# 9.5 Singular Value Decomposition

Extension of the Diagonalization theory for square matrices $n \times n$ to non-square matrices $m \times n$.

In previous sections, we have found that

I) If $A$ is Symm.

$$A = PDP^T \quad \text{(Eigenvalue Decomp.)}$$

$D$: diagonal and columns of $\underline{P}$ eigenvectors of $A$. orthogonal matrix.

II) If $A$ nonsymm.

$$A = PHP^T \quad \text{Hessenberg Decomp.}$$

$P$ orthog matrix and $H$ upper Hessenberg.

III) If eigenvalues of $A$ are real

$$A = PSP^T$$

$P$ orthog and $S$ upper triangular.

Importance of orthog matrix in numerical computation

If $\|\vec{x} - \hat{x}\| \leq e$

$\Rightarrow \|P(\vec{x} - \hat{x})\| = \|\vec{x} - \hat{x}\| \leq e$ errors are not magnified.

In this Section, we will find that for an arbitrary $A$ $m\times n$ matrix

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

$U, V$ orthog. and $\Sigma$ is like-diag. for nonsquare matrices.

Thm 9.5.1 $A$ $m\times n$ matrix

a) $A$ and $A^TA$ have same null space.

b) $A$ and $A^TA$ " " row space

c) $A^T$ and $A^TA$ " " Column space

d) $A$ and $A^TA$ " " rank

Proof:- $(\rightarrow)$ We want to prove
if $A\vec{x} = \vec{0} \Rightarrow (A^TA)\vec{x} = \vec{0}$

In fact, if $A\vec{x} = 0 \Rightarrow (A^TA)\vec{x} = A^T(A x) = A^T(\vec{0}) = \vec{0}$ ✓

$(\leftarrow)$ Want to prove.
if $A^TA\vec{x} = \vec{0} \Rightarrow A\vec{x} = \vec{0}$.

In fact, $A^TA\vec{x} = 0 \Rightarrow \vec{x} \in Null(A^TA) \overset{\text{Thm } 4.8.9(a)}{=} [row(A^TA)]^{\perp} \overset{\text{Symmetry of } A^TA}{=} [Col(A^TA)]^{\perp}$

Also, $A^TA\vec{x} \in Col(A^TA) \Rightarrow \vec{x} \cdot A^TA\vec{x} = \vec{0}$

or $0 = (A^TA\vec{x})^T\vec{x} = \vec{x}^T(A^TA)^T\vec{x} = \vec{x}^T A^TA\vec{x} = (\vec{x}^TA^T)(A\vec{x}) =$
$= (A\vec{x})^T A\vec{x} = A\vec{x} \cdot A\vec{x} = \|A\vec{x}\|^2$

Thus, $\|A\vec{x}\|^2 = 0 \Rightarrow A\vec{x} = \vec{0}$. ✓

Thm 9.5.2  $A$ $m \times n$ matrix, then

a) $A^T A$ is orthogonally diagonalizable

b) The eigenvalues of $A^T A$ are nonnegative

Proof.-

a) $A^T A$ is Symm since $(A^T A)^T = A^T (A^T)^T = A^T A$
then by theorem 7.2.1 $A^T A$ is orthog. diag.

b) First notice that

$$A^T A \vec{v} \cdot \vec{v} = \vec{v}^T A^T A \vec{v} = (A\vec{v})^T A\vec{v} = A\vec{v} \cdot A\vec{v} = \|A\vec{v}\|^2$$

Then if $\vec{v}$ is a unit eigenvector of $A^T A$ with eigenvalue $\lambda$

$$\|A\vec{v}\|^2 = (A^T A)\vec{v} \cdot \vec{v} = \lambda \vec{v} \cdot \vec{v} = \lambda \|v\|^2 = \lambda.$$

Set of eigenvectors of $A^T A$ is orthonormal

$$\Rightarrow \lambda \geq 0. \quad \checkmark$$

---

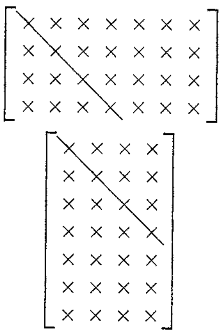
Def.- For $A$ $m \times n$ matrix
If $\lambda_1, \lambda_2, .., \lambda_n$ are the eigenvalues of $A^T A$, then
the real numbers

$$\sigma_1 = \sqrt{\lambda_1}, \quad \sigma_2 = \sqrt{\lambda_2}, \quad \dots, \quad \sigma_n = \sqrt{\lambda_n}$$

are called <u>Singular Values of A</u>.

### Singular Value Decomposition

Before turning to the main result in this section, we will find it useful to extend the notion of a "main diagonal" to matrices that are not square. We define the ***main diagonal*** of an $m \times n$ matrix to be the line of entries shown in Figure 9.5.1—it starts at the upper left corner and extends diagonally as far as it can go. We will refer to the entries on the main diagonal as the ***diagonal entries.***



Main diagonal

▲ **Figure 9.5.1**

We are now ready to consider the main result in this section, which is concerned with a specific way of factoring a general $m \times n$ matrix $A$. This factorization, called ***singular value decomposition*** (abbreviated SVD) will be given in two forms, a brief form that captures the main idea, and an expanded form that spells out the details. The proof is given at the end of this section.

**THEOREM 9.5.3 Singular Value Decomposition**

*If $A$ is an $m \times n$ matrix, then $A$ can be expressed in the form*

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

*where $U$ and $V$ are orthogonal matrices and $\Sigma$ is an $m \times n$ matrix whose diagonal entries are the singular values of $A$ and whose other entries are zero.*

**THEOREM 9.5.4 Singular Value Decomposition (Expanded Form)**

*If $A$ is an $m \times n$ matrix of rank $k$, then $A$ can be factored as*

$$A = U\Sigma V^T = [\mathbf{u}_1 \;\; \mathbf{u}_2 \;\; \cdots \;\; \mathbf{u}_k \mid \mathbf{u}_{k+1} \;\; \cdots \;\; \mathbf{u}_m] \left[\begin{array}{cccc|c} \sigma_1 & 0 & \cdots & 0 & \\ 0 & \sigma_2 & \cdots & 0 & 0_{k\times(n-k)} \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & \sigma_k & \\ \hline & & 0_{(m-k)\times k} & & 0_{(m-k)\times(n-k)} \end{array}\right] \left[\begin{array}{c} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_k^T \\ \hline \mathbf{v}_{k+1}^T \\ \vdots \\ \mathbf{v}_n^T \end{array}\right]$$

*in which $U$, $\Sigma$, and $V$ have sizes $m \times m$, $m \times n$, and $n \times n$, respectively, and in which*

(*a*) $V = [\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \cdots \;\; \mathbf{v}_n]$ *orthogonally diagonalizes* $A^TA$.

(*b*) *The nonzero diagonal entries of* $\Sigma$ *are* $\sigma_1 = \sqrt{\lambda_1}$, $\sigma_2 = \sqrt{\lambda_2}, \ldots, \sigma_k = \sqrt{\lambda_k}$, *where* $\lambda_1, \lambda_2, \ldots, \lambda_k$ *are the nonzero eigenvalues of* $A^TA$ *corresponding to the column vectors of* $V$.

(*c*) *The column vectors of* $V$ *are ordered so that* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > 0$.

(*d*) $\mathbf{u}_i = \dfrac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \dfrac{1}{\sigma_i}A\mathbf{v}_i \qquad (i = 1, 2, \ldots, k)$

(*e*) $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ *is an orthonormal basis for* $\text{col}(A)$.

(*f*) $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_m\}$ *is an extension of* $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ *to an orthonormal basis for* $R^m$.

**Harry Bateman (1882–1946)**

**Historical Note** The term *singular value* is apparently due to the British-born mathematician Harry Bateman, who used it in a research paper published in 1908. Bateman emigrated to the United States in 1910, teaching at Bryn Mawr College, Johns Hopkins University, and finally at the California Institute of Technology. Interestingly, he was awarded his Ph.D. in 1913 by Johns Hopkins at which point in time he was already an eminent mathematician with 60 publications to his name.

[*Image: Courtesy of the Archives, California Institute of Technology*]

The vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ are called the ***left singular vectors*** of $A$, and the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are called the ***right singular vectors*** of $A$.

Def.: For an $m \times n$ matrix $A$, we define the main diagonal to be the line of entries that starts at the upper left corner and extends diagonally as far as it can go. The entries on this main diagonal are called diagonal entries. (See Fig 9.5.1).

For the theorem look at page 509 book.

Proof.: Hypothesis are

① $A$ $m \times n$ matrix

② $\operatorname{rank}(A) = K$

The first thing to consider is that : $\boxed{A^T A = V D V^T}$

$A^T A$ is symm., then thm 7.2.1 $\nearrow$ $\quad (A^T)_{n \times m} A_{m \times n} = (A^T A)_{n \times n}$

where $V$ is orthogonal and the column vectors of $V$ are orthonormal eigenvectors of $A^T A$,

$$V = [\ \vec{v}_1\ \vec{v}_2 \dots \vec{v}_n\ ]$$

Corresponding to the eigenvalues $\lambda_1, \lambda_2, \dots \lambda_n$ that may be repeated and also some of them may be zero.

Since $\underline{\operatorname{rank}(A) = K}$, then (thm 9.5.1) $\underline{\operatorname{rank}(A^T A) = K}$

This also implies that $\underline{\operatorname{rank}(D) = K}$, because $D$ is similar to $A^T A$.

Then,

$$D = \begin{bmatrix} \lambda_1 & & & & & \\ & \lambda_2 & & & 0 & \\ & & \ddots & & & \\ & & & \lambda_k & & \\ & 0 & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} \quad \text{where } \lambda_j \neq 0,\ j = 1,..,k.$$

From the orthogonality of $\vec{v}_i$'s,
it can be shown that the set $\{A\vec{v}_1, .., A\vec{v}_n\}$ is a set of orthogonal vectors. In fact,

$$A\vec{v}_i \cdot A\vec{v}_j = \vec{v}_j^{\,T} A^T A \vec{v}_i = \left(\vec{v}_j^{\,T} A^T A\right) \vec{v}_i =$$

$$\vec{v}_i \cdot (A^T A)^T \vec{v}_j = \vec{v}_i \cdot A^T A \vec{v}_j = \vec{v}_i \cdot \lambda_j v_j = \lambda_j \underbrace{\vec{v}_i \cdot \vec{v}_j}_{\text{because } \vec{v}_i \cdot v_j = 0} = 0$$

We showed that

$$\|A\vec{v}_i\|^2 = \lambda_i, \quad i = 1, 2, .., k, .. n$$

$\Rightarrow$ If $\|A\vec{v}_i\| \neq 0, \quad i = 1, 2, .., k$

Then $S = \{A\vec{v}_1, .., A\vec{v}_k\}$ is an orthog. set of nonzero vectors

and because each $A\vec{v}_i \in Col(A) \Rightarrow S \subset Col(A)$

Also, $rank(A) = dim(col(A)) = k$, then

$S$ is a basis for $Col(A)$. An orthonormal basis for $Col(A)$ is given by

$$\{\vec{u}_1, .., \vec{u}_k\} = \left\{ \frac{A\vec{v}_1}{\|A\vec{v}_1\|}, \frac{A\vec{v}_2}{\|A\vec{v}_2\|}, \dots, \frac{A\vec{v}_k}{\|A\vec{v}_k\|} \right\}$$

$$\frac{1}{\sqrt{\lambda_1}} A\vec{v}_1, \frac{1}{\sqrt{\lambda_2}} A\vec{v}_2, \dots, \frac{1}{\sqrt{\lambda_k}} A\vec{v}_k$$

or $\boxed{A\vec{v}_1 = \sqrt{\lambda_1}\,\vec{u}_1 = \sigma_1 \vec{u}_1, \ldots, A\vec{v}_k = \sigma_k \vec{u}_k}$ (*)

Since $A_{m\times n}$ $(Av_i)_{m\times 1} \in \mathbb{R}^m$, $i = 1, 2, .., k$

Thus, the set $\{\vec{u}_1, \vec{u}_2, .., \vec{u}_k\} \subset \operatorname{Col}(A)$ is lin. indep in $\mathbb{R}^m$

and can be extended to an orthonormal basis

$\{\vec{u}_1, \vec{u}_2, .., \vec{u}_k, \underbrace{\vec{u}_{k+1}, .., \vec{u}_m}\}$ → Not of the form $Av_i$, $i = k+1, .., m$ they need to be obtained by algebraic methods.

Define $U$ as the orthogonal matrix

$$U = \left[ \vec{u}_1 \; \vec{u}_2 \; \cdot \; u_k \; \cdot \; \vec{u}_m \right]_{m\times m}$$

and let $\Sigma$ be the $m\times n$ matrix

$$\left[\begin{array}{c|c} \begin{matrix} \sigma_1 & & 0 \\ & \sigma_2 \ddots & \\ 0 & & \sigma_k \end{matrix} & 0_{k\times(n-k)} \\ \hline 0_{(m-k)\times k} & 0_{(m-k)\times(n-k)} \end{array}\right]_{m\times n}$$

then

$$\underset{m\times m}{U} \underset{m\times n}{\Sigma} = \left[ \sigma_1\vec{u}_1 \; \sigma_2\vec{u}_2 \; \cdots \; \sigma_k\vec{u}_k \; 0 \; 0 \cdots 0 \right]_{m\times n} =$$

$$= \left[ A\vec{v}_1 \; A\vec{v}_2 \cdots A\vec{v}_k \; A\vec{v}_{k+1} \cdots A\vec{v}_n \right]_{m\times n} = AV$$

$$\Rightarrow \quad \boxed{A = U \Sigma V^{-1} = U \Sigma V^{T}.}$$

Sect 9.5
Problem # 9) Find S.V.D. of

$$A = \begin{bmatrix} -2 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix}$$

1) Find $A^TA$

$$A^TA = \begin{bmatrix} -2 & -1 & 2 \\ 2 & 1 & -2 \end{bmatrix} \begin{bmatrix} -2 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$$ obviously Symm.

2) Orthogonally Diagonalizes $A^TA$

$\boxed{\text{Eeigenvalues}}$: $|\lambda I - A^TA| = \begin{vmatrix} \lambda-9 & 9 \\ 9 & \lambda-9 \end{vmatrix} = 0 \Leftrightarrow (\lambda-9)^2 - 81 =$

$\lambda^2 - 18\lambda = 0$

or

$\lambda(\lambda - 18) = 0$

$\Rightarrow \quad \lambda_1 = 18, \quad \lambda_2 = 0.$

$\boxed{\text{Eigenvectors for } \lambda_1 = 18}$

$$\begin{array}{c} 18-9 \\ \begin{bmatrix} 9 & 9 \\ 9 & 9 \end{bmatrix} \\ 18-9 \end{array} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \quad x_1 = -x_2$$

or $\vec{x} = \begin{pmatrix} -x_2 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

So $\vec{v}_1' = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ eigenvector associated to $\lambda_1 = 18$.

For $\boxed{\begin{array}{c}\text{Eigenvectors}\\ \lambda_2 = 0\end{array}}$,

$$\begin{bmatrix} -9 & 9 \\ 9 & -9 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{l} x_1 = x_2 \\ \vec{x} = \begin{pmatrix} x_2 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{array}$$

then $\vec{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ eigenvector corresponding to $\lambda_2 = 0$.

Therefore, if we define

$$P = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$ then $A^TA$ can be orthogonally diagonalized as

$$\overset{P}{\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}} \overset{A^TA}{\begin{bmatrix} -9 & 9 \\ 9 & -9 \end{bmatrix}} \overset{P^T}{\begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}} = \overset{D}{\begin{bmatrix} 18 & 0 \\ 0 & 0 \end{bmatrix}}$$

$$\underline{P\,(A^TA)\,P^T = D}$$

3) <u>Construct U and V of the Sing. val. decomp. theorem</u>

According to <u>thm 9.5.4.</u>

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

and $$\Sigma = \begin{bmatrix} \overset{=\sqrt{\lambda_1}}{\sigma_1} & 0 \\ 0 & \overset{=\sqrt{\lambda_2}}{\sigma_2} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

To construct the matrix $U$, we need to consider $A\vec{v}_i$ for $\vec{v}_i$ such that $A\vec{v}_i \neq \vec{0}$.

thus,
$$A\vec{v}_1 = \begin{bmatrix} -2 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 4/\sqrt{2} \\ 2/\sqrt{2} \\ -4/\sqrt{2} \end{bmatrix}$$

also,
$$\|A\vec{v}_1\| = \sqrt{\frac{16}{2} + \frac{4}{2} + \frac{16}{2}} = \frac{\sqrt{36}}{\sqrt{2}} = \frac{6}{\sqrt{2}} = \frac{6\sqrt{2}}{2} = 3\sqrt{2}$$

$$\Rightarrow \quad \frac{A\vec{v}_1}{\|A\vec{v}_1\|} = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}. \quad \text{According to thm 9.5.4} \quad \vec{u}_1 = \begin{pmatrix} 2/3 \\ 1/3 \\ -2/3 \end{pmatrix}$$

Since $A\vec{v}_2 = 0\vec{v}_2 = \vec{0}$, there is not a $\vec{u}_2$ obtained this way.

4) To complete the construction of an orthogonal matrix $U$, we need to extend the set $S = \{\hat{u}_1\}$ to an orthonormal basis of $\mathbb{R}^{3=m}$. (Notice that this is not unique!)

This is not hard, just consider

$$\vec{u}_1 \cdot \vec{u}_2' = 0 \quad \Leftrightarrow \quad \left(\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right) \cdot (u_{21}, u_{22}, u_{23}) = 0$$

$$\text{or} \quad \frac{2}{3}u_{21} + \frac{1}{3}u_{22} - \frac{2}{3}u_{23} = 0 \Rightarrow u_{21} = u_{23} - \frac{1}{2}u_{22}$$

$$\text{or} \quad \vec{u}_2' = \begin{pmatrix} u_{23} - \frac{1}{2}u_{22} \\ u_{22} \\ u_{23} \end{pmatrix} = \begin{pmatrix} u_{23} \\ 0 \\ u_{23} \end{pmatrix} + \begin{pmatrix} -\frac{1}{2}u_{22} \\ u_{22} \\ 0 \end{pmatrix} = u_{23}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + u_{22}\begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix}$$

One possibility is to choose $u_{23} = 1$, $u_{22} = 0$

then $\vec{u}_2' = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, obviously not a unit vector

thus, $\vec{u}_2 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}$

That way we have an orthonormal set $S' = \left\{ \underset{\vec{u}_1}{\begin{pmatrix} 2/3 \\ 1/3 \\ -2/3 \end{pmatrix}}, \underset{\vec{u}_2}{\begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}} \right\}$

Our purpose is to extend this to a basis in $\mathbb{R}^3$ which is orthonormal. So, we look for a 3rd vector $\vec{u}_3'$ such that

$$\begin{aligned} \vec{u}_1 \perp \vec{u}_3' &\Longrightarrow \\ \vec{u}_2 \perp \vec{u}_3' &\Longrightarrow \end{aligned} \begin{cases} \frac{2}{3} u_{31} + \frac{1}{3} u_{32} - \frac{2}{3} u_{33} = 0 \\ \frac{1}{\sqrt{2}} u_{31} + 0\, u_{32} + \frac{1}{\sqrt{2}} u_{33} = 0 \end{cases}$$

$$\Rightarrow \boxed{u_{31} = -u_{33}}$$

$$\Rightarrow \quad 4u_{33} + u_{32} = 0 \Rightarrow \boxed{u_{32} = -4u_{33}}$$

$$\Rightarrow \quad \vec{u}_3' = \begin{pmatrix} -u_{33} \\ -4u_{33} \\ u_{33} \end{pmatrix} = u_{33} \begin{pmatrix} -1 \\ -4 \\ 1 \end{pmatrix} \Rightarrow \vec{u}_3 = \begin{pmatrix} -\frac{1}{\sqrt{18}} \\ \frac{-4}{\sqrt{18}} \\ \frac{1}{\sqrt{18}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{3\sqrt{2}} \\ \frac{-4}{3\sqrt{2}} \\ \frac{1}{3\sqrt{2}} \end{pmatrix}$$

$$\text{or } \vec{u}_3 = \begin{pmatrix} \frac{-\sqrt{2}}{6} \\ -\frac{2\sqrt{2}}{3} \\ \sqrt{2}/6 \end{pmatrix}$$

Finally, the orthogonal matrix $U$ sought is formed by orthonormal vectors

$$U = [\,\underset{\downarrow}{\vec{u}_1}\ \ \underset{\downarrow}{\vec{u}_2}\ \ \underset{\downarrow}{\vec{u}_3}\,],$$

where $\hat{u}_1 = \dfrac{A\vec{v}_1}{\|A\vec{v}_1\|}$

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

where $$U = \begin{bmatrix} \frac{2}{3} & \frac{1}{\sqrt{2}} & -\frac{\sqrt{2}}{6} \\ \frac{1}{3} & 0 & -\frac{2\sqrt{2}}{3} \\ -\frac{2}{3} & \frac{1}{\sqrt{2}} & \sqrt{2}/6 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and $$V = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Notice that the matrix $U$ is not unique. There are infinitely many ways to complete $S = \{\vec{u}_1\}$ as an orthonormal basis for $\mathbb{R}^3$.

<u>The Singular Value Decomposition reduces to the orthogonal diagonalization of a matrix $A$ if $A$ is symmetric.</u>

In fact, for any matrix $A$, the Singular Value Decomposition is given by

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

where $\Sigma$ has nonzero entries only along its main diagonal, which are the singular values of $A$, or

$$\sigma_i = \sqrt{\lambda_i}, \quad \text{where } A^TA\vec{v}_i = \lambda_i \vec{v}_i, \quad i=1,2,\dots,k,\dots,n \quad \lambda_i \neq 0,\ i=1,\dots,k$$

Also, $V = [\vec{v}_1\ \vec{v}_2 \dots \vec{v}_n]$ are the orthonormal eigenvectors of $A^TA$ corresp. to $\lambda_i$, $i=1,2,\dots,n$

and $U = [\vec{u}_1\ \vec{u}_2 \dots \vec{u}_m] = \left[\frac{A\vec{v}_1}{\|A\vec{v}_1\|}\ \frac{A\vec{v}_2}{\|A\vec{v}_2\|} \dots \frac{A\vec{v}_k}{\|A\vec{v}_k\|} \dots u_{k+1} \dots u_m\right]$

Now, if $A_{n\times n}$ is Symm. $A = A^T$

Sing. values of $A$

Then, $A^TA\vec{v}_i = \lambda_i \vec{v}_i \Longrightarrow A^2\vec{v}_i = \lambda_i \vec{v}_i \Longrightarrow \sqrt{\lambda_i}$ is an eigenvalue for $A$ with the same correspond. eigenvector $\vec{v}_i$

It means $\sigma_i = \sqrt{\lambda_i}$ is an eigenvalue for $A$.

$$\Rightarrow \boxed{A\vec{v}_i = \sqrt{\lambda_i}\,\vec{v}_i} \Rightarrow \|A\vec{v}_i\| = \sqrt{\lambda_i}\,\overbrace{\|\vec{v}_i\|}^{=1} = \sqrt{\lambda_i} = \sigma_i, \quad i=1,2,\dots,k$$

$$\Rightarrow U = \left[\frac{\cancel{\sqrt{\lambda_1}}\,\vec{v}_1}{\cancel{\sqrt{\lambda_1}}}\ \frac{\cancel{\sqrt{\lambda_2}}\,\vec{v}_2}{\cancel{\sqrt{\lambda_2}}} \dots \frac{\cancel{\sqrt{\lambda_k}}\,\vec{v}_k}{\cancel{\sqrt{\lambda_k}}}\ u_{k+1} \dots u_m\right]$$

$$\text{or } U = [\vec{v}_1\ \vec{v}_2 \dots \vec{v}_k\ \hat{u}_{k+1} \dots \hat{u}_n]$$

A possible orthonormal $V$ is given by

$$V = \left[\hat{v}_1 \ldots \hat{v}_k \underbrace{\vec{v}_{k+1} \ldots \vec{v}_n}\right]$$

↳ eigenvectors corresp. to $\lambda = 0$. eigenvalue.

therefore, the matrix $A$ can be decomposed as

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

where

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & & & 0 \\ & \sigma_2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_k & & & \\ & & & & 0 & & \\ & 0 & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{\lambda_1} & & & & & & 0 \\ & \sqrt{\lambda_2} & & & & & \\ & & \ddots & & & & \\ & & & \sqrt{\lambda_k} & & & \\ & & & & 0 & & \\ & 0 & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}$$

$$= D$$

which is the orthogonal diagonalization of $A$.