CIS 5020 Analysis of Algorithms Fall 2024

Assignment 3

Due: Oct 30

Name: Yujie Xu, Richard Zhang

PennID: 51233809, 19331985

1 Question 1

1. Let's first assume all the eigenvalues of A are non-negative. Because we know A is symmetric, we know it can be diagonalized to $A = U\Lambda U^\top$ where U is an orthogonal matrix and Λ is a diagonal matrix which contains eigenvalues $\lambda_i \geq 0$. We also know $U^\top U = I$.

   Then, for any vector $x \in \mathbb{R}^n$, we have:

   $$x^\top Ax = x^\top U\Lambda U^\top x = (U^\top x)^\top \Lambda(U^\top x)$$

   We set $y = Q^\top x$, and the equation would become $(U^\top x)^\top \Lambda(U^\top x) = y^\top \Lambda y$

   Because we know Λ is a diagonal matrix which contains eigenvalues $\lambda_i \geq 0$, we know all of its entries are non-negative. Thus, we can deduce:

   $$x^\top Ax = y^\top \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2 \geq 0$$

   $$x^\top Ax \geq 0$$

   Proof by contradiction: if A has a negative eigenvalue $\lambda_x < 0$, and its eigenvector is $v_x \neq 0$, we have:

   $$v_x^\top Av_x = v_x^\top(\lambda_x v_x) = \lambda_x v_x^\top v_x = \lambda_x \|v_x\|^2 < 0$$

   This contradicts with the conclusion $v^\top Av \geq 0$, so all the eigenvalues of A must be non-negative.

2. Suppose we have a random vector $x \in \mathbf{R}^m$, we want to put it in the matrix $BAB^\top$:

   $$x^\top(BAB^\top)x = (B^\top x)^\top A(B^\top x)$$

   Suppose we have another vector $y = B^\top x \in \mathbf{R}^n$, because we know A is psd, $y^\top Ay \geq 0$. Thus, we have:

   $$x^\top(BAB^\top)x = y^\top Ay \geq 0$$

   $$x^\top(BAB^\top)x \geq 0$$

   This holds for all $x \in \mathbb{R}^m$. Therefore, $BAB^\top$ is psd.

2 Question 2

$$A = \sum_{i=1}^{n} \sigma_i \, \mathbf{u}_i \, \mathbf{v}_i^\top$$

Alternatively, we can rewrite in using the following form:

$$A = U \, \Sigma \, V^\top$$

where U is an orthogonal matrix made up of $[u_1, u_2, ..., u_n]$, V is an orthogonal matrix made up of $[v_1, v_2, ..., v_n]$, and Σ is a diagonal matrix made up of $[\sigma_1, \sigma_2, ..., \sigma_n]$. Because all $\sigma_i \neq 0$, Σ is invertible with its inverse:

$$\Sigma^{-1} = \mathrm{diag}(\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_n^{-1})$$

The inverse of A is then $A^{-1} = V \, \Sigma^{-1} \, U^\top$. We also know $\Sigma^{-1} = \sum_{i=1}^{n} \sigma_i^{-1} \, e_i \, e_i^\top$ where e_i are the standard basis vectors ($V e_i = v_i$, and $(U e_i)^\top = u_i^\top$). If we expand it, we would get:

$$\begin{aligned} A^{-1} &= V \, \Sigma^{-1} \, U^\top \\ &= V (\sum_{i=1}^{n} \sigma_i^{-1} \, e_i \, e_i^\top) U^\top \\ &= \sum_{i=1}^{n} \sigma_i^{-1} \, V \, e_i \, e_i^\top \, U^\top \\ &= \sum_{i=1}^{n} \sigma_i^{-1} \, v_i \, u_i^\top \end{aligned}$$

Therefore, the inverse of A is given by $\sum_{i=1}^{n} \sigma_i^{-1} \, v_i \, u_i^\top$.

3 Question 3

1. Let's first assume all the eigenvalues of A are non-negative. Given that A is symmetric, we know it can be diagonalized to $A = U \Lambda U^\top$ where U is an orthogonal matrix and Λ is a diagonal matrix which contains eigenvalues $\lambda_i \geq 0$. We also know $U^\top U = I$.

   $$A^k = (U \Lambda U^\top)^k = U \Lambda U^\top \cdots U \Lambda U^\top = U \Lambda^k U^\top$$

   Λ^k is a diagonal matrix with entries $\lambda_1^k \ldots \lambda_n^k$

   The trace of a matrix is invariant under orthogonal matrix.

   $$\mathrm{Tr}(A^k) = \mathrm{Tr}(U \Lambda^k U^\top) = \mathrm{Tr}(\Lambda^k)$$

   Therefore $\mathrm{Tr}(A^k) = \sum_i^n \lambda_i^k$

2. We first deduce the expression of $Tr(A^4)$ in terms of the entries of A:

$$\text{Tr}(A^4) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l=1}^{n}\sum_{l=1}^{n} A_{ij}A_{jk}A_{kl}A_{li}$$

The index is a close loop, since the trace is the sum of the diagonal position.

Based on the linearity of expectation:

$$\mathbb{E}[\text{Tr}(A^4)] = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} \mathbb{E}[A_{ij}A_{jk}A_{kl}A_{li}]$$

Note that A is symmetric and the number on each entry is randomly chosen and has an expectation of 0. Therefore the expectation of each entry is non-zero when every entry appears an even number of times.

There are 3 scenarios:

1. all the 4 indices are the same, i.e $i = j = k = l$. The expectation of this term is:

$$\sum_{i=1}^{n} \mathbb{E}[A_{ii}^4] = n$$

2. 3 indices are the same, e.g. $j = k = l$. This pattern appears 4 times. The expectation of this term is:

$$4\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} \mathbb{E}[A_{jj}^2 A_{ij}^2] = 4n^2 - 4n$$

3. 2 indices are the same, i.e. $i = k, j = l$ and $i = j, k = l$ and $i = l, j = k$. This pattern appears 3 times. The expectation of this term is:

$$\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} \mathbb{E}[A_{ij}^4] + 2\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} \mathbb{E}[A_{jj}^2 A_{ij}^2] = 3n^2 - 3n$$

Therefore the explicit expression of expectation is:

$$\mathbb{E}[\text{Tr}(A^4)] = 7n^2 - 6n$$

By Markov's inequality, let $Z = \text{Tr}(A^4), t = \frac{n^4}{7n^2-6n}$:

$$\Pr[Z \geq t\mathbb{E}[Z]] \leq \frac{1}{t}$$
$$\Pr[Z \geq n^4] \leq \frac{7n - 6}{n^3}$$

Noted that:

$$Z = \text{Tr}(A^4) = \sum_{i}^{n} \lambda_i^4 \leq n \cdot \max_i |\lambda_i|^4$$

$\max_i |\lambda_i|^4$ is also a random variable, so we can determine its inequality by the above Markov's inequality:

$$\Pr[n \cdot \max_i |\lambda_i|^4 \geq n^4] \leq \frac{7n-6}{n^3}$$
$$\Pr[\max_i |\lambda_i| \geq n^{\frac{3}{4}}] \leq \frac{7n-6}{n^3}$$

When n is large enough, without the loss of generality, let $n = 100$, $\Pr[\max_i |\lambda_i| \geq n^{\frac{3}{4}}] \leq 0.0007$

Therefore $\max_i |\lambda_i| = O(n^{3/4})$ with probability at least 0.99.