Analysis of Algorithms Assignment 3

Umang Sharma & Amogh Sarangdhar

October 2024

Q1 (i) Suppose A is a symmetric matrix. Then prove that all the eigenvalues of A are non-negative if and only if $x^\top Ax \geq 0$ for every $x \in \mathbb{R}^n$. (Remark: such a matrix A is called a PSD matrix).

Solution

Since A is symmetric, we can use the Spectral theorem,

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^\top$$

where $v_1, ..., v_n$ are orthonormal unit vectors and $Av_i = \lambda_i v_i \forall 1 \leq i \leq n$. This equation can be written in it's matrix form as, x

$$A = Q^\top \Lambda Q$$

where Λ is a diagonal matrix containing the eigenvalues in A.

$$x^\top Ax = x^\top Q^\top \Lambda Qx = Y^\top \Lambda Y$$

This can be re-written as,

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

Now, y_i^2 will always be positive. So, $x^\top Ax \geq 0$ then all the eigenvalues of A will also be greater than or equal to 0 i.e. $\lambda_i \geq 0 \forall i$

Q1 (ii) Suppose $A \in \mathbb{R}^{n\times n}$ is a PSD matrix. Prove that for any matrix $B \in \mathbb{R}^{m\times n}$, $BAB^\top$ is also PSD.

Solution

We know that,

$$A = Q^\top \Lambda Q$$

So,

$$BAB^\top = BQ^\top \Lambda QB^\top$$

Here, $B \in \mathbb{R}^{m\times n}$ and $Q^\top \in \mathbb{R}^{n\times n}$. Let,

$$QB^\top = Y$$

where $Y \in \mathbb{R}^{n\times m}$. Then

$$(QB^\top)^\top = BQ^\top$$

So we can rewrite equation as,

$$BAB^\top = BQ^\top \Lambda QB^\top = Y^\top \Lambda Y$$

Which is again, nothing but the spectral theorem. But, notice that the Diagonal Eigenvalue matrix Λ is the same for matrix A and $BAB^\top$. So, if matrix A is a PSD i.e. all it's eigenvalues are greater than or equal to 0, then matrix $BAB^\top$ is also PSD, because they have the same eigenvalues.

Q2 Suppose A is a square matrix of size $n \times n$ (not necessarily symmetric). Let the SVD of A be given by $\sum_{i=1}^{n} \sigma_i u_i v_i^{\top}$ where $\sigma_i \neq 0$ (for all i). Prove the the inverse of A is given by $\sum_{i=1}^{n} \sigma_i^{-1} v_i u_i \top$. (Recall that the inverse of a square matrix A is the unique matrix B such that $BA = AB = I$).

Solution
We know that, a the inverse of a matrix A is given by A^{-1} such that,

$$AA^{-1} = A^{-1}A = I$$

The SVD of A is given by

$$A = \sum_{i=1}^{n} \sigma_i u_i v_i^{\top}$$

This can be written in the matrix form as,

$$A = U\Sigma V^{\top}$$

where

- $U = [u_1, u_2, ..., u_n]$ is an $n \times n$ orthogonal matrix whose columns u_i are the left singular vectors of A
- $V = [v_1, v_2, ..., v_n]$ is an $n \times n$ orthogonal matrix whose columns v_i are the right singular vectors of A
- $\Sigma = diag(\sigma_1, \sigma_2, ..., \sigma_n)$ s a diagonal matrix containing the non-zero singular values σ_i of A.

Let the inverse of A be denoted by B

$$A^{-1} = B = (U\Sigma V^{\top})^{-1} = V^{-T}\Sigma^{-1}U^{-1} \tag{1}$$

This is because $(XY)^{-1} = Y^{-1}X^{-1}$
Now, V and U are orthogonal, so their inverse is the same as their transpose.

$$U^{-1} = U^{\top} \tag{2}$$
$$V^{-T} = (V^{\top})^{-1} = (V^{-1})^{-1} = V \tag{3}$$

Substituting Equations 2 and 3 back into 1 we get,

$$B = V^{-T}\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^{\top} \tag{4}$$

Which can be written as,

$$B = V\Sigma^{-1}U^{\top} = \sum_{i=1}^{n} \sigma^{-1} v_i u_i^{\top} \tag{5}$$

We can verify that B is the inverse of A by multiplying them and seeing their product,

$$\begin{aligned} AB &= (U\Sigma V^{\top})(V\Sigma^{-1}U^{\top}) \\ &= U\Sigma(V^{\top}V)\Sigma^{-1}U^{\top} \\ &= U(\Sigma\Sigma^{-1})U^{\top} \\ &= UU^{\top} \\ &= I \end{aligned}$$

Thus, we have shown that since multiplying matrix A with B gives us the Identity matrix, B is the inverse of matrix A with its SVD as $V\Sigma^{-1}U^{\top} = \sum_{i=1}^{n} \sigma^{-1} v_i u_i^{\top}$.
Hence Proved.

Q3 (a) Suppose A is symmetric matrix. We know from standard linear algebra that $Tr(A) = \sum_{i=1}^{n} \lambda_i$ where $\lambda_1, ..., \lambda_n$ are the eigenvalues of A. Prove that $Tr(A^k) = \sum_{i=1}^{n} \lambda_i^k$

Solution

To prove that $\text{Tr}(A^k) = \sum_{i=1}^{n} \lambda_i^k$ for a symmetric matrix A, we can utilize the fact that symmetric matrices are diagonalizable.

Since A is symmetric, it can be diagonalized by an orthogonal matrix U:

$$A = UDU^\top$$

where U is an orthogonal matrix ($U^\top U = UU^\top = I$), and D is a diagonal matrix containing the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A.

Now we raise both sides of the diagonalization to the k-th power:

$$A^k = (UDU^\top)^k = UD^kU^\top$$

This is because $UU^\top = I$, and powers distribute over the diagonal matrix D.

Computing the Trace of A^k:

$$\text{Tr}(A^k) = \text{Tr}(UD^kU^\top)$$

The trace of a matrix is invariant under cyclic permutations:

$$\text{Tr}(UD^kU^\top) = \text{Tr}(D^kU^\top U) = \text{Tr}(D^kI) = \text{Tr}(D^k)$$

Since $U^\top U = I$.

$$\text{Tr}(D^k) = \sum_{i=1}^{n} (D^k)_{ii} = \sum_{i=1}^{n} (\lambda_i)^k$$

Because D^k is a diagonal matrix with entries $(\lambda_1)^k, (\lambda_2)^k, \ldots, (\lambda_n)^k$.

Since A is symmetric, it can be diagonalized as $A = UDU^\top$, where U is orthogonal and D is diagonal with eigenvalues λ_i. Then

$$\text{Tr}(A^k) = \text{Tr}(UD^kU^\top) = \text{Tr}(D^kU^\top U) = \text{Tr}(D^kI) = \text{Tr}(D^k) = \sum_{i=1}^{n} \lambda_i^k.$$

Thus, we have proven the statement.

$$\text{Tr}(A^k) = \sum_{i=1}^{n} \lambda_i^k$$

Q3 (b) Consider a random matrix $A \in \mathbb{R}^n$ defined as follows: entry $A(i, j)$ (where $i \leq j$) is set to ± 1 with probability $1/2$. If $i > j$, then $A(i, j)$ is set to be equal to $A(j, i)$. Compute $\mathbb{E}[Tr(A^4)]$ and use that to prove that $max_i |\lambda_i| = O(n^{3/4})$ with probability 0.99.

Solution

To compute $\mathbb{E}[\text{Tr}(A^4)]$ and estimate $\max_i |\lambda_i|$, we'll set $B = A^2$. This approach will help us compute $\text{Tr}(A^4)$ as $\text{Tr}(B^2)$, simplifying our calculations.

Since A is a symmetric matrix, A^2 is also symmetric, and we can set:

$$B = A^2.$$

Our goal is to compute $\mathbb{E}[\mathrm{Tr}(B^2)] = \mathbb{E}[\mathrm{Tr}(A^4)]$.

The entries of B are given by:

$$B_{ij} = (A^2)_{ij} = \sum_{k=1}^{n} A_{ik}A_{kj} = \sum_{k=1}^{n} A_{ik}A_{jk}.$$

Since A is symmetric ($A_{kj} = A_{jk}$), the expression simplifies.

- When $i = j$:

$$B_{ii} = \sum_{k=1}^{n} A_{ik}^2 = \sum_{k=1}^{n} 1 = n.$$

So,

$$\mathbb{E}[B_{ii}] = n.$$

- When $i \neq j$:

$$B_{ij} = \sum_{k=1}^{n} A_{ik}A_{jk}.$$

Since A_{ik} and A_{jk} are independent random variables with mean zero (except when $k = i$ or $k = j$), their product has mean zero. Therefore,

$$\mathbb{E}[B_{ij}] = 0 \quad \text{for } i \neq j.$$

We need $\mathbb{E}[B_{ij}^2]$ to compute $\mathbb{E}[\mathrm{Tr}(B^2)]$.

- When $i = j$:

$$B_{ii} = n \implies B_{ii}^2 = n^2.$$

So,

$$\mathbb{E}[B_{ii}^2] = n^2.$$

- When $i \neq j$:

$$B_{ij} = \sum_{k=1}^{n} A_{ik}A_{jk}.$$

B_{ij} is a sum of n products of random variables. Since each $A_{ik}A_{jk}$ has mean zero and variance 1 (except when $k = i$ or $k = j$, but these contribute negligibly for large n), B_{ij} has mean zero and variance approximately n. Therefore,

$$\mathbb{E}[B_{ij}^2] \approx n \quad \text{for } i \neq j.$$

$$\mathrm{Tr}(B^2) = \sum_{i=1}^{n} B_{ii}^2 + \sum_{i,j=1}^{n} B_{ij}^2.$$

So,

$$\mathbb{E}[\mathrm{Tr}(B^2)] = \sum_{i=1}^{n} \mathbb{E}[B_{ii}^2] + \sum_{i \neq j} \mathbb{E}[B_{ij}^2] = n \cdot n^2 + n(n-1) \cdot n = n^3 + n^3 - n^2 = 2n^3 - n^2.$$

For large n, this simplifies to:

$$\mathbb{E}[\mathrm{Tr}(B^2)] \approx 2n^3.$$

Since $\mathrm{Tr}(A^4) = \mathrm{Tr}(B^2)$, we have:

$$\mathbb{E}[\mathrm{Tr}(A^4)] \approx 2n^3.$$

Because A is symmetric, it has real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, and:

$$\mathrm{Tr}(A^4) = \sum_{i=1}^{n} \lambda_i^4.$$

Therefore,

$$\mathbb{E}\left[\sum_{i=1}^{n} \lambda_i^4\right] \approx 2n^3.$$

We can write:

$$\max_i \lambda_i^4 \leq \sum_{i=1}^{n} \lambda_i^4 = \mathrm{Tr}(A^4).$$

Using Markov's inequality,

$$P\left(\max_i |\lambda_i| \geq t\right) = P\left(\max_i \lambda_i^4 \geq t^4\right) \leq \frac{\mathbb{E}\left[\sum_{i=1}^{n} \lambda_i^4\right]}{t^4} = \frac{2n^3}{t^4}.$$

Set $t = cn^{3/4}$. Then,

$$P\left(\max_i |\lambda_i| \geq cn^{3/4}\right) \leq \frac{2n^3}{(cn^{3/4})^4} = \frac{2n^3}{c^4 n^3} = \frac{2}{c^4}.$$

To have a probability of at least 0.99 that $\max_i |\lambda_i| \leq cn^{3/4}$, we need:

$$\frac{2}{c^4} \leq 0.01 \implies c^4 \geq 200.$$

Thus, $c \geq \sqrt[4]{200} \approx 3.76$.

Hene we proved that with probability at least 0.99, the largest eigenvalue satisfies:

$$\max_i |\lambda_i| = O(n^{3/4}).$$