INF556 Topological Data Analysis
Final Exam — 3 hours

December 17, 2021

Note: The text of the exam is written in English. Your answers can be written indifferently in French or in English. All printed documents are allowed. Computers and pocket calculators are forbidden.

1 Euler characteristic

Given a topological space X and a field $\mathbf{k}$, the *Euler characteristic* is the quantity:

$$\chi(X;\mathbf{k}) = \sum_{i=0}^{+\infty} (-1)^i \dim \mathrm{H}_i(X;\mathbf{k}).$$

Question 1. Show that χ is a topological invariant, that is: for any spaces X, Y that are homotopy equivalent, $\chi(X;\mathbf{k}) = \chi(Y;\mathbf{k})$.
Hint: look at what happens to each individual homology group.

Now we want to prove the Euler-Poincaré theorem:

Theorem 1. *For any simplicial complex X and any field $\mathbf{k}$:*

$$\chi(X;\mathbf{k}) = \sum_{i=0}^{+\infty} (-1)^i n_i(X),$$

where $n_i(X)$ denotes the number of simplices of X of dimension i.

For this we will use topological persistence. Consider an arbitrary filtration of X:

$$\emptyset = X_0 \subseteq X_1 \subseteq \cdots \subseteq X_m = X.$$

Assume without loss of generality that a single simplex σ_j is inserted at each step j:

$$\forall j = 1, \cdots, m,\ X_j \setminus X_{j-1} = \{\sigma_j\}.$$

Note that m is then equal to the number of simplices of X, that is:

$$m = \sum_{i=0}^{+\infty} n_i(X).$$

Let us apply the persistence algorithm to this simplicial filtration. Recall from lecture 5 that we have the following property:

Lemma 2. *At each step j, the insertion of simplex σ_j either creates an independent d_j-dimensional cycle (i.e. increases the dimension of $\mathrm{H}_{d_j}(X_{j-1};\mathbf{k})$ by 1) or kills a (d_j-1)-dimensional cycle (i.e. decreases the dimension of $\mathrm{H}_{d_j-1}(X_{j-1};\mathbf{k})$ by 1), where d_j is the dimension of σ_j.*

Question 2. Using Lemma 2, prove Theorem 1.
Hint: proceed by induction on j.

Question 3. Deduce that the Euler characteristic of a triangulable space is independent of the choice of field $\mathbf{k}$.

2 The Dunce Hat

Recall that the Dunce Hat is obtained by indentifying the three edges of a triangle as shown in Figure 1.

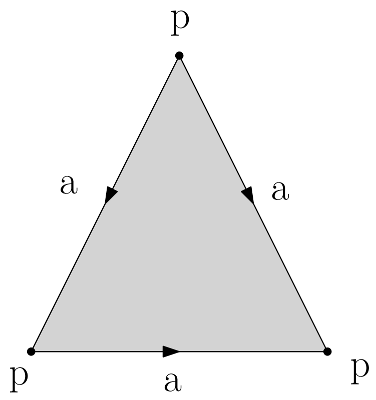

Figure 1: The Dunce Hat.

Question 4. Build a triangulation of the Dunce Hat (you may draw a picture to represent it). Beware that your triangulation must be a simplicial complex, not a general cell complex.

Question 5. Use your simplicial complex to compute the homology of the Dunce Hat.
Hint: to avoid tedious calculations, you can proceed as in exercise 1: pick a filtration of your complex then apply the persistence algorithm; for each simplex σ_j inserted, use Lemma 2 to predict its effect on the homology (identify the created d_j-cycle or the killed (d_j-1)-cycle).

3 Eccentricity-based signatures

Let (X, d_X) be a finite metric space. Define the eccentricity as follows:

$$\forall x \in X,\ \mathrm{ecc}(x) = \frac{1}{2}\max\{\mathrm{d}_X(x,x') \mid x' \in X\}.$$

This function takes its values in $\mathbb{R}^+$. For any $t \in \mathbb{R}^+$, let X_t denote the t-sublevel set of ecc, that is:

$$X_t = \mathrm{ecc}^{-1}([0,t]) = \{x \in X \mid \mathrm{ecc}(x) \leq t\}.$$

Consider the filtration $\mathcal{E}(X, \mathrm{d}_X)$ defined by:

$$\forall t \in \mathbb{R}^+,\ E_t = R_t(X_t, \mathrm{d}_X).$$

where $R_t(X_t, \mathrm{d}_X)$ denotes the Rips complex of X_t of parameter t. Our goal here is to show that this filtration defines a stable signature, that is:

Theorem 3. *For any finite metric spaces (X, d_X) and (Y, d_Y), we have*

$$\mathrm{d}_\mathrm{b}^\infty(\mathrm{Dgm}\ \mathcal{E}(X, \mathrm{d}_X),\ \mathrm{Dgm}\ \mathcal{E}(Y, \mathrm{d}_Y)) \leq 2\, \mathrm{d}_{\mathrm{GH}}(X, Y).$$

We will use the following embedding result, which we saw in PC 8:

Lemma 4. *Any finite metric space (Z, d_Z) embeds isometrically into $(\mathbb{R}^n, \ell^\infty)$, where n denotes the cardinality of Z.*

Let (X, d_X) and (Y, d_Y) be two finite metric spaces, and let $\varepsilon > \mathrm{d}_{\mathrm{GH}}(X, Y)$. As we saw in PC 8, (X, d_X) and (Y, d_Y) can be jointly embedded isometrically into $(\mathbb{R}^d, \ell^\infty)$, for some $d > 0$, such that the Hausdorff distance between their images is at most ε. The construction of the joint embedding is illustrated in Figure 2 and relies on Lemma 4 (for the second step in the construction).

We call respectively $X' = \gamma \circ \gamma_X(X)$ and $Y' = \gamma \circ \gamma_Y(Y)$ the images of X and Y through the embedding.

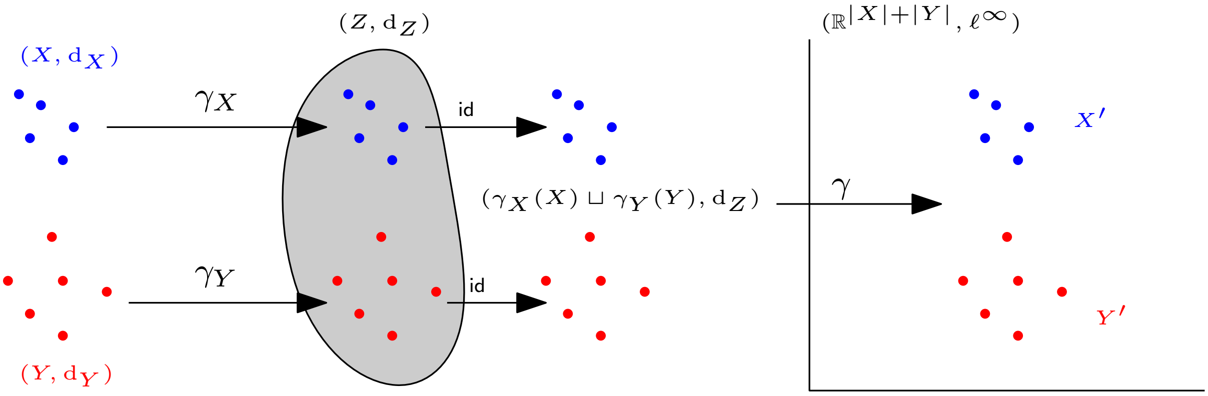

Figure 2: Outline of the embedding for the proof of Theorem 3.

Question 6. Show that $\mathcal{E}(X', \ell^\infty)$ is isomorphic to $\mathcal{E}(X, \mathrm{d}_X)$ as a simplicial filtration.
Hint: this means that there is a bijection $X \to X'$ that induces a bijection between the simplices of the two filtrations, such that the times of appearance of the simplices are preserved.

Similarly, $\mathcal{E}(Y', \ell^\infty)$ is isomorphic to $\mathcal{E}(Y, \mathrm{d}_Y)$. Thus, we have:

$$\mathrm{d}_\mathrm{b}^\infty(\mathrm{Dgm}\ \mathcal{E}(X, \mathrm{d}_X), \mathrm{Dgm}\ \mathcal{E}(Y, \mathrm{d}_Y)) = \mathrm{d}_\mathrm{b}^\infty(\mathrm{Dgm}\ \mathcal{E}(X', \ell^\infty), \mathrm{Dgm}\ \mathcal{E}(Y', \ell^\infty)).$$

For any finite set $S \subset \mathbb{R}^d$ and any $t \geq 0$, let S^t denote the t-offset of S in the ℓ^∞-norm, that is:

$$S^t = \{x \in \mathbb{R}^d \mid \min_{s \in S} \|x - s\|_\infty \leq t\}.$$

Question 7. Show that $X'^t_t \subseteq Y'^{t+\varepsilon}_{t+\varepsilon}$ and $Y'^t_t \subseteq X'^{t+\varepsilon}_{t+\varepsilon}$ for any $t \geq 0$.

Question 8. Define a function $f_{X'} : \mathbb{R}^d \to \mathbb{R}$ whose t-sublevel set is X'^t_t for every $t \in \mathbb{R}^+$. Similarly, define a function $f_{Y'} : \mathbb{R}^d \to \mathbb{R}$ whose t-sublevel set is Y'^t_t for every $t \in \mathbb{R}^+$.

Question 9. Deduce that $\|f_{X'} - f_{Y'}\|_\infty \leq \varepsilon$.

Question 10. Deduce now that $d_b^\infty(\mathrm{Dgm}\, f_{X'}, \mathrm{Dgm}\, f_{Y'}) \le \varepsilon$, where $\mathrm{Dgm}\, h$ denotes the persistence diagram of the filtration of the sublevel sets of h.

Question 11. Deduce now that $d_b^\infty(\mathrm{Dgm}\ \mathcal{EC}(X', \ell^\infty), \mathrm{Dgm}\ \mathcal{EC}(Y', \ell^\infty)) \le \varepsilon$, where the filtration $\mathcal{EC}(Z', \ell^\infty)$ has the space $C_t(Z'_t, \ell^\infty)$ for every $t \in \mathbb{R}^+$ — here C_t stands for the Čech complex of parameter t.

Question 12. Deduce finally that $d_b^\infty(\mathrm{Dgm}\, \mathcal{E}(X', \ell^\infty), \mathrm{Dgm}\, \mathcal{E}(Y', \ell^\infty)) \le 2\varepsilon$.

Question 13. Conclude.

4 Matchings induced by morphisms of persistence modules

Let $\phi\colon M \to N$ be a morphism between two persistence modules indexed over the real line $\mathbb{R}$. For simplicity we will assume that M and N decompose into finite direct sums of interval modules:

$$M \simeq \bigoplus_{i=1}^{n} \mathbf{k}_{[a_i,b_i)} \quad N \simeq \bigoplus_{j=1}^{m} \mathbf{k}_{[c_j,d_j)},$$

where $\mathbf{k}_{[a,b)}$ denotes the interval module supported on the interval $[a,b)$. As you can see, we are also assuming that all the intervals are right-open. Finally, to simplify things further, we are assuming that all the a_i's and b_i's are different from one another, so that the cardinality of the set $\bigcup_{i=1}^n \{a_i, b_i\}$ is exactly $2n$. This is what happens e.g. when M is computed from a finite simplicial filtration using the persistence algorithm. We make the same assumption on the c_j's and d_j's, but note that we may have $a_i = c_j$ or $b_i = d_j$ for some pairs (i,j).

For now, let us pick an arbitrary pair (i,j) and consider a morphism $\phi_{ij}\colon \mathbf{k}_{[a_i,b_i)} \to \mathbf{k}_{[c_j,d_j)}$.

Question 14. Assume that $\phi_{ij} \neq 0$. Show then that $c_j \le a_i < d_j \le b_i$.

Question 15. Assume that ϕ_{ij} is surjective. Show then that $c_j = a_i < d_j \le b_i$.

Dually, we have $c_j \le a_i < d_j = b_i$ whenever ϕ_{ij} is injective.

Let us call $\mathrm{Dgm}\, M$ and $\mathrm{Dgm}\, N$ respectively the persistence barcodes of M and N. We have $\mathrm{Dgm}\, M = \{[a_i, b_i) \mid 1 \le i \le n\}$ and $\mathrm{Dgm}\, N = \{[c_j, d_j) \mid 1 \le j \le m\}$, which are sets since all the bars' endpoints are different in each barcode.

Question 16. Assume that $\phi\colon M \to N$ is surjective. Show then that there exists an injection $\phi^*\colon \mathrm{Dgm}\, N \hookrightarrow \mathrm{Dgm}\, M$ such that, for every $1 \le j \le m$, we have $\phi^*([c_j, d_j)) = [c_j, b)$ for some $b \ge d_j$.

Dually, when $\phi\colon M \to N$ is injective, there exists an injection $\phi^*\colon \mathrm{Dgm}\, M \twoheadrightarrow \mathrm{Dgm}\, N$ such that, for every $1 \le i \le n$, we have $\phi^*([a_i, b_i)) = [c, b_i)$ for some $c \le a_i$.

We now decompose an arbitrary morphism $\phi\colon M \to N$ as follows:

$$\begin{tikzcd} M \arrow[rr, "\phi"] \arrow[dr, "\pi"'] & & N \\ & \mathrm{Im}\ \phi \arrow[ur, hook, "\iota"'] & \end{tikzcd}$$

where $\pi = \phi$ with codomain $\mathrm{Im}\,\phi$, and where ι is the inclusion of $\mathrm{Im}\,\phi$ as a submodule of N.

Let us take for granted the fact that $\mathrm{Im}\,\phi$ itself decomposes as a finite direct sum of interval modules supported on intervals of the form $[e, f)$ with pairwise different endpoints (the proof of this fact is omitted here). The preceding questions show then that π and ι induce injections $\pi^*\colon \mathrm{Dgm}\ \mathrm{Im}\,\phi \hookrightarrow \mathrm{Dgm}\ M$ and $\iota^*\colon \mathrm{Dgm}\ \mathrm{Im}\,\phi \to \mathrm{Dgm}\ N$. Consider the partial matching Γ_ϕ between $\mathrm{Dgm}\ M$ and $\mathrm{Dgm}\ N$ defined by:

$$([a_i, b_i),\ [c_j, d_j)) \in \Gamma_\phi \Longleftrightarrow \exists\, [e, f) \in \mathrm{Dgm}\ \mathrm{Im}\,\phi\ s.t.\ \pi^*([e, f)) = [a_i, b_i) \text{ and } \iota^*([e, f)) = [c_j, d_j).$$

This matching is called the *canonical matching induced by ϕ*.

Question 17. Show that we have $c_j \leq a_i < d_j \leq b_i$ for every matched pair $([a_i, b_i),\ [c_j, d_j)) \in \Gamma_\phi$.

This result is the first step of the so-called *Induced Matching Theorem*, originally proven by Bauer and Lesnick, which is at the heart of an alternate proof of the algebraic stability theorem for persistence modules. The details can be found in Bauer, Lesnick (2015): Induced matchings and the algebraic stability of persistence barcodes, *Journal of Computational Geometry*, 6(2):162–191.