$$\begin{aligned} A &= A_1 + A_2 \\ A_1 &= \tfrac{1}{2}(A + A^t) \\ A_2 &= \tfrac{1}{2}(A - A^t). \end{aligned}$$

EXAMPLE 13. Let T be any linear operator on a finite-dimensional space V. Let $c_1, \ldots, c_k$ be the distinct characteristic values of T, and let W_i be the space of characteristic vectors associated with the characteristic value c_i. Then $W_1, \ldots, W_k$ are independent. See the lemma before Theorem 2. In particular, if T is diagonalizable, then $V = W_1 \oplus \cdots \oplus W_k$.

Definition. *If* V *is a vector space, a* **projection** *of* V *is a linear operator* E *on* V *such that* $E^2 = E$.

Suppose that E is a projection. Let R be the range of E and let N be the null space of E.

1. The vector β is in the range R if and only if $E\beta = \beta$. If $\beta = E\alpha$, then $E\beta = E^2\alpha = E\alpha = \beta$. Conversely, if $\beta = E\beta$, then (of course) β is in the range of E.
2. $V = R \oplus N$.
3. The unique expression for α as a sum of vectors in R and N is $\alpha = E\alpha + (\alpha - E\alpha)$.

From (1), (2), (3) it is easy to see the following. If R and N are subspaces of V such that $V = R \oplus N$, there is one and only one projection operator E which has range R and null space N. That operator is called the **projection on R along N**.

Any projection E is (trivially) diagonalizable. If $\{\alpha_1, \ldots, \alpha_r\}$ is a basis for R and $\{\alpha_{r+1}, \ldots, \alpha_n\}$ a basis for N, then the basis $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ diagonalizes E:

$$[E]_{\mathcal{B}} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$$

where I is the $r \times r$ identity matrix. That should help explain some of the terminology connected with projections. The reader should look at various cases in the plane R^2 (or 3-space, R^3), to convince himself that the projection on R along N sends each vector into R by projecting it parallel to N.

Projections can be used to describe direct-sum decompositions of the space V. For, suppose $V = W_1 \oplus \cdots \oplus W_k$. For each j we shall define an operator E_j on V. Let α be in V, say $\alpha = \alpha_1 + \cdots + \alpha_k$ with α_i in W_i. Define $E_j\alpha = \alpha_j$. Then E_j is a well-defined rule. It is easy to see that E_j is linear, that the range of E_j is W_j, and that $E_j^2 = E_j$. The null space of E_j is the subspace

$$(W_1 + \cdots + W_{j-1} + W_{j+1} + \cdots + W_k)$$

for, the statement that $E_j\alpha = 0$ simply means $\alpha_j = 0$, i.e., that α is actually a sum of vectors from the spaces W_i with $i \neq j$. In terms of the projections E_j we have

$$\alpha = E_1\alpha + \cdots + E_k\alpha \tag{6-13}$$

for each α in V. What (6-13) says is that

$$I = E_1 + \cdots + E_k.$$

Note also that if $i \neq j$, then $E_iE_j = 0$, because the range of E_j is the subspace W_j which is contained in the null space of E_i. We shall now summarize our findings and state and prove a converse.

Theorem 9. *If* $V = W_1 \oplus \cdots \oplus W_k$, *then there exist* k *linear operators* $E_1, \ldots, E_k$ *on* V *such that*

(i) *each* E_i *is a projection* ($E_i^2 = E_i$);
(ii) $E_iE_j = 0$, *if* $i \neq j$;
(iii) $I = E_1 + \cdots + E_k$;
(iv) *the range of* E_i *is* W_i.

Conversely, if $E_1, \ldots, E_k$ *are* k *linear operators on* V *which satisfy conditions* (i), (ii), *and* (iii), *and if we let* W_i *be the range of* E_i, *then* $V = W_1 \oplus \cdots \oplus W_k$.

Proof. We have only to prove the converse statement. Suppose $E_1, \ldots, E_k$ are linear operators on V which satisfy the first three conditions, and let W_i be the range of E_i. Then certainly

$$V = W_1 + \cdots + W_k;$$

for, by condition (iii) we have

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

for each α in V, and $E_i\alpha$ is in W_i. This expression for α is unique, because if

$$\alpha = \alpha_1 + \cdots + \alpha_k$$

with α_i in W_i, say $\alpha_i = E_i\beta_i$, then using (i) and (ii) we have

$$\begin{aligned} E_j\alpha &= \sum_{i=1}^{k} E_j\alpha_i \\ &= \sum_{i=1}^{k} E_jE_i\beta_i \\ &= E_j^2\beta_j \\ &= E_j\beta_j \\ &= \alpha_j. \end{aligned}$$

This shows that V is the direct sum of the W_i. ▌

Exercises

1. Let V be a finite-dimensional vector space and let W_1 be any subspace of V. Prove that there is a subspace W_2 of V such that $V = W_1 \oplus W_2$.

2. Let V be a finite-dimensional vector space and let $W_1, \ldots, W_k$ be subspaces of V such that

$$V = W_1 + \cdots + W_k \quad \text{and} \quad \dim V = \dim W_1 + \cdots + \dim W_k.$$

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

3. Find a projection E which projects R^2 onto the subspace spanned by $(1, -1)$ along the subspace spanned by $(1, 2)$.

4. If E_1 and E_2 are projections onto independent subspaces, then $E_1 + E_2$ is a projection. True or false?

5. If E is a projection and f is a polynomial, then $f(E) = aI + bE$. What are a and b in terms of the coefficients of f?

6. True or false? If a diagonalizable operator has only the characteristic values 0 and 1, it is a projection.

7. Prove that if E is the projection on R along N, then $(I - E)$ is the projection on N along R.

8. Let $E_1, \ldots, E_k$ be linear operators on the space V such that $E_1 + \cdots + E_k = I$.

(a) Prove that if $E_iE_j = 0$ for $i \neq j$, then $E_i^2 = E_i$ for each i.

(b) In the case $k = 2$, prove the converse of (a). That is, if $E_1 + E_2 = I$ and $E_1^2 = E_1$, $E_2^2 = E_2$, then $E_1E_2 = 0$.

9. Let V be a real vector space and E an idempotent linear operator on V, i.e., a projection. Prove that $(I + E)$ is invertible. Find $(I + E)^{-1}$.

10. Let F be a subfield of the complex numbers (or, a field of characteristic zero). Let V be a finite-dimensional vector space over F. Suppose that $E_1, \ldots, E_k$ are projections of V and that $E_1 + \cdots + E_k = I$. Prove that $E_iE_j = 0$ for $i \neq j$ (*Hint:* Use the trace function and ask yourself what the trace of a projection is.)

11. Let V be a vector space, let $W_1, \ldots, W_k$ be subspaces of V, and let

$$V_j = W_1 + \cdots + W_{j-1} + W_{j+1} + \cdots + W_k.$$

Suppose that $V = W_1 \oplus \cdots \oplus W_k$. Prove that the dual space V^* has the direct-sum decomposition $V^* = V_1^0 \oplus \cdots \oplus V_k^0$.

6.7. Invariant Direct Sums

We are primarily interested in direct-sum decompositions $V = W_1 \oplus \cdots \oplus W_k$, where each of the subspaces W_i is invariant under some given linear operator T. Given such a decomposition of V, T induces a linear operator T_i on each W_i by restriction. The action of T is then this.

If α is a vector in V, we have unique vectors $\alpha_1, \ldots, \alpha_k$ with α_i in W_i such that

$$\alpha = \alpha_1 + \cdots + \alpha_k$$

and then

$$T\alpha = T_1\alpha_1 + \cdots + T_k\alpha_k.$$

We shall describe this situation by saying that T is the **direct sum** of the operators $T_1, \ldots, T_k$. It must be remembered in using this terminology that the T_i are not linear operators on the space V but on the various subspaces W_i. The fact that $V = W_1 \oplus \cdots \oplus W_k$ enables us to associate with each α in V a unique k-tuple $(\alpha_1, \ldots, \alpha_k)$ of vectors α_i in W_i (by $\alpha = \alpha_1 + \cdots + \alpha_k$) in such a way that we can carry out the linear operations in V by working in the individual subspaces W_i. The fact that each W_i is invariant under T enables us to view the action of T as the independent action of the operators T_i on the subspaces W_i. Our purpose is to study T by finding invariant direct-sum decompositions in which the T_i are operators of an elementary nature.

Before looking at an example, let us note the matrix analogue of this situation. Suppose we select an ordered basis $\mathcal{B}_i$ for each W_i, and let $\mathcal{B}$ be the ordered basis for V consisting of the union of the $\mathcal{B}_i$ arranged in the order $\mathcal{B}_1, \ldots, \mathcal{B}_k$, so that $\mathcal{B}$ is a basis for V. From our discussion concerning the matrix analogue for a single invariant subspace, it is easy to see that if $A = [T]_{\mathcal{B}}$ and $A_i = [T_i]_{\mathcal{B}_i}$, then A has the block form

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_k \end{bmatrix}. \tag{6-14}$$

In (6-14), A_i is a $d_i \times d_i$ matrix ($d_i = \dim W_i$), and the 0's are symbols for rectangular blocks of scalar 0's of various sizes. It also seems appropriate to describe (6-14) by saying that A is the **direct sum** of the matrices $A_1, \ldots, A_k$.

Most often, we shall describe the subspace W_i by means of the associated projections E_i (Theorem 9). Therefore, we need to be able to phrase the invariance of the subspaces W_i in terms of the E_i.

Theorem 10. *Let* T *be a linear operator on the space* V, *and let* $W_1, \ldots, W_k$ *and* $E_1, \ldots, E_k$ *be as in Theorem 9. Then a necessary and sufficient condition that each subspace* W_i *be invariant under* T *is that* T *commute with each of the projections* E_i, *i.e.,*

$$TE_i = E_iT, \qquad i = 1, \ldots, k.$$

Proof. Suppose T commutes with each E_i. Let α be in W_j. Then $E_j\alpha = \alpha$, and

$$\begin{aligned} T\alpha &= T(E_j\alpha) \\ &= E_j(T\alpha) \end{aligned}$$

which shows that $T\alpha$ is in the range of E_j, i.e., that W_j is invariant under T.

Assume now that each W_i is invariant under T. We shall show that $TE_j = E_jT$. Let α be any vector in V. Then

$$\begin{aligned} \alpha &= E_1\alpha + \cdots + E_k\alpha \\ T\alpha &= TE_1\alpha + \cdots + TE_k\alpha. \end{aligned}$$

Since $E_i\alpha$ is in W_i, which is invariant under T, we must have $T(E_i\alpha) = E_i\beta_i$ for some vector β_i. Then

$$\begin{aligned} E_jTE_i\alpha &= E_jE_i\beta_i \\ &= \begin{cases} 0, & \text{if} \quad i \neq j \\ E_j\beta_j, & \text{if} \quad i = j. \end{cases} \end{aligned}$$

Thus

$$\begin{aligned} E_jT\alpha &= E_jTE_1\alpha + \cdots + E_jTE_k\alpha \\ &= E_j\beta_j \\ &= TE_j\alpha. \end{aligned}$$

This holds for each α in V, so $E_jT = TE_j$. ▮

We shall now describe a diagonalizable operator T in the language of invariant direct sum decompositions (projections which commute with T). This will be a great help to us in understanding some deeper decomposition theorems later. The reader may feel that the description which we are about to give is rather complicated, in comparison to the matrix formulation or to the simple statement that the characteristic vectors of T span the underlying space. But, he should bear in mind that this is our first glimpse at a very effective method, by means of which various problems concerned with subspaces, bases, matrices, and the like can be reduced to algebraic calculations with linear operators. With a little experience, the efficiency and elegance of this method of reasoning should become apparent.

Theorem 11. *Let* T *be a linear operator on a finite-dimensional space* V. *If* T *is diagonalizable and if* $c_1, \ldots, c_k$ *are the distinct characteristic values of* T, *then there exist linear operators* $E_1, \ldots, E_k$ *on* V *such that*

(i) $T = c_1E_1 + \cdots + c_kE_k$;
(ii) $I = E_1 + \cdots + E_k$;
(iii) $E_iE_j = 0,\ i \neq j$;
(iv) $E_i^2 = E_i$ (E_i *is a projection*);
(v) *the range of* E_i *is the characteristic space for* T *associated with* c_i.

Conversely, if there exist k *distinct scalars* $c_1, \ldots, c_k$ *and* k *non-zero linear operators* $E_1, \ldots, E_k$ *which satisfy conditions* (i), (ii), *and* (iii), *then* T *is diagonalizable,* $c_1, \ldots, c_k$ *are the distinct characteristic values of* T, *and conditions* (iv) *and* (v) *are satisfied also.*

Proof. Suppose that T is diagonalizable, with distinct charac-