Lecture 7

6.1 Introduction

Consider a basis $S = \{u_1, u_2, \ldots, u_n\}$ of a vector space V over a field K. For any vector $v \in V$, suppose

$$v = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$$

Then the coordinate vector of v relative to the basis S, which we assume to be a column vector (unless otherwise stated or implied), is denoted and defined by

$$[v]_S = [a_1, a_2, \ldots, a_n]^T$$

Recall (Section 4.11) that the mapping $v \mapsto [v]_S$, determined by the basis S, is an isomorphism between V and K^n.

This chapter shows that there is also an isomorphism, determined by the basis S, between the algebra $A(V)$ of linear operators on V and the algebra M of n-square matrices over K. Thus, every linear mapping $F: V \to V$ will correspond to an n-square matrix $[F]_S$ determined by the basis S. We will also show how our matrix representation changes when we choose another basis.

6.2 Matrix Representation of a Linear Operator

Let T be a linear operator (transformation) from a vector space V into itself, and suppose $S = \{u_1, u_2, \ldots, u_n\}$ is a basis of V. Now $T(u_1), T(u_2), \ldots, T(u_n)$ are vectors in V, and so each is a linear combination of the vectors in the basis S; say,

$$\begin{aligned} T(u_1) &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ T(u_2) &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T(u_n) &= a_{n1}u_1 + a_{n2}u_2 + \cdots + a_{nn}u_n \end{aligned}$$

The following definition applies.

DEFINITION: The transpose of the above matrix of coefficients, denoted by $m_S(T)$ or $[T]_S$, is called the *matrix representation* of T relative to the basis S, or simply the matrix of T in the basis S. (The subscript S may be omitted if the basis S is understood.)

Using the coordinate (column) vector notation, the matrix representation of T may be written in the form

$$m_S(T) = [T]_S = \left[[T(u_1)]_S,\ [T(u_2)]_S,\ \ldots,\ [T(u_1)]_S \right]$$

That is, the columns of $m(T)$ are the coordinate vectors of $T(u_1), T(u_2), \ldots, T(u_n)$, respectively.

EXAMPLE 6.1 Let $F: \mathbf{R}^2 \to \mathbf{R}^2$ be the linear operator defined by $F(x, y) = (2x + 3y,\ 4x - 5y)$.

(a) Find the matrix representation of F relative to the basis $S = \{u_1, u_2\} = \{(1,2),\ (2,5)\}$.

(1) First find $F(u_1)$, and then write it as a linear combination of the basis vectors u_1 and u_2. (For notational convenience, we use column vectors.) We have

$$F(u_1) = F\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 8 \\ -6 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 2 \\ 5 \end{bmatrix} \quad \text{and} \quad \begin{aligned} x + 2y &= 8 \\ 2x + 5y &= -6 \end{aligned}$$

Solve the system to obtain $x = 52$, $y = -22$. Hence, $F(u_1) = 52u_1 - 22u_2$.

(2) Next find $F(u_2)$, and then write it as a linear combination of u_1 and u_2:

$$F(u_2) = F\left(\begin{bmatrix} 2 \\ 5 \end{bmatrix}\right) = \begin{bmatrix} 19 \\ -17 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 2 \\ 5 \end{bmatrix} \quad \text{and} \quad \begin{aligned} x + 2y &= 19 \\ 2x + 5y &= -17 \end{aligned}$$

Solve the system to get $x = 129$, $y = -55$. Thus, $F(u_2) = 129u_1 - 55u_2$.

Now write the coordinates of $F(u_1)$ and $F(u_2)$ as columns to obtain the matrix

$$[F]_S = \begin{bmatrix} 52 & 129 \\ -22 & -55 \end{bmatrix}$$

(b) Find the matrix representation of F relative to the (usual) basis $E = \{e_1, e_2\} = \{(1,0),\ (0,1)\}$.

Find $F(e_1)$ and write it as a linear combination of the usual basis vectors e_1 and e_2, and then find $F(e_2)$ and write it as a linear combination of e_1 and e_2. We have

$$\begin{aligned} F(e_1) &= F(1,0) = (2,2) \quad\ \ = 2e_1 + 4e_2 \\ F(e_2) &= F(0,1) = (3,-5) = 3e_1 - 5e_2 \end{aligned} \quad \text{and so} \quad [F]_E = \begin{bmatrix} 2 & 3 \\ 4 & -5 \end{bmatrix}$$

Note that the coordinates of $F(e_1)$ and $F(e_2)$ form the columns, not the rows, of $[F]_E$. Also, note that the arithmetic is much simpler using the usual basis of $\mathbf{R}^2$.

EXAMPLE 6.2 Let V be the vector space of functions with basis $S = \{\sin t, \cos t, e^{3t}\}$, and let $\mathbf{D}: V \to V$ be the differential operator defined by $\mathbf{D}(f(t)) = d(f(t))/dt$. We compute the matrix representing $\mathbf{D}$ in the basis S:

$$\begin{aligned}
\mathbf{D}(\sin t) &= \quad \cos t = \quad 0(\sin t) + 1(\cos t) + 0(e^3 t) \\
\mathbf{D}(\cos\ t) &= -\sin t = -1(\sin t) + 0(\cos t) + 0(e^{3t}) \\
\mathbf{D}(e^{3t}) &= \quad 3e^{3t} = \quad 0(\sin t) + 0(\cos t) + 3(e^{3t})
\end{aligned}$$

and so

$$[\mathbf{D}] = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Note that the coordinates of $\mathbf{D}(\sin t)$, $\mathbf{D}(\cos t)$, $\mathbf{D}(e^{3t})$ form the columns, not the rows, of $[\mathbf{D}]$.

Matrix Mappings and Their Matrix Representation

Consider the following matrix A, which may be viewed as a linear operator on $\mathbf{R}^2$, and basis S of $\mathbf{R}^2$:

$$A = \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix} \quad \text{and} \quad S = \{u_1, u_2\} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix} \right\}$$

(We write vectors as columns, because our map is a matrix.) We find the matrix representation of A relative to the basis S.

(1) First we write $A(u_1)$ as a linear combination of u_1 and u_2. We have

$$A(u_1) = \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -6 \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \end{bmatrix} + y \begin{bmatrix} 2 \\ 5 \end{bmatrix} \quad \text{and so} \quad \begin{aligned} x + 2y &= -1 \\ 2x + 5y &= -6 \end{aligned}$$

Solving the system yields $x = 7, y = -4$. Thus, $A(u_1) = 7u_1 - 4u_2$.

(2) Next we write $A(u_2)$ as a linear combination of u_1 and u_2. We have

$$A(u_2) = \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} -4 \\ -7 \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \end{bmatrix} + y \begin{bmatrix} 2 \\ 5 \end{bmatrix} \quad \text{and so} \quad \begin{aligned} x + 2y &= -4 \\ 2x + 5y &= -7 \end{aligned}$$

Solving the system yields $x = -6$, $y = 1$. Thus, $A(u_2) = -6u_1 + u_2$. Writing the coordinates of $A(u_1)$ and $A(u_2)$ as columns gives us the following matrix representation of A:

$$[A]_S = \begin{bmatrix} 7 & -6 \\ -4 & 1 \end{bmatrix}$$

Remark: Suppose we want to find the matrix representation of A relative to the usual basis $E = \{e_1, e_2\} = \{[1, 0]^T, [0, 1]^T\}$ of $\mathbf{R}^2$. We have

$$\begin{aligned} A(e_1) &= \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 3e_1 + 4e_2 \\ A(e_2) &= \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -5 \end{bmatrix} = -2e_1 - 5e_2 \end{aligned} \quad \text{and so} \quad [A]_E = \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix}$$

Note that $[A]_E$ is the original matrix A. This result is true in general:

Note that $[A]_E$ is the original matrix A. This result is true in general:

> The matrix representation of any $n \times n$ square matrix A over a field K relative to the usual basis E of K^n is the matrix A itself; that is,
> $$[A]_E = A$$

Algorithm for Finding Matrix Representations

Next follows an algorithm for finding matrix representations. The first Step 0 is optional. It may be useful to use it in Step 1(b), which is repeated for each basis vector.

ALGORITHM 6.1: The input is a linear operator T on a vector space V and a basis $S = \{u_1, u_2, \ldots, u_n\}$ of V. The output is the matrix representation $[T]_S$.

Step 0. Find a formula for the coordinates of an arbitrary vector v relative to the basis S.

Step 1. Repeat for each basis vector u_k in S:

(a) Find $T(u_k)$.

(b) Write $T(u_k)$ as a linear combination of the basis vectors $u_1, u_2, \ldots, u_n$.

Step 2. Form the matrix $[T]_S$ whose columns are the coordinate vectors in Step 1(b).

EXAMPLE 6.3 Let $F\colon \mathbf{R}^2 \to \mathbf{R}^2$ be defined by $F(x,y) = (2x + 3y,\ 4x - 5y)$. Find the matrix representation $[F]_S$ of F relative to the basis $S = \{u_1, u_2\} = \{(1, -2),\ (2, -5)\}$.

(Step 0) First find the coordinates of $(a, b) \in \mathbf{R}^2$ relative to the basis S. We have

$$\begin{bmatrix} a \\ b \end{bmatrix} = x\begin{bmatrix} 1 \\ -2 \end{bmatrix} + y\begin{bmatrix} 2 \\ -5 \end{bmatrix} \quad \text{or} \quad \begin{aligned} x + 2y &= a \\ -2x - 5y &= b \end{aligned} \quad \text{or} \quad \begin{aligned} x + 2y &= a \\ -y &= 2a + b \end{aligned}$$

Solving for x and y in terms of a and b yields $x = 5a + 2b$, $y = -2a - b$. Thus,

$$(a, b) = (5a + 2b)u_1 + (-2a - b)u_2$$

(Step 1) Now we find $F(u_1)$ and write it as a linear combination of u_1 and u_2 using the above formula for (a, b), and then we repeat the process for $F(u_2)$. We have

$$F(u_1) = F(1, -2) = (-4, 14) = 8u_1 - 6u_2$$
$$F(u_2) = F(2, -5) = (-11, 33) = 11u_1 - 11u_2$$

(Step 2) Finally, we write the coordinates of $F(u_1)$ and $F(u_2)$ as columns to obtain the required matrix:

$$[F]_S = \begin{bmatrix} 8 & 11 \\ -6 & -11 \end{bmatrix}$$

Properties of Matrix Representations

This subsection gives the main properties of the matrix representations of linear operators T on a vector space V. We emphasize that we are always given a particular basis S of V.

Our first theorem, proved in Problem 6.9, tells us that the "action" of a linear operator T on a vector v is preserved by its matrix representation.

THEOREM 6.1: Let $T: V \to V$ be a linear operator, and let S be a (finite) basis of V. Then, for any vector v in V, $[T]_S[v]_S = [T(v)]_S$.

EXAMPLE 6.4 Consider the linear operator F on R^2 and the basis S of Example 6.3; that is,

$$F(x,y) = (2x + 3y, \quad 4x - 5y) \qquad \text{and} \qquad S = \{u_1, u_2\} = \{(1,-2), \quad (2,-5)\}$$

Let

$$v = (5,-7), \qquad \text{and so} \qquad F(v) = (-11, 55)$$

Using the formula from Example 6.3, we get

$$[v] = [11, -3]^T \qquad \text{and} \qquad [F(v)] = [55, -33]^T$$

We verify Theorem 6.1 for this vector v (where $[F]$ is obtained from Example 6.3):

$$[F][v] = \begin{bmatrix} 8 & 11 \\ -6 & -11 \end{bmatrix} \begin{bmatrix} 11 \\ -3 \end{bmatrix} = \begin{bmatrix} 55 \\ -33 \end{bmatrix} = [F(v)]$$

Given a basis S of a vector space V, we have associated a matrix $[T]$ to each linear operator T in the algebra $A(V)$ of linear operators on V. Theorem 6.1 tells us that the "action" of an individual linear operator T is preserved by this representation. The next two theorems (proved in Problems 6.10 and 6.11) tell us that the three basic operations in $A(V)$ with these operators—namely (i) addition, (ii) scalar multiplication, and (iii) composition—are also preserved.

THEOREM 6.2: Let V be an n-dimensional vector space over K, let S be a basis of V, and let $\mathbf{M}$ be the algebra of $n \times n$ matrices over K. Then the mapping

$$m: A(V) \to \mathbf{M} \qquad \text{defined by} \qquad m(T) = [T]_S$$

is a vector space isomorphism. That is, for any $F, G \in A(V)$ and any $k \in K$,

(i) $m(F+G) = m(F) + m(G)$ or $[F+G] = [F] + [G]$

(ii) $m(kF) = km(F)$ or $[kF] = k[F]$

(iii) m is bijective (one-to-one and onto).

THEOREM 6.3: For any linear operators $F, G \in A(V)$,

$$m(G \circ F) = m(G)m(F) \text{ or } [G \circ F] = [G][F]$$

(Here $G \circ F$ denotes the composition of the maps G and F.)

6.3 Change of Basis

Let V be an n-dimensional vector space over a field K. We have shown that once we have selected a basis S of V, every vector $v \in V$ can be represented by means of an n-tuple $[v]_S$ in K^n, and every linear operator T in $A(V)$ can be represented by an $n \times n$ matrix over K. We ask the following natural question:

How do our representations change if we select another basis?

In order to answer this question, we first need a definition.

DEFINITION: Let $S = \{u_1, u_2, \ldots, u_n\}$ be a basis of a vector space V, and let $S' = \{v_1, v_2, \ldots, v_n\}$ be another basis. (For reference, we will call S the "old" basis and S' the "new" basis.) Because S is a basis, each vector in the "new" basis S' can be written uniquely as a linear combination of the vectors in S; say,

$$
\begin{aligned}
v_1 &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\
v_2 &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
v_n &= a_{n1}u_1 + a_{n2}u_2 + \cdots + a_{nn}u_n
\end{aligned}
$$

Let P be the transpose of the above matrix of coefficients; that is, let $P = [p_{ij}]$, where $p_{ij} = a_{ji}$. Then P is called the *change-of-basis matrix* (or *transition matrix*) from the "old" basis S to the "new" basis S'.

The following remarks are in order.

Remark 1: The above change-of-basis matrix P may also be viewed as the matrix whose columns are, respectively, the coordinate column vectors of the "new" basis vectors v_i relative to the "old" basis S; namely,

$$P = [[v_1]_S, [v_2]_S, \ldots, [v_n]_S]$$

Remark 2: Analogously, there is a change-of-basis matrix Q from the "new" basis S' to the "old" basis S. Similarly, Q may be viewed as the matrix whose columns are, respectively, the coordinate column vectors of the "old" basis vectors u_i relative to the "new" basis S'; namely,

$$Q = [[u_1]_{S'}, [u_2]_{S'}, \ldots, [u_n]_{S'}]$$

Remark 3: Because the vectors $v_1, v_2, \ldots, v_n$ in the new basis S' are linearly independent, the matrix P is invertible (Problem 6.18). Similarly, Q is invertible. In fact, we have the following proposition (proved in Problem 6.18).

PROPOSITION 6.4: Let P and Q be the above change-of-basis matrices. Then $Q = P^{-1}$.

Now suppose $S = \{u_1, u_2, \ldots, u_n\}$ is a basis of a vector space V, and suppose $P = [p_{ij}]$ is any nonsingular matrix. Then the n vectors

$$v_i = p_{1i}u_i + p_{2i}u_2 + \cdots + p_{ni}u_n, \qquad i = 1, 2, \ldots, n$$

corresponding to the columns of P, are linearly independent [Problem 6.21(a)]. Thus, they form another basis S' of V. Moreover, P will be the change-of-basis matrix from S to the new basis S'.

EXAMPLE 6.5 Consider the following two bases of $\mathbf{R}^2$:

$$S = \{u_1, u_2\} = \{(1,2),\ (3,5)\} \qquad \text{and} \qquad S' = \{v_1, v_2\} = \{(1,-1),\ (1,-2)\}$$

(a) Find the change-of-basis matrix P from S to the "new" basis S'.

Write each of the new basis vectors of S' as a linear combination of the original basis vectors u_1 and u_2 of S. We have

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 3 \\ 5 \end{bmatrix} \quad \text{or} \quad \begin{aligned} x + 3y &= 1 \\ 2x + 5y &= -1 \end{aligned} \quad \text{yielding} \quad x = -8, \quad y = 3$$

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix} + y\begin{bmatrix} 3 \\ 5 \end{bmatrix} \quad \text{or} \quad \begin{aligned} x + 3y &= 1 \\ 2x + 5y &= -1 \end{aligned} \quad \text{yielding} \quad x = -11, \quad y = 4$$

Thus,

$$\begin{aligned} v_1 &= -8u_1 + 3u_2 \\ v_2 &= -11u_1 + 4u_2 \end{aligned} \qquad \text{and hence,} \qquad P = \begin{bmatrix} -8 & -11 \\ 3 & 4 \end{bmatrix}.$$

Note that the coordinates of v_1 and v_2 are the columns, not rows, of the change-of-basis matrix P.

(b) Find the change-of-basis matrix Q from the "new" basis S' back to the "old" basis S.

Here we write each of the "old" basis vectors u_1 and u_2 of S' as a linear combination of the "new" basis vectors v_1 and v_2 of S'. This yields

$$\begin{aligned} u_1 &= 4v_1 - 3v_2 \\ u_2 &= 11v_1 - 8v_2 \end{aligned} \qquad \text{and hence,} \qquad Q = \begin{bmatrix} 4 & 11 \\ -3 & -8 \end{bmatrix}$$

As expected from Proposition 6.4, $Q = P^{-1}$. (In fact, we could have obtained Q by simply finding P^{-1}.)

EXAMPLE 6.6 Consider the following two bases of $\mathbf{R}^3$:

$$E = \{e_1, e_2, e_3\} = \{(1,0,0), \quad (0,1,0), \quad (0,0,1)\}$$

and

$$S = \{u_1, u_2, u_3\} = \{(1,0,1), \quad (2,1,2), \quad (1,2,2)\}$$

(a) Find the change-of-basis matrix P from the basis E to the basis S.

Because E is the usual basis, we can immediately write each basis element of S as a linear combination of the basis elements of E. Specifically,

$$\begin{aligned} u_1 &= (1,0,1) = e_1 + e_3 \\ u_2 &= (2,1,2) = 2e_1 + e_2 + 2e_3 \\ u_3 &= (1,2,2) = e_1 + 2e_2 + 2e_3 \end{aligned} \qquad \text{and hence,} \qquad P = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

Again, the coordinates of u_1, u_2, u_3 appear as the columns in P. Observe that P is simply the matrix whose columns are the basis vectors of S. This is true only because the original basis was the usual basis E.

(b) Find the change-of-basis matrix Q from the basis S to the basis E.

The definition of the change-of-basis matrix Q tells us to write each of the (usual) basis vectors in E as a linear combination of the basis elements of S. This yields

$$\begin{aligned} e_1 &= (1,0,0) = -2u_1 + 2u_2 - u_3 \\ e_2 &= (0,1,0) = -2u_1 + u_2 \\ e_3 &= (0,0,1) = 3u_1 - 2u_2 + u_3 \end{aligned} \qquad \text{and hence,} \qquad Q = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix}$$

We emphasize that to find Q, we need to solve three 3×3 systems of linear equations—one 3×3 system for each of e_1, e_2, e_3.

Alternatively, we can find $Q = P^{-1}$ by forming the matrix $M = [P, I]$ and row reducing M to row canonical form:

$$M = \begin{bmatrix} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 2 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -2 & -2 & 3 \\ 0 & 1 & 0 & 2 & 1 & -2 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{bmatrix} = [I, P^{-1}]$$

thus,

$$Q = P^{-1} = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix}$$

(Here we have used the fact that Q is the inverse of P.)

The result in Example 6.6(a) is true in general. We state this result formally, because it occurs often.

PROPOSITION 6.5: The change-of-basis matrix from the usual basis E of K^n to any basis S of K^n is the matrix P whose columns are, respectively, the basis vectors of S.

Applications of Change-of-Basis Matrix

First we show how a change of basis affects the coordinates of a vector in a vector space V. The following theorem is proved in Problem 6.22.

THEOREM 6.6: Let P be the change-of-basis matrix from a basis S to a basis S' in a vector space V. Then, for any vector $v \in V$, we have

$$P[v]_{S'} = [v]_S \qquad \text{and hence,} \qquad P^{-1}[v]_S = [v]_{S'}$$

Namely, if we multiply the coordinates of v in the original basis S by P^{-1}, we get the coordinates of v in the new basis S'.

Remark 1: Although P is called the change-of-basis matrix from the old basis S to the new basis S', we emphasize that P^{-1} transforms the coordinates of v in the original basis S into the coordinates of v in the new basis S'.

Remark 2: Because of the above theorem, many texts call $Q = P^{-1}$, not P, the transition matrix from the old basis S to the new basis S'. Some texts also refer to Q as the *change-of-coordinates* matrix.

We now give the proof of the above theorem for the special case that dim $V = 3$. Suppose P is the change-of-basis matrix from the basis $S = \{u_1, u_2, u_3\}$ to the basis $S' = \{v_1, v_2, v_3\}$; say,

$$\begin{aligned} v_1 &= a_1u_1 + a_2u_2 + a_3a_3 \\ v_2 &= b_1u_1 + b_2u_2 + b_3u_3 \\ v_3 &= c_1u_1 + c_2u_2 + c_3u_3 \end{aligned} \qquad \text{and hence,} \qquad P = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

Now suppose $v \in V$ and, say, $v = k_1v_1 + k_2v_2 + k_3v_3$. Then, substituting for v_1, v_2, v_3 from above, we obtain

$$\begin{aligned} v &= k_1(a_1u_1 + a_2u_2 + a_3u_3) + k_2(b_1u_1 + b_2u_2 + b_3u_3) + k_3(c_1u_1 + c_2u_2 + c_3u_3) \\ &= (a_1k_1 + b_1k_2 + c_1k_3)u_1 + (a_2k_1 + b_2k_2 + c_2k_3)u_2 + (a_3k_1 + b_3k_2 + c_3k_3)u_3 \end{aligned}$$

Thus,

$$[v]_{S'} = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} \qquad \text{and} \qquad [v]_S = \begin{bmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \\ a_3k_1 + b_3k_2 + c_3k_3 \end{bmatrix}$$

Accordingly,

$$P[v]_{S'} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} a_1k_1 + b_1k_2 + c_1k_3 \\ a_2k_1 + b_2k_2 + c_2k_3 \\ a_3k_1 + b_3k_2 + c_3k_3 \end{bmatrix} = [v]_S$$

Finally, multiplying the equation $[v]_S = P[v]_S$, by P^{-1}, we get

$$P^{-1}[v]_S = P^{-1}P[v]_{S'} = I[v]_{S'} = [v]_{S'}$$

The next theorem (proved in Problem 6.26) shows how a change of basis affects the matrix representation of a linear operator.

THEOREM 6.7: Let P be the change-of-basis matrix from a basis S to a basis S' in a vector space V. Then, for any linear operator T on V,

$$[T]_{S'} = P^{-1}[T]_S P$$

That is, if A and B are the matrix representations of T relative, respectively, to S and S', then

$$B = P^{-1}AP$$

EXAMPLE 6.7 Consider the following two bases of $\mathbf{R}^3$:

$$E = \{e_1, e_2, e_3\} = \{(1,0,0), \quad (0,1,0), \quad (0,0,1)\}$$

and

$$S = \{u_1, u_2, u_3\} = \{(1,0,1), \quad (2,1,2), \quad (1,2,2)\}$$

The change-of-basis matrix P from E to S and its inverse P^{-1} were obtained in Example 6.6.

(a) Write $v = (1,3,5)$ as a linear combination of u_1, u_2, u_3, or, equivalently, find $[v]_S$.

One way to do this is to directly solve the vector equation $v = xu_1 + yu_2 + zu_3$; that is,

$$\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + y \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + z \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \qquad \text{or} \qquad \begin{aligned} x + 2y + z &= 1 \\ y + 2z &= 3 \\ x + 2y + 2z &= 5 \end{aligned}$$

The solution is $x = 7, \quad y = -5, \quad z = 4$, so $v = 7u_1 - 5u_2 + 4u_3$.

On the other hand, we know that $[v]_E = [1,3,5]^T$, because E is the usual basis, and we already know P^{-1}. Therefore, by Theorem 6.6,

$$[v]_S = P^{-1}[v]_E = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 7 \\ -5 \\ 4 \end{bmatrix}$$

Thus, again, $v = 7u_1 - 5u_2 + 4u_3$.

(b) Let $A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & -4 & 1 \\ 3 & -1 & 2 \end{bmatrix}$, which may be viewed as a linear operator on $\mathbf{R}^3$. Find the matrix B that represents A relative to the basis S.

The definition of the matrix representation of A relative to the basis S tells us to write each of $A(u_1)$, $A(u_2)$, $A(u_3)$ as a linear combination of the basis vectors u_1, u_2, u_3 of S. This yields

$$\begin{aligned} A(u_1) &= (-1, 3, 5) = 11u_1 - 5u_2 + 6u_3 \\ A(u_2) &= (1, 2, 9) \quad = 21u_1 - 14u_2 + 8u_3 \\ A(u_3) &= (3, -4, 5) = 17u_1 - 8e_2 + 2u_3 \end{aligned} \qquad \text{and hence,} \qquad B = \begin{bmatrix} 11 & 21 & 17 \\ -5 & -14 & -8 \\ 6 & 8 & 2 \end{bmatrix}$$

We emphasize that to find B, we need to solve three 3×3 systems of linear equations—one 3×3 system for each of $A(u_1)$, $A(u_2)$, $A(u_3)$.

On the other hand, because we know P and P^{-1}, we can use Theorem 6.7. That is,

$$B = P^{-1}AP = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & -2 \\ 2 & -4 & 1 \\ 3 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 11 & 21 & 17 \\ -5 & -14 & -8 \\ 6 & 8 & 2 \end{bmatrix}$$

This, as expected, gives the same result.

6.4 Similarity

Suppose A and B are square matrices for which there exists an invertible matrix P such that $B = P^{-1}AP$; then B is said to be *similar* to A, or B is said to be obtained from A by a *similarity transformation.* We show (Problem 6.29) that similarity of matrices is an equivalence relation.

By Theorem 6.7 and the above remark, we have the following basic result.

THEOREM 6.8: Two matrices represent the same linear operator if and only if the matrices are similar.

That is, all the matrix representations of a linear operator T form an equivalence class of similar matrices.

A linear operator T is said to be *diagonalizable* if there exists a basis S of V such that T is represented by a diagonal matrix; the basis S is then said to *diagonalize* T. The preceding theorem gives us the following result.

THEOREM 6.9: Let A be the matrix representation of a linear operator T. Then T is diagonalizable if and only if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

That is, T is diagonalizable if and only if its matrix representation can be diagonalized by a similarity transformation.

We emphasize that not every operator is diagonalizable. However, we will show (Chapter 10) that every linear operator can be represented by certain ''standard'' matrices called its *normal* or *canonical* forms. Such a discussion will require some theory of fields, polynomials, and determinants.

Functions and Similar Matrices

Suppose f is a function on square matrices that assigns the same value to similar matrices; that is, $f(A) = f(B)$ whenever A is similar to B. Then f induces a function, also denoted by f, on linear operators T in the following natural way. We define

$$f(T) = f([T]_S)$$

where S is any basis. By Theorem 6.8, the function is well defined.

The determinant (Chapter 8) is perhaps the most important example of such a function. The trace (Section 2.7) is another important example of such a function.

EXAMPLE 6.8 Consider the following linear operator F and bases E and S of $\mathbf{R}^2$:

$$F(x, y) = (2x + 3y,\ 4x - 5y), \qquad E = \{(1, 0),\ (0, 1)\}, \qquad S = \{(1, 2),\ (2, 5)\}$$

By Example 6.1, the matrix representations of F relative to the bases E and S are, respectively,

$$A = \begin{bmatrix} 2 & 3 \\ 4 & -5 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 52 & 129 \\ -22 & -55 \end{bmatrix}$$

Using matrix A, we have

(i) Determinant of $F = \det(A) = -10 - 12 = -22$; (ii) Trace of $F = \operatorname{tr}(A) = 2 - 5 = -3$.

On the other hand, using matrix B, we have

(i) Determinant of $F = \det(B) = -2860 + 2838 = -22$; (ii) Trace of $F = \operatorname{tr}(B) = 52 - 55 = -3$.

As expected, both matrices yield the same result.

6.5 Matrices and General Linear Mappings

Last, we consider the general case of linear mappings from one vector space into another. Suppose V and U are vector spaces over the same field K and, say, $\dim V = m$ and $\dim U = n$. Furthermore, suppose

$$S = \{v_1, v_2, \ldots, v_m\} \qquad \text{and} \qquad S' = \{u_1, u_2, \ldots, u_n\}$$

are arbitrary but fixed bases, respectively, of V and U.

Suppose $F: V \to U$ is a linear mapping. Then the vectors $F(v_1), F(v_2), \ldots, F(v_m)$ belong to U, and so each is a linear combination of the basis vectors in S'; say,

$$\begin{aligned} F(v_1) &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ F(v_2) &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ F(v_m) &= a_{m1}u_1 + a_{m2}u_2 + \cdots + a_{mn}u_n \end{aligned}$$

DEFINITION: The transpose of the above matrix of coefficients, denoted by $m_{S,S'}(F)$ or $[F]_{S,S'}$, is called the *matrix representation* of F relative to the bases S and S'. [We will use the simple notation $m(F)$ and $[F]$ when the bases are understood.]

The following theorem is analogous to Theorem 6.1 for linear operators (Problem 6.67).

THEOREM 6.10: For any vector $v \in V$, $[F]_{S,S'}[v]_S = [F(v)]_{S'}$.

That is, multiplying the coordinates of v in the basis S of V by $[F]$, we obtain the coordinates of $F(v)$ in the basis S' of U.

Recall that for any vector spaces V and U, the collection of all linear mappings from V into U is a vector space and is denoted by $\text{Hom}(V, U)$. The following theorem is analogous to Theorem 6.2 for linear operators, where now we let $\mathbf{M} = \mathbf{M}_{m,n}$ denote the vector space of all $m \times n$ matrices (Problem 6.67).

THEOREM 6.11: The mapping $m\colon \text{Hom}(V, U) \to \mathbf{M}$ defined by $m(F) = [F]$ is a vector space isomorphism. That is, for any $F, G \in \text{Hom}(V, U)$ and any scalar k,

(i) $m(F + G) = m(F) + m(G)$ or $[F + G] = [F] + [G]$

(ii) $m(kF) = km(F)$ or $[kF] = k[F]$

(iii) m is bijective (one-to-one and onto).

Our next theorem is analogous to Theorem 6.3 for linear operators (Problem 6.67).

THEOREM 6.12: Let S, S', S'' be bases of vector spaces V, U, W, respectively. Let $F\colon V \to U$ and $G \circ U \to W$ be linear mappings. Then

$$[G \circ F]_{S,S''} = [G]_{S',S''}[F]_{S,S'}$$

That is, relative to the appropriate bases, the matrix representation of the composition of two mappings is the matrix product of the matrix representations of the individual mappings.

Next we show how the matrix representation of a linear mapping $F\colon V \to U$ is affected when new bases are selected (Problem 6.67).

THEOREM 6.13: Let P be the change-of-basis matrix from a basis e to a basis e' in V, and let Q be the change-of-basis matrix from a basis f to a basis f' in U. Then, for any linear map $F: V \to U$,

$$[F]_{e',f'} = Q^{-1}[F]_{e,f}P$$

In other words, if A is the matrix representation of a linear mapping F relative to the bases e and f, and B is the matrix representation of F relative to the bases e' and f', then

$$B = Q^{-1}AP$$

Our last theorem, proved in Problem 6.36, shows that any linear mapping from one vector space V into another vector space U can be represented by a very simple matrix. We note that this theorem is analogous to Theorem 3.18 for $m \times n$ matrices.

THEOREM 6.14: Let $F: V \to U$ be linear and, say, $\text{rank}(F) = r$. Then there exist bases of V and U such that the matrix representation of F has the form

$$A = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

where I_r is the r-square identity matrix.

The above matrix A is called the *normal* or *canonical* form of the linear map F.