Lecture 21: Characteristic polynomials, generalized eigenspaces, and Jordan canonical form over $\mathbf{F} = \mathbb{C}$ (1)

Travis Schedler

Tue, Nov 30, 2010 (version: Tue, Dec 7, 10:53 AM)

Goals (2)

• Characteristic polynomial
• Generalized eigenspaces
• The decomposition theorem
• Cayley-Hamilton theorem
• Jordan canonical form
• Minimal polynomial

Note: we are back to an arbitrary finite-dimensional vector space $V$ (not inner product) together with $T \in \mathcal{L}(V)$. We let $\mathbf{F} = \mathbb{C}$ throughout the lecture. Let $\dim V = n$.

Characteristic polynomial (3)

• Pick a basis in which $\mathcal{M}(T)$ is upper-triangular.
• Let $\lambda_1, \ldots, \lambda_n$ be the diagonal entries.
• Definition: the characteristic polynomial $\chi_T(x) = (x - \lambda_1) \cdots (x - \lambda_n)$.
• Theorem: The characteristic polynomial is well-defined: it does not depend on the choice of basis.
• We will first give a better definition of characteristic polynomial (that doesn’t depend on a basis) and then show it coincides with this one.
Generalized eigenspaces (4)

Definition 1. A generalized eigenvector $v$ of $T$ of eigenvalue $\lambda$ is one such that $(T - \lambda I)^k v = 0$ for some $k \geq 1$.

- The generalized eigenvectors of eigenvalue $\lambda$ form a vector subspace $V(\lambda)$ containing the eigenspace $U(\lambda)$ of $\lambda$.
- $V(\lambda) \neq 0$ iff $U(\lambda) \neq 0$.
- $\dim V(\lambda) = \dim \text{null} (T - \lambda I)^k$ for large enough $k$ (e.g., one that works for every element of a basis of $V(\lambda)$).
- Lemma: $V(\lambda) = \text{null} (T - \lambda I)^{\dim V(\lambda)}$.
- Proof: $S := T - \lambda I$ is nilpotent restricted to $V(\lambda)$. Then, $S^k V(\lambda) = 0$ for some $k$ implies $S^{\dim V(\lambda)} V(\lambda) = 0$: $S^j V(\lambda) \supseteq S^{j+1} V(\lambda)$ if $S^j V(\lambda) \neq 0$.

The decomposition theorem (5)

Theorem 2 (The decomposition theorem, Theorem 8.23). $V = \bigoplus \lambda V(\lambda)$.

- Claim: $V = \text{null} (T - \lambda I)^{\dim V(\lambda)} \oplus \text{range} (T - \lambda I)^{\dim V(\lambda)}$.
- Proof: Enough to show $\text{null} (T - \lambda I)^{\dim V(\lambda)} \cap \text{range} (T - \lambda I)^{\dim V(\lambda)} = 0$.
- If $v$ is in the intersection, $v = (T - \lambda I)^{\dim V(\lambda)} u$ for some $u \in \text{null} (T - \lambda I)^{2 \dim V(\lambda)}$. So $u$ is a generalized eigenvector, in $V(\lambda) = \text{null} (T - \lambda I)^{\dim V(\lambda)}$. Thus $v = 0$.
- Let $\lambda$ be an eigenvalue (which exists since $\mathbf{F} = \mathbb{C}$). Then $V(\lambda) \neq 0$. By induction on $\dim V$, induction, $\text{range} (T - \lambda I)^{\dim V(\lambda)}$ is a direct sum of generalized eigenspaces of $T|_{\text{range} (T - \lambda I)^{\dim V(\lambda)}}$.

Completion of proof of the decomposition theorem (6)

- It remains to show that, if $\nu \neq \lambda$, then $V(\nu) \subseteq \text{range} (T - \lambda I)^{\dim V(\lambda)}$.
- Since $V(\nu)$ is $T$-invariant, $V(\nu) \supseteq (T - \lambda I)^{\dim V(\lambda)} V(\nu) = \text{range} (T - \lambda I)^{\dim V(\lambda)}|_{V(\nu)}$.
- So we need to show $\text{null} (T - \lambda I)^{\dim V(\lambda)}|_{V(\nu)} = 0$.
- This nullspace is $V(\nu) \cap V(\lambda)$. If this intersection is nonzero, then it contains both a nonzero $\nu$-eigenvector and a nonzero $\lambda$-eigenvector.
- But no nonzero $\nu$-eigenvector $v$ can be in $V(\lambda)$, since $(T - \lambda I)^k v = (\nu - \lambda)^k v \neq 0$ for all $k \geq 0$. So the intersection is zero. □
The Cayley-Hamilton theorem (7)
Second definition of characteristic polynomial: $$\hat{\chi}_T(x) = \prod \lambda(x - \lambda)^{\dim V(\lambda)}.$$ (We will prove later this coincides with the previous definition).

**Theorem 3** (Cayley-Hamilton theorem; Theorem 8.20 (for $$\chi_T(T)$$)). $$\hat{\chi}_T(T) = 0.$$  

**Proof.**
- Since $$V = \bigoplus \lambda V(\lambda)$$, it suffices to show that $$\chi_T(T)|_{V(\lambda)} = 0$$ for all $$\lambda$$.
  - But $$(T - \lambda I)^{\dim V(\lambda)}$$ is already zero on $$V(\lambda)$$, and this is a factor of $$\chi_T(T)$$.

Equality of characteristic polynomials (8)

**Theorem 4** (Theorem 8.10). $$\chi_T(T) = \hat{\chi}_T(T).$$  

So $$\chi_T(T)$$ does not depend on the choice of basis! Let $$M_\lambda$$ denote the number of times $$\lambda$$ appears on the diagonal of $$M(T)$$ (for a fixed basis) and let $$N_\lambda$$ denote the dimension of $$V(\lambda)$$. We need to show $$M_\lambda = N_\lambda$$.

**Proof.**
- By Gaussian elimination, if $$A$$ is upper-triangular, $$\dim \text{null}(A) \leq$$ the number of zeroes on the diagonal.
  - Now, for all $$k$$, 0 appears on the diagonal of $$M(T - \lambda I)^k$$ the same number of times it appears on the diagonal of $$M(T - \lambda I)$$, which is $$M_\lambda$$. So, $$\dim V(\lambda) = N_\lambda \leq M_\lambda$$ for all $$\lambda$$.
  - We know from the decomposition theorem that $$\sum \lambda N_\lambda = \dim V$$. But this also equals $$\sum \lambda M_\lambda$$. Since $$N_\lambda \leq M_\lambda$$ for all $$\lambda$$, we conclude $$N_\lambda = M_\lambda$$ for all $$\lambda$$.

Block diagonal form (9)

What happens if we put $$T$$ in a basis obtained from bases of each of the $$V(\lambda)$$?

**Theorem 5** (Theorem 8.28). In this basis, $$M(T)$$ is block diagonal. We can pick the basis so that the block corresponding to $$V(\lambda)$$ is upper triangular with $$\lambda$$ on the diagonal.

**Proof.**
- Since $$V(\lambda)$$ is $$T$$-invariant, $$M(T)$$ is indeed block diagonal with blocks corresponding to each $$V(\lambda)$$ (of size $$\dim V(\lambda)$$).
  - Since $$F = \mathbb{C}$$, we can pick the bases so that $$M(T|_{V(\lambda)})$$ is upper-triangular.
  - Then since the diagonal entries are the eigenvalues, these must all be $$\lambda$$, as $$T|_{V(\lambda)}$$ has only $$\lambda$$ as an eigenvalue.
Jordan canonical form (10)
We can do much better than block-diagonal with upper-triangular blocks having the same diagonal entries (still $\mathbb{F} = \mathbb{C}$):

**Theorem 6.** In some basis, $T$ has block diagonal form with blocks
\[
\begin{pmatrix}
\lambda & 1 \\
& \ddots & \ddots \\
& & \ddots & 1 \\
& & & \lambda
\end{pmatrix}.
\]

In the case that $\mathbb{F} \neq \mathbb{C}$, we have the more general results we won’t be able to prove:

- For $\mathbb{F} = \mathbb{R}$, we can say the same thing if we also allow $\lambda$ to be \(\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\), and then 1 becomes $I$ (alternatively, \(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\)).
- For general $\mathbb{F}$, we can say the same if we allow $\lambda$ to be a matrix with irreducible characteristic polynomial, and replace 1 by \(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\).

Minimal polynomial (11)

**Definition 7.** The minimal polynomial of $T$ is the monic polynomial $p(x)$ of smallest degree such that $p(T) = 0$.

Why does the minimal polynomial exist? If $p, q$ are two monic polynomials of the same degree such that $p(T) = 0 = q(T)$, then $(p - q)(T) = 0$ as well, so $p = q$.

**Theorem 8** (Theorem 8.34). A polynomial $q(x)$ has the property that $q(T) = 0$ iff $q$ is a multiple of the minimal polynomial $p(x)$.

**Proof.** This is a consequence of a general fact about polynomials: Given polynomials $p(x)$ and $q(x)$, we can always write $q(x) = a(x)p(x) + b(x)$ where $\deg b(x) < \deg p(x)$.

- Now, $q(T) = 0$ iff $b(T) = 0$, which is true iff $b(x)$ is the zero polynomial (by definition of $p(x)$).

More on minimal polynomials (12)
We saw that $\chi_T(T) = 0$. Hence the minimal polynomial $p(x)$ divides the characteristic polynomial $\chi_T(x)$.

**Theorem 9** (Theorem 8.36). The roots of $p(x)$ are exactly the eigenvalues of $T$.

**Proof.** Since $p(x)$ divides $\chi_T(x)$, only the eigenvalues can appear as roots.

- On the other hand, if $\lambda$ is an eigenvalue of $T$, and $v$ is a nonzero eigenvector, then $p(T)v = p(\lambda)v$, so this is zero iff $\lambda$ is a root of $p(x)$. 

4
Proof of Jordan canonical form (13)

Jordan canonical form is based on the following key result. Let $N \in L(V)$ be nilpotent and for nonzero $v \in V$, let $m(v)$ be the maximum positive integer such that $N^{m(v)}v \neq 0$.

**Lemma 10** (Lemma 8.40). There exist vectors $v_1, \ldots, v_k \in V$ such that $(v_1, Nv_1, \ldots, N^{m(v_1)}v_1, \ldots, v_k, Nv_k, \ldots)$ is a basis of $V$.

Note: the condition implies that $(N^{m(v_1)}v_1, \ldots, N^{m(v_k)}v_k)$ is a basis of $\text{null } N$ (consider any linear combination sent to zero by $N$).

- We prove by induction on $\dim V$. Since range $N \subseteq V$, we can assume that $N|_{\text{range } N}$ has such vectors $u_1, \ldots, u_j$.
- Since $u_1, \ldots, u_j \in \text{range } N$, we can pick $v_1, \ldots, v_j$ with $Nv_i = u_i$.
- Furthermore, let $v_{j+1}, \ldots, v_k$ be vectors which extend $(N^{m(u_1)}u_1, \ldots, N^{m(u_j)}u_j)$ to a basis of $\text{null } N$.

Completion of proof of the lemma (14)

- Claim: $v_1, \ldots, v_k$ give a basis of the desired form.
- We show linear independence. Suppose $\sum_{i,j} a_{i,j} N^j v_i = 0$. Applying $N$ yields $\sum_{i,j} a_{i,j} N^j u_i = 0$.
- By assumption that the $u_i$ give a basis of range $N$, we have $a_{i,j} = 0$ whenever $j \leq m(u_i) = m(v_i) - 1$.
- Thus the only nonzero coefficients are those of $(N^{m(v_1)}v_1, \ldots, N^{m(v_k)}v_k)$.
- But, these form a basis of null $N$. So all the coefficients are zero.
- Now, the length of the linearly independent list is $\dim \text{range } N + k = \dim \text{range } N + \dim \text{null } N = \dim V$, so it must be a basis. 

Proof of Jordan canonical form (Theorem 8.47) (15)

- In the basis of the lemma, $M(N)$ is block diagonal with $k$ blocks of sizes $m(v_1), \ldots, m(v_k)$, each of the form

\[
\begin{pmatrix}
0 & 1 \\
\vdots & \ddots \\
\vdots & \ddots & 1 \\
0 & & & & & & & & \end{pmatrix},
\]

- Now, for a general operator $T$, we can choose a basis as above for each $V(\lambda)$, in terms of the nilpotent operator $T - \lambda I$. Then, the matrix has the desired form.
Uniqueness of Jordan canonical form (16)

**Theorem 11.** For every $T$, there is a unique Jordan canonical form $\mathcal{M}(T)$, up to rearranging the order of the blocks.

**Proof.**

- Any Jordan matrix is obtained from bases of each of the $V(\lambda)$. So, it suffices to assume there is only one eigenvalue $\lambda$.
- Up to replacing $T$ with $T - \lambda I$, we can assume $T = N$ is nilpotent.
- Now, we need only show that the positive integers $m(v_1), \ldots, m(v_k)$ are unique up to permutation.
- But, $\dim \ker N = k$, $\dim \ker N^2 = |\{j : m(v_j) \geq 2\}| + k$, and in general, $\dim \ker N^r = \sum_{s=0}^{r} |\{j : m(v_j) \geq s\}|$.
- These determine the $m(v_j)$ and are independent of bases. 

Corollary: two matrices are conjugate iff they have the same Jordan canonical form (up to permuting blocks).