Data Analysis and Matrix Computations Assignment 3

Exercise 1.

Prove the Bernstein's inequality: Let $X_1, \ldots, X_N$ be independent, mean zero, sub-exponential random variables. Then, for every $t \geq 0$ we have

$$\mathbb{P}\left\{\left|\sum_{i=1}^{N} X_i\right| \geqslant t\right\} \leqslant 2\exp\left[-c\min\left(\frac{t^2}{\sum_{i=1}^{N}\|X_i\|_{\psi_1}^2}, \frac{t}{\max_i \|X_i\|_{\psi_1}}\right)\right].$$

Exercise 2.

Complete the proof of Theorem 10 (Johnson-Lindenstrauss Lemma) of Lecture 11.

Exercise 3.

Prove the matrix Bernstein's inequality: Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N$ be independent, mean zero, $n \times n$ symmetric random matrices, such that

$$\|\boldsymbol{X}_i\|_2 \leq K$$

almost surely for all i. Then, for every $t \geq 0$ we have

$$\mathbb{P}\left\{\left\|\sum_{i=1}^{N} \boldsymbol{X}_i\right\|_2 \geqslant t\right\} \leqslant 2n \cdot \exp\left(-\frac{t^2/2}{\sigma^2 + Kt/3}\right).$$

Here

$$\sigma^2 = \left\|\sum_{i=1}^{N} \mathbb{E}\boldsymbol{X}_i^2\right\|_2$$

is the norm of the "matrix variance" of the sum.