Probabilistic Methods for Large Scale Signal Processing and Learning (Syllabus)

The analysis of large-scale signal processing, statistics, and machine learning algorithms often entails analyzing probability distributions over high-dimensional spaces. The randomness may come from stochastic modeling assumptions but can also arise from using randomized algorithms. This course is about building some of the analytical toolkit needed to understand these phenomena. A key component will be understanding the concentration of measure phenomenon and applying it to random vectors and random matrices.

Note: this class has been approved as a regular offering under course number 16:332:531

1 Administration

The instructor for this class is:

Prof. Anand Sarwate
CoRE 517
anand.sarwate@rutgers.edu
https://adsarwate.github.io/

The best method of contact is email or in person before or after class.

1.1 Prerequisites

There are two major prerequisites for the course:

- ECE 541 (Stochastic Signals and Systems)
- Linear Algebra, preferably at the level of 16:642:550 (Linear Algebra and Applications)

This is essentially a mathematics/probability course and we will be looking at theorems and proofs. Registration is restricted by “special permission number” (SPN).

Students who are not in ECE are welcome to attend assuming they have the right background in probability and linear algebra.

1.2 Statement on inclusiveness

We must treat every individual with respect. We are diverse in many ways, and this diversity is fundamental to building and maintaining an equitable and inclusive campus community. Diversity can refer to multiple ways that we identify ourselves, including but not limited to race, color, national origin, language, sex, disability, age, sexual orientation, gender identity, religion, creed, ancestry, belief, veteran status, or genetic information. Each of these diverse identities, along with many others not mentioned here, shape the perspectives our students, faculty, and staff bring to our campus.

In engineering, accounting for and incorporating these diverse perspectives fuels excellence and innovation. In turn, our goal is to develop the tools to let us build technologies to support and sustain a more just and equitable future for all. Accommodations

Rutgers University welcomes students with disabilities into all of the University’s educational programs. In order to receive consideration for reasonable accommodations, a student with a disability should contact the Office of Disability Services (ODS, `https://ods.rutgers.edu/`), participate in an intake interview, and provide documentation. ODS may then provide you with a Letter of Accommodations. Share this letter as soon as you can with the instructors for your classes and discuss the accommodations if needed.

2 Course material

2.1 Learning outcomes

In this course, students should develop knowledge and skills that are integral to reading and writing theoretical and algorithmic research that uses probability theory. In particular, by the end of this course, students should be able

- to understand theoretical papers which use the mathematical tools and techniques we will learn and discuss,
- to apply key results and mathematical tools to analyze algorithms and experiments in a variety of applications,
- to write a clear and rigorous mathematical proof using these tools (as part of the exercises),
- to present/teach the key mathematical ideas and arguments used in a research paper to their peers,
- to work effectively with peers from different labs/research groups by learning a common mathematical "language", and
- to practice methods for investigating and explaining to themselves and others advanced mathematical tools which they may find useful in research.

2.2 Textbook

The primary textbook for the course is:

- Roman Vershynin, High Dimensional Probability: An Introduction with Applications in Data Science , Cambridge University Press, 2018. [8]

A preprint verson is available for free online, linked to the title above.

Some other resources which we may use include:

- Boucheron, Lugosi, and Massart, Concentration Inequalities: A Nonasymptotic Theory of Independence, Oxford University Press, 2013. [3]
- Devdatt Dubhashi and Alessandro Panconesi, Concentration of Measure for the Analysis of Randomised Algorithms , Cambridge University Press, 2009. [4]
- Martin J. Wainwright, High-Dimensional Statistics: A Non-Asymptotic Viewpoint, Cambridge University Press, 2019. [5]
- Joel A. Tropp, An Introduction to Matrix Concentration Inequalities , Foundations and Trends in Machine Learning, Vol. 8: No. 1-2, pp 1-230. [6]
- Keith Ball, An Elementary Introduction to Modern Convex Geometry, MSRI Publications,Volume 31, 1997. [2]
- Ramon Van Handel, Probability in High Dimension, Lecture notes for Princeton course APC 550. [7]
- Horn and Johnson, Matrix Analysis, Cambridge University Prress, 1985. [1]

We will use the following abbreviations for references.

- HDP = Vershynin [8]
- HDS - Wainwright [5]
- CMARA - Dubhashi and Panconesi [4]
- MCI - Tropp [6]
- MCG - Ball [2]
- HJ - Horn and Johnson [1]

2.3 Schedule of topics

This is a tentative list of topics to be covered in the semester.

Week	Topics	Reading
1	Introduction to the course	
	Review of classical limit theorems and inequalities	
	Concentration for Bernoulli random variables	HDP 1, 2.1
2	Chernoff and Hoeffding inequalities	
	Applications to random graphs and network analysis	HDP 1, 2.1
3	Subgaussian distributions	
	Subgaussian norm and properties	HDP 2.2-2.4
4	Subexponential distributions	
	Bernstein's inequality	HDP 2.2-2.4
5	Applications to random projections for dimension reduction	
	Geometry of high dimensional vectors	HDP 2.2-2.4
6	High dimensional spheres and spherical caps	
	Applications in communications	MCG
7	Packings and coverings	
	Applications in network community detection	HDP 4.1-4.5, HJ Chapter 5
8	Concentration for Lipschitz functions	
	Applications to covariance estimation	HDP 5.1, 5.4, 5.6
9	Introduction to martingales and Hoeffding-Azuma	
	McDiarmid's inequality	supplementary notes
10	Matrix Bernstein	
	Matrix completion	HDP 6.1-6.6
11	Random processes and iterative algorithms	
	Stochastic approximation	HDP 7.1-7.4
12	Introduction to empirical process theory	
	Glivenko-Cantellli and DKW	HDP 8.1-8.2
13	Final project presentations or sparse recovery	HDP 10
14	Final project presentations	

3 Assessment

This is a graduate course aimed at PhD students. We expect students to engage actively and meaningfully with the material.

3.1 Components of the grade

This course will be graded on:

- 30% participation and activity in class during solution presentations
- 40% written solutions to exercises
- 20% final project presentation
- 10% final project report

The goal of this approach (homework exercises and projects) is to teach students how to collaborate with peers on technical/mathematical work, how to present logical/mathematical arguments in a clear manner that conveys also the intuition behind the solution, how to read and write theoretical results at a level suitable for publication, and how to independently investigate new mathematical tools on their own.

3.2 Homework and problem presentations

Each week will have a set of assigned problem and students in the course will be assigned in pairs to solve the problems, with one student designated as "presenter." On Monday classes there will be a lecture. On Tuesday students will present their assigned problem to the rest of the class, describing the problem, why it is interesting or important, and the solutions.

In the following week, the pairs will have to write up, in LaTeX, the solutions/proofs for their problems using the template provided in class. The instructor will mark corrections and issues to fix and return them. Students will then have an opportunity to correct or improve their writing. This method is akin to "mastery-based grading" in which students will work on the exercise until it is correct.

3.3 Final project

The goal of the course project is to get students to: explore how the tools we are learning get used in contemporary research learn new tools that we didn't have time to cover get practice in presenting research teach each other about new topics

To that end, here are some guidelines:

- you can work alone or with a partner
- pick a topic or set of papers which uses high dimensional probability, concentration of measure, or other tools
- develop a set of lecture notes/educational material to explain this (in LaTeX) – think maybe 5-10 pages in our template
- make a slide presentation of ~ 20 minutes to be presented at the end of class

One way to do things is to describe a problem/application and then introduce the relevant tool/technical result and how it gets used (or the same in reverse).

The deadline for a project proposal will the end of spring break. The presentations will be in the last two weeks of classes and the report will be due in the final exam period.

3.4 Academic integrity

Please be sure to follow the Academic Integrity guidelines. These are very important, especially for the project: be sure to cite your sources and do not use any figures or other materials without explicit written permission. Plagiarism will be grounds for failing the course.

Students should familiarize themselves with the Academic Integrity Policy, available online (http://nbacademicintegrity.rutgers.edu
Quoting from these guidelines:

The principles of academic integrity require that a student:

- make sure that all work submitted in a course, academic research, or other activity is the student's own and created without the aid of impermissible technologies, materials, or collaborations.
- properly acknowledge and cite all use of the ideas, results, images, or words of others.
- properly acknowledge all contributors to a given piece of work.
- obtain all data or results by ethical means and report them accurately without suppressing any results inconsistent with the student's interpretation or conclusions.
- treat all other students ethically, respecting their integrity and right to pursue their educational goals without interference. This principle requires that a student neither facilitate academic dishonesty by others nor obstruct their academic progress.
- uphold the ethical standards and professional code of conduct in the field for which the student is preparing.

Adherence to these principles is necessary to ensure that:

- proper credit for ideas, words, images, results, and other scholarly work, no matter the form or media, is attributed to the appropriate individual(s).
- all student research and work are fairly evaluated, and no student has an inappropriate advantage over others.
- the academic and ethical development of all students is fostered.
- the reputation of the University for integrity, ethics, scholarship, and professionalism is maintained and enhanced.

Any violations to this policy will be reported to Office of Student Conduct (New Brunswick). Violations of academic integrity will be treated in accordance with university policy, and sanctions for violations may range from no credit for the assignment, to a failing course grade to (for the most severe violations) dismissal from the university.

References

[1] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.

[2] K. Ball. "An elementary introduction to modern convex geometry". In: *MSRI Publications: Flavors of Geometry* 31.1–58 (1997), p. 26.

[3] S. Boucheron, G. Lugosi, and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford, UK: Oxford University Press, 2013.

[4] D. Dubhashi and A. Panconesi. *Concentration of Measure for the Analysis of Randomised Algorithms*. Cambridge, UK: Cambridge University Press, 2009.

[5] M. J. Wainwright. *High-Dimensioinal Statistics: A Non-Asymptotic Viewpoint*. Cambridge, UK: Cambridge University Press, 2019.

[6] J. A. Tropp et al. "An introduction to matrix concentration inequalities". In: *Foundations and Trends® in Machine Learning* 8.1-2 (2015), pp. 1–230.

[7] R. Van Handel. "Probability in High Dimension". Lecture notes for Princeton course APC 550. 2021.

[8] R. Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge, UK: Cambridge University Press, 2018.