SVD and eigenvalue decomposition :-

Eigendecomposition :-

If $\underset{\sim}{A} \in \mathbb{R}^{m\times m}$ is non-defective, then it has complete set of m linearly independent eigenvectors.

$$\left.\begin{aligned} \underset{\sim}{A}\,\underset{\sim}{x}_1 &= \lambda_1 \underset{\sim}{x}_1 \\ \underset{\sim}{A}\,\underset{\sim}{x}_2 &= \lambda_2 \underset{\sim}{x}_2 \\ &\vdots \\ \underset{\sim}{A}\,\underset{\sim}{x}_m &= \lambda_m \underset{\sim}{x}_m \end{aligned}\right\} - (1)$$

$$\underset{\sim}{A}\begin{bmatrix} | & | & | & & | \\ \underset{\sim}{x}_1 & \underset{\sim}{x}_2 & \underset{\sim}{x}_3 & \cdots & \underset{\sim}{x}_m \\ | & | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \underset{\sim}{x}_1 & \underset{\sim}{x}_2 & \cdots & \underset{\sim}{x}_m \\ | & | & & | \end{bmatrix}\begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots\ \lambda_m \end{bmatrix}$$

$$\underset{\sim}{A}\,\underset{\sim}{X} = \underset{\sim}{X}\,\underset{\sim}{\Lambda} \quad - (2)$$

$$\Rightarrow \underset{\sim}{A} = \underset{\sim}{X}\,\underset{\sim}{\Lambda}\,\underset{\sim}{X}^{-1} \quad \Rightarrow \underset{\sim}{\Lambda} = \underset{\sim}{X}^{-1}\underset{\sim}{A}\,\underset{\sim}{X} \quad - (3)$$

For $\underset{\sim}{A} \in \mathbb{R}^{m\times m}$, $\underset{\sim}{A}\,\underset{\sim}{x} = \underset{\sim}{b}$ and we expand $\underset{\sim}{b}, \underset{\sim}{x} \in \mathbb{R}^m$ in the basis of eigenvectors, $\underset{\sim}{b} = \sum_{i=1}^{m} b_i' \underset{\sim}{x}_i$

$$\underset{\sim}{b} = \underset{\sim}{X}\,\underset{\sim}{b}', \quad \underset{\sim}{b}' = \begin{bmatrix} b_1' \\ b_2' \\ \vdots \\ b_m' \end{bmatrix}$$

and $\underset{\sim}{x} = \sum_{i=1}^{m} x_i' \underset{\sim}{x}_i$

$$\Rightarrow \underset{\sim}{x} = \underset{\sim}{X}\,\underset{\sim}{x}'$$

where $\underset{\sim}{x}' = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_m' \end{bmatrix}$

$$\underset{\sim}{A}\,\underset{\sim}{x} = \underset{\sim}{b}$$

$$\underset{\sim}{A}\,\underset{\sim}{X}\,\underset{\sim}{x}' = \underset{\sim}{X}\,\underset{\sim}{b}'$$

$$\Rightarrow \underbrace{\underset{\sim}{X}^{-1}\underset{\sim}{A}\,\underset{\sim}{X}}\,\underset{\sim}{x}' = \underset{\sim}{b}'$$

$$\underset{\sim}{\Lambda}\,\underset{\sim}{x}' = \underset{\sim}{b}'$$

<u>Comparison</u> of <u>SVD</u> and <u>eigen decomp</u>:-

* SVD uses two different bases (left singular vectors basis and right singular vectors basis) whereas eigen decomposition uses just one basis ie basis of eigen vectors

* SVD results in orthonormal bases whereas $\underset{\sim}{X}$ in eigendecomposition is generally not orthogonal.

* SVD exists for all matrices no matter dimension; eigendecomposition

exists for square non-defective matrices.

<u>Thm 4</u>:- Non zero singular values of $\underset{\sim}{A} \in \mathbb{R}^{m \times n}$ are square roots of non-zero eigenvalues of $\underset{\sim}{A}^T \underset{\sim}{A}$ or $\underset{\sim}{A}\underset{\sim}{A}^T$.

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

Pf:-
$$\underset{\sim}{A}^T \underset{\sim}{A} = (\underset{\sim}{U}\underset{\sim}{\Sigma}\underset{\sim}{V}^T)^T (\underset{\sim}{U}\underset{\sim}{\Sigma}\underset{\sim}{V}^T)$$
$$= (\underset{\sim}{V}^T)^T \underset{\sim}{\Sigma}^T \underbrace{\underset{\sim}{U}^T\underset{\sim}{U}} \underset{\sim}{\Sigma}\underset{\sim}{V}^T$$
$$= \underset{n\times n}{\underset{\sim}{V}} \underset{n\times m}{\underset{\sim}{\Sigma}^T} \underset{m\times n}{\underset{\sim}{\Sigma}} \underset{n\times n}{\underset{\sim}{V}^T} = \underset{\sim}{V} \tilde{\underset{\sim}{\Sigma}} \underset{\sim}{V}^T$$

where $\tilde{\Sigma}$ is diag matrix of squares of singular values

$\underset{\sim}{V} \tilde{\underset{\sim}{\Sigma}} \underset{\sim}{V}^T$ is eigendecomp of $\underset{\sim}{A}^T\underset{\sim}{A}$

comprising $\sigma_1^2, \sigma_2^2, \ldots \sigma_p^2$ as diagonal of entries of $\tilde{\Sigma}$ with $n-p$ additional zero eigenvalues if $n > p$

<u>Thm 5</u>:- If $\underset{\sim}{A} = \underset{\sim}{A}^T$ singular values of $\underset{\sim}{A}$ are absolute values of eigenvalues of $\underset{\sim}{A}$.

Pf:-
$$\underset{\sim}{A} = \underset{\sim}{X}\, \underset{\sim}{\Lambda}\, \underset{\sim}{X}^{-1}$$
$$= \underset{\sim}{Q}\, \underset{\sim}{\Lambda}\, \underset{\sim}{Q}^{-1}$$
(where $\underset{\sim}{Q}$ is an orthogonal matrix for matrices which are symmetric)
$$= \underset{\sim}{Q}\, \underset{\sim}{\Lambda}\, \underset{\sim}{Q}^{T}$$
$$= \underset{\sim}{Q}\, |\underset{\sim}{\Lambda}|\, \underline{\mathrm{sgn}(\Lambda)}\, \underset{\sim}{Q}^{T}$$

where $|\Lambda|$ and $\mathrm{sgn}(\Lambda)$ are diagonal matrices whose entries are $|\lambda_j|$ and $\mathrm{sgn}(\lambda_j)$.

$$\underset{\sim}{A} = \underset{\sim}{Q}\, |\Lambda|\, \underbrace{\mathrm{sgn}(\Lambda)\, \underset{\sim}{Q}^{T}}$$
$$= \underset{\sim}{Q}\, |\Lambda|\, \bar{\underset{\sim}{Q}}^{T}$$
$$\underset{\sim}{A} = \underset{\sim}{Q}\, |\Lambda|\, \bar{\underset{\sim}{Q}}^{T} \text{ is SVD of } A$$

Thm 6:- $\underset{\sim}{A} \in \mathbb{R}^{m\times m}$; $|\det(A)| = \prod_{i=1}^{m} \sigma_i$

$$|\det(A)| = |\det(\underset{\sim}{U}\, \underset{\sim}{\Sigma}\, \underset{\sim}{V}^{T})|$$
$$= |\det U|\, |\det \Sigma|\, |\det V^{T}|$$
$$= |\det \Sigma| = \prod_{i=1}^{m} \sigma_i$$
$$[|\det V| = |\det U| = 1]$$

Low rank approximations:-

Thm 7:- $A \in \mathbb{R}^{m \times n}$ of rank "r" can be written as sum of "r" rank-one matrices of the form $A = \sum_{j=1}^{r} \sigma_j u_j v_j^T$

where $\{\sigma_j\}$ are singular values and $\{u_j\}$, $\{v_j\}$ are the appropriate singular vectors.

Pf:- Recall $u v^T$ is a rank-one matrix

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

$$\Sigma = \text{Sum}_{j=1}^{r} \Sigma_j \quad \text{where}$$

$$\Sigma_j = \begin{bmatrix} 0 & 0 & 0 & \cdots & \sigma_j & \cdots & 0 \\ \hline & & & 0 & & & \end{bmatrix}_{m \times n}$$

$$\begin{aligned} A &= U \Sigma V^T \\ &= U \left\{ \text{Sum}_{j=1}^{r} \Sigma_j \right\} V^T \\ &= \text{Sum}_{j=1}^{r} U \Sigma_j V^T = \sum_{j=1}^{r} \sigma_j u_j v_j^T \end{aligned}$$

$$A = \sum_{j=1}^{r} \sigma_j u_j v_j^T \qquad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$$

Then k^{th} partial sum $\sum_{j=1}^{k} \sigma_j u_j v_j^T$ $\quad k \leq r$

has as much energy (information) of A as possible

Thm 8:- For any k with $1 \leq k \leq r$

define $\boxed{A_k = \sum_{j=1}^{k} \sigma_j u_j v_j^T}$

then $\|A - A_k\|_2 = \min_{\substack{B \in \mathbb{R}^{m \times n} \\ \text{rank}(B) \leq k}} \|A - B\|_2 = \sigma_{k+1}$

Eckhart - Young Theorem

Proof:- Let there is some (B) whose $\text{rank}(B) \leq k$ such that $\|A - B\|_2 < \|A - A_k\|_2 = \sigma_{k+1}$

$\dim(N(B)) \geq n-k$

as $\text{rank}(B) \leq k$

Consider the subspaces

(i) W_1 : The null space of (B) which is of dimension of atleast $n-k$

(ii) W_2 : The space spanned by $k+1$ right singular vectors of A i.e $v_1, v_2, v_3 \cdots v_{k+1}$

These two subspaces have to intersect? why?

Dimensions of the two subspace add to $(n-k)+(k+1)$ i.e the subspaces must atleast have 1 common vector.

Let such a non-zero vector be $\underset{\sim}{x}$ i.e $\underset{\sim}{x} \in W_1 \cap W_2$

$\underset{\sim}{x} \neq 0$; $\underset{\sim}{x} \in N(B)$ i.e $\underset{\sim}{B}\underset{\sim}{x} = \underset{\sim}{0}$ and $\underset{\sim}{x} \in W_2$

$$\underset{\sim}{x} = \sum_{i=1}^{k+1} c_i \underset{\sim}{v_i}$$

$$\|\underset{\sim}{A}\underset{\sim}{x}\|_2 = \|(\underset{\sim}{A}-\underset{\sim}{B})\underset{\sim}{x}\|_2 \leq \|A-B\|_2 \|\underset{\sim}{x}\|_2 < \sigma_{k+1} \|\underset{\sim}{x}\|_2 \quad - (1)$$

$$\underset{\sim}{A}\underset{\sim}{v_i} = \underbrace{\sigma_i \underset{\sim}{u_i}}$$

$$\|\underset{\sim}{A}\underset{\sim}{x}\|_2^2 = \|\underset{\sim}{A}\sum_{i=1}^{k+1} c_i \underset{\sim}{v_i}\|_2^2 \qquad \|\underset{\sim}{a}\|_2^2 = \underset{\sim}{a}^T\underset{\sim}{a}$$

$$= \|\sum_{i=1}^{k+1} c_i \sigma_i \underset{\sim}{u_i}\|_2^2 = \left[\sum_{i=1}^{k+1} c_i \sigma_i \underset{\sim}{u_i}\right]^T \left[\sum_{i=1}^{k+1} c_i \sigma_i \underset{\sim}{u_i}\right]$$

$$= \sum_{i=1}^{k+1} c_i^2 \sigma_i^2 \quad \text{(use ortho normality of } \underset{\sim}{u_i} \text{]}$$

$$\underbrace{\sum_{i=1}^{k+1} c_i^2 \sigma_i^2} \geq \underbrace{\sum_{i=1}^{k+1} c_i^2 \sigma_{k+1}^2}$$

$$= \left(\sum_{i=1}^{k+1} c_i^2\right) \sigma_{k+1}^2$$

$$= \|\underset{\sim}{x}\|^2 \sigma_{k+1}^2$$

$$\|\underset{\sim}{A}\underset{\sim}{x}\|_2^2 \geq \|\underset{\sim}{x}\|^2 \sigma_{k+1}^2 \text{ i.e } \|\underset{\sim}{A}\underset{\sim}{x}\|_2 \geq \sigma_{k+1} \|x\|_2 \quad - (2)$$

① & ② is a contradiction which means you cannot have a matrix B with $\text{rank}(B) \le k$ such that $\|A-B\|_2 < \|A-A_k\|_2$

<u>Eckhart - Young</u> in <u>Frobenius norm</u>:-

Thm 9: For any k, with $1 \le k \le r$, the matrix $A_k = \sum_{j=1}^{k} \sigma_j u_j v_j^T$ also satisfies

$$\|A - A_k\|_F = \min_{\substack{B \in \mathbb{R}^{m \times n} \\ \text{rank}(B) \le k}} \|A - B\|_F$$

$$= \sqrt{\sigma_{k+1}^2 + \sigma_{k+2}^2 + \dots \sigma_r^2}$$