Linear Algebra for Machine Learning: Core Concepts and Applications

Introduction and Class Flow

- The class opens with technical checks, a discussion about attendance, and a brief on the schedule. The main topic is linear algebra as foundational for machine learning.

Linear Algebra Fundamentals

- Explanation of scalars (quantities with only magnitude) and vectors (quantities with magnitude and direction).

- Data points in datasets can be visualized as points with coordinates (features/variables/dimensions) on mathematical graphs.

- Concepts such as rows and columns in data are introduced as axes, features, or variables.

Matrix Operations

- Matrices (collections of numbers arranged in rows and columns) represent data, with applications in image processing (e.g., pixels in filters/resolutions).

- Definitions of shape (number of rows × number of columns), square matrices (rows = columns), and rectangular matrices.

- Matrix addition: Element-wise (corresponding positions) addition of two matrices of the same shape.

- Matrix multiplication: Row × column operation, only possible when the number of columns in the first equals the number of rows in the second; output matrix shape is (rows of first × columns of second).

- Practical examples and student exercises for both addition and multiplication.

Distances in Mathematics

- Overview of how to calculate distance from origin and between two points in any dimension using the Pythagorean theorem.

- Euclidean distance: Standard straight-line distance (root of sum of squared differences across each feature).

- Limitations of Euclidean distance in high dimensions (all points can seem nearly equidistant).

- Manhattan distance: Distance measured along axes/grid lines (sum of absolute differences), appropriate for "street grid" scenarios.

- Minkowski distance: Generalized formula encompassing both Euclidean and Manhattan distances by varying a parameter.

Vector Multiplication and Projections

- Types of vector multiplication:

  - Dot product: Combines two vectors, output is a scalar (sum of products of corresponding elements), foundational for ML computations, and requires one vector to be transposed for valid multiplication.

  - Cross product: Not used for most ML applications, results in a vector (direction).

- Projections: Concept of projecting one vector onto another (shadow analogy), with formulas based on dot product and vector magnitudes.

- Unit vectors: Vectors with magnitude 1, used to normalize direction for transformation and calculations.

Eigenvalues and Eigenvectors

- Eigenvector: A vector whose direction remains unchanged after a linear transformation (only its magnitude may change).
- Eigenvalue: The factor by which the eigenvector length is scaled by transformation.
- Applications: Essential in image processing, dimensionality reduction, and data analysis for identifying principal directions of change or variance.
- Illustration and interactive explanation clarifying non-eigenvectors versus eigenvectors after transformation.

Lines, Planes, and Hyperplanes in ML

- Equation of a line ($y = mx + c$) and its interpretation in data/machine learning (e.g., regression).
- Slope/intercept as weights in ML modeling.
- Extension from lines (2D) to planes (3D) and hyperplanes (nD) for separating/classifying data.
- The significance of hyperplanes (generalization of plane) in high-dimensional ML tasks and classifiers.
- Use of dot products in expressing these equations.
- Example of applying these models to classify data (e.g., diabetic vs. non-diabetic points).

Applications and Industry Tools

- Explanation of practical tools needed for ML/data science: Pandas, NumPy, Scikit-learn; SciPy for advanced analytics, plus modules like logging or os for coding tasks. Emphasis on core module proficiency over lesser-used libraries.

- Real-world analogies: use of pixel matrices in image filtering, ML classifier separation boundaries, navigation distances, etc.

Session Wrap-Up

- Summary of progression from basic concepts to applications.

- Next class will pick up remaining topics; students are encouraged to study definitions and explore online tools for practice.

- Encouragement and advice regarding learning path and necessary module knowledge.