Linear Algebra II - Cyclic subspaces and Cayley Hamilton Theorem

Ming-Hsuan Kang

March 4, 2020

Cyclic subspaces

Let $T : V \mapsto V$ be a linear transformation. For a nonzero vector $\vec{v}$ of V, there is a simple way to find the smallest T-invariant subspace containing $\vec{v}$ as follows. Let

$$S = \{\vec{v}, T(\vec{v}), T^2(\vec{v}), \cdots\}.$$

It is clear that every T-invariant subspace containing $\vec{v}$ must also contain the set S and hence it also contains $\mathrm{span}(S)$. The subspace $\mathrm{span}(S)$ is called the cyclic subspace generated by $\vec{v}$.

Theorem 1. *Let W be the cyclic subspace generated by $\vec{v}$. Then*

1. *W is T-invariant. (Especially, this implies that W is the smallest T-invariant subspace containing $\vec{v}$.)*

2. *Let m be the largest integer satisfying the condition that $\alpha = \{\vec{v}, T(\vec{v}), \cdots, T^{m-1}(\vec{v})\}$ is linearly independent. Then α forms a basis of W and $\dim W = m$.*

3. *Suppose*
$$T^m(\vec{v}) = a_0\vec{v} + a_1T(\vec{v}) + \cdots + a_{m-1}T^{m-1}(\vec{v}).$$
Then the matrix representation of $T\big|_W$ is
$$\mathrm{Rep}_\alpha(T\big|_W) = \begin{pmatrix} 0 & 0 & \cdots & \cdots & a_0 \\ 1 & 0 & \ddots & \ddots & a_1 \\ 0 & 1 & \ddots & \ddots & a_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & a_{m-1} \end{pmatrix}$$
and the characteristic polynomial of $T\big|_W$ is
$$f_{T|_W}(x) = x^m - a_{m-1}x^{m-1} - \cdots - a_1x - a_0.$$

Proof.
(1) It is sufficient to show that $T(S) \subseteq W$, which is trivial.

(2) We will show that every elements in S is contained in $\mathrm{span}(\alpha)$. Then α spans W and α is

a basis of W. By the definition of m, $\{\vec{v}, T(\vec{v}), \cdots, T^m(\vec{v})\}$ is linearly dependent, so there exists a non-trivial linear relation:

$$c_0\vec{v} + c_1 T(\vec{v}) + \cdots + c_m T^m(\vec{v}) = 0.$$

On the other hand, the first m terms of the right hand side are linearly independent, we must have $c_m \neq 0$. We conclude that $T^m(\vec{v})$ can be written as a linear combination of $\{\vec{v}, T(\vec{v}), \cdots, T^{m-1}(\vec{v})\}$ and hence it is contained in $\mathrm{span}(\alpha)$. Note that we have shown that $T(\alpha) = \{T(\vec{v}), \cdots, T^m(\vec{v})\} \subseteq \mathrm{span}(\alpha)$ which means $\mathrm{span}(\alpha)$ is T-invariant. By (1), $\mathrm{span}(\alpha) = W$.

(3) The matrix representation can be obtained directly from the definition. For example, the first column of the matrix is given by

$$\mathrm{Rep}_\alpha\Big(T(\vec{v})\Big) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

and the last column of the matrix is given by

$$\mathrm{Rep}_\alpha\Big(T\big(T^{m-1}(\vec{v})\big)\Big) = \mathrm{Rep}_\alpha\Big(T^m(\vec{v})\Big) = \mathrm{Rep}_\alpha\Big(\sum_{i=0}^{m-1} a_i T^i(\vec{v})\Big) = \begin{pmatrix} a_0 \\ \vdots \\ a_{m-1} \end{pmatrix}.$$

Finally, let us compute the characteristic polynomial

$$f_{T|_W}(x) = \det\Big(xI - \mathrm{Rep}_\alpha(T|_W)\Big) = \det\begin{pmatrix} x & 0 & \cdots & \cdots & -a_0 \\ -1 & x & \ddots & \ddots & -a_1 \\ 0 & -1 & \ddots & \ddots & -a_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & x - a_{m-1} \end{pmatrix}$$

The above determinant is left for students as an exercise. $\square$

For a polynomial $g(x) = b_0 + b_1 x + \cdots + b_n x^n \in F[x]$, set

$$g(T) = b_0 I + b_1 T + \cdots + b_n T^n,$$

which is still a linear transformation from V to V. Here I standards for the identity map. Immediately followed by the above theorem, we have

$$\begin{aligned} f_{T|_W}(T)(\vec{v}) &= \big(T^m - a_{m-1}T^{m-1} - \cdots - a_1 T - a_0 I\big)\,\vec{v} \\ &= T^m(\vec{v}) - a_{m-1}T^{m-1}(\vec{v}) - \cdots - a_1 T(\vec{v}) - a_0\vec{v} = \vec{0}. \end{aligned}$$

We summarize the above result as the following.

Corollary 2. $f_{T|_W}(T)(\vec{v}) = \vec{0}$.

Cayley-Hamilton Theorem

Now we are ready for proving the well-known Cayley-Hamilton Theorem.

Theorem 3 (Cayley-Hamilton Theorem)**.** *For any linear transformation $T : V \mapsto V$, we have $f_T(T) \equiv 0$. In other words, $f_T(T)$ is equal to the zero transformation on V.*

Proof. To show $f_T(T)$ is the zero transformation, we shall show that $f_T(T)(\vec{v}) = \vec{0}$ for all $\vec{v} \in V$. If $\vec{v} = \vec{0}$, then $f_T(T)(\vec{0}) = \vec{0}$. (This holds for any linear transformation.) If $\vec{v} \neq \vec{0}$, let W be the cyclic subspace spanned by $\vec{v}$. Since W is T-invariant, there exists some $g(x) \in F[x]$ such that

$$f_T(x) = g(x) f_{T|_W}(x).$$

Then by the above corollary,

$$f_T(T)(\vec{v}) = g(T) f_{T|_W}(T)(\vec{v}) = g(T)(\vec{0}) = \vec{0}.$$

$\square$

Cayley-Hamilton Theorem shows that there always exists some polynomial $f(x)$ such that $f(T)$ is the zero transform. Consider the collection of such polynomials

$$\mathrm{Ann}(T) = \{f(x) \in F[x] | f(T) \equiv 0\},$$

called the annihilator of T. Then we can rewrite the Cayley Hamilton Theorem as

Theorem 4 (Cayley-Hamilton Theorem)**.** *The characteristic polynomial $f_T(x)$ of T is contained in its annihilator* $\mathrm{Ann}(T)$.