Topological Data Analysis - Exercises

Lucas Moschen, EMAp/FGV February 10, 2021

Contents

1 General topology

1.1 Important definitions

Definition 1.1.1. *A topological space is a pair $(X, \mathcal{T})$ where X is a set and $\mathcal{T}$ is a collection of subsets of X such that:*

1. $\emptyset \in \mathcal{T}, X \in \mathcal{T}$.
2. *for every infinite collection* $\{O_\alpha\}_{\alpha \in A} \subset \mathcal{T}$, *we have* $\bigcup_{\alpha \in A} O_\alpha \in \mathcal{T}$.
3. *for every finite collection* $\{O_i\}_{1 \leq i \leq n} \subset \mathcal{T}$, *we have* $\bigcap_{1 \leq i \leq n} O_i \in \mathcal{T}$.

Definition 1.1.2. *Let $x \in \mathbb{R}^n$ and $r > 0$. The open ball of center x and radius r, denoted $\mathcal{B}(x, r)$, is defined as: $\mathcal{B}(x, r) = \{y \in \mathbb{R}^n, ||x - y|| < r\}$.*

Definition 1.1.3. *Let $A \subset \mathbb{R}$ and $x \in A$. We say that A is open around x if there exists $r > 0$ such that $\mathcal{B}(x, r) \subset A$. We say that A is open if for every $x \in A$, A is open around x.*

Definition 1.1.4. *Let $(X, \mathcal{T})$ be a topological space, and $Y \subset X$. We define the subspace topology on Y as the following set:*

$$T_{|Y} = \{O \cap Y, O \in \mathcal{T}\}$$

Definition 1.1.5. *Let $f : X \to Y$ be a map. We say that f is continuous if for every $O \in U$, the preimage $f^{-1}(O) = \{x \in X, f(x) \in O\}$ is in $\mathcal{T}$.*

1.2 Exercises

Exercise 1. *Let $X = \{0, 1, 2\}$ be a set with three elements. What are the different topologies that X admits?*

Proof. As we know every Topology contains $\emptyset$ and $\{0, 1, 2\}$, so we can disconsider when writing the topologies, that is, all below contain these subsets.

- (2) Basic: $\{\emptyset, \{0, 1, 2\}\}$ - $\mathcal{P}(\{0, 1, 2\})$.
- (8) With $\{0\}$: $\{\{0\}\} - \{\{0\}, \{0, 1\}\} - \{\{0\}, \{1, 2\}\} - \{\{0\}, \{0, 2\}\} - \{\{0\}, \{0, 2\}, \{0, 1\}\}$ $\{\{0\}, \{2\}, \{0, 2\}\} - \{\{0\}, \{2\}, \{0, 2\}, \{1, 2\}\} - \{\{0\}, \{2\}, \{0, 2\}, \{0, 1\}\}$
- (8) With $\{1\}$: $\{\{1\}\} - \{\{1\}, \{0, 1\}\} - \{\{1\}, \{1, 2\}\} - \{\{1\}, \{0, 2\}\} - \{\{1\}, \{1, 2\}, \{0, 1\}\}$ $\{\{0\}, \{1\}, \{0, 1\}\} - \{\{0\}, \{1\}, \{0, 1\}, \{1, 2\}\} - \{\{0\}, \{1\}, \{0, 1\}, \{0, 2\}\}$
- (8) With $\{2\}$: $\{\{2\}\} - \{\{2\}, \{0, 1\}\} - \{\{2\}, \{1, 2\}\} - \{\{2\}, \{0, 2\}\} - \{\{2\}, \{0, 2\}, \{1, 2\}\}$ $\{\{1\}, \{2\}, \{1, 2\}\} - \{\{1\}, \{2\}, \{1, 2\}, \{0, 1\}\} - \{\{1\}, \{2\}, \{1, 2\}, \{0, 2\}\}$
- (3) No singleton: $\{\{0, 1\}\} - \{\{1, 2\}\} - \{\{0, 2\}\}$

□

Exercise 2. *Let $\mathbb{Z}$ be the set of integers. Consider the cofinite topology $\mathcal{T}$ on $\mathbb{Z}$, defined as follows: a subset $O \subset \mathbb{Z}$ is an open set if and only if $O = \emptyset$ or ${}^cO$ is finite. Here, ${}^cO = \{x \in \mathbb{Z}, x \notin O\}$ represents the complementary of O in $\mathbb{Z}$*

Proof. 1. Show that $\mathcal{T}$ is a topology on $\mathbb{Z}$.

Let's verify the three axioms:

(a) $\emptyset$ is an open set by definition and $\mathbb{Z}$ is open set because ${}^c\mathbb{Z} = \emptyset$ is finite.

(b) Let $\{O_\alpha\}_{\alpha \in A} \subset \mathcal{T}$. So ${}^cO = {}^c\left(\bigcup_{\alpha \in A} O_\alpha\right) = \bigcap_{\alpha \in A} {}^cO_\alpha \implies {}^cO \subset {}^cO_\alpha, \forall \alpha \in A$. If $\forall \alpha, O_\alpha = \emptyset$, then ${}^cO = {}^c\emptyset \implies O = \emptyset$ and O is open. On the other hand, if there exists $\alpha \in A$ such that $O_\alpha \neq \emptyset$ we have ${}^cO_\alpha$ being finite, so is ${}^cO$, given the inclusion. We conclude O is open set.

(c) Let $\{O_i\}_{1 \le i \le n} \subset \mathcal{T}$. So ${}^cO = {}^c\left(\bigcap_{1 \le i \le n} O_i\right) = \bigcup_{1 \le i \le n} {}^cO_i$. If $O_i = \emptyset$ for some $1 \le i \le n$, $O = \emptyset$ because of the intersection. Alternatively, if $\forall i, O_i \neq \emptyset$ we have that ${}^cO_i$ is finite and a finite union of finites is finite. We conclude that O is open set.

By (a), (b) and (c), $\mathcal{T}$ is a topology on $\mathbb{Z}$.

2. Exhibit an sequence of open sets $\{O_n\}_{n \in \mathbb{N}} \subset \mathcal{T}$ such that $\bigcap_{n \in \mathbb{N}} O_n$ is not an open set.

Let $O_n = {}^c\{1, ..., n\}$. Thus ${}^cO_n = \{1, ..., n\}$ is finite and

$$ {}^c\left(\bigcap_{n \in \mathbb{N}} O_n\right) = \bigcup_{n \in \mathbb{N}} {}^cO_n = \bigcup_{n \in \mathbb{N}} \{1, ..., n\} = \mathbb{N}, $$

that is not finite. Therefore, this intersection is not an open set.

□

Exercise 3. *Let $x \in \mathbb{R}^n$, and $r > 0$. Let $y \in \mathcal{B}(x, r)$. Show that*

$$ \mathcal{B}(y, r - ||x - y||) \subset \mathcal{B}(x, r) $$

Proof. Let $z \in \mathcal{B}(y, r - ||x - y||)$, so $||z - y|| < r - ||x - y|| \implies ||z - y|| + ||x - y|| < r$. We can conclude that, by the triangular inequality,

$$ ||x - z|| \le ||x - y|| + ||z - y|| < r. $$

In that sense, $z \in \mathcal{B}(x, r)$ and $\mathcal{B}(y, r - ||x - y||) \subset \mathcal{B}(x, r)$.

Remark. In the notes, the exercise is to prove $\mathcal{B}(y, ||x - y||) \subset \mathcal{B}(x, r)$, however, this does not hold, because if we take y next the border of $\mathcal{B}(x, r)$, $||x - y|| \approx r$ and $B(y, r - \epsilon) \not\subset B(x, r)$.

□

Exercise 4. *Let $x, y \in \mathbb{R}^n$, and $r = ||x - y||$. Show that*

$$ \mathcal{B}\left(\frac{x + y}{2}, \frac{r}{2}\right) \subset \mathcal{B}(x, r) \cap \mathcal{B}(y, r) $$

Proof. Denote $m = \dfrac{x+y}{2}$. Take $z \in \mathcal{B}\left(m, \dfrac{r}{2}\right)$. Thus, using the triangular inequality,

$$||x - z|| \leq ||x - m|| + ||m - z|| = \frac{1}{2}||x - y|| + ||m - z|| < r/2 + r/2 = r$$

$$||y - z|| \leq ||y - m|| + ||m - z|| = \frac{1}{2}||y - x|| + ||m - z|| < r/2 + r/2 = r$$

So $z \in \mathcal{B}(x, r)$, $z \in \mathcal{B}(y, r)$ and $z \in \mathcal{B}(x, r) \cap \mathcal{B}(y, r)$. Therefore $\mathcal{B}(m, \frac{r}{2}) \subset \mathcal{B}(x, r) \cap \mathcal{B}(y, r)$. □

EXERCISE 5. *Show that the open balls $\mathcal{B}(x, r)$ of R^n are open sets (with respect to the Euclidean topology).*

Proof. We have to prove that for every $y \in \mathcal{B}(x, r)$, there exists $\epsilon > 0$ such that $\mathcal{B}(y, \epsilon) \subset \mathcal{B}(x, r)$. Put $\epsilon = r - ||x - y||$. As we have proved in exercise 3, $\mathcal{B}(y, \epsilon) \subset \mathcal{B}(x, r)$. So $\mathcal{B}(x, r)$ is open set. □

EXERCISE 6. *Consider $X = \mathbb{R}$ endowed with the Euclidean topology. Are the following sets open? Are they closed?*

Proof.

1. $[0, 1]$. It's not open set because for every $\epsilon > 0, \mathcal{B}(0, \epsilon) = (-\epsilon, \epsilon) \not\subset [0, 1]$. It's closed because $[0, 1]^c = (-\infty, 0) \cup (1, \infty)$ is an union of two open sets, as we prove in item 3.
2. $[0, 1)$. It's not open for the same reason as before. It's not closed because $B(1, \epsilon) = (1 - \epsilon, 1 + \epsilon) \not\subset (-\infty, 0) \cup [1, \infty]$.
3. $(-\infty, 1)$. It's open because: take $x < 1$. Put $r = 1 - x$ and take $z \in \mathcal{B}(x, r)$. If $z > x$, $|x - z| < 1 - x \implies z < 1$. If $z < x$, it follows $z < 1$. It proves $z < 1$ and $(-\infty, 1)$ is open. It's not closed cause $\forall \epsilon > 0, \mathcal{B}(1, \epsilon) \not\subset [1, \infty)$.
4. the singletons. It's not open cause $\forall \epsilon > 0, x + \epsilon/2 \in \mathcal{B}(x, \epsilon)$. It's close cause $(-\infty, x) \cup (x, \infty)$ is union of open sets.
5. $\mathbb{Q}$. It's not open because for every open ball around a rational, there are irrationals, that is, let $x \in \mathbb{Q}$ and take $\epsilon > 0$, then there exists $y \in (\mathbb{R} - \mathbb{Q}) \cap \mathcal{B}(x, \epsilon)$. It's not closed for the same reason, for every irrational, there is rationals for every open ball.

   Remark. We shall prove the rationals are dense in the reals. Let $x \in \mathbb{Q}$ and $\epsilon > 0$. If ϵ is irrational, take $x - \epsilon/2 \subset (x - \epsilon, x + \epsilon)$. Suppose $x - \epsilon/2$ is rational, then $\frac{2x - \epsilon}{2} = \frac{m}{n}$ for some integers m and n, that is, $2x - \epsilon = 2m/n$ and $\epsilon = 2(x - \frac{m}{n}) \in \mathbb{Q}$, contradiction. So there is an irrational in $\mathcal{B}(x, \epsilon)$. If ϵ is rational, consider

   $$y = \frac{1}{\sqrt{2}}(x - \epsilon) + (1 - \frac{1}{\sqrt{2}})(x + \epsilon) = (x + \epsilon) - \epsilon\sqrt{2}$$

   That is a convex combination, so $y \in \mathcal{B}(x, \epsilon)$. Moreover, y is irrational, with a similar proof by contradiction. This proves the statement.

On the other hand, we must prove for every two irrationals (a, b), there is a rational between them. Denote $c = b - a > 0$. Let $n \in \mathbb{N}$ such that $n > \frac{1}{c} \implies cn > 1 \implies (bn - an) > 1$. So exists $m \in (an, bn) \implies m/n \in (a, b)$. This proves the second statement.

$\square$

Exercise 7. *A map is continuous if and only if the preimage of closed sets are closed sets.*

Proof. First we shall prove that $f^{-1}({}^{c}A) = {}^{c}(f^{-1}(A))$. Let's prove the double inclusion. Take $x \in f^{-1}({}^{c}A)$. So there exists $y \in {}^{c}A$ such that $f(x) = y$. Suppose that $x \in f^{-1}(A)$. It implies the existence of $z \in A$ such that $y = f(x) = z$, absurd. So $x \in {}^{c}(f^{-1}(A))$.

Now take $x \in {}^{c}(f^{-1}(A))$. Therefore, $\forall y \in A, f(x) \neq y$. In that case, $f(x) \in {}^{c}A \implies x \in f^{-1}({}^{c}A)$. Then we have showed the equality.

Now let's prove the equivalence. Suppose f is a continuos map and take a closed set F. We shall prove that $f^{-1}(F)$ is closed. Well, ${}^{c}(f^{-1}(F)) = f^{-1}({}^{c}F)$ is open, because ${}^{c}F$ is open, by the continuity. We conclude that $f^{-1}(F)$ is closed.

Suppose that for every closed set F, we have $f^{-1}(F)$ being closed. We will use that A is open if ${}^{c}A$ is closed. This is true because ${}^{c}({}^{c}(A)) = A$. Take an open set A. ${}^{c}(f^{-1}(A)) = f^{-1}({}^{c}A)$ is closed, because ${}^{c}A$ is. Thus $f^{-1}(A)$ is open and we have proved the continuity of f.

$\square$

2 Homeomorphisms

2.1 Important definitions

Definition 2.1.1. *Let* $(X, \mathcal{T})$ *and* $(Y, \mathcal{U})$ *be two topological spaces, and* $f : X \to Y$ *a map. We say that* f *is a homeomorphism if*

1. f *is a bijection,*
2. $f : X \to Y$ *is continuos,*
3. $f^{-1} : Y \to X$ *is continuos.*

If there exists such a homeomorphism, we say that the two topological spaces are homeomorphic.

Definition 2.1.2. *Let* $(X, \mathcal{T})$ *be a topological space. We say that* X *is connected if for every open sets* $O, O' \in \mathcal{T}$ *such that* $O \cap O' = \emptyset$ *(i.e., they are disjoint), we have*

$$X = O \cup O' \implies O = \emptyset \text{ or } O' = \emptyset.$$

Definition 2.1.3. *Let* $(X, \mathcal{T})$ *be a topological space. Suppose that there exists a collection of* n ***non-empty**, **disjoint** and **connected open sets*** $(O_1, ..., O_n)$ *such that*

$$\bigcup_{1 \leq i \leq n} O_i = X.$$

Then we say that X *admits* n *connected components.*

Definition 2.1.4. *Let* $(X, \mathcal{T})$ *be a topological space, and* $n \geq 0$. *We say that it has dimension* n *if the following is true: for every* $x \in X$, *there exists an open set* O *such that* $x \in O$, *and a homeomorphism* $O \to \mathbb{R}^n$.

2.2 Exercises

Exercise 8. *Show that the topological spaces* $\mathbb{R}^n$ *and* $\mathcal{B}(0, 1) \subset \mathbb{R}^n$ *are homeomorphic.*

Proof. Let $f : \mathcal{B}(0, 1) \to \mathbb{R}^n$ be defined as $f(x) = \dfrac{x}{1 - ||x||}$. I observe it's well defined because $||x|| < 1$. We shall prove f is a homeomorphism.

1. **Injective:** Take $x, y \in \mathcal{B}(0, 1)$ and suppose that

   $$\frac{x}{1 - ||x||} = \frac{y}{1 - ||y||}.$$

   Applying the norm in both sides, we obtain the equation

   $$||x||(1 - ||y|||) = ||y||(1 - ||x||) \implies ||x|| = ||y||.$$

   On the other side x and y points to the same direction, given that

   $$y = \frac{1 - ||y||}{1 - ||x||} x = \alpha x,$$

   with $\alpha = 1$ because of the same norm. We conclude $x = y$.

2. **Surjective:** Take $y \in \mathbb{R}^n$. We shall prove that there exists $x \in \mathcal{B}(0,1)$ such that $f(x) = y$, that is,

$$\frac{x}{1 - ||x||} = y$$

Applying the norm we observe that if that is true, $||x|| = ||y|| - ||y||||x|| \implies ||x|| = \frac{||y||}{1+||y||}$. And $x = (1 - ||x||)y = \frac{1}{1+||y||}y$. We conclude that for every $y \in \mathbb{R}^n$, if we take $x = \frac{y}{1+||y||}$,

$$f(x) = \frac{y/(1 + ||y||)}{1 - ||y||/(1 + ||y||)} = y$$

3. **Continuity of f:** Consider an open set $A \subset \mathbb{R}^n$. Let $B = f^{-1}(A)$. We shall prove B is open, that is, for every $x \in B$, exists $r > 0$ such that $\mathcal{B}(x, r) \subset B$. Take $x = f^{-1}(y) \in B$. Because A is open, there is $\epsilon > 0$ such that $\mathcal{B}(y, \epsilon) \subset A$. Take δ such that

$$\frac{\delta}{1 - ||x|| - \delta}(1 + ||y||) < \epsilon$$

and $z = f^{-1}(w) \in \mathcal{B}(x, \delta)$.

$$\begin{aligned}
||y - w|| &= \left|\left|\frac{x}{1 - ||x||} - \frac{z}{1 - ||z||}\right|\right| = \frac{1}{1 - ||x||}\left|\left|x - \frac{1 - ||x||}{1 - ||z||}z\right|\right| \\
&= \frac{1}{1 - ||x||}\left|\left|x - z + z - \frac{1 - ||x||}{1 - ||z||}z\right|\right| \\
&\leq \frac{||x - z||}{1 - ||x||} + \frac{1}{1 - ||x||}\left(1 - \frac{1 - ||x||}{1 - ||z||}||z||\right) \\
&= \frac{||x - z||}{1 - ||x||} + \frac{||z||}{1 - ||x||}\frac{||x|| - ||z||}{1 - ||z||} \\
&\leq \frac{1}{1 - ||x||}||x - z||(1 + ||w||) \\
&\leq \frac{1}{1 - ||x||}||x - z||(1 + ||y - w|| + ||y||) \\
\implies ||y - w|| &\leq \frac{||x - z||}{1 - ||x|| - ||x - z||}(1 + ||y||) \\
&< \frac{\delta}{1 - ||x|| - \delta}(1 + ||y||) < \epsilon
\end{aligned}$$

So $w \in \mathcal{B}(y, \epsilon) \subset A \implies z \in B$, what proves B is open. It concludes the continuity of f.

4. **Continuity of f^{-1}:** The inverse is given by

$$f^{-1}(y) = \frac{y}{1 + ||y||}$$

The demonstration is quite similar to the previous item, given that the only difference is the signal.

By items (1) - (4), we conclude f is a homeomorphism and $\mathcal{B}(0,1) \simeq \mathbb{R}^n$.

□

Exercise 9. *Show that $\mathcal{B}(x, r)$ and $\mathcal{B}(y, s)$ are homeomorphic.*

Proof. Consider the function $f : \mathcal{B}(0,1) \to \mathcal{B}(c,r)$ given by $f(x) = r \cdot x + c$. Let's prove f is a homeomorphism.

1. **Injective:** If $x, y \in \mathcal{B}(0,1)$ and $rx + c = ry + c \implies x = y$, because $r > 0$ by definition. So f is injective.

2. **Surjective:** Let $y \in \mathcal{B}(c,r)$ and $x = (y-c)/r$. So $||x|| = ||y-c||/r < 1$, by definition. So $x \in \mathcal{B}(0,1)$ and $f(x) = y$ what proves this function is surjective.

3. **Continuity of f:** Let $A \subset \mathcal{B}(c,r)$ open set and denote $B = f^{-1}(A)$. Take $x = f^{-1}(y) \in B$. We know there exists $\epsilon > 0$ such that $\mathcal{B}(y,\epsilon) \subset A$. Define $\delta = \epsilon/r$ and take $z = f^{-1}(w) \in \mathcal{B}(x,\delta)$.
$$||y - w|| = ||rx + c - (rz + c)|| = r||x - z|| < r\delta = \epsilon$$
Therefore $w \in \mathcal{B}(y,\epsilon) \subset A \implies z \in B$. So $\mathcal{B}(x,\delta) \subset B$, what proves B is open. This concludes the continuity of f.

4. **Continuity of f^{-1}:** The inverse is given by
$$f^{-1}(y) = \frac{y-c}{r}$$
This function is continuos for the same argument as before.

By items (1) - (4), we conclude f is a homeomorphism and $\mathcal{B}(0,1) \simeq \mathcal{B}(c,r)$. Since this is an equivalence relation, we have that

$$\mathcal{B}(0,1) \simeq \mathcal{B}(x,r) \text{ and } \mathcal{B}(0,1) \simeq \mathcal{B}(y,s) \text{ implies } \mathcal{B}(x,r) \simeq \mathcal{B}(y,s).$$

□

EXERCISE 10. *Show that* $\mathbb{S}(0,1)$, *the unit circle of* $\mathbb{R}^2$, *is homeomorphic to the ellipse*

$$\mathcal{S}(a,b) = \left\{ (x,y) \in \mathbb{R}^2, \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1 \right\},$$

for any $a, b > 0$.

Proof. Consider the function $f : \mathbb{S}(0,1) \to \mathcal{S}(a,b)$ defined as $f(x,y) = (ax, by)$. Let's prove it is a homeomorphism.

1. **Injective:** Let $(x_1,y_1),(x_2,y_2) \in \mathbb{S}(0,1)$ such that $(ax_1, by_1) = (ax_2, by_2)$. Since $a, b > 0$, we have $x_1 = x_2$ and $y_1 = y_2$. It proves f is injective.

2. **Surjective:** Let $(z,w) \in \mathcal{S}(a,b)$ and $(x,y) = \left(\frac{z}{a}, \frac{w}{b}\right)$. It's clear that $f(x,y) = (z,w)$ and $x^2 + y^2 = \frac{z^2}{a^2} + \frac{w^2}{b^2} = 1$, so $(x,y) \in \mathbb{S}(0,1)$. It proves f is surjective.

3. **Continuity of f:** Let $A \subset \mathcal{S}(a,b)$ open set and denote $B = f^{-1}(A)$. Take $(x,y) = f^{-1}(z,w) \in B$. We know there exists $\epsilon > 0$ such that $\mathcal{B}((z,w),\epsilon) \subset A$. Put δ as defined

below and take $(x',y') = f^{-1}((z',w')) \in \mathcal{B}((x,y),\delta)$. Consider the norm 1

$$\begin{aligned}||(z',w')-(z,w)||_1 &= ||(ax',by')-(ax,by)||_1 = ||\left(a(x'-x),b(y'-y)\right)||_1 \\ &= a|x'-x|+b|y'-y|, \text{ define } c=\max\{a,b\} \\ &\le c(|x'-x|+|y'-y|) = c||(x'-x,y'-y)||_1\end{aligned}$$

By the equivalente of the norms, there exists constants k_1,k_2 such that

$$||(z',w')-(z,w)|| \le k_1||(z',w')-(z,w)||_1 \le ck_1||(x'-x,y'-y)||_1 \le ck_1k_2||(x'-x,y'-y)||$$

Then we need $\delta = \frac{\epsilon}{ck_1k_2}$ in order to prove that $(z',w') \in \mathcal{B}((z,w),\epsilon) \subset A \implies (x',y') \in B$. So $\mathcal{B}((x,y),\delta) \subset B$, what proves B is open. This concludes the continuity of f.

4. **Continuity of f^{-1}:** The inverse is given by

$$f^{-1}((z,w)) = (z/a, w/b)$$

This function is continuos for the same argument as before.

By items (1) - (4), we conclude f is a homeomorphism and $\mathbb{S}(0,1) \simeq \mathcal{S}(a,b)$.

□

Exercise 11. *Show that* $[0,1)$ *and* $(0,1)$ *are not homeomorphic.*

Proof. We shall prove by contradiction. Suppose these exists a homeomorphism $f : [0,1) \to (0,1)$. Let $0 < z = f(0) < 1$ and define the following function

$$\begin{aligned} g : (0,1) &\to (0,z) \cup (z,1) \\ x &\mapsto g(x) = f(x)\end{aligned}$$

This function is well defined given that z is not image of other point but 0. The function is injective because if $g(y) = g(x) \implies f(y) = f(x) \implies x = y$, given that f is injective. This function is also surjective since f is and $0 < w < 1$ and $w \neq z$, it's clear that $f(0) \neq w$. As g is an induced map of a continuos function, by Proposition 1.21 from the notes, it's continuos and so is its inverse. We conclude g is a homeomorphism.

Now I will prove that $(0,1)$ admits only 1 connected component, that is, it's connected. Suppose it's not and there exists $O, O' \subset (0,1)$ open disjoint sets such that $(0,1) = O \cup O'$ and none of them are empty sets. Let $a \in O, b \in O'$ with $a < b$ without loss of generality. Define $\alpha = \sup\{x \in \mathbb{R} : [a,x) \subset O\}$. It's well defined because this set is not empty, given O is open and b is an upper bound. Then $\alpha \le b$. Suppose $\alpha \in O'$, then there exists $r > 0$ such that $(\alpha - r, \alpha + r) \subset O'$. We know that for every $\epsilon > 0$, there exists $w \in (\alpha - \epsilon, \alpha]$ such that $[a,w) \subset O$. That is a contradiction since there exists $w \in (\alpha - r, \alpha)$ such that $[a,w) \subset O$. So $\alpha \in O \implies (\alpha - r, \alpha + r) \subset O$, for some r. We infer that $[a, \alpha + r) \subset O$, what is an absurd. Therefore $(0,1)$ is connected.

For a similar argument, we prove that $(0,z)$ and $(z,1)$ are connected. This implies that the union admits 2 connected components.

In that sense, we have a homeomorphism between a topological space with 1 connected component and other with 2 connected components, what is a contradiction by Proposition 2.14 from the notes. We conclude that $[0, 1)$ and $(0, 1)$ are not homeomorphic.

$\square$

3 Homotopies

3.1 Important definitions

DEFINITION 3.1.1. *Let* $(X, \mathcal{T})$ *and* $(Y, \mathcal{U})$ *be two topological spaces, and* $f, g : X \to Y$ *two continuous maps. A homotopy between* f *and* g *is a map* $F : X \times [0,1] \to Y$ *such that:*

1. $F(\cdot, 0)$ *is equal to* f*,*

2. $F(\cdot, 1)$ *is equal to* g*,*

3. $F : X \times [0,1] \to Y$ *is continuous.*

If such a homotopy exists, we say that the maps f *and* g *are homotopic.*

Remark. Before asking for $F : X \times [0,1] \to Y$ to be continuous, we have to give $X \times [0,1]$ a topology. The topology we choose is the product topology. Consider the topological space $(X, \mathcal{T})$, and endow $[0,1]$ with the subspace topology of $\mathbb{R}$, denoted $T_{|[0,1]}$. The product topology on $X \times [0,1]$, denoted $T \otimes T_{|[0,1]}$, is defined as follows: a set $O \subset X \times [0,1]$ is open if and only if it can be written as a union $U_{\alpha \in A} O_\alpha \times O'_\alpha$ where every O_α is an open set of X and O'_α is an open set of $[0,1]$.

DEFINITION 3.1.2. *Let* $(X, \mathcal{T})$ *and* $(Y, \mathcal{U})$ *be two topological spaces. A homotopy equivalence between* X *and* Y *is a pair of continuous maps* $f : X \to Y$ *and* $g : Y \to X$ *such that:*

1. $g \circ f : X \to X$ *is homotopic to the identity map* $id : X \to X$*,*

2. $f \circ g : Y \to Y$ *is homotopic to the identity map* $id : Y \to Y$*,*

If such a homotopy equivalence exists, we say that X *and* Y *are homotopy equivalent.*

DEFINITION 3.1.3. *Let* $(X, \mathcal{T})$ *be a topological space and* $Y \subset X$ *a subset, endowed with the subspace topology* $T_{|Y}$*. A retraction is a continuous map* $r : X \to Y$ *such that* $\forall y \in Y, r(y) = y$*.*

A deformation retraction is a homotopy $F : X \times [0,1] \to Y$ *between the identity map* $id : X \to X$ *and a retraction* $r : X \to Y$*.*

3.2 Exercises

EXERCISE 12. *Let* $f : \mathbb{R}^n \to X$ *be a continuous map. Then* f *is homotopic to a constant map.*

Proof. I must prove that there exists a homotopy between f and a constant map. Consider the function $F : \mathbb{R}^n \times [0,1] \to X$ defined as

$$F(x,t) = f(tx)$$

It's clear that $F(x,0) = f(0)$, for every $x \in \mathbb{R}^n$. So it's the constant map $f(0)$. We also have that $F(x,1) = f(x), \forall x \in \mathbb{R}^n$. Moreover, let's prove F is continuos. Denote $F' : \mathbb{R}^n \times \mathbb{R} \to X$ the function $F'(x,t) = f(xt)$ and $g : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ the function $g(x,t) = xt$. So $F' = f \circ g$.

Let's prove g is a continuous function. As we are dealing with a real-valued function, by Proposition 1.19 from the notes, I can use the $\epsilon - \delta$ proof. Let $(x,t) \in \mathbb{R}^{n+1}$ and $\epsilon > 0$. In the proof I use the norm 1, without loss of generality because of the equivalence of norms in $\mathbb{R}^n$. Put $\delta = \min\{1, \frac{\epsilon}{\max\{||x||, |t|+1\}}\}$ and suppose $||(x,t) - (x',t')|| = ||x - x'|| + |t - t'| < \delta$. So,

$$\begin{aligned} ||xt - x't'|| &= ||xt - xt' + xt' - x't'|| \\ &\leq |t - t'|||x|| + |t'|||x - x'|| \\ &\leq |t - t'|||x|| + (|t| + \delta)||x - x'|| \\ &< \max\{||x||, |t| + \delta\}\delta \\ &\leq \max\{||x||, |t| + 1\}\delta \leq \epsilon \end{aligned}$$

By this, g is a continuos function. Since f is also continuos, the composition F' is also continuos, by Proposition 1.18. By Proposition 1.21, when we endow F' in $\mathbb{R}^n \times [0,1]$, we obtain a continuos function, that is F is continuos. Then we conclude that f is homotopic to a constant function.

□

EXERCISE 13. *Let $f : \mathbb{S}_1 \to \mathbb{S}_2$ be a continuous map which is not surjective. Prove that it is homotopic to a constant map where the unit sphere $\mathbb{S}_{n-1} = \{x \in \mathbb{R}^n, ||x|| = 1\}$.*

Proof. Let $x_0 \in \mathbb{S}_2$ such that $x_0 \notin f(\mathbb{S}_1)$ and consider the constant map $g(x) = -x_0$, for every $x \in \mathbb{S}_1$. Let $F : \mathbb{S}_1 \times [0,1] \to \mathbb{S}_2$ be defined as

$$F(x,t) = \frac{(1-t)f(x) - tx_0}{||(1-t)f(x) - tx_0||}$$

The first thing we must prove it's well defined. Suppose that $(1-t)f(x) - tx_0 = 0$. If $t = 1$, then $x_0 = 0$, an absurd given that $||x_0|| = 1$. If $t < 1$, $f(x) = \frac{t}{1-t}x_0$ and applying the norm on both sides $1 = ||f(x)|| = \frac{t}{1-t}||x_0|| = \frac{t}{1-t} \implies t = 1/2$. If that is the case, $f(x) - x_0 = 0 \implies f(x) = x_0$, contradiction. Moreover, for all x and t, $||F(x,t)|| = 1 \implies F(\mathbb{S}_1, [0,1]) \subset \mathbb{S}_2$.

Now let's prove it's a homotopy:

1. $F(x,0) = \frac{f(x)}{||f(x)||} = f(x), \forall x \in \mathbb{S}_1$.
2. $F(x,1) = \frac{-x_0}{||x_0||} = -x_0, \forall x \in \mathbb{S}_1$.
3. Consider the extension of the function $F' : \mathbb{S}_1 \times [0,1] \to \mathbb{R}^3$. This function is continuous because it's a combination os continuos function. So F is continuos because it's a restriction of F'. I needed to extend the function because $(1-t)f(x)$ is not necessary in the sphere, so I couldn't prove it's continuos. However, when extended we see each part is continuos.

By (1) - (3), we have proved F is a homotopy and f are homotopic to a constant function.

Remark. If the function is not surjective, it's harder to prove, and I couldn't yet. For instance, this is a reference[1] (but the answers use specialized tools)

□

---

[1] https://math.stackexchange.com/questions/3807715/

Exercise 14. *Show that being homotopic is a transitive relation between maps: for every triplet of maps $f, g, h : X \to Y$, if f, g are homotopic and g, h are homotopic, then f, h are homotopic.*

Proof. We shall prove there exist a homotopy H between f and h. By assumption, there exists a homotopy F between f and g and a homotopy G between g and h. Define $H : X \times [0,1] \to Y$ such that

$$H(x,t) = \begin{cases} F(x,2t), & 0 \le t \le 1/2 \\ G(x,2t-1), & 1/2 < t \le 1 \end{cases}$$

that is, H behaves as F until it reaches a half. When that occurs, $H(x,1/2) = F(x,1) = g(x) = G(x,0)$. After that, H follows G until the end of the interval. So, it's clear that $H(x,0) = F(x,0) = f(x), \forall x \in X$ and $H(x,1) = G(x,1) = h(x), \forall x \in X$. Moreover, let's prove H is continuos. Let $A \subset Y$ closed set. So

$$\begin{aligned} H^{-1}(A) &= \{(x,t) : F(x,2t) \in A, t \le 1/2\} \cup \{(x,t) : G(x,2t-1) \in A, t \ge 1/2\} \\ &= \tilde{F}^{-1}(A) \cup \tilde{G}^{-1}(A), \end{aligned}$$

where $\tilde{F}(x,t) = F(x,2t)$ and $\tilde{G}(x,t) = G(x,2t-1)$, with domain being, respectively, $X \times [0,1/2]$ and $X \times [1/2,1]$. As we see in the following remark, both functions are continuos, so $\tilde{F}^{-1}(A) \cup \tilde{G}^{-1}(A)$ is closed, what proves H is continuos and, therefore, f and h are homotopic.

Remark. The map $(x,t) \mapsto (x,2t)$ is continuos. In order to see that, take $A \subset X \times [0,1]$ open. So it can be written as $\bigcup_a O_a \times I_a$, where $I_a \subset [0,1]$ is an open interval and $O_a \subset X$ is open. The pre-image of A is given by $\bigcup_a O_a \times \frac{1}{2} I_a$, that is still open. With that and using the fact that combination of continuos functions is continuos, $\tilde{F}$ is continuos. For a similar argument, $\tilde{G}$ is continuos.

□

Exercise 15. *Show that being homotopy equivalent is an equivalence relation (reflexive, symmetric and transitive).*

Proof.

1. (*reflexive*): Consider the identity map $id : X \to X$, that is continuos. We shall prove that this function is homotopic to itself. Consider $F : X \times [0,1] \to X$ given by $F(x,t) = x$ for every x and t. It's clear this is a homotopy because $F(x,0) = F(x,1) = x$ and it's continuos. Moreover $id \circ id = id$ by definition of identity. Therefore, there exists a homotopy equivalence between id and itself. We conclude $X \approx X$.

2. (*symmetric*): Suppose $X \approx Y$. So, there exists continuos functions $f : X \to Y$ and $g : Y \to X$ that form a homotopy equivalence. This means that $g : Y \to X$ and $f : X \to Y$ are a homotopy equivalence as well. So $Y \approx X$.

3. (*transitive*): Suppose $X \approx Y$, and let $f_1 : X \to Y$ and $g_1 : Y \to X$ form a homotopy equivalence. Also suppose $Y \approx Z$ and let $f_2 : Y \to Z$ and $g_2 : Z \to Y$ form a homotopy equivalence. Define $f_3 = f_2 \circ f_1$ and $g_3 = g_1 \circ g_2$. Let's proof this is a homotopy equivalence. Both functions are continuos given that they are a composition of continuos functions.

   (a) $g_3 \circ f_3 = g_1 \circ g_2 \circ f_2 \circ f_1$ is homotopic to $id : X \to X$.

Let F_1 be a homotopy between $g_1 \circ f_1$ and id and F_2 a homotopy between $g_2 \circ f_2$ and id. Define

$$F_3(x,t) = \begin{cases} g_1 \circ F_2(\cdot, 2t) \circ f_1(x), & 0 \leq t \leq 1/2 \\ F_1(x, 2t-1), & 1/2 < t \leq 1 \end{cases}$$

So $F_3(x,0) = g_1(F_2(f_1(x),0)) = g_1(g_2(f_2(f_1(x)))) = g_3 \circ f_3(x)$, for every x and $F_3(x,1) = F_1(x,1) = x$, for every x. When $t = 1/2$,

$$F_3(x,1/2) = g_1(F_2(f_1(x),1)) = g_1(f_1(x)) = F_1(x,0)$$

F_3 is continuos by a similar proof written in the last exercise. This implies that $g_3 \circ f_3$ is homotopic to the identity.

(b) $f_3 \circ g_3 = f_1 \circ f_2 \circ g_2 \circ g_1$ is homotopic to $id : Z \to Z$.
This follows a quite similar demonstration and can be omitted.

By the points above f_3 and g_3 is a homotopy equivalence what proves $X \approx Z$. Consequently, homotopy equivalence is an equivalence relation.

□

Exercise 16. *Classify the letters of the alphabet into homotopy equivalence classes.*

Proof. I will consider the upper case alphabet and each letter will be considered as a topological space (a subset from $\mathbb{R}^2$), for example the letter O is homotopy equivalent to a circle, while L is to an interval, or equivalently, a point. Observe that most of the letters are equivalent to a point, because we can think in a continuos reduction. When we have a hole, such as A, D, R, O, P, Q, this continuity is impossible since we'll have a point break. B is a special case because we can't deform into a point without breaking points and also we cannot join the holes in one. So there is three classes, given by its representatives

1. O
2. B
3. I

A	B	C	D	E	F	G	H	I	J	K	L	M
1	2	3	1	3	3	3	3	3	3	3	3	3

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
3	1	1	1	1	3	3	3	3	3	3	3	3

□

4 Simplicial complexes

4.1 Important definitions

Definition 4.1.1. *The standard simplex of dimension n is the following subset of R^{n+1},*

$$\Delta_n = \{x = (x_1, ..., x_{n+1}) \in \mathbb{R}^{n+1} | \forall i, x_i \geq 0 \text{ and } \sum_{i=1}^{n+1} x_1 = 1\}.$$

For any collection of points $a_1, ..., a_k \in \mathbb{R}^n$, we define their convex hull as:

$$conv(\{a_1, ..., a_k\}) = \left\{ \sum_{1 \leq i \leq k} t_i a_i | \sum_{1 \leq i \leq k} t_i = 1, t_1, ..., t_k \geq 0 \right\}$$

Definition 4.1.2. *Let V be a set (called the set of vertices). A simplicial complex over V is a set K of subsets of V (called the simplices) such that, for every $\sigma \in K$ and every non-empty $\tau \subset \sigma$, we have $\tau \in K$. If $\sigma \in K$ is a simplex, its non-empty subsets $\tau \subset \sigma$ are called faces of σ, and σ is called a coface of τ. Moreover, its dimension is $|\sigma| - 1$ and the dimension of a simplicial is the maximum dimension of its simplices.*

Definition 4.1.3. *Let K be a simplicial complex, with vertex $V = [\![1, n]\!]$. In $\mathbb{R}^n$, consider, for every $i \in [\![1, n]\!]$, the vector $e_i = (0, ..., 1, 0, ..., 0)$ (i^{th} coordinate 1, the other ones 0). Let $|K|$ be the subset of $\mathbb{R}^n$ defined as:*

$$|K| = \bigcup_{\sigma \in K} conv(\{e_j, j \in \sigma\}).$$

Endowed with the subspace topology, $(|K|, T_{||K|})$ is a topological space, that we call the ***topological realization*** *of K.*

Definition 4.1.4. *Let $(X, \mathcal{T})$ be a topological space. A triangulation of X is a simplicial complex K such that its topological realization $(|K|, T_{||K|})$ is homeomorphic to $(X, \mathcal{T})$.*

Definition 4.1.5. *Let K be a simplicial complex of dimension n. Its Euler characteristic is the integer*

$$\chi(K) = \sum_{0 \leq i \leq n} (-1)^i \cdot \text{(number of simplices of dimension } i\text{)}.$$

Definition 4.1.6. *The Euler characteristic of a topological space is the Euler characteristic of any triangulation of it.*

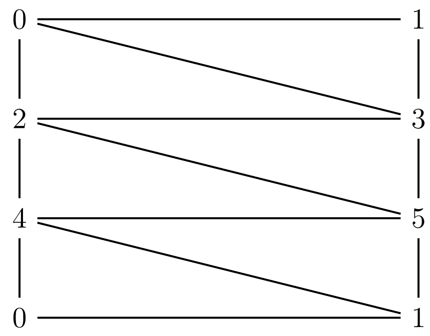
4.2 Exercises

Exercise 17. *Give a triangulation of the cylinder.*

Proof. We can think a triangulation of the cylinder in that following form: each circular section is mapped into a triangle graph. On the other hand, the line can be mapped to an edge. Since a cylinder can be written as $\mathbb{S}_1 \times \mathbb{R}$, the triangulation as well.

Let's write down:

$$K = \{[0, 1, 3], [0, 2, 3], [2, 3, 5], [2, 4, 5], [4.5, 1], [4, 1, 0]\}$$



□

Exercise 18. *What are the Euler characteristics of Examples 4.5 and 4.6? What is the Euler characteristic of the icosahedron?*

Proof. **Exemplo 4.5:**

$$K = \{[0],[1],[2],[3],[0,1],[1,2],[2,3],[3,0],[0,2],[1,3],[0,1,2],[0,1,3],[0,2,3],[1,2,3]\}.$$

$$\chi(K) = 4 - 6 + 4 = 2$$

Exemplo 4.6:

$$K = \{[0],[1],[2],[3],[0,1],[1,2],[1,3],[0,2],[2,3],[3,0],$$
$$[0,1,2],[0,1,3],[0,2,3],[1,2,3],[0,1,2,3]\}.$$

$$\chi(K) = 4 - 6 + 4 - 1 = 1$$

D20: It has 20 faces (dimension 2), 30 edges (dimension 1) and 12 vertices (dimension 0), its Euler characteristic is $12 - 30 + 20 = 2$ (Euler relation).

□

Exercise 19. *Let K be a simplicial complex (with vertex set V). A sub-complex of K is a set $M \subset K$ that is a simplicial complex. Suppose that there exists two sub-complexes M and N of K such that $K = M \cup N$. Show the inclusion-exclusion principle:*

$$\chi(K) = \chi(M) + \chi(N) - \chi(M \cap N)$$

Proof. Denote k_i the number of simplices of dimension i, that is, $\chi(K) = \sum_{0\le i\le k}(-1)^i k_i$, with k being its dimension. Each simplex of dimension i belongs to M, to N or both. Denote k_i^M, k_i^N and k_i^{MN} the number od simplices of dimension i in M, N and $M \cap N$ respectively. So

$$k_i = k_i^M + k_i^N - k_i^{MN}.$$

Therefore,

$$\chi(K) = \sum_{0\le i\le k} (-1)^i (k_i^M + k_i^N - k_i^{MN}) = \sum_{0\le i\le k} (-1)^i k_i^M + \sum_{0\le i\le k} (-1)^i k_i^N - \sum_{0\le i\le k} (-1)^i k_i^{MN}$$

Let m and n be the dimension of M and N, respectively. Now I shall prove $M \cap N$ is a simplicial complex. Take $\sigma \subset M\cap N$ and $\tau \subset \sigma$, then $\tau \subset \sigma \in M$ and $\tau \subset \sigma \in N$, what implies $\tau \in M \cap N$. It proves its a simplicial complex with dimension p. We know the dimension of M is m, so it implies that for all $i > m$, $k_i^M = 0$. Suppose not and let $k_j^M > 0$ for some $j > m$, we would have a simplex with dimension greater or equal than m. This is a contradiction, because the the dimension of M is the maximum dimension of its simplices. This holds for N and $M \cap N$.

$$\chi(K) = \sum_{0\le i\le k} (-1)^i (k_i^M + k_i^N - k_i^{MN}) = \sum_{0\le i\le k} (-1)^i k_i^M + \sum_{0\le i\le k} (-1)^i k_i^N - \sum_{0\le i\le k} (-1)^i k_i^{MN}$$

We conclude that

$$\chi(K) = \sum_{0\le i\le m} (-1)^i k_i^M + \sum_{0\le i\le n} (-1)^i k_i^N - \sum_{0\le i\le p} (-1)^i k_i^{MN} = \chi(M) + \chi(N) - \chi(M \cap N)$$

□

EXERCISE 20. *What is the Euler characteristic of a sphere of dimension 1? 2? 3?*

Proof. We may find the Euler characteristic of one triangulation of the sphere. So we first need to find a triangulation for the sphere $\mathbb{S}_n \subset \mathbb{R}^{n+1}$. We can think in the simplex in this space. In $\mathbb{R}^2$ it's the triangle, in $\mathbb{R}^3$ it's the tetrahedron, in $\mathbb{R}^4$ it's the 5-cell and so on. The simplex in $\mathbb{R}^{n+1}$ has $n+2$ vertices and each vertex connect to all the other. We have also $n+2$ simplices of dimension n, because each has $n+1$ points, that is, $\binom{n+2}{n+1} = n+2$. For each of them, we must include all its subsets.

Now we can calculate the Euler characteristic for each triangulation and therefore each sphere.

$$\chi(\mathbb{S}_1) = -3 + 3 = 0$$

$$\chi(\mathbb{S}_2) = 4 - 6 + 4 = 2$$

$$\chi(\mathbb{S}_3) = -5 + 10 - 10 + 5 = 0$$

$$\chi(\mathbb{S}_4) = 6 - 15 + 20 - 15 + 6 = 2$$

□

EXERCISE 21. *Using the previous exercise, show that $\mathbb{R}^3 - \{0\}$ and $\mathbb{R}^4 - \{0\}$ are not homotopy equivalent.*

Proof. Suppose that $\mathbb{R}^3 - \{0\}$ and $\mathbb{R}^4 - \{0\}$ are homotopy equivalent. By Example 3.15, $\mathbb{S}_{n-1}$ is homotopic equivalent to $\mathbb{R}^n - \{0\}$. By this and using the transitive property we conclude that $\mathbb{S}_2$ and $\mathbb{S}_3$ are homotopic equivalent. If that is true, we infer that they have the same Euler characteristic, what is a contraction by the last exercise. Hence $\mathbb{R}^3 - \{0\}$ and $\mathbb{R}^4 - \{0\}$ are not homotopy equivalent.

□

The computational exercises (22 - 26) can be found in the Github[2]

Exercise 22. *Build triangulations of the letters of the alphabet, and compute their Euler characteristic.*

Exercise 23. *For every n, triangulate the bouquet of n circles (see below). Compute their Euler characteristic.*

Exercise 24. *Implement the following triangulation of the torus.*

Exercise 25. *Consider the following dataset of 30 points $x_0, ..., x_{29}$ in R^2.*

Write a function that takes as an input a parameter $r \geq 0$, and returns the simplicial complex $\mathcal{G}(r)$ defined as follows:

1. *the vertices of $\mathcal{G}(r)$ are the points $x_0, ..., x_{29}$,*
2. *for all $i, j \in [[0, 29]]$ with $i \neq j$, the edge $[i, j]$ belongs to $\mathcal{G}(r)$ if and only if $||x_i - x_j|| \leq r$.*

Compute the number of connected components of $\mathcal{G}(r)$ for several values of r. What do you observe?

Exercise 26. *A Erdos–Renyi random graph $\mathcal{G}(n, p)$ is a simplicial complex obtained as follows:*

1. *add n vertices $1, ..., n$,*
2. *for every $a, b \in [[1, n]]$, add the edge $[a, b]$ to the complex with probability p.*

Builds a function that, given n and p, outputs a simplicial complex $\mathcal{G}(n, p)$. Observe the influence of p on the number of connected components of $\mathcal{G}(10, p)$ and $\mathcal{G}(100, p)$.

[2] github.com/lucasmoschen/topological-data-analysis/blob/main/tutorials/tutorial-1.ipynb

5 Homological algebra

5.1 Important Definitions

DEFINITION 5.1.1. *Let K be a simplicial complex. For any $n \geq 0$, define the sets*

$$K_n = \{\sigma \in K, dim(\sigma) \leq n\}$$

$$K_{(n)} = \{\sigma \in K, dim(\sigma) = n\}.$$

The first set is a simplicial complex, called the n-skeleton of K.

DEFINITION 5.1.2. *Let $n \geq 0$. The n-chains of K is the set $C_n(K)$ whose elements are the formal sums*

$$\sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \text{ where } \forall \sigma \in K_{(n)}, \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z}.$$

We can give $C_n(K)$ a group structure via

$$\sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma + \sum_{\sigma \in K_{(n)}} \eta_\sigma \cdot \sigma = \sum_{\sigma \in K_{(n)}} (\epsilon_\sigma + \eta_\sigma) \cdot \sigma.$$

DEFINITION 5.1.3. *Let $n \geq 1$, and $\sigma \in K_{(n)}$ a simplex of dimension n. We define its boundary as the following element of $C_{n-1}(K)$:*

$$\partial_n \sigma = \sum_{\tau \subset \sigma, |\tau| = |\sigma| - 1} \tau$$

and

$$\partial_n \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma = \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \partial_n \sigma$$

DEFINITION 5.1.4. *We define*

1. *The n-cycles: $Z_n(K) = Ker(\partial_n)$,*

2. *The n-boundaries: $B_n(K) = Im(\partial_{n+1})$.*

We say that two chains $c, c' \in C_n(K)$ are homologous if there exists $b \in B_n(K)$ such that $c = c' + b$.

DEFINITION 5.1.5. *The n^{th} homology group of K is $H_n(K) = Z_n(K)/B_n(K)$.*

DEFINITION 5.1.6. *Let K be a simplicial complex and $n \geq 0$. Its n^{th} Betti number is the integer $\beta_n(K) = dim\ H_n(K)$.*

DEFINITION 5.1.7. *The homology groups of a topological space are the homology groups of any triangulation of it. We define their Betti numbers similarly.*

5.2 Exercises

EXERCISE 27. *Let V be a $\mathbb{Z}/2\mathbb{Z}$-vector space, and W a linear subspace. Prove that*

$$dim\ V/W = dim\ V - dim\ W.$$

Proof. Suppose V is finite-dimensional, such that $dim\ V = n+m$ and $dim\ W = m$. As we have a finite basis and a finite field, we only have finite number of combinations from the vector of this basis. In special V is finite. By Proposition 5.2, V has cardinal 2^{m+n} and W has cardinal 2^m. Let's prove that V/W has cardinal 2^n. We know the quotient space has finite dimension k, because it's finite. So it has 2^k elements. Let $V/W = \{[v_1], ..., [v_{2^k}]\}$ where v_i represents an equivalence class. We know for each $v \in V$ there exists i such that $v \in v_i + W$ and there is $w \in W$ with $x = v_i + w$. As each class is disjoint, we first choose one of the 2^k classes v_i. After we pick out one element of this class. We have 2^m elements in W. Let's see $|v_i + W| = 2^m$. Take $w, w' \in W$ and suppose $v_i + w = v_i + w'$. Summing v_i in both sides we obtaing that $w = w'$. So we have $2^k 2^m$ ways of forming elements os V. We conclude k must be n, what finishes our proof.

Remark. Consider the following proof slightly more general.

Suppose V is finite-dimensional and let $\{w_1, ..., w_m\}$ be a basis of W. So we can extend this to a basis on V, namely, $\{w_1, ..., w_m, v_1, ..., v_n\}$, where, dim $V = m+n$. Consider the set $\{v_1 + W, ..., v_n + W\}$. First, let's prove it's free.

$$0 + W = \sum_{j=1}^{n} \lambda_j (v_j + W) = \left(\sum_{j=1}^{n} \lambda_j v_j \right) + W,$$

So $\left(\sum_{j=1}^{n} \lambda_j v_j\right) \in W$. But it implies it can be written as a linear combination of $\{w_1, ..., w_m\}$, what contradicts the fact that $\{w_1, ..., w_m, v_1, ..., v_n\}$ is linear independent.

Now take $v + W \in V/W$, where $v = \sum_{j=1}^{m} \lambda_j w_j + \sum_{j=1}^{n} \lambda_{j+m} v_j$. So

$$\begin{aligned} v + W &= \left[\sum_{j=1}^{m} \lambda_j w_j + \sum_{j=1}^{n} \lambda_{j+m} v_j \right] + W \\ &= \sum_{j=1}^{m} \lambda_j (w_j + W) + \sum_{j=1}^{n} \lambda_{j+m} (v_j + W) \\ &= \sum_{j=1}^{n} \lambda_{j+m} (v_j + W) \end{aligned}$$

We conclude $\{v_1, ..., v_n\}$ is a basis what implies dim $V/W = n = m + n - m$.

□

Exercise 28. *Let $(G, +)$ be a group, potentially non-commutative. Prove that*

$$\forall g \in G, g + g = 0 \implies G \textit{ is commutative.}$$

Proof. Let $u, v \in G$. So $u + v + v = u + (v + v) = u + 0 = u$. With that in mind, we see that

$$u + v + v + u = 0 \implies (u + v) + (v + u) = 0$$

We know $v + u$ and $u + v$ are elements of G. So we can add to each side and obtain

$$(u + v) + (v + u) + (v + u) = (v + u) \implies u + v = v + u$$

As u and v are arbitrary, G is commutative.

□

Exercise 29. *Compute the Betti numbers $\beta_0(K), \beta_1(K)$ and $\beta_2(K)$ of the following simplicial complex:*

$$K = \{[0],[1],[2],[3],[0,1],[1,2],[2,3],[3,0]\}.$$

Proof. $Z_0(K) = C_0(K)$

$B_0(K) = \{[0]+[1],[1]+[2],[2]+[3],[3]+[0],[0]+[2],[1]+[3],[0]+[1]+[2]+[3],0\}$

As B_0 have 2^3 elements and Z_0 has 2^4, we deduce that

$$\beta_0(K) = dim\ Z_0(K) - dim\ B_0(K) = 4 - 3 = 1$$

$Z_1(K) = \{[0,1]+[1,2]+[2,3]+[3,0],0\}$

$B_1(K) = \{0\}$

Nesse caso, observamos que, utilizando a mesma ideia do ponto anterior

$$\beta_1(K) = 1 - 0 = 1.$$

$$K = \{[0],[1],[2],[3],[0,1],[1,2],[2,3],[3,0]\}.$$

$Z_2(K) = \{0\}$

$B_2(K) = \{0\}$

Portanto $\beta_2(K) = 0$.

□

Exercise 30. *Compute the Betti numbers $\beta_0(K), \beta_1(K)$ and $\beta_2(K)$ of the following simplicial complex:*

$$K = \{[0],[1],[2],[3],[0,1],[1,2],[2,3],[3,0],[0,2],[1,3],[0,1,2],[0,1,3],[0,2,3],[1,2,3]\}.$$

Proof. $Z_0(K) = C_0(K)$

$B_0(K) = \{[0]+[1],[1]+[2],[2]+[3],[3]+[0],[0]+[2],[1]+[3],[0]+[1]+[2]+[3],0\}$

As B_0 have 2^3 elements and Z_0 has 2^4, we deduce that

$$\beta_0(K) = dim\ Z_0(K) - dim\ B_0(K) = 4 - 3 = 1$$

$$\begin{aligned} Z_1(K) = \{&[0,1]+[1,2]+[0,2];[0,1]+[0,3]+[1,3];[0,2]+[2,3]+[0,3];\\ &[1,2]+[2,3]+[1,3];[0,1]+[1,2]+[2,3]+[0,3];[0,1]+[1,3]+[2,3]+[2,0];\\ &[0,2]+[1,2]+[1,3]+[0,3];0\} \end{aligned} \quad (1)$$

$$B_1(K) = \{[0,1]+[1,2]+[0,2]; [0,1]+[0,3]+[1,3]; [0,2]+[2,3]+[0,3];$$
$$[1,2]+[2,3]+[1,3]; [0,1]+[1,2]+[2,3]+[0,3]; [0,1]+[1,3]+[2,3]+[0,2];$$
$$[0,2]+[1,2]+[1,3]+[0,3]; 0\} \quad (2)$$

Observe que os conjuntos são iguais e, portanto,

$$\beta_1(K) = 0.$$

$Z_2(K) = \{[0,1,2]+[0,1,3]+[0,2,3]+[1,2,3], 0\}$
$B_2(K) = \{0\}$
Portanto $\beta_2(K) = 1$.

$\square$

6 Incremental algorithm

6.1 Important definitions

DEFINITION 6.1.1. *Let $i \in [[1, n]]$, and $d = dim(\sigma_i)$. The simplex σ_i is positive if there exists a cycle $c \in Z_d(K^i\ )$ that contains σ_i. Otherwise, σ_i is negative.*

DEFINITION 6.1.2. *Defines for a set $V = \{0, 1, ..., n\}$ a simplicial complex*

$$\Delta_n = \{S \subset C, S \neq \emptyset\}$$

and call it simplicial standard n-simplex with boundary

$$\partial\Delta_n = \Delta_n/V$$

DEFINITION 6.1.3. *Define the boundary matrix of K, denoted Δ as follows: Δ is a $n \times n$ matrix, whose*

$$\Delta_{i,j} = \begin{cases} 1, & \text{if } \sigma_i \subset \sigma_j \text{ and } |\sigma^i| = |\sigma^j| - 1 \\ 0, & \text{else} \end{cases}$$

6.2 Exercise

EXERCISE 31. *Compute again the Betti numbers of the simplicial complexes of Exercises 29 and 30, using the incremental algorithm.*

Proof. 1. $K = \{[0], [1], [2], [3], [0, 1], [1, 2], [2, 3], [3, 0]\}$.

First we determine the ordering to be as placed in the set. It fulfills the required property. After, we find the sign s for it σ^i. The first four elements are positive, because, ∂_0 has $C_0(K^i)$ as kernel. On the other hand $[0, 1]$, $[1, 2]$ and $[2, 3]$ are negatives. At last $[3, 0]$ is positive, because $[0, 1] + [1, 2] + [2, 3] + [3, 0]$ belongs to $Z_1(K^8)$. Now we can follow the algorithm thorough a table:

-	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	σ^8
Sign	+	+	+	+	-	-	-	+
$\beta_0(K)$	1	2	3	4	3	2	1	1
$\beta_1(K)$	0	0	0	0	0	0	0	1
$\beta_2(K)$	0	0	0	0	0	0	0	0

2. $K = \{[0], [1], [2], [3], [0, 1], [1, 2], [2, 3], [3, 0], [0, 2], [1, 3], [0, 1, 2], [0, 1, 3], [0, 2, 3], [1, 2, 3]\}$.

First we determine the ordering to be as placed in the set. It fulfills the required property. After, we find the signs for it σ^i. The vertices have positive sign. The following three edges cannot form any cycle, so they are negative. The last three edges are part of a cycle considering three other already placed of its dimension. When we achieve the simplices with dimension 2, the first three must be negative, because when we sum every combination of them, the boundary has image different from 0. The last, however will be positive. Now we can follow the algorithm thorough a table:

-	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	σ^8	σ^9	σ^{10}	σ^{11}	σ^{12}	σ^{13}	σ^{14}
Sign	+	+	+	+	-	-	-	+	+	+	-	-	-	+
$\beta_0(K)$	1	2	3	4	3	2	1	1	1	1	1	1	1	1
$\beta_1(K)$	0	0	0	0	0	0	0	1	2	3	2	1	0	0
$\beta_2(K)$	0	0	0	0	0	0	0	0	0	0	0	0	0	1

The result corroborates with those found previously.

□

EXERCISE 32. *Prove that $\partial\Delta_n$ is a triangulation of the $(n-1)$-sphere $\mathbb{S}_{n-1} \subset \mathbb{R}^n$.*

Proof. It's clear that $\partial\Delta_n$ is a simplicial complex, because, for every simplex $\sigma \in \partial\Delta_n$ and $\tau \subset \sigma$, $\tau \in \Delta_n$ - since it's a simplicial complex - and $\tau \neq V$, then $\tau \in \partial\Delta_n$. I shall prove that the topological realization of Δ_n, denoted as B_n, is homeomorphic to the $\mathbb{S}_{n-1}$. We can describe it

$$B_n = \{(\alpha_0, ..., \alpha_n) \in [0,1]^{n+1}, \sum_{i=0}^{n} \alpha_i = 1 \text{ and for some } i, \alpha_i = 0\}$$

Let $H = \{(x_0, ..., x_n) \in \mathbb{R}^{n+1} | \sum_{i=0}^{n} x_i = 1\}$. It's clear that $B_n \subset H$. Define $C_n := \partial\mathcal{B}(b, r) \cap H$ such that $C_n \subset$ interior Δ_n. We'll show it's homeomorphic to B_n.

Let $b = \frac{1}{n+1}(1, ..., 1) \in \mathbb{R}^{n+1}$ and take $x \in C_n$. Consider the half line which start in b and crosses x. Its equation can be described as $b + \alpha(x - b)$. It intersects B_n when at least one of its coordinates is zero and the other are between 0 and 1. The sum will be always 1 because we are in H. Denote $x_{(m)} = \min\{x_i : i^{th}\text{coordinate of } x\}$. Let's take α such that

$$b_{(m)} + \alpha(x_{(m)} - b_{(m)}) = 0 \implies \alpha = \frac{1}{1 - (n+1)x_{(m)}}$$

For every other $i \neq (m)$,

$$\begin{aligned} b_i + \alpha(x_i - b_i) &= \frac{1}{n+1}\left(1 + \frac{1}{1-(n+1)x_{(m)}}(x_i(n+1) - 1)\right) \\ &= \frac{1}{n+1}\left(1 - \frac{1}{1-(n+1)x_{(m)}}(1 - x_i(n+1))\right) \\ &= \frac{1}{n+1}\left((n+1)\frac{x_i - x_m}{1-(n+1)x_{(m)}}\right) \\ &= \frac{x_i - x_m}{1-(n+1)x_{(m)}} \geq 0 \end{aligned}$$

since $x_{(m)} < 1/(n+1)$. Note that $1 = \sum_{j=0}^{m} x_j \geq x_i + nx_{(m)} \implies 1 - (n+1)x_{(m)} \geq x_i - x_m$ and we conclude that $b_i + \alpha(x_i - b_i) \in [0,1]$, that is, $b - \alpha(x - b) \in B_n$. Denote this map f. Let's prove it's a homeomorphism.

1. f is injective: suppose not. In this case, there is some point $p \in B_n$ and two points $c_1, c_2 \in C_n$ such that there are two segments starting at p and crossing B_n at p, however it one passes thorough c_1 and c_2 respectively. So both segments have two points in common, what implies they are equal. Suppose, no loss of generality, that c_1 is between p and c_2. Then $1 = ||c_2 - b|| < ||c_1 - b|| = 1$, a contradiction and f is injective.

2. f is surjective: Take y from B_n and make the half segment b to y. It's clear it crosses C_n because p is in its interior.

3. f is continuos: First we note that the map $x \mapsto x_{(m)}$ is continuos. α can be seen as map from B_n to $[1, +\infty)$. If we extend this map to a map $\alpha' : \mathbb{R}^{n+1} \to \mathbb{R}$, we see that α' is continuous by composing continuous functions and its restriction α will be continuos as well. Using the composition property, we infer f is continuos.

4. f^{-1} is continuos: The inverse is quite similar. We consider $y \in B_n$ and the half line $b + \alpha(y - b)$. We map y to the value such $||b + \alpha(y - b)||^2 = r^2$. We obtain, using the norm 2,
$$\alpha = \frac{r^2(n+1) - 1}{(n+1)||y - b||}$$
what is a continuous function of y.

Therefore B_n and C_n are homeomorphic by f. Now we must prove that C_n is homeomorphic to $\mathbb{S}_{n-1}$, because if this is true, by transitivity, $B_n \simeq \mathbb{S}_{n-1}$.

□

EXERCISE 33. *Show that the algorithm stops after a finite number of steps.*

Proof. Consider the algorithm

Algorithm 1: Reduction of the boundary matrix

Input: a boundary matrix Δ
Output: a reduced matrix $\tilde{\Delta}$
for $j \leftarrow 1$ *to* n **do**
    while $\exists i < j$ *such that* $\delta(i) = \delta(j)$ **do**
        add column i to column j
    end
end

Remark. I'd add a condition in the algorithm for undefined $\delta(i)$, that is, we should consider the values in the domain of δ.

I have to prove that for every $1 \leq j \leq n$, the `while` block stops after a finite number of steps. That is, in a finite number of steps $\forall i < j, \delta(i) \neq \delta(j)$. Let's prove by induction on j. If $j = 1$, there is no $i < 1 = j$, then the algorithm goes to $j = 2$. Now suppose by finite number of operations, $\forall i < j, \delta(i) \neq \delta(j)$. Then the process goes to $j + 1$. Suppose it's defined and there is $k = \delta(i) = \delta(j + 1)$. That means $1 = \Delta_{k,i} = \Delta_{k,j+1}$ and $\Delta_{s,i} = \Delta_{s,j+1} = 0, s > k$. When we add columns i and $j + 1$, we will have $\Delta_{s,j+1} = 0$ for all $s \geq k$. Therefore $\delta(j + 1) < \delta(i)$ or it's not defined (the latter implies the while block ends). As we can decrease one unit at most $\delta(j + 1)$ times, we only have a finite number of steps, in this case, a finite number of additions. Since we have only finite values for j, the algorithm ends in a finite number of steps.

□

EXERCISE 34. *Apply Algorithm to solve Exercise 31.*

Proof. So as to solve both exercises, the order will be as placed in the set for each case. After we build the boundary matrix and apply the Algorithm from the last exercise to obtain the reduced matrix and with it, define σ^i as negative if $\delta(i)$ is defined and positive, otherwise.

1. $K = \{[0], [1], [2], [3], [0, 1], [1, 2], [2, 3], [3, 0]\}$.

   This is the boundary matrix

-	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8
σ_1	0	0	0	0	1	0	0	1
σ_2	0	0	0	0	1	1	0	0
σ_3	0	0	0	0	0	1	1	0
σ_4	0	0	0	0	0	0	1	1
σ_5	0	0	0	0	0	0	0	0
σ_6	0	0	0	0	0	0	0	0
σ_7	0	0	0	0	0	0	0	0
σ_8	0	0	0	0	0	0	0	0

   When we apply the algorithm, we start in $i = 5$. However we do not enter the `while` block until $j = 8$ when $\delta(7) = \delta(8) = 4$. So we sum this columns, after σ_6 to σ_8 and at last σ_5 to σ_8. So σ_8 will be the 0-column, while the other remain unaltered. By this, we obtain the signs in order to calculate the Betti numbers.

-	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	σ^8
Sign	+	+	+	+	-	-	-	+
$\beta_0(K)$	1	2	3	4	3	2	1	1
$\beta_1(K)$	0	0	0	0	0	0	0	1
$\beta_2(K)$	0	0	0	0	0	0	0	0

2. $K = \{[0], [1], [2], [3], [0, 1], [1, 2], [2, 3], [3, 0], [0, 2], [1, 3], [0, 1, 2], [0, 1, 3], [0, 2, 3], [1, 2, 3]\}$.

   The boundary matrix is larger than the previous one.

-	$\sigma_{1,2,3,4}$	σ_5	σ_6	σ_7	σ_8	σ_9	σ_{10}	σ_{11}	σ_{12}	σ_{13}	σ_{14}
σ_1	0	1	0	0	1	1	0	0	0	0	0
σ_2	0	1	1	0	0	0	1	0	0	0	0
σ_3	0	0	1	1	0	1	0	0	0	0	0
σ_4	0	0	0	1	1	0	1	0	0	0	0
σ_5	0	0	0	0	0	0	0	1	1	0	0
σ_6	0	0	0	0	0	0	0	1	0	0	1
σ_7	0	0	0	0	0	0	0	0	0	1	1
σ_8	0	0	0	0	0	0	0	0	1	1	0
σ_9	0	0	0	0	0	0	0	1	0	1	0
σ_{10}	0	0	0	0	0	0	0	0	1	0	1
σ_{11}	0	0	0	0	0	0	0	0	0	0	0
σ_{12}	0	0	0	0	0	0	0	0	0	0	0
σ_{13}	0	0	0	0	0	0	0	0	0	0	0
σ_{14}	0	0	0	0	0	0	0	0	0	0	0

This matrix has several more things to deal. First we make the column σ_8 be a 0-column again. After we make the same to σ_9, being replaced by $\sigma_5+\sigma_6+\sigma_9$ and σ_{10} by $\sigma_6+\sigma_7+\sigma_{10}$. We will have

-	$\sigma_{1,2,3,4}$	σ_5	σ_6	σ_7	σ_8	σ_9	σ_{10}	σ_{11}	σ_{12}	σ_{13}	σ_{14}
σ_1	0	1	0	0	0	0	0	0	0	0	0
σ_2	0	1	1	0	0	0	0	0	0	0	0
σ_3	0	0	1	1	0	0	0	0	0	0	0
σ_4	0	0	0	1	0	0	0	0	0	0	0
σ_5	0	0	0	0	0	0	0	1	1	0	0
σ_6	0	0	0	0	0	0	0	1	0	0	1
σ_7	0	0	0	0	0	0	0	0	0	1	1
σ_8	0	0	0	0	0	0	0	0	1	1	0
σ_9	0	0	0	0	0	0	0	1	0	1	0
σ_{10}	0	0	0	0	0	0	0	0	1	0	1
σ_{11}	0	0	0	0	0	0	0	0	0	0	0
σ_{12}	0	0	0	0	0	0	0	0	0	0	0
σ_{13}	0	0	0	0	0	0	0	0	0	0	0
σ_{14}	0	0	0	0	0	0	0	0	0	0	0

The next column to deal is σ_{13}. Summing it with σ_{11} will solve the problem. With σ_{14} we need to replace with $\sigma_{12}+\sigma_{13}\sigma_{14}$, considering the new σ_{13}. But this will turn σ_{14} to be a 0-column. We know, then, the signs for each column.

-	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	σ^8	σ^9	σ^{10}	σ^{11}	σ^{12}	σ^{13}	σ^{14}
Sign	+	+	+	+	-	-	-	+	+	+	-	-	-	+
$\beta_0(K)$	1	2	3	4	3	2	1	1	1	1	1	1	1	1
$\beta_1(K)$	0	0	0	0	0	0	0	1	2	3	2	1	0	0
$\beta_2(K)$	0	0	0	0	0	0	0	0	0	0	0	0	0	1

□

7 Topological inference

7.1 Important definitions

Definition 7.1.1. *For every $t \geq 0$, the t-thickening of the set X, denoted X^t, is the set of points of the ambient space with distance at most t from X:*

$$X^t = \{y \in \mathbb{R}^n, \exists x \in X, ||x - y|| \leq t\}.$$

Equivalently, X^t can be seen as a union of closed balls centered around every point of X.

Definition 7.1.2 (Hausdorff distance). *Let X be any subset of $\mathbb{R}^n$. The function distance to X is the map*

$$\begin{aligned} dist\ (\cdot, X) : \mathbb{R}^n &\to \mathbb{R} \\ y &\mapsto dist\ (y, X) = \inf\{||y - x||, x \in X\} \end{aligned}$$

A projection of $y \in \mathbb{R}^n$ on X is a point $x \in X$ which attains this infimum. If such a point x exists and is unique, we denote it proj (y, X).

Then the Hausdorff distance cna be written as

$$d_H(X, Y) = \max(\sup_{y \in Y} dist\ (y, X), \sup_{x \in X} dist\ (x, Y))$$

Definition 7.1.3. *Let X be any subset of $\mathbb{R}^n$. The medial axis of X is the subset $med(X) \subset \mathbb{R}^n$ which consists of points $y \in \mathbb{R}^n$ that admit at least two projections on X:*

$$med(X) = \{y \in \mathbb{R}^n, \exists x, x' \in X, x \neq x', ||y - x|| = ||y - x'|| = dist(y, X)\}.$$

Definition 7.1.4. *Now, we define the reach of X as its proximity from its medial axis:*

$$reach(X) = \inf\{dist(y, X), y \in med(X)\} = \inf\{||x - y||, x \in X, y \in med(X)\}$$

Equivalently

$$reach(X) = \sup\{t \geq 0, X^t \cap med(X) = \emptyset\}$$

Definition 7.1.5. *Let X be a topological space, and $U = \{U_i\}_{1 \leq i \leq N}$ a cover of X. The nerve of U is the simplicial complex with vertex set $\{1, ..., N\}$ and whose m-simplices are the subsets $\{i_1, ..., i_m\} \subset \{1, ..., N\}$ such that $\cap_{k=0}^{m} U_{i_k} \neq \emptyset$. It is denoted $\mathcal{N}(U)$.*

Definition 7.1.6. *Let $t \geq 0$ and consider the collection $\mathcal{V}^t = \{\bar{\mathcal{B}}(x, t), x \in X\}$. Its nerve is denoted $Cech^t(X)$ and is called the Cech complex of X at time t.*

Definition 7.1.7. *Given a graph G, the corresponding clique complex is the simplicial complex whose vertices are the vertices of G, and whose simplices are the sets of vertices of the cliques of G. Some authors also call it the expansion of G .*

Definition 7.1.8. *The Rips complex of X at time t is the clique complex of the graph G^t defined above. We denote it $Rips^t(X)$.*

7.2 Exercises

EXERCISE 35. *Prove that, when X is closed and bounded, a projection always exists. A set is bounded if there exists $R > 0$ such that $X \subset \mathcal{B}(R, 0)$.*

Proof. Let X a closed and bounded subset of $\mathbb{R}^n$. So it's compact by the Heine-Borel theorem. The map $y \mapsto ||y - x||$ is continuos because for $y \in \mathbb{R}^n$ and $\epsilon > 0$, if we take $\delta = \epsilon$, and $||y - w|| < \delta$,

$$\epsilon > ||y - w|| = ||(y - x) - (w - x)|| \geq |||y - x|| - ||w - x|||$$

By the Weierstrass theorem, the function $y \mapsto ||y - x||$ has a global minimum in X, that is, there exist $x^* \in X$ such that $||y - x^*|| \leq ||y - x||, \forall x \in X$. So the infimum is well defined and a projection always exist. However the uniqueness is not guaranteed.

□

EXERCISE 36. *Let $||\cdot||+\infty$ be the sup norm of function $f : \mathbb{R}^n \to \mathbb{R}^m : ||f||_\infty = \sup_{x\in\mathbb{R}^n} ||f(x)||$. Prove that $d_H(X, Y) = ||dist(\cdot, X) - dist(\cdot, Y)||_\infty$.*

Proof. We have that

$$\begin{aligned} ||\text{dist } (\cdot, X) - \text{dist } (\cdot, Y)||_\infty &= \sup_{x\in\mathbb{R}^n} |\text{dist}(x, X) - \text{dist}(x, Y)| \\ &\geq \sup_{x\in\mathbb{R}^n} (\text{dist}(x, X) - \text{dist}(x, Y)) \\ &\geq sup_{y\in Y} dist(y, X) \end{aligned}$$

given that $Y \subset \mathbb{R}^n$ and $dist(y, Y) = 0$. Also

$$\begin{aligned} ||\text{dist } (\cdot, X) - \text{dist } (\cdot, Y)||_\infty &= \sup_{x\in\mathbb{R}^n} |\text{dist}(x, X) - \text{dist}(x, Y)| \\ &\geq \sup_{x\in\mathbb{R}^n} (\text{dist}(x, Y) - \text{dist}(x, X)) \\ &\geq sup_{x\in X} dist(x, Y) \end{aligned}$$

In special, $||\text{dist } (\cdot, X) - \text{dist } (\cdot, Y)||_\infty \geq d_H(X, Y)$.

The other inequality can be seen as follows. If $dist(x, Y) \leq k, \forall x \in X$, then for all $w \in \mathbb{R}^n$ and $x \in X$,

$$dist(w, Y) \leq dist(w, x) + dist(x, Y) \leq dist(w, x) + k$$

and hence $dist(w, Y) \leq dist(w, X) + k$, taking the infimum. Likewise, if $dist(y, X) \leq k, \forall y \in Y$ we obtain $dist(w, X) \leq dist(w, Y) + k$. Then $|dist(w, X) - dist(w, Y)| \leq k$, for all $w \in \mathbb{R}^n$. In particular $d_H(X, Y) \geq dist(x, Y), \forall x \in X$ and $d_H(X, Y) \geq dist(y, X), \forall y \in Y$. By the fact we have proved

$$d_H(X, Y) \geq \sup_{w\in R^n} |dist(w, X) - dist(w, Y)| = ||dist(\cdot, X) - dist(\cdot, Y)||_\infty$$

We conclude $d_H(X, Y) = ||dist(\cdot, X) - dist(\cdot, Y)||_\infty$.

□

EXERCISE 37. *Let X, Y be two closed and bounded subsets of $\mathbb{R}^n$. Show that for every $t \geq 0$, the thickenings satisfy*

$$d_H(X^t, Y^t) \leq d_H(X, Y).$$

Give an example for which $d_H(X^t, Y^t) < d_H(X, Y)$.

Proof. Let's do it by steps, because dealing with infimum and supremum can be trick. Let $t \geq 0$ and denote *dist* by d.

1. Let's prove $d(w, Y^t) \leq d(w, Y) - t, \forall w \in X^t / Y^t$.

   Take $y \in Y$ and denote $\beta = ||w - y|| > t$. Take the point $\alpha w + (1-\alpha)y$ such that, for some $\alpha \in [0,1]$, $t = ||\alpha y + (1-\alpha)y - y|| = \alpha\beta$, that is, $\alpha = t/\beta < 1$. So we will have $||w - \alpha w - (1-\alpha)y|| = (1-\alpha)\beta = \beta - t$. It implies $d(w, Y^t) \leq d(w, Y) - t$. After we will see the restriction over the domain of w is not restrictive.

2. $\sup_{w \in X^t} d(w, Y) \leq \sup_{x \in X} d(x, Y) + t$

   If $w \in X^t$, there exists $x \in X$ such that $||w - x|| \leq t$. So $d(w, Y) \leq d(x, Y) + t \leq \sup_{x \in X} d(x, Y) + t$. It values for all w, then $\sup_{w \in X^t} d(w, Y) \leq \sup_{x \in X} d(x, Y) + t$.

3. $\sup_{w \in X^t} d(w, Y^t) \leq \sup_{x \in X} d(x, Y)$.

   For every $w \in X^t / Y^t$, we have $d(w, Y^t) \leq \sup_{w \in X^t} d(w, Y) - t$, because of what we've proved on point 1. Then

   $$\sup_{w \in X^t / Y^t} d(w, Y^t) \leq \sup_{w \in X^t} d(w, Y) - t \leq \sup_{x \in X} d(x, Y),$$

   as argued on point 2. Otherwise suppose $w \in Y^t \cap X^t$, then $d(w, Y^t) = 0$ and the inequality $\sup_{w \in X^t \cap Y^t} d(w, Y^t) = 0 \leq \sup_{x \in X} d(x, Y)$ is trivial, because d is a metric. Therefore

   $$\sup_{w \in X^t} d(w, Y^t) = \max\{\sup_{w \in X^t \cap Y^t} d(w, Y^t), \sup_{w \in X^t / Y^t} d(w, Y^t)\} \leq \sup_{x \in X} d(x, Y)$$

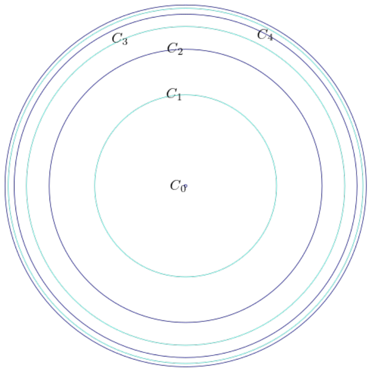
4. We can show similarly that $\sup_{z \in Y^t} d(z, X^t) \leq \sup_{y \in Y} d(y, X)$.

5. Therefore, the last two points imply $d_H(X^t, Y^t) \leq d_H(X, Y)$.

Example

Consider the following construction. Let C_0 be the origen and C_i be the circumference of center 0 and radius $\sum_{j=0}^{i-1} 2^{-j} = 2 - 2^{1-i}$. Each time the adicional value to the radius of the next circle is decreasing by half. All the circles are contained in $\mathcal{B}(0, 2)$. We will define

$$X = \bigcup_{i=0}^{\infty} C_{2i} \text{ and } Y = \bigcup_{i=0}^{\infty} C_{2i+1}$$



I claim that $d_H(X,Y) = 1$. For every $x \in C_i$, we have that $d(x,Y) = 2^{-i}$, because the closest point in Y is in C^{i+1}. We can argue the same for $d(y,X)$. Therefore $d_H(X,Y) = \sup_i 2^{-i} = 1$. Consider now X^t and Y^t for some $t > 0$. We are creating rings with thickness $2t$. Some rings of different sets are being joined. Let's calculate $d_H(X^t, Y^t)$. In general, $d(x, Y^t)$ will decreased, be unaltered (when we pick some point in a border of C_i^t that maps to C_{i+1}^t or go to 0, when two rings join. But it shall not increase. However if we take the origin, $d(0,Y) = 1 - t$, but no other can have a greater distance. Thus, $d_H(X^t, Y^t) = 1 - t < 1 = d_H(X,Y)$.

□

Exercise 38. *Show that the Hausdorff distance is equal to*

$$inf\{t \geq 0, X \subset Y^t \text{ and } Y \subset X^t\}$$

Proof. Let t such that $X \subset Y^t$ and $Y \subset X^t$. Take $x \in X$ and consider $d(x,Y)$. By the choice of $t, x \in Y^t$, what implies $d(x,Y) \leq t$, because we can find $y \in Y$ such that $||x - y|| \leq t$. As x is arbitrary, $\sup_{x \in X} d(x,Y) \leq t$. The same can be told about $\sup_{y \in Y} d(y,X)$. That implies $d_H(X,Y) \leq t$. As we took arbitrary t, applying the infimum we obtain

$$d_H(X,Y) \leq \inf\{t \geq 0, X \subset Y^t, Y \subset X^t\}.$$

Suppose the above inequality is strict ($<$). Then there is $\epsilon > 0$ such that $d_H(X,Y) + \epsilon$ is still less. Define $s = d_H(X,Y) + \epsilon$. Then $s > d(x,Y), \forall x \in X$ and $s > d(y,X), \forall y \in Y$. As $s > d(x,Y)$, there is $y \in Y$ such that $||x - y|| < s \implies x \in \bar{\mathcal{B}}(x,y) \implies x \in Y^s$. This holds for every $x \in X$, so $X \subset Y^s$. Likewise we prove $Y \subset X^s$. That is a contradiction, since $s < \inf\{t \geq 0, X \subset Y^t, Y \subset X^t\}$. We conclude that $d_H(X,Y) = \inf\{t \geq 0, X \subset Y^t, Y \subset X^t\}$.

□

Exercise 39. *Compute the reach of the following subsets of* R^2*:*

1. *the set* $\{(0,n), n \in \mathbb{Z}\}$,

2. *the segment* $\{(t,0), t \in [0,1]\}$,

3. *the unit circle with origin* $\mathbb{S}_1 \cup \{(0,0)\}$,

4. *the square* $\{(x,y) \in R^2, \max\{|x|,|y|\} = 1\}$,

5. *the ellipse* $\{(x_1, x_2) \in \mathbb{R}^2, (\frac{x_1}{a})^2 + (\frac{x_2}{b})^2 = 1\}$

Proof. 1. Define this set as Z. So $reach(Z) = \inf\{d(y, Z), y \in med(Z)\}$. We have that $med(Z) = \{(x, n + 1/2), x \in \mathbb{R}, n \in \mathbb{Z}\}$, the bisector os two sequential points. For each point y in one of these lines, $dist(y, Z) = ||(x, n+1/2) - (0, n)|| = ||(x, 1/2)|| = \sqrt{x^2 + 1/4}$. The infimum over y occurs when $x = 0$ and we conclude $reach(Z) = 1/2$.

2. Denote this set as I. Take $y = (y_1, y_2) \in \mathbb{R}^2$. We want the values on I which distance to y is $d(y, I) = \inf\{\sqrt{(y_1 - t)^2 + y_2^2}, t \in [0, 1]\} = \min\{\sqrt{(y_1 - t)^2 + y_2^2}, t \in [0, 1]\}$, since I is closed and bounded, there exists a projection. We must minimize $(y_1 - t)^2$ in the interval $[0, 1]$, but this clearly have only one solution. We conclude $med(I) = \emptyset$ and $reach(I) = \infty$.

3. We have that $med(\mathbb{S}_1 \cap 0) = \partial\mathcal{B}(0, 1/2)$. For each $y \in \partial\mathcal{B}(0, 1/2), d(y, Z) = 1/2$, the distance to the origin. Therefore $reach(\mathbb{S} \cap 0) = 1/2$.

4. The medial axis of the square is an "X" crossing the diagonals. Because each point of the diagonal forms a smaller square which minimizes distance the adjacency borders (we have two). Then $reach(square) = 0$.

5. Suppose $a > b$ for simplification, but in the other case, we only need to change the axis, what does not affect the reach. Let $(w, z) \in \mathbb{R}^2$ in the exterior of the ellipse. Suppose we have two different projections of this point in the ellipse and consider the circle centered in (w, z) with distance being this minimum. Both figures are convex, so all points in the segment between the projections are in the interior of the circle and the ellipse, however this is an absurd, because the arc in the ellipse between the projection points are, therefore, inside the circle, meaning, their distance too (w, z) is smaller. So there cannot be two projections. The existence of a projection can be given by the ellipse being closed and bounded. I can say the same argument for points above and below the $x - axis$. I infer, then, that an open interval in the $x - axis$ is the median axis. Let's find this interval. Let $(t, 0) \in \mathbb{R}^2$.

Consider the problem

$$\min_{x,y} (x - t)^2 + y^2$$
$$s.t. \ \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

I can translate this problem into

$$\min_{x \in [-a,a]} f(x) = (x - t)^2 + b^2 - \frac{b^2}{a^2}x^2$$

Using calculus,

$$f'(x) = 2(x - t) - 2\frac{b^2}{a^2}x = 0 \implies x = \frac{t}{1 - b^2/a^2}$$

Then our answer is $x \in \{-a, a, \frac{t}{1-b^2/a^2}\}$ that minimizes f. If $x \neq \pm a$, we have two values of y which address the minimum and, therefore, there are two projections. Denote E to refer the ellipse. We then have that $med(E) = (-s, s)$ such that, if $x > s$, $f(x)$ is minimized at a and $if x < -s$, at $-a$. In order to find s, we must calculate $f(\frac{t}{1-b^2/a^2})$ and compare with

$(a-t)^2$. Then

$$\left(\frac{a^2 t}{a^2-b^2}-t\right)^2+b^2-b^2a^2\frac{t^2}{(a^2-b^2)^2}=t^2b^2\left(b^2-\frac{a^2}{(a^2-b^2)^2}\right)+b^2>(a-t)^2$$

Solving this inequality we find $s=\sqrt{a^2-b^2}$ and, then, the reach will be the distance $a-s$.

□

EXERCISE 40. *Compute the reaches of the subsets of Exercise 39.*

EXERCISE 41. *Verify that the clique complex of a graph is a simplicial complex. If the graph contains n vertices, give an upper bound on the number of simplices of the clique complex.*

Proof. Denote the clique complex K. Let $\sigma \in K$ and $\tau \subset \sigma$. Then σ is a clique of the graph. Since τ is subset, every node it contais connects to each other, because they connect in σ. So τ is a clique, what implies $\tau \in K$.

In the worst case, all the vertices connects to each other, so we'll have

$$\sum_{i=1}^{n}\binom{n}{i}=2^n-1$$

simplices.

□

EXERCISE 42. *Improve the previous proposition as follows: $Cech^t(X) \subset Rips^t(X) \subset Cech^{ct}(X)$, where $c=\sqrt{\frac{2n}{n+1}}$.*

8 Datasets have topology

8.1 Important definitions

Definition 8.1.1. *Let $X \subset \mathbb{R}^n$ and $i \geq 0$. The i^{th} Betti curve of X is the map*

$$\begin{aligned} \beta_i(t) :& \mathbb{R}^+ \to \mathbb{N} \\ & t \mapsto \beta_i(X^t) \end{aligned}$$

As a consequence of the nerve theorem, the map $t \to \beta_i(t)$ is equal to $t \mapsto \beta_i(Cech^t(X))$. In practice, we may use the following map, called the i^{th} Betti curve of the Rips complex of X

$$\begin{aligned} \beta_i^{Rips}(t) :& \mathbb{R}^+ \to \mathbb{N} \\ & t \mapsto \beta_i(Rips^t(X)) \end{aligned}$$

8.2 Exercises

Exercise 43. *Show that $t \mapsto \beta_0(t)$ is non-increasing. Show that $t \mapsto \beta_0^{Rips}(t)$ is also non-increasing.*

Proof. In order to prove the first statement, I shall find for each t the value $\beta_i(X^t)$, that is, we need to find $\beta_i(K)$, where K is a triangulation of X^t. Actually, I'll the weak version, where $|K|$ must be homotopy equivalent to X^t. In particular, a consider the $Cech^t(X)$. Take $s < t$ and consider $\mathcal{N}(\mathcal{V}^s), \mathcal{N}(\mathcal{V}^t)$ the nerves respectively. If $\{x_1, ..., x_m\}$ is m-simplex of $\mathcal{N}(\mathcal{V}^s)$, we have that $\bigcap_{k=1}^m \bar{\mathcal{B}}(x_k, s) \neq \emptyset$. For each $k \in [[1, m]], \bar{\mathcal{B}}(x_k, s) \subset \bar{\mathcal{B}}(x_k, t)$ and for that reason $\bigcap_{k=1}^m \bar{\mathcal{B}}(x_k, t) \neq \emptyset \implies \{x_1, ..., x_m\} \in \mathcal{N}(\mathcal{V}^t)$. Therefore as t increases, no simplices are removed and, maybe, more can be added.

Remember that $\beta_0(K) = dim Z_0(K) - dim B_0(K)$. We already know that $Z_0(K) = C_0(K)$, so its dimension do not change. Otherwise, observe that $C_1(\mathcal{N}(\mathcal{V}^s)) \subset C_1(\mathcal{N}(\mathcal{V}^t))$, as we proved above. In that sense, no element can be removed from $B_0(\mathcal{N}(\mathcal{V}^s))$ and hence its dimension cannot decrease (and it can increase if we add new elements to it, when new simplices in $C_1(\mathcal{N}(\mathcal{V}^t))$ are not mapped in the existing images). We conclude $t \mapsto \beta_0(\mathcal{N}(\mathcal{V}^t))$ is non-increasing, then $t \mapsto \beta_0(X^t)$ is non-increasing.

We can use the same argument to prove $t \mapsto \beta_0^{Rips}(t)$ is non-increasing. When t increases, we only can add edges, but not remove, as the considered difference increases (note the condition to add edges is x_i, x_j such that $||x_i - x_j|| \leq 2t$).

□

The computational exercises (44 - 46) can be found in the Github[3]

Exercise 44. *In the notebook is given a subset of $\mathbb{R}^4$ with 200 elements. It has been sampled on a famous 2-dimensional object. Compute the Betti curves of its Rips complex on $[0, 1]$. Can you recognize which surface it is?*

---

[3] github.com/lucasmoschen/topological-data-analysis/blob/main/tutorials/tutorial-2.ipynb

Exercise 45. *In the notebook is given a collection of images from* https://www.cs.columbia.edu/CAVE/software/softlib/coil-20.php. *It consists of 20 objects, for each of which 72 pictures have been taken. Each image has 128×128 pixels. Embed each collection of 72 images in R 128×128 , and compute the Betti curves of the corresponding Rips complex.*

Exercise 46. *We are given the data of this paper*[4]. *It consists in 14 correlation matrices, each matrix representing correlations between the components of a protein. Transform the matrices of correlations into matrices of distances. Then, compute the 1-Betti curves of the Rips complex for each of these matrices of distances. Compare the Betti curves of the different proteins. Do you recognize two different types of proteins (open and closed)?*

[4] https://pubmed.ncbi.nlm.nih.gov/26812805/

9 Decomposition of persistence modules

9.1 Important definitions

Definition 9.1.1. *Let K and L be two simplicial complexes, and V_K, V_L their set of vertices. A simplicial map between K and L is a map $f : V_K \to V_L$ such that $\forall \sigma \in K, f(\sigma) \in L$. When there is no risk of confusion, we may denote a simplicial map $f : K \to L$ instead of $f : V_K \to V_L$.*

We define a linear map $f_n : C_n(K) \to C_n(L)$ as follows (definition on the simplices, and extended by linearity):

$$f_n : \sigma \mapsto f(\sigma)\mathbb{1}\{dim(f(\sigma)) = n\}$$

Definition 9.1.2. *By definition of the homology groups, we have defined a map*

$$(f_n)_* : H_n(K) \to H_n(L).$$

It is called the induced map in homology. Explicitly, the map $(f_n)_*$ can be described as follows (the following formula is to be read modulo boundaries):

$$c = \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \mapsto \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot f_n(\sigma).$$

Definition 9.1.3. *Let $i \geq 0$, $t_0 \geq 0$ and consider a cycle $c \in H_i(Cech^{t_0}(X))$. Its death time is*

$$\sup\{t \geq t_0, (i_{t_0}^t)(c) \neq 0\}$$

and its birth time is

$$\inf\{t \geq t0, (i_t^{t_0})^{-1}(\{c\}) \neq \emptyset\}$$

$$\textit{persistence} = \textit{death time - birth time}$$

Definition 9.1.4. *A persistence module $\mathbb{V}$ over $\mathbb{R}^+$ with coefficients in $\mathbb{Z}/2\mathbb{Z}$ is a pair $(\mathbb{V}, \mathbb{v})$ where $\mathbb{V} = (V^t)_{t\in\mathbb{R}^+}$ is a family of $\mathbb{Z}/2\mathbb{Z}$-vector spaces, and $\mathbb{v} = (v_s^t : V^s \to V^t)_{s\leq t\in\mathbb{R}^+}$ a family of linear maps such that:*

1. *for every $t \in \mathbb{R}^+, v_t^t$ is the identity map,*
2. *for every $r, s, t \in \mathbb{R}^+$ such that $r \leq s \leq t$, we have $v_s^t \circ v_r^s = v_r^t$*

When the context is clear, we may denote $\mathbb{V}$ instead of $(\mathbb{V}, \mathbb{v})$.

Definition 9.1.5. *An isomorphism between two persistence modules $\mathbb{V}$ and $\mathbb{W}$ is a family of isomorphisms of vector spaces $\phi = (\phi_t : V^t \to W^t)_{t\in\mathbb{R}^+}$ such that the following diagram commutes for every $s \leq t \in \mathbb{R}^+$.*

$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \phi_s \downarrow & & \downarrow \phi_t \\ W^s & \xrightarrow[w_s^t]{} & W^t \end{array}$$

Definition 9.1.6. *Let $(\mathbb{V}, \mathbb{v})$ and $(\mathbb{W}, \mathbb{w})$ be two persistence modules. Their sum is the persistence module $V \oplus W$ defined with the vector spaces $(V \oplus W)^t = V^t \oplus W^t$ and the linear maps*

$$(v \oplus w)_s^t : (x, y) \in (V \oplus W)^s \mapsto (v_s^t(x), w_s^t(y)) \in (V \oplus W)^t.$$

A persistence module $\mathbb{U}$ is indecomposable if for every pair of persistence modules $\mathbb{V}$ and $\mathbb{W}$ such that $\mathbb{U}$ is isomorphic to the sum $\mathbb{V} \oplus \mathbb{W}$, then one of the summand has to be a trivial persistence module, that is, equal to zero for every $t \in \mathbb{R}^+$. Otherwise, $\mathbb{U}$ is said decomposable.

Definition 9.1.7. *Let $I \subset \mathbb{R}^+$ be an interval. Intervals have the form $[a, b]$, $(a, b]$, $[a, b)$ or (a, b), with $a, b \in \mathbb{R}^+$ such that $a \leq b$, and potentially $b = +\infty$. The interval module associated to I is the persistence module $\mathbb{B}[I]$ with vector spaces $\mathbb{B}^t[I]$ and linear maps $v_s^t : \mathbb{B}^s[I] \to \mathbb{B}^t[I]$ defined as*

$$\mathbb{B}^t[I] = \begin{cases} \mathbb{Z}/2\mathbb{Z} & \text{if } t \in I, \\ 0 & \text{otherwise} \end{cases} \quad \text{and } v_s^t = \begin{cases} id & \text{if } s, t \in I \\ 0 & \text{otherwise} \end{cases}$$

Definition 9.1.8. *A persistence module $\mathbb{V}$ decomposes into interval module if there exists a set $\{\mathbb{B}_i, i \in I\}$ of interval modules such that $\mathbb{V}$ is isomorphic to the sum $\bigoplus_{i \in I} \mathbb{B}_i$;. In other words, there exists a multiset $\mathcal{I}$ of intervals of T such that*

$$\mathbb{V} = \bigoplus_{I \in \mathcal{I}} \mathbb{B}[I].$$

Multiset means that $\mathcal{I}$ may contain several copies of the same interval I. Such a module is said decomposable into interval modules, or simply decomposable when the context is clear.

9.2 Exercises

Exercise 47. *Consider a linear map $f : V \to W$ between vector spaces. Suppose that there exists linear subspaces $A \subset V$ and $B \subset W$ such that $f(A) \subset B$. Show that one can define a map $f_* : V/A \to W/B$ as follows: to any equivalence class $v + A$ of V/A, let $f_*(v + A) = f(v) + B$.*

Proof. We have to prove f_* is well-defined. In other words, if $v_1 + A = v_2 + A$, that is, $v_1 - v_2 \in A$, then $f(v_1) + B = f(v_2) + B$. In that sense, we must prove that $f(v_1) - f(v_2) = f(v_1 - v_2) \in B$. But this is true since by hypotheses $v_1 - v_2 \in A \implies f(v_1 - v_2) \in f(A) \subset B$. Then f_* is indeed well defined.

Let's prove it's linear. Let $\alpha \in \mathbb{F}$, and $v_1, v_2 \in V$. So

$$\begin{aligned} f_*((\alpha v_1 + v_2) + A) &= f(\alpha v_1 + v_2) + B \\ &= \alpha f(v_1) + f(v_2) + B \\ &= \alpha(f(v_1) + B) + (f(v_2) + B) \\ &= \alpha f_*(v_1 + A) + f_*(v_2 + A) \end{aligned}$$

The third equation is true because $B = \alpha B + B$, that is, if $b \in B$, then $b = \alpha 0 + b \in \alpha B + B$ and in $b \in \alpha B + B$, one can find $b_1, b_2 \in B$ such that $b = \alpha b_1 + b_2$. Because of the linearity of B, we have $b \in B$. We conclude f_* is a linear map induced by f between quotient linear spaces. □

Exercise 48. *Let the simplicial* $K = \{[0],[1],[2],[3],[4],[5],[0,1],[1,2],[2,3],[3,4],[4,5],[5,0]\}$ *and* $L = \{[0],[1],[2],[0,1],[1,2],[2,0]\}$. *Consider the simplicial map* $f : i \mapsto i$ *modulo 3. Show that the induced map* $(f_1)_*$ *is zero.*

Proof. We have that $(f_1)_* : H_1(K) \to H_1(L)$, such that

$$H_1(K) = \{[0,1]+[1,2]+[2,3]+[3,4]+[4,5]+[5,0], 0\}.$$

This happens because $Z_1(K)$ has only two elements not equivalents. The same can be told about $H_1(L)$. Then

$$\begin{aligned}(f_1)_*([0,1]+...+[5,0]) &= f_1([0,1])+...+f_1([5,0])\\ &= f([0,1])+...+f([5,0])\\ &= [0,1]+[1,2]+[2,0]+[0,1]+[1,2]+[2,0] = 0\end{aligned}$$

And its trivial that $(f_1)_*(0) = 0$ (linear map), what implies the induced map $(f_1)_*$ is zero. □

Exercise 49. *Fill the empty spaces ( ) in the following proof of Brouwer's fixed point theorem.*

Proof. Let $f : \bar{\mathcal{B}}(0,1) \to \bar{\mathcal{B}}(0,1)$ be a continuous map, where $\bar{\mathcal{B}}(0,1)$ denotes the closed unit ball of $\mathbb{R}^n$. Let us show that f admits a fixed point (i.e., an element $x \in \bar{\mathcal{B}}(0,1)$ such that $f(x) = x$). By contradiction, suppose that it is not the case. We can build an application $F : \bar{\mathcal{B}}(0,1) \to \mathbb{S}(0,1) \subset \bar{\mathcal{B}}(0,1)$, such that F restricted to $\mathbb{S}(0,1)$ is the identity. To do so, define $F(x)$ as the first intersection point between the half-line $[x, f(x))$ and $\mathbb{S}(0,1)$.

Denote the inclusion $i : \mathbb{S}(0,1) \to \mathcal{B}(0,1)$. We have that $F \circ i : \mathbb{S}(0,1) \to \mathbb{S}(0,1)$ is the identity. By functoriality of homology, we obtain, for all $i \geq 0$, the commutative diagrams.

$$\mathbb{S}(0,1) \xrightarrow{i} \bar{\mathcal{B}}(0,1) \xrightarrow{F} \mathbb{S}(0,1), \quad F \circ i = id \qquad\qquad H_i(\mathbb{S}(0,1)) \xrightarrow{i_*} H_i(\bar{\mathcal{B}})(0,1) \xrightarrow{F_*} H_i(\mathbb{S}(0,1)), \quad F_* \circ i_* = (id_*)$$

But choosing $i = n-1$, we have $H_i(\mathbb{S}(0,1)) \simeq (\mathbb{Z}/2\mathbb{Z}), H_i(\bar{\mathcal{B}}(0,1)) \simeq (0)$ and the following diagram cannot commute:

$$(\mathbb{Z}/2\mathbb{Z}) \longrightarrow 0 \longrightarrow (\mathbb{Z}/2\mathbb{Z}), \quad id$$

□

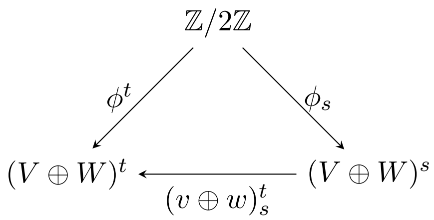
Exercise 50. *Show that the interval modules are indecomposable.*

Proof. WLet $\mathbb{V}$ and $\mathbb{W}$ two persistence modules whose sum is isomorphic to $\mathbb{B}[I]$. Then there is $\phi = (\phi^t)_{t\in\mathbb{R}^+}$ such that $\mathbb{B}^t[I] \overset{\phi^t}{\to} (V \oplus W)^t$ are isomorphic. If $t \notin I$, We have that $\mathbb{B}^t[I] = 0$ is isomorphic to $(V \oplus W)^t$, what implies $(V \oplus W)^t = (0,0)$ and we are done. Otherwise, if $t \in I$, we have $(V \oplus W)^t$ and $\mathbb{Z}/2\mathbb{Z}$ are isomorphic. It's clear that $dim(\mathbb{Z}/2\mathbb{Z}) = 1$, because the bases only need one element. We claim that $dim((V \oplus W)^t) = 1$. We already know $\{\phi^t(1+\mathbb{Z})\}$ is a set of independent sets of $(V \oplus W)^t$ and we shall prove it spans the space. Take $(v, w) \in (V \oplus W)^t$. Then, for some $e \in \mathbb{Z}/2\mathbb{Z}$,

$$(v, w) = \phi^t(e) = \phi^t((1 + \mathbb{Z})\alpha) = \alpha\phi^t(1 + \mathbb{Z}).$$

Therefore $1 = dim((V + W)^t) = dim(V^t) + dim(W^t) \implies dim(V^t) = 0$ xor $dim(W^t) = 0$, that is, $V^t = 0$ xor $W^t = 0$. In special we have for all t, we have $V^t = 0$ or $W^t = 0$. Suppose $V^t \neq 0$ and $W^s \neq 0$, for some $t, s \in \mathbb{R}^+$. We already proved that if $t \notin I$, then $V^t = 0$ and if $s \notin I$, then $W^s = 0$, so we don't need to consider these cases.

We know $W^t = 0$ and $V^s = 0$ by what we've proved in the first paragraph. By the commutative property and supposing $s \leq t$ with no loss of generality,



Therefore,

$$\phi^t(1 + \mathbb{Z}) = (v, 0) \text{ and } \phi^s(1 + \mathbb{Z}) = (0, w), \text{ then } v_s^t(0) = v, w_s^t(w) = 0$$

Alternatively

$$\phi^t(0 + \mathbb{Z}) = (0, 0) \text{ and } \phi^s(0 + \mathbb{Z}) = (0, 0), \text{ then } v_s^t(0) = 0$$

This is clearly a contradiction. □

EXERCISE 51. *Let $\mathcal{M}$ be the unit circle of $\mathbb{R}^2$, and $X \subset \mathbb{R}^2$ a finite subset. Denote the Hausdorff distance $\epsilon = d_H(X, \mathcal{M})$. Suppose that ϵ is small enough. Let $\mathbb{U}$ denote the persistence module of the 1st homology of the Cech filtration of X. Using Theorem 7.3, shows that there exists an interval I on which $\mathbb{U}$ is constant and equal to $\mathbb{Z}/2\mathbb{Z}$. Deduce the existence of a bar in the barcode, and give a lower bound on its persistence. Do the same with Theorem 7.4.*

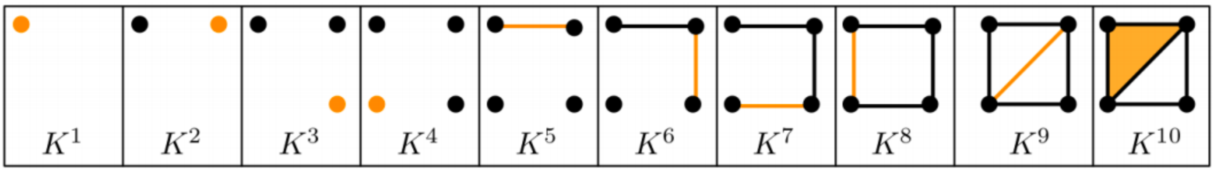
Proof. Let's denote $\mathbb{U} = (H_1(Cech^t(X)))_{t \in \mathbb{R}^+}$. By theorem 7.3, if we take

$$t \in [4d_H(X, \mathcal{M}), reach(\mathcal{M}) - 3d_H(X, \mathcal{M}))$$

we have $X^t \approx \mathcal{M}$. Then, by Proposition 5.14, we have isomorphic homology groups $H_1(X)$ and $H_1(\mathcal{M})$. We know $H_1(circle) = \mathbb{Z}/2\mathbb{Z}$. Then we take $I = [4d_H(X, \mathcal{M}), reach(\mathcal{M}) - 3d_H(X, \mathcal{M}))$.

To use Theorem 7.4, we need the additional supposition that $X \subset \mathcal{M}$ is finite. If we take $I = [2d_H(X, \mathcal{M}), \sqrt{\frac{3}{5}} reach(\mathcal{M}))$, we have the result. □

Exercise 52. *Compute the barcode of the filtration of Subsection 6.1:*



with the following filtration values: $t(\sigma) = 0$ *for the vertices,* $t(\sigma) = \frac{1}{2}$ *for the edges of the square, and* $t(\sigma) = \frac{\sqrt{2}}{2}$ *for the diagonal edge and the triangle.*

First let's build the boundary matrix.

-	$\sigma_{1,2,3,4}$	σ_5	σ_6	σ_7	σ_8	σ_9	σ_{10}
σ_1	0	1	0	0	1	0	0
σ_2	0	1	1	0	0	1	0
σ_3	0	0	1	1	0	0	0
σ_4	0	0	0	1	1	1	0
σ_5	0	0	0	0	0	0	1
σ_6	0	0	0	0	0	0	0
σ_7	0	0	0	0	0	0	0
σ_8	0	0	0	0	0	0	1
σ_9	0	0	0	0	0	0	1
σ_{10}	0	0	0	0	0	0	0

as already placed in the notes. After apply Gauss reduction we obtain

-	$\sigma_{1,2,3,4}$	σ_5	σ_6	σ_7	σ_8	σ_9	σ_{10}
σ_1	0	1	0	0	0	0	0
σ_2	0	1	1	0	0	0	0
σ_3	0	0	1	1	0	0	0
σ_4	0	0	0	1	0	0	0
σ_5	0	0	0	0	0	0	1
σ_6	0	0	0	0	0	0	0
σ_7	0	0	0	0	0	0	0
σ_8	0	0	0	0	0	0	1
σ_9	0	0	0	0	0	0	1
σ_{10}	0	0	0	0	0	0	0

Consider the pairs $(\sigma^{\delta(j)}, \sigma^j)$. We do not have $\delta(j) = 1$, then $(+\infty, \sigma^1)$ is pair. Since $\delta(5) = 2$, (σ^2, σ^5) is another pair. Continuing this reasoning we obtain the pairs

$$(\sigma^3, \sigma^6), (\sigma^4, \sigma^7), (\sigma^8, +\infty), (\sigma^9, \sigma^{10}).$$

We conclude the barcodes will be

$$\mathcal{I}_1 = \{(0.5, +\infty)\}, \mathcal{I}_0 = \{(0, 1/2), (0, 1/2), (0, 1/2), (0, +\infty)\}$$

10 Stability of persistence modules

10.1 Important definitions

Definition 10.1.1. *Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $\bar{\mathbb{R}^+}^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$. Here, $\bar{\mathbb{R}^+}$ represent the extended real line $\mathbb{R}^+ \cup \{+\infty\}$. A partial matching between the barcodes is a subset $M \subset P \times Q$ such that*

- *for every $p \in P$, there exists at most one $q \in Q$ such that $(p, q) \in M$,*
- *for every $q \in Q$, there exists at most one $p \in P$ such that $(p, q) \in M$.*

The points $p \in P$ (resp. $q \in Q$) such that there exists $q \in Q$ (resp. $p \in P$) with $(p, q) \in M$ are said matched by M. If a point $p \in P$ (resp. $q \in Q$) is not matched by M, we consider that it is matched with the singleton $\bar{p} = [\frac{p_1+p_2}{2}, \frac{p_1+p_2}{2}]$ (resp. $\bar{q} = [\frac{q_1+q_2}{2}, \frac{q_1+q_2}{2}]$). The cost of a matched pair (p, q) (resp. $(p, \bar{p})$, resp. $(q, \bar{q})$) is the sup norm $||p - q||_\infty = \sup\{|p_1 - q_1|, |p_2 - q_2|\}$. The cost of the partial matching M, denoted $cost(M)$, is the supremum of all such costs.

Definition 10.1.2. *The bottleck distance between two barcodes P and Q is defined as the infimum of costs over all the partial matchings:*

$$d_b(P, Q) = \inf\{cost(M), M \text{ is a partial matching between } P \text{ and } Q\}.$$

If U and V are two decomposable persistence modules, we define their bottleneck distance as

$$d_b(U, V) = d_b(Diagram(U), Diagram(V)).$$

Definition 10.1.3. *We now define an algebraic-flavored distance. Consider two persistence modules V and W. Given $\epsilon \geq 0$, an ϵ-morphism between V and W is a family of linear maps $\phi = (\phi_t : V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagram commutes for every $s \leq t \in \mathbb{R}^+$:*

$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \phi_s \downarrow & & \downarrow \phi_t \\ W^{s+\epsilon} & \xrightarrow[w_{s+\epsilon}^{t+\epsilon}]{} & W^{t+\epsilon} \end{array}$$

Definition 10.1.4. *An ϵ-interleaving between V and W is a pair of ϵ-morphisms $(\phi_t : V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ and $(\psi_t : W^t \to V^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagrams commute for every $t \in \mathbb{R}^+$:*

$$\begin{array}{ccccc} V^t & & \xrightarrow{v_t^{t+2\epsilon}} & & V^{t+2\epsilon} \\ & \phi_t \searrow & & \nearrow \psi_{t+\epsilon} & \\ & & W^{t+\epsilon} & & \end{array} \qquad \begin{array}{ccccc} & & V^{t+\epsilon} & & \\ & \psi_t \nearrow & & \searrow \phi_{t+\epsilon} & \\ W^t & & \xrightarrow[w_t^{t+2\epsilon}]{} & & W^{t+2\epsilon} \end{array}$$

Definition 10.1.5. *The interleaving distance between two persistence modules V and W is defined as*

$$d_i(V, W) = \inf\{\epsilon \geq 0, V \text{ and } W \text{ are } \epsilon - \text{ interleaved}\}.$$

10.2 Exercises

Exercise 53. *Let $\mathcal{M}$ be the unit circle of $\mathbb{R}^2$, and $X \subset \mathbb{R}^2$ a finite subset. Denote the Hausdorff distance $\epsilon = d_H(X, \mathcal{M})$. Suppose that ϵ is small enough. Let $\mathbb{U}$ denote the persistence module of the 1st homology of the Cech filtration of X. Using the stability theorem, deduce the existence of a bar in the barcode, and give a lower bound on its persistence. Compare your result with Exercise 51.*

11 Python tutorial

The tutorial can be found in the Github on the link:
github.com/lucasmoschen/topological-data-analysis/blob/main/tutorials/tutorial-3.ipynb