Graduate Topological Data Analysis Seminar

October 17, 2022

Abstract

We'll be working together as people with different backgrounds in mathematics to try to gain a working understanding of topological data analysis. The plan is to work through the core chapters of the main reference, and then (probably late next semester) pick some papers on applications that we find interesting and give talks on them. We'll hopefully have those of us with a geometry/topology/homotopy theory background present on the first few chapters of the main reference this semester to bring those of us with other backgrounds up to speed, and then they in turn can start off next semester presenting on the next few chapters which focus more on MCS/STAT stuff.

I've tried my best to find references at varying levels of difficulty/detail in each area, and hopefully most of them are available for free online, but please let me know if you have trouble finding any of them (or know of other good sources to add)! This blog post also has a huge list of references on data science for mathematicians, some of which I included below. **Please don't feel overwhelmed by this list; we're definitely going to mostly stick to the main reference and the overview articles. This is just intended as a repository of sources for filling in background information from the main reference as needed, or for further reading if anything sparks your interest.**

1 Main reference

- Zomorodian: "Topology and Computing"

This is a book on topological data analysis that I've heard is the best source to learn it from. It probably leans a little more towards people with an MCS background, so us homotopy theorists and geometry/topology people may have to rely a little more heavily on the overview articles and some of the background materials in MCS/STAT topics.

2 Other references

2.1 Overview articles

- Carlsson: "Topology and Data"
- Carlsson: "Persistent Homology and Applied Homotopy Theory"
- Otter et al.: "A roadmap for the computation of persistent homology"

These are three overview articles that we should all try to read at least one of, preferably before we get started on the main text. The two Carlsson articles might be better suited to those with a geometry/topology background and homotopy theory background (respectively), while the third article is probably best for those with other backgrounds.

2.2 Background References

2.2.1 Algebra

- Fraleigh: "A First Course in Abstract Algebra"
- Aluffi: "Algebra: Chapter 0"

These are great resources for abstract and homological algebra at the undergraduate and graduate level, respectively. Fraleigh's "Groups in Topology" section is also incidentally my favorite introduction to homological algebra and algebraic topology at the undergraduate level. Aluffi also covers a considerable amount of category theory, and is an excellent reference on the topic for those with some algebra experience. It's also the standard textbook for the graduate algebra prelim sequence at UIC. The main reference goes through homology with coefficients in general abelian groups and R-modules (these things will be defined in the Groups chapter), but if those things scare you, you can certainly pretend everything is the category of real vector spaces to keep the algebra prerequisites to a minimum (i.e. an undergraduate linear algebra course) for those of us without an abstract algebra background. I can't find a (legal) free version of Aluffi, but there are several copies in the grad lounge, and feel free to email me for a digital copy.

2.2.2 Category Theory

- Riehl: "Category Theory in Context"
- Spivak: "Category Theory for Scientists"
- Barr and Wells: "Category Theory for Computing Science"

Riehl is the best introductory course in category theory that I know of, with lots of exercises. It is probably more geared toward the pure math-minded though. Spivak is an introductory course for those outside of pure math; I'm not super familiar with it, but it looks great, and I've heard Spivak is a great author. Barr and Wells seems to be really focused on MCS-type stuff, and it has solutions to at least some of the exercises, which is nice.

2.2.3 Topology and Homotopy Theory

- Munkres: "Topology"
- Hatcher: "Algebraic Topology"
- Bredon: "Topology and Geometry"

Munkres is the standard reference on point-set topology at the undergraduate level. It also covers some of the basics of homotopy theory in the later chapters, though this may not be the best presentation (although it may be somewhat more accessible to those without a background in geometry/topology given that it is an undergraduate text). Hatcher is the standard textbook on algebraic topology (and is used in the geometry/topology prelim course at UIC), but I personally hate it (other homotopy theorists do not necessarily feel this way). Bredon is one of my personal favorite books on algebraic topology, and also covers some material on differentiable manifolds which may come up at some point.

2.2.4 Data Science and CS

- Irizarry: "Introduction to Data Science"
- Aggarwal: "Neural Networks and Deep Learning"
- Kleinberg and Tardos: "Algorithm Design"

The first two books are taken from the blog post referenced above, but I know nothing about them. Irizarry has to be "purchased" through an online bookstore, but it is pay-what-you-want, so you can get it for free, and they'll email you a pdf link. The third book has been used for the introductory graduate algorithms course at UIC in the past, but again, I don't know anything about this stuff.

2.2.5 Statistics

- Ross: "A First Course in Probability"
- Keener: "Theoretical Statistics: Topics for a Core Course"

Ross is the probability theory book used at my undergrad, and seems to be the standard introduction. I'm not super familiar with it as I took probability theory at another school (which didn't use a textbook), but I have TA'd a class out of this book and it seemed good. Keener is the textbook used in the introductory statistics prelim graduate course at UIC, and also gives an overview of probability theory from a measure-theoretic viewpoint.

2.3 Applications

2.3.1 General

- iMSi Topological Data Analysis workshop

This is a workshop on TDA that was held at the iMSi in Chicago during Spring 2021, with video recordings of all of the talks. It may be a good place to find applications/current research in the field.

2.3.2 Machine Learning and Neural Networks

- Hensel et al.: "A Survey of Topological Machine Learning Methods"
- Magai and Ayzenberg: "Topology and geometry of data manifold in deep learning"
- Gerken et al.: "Geometric Deep Learning and Equivariant Neural Networks"

I figured many of you would be interested in applications of TDA to machine learning and neural networks, so I tried to find a few introductory/survey papers on those areas. However, I know nothing about this stuff, so I'm very open to suggestions!

2.3.3 Medicine and Biology

- Skaf and Laubenbacher: "Topological data analysis in biomedicine: A review"
- Salch et al.: "From mathematics to medicine: A practical primer on topological data analysis (TDA) and the development of related analytic tools for the functional discovery of latent structure in fMRI data"
- Bussola et al.: "Quantification of the Immune Content in Neuroblastoma: Deep Learning and Topological Data Analysis in Digital Pathology"
- Amézquita et al. "The shape of things to come: Topological data analysis and biology, from molecules to organisms"

My particular interest in TDA (besides its obvious relation to homotopy theory) lies in biomedical applications. These are a few survey articles that I found that look cool.