Eigenvectors and Upper-Triangular Matrices

Lecture 14

Dept. of Math., SUSTech

2024.4

Eigenvalues, Eigenvectors, and Invariant Subspaces

Polynomials Applied to Operators

The main reason that a richer theory exists for operators (which map a vector space into itself) than for more general linear maps is that operators can be raised to powers. We begin this section by defining that notion and the key concept of applying a polynomial to an operator.

5.16 Definition T^m

Suppose $T \in \mathcal{L}(V)$ and m is a positive integer.

- T^m is defined by
$$T^m = \underbrace{T \cdots T}_{m \text{ times}}.$$
- T^0 is defined to be the identity operator I on V.
- If T is invertible with inverse T^{-1}, then T^{-m} is defined by
$$T^{-m} = (T^{-1})^m.$$

Polynomials applied to operators

5.17 Definition $p(T)$

Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbf{F})$ is a polynomial given by

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

for $z \in \mathbf{F}$. Then $p(T)$ is the operator defined by

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m.$$

Example.

5.18 Example Suppose $D \in \mathcal{L}\big(\mathcal{P}(\mathbf{R})\big)$ is the differentiation operator defined by $Dq = q'$ and p is the polynomial defined by $p(x) = 7 - 3x + 5x^2$. Then $p(D) = 7I - 3D + 5D^2$; thus

$$\big(p(D)\big)q = 7q - 3q' + 5q''$$

for every $q \in \mathcal{P}(\mathbf{R})$.

Multiplicative properties

If we fix an operator $T \in \mathscr{L}(V)$, then the function from $\mathscr{P}(\mathbb{F})$ to $\mathscr{L}(V)$ given by $p \mapsto p(T)$ is linear, as you should verify:

5.19 **Definition** ***product of polynomials***

If $p, q \in \mathcal{P}(\mathbf{F})$, then $pq \in \mathcal{P}(\mathbf{F})$ is the polynomial defined by

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

for $z \in \mathbf{F}$.

5.20

Any two polynomials of an operator commute, as shown below.

5.20 Multiplicative properties

Suppose $p, q \in \mathcal{P}(\mathbf{F})$ and $T \in \mathcal{L}(V)$. Then

(a) $(pq)(T) = p(T)q(T)$;

(b) $p(T)q(T) = q(T)p(T)$.

Part (a) holds because when expanding a product of polynomials using the distributive property, it does not matter whether the symbol is z or T.

Proof

Proof. (a) Suppose $p(z) = \sum_{j=0}^{m} a_j z^j$ and $q(z) = \sum_{k=0}^{n} b_k z^k$ for $z \in \mathbb{F}$. Then

$$(pq)(z) = \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k z^{j+k}.$$

Thus

$$(pq)(T) = \sum_{j=0}^{m} \sum_{k=0}^{n} a_j b_k T^{j+k} = \left(\sum_{j=0}^{m} a_j T^j \right) \left(\sum_{k=0}^{n} b_k T^k \right) = p(T)q(T).$$

(b) Part (a) implies

$$p(T)q(T) = (pq)(T) = (qp)(T) = q(T)p(T).$$

Eigenvalues and Eigenvectors

Now we come to one of the central results about operators on complex vector spaces.

5.21 Operators on complex vector spaces have an eigenvalue

Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.

Proof. Suppose V is a complex vector space with dimension $n > 0$ and $T \in \mathscr{L}(V)$. Choose $v \in V$ with $v \neq 0$. Then

$$v, Tv, T^2v, \cdots, T^nv$$

is not linearly independent, because V has dimension n and we have $n+1$ vectors. Thus there exist complex numbers $a_0, a_1, \cdots, a_n$ not all 0, such that

$$0 = a_0v + a_1Tv + \cdots + a_nT^nv.$$

Proof.

Note that $a_1, a_2, \cdots, a_n$ cannot all be 0, because otherwise the equation above would become $0 = a_0 v$, which would force a_0 also to be 0.

Make the a's the coefficients of a polynomial, which by the Fundamental Theorem of Algebra (4.14) has a factorization

$$a_0 + a_1 z + \cdots + a_n z^n = c(z - \lambda_1) \cdots (z - \lambda_m),$$

where c is a nonzero complex number, each λ_j is in $\mathbb{C}$, and the equation holds for all $z \in \mathbb{C}$ (here m is not necessarily equal to n, because a_n may equal 0). We then have

$$\begin{aligned} 0 &= a_0 v + a_1 T v + \cdots + a_n T^n v = (a_0 I + a_1 T + \cdots + a_n T^n) v \\ &= c(T - \lambda_1 I) \cdots (T - \lambda_m I) v. \end{aligned}$$

proof

Thus $T - \lambda_j I$ is not injective for at least one j. In other words, T has an eigenvalue.

Comment: See Exercises 16 and 17 for possible ways to rewrite the proof given in the textbook using the idea of the proof in a slightly different form.

Upper-Triangular Matrices

Now that we are studying operators, which map a vector space to itself, the emphasis is on using only one basis.

5.22 **Definition** ***matrix of an operator,*** $\mathcal{M}(T)$

Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. The ***matrix of*** T with respect to this basis is the n-by-n matrix

$$\mathcal{M}(T) = \begin{pmatrix} A_{1,1} & \dots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \dots & A_{n,n} \end{pmatrix}$$

whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}v_1 + \dots + A_{n,k}v_n.$$

If the basis is not clear from the context, then the notation $\mathcal{M}\big(T, (v_1, \dots, v_n)\big)$ is used.

Diagonal of a Matrix

A central goal of linear algebra is to show that given an operator $T \in \mathscr{L}(V)$, there exists a basis of V with respect to which T has a reasonably simple matrix. To make this vague formulation a bit more precise, we might try to choose a basis of V such that $\mathscr{M}(T)$ has many 0's. Soon we will see that we can choose a basis of V with respect to which the matrix of T has even more 0's.

5.24 Definition *diagonal of a matrix*

The ***diagonal*** of a square matrix consists of the entries along the line from the upper left corner to the bottom right corner.

Upper Triangular Matrix

upper-triangular matrix

5.25 **Definition** ***upper-triangular matrix***

A matrix is called ***upper triangular*** if all the entries below the diagonal equal 0.

Conditions for upper-triangular matrix

The following proposition demonstrates a useful connection between upper-triangular matrices and invariant subspaces.

5.26 Conditions for upper-triangular matrix

Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Then the following are equivalent:

(a) the matrix of T with respect to $v_1, \ldots, v_n$ is upper triangular;

(b) $Tv_j \in \operatorname{span}(v_1, \ldots, v_j)$ for each $j = 1, \ldots, n$;

(c) $\operatorname{span}(v_1, \ldots, v_j)$ is invariant under T for each $j = 1, \ldots, n$.

Proof.

Only need to prove (b) implies (c). □

Over $\mathbb{C}$, every operator has an upper-triangular matrix

The next result does not hold on real vector spaces, because the first vector in a basis with respect to which an operator has an upper-triangular matrix is an eigenvector of the operator. Thus if an operator on a real vector space has no eigenvalues, then there is no basis with respect to which the operator has an upper-triangular matrix.

5.27 Over $\mathbf{C}$, every operator has an upper-triangular matrix

Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Then T has an upper-triangular matrix with respect to some basis of V.

Proof.

Proof. We will use induction on the dimension of V. Clearly the desired result holds if $\dim\ V = 1$.

Induction Hypothesis:

Suppose now that $\dim\ V > 1$ and the desired result holds for all complex vector spaces whose dimension is less than the dimension of V.

Let λ be any eigenvalue of T (5.21 guarantees that T has an eigenvalue). Let

$$U = \text{range } (T - \lambda I).$$

Because $T - \lambda I$ is not surjective (see 3.69), $\dim U < \dim V$.

Proof

Furthermore, U is invariant under T. To prove this, suppose $u \in U$. Then

$$Tu = (T - \lambda I)u + \lambda u.$$

Obviously $(T - \lambda I)u \in U$ (because U equals the range of $T - \lambda I$) and $\lambda u \in U$. Thus the equation above shows that $Tu \in U$. Hence U is invariant under T, as claimed.

Thus $T|_U$ is an operator on U. By our **induction hypothesis**, there is a basis $u_1, u_2, \cdots, u_m$ of U with respect to which $T|_U$ has an upper triangular matrix. Thus for each j we have (using 5.26)

$$T(u_j) = T|_U(u_j) \in \text{span } (u_1, \cdots, u_j).$$

Proof

Extend $u_1, \cdots, u_m$ to a basis $u_1, \cdots, u_m, v_1, \cdots, v_n$ of V. For each k, we have

$$Tv_k = (T - \lambda I)v_k + \lambda v_k.$$

The definition of U shows that $(T - \lambda I)v_k \in U = \text{span }(u_1, \cdots, u_m)$. Thus the equation above shows that

$$Tv_k \in \text{span }(u_1, \cdots, u_m, v_1, \cdots, v_k).$$

Therefore, we conclude (using 5.26) that T has an upper triangular matrix with respect to the basis $u_1, \cdots, u_m, v_1, \cdots, v_n$ of V, as desired.

Determination of invertibility from upper-triangular matrix

How does one determine from looking at the matrix of an operator whether the operator is invertible? If we are fortunate enough to have a basis with respect to which the matrix of the operator is upper triangular, then this problem becomes easy, as the following proposition shows.

5.30 Determination of invertibility from upper-triangular matrix

Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then T is invertible if and only if all the entries on the diagonal of that upper-triangular matrix are nonzero.

Proof

Proof. Suppose $v_1, v_2, \cdots, v_n$ is a basis of V with respect to which T has an upper-triangular matrix

$$\mathscr{M}(T) = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \cdots\cdots(*).$$

We need to prove that T is invertible if and only if all the λ_j's are nonzero.

First suppose the diagonal entries $\lambda_1, \lambda_2, \cdots, \lambda_n$ are all nonzero.

Proof

The upper triangular matrix in $(*)$ implies that $Tv_1 = \lambda_1 v_1$. Because $\lambda_1 \neq 0$, we have $T(v_1/\lambda_1) = v_1$; thus $v_1 \in$ range T.

Now

$$T(v_2/\lambda_2) = av_1 + v_2$$

for some $a \in \mathbb{F}$.

The left side of the equation above and av_1 are both in range T; thus $v_2 \in$ range T. Similarly, we see that

$$T(v_3/\lambda_3) = bv_1 + cv_2 + v_3$$

for some $b,\ c \in \mathbb{F}$.

Proof

The left side of the equation above and bv_1, cv_2 are all in range T; thus

$$v_3 \in \text{range } T.$$

Continuing in this fashion, we conclude that $v_1, v_2, \cdots, v_n \in$ range T. Because $v_1, v_2, \cdots, v_n$ is a basis of V, this implies that range $T = V$. In other words, T is surjective. Hence T is invertible (by 3.69), as desired.

To prove the other direction, now suppose that T is invertible. This implies that $\lambda_1 \neq 0$, because otherwise we would have $Tv_1 = 0$.

Proof.

Let $1 < j \leq n$, and suppose $\lambda_j = 0$. Then $(*)$ implies that T maps span$(v_1, \cdots, v_j)$ into span$(v_1, \cdots, v_{j-1})$. Because

$$\dim \text{ span}(v_1, \cdots, v_j) = j \quad \text{and} \quad \dim \text{ span}(v_1, \cdots, v_{j-1}) = j - 1,$$

This implies that T restricted to span$(v_1, \cdots, v_j)$ is not injective (by 3.23). Thus there exists $v \in$ span$(v_1, \cdots, v_j)$ such that $v \neq 0$ and $Tv = 0$. Thus T is not injective, which contradicts our hypothesis (for this direction) that T is invertible. This contradiction means that our assumption that $\lambda_j = 0$ must be false. Hence $\lambda_j \neq 0$, as desired.

Determination of eigenvalues from upper-triangular matrix

Unfortunately no method exists for exactly computing the eigenvalues of an operator from its matrix. However, if we are fortunate enough to find a basis with respect to which the matrix of the operator is upper triangular, then the problem of computing the eigenvalues becomes trivial, as the following proposition shows.

5.32 Determination of eigenvalues from upper-triangular matrix

Suppose $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V. Then the eigenvalues of T are precisely the entries on the diagonal of that upper-triangular matrix.

Powerful numeric techniques exist for finding good approximations to the eigenvalues of an operator from its matrix.

Homework Assignment 14

5.B: 4, 9, 11, 14, 17, 20.