Singular Value Decomposition –

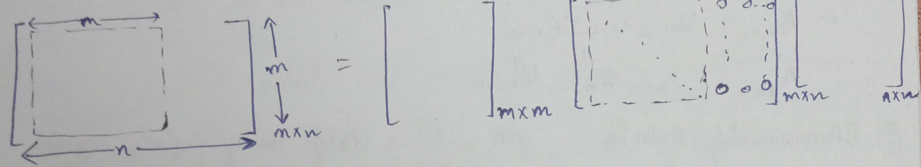
SVD is a matrix decomposition method for reducing a matrix to its constituent parts in order to make certain subsequent matrix calculations simpler.

A is the real $m \times n$ matrix that we wish to decompose

U is $m \times m$ matrix, Σ is $m \times n$ diagonal matrix and

V^T is $n \times n$ matrix.

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



Applications of SVD – ① Calculation of other matrix operations, such as, matrix inverse, ② Data reduction method in ML. ③ least squares linear regression, image compression & denoising data.

① **SVD for Pseudoinverse**: The pseudoinverse is the generalization of the matrix inverse for square matrices to rectangular matrices where the no. of col. & rows are not equal.

$$\cancel{A^+ = (U\Sigma V^T)^+ = (V^T)^+ \Sigma^+ U^+}$$

$$A^{-1} = (U \Sigma V^T)^{-1}$$
$$= (V^T)^{-1} \Sigma^{-1} U^{-1} = V \Sigma^{-1} U^T$$

($\because$ U & V are orthogonal matrix $\therefore$ $U^TU = UU^T = 1$, $VV^T = V^TV = 1$)

→ order of A is $m \times n$
" " U is $m \times m$
" " Σ is $m \times n$
" " V^T is $n \times n$
} SVD

→ order of A^+ is $n \times m$, order of V is $n \times n$, order of Σ^+ is $n \times m$ and order of U^T is $m \times m$

sigma inverse, Σ^{-1} is transpose of Σ matrix with the singular values reciprocal.

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \vdots & 0 & 0 & \cdots & 0 \\ & \sigma_2 & & \vdots & \vdots & & & \vdots \\ & & \ddots & \vdots & 0 & 0 & \cdots & 0 \\ & & & \sigma_k & 0 & 0 & \cdots & 0 \end{bmatrix}_{m \times n}$$

$$\Sigma^{-1} = \begin{bmatrix} 1/\sigma_1 & & & \\ & 1/\sigma_2 & & \\ & & \ddots & 1/\sigma_k \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & \cdots & 0 \end{bmatrix}_{n \times m}$$

$\because$ $A_{m\times n} = U_{m\times m} \Sigma_{m\times n} V^T_{n\times n}$

$A^+_{n\times m} = V_{n\times n} \Sigma^{-1}_{n\times m} U^T_{m\times m}$

(2) Dimensionality Reduction - Data with a large no. of features, such as more features (columns), than observations (rows) [3 x 10 matrix]

hair	colour	size	quality	thickness . . . .
	red	x	bad	1.3 - - -
	Yellow	y	good	1.4 - - -
	Brown	z	good	1.5 - - -

so, we can reduce the dimension i.e. Data may be reduced to a smaller subset of features that are most relevant to the prediction problem.

The result is a matrix with lower rank that is said to approximate the original matrix.

firstly, we perform SVD operation on the original data set & select top k largest singular values in sigma.
These columns can be selected from sigma & the rows selected from V^T

$$\boxed{B = U \Sigma_k V^T_k}$$

$T = U \cdot \Sigma_k$, $T = V_k \cdot A$

$$\underset{m\times n}{A} = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}\,\underset{n\times n}{V^T}$$

$$\underset{m\times n}{A} = \underset{m\times m}{U}\,\underset{m\times k}{\Sigma_k}\,\underset{k\times n}{V_k^T}, \quad \text{where } k<n$$

→ While choosing these k singular values, one may come accross singular values that are relatively close to zero and it is not possible to accurately compute them. In such we call the matrix ill-conditioned, because dividing by the singular values (that are close to zero will result in numerical errors.

→ The degree to which ill-conditioning prevents a matrix from being inverted accurately depends on the ratio of its largest to smallest singular value, a quantity known as the condition no.

condition no. $= \Sigma_{11}/\Sigma_{nn}$.

The larger the condition no., the more practical non-invertibility is.

Suppose there is a document matrix of $n\times d$ order where there are n documents, d terms (eg. theorems, proofs, etc) & it is represented by a matrix, 'A'

$$\begin{array}{cc} & d \text{ terms (eg: thm, proofs, etc)} \\ n \text{ documents} & \begin{bmatrix} A \\ A_{ij} = \text{frequency of } j\text{th term in } i\text{th document} \end{bmatrix} \end{array}$$

i.e. A_{ij} represents, frequency of j'a theorems' in i'th document. like in 10th document there are theorems i.e. $A(10)(5) = 6$ (6 thms in 10th document)

so, finding a subset of terms that accurately clusters the document (this is done by SVD)

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \underset{U}{\begin{bmatrix} 0.18 & 0 \\ 0.36 & 0 \\ 0.18 & 0 \\ 0.90 & 0 \\ 0 & 0.53 \\ 0 & 0.80 \\ 0 & 0.27 \end{bmatrix}_{7\times 2}} \times \underset{\Sigma}{\begin{bmatrix} 9.64 & 0 \\ 0 & 5.29 \end{bmatrix}_{2\times 2}} \times \underset{V}{\begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0.71 \end{bmatrix}_{2\times 5}}$$

$$A_{7\times 5} = U \Sigma V^T$$

A has rank 2 (clearly seen), we get 2 singular values & thus we formed U, Σ, V^T.

now, for dim. reduction, we take first singular value & set the other as 0 i.e. 5.29 replaced by 0.

then Σ is reduced to 1×1 ∴ U to 7×1 & V to 1×5

$$\begin{bmatrix} 0.18 \\ 0.36 \\ 0.18 \\ 0.90 \\ 0 \\ 0 \\ 0 \end{bmatrix}_{U_k} \times \begin{bmatrix} 9.64 \end{bmatrix}_{\Sigma_k} \times \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \end{bmatrix}_{V_k}$$

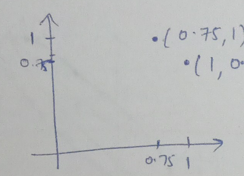
$$= \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \text{A matrix.}$$

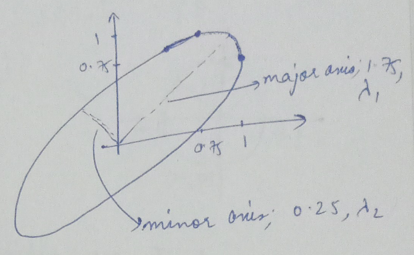
Geometrical Interpretation of SVD -

Basically, given any unit circle, if we apply SVD on it, we get an ellipse (ellipsoid for a sphere, in 3-D)

suppose, we are given a covariance matrix $A = \begin{bmatrix} 1 & 0.75 \\ 0.75 & 1 \end{bmatrix}$

then, $\lambda_1 = 1.75$ and $\lambda_2 = 0.25$





ellipse has major axis as largest eigen value & minor axis as smaller eigen value

→ SVD tells us the action of A on any vector x, in three steps

1. Rotation (multiply by V, that doesn't change vector length of x)
2. Scaling / stretching (the ith component is stretched by s_i) ↳ ith singular value.
3. another Rotation (multiplication by U)

$$\vec{y} = A\vec{x} = (USV^T)\vec{x}$$

↓ action of A upon $\vec{x}$.

$$\vec{x} \xrightarrow[\text{rotate}]{V^T} V^T\vec{x} \xrightarrow[\text{stretch}]{S} SV^T\vec{x} \xrightarrow[\text{rotate}]{U} USV^T\vec{x} = \vec{y}$$

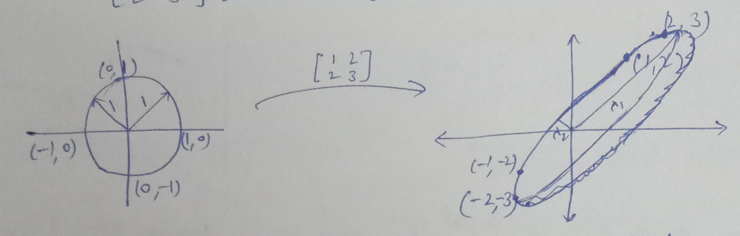
if $\vec{x}$'s are $\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ then $A\vec{x}$ will give ellipse $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \end{bmatrix}, \begin{bmatrix} -2 \\ -3 \end{bmatrix}$
is a circle

where $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$

$A\vec{x} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$,

$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$

$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$, same as $(V\Sigma V^T)x$.



we need to see which vector is transformed to the main axes of ellipse.

→ here, A is just a scaling matrix.
the eigen decomposition of A i.e. $Av_i = \lambda_i v_i$ tells us ~~for~~ which orthogonal axes (v_i) it scales, & by how much (λ_i)

V is orthogonal

$\boxed{\begin{matrix} AV = V\Sigma \\ AVV^T = V\Sigma V^T \\ A = V\Sigma V^T \end{matrix}}$

$Av_i = \lambda_i v_i$

$A = \lambda V V^T$

$A = [v_1\ v_2\ v_3 \dots v_n]\begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots\ \lambda_n \end{bmatrix}[v_1\ v_2 \dots v_n]^T$

$\boxed{A = V\Sigma V^T}$, we see the eigen vectors of A are the axes of the ellipse.

Note: A also contain rotation.

$Av_i = \sigma_i u_i \ (\sigma_i \geq 0)$

$\boxed{AV = U\Sigma}$

→ $\boxed{A = U\Sigma V^T}$

SVD computation / calculating σ given A (m×n) matrix :

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

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

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

$\sqrt{\text{eigen values of } AA^T}$ = singulaar values of A, $\sqrt{\text{e. values of } A^TA}$ = singular values of A

$$\sqrt{\lambda(AA^T)} = \sigma(A) \qquad \sqrt{\lambda(A^TA)} = \sigma(A)$$

also, $\boxed{AA^T,\ A^TA \text{ is symmetric}}$

SVD example:

find SVD of the matrix $A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

Soln:

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

$\lambda_1, \lambda_2 = 2, 2$ (clearly seen)

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

$$(A^TA - \lambda I) = \begin{bmatrix} 1-\lambda & 0 & 1 & 0 \\ 0 & 1-\lambda & 0 & 1 \\ 1 & 0 & 1-\lambda & 0 \\ 0 & 1 & 0 & 1-\lambda \end{bmatrix} \quad , \quad \det(A^TA - \lambda I) = 0$$

$(1-\lambda)\left[(1-\lambda)^3 + 1(\lambda-1)\right] + 1\left[0(0-0) - (1-\lambda)(0-0) + 1(1-0)\right] = 0$

$\Rightarrow (1-\lambda)^4 - (1-\lambda)^2 + 1 = 0$

$\lambda^2 - \lambda + 1 = 0$

$\cancel{\Rightarrow [(1-\lambda)^2]^2 + 1 = 0 \Rightarrow (1+\lambda^2 - 2\lambda)^2 + 1 = 0}$

$\cancel{1 + \lambda^4 + 4\lambda^2 + 2\lambda^2 - 4\lambda^3 - 4\lambda + 1 = 0}$

$\cancel{\lambda^4 - 4\lambda^3 + 6\lambda^2 - 4\lambda + 2 = 0}$

$\cancel{(1-\lambda)\left[(1-\lambda)(\lambda^2 - 0) + 1(0)\right] + 1[0] = 0}$

$\cancel{(1-\lambda)\left[\lambda^2(1-\lambda) + 1\right] = 0}$

$\cancel{\lambda^2(1-\lambda) + 1 = 0}$

$\cancel{\lambda^2 - \lambda^3 + 1 = 0}$

$\det(A^TA - \lambda I) = 0$

$$\det\begin{bmatrix} 1-\lambda & 0 & 1 & 0 \\ 0 & 1-\lambda & 0 & 1 \\ 0 & 0 & 1-\lambda & 0 \\ 0 & 1 & 0 & 1-\lambda \end{bmatrix} = 0 \Rightarrow (1-\lambda)\begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & 1-\lambda \end{vmatrix} + 1\begin{vmatrix} 0 & 1-\lambda & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1-\lambda \end{vmatrix} = 0$$

$$\Rightarrow (1-\lambda)\left[(1-\lambda)\{(1-\lambda)^2 - 0\} + 1(0-(1-\lambda))\right] + 1\left[(1-\lambda)\{(1-\lambda)-0\} + 1\{1-0\}\right]$$

$$\Rightarrow (1-\lambda)\left[(1-\lambda)^3 - (1-\lambda)\right] + \left(-(1-\lambda)^2 + 1\right) = 0$$

$$(1-\lambda)^4 - (1-\lambda)^2 - (1-\lambda)^2 + 1 = 0$$

$$(1-\lambda)^4 - 2(1-\lambda)^2 + 1 = 0$$

let $(1-\lambda)^2 = x$, $x^2 - 2x + 1 = 0$

$(x-1)^2 = 0$

$\boxed{x_1 = 1}$, $\boxed{x_2 = 1}$

for $x_1 = 1$ $\therefore$ $(1-\lambda)^2 = 1$

$1 - \lambda = \pm 1$

$\lambda = 0, 2$

for $x_2 = 1$, $\therefore$ $\lambda = 0, 2$

$\therefore$ eigen values are $0, 0, 2, 2$

so, singular values are $\sigma_1 = \sigma_2 = \sqrt{2}$

now, we find eigen vector corresponding to these eigen values,

$AA^T = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ $\qquad (AA^T)v = \lambda v$

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cancel{= \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = \lambda\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}$$

$$\begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \Rightarrow \begin{matrix} x_1 = x_1 \\ x_2 = x_2 \end{matrix}$$

$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in R^2$ where $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is orthonormal $\therefore$ let us take eigen vectors as $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ & $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ the standard basis of R^2.

$$\therefore U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

now, we need to find V^T.

V^T can be find either by finding eigen vectors of A^TA or

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

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

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

$$\therefore \boxed{V = A^T U \Sigma}$$

hence, $V = [v_1 \ v_2 \ v_3 \ v_4]$

then $v_1 = \frac{1}{\sigma_1} A^T u_1$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$v_2 = \frac{1}{\sigma_2} A^T u_2$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

σ_3, σ_4 are zeros, $\therefore$ v_3, v_4 we need to find by some other method we know v_1, v_2, v_3, v_4 forms orthonormal basis, $\therefore$ we choose v_3, v_4 as $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$ & $\frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

$\cancel{A = P\Sigma Q^T}$ $A = U\Sigma V^T$

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

Q.2 $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$

$$AA^T = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 9+4+4 & 6+6-4 \\ 6+6-4 & 4+9+4 \end{bmatrix}$$

$$= \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

$\det(AA^T - \lambda I) = 0$

$$\det \begin{bmatrix} 17-\lambda & 8 \\ 8 & 17-\lambda \end{bmatrix} = 0$$

$(17-\lambda)^2 - 64 = 0$

$289 + \lambda^2 - 34\lambda - 64 = 0$

$\lambda^2 - 34\lambda + 225 = 0$

$\lambda^2 - 25\lambda - 9\lambda + 225 = 0$ $3 \times 5 \times 3 \times 5$

$(\lambda - 25)(\lambda - 9) = 0$

$\lambda = 25, 9$

$\therefore \sigma_1 = 5, \sigma_2 = 3$

now, $Ax = \lambda x$

$$\therefore \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 25 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow \begin{matrix} 17x_1 + 8x_2 = 25x_1 \\ 8x_1 + 17x_2 = 25x_2 \end{matrix}$$

$\Rightarrow -8x_1 + 8x_2 = 0$

$8x_1 - 8x_2 = 0$

$\Rightarrow 12x_1 = -8x_2$

$\cancel{\boxed{3x_1 = -2x_2}}$

$\cancel{3x_1 = -12x_2}$

$\cancel{\boxed{x_1 = -4x_2}}$

$9x_1 + 6x_2 = 0$

$\begin{bmatrix} x_1 \end{bmatrix} = \begin{bmatrix} -2 \end{bmatrix}$

$\sqrt{x_1^2+x_2^2} = \sqrt{(+1)^2+1^2} = \sqrt{2}$

$-x_1 + x_2 = 0$

$\rightarrow x_1 = x_2$

$\therefore x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$u_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} +1 \\ 1 \end{bmatrix}$

for $\lambda_2 = 9$

$$\begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 9 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$17x_1 + 8x_2 = 9x_1 \Rightarrow 8x_1 = -8x_2$

$8x_1 + 17x_2 = 9x_2 \qquad 8x_1 = -8x_2$

$\therefore \boxed{x_1 = -x_2}$

$\therefore x = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$\therefore u_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$\therefore U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

now, we need to find V,

$$A^TA = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

$$\det(A^TA - \lambda I) = 0 \Rightarrow \begin{vmatrix} 13-\lambda & 12 & 2 \\ 12 & 13-\lambda & -2 \\ 2 & -2 & 8-\lambda \end{vmatrix} = 0$$

$= (13-\lambda)\left[(13-\lambda)(8-\lambda) - 4\right] - 12\left[12(8-\lambda) + 4\right] + 2\left[-24 - 2(13-\lambda)\right] = 0$

$= (13-\lambda)\left[(13-\lambda)(8-\lambda) - 4\right] - 12\left[96 - 12\lambda + 4\right] + 2\left[-24 - 26 + 2\lambda\right] = 0$

$(13-\lambda)\left[104 - 21\lambda + \lambda^2 - 4\right] - 12\left[-12\lambda + 100\right] + 2\left[-50 + 2\lambda\right] = 0$

$\lambda = 25, 9, 0$

$A^T A v = \lambda v$

$$\begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 25 \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$13x + 12y + 2z = 25x$ $\Rightarrow -12x + 12y + 2z = 0$

$12x + 13y - 2z = 25y$ $\quad 12x - 12y - 2z = 0$

$2x - 2y + 8z = 25z$ $\quad 2x - 2y - 17z = 0$

$\Rightarrow \dfrac{x_1}{204-4} = \dfrac{-x_2}{-204+4} = \dfrac{x_3}{-24+24}$

$\Rightarrow \dfrac{x_1}{200} = \dfrac{-x}{-200} = \dfrac{x_3}{0}$

$\cancel{12x - 12y = 2z}$

$\cancel{6x - 6y = 2z}$

$\cancel{x - y = z/3}$

$\cancel{x - y = 17z/2}$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 200 \\ 200 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

or

$v_1 = \dfrac{1}{\sigma_1} A^T u_1$

$$= \frac{1}{5} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{5\sqrt{2}} \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$v_2 = \dfrac{1}{\sigma_2} A^T u_2$

$$= \frac{1}{3} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{3\sqrt{2}} \begin{bmatrix} -3+2 \\ -2+3 \\ -2-2 \end{bmatrix} = \frac{1}{3\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ -4 \end{bmatrix}$$

$$v_3 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\cancel{V} = \cancel{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1/3 & x_1\sqrt{2} \\ 1 & 1/3 & x_2\sqrt{2} \\ 0 & -4/3 & x_3\sqrt{2} \end{bmatrix}}$$

for $\sigma_3 = 0$,

$$\begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$13x + 12y + 2z = 0$

$12x + 13y - 2z = 0$

$2x - 2y + 8z = 0$

$$\frac{x_1}{104-4} = \frac{-x_2}{96+4} = \frac{x_3}{-24-26}$$

$$\frac{x_1}{100} = \frac{x_2}{-100} = \frac{x_3}{-50}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}$$

$$\sqrt{2^2+(-2)^2+(-1)^2} = \sqrt{4+4+1} = 3$$

$$\therefore \quad v_3 = \frac{1}{3}\begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}$$

$$V = \begin{bmatrix} 1/\sqrt{2} & -1/3\sqrt{2} & 2/3 \\ 1/\sqrt{2} & 1/3\sqrt{2} & -2/3 \\ 0 & -4/3\sqrt{2} & -1/3 \end{bmatrix}$$

$$V^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/3\sqrt{2} & 1/3\sqrt{2} & -4/3\sqrt{2} \\ 2/3 & -2/3 & -1/3 \end{bmatrix}$$

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = U \Sigma V^T$$

$$= \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/3\sqrt{2} & 1/3\sqrt{2} & -4/3\sqrt{2} \\ 2/3 & -2/3 & -1/3 \end{bmatrix}$$

The S.V.D decomposition of A.

kth approximation :

$$A = U\Sigma V^T = \sigma_1 u_1 v_1^T + \dots + \sigma_r u_r v_r^T$$

$$A_k = U_k \Sigma_k V_k^T = \sigma_1 u_1 v_1^T + \dots + \sigma_k u_k v_k^T$$

(k is largest singular value (or the ~~best~~ k pieces gives best approximation of A))

If B has rank k, then $\|A - B\| \geq \|A - A_k\|$

Eckart Young theorem.

Norms :-

$l2 \rightarrow \|A\|_2 = \sigma_1$ → largest singular value — ①

$\|A\|_F = \sqrt{|a_{11}|^2 + |a_{12}|^2 + \dots |a_{1n}|^2 + \dots |a_{m1}|^2 + \dots |a_{mn}|^2}$ — ②

Frobenius norm

l_2 $\|V\|_2 = \sqrt{v_1^2 + v_2^2 + \dots v_n^2}$

l_1 $\|V\|_1 = |v_1| + |v_2| + |v_3| \dots$

l^∞ $\|V\|_\infty = \max |v_i|$

$\|A\|_{Nuclear} = \sigma_1 + \sigma_2 + \dots \sigma_r$ (sum of all singular values)

eg- $k = 2$,

$$A = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3.5 & 3.5 & 0 & 0 \\ 3.5 & 3.5 & 0 & 0 \\ 0 & 0 & 1.5 & 1.5 \\ 0 & 0 & 1.5 & 1.5 \end{bmatrix}$$

$$A_2^k = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

rank 2,

so $4 \rightarrow 3.5$ $\quad 2 \rightarrow 1.5$

$3 \rightarrow 3.5$ $\quad 1 \rightarrow 1.5$

we differ only by $\frac{1}{2}$

but to put up with rank 2, we also replaced $B_{12}, B_{21}, B_{34}, B_{43}$ by respective values.

Thus making A_2^k the best approximation of A where

$$\|A - A_2\| \leq \|A - B\|$$