A study on persistent homology with integer coefficients

Javier Perera-Lago

US

Sixth EACA International School on Computer Algebra and its Applications

- Persistent homology is a technique used to analyse the evolution of a certain algebraic property, called homology, on a topological space built step by step.
- It is useful in applied areas such as data analysis, computer vision and signal processing.
- But first, what is homology?

Homology

Chain complex $\mathcal{C}$:

$$\cdots \longrightarrow C_{p+1} \xrightarrow{d_{p+1}} C_p \xrightarrow{d_p} C_{p-1} \longrightarrow \cdots \qquad d_p \circ d_{p+1} \equiv 0$$

- C_p is a R-module $\forall p$
- $c \in C_p$ is called a p-chain
- $c \in Z_p$ (i.e., $d_p(c) = 0$) is called a p-cycle
- $c \in B_p$ (i.e., $\exists c' \in C_{p+1} | d_{p+1}(c') = c$) is called a p-boundary

Homology

Since $d_p \circ d_{p+1} \equiv 0$, we have

$$B_p \subset Z_p \subset C_p$$

and then we can define p-homology as the quotient R-module:

$$H_p = \frac{Z_p}{B_p}$$

Elements in H_p are called p-homology classes.

Homology

The choice of the coefficient ring R affects to the final results.

- When $R = \mathbb{Z}$, the modules $C_p(K), Z_p(K), B_p(K), H_p(K)$ are abelian groups.
- When R is a field such as $\mathbb{R}$ or $\mathbb{Z}_2$, the modules $C_p(K), Z_p(K), B_p(K), H_p(K)$ are vector spaces.

Roughly speaking, a 0-homology class represents a connected component on the simplicial complex, a 1-homology class represents a non-contractible loop, a 2-homology class represents a void, and so on.

Homology

We use the following classification theorem to compare homology groups:

Theorem (fundamental theorem of finitely generated modules)

Let be R a PID and let be M a R-module finitely generated. Then M is isomorphic to a direct sum:

$$M \cong R^{\beta} \oplus \bigoplus_{j=1}^{m} R/d_j R,$$

where β is a non-negative integer called Betti number and d_j are non null and non unit elements of R such $d_j | d_{j+1} \forall j$. This direct sum is unique up to product of d_j by an unit of R.

Persistent homology

Now imagine we don't have a single chain complex but a sequence of increasing chain complexes like this one:

$$\emptyset = \mathcal{C}_0 \hookrightarrow \mathcal{C}_1 \hookrightarrow \mathcal{C}_2 \hookrightarrow \cdots \hookrightarrow \mathcal{C}_m$$

Every chain complex has its own p-homology group, and we can consider the persistence Rmodule:

$$0 = H_0 \xrightarrow{\rho_{0,1}} H_1 \xrightarrow{\rho_{1,2}} H_2 \xrightarrow{\rho_{2,3}} \cdots \xrightarrow{\rho_{m-1,m}} H_m \tag{1}$$

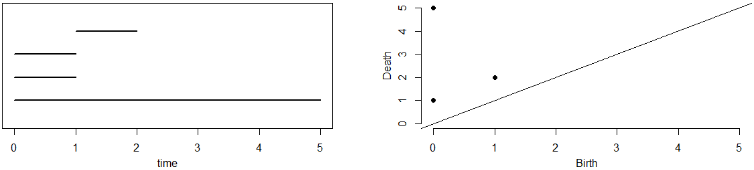
where H_i is the p-homology group in step i and $\rho_{i,j} = \rho_{j-1,j} \circ \cdots \circ \rho_{i,i+1}$ are homomorphisms induced by the inclusion map.

Persistent homology

When R is a field F, the polynomial ring $F[x]$ is a PID and we can use the previous classification theorem to relate the persistent homology module with the graded module:

$$\bigoplus_{i=1}^{n} \Sigma^{\alpha_i} F[x] \oplus \bigoplus_{j=1}^{m} \Sigma^{\gamma_j} F[x]/(x^{n_j}),$$

If for each $\Sigma^{\alpha_i} F[x]$ we consider the interval $[\alpha_i, \infty)$, and for each $\Sigma^{\gamma_j} F[x]/(x^{n_j})$ we consider $[\gamma_j, \gamma_j + n_j)$, we can define a barcode or equivalently a persistent diagram.



Persistent homology

- Barcodes and persistent diagrams are complete invariants for persistent homology.
- An interval $[i, j)$ in the barcode shows that there is a homology class that is born in step i and dies in step j.
- They are stable, that is, little perturbations on the filtered simplicial complex only induce little perturbations on barcodes and persistent diagrams.

When we use coefficients in $R = \mathbb{Z}$, $\mathbb{Z}[x]$ is not a PID so we cannot use this classification theorem, we don't have invariants and we don't know any stability results.

Persistent homology

If there is no classification theorem for persistent homology with coefficients in $\mathbb{Z}$, how can we work with it?

1. We introduce the *BD* groups, defined in *Defining and computing persistent Z-homology in the general case, 2014 (Romero, Ana and Heras, Jónathan and Rubio, Julio and Sergeraert, Francis)* for persistent homology with integer coefficients.
2. We introduce the *V* groups, defined in *Decomposition of pointwise finite-dimensional persistence modules, 2015 (Crawley-Boevey, William)* for persistent homology with field coefficients.
3. We look for connections between the two theories.

BD groups

From the persistence module:

$$0 = H_0 \xrightarrow{\rho_{0,1}} H_1 \xrightarrow{\rho_{1,2}} H_2 \xrightarrow{\rho_{2,3}} \cdots \xrightarrow{\rho_{m-1,m}} H_m$$

we define:

- $H_{i,j} = \operatorname{Im} \rho_{i,j} = \rho_{i,j}(H_i) \subset H_j$ for $i \leq j$
- $H_{i,k,j} = H_{i,k} \cap (\rho_{k,j})^{-1}(H_{i-1,j}) \subset H_k$ for $i \leq k \leq j$
- $BD_{i,j} = \frac{H_{i,i,j}}{H_{i,i,j-1}} = \frac{H_{i,i+1,j}}{H_{i,i+1,j-1}} = \cdots = \frac{H_{i,j-2,j}}{H_{i,j-2,j-1}} = \frac{H_{i,j-1,j}}{H_{i-1,j-1}}$

A non trivial $BD_{i,j}$ group is meant to show that there is a homology class that is born in step i and dies in step j

BD groups

In order to connect the $BD_{i,j}$ groups with V groups, defined in terms of intervals $I = [i, j)$, we propose the alternate notation

$$BD_{I,k} = \frac{H_{i,k,j}}{H_{i,k,j-1}}$$

And we give a new definition for BD groups with infinite intervals $I = [i, \infty)$, given by

$$BD_I = \frac{H_{i,m}}{H_{i-1,m}}$$

V groups

From the persistence module:

$$0 = H_0 \xrightarrow{\rho_{0,1}} H_1 \xrightarrow{\rho_{1,2}} H_2 \xrightarrow{\rho_{2,3}} \cdots \xrightarrow{\rho_{m-1,m}} H_m$$

we define:

- $V_{I,k}^{+} = \operatorname{Im} \rho_{i,k} \cap \ker \rho_{k,j}$
- $V_{I,k}^{-} = (\operatorname{Im} \rho_{i,k} \cap \ker \rho_{k,j-1}) + (\operatorname{Im} \rho_{i-1,k} \cap \ker \rho_{k,j})$
- $V_{I,k} = V_{I,k}^{+} / V_{I,k}^{-}$

A non trivial $V_{I,k}$ group is meant to show that there is a homology class being born at the beggining of I and dying at its end.

Relation between BD and V groups

Theorem

Given a persistent homology module with coefficients over a field F, the vector spaces $BD_{I,k}$ and $V_{I,k}$ are isomorphic for each interval I and for each $k \in I$.

Conjecture

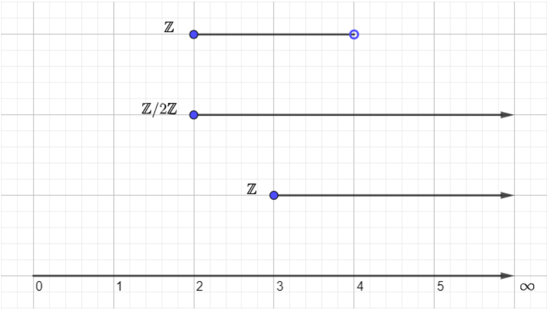
Given a persistent homology module with integer coefficients, the groups $BD_{I,k}$ and $V_{I,k}$ are isomorphic for each interval I and for each $k \in I$.

Proposition

Given a persistent homology module, independently on the choice of coefficients, $V^{+}_{I,k} \subset H_{i,k,j}$ and $V^{-}_{I,k} \subset H_{i,k,j-1}$.

Integer barcodes

Using BD groups, we can define an "integer" barcode that generalizes barcodes previously seen. For each non trivial $BD_{I,k}$ we draw one or many bars labelled with $\mathbb{Z}$ or $\mathbb{Z}/d\mathbb{Z}$.



We could define similar barcodes using V groups.

After that, we showed some stability results, but we needed to extend this theory. Until now, we were using a simplicial complex built on m steps. Now we will use a simplicial complex that grows continuously.
Consider a totally ordered set T such as $\mathbb{R}$. Then, the persistent homology module has a group H_i for each $i \in T$ and maps $\rho_{i,j}$ for each $i \leq j$ such that $\rho_i, j = \rho_{k,j} \circ \rho_{i,k}$.

BD and V groups with general index

Definition

Let be $T \subset \mathbb{R}$. *A cut in* T *is a pair of subsets* $c = (c^-, c^+)$ *such as* $T = c^- \cup c^+$ *y* $s < t$ *for any* $s \in c^-$, $t \in c^+$.

Definition

Let be the persistence $\mathbb{Z}$*-module* H_n *and let be ca cut in* T.

$t \in c^+$	$t \in c^-$
$\mathrm{Im}^-_{c,t} = \bigcup_{s \in c^-} \mathrm{Im}\, \rho_{s,t}$	$\ker^-_{c,t} = \bigcup_{r \in c^-, t \leq r} \ker \rho_{t,r}$
$\mathrm{Im}^+_{c,t} = \bigcap_{s \in c^+, s \leq t} \mathrm{Im}\, \rho_{s,t}$	$\ker^-_{c,t} = \bigcap_{r \in c^+} \ker \rho_{t,r}$

BD and V groups with general index

Definition

Let be the persistence $\mathbb{Z}$-module H_n and let be l, u two cuts in T and let be I the interval given by $I = l^+ \cap u^-$. Given $k \in I$, we define the submodules of H_t:

$$V_{I,t}^+ = \operatorname{Im}_{l,t}^+ \cap \ker_{u,t}^+$$
$$V_{I,t}^- = (\operatorname{Im}_{l,t}^- \cap \ker_{u,t}^+) + (\operatorname{Im}_{l,t}^+ \cap \ker_{u,t}^-)$$

Since $V_{I,t}^- \subset V_{I,t}^+$, we can define the quotient module:

$$V_{I,t} = V_{I,t}^+ / V_{I,t}^-$$

When using $T = \mathbb{N}$, we get the same definitions as before.

BD and V groups with general index

The generalization of the definition of $BD_n^{l,k}$ is more difficult:

Definition (original)

Let be the persistence $\mathbb{Z}$-module H_n and let be l, u two cuts in T and let be I the interval given by $I = l^+ \cap u^-$. Given a $k \in I$, we define $BD_{I,k}$ as:

$$BD_{I,k} = \frac{\mathrm{Im}_{l,k}^{+} \cap \bigcap_{t \in u^+} \rho_{k,t}^{-1}(\mathrm{Im}_{l,t}^{-})}{\mathrm{Im}_{l,k}^{+} \cap \bigcup_{k \leq t, t \in u^-} \rho_{k,t}^{-1}(\mathrm{Im}_{l,t}^{-})}$$

When using $T = \mathbb{N}$, we get the same definitions as before.

BD and V groups with general index

Theorem (original)

Let be H_n a persistence F-module with index T, let be l, u two cuts in T and let be I the interval $l^+ \cap u^-$. Given $k \in I$, we have a vector space isomorphism:

$$BD_{I,k} \cong V_{I,k}$$

Conjecture

Let be H_n a persistence $\mathbb{Z}$-module with index T, let be l, u two cuts in T and let be I the interval $l^+ \cap u^-$. Given $k \in I$, we have an abelian group isomorphism:

$$BD_{I,k} \cong V_{I,k}$$

Stability

Definition

Let be $c = (c^-, c^+)$ *a cut in* $\mathbb{R}$. *Given* $\varepsilon \in \mathbb{R}$, *we define* $c + \varepsilon$ *as the cut defined by the subsets:* $(c+\varepsilon)^- = \{a \in \mathbb{R} | a - \varepsilon \in c^-\}$, $(c+\varepsilon)^+ = \{a \in \mathbb{R} | a - \varepsilon \in c^+\}$.

Definition

Let be H_n *a persistent* $\mathbb{Z}$*-module and let be* $\varepsilon \in \mathbb{R}$. *We define* $1_\varepsilon(H_n)$ *as the persistence module given by the groups* $1_\varepsilon(H_s) = \rho_{s,s+\varepsilon}(H_s)$ *and structure maps* $1_\varepsilon(\rho_{s,t}) = \rho_{s+\varepsilon,t+\varepsilon}|_{\rho_{s,s+\varepsilon}(H_s)}$

Stability

Theorem (original)

Let be H_n a persistent $\mathbb{Z}$-module and let be $\varepsilon \in \mathbb{R}$. Let be $H'_n = 1_\varepsilon(H_n)$, l, u two cuts in $\mathbb{R}$ and $I = l^+ \cap u^-$. Given $t \in I$:

$$V'_{I,t} = V_{J,t+\varepsilon}$$

$$BD'_{I,t} = BD_{J,t+\varepsilon}$$

where $J = l^+ \cap (u+\varepsilon)^-$.

This is only a small achievement but it makes us believe that it is possible to develop a full stability theory for persistent homology with coefficients in $\mathbb{Z}$ similar to that of the book: *F. Chazal, V. De Silva, M. Glisse, and S. Oudot. The structure and stability of persistence modules. Springer, 2016.*

Summary

- It is possible to find connections between a theory made for field persistent homology (V groups) and a theory made for integer persistent homology (BD groups).
- It is possible to define (non complete) invariants for integer persistent homology.
- It is possible to extend BD groups for a continuous context.
- It is possible to find some partial stability results for integer persistent homology.

References

Perera-Lago, Javier
Un estudio sobre la homología persistente con coeficientes enteros, *2021*, Advisor: Prof. Rocio Gonzalez-Diaz
https://idus.us.es/handle/11441/130282.

Romero, Ana and Heras, Jónathan and Rubio, Julio and Sergeraert, Francis
Defining and computing persistent Z-homology in the general case, *2014*
https://arxiv.org/abs/1403.7086

Crawley-Boevey, William
Decomposition of pointwise finite-dimensional persistence modules, *2015*
https://www.worldscientific.com/doi/abs/10.1142/S0219498815500668?casa_token=ztFcpRY50T4AAAAA:rd96PvT-f6KrBofvPGAfHhK-VLs7FmoJ0nzD1YRTe_Jxpl0r46FPcoaN8Q2UvhVIh0KDPtgNlg