

Indian Institute of Technology Kharagpur
Centre of Excellence in Artificial Intelligence
AI61003 Linear Algebra for AI and ML
Assignment 1, Due on: September 15

ANSWER ALL THE QUESTIONS

1. Let $P_n(\mathbb{R})$ denote the set of all polynomials in indeterminate x with real coefficients.

- (a) Prove that $P_n(\mathbb{R})$ is a real vector space.
- (b) Define a function $\mathcal{F} : P_n(\mathbb{R}) \rightarrow \mathbb{R}$ as

$$\mathcal{F}(p(x)) = \left. \frac{d}{dx} p(x) \right|_{x=0}$$

In other words, the function assigns every polynomial with the value of its derivative at $x = 0$. Prove that this is a linear functional.

- (c) Find an inner product representation for the linear functional in the above question.
2. Let $x \in \mathbb{R}^n$ and $\mathbf{1}_n$ be the n -vector with all entries 1. Let $\text{avg}(x)$ and $\text{std}(x)$ be as defined in the class. Then for any $\alpha, \beta \in \mathbb{R}$ prove the following.
- (a) $\text{avg}(\alpha x + \beta \mathbf{1}_n) = \alpha \text{avg}(x) + \beta$
 - (b) $\text{std}(\alpha x + \beta \mathbf{1}_n) = |\alpha| \text{std}(x)$
3. Let $w \in \mathbb{R}^n$ be a given vector with $w_i > 0$ for $i = 1, 2, \dots, n$. Then for any $x \in \mathbb{R}^n$, define the function

$$\|x\|_w = \sqrt{\sum_{i=1}^n w_i x_i^2}$$

Show that the function $\|\cdot\|_w$ defines a norm called as weighted norm.

4. Discuss in details the existence and uniqueness conditions for system of linear equations where $Ax = b$ where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ are given matrices.
5. Prove that matrix multiplication is associative; however matrix multiplication is not commutative. Further, let $A \in \mathbb{R}^{p \times q}, B \in \mathbb{R}^{q \times r}, C \in \mathbb{R}^{r \times t}$ be given matrices. Derive conditions on p, q, r, t such that $(AB)C$ is computationally more efficient than $A(BC)$.
6. Define left inverse of a matrix A . Determine in which of the following cases left inverse exists. In each of the following cases, is it possible to characterize all the left inverses if one left inverse exists.

$$(a) A = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$(b) A = \begin{bmatrix} 2 & 0 \\ 0 & -2 \\ 3 & 3 \end{bmatrix}$$

7. Let $A \in \mathbb{R}^{n \times n}$ be an invertible matrix. Construct the matrices L_{ij} discussed in the class and write down the procedure to compute LU decomposition of A .
8. Consider k -means clustering algorithm as follows with the standard terminology and notation introduced in the class as follows.
- Input: $x_1, x_2, \dots, x_N \in \mathbb{R}^n$. Initial list of k cluster representatives z_1, \dots, z_k .
- Output: Cluster assignment c_1, c_2, \dots, c_N
- Repeat until convergence
1. Cluster assignment based on cluster representatives.
 2. Update cluster representatives.
- (a) In Step 1, what is the computational complexity?
 - (b) In Step 2, what is the computational complexity?
 - (c) Assuming 10 iterations are performed, how many number of computations are involved to obtain the cluster assignment for the given data points?
9. **Image Clustering:** Consider the MNIST database of handwritten digits. Choose 100 images of each digit from this data set. In the notation of Problem 8Item.15, determine values N and n . Fix a reasonable convergence criterion. Perform the following exercises (a),(b) and (c) in two cases:
 case (i) random initialization of cluster representatives;
 case(ii) choose cluster representatives from the given data set.
- (a) For $k = 20$, run the above algorithm to cluster the given images into 20 clusters. Plot the cluster representatives after the algorithm converges. Count the number of iterations.
 - (b) Choose 50 images (not chosen previously) from the MNIST data set randomly and assign the clusters to these *test* images. What is the accuracy of cluster assignment?
 - (c) For $k = 5$ to $k = 20$, tabulate the values of J^{clust} and discuss what may be the optimal size of number of clusters.
- Does the choice of initial condition have any effect on the performance of k -means clustering algorithm?
10. **Document Clustering:** Collect the set of documents form Wikipedia with the following titles:

- Linear algebra
- Data Science
- Artificial intelligence
- European Central Bank
- Financial technology
- International Monetary Fund
- Basketball
- Swimming
- Cricket

Vectorize the documents with TF-IDF vectorizer (Word Frequency Vectors) and perform the following operations.

- (a) Apply k -means clustering on the vectorized data for $k = 4, k = 8$ and $k = 12$.
- (b) Find the document-cluster association (Which document belongs to which cluster) for each k in the above clustering.
- (c) Out of the three k values which one is better for the given data and why?

Link for tutorial video for this problem:

<https://www.youtube.com/watch?v=YcaMlHtBdp4>

***** THE END *****