lyryx with Open Texts

LINEAR ALGEBRA with Applications

Open Edition

Adapted for

Emory University

Math 221

Linear Algebra

Sections 1 & 2

Lectured and adapted by

Le Chen

April 15, 2021

le.chen@emory.edu

Course page

http://math.emory.edu/~lchen41/teaching/2021_Spring_Math221

by W. Keith Nicholson

Contents

2.6 Linear Transformations

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

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

is called the *matrix transformation induced* by A. In Section 2.2, we saw that many important geometric transformations were in fact matrix transformations. These transformations can be characterized in a different way. The new idea is that of a linear transformation, one of the basic notions in linear algebra. We define these transformations in this section, and show that they are really just the matrix transformations looked at in another way. Having these two ways to view them turns out to be useful because, in a given situation, one perspective or the other may be preferable.

Linear Transformations

Definition 2.13 Linear Transformations $\mathbb{R}^n \to \mathbb{R}^m$

A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a **linear transformation** *if it satisfies the following two conditions for all vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ and all scalars a:*

T1 $\quad T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$

T2 $\quad T(a\mathbf{x}) = aT(\mathbf{x})$

Of course, $\mathbf{x}+\mathbf{y}$ and $a\mathbf{x}$ here are computed in $\mathbb{R}^n$, while $T(\mathbf{x})+T(\mathbf{y})$ and $aT(\mathbf{x})$ are in $\mathbb{R}^m$. We say that T *preserves addition* if T1 holds, and that T *preserves scalar multiplication* if T2 holds. Moreover, taking $a = 0$ and $a = -1$ in T2 gives

$$T(\mathbf{0}) = \mathbf{0} \quad \text{and} \quad T(-\mathbf{x}) = -T(\mathbf{x}) \quad \text{for all } \mathbf{x}$$

Hence T preserves the zero vector and the negative of a vector. Even more is true.

Recall that a vector $\mathbf{y}$ in $\mathbb{R}^n$ is called a **linear combination** of vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ if $\mathbf{y}$ has the form

$$\mathbf{y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k$$

for some scalars $a_1, a_2, \dots, a_k$. Conditions T1 and T2 combine to show that every linear transformation T *preserves linear combinations* in the sense of the following theorem. This result is used repeatedly in linear algebra.

Theorem 2.6.1: Linearity Theorem

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then for each $k = 1, 2, \dots$

$$T(a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k) = a_1T(\mathbf{x}_1) + a_2T(\mathbf{x}_2) + \cdots + a_kT(\mathbf{x}_k)$$

for all scalars a_i and all vectors $\mathbf{x}_i$ in $\mathbb{R}^n$.

Proof. If $k=1$, it reads $T(a_1\mathbf{x}_1)=a_1T(\mathbf{x}_1)$ which is Condition T1. If $k=2$, we have

$$\begin{array}{rcll} T(a_1\mathbf{x}_1+a_2\mathbf{x}_2) & = & T(a_1\mathbf{x}_1)+T(a_2\mathbf{x}_2) & \text{by Condition T1} \\ & = & a_1T(\mathbf{x}_1)+a_2T(\mathbf{x}_2) & \text{by Condition T2} \end{array}$$

If $k=3$, we use the case $k=2$ to obtain

$$\begin{array}{rcll} T(a_1\mathbf{x}_1+a_2\mathbf{x}_2+a_3\mathbf{x}_3) & = & T\left[(a_1\mathbf{x}_1+a_2\mathbf{x}_2)+a_3\mathbf{x}_3\right] & \text{collect terms} \\ & = & T(a_1\mathbf{x}_1+a_2\mathbf{x}_2)+T(a_3\mathbf{x}_3) & \text{by Condition T1} \\ & = & \left[a_1T(\mathbf{x}_1)+a_2T(\mathbf{x}_2)\right]+T(a_3\mathbf{x}_3) & \text{by the case } k=2 \\ & = & \left[a_1T(\mathbf{x}_1)+a_2T(\mathbf{x}_2)\right]+a_3T(\mathbf{x}_3) & \text{by Condition T2} \end{array}$$

The proof for any k is similar, using the previous case $k-1$ and Conditions T1 and T2. □

The method of proof in Theorem 2.6.1 is called *mathematical induction* (Appendix **??**).

Theorem 2.6.1 shows that if T is a linear transformation and $T(\mathbf{x}_1)$, $T(\mathbf{x}_2)$, $\dots$, $T(\mathbf{x}_k)$ are all known, then $T(\mathbf{y})$ can be easily computed for any linear combination $\mathbf{y}$ of $\mathbf{x}_1$, $\mathbf{x}_2$, $\dots$, $\mathbf{x}_k$. This is a very useful property of linear transformations, and is illustrated in the next example.

Example 2.6.1

If $T:\mathbb{R}^2\to\mathbb{R}^2$ is a linear transformation, $T\left[\begin{array}{c}1\\1\end{array}\right]=\left[\begin{array}{c}2\\-3\end{array}\right]$ and $T\left[\begin{array}{c}1\\-2\end{array}\right]=\left[\begin{array}{c}5\\1\end{array}\right]$, find $T\left[\begin{array}{c}4\\3\end{array}\right]$.

Solution. Write $\mathbf{z}=\left[\begin{array}{c}4\\3\end{array}\right]$, $\mathbf{x}=\left[\begin{array}{c}1\\1\end{array}\right]$, and $\mathbf{y}=\left[\begin{array}{c}1\\-2\end{array}\right]$ for convenience. Then we know $T(\mathbf{x})$ and $T(\mathbf{y})$ and we want $T(\mathbf{z})$, so it is enough by Theorem 2.6.1 to express $\mathbf{z}$ as a linear combination of $\mathbf{x}$ and $\mathbf{y}$. That is, we want to find numbers a and b such that $\mathbf{z}=a\mathbf{x}+b\mathbf{y}$. Equating entries gives two equations $4=a+b$ and $3=a-2b$. The solution is, $a=\frac{11}{3}$ and $b=\frac{1}{3}$, so $\mathbf{z}=\frac{11}{3}\mathbf{x}+\frac{1}{3}\mathbf{y}$. Thus Theorem 2.6.1 gives

$$T(\mathbf{z})=\tfrac{11}{3}T(\mathbf{x})+\tfrac{1}{3}T(\mathbf{y})=\tfrac{11}{3}\left[\begin{array}{c}2\\-3\end{array}\right]+\tfrac{1}{3}\left[\begin{array}{c}5\\1\end{array}\right]=\tfrac{1}{3}\left[\begin{array}{c}27\\-32\end{array}\right]$$

This is what we wanted.

Example 2.6.2

If A is $m\times n$, the matrix transformation $T_A:\mathbb{R}^n\to\mathbb{R}^m$, is a linear transformation.

Solution. We have $T_A(\mathbf{x})=A\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$, so Theorem 2.2.2 gives

$$T_A(\mathbf{x}+\mathbf{y})=A(\mathbf{x}+\mathbf{y})=A\mathbf{x}+A\mathbf{y}=T_A(\mathbf{x})+T_A(\mathbf{y})$$

and

$$T_A(a\mathbf{x})=A(a\mathbf{x})=a(A\mathbf{x})=aT_A(\mathbf{x})$$

hold for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ and all scalars a. Hence T_A satisfies T1 and T2, and so is linear.

The remarkable thing is that the *converse* of Example 2.6.2 is true: Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is actually a matrix transformation. To see why, we define the **standard basis** of $\mathbb{R}^n$ to be the set of columns

$$\{\mathbf{e}_1,\ \mathbf{e}_2,\ \dots,\ \mathbf{e}_n\}$$

of the identity matrix I_n. Then each $\mathbf{e}_i$ is in $\mathbb{R}^n$ and every vector $\mathbf{x} = \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right]$ in $\mathbb{R}^n$ is a linear combination of the $\mathbf{e}_i$. In fact:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$$

as the reader can verify. Hence Theorem 2.6.1 shows that

$$T(\mathbf{x}) = T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n)$$

Now observe that each $T(\mathbf{e}_i)$ is a column in $\mathbb{R}^m$, so

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

is an $m \times n$ matrix. Hence we can apply Definition 2.5 to get

$$T(\mathbf{x}) = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n) = \left[\begin{array}{cccc} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right] = A\mathbf{x}$$

Since this holds for every $\mathbf{x}$ in $\mathbb{R}^n$, it shows that T is the matrix transformation induced by A, and so proves most of the following theorem.

Theorem 2.6.2

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a transformation.

1. *T is linear if and only if it is a matrix transformation.*
2. *In this case $T = T_A$ is the matrix transformation induced by a unique $m \times n$ matrix A, given in terms of its columns by*

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

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

Proof. It remains to verify that the matrix A is unique. Suppose that T is induced by another matrix B. Then $T(\mathbf{x}) = B\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. But $T(\mathbf{x}) = A\mathbf{x}$ for each $\mathbf{x}$, so $B\mathbf{x} = A\mathbf{x}$ for every $\mathbf{x}$. Hence $A = B$ by Theorem 2.2.6. □

Hence we can speak of *the* matrix of a linear transformation. Because of Theorem 2.6.2 we may (and shall) use the phrases "linear transformation" and "matrix transformation" interchangeably.

Example 2.6.3

Define $T : \mathbb{R}^3 \to \mathbb{R}^2$ by $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ for all $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ in $\mathbb{R}^3$. Show that T is a linear transformation and use Theorem 2.6.2 to find its matrix.

Solution. Write $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, so that $\mathbf{x}+\mathbf{y} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ x_3+y_3 \end{bmatrix}$. Hence

$$T(\mathbf{x}+\mathbf{y}) = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = T(\mathbf{x}) + T(\mathbf{y})$$

Similarly, the reader can verify that $T(a\mathbf{x}) = aT(\mathbf{x})$ for all a in $\mathbb{R}$, so T is a linear transformation. Now the standard basis of $\mathbb{R}^3$ is

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

so, by Theorem 2.6.2, the matrix of T is

$$A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Of course, the fact that $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ shows directly that T is a matrix transformation (hence linear) and reveals the matrix.

To illustrate how Theorem 2.6.2 is used, we rederive the matrices of the transformations in Examples 2.2.13 and 2.2.15.

Example 2.6.4

Let $Q_0 : \mathbb{R}^2 \to \mathbb{R}^2$ denote reflection in the x axis (as in Example 2.2.13) and let $R_{\frac{\pi}{2}} : \mathbb{R}^2 \to \mathbb{R}^2$ denote counterclockwise rotation through $\frac{\pi}{2}$ about the origin (as in Example 2.2.15). Use Theorem 2.6.2 to find the matrices of Q_0 and $R_{\frac{\pi}{2}}$.

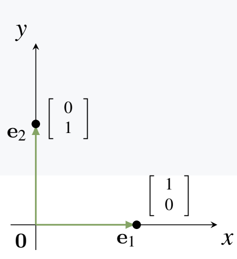


Figure 2.6.1

Solution. Observe that Q_0 and $R_{\frac{\pi}{2}}$ are linear by Example 2.6.2 (they are matrix transformations), so Theorem 2.6.2 applies to them. The standard basis of $\mathbb{R}^2$ is

$\{\mathbf{e}_1,\ \mathbf{e}_2\}$ where $\mathbf{e}_1 = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$ points along the positive x axis, and $\mathbf{e}_2 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$ points along the positive y axis (see Figure 2.6.1).

The reflection of $\mathbf{e}_1$ in the x axis is $\mathbf{e}_1$ itself because $\mathbf{e}_1$ points along the x axis, and the reflection of $\mathbf{e}_2$ in the x axis is $-\mathbf{e}_2$ because $\mathbf{e}_2$ is perpendicular to the x axis. In other words, $Q_0(\mathbf{e}_1) = \mathbf{e}_1$ and $Q_0(\mathbf{e}_2) = -\mathbf{e}_2$. Hence Theorem 2.6.2 shows that the matrix of Q_0 is

$$\left[\begin{array}{cc} Q_0(\mathbf{e}_1) & Q_0(\mathbf{e}_2) \end{array}\right] = \left[\begin{array}{cc} \mathbf{e}_1 & -\mathbf{e}_2 \end{array}\right] = \left[\begin{array}{rr} 1 & 0 \\ 0 & -1 \end{array}\right]$$

which agrees with Example 2.2.13.

Similarly, rotating $\mathbf{e}_1$ through $\frac{\pi}{2}$ counterclockwise about the origin produces $\mathbf{e}_2$, and rotating $\mathbf{e}_2$ through $\frac{\pi}{2}$ counterclockwise about the origin gives $-\mathbf{e}_1$. That is, $R_{\frac{\pi}{2}}(\mathbf{e}_1) = \mathbf{e}_2$ and $R_{\frac{\pi}{2}}(\mathbf{e}_2) = -\mathbf{e}_2$. Hence, again by Theorem 2.6.2, the matrix of $R_{\frac{\pi}{2}}$ is

$$\left[\begin{array}{cc} R_{\frac{\pi}{2}}(\mathbf{e}_1) & R_{\frac{\pi}{2}}(\mathbf{e}_2) \end{array}\right] = \left[\begin{array}{cc} \mathbf{e}_2 & -\mathbf{e}_1 \end{array}\right] = \left[\begin{array}{rr} 0 & -1 \\ 1 & 0 \end{array}\right]$$

agreeing with Example 2.2.15.

Example 2.6.5

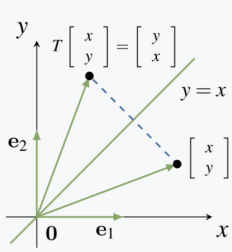


Figure 2.6.2

Let $Q_1 : \mathbb{R}^2 \to \mathbb{R}^2$ denote reflection in the line $y = x$. Show that Q_1 is a matrix transformation, find its matrix, and use it to illustrate Theorem 2.6.2.

Solution. Figure 2.6.2 shows that $Q_1 \left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{c} y \\ x \end{array}\right]$. Hence $Q_1 \left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right] \left[\begin{array}{c} y \\ x \end{array}\right]$, so Q_1 is the matrix transformation induced by the matrix $A = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]$. Hence Q_1 is linear (by Example 2.6.2) and so Theorem 2.6.2 applies. If $\mathbf{e}_1 = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$ and $\mathbf{e}_2 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$ are the standard basis of $\mathbb{R}^2$, then it is clear geometrically that $Q_1(\mathbf{e}_1) = \mathbf{e}_2$ and $Q_1(\mathbf{e}_2) = \mathbf{e}_1$. Thus (by Theorem 2.6.2) the matrix of Q_1 is $\left[\begin{array}{cc} Q_1(\mathbf{e}_1) & Q_1(\mathbf{e}_2) \end{array}\right] = \left[\begin{array}{cc} \mathbf{e}_2 & \mathbf{e}_1 \end{array}\right] = A$ as before.

Recall that, given two "linked" transformations

$$\mathbb{R}^k \xrightarrow{T} \mathbb{R}^n \xrightarrow{S} \mathbb{R}^m$$

we can apply T first and then apply S, and so obtain a new transformation

$$S \circ T : \mathbb{R}^k \to \mathbb{R}^m$$

called the **composite** of S and T, defined by

$$(S \circ T)(\mathbf{x}) = S[T(\mathbf{x})] \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^k$$

If S and T are linear, the action of $S \circ T$ can be computed by multiplying their matrices.

Theorem 2.6.3

Let $\mathbb{R}^k \xrightarrow{T} \mathbb{R}^n \xrightarrow{S} \mathbb{R}^m$ be linear transformations, and let A and B be the matrices of S and T respectively. Then $S \circ T$ is linear with matrix AB.

Proof. $(S \circ T)(\mathbf{x}) = S[T(\mathbf{x})] = A[B\mathbf{x}] = (AB)\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^k$. □

Theorem 2.6.3 shows that the action of the composite $S \circ T$ is determined by the matrices of S and T. But it also provides a very useful interpretation of matrix multiplication. If A and B are matrices, the product matrix AB induces the transformation resulting from first applying B and then applying A. Thus the study of matrices can cast light on geometrical transformations and vice-versa. Here is an example.

Example 2.6.6

Show that reflection in the x axis followed by rotation through $\frac{\pi}{2}$ is reflection in the line $y = x$.

Solution. The composite in question is $R_{\frac{\pi}{2}} \circ Q_0$ where Q_0 is reflection in the x axis and $R_{\frac{\pi}{2}}$ is rotation through $\frac{\pi}{2}$. By Example 2.6.4, $R_{\frac{\pi}{2}}$ has matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and Q_0 has matrix $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Hence Theorem 2.6.3 shows that the matrix of $R_{\frac{\pi}{2}} \circ Q_0$ is $AB = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, which is the matrix of reflection in the line $y = x$ by Example 2.6.3.

This conclusion can also be seen geometrically. Let $\mathbf{x}$ be a typical point in $\mathbb{R}^2$, and assume that $\mathbf{x}$ makes an angle α with the positive x axis. The effect of first applying Q_0 and then applying $R_{\frac{\pi}{2}}$ is shown in Figure 2.6.3. The fact that $R_{\frac{\pi}{2}}[Q_0(\mathbf{x})]$ makes the angle α with the positive y axis shows that $R_{\frac{\pi}{2}}[Q_0(\mathbf{x})]$ is the reflection of $\mathbf{x}$ in the line $y = x$.

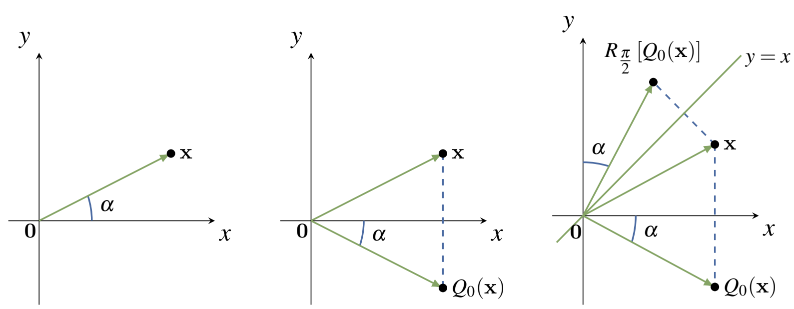


Figure 2.6.3

In Theorem 2.6.3, we saw that the matrix of the composite of two linear transformations is the product of their matrices (in fact, matrix products were defined so that this is the case). We are going to apply this fact to rotations, reflections, and projections in the plane. Before proceeding, we pause to present useful geometrical descriptions of vector addition and scalar multiplication in the plane, and to give a short review of angles and the trigonometric functions.

Some Geometry

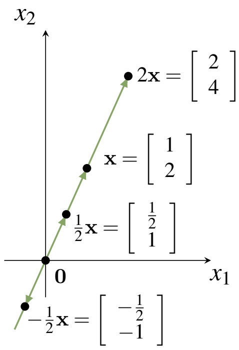


Figure 2.6.4

As we have seen, it is convenient to view a vector $\mathbf{x}$ in $\mathbb{R}^2$ as an arrow from the origin to the point $\mathbf{x}$ (see Section 2.2). This enables us to visualize what sums and scalar multiples mean geometrically. For example consider $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ in $\mathbb{R}^2$. Then $2\mathbf{x} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $\frac{1}{2}\mathbf{x} = \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$ and $-\frac{1}{2}\mathbf{x} = \begin{bmatrix} -\frac{1}{2} \\ -1 \end{bmatrix}$, and these are shown as arrows in Figure 2.6.4.

Observe that the arrow for $2\mathbf{x}$ is twice as long as the arrow for $\mathbf{x}$ and in the same direction, and that the arrows for $\frac{1}{2}\mathbf{x}$ is also in the same direction as the arrow for $\mathbf{x}$, but only half as long. On the other hand, the arrow for $-\frac{1}{2}\mathbf{x}$ is half as long as the arrow for $\mathbf{x}$, but in the *opposite* direction. More generally, we have the following geometrical description of scalar multiplication in $\mathbb{R}^2$:

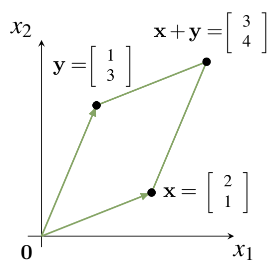


Figure 2.6.5

Scalar Multiple Law

Let $\mathbf{x}$ be a vector in $\mathbb{R}^2$. The arrow for $k\mathbf{x}$ is $|k|$ times[12] as long as the arrow for $\mathbf{x}$, and is in the same direction as the arrow for $\mathbf{x}$ if $k > 0$, and in the opposite direction if $k < 0$.

Now consider two vectors $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ in $\mathbb{R}^2$. They are plotted in Figure 2.6.5 along with their sum $\mathbf{x}+\mathbf{y} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. It is a routine matter to verify that the four points $\mathbf{0}$, $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{x}+\mathbf{y}$ form the vertices of a **parallelogram**–that is opposite sides are parallel and of the same length. (The reader should verify that the side from $\mathbf{0}$ to $\mathbf{x}$ has slope of $\frac{1}{2}$, as does the side from $\mathbf{y}$ to $\mathbf{x}+\mathbf{y}$, so these sides are parallel.) We state this as follows:

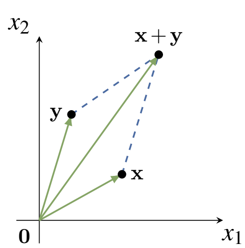


Figure 2.6.6

Parallelogram Law

Consider vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^2$. If the arrows for $\mathbf{x}$ and $\mathbf{y}$ are drawn (see Figure 2.6.6), the arrow for $\mathbf{x}+\mathbf{y}$ corresponds to the fourth vertex of the parallelogram determined by the

[12] If k is a real number, $|k|$ denotes the **absolute value** of k; that is, $|k| = k$ if $k \geq 0$ and $|k| = -k$ if $k < 0$.

Figure 2.6.7

points $\mathbf{x}$, $\mathbf{y}$, *and* $\mathbf{0}$.

We will have more to say about this in Chapter 4.

Before proceeding we turn to a brief review of angles and the trigonometric functions. Recall that an angle θ is said to be in **standard position** if it is measured counterclockwise from the positive x axis (as in Figure 2.6.7). Then θ uniquely determines a point $\mathbf{p}$ on the **unit circle** (radius 1, centre at the origin). The **radian** measure of θ is the length of the arc on the unit circle from the positive x axis to $\mathbf{p}$. Thus $360° = 2\pi$ radians, $180° = \pi$, $90° = \frac{\pi}{2}$, and so on.

The point $\mathbf{p}$ in Figure 2.6.7 is also closely linked to the trigonometric functions **cosine** and **sine**, written $\cos\theta$ and $\sin\theta$ respectively. In fact these functions are *defined* to be the x and y coordinates of $\mathbf{p}$; that is $\mathbf{p} = \left[\begin{array}{c} \cos\theta \\ \sin\theta \end{array}\right]$. This defines $\cos\theta$ and $\sin\theta$ for the arbitrary angle θ (possibly negative), and agrees with the usual values when θ is an acute angle $\left(0 \leq \theta \leq \frac{\pi}{2}\right)$ as the reader should verify. For more discussion of this, see Appendix **??**.

Rotations

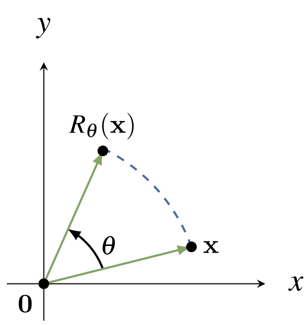


Figure 2.6.8

We can now describe rotations in the plane. Given an angle θ, let

$$R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$$

denote counterclockwise rotation of $\mathbb{R}^2$ about the origin through the angle θ. The action of R_θ is depicted in Figure 2.6.8. We have already looked at $R_{\frac{\pi}{2}}$ (in Example 2.2.15) and found it to be a matrix transformation. It turns out that R_θ is a matrix transformation for *every* angle θ (with a simple formula for the matrix), but it is not clear how to find the matrix. Our approach is to first establish the (somewhat surprising) fact that R_θ is *linear*, and then obtain the matrix from Theorem 2.6.2.

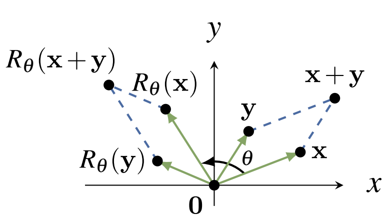


Figure 2.6.9

Let $\mathbf{x}$ and $\mathbf{y}$ be two vectors in $\mathbb{R}^2$. Then $\mathbf{x}+\mathbf{y}$ is the diagonal of the parallelogram determined by $\mathbf{x}$ and $\mathbf{y}$ as in Figure 2.6.9.

The effect of R_θ is to rotate the *entire* parallelogram to obtain the new parallelogram determined by $R_\theta(\mathbf{x})$ and $R_\theta(\mathbf{y})$, with diagonal $R_\theta(\mathbf{x}+\mathbf{y})$. But this diagonal is $R_\theta(\mathbf{x})+R_\theta(\mathbf{y})$ by the parallelogram law (applied to the new parallelogram). It follows that

$$R_\theta(\mathbf{x}+\mathbf{y}) = R_\theta(\mathbf{x}) + R_\theta(\mathbf{y})$$

A similar argument shows that $R_\theta(a\mathbf{x}) = aR_\theta(\mathbf{x})$ for any scalar a, so $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ is indeed a linear transformation.

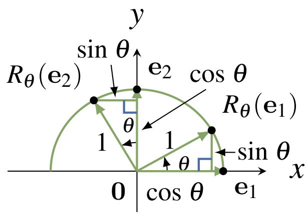


Figure 2.6.10

With linearity established we can find the matrix of R_θ. Let $\mathbf{e}_1 = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$ and $\mathbf{e}_2 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$ denote the standard basis of $\mathbb{R}^2$. By Figure 2.6.10 we see that

$$R_\theta(\mathbf{e}_1) = \left[\begin{array}{c} \cos\theta \\ \sin\theta \end{array}\right] \quad \text{and} \quad R_\theta(\mathbf{e}_2) = \left[\begin{array}{c} -\sin\theta \\ \cos\theta \end{array}\right]$$

Hence Theorem 2.6.2 shows that R_θ is induced by the matrix

$$\left[\begin{array}{cc} R_\theta(\mathbf{e}_1) & R_\theta(\mathbf{e}_2) \end{array}\right] = \left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right]$$

We record this as

Theorem 2.6.4

The rotation $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear transformation with matrix $\left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right]$.

For example, $R_{\frac{\pi}{2}}$ and R_π have matrices $\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right]$ and $\left[\begin{array}{cc} -1 & 0 \\ 0 & -1 \end{array}\right]$, respectively, by Theorem 2.6.4. The first of these confirms the result in Example 2.2.15. The second shows that rotating a vector $\mathbf{x} = \left[\begin{array}{c} x \\ y \end{array}\right]$ through the angle π results in $R_\pi(\mathbf{x}) = \left[\begin{array}{cc} -1 & 0 \\ 0 & -1 \end{array}\right] \left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{c} -x \\ -y \end{array}\right] = -\mathbf{x}$. Thus applying R_π is the same as negating $\mathbf{x}$, a fact that is evident without Theorem 2.6.4.

Example 2.6.7

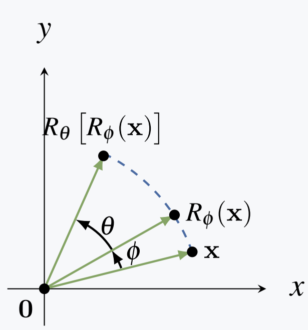


Figure 2.6.11

Let θ and ϕ be angles. By finding the matrix of the composite $R_\theta \circ R_\phi$, obtain expressions for $\cos(\theta+\phi)$ and $\sin(\theta+\phi)$.

Solution. Consider the transformations $\mathbb{R}^2 \xrightarrow{R_\phi} \mathbb{R}^2 \xrightarrow{R_\theta} \mathbb{R}^2$. Their composite $R_\theta \circ R_\phi$ is the transformation that first rotates the plane through ϕ and then rotates it through θ, and so is the rotation through the angle $\theta+\phi$ (see Figure 2.6.11). In other words

$$R_{\theta+\phi} = R_\theta \circ R_\phi$$

Theorem 2.6.3 shows that the corresponding equation holds for the matrices of these transformations, so Theorem 2.6.4 gives:

$$\left[\begin{array}{cc} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{array}\right] = \left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right] \left[\begin{array}{cc} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{array}\right]$$

If we perform the matrix multiplication on the right, and then compare first column entries, we obtain

$$\begin{aligned} \cos(\theta+\phi) &= \cos\theta\cos\phi - \sin\theta\sin\phi \\ \sin(\theta+\phi) &= \sin\theta\cos\phi - \cos\theta\sin\phi \end{aligned}$$

These are the two basic identities from which most of trigonometry can be derived.

Reflections

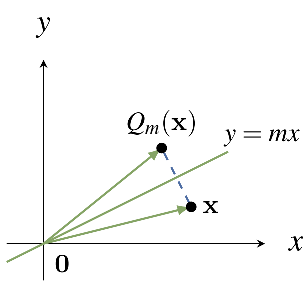


Figure 2.6.12

The line through the origin with slope m has equation $y = mx$, and we let $Q_m : \mathbb{R}^2 \to \mathbb{R}^2$ denote reflection in the line $y = mx$.

This transformation is described geometrically in Figure 2.6.12. In words, $Q_m(\mathbf{x})$ is the "mirror image" of $\mathbf{x}$ in the line $y = mx$. If $m = 0$ then Q_0 is reflection in the x axis, so we already know Q_0 is linear. While we could show directly that Q_m is linear (with an argument like that for R_θ), we prefer to do it another way that is instructive and derives the matrix of Q_m directly without using Theorem 2.6.2.

Let θ denote the angle between the positive x axis and the line $y = mx$. The key observation is that the transformation Q_m can be accomplished in three steps: First rotate through $-\theta$ (so our line coincides with the x axis), then reflect in the x axis, and finally rotate back through θ. In other words:

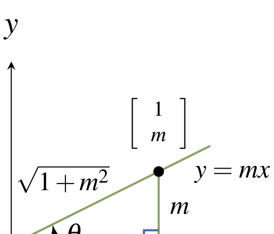
$$Q_m = R_\theta \circ Q_0 \circ R_{-\theta}$$

Since $R_{-\theta}$, Q_0, and R_θ are all linear, this (with Theorem 2.6.3) shows that Q_m is linear and that its matrix is the product of the matrices of R_θ, Q_0, and $R_{-\theta}$. If we write $c = \cos\theta$ and $s = \sin\theta$ for simplicity, then the matrices of R_θ, $R_{-\theta}$, and Q_0 are

$$\left[\begin{array}{rr} c & -s \\ s & c \end{array}\right], \quad \left[\begin{array}{rr} c & s \\ -s & c \end{array}\right], \quad \text{and} \quad \left[\begin{array}{rr} 1 & 0 \\ 0 & -1 \end{array}\right] \text{ respectively.}$$[13]

Hence, by Theorem 2.6.3, the matrix of $Q_m = R_\theta \circ Q_0 \circ R_{-\theta}$ is

$$\left[\begin{array}{rr} c & -s \\ s & c \end{array}\right]\left[\begin{array}{rr} 1 & 0 \\ 0 & -1 \end{array}\right]\left[\begin{array}{rr} c & s \\ -s & c \end{array}\right] = \left[\begin{array}{cc} c^2 - s^2 & 2sc \\ 2sc & s^2 - c^2 \end{array}\right]$$



We can obtain this matrix in terms of m alone. Figure 2.6.13 shows that

$$\cos\theta = \frac{1}{\sqrt{1+m^2}} \text{ and } \sin\theta = \frac{m}{\sqrt{1+m^2}}$$

so the matrix $\left[\begin{array}{cc} c^2 - s^2 & 2sc \\ 2sc & s^2 - c^2 \end{array}\right]$ of Q_m becomes $\frac{1}{1+m^2}\left[\begin{array}{cc} 1 - m^2 & 2m \\ 2m & m^2 - 1 \end{array}\right]$.

Theorem 2.6.5

Let Q_m denote reflection in the line $y = mx$. Then Q_m is a

[13] The matrix of $R_{-\theta}$ comes from the matrix of R_θ using the fact that, for all angles θ, $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin(\theta)$.

linear transformation with matrix $\frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}$.

Note that if $m = 0$, the matrix in Theorem 2.6.5 becomes $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, as expected. Of course this analysis fails for reflection in the y axis because vertical lines have no slope. However it is an easy exercise to verify directly that reflection in the y axis is indeed linear with matrix $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$.[14]

Example 2.6.8

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation through $-\frac{\pi}{2}$ followed by reflection in the y axis. Show that T is a reflection in a line through the origin and find the line.

Solution. The matrix of $R_{-\frac{\pi}{2}}$ is $\begin{bmatrix} \cos(-\frac{\pi}{2}) & -\sin(-\frac{\pi}{2}) \\ \sin(-\frac{\pi}{2}) & \cos(-\frac{\pi}{2}) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and the matrix of reflection in the y axis is $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. Hence the matrix of T is $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ and this is reflection in the line $y = -x$ (take $m = -1$ in Theorem 2.6.5).

Projections

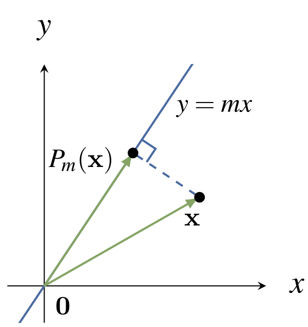


Figure 2.6.14

The method in the proof of Theorem 2.6.5 works more generally. Let $P_m : \mathbb{R}^2 \to \mathbb{R}^2$ denote projection on the line $y = mx$. This transformation is described geometrically in Figure 2.6.14.

If $m = 0$, then $P_0 \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}$ for all $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$, so P_0 is linear with matrix $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Hence the argument above for Q_m goes through for P_m. First observe that

$$P_m = R_\theta \circ P_0 \circ R_{-\theta}$$

as before. So, P_m is linear with matrix

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} c^2 & sc \\ sc & s^2 \end{bmatrix}$$

where $c = \cos\theta = \frac{1}{\sqrt{1+m^2}}$ and $s = \sin\theta = \frac{m}{\sqrt{1+m^2}}$.

[14] Note that $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \lim_{m\to\infty} \frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}$.

This gives:

Theorem 2.6.6

Let $P_m : \mathbb{R}^2 \to \mathbb{R}^2$ be projection on the line $y = mx$. Then P_m is a linear transformation with matrix $\frac{1}{1+m^2}\begin{bmatrix} 1 & m \\ m & m^2 \end{bmatrix}$.

Again, if $m = 0$, then the matrix in Theorem 2.6.6 reduces to $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ as expected. As the y axis has no slope, the analysis fails for projection on the y axis, but this transformation is indeed linear with matrix $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ as is easily verified directly.

Note that the formula for the matrix of Q_m in Theorem 2.6.5 can be derived from the above formula for the matrix of P_m. Using Figure 2.6.12, observe that $Q_m(\mathbf{x}) = \mathbf{x} + 2[P_m(\mathbf{x}) - \mathbf{x}]$ so $Q_m(x) = 2P_m(\mathbf{x}) - \mathbf{x}$. Substituting the matrices for $P_m(\mathbf{x})$ and $1_{\mathbb{R}^2}(\mathbf{x})$ gives the desired formula.

Example 2.6.9

Given $\mathbf{x}$ in $\mathbb{R}^2$, write $\mathbf{y} = P_m(\mathbf{x})$. The fact that $\mathbf{y}$ lies on the line $y = mx$ means that $P_m(\mathbf{y}) = \mathbf{y}$. But then

$$(P_m \circ P_m)(\mathbf{x}) = P_m(\mathbf{y}) = \mathbf{y} = P_m(\mathbf{x}) \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^2, \quad \text{that is, } P_m \circ P_m = P_m.$$

In particular, if we write the matrix of P_m as $A = \frac{1}{1+m^2}\begin{bmatrix} 1 & m \\ m & m^2 \end{bmatrix}$, then $A^2 = A$. The reader should verify this directly.

Exercises for 2.6

Exercise 2.6.1 Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation.

a. Find $T\begin{bmatrix} 8 \\ 3 \\ 7 \end{bmatrix}$ if $T\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $T\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$.

b. Find $T\begin{bmatrix} 5 \\ 6 \\ -13 \end{bmatrix}$ if $T\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ and $T\begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

---

b. $\begin{bmatrix} 5 \\ 6 \\ -13 \end{bmatrix} = 3\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} - 2\begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix}$, so

$$T\begin{bmatrix} 5 \\ 6 \\ -13 \end{bmatrix} = 3T\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} - 2T\begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix} = 3\begin{bmatrix} 3 \\ 5 \end{bmatrix} - 2\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 11 \\ 11 \end{bmatrix}$$

Exercise 2.6.2 Let $T:\mathbb{R}^4\to\mathbb{R}^3$ be a linear transformation.

a. Find $T\begin{bmatrix}1\\3\\-2\\-3\end{bmatrix}$ if $T\begin{bmatrix}1\\1\\0\\-1\end{bmatrix}=\begin{bmatrix}2\\3\\-1\end{bmatrix}$ and $T\begin{bmatrix}0\\-1\\1\\1\end{bmatrix}=\begin{bmatrix}5\\0\\1\end{bmatrix}$.

b. Find $T\begin{bmatrix}5\\-1\\2\\-4\end{bmatrix}$ if $T\begin{bmatrix}1\\1\\1\\1\end{bmatrix}=\begin{bmatrix}5\\1\\-3\end{bmatrix}$ and $T\begin{bmatrix}-1\\1\\0\\2\end{bmatrix}=\begin{bmatrix}2\\0\\1\end{bmatrix}$.

---

b. As in 1(b), $T\begin{bmatrix}5\\-1\\2\\-4\end{bmatrix}=\begin{bmatrix}4\\2\\-9\end{bmatrix}$.

Exercise 2.6.3 In each case assume that the transformation T is linear, and use Theorem 2.6.2 to obtain the matrix A of T.

a. $T:\mathbb{R}^2\to\mathbb{R}^2$ is reflection in the line $y=-x$.

b. $T:\mathbb{R}^2\to\mathbb{R}^2$ is given by $T(\mathbf{x})=-\mathbf{x}$ for each $\mathbf{x}$ in $\mathbb{R}^2$.

c. $T:\mathbb{R}^2\to\mathbb{R}^2$ is clockwise rotation through $\frac{\pi}{4}$.

d. $T:\mathbb{R}^2\to\mathbb{R}^2$ is counterclockwise rotation through $\frac{\pi}{4}$.

---

b. $T(\mathbf{e}_1)=-\mathbf{e}_2$ and $T(\mathbf{e}_2)=-\mathbf{e}_1$. So $A\begin{bmatrix}T(\mathbf{e}_1) & T(\mathbf{e}_2)\end{bmatrix}=\begin{bmatrix}-\mathbf{e}_2 & -\mathbf{e}_1\end{bmatrix}=\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$.

d. $T(\mathbf{e}_1)=\begin{bmatrix}\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\end{bmatrix}$ and $T(\mathbf{e}_2)=\begin{bmatrix}-\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\end{bmatrix}$ So $A=\begin{bmatrix}T(\mathbf{e}_1) & T(\mathbf{e}_2)\end{bmatrix}=\frac{\sqrt{2}}{2}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}$.

Exercise 2.6.4 In each case use Theorem 2.6.2 to obtain the matrix A of the transformation T. You may assume that T is linear in each case.

a. $T:\mathbb{R}^3\to\mathbb{R}^3$ is reflection in the $x-z$ plane.

b. $T:\mathbb{R}^3\to\mathbb{R}^3$ is reflection in the $y-z$ plane.

---

b. $T(\mathbf{e}_1)=-\mathbf{e}_1$, $T(\mathbf{e}_2)=\mathbf{e}_2$ and $T(\mathbf{e}_3)=\mathbf{e}_3$. Hence Theorem 2.6.2 gives $A\begin{bmatrix}T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3)\end{bmatrix}=\begin{bmatrix}-\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\end{bmatrix}=\begin{bmatrix}-1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$.

Exercise 2.6.5 Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be a linear transformation.

a. If $\mathbf{x}$ is in $\mathbb{R}^n$, we say that $\mathbf{x}$ is in the *kernel* of T if $T(\mathbf{x})=\mathbf{0}$. If $\mathbf{x}_1$ and $\mathbf{x}_2$ are both in the kernel of T, show that $a\mathbf{x}_1+b\mathbf{x}_2$ is also in the kernel of T for all scalars a and b.

b. If $\mathbf{y}$ is in $\mathbb{R}^n$, we say that $\mathbf{y}$ is in the *image* of T if $\mathbf{y}=T(\mathbf{x})$ for some $\mathbf{x}$ in $\mathbb{R}^n$. If $\mathbf{y}_1$ and $\mathbf{y}_2$ are both in the image of T, show that $a\mathbf{y}_1+b\mathbf{y}_2$ is also in the image of T for all scalars a and b.

---

b. We have $\mathbf{y}_1=T(\mathbf{x}_1)$ for some $\mathbf{x}_1$ in $\mathbb{R}^n$, and $\mathbf{y}_2=T(\mathbf{x}_2)$ for some $\mathbf{x}_2$ in $\mathbb{R}^n$. So $a\mathbf{y}_1+b\mathbf{y}_2=aT(\mathbf{x}_1)+bT(\mathbf{x}_2)=T(a\mathbf{x}_1+b\mathbf{x}_2)$. Hence $a\mathbf{y}_1+b\mathbf{y}_2$ is also in the image of T.

Exercise 2.6.6 Use Theorem 2.6.2 to find the matrix of the **identity transformation** $1_{\mathbb{R}^n}:\mathbb{R}^n\to\mathbb{R}^n$ defined by $1_{\mathbb{R}^n}(\mathbf{x})=\mathbf{x}$ for each $\mathbf{x}$ in $\mathbb{R}^n$.

Exercise 2.6.7 In each case show that $T:\mathbb{R}^2\to\mathbb{R}^2$ is not a linear transformation.

a) $T\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}xy\\0\end{bmatrix}$ b) $T\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\y^2\end{bmatrix}$

---

b. $T\left(2\begin{bmatrix}0\\1\end{bmatrix}\right) \neq 2\begin{bmatrix}0\\-1\end{bmatrix}$.

Exercise 2.6.8 In each case show that T is either reflection in a line or rotation through an angle, and find the line or angle.

a. $T\begin{bmatrix}x\\y\end{bmatrix} = \frac{1}{5}\begin{bmatrix}-3x+4y\\4x+3y\end{bmatrix}$

b. $T\begin{bmatrix}x\\y\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}x+y\\-x+y\end{bmatrix}$

c. $T\begin{bmatrix}x\\y\end{bmatrix} = \frac{1}{\sqrt{3}}\begin{bmatrix}x-\sqrt{3}y\\\sqrt{3}x+y\end{bmatrix}$

d. $T\begin{bmatrix}x\\y\end{bmatrix} = -\frac{1}{10}\begin{bmatrix}8x+6y\\6x-8y\end{bmatrix}$

---

b. $A = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\-1 & 1\end{bmatrix}$, rotation through $\theta = -\frac{\pi}{4}$.

d. $A = \frac{1}{10}\begin{bmatrix}-8 & -6\\-6 & 8\end{bmatrix}$, reflection in the line $y = -3x$.

Exercise 2.6.9 Express reflection in the line $y = -x$ as the composition of a rotation followed by reflection in the line $y = x$.

Exercise 2.6.10 Find the matrix of $T : \mathbb{R}^3 \to \mathbb{R}^3$ in each case:

a. T is rotation through θ about the x axis (from the y axis to the z axis).

b. T is rotation through θ about the y axis (from the x axis to the z axis).

---

b. $\begin{bmatrix}\cos\theta & 0 & -\sin\theta\\0 & 1 & 0\\\sin\theta & 0 & \cos\theta\end{bmatrix}$

Exercise 2.6.11 Let $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ denote reflection in the line making an angle θ with the positive x axis.

a. Show that the matrix of T_θ is $\begin{bmatrix}\cos 2\theta & \sin 2\theta\\\sin 2\theta & -\cos 2\theta\end{bmatrix}$ for all θ.

b. Show that $T_\theta \circ R_{2\phi} = T_{\theta-\phi}$ for all θ and ϕ.

Exercise 2.6.12 In each case find a rotation or reflection that equals the given transformation.

a. Reflection in the y axis followed by rotation through $\frac{\pi}{2}$.

b. Rotation through π followed by reflection in the x axis.

c. Rotation through $\frac{\pi}{2}$ followed by reflection in the line $y = x$.

d. Reflection in the x axis followed by rotation through $\frac{\pi}{2}$.

e. Reflection in the line $y = x$ followed by reflection in the x axis.

f. Reflection in the x axis followed by reflection in the line $y = x$.

---

b. Reflection in the y axis

d. Reflection in $y = x$

f. Rotation through $\frac{\pi}{2}$

Exercise 2.6.13 Let R and S be matrix transformations $\mathbb{R}^n \to \mathbb{R}^m$ induced by matrices A and B respectively. In each case, show that T is a matrix transformation and describe its matrix in terms of A and B.

a. $T(\mathbf{x}) = R(\mathbf{x}) + S(\mathbf{x})$ for all $\mathbf{x}$ in $\mathbb{R}^n$.

b. $T(\mathbf{x}) = aR(\mathbf{x})$ for all $\mathbf{x}$ in $\mathbb{R}^n$ (where a is a fixed real number).

---

b. $T(\mathbf{x}) = aR(\mathbf{x}) = a(A\mathbf{x}) = (aA)\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}$. Hence T is induced by aA.

Exercise 2.6.14 Show that the following hold for all linear transformations $T : \mathbb{R}^n \to \mathbb{R}^m$:

a) $T(\mathbf{0}) = \mathbf{0}$

b) $T(-\mathbf{x}) = -T(\mathbf{x})$ for all $\mathbf{x}$ in $\mathbb{R}^n$

---

b. If $\mathbf{x}$ is in $\mathbb{R}^n$, then $T(-\mathbf{x}) = T[(-1)\mathbf{x}] = (-1)T(\mathbf{x}) = -T(\mathbf{x})$.

Exercise 2.6.15 The transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x}$ in $\mathbb{R}^n$ is called the **zero transformation**.

a. Show that the zero transformation is linear and find its matrix.

b. Let $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ denote the columns of the $n \times n$ identity matrix. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear and $T(\mathbf{e}_i) = \mathbf{0}$ for each i, show that T is the zero transformation. [*Hint*: Theorem 2.6.1.]

Exercise 2.6.16 Write the elements of $\mathbb{R}^n$ and $\mathbb{R}^m$ as rows. If A is an $m \times n$ matrix, define $T : \mathbb{R}^m \to \mathbb{R}^n$ by $T(\mathbf{y}) = \mathbf{y}A$ for all rows $\mathbf{y}$ in $\mathbb{R}^m$. Show that:

a. T is a linear transformation.

b. the rows of A are $T(\mathbf{f}_1), T(\mathbf{f}_2), \dots, T(\mathbf{f}_m)$ where $\mathbf{f}_i$ denotes row i of I_m. [*Hint*: Show that $\mathbf{f}_iA$ is row i of A.]

Exercise 2.6.17 Let $S : \mathbb{R}^n \to \mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear transformations with matrices A and B respectively.

a. Show that $B^2 = B$ if and only if $T^2 = T$ (where T^2 means $T \circ T$).

b. Show that $B^2 = I$ if and only if $T^2 = 1_{\mathbb{R}^n}$.

c. Show that $AB = BA$ if and only if $S \circ T = T \circ S$.

[*Hint*: Theorem 2.6.3.]

---

b. If $B^2 = I$ then $T^2(\mathbf{x}) = T[T(\mathbf{x})] = B(B\mathbf{x}) = B^2\mathbf{x} = I\mathbf{x} = \mathbf{x} = 1_{\mathbb{R}^2}(\mathbf{x})$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Hence $T^2 = 1_{\mathbb{R}^2}$. If $T^2 = 1_{\mathbb{R}^2}$, then $B^2\mathbf{x} = T^2(\mathbf{x}) = 1_{\mathbb{R}^2}(\mathbf{x}) = \mathbf{x} = I\mathbf{x}$ for all $\mathbf{x}$, so $B^2 = I$ by Theorem 2.2.6.

Exercise 2.6.18 Let $Q_0 : \mathbb{R}^2 \to \mathbb{R}^2$ be reflection in the x axis, let $Q_1 : \mathbb{R}^2 \to \mathbb{R}^2$ be reflection in the line $y = x$, let $Q_{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ be reflection in the line $y = -x$, and let $R_{\frac{\pi}{2}} : \mathbb{R}^2 \to \mathbb{R}^2$ be counterclockwise rotation through $\frac{\pi}{2}$.

a. Show that $Q_1 \circ R_{\frac{\pi}{2}} = Q_0$.

b. Show that $Q_1 \circ Q_0 = R_{\frac{\pi}{2}}$.

c. Show that $R_{\frac{\pi}{2}} \circ Q_0 = Q_1$.

d. Show that $Q_0 \circ R_{\frac{\pi}{2}} = Q_{-1}$.

---

b. The matrix of $Q_1 \circ Q_0$ is $\left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]\left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right] = \left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right]$, which is the matrix of $R_{\frac{\pi}{2}}$.

d. The matrix of $Q_0 \circ R_{\frac{\pi}{2}}$ is $\left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right]\left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right] = \left[\begin{array}{cc} 0 & -1 \\ -1 & 0 \end{array}\right]$, which is the matrix of Q_{-1}.

Exercise 2.6.19 For any slope m, show that:

a) $Q_m \circ P_m = P_m$

b) $P_m \circ Q_m = P_m$

Exercise 2.6.20 Define $T : \mathbb{R}^n \to \mathbb{R}$ by $T(x_1, x_2, \dots, x_n) = x_1 + x_2 + \cdots + x_n$. Show that T is a linear transformation and find its matrix.

We have $T(\mathbf{x}) = x_1 + x_2 + \cdots + x_n = \left[\begin{array}{cccc} 1 & 1 & \cdots & 1 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right]$, so T is the matrix transformation induced by the matrix $A = \left[\begin{array}{cccc} 1 & 1 & \cdots & 1 \end{array}\right]$. In particular, T is linear. On the other hand, we can use Theorem 2.6.2 to get A, but to do this we must first show directly that T is linear. If we write $\mathbf{x} = \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right]$ and $\mathbf{y} = \left[\begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_n \end{array}\right]$.

Then

$$\begin{aligned}T(\mathbf{x}+\mathbf{y}) &= T\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} \\ &= (x_1+y_1)+(x_2+y_2)+\cdots+(x_n+y_n) \\ &= (x_1+x_2+\cdots+x_n)+(y_1+y_2+\cdots+y_n) \\ &= T(\mathbf{x})+T(\mathbf{y})\end{aligned}$$

Similarly, $T(a\mathbf{x}) = aT(\mathbf{x})$ for any scalar a, so T is linear. By Theorem 2.6.2, T has matrix $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$, as before.

Exercise 2.6.21 Given c in $\mathbb{R}$, define $T_c : \mathbb{R}^n \to \mathbb{R}$ by $T_c(\mathbf{x}) = c\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Show that T_c is a linear transformation and find its matrix.

Exercise 2.6.22 Given vectors $\mathbf{w}$ and $\mathbf{x}$ in $\mathbb{R}^n$, denote their dot product by $\mathbf{w}\cdot\mathbf{x}$.

a. Given $\mathbf{w}$ in $\mathbb{R}^n$, define $T_\mathbf{w} : \mathbb{R}^n \to \mathbb{R}$ by $T_\mathbf{w}(\mathbf{x}) = \mathbf{w}\cdot\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Show that $T_\mathbf{w}$ is a linear transformation.

b. Show that *every* linear transformation $T : \mathbb{R}^n \to \mathbb{R}$ is given as in (a); that is $T = T_\mathbf{w}$ for some $\mathbf{w}$ in $\mathbb{R}^n$.

---

b. If $T : \mathbb{R}^n \to \mathbb{R}$ is linear, write $T(\mathbf{e}_j) = w_j$ for each $j = 1, 2, \dots, n$ where $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$. Since $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$, Theorem 2.6.1 gives

$$\begin{aligned}T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) \\ &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n) \\ &= x_1w_1 + x_2w_2 + \cdots + x_nw_n \\ &= \mathbf{w}\cdot\mathbf{x} = T_\mathbf{w}(\mathbf{x})\end{aligned}$$

where $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$. Since this holds for all $\mathbf{x}$ in $\mathbb{R}^n$, it shows that $T = T_\mathbf{W}$. This also follows from Theorem 2.6.2, but we have first to verify that T is linear. (This comes to showing that $\mathbf{w}\cdot(\mathbf{x}+\mathbf{y}) = \mathbf{w}\cdot\mathbf{s} + \mathbf{w}\cdot\mathbf{y}$ and $\mathbf{w}\cdot(a\mathbf{x}) = a(\mathbf{w}\cdot\mathbf{x})$ for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ and all a in $\mathbb{R}$.) Then T has matrix $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{bmatrix} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}$ by Theorem 2.6.2. Hence if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ in $\mathbb{R}$, then $T(\mathbf{x}) = A\mathbf{x} = \mathbf{w}\cdot\mathbf{x}$, as required.

Exercise 2.6.23 If $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$, show that there is a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ such that $T(\mathbf{x}) = \mathbf{y}$. [*Hint*: By Definition 2.5, find a matrix A such that $A\mathbf{x} = \mathbf{y}$.]

---

b. Given $\mathbf{x}$ in $\mathbb{R}$ and a in $\mathbb{R}$, we have

$$\begin{array}{rcll} (S\circ T)(a\mathbf{x}) &=& S[T(a\mathbf{x})] & \text{Definition of } S\circ T \\ &=& S[aT(\mathbf{x})] & \text{Because } T \text{ is linear.} \\ &=& a[S[T(\mathbf{x})]] & \text{Because } S \text{ is linear.} \\ &=& a[S\circ T(\mathbf{x})] & \text{Definition of } S\circ T \end{array}$$

Exercise 2.6.24 Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m \xrightarrow{S} \mathbb{R}^k$ be two linear transformations. Show directly that $S\circ T$ is linear. That is:

a. Show that $(S\circ T)(\mathbf{x}+\mathbf{y}) = (S\circ T)\mathbf{x} + (S\circ T)\mathbf{y}$ for all $\mathbf{x}$, $\mathbf{y}$ in $\mathbb{R}^n$.

b. Show that $(S\circ T)(a\mathbf{x}) = a[(S\circ T)\mathbf{x}]$ for all $\mathbf{x}$ in $\mathbb{R}^n$ and all a in $\mathbb{R}$.

Exercise 2.6.25 Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m \xrightarrow{S} \mathbb{R}^k \xrightarrow{R} \mathbb{R}^k$ be linear. Show that $R\circ(S\circ T) = (R\circ S)\circ T$ by showing directly that $[R\circ(S\circ T)](\mathbf{x}) = [(R\circ S)\circ T)](\mathbf{x})$ holds for each vector $\mathbf{x}$ in $\mathbb{R}^n$.