Seminar in Advance Application of Statistics: Random Matrix Theory and Applications (STAT 9910-303, Spring 2023)

1 Basic course information

Instructor: Jiaoyang Huang
https://statistics.wharton.upenn.edu/profile/huangjy
E-mail: huangjy@wharton.upenn.edu
Lecture: Thursday 3:30-5:00pm.
Office hours: By appointment, location: Academic Research Building 309.

Since the pioneering work of Wigner, the tools created in random matrix theory have found applications in many other disciplines of science and engineering. This Ph.D. topic course aims to cover important aspects of random matrix theory, and its applications in high dimensional statistics and machine learning. Our main goal is to expose students to new tools and techniques in random matrix theory and applications in statistics, and provide motivated students with the core toolkits for working on related research.

Prerequisites: This course is intended for students with strong mathematical background. There are no formal prerequisites, but students are expected to be familiar with linear algebra, and probability tools such as concentration inequalities.

2 Course descriptions

In this class, we will cover the following four aspects of random matrix theory:

Universality of eigenvalue distributions: Random matrices have been introduced by Physicist Wigner in 1950's to model the Hamiltonian of heavy atoms. The success of random matrices in modeling physical systems lies in the universality phenomenon of their eigenvalue and eigenvector statistics. Under mild conditions, the empirical eigenvalue distributions of Wigner matrices and sample covariance matrices converge to the *Semi-circle distribution* and *Marchenko-Pastur distribution*. We will discuss different approaches to prove this universality phenomenon.

Non-asymptotic analysis of random matrices: Sample covariance matrices play a central role in statistics. For high-dimensional data, when the dimension is comparable to the sample size, it falls into the regime of random matrix theory. We will discuss tools for the analysis of the extreme singular values of random matrices with possibly dependent rows or columns. Several applications will be covered, particularly for the problem of estimating covariance matrices in statistics and for probabilistic constructions of measurement matrices in compressed sensing.

Signal detection in spiked random matrix models: Many statistical models are in the form of "signal"+"noise", and the signal-to-noise ratio measures the size of noise. For the spiked random matrix model, the signal is a deterministic low rank matrix and noise is some random matrix, and the goal is to detect/recover the signal from the noisy observation. There is a critical signal-to-noise ratio. If the signal-to-noise ratio is below this threshold, it is information theoretically impossible to detect the signal; If the signal-to-noise ratio is above this threshold, simple algorithms can be used to detect the signal. We will discuss this phenomenon for spiked Wigner matrices and sample covariance matrices.

Deep neural networks: Deep neural networks are built out of matrices and non-linear activation functions. The weights of deep neural networks are initialized to random numbers, and then trained using stochastic gradient descent. The random initialized deep neural networks are compositions of random matrices and nonlinear activation functions. We will discuss the properties of random initialized deep neural networks, and the neural tangent kernel associated with deep neural networks. When the neural network is wide enough, starting from the random initialization, the training dynamics of neural networks are equivalent to kernel regression using the neural tangent kernel.

Besides the above materials, we will also have several guest lectures on the applications of random matrix theory to high dimensional statistics, computer science and biostatistics.

Textbooks and References

1. Terence Tao, "Topics in random matrix theory" `https://terrytao.files.wordpress.com/2011/02/matrix-book.pdf`
2. Roman Vershynin, "Introduction to the non-asymptotic analysis of random matrices" `https://arxiv.org/pdf/1011.3027.pdf`
3. Martin J. Wainwright, "High-Dimensional Statistics A Non-Asymptotic Viewpoint", online version available from Penn Library.
4. Jinho Baik, Gerard Ben Arous, Sandrine Peche, "Phase transition of the largest eigenvalue for non-null complex sample covariance matrices" `https://arxiv.org/abs/math/0403022`
5. Arthur Jacot, Franck Gabriel, Clment Hongler, "Neural Tangent Kernel: Convergence and Generalization in Neural Networks" `https://arxiv.org/abs/1806.07572`

3 Grading Policy

There will be optional homework assignments, which will be posted on Canvas. Each enrolled student is expected to scribe the notes for one lecture, which is due in two weeks from the lecture. No mid/final exam.