Comprehensive and Beginner-Friendly Guide to Vector Spaces, Eigenvalues, and Matrix Factorizations

This guide breaks down the core concepts of **vector spaces, eigenvalues, and matrix factorizations** into simple and detailed explanations, including practical examples and real-world applications. These topics are crucial for understanding machine learning, and the content is tailored for beginners.

---

2. Vector Spaces

Vector spaces are foundational in mathematics and machine learning, providing the structure to represent and analyze data, features, and models.

2.1 What is a Vector Space?

A **vector space** is a set of vectors where two key operations are defined:

1. **Addition**: Adding two vectors produces another vector in the same space.
2. **Scalar Multiplication**: Multiplying a vector by a scalar (a number) produces another vector in the same space.

Key Properties of a Vector Space

- **Zero Vector**: The space contains a zero vector, denoted as $\vec{0}$, such that:

$$\vec{v} + \vec{0} = \vec{v}.$$

- **Closure**: The addition and scalar multiplication of vectors remain within the space.
- **Distributive Properties**: Addition and scalar multiplication distribute over scalars and vectors.

2.2 Key Concepts in Vector Spaces

Basis

- A **basis** is a minimal set of vectors that can represent any vector in the space through a linear combination.
- Example: In R^2, the standard basis is:

$$\{[1, 0], [0, 1]\}.$$

  Any vector in R^2, such as $[3, 5]$, can be expressed as:

$$[3, 5] = 3 \cdot [1, 0] + 5 \cdot [0, 1].$$

Dimension

- The **dimension** of a vector space is the number of vectors in its basis.
- Example: R^3 has dimension 3 because it requires three basis vectors, e.g., $\{[1, 0, 0], [0, 1, 0], [0, 0, 1]\}$.

Linear Independence

- A set of vectors $\{v_1, v_2, \ldots, v_k\}$ is **linearly independent** if:

$$c_1 v_1 + c_2 v_2 + \ldots + c_k v_k = 0 \implies c_1 = c_2 = \ldots = c_k = 0.$$

- Example: $\{[1, 0], [0, 1]\}$ is independent because no vector can be written as a combination of the other.

2.3 Practical Applications of Vector Spaces

Feature Representation

- In machine learning, data points are often represented as vectors.
- Example: An image with 28×28 pixels can be represented as a vector in R^{784}.

Dimensionality Reduction

- High-dimensional data is projected onto a lower-dimensional subspace while preserving important features.
- Example: **Principal Component Analysis (PCA)** reduces data dimensions by finding new axes (principal components) that explain the most variance.

---

3. Eigenvalues and Eigenvectors

3.1 What Are Eigenvalues and Eigenvectors?

An **eigenvalue** and its corresponding **eigenvector** provide insight into how a matrix transforms data. They satisfy the equation:

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

where:

- A: A square matrix.
- $\vec{v}$: An eigenvector (a non-zero vector).
- λ: An eigenvalue (a scalar).

3.2 Understanding Eigenvalues and Eigenvectors

- The matrix A scales the eigenvector $\vec{v}$ by the factor λ.
- Example: In PCA, eigenvectors define directions of maximum variance, and eigenvalues measure the magnitude of that variance.

3.3 Practical Example

Find the eigenvalues and eigenvectors for:

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Step 1: Compute the characteristic equation:

$$\det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = 0.$$

$$(2-\lambda)^2 - 1 = 0 \implies \lambda^2 - 4\lambda + 3 = 0.$$

$$\lambda = 3,\ 1.$$

Step 2: Solve for eigenvectors:

- For $\lambda = 3$:

$$(A - 3I)\vec{v} = 0 \implies \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

- For $\lambda = 1$:

$$(A - I)\vec{v} = 0 \implies \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

3.4 Applications

1. **Principal Component Analysis (PCA)**:
   - Eigenvectors of the covariance matrix define new axes for data projection.

2. **Spectral Clustering**:
   - Eigenvectors of graph Laplacians partition data into clusters.

---

4. Matrix Factorizations

4.1 What is Matrix Factorization?

Matrix factorization decomposes a matrix into simpler matrices to make computations more efficient.

4.2 Types of Matrix Factorizations

LU Decomposition

- $A = LU$, where L is lower triangular, and U is upper triangular.
- Application: Efficiently solves linear systems using forward and backward substitution.

Example: For:

$$A = \begin{bmatrix} 4 & 3 \\ 6 & 3 \end{bmatrix},$$

LU decomposition gives:

$$L = \begin{bmatrix} 1 & 0 \\ 1.5 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 4 & 3 \\ 0 & -1.5 \end{bmatrix}.$$

QR Decomposition

- $A = QR$, where:
  - Q: Orthogonal matrix ($Q^T Q = I$).
  - R: Upper triangular matrix.
- Application: Used in least-squares regression.

Singular Value Decomposition (SVD)

- $A = U\Sigma V^T$, where:
  - U, V: Orthogonal matrices.
  - Σ: Diagonal matrix of singular values.
- Application: Dimensionality reduction (e.g., PCA).

Example: For $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, SVD gives:

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

4.3 Applications

1. **Dimensionality Reduction**:
   - PCA uses SVD to reduce high-dimensional data while preserving variance.
2. **Collaborative Filtering**:
   - Matrix factorization is used in recommendation systems (e.g., Netflix).