Topological Data Analysis

1 December 2022

Exercises

1. Consider the functions $f, g\colon [-1,1] \to \mathbb{R}$ given by
$$f(x) = x^5 - x, \qquad g(x) = \frac{1}{5}(x^9 + 7x^5 - 10x).$$
   (a) Find the persistence modules $V(f)$ and $V(g)$ and the spectrum of each.
   (b) Compute the interleaving distance $d_{\text{int}}(V(f), V(g))$.
   (c) Check that $d_{\text{int}}(V(f), V(g)) < \|f - g\|_\infty$ on $[-1,1]$.

2. Consider the following point clouds in $\mathbb{R}^2$:
$$X = \{(0.81, 2.87), (2.15, 1.18), (3.19, 3.62), (4.17, 2.01), (5.32, 4.88), (6.21, 3.13)\},$$
$$Y = \{(0.75, 2.80), (2.33, 1.25), (3.28, 3.66), (4.15, 2.15), (5.24, 4.78), (6.34, 3.12)\}.$$
   (a) Compute the Hausdorff distance $d_H(X,Y)$ and the Gromov–Hausdorff distance $d_{GH}(X,Y)$.
   (b) Compute the bottleneck distance $W_\infty(D(X), D(Y))$ between the Vietoris–Rips persistence diagrams of X and Y.
   (c) Check that $W_\infty(D(X), D(Y)) < 2\, d_{GH}(X,Y)$.

3. Prove that the Gromov–Hausdorff distance between a single point and a non-empty compact subset K of a metric space is equal to half the diameter of K.

Please deliver through Campus Virtual as a pdf file before December 15 at 10:00.

Longer exercises (optional)

1. Let X_1 be the following point cloud:

$(0.83089090, 0.92139106)$	$(0.14943596, 0.59539077)$	$(-0.42530458, 0.49917853)$
$(1.05135118, 0.69658666)$	$(1.40450974, 1.63377801)$	$(1.20266640, 0.79526767)$
$(1.40570229, 1.45601583)$	$(0.81189552, 1.27512527)$	$(1.23633497, 1.81306232)$
$(0.05738789, 1.58270914)$	$(2.08298134, 0.87369037)$	$(0.42736194, 1.51483777)$
$(1.83601599, 2.31975042)$	$(2.53595820, 1.29246603)$	$(0.86969203, 0.99283119)$
$(0.47738787, 0.41686889)$	$(1.01228192, 0.15565445)$	$(2.11973316, 1.44747695)$
$(0.58929427, 1.24252248)$	$(2.05421121, 2.05534683)$	$(1.37592899, 0.96259255)$
$(1.65529594, 0.98534007)$	$(0.75427762, 0.75383665)$	$(0.99710074, 1.67116645)$
$(0.82368247, 1.00514813)$	$(-0.14872117, 0.02096184)$	$(0.61740303, 1.30389238)$
$(1.49864810, 0.92041849)$	$(1.47630053, 0.87510234)$	$(1.81213878, 1.91489774)$

and let X_2 be the following point cloud:

$(1.40754745, 2.647683)$	$(1.90579751, 3.350476)$	$(1.86389493, 3.231760)$
$(1.24523782, 2.820778)$	$(2.33118508, 3.963780)$	$(2.58430981, 3.650785)$
$(1.55517672, 3.303610)$	$(0.60439609, 2.430479)$	$(2.20839977, 3.496688)$
$(1.97133972, 2.321545)$	$(1.98517772, 2.459581)$	$(3.53630008, 3.469574)$
$(1.39970730, 2.477310)$	$(2.81785877, 3.541370)$	$(2.13849006, 3.351805)$
$(1.15300714, 3.251407)$	$(-0.02755078, 1.812042)$	$(2.78733827, 3.396214)$
$(2.04172416, 4.218545)$	$(2.06339041, 3.167594)$	$(1.92692528, 2.975792)$
$(2.00200034, 2.743398)$	$(0.75285330, 3.240254)$	$(3.52340370, 3.839808)$
$(1.82156736, 3.471864)$	$(2.15114208, 2.660630)$	$(1.95290166, 3.071435)$
$(1.75719936, 3.097291)$	$(2.28021529, 2.584600)$	$(0.99651700, 2.791795)$

Find out whether the next set of points belongs to the same source as X_1 or to that of X_2. (*Hint:* Compute persistence diagrams of X_1 and X_2 with and without this additional set of data, and compare distances.)

$(2.29024248, 2.526838)$	$(2.94471778, 3.495130)$	$(1.75890949, 2.685264)$
$(2.13586188, 2.792384)$	$(1.59348125, 3.347204)$	$(0.93981079, 2.512890)$
$(2.64627954, 2.317721)$	$(2.50426440, 3.819559)$	$(2.37635825, 2.883139)$
$(2.01157856, 2.221607)$		

2. Find an article in the Internet in which persistent homology is used as a tool for data analysis in biomedical sciences, social sciences, finance, complex networks, neural networks, etc. Describe the objectives of the article and summarize its conclusions.

3. Find out how to associate a barcode to any tree (i.e., a graph with no cycles) in Euclidean space $\mathbb{R}^3$ by means of the algorithm called *topological morphology descriptor* (TMD), and provide examples. See L. Kanari et al., A topological representation of branching neuronal morphologies, *Neuroinformatics* 16 (2018), 3–13 (open access), https://doi.org/10.1007/s12021-017-9341-1.

4. Read and explain the definition of *multiparametric persistence modules* and the corresponding interleaving distance. See M. Lesnick, The theory of the interleaving distance on multidimensional persistence modules, *Found. Comput. Math.* 15 (2015), 613–650, also available from arXiv:1106.5305, 2015.

5. Read and explain the definition of *persistent homology dimension* (PH dimension) from J. Jaquette, B. Schweinhart, Fractal dimension estimation with persistent homology: A comparative study, *Commun. Nonlinear Sci. Numer. Simulat.* 84 (2020), 105163, also available from arXiv:1907.11182, 2019. Discuss its use for practical estimations of fractal dimensions of sets such as the Sierpiński triangle using random point samples.

Longer exercises can be delivered until December 20.