Introduction to Linear Algebra with Applications to AI and Machine Learning

Johar M. Ashfaque

March 24, 2025

1 Introduction

Linear Algebra is a branch of mathematics that deals with vectors, matrices, and linear transformations. It provides the foundational tools for many areas of mathematics, physics, engineering, economics, and computer science. In this document, we will cover the basic concepts of linear algebra and explore their applications in Artificial Intelligence (AI) and Machine Learning (ML).

2 Basic Concepts of Linear Algebra

2.1 Vectors

A **vector** is an ordered list of numbers, which can be represented as a column or row. A vector in n-dimensional space can be written as:

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

where $v_1, v_2, \ldots, v_n$ are the components of the vector. Vectors can be added or scaled by scalar values. For two vectors $\mathbf{u}$ and $\mathbf{v}$ in n-dimensional space:

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix}$$

A scalar multiple of a vector $\mathbf{v}$ is:

$$
c \cdot \mathbf{v} = \begin{pmatrix} c \cdot v_1 \\ c \cdot v_2 \\ \vdots \\ c \cdot v_n \end{pmatrix}
$$

2.2 Matrices

A **matrix** is a rectangular array of numbers arranged in rows and columns. A matrix A with m rows and n columns is denoted as $A \in \mathbb{R}^{m \times n}$, where:

$$
A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}
$$

2.3 Matrix Operations

Matrices can be added, multiplied by scalars, and multiplied by other matrices. For matrix multiplication, if A is an $m \times n$ matrix and B is an $n \times p$ matrix, their product AB will be an $m \times p$ matrix:

$$
C = AB, \quad C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}
$$

2.4 Linear Independence and Span

A set of vectors is **linearly independent** if no vector in the set can be written as a linear combination of the others. If a set of vectors is linearly dependent, at least one vector can be expressed as a linear combination of the others. The **span** of a set of vectors is the set of all possible linear combinations of those vectors.

2.5 Eigenvalues and Eigenvectors

An **eigenvector** of a matrix A is a non-zero vector $\mathbf{v}$ such that:

$$
A\mathbf{v} = \lambda \mathbf{v}
$$

where λ is a scalar known as the **eigenvalue** corresponding to the eigenvector $\mathbf{v}$.

3 Applications of Linear Algebra in AI and Machine Learning

Linear algebra is the backbone of many algorithms used in AI and Machine Learning. Some of the key areas where linear algebra plays a role include:

3.1 Data Representation

In AI, data is often represented as matrices or tensors. For instance, images can be represented as 2D matrices of pixel values, and color images as 3D tensors (height, width, and color channels). Linear algebra allows for efficient manipulation of such data.

3.2 Principal Component Analysis (PCA)

PCA is a dimensionality reduction technique commonly used in machine learning. The goal is to transform the data into a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinates (called the principal components). The computation of the principal components involves eigenvectors and eigenvalues of the covariance matrix of the data.

3.3 Linear Regression

Linear regression is a fundamental technique in statistics and machine learning for predicting a target variable from a set of features. The relationship between the dependent variable $\mathbf{y}$ and the independent variables $\mathbf{X}$ is modeled as:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon$$

where β is the vector of coefficients, and ϵ is the error term. Finding the optimal β involves solving a system of linear equations, typically using matrix operations.

3.4 Neural Networks

Neural networks, a key component of deep learning, heavily rely on matrix operations. In a neural network, the data is passed through multiple layers, each performing linear transformations (matrix multiplications) followed by nonlinear activations. Backpropagation, the algorithm used to train neural networks, involves computing gradients using matrix derivatives.

3.5 Singular Value Decomposition (SVD)

SVD is used in many machine learning algorithms, such as in recommendation systems and image compression. SVD decomposes a matrix A into three other matrices U, Σ, and V such that:

$$A = U\Sigma V^T$$

where U and V are orthogonal matrices, and Σ is a diagonal matrix. This decomposition is useful for dimensionality reduction and noise reduction in data.

4 Conclusion

Linear algebra is an essential tool for understanding and implementing AI and machine learning algorithms. Its applications range from data representation and dimensionality reduction to solving optimization problems and training deep neural networks. A solid understanding of vectors, matrices, eigenvalues, eigenvectors, and matrix decompositions is crucial for anyone looking to delve into these fields.