Math 672 Lecture 32

<u>Recall:</u> A <u>monic</u> polynomial is a polynomial whose leading coefficient is 1.

<u>Theorem</u>: Let V be a finite-dimensional vector space over a field F and let $T \in \mathcal{L}(V)$. Then there exists a unique monic polynomial m_T of smallest degree such that $m_T(T) = 0$. We call m_T the <u>minimal polynomial</u> of T.

<u>Proof:</u> Let $n = \dim(V)$. Then

$$\{I, T, T^2, \dots, T^{n^2}\}$$

is linearly dependent (since $\dim(\mathcal{L}(V)) = n^2$) and hence there exists $d \in \mathbb{Z}^+$ such that $\{I, T, \dots, T^{d-1}\}$ is linearly independent and $T^d \in \operatorname{span}(I, T, \dots, T^{d-1})$. We can thus write

$$T^d = \sum_{j=0}^{d-1} \alpha_j T^j \quad \text{for some } \alpha_0, \alpha_1, \dots, \alpha_{d-1} \in F$$

$$\Rightarrow T^d - \sum_{j=1}^{d-1} \alpha_j T^j = 0$$

$$\Rightarrow m_T(T) = 0, \; m_T(x) = x^d - \sum_{j=0}^{d-1} \alpha_j T^j.$$

Thus there exists a monic polynomial m_T such that $m_T(T) = 0$.

If $p \in \mathcal{P}(F)$ and $\deg(p) = t < d$, then $p(T) \neq 0$, since otherwise

$\{I, T, \ldots, T^t\} \subseteq \{I, T, \ldots, T^{d-1}\}$ is linearly dependent. If $p \in \mathcal{P}(F)$ is a monic polynomial satisfying $\deg(p) = d$ and $p(T) = 0$, then

$$m_T(T) - p(T) = 0 - 0 = 0$$
$$\Rightarrow (m_T - p)(T) = 0.$$

Since m_T and p are both monic and of degree d,

$$\deg(m_T - p) < d \quad \text{or} \quad m_T - p = 0.$$

The first possibility is impossible (since $\{I, T, \ldots, T^{d-1}\}$ is linearly independent) and hence $m_T - p = 0$ must hold. Thus m_T is unique. //

<u>Theorem</u>: Let V be a finite-dimensional vector space over a field F, let $T \in \mathcal{L}(V)$, and let m_T be the minimal polynomial of T. Then $\lambda \in F$ is an eigenvalue of T iff λ is a root of m_T.

<u>Proof</u>: First suppose λ is a root of m_T; then we can write

$$m_T(x) = (x - \lambda) q(x),$$

where $q \in \mathcal{P}(F)$ has degree less than $\deg(m_T)$. It follows that

$$q(T) \neq 0$$
$$\Rightarrow q(T)(v) \neq 0 \text{ for some } v \in V, v \neq 0$$
$$\Rightarrow (T - \lambda I) q(T)(v) = 0 \quad (\text{since } (T - \lambda I) q(T) = m_T(T) = 0)$$

$\Rightarrow \lambda$ is an eigenvalue of T (with eigenvector $q(T)(v)$).

Conversely, suppose λ is an eigenvalue of T, say

$$T(v) = \lambda v, \quad v \neq 0.$$

It follows that

$$m_T(T)(v) = m_T(\lambda)\, v$$

(since $T^j(v) = \lambda^j v \quad \forall j \geq 0$). But then

$$\begin{aligned} m_T(T) = 0 &\Rightarrow m_T(T)(v) = 0 \\ &\Rightarrow m_T(\lambda) v = 0 \\ &\Rightarrow m_T(\lambda) = 0 \quad (\text{since } v \neq 0). \end{aligned}$$

Thus λ is a root of m_T. //

<u>Theorem</u>: Let V be a finite-dimensional vector space over a field F, let $T \in \mathcal{L}(V)$, and let m_T be the minimal polynomial of T. Then $p \in \mathcal{P}(F)$ satisfies $p(T) = 0$ iff $p(x)$ is a multiple of $m_T(x)$.

<u>Proof</u>: One direction is obvious: if $p(x) = q(x) m_T(x)$, then

$$p(T) = q(T) m_T(T) = q(T) 0 = 0.$$

Conversely, suppose $p(T) = 0$. By the division algorithm for polynomials, there exist $q, r \in \mathcal{P}(F)$ such that

$$p(x) = q(x)m_T(x) + r(x) \text{ and } (r(x) = 0 \text{ or } \deg(r) < \deg(m_T)).$$

But then

$$p(T) = q(T)m_T(T) + r(T) \Rightarrow r(T) = 0$$

$$\Rightarrow r(x) = 0 \text{ (since } r(T) = 0 \text{ and } \deg(r) < \deg(m_T) \text{ is impossible).}$$

Thus $p(x) = q(x)m_T(x)$. //

What is m_T? If $F = \mathbb{C}$, so that T has a Jordan form, then we can identify m_T explicitly. For the rest of the lecture, assume that $F = \mathbb{C}$.

Recall:

- $V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_k, T)$, where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T.
- Each $G(\lambda_j, T)$ is invariant under T.
- For each $j = 1, \ldots, k$, there exists a smallest positive integer r_j such that $G(\lambda_j, T) = \mathcal{N}((T - \lambda_j I)^{r_j})$. (Note: I previously called this integer m_j.)
- $m_j = \dim(G(\lambda_j, T))$ is called the <u>algebraic multiplicity</u> of λ_j.
- If we choose any basis $\mathcal{B}$ of V consisting of generalized eigenvalues, then $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ is block diagonal. Any such basis is of the form $\mathcal{B} = \bigcup_{j=1}^{k} \mathcal{B}_j$, where $\mathcal{B}_j$ is a basis for $G(\lambda_j, T)$.

- We can always choose each $\mathcal{B}_j$ to be the union of one or more generalized eigenvectors chains:

$$\left\{ (T-\lambda_j I)^{r_{1j}-1}(v_{1j}), (T-\lambda_j I)^{r_{1j}-2}(v_{1j}), \dots, v_{1j}, \quad 1 \le r_{i,j} \le r_j \;\forall i=1,\dots,s_j \right.$$
$$\left. \dots, (T-\lambda_j I)^{r_{s_j,j}-1}(v_{s_j,j}), (T-\lambda_j I)^{r_{s_j,j}-2}(v_{s_j,j}), \dots, v_{s_j,j} \right\}$$

Then the block of $\mathfrak{M}_{\mathcal{B},\mathcal{B}}(T)$ corresponding to $G(\lambda_j, T)$ is itself block diagonal, and each block is a Jordan block:

$$J_{ij} = \begin{bmatrix} \lambda_j & 1 & & \\ & \lambda_j & 1 & \\ & & \ddots & 1 \\ & & & \lambda_j \end{bmatrix} \in \mathbb{C}^{r_{ij} \times r_{ij}}, \quad i=1,\dots,s_j$$

<u>Lemma</u>: Let J be an $r \times r$ Jordan block:

$$J = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{C}^{r \times r}.$$

Then $(J-\lambda I)^t = 0$ iff $t \ge r$.

<u>Proof</u>: Let $\{e_1, \dots, e_r\}$ be the standard basis for $\mathbb{C}^r$. Note that

$$(J-\lambda I)e_1 = 0,$$
$$(J-\lambda I)e_j = e_{j-1}, \quad j = 2, \dots, r.$$

By induction, it is easy to show that

$$(J-\lambda I)^t e_j = 0, \quad j=1,\dots,t,$$

$$(J-\lambda I)^t e_j = e_{j-t}, \quad j=t+1,\dots,r.$$

The result follows. //

<u>Theorem</u>: Let V be a finite-dimensional vector space over $\mathbb{C}$, let $T \in \mathcal{L}(V)$, let $\lambda_1,\dots,\lambda_k$ be the distinct eigenvalues of T, and let

$$G(\lambda_j, T) = \mathcal{N}((T-\lambda_j I)^{r_j}),$$

where r_j is the smallest positive integer for which this holds (that is, where $\mathcal{N}((T-\lambda_j I)^{r_j-1}) \subsetneq \mathcal{N}((T-\lambda_j I)^{r_j})$). Then

$$m_T(x) = (x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2}\cdots(x-\lambda_k)^{r_k}.$$

<u>Proof:</u> Choose a basis $\mathcal{B}$ for V such that $J = \mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ is in Jordan form: $J = \text{diag}(J_1, J_2, \dots, J_\ell)$, where each J_i is a Jordan block. For $p \in \mathcal{P}(\mathbb{C})$, it is easy to show that

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(p(T)) = p(J).$$

($\mathcal{M}_{\mathcal{B},\mathcal{B}}$ is a linear map, so it suffices to prove that $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T^j) = J^j$ for all $j \geq 0$. But $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T^0) = \mathcal{M}_{\mathcal{B},\mathcal{B}}(I) = I = J^0$, $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = J$,

$$\mathcal{M}_{\mathcal{B}}(T^2(v)) = \mathcal{M}_{\mathcal{B}}(T(T(v)))$$

$$= J\mathcal{M}_{\mathcal{B}}(T(v))$$

$$= J^2\mathcal{M}_{\mathcal{B}}(v) \quad \forall v \in V$$

$$\Rightarrow \mathcal{M}_{\mathcal{B},\mathcal{B}}(T^2) = J^2,$$

etc.) Also, it is easy to show that

$$p(J) = \text{diag}(p(J_1), p(J_2), \ldots, p(J_\ell))$$

(similarly, it suffices to prove that $J^j = \text{diag}(J_1^j, J_2^j, \ldots, J_\ell^j)$).

Now, if $\lambda \neq \lambda_j$, then $J_j - \lambda I$ is invertible (upper triangular with nonzeros on the diagonal); hence

$$(J_j - \lambda_i I)^t \text{ is nonsingular } \forall t \geq 0 \ \forall i \neq j.$$

Now, since λ is a root of m_T iff λ is an eigenvalue of T, and since m_T can be fully factored over $\mathbb{C}$, we have

$$m_T(x) = (x-\lambda_1)^{c_1}(x-\lambda_2)^{c_2}\cdots(x-\lambda_k)^{c_k} \text{ for some } c_1, \ldots, c_k \in \mathbb{Z}^+.$$

It follows that

$$m_T(J_j) = \prod_{i=1}^{k}(J_j - \lambda_i I)^{c_i}$$

$$= \left(\prod_{\substack{i=1 \\ i\neq j}}^{k}(J_j - \lambda_i I)^{c_i}\right)(J_j - \lambda_j I)^{c_j},$$

So

$$m_T(J_j) = 0 \iff \left(\prod_{\substack{i=1 \\ i \neq j}}^{k} (J_j - \lambda_i I)^{c_i} \right) (J_j - \lambda_j I)^{c_j}$$

$$\iff (J_j - \lambda_j I)^{c_j} = 0 \quad \Big(\text{since } \prod_{\substack{i=1 \\ i \neq j}}^{k} (J_j - \lambda_i I)^{c_i} \text{ is invertible}\Big)$$

$$\iff c_j \geq r_j.$$

Since m_T has the minimal degree of any polynomial satisfying $m_T(T) = 0$, we must have $c_j = r_j$ for all j, that is,

$$m_T(x) = (x-\lambda_1)^{r_1}(x-\lambda_2)^{r_2} \cdots (x-\lambda_k)^{r_k}. \quad //$$