Math 672 Lecture 31

Theorem: Let V be a finite-dimensional vector space over a field F and let $N \in \mathcal{L}(V)$ be nilpotent. Then there exist nonnegative integers $m_1, \ldots, m_k$ and vectors $v_1, \ldots, v_k \in V$ such that

$$\{v_1, N(v_1), \ldots, N^{m_1}(v_1), v_2, N(v_2), \ldots, N^{m_2}(v_2), \ldots, v_k, N(v_k), \ldots, N^{m_k}(v_k)\}$$

is a basis for V and

$$N^{m_j+1}(v_j) = 0 \quad \forall j = 1, 2, \ldots, k.$$

Proof: We argue by induction on $n = \dim(V)$. If $n=1$, then the only nilpotent operator is the zero operator and we can take $m_1 = 0$ and v_1 to be any nonzero vector in V. (Why must N be the zero operator? Suppose $V = \mathrm{span}(v)$ and $N(v) \neq 0$. Then $N(v)$ must equal λv for some nonzero λ (since $V = \mathrm{span}(v)$) and hence $N^k(v) \neq 0$ $\forall k \geq 1$, a contradiction.)

Now suppose the result holds for all vector spaces over F with dimension at most $n-1$, where $n \geq 2$. Let V be a vector space over F having dimension n and let $N \in \mathcal{L}(V)$ be nilpotent. Note that N is not injective (because it is nilpotent) and hence N is not surjective. It follows that, if $U = R(N)$, then $\dim(U) < n$. Recall that

$R(N)$ is invariant under N, so we can define $S \in \mathcal{L}(U)$ by

$S = N|_U$. Then S is nilpotent and, by the induction hypothesis,

there exist nonnegative integers $m_1', \ldots, m_t'$ and $u_1, \ldots, u_t \in U$ such that

$$u_1, N(u_1), \ldots, N^{m_1'}(u_1), \; \ldots \;, u_t, N(u_t), \ldots, N^{m_t'}(u_t)$$

(using the fact that $S = N|_U$)

form a basis for U. Moreover, each $u_j \in R(N)$, so

there exist $v_1, \ldots, v_t \in V$ such that $u_j = N(v_j)$ for $j = 1, \ldots, t$.

Define $m_j = m_j' + 1$ and consider

$$\mathcal{B}' = \{ v_1, u_1, N(u_1), \ldots, N^{m_1'}(u_1), \; \ldots \;, v_t, u_t, N(u_t), \ldots, N^{m_t'}(u_t) \}$$

$$= \{ v_1, N(v_1), \ldots, N^{m_1}(v_1), \; \ldots \;, v_t, N(v_t), \ldots, N^{m_t}(u_t) \}.$$

We will prove that $\mathcal{B}'$ is linearly independent and that $\mathcal{B}'$ can be extended to a basis $\mathcal{B}$ for V of the type described in the theorem.

Suppose $\alpha_{ij} \in F$, $0 \le i \le m_j$, $1 \le j \le t$, satisfy

$$\sum_{j=1}^{t} \sum_{i=0}^{m_j} \alpha_{ij} N^i(v_j) = 0.$$

Applying N to both sides yields

$$\sum_{j=1}^{t} \sum_{i=0}^{m_j} \alpha_{ij} N^{i+1}(v_j) = 0$$

$$\Rightarrow \sum_{j=1}^{t} \sum_{i=0}^{m_j'} \alpha_{ij} N^i(u_j) = 0 \quad (\text{since } N^{i+1}(v_j) = N^i(u_j) \text{ and } N^{m_j'+1}(u_j) = 0 \ \forall j)$$

$$\Rightarrow \alpha_{ij} = 0, \ 0 \le i \le m_j - 1, \ 1 \le j \le t$$

$$\Rightarrow \sum_{j=1}^{t} \alpha_{m_j, j} N^{m_j}(v_j) = 0$$

$$\Rightarrow \sum_{j=1}^{t} \alpha_{m_j, j} N^{m_j'}(u_j) = 0$$

$$\Rightarrow \alpha_{m_j, j} = 0, \ 1 \le j \le t \quad (\text{a subset of a linearly independent set is linearly independent}).$$

Thus $\alpha_{ij} = 0$, $0 \le i \le m_j$, $1 \le j \le t$, and we have shown that $\mathcal{B}'$ is linearly independent.

If $|\mathcal{B}'| < n$, then extend $\mathcal{B}'$ to a basis $\mathcal{B}''$ of V by adding vectors $w_{t+1}, \dots, w_k$. (Otherwise, just define $k = t$ and the proof is complete). For each $\ell = t+1, \dots, k$,

$$N(w_\ell) \in R(N)$$

$$\Rightarrow N(w_\ell) = \sum_{j=1}^{t} \sum_{i=1}^{m_j} \alpha_{ij} N^i(v_j) \text{ for some } \alpha_{ij} \in F$$

Starts at 1 because we don't need $v_1, \dots, v_t$ to span $R(N)$

$$\Rightarrow N(v_\ell) = 0, \text{ where } v_\ell = w_\ell - \sum_{j=1}^{t} \sum_{i=1}^{m_j} \alpha_{ij} N^{i-1}(v_j)$$

Now define

$$\mathcal{B} = \{v_1, N(v_1), \ldots, N^{m_1}(v_1), \ldots, v_t, N(v_t), \ldots, N^{m_t}(v_t), v_{t+1}, \ldots, v_n\}.$$

Then $\mathcal{B}$ spans V, since $\mathcal{B}' \subseteq \mathcal{B}$ and $w_\ell \in \operatorname{span}(\mathcal{B})$ for $\ell = t+1, \ldots, n$.

Thus, since $|\mathcal{B}| = |\mathcal{B}''|$, we see that $\mathcal{B}$ is a basis for V,

and $\mathcal{B}$ has the desired form (note that $m_j = 0$ and $N^{0+1}(v_j) = 0$

for $j = t+1, \ldots, n$). This completes the proof by induction. //

Now let V be a finite-dimensional complex vector space, let $T \in \mathcal{L}(V)$,

and let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. We know the following:

- $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_k, T)$
- Each $G(\lambda_j, T)$ is invariant under T.
- $(T - \lambda_j I)\big|_{G(\lambda_j, T)}$ is nilpotent.

Let us assume that m_j is the smallest positive integer such that

$$\left[(T - \lambda_j I)\big|_{G(\lambda_j, T)}\right]^{m_j} = 0.$$

Suppose s_j, t_j are the geometric and algebraic multiplicities of

λ_j:

$$s_j = \dim(E(\lambda_j, T)) \leq \dim(G(\lambda_j, T)) = t_j$$

There is a basis for $G(\lambda_j, T)$ consisting of s_j generalized eigenvector chains (some of which may contain only one vector, which would be an eigenvector):

$$\begin{aligned}\mathcal{B}_j\ \{ & (T-\lambda_j I)^{r_{1j}-1}(v_{1j}), (T-\lambda_j I)^{r_{1j}-2}(v_{1j}), \ldots, v_{1j}, \\ & (T-\lambda_j I)^{r_{2j}-1}(v_{2j}), (T-\lambda_j I)^{r_{2j}-2}(v_{2j}), \ldots, v_{2j}, \\ & \ldots, (T-\lambda_j)^{r_{s_j,j}-1}(v_{s_j,j}), (T-\lambda_j I)^{r_{s_j,j}-2}(v_{s_j,j}), \ldots, v_{s_j,j} \}\end{aligned}$$

$1 \le r_{ij} \le m_j$ for $i = 1, \ldots, s_j$.

(this follows from the previous theorem). If we define

$$\mathcal{B} = \bigcup_{j=1}^{k} \mathcal{B}_j,$$

Then $A = \mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ is block diagonal, where the jth block is itself block diagonal, with s_j Jordan blocks (some of which may be 1×1. When $\mathcal{B}$ is chosen this way, A is called the <u>Jordan form</u> (or <u>Jordan canonical form</u> or <u>Jordan normal form</u>) of T.

<u>Notes</u>:

- The Jordan form is generally not unique (since the blocks can be reordered by reordering the chains).
- The values of k, $\lambda_1, \lambda_2, \ldots, \lambda_k$, $m_1, \ldots, m_k$, $t_1, \ldots, t_k$, and $s_1, \ldots, s_k$ may or may not uniquely (up to reordering) define the Jordan form.

__Example__: Suppose $k=1$,

$$(T-\lambda I)^3 = 0 \qquad (m=3)$$

$$\dim(\mathcal{N}((T-\lambda I)^3)) = 7 \qquad (t=7)$$

$$\dim(\mathcal{N}(T-\lambda I)) = 3 \qquad (s=3)$$

There must be one chain of length 3:

$$(T-\lambda I)^2(v_1),\ (T-\lambda I)(v_1),\ v_1$$

There could be a second chain of length 3, plus a chain of length 1:

$$(T-\lambda I)^2(v_2),\ (T-\lambda I)(v_2),\ v_2$$

$$v_3$$

In this case, the Jordan form is

$$\begin{bmatrix} \lambda & 1 & & & & & \\ & \lambda & 1 & & & & \\ & & \lambda & & & & \\ & & & \lambda & 1 & & \\ & & & & \lambda & 1 & \\ & & & & & \lambda & \\ & & & & & & \lambda \end{bmatrix}$$

Or (in addition to the first chain of length 3), we could have two chains of length 2:

$$(T-\lambda I)(v_2),\ v_2$$

$(T-\lambda I)(v_3), v_3.$

Then the Jordan form would be

$$\left[\begin{array}{ccc|cc|cc} \lambda & 1 & & & & & \\ & \lambda & 1 & & & & \\ & & \lambda & & & & \\ \hline & & & \lambda & 1 & & \\ & & & & \lambda & & \\ \hline & & & & & \lambda & 1 \\ & & & & & & \lambda \end{array}\right]$$

What information uniquely determines the Jordan form (again, up to reordering)?

<u>Answer</u>: $\dim(\mathcal{N}((T-\lambda I)^j)),\ 1 \le j \le m.$

	Case 1	Case 2
$\dim(\mathcal{N}(T-\lambda I))$	3	3
$\dim(\mathcal{N}((T-\lambda I)^2))$	5	6
$\dim(\mathcal{N}((T-\lambda I)^3))$	7	7