3. Linear Transformations

3.1. Linear Transformations

We shall now introduce linear transformations, the objects which we shall study in most of the remainder of this book. The reader may find it helpful to read (or reread) the discussion of functions in the Appendix, since we shall freely use the terminology of that discussion.

Definition. *Let* V *and* W *be vector spaces over the field* F. *A* **linear transformation from** V **into** W *is a function* T *from* V *into* W *such that*

$$T(c\alpha + \beta) = c(T\alpha) + T\beta$$

for all α *and* β *in* V *and all scalars* c *in* F.

Example 1. If V is any vector space, the identity transformation I, defined by $I\alpha = \alpha$, is a linear transformation from V into V. The **zero transformation** 0, defined by $0\alpha = 0$, is a linear transformation from V into V.

Example 2. Let F be a field and let V be the space of polynomial functions f from F into F, given by

$$f(x) = c_0 + c_1x + \cdots + c_kx^k.$$

Let

$$(Df)(x) = c_1 + 2c_2x + \cdots + kc_kx^{k-1}.$$

Then D is a linear transformation from V into V—the differentiation transformation.

EXAMPLE 3. Let A be a fixed $m \times n$ matrix with entries in the field F. The function T defined by $T(X) = AX$ is a linear transformation from $F^{n\times 1}$ into $F^{m\times 1}$. The function U defined by $U(\alpha) = \alpha A$ is a linear transformation from F^m into F^n.

EXAMPLE 4. Let P be a fixed $m \times m$ matrix with entries in the field F and let Q be a fixed $n \times n$ matrix over F. Define a function T from the space $F^{m\times n}$ into itself by $T(A) = PAQ$. Then T is a linear transformation from $F^{m\times n}$ into $F^{m\times n}$, because

$$\begin{aligned} T(cA + B) &= P(cA + B)Q \\ &= (cPA + PB)Q \\ &= cPAQ + PBQ \\ &= cT(A) + T(B). \end{aligned}$$

EXAMPLE 5. Let R be the field of real numbers and let V be the space of all functions from R into R which are *continuous*. Define T by

$$(Tf)(x) = \int_0^x f(t)\, dt.$$

Then T is a linear transformation from V into V. The function Tf is not only continuous but has a continuous first derivative. The linearity of integration is one of its fundamental properties.

The reader should have no difficulty in verifying that the transformations defined in Examples 1, 2, 3, and 5 are linear transformations. We shall expand our list of examples considerably as we learn more about linear transformations.

It is important to note that if T is a linear transformation from V into W, then $T(0) = 0$; one can see this from the definition because

$$T(0) = T(0 + 0) = T(0) + T(0).$$

This point is often confusing to the person who is studying linear algebra for the first time, since he probably has been exposed to a slightly different use of the term 'linear function.' A brief comment should clear up the confusion. Suppose V is the vector space R^1. A linear transformation from V into V is then a particular type of real-valued function on the real line R. In a calculus course, one would probably call such a function linear if its graph is a straight line. A linear transformation from R^1 into R^1, according to our definition, will be a function from R into R, the graph of which is a straight line *passing through the origin*.

In addition to the property $T(0) = 0$, let us point out another property of the general linear transformation T. Such a transformation 'preserves' linear combinations; that is, if $\alpha_1, \ldots, \alpha_n$ are vectors in V and $c_1, \ldots, c_n$ are scalars, then

$$T(c_1\alpha_1 + \cdots + c_n\alpha_n) = c_1(T\alpha_1) + \cdots + c_n(T\alpha_n).$$

This follows readily from the definition. For example,

$$\begin{aligned} T(c_1\alpha_1 + c_2\alpha_2) &= c_1(T\alpha_1) + T(c_2\alpha_2) \\ &= c_1(T\alpha_1) + c_2(T\alpha_2). \end{aligned}$$

Theorem 1. *Let V be a finite-dimensional vector space over the field F and let $\{\alpha_1, \ldots, \alpha_n\}$ be an ordered basis for V. Let W be a vector space over the same field F and let $\beta_1, \ldots, \beta_n$ be any vectors in W. Then there is precisely one linear transformation T from V into W such that*

$$T\alpha_j = \beta_j, \qquad j = 1, \ldots, n.$$

Proof. To prove there is some linear transformation T with $T\alpha_j = \beta_j$ we proceed as follows. Given α in V, there is a unique n-tuple $(x_1, \ldots, x_n)$ such that

$$\alpha = x_1\alpha_1 + \cdots + x_n\alpha_n.$$

For this vector α we define

$$T\alpha = x_1\beta_1 + \cdots + x_n\beta_n.$$

Then T is a well-defined rule for associating with each vector α in V a vector $T\alpha$ in W. From the definition it is clear that $T\alpha_j = \beta_j$ for each j. To see that T is linear, let

$$\beta = y_1\alpha_1 + \cdots + y_n\alpha_n$$

be in V and let c be any scalar. Now

$$c\alpha + \beta = (cx_1 + y_1)\alpha_1 + \cdots + (cx_n + y_n)\alpha_n$$

and so by definition

$$T(c\alpha + \beta) = (cx_1 + y_1)\beta_1 + \cdots + (cx_n + y_n)\beta_n.$$

On the other hand,

$$\begin{aligned} c(T\alpha) + T\beta &= c\sum_{i=1}^{n} x_i\beta_i + \sum_{i=1}^{n} y_i\beta_i \\ &= \sum_{i=1}^{n} (cx_i + y_i)\beta_i \end{aligned}$$

and thus

$$T(c\alpha + \beta) = c(T\alpha) + T\beta.$$

If U is a linear transformation from V into W with $U\alpha_j = \beta_j$, $j = 1, \ldots, n$, then for the vector $\alpha = \sum_{i=1}^{n} x_i\alpha_i$ we have

$$\begin{aligned} U\alpha &= U\left(\sum_{i=1}^{n} x_i\alpha_i\right) \\ &= \sum_{i=1}^{n} x_i(U\alpha_i) \\ &= \sum_{i=1}^{n} x_i\beta_i \end{aligned}$$

so that U is exactly the rule T which we defined above. This shows that the linear transformation T with $T\alpha_j = \beta_j$ is unique. ∎

Theorem 1 is quite elementary; however, it is so basic that we have stated it formally. The concept of function is very general. If V and W are (non-zero) vector spaces, there is a multitude of functions from V into W. Theorem 1 helps to underscore the fact that the functions which are linear are extremely special.

EXAMPLE 6. The vectors

$$\alpha_1 = (1, 2)$$
$$\alpha_2 = (3, 4)$$

are linearly independent and therefore form a basis for R^2. According to Theorem 1, there is a unique linear transformation from R^2 into R^3 such that

$$T\alpha_1 = (3, 2, 1)$$
$$T\alpha_2 = (6, 5, 4).$$

If so, we must be able to find $T(\epsilon_1)$. We find scalars c_1, c_2 such that $\epsilon_1 = c_1\alpha_1 + c_2\alpha_2$ and then we know that $T\epsilon_1 = c_1T\alpha_1 + c_2T\alpha_2$. If $(1, 0) = c_1(1, 2) + c_2(3, 4)$ then $c_1 = -2$ and $c_2 = 1$. Thus

$$\begin{aligned} T(1, 0) &= -2(3, 2, 1) + (6, 5, 4) \\ &= (0, 1, 2). \end{aligned}$$

EXAMPLE 7. Let T be a linear transformation from the m-tuple space F^m into the n-tuple space F^n. Theorem 1 tells us that T is uniquely determined by the sequence of vectors $\beta_1, \ldots, \beta_m$ where

$$\beta_i = T\epsilon_i, \quad i = 1, \ldots, m.$$

In short, T is uniquely determined by the images of the standard basis vectors. The determination is

$$\alpha = (x_1, \ldots, x_m)$$
$$T\alpha = x_1\beta_1 + \cdots + x_m\beta_m.$$

If B is the $m \times n$ matrix which has row vectors $\beta_1, \ldots, \beta_m$, this says that

$$T\alpha = \alpha B.$$

In other words, if $\beta_i = (B_{i1}, \ldots, B_{in})$, then

$$T(x_1, \ldots, x_n) = [x_1 \cdots x_m] \begin{bmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & & \vdots \\ B_{m1} & \cdots & B_{mn} \end{bmatrix}.$$

This is a very explicit description of the linear transformation. In Section 3.4 we shall make a serious study of the relationship between linear transformations and matrices. We shall not pursue the particular description $T\alpha = \alpha B$ because it has the matrix B on the right of the vector α, and that can lead to some confusion. The point of this example is to show that we can give an explicit and reasonably simple description of all linear transformations from F^m into F^n.

If T is a linear transformation from V into W, then the range of T is not only a subset of W; it is a subspace of W. Let R_T be the range of T, that is, the set of all vectors β in W such that $\beta = T\alpha$ for some α in V. Let β_1 and β_2 be in R_T and let c be a scalar. There are vectors α_1 and α_2 in V such that $T\alpha_1 = \beta_1$ and $T\alpha_2 = \beta_2$. Since T is linear

$$\begin{aligned} T(c\alpha_1 + \alpha_2) &= cT\alpha_1 + T\alpha_2 \\ &= c\beta_1 + \beta_2, \end{aligned}$$

which shows that $c\beta_1 + \beta_2$ is also in R_T.

Another interesting subspace associated with the linear transformation T is the set N consisting of the vectors α in V such that $T\alpha = 0$. It is a subspace of V because

(a) $T(0) = 0$, so that N is non-empty;
(b) if $T\alpha_1 = T\alpha_2 = 0$, then

$$\begin{aligned} T(c\alpha_1 + \alpha_2) &= cT\alpha_1 + T\alpha_2 \\ &= c0 + 0 \\ &= 0 \end{aligned}$$

so that $c\alpha_1 + \alpha_2$ is in N.

Definition. *Let* V *and* W *be vector spaces over the field* F *and let* T *be a linear transformation from* V *into* W. *The* **null space** *of* T *is the set of all vectors* α *in* V *such that* $T\alpha = 0$.

If V *is finite-dimensional, the* **rank** *of* T *is the dimension of the range of* T *and the* **nullity** *of* T *is the dimension of the null space of* T.

The following is one of the most important results in linear algebra.

Theorem 2. *Let* V *and* W *be vector spaces over the field* F *and let* T *be a linear transformation from* V *into* W. *Suppose that* V *is finite-dimensional. Then*

$$\text{rank}\ (T) + \text{nullity}\ (T) = \dim V.$$

Proof. Let $\{\alpha_1, \ldots, \alpha_k\}$ be a basis for N, the null space of T. There are vectors $\alpha_{k+1}, \ldots, \alpha_n$ in V such that $\{\alpha_1, \ldots, \alpha_n\}$ is a basis for V. We shall now prove that $\{T\alpha_{k+1}, \ldots, T\alpha_n\}$ is a basis for the range of T. The vectors $T\alpha_1, \ldots, T\alpha_n$ certainly span the range of T, and since $T\alpha_j = 0$, for $j \leq k$, we see that $T\alpha_{k+1}, \ldots, T\alpha_n$ span the range. To see that these vectors are independent, suppose we have scalars c_i such that

$$\sum_{i=k+1}^{n} c_i(T\alpha_i) = 0.$$

This says that

$$T\left(\sum_{i=k+1}^{n} c_i\alpha_i\right) = 0$$

and accordingly the vector $\alpha = \sum_{i=k+1}^{n} c_i\alpha_i$ is in the null space of T. Since $\alpha_1, \ldots, \alpha_k$ form a basis for N, there must be scalars $b_1, \ldots, b_k$ such that

$$\alpha = \sum_{i=1}^{k} b_i\alpha_i.$$

Thus

$$\sum_{i=1}^{k} b_i\alpha_i - \sum_{j=k+1}^{n} c_j\alpha_j = 0$$

and since $\alpha_1, \ldots, \alpha_n$ are linearly independent we must have

$$b_1 = \cdots = b_k = c_{k+1} = \cdots = c_n = 0.$$

If r is the rank of T, the fact that $T\alpha_{k+1}, \ldots, T\alpha_n$ form a basis for the range of T tells us that $r = n - k$. Since k is the nullity of T and n is the dimension of V, we are done. ∎

Theorem 3. *If* A *is an* m × n *matrix with entries in the field* F, *then*

row rank (A) = *column rank* (A).

Proof. Let T be the linear transformation from $F^{n\times 1}$ into $F^{m\times 1}$ defined by $T(X) = AX$. The null space of T is the solution space for the system $AX = 0$, i.e., the set of all column matrices X such that $AX = 0$. The range of T is the set of all $m \times 1$ column matrices Y such that $AX = Y$ has a solution for X. If $A_1, \ldots, A_n$ are the columns of A, then

$$AX = x_1A_1 + \cdots + x_nA_n$$

so that the range of T is the subspace spanned by the columns of A. In other words, the range of T is the column space of A. Therefore,

$$\text{rank } (T) = \text{column rank } (A).$$

Theorem 2 tells us that if S is the solution space for the system $AX = 0$, then

$$\dim S + \text{column rank } (A) = n.$$

We now refer to Example 15 of Chapter 2. Our deliberations there showed that, if r is the dimension of the row space of A, then the solution space S has a basis consisting of $n - r$ vectors:

$$\dim S = n - \text{row rank } (A).$$

It is now apparent that

$$\text{row rank } (A) = \text{column rank } (A). \quad ∎$$

The proof of Theorem 3 which we have just given depends upon explicit calculations concerning systems of linear equations. There is a more conceptual proof which does not rely on such calculations. We shall give such a proof in Section 3.7.

Exercises

1. Which of the following functions T from R^2 into R^2 are linear transformations?

(a) $T(x_1, x_2) = (1 + x_1, x_2)$;
(b) $T(x_1, x_2) = (x_2, x_1)$;
(c) $T(x_1, x_2) = (x_1^2, x_2)$;
(d) $T(x_1, x_2) = (\sin x_1, x_2)$;
(e) $T(x_1, x_2) = (x_1 - x_2, 0)$.

2. Find the range, rank, null space, and nullity for the zero transformation and the identity transformation on a finite-dimensional space V.

3. Describe the range and the null space for the differentiation transformation of Example 2. Do the same for the integration transformation of Example 5.

4. Is there a linear transformation T from R^3 into R^2 such that $T(1, -1, 1) = (1, 0)$ and $T(1, 1, 1) = (0, 1)$?

5. If

$$\begin{aligned} \alpha_1 &= (1, -1), & \beta_1 &= (1, 0) \\ \alpha_2 &= (2, -1), & \beta_2 &= (0, 1) \\ \alpha_3 &= (-3, 2), & \beta_3 &= (1, 1) \end{aligned}$$

is there a linear transformation T from R^2 into R^2 such that $T\alpha_i = \beta_i$ for $i = 1, 2$ and 3?

6. Describe explicitly (as in Exercises 1 and 2) the linear transformation T from F^2 into F^2 such that $T\epsilon_1 = (a, b)$, $T\epsilon_2 = (c, d)$.

7. Let F be a subfield of the complex numbers and let T be the function from F^3 into F^3 defined by

$$T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2, -x_1 - 2x_2 + 2x_3).$$

(a) Verify that T is a linear transformation.

(b) If (a, b, c) is a vector in F^3, what are the conditions on a, b, and c that the vector be in the range of T? What is the rank of T?

(c) What are the conditions on a, b, and c that (a, b, c) be in the null space of T? What is the nullity of T?

8. Describe explicitly a linear transformation from R^3 into R^3 which has as its range the subspace spanned by $(1, 0, -1)$ and $(1, 2, 2)$.

9. Let V be the vector space of all $n \times n$ matrices over the field F, and let B be a fixed $n \times n$ matrix. If

$$T(A) = AB - BA$$

verify that T is a linear transformation from V into V.

10. Let V be the set of all complex numbers regarded as a vector space over the