Random matrices and covariance estimation

Kaveh S. Nobari

Lectures in High-Dimensional Statistics

Department of Mathematics and Statistics
Lancaster University

Contents

Motivation

The issue of covariance estimation is intertwined with random matrix theory, since sample covariance is a particular type of random matrix. These slides follow the structure of chapter 6 of Wainwright (2019) to shed light on random matrices in a non-asymptotic setting, with the aim of obtaining explicit deviation inequalities that hold for all sample sizes and matrix dimensions.

In the classical framework of covariance matrix estimation the sample size n tends to infinity while the matrix dimension d is fixed; in this setting the behaviour of sample covariance matrix is characterized by the usual limit theory. In contrast, in high-dimensional settings the data dimension is either comparable to the sample size ($d \asymp n$) or possibly much larger than the sample size $d \gg n$.

Motivation

We begin with the simplest case, namely ensembles of Gaussian random matrices, and we then discuss more general sub-Gaussian ensembles, before moving to milder tail conditions.

Consider a rectangular matrix $A \in \mathbb{R}^{n \times m}$ with $n \geq m$, the ordered singular values are written as follows

$$\sigma_{\max}(A) = \sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_m(A) = \sigma_{\min}(A) \geq 0$$

References

Wainwright, M. J. (2019). *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press.