Linear Algebra II - Nilpotent Linear Transformation

Ming-Hsuan Kang

March 19, 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.*

Definition For a nilpotent linear transform T, the smallest positive integer k satisfying $T^k \equiv= 0$ is called the index of T.

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. *Let T be a nilpotent linear transformation on V of index k. Let $\vec{v}_1, \cdots, \vec{v}_m$ be vectors in V such that $T^{k-1}(\vec{v}_1), \cdots, T^{k-1}(\vec{v}_m)$ forms a basis of $\mathrm{Im}(T^{k-1})$. The following hold.*

1. $V = \ker(T^{k-1}) \bigoplus F\vec{v}_i$.
2. $\sum \mathrm{cyclic}(\vec{v}_i) = \bigoplus \mathrm{cyclic}(\vec{v}_i)$.
3. *There exists a T-invariant subspace W contained in $\ker(T^{k-1})$ such that $V = W \bigoplus \mathrm{cyclic}(\vec{v}_i)$.*

Proof. 1) First, we show that

$$\ker(T^{k-1}) + \sum_{i=1}^m F\vec{v}_i = \ker(T^{k-1}) \bigoplus_{i=1}^m F\vec{v}_i.$$

Suppose $\vec{0} = \vec{v} + \sum a_i\vec{v}_i$ for some $\vec{v} \in \ker(T^{k-1})$ and $a_i \in F$. Apply T^{k-1} to the both hand sides, we obtain

$$\vec{0} = \vec{0} + \sum_{i=1}^m a_i T^{k-1}(\vec{v}_i).$$

By assumption of $\vec{v}_i$, we have $a_1 = \cdots = a_m = 0$ which implies that $\vec{v} = 0$. Therefore, $\ker(T^{k-1}) + \sum F\vec{v}_i = \ker(T^{k-1}) \bigoplus F\vec{v}_i$.

On the other hand, by dimensional formula, we have

$$\dim\left(\ker(T^{k-1}) \bigoplus F\vec{v}_i\right) = \dim\ker(T^{k-1}) + \dim\mathrm{Im}(T^{k-1}) = \dim V.$$

We conclude that $\ker(T^{k-1}) \bigoplus F\vec{v}_i = V$.

2) By assumption of $\vec{v}_i$, the subspace $\mathrm{cyclic}(\vec{v}_i)$ admits a basis $\alpha_i = \{\vec{v}_i, T(\vec{v}_i) \cdots, T^{k-1}(\vec{v}_i)\}$. It

remains to show that $\alpha_1 \cup \cdots \alpha_m$ is linearly independent. Suppose $\sum_{i,j} a_{ij} T^i(\vec{v}_j) = \vec{0}$ is a non-trivial linear relation. Let r be the smallest integer satisfying $a_{rj} \neq 0$ for some j. Then applying T^{k-r-1} to the above equation, we obtain

$$\vec{0} = T^{k-r-1}\left(\sum_{i=r}^{k}\sum_{j=1}^{m} a_{ij}T^i(\vec{v}_j)\right) = \left(\sum_{i=r}^{k}\sum_{j} a_{ij}T^{k-r-1+i}(\vec{v}_j)\right) = \left(\sum_{j=1}^{m} a_{ij}T^{k-1}(\vec{v}_j)\right).$$

Thus, we obtain a non-trivial linear relation of $T^{k-1}(\vec{v}_1), \cdots, T^{k-1}(\vec{v}_m)$, which is a contradiction.

3) Let us prove the theorem by induction on the index k. If $k = 1$, then T is the zero transform and T^0 is the identity map. In this case, $\vec{v}_1, \cdots, \vec{v}_m$ forms a basis of the whole V and $W = \{\vec{0}\}$. Now suppose $k > 1$ and the theorem holds for index $< k$. Denote T restricted on $\ker(T^{k-1})$ by $\tilde{T}$ for short. Then $\tilde{T}$ is a nilpotent linear transform of index $k-1$. Note that $T(\vec{v}_1), \cdots, T(\vec{v}_m)$ are contained in $\ker(T^{k-1})$ and

$$\{\tilde{T}^{k-2}(T(\vec{v}_1)), \cdots, \tilde{T}^{k-2}(T(\vec{v}_m))\} = \{T^{k-1}(\vec{v}_1), \cdots, T^{k-1}(\vec{v}_m)\}$$

is linearly independent. Therefore, we can extent it to a basis of $\mathrm{Im}(\tilde{T}^{k-2})$, namely

$$\{\tilde{T}^{k-2}(T(\vec{v}_1)), \cdots, \tilde{T}^{k-2}(T(\vec{v}_m))\} \cup \{\tilde{T}^{k-2}(\vec{u}_1), \cdots, \tilde{T}^{k-2}(\vec{u}_r)\}.$$

By induction, the part 3) holds for $\tilde{T}$ so that there exists a $\tilde{T}$-invariant subspace W_0 in $\ker(\tilde{T}^{k-2})$ such that

$$\ker(T^{k-1}) = W_0 \bigoplus_{i=1}^{m} \mathrm{cyclic}(T(\vec{v}_i)) \bigoplus_{j=1}^{r} \mathrm{cyclic}(\vec{u}_j).$$

Finally, set $W = W_0 \bigoplus_{j=1}^{r} \mathrm{cyclic}(\vec{u}_j)$, which is a $\tilde{T}$-invariant subspace in $\ker(T^{k-1})$ and also a T-invariant subspace in V. Together with the part 1), we have

$$V = \ker(T^{k-1}) \bigoplus_{i=1}^{m} F\vec{v}_i = W \bigoplus_{i=1}^{m} \mathrm{cyclic}(T(\vec{v}_i)) \bigoplus_{i=1}^{m} F\vec{v}_i = W \bigoplus_{i=1}^{m} \mathrm{cyclic}(\vec{v}_i).$$

□

Corollary 3. *For a nilpotent linear transform T on V, V can be decomposed as a direct sum of cyclic subspaces of T.*

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

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

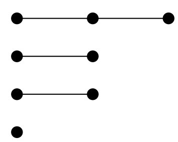
From the above diagram, the red parts form the basis of $\ker(T)$.

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

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 $\text{cyclic}(\vec{v}_i)$. 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 $V = \oplus_{i=1}^{k}\text{cyclic}(\vec{v}_i)$. Let $d_i = \dim\text{cyclic}(\vec{v}_i)$. 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 4 (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 $\dim W_i \geq dimW_j$ and $\dim U_i \geq dimU_j$ 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

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

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

$$\begin{matrix} \bullet\!\!-\!\!\!-\!\!\!-\!\!\bullet \\ \bullet \end{matrix}$$

In this case, V admits a basis of the form

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

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

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

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

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

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

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

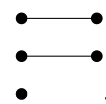
Example Let

$$A = \begin{pmatrix} 2 & 4 & 0 & 0 & 4 \\ -1 & -2 & 1 & 2 & 1 \\ 0 & 0 & 3 & 6 & 9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -2 & -3 \end{pmatrix}.$$

By direct computation, we have

$$\dim\ker(A) = 3 \text{ and } \dim\ker(A^2) = 5,$$

which implies that its dot diagram is



In this case, V admits a basis of the form

$$\begin{array}{ccccc} \vec{0} & \leftarrow & A(\vec{v}_1) & \leftarrow & \vec{v}_1 \\ \vec{0} & \leftarrow & A(\vec{v}_2) & \leftarrow & \vec{v}_2 \\ \vec{0} & \leftarrow & \vec{v}_3. & & \end{array}$$

First, let us find cyclic subspaces of dimension two. Since

$$\mathrm{span}\{A(\vec{v}_1), A(\vec{v}_2)\} = \mathrm{Im}(A),$$

we can choose $A(\vec{v}_1) = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ and $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$; we can choose $A(\vec{v}_2) = \begin{pmatrix} 0 \\ 1 \\ 3 \\ 0 \\ -1 \end{pmatrix}$ and $\vec{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

To find $\vec{v}_3$, note that

$$\mathrm{span}\{\vec{v}_3, A(\vec{v}_1), A(\vec{v}_2)\} = \ker(A) = \mathrm{span}\left\{ \begin{pmatrix} 0 \\ -1 \\ -3 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

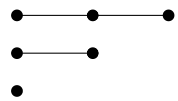
Therefore, we can choose $\vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}$. Finally let $\alpha = \{\vec{v}_1, A(\vec{v}_1), \vec{v}_2, A(\vec{v}_2), \vec{v}_3\}$ and $P = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix}$ be the matrix of change basis. Then

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

Example Let $A = \begin{pmatrix} 4 & 6 & -2 & 2 & 0 & 2 \\ -2 & -4 & 2 & -2 & 2 & -2 \\ 3 & -1 & 0 & 0 & 11 & -8 \\ 2 & -1 & 0 & 0 & 8 & -6 \\ -2 & -2 & 1 & -1 & -2 & 1 \\ -1 & 0 & 0 & 0 & -3 & 2 \end{pmatrix}$. By direct computation, we have

$$\dim\ker(A) = 3, \dim\ker(A^2) = 5 \text{ ,and } \dim\ker(A^3) = 6$$

which implies that the dot diagram is



In this case, V admits a basis of the form

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

First, let us find the cyclic subspace of dimension three. Since

$$\mathrm{span}\{A^2(\vec{v}_1)\} = \mathrm{Im}(A^2) = \mathrm{Im}(\begin{pmatrix} 0 & 0 & 4 & -4 & 0 & 4 \\ 0 & 0 & -2 & 2 & 0 & -2 \\ 0 & 0 & 3 & -3 & 0 & 3 \\ 0 & 0 & 2 & -2 & 0 & 2 \\ 0 & 0 & -2 & 2 & 0 & -2 \\ 0 & 0 & -1 & 1 & 0 & -1 \end{pmatrix}),$$

we can choose $A^2(\vec{v}_1) = \begin{pmatrix} 4 \\ -2 \\ 3 \\ 2 \\ -2 \\ -1 \end{pmatrix}$ and $\vec{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. Next, let us find the cyclic subspace of dimension two. Note that

$$\mathrm{span}\Big\{A^2(\vec{v}_1), A(\vec{v}_2)\Big\} = A(\ker(A^2)) = \mathrm{span}\Big\{\begin{pmatrix} 4 \\ -4 \\ -8 \\ -6 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 11 \\ 8 \\ -2 \\ -3 \end{pmatrix}\Big\}.$$

Therefore, we can choose $A(\vec{v}_2) = \begin{pmatrix} 4 \\ -4 \\ -8 \\ -6 \\ 0 \\ 2 \end{pmatrix}$ and $\vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}$. Third, let us find the cyclic subspace of dimension one. Note that

$$\mathrm{span}\Big\{A^2(\vec{v}_1), A(\vec{v}_2), \vec{v}_3\Big\} = \ker(A) = \mathrm{span}\Big\{\begin{pmatrix} 2 \\ -2 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -3 \\ 2 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}\Big\}.$$

Therefore, we can choose $\vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

Finally let $\alpha = \{\vec{v}_1, A(\vec{v}_1), A^2(\vec{v}_1), \vec{v}_2, A(\vec{v}_2), \vec{v}_3\}$ and P be the corresponding matrix of change basis. Then

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