Singular Value Decomposition.

The goal is to obtain a factorization of an arbitrary matrix $A_{m\times n}$ similar to those obtained for diagonalizable square matrices

It means to express

$$\boxed{A_{m\times n} = U_{m\times m} \Sigma_{m\times n} V^T_{n\times n}} \qquad (1.1)$$

where U and V are orthogonal matrices

and $\Sigma_{m\times n}$ is a quasi-diagonal matrix.

For example,

$$\Sigma_{3\times 2} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} \text{ or } \Sigma_{2\times 3} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix}$$

$$\Sigma_{5\times 3} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Sigma_{3\times 5} = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 & 0 \end{bmatrix}$$

To define U, Σ and V, we consider the matrix A^TA.

<u>Theorem.</u> For $A_{m\times n}$

i) A^TA is a square matrix: $(A^TA)_{n\times n}$.

ii) A^TA is symmetric.

iii) Then, A^TA is orthogonal diagonalizable.

iv) The eigenvalues λ_i of A^TA are non-negative, i.e., $\lambda_i \geq 0$, $i=1,..,n$.

<u>Proof.</u>

If $A_{m\times n} \Rightarrow A^T_{n\times m}$

i) and $(A^T)_{n\times m} A_{m\times n} = (A^TA)_{n\times n}$

ii) $(A^TA)^T = A^T(A^T)^T = A^TA$.

iii) Since A^TA is Symmetric then
A^TA is orthogonally diagonalizable
In fact,

$$\boxed{A^TA = PDP^T}$$

Where $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$ and $P = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$

and $\lambda_i \in \mathbb{R}$ are the eigenvalues of A^TA.
They may be repeated but they are real.
Also, $V_1, .., \vec{V}_n$ are the corresponding eigenvectors of $\lambda_1, \lambda_2 .., \lambda_n$. They are orthogonal and unit vectors.
It means $\vec{V}_i \cdot \vec{V}_j = 0$ if $i \neq j$
and $\vec{V}_i \cdot \vec{V}_i = \|\vec{V}_i\|^2 = 1$.

iv) For all $i = 1, .., n$

$$\lambda_i = \|A\vec{v}_i\|^2 \geq 0.$$

In fact,
$$\|A\vec{v}_i\|^2 = A\vec{v}_i \cdot A\vec{v}_i = (A\vec{v}_i)^T A\vec{v}_i = \vec{V}_i^T A^T A \vec{V}_i$$
$$= \vec{V}_i^T (A^TA\vec{v}_i) = \vec{V}_i^T \lambda_i \vec{V}_i$$
$$= \lambda_i \vec{V}_i \cdot \vec{V}_i = \lambda_i \|\vec{V}_i\|^2 = \lambda_i$$

Also,
$$\|A\vec{v}_i\| = \sqrt{\lambda_i}, \quad i = 1, .., n.$$

Now, since A^TA is symmetric

$$\boxed{A^TA = PDP^T} \qquad (4.1)$$

$$P = \underbrace{\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix}}_{\text{orthonormal Eigenvectors}} \quad \text{and} \quad D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

Corresponding λ's eigenvalues

Also, from (1.1)

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

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

$$= V\Sigma^T \underbrace{(U^TU)}_{=I} \Sigma V^T$$

or

$$\boxed{A^TA = V(\Sigma^T\Sigma)V^T} \qquad (4.2)$$

From (4.1) and (4.2)

$$\boxed{PDP^T = A^TA = V(\Sigma^T\Sigma)V^T} \qquad (*)$$

So, ~~it~~ will be natural to define

$$\boxed{V = P} \quad \text{and} \quad \boxed{\Sigma^T\Sigma = D} \qquad (4.3)$$

As an example, consider

$$A_{5\times 3} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ \vdots & \vdots & \vdots \\ a_{51} & a_{52} & a_{53} \end{bmatrix}$$

Then $(A^T A)_{3\times 3}$

with eigenvalues: $\lambda_1, \lambda_2, \lambda_3$ (maybe repeated)

and $\vec{V}_1, \vec{V}_2, \vec{V}_3$ are a set of orthonormal eigenvectors corresponding to $\lambda_1, \lambda_2, \lambda_3$ respectivamente.

and

$$A^T A = PDP^T = \begin{bmatrix} \vec{V}_1 & \vec{V}_2 & \vec{V}_3 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} \vec{V}_1^T \rightarrow \\ \vec{V}_2^T \rightarrow \\ \vec{V}_3^T \rightarrow \end{bmatrix}$$

Then, from (4.3), it will be natural to define

$$\boxed{V = [\underbrace{\vec{V}_1 \ \vec{V}_2 \ \vec{V}_3}_{\text{orthonormal eigenvectors of } A^TA.}]} \quad \text{and} \qquad (5.1)$$

$$\Sigma^T_{3\times 5} \Sigma_{5\times 3} = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix} = D = \underbrace{\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}}_{\lambda\text{'s eigenvalues of } A^TA.}$$

Therefore, it will be natural to define:

$$\sigma_i^2 = \lambda_i, \quad i = 1, 2, 3$$ $\longrightarrow$ Eigenvalues of A^TA.

or

$$\boxed{\sigma_i = \sqrt{\lambda_i},} \quad i = 1, 2, 3 \qquad (6.1)$$

The scalars σ_i are called <u>Singular values of A</u>.

So far, we have constructed Σ and V.

In fact,

$$\boxed{A_{5\times3} = U_{5\times5}\ \Sigma_{5\times3}\ V^T_{3\times3}}$$

$$= \underbrace{\begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_5 \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}_{5\times5}}_{?} \begin{bmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \sqrt{\lambda_2} & 0 \\ 0 & 0 & \sqrt{\lambda_3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \vec{V}_1 & \vec{V}_2 & \vec{V}_3 \end{bmatrix}$$

The column vectors $\vec{u}_i$ are defined as: (6.2)

$$\boxed{\vec{u}_i = \frac{A\vec{v}_i}{\sigma_i}, \quad i = 1, \dots, 3}$$

and $\vec{u}_4, \vec{u}_5$ are arbitrary vectors which form an orthonormal set with $\vec{u}_1, \vec{u}_2$ and $\vec{u}_3$.

If $\sigma_i = 0$, for any $i = 1 \to 3$.
then u_i is also an arbitrary orthog. vector to others and $\|v_i\| = 1$.

Therefore, if $\sigma_1, \sigma_2, \sigma_3 \neq 0$

$$\vec{u}_1 = \frac{A\vec{v}_1}{\sigma_1}, \quad \vec{u}_2 = \frac{A\vec{v}_2}{\sigma_2}, \quad \vec{u}_3 = \frac{A\vec{v}_3}{\sigma_3}$$

It will be shown that

$$\|A\vec{v}_i\| = \sigma_i, \quad \text{then} \quad \|\vec{u}_i\| = 1.$$

and also that $A\vec{v}_i \cdot A\vec{v}_j = 0$, for $i \neq j$.

Thus, the orthogonal matrix U for the SVD can be defined as

$$\boxed{U = \left[\frac{A\vec{v}_1}{\sigma_1} \;\; \frac{A\vec{v}_2}{\sigma_2} \;\; \frac{A\vec{v}_3}{\sigma_3} \;\; \vec{u}_4 \;\; \vec{u}_5 \right]_{5\times 5}} \tag{7.1}$$

Finally, the SVD of $A_{5\times 3}$ is given by

$$A_{5\times 3} = \left[\frac{A\vec{v}_1}{\sigma_1} \;\; \frac{A\vec{v}_2}{\sigma_2} \;\; \frac{A\vec{v}_3}{\sigma_3} \;\; \vec{u}_4 \;\; \vec{u}_5 \right]_{5\times 5} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{5\times 3} \left[\vec{v}_1 \;\; \vec{v}_2 \;\; \vec{v}_3 \right]_{3\times 3}$$

Remark: $\vec{u}_4$ and $\vec{u}_5$ are orthogonal vectors to $\frac{A\vec{v}_1}{\sigma_1}, \frac{A\vec{v}_2}{\sigma_2}, \frac{A\vec{v}_3}{\sigma_3}$. They are also unit vectors.

They can be obtained by finding a basis for the null space of matrix

$$\begin{bmatrix} (A\vec{v}_1/\sigma_1)^T \longrightarrow \\ (A\vec{v}_2/\sigma_2)^T \longrightarrow \\ (A\vec{v}_3/\sigma_3)^T \longrightarrow \end{bmatrix}.$$

Example

For the matrix $A = \begin{bmatrix} -2 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix}_{3\times 2}$

Find its Singular values and corresponding Singular vectors.

Answ: By definition the singular values of A are the square roots of the eigenvalues of A^TA, and right singular vectors are any unit eigenvector $\vec{v}_i$ of A^TA corresponding to an eigenvalue λ_i.

Now,

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

Obviously Symmetric as expected why?

Eigenvalues of A^TA:

$$|A-\lambda I| = 0 \iff \begin{vmatrix} 9-\lambda & -9 \\ -9 & 9-\lambda \end{vmatrix} = (\lambda-9)^2 - 81$$

$$= \lambda^2 - 18\lambda = \lambda(\lambda - 18) \implies \boxed{\lambda_1 = 18, \ \lambda_2 = 0}$$

$\implies$ Singular Values of A: $\sigma_1 = 3\sqrt{2}, \quad \sigma_2 = 0$

Eigenvectors of A^TA:

$\lambda_1 = 18$

$$(A^TA - 18I)\vec{x} = \vec{0} \iff \begin{bmatrix} 9-18 & -9 \\ -9 & 9-18 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$-9x_1 - 9x_2 = 0 \Rightarrow x_1 = -x_2.$$

or $\vec{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \Rightarrow \vec{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ is a unit eigenvector of A^TA and a right singular vector of A corresponding to $\sigma_1 = \sqrt{18} = 3\sqrt{2}$.

Similarly, for

$\lambda_2 = 0$

$$(A^TA - 0I)\vec{x} = 0 \iff \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or $9x_1 = 9x_2 \Rightarrow x_1 = x_2 \Rightarrow \vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

or $\vec{V}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ right singular vector of A for $\sigma_2 = 0$.

Therefore,

Eigenvalues	Corresponding Eigenvectors
$\lambda_1 = 18$	$\vec{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$
$\lambda_2 = 0$	$\vec{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

Then the SVD so far is given by

$$A = \underbrace{[\vec{u}_1\ \vec{u}_2\ \vec{u}_3]_{3\times3}}_{\text{unknowns}} \begin{bmatrix} \overset{=\sigma_1=\sqrt{18}}{3\sqrt{2}} & 0 \\ 0 & 0^{=\sigma_2} \\ 0 & 0 \end{bmatrix}_{3\times2} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}_{2\times2}$$

To determine u_i $(i=1,2,3)$, we will use the definition

$$u_i = \frac{A\vec{v}_i}{\sigma_i} \quad \text{if } \sigma_i \neq 0.$$

Now, $\sigma_1 = \sqrt{18} = 3\sqrt{2}$ is the only singular value nonzero.

then,

$$U = [\vec{u}_1\ \vec{u}_2\ \vec{u}_3] = \left[\frac{A\vec{v}_1}{\sigma_1}\ \ \underset{?}{\vec{u}_2}\ \ \underset{?}{\vec{u}_3} \right]$$

and

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

and

$$u_1 = \frac{A\vec{v}_1}{\sigma_1} = \frac{1}{3\sqrt{2}} \begin{bmatrix} 4/\sqrt{2} \\ 2/\sqrt{2} \\ -4/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$$

An updated SVD of A

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

is given by

$$\underset{3\times 2}{A} = \begin{array}{c} \begin{array}{ccc} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{array} \\ \begin{bmatrix} 2/3 & & \\ 1/3 & ? & ? \\ -2/3 & & \end{bmatrix}_{3\times 3} \end{array} \begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}_{3\times 2} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}_{2\times 2}$$

$$U \qquad \Sigma \qquad V^T$$

We still need $\vec{u}_2$ and $\vec{u}_3$ unit vectors such that

$$\vec{u}_2 \perp \vec{u}_1, \quad \vec{u}_3 \perp \vec{u}_1 \text{ and } \vec{u}_3 \perp \vec{u}_2$$

$$\vec{u}_2 \cdot \vec{u}_1 = 0 \qquad \vec{u}_3 \cdot \vec{u}_1 = 0 \qquad \vec{u}_3 \cdot \vec{u}_2 = 0$$

Any vector $\vec{w}$ orthogonal to $\vec{u}_1$ should satisfy

$$\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

$$\vec{w} \cdot \vec{u}_1 = 0 \iff \frac{2}{3} w_1 + \frac{1}{3} w_2 - \frac{2}{3} w_3 = 0 \iff$$

$$\vec{w}_1 = w_3 - \frac{1}{2} w_2 \quad \text{or} \quad \vec{w} = \begin{bmatrix} w_3 - \frac{1}{2} w_2 \\ w_2 \\ w_3 \end{bmatrix} =$$

$$= w_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + w_2 \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}.$$

Then $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} \right\}$ is a basis in $\mathbb{R}^3$ for

vectors $\vec{w}$ orthogonal to $\vec{u}_1$.

Since they are not orthogonal between them, we apply G-S.

to obtain an orthogonal basis.

We start defining,

$$\vec{u}_2' = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \Rightarrow \vec{u}_3' = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} \cdot \vec{u}_2'}{\vec{u}_2' \cdot \vec{u}_2'} \vec{u}_2' =$$

$$= \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \frac{\left(-\frac{1}{2}\right)}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1/4 \\ 0 \\ 1/4 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix} = \vec{u}_3'$$

$$\Rightarrow u_2 = \frac{\vec{u}_2'}{\|\vec{u}_2'\|} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \vec{u}_3 = \frac{\vec{u}_3'}{\|\vec{u}_3'\|} = \frac{4}{3\sqrt{2}} \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix} = \begin{bmatrix} -\sqrt{2}/6 \\ \frac{2\sqrt{2}}{3} \\ \sqrt{2}/6 \end{bmatrix}$$

Since, $\|\vec{u}_3'\| = \sqrt{\frac{1}{16} + 1 + \frac{1}{16}} = \sqrt{\frac{18}{16}} = \frac{3\sqrt{2}}{4}$

and finally.

$$A = \underset{3\times 3}{\begin{array}{ccc} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ \end{array}} \quad \underset{3\times 2}{} \quad \underset{2\times 2}{\begin{array}{cc} \vec{v}_1 & \vec{v}_2 \end{array}}$$

$$A = \underbrace{\begin{bmatrix} 2/3 & 1/\sqrt{2} & -\sqrt{2}/6 \\ 1/3 & 0 & 2\sqrt{2}/3 \\ -2/3 & 1/\sqrt{2} & \sqrt{2}/6 \end{bmatrix}_{3\times 3}}_{U} \underbrace{\begin{bmatrix} \overset{=\sigma_1 = 3\sqrt{2}}{\sqrt{18}} & 0 \\ 0 & \overset{=\sigma_2}{0} \\ 0 & 0 \end{bmatrix}_{3\times 2}}_{\Sigma} \underbrace{\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}_{2\times 2}}_{V^T}$$

Verify that A is actually equal to

$$\underline{U \Sigma V^T.}$$

In fact,

$$\Sigma V^T = \begin{bmatrix} -\frac{3\cancel{\sqrt{2}}}{\cancel{\sqrt{2}}} & \frac{3\cancel{\sqrt{2}}}{\cancel{\sqrt{2}}} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$U\Sigma V^T = U(\Sigma V^T) = \begin{bmatrix} 2/3 & 1/\sqrt{2} & -\sqrt{2}/6 \\ 1/3 & 0 & 2\sqrt{2}/3 \\ -2/3 & 1/\sqrt{2} & \sqrt{2}/6 \end{bmatrix} \begin{bmatrix} -3 & 3 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ -1 & 1 \\ 2 & -2 \end{bmatrix} \checkmark = A$$

Two important properties to prove to define SVD.

① $\|A\vec{v}_i\|^2 = \lambda_i \overbrace{\|\vec{v}_i\|}^{=1} = \lambda_i$, for $\vec{v}_i$ eigenvector of A^TA with eigenvalue λ.

defining $\sigma_i = \sqrt{\lambda_i}$ Singular Values

leads to $\boxed{\|A\vec{v}_i\| = \sigma_i \overbrace{\|\vec{v}_i\|}^{=1} = \sigma_i.}$

② If v_i and v_j are two Columns of V

then, $\boxed{\vec{v}_i \cdot \vec{v}_j = 0, \text{ if } i \neq j}$

Because, the matrix V is defined as

$$V = [\vec{v}_1\ \vec{v}_2 \ldots \vec{v}_n], \text{ where}$$

the set $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ forms an orthonormal basis of eigenvectors for A^TA. This basis always exists because A^TA is symmetric.

But also, $\vec{v}_i \cdot \vec{v}_j = 0 \Rightarrow \boxed{A\vec{v}_i \cdot A\vec{v}_j = 0, \quad i \neq j.}$

Proof (1)

$$A\vec{v}_i \cdot A\vec{v}_i = (A\vec{v}_i)^T A\vec{v}_i = \vec{v}_i^T A^T A \vec{v}_i$$
$$= \vec{v}_i^T \lambda_i v_i = \lambda_i \vec{v}_i^T \vec{v}_i = \lambda_i (\vec{v}_i \cdot \vec{v}_i)$$
$$= \lambda_i \|v_i\|^2 = \lambda_i (1) = \lambda_i, \text{ for all } i = 1, .., n.$$

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

or $\|A\vec{v}_i\| = \sqrt{\lambda_i} \overset{def}{=} \sigma_i, \quad i = 1, .., n$ $\longrightarrow$ Singular Values.

Proof (2)

$$A\vec{v}_i \cdot A\vec{v}_j = \vec{v}_i^T A^T A \vec{v}_j = \vec{v}_i^T \lambda_j \vec{v}_j$$
$$= \lambda_j (\vec{v}_i^T \vec{v}_j) = \lambda_j \vec{v}_i \cdot \vec{v}_j$$
$$= \lambda_j (0) = 0$$

General Theory of SVD. Determination of U.

The vectors $A\vec{v}_i$ $(i=1,..,r)$ for i such that $\lambda_i \neq 0$.

play an important role in the SVD of A.

They are used to define the matrix U of the SVD of A

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

They also have very nice properties as next thm shows.

Theorem:

Hypothesis: i) If $\{\vec{V}_1, .., \vec{V}_n\}$ orthonormal eigenvectors of A^TA (Then a basis for $\mathbb{R}^n$)

ii) If Corresponding eigenvalues of A^TA satisfy

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \overset{\neq 0}{\lambda_r} \geq \overset{=0}{\lambda_{r+1}} \geq \dots \geq \overset{=0}{\lambda_n}$$

Then

$\{A\vec{V}_1, .., A\vec{V}_r\}$ is an orthogonal basis for Col A

which implies rank $A = r$.

Proof.- Recall that $\|A\vec{v}_i\| = \sqrt{\lambda_i}$,
then if $\lambda_i \neq 0 \Rightarrow A\vec{v}_i \neq 0$, for $i = 1,..,r$.

Also
$$A\vec{v}_j \cdot A\vec{v}_i = \vec{v}_j^T A^T A \vec{v}_i = \vec{v}_j^T (A^TA\vec{v}_i) =$$
$$= \vec{v}_j^T \lambda_i \vec{v}_i = \lambda_i \vec{v}_j^T v_i = \lambda_i \vec{v}_j \cdot \vec{v}_i$$
$$= \begin{cases} \lambda_i \|\vec{v}_i\|^2 = \lambda_i \cdot 1 = \lambda_i, & i=j \\ \lambda_i (0) = 0, & i \neq j \end{cases}$$

Therefore, $S = \{A\vec{v}_1, .., A\vec{v}_r\}$ is an <u>orthogonal set of Col A</u>
each $A\vec{v}_i$ is in Col(A).

Since none of this vectors is zero, ($\|A\vec{v}_i\| = \sqrt{\lambda_i} \neq 0$) this set is also linearly independent.

It can be shown that S also spans Col A.

In fact,

For any $\vec{y} \in \text{Col} A \subset \mathbb{R}^m$, there is an $x \in \mathbb{R}^n$ such that $\vec{y} = A\vec{x}$

and $\vec{x} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$, since $\{\vec{v}_1, .., \vec{v}_n\}$ basis of $\mathbb{R}^n$.

$$\Rightarrow \vec{y} = A\vec{x} = c_1 A\vec{v}_1 + \cdots + c_r A\vec{v}_r + \underbrace{\overset{=0}{\cdots} + c_n \overset{=0}{A\vec{v}_n}}_{\text{Since } \|A\vec{v}_j\| = \lambda_j^{1/2} = 0,\ j = r+1, .., n}$$

$\Rightarrow \vec{y} \in \text{spans} \{A\vec{v}_1, .., A\vec{v}_r\}$

$\Rightarrow \quad S = \{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an orthogonal basis for Col A

$\Rightarrow \quad \text{rank } A = r.$

Determining U orthogonal matrix in the decomposition $\underset{m\times n}{A} = \underset{m\times m}{U} \underset{m\times n}{\Sigma} \underset{n\times n}{V^T}$

Multiplying by V both right sides of the decomposition

$$AV = U\Sigma.$$

or

For $r<n$, $r<m$

$$\left[A\hat{v}_1 \; A\hat{v}_2 \ldots A\hat{v}_r \; 0 \ldots 0 \right]_{m\times n} = \left[\vec{u}_1 \; \vec{u}_2 \cdots \vec{u}_r \cdots \vec{u}_m \right]_{m\times m} \begin{bmatrix} \sqrt{\lambda_1} & & & & & \\ & \sqrt{\lambda_2} & & & 0 & \\ & & \ddots & & & \\ & & & \sqrt{\lambda_r} & & \\ & 0 & & & 0 & \\ & & & & & \ddots & 0 \end{bmatrix}_{m\times n}$$

$$\Rightarrow \quad A\hat{v}_1 = \sqrt{\lambda_1}\,\vec{u}_1, \ldots, \quad A\hat{v}_r = \sqrt{\lambda_r}\,\vec{u}_r, \quad \lambda_i \neq 0, \; i=1,\ldots,r$$

So an appropriate definition for the entries of U is

$$\boxed{u_i = \frac{1}{\sqrt{\lambda_i}} A\vec{v}_i = \frac{1}{\sigma_i} A\hat{v}_i, \quad i=1,\ldots,r} \qquad (14.1)$$

We have already seen that $\|A\hat{v}_i\| = \sqrt{\lambda_i} = \sigma_i$, $i=1,\ldots,r$.

Therefore, $\|\hat{u}_i\| = \dfrac{\|Av_i\|}{\sigma_i} = 1$.

And we have also proved that

$\{A\hat{v}_1, \ldots, A\hat{v}_r\}$ is an orthonormal basis of $col(A)$.

For the remainder U entries $\vec{u}_{r+1}, \ldots, \vec{u}_m$

It is appropriate to define them as orthonormal vectors which are also orthogonal to $\vec{u}_1, \ldots, \vec{u}_r$.

As a consequence, an appropriate definition of U is

$$\boxed{U = \begin{bmatrix} \underbrace{\frac{A\vec{v}_1}{\sigma_1}}_{=\vec{u}_1} & \underbrace{\frac{A\vec{v}_2}{\sigma_2}}_{=\vec{u}_2} & \cdots & \underbrace{\frac{A\vec{v}_r}{\sigma_r}}_{=\vec{u}_r} & \vec{u}_{r+1} & \cdots & \vec{u}_m \\ \downarrow & \downarrow & & \downarrow & \downarrow & & \downarrow \end{bmatrix}_{m\times m} \quad (15.1)}$$

Bringing together (13.3), (13.4) and (15.1), we arrive to the desired decomposition:

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

or

$$A = \begin{bmatrix} \vec{u}_1 \, \vec{u}_2 \cdots \vec{u}_r \; \vec{u}_{r+1} \cdots \vec{u}_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & 0 \\ & \ddots & & 0 \cdots & 0 \\ & & \sigma_r & & \\ & 0 & & 0 \; 0 \cdots & 0 \\ & & & \ddots & \\ & & & 0 \cdots & 0 \end{bmatrix} \begin{bmatrix} \vec{v}_1 \, \vec{v}_2 \cdots \vec{v}_n \end{bmatrix}^T \qquad (15.2)$$

This special factorization of A is called

<u>THE SINGULAR VALUE DECOMPOSITION.</u>