B24 Aug 4 Lec 1 Notes

Definition:

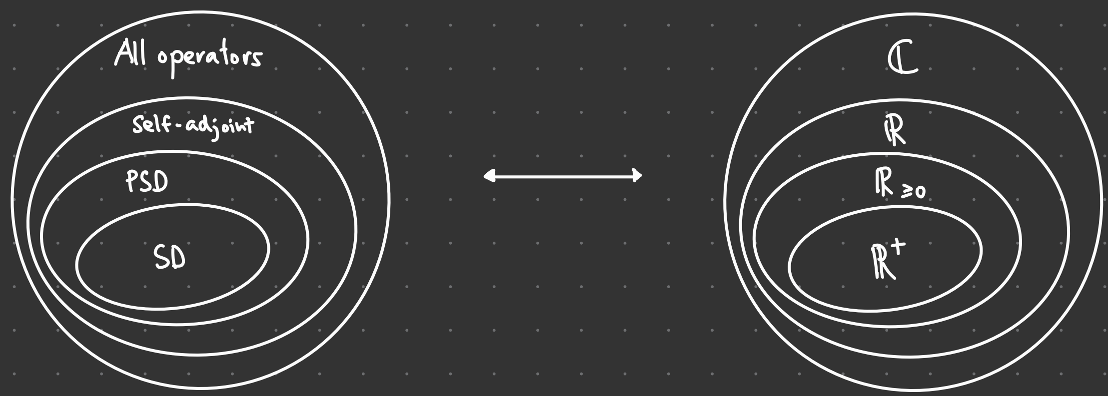
A self-adjoint L.T. $A: X \to X$ is called positive definite if

$$\langle Ax, x \rangle > 0 \ , \ \forall x \in X, \quad x \neq 0$$

and $A: X \to X$ is called positive semidefinite if

$$\langle Ax, x \rangle \geq 0 \ , \ \forall x \in X$$

Useful Analogy:



Adjoint　　　　　　complex conjugate

Definition:

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

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

Remark:

A^*A is self-adjoint

$$(A^*A)^* = A^* (A^*)^* = A^* A \ ,$$

and moreover A^*A is PSD:

$$\langle A^*Ax, x \rangle = \langle Ax, Ax \rangle = \|Ax\|^2 \geq 0 \quad , \ \forall x \in X$$

We will now prove that therefore A^*A has a "square root".

Theorem:

Let $A : X \to X$ be self-adjoint. Then:

(i) A is PD iff all eigenvalues of A are positive.
(ii) A is PSD iff all eigenvalues of A are non-negative.

Proof:

By:

> 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.

there exists an orthonormal basis

$$u_1, \dots, u_n \text{ for } X$$

consisting of eigenvectors for A, with eigenvalues $\lambda_1, \dots, \lambda_n$.

Assume A is PD. Then

$$0 < \langle A u_i, u_i \rangle = \langle \lambda_i u_i, u_i \rangle = \lambda_i \langle u_i, u_i \rangle$$
$$= \lambda_i \|u_i\|^2$$

So $\lambda_i \|u_i\|^2 > 0 \Rightarrow \lambda_i > 0$

Now assume $\lambda_1, \dots, \lambda_n$ are all positive. Let $x \in X \setminus \{0\}$. Then $x = \sum_{i=1}^n \alpha_i u_i$. So:

$$\begin{aligned}
\langle Ax, x \rangle &= \left\langle A\left(\sum_{i=1}^n \alpha_i u_i\right), \sum_{i=1}^n \alpha_i u_i \right\rangle \\
&= \left\langle \sum_{i=1}^n \alpha_i A(u_i), \sum_{i=1}^n \alpha_i u_i \right\rangle \quad \text{Since } A \text{ is linear} \\
&= \left\langle \sum_{i=1}^n \alpha_i \lambda_i u_i, \sum_{i=1}^n \alpha_i u_i \right\rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n \langle \alpha_i \lambda_i u_i, \alpha_j u_j \rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \lambda_i \overline{\alpha_j} \underbrace{\langle u_i, u_j \rangle}_{= \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}} \\
&= \sum_{i=1}^n \alpha_i \lambda_i \overline{\alpha_i} \langle u_i, u_i \rangle \\
&= \sum_{i=1}^n |\alpha_i|^2 \lambda_i \|u_i\|^2 \\
&> 0 \quad \text{since } \lambda_i > 0 \text{ and at least one } \alpha_i > 0
\end{aligned}$$

So A is PD. This proves (i). Proof of (ii) is similar. ☑

Corollary:

Let $A: X \to X$ be PSD. Then there exists a unique PSD $: X \to X$ s.t.

$$B^2 := BB = A$$

Proof:

As in the previous proof, there exists an orthogonal basis

$$u_1, \dots, u_n \text{ for } X$$

Consisting of eigenvectors for A, with eigenvalues $\lambda_1, \dots, \lambda_n$

Moreover, since A is PSD, we know by the previous result that $\lambda_1 \geq 0, \dots, \lambda_n \geq 0$.

So

$$[A]^{u_1, \dots, u_n}_{u_1, \dots, u_n} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

Define $B: X \to X$ by

$$[B]^{u_1, \dots, u_n}_{u_1, \dots, u_n} = \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{bmatrix}$$

Then B is a PSD by the previous result, since the eigenvalues of B are $\sqrt{\lambda_1} \geq 0, \dots, \sqrt{\lambda_n} \geq 0$

Furthermore,

$$\begin{aligned} [B^2]^{u_1, \dots, u_n}_{u_1, \dots, u_n} &= [B]^{u_1, \dots, u_n}_{u_1, \dots, u_n} [B]^{u_1, \dots, u_n}_{u_1, \dots, u_n} \\ &= \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \\ &= [A]^{u_1, \dots, u_n}_{u_1, \dots, u_n} \end{aligned}$$

$\Rightarrow B^2 = A$

Now suppose that $C: X \to X$ is PSD s.t.

$$C^2 = A$$

Proof (continued...):

We need to show $C = B$.

Because C is PSD, there exists an orthonormal basis

$$v_1, \dots, v_n \text{ for } X$$

consisting of eigenvalues for C, with eigenvalues $\mu_1, \dots, \mu_n$.

So

$$[C]_{v_1,\dots,v_n}^{v_1,\dots,v_n} = \begin{bmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{bmatrix}$$

Since $C^2 = A$,

$$\begin{aligned}
[C^2]_{v_1,\dots,v_n}^{v_1,\dots,v_n} &= [C]_{v_1,\dots,v_n}^{v_1,\dots,v_n} [C]_{v_1,\dots,v_n}^{v_1,\dots,v_n} \\
&= \begin{bmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{bmatrix} \begin{bmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{bmatrix} \\
&= \begin{bmatrix} \mu_1^2 & & 0 \\ & \ddots & \\ 0 & & \mu_n^2 \end{bmatrix} \\
&= [A]_{v_1,\dots,v_n}^{v_1,\dots,v_n}
\end{aligned}$$

So $\mu_1^2, \dots, \mu_n^2$ are the eigenvalues of A, so up to relabelling $\mu_i^2 = \lambda_i$ for $1 \le i \le n$,

i.e. $\sqrt{\lambda_i} = \mu_i$ for $1 \le i \le n$ and $Av_i = \mu_i^2 v_i$ for $1 \le i \le n$

So

$$\begin{aligned}
[A]_{v_1,\dots,v_n}^{v_1,\dots,v_n} &= \begin{bmatrix} \mu_1^2 & & 0 \\ & \ddots & \\ 0 & & \mu_n^2 \end{bmatrix} \\
&= [B^2]_{v_1,\dots,v_n}^{v_1,\dots,v_n} = [C^2]_{v_1,\dots,v_n}^{v_1,\dots,v_n}
\end{aligned}$$

and

$$\begin{aligned}
[C]_{v_1,\dots,v_n}^{v_1,\dots,v_n} &= \begin{bmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{bmatrix} \\
&= [B]_{v_1,\dots,v_n}^{v_1,\dots,v_n}
\end{aligned}$$

$\Rightarrow C = B$ ▨