an illuminating and far from trivial application of the general theory of determinants developed in Chapter 5.

Let K be the commutative ring with identity consisting of all polynomials in T. Of course, K is actually a commutative algebra with identity over the scalar field. Choose an ordered basis $\{\alpha_1, \ldots, \alpha_n\}$ for V, and let A be the matrix which represents T in the given basis. Then

$$T\alpha_i = \sum_{j=1}^{n} A_{ji}\alpha_j, \qquad 1 \leq i \leq n.$$

These equations may be written in the equivalent form

$$\sum_{j=1}^{n} (\delta_{ij}T - A_{ji}I)\alpha_j = 0, \qquad 1 \leq i \leq n.$$

Let B denote the element of $K^{n\times n}$ with entries

$$B_{ij} = \delta_{ij}T - A_{ji}I.$$

When $n = 2$

$$B = \begin{bmatrix} T - A_{11}I & -A_{21}I \\ -A_{12}I & T - A_{22}I \end{bmatrix}$$

and

$$\begin{aligned} \det B &= (T - A_{11}I)(T - A_{22}I) - A_{12}A_{21}I \\ &= T^2 - (A_{11} + A_{22})T + (A_{11}A_{22} - A_{12}A_{21})I \\ &= f(T) \end{aligned}$$

where f is the characteristic polynomial:

$$f = x^2 - (\text{trace } A)x + \det A.$$

For the case $n > 2$, it is also clear that

$$\det B = f(T)$$

since f is the determinant of the matrix $xI - A$ whose entries are the polynomials

$$(xI - A)_{ij} = \delta_{ij}x - A_{ji}.$$

We wish to show that $f(T) = 0$. In order that $f(T)$ be the zero operator, it is necessary and sufficient that $(\det B)\alpha_k = 0$ for $k = 1, \ldots, n$. By the definition of B, the vectors $\alpha_1, \ldots, \alpha_n$ satisfy the equations

$$\sum_{j=1}^{n} B_{ij}\alpha_j = 0, \qquad 1 \leq i \leq n. \tag{6-6}$$

When $n = 2$, it is suggestive to write (6-6) in the form

$$\begin{bmatrix} T - A_{11}I & -A_{21}I \\ -A_{12}I & T - A_{22}I \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In this case, the classical adjoint, adj B is the matrix

$$\tilde{B} = \begin{bmatrix} T - A_{22}I & A_{21}I \\ A_{12}I & T - A_{11}I \end{bmatrix}$$

and

$$\tilde{B}B = \begin{bmatrix} \det B & 0 \\ 0 & \det B \end{bmatrix}.$$

Hence, we have

$$\begin{aligned}(\det B)\begin{bmatrix}\alpha_1 \\ \alpha_2\end{bmatrix} &= (\tilde{B}B)\begin{bmatrix}\alpha_1 \\ \alpha_2\end{bmatrix} \\ &= \tilde{B}\left(B\begin{bmatrix}\alpha_1 \\ \alpha_2\end{bmatrix}\right) \\ &= \begin{bmatrix}0 \\ 0\end{bmatrix}.\end{aligned}$$

In the general case, let $\tilde{B} = \operatorname{adj} B$. Then by (6-6)

$$\sum_{j=1}^{n} \tilde{B}_{ki}B_{ij}\alpha_j = 0$$

for each pair k, i, and summing on i, we have

$$\begin{aligned}0 &= \sum_{i=1}^{n}\sum_{j=1}^{n} \tilde{B}_{ki}B_{ij}\alpha_j \\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} \tilde{B}_{ki}B_{ij}\right)\alpha_j.\end{aligned}$$

Now $\tilde{B}B = (\det B)I$, so that

$$\sum_{i=1}^{n} \tilde{B}_{ki}B_{ij} = \delta_{kj}\det B.$$

Therefore

$$\begin{aligned}0 &= \sum_{j=1}^{n} \delta_{kj}(\det B)\alpha_j \\ &= (\det B)\alpha_k, \qquad 1 \leq k \leq n. \quad \blacksquare\end{aligned}$$

The Cayley-Hamilton theorem is useful to us at this point primarily because it narrows down the search for the minimal polynomials of various operators. If we know the matrix A which represents T in some ordered basis, then we can compute the characteristic polynomial f. We know that the minimal polynomial p divides f and that the two polynomials have the same roots. There is no method for computing precisely the roots of a polynomial (unless its degree is small); however, if f factors

(6-7) $\quad f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}, \qquad c_1, \ldots, c_k$ distinct, $d_i \geq 1$

then

(6-8) $\quad p = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}, \qquad 1 \leq r_j \leq d_j.$

That is all we can say in general. If f is the polynomial (6-7) and has degree n, then for every polynomial p as in (6-8) we can find an $n \times n$ matrix which has f as its characteristic polynomial and p as its minimal polynomial. We shall not prove this now. But, we want to emphasize the fact that the knowledge that the characteristic polynomial has the form (6-7) tells us that the minimal polynomial has the form (6-8), and it tells us nothing else about p.

EXAMPLE 5. Let A be the 4×4 (rational) matrix

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

The powers of A are easy to compute:

$$A^2 = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{bmatrix}.$$

Thus $A^3 = 4A$, i.e., if $p = x^3 - 4x = x(x + 2)(x - 2)$, then $p(A) = 0$. The minimal polynomial for A must divide p. That minimal polynomial is obviously not of degree 1, since that would mean that A was a scalar multiple of the identity. Hence, the candidates for the minimal polynomial are: p, $x(x + 2)$, $x(x - 2)$, $x^2 - 4$. The three quadratic polynomials can be eliminated because it is obvious at a glance that $A^2 \neq -2A$, $A^2 \neq 2A$, $A^2 \neq 4I$. Therefore p is the minimal polynomial for A. In particular 0, 2, and -2 are the characteristic values of A. One of the factors x, $x - 2$, $x + 2$ must be repeated twice in the characteristic polynomial. Evidently, rank $(A) = 2$. Consequently there is a two-dimensional space of characteristic vectors associated with the characteristic value 0. From Theorem 2, it should now be clear that the characteristic polynomial is $x^2(x^2 - 4)$ and that A is similar over the field of rational numbers to the matrix

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

Exercises

1. Let V be a finite-dimensional vector space. What is the minimal polynomial for the identity operator on V? What is the minimal polynomial for the zero operator?

2. Let a, b, and c be elements of a field F, and let A be the following 3×3 matrix over F:

$$A = \begin{bmatrix} 0 & 0 & c \\ 1 & 0 & b \\ 0 & 1 & a \end{bmatrix}.$$

Prove that the characteristic polynomial for A is $x^3 - ax^2 - bx - c$ and that this is also the minimal polynomial for A.

3. Let A be the 4×4 real matrix

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

Show that the characteristic polynomial for A is $x^2(x-1)^2$ and that it is also the minimal polynomial.

4. Is the matrix A of Exercise 3 similar over the field of complex numbers to a diagonal matrix?

5. Let V be an n-dimensional vector space and let T be a linear operator on V. Suppose that there exists some positive integer k so that $T^k = 0$. Prove that $T^n = 0$.

6. Find a 3×3 matrix for which the minimal polynomial is x^2.

7. Let n be a positive integer, and let V be the space of polynomials over R which have degree at most n (throw in the 0-polynomial). Let D be the differentiation operator on V. What is the minimal polynomial for D?

8. Let P be the operator on R^2 which projects each vector onto the x-axis, parallel to the y-axis: $P(x, y) = (x, 0)$. Show that P is linear. What is the minimal polynomial for P?

9. Let A be an $n \times n$ matrix with characteristic polynomial

$$f = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}.$$

Show that

$$c_1d_1 + \cdots + c_kd_k = \text{trace } (A).$$

10. Let V be the vector space of $n \times n$ matrices over the field F. Let A be a fixed $n \times n$ matrix. Let T be the linear operator on V defined by

$$T(B) = AB.$$

Show that the minimal polynomial for T is the minimal polynomial for A.

11. Let A and B be $n \times n$ matrices over the field F. According to Exercise 9 of Section 6.1, the matrices AB and BA have the same characteristic values. Do they have the same characteristic polynomial? Do they have the same minimal polynomial?

6.4. Invariant Subspaces

In this section, we shall introduce a few concepts which are useful in attempting to analyze a linear operator. We shall use these ideas to obtain characterizations of diagonalizable (and triangulable) operators in terms of their minimal polynomials.

Definition. *Let* V *be a vector space and* T *a linear operator on* V. *If* W *is a subspace of* V, *we say that* W *is* **invariant under** T *if for each vector* α *in* W *the vector* $T\alpha$ *is in* W, *i.e., if* T(W) *is contained in* W.

EXAMPLE 6. If T is any linear operator on V, then V is invariant under T, as is the zero subspace. The range of T and the null space of T are also invariant under T.

EXAMPLE 7. Let F be a field and let D be the differentiation operator on the space $F[x]$ of polynomials over F. Let n be a positive integer and let W be the subspace of polynomials of degree not greater than n. Then W is invariant under D. This is just another way of saying that D is 'degree decreasing.'

EXAMPLE 8. Here is a very useful generalization of Example 6. Let T be a linear operator on V. Let U be any linear operator on V which commutes with T, i.e., $TU = UT$. Let W be the range of U and let N be the null space of U. Both W and N are invariant under T. If α is in the range of U, say $\alpha = U\beta$, then $T\alpha = T(U\beta) = U(T\beta)$ so that $T\alpha$ is in the range of U. If α is in N, then $U(T\alpha) = T(U\alpha) = T(0) = 0$; hence, $T\alpha$ is in N.

A particular type of operator which commutes with T is an operator $U = g(T)$, where g is a polynomial. For instance, we might have $U = T - cI$, where c is a characteristic value of T. The null space of U is familiar to us. We see that this example includes the (obvious) fact that the space of characteristic vectors of T associated with the characteristic value c is invariant under T.

EXAMPLE 9. Let T be the linear operator on R^2 which is represented in the standard ordered basis by the matrix

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

Then the only subspaces of R^2 which are invariant under T are R^2 and the zero subspace. Any other invariant subspace would necessarily have dimension 1. But, if W is the subspace spanned by some non-zero vector α, the fact that W is invariant under T means that α is a characteristic vector, but A has no real characteristic values.

When the subspace W is invariant under the operator T, then T induces a linear operator T_W on the space W. The linear operator T_W is defined by $T_W(\alpha) = T(\alpha)$, for α in W, but T_W is quite a different object from T since its domain is W not V.

When V is finite-dimensional, the invariance of W under T has a simple matrix interpretation, and perhaps we should mention it at this point. Suppose we choose an ordered basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ for V such that $\mathcal{B}' = \{\alpha_1, \ldots, \alpha_r\}$ is an ordered basis for W ($r = \dim W$). Let $A = [T]_{\mathcal{B}}$ so that

$$T\alpha_j = \sum_{i=1}^{n} A_{ij}\alpha_i.$$

Since W is invariant under T, the vector $T\alpha_j$ belongs to W for $j \leq r$. This means that

$$T\alpha_j = \sum_{i=1}^{r} A_{ij}\alpha_i, \qquad j \leq r. \tag{6-9}$$

In other words, $A_{ij} = 0$ if $j \leq r$ and $i > r$.

Schematically, A has the block form

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \tag{6-10}$$

where B is an $r \times r$ matrix, C is an $r \times (n - r)$ matrix, and D is an $(n - r) \times (n - r)$ matrix. The reader should note that according to (6-9) the matrix B is precisely the matrix of the induced operator T_W in the ordered basis $\mathcal{B}'$.

Most often, we shall carry out arguments about T and T_W without making use of the block form of the matrix A in (6-10). But we should note how certain relations between T_W and T are apparent from that block form.

Lemma. *Let* W *be an invariant subspace for* T. *The characteristic polynomial for the restriction operator* T_W *divides the characteristic polynomial for* T. *The minimal polynomial for* T_W *divides the minimal polynomial for* T.

Proof. We have

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

where $A = [T]_{\mathcal{B}}$ and $B = [T_W]_{\mathcal{B}'}$. Because of the block form of the matrix

$$\det(xI - A) = \det(xI - B)\det(xI - D).$$

That proves the statement about characteristic polynomials. Notice that we used I to represent identity matrices of three different sizes.

The kth power of the matrix A has the block form

$$A^k = \begin{bmatrix} B^k & C_k \\ 0 & D^k \end{bmatrix}$$

where C_k is some $r \times (n - r)$ matrix. Therefore, any polynomial which annihilates A also annihilates B (and D too). So, the minimal polynomial for B divides the minimal polynomial for A. ∎

Example 10. Let T be any linear operator on a finite-dimensional space V. Let W be the subspace spanned by *all* of the characteristic vectors