EXERCISE 4 – Randomized matrix computations, Fall'24

Prof. D. Kressner
H. Lam

1 ▶ Operator monotonicity

a) Show that the logarithm is operator monotone by the integral representation:

$$\log(a) = \int_0^\infty [(1+x)^{-1} - (a+x)^{-1}]\mathrm{d}x, \qquad \text{for } a > 0.$$

b) Let $f : \mathbb{R}^{n\times n} \to \mathbb{R}$ be defined as $f(B) = \operatorname{trace}(\exp(B))$. Prove that f is operator monotone, i.e., if $A, B \in \mathbb{R}^{n\times n}$ and $B - A \in \mathbb{R}^{n\times n}$ are symmetric positive semidefinite matrics, then $f(A) \leq f(B)$.

2 ▶ Largest eigenvalue of sum of symmetric independent matrices

Let $X_1, \dots, X_s$ be symmetric independent random matrices. Prove that

$$\mathbb{E}\Big[\lambda_{\max}\Big(\sum_{i=1}^s X_i\Big)\Big] \leq \inf_{\theta>0} \frac{1}{\theta} \log \operatorname{trace}\Big(\exp\Big(\sum_{i=1}^s \log\left(\mathbb{E}e^{\theta X_i}\right)\Big)\Big).$$

Hint: combine the proof and the Lemma from Lecture 4 slide 6.

3 ▶ Matrix sparsification

Let $A \in \mathbb{R}^{m\times n}$ be a matrix. In this question, we design and analyze a sampling approach for approximating A by a sparse matrix.

a) Express A as a sum of mn matrices, each with at most one nonzero entry.

b) Show how to construct an unbiased estimator X of A by uniform sampling.

c) Define

$$\tilde{X}_s = \frac{1}{s}\sum_{k=1}^s X_i \quad \text{where each } X_i \text{ is an independent copy of } X.$$

For $\epsilon \in [0, 1]$, using the Matrix Bernstein inequality to give an upper bound on the number of s needed to obtain

$$\mathbb{E}\left[\|\tilde{X}_s - A\|_2\right] \leq 2\epsilon\|A\|_2.$$

d) Define the probability mass

$$p_{ij} = \frac{1}{2}\left[\frac{|a_{ij}|^2}{\|A\|_F^2} + \frac{|a_{ij}|}{\|A\|_{\ell_1}}\right] \quad \text{for} \quad i = 1, \dots, m \quad \text{and} \quad j = 1, \dots, n.$$

Here, $\|\cdot\|_{\ell_1}$ is the entrywise ℓ_1 norm, i.e $\|A\|_{\ell_1} := \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|$. Using the Matrix Bernstein inequality to provide a bound on the number s of samples needed to achieve $\mathbb{E}\left[\|\tilde{X}_s - A\|_2\right] \leq 2\epsilon\|A\|_2$. Express the result in terms of the stable rank of A.

e) Implement both procedures and apply them to the RBF kernel matrix, i.e., for $h > 0$,

$$a_{ij} = \exp(-\|x_i - x_j\|_2^2/(2h)) \quad \text{for } i, j = 1, \dots, n$$

associated with randomly generated $x_1, \dots, x_n$, uniformly drawn from the unit cube $[0, 1]^d$. Plot the sampling distribution of the spectral norm error as a function of the number s of samples.