Topic#12

Invariant subspace and Cayley-Hamilton theorem

The goal of this topic is to show

> **Thm (Cayley-Hamilton).** Let $T \in \mathcal{L}(V)$ with $\dim(V) < \infty$, and $f(t)$ be the c.p. of T. Then, T satisfies the characteristic equation in the sense that
>
> $$f(T) = T_0,$$
>
> where T_0 is the zero transformation, i.e., $f(T)$ is a zero transformation.

Note:

- If one has $f(t) = \sum_{k=0}^{n} a_k t^k$, then $f(T)$ means

$$f(T) = \sum_{k=0}^{n} a_k T^k \in \mathcal{L}(V).$$

- It is also convenient to write the zero transformation T_0 as 0 and hence $f(T) = T_0$ as $f(T) = 0$.

Def. Let $T \in \mathcal{L}(V)$, and W be a subspace of V. Then, W is T-**invariant** if $T(W) \subseteq W$, i.e.

$$T(v) \in W, \ \forall v \in W.$$

Lemma#1. Let $T \in \mathcal{L}(V)$, $0 \neq x \in V$. Then

$$W \overset{def}{=} \text{span}(\{x, T(x), T^2(x), \cdots\})$$

is T-invariant. And, W is the smallest T-invariant subspace of V containing x in the sense that any T-invariant subspace of V containing x must contain W.

Proof. $T^k(x) \in V$ for $k = 0, 1, \cdots$, so, W is a subspace of V. To show W is T-invariant, take $v \in W$, then $\exists m \geq 1$ & $a_0, a_1, \cdots, a_m \in \mathbb{F}$ s.t.

$$v = a_0 x + a_1 T(x) + \cdots + a_m T^m(x).$$

$$\begin{aligned} \therefore T(v) &\overset{T \in \mathcal{L}}{=} T(a_0 x + a_1 T(x) + \cdots + a_m T^m(x)) \\ &= a_0 T(x) + a_1 T^2(x) + \cdots + a_m T^{m+1}(x) \in W. \end{aligned}$$

$\therefore W$ is T-invariant.

Let U be T-invariant with $x \in U$. To show $W \subset U$, take $v \in W$. As before, one can write $v = a_0 x + a_1 T(x) + \cdots + a_m T^m(x)$. Since $x \in U$ and U is T-invariant, all vectors x, $T(x), \ldots, T^m(x)$ are in U. Noting that U is a subspace of V, the linear combination $v = a_0 x + a_1 T(x) + \cdots + a_m T^m(x)$ is still in U. This shows $W \subset U$. □

Due to the above lemma, we introduce

Def. For $0 \neq x \in V$ and $T \in \mathcal{L}(V)$,

$$\text{span}(\{x, T(x), T^2(x), \cdots\})$$

is called the T-**cyclic subspace** of V generated by x.

Note: We let $x \neq 0$ to avoid the trivial case.

Lemma#2. Let $T \in \mathcal{L}(V)$ with $n = \dim(V) < \infty$, and $W = \text{span}(\{v, T(v), T^2(v), \cdots\})$ be the T-cyclic subspace of V generated by $0 \neq v \in V$. Let $k = \dim(W) \leq n$, then

$$\{v, T(v), T^2(v), \cdots, T^{k-1}(v)\}$$

is a basis for W.

Proof. Recall that W is the smallest T-invariant subspace of V containing v. Let

$$j \overset{def}{=} \max\{m \geq 1 : \gamma = \{v, T(v), \cdots, T^{m-1}(v)\} \text{ is l.indep't }\}.$$

Note $\sharp\gamma = m \leq k$, then j is well defined with $1 \leq j \leq k$.

We write

$$\beta = \{v, T(v), \cdots, T^{j-1}(v)\}$$

that is l.indep subset of W, and define $Z \overset{def}{=} \text{span}(\beta)$, then Z is a subspace of W with the basis β.

Claim: $Z = W$, i.e.

$$\text{span}(\{v, T(v), \cdots, T^{j-1}(v)\}) = \text{span}(\{v, T(v), \cdots\}),$$

then $j = k$. (think about why!)

Proof of Claim:

"$\subseteq$": Direct to see.

"$\supseteq$": It suffices to show $Z = $span$\{v, T(v), \cdots, T^{j-1}(v)\}$ is a T-invariant subspace containing v. (why?)

Let $z \in Z$, then $z = c_0 v + c_1 T(v) + \cdots + c_{j-1} T^{j-1}(v)$.

$\therefore T(z) = c_0 T(v) + c_1 T^2(v) + \cdots + c_{j-1} T^j(v)$.

By def of j, $\beta \cup \{T^j(v)\}$ is l. dep., then $T^j(v) \in$span$(\beta) = Z$.

$\therefore T(z)$ is a linear combination of vectors in Z

Then, $T(z) \in Z$, since Z is a v.s.

This proves Z is T-invariant. □

We need one more lemma to prove CH theorem.

Note: For $T \in \mathcal{L}(V)$, let W be a T-invariant subspace of V. Then, $T_W \in \mathcal{L}(W, W) = \mathcal{L}(W)$. (It is well defined because W is T invariant, $T(W) \subset W$.)

Lemma#3. Let $T \in \mathcal{L}(V)$ with $\dim(V) < \infty$, and W be a T-invariant subspace of V. Then the c.p. of T_W divides the c.p. of T.

Proof. Set dim(W)=$k \leq n < \infty$. Let $\gamma \stackrel{def}{=} \{v_1, \cdots, v_k\}$: o.b. for W,
extend it to an o.b. $\beta = \{v_1, \cdots, v_k, v_{k+1}, \cdots, v_n\}$ for V.
Set $[T]_\beta = A$, $[T_W]_\gamma = B$. Then $A = ([T(v_1)]_\beta | \cdots | [T(v_k)]_\beta | \cdots)$
$= \begin{pmatrix} B & B_1 \\ 0 & B_2 \end{pmatrix}$.
Let $f(t)$: c.p. of T, $g(t)$: c.p. of T_W, then

$$\begin{aligned} f(t) = \det(A - tI_n) &= \det \begin{pmatrix} B - tI_k & B_1 \\ 0 & B_2 - tI_{n-k} \end{pmatrix} \\ &= \det(B - tI_k) \cdot \det(B_2 - tI_{n-k}) \\ &= g(t) \cdot \det(B_2 - tI_{n-k}) \end{aligned}$$

$\therefore g(t)$ divides $f(t)$ where g(t) is the c.p. of T_W . □

Proof of Cayley-Hamilton Thm: Let $f(t)$ be the c.p. of T.
To show: $f(T) \in \mathcal{L}(V)$ is a zero transformation, i.e.

$$f(T)(v) = 0_v \text{ for ANY } v \in V.$$

Case $v = 0$: TRUE, since $f(T)$ is linear.
Case $v \neq 0$: Note that from now on we FIX such nonzero v.
Let $W \stackrel{def}{=} \text{span}(\{v, T(v), \cdots\})$ be the T-cyclic subspace of V generated by v with $k = \dim(W) \leq n = \dim(V)$. By Lemma#2, $\beta = \{v, T(v), \cdots, T^{k-1}(v)\}$ is an o.b. for W. By $T^k(v) \in W$, we see that there are $a_0, \cdots, a_{k-1} \in \mathbb{F}$ such that

$$a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + T^k(v) = 0. \qquad (*)$$

Then, $[T_W]_\beta = ([T_W(v)]_\beta | \cdots | [T_W(T^{k-1}(v)]_\beta) =$
$([T(v)]_\beta | \cdots | [T(T^{k-1}(v)]_\beta) = ([T(v)]_\beta | \cdots | [T^k(v)]_\beta)$

$$= \begin{pmatrix} 0 & 0 & -a_0 \\ 1 & \vdots & -a_1 \\ \vdots \ \ddots & \vdots & \vdots \\ 0 & 1 & -a_{k-1} \end{pmatrix}.$$

Let $g(t) := \det([T_W]_\beta - tI_k)$ be the c.p. of T_W, then

$$g(t) = (-1)^k(a_0 + a_1 t + \cdots + a_{k-1}t^{k-1} + t^k). \text{ (Exercise!)}$$

(Hint: multiply the k-th row by t, added to the $(k-1)$-th row, then repeat it.)

By $(*)$, we have $g(T)(v) = 0_v$.

Moreover, by Lemma#3, $g(t)$ divides $f(t)$, i.e., $\exists$ poly $q(t)$ s.t. $f(t) = q(t)g(t)$, so that $f(T) = q(T) \circ g(T)$.

Therefore,

$$f(T)(v) = [q(T) \circ g(T)](v) = q(T)(g(T)(v)) = q(T)(0_v) = 0_v.$$

□