Exercise 3 – Randomized matrix computations, Fall'24

Prof. D. Kressner
H. Lam

1 ► Operations on sub-Gaussian variables

Suppose that X_1 and X_2 are zero-mean and sub-Gaussian with parameters σ_1 and σ_2, respectively.

(a) If X_1 and X_2 are independent, show that the random variable $X_1 + X_2$ is sub-Gaussian with parameter $\sqrt{\sigma_1^2 + \sigma_2^2}$.

(b) Show that, in general (without assuming independence), the random variable $X_1 + X_2$ is sub-Gaussian with parameter at most $\sqrt{2}\sqrt{\sigma_1^2 + \sigma_2^2}$.

(c) In the same setting as part (b), show that $X_1 + X_2$ is sub-Gaussian with parameter at most $\sigma_1 + \sigma_2$.

(d) If X_1 and X_2 are independent, show that $X_1 X_2$ is sub-exponential with parameters $(v, b) = (\sqrt{2}\sigma_1\sigma_2, \sqrt{2}\sigma_1\sigma_2)$.

2 ► Quadratic Forms of a Gaussian random vector

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix, and let $X \sim N(0, I_n)$. Prove that $X^T A X$ is a sub-exponential random variable.

3 ► Application of Trace Estimation

As discussed in the lecture, the Girard-Hutchinson estimator is a Monte Carlo method used to estimate the trace of a symmetric matrix. A significant advantage of this estimator is that it doesn't require explicitly storing the full matrix A; instead, only the action $x \to Ax$ needs to be provided. In this exercise we will estimate $\log(\det(A))$ via trace estimation. This quantity plays a significant role in Bayesian optimal experimental design.

(a) Let $A \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix. Prove that

$$\log(\det(A)) = \operatorname{trace}(\log(A)),$$

where $\log(A)$ is a matrix X such that $\exp(X) = A$, with

$$\exp(X) = \sum_{k=0}^{\infty} \frac{1}{k!} X^k.$$

(b) Implement the Girard-Hutchinson estimator to estimate $\log(\det(A))$ with A is the matrix from the file `thermomech_TC.mat` in Moodle, using 10, 20, ..., 100 Gaussian sample vectors. Repeat the experiments and plot the average relative error vs sample size.

Hint: Study and modify the `main.m`, the action $x \to \log(A)x$ can be estimated using the Lanczos method, as implemented in the script `lanczos.m`. All files are available on Moodle.

4 ► Concentration around medians

Given a real random variable X, suppose that there are positive constants c_1, c_2 such that

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq c_1 e^{-c_2 t^2} \quad \text{for all } t \geq 0. \tag{1}$$

(a) Prove that $\operatorname{Var}(X) \leq \frac{c_1}{c_2}$.

(b) Show that whenever the mean concentration bound (1) holds, then for any median m_X i.e. any number such that $\mathbb{P}(X \geq m_X) \geq 1/2$ and $\mathbb{P}(X \leq m_X) \geq 1/2$, we have

$$\mathbb{P}(|X - m_X| \geq t) \leq c_3 e^{-c_4 t^2} \quad \text{for all } t \geq 0, \tag{2}$$

where $c_3 := 4c_1$ and $c_4 := \frac{c_2}{8}$.

(c) Conversely, show that whenever the median concentration bound (2) holds, then mean concentration (1) holds with $c_1 = 2c_3$ and $c_2 = \frac{c_4}{4}$.

5 ► Girard-Hutchinson trace estimation with Rademacher vector

Suppose we estimate the trace of a symmetric positive semidefinite matrix $A \in \mathbb{R}^{n\times n}$ using Girard-Hutchinson trace estimator with N independent Rademacher vectors.

(a) Show that Hoeffding's inequality leads to

$$\Pr\{|\mathrm{trace}_N(A) - \mathrm{trace}(A)| \geq \epsilon\} \leq 2\exp\Big(-\frac{\epsilon^2 N}{2\|A\|_2^2 n^2}\Big).$$

(b) Show that Bernstein's inequality leads to

$$\Pr\{|\mathrm{trace}_N(A) - \mathrm{trace}(A)| \geq \epsilon\} \leq 2\exp\Big(-\frac{\epsilon^2 N}{4\|A\|_F^2 + 4/3 \cdot \epsilon n \|A\|_2}\Big).$$