Math 24: Winter 2021
Lecture 20

Dana P. Williams

Dartmouth College

Monday, February 22, 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. But first, are there any questions from last time?

Eigenspaces

Definition

Suppose that $T \in \mathcal{L}(V)$ is a linear operator on a vector space V. If λ is an eigenvalue for T, then $E_\lambda = \{ v \in V : T(v) = \lambda v \}$ is called that λ-eigenspace for T. If $A \in M_{n\times n}(\mathbf{F})$ and λ is an eigenvalue for A, then the λ-eigenspace for A is just the λ-eigenspace for L_A.

Theorem

Suppose that V is a vector space and $T \in \mathcal{L}(V)$. Suppose that $\lambda_1, \dots, \lambda_k$ are distinct eigenvalues for T. For each $1 \le i \le k$, let S_i be a finite linear independent in E_{λ_i}. Then $S = S_1 \cup \dots \cup S_k$ is linear independent.

Corollary

Suppose that V is n-dimensional and that $T \in \mathcal{L}(V)$ has n-distinct eigenvalues. Then T is diagonalizable.

Polynomials

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

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

Example

For example, the polynomial $p(\lambda) = -\lambda^3 + 6\lambda^2 - 11\lambda + 6$ factors as $-(\lambda - 1)(\lambda - 2)(\lambda - 3)$, so it splits over $\mathbf{R}$. (Unfortunately, there is no easy way to see this other than multiplying out the right-hand side.) But the polynomial $q(\lambda) = \lambda^3 - \lambda^2 + \lambda - 1 = \lambda^2(\lambda - 1) + \lambda - 1 = (\lambda - 1)(\lambda^2 + 1)$ does not split over $\mathbf{R}$. However, it does split over $\mathbf{C}$: $q(\lambda) = (\lambda - 1)(\lambda - i)(\lambda + i)$.

The Joy of Splitting

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 $\mathbf{F}$.

Proof.

Since T is diagonalizable, there is a basis $\beta = \{ v_1, \dots, v_n \}$ of eigenvectors for T. Then $[T]_\beta = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$ is diagonal. Then the characteristic polynomial of T is

$$p(\lambda) = \det(D - \lambda I_n) = \det \begin{pmatrix} \lambda_1 - \lambda & 0 & \cdots & 0 \\ 0 & \lambda_2 - \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n - \lambda \end{pmatrix}$$
$$= (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n).$$

Therefore $p(\lambda)$ splits over $\mathbf{F}$. □

Converse

Remark

The converse of the preceding theorem does not hold. We know that

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

is not diagonalizable, but its characteristic polynomial $p(\lambda) = (\lambda - 1)^2$ certainly splits.

Multiplicity

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.

Example

Let $A = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 2 \end{pmatrix}$. Then the characteristic polynomial of A is $p(\lambda) = (\lambda - 3)(\lambda - 2)^3$. Thus $\lambda = 3$ has multiplicity 1 while $\lambda = 2$ has multiplicity 3.

Theorem on Multiplicity

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.

Remark

For the proof, we need a little lemma on determinants of block matrices that is of some interest in itself.

A Lemma

Lemma

Let $D \in M_{n\times n}(\mathbf{F})$ have the block form

$$D = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

where $A \in M_{k\times k}(\mathbf{F})$ and $C \in M_{(n-k)\times(n-k)}(\mathbf{F})$. Then $\det(D) = \det(A)\det(C)$*.*

Proof of the Lemma.

We will use the result that can compute $\det(D)$ be expanding down the first column. Then the result is clear if $k = 1$. So we assume we know the result if $k - 1 \geq 1$. Then

$$\det(D) = \sum_{j=1}^{n} (-1)^{i+1} D_{i1} \det(\widetilde{D}_{i1}) = \sum_{j=1}^{k} (-1)^{i+1} A_{i1} \det(\widetilde{D}_{i1}).$$

Proof

Proof Continued.

But $\widetilde{D}_{i1}$ is of the form

$$D_{i1} = \begin{pmatrix} \widetilde{A}_{i1} & B(i) \\ 0 & C \end{pmatrix}$$

for an $(k-1) \times (n-k)$ matrix $B(i)$. Since $\widetilde{A}_{i1}$ is $(n-1) \times (n-1)$, our induction hypothesis implies that,

$$\begin{aligned} \det(D) &= \Big(\sum_{j=1}^{k} (-1)^{i+1} A_{i1} \det(\widetilde{A}_{i1})\Big) \det(C) \\ &= \det(A) \det(C) \end{aligned}$$

□

Proof of the Theorem

Proof of the Theorem.

We need to prove the $1 \leq \dim(E_\lambda) \leq m$ where m is the multiplicity of λ. Let $\alpha = \{ v_1, \dots, v_p \}$ be an ordered basis for E_λ. (Since λ is an eigenvalue, $p \geq 1$!) Now extend α to a basis $\beta = \{ v_1, \dots, v_p, v_{p+1}, \dots, v_n \}$ for V. Since $[T(v_i)]_\beta = [\lambda v_i]_\beta = \lambda e_i$ if $1 \leq i \leq p$, $[T]_\beta$ has the form

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

Therefore the characteristic polynomial of T is

$$p(t) = \det(A - tI_n) = \det \begin{pmatrix} \lambda I_p - tI_p & B \\ O & C - tI_{n-p} \end{pmatrix}$$
$$= \det \begin{pmatrix} (\lambda - t) I_p & B \\ O & C - tI_{n-p} \end{pmatrix}$$

Proof

Proof Continued.

By our lemma,

$$\begin{aligned} p(t) &= \det\big((\lambda - t)I_p\big)\det(C - tI_{n-p}) \\ &= (\lambda - t)^p g(t) = (-1)^p (t-\lambda)^p g(t) \end{aligned}$$

where $g(t)$ is a polynomial. Thus $(t-\lambda)^p$ is a factor of $p(t)$. Therefore $p \leq m$. □

Break Time

Time for a break and some questions.

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}, then $\beta = \beta_1 \cup \cdots \cup \beta_k$ is a basis for V of eigenvectors for T.*

Proof

Proof.

Let $n = \dim(V)$ and $d_i = \dim(E_{\lambda_i})$ for $1 \leq i \leq k$. Observe that since the characteristic polynomial splits, $\sum_{i=1}^{k} m_i = n$.

Suppose that T is diagonalizable and that β is a basis for V consisting of eigenvectors for T. Note that every element of β is in some E_{λ_i}. Let n_i be the number of vectors in $\beta \cap E_{\lambda_i}$. Since any subset of β is linearly independent and $\dim(E_{\lambda_i}) = d_i$, we have $n_i \leq d_i$, and $d_i \leq m_i$ be our theorem. Thus

$$n = \sum_{i=1}^{k} n_i \leq \sum_{i=1}^{k} d_i \leq \sum_{k=1}^{k} m_i = n.$$

Thus each sum equals n and

$$\sum_{i=1}^{k} (m_i - d_i) = 0.$$

Since $m_i - d_i \geq 0$ for each i, we have $m_i = d_i$ for all i. This proves the $\implies$ implication of item (1).

Proof

Proof Continued.

Conversely, assume that $d_i = m_i$ for all i. Let β_i be a basis for E_{λ_i} and let $\beta = \bigcup_{i=1}^k \beta_i$. We proved that β must be linearly independent. But β has

$$\sum_{i=1}^{k} d_i = \sum_{i=1}^{k} m_i = n$$

elements. Since $\dim(V) = n$ and β is linearly independent, it is basis of eigenvalues. This shows T is diagonalizable.

This proves the other implication for item (1) as well as item (2). □

Practice

Remark

Now given a linear operator on a n-dimensional vector space V, we are now prepared to decide whether T is diagonalizable or not.

1. First, the characteristic polynomial of T must split. If it does not, then T is not diagonalizable.
2. Second, for every eigenvalue λ of T the dimension of the eigenspace E_λ must equal the multiplicity m of λ. Since $\dim(E_\lambda) = \mathrm{N}(T - \lambda I_V)$, the multiplicity m must equal $n - \mathrm{rank}(T - \lambda I_V)$.

Of course, these considerations carry over almost word for word for a square matrix A.

Example

Example

Consider the matrix $A = \begin{pmatrix} 2 & a & b \\ 0 & 2 & c \\ 0 & 0 & 3 \end{pmatrix}$. The characteristic polynomial is $p(\lambda) = -(\lambda - 3)(\lambda - 2)^2$. Hence the diagonalizability of A depends entirely on what happens with $\lambda = 2$. Why? Since $A - 2I_3 = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 1 \end{pmatrix}$ has rank equal to 1 only if $a = 0$, it follows that A is diagonalizable if and only if $a = 0$.

Matrix Powers

Example

If $A \in M_{n\times n}(\mathbf{F})$, then it can be useful to compute matrix powers A^k without the hassle of repeated matrix multiplication. But this is routine if A is diagonalizable. Then $A = QDQ^{-1}$ for an invertible matrix Q and a diagonal matrix D. Computing D^k is easy! But

$$A^k = (QDQ^{-1})(QDQ^{-1})\cdots(QDQ^{-1}) = QD^kQ^{-1}.$$

For example, let $A = \left(\begin{smallmatrix} -3 & 6 \\ -4 & 7 \end{smallmatrix}\right)$. I leave it to you go check that A is diagonalizable with $D = Q^{-1}AQ$ for $D = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 3 \end{smallmatrix}\right)$ and $Q = \left(\begin{smallmatrix} 3 & 1 \\ 2 & 1 \end{smallmatrix}\right)$. Then

$$\begin{aligned} A^k = QD^kQ^{-1} &= \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} -3 & 6 \\ -4 & 7 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 3 - 2\cdot 3^k & 3^{k+1} - 3 \\ 2 - 2\cdot 3^k & 3^{k+1} - 2 \end{pmatrix}. \end{aligned}$$

Example

Example

Let $A = \begin{pmatrix} -10 & 18 \\ -9 & 17 \end{pmatrix}$. The goal here is to find a matrix B such that $B^3 = A$. We start by diagonalizing A. Its characteristic polynomial is $p(\lambda) = (-10 - \lambda)(17 - \lambda) + 9 \cdot 18 = \lambda^2 - 7\lambda - 170 + 162 = \lambda^2 - 7\lambda - 8 = (\lambda + 1)(\lambda - 8)$. Then $v_1 = (2, 1)$ is an eigenvector with eigenvalue -1 and $v_2 = (1, 1)$ is an eigenvector with eigenvalue 8. Thus $\begin{pmatrix} -1 & 0 \\ 0 & 8 \end{pmatrix} = Q^{-1}AQ$ where $Q = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Clearly, $\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}^3 = \begin{pmatrix} -1 & 0 \\ 0 & 8 \end{pmatrix}$. Since $A = Q\begin{pmatrix} -1 & 0 \\ 0 & 8 \end{pmatrix}Q^{-1}$. We can let $B = Q\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}Q^{-1}$. That is,

$$B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -4 & 6 \\ -3 & 5 \end{pmatrix}.$$

Enough

1. That is enough for today.