To do : SVD basics & intuition

1. Motivation
2. Spectral Theorem
3. Computing the SVD
4. Conceptualizing SVD

Motivation : Fundamental in many aspects of EECS — PCA, ML, classification, etc.

Spectral Theorem : Any symmetric matrix is orthogonally diagonalizable

→ Symmetric Matrix : $A = A^T$ e.g. $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

→ Orthogonally diagonalizable : $A = \underbrace{V \Lambda V^T}$ eigenvectors of A

→ the eigenvectors are mutually orthogonal and has a norm of 1

I can create a symmetric matrix from any given matrix A via $\begin{cases} AA^T \\ A^TA \end{cases}$

→ eigenvalues of A^TA / AA^T are non negative ($\lambda_i \geq 0$)

Quick Proof

let $\vec{v}$ be the eigenvector of A^TA

$\therefore A^TA\vec{v} = \lambda\vec{v}$ Multiply $\vec{v}^T$ both sides

$\vec{v}^T A^T A\vec{v} = \vec{v}^T \lambda \vec{v}$

$(A\vec{v})^T A\vec{v} = \lambda \vec{v}^T \vec{v}$ $\because (AB)^T = B^T A^T$

<u>Recall</u>: $\|\vec{x}\|_2^2 = \vec{x}^T x$ [Monday's Disc]

$\therefore \|A\vec{v}\|_2^2 = \lambda \|\vec{v}\|_2^2$

$\therefore \lambda = \dfrac{\|A\vec{v}\|_2^2}{\|\vec{v}\|_2^2} \geq 0$ because the norms of vectors are always ≥ 0

Check if $A^T A = (A^T A)^T$

$= A^T \cdot (A^T)^T = A^T A$ ✓ $\therefore A^T A$ is symmetric

<u>Steps to computing SVD</u>: $A = U \Sigma V^T$

⓪ Understand dimensionally what matrix size each component is.

$\underset{m\times n}{A} = \underset{m\times m}{U} \underset{m\times n}{\Sigma} \underset{n\times n}{V^T}$ [full form SVD]

① Pick $A^T A$ or AA^T

② If $A^T A$, find the eigenvectors of $A^T A$ and order them:

$A^T A \vec{v_i} = \lambda_i \vec{v_i}$, $i = 1$ to r

for $\lambda_1 \geq \dots \lambda_r \geq 0$ $r = \text{rank}(A)$

$\vec{v_i}$ from forms the V basis

If AA^T, find the eigenvector of AA^T " " "

$\rightarrow AA^T \vec{u_i} = \lambda_i \vec{u_i}$

$\therefore \vec{u_i}$ forms the U matrix

③ $\sigma_i = \sqrt{\lambda_i}$

④ If $A^T A$, obtain $\vec{u_i}$ from $\vec{u_i} = \frac{1}{\sigma_i} A\vec{v_i}$, $i = 1$ to r

If AA^T, obtain $\vec{v_i}$ from $\vec{v_i} = \frac{1}{\sigma_i} A^T \vec{u_i}$, $i = 1$ to r

★ ⑤ To completely reconstruct U or V, you need to G.S. with appropriate basis vectors to get full matrix

Forms of SVD

① Full Form $A = U \Sigma V^T$

② Compact Form $A = U_r \Sigma_r V_r^T$, $r = \text{rank}(A)$

③ Outer Product Form $A = \sum_{i=1}^{r} \sigma_i u_i \vec{v_i}^T$

Q1 : $\underset{3\times 2}{\overset{(m\times n)}{A}} = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \underset{3\times 3}{U} \; \underset{3\times 2}{\Sigma} \; \underset{2\times 2}{V^T}$, $\text{rank}(A) = 1$

(a)(i) $A^T A = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$ $\quad \lambda_1 = 18 \longrightarrow \vec{v_1} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

$\lambda_2 = 0 \longrightarrow \vec{v_2} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

(ii) Unpack $V = [V_r , V_{n-r}] \longrightarrow V_{n-r} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

$V_r = \vec{v_1} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

Unpack $\Lambda \longrightarrow \Lambda_r$ = diagonal matrix of $\sqrt{\lambda_i}$ for $i = 1$ to r

$\therefore \Lambda_r = [18]$

(iii) $\Sigma_r = \Lambda_r^{1/2} = [\sqrt{18}]$

$\Sigma = \begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ (fill in 0s such that dimensionally matches with A)

(iv) $U_r = A V_r \Sigma_r^{-1} = \begin{bmatrix} -1/3 \\ 2/3 \\ -2/3 \end{bmatrix}$

$\therefore$ Construct U using gram-schmidt

$$\begin{bmatrix} -1/3 & 1 & 0 & 0 \\ 2/3 & 0 & 1 & 0 \\ -2/3 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{G.S.} \begin{bmatrix} -1/3 & \frac{\sqrt{8}}{3} & 0 & \cancel{0} \\ 2/3 & \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} & \cancel{0} \\ -2/3 & \frac{-1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} & \cancel{0} \end{bmatrix}$$

U

(a) $\underset{3\times2}{A} = U\Sigma V^T = \underset{3\times3}{\begin{bmatrix} -1/3 & \sqrt{8}/3 & 0 \\ 2/3 & \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -2/3 & \frac{-1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}} \underset{3\times2}{\begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}} \underset{2\times2}{\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}}^T$

$\underset{3\times2}{A} = U_r\Sigma_r V_r^T$

$= \underset{3\times1}{\begin{bmatrix} -1/3 \\ 2/3 \\ -2/3 \end{bmatrix}} \underset{1\times1}{\begin{bmatrix} \sqrt{18} \end{bmatrix}} \underset{1\times2}{\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}} \rightarrow 3\times2$ (compact form)

(b) $\overset{3\times2}{A} = U\Sigma V^T$

$\underset{2\times3}{A^T} = (U\Sigma V^T)^T = V\Sigma^T U^T$

$= \underset{2\times2}{\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}} \underset{2\times3}{\begin{bmatrix} \sqrt{18} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}} \underset{3\times3}{\begin{bmatrix} -1/3 & \sqrt{8}/3 & 0 \\ 2/3 & \frac{1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -2/3 & \frac{-1}{3\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}^T$

(c) $\text{Null}(A) = \text{Col}(V_{n-r})$

We can write $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$

Null(A) means $A\vec{x} = \vec{0}$, $\vec{x} \neq \vec{0}$

$A\vec{x} = \sum_{i=1}^{r} \sigma_i u_i \underline{\vec{v}_i^T \vec{x}}$

$\longrightarrow$ What vector of $\vec{x}$ makes $\vec{v}_i^T\vec{x} = 0$?

$\longrightarrow$ orthogonal vector to $\vec{v}_i$

$\therefore$ $\vec{x}$ to be $\perp$ to $\vec{v}_i$ for $i=1$ to r

$\therefore$ $\vec{x} \in \text{Col}(V_{n-r}) \longrightarrow V$ is orthonormal

$[\underline{V_r}, V_{n-r}]$

orthogonal to V_r

$\therefore$ $\text{Null}(A) = \text{Col}(V_{n-r})$

(d) $\text{Col}(A) = \text{Col}(U_r)$

$A = \sum_{i=1}^{r} \sigma_i u_i \vec{v}_i^T$

Col(A) means that for a given vector $\vec{b}$, $A\vec{x} = \vec{b}$

$$\therefore \quad A\vec{x} = \sum_{i=1}^{r} \sigma_i \vec{u_i} \underbrace{\vec{v_i}^T \vec{x}}_{\neq 0,\ \vec{v_i}^T\vec{x} \text{ is a scalar}}$$

$$= \sum_{i=1}^{r} \underbrace{(\vec{v_i}^T \vec{x}) \sigma_i}_{\text{scalar}} \vec{u_i}$$

$$= \sum_{i=1}^{r} \alpha_i \vec{u_i} \quad \Rightarrow \text{linear combination of } \vec{u_1}, \vec{u_2}, \ldots \vec{u_r}$$

$\therefore \ \vec{b} \in Col(U_r) \quad \therefore \ A\vec{x} \in Col(U_r)$

$\therefore \ Col(A) = Col(U_r) \qquad \therefore \ \vec{b} \in Col(A)$

$\therefore \ Col(A) = Col(U_r$

(f) AA^T

$r = \text{rank}(A)$

$(U, \Lambda) = \text{diagonalize}\ (AA^T)$

Unpack $\Lambda = \begin{bmatrix} \Lambda_r & 0_{r \times (m-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (m-r)} \end{bmatrix}$

$\Sigma_r = \Lambda_r^{1/2}$

Pack Σ

$V_r = A^T U_r \Sigma_r^{-1}$

$V = \text{extend basis}(V_r, \mathbb{R}^n)$

return (U, Σ, V)