A Linear Algebra Proof of Fundamental Theorem of Algebra

Arindama Singh
Department of Mathematics
IIT Madras

Abstract: We try to understand and rewrite Derksen's proof of the Fundamental theorem of Algebra basing on the techniques of Linear Algebra.

1 Introduction

The fundamental theorem of Algebra states that every polynomial with complex coefficients has a complex zero. In [1], Derksen proves this result via results involving a finite number of commuting endomorphisms on a vector space. It is easy to observe from the paper that only a pair of commuting endomorphisms is enough to prove the result.

We will break the proof into some lemmas and introduce some terminology which will help us in understanding the crucial aspects of Derksen's proof.

2 Preliminaries

We will write $\mathbb{R}$ for the field of real numbers, $\mathbb{C}$ for the field of complex numbers, and $\mathbb{F}$ for a field, which is either $\mathbb{R}$ or $\mathbb{C}$.

Let V be a vector space over $\mathbb{F}$ of dimension $n \in \mathbb{N}$. Elements of V are called vectors, and elements of $\mathbb{F}$ are called scalars. When $\mathbb{F} = \mathbb{R}$, V is called a *real vector space*; and when $\mathbb{F} = \mathbb{C}$, we say that V is a *complex vector space*. If n is odd, we say that V is an *odd dimensional vector space*.

Let $T : V \to V$ be a linear operator. If a scalar $\lambda \in \mathbb{F}$ and a nonzero vector $v \in V$ exist such that $T(v) = \lambda v$, then such a scalar λ is called an *eigenvalue* of T and such a vector v is called an *eigenvector* of T. When we say that T *has an eigenvector* we mean that the equation $T(v) = \lambda v$ is satisfied for some scalar λ. In this case, we say that λ is an *eigenvalue* of T corresponding to the eigenvector v.

The *null space* of T consists of all vectors $x \in V$ such that $T(x) = 0$; we denote the null space of T by $N(T)$. The *range space* of T, denoted by $R(T)$, consists of all vectors $T(x)$, where $x \in V$. Recall that $N(T)$ and $R(T)$ are subspaces of V and the rank-nullity theorem says that $\dim(N(T)) + \dim(R(T)) = \dim(V)$.

If $Tv = \lambda v$ is satisfied for some scalar λ and some nonzero vector v, then $(\lambda I - T)v = 0$. It implies that $\dim(N(\lambda I - T)) \geq 1$. Thus, $\dim(R(\lambda I - T)) < n$. Fix any ordered basis for V. With respect to this basis, let M be the matrix representation of T. Then, $\lambda I - M$ has rank less than n. As $\lambda I - M$ is a matrix of order n, it is singular. Hence, its determinant is 0. We see that $t = \lambda$ satisfies $\det(tI - M) = 0$.

However, $\det(tI-M)$ is a monic polynomial in t having degree n. This polynomial $\det(tI-M)$ does not depend upon the particular ordered basis chosen for V since a change of basis will yield the matrix representation of $tI - T$ as $P^{-1}(tI - M)P$ for some invertible $n \times n$ matrix P. The polynomial $\det(tI - M)$ is called the *characteristic polynomial* of T.

So, an eigenvalue of T is a zero of the characteristic polynomial of T. Conversely, if a zero of the characteristic polynomial of T happens to be in the underlying field $\mathbb{F}$, then it is an eigenvalue of T. The same are true when T is a square matrix with entries from $\mathbb{F}$ since a square matrix of order n is a linear operator on $\mathbb{F}^{n\times 1}$. Further, with respect to an ordered basis of V if A is the matrix representation of T, then a scalar is an eigenvalue of T if and only if it is an eigenvalue of A.

3 The lemmas

Our first result is an application of the intermediate value theorem.

Lemma 1 *Each linear operator on an odd dimensional real vector space has an eigenvector.*

Proof. Let V be a real vector space with $\dim(V) = n$, where n is odd. Let T be a linear operator on V. The characteristic polynomial $p(t)$ of T is a monic polynomial of degree n with real coefficients. Since n is odd, $\lim_{t\to-\infty} p(t) = -\infty$ and $\lim_{t\to\infty} p(t) = \infty$. As $p(t)$ is a continuous function, by the Intermediate value theorem $p(t)$ has a real zero, say, λ. Then λ is a real zero of $p(t)$ so that it is an eigenvalue of T. So, there exists a nonzero vector $v \in V$ such that $T(v) = \lambda v$. Such a vector v is an eigenvector of T. □

Let S and T be two linear operators on a vector space V over $\mathbb{F}$. We say that they are *commuting operators* if $S(T(x)) = T(S(x))$ for each $x \in V$. A nonzero vector $v \in V$ that satisfies the equations $S(v) = \lambda v$ and $T(v) = \mu v$ for some scalars λ and μ is called a *common eigenvector* of S and T. Once such a vector v exists, we say that S and T *have a common eigenvector*.

Lemma 2 *Let $k \in \mathbb{N}$. If each linear operator on a finite dimensional nonzero vector space over $\mathbb{F}$ with its dimension not divisible by k has an eigenvector, then any two commuting linear operators on a finite dimensional nonzero vector space over $\mathbb{F}$ with its dimension not divisible by k have a common eigenvector.*

Proof. Assume that each linear operator on a finite dimensional nonzero vector space over $\mathbb{F}$ with its dimension not divisible by k has an eigenvector. Let V be a finite dimensional nonzero vector space over $\mathbb{F}$ such that k does not divide $\dim(V)$. Let S and T be two commuting operators on V. We show by induction on $\dim(V)$ that S and T have a common eigenvector.

Suppose $\dim(V) = 1$. Let $\{v\}$ be a basis for V. Now, $S(v) \in V$ implies that there exists $\alpha \in \mathbb{F}$ such that $S(v) = \alpha v$. Similarly, $T(v) = \beta v$ for some $\beta \in \mathbb{F}$. Then, v is a common eigenvector of S and T.

Assume the induction hypothesis that if U is any vector space over $\mathbb{F}$ of dimension less than n, where k does not divide $\dim(U)$, then any two commuting linear operators on U have a common eigenvector. Let V be a vector space over $\mathbb{F}$ of dimension n, where k does not divide n. Let v be an eigenvector of S. There exists $\alpha \in \mathbb{F}$ such that $S(v) = \alpha v$. Let $\mathcal{N} = N(S - \alpha I)$, the null space of $S - \alpha I$. Since $v \in \mathcal{N}$, $\dim(\mathcal{N}) \geq 1$.

Case 1: Suppose $\mathcal{N} = V$. Let w be an eigenvector of T. Then, $w \in N(S - \alpha I)$. Now, w is a common eigenvector of S and T.

Case 2: Suppose $\mathcal{N}$ is a proper subspace of V. Then $1 \leq \dim(\mathcal{N}) < n$. We first show that both S and T are functions on $\mathcal{N}$. For this, let $x \in \mathcal{N}$. Then, $(S - \alpha I)(x) = 0$ so that

$$\begin{aligned}(S - \alpha I)S(x) &= S^2(x) - \alpha S(x) = S(S - \alpha I)(x) = S(0) = 0,\\ (S - \alpha I)T(x) &= S(T(x)) - \alpha T(x) = T(S(x)) - T(\alpha x) = T(S - \alpha I)(x) = T(0) = 0.\end{aligned}$$

That is, $S(x) \in \mathcal{N}$ and $T(x) \in \mathcal{N}$. Hence, both S and T are functions of $\mathcal{N}$. Since S and T are commuting operators on V, it follows that their restrictions to $\mathcal{N}$ are commuting operators.

If k does not divide $\dim(\mathcal{N})$, then by the induction hypothesis, the restriction operators of S and T to $\mathcal{N}$ have a common eigenvector in $\mathcal{N}$. This common eigenvector is also a common eigenvector of S and T as linear operators on V.

So, suppose that k divides $\dim(\mathcal{N})$. Consider $\mathcal{R} = R(S - \alpha I)$, the range space of the linear operator $S - \alpha I$. We first show that S and T are functions on $\mathcal{R}$. For this, let $y \in \mathcal{R}$. Then $y = (S - \alpha I)(x)$ for some $x \in V$. Now,

$$\begin{aligned}S(y) &= S(S - \alpha I)(x) = S^2(x) - \alpha S(x) = (S - \alpha I)S(x) \in \mathcal{R},\\ T(y) &= T(S - \alpha I)(x) = T(S(x)) - \alpha T(x) = S(T(x)) - \alpha T(x) = (S - \alpha I)T(x) \in \mathcal{R}.\end{aligned}$$

That is, both S and T are functions on $\mathcal{R}$. Since S and T are commuting operators on V, their restrictions to $\mathcal{R}$ are commuting operators. As $\dim(\mathcal{N}) + \dim(\mathcal{R}) = n$, k divides $\dim(\mathcal{N})$ and k does not divide n, it follows that k does not divide $\dim(\mathcal{R})$. Also, $\dim(\mathcal{R}) < n$. By the induction hypothesis, the restriction operators of S and T to $\mathcal{R}$ have a common eigenvector w. Now, w is a common eigenvector of S and T as linear operators on V. □

We use Lemmas 1-2 to extend the result of Lemma 1 to linear operators on odd dimensional complex vector spaces.

Lemma 3 *Each linear operator on an odd dimensional vector space over $\mathbb{F}$ has an eigenvector.*

Proof. Due to Lemma 1, we need to consider the case $\mathbb{F} = \mathbb{C}$. So, let V be a complex vector space with $\dim(V) = n$, where n is odd. Let T be a linear operator on V. Fix an ordered basis for V. With respect to this ordered basis, let A be the matrix representation of T. Then, A is a square matrix of order n with complex entries. We need to show that there exist $\lambda \in \mathbb{C}$ and a nonzero vector $v \in \mathbb{C}^{n\times 1}$ such that $Av = \lambda v$.

Consider $H_n = \{X \in \mathbb{C}^{n\times n} : X^* = X\}$, the set of all hermitian matrices of order n. Then, H_n is a real vector space with $\dim(H_n) = n^2$, which is odd. Write

$$L_1(X) = \frac{AX + XA^*}{2}, \quad L_2(X) = \frac{AX - XA^*}{2i} \quad \text{for } X \in H_n.$$

If $X \in H_n$, then $X^* = X$, so that

$$[L_1(X)]^* = \frac{X^*A^* + AX^*}{2} = L_1(X), \quad [L_2(X)]^* = \frac{X^*A^* - AX^*}{2i} = L_2(X).$$

That is, L_1 and L_2 as given by the above equations are functions on H_n. Further, if $b \in \mathbb{R}$, $X, Y \in H_n$, then

$$\begin{aligned}
L_1(bX+Y) &= \frac{A(bX+Y) + (bX+Y)A^*}{2} \\
&= b\,\frac{AX + XA^*}{2} + \frac{AY + YA^*}{2} = bL_1(X) + L_1(Y), \\
L_2(bX+Y) &= \frac{A(bX+Y) - (bX+Y)A^*}{2i} \\
&= b\,\frac{AX - XA^*}{2i} + \frac{AY - YA^*}{2i} = bL_2(X) + L_2(Y), \\
L_1(L_2(X)) &= L_1\Big(\frac{AX - XA^*}{2i}\Big) = \frac{1}{2}\Big[A\Big(\frac{AX - XA^*}{2i}\Big) + \Big(\frac{AX - XA^*}{2i}\Big)A^*\Big] \\
&= \frac{1}{4i}\Big(A^2X - AXA^* + AXA^* - X(A^*)^2\Big) = \frac{A^2X - X(A^*)^2}{4i}. \\
L_2(L_1(X)) &= L_2\Big(\frac{AX + XA^*}{2}\Big) = \frac{1}{2i}\Big[A\Big(\frac{AX + XA^*}{2}\Big) + \Big(\frac{AX + XA^*}{2}\Big)A^*\Big] \\
&= \frac{1}{4i}\Big(A^2X + AXA^* - AXA^* - X(A^*)^2\Big) = \frac{A^2X - X(A^*)^2}{4i}.
\end{aligned}$$

Hence, L_1 and L_2 are commuting operators on the odd dimensional real vector space H_n.

By Lemma 1, each linear operator on a finite dimensional nonzero real vector space with dimension not divisible by 2 has an eigenvector. By Lemma 2, any two commuting operators on a finite dimensional nonzero real vector space with dimension not divisible by 2 have a common eigenvector. Thus, L_1 and L_2 have a common eigenvector, say $B \in H_n$. That is,

$$L_1(B) = \alpha B, \quad L_2(B) = \beta B \quad \text{for some } \alpha, \beta \in \mathbb{R}.$$

We find that

$$\begin{aligned}
L_1(B) + iL_2(B) &= \frac{AB + BA^*}{2} + \frac{AB - BA^*}{2} = AB, \\
L_1(B) + iL_2(B) &= \alpha B + i\beta B = (\alpha + i\beta)B.
\end{aligned}$$

Thus, $AB = (\alpha + i\beta)B$. As $B \neq 0$, let v be a nonzero column of B. It follows that $Av = (\alpha + i\beta)v$. This completes the proof. □

Let $m \in \mathbb{N} \cup \{0\}$. A natural number n is said to be *of evenness* m if 2^m divides n but 2^{m+1} does not divide n. Also, we agree to say that a linear operator is of evenness m if it is a linear operator on a finite dimensional nonzero complex vector space whose dimension is of evenness m. Notice that a linear operator of evenness 0 is simply a linear operator on an odd dimensional complex vector space.

Lemma 4 *Let $m \in \mathbb{N} \cup \{0\}$. If each linear operator of evenness less than or equal to m has an eigenvector, then each linear operator of evenness $m+1$ also has an eigenvector.*

Proof. Assume that each linear operator of evenness less than or equal to m has an eigenvector. Let $T : V \to V$ be a linear operator of evenness $m+1$. Write $\dim(V) = n$. Then, n is of evenness $m+1$. Fix an ordered basis for V. Let A be the matrix representation of T with respect to this ordered basis. Then, A is a matrix of order n with complex entries. We need to prove that there exist $\lambda \in \mathbb{C}$ and a nonzero vector $v \in \mathbb{C}^{n\times 1}$ such that $Av = \lambda v$.

Consider $S_n = \{X \in \mathbb{C}^{n\times n} : X^t = X\}$, the set of all symmetric matrices with complex entries of order n. Then, S_n is a complex vector space of dimension $n(n+1)/2$. Notice that $n = 2^{m+1}k$ for an odd integer k and $m \geq 0$. Thus, $n+1$ is odd so that $n(n+1)/2 = 2^m\ell$ for some odd integer ℓ. Thus, $\dim(S_n)$ is of evenness m so that any linear operator on S_n is of evenness m. Define

$$L_3(X) = AX + XA^t, \quad L_4(X) = AXA^t \quad \text{for } X \in S_n.$$

For $X \in S_n$, we have

$$[L_3(X)]^t = (AX + XA^t)^t = X^tA^t + AX^t = XA^t + AX = L_3(X),$$

$$[L_4(X)]^t = (AXA^t)^t = AX^tA^t = AXA^t = L_4(X).$$

Thus, $L_3, L_4 : S_n \to S_n$ as given above are functions. Again, for $X, Y \in S_n$ and $b \in \mathbb{C}$,

$$\begin{aligned}
L_3(bX+Y) &= A(bX+Y)A^t = bAXA^t + AYA^t = bL_3(X) + L_3(Y),\\
L_4(bX+Y) &= A(bX+Y)A^t = bAXA^t + AYA^t = bL_4(X) + L_4(Y),\\
L_3(L_4(X)) &= L_3(AXA^t) = A(AXA^t) + (AXA^t)A^t = A^2XA^t + AX(A^t)^2,\\
L_4(L_3(X)) &= L_4(AX + XA^t) = A(AX + XA^t)A^t = A^2XA^t + AX(A^t)^2.
\end{aligned}$$

That is, L_3 and L_4 are commuting operators. But they also have evenness m. Thus, L_3 and L_4 are commuting operators on a finite dimensional nonzero complex vector space of dimension not divisible by 2^{m+1}.

Our assumption that each linear operator of evenness less than or equal to m has an eigenvector implies that each linear operator on a finite dimensional nonzero complex vector space of dimension not divisible by 2^{m+1} has an eigenvector. By Lemma 2, L_3 and L_4 have a common eigenvector. That is, there exists $B \in S_n$, $B \neq 0$ and $\alpha,\ \beta \in \mathbb{C}$ such that

$$L_3(B) = AB + BA^t = \alpha B, \quad L_4(B) = ABA^t = \beta B.$$

Then, $\beta B = ABA^t = A(\alpha B - AB) = \alpha AB - A^2B$. It implies $A^2B - \alpha AB + \beta B = 0$.

Let $\gamma = \left(\alpha + \sqrt{\alpha^2 - 4\beta}\right)/2$ and $\delta = \left(\alpha - \sqrt{\alpha^2 - 4\beta}\right)/2$. Then,

$$\begin{aligned}
(A - \gamma I)(A - \delta I) &= A^2 - (\gamma + \delta)A + \gamma\delta I\\
&= A^2 - \alpha A + \tfrac{1}{4}\left[\alpha^2 - (\alpha^2 - 4\beta)\right] I = A^2 - \alpha A + \beta I.
\end{aligned}$$

Thus,

$$(A - \gamma I)(A - \delta I)B = A^2B - \alpha AB + \beta B = 0.$$

If $(A - \delta I)B = 0$, then any nonzero column u of B satisfies $(A - \delta I)u = 0$. So, we take $\lambda = \delta$ and $v = u$. If $(A - \delta I)B \neq 0$, then any nonzero column w of $(A - \delta I)B$ satisfies $(A - \gamma I)w = 0$. In this case, we take $\lambda = \gamma$ and $v = w$. □

4 The main result

Using Lemmas 1-4, we prove our main result as follows.

Theorem 1 *Every linear operator on a finite dimensional nonzero complex vector space has an eigenvector.*

Proof. We prove the theorem by induction on the evenness of a linear operator. Let V be a complex vector space. In the basis step, if a linear operator $T : V \to V$ is of evenness 0, then $2^{0+1} = 2$ does not divide $\dim(V)$. By Lemma 3, T has an eigenvector.

Assume the induction hypothesis that each linear operator of evenness less than or equal to $m \in \mathbb{N} \cup \{0\}$ has an eigenvector. Let $T : V \to V$ be a linear operator with evenness $m + 1$. By Lemma 4, T has an eigenvector. □

In particular, each square matrix with complex entries has an eigenvector. Using this result, we prove the fundamental theorem of algebra.

Theorem 2 *Every polynomial in a single variable with complex coefficients has a zero in the field of complex numbers.*

Proof. Let $p_n(t)$ be a polynomial of degree $n \in \mathbb{N}$. Then, the coefficient a of t^n in $p_n(t)$ is nonzero. Let $p(t) = p_n(t)/a$. Then, the coefficient of t^n in $p(t)$ is 1. Clearly a complex number is a zero of $p_n(t)$ if and only if it is a zero of $p(t)$. Let

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_{n-1} t^{n-1} + t_n.$$

Let A be the companion matrix of $p(t)$, that is the square matrix of order n whose ith column is e_{i+1} for $1 \leq i \leq n-1$, and the nth column is

$$\begin{bmatrix} -a_0 & -a_1 & \cdots & -a_{n-1} \end{bmatrix}^t.$$

Then, the characteristic polynomial of A is

$$\det(tI - A) = \begin{vmatrix} t & 0 & 0 & \cdots & 0 & a_0 \\ -1 & t & 0 & \cdots & 0 & a_1 \\ 0 & -1 & t & \cdots & 0 & a_2 \\ & & \ddots & \vdots & & \\ 0 & 0 & 0 & \ddots & t & a_{n-2} \\ 0 & 0 & 0 & \cdots & -1 & t + a_{n-1} \end{vmatrix}$$

If $n = 1$, then this determinant is $t + a_0$. When $n = 2$, the determinant is $t(t + a_1) + a_0 = t^2 + a_1 t + a_0$. For $n > 2$, expand the determinant on its first row to get the determinant equal to

$$t \begin{vmatrix} t & 0 & 0 & \cdots & 0 & a_1 \\ -1 & t & 0 & \cdots & 0 & a_2 \\ 0 & -1 & t & \cdots & 0 & a_3 \\ & & \ddots & \vdots & & \\ 0 & 0 & 0 & \ddots & t & a_{n-2} \\ 0 & 0 & 0 & \cdots & -1 & t + a_{n-1} \end{vmatrix} + (-1)^{n-1} a_0 \begin{vmatrix} -1 & t & 0 & \cdots & 0 \\ 0 & -1 & t & \cdots & 0 \\ & & \ddots & \vdots & \\ 0 & 0 & 0 & \ddots & t \\ 0 & 0 & 0 & \cdots & -1 \end{vmatrix}$$

By induction, the first one evaluates to $t(t^{n-1} + a_{n-1}t^{n-2} + \cdots + a_2 t + a_1)$. The second one is the product of the diagonal entries, as the underlying matrix is upper triangular; so it evaluates to $(-1)^{n-1}$. Therefore,

$$\det(tI - A) = t(t^{n-1} + a_{n-1}t^{n-2} + \cdots + a_2 t + a_1) + (-1)^{2(n-1)}a_0 = p(t).$$

By Theorem 1, the matrix A has an eigenvector. Then, the corresponding eigenvalue is a zero of the polynomial $p(t)$. □

Sometimes the following result is mentioned as the fundamental theorem of algebra.

Theorem 3 *A polynomial in a single variable of degree n with complex coefficients has exactly n complex zeros, counting multiplicities.*

Proof. Without loss of generality, let $p(t)$ be a monic polynomial of degree $n \geq 1$ with complex coefficients, say

$$p(t) = a_0 + a_1 t + \cdots + a_{n-1}t^{n-1} + t^n \quad \text{for } a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}.$$

If $n = 1$, then $p(t) = a_0 + t$ and it has exactly one zero, namely, $-a_0$. Assume the induction hypothesis that for each monic polynomial of degree $n-1$ with complex coefficients has exactly $n-1$ complex zeros. Due to Theorem 2, $p(t)$ has a complex zero, say α. Now,

$$\begin{aligned} p(t) &= p(t) - p(\alpha) \\ &= a_1(t-\alpha) + a_2(t^2 - \alpha^2) + \cdots + a_{n-1}(t^{n-1} - \alpha^{n-1}) + (t^n - \alpha^n) \\ &= (t-\alpha)\big[a_1 + a_2(t+\alpha) + \cdots + a_{n-1}(t^{n-2} + \cdots + \alpha^{n-2}) + (t^{n-1} + \alpha t^{n-2} + \cdots + \alpha^{n-1})\big]. \end{aligned}$$

The bracketed term here is a polynomial of degree $n-1$. By the induction hypothesis, it has exactly $n-1$ complex zeros. It follows that $p(t)$ has exactly n complex zeros. □

References

[1] H. Derksen, The fundamental theorem of algebra and linear algebra, *The American Mathematical Monthly*, 110 : 7 (620-623) 2003.