Topic#17

Normal operator & Self-adjoint operator

Goal: Recall that for $A \in M_{n\times n}(F)$ ($F = \mathbb{C}$ or $\mathbb{R}$),

$$A \text{ is normal} \Longleftrightarrow AA^* = A^*A.$$

1°. Define a normal operator $T \in \mathcal{L}(V)$?

2°. Characterize a normal operator $T \in \mathcal{L}(V)$?

3°. A self-adjoint matrix (i.e. $A = A^*$) is normal. Can we do a similar extension as well as its characterization?

Other terminology: A complex self-adjoint matrix is also usually called a **Hermitian** matrix. Hermitian matrices can be understood as the complex extension of real symmetric matrices.

Throughout this topic, we always let $T \in \mathcal{L}(V)$, where V is an i.p.s. (dim can be finite or infinite). Assume that $T^* \in \mathcal{L}(V)$ exists.

Def.

T is **normal** if $TT^* = T^*T$.

T is **self-adjoint** if $T = T^*$.

1^{st} goal is to show:

Theorem. Let $T \in \mathcal{L}(V)$, where V is a complex i.p.s. with $dim(V) < \infty$. Then T is normal **iff** $\exists$ an orthonormal basis for V consisting of eigenvectors of T.

We divide the proof by a few steps.

Step 1. Proof of " $\Leftarrow$ ":

Let $n = \dim(V)$ and $\beta = \{v_1, ..., v_n\}$ be an orthonormal basis for V of eigenvectors of T, with

$$T(v_i) = \lambda_i v_i, \quad \lambda_i \in \mathbb{C}, 1 \le i \le n.$$

Then, $[T]_\beta = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ is diagonal, and hence $[T^*]_\beta = ([T]_\beta)^* = \mathrm{diag}(\overline{\lambda}_1, ..., \overline{\lambda}_n)$ is also diagonal. Note: $\lambda_i \overline{\lambda_i} = |\lambda_i|^2$, then

$$[TT^*]_\beta = [T]_\beta [T^*]_\beta = \begin{pmatrix} |\lambda_1|^2 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & |\lambda_n|^2 \end{pmatrix} = [T^*]_\beta [T]_\beta = [T^* T]_\beta.$$

So, it follows $[TT^*]_\beta = [T^* T]_\beta$. One then has $TT^* = T^* T$. □

Remark: "$\Leftarrow$" is also true if V is a finite-dim real i.p.s.

But, the converse statement "$\Rightarrow$" may not be true in the following cases:

(a) V is a finite-dim real i.p.s.

(b) V is an infinite-dim complex i.p.s.

Counterexample to treat case

(a) V is a finite-dim real i.p.s.:

In the previous lecture we showed that the rotation $T_{\pi/2} \in \mathcal{L}(\mathbb{R}^2)$ has no eigenvector. But,

$$T_{\pi/2} = L_A, \quad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, T^*_{\pi/2} = L_{A^*}, \quad A^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Note: $AA^* = I_2 = A^*A$ (Exercise), $\therefore T_{\pi/2}T^*_{\pi/2} = T^*_{\pi/2}T_{\pi/2}$

$\therefore T_{\pi/2}$ is normal. But $T_{\pi/2}$ has no eigenvetor. □

Counterexample to treat case

(b): V is an infinite-dim complex i.p.s.

Recall: $H =$ set of continuous complex-valued functions on $[0, 2\pi]$.

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}dt$$

$S = \{f_n : n = 0, \pm 1, \dots\}$ with $f_n \overset{def}{=} e^{int}$ is orthonormal.

$$V \overset{def}{=} span(S) \text{ is an infinite-dim complex i.p.s.}$$

Claim. $\exists$ a normal operator $T \in \mathcal{L}(V)$ which has no eigenvector.

Pf. Def $T, U \in \mathcal{L}(V)$ as $T(f) \overset{def}{=} f_1 f, \quad U(f) \overset{def}{=} f_{-1} f$.
Then, $T(f_n) = f_{n+1}, \quad U(f_n) = f_{n-1}, \quad n = 0, \pm 1, \dots$.
Thus, $\langle T(f_m), f_n\rangle = \langle f_{m+1}, f_n\rangle = \delta_{m+1,n} = \delta_{m,n-1}$
$= \langle f_m, f_{n-1}\rangle = \langle f_m, U(f_n)\rangle$
$\therefore$ $T^* = U$ exists (think about why),
and $TT_* = TU = I = UT = T^*T$, i.e. T is normal. $\square$

But T has no eigenvectors.

Otherwise, let $f \in V$ be an eigenvector of T, i.e. $T(f) = \lambda f$ for some $\lambda \in \mathbb{C}$. As $V = span(S)$, we may write

$$f = \sum_{i=n}^{m} a_i f_i, \quad a_m \neq 0, \quad n \leq m.$$

Thus,

$$T(f) \overset{T \in \mathcal{L}}{=} \sum_{i=n}^{m} a_i T(f_i) = \sum_{i=n}^{m} a_i f_{i+1} = \lambda f = \sum_{i=n}^{m} \lambda a_i f_i.$$

By this identity and $a_m \neq 0$, we see

f_{m+1} is a linear combination of $f_n, f_{n+1}, \ldots, f_m$,

which is a contradiction with the fact that S is l. indep. □□

Step 2. To show " $\Rightarrow$ ", we need to make two preparations.

In this step, we make the 1^{st} preparation.

Note: V can be either complex or real i.p.s.

Thm (Schur Lemma). Let $T \in \mathcal{L}(V)$, where V is a finite-dim i.p.s. Aussume further that the c.p. of T splits over $\mathbb{F}$. Then, $\exists$ an orthonormal o.b. β for V such that $[T]_\beta$ is upper triangular.

Proof of Theorem. As a preparation, we need to

Claim. Let $T \in \mathcal{L}(V)$ for a finite-dim i.p.s. V. If T has an e.v., then so does T^*.

Proof of Claim. Let $T(v) = \lambda v$, $0 \neq v \in V$, $\lambda \in \mathbb{C}$.
Then, $\forall\, x \in V$,

$$\begin{aligned} 0 = \langle 0, x\rangle &= \langle (T - \lambda I)v, x\rangle \\ &= \langle v, (T - \lambda I)^*(x)\rangle \\ &= \langle v, (T^* - \overline{\lambda} I)(x)\rangle, \quad \therefore\ v \perp R(T^* - \overline{\lambda} I). \end{aligned}$$

As $v \neq 0$, $R(T^* - \overline{\lambda} I) \neq V$.
$\therefore$ $T^* - \overline{\lambda} I$ is not onto and hence not one-to-one.
$\therefore$ $N(T^* - \overline{\lambda} I)$ contains at least one nonzero vextor, call it u.
$(T^* - \overline{\lambda})(u) = 0$ i.e. $T^*(u) = \overline{\lambda} u$. $0 \neq u \in V$
$\therefore$ u is an eigenvector of T^* associated with $\overline{\lambda}$. □

We continue: Induction in $n \overset{def}{=} dim(V)$.

$n = 1$: true obviously.

Assume "true" for $n - 1 (n \geq 2)$, to show "true" for n,
i.e., let $T \in \mathcal{L}(V)$ split with $dim(V) = n$, to find the desired β.

As T splits, T has an eigenvector, so T^* also has an eigenvector by the previous claim. Let $T^*(z) = \lambda z$ for some unit eigenvector z and for some $\lambda \in \mathbb{F}$. Set $W = span(\{z\})$.

Claim. $W^\perp$ is T-invariant.
Proof of claim. Let $y \in W^\perp$, to show $T(y) \in W^\perp$, i.e. to show

$$\langle T(y), x \rangle = 0, \forall x \in W.$$

Take $x = cz \in W$, then

$$\begin{aligned} \langle T(y), x \rangle = \langle T(y), cz \rangle = \langle y, T^*(cz) \rangle = \langle y, cT^*(z) \rangle \\ = \langle y, c\lambda z \rangle = \overline{c\lambda} \langle y, z \rangle = 0. \end{aligned}$$

□

By this claim,

$T_{W^\perp} \in \mathcal{L}(W^\perp)$ is well-defined and c.p. of $T_{W^\perp}$ divides c.p. of T.

As T splits, so does $T_{W^\perp}$. So, $T_{W^\perp} \in \mathcal{L}(W^\perp)$ splits, where $W^\perp$ is an $(n-1)$-dim i.p.s. for $V = W \bigoplus W^\perp$ where $\dim W = 1$.

Induction assumption implies that

$\exists$ an orthonormal basis γ for $W^\perp$ s.t. $[T_{W^\perp}]_\gamma$ is upper triangular.

then we see

$$\beta \overset{def}{=} \gamma \cup \{z\} \text{ is an orthonormal basis for } V$$

s.t. $[T]_\beta = \begin{pmatrix} \textit{an upper} & * \\ \textit{triangular matrix} & \vdots \\ 0 \cdots 0 & * \end{pmatrix}$ is upper triangular.

Note: The 1st to the $(n-1)$th entries in the last row are zeros because each entry corresponds to the nth component of β-coordinates of each basis vector in γ acted by T.

Step 3: We make the 2^{nd} preparation.

Note: Below V can be either complex or real i.p.s. and it can be either finite-dim or ∞-dim.

Theorem. Let $T \in \mathcal{L}(V)$ be normal for an i.p.s. V. Then,

(a) $\|T(x)\| = \|T^*(x)\|, \quad \forall x \in V$.

(b) $T - cI$ is normal for any $c \in F$.

(c) If $x \neq 0$ is a λ-e.v. of T, then x is also a $\overline{\lambda}$-e.v. of T^*.

(d) Two e-vectors associated with two distinct e-values of T must be orthogonal.

Proof.

(a) Let $x \in V$,

$$\|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, T^*T(x)\rangle = \langle x, TT^*(x)\rangle$$
$$= \langle T^*(x), T^*(x)\rangle = \|T^*(x)\|^2.$$

(b) Let $c \in F$, check

$(T - cI)^*(T - cI) = (T^* - \bar{c}I)(T - cI) \overset{ok}{=} (T - cI)(T - cI)^*$.

Exercise: Use $(T - cI)^* = T^* - \bar{c}I$, and $TT^* = T^*T$.

(c) Let $T(x) = \lambda x$, $0 \neq x \in V$, i.e. $(T - \lambda I)(x) = 0$.

Note: $T - \lambda I$ is also normal, then

$0 = \|(T - \lambda I)(x)\| = \|(T - \lambda I)^*(x)\| \overset{(a)(b)}{=} \|(T^* - \overline{\lambda}I)(x)\|$.

$\therefore (T^* - \overline{\lambda}I)(x) = 0$, i.e. $T^*(x) = \overline{\lambda}x$, $0 \neq x \in V$.

(d) Let $T(x_1) = \lambda_1 x_1, T(x_2) = \lambda_2 x_2, x_1 \neq 0, x_2 \neq 0, \lambda_1 \neq \lambda_2$.

By (c), $T^*(x_2) = \overline{\lambda_2}x_2$. Then

$$\lambda_1\langle x_1, x_2\rangle = \langle \lambda_1 x_1, x_2\rangle = \langle T(x_1), x_2\rangle = \langle x_1, T^*(x_2)\rangle$$
$$= \langle x_1, \overline{\lambda_2}x_2\rangle = \lambda_2\langle x_1, x_2\rangle$$

$\because \lambda_1 \neq \lambda_2$ and $(\lambda_1 - \lambda_2)\langle x_1, x_2\rangle = 0 \therefore \langle x_1, x_2\rangle = 0$. □

Step 4: This last step is to give the proof of "$\Rightarrow$":
Assume: T is normal.
$\because F = \mathbb{C}$, $\therefore$ the c.p. of T splits,
then by Schur's lemma,

$\exists$ an orthonormal basis β such that $[T]_\beta$ is upper triangular.

Set $\beta = \{v_1, \ldots, v_n\}$, and $A = [T]_\beta$.

Claim. All vectors in β are eigenvectors of T.

Proof of claim.

1^{st} column, $[T(v_1)]_\beta = 1^{st}$ column of A. For A is upper triangular, $T(v_1) = A_{11}v_1 + 0v_2 + \cdots + 0v_n = A_{11}v_1$. So, $v_1 \neq 0$ is an e-vector of T with e-value A_{11}.

2^{nd} column: $[T(v_2)]_\beta = 2^{nd}$ column of A. Keep in mind, to show $A_{21} = 0$. $\because T(v_2) = A_{21}v_1 + A22v_2$ and $\|v_1\| = 1$ and $\langle v_2, v_1\rangle = 0$ $\therefore \langle T(v_2), v_1\rangle = \langle A_{21}v1 + A_{22}v_2, v_1\rangle = A_{21}\langle v_1, v_1\rangle = A_{21}$

On the other hand,
$LHS = \langle T(v_2), v_1\rangle = \langle v_2, T^*(v_1)\rangle = \langle v_2, \overline{A_{11}}v_1\rangle = A_{11}\langle v_2, v_1\rangle = 0$
$\therefore A_{21} = 0$.

Similarly, 3^{rd} column: one can shows $A_{31} = A_{32} = 0 \cdots$. Remark: you may use induction argument to show: $A_{ij} = 0$, $i > j$ (Exercise).
$\therefore$ the upper-triangular matrix A becomes diagonal!

□

2^{nd} goal:

Theorem. Let $T \in \mathcal{L}(V)$, where V is a real i.p.s. with $dim(V) < \infty$. Then, T is self-adjoint **iff** $\exists$ an orthonormal basis β for V consisting of e-vectors of T.

Proof of " $\Leftarrow$ " :

Assume:

$\exists$ an orthonormal basis β for V consisting of e-vectors of T.

Then $[T]_\beta$ is a diagonal real matrix, thus $[T]_\beta$ is real symmetric and hence self-adjoint, so T is self-adjoint.

$(\because [T - T^*]_\beta = [T]_\beta - [T^*]_\beta = [T]_\beta - [T]_\beta{}^* = [T]_\beta - [T]_\beta = 0)$

To show " $\Rightarrow$ ", we need a

Lemma. Let $T \in \mathcal{L}(V)$ be self-adjoint, where V is a finite-dim i.p.s. (either complex or real). Then

(a) Any eigenvalue of T is real.

(b) If $F = \mathbb{R}$, then the c.p. of T splits over R.

Proof of lemma.

(a) Let $T(x) = \lambda(x)$, $x \neq 0$, $\lambda \in F$. Then

$$\lambda x = T(x) \overset{(T=T^*)}{=} T^*(x) = \bar{\lambda} x.$$

$\because$ x is a nonzero vector, and $(\lambda - \bar{\lambda})x = 0$ $\therefore \lambda = \bar{\lambda}$, i.e. λ is real. □

(b) Let $n = dim(V), F = \mathbb{R}$.
Let β be an orthonormal basis for V and $A = [T]_\beta$.
Note: A is self-adjoint (indeed, real symmetric).
Also note: $L_A \in \mathcal{L}(\mathbb{C}^n)$ is self-adjoint
($\because [L_A]_\gamma = A$ for the s.o.b. orthonormal γ for $\mathbb{C}^n$).
Note: Fundamental theorem of algebra tells: the c.p. of L_A

$$= det(L_A - tI) = (t - \lambda_1)^{m_1} \cdots (t - \lambda_k)^{m_k} \text{ each } \lambda_i \in \mathbb{C}.$$

By (a), each λ_i is real.
$\therefore \lambda_1, \ldots \lambda_n \in \mathbb{R}$, it means that the c.p. of L_A splits over $\mathbb{R}$.
Note: $T \& L_A$ have the same c.p.
$\therefore$ the c.p. of T splits over $F = \mathbb{R}$. □□

Proof of " $\Rightarrow$ " in thm.

Assume: T is self-adjoint. As $F = \mathbb{R}$, the previous lemma tells the c.p. of T splits. Apply the Schur's theorem, then $\exists$ an orthonormal basis β for V such that $[T]_\beta$ is upper triangular. Note:

$$([T]_\beta)^* = [T^*]_\beta \overset{(T^*=T)}{=} [T]_\beta,$$

i.e. $[T]_\beta$ is real symmetric, but it is also upper triangular, hence $[T]_\beta$ is real diagonal.

$\therefore$ all vector in β must be eigenvectors of T. □□

Last remark: for $A \in M_{n\times n}\mathbb{F}$

If A is real-symmetric, then A is self-adjoint and hence normal.

But, if A is complex-symmetric, then A may NOT be self-adjoint and A may NOT be normal.

Example:

$$A = \begin{pmatrix} i & i \\ i & 1 \end{pmatrix} \in M_{2\times 2}(\mathbb{C})$$

$\because A^t = A$

$\therefore$ A is complex symmetric.

Note $A^* = \overline{A}^t = \begin{pmatrix} -i & -i \\ -i & 1 \end{pmatrix} \neq A$ then A is NOT self-adjoint.

$$AA^* = \begin{pmatrix} i & i \\ i & 1 \end{pmatrix}\begin{pmatrix} -i & -i \\ -i & 1 \end{pmatrix} = \begin{pmatrix} -i^2 - i^2 & -i^2 + i \\ -i^2 - i & -i^2 + 1 \end{pmatrix} = \begin{pmatrix} 2 & 1+i \\ 1-i & 2 \end{pmatrix},$$

$$A^*A = \begin{pmatrix} -i & -i \\ -i & 1 \end{pmatrix}\begin{pmatrix} i & i \\ i & 1 \end{pmatrix} = \begin{pmatrix} -i^2 - i^2 & -i^2 - i \\ -i^2 + i & -i^2 + 1 \end{pmatrix} = \begin{pmatrix} 2 & 1-i \\ 1+i & 2 \end{pmatrix}.$$

$\therefore AA^* \neq A^*A$, i.e. A is **NOT** normal. □