

CS 556-B: Mathematical Foundations of Machine Learning

Homework 1: Linear Algebra (100 points)

Note: All solutions methods must be fully explained. If a problem requires you to submit your code (this will be explicitly mentioned in the question), please ensure that your code (in the form of a jupyter notebook) is developed in a python3 environment and appropriately commented.

Vectors

1. (1 point) Find the magnitude of the vector $\mathbf{x} = \begin{bmatrix} 1 \\ -3 \\ 8 \\ -5 \\ -1 \end{bmatrix}$

Answer: $\|\mathbf{x}\| = \sqrt{1^2 + (-3)^2 + 8^2 + (-5)^2 + (-1)^2} = \sqrt{1 + 9 + 64 + 25 + 1} = 10$ units

2. (2 points) Consider two vectors represented by the set $S = \left\{ \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \end{bmatrix} \right\}$ (with each in vector space \mathbb{R}^2). What is the span of S ? Does the vector $\begin{bmatrix} 20 \\ 4 \end{bmatrix}$ belong to the span of two vectors. Briefly explain your reasoning leveraging the definition of the *span* of a set of vectors.

Answer: The span of vectors belonging to set $S = \left\{ \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \end{bmatrix} \right\}$ is a set of all vectors in space \mathbb{R}^2 that can be reached using only the vectors in set S . The given vector $\begin{bmatrix} 20 \\ 4 \end{bmatrix}$ does not belong to the span of set S because the vector cannot be reached through a linear combination of the given vectors.

Dot Product

3. (2 points) If two vectors \mathbf{a} , \mathbf{b} have magnitudes 10 and 6 respectively and the angle between them is $\frac{\pi}{3}$ radians, what is their dot product?

Answer: $\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \|\mathbf{b}\| \cos \theta$
 $= 10 \times 6 \times \cos \pi/3$
 $= 60 \times \frac{1}{2} = 30$

4. (2 points) Let vector $\mathbf{u} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, calculate the dot product of \mathbf{u} and \mathbf{v} also calculate the angle between (i.e., not the cosine of the angle but the actual angle in radians or degrees) \mathbf{u} and \mathbf{v} .

Answer: $\mathbf{u} \cdot \mathbf{v} = 5 \times 2 + (-3) \times 1 = 7$
 $7 = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta = \sqrt{5^2 + (-3)^2} \cdot \sqrt{2^2 + 1^2} \cdot \cos \theta$
 $= \sqrt{34} \times 5 \cdot \cos \theta \Rightarrow \cos \theta = \frac{7}{\sqrt{170}} = 57.53 \text{ degrees}$

Linear Independence

5. (3 points) Check if the vectors $\mathbf{x} = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$, $\mathbf{z} = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$

are linearly independent. Note: The condition for linear independence is that given a set S of vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ and coefficients a, b, c , $a\mathbf{x} + b\mathbf{y} + c\mathbf{z} = \mathbf{0}$ if and only if $a = b = c = 0$.

Answer: From the surface, it looks like the vectors are linearly independent because it doesn't look like we can just multiply one of them by a scalar to reach another. But, let's try reducing the matrix containing the three vectors into Reduced Row Echelon Form:

$$R3 = R3 - 3R1 = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 3 & -1 \\ 0 & 2 & 0 \end{bmatrix}$$

Now if we go through steps: $R2 = R2 - 4R1$, $R2 = -R2$, $R1 = R1 - R2$, $R3 = R3 - 2R2$, $R3 = -1/10 * R3$, $R1 = R1 + 4R3$ then $R2 = R2 - 5R3$, we will reach the Reduced Row Echelon Form without having two rows either become equal or one of the rows becoming 0. Which means, that, only when $a = b = c = 0$ will the equation formed by those vectors be 0. That means that the 3 given vectors are linearly independent.

6. (10 points) Given a subset of vectors $S = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ for $k \in \mathbb{N}$ of a vector space V , prove that S is linearly independent iff a linear combination of elements of S with non-zero coefficients does not yield $\mathbf{0}$. **Hint:** To prove *iff* statements, i.e., $A \text{ iff } B$ ($A \iff B$), first prove $A \rightarrow B$, then prove $B \leftarrow A$.

Answer: If S is linearly independent, then the linear combination: $\mathbf{a}_1\mathbf{x}_1 + \mathbf{a}_2\mathbf{x}_2 + \dots + \mathbf{a}_k\mathbf{x}_k = \mathbf{0}$ has only one possible solution and that is $\mathbf{a}_1 = \mathbf{a}_2 = \mathbf{a}_3 = \dots = \mathbf{a}_k = \mathbf{0}$ because S is linearly independent. Therefore, a linear combination of vectors in S with non zero coefficients cannot yield the 0 vector. This means that there is no non-trivial linear combination of vectors within S that equals the zero vector. Thus $A \rightarrow B$

Now let's try it the other way, let's say that there is an $a_1 \neq 0$, this means that vector

$x_1 = \frac{a_2}{a_1}x_2 - \frac{a_3}{a_1}x_3 - \dots - \frac{a_k}{a_1}x_k$ which means that all vectors in S are not linearly independent, which is

contradictory. Thus, if $a_1 = a_2 = a_3 = \dots = a_k = 0$ then vectors in the given span S are linearly independent. Therefore $B \rightarrow A$

Since $A \rightarrow B$ and $B \rightarrow A \implies A \iff B$

Matrices

7. (3 points) Demonstrate the distributive property of matrix multiplication over addition.

Given $A = \begin{bmatrix} -1 & 5 \\ -7 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 2 & -3 \\ 1 & 8 \end{bmatrix}$, $C = \begin{bmatrix} 3 & 7 \\ -4 & 8 \end{bmatrix}$, demonstrate: $A(B+C) = AB + AC$

$$\text{Answer: } B + C = \begin{bmatrix} 2 & -3 \\ 1 & 8 \end{bmatrix} + \begin{bmatrix} 3 & 7 \\ -4 & 8 \end{bmatrix} = \begin{bmatrix} 2+3 & -3+7 \\ 1+(-4) & 8+8 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ -3 & 16 \end{bmatrix}$$

$$A(B+C) = \begin{bmatrix} -1 & 5 \\ -7 & 3 \end{bmatrix} * \begin{bmatrix} 5 & 4 \\ -3 & 16 \end{bmatrix} = \begin{bmatrix} -1*5+5*-3 & -1*4+5*16 \\ -7*5+3*-3 & -7*4+3*16 \end{bmatrix} = \begin{bmatrix} -20 & 76 \\ -44 & 20 \end{bmatrix}$$

$$AB + AC = \begin{bmatrix} -1 & 5 \\ -7 & 3 \end{bmatrix} * \begin{bmatrix} 2 & -3 \\ 1 & 8 \end{bmatrix} + \begin{bmatrix} -1 & 5 \\ -7 & 3 \end{bmatrix} * \begin{bmatrix} 3 & 7 \\ -4 & 8 \end{bmatrix} = \begin{bmatrix} 3 & 43 \\ -11 & 45 \end{bmatrix} + \begin{bmatrix} -23 & 33 \\ -33 & -25 \end{bmatrix} \\ = \begin{bmatrix} -20 & 76 \\ -44 & 20 \end{bmatrix} = A(B+C)$$

8. (10 points) Calculate the inverse of matrix $A = \begin{bmatrix} 1 & -3 & 3 \\ 8 & 5 & -2 \\ -4 & 7 & 2 \end{bmatrix}$. Note: It is acceptable to leave the final solution with fractional entities in the matrix (i.e., no requirement to convert fractions to decimal numbers).

$$\begin{aligned}
\text{Answer: 1) Minors} &= \begin{bmatrix} 24 & 8 & 76 \\ 8 & 5 & -2 \\ -4 & 7 & 2 \end{bmatrix} \\
2) \text{ Cofactors} &= 1) * \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 24 & -8 & 76 \\ 27 & 14 & 5 \\ -9 & 26 & 29 \end{bmatrix} \\
3) \text{ Transpose of 2)} &= \begin{bmatrix} 24 & 27 & -9 \\ -8 & 14 & 26 \\ 76 & 5 & 29 \end{bmatrix} \\
4) \text{ Inverse} &= 3) * \frac{1}{\det(A)} \\
\det(A) &= 1 * (10 + 14) + 3 * (16 - 8) + 3 * 76 = 276 \\
A^{-1} &= \frac{1}{276} * \begin{bmatrix} 24 & 27 & -9 \\ -8 & 14 & 26 \\ 76 & 5 & 29 \end{bmatrix}
\end{aligned}$$

Change of Bases

9. (10 points) Consider the three columns in matrix A (problem 8) to be our new basis of interest in R^3 . If a vector $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ defined on the natural basis in R^3 (i.e., $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$), how would vector \mathbf{x} be represented in the basis defined by the matrix A in problem 8?

Matrices

10. (3 points) Which of the following matrices (without being altered) is in Reduced Row Echelon Form (RREF)? For matrices that are not in RREF (if any) please state ALL the violations due to which they are not in RREF.

$$A_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 8 \\ 0 & 0 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

11. (4 points) What is the rank of each matrix below? Demonstrate intermediate steps employed to obtain the rank?

$$\begin{aligned}
(a) \quad A_1 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
(b) \quad A_2 &= \begin{bmatrix} 4 & -1 & 9 \\ 1 & 4 & 4 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

Subspaces & Projections

12. (5 points) If a matrix of dimensions 5×9 has rank 3, what are the dimensions of the 4 fundamental subspaces (i.e., row space, column space, null space, left null space)? What is the sum of all 4 dimensions i.e., add the dimension of each subspace and report the total?
13. (5 points) If a 4×5 matrix has rank 4, what are the dimensions of its columnspace (e.g., which of $\mathbb{R}^1, \mathbb{R}^2, \dots, \mathbb{R}^n$ represents the column space) and left nullspace (i.e., for a matrix $A_{m \times n}$, the left null space is the set of all vectors \mathbf{x} such that $A^T \mathbf{x} = 0$)?

14. (10 points) Find the set of vectors that form the null space of the matrix $A = \begin{bmatrix} 3 & 3 & 3 & 3 \\ 5 & 4 & 3 & 2 \\ 3 & 3 & 4 & 1 \end{bmatrix}$

15. (10 points) Find the *complete* set of solutions for the system of linear equations ($Ax=b$) defined below:

$$\begin{aligned} x_1 - 2x_2 - 2x_3 &= b_1 \\ 2x_1 - 5x_2 - 4x_3 &= b_2 \\ 4x_1 - 9x_2 - 8x_3 &= b_3 \end{aligned}$$

Note: First find the solution(s) that satisfy the *column space* relationship, then find the set of solutions to the null space and add the two to obtain the final solution.

16. (10 points) Calculate the Projection of \mathbf{b} onto the column space defined by matrix A . In each case, also calculate the magnitude of the vector from \mathbf{b} perpendicular to the projection.

(a) $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad b = \begin{bmatrix} 2 \\ 5 \\ 7 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \quad b = \begin{bmatrix} 3 \\ 3 \\ 7 \end{bmatrix}$

17. (10 points) Suppose matrix M is defined as follows: $M = \begin{bmatrix} 1 & 0 & 4 \\ 2 & 3 & 4 \\ 0 & -1 & 2 \end{bmatrix}$ Convert the matrix to form an equivalent orthonormal basis via. **orthogonalization** using the Gram-Schmidt procedure.

Answer: Finding A, B and C:

$$\mathbf{A} = \mathbf{a} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\mathbf{B} = \mathbf{b} - \frac{\mathbf{A}^T \cdot \mathbf{b}}{\mathbf{A}^T \cdot \mathbf{A}} \cdot \mathbf{A} = \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix} - \frac{\begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}}{\begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}} \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} -6/5 \\ 3/5 \\ -1 \end{bmatrix}$$

$$\begin{aligned} \mathbf{C} &= \mathbf{c} - \frac{\mathbf{A}^T \cdot \mathbf{c}}{\mathbf{A}^T \cdot \mathbf{A}} \cdot \mathbf{A} - \frac{\mathbf{B}^T \cdot \mathbf{c}}{\mathbf{B}^T \cdot \mathbf{B}} \cdot \mathbf{B} = \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix} - \frac{\begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix}}{\begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}} \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 0 & 3 & -1 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix}}{\begin{bmatrix} 0 & 3 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}} \cdot \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix} - \begin{bmatrix} 12/5 \\ 24/5 \\ 0 \end{bmatrix} - \begin{bmatrix} 66/35 \\ -33/35 \\ 11/7 \end{bmatrix} = \begin{bmatrix} -2/7 \\ 1/7 \\ 3/7 \end{bmatrix} \end{aligned}$$

Orthonormal vectors =

$$\mathbf{q1} = \frac{\mathbf{A}}{\|\mathbf{A}\|} = \frac{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}}{\sqrt{1^2+2^2}} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}$$

$$\mathbf{q2} = \frac{\mathbf{B}}{\|\mathbf{B}\|} = \frac{\begin{bmatrix} -6/5 \\ 3/5 \\ -1 \end{bmatrix}}{\sqrt{(-6/5)^2 + (3/5)^2 + (-1)^2}} = \begin{bmatrix} -6/\sqrt{70} \\ 3/\sqrt{70} \\ 5/\sqrt{70} \end{bmatrix}$$

$$\mathbf{q3} = \frac{\mathbf{C}}{\|\mathbf{C}\|} = \frac{\begin{bmatrix} -2/7 \\ 1/7 \\ 3/7 \end{bmatrix}}{\sqrt{(-2/7)^2 + (1/7)^2 + (3/7)^2}} = \begin{bmatrix} -2/\sqrt{14} \\ 1/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}$$

$$\text{Orthonormal Basis} = \begin{bmatrix} 1/\sqrt{5} & -6/\sqrt{70} & -2/\sqrt{14} \\ 2/\sqrt{5} & 3/\sqrt{70} & 1/\sqrt{14} \\ 0 & 5/\sqrt{70} & 3/\sqrt{14} \end{bmatrix}$$