Machine Learning

Dimensionality Reduction using SVD

Satishkumar L. Varma

Department of Information Technology
SVKM's Dwarkadas J. Sanghvi College of Engineering, Vile Parle, Mumbai.
ORCID | Scopus | Google Scholar | Google Site | Website

Outline

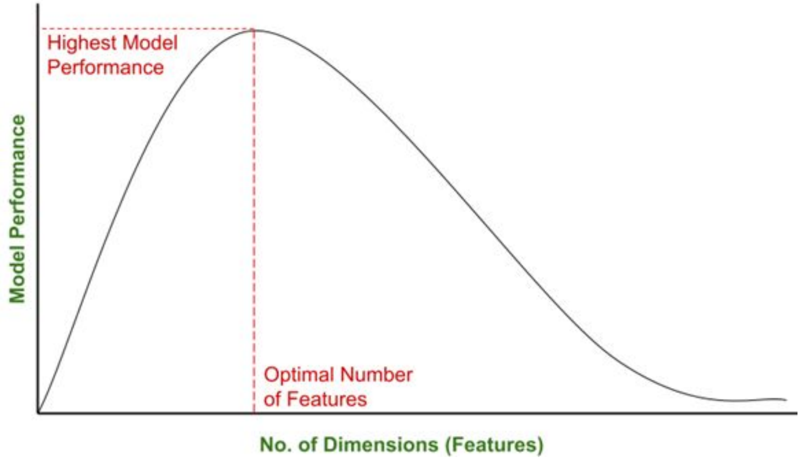
- Dimensionality Reduction
  - Dimensionality Reduction Techniques
    - Principal Components Analysis (Eigenvalues, Eigen vectors, Orthogonality)
    - Independent Component Analysis (ICA)
    - Single Value Decomposition (SVD)

Singular Value Decomposition

- Data Science and ML helps to solve several complex regression and classification problems.
- However, the performance of all these models depends on the input dataset.
- So, It is important to provide optimal dataset to ML models.
- Large dataset leads to increased computational demands
- Large dataset leads to overfitting.
- If we provide large dataset (with a large no. of features/columns) to ML models
  - it gives rise to the problem of overfitting,
  - wherein the model starts getting influenced by outlier values and noise.
  - This is called the Curse of Dimensionality.

Dimensionality Reduction

- Dimensionality Reduction is a statistical/ML-based technique
- It helps to reduce the number of features in our dataset.
- It helps to obtain a dataset with an optimal number of dimensions.
- It is useful in Feature Extraction
  - By reduce the number of dimensions by mapping a higher dimensional feature space to a lower-dimensional feature space.
- Effect of the change in model performance with the increase in the number of dimensions of the dataset.
- The model performance is best only at an option dimension, beyond which it starts decreasing.
- Technique of Dimensionality Reduction (Feature Extraction)
  - Principal Component Analysis (PCA)
  - Independent Component Analysis (ICA)
  - Singular Value Decomposition (SVD)
  - Linear Discriminant Analysis (LDA)



Singular Value Decomposition

- Singular Value Decomposition (SVD)
- It is a method used in linear algebra to decompose a matrix into three simpler matrices,
  - making it easier to analyze and manipulate.
- SVD is one of the important concepts in linear algebra.
- To understand the meaning of SVD, one must be aware of the related concepts such as matrix, types of matrices, transformations of a matrix, etc.
- SVD concept is connected to various concepts of linear algebra,

Singular Value Decomposition

- Example: Understanding SVD
- Imagine you have a table of data, like a set of ratings where rows are people, and columns are products.
- The numbers in the table show how much each person likes each product.
- SVD helps you split that table into three parts:
  - U: This part tells you about the people people's preferences.
  - Σ: This part shows how important each book is (how much each rating matters).
  - V^T: This part tells you about the books (how similar books are to each other)
- So, mathematically, the SVD of a matrix A (of size m×n) is represented as: $A = U \Sigma V^T$
  - U: An m×m orthogonal matrix whose columns are the left singular vectors of A.
  - Σ: A diagonal m×n matrix containing the singular values of A in descending order.
  - V^T: The transpose of an n×n orthogonal matrix, where the columns are the right singular vectors of A.

Rating	Book 1	Book 2
Satish	5	3
Harsh	4	2
Saransh	2	5

Singular Value Decomposition

- Steps to Perform Singular Value Decomposition
  - Step 1: Compute AA^T
  - Step 2: Find the Eigenvalues of AA^T
  - Step 3: Find the Right Singular Vectors (Eigenvectors of A^TA)
  - Step 4: Compute the Left Singular Vectors (Matrix U)
  - Step 5: Final SVD Equation

Singular Value Decomposition

- **Example 1:**
- Step 1: Compute AA^T
- Step 2: Find the Eigenvalues of AA^T
- Thus, the eigenvalues are $\lambda_1=25$ & $\lambda_2=9$
- These eigenvalues correspond to the singular values $\sigma_1=5$ and $\sigma_2=3$
- Since the singular values are the square roots of the eigenvalues.

Step 1: Compute AA^T

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \quad A^T = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \quad AA^T = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

Step 2: Find the Eigenvalues of AA^T

$$\det(AA^T - \lambda I) = 0 \quad \det \begin{bmatrix} 17-\lambda & 8 \\ 8 & 17-\lambda \end{bmatrix} = 0 \quad (\lambda-25)(\lambda-9) = 0$$

Step 3: Find the Right Singular Vectors (Eigenvectors of A^TA for $\lambda=25$ and $\lambda=9$).

For $\lambda = 25$: Solve $(A^TA - 25I)v = 0$:

$$A^TA - 25I = \begin{bmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{bmatrix}$$ Row-reduce this matrix to: $$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The eigenvector corresponding to $\lambda = 25$ is: $v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$

For $\lambda = 9$: Solve $(A^TA - 9I)v = 0$: The eigenvector corresponding to $\lambda = 9$ is: $v_2 = \begin{bmatrix} \frac{1}{\sqrt{18}} \\ \frac{-1}{\sqrt{18}} \\ \frac{4}{\sqrt{18}} \end{bmatrix}$

For the third eigenvector v_3:

Since v_3 must be perpendicular to v_1 and v_2, solve $v_1^T v_3 = 0$ and $v_2^T v_3 = 0$, : $v_3 = \begin{bmatrix} \frac{2}{3} \\ \frac{-2}{3} \\ \frac{-1}{3} \end{bmatrix}$

Singular Value Decomposition

- Example 1:
- Steps to Perform Singular Value Decomposition
  - Step 4: Compute the Left Singular Vectors (Matrix U)
  - Step 5: Final SVD Equation

Step 4: Compute the Left Singular Vectors (Matrix U)

To compute the left singular vectors U, use the formula $u_i = \frac{1}{\sigma_i} A v_i$: $U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$

Step 5: Final SVD Equation

Finally, the Singular Value Decomposition of matrix A is: $A = U\Sigma V^T$

Where: $U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$ $\Sigma = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$ $V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{18}} & \frac{-1}{\sqrt{18}} & \frac{4}{\sqrt{18}} \\ \frac{2}{3} & \frac{-2}{3} & \frac{1}{3} \end{bmatrix}$

Thus, the SVD of matrix A is: $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{18}} & \frac{-1}{\sqrt{18}} & \frac{4}{\sqrt{18}} \\ \frac{2}{3} & \frac{-2}{3} & \frac{1}{3} \end{bmatrix}$

Singular Value Decomposition

- Applications of Singular Value Decomposition (SVD)
- **Rank, Range, and Null Space**
  - The rank, range, and null space of a matrix M can be derived from its SVD.
  - Rank: The rank of matrix M is the number of non-zero singular values in Σ.
  - Range: The range of matrix M is the span of the left singular vectors in matrix U corresponding to the non-zero singular values.
  - Null Space: The null space of matrix M is the span of the right singular vectors in matrix V corresponding to the zero singular values.
- **Digital Signal Processing**
  - SVD can be used to analyze signals and filter noise.
- **Digital Image Processing**
  - SVD is used for image compression and denoising.
  - It helps in reducing the dimensionality of image data by preserving the most significant singular values and discarding the rest.

Singular Value Decomposition

- Key points:
- SVD decomposition function simplifies complex data by breaking it into three smaller parts.
- SVD helps uncover hidden patterns and relationships
- It help to analyze and work with large datasets.
- SVD is useful in tasks like
  - recommendations,
  - data compression, and
  - finding important features,
  - making data simpler

Singular Value Decomposition

- Example 2
- The Singular Value Decomposition of a matrix is a factorization of the matrix into three matrices.
- The SVD of matrix A can be expressed in terms of the factorization of A into the product of three matrices as
  - $A = UDV^T$
- The columns of U and V are orthonormal, and the matrix D is diagonal with real positive entries.

Singular Value Decomposition

- Example 2: Singular Value Decomposition 2×2 Matrix Example

$$A = \begin{bmatrix} -4 & -7 \\ 1 & 4 \end{bmatrix} \quad A^T = \begin{bmatrix} -4 & 1 \\ -7 & 4 \end{bmatrix} \quad AA^T = \begin{bmatrix} -4 & -7 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} -4 & 1 \\ -7 & 4 \end{bmatrix} = \begin{bmatrix} 65 & -32 \\ -32 & 17 \end{bmatrix}$$

Find the Eigenvalues of AA^T

$$|A \cdot A' - \lambda I| = 0 \qquad \begin{vmatrix} (65 - \lambda) & -32 \\ -32 & (17 - \lambda) \end{vmatrix} = 0$$

$$(65 - \lambda) \times (17 - \lambda) - (-32) \times (-32) = 0$$

$$\left(1105 - 82\lambda + \lambda^2\right) - 1024 = 0 \qquad \left(\lambda^2 - 82\lambda + 81\right) = 0$$

$$(\lambda - 1)(\lambda - 81) = 0 \qquad (\lambda - 1) = 0 \text{ or } (\lambda - 81) = 0$$

The eigenvalues of the matrix $A \cdot A'$ are given by $\lambda = 1, 81$.

Eigenvectors for $\lambda = 81$: $v_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ Eigenvectors for $\lambda = 1$: $v_2 = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$

Similarly, find the eigenvectors for A'.A as:

Eigenvectors for $\lambda = 81$: $v_1 = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$ Eigenvectors for $\lambda = 1$: $v_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

Singular Value Decomposition

- Normalize Eigenvectors

Normalize Eigenvectors:

Eigenvectors for λ = 1 : $(0.5, 1)$, Length L = $\sqrt{0.5^2 + 1^2} = 1.118$

So, normalizing gives $v_1 = \left(\frac{0.5}{1.118}, \frac{1}{1.118}\right) = (0.447, 0.894)$

Eigenvectors for λ = 81 : $(-2, 1)$, Length L = $\sqrt{(-2)^2 + 1^2} = 2.236$

So, normalizing gives $v_2 = \left(\frac{-2}{2.236}, \frac{1}{2.236}\right) = (-0.894, 0.447)$

Singular Value Decomposition

- Using these values, we can write the solution as:

$$\Sigma = \begin{bmatrix} \sqrt{81} & 0 \\ 0 & \sqrt{1} \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}$$

$$U = [u_1, u_2] = \begin{bmatrix} -0.894 & 0.447 \\ 0.447 & 0.894 \end{bmatrix}$$

V is found using formula $v_i = \frac{1}{\sigma_i} A^T \cdot u_i$

$$V = \begin{bmatrix} 0.447 & -0.894 \\ 0.894 & 0.447 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{81} & 0 \\ 0 & \sqrt{1} \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}$$

$$V = [v_1, v_2] = \begin{bmatrix} 0.447 & -0.894 \\ 0.894 & 0.447 \end{bmatrix}$$

U is found using formula $u_i = \frac{1}{\sigma_i} A \cdot v_i$

$$U = \begin{bmatrix} -0.894 & 0.447 \\ 0.447 & 0.894 \end{bmatrix}$$

References

Text books:

1. Ethem Alpaydin, “Introduction to Machine Learning , 4th Edition, The MIT Press, 2020.
2. Peter Harrington, “Machine Learning in Action”, 1st Edition, Dreamtech Press, 2012."
3. Tom Mitchell, “Machine Learning”, 1st Edition, McGraw Hill, 2017.
4. Andreas C, Müller and Sarah Guido, “Introduction to Machine Learning with Python: A Guide for Data Scientists”, 1ed, O'reilly, 2016.
5. Kevin P. Murphy, “Machine Learning: A Probabilistic Perspective”, 1st Edition, MIT Press, 2012."

Reference Books:

6. Aurélien Géron, “Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow”, 2nd Edition, Shroff/O'Reilly, 2019.
7. Witten Ian H., Eibe Frank, Mark A. Hall, and Christopher J. Pal., “Data Mining: Practical machine learning tools and techniques”, 1st Edition, Morgan Kaufmann, 2016.
8. Han, Kamber, “Data Mining Concepts and Techniques”, 3rd Edition, Morgan Kaufmann, 2012.
9. Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar, “Foundations of Machine Learning”, 1ed, MIT Press, 2012.
10. H. Dunham, “Data Mining: Introductory and Advanced Topics”, 1st Edition, Pearson Education, 2006.

Thank You.