DEPARTMENT OF MATHEMATICS

TECHNICAL REPORT

SINGULAR VALUE DECOMPOSITION

Andrew Lounsbury

September 2018

No. 2018-2

TENNESSEE TECHNOLOGICAL UNIVERSITY
Cookeville, TN 38505

Singular Value Decomposition*

Andrew Lounsbury
Department of Mathematics
Tennessee Technological University
Cookeville, TN, 38505
awlounsbur42@students.tntech.edu

September 28, 2018

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

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*

$$A = U\Sigma V^T, \tag{1}$$

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

$$A = U\Sigma V^T = \underbrace{\left[\begin{array}{c|c|c|c} u_1 & u_2 & \cdots & u_m \end{array}\right]}_{U(m\times m)} \underbrace{\begin{bmatrix} \sigma_1 & 0 & \dots & & \dots & 0 & \dots & 0 \\ 0 & \ddots & & & & \vdots & & \vdots \\ \vdots & & \sigma_r & & & & \ddots & \\ 0 & & & 0 & & \vdots & & \vdots \\ 0 & 0 & \dots & & \ddots & 0 & \dots & 0 \end{bmatrix}}_{\Sigma(m\times n)} \underbrace{\begin{bmatrix} v_1^T \\ \hline v_2^T \\ \hline \vdots \\ \hline v_n^T \end{bmatrix}}_{V^T(n\times n)}, \tag{2}$$

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 $A^T A$.

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

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

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

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

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

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

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

If $\operatorname{rank}(A) = r < \min(m, n)$, then there are $n - r$ zero columns and rows in Σ, rendering the

$r+1,\ldots,m$ columns of U and $r+1,\ldots,n$ rows in V^T unnecessary to recover A. Therefore,

$$
\begin{aligned}
A = U\Sigma V^T &= \left[\begin{array}{c|c|c|c|c|c} u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \end{array}\right]
\begin{bmatrix}
\sigma_1 & 0 & \ldots & & \ldots & 0 & \ldots & 0 \\
0 & \ddots & & & & \vdots & & \vdots \\
\vdots & & \sigma_r & & & & \ddots & \\
0 & & & 0 & & \vdots & & \vdots \\
0 & 0 & \ldots & & \ddots & 0 & \ldots & 0
\end{bmatrix}
\begin{bmatrix}
v_1^T \\ \hline \vdots \\ \hline v_r^T \\ \hline v_{r+1}^T \\ \hline \vdots \\ \hline v_n^T
\end{bmatrix} \\
&= \underbrace{\left[\begin{array}{c|c|c} u_1 & \cdots & u_r \end{array}\right]}_{m\times r}
\underbrace{\begin{bmatrix}
\sigma_1 & 0 & \ldots & 0 \\
0 & \sigma_2 & & \vdots \\
\vdots & & \ddots & \\
0 & \ldots & & \sigma_r
\end{bmatrix}}_{r\times r}
\underbrace{\begin{bmatrix}
v_1^T \\ \hline \vdots \\ \hline v_r^T
\end{bmatrix}}_{r\times n} \\
&= u_1\sigma_1 v_1 + \cdots + u_r\sigma_r v_r. \qquad (6)
\end{aligned}
$$

2 Steps for Calculation of SVD

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

1. Compute A^TA of a real $m \times n$ matrix A of rank r.

2. Compute the singular values of A^TA.

   Solve the characteristic equation $\Delta_{A^TA}(\lambda) = |A^TA - \lambda I| = 0$ of A^TA for the eigenvalues $\lambda_1,\ldots,\lambda_r$ of A^TA. 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,

   $$\sigma_i = +\sqrt{\lambda_i}, \quad i = 1,\ldots,r. \qquad (7)$$

3. Sort the singular values, possibly renaming them, so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$.

4. Construct the Σ 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^TA.

   Find a basis for $\text{Null}(A^TA - \lambda_i I)$. That is, solve $(A^TA - \lambda_i I)s_i = 0$ for s_i, an eigenvector of A corresponding to λ_i, for each eigenvalue λ_i. Since A^TA 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

   $$v_i = \frac{1}{||s_i||}s_i, \quad i = 1,\ldots,r. \qquad (8)$$

If $n > r$, the additional $n-r$ vectors $v_{r+1}, \dots, v_n$ need to be chosen as an orthonormal basis in Null(A). Note that since $Av_i = \sigma_i u_i$ for $i = 1, \dots,$ vectors $v_1, \dots, v_r$ provide an orthonormal basis for Row(A) while the vectors $u_1, \dots, u_r$ provide an orthonormal basis for Col(A). In particular,

$$\mathbb{R}^n = \text{Row}(A) \perp \text{Null}(A) = \text{span}\{v_1, \dots, v_r\} \perp \text{span}\{v_{r+1}, \dots, v_{r+(n-r)}\}. \tag{9}$$

7. Construct the orthogonal matrix $V = [v_1|\cdots|v_n]$.

8. Verify $V^T V = I$.

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

$$Av_i = \sigma_i u_i \implies u_i = \frac{Av_i}{\sigma_i}, \; i = 1 \dots r. \tag{10}$$

In this method, $u_1, \dots, u_r$ are orthogonal by Lemma 5.

Alternatively,

(i) Note that $AA^T = U(\Sigma\Sigma^T)U^T$ 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 $\Delta_{A^T A}(\lambda) = \Delta_{AA^T}(\lambda)$ by Lemma 8, $\sigma_1, \dots, \sigma_r$ are also the square roots of the eigenvalues of AA^T.

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

$$\mathbb{R}^m = \text{Col}(A) \perp \text{Null}(A^T) = \text{span}\{u_1, \dots, u_r\} \perp \text{span}\{u_{r+1}, \dots, u_{r+(m-r)}\}. \tag{11}$$

10. Construct $U = [u_1|\cdots|u_m]$.

11. Verify $U^T U = I$.

12. Verify $A = U\Sigma V^T$.

13. Construct the dyadic decomposition[1] of A, as described in Thm. 13:

$$A = U\Sigma V^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + u_r \sigma_r v_r^T. \tag{12}$$

[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*

$$\begin{cases} Av_i = \sigma_i u_i, & i = 1, \dots, r \\ Av_i = 0, & i = r+1, \dots, r+(n-r) \end{cases} \implies \begin{cases} \text{Row}(A) &= \text{span}\{v_1, \dots, v_r\} \\ \text{Null}(A) &= \text{span}\{v_{r+1}, \dots, v_{r+(n-r)}\} \end{cases}$$

2. $A^TA = V(\Sigma^T\Sigma)V^T : \mathbb{R}^n \to \mathbb{R}^n$
3. $A^TAV = V(\Sigma^T\Sigma)$ *and*

$$\begin{cases} A^TAv_i = \sigma_i^2 v_i, & i = 1, \dots, r \\ A^TAv_i = 0, & i = r+1, \dots, r+(n-r) \end{cases} \implies \begin{cases} \text{Row}(A^TA) &= \text{span}\{v_1, \dots, v_r\} \\ \text{Null}(A^TA) &= \text{span}\{v_{r+1}, \dots, v_{r+(n-r)}\} \end{cases}$$

4. $U^TA = \Sigma V^T$ *and*

$$\begin{cases} u_i^TA = \sigma_i v_i^T, & i = 1, \dots, r \\ u_i^TA = 0, & i = r+1, \dots, r+(m-r) \end{cases} \implies \begin{cases} \text{Col}(A) &= \text{span}\{u_1, \dots, u_r\} \\ \text{Null}(A^T) &= \text{span}\{u_{r+1}, \dots, u_{r+(m-r)}\} \end{cases}$$

5. $AA^T = U(\Sigma\Sigma^T)U^T : \mathbb{R}^m \to \mathbb{R}^m$
6. $AA^TU = U(\Sigma\Sigma^T)$ *and*

$$\begin{cases} AA^Tu_i = \sigma_i^2 u_i, & i = 1, \dots, r \\ AA^Tu_i = 0, & i = r+1, \dots, r+(m-r) \end{cases} \implies \begin{cases} \text{Row}(AA^T) &= \text{span}\{u_1, \dots, u_r\} \\ \text{Null}(AA^T) &= \text{span}\{u_{r+1}, \dots, u_{r+(m-r)}\} \end{cases}$$

Proof of (1).

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

So,

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

Hence,

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

Proof of (2).

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

Proof of (3).

$$A^TAV = (V\Sigma^T U^T)(U\Sigma V^T)V = V\Sigma^T(U^TU)\Sigma(V^TV) = V(\Sigma^T\Sigma)$$

So,

$$\begin{aligned}
A^TAV &= \left[\ A^TAv_1 \mid \cdots \mid A^TAv_r \mid A^TAv_{r+1} \mid \cdots \mid A^TAv_n\ \right] \\
&= \left[\ \lambda_1 v_1 \mid \cdots \mid \lambda_r v_r \mid \lambda_{r+1} v_{r+1} \mid \cdots \mid \lambda_n v_n\ \right] \\
&= \left[\ v_1 \mid \cdots \mid v_r \mid v_{r+1} \mid \cdots \mid v_n\ \right] \begin{bmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \lambda_r & & \\ 0 & & & 0 & \\ 0 & 0 & \ldots & & \ddots \end{bmatrix} \\
&= \left[\ v_1 \mid \cdots \mid v_r \mid v_{r+1} \mid \cdots \mid v_n\ \right] \begin{bmatrix} \sigma_1^2 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \sigma_r^2 & & \\ 0 & & & 0 & \\ 0 & 0 & \ldots & & \ddots \end{bmatrix} \\
&= \left[\ \sigma_1^2 v_1 \mid \cdots \mid \sigma_r^2 v_r \mid 0 \mid \cdots \mid 0\ \right].
\end{aligned}$$

Hence,

1. $A^TAv_1 = \sigma_1^2 v_1,\ \ldots,\ A^TAv_r = \sigma_r^2 v_r$, and
2. $A^TAv_{r+1} = 0,\ \ldots,\ A^TAv_{r+(n-r)} = 0$.

Proof of (4).

$$U^TA = U^T(U\Sigma V^T) = (U^TU)\Sigma V^T = \Sigma V^T$$

So,

$$U^T A = \begin{bmatrix} u_1^T \\ \hline \vdots \\ \hline u_r^T \\ \hline u_{r+1}^T \\ \hline \vdots \\ \hline u_m^T \end{bmatrix} A = \begin{bmatrix} u_1^T A \\ \hline \vdots \\ \hline u_r^T A \\ \hline u_{r+1}^T A \\ \hline \vdots \\ \hline u_m^T A \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 & 0 & \dots & & \dots & 0 & \dots & 0 \\ 0 & \ddots & & & & \vdots & & \vdots \\ \vdots & & \sigma_r & & & & \ddots & \\ 0 & & & 0 & & \vdots & & \vdots \\ 0 & 0 & \dots & & \ddots & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ \hline \vdots \\ \hline v_r^T \\ \hline v_{r+1}^T \\ \hline \vdots \\ \hline v_n^T \end{bmatrix} = \begin{bmatrix} \sigma_1 v_1^T \\ \hline \vdots \\ \hline \sigma_r v_r^T \\ \hline 0 \\ \hline \vdots \\ \hline 0 \end{bmatrix}$$

Hence,

1. $u_1^T A = \sigma_1 v_1^T, \dots, u_r^T A = \sigma_r v_r^T$, and
2. $u_{r+1}^T A = 0, \dots, u_{r+(m-r)}^T A = 0$.

Proof of (5).

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

Proof of (6).

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

So,

$$\begin{aligned}
AA^TU &= \left[\ AA^Tu_1 \mid \cdots \mid AA^Tu_r \mid AA^Tu_{r+1} \mid \cdots \mid AA^Tu_m\ \right] \\
&= \left[\ \lambda_1 u_1 \mid \cdots \mid \lambda_r u_r \mid \lambda_{r+1}u_{r+1} \mid \cdots \mid \lambda_n u_m\ \right] \\
&= \left[\ u_1 \mid \cdots \mid u_r \mid u_{r+1} \mid \cdots \mid u_m\ \right] \begin{bmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \lambda_r & & \\ 0 & & & 0 & \\ 0 & 0 & \dots & & \ddots \end{bmatrix} \\
&= \left[\ u_1 \mid \cdots \mid u_r \mid u_{r+1} \mid \cdots \mid u_m\ \right] \begin{bmatrix} \sigma_1^2 & 0 & \dots & \dots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \sigma_r^2 & & \\ 0 & & & 0 & \\ 0 & 0 & \dots & & \ddots \end{bmatrix} \\
&= \left[\ \sigma_1^2 u_1 \mid \cdots \mid \sigma_r^2 u_r \mid 0 \mid \cdots \mid 0\ \right].
\end{aligned}$$

Hence,

1. $AA^Tu_1 = \sigma_1^2 u_1, \dots, AA^Tu_r = \sigma_r^2 u_r$, and
2. $AA^Tu_{r+1} = 0, \dots, AA^Tu_{r+(m-r)} = 0$.

■

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

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

Proof.

$$
\begin{aligned}
A &= U\Sigma V^T \\
&= \begin{bmatrix} \sigma_1 u_{11} & \sigma_2 u_{12} & \dots & \sigma_r u_{1r} \\ \sigma_1 u_{21} & \sigma_2 u_{22} & \dots & \sigma_r u_{2r} \\ \vdots & & & \vdots \\ \sigma_1 u_{(r-1)1} & & & \\ \sigma_1 u_{r1} & \sigma_2 u_{r2} & \dots & \sigma_r u_{rr} \end{bmatrix} \begin{bmatrix} v_{11}^T & v_{12}^T & \dots & v_{1r}^T \\ v_{21}^T & \ddots & & \\ \vdots & & & \vdots \\ v_{r1}^T & v_{r2}^T & \dots & v_{rr}^T \end{bmatrix} \\
&= \begin{bmatrix} (\sigma_1 u_{11} v_{11}^T + \cdots + \sigma_r u_{1r} v_{r1}^T) & (\sigma_1 u_{11} v_{12}^T + \cdots + \sigma_r u_{1r} v_{r2}^T) & \dots & (\sigma_1 u_{11} v_{1r}^T + \cdots + \sigma_r u_{1r} v_{rr}^T) \\ (\sigma_1 u_{21} v_{11}^T + \cdots + \sigma_r u_{2r} v_{r1}^T) & & & \\ \vdots & & & \\ (\sigma_1 u_{r1} v_1^T + \cdots + \sigma_r u_{rr} v_{r1}^T) & \dots & \dots & (\sigma_1 u_{r1} v_{1r}^T + \cdots + \sigma_r u_{rr} v_{rr}^T) \end{bmatrix} \\
&= \sigma_1 u_1 v_1^T + \\
&\quad \begin{bmatrix} (\sigma_2 u_{12} v_{21}^T + \cdots + \sigma_r u_{1r} v_{r1}^T) & (\sigma_2 u_{12} v_{22}^T + \cdots + \sigma_r u_{1r} v_{r2}^T) & \dots & (\sigma_2 u_{12} v_{2r}^T + \cdots + \sigma_r u_{1r} v_{rr}^T) \\ (\sigma_2 u_{22} v_{21}^T + \cdots + \sigma_r u_{2r} v_{r1}^T) & & & \vdots \\ \vdots & & & \\ (\sigma_2 u_{r2} v_{21}^T + \cdots + \sigma_r u_{rr} v_{2r}^T) & \dots & \dots & (\sigma_2 u_{rr} v_{2r}^T + \cdots + \sigma_r u_{rr} v_{rr}^T) \end{bmatrix} \\
&= \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T \\
&= u_1 \sigma_1 v_1^T + u_2 \sigma_2 v_2^T + \cdots + u_r \sigma_r v_r^T \\
&= \sum_{i=1}^{r} u_i \sigma_i v_i^T
\end{aligned}
$$

■

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*

$$\Delta_{A^TA}(\lambda) = \begin{vmatrix} 21-\lambda & 10 & 11 & 12 \\ 10 & 5-\lambda & 5 & 5 \\ 11 & 5 & 6-\lambda & 7 \\ 12 & 5 & 7 & 9-\lambda \end{vmatrix} = \lambda^4 - 41\lambda^3 + 85\lambda^2 \implies \begin{matrix} \lambda_1 = \frac{41+3\sqrt{149}}{2} \\ \lambda_2 = \frac{41-3\sqrt{149}}{2} \end{matrix} \implies$$

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

$$S^T = \begin{bmatrix} \frac{71+3\sqrt{149}}{50} & 1 & \frac{21+3\sqrt{149}}{50} & \frac{-4+3\sqrt{149}}{25} \\ \frac{71-3\sqrt{149}}{50} & 1 & \frac{21-3\sqrt{149}}{50} & \frac{-4-3\sqrt{149}}{25} \\ -1 & 1 & 1 & 0 \\ -2 & 3 & 0 & 1 \end{bmatrix} \qquad \begin{aligned} ||s_1|| &= \frac{\sqrt{16{,}092+456\sqrt{149}}}{50} \\ ||s_2|| &= \frac{\sqrt{16{,}092-456\sqrt{149}}}{50} \\ ||s_3|| &= \sqrt{3} \\ ||s_4|| &= \sqrt{14} \end{aligned}$$

S^T contains the transposed eigenvectors of A^TA.

$$V^T = \begin{bmatrix} \frac{71+3\sqrt{149}}{\sqrt{16{,}092+456\sqrt{149}}} & \frac{50}{\sqrt{16{,}092+456\sqrt{149}}} & \frac{21+3\sqrt{149}}{\sqrt{16{,}092+456\sqrt{149}}} & \frac{-8+6\sqrt{149}}{\sqrt{16{,}092+456\sqrt{149}}} \\ \frac{71-3\sqrt{149}}{\sqrt{16{,}092-456\sqrt{149}}} & \frac{50}{\sqrt{16{,}092-456\sqrt{149}}} & \frac{21-3\sqrt{149}}{\sqrt{16{,}092-456\sqrt{149}}} & \frac{-8-6\sqrt{149}}{\sqrt{16{,}092-456\sqrt{149}}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 \\ -\frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} & 0 & \frac{1}{\sqrt{14}} \end{bmatrix}$$

$$u_1 = \frac{Av_1}{\sigma_1} = \left(\frac{2}{\sqrt{82+6\sqrt{149}}}\right) \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 \\ 4 & 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 71+3\sqrt{149} \\ 50 \\ 21+3\sqrt{149} \\ -8+6\sqrt{149} \end{bmatrix} \left(\frac{1}{\sqrt{16{,}092+456\sqrt{149}}}\right) = \begin{bmatrix} \frac{152+36\sqrt{149}}{\sqrt{1{,}727{,}208+133{,}944\sqrt{149}}} \\ \frac{410+30\sqrt{149}}{\sqrt{1{,}727{,}208+133{,}944\sqrt{149}}} \\ \frac{820+60\sqrt{149}}{\sqrt{1{,}727{,}208+133{,}944\sqrt{149}}} \end{bmatrix}$$

$$u_2 = \frac{Av_2}{\sigma_2} = \left(\frac{2}{\sqrt{82-6\sqrt{149}}}\right) \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 1 \\ 4 & 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} 71-3\sqrt{149} \\ 50 \\ 21-3\sqrt{149} \\ -8-6\sqrt{149} \end{bmatrix} \left(\frac{1}{\sqrt{16{,}092-456\sqrt{149}}}\right) = \begin{bmatrix} \frac{152-36\sqrt{149}}{\sqrt{1{,}727{,}208-133{,}944\sqrt{149}}} \\ \frac{410-30\sqrt{149}}{\sqrt{1{,}727{,}208-133{,}944\sqrt{149}}} \\ \frac{820-60\sqrt{149}}{\sqrt{1{,}727{,}208-133{,}944\sqrt{149}}} \end{bmatrix}$$

Since A is 3×4, U should be a 3×3 matrix. However, there is no σ_3, so we cannot use $u_i = \frac{Av_i}{\sigma_i}$ to find u_3. Instead, we use the left null space of A.

$$\text{Null}(A^T) = \left[\begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 1 & 1 & 2 & 0 \\ 2 & 1 & 2 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & -3 & -6 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 2 & 4 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \implies$$

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

Normalizing y to obtain u_3 :

$$u_3 = \frac{1}{||y||}y = \frac{1}{\sqrt{5}} \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

$$So,\ U = \begin{bmatrix} \frac{152+36\sqrt{149}}{\sqrt{1{,}727{,}208+133{,}944\sqrt{149}}} & \frac{152-36\sqrt{149}}{\sqrt{1{,}727{,}208-133{,}944\sqrt{149}}} & 0 \\ \frac{410+30\sqrt{149}}{\sqrt{1{,}727{,}208+133{,}944\sqrt{149}}} & \frac{410-30\sqrt{149}}{\sqrt{1{,}727{,}208-133{,}944\sqrt{149}}} & \frac{-2}{\sqrt{5}} \\ \frac{820+60\sqrt{149}}{\sqrt{1{,}727{,}208+133{,}944\sqrt{149}}} & \frac{820-60\sqrt{149}}{\sqrt{1{,}727{,}208-133{,}944\sqrt{149}}} & \frac{1}{\sqrt{5}} \end{bmatrix}.$$

Thus, $A =$

$$\begin{bmatrix} \frac{152+36\sqrt{149}}{\sqrt{1{,}727{,}208+133{,}944\sqrt{149}}} & \frac{152-36\sqrt{149}}{\sqrt{1{,}727{,}208-133{,}944\sqrt{149}}} & 0 \\ \frac{410+30\sqrt{149}}{\sqrt{1{,}727{,}208+133{,}944\sqrt{149}}} & \frac{410-30\sqrt{149}}{\sqrt{1{,}727{,}208-133{,}944\sqrt{149}}} & \frac{-2}{\sqrt{5}} \\ \frac{820+60\sqrt{149}}{\sqrt{1{,}727{,}208+133{,}944\sqrt{149}}} & \frac{820-60\sqrt{149}}{\sqrt{1{,}727{,}208-133{,}944\sqrt{149}}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{\sqrt{82+6\sqrt{149}}}{2} & 0 & 0 & 0 \\ 0 & \frac{\sqrt{82-6\sqrt{149}}}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \frac{71+3\sqrt{149}}{\sqrt{16{,}092+456\sqrt{149}}} & \frac{50}{\sqrt{16{,}092+456\sqrt{149}}} & \frac{21+3\sqrt{149}}{\sqrt{16{,}092+456\sqrt{149}}} & \frac{-8+6\sqrt{149}}{\sqrt{16{,}092+456\sqrt{149}}} \\ \frac{71-3\sqrt{149}}{\sqrt{16{,}092-456\sqrt{149}}} & \frac{50}{\sqrt{16{,}092-456\sqrt{149}}} & \frac{21-3\sqrt{149}}{\sqrt{16{,}092-456\sqrt{149}}} & \frac{-8-6\sqrt{149}}{\sqrt{16{,}092-456\sqrt{149}}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 \\ -\frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} & 0 & \frac{1}{\sqrt{14}} \end{bmatrix} = U\Sigma V^T.$$

Example 2. *Let* $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

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

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

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

$\Delta_{A^TA}(\lambda) = \Delta_{AA^T}(\lambda) = |A^TA - \lambda I| = |AA^T - \lambda I| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix}$

$= (2-\lambda)(1-\lambda) - 1^2 = \lambda^2 - 3\lambda + 1 \implies \begin{matrix} \lambda_1 = \frac{3+\sqrt{5}}{2} \\ \lambda_2 = \frac{3-\sqrt{5}}{2} \end{matrix}$

$\begin{matrix} \sigma_1 = \sqrt{\lambda_1} = \sqrt{\frac{3+\sqrt{5}}{2}} = \frac{\sqrt{6+2\sqrt{5}}}{2} = \frac{1+\sqrt{5}}{2} \\ \sigma_2 = \sqrt{\lambda_2} = \sqrt{\frac{3-\sqrt{5}}{2}} = \frac{\sqrt{6-2\sqrt{5}}}{2} = \frac{1-\sqrt{5}}{2} \end{matrix} \implies \Sigma = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}$

$\boxed{v_1\ \&\ u_1:}$

$[A^TA - \lambda_1 I]x_1 = [AA^T - \lambda_1 I]s_1 = \left[\begin{array}{cc|c} 2 - \frac{3+\sqrt{5}}{2} & 1 & 0 \\ 1 & 1 - \frac{3+\sqrt{5}}{2} & 0 \end{array}\right] = \left[\begin{array}{cc|c} \frac{1-\sqrt{5}}{2} & 1 & 0 \\ 1 & \frac{-1-\sqrt{5}}{2} & 0 \end{array}\right] \rightarrow$

$\left[\begin{array}{cc|c} 1 & \frac{-1-\sqrt{5}}{2} & 0 \\ \frac{1-\sqrt{5}}{2} & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & \frac{-1-\sqrt{5}}{2} & 0 \\ 0 & 0 & 0 \end{array}\right] \implies s_1 = \begin{bmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{bmatrix}$

$||s_1|| = \sqrt{\frac{6+2\sqrt{5}}{4} + 1} = \sqrt{\frac{10+2\sqrt{5}}{4}} = \frac{\sqrt{10+2\sqrt{5}}}{2}$

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

$\boxed{v_2\ \&\ u_2:}$

$$[A^TA-\lambda_2 I]s_2=[AA^T-\lambda_2 I]s_2=\left[\begin{array}{cc|c} 2-\frac{3-\sqrt{5}}{2} & 1 & 0 \\ 1 & 1-\frac{3-\sqrt{5}}{2} & 0 \end{array}\right]=\left[\begin{array}{cc|c} \frac{1+\sqrt{5}}{2} & 1 & 0 \\ 1 & \frac{-1+\sqrt{5}}{2} & 0 \end{array}\right]\rightarrow$$

$$\left[\begin{array}{cc|c} 1 & \frac{-1+\sqrt{5}}{2} & 0 \\ \frac{1+\sqrt{5}}{2} & 1 & 0 \end{array}\right]\rightarrow\left[\begin{array}{cc|c} 1 & \frac{-1+\sqrt{5}}{2} & 0 \\ 0 & 0 & 0 \end{array}\right]\implies s_2=\begin{bmatrix}\frac{1-\sqrt{5}}{2}\\ 1\end{bmatrix}$$

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

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

$$V^T=\begin{bmatrix}\frac{1+\sqrt{5}}{\sqrt{10+2\sqrt{5}}} & \frac{2}{\sqrt{10+2\sqrt{5}}}\\ \frac{1-\sqrt{5}}{\sqrt{10-2\sqrt{5}}} & \frac{2}{\sqrt{10-2\sqrt{5}}}\end{bmatrix}, U=\begin{bmatrix}\frac{1+\sqrt{5}}{\sqrt{10+2\sqrt{5}}} & \frac{1-\sqrt{5}}{\sqrt{10-2\sqrt{5}}}\\ \frac{2}{\sqrt{10+2\sqrt{5}}} & \frac{2}{\sqrt{10-2\sqrt{5}}}\end{bmatrix}$$

To verify that $v_1=u_1$ and $v_2=u_2$:

$$Av_1=\begin{bmatrix}1&1\\1&0\end{bmatrix}\begin{bmatrix}\frac{1+\sqrt{5}}{\sqrt{10+2\sqrt{5}}}\\ \frac{2}{\sqrt{10+2\sqrt{5}}}\end{bmatrix}=\begin{bmatrix}\frac{3+\sqrt{5}}{\sqrt{10+2\sqrt{5}}}\\ \frac{1+\sqrt{5}}{\sqrt{10+2\sqrt{5}}}\end{bmatrix}=\begin{bmatrix}\frac{6+2\sqrt{5}}{2\sqrt{10+2\sqrt{5}}}\\ \frac{1+\sqrt{5}}{\sqrt{10+2\sqrt{5}}}\end{bmatrix}=\frac{1+\sqrt{5}}{2}\begin{bmatrix}\frac{1+\sqrt{5}}{\sqrt{10+2\sqrt{5}}}\\ \frac{2}{\sqrt{10+2\sqrt{5}}}\end{bmatrix}=\sigma_1u_1=\sigma_1v_1$$

$$Av_2=\begin{bmatrix}1&1\\1&0\end{bmatrix}\begin{bmatrix}\frac{1-\sqrt{5}}{\sqrt{10-2\sqrt{5}}}\\ \frac{2}{\sqrt{10-2\sqrt{5}}}\end{bmatrix}=\begin{bmatrix}\frac{3-\sqrt{5}}{\sqrt{10-2\sqrt{5}}}\\ \frac{1-\sqrt{5}}{\sqrt{10-2\sqrt{5}}}\end{bmatrix}=\begin{bmatrix}\frac{6-2\sqrt{5}}{2\sqrt{10-2\sqrt{5}}}\\ \frac{1-\sqrt{5}}{\sqrt{10-2\sqrt{5}}}\end{bmatrix}=\frac{1-\sqrt{5}}{2}\begin{bmatrix}\frac{1-\sqrt{5}}{\sqrt{10-2\sqrt{5}}}\\ \frac{2}{\sqrt{10-2\sqrt{5}}}\end{bmatrix}=\sigma_2u_2=\sigma_2v_2$$

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,

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

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 `img_r`,`img_g`, 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. `posint` indicates that `n` must be a positive integer.

```
approx:=proc(img_r,img_g,img_b,n::posint)local
Singr,Singg,Singb,Ur,Ug,Ub,Vtr,Vtg,Vtb,singr,
```

```
singg,singb,ur,ug,ub,vr,vg,vb,Mr,Mg,Mb,i,img_rgb;
```

In place of a Σ matrix, we create a list of the σ'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.

```
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 σ_i, u_i, and v_i^T to create the $\sigma_i u_i v_i^T$ dyads for $i = 1 \ldots r$:

```
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, $rank(A) = r = \min(m, n)$.

Adding the dyads to produce the approximations of each layer:

```
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:

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

Displaying the result in 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 $100 \times 200 \times 3$, 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 $\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T$.

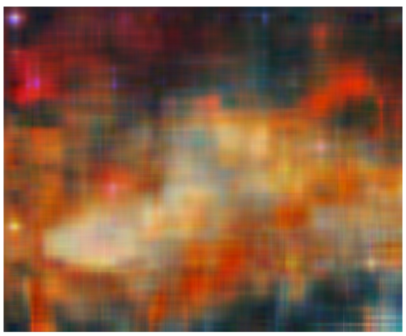
$$\sigma_1 \begin{bmatrix} \\ u_1 \\ \\ \end{bmatrix} \begin{bmatrix} - & v_1^T & - \end{bmatrix} + \sigma_2 \begin{bmatrix} \\ u_2 \\ \\ \end{bmatrix} \begin{bmatrix} - & v_2^T & - \end{bmatrix}$$

The σ'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 σ_1 and σ_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 σ_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 σ 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

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

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,

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

and,

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

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

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

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

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 $A^T(Ax) = (A^T A)x = 0$.
   $\therefore x \in \text{Null}(A^T A)$.

Proof of ($\supseteq$). 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,

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

   ∎

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

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. $\text{rank}(A^T) = \text{rank}(AA^T) = \text{rank}(A)$, 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 $\text{Null}(A^T A) = \text{Null}(A)$, 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. $\sigma_i \neq 0$ *for* $i = 1 \ldots r$.

Lemma 5. *Since the vectors $\{v_1, \dots, v_r\}$ are orthonormal, vectors $\{u_1, \dots, u_r\}$ computed as*

$$u_i = \frac{Av_i}{\sigma_i}, \quad i = 1 \dots r,$$

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^Tv has rank 1.*

Proof. We compute the following:

$$vv^T = \underbrace{\begin{bmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{bmatrix}}_{n\times 1} \underbrace{\begin{bmatrix} v_{11}^T & v_{21}^T & \dots & v_{n1}^T \end{bmatrix}}_{1\times n} = \underbrace{\begin{bmatrix} v_{11}v_{11}^T & v_{11}v_{12}^T & \dots & v_{11}v_{1n}^T \\ v_{21}v_{11}^T & v_{21}v_{12}^T & \dots & v_{21}v_{1n}^T \\ \vdots & & \ddots & \vdots \\ v_{n1}v_{11}^T & v_{n1}v_{1n}^T & \dots & v_{n1}v_{1n}^T \end{bmatrix}}_{n\times n} = \begin{bmatrix} - & v_{11}v^T & - \\ - & v_{21}v^T & - \\ & \vdots & \\ - & v_{n1}v^T & - \end{bmatrix}.$$

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 $A^TAv = A^T\lambda v = \lambda A^Tv = \lambda Av = \lambda^2 v$. Hence, λ^2 is an eigenvalue of A^TA. Therefore, λ is a singular value of A.

On the other hand, let $\sigma > 0$ be a singular value of A. So, $A^TAv = \sigma^2 v$ for some nonzero eigenvector v. $\Delta_{A^TA}(\sigma^2) = \det(A^TA - \sigma^2 I) = \det(A^2 - \sigma^2 I) = \det((A - \sigma I)(A + \sigma I)) = \det(A - \sigma I)\det(A + \sigma I) = 0$, 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 σ is an eigenvalue of A. ∎

Lemma 8. *Let A be an $m \times n$ matrix. Then, A^TA and AA^T share the same nonzero eigenvalues and, therefore, both provide the singular values of A. In the case where $m = n$, $\Delta_{A^TA}(t) = \Delta_{AA^T}(t)$.*

Proof. From parts 2 and 5 of Theorem (1), we have $A^TA = V\Sigma^T\Sigma V^T = V\Lambda_n V^{-1}$, and $AA^T = U\Sigma\Sigma^T U^T = U\Lambda_m U^{-1}$, where Λ_n is $n \times n$, and Λ_m is $m \times m$. Hence, A^TA is similar to Λ_n, and AA^T is similar to Λ_m.

We cannot show that Λ_n is similar to Λ_m. However, from Definition (1), we know what Σ and Σ^T look like, so we know that when we multiply them together to obtain Λ_n and Λ_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.

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

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

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

$$\begin{aligned} \Delta_{A^TA}(t) = \det(\Lambda_n - tI) &= \underbrace{(\lambda_1 - t)(\lambda_2 - t)\ldots(\lambda_r - t)}_{r \text{ factors}}\underbrace{(-t)\ldots(-t)}_{n-r \text{ factors}} \\ &= (-t)^{n-r}(\lambda_1 - t)(\lambda_2 - t)\ldots(\lambda_r - t) \\ &= (-t)^{n-r}h(t), \end{aligned}$$

and

$$\begin{aligned} \Delta_{AA^T}(t) = \det(\Lambda_m - tI) &= \underbrace{(\lambda_1 - t)(\lambda_2 - t)\ldots(\lambda_r - t)}_{r \text{ factors}}\underbrace{(-t)\ldots(-t)}_{m-r \text{ factors}} \\ &= (-t)^{m-r}(\lambda_1 - t)(\lambda_2 - t)\ldots(\lambda_r - t) \\ &= (-t)^{m-r}h(t), \end{aligned}$$

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 $\Delta_{A^TA}(t) = \Delta_{AA^T}(t) = (-t)^{n-r}h(t)$, and $\Lambda_m = \Lambda_n$. ■

References

[1] Maple (2018). Maplesoft, A division of Waterloo Maple Inc., Waterloo, Ontario, `http://maplesoft.com`.

[2] APOD, 23 April 2018, "The Blue Horsehead Nebula in Infrared." *Astronomy Picture of the Day,* NASA, April 2018.

[3] D. Poole, *Linear Algebra: A Modern Introduction*, Stamford, Cengage Learning, 2015.

[4] G. Strang, *Introduction to Linear Algebra*, Wellesley, Wellesley-Cambridge Press, 2010.