Alessio Pellegrini
Ana Zegarac

Student Seminar in Topological Data Analysis

Spring Semester 2023

Topics

0 Introduction

The format and topics of seminar are explained and the distribution of the talks up to the Easter break take place.

Below you can find an overview with references and the content of each week's talks. Check out the TDA seminar[1] run by Marius Thaule and Melvin Vaupel for an extensive list of references.

1 Simplicial Complexes, 28.02.23

References: The relevant definitions appear across the work of Carlsson et al [Car09; CVJ21; Car14].

Content: An *(abstract) simplicial complex* is a combinatorial object consisting of vertices and simplices with certain properties. In TDA[2] there are several natural ways to cook up simplicial complexes associated to some real-life data (coming in the form of a finite metric space X, some subset of $\mathbb{R}^n$...). We will mostly work with *Alpha, Čech, and Vietoris–Rips complexes*. The goal of the talk is to introduce and explain these different notions and put an emphasis on their difference using examples.

2 Homology Groups, 07.03.23

References: The standard references are [Bre97; Hat02; May99; Spa95]. Some TDA papers also contain "quick and dirty" introductions to homology, see the references in Week 1.

Content: *Homology* associates to any topological space X a sequence of abelian groups denoted by $H_n(X)$, $n \in \mathbb{Z}$. There are several equivalent definitions of homology, but most of them are not well suited for explicit computations. In the case where X comes in the form of a simplicial complex however, one can compute the homology by hand, or using a computer program. In this talk we would like to revisit/introduce the key notions of every homology theory, that is chains, boundaries and cycles, recall

[1] In case you are not using a PDF–viewer: https://wiki.math.ntnu.no/ma3001/2021h/tda/start
[2] Short for topological data analysis.

functoriality and Betti numbers, and refresh some distinctive features of homology. Examples are key here.

3 Introduction to Javaplex, 14.03.23

References: The tutorial [AT+11] is well written and easy to "code–along". The programming language used is mathlab.[3]

Content: Implementing some simplicial complexes using Javaplex in class and compute their Betti numbers.

4 Persistent Homology, 21.03.2023

References: Carlsson's papers are good and relevant for TDA purposes with emphasis on intuition, whereas Polterovich's book is rigorous and probably easier to read for those with the pure math background [Pol+20; Car09; CVJ21].

Content: The notions of *persistence vector spaces* and their associated *persistent homology* are at the center of TDA. The general theory is very well explained in the book of Polterovich et al. The idea in TDA is to refine the homology invariant by introducing a meaningful filtration and observing how the homology changes as one follows the filtration. The goal of the talk is to define and introduce the necessary material to make sense of the definition of persistent homology, and explain the concepts using several examples of filtered simplicial complexes introduced in Week 1.

5 Barcodes, Classification Theorem, 28.03.2023

References: [Pol+20]

Content: *Barcodes* are examples of persistence vector spaces that are relatively easy to understand and of major importance due to the *Classification Theorem*, which asserts that every persistence vector space is isomorphic to a "unique" barcode. The goal of the talk is to get a good grasp of barcodes and explain the Classification Theorem with a proof, if time permits.

[3] Since you've all had a course in numerics we assume that you have access to a computer with a running version of mathlab. If not, let us know.

Easter Break, no seminar on the 11.04.23

6 Natural Image Patches, 18.04.23

References: [Car+08; Car14]

Content: Topological Data Analysis can be used in image analysis. The goal of this talk would be to explain how the theory we've seen so far can be used in image analysis by summarising the paper [Car+08].

7 Stability of Persistent Homology, 25.04.23

References: [Pol+20]

Content: It is possible to define a metric on the space of barcodes. The goal of this talk is to define the bottleneck and the interleaving distance and to state the Isometry Theorem.

8 Reeb Graphs and Mapper, 16.05.23

References: [DW22; Car09; SMC+07]

Content: The goal of this talk is to define the Reeb graph and the (topological and the statistical) Mapper and show examples of both.

9 Applications in the Analysis of Breast Cancer Data, 23.05.23

References: [NLC11]

Content: In [NLC11], Mapper was used to identify a new subgroup of breast cancers that was invisible to cluster analysis. The goal of this talk is to summarlise the paper [NLC11].

References

[AT+11] H. Adams, A. Tausz, et al. "Javaplex tutorial". *Google Scholar* (2011).

[Bre97] G. E. Bredon. *Topology and geometry.* Vol. 139. Graduate Texts in Mathematics. Corrected third printing of the 1993 original. Springer-Verlag, New York, 1997, pp. xiv+557.

[Car14] G. Carlsson. "Topological pattern recognition for point cloud data". *Acta Numerica* 23 (2014), pp. 289–368.

[Car09] G. Carlsson. "Topology and data". *Bulletin of the American Mathematical Society* 46.2 (2009), pp. 255–308.

[Car+08] G. Carlsson, T. Ishkhanov, V. De Silva, and A. Zomorodian. “On the local behavior of spaces of natural images”. *International journal of computer vision* 76 (2008), pp. 1–12.

[CVJ21] G. Carlsson and M. Vejdemo-Johansson. *Topological data analysis with applications*. Cambridge University Press, 2021.

[DW22] T. K. Dey and Y. Wang. *Computational topology for data analysis*. Cambridge University Press, 2022.

[Hat02] A. Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002, pp. xii+544.

[May99] J. P. May. *A concise course in algebraic topology*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1999, pp. x+243.

[NLC11] M. Nicolau, A. J. Levine, and G. Carlsson. “Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival”. *Proceedings of the National Academy of Sciences* 108.17 (2011), pp. 7265–7270.

[Pol+20] L. Polterovich, D. Rosen, K. Samvelyan, and J. Zhang. *Topological persistence in geometry and analysis*. Vol. 74. American Mathematical Soc., 2020.

[SMC+07] G. Singh, F. Mémoli, G. E. Carlsson, et al. “Topological methods for the analysis of high dimensional data sets and 3d object recognition.” *PBG@ Eurographics* 2 (2007), pp. 091–100.

[Spa95] E. H. Spanier. *Algebraic topology*. Corrected reprint of the 1966 original. Springer-Verlag, New York, 1995, pp. xvi+528.