HOMEWORK 1 – Randomized matrix computations, Fall'24

Prof. D. Kressner
H. Lam

Sciper: **Name:**

Please read the following instructions carefully

- This homework consists of 3 questions, and the submission deadline is October 18 at 23h59.
- You can either typeset or scan your handwritten solution. Please also include this page with your signature.
- Submit your homework by email to hysan.lam@epfl.ch.
- **The submitted homework must be your own personal work and must not be copied from elsewhere.**

Your signature:

1 ► Power method

The power method is a simple algorithm used to approximate the largest (in magnitude) eigenvalue of a matrix:

Algorithm 1 Power method

```
Input: Matrix A ∈ ℝ^{n×n}.
Output: Approximate eigenpair (μ^(k), x^(k)) of matrix A.
Choose starting vector x^(0) ∈ ℝ^n.
k = 0.
repeat
    Set k := k + 1.
    Compute y^(k) := A x^(k−1).
    Normalize x^(k) := y^(k) / ‖y^(k)‖_2.
    μ_k := (x^(k))^⊤ A x^(k).
until convergence is detected
```

a) Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive semi-definite matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ and an associate orthonormal basis of eigenvectors $u_1, \ldots, u_n$. Suppose that k steps of the power method with starting vector $x^{(0)} \in \mathbb{R}^n$, $\|x^{(0)}\|_2 = 1$, are carried out, leading to the approximation μ_k of λ_1. Prove that

$$\lambda_1 \geq \mu_k \geq \lambda_1 \langle x^{(0)}, u_1 \rangle^{1/k}.$$

Hint: Express μ_k in terms of $x^{(0)}$ and apply Jensen's inequality.

b) Suppose $n = 14400$ and we use the random starting vector $x^{(0)} = z/\|z\|_2$ where $z \sim N(0, I_n)$. Using a result from the lecture, how many iterations are required to guarantee that

$$\lambda_1 \geq \mu_k \geq \frac{1}{2}\lambda_1$$

holds with probability at least 99.99%?

2 ► Outer product of Gaussian random vectors

a) Let x be a chi-square random variable with k degrees of freedom, i.e., $x = \sum_{i=1}^{k} z_i^2$ where z_i are independent standard Gaussian random variables. Using the Chernoff bound, prove that for $\gamma > 1$ and $t/k < 1/2$, it holds that

$$\mathbb{P}\{x > k\gamma\} \leq (1 - 2t/k)^{-k/2} e^{-t\gamma}.$$

Furthermore, prove that

$$\mathbb{P}\{x > k\gamma\} \leq (\gamma e^{1-\gamma})^{\frac{k}{2}}.$$

b) Let $x \sim N(0, I_n)$ and $y \sim N(0, I_n)$ be independent standard Gaussian random vectors. Using part a), prove that the Frobenius norm of the outer product of x and y is not too large: For $\gamma > 1$, it holds that

$$\mathbb{P}\{\|xy^\top\|_F > n\gamma\} \leq 2\exp(-n(\gamma - \ln\gamma - 1)/2).$$

Bonus question: Try to improve the bound in (b).

3 ► Method of characteristic function

An important tool in probability theory is the characteristic function. For a random variable z, the characteristic function is defined as

$$\varphi_z(t) := \mathbb{E}[\exp(itz)], \quad i := \sqrt{-1}.$$

The characteristic function characterizes the probability distribution. A way to see this property is through the Lévy's theorem (you can use it without proof):

Theorem 1 (Lévy) *For a random variable z, let $\varphi_z(t)$ denote its characteristic function and F_z denote its cumulative distribution function. Assume that $\varphi_z(t)$ and the density of z are integrable, then for $h > 0$, we have*

$$F(x+h) - F(x-h) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin(ht)}{t} e^{-ixt}\varphi_z(t)\mathrm{d}t.$$

Using this result, the following problems establish relations between the bilinear random form $x^\top Ay$ and the Frobenius norm of a general (square) matrix A.

a) Suppose that $x_1, \ldots, x_n$ are independent real random variables (over the same probability space). Given (fixed) real numbers $\alpha_1, \ldots, \alpha_n$, prove that

$$\varphi_{(\alpha_1 x_1 + \ldots + \alpha_n x_n)}(t) = \prod_{j=1}^{n} \varphi_{x_j}(a_j t_j).$$

b) Let $A \in \mathbb{R}^{n\times n}$ and consider independent random Gaussian vectors $x \sim N(0, I_n)$ and $y \sim N(0, I_n)$. Show that

$$\mathbb{E}[(x^\top Ay)^2] = \|A\|_F^2.$$

c) Let $A \in \mathbb{R}^{n\times n}$ and consider independent random Gaussian vectors $x \sim N(0, I_n)$ and $y \sim N(0, I_n)$. Show that

$$\mathbb{P}\{\|A\|_F > \gamma \cdot x^\top Ay\} = \Pr\left\{(x^\top Ay)^2 < \frac{\|A\|_F^2}{\gamma^2}\right\} \le \frac{2}{\pi}\int_0^{\infty} \frac{|\sin(t/\gamma)|}{t}\frac{1}{\sqrt{1+t^2}}\mathrm{d}t$$

holds for $\gamma > 1$. Furthermore, using the bound

$$|\sin(t)| \le \begin{cases} t \text{ for } t \in [0,1], \\ 1 \text{ elsewhere,} \end{cases}$$

conclude that

$$\mathbb{P}\{\|A\|_F > \gamma \cdot x^\top Ay\} \le \frac{2}{\pi}\gamma^{-1}(2 + \ln(1+2\gamma)). \tag{1}$$

Hint: The characteristic function of the product of two independent standard Gaussian random variables is $1/\sqrt{1+t^2}$.

Bonus: Explore the tightness of the bound (1) numerically for different A (with small and large stable ranks).