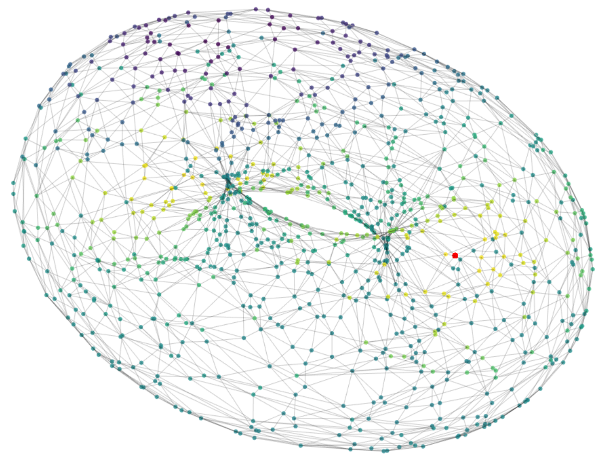
Topological Data Analysis and Persistence Theory



NSF/CBMS Conference

Valdosta State University

August 8-12, 2022

Peter Bubenik, University of Florida

Lecture 3: Combinatorics and Geometry

Outline:
1. Möbius Inversion and Persistence
2. Gromov-Hausdorff Stability
3. Nerve Theorem

Please interrupt me!!!

1. Möbius Inversion and Persistence

1.1 Intervals and Möbius Inversion

Let P be a poset.

interval in P

For $i, j \in P$ with $i \leq j$, let $[i,j] = \{k \in P \mid i \leq k \leq j\}$.

Let $\mathrm{Int}\, P$ denote the set of intervals in P

We are interested in functions $f : \mathrm{Int}\, P \to \mathbb{Z}$

The zeta function $\zeta : \mathrm{Int}\, P \to \mathbb{Z}$ is given by the constant function 1.

If P has a largest element ω then any function $h : P \to \mathbb{Z}$ may be extended to $h : \mathrm{Int}\, P \to \mathbb{Z}$ by defining $h\,[x,y] = \begin{cases} h(x) & \text{if } y = \omega \\ 0 & \text{otherwise} \end{cases}$.

The set of functions, $\mathrm{Int}\, P \to \mathbb{Z}$, has a binary operator, $*$, called convolution, given by

$$f * g\,[x,y] = \sum_{c \in [x,y]} f(x,c)\, g(c,y)$$

Proposition The zeta function has an inverse with respect to $*$ μ, called the Möbius function, which may be defined/computed recursively.

1.2 The Rank function

Consider a persistence module $M : [n] \to \mathrm{Vect}$

$$M_0 \to M_1 \to M_2 \to \cdots \to M_n$$

We extend this to $M : [n+1] \to \mathrm{Vect}$, by $M_{n+1} := 0$

$$M_0 \to M_1 \to M_2 \to \cdots \to M_n \to 0$$

$(\{0, 1, \cdots, n+1\}, \leq)$

Consider half-open intervals in $[n+1]$:

$$[i, j) := \{ k \in [n+1] \mid i \leq k < j \}$$

Let $[n+1]^2_<$ denote the set of half-open intervals in $[n+1]$ with partial order given by inclusion

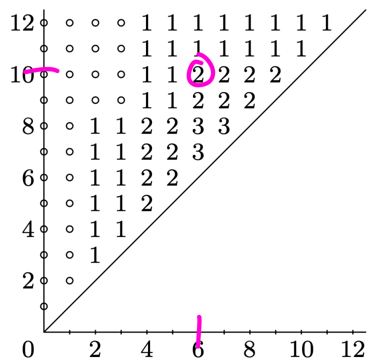
Define $\mathrm{Rank} : [n+1]^2_< \to \mathbb{Z}$

$$[i, j) \mapsto \operatorname{rank} M(i \leq j-1)$$

$\circ = 0$

Example $n = 11$

$\mathrm{Rank} : [12]^2_< \to \mathbb{Z}$



$\mathrm{Rank}\,[6, 10)$

$= \operatorname{rank} M(6 \leq 9)$

$= 2$

1.3 Persistence Diagrams via Möbius Inversion

Given a persistence module $M : [n] \to \mathrm{Vect}$

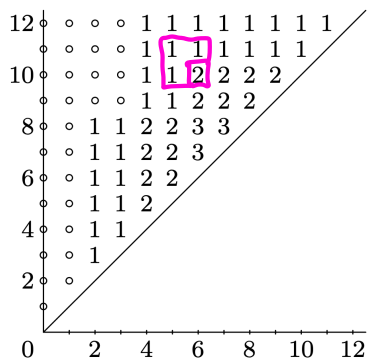
Define $PD : [n+1]^2_< \to \mathbb{Z}$ by

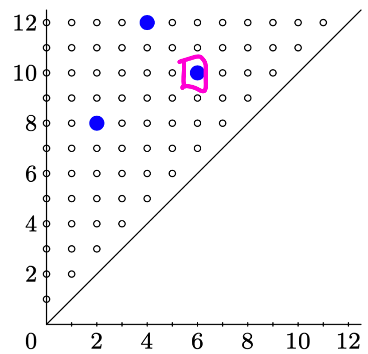
$$PD := \mu * \mathrm{Rank}\, M$$

It turns out that

$$PD\,[i,j) = \mathrm{Rank}\,[i,j) - \mathrm{Rank}\,[i-1,j) - \mathrm{Rank}\,[i,j+1) + \mathrm{Rank}\,[i-1,j+1)$$

$$\mathrm{Rank} \xrightarrow[\text{Inversion}]{\text{Möbius}} PD$$





$\bullet \equiv 1$

$\circ \equiv 0$

$$PD\,[6,10) = \mathrm{Rank}\,[6,10) - \mathrm{Rank}\,[5,10) - \mathrm{Rank}\,[6,11) + \mathrm{Rank}\,[5,11)$$

$$= 2 - 1 - 1 + 1$$

$$= 1$$

<u>Proposition</u> Given $K : [n] \to \mathrm{Simp}$. Let $M = H_j K$

Then PD is the persistence diagram of M

<u>Definition</u> Given $M : [n] \to \mathrm{Vect}$

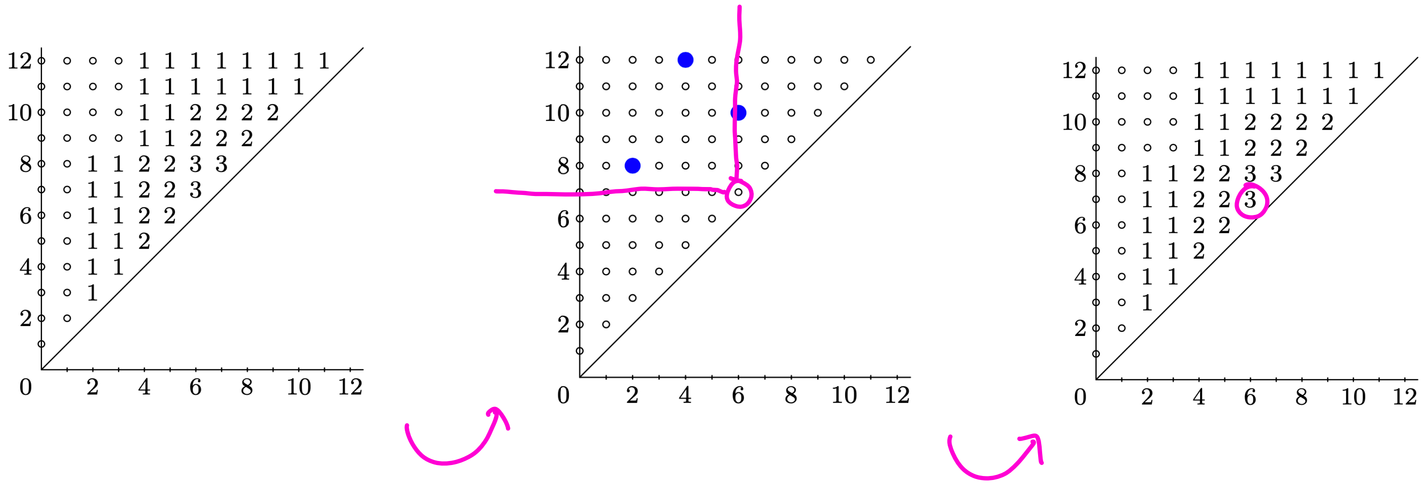
Call PD the persistence diagram of M.

Lemma $\quad \mathrm{Rank} = \zeta * \mu * \mathrm{Rank}$

$$= \zeta * \mathrm{PD}$$

That is, $\quad \mathrm{Rank}\,[i,j) = \sum_{\substack{[k,\ell) \\ k \le i < j \le \ell}} \mathrm{PD}\,[k,\ell)$

Rank $\qquad\qquad$ PD $\qquad\qquad$ Rank



$\mu * -$ $\qquad\qquad$ $\zeta * -$

1.4 Graded Rank and Graded Persistence Diagram

Let $k \in \mathbb{N}$. Let $\mathrm{Rank}_k : [n+1]^2_< \to \mathbb{Z}$

be given by $\mathrm{Rank}_k[i,j] = \begin{cases} 1 & \text{if } \mathrm{Rank}[i,j] \geq k \\ 0 & \text{otherwise} \end{cases}$

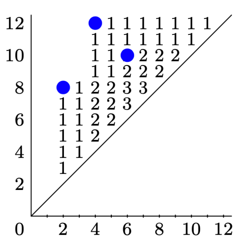
Lemma $\mathrm{Rank} = \sum_k \mathrm{Rank}_k$

Recall $PD = \mu * \mathrm{Rank}$

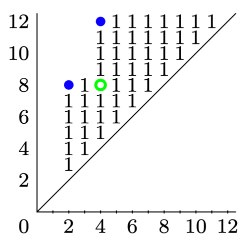
Define $PD_k = \mu * \mathrm{Rank}_k$

Call this the graded persistence diagram.

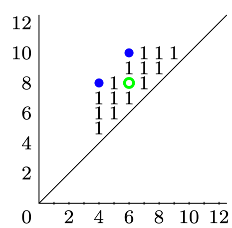
$\mathrm{Rank} \to PD$



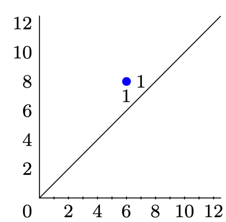
$\mathrm{Rank}_1 \to PD_1$



$\mathrm{Rank}_2 \to PD_2$



$\mathrm{Rank}_3 \to PD_3$



$\bullet \equiv 1$

$\circ \equiv -1$

Theorem $PD = \sum_k PD_k$

That is, the graded persistence diagram has more information.

2. Hausdorff and Gromov-Hausdorff Stability

2.1 Hausdorff stability of the Čech complex

Consider $X = \{x_1, \dots, x_n\} \subset \mathbb{R}^d$

$Y = \{y_1, \dots, y_m\} \subset \mathbb{R}^d$

Hausdorff distance:

$$d_H(X, Y) = \max\Big(\max_i \overbrace{d(x_i, Y)}^{= \min_j d(x_i, y_j)},\ \max_j d(y_j, X) \Big)$$

<u>Theorem</u> $W_\infty(\mathrm{Dgm}\, H\, \check{C}X,\ \mathrm{Dgm}\, H\, \check{C}Y) \leq d_H(X, Y)$

2.2 Gromov-Hausdorff stability of V-R complex

Consider (X, d_X) where $X = \{x_1, \dots, x_n\}$

(Y, d_Y) where $Y = \{y_1, \dots, y_m\}$

Consider (Z, d_Z) where $Z = X \sqcup Y$, $d_Z|_X = d_X$, $d_Z|_Y = d_Y$

Gromov-Hausdorff distance:

$$d_{GH}((X, d_X), (Y, d_Y)) = \inf_{(Z, d_Z)} d_H(X, Y)$$

<u>Theorem</u> $W_\infty(\mathrm{Dgm}\, H\, VR\, X,\ \mathrm{Dgm}\, H\, VR\, Y) \leq d_{GH}(X, Y)$

3. Nerve and Persistent Nerve Theorems

Consider a topological space Y with cover $\{A_i\}_{i \in I}$.

That is, $\forall i \quad A_i \subset Y$ and $Y = \bigcup_i A_i$

The nerve of this cover is the simplicial complex

K consisting of nonempty subsets of I, $\sigma = \{i_0, \dots i_n\}$

s.t. $A_{i_0} \cap \dots \cap A_{i_n} \neq \emptyset$.

Example $\quad X = \{x_1, \dots, x_n\} \subset \mathbb{R}^d$, $r \geq 0$

$Y_r = \bigcup_{i=1}^n B_r(x_i)$ has cover $\{B_r(x_i)\}_{i=1}^n$.

The nerve of this cover is the Čech complex $\check{C}_r X$.

Nerve Theorem $\quad$ In this example, $H(\check{C}_r X) \cong H Y_r$

simplicial homology $\quad$ singular homology.

Persistent Nerve Theorem $\quad$ In the above example,

the persistence modules $H \check{C}_\bullet X$ and $H Y_\bullet$

are isomorphic.