Theory Assignment :-

problem 1 : The proof of Linearity.

In DL, an affine layer $Z = Wx + b$ is often called a Linear Layer, but without a non linear activation function (Like Relu), stacking multiple layers simply Collapses into one single Linear transformation. Let us prove what makes a function truly Linear.

Condition two transformations T and S:

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x + y \\ x - 3y \end{bmatrix}$$

$$S\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x^2 \\ y \end{bmatrix}$$

Task A: Formally prove that transformation T is Linear by showing it satisfies both Additivity and homogeneity.

TaskB: Formally prove that the transformation S is not linear by providing a specific counter-example where the rules of linearity fail.

<u>Solutions</u>:

TaskA: $T(u+v) = T(u) + T(v)$.

let $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

$$T(u+v) = T\left(\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix}\right) = \begin{bmatrix} 2(u_1+v_1) + (u_2+v_2) \\ (u_1+v_1) - 3(u_2+v_2) \end{bmatrix}$$

$$= \begin{bmatrix} (2u_1+u_2) + (2v_1+v_2) \\ (u_1-3u_2) + (v_1-3v_2) \end{bmatrix} = \begin{bmatrix} 2u_1+u_2 \\ u_1-3u_2 \end{bmatrix} + \begin{bmatrix} 2v_1+v_2 \\ v_1-3v_2 \end{bmatrix} = T(u)+T(v)$$

Homogeneity $T(cu) = cT(u)$ is similar.

<u>TaskB</u>: Let $u = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and Scalar $c=2$

$$S(2u) = S\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 4^2 \\ 2 \end{bmatrix} = \begin{bmatrix} 16 \\ 2 \end{bmatrix}$$

However, $2S(u) = 2\begin{bmatrix} 2^2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \end{bmatrix}$ since $\begin{bmatrix} 16 \\ 2 \end{bmatrix} \neq \begin{bmatrix} 8 \\ 1 \end{bmatrix}$ non linear.

Problem 2: The Geometry of the Determinant.

The determinant of a 2X2 matrix is not just a calculation $(ad - bc)$; it represents the factor by which a linear transformation scales area. Consider the unit square defined by the standard basis vector $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Its area is exactly 1.

Consider the following two transformation matrices:

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Task A: Calc Det(A), and geometrically explain what is this matrix does the unit square

Task B: Calc Det(B), Matrix B is called a "Shear" transformation. explain conceptually how a transformation can severly warp the shape of the grid: but still result in a determint 1.

Solution

Task A: $\det(A) = 3 \cdot 2 - 0 \cdot 0 = 6$.

Geometrically ~~this~~ ~~transformation~~ stretchs the x axis by 3 and y axis by 2. The unit square of an area is 6.

Task B: $\det(B) = 1 \cdot 1 - 1 \cdot 0 = 1$.

While the square is sheared into a slanted parallelogram ~~that~~ the base width remain 1, and perpendicular hights remain exactly 1.

Area = base × height

= 1 × 1

= 1

The transformation distroy the shape but preserves the total area of the space perfectly.

Problem 3: The Null (Kernel) Space and Information Loss.

In ML, if your i/p data has 100 features but the transformation matrix W collapses them into a 10 Dim output, you loss information.

The Null Space (Kernal) of a matrix C is the set of all vectors x that get mapped to zero vector: $Cx=0$.
If non-zero vector get crushed to zero, the transformation is not invertible.

Consider the matrix:

$$C=\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$$

Task A: Find the non zero vector $x=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ such that $Cx=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

TaskB: let $v=\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ compute $y=Cv$, Then Compute $C(v+x)$, where x is the vector you found in

TASK A

TASK C: Based on the result of TASK B, logically explain why matrix C can't have an inverse matrix C^{-1}

Solutions:

TASK A: Solve $CX = 0 \Rightarrow \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$x_1 - x_2 = 0 \Rightarrow x_1 = x_2$

Therefore, any vector where coordinates are equals, such as $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. maps to the zero. vector.

TASK B: $Cv = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

$C(v+x) = C\begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

TASK C: A matrix is only invertible if every output traces back to exactly one input unique input.

Since both input vector $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 5 \end{bmatrix}$ map to the exact same output $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$, there is no way for an inverse function to know which one was the original input. The matrix destroys information.