Week 8 : Normal and Self Adjoint Operators (textbook §6.4)

Review of Diagonalizability

Let V be a finite dimensional vector space (over $\mathbb{F}$).

Recall the definition of diagonalizability :

(operator form) : $T : V \to V$ linear operator is diagonalizable $\iff$ there exists an "eigenbasis" for V (i.e. a basis consisting of eigenvectors)

||

(matrix form) : $A \in M_{n\times n}(\mathbb{F})$ is diagonalizable $\iff$ there exists an invertible $Q \in M_{n\times n}(\mathbb{F})$ s.t $Q^{-1}AQ$ is a diagonal matrix.

Characterization of diagonalizability

A matrix $A \in M_{n\times n}(\mathbb{F})$ (or an operator $T : V \to V$) is diagonalizable

if and only if (i) The characteristic polynomial $f(t)$ splits over $\mathbb{F}$

i.e. $f(t) = (-1)^n (t-\lambda_1)^{m_1} \cdots (t-\lambda_k)^{m_k}$

AND (ii) $\boxed{m_i = \dim_{\mathbb{F}} E_{\lambda_i}}$ for each $i=1,\dots,k$.

The characterization above is our last resort, since we have to do ALL the calculations to check the conditions!

Fortunately, sometimes we can be a bit lazy......

1st Sufficient Test : there exist n distinct eigenvalues $\Rightarrow$ diagonalizable.

2nd Sufficient Test : Symmetric real matrices $\Rightarrow$ diagonalizable.

$\boxed{A = A^t}$

Question: (1) Why is 2^{nd} Sufficient test true?

(2) What about complex matrices?

Ans: "Spectral Theorems"!!

$\boxed{\text{Diagonalizability and } \langle , \rangle}$

In "2^{nd} Sufficient Test", we need to take transpose of a matrix.

Remember that taking (conjugate) transpose of a matrix is essentially the same as taking the adjoint of a linear operator:

$$\boxed{[T^*]_\beta = [T]_\beta^*}$$

β : orthonormal basis

When $\mathbb{F} = \mathbb{R}$, it is simply the transpose!

$[T]_\beta$ is a real symmetric matrix $\iff T^* = T$

From this, we can restate "2^{nd} Sufficient Test" in operator form:

Real Spectral Theorem:

Let $(V, \langle , \rangle)$ be a finite dimensional inner product space over $\mathbb{R}$.

Suppose $T : V \to V$ is a linear operator.

$\boxed{T^* = T} \iff$ there exists an orthonormal eigenbasis for V.

Remark: In fact this says a lot more than the "2^{nd} Sufficient Test". This is an "if and only if" statement, but we ask more - we need the eigenbasis to be orthonormal as well!

This is a natural requirement. Remember that we always prefer orthonormal basis to just a general basis whenever we work with inner product space $(V, \langle , \rangle)$.

Given a linear operator $T: V \to V$ on a finite dimensional inner product space $(V, \langle , \rangle)$ over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$,

<u>Question 1</u>: $\exists$ eigenbasis ? (ie is T diagonalizable ?)

<u>Question 2</u>: $\exists$ orthonormal eigenbasis ?

We already knew that Question 1 is <u>rather subtle</u>.

With $\langle , \rangle$, we are in fact asking for <u>MORE</u> in Question 2.

Of course, "Yes in Question 2" $\Longrightarrow$ "Yes in Question 1"

but not vice versa!

Surprisingly, Question 2 has a much "cleaner" answer:

- If $\mathbb{F} = \mathbb{R}$, this is answered completely by "Real Spectral Theorem".

When $\mathbb{F} = \mathbb{C}$, we have the following:

<u>Complex Spectral Theorem</u>:

Let $(V, \langle , \rangle)$ be a finite dimensional inner product space over $\mathbb{C}$.

Suppose $T: V \to V$ is a linear operator.

$\boxed{TT^* = T^*T} \iff$ there exists an orthonormal eigenbasis for V

$\boxed{\text{Normal \& Self Adjoint Operators}}$

Hence, operators T satisfying the conditions in the Spectral Theorems are very special, just like symmetric matrices are special. They deserve some names.

Def$^{\underline{n}}$: Let $T: V \to V$ be a linear operator on an inner product space (over $\mathbb{R}$ or $\mathbb{C}$).

(i) T is normal $\iff$ $\boxed{TT^* = T^*T}$ i.e. T and T^* "commutes".

(ii) T is self-adjoint (Hermitian) $\iff$ $\boxed{T^* = T}$

As before, everything has a "matrix" version:

$$A \in M_{n\times n}(\mathbb{F}),\ \mathbb{F} = \mathbb{R} \text{ or } \mathbb{C} \text{ is } \begin{cases} \text{normal} & \iff AA^* = A^*A \\ \text{self-adjoint} & \iff A^* = A \end{cases}$$

We will study some examples and properties of such operators or matrices. Let's start with a simple (but important) observation.

$\boxed{\textbf{Prop:}\ \text{self-adjoint} \Rightarrow \text{normal}}$

Pf: $T^* = T \Rightarrow TT^* = T^2 = T^*T$. $\square$

Remark: The Spectral Theorems then tell us that it is easier to diagonalize (by orthonormal eigenbasis) a linear operator on a complex inner product space!

Clearly, normal $\not\Rightarrow$ self-adjoint.

Example 1: Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be (counterclockwise) rotation by the angle $\theta \in (0, \pi)$.

We know that the matrix representation of T in the standard basis β (which is orthonormal!!) is

$$[T]_\beta = A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$$

Clearly A is NOT self-adjoint.

$$A^t = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \neq A.$$

But A is normal.

$$A^t A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = AA^t.$$

(In fact, A is "orthogonal".)

[Recall: A is diagonalizable over $\mathbb{C}$ but NOT over $\mathbb{R}$.]

Example 2: Any real skew-symmetric matrix A, i.e. $A \in M_{n\times n}(\mathbb{R})$ and $\boxed{A^t = -A}$, is normal but NOT self-adjoint (unless $A = 0$).

(Reason: $A^t A = -A^2 = AA^t$.)

e.g. $A = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix}$

Corollary: Any real skew-symmetric matrix is diagonalizable over $\mathbb{C}$.

We now establish some general properties for normal and self-adjoint operators.

Theorem: Let $(V, \langle , \rangle)$ be an inner product space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

If $T: V \to V$ is a normal linear operator, then

(a) $\|Tx\| = \|T^*x\|$ for $\underline{\underline{\text{all}}}$ $x \in V$.

(b) $T - cI$ is also normal for $\underline{\underline{\text{all}}}$ $c \in \mathbb{F}$.

☆ → (c) $Tx = \lambda x \implies T^*x = \bar{\lambda}x$.

(d) $x_1 \in E_{\lambda_1}(T),\ x_2 \in E_{\lambda_2}(T) \implies \langle x_1, x_2 \rangle = 0$.

$\lambda_1 \neq \lambda_2$

In other words, eigenvectors of T in different eigenspaces are orthogonal to each other.

If, in addition, T is self-adjoint, then

(e) all the eigenvalues of T are $\underline{\underline{\text{real}}}$.

Proof: Assume T is normal, i.e. $\boxed{TT^* = T^*T}$

(a) $$\|Tx\|^2 \overset{\text{norm}}{=} \langle Tx, Tx \rangle \overset{\text{adj.}}{=} \langle x, T^*Tx \rangle \overset{\text{normal}}{=} \langle x, TT^*x \rangle$$
$$\overset{\text{adj.}}{=} \langle T^*x, T^*x \rangle \overset{\text{norm.}}{=} \|T^*x\|^2.$$

(b) $$(T - cI)^*(T - cI) = (T^* - \bar{c}I)(T - cI) = \underset{\parallel\ \text{normal}}{\boxed{T^*T}} - \bar{c}T - cT^* + |c|^2$$
$$(T - cI)(T - cI)^* = (T - cI)(T^* - \bar{c}I) = \boxed{TT^*} - \bar{c}T - cT^* + |c|^2 I$$

Hence, $T - cI$ is also normal.

(c) $Tx = \lambda x \Rightarrow (T - \lambda I)x = 0$

$$\Rightarrow \|(T-\lambda I)x\| = 0$$

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

$$\Rightarrow \|(T^* - \bar{\lambda} I)x\| = 0$$

$$\Rightarrow (T^* - \bar{\lambda} I)x = 0$$

$$\Rightarrow T^* x = \bar{\lambda} x .$$

(d) By assumption, $Tx_1 = \lambda_1 x_1$ and $Tx_2 = \lambda_2 x_2$.

$$\begin{aligned} \lambda_1 \langle x_1, x_2 \rangle &= \langle \lambda_1 x_1, x_2 \rangle \\ &= \langle Tx_1, x_2 \rangle \\ &\overset{\text{adj.}}{=} \langle x_1, T^* x_2 \rangle \\ &\overset{(c)}{=} \langle x_1, \bar{\lambda}_2 x_2 \rangle \\ &= \lambda_2 \langle x_1, x_2 \rangle \end{aligned}$$

Thus $\lambda_1 \langle x_1, x_2 \rangle = \lambda_2 \langle x_1, x_2 \rangle$. Since $\lambda_1 \neq \lambda_2$, $\langle x_1, x_2 \rangle = 0$.

(e) Now, assume further that T is self-adjoint, i.e. $\boxed{T^* = T}$.

Suppose $\lambda \in \mathbb{C}$ is an eigenvalue of T with eigenvector x.

$$\lambda x = Tx \underset{\text{adj.}}{\overset{\text{self}}{=}} T^* x \overset{(c)}{=} \bar{\lambda} x$$

Since $x \neq 0$, we have $\lambda = \bar{\lambda}$, i.e. $\lambda \in \mathbb{R}$.

$\square$

Now, we come to the proofs of the two "Spectral Theorems".

The most important ingredient is the following:

Schur's Lemma:

Let $T : V \to V$ be a linear operator on a finite dimensional inner product space $(V, \langle , \rangle)$ over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

If the characteristic polynomial of T splits over $\mathbb{F}$, then there exists an orthonormal basis β for V s.t.

$$[T]_\beta = \begin{pmatrix} \ddots & * \\ 0 & \ddots \end{pmatrix} \quad \text{is upper triangular.}$$

Remark: When $\mathbb{F} = \mathbb{C}$, the hypothesis is always satisfied by the Fundamental Theorem of Algebra.

Proof of Schur's Lemma: By induction on $\dim V = n$.

For $n = 1$, the result is trivial.

Assume the result holds for $n = k-1$. We need to show that it is true for $n = k$ as well.

Suppose $\dim V = k$. Since the characteristic polynomial of T splits over $\mathbb{F}$, there must be at least one eigenvalue $\lambda \in \mathbb{F}$ for T.

Claim: T^* has at least one eigenvalue as well.

Proof of Claim: Fix any orthonormal basis β for V, recall that $[T^*]_\beta = [T]_\beta^*$.

As λ is an eigenvalue for T, we have $\det(T-\lambda I)=0$.

By Exercise, $\det(A^*)=\overline{\det A}$ for any $A\in M_{n\times n}(\mathbb{F})$.

Therefore,

$$\begin{aligned}\det(T^*-\bar{\lambda}I) &= \det([T^*]_\beta-\bar{\lambda}I)\\ &= \det([T]_\beta^*-\bar{\lambda}I)\\ &= \det(([T]_\beta-\lambda I)^*)\\ &= \overline{\det([T]_\beta-\lambda I)} \underset{\uparrow}{=} 0\end{aligned}$$

λ is an eigenvalue of T

Hence, $\bar{\lambda}$ is an eigenvalue for T^*. —— □

(eigenvalue $\bar{\lambda}$)

By Claim, we can pick a unit eigenvector $z\in V$ for T^*.

Define the subspace $W=\operatorname{span}\{z\}^\perp$. Note: $\boxed{\dim W=k-1}$.

Claim: W is a T-invariant subspace, i.e. $T(W)\subseteq W$.

Proof of Claim: Let $w\in W$, i.e. $\langle w,z\rangle=0$.

Then $\langle Tw,z\rangle=\langle w,T^*z\rangle=\langle w,\bar{\lambda}z\rangle=\lambda\langle w,z\rangle=0$.

Hence, $Tw\in W$ as well. —— □

Now, we can consider $T_W:W\to W$, the restriction of T to W, with $\dim W=k-1$. By induction hypothesis, there exists an orthonormal basis for W, say γ, s.t. $[T_W]_{\gamma}$ is upper triangular.

Note: Since char. poly of $T_W \mid$ char. poly. of T ← splits over $\mathbb{F}$

this also splits over $\mathbb{F}$ as well.

By taking $\beta = \gamma \cup \{z\}$, since $V = W \oplus W^{\perp}$, we have

$$[T]_\beta = \left(\begin{array}{ccc|c} & [T_W]_\gamma & & * \\ & & & \vdots \\ & & & * \\ \hline 0 & \cdots & 0 & * \end{array}\right)$$

is upper-triangular since $[T_W]_\gamma$ is!

Moreover, β is clearly an orthonormal basis for V. We have thus proved the lemma by induction.

$\square$

Using Schur's Lemma, we can now prove the Spectral Theorems, which we restate below:

Complex Spectral Theorem: Let $T: V \to V$ be a linear operator on a finite dimensional inner product space over $\mathbb{C}$.

T is normal $\iff$ $\exists$ orthonormal eigenbasis for V

Real Spectral Theorem: Let $T: V \to V$ be a linear operator on a finite dimensional inner product space over $\mathbb{R}$.

T is self-adjoint $\iff$ $\exists$ orthonormal eigenbasis for V

Proof of $\mathbb{C}$ Spectral Theorem:

"$\Leftarrow$" trivial since diagonal matrices are normal.

"$\Rightarrow$" Assume T is normal. Since any polynomial splits over $\mathbb{C}$, we can apply Schur's Lemma to obtain an orthonormal basis β for V s.t.

$$[T]_\beta = \begin{pmatrix} \ddots & * \\ 0 & \ddots \end{pmatrix} = A \quad \text{is upper-triangular.}$$

Claim: A is indeed diagonal.

Proof of Claim: Let $\beta = \{v_1, v_2, \dots, v_n\}$.

Note that β orthonormal $\iff$ $\boxed{\langle v_i, v_j \rangle = \delta_{ij}.}$ (*)

Now, $[T]_\beta$ is upper triangular $\Rightarrow$ v_1 is an eigenvector.

say $T v_1 = \lambda_1 v_1$

$$\text{i.e.} \quad [T]_\beta = \begin{pmatrix} \lambda_1 & * & * & & \\ 0 & * & * & * & \\ \vdots & & * & & \\ & 0 & & \ddots & \\ 0 & & & & * \end{pmatrix} = A$$

$$[Tv_1]_\beta = \text{first column}$$

Claim: This entry is zero!

By definition, that entry is

$$A_{12} = \langle T v_2, v_1 \rangle = \langle v_2, T^* v_1 \rangle \overset{T \text{ normal}}{=} \langle v_2, \bar{\lambda}_1 v_1 \rangle \overset{(*)}{=} 0 .$$

Similarly, we can prove that all the entries above diagonal are zero (Exercise: Prove this by induction!). Hence A is diagonal.

$\square$

Proof of $\mathbb{R}$ Spectral Theorem:

"$\Leftarrow$" trivial exercise.

"$\Rightarrow$" Assume T is self-adjoint.

Then, $[T]_\beta$ is a real symmetric matrix in ANY orthonormal basis β.

Regarding $[T]_\beta \in M_{n\times n}(\mathbb{C})$ as a "complex" matrix, it is of course normal & self-adjoint.

By previous theorem, all the eigenvalues of $[T]_\beta \in M_{n\times n}(\mathbb{C})$ are in fact real.

This implies that the char. poly. of T splits over $\mathbb{R}$.

Schur's Lemma applies and there exists an orthonormal basis β for V s.t.

$$[T]_\beta = A = \begin{pmatrix} \diagdown & * \\ 0 & \diagdown \end{pmatrix} \in M_{n\times n}(\mathbb{R})$$

upper-triangular

Since T is self-adjoint, the matrix $A = [T]_\beta$ is a real symmetric matrix (as β is orthonormal).

The only upper-triangular & symmetric matrices are diagonal matrices. We are done!

$\square$