

$$\text{i.e. } A\vec{p}_i = \lambda_i \vec{p}_i \text{ for } i=1, 2, \dots, n.$$

$\Rightarrow \lambda_1, \lambda_2, \dots, \lambda_n$ are all the eigenvalues of A .

As P is unitary, $\{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_n\}$ is an orthonormal basis of \mathbb{C}^n . i.e. $\vec{p}_i^* \vec{p}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$

" \Rightarrow " If A is positive definite, then $\forall \vec{x} \neq 0: \vec{x}^* A \vec{x} > 0$

In particular, $0 < \vec{p}_i^* A \vec{p}_i = \vec{p}_i^* \lambda_i \vec{p}_i = \lambda_i \vec{p}_i^* \vec{p}_i = \lambda_i$ for each $i=1, 2, \dots, n$.

" \Leftarrow " Suppose $\lambda_i > 0$ for all $i=1, 2, \dots, n$,

$\forall \vec{x} \in \mathbb{C}^n$, as $\{\vec{p}_1, \dots, \vec{p}_n\}$ is an orthonormal basis of \mathbb{C}^n : $\vec{x} = \sum_{i=1}^n a_i \vec{p}_i$

$$\begin{aligned} \text{then } \vec{x}^* A \vec{x} &= (A \vec{x}, \vec{x}) = \left(A \left(\sum_{i=1}^n a_i \vec{p}_i \right), \sum_{k=1}^n a_k \vec{p}_k \right) \\ &= \left(\sum_{i=1}^n a_i (A \vec{p}_i), \sum_{k=1}^n a_k \vec{p}_k \right) \\ &= \left(\sum_{i=1}^n a_i \lambda_i \vec{p}_i, \sum_{k=1}^n a_k \vec{p}_k \right) \\ &= \sum_{j=1}^n a_j \lambda_j \bar{a}_j (\vec{p}_j, \vec{p}_j) \\ &= \sum_{j=1}^n \lambda_j |a_j|^2 > 0 \end{aligned}$$

■

Singular Value Decomposition.

Theorem: Let $A \in \mathbb{C}^{n \times p}$. Suppose $\text{rank}(A)=r$. There exists unitary matrices $U \in \mathbb{C}^{n \times n}$ and

$$V \in \mathbb{C}^{p \times p} \text{ and } \Sigma \in \mathbb{R}^{n \times p} \text{ with } \Sigma = \begin{pmatrix} D_{rr} & 0 \\ 0 & 0 \end{pmatrix}$$

where $D = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}$ with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$. Such that

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

Note that the above decomposition is called a singular value decomposition. And $\sigma_1, \sigma_2, \dots, \sigma_r$ are called singular values of A .

Construction of V :

As $A \in \mathbb{C}^{n \times p}$, $A^*A \in \mathbb{C}^{p \times p}$. Since $(A^*A)^* = A^*A \Rightarrow A^*A$ is Hermitian.

By Spectral Theorem, there exists a unitary matrix V and a diagonal matrix $C = (\lambda_1 \ \dots \ \lambda_p) \in \mathbb{C}^{p \times p}$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$. Such that $A^*A = VCV^*$.

Denote $V = (\vec{v}_1 \ \dots \ \vec{v}_p)$, then $A^*A \vec{v}_i = \lambda_i \vec{v}_i$.

Construction of Σ :

Lemma: For any $A \in \mathbb{C}^{n \times p}$: (1). $\text{Null}(A) = (\text{Range}(A^*))^\perp$, (2) $\text{Null}(A^*A) = \text{null}(A)$

(3). $\text{Rank}(A^*A) = \text{Rank}(A) = \text{Rank}(A^*)$, (4). $\text{Range}(A^*A) = \text{Range}(A)$.

Lemma: Suppose $\text{rank}(A) = r$. Then $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$, $\lambda_{r+1} = \dots = \lambda_p = 0$.

Proof: Claim: $\lambda_1 \geq \dots \geq \lambda_p \geq 0$.

$$\text{For each } \lambda_i: \quad A^*A \vec{v}_i = \lambda_i \vec{v}_i$$

$$\Rightarrow \vec{v}_i^* A^* A \vec{v}_i = \lambda_i \vec{v}_i^* \vec{v}_i = \lambda_i \|\vec{v}_i\|^2$$

$$\text{On the other hand } \vec{v}_i^* A^* A \vec{v}_i = (\vec{A}\vec{v}_i)^* \vec{A}\vec{v}_i = \|\vec{A}\vec{v}_i\|^2$$

$$\Rightarrow \lambda_i \|\vec{v}_i\|^2 = \|\vec{A}\vec{v}_i\|^2 \geq 0$$

$$\text{As } \|\vec{v}_i\|^2 > 0 \Rightarrow \lambda_i \geq 0.$$

Suppose $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$, $\lambda_{k+1} = \dots = \lambda_p = 0$. We will prove that $k=r$.

As $A^*A \vec{v}_i = \lambda_i \vec{v}_i$, for any $\lambda_i \neq 0$: $A^*A(\pm \vec{v}_i) = \vec{v}_i \in \text{Range}(A^*A)$.

$$\Rightarrow \{\vec{v}_1, \dots, \vec{v}_k\} \subseteq \text{Range}(A^*A).$$

Since $\{\vec{v}_1, \dots, \vec{v}_k\}$ is linearly independent, $k \leq \text{rank}(A^*A) = \text{rank}(A) = r$.

As $\lambda_{k+1} = \dots = \lambda_p = 0$ and $A^*A \vec{v}_i = \lambda_i \vec{v}_i = \vec{0}$ for all $i = k+1, \dots, p$

$$\Rightarrow \{\vec{v}_{k+1}, \dots, \vec{v}_p\} \subseteq \text{null}(A^*A) = \text{null}(A)$$

$$\text{since } \dim \text{null}(A) + \text{rank}(A) = p \Rightarrow \dim \text{null}(A) = p - r$$

$$\Rightarrow p-k \leq p-r$$

$$\Rightarrow k \geq r$$

Since $k \geq r$ and $k \leq r \Rightarrow k=r$.

Definition: Define $\sigma_i = \sqrt{\lambda_i}$ for all $i=1, 2, \dots, r$.

$$\text{Take } D = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}_{n \times r} \Rightarrow \Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}_{n \times p} \in \mathbb{R}^{n \times p}.$$

Corollary: $\{\vec{v}_{r+1}, \vec{v}_{r+2}, \dots, \vec{v}_p\}$ is a basis of $\text{null}(A^*A) = \text{null}(A)$.

Proof: $\dim(\text{null}(A)) = p-r$ and $\{\vec{v}_{r+1}, \dots, \vec{v}_p\} \subseteq \text{null}(A^*A) = \text{null}(A)$ is linearly independent
 $\Rightarrow \{\vec{v}_{r+1}, \dots, \vec{v}_p\}$ is a basis of $\text{null}(A)$.

Construction of U:

Lemma: for each $i=1, 2, \dots, r$, define $\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i$. Then $\{\vec{u}_1, \dots, \vec{u}_r\}$ is orthonormal.

$$\begin{aligned} \text{Proof: } (\vec{u}_i, \vec{u}_j) &= \left(\frac{1}{\sigma_i} A \vec{v}_i, \frac{1}{\sigma_j} A \vec{v}_j \right) \\ &= \frac{1}{\sigma_i \sigma_j} (A \vec{v}_i, A \vec{v}_j) \\ &= \frac{1}{\sigma_i \sigma_j} (\vec{v}_i, A^* A \vec{v}_j) \\ &= \frac{1}{\sigma_i \sigma_j} (\vec{v}_i, \lambda_j \vec{v}_j) \\ &= \frac{\lambda_j}{\sigma_i \sigma_j} (\vec{v}_i, \vec{v}_j) = \begin{cases} 0, & i \neq j \\ \frac{\lambda_j}{\sigma_j^2} = 1, & i = j. \end{cases} \end{aligned}$$

Definition: Extend $\{\vec{u}_1, \dots, \vec{u}_r\}$ to an orthonormal basis $\{\vec{u}_1, \dots, \vec{u}_r, \vec{u}_{r+1}, \dots, \vec{u}_n\}$.

Then $U = (\vec{u}_1 \ \vec{u}_2 \ \dots \ \vec{u}_n) \in \mathbb{C}^{n \times n}$ unitary.

Remark: As $\{\vec{u}_1, \dots, \vec{u}_r\} \subseteq \text{Range}(A)$, and $\text{rank}(A) = r$

$\Rightarrow \{\vec{u}_1, \dots, \vec{u}_r\}$ is an orthonormal basis of $\text{Range}(A)$.

$\Rightarrow \{\vec{u}_1, \dots, \vec{u}_r\}$ is an orthonormal basis of $(\text{Range}(A))^\perp$.

Since $(\text{Range}(A))^\perp = \text{Null}(A^*)$

\Rightarrow To get $\{\vec{v}_m, \dots, \vec{v}_n\}$, we first find a basis of the solution set $A^* \vec{x} = \vec{0}$.

Then Gram-Schmidt and normalize the basis.

Theorem: With the construction of V , Σ and U above, $A = U\Sigma V^*$.

Proof: It is equivalent to prove $AV = U\Sigma$.

Denote $\vec{V} = (\vec{v}_1 \ \vec{v}_2 \ \dots \ \vec{v}_p)$. By the corollary: $\{\vec{v}_m, \dots, \vec{v}_n\} \subseteq \text{null}(A)$

$$\Rightarrow AV = (A\vec{v}_1 \ A\vec{v}_2 \ \dots \ A\vec{v}_r \ A\vec{v}_m \ \dots \ A\vec{v}_p)$$

$$= (A\vec{v}_1 \ A\vec{v}_2 \ \dots \ A\vec{v}_r \ \underbrace{\vec{0} \ \dots \ \vec{0}}_{\text{per columns}})_{n \times p}$$

$$U\Sigma = (\vec{u}_1 \ \vec{u}_2 \ \dots \ \vec{u}_r \ \vec{u}_m \ \dots \ \vec{u}_n) \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}_{n \times p}$$

$$= (\sigma_1 \vec{u}_1 \ \sigma_2 \vec{u}_2 \ \dots \ \sigma_r \vec{u}_r \ \underbrace{\vec{0} \ \dots \ \vec{0}}_{\text{per columns}})_{n \times p}$$

$$\downarrow \vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$$

$$= (A\vec{v}_1 \ A\vec{v}_2 \ \dots \ A\vec{v}_r \ \underbrace{\vec{0} \ \dots \ \vec{0}}_{\text{per columns}})$$

$$= AV. \quad \blacksquare$$

Corollary: Define $U_r = (\vec{u}_1 \ \vec{u}_2 \ \dots \ \vec{u}_r) \in \mathbb{C}^{n \times r}$

$$V_r = (\vec{v}_1 \ \vec{v}_2 \ \dots \ \vec{v}_r) \in \mathbb{C}^{p \times r}$$

$$D = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} \in \mathbb{C}^{r \times r}$$

$$\text{then } A = U_r D V_r^*$$

Remark: $A = U_r D V_r^*$ is called reduced SVD of A .

Proof. By singular value decomposition $A = U \Sigma V^*$

$$A = (\vec{u}_1 \vec{u}_2 \cdots \vec{u}_r \vec{0} \cdots \vec{0}) \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r & 0 \\ 0 & & & 0 \end{pmatrix} \begin{pmatrix} \vec{v}_1^* \\ \vec{v}_2^* \\ \vdots \\ \vec{v}_r^* \\ \vec{0}^* \\ \vdots \\ \vec{0}^* \end{pmatrix}$$

$$= (\sigma_1 \vec{u}_1 \vec{u}_2^* \cdots \sigma_r \vec{u}_r \vec{0}^* \cdots \vec{0}^*) \begin{pmatrix} \vec{v}_1^* \\ \vec{v}_2^* \\ \vdots \\ \vec{v}_r^* \\ \vec{0}^* \\ \vdots \\ \vec{0}^* \end{pmatrix}$$

$$= \sigma_1 \vec{u}_1 \vec{v}_1^* + \sigma_2 \vec{u}_2 \vec{v}_2^* + \cdots + \sigma_r \vec{u}_r \vec{v}_r^*. \quad (\text{Note } \vec{u}_i \in \mathbb{C}^{n \times 1}, \vec{v}_i^* \in \mathbb{C}^{m \times p}).$$

$$U_r D V_r^* = (\vec{u}_1 \vec{u}_2 \cdots \vec{u}_r) \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} \begin{pmatrix} \vec{v}_1^* \\ \vec{v}_2^* \\ \vdots \\ \vec{v}_r^* \\ \vec{0}^* \\ \vdots \\ \vec{0}^* \end{pmatrix}$$

$$= (\sigma_1 \vec{u}_1 \vec{u}_2^* \cdots \sigma_r \vec{u}_r^*) \begin{pmatrix} \vec{v}_1^* \\ \vec{v}_2^* \\ \vdots \\ \vec{v}_r^* \\ \vec{0}^* \\ \vdots \\ \vec{0}^* \end{pmatrix}$$

$$= \sigma_1 \vec{u}_1 \vec{v}_1^* + \cdots + \sigma_r \vec{u}_r \vec{v}_r^* = A. \quad \blacksquare$$

Theorem (Moore-Penrose Inverse). Let $A \in \mathbb{C}^{m \times p}$, and let $A = U_r D V_r^*$ be the reduced SVD of A . Then $A^+ = V_r D^* U_r^*$ is called the Moore-Penrose inverse of A , and A^+ satisfies the following conditions:

- 1) $AA^+A = A$
- 2) $A^+AA^+ = A^+$
- 3) $(AA^+)^* = AA^+$
- 4) $(A^+A)^* = A^+A$

Remark: One may check that if $A \in \mathbb{C}^{n \times n}$ and A is invertible, then $A^+ = A^{-1}$.

Proof: Note that $V_r^* V_r = \begin{pmatrix} \vec{v}_1^* \\ \vdots \\ \vec{v}_r^* \end{pmatrix} (\vec{v}_1 \vec{v}_2 \cdots \vec{v}_r) = \begin{pmatrix} \vec{v}_1^* \vec{v}_1 & \vec{v}_1^* \vec{v}_2 & \cdots & \vec{v}_1^* \vec{v}_r \\ \vdots & \vdots & \ddots & \vdots \\ \vec{v}_r^* \vec{v}_1 & \vec{v}_r^* \vec{v}_2 & \cdots & \vec{v}_r^* \vec{v}_r \end{pmatrix} = I_r \in \mathbb{C}^{r \times r}$
similarly $U_r^* U_r = I_r$.

- 1) $AA^+A = U_r D V_r^* V_r D^* U_r^* U_r D V_r^* = U_r D V_r^* = A.$
- 2) Similar to 1).
- 3) $AA^+ = U_r D V_r^* V_r D^* U_r^* = U_r U_r^* \Rightarrow AA^+ = U_r U_r^* = (U_r U_r^*)^* = (A^+ A)^*$
- 4) $A^+ A = V_r D^* U_r^* U_r D V_r^* = V_r V_r^* \Rightarrow (A^+ A)^* = A^+ A = V_r V_r^*. \quad \text{end of March 3}$