8. Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) is a fundamental concept in linear algebra used extensively in data science and machine learning. It provides a way to decompose a matrix into three simpler matrices, revealing important properties of the original matrix.

What is SVD?

Given a matrix A, SVD decomposes it as:

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

where:

- A is the original $m \times n$ matrix.
- U is an $m \times m$ orthogonal matrix (columns of U are left singular vectors of A).
- Σ is an $m \times n$ diagonal matrix containing singular values (sorted in decreasing order).
- V^T is the transpose of an $n \times n$ orthogonal matrix (rows of V^T are right singular vectors of A).

Breaking Down the Components

1. **Orthogonal Matrix U:**
   - Columns of U form an orthonormal basis for the column space of A.
   - If U has orthogonal columns, multiplying it does not distort angles or magnitudes.
2. **Diagonal Matrix Σ:**
   - The singular values in Σ are the square roots of the eigenvalues of A^TA or AA^T.
   - These values represent the "strength" or importance of each dimension.
3. **Orthogonal Matrix V:**
   - Columns of V form an orthonormal basis for the row space of A.

Step-by-Step Example: SVD

Let's compute the SVD of a simple matrix:

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$$

Step 1: Compute $A^T A$

$$A^T A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}^T \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 13 & 12 \\ 12 & 13 \end{bmatrix}$$

Step 2: Find the Eigenvalues of $A^T A$ Solve:

$$\det(A^T A - \lambda I) = 0$$

$$\det \begin{bmatrix} 13 - \lambda & 12 \\ 12 & 13 - \lambda \end{bmatrix} = 0$$

$$(13 - \lambda)^2 - 12^2 = 0 \quad \Rightarrow \quad \lambda^2 - 26\lambda + 25 = 0$$

$$\lambda = 25, 1$$

Step 3: Compute Singular Values The singular values are the square roots of the eigenvalues:

$$\Sigma = \begin{bmatrix} \sqrt{25} & 0 \\ 0 & \sqrt{1} \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}$$

Step 4: Find Eigenvectors for $A^T A$ For $\lambda = 25$:

$$(A^T A - 25I)v = 0$$

$$\begin{bmatrix} -12 & 12 \\ 12 & -12 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \quad \Rightarrow \quad v = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

For $\lambda = 1$:

$$(A^T A - I)v = 0$$

$$\begin{bmatrix} 12 & 12 \\ 12 & 12 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0 \quad \Rightarrow \quad v = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

Thus:

$$V = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

Step 5: Compute U U is obtained using:

$$U = A \cdot V \cdot \Sigma^{-1}$$

---

9. Applications in Machine Learning

SVD is widely used in machine learning for tasks like dimensionality reduction, noise removal, and data compression.

1. Dimensionality Reduction

Principal Component Analysis (PCA) uses SVD to project data onto the directions of maximum variance. The eigenvectors of $A^T A$ define the principal components.

Example: Given a dataset with n features:

- Use SVD to compute Σ, U, and V.
- Retain only the top k singular values in Σ, reducing the data to k-dimensions.

2. Noise Reduction

In image processing:

- Compute the SVD of an image matrix.

- Set smaller singular values to zero (removing noise).
- Reconstruct the matrix using the modified Σ.

3. Least Squares Regression

For a linear regression problem $Ax = b$:

- SVD helps solve the normal equations efficiently:

$$x = V\Sigma^{-1}U^T b$$

Step-by-Step Guide to Applying SVD

1. **Compute SVD:** Decompose A into $U\Sigma V^T$.
2. **Analyze Σ:** Singular values reveal the importance of each component.
3. **Modify Σ:** For dimensionality reduction or noise removal, truncate small singular values.
4. **Reconstruct Data:** Use the truncated matrices to approximate the original matrix.

This detailed explanation ensures that beginners can grasp the concepts of SVD and its applications in machine learning