Minimum Polynomials of Linear Transformations

Spencer De Chenne

20 April 2014

1 Introduction and Preliminaries

Throughout algebra, we occupy ourselves by studying the structure of various objects, even when a proper definition of "structure" is elusive. In essence, the goal of analyzing the structure of an algebraic object is to know as much about the object provided as little information as possible. In this quest, we have developed many different techniques to do so. Tools in our current linear algebra arsenal include understanding the implications of two vector spaces being isomorphic, two linear transformations being similar, etc. One specific and useful tool used frequently in various areas of algebraic study which we have largely left untouched is the polynomial.

Polynomials

The focus of this paper is on the minimum polynomial of a linear transformation, and the various consequences which arise when studying the minimum polynomial. However, before we begin, we shall review basic information about polynomials in general. While typical first courses in linear algebra focus on vectors spaces with scalars from the field $\mathbb{C}$, we shall focus on vector spaces with scalars from any field $\mathbb{F}$. Thus, in the course of this paper, the polynomials in question will always be elements of $\mathbb{F}[x]$. Recall that one of the primary differences between studying vector spaces over $\mathbb{C}$ and vector spaces over $\mathbb{F}$ comes from irreducible factors, and consequences of the field $\mathbb{C}$ being algebraically closed. Hence, any polynomial in $\mathbb{C}[x]$ can be factored into linear factors, and so irreducible elements in $\mathbb{C}[x]$ are trivial. This, of course, is not true in the general $\mathbb{F}[x]$ case. We shall review some basic information about polynomials from $\mathbb{F}[x]$, but we omit the proofs of theorems as the material should be review.

Definition A polynomial $f(x) \in \mathbb{F}[x]$, given by $f(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$, is monic if $a_n = 1$.

Recall that given any polynomial $p(x) \in \mathbb{F}[x]$, we can "make the polynomial monic" by dividing by the leading term. This process will preserve roots of the polynomial while allowing for important properties, such as uniqueness, which we will rely on extensively.

Theorem 1.1. (Division Theorem) *Given $f(x), g(x) \in \mathbb{F}[x]$, with $g(x)$ monic, there exist unique polynomials $q(x)$ and $r(x)$, with $\deg r(x) < \deg g(x)$, such that*

$$f(x) = q(x)g(x) + r(x).$$

The Division Theorem is reminiscent of the division algorithm for integers. Similarly, the following definition and theorem are similar to the Fundamental Theorem of Arithmetic for integers.

Definition Polynomials $f(x)$ and $g(x)$ are relatively prime if their greatest common divisor is 1. A polynomial $f(x)$ is irreducible if the only monic non-constant polynomial dividing f is a scalar multiple of f.

Theorem 1.2. *Let $p(x) \in \mathbb{F}[x]$ be a non-constant monic polynomial. Then $p(x)$ is a unique (up to order) product of irreducible polynomials.*

We have now provided the equipment to discuss polynomials in general. Before we tackle minimum polynomials, we turn our attention to preliminaries for linear transformations.

Linear Transformations

Given a vectors space V over a field $\mathbb{F}$, we can define an endomorphism on V as a linear map from V to V. It may come as no surprise that the set of all endomorphisms on V, $\mathrm{End}(V)$, forms a vector space, if we define scalar multiplication in the canonical way. If V has dimension n, then $\mathrm{End}(V)$ has dimension n^2 (which can easily be seen by noticing the correspondence between $\mathrm{End}(V)$ and the set of all $n \times n$ matrices). We shall now define a new algebraic structure, an *algebra*, and show that $\mathrm{End}(V)$ is an algebra.

Definition An algebra is a vector space V with a bilinear product

$$\begin{aligned} V \times V &\to V \\ (v, w) &\to vw \end{aligned}$$

that distributes over vector addition

$$\begin{aligned} u(v+w) &= uv + uw, \quad \text{and} \\ (v+w)u &= vu + wu \end{aligned}$$

and satisfies $v(aw) = a(vw) = (av)w$ for all $a \in \mathbb{F}$.

As the reader may notice, the definition of an algebra is similar to that of a ring. In fact, every algebra forms a ring. Rather than referring to this new operator as the bilinear product, we will simply refer to it as multiplication, although this is not to be confused with scalar multiplication.

We shall now show that $\mathrm{End}(V)$ is, in fact, an algebra. For $T, R \in \mathrm{End}(V)$, we define multiplication via

$$TR = T \circ R.$$

Clearly, for every $v \in V$ and $R, S, T \in \mathrm{End}(V)$, we can see

$$(R(S+T))(v) = R(S(v) + T(v)) = (RS)(v) + (RT)(v) = (RS + RT)(v),$$

and

$$((S+T)R)(v) = (S+T)R(v) = (SR)(v) + (TR)(v) = (SR+TR)(v),$$

due to the linearity of R, S, and T. Then, for $a \in \mathbb{F}$, again due to the linearity of each element, we can see that $R(aT) = a(RT) = (aR)T$. Therefore, as we have already claimed $\text{End}(V)$ is a vector space, we can see that $\text{End}(V)$ is also an algebra.

Now that we have defined multiplication for endomorphisms, we can combine the theory of endomorphisms and polynomials. If we let $p(x)$ be a polynomial in $\mathbb{F}[x]$ and T be an endomorphism, then we can see that $p(T)$ is again an endomorphism, owing to $\text{End}(V)$ being an algebra. That is, for $v, w \in V$ and $a \in \mathbb{F}$,

$$\begin{aligned} p(T)(v+w) &= p(T)(v) + p(T)(w), \quad \text{and} \\ p(T)(av) &= ap(T)(v). \end{aligned}$$

We have now constructed the tools necessary to study minimum polynomials of linear transformations.

2 Annihilating, Minimum, and Characteristic Polynomials

Let V be a finite-dimensional vector space and T an endomorphism of V. We are already familiar with the characteristic polynomial of T, $c_T(x) = \det(xI - [T]_B)$, where $[T]_B$ is the square matrix representation of T relative to a basis B. We shall introduce two other polynomials associated with T: for any nonzero $v \in V$ there is a T-annihilator polynomial for v, and there is also the minimum polynomial of T.

The reader should also note that throughout the paper, we will often state that a polynomial $p(x)$ is satisfied by $T(v)$, for some appropriate endomorphism T and vector v. This, however, is a misnomer, as $T(v)$ is a vector, and we have not provided a definition for vector multiplication other than for a vector space of endomorphisms. Rather, what is meant by this statement is a two-step process: first, the polynomial $p(x)$ is evaluated at T, producing an endomorphism $p(T)$, which we then evaluate at v, producing $p(T)(v)$. That is, $T(v)$ satisfying $p(x)$ is to imply $p(T)(v) = 0$, and not $p(T(v)) = 0$.

Theorem 2.1. (Annihilator Polynomial) *Let V be an n-dimensional vector space, T an endomorphism of V, and $v \in V$ a non-zero vector. Then there is a unique monic polynomial of minimum degree, $m_{T,v}(x)$ such that $m_{T,v}(T)(v) = 0$. This polynomial has degree at most n.*

Proof. Consider the set $A = \{v, T(v), ..., T^n(v)\}$. Then A is a set of $n+1$ vectors in an n-dimensional vector space, and must be linearly dependent. Therefore, there exist scalars $a_0, a_1, ..., a_n$, not all zero, such that

$$a_n T^n(v) + a_{n-1} T^{n-1}(v) + ... + a_0 v = 0.$$

We can express this linear combination of vectors as a polynomial $p(x)$ evaluated at $T(v)$, where $p(x) = a_n x^n + ... + a_1 x + a_0$ is such a polynomial. Let a_s denote the coefficient on the leading term; that is, the coefficient on the highest power term with a nonzero coefficient. If we define $m_{T,v}(x) = b_n x^n + b_{n-1} x^{n-1} + ... + b_0$, where $b_i = a_i / a_s$, then $m_{T,v}$ is a monic polynomial. Because

$m_{T,v}(T)(v) = \frac{1}{a_s}p(T)(v) = 0$, then $m_{T,v}(x)$ is satisfied by $T(v)$, and therefore there exists a monic polynomial $m_{T,v}(x)$ with degree at most n such that $m_{T,v}(T)(v) = 0$. Because n is finite, then there must exist a monic polynomial of minimal degree that $T(v)$ satisfies. To show uniqueness, let m and m' be two monic polynomials of minimum degree that $T(v)$ satisfies. Then clearly, $T(v)$ must satisfy $(m - m')$. However, because m and m' are both monic, $\deg(m - m') < \deg m$, which contradicts the minimality of $\deg m$. Thus, $m_{T,v}(x)$ is unique. □

Definition The polynomial from the previous theorem, $m_{T,v}(x)$, is the T-annihilator polynomial of v.

After long delay, we shall now define the minimum polynomial for an endomorphism. The proof relies on $\mathrm{End}(V)$ forming an algebra as well as properties of polynomials.

Theorem 2.2. (Minimum Polynomial) *Let V be a finite-dimensional vector space with dimension n, and T an endomorphism of V. Then there exists a unique monic polynomial of minimum degree, $m_T(x)$, such that $m_T(T)(v) = 0$ for every $v \in V$. This polynomial has a degree less than or equal to n^2.*

Proof. Choose a basis $B = \{b_1, b_2, ..., b_n\}$ for V, and let $p_i(x) = m_{T,b_i}(x)$, the T-annihilator polynomial for each vector b_i. Define $q(x)$ to be the least common multiple of the polynomials $p_1(x), p_2(x), ..., p_n(x)$. Because each of the $p_i(x)$'s is a monic polynomial, then $q(x)$ must too be monic. Recall that if $f(x)$ and $g(x)$ are polynomials, then $\deg fg = \deg f + \deg g$. Now, because $q(x)$ is the least common multiple of n polynomials each with degree at most n, then

$$\deg q(x) \leq \sum_{i=1}^{n} \deg p_i(x) \leq n^2.$$

For each $v \in V$, v is a linear combination of the basis elements; i.e.,

$$v = \sum_{i=1}^{n} \alpha_i b_i.$$

Because $q(T)$ is again an endomorphism, we can see that

$$\begin{aligned} q(T)(v) &= q(T)\left(\sum_{i=1}^{n} \alpha_i b_i\right) \\ &= \sum_{i=1}^{n} \alpha_i q(T)(b_i). \end{aligned}$$

Because $q(x)$ is a multiple of each $p_i(x)$, then $q(b_i) = 0$ for each b_i. Therefore,

$$q(T)(v) = \sum_{i=1}^{n} \alpha_i q(T)(b_i) = 0,$$

and we have constructed a monic polynomial satisfied by $T(v)$ for every $v \in V$. Thus, because the degree of $q(x)$ is bounded, there exists a monic polynomial of minimum degree satisfied by $T(v)$ for every $v \in V$. The proof of uniqueness is similar to that in the previous theorem, and is omitted. The polynomial $m_T(x) = q(x)$ satisfies our hypotheses, concluding our proof. □

Definition The polynomial from the previous theorem, $m_T(x)$, is the minimum polynomial of the endomorphism T.

We have shown that the minimum polynomial of an endomorphism T, $m_T(x)$, has degree at most n^2. We shall later show that, in fact, $\deg m_T(x) \leq n$. However, before we show this, we must introduce one more lemma.

Lemma 2.3. *Let V be a vector space, and T an endomorphism of V. For $v_1, v_2, ..., v_k \in V$, let $p_i(x) = m_{T,v_i}(x)$. Suppose the polynomials $p_1(x), p_2(x), ..., p_k(x)$ are pairwise relatively prime. Then, if $v = v_1 + v_2 + ... + v_k$, the T-annihilator polynomial of v is given by*

$$m_{T,v}(x) = p_1(x)p_2(x) \cdots p_k(x).$$

Proof. See Weintraub's *A Guide to Advance Linear Algebra*, pg. 112. □

We are now ready to show that the minimum polynomial of an endomorphism of an n-dimensional vector space has degree at most n. We shall show that there is a vector $v \in V$ such that $m_{T,v}(x) = m_T(x)$. Then, because $m_T(x)$ is the T-annihilator of some vector v, and every T-annihilator polynomial has degree at most n, then the minimum polynomial of T must also have degree at most n.

Theorem 2.4. *Let V be an n-dimensional vector space, and T an endomorphism of V. Then there is some $v \in V$ such that $m_T(x) = m_{T,v}(x)$.*

Proof. Choose $B = \{v_1, v_2, ..., v_n\}$ to be a basis of V. As we have shown in the theorem proving the existence of $m_T(x)$, $m_T(x)$ is the least common multiple of m_{T,v_i} for $v_i \in B$. We can factor

$$m_T(x) = p_1(x)^{f_1} \cdots p_k(x)^{f_k}$$

into powers of distinct irreducible polynomials. As each $p_i(x)$ are pairwise relatively prime, so are each $p_i(x)^{f_i}$. For each i, $p_i(x)^{f_i}$ must appear as a factor of $m_{T,v_j}(x)$ for some $v_j \in B$, as $m_T(x)$ is the least common multiple of a set of polynomials. We can write $m_{T,v_j}(x) = p_i(x)^{f_i}q(x)$. Then the vector $u_i = q(T)(v_j)$ has annihilator $p_i(x)^{f_i}$. By our previous theorem, the vector $v = u_1 + u_2 + ... + u_k$ has a T-annihilator polynomial $m_{T,v}(x) = p_1(x)^{f_1} \cdots p_k(x)^{f_k}$. Therefore, $m_{T,v}(x) = m_T(x)$. □

Characteristic Polynomials

As linear transformations can be represented via matrices, we now briefly turn our attention to matrices. While characteristic polynomials can be defined for linear transformations, they are generally computed from matrix representations. First, we will show that similar matrices have identical characteristic polynomials, and hence every matrix representation of a linear transformation has the same characteristic polynomial. Then, while we can easily compute the characteristic polynomial of a matrix, it is less clear how to compute a matrix representation of a linear transformation given a characteristic polynomial. However, we will show that we, in fact, can do so.

Theorem 2.5. *Let A and B be similar matrices. Then the characteristic polynomials of A and B, $c_A(x)$ and $c_B(x)$, are equal.*

Proof. First, let $B = P^{-1}AP$. Then we can see that for any $x \in \mathbb{F}$,

$$\begin{aligned} xI - B &= xP^{-1}P - P^{-1}AP \\ &= P^{-1}xP - P^{-1}AP \\ &= P^{-1}(x - A)P. \end{aligned}$$

Therefore,

$$\begin{aligned} \det(xI - B) &= \det(P^{-1}(xI - A)P) \\ &= \det(P^{-1})\det(xI - A)\det(P) \\ &= \det(P)^{-1}\det(xI - A)\det(P) \\ &= \det(xI - A). \end{aligned}$$

□

Thus, we can see that if $[T]_B$ and $[T]_A$ are two matrix representations of the endomorphism T with respect to bases B and A, respectively, then $c_A(x) = c_B(x)$. Thus, $c_T(x)$ is a well-defined polynomial that may be easily computed, unlike $m_T(x)$. Now, given a monic polynomial $f(x)$, we will introduce a way to create a matrix whose characteristic polynomial is $f(x)$.

Definition Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots a_1 x + a_0$ be a monic polynomial in $\mathbb{F}[x]$ of degree $n \geq 1$. Then the companion matrix, $C(f(x))$, of $f(x)$ is the $n \times n$ matrix

$$C(f(x)) = \begin{bmatrix} -a_{n-1} & 1 & 0 & \cdots & 0 \\ -a_{n-2} & 0 & 1 & \cdots & 0 \\ & & \vdots & \ddots & \\ -a_1 & 0 & 0 & \cdots & 1 \\ -a_0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

where the 1's are located on the super-diagonal.

While the definition of this matrix may have seemed unmotivated, we shall see that it has convenient properties. First, we shall prove a theorem that transitions into later theory. The techniques used in later proofs will be reminiscent of techniques used here, and so it is for the convenience of the reader that the following theorem is proved. The proof is directly from Weintraub's *A Guide to Advanced Linear Algebra*.

Theorem 2.6. *Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be a monic polynomial and let $A = C(f(x))$ be its companion matrix. Let $V = \mathbb{F}^n$ and let $T = T_A$ be an endomorphism of V defined by $T(v) = Av$. Let $v = e_n$ be the nth standard basis vector. Then the subspace W of V defined by $W = \{g(T)(v) : g(x) \in \mathbb{F}[x]\}$ is V. Furthermore, $m_T(x) = m_{T,v}(x) = f(x)$.*

Proof. We see that $T(e_n) = e_{n-1}$, $T^2(e_n) = e_{n-2}$, and in general $T^k(e_n) = e_{n-k}$ for $k \leq n-1$. Thus, the subspace W of V contains the subspace spanned by $\{T^{n-1}(v), ..., T(v), v\} = \{e_1, ..., e_{n-1}, e_n\}$, the standard basis vectors of V, which is all of V. We also see that this set is linearly independent, and hence there is no nonzero polynomial $p(x)$ with degree less than or equal to $n-1$ with $p(T)(v)=0$. From

$$\begin{aligned} T^n(v) &= T(e_1) = -a_{n-1}e_1 - a_{n-2}e_2 \cdots - a_o e_n \\ &= -a_{n-1}T^{n-1}(v) - a_{n-2}T^{n-2}(v) - \cdots - a_1 T(v) - a_0 v \end{aligned}$$

we see that

$$0 = a_n T^n(v) + \cdots + a_1 T(v) + a_0 v,$$

which is to say $f(T)(v)=0$. Therefore, $m_{T,v}(x)=f(x)$.

On the one hand, $m_{T,v}(x)$ divides $m_T(x)$, an obvious consequence of both polynomials being satisfied by $T(v)$. On the other hand, since every $w \in V$ is $w = g(T)(v)$ for some polynomial $g(x)$,

$$\begin{aligned} m_{T,v}(T)(w) &= m_{T,v}(T)g(T)(v) \\ &= g(T)m_{T,v}(T)(v) \\ &= g(T)(0) \\ &= 0, \end{aligned}$$

for every $w \in V$. Thus, $m_T(x)$ divides $m_{T,v}(x)$. Therefore, we can see that

$$m_T(x) = m_{T,v}(x) = f(x).$$

□

We will present one more lemma involving the companion matrix but will not include the full proof.

Lemma 2.7. *Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ be a monic polynomial of degree $n \geq 1$ and let $A = C(f(x))$ be its companion matrix. Then $c_A(x) = \det(xI - A) = f(x)$.*

Proof. The proof follows from an induction argument on n. The reader can find the proof in Weintraub's *A Guide to Advanced Linear Algebra*. □

Starting with a monic polynomial $p(x)$, we can now create a linear transformation whose minimum polynomial is $p(x)$. This reverse-engineering process can be useful for obtaining linear transformations with specific properties, specific eigenvalues, and the like. As we will later see, roots of the minimum polynomial and characteristic polynomial are deeply related, and hence the obvious application of the companion matrix is to construct linear transformations with specific eigenvalues. We now turn our attention to invariant subspaces and the relationship between the characteristic polynomial and minimum polynomial.

Minimum Polynomial vs. Characteristic Polynomial

In this section, we shall investigate the relationship between the minimum and characteristic polynomial of a linear transformation. We will also investigate invariant subspaces and ways to express a vector space as the direct sum of invariant subspaces found using the minimum polynomial. We begin with simple material.

Lemma 2.8. *Let V be a vector space, T be an endomorphism of V, and $p(x)$ a polynomial in $\mathbb{F}[x]$. If W is a T-invariant subspace of V, then W is also an invariant subspace of V under $p(T)$.*

Proof. A careful proof of this lemma is a trivial exercise, and so it suffices to say that W is clearly invariant under T^n for any n. Hence, it must too be invariant under linear combinations of powers of T, and thus is invariant under any $p(T)$. □

The following definition should remind the reader of theory developed from the companion matrix. We shall generalize the span of a single vector over all polynomials to the span of a set of vectors over all possible polynomials.

Definition Let V be a vector space over the field $\mathbb{F}$ and T be an endomorphism of V. If $B = \{v_1, v_2, ..., v_k\}$ is a set of vectors in V, then the T-span of B is the subspace

$$W = \left\{ \sum_{i=1}^{k} p_i(T)(v_i) : p_i(x) \in \mathbb{F}[x] \right\}.$$

If $V = W$, then we shall say that V is T-generated by B. We need to show that the T-span of a set of vectors does form a subspace of V.

Theorem 2.9. *Let V be a vector space and T be an endomorphism of V. For a set of vectors B in V, the T-span of B, W, is the smallest T-invariant subspace of V containing B.*

Before we prove this theorem, we should first clarify what is meant by the "smallest" subspace. This simply means that W is contained in any T-invariant subspace that contains B.

Proof. Let W be the T-span of B. We can easily show that W is a subspace of V. As $0 \in \mathbb{F}[x]$, then $0 \in W$. Further, for $w_1, w_2 \in W$, and $\alpha \in \mathbb{F}$, where

$$w_1 = \sum_{i=1}^{k} p_{1,i}(T)(v_i) \quad \text{and}$$
$$w_2 = \sum_{i=1}^{k} p_{2,i}(T)(v_i),$$

then $w_1 + w_2 = \sum_{i=1}^{k}(p_{1,i} + p_{2,i})(T)(v_i) \in W$. Then,

$$\begin{aligned}\alpha w_1 &= \alpha \sum_{i=1}^{k} p_{1,i}(T)(v_i) \\ &= \sum_{i=1}^{k} \alpha p_{1,i}(T)(v_i) \in W.\end{aligned}$$

To show W is T-invariant, consider

$$\begin{aligned}T(w_1) &= T\left(\sum_{i=1}^{k} p_{1,i}(T)(v_i)\right) \\ &= \sum_{i=1}^{k} T(p_{1,i}(T)(v_i)) \\ &= \sum_{i=1}^{k} q_{1,i}(T)(v_i) \in W,\end{aligned}$$

where $q_{1,i}(x) = xp_{1,i}(x)$ for all i. We must now show that W is contained in every other T-invariant subspace containing B.

Let $\hat{W}$ be a T-invariant subspace containing B. It is convenient to recall that $\hat{W}$ is also invariant under $p(T)$ for any polynomial $p(x)$. Thus, for vectors $w_1, w_2, ..., w_k \in \hat{W}$,

$$\sum_{i=1}^{k} p_i(T)(w_i)$$

must also be in $\hat{W}$. Therefore, $W \subseteq \hat{W}$, and W must be the smallest T-invariant subspace of V containing B. □

Our following lemma will link together T-invariant subspaces and annihilator polynomials.

Lemma 2.10. *Let V be a finite-dimensional vector space, and T an endomorphism of V. For $w \in V$, let W be the subspace of V T-generated by w. Then the dimension of W is equal to the degree of the T-annihilator $m_{T,w}(x)$ of w.*

Proof. We shall prove this lemma by showing $m_{T,w}(x)$ has degree k if and only if $\{w, T(w), ..., T^{k-1}(w)\}$ is a basis of W.

Because W is a finite-dimensional vector space, let $\dim(W) = k$. Now suppose $\{w, T(w), ..., T^{k-1}(w)\}$ is basis for W. Then there is no non-trivial linear combination of the basis vectors that is equal to 0, and hence no polynomial of degree $k-1$ or less that is satisfied by $T(w)$. However, the set of vectors $\{w, T(w), ..., T^{k-1}(w), T^k(w)\}$ is a set of $k+1$ vectors in a k-dimension vector space, and hence there is a polynomial with degree k that is satisfied by $T(w)$. Therefore, the T-annihilator of w, $m_{T,w}(x)$ has degree k.

Now suppose $\deg m_{T,w}(x) = k$. Then there is no subset of $\{w, T(w), ..., T^{k-1}(w)\}$ that forms a linearly dependent set. Now, $T^k(w)$ can be expressed as a non-trivial linear combination of the vectors $\{w, T(w), ..., T^{k-1}(w)\}$, and inductively so can each T^{k+i} for every $i \in \mathbb{N}$. Thus, $\{w, T(w), ..., T^{k-1}(w)\}$ is a basis for W. This concludes our proof. □

The following theorem will be helpful in expressing a vector space as a direct sum of invariant subspaces found via minimum polynomials.

Theorem 2.11. *Let V be a vector space, T and endomorphism of V, and $p(x) \in \mathbb{F}[x]$. Then*

$$\ker(p(T)) = \{v \in V : p(T)(v) = 0\}$$

is a T-invariant subspace of v.

Proof. It is a trivial matter to show that $\ker(p(T))$ is a subspace of V. Thus, we focus our attention on showing $\ker(p(T))$ is T-invariant. To show $\ker(p(T))$ is T-invariant, we must show $T(v) \in \ker(p(T))$ for all $v \in \ker(p(T))$. Clearly,

$$\begin{aligned} p(T)(T(v)) &= T(p(T)(v)) \\ &= T(0) \\ &= 0. \end{aligned}$$

Therefore, $T(v) \in \ker(p(T))$, and $\ker(p(T))$ is a T-invariant subspace. □

The following theorem, proved earlier in many texts, discusses the direct relationship between the minimum polynomial and characteristic polynomial of an endomorphism T. Our proof also uses the Cayley-Hamilton theorem. Although some texts state the Cayley-Hamilton theorem as a corollary of the following proof, it was originally proven using completely different methods. Thus, it is not unreasonable to assume the Cayley-Hamilton theorem. However, a proof of the following theorem that uses the Cayley-Hamilton theorem can be found in Weintraub [1].

Theorem 2.12. *Let V be a finite-dimensional vector space and T an endomorphism of V. Let $m_T(x)$ and $c_T(x)$ be the minimum polynomial and characteristic polynomial of T, respectively. Then*

1. *$m_T(x)$ divides $c_T(x)$, and*
2. *every irreducible factor of $c_T(x)$ is an irreducible factor of $m_T(x)$.*

Proof. 1.) From the Cayley-Hamilton theorem, we know that $c_T(T) = 0$. Thus, by the division theorem, there exist unique polynomials $q(x)$ and $r(x)$ such that

$$c_T(x) = q(x)m_T(x) + r(x),$$

where $\deg r(x) < \deg m_T(x)$. However, we can clearly see that $c_T(T) = 0$ implies $r(T) = 0$. Because $r(x)$ has degree strictly less than $m_T(x)$, this violates the minimality of the degree of $m_T(x)$ unless $r(x) = 0$. Thus, $m_T(x)$ divides $c_T(x)$.

The proof of 2.) can be found in Weintraub's *A Guide to Advanced Linear Algebra*.

□

We shall concretely restate this theorem as follows.

Theorem 2.13. *Let $m_T(x) = p_1(x)^{e_1} p_2(x)^{e_2} \cdots p_k(x)^{e_k}$ factor into distinct irreducible polynomials $p_1(x), p_2(x), ..., p_k(x)$. Then $c_T(x) = p_1(x)^{f_1} p_2(x)^{f_2} \cdots p_k(x)^{f_k}$, where $f_i \geq e_i$ for every i.*

The following special case tells us much about the structure of both a vector space and an endomorphism.

Corollary 2.14. *Let V be an n-dimensional vector space and T an endomorphism of V. Then V is T-generated by a single element if and only if $m_T(x)$ is a polynomial of degree n, or, equivalently, if $m_T(x) = c_T(x)$.*

Proof. We shall begin by showing if the degree of $m_T(x)$ is n, then V is T-generated by a single element. For $v \in v$, we know that the T-generated subspace of V has dimension equal to $\deg m_{T,v}(x)$. Because there exists $\hat{v} \in V$ such that $m_{T,\hat{v}}(x) = m_T(x)$, then the dimension of the subspace T-generated by $\hat{v}$ is n, and hence is V.

Now, we suppose V is T-generated by a single element, v. Consider the set
$B = \{v, T(v), ..., T^{n-1}(v)\}$. If any of the vectors in B can be expressed as a linear combination of the other vectors in B, then there exists an $i < n$ such that the set $\{v, T(v), ..., T^{n-i}(v)\}$ is linearly independent and spans V. However, this is impossible unless $i = 1$. Therefore, $m_T(x)$ has degree n.

Because $m_T(x)$ and $c_T(x)$ are both monic polynomials of degree n, and $m_T(x)$ divides $c_T(x)$, then $m_T(x) = c_T(x)$. □

Our final theorem ties together the material accumulated thus far. The theorem relates invariant subspaces found using the minimum polynomial and the structure of a linear transformation.

Theorem 2.15. *Let V be a vector space and let T be an endomorphism of V. Let T have the minimum polynomial $m_T(x)$ that factors into the product of pairwise relatively prime polynomials, $m_T(x) = p_1(x)p_2(x)\cdots p_k(x)$. For each i, let $W_i = \ker(p_i(T))$. Then each W_i is T-invariant, and $V = W_1 \oplus \cdots \oplus W_k$.*

Proof. See Weintraub, page 125. □

We have seen various ways of using the minimum polynomial of a linear transformation. Not only are they helpful in analyzing the structure of the linear transformation itself, but they also display underlying structure of a linear transformation. Further, there are many topics that rely on the minimum polynomial, such as rational canonical form, that were not addressed in this paper. However, when the reader studies such topics, the material presented here will be an extremely useful aid.

Bibliography

[1] Weintraub, Steven H. *A Guide to Advanced Linear Algebra.* United States of America: The Mathematical Association of America, 2011.

[2] Curtis, Morton L. *Abstract Linear Algebra.* New York: Springer-Verlag, 1990.