17.08.2022

CLASS - IV

Theorem (Singular Value Decomposition) : If $A \in M_{m,n}(\mathbb{R})$, $\exists !$ orthogonal matrices $U = [u_1 | \dots | u_m] \in O_m(\mathbb{R})$ & $V = [v_1 | \dots | v_n] \in O_n(\mathbb{R})$ such that $U^T A V = \Sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_p)$

$p = \min\{m, n\}$, where $\sigma_1 \geq \dots \geq \sigma_p \geq 0$

$\uparrow$

1st singular value

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

$A^T A \in M_n(\mathbb{R})$; $R(A) \to$ Range of A

$R(A^T) \to$ Range of A^T

$N(A) \to$ Null space of A

$N(A^T) \to$ Null space of A^T.

v_i's are eigenvectors of $A^T A$

$u_i = \frac{1}{\sigma_i} A v_i \quad (\sigma_i > 0)$ $\qquad \{u_1, \dots, u_r\}$ span the range of A.

$(AV)_i = A v_i = \sigma_i u_i = (U\Sigma)_i$

$AV = U\Sigma$

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

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

Geometric Interpretation of SVD

$A \in M_{m,n}(\mathbb{R})$ $\{Ax ; \|x\|_2 \leq 1\}$ What does it look like geometrically?

$\to$ hyperellipsoid.

The semi-axis directions of $A(\mathbb{R}^n)_1$ are defined by u_i's and lengths are the corresponding singular values.

Cor 1 If $A \in M_{m,n}(\mathbb{R})$, then $\|A\|_2 = \sigma_1$

$$\|A\|_F = \sqrt{\sigma_1^2 + \dots + \sigma_p^2}, \quad p = \min(m,n)$$
$$= \sqrt{\langle A, A \rangle_F}$$

where $\langle A, B \rangle_F = tr(B^T A)$.

Proof:-

$$\|A\|_2 = \sup_{\|x\|_2 \leq 1} (\|Ax\|_2)$$

$$\|A\|_2 = \|U \Sigma V^T\|_2 = \|\Sigma\|_2 = \sigma_1.$$

$$\|A\|_F^2 = tr(A^T A) = tr(\Sigma^T \Sigma) = \sum_{i=1}^{p} \sigma_i^2 \text{ (sum of eigenvalues of } A^T A)$$

Cor 2 If $A \in M_{m,n}(\mathbb{R})$ and $E \in M_{m,n}(\mathbb{R})$, then

$$\sigma_{max}(A+E) \leq \sigma_{max}(A) + \|E\|_2$$

$$\sigma_{min}(A+E) \geq \sigma_{min}(A) - \|E\|_2.$$

Proof:-

$$\sigma_{max}(A+E) = \|A+E\|_2 \leq \|A\|_2 + \|E\|_2.$$

$$\sigma_{min}(A+E) = \inf_{\|x\|_2 = 1} \|(A+E)x\|_2.$$

$$= \inf_{\|x\|_2 = 1} \|Ax - (-Ex)\|_2$$

$$\geq \inf_{\|x\|_2 = 1} \|Ax\|_2 - \inf_{\|x\|_2 = 1} \|Ex\|_2$$

$$\geq \sigma_{min}(A) - \|E\|_2$$

Cor 3) If $A \in M_{m,n}(\mathbb{R})$ $m > n$, $z \in \mathbb{R}^m$. Then

$$\sigma_{max}[A|z] \geq \sigma_{max}(A)$$

$$\sigma_{min}[A|z] \leq \sigma_{min}(A)$$

Proof :- $\sigma_{max}(A) = \sup_{\|x\|_2=1} \|Ax\|_2 = \left\| [A|z]\begin{bmatrix} x \\ 0 \end{bmatrix} \right\|_2 \leq \sigma_{max}([A|z])$

- $\sigma_{min}(A) = \inf_{\|x\|_2=1} \|Ax\|_2 = \left\| [A|z]\begin{bmatrix} x \\ 0 \end{bmatrix} \right\|_2 \geq \sigma_{min}([A|z])$

Example. $A, B \in H_n(\mathbb{C})$: Given Spectral Distributions of A & B what can you say about the spectral distribution of $(A+B)$.

$$\langle (A+B)x, x \rangle = \langle Ax + x \rangle + \langle Bx, x \rangle$$

$$\sup_{\|x\|_2=1} \langle (A+B)x, x \rangle \leq \sup_{\|x\|_2=1} \langle Ax, x \rangle + \sup_{\|x\|_2=1} \langle Bx, x \rangle$$

$$\lambda_{A+B}(1) \leq \lambda_A(1) + \lambda_B(1)$$

$$\lambda_{A+B}(n) \geq \lambda_A(n) + \lambda_B(n)$$

Cor 4) If $A \in M_{m,n}(\mathbb{R})$ and ~~rank (A) = r then~~ A has r strictly positive singular values, then rank $(A) = r$ and Null $(A) = \text{Span}\{v_{r+1}, \dots, v_n\}$.

Range $(A) = \text{span}\{u_1, \dots, u_r\}$.

Cor 5) If $A \in M_{m,n}(\mathbb{R})$ & rank $(A) = r$, then

$$A = \sum_{i=1}^{r} \sigma_i \underbrace{u_i v_i^T}_{\text{rank 1 matrix}}$$

Proof:- $A = U \Sigma V^T$

$$= [\sigma_1 u_1 | \dots | \sigma_r u_r | 0 \dots | 0] \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$= \sum_{i=1}^{r} \sigma_i \underbrace{u_i v_i^T}_{\text{rank 1 matrix}}$$

$\sigma_1 u_1 v_1^T$ is the best rank 1 approximation to A.

Similarly, $\sigma_1 u_1 v_1^T + \dots + \sigma_k u_k v_k^T$ is the best rank k approximation to A. (In $\|\cdot\|_2$ sense)

Theorem (Eckhart - Young Theorem)

If $k < r = \text{rank}(A)$ and $A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$, then

$$\min_{\text{rank}(B)=k} \|A-B\|_2 = \|A - A_k\|_2 = \sigma_{k+1}$$

Pf:- $U^T A_k V = \text{diag}(\sigma_1, \dots, \sigma_k, 0, \dots, 0)$

A_k has rank k.

$U^T(A - A_k)V = \text{diag}(0, \dots, 0, \sigma_{k+1}, \dots, \sigma_p)$

$\|A - A_k\|_2 = \sigma_{k+1}$

Let $B = M_{m,n}(\mathbb{R})$ with $\text{rank}(B) = k$.

Choose $x_1, \dots, x_{n-k}$ unit vectors such that (orthonormal vectors)

$\text{Null}(B) = \text{Span}\{x_1, \dots, x_{n-k}\}$

$z \in \text{Span}\{x_1, \dots, x_{n-k}\} \cap \text{Span}\{v_1, \dots, v_{k+1}\} \neq \phi$.

$\|z\|_2 = 1$

$\therefore Bz = 0$.

$$Az = \sum_{i=1}^{n} \sigma_i (v_i^T z) u_i$$

$$\|A-B\|_2^2 \geq \|(A-B)z\|_2^2 = \|Az\|_2^2 = \sum_{i=1}^{k+1} \sigma_i^2 (v_i^T z)^2$$

$$\geq \sigma_{k+1}^2 \left(\sum_{i=1}^{k+1} |\langle v_i, z\rangle|^2 \right)$$

$\left[(\text{as } z \in \text{Span}\{v_1, \dots, v_{k+1}\})\right]$ $= \sigma_{k+1}^2 \|z\|^2$

$= \sigma_{k+1}^2$

$\Rightarrow \|A-B\|_2 \geq \sigma_{k+1}$ ∎