Instructor: Nikhil Muralidhar
October 5, 2022

CS 556-B: Mathematical Foundations of Machine Learning
Homework 1: Linear Algebra (100 points)

Note: Calculators allowed for trigonometric operations & arithmetic operations (i.e., addition, subtraction, multiplication or division of *scalars*). All solutions methods must be full explained.

Vectors

.

1. (5 points) Find the magnitude of the vector $\mathbf{x} = \begin{bmatrix} 2 \\ -3 \\ 2 \\ 4 \\ -4 \end{bmatrix}$

   Solution: The magnitude of a vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

   is calculated using the formula:

$$||\mathbf{x}|| = \sqrt{x_1^2 + x_2^2 + ... + x_m^2}$$

   Hence,

$$||\mathbf{x}|| = \sqrt{2^2 + (-3)^2 + 2^2 + 4^2 + (-4)^2} = \sqrt{49} = 7$$

2. (5 points) Consider two vectors $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (in vector space $\mathbb{R}^2$), what is their span? Briefly explain your reasoning leveraging the definition of the span of a set of vectors.

   Solution: The span of a set of vectors in set S is defined as all possible vectors that are *reachable* using only the vectors in S. Specifically, with linear combinations of vectors in S.

   In the given set, we have two vectors $S = \{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \}$ and the space of interest is $\mathbb{R}^2$. Since the vectors in the given set are linearly independent (as they are orthogonal - implied by a dot product of zero between the two vectors), it implies that all vectors in $\mathbb{R}^2$ are reachable using linear combinations of vectors in S (i.e., S forms a basis in $\mathbb{R}^2$). Hence the span is the entire vector space $\mathbb{R}^2$.

Dot Product

3. (10 points) If two vectors $\mathbf{a}$, $\mathbf{b}$ have magnitudes 3 and 5 respectively and the angle between them is $\frac{\pi}{2}$ radians, what is their dot product?

Solution: We know that the dot product of two vectors is given by $\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}||\,||\mathbf{v}||\cos(\theta)$

We also know that the cosine of the angle between two perpendicular vectors (i.e., $\cos(\frac{\pi}{2})$) is zero. Hence it must be the case that the dot product of vectors $\mathbf{u}$ and $\mathbf{v}$ is zero. i.e., $\mathbf{u} \cdot \mathbf{v} = 0$.

4. (10 points) Let vector $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$, 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 $\mathbf{v}$.

Solution: We know that the dot product of two vectors is given by $\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}||\,||\mathbf{v}||\cos(\theta)$

$\mathbf{u} \cdot \mathbf{v} = 1 \times (-2) + 3 \times 7 = 19$

$||\mathbf{u}|| = \sqrt{1^2 + 3^2} = \sqrt{10}$

$||\mathbf{v}|| = \sqrt{(-2)^2 + 7^2} = \sqrt{53}$

$\theta = \frac{19}{\sqrt{10} \times \sqrt{53}} = \mathbf{0.6005}$ radians

or

Answer in Degrees: $\theta = \mathbf{34.3804}°$

Linear Independence

5. (15 points) Check if the vectors $\mathbf{x} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 3 \\ -2 \\ 2 \end{bmatrix}$, $\mathbf{z} = \begin{bmatrix} 5 \\ 2 \\ -6 \end{bmatrix}$ 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, $a\mathbf{x} + b\mathbf{y} + c\mathbf{z} = 0$ if and only if $a = b = c = 0$.

Solution: We can re-write the condition for linear independence using vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ as follows:

a$\mathbf{x}$ + b$\mathbf{y}$ + c$\mathbf{z}$ = $\mathbf{0}$

An alternate method to represent the relationship would be as a multiplication of a matrix by a vector to yield the $\mathbf{0}$ vector.

$$\begin{bmatrix} -1 & 3 & 5 \\ 0 & -2 & 2 \\ 2 & 2 & -6 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Essentially, we need to demonstrate that the above matrix vector produce has a = b = c = 0 as its solution. To do this, we can employ Gaussian Elimination.

We first re-write the above matrix vector produce in augmented matrix form.

$$\left[\begin{array}{ccc:c} -1 & 3 & 5 & 0 \\ 0 & -2 & 2 & 0 \\ 2 & 2 & -6 & 0 \end{array}\right]$$

Our goal in applying Gaussian elimination is to reduce the above matrix vector product to row-echelon form so that a solution can be obtained for a, b, c.

$$\left[\begin{array}{ccc:c} -1 & 3 & 5 & 0 \\ 0 & -2 & 2 & 0 \\ 2 & 2 & -6 & 0 \end{array}\right] \xrightarrow{R_3+2R_1} \left[\begin{array}{ccc:c} -1 & 3 & 5 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & 8 & 4 & 0 \end{array}\right] \xrightarrow{\frac{1}{2}R_2} \left[\begin{array}{ccc:c} -1 & 3 & 5 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 8 & 4 & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc:c} -1 & 3 & 5 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 8 & 4 & 0 \end{array}\right] \xrightarrow{R_1+3R_2} \left[\begin{array}{ccc:c} -1 & 0 & 8 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 8 & 4 & 0 \end{array}\right] \xrightarrow{R_3+8R_2} \boxed{\left[\begin{array}{ccc:c} -1 & 0 & 8 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 12 & 0 \end{array}\right]}$$

The final state (matrix highlighted with a black frame) is in *row echelon* form and we can derive that values of $c = 0$ (from row R_3), $b = 0$ (from row R_2) and $a = 0$ (from row R_1).

Hence we can conclude that the vectors $\mathbf{x}$, $\mathbf{y}, \mathbf{z}$ are independent.

6. (15 points) Given a subset of vectors $S = \{\mathbf{x}_1, \mathbf{x}_2, .., \mathbf{x}_k\}$ 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 **iff** B (A $\iff$ B), first prove A $\rightarrow$ B, then prove B $\leftarrow$ A.

Solution: The given condition implies an if and only if condition requiring us to prove both directions of the statement.

Part 1: Set S of vectors is linearly independent $\rightarrow$ a linear combination of elements of S with non-zero coefficients does not yield $\mathbf{0}$.

We shall prove part 1 by contradiction.

Condition 1: Set S of vectors is linearly independent $\rightarrow$ a linear combination of elements of S with non-zero coefficients yields $\mathbf{0}$

Mathematically, we can write the right hand side of *condition 1* as follows:

$$\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + ... + \lambda_k\mathbf{x}_k = 0 \tag{1}$$

Further, according to *condition 1*, let us assume some scalar $\lambda_i \neq 0$. Hence Eq. 1 can be re-written as follows:

$$\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + ... + \lambda_{i-1}\mathbf{x}_{i-1} + \lambda_{i+1}\mathbf{x}_{i+1} + ... + \lambda_k\mathbf{x}_k = -\lambda_i\mathbf{x}_i \tag{2}$$

Dividing both sides of Eq. 2 by $-\lambda_i$ we have:

$$\frac{\lambda_1}{-\lambda_i}\mathbf{x}_1 + \frac{\lambda_2}{-\lambda_i}\mathbf{x}_2 + ... + \frac{\lambda_{i-1}}{-\lambda_i}\mathbf{x}_{i-1} + \frac{\lambda_{i+1}}{-\lambda_i}\mathbf{x}_{i+1} + ... + \frac{\lambda_k}{-\lambda_i}\mathbf{x}_k = \mathbf{x}_i \tag{3}$$

However Eq. 3 implies that $\mathbf{x}_i$ is a linear combinations of the other vectors present in S, which is a contradiction of *condition 1.*

Part 2: If a linear combinations of vectors in S with non-zero coefficients does not yield $\mathbf{0}$ $\rightarrow$ vectors in S are linearly independent.

We shall once again prove part 2 by contradiction.

Condition 2: If a linear combinations of vectors in S with non-zero coefficients does not yield $\mathbf{0}$ $\rightarrow$ vectors in S are linearly dependent.

Let us assume some vector $\mathbf{x}_i \in S$ is linearly dependent on the rest of the vectors in S. This condition can be mathematically represented as follows:

$$\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + ... + \lambda_{i-1}\mathbf{x}_{i-1} + \lambda_{i+1}\mathbf{x}_{i+1} + ... + \lambda_k\mathbf{x}_k = \mathbf{x}_i \tag{4}$$

Adding $-x_i$ to both sides of Eq. 4 we get:

$$\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + ... + \lambda_{i-1}\mathbf{x}_{i-1} + \lambda_{i+1}\mathbf{x}_{i+1} + ... + \lambda_k\mathbf{x}_k - \mathbf{x}_i = 0 \tag{5}$$

Eq. 5 has the vector $\mathbf{x}_i$ with a non-zero coefficient i.e., -1 , yet Eq. 5 yields $\mathbf{0}$ which is a contradiction of condition 2. Hence by part 1, 2 we have proved the original statement of linear independence.

Matrices

7. (10 points) Demonstrate the distributive property of matrix multiplication over addition.

   Given $A = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 6 & 2 \\ 3 & 2 \end{bmatrix}$, $C = \begin{bmatrix} 1 & -2 \\ 4 & -1 \end{bmatrix}$, demonstrate: A(B+C) = AB + AC

   Solution: Given the matrices A, B, C, we first calculate AB, AC and B + C as follows.

$$AB = \begin{bmatrix} 3 \times 6 + 4 \times 3 & 3 \times 2 + 4 \times 2 \\ 1 \times 6 + 2 \times 3 & 1 \times 2 + 2 \times 2 \end{bmatrix} = \begin{bmatrix} 30 & 14 \\ 12 & 6 \end{bmatrix}$$

$$AC = \begin{bmatrix} 3 \times 1 + 4 \times 4 & 3 \times -2 + 4 \times -1 \\ 1 \times 1 + 2 \times 4 & 1 \times -2 + 2 \times -1 \end{bmatrix} = \begin{bmatrix} 19 & -10 \\ 9 & -4 \end{bmatrix}$$

$$\boxed{\text{AB + AC} = \begin{bmatrix} 49 & 4 \\ 21 & 2 \end{bmatrix}}$$

$$B + C = \begin{bmatrix} 6+1 & 2+(-2) \\ 3+4 & 2+(-1) \end{bmatrix} = \begin{bmatrix} 7 & 0 \\ 7 & 1 \end{bmatrix}$$

$$\boxed{\text{A(B + C)} = \begin{bmatrix} 3 \times 7 + 4 \times 7 & 3 \times 0 + 4 \times 1 \\ 1 \times 7 + 2 \times 7 & 1 \times 0 + 2 \times 1 \end{bmatrix} = \begin{bmatrix} 49 & 4 \\ 21 & 2 \end{bmatrix}}$$

   Thus we have demonstrated that the distributive property holds for distribution of matrix multiplication over addition.

8. (15 points) Calculate the inverse of matrix $A = \begin{bmatrix} 3 & 1 & 2 \\ -2 & -4 & 1 \\ 5 & -3 & 2 \end{bmatrix}$. 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).

   Solution: The procedure we follow for calculating the inverse of a matrix is to first calculate the *minors* and *cofactors* followed by the *adjugate* matrix. A^{-1} is then obtained by multiplying the inverse of the determinant of matrix A with the adjugate matrix.

   Minors Calculation The minors matrix is first calculated by considering a determinant of each 2×2 matrix obtained by disregarding the row and column of the matrix element under consideration.

$$\begin{bmatrix} \begin{vmatrix} -4 & 1 \\ -3 & 2 \end{vmatrix} & \begin{vmatrix} -2 & 1 \\ 5 & 2 \end{vmatrix} & \begin{vmatrix} -2 & -4 \\ 5 & -3 \end{vmatrix} \\ \begin{vmatrix} 1 & 2 \\ -3 & 2 \end{vmatrix} & \begin{vmatrix} 3 & 2 \\ 5 & 2 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 5 & -3 \end{vmatrix} \\ \begin{vmatrix} 1 & 2 \\ -4 & 1 \end{vmatrix} & \begin{vmatrix} 3 & 2 \\ -2 & 1 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ -2 & -4 \end{vmatrix} \end{bmatrix} \xrightarrow{Minors} \begin{bmatrix} -5 & -9 & 26 \\ 8 & -4 & -14 \\ 9 & 7 & -10 \end{bmatrix}$$

   Cofactors: The cofactors matrix is obtained by multiplying the minors matrix (elementwise) with the following matrix.

$$\begin{bmatrix} +1 & -1 & +1 \\ -1 & +1 & -1 \\ +1 & -1 & +1 \end{bmatrix}$$

The co-factors matrix is as follows:

$$\begin{bmatrix} -5 & 9 & 26 \\ -8 & -4 & 14 \\ 9 & -7 & -10 \end{bmatrix}$$

Adjugate Matrix: The adjugate matrix is obtained as a transpose of the cofactors matrix.

$$\boxed{\begin{bmatrix} -5 & -8 & 9 \\ 9 & -4 & -7 \\ 26 & 14 & -10 \end{bmatrix}}$$

Determinant of A: The determinant of matrix A i.e., $|A|$ is obtained by applying the *Laplace expansion* to the top row of A, which can be achieved by multiplying element-wise the top row of the cofactors matrix with the top row of A. $\boxed{|A| = 3\times(-5) + 1 \times (9) + 2 \times 26 = -15 + 9 + 52 = \mathbf{46}}$

$\mathbf{A^{-1}}$: The inverse of A is given by $\frac{1}{|A|}\text{adj}(A)$ where adj(A) represents the adjugate matrix of A.

$$A^{-1} = \frac{1}{46}\begin{bmatrix} -5 & -8 & 9 \\ 9 & -4 & -7 \\ 26 & 14 & -10 \end{bmatrix} \rightarrow \begin{bmatrix} \frac{-5}{46} & \frac{-4}{23} & \frac{9}{46} \\ \frac{9}{46} & \frac{-2}{23} & \frac{-7}{46} \\ \frac{13}{23} & \frac{7}{23} & \frac{-5}{23} \end{bmatrix}$$

Change of Bases

9. (15 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., $\mathbf{e_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{e_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{e_1} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$), how would vector $\mathbf{x}$ be represented in the basis defined by the matrix A in problem 8.

Solution: Given the vector $\mathbf{x}$ defined on the standard basis and given the new basis defined by matrix A (from problem 8), we know the relationship that

$A\mathbf{x}_A = \mathbf{x}$ holds. Our goal is to find the value of vector $\mathbf{x}_A$ which is the representation of $\mathbf{x}$ in the basis defined by A.

$$A\mathbf{x}_A = \mathbf{x}$$
$$\mathbf{x}_A = A^{-1}\mathbf{x}$$

substituting in values of x and A^{-1} we have the following relationship.

$$\mathbf{x}_A = \begin{bmatrix} \frac{-5}{46} & \frac{-4}{23} & \frac{9}{46} \\ \frac{9}{46} & \frac{-2}{23} & \frac{-7}{46} \\ \frac{13}{23} & \frac{7}{23} & \frac{-5}{23} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{-5-16+27}{46} \\ \frac{9-8-21}{46} \\ \frac{26+28-30}{46} \end{bmatrix} = \boxed{\begin{bmatrix} \frac{6}{46} \\ \frac{-20}{46} \\ \frac{24}{46} \end{bmatrix}}$$

Hence we have (the vector highlighted in the black box) as the value of the vector $\mathbf{x}$ in the new basis defined by matrix A. This new vector is represented as vector $\mathbf{x}_A$.