Tutorial on Topological Data Analysis

Kunlin

July 15, 2020

1 Metric Topology

1.1 Motivation

Question: How to do "*smooth deformation*" of an object.

Definition 1.1. *Smooth deformation will not change the shape of an object.*

If we define a shape by counting the number of holes in the object...

You cannot change a ball into a donut, because a ball has no holes a donut has one hole.

In TDA:

- Start from a data set
- Construct topological objects
- Study their properties that are not changed by smooth deformation

Challenges:

- How to define shapes (number of holes)?
- How to count holes when dimension $n = 2, 3, 4, \cdots$

Idea: dimension reduction while maintaining homology.

1.2 Open Sets

Definition 1.2. *Let* (X, d) *be a metric space. For each element* $a \in X$ *and each* $r \in (0, \infty) \subseteq \mathbb{R}$*, the* open ball with a center a and radius r *is the subset*

$$B_{a,r} = \{x \in X | d(a, x) < r\} \subseteq X \tag{1}$$

Example 1.3. *Let (X, d) be a metric space. If $r > 0$, then*

$$B_{a,r} = \{x \in X | d(a, x) < r\} \subseteq X \tag{2}$$

is open.

Solution. If $y \in B(a, r)$, then $\delta = r - d(a, y) > 0$ and, whenever $d(z, y) < \delta$, the triangle inequality gives us

$$d(a, z) \leq d(a, y) + d(y, z) < r$$

so $z \in B(a, r)$. Thus $B(a, r)$ is open.

Theorem 1.4. *If (X, d) is a metric space, then the following statements are true.*

(i) The empty set $\emptyset$ and the space X are open.

(ii) If U_α is open for all $\alpha \in A$, then $\bigcup_{\alpha \in A} U_\alpha$ is open. (In other words, the union of open sets is open.)

(iii) If U_j is open for all $1 \leq j \leq n$, then $\bigcap_{j=1}^{n} U_j$ is open.

Proof. (i) Since there are no points e in $\emptyset$, the statement

$$x \in \emptyset \text{ whenever } d(x, e) < 1$$

holds for all $e \in \emptyset$. Since every point x belongs to X, the statement

$$x \in X \text{ whenever } d(x, e) < 1$$

holds for all $e \in X$.

(ii) If $e \in \bigcup_{\alpha \in A} U_\alpha$, then we can find a particular $\alpha_1 \in A$ with $e \in U_{\alpha_1}$. Since U_{α_1} is open, we can find a $\delta > 0$ such that

$$x \in U_{\alpha_1} \text{ whenever } d(x, e) < \delta.$$

Since $U_{\alpha_1} \subseteq \bigcup_{\alpha \in A} U_\alpha$,

$$x \in \bigcup_{\alpha \in A} U_\alpha \text{whenever } d(x, e) < \delta.$$

Thus $\bigcup_{\alpha \in A} U_\alpha$ is open.

(iii) If $e \in \bigcap_{j=1}^{n} U_j$, then $e \in U_j$ for each $1 \leq j \leq n$. Since U_j is open, we can find a $\delta_j > 0$ such that

$$x \in U_j \text{ whenever } d(x, e) < \delta_j.$$

Setting $\delta = \min_{1 \le j \le n} \delta_j$, we have $\delta > 0$ and

$$x \in U_j \text{ whenever } d(x, e) < \delta$$

for all $1 \le j \le n$. Thus

$$x \in \bigcap_{j=1}^{n} U_j \text{ whenever } d(x, e) < \delta$$

and we have shown that $\bigcap_{j=1}^{n} U_j$ is open. □

Example 1.5. *Here is an example of how property (iii) faiuls when we consider the intersection of an infinite family of open subsets. Let $X = \mathbb{R}$ with the Euclidean metric, $d(x, y) = |x - y|$, and let $U_i = (-\frac{1}{i}, \frac{1}{i})$ for all $i > 1$. Then, $\bigcap_{i=1}^{\infty} U_i = \{0\}$, which is not open.*

Example 1.6. *If we work in $\mathbb{R}^n$ with the Euclidean metric, then the one point set $\{\mathbf{x}\}$ is not open.*

Proof. Choose $\mathbf{e} \in \mathbb{R}^n$ with $\|\mathbf{e}\|_2 = 1$. (We could take $\mathbf{e} = (1, 0, 0, \ldots, 0)$.) If $\delta > 0$, then, setting $\mathbf{y} = \mathbf{x} + (\delta/2)\mathbf{e}$, we have $\|\mathbf{x} - \mathbf{y}\|_2 < \delta$, yet $\mathbf{y} \notin \{\mathbf{x}\}$. Thus $\{\mathbf{x}\}$ is not open. □

1.3 Continuous functions between metric spaces

Definition 1.7. *Lext (X, d_X) and (Y, d_Y) be metric spaces. A map $f : X \to Y$ such that*

$$d_Y(f(x), f(x^{'})) = d_X(x, x^{'}) \tag{3}$$

for all $x, x^{'} \in X$ is called isometry between X and Y.

Definition 1.8. Isometry *is a distance-preserving transformation between metric spaces, usually assumed to be bijective.*

Example 1.9. *Conisder the sets $\mathbb{R}^2$ and $\mathbb{C}$, both with Euclidean metric. Then, the map $f : \mathbb{C} \to \mathbb{R}^2$ defined by $f(a + bi) = (a, b)$, which is isometry.*

Example 1.10. *Is $f(x) = x^2$ an isometry on Euclidean distance? $d(3, 4) = \sqrt{9 + 16} = 5$ but $d(f(3), f(4)) = d(9, 16) = \sqrt{81 + 256} = \sqrt{337}$*

Theorem 1.11. *Let (X, d) and (Y, ρ) be metric spaces. A function $f : X \to Y$ is continuous if and only if $f^{-1}(U)$ is open in X whenever U is open in Y.*

Proof. Suppose first that f is continuous and that U is open in Y. If $x \in f^{-1}(U)$, then we can find a $y \in U$ with $f(x) = y$. Since U is open in Y, we can find an $\epsilon > 0$ such that

$$z \in U \text{ whenever } \rho(y, z) < \epsilon.$$

Since f is continuous, we can find a $\delta > 0$ such that

$$\rho(y, f(w)) = \rho(f(x), f(w)) < \epsilon \text{ whenever } d(x, w) < \delta.$$

Thus

$$f(w) \in U \text{ whenever } d(x, w) < \delta.$$

In other words,

$$w \in f^{-1}(U) \text{ whenever } d(x, w) < \delta.$$

We have shown that $f^{-1}(U)$ is open.

We now seek the converse result. Suppose that $f^{-1}(U)$ is open in X whenever U is open in Y. Suppose $x \in X$ and $\epsilon > 0$. We know that the open ball

$$B(f(x), \epsilon) = \{y \in Y \,:\, \rho(f(x), y) < \epsilon\}$$

is open. Thus $x \in f^{-1}\big(B(f(x), \epsilon)\big)$ and $f^{-1}\big(B(f(x), \epsilon)\big)$ is open. It follows that there is a $\delta > 0$ such that

$$w \in f^{-1}\big(B(f(x), \epsilon)\big) \text{ whenever } d(x, w) < \delta,$$

so, in other words,

$$\rho(f(x), f(w)) < \epsilon \text{ whenever } d(x, w) < \delta.$$

Thus f is continuous. □

Theorem 1.12. *If (X, d), (Y, ρ), (Z, σ) are metric spaces and $g : X \to Y$, $f : Y \to Z$ are continuous, then so is the composition $f * g$*

Proof. If U is open in Z, then, by continuity, $f^{-1}(U)$ is open in Y and so, by continuity, $(fg)^{-1}(U) = g^{-1}(f^{-1}(U))$ is open in X. Thus $f * g$ is continuous. □