teristic values $c_1, \ldots, c_k$. Let W_i be the space of characteristic vectors associated with the characteristic value c_i. As we have seen,

$$V = W_1 \oplus \cdots \oplus W_k.$$

Let $E_1, \ldots, E_k$ be the projections associated with this decomposition, as in Theorem 9. Then (ii), (iii), (iv) and (v) are satisfied. To verify (i), proceed as follows. For each α in V,

$$\alpha = E_1\alpha + \cdots + E_k\alpha$$

and so

$$\begin{aligned} T\alpha &= TE_1\alpha + \cdots + TE_k\alpha \\ &= c_1E_1\alpha + \cdots + c_kE_k\alpha. \end{aligned}$$

In other words, $T = c_1E_1 + \cdots + c_kE_k$.

Now suppose that we are given a linear operator T along with distinct scalars c_i and non-zero operators E_i which satisfy (i), (ii) and (iii). Since $E_iE_j = 0$ when $i \neq j$, we multiply both sides of $I = E_1 + \cdots + E_k$ by E_i and obtain immediately $E_i^2 = E_i$. Multiplying $T = c_1E_1 + \cdots + c_kE_k$ by E_i, we then have $TE_i = c_iE_i$, which shows that any vector in the range of E_i is in the null space of $(T - c_iI)$. Since we have assumed that $E_i \neq 0$, this proves that there is a non-zero vector in the null space of $(T - c_iI)$, i.e., that c_i is a characteristic value of T. Furthermore, the c_i are all of the characteristic values of T; for, if c is any scalar, then

$$T - cI = (c_1 - c)E_1 + \cdots + (c_k - c)E_k$$

so if $(T - cI)\alpha = 0$, we must have $(c_i - c)E_i\alpha = 0$. If α is not the zero vector, then $E_i\alpha \neq 0$ for some i, so that for this i we have $c_i - c = 0$.

Certainly T is diagonalizable, since we have shown that every non-zero vector in the range of E_i is a characteristic vector of T, and the fact that $I = E_1 + \cdots + E_k$ shows that these characteristic vectors span V. All that remains to be demonstrated is that the null space of $(T - c_iI)$ is exactly the range of E_i. But this is clear, because if $T\alpha = c_i\alpha$, then

$$\sum_{j=1}^{k} (c_j - c_i)E_j\alpha = 0$$

hence

$$(c_j - c_i)E_j\alpha = 0 \qquad \text{for each } j$$

and then

$$E_j\alpha = 0, \qquad j \neq i.$$

Since $\alpha = E_1\alpha + \cdots + E_k\alpha$, and $E_j\alpha = 0$ for $j \neq i$, we have $\alpha = E_i\alpha$, which proves that α is in the range of E_i. ∎

One part of Theorem 9 says that for a diagonalizable operator T, the scalars $c_1, \ldots, c_k$ and the operators $E_1, \ldots, E_k$ are uniquely determined by conditions (i), (ii), (iii), the fact that the c_i are distinct, and the fact that the E_i are non-zero. One of the pleasant features of the decomposition $T = c_1E_1 + \cdots + c_kE_k$ is that if g is any polynomial over the field F, then

$$g(T) = g(c_1)E_1 + \cdots + g(c_k)E_k.$$

We leave the details of the proof to the reader. To see how it is proved one need only compute T^r for each positive integer r. For example,

$$\begin{aligned} T^2 &= \sum_{i=1}^k c_iE_i \sum_{j=1}^k c_jE_j \\ &= \sum_{i=1}^k \sum_{j=1}^k c_ic_jE_iE_j \\ &= \sum_{i=1}^k c_i^2E_i^2 \\ &= \sum_{i=1}^k c_i^2E_i. \end{aligned}$$

The reader should compare this with $g(A)$ where A is a diagonal matrix; for then $g(A)$ is simply the diagonal matrix with diagonal entries $g(A_{11})$, $\ldots$, $g(A_{nn})$.

We should like in particular to note what happens when one applies the Lagrange polynomials corresponding to the scalars $c_1, \ldots, c_k$:

$$p_j = \prod_{i \neq j} \frac{(x - c_i)}{(c_j - c_i)}.$$

We have $p_j(c_i) = \delta_{ij}$, which means that

$$\begin{aligned} p_j(T) &= \sum_{i=1}^k \delta_{ij}E_i \\ &= E_j. \end{aligned}$$

Thus the projections E_j not only commute with T but are polynomials in T.

Such calculations with polynomials in T can be used to give an alternative proof of Theorem 6, which characterized diagonalizable operators in terms of their minimal polynomials. The proof is entirely independent of our earlier proof.

If T is diagonalizable, $T = c_1E_1 + \cdots + c_kE_k$, then

$$g(T) = g(c_1)E_1 + \cdots + g(c_k)E_k$$

for every polynomial g. Thus $g(T) = 0$ if and only if $g(c_i) = 0$ for each i. In particular, the minimal polynomial for T is

$$p = (x - c_1) \cdots (x - c_k).$$

Now suppose T is a linear operator with minimal polynomial $p = (x - c_1) \cdots (x - c_k)$, where $c_1, \ldots, c_k$ are distinct elements of the scalar field. We form the Lagrange polynomials

$$p_j = \prod_{i \neq j} \frac{(x - c_i)}{(c_j - c_i)}.$$

We recall from Chapter 4 that $p_j(c_i) = \delta_{ij}$ and for any polynomial g of degree less than or equal to $(k - 1)$ we have

$$g = g(c_1)p_1 + \cdots + g(c_k)p_k.$$

Taking g to be the scalar polynomial 1 and then the polynomial x, we have

$$\begin{aligned} 1 &= p_1 + \cdots + p_k \\ x &= c_1p_1 + \cdots + c_kp_k. \end{aligned} \tag{6-15}$$

(The astute reader will note that the application to x may not be valid because k may be 1. But if $k = 1$, T is a scalar multiple of the identity and hence diagonalizable.) Now let $E_j = p_j(T)$. From (6-15) we have

$$\begin{aligned} I &= E_1 + \cdots + E_k \\ T &= c_1E_1 + \cdots + c_kE_k. \end{aligned} \tag{6-16}$$

Observe that if $i \neq j$, then p_ip_j is divisible by the minimal polynomial p, because p_ip_j contains every $(x - c_r)$ as a factor. Thus

$$E_iE_j = 0, \qquad i \neq j. \tag{6-17}$$

We must note one further thing, namely, that $E_i \neq 0$ for each i. This is because p is the minimal polynomial for T and so we cannot have $p_i(T) = 0$ since p_i has degree less than the degree of p. This last comment, together with (6-16), (6-17), and the fact that the c_i are distinct enables us to apply Theorem 11 to conclude that T is diagonalizable. ▌

Exercises

1. Let E be a projection of V and let T be a linear operator on V. Prove that the range of E is invariant under T if and only if $ETE = TE$. Prove that both the range and null space of E are invariant under T if and only if $ET = TE$.

2. Let T be the linear operator on R^2, the matrix of which in the standard ordered basis is

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

Let W_1 be the subspace of R^2 spanned by the vector $\epsilon_1 = (1, 0)$.

(a) Prove that W_1 is invariant under T.

(b) Prove that there is no subspace W_2 which is invariant under T and which is complementary to W_1:

$$R^2 = W_1 \oplus W_2.$$

(Compare with Exercise 1 of Section 6.5.)

3. Let T be a linear operator on a finite-dimensional vector space V. Let R be the range of T and let N be the null space of T. Prove that R and N are independent if and only if $V = R \oplus N$.

4. Let T be a linear operator on V. Suppose $V = W_1 \oplus \cdots \oplus W_k$, where each W_i is invariant under T. Let T_i be the induced (restriction) operator on W_i.

(a) Prove that $\det(T) = \det(T_1) \cdots \det(T_k)$.

(b) Prove that the characteristic polynomial for f is the product of the characteristic polynomials for $f_1, \ldots, f_k$.

(c) Prove that the minimal polynomial for T is the least common multiple of the minimal polynomials for $T_1, \ldots, T_k$. (*Hint:* Prove and then use the corresponding facts about direct sums of matrices.)

5. Let T be the diagonalizable linear operator on R^3 which we discussed in Example 3 of Section 6.2. Use the Lagrange polynomials to write the representing matrix A in the form $A = E_1 + 2E_2$, $E_1 + E_2 = I$, $E_1E_2 = 0$.

6. Let A be the 4×4 matrix in Example 6 of Section 6.3. Find matrices E_1, E_2, E_3 such that $A = c_1E_1 + c_2E_2 + c_3E_3$, $E_1 + E_2 + E_3 = I$, and $E_iE_j = 0$, $i \neq j$.

7. In Exercises 5 and 6, notice that (for each i) the space of characteristic vectors associated with the characteristic value c_i is spanned by the column vectors of the various matrices E_j with $j \neq i$. Is that a coincidence?

8. Let T be a linear operator on V which commutes with every projection operator on V. What can you say about T?

9. Let V be the vector space of continuous real-valued functions on the interval $[-1, 1]$ of the real line. Let W_e be the subspace of even functions, $f(-x) = f(x)$, and let W_o be the subspace of odd functions, $f(-x) = -f(x)$.

(a) Show that $V = W_e \oplus W_o$.

(b) If T is the indefinite integral operator

$$(Tf)(x) = \int_0^x f(t)\,dt$$

are W_e and W_o invariant under T?

6.8. The Primary Decomposition Theorem

We are trying to study a linear operator T on the finite-dimensional space V, by decomposing T into a direct sum of operators which are in some sense elementary. We can do this through the characteristic values and vectors of T in certain special cases, i.e., when the minimal polynomial for T factors over the scalar field F into a product of distinct monic polynomials of degree 1. What can we do with the general T? If we try to study T using characteristic values, we are confronted with two problems. First, T may not have a single characteristic value; this is really a deficiency in the scalar field, namely, that it is not algebraically closed. Second, even if the characteristic polynomial factors completely over F into a product of polynomials of degree 1, there may not be enough characteristic vectors for T to span the space V; this is clearly a deficiency in T. The second situation is illustrated by the operator T on F^3 (F any field) represented in the standard basis by

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

The characteristic polynomial for A is $(x - 2)^2(x + 1)$ and this is plainly also the minimal polynomial for A (or for T). Thus T is not diagonalizable. One sees that this happens because the null space of $(T - 2I)$ has dimension 1 only. On the other hand, the null space of $(T + I)$ and the null space of $(T - 2I)^2$ together span V, the former being the subspace spanned by ϵ_3 and the latter the subspace spanned by ϵ_1 and ϵ_2.

This will be more or less our general method for the second problem. If (remember this is an assumption) the minimal polynomial for T decomposes

$$p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}$$

where $c_1, \ldots, c_k$ are distinct elements of F, then we shall show that the space V is the direct sum of the null spaces of $(T - c_iI)^{r_i}$, $i = 1, \ldots, k$. The hypothesis about p is equivalent to the fact that T is triangulable (Theorem 5); however, that knowledge will not help us.

The theorem which we prove is more general than what we have described, since it works with the primary decomposition of the minimal polynomial, whether or not the primes which enter are all of first degree. The reader will find it helpful to think of the special case when the primes are of degree 1, and even more particularly, to think of the projection-type proof of Theorem 6, a special case of this theorem.

Theorem 12 (Primary Decomposition Theorem). *Let* T *be a linear operator on the finite-dimensional vector space* V *over the field* F. *Let* p *be the minimal polynomial for* T,

$$p = p_1^{r_1} \cdots p_k^{r_k}$$

where the p_i *are distinct irreducible monic polynomials over* F *and the* r_i *are positive integers. Let* W_i *be the null space of* $p_i(T)^{r_i}$, $i = 1, \ldots, k$. *Then*

(i) $V = W_1 \oplus \cdots \oplus W_k$;
(ii) *each* W_i *is invariant under* T;
(iii) *if* T_i *is the operator induced on* W_i *by* T, *then the minimal polynomial for* T_i *is* $p_i^{r_i}$.

Proof. The idea of the proof is this. If the direct-sum decomposition (i) is valid, how can we get hold of the projections $E_1, \ldots, E_k$ associated with the decomposition? The projection E_i will be the identity on W_i and zero on the other W_j. We shall find a polynomial h_i such that $h_i(T)$ is the identity on W_i and is zero on the other W_j, and so that $h_1(T) + \cdots + h_k(T) = I$, etc.