Linear Algebra II - Jordan Canonical Form

Ming-Hsuan Kang

March 16, 2020

Nilpotent Linear Transformations

In the last lecture, we have shown that for the linear transform T on V , when $f_T(x)$ splits, V can be decomposed as a direct sum of generalized eigenspaces. Since each generalized eigenspace is T-invariant, it remains to find a nice matrix representation of T restricted on each generalized eigenspace.

Therefore, we may assume that V is equal to a single generalized eigenspace $E_\infty(\lambda)$. On the other hand, we can also replace T by $T - \lambda I$ so that we can further assume that $\lambda = 0$ and $V = E_\infty(0) = \ker_\infty(T)$.

Definition A linear transform T is nilpotent if $T^k \equiv 0$ for some positive integer k.

The following are equivalent definitions of nilpotency and its proof is left to readers as exercise.

Theorem 1. *The following are equivalent.*

1. *T is nilpotent.*
2. *$T^n \equiv 0$ where $n = \dim V$.*
3. *$T^k \equiv 0$ for some positive integer k.*

Now suppose that T is nilpotent. Let $\vec{v}$ be a nonzero vector. Let $\alpha = \{\vec{v}, T(\vec{v}), \cdots, T^{m-1}(\vec{v})\}$ is the basis of the cyclic subspace W generated by $\vec{v}$. Then we have

$$T^m(\vec{v}) = a_0\vec{v} + \cdots + a_{m-1}T^{m-1}(\vec{v})$$

for some $a_0, \cdots, a_{m-1}$ in F such that

$$f_{T\big|_W}(x) = x^m - a_{m-1}x^{m-1} - \cdots - a_0,$$

the characteristic polynomial of T restricted on W. On the other hand, $f_{T\big|_W}(x)$ divides $f_T(x) = x^n$, which means $f_{T\big|_W}(x) = x^m$ and $T^m(\vec{v}) = \vec{0}$. In this case, we have

$$\mathrm{Rep}_\alpha(T\big|_W) = \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 1 & 0 & \ddots & \ddots & 0 \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Besides, it is convenient to express the action of T on W by the diagram

$$\vec{0} \leftarrow T^{m-1}(\vec{v}) \leftarrow \cdots \leftarrow T(\vec{v}) \leftarrow \vec{v}.$$

The goal of this lecture is the following theorem.

Theorem 2. *For a nilpotent linear transformation T on V, V can be decomposed as a direct sum of cyclic subspaces of T.*

Proof. Let us prove the theorem by induction on $n = dimV$. If $n = 1$, then V itself is the cyclic subspace of any nonzero vector. Suppose theorem holds for cases of dimension $< n$.

Case a): Suppose $\ker(T) \not\subseteq \mathrm{Im}(T)$, then there exists some $\vec{v} \in \ker(T)$ with $\vec{v} \notin \mathrm{Im}(T)$. We will decompose V as a direct sum of two nonzero T-invariant subspaces, then by induction, each of them can be decomposed as a directed sum of cyclic subspaces and so is V.

Note that

$$F\vec{v} + \mathrm{Im}(T) = F\vec{v} \oplus \mathrm{Im}(T),$$

so we can decompose V as

$$V = F\vec{v} \oplus \mathrm{Im}(T) \oplus W$$

for some subspace W. Since

$$T(F\vec{v}) = \{\vec{0}\} \quad \text{and} \quad T\bigg(\mathrm{Im}(T) \oplus W\bigg) \subseteq T(V) = \mathrm{Im}(T),$$

$F\vec{v}$ and $\mathrm{Im}(T) \oplus W$ are both T-invariant.

Case b): Suppose $\ker(T) \subseteq \mathrm{Im}(T)$. Since 0 is always an eigenvalue of T, $\ker(T)$ is nonzero and the dimension of $\mathrm{Im}(T)$ is smaller then the dimensino of V. By induction,

$$\mathrm{Im}(T) = \bigoplus_{i=1}^{k} \mathrm{cyclic}(\tilde{\mathrm{v}}_{\mathrm{i}}).$$

for some nonzero $\vec{v}_1, \cdots, \vec{v}_k \in V$. From the above decomposition, it is easy to see that $\dim\ker(T) = k$ and

$$\mathrm{Im}(T) = \left(\bigoplus_{i=1}^{k} F\vec{v}_i\right) \bigoplus \mathrm{Im}(T^2).$$

Since $\vec{v}_1, \cdots, \vec{v}_2$ all lie in $\mathrm{Im}(T)$, for each i, we can find $\vec{u}_i \in V$ such that $T(\vec{u}_i) = \vec{v}_i$. To complete the proof, we shall show that $V = \bigoplus_{i=1}^{k} \mathrm{cyclic}(\vec{u}_i)$, or equivalent

$$V = \left(\bigoplus_{i=1}^{k} F\vec{w}_i\right) \bigoplus \mathrm{Im}(T).$$

First, by the dimension formula,

$$\dim V = \dim\ker(T) + \dim\mathrm{Im}(T) = k + \dim\mathrm{Im}(T).$$

It remains to show $\left(\sum_{i=1}^{k} F\vec{w}_i\right) + \mathrm{Im}(T)$ is a direct sum. Suppose $\sum_{i=1}^{k} a_i\vec{w}_i + \vec{v} = \vec{0}$ for some $a_i \in F$ and $\vec{v} \in \mathrm{Im}(T)$. Applying T to this equation, we obtain

$$\vec{0} = T\left(\sum_{i=1}^{k} a_i\vec{w}_i + \vec{v}\right) = \sum_{i=1}^{k} a_i\vec{v}_i + T(\vec{v}) \in \left(\bigoplus_{i=1}^{k} F\vec{v}_i\right) \bigoplus \mathrm{Im}(T^2).$$

We conclude that $a_1 = \cdots = a_k = 0$ and hence $\vec{v} = \vec{0}$. Thus, $\left(\sum_{i=1}^{k} F\vec{w}_i\right) + \mathrm{Im}(T)$ is a direct sum. $\square$

Next, let us discuss the uniqueness of cyclic subspace decomposition. Consider the following example. Suppose

$$V = \mathrm{cyclic}(\vec{v}_1) \oplus \mathrm{cyclic}(\vec{v}_2) \oplus \mathrm{cyclic}(\vec{v}_3) \oplus \mathrm{cyclic}(\vec{v}_4)$$

and the action of T can be characterized by

$$\begin{array}{ccccccc} \vec{0} & \leftarrow & T^2(\vec{v}_1) & \leftarrow & T(\vec{v}_1) & \leftarrow & \vec{v}_1 \\ \vec{0} & \leftarrow & T(\vec{v}_2) & \leftarrow & \vec{v}_2 & & \\ \vec{0} & \leftarrow & T(\vec{v}_3) & \leftarrow & \vec{v}_3 & & \\ \vec{0} & \leftarrow & \vec{v}_4. & & & & \end{array}$$

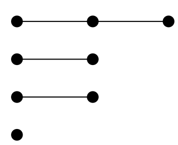
From the above diagram, the red parts form the basis of $\ker(T)$.

$$\begin{array}{ccccccc} \vec{0} & \leftarrow & T^2(\vec{v}_1) & \leftarrow & T(\vec{v}_1) & \leftarrow & \vec{v}_1 \\ \vec{0} & \leftarrow & T(\vec{v}_2) & \leftarrow & \vec{v}_2 & & \\ \vec{0} & \leftarrow & T(\vec{v}_3) & \leftarrow & \vec{v}_3 & & \\ \vec{0} & \leftarrow & \vec{v}_4. & & & & \end{array}$$

Similarly, the blue parts form the basis of $\ker(T^2)$.

$$\begin{array}{ccccccc} \vec{0} & \leftarrow & T^2(\vec{v}_1) & \leftarrow & T(\vec{v}_1) & \leftarrow & \vec{v}_1 \\ \vec{0} & \leftarrow & T(\vec{v}_2) & \leftarrow & \vec{v}_2 & & \\ \vec{0} & \leftarrow & T(\vec{v}_3) & \leftarrow & \vec{v}_3 & & \\ \vec{0} & \leftarrow & \vec{v}_4. & & & & \end{array}$$

Remark. If we only need care about the sizes of cyclic subspaces then it is common to use "the dot diagram" as



Let d_i be the dimension of $\mathrm{cyclic}(\vec{v}_i)$. We conclude that

$$\dim\ker(T) = \#\{i | d_i \geq 1\}.$$

and

$$\dim\ker(T^2) = \#\{i | d_i \geq 1\} + \#\{j | d_j \geq 2\}.$$

In general, suppose $V = \oplus_{i=1}^{k} \mathrm{cyclic}(\vec{v}_i)$. Let $d_i = \dim \mathrm{cyclic}(\vec{v}_i)$. We have

$$\begin{array}{rcl} \dim\ker(T) & = & \#\{i | d_i \geq 1\} \\ \dim\ker(T^2) - \dim\ker(T) & = & \#\{i | d_i \geq 2\} \\ & \vdots & \\ \dim\ker(T^r) - \dim\ker(T^{r-1}) & = & \#\{i | d_i \geq r\}. \end{array}$$

Corollary 3 (Uniqueness of cyclic decomposition)**.** *For a nilpotent linear transformation T on V, let $V = \oplus_{i=1}^{k} W_i = \oplus_{j=1}^{r} U_j$ be two cyclic subspace decompositions. If none of subspace is the zero subspace, then $r = k$. Moreover, if $\dim W_i \geq dim W_j$ and $\dim U_i \geq dim U_j$ for all $i > j$, then $\dim W_i = \dim U_i$ for all i.*

Explicit Cyclic Subspace Decomposition

The basis principle to compute cyclic subspace decompositions is to find the highest dimensional cyclic subspaces first.

Example Let

$$A = \begin{pmatrix} 3 & 12 & -9 \\ -1 & -6 & 6 \\ -1 & -4 & 3 \end{pmatrix}.$$

By direct computation, we have

$$\dim\ker(A) = 1, \dim\ker(A^2) = 2, \text{ and } \dim\ker(A^3) = 3$$

which implies that the dot diagram is

$$\bullet \!-\!\!-\!\!-\! \bullet \!-\!\!-\!\!-\! \bullet$$

In this case, V admits a basis of the form

$$\vec{0} \;\leftarrow A^2(\vec{v}_1) \;\;\leftarrow\;\; A(\vec{v}_1) \;\;\leftarrow\;\; \vec{v}_1$$

To find $\vec{v}_1$, note that

$$\mathrm{Im}(A^2) = \mathrm{Im}(\begin{pmatrix} 6 & 0 & 18 \\ -3 & 0 & -9 \\ -2 & 0 & -6 \end{pmatrix}) = \mathrm{span}\{A^2(\vec{v}_1)\}.$$

Therefore, we can choose $A^2(\vec{v}_1) = \left(\begin{smallmatrix} 6 \\ -3 \\ -2 \end{smallmatrix}\right)$ and $\vec{v}_1 = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$. . Finally let $\alpha = \{\vec{v}_1, A(\vec{v}_1), A^2(\vec{v}_1)\}$ and $P = \left(\begin{smallmatrix} 1 & 3 & 6 \\ 0 & -1 & -3 \\ 0 & -1 & -2 \end{smallmatrix}\right)$ be the matrix of change basis. Then

$$P^{-1}AP = \left(\begin{smallmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{smallmatrix}\right).$$

Example Let

$$A = \begin{pmatrix} 3 & 6 & -3 \\ -2 & -4 & 2 \\ -1 & -2 & 1 \end{pmatrix}.$$

By direct computation, we have

$$\dim\ker(A) = 2 \text{ and } \dim\ker(A^2) = 3$$

which implies that the dot diagram is

$$\begin{array}{l} \bullet \!-\!\!-\!\!-\! \bullet \\ \bullet \end{array}$$

In this case, V admits a basis of the form

$$\begin{array}{ccccc} \vec{0} & \leftarrow & A(\vec{v}_1) & \leftarrow & \vec{v}_1 \\ \vec{0} & \leftarrow & \vec{v}_2. & & \end{array}$$

First, let us find $\vec{v}_1$. Note that

$$\mathrm{Im}(A) = \mathrm{span}\{A(\vec{v}_1)\}.$$

Therefore, we can choose $A(\vec{v}_1) = \begin{pmatrix} 3 \\ -2 \\ -1 \end{pmatrix}$ and $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. To find $\vec{v}_2$, note that

$$\mathrm{span}\{A(\vec{v}_1), \vec{v}_2\} = \ker(A) = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \right\}.$$

Therefore, we can choose $v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Finally let $\alpha = \{\vec{v}_1, A(\vec{v}_1), \vec{v}_2\}$ and $P = \begin{pmatrix} 1 & 3 & 1 \\ 0 & -2 & 0 \\ 0 & -1 & 1 \end{pmatrix}$ be the matrix of change basis. Then

$$P^{-1}AP = \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 1 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array} \right).$$