Chapter 3

Linear Transformations

In your previous mathematics courses you undoubtedly studied real-valued functions of one or more variables. For example, when you discussed parabolas the function $f(x) = x^2$ appeared, or when you talked abut straight lines the function $f(x) = 2x$ arose. In this chapter we study functions of several variables, that is, functions of vectors. Moreover, their values will be vectors rather than scalars. The particular transformations that we study also satisfy a "linearity" condition that will be made precise later.

3.1 Definition and Examples

Before defining a linear transformation we look at two examples. The first is not a linear transformation and the second one is.

Example 1. Let $V = \mathbb{R}^2$ and let $W = \mathbb{R}$. Define $f\colon\ V \to W$ by $f(x_1, x_2) = x_1 x_2$. Thus, f is a function defined on a vector space of dimension 2, with values in a one-dimensional space. The notation is highly suggestive; that is, $f\colon\ V \to W$ indicates that f does something to a vector in V to get a vector in W. For example, $f(1, -1) = -1$, $f(1, 2) = 2$, etc. We will see later that this function does not satisfy the "linearity" condition and hence is not a linear transformation. $\square$

Example 2. Let $V = \mathbb{R}^2$ and $W = \mathbb{R}^3$. Define $L\colon\ V \to W$ by $L(x_1, x_2) = (x_1, x_2 - x_1, x_2)$. Here the function L takes a vector in $\mathbb{R}^2$ and transforms it into a vector in $\mathbb{R}^3$. For example, $L(1, -1) = (1, -2, -1)$ and $L(2, 6) = (2, 4, 6)$. This particular function satisfies the linearity condition below, and so would be called a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$. $\square$

Definition 3.1. Let V and W be two vector spaces. Let L be a function defined on V with values in W. L will be called a linear transformation if it satisfies the following two conditions:

1. $L(\boldsymbol{x} + \boldsymbol{y}) = L(\boldsymbol{x}) + L(\boldsymbol{y})$, for any two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in V.

2. $L(c\boldsymbol{x}) = cL(\boldsymbol{x})$, for any scalar c and vector $\boldsymbol{x}$ in V.

Let's go back to Example 2 and verify that we did indeed have a linear transformation. For any $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{y} = (y_1, y_2)$, we have

$$\begin{aligned} L(\boldsymbol{x}+\boldsymbol{y}) &= L[(x_1+y_1, x_2+y_2)] = [x_1+y_1, (x_2+y_2)-(x_1+y_1), x_2+y_2] \\ &= (x_1, x_2-x_1, x_2) + (y_1, y_2-y_1, y_2) = L(\boldsymbol{x}) + L(\boldsymbol{y}). \end{aligned}$$

Thus, condition 1 holds. Moreover

$$L(c\boldsymbol{x}) = L[(cx_1, cx_2)] = (cx_1, cx_2 - cx_1, cx_2) = c(x_1, x_2 - x_1, x_2) = cL(\boldsymbol{x})$$

Since 1 and 2 hold, L is a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$. The reader should now check that the function in Example 1 does not satisfy either of these two conditions.

Example 3. Define $L\colon\ \mathbb{R}^3 \to \mathbb{R}^2$ by $L(x_1, x_2, x_3) = (x_3 - x_1, x_1 + x_2)$.

a. Compute $L(\boldsymbol{e}_1), L(\boldsymbol{e}_2)$, and $L(\boldsymbol{e}_3)$.

b. Show L is a linear transformation.

c. Show $L(x_1, x_2, x_3) = x_1L(\boldsymbol{e}_1) + x_2L(\boldsymbol{e}_2) + x_3L(\boldsymbol{e}_3)$.

a. $L[(1,0,0)] = (-1,1), L[(0,1,0)] = (0,1), L[(0,0,1)] = (1,0)$.

b.
$$\begin{aligned} L(\boldsymbol{x}+\boldsymbol{y}) &= L[(x_1+y_1, x_2+y_2, x_3+y_3)] \\ &= ((x_3+y_3)-(x_1+y_1), (x_1+y_1)+(x_2+y_2)) \\ &= (x_3-x_1, x_1+x_2) + (y_3-y_1, y_1+y_2) \\ &= L(\boldsymbol{x}) + L(\boldsymbol{y}) \\ L(c\boldsymbol{x}) &= L[(cx_1, cx_2, cx_3)] = (cx_3 - cx_1, cx_1 + cx_2) \\ &= c(x_3 - x_1, x_1 + x_2) = cL(\boldsymbol{x}) \end{aligned}$$

Thus L satisfies conditions 1 and 2 of Definition 3.1, and it is a linear transformation.

c.
$$\begin{aligned} L(x_1, x_2, x_3) &= L(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3) \\ &= L(x_1\boldsymbol{e}_1) + L(x_2\boldsymbol{e}_2) + L(x_3\boldsymbol{e}_3) \\ &= x_1L(\boldsymbol{e}_1) + x_2L(\boldsymbol{e}_2) + x_3L(\boldsymbol{e}_3) \end{aligned}$$

Notice that c implies that once $L(\boldsymbol{e}_k)$, $k = 1, 2, 3$, are known, the fact that L is a linear transformation completely determines $L(\boldsymbol{x})$ for any vector $\boldsymbol{x}$ in $\mathbb{R}^3$. □

We collect a few facts about linear transformations in the next theorem.

Theorem 3.1. *Let L be a linear transformation from a vector space V into a vector space W. Then*

1. $L(\mathbf{0}) = \mathbf{0}$

2. $L(-\boldsymbol{x}) = -L(\boldsymbol{x})$

3. $L\left(\sum_{k=1}^{n} a_k\boldsymbol{x}_k\right) = \sum_{k=1}^{n} a_k L(\boldsymbol{x}_k)$

Proof.

1. Let $\boldsymbol{x}$ be any vector in V. Then $L(\boldsymbol{x}) = L(\boldsymbol{x}+\mathbf{0}) = L(\boldsymbol{x}) + L(\mathbf{0})$. Adding $-L(\boldsymbol{x})$ to both sides, we have $L(\mathbf{0}) = \mathbf{0}$, where the zero vector on the left-hand side is in V while the zero vector on the right-hand side is in W.

2. $\mathbf{0} = L(\mathbf{0}) = L(\boldsymbol{x}-\boldsymbol{x}) = L(\boldsymbol{x}) + L(-\boldsymbol{x})$. Thus $L(-\boldsymbol{x}) = -L(\boldsymbol{x})$.

3. We show that this formula is true for $n = 3$ and leave the details of an induction argument to the reader.

$$\begin{aligned} L(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + a_3\boldsymbol{x}_3) &= L(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2) + L(a_3\boldsymbol{x}_3) \\ &= L(a_1\boldsymbol{x}_1) + L(a_2\boldsymbol{x}_2) + L(a_3\boldsymbol{x}_3) \\ &= a_1L(\boldsymbol{x}_1) + a_2L(\boldsymbol{x}_2) + a_3L(\boldsymbol{x}_3) \end{aligned}$$ □

Example 4. Let $L\colon\ \mathbb{R}^3 \to \mathbb{R}^4$ be a linear transformation. Suppose we know that $L(1,0,1) = (-1,1,0,2)$, $L(0,1,1) = (0,6,-2,0)$, and $L(-1,1,1) = (4,-2,1,0)$. Determine $L(1,2,-1)$.

Solution. The trick is to realize that the three vectors for which we know L form a basis F of $\mathbb{R}^3$. Thus, all we need to do is find the coordinates of $(1,2,-1)$ with respect to F, and then use 3 of Theorem 3.1. The change of basis matrix P below is such that $[\boldsymbol{x}]_F^T = P[\boldsymbol{x}]_S^T$.

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

Using this matrix to find the coordinates of $(1,2,-1)$ with respect to F, we have

$$[1,2,-1]_F^T = P[1,2,-1]_S^T = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 2 & -1 \\ -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -3 \\ 6 \\ -4 \end{bmatrix}$$

Thus

$$(1,2,-1) = -3(1,0,1) + 6(0,1,1) + (-4)(-1,1,1)$$

and

$$\begin{aligned} L(1,2,-1) &= -3L(1,0,1) + 6L(0,1,1) + (-4)L(-1,1,1) \\ &= -3(-1,1,0,2) + 6(0,6,-2,0) - 4(4,-2,1,0) \\ &= (-13,41,-16,-6) \end{aligned}$$ □

A standard method of defining a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is by matrix multiplication. Thus, if $\boldsymbol{x} = (x_1, \dots, x_n)$ is any vector in $\mathbb{R}^n$ and $A = [a_{jk}]$ is an $m \times n$ matrix, define $L(\boldsymbol{x}) = A\boldsymbol{x}^T$. Then $L(\boldsymbol{x})$ is an $m \times 1$ matrix that we think of as a vector in $\mathbb{R}^m$. The various properties of matrix multiplication that were proved in Theorem 1.3 are just the statements that L is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Example 5. Let $A = \begin{bmatrix} 1 & -1 & 2 \\ 4 & 1 & 3 \end{bmatrix}$. If L is the linear transformation defined by A, compute the following:

a. $L(x_1, x_2, x_3)$ b. $L(1,0,0), L(0,1,0), L(0,0,1)$

$$\begin{aligned} L(x_1, x_2, x_3) &= \begin{bmatrix} 1 & -1 & 2 \\ 4 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \begin{bmatrix} x_1 - x_2 + 2x_3 \\ 4x_1 + x_2 + 3x_3 \end{bmatrix} \end{aligned}$$

$$L(1,0,0) = (1,4)^T \qquad L(0,1,0) = (-1,1)^T \qquad L(0,0,1) = (2,3)^T$$

The reader should note that $L(\boldsymbol{e}_1)$ is the first column of A, $L(\boldsymbol{e}_2)$ is the second column of A, and $L(\boldsymbol{e}_3)$ is the third column. □

In general, if A is an $m \times n$ matrix and $L(\boldsymbol{x}) = A\boldsymbol{x}$, then $L(\boldsymbol{e}_k)$ will be the kth column of the matrix A.

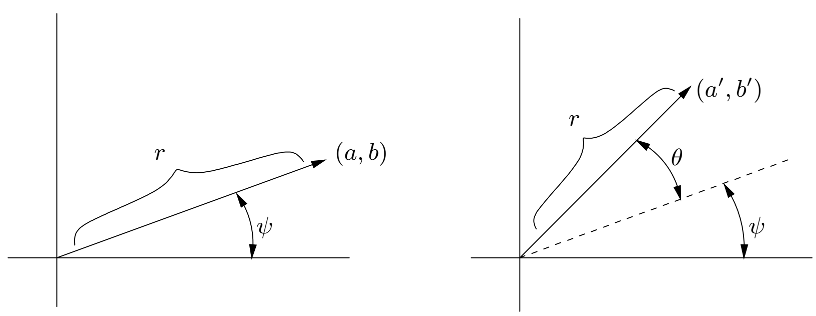

Figure 3.1

Until now we've thought of a linear transformation as an expression combining n variables to produce a vector in R^m. If we limit ourselves to this algebraic viewpoint we miss a fuller appreciation of linear transformations. For example, consider the mapping that rotates the points in the plane through an angle θ about the origin. Thus, if the point (a, b) is rotated through an angle θ to the position (a', b'), it turns out that the formulas relating (a', b') to (a, b) imply that this is a linear transformation. In fact (cf. Figure 3.1), setting

$$r = (a^2 + b^2)^{1/2} = [(a')^2 + (b')^2]^{1/2},$$

we have that

$$\begin{aligned} a' &= r\cos(\theta+\psi) = r(\cos\theta\cos\psi - \sin\theta\sin\psi) \\ &= a\cos\theta - b\sin\theta \\ b' &= r\sin(\theta+\psi) = r(\sin\theta\cos\psi + \sin\psi\cos\theta) \\ &= a\sin\theta + b\cos\theta \end{aligned}$$

Thus,

$$\begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Now whenever we see a matrix A of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where $a^2 + b^2 = 1$, we can think of A as defining a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ that rotates the plane about the origin through an angle θ, where $\cos\theta = a$, $\sin\theta = b$. Note that $A^T = A^{-1}$ corresponds to a rotation of $-\theta$.

In the succeeding pages we sometimes describe a linear transformation in a geometrical manner as well as algebraically, and the reader should try to visualize what the particular transformation is doing.

Example 6. Describe in geometrical terms the linear transformation defined by the following matrices:

a. $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. This is a clockwise rotation of the plane about the origin through 90 degrees.

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

$$A[x_1, x_2]^T = \left(2x_1, \frac{1}{3}x_2\right)^T$$

This linear transformation stretches the vectors in the subspace $S[\boldsymbol{e}_1]$ by a factor of 2 and at the same time compresses the vectors in the subspace $S[\boldsymbol{e}_2]$ by a factor of $\frac{1}{3}$. See Figure 3.2.

c. $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. For this A, the pair (a, b) gets sent to the pair $(-a, b)$. Hence this linear transformation reflects $\mathbb{R}^2$ through the x_2 axis. See Figure 3.3. $\square$

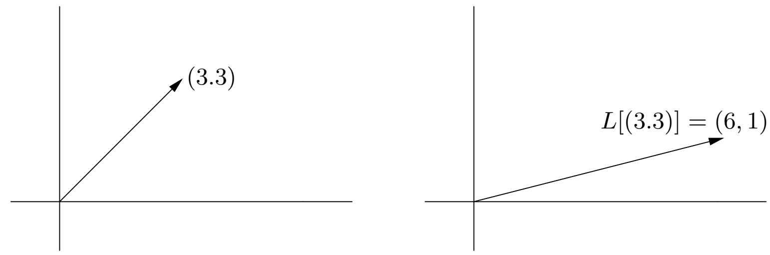

Figure 3.2

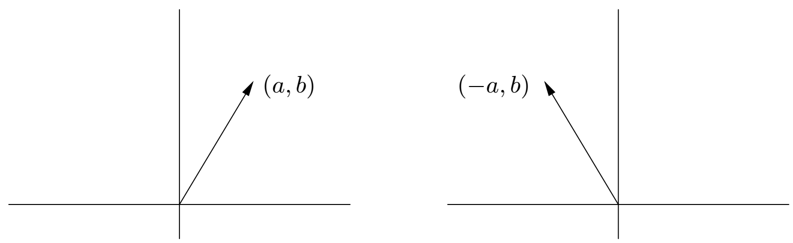

Figure 3.3

Problem Set 3.1

1. Let $L(x_1, x_2, x_3) = x_1 - x_2 + x_3$.
   a. Show that L is a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}$.
   b. Find a 1×3 matrix A such that $L(\boldsymbol{x}) = A\boldsymbol{x}^T$ for every $\boldsymbol{x}$ in $\mathbb{R}^3$.
   c. Compute $L(\boldsymbol{e}_k)$ for $k = 1, 2, 3$.
   d. Find a basis for the subspace $K = \{\boldsymbol{x}\colon\ A\boldsymbol{x}^T = 0\}$.

2. Let L be a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$ such that $L(\boldsymbol{e}_1) = (-1, 6)$, $L(\boldsymbol{e}_2) = (0, 2)$, $L(\boldsymbol{e}_3) = (8, 1)$.
   a. $L(1, 2, -6) =?$
   b. $L(x_1, x_2, x_3) =?$
   c. Find a matrix A such that $L(\boldsymbol{x}) = A\boldsymbol{x}^T$.

3. Let L be a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^5$. Suppose that $L(1, 0, 1) = (0, 1, 0, 2, 0)$, $L(0, -1, 2) = (-1, 6, 2, 0, 1)$, and $L(1, 1, 2) = (2, -3, 1, 4, 0)$. Notice that the three vectors for which we know L form a basis of $\mathbb{R}^3$.
   a. Compute $L(\boldsymbol{e}_k)$ for $k = 1, 2, 3$.
   b. $L(x_1, x_2, x_3) =?$
   c. Find a matrix A such that $L(\boldsymbol{x}) = A\boldsymbol{x}^T$.

4. For each of the following functions f determine an appropriate V and W. Then decide if f is a linear transformation from V to W.

    a. $f(x_1, x_2) = (x_1, 0, 1)$
    b. $f(x_1, x_2) = (x_1 - x_2, x_1)$
    c. $f(x) = (x, x)$
    d. $f(x_1, x_2, x_3) = (x_1, x_2, x_2, x_3, x_3, x_1)$

5. Let $L\colon\ V \to W$ be a linear transformation. Let K be any subspace of V. Define $L(K) = \{L(\boldsymbol{x})\colon\ \boldsymbol{x}$ is any vector in $K\}$. Show that $L(K)$ is a subspace of W.

6. Let $L\colon\ V \to W$ be a linear transformation. Let H be any subspace of W. Define $L^{-1}(H) = \{\boldsymbol{x}\colon\ L(\boldsymbol{x})$ is in $H\}$. Show that $L^{-1}(H)$ is a subspace of V.

7. Show that the function defined in Example 1 is not a linear transformation.

8. Let L_1 and L_2 both be linear transformations from V into W. Let $B = \{\boldsymbol{f}_k, k = 1, \ldots, n\}$ be any basis of V. Suppose that $L_1(\boldsymbol{f}_k) = L_2(\boldsymbol{f}_k)$ for each k. Show that $L_1(\boldsymbol{x}) = L_2(\boldsymbol{x})$ for every vector $\boldsymbol{x}$ in V.

9. Let $A = [a_{jk}]$ be an $m \times n$ matrix. If $L(\boldsymbol{x}) = A\boldsymbol{x}^T$, show that $L(\boldsymbol{e}_k)$ is the kth column of A.

10. Let $S = cI_2$, be an arbitrary 2×2 scalar matrix. Describe the geometrical effect that the linear transformation $S\boldsymbol{x}^T$ has on $\mathbb{R}^2$.

11. Suppose $D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$ is an arbitrary 2×2 diagonal matrix. Describe what happens to $\boldsymbol{x}$ under the linear transformation $D\boldsymbol{x}^T$ for various values of d_j.

12. If D is any invertible 2×2 diagonal matrix, describe geometrically the effects of the linear transformations defined by the two matrices D^{-1} and $D^{-1}D$.

13. Let P_n be the vector space of all polynomials of degree at most n. Define $L(\boldsymbol{p}) = t^2\boldsymbol{p}$ for each $\boldsymbol{p}$ in P_n. Then L can be thought of as a mapping from P_n to some vector space W. List some of these vector spaces, and then show that L is a linear transformation for each of your W's.

14. Let $V = C[0, 1]$, the vector space of continuous functions defined on [0,1].

    a. Define $L[\boldsymbol{f}] = \int_0^1 \boldsymbol{f}(t)dt$. Show that L is a linear transformation from V to $\mathbb{R}^1$.
    b. Define $T[\boldsymbol{f}](t) = \int_0^t \boldsymbol{f}(s)ds$, for each t in [0,1]. Show that T is a linear transformation from V to V.

15. Show that the operation of differentiation can be viewed as a linear transformation from P_n to P_{n-1}.

16. Let $V = C[0,1]$.

    a. Let $L\colon\ V \to V$ be defined by $L[\boldsymbol{f}](x) = \boldsymbol{f}(x)\sin x$. Is L a linear transformation?

    b. Let $L\colon\ V \to V$ be defined by $L[\boldsymbol{f}](x) = \sin x + \boldsymbol{f}(x)$. Is L a linear transformation?

17. Let $A = \begin{bmatrix} 2 & -3 \\ 1 & 5 \end{bmatrix}$. Let $V = M_{22}$.

    a. Define $L\colon\ V \to V$ by $L(\boldsymbol{x}) = \boldsymbol{x}A$ (matrix multiplication). Compute $L(\boldsymbol{e}_j)$ for $j = 1,2,3,4$, where the $\boldsymbol{e}_j$'s denote the standard basis vectors of M_{22}.

    b. Show that L is a linear transformation.

    c. Repeat parts a and b for $L\colon\ V \to V$ defined by $L(\boldsymbol{x}) = A\boldsymbol{x}$.

3.2 Matrix Representations

In the preceding section, matrices were used to define linear transformations. In this section we show that every linear transformation between two finite-dimensional vector spaces can be represented by a matrix. Suppose first that L is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. To find a matrix that can be used to represent L we do the following: let $\{\boldsymbol{e}_k\}$, $k = 1,2,\ldots,n$ be the standard basis of $\mathbb{R}^n$. Then

$$L(\boldsymbol{e}_k) = (a_{1k}, a_{2k}, \ldots, a_{mk}) \tag{3.1}$$

for some constants a_{1k}, a_{2k}. etc. The subscript convention is important to remember when forming the matrix A, that will represent L. Thus, $A = [a_{jk}]$ is an $m \times n$ matrix, and the entries in the kth column of A are the coordinates of $L(\boldsymbol{e}_k)$ with respect to the standard basis in R^m.

Example 1. Let $L\colon\ \mathbb{R}^3 \to \mathbb{R}^4$ be defined by

$$\begin{aligned} L(x_1,x_2,x_3) = (&-6x_2 + 2x_3, x_1 - x_2 + x_3, \\ &- x_1 + x_2 - 6x_3, 3x_1 - x_2 + 4x_3) \end{aligned}$$

Then

$$\begin{aligned} L(\boldsymbol{e}_1) &= L(1,0,0) = (0,1,-1,3) = (a_{11}, a_{21}, a_{31}, a_{41}) \\ L(\boldsymbol{e}_2) &= L(0,1,0) = (-6,-1,1,-1) = (a_{12}, a_{22}, a_{32}, a_{42}) \\ L(\boldsymbol{e}_3) &= L(0,0,1) = (2,1,-6,4) = (a_{13}, a_{23}, a_{33}, a_{34}) \end{aligned}$$

Thus,

$$A = \begin{bmatrix} 0 & -6 & 2 \\ 1 & -1 & 1 \\ -1 & 1 & -6 \\ 3 & -1 & 4 \end{bmatrix}$$

The next task is to show how to use this matrix in computing $L(\boldsymbol{x})$. Let $\boldsymbol{x} = (x_1, \ldots, x_n)$. Then, assuming $L\colon\ \mathbb{R}^n \to \mathbb{R}^m$,

$$\begin{aligned} L(\boldsymbol{x}) &= L\left(\sum_{k=1}^{n} x_k \boldsymbol{e}_k\right) = \sum_{k=1}^{n} x_k L(\boldsymbol{e}_k) \\ &= \sum_{k=1}^{n} x_k \left[\sum_{j=1}^{m} a_{jk}\boldsymbol{e}_j\right] = \sum_{j=1}^{m}\left[\sum_{k=1}^{n} a_{jk}x_k\right]\boldsymbol{e}_j \end{aligned}$$

Note: The coordinates of $L(\boldsymbol{x})$ with respect to the standard basis in $\mathbb{R}^m$ can be found by computing the matrix product $A\boldsymbol{x}^T$, where $[x_1, \ldots, x_n] = [\boldsymbol{x}]_S$. □

Example 2. In Example 1 we had

$$L(x_1, x_2, x_3) = (-6x_2 + 2x_3, x_1 - x_2 + x_3, -x_1 + x_2 - 6x_3, 3x_1 - x_2 + 4x_3)$$

with matrix representation:

$$A = \begin{bmatrix} 0 & -6 & 2 \\ 1 & -1 & 1 \\ -1 & 1 & -6 \\ 3 & -1 & 4 \end{bmatrix}$$

Computing $A\boldsymbol{x}$, we have

$$\begin{bmatrix} 0 & -6 & 2 \\ 1 & -1 & 1 \\ -1 & 1 & -6 \\ 3 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -6x_2 + 2x_3 \\ x_1 - x_2 + x_3 \\ -x_1 + x_2 - 6x_3 \\ 3x_1 - x_2 + 4x_3 \end{bmatrix}$$

Thus, $A\boldsymbol{x}$ gives us the coordinates of $L(\boldsymbol{x})$ in $\mathbb{R}^4$. □

The preceding computations were based upon the vector spaces being $\mathbb{R}^n$ and $\mathbb{R}^m$ and using the standard bases. None of this is necessary in order for us to interpert L as matrix multiplication.

Let $L\colon\ V \to W$ be a linear transformation from V into W. Let $F = \{\boldsymbol{f}_k\colon\ k = 1, \ldots, n\}$, and $B = \{\boldsymbol{g}_j\colon\ j = 1, 2, \ldots, m\}$ be bases of V and W, respectively. Proceeding as before, we write $L(\boldsymbol{f}_k)$ as a linear combination of the basis vectors in G.

$$\begin{aligned} L(\boldsymbol{f}_k) &= a_{1k}\boldsymbol{g}_1 + a_{2k}\boldsymbol{g}_2 + \cdots + a_{mk}\boldsymbol{g}_m \\ &= \sum_{j=1}^{m} a_{jk}\boldsymbol{g}_j \end{aligned} \tag{3.2}$$

Let $A = [a_{jk}]$ be the $m \times n$ matrix whose kth column is $[L(\boldsymbol{f}_k)]_G^T$, the *coordinates* of $L[\boldsymbol{f}_k]$ with respect to the basis G. This matrix A depends upon three things:

1. The linear transformation L
2. The basis F in V
3. The basis G in W

If we change either of the bases picked, the matrix representation A will also change.

The next calculation illustrates how to use the representation A to calculate the coordinates of $L(\boldsymbol{x})$. Let $\boldsymbol{x}$ be any vector in V. Let$[\boldsymbol{x}]_F = [x_1, \ldots, x_n]$. We want to show that the coordinates of $L[\boldsymbol{x}]$ with respect to G are given by the matrix product $A[\boldsymbol{x}]_F^T$. Computing $L[\boldsymbol{x}]$ we have

$$\begin{aligned} L(\boldsymbol{x}) &= L\left[\sum_{k=1}^{n} x_k \boldsymbol{f}_k\right] = \sum_{k=1}^{n} x_k L(\boldsymbol{f}_k) \\ &= \sum_{k=1}^{n} x_k \left(\sum_{j=1}^{m} a_{jk} \boldsymbol{g}_j\right) \\ &= \sum_{j=1}^{m} \left(\sum_{k=1}^{n} a_{jk} x_k\right) \boldsymbol{g}_j \end{aligned}$$

This equation says that the jth coordinate of $L(\boldsymbol{x})$ with respect to the basis G is $\sum_{k=1}^{n} a_{jk} x_k$, but this is just the jth row in the $m \times 1$ matrix $A[\boldsymbol{x}]_F^T$. *In other words, when using A we do not compute $L[\boldsymbol{x}]$ directly, but rather the coordinates of $L[\boldsymbol{x}]$ with respect to the basis G, that is,*

$$[L(\boldsymbol{x})]_G^T = A[\boldsymbol{x}]_F^T \tag{3.3}$$

Example 3. Let $V = \mathbb{R}^2$ and $W = \mathbb{R}^3$. Define $L\colon\ V \to W$ by $L(x_1, x_2) = (x_1 - x_2, x_1, x_2)$. Let $F = \{(1,1), (-1,1)\}$, and let $G = \{(1,0,1),\ (0,1,1),\ (1,1,0)\}$.

a. Find the matrix representation of L using the standard bases in both V and W.

b. Find the matrix representation of L using the standard basis in V and the basis G in W.

c. Find the matrix representation of L using the basis F in $\mathbb{R}^2$ and the standard basis in $\mathbb{R}^3$.

d. Find the matrix representation of L using the bases F and G.

Solution.

a. $L(\boldsymbol{e}_1) = L(1,0) = (1,1,0) = \boldsymbol{e}_1 + \boldsymbol{e}_2$
$L(\boldsymbol{e}_2) = L(0,1) = (-1,0,1) = -\boldsymbol{e}_1 + \boldsymbol{e}_3$

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

b. $L(\boldsymbol{e}_1) = (1,1,0) = 0(1,0,1) + 0(0,1,1) + (1,1,0) = 0\boldsymbol{g}_1 + 0\boldsymbol{g}_2 + \boldsymbol{g}_3$
$L(\boldsymbol{e}_2) = (-1,0,1) = 0(1,0,1) + (0,1,1) - (1,1,0) = 0\boldsymbol{g}_1 + \boldsymbol{g}_2 - \boldsymbol{g}_3$

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & -1 \end{bmatrix}$$

c. $L(\boldsymbol{f}_1) = L(1,1) = (0,1,1) = \boldsymbol{e}_2 + \boldsymbol{e}_3$
$L(\boldsymbol{f}_2) = L(-1,1) = (-2,-1,1) = -2\boldsymbol{e}_1 - \boldsymbol{e}_2 + \boldsymbol{e}_3$

$$A = \begin{bmatrix} 0 & -2 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$$

d. $L(\boldsymbol{f}_1) = 0\boldsymbol{g}_1 + \boldsymbol{g}_2 + 0\boldsymbol{g}_3$
$L(\boldsymbol{f}_2) = 0\boldsymbol{g}_1 + \boldsymbol{g}_2 - 2\boldsymbol{g}_3$

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & -2 \end{bmatrix} \qquad \square$$

It is clear from this example that the matrix representation of a linear transformation depends upon which bases are used in V and in W. If V and W are the same vector spaces, then normally (in this text always) only one basis is used, rather than two different bases for the same vector space.

Example 4. Let $L\colon\ \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation. Let $F = \{(1,6), (-2,3)\}$. Suppose the matrix representation of L with respect to F is

$$A = \begin{bmatrix} 2 & 8 \\ -1 & -4 \end{bmatrix}$$

Compute $L(\boldsymbol{x})$ for any vector $\boldsymbol{x}$ in $\mathbb{R}^2$.

Solution. Let $\boldsymbol{x} = (x_1, x_2)$ be any vector in $\mathbb{R}^2$. To compute $L(\boldsymbol{x})$ using the matrix A we need to find $[\boldsymbol{x}]_F$, the coordinates of $\boldsymbol{x}$ with respect to the basis F. The change of basis matrix P below gives the basis F in terms of the standard basis

$$P = \begin{bmatrix} 1 & -2 \\ 6 & 3 \end{bmatrix}$$

Using P^{-1} we calculate $[\boldsymbol{x}]_F^T$

$$\begin{aligned}[\boldsymbol{x}]_F = P^{-1}\begin{bmatrix}x_1\\x_2\end{bmatrix} &= \frac{1}{15}\begin{bmatrix}3 & 2\\-6 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}\\ &= \frac{1}{15}\begin{bmatrix}3x_1+2x_2\\-6x_1+\ \ x_2\end{bmatrix}\end{aligned}$$

Thus, the coordinates of $L(\boldsymbol{x})$ with respect to F are

$$\begin{aligned}A[\boldsymbol{x}]_F &= \begin{bmatrix}2 & 8\\-1 & -4\end{bmatrix}\begin{bmatrix}(3x_1+2x_2)/15\\(-6x_1+x_2)/15\end{bmatrix}\\ &= \frac{1}{15}\begin{bmatrix}-42x_1+12x_2\\21x_1-\ \ 6x_2\end{bmatrix}\end{aligned}$$

Thus

$$\begin{aligned}L(\boldsymbol{x}) &= \left(\frac{1}{15}\right)(-42x_1+12x_2)\boldsymbol{f}_1 + \left(\frac{1}{15}\right)(21x_1-6x_2)\boldsymbol{f}_2\\ &= \left(\frac{1}{15}\right)(-42x_1+12x_2)(1,6) + \left(\frac{1}{15}\right)(21x_1-6x_2)(-2,3)\\ &= \frac{6x_2-21x_1}{15}\{2(1,6)-(-2,3)\}\\ &= \frac{6x_2-21x_1}{15}(4,9)\end{aligned}$$

□

Example 5. Let $F = \left\{\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\right\}$. Thus F is the standard basis of M_{22}. Let $B = \begin{bmatrix}-2 & 1\\3 & 4\end{bmatrix}$. Define $L\colon\ M_{22} \to M_{22}$ by $L(\boldsymbol{x}) = B\boldsymbol{x}$. Find the matrix representation of L with respect to the standard basis F of M_{22}.

$$\begin{aligned}L(\boldsymbol{f}_1) = B\boldsymbol{f}_1 &= \begin{bmatrix}-2 & 1\\3 & 4\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} = \begin{bmatrix}-2 & 0\\3 & 0\end{bmatrix}\\ &= -2\boldsymbol{f}_1 + 3\boldsymbol{f}_3\\ L(\boldsymbol{f}_2) = B\boldsymbol{f}_2 &= \begin{bmatrix}-2 & 1\\3 & 4\end{bmatrix}\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} = \begin{bmatrix}0 & -2\\0 & 3\end{bmatrix}\\ &= -2\boldsymbol{f}_2 + 3\boldsymbol{f}_4\\ L(\boldsymbol{f}_3) = B\boldsymbol{f}_3 &= \begin{bmatrix}-2 & 1\\3 & 4\end{bmatrix}\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\4 & 0\end{bmatrix}\\ &= \boldsymbol{f}_1 + 4\boldsymbol{f}_3\\ L(\boldsymbol{f}_4) = B\boldsymbol{f}_4 &= \begin{bmatrix}-2 & 1\\3 & 4\end{bmatrix}\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}0 & 1\\0 & 4\end{bmatrix}\\ &= \boldsymbol{f}_2 + 4\boldsymbol{f}_4\end{aligned}$$

Thus, the matrix representation of L is

$$\begin{bmatrix} -2 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 \\ 3 & 0 & 4 & 0 \\ 0 & 3 & 0 & 4 \end{bmatrix}$$

□

In the following pages, when we say A, an $m \times n$ matrix, is a linear transformation or represents a linear transformation without specifically mentioning a basis or vector spaces, it is to be understood that $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, and that the standard bases in both V and W are being used.

Problem Section 3.2

1. Let $L(x_1, x_2) = (3x_1 + 6x_2, -2x_1 + x_2)$

   a. Find the matrix representation of L using the standard bases.

   b. Find the matrix representation of L using the basis $F = \{(-4, 1), (2, 3)\}$.

2. Let $L\colon\ \mathbb{R}^2 \to \mathbb{R}^4$ have matrix representation $A = \begin{bmatrix} 6 & 1 \\ -4 & 0 \\ 2 & 9 \\ 8 & -3 \end{bmatrix}$ with respect to the standard bases.

   a. $L(\boldsymbol{e}_1) =?, L(\boldsymbol{e}_2) =?$

   b. $L(-3, 7) =?$ c. $L(x_1, x_2) =?$

3. Let L be a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^2$. Define $L^2(\boldsymbol{x}) = L(L(\boldsymbol{x}))$, $\boldsymbol{L}^3(\boldsymbol{x}) = L(L^2(\boldsymbol{x}))$, and $L^{n+1}(\boldsymbol{x}) = L(L^n(\boldsymbol{x}))$. Suppose $L(x_1, x_2) = (ax_1 + bx_2, cx_1 + dx_2)$.

   a. Find the matrix representation A of L with respect to the standard bases.

   b. Show that the matrix representation of L^2 with respect to the standard bases is A^2, and in general the matrix representation of L^n with respect to the standard bases is A^n.

   c. What can you say if some basis other than the standard basis is used?

4. Let $V = \mathbb{R}^3$ and let $F = \{(1, 2, -3), (1, 0, 0), (0, 1, 0)\}$. Suppose that the matrix A represents a linear transformation $L\colon\ \mathbb{R}^3 \to \mathbb{R}^3$ with respect to the basis F, where

$$A = \begin{bmatrix} 0 & 1 & -2 \\ 2 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$$

   a. $L(1, 2, -3) =?$ b. $L(1, 0, 0) =?$ c. $L(x_1, x_2, x_3) =?$

5. Let V be a vector space of dimension n. Define $L\colon V \to V$ by $L(\boldsymbol{x}) = c\boldsymbol{x}$, where c is any constant. Let F be any basis of V. What is the matrix representation of L with respect to this basis?

6. Let $L_1(x_1, x_2) = (x_1 - x_2, 2x_1 + 3x_2)$ and let $L_2(x_1, x_2) = (2x_1 - 5x_2, 3x_1 - x_2)$. Define $(L_1 + L_2)(\boldsymbol{x}) = L_1(\boldsymbol{x}) + L_2(\boldsymbol{x})$.

   a. Find the matrix representations A_1 and A_2 of L_1 and L_2, respectively, with respect to the standard basis of R^2.
   b. Show that the matrix representation of $L_1 + L_2$ is $A_1 + A_2$.

7. Let L_1 and L_2 be two linear transformations mapping $\mathbb{R}^2$ into $\mathbb{R}^2$. Let F be any basis of $\mathbb{R}^2$. Let A_1 and A_2 be the matrix representations of L_1 and L_2, with respect to the basis F, respectively. Show that the matrix representation of $L_1 + L_2$ with respect to F is $A_1 + A_2$.

8. Let $L(x_1, x_2, x_3) = (x_2 + x_3, 6x_1 - x_2 + 3x_3, 2x_1 + 3x_2 - 7x_3, 2x_1 + 6x_3)$.

   a. Compute $L[\boldsymbol{e}_k]$ for $k = 1, 2, 3$.
   b. Find the matrix representation A, of L, with respect to the standard bases in $\mathbb{R}^3$ and $\mathbb{R}^4$.
   c. Compute $A[\boldsymbol{x}]$ for any vector $\boldsymbol{x}$ in $\mathbb{R}^3$.

9. Define $L\colon \mathbb{R}^4 \to \mathbb{R}^2$ by $L(x_1, x_2, x_3, x_4) = (x_2 + 2x_3 + 3x_4, 2x_1 - 6x_4)$.

   a. Compute $L(\boldsymbol{e}_k)$ for $k = 1, 2, 3, 4$.
   b. Find the matrix representation A, of L, with respect to the standard bases in $\mathbb{R}^4$ and $\mathbb{R}^2$.
   c. Compute $A[\boldsymbol{x}]$ for any vector $\boldsymbol{x}$ in $\mathbb{R}^4$.

10. Let L be a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$. Let $F = \{(1,1,1), (0,1,1), (1,1,0)\} = \{\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3\}$. Let $G = \{(1,2), (2,3)\} = \{\boldsymbol{g}_1, \boldsymbol{g}_2\}$. Suppose that the matrix representation of L with respect to these bases is $\begin{bmatrix} 2 & 0 & 4 \\ 1 & -2 & 1 \end{bmatrix}$.

    a. For $k = 1, 2, 3$, $[L(\boldsymbol{f}_k)]_G = ?$
    b. Compute $L(\boldsymbol{f}_k)$ for $k = 1, 2, 3$.
    c. Find the matrix representation of L using the standard basis S in $\mathbb{R}^3$ and the basis G in $\mathbb{R}^2$.
    d. Find the matrix representation of L using the standard basis S in both $\mathbb{R}^3$ and $\mathbb{R}^2$.

11. Let $L\colon \mathbb{R}^2 \to \mathbb{R}^2$. Let $F = \{\boldsymbol{f}_1, \boldsymbol{f}_2\}$ be a basis of $\mathbb{R}^2$. Suppose that

$$A = \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$$

is the matrix representation of L with respect to the basis F. What is $L(\boldsymbol{f}_k)$ for $k = 1, 2$?

12. Let L be the linear transformation that rotates the plane through an angle of θ degrees. Let A be the matrix representation of L. Then as we saw in Section 3.1
$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$
Find the matrix representations of $L^2, L^3, \ldots, L^n$. (Hint: What is L^2 geometrically?)

13. Let L_1 and L_2 be two linear transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$. Define the composition of L_1 with L_2 by $L_1 \circ L_2(\boldsymbol{x}) = L_1(L_2(\boldsymbol{x}))$.

    a. Show that the composition of two linear transformations is also a linear transformation.

    b. If A_1 and A_2 are the matrix representations of L_1 and L_2 with respect to the same basis, respectively, show that the matrix representation of the composition $L_1 \circ L_2$ is given by the matrix product $A_1 A_2$.

14. Find the matrix representations for each of the following linear transformations with respect to the standard basis of the vector space in question:

    a. $L\colon\ P_n \to P_{n-1}$ by $L(\boldsymbol{p}) = \boldsymbol{p}'$, i.e., $L(\boldsymbol{p})$ is the derivative of the polynomial $\boldsymbol{p}$.

    b. $L\colon\ P_n \to P_n$ by $L(\boldsymbol{p}) = \boldsymbol{p}'$.

    c. $L\colon\ P_n \to P_{n+2}$ by $L(\boldsymbol{p}) = t^2\boldsymbol{p}$.

15. Define $L[\boldsymbol{p}](t) = \int_0^t \boldsymbol{p}(s)ds$, for each t in [0,1]. Then $L\colon\ P_n \to P_{n+1}$. Find a matrix representation for L using the standard bases.

16. If A is an $m \times n$ matrix, we can think of A as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. What spaces are appropriate for each of the following matrices to be thought of as a linear transformation?
a. A^T b. $A^T A$ c. AA^T

17. Let L be a linear transformation from a vector space V to a vector space W. Suppose that $L(\boldsymbol{x}) = \mathbf{0}$ for every vector $\boldsymbol{x}$ in V. What must any matrix representation of L equal?

18. Let $V = \{\sum_{j=1}^2 (a_j \cos jx + b_j \sin jx)\colon\ a_j$ and b_j arbitrary$\}$. Define $L\colon\ V \to V$ by
$$L\left(\sum_{j=1}^2 a_j \cos jx + b_j \sin jx\right) = \sum_{j=1}^2 (-ja_j \sin jx + jb_j \cos jx)$$

    a. Find a basis F for the vector space V.

    b. Find the matrix representation A of L with respect to your basis.

19. Using the same notation as in Example 5, define $L\colon\ M_{22} \to M_{22}$ by $L(\boldsymbol{x}) = \boldsymbol{x}B$. Find the matrix representation of L with respect to the standard basis of M_{22}.

20. Let $G = \left\{\begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}\right\}$. Define $L\colon\ M_{22} \to M_{23}$ by $L[\boldsymbol{x}] = \boldsymbol{x}B = \boldsymbol{x}\begin{bmatrix}-1 & 3 & 4\\2 & 0 & 1\end{bmatrix}$. Using the standard basis in M_{23} and the basis G in M_{22}, find the matrix representation of L.

21. Let B and G be as in problem 20. Define $L\colon\ M_{32} \to M_{22}$ by $L[\boldsymbol{x}] = B\boldsymbol{x}$. Using the standard basis in M_{32} and the basis G in M_{22}, find the matrix representation of L.

22. Let $L\colon\ M_{22} \to M_{22}$ be defined by

$$L\left(\begin{bmatrix}a & b\\c & d\end{bmatrix}\right) = \begin{bmatrix}a & 0\\0 & 0\end{bmatrix}$$

   a. Show that L is a linear transformation.

   b. Find the matrix representation of L with respect to the standard basis of M_{22}.

   c. Show that there is no 2×2 matrix B such that $L[\boldsymbol{x}] = B\boldsymbol{x}(L[\boldsymbol{x}] = \boldsymbol{x}B)$ for all $\boldsymbol{x}$ in M_{22}.

23. Let V be the vector space in problem 18. For each $\boldsymbol{f}$ in V define $L[\boldsymbol{f}](t) = \int_0^t f(s)ds$. Show that L is a linear transformation from V to V. Find its matrix representation with respect to the basis found in problem 18. Show that the product of the matrix found in problem 18 with the matrix found in this problem equals I_4.

3.3 Kernel and Range of a Linear Transformation

For any linear transformation L, mapping V into W, there are two important subspaces associated with L. The first is a subspace of V called the kernel of L; the second is a subspace of W called the range of L. In this section we define these two subspaces and describe their relation to the solution set of a system of equations.

Definition 3.2. Let $L\colon\ V \to W$. The kernel of L is the set of vectors $\boldsymbol{x}$ in V for which $L(\boldsymbol{x}) = \mathbf{0}$. Letting $\ker(L)$ represent the kernel of L, we have $\ker(L) = \{\boldsymbol{x}\colon\ L(\boldsymbol{x}) = \mathbf{0}\}$.

Example 1. Let $A = \begin{bmatrix}2 & -6 & 4\\1 & -1 & 2\end{bmatrix}$ be the matrix representation of L. Find the kernel K of this linear transformation.

Solution. Since A is a 2×3 matrix, $A\colon\ \mathbb{R}^3 \to \mathbb{R}^2$. We are asked to find those $\boldsymbol{x} = (x_1, x_2, x_3)$ such that

$$A\boldsymbol{x} = \begin{bmatrix} 2 & -6 & 4 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 - 6x_2 + 4x_3 \\ x_1 - x_2 + 2x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Thus, $\boldsymbol{x}$ is in the kernel of A if and only if $2x_1 - 6x_2 + 4x_3 = 0 = x_1 - x_2 + 2x_3$. Hence $K = \{(x_1, x_2, x_3)\colon\ x_1 + 2x_3 = 0 = x_2\}$. □

This example demonstrates that the kernel is just the solution set of a homogeneous system of linear equations. We note that K has dimension equal to 1 and that $(-2, 0, 1)$ is a basis of K.

Definition 3.3. Let L be a linear transformation mapping V into W. The range of L is the set of vectors $\boldsymbol{w}$ in W such that $L(\boldsymbol{x}) = \boldsymbol{w}$, for some vector $\boldsymbol{x}$ in V. Thus, $\mathrm{Rg}(L) = \mathrm{range}(L) = \{\boldsymbol{w}\colon\ \boldsymbol{w} = L(\boldsymbol{x}) \text{ for some } \boldsymbol{x} \text{ in } V\}$.

Example 2. Find the range of the linear transformation in Example 1.

Solution. Since $A\colon\ \mathbb{R}^3 \to \mathbb{R}^2$, the range of A consists of those $\boldsymbol{w}$ in $\mathbb{R}^2$ such that $A\boldsymbol{x} = \boldsymbol{w}$ has a solution. The augmented matrix of the associated system is

$$\begin{bmatrix} 2 & -6 & 4 & w_1 \\ 1 & -1 & 2 & w_2 \end{bmatrix}$$

It is clear that this system has a solution regardless of the values of w_1 and w_2, e.g., $x_1 = (6w_2 - w_1)/4$, $x_2 = (2w_2 - w_1)/4$, and $x_3 = 0$. Thus, $\mathrm{Rg}(L) = \mathbb{R}^2$. □

Theorem 3.2. *Let $L\colon\ V \to W$ be a linear transformation. Then*

a. $\ker(L)$ *is a subspace of V.*

b. $\mathrm{Rg}(L)$ *is a subspace of W.*

Proof. Since $L(\mathbf{0}) = \mathbf{0}$, we know that both the kernel and the range are nonempty. Thus, to show that these two sets are subspaces we may use Theorem 2.6. Hence, suppose that $\boldsymbol{x}$ and $\boldsymbol{y}$ are in $K = \ker(L)$. Then $L(\boldsymbol{x} + \boldsymbol{y}) = L(\boldsymbol{x}) + L(\boldsymbol{y}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$. Thus, K is closed under vector addition. Now let a be any scalar; then $L(a\boldsymbol{x}) = aL(\boldsymbol{x}) = a\mathbf{0} = \mathbf{0}$, and we see that K is also closed under scalar multiplication. This shows that K is a subspace. To see that $\mathrm{Rg}(L)$ is a subspace, suppose that $\boldsymbol{u}$ and $\boldsymbol{v}$ are any two vectors in $\mathrm{Rg}(L)$. Then there are two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in V such that $L(\boldsymbol{x}) = \boldsymbol{u}$ and $L(\boldsymbol{y}) = \boldsymbol{v}$. Then $L(\boldsymbol{x} + \boldsymbol{y}) = L(\boldsymbol{x}) + L(\boldsymbol{y}) = \boldsymbol{u} + \boldsymbol{v}$ and $\mathrm{Rg}(L)$ is closed under addition. Similarly if a is any constant, then $a\boldsymbol{u} = aL(\boldsymbol{x}) = L(a\boldsymbol{x})$. Since $\mathrm{Rg}(L)$ is closed under vector addition and scalar multiplication, it is a subspace of W.

Consider a system of m linear equations in n unknowns

$$\begin{gathered} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{gathered} \tag{3.4}$$

Let L be the linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by $L[\boldsymbol{x}] = A\boldsymbol{x}$, where A is the coefficient matrix $[a_{jk}]$ of (3.4). Then the kernel of L is just the solution set of the homogeneous system associated with (3.4). For $\boldsymbol{x}$ is in $\ker(L)$ if and only if $L(\boldsymbol{x}) = \mathbf{0}$, but $L(\boldsymbol{x}) = \mathbf{0}$ if and only if $A\boldsymbol{x} = \mathbf{0}$. That is, $\boldsymbol{x}$ is in $\ker(L)$ if and only if x is a solution of (3.4) with $b_j = 0$ for $j = 1, 2, \ldots, m$. The range of L consists of those vectors $\boldsymbol{b}$ in $\mathbb{R}^m$ such that there is an $\boldsymbol{x}$ in $\mathbb{R}^n$ for which $L(\boldsymbol{x}) = \boldsymbol{b}$, i.e., $A\boldsymbol{x} = \boldsymbol{b}$. That is, $\boldsymbol{b}$ is in the range of L if and only if (3.4) has a solution. □

Example 3. Consider the following system of equations:

$$\begin{aligned} -x_1 + 2x_2 + \ 3x_4 &= b_1 \\ 2x_1 + 3x_2 + 7x_3 + 8x_4 &= b_2 \\ 4x_1 - 2x_2 + 6x_3 &= b_3 \end{aligned} \tag{3.5}$$

Find the kernel and range of the coefficient matrix of the above system of equations. Then determine the dimensions of these two subspaces.

Solution. The coefficient matrix A equals

$$\begin{bmatrix} -1 & 2 & 0 & 3 \\ 2 & 3 & 7 & 8 \\ 4 & -2 & 6 & 0 \end{bmatrix}$$

and is row equivalent to the matrix

$$\begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, $\boldsymbol{x}$ is a solution to the homogeneous system, i.e., $\boldsymbol{x}$ is in $\ker(A)$ if and only if $x_2 = -x_3 - 2x_4$ and $x_1 = -2x_3 - x_4$. Thus, $\ker(A) = \{(x_1, x_2, x_3, x_4)\colon\ x_1 = -2x_3 - x_4,\ x_2 = -x_3 - 2x_4\}$. A basis for $\ker(A)$ is $\{(-2, -1, 1, 0), (-1, -2, 0, 1)\}$. Thus, $\dim(\ker(A)) = 2$.

The augmented matrix of (3.5)

$$\begin{bmatrix} -1 & 2 & 0 & 3 & b_1 \\ 2 & 3 & 7 & 8 & b_2 \\ 4 & -2 & 6 & 0 & b_3 \end{bmatrix}$$

is row equivalent to

$$\begin{bmatrix} -1 & 2 & 0 & 3 & b_1 \\ 0 & 1 & 1 & 2 & (b_2 + 2b_1)/7 \\ 0 & 0 & 0 & 0 & (26b_1 - 2b_2 + 7b_3)/14 \end{bmatrix}$$

(3.5) has a solution if and only if

$$26b_1 - 2b_2 + 7b_3 = 0$$

Thus, $\mathrm{Rg}(A) = \{(b_1, b_2, b_3)\colon\ 26b_1 - 2b_2 + 7b_3 = 0\}$. A basis for $\mathrm{Rg}(A)$ is $\{(1, 13, 0),\ (7, 0, -26)\}$ and $\dim(\mathrm{Rg}(A)) = 2$. □

In the previous example $A\colon\ \mathbb{R}^4 \to \mathbb{R}^3$, $\dim(\ker(A)) = 2$, and $\dim(\mathrm{Rg}(A)) = 2$. It is not a coincidence that we have the following relationship: $\dim(\ker(A)) + \dim(\mathrm{Rg}(A)) = \dim(\mathbb{R}^4)$.

Theorem 3.3. *Let L be a linear transformation from V to W, where V is a finite dimensional vector space. Then*

$$\dim(\ker(L)) + \dim(\mathrm{Rg}(L)) = \dim(V) \tag{3.6}$$

Proof. Let $\dim(V) = n$. Suppose that $\dim(\ker(L)) = k$, where we assume first that $0 < k < n$. Let $\boldsymbol{x}_j$, $j = 1, \dots, k$ be a basis for $\ker(L)$ and let $\boldsymbol{y}_j$, $j = 1, \dots, n-k$ be a set of $n-k$ linearly independent vectors in V such that $S = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_k,\ \boldsymbol{y}_1, \dots, \boldsymbol{y}_{n-k}\}$ is a basis of V. Then $\{L(\boldsymbol{x}_1), \dots, L(\boldsymbol{y}_1), \dots, L(\boldsymbol{y}_{n-k})\}$ is a spanning set of $\mathrm{Rg}(L)$. Since the $\boldsymbol{x}_j$ are in $\ker(L)$, we have $L(\boldsymbol{x}_j) = \boldsymbol{0}$ for $j = 1, \dots, k$. Thus, $\{L(\boldsymbol{y}_1), \dots, L(\boldsymbol{y}_{n-k})\}$ must span $\mathrm{Rg}(L)$. We wish to show that this set is linearly independent, and hence forms a basis for $\mathrm{Rg}(L)$. To this end suppose that

$$c_1 L(\boldsymbol{y}_1) + \cdots + c_{n-k} L(\boldsymbol{y}_{n-k}) = \boldsymbol{0}$$

Setting $\boldsymbol{z} = c_1\boldsymbol{y}_1 + \cdots + c_{n-k}\boldsymbol{y}_{n-k}$, we have $L(\boldsymbol{z}) = \boldsymbol{0}$. Thus, $\boldsymbol{z}$ is in $\ker(L)$ and there are constants a_j such that

$$a_1\boldsymbol{x}_1 + \cdots + a_k\boldsymbol{x}_k = \boldsymbol{z} = c_1\boldsymbol{y}_1 + \cdots + c_{n-1}\boldsymbol{y}_{n-k}$$

Since the set S is linearly independent, every one of the constants must be zero. In particular $c_1 = c_2 = \cdots = c_{n-k} = 0$, and we conclude that the set $\{L(\boldsymbol{y}_1), \dots, L(\boldsymbol{y}_{n-k})\}$ is a basis for $\mathrm{Rg}(L)$. Hence we have

$$\dim(\ker(L)) + \dim(\mathrm{Rg}(L)) = k + (n-k) = n = \dim(V)$$

This equation remains to be verified in the two cases $k = 0$ and $k = n$. We leave the details as an exercise for the reader $\square$

In the next section we show how one can easily determine the dimension of $\mathrm{Rg}(L)$. This technique coupled with the above theorem gives us an effective means of determining how large the solution space of a set of homogeneous linear equations is, and hence the size of the solution set for any system of linear equations, homogeneous or not; cf. Theorem 1.10.

Before starting the next section, we define several items.

Definition 3.4. Let $L\colon\ V \to W$ be a linear transformation. L is said to be one-to-one if $L(\boldsymbol{x}) = L(\boldsymbol{y})$ implies that $\boldsymbol{x} = \boldsymbol{y}$.

Definition 3.5. Let $L\colon\ V \to W$ be a linear transformation. L is said to be onto if $\mathrm{Rg}(L) = W$.

Example 4. Let $L\colon\ \mathbb{R}^3 \to \mathbb{R}^2$ have matrix representation A, where A is given below. Show L is onto but not one-to-one.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix}$$

Solution. To see that L is not one-to-one we observe that the vector $(1, -2, 1)$ is in $\ker(L)$; that is, $L(1, -2, 1) = (0, 0) = L(0, 0, 0)$, but $(1, -2, 1) \neq (0, 0, 0)$. Hence, L is not one-to-one. To see that L is onto we have to show that $\mathrm{Rg}(L) = \mathbb{R}^2$. Thus, let (w_1, w_2) be any vector in $\mathbb{R}^2$. Our task is to find $\boldsymbol{x} = (x_1, x_2, x_3)$ such that $L(\boldsymbol{x}) = (w_1, w_2)$. Equivalently we need to solve the following system of equations:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= w_1 \\ x_2 + 2x_3 &= w_2 \end{aligned}$$

A solution to this system is $x_1 = w_1 - 2w_2$, $x_2 = w_2$, and $x_3 = 0$. □

Theorem 3.4. *Let $L\colon\ V \to W$ be a linear transformation. Assume, for parts b and c, that V and W are finite dimensional. Then*

a. *L is one-to-one if and only if* $\ker(L) = \{\mathbf{0}\}$.

b. *L is one-to-one if and only if* $\dim(V) = \dim(\mathrm{Rg}(L))$.

c. *L is onto if and only if* $\dim(\mathrm{Rg}(L)) = \dim(W)$.

Proof. Suppose L is one-to-one. We want to show that $K = \ker(L) = \{\mathbf{0}\}$. Thus, suppose that $\boldsymbol{x}$ is in K. Then $L(\boldsymbol{x}) = \mathbf{0} = L(\mathbf{0})$ and we must have $\boldsymbol{x} = \mathbf{0}$. Conversely, suppose $K = \{\mathbf{0}\}$. Then if $L(\boldsymbol{x}) = L(\boldsymbol{y})$, we must have $L(\boldsymbol{x} - \boldsymbol{y}) = 0$. Hence, $\boldsymbol{x} - \boldsymbol{y}$ is in K, and we conclude that $\boldsymbol{x} = \boldsymbol{y}$. Part b of this theorem is an easy consequence of part a and Theorem 3.3. Suppose that L is one-to-one. Then we have $\dim(V) = \dim(\mathrm{Rg}(L)) + \dim(\ker(L)) = \dim(\mathrm{Rg}(L)) + 0 = \dim(Rg(L))$. Conversely, if $\dim(\mathrm{Rg}(L)) = \dim(V)$, then $\dim(\ker(L)) = 0$, and we have $\ker(L) = \{0\}$. The verification of part c is left to the reader as an exercise. □

Problem Set 3.3

1. Each of the matrices below represents a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. Determine the values of n and m for each matrix. Then determine their kernels and ranges and find a basis for each of these subspaces.

   a. $\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ c. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ d. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$

2. Each of the matrices below represents a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. Determine the values of n and m for each matrix, and their respective kernels and ranges.

   a. $\begin{bmatrix} 1 & 2 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}$ c. $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

3. Let A be the coefficient matrix of the system of equations below. If L is the linear transformation defined by A, what is the range of L and what is its kernel? Does this particular equation have a solution; i.e., is $(-2, 1)$ in the range of L?
$$\begin{aligned} 2x_1 - 5x_2 + 3x_3 &= -2 \\ x_1 + 3x_2 \qquad &= \ \ 1 \end{aligned}$$

4. For each of the matrices below determine the dimension of its range and the dimension of its kernel. Then decide if the linear transformations represented by these matrices are onto and/or one-to-one.

   a. $\begin{bmatrix} 1 & 2 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}$ c. $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

5. For each of the matrices below determine the dimensions of their range and kernel. Then determine if the linear transformations represented by these matrices are onto and/or one-to-one.

   a. $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 4 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 2 & -1 & 3 \\ 1 & -1 & 1 & -1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$

6. Verify part c of Theorem 3.4. Remember, the range of L is always a subspace of W.

7. Let $L\colon\ V \to W$ be a linear transformation. Let $\{\boldsymbol{x}_j\colon\ j = 1, \ldots, n\}$ be a basis of V. Show that the set $\{L(\boldsymbol{x}_j)\colon\ j = 1, \ldots, n\}$ is a spanning set for $\mathrm{Rg}(L)$.

8. Let $L\colon\ \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

   a. Show that if $n > m$, then L must have a nontrivial kernel, i.e., $\dim(\ker(L)) > 0$.

   b. If $n \leq m$ does L have to be one-to-one?

9. Let $L\colon\ \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

   a. If $n < m$, show that L cannot be onto.

   b. If $n \geq m$, must L be onto?

10. Let $L\colon V \to W$ be a linear transformation. Let Q be that subset of W that contains all vectors in W not in the range of L, i.e., $Q = W \backslash \mathrm{Rg}(L)$. Is Q a subspace of W?

11. Let S be any $n \times n$ scalar matrix, i.e., $S = cI_n$ for some constant c. Determine the kernel and range of S for various values of c.

12. Let D be any $n \times n$ diagonal matrix. Determine the kernel and range of D for various values of the diagonal entries. For example, what happens if the entry in the 1,1 position of D is zero? Is nonzero?

13. Characterize the kernel and range for the linear transformations in problems 13, 14, and 15 in Problem Set 3.1.

14. For each of the matrices A in problem 1, compute A^T. Then determine the kernel and range of A^T.

15. For each of the matrices A in problem 1, compute A^TA. Determine if these product matrices are one-to-one and/or onto. Compare the kernels of A^TA and A.

16. Let $L\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Show that if L is both onto and one-to-one, then $m = n$.

17. Verify Theorem 3.3 for the two cases $\ker(L) = \{\mathbf{0}\}$ and $\ker(L) = V$.

18. Let $L\colon P_2 \to P_3$ be defined by $L[\boldsymbol{p}](t) = \int_0^t \boldsymbol{p}(s)ds$. Find a matrix representation for L using the standard basis in each of the vector spaces. Find a basis for the range and kernel of L.

19. Let $L\colon P_3 \to P_3$ be defined by $L[\boldsymbol{p}](t) = \boldsymbol{p}'(t)$. Find a matrix representation for L using the standard basis in each of the vector spaces. Find a basis for the range and kernel of L.

20. Let $B = \begin{bmatrix} -1 & 2 & 5 \\ 2 & 3 & 1 \end{bmatrix}$.

    a. Let $L[\boldsymbol{x}] = \boldsymbol{x}B$ for any $\boldsymbol{x}$ in M_{22}. Then L is a linear transformation from M_{22} to M_{23}. Find a basis for the kernel of L and also a basis for the range of L.

    b. Let $L[\boldsymbol{x}] = B\boldsymbol{x}$ for any $\boldsymbol{x}$ in M_{32}. Then L is a linear transformation from M_{32} to M_{22}. Find a basis for the kernel of L and also a basis for the range of L.

    c. Find a matrix representation for each of the above two linear transformations. Use the standard basis.

3.4 Rank of a Matrix

In the last section we proved a theorem that said the dimensions of the kernel, range, and domain of a linear transformation are related by the equation $\dim(V) = \dim(\ker) + \dim(\mathrm{Rg})$. We have also seen that the kernel is just the solution set for a system of homogeneous equations. In this section we show that the dimension of the range of L is the same as the maximum number of linearly independent rows or columns in a matrix representation of L. Since this number is easy to calculate, we have an efficient method for computing the dimension of $\mathrm{Rg}(L)$ and hence an efficient method of determining the size of the solution set of a system of linear equations.

Definition 3.6. Let $A = [a_{jk}]$ be an $m \times n$ matrix. Each row of A can be thought of as a vector in $\mathbb{R}^n$ and each column of A can be considered a vector in $\mathbb{R}^m$. The row space of A is that subspace of $\mathbb{R}^n$ spanned by the row vectors of A, and the column space of A is that subspace of $\mathbb{R}^m$ spanned by the column vectors of A.

Definition 3.7. Let $A = [a_{jk}]$ be an $m \times n$ matrix. The row rank of A is the dimension of the row space of A and the column rank of A is the dimension of the column space of A.

Example 1. Let $A = \begin{bmatrix} -2 & -1 \\ 1 & 1 \\ 0 & 4 \end{bmatrix}$. The row space of A is that subspace of $\mathbb{R}^2$ spanned by the set $\{(-2,-1),\ (1,1),\ (0,4)\}$. Clearly this set spans a subspace of $\mathbb{R}^2$ of dimension 2. Hence the row space of A is $\mathbb{R}^2$, and the row rank is 2. The column space of A is that subspace of $\mathbb{R}^3$ spanned by the set $\{(-2,1,0),(-1,1,4)\}$. Since this set is linearly independent the column space has dimension 2. Thus the column rank is 2. $\square$

The fact that the row rank of A was equal to its column rank was no accident, as the following theorem shows.

Theorem 3.5. *Let $A = [a_{ik}]$ be an $m \times n$ matrix. Then the column rank and the row rank of A are equal.*

Proof. Suppose the column rank of A is p. Then $0 \le p \le n$. If $p = 0$, every column of A is the zero vector in $\mathbb{R}^m$, and hence every row of A is the zero vector in $\mathbb{R}^n$. Thus the row space is the zero vector and we have row rank equal to 0 also. Now assume that $p > 0$. Let $\{\boldsymbol{z}_j\colon\ j = 1, \ldots, p\}$ be a basis for the column space of A, where

$$\boldsymbol{z}_j = (z_{1j}, z_{2j}, \ldots, z_{mj})$$

Then if $\boldsymbol{C}_k$ is the kth column of A, i.e., $\boldsymbol{C}_k = (a_{1k}, a_{2k}, \ldots, a_{mk})^T$, there are constants b_{jk}, $1 \le j \le p$, such that

$$\boldsymbol{C}_k = \sum_{j=1}^{p} b_{jk}\boldsymbol{z}_j^T$$

Equating components, we have the following:

$$a_{ik} = \sum_{j=1}^{p} b_{jk} z_{ij} \qquad 1 \leq k \leq n, 1 \leq i \leq m$$

Thus if $\boldsymbol{r}_i$ is the ith row of A, we have

$$\begin{aligned} \boldsymbol{r}_i &= (a_{i1}, a_{i2}, a_{i3}, \ldots, a_{ij}) \\ &= \left(\sum_{j=1}^{p} b_{j1} z_{ij}, \sum_{j=1}^{p} b_{j2} z_{ij}, \ldots, \sum_{j=1}^{p} b_{jn} z_{ij} \right) \\ &= \sum_{j=1}^{p} z_{ij} (b_{j1}, b_{j2}, \ldots, b_{jn}) \end{aligned}$$

Thus, the p row vectors $(b_{j1}, \ldots, b_{jn})$, $1 \leq j \leq p$, form a spanning set for the row space of A. Hence, row rank $\leq p$. This shows that the row rank of any matrix must be less than or equal to its column rank. Since the row rank of A^T is the column rank of A, and the column rank of A^T is the row rank of A, we also have that the column rank is less than or equal to the row rank. Hence the row and column ranks are equal. $\square$

Definition 3.8. The rank of a matrix is the common value of its row and column rank.

Example 2. Compute the rank of the following matrix:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 1 \\ 0 & 2 & 3 & 2 \end{bmatrix}$$

Solution. This matrix is easily seen to be row equivalent to the matrix.

$$B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 5 & 0 \end{bmatrix}$$

This last matrix has rank equal to 3. Since the rows of B were obtained from linear combinations of the rows of A and we can also obtain the rows of A from linear combinations of the rows of B, the row spaces of A and B must be the same. Hence, A and B have the same row rank and thus the same rank, namely, 3. $\square$

In the preceding example we computed the row rank of A by first finding the row rank of a matrix that was row equivalent to A and then used the fact that their row ranks were equal. We formalize this in the next theorem.

Theorem 3.6. *If A and B are two matrices that are row or column equivalent, then the rank of A is equal to the rank of B.*

Proof. See problem 10 at the end of this section. $\square$

We now need to relate these ideas to the problem of describing the solution space of a system of equations. Consider the linear system of equations $A\boldsymbol{x} = \boldsymbol{b}$, where $A = [a_{jk}]$ is an $m \times n$ matrix. The matrix A defines a linear transformation L from $\mathbb{R}^n$ to $\mathbb{R}^m$. Since $L(\boldsymbol{e}_k)$ equals the kth column of A, it is clear that the column space of A is the same as the range of L. In fact if we let $\boldsymbol{C}_k$ be the kth column of A, and $\boldsymbol{x} = (x_1, \ldots, x_n)^T$, then $A\boldsymbol{x} = x_1\boldsymbol{C}_1 + \cdots + x_n\boldsymbol{C}_n$; that is, $A\boldsymbol{x}$ is just a linear combination of the columns of A. Thus, if A is

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & -2 & 8 \end{bmatrix}$$

we may write $A[x_1, x_2, x_3]^T$ as

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & -2 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 4 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ -2 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 8 \end{bmatrix}$$

Clearly this is a linear combination of the columns of A.

These remarks make it clear that the column rank of A (the dimension of the column space) is equal to the dimension of the range of L. Hence we have

$$\dim(\ker(L)) = n - \dim(\mathrm{Rg}(L)) = n - \text{rank of } A$$

Remembering that $\ker(L)$ is just the solution space of the homogeneous system $A\boldsymbol{x} = \boldsymbol{0}$, we have our promised algorithm for computing the size of the solution space.

What can be said about nonhomogeneous equations? Essentially the same thing, as the following theorem indicates.

Theorem 3.7. *Suppose $\boldsymbol{x}_0$ satisfies $A\boldsymbol{x}_0 = \boldsymbol{b}$. Then the solution space of $A\boldsymbol{x} = \boldsymbol{b}$ equals $\boldsymbol{x}_0 + \ker(A)$; that is, every solution of $A\boldsymbol{x} = \boldsymbol{b}$ is equal to $\boldsymbol{x}_0$ plus some vector in the kernel. Note that this is just Theorem 1.10 restated.*

Proof. Suppose $A\boldsymbol{x} = \boldsymbol{b}$. Then

$$\boldsymbol{0} = \boldsymbol{b} - \boldsymbol{b} = A\boldsymbol{x} - A\boldsymbol{x}_0 = A(\boldsymbol{x} - \boldsymbol{x}_0)$$

Thus $\boldsymbol{x} - \boldsymbol{x}_0$ is in the kernel of A, and $\boldsymbol{x} = \boldsymbol{x}_0 + (\boldsymbol{x} - \boldsymbol{x}_0)$. Conversely if $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{z}$ for some $\boldsymbol{z}$ in the kernel, then

$$A\boldsymbol{x} = A(\boldsymbol{x}_0 + \boldsymbol{z}) = A\boldsymbol{x}_0 + A\boldsymbol{z} = \boldsymbol{b} + \boldsymbol{0} = \boldsymbol{b}$$ $\square$

Example 3. Describe the solution set of the following system of equations:

$$\begin{aligned} 2x_1 \qquad\quad + 4x_3 + \ \ 5x_4 &= \ \ 8 \\ x_1 + 2x_2 \qquad\quad + \ \ 5x_4 &= \ \ 4 \\ x_1 + 6x_2 + 3x_3 + 10x_4 &= 11 \end{aligned}$$

Solution. The coefficient matrix A, which equals

$$\begin{bmatrix} 2 & 0 & 4 & 5 \\ 1 & 2 & 0 & 5 \\ 1 & 6 & 3 & 10 \end{bmatrix}$$

is row equivalent to the matrix

$$\begin{bmatrix} 1 & 0 & 0 & \frac{5}{2} \\ 0 & 1 & 0 & \frac{5}{4} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Hence, $\dim(\mathrm{Rg}(A)) =$ rank of $A = 3$. Since $4 - 3 = 1$, $\dim(\ker A) = 1$. A basis for the kernel is $(-\frac{5}{2}, -\frac{5}{4}, 0, 1)$. Thus, the solution space of the equation $A\boldsymbol{x} = (8, 4, 11)^T$ is of the form $\boldsymbol{x} = \boldsymbol{x}_0 + c(-\frac{5}{2}, -\frac{5}{4}, 0, 1)$, for any constant c, assuming of course that there is at least one particular solution $\boldsymbol{x}_0$. There is a solution if and only if (8,4,11) is in the range of A and equivalently if and only if $(8, 4, 11)^T$ is in the column space of A. A basis of the column space is $\{(2, 1, 1)^T, (0, 2, 6)^T, (4, 0, 3)^T\}$. One easily sees that $(8, 4, 11)^T = 2(2, 1, 1)^T + (0, 2, 6)^T + (4, 0, 3)^T$; that is, $\boldsymbol{b}$ equals 2(first column) + (second column) + (third column) + 0(fourth column). Thus, $\boldsymbol{x}_0 = (2, 1, 1, 0)^T$ is a particular solution, and every solution is of the form

$$\boldsymbol{x} = (2, 1, 1, 0) + c\left(-\frac{5}{2}, -\frac{5}{4}, 0, 1\right) \qquad \square$$

A common problem has to do with data fitting. In its simplest form, this problem can be viewed in the following manner. Assume we have a collection of n points (x_k, y_k), $1 \leq k \leq n$, where $x_j \neq x_k$ if $j \neq k$. We wish to find a polynomial $\boldsymbol{p}(t)$ of degree m such that $\boldsymbol{p}(x_k) = y_k$ for each k. In fact, we would like to find the smallest value of m such that regardless of the values y_k such a polynomial will exist. Since there are n data points and a polynomial of degree m has $m + 1$ arbitrary coefficients we might conjecture m equal to $n - 1$ is the smallest value of m for which we are guaranteed a solution. We now recast this problem in terms of linear transformations. Thus, let (x_j, y_j), $j = 1, 2, \ldots, n$, be given. Let $\boldsymbol{p}(t)$ be any polynomial in P_m. Define $L\colon\ P_m \to \mathbb{R}^n$ by

$$L(\boldsymbol{p}) = (\boldsymbol{p}(x_1), \boldsymbol{p}(x_2), \ldots, \boldsymbol{p}(x_n))$$

That is, we evaluate our polynomial at the n numbers x_j. For example, if we had the three points $(-1, -2)$, (0,6), and (1,0) and $\boldsymbol{p}(t) = 8 + 2t - 8t^3 + t^4$, then

$$L(\boldsymbol{p}) = (\boldsymbol{p}(-1), \boldsymbol{p}(0), \boldsymbol{p}(1)) = (15, 8, 3)$$

The question then, as to whether or not a polynomial $\boldsymbol{p}$ of degree m can be picked so that $\boldsymbol{p}(x_j) = y_j$, amounts to deciding if the point $(y_1, y_2, \ldots, y_n)$ lies in the range of the linear transformation L. We next find a matrix representation for this linear transformation using the natural basis $\{1, t, t^2, \dots, t^m\}$ in P_m and the standard basis $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ in $\mathbb{R}^n$.

$$\begin{aligned} L(1) &= (1, 1, \dots, 1) \\ L(t) &= (x_1, x_2, \dots, x_n) \\ L(t^k) &= (x_1^k, \dots, x_n^k) \qquad k = 1, 2, \dots, m \end{aligned}$$

Thus, we have the $n \times (m+1)$ matrix

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m \\ 1 & x_2 & x_2^2 & \dots & x_2^m \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^m \end{bmatrix}$$

If all the x_j's are different, it can be shown that the rank of A is the smaller of the two numbers $m+1$ and n. Thus, if we wish to always be able to solve $L(\boldsymbol{p}) = (y_1, \dots, y_n)$ we need rank $A = n$. Clearly, the smallest value of m that works is $m = n - 1$; i.e., A is a square matrix. Another way of stating this is that for any n distinct numbers $x_1, \dots, x_n$ the linear transformation L maps P_{n-1} (polynomials of degree $n-1$ or less) onto $\mathbb{R}^n$ in a one-to-one fashion.

Example 4. Given the three points $(1, -6)$, (2,0), and (3,6), we know from the above discussion that there is a polynomial $\boldsymbol{p}(t)$, of degree $3 - 1 = 2$, such that $\boldsymbol{p}(1) = -6$, $\boldsymbol{p}(2) = 0$, and $\boldsymbol{p}(3) = 6$. Find this polynomial.

Solution. The transformation $L\colon\ P_2 \to \mathbb{R}^3$ defined by the abscissas of these three points satisfies

$$L(\mathbf{1}) = (1, 1, 1) \qquad L(\boldsymbol{t}) = (1, 2, 3) \qquad L(\boldsymbol{t}^2) = (1, 4, 9)$$

The matrix representation for this transformation is

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix}$$

This matrix has rank equal to 3. We wish to find $\boldsymbol{p}(t)$ such that $L(\boldsymbol{p}) = (-6, 0, 6)$. In terms of the coefficients a_j of $\boldsymbol{p}(t) = a_0 + a_1 t + a_2 t^2$, we want a solution to the equation

$$A \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -6 \\ 0 \\ 6 \end{bmatrix}$$

The unique solution to this equation is $a_0 = -12$, $a_1 = 6$, $a_2 = 0$. Thus, $\boldsymbol{p}(t) = -12 + 6t$ is the unique polynomial of degree 2 or less that fits the data. We know that the polynomial is unique, since the matrix A has rank 3, which implies that $\dim(\ker(L)) = 3 - 3 = 0$. See Figure 3.4. $\square$

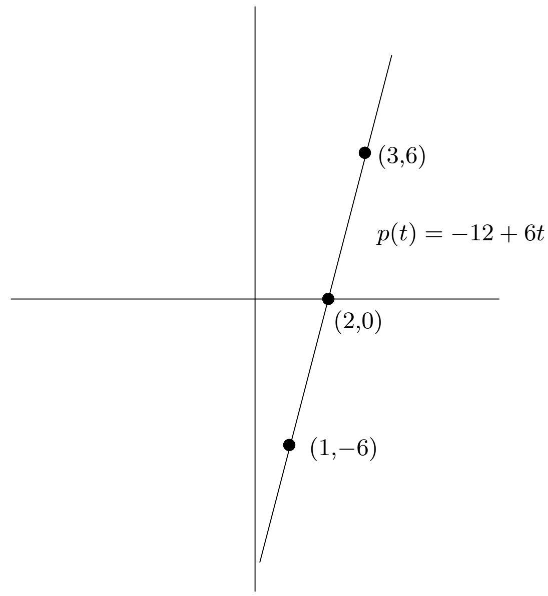

Figure 3.4

Problem Set 3.4

1. Calculate the rank of each of the following matrices:

   a. $\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ c. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ d. $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$

2. Calculate the rank of each of the following matrices:

   a. $\begin{bmatrix} 1 & 2 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}$ c. $\begin{bmatrix} 7 \\ 5 \end{bmatrix}$

3. Each of the matrices below represents a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. Determine the values of n and m for each matrix. Then determine the dimensions of the range and kernel of L

   a. $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 4 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 2 & -1 & 3 \\ 1 & -1 & 1 & -1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$

4. For each matrix below, determine the dimensions of the range and kernel. Then decide if the linear transformation it represents is onto and/or one-to-one.

   a. $\begin{bmatrix} 1 & 2 \end{bmatrix}$ b. $\begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}$ c. $\begin{bmatrix} 7 \\ 5 \end{bmatrix}$

5. For each matrix below determine the dimensions of the range and kernel. Then decide if the linear transformation it represents is onto and/or one-to-one.

   a. $\begin{bmatrix} 1 & 1 & -1 \\ 5 & 0 & 4 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ b. $\begin{bmatrix} 2 & 3 & 5 & 7 \\ 1 & -1 & 1 & -1 \\ 5 & 0 & 5 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 1 & 0 & 0 \end{bmatrix}$

6. Compute the row rank and column rank of each of the following matrices:

   a. $\begin{bmatrix} 1 & 6 & 0 & 3 \\ 2 & -1 & 1 & 0 \end{bmatrix}$ b. $\begin{bmatrix} -3 & 6 & 4 & 1 \\ 2 & 8 & 4 & 3 \\ -4 & 1 & 0 & 0 \end{bmatrix}$ c. $\begin{bmatrix} 1 & 2 & 1 \\ 4 & 5 & 6 \\ 8 & -1 & 2 \\ 6 & 1 & 0 \end{bmatrix}$

7. Consider the following system of linear equations:

$$\begin{aligned} 2x_1 \quad - 6x_3 &= -6 \\ x_2 + \ x_3 &= \ \ 1 \end{aligned}$$

   Let A be the coefficient matrix of this system.

   a. Compute the rank of A.
   b. $\dim(\ker(A)) =$? Find a basis for $\ker(A)$.
   c. $\dim(\mathrm{Rg}(A)) =$? Find a basis for $\mathrm{Rg}(A)$.
   d. Is A a one-to-one linear transformation?
   e. Is A onto?
   f. Does the above system of equations have a solution? If yes, characterize the solution set in terms of the kernel of A and one particular solution.

8. Let $L\colon\ \mathbb{R}^3 \to \mathbb{R}^2$ have the matrix representation

$$A = \begin{bmatrix} -2 & 1 & 3 \\ 4 & -1 & 0 \end{bmatrix}$$

   Show that the range of L and the column space of A are the same subspace of $\mathbb{R}^2$.

9. Consider the following system of linear equations:

$$\begin{aligned} 4x_1 \phantom{{}-x_2} + 2x_3 + x_4 &= 0 \\ 2x_1 - x_2 + x_3 + 3x_4 &= 1 \\ -8x_1 - 2x_2 - 4x_3 + 3x_4 &= 2 \end{aligned}$$

Let A be the coefficient matrix of this system.

a. Compute the rank of A.

b. $\dim(\ker(A)) =$? Find a basis for $\ker(A)$.

c. $\dim(\mathrm{Rg}(A)) =$? Find a basis for $\mathrm{Rg}(A)$.

d. Is A a one-to-one linear transformation?

e. Is A onto?

f. Does the above system of equations have a solution? If yes, characterize the solution set in terms of the kernel of A and a particular solution.

10. Let $\boldsymbol{x}_k$ for $k = 1, 2, \ldots, p$ be vectors in a vector space. Let $V = S[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p]$. If c is any nonzero constant show that

a. $V = S[c\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p]$.

b. $V = S[\boldsymbol{x}_1, c\boldsymbol{x}_1 + \boldsymbol{x}_2, \ldots, \boldsymbol{x}_p]$.

The reader should note that the result proved in this problem is what is needed to verify Theorem 3.6, the vectors then representing the rows or columns of a matrix.

11. Given data points (1,0) and (2,1), define $L\colon\ P_2 \to \mathbb{R}^2$ by $L(\boldsymbol{p}) = (\boldsymbol{p}(1), \boldsymbol{p}(2))$. Here $\boldsymbol{p} = \boldsymbol{p}(t)$ is any polynomial of degree 2 or less. Find the matrix representation of L with respect to the standard bases in P_2 and $\mathbb{R}^2$. Show that the rank of A equals 2. What does this say about fitting polynomials of degree at most 2 through the data points (1,0) and (2,1)?

12. Let $\boldsymbol{p}$ be any polynomial in P_2. Define $L\colon\ P_2 \to \mathbb{R}^3$ by $L(\boldsymbol{p}) = (\boldsymbol{p}(-2), \boldsymbol{p}(0), \boldsymbol{p}(1))$. Find the matrix representation A of L with respect to the standard bases in P_2 and $\mathbb{R}^3$. Show that A has rank equal to 3. What does this say about fitting polynomials of degree 2 or less through three points in the plane with x coordinates -2, 0, and 1?

13. Find a polynomial in P_1, if possible, that fits the following data:

a (2,6), (3,6) b. (2,0), $(-1, 4)$ c. (2,6), (3,6), (4,7)

14. Find a polynomial in P_2, if possible, that fits the following data:

a. $(-2, 6)$, (3,7) b. $(-2, 6)$, (3,7), (4,7)
c. $(-2, 6)$, (3,7), (4,7), (5,8)

15. Let $V = C[0,1]$, the vector space of real-valued continuous functions defined on [0,1]. Define the mapping L by

$$L[\boldsymbol{f}] = \left[\frac{\boldsymbol{f}(0)}{2} + \boldsymbol{f}\left(\frac{1}{2}\right) + \frac{\boldsymbol{f}(1)}{2}\right]\left(\frac{1}{2}\right)$$

Thus, if $\boldsymbol{f}(1) = \sin t$, we have $L[\boldsymbol{f}] = [(\sin 0)/2 + \sin\frac{1}{2} + (\sin 1)/2](\frac{1}{2})$. This formula is just the trapezoid rule for the approximate evaluation of integrals over the interval [0,1] using the points 0, $\frac{1}{2}$, and 1.

   a. Show that L is a linear transformation, after deciding of course what W should be. Characterize the range and kernel of L.

   b. Let $L_1[\boldsymbol{f}] = L[\boldsymbol{f}] - \int_0^1 \boldsymbol{f}(t)dt$. Show that L_1 is also a linear transformation. How would you describe its kernel?

16. For each of the following matrices A compute the rank of A, A^T, AA^T, and A^TA. You should get rank A^TA equals rank A and rank AA^T equals rank A^T. Hence all four numbers are equal.

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

17. Let A be any $m \times n$ matrix. How are the row space of A and the range of A^T related?

18. Let x_1, x_2, and x_3 be three different numbers.

   a. Show rank $\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix} = 2$.

   b. Show rank $\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix} = 3$.

   Hint: The matrix in a is row equivalent to $\begin{bmatrix} 1 & x_1 \\ 0 & x_2 - x_1 \end{bmatrix}$.

19. Define $L\colon\ P_2 \to \mathbb{R}^3$ by $L[\boldsymbol{p}] = (\boldsymbol{p}(x_1), \boldsymbol{p}(x_2), \boldsymbol{p}(x_3))$, where the x_j's are all different. Show that L is one-to-one and onto.

20. Let $x_1, x_2, \ldots, x_n$ be n pairwise distinct numbers. Let

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{n-1} \end{bmatrix}$$

Show that rank $A = n$.

3.5 Change of Basis Formulas

In Section 3.3 we learned how to represent a linear transformation as a matrix, and in the last section we saw how the matrix can be used to tell us some facts about the kernel and range of the linear transformation. Clearly the "simpler" the matrix representation the easier it is to understand the linear transformation. Since Gaussian elimination is a method used to obtain a "simple" matrix representation for a system of equations, we have already seen the utility of finding nice representations.

Once we have a fixed linear transformation, that maps a vector space V into itself, the only variable in determining its matrix representation is our choice of basis.

Before studying how to pick such a basis (cf. Chapter 5), we need to learn how the matrix representations of the same linear transformation with respect to different bases are related to each other. Thus, let L be a linear transformation from V into V, where $\dim(V) = n$. Let $F = \{\boldsymbol{f}_j : \ j = 1, \ldots, n\}$ and $G = \{\boldsymbol{g}_j : \ j = 1, \ldots, n\}$ be two bases of V. Let $A = [a_{jk}]$ be the matrix representation of L with respect to F and let $B = [b_{jk}]$ be the matrix representation of L with respect to G. This means that the following equations hold:

$$L[\boldsymbol{f}_j] = \sum_{k=1}^{n} a_{kj} \boldsymbol{f}_k \tag{3.7}$$

$$L[\boldsymbol{g}_j] = \sum_{k=1}^{n} b_{kj} \boldsymbol{g}_k \tag{3.8}$$

Let $P = [p_{jk}]$ be the change of basis matrix that satisfies $[\boldsymbol{x}]_G^T = P[\boldsymbol{x}]_F^T$, that is

$$\boldsymbol{f}_j = \sum_{k=1}^{n} p_{kj} \boldsymbol{g}_k \tag{3.9}$$

and if $P^{-1} = [q_{jk}]$, then

$$\boldsymbol{g}_j = \sum_{k=1}^{n} q_{kj} \boldsymbol{f}_k \tag{3.10}$$

We refer the reader to Theorem 2.11 and the material preceding it. Computing $L[\boldsymbol{f}_j]$, we have

$$\begin{aligned}
L[\boldsymbol{f}_j] &= L\left[\sum_{k=1}^{n} p_{kj} \boldsymbol{g}_k\right] = \sum_{k=1}^{n} p_{kj} L[\boldsymbol{g}_k] \\
&= \sum_{k=1}^{n} p_{kj} \left(\sum_{s=1}^{n} b_{sk} \boldsymbol{g}_s\right) = \sum_{k=1}^{n} \sum_{s=1}^{n} p_{kj} b_{sk} \left(\sum_{m=1}^{n} q_{ms} \boldsymbol{f}_m\right) \\
&= \sum_{m=1}^{n} \left(\sum_{k=1}^{n} \sum_{s=1}^{n} (q_{ms} b_{sk} p_{kj})\right) \boldsymbol{f}_m
\end{aligned}$$

Comparing this expression with (3.7) and equating coefficients, we have

$$a_{mj} = \sum_{k=1}^{n} \sum_{s=1}^{n} q_{ms} b_{sk} p_{kj} \tag{3.11}$$

for m and j varying from 1 through n. However, the right-hand side of (3.11) is the m, j entry of the matrix $P^{-1}BP$. Thus we have

$$A = P^{-1}BP \quad \text{and} \quad B = PAP^{-1} \tag{3.12}$$

A better way to remember how the matrix representations A and B are related is to utilize formulas (3.3), that is,

$$[L(\boldsymbol{x})]_F^T = A[\boldsymbol{x}]_F^T \qquad [L(\boldsymbol{x})]_G^T = B[\boldsymbol{x}]_G^T$$

Thus,

$$\begin{aligned}[L(\boldsymbol{x})]_G^T &= P[L(\boldsymbol{x})]_F^T = P(A[\boldsymbol{x}]_F^T) \\ &= (PA)(P^{-1}[\boldsymbol{x}]_G^T) = (PAP^{-1})[\boldsymbol{x}]_G^T\end{aligned}$$

Hence, we must have $B = PAP^{-1}$. The reader is referred to problem 11 in Section 2.6 for a justification of this last step.

Example 1. Let L be a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ defined by $L(x_1, x_2) = (2x_1 + x_2, 3x_1 - 2x_2)$. Verify formulas (3.12) for $F = \{(1,1), (-1,2)\}$ and $G = \{(-1,-1), (2,0)\}$.

Solution. Writing the vectors in F as linear combinations of the vectors in G, we have

$$\begin{aligned}(1,1) &= -(-1,-1) + 0(2,0) \\ (-1,2) &= -2(-1,-1) + \left(-\frac{3}{2}\right)(2,0)\end{aligned}$$

Thus,

$$P = \begin{bmatrix} -1 & -2 \\ 0 & -\frac{3}{2} \end{bmatrix} \qquad P^{-1} = \begin{bmatrix} -1 & \frac{4}{3} \\ 0 & -\frac{2}{3} \end{bmatrix}$$

To determine A, we compute

$$\begin{aligned}L(1,1) = (3,1) &= \frac{7}{3}(1,1) - \frac{2}{3}(-1,2) \\ L(-1,2) = (0,-7) &= -\frac{7}{3}(1,1) - \frac{7}{3}(-1,2)\end{aligned}$$

Thus,

$$A = \begin{bmatrix} \frac{7}{3} & -\frac{7}{3} \\ -\frac{2}{3} & -\frac{7}{3} \end{bmatrix}$$

To determine B, we compute

$$L(-1,-1) = (-3,-1) = (-1,-1) - (2,0)$$
$$L(2,0) = (4,6) = -6(-1,-1) - (2,0)$$

Thus,

$$B = \begin{bmatrix} 1 & -6 \\ -1 & -1 \end{bmatrix}$$

Computing $P^{-1}BP$ we have

$$\begin{bmatrix} -1 & \frac{4}{3} \\ 0 & -\frac{2}{3} \end{bmatrix} \begin{bmatrix} 1 & -6 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ 0 & -\frac{3}{2} \end{bmatrix} = \begin{bmatrix} \frac{7}{3} & -\frac{7}{3} \\ -\frac{2}{3} & -\frac{7}{3} \end{bmatrix} = A$$

which is formula (3.12). □

Example 2. Let $L\colon\ \mathbb{R}^4 \to \mathbb{R}^4$ be a linear transformation defined by

$$L(x_1, x_2, x_3, x_4) = (2x_3 + x_4, -3x_1 + x_2 - x_4, x_1 - x_3 + 6x_4, x_2 - x_3)$$

Let S be the standard basis of $\mathbb{R}^4$ and let $G = \{(-1,0,1,1), (0,1,-1,0), (0,0,1,1), (1,0,1,0)\}$. Find the matrix representations of L with respect to S, and then G, by employing (3.12).

Solution. Let A be the matrix representation of L with respect to S. By inspection we have

$$A = \begin{bmatrix} 0 & 0 & 2 & 1 \\ -3 & 1 & 0 & -1 \\ 1 & 0 & -1 & 6 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

If P^{-1} is the change of basis matrix that satisfies $[x]_S^T = P^{-1}[x]_G^T$ then,

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

An easy computation gives us

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

Using (3.12), where B is the matrix representation of L with respect to the basis

G, we have

$$\begin{aligned}
B &= PAP^{-1} \\
&= \begin{bmatrix} -1 & 1 & 1 & -1 \\ 0 & 1 & 0 & 0 \\ 1 & -1 & -1 & 2 \\ 0 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 2 & 1 \\ -3 & 1 & 0 & -1 \\ 1 & 0 & -1 & 6 \\ 0 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 4 & 2 & 2 & -1 \\ 2 & 1 & -1 & -3 \\ -5 & 0 & -3 & 3 \\ 7 & 0 & 5 & -2 \end{bmatrix}
\end{aligned}$$

As a check on our computations we compute $L(-1,0,1,1)$ using the matrix B, and then compare this with our original definition of L. The vector $(-1,0,1,1)$ is the first vector in the basis G. Thus, the first column of B contains the coordinates of $L(-1,0,1,1)$ with respect to the basis G. Hence,

$$\begin{aligned}
L(-1,0,1,1) &= 4(-1,0,1,1) + 2(0,1,-1,0) - 5(0,0,1,1) + 7(1,0,1,0) \\
&= (3,2,4,-1)
\end{aligned}$$

This is exactly what we get when we compute $L(-1,0,1,1)$ directly. □

The preceding calculations have shown us that if $L\colon V \to V$ is a linear transformation, and A and B are two matrix representations of L with respect to two different bases of V, then there is a matrix P such that $A = P^{-1}BP$.

Definition 3.9. Let A and B be two $n \times n$ matrices. We say that A is similar to B, if there is a nonsingular matrix P such that $A = P^{-1}BP$.

As we have seen, the matrix P is nothing more than the matrix relating two different bases of our vector space. Moreover, we understand that similar matrices are just different matrix representations of the same linear transformation.

The change of basis discussion assumed that the linear transformation L mapped V into V. What happens when $L\colon V \to W$, and we change bases in V and W? Without going through the calculations, we state the appropriate theorem.

Theorem 3.8. *Let $L\colon V \to W$ be a linear transformation. Let F and $\mathscr{F}$ be two bases of V. Let G and $\mathscr{G}$ be two bases of W. Let A be the matrix representation of L using the bases F and G while B is the representation using the bases $\mathscr{F}$ and $\mathscr{G}$. Let $P = [p_{jk}]$ be the change of basis matrix that satisfies $[\boldsymbol{x}]^T_{\mathscr{F}} = P[\boldsymbol{x}]^T_F$. Let $Q = [q_{jk}]$ be the matrix that satisfies $[\boldsymbol{y}]^T_{\mathscr{G}} = Q[\boldsymbol{y}]^T_G$. Then*

$$A = Q^{-1}BP \tag{3.13}$$

The reader should look at problem 9 at the end of this section for one method of proving this result.

Problem Set 3.5

1. Let $L\colon\ \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $L(x_1, x_2) = (2x_1, 2x_2)$. Let $F = \{(1, -1), (2, 5)\}$.

   a. Find the matrix representation A of L with respect to the standard basis S.

   b. Let P be the change of basis matrix that satisfies $[\boldsymbol{x}]_F^T = P[\boldsymbol{x}]_S^T$. Find P and P^{-1}.

   c. Find the matrix representation B of L with respect to the basis F by using (3.12).

2. Let A be any 2×2 scalar matrix, that is

$$A = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} = cI_2$$

   Let P be any 2×2 nonsingular matrix. Show that PAP^{-1} equals A; cf. problem 1.

3. Let $L\colon\ \mathbb{R}^3 \to \mathbb{R}^3$. Let $F = \{(1, 2, -3), (1, 0, 0), (0, 1, 0)\}$. Suppose

$$A = \begin{bmatrix} 0 & 1 & -2 \\ 2 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$$

   is the matrix representation of L with respect to the basis F. Use (3.12) to find the matrix representation of L with respect to the standard basis; cf. problem 4 of Section 3.2.

4. Suppose $A = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$ is the matrix representation of a linear transformation with respect to the standard basis. Let $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and set $B = PAP^{-1}$. Then B can be thought of as the matrix representation of L with respect to some other basis. What is this basis?

5. Let A, B, and C be three $n \times n$ matrices. Show that

   a. A is similar to itself.

   b. If A is similar to B, then B is similar to A.

   c. If A is similar to B and B is similar to C, then A is similar to C.

6. Let A be any 2×2 matrix. Show that if A is similar to I_2, then $A = I_2$.

7. Let $F = \{(1, 2), (1, 0)\}$ and $G = \{(1, -1), (0, 1)\}$. Let $L(x_1, x_2) = (x_1 + x_2, 2x_1 - x_2)$. Find the matrix representations of L with respect to bases F and G. Verify (3.12).

8. Let $F = \{(1,0,1),(0,1,1),(1,1,0)\}$ and $G = \{(1,0,-1),(-1,1,0),(0,0,1)\}$. Let $L(x_1,x_2,x_3) = (x_1+x_2+x_3, x_2+x_3, x_3)$. Verify (3.12) for the above bases and linear transformation L.

9. Prove Theorem 3.8. Hint: Use problem 11 in Section 2.6 and formula (3.3).

10. Let $F_1 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

    Let $F_2 = \left\{ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\}$

    Let $G_1 = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}$

    Let $G_2 = \left\{ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \right\}$

    Let ${}_jA_k$ be the matrix representation of a linear transformation $L\colon M_{22} \to M_{23}$ with respect to the bases G_j and F_k. That is,

    $$[L(\boldsymbol{x})]^T_{G_j} = {}_jA_k[\boldsymbol{x}]^T_{F_k}$$

    Suppose that

    $${}_1A_1 = \begin{bmatrix} -2 & 0 & -3 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 4 & 1 \\ 3 & 1 & -5 & 0 \\ 2 & -2 & 0 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

    Find the other three matrix representations.

11. Suppose A and B are similar $n \times n$ matrices, i.e., there is a matrix P such that $A = PBP^{-1}$.

    a. How are $\ker(A)$ and $\ker(B)$ related?
    b. How are $\mathrm{Rg}(A)$ and $\mathrm{Rg}(B)$ related?

12. Let $L\colon P_2 \to P_2$ be a linear transformation. Let $F = \{t^2+t-1, t^2+2, t-6\}$. Suppose the matrix representation of L with respect to F is

    $$A = \begin{bmatrix} -14 & -2 & -18 \\ 23 & 11 & 18 \\ 11 & 2 & 15 \end{bmatrix}$$

    Find the matrix representation of L with respect to the standard basis of P_2.

Supplementary Problems

1. Define and give examples of each of the following:

   a. Linear transformation

   b. Range and kernel of a linear transformation

   c. Rank of a matrix

   d. Column (row) rank

2. Let $A = \begin{bmatrix} 2 & 6 \\ 1 & -2 \\ 3 & 5 \end{bmatrix}$ be the matrix representation of a linear transformation from $\mathbb{R}^k$ to $\mathbb{R}^P$.

   a. $k = ? p = ?$

   b. Determine the dimensions of ker(A) and Rg(A).

3. For each of the following matrices determine the rank and find a basis for the kernel:

   a. $\begin{bmatrix} 3 & 1 & 2 \\ 1 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix}$ b. $\begin{bmatrix} -1 & 2 & 0 & 6 \\ 4 & 3 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ c. $\begin{bmatrix} 4 & 2 & 4 & 6 \\ 6 & 3 & 6 & 9 \\ 2 & 1 & 2 & 1 \end{bmatrix}$

4. Let A and B be two matrices. Show that

$$\text{Rank}(AB) \leq \text{minimum}\{\text{rank}(A), \text{rank}(B)\}$$

   Hint: Rank (A) equals dim(Rg(A)).

5. Define $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ by $L(x_1, x_2) = (x_1 + x_2, -x_2)$. Show that L maps the straight line $y = mx$ onto the straight line $y = -[m/(m+1)]x$. What happens to the line $y = -x$?

6. Describe, geometrically, the following linear transformations, and in each case determine the kernel and range:

   a. $L(x_1, x_2) = (x_1 - x_2, x_1 + x_2)$

   b. $L(x_1, x_2) = (2x_1 - x_2, 4x_1 - 2x_2)$

7. Let V_1, V_2, and V_3 be vector spaces with bases F_1, F_2, and F_3, respectively. Suppose $L_1\colon V_1 \to V_2$ and $L_2\colon V_2 \to V_3$ are linear transformations. Let A_1 and A_2 be their corresponding matrix representations with respect to the given bases. Define $L\colon V_1 \to V_3$ by $L(\boldsymbol{x}) = L_2(L_1(\boldsymbol{x}))$.

   a Show that L is a linear transformation from V_1 to V_3.

   b. Show that the matrix representation of L with respect to the bases F_1 and F_3 is A_2A_1.

8. A linear transformation $L\colon\ \mathbb{R}^2 \to \mathbb{R}^2$ is said to be positive if it maps the first quadrant into the first quadrant; that is, if x_1 and x_2 are both positive, then so are y_1 and y_2, where $L(x_1, x_2) = (y_1, y_2)$.

   a. Let $L(x_1, x_2) = (ax_1 + bx_2, cx_1 + dx_2)$. Show that L is positive if and only if a, b, c, and d are all non-negative and at least one of $\{a, b\}$ is positive and one of $\{c, d\}$ is positive.

   b. Find a positive linear transformation whose matrix representation with respect to some basis has at least one negative entry.

   c. Find an example of a positive linear transformation whose kernel has dimension equal to 1.

9. A mapping T from a vector space V into V is said to be an affine transformation if

$$T(\boldsymbol{x}) = L(\boldsymbol{x}) + \boldsymbol{a}$$

   where L is a linear transformation from V into V and a is any fixed vector in V.

   a. Show that $T(\boldsymbol{x}) = L(\boldsymbol{x}) + \boldsymbol{a}$ is a linear transformation if and only if a equals the zero vector.

   b. Given any straight line in $\mathbb{R}^2$, show that there is an affine transformation that maps the line $x_2 = 0$ onto the given line. Hint: Rotate and then translate.

10. Two vector spaces V_1 and V_2 are said to be isomorphic if there is a linear transformation $L\colon\ V_1 \to V_2$ that is both one-to-one and onto.

   a. Suppose $\dim(V_1) = m$ and V_1 is isomorphic to V_2. Show $\dim(V_2) = m$.

   b. Show that two finite-dimensional vector spaces are isomorphic if and only if they have the same dimension.

   c. Let $V_1 = \mathbb{R}^1$ and let V_2 be the vector space defined in Example 5 Section 2.2. Define $L\colon\ V_1 \to V_2$ by $L(x) = e^x$. Show that L is a linear transformation that is one-to-one and onto.

11. Let $V = \{\sum_{k=1}^{n} a_k \sin(kx)e^{-k^2 t}, n$ any positive integer and a_k arbitrary numbers$\}$.

   a. Under ordinary addition and multiplication show that V is a vector space. Define $L\colon\ V \to V$ by $L[u] = \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2}$.

   b. Show L is a linear transformation.

   c. Find the kernel and range of L.

12. If V is a finite-dimensional space and L is a linear transformation from V into V, then L is one-to-one if and only if L is onto. This result, as the following shows, may not be true if V is infinite-dimensional. Let V be the vector space of all polynomials in the variable t.

    a. Define $L[\boldsymbol{p}](t) = t\boldsymbol{p}(t)$. Thus, $L[\boldsymbol{t}] = \boldsymbol{t}^2$, $L[\boldsymbol{t}^2] = \boldsymbol{t}^3$, and $L[\boldsymbol{t}^n] = \boldsymbol{t}^{n+1}$. Show that L is one-to-one but not onto.

    b. Define $L[\boldsymbol{p}](t) = \boldsymbol{p}'(t)$. Thus, $L[\boldsymbol{t}^n] = n\boldsymbol{t}^{n-1}$. Show that L is onto but not one-to-one. What is the kernel of L?

13. Let L be a linear transformation from V into V. A subspace W of V is said to be invariant under L if, for every $\boldsymbol{x}$ in W, $L[\boldsymbol{x}]$ is also in W. Thus, we may also consider L to be a linear transformation from W to W. Show that for any linear transformation L each of the following is an invariant subspace:

    a. $W = \{0\}, W = V$.

    b. For any constant λ, show $W_\lambda = \{\boldsymbol{x}\colon\ L[\boldsymbol{x}] = \lambda x\}$ is an invariant subspace.

14. Let $A = \begin{bmatrix} 2 & 1 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ be the matrix representation of a linear transformation L with respect to the standard basis. Show that $S[\boldsymbol{e}_1]$ is not an invariant subspace but that $S[\boldsymbol{e}_1, \boldsymbol{e}_2]$ and $S[\boldsymbol{e}_3]$ are invariant subspaces; cf. problem 13.

15. Let $F = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ be a basis for a vector space V. Let L be a linear transformation from V into V for which $S[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k]$ and $S[\boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_n]$ are both invariant subspaces of L. Let A be the matrix representation of L with respect to the basis F. Show that

$$A = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}$$

    where B_1 is a $k \times k$ matrix and B_2 is an $(n-k)\times(n-k)$ matrix. Conversely, suppose that A has the above form. Show that the above two subspaces must be invariant.