Directed Reading Program

Spring 2021

Week 1

Readings Section 1 – 3 of [Tausz and Carlsson, 2011]

Questions

1. (revision) Recall that p.152 – 156 of [Edelsbrunner and Harer, 2010] gives an algorithm to compute the persistent homology of a filtration. Compute the persistent homology of the solid tetrahedron (with respect to a sensible filtration).
2. (zig-zag persistence)
   (a) Explain how zig-zag persistence helps decompose a large dataset into smaller ones whose persistence homology can be computed in practice.
   (b) Explain what a bar in the figure on p.6 of [Tausz and Carlsson, 2011] means.

Week 2

Readings Section 7 of [Otter et al., 2017]

Questions

1. (revision) Recall that p.152 – 156 of [Edelsbrunner and Harer, 2010] gives an algorithm to compute the persistent homology of a filtration. Compute the persistent homology of the solid tetrahedron (with respect to a sensible filtration).
2. (algorithm comparison)Among algorithms listed in the paper, which is the most efficient for generic purposes? Which are the runners-up? What are their limitations?

Week 3

Readings [Cavanna et al., 2015b] and the video described in its abstract

Questions

1. (revision) Recall that p.152 – 156 of [Edelsbrunner and Harer, 2010] gives an algorithm to compute the persistent homology of a filtration. Compute the persistent homology of the solid tetrahedron with the following filtration. For notational simplicity, we identify a vertex with the singleton of its index and a simplex with the set of the indices of the vertices.

$$\begin{aligned}
f(\{1\}) &= 1 \\
f(\{2\}) &= 2 \\
f(\{3\}) &= 3 \\
f(\{1,2\}) &= 4 \\
f(\{4\}) &= 5 \\
f(\{1,3\}) &= 6 \\
f(\{2,3\}) &= 7 \\
f(\{1,2,3\}) &= 8 \\
f(\{1,4\}) &= 9 \\
f(\{2,4\}) &= 10 \\
f(\{3,4\}) &= 11 \\
f(\{1,3,4\}) &= 12 \\
f(\{2,3,4\}) &= 13 \\
f(\{1,2,4\}) &= 14 \\
f(\{1,2,3,4\}) &= 15
\end{aligned}$$

   Below is a description of the filtration to make sure we are on the same page. In this filtration, the first three connected components appear at filtration value $1, 2, 3$, while the second connected component dies at value 4 when vertices 1 and 2 are joined. The first 1-dimensionoal homological feature appears at filtration value 7 when the hollow triangle with vertices $1, 2, 3$ and dies immediately afterwards at filtration value 8 when the triangular simplex is added.

2. (ball truncation) Convince yourself that very few balls will be covered by other balls if not the growth of balls is not stopped.

3. (filtration approximation) Identify the parts of Lemma 1 that
   - guarantees the new filtration is indeed a filtration, and
   - guarantees the new filtration approximates the old one.

4. (higher-dimensional perspective) Convince yourself the nerve of the higher-dimensional object is indeed the same as those in filtration.
5. (ball removal) Draw a picture for the higher-dimensional perspective for a 1D dataset where the removal of balls does reduce the size of nerves.
6. (optional; choice of threshold) Convince yourself the choice of λ_i's are natural.
7. (optional; embeddinng in ℓ^∞) Show that a metric space with n points can be embedded in $\ell^\infty(\mathbb{R}^n)$. (This is called Frechet's theorem, and is Proposition 15.6.1 of [Matoušek, 2002].)

Week 4

Readings [Cavanna et al., 2015a] and the video described in its abstract; [Moitra et al., 2018] if you have extra time

Questions

1. (revision) Recall that p.152 – 156 of [Edelsbrunner and Harer, 2010] gives an algorithm to compute the persistent homology of a filtration. Compute the persistent homology of the solid tetrahedron with the following filtration. For notational simplicity, we identify a vertex with the singleton of its index and a simplex with the set of the indices of the vertices.

$$\begin{aligned}
f(\{1\}) &= 1\\
f(\{2\}) &= 2\\
f(\{3\}) &= 3\\
f(\{1,2\}) &= 4\\
f(\{4\}) &= 5\\
f(\{1,3\}) &= 6\\
f(\{2,3\}) &= 7\\
f(\{1,2,3\}) &= 8\\
f(\{1,4\}) &= 9\\
f(\{2,4\}) &= 10\\
f(\{3,4\}) &= 11\\
f(\{1,3,4\}) &= 12\\
f(\{2,3,4\}) &= 13\\
f(\{1,2,4\}) &= 14\\
f(\{1,2,3,4\}) &= 15
\end{aligned}$$

Below is a description of the filtration to make sure we are on the same page. In this filtration, the first three connected components appear at filtration value $1, 2, 3$, while the second connected component dies at value 4 when vertices 1 and 2 are joined. The first 1-dimensionoal homological feature appears at filtration value 7 when the hollow triangle with vertices $1, 2, 3$ and dies immediately afterwards at filtration value 8 when the triangular simplex is added.

2. (ball truncation) Convince yourself that very few balls will be covered by other balls if not the growth of balls is not stopped.
3. (filtration approximation) Identify the parts of Lemma 1 that
   - guarantees the new filtration is indeed a filtration, and
   - guarantees the new filtration approximates the old one.
4. (higher-dimensional perspective) Convince yourself the nerve of the higher-dimensional object is indeed the same as those in filtration.
5. (ball removal) Draw a picture for the higher-dimensional perspective for a 1D dataset where the removal of balls does reduce the size of nerves.
6. (optional; choice of threshold) Convince yourself the choice of λ_i's are natural.
7. (optional; embeddinng in ℓ^∞) Show that a metric space with n points can be embedded in $\ell^\infty(\mathbb{R}^n)$. (This is called Frechet's theorem, and is Proposition 15.6.1 of [Matoušek, 2002].)
8. (optional for this week) Summarize the technique used in [Moitra et al., 2018] to reduce the number of points in the dataset. Identify the theoretical guarantee.

Week 5

Readings [Moitra et al., 2018, Bauer, 2019]

Questions

1. Summarize the technique used in [Moitra et al., 2018] to reduce the number of points in the dataset. Identify the theoretical guarantee.
2. Identify tricks Ripser used to speed up the computation of persistence diagrams. (We will look into each of them in detail in weeks to come. You are not expected to understand each of them now.)

Week 6

Readings [Chen and Kerber, 2011]

Questions

1. Illustrate the technique used in [Chen and Kerber, 2011] to speed up persistent homology calculation with the (now updated) example from week 4.
2. Recall from Section 3.4 of [Bauer, 2019] that one column of R needs to be stored when implementing the algorithm in [Chen and Kerber, 2011]. Illustrate this with the example from week 4.

Week 7

Readings [Chen and Kerber, 2011, de Silva et al., 2011]

Questions

1. Illustrate the technique used in [Chen and Kerber, 2011] to speed up persistent homology calculation with the (now updated) example from week 4.
2. Recall from Section 3.4 of [Bauer, 2019] that one column of R needs to be stored when implementing the algorithm in [Chen and Kerber, 2011]. Illustrate this with the example from week 4.
3. Explain in your own words why homology and cohomology have the same barcodes.

Week 8

Readings [Chen and Kerber, 2011, de Silva et al., 2011]; skip parts about relative homology and cohomology

Questions

1. Recall from Section 3.4 of [Bauer, 2019] that one column of R needs to be stored when implementing the algorithm in [Chen and Kerber, 2011]. Illustrate this with the example from week 4.
2. Construct D and $D^{\perp}$ for the filtration defined in week 4. (D is just the boundary matrix; for $D^{\perp}$, see Section 2.6 and 3.1 of [de Silva et al., 2011]).
3. Illustrate the computation of persistent cohomology described in [de Silva et al., 2011] by applying `pHcol` (Algorithm 1) to $D^{\perp}$. (Do *not* use twist in [Chen and Kerber, 2011]). What do zero columns and lowest ones now mean? Is this result consistent with persistent homology? Are there exceptions to your interpretations?
4. Implement `pHrow` (Algorithm 2) on D and $D^{\perp}$. Explain in your own words how the algorithm "examine the future" (see Section 4.2). How does it speed up the computation?
5. Is twist in [Chen and Kerber, 2011] compatible with `pHrow` (Algorithm 2)? If so, how?

Week 9

Readings Reread [Bauer, 2019]; skip parts about emergent pairs (mainly Section 3.5).

Questions

1. Make sense of the discussion at the end of Section 3.2 about the combination of persistent cohomology and clearing. Illustrate the combination for the filtration of week 4.
2. This is your second reading of [Bauer, 2019], after diving into the papers it cites. Does it make more sense now? Summarize the paper except the part about emergent pairs.

Notes on Point-Set Topology [Arnold, 2011] is a gentle introduction to the intuitive concepts in topology. Prof. Hatcher's note [Hatcher, 2005] strikes a good balance between rigor and intuition (and length!). [Munkres, 2000] is a standard textbook of the field. [Engelking, 1989] is a standard reference of the field.

Week 10

Readings Reread [Bauer, 2019]; pay attention to parts with identities about binomial coefficient (which counts the number of operations needed in persistent homology computation) and emergent pairs (mainly Section 3.5).

Questions

1. Make sense of the discussion in Section 3.2 and 3.3 about the number of reductions when different speedups are used.
2. Give an example of a pair that is
    - both an emergent face pair and an emergent coface pair.
    - an emergent face pair but not an emergent coface pair.
    - an emergent coface pair but not an emergent face pair.
    - neither an emergent face pair nor an emergent coface pair.
3. How does emergent pairs enter the algorithm? (Hint: "Column Addition" part of Section 4)
4. How is Section 3.6 related to the rest of the paper?

Week 11

Readings [Mendoza-Smith and Tanner, 2017]

Questions

1. Identify the main contribution of the paper.
2. Do the pictures look pretty? Why?
3. Identify the major prior results the paper builds on.
4. What is the main parallelization strategy?

References

[Arnold, 2011] Arnold, B. (2011). *Intuitive Concepts in Elementary Topology.* Dover Books on Mathematics. Dover Publications.

[Bauer, 2019] Bauer, U. (2019). Ripser: efficient computation of Vietoris-Rips persistence barcodes. Preprint.

[Cavanna et al., 2015a] Cavanna, N. J., Jahanseir, M., and Sheehy, D. R. (2015a). A geometric perspective on sparse filtrations. In *Proceedings of the Canadian Conference on Computational Geometry.*

[Cavanna et al., 2015b] Cavanna, N. J., Jahanseir, M., and Sheehy, D. R. (2015b). Visualizing sparse filtrations. In *Proceedings of the 31st International Symposium on Computational Geometry.*

[Chen and Kerber, 2011] Chen, C. and Kerber, M. (2011). Persistent homology computation with a twist. In *Proceedings 27th European Workshop on Computational Geometry.*

[de Silva et al., 2011] de Silva, V., Morozov, D., and Vejdemo-Johansson, M. (2011). Dualities in persistent (co)homology. *Inverse Problems*, 27(12):124003.

[Edelsbrunner and Harer, 2010] Edelsbrunner, H. and Harer, J. (2010). *Computational Topology: An Introduction.* Applied Mathematics. American Mathematical Society, Providence, RI.

[Engelking, 1989] Engelking, R. (1989). *General Topology.* Sigma series in pure mathematics. Heldermann.

[Hatcher, 2005] Hatcher, A. (2005). Notes on introductory point-set topology. `http://pi.math.cornell.edu/~hatcher/Top/Topdownloads.html`.

[Matoušek, 2002] Matoušek, J. (2002). *Lectures on Discrete Geometry.* Sprinnger-Verlag, 1st edition.

[Mendoza-Smith and Tanner, 2017] Mendoza-Smith, R. and Tanner, J. (2017). Parallel multi-scale reduction of persistent homology filtrations.

[Moitra et al., 2018] Moitra, A., Malott, N. O., and Wilsey, P. A. (2018). Cluster-based data reduction for persistent homology. In *2018 IEEE International Conference on Big Data (Big Data)*, pages 327–334.

[Munkres, 2000] Munkres, J. (2000). *Topology*. Featured Titles for Topology. Prentice Hall, Incorporated.

[Otter et al., 2017] Otter, N., Porter, M. A., Tillmann, U., Peter, G., and Harrington, H. A. (2017). A roadmap for the computation of persistent homology. *EPJ Data Science*, 6:17.

[Tausz and Carlsson, 2011] Tausz, A. and Carlsson, G. (2011). Applications of Zigzag Persistence to Topological Data Analysis. *arXiv e-prints*, page arXiv:1108.3545.