An Introduction to RIVET

CSE 5339 Lecture 4

February 1, 2022

Main Ideas

Recall:

- One-dimensional persistence modules decompose nicely into a direct sum of interval modules, $M = \oplus_{I \in \mathcal{I}} k^I$.
- Multiparameter persistence modules do not always decompose as direct sum of interval modules.
- [Carlsson and Zomorodian, 2009]: there is no complete, discrete invariant for multiparameter persistence modules.

Main Ideas

Recall:

- One-dimensional persistence modules decompose nicely into a direct sum of interval modules, $M = \oplus_{I \in \mathcal{I}} k^I$.
- Multiparameter persistence modules do not always decompose as direct sum of interval modules.
- [Carlsson and Zomorodian, 2009]: there is no complete, discrete invariant for multiparameter persistence modules.
- Many invariants exist for multiparameter persistence modules, losing some information but still capturing important features of the modules: see Lecture 3.
- RIVET [Lesnick and Wright, 2015b] computes and visualizes some of these invariants for 2D persistence modules.

RIVET

- RIVET: Rank Invariant Visualization and Exploration Tool. For M a 2D persistence module, RIVET computes:
  - Hilbert (dimension) function of M,
  - Fibered barcode of M and,
  - (multigraded) Betti numbers of M.
- In-depth mathematical explanations can be found at [Lesnick and Wright, 2015a].

Invariants: Hilbert Function

Definition (Hilbert Function)

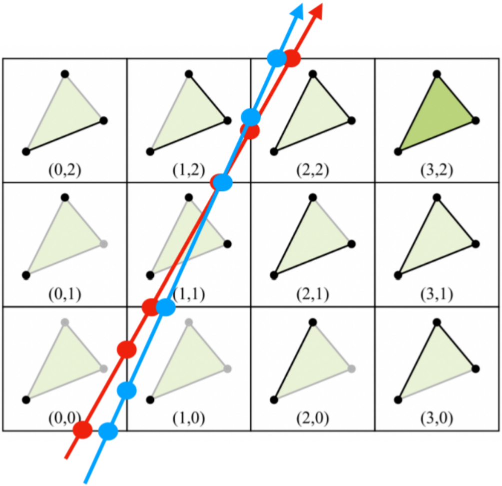
For M a persistence module over $\mathbb{R}^n$, the Hilbert function is $\mathrm{Hil}(M) : \mathbb{R}^n \to \mathbb{Z}_{\geq 0}$ given by

$$\mathrm{Hil}(M)(a) = \dim(M(a))$$

- Visualized in RIVET via shaded rectangles: darker shade $\to$ higher value of $\mathrm{Hil}(M)$.

Invariants: Fibered Barcode

- Recall *fibered barcode*: consists of barcodes of 1D persistence modules found by slicing M over all lines (of non-negative slope).



Invariants: Betti Numbers

- In 1D standard homology, Betti numbers are the same as the Hilbert function. In higher dimensional persistence, multigraded Betti numbers contain more information.
- For $i \geq 0$, M over $\mathbb{R}^n$, have multigraded Betti numbers $\xi_i(M) : \mathbb{R}^n \to \mathbb{Z}_{\geq 0}$.

Invariants: Betti Numbers

- In 1D standard homology, Betti numbers are the same as the Hilbert function. In higher dimensional persistence, multigraded Betti numbers contain more information.
- For $i \geq 0$, M over $\mathbb{R}^n$, have multigraded Betti numbers $\xi_i(M) : \mathbb{R}^n \to \mathbb{Z}_{\geq 0}$.
- RIVET focuses on $\xi_0, \ \xi_1$. For $a \in \mathbb{R}^n$, $\xi_0(M)(a)$, $\xi_1(M)(a)$ are the number of generators and relations, respectively, at index a in a minimal presentation for M.
- In RIVET, ξ_0, ξ_1 represented by green and red dots, respectively.

RIVET Pipeline

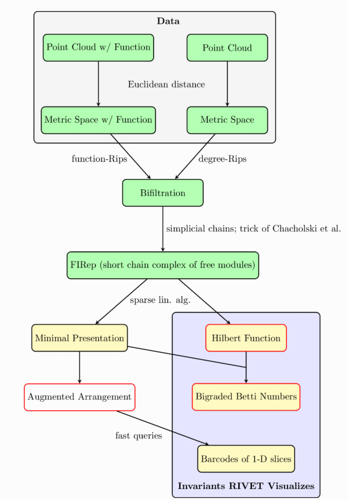


Figure: Pipeline for RIVET computation.

RIVET Inputs

- User inputs a point cloud (Euclidean coordinates) or a metric space (distance matrix), and inputs a function or chooses a built-in function.
- With underlying metric space X and $\gamma : X \to \mathbb{R}$ a function, the function-Rips bifiltration is given by:

$$FR(\gamma)_{a,b} := R(\gamma^{-1}(-\infty, a])_b.$$

Built-in Functions

RIVET has the following functions built-in:

- Gaussian density function:

$$\gamma(x) = C \sum_{y \in X} e^{\frac{-d(x,y)^2}{2\sigma}}$$

- C normalization constant, σ "standard deviation". Higher values for "central" points and points in denser regions.

Built-in Functions

RIVET has the following functions built-in:

- Gaussian density function:

$$\gamma(x) = C \sum_{y \in X} e^{\frac{-d(x,y)^2}{2\sigma}}$$

- C normalization constant, σ "standard deviation". Higher values for "central" points and points in denser regions.
- Eccentricity function:

$$\gamma(x) = \left(\frac{\sum_{y \in X} d(x,y)^q}{|X|} \right)^{\frac{1}{q}}$$

- $q \in [1, \infty)$ selectable parameter. Higher values for points in exterior of space.

Built-in Functions

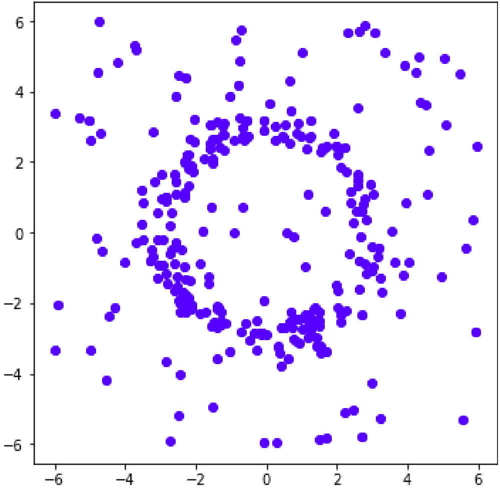
- Ball density function:

$$\gamma(x) = C \cdot (\#\ \text{points in}\ X\ \text{within distance}\ r\ \text{of}\ x)$$

- C normalization constant. Higher values for points in denser regions.

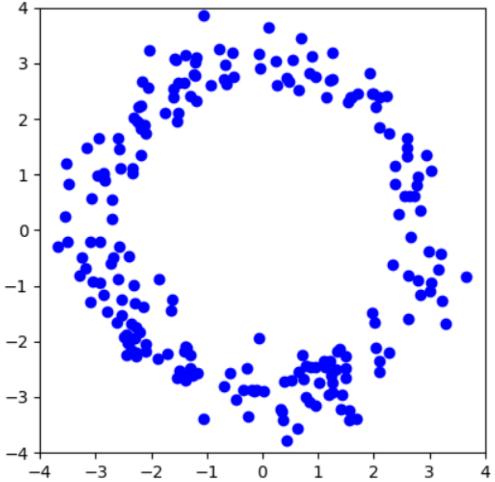
Built-in Functions

Example of ball density function:



Built-in Functions

γ ball density function with $C = 1$, $r = 1$, $\gamma^{-1}([10, \infty))$

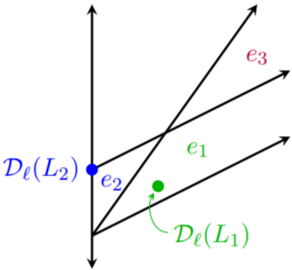


Augmented Arrangement

- After inputs, RIVET computes 2D persistence module. Then computes *augmented arrangement* for fast querying of fibered barcode.

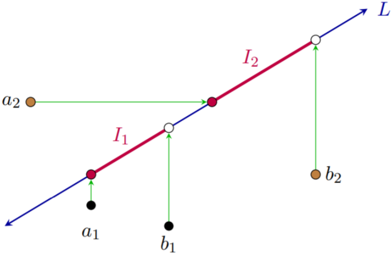
Augmented Arrangement

- After inputs, RIVET computes 2D persistence module. Then computes *augmented arrangement* for fast querying of fibered barcode.
- RIVET introduces the augmented arrangement $A^{\star}(M)$. $A^{\star}(M)$ consists of a decomposition of the plane into 2-cells, and a collection of pairs $T^e = \{(a,b),\ a,b \in \mathbb{R}^2\}$ for each 2-cell e.



Augmented Arrangement

- Idea: when given a line L, determine a 2-cell e, then project each pair $(a, b) \in T^e$ onto L either rightwards or upwards:



- $\mathcal{B}(M^L) = \{\{(p(a), p(b)) \mid (a, b) \in T^e\}\}$. Could occur that $p(a) \geq p(b)$, which yields no corresponding interval in barcode.

Runtime Analysis

To analyze runtime, we need the following constants:

- κ, which is a constant based on the size of the support of ξ_i. n, the size of a bifiltration, which is the number of minimal indices at which a simplex appears.

Runtime Analysis

To analyze runtime, we need the following constants:

- κ, which is a constant based on the size of the support of ξ_i. n, the size of a bifiltration, which is the number of minimal indices at which a simplex appears.

Theorem

For $\mathcal{F}$ a bifiltration of size n, and $M = H_i(\mathcal{F})$, computation of $A^\star(M)$ from $\mathcal{F}$ takes $O(n^3\kappa + n\kappa^2 + \kappa^2 \log \kappa)$ and storage requires $O(n^2 + n\kappa^2)$.

Innovation of $A^\star(M)$ gives us:

Runtime Analysis

To analyze runtime, we need the following constants:

- κ, which is a constant based on the size of the support of ξ_i. n, the size of a bifiltration, which is the number of minimal indices at which a simplex appears.

Theorem

For $\mathcal{F}$ a bifiltration of size n, and $M = H_i(\mathcal{F})$, computation of $A^\star(M)$ from $\mathcal{F}$ takes $O(n^3\kappa + n\kappa^2 + \kappa^2 \log \kappa)$ and storage requires $O(n^2 + n\kappa^2)$.

Innovation of $A^\star(M)$ gives us:

Theorem

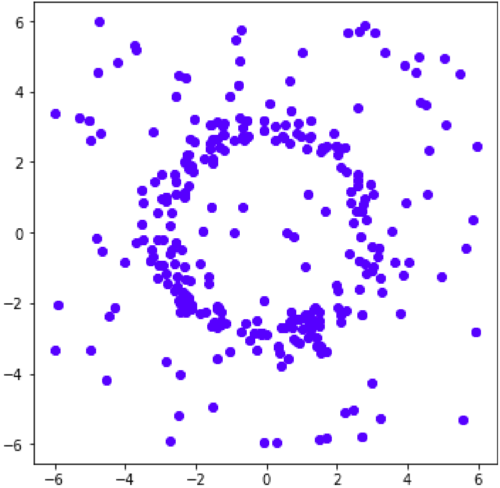
For most lines L of non-negative slope, we can query $A^\star(M)$ for $\mathcal{B}(M^L)$ in $O(\log \kappa + |\mathcal{B}(M^L)|)$ time.

Setting up RIVET

- RIVET is freely available from github at `https://github.com/rivetTDA/rivet`.
- Documentation: `https://rivet.readthedocs.io/en/latest/`
- RIVET is supported on Linux (Ubuntu) and Mac OS X. Can be run on Windows via WSL.

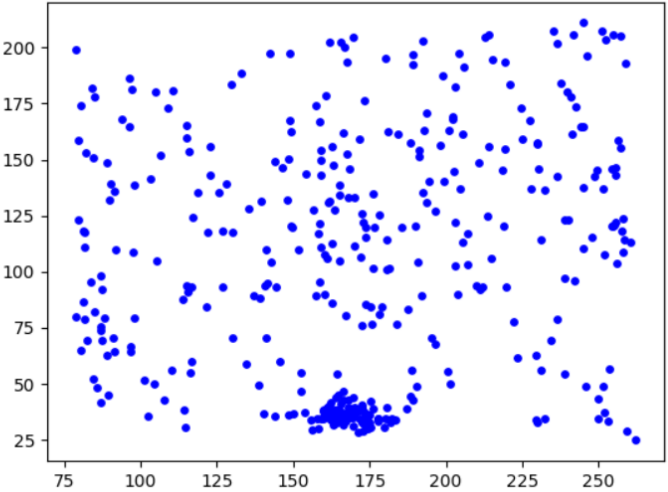
Example dataset

Noisy circle dataset comes with RIVET install:



Rat Motion Example:

Data on rat trajectory, `https://crcns.org/data-sets/hc/hc-2/`



References

Carlsson, G. and Zomorodian, A. (2009).
The theory of multidimensional persistence.
Discrete & Computational Geometry, 42(1):71–93.

Lesnick, M. and Wright, M. (2015a).
Interactive visualization of 2-d persistence modules.
arXiv preprint arXiv:1512.00180.

Lesnick, M. and Wright, M. (2015b).
Rivet tda.
https://github.com/rivetTDA/rivet.