Computation of (persistent) homology

Computing homology : $H_p(K) = Z_p(K) / B_p(K)$

Z_p, B_p : $\mathbb{Z}_2$-vector spaces

H_p : quotient space

Computing basis for quotient space:

- Find basis for B_p (boundaries)
- Extend to basis for Z_p (cycles)
- new non-boundary basis cycles
  generate quotient Z_p / B_p $= H_p$

Notation:

- D : boundary matrix
- R_i : i^{th} col of matrix R
- pivot R_i : largest index w/ nonzero entry

Algorithm (matrix reduction, $\sim$ Gauss elim)

$R = D$, $V = I$

while $\exists i < j$ with pivot $R_i =$ pivot R_j
  add R_i to R_j, add V_i to V_j

[Cohen-Steiner / Edels. / Morozov 2007]

Result :

- $R = D \cdot V$ is __reduced__ ; unique pivots
  in nonzero columns
- V is full rank upper triangular

Properties :

- $\Sigma_B = \{ R_j \mid R_j \neq 0 \}$: basis for bdries
- $\Sigma_Z = \{ V_i \mid R_i = D \cdot V_i = 0 \}$: basis for cycles

But : $\Sigma_B \nsubseteq \Sigma_Z$ in general!

$\rightsquigarrow$ exchange basis elements in Σ_Z by Σ_B

Adapted basis for cycles

swap any $R_j \in \Sigma_B$
  replacing $V_i \in \Sigma_Z$
  where $i =$ pivot R_j $(=$ pivot $V_i)$

- $\widetilde{\Sigma}_Z = \Sigma_B \cup \underbrace{\{ V_i \in \Sigma_Z \mid i \notin \text{pivots } R \}}_{\Sigma_E}$
  is another basis for cycles
  (proof: $\sim$ Steinitz basis exchange lemma)

- Σ_E yields a basis for homology

Filtrations

Assumption for computation:
$$\emptyset = K_0 \subseteq K_1 \subseteq \dots \subseteq K_n = K$$
simplicial complexes, $\quad K_i \setminus K_{i-1} = \{\sigma_i\}$
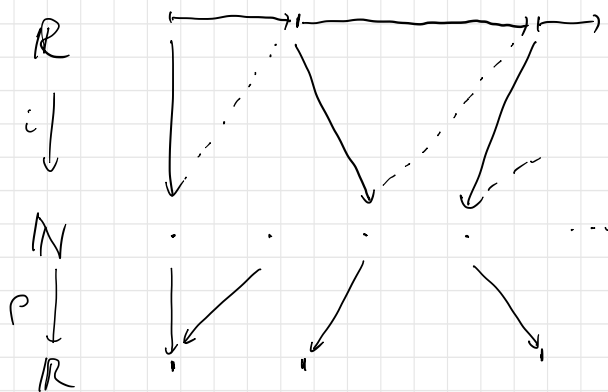
For alpha shapes, Vietoris-Rips:
- indexing by $\mathbb{R}$ (instead of $\mathbb{N}$)
- simultaneous simplices

Reindexing function $i: \mathbb{R} \to \mathbb{N}$
- weakly order preserving (poset map)
- right inverse of filtration value: $\rho \circ i = id_{\mathbb{R}}$

$\mathbb{R}$

$i \downarrow$

$\mathbb{N}$

$\rho \downarrow$

$\mathbb{R}$

Persistent homology

D: bdry. matrix wrt. filtration-ordered basis

- $\Sigma_{B,k} = \{ R_j \in \Sigma_B \mid j \leq k \}$: basis of $B_*(K_k)$

- $\widetilde{\Sigma}_{Z,k} = \Sigma_{B,k} \cup \{ R_j \in \Sigma_B \mid k \in [i,j), i = \text{pivot } R_j \}$
  $$\cup \{ V_i \in \Sigma_E \mid k \in [i, \infty) \}$$

$\underbrace{}_{\Sigma_{H,k}}$

Thm The persistent homology of the filtration $(K_i)_i$ is described up to isomorphism by the persistence barcode
$$\{ [i,j) \mid i = \text{pivot } R_j \} \cup$$
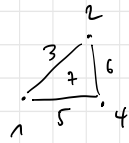$$\{ [i, \infty) \mid i \notin \text{pivots } R \}.$$
and generated by the cycles
$$\Sigma_B \cup \Sigma_E$$
$$\quad \uparrow \qquad \uparrow$$

eventually   essential
boundaries   cycles

Example

$$D = \begin{pmatrix} & 1 & 1 & & & & \\ & & 1 & & 1 & & \\ & & & & & 1 & \\ & & & 1 & 1 & & \\ & & & & & 1 & \\ & & & & & & 1 \\ & & & & & & \end{pmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix}$$

(columns labeled: 1 2 3 4 5 6 7)

$$R = \begin{pmatrix} & 1 & 1 & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ & & & & & 1 & \\ & & & & & & \end{pmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix} \quad V = \begin{pmatrix} 1 & & & & & & \\ & & 1 & & & & \\ & & & 1 & & 1 & \\ & & & & 1 & & \\ & & & & & 1 & 1 \\ & & & & & & 1 \\ & & & & & & 1 \end{pmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix}$$

(here: all dims in one matrix;
more efficient: dimensions separately)

Optimization: clearing [Chen/Kerber 2011]

- exchanged cycles

$$\{ V_i \in \Sigma_2 \mid i = \text{pivot } R_j \quad \text{for some } j\}$$

are not used $(R_i = D \cdot V_i = 0$, since V_i is cycle)

- in practice those are

  - many

  - expensive to compute

$\to$ avoid reducing those cols

Note: if $i = \text{pivot } R_j$ then simplices have

$$\dim \sigma_i = \dim \sigma_j - 1$$

$\Rightarrow$ after reducing bdry matrix for dim d,
we know the cols to skip in dim $d-1$

- setting $V_i = R_j$ for $i = \text{pivot } R_j$
  still yields V full rk. u. tri,
  $$R = D \cdot V \quad \text{reduced}$$

Cohomology

- co chains $C^p(K) = C_p(K)^* = \text{Hom}(C_p(K), \mathbb{F})$
  (dual vector space)

- coboundary map $\delta_p : C^{p-1} \to C^p$:
  adjoint of $\partial_p : C_p \to C_{p-1}$
  $$\delta^p(\eta)(c) = \eta(\partial_p c) \quad \text{for } c \in C_p$$
  $$\eta \in C^{p-1}$$

- $\delta^p \circ \delta^{p-1} = 0$

- cocycles: $Z^p = \ker \delta^{p+1}$
- coboundaries: $B^p = \text{im } \delta^p$
- cohomology: Z^p / B^p

Isomorphism of homology / cohomology

Prop. $\dim H_p(K) = \dim H^p(K)$.

Proof · $\dim H_p = \dim Z_p - \dim B_p$
- $\dim H^p = \dim Z^p - \dim B^p$
- $\underbrace{\dim Z_p}_{\dim \ker \partial_p} + \underbrace{\dim B^p}_{\substack{\dim \text{ im } \delta_p \\ = \dim \text{ im } \partial_p}} = \dim C_p$

$= \dim C^p = \dim Z^p + \dim B_p . \quad \square$

Computing homology vs cohomology

Observation [de Silva et al. 2011]:
(persistent) cohomology is faster than homology

Counting columns (n vertices, k-skeleton of full simplex)

- homology:
$$\underbrace{\sum_{p=1}^{k} \binom{n}{p+1}}_{\dim C_p} \text{columns} = \sum_{p=1}^{k} \left(\underbrace{\binom{n-1}{p}}_{\dim B_{p-1}} + \underbrace{\binom{n-1}{p+1}}_{\dim Z_p} \right)$$

- with clearing:
$$\sum_{p=1}^{k} \binom{n-1}{p} + \binom{n-1}{k+1} = \sum_{p=1}^{k+1} \binom{n-1}{p}$$

- cohomology:
$$\underbrace{\sum_{p=0}^{k-1} \binom{n}{p+1}}_{\dim C^p} \text{cols} = \sum_{p=0}^{k-1} \left(\underbrace{\binom{n-1}{p+1}}_{\dim B^{p+1}} + \underbrace{\binom{n-1}{p}}_{\dim Z^p} \right)$$

- with clearing:
$$\sum_{p=0}^{k-1} \binom{n-1}{p+1} + \binom{n-1}{0} = \sum_{p=0}^{k-1} \binom{n-1}{p}$$

- example ($k=3$, $n=192$)
  54.8 mio (homology)
  1.1 mio (cohomology)

Persistence modules

Filtration:

$$K_1 \hookrightarrow K_2 \hookrightarrow \ldots \hookrightarrow K_n = K$$

Diagram of topological spaces, indexed by $\mathbb{N}$
  (or $\mathbb{R}$; e.g. alpha shapes, Vietoris-Rips)
$(\mathbb{R} \to) \mathbb{N} \to$ Top   (diagram = functor)

Apply homology:

$$H_*(K_1) \to H_*(K_2) \to \ldots$$

Diagram of vector spaces
$(\mathbb{R} \to) \mathbb{N} \to$ vect   :   <u>persistence module</u>

<u>morphism</u> of persistence modules $f: V \to W$:
natural transformation : family of
linear maps such that the diagram

$$V_1 \longrightarrow V_2 \longrightarrow \ldots \longrightarrow$$
$$\downarrow f_1 \qquad \downarrow f_2$$
$$W_1 \longrightarrow W_2 \longrightarrow \ldots \longrightarrow$$

commutes.

Structure

A pers. mod. is <u>indecomposable</u> if
can't be written as nontriv. direct sum.

<u>Thm</u> (Krull - Schmidt - Remak (for $\mathbb{N} \to$ vect);
    Crawley Boevey 2015 (for $\mathbb{R} \to$ vect))
Any persistence module is a direct sum
of <u>interval modules</u>

$$0 \to 0 \to \underbrace{\mathbb{F} \to \ldots \to \mathbb{F}}_{} \to 0 \to \ldots$$

(in an essentially <u>unique</u> way).

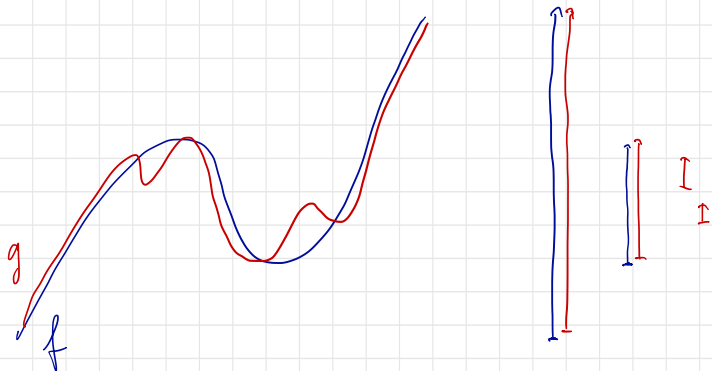$\Rightarrow$ the intervals defining (the support of)
the interval modules describe the isomorphism
type ($\to$ <u>persistence barcode</u>)

Note: persistence modules $\mathbb{N} \to$ vect
   $\Longleftrightarrow$ fin. gen. modules over poly. ring $\mathbb{F}[t]$
   (index $\Longleftrightarrow$ poly. degree)

Stability of persistence

Thm [Cohen-St. , Edels. , Haves 2005]

Let $\|f-g\|_\infty = \delta$. Then $\exists$ matching between intervals of barcodes $B(f)$, $B(g)$ s.t.

- matched intervals have δ-close endpoints
- unmatched intervals have length $\le 2\delta$

$\inf_{\substack{\delta \\ \text{matchings} \\ B(f) \leftrightarrow B(g)}} \rightsquigarrow$ Bottleneck distance

Stability step by step:

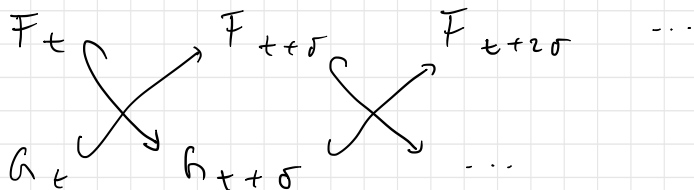function $\qquad\qquad X \rightarrow \mathbb{R}$

① $\downarrow$ sublevel set

topol. spaces $\quad$ (filtration) $\quad \mathbb{R} \rightarrow$ Top

② $\downarrow$ homology

vector spaces $\quad$ (persistence module) $\mathbb{R} \rightarrow$ vect

③ $\downarrow$ persistence diagram / barcode

collection of intervals $\qquad \mathbb{R} \rightarrow$ Mch

① ② ③ are
1-Lipschitz!
(which metric?
$\rightsquigarrow$ interleavings)

Interleavings

Let $\delta = \|f - g\|_\infty$. Write

$$F_t = f^{-1}(-\infty, t] \quad , \quad G_t = \ldots$$

Then

$F, G : \mathbb{R} \to \text{Top}$ are $\underline{\delta\text{-interleaved}}$.
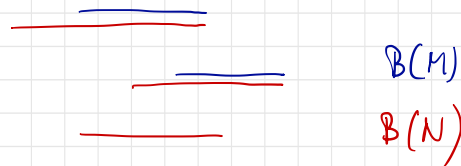
Define interleaving distance:

$$d_I(F, G) = \inf_\delta : \exists \ \delta\text{-interleaving } F \bowtie G$$

Apply homology (functor) $\rightsquigarrow$ preserves commut. $H_*(F)$, $H_*(G) : \mathbb{R} \to \text{Vect}$ are δ-intel.

Structure of persistence submodules/ quotients

__Thm__. Let M, N persistence modules with $f : M \to N$ an injection (surj.) (i.e., $f_t : M_t \to N_t$ injective $\forall t$).

Then

$B(M)$

$B(N)$

$\forall$ interval in $B(M)$
$\exists !$ interval in $B(M)$
with large or same left (right)
  and same right end (left)

__Proof__ (sketch)

· Bar I of $B(M)$ is mapped to lin. comb. of bars of $B(N)$

· Naturality: I overlaps all those bars to the right (at both ends)

· Injectivity: one of them ends tog. w. I

Induced matchings [B, Lesnick 2015]

For a general morphism $f: L \to N$
consider epi-mono fact:

$$L \twoheadrightarrow \operatorname{im} f \hookrightarrow N$$

Note:
- injection $M \hookrightarrow N$ induces
injection of intervals $B(M) \hookrightarrow B(N)$
- surjection $L \twoheadrightarrow M$ induces
injection $B(M) \hookrightarrow B(L)$

- compose to a matching $B(L) \nrightarrow B(N)$

$$
\begin{array}{l}
\underline{\quad\quad\quad} \quad B(L) \\
\underline{\quad\quad\quad} \quad B(M) \\
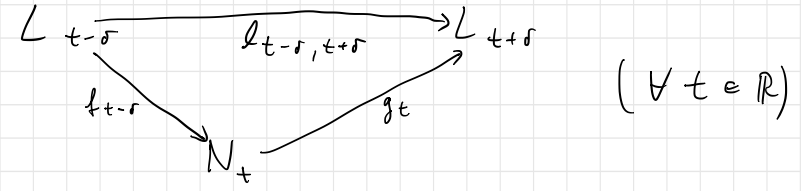\underline{\quad\quad\quad} \quad B(N)
\end{array}
$$

- Bar in $B(L)$ overlaps matched
bar in $B(N)$ to the right

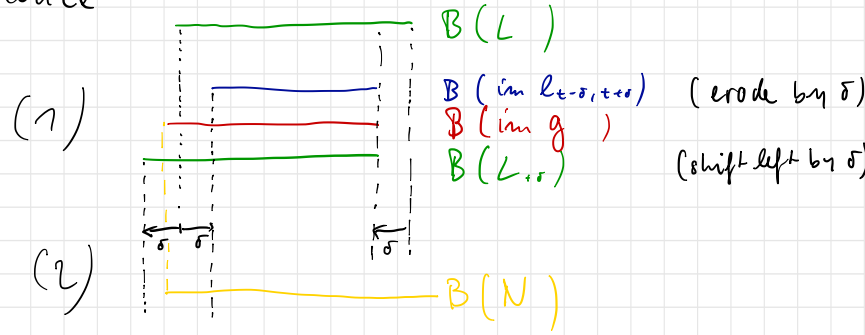Apply this to interleavings ...

Stability from interleavings

Consider interleaving: $f_t: L_t \to N_{t+\delta}$, $g_t: N_t \to L_{t+\delta}$

$$
\begin{array}{ccc}
L_{t-\delta} & \xrightarrow{\quad \ell_{t-\delta, t+\delta} \quad} & L_{t+\delta} \\
\searrow_{f_{t-\delta}} & & \nearrow_{g_t} \\
& N_t &
\end{array}
\qquad (\forall t \in \mathbb{R})
$$

Then
$$\operatorname{im} \ell_{t-\delta, t+\delta} \hookrightarrow \operatorname{im} g_t \hookrightarrow L_{t+\delta} \qquad (1)$$

Moreover
$$N_t \twoheadrightarrow \operatorname{im} g_t \qquad (2)$$

Hence

$B(L)$
$B(\operatorname{im} \ell_{t-\delta, t+\delta})$ (erode by δ)
$B(\operatorname{im} g)$
$B(L_{+\delta})$ (shift left by δ)
$B(N)$

(1)
(2)

- Matched bars of L, N have δ-close
left endpoints! (Dually: right ends)

- If a bar of L has length $> 2\delta$
then it must be matched
$\Rightarrow$ unmatched bars have length $\leq 2\delta$