Numerical Linear Algebra

Singular Value Decomposition of a Matrix[1]

[1] BYJU'S, https://byjus.com/maths/singular-value-decomposition/

Singular Value Decomposition:

The Singular Value Decomposition of a matrix is a factorization of the matrix into three matrices. Thus, the singular value decomposition of matrix $\mathbf{A}$ can be expressed in terms of the factorization of $\mathbf{A}$ into the product of three matrices as $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\intercal$. Here, the columns of $\mathbf{U}$ and $\mathbf{V}$ are orthonormal, and the matrix $\mathbf{D}$ is diagonal with real positive entries.

Mathematically, the singular value decomposition of a matrix can be explained as follows:
Consider a matrix $\mathbf{A}$ of order $m \times n$. This can be uniquely decomposed as:

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\intercal$$

$\mathbf{U}$ is $m \times n$ and column orthogonal (that means its columns are eigenvectors of $\mathbf{A}\mathbf{A}^\intercal$)

$$\mathbf{A}\mathbf{A}^\intercal = \mathbf{U}\mathbf{D}\mathbf{V}^\intercal\mathbf{V}\mathbf{D}\mathbf{U}^\intercal = \mathbf{U}\mathbf{D}^2\mathbf{U}^\intercal$$

$\mathbf{V}$ is $n \times n$ and column orthogonal (that means its columns are eigenvectors of $\mathbf{A}^\intercal\mathbf{A}$)

$$\mathbf{A}^\intercal\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{U}^\intercal\,\mathbf{U}\mathbf{D}\mathbf{V}^\intercal = \mathbf{V}\mathbf{D}^2\mathbf{V}^\intercal$$

$\mathbf{D}$ is $n \times n$ diagonal, where non-negative real values are called singular values.

Let $\mathbf{D} = diag(\sigma_1, \sigma_2, \cdots, \sigma_n)$ ordered such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$. If σ is a singular value of $\mathbf{A}$, its square is an eigenvalue of $\mathbf{A}^\intercal\mathbf{A}$. Also, let $\mathbf{U} = (\mathbf{u}_1\,\mathbf{u}_2\,\cdots\,\mathbf{u}_n)$ and $\mathbf{V} = (\mathbf{v}_1\,\mathbf{v}_2\,\cdots\,\mathbf{v}_n)$. Therefore,

$$A = \sum_{i=1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^\intercal$$

Here, the sum can be given from 1 to r so that r is the rank of matrix $\mathbf{A}$.

Singular Value Decomposition 2×2 Matrix Example:

Question: Find the singular value decomposition of a matrix

$$\mathbf{A} = \begin{bmatrix} -4 & -7 \\ 1 & 4 \end{bmatrix}$$

Solution: Given,

$$\mathbf{A} = \begin{bmatrix} -4 & -7 \\ 1 & 4 \end{bmatrix} \quad \Rightarrow \quad \mathbf{A}^{\intercal} = \begin{bmatrix} -4 & 1 \\ -7 & 4 \end{bmatrix}$$

So,

$$\mathbf{A}^{\intercal}\mathbf{A} = \begin{bmatrix} -4 & 1 \\ -7 & 4 \end{bmatrix} \begin{bmatrix} -4 & -7 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 17 & 32 \\ 32 & 65 \end{bmatrix}$$

and,

$$\mathbf{A}\mathbf{A}^{\intercal} = \begin{bmatrix} -4 & -7 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} -4 & 1 \\ -7 & 4 \end{bmatrix} = \begin{bmatrix} 65 & -32 \\ -32 & 17 \end{bmatrix}$$

Finding the eigenvelues for $\mathbf{A}^{\mathsf{T}}\mathbf{A}$.

$$|\mathbf{A}^{\mathsf{T}}\mathbf{A} - \lambda I| = 0 \quad \Rightarrow \quad \begin{vmatrix} 17-\lambda & 32 \\ 32 & 65-\lambda \end{vmatrix} = 0$$

$$\begin{aligned}
&\Rightarrow \quad (17-\lambda)\times(65-\lambda) - (32)\times(32) = 0 \\
&\Rightarrow \quad (1105 - 82\lambda + \lambda^2) - 1024 = 0 \\
&\Rightarrow \quad (\lambda^2 - 82\lambda + 81) = 0 \\
&\Rightarrow \quad (\lambda - 1)(\lambda - 81) = 0 \\
&\Rightarrow \quad (\lambda - 1) = 0 \quad \text{or} \quad (\lambda - 81) = 0
\end{aligned}$$

$\therefore$ The eigenvalues of the matrix $\mathbf{A}^{\mathsf{T}}\mathbf{A}$ are given by $\lambda_1 = 81$ and $\lambda_2 = 1$.

Now, the eigenvector of $\mathbf{A}^\intercal\mathbf{A}$ for $\lambda_1 = 81$ is:

$$(\mathbf{A}^\intercal\mathbf{A} - 81I)\mathbf{v}_1 = 0$$

$$\Rightarrow \begin{pmatrix} 17-81 & 32 \\ 32 & 65-81 \end{pmatrix} \mathbf{v}_1 = 0 \qquad \Rightarrow \begin{pmatrix} -64 & 32 \\ 32 & -16 \end{pmatrix} \mathbf{v}_1 = 0$$

$$\Rightarrow \begin{pmatrix} -2 & 1 \\ 0 & 0 \end{pmatrix} \mathbf{v}_1 = 0$$

So the eigenvector corresponding to $\lambda_1 = 81$ is $(0.5, 1)^\intercal$ and its length is $\sqrt{0.5^2 + 1^2} = 1.118$

$\therefore$ Normalizing the eigenvector for $\lambda_1 = 81$ we get

$$\mathbf{v}_1 = \left(\frac{0.5}{1.118}, \frac{1}{1.118}\right)^\intercal = (0.4472, 0.8944)^\intercal$$

And, the eigenvector of $\mathbf{A}^\intercal\mathbf{A}$ for $\lambda_2 = 1$ is:

$$(\mathbf{A}^\intercal\mathbf{A} - 1I)\mathbf{v}_2 = 0$$

$$\Rightarrow \begin{pmatrix} 17-1 & 32 \\ 32 & 65-1 \end{pmatrix} \mathbf{v}_2 = 0 \qquad \Rightarrow \begin{pmatrix} 16 & 32 \\ 32 & 64 \end{pmatrix} \mathbf{v}_2 = 0$$

$$\Rightarrow \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \mathbf{v}_2 = 0$$

So the eigenvector corresponding to $\lambda_2 = 1$ is $(-2, 1)^\intercal$ and its length is $\sqrt{(-2)^2 + 1^2} = 2.236$

$\therefore$ Normalizing the eigenvector for $\lambda_2 = 1$ we get

$$\mathbf{v}_2 = \left(\frac{-2.0}{2.236}, \frac{1}{2.236}\right)^\intercal = (-0.8944, 0.4472)^\intercal$$

The eigenvalues of the matrix $\mathbf{AA}^\intercal$ are also $\lambda_1 = 81$ and $\lambda_2 = 1$.

So, the eigenvector of $\mathbf{AA}^\intercal$ for $\lambda_1 = 81$ is:

$$(\mathbf{AA}^\intercal - 81I)\mathbf{u}_1 = 0$$

$$\Rightarrow \begin{pmatrix} 65-81 & -32 \\ -32 & 17-81 \end{pmatrix} \mathbf{u}_1 = 0 \qquad \Rightarrow \begin{pmatrix} -16 & -32 \\ -32 & -64 \end{pmatrix} \mathbf{u}_1 = 0$$

$$\Rightarrow \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \mathbf{u}_1 = 0$$

So the eigenvector corresponding to $\lambda_1 = 81$ is $(-2, 1)^\intercal$ and its length is $\sqrt{(-2)^2 + 1^2} = 2.236$

$\therefore$ Normalizing the eigenvector for $\lambda_1 = 81$ we get

$$\mathbf{u}_1 = \left(\frac{-2}{2.236}, \frac{1}{2.236}\right)^\intercal = (-0.8944, 0.4472)^\intercal$$

So, the eigenvector of $\mathbf{AA}^\intercal$ for $\lambda_2 = 1$ is:

$$(\mathbf{AA}^\intercal - 1I)\mathbf{u}_2 = 0$$

$$\Rightarrow \begin{pmatrix} 65-1 & -32 \\ -32 & 17-1 \end{pmatrix} \mathbf{u}_2 = 0 \qquad \Rightarrow \begin{pmatrix} 64 & -32 \\ -32 & 16 \end{pmatrix} \mathbf{u}_2 = 0$$

$$\Rightarrow \begin{pmatrix} -2 & 1 \\ 0 & 0 \end{pmatrix} \mathbf{u}_2 = 0$$

So the eigenvector corresponding to $\lambda_2 = 1$ is $(0.5, 1)^\intercal$ and its length is $\sqrt{0.5^2 + 1^2} = 1.118$

$\therefore$ Normalizing the eigenvector for $\lambda_2 = 1$ we get

$$\mathbf{u}_2 = \left(\frac{0.5}{1.118}, \frac{1}{1.118}\right)^\intercal = (0.4472, 0.8944)^\intercal$$

Therefore,

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\intercal}$$

where

$$\mathbf{A} = \begin{bmatrix} -4 & -7 \\ 1 & 4 \end{bmatrix}$$

$$\mathbf{\Sigma} = diag(\sqrt{\lambda_1}, \sqrt{\lambda_2}) = \begin{bmatrix} \sqrt{81} & 0 \\ 0 & \sqrt{1} \end{bmatrix} = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{U} = [\mathbf{u}_1 | \mathbf{u}_2] = \begin{bmatrix} -0.8944 & 0.4472 \\ 0.4472 & 0.8944 \end{bmatrix}$$

$$\mathbf{V} = [\mathbf{v}_1 | \mathbf{v}_2] = \begin{bmatrix} 0.4472 & -0.8944 \\ 0.8944 & 0.4472 \end{bmatrix}$$

Also, if any one of the **V** and **U** is known, then the other can be found, by the following equations

$$\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}^{\intercal} \cdot \mathbf{v}_i \quad and \quad \mathbf{v}_i = \frac{1}{\sigma_i}\mathbf{A} \cdot \mathbf{u}_i$$