Mathematical Foundations for Machine Learning (S1-24_AIMLCZC416): Lecture 2 Overview

In the second lecture of the course *Mathematical Foundations for Machine Learning*, we delved into several critical concepts that form the backbone of machine learning algorithms and their mathematical principles. Here's a detailed breakdown of the topics covered, including key points, practical examples, and recommendations for further study.

Lecture Objectives

1. **To understand the basics of solving systems of linear equations.**
2. **To explore concepts of vector spaces, norms, and inner products.**
3. **To examine the relevance of these mathematical tools in machine learning contexts.**

Key Topics Covered

1. Systems of Linear Equations

- **Definition**: A set of equations with multiple variables, which are solved to find a common solution.
- **Mathematical Representation**:

$$A\mathbf{x} = \mathbf{b}$$

  where:

  - A is a coefficient matrix.
  - $\mathbf{x}$ is a vector of variables.

  - $\mathbf{b}$ is a vector of constants.
- **Techniques Discussed**:
  - Gaussian Elimination
  - Row Echelon Form (REF) and Reduced Row Echelon Form (RREF)
  - Matrix Inversion Method
- **Practical Example**: Solve the following system:

$$\begin{gathered} 2x + y - z = 8 \\ -3x - y + 2z = -11 \\ -2x + y + 2z = -3 \end{gathered}$$

  Solution steps were demonstrated using Gaussian elimination.
- **Applications in Machine Learning**:
  - Used in linear regression models.
  - Forms the basis for optimization problems.

---

2. Vector Spaces

- **Key Concepts**:
  - **Linear Independence**: A set of vectors is linearly independent if no vector in the set is a linear combination of the others.
  - **Basis and Rank**:
    - **Basis**: A minimal set of linearly independent vectors that span the space.
    - **Rank**: The dimension of the vector space spanned by the rows or columns of a matrix.
  - **Affine Spaces**: A geometric structure that generalizes the properties of parallel lines.
- **Relevance to Machine Learning**:

- Vector spaces are foundational to understanding feature spaces in algorithms like PCA and SVM.

---

3. Norms and Inner Products

- **Norms**:
  - Definition: A measure of the "length" or "size" of a vector.
  - Common Types:
    - L^2-**Norm**: $\|x\|_2 = \sqrt{\sum_i x_i^2}$
    - L^1-**Norm**: $\|x\|_1 = \sum_i |x_i|$
- **Inner Products**:
  - Definition: A generalization of the dot product that defines angles between vectors.
  - Formula:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} u_i v_i$$

- **Applications in Machine Learning**:
  - Norms are used in regularization techniques like Lasso (L^1) and Ridge (L^2).
  - Inner products are used to calculate similarity between data points.

---

4. Analytic Geometry

- **Geometric Interpretation**:
  - Understanding lengths (L^2-norm), distances, and angles.

- Orthonormal basis: A set of vectors that are orthogonal and of unit length.
- **Use in Machine Learning**:
  - Feature scaling and transformation.
  - Geometric interpretation of optimization problems.

Highlights and Practical Insights

1. **Interactive Problem-Solving**:
   - The session included solving systems of equations live using Gaussian elimination and discussing the significance of solutions in optimization.
2. **Conceptual Connections**:
   - Explained how vector spaces underpin machine learning models like linear regression and PCA.
3. **Real-World Applications**:
   - Demonstrated the relevance of norms in penalizing large coefficients during regression.

Challenges Addressed

1. **Difficulty in Understanding RREF**:
   - A step-by-step approach was provided to convert matrices into their reduced forms, with examples.
2. **Linking Theory to Practice**:
   - Made connections between abstract mathematical concepts and practical machine learning applications.
3. **Managing Lecture Pace**:

- Slides were used to maintain clarity and speed, ensuring better understanding compared to the previous session.