Proof of the Singular Value Decomposition (SVD)

Zakk Heile

December 2024

1 Introduction

The Singular Value Decomposition (SVD) is a fundamental theorem in linear algebra with numerous applications in areas such as statistics, signal processing, and machine learning. This document provides a detailed proof of the SVD, demonstrating that any $m \times n$ matrix A can be decomposed into the product of three matrices: $A = U\Sigma V^\top$, where U and V are orthogonal matrices, and Σ is a diagonal matrix containing the singular values of A.

2 Preliminaries

Let A be an $m \times n$ matrix. Consider the Gram matrices $AA^\top$ and $A^\top A$. Both of these matrices are symmetric and positive semi-definite (PSD), implying that their eigenvalues are non-negative.

2.1 Positive Semi-Definiteness of Gram Matrices

Proof of Non-Negative Eigenvalues. A Gram matrix $G = M^\top M$ is always positive semi-definite. For any vector $x \in \mathbb{R}^n$, we have:

$$x^\top G x = x^\top M^\top M x = (Mx)^\top (Mx) = \|Mx\|^2 \geq 0.$$

Here, $\|Mx\|^2$ represents the squared Euclidean norm of the vector Mx, which is always non-negative. $\square$

3 Eigenvalues of $A^\top A$ and $AA^\top$

Both $A^\top A$ and $AA^\top$ are symmetric and PSD matrices, hence they have real, non-negative eigenvalues. Let us denote the eigenvalues of $A^\top A$ by σ_i^2 and those of $AA^\top$ by the same σ_i^2. The non-zero eigenvalues of these two Gram matrices are identical, and the remaining eigenvalues are zero.

3.1 Equivalence of Non-Zero Eigenvalues

Direction 1: From $A^\top A$ to $AA^\top$

Assume $A^\top A x = \lambda x$ with $\lambda \neq 0$ and $x \neq 0$.

$$AA^\top (Ax) = A(A^\top A x) = A(\lambda x) = \lambda(Ax).$$

Define $y = Ax$. Then:

$$AA^\top y = \lambda y.$$

Since $\lambda \neq 0$ and $x \neq 0$, $y \neq 0$. Thus, y is an eigenvector of $AA^\top$ corresponding to the eigenvalue λ.

Direction 2: From $AA^\top$ to $A^\top A$

Assume $AA^\top y = \mu y$ with $\mu \neq 0$ and $y \neq 0$.

$$A^\top A(A^\top y) = A^\top (AA^\top y) = A^\top(\mu y) = \mu(A^\top y).$$

Define $x = A^\top y$. Then:

$$A^\top A x = \mu x.$$

Since $\mu \neq 0$ and $y \neq 0$, $x \neq 0$. Thus, x is an eigenvector of $A^\top A$ corresponding to the eigenvalue μ.

3.2 Multiplicity of Zero Eigenvalues

The eigenvalue 0 corresponds to the nullspace of A and $A^\top$. The multiplicity of the eigenvalue 0 in $A^\top A$ is equal to the dimension of the nullspace of A, and similarly for $AA^\top$.

4 Constructing the SVD

Consider an eigenvector-eigenvalue pair of $A^\top A$:

$$A^\top A v_i = \sigma_i^2 v_i.$$

Define $u_i = \dfrac{Av_i}{\sigma_i}$. We claim that u_i is a unit eigenvector of $AA^\top$. Note that we are assuming the singular value is not 0.

Proof that u_i is an Eigenvector of $AA^\top$.

$$AA^\top u_i = AA^\top \left(\frac{Av_i}{\sigma_i} \right) = \frac{AA^\top A v_i}{\sigma_i} = \frac{A\sigma_i^2 v_i}{\sigma_i} = \sigma_i A v_i = \sigma_i^2 u_i.$$

Thus, u_i satisfies $AA^\top u_i = \sigma_i^2 u_i$, making it an eigenvector of $AA^\top$ with eigenvalue σ_i^2. □

Proof that u_i is a Unit Vector.

$$u_i^\top u_i = \left(\frac{Av_i}{\sigma_i} \right)^\top \left(\frac{Av_i}{\sigma_i} \right) = \frac{v_i^\top A^\top A v_i}{\sigma_i^2} = \frac{v_i^\top (\sigma_i^2 v_i)}{\sigma_i^2} = v_i^\top v_i = 1.$$

Thus, u_i is a unit vector. □

4.1 Forming the Orthogonal Matrices

From the above constructions, we have:

$$U = AV\Sigma^{-1},$$

where V is the matrix whose columns are the eigenvectors v_i of $A^\top A$, and Σ is the diagonal matrix with entries σ_i. Since U consists of orthonormal vectors u_i, it is an orthogonal matrix.

Similarly, V is orthogonal by construction.

4.2 A Note on Rank

How do we know that the matrices are orthogonal and sufficient rank, namely that V has an inverse? Additionally, we are still assuming no singular values are 0 so Σ has an inverse.

If A is injective then $v_i \mapsto u_i$ is an injective map.

Formally, suppose for the sake of contradiction:

$$A^\top A v_1 = \sigma_1^2 v_1$$

$$A^\top A v_2 = \sigma_2^2 v_2$$

$$Av_1 = Av_2$$

Then, $Av_1 - Av_2 = 0$ and as they are equal, both of them must be 0, meaning both v_1 and v_2 are in the null space of A. This means they are 0-eigenvalue eigenvectors of A, meaning they are $\sigma_i^2 = 0^2 = 0$ eigenvalue eigenvectors of their Gram matrices.

Thus, we only have to worry about this if the diagonal matrix has non-trivial null space.

If Σ has a non-trivial null space, we extend U and V to full orthogonal matrices by adding orthonormal vectors that span the nullspaces of $A^\top$ and A, respectively. These additional vectors correspond to singular values of zero, ensuring that U and V remain orthogonal regardless of the rank of A.

5 Conclusion: The SVD

Putting everything together, we obtain the Singular Value Decomposition of A:

$$A = U\Sigma V^\top,$$

where:

- U is an $m \times m$ orthogonal matrix whose columns are the eigenvectors of $AA^\top$,
- Σ is an $m \times n$ diagonal matrix with non-negative real numbers σ_i on the diagonal,
- V is an $n \times n$ orthogonal matrix whose columns are the eigenvectors of $A^\top A$.

This decomposition reveals the intrinsic geometric structure of the matrix A, facilitating various applications in numerical analysis, data compression, and beyond.

6 Rank Considerations

Understanding the rank of a matrix and its Gram matrices is crucial, especially when dealing with matrices that are not full rank. The following section elucidates the relationship between the ranks of A, $A^\top A$, and $AA^\top$.

6.1 Equivalence of Ranks

If we can show that $\text{rank}(A) = \text{rank}(A^\top A)$, then because $\text{rank}(A) = \text{rank}(A^\top)$, it follows that:

$$\text{rank}(A) = \text{rank}(A^\top) = \text{rank}(AA^\top).$$

Proof that rank$(A) =$ *rank*$(A^\top A)$

We need to show that the solution sets of $Ax = 0$ and $A^\top Ax = 0$ are identical, i.e., $Ax = 0 \iff A^\top Ax = 0$.

Proof. **Forward Direction:** Assume $Ax = 0$. Then:

$$A^\top(Ax) = A^\top 0 = 0,$$

which shows $A^\top Ax = 0$.

Conversely: Assume $A^\top Ax = 0$. Then:

$$x^\top(A^\top Ax) = 0.$$

Expanding this, we have:

$$(Ax)^\top(Ax) = \|Ax\|^2 = 0.$$

By the definition of the inner product, for any vector v, $v^\top v = 0 \iff v = 0$. Thus, $Ax = 0$.

Therefore, $Ax = 0 \iff A^\top Ax = 0$, which implies that:

$$\text{null}(A) = \text{null}(A^\top A).$$

□

Since the nullspaces are identical, their dimensions are equal. Consequently, the ranks satisfy:

$$\text{rank}(A) = \text{rank}(A^\top A).$$

Similarly, by considering $A^\top$, we can show that:

$$\text{rank}(A) = \text{rank}(A^\top) = \text{rank}(AA^\top).$$

7 Singular Values and Rank

Given that $\text{rank}(A) = r$, where $r \leq \min(m, n)$, the SVD of A can be expressed as:

$$A = U\Sigma V^\top,$$

where:

- U is an $m \times m$ orthogonal matrix whose first r columns are the eigenvectors of $AA^\top$ corresponding to the non-zero singular values, and the remaining $m - r$ columns span the nullspace of $A^\top$.
- Σ is an $m \times n$ diagonal matrix with the first r diagonal entries being the positive singular values $\sigma_1, \sigma_2, \ldots, \sigma_r$ and the remaining entries being zero.
- V is an $n \times n$ orthogonal matrix whose first r columns are the eigenvectors of $A^\top A$ corresponding to the non-zero singular values, and the remaining $n - r$ columns span the nullspace of A.