B24 Aug 6 Lec 2 Notes

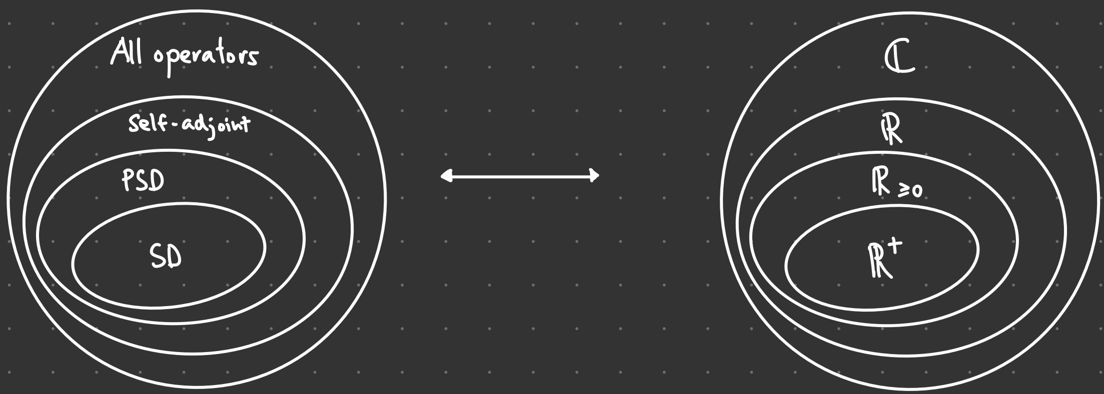
Definition:

Let $A : X \to Y$ be a L.T. The Hermitan square of A is defined by

$$A^* A : X \to X$$

By the last result from last class, A^*A has a unique PSD "square root", which we denote by $|A| = \sqrt{A^*A}$ and we call this the modulus of A.

Useful Analogy:



Adjoint	$\longleftrightarrow$	complex conjugate
Hermitan square of A is A^*A	$\longleftrightarrow$	$z \mapsto \bar{z}z = \lvert z\rvert^2 \in \mathbb{R}_{\geq 0}$
$\lvert A\rvert = \sqrt{A^*A}$	$\longleftrightarrow$	$\lvert z\rvert = \sqrt{\bar{z}z} = \sqrt{\lvert z\rvert^2}$
Unitary operators	$\longleftrightarrow$	$\{ z \in \mathbb{C} \mid \lvert z\rvert = 1 \}$

Proposition:

If $A : X \to Y$ is a L.T., then:

$$\| \, |A| x \, \| = \| Ax \| , \quad \forall x \in X$$

Proof:

$$\begin{aligned} \| \, |A| x \, \|^2 &= \langle |A|x, |A|x \rangle \\ &= \langle |A|^*|A| x, x \rangle \\ &= \langle |A|^2 x, x \rangle \\ &= \langle A^*Ax, x \rangle \\ &= \langle Ax, Ax \rangle \\ &= \|Ax\|^2 \end{aligned}$$

$|A|$ is PSD $\Rightarrow$ $|A|$ is self-adjoint

$|A| = \sqrt{A^*A}$

$\blacksquare$

Corollary:

If $A: X \to Y$ is a L.T., then:

$$\operatorname{Ker} A = \operatorname{Ker} |A| = (\operatorname{ran} |A|)^{\perp}$$

Proof:

We have

$$\| \, |A| x \, \| = \| Ax \| \quad , \forall x \in X$$

i.e. $Ax = 0$ iff $\|Ax\| = 0$
iff $\| \, |A| x \, \| = 0$
iff $|A| x = 0$

So $\operatorname{Ker}(A) = \operatorname{Ker}(|A|)$

Moreover, for any L.T. $T: X \to Y$, recall $\operatorname{Ker}(T) = (\operatorname{ran}(T^*))^{\perp}$

So

$$\begin{aligned} \operatorname{Ker}(|A|) &= (\operatorname{ran}(|A|^*))^{\perp} \\ &= (\operatorname{ran}(|A|))^{\perp} \end{aligned}$$

Since $|A|$ is self-adjoint

$\boxtimes$

Theorem: Polar Decomposition

Let $A: X \to X$ be a L.T.. Then there exists a unitary $U: X \to X$ s.t.

$$A = \underbrace{U}_{\text{unitary}} \underbrace{|A|}_{\text{PSD}}$$

Proof:

We first define $U_0 : \operatorname{ran}(|A|) \to X$ as follows

If $x \in \operatorname{ran}(|A|)$, then $x = |A| v$ for some $v \in X$, and we define $U_0 x := Av$

In order to prove that U_0 is well-defined consider $v' \in X$ s.t. $|A| v' = x$

Definition:

Let $x \in \operatorname{ran}(A)$. Then define $U_0 x := Av$, where $v \in |A|^{-1}(x)$.

Then $|A|(v - v') = |A| v - |A| v' = x - x = 0$

i.e. $v - v' \in \operatorname{Ker}(|A|) = \operatorname{Ker}(A)$

by previous corollary

So $v = v' + v''$ where $v'' \in \operatorname{Ker}(A)$

Proof (continued...):

So $Av = A(v' + v'') = Av' + \cancel{Av''}^{\,0}$

and $v - v' \in \ker(A) \Rightarrow Av = Av'$

i.e. U_0 is well-defined

Moreover, U_0 is linear, and:

$$\|U_0 x\| = \|Av\| \qquad v \text{ is s.t. } |A|v = x$$

$$= \|\,|A|v\| \qquad \text{By previous proposition}$$

$$= \|x\| \quad , \forall x \in \operatorname{ran}|A|$$

So U_0 is isometric, and we have the formula:

$$U_0 |A| v = Av \ , \ \forall v \in X$$

It remains to extend the definition of U_0 to all of X

Recall $X = \operatorname{ran}|A| \oplus (\operatorname{ran}|A|)^{\perp}$

i.e for any $x \in X$ there exists unique $v_1 \in \operatorname{ran}|A|$, $v_2 \in (\operatorname{ran}|A|)^{\perp}$ s.t.

$$x = v_1 + v_2$$

Moreover, $(\operatorname{ran}|A|)^{\perp} = \operatorname{Ker}|A|^*$ fundamental subspace result

$= \operatorname{Ker}|A|$ $\quad |A|$ is self-adjoint

$= \operatorname{Ker} A$ $\quad$ Previous corollary

Define $U_1 : \operatorname{Ker} A \to (\operatorname{ran}(A))^{\perp}$ as any unitary map. we can do this since

$$\dim(\operatorname{Ker}(A)) = \dim((\operatorname{Ran}(A))^{\perp})$$

since

$$\dim(\operatorname{Ker}(A)) + \cancel{\dim(\operatorname{ran}(A))} = \dim(X)$$
$$= \dim((\operatorname{ran}(A))^{\perp}) + \cancel{\dim(\operatorname{ran}(A))}$$

We define

$$U : X = \operatorname{ran}|A| \oplus (\operatorname{ran}|A|)^{\perp} \longrightarrow X \quad \text{by} \quad x = \underbrace{v_1}_{v_1 \in \operatorname{ran}|A|} + \underbrace{v_2}_{v_2 \in (\operatorname{ran}|A|)^{\perp}} \mapsto U_0(v_1) + U_1(v_2)$$

Proof (continued...):

Now we check that U is an isometry:

$$\| Ux \|^2 = \| U_0 v_1 + U_1 v_2 \|^2$$

$$= \| U_0 v_1 \|^2 + \| U_1 v_2 \|^2$$ Since $U_0 v_1 \in \operatorname{ran}(A)$ and $U_1 v_2 \in (\operatorname{ran}(A))^{\perp}$ i.e. $U_0 v_1 \perp U_1 v_2$

$$= \| v_1 \|^2 + \| v_2 \|^2$$ Since U_0 and U_1 are isometries

$$= \| v_1 + v_2 \|^2$$ $v_1 \perp v_2$

$$= \| x \|^2 , \quad \forall x \in X$$

i.e. U is an isometry, hence unitary since any isometry from one space to another with the same dimension is unitary ☒

Definition:

If $A: X \to Y$ is a L.T., then the singular values of A are defined as the eigenvalues of $|A|$.

Let $A: X \to Y$ be a L.T., and denote by $\sigma_1, \dots, \sigma_n$ the singular values of A, where $\sigma_1, \dots, \sigma_n$ are non-zero, and $\sigma_{r+1} = \dots = \sigma_n = 0$

Let $v_1, \dots, v_n$ be an orthonormal basis of eigenvectors for A^*A (recall that A^*A is self-adjoint, and:

> Theorem: B24 July 30 Lec 2 Notes
>
> Let $A: X \to X$ be self-adjoint. Then A has $\dim(X)$ many eigenvalues (counting multiplicity), all eigenvalues are real, and there exists an orthonormal basis for X consisting of eigenvectors for A.

So

$$A^*A v_k = |A|^2 v_k$$ $|A| = \sqrt{A^*A}$

$$= |A||A| v_k$$

$$= |A| \sigma_k v_k$$

$$= \sigma_k^2 v_k , \quad \text{for } 1 \le k \le n$$

Let $w_k := \frac{1}{\sigma_k} A v_k$, for $1 \le k \le r$

Proposition:

$w_1, \dots, w_r$ is an orthonormal system in Y.

Proof:

$$\langle w_k, w_j \rangle = \left\langle \frac{1}{\sigma_k} A v_k , \frac{1}{\sigma_j} A v_j \right\rangle$$

$$= \frac{1}{\sigma_k \sigma_j} \langle A v_k , A v_j \rangle$$

$\sigma_1, \dots, \sigma_n$ are e.v. of the self-adjoint L.T. $|A|$, hence all are real.

$$= \frac{1}{\sigma_k \sigma_j} \langle A^* A v_k , v_j \rangle$$

$$= \frac{1}{\sigma_k \sigma_j} \langle |A|^2 v_k , v_j \rangle$$

$$= \frac{1}{\sigma_k \sigma_j} \langle \sigma_k^2 v_k , v_j \rangle$$

$$= \frac{\sigma_k^2}{\sigma_k \sigma_j} \langle v_k , v_j \rangle = \begin{cases} 1 & \text{if } k=j \\ 0 & \text{if } k \neq j \end{cases}$$

☑

Proposition: Schmidt decomposition

For any $x \in X$, we have

$$\bigstar \quad Ax = \sum_{k=1}^{r} \sigma_k \langle x, v_k \rangle w_k$$

Proof:

The function

$$x \mapsto \sum_{k=1}^{r} \sigma_k \langle x, v_k \rangle w_k$$

is a L.T., and hence in order to prove $\bigstar$, it suffices to show

$$A v_j = \sum_{k=1}^{r} \sigma_k \langle v_j, v_k \rangle w_k , \quad \text{for } 1 \leq j \leq n$$

If $j \leq r$ (i.e. if $\sigma_j \neq 0$), then:

$$\sum_{k=1}^{r} \sigma_k \langle v_j, v_k \rangle w_k = \sigma_j \langle v_j , v_j \rangle w_j$$

$v_1, \dots, v_n$ are orthonormal

$$= \cancel{\sigma_j} \underbrace{\langle v_j , v_j \rangle}_{=1} \frac{1}{\cancel{\sigma_j}} A v_j$$

$$= A v_j$$

If $j > r$ (i.e. if $\sigma_j = 0$), then:

$$\sum_{k=1}^{r} \sigma_k \langle v_j , v_k \rangle w_k = 0$$

We claim $A v_j = 0$. Indeed $A^* A v_j = \sigma_j^2 v_j = 0 v_j = 0$ ☑

Theorem: Cayley - Hamilton Theorem

Let A be a square matrix, and $p(\lambda) : \det(A - \lambda I)$ the characteristic polynomial of A. Then:

$$P(A) = 0$$

Ex 1:

$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$. Then

$$\begin{aligned} p(\lambda) = \det(A - \lambda I) &= (1-\lambda)(3-\lambda) \\ &= 3\lambda^0 - 4\lambda^1 + \lambda^2 \end{aligned}$$

$$P(A) = 3\underbrace{A^0}_{= I \text{ by convention}} - 4A^1 + A^2$$

$$= \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} - 4\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}^2$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$