Work Integrated Learning Programmes Division
Cluster Programme in AIML

Assignment 1

AIMLC ZC416

Mathematical Foundations for Machine Learning

Instructions

1. This is not a group activity. Each student should do the problems and submit individually.

2. Assignments have to be handwritten and uploaded as a single pdf file with name BITSID.pdf

3. Submissions beyond 5th of July, 2023 23.59 hrs would not be graded

4. Assignments sent via email would not be accepted

5. Copying is strictly prohibited. Adoption of unfair means would lead to disciplinary action.

Answer all the questions

Q1) We have a system $L\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{b} \neq \boldsymbol{0}$ where L is a 100000×100000 lower-triangular matrix and $\boldsymbol{b}$ is a 100000×1 vector. The computer system on which we would like to solve this system has two constraints - one, we cannot store the matrix L in memory, but can only store individual rows of the matrix in memory, and two, writes into memory are much more expensive than reads from memory. Can you write an algorithm to solve this system respecting the first constraint and ensuring that you make as few writes to memory as possible. (2)

Q2) Let E_{ij} denote a 50×50 elimination matrix where the entry in the ith row and jth column below the main diagonal is non-zero and all other off-diagonal elements are zero. For $i = 5, j = 2$ find the set of elimination matrices E_{pq} where $E_{pq}E_{ij} = E_{ij}E_{pq}$. Using the action of an elimination matrix when it is applied on another matrix, can you explain why commutativity works for some pairs of elimination matrices and not for others? Give adequate justifications. (2)

Q3) Find a 2×2 matrix A with the following properties: Each eigenvalue is a complex number with modulus equal to 1. The eigenvalues remain complex for A^n for $0 \leq n \leq 49$ and become real for A^{50}. (2)

Q4) A student A is given the task of computing the eigenvalues of a $n \times n$ matrix where n is atleast 10^6. He approaches his friend B for help. B says that the job is an extremely time-consuming one. A then says the matrix is an upper-triangular matrix. At this A says that he can obtain all the eigenvalues of the system by querying only n elements in the matrix. He claims to be able to solve for the eigenvalues of the system in $O(n)$ time. Do you think B's confidence is justified? If so, what is the algorithm that B is going to use? If not, explain why B's confidence is misplaced. (2)

Q5) Consider the following argument: let $C \in \mathbb{R}^{n\times n}$ be such that $C^TC = CC^T$. Using the SVD we can write $C = U\Sigma V^T$, and since $C^TC = CC^T$ we see that $V\Sigma^T U^T U\Sigma V^T = U\Sigma V^T V\Sigma U^T$ and thus $V\Sigma^T\Sigma V^T = U\Sigma\Sigma^T U^T$. Therefore $U = V$ and C is a symmetric matrix. Can you say something on the soundness of this argument with appropriate justifications? (2)