Topological Data Analysis and Clustering Algorithms in Machine Learning

İsmail GÜZEL

Mathematical Engineering - İTÜ

March 13, 2022

Outline

1. Introduction
2. Hierarchical Clustering
3. Topological Data Analysis
4. Theoretical Contributions
5. Experimental Contributions

Outline

Hierarchical Clustering and Zeroth Persistent Homology

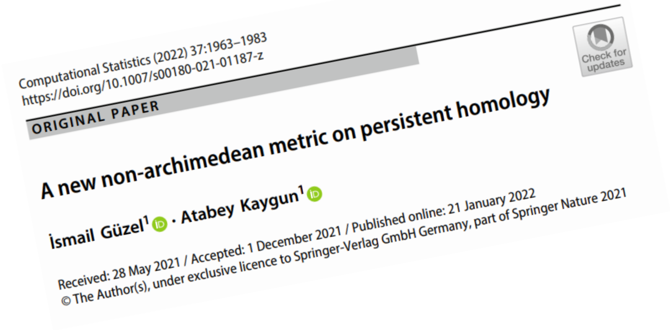
İsmail GÜZEL [1] Atabey KAYGUN [2]
[1]iguzel@itu.edu.tr [2]kaygun@itu.edu.tr
[1,2]Istanbul Technical University

İTÜ

1. Introduction
2. Hierarchical Clustering
3. Topological Data Analysis
4. Theoretical Contributions
5. Experimental Contributions

Outline

1. Introduction
2. Hierarchical Clustering
3. Topological Data Analysis
4. Theoretical Contributions
5. Experimental Contributions


Computational Statistics (2022) 37:1963–1983
https://doi.org/10.1007/s00180-021-01187-z

Check for updates

ORIGINAL PAPER

A new non-archimedean metric on persistent homology

İsmail Güzel[1] · Atabey Kaygun[1]

Received: 28 May 2021 / Accepted: 1 December 2021 / Published online: 21 January 2022

Outline

1. Introduction
2. Hierarchical Clustering
3. Topological Data Analysis
4. Theoretical Contributions
5. Experimental Contributions

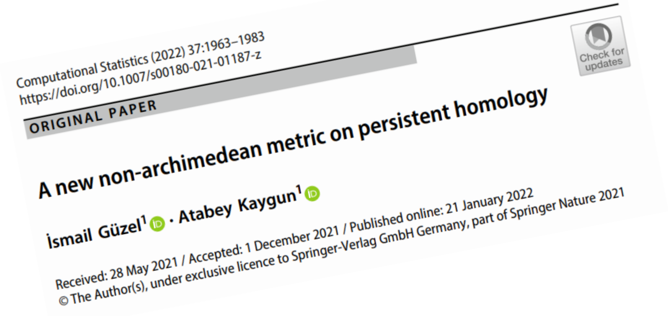
Hierarchical Clustering and Zeroth Persistent Homology

İsmail GÜZEL [1] Atabey KAYGUN [2]

[1]iguzel@itu.edu.tr [2]kaygun@itu.edu.tr

[1,2]Istanbul Technical University

İTÜ



Computational Statistics (2022) 37:1963–1983

https://doi.org/10.1007/s00180-021-01187-z

ORIGINAL PAPER

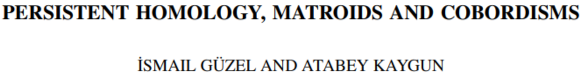
Check for updates

A new non-archimedean metric on persistent homology

İsmail Güzel[1] · Atabey Kaygun[1]

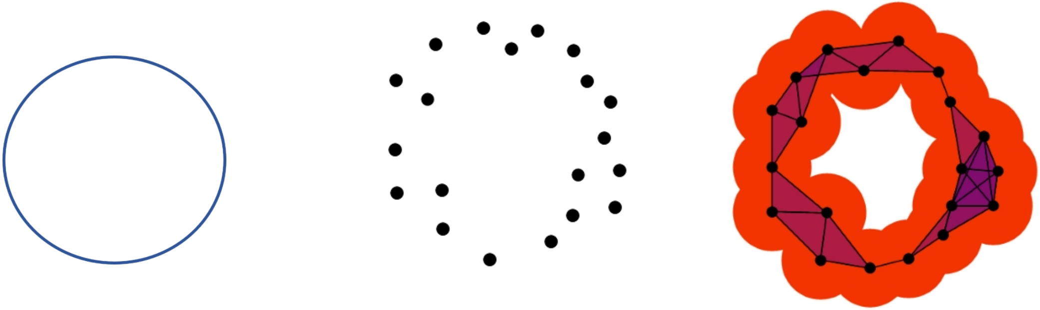
Received: 28 May 2021 / Accepted: 1 December 2021 / Published online: 21 January 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021



PERSISTENT HOMOLOGY, MATROIDS AND COBORDISMS

İSMAIL GÜZEL AND ATABEY KAYGUN



> *"Data has shape, and shape has meaning."*

Prof. Gunnar Carlsson

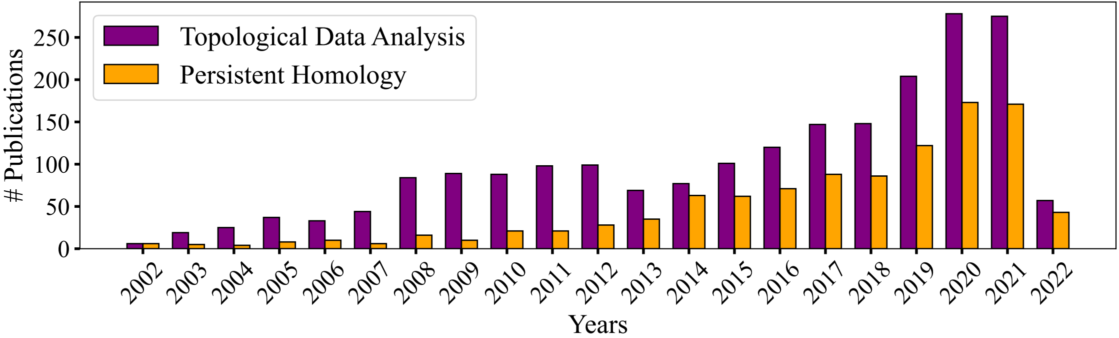
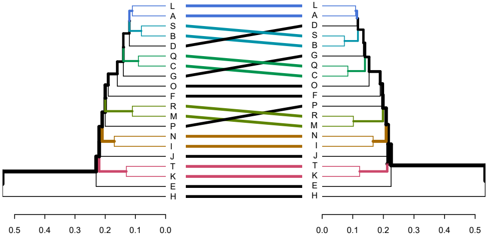
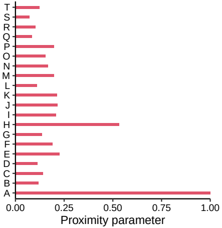


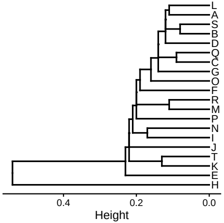
Figure: The data taken from SCOPUS

Questions answered

1. What are the similarities and differences between hierarchical clustering and 0-th persistent homology?
2. What is the difference from the persistence barcode?
3. What about higher dimensional persistent homology?
4. What are the experiments on real datasets?

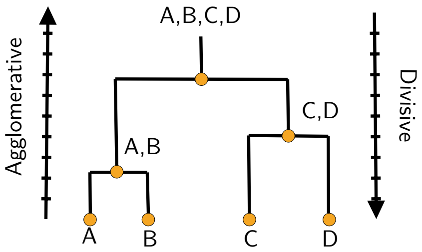


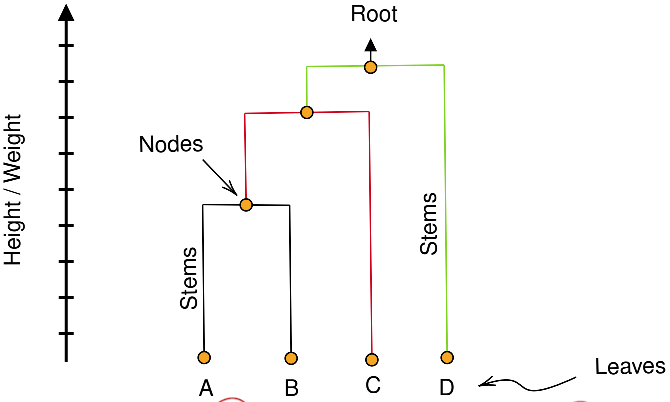


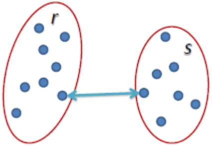


Hierarchical Clustering

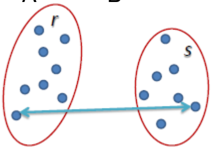
- Unsupervised machine learning algorithm
- Aim: divide the data set into disjoint subsets such that
  - homogeneous in cluster
  - heterogeneous between clusters
- The metric structure of ambient space from data set.
- A nice tree-based representation, called *a dendrogram*



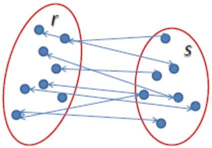




Single-Linkage



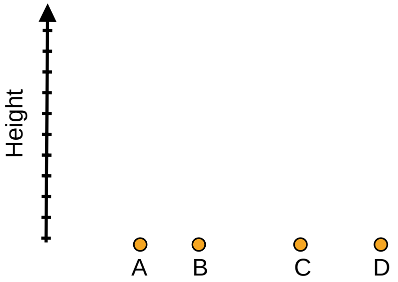
Complete-Linkage



Average-Linkage

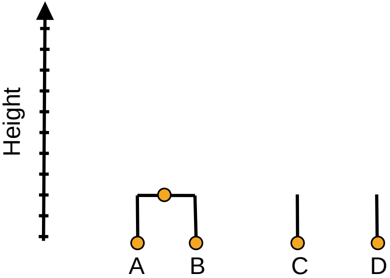
Cophenetic Matrix

$$
\begin{array}{c|c}
 & \begin{matrix} A & B & C & D \end{matrix} \\
\hline
\begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{bmatrix} & 2 & 5 & 6 \\ & & 4 & 5 \\ & & & 3 \\ & & & \end{bmatrix}
\end{array}
$$



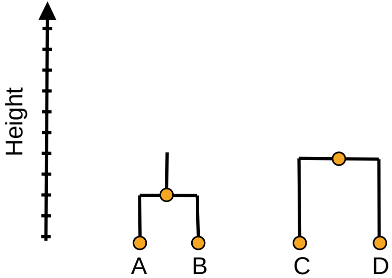
Cophenetic Matrix

$$
\begin{array}{c c}
 & \begin{array}{cccc} A & B & C & D \end{array} \\
\begin{array}{c} A \\ B \\ C \\ D \end{array} &
\begin{bmatrix}
 & 2 & 5 & 6 \\
 & & 4 & 5 \\
 & & & 3 \\
 & & &
\end{bmatrix}
\end{array}
$$



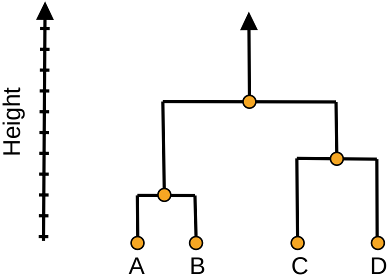
Cophenetic Matrix

$$\begin{array}{cc} & \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{bmatrix} & 2 & 5 & 6 \\ & & 4 & 5 \\ & & & 3 \\ & & & \end{bmatrix} \end{array}$$



Cophenetic Matrix

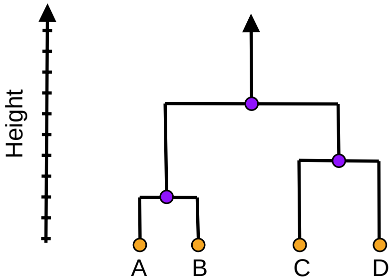
$$\begin{array}{c|c} & \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{bmatrix} & 2 & 5 & 6 \\ & & 4 & 5 \\ & & & 3 \\ & & & \end{bmatrix} \end{array}$$



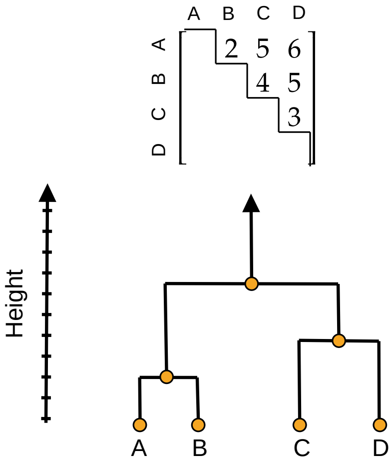
Cophenetic Matrix

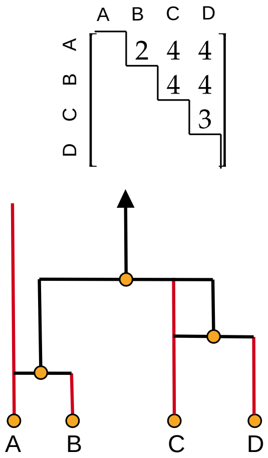
$$\begin{array}{c c} & \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{bmatrix} & 2 & 5 & 6 \\ & & 4 & 5 \\ & & & 3 \\ & & & \end{bmatrix} \end{array}$$

$$\begin{array}{c c} & \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \end{array} & \begin{bmatrix} & 2 & 4 & 4 \\ & & 4 & 4 \\ & & & 3 \\ & & & \end{bmatrix} \end{array}$$



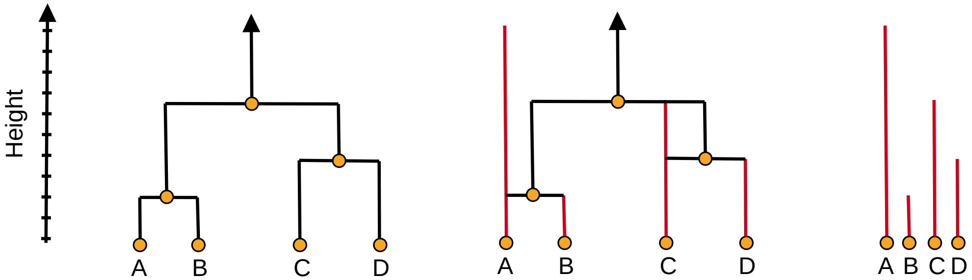
Cophenetic Matrix





Cophenetic Matrix

$$
\begin{array}{c|c}
 & \begin{matrix} A & B & C & D \end{matrix} \\
\begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{bmatrix} & 2 & 5 & 6 \\ & & 4 & 5 \\ & & & 3 \\ & & & \end{bmatrix}
\end{array}
\qquad
\begin{array}{c|c}
 & \begin{matrix} A & B & C & D \end{matrix} \\
\begin{matrix} A \\ B \\ C \\ D \end{matrix} & \begin{bmatrix} & 2 & 4 & 4 \\ & & 4 & 4 \\ & & & 3 \\ & & & \end{bmatrix}
\end{array}
$$



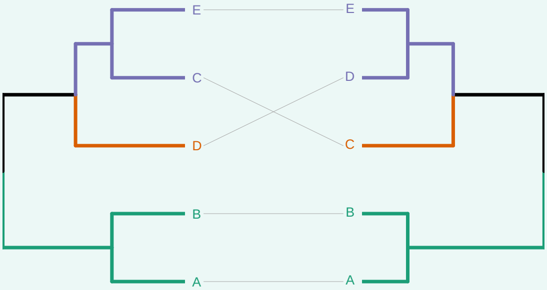
Compare Dendrograms

- Two Dendrograms
- Two cophenetic matrix

Mantel Test

- Non-parametric Test
- Like correlation coefficient
- Mantel statistic $r \in [-1, 1]$

Tanglegrams



Entanglement values ranges 0 and 1.

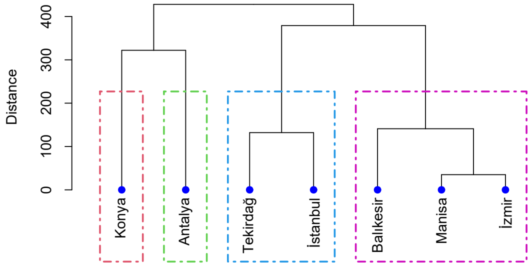
Silhouette scores

Silhouette scores

$$s(x) = \frac{b(x) - a(x)}{\max(a(x), b(x))},$$

$$a(x) = d(x, U(x)) \quad \text{and} \quad b(x) = \min_{C \neq U(x)} d(x, C).$$

	Tekirdağ	İstanbul	Balıkesir	Manisa	İzmir	Konya	Antalya
Tekirdağ	0						
İstanbul	132	0					
Balıkesir	379	390	0				
Manisa	511	529	141	0			
İzmir	506	564	176	35	0		
Konya	794	662	551	534	550	0	
Antalya	850	718	505	428	444	322	0



Points	Cohesion	Separation	Silhouette
Tekirdağ	132	465.3	0.72
İstanbul	132	494.5	0.73
Balıkesir	158.5	384.5	0.59
Manisa	88	428	0.79
İzmir	105.5	444	0.76
Konya	0	322	1
Antalya	0	322	1

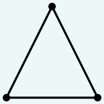
Overall silhouette score is $s \approx 0.80$

Simplicial Technology

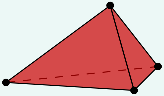
$k-$simplex

0-simplex
vertex

1-simplex
edge

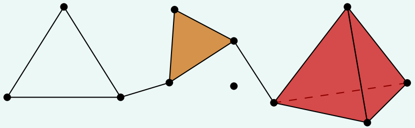


2-simplex
triangle

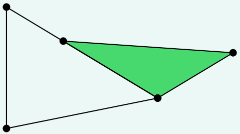


3-simplex
tetrahedron

Simplicial complex



Simplicial complex



Not simplicial complex

Point Cloud to Complex

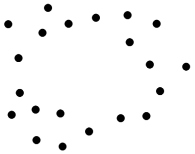
Čech Complex

$$\mathscr{C}_\varepsilon = \left\{ \sigma \subseteq \mathscr{D} \quad \middle| \quad \bigcap_{x \in \sigma} B_\varepsilon(x) \neq \varnothing \right\}, \quad B_\varepsilon(x) = \{ y \mid d(x,y) < \varepsilon \}$$

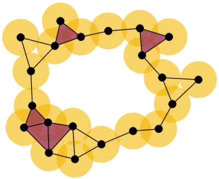
Vietoris Rips Complex

$$\mathscr{R}_\varepsilon = \{ \sigma \subset \mathscr{D} \quad | \quad \|x - y\| \leq \varepsilon, \text{ for all } x, y \in \sigma \}$$

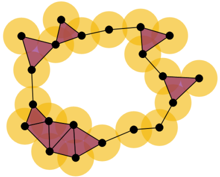
$\mathscr{D}$



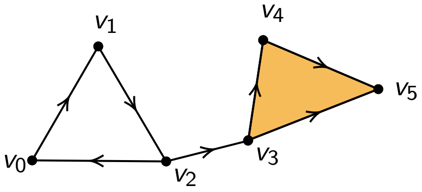
$\mathscr{C}_\varepsilon$



$\mathscr{R}_\varepsilon$



Example: A Simplicial complex $\mathscr{K}$



$$\mathscr{C}_0 = \{v_0, v_1, v_2, v_3, v_4, v_5\},$$
$$\mathscr{C}_1 = \begin{array}{r}\{[v_0, v_1], [v_1, v_2], [v_2, v_0], [v_2, v_3],\\ [v_3, v_4], [v_4, v_5], [v_3, v_5]\}\end{array},$$
$$\mathscr{C}_2 = \{[v_3, v_4, v_5]\}.$$

$$0 \xrightarrow{\partial_3} \mathscr{C}_2 \xrightarrow{\partial_2} \mathscr{C}_1 \xrightarrow{\partial_1} \mathscr{C}_0 \xrightarrow{\partial_0} 0$$

Homology

The k^{th} homology group of $\mathscr{K}$ is defined by

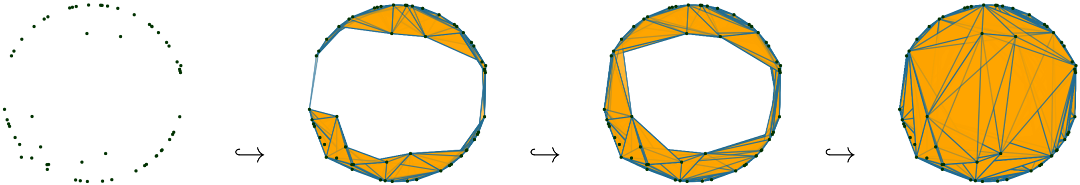
$$H_k(K) := {}^{\ker(\partial_k)}\!\big/_{\operatorname{im}(\partial_{k+1})}$$

$$\beta_0 = \dim(\ker(\partial_0)) - \dim(\operatorname{im}(\partial_1)) = 6 - 5 = 1$$

$$\beta_1 = \dim(\ker(\partial_1)) - \dim(\operatorname{im}(\partial_2)) = 2 - 1 = 1$$

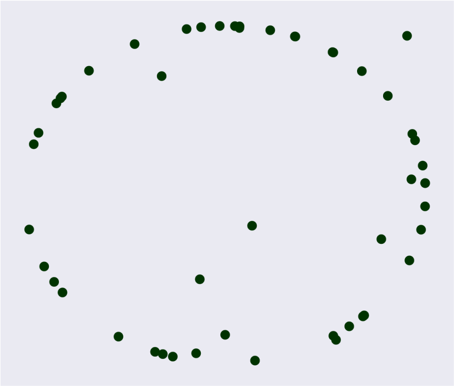
Vietoris-Rips Filtration

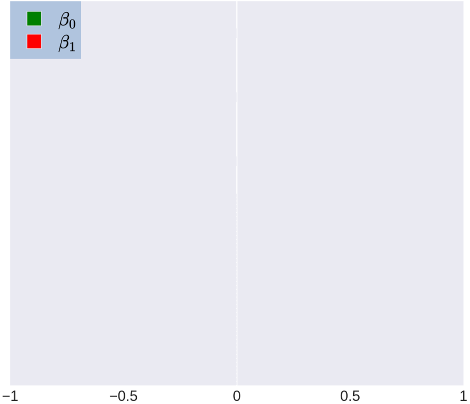
Filtration of a simplicial complex K is a collection of subcomplexes $\mathbb{K} = \{K_\varepsilon : \varepsilon \in \mathbb{R}^+\}$ that satisfy $K_{\varepsilon_1} \subseteq K_{\varepsilon_2}$ whenever $\varepsilon_1 \leq \varepsilon_2$.



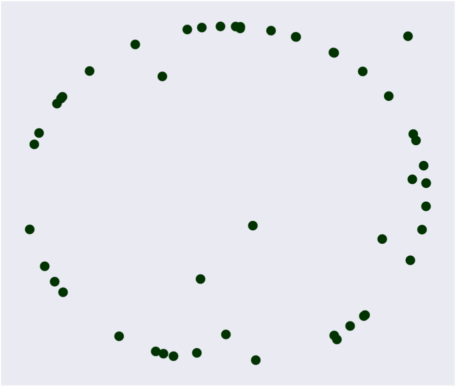
$\beta_0 = 50$	$\beta_0 = 1$	$\beta_0 = 1$	$\beta_0 = 1$
$\beta_1 = 0$	$\beta_1 = 1$	$\beta_1 = 1$	$\beta_1 = 0$

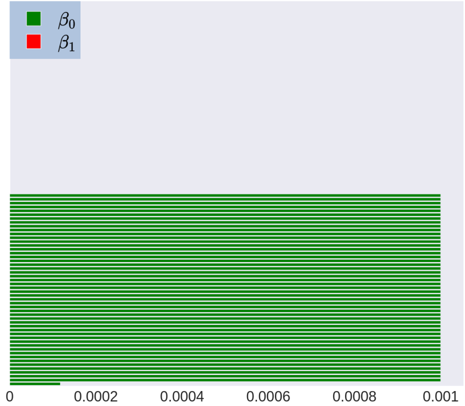
Persistent Homology and Barcodes



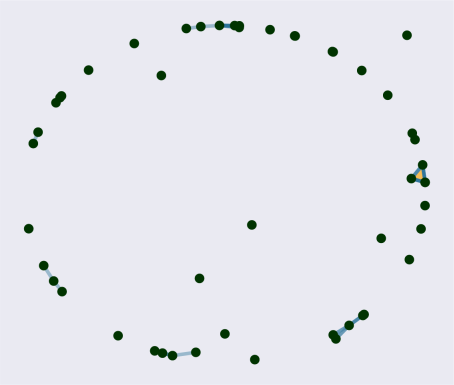


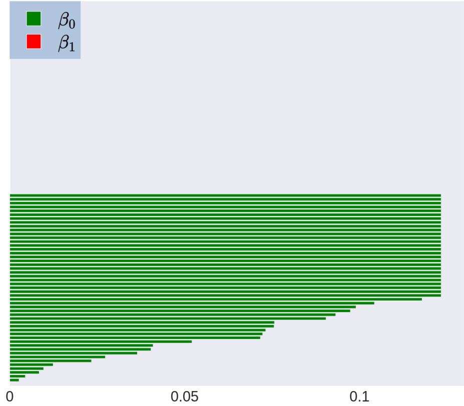
Persistent Homology and Barcodes



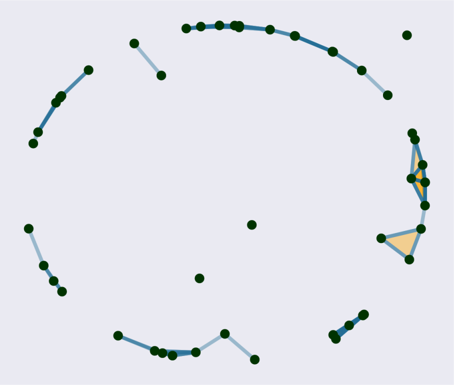


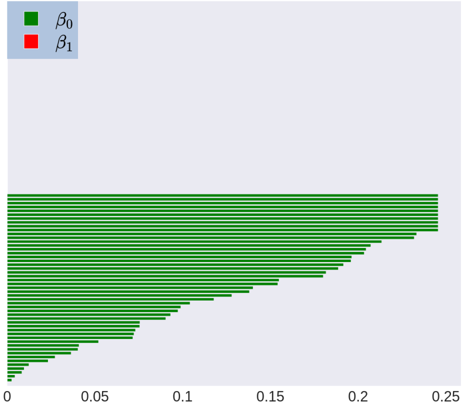
Persistent Homology and Barcodes



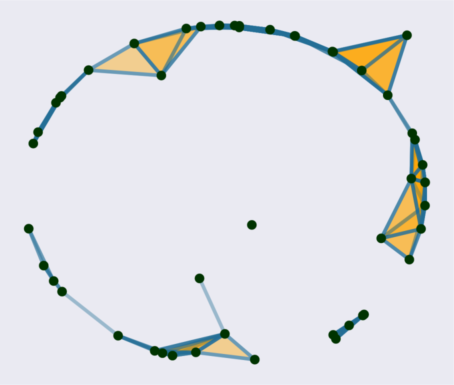


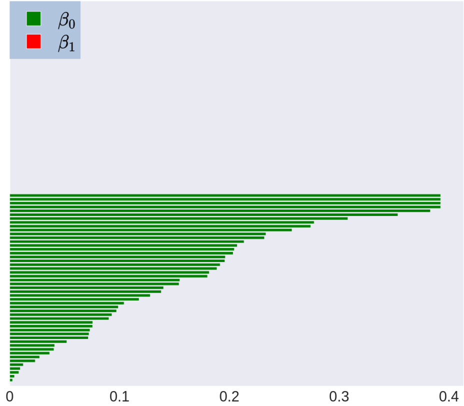
Persistent Homology and Barcodes



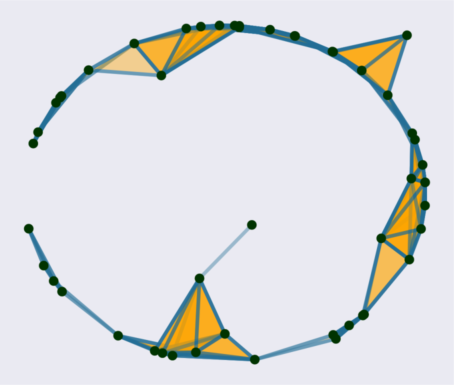


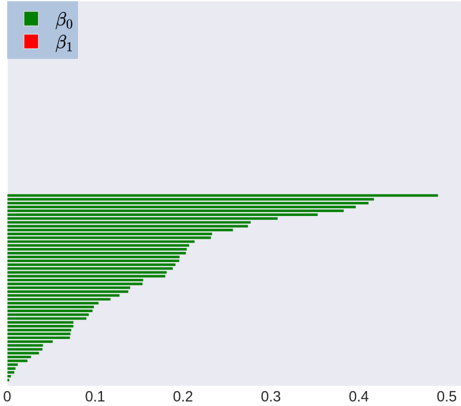
Persistent Homology and Barcodes



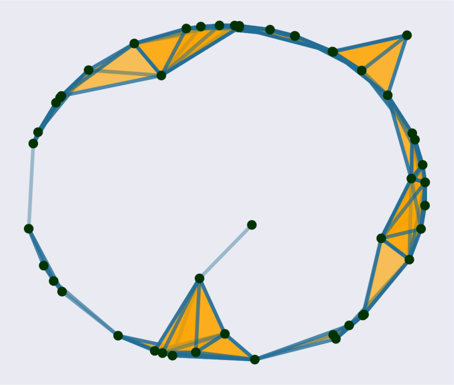


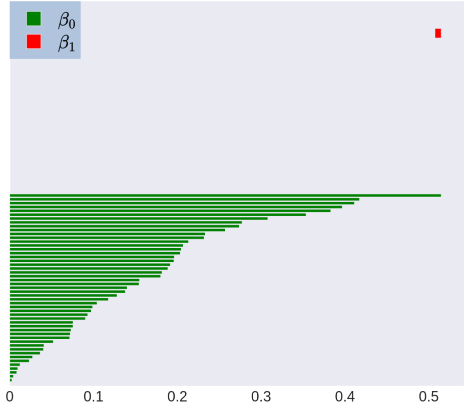
Persistent Homology and Barcodes



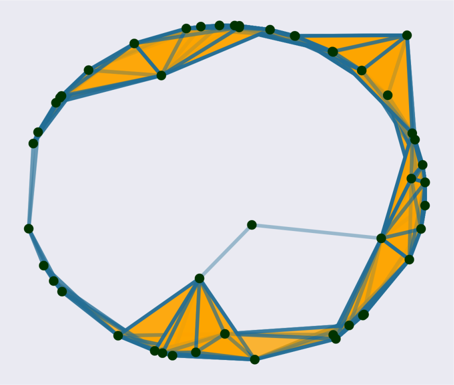


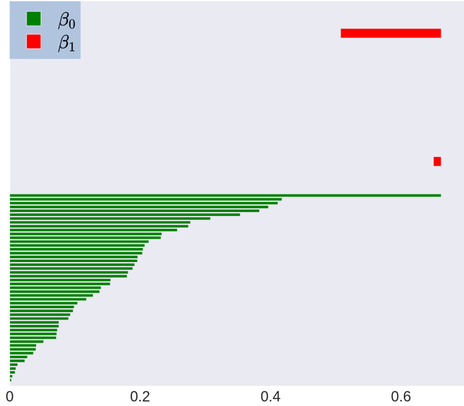
Persistent Homology and Barcodes



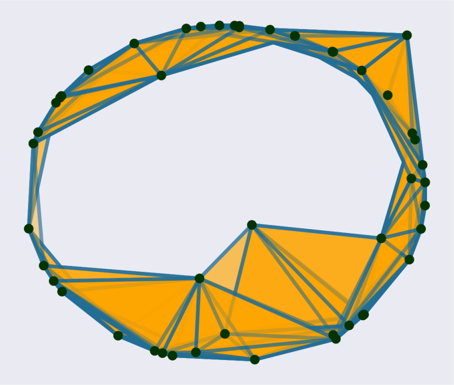


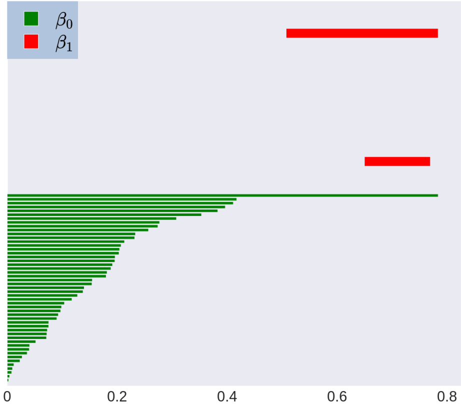
Persistent Homology and Barcodes



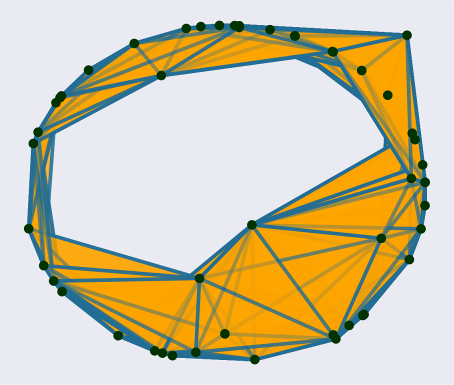


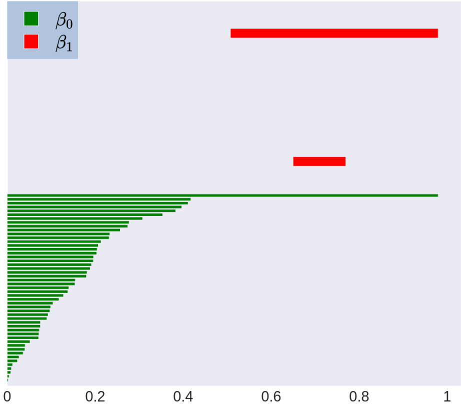
Persistent Homology and Barcodes



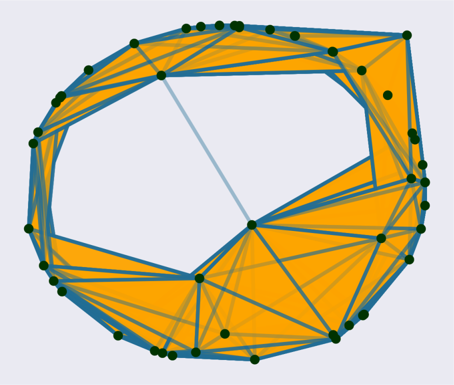


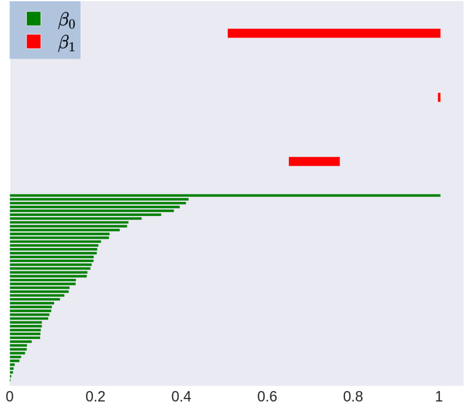
Persistent Homology and Barcodes



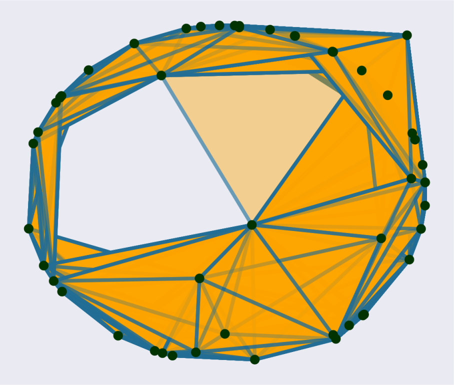


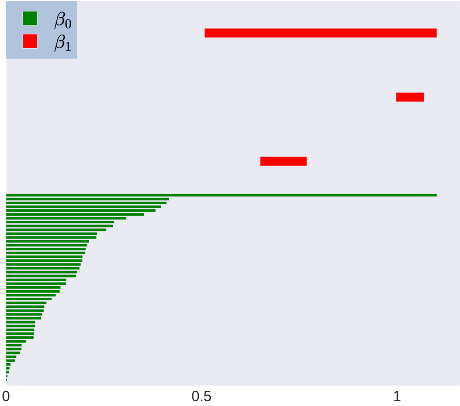
Persistent Homology and Barcodes



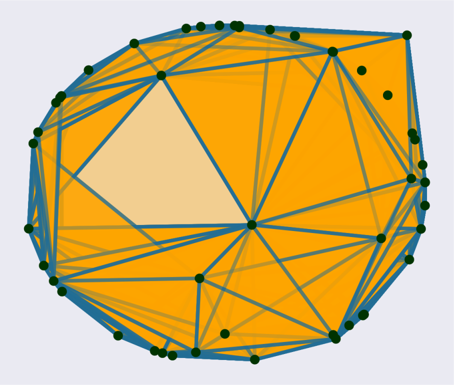


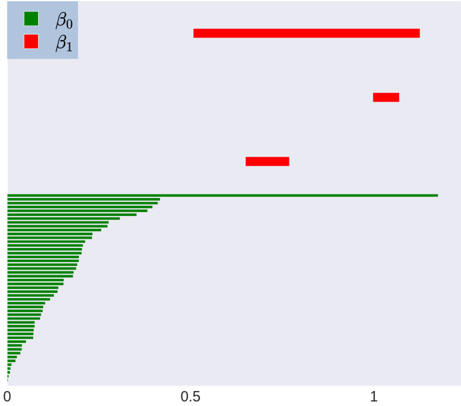
Persistent Homology and Barcodes





Persistent Homology and Barcodes





Persistent Homology

Filtered Vietoris-Rips complex

$$R_{\varepsilon_1} \hookrightarrow R_{\varepsilon_2} \hookrightarrow \cdots \hookrightarrow R_{\varepsilon_i} \hookrightarrow \cdots \hookrightarrow R_{\varepsilon_j} \hookrightarrow \cdots\cdots \hookrightarrow R_{\varepsilon_{max}}$$

After applying the homology functor,

$$H_k(R_{\varepsilon_1}) \to H_k(R_{\varepsilon_2}) \to \cdots \to H_k(R_{\varepsilon_i}) \to \cdots \to H_k(R_{\varepsilon_j}) \to \cdots \to H_k(R_{\varepsilon_{max}})$$

For every pair $\varepsilon_i, \varepsilon_j$

$$\psi^k_{\varepsilon_i,\varepsilon_j} : H_k(R_{\varepsilon_i}) \to H_k(R_{\varepsilon_j})$$

Definition

The $k-$th persistent homology group is

$$PH_k := \operatorname{im} \psi^k_{\varepsilon_i,\varepsilon_j}.$$

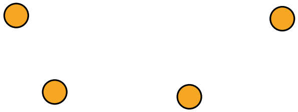
Go Forward

Persistent Homology - Big Picture

Data	Point Cloud	$P \subset \mathbb{R}^d$
↓	↓ Simplicial Technology	
Topology	Topological Spaces (Filtration)	$K\colon \ \mathbf{R} \to \mathbf{Top}$
↓	↓ Homology	
Algebra	Vector Spaces (Persistence module)	$M\colon \ \mathbf{R} \to \mathbf{Vect}$
↓	↓ Structure Theorem	
Combinatoric	Intervals (Persistence barcode)	$B\colon \ \mathbf{R} \to \mathbf{Mch}$

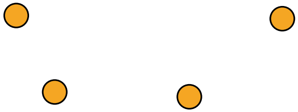
Relation between homology classes

How do we connect points ?



Relation between homology classes

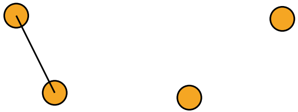
How do we connect points ?



$$\partial_1 = \begin{bmatrix} 0 \end{bmatrix} \qquad \partial_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Relation between homology classes

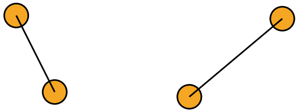
How do we connect points ?



$$\partial_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \partial_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Relation between homology classes

How do we connect points ?



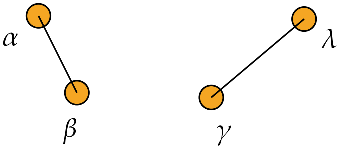
$$\partial_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \qquad \partial_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Relation between homology classes

How do we connect points ?

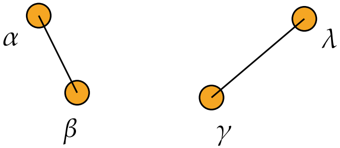
$$D = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Rank = 2



Relation between homology classes

How do we connect points ?



$$D = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

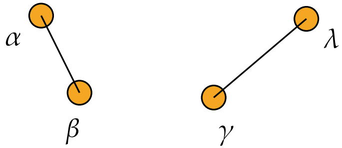
Rank = 2

$$D_{\alpha,\beta} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

Rank = 2

Relation between homology classes

How do we connect points ?



$$D = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Rank = 2

$$D_{\alpha,\beta} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

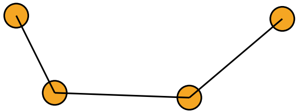
Rank = 2

$$D_{\alpha,\gamma} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

Rank = 3

Relation between homology classes

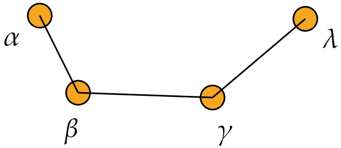
How do we connect points ?



$$\partial_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad \partial_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Relation between homology classes

How do we connect points ?



$$D = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

Rank = 3

$$D_{\beta,\gamma} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

Rank = 3

$$D_{\alpha,\gamma} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

Rank = 3

Homological Distance

1. Given a point cloud P, construct the Rips Complex

$$R_\varepsilon(P) = \{\sigma \subset P \mid \|x - y\| \leq \varepsilon, \text{ for all } x, y \in \sigma\}$$

2. $R_{\varepsilon_1} \subseteq R_{\varepsilon_2}$ for a given pair ε_1 and ε_2 values with $\varepsilon_1 \leq \varepsilon_2$
3. For the pair $\varepsilon_1 \leq \varepsilon_2$ we have the natural map in homology,

$$\psi^k_{\varepsilon_1,\varepsilon_2} : H_k(R_{\varepsilon_1}) \to H_k(R_{\varepsilon_2})$$

4. Take two cycles $\alpha,\ \beta \in H_k(\mathscr{R}_{\varepsilon_1})$
5. $\alpha' := \psi^k_{\varepsilon_1,\varepsilon_2}(\alpha)$ and $\beta' := \psi^k_{\varepsilon_1,\varepsilon_2}(\beta)$ by padding α and β with suitable number of 0's.

Homological Distance

- Add α', β' and α', β' together to the differential matrix $\mathscr{D}^t$ at ε_2.
- Calculate rank of the differential matrices.
- Check whether $\psi^k_{\varepsilon_1,\varepsilon_2}(\alpha)$ and $\psi^k_{\varepsilon_1,\varepsilon_2}(\beta)$ are linearly dependent.

Homological Distance

k-th homological cophenetic distance $D_k(\alpha,\beta)$ between homology classes α and β is defined as

$$\inf\left\{\eta-\varepsilon \geq 0 \mid \psi^k_{\varepsilon,\eta}(\alpha), \psi^k_{\varepsilon,\eta}(\beta) \text{ non-zero and lin. dep.}\right\}$$

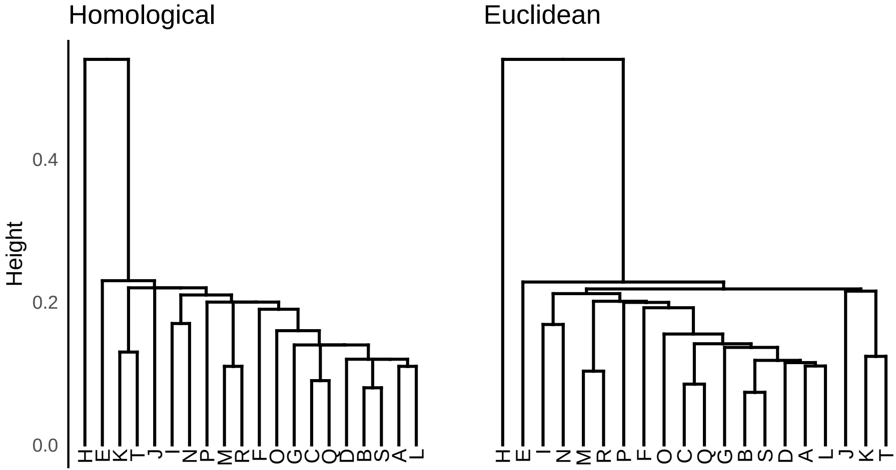
Experiments

A synthetic point cloud

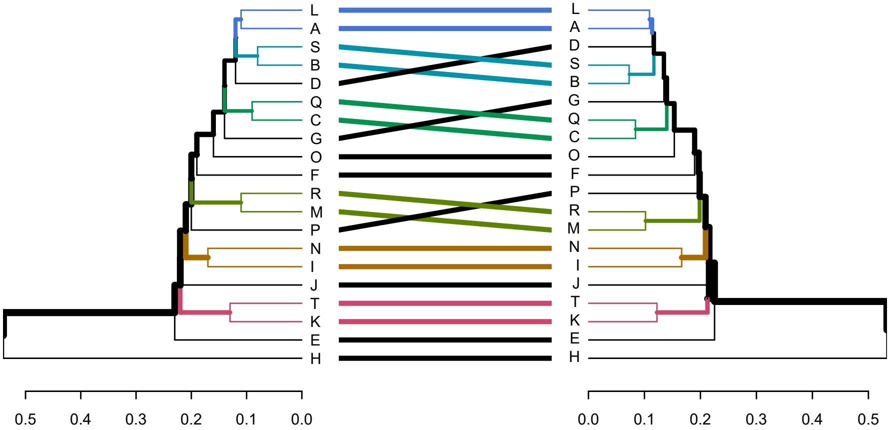
- $D \in \mathbb{R}^2$ and $|D| = 20$,
- Uniform distribution over $[0,1)$,
- Labeled with the first 20 letters.

Input: A point cloud D, $|D| = 20$ and
a list $\varepsilon = \{\varepsilon_1 = 0, 0.05, \ldots, 0.95, \varepsilon_{max} = 1\}$.
Output: Dendrograms
begin
  HomDist $= [\ \ ]_{20\times 20}$
  $\mathscr{R}_\varepsilon(P) \longleftarrow$ Vietoris-Rips filtration ;
  for *each* ε **do**
    for *every* $\alpha_i, \alpha_j \in H_0(\mathscr{R}_\varepsilon)$ **do**
      HomDist$_{i,j} \longleftarrow \inf\{\varepsilon \,|\, \textit{check lin. dep.}\}$;
    end
  end
  $E(D) \longleftarrow$ EuclideanDist(D);
  $Dend_1 \longleftarrow$ HierarchicalClustering(HomDist(D)) ;
  $Dend_2 \longleftarrow$ HierarchicalClustering$(E(D))$;
  Compare$(Dend_1, Dend_2)$
end

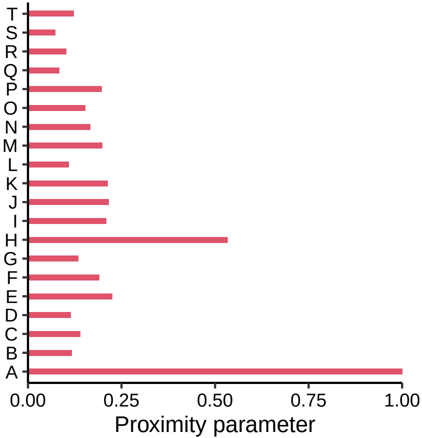
Two dendrograms with single-linkage

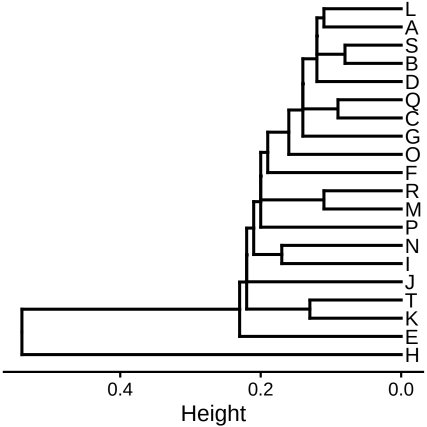


Tanglegram with the entanglement of 0.01



Barcode and Enriched Barcode





Turkish Cities and Mantel Statistics

- 24 Türkiye cities
- Single-linkage
- Different metrics
- Dendrograms

Metrics	Bray-Curtis	Cosine	Manhattan	Euclidean	Minkowski	Homological
Bray-Curtis	1.00	0.64	0.96	0.90	0.90	0.90
Cosine		1.00	0.61	0.52	0.69	0.59
Manhattan			1.00	0.96	0.87	0.97
Euclidean				1.00	0.75	0.98
Minkowski					1.00	0.78
Homological						1.00

Turkish Cities and Mantel Statistics

- 24 Türkiye cities
- Single-linkage
- Different metrics
- Dendrograms

Metrics	Bray-Curtis	Cosine	Manhattan	Euclidean	Minkowski	Homological
Bray-Curtis	1.00	0.64	0.96	0.90	0.90	0.90
Cosine		1.00	0.61	0.52	0.69	0.59
Manhattan			1.00	0.96	0.87	0.97
Euclidean				1.00	0.75	0.98
Minkowski					1.00	0.78
Homological						1.00

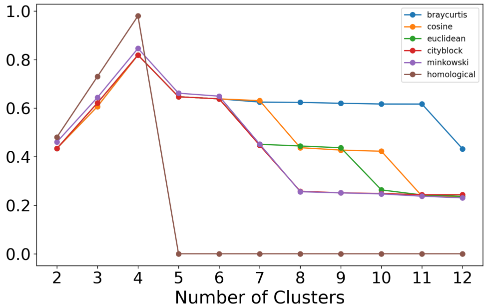
Datasets

Table: Datasets used and their properties.

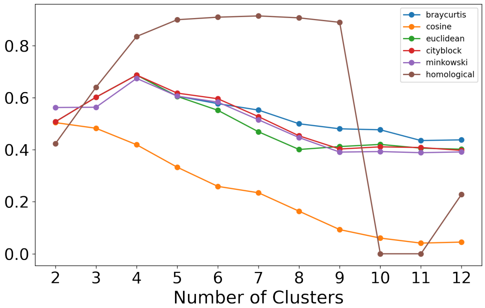
Dataset	#Instances	#Attributes	Supervised	#Classes
Turkish Cities	82	2	No	-
Iris	150	4	Yes	3
Cancer Coimbra	116	10	Yes	2
Synthetic (total separation)	100	100	Yes	4
Synthetic (with mixture)	100	2	Yes	4

Silhouette Scores

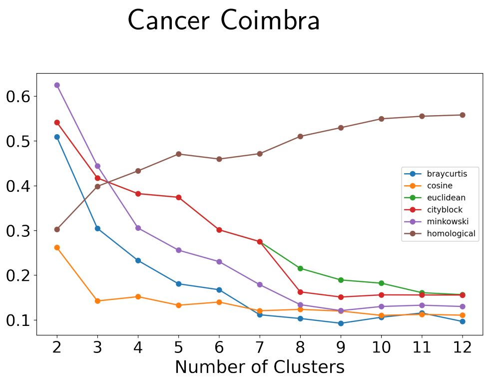
Synthetic Perfect

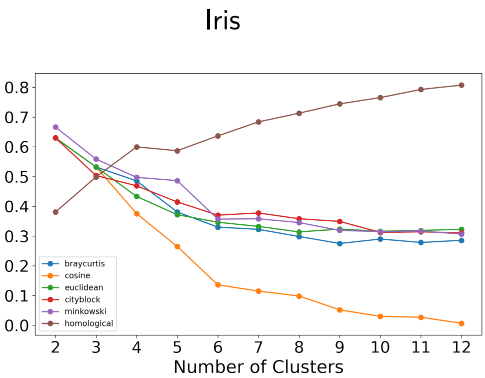


Synthetic Mixed



Silhouette Scores





A comparison of metrics on Mixed Synthetic datasets.

Synthetic Dataset

Metric	F1		Acc.		Hom.		Comp.		M.Info		Rand	
Bray-Curtis	1.00	A	1.00	A	1.00	A	1.00	A	1.38	A	1.00	A
Cosine	0.83	A	0.91	A	0.72	C	0.77	S	1.38	C	0.87	A
Manhattan	1.00	S	1.00	S	1.00	S	1.00	S	0.99	S	1.00	S
Euclidean	1.00	A	1.00	A	1.00	A	1.00	A	1.38	A	1.00	A
Minkowski	1.00	C	1.00	C	1.00	C	1.00	C	1.38	C	1.00	C
Homological	0.98	A	0.99	A	0.95	A	1.00	S	1.31	A	0.98	A

A comparison of metrics on Cancer datasets.

Real Dataset

Metric	F1		Acc.		Hom.		Comp.		M.Info		Rand	
Bray-Curtis	0.56	S	0.56	S	0.02	A	0.14	S	0.02	A	0.50	S
Cosine	0.55	C	0.55	C	0.01	S	0.12	S	0.01	S	0.50	C
Manhattan	0.53	S	0.53	S	0.02	A	0.13	A	0.02	A	0.50	S
Euclidean	0.54	W	0.54	W	0.02	A	0.13	A	0.02	A	0.50	W
Minkowski	0.53	S	0.53	S	0.02	A	0.13	A	0.02	A	0.50	S
Homological	0.61	W	0.61	W	0.03	W	1.00	S	0.02	W	0.52	W

BAP, TÜBİTAK, Michigan State University, Other Works

BASARIM2022

Classification of Stochastic Processes with Topological Data Analysis

İsmail Güzel
Mathematical Engineering
İstanbul Technical University
İstanbul, Türkiye
iguzel@itu.edu.tr

Atabey Kaygun
Mathematical Engineering
İstanbul Technical University
İstanbul, Türkiye
kaygun@itu.edu.tr

SIAM Data Mining: TDA, ML

A Case Study on Identifying Bifurcation and Chaos with CROCKER Plots

İsmail Güzel * Elizabeth Munch † Firas Khasawneh ‡

Abstract

The CROCKER plot is a coarsened but easy to visualize representation of the data in a one-parameter varying family of persistence barcodes. In this paper, we use the CROCKER plot to view changes in the persistence under a varying bifurcation parameter. We perform experiments to support our methods using the Rössler and Lorenz system and show the relationship with common methods for bifurcation analysis such as the Lyapunov exponent.

Erzincan Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
2022, 15(ÖZEL SAYI I), 1-13
ISSN: 1307-9085, e-ISSN: 2149-4584
Araştırma Makalesi

Erzincan University
Journal of Science and Technology
2022, 15(SPECIAL ISSUE I), 1-13
DOI: 10.18185/erzifbed.1199660
Research Article

Attitudes and Behaviors of Turkish Consumers Regarding the Olive Oil Consumption

Ümit ALTUNTAŞ[1,2*], İsmail GÜZEL[3], Özlem YILMAZ[4], Sibel ULUATA[5],

Beraat ÖZÇELİK[1]

[1]Istanbul Technical University, Department of Food Engineering, 34469, Istanbul, Türkiye
[2]Gümüşhane University, Department of Food Engineering, 29100, Gümüşhane, Türkiye
[3]Istanbul Technical University, Mathematical Engineering Department, 34469, Istanbul, Türkiye
[4]Bayburt University, Hotel, Restaurant and Catering Department, 69000, Bayburt, Türkiye
[5]Inonu University, Department of Food Engineering, 44280, Malatya, Türkiye

Geliş / Received: 04/11/2022, Kabul / Accepted: 09/12/2022

molecules MDPI

Article

Phenolic Constituents, Antioxidant and Antimicrobial Activity and Clustering Analysis of Propolis Samples Based on PCA from Different Regions of Anatolia

Ümit Altuntaş [1,2,*], İsmail Güzel [3] and Beraat Özçelik [1,4]

1 Department of Food Engineering, İstanbul Technical University, 34469 İstanbul, Turkey
2 Department of Food Engineering, Gümüşhane University, 29000 Gümüşhane, Turkey
3 Department of Mathematical Engineering, İstanbul Technical University, 34469 İstanbul, Turkey
4 Bioactive Research & Innovation, Food Manuf. Indust. Trade Ltd., Teknokent ARI-3, 34469 Istanbul, Turkey
* Correspondence: ualtuntas@itu.edu.tr

BAP, TÜBİTAK, Michigan State University, Other Works

BASARIM2022

Classification of Stochastic Processes with
Topol

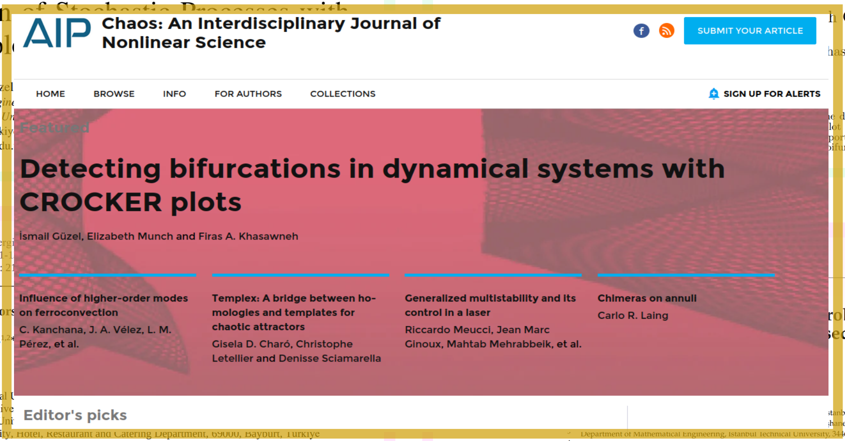
İsmail Güzel
Mathematical Engine
İstanbul Technical Ur
İstanbul, Türkiy
iguzel@itu.edu.

SIAM Data Mining: TDA, ML

h CROCKER Plots

hasawneh ‡

e data in a one-parameter
lot to view changes in the
port our methods using the
bifurcation analysis such as



Erzincan Üniversitesi
Fen Bilimleri Enstitüsü Dergi
2022, 15(ÖZEL SAYI I), 1-1
ISSN: 1307-9085, e-ISSN: 21
Araştırma Makalesi

Attitudes and Behavior

Ümit ALTUNTAŞ[1,2*]

[1]Istanbul Technical U
[2]Gümüşhane Unive
[3]Istanbul Technical Uni
[4]Bayburt University, Hotel, Restaurant and Catering Department, 69000, Bayburt, Türkiye
[5]Inonu University, Department of Food Engineering, 44280, Malatya, Türkiye

Geliş / Received: 04/11/2022, Kabul / Accepted: 09/12/2022

MDPI

robial Activity
sed on PCA

stanbul, Turkey
hane, Turkey
Department of Mathematical Engineering, Istanbul Technical University, 34469 İstanbul, Turkey
[4] Bioactive Research & Innovation, Food Manuf. Indust. Trade Ltd., Teknokent ARI-3, 34469 Istanbul, Turkey
[*] Correspondence: ualtuntas@itu.edu.tr

Future Work

On this thesis

- Apply real-world dataset for the first degree of homology.
- Visualization tools for the first degree of homology.
- Apply to the categorical dataset.
- Deal with problems about computational power and memory.

Other tasks

- Relation between Lyapunov exponent and persistent homology
- Two dimension bifurcation and CROCKER
- Classification Alpha-stable processes via TDA

Thank You!

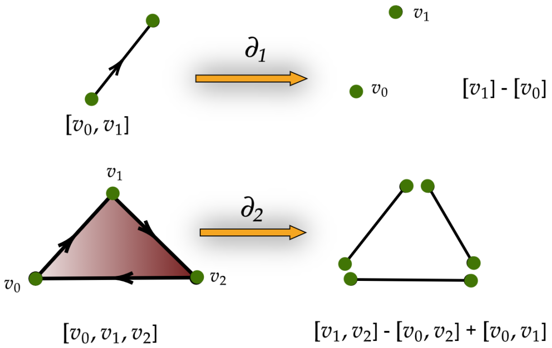
Seperation of complex

Chain Complex

A chain complex of a simplicial complex $\mathscr{K}$ is a sequence of abelian groups or modules $\mathscr{C}_k$ connected by homomorphisms $\partial_k : \mathscr{C}_k \to \mathscr{C}_{k-1}$ such that $\partial_{k-1} \circ \partial_k = 0$ for $k \in \mathbb{Z}$.

$$\dots \xrightarrow{\partial_{k+2}} \mathscr{C}_{k+1} \xrightarrow{\partial_{k+1}} \mathscr{C}_k \xrightarrow{\partial_k} \dots$$

$$\partial_k \sigma = \sum_{i=0}^{k} (-1)^i [v_0, v_1, \dots, \hat{v}_i, \dots, v_k]$$

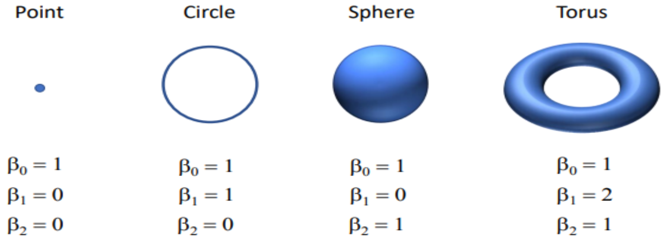


Homology and Betti Number

The k^{th} homology group of a simplicial complex K is defined by

$$H_k(K) := {}^{\ker(\partial_k)}\!\big/_{\mathrm{im}(\partial_{k+1})}.$$

The dimension of the k^{th} homology group of K is called the k^{th} *Betti number* $\beta_k(K)$.



Well-defined Persistence barcode

- Persistence modules

$$V_1 \to V_2 \to \cdots \to V_r \to \cdots \to V_s \to \cdots \to V_n$$

- Decompose persistence module into interval modules,

$$0 \to 0 \to \cdots \to 0 \to \mathbb{Z}_2 \to \cdots \to \mathbb{Z}_2 \to 0 \to \cdots 0$$

- A persistence module V indexed by $T \subset \mathbb{R}$ is $q - \textit{tame}$ if for any $v < s$ in T, the rank of the linear map $v_r^s : V_r \to V_s$ is finite.
- If $\mathbb{V}$ is a $q - \textit{tame}$ persistence module, then it has a well-defined persistence barcode.

Go Back

Combinatoric Possibilities

- Check whether $\psi^k_{\varepsilon_1,\varepsilon_2}(\alpha)$ and $\psi^k_{\varepsilon_1,\varepsilon_2}(\beta)$ are linearly independent by evaluating the rank of the differential matrix $\mathscr{D}$ at ε_2 with appending α', β' and α', β' together.
- $$\begin{aligned} r_\alpha &:= \operatorname{rank}(\mathscr{D}_\alpha) - \operatorname{rank}(\mathscr{D}) \\ r_\beta &:= \operatorname{rank}(\mathscr{D}_\beta) - \operatorname{rank}(\mathscr{D}) \\ r_{\alpha,\beta} &:= \operatorname{rank}(\mathscr{D}_{\alpha,\beta}) - \operatorname{rank}(\mathscr{D}) \end{aligned}$$

  with the following cases:

  $$\begin{cases} \alpha \text{ and } \beta \text{ both die} & \text{if } r_{\alpha,\beta} = 0, \\ \alpha \text{ and } \beta \text{ both live} & \text{if } r_{\alpha,\beta} = 2, \\ \alpha \text{ dies and } \beta \text{ lives} & \text{if } r_{\alpha,\beta} = 1 \text{ and } r_\alpha = 0, \\ \alpha \text{ lives and } \beta \text{ dies} & \text{if } r_{\alpha,\beta} = 1 \text{ and } r_\beta = 0, \\ \alpha \text{ and } \beta \text{ merge} & \text{if } r_{\alpha,\beta} = 1,\ r_\alpha = 1 \text{ and } r_\beta = 1. \end{cases}$$

Matroids

Definition

A partially ordered set is defined as an ordered pair $P = (X, \leq)$.
Here, X is called the ground set of P and $\leq$ is the partial order of P

Definition

A matroid $M = (S, \mathbb{I})$ is a finite ground set S together with a collection of sets $\mathbb{I} \subset 2^S$ satisfying

- Downward closed: $A \in \mathbb{I}$ and $B \subseteq A \Rightarrow B \in \mathbb{I}$
- Exchange property: $A,\ B, \in \mathbb{I}$ and $|B| < |A| \Rightarrow \exists x \in A \setminus B$ such that $\{x\} \bigcup B \in \mathbb{I}$.

Matroids

Terminology

- Independent set: $I \in \mathbb{I}$
- Circuit: Minimal dependent set of M
- Basis: Maximal independent set of M
- Span: Basis B and $B \subseteq \mathbb{I} \Rightarrow \mathbb{I}$ is a spanning set.

- Ground set $\mathbb{V}$: set of vectors spanning $\mathbb{R}^d$
- Independent set $\mathbb{I}$: bases of $\mathbb{R}^d$ in $\mathbb{V}$
- Matroid: $(\mathbb{V}, \mathbb{I})$

The Rank Function of a Matroid

Definition

Let M be a matroid on a finite ground set E. The rank $r(X)$ of a subset $X \subseteq E$ is the cardinality of the largest independent set contained in X. In other words

$$r(X) = \max\{|A| \in N \mid A \subseteq X \text{ and } A \in \mathscr{I}\}$$

Cobordisms

For two linearly independent pair of homology classes α and β in $H_1(R_\varepsilon)$, one can see $\psi^1_{\varepsilon,\eta}(\alpha)$ and $\psi^1_{\varepsilon,\eta}(\beta)$ are linearly dependent in $H_1(R_\eta)$. We, then, visualize that two classes α and β merged at time η .

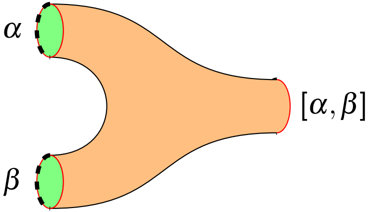


Figure: Cobordism in the merging case from $S^1 \sqcup S^1$ to S^1 representing two cycles α and β evolve in $[\alpha, \beta]$ from ε to η.