Chapter 9

Change of Basis

If A is an $m \times n$ matrix, the corresponding **matrix transformation** $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is defined by

$$T_A(\mathbf{x}) = A\mathbf{x} \quad \text{for all columns } \mathbf{x} \text{ in } \mathbb{R}^n$$

It was shown in Theorem 2.6.2 that every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation; that is, $T = T_A$ for some $m \times n$ matrix A. Furthermore, the matrix A is uniquely determined by T. In fact, A is given in terms of its columns by

$$A = \left[\begin{array}{cccc} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{array} \right]$$

where $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$.

In this chapter we show how to associate a matrix with *any* linear transformation $T : V \to W$ where V and W are finite-dimensional vector spaces, and we describe how the matrix can be used to compute $T(\mathbf{v})$ for any $\mathbf{v}$ in V. The matrix depends on the choice of a basis B in V and a basis D in W, and is denoted $M_{DB}(T)$. The case when $W = V$ is particularly important. If B and D are two bases of V, we show that the matrices $M_{BB}(T)$ and $M_{DD}(T)$ are similar, that is $M_{DD}(T) = P^{-1}M_{BB}(T)P$ for some invertible matrix P. Moreover, we give an explicit method for constructing P depending only on the bases B and D. This leads to some of the most important theorems in linear algebra, as we shall see in Chapter 11.

9.1 The Matrix of a Linear Transformation

Let $T : V \to W$ be a linear transformation where $\dim V = n$ and $\dim W = m$. The aim in this section is to describe the action of T as multiplication by an $m \times n$ matrix A. The idea is to convert a vector $\mathbf{v}$ in V into a column in $\mathbb{R}^n$, multiply that column by A to get a column in $\mathbb{R}^m$, and convert this column back to get $T(\mathbf{v})$ in W.

Converting vectors to columns is a simple matter, but one small change is needed. Up to now the *order* of the vectors in a basis has been of no importance. However, in this section, we shall speak of an **ordered basis** $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$, which is just a basis where the order in which the vectors are listed is taken into account. Hence $\{\mathbf{b}_2, \mathbf{b}_1, \mathbf{b}_3\}$ is a different *ordered* basis from $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$.

If $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ is an ordered basis in a vector space V, and if

$$\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n, \quad v_i \in \mathbb{R}$$

is a vector in V, then the (uniquely determined) numbers $v_1, v_2, \dots, v_n$ are called the **coordinates** of $\mathbf{v}$ with respect to the basis B.

Definition 9.1 Coordinate Vector $C_B(\mathbf{v})$ of $\mathbf{v}$ for a basis B

*The **coordinate vector** of $\mathbf{v}$ with respect to B is defined to be*

$$C_B(\mathbf{v}) = (v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n) = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

The reason for writing $C_B(\mathbf{v})$ as a column instead of a row will become clear later. Note that $C_B(\mathbf{b}_i) = \mathbf{e}_i$ is column i of I_n.

Example 9.1.1

The coordinate vector for $\mathbf{v} = (2,\ 1,\ 3)$ with respect to the ordered basis $B = \{(1,\ 1,\ 0),\ (1,\ 0,\ 1),\ (0,\ 1,\ 1)\}$ of $\mathbb{R}^3$ is $C_B(\mathbf{v}) = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ because

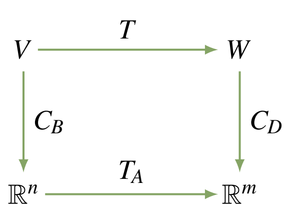
$$\mathbf{v} = (2,\ 1,\ 3) = 0(1,\ 1,\ 0) + 2(1,\ 0,\ 1) + 1(0,\ 1,\ 1)$$

Theorem 9.1.1

If V has dimension n and $B = \{\mathbf{b}_1,\ \mathbf{b}_2,\ \dots,\ \mathbf{b}_n\}$ is any ordered basis of V, the coordinate transformation $C_B : V \to \mathbb{R}^n$ is an isomorphism. In fact, $C_B^{-1} : \mathbb{R}^n \to V$ is given by

$$C_B^{-1}\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n \quad \textit{for all} \quad \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \textit{ in } \mathbb{R}^n.$$

Proof. The verification that C_B is linear is Exercise 9.1.13. If $T : \mathbb{R}^n \to V$ is the map denoted C_B^{-1} in the theorem, one verifies (Exercise 9.1.13) that $TC_B = 1_V$ and $C_BT = 1_{\mathbb{R}^n}$. Note that $C_B(\mathbf{b}_j)$ is column j of the identity matrix, so C_B carries the basis B to the standard basis of $\mathbb{R}^n$, proving again that it is an isomorphism (Theorem 7.3.1) □



Now let $T : V \to W$ be any linear transformation where $\dim V = n$ and $\dim W = m$, and let $B = \{\mathbf{b}_1,\ \mathbf{b}_2,\ \dots,\ \mathbf{b}_n\}$ and D be ordered bases of V and W, respectively. Then $C_B : V \to \mathbb{R}^n$ and $C_D : W \to \mathbb{R}^m$ are isomorphisms and we have the situation shown in the diagram where A is an $m \times n$ matrix (to be determined). In fact, the composite

$$C_DTC_B^{-1} : \mathbb{R}^n \to \mathbb{R}^m \text{ is a linear transformation}$$

so Theorem 2.6.2 shows that a unique $m \times n$ matrix A exists such that

$$C_DTC_B^{-1} = T_A, \qquad \text{equivalently } C_DT = T_AC_B$$

T_A acts by left multiplication by A, so this latter condition is

$$C_D[T(\mathbf{v})] = AC_B(\mathbf{v}) \text{ for all } \mathbf{v} \text{ in } V$$

This requirement completely determines A. Indeed, the fact that $C_B(\mathbf{b}_j)$ is column j of the identity matrix gives

$$\text{column } j \text{ of } A = AC_B(\mathbf{b}_j) = C_D[T(\mathbf{b}_j)]$$

for all j. Hence, in terms of its columns,

$$A = \left[\begin{array}{cccc} C_D[T(\mathbf{b}_1)] & C_D[T(\mathbf{b}_2)] & \cdots & C_D[T(\mathbf{b}_n)] \end{array}\right]$$

Definition 9.2 Matrix $M_{DB}(T)$ of $T : V \to W$ for bases D and B

This is called the ***matrix of*** T ***corresponding to the ordered bases*** B ***and*** D*, and we use the following notation:*

$$M_{DB}(T) = \left[\begin{array}{cccc} C_D[T(\mathbf{b}_1)] & C_D[T(\mathbf{b}_2)] & \cdots & C_D[T(\mathbf{b}_n)] \end{array}\right]$$

This discussion is summarized in the following important theorem.

Theorem 9.1.2

Let $T : V \to W$ be a linear transformation where $\dim V = n$ and $\dim W = m$, and let $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and D be ordered bases of V and W, respectively. Then the matrix $M_{DB}(T)$ just given is the unique $m \times n$ matrix A that satisfies

$$C_D T = T_A C_B$$

Hence the defining property of $M_{DB}(T)$ is

$$C_D[T(\mathbf{v})] = M_{DB}(T)C_B(\mathbf{v}) \text{ for all } \mathbf{v} \text{ in } V$$

The matrix $M_{DB}(T)$ is given in terms of its columns by

$$M_{DB}(T) = \left[\begin{array}{cccc} C_D[T(\mathbf{b}_1)] & C_D[T(\mathbf{b}_2)] & \cdots & C_D[T(\mathbf{b}_n)] \end{array}\right]$$

The fact that $T = C_D^{-1} T_A C_B$ means that the action of T on a vector $\mathbf{v}$ in V can be performed by first taking coordinates (that is, applying C_B to $\mathbf{v}$), then multiplying by A (applying T_A), and finally converting the resulting m-tuple back to a vector in W (applying C_D^{-1}).

Example 9.1.2

Define $T : \mathbf{P}_2 \to \mathbb{R}^2$ by $T(a + bx + cx^2) = (a + c, \ b - a - c)$ for all polynomials $a + bx + cx^2$. If $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ and $D = \{\mathbf{d}_1, \mathbf{d}_2\}$ where

$$\mathbf{b}_1 = 1, \ \mathbf{b}_2 = x, \ \mathbf{b}_3 = x^2 \quad \text{and} \quad \mathbf{d}_1 = (1, \ 0), \ \mathbf{d}_2 = (0, \ 1)$$

compute $M_{DB}(T)$ and verify Theorem 9.1.2.

Solution. We have $T(\mathbf{b}_1) = \mathbf{d}_1 - \mathbf{d}_2$, $T(\mathbf{b}_2) = \mathbf{d}_2$, and $T(\mathbf{b}_3) = \mathbf{d}_1 - \mathbf{d}_2$. Hence

$$M_{DB}(T) = \left[\begin{array}{ccc} C_D[T(\mathbf{b}_1)] & C_D[T(\mathbf{b}_2)] & C_D[T(\mathbf{b}_n)] \end{array}\right] = \left[\begin{array}{rrr} 1 & 0 & 1 \\ -1 & 1 & -1 \end{array}\right]$$

If $\mathbf{v} = a + bx + cx^2 = a\mathbf{b}_1 + b\mathbf{b}_2 + c\mathbf{b}_3$, then $T(\mathbf{v}) = (a+c)\mathbf{d}_1 + (b-a-c)\mathbf{d}_2$, so

$$C_D[T(\mathbf{v})] = \left[\begin{array}{c} a+c \\ b-a-c \end{array}\right] = \left[\begin{array}{rrr} 1 & 0 & 1 \\ -1 & 1 & -1 \end{array}\right] \left[\begin{array}{c} a \\ b \\ c \end{array}\right] = M_{DB}(T)C_B(\mathbf{v})$$

as Theorem 9.1.2 asserts.

The next example shows how to determine the action of a transformation from its matrix.

Example 9.1.3

Suppose $T : \mathbf{M}_{22}(\mathbb{R}) \to \mathbb{R}^3$ is linear with matrix $M_{DB}(T) = \left[\begin{array}{rrrr} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right]$ where

$$B = \left\{ \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array}\right] \right\} \text{ and } D = \{(1,\ 0,\ 0),\ (0,\ 1,\ 0),\ (0,\ 0,\ 1)\}$$

Compute $T(\mathbf{v})$ where $\mathbf{v} = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$.

Solution. The idea is to compute $C_D[T(\mathbf{v})]$ first, and then obtain $T(\mathbf{v})$. We have

$$C_D[T(\mathbf{v})] = M_{DB}(T)C_B(\mathbf{v}) = \left[\begin{array}{rrrr} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \left[\begin{array}{c} a \\ b \\ c \\ d \end{array}\right] = \left[\begin{array}{c} a-b \\ b-c \\ c-d \end{array}\right]$$

$$\begin{aligned} \text{Hence } T(\mathbf{v}) &= (a-b)(1,\ 0,\ 0) + (b-c)(0,\ 1,\ 0) + (c-d)(0,\ 0,\ 1) \\ &= (a-b,\ b-c,\ c-d) \end{aligned}$$

The next two examples will be referred to later.

Example 9.1.4

Let A be an $m \times n$ matrix, and let $T_A : \mathbb{R}^n \to \mathbb{R}^m$ be the matrix transformation induced by $A : T_A(\mathbf{x}) = A\mathbf{x}$ for all columns $\mathbf{x}$ in $\mathbb{R}^n$. If B and D are the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively (ordered as usual), then

$$M_{DB}(T_A) = A$$

In other words, the matrix of T_A corresponding to the standard bases is A itself.

Solution. Write $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$. Because D is the standard basis of $\mathbb{R}^m$, it is easy to verify that $C_D(\mathbf{y}) = \mathbf{y}$ for all columns $\mathbf{y}$ in $\mathbb{R}^m$. Hence

$$M_{DB}(T_A) = \left[\begin{array}{cccc} T_A(\mathbf{e}_1) & T_A(\mathbf{e}_2) & \cdots & T_A(\mathbf{e}_n) \end{array}\right] = \left[\begin{array}{cccc} A\mathbf{e}_1 & A\mathbf{e}_2 & \cdots & A\mathbf{e}_n \end{array}\right] = A$$

because $A\mathbf{e}_j$ is the jth column of A.

Example 9.1.5

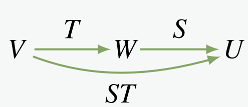
Let V and W have ordered bases B and D, respectively. Let $\dim V = n$.

1. The identity transformation $1_V : V \to V$ has matrix $M_{BB}(1_V) = I_n$.
2. The zero transformation $0 : V \to W$ has matrix $M_{DB}(0) = 0$.

The first result in Example 9.1.5 is false if the two bases of V are not equal. In fact, if B is the standard basis of $\mathbb{R}^n$, then the basis D of $\mathbb{R}^n$ can be chosen so that $M_{DB}(1_{\mathbb{R}^n})$ turns out to be any invertible matrix we wish (Exercise 9.1.14).

The next two theorems show that composition of linear transformations is compatible with multiplication of the corresponding matrices.

Theorem 9.1.3



Let $V \xrightarrow{T} W \xrightarrow{S} U$ be linear transformations and let B, D, and E be finite ordered bases of V, W, and U, respectively. Then

$$M_{EB}(ST) = M_{ED}(S) \cdot M_{DB}(T)$$

Proof. We use the property in Theorem 9.1.2 three times. If $\mathbf{v}$ is in V,

$$M_{ED}(S)M_{DB}(T)C_B(\mathbf{v}) = M_{ED}(S)C_D[T(\mathbf{v})] = C_E[ST(\mathbf{v})] = M_{EB}(ST)C_B(\mathbf{v})$$

If $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, then $C_B(\mathbf{e}_j)$ is column j of I_n. Hence taking $\mathbf{v} = \mathbf{e}_j$ shows that $M_{ED}(S)M_{DB}(T)$ and $M_{EB}(ST)$ have equal jth columns. The theorem follows. □

Theorem 9.1.4

Let $T : V \to W$ be a linear transformation, where $\dim V = \dim W = n$. The following are equivalent.

1. *T is an isomorphism.*
2. *$M_{DB}(T)$ is invertible for all ordered bases B and D of V and W.*
3. *$M_{DB}(T)$ is invertible for some pair of ordered bases B and D of V and W.*

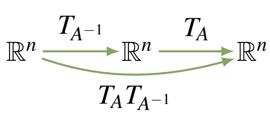
When this is the case, $[M_{DB}(T)]^{-1} = M_{BD}(T^{-1})$.

Proof. (1) $\Rightarrow$ (2). We have $V \xrightarrow{T} W \xrightarrow{T^{-1}} V$, so Theorem 9.1.3 and Example 9.1.5 give

$$M_{BD}(T^{-1})M_{DB}(T) = M_{BB}(T^{-1}T) = M_{BB}(1v) = I_n$$

Similarly, $M_{DB}(T)M_{BD}(T^{-1}) = I_n$, proving (2) (and the last statement in the theorem).

(2) $\Rightarrow$ (3). This is clear.



(3) $\Rightarrow$ (1). Suppose that $T_{DB}(T)$ is invertible for some bases B and D and, for convenience, write $A = M_{DB}(T)$. Then we have $C_D T = T_A C_B$ by Theorem 9.1.2, so

$$T = (C_D)^{-1} T_A C_B$$

by Theorem 9.1.1 where $(C_D)^{-1}$ and C_B are isomorphisms. Hence (1) follows if we can demonstrate that $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is also an isomorphism. But A is invertible by (3) and one verifies that $T_A T_{A^{-1}} = 1_{\mathbb{R}^n} = T_{A^{-1}} T_A$. So T_A is indeed invertible (and $(T_A)^{-1} = T_{A^{-1}}$). □

In Section 7.2 we defined the rank of a linear transformation $T : V \to W$ by $\operatorname{rank} T = \dim(\operatorname{im} T)$. Moreover, if A is any $m \times n$ matrix and $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is the matrix transformation, we showed that $\operatorname{rank}(T_A) = \operatorname{rank} A$. So it may not be surprising that $\operatorname{rank} T$ equals the rank of any matrix of T.

Theorem 9.1.5

Let $T : V \to W$ be a linear transformation where $\dim V = n$ and $\dim W = m$. If B and D are any ordered bases of V and W, then $\operatorname{rank} T = \operatorname{rank}[M_{DB}(T)]$.

Proof. Write $A = M_{DB}(T)$ for convenience. The column space of A is $U = \{A\mathbf{x} \mid \mathbf{x} \text{ in } \mathbb{R}^n\}$. This means $\operatorname{rank} A = \dim U$ and so, because $\operatorname{rank} T = \dim(\operatorname{im} T)$, it suffices to find an isomorphism $S : \operatorname{im} T \to U$. Now every vector in $\operatorname{im} T$ has the form $T(\mathbf{v})$, $\mathbf{v}$ in V. By Theorem 9.1.2, $C_D[T(\mathbf{v})] = AC_B(\mathbf{v})$ lies in U. So define $S : \operatorname{im} T \to U$ by

$$S[T(\mathbf{v})] = C_D[T(\mathbf{v})] \text{ for all vectors } T(\mathbf{v}) \in \operatorname{im} T$$

The fact that C_D is linear and one-to-one implies immediately that S is linear and one-to-one. To see that S is onto, let $A\mathbf{x}$ be any member of U, $\mathbf{x}$ in $\mathbb{R}^n$. Then $\mathbf{x} = C_B(\mathbf{v})$ for some $\mathbf{v}$ in V because C_B is onto. Hence $A\mathbf{x} = AC_B(\mathbf{v}) = C_D[T(\mathbf{v})] = S[T(\mathbf{v})]$, so S is onto. This means that S is an isomorphism. □

Example 9.1.6

Define $T : \mathbf{P}_2 \to \mathbb{R}^3$ by $T(a+bx+cx^2) = (a-2b,\ 3c-2a,\ 3c-4b)$ for $a, b, c \in \mathbb{R}$. Compute rank T.

Solution. Since rank T = rank $[M_{DB}(T)]$ for *any* bases $B \subseteq \mathbf{P}_2$ and $D \subseteq \mathbb{R}^3$, we choose the most convenient ones: $B = \{1,\ x,\ x^2\}$ and $D = \{(1,\ 0,\ 0),\ (0,\ 1,\ 0),\ (0,\ 0,\ 1)\}$. Then $M_{DB}(T) = \left[\ C_D[T(1)]\quad C_D[T(x)]\quad C_D[T(x^2)]\ \right] = A$ where

$$A = \left[\begin{array}{rrr} 1 & -2 & 0 \\ -2 & 0 & 3 \\ 0 & -4 & 3 \end{array}\right]. \quad \text{Since } A \to \left[\begin{array}{rrr} 1 & -2 & 0 \\ 0 & -4 & 3 \\ 0 & -4 & 3 \end{array}\right] \to \left[\begin{array}{rrr} 1 & -2 & 0 \\ 0 & 1 & -\frac{3}{4} \\ 0 & 0 & 0 \end{array}\right]$$

we have rank $A = 2$. Hence rank $T = 2$ as well.

We conclude with an example showing that the matrix of a linear transformation can be made very simple by a careful choice of the two bases.

Example 9.1.7

Let $T : V \to W$ be a linear transformation where dim $V = n$ and dim $W = m$. Choose an ordered basis $B = \{\mathbf{b}_1,\ \dots,\ \mathbf{b}_r,\ \mathbf{b}_{r+1},\ \dots,\ \mathbf{b}_n\}$ of V in which $\{\mathbf{b}_{r+1},\ \dots,\ \mathbf{b}_n\}$ is a basis of ker T, possibly empty. Then $\{T(\mathbf{b}_1),\ \dots,\ T(\mathbf{b}_r)\}$ is a basis of im T by Theorem 7.2.5, so extend it to an ordered basis $D = \{T(\mathbf{b}_1),\ \dots,\ T(\mathbf{b}_r),\ \mathbf{f}_{r+1},\ \dots,\ \mathbf{f}_m\}$ of W. Because $T(\mathbf{b}_{r+1}) = \cdots = T(\mathbf{b}_n) = \mathbf{0}$, we have

$$M_{DB}(T) = \left[\ C_D[T(\mathbf{b}_1)]\ \cdots\ C_D[T(\mathbf{b}_r)]\ \ C_D[T(\mathbf{b}_{r+1})]\ \cdots\ C_D[T(\mathbf{b}_n)]\ \right] = \left[\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array}\right]$$

Incidentally, this shows that rank $T = r$ by Theorem 9.1.5.

Exercises for 9.1

Exercise 9.1.1 In each case, find the coordinates of $\mathbf{v}$ with respect to the basis B of the vector space V.

a. $V = \mathbf{P}_2$, $\mathbf{v} = 2x^2 + x - 1$, $B = \{x+1,\ x^2,\ 3\}$

b. $V = \mathbf{P}_2$, $\mathbf{v} = ax^2 + bx + c$, $B = \{x^2,\ x+1,\ x+2\}$

c. $V = \mathbb{R}^3$, $\mathbf{v} = (1,\ -1,\ 2)$,
$B = \{(1,\ -1,\ 0),\ (1,\ 1,\ 1),\ (0,\ 1,\ 1)\}$

d. $V = \mathbb{R}^3$, $\mathbf{v} = (a,\ b,\ c)$,
$B = \{(1,\ -1,\ 2),\ (1,\ 1,\ -1),\ (0,\ 0,\ 1)\}$

e. $V = \mathbf{M}_{22}$, $\mathbf{v} = \left[\begin{array}{rr} 1 & 2 \\ -1 & 0 \end{array}\right]$,
$B = \left\{ \left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right], \left[\begin{array}{cc} 1 & 0 \\ 1 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 0 \\ 1 & 1 \end{array}\right], \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] \right\}$

Exercise 9.1.2 Suppose $T : \mathbf{P}_2 \to \mathbb{R}^2$ is a linear transformation. If $B = \{1,\ x,\ x^2\}$ and $D = \{(1,\ 1),\ (0,\ 1)\}$, find the action of T given:

a. $M_{DB}(T) = \left[\begin{array}{rrr} 1 & 2 & -1 \\ -1 & 0 & 1 \end{array}\right]$

b. $M_{DB}(T) = \left[\begin{array}{rrr} 2 & 1 & 3 \\ -1 & 0 & -2 \end{array}\right]$

Exercise 9.1.3 In each case, find the matrix of the linear transformation $T : V \to W$ corresponding to the bases B and D of V and W, respectively.

a. $T: \mathbf{M}_{22} \to \mathbb{R}$, $T(A) = \text{tr } A$;
$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$,
$D = \{1\}$

b. $T: \mathbf{M}_{22} \to \mathbf{M}_{22}$, $T(A) = A^T$;
$B = D = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

c. $T: \mathbf{P}_2 \to \mathbf{P}_3$, $T[p(x)] = xp(x)$; $B = \{1, x, x^2\}$ and $D = \{1, x, x^2, x^3\}$

d. $T: \mathbf{P}_2 \to \mathbf{P}_2$, $T[p(x)] = p(x+1)$;
$B = D = \{1, x, x^2\}$

Exercise 9.1.4 In each case, find the matrix of $T: V \to W$ corresponding to the bases B and D, respectively, and use it to compute $C_D[T(\mathbf{v})]$, and hence $T(\mathbf{v})$.

a. $T: \mathbb{R}^3 \to \mathbb{R}^4$, $T(x, y, z) = (x+z, 2z, y-z, x+2y)$; B and D standard; $\mathbf{v} = (1, -1, 3)$

b. $T: \mathbb{R}^2 \to \mathbb{R}^4$, $T(x, y) = (2x-y, 3x+2y, 4y, x)$; $B = \{(1, 1), (1, 0)\}$, D standard; $\mathbf{v} = (a, b)$

c. $T: \mathbf{P}_2 \to \mathbb{R}^2$, $T(a+bx+cx^2) = (a+c, 2b)$;
$B = \{1, x, x^2\}$, $D = \{(1, 0), (1, -1)\}$;
$\mathbf{v} = a+bx+cx^2$

d. $T: \mathbf{P}_2 \to \mathbb{R}^2$, $T(a+bx+cx^2) = (a+b, c)$;
$B = \{1, x, x^2\}$, $D = \{(1, -1), (1, 1)\}$;
$\mathbf{v} = a+bx+cx^2$

e. $T: \mathbf{M}_{22} \to \mathbb{R}$, $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a+b+c+d$;
$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$,
$D = \{1\}$; $\mathbf{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

f. $T: \mathbf{M}_{22} \to \mathbf{M}_{22}$,
$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b+c \\ b+c & d \end{bmatrix}$;
$B = D =$
$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$;
$\mathbf{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Exercise 9.1.5 In each case, verify Theorem 9.1.3. Use the standard basis in $\mathbb{R}^n$ and $\{1, x, x^2\}$ in $\mathbf{P}_2$.

a. $\mathbb{R}^3 \xrightarrow{T} \mathbb{R}^2 \xrightarrow{S} \mathbb{R}^4$; $T(a, b, c) = (a+b, b-c)$, $S(a, b) = (a, b-2a, 3b, a+b)$

b. $\mathbb{R}^3 \xrightarrow{T} \mathbb{R}^4 \xrightarrow{S} \mathbb{R}^2$;
$T(a, b, c) = (a+b, c+b, a+c, b-a)$,
$S(a, b, c, d) = (a+b, c-d)$

c. $\mathbf{P}_2 \xrightarrow{T} \mathbb{R}^3 \xrightarrow{S} \mathbf{P}_2$; $T(a+bx+cx^2) = (a, b-c, c-a)$,
$S(a, b, c) = b+cx+(a-c)x^2$

d. $\mathbb{R}^3 \xrightarrow{T} \mathbf{P}_2 \xrightarrow{S} \mathbb{R}^2$;
$T(a, b, c) = (a-b)+(c-a)x+bx^2$,
$S(a+bx+cx^2) = (a-b, c)$

Exercise 9.1.6 Verify Theorem 9.1.3 for $\mathbf{M}_{22} \xrightarrow{T} \mathbf{M}_{22} \xrightarrow{S} \mathbf{P}_2$ where $T(A) = A^T$ and $S\begin{bmatrix} a & b \\ c & d \end{bmatrix} = b+(a+d)x+cx^2$. Use the bases
$B = D = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$
and $E = \{1, x, x^2\}$.

Exercise 9.1.7 In each case, find T^{-1} and verify that $[M_{DB}(T)]^{-1} = M_{BD}(T^{-1})$.

a. $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T(a, b) = (a+2b, 2a+5b)$;
$B = D =$ standard

b. $T: \mathbb{R}^3 \to \mathbb{R}^3$, $T(a, b, c) = (b+c, a+c, a+b)$;
$B = D =$ standard

c. $T: \mathbf{P}_2 \to \mathbb{R}^3$, $T(a+bx+cx^2) = (a-c, b, 2a-c)$;
$B = \{1, x, x^2\}$, $D =$ standard

d. $T: \mathbf{P}_2 \to \mathbb{R}^3$,
$T(a+bx+cx^2) = (a+b+c, b+c, c)$;
$B = \{1, x, x^2\}$, $D =$ standard

Exercise 9.1.8 In each case, show that $M_{DB}(T)$ is invertible and use the fact that $M_{BD}(T^{-1}) = [M_{BD}(T)]^{-1}$ to determine the action of T^{-1}.

a. $T: \mathbf{P}_2 \to \mathbb{R}^3$, $T(a+bx+cx^2) = (a+c, c, b-c)$;
$B = \{1, x, x^2\}$, $D =$ standard

b. $T: \mathbf{M}_{22} \to \mathbb{R}^4$,
$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = (a+b+c, b+c, c, d)$;
$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$,
$D =$ standard

Exercise 9.1.9 Let $D : \mathbf{P}_3 \to \mathbf{P}_2$ be the differentiation map given by $D[p(x)] = p'(x)$. Find the matrix of D corresponding to the bases $B = \{1, x, x^2, x^3\}$ and $E = \{1, x, x^2\}$, and use it to compute $D(a + bx + cx^2 + dx^3)$.

Exercise 9.1.10 Use Theorem 9.1.4 to show that $T : V \to V$ is not an isomorphism if $\ker T \neq 0$ (assume $\dim V = n$). [*Hint*: Choose any ordered basis B containing a vector in $\ker T$.]

Exercise 9.1.11 Let $T : V \to \mathbb{R}$ be a linear transformation, and let $D = \{1\}$ be the basis of $\mathbb{R}$. Given any ordered basis $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ of V, show that $M_{DB}(T) = [T(\mathbf{e}_1) \cdots T(\mathbf{e}_n)]$.

Exercise 9.1.12 Let $T : V \to W$ be an isomorphism, let $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be an ordered basis of V, and let $D = \{T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)\}$. Show that $M_{DB}(T) = I_n$—the $n \times n$ identity matrix.

Exercise 9.1.13 Complete the proof of Theorem 9.1.1.

Exercise 9.1.14 Let U be any invertible $n \times n$ matrix, and let $D = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ where $\mathbf{f}_j$ is column j of U. Show that $M_{BD}(1_{\mathbb{R}^n}) = U$ when B is the standard basis of $\mathbb{R}^n$.

Exercise 9.1.15 Let B be an ordered basis of the n-dimensional space V and let $C_B : V \to \mathbb{R}^n$ be the coordinate transformation. If D is the standard basis of $\mathbb{R}^n$, show that $M_{DB}(C_B) = I_n$.

Exercise 9.1.16 Let $T : \mathbf{P}_2 \to \mathbb{R}^3$ be defined by $T(p) = (p(0), p(1), p(2))$ for all p in $\mathbf{P}_2$. Let $B = \{1, x, x^2\}$ and $D = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

a. Show that $M_{DB}(T) = \left[\begin{array}{rrr} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{array} \right]$ and conclude that T is an isomorphism.

b. Generalize to $T : \mathbf{P}_n \to \mathbb{R}^{n+1}$ where $T(p) = (p(a_0), p(a_1), \dots, p(a_n))$ and $a_0, a_1, \dots, a_n$ are distinct real numbers. [*Hint*: Theorem 3.2.7.]

Exercise 9.1.17 Let $T : \mathbf{P}_n \to \mathbf{P}_n$ be defined by $T[p(x)] = p(x) + xp'(x)$, where $p'(x)$ denotes the derivative. Show that T is an isomorphism by finding $M_{BB}(T)$ when $B = \{1, x, x^2, \dots, x^n\}$.

Exercise 9.1.18 If k is any number, define $T_k : \mathbf{M}_{22} \to \mathbf{M}_{22}$ by $T_k(A) = A + kA^T$.

a. If $B = \left\{ \left[\begin{array}{rr} 1 & 0 \\ 0 & 0 \end{array} \right], \left[\begin{array}{rr} 0 & 0 \\ 0 & 1 \end{array} \right], \left[\begin{array}{rr} 0 & 1 \\ 1 & 0 \end{array} \right], \left[\begin{array}{rr} 0 & 1 \\ -1 & 0 \end{array} \right] \right\}$ find $M_{BB}(T_k)$, and conclude that T_k is invertible if $k \neq 1$ and $k \neq -1$.

b. Repeat for $T_k : \mathbf{M}_{33} \to \mathbf{M}_{33}$. Can you generalize?

The remaining exercises require the following definitions. If V and W are vector spaces, the set of all linear transformations from V to W will be denoted by

$$\mathbf{L}(V, W) = \{T \mid T : V \to W \text{ is a linear transformation }\}$$

Given S and T in $\mathbf{L}(V, W)$ and a in $\mathbb{R}$, define $S + T : V \to W$ and $aT : V \to W$ by

$$\begin{aligned} (S + T)(\mathbf{v}) &= S(\mathbf{v}) + T(\mathbf{v}) && \text{for all } \mathbf{v} \text{ in } V \\ (aT)(\mathbf{v}) &= aT(\mathbf{v}) && \text{for all } \mathbf{v} \text{ in } V \end{aligned}$$

Exercise 9.1.19 Show that $\mathbf{L}(V, W)$ is a vector space.

Exercise 9.1.20 Show that the following properties hold provided that the transformations link together in such a way that all the operations are defined.

a. $R(ST) = (RS)T$

b. $1_W T = T = T 1_V$

c. $R(S + T) = RS + RT$

d. $(S + T)R = SR + TR$

e. $(aS)T = a(ST) = S(aT)$

Exercise 9.1.21 Given S and T in $\mathbf{L}(V, W)$, show that:

a. $\ker S \cap \ker T \subseteq \ker (S + T)$

b. $\operatorname{im} (S + T) \subseteq \operatorname{im} S + \operatorname{im} T$

Exercise 9.1.22 Let V and W be vector spaces. If X is a subset of V, define

$$X^0 = \{T \text{ in } \mathbf{L}(V, W) \mid T(\mathbf{v}) = 0 \text{ for all } \mathbf{v} \text{ in } X\}$$

a. Show that X^0 is a subspace of $\mathbf{L}(V, W)$.

b. If $X \subseteq X_1$, show that $X_1^0 \subseteq X^0$.

c. If U and U_1 are subspaces of V, show that $(U + U_1)^0 = U^0 \cap U_1^0$.

Exercise 9.1.23 Define $R : \mathbf{M}_{mn} \to \mathbf{L}(\mathbb{R}^n, \mathbb{R}^m)$ by $R(A) = T_A$ for each $m \times n$ matrix A, where $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is given by $T_A(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Show that R is an isomorphism.

Exercise 9.1.24 Let V be any vector space (we do not assume it is finite dimensional). Given $\mathbf{v}$ in V, define $S_\mathbf{v} : \mathbb{R} \to V$ by $S_\mathbf{v}(r) = r\mathbf{v}$ for all r in $\mathbb{R}$.

a. Show that $S_\mathbf{v}$ lies in $\mathbf{L}(\mathbb{R}, V)$ for each $\mathbf{v}$ in V.

b. Show that the map $R : V \to \mathbf{L}(\mathbb{R}, V)$ given by $R(\mathbf{v}) = S_\mathbf{v}$ is an isomorphism. [*Hint*: To show that R is onto, if T lies in $\mathbf{L}(\mathbb{R}, V)$, show that $T = S_\mathbf{v}$ where $\mathbf{v} = T(1)$.]

Exercise 9.1.25 Let V be a vector space with ordered basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$. For each $i = 1, 2, \dots, m$, define $S_i : \mathbb{R} \to V$ by $S_i(r) = r\mathbf{b}_i$ for all r in $\mathbb{R}$.

a. Show that each S_i lies in $\mathbf{L}(\mathbb{R}, V)$ and $S_i(1) = \mathbf{b}_i$.

b. Given T in $\mathbf{L}(\mathbb{R}, V)$, let $T(1) = a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \cdots + a_n\mathbf{b}_n$, a_i in $\mathbb{R}$. Show that $T = a_1S_1 + a_2S_2 + \cdots + a_nS_n$.

c. Show that $\{S_1, S_2, \dots, S_n\}$ is a basis of $\mathbf{L}(\mathbb{R}, V)$.

Exercise 9.1.26 Let $\dim V = n$, $\dim W = m$, and let B and D be ordered bases of V and W, respectively. Show that $M_{DB} : \mathbf{L}(V, W) \to \mathbf{M}_{mn}$ is an isomorphism of vector spaces. [*Hint*: Let $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $D = \{\mathbf{d}_1, \dots, \mathbf{d}_m\}$. Given $A = [a_{ij}]$ in $\mathbf{M}_{mn}$, show that $A = M_{DB}(T)$ where $T : V \to W$ is defined by $T(\mathbf{b}_j) = a_{1j}\mathbf{d}_1 + a_{2j}\mathbf{d}_2 + \cdots + a_{mj}\mathbf{d}_m$ for each j.]

Exercise 9.1.27 If V is a vector space, the space $V^* = \mathbf{L}(V, \mathbb{R})$ is called the **dual** of V. Given a basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ of V, let $E_i : V \to \mathbb{R}$ for each $i = 1, 2, \dots, n$ be the linear transformation satisfying

$$E_i(\mathbf{b}_j) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

(each E_i exists by Theorem 7.1.3). Prove the following:

a. $E_i(r_1\mathbf{b}_1 + \cdots + r_n\mathbf{b}_n) = r_i$ for each $i = 1, 2, \dots, n$

b. $\mathbf{v} = E_1(\mathbf{v})\mathbf{b}_1 + E_2(\mathbf{v})\mathbf{b}_2 + \cdots + E_n(\mathbf{v})\mathbf{b}_n$ for all $\mathbf{v}$ in V

c. $T = T(\mathbf{b}_1)E_1 + T(\mathbf{b}_2)E_2 + \cdots + T(\mathbf{b}_n)E_n$ for all T in V^*

d. $\{E_1, E_2, \dots, E_n\}$ is a basis of V^* (called the **dual basis** of B).

Given $\mathbf{v}$ in V, define $\mathbf{v}^* : V \to \mathbb{R}$ by $\mathbf{v}^*(\mathbf{w}) = E_1(\mathbf{v})E_1(\mathbf{w}) + E_2(\mathbf{v})E_2(\mathbf{w}) + \cdots + E_n(\mathbf{v})E_n(\mathbf{w})$ for all $\mathbf{w}$ in V. Show that:

e. $\mathbf{v}^* : V \to \mathbb{R}$ is linear, so $\mathbf{v}^*$ lies in V^*.

f. $\mathbf{b}_i^* = E_i$ for each $i = 1, 2, \dots, n$.

g. The map $R : V \to V^*$ with $R(\mathbf{v}) = \mathbf{v}^*$ is an isomorphism. [*Hint*: Show that R is linear and one-to-one and use Theorem 7.3.3. Alternatively, show that $R^{-1}(T) = T(\mathbf{b}_1)\mathbf{b}_1 + \cdots + T(\mathbf{b}_n)\mathbf{b}_n$.]

9.2 Operators and Similarity

While the study of linear transformations from one vector space to another is important, the central problem of linear algebra is to understand the structure of a linear transformation $T : V \to V$ from a space V to itself. Such transformations are called **linear operators**. If $T : V \to V$ is a linear operator where $\dim(V) = n$, it is possible to choose bases B and D of V such that the matrix $M_{DB}(T)$ has a very simple form: $M_{DB}(T) = \left[\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array} \right]$ where $r = \operatorname{rank} T$ (see Example 9.1.7). Consequently, only the rank of T is revealed by determining the simplest matrices $M_{DB}(T)$ of T where the bases B and D can be chosen arbitrarily. But if we insist that $B = D$ and look for bases B such that $M_{BB}(T)$ is as simple as possible, we learn a great deal about the operator T. We begin this task in this section.

The B-matrix of an Operator

Definition 9.3 Matrix $M_{DB}(T)$ of $T: V \to W$ for basis B

*If $T: V \to V$ is an operator on a vector space V, and if B is an ordered basis of V, define $M_B(T) = M_{BB}(T)$ and call this the **B-matrix** of T.*

Recall that if $T: \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator and $E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$, then $C_E(\mathbf{x}) = \mathbf{x}$ for every $\mathbf{x} \in \mathbb{R}^n$, so $M_E(T) = [T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_n)]$ is the matrix obtained in Theorem 2.6.2. Hence $M_E(T)$ will be called the **standard matrix** of the operator T.

For reference the following theorem collects some results from Theorem 9.1.2, Theorem 9.1.3, and Theorem 9.1.4, specialized for operators. As before, $C_B(\mathbf{v})$ denoted the coordinate vector of $\mathbf{v}$ with respect to the basis B.

Theorem 9.2.1

Let $T: V \to V$ be an operator where $\dim V = n$, and let B be an ordered basis of V.

1. $C_B(T(\mathbf{v})) = M_B(T)C_B(\mathbf{v})$ *for all* $\mathbf{v}$ *in* V.
2. *If* $S: V \to V$ *is another operator on* V, *then* $M_B(ST) = M_B(S)M_B(T)$.
3. T *is an isomorphism if and only if* $M_B(T)$ *is invertible. In this case* $M_D(T)$ *is invertible for every ordered basis* D *of* V.
4. *If* T *is an isomorphism, then* $M_B(T^{-1}) = [M_B(T)]^{-1}$.
5. *If* $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$, *then* $M_B(T) = \left[\begin{array}{cccc} C_B[T(\mathbf{b}_1)] & C_B[T(\mathbf{b}_2)] & \cdots & C_B[T(\mathbf{b}_n)] \end{array}\right]$.

For a fixed operator T on a vector space V, we are going to study how the matrix $M_B(T)$ changes when the basis B changes. This turns out to be closely related to how the coordinates $C_B(\mathbf{v})$ change for a vector $\mathbf{v}$ in V. If B and D are two ordered bases of V, and if we take $T = 1_V$ in Theorem 9.1.2, we obtain

$$C_D(\mathbf{v}) = M_{DB}(1_V)C_B(\mathbf{v}) \quad \text{for all } \mathbf{v} \text{ in } V$$

Definition 9.4 Change Matrix $P_{D \leftarrow B}$ for bases B and D

*With this in mind, define the **change matrix** $P_{D \leftarrow B}$ by*

$$P_{D \leftarrow B} = M_{DB}(1_V) \quad \text{for any ordered bases } B \text{ and } D \text{ of } V$$

This proves equation 9.2 in the following theorem:

Theorem 9.2.2

Let $B = \{\mathbf{b}_1,\ \mathbf{b}_2,\ \dots,\ \mathbf{b}_n\}$ and D denote ordered bases of a vector space V. Then the change matrix $P_{D \leftarrow B}$ is given in terms of its columns by

$$P_{D \leftarrow B} = \left[\begin{array}{cccc} C_D(\mathbf{b}_1) & C_D(\mathbf{b}_2) & \cdots & C_D(\mathbf{b}_n) \end{array}\right] \tag{9.1}$$

and has the property that

$$C_D(\mathbf{v}) = P_{D \leftarrow B} C_B(\mathbf{v}) \text{ for all } \mathbf{v} \text{ in } V \tag{9.2}$$

Moreover, if E is another ordered basis of V, we have

1. $P_{B \leftarrow B} = I_n$
2. $P_{D \leftarrow B}$ *is invertible and* $(P_{D \leftarrow B})^{-1} = P_{B \leftarrow D}$
3. $P_{E \leftarrow D} P_{D \leftarrow B} = P_{E \leftarrow B}$

Proof. The formula 9.2 is derived above, and 9.1 is immediate from the definition of $P_{D \leftarrow B}$ and the formula for $M_{DB}(T)$ in Theorem 9.1.2.

1. $P_{B \leftarrow B} = M_{BB}(1_V) = I_n$ as is easily verified.
2. This follows from (1) and (3).
3. Let $V \xrightarrow{T} W \xrightarrow{S} U$ be operators, and let B, D, and E be ordered bases of V, W, and U respectively. We have $M_{EB}(ST) = M_{ED}(S) M_{DB}(T)$ by Theorem 9.1.3. Now (3) is the result of specializing $V = W = U$ and $T = S = 1_V$. □

Property (3) in Theorem 9.2.2 explains the notation $\mathbf{P}_{D \leftarrow B}$.

Example 9.2.1

In $\mathbf{P}_2$ find $P_{D \leftarrow B}$ if $B = \{1,\ x,\ x^2\}$ and $D = \{1,\ (1-x),\ (1-x)^2\}$. Then use this to express $p = p(x) = a + bx + cx^2$ as a polynomial in powers of $(1-x)$.

Solution. To compute the change matrix $P_{D \leftarrow B}$, express $1,\ x,\ x^2$ in the basis D:

$$\begin{aligned} 1 &= 1 + 0(1-x) + 0(1-x)^2 \\ x &= 1 - 1(1-x) + 0(1-x)^2 \\ x^2 &= 1 - 2(1-x) + 1(1-x)^2 \end{aligned}$$

Hence $P_{D\leftarrow B} = \left[C_D(1),\ C_D(x),\ C_D(x)^2\right] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$. We have $C_B(p) = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, so

$$C_D(p) = P_{D\leftarrow B}C_B(p) = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a+b+c \\ -b-2c \\ c \end{bmatrix}$$

Hence $p(x) = (a+b+c)-(b+2c)(1-x)+c(1-x)^2$ by Definition 9.1.[1]

Now let $B = \{\mathbf{b}_1,\ \mathbf{b}_2,\ \dots,\ \mathbf{b}_n\}$ and B_0 be two ordered bases of a vector space V. An operator $T : V \to V$ has different matrices $M_B[T]$ and $M_{B_0}[T]$ with respect to B and B_0. We can now determine how these matrices are related. Theorem 9.2.2 asserts that

$$C_{B_0}(\mathbf{v}) = P_{B_0\leftarrow B}C_B(\mathbf{v}) \text{ for all } \mathbf{v} \text{ in } V$$

On the other hand, Theorem 9.2.1 gives

$$C_B[T(\mathbf{v})] = M_B(T)C_B(\mathbf{v}) \text{ for all } \mathbf{v} \text{ in } V$$

Combining these (and writing $P = P_{B_0\leftarrow B}$ for convenience) gives

$$\begin{aligned} PM_B(T)C_B(\mathbf{v}) &= PC_B[T(\mathbf{v})] \\ &= C_{B_0}[T(\mathbf{v})] \\ &= M_{B_0}(T)C_{B_0}(\mathbf{v}) \\ &= M_{B_0}(T)PC_B(\mathbf{v}) \end{aligned}$$

This holds for all $\mathbf{v}$ in V. Because $C_B(\mathbf{b}_j)$ is the jth column of the identity matrix, it follows that

$$PM_B(T) = M_{B_0}(T)P$$

Moreover P is invertible (in fact, $P^{-1} = P_{B\leftarrow B_0}$ by Theorem 9.2.2), so this gives

$$M_B(T) = P^{-1}M_{B_0}(T)P$$

This asserts that $M_{B_0}(T)$ and $M_B(T)$ are similar matrices, and proves Theorem 9.2.3.

Theorem 9.2.3: Similarity Theorem

Let B_0 and B be two ordered bases of a finite dimensional vector space V. If $T : V \to V$ is any linear operator, the matrices $M_B(T)$ and $M_{B_0}(T)$ of T with respect to these bases are similar. More precisely,

$$M_B(T) = P^{-1}M_{B_0}(T)P$$

where $P = P_{B_0\leftarrow B}$ is the change matrix from B to B_0.

[1] This also follows from Taylor's theorem (Corollary 6.5.3 of Theorem 6.5.1 with $a = 1$).

Example 9.2.2

Let $T:\mathbb{R}^3\to\mathbb{R}^3$ be defined by $T(a,\ b,\ c)=(2a-b,\ b+c,\ c-3a)$. If B_0 denotes the standard basis of $\mathbb{R}^3$ and $B=\{(1,\ 1,\ 0),\ (1,\ 0,\ 1),\ (0,\ 1,\ 0)\}$, find an invertible matrix P such that $P^{-1}M_{B_0}(T)P=M_B(T)$.

Solution. We have

$$M_{B_0}(T)=\left[\begin{array}{ccc} C_{B_0}(2,\ 0,\ -3) & C_{B_0}(-1,\ 1,\ 0) & C_{B_0}(0,\ 1,\ 1)\end{array}\right]=\left[\begin{array}{rrr} 2 & -1 & 0\\ 0 & 1 & 1\\ -3 & 0 & 1\end{array}\right]$$

$$M_B(T)=\left[\begin{array}{ccc} C_B(1,\ 1,\ -3) & C_B(2,\ 1,\ -2) & C_B(-1,\ 1,\ 0)\end{array}\right]=\left[\begin{array}{rrr} 4 & 4 & -1\\ -3 & -2 & 0\\ -3 & -3 & 2\end{array}\right]$$

$$P=P_{B_0\leftarrow B}=\left[\begin{array}{ccc} C_{B_0}(1,\ 1,\ 0) & C_{B_0}(1,\ 0,\ 1) & C_{B_0}(0,\ 1,\ 0)\end{array}\right]=\left[\begin{array}{rrr} 1 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0\end{array}\right]$$

The reader can verify that $P^{-1}M_{B_0}(T)P=M_B(T)$; equivalently that $M_{B_0}(T)P=PM_B(T)$.

A square matrix is diagonalizable if and only if it is similar to a diagonal matrix. Theorem 9.2.3 comes into this as follows: Suppose an $n\times n$ matrix $A=M_{B_0}(T)$ is the matrix of some operator $T:V\to V$ with respect to an ordered basis B_0. If another ordered basis B of V can be found such that $M_B(T)=D$ is diagonal, then Theorem 9.2.3 shows how to find an invertible P such that $P^{-1}AP=D$. In other words, the "algebraic" problem of finding P such that $P^{-1}AP$ is diagonal comes down to the "geometric" problem of finding a basis B such that $M_B(T)$ is diagonal. This shift of emphasis is one of the most important techniques in linear algebra.

Each $n\times n$ matrix A can be easily realized as the matrix of an operator. In fact, (Example 9.1.4),

$$M_E(T_A)=A$$

where $T_A:\mathbb{R}^n\to\mathbb{R}^n$ is the matrix operator given by $T_A(\mathbf{x})=A\mathbf{x}$, and E is the standard basis of $\mathbb{R}^n$. The first part of the next theorem gives the converse of Theorem 9.2.3: Any pair of similar matrices can be realized as the matrices of the same linear operator with respect to different bases. This is part 1 of the following theorem.

Theorem 9.2.4

Let A be an $n\times n$ matrix and let E be the standard basis of $\mathbb{R}^n$.

1. *Let A' be similar to A, say $A'=P^{-1}AP$, and let B be the ordered basis of $\mathbb{R}^n$ consisting of the columns of P in order. Then $T_A:\mathbb{R}^n\to\mathbb{R}^n$ is linear and*

$$M_E(T_A)=A \quad\text{and}\quad M_B(T_A)=A'$$

2. *If B is any ordered basis of $\mathbb{R}^n$, let P be the (invertible) matrix whose columns are the vectors in B in order. Then*

$$M_B(T_A)=P^{-1}AP$$

Proof.

1. We have $M_E(T_A) = A$ by Example 9.1.4. Write $P = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{bmatrix}$ in terms of its columns so $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis of $\mathbb{R}^n$. Since E is the standard basis,

$$P_{E \leftarrow B} = \begin{bmatrix} C_E(\mathbf{b}_1) & \cdots & C_E(\mathbf{b}_n) \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{bmatrix} = P$$

   Hence Theorem 9.2.3 (with $B_0 = E$) gives $M_B(T_A) = P^{-1}M_E(T_A)P = P^{-1}AP = A'$.

2. Here P and B are as above, so again $P_{E \leftarrow B} = P$ and $M_B(T_A) = P^{-1}AP$. □

Example 9.2.3

Given $A = \begin{bmatrix} 10 & 6 \\ -18 & -11 \end{bmatrix}$, $P = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$, and $D = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$, verify that $P^{-1}AP = D$ and use this fact to find a basis B of $\mathbb{R}^2$ such that $M_B(T_A) = D$.

Solution. $P^{-1}AP = D$ holds if $AP = PD$; this verification is left to the reader. Let B consist of the columns of P in order, that is $B = \left\{ \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right\}$. Then Theorem 9.2.4 gives $M_B(T_A) = P^{-1}AP = D$. More explicitly,

$$M_B(T_A) = \left[C_B\left(T_A \begin{bmatrix} 2 \\ -3 \end{bmatrix}\right) \quad C_B\left(T_A \begin{bmatrix} -1 \\ 2 \end{bmatrix}\right) \right] = \left[C_B \begin{bmatrix} 2 \\ -3 \end{bmatrix} \quad C_B \begin{bmatrix} 2 \\ -4 \end{bmatrix} \right] = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} = D$$

Let A be an $n \times n$ matrix. As in Example 9.2.3, Theorem 9.2.4 provides a new way to find an invertible matrix P such that $P^{-1}AP$ is diagonal. The idea is to find a basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ of $\mathbb{R}^n$ such that $M_B(T_A) = D$ is diagonal and take $P = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$ to be the matrix with the $\mathbf{b}_j$ as columns. Then, by Theorem 9.2.4,

$$P^{-1}AP = M_B(T_A) = D$$

As mentioned above, this converts the algebraic problem of diagonalizing A into the geometric problem of finding the basis B. This new point of view is very powerful and will be explored in the next two sections.

Theorem 9.2.4 enables facts about matrices to be deduced from the corresponding properties of operators. Here is an example.

Example 9.2.4

1. If $T : V \to V$ is an operator where V is finite dimensional, show that $TST = T$ for some invertible operator $S : V \to V$.

2. If A is an $n \times n$ matrix, show that $AUA = A$ for some invertible matrix U.

Solution.

1. Let $B = \{\mathbf{b}_1, \dots, \mathbf{b}_r, \mathbf{b}_{r+1}, \dots, \mathbf{b}_n\}$ be a basis of V chosen so that $\ker T = \operatorname{span}\{\mathbf{b}_{r+1}, \dots, \mathbf{b}_n\}$. Then $\{T(\mathbf{b}_1), \dots, T(\mathbf{b}_r)\}$ is independent (Theorem 7.2.5), so complete it to a basis $\{T(\mathbf{b}_1), \dots, T(\mathbf{b}_r), \mathbf{f}_{r+1}, \dots, \mathbf{f}_n\}$ of V.

By Theorem 7.1.3, define $S : V \to V$ by

$$S[T(\mathbf{b}_i)] = \mathbf{b}_i \quad \text{for } 1 \leq i \leq r$$
$$S(\mathbf{f}_j) = \mathbf{b}_j \quad \text{for } r < j \leq n$$

Then S is an isomorphism by Theorem 7.3.1, and $TST = T$ because these operators agree on the basis B. In fact,

$$(TST)(\mathbf{b}_i) = T[ST(\mathbf{b}_i)] = T(\mathbf{b}_i) \text{ if } 1 \leq i \leq r, \text{ and}$$
$$(TST)(\mathbf{b}_j) = TS[T(\mathbf{b}_j)] = TS(\mathbf{0}) = \mathbf{0} = T(\mathbf{b}_j) \text{ for } r < j \leq n$$

2. Given A, let $T = T_A : \mathbb{R}^n \to \mathbb{R}^n$. By (1) let $TST = T$ where $S : \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism. If E is the standard basis of $\mathbb{R}^n$, then $A = M_E(T)$ by Theorem 9.2.4. If $U = M_E(S)$ then, by Theorem 9.2.1, U is invertible and

$$AUA = M_E(T)M_E(S)M_E(T) = M_E(TST) = M_E(T) = A$$

as required.

The reader will appreciate the power of these methods if he/she tries to find U directly in part 2 of Example 9.2.4, even if A is 2×2.

A property of $n \times n$ matrices is called a **similarity invariant** if, whenever a given $n \times n$ matrix A has the property, every matrix similar to A also has the property. Theorem 5.5.1 shows that rank, determinant, trace, and characteristic polynomial are all similarity invariants.

To illustrate how such similarity invariants are related to linear operators, consider the case of rank. If $T : V \to V$ is a linear operator, the matrices of T with respect to various bases of V all have the same rank (being similar), so it is natural to regard the common rank of all these matrices as a property of T itself and not of the particular matrix used to describe T. Hence the rank of T could be *defined* to be the rank of A, where A is *any* matrix of T. This would be unambiguous because rank is a similarity invariant. Of course, this is unnecessary in the case of rank because rank T was defined earlier to be the dimension of im T, and this was *proved* to equal the rank of every matrix representing T (Theorem 9.1.5). This definition of rank T is said to be *intrinsic* because it makes no reference to the matrices representing T. However, the technique serves to identify an intrinsic property of T with *every* similarity invariant, and some of these properties are not so easily defined directly.

In particular, if $T : V \to V$ is a linear operator on a finite dimensional space V, define the **determinant** of T (denoted $\det T$) by

$$\det T = \det M_B(T), \quad B \text{ any basis of } V$$

This is independent of the choice of basis B because, if D is any other basis of V, the matrices $M_B(T)$ and $M_D(T)$ are similar and so have the same determinant. In the same way, the **trace** of T (denoted $\text{tr } T$) can be defined by

$$\text{tr } T = \text{tr } M_B(T), \quad B \text{ any basis of } V$$

This is unambiguous for the same reason.

Theorems about matrices can often be translated to theorems about linear operators. Here is an example.

Example 9.2.5

Let S and T denote linear operators on the finite dimensional space V. Show that

$$\det(ST) = \det S \det T$$

Solution. Choose a basis B of V and use Theorem 9.2.1.

$$\begin{aligned}\det(ST) = \det M_B(ST) &= \det[M_B(S)M_B(T)] \\ &= \det[M_B(S)]\det[M_B(T)] = \det S \det T\end{aligned}$$

Recall next that the characteristic polynomial of a matrix is another similarity invariant: If A and A' are similar matrices, then $c_A(x) = c_{A'}(x)$ (Theorem 5.5.1). As discussed above, the discovery of a similarity invariant means the discovery of a property of linear operators. In this case, if $T : V \to V$ is a linear operator on the finite dimensional space V, define the **characteristic polynomial** of T by

$$c_T(x) = c_A(x) \text{ where } A = M_B(T),\ B \text{ any basis of } V$$

In other words, the characteristic polynomial of an operator T is the characteristic polynomial of *any* matrix representing T. This is unambiguous because any two such matrices are similar by Theorem 9.2.3.

Example 9.2.6

Compute the characteristic polynomial $c_T(x)$ of the operator $T : \mathbf{P}_2 \to \mathbf{P}_2$ given by $T(a+bx+cx^2) = (b+c)+(a+c)x+(a+b)x^2$.

Solution. If $B = \{1,\ x,\ x^2\}$, the corresponding matrix of T is

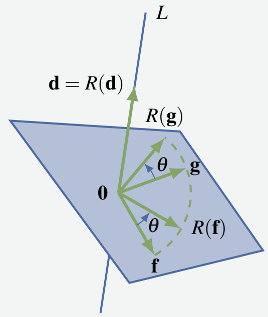
$$M_B(T) = \left[\begin{array}{ccc} C_B[T(1)] & C_B[T(x)] & C_B[T(x^2)] \end{array}\right] = \left[\begin{array}{ccc} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{array}\right]$$

Hence $c_T(x) = \det[xI - M_B(T)] = x^3 - 3x - 2 = (x+1)^2(x-2)$.

In Section 4.4 we computed the matrix of various projections, reflections, and rotations in $\mathbb{R}^3$. However, the methods available then were not adequate to find the matrix of a rotation about a line through the origin. We conclude this section with an example of how Theorem 9.2.3 can be used to compute such a matrix.

Example 9.2.7

Let L be the line in $\mathbb{R}^3$ through the origin with (unit) direction vector $\mathbf{d} = \frac{1}{3}\left[\begin{array}{ccc} 2 & 1 & 2 \end{array}\right]^T$. Compute the matrix of the rotation about L through an angle θ measured counterclockwise when viewed in the direction of $\mathbf{d}$.



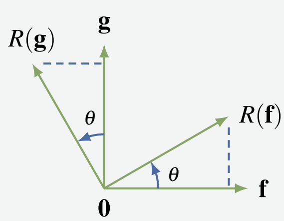
Solution. Let $R : \mathbb{R}^3 \to \mathbb{R}^3$ be the rotation. The idea is to first find a basis B_0 for which the matrix of $M_{B_0}(R)$ of R is easy to compute, and then use Theorem 9.2.3 to compute the "standard" matrix $M_E(R)$ with respect to the standard basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of $\mathbb{R}^3$.

To construct the basis B_0, let K denote the plane through the origin with $\mathbf{d}$ as normal, shaded in the diagram. Then the vectors $\mathbf{f} = \frac{1}{3}\left[\begin{array}{ccc} -2 & 2 & 1 \end{array}\right]^T$ and $\mathbf{g} = \frac{1}{3}\left[\begin{array}{ccc} 1 & 2 & -2 \end{array}\right]^T$ are both in K (they are orthogonal to $\mathbf{d}$) and are independent (they are orthogonal to each other).

Hence $B_0 = \{\mathbf{d}, \mathbf{f}, \mathbf{g}\}$ is an orthonormal basis of $\mathbb{R}^3$, and the effect of R on B_0 is easy to determine. In fact $R(\mathbf{d}) = \mathbf{d}$ and (as in Theorem 2.6.4) the second diagram gives

$$R(\mathbf{f}) = \cos\theta\mathbf{f} + \sin\theta\mathbf{g} \quad \text{and} \quad R(\mathbf{g}) = -\sin\theta\mathbf{f} + \cos\theta\mathbf{g}$$

because $\|\mathbf{f}\| = 1 = \|\mathbf{g}\|$. Hence



$$M_{B_0}(R) = \left[\begin{array}{ccc} C_{B_0}(\mathbf{d}) & C_{B_0}(\mathbf{f}) & C_{B_0}(\mathbf{g}) \end{array}\right] = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{array}\right]$$

Now Theorem 9.2.3 (with $B = E$) asserts that $M_E(R) = P^{-1}M_{B_0}(R)P$ where

$$P = P_{B_0 \leftarrow E} = \left[\begin{array}{ccc} C_{B_0}(\mathbf{e}_1) & C_{B_0}(\mathbf{e}_2) & C_{B_0}(\mathbf{e}_3) \end{array}\right] = \frac{1}{3}\left[\begin{array}{rrr} 2 & 1 & 2 \\ -2 & 2 & 1 \\ 1 & 2 & -2 \end{array}\right]$$

using the expansion theorem (Theorem 5.3.6). Since $P^{-1} = P^T$ (P is orthogonal), the matrix of R with respect to E is

$$\begin{aligned} M_E(R) &= P^T M_{B_0}(R)P \\ &= \frac{1}{9}\left[\begin{array}{ccc} 5\cos\theta + 4 & 6\sin\theta - 2\cos\theta + 2 & 4 - 3\sin\theta - 4\cos\theta \\ 2 - 6\sin\theta - 2\cos\theta & 8\cos\theta + 1 & 6\sin\theta - 2\cos\theta + 2 \\ 3\sin\theta - 4\cos\theta + 4 & 2 - 6\sin\theta - 2\cos\theta & 5\cos\theta + 4 \end{array}\right] \end{aligned}$$

As a check one verifies that this is the identity matrix when $\theta = 0$, as it should.

Note that in Example 9.2.7 not much motivation was given to the choices of the (orthonormal) vectors $\mathbf{f}$ and $\mathbf{g}$ in the basis B_0, which is the key to the solution. However, if we begin with *any* basis containing $\mathbf{d}$ the Gram-Schmidt algorithm will produce an orthogonal basis containing $\mathbf{d}$, and the other two vectors will automatically be in $L^\perp = K$.

Exercises for 9.2

Exercise 9.2.1 In each case find $P_{D\leftarrow B}$, where B and D are ordered bases of V. Then verify that $C_D(\mathbf{v}) = P_{D\leftarrow B}C_B(\mathbf{v})$.

a. $V = \mathbb{R}^2$, $B = \{(0, -1), (2, 1)\}$, $D = \{(0, 1), (1, 1)\}$, $\mathbf{v} = (3, -5)$

b. $V = \mathbf{P}_2$, $B = \{x, 1+x, x^2\}$, $D = \{2, x+3, x^2-1\}$, $\mathbf{v} = 1+x+x^2$

c. $V = \mathbf{M}_{22}$,
$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$,
$D = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}$,
$\mathbf{v} = \begin{bmatrix} 3 & -1 \\ 1 & 4 \end{bmatrix}$

Exercise 9.2.2 In $\mathbb{R}^3$ find $P_{D\leftarrow B}$, where $B = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ and $D = \{(1, 0, 1), (1, 0, -1), (0, 1, 0)\}$. If $\mathbf{v} = (a, b, c)$, show that $C_D(\mathbf{v}) = \frac{1}{2}\begin{bmatrix} a+c \\ a-c \\ 2b \end{bmatrix}$ and $C_B(\mathbf{v}) = \begin{bmatrix} a-b \\ b-c \\ c \end{bmatrix}$, and verify that $C_D(\mathbf{v}) = P_{D\leftarrow B}C_B(\mathbf{v})$.

Exercise 9.2.3 In $\mathbf{P}_3$ find $P_{D\leftarrow B}$ if $B = \{1, x, x^2, x^3\}$ and $D = \{1, (1-x), (1-x)^2, (1-x)^3\}$. Then express $p = a+bx+cx^2+dx^3$ as a polynomial in powers of $(1-x)$.

Exercise 9.2.4 In each case verify that $P_{D\leftarrow B}$ is the inverse of $P_{B\leftarrow D}$ and that $P_{E\leftarrow D}P_{D\leftarrow B} = P_{E\leftarrow B}$, where B, D, and E are ordered bases of V.

a. $V = \mathbb{R}^3$, $B = \{(1, 1, 1), (1, -2, 1), (1, 0, -1)\}$, $D =$ standard basis, $E = \{(1, 1, 1), (1, -1, 0), (-1, 0, 1)\}$

b. $V = \mathbf{P}_2$, $B = \{1, x, x^2\}$, $D = \{1+x+x^2, 1-x, -1+x^2\}$, $E = \{x^2, x, 1\}$

Exercise 9.2.5 Use property (2) of Theorem 9.2.2, with D the standard basis of $\mathbb{R}^n$, to find the inverse of:

a. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ b. $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 0 \\ -1 & 0 & 2 \end{bmatrix}$

Exercise 9.2.6 Find $P_{D\leftarrow B}$ if $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4\}$ and $D = \{\mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_1, \mathbf{b}_4\}$. Change matrices arising when the bases differ only in the *order* of the vectors are called **permutation matrices**.

Exercise 9.2.7 In each case, find $P = P_{B_0\leftarrow B}$ and verify that $P^{-1}M_{B_0}(T)P = M_B(T)$ for the given operator T.

a. $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(a, b, c) = (2a-b, b+c, c-3a)$; $B_0 = \{(1, 1, 0), (1, 0, 1), (0, 1, 0)\}$ and B is the standard basis.

b. $T : \mathbf{P}_2 \to \mathbf{P}_2$, $T(a+bx+cx^2) = (a+b)+(b+c)x+(c+a)x^2$; $B_0 = \{1, x, x^2\}$ and $B = \{1-x^2, 1+x, 2x+x^2\}$

c. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$,
$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+d & b+c \\ a+c & b+d \end{bmatrix}$;
$B_0 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$,
and
$B = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \right\}$

Exercise 9.2.8 In each case, verify that $P^{-1}AP = D$ and find a basis B of $\mathbb{R}^2$ such that $M_B(T_A) = D$.

a. $A = \begin{bmatrix} 11 & -6 \\ 12 & -6 \end{bmatrix}$ $P = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$ $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

b. $A = \begin{bmatrix} 29 & -12 \\ 70 & -29 \end{bmatrix}$ $P = \begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}$ $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Exercise 9.2.9 In each case, compute the characteristic polynomial $c_T(x)$.

a. $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T(a, b) = (a-b, 2b-a)$

b. $T : \mathbb{R}^2 \to \mathbb{R}^2$, $T(a, b) = (3a+5b, 2a+3b)$

c. $T : \mathbf{P}_2 \to \mathbf{P}_2$,
$T(a+bx+cx^2)$
$= (a-2c)+(2a+b+c)x+(c-a)x^2$

d. $T : \mathbf{P}_2 \to \mathbf{P}_2$,
$T(a+bx+cx^2)$
$= (a+b-2c)+(a-2b+c)x+(b-2a)x^2$

e. $T : \mathbb{R}^3 \to \mathbb{R}^3$, $T(a, b, c) = (b, c, a)$

f. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$, $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a-c & b-d \\ a-c & b-d \end{bmatrix}$

Exercise 9.2.10 If V is finite dimensional, show that a linear operator T on V has an inverse if and only if $\det T \neq 0$.

Exercise 9.2.11 Let S and T be linear operators on V where V is finite dimensional.

a. Show that $\text{tr}\,(ST) = \text{tr}\,(TS)$. [*Hint*: Lemma 5.5.1.]

b. [See Exercise 9.1.19.] For a in $\mathbb{R}$, show that $\text{tr}\,(S+T) = \text{tr}\, S + \text{tr}\, T$, and $\text{tr}\,(aT) = a\,\text{tr}\,(T)$.

Exercise 9.2.12 If A and B are $n \times n$ matrices, show that they have the same null space if and only if $A = UB$ for some invertible matrix U. [*Hint*: Exercise 7.3.28.]

Exercise 9.2.13 If A and B are $n \times n$ matrices, show that they have the same column space if and only if $A = BU$ for some invertible matrix U. [*Hint*: Exercise 7.3.28.]

Exercise 9.2.14 Let $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard ordered basis of $\mathbb{R}^n$, written as columns. If $D = \{\mathbf{d}_1, \dots, \mathbf{d}_n\}$ is any ordered basis, show that $P_{E \leftarrow D} = \begin{bmatrix} \mathbf{d}_1 & \cdots & \mathbf{d}_n \end{bmatrix}$.

Exercise 9.2.15 Let $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ be any ordered basis of $\mathbb{R}^n$, written as columns. If $Q = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix}$ is the matrix with the $\mathbf{b}_i$ as columns, show that $QC_B(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v}$ in $\mathbb{R}^n$.

Exercise 9.2.16 Given a complex number w, define $T_w : \mathbb{C} \to \mathbb{C}$ by $T_w(z) = wz$ for all z in $\mathbb{C}$.

a. Show that T_w is a linear operator for each w in $\mathbb{C}$, viewing $\mathbb{C}$ as a real vector space.

b. If B is any ordered basis of $\mathbb{C}$, define $S : \mathbb{C} \to \mathbf{M}_{22}$ by $S(w) = M_B(T_w)$ for all w in $\mathbb{C}$. Show that S is a one-to-one linear transformation with the additional property that $S(wv) = S(w)S(v)$ holds for all w and v in $\mathbb{C}$.

c. Taking $B = \{1, i\}$ show that $S(a+bi) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ for all complex numbers $a+bi$. This is called the **regular representation** of the complex numbers as 2×2 matrices. If θ is any angle, describe $S(e^{i\theta})$ geometrically. Show that $S(\overline{w}) = S(w)^T$ for all w in $\mathbb{C}$; that is, that conjugation corresponds to transposition.

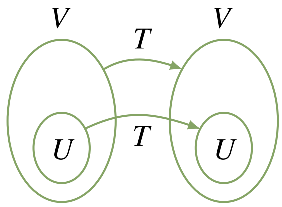
Exercise 9.2.17 Let $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ and $D = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_n\}$ be two ordered bases of a vector space V. Prove that $C_D(\mathbf{v}) = P_{D \leftarrow B} C_B(\mathbf{v})$ holds for all $\mathbf{v}$ in V as follows: Express each $\mathbf{b}_j$ in the form $\mathbf{b}_j = p_{1j}\mathbf{d}_1 + p_{2j}\mathbf{d}_2 + \cdots + p_{nj}\mathbf{d}_n$ and write $P = [p_{ij}]$. Show that $P = \begin{bmatrix} C_D(\mathbf{b}_1) & C_D(\mathbf{b}_1) & \cdots & C_D(\mathbf{b}_1) \end{bmatrix}$ and that $C_D(\mathbf{v}) = PC_B(\mathbf{v})$ for all $\mathbf{v}$ in B.

Exercise 9.2.18 Find the standard matrix of the rotation R about the line through the origin with direction vector $\mathbf{d} = \begin{bmatrix} 2 & 3 & 6 \end{bmatrix}^T$. [*Hint*: Consider $\mathbf{f} = \begin{bmatrix} 6 & 2 & -3 \end{bmatrix}^T$ and $\mathbf{g} = \begin{bmatrix} 3 & -6 & 2 \end{bmatrix}^T$.]

9.3 Invariant Subspaces and Direct Sums

A fundamental question in linear algebra is the following: If $T : V \to V$ is a linear operator, how can a basis B of V be chosen so the matrix $M_B(T)$ is as simple as possible? A basic technique for answering such questions will be explained in this section. If U is a subspace of V, write its image under T as

$$T(U) = \{T(\mathbf{u}) \mid \mathbf{u} \text{ in } U\}$$



Definition 9.5 T-invariant Subspace

*Let $T : V \to V$ be an operator. A subspace $U \subseteq V$ is called **T-invariant** if $T(U) \subseteq U$, that is, $T(\mathbf{u}) \in U$ for every vector $\mathbf{u} \in U$. Hence T is a linear operator on the vector space U.*

This is illustrated in the diagram, and the fact that $T : U \to U$ is an operator on U is the primary reason for our interest in T-invariant subspaces.

Example 9.3.1

Let $T : V \to V$ be any linear operator. Then:

1. $\{\mathbf{0}\}$ and V are T-invariant subspaces.
2. Both $\ker T$ and $\operatorname{im} T = T(V)$ are T-invariant subspaces.
3. If U and W are T-invariant subspaces, so are $T(U)$, $U \cap W$, and $U + W$.

Solution. Item 1 is clear, and the rest is left as Exercises 9.3.1 and 9.3.2.

Example 9.3.2

Define $T : \mathbb{R}^3 \to \mathbb{R}^3$ by $T(a,\, b,\, c) = (3a + 2b,\, b - c,\, 4a + 2b - c)$. Then $U = \{(a,\, b,\, a) \mid a,\, b \text{ in } \mathbb{R}\}$ is T-invariant because

$$T(a,\, b,\, a) = (3a + 2b,\, b - a,\, 3a + 2b)$$

is in U for all a and b (the first and last entries are equal).

If a spanning set for a subspace U is known, it is easy to check whether U is T-invariant.

Example 9.3.3

Let $T : V \to V$ be a linear operator, and suppose that $U = \operatorname{span}\{\mathbf{u}_1,\, \mathbf{u}_2,\, \dots,\, \mathbf{u}_k\}$ is a subspace of V. Show that U is T-invariant if and only if $T(\mathbf{u}_i)$ lies in U for each $i = 1,\, 2,\, \dots,\, k$.

Solution. Given $\mathbf{u}$ in U, write it as $\mathbf{u} = r_1\mathbf{u}_1 + \cdots + r_k\mathbf{u}_k$, r_i in $\mathbb{R}$. Then

$$T(\mathbf{u}) = r_1 T(\mathbf{u}_1) + \cdots + r_k T(\mathbf{u}_k)$$

and this lies in U if each $T(\mathbf{u}_i)$ lies in U. This shows that U is T-invariant if each $T(\mathbf{u}_i)$ lies in U; the converse is clear.

Example 9.3.4

Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(a,\ b) = (b,\ -a)$. Show that $\mathbb{R}^2$ contains no T-invariant subspace except 0 and $\mathbb{R}^2$.

Solution. Suppose, if possible, that U is T-invariant, but $U \neq 0$, $U \neq \mathbb{R}^2$. Then U has dimension 1 so $U = \mathbb{R}\mathbf{x}$ where $\mathbf{x} \neq \mathbf{0}$. Now $T(\mathbf{x})$ lies in U—say $T(\mathbf{x}) = r\mathbf{x}$, r in $\mathbb{R}$. If we write $\mathbf{x} = (a,\ b)$, this is $(b,\ -a) = r(a,\ b)$, which gives $b = ra$ and $-a = rb$. Eliminating b gives $r^2a = rb = -a$, so $(r^2+1)a = 0$. Hence $a = 0$. Then $b = ra = 0$ too, contrary to the assumption that $\mathbf{x} \neq \mathbf{0}$. Hence no one-dimensional T-invariant subspace exists.

Definition 9.6 Restriction of an Operator

Let $T : V \to V$ be a linear operator. If U is any T-invariant subspace of V, then

$$T : U \to U$$

is a linear operator on the subspace U, called the ***restriction*** *of T to U.*

This is the reason for the importance of T-invariant subspaces and is the first step toward finding a basis that simplifies the matrix of T.

Theorem 9.3.1

Let $T : V \to V$ be a linear operator where V has dimension n and suppose that U is any T-invariant subspace of V. Let $B_1 = \{\mathbf{b}_1,\ \dots,\ \mathbf{b}_k\}$ be any basis of U and extend it to a basis $B = \{\mathbf{b}_1,\ \dots,\ \mathbf{b}_k,\ \mathbf{b}_{k+1},\ \dots,\ \mathbf{b}_n\}$ of V in any way. Then $M_B(T)$ has the block triangular form

$$M_B(T) = \left[\begin{array}{cc} M_{B_1}(T) & Y \\ 0 & Z \end{array} \right]$$

where Z is $(n-k) \times (n-k)$ and $M_{B_1}(T)$ is the matrix of the restriction of T to U.

Proof. The matrix of (the restriction) $T : U \to U$ with respect to the basis B_1 is the $k \times k$ matrix

$$M_{B_1}(T) = \left[\begin{array}{cccc} C_{B_1}[T(\mathbf{b}_1)] & C_{B_1}[T(\mathbf{b}_2)] & \cdots & C_{B_1}[T(\mathbf{b}_k)] \end{array} \right]$$

Now compare the first column $C_{B_1}[T(\mathbf{b}_1)]$ here with the first column $C_B[T(\mathbf{b}_1)]$ of $M_B(T)$. The fact that $T(\mathbf{b}_1)$ lies in U (because U is T-invariant) means that $T(\mathbf{b}_1)$ has the form

$$T(\mathbf{b}_1) = t_1\mathbf{b}_1 + t_2\mathbf{b}_2 + \cdots + t_k\mathbf{b}_k + 0\mathbf{b}_{k+1} + \cdots + 0\mathbf{b}_n$$

Consequently,

$$C_{B_1}[T(\mathbf{b}_1)] = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_k \end{bmatrix} \text{ in } \mathbb{R}^k \quad \text{whereas} \quad C_B[T(\mathbf{b}_1)] = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ in } \mathbb{R}^n$$

This shows that the matrices $M_B(T)$ and $\begin{bmatrix} M_{B_1}(T) & Y \\ 0 & Z \end{bmatrix}$ have identical first columns.

Similar statements apply to columns 2, 3, ..., k, and this proves the theorem. □

The block upper triangular form for the matrix $M_B(T)$ in Theorem 9.3.1 is very useful because the determinant of such a matrix equals the product of the determinants of each of the diagonal blocks. This is recorded in Theorem 9.3.2 for reference, together with an important application to characteristic polynomials.

Theorem 9.3.2

Let A be a block upper triangular matrix, say

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1n} \\ 0 & A_{22} & A_{23} & \cdots & A_{2n} \\ 0 & 0 & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & A_{nn} \end{bmatrix}$$

where the diagonal blocks are square. Then:

1. $\det A = (\det A_{11})(\det A_{22})(\det A_{33})\cdots(\det A_{nn})$.
2. $c_A(x) = c_{A_{11}}(x)c_{A_{22}}(x)c_{A_{33}}(x)\cdots c_{A_{nn}}(x)$.

Proof. If $n = 2$, (1) is Theorem 3.1.5; the general case (by induction on n) is left to the reader. Then (2) follows from (1) because

$$xI - A = \begin{bmatrix} xI - A_{11} & -A_{12} & -A_{13} & \cdots & -A_{1n} \\ 0 & xI - A_{22} & -A_{23} & \cdots & -A_{2n} \\ 0 & 0 & xI - A_{33} & \cdots & -A_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & xI - A_{nn} \end{bmatrix}$$

where, in each diagonal block, the symbol I stands for the identity matrix of the appropriate size. □

Example 9.3.5

Consider the linear operator $T : \mathbf{P}_2 \to \mathbf{P}_2$ given by

$$T(a+bx+cx^2) = (-2a-b+2c)+(a+b)x+(-6a-2b+5c)x^2$$

Show that $U = \text{span}\{x,\ 1+2x^2\}$ is T-invariant, use it to find a block upper triangular matrix for T, and use that to compute $c_T(x)$.

Solution. U is T-invariant by Example 9.3.3 because $U = \text{span}\{x,\ 1+2x^2\}$ and both $T(x)$ and $T(1+2x^2)$ lie in U:

$$\begin{aligned} T(x) &= -1+x-2x^2 = x-(1+2x^2) \\ T(1+2x^2) &= 2+x+4x^2 = x+2(1+2x^2) \end{aligned}$$

Extend the basis $B_1 = \{x,\ 1+2x^2\}$ of U to a basis B of $\mathbf{P}_2$ in any way at all—say, $B = \{x,\ 1+2x^2,\ x^2\}$. Then

$$\begin{aligned} M_B(T) &= \left[\begin{array}{ccc} C_B[T(x)] & C_B[T(1+2x^2)] & C_B[T(x^2)] \end{array}\right] \\ &= \left[\begin{array}{ccc} C_B(-1+x-2x^2) & C_B(2+x+4x^2) & C_B(2+5x^2) \end{array}\right] \\ &= \left[\begin{array}{rr|r} 1 & 1 & 0 \\ -1 & 2 & 2 \\ \hline 0 & 0 & 1 \end{array}\right] \end{aligned}$$

is in block upper triangular form as expected. Finally,

$$c_T(x) = \det \left[\begin{array}{cc|c} x-1 & -1 & 0 \\ 1 & x-2 & -2 \\ \hline 0 & 0 & x-1 \end{array}\right] = (x^2-3x+3)(x-1)$$

Eigenvalues

Let $T : V \to V$ be a linear operator. A one-dimensional subspace $\mathbb{R}\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$, is T-invariant if and only if $T(r\mathbf{v}) = rT(\mathbf{v})$ lies in $\mathbb{R}\mathbf{v}$ for all r in $\mathbb{R}$. This holds if and only if $T(\mathbf{v})$ lies in $\mathbb{R}\mathbf{v}$; that is, $T(\mathbf{v}) = \lambda\mathbf{v}$ for some λ in $\mathbb{R}$. A real number λ is called an **eigenvalue** of an operator $T : V \to V$ if

$$T(\mathbf{v}) = \lambda\mathbf{v}$$

holds for some nonzero vector $\mathbf{v}$ in V. In this case, $\mathbf{v}$ is called an **eigenvector** of T corresponding to λ. The subspace

$$E_\lambda(T) = \{\mathbf{v} \text{ in } V \mid T(\mathbf{v}) = \lambda\mathbf{v}\}$$

is called the **eigenspace** of T corresponding to λ. These terms are consistent with those used in Section 5.5 for matrices. If A is an $n \times n$ matrix, a real number λ is an eigenvalue of the matrix operator $T_A : \mathbb{R}^n \to \mathbb{R}^n$ if and only if λ is an eigenvalue of the matrix A. Moreover, the eigenspaces agree:

$$E_\lambda(T_A) = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x}\} = E_\lambda(A)$$

The following theorem reveals the connection between the eigenspaces of an operator T and those of the matrices representing T.

Theorem 9.3.3

Let $T : V \to V$ be a linear operator where $\dim V = n$, let B denote any ordered basis of V, and let $C_B : V \to \mathbb{R}^n$ denote the coordinate isomorphism. Then:

1. *The eigenvalues λ of T are precisely the eigenvalues of the matrix $M_B(T)$ and thus are the roots of the characteristic polynomial $c_T(x)$.*
2. *In this case the eigenspaces $E_\lambda(T)$ and $E_\lambda[M_B(T)]$ are isomorphic via the restriction $C_B : E_\lambda(T) \to E_\lambda[M_B(T)]$.*

Proof. Write $A = M_B(T)$ for convenience. If $T(\mathbf{v}) = \lambda\mathbf{v}$, then $\lambda C_B(\mathbf{v}) = C_B[T(\mathbf{v})] = AC_B(\mathbf{v})$ because C_B is linear. Hence $C_B(\mathbf{v})$ lies in $E_\lambda(A)$, so we do have a function $C_B : E_\lambda(T) \to E_\lambda(A)$. It is clearly linear and one-to-one; we claim it is onto. If $\mathbf{x}$ is in $E_\lambda(A)$, write $\mathbf{x} = C_B(\mathbf{v})$ for some $\mathbf{v}$ in V (C_B is onto). This $\mathbf{v}$ actually lies in $E_\lambda(T)$. To see why, observe that

$$C_B[T(\mathbf{v})] = AC_B(\mathbf{v}) = A\mathbf{x} = \lambda\mathbf{x} = \lambda C_B(\mathbf{v}) = C_B(\lambda\mathbf{v})$$

Hence $T(\mathbf{v}) = \lambda\mathbf{v}$ because C_B is one-to-one, and this proves (2). As to (1), we have already shown that eigenvalues of T are eigenvalues of A. The converse follows, as in the foregoing proof that C_B is onto. □

Theorem 9.3.3 shows how to pass back and forth between the eigenvectors of an operator T and the eigenvectors of any matrix $M_B(T)$ of T:

$$\mathbf{v} \text{ lies in } E_\lambda(T) \quad \text{if and only if} \quad C_B(\mathbf{v}) \text{ lies in } E_\lambda[M_B(T)]$$

Example 9.3.6

Find the eigenvalues and eigenspaces for $T : \mathbf{P}_2 \to \mathbf{P}_2$ given by

$$T(a+bx+cx^2) = (2a+b+c) + (2a+b-2c)x - (a+2c)x^2$$

Solution. If $B = \{1,\ x,\ x^2\}$, then

$$M_B(T) = \left[\begin{array}{ccc} C_B[T(1)] & C_B[T(x)] & C_B[T(x^2)] \end{array}\right] = \left[\begin{array}{rrr} 2 & 1 & 1 \\ 2 & 1 & -2 \\ -1 & 0 & -2 \end{array}\right]$$

Hence $c_T(x) = \det[xI - M_B(T)] = (x+1)^2(x-3)$ as the reader can verify.

Moreover, $E_{-1}[M_B(T)] = \mathbb{R}\left[\begin{array}{r} -1 \\ 2 \\ 1 \end{array}\right]$ and $E_3[M_B(T)] = \mathbb{R}\left[\begin{array}{r} 5 \\ 6 \\ -1 \end{array}\right]$, so Theorem 9.3.3 gives $E_{-1}(T) = \mathbb{R}(-1+2x+x^2)$ and $E_3(T) = \mathbb{R}(5+6x-x^2)$.

Theorem 9.3.4

Each eigenspace of a linear operator $T : V \to V$ is a T-invariant subspace of V.

Proof. If $\mathbf{v}$ lies in the eigenspace $E_\lambda(T)$, then $T(\mathbf{v}) = \lambda\mathbf{v}$, so $T[T(\mathbf{v})] = T(\lambda\mathbf{v}) = \lambda T(\mathbf{v})$. This shows that $T(\mathbf{v})$ lies in $E_\lambda(T)$ too. □

Direct Sums

Sometimes vectors in a space V can be written naturally as a sum of vectors in two subspaces. For example, in the space $\mathbf{M}_{nn}$ of all $n \times n$ matrices, we have subspaces

$$U = \{P \text{ in } \mathbf{M}_{nn} \mid P \text{ is symmetric } \} \quad \text{and} \quad W = \{Q \text{ in } \mathbf{M}_{nn} \mid Q \text{ is skew symmetric}\}$$

where a matrix Q is called **skew-symmetric** if $Q^T = -Q$. Then every matrix A in $\mathbf{M}_{nn}$ can be written as the sum of a matrix in U and a matrix in W; indeed,

$$A = \tfrac{1}{2}(A + A^T) + \tfrac{1}{2}(A - A^T)$$

where $\frac{1}{2}(A + A^T)$ is symmetric and $\frac{1}{2}(A - A^T)$ is skew symmetric. Remarkably, this representation is unique: If $A = P + Q$ where $P^T = P$ and $Q^T = -Q$, then $A^T = P^T + Q^T = P - Q$; adding this to $A = P + Q$ gives $P = \frac{1}{2}(A + A^T)$, and subtracting gives $Q = \frac{1}{2}(A - A^T)$. In addition, this uniqueness turns out to be closely related to the fact that the only matrix in both U and W is 0. This is a useful way to view matrices, and the idea generalizes to the important notion of a direct sum of subspaces.

If U and W are subspaces of V, their *sum* $U + W$ and their *intersection* $U \cap W$ were defined in Section 6.4 as follows:

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \text{ in } U \text{ and } \mathbf{w} \text{ in } W\}$$
$$U \cap W = \{\mathbf{v} \mid \mathbf{v} \text{ lies in both } U \text{ and } W\}$$

These are subspaces of V, the sum containing both U and W and the intersection contained in both U and W. It turns out that the most interesting pairs U and W are those for which $U \cap W$ is as small as possible and $U + W$ is as large as possible.

Definition 9.7 Direct Sum of Subspaces

A vector space V is said to be the ***direct sum*** *of subspaces U and W if*

$$U \cap W = \{\mathbf{0}\} \quad \text{and} \quad U + W = V$$

In this case we write $V = U \oplus W$. Given a subspace U, any subspace W such that $V = U \oplus W$ is called a ***complement*** *of U in V.*

Example 9.3.7

In the space $\mathbb{R}^5$, consider the subspaces $U = \{(a, b, c, 0, 0) \mid a, b, \text{ and } c \text{ in } \mathbb{R}\}$ and $W = \{(0, 0, 0, d, e) \mid d \text{ and } e \text{ in } \mathbb{R}\}$. Show that $\mathbb{R}^5 = U \oplus W$.

Solution. If $\mathbf{x} = (a, b, c, d, e)$ is any vector in $\mathbb{R}^5$, then $\mathbf{x} = (a, b, c, 0, 0) + (0, 0, 0, d, e)$, so $\mathbf{x}$ lies in $U + W$. Hence $\mathbb{R}^5 = U + W$. To show that $U \cap W = \{\mathbf{0}\}$, let $\mathbf{x} = (a, b, c, d, e)$ lie in $U \cap W$. Then $d = e = 0$ because $\mathbf{x}$ lies in U, and $a = b = c = 0$ because $\mathbf{x}$ lies in W. Thus $\mathbf{x} = (0, 0, 0, 0, 0) = \mathbf{0}$, so $\mathbf{0}$ is the only vector in $U \cap W$. Hence $U \cap W = \{\mathbf{0}\}$.

Example 9.3.8

If U is a subspace of $\mathbb{R}^n$, show that $\mathbb{R}^n = U \oplus U^\perp$.

Solution. The equation $\mathbb{R}^n = U + U^\perp$ holds because, given $\mathbf{x}$ in $\mathbb{R}^n$, the vector $\text{proj}_U\ \mathbf{x}$ lies in U and $\mathbf{x} - \text{proj}_U\ \mathbf{x}$ lies in $U^\perp$. To see that $U \cap U^\perp = \{\mathbf{0}\}$, observe that any vector in $U \cap U^\perp$ is orthogonal to itself and hence must be zero.

Example 9.3.9

Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be a basis of a vector space V, and partition it into two parts: $\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ and $\{\mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$. If $U = \text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ and $W = \text{span}\{\mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$, show that $V = U \oplus W$.

Solution. If $\mathbf{v}$ lies in $U \cap W$, then $\mathbf{v} = a_1\mathbf{e}_1 + \cdots + a_k\mathbf{e}_k$ and $\mathbf{v} = b_{k+1}\mathbf{e}_{k+1} + \cdots + b_n\mathbf{e}_n$ hold for some a_i and b_j in $\mathbb{R}$. The fact that the $\mathbf{e}_i$ are linearly independent forces all $a_i = b_j = 0$, so $\mathbf{v} = \mathbf{0}$. Hence $U \cap W = \{\mathbf{0}\}$. Now, given $\mathbf{v}$ in V, write $\mathbf{v} = v_1\mathbf{e}_1 + \cdots + v_n\mathbf{e}_n$ where the v_i are in $\mathbb{R}$. Then $\mathbf{v} = \mathbf{u} + \mathbf{w}$, where $\mathbf{u} = v_1\mathbf{e}_1 + \cdots + v_k\mathbf{e}_k$ lies in U and $\mathbf{w} = v_{k+1}\mathbf{e}_{k+1} + \cdots + v_n\mathbf{e}_n$ lies in W. This proves that $V = U + W$.

Example 9.3.9 is typical of all direct sum decompositions.

Theorem 9.3.5

Let U and W be subspaces of a finite dimensional vector space V. The following three conditions are equivalent:

1. *$V = U \oplus W$.*
2. *Each vector $\mathbf{v}$ in V can be written uniquely in the form*

$$\mathbf{v} = \mathbf{u} + \mathbf{w} \qquad \mathbf{u} \text{ in } U,\ \mathbf{w} \text{ in } W$$

3. *If $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ and $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ are bases of U and W, respectively, then $B = \{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{w}_1, \dots, \mathbf{w}_m\}$ is a basis of V.*

(The uniqueness in (2) means that if $\mathbf{v} = \mathbf{u}_1 + \mathbf{w}_1$ is another such representation, then $\mathbf{u}_1 = \mathbf{u}$ and $\mathbf{w}_1 = \mathbf{w}$.)

Proof. Example 9.3.9 shows that (3) $\Rightarrow$ (1).

(1) $\Rightarrow$ (2). Given $\mathbf{v}$ in V, we have $\mathbf{v} = \mathbf{u} + \mathbf{w}$, $\mathbf{u}$ in U, $\mathbf{w}$ in W, because $V = U + W$. If also $\mathbf{v} = \mathbf{u}_1 + \mathbf{w}_1$, then $\mathbf{u} - \mathbf{u}_1 = \mathbf{w}_1 - \mathbf{w}$ lies in $U \cap W = \{\mathbf{0}\}$, so $\mathbf{u} = \mathbf{u}_1$ and $\mathbf{w} = \mathbf{w}_1$.

(2) $\Rightarrow$ (3). Given $\mathbf{v}$ in V, we have $\mathbf{v} = \mathbf{u} + \mathbf{w}$, $\mathbf{u}$ in U, $\mathbf{w}$ in W. Hence $\mathbf{v}$ lies in span B; that is, $V = \text{span } B$. To see that B is independent, let $a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k + b_1\mathbf{w}_1 + \cdots + b_m\mathbf{w}_m = \mathbf{0}$. Write $\mathbf{u} = a_1\mathbf{u}_1 + \cdots + a_k\mathbf{u}_k$ and $\mathbf{w} = b_1\mathbf{w}_1 + \cdots + b_m\mathbf{w}_m$. Then $\mathbf{u} + \mathbf{w} = \mathbf{0}$, and so $\mathbf{u} = \mathbf{0}$ and $\mathbf{w} = \mathbf{0}$ by the uniqueness in (2). Hence $a_i = 0$ for all i and $b_j = 0$ for all j. □

Condition (3) in Theorem 9.3.5 gives the following useful result.

Theorem 9.3.6

If a finite dimensional vector space V is the direct sum $V = U \oplus W$ of subspaces U and W, then

$$\dim V = \dim U + \dim W$$

These direct sum decompositions of V play an important role in any discussion of invariant subspaces. If $T : V \to V$ is a linear operator and if U_1 is a T-invariant subspace, the block upper triangular matrix

$$M_B(T) = \begin{bmatrix} M_{B_1}(T) & Y \\ 0 & Z \end{bmatrix} \tag{9.3}$$

in Theorem 9.3.1 is achieved by choosing any basis $B_1 = \{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ of U_1 and completing it to a basis $B = \{\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ of V in any way at all. The fact that U_1 is T-invariant ensures that the first k columns of $M_B(T)$ have the form in (9.3) (that is, the last $n-k$ entries are zero), and the question arises whether the additional basis vectors $\mathbf{b}_{k+1}, \dots, \mathbf{b}_n$ can be chosen such that

$$U_2 = \text{span}\{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$$

is *also* T-invariant. In other words, does each T-invariant subspace of V have a T-invariant complement? Unfortunately the answer in general is no (see Example 9.3.11 below); but when it is possible, the matrix $M_B(T)$ simplifies further. The assumption that the complement $U_2 = \text{span}\{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ is T-invariant too means that $Y = 0$ in equation 9.3 above, and that $Z = M_{B_2}(T)$ is the matrix of the restriction of T to U_2 (where $B_2 = \{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$). The verification is the same as in the proof of Theorem 9.3.1.

Theorem 9.3.7

Let $T : V \to V$ be a linear operator where V has dimension n. Suppose $V = U_1 \oplus U_2$ where both U_1 and U_2 are T-invariant. If $B_1 = \{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ and $B_2 = \{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ are bases of U_1 and U_2 respectively, then

$$B = \{\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$$

is a basis of V, and $M_B(T)$ has the block diagonal form

$$M_B(T) = \begin{bmatrix} M_{B_1}(T) & 0 \\ 0 & M_{B_2}(T) \end{bmatrix}$$

where $M_{B_1}(T)$ and $M_{B_2}(T)$ are the matrices of the restrictions of T to U_1 and to U_2 respectively.

Definition 9.8 Reducible Linear Operator

The linear operator $T : V \to V$ is said to be **reducible** *if nonzero T-invariant subspaces U_1 and U_2 can be found such that $V = U_1 \oplus U_2$.*

Then T has a matrix in block diagonal form as in Theorem 9.3.7, and the study of T is reduced to studying its restrictions to the lower-dimensional spaces U_1 and U_2. If these can be determined, so can T. Here is an example in which the action of T on the invariant subspaces U_1 and U_2 is very simple indeed. The result for operators is used to derive the corresponding similarity theorem for matrices.

Example 9.3.10

Let $T : V \to V$ be a linear operator satisfying $T^2 = 1_V$ (such operators are called **involutions**). Define

$$U_1 = \{\mathbf{v} \mid T(\mathbf{v}) = \mathbf{v}\} \quad \text{and} \quad U_2 = \{\mathbf{v} \mid T(\mathbf{v}) = -\mathbf{v}\}$$

a. Show that $V = U_1 \oplus U_2$.

b. If $\dim V = n$, find a basis B of V such that $M_B(T) = \left[\begin{array}{cc} I_k & 0 \\ 0 & -I_{n-k} \end{array}\right]$ for some k.

c. Conclude that, if A is an $n \times n$ matrix such that $A^2 = I$, then A is similar to $\left[\begin{array}{cc} I_k & 0 \\ 0 & -I_{n-k} \end{array}\right]$ for some k.

Solution.

a. The verification that U_1 and U_2 are subspaces of V is left to the reader. If $\mathbf{v}$ lies in $U_1 \cap U_2$, then $\mathbf{v} = T(\mathbf{v}) = -\mathbf{v}$, and it follows that $\mathbf{v} = \mathbf{0}$. Hence $U_1 \cap U_2 = \{\mathbf{0}\}$. Given $\mathbf{v}$ in V, write

$$\mathbf{v} = \tfrac{1}{2}\{[\mathbf{v} + T(\mathbf{v})] + [\mathbf{v} - T(\mathbf{v})]\}$$

Then $\mathbf{v} + T(\mathbf{v})$ lies in U_1, because $T[\mathbf{v} + T(\mathbf{v})] = T(\mathbf{v}) + T^2(\mathbf{v}) = \mathbf{v} + T(\mathbf{v})$. Similarly, $\mathbf{v} - T(\mathbf{v})$ lies in U_2, and it follows that $V = U_1 + U_2$. This proves part (a).

b. U_1 and U_2 are easily shown to be T-invariant, so the result follows from Theorem 9.3.7 if bases $B_1 = \{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ and $B_2 = \{\mathbf{b}_{k+1}, \dots, \mathbf{b}_n\}$ of U_1 and U_2 can be found such that $M_{B_1}(T) = I_k$ and $M_{B_2}(T) = -I_{n-k}$. But this is true for *any* choice of B_1 and B_2:

$$\begin{aligned} M_{B_1}(T) &= \left[\begin{array}{cccc} C_{B_1}[T(\mathbf{b}_1)] & C_{B_1}[T(\mathbf{b}_2)] & \cdots & C_{B_1}[T(\mathbf{b}_k)] \end{array}\right] \\ &= \left[\begin{array}{cccc} C_{B_1}(\mathbf{b}_1) & C_{B_1}(\mathbf{b}_2) & \cdots & C_{B_1}(\mathbf{b}_k) \end{array}\right] \\ &= I_k \end{aligned}$$

A similar argument shows that $M_{B_2}(T) = -I_{n-k}$, so part (b) follows with $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$.

c. Given A such that $A^2 = I$, consider $T_A : \mathbb{R}^n \to \mathbb{R}^n$. Then $(T_A)^2(\mathbf{x}) = A^2\mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$, so $(T_A)^2 = 1_V$. Hence, by part (b), there exists a basis B of $\mathbb{R}^n$ such that

$$M_B(T_A) = \begin{bmatrix} I_r & 0 \\ 0 & -I_{n-r} \end{bmatrix}$$

But Theorem 9.2.4 shows that $M_B(T_A) = P^{-1}AP$ for some invertible matrix P, and this proves part (c).

Note that the passage from the result for operators to the analogous result for matrices is routine and can be carried out in any situation, as in the verification of part (c) of Example 9.3.10. The key is the analysis of the operators. In this case, the involutions are just the operators satisfying $T^2 = 1_V$, and the simplicity of this condition means that the invariant subspaces U_1 and U_2 are easy to find.

Unfortunately, not every linear operator $T : V \to V$ is reducible. In fact, the linear operator in Example 9.3.4 has *no* invariant subspaces except 0 and V. On the other hand, one might expect that this is the only type of nonreducible operator; that is, if the operator *has* an invariant subspace that is not 0 or V, then *some* invariant complement must exist. The next example shows that even this is not valid.

Example 9.3.11

Consider the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by $T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+b \\ b \end{bmatrix}$. Show that $U_1 = \mathbb{R}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is T-invariant but that U_1 has not T-invariant complement in $\mathbb{R}^2$.

Solution. Because $U_1 = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$ and $T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, it follows (by Example 9.3.3) that U_1 is T-invariant. Now assume, if possible, that U_1 has a T-invariant complement U_2 in $\mathbb{R}^2$. Then $U_1 \oplus U_2 = \mathbb{R}^2$ and $T(U_2) \subseteq U_2$. Theorem 9.3.6 gives

$$2 = \dim \mathbb{R}^2 = \dim U_1 + \dim U_2 = 1 + \dim U_2$$

so $\dim U_2 = 1$. Let $U_2 = \mathbb{R}\mathbf{u}_2$, and write $\mathbf{u}_2 = \begin{bmatrix} p \\ q \end{bmatrix}$. We claim that $\mathbf{u}_2$ is not in U_1. For if $\mathbf{u}_2 \in U_1$, then $\mathbf{u}_2 \in U_1 \cap U_2 = \{\mathbf{0}\}$, so $\mathbf{u}_2 = \mathbf{0}$. But then $U_2 = \mathbb{R}\mathbf{u}_2 = \{\mathbf{0}\}$, a contradiction, as $\dim U_2 = 1$. So $\mathbf{u}_2 \notin U_1$, from which $q \neq 0$. On the other hand, $T(\mathbf{u}_2) \in U_2 = \mathbb{R}\mathbf{u}_2$ (because U_2 is T-invariant), say $T(\mathbf{u}_2) = \lambda\mathbf{u}_2 = \lambda\begin{bmatrix} p \\ q \end{bmatrix}$.

Thus

$$\begin{bmatrix} p+q \\ q \end{bmatrix} = T\begin{bmatrix} p \\ q \end{bmatrix} = \lambda\begin{bmatrix} p \\ q \end{bmatrix} \quad \text{where } \lambda \in \mathbb{R}$$

Hence $p+q = \lambda p$ and $q = \lambda q$. Because $q \neq 0$, the second of these equations implies that $\lambda = 1$, so the first equation implies $q = 0$, a contradiction. So a T-invariant complement of U_1 does not exist.

This is as far as we take the theory here, but in Chapter 11 the techniques introduced in this section will be refined to show that every matrix is similar to a very nice matrix indeed—its Jordan canonical form.

Exercises for 9.3

Exercise 9.3.1 If $T : V \to V$ is any linear operator, show that $\ker T$ and $\operatorname{im} T$ are T-invariant subspaces.

Exercise 9.3.2 Let T be a linear operator on V. If U and W are T-invariant, show that

a. $U \cap W$ and $U + W$ are also T-invariant.

b. $T(U)$ is T-invariant.

Exercise 9.3.3 Let S and T be linear operators on V and assume that $ST = TS$.

a. Show that $\operatorname{im} S$ and $\ker S$ are T-invariant.

b. If U is T-invariant, show that $S(U)$ is T-invariant.

Exercise 9.3.4 Let $T : V \to V$ be a linear operator. Given $\mathbf{v}$ in V, let U denote the set of vectors in V that lie in every T-invariant subspace that contains $\mathbf{v}$.

a. Show that U is a T-invariant subspace of V containing $\mathbf{v}$.

b. Show that U is contained in every T-invariant subspace of V that contains $\mathbf{v}$.

Exercise 9.3.5

a. If T is a scalar operator (see Example 7.1.1) show that every subspace is T-invariant.

b. Conversely, if every subspace is T-invariant, show that T is scalar.

Exercise 9.3.6 Show that the only subspaces of V that are T-invariant for every operator $T : V \to V$ are 0 and V. Assume that V is finite dimensional. [*Hint*: Theorem 7.1.3.]

Exercise 9.3.7 Suppose that $T : V \to V$ is a linear operator and that U is a T-invariant subspace of V. If S is an invertible operator, put $T' = STS^{-1}$. Show that $S(U)$ is a T'-invariant subspace.

Exercise 9.3.8 In each case, show that U is T-invariant, use it to find a block upper triangular matrix for T, and use that to compute $c_T(x)$.

a. $T : \mathbf{P}_2 \to \mathbf{P}_2$,
$T(a + bx + cx^2)$
$= (-a + 2b + c) + (a + 3b + c)x + (a + 4b)x^2$,
$U = \operatorname{span}\{1,\ x + x^2\}$

b. $T : \mathbf{P}_2 \to \mathbf{P}_2$,
$T(a + bx + cx^2)$
$= (5a - 2b + c) + (5a - b + c)x + (a + 2c)x^2$,
$U = \operatorname{span}\{1 - 2x^2,\ x + x^2\}$

Exercise 9.3.9 In each case, show that $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ has no invariant subspaces except 0 and $\mathbb{R}^2$.

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

b. $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, 0 < \theta < \pi$

Exercise 9.3.10 In each case, show that $V = U \oplus W$.

a. $V = \mathbb{R}^4$, $U = \operatorname{span}\{(1,\ 1,\ 0,\ 0),\ (0,\ 1,\ 1,\ 0)\}$,
$W = \operatorname{span}\{(0,\ 1,\ 0,\ 1),\ (0,\ 0,\ 1,\ 1)\}$

b. $V = \mathbb{R}^4$, $U = \{(a,\ a,\ b,\ b) \mid a,\ b \text{ in } \mathbb{R}\}$,
$W = \{(c,\ d,\ c,\ -d) \mid c,\ d \text{ in } \mathbb{R}\}$

c. $V = \mathbf{P}_3$, $U = \{a + bx \mid a,\ b \text{ in } \mathbb{R}\}$,
$W = \{ax^2 + bx^3 \mid a,\ b \text{ in } \mathbb{R}\}$

d. $V = \mathbf{M}_{22}$, $U = \left\{ \begin{bmatrix} a & a \\ b & b \end{bmatrix} \middle| a,\ b \text{ in } \mathbb{R} \right\}$,
$W = \left\{ \begin{bmatrix} a & b \\ -a & b \end{bmatrix} \middle| a,\ b \text{ in } \mathbb{R} \right\}$

Exercise 9.3.11 Let $U = \operatorname{span}\{(1,\ 0,\ 0,\ 0),\ (0,\ 1,\ 0,\ 0)\}$ in $\mathbb{R}^4$. Show that $\mathbb{R}^4 = U \oplus W_1$ and $\mathbb{R}^4 = U \oplus W_2$, where $W_1 = \operatorname{span}\{(0,\ 0,\ 1,\ 0),\ (0,\ 0,\ 0,\ 1)\}$ and $W_2 = \operatorname{span}\{(1,\ 1,\ 1,\ 1),\ (1,\ 1,\ 1,\ -1)\}$.

Exercise 9.3.12 Let U be a subspace of V, and suppose that $V = U \oplus W_1$ and $V = U \oplus W_2$ hold for subspaces W_1 and W_2. Show that $\dim W_1 = \dim W_2$.

Exercise 9.3.13 If U and W denote the subspaces of even and odd polynomials in $\mathbf{P}_n$, respectively, show that $\mathbf{P}_n = U \oplus W$. (See Exercise 6.3.36.) [*Hint*: $f(x) + f(-x)$ is even.]

Exercise 9.3.14 Let E be an $n \times n$ matrix with $E^2 = E$. Show that $\mathbf{M}_{nn} = U \oplus W$, where $U = \{A \mid AE = A\}$ and $W = \{B \mid BE = 0\}$. [*Hint*: XE lies in U for every matrix X.]

Exercise 9.3.15 Let U and W be subspaces of V. Show that $U \cap W = \{\mathbf{0}\}$ if and only if $\{\mathbf{u}, \mathbf{w}\}$ is independent for all $\mathbf{u} \neq \mathbf{0}$ in U and all $\mathbf{w} \neq \mathbf{0}$ in W.

Exercise 9.3.16 Let $V \xrightarrow{T} W \xrightarrow{S} V$ be linear transformations, and assume that $\dim V$ and $\dim W$ are finite.

a. If $ST = 1_V$, show that $W = \text{im } T \oplus \ker S$. [*Hint*: Given $\mathbf{w}$ in W, show that $\mathbf{w} - TS(\mathbf{w})$ lies in $\ker S$.]

b. Illustrate with $\mathbb{R}^2 \xrightarrow{T} \mathbb{R}^3 \xrightarrow{S} \mathbb{R}^2$ where $T(x, y) = (x, y, 0)$ and $S(x, y, z) = (x, y)$.

Exercise 9.3.17 Let U and W be subspaces of V, let $\dim V = n$, and assume that $\dim U + \dim W = n$.

a. If $U \cap W = \{\mathbf{0}\}$, show that $V = U \oplus W$.

b. If $U + W = V$, show that $V = U \oplus W$. [*Hint*: Theorem 6.4.5.]

Exercise 9.3.18 Let $A = \left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array} \right]$ and consider $T_A : \mathbb{R}^2 \to \mathbb{R}^2$.

a. Show that the only eigenvalue of T_A is $\lambda = 0$.

b. Show that $\ker(T_A) = \mathbb{R} \left[\begin{array}{c} 1 \\ 0 \end{array} \right]$ is the unique T_A-invariant subspace of $\mathbb{R}^2$ (except for 0 and $\mathbb{R}^2$).

Exercise 9.3.19 If $A = \left[\begin{array}{cccc} 2 & -5 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 1 & 1 \end{array} \right]$, show that $T_A : \mathbb{R}^4 \to \mathbb{R}^4$ has two-dimensional T-invariant subspaces U and W such that $\mathbb{R}^4 = U \oplus W$, but A has no real eigenvalue.

Exercise 9.3.20 Let $T : V \to V$ be a linear operator where $\dim V = n$. If U is a T-invariant subspace of V, let $T_1 : U \to U$ denote the restriction of T to U (so $T_1(\mathbf{u}) = T(\mathbf{u})$ for all $\mathbf{u}$ in U). Show that $c_T(x) = c_{T_1}(x) \cdot q(x)$ for some polynomial $q(x)$. [*Hint*: Theorem 9.3.1.]

Exercise 9.3.21 Let $T : V \to V$ be a linear operator where $\dim V = n$. Show that V has a basis of eigenvectors if and only if V has a basis B such that $M_B(T)$ is diagonal.

Exercise 9.3.22 In each case, show that $T^2 = 1$ and find (as in Example 9.3.10) an ordered basis B such that $M_B(T)$ has the given block form.

a. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$ where $T(A) = A^T$, $M_B(T) = \left[\begin{array}{cc} I_3 & 0 \\ 0 & -1 \end{array} \right]$

b. $T : \mathbf{P}_3 \to \mathbf{P}_3$ where $T[p(x)] = p(-x)$, $M_B(T) = \left[\begin{array}{cc} I_2 & 0 \\ 0 & -I_2 \end{array} \right]$

c. $T : \mathbb{C} \to \mathbb{C}$ where $T(a + bi) = a - bi$, $M_B(T) = \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right]$

d. $T : \mathbb{R}^3 \to \mathbb{R}^3$ where $T(a, b, c) = (-a + 2b + c, b + c, -c)$, $M_B(T) = \left[\begin{array}{cc} 1 & 0 \\ 0 & -I_2 \end{array} \right]$

e. $T : V \to V$ where $T(\mathbf{v}) = -\mathbf{v}$, $\dim V = n$, $M_B(T) = -I_n$

Exercise 9.3.23 Let U and W denote subspaces of a vector space V.

a. If $V = U \oplus W$, define $T : V \to V$ by $T(\mathbf{v}) = \mathbf{w}$ where $\mathbf{v}$ is written (uniquely) as $\mathbf{v} = \mathbf{u} + \mathbf{w}$ with $\mathbf{u}$ in U and $\mathbf{w}$ in W. Show that T is a linear transformation, $U = \ker T$, $W = \text{im } T$, and $T^2 = T$.

b. Conversely, if $T : V \to V$ is a linear transformation such that $T^2 = T$, show that $V = \ker T \oplus \text{im } T$. [*Hint*: $\mathbf{v} - T(\mathbf{v})$ lies in $\ker T$ for all $\mathbf{v}$ in V.]

Exercise 9.3.24 Let $T : V \to V$ be a linear operator satisfying $T^2 = T$ (such operators are called **idempotents**). Define $U_1 = \{\mathbf{v} \mid T(\mathbf{v}) = \mathbf{v}\}$, $U_2 = \ker T = \{\mathbf{v} \mid T(\mathbf{v}) = 0\}$.

a. Show that $V = U_1 \oplus U_2$.

b. If $\dim V = n$, find a basis B of V such that $M_B(T) = \left[\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array} \right]$, where $r = \text{rank } T$.

c. If A is an $n \times n$ matrix such that $A^2 = A$, show that A is similar to $\left[\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array} \right]$, where $r = \text{rank } A$. [*Hint*: Example 9.3.10.]

Exercise 9.3.25 In each case, show that $T^2 = T$ and find (as in the preceding exercise) an ordered basis B such that $M_B(T)$ has the form given (0_k is the $k \times k$ zero matrix).

a. $T : \mathbf{P}_2 \to \mathbf{P}_2$ where $T(a+bx+cx^2) = (a-b+c)(1+x+x^2)$, $M_B(T) = \left[\begin{array}{cc} 1 & 0 \\ 0 & 0_2 \end{array} \right]$

b. $T : \mathbb{R}^3 \to \mathbb{R}^3$ where $T(a,\ b,\ c) = (a+2b,\ 0,\ 4b+c)$, $M_B(T) = \left[\begin{array}{cc} I_2 & 0 \\ 0 & 0 \end{array} \right]$

c. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$ where $T \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] = \left[\begin{array}{rr} -5 & -15 \\ 2 & 6 \end{array} \right] \left[\begin{array}{cc} a & b \\ c & d \end{array} \right]$, $M_B(T) = \left[\begin{array}{cc} I_2 & 0 \\ 0 & 0_2 \end{array} \right]$

Exercise 9.3.26 Let $T : V \to V$ be an operator satisfying $T^2 = cT$, $c \neq 0$.

a. Show that $V = U \oplus \ker T$, where $U = \{\mathbf{u} \mid T(\mathbf{u}) = c\mathbf{u}\}$. [*Hint*: Compute $T(\mathbf{v} - \frac{1}{c}T(\mathbf{v}))$.]

b. If $\dim V = n$, show that V has a basis B such that $M_B(T) = \left[\begin{array}{cc} cI_r & 0 \\ 0 & 0 \end{array} \right]$, where $r = \text{rank } T$.

c. If A is any $n \times n$ matrix of rank r such that $A^2 = cA$, $c \neq 0$, show that A is similar to $\left[\begin{array}{cc} cI_r & 0 \\ 0 & 0 \end{array} \right]$.

Exercise 9.3.27 Let $T : V \to V$ be an operator such that $T^2 = c^2$, $c \neq 0$.

a. Show that $V = U_1 \oplus U_2$, where $U_1 = \{\mathbf{v} \mid T(\mathbf{v}) = c\mathbf{v}\}$ and $U_2 = \{\mathbf{v} \mid T(\mathbf{v}) = -c\mathbf{v}\}$. [*Hint*: $\mathbf{v} = \frac{1}{2c}\{[T(\mathbf{v}) + c\mathbf{v}] - [T(\mathbf{v}) - c\mathbf{v}]\}$.]

b. If $\dim V = n$, show that V has a basis B such that $M_B(T) = \left[\begin{array}{cc} cI_k & 0 \\ 0 & -cI_{n-k} \end{array} \right]$ for some k.

c. If A is an $n \times n$ matrix such that $A^2 = c^2 I$, $c \neq 0$, show that A is similar to $\left[\begin{array}{cc} cI_k & 0 \\ 0 & -cI_{n-k} \end{array} \right]$ for some k.

Exercise 9.3.28 If P is a fixed $n \times n$ matrix, define $T : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$ by $T(A) = PA$. Let U_j denote the subspace of $\mathbf{M}_{nn}$ consisting of all matrices with all columns zero except possibly column j.

a. Show that each U_j is T-invariant.

b. Show that $\mathbf{M}_{nn}$ has a basis B such that $M_B(T)$ is block diagonal with each block on the diagonal equal to P.

Exercise 9.3.29 Let V be a vector space. If $f : V \to \mathbb{R}$ is a linear transformation and $\mathbf{z}$ is a vector in V, define $T_{f,\ \mathbf{z}} : V \to V$ by $T_{f,\ \mathbf{z}}(\mathbf{v}) = f(\mathbf{v})\mathbf{z}$ for all $\mathbf{v}$ in V. Assume that $f \neq 0$ and $\mathbf{z} \neq \mathbf{0}$.

a. Show that $T_{f,\ \mathbf{z}}$ is a linear operator of rank 1.

b. If $f \neq 0$, show that $T_{f,\ \mathbf{z}}$ is an idempotent if and only if $f(\mathbf{z}) = 1$. (Recall that $T : V \to V$ is called an idempotent if $T^2 = T$.)

c. Show that every idempotent $T : V \to V$ of rank 1 has the form $T = T_{f,\ \mathbf{z}}$ for some $f : V \to \mathbb{R}$ and some $\mathbf{z}$ in V with $f(\mathbf{z}) = 1$. [*Hint*: Write $\text{im } T = \mathbb{R}\mathbf{z}$ and show that $T(\mathbf{z}) = \mathbf{z}$. Then use Exercise 9.3.23.]

Exercise 9.3.30 Let U be a fixed $n \times n$ matrix, and consider the operator $T : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$ given by $T(A) = UA$.

a. Show that λ is an eigenvalue of T if and only if it is an eigenvalue of U.

b. If λ is an eigenvalue of T, show that $E_\lambda(T)$ consists of all matrices whose columns lie in $E_\lambda(U)$:
$$E_\lambda(T) = \{ \left[\begin{array}{cccc} P_1 & P_2 & \cdots & P_n \end{array} \right] \mid P_i \text{ in } E_\lambda(U) \text{ for each } i \}$$

c. Show if $\dim [E_\lambda(U)] = d$, then $\dim [E_\lambda(T)] = nd$. [*Hint*: If $B = \{\mathbf{x}_1,\ \dots,\ \mathbf{x}_d\}$ is a basis of $E_\lambda(U)$, consider the set of all matrices with one column from B and the other columns zero.]

Exercise 9.3.31 Let $T : V \to V$ be a linear operator where V is finite dimensional. If $U \subseteq V$ is a subspace, let $\overline{U} = \{\mathbf{u}_0 + T(\mathbf{u}_1) + T^2(\mathbf{u}_2) + \cdots + T^k(\mathbf{u}_k) \mid \mathbf{u}_i \text{ in } U,\ k \geq 0\}$. Show that $\overline{U}$ is the smallest T-invariant subspace containing U (that is, it is T-invariant, contains U, and is contained in every such subspace).

Exercise 9.3.32 Let $U_1, \dots, U_m$ be subspaces of V and assume that $V = U_1 + \cdots + U_m$; that is, every $\mathbf{v}$ in V can be written (in at least one way) in the form $\mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_m$, $\mathbf{u}_i$ in U_i. Show that the following conditions are equivalent.

i. If $\mathbf{u}_1 + \cdots + \mathbf{u}_m = \mathbf{0}$, $\mathbf{u}_i$ in U_i, then $\mathbf{u}_i = \mathbf{0}$ for each i.

ii. If $\mathbf{u}_1 + \cdots + \mathbf{u}_m = \mathbf{u}'_1 + \cdots + \mathbf{u}'_m$, $\mathbf{u}_i$ and $\mathbf{u}'_i$ in U_i, then $\mathbf{u}_i = \mathbf{u}'_i$ for each i.

iii. $U_i \cap (U_1 + \cdots + U_{i-1} + U_{i+1} + \cdots + U_m) = \{\mathbf{0}\}$ for each $i = 1, 2, \dots, m$.

iv. $U_i \cap (U_{i+1} + \cdots + U_m) = \{\mathbf{0}\}$ for each $i = 1, 2, \dots, m-1$.

When these conditions are satisfied, we say that V is the **direct sum** of the subspaces U_i, and write $V = U_1 \oplus U_2 \oplus \cdots \oplus U_m$.

Exercise 9.3.33

a. Let B be a basis of V and let $B = B_1 \cup B_2 \cup \cdots \cup B_m$ where the B_i are pairwise disjoint, nonempty subsets of B. If $U_i = \operatorname{span} B_i$ for each i, show that $V = U_1 \oplus U_2 \oplus \cdots \oplus U_m$ (preceding exercise).

b. Conversely if $V = U_1 \oplus \cdots \oplus U_m$ and B_i is a basis of U_i for each i, show that $B = B_1 \cup \cdots \cup B_m$ is a basis of V as in (a).

Exercise 9.3.34 Let $T : V \to V$ be an operator where $T^3 = 0$. If $\mathbf{u} \in V$ and $U = \operatorname{span}\{\mathbf{u},\ T(\mathbf{u}),\ T^2(\mathbf{u})\}$, show that U is T-invariant and has dimension 3.