Singular Value Decomposition (SVD)

- Diagonalisable matrix:

$$M = P \cdot \underbrace{\Lambda}_{\text{diagonal.}} \cdot P^{-1}$$

- Symmetric: $P^{-1} = P^{T}$

- Diagonal of Λ = eig values of M
- Columns of P = eig-vectors of M

SVD = "diagonalisation" for non-square matrices

THM: $M \in \mathbb{R}^{m \times n}$ then:

(#) $$M = U \cdot \Sigma \cdot V^{T}$$

where:

- $U \in \mathbb{R}^{m \times m}$, $U^{T}U = I_{m \times m}$
- $V \in \mathbb{R}^{n \times n}$, $V^{T}V = I_{n \times n}$
- $\Sigma \in \mathbb{R}^{m \times n}$ – "diagonal"

$$\downarrow$$

$$\Sigma_{i,j} = 0 \text{ if } i \neq j.$$

$m < n$:

$$\Sigma = \begin{array}{cc} m & n-m \\ \left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \sigma_2 & & 0 \\ 0 & & \sigma_m & \end{array}\right) \end{array} \Big\} m$$

$m > n$:

$$\Sigma = \begin{array}{c} \left(\begin{array}{ccc} \sigma_1 & & 0 \\ 0 & & \sigma_n \\ \hline & 0 & \end{array}\right) \\ n \end{array} \begin{array}{l} \} n \\ \} m-n \end{array}$$

⊗ Assume: $\sigma_1 \geq \sigma_2 \geq \sigma_3 \cdots \geq \sigma_r$ $r = \min(m,n)$

⊕ Can be written as:

$$M = \sum_{i=1}^{r} \sigma_i \underbrace{\underline{u}_i}_{m\times 1} \cdot \underbrace{\underline{v}_i^T}_{1\times n}$$

$m \times n$

$\underline{u}_i$ = i-th column of U

$\underline{v}_i$ = i-th column of V

Ex.

$$M = \begin{pmatrix} 4 & 0 & -6 \\ 3 & 0 & 8 \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}}_{U} \cdot \underbrace{\begin{pmatrix} 10 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}}_{\Sigma} \cdot \underbrace{\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{V^T}$$

SINGULAR VALUES / VECTORS

Let $M \in \mathbb{R}^{m\times n}$.

σ is a singular value of M if we can find $\underline{u} \in \mathbb{R}^m$, $\underline{v} \in \mathbb{R}^n$ such that:

$$M \cdot \underline{v} = \sigma \underline{u} \quad \& \quad \underline{u}^T M = \sigma \underline{v}^T$$

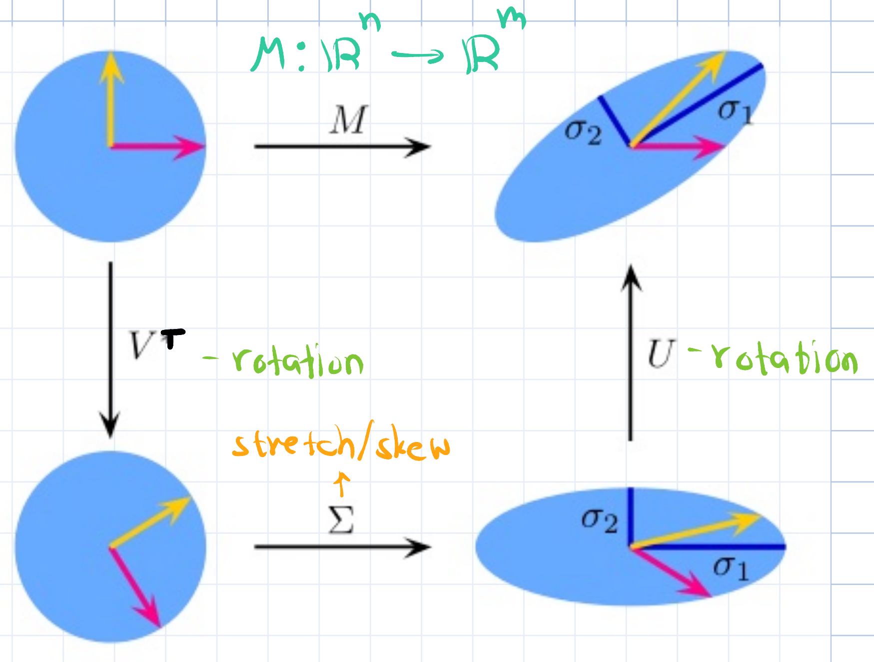
- $\underline{u}$ = left singular vector
- $\underline{v}$ = right singular vector

Conclusion:

(1) The values $\sigma_1, \_\_\_, \sigma_r$ in Σ (SVD) are all the singular values of M.

(2) Columns of U = left-singular vectors

Columns of V = right-singular vectors



$$M = U \cdot \Sigma \cdot V^*$$

Why?

$$\boxed{V \cdot V^T = I}$$

$$\underbrace{M \cdot v_i}_{i\text{-th column of } V} = (U \cdot \Sigma \cdot \overbrace{V^T}^{n \times n}) \cdot \overbrace{v_i}^{n \times 1} = U \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sigma_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \sigma_i \cdot u_i$$

$$V^T \cdot v_i = \begin{array}{c} \\ \\ \\ i \rightarrow \\ \\ \\ \end{array} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Ex.

$$c = \frac{1}{\sqrt{2}}$$

$$M = \begin{pmatrix} -1 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & -1 \end{pmatrix} \overset{SVD}{=} \underbrace{\begin{pmatrix} -c & c & 0 \\ 0 & 0 & 1 \\ c & c & 0 \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} c & 0 & -c \\ c & 0 & c \\ 0 & -1 & 0 \end{pmatrix}}_{V^T}$$

diagonalisation

$$= \underbrace{\begin{pmatrix} c & c & 0 \\ 0 & 0 & 1 \\ -c & c & 0 \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} -3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\Lambda} \underbrace{\begin{pmatrix} c & 0 & -c \\ c & 0 & c \\ 0 & 1 & 0 \end{pmatrix}}_{P^{-1}}$$

SVD vs. Diagonalisation

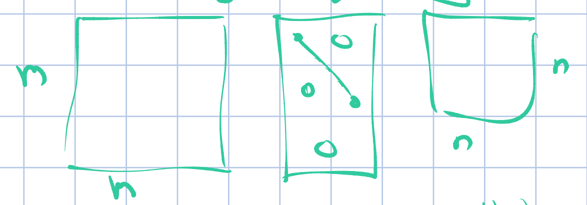
(1) eigenvalues – only square matrix.

(2) diagonalisation – only square matrix
+ diagonalisable.

(*) SVD = all matrices.

(3) singular values ≥ 0
eigenvalues – not necessarily.

Suppose $M = U\cdot\Sigma\cdot V^T$ (SVD), $\underline{m \geq n}$



Take $M^T M \in \mathbb{R}^{n\times n}$ → (small)

$$(U\Sigma V^T)^T\cdot(U\Sigma V^T)$$

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

$$\underbrace{\begin{pmatrix} \sigma_1 & 0 & \vdots & \\ & \ddots & \vdots & 0 \\ 0 & \sigma_n & \vdots & \end{pmatrix}}_{\Sigma^T} \underbrace{\begin{pmatrix} \sigma_1 & 0 \\ 0 & \ddots \ \sigma_n \\ \hline & 0 \end{pmatrix}}_{\Sigma} = \begin{pmatrix} \sigma_1^2 & 0 \\ & \ddots \\ 0 & \sigma_n^2 \end{pmatrix} = \Lambda$$

$$\Downarrow$$

diagonalisation of $M^T M$.

$$M^T M = \underbrace{V\cdot\Lambda\cdot V^T}, \quad VV^T = I$$

Λ → diagonal.

(1) $M^T M$ - diagonalisable

(2) $\sigma_1^2, \ldots, \sigma_n^2$ - eigen-values of $M^T M$ $M \cdot M^T$

(3) $v_1, \ldots, v_n$ (columns of V) are the eigenvectors of $M^T M$.

$u_1, \ldots, u_m$ U $M \cdot M^T$

⊛ Same with $M \cdot M^T$

⊛ $\text{rank}(M) = \#$ non-zero singular values

Application - Reduced Rank Approximation

$M \in \mathbb{R}^{m \times n}$, $\text{rank}(M) = R$.

We look $\tilde{M} \in \mathbb{R}^{m \times n}$ with $\text{rank}(\tilde{M}) = r < R$

Closest to M in Frobenius norm:

$$\|M - \tilde{M}\|_F^2 = \sum_{i,j} (M_{i,j} - \tilde{M}_{i,j})^2$$

Claim:

$$M = U \cdot \Sigma \cdot V^T \quad \text{(SVD)}$$

sing. values: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_R$

$$\Sigma = \left(\begin{array}{ccc|c} \sigma_1 & 0 & 0 & \\ 0 & \sigma_R & & 0 \\ & & 0 & \\ 0 & & \ddots\ 0 & \end{array} \right)$$

Define: $\tilde{\Sigma} = \begin{pmatrix} \sigma_1 & 0 & & \\ 0 & \ddots & & 0 \\ & \sigma_r & & \\ & & 0 & \\ 0 & & & \ddots \; 0 \end{pmatrix}$

Then: $\tilde{M} = U \cdot \underbrace{\tilde{\Sigma}}_{\text{"truncated"}} \cdot V^T$

Same as SVD of M.

Theorem by Echart-Young-Mirsky

Alternatively:

$$M = \sum_{i=1}^{R} \sigma_i \underline{u}_i \underline{v}_i^T$$

$$\downarrow$$

$$\tilde{M} = \sum_{i=1}^{r} \sigma_i \underline{u}_i \underline{v}_i^T$$

Ex.

$$M = \begin{pmatrix} 4 & 0 & -6 \\ 3 & 0 & 8 \end{pmatrix} \rightarrow \text{rank} = 2 = R$$

$$= \underbrace{\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}}_{U} \cdot \underbrace{\begin{pmatrix} 10 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}}_{\Sigma} \cdot \underbrace{\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{V^T}$$

$$\tilde{M} = \underbrace{\begin{pmatrix} \frac{3}{5} & \frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}}_{U} \cdot \underbrace{\begin{pmatrix} 10 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\tilde{\Sigma}} \cdot \underbrace{\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{V^T}$$

$$= \begin{pmatrix} 0 & 0 & -6 \\ 0 & 0 & 8 \end{pmatrix} \rightarrow \underline{\underline{\text{rank} = 1}}$$