

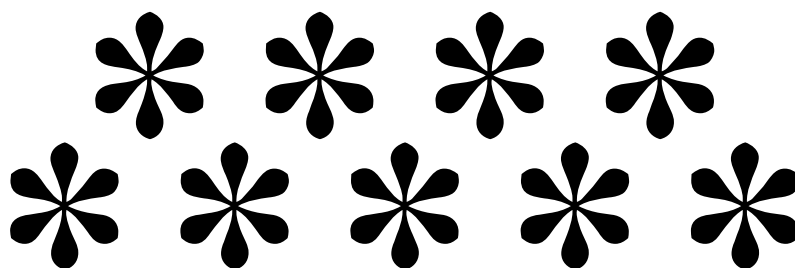
CHAPTER 9

Operators on Real Vector Spaces

In this chapter we delve deeper into the structure of operators on real vector spaces. The important results here are somewhat more complex than the analogous results from the last chapter on complex vector spaces.

Recall that F denotes \mathbf{R} or \mathbf{C} .
Also, V is a finite-dimensional, nonzero vector space over F .

Some of the new results in this chapter are valid on complex vector spaces, so we have not assumed that V is a real vector space.



Eigenvalues of Square Matrices

We have defined eigenvalues of operators; now we need to extend that notion to square matrices. Suppose A is an n -by- n matrix with entries in \mathbf{F} . A number $\lambda \in \mathbf{F}$ is called an **eigenvalue** of A if there exists a nonzero n -by-1 matrix x such that

$$Ax = \lambda x.$$

For example, 3 is an eigenvalue of $\begin{bmatrix} 7 & 8 \\ 1 & 5 \end{bmatrix}$ because

$$\begin{bmatrix} 7 & 8 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 6 \\ -3 \end{bmatrix} = 3 \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

As another example, you should verify that the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has no eigenvalues if we are thinking of \mathbf{F} as the real numbers (by definition, an eigenvalue must be in \mathbf{F}) and has eigenvalues i and $-i$ if we are thinking of \mathbf{F} as the complex numbers.

We now have two notions of eigenvalue—one for operators and one for square matrices. As you might expect, these two notions are closely connected, as we now show.

9.1 Proposition: *Suppose $T \in \mathcal{L}(V)$ and A is the matrix of T with respect to some basis of V . Then the eigenvalues of T are the same as the eigenvalues of A .*

PROOF: Let (v_1, \dots, v_n) be the basis of V with respect to which T has matrix A . Let $\lambda \in \mathbf{F}$. We need to show that λ is an eigenvalue of T if and only if λ is an eigenvalue of A .

First suppose λ is an eigenvalue of T . Let $v \in V$ be a nonzero vector such that $Tv = \lambda v$. We can write

$$\mathbf{9.2} \quad v = a_1 v_1 + \dots + a_n v_n,$$

where $a_1, \dots, a_n \in \mathbf{F}$. Let x be the matrix of the vector v with respect to the basis (v_1, \dots, v_n) . Recall from Chapter 3 that this means

$$\mathbf{9.3} \quad x = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

We have

$$Ax = \mathcal{M}(T)\mathcal{M}(\nu) = \mathcal{M}(T\nu) = \mathcal{M}(\lambda\nu) = \lambda\mathcal{M}(\nu) = \lambda x,$$

where the second equality comes from 3.14. The equation above shows that λ is an eigenvalue of A , as desired.

To prove the implication in the other direction, now suppose λ is an eigenvalue of A . Let x be a nonzero n -by-1 matrix such that $Ax = \lambda x$. We can write x in the form 9.3 for some scalars $a_1, \dots, a_n \in \mathbb{F}$. Define $\nu \in V$ by 9.2. Then

$$\mathcal{M}(T\nu) = \mathcal{M}(T)\mathcal{M}(\nu) = Ax = \lambda x = \mathcal{M}(\lambda\nu).$$

where the first equality comes from 3.14. The equation above implies that $T\nu = \lambda\nu$, and thus λ is an eigenvalue of T , completing the proof. ■

Because every square matrix is the matrix of some operator, the proposition above allows us to translate results about eigenvalues of operators into the language of eigenvalues of square matrices. For example, every square matrix of complex numbers has an eigenvalue (from 5.10). As another example, every n -by- n matrix has at most n distinct eigenvalues (from 5.9).

Block Upper-Triangular Matrices

Earlier we proved that each operator on a complex vector space has an upper-triangular matrix with respect to some basis (see 5.13). In this section we will see that we can almost do as well on real vector spaces.

In the last two chapters we used block diagonal matrices, which extend the notion of diagonal matrices. Now we will need to use the corresponding extension of upper-triangular matrices. A **block upper-triangular matrix** is a square matrix of the form

$$\begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where A_1, \dots, A_m are square matrices lying along the diagonal, all entries below A_1, \dots, A_m equal 0, and the $*$ denotes arbitrary entries. For example, the matrix

As usual, we use an asterisk to denote entries of the matrix that play no important role in the topics under consideration.

$$A = \begin{bmatrix} 4 & 10 & 11 & 12 & 13 \\ 0 & -3 & -3 & 14 & 25 \\ 0 & -3 & -3 & 16 & 17 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 0 & 0 & 5 & 5 \end{bmatrix}$$

is a block upper-triangular matrix with

$$A = \begin{bmatrix} A_1 & & * \\ & A_2 & \\ 0 & & A_3 \end{bmatrix},$$

where

$$A_1 = \begin{bmatrix} 4 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -3 & -3 \\ -3 & -3 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}.$$

Every upper-triangular matrix is also a block upper-triangular matrix with blocks of size 1-by-1 along the diagonal. At the other extreme, every square matrix is a block upper-triangular matrix because we can take the first (and only) block to be the entire matrix. Smaller blocks are better in the sense that the matrix then has more 0's.

Now we prove that for each operator on a real vector space, we can find a basis that gives a block upper-triangular matrix with blocks of size at most 2-by-2 on the diagonal.

9.4 Theorem: Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Then there is a basis of V with respect to which T has a block upper-triangular matrix

$$\begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues.

PROOF: Clearly the desired result holds if $\dim V = 1$.

Next, consider the case where $\dim V = 2$. If T has an eigenvalue λ , then let $v_1 \in V$ be any nonzero eigenvector. Extend (v_1) to a basis (v_1, v_2) of V . With respect to this basis, T has an upper-triangular matrix of the form

$$\begin{bmatrix} \lambda & a \\ 0 & b \end{bmatrix}.$$

In particular, if T has an eigenvalue, then there is a basis of V with respect to which T has an upper-triangular matrix. If T has no eigenvalues, then choose any basis (v_1, v_2) of V . With respect to this basis,

the matrix of T has no eigenvalues (by 9.1). Thus regardless of whether T has eigenvalues, we have the desired conclusion when $\dim V = 2$.

Suppose now that $\dim V > 2$ and the desired result holds for all real vector spaces with smaller dimension. If T has an eigenvalue, let U be a one-dimensional subspace of V that is invariant under T ; otherwise let U be a two-dimensional subspace of V that is invariant under T (5.24 guarantees that we can choose U in this fashion). Choose any basis of U and let A_1 denote the matrix of $T|_U$ with respect to this basis. If A_1 is a 2-by-2 matrix, then T has no eigenvalues (otherwise we would have chosen U to be one-dimensional) and thus $T|_U$ has no eigenvalues. Hence if A_1 is a 2-by-2 matrix, then A_1 has no eigenvalues (see 9.1).

Let W be any subspace of V such that

$$V = U \oplus W;$$

2.13 guarantees that such a W exists. Because W has dimension less than the dimension of V , we would like to apply our induction hypothesis to $T|_W$. However, W might not be invariant under T , meaning that $T|_W$ might not be an operator on W . We will compose with the projection $P_{W,U}$ to get an operator on W . Specifically, define $S \in \mathcal{L}(W)$ by

$$S\mathbf{w} = P_{W,U}(T\mathbf{w})$$

for $\mathbf{w} \in W$. Note that

$$\begin{aligned} T\mathbf{w} &= P_{U,W}(T\mathbf{w}) + P_{W,U}(T\mathbf{w}) \\ &= P_{U,W}(T\mathbf{w}) + S\mathbf{w} \end{aligned}$$

9.6

for every $\mathbf{w} \in W$.

By our induction hypothesis, there is a basis of W with respect to which S has a block upper-triangular matrix of the form

$$\begin{bmatrix} A_2 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues. Adjoin this basis of W to the basis of U chosen above, getting a basis of V . A minute's thought should convince you (use 9.6) that the matrix of T with respect to this basis is a block upper-triangular matrix of the form 9.5, completing the proof. ■

Recall that if $\mathbf{v} = \mathbf{w} + \mathbf{u}$, where $\mathbf{w} \in W$ and $\mathbf{u} \in U$, then $P_{W,U}\mathbf{v} = \mathbf{w}$.

The Characteristic Polynomial

For operators on complex vector spaces, we defined characteristic polynomials and developed their properties by making use of upper-triangular matrices. In this section we will carry out a similar procedure for operators on real vector spaces. Instead of upper-triangular matrices, we will have to use the block upper-triangular matrices furnished by the last theorem.

In the last chapter, we did not define the characteristic polynomial of a square matrix with complex entries because our emphasis is on operators rather than on matrices. However, to understand operators on real vector spaces, we will need to define characteristic polynomials of 1-by-1 and 2-by-2 matrices with real entries. Then, using block-upper triangular matrices with blocks of size at most 2-by-2 on the diagonal, we will be able to define the characteristic polynomial of an operator on a real vector space.

To motivate the definition of characteristic polynomials of square matrices, we would like the following to be true (think about the Cayley-Hamilton theorem; see 8.20): if $T \in \mathcal{L}(V)$ has matrix A with respect to some basis of V and q is the characteristic polynomial of A , then $q(T) = 0$.

Let's begin with the trivial case of 1-by-1 matrices. Suppose V is a real vector space with dimension 1 and $T \in \mathcal{L}(V)$. If $[\lambda]$ equals the matrix of T with respect to some basis of V , then T equals λI . Thus if we let q be the degree 1 polynomial defined by $q(x) = x - \lambda$, then $q(T) = 0$. Hence we define the characteristic polynomial of $[\lambda]$ to be $x - \lambda$.

Now let's look at 2-by-2 matrices with real entries. Suppose V is a real vector space with dimension 2 and $T \in \mathcal{L}(V)$. Suppose

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

is the matrix of T with respect to some basis (v_1, v_2) of V . We seek a monic polynomial q of degree 2 such that $q(T) = 0$. If $b = 0$, then the matrix above is upper triangular. If in addition we were dealing with a complex vector space, then we would know that T has characteristic polynomial $(z - a)(z - d)$. Thus a reasonable candidate might be $(x - a)(x - d)$, where we use x instead of z to emphasize that now we are working on a real vector space. Let's see if the polynomial

$(x - a)(x - d)$, when applied to T , gives 0 even when $b \neq 0$. We have

$$(T - aI)(T - dI)v_1 = (T - dI)(T - aI)v_1 = (T - dI)(bv_2) = bc v_1$$

and

$$(T - aI)(T - dI)v_2 = (T - aI)(cv_1) = bc v_2.$$

Thus $(T - aI)(T - dI)$ is not equal to 0 unless $bc = 0$. However, the equations above show that $(T - aI)(T - dI) - bcI = 0$ (because this operator equals 0 on a basis, it must equal 0 on V). Thus if $q(x) = (x - a)(x - d) - bc$, then $q(T) = 0$.

Motivated by the previous paragraph, we define the **characteristic polynomial** of a 2-by-2 matrix $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$ to be $(x - a)(x - d) - bc$. Here we are concerned only with matrices with real entries. The next result shows that we have found the only reasonable definition for the characteristic polynomial of a 2-by-2 matrix.

9.7 Proposition: Suppose V is a real vector space with dimension 2 and $T \in \mathcal{L}(V)$ has no eigenvalues. Let $p \in \mathcal{P}(\mathbf{R})$ be a monic polynomial with degree 2. Suppose A is the matrix of T with respect to some basis of V .

- (a) If p equals the characteristic polynomial of A , then $p(T) = 0$.
- (b) If p does not equal the characteristic polynomial of A , then $p(T)$ is invertible.

PROOF: We already proved (a) in our discussion above. To prove (b), let q denote the characteristic polynomial of A and suppose that $p \neq q$. We can write $p(x) = x^2 + \alpha_1 x + \beta_1$ and $q(x) = x^2 + \alpha_2 x + \beta_2$ for some $\alpha_1, \beta_1, \alpha_2, \beta_2 \in \mathbf{R}$. Now

$$p(T) = p(T) - q(T) = (\alpha_1 - \alpha_2)T + (\beta_1 - \beta_2)I.$$

If $\alpha_1 = \alpha_2$, then $\beta_1 \neq \beta_2$ (otherwise we would have $p = q$). Thus if $\alpha_1 = \alpha_2$, then $p(T)$ is a nonzero multiple of the identity and hence is invertible, as desired. If $\alpha_1 \neq \alpha_2$, then

$$p(T) = (\alpha_1 - \alpha_2)\left(T - \frac{\beta_2 - \beta_1}{\alpha_1 - \alpha_2}I\right),$$

which is an invertible operator because T has no eigenvalues. Thus (b) holds. ■

Part (b) of this proposition would be false without the hypothesis that T has no eigenvalues. For example, define $T \in \mathcal{L}(\mathbf{R}^2)$ by $T(x_1, x_2) = (0, x_2)$. Take $p(x) = x(x - 2)$. Then p is not the characteristic polynomial of the matrix of T with respect to the standard basis, but $p(T)$ is not invertible.

Suppose V is a real vector space with dimension 2 and $T \in \mathcal{L}(V)$ has no eigenvalues. The last proposition shows that there is precisely one monic polynomial with degree 2 that when applied to T gives 0. Thus, though T may have different matrices with respect to different bases, each of these matrices must have the same characteristic polynomial. For example, consider $T \in \mathcal{L}(\mathbf{R}^2)$ defined by

$$\mathbf{9.8} \quad T(x_1, x_2) = (3x_1 + 5x_2, -2x_1 - x_2).$$

The matrix of T with respect to the standard basis of \mathbf{R}^2 is

$$\begin{bmatrix} 3 & 5 \\ -2 & -1 \end{bmatrix}.$$

The characteristic polynomial of this matrix is $(x - 3)(x + 1) + 2 \cdot 5$, which equals $x^2 - 2x + 7$. As you should verify, the matrix of T with respect to the basis $((-2, 1), (1, 2))$ equals

$$\begin{bmatrix} 1 & -6 \\ 1 & 1 \end{bmatrix}.$$

The characteristic polynomial of this matrix is $(x - 1)(x - 1) + 1 \cdot 6$, which equals $x^2 - 2x + 7$, the same result we obtained by using the standard basis.

When analyzing upper-triangular matrices of an operator T on a complex vector space V , we found that subspaces of the form

$$\text{null}(T - \lambda I)^{\dim V}$$

played a key role (see 8.10). Those spaces will also play a role in studying operators on real vector spaces, but because we must now consider block upper-triangular matrices with 2-by-2 blocks, subspaces of the form

$$\text{null}(T^2 + \alpha T + \beta I)^{\dim V}$$

will also play a key role. To get started, let's look at one- and two-dimensional real vector spaces.

First suppose that V is a one-dimensional real vector space and that $T \in \mathcal{L}(V)$. If $\lambda \in \mathbf{R}$, then $\text{null}(T - \lambda I)$ equals V if λ is an eigenvalue of T and $\{0\}$ otherwise. If $\alpha, \beta \in \mathbf{R}$ with $\alpha^2 < 4\beta$, then

$$\text{null}(T^2 + \alpha T + \beta I) = \{0\}.$$

(Proof: Because V is one-dimensional, there is a constant $\lambda \in \mathbf{R}$ such that $Tv = \lambda v$ for all $v \in V$. Thus $(T^2 + \alpha T + \beta I)v = (\lambda^2 + \alpha\lambda + \beta)v$. However, the inequality $\alpha^2 < 4\beta$ implies that $\lambda^2 + \alpha\lambda + \beta \neq 0$, and thus $\text{null}(T^2 + \alpha T + \beta I) = \{0\}$.)

Now suppose V is a two-dimensional real vector space and $T \in \mathcal{L}(V)$ has no eigenvalues. If $\lambda \in \mathbf{R}$, then $\text{null}(T - \lambda I)$ equals $\{0\}$ (because T has no eigenvalues). If $\alpha, \beta \in \mathbf{R}$ with $\alpha^2 < 4\beta$, then $\text{null}(T^2 + \alpha T + \beta I)$ equals V if $x^2 + \alpha x + \beta$ is the characteristic polynomial of the matrix of T with respect to some (or equivalently, every) basis of V and equals $\{0\}$ otherwise (by 9.7). Note that for this operator, there is no middle ground—the null space of $T^2 + \alpha T + \beta I$ is either $\{0\}$ or the whole space; it cannot be one-dimensional.

Now suppose that V is a real vector space of any dimension and $T \in \mathcal{L}(V)$. We know that V has a basis with respect to which T has a block upper-triangular matrix with blocks on the diagonal of size at most 2-by-2 (see 9.4). In general, this matrix is not unique— V may have many different bases with respect to which T has a block upper-triangular matrix of this form, and with respect to these different bases we may get different block upper-triangular matrices.

We encountered a similar situation when dealing with complex vector spaces and upper-triangular matrices. In that case, though we might get different upper-triangular matrices with respect to the different bases, the entries on the diagonal were always the same (though possibly in a different order). Might a similar property hold for real vector spaces and block upper-triangular matrices? Specifically, is the number of times a given 2-by-2 matrix appears on the diagonal of a block upper-triangular matrix of T independent of which basis is chosen? Unfortunately this question has a negative answer. For example, the operator $T \in \mathcal{L}(\mathbf{R}^2)$ defined by 9.8 has two different 2-by-2 matrices, as we saw above.

Though the number of times a particular 2-by-2 matrix might appear on the diagonal of a block upper-triangular matrix of T can depend on the choice of basis, if we look at characteristic polynomials instead of the actual matrices, we find that the number of times a particular characteristic polynomial appears is independent of the choice of basis. This is the content of the following theorem, which will be our key tool in analyzing the structure of an operator on a real vector space.

Recall that $\alpha^2 < 4\beta$ implies that $x^2 + \alpha x + \beta$ has no real roots; see 4.11.

9.9 Theorem: Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose that with respect to some basis of V , the matrix of T is

$$9.10 \quad \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues.

(a) If $\lambda \in \mathbf{R}$, then precisely $\dim \text{null}(T - \lambda I)^{\dim V}$ of the matrices A_1, \dots, A_m equal the 1-by-1 matrix $[\lambda]$.

This result implies that $\text{null}(T^2 + \alpha T + \beta I)^{\dim V}$ must have even dimension.

(b) If $\alpha, \beta \in \mathbf{R}$ satisfy $\alpha^2 < 4\beta$, then precisely

$$\frac{\dim \text{null}(T^2 + \alpha T + \beta I)^{\dim V}}{2}$$

of the matrices A_1, \dots, A_m have characteristic polynomial equal to $x^2 + \alpha x + \beta$.

This proof uses the same ideas as the proof of the analogous result on complex vector spaces (8.10). As usual, the real case is slightly more complicated but requires no new creativity.

PROOF: We will construct one proof that can be used to prove both (a) and (b). To do this, let $\lambda, \alpha, \beta \in \mathbf{R}$ with $\alpha^2 < 4\beta$. Define $p \in \mathcal{P}(\mathbf{R})$ by

$$p(x) = \begin{cases} x - \lambda & \text{if we are trying to prove (a);} \\ x^2 + \alpha x + \beta & \text{if we are trying to prove (b).} \end{cases}$$

Let d denote the degree of p . Thus $d = 1$ if we are trying to prove (a) and $d = 2$ if we are trying to prove (b).

We will prove this theorem by induction on m , the number of blocks along the diagonal of 9.10. If $m = 1$, then $\dim V = 1$ or $\dim V = 2$; the discussion preceding this theorem then implies that the desired result holds. Thus we can assume that $m > 1$ and that the desired result holds when m is replaced with $m - 1$.

For convenience let $n = \dim V$. Consider a basis of V with respect to which T has the block upper-triangular matrix 9.10. Let U_j denote the span of the basis vectors corresponding to A_j . Thus $\dim U_j = 1$ if A_j is a 1-by-1 matrix and $\dim U_j = 2$ if A_j is a 2-by-2 matrix. Let $U = U_1 + \dots + U_{m-1}$. Clearly U is invariant under T and the matrix of $T|_U$ with respect to the obvious basis (obtained from the basis vectors corresponding to A_1, \dots, A_{m-1}) is

$$9.11 \quad \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_{m-1} \end{bmatrix}.$$

Thus, by our induction hypothesis,

9.12 precisely $(1/d) \dim \text{null } p(T|_U)^n$ of the matrices A_1, \dots, A_{m-1} have characteristic polynomial p .

Actually the induction hypothesis gives 9.12 with exponent $\dim U$ instead of n , but then we can replace $\dim U$ with n (by 8.6) to get the statement above.

Suppose $u_m \in U_m$. Let $S \in \mathcal{L}(U_m)$ be the operator whose matrix (with respect to the basis corresponding to U_m) equals A_m . In particular, $Su_m = P_{U_m, U} T u_m$. Now

$$\begin{aligned} T u_m &= P_{U, U_m} T u_m + P_{U_m, U} T u_m \\ &= *_U + S u_m, \end{aligned}$$

where $*_U$ denotes a vector in U . Note that $Su_m \in U_m$; thus applying T to both sides of the equation above gives

$$T^2 u_m = *_U + S^2 u_m,$$

where again $*_U$ denotes a vector in U , though perhaps a different vector than the previous usage of $*_U$ (the notation $*_U$ is used when we want to emphasize that we have a vector in U but we do not care which particular vector—each time the notation $*_U$ is used, it may denote a different vector in U). The last two equations show that

$$\mathbf{9.13} \quad p(T)u_m = *_U + p(S)u_m$$

for some $*_U \in U$. Note that $p(S)u_m \in U_m$; thus iterating the last equation gives

$$\mathbf{9.14} \quad p(T)^n u_m = *_U + p(S)^n u_m$$

for some $*_U \in U$.

The proof now breaks into two cases. First consider the case where the characteristic polynomial of A_m does not equal p . We will show that in this case

$$\mathbf{9.15} \quad \text{null } p(T)^n \subset U.$$

Once this has been verified, we will know that

$$\text{null } p(T)^n = \text{null } p(T|_U)^n,$$

and hence 9.12 will tell us that precisely $(1/d) \dim \text{null } p(T)^n$ of the matrices A_1, \dots, A_m have characteristic polynomial p , completing the proof in the case where the characteristic polynomial of A_m does not equal p .

To prove 9.15 (still assuming that the characteristic polynomial of A_m does not equal p), suppose $v \in \text{null } p(T)^n$. We can write v in the form $v = u + u_m$, where $u \in U$ and $u_m \in U_m$. Using 9.14, we have

$$0 = p(T)^n v = p(T)^n u + p(T)^n u_m = p(T)^n u + *_{U} + p(S)^n u_m$$

for some $*_{U} \in U$. Because the vectors $p(T)^n u$ and $*_{U}$ are in U and $p(S)^n u_m \in U_m$, this implies that $p(S)^n u_m = 0$. However, $p(S)$ is invertible (see the discussion preceding this theorem about one- and two-dimensional subspaces and note that $\dim U_m \leq 2$), so $u_m = 0$. Thus $v = u \in U$, completing the proof of 9.15.

Now consider the case where the characteristic polynomial of A_m equals p . Note that this implies $\dim U_m = d$. We will show that

$$\mathbf{9.16} \quad \dim \text{null } p(T)^n = \dim \text{null } p(T|_U)^n + d,$$

which along with 9.12 will complete the proof.

Using the formula for the dimension of the sum of two subspaces (2.18), we have

$$\begin{aligned} \dim \text{null } p(T)^n &= \dim(U \cap \text{null } p(T)^n) + \dim(U + \text{null } p(T)^n) - \dim U \\ &= \dim \text{null } p(T|_U)^n + \dim(U + \text{null } p(T)^n) - (n - d). \end{aligned}$$

If $U + \text{null } p(T)^n = V$, then $\dim(U + \text{null } p(T)^n) = n$, which when combined with the last formula above for $\dim \text{null } p(T)^n$ would give 9.16, as desired. Thus we will finish by showing that $U + \text{null } p(T)^n = V$.

To prove that $U + \text{null } p(T)^n = V$, suppose $u_m \in U_m$. Because the characteristic polynomial of the matrix of S (namely, A_m) equals p , we have $p(S) = 0$. Thus $p(T)u_m \in U$ (from 9.13). Now

$$p(T)^n u_m = p(T)^{n-1}(p(T)u_m) \in \text{range } p(T|_U)^{n-1} = \text{range } p(T|_U)^n,$$

where the last equality comes from 8.9. Thus we can choose $u \in U$ such that $p(T)^n u_m = p(T|_U)^n u$. Now

$$\begin{aligned} p(T)^n(u_m - u) &= p(T)^n u_m - p(T)^n u \\ &= p(T)^n u_m - p(T|_U)^n u \\ &= 0. \end{aligned}$$

Thus $u_m - u \in \text{null } p(T)^n$, and hence u_m , which equals $u + (u_m - u)$, is in $U + \text{null } p(T)^n$. In other words, $U_m \subset U + \text{null } p(T)^n$. Therefore $V = U + U_m \subset U + \text{null } p(T)^n$, and hence $U + \text{null } p(T)^n = V$, completing the proof. ■

As we saw in the last chapter, the eigenvalues of an operator on a complex vector space provide the key to analyzing the structure of the operator. On a real vector space, an operator may have fewer eigenvalues, counting multiplicity, than the dimension of the vector space. The previous theorem suggests a definition that makes up for this deficiency. We will see that the definition given in the next paragraph helps make operator theory on real vector spaces resemble operator theory on complex vector spaces.

Suppose V is a real vector space and $T \in \mathcal{L}(V)$. An ordered pair (α, β) of real numbers is called an **eigenpair** of T if $\alpha^2 < 4\beta$ and

$$T^2 + \alpha T + \beta I$$

is not injective. The previous theorem shows that T can have only finitely many eigenpairs because each eigenpair corresponds to the characteristic polynomial of a 2-by-2 matrix on the diagonal of 9.10 and there is room for only finitely many such matrices along that diagonal. Guided by 9.9, we define the **multiplicity** of an eigenpair (α, β) of T to be

$$\frac{\dim \text{null}(T^2 + \alpha T + \beta I)^{\dim V}}{2}.$$

From 9.9, we see that the multiplicity of (α, β) equals the number of times that $x^2 + \alpha x + \beta$ is the characteristic polynomial of a 2-by-2 matrix on the diagonal of 9.10.

As an example, consider the operator $T \in \mathcal{L}(\mathbb{R}^3)$ whose matrix (with respect to the standard basis) equals

$$\begin{bmatrix} 3 & -1 & -2 \\ 3 & 2 & -3 \\ 1 & 2 & 0 \end{bmatrix}.$$

You should verify that $(-4, 13)$ is an eigenpair of T with multiplicity 1; note that $T^2 - 4T + 13I$ is not injective because $(-1, 0, 1)$ and $(1, 1, 0)$ are in its null space. Without doing any calculations, you should verify that T has no other eigenpairs (use 9.9). You should also verify that 1 is an eigenvalue of T with multiplicity 1, with corresponding eigenvector $(1, 0, 1)$, and that T has no other eigenvalues.

*Though the word **eigenpair** was chosen to be consistent with the word **eigenvalue**, this terminology is not in widespread use.*

In the example above, the sum of the multiplicities of the eigenvalues of T plus twice the multiplicities of the eigenpairs of T equals 3, which is the dimension of the domain of T . The next proposition shows that this always happens on a real vector space.

This proposition shows that though an operator on a real vector space may have no eigenvalues, or it may have no eigenpairs, it cannot be lacking in both these useful objects. It also shows that an operator on a real vector space V can have at most $(\dim V)/2$ distinct eigenpairs.

9.17 Proposition: *If V is a real vector space and $T \in \mathcal{L}(V)$, then the sum of the multiplicities of all the eigenvalues of T plus the sum of twice the multiplicities of all the eigenpairs of T equals $\dim V$.*

PROOF: Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Then there is a basis of V with respect to which the matrix of T is as in 9.9. The multiplicity of an eigenvalue λ equals the number of times the 1-by-1 matrix $[\lambda]$ appears on the diagonal of this matrix (from 9.9). The multiplicity of an eigenpair (α, β) equals the number of times $x^2 + \alpha x + \beta$ is the characteristic polynomial of a 2-by-2 matrix on the diagonal of this matrix (from 9.9). Because the diagonal of this matrix has length $\dim V$, the sum of the multiplicities of all the eigenvalues of T plus the sum of twice the multiplicities of all the eigenpairs of T must equal $\dim V$. ■

Suppose V is a real vector space and $T \in \mathcal{L}(V)$. With respect to some basis of V , T has a block upper-triangular matrix of the form

$$\begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix}, \quad \text{9.18}$$

where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues (see 9.4). We define the **characteristic polynomial** of T to be the product of the characteristic polynomials of A_1, \dots, A_m . Explicitly, for each j , define $q_j \in \mathcal{P}(\mathbf{R})$ by

$$\text{9.19} \quad q_j(x) = \begin{cases} x - \lambda & \text{if } A_j \text{ equals } [\lambda]; \\ (x - a)(x - d) - bc & \text{if } A_j \text{ equals } \begin{bmatrix} a & c \\ b & d \end{bmatrix}. \end{cases}$$

Note that the roots of the characteristic polynomial of T equal the eigenvalues of T , as was true on complex vector spaces.

Then the characteristic polynomial of T is

$$q_1(x) \dots q_m(x).$$

Clearly the characteristic polynomial of T has degree $\dim V$. Furthermore, 9.9 insures that the characteristic polynomial of T depends only on T and not on the choice of a particular basis.

Now we can prove a result that was promised in the last chapter, where we proved the analogous theorem (8.20) for operators on complex vector spaces.

9.20 Cayley-Hamilton Theorem: *Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Let q denote the characteristic polynomial of T . Then $q(T) = 0$.*

PROOF: Choose a basis of V with respect to which T has a block upper-triangular matrix of the form 9.18, where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues. Suppose U_j is the one- or two-dimensional subspace spanned by the basis vectors corresponding to A_j . Define q_j as in 9.19. To prove that $q(T) = 0$, we need only show that $q(T)|_{U_j} = 0$ for $j = 1, \dots, m$. To do this, it suffices to show that

This proof uses the same ideas as the proof of the analogous result on complex vector spaces (8.20).

$$9.21 \quad q_1(T) \dots q_j(T)|_{U_j} = 0$$

for $j = 1, \dots, m$.

We will prove 9.21 by induction on j . To get started, suppose that $j = 1$. Because $\mathcal{M}(T)$ is given by 9.18, we have $q_1(T)|_{U_1} = 0$ (obvious if $\dim U_1 = 1$; from 9.7(a) if $\dim U_1 = 2$), giving 9.21 when $j = 1$.

Now suppose that $1 < j \leq n$ and that

$$\begin{aligned} 0 &= q_1(T)|_{U_1} \\ 0 &= q_1(T)q_2(T)|_{U_2} \\ &\vdots \\ 0 &= q_1(T) \dots q_{j-1}(T)|_{U_{j-1}}. \end{aligned}$$

If $v \in U_j$, then from 9.18 we see that

$$q_j(T)v = u + q_j(S)v,$$

where $u \in U_1 + \dots + U_{j-1}$ and $S \in \mathcal{L}(U_j)$ has characteristic polynomial q_j . Because $q_j(S) = 0$ (obvious if $\dim U_j = 1$; from 9.7(a) if $\dim U_j = 2$), the equation above shows that

$$q_j(T)v \in U_1 + \dots + U_{j-1}$$

whenever $v \in U_j$. Thus, by our induction hypothesis, $q_1(T) \dots q_{j-1}(T)$ applied to $q_j(T)v$ gives 0 whenever $v \in U_j$. In other words, 9.21 holds, completing the proof. ■

Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Clearly the Cayley-Hamilton theorem (9.20) implies that the minimal polynomial of T has degree at most $\dim V$, as was the case on complex vector spaces. If the degree of the minimal polynomial of T equals $\dim V$, then, as was also the case on complex vector spaces, the minimal polynomial of T must equal the characteristic polynomial of T . This follows from the Cayley-Hamilton theorem (9.20) and 8.34.

Finally, we can now prove a major structure theorem about operators on real vector spaces. The theorem below should be compared to 8.23, the corresponding result on complex vector spaces.

9.22 Theorem: *Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T , with U_1, \dots, U_m the corresponding sets of generalized eigenvectors. Let $(\alpha_1, \beta_1), \dots, (\alpha_M, \beta_M)$ be the distinct eigenpairs of T and let $V_j = \text{null}(T^2 + \alpha_j T + \beta_j I)^{\dim V}$. Then*

- (a) $V = U_1 \oplus \dots \oplus U_m \oplus V_1 \oplus \dots \oplus V_M$;
- (b) each U_j and each V_j is invariant under T ;
- (c) each $(T - \lambda_j I)|_{U_j}$ and each $(T^2 + \alpha_j T + \beta_j I)|_{V_j}$ is nilpotent.

This proof uses the same ideas as the proof of the analogous result on complex vector spaces (8.23).

PROOF: From 8.22, we get (b). Clearly (c) follows from the definitions.

To prove (a), recall that $\dim U_j$ equals the multiplicity of λ_j as an eigenvalue of T and $\dim V_j$ equals twice the multiplicity of (α_j, β_j) as an eigenpair of T . Thus

$$\mathbf{9.23} \quad \dim V = \dim U_1 + \dots + \dim U_m + \dim V_1 + \dots + \dim V_M;$$

this follows from 9.17. Let $U = U_1 + \dots + U_m + V_1 + \dots + V_M$. Note that U is invariant under T . Thus we can define $S \in \mathcal{L}(U)$ by

$$S = T|_U.$$

Note that S has the same eigenvalues, with the same multiplicities, as T because all the generalized eigenvectors of T are in U , the domain of S . Similarly, S has the same eigenpairs, with the same multiplicities, as T . Thus applying 9.17 to S , we get

$$\dim U = \dim U_1 + \dots + \dim U_m + \dim V_1 + \dots + \dim V_M.$$

This equation, along with 9.23, shows that $\dim V = \dim U$. Because U is a subspace of V , this implies that $V = U$. In other words,

$$V = U_1 + \cdots + U_m + V_1 + \cdots + V_M.$$

This equation, along with 9.23, allows us to use 2.19 to conclude that (a) holds, completing the proof. ■

Exercises

1. Prove that 1 is an eigenvalue of every square matrix with the property that the sum of the entries in each row equals 1.
2. Consider a 2-by-2 matrix of real numbers

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

Prove that A has an eigenvalue (in \mathbf{R}) if and only if

$$(a - d)^2 + 4bc \geq 0.$$

3. Suppose A is a block diagonal matrix

$$A = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is a square matrix. Prove that the set of eigenvalues of A equals the union of the eigenvalues of A_1, \dots, A_m .

Clearly Exercise 4 is a stronger statement than Exercise 3. Even so, you may want to do Exercise 3 first because it is easier than Exercise 4.

4. Suppose A is a block upper-triangular matrix

$$A = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is a square matrix. Prove that the set of eigenvalues of A equals the union of the eigenvalues of A_1, \dots, A_m .

5. Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose $\alpha, \beta \in \mathbf{R}$ are such that $T^2 + \alpha T + \beta I = 0$. Prove that T has an eigenvalue if and only if $\alpha^2 \geq 4\beta$.
6. Suppose V is a real inner-product space and $T \in \mathcal{L}(V)$. Prove that there is an orthonormal basis of V with respect to which T has a block upper-triangular matrix

$$\begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each A_j is a 1-by-1 matrix or a 2-by-2 matrix with no eigenvalues.

7. Prove that if $T \in \mathcal{L}(V)$ and j is a positive integer such that $j \leq \dim V$, then T has an invariant subspace whose dimension equals $j - 1$ or j .
8. Prove that there does not exist an operator $T \in \mathcal{L}(\mathbf{R}^7)$ such that $T^2 + T + I$ is nilpotent.
9. Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^7)$ such that $T^2 + T + I$ is nilpotent.
10. Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose $\alpha, \beta \in \mathbf{R}$ are such that $\alpha^2 < 4\beta$. Prove that

$$\text{null}(T^2 + \alpha T + \beta I)^k$$

has even dimension for every positive integer k .

11. Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose $\alpha, \beta \in \mathbf{R}$ are such that $\alpha^2 < 4\beta$ and $T^2 + \alpha T + \beta I$ is nilpotent. Prove that $\dim V$ is even and

$$(T^2 + \alpha T + \beta I)^{\dim V/2} = 0.$$

12. Prove that if $T \in \mathcal{L}(\mathbf{R}^3)$ and 5, 7 are eigenvalues of T , then T has no eigenpairs.
13. Suppose V is a real vector space with $\dim V = n$ and $T \in \mathcal{L}(V)$ is such that

$$\text{null } T^{n-2} \neq \text{null } T^{n-1}.$$

Prove that T has at most two distinct eigenvalues and that T has no eigenpairs.

14. Suppose V is a vector space with dimension 2 and $T \in \mathcal{L}(V)$. Prove that if

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

is the matrix of T with respect to some basis of V , then the characteristic polynomial of T equals $(z - a)(z - d) - bc$.

15. Suppose V is a real inner-product space and $S \in \mathcal{L}(V)$ is an isometry. Prove that if (α, β) is an eigenpair of S , then $\beta = 1$.

You do not need to find the eigenvalues of T to do this exercise. As usual unless otherwise specified, here V may be a real or complex vector space.