Generalized Persistence Diagrams for Persistence Modules over Posets

Woojin Kim

joint with Facundo Mémoli

The Ohio State University

UFTDA

01/30/2020

NSF IIS-1422400, CCF-1526513, DMS-1723003, CCF-1740761

Motivation

$F : \mathbb{Z} \to$ **vec** is given.

Let $i \leq j$ in $\mathbb{Z}$.

- $F(i \leq j)$: The linear map from $F(i)$ to $F(j)$.
- **rk**$F(i \leq j)$: The **rank** of $F(i \leq j)$.
- **barc**(F): The **barcode** of F, which is a multiset of intervals of $\mathbb{Z}$.

Fact

$$\mathbf{rk}F(i \leq j) = \#\, \{\!\!\{ I \in \mathbf{barc}(F) : [i,j] \subset I \}\!\!\}.$$

Motivation

$$\mathbf{rk}F(i \leq j) = \#\{\!\{I \in \mathbf{barc}(F) : [i,j] \subset I\}\!\}.$$

(# of $[1,2]$ in $\mathbf{barc}(F)$)

$$= \#\{\!\{I \in \mathbf{barc}(F) : [1,2] \subset I\}\!\} \quad - \#\{\!\{I \in \mathbf{barc}(F) : [1,2] \subsetneq I\}\!\}.$$

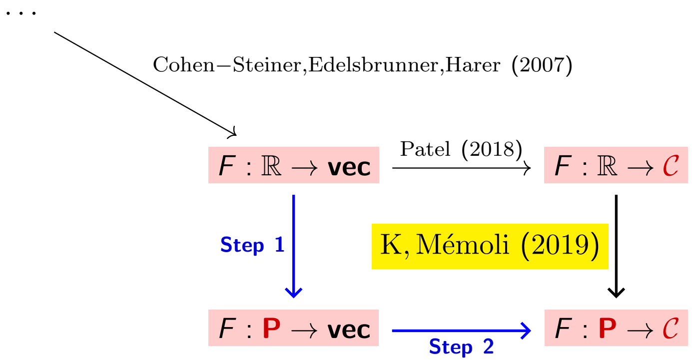
$$= \mathbf{rk}F(1 \leq 2) \quad - \quad (\mathbf{rk}F(0 \leq 2) + \mathbf{rk}F(1 \leq 3) - \mathbf{rk}F(0 \leq 3)).$$

Observation

$$[1,2] \subsetneq I \quad \Leftrightarrow \quad [0,2] \subset I \quad \text{or} \quad [1,3] \subset I.$$

Goal: **Use "this idea" to generalize the notion of persistence diagrams.**

Goal: Define persistence diagrams for generalized persistence modules $F : \mathbf{P} \to \mathcal{C}$ using **the principle of inclusion-exclusion**.



$\mathcal{C}$: a category (satisfying mild conditions).
$\mathbf{P}$: a poset (satisfying mild conditions).

Road map in Step 1

$$F : \mathbb{R} \to \textbf{vec}$$

$$\Big\downarrow \textbf{Step 1}$$

$$F : \mathbf{P} \to \textbf{vec}$$

We will define
1. Rank
2. Rank invariant
3. Persistence diagram

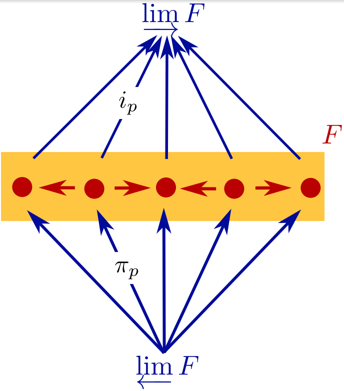
of $F : \mathbf{P} \to \textbf{vec}$.

Rank

Let $\mathbf{P}$ be a **connected** poset, i.e. $\forall p, q \in \mathbf{P}$ there exists a sequence $p = p_0, p_1, \ldots, p_n = q$ in $\mathbf{P}$ s.t. p_i and p_{i+1} are comparable.

Given $F : \mathbf{P} \to \mathbf{vec}$,

$$\mathbf{rank}(F) := \mathbf{rank}\left(\phi_F : \varprojlim F \to \varinjlim F\right)$$



$\phi_F = i_p \circ \pi_p$ for any $p \in \mathbf{P}$.

Studying the Limit-to-Colimit map in TDA is **not** a new idea.

1. Work by **Patel and MacPherson** circa 2012.
2. **Chambers and Letscher** *"Persistent Homology Over Directed Acyclic Graphs" (arXived in 2014, published in 2018)*.
3. **Botnan, Oppermann and Steen**, a work in progress. See Botnan's AATRN online seminar on 04/05/2017 (available on Youtube).

We assume that $\mathbf{P}$ is a finite and connected poset (for simplicity).

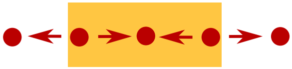
- **Hasse**($\mathbf{P}$): the Hasse diagram of $\mathbf{P}$.

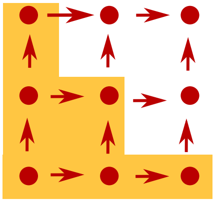
$$\text{For } p, q \in \mathbf{P}, \quad [p \to q] \Leftrightarrow p < q \text{ and no } r \text{ s.t. } p < r < q$$

- $\overline{\textbf{Hasse}}(\mathbf{P})$: Undirected version of **Hasse**($\mathbf{P}$).

Definition

A subset $K \subset \mathbf{P}$ will be called a **path-connected subposet** of $\mathbf{P}$ if the induced subgraph of $\overline{\textbf{Hasse}}(\mathbf{P})$ on K is path-connected.





Con($\mathbf{P}$): the set of all path-connected subposets of $\mathbf{P}$.

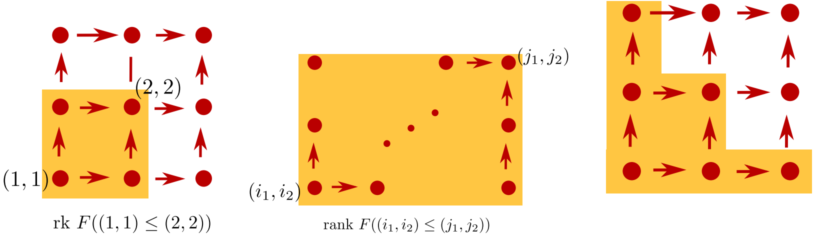
Con(P): the set of all path-connected subposets of **P**.

Rank invariant of $F : \mathbf{P} \to \mathbf{vec}$

$$\mathbf{rk}_F : \mathbf{Con}(\mathbf{P}) \to \mathbb{Z}_{\geq 0}$$

$$K \mapsto \mathbf{rank}(F|_K)$$

Remark. For multidimensional persistence, this definition is a refinement of the standard rank invariant.



We will see that

The rank invariant is a complete invariant for **interval decomposable persistence modules.**

Let **P** be a finite and connected poset.

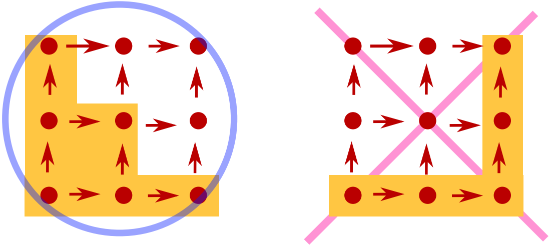
Definition (Botnan & Lesnick)

A non-empty subset $K \subset \mathbf{P}$ is called an **interval** of **P** if K is path-connected and <u>convex</u>, i.e.

$$[p, q \in K,\ r \in \mathbf{P} \text{ and } p \leq r \leq q] \Rightarrow r \in K.$$

$\mathbf{Int}(\mathbf{P})$: the set of all intervals of **P**.

Remark. In general, $\mathbf{Int}(\mathbf{P}) \subsetneq \mathbf{Con}(\mathbf{P})$.



If **P** is linear ($\bullet \longleftrightarrow \bullet \longleftrightarrow \bullet \longleftrightarrow \bullet$), then $\mathbf{Con}(\mathbf{P}) = \mathbf{Int}(\mathbf{P})$.

Let $K \in \mathbf{Int}(\mathbf{P})$. The interval module $I^K : \mathbf{P} \to \mathbf{vec}$ is defined as

$$I^K(p) = \begin{cases} k, & p \in K \\ 0, & \text{otherwise,} \end{cases} \qquad I^K(p \leq q) = \begin{cases} id_k, & p, q \in K \\ 0, & \text{otherwise.} \end{cases}$$

Definition

$F : \mathbf{P} \to \mathbf{vec}$ is called **interval decomposable** if there exists a multiset $\mathbf{barc}(F)$ of intervals of $\mathbf{P}$ such that

$$F \cong \bigoplus_{K \in \mathbf{barc}(F)} I^K.$$

Proposition 1 (K,Mémoli 19)

Assume that F is interval decomposable. For $K \in \mathbf{Con}(\mathbf{P})$,

$$\mathbf{rk}_F(K) = \# \{\!\{ I \in \mathbf{barc}(F) : K \subset I \}\!\}$$

Using this proposition, we will compute $\mathbf{barc}(F)$ from $\mathbf{rk}_F$.

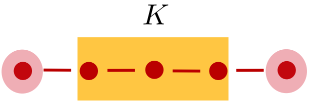
Let $K \in \mathbf{Con}(\mathbf{P})$.

Definition (Neighborhood of K)

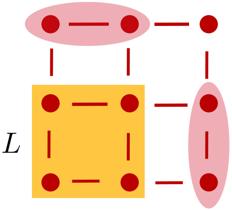
$$\mathbf{nbd}(K) := \{p \in \mathbf{P} \setminus K : \exists q \in K,\ p \text{ and } q \text{ are adjacent in } \overline{\mathbf{Hasse}}(\mathbf{P}).\}$$

Also, $o_K := |\mathbf{nbd}(K)|$ will be called the **perimeter** of K.

Example.



$o_K = 2$



$o_L = 4$

Definition (n-th entourage of K)

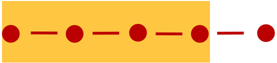
Let $K \in \mathbf{Con}(\mathbf{P})$.

$$K^n := \{L \subset \mathbf{P} : K \subset L \subset K \cup \mathbf{nbd}(K) \text{ with } |L \cap \mathbf{nbd}(K)| = n\}.$$

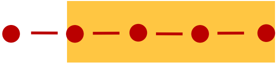
Example.

K

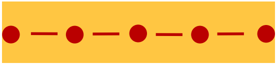
K^1:



and



K^2:



$K^n = \emptyset$ for $n > 2$.

Theorem 2 (K,Mémoli 19)

Assume that $F : \mathbf{P} \to \mathbf{vec}$ is interval decomposable. For $I \in \mathbf{Int}(\mathbf{P})$,

$$(\#\text{ of } I \text{ in } \mathbf{barc}(F)) =$$

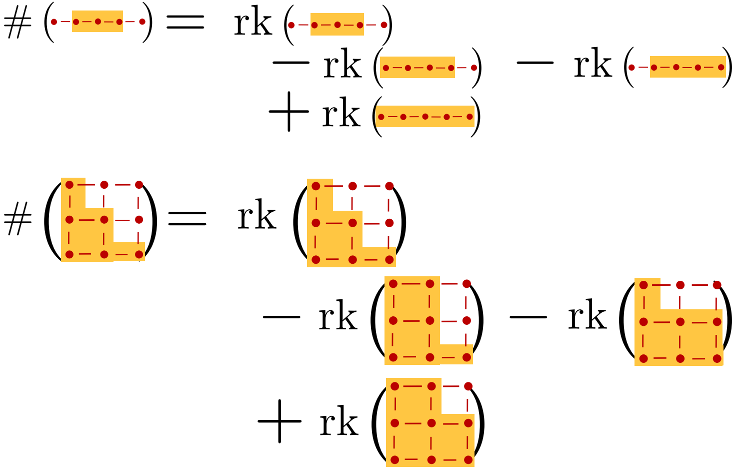
$$\mathbf{rk}_F(I) - \left(\sum_{K \in I^1} \mathbf{rk}_F(K) - \sum_{L \in I^2} \mathbf{rk}_F(L) - \ldots + (-1)^{o_I} \sum_{M \in I^{o_I}} \mathbf{rk}_F(M) \right).$$

By Proposition 1, $\mathbf{rk}_F(I) = \# \{\!\{ J \in \mathbf{barc}(F) : I \subset J \}\!\}$, and

$$\sum_{K \in I^1} \mathbf{rk}_F(K) - \sum_{L \in I^2} \mathbf{rk}_F(L) - \ldots - (-1)^{o_I} \sum_{M \in I^{o_I}} \mathbf{rk}_F(M)$$

is $\# \{\!\{ J \in \mathbf{barc}(F) : I \subsetneq J \}\!\}$.

Examples



Corollary 3

The rank invariant $\mathbf{rk}_F : \mathbf{Con}(\mathbf{P}) \to \mathbb{Z}_{\geq 0}$ is a complete invariant for interval decomposable persistence modules $F : \mathbf{P} \to \mathbf{vec}$.

Theorem 4 (K,Mémoli 19)

Let $F : \mathbf{P} \to \mathbf{vec}$. The map $\mathbf{dgm}_F : \mathbf{Con}(\mathbf{P}) \to \mathbb{Z}$ sending I to

$$\mathbf{rk}_F(I) - \sum_{K \in I^1} \mathbf{rk}_F(K) + \sum_{L \in I^2} \mathbf{rk}_F(L) + \ldots + (-1)^{o_I} \sum_{M \in K^{o_I}} \mathbf{rk}_F(M).$$

is the Möbius inversion of $\mathbf{rk}_F : \mathbf{Con}(\mathbf{P}) \to \mathbb{Z}_{\geq 0}$.

Proof idea. Compute the Möbius function of $(\mathbf{Con}(\mathbf{P}), \supset)$ – Make use of several classical results in [Rota1964].

Definition (Persistence diagram of $F : \mathbf{P} \to \mathbf{vec}$)

The map $\mathbf{dgm}_F : \mathbf{Con}(\mathbf{P}) \to \mathbb{Z}$ is the persistence diagram of F.

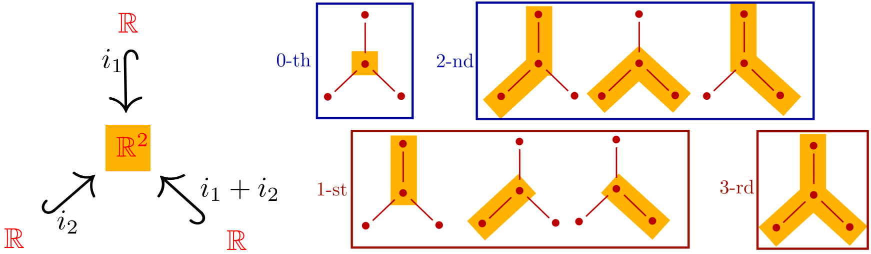
Remark.

1. We are adopting ideas in [Patel18].
2. $\mathbf{dgm}_F$ is defined regardless of the interval decomposability of F.

Corollary 5

Given $F : \mathbf{P} \to \mathbf{vec}$, if $\mathbf{dgm}_F \not\geq 0$, then F is **not** interval decomposable.

Ex(thanks to Patel).



Next step

$$\boxed{F : \mathbf{P} \to \mathbf{vec}} \xrightarrow[\text{Step 2}]{} \boxed{F : \mathbf{P} \to \mathcal{C}}$$

We will adopt other ideas in [Patel18].

Let $(\mathcal{C}, \Box)$ be any essentially small, symmetric monoidal, (finitely) **bicomplete** category with images –[Patel18].

Ex. $(\mathbf{vec}, \bigoplus)$, $(\mathbf{ab}, \bigoplus)$, $(\mathbf{set}, \bigsqcup)$.

$\mathcal{I}(\mathcal{C})$: the set of all isomorphism classes of $\mathcal{C}$.

$\mathcal{I}(\mathcal{C})$ is a symmetric monoid: $[A]\Box[B] = [A\Box B]$.

$\mathcal{A}(\mathcal{C})$: the group completion of $\mathcal{I}(\mathcal{C})$. (the **Grothendieck group** of $\mathcal{C}$).

$$\left(\mathcal{I}(\mathbf{vec}), \bigoplus\right) \cong (\mathbb{Z}_{\geq 0}, +) \qquad \Rightarrow \mathcal{A}(\mathbf{vec}) \cong (\mathbb{Z}, +).$$

$$(\mathcal{I}(\mathbf{set}), \sqcup) \cong (\mathbb{Z}_{\geq 0}, +) \qquad \Rightarrow \mathcal{A}(\mathbf{set}) \cong (\mathbb{Z}, +).$$

$$\left(\mathcal{I}(\mathbf{ab}), \bigoplus\right) \cong \left(\mathbb{Z}_{\geq 0} \oplus \bigoplus_{(m,p)} \mathbb{Z}_{\geq 0}, +\right) \qquad \Rightarrow \mathcal{A}(\mathbf{ab}) \cong \left(\mathbb{Z} \oplus \bigoplus_{(m,p)} \mathbb{Z}, +\right).$$

Examples of $\mathbf{P} \to \mathcal{C}$

1. $F : \mathbf{P} \to \mathbf{ab}$: by applying $\mathrm{H}(-;\mathbb{Z})$ to a **P**-indexed filtration. —Torsion in homology group is captured. **(One of Patel's motivations)**

2. $F : \mathbb{Z} \to \mathbf{set}$ arises from **merge trees**.

Let $\mathbb{ZZ} = \cdots \quad \bullet \longrightarrow \bullet \longleftarrow \bullet \longrightarrow \bullet \longleftarrow \bullet \quad \cdots$

3. $F : \mathbb{ZZ} \to \mathbf{set}$ arises from **Reeb graphs**. [de Silva, Munch, Patel16], [Curry,Patel16].

Let $F : \mathbf{P} \to \mathcal{C}$.

Definition (Rank of $F : \mathbf{P} \to \mathcal{C}$)

$$\mathbf{rank}(F) := \left[\mathbf{im}\left(\phi_F : \varprojlim F \to \varinjlim F\right)\right] \in \mathcal{I}(\mathcal{C}).$$

Definition (Rank invariant of F)

$$\mathbf{rk}_F : \mathbf{Con}(\mathbf{P}) \to \mathcal{I}(\mathcal{C})$$

$$K \mapsto \mathbf{rank}(F|_K).$$

Definition (Persistence diagram of F)

$$\mathbf{dgm}_F : \mathbf{Con}(\mathbf{P}) \to \mathcal{A}(\mathcal{C})$$

$K \mapsto$

$$\mathbf{rk}_F(K) - \sum_{L \in K^1} \mathbf{rk}_F(L) + \sum_{M \in K^2} \mathbf{rk}_F(M) + \ldots + (-1)^{o_K} \sum_{N \in K^{o_K}} \mathbf{rk}_F(N).$$

Open questions and future work

1. For which $\mathbf{P}$ and $\mathcal{C}$,

$$[F : \mathbf{P} \to \mathcal{C}] \longrightarrow [\mathbf{dgm}_F]$$

   is stable? For which metrics?

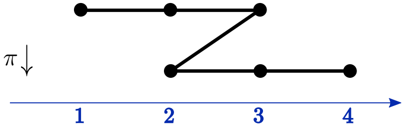
   Known: [Classics], [McCleary,Patel18], [Botnan,Lesnick18]

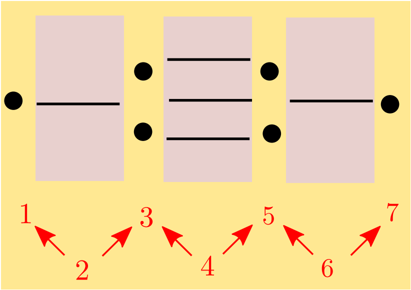
2. For persistence modules $\mathbf{P} \to \mathbf{vec}$ which are not interval decomposable, how faithful is our rank invariant or persistence diagram?
3. Characterization of dynamic metric spaces.

Thank you for paying attention.

https://arxiv.org/abs/1810.11517

Example.





P is a finite, connected poset.

Proposition 1(K,Mémoli 19)

Assume that F is interval decomposable. For $K \in \mathbf{Con}(\mathbf{P})$,

$$\mathbf{rk}_F(K) = (\#\text{ of intervals in } \mathbf{barc}(F) \text{ containing } K)$$

Proof sketch (Thanks to Botnan & Lesnick).

Let $F \cong \bigoplus_{L \in \mathbf{barc}(F)} I^L$.

$$\begin{aligned}\mathbf{rk}_F(K) &= \mathbf{rank}(F|_K) \\ &= \mathbf{rank}\left(\bigoplus_{L \in \mathbf{barc}(F)} I^L|_K\right) \\ &= \sum_{L \in \mathbf{barc}(F)} \mathbf{rank}\left(I^L|_K\right) \\ &= \sum_{L \in \mathbf{barc}(F)} \mathbf{1}_{\{L \supset K\}}. \square\end{aligned}$$

Möbius inversion formula

Let $\mathbf{Q}$ be a locally finite poset and let k be a field. Suppose that an element $0 \in \mathbf{Q}$ exists with the property that $0 \leq q$ for all $q \in \mathbf{Q}$. Consider a pair of functions $f, g : \mathbf{Q} \to k$ with the property that

$$g(q) = \sum_{r \leq q} f(r).$$

Then, $f(q) = \sum_{r \leq q} g(r) \cdot \mu_{\mathbf{Q}}(r, q)$ for $q \in \mathbf{Q}$.

Given $F : \mathbf{P} \to \mathbf{vec}$,

$$\mathbf{Q} = \mathbf{Con}(\mathbf{P}),\ g = \mathbf{rk}_F,\ f = \mathbf{dgm}_F.$$