Weekly Assignment 4 Solutions

Jadyn V. Breland
MATH 117: Advanced Linear Algebra

August 30, 2023

In this assignment, you will prove the Cayley-Hamilton theorem. Throughout, V is a finite-dimensional vector space and $L : V \to V$ is a linear operator. See the weekly assignment webpage for due dates, templates, and assignment description. Make sure to justify any claims you make. You may not appeal to any results that we have not discussed in class.

1. Suppose that W is an L-invariant[1] subspace of V. Then we obtain by restriction a linear operator $L|_W : W \to W$. Prove that the characteristic polynomial of $L|_W$ divides the characteristic polynomial of L.[2]

Proof. Let $A = (b_1, \ldots, b_k)$ be a basis for W and extend to a basis $B = (b_1, \ldots b_n)$ for V. For each $i = 1, \ldots, n$, write

$$[L(b_i)]_B = (\alpha_{1i}, \ldots, \alpha_{ni}).$$

Then since W is L-invariant, $L(b_i) \in W$ for all $i = 1, \ldots, k$ which implies that $\alpha_{ji} = 0$ for all $k < j \leq n$. Thus, there exist matrices $C \in F^{(n-k)\times k}$, and $D \in F^{(n-k)\times(n-k)}$ such that

$$[L]_B = \begin{pmatrix} [L\ |_W]_A & C \\ 0 & D \end{pmatrix}.$$

Then we see that

$$\begin{aligned} c_L(x) &= \det(xI_n - [L]_B) \\ &= \begin{pmatrix} xI_k - [L|_W]_A & -C \\ 0 & xI_{n-k} - D \end{pmatrix} \\ &= c_{L|_W}(x) c_D(x). \end{aligned}$$

This proves that $c_{L|_W}(x)$ divides $c_L(x)$, as claimed. Hopefully, you recognize this argument from the notes. □

[1] See Def 4.3.6

[2] Hint: start with a basis for W and extend to a basis B for V. Argue that $[L]_B = \left(\begin{smallmatrix} X & Y \\ 0 & Z \end{smallmatrix}\right)$ for some matrices X, Y, Z. Compute $c_L(x)$ using this fact.

2. For $v \neq 0$, define $\operatorname{span}(L, v) := \operatorname{span}\{L^i(v) : i \in \mathbb{N}_0\}$. Here $L^0 := \mathrm{Id}_V$.

(a) Prove that $\operatorname{span}(L, v)$ is an L-invariant subspace of V.

(b) Prove that there exists a largest integer k such that $B_k := \{v, L(v), \ldots, L^{k-1}(v)\}$ is independent. Moreover, show that B_k is a basis for $\operatorname{span}(L, v)$.

Proof. (a) We need to show that $L(\operatorname{span}(L, v)) \subseteq \operatorname{span}(L, v)$. Since L is a linear map, it suffices to show that $L(\{L^i(v) : i \in \mathbb{N}_0\}) \subseteq \operatorname{span}(L, v)$. But this is obvious because $L(L^i(v)) = L^{i+1}(v) \in \operatorname{span}(L, v)$ for any $i \in \mathbb{N}_0$.

(b) The set

$$\{n \in \mathbb{N} : B_n \text{ is independent}\}$$

is nonempty since B_1 is independent and is bounded from above since V is finite-dimensional. Therefore, the set has a maximum element k.

To see that B_k spans $\operatorname{span}(L, v)$, it suffices to show that $L^i(v) \in \operatorname{span}(B_k)$ for all $i \geq k$. We prove this claim by induction on i. The claim is true for $i = k$, because B_k is independent, while the set

$$B_{k+1} = \{v, L(v), \ldots, L^{k-1}(v), L^k(v)\}$$

is dependent (by maximality of K). In particular, $\operatorname{span}(B_{k+1}) = \operatorname{span}(B_k)$. Let $i \geq k$ and assume that $L^{i-1}(v) \in \operatorname{span}(B_k)$. Then

$$L^{i-1}(v) = \sum_{j=0}^{k-1} \alpha_i L^i(v)$$

for some $\alpha_0, \ldots, \alpha_{k-1} \in F$. Thus,

$$L^i(v) = L\left(\sum_{j=0}^{k-1} \alpha_j L^j(v)\right) = \sum_{j=0}^{k-1} L^{j+1}(v) = \sum_{j=1}^{k} L^j(v).$$

Thus, $L_i(v) \in \operatorname{span}(B_{k+1}) = \operatorname{span}(B_k)$.

□

3. Let $v \neq 0$ and let $B := B_k$ be the basis for $W := \operatorname{span}(L, v)$ from Problems 1 & 2. Define $m_{L,v}(x) = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} + x^k \in F[x]$ where $a_0, a_1, \ldots, a_{k-1}$ are the unique coefficients such that

$$a_0 v + a_1 L(v) + \cdots + a_{k-1}L^{k-1}(v) + L^k(v) = 0.$$

Since W is L-invariant, we obtain by restriction a linear operator $L|_W : W \to W$. The goal of this problem is to show that the characteristic polynomial of $L|_W$ is equal to $m_{L,v}(x)$.

(a) Compute the matrix $xI_k - [L|_W]_B$.

(b) Show that $\det(xI_k - [L|_W]_B) = m_{L,v}(x)$ using induction on k. [3]

[3] Hint: For the inductive step, start by using cofactor expansion along the first row.

Proof. (a) This part is a straightforward computation. Observe that $L(L^i(v)) = L^{i+1}(v)$ for all $0 \le i \le k-1$. Thus,

$$[L\mid_W]_B = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & -a_0 \\ 1 & 0 & 0 & & 0 & 0 & -a_1 \\ 0 & 1 & 0 & & 0 & 0 & -a_2 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & & 1 & 0 & -a_{k-2} \\ 0 & 0 & 0 & \cdots & 0 & 1 & -a_{k-1} \end{pmatrix}.$$

Thus,

$$xI_k - [L\mid_W]_B = \begin{pmatrix} x & 0 & 0 & \cdots & 0 & 0 & a_0 \\ -1 & x & 0 & & 0 & 0 & a_1 \\ 0 & -1 & x & & 0 & 0 & a_2 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & & -1 & x & a_{k-2} \\ 0 & 0 & 0 & \cdots & 0 & -1 & x+a_{k-1} \end{pmatrix}.$$

(b) The base case is $k = 1$. In that case, $\det(xI_1 - [L\mid_W]_B) = \det(x + a_0) = x + a_0$. Now, let $k \ge 1$ and assume that the statement holds for $k-1$. Then, using cofactor expansion along the first row, we obtain

$$\begin{aligned} \det(xI_k - [L\mid_W]_B) &= x\begin{vmatrix} x & 0 & & 0 & 0 & a_1 \\ -1 & x & & 0 & 0 & a_2 \\ \vdots & & \ddots & & & \vdots \\ 0 & 0 & & -1 & x & a_{k-2} \\ 0 & 0 & \cdots & 0 & -1 & x+a_{k-1} \end{vmatrix} + (-1)^{k-1}a_0 \begin{vmatrix} -1 & x & 0 & & 0 & 0 \\ 0 & -1 & x & & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & & -1 & x \\ 0 & 0 & 0 & \cdots & 0 & -1 \end{vmatrix} \\ &= x(x^{k-1} + a_{k-1}x^{k-2} + \cdots + a_2x + a_1) + (-1)^{2k-2}a_0 \\ &= x^k + a_{k-1}x^{k-1} + \cdots + a_1x + a_0. \end{aligned}$$

We used the inductive hypothesis to compute the determinant of the first cofactor.

□

4. Use Problems 1-3 to prove the Cayley-Hamilton theorem: the linear operator L is a root of its characteristic polynomial, that is, $\mathrm{c}_L(L)$ is the zero operator in $\mathrm{End}(V)$.[4]

Proof. We need to show that $c_L(L)(v) = 0$ for all $v \in V$. If $v = 0$, then this is trivial because $c_L(L)$ is a linear map. Assume $v \ne 0$. Then $W = \mathrm{span}(L, v)$ is an L-invariant subspace, so L restricts to a linear operator $L\mid_W$ on W. By Problem 3, $c_{L|_W}(x) = m_{L,v}(x)$. Notice that $m_{L,v}(L)(v) = 0$ by definition of $m_{L,v}(x)$. Finally, using Problem 1, $m_{L,v}(x)$ divides $c_L(x)$, so there exists $p(x) \in F[x]$ such that $c_L(x) = p(x)m_{L,v}(x)$. Then

$$c_L(L)(v) = p(L)(m_{L,v}(v)) = p(L)(0) = 0.$$

This proves that $c_L(L)$ is the zero operator.

□

[4]Hint: need to show $c_L(L)(v) = 0$ for all $v \in V$. Case 1: $v = 0$. Case 2: $v \ne 0$, invoke Problems 1 & 3 to write $c_L(x)$ as a product of two polynomials. Use this to show $c_L(L)(v) = 0$.