Topological Descriptors for Geometric Data	TUM 07/2016

① Existence of barcodes:

Def: Filtration over $T \subseteq \mathbb{R}$:

Nested family $(F_t)_{t \in T}$ of topological spaces:

$$F_s \subseteq F_t \quad \forall s \leq t \in T.$$

Note: recall def. of simplicial complex: vertex set $V \to K \subseteq 2^V$ closed under subsets.

Def: Persistence module over $T \subseteq \mathbb{R}$ (fixed field k):

- Family $(V_t)_{t \in T}$ of k-vector spaces
- Family $(v_s^t)_{s \leq t \in T}$ of k-linear maps

s.t.:
- $v_t^t = \mathrm{id}_{V_t} \quad \forall t \in T$
- $$\begin{tikzcd} V_s \arrow[rr, "v_s^u"] \arrow[dr, "v_s^t"'] & & V_u \\ & V_t \arrow[ur, "v_t^u"'] & \end{tikzcd}$$ commutes $\forall s \leq t \leq u \in T$.
  (i.e. $v_s^u = v_t^u \circ v_s^t$).

Q Why do we say "module"?

Example: $T = \mathbb{N}$

$$\begin{tikzcd} V_0 \arrow[r, "v_0^1"] \arrow[rr, bend right, "v_0^2\ (\text{redundant})"'] & V_1 \arrow[r, "v_1^2"] & V_2 \arrow[r] & \cdots \end{tikzcd}$$

$\mathbb{V} := \bigoplus_{i \in \mathbb{N}} V_i \qquad \mathbb{V} \ni x = (x_0, x_1, x_2, \ldots)$ (finitely many nonzero coordinates).

k-vector space $\xrightarrow{\text{lift}}$ module over $k[t]$:

$$t \cdot (x_0, x_1, x_2, \ldots) = (0, v_0^1(x_0), v_1^2(x_1), \ldots)$$

⚠ **Note:** More generally, every persistence module can be viewed as a left module over a certain algebra generated by the paths in the poset $(T, \leq)$.

we can decompose it as a direct sum ...

Note: A persistence module can also be viewed as a <u>functor</u> from the poset $(T, \leq)$ to the category Vect_k.

Thm: (decomposition of finitely generated persistence modules):

hyp: $\dim \bigoplus_{t \in T} V_t \left(= \sum_{t \in T} \dim V_t \right) < +\infty$.

Then, we can view $\mathbb{V}$ as a persistence module over some finite set $T' \subset \mathbb{N}$ ($\mathbb{V}$ has only <u>finitely many</u> nonzero constituent spaces).

$$0 \longrightarrow 0 \longrightarrow \underset{\in T'}{V_{t_1}} \longrightarrow 0 \longrightarrow \underset{\in T'}{V_{t_2}} \longrightarrow \underset{\in T'}{V_{t_3}} \longrightarrow 0 \longrightarrow \cdots$$

Hence, $\mathbb{V}$ can be viewed as a <u>left module</u> over $\underset{(PID)}{k[t]}$ (cf. previous example), finitely generated

→ decomposition theorem:

$$\mathbb{V} \simeq \underbrace{\bigoplus_{j \in J} t^{b_j} k[t]}_{\substack{\text{free part} \\ (=0 \text{ here because} \\ \text{no vector lasts indefinitely})}} \oplus \underbrace{\bigoplus_{l \in L} t^{b_l} k[t] / t^{r_l} k[t]}_{\text{torsion part}} .$$

torsion part (lifespan: $\underbrace{[b_l, b_l + r_l)}_{\text{interval}}$).

Note: This special case is interesting because it is the one that occurs in computations (filtrations are made of <u>finitely many</u> (finite) <u>simplicial complexes</u>).

Decomposition Theorem: (reproduced with the 2 relevant items)

A persistence module $\mathbb{V}$ over $T \subseteq \mathbb{R}$ is <u>interval-decomposable</u> in the following cases (at least):

- when T is finite,
- when $\mathbb{V}$ is ~~finite~~ pointwise finite-dimensional (<u>pfd</u>).

Moreover, the decomposition is essentially unique.

② Computation of barcodes from filtrations:

Method:

Input: simplicial filtration $\mathcal{K}$:

- $T = \{0, 1, \dots, m\}$
- $K_0 = \emptyset$, $K_m = K$ (simplicial complex)
- $\forall i$, K_i is a subcomplex of K_{i+1}.

⊕ technical condition: $\forall i$, $K_i = K_{i-1} \cup \{\sigma_i\}$ (single simplex inserted at a time).

Algo: ① boundary matrix M (binary):

$$M_{ij} = \begin{cases} 1 & \text{if } \sigma_i \in \partial \sigma_j \ \& \ \dim \sigma_i = \dim \sigma_j - 1 \\ 0 & \text{otherwise} \end{cases}$$

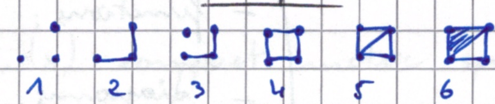
② $\mathrm{low}(j) = \begin{cases} \max\{i \mid M_{ij} \neq 0\} \\ 0 \text{ if } M_{ij} = 0 \ \forall i \end{cases}$

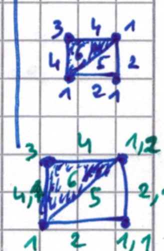
③ reduce to column-echelon form (by scanning columns from left to right):

For $j = 1$ to m do:
  While $\exists\, \ell < j$ s.t. $\mathrm{low}(\ell) = \mathrm{low}(j) \neq 0$:
    $c_j \leftarrow c_j + c_\ell$.

④ Interpretation: – every <u>pivot</u> pairs 2 simplices (σ_i, σ_j) and gives $[i, j)$.
– every <u>unpaired simplex</u> σ_i gives $[i, +\infty)$.

Example:





→ ~~enforce~~ technical condition:

→ t increasing

$$\begin{array}{c|cccccccccc} & 1 & 1,1 & 1,2 & 2 & 2,1 & 3 & 4 & 4,1 & 5 & 6 \\ \hline 1 & & & & 1 & & & & 1 & 1 & \\ 1,1 & & & & \boxed{1} & 1 & & & & & \\ 1,2 & & & & & \boxed{1} & & 1 & & \boxed{1} & \\ 2 & & & & & & & & & & \\ 2,1 & & & & & & & & & & \\ 3 & & & & & & & \boxed{1} & \boxed{1} & & \\ 4 & & & & & & & & & & 1 \\ 4,1 & & & & & & & & & & 1 \\ 5 & & & & & & & & & & \boxed{1} \\ 6 & & & & & & & & & & \end{array}$$

columns to reduce.

– – – reduction – – –

$$\begin{array}{c|cccccccccc} & 1 & 1,1 & 1,2 & 2 & 2,1 & 3 & 4 & 4,1 & 5 & 6 \\ \hline \boxed{1} & & & & 1 & & & & & & \\ 1,1 & & & & \boxed{1} & 1 & & & & & \\ 1,2 & & & & & \boxed{1} & & 1 & & & \\ 2 & & & & & & & & & & \\ 2,1 & & & & & & & & & & \\ 3 & & & & & & & \boxed{1} & & & \\ 4 & & & & & & & & & & 1 \\ \boxed{4,1} & & & & & & & & & & 1 \\ 5 & & & & & & & & & & \boxed{1} \\ 6 & & & & & & & & & & \end{array}$$

$\Rightarrow H(\mathcal{K}) \simeq \boxed{\mathbb{I}_{[1,2)}} \oplus \boxed{\mathbb{I}_{[1,2)}} \oplus \boxed{\mathbb{I}_{[3,4)}} \oplus \boxed{\mathbb{I}_{[5,6)}} \oplus \boxed{\mathbb{I}_{[1,\infty)}} \oplus \boxed{\mathbb{I}_{[4,\infty)}}$

$\boxed{\text{Complexity:}}$ $\boxed{O(m^3)}$ (~~pass~~ near-linear in practice).

- divide-and-conquer (fast matrix multiplication) : $O(m^\omega)$.
- random projections : ?

---

③ stability: [functional setting, easier to grasp]

Operator D_g: $\underset{\substack{\text{(function}\\ X\to\mathbb{R})}}{f} \mapsto \underset{\text{(diagram)}}{D_g(f)}$ or $\underset{\text{(barcode)}}{\mathcal{B}(f)}$

$\boxed{Q}$ show that this operator is Lipschitz-continuous.

↳ metrics:

- functions: $\|f-g\|_\infty$
- diagrams: d_b^∞ (bottleneck distance) (modified W_∞)

$\boxed{\text{Def:}}$ partial matching $M: A \leftrightarrow B$:

$M \subseteq A\times B$ s.t.:

$$\begin{cases} \forall a\in A, \ \exists \text{ at most } 1\ b\in B \text{ s.t. } (a,b)\in M \\ \forall b\in B, \ \exists \text{ at most } 1\ a\in A \text{ s.t. } (a,b)\in M. \end{cases}$$

$\boxed{\text{Def:}}$ Cost of a matching $M: A\leftrightarrow B$:

$$c(M) := \max\left\{ \sup_{(a,b)\in M} \|a-b\|_\infty \ ; \sup_{\substack{s\in A\sqcup B\\ s \text{ unmatched}}} \|s-\bar{s}\|_\infty \right\}.$$

orthogonal projection onto the diagonal $\Delta : y=x$.

$\boxed{\text{Def:}}$ $d_b^\infty(A,B) := \inf_{M:A\leftrightarrow B} c(M)$.

Exercise: show that d_b^∞ is an extended pseudo-distance.

$\boxed{\text{Thm:}}$ (Stability) suppose $f,g: X\to\mathbb{R}$ are pfd.

Then, $d_b^\infty(D_g f, D_g g) \leq \|f-g\|_\infty$.

Notes: ⓐ no further hyp. on X, f, g
ⓑ f, g may be def. on $\neq$ spaces
ⓒ functions not even neces- sary ...