Foundational Mathematics for ML (MA2221)

Assignment (NumPy & Linear Algebra)

Instructor: **Biswarup Biswas**

Instructions:

- All problems must be solved using **NumPy**.
- Use **functions, loops, and conditional statements** wherever appropriate.
- Avoid using advanced routines not discussed in the notebook.
- Problems are marked with difficulty levels: Easy ([E]), Moderate ([M]), Challenging ([C]).

1. [E] Create a NumPy array x of length 10 containing equally spaced values between 0 and 1.
   (a) Compute $x^2 + 2x + 1$ using vectorized operations.
   (b) Repeat the computation using a `for` loop.
   (c) Verify that both results are identical.

2. [E] Write a function `safe_normalize(x)` that:
   - computes the Euclidean norm of x,
   - normalizes x only if $\|x\|_2 > 1$,
   - otherwise returns x unchanged.

   Test the function on vectors of different magnitudes.

3. [M] Given vectors $x, y \in \mathbb{R}^n$:
   (a) Compute their dot product using a loop.
   (b) Compute the dot product using `np.dot`.
   (c) Compute the Euclidean distance between x and y.

   Compare the results numerically.

4. [E] Given the matrix
$$X = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix},$$
   compute the distance between the first row and every other row using a loop.

5. [M] Let $v \in \mathbb{R}^n$ be a fixed vector and $X \in \mathbb{R}^{m \times n}$.
   (a) Compute the cosine similarity between v and each row of X.
   (b) Print only those rows whose similarity exceeds 0.8.

   Use a loop and an `if` condition.

6. [M] Given vectors $x, y \in \mathbb{R}^n$:
   (a) Compute the projection of x onto y.
   (b) Compute the residual vector.
   (c) Verify numerically that the residual is orthogonal to y.

7. [E] Generate random 2×2 matrices.
   (a) Compute their determinants using `np.linalg.det`.
   (b) Use an `if` condition to identify nearly singular matrices.

8. [M] For an invertible matrix A and vector b:
   (a) Solve $Ax = b$ using `np.linalg.solve`.
   (b) Solve the same system using $A^{-1}b$.
   (c) Compare the two solutions numerically.

9. [M] Let $x, y \in \mathbb{R}^n$ and consider the model $y \approx wx$.
   (a) Compute the error $\|y - wx\|_2^2$ for different values of w.
   (b) Find the value of w that minimizes the error.
   (c) Plot the error as a function of w.

10. [C] Write a program that:
    - generates a random dataset,
    - normalizes each vector if needed,
    - computes pairwise cosine similarities,
    - prints a warning if highly similar vectors are detected.

::: End of Assignment Sheet :::