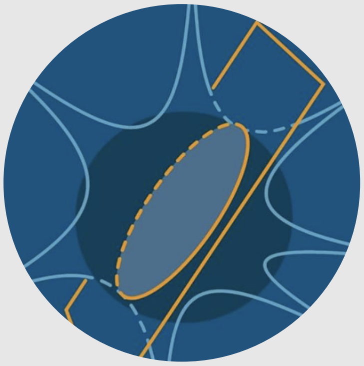

High Dimensional Probability

CS 12: Section 1

Instructor Info —

Ritvik Teegavarapu

OH: Thursdays 9:00 - 10:00 PM

Location: ANB 104

Website

rteegava@caltech.edu

Course Info —

Prereqs: ACM 104, ACM 116

Lecture: Thursdays 7:30 - 9:00 PM

Location: ANB 104

Overview

This course introduces the foundations of high-dimensional probability (HDP), a field that extends classical probability theory to settings involving large collections of random variables or random variables in high-dimensional spaces. While classical probability provides powerful asymptotic results under the assumption of *infinite* repetitions of an experiment, HDP focuses on non-asymptotic analysis by quantifying how randomness behaves when only a finite (but potentially large) number of observations or dimensions are available, albeit not as laconic.

High-dimensional probability is especially relevant in modern applications across computational mathematics, data science, and machine learning. The course emphasizes tools and techniques that are both broadly applicable and conceptually unified, forming a versatile probabilistic framework for analyzing complex random structures. Students will develop a core toolbox of HDP methods and gain an understanding of how these methods interact with and reinforce one another in practical and theoretical settings.

Material

Seperate lecture notes will be posted on `Piazza`. However, the lectures will be roughly based on the following primary source, which is Joel Tropp's notes from a previous iteration of the course. I have retooled the content to avoid measure theory, so read at your own risk[1]! Here are some other references that might be of interest:

1. Roman Vershynin's book is a secondary source that will highlight some of the applications of the techniques we will discuss.
2. Ramon von Handel's notes and Boucheron's book are comprehensive measure-theoretic treatments of the content.

Grading Scheme

This class is offered pass-fail, and there will be weekly problem sets due Thursday midnight, each with an Exercise, Problem, and Application.

- At least 15 out of the 30 questions must be completed to receive a Pass in the class.
- All 10 of the Exercises must be completed, and you may select a subset of 5 questions across the Problems and Applications across the term to complete.

You are free to use my office hours or `Piazza` to ask questions about the problem sets, or content in general!

Learning Objectives

This class will attempt to rigorously answer how close (and how fast) do we expect random variables to converge to some point estimate, and how likely is it for a random variable $Z = (X_1, \ldots, X_n)$ to deviate from it? Formally speaking, we will investigate the statement below.

$$\mathbb{P}\left\{|f(Z) - \mathbb{E}[f(Z)]| \geq t\right\} \text{ for } t \in \mathbb{R}^+$$

We will spend roughly the first half of the class developing fundamental statements in probability theory known as *functional inequalities*, and the second half on extending these techniques to random vectors/matrices with some remarks on the end on how to understand more generic forms of concentration.

We will also ground this content with many real-world applications, ranging from statistics & data science, machine learning, randomized algorithms, and many more!

[1] I am, of course, happy to clarify details, but the focus of the class is not on measure theory.

Class Schedule

The following is the weekly breakdown of the content covered in this class, with Week 0 being a probability primer document available on `Piazza` I expect you to read ahead of time. While the lecture notes will be self-contained and have all the content needed to complete the homework, I encourage you to read portions of the listed chapters time permitting. I will have "Appendices" in the lecture notes to add additional details/derivations that I did not have time to present in lecture.

Week 0	Review of Probability Theory	Lecture Notes 0 | Ch. 1 of Vershynin
Week 1	Exponential Concentration	Lecture Notes 1 | Ch. 4 of Tropp, Ch. 2 of Vershynin
Week 2	Sub-Gaussian Concentration	Lecture Notes 2 | Ch. 4 of Tropp, Ch. 2 of Vershynin
Week 3	Variance Bounds	Lecture Notes 3 | Ch. 2 of Tropp
Week 4	Poincaré Inequalities	Lecture Notes 4 | Ch. 3 of Tropp
Week 5	Entropy Bounds	Lecture Notes 5 | Ch. 5 of Tropp
Week 6	Log-Sobolev Inequalities	Lecture Notes 6 | Ch. 6 of Tropp
Week 7	Metric Space Concentration	Lecture Notes 7 | Ch. 5 of Vershynin
Week 8	Vector Concentration	Lecture Notes 8 | Ch. 3 of Vershynin
Week 9	Matrix Concentration	Lecture Notes 9 | Ch. 5 of Vershynin, Ch. 8 of Tropp
Week 10	Concentration Toolkit	Lecture Notes 10 | Ch. 6 of Vershynin
Week 11	Anti-Concentration[2]	Slides by Vershynin

[2] This is an optional lecture!