Singular Value Decomposition (SVD)

* why SVD?

⇒ If A is a symmetric matrix and real it can be decomposed as

$$A = Q \Lambda Q^T \quad \text{(spectral theorem)}$$

⇒ If A is not square, what can we do?

SVD can be used to decompose A & find useful information from A.

* Key points:

⇒ We need two sets of singular vectors: $U = \{u_1, u_2, \ldots, u_m\}$ & $V = \{v_1, v_2 \ldots v_n\}$ for $A_{m \times n}$.

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

the columns of U & V are orthonormal.

Σ is the diagonal matrix.

* For singular vectors

$$A v_i = \sigma u_i, \quad \text{for } i \leq r \text{ (rank of } A)$$

$$A v_{r+1} = 0 = A v_n.$$

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$$

Matrix form

$$AV = U\Sigma \qquad (V^T = V^{-1},\ U^T = U^{-1})$$

$$A[v_1\ v_2 \cdots v_n] = [u_1\ u_2 \cdots u_m] \begin{bmatrix} \sigma_1 & & & & 0 \\ & \sigma_2 & & & 0 \\ & & \ddots & \sigma_r & 0 \\ 0 & 0 & \cdots & & 0 \end{bmatrix}$$

* Reduced form of SVD

$AV = U\Sigma$ can have lot of zeros.

It can happen if A has smaller rank.

$$AV_r = U_r \Sigma_r \qquad (r = \text{rank})$$

$$A[v_1\ v_2 \cdots v_r] = [u_1\ u_2 \cdots u_r] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 \cdots & 0 \\ 0 & \cdots & & \sigma_r \end{bmatrix}$$

row space   column space.

* Importance of SVD:

⇒ Separates matrix in rank-1 pieces.

⇒ Pieces comes in order of importance.

⇒ The first piece $\sigma_1 u_1 v_1^T$ is the closest rank-1 matrix to A.

⇒ In general, sum of first k pieces is the best possible ~~for~~ rank k. approx to A.

$$A_K = \sum_{i=1}^{K} \sigma_i u_i v_i^T$$

~~is~~

$\boxed{\text{ECKART-YOUNG THEOREM: If } B \text{ has rank } k \text{ then } \|A - A_K\| \le \|A - B\|}$

Proof of SVD:

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

we want to find two sets of singular vectors u's & v's.

Let us try it using symmetric matrices:

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

* V contains orthonormal eigenvectors of $A^T A$
* U " " " " $A A^T$
* $\sigma_1^2, \sigma_2^2 \ldots, \sigma_r^2$ are non-zero eigenvalues of both $A^T A$ & $A A^T$.

SVD requires: $A v_k = \sigma_k u_k$.

Now choose orthonormal eigenvectors $v_1, v_2 \ldots, v_r$ of $A^T A$ & then choose $\sigma_k = \sqrt{\lambda_k}$.

Thus
$$A^T A v_k = \sigma_k^2 v_k$$
$$u_k = \frac{A v_k}{\sigma_k}$$

Now prove u's are eigenvectors of $A A^T$.

$$A A^T u_k = A A^T \frac{A v_k}{\sigma_k} = \sigma_k^2 u_k.$$

Also

$$u_j^T u_k = \left(\frac{A v_j}{\sigma_j}\right)^T \cdot \frac{A v_k}{\sigma_k} = \frac{v_j^T}{\sigma_j \sigma_k}(A^T A v_k) = \begin{cases} 1 & (j = k) \\ 0 & (j \neq k) \end{cases}$$

Geometry of SVD:

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

$U \to$ orthogonal $\to$ Rotation

$\Sigma \to$ diagonal $\to$ stretch

$V^T \to$ orthogonal. $\to$ Rotation.