# 18.700 - LINEAR ALGEBRA, DAY 11
# THE MINIMAL POLYNOMIAL

SAM SCHIAVONE

## CONTENTS

I. Pre-class Planning 1
I.1. Goals for lesson 1
I.2. Methods of assessment 1
I.3. Materials to bring 1
II. Lesson Plan 2
II.1. Last time 2
II.2. 5A: Invariant subspaces and Eigenvectors, cont. 2
II.3. The Minimal Polynomial 4

## I. PRE-CLASS PLANNING

### I.1. Goals for lesson.

(1) Students will learn what it means to evaluate a polynomial at a linear operator.
(2) Students will learn the defintion of the minimal polynomial.
(3) Students will learn that the roots of the minimal polynomial are exactly the eigenvalues.
(4) Students will learn how to compute the eigenvalues and eigenvectors of a linear operator.

### I.2. Methods of assessment.

(1) Student responses to questions posed during lecture
(2) Student responses to worksheet

### I.3. Materials to bring.

(1) Laptop + adapter (2) Worksheets

(0:00)

## II. Lesson Plan

Announcements: • Exam grades posted?

### II.1. Last time.

- $V \cong W \iff \dim(V) = \dim(W)$.
- For a fixed choice of basis $\mathcal{B}$, the coordinate map

$$\begin{aligned} V &\to \mathbb{F}^n \\ v &\mapsto [v]_{\mathcal{B}} \end{aligned}$$

  is an isomorphism.
- Let $V$ and $W$ have dimension $n$ and $m$ with bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Then

$$\begin{aligned} \mathcal{L}(V, W) &\to M_{m\times n}(\mathbb{F}) \\ T &\mapsto {}_{\mathcal{C}}[T]_{\mathcal{B}} \end{aligned}$$

  is an isomorphism.
- $[T(v)]_{\mathcal{C}} = {}_{\mathcal{C}}[T]_{\mathcal{B}}[v]_{\mathcal{B}}$

$$\begin{array}{ccc} V & \xrightarrow{\quad T \quad} & W \\ {\scriptstyle \varphi_{\mathcal{B}}}\downarrow & & \downarrow{\scriptstyle \varphi_{\mathcal{C}}} \\ \mathbb{F}^n & \xrightarrow[\quad L_{{}_{\mathcal{C}}[T]_{\mathcal{B}}} \quad]{} & \mathbb{F}^m \end{array}$$

- **Proposition 1** (Change of basis formula)**.** *Suppose $\mathcal{B}$ and $\mathcal{C}$ are both bases of $V$. Given $T \in \mathcal{L}(V)$, then*

$${}_{\mathcal{B}}[T]_{\mathcal{B}} = ({}_{\mathcal{C}}[I]_{\mathcal{B}})^{-1}{}_{\mathcal{C}}[T]_{\mathcal{C}}\ {}_{\mathcal{C}}[I]_{\mathcal{B}}\,.$$

- Defined eigenvalues and eigenvectors.

### II.2. 5A: Invariant subspaces and Eigenvectors, cont.

**Definition 2.** Let $T \in \mathcal{L}(V)$. A scalar $\lambda \in \mathbb{F}$ is an *eigenvalue* of $T$ if there exists $v \in V$ with $v \neq 0$ such that $T(v) = \lambda v$. Such a $v$ is called an *eigenvector* corresponding to $\lambda$.

**Remark 3.**

- "eigen-" means "self" or "own". An eigenvector maps into its own span under $T$.
- We require that $v \neq 0$ because $T(0) = \lambda 0$ for all $\lambda \in \mathbb{F}$.

[Show gif depicting eigenvectors in $\mathbb{R}^2$: `https://upload.wikimedia.org/wikipedia/commons/a/ad/Eigenvectors-extended.gif`.]

**Theorem 4.** *Suppose $V$ is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{F}$. TFAE.*

*(i) $\lambda$ is an eigenvalue of $T$.*
*(ii) $T - \lambda I$ is not injective.*
*(iii) $T - \lambda I$ is not surjective.*
*(iv) $T - \lambda I$ is not invertible.*

*Proof.* (a) $\implies$ (b): Assume $\lambda$ is an eigenvalue of $T$ with corresponding eigenvector $v \neq 0$, so $T(v) = \lambda v$. Then

$$0 = T(v) - \lambda v = (T - \lambda I)(v)$$

so $0 \neq v \in \ker(T)$. Thus $T$ is not one-to-one.

(b) $\implies$ (a): Assume $T - \lambda I$ is not injective. Then $\ker(T - \lambda I) \neq \{0\}$ so there exists $0 \neq v \in \ker(T - \lambda I)$. Then

$$0 = (T - \lambda I)(v) = T(v) - \lambda v \implies T(v) = \lambda v$$

so $v$ is an eigenvector with eigenvalue $\lambda$.

We previously showed the equivalence of (b), (c), and (d). □

**Proposition 5.** *Let $T \in \mathcal{L}(V)$. Suppose that $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of $T$ with corresponding eigenvectors $v_1, \ldots, v_k$. Then $v_1, \ldots, v_k$ are linearly independent.*

*Proof.* We proceed by induction on $k$, the number of eigenvalues.

Base case: $k = 1$. An eigenvector is nonzero by definition, so the list $v_1$ is linearly independent by a previous homework problem.

Inductive step: Assume the result holds for $k - 1$ and assume $T$ has $k$ distinct eigenvalues. Suppose that

$$a_1 v_1 + \cdots + a_k v_k = 0 \tag{6}$$

for some $a_1, \ldots, a_k \in \mathbb{F}$. Goal: $a_i = 0$ for all $i$. Note that

$$(T - \lambda_k I)(v_i) = T(v_i) - \lambda_k v_i = \lambda_i v_i - \lambda_k v_i = (\lambda_i - \lambda_k) v_k$$

for all $i = 1, \ldots, k$. Applying $T - \lambda_k I$ to (6), we find

$$\begin{aligned} 0 &= (T - \lambda_k I)(a_1 v_1 + \cdots + a_k v_k) \\ &= a_1(\lambda_1 - \lambda_k) v_1 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1} + \cancelto{0}{a_k(\lambda_k - \lambda_k) v_k}\,. \end{aligned}$$

Since $v_1, \ldots, v_{k-1}$ are linearly independent by the inductive hypothesis, then $a_i(\lambda_i - \lambda_k) = 0$ for all $i = 1, \ldots, k-1$. Since the $\lambda_i$ are distinct, then $a_i = 0$ for all $i = 1, \ldots, k-1$. Then (6) becomes

$$a_k v_k = 0\,.$$

But $v_k$ is an eigenvector, hence is nonzero, so $a_k = 0$ by the base case. □

**Corollary 7.** *If $V$ is finite-dimensional, then every operator $T \in \mathcal{L}(V)$ has at most* $\dim(V)$ *distinct eigenvalues.*

*Proof.* Apply the previous result and LI $\leq$ span. □

II.2.1. *Polynomials applied to linear operators.* Given a linear operator $T : V \to V$, then we can compose $T$ with itself: $T \circ T = T^2$. We similarly define

$$T^m = \begin{cases} \overbrace{T \cdots T}^{m \text{ times}} & \text{if } m > 0; \\ I & \text{if } m = 0; \\ (T^{-1})^{|m|} & \text{if } m < 0 \text{ and } T \text{ is invertible.} \end{cases}$$

**Lemma 8.**

- $T^m T^n = T^{m+n}$
- $(T^m)^n = T^{mn}$

*Proof.* Exercise. □

**Definition 9.** Given $T \in \mathcal{L}(V)$, and $p \in \mathcal{P}(\mathbb{F})$ with

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m \,,$$

define the operator $p(T) \in \mathcal{L}(V)$ by

$$p(T) := a_0 I + a_1 T + \cdots + a_m T^m \,.$$

**Definition 10.** Let $p, q \in \mathcal{P}(\mathbb{F})$. Their product $pq$ is defined pointwise:

$$(pq)(z) := p(z)q(z)$$

for all $z \in \mathbb{F}$.

Note that multiplication of polynomials is commutative. The same is true when we apply polynomials to linear operators.

**Lemma 11.**

*(i)* $(pq)(T) = p(T)q(T)$;
*(ii)* $p(T)q(T) = q(T)p(T)$.

*Proof.* Exercise. □

**Lemma 12.** *Let $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbb{F})$. Then* $\ker(p(T))$ *and* $\operatorname{img}(p(T))$ *are $T$-invariant.*

*Proof.* Exercise. □

II.3. **The Minimal Polynomial.**

**Definition 13.** A polynomial is *monic* if its leading coefficient is 1.

**Example 14.** $4x^3 - 3x + 1$ is *not* monic. $x^5 - 2x^2 + 3$ is monic.

**Theorem 15.** *Let $V$ be a finite-dimensional vector space and $T \in \mathcal{L}(V)$. There is a unique monic polynomial $m \in \mathcal{P}(\mathbb{F})$ of minimum degree such that $m(T) = 0$. Moreover,* $\deg(m) \leq \dim(V)$.

*Proof.* Let $n := \dim(V)$.

Existence: We proceed by strong induction on $n$. Base case: $n = 0$. Then $V = \{0\}$, so $I$ is the zero operator on $V$. Thus we can take $m$ to be the constant polynomial 1.

Inductive step: Now assume that $n \geq 1$ and the result holds for all vector spaces of dimension $< n$. Choose a nonzero $u \in V$ and consider

$$u, T(u), T^2(u), \ldots, T^n(u) \,.$$

Since this list consists of $n + 1$ vectors, then it must be [ask students] linearly dependent. By the Linear Dependence Lemma, then there is a minimal positive integer $d \in \{1, \ldots, n\}$ such that

$$T^d(u) \in \operatorname{span}(u, T(u), \ldots, T^{d-1}(u)) \,.$$

Then

$$c_0 u + c_1 T(u) + \cdots + c_{d-1} T^{d-1}(u) + T^d(u) = 0$$

for some $c_0, \dots, c_{d-1} \in \mathbb{F}$, not all zero. Letting

$$q(z) := c_0 + c_1 z + \cdots + c_{d-1} z^{d-1} + z^d \in \mathcal{P}(\mathbb{F}),$$

then $q(T)u = 0$. Note that

$$q(T)(T^k(u)) = T^k(q(T)(u)) = T^k(0) = 0 \tag{16}$$

for all $k \in \mathbb{Z}_{\geq 0}$. Since we chose $d$ to be minimal, then $u, T(u), \dots, T^{d-1}(u)$ is linearly independent. Since these are all in $\ker(q(T))$ by (16), then $\dim(\ker(q(T))) \geq d$. By Rank-Nullity, then

$$\dim(\operatorname{img}(q(T)) = \dim(V) - \dim(\ker(T)) \leq \dim(V) - d.$$

By a previous result, $\operatorname{img}(q(T))$ is $T$-invariant, so we can apply the inductive hypothesis to the restriction $T|_{\operatorname{img}(q(T))}$. Thus there is a monic polynomial $s \in \mathcal{P}(\mathbb{F})$ such that

$$s(T|_{\operatorname{img}(q(T))}) = 0 \qquad \text{and} \qquad \deg(s) \leq \dim(\operatorname{img}(q(T))) \leq \dim(V) - d.$$

Consider the product $(sq)(z) = s(z)q(z)$. Given $v \in V$, then

$$((sq)(T))(v) = s(T)(q(T)(v)) = 0$$

since $s(T)|_{\operatorname{img}(T)} = s(T|_{\operatorname{img}(T)}) = 0$. Thus $sq$ is a monic polynomial with $(sq)(T) = 0$ and $\deg(sq) \leq \dim(V)$.

Uniqueness: [Leave as exercise if necessary.] Suppose that $m_1$ and $m_2$ are both monic polynomials of smallest degree such that $m_1(T) = 0$ and $m_2(T) = 0$. Consider $m_1 - m_2$. We have $(m_1 - m_2)(T) = 0$ and since both $m_1$ and $m_2$ are monic, then $\deg(m_1 - m_2) < \deg(m_1)$. If $m_1 - m_2 \neq 0$, then we can rescale $m_1 - m_2$ by the reciprocal of its leading coefficient, obtaining a monic polynomial stricly smaller degree, contradiction. Thus $m_1 - m_2 = 0$, i.e., $m_1 = m_2$. □

**Definition 17.** With notation as above, the *minimal polynomial* of $T$ is $m$, i.e., the unique polynomial of smallest degree such that $m(T) = 0$. It is denoted $\operatorname{minpoly}(T)$.

**Corollary 18.** *Let $V$ be a nonzero finite-dimensional $\mathbb{C}$-vector space and $T \in \mathcal{L}(V)$. Then $T$ has an eigenvalue.*

*Proof.* Let $m := \operatorname{minpoly}(T)$. Note that $m$ is nonconstant: if $m = c$ were constant, then we would have $cI = 0$, contradicting the fact that $V \neq \{0\}$.

By the Fundamental Theorem of Algebra, there exists $\lambda \in \mathbb{C}$ such that $m(\lambda) = 0$. Then

$$m(z) = (z - \lambda)q(z)$$

for some monic $q \in \mathcal{P}(\mathbb{C})$. Then

$$0 = m(T) = (T - \lambda I)q(T).$$

Since $\deg(q) < \deg(m)$ and $m$ is the minpoly, then $q(T) \neq 0$. Then there is some vector $v \in V$ such that $q(T)(v) \neq 0$. Then

$$0 = m(T)(v) = (T - \lambda I)(q(T)(v))$$

so $q(T)(v)$ is an eigenvector of $T$ with eigenvalue $\lambda$. □

**Remark 19.** Here we used the fact that $\mathbb{C}$ is algebraically closed in an important way. The result is not true over $\mathbb{R}$!

**Example 20.** Consider the right shift operator

$$R : \mathbb{F}^\infty \to \mathbb{F}^\infty$$
$$(x_1, x_2, \ldots) \mapsto (0, x_1, x_2, \ldots) .$$

Then $R$ has no eigenvectors and no eigenvalues (exercise). [Ask students why this doesn't contradict theorem.]

**Corollary 21.** *With notation as above, the eigenvalues of $T$ are exactly the roots of* $\operatorname{minpoly}(T)$.

*Proof.* We have seen that all the roots of $m := \operatorname{minpoly}(T)$ are eigenvalues of $T$. Conversely, suppose $\lambda \in \mathbb{F}$ is an eigenvalue of $T$. Then there exists $0 \neq v \in V$ such that $T(v) = \lambda v$. Applying $T$ to both sides repeatedly, we see that $T^k(v) = \lambda^k v$ for all $k \in \mathbb{Z}_{\geq 0}$. Taking appropriate linear combinations of these monomials, we have [write "0 = . . . " last]

$$0 = m(T)v = m(\lambda)v .$$

Since $v \neq 0$, then $m(\lambda) = 0$. □

Q: Given a linear operator $T$, how can we compute its eigenvalues and eigenvectors?
A:

(1) To compute $\operatorname{minpoly}(T)$, we need to find the smallest $d$ such that

$$c_0 I + c_1 T + \cdots + c_{d-1} T^{d-1} = -T^d$$

has a solution for $c_0, \ldots, c_{d-1} \in \mathbb{F}$. We can choose a basis $\mathcal{B}$ for $V$ and apply $[\cdot]_\mathcal{B}$ to the above equation. This produces a matrix equation which can be thought of as a linear system of $(\dim(V))^2$ equations in $d$ unknowns.

This yields the following algorithm: for each $d = 1, 2, \ldots$, see if the above system of equations has a solution. By the theorem, this algorithm terminates at the latest when $d = \dim(V)$.

(2) Usually faster, but not guaranteed to always work: choose $v \in V, v \neq 0$ and consider the equation

$$c_0 v + c_1 T(v) + \cdots + c_{n-1} T^{n-1}(v) = -T^n(v)$$

where $n := \dim(V)$. Again, by choosing a basis for $V$ and applying $[\cdot]_\mathcal{B}$, we obtain a system of $n$ equations in the $n$ unknowns $c_0, \ldots, c_{n-1}$. If the solution to this system is unique, this yields the coefficients of $\operatorname{minpoly}(T)$.

**Proposition 22.** *Suppose $V$ is finite-dimensional, $T \in \mathcal{L}(V)$ and $q \in \mathcal{P}(\mathbb{F})$. Then $q(T) = 0$ iff* $\operatorname{minpoly}(T)$ *divides $q$, i.e., $q =$* $\operatorname{minpoly}(T)f$ *for some $f \in \mathcal{P}(\mathbb{F})$.*

*Proof idea.* Use the division algorithm to divide $q$ by $\operatorname{minpoly}(T)$ and consider the remainder. □

**Corollary 23.** *With the same assumptions, suppose $U$ is a $T$-invariant subspace of $V$. Then* $\operatorname{minpoly}(T|_U)$ *divides* $\operatorname{minpoly}(T)$.

**Corollary 24.** *With the same assumptions, $T$ is not invertible iff the constant term of* $\operatorname{minpoly}(T)$ *is* 0.

*Proof.* Let $m := \operatorname{minpoly}(T)$. Then

$$\begin{aligned} T \text{ is not invertible} &\iff 0 \text{ is an eigenvalue of } T \\ &\iff 0 \text{ is a zero of } p \\ &\iff p(0) = 0 \\ &\iff \text{the constant term of } p \text{ is } 0. \end{aligned}$$

$\square$