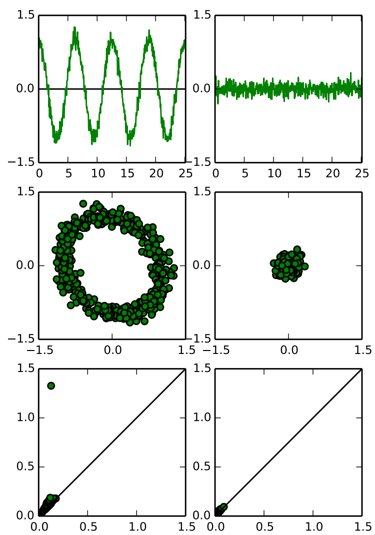
CMSE 491/890: Topological Analysis of Large Datasets

Class: M/W 10:20 - 11:40, 1220 Engineering Building
Credit hours: 3

Instructor Elizabeth Munch, Ph.D.
Asst Prof, Dept of CMSE, Dept of Math
Office 428 South Shaw Lane
Engineering Bldg, Rm 1511
E-mail muncheli@msu.edu (*Preferred method of contact)
Phone 517-432-0619
Office Hours TBD

Course Description:

Topology is the study of shapes. Recently, a great deal of work has gone into the study of using topological methods for problems in science and engineering, particularly in data analysis. This new research topic goes by many names, most often Computational Topology, Applied Topology, or Topological Data Analysis (TDA). We will work with many disparate fields such as algebraic topology, geometry, linear algebra, abstract algebra, algorithms, statistics, and sheaf theory in order to understand recent results in the field. We will study and use efficient software for the computation of things discussed in class, such as persistent homology and Reeb graphs. We will also look at applications in fields such as computer graphics, image analysis, sensor networks, clustering, time series analysis, and genetics.



Prerequisites:

Mathematical prerequisites: Linear algebra. Some familiarity of computer programming with packages such as Python, R, or MATLAB is expected. Individuals with backgrounds in mathematics, engineering, computer science, or other natural sciences with some computational training, will find this class of interest.

Website:

The D2L website, `https://d2l.msu.edu/`, has announcements about the course and any necessary files. Homework assignments and project information will also be announced and saved on this website.

Textbook:

1. *Required:* Herbert Edelsbrunner & John Harer. Computational Topology: An Introduction. American Mathematical Society, 2010.
2. *Recommended:* Munkres, James R. Elements of algebraic topology. Vol. 2. Reading: Addison-Wesley, 1984.

Grading:

- **Reading:** This course will be extremely reading heavy. Since we are looking at state of the art research, this will be a mix of textbook reading and research papers. Other than the textbook, all readings will be posted on blackboard.
- **Homework:** There will be approximately bi-weekly homework sets given which will be turned in for a grade. This homework will be typeset, preferably in LaTeX or no credit will be given.
- **Project:** The main component of this course is a project applying the tools you have learned in the class. More specifics will be given in the first few classes, but the general plan is as follows:
    1. The project can range from the applied (e.g. analyze a particular data set) to the theoretical (e.g. extend a particular theorem), and from the known (e.g. review a published result) to the

completely original. I will provide some examples/ideas/data sets for projects but, again, you have almost limitless freedom.

2. Evaluation will be based on an in class presentation during the last week of classes and/or finals week, as well as a writeup.
3. A project proposal should be presented by mid-semester (specific date to be determined). The proposal is to be at most one page long, describe what your project is going to be, what you need to do in order to complete it, and how it fits with what has been done in applied topology (i.e. does it extend a known theorem, is it another take on a known analysis, etc).

- **Points:** Your grade will be based on the total number of accumulated points from the semester. The *estimated* number of points is below.

	Estimated Points
Homeworks	5 homeworks × 20 points = 100
Project	100
TOTAL:	200

Tentative Schedule:
Note that this schedule is only provided as a guide and is very likely subject to change. Please check D2L regularly for reading assignments and true schedule.

Week		Date	Topic	Other
1	W	Aug 30	Linear Algebra Review	
2	M	Sep 4	No class (Labor Day)	
	W	Sep 6	Intro graph theory	
3	M	Sep 11	Simplicial complexes	
	W	Sep 13	Cpxes for point clouds/ cubical complexes	
4	M	Sep 18	*Computation – Simplicial Cpx*	
	W	Sep 20	Homology	
5	M	Sep 25	Computing Homology	
	W	Sep 27	Relative & local homology	
6	M	Oct 2	*Computation – Homology*	
	W	Oct 4	Union f nd, MSTs, 0-dim persistence	
7	M	Oct 9	Persistent Homology	
	W	Oct 11	Persistent Homology	
8	M	Oct 16	*Computation – Pers Hom*	
	W	Oct 18	*Computation – Pers Hom*	
9	M	Oct 23	Metrics & the stability theorem	Proposal Due
	W	Oct 25	Applications – Time series	
10	M	Oct 30	Applications – Image processing	
	W	Nov 1	Applications – Neuroscience	
11	M	Nov 6	Reeb graphs / Mapper	
	W	Nov 8	*Computation – Mapper*	
12	M	Nov 13	Mapper applications	
	W	Nov 15	Variations on Persistence	
13	M	Nov 20	Stats & Persistence – Landscapes, Images	Rough draft due
	W	Nov 22	Stats & Persistence – Means, Cis, ML	
14	M	Nov 27	Interleaving distance	
	W	Nov 29	Sensor Networks	
15	M	Dec 4		Presentations
	W	Dec 6		Presentations