

High-dimensional Probability with Applications to Big Data Sciences

Lecture Summer Semester 2019/2020

With the fast growth of data sciences, there was a dramatic surge of interest and activity over the past two decades in high-dimensional probability that provides vital methods and tools for a wide range of applications. High-dimensional probability is the area of probability theory that studies random objects in \mathbb{R}^n , where the dimension n can be very large. As classical probabilistic tools are no longer sufficient for most of the modern applications in data sciences, these lectures intend to cover *partially* this gap. The focus of the lectures is the non-asymptotic theory in high-dimensional probability with a view towards modern applications in big data sciences. Here is an *incomplete* list of topics that will be covered:

1. Basic concentration inequalities: Hoeffding, Bernstein, McDiarmid and Khintchine's inequalities.
2. Random vectors in high dimensions
3. Matrix concentration inequalities
4. Sub-Gaussian processes

Applications:

1. Dimension reduction with Johnson-Lindenstrauss lemma
2. Community detection
3. Covariance estimation and clustering
4. Matrix completion

The lectures are self-contained and are open for students who have a good knowledge in linear algebra, measure theory and have succeeded Stochastics I. We will primarily rely on the following textbooks:

1. Roman Vershynin, *High-dimensional probability: An introduction with applications in data science*.
2. Martin J. Wainwright, *High-dimensional statistics: A non-asymptotic viewpoint*.

Time and Place: Thursday, 10 – 12, SR 10, Geb E2 4

For further information, please contact Marwa Banna (banna@math.uni-sb.de, room 223 in building E2 4). See also:

<https://www.math.uni-sb.de/ag/speicher/lehre.html>