Spetral Theorem for Hermitian Matrices.

A matrix $A \in \mathbb{C}^{n\times n}$ is Hermitian if and only if there exists $P \in \mathbb{C}^{n\times n}$ unitary $D \in \mathbb{R}^{n\times n}$ diagonal such that $A = PDP^*$.

<u>Proof</u>: "$\Leftarrow$": If $A = PDP^*$ for some P unitary and $D \in \mathbb{R}^{n\times n}$ diagonal, then

$$A^* = (PDP^*)^* = (P^*)^* D^* P^* = PD^*P^* = PDP^* = A$$

$\Rightarrow$ A is Hermitian.

"$\Rightarrow$": If A is Hermitian. Then by spectral theorem, there exists P unitary and D diagonal such that $A = PDP^*$.

Since $A^* = (PDP^*)^* = PD^*P^* = A = PDP^* \Rightarrow D = D^*$

$\Rightarrow D \in \mathbb{R}^{n\times n}$.

Spetral Theorem for symmetric matrices.

$A \in \mathbb{R}^{n\times n}$ is symmetric if and only if there exists an orthogonal matrix P and a diagonal matrix $D \in \mathbb{R}^{n\times n}$ such that $A = PDP^T$

// end of Feb26.

Theorem. Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix. Then A is positive definite if and only if all the eigenvalues of A are positive.

<u>Proof</u>. By Spetrial Theorem, since A is Hermitian, there exists a unitary matrix $P = (\vec{p}_1 \vec{p}_2 \cdots \vec{p}_n) \in \mathbb{C}^{n\times n}$ and a diagonal matrix $D = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ such that $A = PDP^*$

$\Rightarrow AP = PD \Rightarrow A(\vec{p}_1\ \vec{p}_2 \cdots \vec{p}_n) = (\vec{p}_1\ \vec{p}_2 \cdots \vec{p}_n)\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix}$

$\Rightarrow (A\vec{p}_1\ A\vec{p}_2 \cdots A\vec{p}_n) = (\lambda_1\vec{p}_1\ \lambda_2\vec{p}_2 \cdots \lambda_n\vec{p}_n)$

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

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

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

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

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 $i=1,2,\cdots,n$.

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

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

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

$\blacksquare$

Singular Value Decomposition.

<u>Theorem</u>: Let $A\in\mathbb{C}^{n\times p}$. Suppose $\operatorname{rank}(A)=r$. There exists unitary matrices $U\in\mathbb{C}^{n\times n}$ and $V\in\mathbb{C}^{p\times p}$ and $\Sigma\in\mathbb{R}^{n\times p}$ with $\Sigma = \begin{pmatrix} D_{r\times r} & 0 \\ 0 & 0\end{pmatrix}$

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

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

Note that the above decomposition is called a singular value decomposition. And $\sigma_1, \sigma_2, \cdots, \sigma_r$ are 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 $C = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix} \in \mathbb{C}^{p\times p}$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$. Such that $A^*A = VCV^*$.

Denote $V = (\vec{v}_1 \cdots \vec{v}_p)$, 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). $\text{Range}(A^*A) = \text{Range}(A)$.

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

<u>Proof</u>: Claim: $\lambda_1 \geq \cdots \geq \lambda_p \geq 0$.

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

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

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

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

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

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

As $A^*A\vec{v}_i = \lambda_i \vec{v}_i$, for any $\lambda_i \neq 0$: $A^*A(\frac{1}{\lambda_i}\vec{v}_i) = \vec{v}_i \in \text{Range}(A^*A)$.

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

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

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

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

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

$$\Rightarrow \quad p-k \le p-r$$

$$\Rightarrow \quad k \ge r$$

since $k \ge r$ and $k \le r \Rightarrow k = r$.

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

Take $D = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix}_{r \times r} \Rightarrow \Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}_{n \times p} \in \mathbb{R}^{n \times p}$.

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

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

$\Rightarrow \{\vec{v}_{r+1}, \cdots, \vec{v}_p\}$ is a basis of $\text{null}(A)$.

Construction of U:

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

Proof:

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

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

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

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

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

$\Rightarrow \{\vec{u}_{r+1}, \cdots, \vec{u}_n\}$ is an orthonormal basis of $(\text{Range}(A))^\perp$.

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

$\Rightarrow$ To get $\{\vec{u}_{r+1}, \cdots, \vec{u}_n\}$, 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 $\vec{V} = (\vec{v}_1 \ \vec{v}_2 \ \cdots \ \vec{v}_p)$. By the corollary, $\{\vec{v}_{r+1}, \cdots, \vec{v}_p\} \subseteq \text{null}(A)$

$$\Rightarrow AV = (A\vec{v}_1 \ \ A\vec{v}_2 \ \cdots \ A\vec{v}_r \ \ A\vec{v}_{r+1} \ \cdots \ A\vec{v}_p)$$

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

$$U\Sigma = (\vec{u}_1 \ \vec{u}_2 \ \cdots \ \vec{u}_r \ \vec{u}_{r+1} \ \cdots \ \vec{u}_n)_{n\times n} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}_{n\times p}$$

$$= (\sigma_1\vec{u}_1 \ \ \sigma_2\vec{u}_2 \ \cdots \ \sigma_r\vec{u}_r, \ \underbrace{\vec{0} \ \cdots \ \vec{0}}_{p-r \text{ columns}})_{n\times p}$$

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

$$= (A\vec{v}_1 \ \ A\vec{v}_2 \ \cdots \ A\vec{v}_r \ \ \vec{0} \ \cdots \ \vec{0})$$

$$= AV. \qquad \blacksquare$$

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

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

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

Then $A = U_r D V_r^*$

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