Lecture 18: Cayley-Hamilton and bilinear forms

We begin this lecture with the proof of Cayley-Hamilton. Recall the statement.

Theorem 0.1 (Cayley-Hamilton)**.** *Let $T : V \to V$ be linear with dim $V < \infty$. Writing c_T for the characteristic polynomial of T*

$$c_T(T) = 0 \ .$$

Proof. We will only give a proof in the case that $\mathbb{F}$ is algebraically closed. The general case follows by doing a field extension. (You can see many different proofs online or in textbooks though.) Since $\mathbb{F}$ is algebraically closed we may factor the characteristic polynomial

$$c_T(x) = (x - \lambda_1)^{n_1} \cdots (x - \lambda_k)^{n_k} \ ,$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues. Let B be a basis such that $[T]_B^B$ is in Jordan form. Last lecture it was shown that n_i is equal to the dimension of $\hat{E}_{\lambda_i}$.

So taking $v \in B$ such that $v \in \hat{E}_{\lambda_i}$, we get

$$(c_T(T))(v) = (T - \lambda_1 I)^{n_1} \cdots (T - \lambda_k I)^{n_k}(v) = \left(\prod_{j \neq i} (T - \lambda_j I)^{n_j} \right) (T - \lambda_i I)^{n_i}(v) \ .$$

In last lecture it was shown that if $(T - \lambda_i I)^{\dim \hat{E}_{\lambda_i}}$ restricted to $\hat{E}_{\lambda_i}$ is zero. Therefore $c_T(T)(v) = \vec{0}$. Since $c_T(T)$ sends all basis vectors to $\vec{0}$, it must be the zero operator. $\square$

We can now finish with the uniqueness of Jordan form.

Theorem 0.2 (Uniqueness of Jordan form)**.** *Let $T : V \to V$ be linear with $\mathbb{F}$ algebraically closed. If B and B' are bases such that $[T]_B^B$ and $[T]_{B'}^{B'}$ are in Jordan form, these matrices are equal up to permutation of blocks.*

Proof. Write $\lambda_1, \ldots, \lambda_k$ for the distinct eigenvalues of T. Setting S_i as the generalized eigenvectors in B corresponding to λ_i and S_i' the same for B', we have from last lecture that

$$S_i, S_i' \text{ are bases for } \hat{E}_{\lambda_i} \ .$$

Therefore both S_i and S_i' are chain bases for $T - \lambda_i I$ restricted to $\hat{E}_{\lambda_i}$. By the uniqueness of nilpotent form, the number of chains of each length is the same in S_i, S_i'. This means the number of Jordan blocks of each size for λ_i is the same in B and B'. This is true for all i, and proves the theorem. $\square$

We now move to bilinear forms, which are 2-linear functions.

Definition 0.3. *A 2-linear function $f : V \times V \to \mathbb{F}$ is called a bilinear form. The set $Bil(V, \mathbb{F})$ of bilinear forms on V forms a vector space.*

If V is finite dimensional then we can find a matrix form for f in terms of a basis.

Theorem 0.4. *Let f be a bilinear form on V and B a basis for V. Define the matrix $[f]_B^B$ by*

$$\left([f]_B^B\right)_{i,j} = f(v_j, v_i) \ .$$

Then for all $v, w \in V$, we have

$$[w]_B^t [f]_B^B [v]_B \ .$$

Furthermore, $[f]_B^B$ is the unique matrix such that this equation holds for all $v, w \in V$.

Proof. To show this, let B be a basis for V and write $B = \{v_1, \ldots, v_n\}$. If $v, w \in V$, write

$$v = a_1 v_1 + \cdots + a_n v_n \text{ and } w = b_1 v_1 + \cdots + b_n v_n \ .$$

Then

$$\begin{aligned} f(v,w) = f(a_1 v_1 + \cdots + a_n v_n, w) = \sum_{i=1}^n a_i f(v_i, w) &= \sum_{i=1}^n a_i f(v_i, b_1 v_1 + \cdots + b_n v_n) \\ &= \sum_{i=1}^n a_i \left[\sum_{j=1}^n b_j f(v_i, v_j) \right] \\ &= \sum_{i=1}^n ([v]_B)_{i,1} \sum_{j=1}^n ([w]_B)_{1,j}^t \left([f]_B^B\right)_{j,i} \\ &= \sum_{i=1}^n \left([w]_B^t [f]_B^B\right)_{1,i} ([v]_B)_{i,1} \\ &= [w]_B^t [f]_B^B [v]_B \ . \end{aligned}$$

For uniqueness, note that if A is any matrix such that $f(v,w) = [w]_B^t A [v]_B$ for all v, w, we apply this to v_i, v_j:

$$f(v_j, v_i) = [v_i]_B^t A [v_j]_B = e_i^t A e_j = A_{i,j} \ .$$

Thus A has the same entries as those of $[f]_B^B$. $\square$

Remarks.

- If we define the *standard dot product* of two vectors $\vec{a}, \vec{b} \in \mathbb{F}^n$ by

  $$\vec{a} \cdot \vec{b} = a_1 b_1 + \cdots + a_n b_n \ ,$$

  where we have written the vectors $\vec{a} = (a_1, \ldots, a_n)$ and $\vec{b} = (b_1, \ldots, b_n)$, then we can write the above result as

  $$f(v,w) = \left([f]_B^B [v]_B\right) \cdot [w]_B \ .$$

- Given any $A \in M_{n,n}(\mathbb{F})$ and basis B of V (of size n), the function f_A, given by

  $$f_A(v,w) = [w]_B^t A [v]_B$$

  is bilinear (check this!) and has A for its matrix relative to B. To see this, we compute the (i,j)-th entry of $[f_A]_B^B$: it is

  $$f_A(v_j, v_i) = [v_i]_B^t A [v_j]_B = e_i^t A e_j = A_{i,j} \ .$$

- An important example comes from $A = I$. This corresponds to the dot product in basis B:

  $$(v,w) \mapsto [w]_B^t [v]_B = [v]_B \cdot [w]_B \ .$$

- Given B a basis of V, the map $f \mapsto [f]_B^B$ is an isomorphism.

  Proof. If $f, g \in Bil(V, \mathbb{F})$ and $c \in \mathbb{F}$,

  $$\left([cf+g]_B^B\right)_{i,j} = (cf+g)(v_j, v_i) = cf(v_j, v_i) + g(v_j, v_i) = c\left([f]_B^B\right)_{i,j} + \left([g]_B^B\right)_{i,j} \ ,$$

  so $[cf+g]_B^B = c[f]_B^B + [g]_B^B$, proving linearity. If $[f]_B^B$ is the zero matrix, then

  $$f(v,w) = [w]_B^t [f]_B^B [v]_B = 0 \text{ for all } v, w \in V \ ,$$

  so f is zero, proving injectivity. Last, if A is a given matrix in $M_{n,n}(\mathbb{F})$, we showed above that $[f_A]_B^B = A$. This means that the map is surjective and we are done. $\square$

Definition 0.5. *The rank of a bilinear form f on a finite-dimensional vector space V is defined as rank$([f]_B^B)$ in any basis B.*

To show this is well-defined will take a bit of work. Given a bilinear form f on V we can fix a vector $v \in V$ and define

$$L_f(v) : V \to \mathbb{F} \text{ by } (L_f(v))(w) = f(v,w) \ .$$

Since f is bilinear, this is a linear functional and thus lives in V^*. Let $Bil(V, \mathbb{F})$ be the vector space of bilinear forms on V.

We can nicely represent the matrix of L_f relative to the basis B and the dual basis B^*.

Proposition 0.6. *$L_f : V \to V^*$ is a linear transformation and*

$$[L_f]^B_{B^*} = [f]^B_B \ .$$

Proof. Let $v_1, v_2 \in V$ and $c \in \mathbb{F}$. To show that $L_f(cv_1 + v_2) = cL_f(v_1) + L_f(v_2)$, we will need to apply these functionals to vectors in V. So let $w \in V$ and compute

$$\begin{aligned} L_f(cv_1 + v_2)(w) = f(cv_1 + v_2, w) &= cf(v_1, w) + f(v_2, w) \\ &= cL_f(v_1)(w) + L_f(v_1)(w) = (cL_f(v_1) + L_f(v_2))(w) \ . \end{aligned}$$

This is true for all w, so L_f is linear.

The i-th column of $[L_f]^B_{B^*}$ is obtained by writing

$$B = \{v_1, \ldots, v_n\}, \ B^* = \{v_1^*, \ldots, v_n^*\}$$

and writing $L_f(v_i)$ in terms of B^*. Recall that if g is a linear functional then its representation in terms of the dual basis can be written

$$f = f(v_1)v_1^* + \cdots + f(v_n)v_n^* \ .$$

Therefore we can write

$$\begin{aligned} L_f(v_i) &= L_f(v_i)(v_1)v_1^* + \cdots + L_f(v_i)(v_n)v_n^* \\ &= f(v_i, v_1)v_1^* + \cdots + f(v_i, v_n)v_n^* \ . \end{aligned}$$

This means that the (j,i)-th entry of $[L_f]^B_{B^*}$ is (v_i, v_j), the (j,i)-th entry of $[f]^B_B$. $\square$