

5.2 DIAGONALIZABILITY

In Section 5.1, we presented the diagonalization problem and observed that not all linear operators or matrices are diagonalizable. Although we are able to diagonalize operators and matrices and even obtain a necessary and sufficient condition for diagonalizability (Theorem 5.1 p. 246), we have not yet solved the diagonalization problem. What is still needed is a simple test to determine whether an operator or a matrix can be diagonalized, as well as a method for actually finding a basis of eigenvectors. In this section, we develop such a test and method.

In Example 6 of Section 5.1, we obtained a basis of eigenvectors by choosing one eigenvector corresponding to each eigenvalue. In general, such a procedure does not yield a basis, but the following theorem shows that any set constructed in this manner is linearly independent.

Theorem 5.5. *Let T be a linear operator on a vector space V , and let $\lambda_1, \lambda_2, \dots, \lambda_k$ be distinct eigenvalues of T . If $v_1, v_2, \dots, v_k$ are eigenvectors of T such that λ_i corresponds to v_i ($1 \leq i \leq k$), then $\{v_1, v_2, \dots, v_k\}$ is linearly independent.*

Proof. The proof is by mathematical induction on k . Suppose that $k = 1$. Then $v_1 \neq 0$ since v_1 is an eigenvector, and hence $\{v_1\}$ is linearly independent. Now assume that the theorem holds for $k - 1$ distinct eigenvalues, where $k - 1 \geq 1$, and that we have k eigenvectors $v_1, v_2, \dots, v_k$ corresponding to the distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$. We wish to show that $\{v_1, v_2, \dots, v_k\}$ is linearly independent. Suppose that $a_1, a_2, \dots, a_k$ are scalars such that

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = 0. \quad (1)$$

Applying $T - \lambda_k I$ to both sides of (1), we obtain

$$a_1(\lambda_1 - \lambda_k)v_1 + a_2(\lambda_2 - \lambda_k)v_2 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

By the induction hypothesis $\{v_1, v_2, \dots, v_{k-1}\}$ is linearly independent, and hence

$$a_1(\lambda_1 - \lambda_k) = a_2(\lambda_2 - \lambda_k) = \cdots = a_{k-1}(\lambda_{k-1} - \lambda_k) = 0.$$

Since $\lambda_1, \lambda_2, \dots, \lambda_k$ are distinct, it follows that $\lambda_i - \lambda_k \neq 0$ for $1 \leq i \leq k - 1$. So $a_1 = a_2 = \cdots = a_{k-1} = 0$, and (1) therefore reduces to $a_k v_k = 0$. But $v_k \neq 0$ and therefore $a_k = 0$. Consequently $a_1 = a_2 = \cdots = a_k = 0$, and it follows that $\{v_1, v_2, \dots, v_k\}$ is linearly independent. ■

Corollary. *Let T be a linear operator on an n -dimensional vector space V . If T has n distinct eigenvalues, then T is diagonalizable.*

Proof. Suppose that T has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$. For each i choose an eigenvector v_i corresponding to λ_i . By Theorem 5.5, $\{v_1, \dots, v_n\}$ is linearly independent, and since $\dim(V) = n$, this set is a basis for V . Thus, by Theorem 5.1 (p. 246), T is diagonalizable. ■

Example 1

Let

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in M_{2 \times 2}(R).$$

The characteristic polynomial of A (and hence of L_A) is

$$\det(A - tI) = \det \begin{pmatrix} 1-t & 1 \\ 1 & 1-t \end{pmatrix} = t(t-2),$$

and thus the eigenvalues of L_A are 0 and 2. Since L_A is a linear operator on the two-dimensional vector space R^2 , we conclude from the preceding corollary that L_A (and hence A) is diagonalizable. ◆

The converse of Theorem 5.5 is false. That is, it is not true that if T is diagonalizable, then it has n distinct eigenvalues. For example, the identity operator is diagonalizable even though it has only one eigenvalue, namely, $\lambda = 1$.

We have seen that diagonalizability requires the existence of eigenvalues. Actually, diagonalizability imposes a stronger condition on the characteristic polynomial.

Definition. A polynomial $f(t)$ in $P(F)$ **splits over** F if there are scalars $c, a_1, \dots, a_n$ (not necessarily distinct) in F such that

$$f(t) = c(t - a_1)(t - a_2) \cdots (t - a_n).$$

For example, $t^2 - 1 = (t + 1)(t - 1)$ splits over R , but $(t^2 + 1)(t - 2)$ does not split over R because $t^2 + 1$ cannot be factored into a product of linear factors. However, $(t^2 + 1)(t - 2)$ does split over C because it factors into the product $(t + i)(t - i)(t - 2)$. If $f(t)$ is the characteristic polynomial of a linear operator or a matrix over a field F , then the statement that $f(t)$ splits is understood to mean that it splits over F .

Theorem 5.6. *The characteristic polynomial of any diagonalizable linear operator splits.*

Proof. Let T be a diagonalizable linear operator on the n -dimensional vector space V , and let β be an ordered basis for V such that $[T]_\beta = D$ is a

diagonal matrix. Suppose that

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

and let $f(t)$ be the characteristic polynomial of T . Then

$$\begin{aligned} f(t) &= \det(D - tI) = \det \begin{pmatrix} \lambda_1 - t & 0 & \cdots & 0 \\ 0 & \lambda_2 - t & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n - t \end{pmatrix} \\ &= (\lambda_1 - t)(\lambda_2 - t) \cdots (\lambda_n - t) = (-1)^n (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n). \quad \blacksquare \end{aligned}$$

From this theorem, it is clear that if T is a diagonalizable linear operator on an n -dimensional vector space that fails to have distinct eigenvalues, then the characteristic polynomial of T must have repeated zeros.

The converse of Theorem 5.6 is false; that is, the characteristic polynomial of T may split, but T need not be diagonalizable. (See Example 3, which follows.) The following concept helps us determine when an operator whose characteristic polynomial splits is diagonalizable.

Definition. Let λ be an eigenvalue of a linear operator or matrix with characteristic polynomial $f(t)$. The **(algebraic) multiplicity** of λ is the largest positive integer k for which $(t - \lambda)^k$ is a factor of $f(t)$.

Example 2

Let

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{pmatrix},$$

which has characteristic polynomial $f(t) = -(t - 3)^2(t - 4)$. Hence $\lambda = 3$ is an eigenvalue of A with multiplicity 2, and $\lambda = 4$ is an eigenvalue of A with multiplicity 1. $\blacklozenge$

If T is a diagonalizable linear operator on a finite-dimensional vector space V , then there is an ordered basis β for V consisting of eigenvectors of T . We know from Theorem 5.1 (p. 246) that $[T]_\beta$ is a diagonal matrix in which the diagonal entries are the eigenvalues of T . Since the characteristic polynomial of T is $\det([T]_\beta - tI)$, it is easily seen that each eigenvalue of T must occur as a diagonal entry of $[T]_\beta$ exactly as many times as its multiplicity. Hence

β contains as many (linearly independent) eigenvectors corresponding to an eigenvalue as the multiplicity of that eigenvalue. So the number of linearly independent eigenvectors corresponding to a given eigenvalue is of interest in determining whether an operator can be diagonalized. Recalling from Theorem 5.4 (p. 250) that the eigenvectors of T corresponding to the eigenvalue λ are the nonzero vectors in the null space of $T - \lambda I$, we are led naturally to the study of this set.

Definition. Let T be a linear operator on a vector space V , and let λ be an eigenvalue of T . Define $E_\lambda = \{x \in V : T(x) = \lambda x\} = N(T - \lambda I_V)$. The set E_λ is called the **eigenspace** of T corresponding to the eigenvalue λ . Analogously, we define the **eigenspace** of a square matrix A to be the eigenspace of L_A .

Clearly, E_λ is a subspace of V consisting of the zero vector and the eigenvectors of T corresponding to the eigenvalue λ . The maximum number of linearly independent eigenvectors of T corresponding to the eigenvalue λ is therefore the dimension of E_λ . Our next result relates this dimension to the multiplicity of λ .

Theorem 5.7. Let T be a linear operator on a finite-dimensional vector space V , and let λ be an eigenvalue of T having multiplicity m . Then $1 \leq \dim(E_\lambda) \leq m$.

Proof. Choose an ordered basis $\{v_1, v_2, \dots, v_p\}$ for E_λ , extend it to an ordered basis $\beta = \{v_1, v_2, \dots, v_p, v_{p+1}, \dots, v_n\}$ for V , and let $A = [T]_\beta$. Observe that v_i ($1 \leq i \leq p$) is an eigenvector of T corresponding to λ , and therefore

$$A = \begin{pmatrix} \lambda I_p & B \\ O & C \end{pmatrix}.$$

By Exercise 21 of Section 4.3, the characteristic polynomial of T is

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

where $g(t)$ is a polynomial. Thus $(\lambda - t)^p$ is a factor of $f(t)$, and hence the multiplicity of λ is at least p . But $\dim(E_\lambda) = p$, and so $\dim(E_\lambda) \leq m$. ■

Example 3

Let T be the linear operator on $P_2(R)$ defined by $T(f(x)) = f'(x)$. The matrix representation of T with respect to the standard ordered basis β for

$P_2(R)$ is

$$[T]_\beta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Consequently, the characteristic polynomial of T is

$$\det([T]_\beta - tI) = \det \begin{pmatrix} -t & 1 & 0 \\ 0 & -t & 2 \\ 0 & 0 & -t \end{pmatrix} = -t^3.$$

Thus T has only one eigenvalue ($\lambda = 0$) with multiplicity 3. Solving $T(f(x)) = f'(x) = 0$ shows that $E_\lambda = N(T - \lambda I) = N(T)$ is the subspace of $P_2(R)$ consisting of the constant polynomials. So $\{1\}$ is a basis for E_λ , and therefore $\dim(E_\lambda) = 1$. Consequently, there is no basis for $P_2(R)$ consisting of eigenvectors of T , and therefore T is not diagonalizable. $\blacklozenge$

Example 4

Let T be the linear operator on $\mathbb{R}^3$ defined by

$$T \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 4a_1 & + & a_3 \\ 2a_1 + 3a_2 + 2a_3 \\ a_1 & + & 4a_3 \end{pmatrix}.$$

We determine the eigenspace of T corresponding to each eigenvalue. Let β be the standard ordered basis for $\mathbb{R}^3$. Then

$$[T]_\beta = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix},$$

and hence the characteristic polynomial of T is

$$\det([T]_\beta - tI) = \det \begin{pmatrix} 4-t & 0 & 1 \\ 2 & 3-t & 2 \\ 1 & 0 & 4-t \end{pmatrix} = -(t-5)(t-3)^2.$$

So the eigenvalues of T are $\lambda_1 = 5$ and $\lambda_2 = 3$ with multiplicities 1 and 2, respectively.

Since

$$E_{\lambda_1} = N(T - \lambda_1 I) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\},$$

E_{λ_1} is the solution space of the system of linear equations

$$\begin{array}{rcl} -x_1 & + & x_3 = 0 \\ 2x_1 - 2x_2 + 2x_3 & = & 0 \\ x_1 & - & x_3 = 0. \end{array}$$

It is easily seen (using the techniques of Chapter 3) that

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$$

is a basis for E_{λ_1} . Hence $\dim(E_{\lambda_1}) = 1$.

Similarly, $E_{\lambda_2} = N(T - \lambda_2 I)$ is the solution space of the system

$$\begin{array}{rcl} x_1 + x_3 & = & 0 \\ 2x_1 + 2x_3 & = & 0 \\ x_1 + x_3 & = & 0. \end{array}$$

Since the unknown x_2 does not appear in this system, we assign it a parametric value, say, $x_2 = s$, and solve the system for x_1 and x_3 , introducing another parameter t . The result is the general solution to the system

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = s \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \text{ for } s, t \in R.$$

It follows that

$$\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

is a basis for E_{λ_2} , and $\dim(E_{\lambda_2}) = 2$.

In this case, the multiplicity of each eigenvalue λ_i is equal to the dimension of the corresponding eigenspace E_{λ_i} . Observe that the union of the two bases just derived, namely,

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\},$$

is linearly independent and hence is a basis for R^3 consisting of eigenvectors of T . Consequently, T is diagonalizable. $\blacklozenge$