Singular Value Decomposition Example

Paul Skoufranis

March 11, 2010

We shall explore a concept known as the singular value decomposition of a matrix.

Definition Let A be an $n \times m$ matrix. The singular values of A are the square roots of the eigenvalues of $m \times m$ matrix $A^T A$ listed according to their algebraic multiplicity. We should note that the eigenvalues of $A^T A$ will always be positive numbers so it makes sense to take the square root. It is customary to denote the singular values by σ_1, σ_2, …, σ_m with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m$.

Theorem Any $n \times m$ matrix A can be written in the form $A = U\Sigma V^T$ where U is an orthogonal $n \times n$ matrix, V is an orthogonal $m \times m$ matrix, and Σ is an $n \times m$ matrix whose first r diagonal entries are the non-zero singular values σ_1, σ_2, …, σ_r of A and whose other entries are all zeros. The expression $U\Sigma V^T$ is known as the Singular Value Decomposition of A.

Now we shall go through an example of computing the Singular Value Decomposition (SVD) of a matrix A, why it works, and how we can use the SVD to determine the action of A on the unit circle $x^2+y^2 = 1$.

Example Compute the SVD of the matrix

$$A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ \sqrt{3} & 0 \end{bmatrix}$$

and use this decomposition to describe the action of A on the unit circle.

Solution: The first step in finding the SVD of A is to find the singular values of A. To do this, we simply need to compute the eigenvalues of $A^T A$. Thus let us first compute $A^T A$;

$$A^T A = \begin{bmatrix} 1 & 1 & \sqrt{3} \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ \sqrt{3} & 0 \end{bmatrix} = \begin{bmatrix} 5 & -2 \\ -2 & 2 \end{bmatrix}$$

Therefore the characteristic polynomial of $A^T A$ is

$$f_{A^T A}(\lambda) = det(A^T A - \lambda I) = (5 - \lambda)(2 - \lambda) - 4 = \lambda^2 - 7\lambda + 6 = (\lambda - 1)(\lambda - 6)$$

Therefore the eigenvalues of $A^T A$ are 1 and 6. Thus the singular values of A are $\sigma_1 = 1$ and $\sigma_2 = \sqrt{6}$ (in general, at this step you would take the square root of each eigenvalue λ to get a singular value σ and include it $am(\lambda)$ times). This tells us that the matrix Σ will be the 3×2 matrix

$$\Sigma = \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

where we place $\sigma_1 = \sqrt{6}$ in the first column and $\sigma_2 = 1$ in the second column since $\sqrt{6} > 1$ (normally, if $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are your non-zero singular values (including repetitions), we would place these entries along the diagonal of Σ and place zeros elsewhere).

Our next goal is to determine the matrix V. We know that since A^TA is a symmetric matrix, A^TA will have an orthonormal basis of eigenvectors. It turns out that the matrix V will be 2×2 matrix whose columns are an orthonormal eigenbasis for A^TA. Therefore we compute the eigenspaces of A^TA. We notice that

$$\mathcal{E}_6 = ker(A^TA - 6I) = ker\left(\begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix}\right) = ker\left(\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}\right) = span\left\{\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right\}$$

and

$$\mathcal{E}_1 = ker(A^TA - I) = ker\left(\begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}\right) = ker\left(\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix}\right) = span\left\{\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$$

Let

$$\vec{v}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Then we let V be the 2×2 matrix

$$V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

where we place the eigenvector for the eigenvalue 6 in the first column and the eigenvector for the eigenvalue 1 in the second column since we placed $\sqrt{6}$ in first column of Σ and $\sqrt{1}$ in the second column of Σ.

Now our question is how do we find the orthogonal 3×3 matrix U? Well, if we could find an orthogonal 3×3 matrix U such that $AV = U\Sigma$, we would be done since $V^{-1} = V^T$ as V is an orthogonal matrix. Let us consider the actions of AV and Σ on the standard basis $\vec{e}_1$ and $\vec{e}_2$ of $\mathbb{R}^2$. Well

$$\Sigma\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \sigma_1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \Sigma\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \sigma_2\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

and

$$AV\begin{bmatrix} 1 \\ 0 \end{bmatrix} = A\vec{v}_1 \qquad AV\begin{bmatrix} 0 \\ 1 \end{bmatrix} = A\vec{v}_2$$

Therefore, if U were an orthogonal 3×3 matrix such that $AV = U\Sigma$ and $\vec{u}_1, \vec{u}_2$, and $\vec{u}_3$ were the columns of U, then $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ would be an orthonormal basis of $\mathbb{R}^3$ and

$$A\vec{v}_1 = AV\begin{bmatrix} 1 \\ 0 \end{bmatrix} = U\Sigma\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \sqrt{6}U\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \sigma_1\vec{u}_1$$

and

$$A\vec{v}_2 = AV\begin{bmatrix} 0 \\ 1 \end{bmatrix} = U\Sigma\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \sqrt{1}U\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \sigma_2\vec{u}_2$$

Thus we should let

$$\vec{u}_1 = \frac{1}{\sigma_1}A\vec{v}_1 = \frac{1}{6}\begin{bmatrix} 1 & -1 \\ 1 & -1 \\ \sqrt{3} & 0 \end{bmatrix}\left(\frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right) = \frac{1}{\sqrt{30}}\begin{bmatrix} 3 \\ 3 \\ 2\sqrt{3} \end{bmatrix}$$

and

$$\vec{u}_2 = \frac{1}{\sigma_2}A\vec{v}_2 = \frac{1}{1}\begin{bmatrix} 1 & -1 \\ 1 & -1 \\ \sqrt{3} & 0 \end{bmatrix}\left(\frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 \\ -1 \\ \sqrt{3} \end{bmatrix}$$

If we check, we see that $\{\vec{u}_1, \vec{u}_2\}$ are indeed orthonormal vectors. The question is, why does this work? Well, just using the facts that $\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1$, $\vec{u}_2 = \frac{1}{\sigma_2} A\vec{v}_2$, $A^T A\vec{v}_1 = \sigma_1^2 \vec{v}_1$ (since $\vec{v}_1$ is an eigenvector for $A^T A$ with eigenvalue σ_1^2), $A^T A\vec{v}_2 = \sigma_2^2 \vec{v}_2$ ($\vec{v}_2$ is an eigenvector for $A^T A$ with eigenvalue σ_2^2), and $\{\vec{v}_1, \vec{v}_2\}$ is an orthonormal set, we see that

$$\begin{aligned}
\vec{u}_1 \cdot \vec{u}_1 &= \left(\frac{1}{\sigma_1} A\vec{v}_1\right) \cdot \left(\frac{1}{\sigma_1} A\vec{v}_1\right) \\
&= \left(\frac{1}{\sigma_1} A\vec{v}_1\right)^T \left(\frac{1}{\sigma_1} A\vec{v}_1\right) \\
&= \frac{1}{\sigma_1^2} \vec{v}_1^T A^T A\vec{v}_1 \\
&= \frac{1}{\sigma_1^2} \vec{v}_1^T (A^T A\vec{v}_1) \\
&= \frac{1}{\sigma_1^2} \vec{v}_1^T (\sigma_1^2 \vec{v}_1) \\
&= \vec{v}_1^T \vec{v}_1 \\
&= \vec{v}_1 \cdot \vec{v}_1 = 1
\end{aligned}$$

and

$$\begin{aligned}
\vec{u}_2 \cdot \vec{u}_2 &= \left(\frac{1}{\sigma_2} A\vec{v}_2\right) \cdot \left(\frac{1}{\sigma_2} A\vec{v}_2\right) \\
&= \left(\frac{1}{\sigma_2} A\vec{v}_2\right)^T \left(\frac{1}{\sigma_2} A\vec{v}_2\right) \\
&= \frac{1}{\sigma_2^2} \vec{v}_2^T A^T A\vec{v}_2 \\
&= \frac{1}{\sigma_2^2} \vec{v}_2^T (A^T A\vec{v}_2) \\
&= \frac{1}{\sigma_2^2} \vec{v}_2^T (\sigma_2^2 \vec{v}_2) \\
&= \vec{v}_2^T \vec{v}_2 \\
&= \vec{v}_2 \cdot \vec{v}_2 = 1
\end{aligned}$$

and

$$\begin{aligned}
\vec{u}_1 \cdot \vec{u}_2 &= \left(\frac{1}{\sigma_1} A\vec{v}_1\right) \cdot \left(\frac{1}{\sigma_2} A\vec{v}_2\right) \\
&= \left(\frac{1}{\sigma_1} A\vec{v}_1\right)^T \left(\frac{1}{\sigma_2} A\vec{v}_2\right) \\
&= \frac{1}{\sigma_1 \sigma_2} \vec{v}_1^T A^T A\vec{v}_2 \\
&= \frac{1}{\sigma_1 \sigma_2} \vec{v}_1^T (A^T A\vec{v}_2) \\
&= \frac{1}{\sigma_2 \sigma_2} \vec{v}_1^T (\sigma_2^2 \vec{v}_2) \\
&= \frac{\sigma_2}{\sigma_1} \vec{v}_1^T \vec{v}_2 \\
&= \frac{\sigma_2}{\sigma_1} \vec{v}_1 \cdot \vec{v}_2 = 0
\end{aligned}$$

which is want we wanted.

However, to find an orthogonal matrix using $\vec{u}_1$ and $\vec{u}_2$, we need a third orthonormal vector as we need U to be a 3×3 matrix and we only have two orthonormal vectors (in some cases, you will not need this step). To find $\vec{u}_3$, write

$$\vec{u}_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Then we want $a^2 + b^2 + c^2 = 1$ so $\vec{u}_3$ is unit length, and we want

$$0 = \vec{u}_1 \cdot \vec{u}_3 = \frac{1}{\sqrt{30}}(3a + 3b + 2\sqrt{3}c)$$

and

$$0 = \vec{u}_2 \cdot \vec{u}_3 = \frac{1}{\sqrt{5}}(-a - b + \sqrt{3}c)$$

By solving a system of linear equations, we obtain that

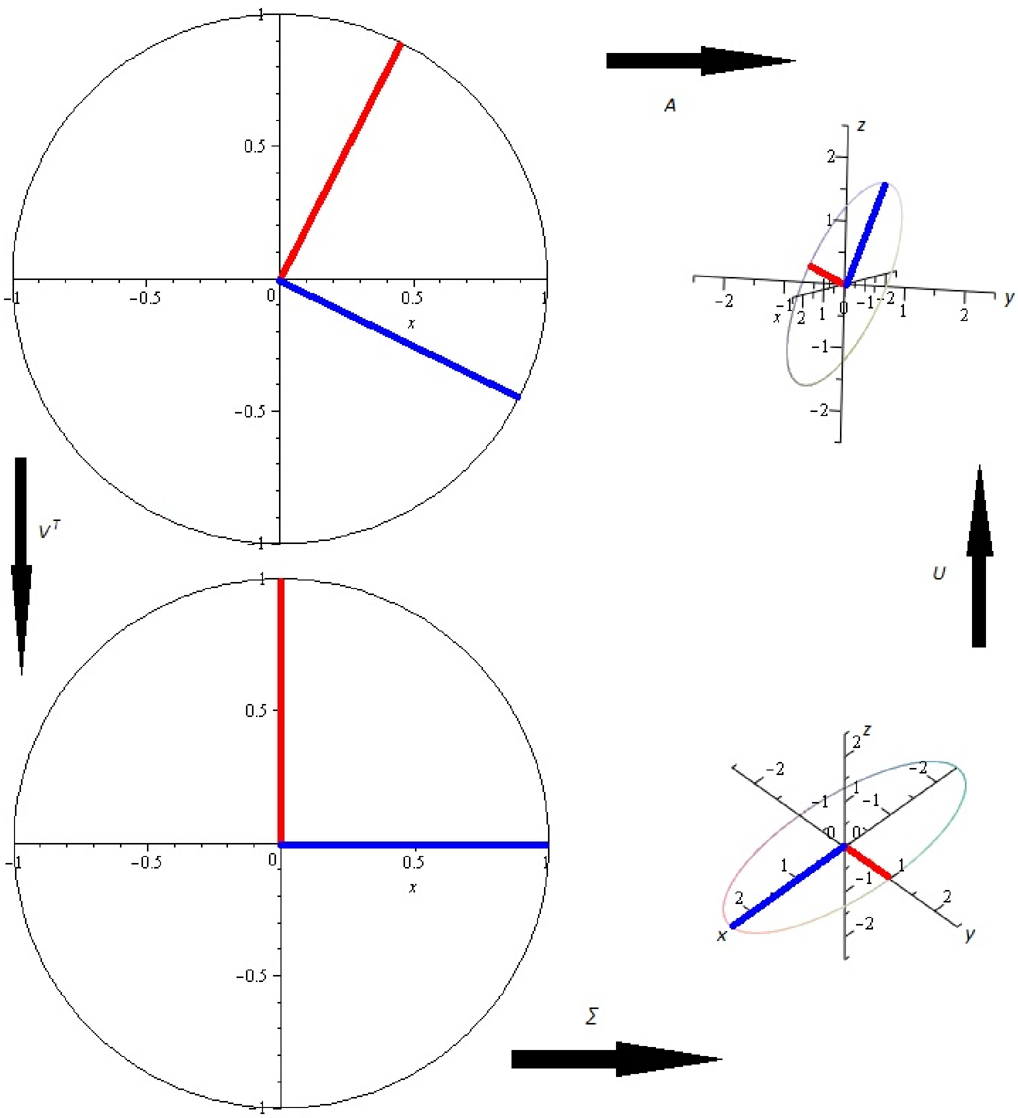
$$\alpha \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

where $\alpha \in \mathbb{R}$ is the set of all solutions to the equations $0 = \frac{1}{\sqrt{30}}(3a + 3b + 2\sqrt{3}c)$ and $0 = \frac{1}{\sqrt{5}}(-a - b + \sqrt{3}c)$. Therefore, if we let $\vec{u}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\vec{u}_3$ is a normal vector that is orthogonal to $\vec{u}_1$ and $\vec{u}_2$. Thus $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthonormal basis for $\mathbb{R}^3$. Then, if we let

$$U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{30}} & -\frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} \\ \frac{3}{\sqrt{30}} & -\frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{10}} & \frac{\sqrt{3}}{\sqrt{5}} & 0 \end{bmatrix}$$

then U is an orthogonal matrix and $A = U\Sigma V^T$ as desired. If you are not confident in your computations, you could easily check your answer by multiplying this out.

Now we desire to determine the action of A on the unit circle. To see this, we will first apply V^T to the unit circle, then Σ to that image, and then we will apply U. The following is a diagram representing the actions of each map:



In the diagram, we start with the top left diagram. The vector $\vec{v}_1$ and its images are shown in blue and the vector $\vec{v}_2$ and its images are shown in red. First we know that V take $\vec{e}_1$ to $\vec{v}_1$ and $\vec{e}_2$ to $\vec{v}_2$. Since $V^T = V^{-1}$ as V is an orthogonal matrix, V^T takes $\vec{v}_1$ to $\vec{e}_1$ and $\vec{v}_2$ to $\vec{e}_2$. Then Σ adds a z-axis and dilates the x-axis by $\sqrt{6}$. Then U rotates the whole diagram in an awkward manner (you probably do not need to know exactly what this rotation is but, it takes the x-axis to the line through $\vec{u}_1$, it takes the y-axis to the line through $\vec{u}_2$, and it takes the z-axis to the line through $\vec{u}_3$). It is hard to see what the curve in the final diagram looks like. The blue line goes to the point $\frac{\sqrt{6}}{\sqrt{30}}(3, 3, 2\sqrt{3})$ and the red line goes to $\frac{1}{\sqrt{5}}(-1, -1, \sqrt{3})$.