SINGULAR VALUE DECOMPOSITION

Andrew Lounsbury

September 2018

No. 2018-2
Abstract

The Singular Value Decomposition (SVD) provides a cohesive summary of a handful of topics introduced in basic linear algebra. SVD may be applied to digital photographs so that they may be approximated and transmitted with a concise computation.

Mathematics Subject Classification. Primary 15A23, 15A24
Keywords. singular value, matrix factorization

Contents

1 Introduction 1
2 Steps for Calculation of SVD 3
3 Theory 5
4 Examples 9
5 Maple 12
6 Conclusions 16
A Appendix 16

1 Introduction

This paper begins with a definition of SVD and instructions on how to compute it, which includes calculating eigenvalues, singular values, eigenvectors, left and right singular vectors, or, alternatively, orthonormal bases for the four fundamental spaces of a matrix. We present two theorems

*Completed in partial fulfillment of the requirements for MATH 4993 Mathematical Research (3 cr) in spring 2018 under supervision of Dr. R. Ablamowicz.
that result from SVD with corresponding proofs. We provide examples of matrices and their singular value decompositions. There is also a section involving Maple that includes examples of photographs. It is demonstrated how the inclusion of more and more information from the SVD allows one to construct accurate approximations of a color image.

Definition 1. Let $A$ be an $m \times n$ real matrix of rank $r \leq \min(m, n)$. A **Singular Value Decomposition (SVD)** is a way to factor $A$ as

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

where $U$ and $V$ are orthogonal matrices such that $U^TU = I_m$ and $V^TV = I_n$. The $\Sigma$ matrix contains the singular values of $A$ on its pseudo-diagonal, with zeros elsewhere. Thus,

$$A = U \Sigma V^T = \begin{bmatrix} u_1 & u_2 & \cdots & u_m \end{bmatrix}_{U(m \times m)} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \sigma_r & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \ddots & \vdots \\ \end{bmatrix}_{\Sigma(m \times n)} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \\ \end{bmatrix}_{V^T(n \times n)},$$

with $u_1, \ldots, u_m$ being the orthonormal columns of $U$, $\sigma_1, \ldots, \sigma_r$ being the singular values of $A$ satisfying $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and $v_1, \ldots, v_n$ being the orthonormal columns of $V^T$. Singular values are defined as the positive square roots of the eigenvalues of $A^TA$.

Note that since $A^TA$ of size $n \times n$ is real and symmetric of rank $r$, $r$ of its eigenvalues $\sigma_i^2$, $i = 1, \ldots, r$, are positive and therefore real, while the remaining $n - r$ eigenvalues are zero. In particular,

$$A^TA = V(\Sigma^T \Sigma)V^T, \quad A^TA v_i = \sigma_i^2 v_i, \quad i = 1, \ldots, r, \quad A^TA v_i = 0, \quad i = r + 1, \ldots, n. \quad (3)$$

Thus, the first $r$ vectors $v_i$ are the eigenvectors of $A^TA$ with the eigenvalues $\sigma_i^2$. Likewise, we have

$$AA^T = U(\Sigma \Sigma^T)U^T, \quad AA^T u_i = \sigma_i^2 u_i, \quad i = 1, \ldots, r, \quad AA^T u_i = 0, \quad i = r + 1, \ldots, m. \quad (4)$$

Thus, the first $r$ vectors $u_i$ are the eigenvectors of $AA^T$ with the eigenvalues $\sigma_i^2$.

Furthermore, it can be shown (see Lemmas 1 and 2) that

$$Av_i = \sigma_i u_i, \quad i = 1, \ldots, r, \quad \text{and} \quad Av_i = 0, \quad i = r + 1, \ldots, n. \quad (5)$$

If rank$(A) = r < \min(m, n)$, then there are $n - r$ zero columns and rows in $\Sigma$, rendering the
\[ A = U \Sigma V^T = \begin{bmatrix} u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \sigma_r & \ddots & \vdots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & \sigma_r \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_r^T \\ v_{r+1}^T \\ \vdots \\ v_n^T \end{bmatrix} \]

\[ = u_1 \sigma_1 v_1 + \cdots + u_r \sigma_r v_r. \tag{6} \]

### 2 Steps for Calculation of SVD

Here, we provide an algorithm to calculate a singular value decomposition of a matrix.

1. Compute \( A^T A \) of a real \( m \times n \) matrix \( A \) of rank \( r \).
2. Compute the singular values of \( A^T A \).
   
   Solve the characteristic equation \( \Delta_{A^T A}(\lambda) = |A^T A - \lambda I| = 0 \) of \( A^T A \) for the eigenvalues \( \lambda_1, \ldots, \lambda_r \) of \( A^T A \). These eigenvalues will be positive. Take their square roots to obtain \( \sigma_1, \ldots, \sigma_r \) which are the singular values of \( A \), that is,
   \[ \sigma_i = +\sqrt{\lambda_i}, \quad i = 1, \ldots, r. \tag{7} \]
3. Sort the singular values, possibly renaming them, so that \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \).
4. Construct the \( \Sigma \) matrix of size \( m \times n \) such that \( \Sigma_{ii} = \sigma_i \) for \( i = 1, \ldots, r \), and \( \Sigma_{ij} = 0 \) when \( i \neq j \).
5. Compute the eigenvectors of \( A^T A \).
   
   Find a basis for \( \text{Null}(A^T A - \lambda_i I) \). That is, solve \( (A^T A - \lambda_i I)s_i = 0 \) for \( s_i \), an eigenvector of \( A \) corresponding to \( \lambda_i \), for each eigenvalue \( \lambda_i \). Since \( A^T A \) is symmetric, its eigenvectors corresponding to different eigenvalues are already orthogonal (but likely not orthonormal). See Lemma 1.
6. Compute the (right singular) vectors \( v_1, \ldots, v_r \) by normalizing each eigenvector \( s_i \) by multiplying it by \( \frac{1}{\|s_i\|} \). That is, let
   \[ v_i = \frac{1}{\|s_i\|} s_i, \quad i = 1, \ldots, r. \tag{8} \]
If \( n > r \), the additional \( n - r \) vectors \( v_{r+1}, \ldots, v_n \) need to be chosen as an orthonormal basis in \( \text{Null}(A) \). Note that since \( Av_i = \sigma_i u_i \) for \( i = 1, \ldots, r \), vectors \( v_1, \ldots, v_r \) provide an orthonormal basis for \( \text{Row}(A) \) while the vectors \( u_1, \ldots, u_r \) provide an orthonormal basis for \( \text{Col}(A) \). In particular,

\[
\mathbb{R}^n = \text{Row}(A) \perp \text{Null}(A) = \text{span}\{v_1, \ldots, v_r\} \perp \text{span}\{v_{r+1}, \ldots, v_{r+(n-r)}\}.
\]

7. Construct the orthogonal matrix \( V = [v_1 \cdots v_n] \).

8. Verify \( V^TV = I \).

9. Compute the (left singular) vectors \( u_1, \ldots, u_r \) as

\[
Av_i = \sigma_i u_i \implies u_i = \frac{Av_i}{\sigma_i}, \quad i = 1 \ldots r.
\]

In this method, \( u_1, \ldots, u_r \) are orthogonal by Lemma 5.

Alternatively,

(i) Note that \( AA^T = U(S\Sigma S^T)U^T \) suggests the vectors of \( U \) can be calculated as the eigenvectors of \( AA^T \). In using this method, the vectors need to be normalized first. Namely, \( u_i = \frac{1}{||s_i||} s_i \), where \( s_i \) is an eigenvector of \( AA^T \).

(ii) Since \( \Delta_{AA^T}(\lambda) = \Delta_{AA^T}(\lambda) \) by Lemma 8, \( \sigma_1, \ldots, \sigma_r \) are also the square roots of the eigenvalues of \( AA^T \).

If \( m > r \), the additional \( m - r \) vectors \( u_{r+1}, \ldots, u_m \) need to be chosen as an orthonormal basis in \( \text{Null}(A^T) \). Note that since \( Av_i = \sigma_i u_i \) for \( i = 1, \ldots, r \), vectors \( u_1, \ldots, u_r \) provide an orthonormal basis for \( \text{Col}(A^T) \) while the vectors \( u_{r+1}, \ldots, u_m \) provide an orthonormal basis for the left null space \( \text{Null}(A^T) \). In particular,

\[
\mathbb{R}^m = \text{Col}(A) \perp \text{Null}(A^T) = \text{span}\{u_1, \ldots, u_r\} \perp \text{span}\{u_{r+1}, \ldots, u_{r+(m-r)}\}.
\]

10. Construct \( U = [u_1 \cdots u_m] \).

11. Verify \( U^TU = I \).

12. Verify \( A = USV^T \).

13. Construct the dyadic decomposition\(^1\) of \( A \), as described in Thm. 13:

\[
A = USV^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + u_r \sigma_r v_r^T.
\]

\(^1\)A dyad is a product of an \( n \times 1 \) column vector with another \( 1 \times n \) row vector, e.g., \( u_1 v_1^T \), resulting in a square \( n \times n \) matrix whose rank is 1 by Lemma 6.
3 Theory

In this section, we provide the two theorems related to SVD along with their proofs.

**Theorem 1.** Let \( A = U \Sigma V^T \) be a singular value decomposition of an \( m \times n \) real matrix of rank \( r \). Then,

1. \( AV = U \Sigma \) and
   \[
   \begin{align*}
   \left\{ \begin{array}{l}
   Av_i = \sigma_i u_i, \quad i = 1, \ldots, r \\
   Av_i = 0, \quad i = r + 1, \ldots, r + (n - r)
   \end{array} \right. \implies \left\{ \begin{array}{l}
   \text{Row}(A) = \text{span}\{v_1, \ldots, v_r\} \\
   \text{Null}(A) = \text{span}\{v_{r+1}, \ldots, v_{r+(n-r)}\}
   \end{array} \right.
   \]

2. \( A^T A = V (\Sigma^T \Sigma) V^T : \mathbb{R}^n \to \mathbb{R}^n \)

3. \( A^T AV = V (\Sigma^T \Sigma) \) and
   \[
   \begin{align*}
   \left\{ \begin{array}{l}
   A^T Av_i = \sigma_i^2 v_i, \quad i = 1, \ldots, r \\
   A^T Av_i = 0, \quad i = r + 1, \ldots, r + (n - r)
   \end{array} \right. \implies \left\{ \begin{array}{l}
   \text{Row}(A^T A) = \text{span}\{v_1, \ldots, v_r\} \\
   \text{Null}(A^T A) = \text{span}\{v_{r+1}, \ldots, v_{r+(n-r)}\}
   \end{array} \right.
   \]

4. \( U^T A = \Sigma V^T \) and
   \[
   \begin{align*}
   \left\{ \begin{array}{l}
   u_i^T A = \sigma_i v_i, \quad i = 1, \ldots, r \\
   u_i^T A = 0, \quad i = r + 1, \ldots, r + (m - r)
   \end{array} \right. \implies \left\{ \begin{array}{l}
   \text{Col}(A) = \text{span}\{u_1, \ldots, u_r\} \\
   \text{Null}(A^T) = \text{span}\{u_{r+1}, \ldots, u_{r+(m-r)}\}
   \end{array} \right.
   \]

5. \( A A^T = U (\Sigma \Sigma^T) U^T : \mathbb{R}^m \to \mathbb{R}^m \)

6. \( A A^T U = U (\Sigma \Sigma^T) \) and
   \[
   \begin{align*}
   \left\{ \begin{array}{l}
   A A^T u_i = \sigma_i^2 u_i, \quad i = 1, \ldots, r \\
   A A^T u_i = 0, \quad i = r + 1, \ldots, r + (m - r)
   \end{array} \right. \implies \left\{ \begin{array}{l}
   \text{Row}(AA^T) = \text{span}\{u_1, \ldots, u_r\} \\
   \text{Null}(AA^T) = \text{span}\{u_{r+1}, \ldots, u_{r+(m-r)}\}
   \end{array} \right.
   \]

**Proof of (1).**

\[
AV = (U \Sigma V^T)V = U \Sigma (V^TV) = U \Sigma
\]

So,

\[
AV = \begin{bmatrix}
    Av_1 & \cdots & Av_r & Av_{r+1} & \cdots & Av_n
\end{bmatrix}
\begin{bmatrix}
    \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\
    0 & \ddots & & \vdots & & \vdots \\
    \vdots & & \ddots & \vdots & & \vdots \\
    0 & 0 & \cdots & \sigma_r & & \vdots \\
    0 & 0 & \cdots & 0 & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
    u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m
\end{bmatrix}
\]

\[
= \begin{bmatrix}
    \sigma_1 u_1 & \cdots & \sigma_r u_r & 0 & \cdots & 0
\end{bmatrix}
\]

Hence,
1. $Av_1 = \sigma_1 u_1$, ..., $Av_r = \sigma_r u_r$, and 
2. $Av_{r+1} = 0$, ..., $Av_{r+(n-r)} = 0$.

**Proof of (2).**

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

**Proof of (3).**

$$A^T AV = (V\Sigma^T U^T) (U\Sigma V^T) V = V\Sigma^T (U^T U) \Sigma (V^T V) = V(\Sigma^T \Sigma)$$

So,

$$A^T AV = \begin{bmatrix} A^T Av_1 | & \cdots | & A^T Av_r | & A^T Av_{r+1} | & \cdots | & A^T Av_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 | & \cdots | & \lambda_r v_r | & \lambda_{r+1} v_{r+1} | & \cdots | & \lambda_n v_n \end{bmatrix}$$

$$= \begin{bmatrix} v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & & \\
\vdots & & \lambda_r \\
0 & 0 & \ddots \\
0 & 0 & \cdots & 0 \\
\sigma_1^2 & 0 & \cdots & 0 \\
0 & \ddots & \vdots & \\
\vdots & \sigma_r^2 & \ddots & \\
0 & 0 & \ddots \\
0 & 0 & \cdots & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1^2 v_1 & \cdots & \sigma_r^2 v_r & 0 & \cdots & 0 \end{bmatrix}.$$ 

Hence,

1. $A^T Av_1 = \sigma_1^2 v_1$, ..., $A^T Av_r = \sigma_r^2 v_r$, and 
2. $A^T Av_{r+1} = 0$, ..., $A^T Av_{r+(n-r)} = 0$.

**Proof of (4).**

$$U^T A = U^T (U\Sigma V^T) = (U^T U) \Sigma V^T = \Sigma V^T$$
So,

\[
U^T A = \begin{bmatrix}
    u_1^T \\
    \vdots \\
    u_r^T \\
    u_{r+1}^T \\
    \vdots \\
    u_m^T
\end{bmatrix} \quad A = \begin{bmatrix}
    u_1^T A \\
    \vdots \\
    u_r^T A \\
    u_{r+1}^T A \\
    \vdots \\
    u_m^T A
\end{bmatrix}
\]

\[
= \begin{bmatrix}
    \sigma_1 & 0 & \cdots & \cdots & 0 & \cdots & 0 \\
    0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
    \vdots & \sigma_r & \ddots & \ddots & \ddots & \ddots & \ddots \\
    0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots \\
    0 & 0 & \cdots & \ddots & 0 & \cdots & 0
\end{bmatrix} \begin{bmatrix}
    v_1^T \\
    \vdots \\
    v_r^T \\
    v_{r+1}^T \\
    \vdots \\
    v_n^T
\end{bmatrix} = \begin{bmatrix}
    \sigma_1 v_1^T \\
    \vdots \\
    \sigma_r v_r^T \\
    0 \\
    \vdots \\
    0
\end{bmatrix}
\]

Hence,

1. \( u_1^T A = \sigma_1 v_1^T, \ldots, u_r^T A = \sigma_r v_r^T \), and

2. \( u_{r+1}^T A = 0, \ldots, u_{r+(m-r)}^T A = 0 \).

**Proof of (5).**

\[
AA^T = (U \Sigma V^T)(U \Sigma V^T)^T = (U \Sigma V^T)(V \Sigma^T U^T) = U \Sigma^T (V^T V) \Sigma U^T = U \Sigma \Sigma U^T
\]

**Proof of (6).**

\[
AA^T U = (U \Sigma V^T)(U \Sigma V^T)^T U = (U \Sigma V^T)(V \Sigma^T U^T) U = U \Sigma (V^T V) \Sigma^T (U^T U) = U (\Sigma \Sigma)
\]
So,

\[
AA^T U = \begin{bmatrix}
AA^T u_1 & \cdots & AA^T u_r & AA^T u_{r+1} & \cdots & AA^T u_m
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\lambda_1 u_1 & \cdots & \lambda_r u_r & \lambda_{r+1} u_{r+1} & \cdots & \lambda_m u_m
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & & \\
\vdots & & \lambda_r & \\
0 & 0 & \ddots & 0 \\
\sigma_1^2 & 0 & \cdots & 0 \\
0 & \ddots & & \\
0 & 0 & \cdots & \sigma_r^2 \\
0 & 0 & \cdots & 0
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\sigma_1^2 u_1 & \cdots & \sigma_r^2 u_r & 0 & \cdots & 0
\end{bmatrix}.
\]

Hence,

1. \(AA^T u_1 = \sigma_1^2 u_1, \ldots, AA^T u_r = \sigma_r^2 u_r\), and
2. \(AA^T u_{r+1} = 0, \ldots, AA^T u_{r+(m-r)} = 0\).

\[\square\]

**Theorem 2.** Let \(A = U \Sigma V^T\) be a singular value decomposition of an \(m \times n\) real matrix of rank \(r\). Then,

\[
A = U \Sigma V^T = \sum_{i=1}^{r} u_i \sigma_i v_i^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T.
\] (13)
Proof.

\[ A = U \Sigma V^T \]

\[

\begin{bmatrix}
\sigma_1 u_{11} & \sigma_2 u_{12} & \ldots & \sigma_r u_{1r} \\
\sigma_1 u_{21} & \sigma_2 u_{22} & \ldots & \sigma_r u_{2r} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_1 u_{(r-1)1} & \sigma_2 u_{r2} & \ldots & \sigma_r u_{rr}
\end{bmatrix}
\begin{bmatrix}
v_{11}^T & v_{12}^T & \ldots & v_{1r}^T \\
v_{21}^T & v_{22}^T & \ldots & v_{2r}^T \\
\vdots & \vdots & \ddots & \vdots \\
v_{r1}^T & v_{r2}^T & \ldots & v_{rr}^T
\end{bmatrix}
\]

\[ = \sigma_1 u_{11}v_{11}^T + \cdots + \sigma_r u_{1r}v_{1r}^T \]

\[ (\sigma_1 u_{11}v_{11}^T + \cdots + \sigma_r u_{1r}v_{1r}^T) (\sigma_1 u_{12}v_{12}^T + \cdots + \sigma_r u_{2r}v_{2r}^T) \cdots (\sigma_1 u_{11}v_{11r}^T + \cdots + \sigma_r u_{1r}v_{rr}^T) \]

\[ = \sigma_1 u_{11}v_{11}^T + \sigma_1 u_{22}v_{22}^T + \cdots + \sigma_r u_{rr}v_{rr}^T \]

\[ = \sum_{i=1}^{r} u_i \sigma_i v_i^T \]

4 Examples

In this section we calculate the singular value decomposition of a few matrices.

Example 1. Let \( A = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 \\ 4 & 2 & 2 & 2 \end{bmatrix}, \) then \( A^T A = \begin{bmatrix} 21 & 10 & 11 & 12 \\ 10 & 5 & 5 & 5 \\ 11 & 5 & 6 & 7 \\ 12 & 5 & 7 & 9 \end{bmatrix}, \) and
\[ \Delta_{A^T A}(\lambda) = \begin{bmatrix} 21 - \lambda & 10 & 11 & 12 \\ 10 & 5 - \lambda & 5 & 5 \\ 11 & 5 & 6 - \lambda & 7 \\ 12 & 5 & 7 & 9 - \lambda \end{bmatrix} = \lambda^4 - 41\lambda^3 + 85\lambda^2 \implies \lambda_1 = \frac{41 + 3\sqrt{149}}{2}, \quad \lambda_2 = \frac{41 - 3\sqrt{149}}{2} \]

\[ \sigma_1 = \frac{\sqrt{41 + 3\sqrt{149}}}{\sqrt{2}} \quad \sigma_2 = \frac{\sqrt{41 - 3\sqrt{149}}}{\sqrt{2}} \implies \Sigma = \begin{bmatrix} \sqrt{\frac{82 + 6\sqrt{149}}{2}} & 0 & 0 \\ 0 & \sqrt{\frac{82 - 6\sqrt{149}}{2}} & 0 \\ 0 & 0 & 0 \end{bmatrix} \]

\[ S^T = \begin{bmatrix} \frac{71 + 3\sqrt{149}}{50} & \frac{21 + 3\sqrt{149}}{50} & \frac{-4 + 3\sqrt{149}}{25} \frac{1}{\sqrt{149}} \\ \frac{71 - 3\sqrt{149}}{50} & \frac{21 - 3\sqrt{149}}{50} & \frac{-4 - 3\sqrt{149}}{25} \frac{1}{\sqrt{149}} \\ 1 & 1 & 0 \quad -2 & 3 & 0 \quad 1 \end{bmatrix} \]

\[ \|s_1\| = \frac{\sqrt{16,092 + 456\sqrt{149}}}{50}, \quad \|s_2\| = \frac{\sqrt{16,092 - 456\sqrt{149}}}{50}, \quad \|s_3\| = \frac{\sqrt{3}}{\sqrt{149}}, \quad \|s_4\| = \frac{\sqrt{3}}{\sqrt{149}} \]

\[ S^T \text{ contains the transposed eigenvectors of } A^T A. \]

\[ V^T = \begin{bmatrix} \frac{71 + 3\sqrt{149}}{\sqrt{16,092 + 456\sqrt{149}}} & \frac{21 + 3\sqrt{149}}{\sqrt{16,092 + 456\sqrt{149}}} & \frac{-4 + 3\sqrt{149}}{\sqrt{16,092 + 456\sqrt{149}}} \\ \frac{71 - 3\sqrt{149}}{\sqrt{16,092 - 456\sqrt{149}}} & \frac{21 - 3\sqrt{149}}{\sqrt{16,092 - 456\sqrt{149}}} & \frac{-4 - 3\sqrt{149}}{\sqrt{16,092 - 456\sqrt{149}}} \\ \frac{1}{\sqrt{149}} & \frac{1}{\sqrt{149}} & 0 \end{bmatrix} \]

\[ u_1 = \frac{A_{v1}}{\sigma_1} = \frac{2}{\sqrt{82 + 6\sqrt{149}}} \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 4 & 2 & 2 \end{bmatrix} \begin{bmatrix} 71 + 3\sqrt{149} \\ 50 \\ \frac{1}{\sqrt{16,092 + 456\sqrt{149}}} \end{bmatrix} = \begin{bmatrix} -152 + 36\sqrt{149} \\ \frac{1,727,208 + 133,944\sqrt{149}}{410 + 30\sqrt{149}} \\ \frac{1,727,208 - 133,944\sqrt{149}}{820 + 60\sqrt{149}} \end{bmatrix} \]

\[ u_2 = \frac{A_{v2}}{\sigma_2} = \frac{2}{\sqrt{82 - 6\sqrt{149}}} \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 4 & 2 & 2 \end{bmatrix} \begin{bmatrix} 71 - 3\sqrt{149} \\ 50 \\ \frac{1}{\sqrt{16,092 - 456\sqrt{149}}} \end{bmatrix} = \begin{bmatrix} -152 - 36\sqrt{149} \\ \frac{1,727,208 - 133,944\sqrt{149}}{410 - 30\sqrt{149}} \\ \frac{1,727,208 + 133,944\sqrt{149}}{820 - 60\sqrt{149}} \end{bmatrix} \]

Since \( A \) is 3 \times 4, \( U \) should be a 3 \times 3 matrix. However, there is no \( \sigma_3 \), so we cannot use \( u_i = \frac{A_{vi}}{\sigma_i} \) to find \( u_3 \). Instead, we use the left null space of \( A \).

\[ \text{Null}(A^T) = \begin{bmatrix} 1 & 2 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 1 & 1 & 2 & 0 \\ 2 & 1 & 2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & -3 & -6 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \]

\[ y = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} \]

Normalizing \( y \) to obtain \( u_3 \):

\[ u_3 = \frac{1}{\|y\|} y = \frac{1}{\sqrt{5}} \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \]
So, 

$$U = \begin{bmatrix}
\sqrt{172.7208 + 133.944\sqrt{149}} & \sqrt{172.7208 - 133.944\sqrt{149}} & 0 \\
\sqrt{172.7208 + 133.944\sqrt{149}} & \sqrt{172.7208 - 133.944\sqrt{149}} & -2 \sqrt{5} \\
\sqrt{172.7208 + 133.944\sqrt{149}} & \sqrt{172.7208 - 133.944\sqrt{149}} & \sqrt{172.7208 - 133.944\sqrt{149}} \\
\sqrt{172.7208 + 133.944\sqrt{149}} & \sqrt{172.7208 - 133.944\sqrt{149}} & 1 \sqrt{5}
\end{bmatrix}.$$ 

Thus, 

$$A = \begin{bmatrix}
\sqrt{172.7208 + 133.944\sqrt{149}} & \sqrt{172.7208 - 133.944\sqrt{149}} & 0 \\
\sqrt{172.7208 + 133.944\sqrt{149}} & \sqrt{172.7208 - 133.944\sqrt{149}} & -2 \sqrt{5} \\
\sqrt{172.7208 + 133.944\sqrt{149}} & \sqrt{172.7208 - 133.944\sqrt{149}} & \sqrt{172.7208 - 133.944\sqrt{149}}
\end{bmatrix} \begin{bmatrix}
\sqrt{82+6\sqrt{149}} & 0 & 0 \\
0 & \sqrt{82-6\sqrt{149}} & 0 \\
0 & 0 & 0
\end{bmatrix} = U\Sigma V^T.$$ 

Example 2. Let

$$A = \begin{bmatrix}
1 & 1 \\
1 & 0
\end{bmatrix}.$$ 

Notice that when $A$ is a symmetric matrix, $A^T A = AA^T$, so $U = V$. Less work is required.

$$A^T A = AA^T = \begin{bmatrix}
1 & 1 \\
1 & 0
\end{bmatrix} \begin{bmatrix}
1 & 1 \\
1 & 0
\end{bmatrix} = \begin{bmatrix}
2 & 1 \\
1 & 1
\end{bmatrix}$$

$$rank(A) = rank(A^T A) = rank(AA^T) = 2$$

$$\Delta_{A^T A}(\lambda) = \Delta_{AA^T}(\lambda) = |A^T A - \lambda I| = |AA^T - \lambda I| = \begin{vmatrix}
2 - \lambda & 1 \\
1 & 1 - \lambda
\end{vmatrix} = (2 - \lambda)(1 - \lambda) - 1^2 = \lambda^2 - 3\lambda + 1 \implies \lambda_1 = \frac{3 + \sqrt{5}}{2}, \lambda_2 = \frac{3 - \sqrt{5}}{2}$$

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{\frac{3 + \sqrt{5}}{2}} = \sqrt{\frac{6 + 2\sqrt{5}}{2}} = \frac{1 + \sqrt{5}}{2}$$

$$\sigma_2 = \sqrt{\lambda_2} = \sqrt{\frac{3 - \sqrt{5}}{2}} = \sqrt{\frac{6 - 2\sqrt{5}}{2}} = \frac{1 - \sqrt{5}}{2} \implies \Sigma = \begin{bmatrix}
\frac{1 + \sqrt{5}}{2} & 0 \\
0 & \frac{1 - \sqrt{5}}{2}
\end{bmatrix}$$

$$v_1 \& u_1 :$$

$$[A^T A - \lambda_1 I]x_1 = [AA^T - \lambda_1 I]s_1 = \begin{bmatrix}
2 & \frac{3 + \sqrt{5}}{2} \\
1 & 1 - \frac{3 + \sqrt{5}}{2}
\end{bmatrix} \begin{bmatrix}
0 \\
0
\end{bmatrix} = \begin{bmatrix}
\frac{1 - \sqrt{5}}{2} \\
\frac{1 + \sqrt{5}}{2}
\end{bmatrix} \implies s_1 = \begin{bmatrix}
\frac{1 + \sqrt{5}}{2} \\
\frac{1 - \sqrt{5}}{2}
\end{bmatrix}$$

$$||s_1|| = \sqrt{\frac{6 + 2\sqrt{5}}{2} + 1} = \sqrt{\frac{10 + 2\sqrt{5}}{2}} = \sqrt{10 + 2\sqrt{5}}$$

$$v_1 = u_1 = \frac{1}{||s_1||} s_1 = \frac{1}{\sqrt{10 + 2\sqrt{5}}} \begin{bmatrix}
\frac{1 + \sqrt{5}}{2} \\
\frac{1 - \sqrt{5}}{2}
\end{bmatrix} = \frac{2}{\sqrt{10 + 2\sqrt{5}}} \begin{bmatrix}
\frac{1 + \sqrt{5}}{2} \\
\frac{1 - \sqrt{5}}{2}
\end{bmatrix} = \begin{bmatrix}
\frac{1 + \sqrt{5}}{\sqrt{10 + 2\sqrt{5}}} \\
\frac{1 - \sqrt{5}}{\sqrt{10 + 2\sqrt{5}}}
\end{bmatrix}$$

$$v_2 \& u_2 :$$
\[ [A^T - \lambda_2 I]s_2 = [AA^T - \lambda_2 I]s_2 = \begin{bmatrix} 2 - \frac{3}{2}\sqrt{5} & 1 & \frac{1}{2} & 0 \\ 1 & 1 - \frac{3}{2}\sqrt{5} & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 + \sqrt{5} & 1 & -\frac{1}{2} & 0 \\ 1 & 1 - \frac{1}{2} \sqrt{5} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \]

\[ ||s_2|| = \sqrt{6 - 2\sqrt{5}} + 1 = \sqrt{\frac{10 - 2\sqrt{5}}{4}} = \sqrt{\frac{10 - 2\sqrt{5}}{2}} \]

\[ v_2 = u_2 = \frac{1}{||s_2||} s_2 = \frac{1}{\sqrt{\frac{10 - 2\sqrt{5}}{2}}} \begin{bmatrix} 1 - \sqrt{5} \\ 2 \\ \sqrt{10 - 2\sqrt{5}} \\ \sqrt{10 - 2\sqrt{5}} \end{bmatrix} = \frac{2}{\sqrt{10 - 2\sqrt{5}}} \begin{bmatrix} 1 - \sqrt{5} \\ 2 \\ \sqrt{10 - 2\sqrt{5}} \\ \sqrt{10 - 2\sqrt{5}} \end{bmatrix} \]

\[ V^T = \begin{bmatrix} 1 + \sqrt{5} \\ \sqrt{10 + 2\sqrt{5}} \\ \sqrt{10 - 2\sqrt{5}} \\ \sqrt{10 - 2\sqrt{5}} \end{bmatrix}, U = \begin{bmatrix} 1 + \sqrt{5} \\ \sqrt{10 + 2\sqrt{5}} \\ \sqrt{10 - 2\sqrt{5}} \\ \sqrt{10 - 2\sqrt{5}} \end{bmatrix} \]

To verify that \( v_1 = u_1 \) and \( v_2 = u_2 \):

\[ Av_1 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1 + \sqrt{5}}{10 + 2\sqrt{5}} \\ \frac{1 + \sqrt{5}}{10 - 2\sqrt{5}} \end{bmatrix} = \begin{bmatrix} \frac{1 + \sqrt{5}}{10 + 2\sqrt{5}} \\ \frac{1 + \sqrt{5}}{10 - 2\sqrt{5}} \end{bmatrix} = \frac{1 + \sqrt{5}}{2} \begin{bmatrix} \frac{1 + \sqrt{5}}{10 + 2\sqrt{5}} \\ \frac{1 + \sqrt{5}}{10 - 2\sqrt{5}} \end{bmatrix} = \sigma_1 u_1 = \sigma_1 v_1 \]

\[ Av_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1 - \sqrt{5}}{10 + 2\sqrt{5}} \\ \frac{1 - \sqrt{5}}{10 - 2\sqrt{5}} \end{bmatrix} = \begin{bmatrix} \frac{1 - \sqrt{5}}{10 + 2\sqrt{5}} \\ \frac{1 - \sqrt{5}}{10 - 2\sqrt{5}} \end{bmatrix} = \frac{1 - \sqrt{5}}{2} \begin{bmatrix} \frac{1 - \sqrt{5}}{10 + 2\sqrt{5}} \\ \frac{1 - \sqrt{5}}{10 - 2\sqrt{5}} \end{bmatrix} = \sigma_2 u_2 = \sigma_2 v_2 \]

Notice that this implies the eigenvalues of \( A \) are equal to the singular values of \( A \). By Lemma 7, every symmetric matrix has this property. Thus,

\[ A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1 + \sqrt{5}}{10 + 2\sqrt{5}} \\ \frac{1 - \sqrt{5}}{10 - 2\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1 + \sqrt{5}}{2} \\ \frac{1 - \sqrt{5}}{2} \end{bmatrix} = U \Sigma V^T = U \Lambda V^T. \]

5 Maple

The purpose in subjecting a color photograph to the Singular Value Decomposition is to greatly reduce the amount of data required to transmit the photograph to or from a satellite, for instance. A digital image is essentially a matrix comprised of three other matrices of identical size. These are the red, green, and blue layers that combine to produce the colors in the original image. Obtain the three layers of the image using the Maple command GetLayers from the ImageTools package. It is on each of these three layers that we perform the SVD. Define each one as \( \text{img}_r, \text{img}_g, \text{and} \ img_b \). Define the singular values of each matrix using the SingularValues command in the Linear Algebra package. Maple will also calculate \( U \) and \( V^T \). Simply set the output of SingularValues = ['U', 'Vt'].

In the argument of the following procedure, the variable \( n \) denotes which approximation the procedure will compute, that is, the number of singular values that it will include. \text{posint} indicates that \( n \) must be a positive integer.

\[ \text{approx} := \text{proc}(	ext{img}_r, \text{img}_g, \text{img}_b, \text{n}::\text{posint}) \text{local} \]
\[ \text{Singr}, \text{Singg}, \text{Singb}, \text{Ur}, \text{Ug}, \text{Ub}, \text{Vtr}, \text{Vtg}, \text{Vtb}, \text{singr}, \]

12
In place of a $\Sigma$ matrix, we create a list of the $\sigma$'s for each red, green and blue layer, as well as the red, green, and blue $U$ and $V^T$ matrices. It is important to note that Maple outputs the transpose of $V$. Hence, it does not need to be transposed.

```maple
Singr:=SingularValues(img_r,output='list');
Singg:=SingularValues(img_g,output='list');
Singb:=SingularValues(img_b,output='list');
Ur,Vtr:=SingularValues(img_r,output=['U','Vt']);
Ug,Vtg:=SingularValues(img_g,output=['U','Vt']);
Ub,Vtb:=SingularValues(img_b,output=['U','Vt']);

Pulling out each individual $\sigma_i$, $u_i$, and $v_i^T$ to create the $\sigma_i u_i v_i^T$ dyads for $i = 1 \ldots r$:

```maple
for i from 1 to n do
    singr[i]:=Singr[i];
singg[i]:=Singg[i];
singb[i]:=Singb[i];
    ur[i]:=LinearAlgebra:-Column(Ur,i..i);
    vr[i]:=LinearAlgebra:-Column(LinearAlgebra:-Transpose(Vtr),i..i);
    ug[i]:=LinearAlgebra:-Column(Ug,i..i);
    vg[i]:=LinearAlgebra:-Column(LinearAlgebra:-Transpose(Vtg),i..i);
    ub[i]:=LinearAlgebra:-Column(Ub,i..i);
    vb[i]:=LinearAlgebra:-Column(LinearAlgebra:-Transpose(Vtb),i..i);
end do;
```

Note that Maple stores data as floating point numbers, so values that would be 0 otherwise are stored as a small decimal number very close to 0, yet still greater than 0. This means that, when working with Maple, $\text{rank}(A) = r = \min(m,n)$.

Adding the dyads to produce the approximations of each layer:

```maple
Mr:=add(singr[i]*ur[i].LinearAlgebra:-Transpose(vr[i]),i=1..n);
Mg:=add(singg[i]*ug[i].LinearAlgebra:-Transpose(vg[i]),i=1..n);
Mb:=add(singb[i]*ub[i].LinearAlgebra:-Transpose(vb[i]),i=1..n);
```

Combining the approximations of each layer:

```maple
img_rgb:=CombineLayers(Mr,Mg,Mb);
```

Displaying the result in Maple:

```maple
Embed(img_rgb);
end proc:
```
Application  Suppose we have a color photograph that is 100 × 200 pixels, or entries. That’s 20,000 pixels. When we separate it into its red, green, and blue layers, the number of entries becomes 100×200×3, or 60,000 entries. Apply SVD to each of these matrices and take enough of the dyad decomposition to obtain a meaningful approximation. For the sake of example, suppose two products from the outer product decomposition suffice. In other words, we have \( \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T \).

\[
\begin{bmatrix}
\sigma_1 \\
\sigma_2
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2
\end{bmatrix}
\begin{bmatrix}
v_1^T \\
v_2^T
\end{bmatrix}
\]

The \( \sigma \)'s are just scalars, the \( u_i \) column vectors are 100 × 1, and the row vectors \( v_i^T \) are 1 × 200. So, it follows that \( \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T \) has 301 + 301 = 602 entries. Multiply this by three to account for the red, green, and blue layers to obtain 1,806 entries, approximately 3% of the original 60,000 entries that we would have had to send had we not utilized the SVD. Now, when the satellite sends the photograph down to Earth, it sends those \( \sigma_1 \) and \( \sigma_2 \), \( u_1 \) and \( u_2 \), and \( v_1^T \) and \( v_2^T \) separately. All that needs to be done to recover the approximation is to multiply these together and add them up back on Earth.

Example 3. Here, we show a progression of approximations of a photograph of a galaxy. The picture has dimensions 780 × 960, and each approximation uses \( n \) singular values.
(a) original image

(b) $n = 1$. This is merely $\sigma_1 u_1 v_1^T$, the first term in the sum of 780 dyads $\sigma_i u_i v_i^T$. Thus, it bares very little resemblance to the original image.

(c) $n = 10$: $\sigma_1 u_1 v_1^T + \cdots + \sigma_{10} u_{10} v_{10}^T$
With only $\frac{10}{780}$ dyads, or 1.28\% of all the data, we can already see the shapes in the original image starting to come together.

(d) $n = 100$: $\sigma_1 u_1 v_1^T + \cdots + \sigma_{100} u_{100} v_{100}^T$
Notice that with only $\frac{100}{780}$, or 13\%, of the information contained in the original image we have an approximation that is close to being indistinguishable from the original.

(e) $n = 500$: $\sigma_1 u_1 v_1^T + \cdots + \sigma_{500} u_{500} v_{500}^T$
This is 64.1\% of all the information in the original image. The lack of a noticeable difference between this picture and the previous one illustrates the fact that the $\sigma_i$ values are getting much smaller as $i$ increases.

(f) $n = 780$: $\sigma_1 u_1 v_1^T + \cdots + \sigma_{780} u_{780} v_{780}^T$
When we use all singular values, the approximation is the same as the original image.
6 Conclusions

In summary, the application of singular value decomposition we have detailed provides a method of calculating very accurate approximations of photographs so that they may be transmitted from satellites to Earth without requiring large amounts of data. SVD provides bases for the Four Fundamental Subspaces of a matrix, and it gets its versatility from the ordering of the $\sigma$ values. SVD is also used in the calculation of pseudoinverses, as illustrated in [3], among other things.

A Appendix

In this appendix, we prove a few results related to symmetric matrices.

Lemma 1. Let $A = A^T$. Then eigenvectors corresponding to different eigenvalues are orthogonal.

Proof. Let $Av_1 = \lambda_1 v_1$ and $Av_2 = \lambda_2 v_2$. We show that $v_1 \cdot v_2 = v_1^T v_2 = 0$. Observe that

$$
\lambda_1 (v_1^T v_2) = (A v_1)^T v_2 = (Av_1) v_2 = v_1^T (Av_2) = v_1^T (\lambda_2 v_2) = \lambda_2 (v_1^T v_2).
$$

Thus, $\lambda_1 (v_1^T v_2) = \lambda_2 (v_1^T v_2) \implies (\lambda_1 - \lambda_2) (v_1^T v_2) = 0$, so $v_1^T v_2 = 0$ since $\lambda_1 \neq \lambda_2$.

Lemma 2. Let $A = A^T$ and $A$ be real. Then, eigenvalues of $A$ are non-negative.

Proof. Suppose $A = A^T$ and $A = \overline{A}$. Let $Av = \lambda v$, where $v \neq 0$. We show that $\lambda = \overline{\lambda}$. Thus,

$$
A \overline{v} = \overline{\lambda} \overline{v} \quad \text{and} \quad (A \overline{v})^T v = \overline{v}^T (A^T v) = \overline{v}^T (Av) = \overline{v}^T (\lambda v) = \overline{\lambda} (v^T v),
$$

and,

$$
(\overline{\lambda} v^T v) = \overline{\lambda} (\overline{v}^T v) = \lambda (v^T v). \quad (14)
$$

Let $v = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$, where $z_i \in \mathbb{C}$. So,

$$
\overline{v}^T v = \overline{z_1, \ldots, z_n} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = z_1 \overline{z}_1 + \cdots + z_n \overline{z}_n = |z_1|^2 + \cdots + |z_n|^2 > 0.
$$

From (14) we have $\overline{\lambda} (v^T v) - \lambda (v^T v) = 0$ since $v \neq 0$. Hence, $(\overline{\lambda} - \lambda) (v^T v) = 0$, so $\overline{\lambda} = \lambda$ since $v^T v > 0$.

Lemma 3. Let $A$ be an $m \times n$ matrix.

1. $\text{Null}(A) = \text{Null}(A^T A)$

Proof of $(\subseteq)$. Let $x \in \text{Null}(A)$. Then $Ax = 0$ so $A^T (Ax) = (A^T A) x = 0$.

$\therefore x \in \text{Null}(A^T A)$. 

16
Proof of \((\subseteq)\). Let \(x \in \text{Null}(A^T A)\). Then \((A^T A)x = 0\). Thus \(A^T (Ax) = 0\), so \(Ax \in \text{Null}(A^T)\). On the other hand, \(Ax \in \text{Col}(A)\). Since \(\mathbb{R}^m = \text{Col}(A) \oplus \text{Null}(A^T)\), we conclude that \(Ax = 0\) since \(\text{Col}(A) \cap \text{Null}(A^T) = \{0\}\). Thus, \(\text{Null}(A^T A) \subseteq \text{Null}(A)\).

\[\therefore \text{Null}(A) = \text{Null}(A^T A)\]

2. \(\text{rank}(A) = \text{rank}(A^T A)\)

Proof. Let \(r = \text{rank}(A)\). Thus,

\[r = n - \dim(\text{Null}(A)) = n - \dim(\text{Null}(A^T A)) = \text{rank}(A^T A)\ 	ext{ since } A^T A \text{ is } n \times n.\]

\[\therefore \text{rank}(A) = \text{rank}(A^T A)\]

3. \(\dim(\text{Null}(A)) = n - r = \dim(\text{Null}(A^T A))\)
\(\dim(\text{Col}(A)) = r = \dim(\text{Row}(A))\)
\(\dim(\text{Col}(A^T)) = \text{rank}(A^T A) = \text{rank}(A) = \dim(\text{Col}(A)) = r\)

4. \(\text{Null}(A^T) = \text{Null}(AA^T)\)

Proof of \((\subseteq)\). Let \(x \in \text{Null}(A^T)\), then \(A^T x = 0\) so \(A(A^T x) = (AA^T)x = 0\).

\[\therefore x \in \text{Null}(AA^T)\]

Proof of \((\supseteq)\). Let \(x \in \text{Null}(AA^T)\). Then \((AA^T)x = 0\). Thus, \(A(A^T x) = 0\) so \(A^T x \in \text{Null}(A)\). On the other hand, \(A^T x \in \text{Col}(A^T)\). Since

\[\mathbb{R}^m = \text{Col}(A^T) \oplus \text{Null}(A),\]

we conclude that \(A^T x = 0\) since \(\text{Col}(A^T) \cap \text{Null}(A) = \{0\}\). Thus, \(\text{Null}(AA^T) \subseteq \text{Null}(A^T)\).

\[\therefore \text{Null}(A^T) = \text{Null}(AA^T)\]

5. \(\text{rank}(A^T) = \text{rank}(AA^T) = \text{rank}(A), \text{ so } \text{rank}(A) = \text{rank}(A^T A) = \text{rank}(A^T) = \text{rank}(AA^T)\).  

6. We have the following:

(i) \(\dim(\text{Null}(A^T)) = m - r = \dim(\text{Null}(AA^T))\),
(ii) \(\dim(\text{Col}(A^T)) = \text{rank}(A^T) = \text{rank}(A) = r = \dim(\text{Row}(A^T))\),
(iii) \(\dim(\text{Col}(AA^T)) = \text{rank}(AA^T) = \text{rank}(A) = \dim(\text{Col}(A)) = r\).

7. Since \(A^T A v_i = \sigma_i^2 v_i, i = 1 \ldots r\), and vectors \(\{v_1, \ldots, v_r\}\) provide a basis for \(\text{Row}(A)\), we have

\[Av_i \neq 0, \quad i = 1 \ldots r.\]

If \(A^T A v_i = 0\), then \(v_i\) would belong to \(\text{Null}(A^T A) \supseteq \text{Null}(A)\), which would give \(Av_i = 0\). \(\therefore\) contradiction. So, \(A^T A v_i = \sigma_i^2 v_i \neq 0\), which implies \(\sigma_i^2 \neq 0\) so \(\sigma_i \neq 0, i = 1 \ldots r\).

Lemma 4. \(\sigma_i \neq 0 \text{ for } i = 1 \ldots r\).
Lemma 5. Since the vectors \( \{v_1, \ldots, v_r\} \) are orthonormal, vectors \( \{u_1, \ldots, u_r\} \) computed as

\[
u_i = \frac{Av_i}{\sigma_i}, \quad i = 1 \ldots r,
\]

are also orthonormal.

Proof. We have the following:

\[
(Av_i)^T(Av_j) = v_i^T(A^TAv_j) = v_i^T \sigma_j^2 v_j = \sigma_j^2 v_i^T v_j = \sigma_j^2 \delta_{ij} = \begin{cases} 0, & i \neq j; \\ \sigma_j^2, & i = j. \end{cases}
\]

Thus, vectors \( \{u_i\}_{i=1}^r \) are orthogonal and orthonormal because:

\[
u_i^T u_j = \left(\frac{Av_i}{\sigma_i}\right)^T \left(\frac{Av_j}{\sigma_j}\right) = \frac{1}{\sigma_i \sigma_j} (Av_i)^T(Av_j) = \left(\frac{1}{\sigma_i \sigma_j}\right) \sigma_j^2 \delta_{ij} = \begin{cases} 0, & i \neq j; \\ 1, & i = j. \end{cases}
\]

Lemma 6. Let \( v \) be a \( n \times 1 \) vector. Then, the dyad \( v^T v \) has rank 1.

Proof. We compute the following:

\[
v^T v = \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix}_{n \times 1} \begin{bmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{bmatrix}_{1 \times n} = \begin{bmatrix} v_{11} v_{11}^T & v_{11} v_{12}^T & \cdots & v_{11} v_{1n}^T \\ v_{21} v_{11}^T & v_{21} v_{12}^T & \cdots & v_{21} v_{1n}^T \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} v_{11}^T & v_{n1} v_{12}^T & \cdots & v_{n1} v_{1n}^T \end{bmatrix}_{n \times n} = \begin{bmatrix} - v_{11} v_{11}^T & - v_{11} v_{12}^T & \cdots & - v_{11} v_{1n}^T \\ - v_{21} v_{11}^T & - v_{21} v_{12}^T & \cdots & - v_{21} v_{1n}^T \\ \vdots & \vdots & \ddots & \vdots \\ - v_{n1} v_{11}^T & - v_{n1} v_{12}^T & \cdots & - v_{n1} v_{1n}^T \end{bmatrix}.
\]

Since the rows of \( vv^T \) are multiples of \( v^T \), they are linearly dependent. Therefore, we have \( \text{rank}(vv^T) = 1 \).

Lemma 7. Let \( A \) be a symmetric matrix. Then, the eigenvalues of \( A \) are equal to the singular values of \( A \).

Proof. Let \( A \) be a matrix such that \( A = A^T \), let \( \lambda \geq 0 \) be an eigenvalue of \( A \), and let \( v \) be an eigenvector of \( A \). Then, \( Av = \lambda v \), and \( A^T Av = A^T \lambda v = \lambda A^T v = \lambda Av = \lambda^2 v \). Hence, \( \lambda^2 \) is an eigenvalue of \( A^T A \). Therefore, \( \lambda \) is a singular value of \( A \).

On the other hand, let \( \sigma > 0 \) be a singular value of \( A \). So, \( A^T Av = \sigma^2 v \) for some nonzero eigenvector \( v \). \( \Delta_{A^T A}(\sigma^2) = \det(A^T A - \sigma^2 I) = \det(A^2 - \sigma^2 I) = \det((A - \sigma I)(A + \sigma I)) = \det(A - \sigma I) \det(A + \sigma I) = 0 \), which implies that \( \det(A - \sigma I) = 0, \det(A + \sigma I) = 0 \), or \( \det(A - \sigma I) = \det(A + \sigma I) = 0 \). Suppose, \( \det(A + \sigma I) = 0 \). Then, \( -\sigma \) is an eigenvalue of \( A \), a contradiction since \( \sigma > 0 \). Therefore, \( \det(A + \sigma I) = 0 \), and \( \sigma \) is an eigenvalue of \( A \).

Lemma 8. Let \( A \) be an \( m \times n \) matrix. Then, \( A^T A \) and \( AA^T \) share the same nonzero eigenvalues and, therefore, both provide the singular values of \( A \). In the case where \( m = n \), \( \Delta_{A^T A}(t) = \Delta_{AA^T}(t) \).
Proof. From parts 2 and 5 of Theorem (1), we have $A^T A = V \Sigma^T \Sigma V^T = V \Lambda_n V^{-1}$, and $A A^T = U \Sigma \Sigma^T U^T = U \Lambda_m U^{-1}$, where $\Lambda_n$ is $n \times n$, and $\Lambda_m$ is $m \times m$. Hence, $A^T A$ is similar to $\Lambda_n$, and $A A^T$ is similar to $\Lambda_m$.

We cannot show that $\Lambda_n$ is similar to $\Lambda_m$. However, from Definition (1), we know what $\Sigma$ and $\Sigma^T$ look like, so we know that when we multiply them together to obtain $\Lambda_n$ and $\Lambda_m$, we get two square matrices with $\sigma_1^2, \ldots, \sigma_r^2$, or $\lambda_1, \ldots, \lambda_r$, and zeros along the diagonal, and zeros elsewhere. The only difference between the two is that one is larger, depending on whether $m < n$ or $m > n$, and has more zeros on its diagonal.

$$\Lambda_n = \begin{bmatrix}
\sigma_1^2 & 0 & \cdots & 0 \\
0 & \ddots & & \vdots \\
\vdots & & \sigma_r^2 & \\
0 & \cdots & 0 & 0
\end{bmatrix}_{n \times n} = \begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & & \vdots \\
\vdots & & \lambda_r & \\
0 & \cdots & 0 & 0
\end{bmatrix}
$$

$$\Lambda_m = \begin{bmatrix}
\sigma_1^2 & 0 & \cdots & 0 \\
0 & \ddots & & \vdots \\
\vdots & & \sigma_r^2 & \\
0 & \cdots & 0 & 0
\end{bmatrix}_{m \times m} = \begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & & \vdots \\
\vdots & & \lambda_r & \\
0 & \cdots & 0 & 0
\end{bmatrix}
$$

Since the eigenvalues of a diagonal matrix are simply the entries on its diagonal, we know that $\Lambda_n$ and $\Lambda_m$ both have $\lambda_1, \ldots, \lambda_r$ as eigenvalues. Because they are similar to $A^T A$ and $A A^T$, respectively, we know that $A^T A$ and $A A^T$ both must also have $\lambda_1, \ldots, \lambda_r$ as eigenvalues.

Moreover, an $n \times n$ matrix has $n$ entries on its diagonal and hence has a characteristic polynomial of degree $n$. Thus,

$$\Delta_{A^T A}(t) = \det(\Lambda_n - tI) = (\lambda_1 - t)(\lambda_2 - t) \cdots (\lambda_r - t)(-t) \cdots (-t)$$

$$= (-t)^{n-r}(\lambda_1 - t) (\lambda_2 - t) \cdots (\lambda_r - t)$$

$$= (-t)^{n-r} h(t),$$

and

$$\Delta_{A A^T}(t) = \det(\Lambda_m - tI) = (\lambda_1 - t)(\lambda_2 - t) \cdots (\lambda_r - t)(-t) \cdots (-t)$$

$$= (-t)^{m-r}(\lambda_1 - t) (\lambda_2 - t) \cdots (\lambda_r - t)$$

$$= (-t)^{m-r} h(t).$$
where \( h(t) \) is a monic polynomial of degree \( r \) with only \( \lambda_1, \ldots, \lambda_r \) as roots.

Therefore, when \( A \) is a square \( n \times n \) matrix, we have the special case where \( \Delta_{A^T A}(t) = \Delta_{AA^T}(t) = (-t)^{n-r}h(t) \), and \( \Lambda_m = \Lambda_n \).

\[
\]

References


