INF556 Topological Data Analysis
Final Exam — 3 hours

December 16, 2022

Important:

- The exercises and problem are independent of one another.
- The text of the exam is written in English. Your answers can be written indifferently in French or in English.
- Please keep in mind that the quality of your answers (completeness of the arguments and clarity of their exposition) will be key for the grading.
- All printed documents are allowed. By contrast, computers, cellphones, tablets, pocket calculators, etc., are forbidden.

1 Some calculations...

Let the field of coefficients be $\mathbb{Z}/2\mathbb{Z}$.

Question 1. Compute the barcode associated with the simplicial filtration depicted in Figure 1.

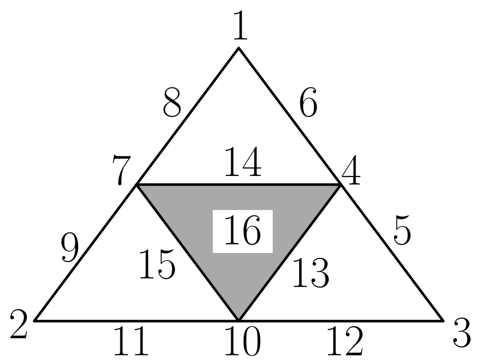

Figure 1: A simplicial filtration.

Question 2. Deduce the homology of the topological space underlying the simplicial complex of Figure 1.

2 Scissors and glue...

Question 3. The projective line $\mathbb{RP}^1$ is obtained from the unit circle $\mathbb{S}^1$ by identifying antipodal points: $x \sim -x$ for every $x \in \mathbb{S}^1$. Show that $\mathbb{RP}^1$ is homeomorphic to the circle $\mathbb{S}^1$ itself. A proof by pictures showing the sequence of gluing and twisting operations will be enough.

Question 4. Show that the quotient space depicted in Figure 2 is homeomorphic to the torus. Again, a proof by pictures showing the sequence of cuttings and gluings will be enough.

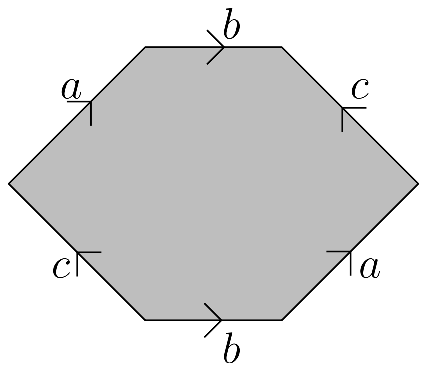

Figure 2: The opposite sides of the hexagon are glued together with the given orientations.

3 Bouquets of circles and spheres

Given topological spaces $X_1, \cdots, X_n$ and fixed basepoints $x_1 \in X_1, \cdots, x_n \in X_n$, define their *bouquet* as the quotient of the disjoint union of the X_i's by the identification of the x_i's, that is:

$$X_1 \vee \cdots \vee X_n := \left(\bigsqcup_{i=1}^{n} X_i \right) / \sim$$

where $\bigsqcup$ denotes the disjoint union and where $\sim$ denotes the equivalence relation induced by the identifications $x_i \sim x_j$, $1 \leq i \leq j \leq n$. See Figure 3 for an illustration with $n = 8$.

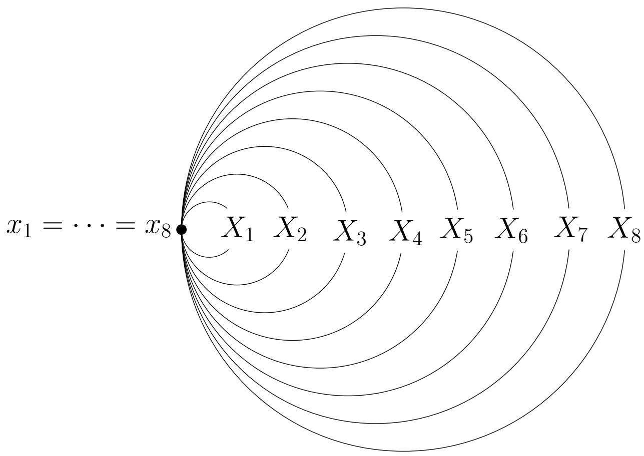

Figure 3: A bouquet of 8 circles.

Question 5. Compute the homology (in $\mathbb{Z}/2\mathbb{Z}$) of a bouquet of n circles with basepoints, for arbitrary n. You can proceed for instance by induction, and do not forget that for calculations we use simplicial homology on simplicial complexes.

Question 6. Assume the bouquet of circles is embedded in the plane as in Figure 3. Compute then the homology of its complement in the plane.

Question 7. Same questions for a bouquet of n 2-spheres with basepoints, embedded in $\mathbb{R}^3$.

4 Problem: the space of persistence diagrams

Let $(\mathcal{X}, d_{\mathcal{X}})$ be a metric space. For $x, y \in \mathcal{X}$, we call a *path* from x to y a continuous map $\gamma : [0,1] \to X$ such that $\gamma(0) = x, \gamma(1) = y$. We note $\Gamma(x, y)$ the set of all such paths. For all the remainder of this problem, we will suppose that $\mathcal{X}$ is always *path-connected*, that is $\Gamma(x, y)$ is never empty for any $x, y \in \mathcal{X}$.

A *subdivision* of $[0,1]$ is given by an integer $n \geqslant 2$ and a n-tuple $t_1 \leqslant .. \leqslant t_n$ such that $t_1 = 0, t_n = 1$. We write $\mathcal{S}$ the set of all subdivisions. We define the *length* of a path γ to be:

$$L(\gamma) := \sup_{(t_1..t_n) \in \mathcal{S}} \sum_{i=1}^{n-1} d_{\mathcal{X}}(\gamma(t_i), \gamma(t_{i+1})) \tag{1}$$

We now define the *geodesic distance* on $\mathcal{X}$ to be:

$$d_g(x, y) := \inf_{\gamma \in \Gamma(x,y)} L(\gamma) \tag{2}$$

We finally say that a path $\gamma \in \Gamma(x, y)$ is a *geodesic* between x and y if it achieves the infimum in (2), and that $(\mathcal{X}, d_{\mathcal{X}})$ is a *geodesic space* if any two points x, y are connected by such a geodesic (i.e. the infimum in (2) is always achieved).

4.1 General facts

Question 8. Let $(\mathcal{X}, d_{\mathcal{X}})$ be a metric space. Show that $\forall x, y \in \mathcal{X}, d_g(x, y) \geqslant d_{\mathcal{X}}(x, y)$

Question 9. Show that $(\mathbb{R}^d, ||.||_2)$ is a geodesic space.

Question 10. Let $(\mathcal{X}, d_{\mathcal{X}})$ be a geodesic space. Show that the geodesic distance introduced in (2) is indeed a distance over $\mathcal{X}$, i.e it satisfies the following axioms:

- $\forall x, y \in \mathcal{X}, d_g(x, y) = 0 \Leftrightarrow x = y$
- $\forall x, y \in \mathcal{X}, d_g(x, y) = d_g(y, x)$
- $\forall x, y, z \in \mathcal{X}, d_g(x, z) \leqslant d_g(x, y) + d_g(y, z)$

4.2 Geodesics for persistence diagrams

In the following, we consider that a persistence diagram is a **finite** multiset of points[1] in $\mathbb{R}^2_> := \{(x_1, x_2) \in \mathbb{R}^2 | x_2 > x_1\}$ (called the *off-diagonal points*), together with a unique virtual point (representing the diagonal) $\{\Delta\}$ with infinite (countable) multiplicity. We say that ϕ is a *matching* between X and Y, noted $\phi \in \Pi(X, Y)$, if ϕ is a bijection between the points in X (including all copies of $\{\Delta\}$) and the points in Y (including all copies of $\{\Delta\}$).

[1]A multiset is a set of points with multiplicities. Given a point x, its copies can be labeled e.g. $x^{(1)}..x^{(n)}$ and treated as different points despite their being located at the same place.

Remark: This is equivalent to the definition given during the lectures but will be more convenient to use here.

We endow the space of persistence diagrams with the d_2 metric: for two diagrams X, Y we have,

$$d_2(X,Y) = \left(\inf_{\phi \in \Pi(X,Y)} \sum_{x \in X} ||x - \phi(x)||_2^2 \right)^{\frac{1}{2}} \tag{3}$$

with the convention that $||x - \Delta||_2 = ||x - \pi_\Delta(x)||_2$, where $\pi_\Delta(x)$ denotes the orthogonal projection of x onto the diagonal, and $||\Delta - \Delta||_2 = 0$.

A matching ϕ between X and Y which minimizes (3) is said to be *optimal.*

Goal: The goal of this subsection is to show that the set of persistence diagrams endowed with the d_2 metric is a geodesic space.

Question 11. Show that $\Pi(X, Y)$ is never empty. Show also that the infimum in (3) is always achieved. Is it unique?

Question 12. Let X, Y be two diagrams and ϕ be an optimal matching between them. We introduce $\gamma : t \mapsto \{(1-t)\, x + t\, \phi(x) \mid x \in X\}$. Show that γ is a geodesic between X and Y. Conclude.

4.3 Curvature of the space of persistence diagrams

An important step towards understanding the structure of a geodesic space is to study its curvature. The goal of this subsection is twofold: first, to prove that there is no upper bound on the curvature of the space of persistence diagrams; second, to prove that the curvature is everywhere lower-bounded by zero.

For the first objective, we use the following characterization of spaces with curvature bounded from above. If a geodesic space $(\mathcal{X}, d_\mathcal{X})$ has curvature upper-bounded by some $\kappa > 0$, then for any $x, y \in \mathcal{X}$ such that $d_\mathcal{X}(x, y) < \frac{1}{\kappa}$ there is a *unique* geodesic between x and y.

Question 13. Using this characterization, exhibit a counterexample (or a family thereof) showing that there is no upper bound on the curvature of the space of persistence diagrams.

For the second objective, we use the following characterization: given a **geodesic** space $(\mathcal{X}, d_\mathcal{X})$, we say that $\mathcal{X}$ *is non-negatively curved* if for all $X, Y \in X$, for all geodesic $\gamma : [0,1] \to \mathcal{X}$ between X and Y, and any $Z \in \mathcal{X}$, we have:

$$\forall t \in [0,1], \quad d_\mathcal{X}(Z, \gamma(t))^2 \geqslant t\, d_\mathcal{X}(Z,Y)^2 + (1-t)\, d_\mathcal{X}(Z,X)^2 - t\,(1-t)\, d_\mathcal{X}(X,Y)^2 \tag{4}$$

When (4) appears to be an equality (for all X, Y, Z, γ, t), we say that $\mathcal{X}$ has *curvature zero.*

Question 14. Prove that $(\mathbb{R}^d, ||.||_2)$ has curvature zero.

We now consider three arbitrary diagrams, say X, Y, Z. Let ϕ_X^Y be an optimal matching between X and Y, and let γ be the corresponding geodesic as per Question 12. For $t \in [0,1]$, we introduce $\psi_t : X \to \mathbb{R}^2_> \cup \{\Delta\}$ defined by $\psi_t(x) = (1-t)\, x + t\, \phi_X^Y(x)$, and we also consider $\phi_Z^t : Z \to \gamma(t)$ an optimal matching between Z and $\gamma(t)$. We finally introduce $\phi_Z^X = (\psi_t)^{-1} \circ \phi_Z^t$ and $\phi_Z^Y = \phi_X^Y \circ \phi_Z^X$.

Question 15. Prove that, for any $x \in X$, if there is some $t \in (0,1)$ such that $\psi_t(x) = \Delta$, then $x = \Delta$ and $\phi_X^Y(x) = \Delta$.

Question 16. Prove the following results:

$$d_2(Z,\gamma(t))^2 = \sum_{z\in Z} ||z - [(1-t)\phi_Z^X(z) + t\phi_Z^Y(z)]||^2 \tag{5}$$

$$d_2(Z,Y)^2 \leqslant \sum_{z\in Z} ||z - \phi_Z^Y(z)||^2 \tag{6}$$

$$d_2(Z,X)^2 \leqslant \sum_{z\in Z} ||z - \phi_Z^X(z)||^2 \tag{7}$$

$$d_2(X,Y)^2 = \sum_{z\in Z} ||\phi_Z^X(z) - \phi_Z^Y(z)||^2 \tag{8}$$

Question 17. Combine Eqs (4-8) to conclude that the space of persistence diagrams is non-negatively curved.