_3[a,b,c,d] = [b,c,d] - [a,c,d] + [a,b,d] - [a,b,c]. \end{aligned}$$

Chain complex

$$0 \xrightarrow{\delta_{k+1}} C_k \xrightarrow{\delta_k} C_{k-1} \xrightarrow{\delta_{k-1}} \cdots \xrightarrow{\delta_2} C_1 \xrightarrow{\delta_1} 0 \tag{7}$$

What are we seeking for?
Filtrations of spaces



$$P \subseteq R_2(P) \subseteq R_4(P)$$

Taken from Tam Thanh Truong: Persistent Homology via Quotient spaces.

Technical details

A **filtration** is a nested sequence of sets

$$\emptyset = K_0 \subseteq K_1 \subseteq K_2 \cdots \subseteq K_{k-1} \subseteq K_k = K, \tag{8}$$

such that each simplicial complex K_i is a valid subcomplex of K_{i+1}. To compute persistent homology is is enough to reduce a *single* boundary matrix of the simplicial complex K in order to get the homology groups along the filtration.

Homology groups

The kth **homology group** of a simplicial complex is defined as

$$H_k(K) = \ker \delta_k C_k(K)/\text{im } \delta_{k+1} C_{k+1}(X). \tag{9}$$

Intuitively, the kernel of the boundary homomorphism of the kth chain group gives all k-cycles, thus the cycle group, from which we quotient out all elements of the kth boundary group, i.e. $H_k(K) = Z_k(K)/B_k(K)$.

Technical details: Both subgroups, the cycle and boundary group, are normal because our chain groups are abelian.

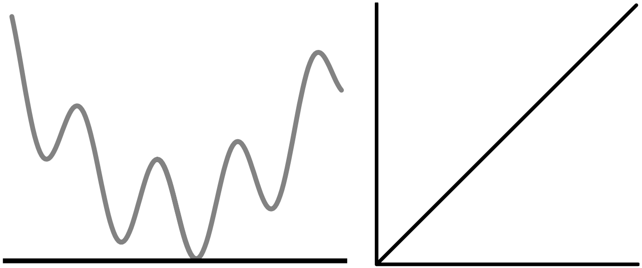
Betti numbers

The kth **Betti** number β_k is the rank of the kth homology group $H_k(K)$ of the topological space K.

We'll track the betti numbers to track the amount of holes along the filtration. Other properties of the vector space could also be helpful.

Example of Persistent Homology

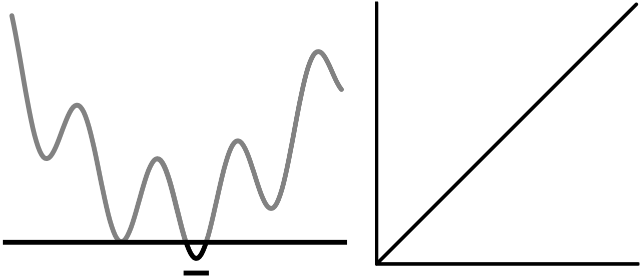
Persistent homology of a continuous function



Taken from Bastian Rieck: An Introduction to Persistent Homology.

Example of Persistent Homology

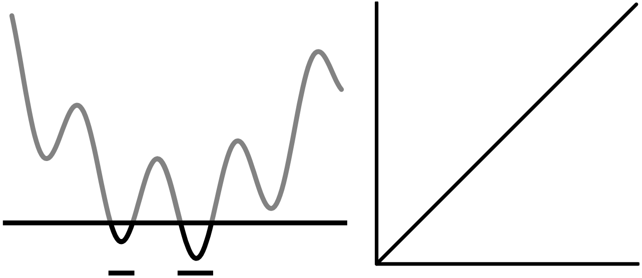
Persistent homology of a continuous function



Taken from Bastian Rieck: An Introduction to Persistent Homology.

Example of Persistent Homology

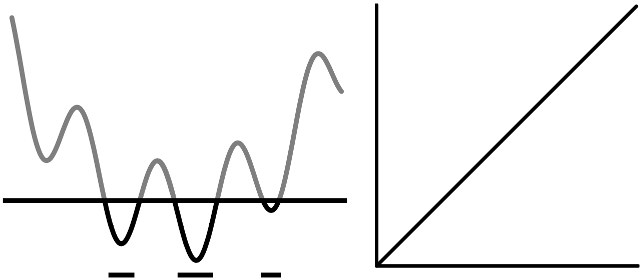
Persistent homology of a continuous function



Taken from Bastian Rieck: An Introduction to Persistent Homology.

Example of Persistent Homology

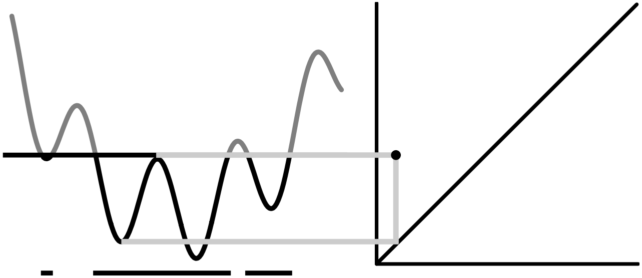
Persistent homology of a continuous function



Taken from Bastian Rieck: An Introduction to Persistent Homology.

Example of Persistent Homology

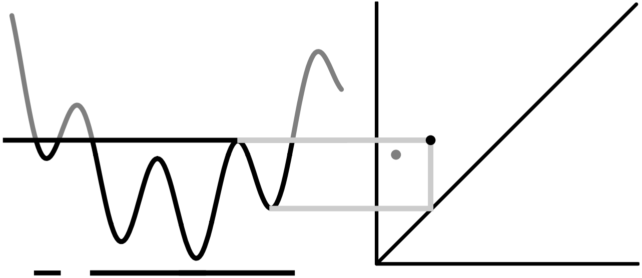
Persistent homology of a continuous function



Taken from Bastian Rieck: An Introduction to Persistent Homology.

Example of Persistent Homology

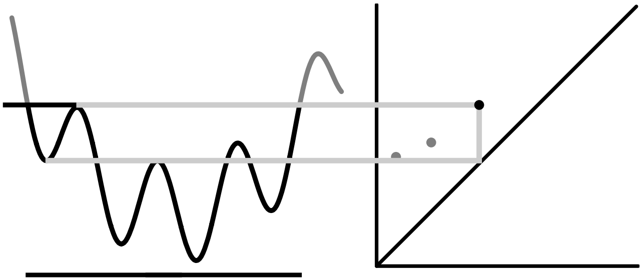
Persistent homology of a continuous function



Taken from Bastian Rieck: An Introduction to Persistent Homology.

Example of Persistent Homology

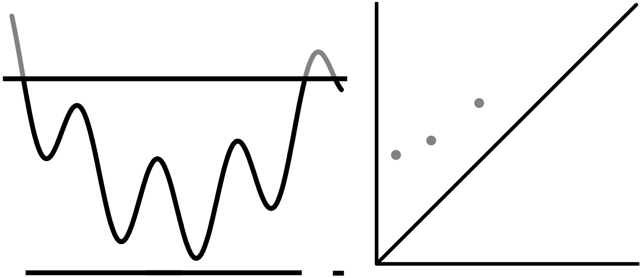
Persistent homology of a continuous function



Taken from Bastian Rieck: An Introduction to Persistent Homology.

Example of Persistent Homology

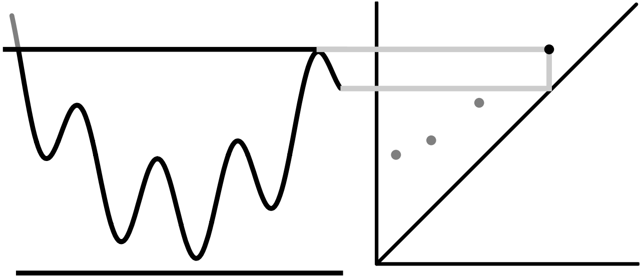
Persistent homology of a continuous function



Taken from Bastian Rieck: An Introduction to Persistent Homology.

Example of Persistent Homology

Persistent homology of a continuous function



Taken from Bastian Rieck: An Introduction to Persistent Homology.

Part IV:
Talks

Topological spaces and groups
Techies preferred

Define and illustrate the concepts topological space, metric space, topology of a metric space, bases and subbases of topologies, continuous mappings and homeomorphisms, compactness, groups, homomorphisms of groups and the most important vector spaces, Hilbert spaces, Banach spaces and Fréchet spaces.

Literature: Jänich, Klaus. Topologie. Springer-Verlag, 2013. Chap. 1 and 2.

Simplicial complexes
Techies preferred

Define simplicial complexes and simplicial maps. Explain the simplicial approximation theorem and the nerve theorem. You do not need to show the proofs in their entirety here, but you should have worked through them. Sketch the core concepts and ideas.

Prerequisites: Topological spaces and groups

Literature: James Munkres: Elements of algebraic topology, CRC Press, 2018. Sections $1-3$ and sections 14 and 16.

Simplicial complexes associated with point clouds

Define and show different simplicial complexes on a set of points in a metric space. In particular, show the Voronoi diagrams, the Delaunay complexes, the duality of the two as well as Alpha complexes and Witness complexes.

Prerequisites: Topological spaces and groups

Literature: Boissonnat, Jean-Daniel, Frédéric Chazal, and Mariette Yvinec. Geometric and topological inference. Vol. 57. Cambridge University Press, 2018. Chap. 4.2, 4.3, 6.1, 6.2.

Simplicial and singular homology groups

Techies preferred

Explain and define the simplicial homology groups on a simplicial complex, the so-called σ-complex. Explain what is meant by singular homology and use the cutout theorem to show that the simplicial homology groups match the singular homology groups on a σ complex.

Prerequisites: Topological spaces and groups, Simplicial Complexes

Literature: Hatcher, Allen. Algebraic topology. Cambridge University Press, 2005. S. 102-106, 128-130.

Persistent homology

Mathies preferred

Define persistent homology, show persistence diagrams and explain persistence diagrams using the relationship between persistent homology classes and the critical values of one-dimensional tame functions. Also explain persis-tent homology on the filtration of point sets in higher-dimensional spaces.

Prerequisites: Topological spaces and groups, Simplicial Complexes

Literature: Edelsbrunner, Herbert, and John Harer. "Persistent homology-asurvey." Contemporary mathematics 453 (2008): 257-282.
Boissonnat, Jean-Daniel, Frédéric Chazal, and Mariette Yvinec. Geometric andtopological inference. Vol. 57. Cambridge University Press, 2018. Chap. 11.5.

Computation of persistent homology

Mathies preferred

Specify the algorithm to calculate persistence diagrams and describe the matrix reduction techniques used. Give examples and make exemplary calculations of persistent homology on the triangulation of basic compact geometric objects.

Prerequisites: Topological spaces and groups, Simplicial Complexes, Simplicial and singular homology groups

Literature: Afra Zomorodian, and Gunnar Carlsson. "Computing persistenthomology." Discrete & Computational Geometry 33.2 (2005): 249-274.
Otter, Nina, et al. "A roadmap for the computation of persistent homology."EPJ Data Science 6.1 (2017): 17.

Persistent homology and cohomology*

Mathies preferred

Remind the audience of the definition of persistent homology. Specify the dual concept of persistent cohomology and explain the proof that both persistent homology and persistent cohomology generate the same barcodes. Why is persistent cohomology more efficient to compute?

Prerequisites: Topological spaces and groups, Simplicial Complexes, Simplicial and singular homology groups, Persistent homology

Literature: De Silva, Vin, Dmitriy Morozov, and Mikael Vejdemo-Johansson. "Dualities in persistent (co) homology." Inverse Problems 27.12 (2011): 124003.

Distances and the stability theorem

Describe the stability of the persistence diagrams in relation to the Hausdorff distance, the bottleneck distance and the Wasserstein distance. Give all three metrics and explain them with illustrative examples. Explain the quadrant lemma without formal proof, so that the listener gets a good understanding of the properties of persistence diagrams.

Prerequisites: Topological spaces and groups, Simplicial Complexes, Simplicial and singular homology groups, Persistent homology

Literature: David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. "Stability of persistence diagrams." Discrete & Computational Geometry 37.1 (2007): 103-120.

Statistics in persistent homology

Define persistence landscapes and make it clear how to obtain them from the persistence diagrams. Explain why the ordinary persistence diagram, unlike the persistence landscape, does not lie in a vector space and explain why this property is important for statistical data analysis.

Prerequisites: Topological spaces and groups, Simplicial Complexes, Simplicial and singular homology groups, Persistent homology

Literature: Bubenik, Peter. "Statistical topological data analysis using persistence landscapes." The Journal of Machine Learning Research 16.1 (2015): 77-102.

More Questions?

Contact me: **luciano.melodia@fau.de**

Thank you for your attention!