Singular Value Decomposition Report

Brett Bono

October 2024

1 Introduction

The Singular Value Decomposition (SVD) is one of many matrix decompositions. It was first discovered by differential geometers who sought to determine whether a real bilinear form could be transformed into another by independent orthogonal transformations of the two spaces on which it acts.

The first proof of SVD for rectangular and complex matrices was presented by Carl Eckart and Gale J. Young in 1936. They viewed it as a generalization of the principal axis transformation for Hermitian matrices.

2 Defining SVD

The Singular Value Decomposition is defined as:

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

SVD's fall under 2 categories, Reduced SVD and Full SVD. The difference between the two is minor, however, the reduced SVD is more applicable when using it within computer algorithms

Full SVD The Full SVD is defined as follows:

$A := m \times n$ arbitrary matrix
$U := m \times n$ unitary matrix
$\Sigma := m \times n$ matrix of singular values
$V^T := n \times n$ unitary matrix

In full SVD, Σ contains zero entries to match the dimensions $m \times n$

Reduced SVD The Reduced SVD is defined as follows:

$A :=$ $m \times n$ arbitrary matrix

$U :=$ $m \times n$ unitary Matrix

$\Sigma :=$ $n \times n$ matrix of singular values

$V^T :=$ $n \times n$ unitary Matrix

In reduced SVD, Σ contains only the non-zero singular values, making it smaller, with dimensions $n \times n$.

3 Geometric View

As mentioned earlier, Singular Value Decomposition was originally discovered by differential geometers. The SVD has a very unique geometric representation. The SVD can be viewed geometrically by seeing it as a series of transformations.

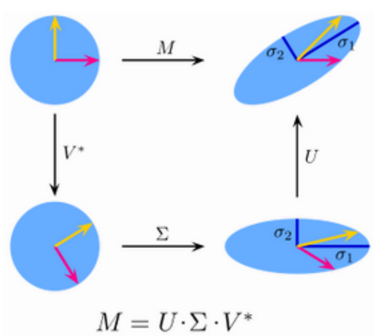
$$A = U\Sigma V^T$$

V^T represents a rotation/reflection in the input space. It maps the original coordinate system of matrix A into a new orthonormal basis.

Applying Σ, the matrix of singular values, scales the space defined by the new axis of V^T

Finally, by applying U the data gets transformed back into the original space.

As seen from a 2×2 matrix where $M = U\Sigma V^*$ is equivalent to $A = U\Sigma V^T$:



4 Applying SVD

Given a matrix A , we will go through the process of getting matrices $U\Sigma V^T$.

Let A be a 2×2 matrix.

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

Step 1: Compute A^TA
First, calculate the matrix A^TA:

$$A^T = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

Now compute:

$$A^TA = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$$

Step 2: Find the Eigenvalues and Eigenvectors of A^TA

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

$$\det\left(\begin{bmatrix} 4-\lambda & 0 \\ 0 & 9-\lambda \end{bmatrix}\right) = (4-\lambda)(9-\lambda) = 0$$

The eigenvalues are:

$$\lambda_1 = 4, \quad \lambda_2 = 9$$

Finding the eigenvector for: $\lambda_1 = 4$:

$$A^TA - 4I = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix}$$

The null space gives the eigenvector:

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Finding the eigenvector for: $\lambda_2 = 9$:

$$A^T A - 9I = \begin{bmatrix} -5 & 0 \\ 0 & 0 \end{bmatrix}$$

The null space gives the eigenvector:

$$v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Step 3: Construct the Matrix V

Matrix V is formed from the normalized eigenvectors of $A^T A$:

$$V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Step 4: Compute AA^T

Now, calculate AA^T:

$$AA^T = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$$

Step 5: Find Eigenvalues and Eigenvectors of AA^T

Using similar steps as for $A^T A$:

The eigenvalues are the same:

$$\lambda_1 = 4, \quad \lambda_2 = 9$$

Finding the eigenvectors for: $\lambda_1 = 4$:

$$AA^T - 4I = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix}$$

The null space gives the eigenvector:

$$u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Finding the eigenvectors for: $\lambda_2 = 9$:

$$AA^T - 9I = \begin{bmatrix} -5 & 0 \\ 0 & 0 \end{bmatrix}$$

The null space gives the eigenvector:

$$u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Step 6: Construct Matrix U
Matrix U is formed from the normalized eigenvectors of AA^T:

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Step 7: Construct Matrix Σ
The singular values are the square roots of the eigenvalues:

$$\sigma_1 = \sqrt{9} = 3, \quad \sigma_2 = \sqrt{4} = 2$$

such that Σ takes the form $\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$ and $\sigma_1 \geq \sigma_2 \geq ...\sigma_n$

Thus, the matrix Σ is:

$$\Sigma = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

Result
Having all matrices, we get:

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}, \quad V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The Singular Value Decomposition is:

$$A = U\Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

5 The Spectral Theorem

The Spectral Theorem states that for any symmetric matrix $A \in \mathbf{R}^{n \times n}$, there exists an orthonormal basis of eigenvectors corresponding to its eigenvalues.

If A is a symmetric matrix, then:

$$A = Q\Lambda Q^T$$

Such that:
Q is an orthogonal matrix whose columns are the normalized eigenvectors of A. Λ is a diagonal matrix containing the eigenvalues of A.

5.1 Connection to Singular Value Decomposition (SVD)

The Singular Value Decomposition (SVD) is a generalization of the Spectral Theorem for any $m \times n$ matrix A. The SVD of A is given by:

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

where:

U is an $m \times m$ orthogonal matrix containing the orthonormal basis for the column space of A

Σ is an $m \times n$ diagonal matrix containing the singular values of A (which are the square roots of the eigenvalues of $A^T A$ or AA^T).

V is an $n \times n$ orthogonal matrix containing the orthonormal basis for the row space of A.

5.2 Relation Between Eigenvalues and Singular Values

For a symmetric matrix A, the eigenvalues obtained from the Spectral Theorem are the same as the singular values obtained from SVD

6 The Fundamental Theorem of Linear Algebra

Given an $m \times n$ matrix A, the fundamental theorem of linear algebra relates to the four fundamental matrix subspaces of A:

1. $\dim R(A) = \dim R(A^T)$ and $\dim R(A) + \dim N(A) = n$, where $R(A)$ denotes the range (or column space) of A, A^T is its transpose, and $N(A)$ is its null space.
2. The null space $N(A)$ is orthogonal to the row space $R(A^T)$.

3. There exist orthonormal bases for both the column space $R(A)$ and the row space $R(A^T)$ of A.

6.1 Connections to SVD

The non-zero singular values in Σ indicate the rank of A.

U and V provide orthonormal bases for the column space and row space.

7 Eckart-Young Theorem

The Eckart-Young theorem states that for any matrix $A \in \mathbf{R}^{m \times n}$ and for any k (where $k < \text{rank}(A)$), the best rank-k approximation of A in terms of the spectral norm is given by the truncated singular value decomposition of A.

Let $A = U\Sigma V^T$ be the SVD of A

Then the best rank-k approximation A_k is given by:

$A_k = U_k \Sigma_k V_k^T$

where,

U_k consists of the first k columns of U

Σ_k is the $k \times k$ diagonal matrix containing the largest k singular values,

V_k consists of the first k columns of V.

The approximation error in the spectral norm can be expressed as:

$$\|A - A_k\|_2 = \sigma_{k+1}$$

where σ_{k+1} is the smallest singular value not included in A_k.

Proof

show that for any matrix $A \in \mathbf{R}^{m\times n}$ and for any $k < \text{rank}(A)$, the best rank-k approximation A_k satisfies:

$$\|A - A_k\|_2 = \sigma_{k+1}$$

$$A_k = U_k \Sigma_k V_k^T$$

where U_k consists of the first k columns of U, Σ_k contains the largest k singular values, and V_k consists of the first k columns of V.

The error can be expressed as:

$$A - A_k = U\Sigma V^T - U_k \Sigma_k V_k^T$$

$$\|A - A_k\|_2 = \|U\Sigma V^T - U_k \Sigma_k V_k^T\|_2$$

We can rewrite $A - A_k$:

$$A - A_k = U\Sigma V^T - U_k \Sigma_k V_k^T = U\Sigma V^T - U_k \Sigma_k V_k^T + U_k \Sigma_k V_k^T - U_k \Sigma_k V_k^T$$

$$A - A_k = U\Sigma V^T - U_k \Sigma_k V_k^T + (U - U_k)\Sigma V^T$$

Applying properties of orthogonal matrices,

$$\|A - A_k\|_2 = \|U(\Sigma - \Sigma_k V^T)\|_2$$

Since U is orthogonal, we have:

$$\|A - A_k\|_2 = \|\Sigma - \Sigma_k V^T\|_2$$

The spectral norm is defined as the largest singular value of the matrix. The remaining singular values after truncation correspond to σ_{k+1}, as the spectral norm of the difference $A - A_k$ will be equal to the smallest singular value of the remaining part, which is:

$$\|A - A_k\|_2 = \sigma_{k+1}$$

■

8 Application

The applications of SVD are large.To almost anything data related you can get its SVD decomposition. There is a large variety of industries who use SVD.

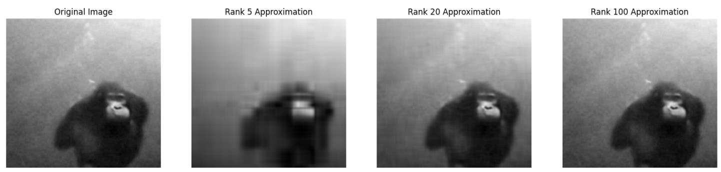
Below are examples of how SVD can be used within different areas.

8.1 Image Compression

Given a digital image, Singular Value Decomposition can be performed on its corresponding numeric matrix. For a colored image, the image is typically represented as three separate matrices: one for the red channel, one for the green channel, and one for the blue channel. Each of these color matrices can undergo SVD independently.

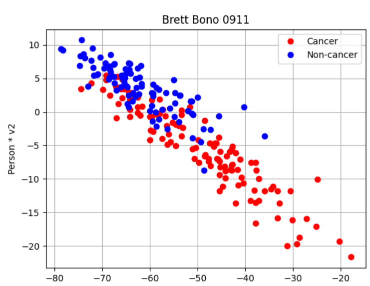
With this decomposition, a rank-k approximation can be constructed by retaining only the k largest singular values, thereby capturing the most significant information in the image. This process allows for a compressed approximation of the image, reducing its complexity while preserving its essential features. The higher the rank k, the closer the approximation is to the original image.

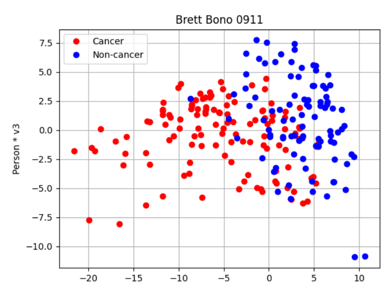
As Shown:



8.2 Ovarian Cancer Data

Given observation data on ovarian cancer, SVD can be applied for dimensionality reduction. By applying SVD, we can project the high-dimensional data onto lower-dimensional spaces, allowing for the visualization in multiple 2D graphs. This allows us to better understand the structure of the data and observe potential patterns, as shown in the following plots:





8.3 Watermarking

Through the *Liu and Tan: Embedding and Extraction* algorithms, a watermark could be embedded into an image with minimal alteration to the host image. Using a specific key, the watermark can later be extracted, ensuring the protection of digital images.

Both algorithms described:

Algorithm 1 Liu & Tan: Embedding $(A_W, K) = E(A, W, \alpha)$

Take the SVD of data matrix: $A \to USV^T$
Add αW to S and compute the SVD of their sum: $S + \alpha W \to U_W S_W V_W^T$
Replace S with S_W to reconstruct watermarked matrix: $A_W \leftarrow U S_W V^T$
Save the keys: $K \leftarrow (S, U_W, V_W, \alpha)$
return (A_W, K)

Algorithm 2 Liu & Tan: Extraction $\tilde{W} = X(\tilde{A_W}, K)$

$(S, U_W, V_W, \alpha) \leftarrow K$
Take the SVD of $\tilde{A_W}$: $\tilde{A_W} \to \tilde{U}\tilde{S_W}\tilde{V}^T$
Construct $\tilde{D}$: $\tilde{D} \leftarrow U_W \tilde{S_W} V_W^T$
Reconstruct watermarked matrix: $\tilde{W} \leftarrow \frac{1}{\alpha}(\tilde{D} - S)$
return $\tilde{W}$

Below are the outputs of the two algorithms:

Figure 1: Original Image

Figure 2: Watermark Image

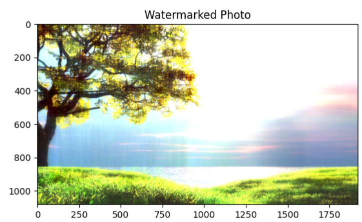


Figure 3: Watermarked Photo with $\alpha = 200$

8.4 Noise Reduction in signal processing (Truncatation of Small Singular Values)

Similar to image approximation, where the quality of the approximation improves rapidly by capturing the largest singular values first, noise reduction follows the same idea. By retaining the most dominant singular values, which represent the core structure of the signal, unwanted noise can be reduced.

Once we truncate the smaller singular values and retain the top k singular values, we can reconstruct the denoised signal matrix:

$$A_k = U_k \Sigma_k V_k^T$$

where:

- U_k contains the first k columns of U.
- Σ_k is a diagonal matrix with the top k singular values.
- V_k^T contains the first k rows of V^T.

By reconstructing the signal with the top k singular values, we remove the noise while preserving the structure of the signal.

Algorithm 3 Noise Reduction Using SVD

Input: Noisy signal matrix A, threshold k
Output: Denoised signal matrix A_k
Perform SVD on matrix A:

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

Retain only the top k singular values:

$$\Sigma_k = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k)$$

Reconstruct the denoised signal matrix:

$$A_k = U_k \Sigma_k V_k^T$$

return A_k

8.5 Observing Stock Trends

By using the same pseudo algorithm for noise reduction but applied to stock market data, SVD helps identify dominant features within the market. The key singular values correspond to principal components or market factors that explain the most variance in stock prices or returns. The smaller singular values, which may correspond to noise or less significant factors, can be discarded to focus on market trends.

By capturing the dominant trends and discarding noise, SVD facilitates clearer insights into market dynamics, aiding in portfolio optimization, risk management, and trend analysis.

It is common for analysts to apply SVD to a matrix of historical stock prices where the rows represent different stocks and the columns represent time periods. The resulting decomposition identifies principal market trends such as overall market growth, and sector-specific movements.