Math 24: Winter 2021
Lecture 21

Dana P. Williams

Dartmouth College

Wednesday, February 24, 2021

Let's Get Started

1. We should be recording.
2. Remember, it is more comfortable for me if you turn on your video so that I feel like I am talking to real people.
3. We will be skipping §5.3 and playing in §5.4 for a bit. Then on to §6.1.
4. But first, are there any questions from last time?

Review

Definition

A polynomial $p(\lambda) \in \mathrm{P}(\mathbf{F})$ splits over $\mathbf{F}$ if there are scalars $c, a_1, \dots, a_n \in \mathbf{F}$ such that

$$p(\lambda) = c(\lambda - a_1) \cdots (\lambda - a_n)$$

Theorem

Suppose that V is a finite-dimensional vector space over $\mathbf{F}$ and that $T \in \mathcal{L}(V)$ is diagonalizable. Then the characteristic polynomial of T splits over $\mathbf{F}$.

Review

Definition

Suppose that V is a finite-dimensional vector space and $T \in \mathcal{L}(V)$. If λ_0 is an eigenvalue for T, then the multiplicity of λ_0 is the largest positive integer k such that $(\lambda - \lambda_0)^k$ is a factor of the characteristic polynomial $p(\lambda)$ of T. We use the same terminology for the eigenvalues of a matrix.

Theorem

Suppose that V is a finite-dimensional vector space and $T \in \mathcal{L}(V)$. If λ is an eigenvalue of T of multiplicity m, then

$$1 \leq \dim(E_\lambda) \leq m$$

where E_λ is the λ-eigenspace for T.

Review

Theorem (Diagonalizability Theorem)

Suppose that V is a finite-dimensional vector space over $\mathbf{F}$ and $T \in \mathcal{L}(V)$ is a linear operator whose characteristic polynomial splits over $\mathbf{F}$. Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T with multiplicities $m_1, \dots, m_k$, respectively.

1. *Then T is diagonalizable if and only if $\dim(E_{\lambda_i}) = m_i$ for $1 \leq i \leq k$.*
2. *If T is diagonalizable and β_i is a basis for E_{λ_i} for $1 \leq i \leq k$, then $\beta = \beta_1 \cup \cdots \cup \beta_k$ is a basis for V of eigenvectors for T.*

Invariant Subspaces

Definition

Let T be a linear operator on a vector space V. A subspace W of V is called a T-invariant subspace, or just invariant, if $v \in W$ implies $T(v) \in W$. Alternatively, $T(W) \subset W$.

Example (Low Hanging Fruit)

If $T \in \mathcal{L}(V)$, then the following are always invariant subspaces.

1. $W = \{0_V\}$.
2. $W = V$.
3. $W = \mathrm{R}(T)$.
4. $W = \mathrm{N}(T)$.
5. $W = E_\lambda$ for any eigenvalue λ for T.

Solution

These are all relatively easy to check and I hope you will do so. Consider this an (easy) unassigned homework problem.

Example

Example

Define $T \in \mathcal{L}(\mathbf{R}^3)$ by $T(x, y, z) = (x + y, y + z, z + x)$ and let $W = \{ (x, y, z) \in \mathbf{R}^3 : x + y + z = 0 \}$. Then if $(a, b, c) \in W$, $T(a, b, c) = (a + b, b + c, c + a)$ and

$$(a + b) + (b + c) + (c + a) = 2(a + b + c) = 0.$$

That is $T(a, b, c) \in W$ and W is an invariant subspace.

Definition

Suppose $T \in \mathcal{L}(V)$ and $v \in V$. Then

$$W = \operatorname{Span}\{\, v, T(v), T^2(v), \dots \}$$

is called the *T-cyclic subspace* of V generated by v.

Remark

It is a homework exercise to verify that the T-cyclic subspace W generated by v is T-invariant and is contained in any invariant subspace that contains v. Otherwise said, W is the smallest T-invariant subspace containing v.

Example

Example

Let $T \in \mathcal{L}(\mathbf{R}^3)$ be given by $T(x, y, z) = (x + y, y + z, z + x)$ as before. Let $v = (1, -1, 0)$. Then $T(v) = (0, -1, 1)$, $T^2(v) = (-1, 0, 1)$, $T^3(v) = (-1, 1, 0) = -v$. Therefore

$$\begin{aligned} W &= \operatorname{Span}\{\, v, T(v), T^2(v), \dots \,\} \\ &= \operatorname{Span}\{\, (1, -1, 0), (0, -1, 1), (-1, 0, 1) \,\} \\ &= \operatorname{Span}\{\, (1, -1, 0), (0, -1, 1) \,\} \\ &= \{\, (x, y, z) \in \mathbf{R}^3 : x + y + z = 0 \,\} \end{aligned}$$

Return

Break Time

Time for a break. Questions?

Restrictions

Remark

Suppose that $T \in \mathcal{L}(V)$ and that W is a T-invariant subspace of V. Then we can define $T_W : W \to W$ by $T_W(v) = T(v)$ for all $v \in W$. It is a straightforward exercise to see that T_W is linear.

Definition

If $T \in \mathcal{L}(V)$ and if W is a T-invariant subspace of V, then the operator $T_W \in \mathcal{L}(W)$ defined above is called the restriction of T to W.

Remark

Note that since the restriction T_W is itself a bonafide linear operator, it has its own characteristic polynomial, eigenvalues, and eigenvectors. Our next result examines how these are related to those for T.

Characteristic Polynomials

Theorem

Suppose that $T \in \mathcal{L}(V)$ for a finite-dimensional vector space V and that W is a T-invariant subspace of V. Then the characteristic polynomial $p_W(\lambda)$ of T_W divides the characteristic polynomial $p(\lambda)$ of T.

Proof.

Let $\gamma = \{ v_1, \dots, v_k \}$ be an ordered basis for W. Let $\beta = \{ v_1, \dots, v_k, v_{k+1}, \dots, v_n \}$ be an extension of γ to an ordered basis of V. Let

$$A = [T]_\beta = \big[[T(v_1)]_\beta \ \cdots \ [T(v_k)]_\beta \ [T(v_{k+1})]_\beta \ \cdots \ [T(v_n)]_\beta \big].$$

Proof

Proof Continued.

Since $T(v_i) \in W$ for $1 \leq i \leq k$, the corresponding coordinate vector $[T(v_i)]_\beta$ is of the form $(\underbrace{a_1, \ldots, a_k}_{[T_W(v_i)]_\gamma}, \underbrace{0, \cdots, 0}_{n-k})$. Hence

$$[T]_\beta = \begin{pmatrix} [T_W]_\gamma & B \\ O & C \end{pmatrix}.$$

Hence, using our cute lemma from last lecture,

$$\begin{aligned} p(\lambda) = \det([T]_\beta - \lambda I_n) &= \det \begin{pmatrix} [T_W]_\gamma - \lambda I_k & B \\ O & C - \lambda I_{n-k} \end{pmatrix} \\ &= \det\big([T]_\beta - \lambda I_n\big) \det\big(C - \lambda I_{n-k}\big) \\ &= p_W(\lambda) g(\lambda) \end{aligned}$$

where $p_W(\lambda)$ is the characteristic polynomial of T_W and $g(\lambda)$ is a polynomial. Thus $p_W(\lambda)$ divides $p(\lambda)$ as claimed. □

Cyclic Subspaces

Theorem

Suppose that $T \in \mathcal{L}(V)$ *for a finite-dimensional vector space* V*. Suppose that* $v \in V$ *is nonzero and let* $W = \operatorname{Span}\{\, v, T(v), \dots \}$ *be the* T*-cyclic subspace generated by* v*. Let* $k = \dim(W)$*.*

1. $\beta = \{\, v, T(v), \dots, T^{k-1}(v) \,\}$ *is an ordered basis for* W*.*
2. *If*
$$a_0 v + a_1 T(v) + \cdots a_{k-1} T^{k-1}(v) + T^k(v) = 0_V,$$
then the characteristic polynomial of T_W *is*
$$p_W(\lambda) = (-1)^k \bigl(a_0 + a_1\lambda + \cdots a_{k-1}\lambda^{k-1} + \lambda^k .\bigr)$$

Proof

Proof.

Since $v \neq 0_V$, $\{v\}$ is linearly independent. Hence there is a largest $j \geq 1$ such that $\beta = \{\, v, T(v), \dots, T^{j-1}(v)\,\}$ is linearly independent. (Note that we could have $j = 1$! Then, here and elsewhere, we interpret $T^0 = I_V$.) Such a j exists because V is finite dimensional. Let $Z = \text{Span}(\beta)$. Note that β is a basis for Z. We must have $T^j(v) \in Z$—otherwise $\beta \cup \{\, T^j(v)\}$ would be linearly independent. I claim that Z is T-invariant. Let $w \in Z$. Then there are scalars a_i such that $w = a_0 v + a_1 T(v) + \cdots a_{j-1} T^{j-1}(v)$. But then

$$T(w) = a_0 T(v) + \cdots + a_{j-2} T^{j-1}(v) + a_{j-1} T^j(v).$$

Since $T^j(v) \in Z = \text{Span}(\beta)$, so it $T(w)$. This proves the claim. Since $v \in Z$ and Z is T-invariant, it follows from homework that $W \subset Z$. But we clearly have $Z \subset W$. Hence $W = Z$ and $\dim(W) = j$. Thus $j = k$ and we have proved item (1).

Proof

Proof Continued.

(2) Let $\beta = \{ v, T(v), \dots, T^{k-1}(v) \}$ be our ordered basis from the first part of the proof. Since $T^k(v) \in W$, there are unique scalars such that

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

Then

$$[T_W]_\beta = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix}.$$

You get to use induction, on homework, to calculate that the characteristic polynomial of T_W is

$$p_W(\lambda) = (-1)^k (a_0 + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k).$$

□

Example

Example

As in our earlier example, let $T(x, y, z) = (x + y, y + z, z + x)$ and let W be the cyclic subspace generated by $v = (1, -1, 0)$. Then $\dim(W) = 2$ and

$$v - T(v) + T^2(v) = (1, -1, 0) - (0, -1, 1) + (-1, 0, 1) = 0_{\mathbf{R}^3}.$$

By the theorem, $p_W(\lambda) = (-1)^2(1 - \lambda + \lambda^2)$.

But $\beta = \{ (1, -1, 0), (0, -1, 1) \}$ is a basis for W and

$$[T_W]_\beta = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}.$$

Thus we also can compute that

$$p_W(\lambda) = \det\begin{pmatrix} -\lambda & -1 \\ 1 & 1-\lambda \end{pmatrix} = \lambda^2 - \lambda + 1.$$

Break Time

Time for a break and questions.

Remark

Suppose that V is a vector space over $\mathbf{F}$ and $T \in \mathcal{L}(V)$. Then if $p \in \mathrm{P}(\mathbf{F})$ is given by $p(t) = a_0 + a_1 t + \cdots + a_n t^n$, then we get a new operator $p(T) \in \mathcal{L}(V)$ by letting

$$p(T) = a_0 I_V + a_1 T + \cdots + a_n T^n.$$

Note the appearance of I_V in the "constant term". As before, we can think of I_V as T^0. Then if $v \in V$,

$$p(T)(v) = a_0 v + a_1 T(v) + \cdots + a_n T^n(v).$$

For example, if $p(t) = t^2 - 1$, then $p(T) = T^2 - I_V$ and $p(T)(v) = T^2(v) - v$. Of course, we can play the same game with matrices $A \in M_{n\times n}(\mathbf{F})$. Then $p(A) = A^2 - I_n$. For example, if $p(t) = t - \lambda$, then the characteristic polynomial of A is $\det(p(A)) = \det(A - \lambda I_n)$.

Killing T

Remark

If $\dim(V) = n$, then since $\mathcal{L}(V)$ is isomorphic to $M_{n\times n}(\mathbf{F})$, $\dim(\mathcal{L}(V)) = n^2$. Therefore the set $\{ I_V, T, T^2, \dots, T^{n^2} \}$ is linearly dependent and there are scalars $a_0, \dots, a_{n^2}$ such that

$$a_0 I_V + a_1 T + \cdots + a_{n^2} T^{n^2} = T_0$$

Thus if we let $p \in P_{n^2}(\mathbf{F})$ be given by $p(t) = a_0 + a_1 t + \cdots + a_{n^2} t^{n^2}$, then

$$p(T) = T_0$$

where as always T_0 is the zero map. Of course, if $A \in M_{n\times n}(\mathbf{F})$, we can do the analogous thing and find $p \in P_{n^2}(\mathbf{F})$ such that $p(A)$ is the zero matrix. We now have the tools to show that we can do the same thing with a polynomial of degree n. In fact, we will show that the characteristic polynomial always does the job.

Cayley-Hamilton Theorem

Theorem (Cayley-Hamilton Theorem)

Suppose that V is a finite-dimensional vector space and that $T \in \mathcal{L}(V)$. If $p(\lambda)$ is the characteristic polynomial of T, then $p(T) = T_0$.

Remark

I the language of the text, we way that T "satisfies" its characteristic polynomial.

Proof

Proof.

It will suffice to show that $p(T)(v) = 0_V$ for all $v \in V$. Since $p(T)$ is a linear operator, this is automatic if $v = 0_V$, so we assume $v \neq 0_V$. We let W be the T-cyclic subspace generated by v and suppose that $\dim(W) = k$. Then as in the proof of our theorem, there are constants a_i such that

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

Then we proved that the characteristic polynomial of T_W is

$$p_W(\lambda) = (-1)^k(a_0 + a_1\lambda + \cdots + a_{k-1}\lambda^{k-1} + \lambda^k).$$

Therefore

$$p_W(T)(v) = (-1)^k\big(a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + T^k(v)\big)$$

$$= (-1)^k 0_V = 0_V.$$

Proof

Proof Continued.

We also proved that $p_W(\lambda)$ divides $p(\lambda)$ so that $p(\lambda) = g(\lambda)p_W(\lambda)$ for some $g \in P_n(\mathbf{F})$. But then,

$$p(T)(v) = g(T)p_W(T)(v) = g(T)\big(p_W(T)(v)\big) = g(T)(0_V) = 0_V.$$

This completes the proof of the theorem. □

Corollary

If $A \in M_{n\times n}(\mathbf{F})$ and $p(\lambda)$ is the characteristic polynomial of A, then $p(A) = O$.

Sketch of the Proof.

By the theorem, $p(L_A) = T_0$. Consider $[p(L_A)]_\sigma$ where σ is the standard basis. □

Example

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then its characteristic polynomial is $p(\lambda) = \lambda^2 - 5\lambda - 2$. Then

$$\begin{aligned} p(A) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^2 - 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

as per the theorem.

Enough

1. That is enough for today.