

MINIMAL POLYNOMIALS AND THE SPECTRAL THEOREM FOR HONORS LINEAR ALGEBRA

THOMAS HUNTER

ABSTRACT. These notes are to be a supplement to Bretscher, which contains no information on the minimal polynomial, and to my theory notes which are both more general and more telegraphic. We give the basic results and the development through a proof of the spectral theorem. The reader should compare this with Bretscher's treatment of the spectral theorem.

1. POLYNOMIALS

Definition 1. If \mathbb{k} is a field and $a_0, \dots, a_d \in \mathbb{k}$ we call the expression

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_dx^d$$

a polynomial with coefficients in \mathbb{k} . If $a_d \neq 0$ we say that d is the degree of $p(x)$ and write $\deg p(x) = d$. If $a_d = 1$ we say that $p(x)$ is monic. When we want to refer to the set of all polynomials with coefficients in \mathbb{k} we will write

$$\mathbb{k}[x] := \{p(x) = a_0 + \dots + a_dx^d \mid d \in \mathbb{Z}_{\geq 0}, a_0, \dots, a_d \in \mathbb{k}\}$$

Remark 2. Some people use “monic” to mean one-to-one. This use is entirely different; when the reader cannot tell from context which is meant, the author should be clear.

Remark 3. \mathbb{k} need not be a field for the definition to make sense. We will not digress here enough to say exactly what generality is possible other than to mention that $\mathbb{k} = \mathbb{Z}$ is a common occurrence. The reader who wants more rigorous details may seek a more rigorous source.

Remark 4. The “usual” arithmetic of polynomials is present in $\mathbb{k}[x]$. We may add and multiply polynomials. In particular, $\mathbb{k}[x]$ is a vector space over \mathbb{k} .

Example 5. $\mathbb{R}[x]$ is the collection of polynomials familiar from a first course in calculus. $\mathbb{Z}[x] \subseteq \mathbb{R}[x]$ are the polynomials with integral coefficients, and so on. Thus the sets

$$\mathbb{Z}[x] \subseteq \mathbb{Q}[x] \subseteq \mathbb{R}[x] \subseteq \mathbb{C}[x]$$

should all be fairly familiar to the reader.

Date: November 15, 2005.

Example 6. A less familiar and more interesting example is $(\mathbb{Z}/2)[x]$. In this domain we have, for example:

$$\begin{aligned}
 (1+x)^2 &= 1 + 2x + x^2 = 1 + x^2 \\
 (1+x)^3 &= 1 + x + x^2 + x^3 \\
 (1+x)^4 &= \left((1+x)^2\right)^2 = (1+x^2)^2 = 1 + (x^2)^2 = 1 + x^4 \\
 (1) \quad (1+x)^{2^n} &= 1 + x^{2^n}.
 \end{aligned}$$

Equation (1) is known as the “freshman binomial formula.”

Remark 7. Since addition and multiplication are defined in \mathbb{k} , if we are given $p(x) \in \mathbb{k}[x]$ and $\alpha \in \mathbb{k}$ we may consider the expression $p(\alpha)$ obtained by substituting α for x to be an element of \mathbb{k} . This is just the usual notion of polynomial as function.

Definition 8. If $p(x) = a_0 + \dots + a_d x^d \in \mathbb{k}[x]$ and $A \in M_{n \times n}(\mathbb{k})$ we define

$$p(A) := a_0 I + a_1 A + \dots + a_d A^d \in M_{n \times n}(\mathbb{k}).$$

Exercise 1. Suppose $p(x), q(x) \in \mathbb{k}[x]$ and $A \in M_{n \times n}(\mathbb{k})$. Let $u(x) = p(x)q(x)$.

$$p(A)q(A) = u(A).$$

That is to say that one achieves the same result by either multiply polynomials first and then plug in A or plug in A first and then multiply the resulting matrix.

2. MINIMAL POLYNOMIALS

Proposition 9. If $A \in M_{n \times n}(\mathbb{k})$ then there is a monic polynomial $p(x) \in \mathbb{k}[x]$ with $p(A) = 0_{n \times n}$.

Proof. $M_{n \times n}(\mathbb{k})$ is an n^2 dimensional vector space, so the elements $I, A, A^2, \dots, A^{n^2}$ must be linearly dependent. That is to say there is a nonzero polynomial $q(x) \in \mathbb{k}[x]$ with $q(A) = 0$. If $q(x) = b_0 + \dots + b_d x^d$ with $b_d \neq 0$, let $p(x) = \frac{1}{b_d} q(x)$. \square

Definition 10. Given $A \in M_{n \times n}(\mathbb{k})$, a minimal polynomial for A is a monic polynomial $p(x) \in \mathbb{k}[x]$ of smallest possible degree for which $p(A) = 0$.

Exercise 2. Given A any two minimal polynomials for A are equal.

Remark 11. Thus, it makes sense to refer to “the” minimal polynomial. Except for in the exercises, we will not need uniqueness of the minimal polynomial in these notes. (See my theory notes, for a more complete discussion.)

Proposition 12. Let $p(x)$ be a minimal polynomial for $A \in M_{n \times n}(\mathbb{k})$. A is invertible if and only if $p(0) \neq 0$. In fact we have the following:

- (1) If there is a B so that $AB = I$ or a C so that $CA = I$ then $p(0) \neq 0$.
 (2) If $p(0) \neq 0$ then there is a D so that $AD = DA = I$.
 (3) If $AB = CA = I$ then $B = C$.

It follows that any left or right inverse is a two-sided inverse and all inverses are equal.

Proof. First note that if $p(x) = a_0 + a_1x + \dots + a_dx^d$ then $a_0 = p(0)$.

Proof of (1): Suppose that $CA = I$ and that $p(0) = 0$. We then would have

$$0_{n \times n} = C0_{n \times n} = Cp(A) = a_1 + \dots + a_dA^{d-1}$$

which would contradict the minimality of $p(x)$. Thus, if $CA = I$, $p(0) \neq 0$. A similar argument works if $AB = I$.

Proof of (2): If $p(0) \neq 0$, we can solve for I in $p(A) = 0_{n \times n}$ to get

$$I = A(-a_1I - a_2A - \dots - a_dA^{d-1})\frac{1}{a_0} = (-a_1I - a_2A - \dots - a_dA^{d-1})\frac{1}{a_0}A.$$

This shows that $AD = DA = I$ is invertible with $D = (-a_1I - a_2A - \dots - a_dA^{d-1})\frac{1}{a_0}$.

Proof of (3): If $AB = CA = I$ then $C = CAB = B$. □

Lemma 13. *Let $A \in M_{n \times n}(\mathbb{k})$. If $p(x)$ is a minimal polynomial for A then $p(x + \lambda)$ is a minimal polynomial for $A - \lambda I$.*

Proof. For any polynomial $p(x)$ let $q(x) = p(x + \lambda)$. Then we also have $q(x - \lambda) = p(x)$. p and q will have the same degree and each will be monic if the other one is monic. Since $q(A - \lambda I) = 0_{n \times n}$ if and only if $p(A) = 0_{n \times n}$. The result follows. □

Corollary 14. *Let $p(x)$ be a minimal polynomial for A . The eigenvalues of A are exactly the roots of $p(x)$.*

Proof. λ is an eigenvalue if and only if $A - \lambda I$ is not invertible. $A - \lambda I$ is not invertible if and only if $p(x + \lambda)|_{x=0} = 0$. □

The following corollary is immediate.

Corollary 15. *If every $p(x) \in \mathbb{k}[x]$ has a root, then every $A \in M_{n \times n}(\mathbb{k})$ has an eigenvalue.*

The following theorem is taken for granted in calculus and proved in a typical abstract algebra or complex analysis class.

Theorem 16 (The “Fundamental Theorem of Algebra”). *Every $p(c) \in \mathbb{C}[x]$ has a root.*

The following corollary is immediate.

Corollary 17. *Every $A \in M_{n \times n}(\mathbb{C})$ has an eigenvalue. Every $A \in M_{n \times n}(\mathbb{R})$ has a (possibly complex) eigenvalue.*

3. COMPLEX NUMBERS AND COMPLEX INNER PRODUCT

See Bretscher for a review of complex arithmetic. We briefly review what we need here. In this section, we omit the arrows over vectors to make room for complex conjugation.

Remark 18. Recall that every $z \in \mathbb{C}$ has a unique representation $z = a + bi$ for $a, b \in \mathbb{R}$. Recall also that the complex conjugate of $z = a + bi$ is defined by $\bar{z} := a - bi$. We have the following facts about complex conjugation:

- (1) z is real if and only if $z = \bar{z}$.
- (2) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$.
- (3) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.
- (4) $\bar{z}z = a^2 + b^2$ is zero if and only if $z = 0$.

Definition 19. If $A = [z_{ij}] \in M_{l \times n}(\mathbb{C})$ we define $\bar{A} := [\bar{z}_{ij}]$ to be the matrix A with each entry replaced by its complex conjugate. In particular, we may apply this construction to a vector $v \in \mathbb{C}^n$.

Definition 20. If $v, w \in \mathbb{C}^n$ define $\langle v, w \rangle := \bar{v}^{\text{tr}} w$.

The following properties of \langle, \rangle are easy to verify. (We make no attempt to be complete here—just list the most important facts we will use.)

Proposition 21. For any $v, w \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$ we have:

- (1) $\langle w, v \rangle = \overline{\langle v, w \rangle}$.
- (2) $\langle v, v \rangle$ is a non-negative real number which is zero if and only if v is the zero vector.
- (3) $\langle \lambda v, w \rangle = \bar{\lambda} \langle v, w \rangle$.
- (4) $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$.

Definition 22. $A \in M_{n \times n}(\mathbb{C})$ is said to be Hermitian if $\bar{A}^{\text{tr}} = A$.

Lemma 23. A is Hermitian if and only if $\langle Av, w \rangle = \langle v, Aw \rangle$ for every $v, w \in \mathbb{C}^n$.

Proof. We prove one implication leaving the other as an exercise for the reader. If $A = \bar{A}^{\text{tr}}$ we have

$$\langle Av, w \rangle = \overline{(\bar{A}v)^{\text{tr}}} w = (\bar{A}\bar{v})^{\text{tr}} w = \bar{v}^{\text{tr}} \bar{A}^{\text{tr}} w = \bar{v}^{\text{tr}} Aw = \langle v, Aw \rangle.$$

□

Exercise 3. Finish the proof of lemma 23.

Proposition 24. Every eigenvalue of a Hermitian matrix is real.

Proof. Let $A \in M_{n \times n}(\mathbb{C})$ and suppose $Av = \lambda v$ for some $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^n$. We then have

$$\bar{\lambda} \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \langle v, \lambda v \rangle = \lambda \langle v, v \rangle.$$

Since $v \neq 0$ it follows that $\lambda = \bar{\lambda}$.

□

The corollary is now immediate.

Corollary 25. *Every (real or complex) eigenvalue of a symmetric matrix is real.*

Corollary 26. *Every symmetric real matrix has a real eigenvalue.*

Proof. Every real matrix has a (real or complex) eigenvalue by Corollary 17. This eigenvalue is real by Corollary 26. \square

4. THE SPECTRAL THEOREM

In this section we will no longer refer to complex numbers, so we revert to the convention of using arrows atop our vectors.

Theorem 27 (“The spectral theorem”). *Every symmetric matrix $A \in M_{n \times n}(\mathbb{R})$ has an orthonormal eigenbasis.*

Proof. In this proof, you may find it convenient to refer to the following diagram, which we will construct as we proceed:

$$\begin{array}{ccc}
 \mathbb{R}^n & \xrightarrow{A} & \mathbb{R}^n \\
 \uparrow S & & \uparrow S \\
 \mathbb{R}^n & \xrightarrow{\begin{bmatrix} \lambda & 0 \\ 0 & A' \end{bmatrix}} & \mathbb{R}^n \\
 \uparrow T & & \uparrow T \\
 \mathbb{R}^n & \xrightarrow{D} & \mathbb{R}^n
 \end{array}$$

$S^{-1}AS$

We prove the theorem by induction on n and organize our proof by minimal criminal: If the theorem fails, then there must be a smallest n for which it is false. Since every 1×1 matrix is diagonal, this n must be bigger than 1. Let A be an $n \times n$ matrix of this smallest dimension, we will proceed to show that A has an orthonormal eigenbasis. This will contradict our hypothesis that our theorem fails for this n proving the theorem.

Since A is a real symmetric matrix there is a nonzero $\vec{v} \in \mathbb{R}^n$ and a $\lambda \in \mathbb{R}$ so that $A\vec{v} = \lambda\vec{v}$. Let $W := (\text{span } \vec{v})^\perp$ and let $\{\vec{w}_1, \dots, \vec{w}_{n-1}\}$ be an orthonormal basis for W . Note that for every $\vec{w} \in W$ we have $A\vec{w} \cdot \vec{v} = \vec{w} \cdot A\vec{v} = \lambda\vec{w} \cdot \vec{v} = \vec{0}$. It follows that $A\vec{w} \in W$. Thus, if we take $S = [\vec{v} \quad \vec{w}_1 \quad \dots \quad \vec{w}_{n-1}]$ we have

$$[A]_{\{\vec{v}, \vec{w}_1, \dots, \vec{w}_{n-1}\}} = \begin{bmatrix} \lambda & 0_{1 \times (n-1)} \\ 0_{(n-1) \times 1} & A' \end{bmatrix} = S^{-1}AS = S^{\text{tr}}AS$$

for some $A' \in M_{(n-1) \times (n-1)}(\mathbb{R})$. Note that $(S^{\text{tr}}AS)^{\text{tr}} = S^{\text{tr}}AS$ so A' is symmetric.

Now, since n was minimal, A' has an orthonormal eigenbasis. That is to say there is a $T' \in O(n-1)$ and a diagonal $D' \in M_{(n-1) \times (n-1)}$ so that $(T')^{-1}A'T' = D'$. Set $T := \begin{bmatrix} 1 & 0_{1 \times (n-1)} \\ 0_{(n-1) \times 1} & T' \end{bmatrix}$. And $D = \begin{bmatrix} 1 & 0_{1 \times (n-1)} \\ 0_{(n-1) \times 1} & D' \end{bmatrix}$. We then have $D = T^{-1}S^{-1}AST$ (see the diagram) so ST diagonalizes A . Finally ST is orthogonal since $(ST)^{-1} = T^{-1}S^{-1} = T^{\text{tr}}S^{\text{tr}} = (ST)^{\text{tr}}$. \square