# Comprehensive Study Guide: Mathematical Foundations for Machine Learning

Mathematical foundations are indispensable for designing and analyzing machine learning algorithms. This guide provides in-depth knowledge of essential topics such as **Linear Algebra**, **Vector Calculus**, and **Optimization**, with real-world applications and examples. The guide is designed to prepare students for lectures and exams while also providing a strong foundation for practical machine learning problems.

---

## Section 1: Vectors and Vector Spaces

### Core Concepts

- **Vector Definition**: A vector is a mathematical entity that has both magnitude and direction. In machine learning, vectors represent features, data points, and even weights of a model. For instance, a 3D vector is written as $\vec{v} = [v_1, v_2, v_3]$, where $v_1, v_2$, and $v_3$ are components along the x, y, and z axes.
- **Vector Operations**:
  - **Addition**: Combines two vectors element-wise: $\vec{u} + \vec{v} = [u_1 + v_1, u_2 + v_2, ..., u_n + v_n]$.
  - **Scalar Multiplication**: Scales each component of a vector by a scalar: $c\vec{v} = [c \cdot v_1, c \cdot v_2, ..., c \cdot v_n]$.
  - **Dot Product**: Measures the similarity of two vectors: $\vec{u} \cdot \vec{v} = \sum_{i=1}^{n} u_i v_i$.
  - **Cross Product** (only in 3D): Produces a vector orthogonal to both input vectors.
- **Linear Independence**: A set of vectors $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_k\}$ is linearly independent if no vector in the set can be written as a linear combination of the others. For example:

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_k\vec{v}_k = \vec{0}$$

  If $c_1, c_2, \ldots, c_k = 0$ is the only solution, the vectors are independent.

- **Basis and Dimension**:
  - A **basis** is a minimal set of linearly independent vectors that spans a vector space.
  - The **dimension** of a vector space is the number of vectors in its basis.

### Applications in Machine Learning

1. **Feature Representation**: Each data point is represented as a vector in a high-dimensional space.
2. **Embeddings**: Word embeddings like Word2Vec represent words as vectors in semantic spaces.
3. **Projections**: Projecting data onto lower-dimensional subspaces for visualization or computation.

### Examples

1. Given $\vec{u} = [1, 2]$ and $\vec{v} = [3, 4]$:
   - $\vec{u} + \vec{v} = [4, 6]$,
   - $2\vec{u} = [2, 4]$.
2. Find if $\vec{v}_1 = [1, 2]$, $\vec{v}_2 = [2, 4]$ are linearly independent:
   - Since $\vec{v}_2 = 2\vec{v}_1$, they are dependent.

---

## Section 2: Matrices and Linear Transformations

Matrices are rectangular arrays of numbers used to encode linear transformations. They are critical for performing operations on data.

### Core Concepts

- **Matrix Definition**: A matrix is written as:

$$A = [\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix}],$$

where $a_{ij}$ represents the element in the $i$-th row and $j$-th column.

- **Matrix Operations**:
  - **Addition**: $A + B = C$, where $c_{ij} = a_{ij} + b_{ij}$.
  - **Scalar Multiplication**: $cA = \begin{bmatrix} ca_{11} & ca_{12} \\ ca_{21} & ca_{22} \end{bmatrix}$.
  - **Matrix Multiplication**: Combines two matrices $A$ and $B$:

$$C = AB, \quad c_{ij} = \sum_k a_{ik} b_{kj}.$$

- **Special Matrices**:
  - **Identity Matrix**: Diagonal entries are 1; all others are 0.
  - **Transpose**: Rows become columns, $A^T$.

### Applications in Machine Learning

1. **Neural Networks**: Weight matrices transform input vectors.
2. **Data Transformations**: Scaling, rotation, and projection of data.

### Examples

1. For $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$:
   - $A + B = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$.

---

## Section 3: Systems of Linear Equations

### Core Concepts

- **Representation**: A system of linear equations can be written in matrix form:

$$Ax = b,$$

  where $A$ is the coefficient matrix, $x$ is the variable vector, and $b$ is the constant vector.

- **Solution Types**:
  1. **Unique Solution**: System is consistent and independent.
  2. **Infinite Solutions**: System has free variables.
  3. **No Solution**: System is inconsistent.

### Solution Techniques

1. **Gaussian Elimination**:
   - Reduce the matrix to Row-Echelon Form (REF) using elementary row operations.
   - Perform back-substitution to find solutions.
2. **Gauss-Jordan Elimination**:
   - Further reduce the matrix to Reduced Row-Echelon Form (RREF) for direct solutions.

### Examples

1. Solve:

$$x + y + z = 6, \quad 2x + 3y + z = 14, \quad 3x + y + 2z = 14.$$

   Steps:

   - Reduce to REF:

$$\begin{bmatrix} 1 & 1 & 1 & 6 \\ 2 & 3 & 1 & 14 \\ 3 & 1 & 2 & 14 \end{bmatrix}.$$

   - Back-substitute to find $x = 2, y = 4, z = 0$.

# Section 4: Eigenvalues and Eigenvectors

### Core Concepts

- **Definition**: Eigenvalues $\lambda$ and eigenvectors $\vec{v}$ satisfy:

$$A\vec{v} = \lambda\vec{v}.$$

### Applications

- **Principal Component Analysis (PCA)**: Reduces dimensionality while preserving variance.

### Example

For $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$:

1. Eigenvalues: $\lambda = 5, 2$.
2. Eigenvectors: $\vec{v}_1 = [1, 2]^T$, $\vec{v}_2 = [-1, 1]^T$.