5 Linear Transformations

Outcome:

5. Understand linear transformations, their compositions, and their application to homogeneous coordinates. Understand representations of vectors with respect to different bases. Understand eigenvalues and eigenspaces, diagonalization.

   Performance Criteria:

   (a) Evaluate a transformation.

   (b) Determine the formula for a transformation in $\mathbb{R}^2$ or $\mathbb{R}^3$ that has been described geometrically.

   (c) Determine whether a given transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$ is linear. If it isn't, give a counterexample; if it is, prove that it is.

   (d) Given the action of a transformation on each vector in a basis for a space, determine the action on an arbitrary vector in the space.

   (e) Give the matrix representation of a linear transformation.

   (f) Find the composition of two transformations.

   (g) Find matrices that perform combinations of dilations, reflections, rotations and translations in $\mathbb{R}^2$ using homogenous coordinates.

   (h) Determine whether a given vector is an eigenvector for a matrix; if it is, give the corresponding eigenvalue.

   (i) Determine eigenvectors and corresponding eigenvalues for linear transformations in $\mathbb{R}^2$ or $\mathbb{R}^3$ that are described geometrically.

   (j) Find the characteristic polynomial for a 2×2 or 3×3 matrix. Use it to find the eigenvalues of the matrix.

   (k) Give the eigenspace E_j corresponding to an eigenvalue λ_j of a matrix.

   (l) Determine the principal stresses and the orientation of the principal axes for a two-dimensional stress element.

   (m) Diagonalize a matrix; know the forms of the matrices P and D from $P^{-1}AP = D$.

   (n) Write a system of linear differential equations in matrix-vector form. Write the initial conditions in vector form.

   (o) Solve a system of two linear differential equations; solve an initial value problem for a system of two linear differential equations.

5.1 Transformations of Vectors

Performance Criteria:

5. (a) Evaluate a transformation.
   (b) Determine the formula for a transformation in $\mathbb{R}^2$ or $\mathbb{R}^3$ that has been described geometrically.

Back in a "regular" algebra class you might have considered a function like $f(x) = \sqrt{x+5}$, and you may have discussed the fact that this function is only valid for certain values of x. When considering functions more carefully, we usually "declare" the function before defining it:

$$\text{Let } f : [-5, \infty) \to \mathbb{R} \text{ be defined by } f(x) = \sqrt{x+5}$$

Here the set $[-5, \infty)$ of allowable "inputs" is called the **domain** of the function, and the set $\mathbb{R}$ is sometimes called the **codomain** or **target set**. Those of you with programming experience will recognize the process of first declaring the function, then defining it. Later you might "call" the function, which in math we refer to as "evaluating" it.

In a similar manner we can define functions from one vector space to another, like

$$\text{Define } T : \mathbb{R}^2 \to \mathbb{R}^3 \text{ by } T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_2 \\ x_1^2 \end{bmatrix}$$

We will call such a function a **transformation**, hence the use of the letter T. (When we have a second transformation, we'll usually call it S.) The word "transformation" implies that one vector is transformed into another vector. It should be clear how a transformation works:

◇ **Example 5.1(a):** Find $T\left(\begin{bmatrix} -3 \\ 5 \end{bmatrix}\right)$ for the transformation defined above.

$$T\left(\begin{bmatrix} -3 \\ 5 \end{bmatrix}\right) = \begin{bmatrix} -3+5 \\ 5 \\ (-3)^2 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 9 \end{bmatrix}$$

It gets a bit tiresome to write both parentheses and brackets, so from now on we will dispense with the parentheses and just write

$$T\begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 9 \end{bmatrix}$$

At this point we should note that you have encountered other kinds of transformations. For example, taking the derivative of a function results in another function,

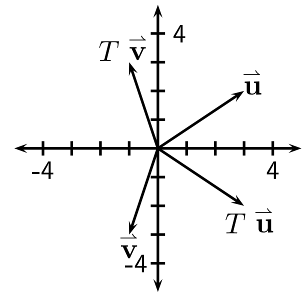
$$\frac{d}{dx}(x^3 - 5x^2 + 2x - 1) = 3x^2 - 10x + 2,$$

so the action of taking a derivative can be thought of as a transformation. Such transformations are often called **operators**.

Sometimes we will wish to determine the formula for a transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ or $\mathbb{R}^3$ to $\mathbb{R}^3$ that has been described geometrically.

◇ **Example 5.1(b):** Determine the formula for the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ that reflects vectors across the x-axis.

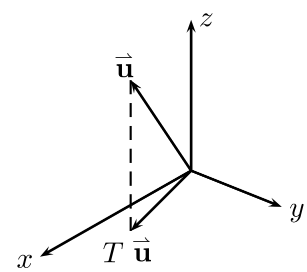
Solution: First we might wish to draw a picture to see what such a transformation does to a vector. To the right we see the vectors $\vec{\mathbf{u}} = [3,2]$ and $\vec{\mathbf{v}} = [-1,-3]$, and their transformations $T\,\vec{\mathbf{u}} = [3,-2]$ and $T\,\vec{\mathbf{v}} = [-1,3]$. From these we see that what the transformation does is change the sign of the second component of a vector. Therefore



$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}$$

◇ **Example 5.1(c):** Determine the formula for the transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ that projects vectors onto the xy-plane.

Solution: It is a little more difficult to draw a picture for this one, but to the right you can see an attempt to illustrate the action of this transformation on a vector $\vec{\mathbf{u}}$. Note that in projecting a vector onto the xy-plane, the x- and y-coordinates stay the same, but the z-coordinate becomes zero. The formula for the transformation is then



$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

Let's now look at the above example in a different way. Note that the xy-plane is a 2-dimensional subspace of $\mathbb{R}^3$ that corresponds (exactly!) with $\mathbb{R}^2$. We can therefore look at the transformation as $T : \mathbb{R}^3 \to \mathbb{R}^2$ that assigns to every point in $\mathbb{R}^3$ its projection onto the xy-plane taken as $\mathbb{R}^2$. The formula for this transformation is then

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

We conclude this section with a very important observation. Consider the matrix

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

and define $T_A\,\vec{\mathbf{x}} = A\,\vec{\mathbf{x}}$ for every vector for which $A\,\vec{\mathbf{x}}$ is defined. This transformation acts on vectors in $\mathbb{R}^2$ and "returns" vectors in $\mathbb{R}^3$. That is, $T_A : \mathbb{R}^2 \to \mathbb{R}^3$. In general, we can use any $m \times n$ matrix A to define a transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ in this manner. In the next section we will see that such transformations have a desirable characteristic, and that every transformation with that characteristic can be represented by multiplication by a matrix.

◇ **Example 5.1(d):** Find $T_A \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, where T_A is defined as above, for the matrix given.

Solution: $T_A \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 & 1 \\ 0 & -3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} -14 \\ -3 \\ 5 \end{bmatrix}$

Section 5.1 Exercises

To Solutions

1. For each of the following a transformation T is declared and defined, and one or more vectors $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ is(are) given. Find the transformation(s) of the vector(s), labelling your answer(s) correctly.

   (a) $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\quad T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 x_2 \\ x_2^2 \end{bmatrix}$, $\quad \vec{\mathbf{u}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

   (b) $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $\quad T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ 3x_2 \\ -x_1 \end{bmatrix}$, $\quad \vec{\mathbf{u}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

   (c) $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $\quad T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_3 + x_1 \end{bmatrix}$, $\quad \vec{\mathbf{v}} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, $\vec{\mathbf{w}} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$

   (d) $T: \mathbb{R}^4 \rightarrow \mathbb{R}^4$, $\quad T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_4 \\ x_1 \\ x_2 \end{bmatrix}$, $\quad \vec{\mathbf{u}} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$

   (e) $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $\quad T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 5 \\ x_2 - 2 \\ x_3 + 1 \end{bmatrix}$, $\quad \vec{\mathbf{u}} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, $\vec{\mathbf{w}} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$

   (f) $T: \mathbb{R}^5 \rightarrow \mathbb{R}^5$, $\quad T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_2 \\ x_4 - x_3 \\ x_5 - x_4 \\ -x_5 \end{bmatrix}$, $\quad \vec{\mathbf{v}} = \begin{bmatrix} 2 \\ -7 \\ 5 \\ 4 \\ -1 \end{bmatrix}$

2. For each of the transformations in Exercise 1, determine whether there is a matrix A for which $T = T_A$, as described in the Example 5.1(d) and the discussion preceeding it.

3. For each of the following, give the transformation T that acts on points/vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ in the manner described. Be sure to include both

   - a "declaration statement" of the form "Define $T : \mathbb{R}^m \to \mathbb{R}^n$ by" and
   - a mathematical formula for the transformation.

   To do this you might find it useful to list a few specific points or vectors and the points or vectors they transform to. *Points on the axes are often useful for this due to the simplicity of working with them.*

   (a) The transformation that reflects every vector in $\mathbb{R}^2$ across the x-axis.

   (b) The transformation that rotates every vector in $\mathbb{R}^2$ 90 degrees clockwise.

   (c) The transformation that translates every point in $\mathbb{R}^2$ three points to the right and one point up. *We will see soon that this is a very important and interesting kind of transformation.*

   (d) The transformation that reflects every vector in $\mathbb{R}^2$ across the line $y = -x$.

   (e) The transformation that projects every vector in $\mathbb{R}^2$ onto the x-axis.

   (f) The transformation that reflects every point in $\mathbb{R}^3$ across the xz-plane.

   (g) The transformation that rotates every point in $\mathbb{R}^3$ counterclockwise 90 degrees, as looking down the positive z-axis, around the z-axis.

   (h) The transformation that rotates every point in $\mathbb{R}^3$ counterclockwise 90 degrees, as looking down the positive y-axis, around the y-axis.

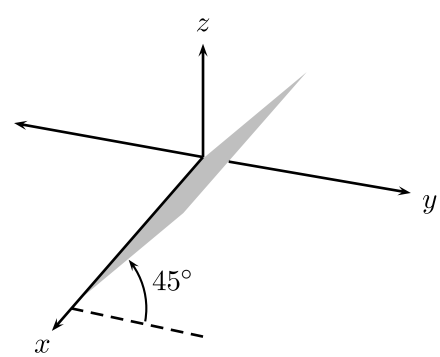
   (i) The transformation that projects every point in $\mathbb{R}^3$ across the xz-plane.

   (j) The transformation that projects every point in $\mathbb{R}^3$ onto the y-axis.

   (k) The transformation that takes every point in $\mathbb{R}^2$ and puts it at the corresponding point in $\mathbb{R}^3$ on the plane $z = 2$.

   (l) The transformation that translates every point in $\mathbb{R}^3$ upward by four units and in the negative y-direction by one unit.

4. The picture to the right shows a plane containing the x-axis and at a 45 degree angle to the xy-plane. Consider a transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ that is performed as follows: Each point in $\mathbb{R}^2$ is transformed to the point in $\mathbb{R}^3$ that is on the 45 degree plane directly above (or below) its location in the xy-plane. Declare the transformation and give its formula. **Hint:** sketch a picture of just the yz-plane.



5. Declare and define a transformation T that reflects every point in $\mathbb{R}^3$ across the plane shown in Exercise 4. If possible, give a matrix A for which $T = T_A$.

6. The transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + ax_2 \\ x_2 \end{bmatrix}$ for any constant a is a type of transformation called a **shear**. Such transformations will become quite important to us soon. Let's let $a = 1$, so the transformation becomes $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 \end{bmatrix}$.

   (a) Describe what the transformation does geometrically to every point on the horizontal line with y-coordinate one.

   (b) Describe what the transformation does geometrically to every point on the horizontal line with y-coordinate two.

   (c) Describe what the transformation does geometrically to every point on the horizontal line with y-coordinate negative one.

   (d) Describe what the transformation does geometrically to every point on the horizontal line with y-coordinate zero.

   (e) What does the transformation do to every point with positive y-coordinate. Be as specific as you can.

   (f) What does the transformation do to every point with negative y-coordinate. Be as specific as you can.

   (g) Give a matrix A for which the transformation $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + ax_2 \\ x_2 \end{bmatrix}$ is T_A.

7. Now consider the transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + ax_3 \\ x_2 + bx_3 \\ x_3 \end{bmatrix}$ for any constants a and b. This is a shear in $\mathbb{R}^3$.

   (a) Suppose that $a = 2$ and $b = -1$. Describe what T does to all points in the plane $z = 1$ in that case.

   (b) Still assuming that $a = 2$ and $b = -1$, give a matrix A for which $T = T_A$ *for just the points in the plane* $z = 1$.

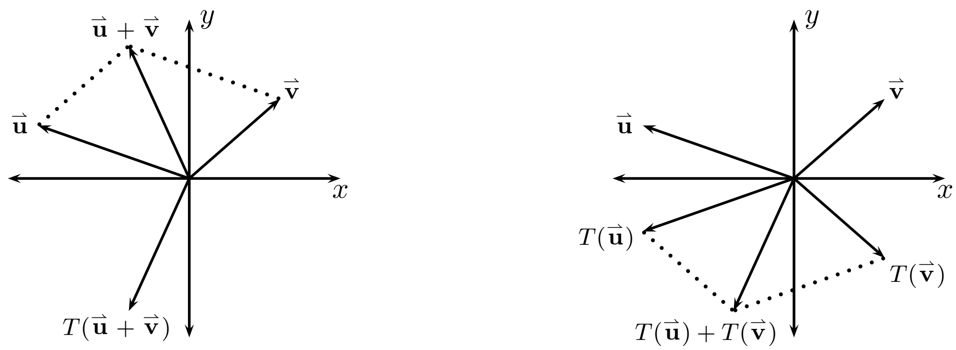
5.2 Linear Transformations

Performance Criteria:

5. (c) Determine whether a given transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$ is linear. If it isn't, give a counterexample; if it is, prove that it is.

   (d) Given the action of a transformation on each vector in a basis for a space, determine the action on an arbitrary vector in the space.

To begin this section, recall the transformation from Example 5.1(b) that reflects vectors in $\mathbb{R}^2$ across the x-axis. In the drawing below and to the left we see two vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ that are added, and then the vector $\vec{\mathbf{u}}+\vec{\mathbf{v}}$ is reflected across the x-axis. In the drawing below and to the right the same vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are reflected across the x-axis *first*, then the resulting vectors $T(\vec{\mathbf{u}})$ and $T(\vec{\mathbf{v}})$ are added.



Note that $T(\vec{\mathbf{u}}+\vec{\mathbf{v}}) = T(\vec{\mathbf{u}})+T(\vec{\mathbf{v}})$. Not all transformations have this property, but those that do have it, along with an additional property, are very important:

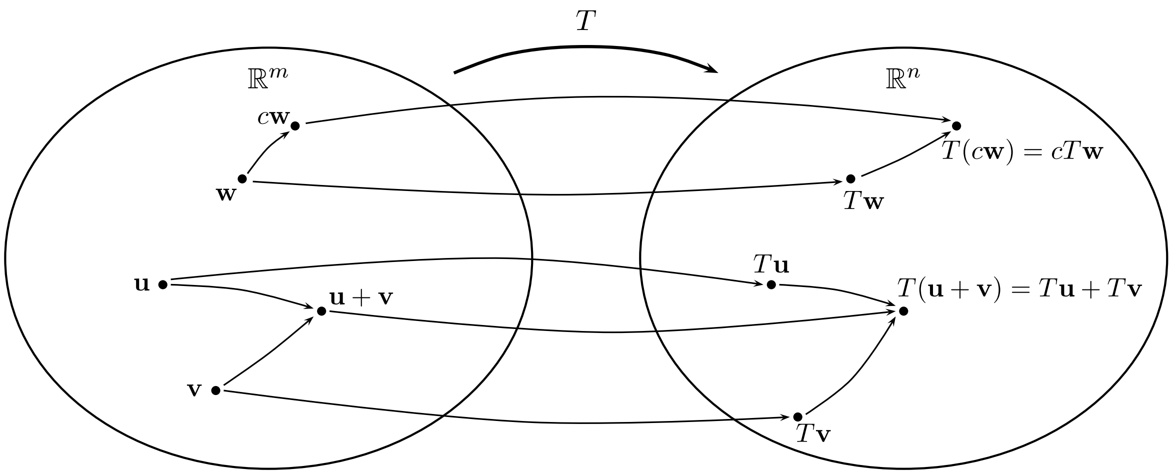
<u>DEFINITION 5.2.1</u>: **Linear Transformation**

A transformation $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is called a **linear transformation** if, for every scalar c and every pair of vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ in $\mathbb{R}^m$

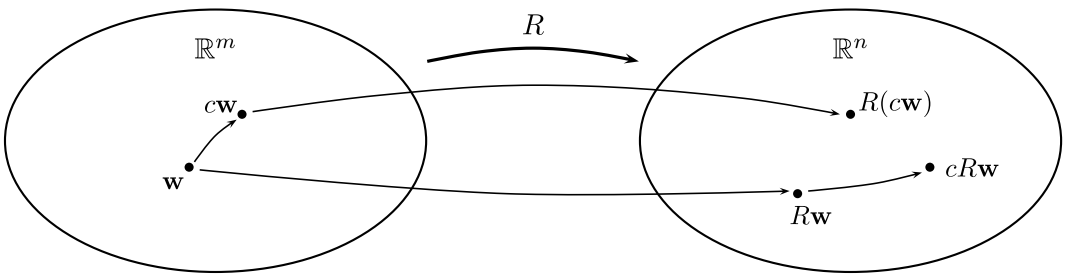
1) $T(\vec{\mathbf{u}}+\vec{\mathbf{v}}) = T(\vec{\mathbf{u}})+T(\vec{\mathbf{v}})$ (**additivity**) and

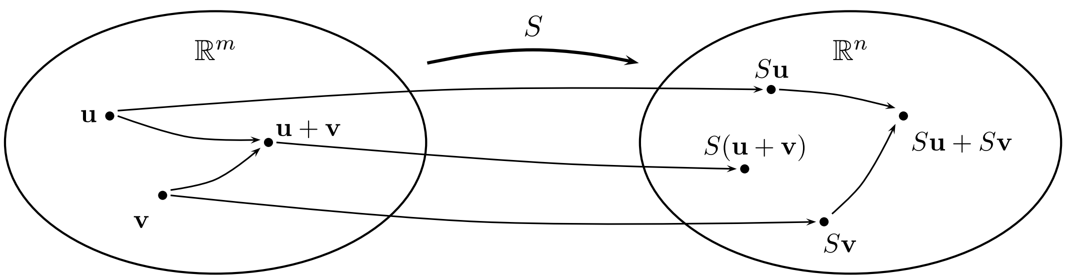
2) $T(c\,\vec{\mathbf{u}}) = c\,T(\vec{\mathbf{u}})$ (**homogeneity**).

Note that the above statement describes how a transformation T interacts with the two operations of vectors, addition and scalar multiplication. It tells us that if we take two vectors in the domain and add them in the domain, then transform the result, we will get the same thing as if we transform the vectors individually first, then add the results in the codomain. We will also get the same thing if we multiply a vector by a scalar and then transform as we will if we transform first, *then* multiply by the scalar.This is illustrated in the *mapping diagram* at the top of the next page.



The following two mapping diagrams are for transformations R and S that *ARE NOT* linear:





◇ **Example 5.2(a):** Let A be an $m \times n$ matrix. Is $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T_A\,\vec{\mathbf{x}} = A\,\vec{\mathbf{x}}$ a linear transformation?

Solution: We know from properties of multiplying a vector by a matrix that

$$T_A(\vec{\mathbf{u}}+\vec{\mathbf{v}}) = A(\vec{\mathbf{u}}+\vec{\mathbf{v}}) = A\,\vec{\mathbf{u}} + A\,\vec{\mathbf{v}} = T_A\,\vec{\mathbf{u}} + T_A\,\vec{\mathbf{v}}, \qquad T_A(c\,\vec{\mathbf{u}}) = A(c\,\vec{\mathbf{u}}) = cA\,\vec{\mathbf{u}} = cT_A\,\vec{\mathbf{u}}\,.$$

Therefore T_A is a linear transformation.

$\diamond$ **Example 5.2(b):** Is $T:\mathbb{R}^2 \to \mathbb{R}^3$ defined by $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1+x_2 \\ x_2 \\ x_1^2 \end{bmatrix}$ a linear transformation? If so, show that it is; if not, give a counterexample demonstrating that.

Solution: A good way to begin such an exercise is to try the two properties of a linear transformation for some specific vectors and scalars. If either condition is not met, then we have our counterexample, and if both hold we need to show they hold in general. Usually it is a bit simpler to check the condition $T(c\vec{\mathbf{u}}) = cT\,\vec{\mathbf{u}}$. In our case, if $c=2$ and $\vec{\mathbf{u}} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$,

$$T\left(2\begin{bmatrix} 3 \\ 4 \end{bmatrix}\right) = T\begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 14 \\ 8 \\ 36 \end{bmatrix} \quad \text{and} \quad 2T\begin{bmatrix} 3 \\ 4 \end{bmatrix} = 2\begin{bmatrix} 7 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 14 \\ 8 \\ 18 \end{bmatrix}$$

Because $T(c\,\vec{\mathbf{u}}) \neq cT\,\vec{\mathbf{u}}$ for our choices of c and u, T is not a linear transformation.

The next example shows the process required to show in general that a transformation is linear.

$\diamond$ **Example 5.2(c):** Determine whether $T:\mathbb{R}^3 \to \mathbb{R}^2$ defined by $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1+x_2 \\ x_2-x_3 \end{bmatrix}$ is linear. If it is, prove it in general; if it isn't, give a specific counterexample.

Solution: First let's check condition (1) of a linear transformation with the two specific vectors $\vec{\mathbf{u}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\vec{\mathbf{v}} = \begin{bmatrix} 4 \\ -5 \\ 6 \end{bmatrix}$. (I threw the negative in there just in case something funny happens when everything is positive.) Then

$$T\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ -5 \\ 6 \end{bmatrix}\right) = T\begin{bmatrix} 5 \\ -3 \\ 9 \end{bmatrix} = \begin{bmatrix} 2 \\ -12 \end{bmatrix}$$

and

$$T\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + T\begin{bmatrix} 4 \\ -5 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ -11 \end{bmatrix} = \begin{bmatrix} 2 \\ -12 \end{bmatrix}$$

so the first condition of linearity *appears* to hold. Let's prove it in general. Let $\vec{\mathbf{u}} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $\vec{\mathbf{v}} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ be arbitrary (that is, randomly selected) vectors in $\mathbb{R}^3$. Then

$$T(\vec{\mathbf{u}}+\vec{\mathbf{v}}) = T\left(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}\right) = T\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ u_3+v_3 \end{bmatrix} = \begin{bmatrix} u_1+v_1+u_2+v_2 \\ (u_2+v_2)-(u_3+v_3) \end{bmatrix} =$$

$$\begin{bmatrix} u_1+u_2+v_1+v_2 \\ (u_2-u_3)+(v_2-v_3) \end{bmatrix} = \begin{bmatrix} u_1+u_2 \\ u_2-u_3 \end{bmatrix} + \begin{bmatrix} v_1+v_2 \\ v_2-v_3 \end{bmatrix} = T\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + T\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = T(\vec{\mathbf{u}})+T(\vec{\mathbf{v}})$$

This shows that the first condition of linearity holds in general. Let $\vec{\mathbf{u}}$ again be arbitrary, along with the scalar c. Then

$$T(c\,\vec{\mathbf{u}}) = T\left(c\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}\right) = T\begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix} = \begin{bmatrix} cu_1+cu_2 \\ cu_2-cu_3 \end{bmatrix} =$$

$$\begin{bmatrix} c(u_1+u_2) \\ c(u_2-u_3) \end{bmatrix} = c\begin{bmatrix} u_1+u_2 \\ u_2-u_3 \end{bmatrix} = cT\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = cT(\vec{\mathbf{u}})$$

so the second condition holds as well, and T is a linear transformation.

There is a handy fact associated with linear transformations:

> THEOREM 5.2.2: If T is a linear transformation, then $T(\vec{\mathbf{0}}) = \vec{\mathbf{0}}$.

Note that this does not say that if $T(\vec{\mathbf{0}}) = \vec{\mathbf{0}}$, then T is a linear transformation, as you will see below. However, the contrapositive of the above statement tells us that *if $T(\vec{\mathbf{0}}) \neq \vec{\mathbf{0}}$, then T is not a linear transformation.*

When working with coordinate systems, one operation we often need to carry out is a **translation**, which means a shift of all points the same distance and direction. The transformation in the following example is a translation in $\mathbb{R}^2$.

◇ **Example 5.2(d):** Let a and b be any real numbers, with not both of them zero. Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1+a \\ x_2+b \end{bmatrix}$. Is T a linear transformation?

Solution: Because $T\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ (since not both a and b are zero), T is not a linear transformation.

We will find that the result of this example is quite unfortunate, because translations are very important in applications and the fact that they are not linear could potentially make them hard to work with. Fortunately there is a clever way around this problem - you'll see that in Section 5.5.

◇ **Example 5.2(e):** Determine whether $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1+x_2 \\ x_1x_2 \end{bmatrix}$ is linear. If it is, prove it in general; if it isn't, give a specific counterexample.

Solution: It is easy to see that $T(\vec{\mathbf{0}})=\vec{\mathbf{0}}$, so we can't immediately rule out T being linear, as we did in the last example. Let's do a quick check of the first condition of the definition of a linear transformation with an example. Let $\vec{\mathbf{u}}=\begin{bmatrix} -3 \\ 2 \end{bmatrix}$ and $\vec{\mathbf{v}}=\begin{bmatrix} 1 \\ 4 \end{bmatrix}$. Then

$$T(\vec{\mathbf{u}}+\vec{\mathbf{v}}) = T\left(\begin{bmatrix} -3 \\ 2 \end{bmatrix}+\begin{bmatrix} 1 \\ 4 \end{bmatrix}\right) = T\begin{bmatrix} -2 \\ 6 \end{bmatrix} = \begin{bmatrix} 4 \\ -12 \end{bmatrix}$$

and

$$T\vec{\mathbf{u}}+T\vec{\mathbf{v}} = T\begin{bmatrix} -3 \\ 2 \end{bmatrix}+T\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ -6 \end{bmatrix}+\begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix}$$

Clearly $T(\vec{\mathbf{u}}+\vec{\mathbf{v}}) \neq T\vec{\mathbf{u}}+T\vec{\mathbf{v}}$, so T is not a linear transformation.

We can "mix" the additivity and homogeneity of the definition of a linear transformation to arrive at the following:

> <u>Theorem 5.2.3</u>: If $T:\mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation if and only if
>
> $$T(c_1\vec{\mathbf{v}}_1+c_2\vec{\mathbf{v}}_2+\cdots+c_k\vec{\mathbf{v}}_k) = c_1T(\vec{\mathbf{v}}_1)+c_2T(\vec{\mathbf{v}}_2)+\cdots+c_kT(\vec{\mathbf{v}}_k)$$
>
> for all $\vec{\mathbf{v}}_1,\vec{\mathbf{v}}_2,...,\vec{\mathbf{v}}_k$ in $\mathbb{R}^m$ and all scalars $c_1,c_2,...,c_k$.

This is deceptively powerful result. Suppose, in particular, that the vectors $\vec{\mathbf{v}}_1,\vec{\mathbf{v}}_2,...,\vec{\mathbf{v}}_k$ in the above theorem constitute a basis for $\mathbb{R}^m$. Then every vector in $\mathbb{R}^m$ can be written as a unique linear combination of those vectors. If we have a linear transformation T and we know what it does to each of $\vec{\mathbf{v}}_1,\vec{\mathbf{v}}_2,...,\vec{\mathbf{v}}_k$, then the above theorem says that we then know what T does to every vector in $\mathbb{R}^m$.

◇ **Example 5.2(f):** The set $\mathcal{S}=\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$ is a basis for $\mathbb{R}^3$. Suppose that $T:\mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation such that

$$T\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \qquad T\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \qquad T\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}.$$

Find $T\begin{bmatrix} 2 \\ -7 \\ 4 \end{bmatrix}$.

Solution: We can see that

$$\begin{bmatrix} 2 \\ -7 \\ 4 \end{bmatrix} = 9\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 11\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 4\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Therefore, by Theorem 2.1.3,

$$\begin{aligned}
T\begin{bmatrix} 2 \\ -7 \\ 4 \end{bmatrix} &= T\left(9\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 11\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 4\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) \\
&= 9T\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 11T\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 4T\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\
&= 9\begin{bmatrix} -1 \\ 4 \end{bmatrix} - 11\begin{bmatrix} 5 \\ 0 \end{bmatrix} + 4\begin{bmatrix} 3 \\ -2 \end{bmatrix} \\
&= \begin{bmatrix} -9 \\ 36 \end{bmatrix} + \begin{bmatrix} -55 \\ 0 \end{bmatrix} + \begin{bmatrix} 12 \\ -8 \end{bmatrix} \\
&= \begin{bmatrix} -52 \\ 28 \end{bmatrix}
\end{aligned}$$

Section 5.2 Exercises To Solutions

1. For each of the following, a transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is given by describing its action on a vector $\vec{\mathbf{x}} = [x_1, x_2]$. For each transformation, determine whether it is linear by
   - finding $T(c\,\vec{\mathbf{u}})$ and $c(T\,\vec{\mathbf{u}})$ and seeing if they are equal,
   - finding $T(\vec{\mathbf{u}} + \vec{\mathbf{v}})$ and $T(\vec{\mathbf{u}}) + T(\vec{\mathbf{v}})$ and seeing if they are equal.

   For any that you find to be linear, say so. For any that are not, say so and produce a **specific** counterexample to one of the two conditions for linearity.

   (a) $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 + x_2 \end{bmatrix}$ (b) $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_1 x_2 \end{bmatrix}$

   (c) $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} |x_1| \\ |x_2| \end{bmatrix}$ (d) $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 \\ x_1 - x_2 \end{bmatrix}$

2. The transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ defined by $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 3x_2 - 5x_1 \\ x_1 \end{bmatrix}$ is linear.

   (a) Show that, for vectors $\vec{\mathbf{u}} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\vec{\mathbf{v}} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, $T(\vec{\mathbf{u}} + \vec{\mathbf{v}}) = T(\vec{\mathbf{u}}) + T(\vec{\mathbf{v}})$.

   Do this via a string of equal expressions, beginning with $T(\vec{\mathbf{u}} + \vec{\mathbf{v}})$ and ending with $T\,\vec{\mathbf{u}} + T\,\vec{\mathbf{v}}$ as done in Example 5.2(c).

   (b) Show that, for scalar c and vector $\vec{\mathbf{u}} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $T(c\,\vec{\mathbf{u}}) = cT(\vec{\mathbf{u}}) + T\,\vec{\mathbf{v}}$. Do this

   via a string of equal expressions, beginning with $T(c\,\vec{\mathbf{u}})$ and ending with $cT(\vec{\mathbf{u}})$.

3. Two transformations from $\mathbb{R}^3$ to $\mathbb{R}^2$ are given below. One is linear and one is not. For the one that is, prove it in the manner of Example 5.2(c). For the one that is not, give a *specific* counterexample showing that the transformation violates the definition of a linear transformation. That is, show that one of $T(c\,\vec{\mathbf{u}}) = cT\,\vec{\mathbf{u}}$ or $T(\vec{\mathbf{u}}+\vec{\mathbf{v}}) = T\,\vec{\mathbf{u}} + T\,\vec{\mathbf{v}}$ fails.

(a) $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_3 \\ x_1 + x_2 + x_3 \end{bmatrix}$ (b) $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1x_2 + x_3 \\ x_1 \end{bmatrix}$

4. For each of the following transformations,

- if it is linear, give a proof that it is, in the manner of Example 5.2(c)
- if it is not linear, demonstrate that it not with an appropriate counterexample.

(a) The transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_3 \\ x_2 + x_3 \\ x_1 + x_2 \end{bmatrix}$.

(b) The transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 1 \\ x_2 - 1 \end{bmatrix}$

5. (a) $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation for which $T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$ and $T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$. Find $T\begin{bmatrix} -5 \\ -2 \end{bmatrix}$.

(b) $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation for which $T\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ and $T\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ -4 \end{bmatrix}$. Find $T\begin{bmatrix} -6 \\ 3 \end{bmatrix}$.

(c) $T : \mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation for which $T\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ -1 \end{bmatrix}$, $T\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$, and $T\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$. Find $T\begin{bmatrix} 2 \\ 7 \\ -1 \end{bmatrix}$.

(d) $T : \mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation for which $T\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix}$, $T\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$, and $T\begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$. Find $T\begin{bmatrix} 11 \\ 3 \\ -5 \end{bmatrix}$.

5.3 Linear Transformations and Matrices

Performance Criteria:

5. (e) Give the matrix representation of a linear transformation.

Recall from Example 5.2(a) that if A is an $m \times n$ matrix, then $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(\vec{\mathbf{x}}) = A\,\vec{\mathbf{x}}$ is a linear transformation. It turns out that the converse of this is true as well:

<u>THEOREM 5.3.1</u>: **Matrix of a Linear Transformation**

If $T : \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation, then there is a matrix A such that $T(\vec{\mathbf{x}}) = A\,\vec{\mathbf{x}}$ for every $\vec{\mathbf{x}}$ in $\mathbb{R}^m$. We will call A the matrix that represents the transformation.

As it is cumbersome and confusing the represent a linear transformation by the letter T and the matrix representing the transformation by the letter A, we will instead adopt the following convention: We'll denote the transformation itself by T, and the matrix of the transformation by $[T]$.

◇ **Example 5.3(a):** Find the matrix $[T]$ of the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ of Example 5.2(c), defined by $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 - x_3 \end{bmatrix}$.

Solution: We can see that $[T]$ needs to have three columns and two rows in order for the multiplication to be defined, and that we need to have

$$\begin{bmatrix} \underline{\quad} & \underline{\quad} & \underline{\quad} \\ \underline{\quad} & \underline{\quad} & \underline{\quad} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 - x_3 \end{bmatrix}$$

From this we can see that the first row of the matrix needs to be $1 \;\; 1 \;\; 0$ and the second row needs to be $0 \;\; 1 \;\; -1$. The matrix representing T is then $[T] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$.

The sort of "visual inspection" method used above can at times be inefficient, especially when trying to find the matrix of a linear transformation based on a geometric description of the action of the transformation. To see a more effective method, let's look at any linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$. Suppose that the matrix of of the transformation is $[T] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then for the two standard basis vectors $\vec{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$,

$$T(\vec{\mathbf{e}}_1) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} \qquad \text{and} \qquad T(\vec{\mathbf{e}}_2) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}.$$

This indicates that the columns of $[T]$ are the vectors $T(\vec{\mathbf{e}}_1)$ and $T(\vec{\mathbf{e}}_2)$. In general we have the following:

THEOREM 5.3.2: **Finding the Matrix of a Linear Transformation**

Let $\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, \ldots, \vec{\mathbf{e}}_m$ be the standard basis vectors of $\mathbb{R}^m$, and suppose that $T : \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation. Then the columns of $[T]$ are the vectors obtained when T acts on each of the standard basis vectors $\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, \ldots, \vec{\mathbf{e}}_m$. We indicate this by

$$[T] = [\, T(\vec{\mathbf{e}}_1)\ T(\vec{\mathbf{e}}_2)\ \cdots\ T(\vec{\mathbf{e}}_m)\,]$$

◇ **Example 5.3(b):** Let T be the transformation in $\mathbb{R}^2$ that rotates all vectors counterclockwise by ninety degrees. This is a linear transformation; use the previous theorem to determine its matrix $[T]$.

Solution: It should be clear that $T(\vec{\mathbf{e}}_1) = T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $T(\vec{\mathbf{e}}_2) = T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$. Then

$$[T] = [\, T(\vec{\mathbf{e}}_1)\ T(\vec{\mathbf{e}}_2)\,] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

The results of this section are particularly powerful from a computational point of view...

Section 5.3 Exercises To Solutions

1. Each of the following transformations T from the Section 5.2 Exercises is linear. Give the matrix $[T]$ of each.

   (a) $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_3 \\ x_1 + x_2 + x_3 \end{bmatrix}$

   (b) $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_3 \\ x_2 + x_3 \\ x_1 + x_2 \end{bmatrix}$

   (c) $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 3x_2 - 5x_1 \\ x_1 \end{bmatrix}$

   (d) $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 + x_2 \end{bmatrix}$

   (e) $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 \\ x_1 - x_2 \end{bmatrix}$

2. In this exercise we'll apply Theorem 5.3.2 to find the matrix $[T]$ of the transformation that reflects every vector in $\mathbb{R}^2$ across the line $y = x$.

   (a) Sketch the $\mathbb{R}^2$ axes, and on that sketch the line $y = x$ and the two standard basis vectors $\vec{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

   (b) Give the vectors $T(\vec{\mathbf{e}}_1)$ and $T(\vec{\mathbf{e}}_2)$.

   (c) Give $[T]$, the matrix of the transformation.

3. Now we'll use Theorem 5.3.2 to find the matrix of the transformation that projects all vectors onto the line through the origin and the point $(4,3)$.

   (a) Sketch the $\mathbb{R}^2$ axes, and then the two standard basis vectors $\vec{\mathbf{e}}_1$ and $\vec{\mathbf{e}}_2$, and the vector $\vec{\mathbf{v}}$ from the origin to the point $(4,3)$.

   (b) Recall that the projection of a vector $\vec{\mathbf{u}}$ onto a vector $\vec{\mathbf{v}}$ is given by

   $$\text{proj}_{\mathbf{v}}\, \vec{\mathbf{u}} = \frac{\vec{\mathbf{u}}\cdot\vec{\mathbf{v}}}{\vec{\mathbf{v}}\cdot\vec{\mathbf{v}}}\,\vec{\mathbf{v}}\,.$$

   Noting that projecting on the line through the origin and $(4,3)$ is the same as projecting on the vector $\vec{\mathbf{v}}$ from the origin to the point $(4,3)$, find $T(\vec{\mathbf{e}}_1)$ and $T(\vec{\mathbf{e}}_2)$.

   (c) You can now give the matrix $[T]$ that projects all vectors onto the line through the origin and $(4,3)$. Do so!

4. Repeat the process from Exercise 3 to find the matrix $[T]$ of the transformation that projects every vector on the line through the origin and the point (a,b).

5. In this exercise we'll determine the matrix of the transformation T that rotates every vector 90 degrees counterclockwise (when looking along the positive z-axis toward the origin) around the z-axis, then 90 degrees counterclockwise around the x-axis (again, with counterclockwise being as one looks along the positive x-axis toward the origin).

   (a) Consider the vector $\vec{\mathbf{v}}=\begin{bmatrix}4\\0\\3\end{bmatrix}$. What is the vector $\vec{\mathbf{v}}_1$ obtained when $\vec{\mathbf{v}}$ is rotated 90 degrees counterclockwise around the z-axis? What is the vector $\vec{\mathbf{v}}_2$ obtained when when $\vec{\mathbf{v}}_1$ is rotated 90 degrees counterclockwise around the x-axis? Note that $\vec{\mathbf{v}}_2 = T(\vec{\mathbf{v}})$.

   (b) The standard basis vectors in $\mathbb{R}^3$ are $\vec{\mathbf{e}}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $\vec{\mathbf{e}}_1 = \begin{bmatrix}0\\1\\0\end{bmatrix}$, and $\vec{\mathbf{e}}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$. Find $T(\vec{\mathbf{e}}_1)$, $T(\vec{\mathbf{e}}_2)$ and $T(\vec{\mathbf{e}}_3)$.

   (c) Give the matrix $[T]$. Multiply it times the vector $\vec{\mathbf{v}}$ from part (a) and see if you get the final result of that part, $\vec{\mathbf{v}}_2$. (You should!)

5.4 Compositions of Transformations

Performance Criterion:

5. (f) Find the composition of two transformations.

It is likely that at some point in your past you have seen the concept of the composition of two functions; if the functions were denoted by f and g, one composition of them is the new function $f \circ g$. We call this new function "f of g", and we must describe how it works. This is simple - for any x, $(f \circ g)(x) = f[g(x)]$. That is, g acts on x, and f then acts on the result. There is another composition, $g \circ f$, which is defined the same way (but, of course, in the opposite order). For specific functions, you were probably asked to find the new rule for these two compositions. Here's a reminder of how that is done:

◇ **Example 5.4(a):** For the functions $f(x) = 2x - 1$ and $g(x) = 4x - x^2$, find the formulas for the composition functions $f \circ g$ and $g \circ f$.

Solution: Basic algebra gives us

$$(f \circ g)(x) = f[g(x)] = f[4x - x^2] = 2(4x - x^2) - 1 = 8x - 2x^2 - 1 = -2x^2 + 8x - 1$$

and

$$(g \circ f)(x) = g[f(x)] = g[2x - 1] = 4(2x - 1) - (2x - 1)^2 =$$

$$(8x - 4) - (4x^2 - 4x + 1) = 8x - 4 - 4x^2 + 4x - 1 = -4x^2 + 12x - 5$$

The formulas are then $(f \circ g)(x) = -2x^2 + 8x - 1$ and $(g \circ f)(x) = -4x^2 + 12x - 5$.

Worthy of note here is that the two compositions $f \circ g$ and $g \circ f$ are not the same!

One thing that was probably glossed over when you first saw this concept was the fact that the range (all possible outputs) of the first function to act must fall within the domain (allowable inputs) of the second function to act. Suppose, for example, that $f(x) = \sqrt{x - 4}$ and $g(x) = x^2$. The function f will be undefined unless x is at least four; we indicate this by writing $f : [4, \infty) \to \mathbb{R}$. This means that we need to restrict g in such a way as to make sure that $g(x) \geq 4$ if we wish to form the composition $f \circ g$. One simple way to do this is to restrict the domain of g to $[2, \infty)$. (We could include the interval $(-\infty, -2]$ also, but for the sake of simplicity we will just use the positive interval.) The range of g is then $[4, \infty)$, which coincides with the domain of f. We now see how these ideas apply to transformations, and we see how to carry out a process like that of Example 5.4(a) for transformations.

◇ **Example 5.4(b):** Let $S : \mathbb{R}^3 \to \mathbb{R}^2$ and $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$S \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1^2 \\ x_2 x_3 \end{bmatrix}, \qquad T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 3x_2 \\ 2x_2 - x_1 \end{bmatrix}$$

Determine whether each of the compositions $S \circ T$ and $T \circ S$ exists, and find a formula for either of them that do.

Solution: Since the domain of S is $\mathbb{R}^3$ and the range of T is a subset of $\mathbb{R}^2$, the composition $S \circ T$ does not exist. The range of S falls within the domain of T, so the composition $T \circ S$ does exist. Its equation is found by

$$(T \circ S)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = T\left(S\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = T\begin{bmatrix} x_1^2 \\ x_2x_3 \end{bmatrix} = \begin{bmatrix} x_1^2 + 3x_2x_3 \\ 2x_2x_3 - x_1^2 \end{bmatrix}$$

Let's formally define what we mean by a composition of two transformations.

DEFINITION 5.4.1 **Composition of Transformations**

Let $S : \mathbb{R}^p \to \mathbb{R}^n$ and $T : \mathbb{R}^m \to \mathbb{R}^p$ be transformations. The **composition** of S and T, denoted by $S \circ T$, is the transformation $S \circ T : \mathbb{R}^m \to \mathbb{R}^n$ defined by

$$(S \circ T)\ \vec{\mathbf{x}} = S(T\ \vec{\mathbf{x}})$$

for all vectors $\vec{\mathbf{x}}$ in $\mathbb{R}^m$.

Although the above definition is valid for compositions of any transformations between vector spaces, we are primarily interested in linear transformations. Recall that any linear transformation between vector spaces can be represented by matrix multiplication for some matrix. Suppose that $S : \mathbb{R}^3 \to \mathbb{R}^3$ and $T : \mathbb{R}^2 \to \mathbb{R}^3$ are linear transformations that can be represented by the matrices

$$[S] = \begin{bmatrix} 3 & -1 & 5 \\ 0 & 2 & 1 \\ 4 & 0 & -3 \end{bmatrix} \qquad \text{and} \qquad [T] = \begin{bmatrix} 2 & 7 \\ -6 & 1 \\ 1 & -4 \end{bmatrix}$$

respectively.

◇ **Example 5.4(c):** For the transformations S and T just defined, find $(S \circ T)\ \vec{\mathbf{x}} = (S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then find the matrix of the transformation $S \circ T$.

Solution: We see that

$$\begin{aligned}(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= S\left(T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = S\left(\begin{bmatrix} 2 & 7 \\ -6 & 1 \\ 1 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) \\ &= S\begin{bmatrix} 2x_1 + 7x_2 \\ -6x_1 + x_2 \\ x_1 - 4x_2 \end{bmatrix} \\ &= \begin{bmatrix} 3 & -1 & 5 \\ 0 & 2 & 1 \\ 4 & 0 & -3 \end{bmatrix}\begin{bmatrix} 2x_1 + 7x_2 \\ -6x_1 + x_2 \\ x_1 - 4x_2 \end{bmatrix}\end{aligned}$$

$$= \begin{bmatrix} 3 & -1 & 5 \\ 0 & 2 & 1 \\ 4 & 0 & -3 \end{bmatrix} \begin{bmatrix} 2x_1 + 7x_2 \\ -6x_1 + x_2 \\ x_1 - 4x_2 \end{bmatrix}$$

$$= \begin{bmatrix} 3(2x_1 + 7x_2) - (-6x_1 + x_2) + 5(x_1 - 4x_2) \\ 0(2x_1 + 7x_2) + 2(-6x_1 + x_2) + (x_1 - 4x_2) \\ 4(2x_1 + 7x_2) + 0(-6x_1 + x_2) - 3(x_1 - 4x_2) \end{bmatrix}$$

$$= \begin{bmatrix} 17x_1 + 0x_2 \\ -11x_1 - 2x_2 \\ 5x_1 + 40x_2 \end{bmatrix}$$

From this we can see that the matrix of $S \circ T$ is $[S \circ T] = \begin{bmatrix} 17 & 0 \\ -11 & -2 \\ 5 & 40 \end{bmatrix}$.

Recall that the linear transformations of this example have matrices $[S]$ and $[T]$, and we find that

$$[S][T] = \begin{bmatrix} 3 & -1 & 5 \\ 0 & 2 & 1 \\ 4 & 0 & -3 \end{bmatrix} \begin{bmatrix} 2 & 7 \\ -6 & 1 \\ 1 & -4 \end{bmatrix} = \begin{bmatrix} 17 & 0 \\ -11 & -2 \\ 5 & 40 \end{bmatrix}.$$

This illustrates the following:

Theorem 5.4.2 Matrix of a Composition

Let $S : \mathbb{R}^p \to \mathbb{R}^n$ and $T : \mathbb{R}^m \to \mathbb{R}^p$ be linear transformations with matrices $[S]$ and $[T]$. Then

$$[S \circ T] = [S][T]$$

Section 5.4 Exercises To Solutions

1. Let $R : \mathbb{R}^2 \to \mathbb{R}^2$, $S : \mathbb{R}^2 \to \mathbb{R}^3$ and $T : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$R\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 x_2 \\ x_1 - x_2 \end{bmatrix}, \quad S\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 \\ x_2 \\ x_2 - 3x_1 \end{bmatrix}, \quad T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 - 5 \\ x_1 + x_3 + 2 \end{bmatrix}.$$

For each of the following compositions, give the declaration statement of the form $transformation : \mathbb{R}^m \to \mathbb{R}^n$ and the formula for the transformation, showing your work as done in Example 5.4(b), and simplify by combining like terms when possible.

(a) $S \circ R$ (b) $T \circ S$ (c) $R \circ T$ (d) $S \circ T$

2. Let $S : \mathbb{R}^3 \to \mathbb{R}^2$ and $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by

$$S\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_3 \\ x_1 + x_2 + x_3 \end{bmatrix} \quad \text{and} \quad T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 1 \\ x_2 - 1 \\ x_1 + x_2 \end{bmatrix}.$$

(a) Give both $S \circ T$ and $T \circ S$ in the same sort of way that S and T are given above. Combine like terms in each component, where possible.

(b) Write statements of the the form $S \circ T : \mathbb{R}^m \to \mathbb{R}^n$ for each composition, with the correct values of m and n.

3. Let $S : \mathbb{R}^2 \to \mathbb{R}^2$ and $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by

$$S \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ x_1 \end{bmatrix} \quad \text{and} \quad T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 1 \\ x_2 - 1 \\ x_1 + x_2 \end{bmatrix}.$$

Only one of $S \circ T$ and $T \circ S$ is possible. Give it in the same sort of way that S and T are given above, and write a statement of the form $transformation : \mathbb{R}^m \to \mathbb{R}^n$, with the correct values of m and n.

4. Consider the linear transformations $S \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ 2x \\ -3y \end{bmatrix}$, $T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5x - y \\ x + 4y \end{bmatrix}$.

(a) Since both of these are linear transformations, there are matrices $[S]$ and $[T]$ representing them. Give those two matrices.

(b) Give equations for either (or both) of the compositions $S \circ T$ and $T \circ S$ that exist.

(c) Give the matrix for either (or both) of the compositions that exist. Label it, with the notation $[S \circ T]$.

(d) Find either (or both) of $[S][T]$ and $[T][S]$ that exist.

(e) What did you notice in parts (c) and (d)? **Answer this with a complete sentence.**

5. Consider the three transformations

$$R \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_3 \\ x_1 + x_2 + x_3 \end{bmatrix}, \quad S \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 x_2 + x_3 \\ x_1 \end{bmatrix}, \quad T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_3 \\ x_2 + x_3 \\ x_1 + x_2 \end{bmatrix}$$

(a) Any time that we have three transformations, there are six potentially possible compositions of them, with $R \circ S$ being the "first." list the other five.

(b) Only some of the transformations that you listed are possible. Give each that is, in the same way that you have been doing, or as is shown in Example 5.4(b) and (c). Be sure to simplify where possible.

(c) Two of the transformations are linear, as is one of the compositions of them. Give the matrices of the two that are linear and the matrix of their composition. Verify that Theorem 5.4.2 holds.

5.5 Transformations and Homogeneous Coordinates

Performance Criteria:

5. (g) Find matrices that perform combinations of dilations, reflections, rotations and translations in $\mathbb{R}^2$ using homogenous coordinates.

We are now equipped to return to the applications of rotation and reflection matrices in the context of linear transformations. We know from Example 5.2(a) that a transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ defined by multiplication by a matrix is a linear transformation. One should be able to convince oneself geometrically as well that rotations and reflections are linear.

If we wanted to perform a rotation T followed by a reflection S, this would be done by the composition $S \circ T$, and we know from the previous section that the matrix of $S \circ T$ is simply $[S][T]$. Using formulas from Chapter 3 to get the matrices $[S]$ and $[T]$, it is then fairly simple to come up with a single matrix to perform the desired composition. Transformations like rotations and reflections are quite useful in areas like robotics and computer graphics, and when using them we often wish to compose several such transformations as just described.

In example 5.2(d) we saw that translations like

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + a \\ x_2 + b \end{bmatrix} \tag{1}$$

are not linear, so they do not have matrix representations. What we will find in this section is that if we work in the two-dimensional plane $z = 1$ in $\mathbb{R}^3$, a translation like (1) becomes a shear in $\mathbb{R}^3$, which *IS* linear. Before looking into how this is done, we first see a method for multiplying a matrix times several vectors all at the same time.

◇ **Example 5.5(a):** Let $A = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$, $\vec{\mathbf{u}}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$, $\vec{\mathbf{u}}_3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$. Find each of $A\vec{\mathbf{u}}_1$, $A\vec{\mathbf{u}}_2$, $A\vec{\mathbf{u}}_3$.

Solution: $A\vec{\mathbf{u}}_1 = \begin{bmatrix} -1 \\ 22 \end{bmatrix}$, $A\vec{\mathbf{u}}_2 = \begin{bmatrix} 23 \\ 4 \end{bmatrix}$, $A\vec{\mathbf{u}}_3 = \begin{bmatrix} 6 \\ 21 \end{bmatrix}$.

◇ **Example 5.5(b):** Let A be as in the previous example, and let $B = \begin{bmatrix} 1 & 7 & 3 \\ 4 & -2 & 3 \end{bmatrix}$, the matrix whose columns are $\vec{\mathbf{u}}_1$, $\vec{\mathbf{u}}_2$ and $\vec{\mathbf{u}}_3$ from Example 5.5(a). Compute AB.

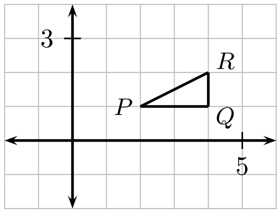
Solution: $AB = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} 1 & 7 & 3 \\ 4 & -2 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 23 & 6 \\ 22 & 4 & 21 \end{bmatrix}$.

We note in the above two examples that the columns of AB are simply the results of multiplying each column of B individually by A. The reason that this is of interest to us is that we will wish to perform a linear transformation, like a rotation, on a geometric object in $\mathbb{R}^2$. Conceptually we would then multiply every point (of which there are infinitely many) of the object by a rotation matrix. In practice, if the object is a polygon all we have to do is transform each of the vertices of the polygon and connect the resulting vertices with line segments in order to transform the entire polygon. Let's demonstrate with an example. We will utilize the fact that the matrix

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

rotates all points in $\mathbb{R}^2$ 90 degrees counterclockwise around the origin.

◇ **Example 5.5(c):** Find the triangle $\triangle P'Q'R'$ obtained by rotating the triangle $\triangle PQR$ shown to the right counterclockwise 90 degrees around the origin.

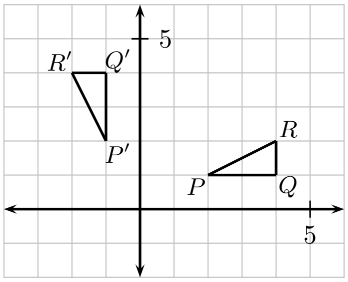


Solution: We can represent the triangle $\triangle PQR$ by the matrix $[PQR] = \begin{bmatrix} 2 & 4 & 4 \\ 1 & 1 & 2 \end{bmatrix}$.

From the above discussion we know we can create the new triangle $\triangle P'Q'R'$ by simply multiplying the matrix $[PQR]$ representing $\triangle PQR$ by the rotation matrix A given above to get a matrix $[P'Q'R']$ whose columns are the points P', Q' and R' of the transformed triangle $\triangle P'Q'R'$. We then simply plot the vertices P', Q' and R' and connect them in order to get $\triangle P'Q'R'$.

$$[P'Q'R'] = A\,[PQR] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 & 4 \\ 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} -1 & -1 & -2 \\ 2 & 4 & 4 \end{bmatrix} \tag{2}$$

On the grid to the right we see the original triangle $\triangle PQR$ and the transformed triangle $\triangle P'Q'R'$ whose vertices are given by the columns of the matrix $[P'Q'R']$ obtained by the multiplication (2) above.



In Example 5.5(a) we found that if $A = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}$ and $\vec{\mathbf{u}}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, then $A\,\vec{\mathbf{u}}_1 = \begin{bmatrix} -1 \\ 22 \end{bmatrix}$.
Note how this result compares with the following example.

◇ **Example 5.5(d):** Let $C = \begin{bmatrix} 3 & -1 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\vec{\mathbf{w}} = \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$. Find the product $C\,\vec{\mathbf{w}}$.

Solution: $C\,\vec{\mathbf{w}} = \begin{bmatrix} 3 & -1 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 22 \\ 1 \end{bmatrix}$

Observe carefully how the matrix C was obtained from A by augmenting with a column of zeros and then adding a row of two zeros and a one, and how $\vec{\mathbf{w}}$ was obtained from $\vec{\mathbf{u}}_1$ by adding a third component of one. The result of $C\vec{\mathbf{w}}$ is then the result of $A\vec{\mathbf{u}}_1$, but also with an additional component of one. We can thus do $\mathbb{R}^2$ transformations, like rotations and reflections, in the plane $z=1$ in this manner.

Why would we want to do this? The following example will show that.

◇ **Example 5.5(e):** What is the result when the matrix $A = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix}$ acts on the vector $\vec{\mathbf{x}}_h = \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$ through multiplication?

Solution: $A\vec{\mathbf{x}}_h = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 + a \\ x_2 + b \\ 1 \end{bmatrix}$

We say that $\vec{\mathbf{x}}_h$ is the **homogeneous coordinate** vector in $\mathbb{R}^3$ for the vector $\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ in $\mathbb{R}^2$, and we can see that the first two components of $A\vec{\mathbf{x}}_h$ are the result of the translation

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + a \\ x_2 + b \end{bmatrix}.$$

A translation is not linear in $\mathbb{R}^2$, but it *IS* linear when performed as a shear in the plane $z=1$ in $\mathbb{R}^3$. This allows us to do a translation with a homogenous matrix in $\mathbb{R}^3$.

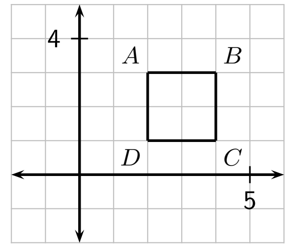
◇ **Example 5.5(f):** Use a homogenous matrix to translate $\vec{\mathbf{x}} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ three units left an one unit up.

Solution: In this case the homogeneous translation matrix is $A = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ and the homogeneous form of $\vec{\mathbf{x}}$ is $\vec{\mathbf{x}}_h = \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}$. Multiplying them gives us

$$A\vec{\mathbf{x}}_h = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix},$$

so the translation of $\vec{\mathbf{x}}$ is $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

Now we're finally ready to do something interesting! Consider the square $ABCD$ shown to the right, and suppose that we wish to rotate it 30 degrees counterclockwise about its center. We know how to obtain a matrix that rotates objects 30 degrees counterclockwise *around the origin*, but not around other points. The idea here is simple, though. Let T be the translation that shifts the square so that its center is at the origin, and let R_{30} be a rotation of 30 degrees counterclockwise around the origin. If we first apply the transformation T, then R_{30}, then T^{-1} (to move the square back after rotating it), we will accomplish what we want. We translate the square to the origin, rotate it there, then move the rotated square back to the original location. This is the composition $T^{-1} \circ R_{30} \circ T$ (remember that the rightmost transformation acts first!), and from the previous section we know that the matrix of this transformation is the product $[T^{-1}][R_{30}][T]$ of the individual transformation matrices. We will need to recall the following from Chapter 3:



Rotation Matrix in $\mathbb{R}^2$

For the matrix $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ and any position vector $\vec{\mathbf{x}}$ in $\mathbb{R}^2$, the product $A\,\vec{\mathbf{x}}$ is the vector resulting when $\vec{\mathbf{x}}$ is rotated counterclockwise around the origin by the angle θ.

◇ **Example 5.5(g):** Create a homogenous matrix to rotate the square $ABCD$ 30 degrees counterclockwise around its center.

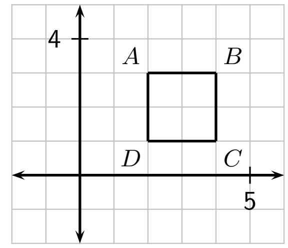
Solution: Note that the transformation T shifts every point three units to the left and two units down, so T^{-1} must shift every point three units to the right and two units up. We can determine the homogeneous matrices of these by using the form demonstrated in Example 5.5(e), and the rotation matrix can be obtained using the formula above, but we need to augment with a column of zeros and add the row $0\ \ 0\ \ 1$. We thus have

$$[T] = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}, \qquad [T^{-1}] = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \qquad [R_{30}] = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The matrix that rotates the square 30 degrees counterclockwise around its center is then the product $[T^{-1}][R_{30}][T]$, computed below:

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & \frac{8-3\sqrt{3}}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & \frac{1-2\sqrt{3}}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

◇ **Example 5.5(h):** Apply the matrix from Example 5.5(g) to the square $ABCD$ shown to the right to rotate it 30 degrees counterclockwise around its center.

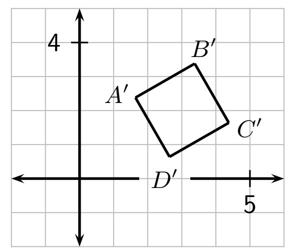


Solution: We can represent the square $ABCD$ with a homogenous matrix, shown below and to the left. Each column gives the coordinates of a vertex, with an extry component of one to represent the point in homogeneous coordinates. Converting the entries of the transformation matrix from Example 5.5(g) to decimal form gives the matrix to the right below.

$$[ABCD] = \begin{bmatrix} 2 & 4 & 4 & 2 \\ 3 & 3 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \qquad [T^{-1} \circ R_{30} \circ T] = \begin{bmatrix} 0.87 & -0.50 & 1.40 \\ 0.50 & 0.87 & -1.23 \\ 0 & 0 & 1 \end{bmatrix}$$

To get the coordinates of the vertices of the new square $A'B'C'D'$ we let our transformation act on the original through the product $[T^{-1} \circ R_{30} \circ T][ABCD]$ to get the homogeneous matrix of the new square, shown below and to the left. The new square is graphed below and to the right, and we can see that it is square $ABCD$ rotated 30 degrees counterclockwise around its center.

$$[A'B'C'D'] = \begin{bmatrix} 1.64 & 3.38 & 4.38 & 2.64 \\ 2.38 & 3.38 & 1.64 & 0.64 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$



We end this section with a comment. Pretty much every calculation we have done all term boils down to adding and multiplying. If one were to be writing code to do the things you've done in this assignment, you would simply do it as a bunch of multiplications and additions, rather than doing it with matrices. For example, to rotate the point (x, y) by an angle of θ around the origin we would simply compute the new coordinates (w, z) by

$$w = x\cos\theta - y\sin\theta, \qquad\qquad z = x\sin\theta + y\cos\theta$$

The advantage of linear algebra for tasks like this is not computational, but conceptual. Without the theory that we have developed, figuring out how to do transformations like the ones you will see in some of the exercises would be far more difficult!

Section 5.5 Exercises To Solutions

For these exercises we will use the following notation, which is not necessarily standard. Those of you who encounter these ideas in a robotics course will see a standard notation that is somewhat more complicated than we need now. Here is what we'll use.

- R_θ will be a rotation of θ, with the rotation being copunterclockwise if θ is positive, and clockwise if θ is negative.
- $T_{(a,b)}$ will be a translation by a units in the x-direction and b units in the y-direction.
- $R_{(a,b)}$ will be a reflection across the line through the origin and the point (a,b).

Note that we are using R for both rotations and reflections, but which it is in each case should be clear from the subscripts.

1. For each of the following, give the 3×3 homogeneous matrix that would be used to perform the given transformation on vectors/points in $\mathbb{R}^2$ expressed in homogeneous coordinates. You should not need a formula for the given reflections.

   (a) R_{-90° (b) $T_{(3,-5)}$ (c) $R_{(1,0)}$

   (d) $T_{(1,2)}$ (e) $R_{(1,-1)}$ (f) $R_{\pi/3}$

2. Use the notation described at the start of these xercises to describe each of the following transformations as a composition of rotations, translations and reflections.

   (a) A reflection across the line $y = \frac{3}{2}x$ followed by a rotation of 50 degrees counterclockwise around the origin.

   (b) A rotation of 50 degrees counterclockwise around the origin followed by a reflection across the line $y = \frac{3}{2}x$.

   (c) A rotation of 25 degrees clockwise around the point $(6,-2)$.

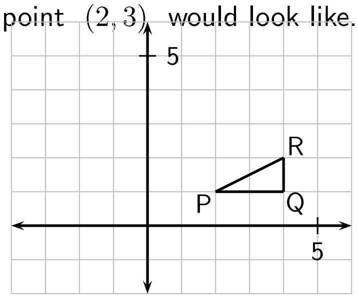
   (d) A reflection across the line $y = x - 3$. (**Hint:** Translate so that the line goes through the origin, reflect, translate back.)

3. (a) Find a homogenous matrix that will rotate all points in $\mathbb{R}^2$ 90° counterclockwise about the point $(2,3)$, using homogeneous coordinates.

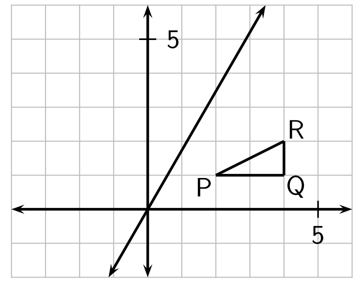
   (b) Consider the triangle $\triangle PQR$ shown to the right below. Sketch what you think the result $\triangle P'Q'R'$ of the rotation 90° counterclockwise about the point $(2,3)$ would look like.

   (c) Use your answer to (a) and a homogenous coordinate representation of the triangle $\triangle PQR$ to find the rotated coordinates P', Q' and R'.

   (d) Plot the rotated points and draw the rotated triangle $\triangle P'Q'R'$ on the grid to the right. If it doesn't look like what you predicted in (b), figure out which is wrong and correct it.



4. Suppose that we wish to reflect $\triangle PQR$ across the line through the origin and at an angle of $60°$ to the positive x-axis, as shown in the picture below and to the right. We already know how to reflect across the x-axis, so we'll take advantage of that fact. What we want to do is rotate the line to the x-axis, reflect, then rotate back.



   (a) Find the single matrix that does this by multiplying some other matrices. **Round the entries of the final matrix to the nearest hundredth, or give them in exact form.**

   (b) Apply the matrix to the homogenous coordinates of P, Q and R to get vertices of a new triangle $\triangle P'Q'R'$.

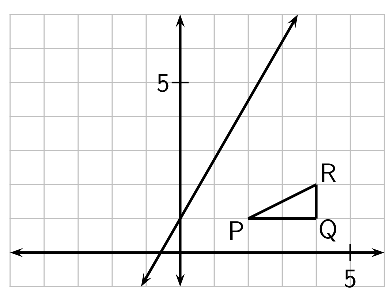
   (c) Draw $\triangle P'Q'R'$ on the graph to the right. If it doesn't look like the reflection of $\triangle PQR$ across the line, find your error and correct it.

5. Use a method like that of the previous exercise to derive the following formula. You will need to use the facts that the cosine of an angle of a right triangle is the adjacent side over the hypotemuse and the sine of the angle is the opposite side over the hypotenuse.

Reflection Matrix in $\mathbb{R}^2$

For the matrix $C = \begin{bmatrix} \dfrac{a^2-b^2}{a^2+b^2} & \dfrac{2ab}{a^2+b^2} \\ \dfrac{2ab}{a^2+b^2} & \dfrac{b^2-a^2}{a^2+b^2} \end{bmatrix}$ and any position vector $\vec{\mathbf{x}}$ in $\mathbb{R}^2$, the product $C\,\vec{\mathbf{x}}$ is the vector resulting when $\vec{\mathbf{x}}$ is reflected across the line containing the origin and the point (a, b).

6. Find a single homogeneous matrix that will reflect $\triangle PQR$ across the line through the point $(0, 1)$ and at an angle of $60°$ to the x-axis, shown on the graph to the right. Test your result as you have been. **Round the entries of the matrix to the nearest hundredth, or give them in exact form.**



5.6 An Introduction to Eigenvalues and Eigenvectors

Performance Criteria:

5. (h) Determine whether a given vector is an eigenvector for a matrix; if it is, give the corresponding eigenvalue.

   (i) Determine eigenvectors and corresponding eigenvalues for linear transformations in $\mathbb{R}^2$ or $\mathbb{R}^3$ that are described geometrically.

Recall that the two main features of a vector in $\mathbb{R}^n$ are direction and magnitude. In general, when we multiply a vector $\vec{\mathbf{x}}$ in $\mathbb{R}^n$ by an $n\times n$ matrix A, the result $A\vec{\mathbf{x}}$ is a new vector in $\mathbb{R}^n$ whose direction and magnitude are different than those of $\vec{\mathbf{x}}$. For every square matrix A there are some vectors whose directions are not changed (other than perhaps having their directions reversed) when multiplied by the matrix. That is, multiplying $\vec{\mathbf{x}}$ by A gives the same result as multiplying $\vec{\mathbf{x}}$ by a scalar. It is very useful for certain applications to identify which vectors those are, and what the corresponding scalar is. Let's use the following example to get started:

◇ **Example 5.6(a):** Multiply $A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$ times $\vec{\mathbf{u}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\vec{\mathbf{v}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and determine whether multiplication by A is the same as multiplying by a scalar in either case.

Solution:

$$A\vec{\mathbf{u}} = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -22 \\ 18 \end{bmatrix}, \quad A\vec{\mathbf{v}} = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

For the first multiplication there appears to be nothing special going on. For the second multiplication, the effect of multiplying $\vec{\mathbf{v}}$ by A is the same as simply multiplying $\vec{\mathbf{v}}$ by -1. Note also that

$$\begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} -6 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ -3 \end{bmatrix}, \qquad \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 8 \\ -4 \end{bmatrix} = \begin{bmatrix} -8 \\ 4 \end{bmatrix}$$

It appears that if we multiply any scalar multiple of $\vec{\mathbf{v}}$ by A the same thing happens; the result is simply the negative of the vector. That is, $A\vec{\mathbf{x}} = (-1)\vec{\mathbf{x}}$ for every scalar multiple of $\vec{\mathbf{x}}$.

We say that $\vec{\mathbf{v}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and all of its scalar multiples are **eigenvectors** of A, with *corresponding* **eigenvalue** -1. Here is the formal definition of an eigenvalue and eigenvector:

Definition 5.6.1: **Eigenvalues and Eigenvectors**

A scalar λ is called an **eigenvalue** of a matrix A if there is a *nonzero* vector $\vec{\mathbf{x}}$ such that

$$A\vec{\mathbf{x}} = \lambda\vec{\mathbf{x}}\,.$$

The vector $\vec{\mathbf{x}}$ is an **eigenvector** corresponding to the eigenvalue λ.

Make special note of this:

An eigenvector must be a nonzero vector, but zero *IS* allowed as an eigenvalue.

One comment is in order at this point. Suppose that $\vec{\mathbf{x}}$ has n components. Then $\lambda\vec{\mathbf{x}}$ does as well, so A must have n rows. However, for the multiplication $A\vec{\mathbf{x}}$ to be possible, A must also have n columns. For this reason, *only square matrices have eigenvalues and eigenvectors*. We now see how to determine whether a vector is an eigenvector of a matrix.

◇ **Example 5.6(b):** Determine whether either of $\vec{\mathbf{w}}_1 = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$ and $\vec{\mathbf{w}}_2 = \begin{bmatrix} -3 \\ 3 \end{bmatrix}$ are eigenvectors for the matrix $A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$ of Example 5.6(a). If either is, give the corresponding eigenvalue.

Solution: We see that

$$A\vec{\mathbf{w}}_1 = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 4 \\ -1 \end{bmatrix} = \begin{bmatrix} -10 \\ 7 \end{bmatrix} \quad \text{and} \quad A\vec{\mathbf{w}}_2 = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} -3 \\ 3 \end{bmatrix} = \begin{bmatrix} -6 \\ 6 \end{bmatrix}$$

$\vec{\mathbf{w}}_1$ is not an eigenvector of A because there is no scalar λ such that $A\vec{\mathbf{w}}_1$ is equal to $\lambda\vec{\mathbf{w}}_1$. $\vec{\mathbf{w}}_2$ *IS* an eigenvector, with corresponding eigenvalue $\lambda = 2$, because $A\vec{\mathbf{w}}_2 = 2\,\vec{\mathbf{w}}_2$.

Note that for the 2×2 matrix A of Examples 5.6(a) and (b) we have seen two eigenvalues now. It turns out that those are the only two eigenvalues, which illustrates the following:

> Theorem 5.6.2: The number of eigenvalues of an $n \times n$ matrix is at most n.

Do not let the use of the Greek letter lambda intimidate you - it is simply some scalar! It is tradition to use λ to represent eigenvalues. Now suppose that $\vec{\mathbf{x}}$ is an eigenvector of an $n \times n$ matrix A, with corresponding eigenvalue λ, and let c be any scalar. Then for the vector $c\vec{\mathbf{x}}$ we have

$$A(c\vec{\mathbf{x}}) = c(A\vec{\mathbf{x}}) = c(\lambda\vec{\mathbf{x}}) = (c\lambda)\vec{\mathbf{x}} = \lambda(c\vec{\mathbf{x}})$$

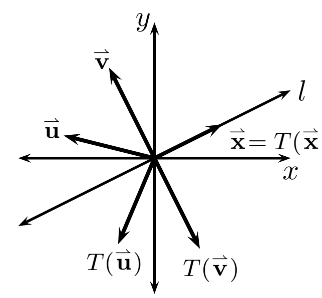
This shows that any scalar multiple of $\vec{\mathbf{x}}$ is also an eigenvector of A with the same eigenvalue λ. We saw this in Example 5.6(a). The set of all scalar multiples of $\vec{\mathbf{x}}$ is of course a subspace of $\mathbb{R}^n$, and we call it the **eigenspace** corresponding to λ. $\vec{\mathbf{x}}$, or any scalar multiple of it, is a basis for the eigenspace. The two eigenspaces you have seen so far have dimension one, but an eigenspace can have a higher dimension.

> Definition 5.6.3: **Eigenspace Corresponding to an Eigenvalue**
>
> For a given eigenvalue λ_j of an $n\times n$ matrix A, the **eigenspace** E_j corresponding to λ is the set of all eigenvectors corresponding to λ_j. It is a subspace of $\mathbb{R}^n$.

So far we have been looking at eigenvectors and eigenvalues from a purely algebraic viewpoint, by looking to see if the equation $A\vec{\mathbf{x}} = \lambda\vec{\mathbf{x}}$ held for some vector $\vec{\mathbf{x}}$ and some scalar λ. It is useful to have some geometric understanding of eigenvectors and eigenvalues as well. In the next two examples we consider eigenvectors and eigenvalues of two linear transformations in $\mathbb{R}^2$ from a geometric standpoint. Although we have defined eigenvalues in terms of matrices, recall that any *linear* transformation T can be represented by a matrix T, so it makes sense to talk about eigenvectors and eigenvalues of a transformation, as long as it is linear. We simply substitute the equation $T(\vec{\mathbf{x}}) = \lambda\vec{\mathbf{x}}$, which tells us that $\vec{\mathbf{x}}$ is an eigenvector if the action of T on it leaves its direction unchanged or opposite of what it was (or, in the case of $\lambda = 0$, "shrinks it to the zero vector").

◇ **Example 5.6(c):** The transformation T that reflects very vector in $\mathbb{R}^2$ over the line l with equation $y = \frac{1}{2}x$ is a linear transformation. Determine the eigenvectors and corresponding eigenvalues for this transformation.

Solution: We begin by observing that any vector that lies on l will be unchanged by the reflection, so it will be an eigenvector, with eigenvalue $\lambda = 1$. These vectors are all the scalar multiples of $\vec{\mathbf{x}} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$; see the picture to the right. A vector not on the line, $\vec{\mathbf{u}}$, is shown along with its reflection $T(\vec{\mathbf{u}})$ as well. We can see that its direction is changed, so it is not an eigenvector. However, for any vector $\vec{\mathbf{v}}$ that is perpendicular to l we have $T(\vec{\mathbf{v}}) = -\vec{\mathbf{v}}$. Therefore any such vector is an eigenvector with eigenvalue $\lambda = -1$. Those vectors are all the scalar multiples of $\vec{\mathbf{x}} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.



◇ **Example 5.6(d):** Let T be the transformation T that rotates every vector in $\mathbb{R}^2$ by thirty degrees counterclockwise; this is a linear transformation. Determine the eigenvectors and corresponding eigenvalues for this transformation.

Solution: Because every vector in $\mathbb{R}^2$ will be rotated by thirty degrees, the direction of every vector will be altered, so there are no eigenvectors for this transformation.

Our conclusion in Example 5.6(d) is correct in one sense, but incorrect in another. Geometrically, in a way that we can see, the conclusion is correct. Algebraically, the transformation has eigenvectors, but their components are complex numbers, and the corresponding eigenvalues are complex numbers as well. In this course we will consider only real eigenvectors and eigenvalues.

Section 5.6 Exercises

To Solutions

1. Consider the matrix $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$.

   (a) Find $A\vec{\mathbf{x}}$ for each of the following vectors:

$$\vec{\mathbf{x}}_1 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}, \quad \vec{\mathbf{x}}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \vec{\mathbf{x}}_3 = \begin{bmatrix} 1 \\ 5 \end{bmatrix}, \quad \vec{\mathbf{x}}_4 = \begin{bmatrix} -3 \\ -3 \end{bmatrix}, \quad \vec{\mathbf{x}}_5 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

(b) Give the vectors from part (a) that are eigenvectors and, for each, give the corresponding eigenvalue.

(c) Give one of the eigenvalues that you have found. Then give the general form for *any* eigenvector corresponding to that eigenvalue.

(d) Repeat part (c) for the other eigenvalue that you have found.

2. For each of the following a matrix is given, along with several vectors. Determine which of the vectors are eigenvectors, and give the corresponding eigenvalue for each.

(a) $A = \begin{bmatrix} 2 & 7 \\ -1 & -6 \end{bmatrix}, \quad \vec{\mathbf{u}}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \vec{\mathbf{u}}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{u}}_3 = \begin{bmatrix} -7 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{u}}_4 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad \vec{\mathbf{v}}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \vec{\mathbf{v}}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{v}}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{v}}_4 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 4 \end{bmatrix}, \quad \vec{\mathbf{w}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{\mathbf{w}}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{w}}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{w}}_4 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$

(d) $A = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}, \quad \vec{\mathbf{u}}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \vec{\mathbf{u}}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{u}}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{u}}_4 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{u}}_5 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$

3. Suppose that a matrix A has eigenvectors $\vec{\mathbf{x}}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$ and $\vec{\mathbf{x}}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ with respective eigenvalues $\lambda_1 = 2$ and $\lambda_2 = -1$. Which of the following are also eigenvectors, and what are their corresponding eigenvalues?

$$\vec{\mathbf{v}}_1 = \begin{bmatrix} -6 \\ 0 \\ -12 \end{bmatrix}, \quad \vec{\mathbf{v}}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{v}}_3 = \begin{bmatrix} 6 \\ 2 \\ -2 \end{bmatrix}, \quad \vec{\mathbf{v}}_4 = \begin{bmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \quad \vec{\mathbf{v}}_5 = \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}$$

4. The previous exercise is based on the idea that if $\vec{\mathbf{x}}$ is an eigenvector of A, then any scalar multiple of $\vec{\mathbf{x}}$ is also an eigenvector, with the same eigenvalue. This exercise will show that there can be linearly independent eigenvectors that share the same eigenvalue. Determine which of the vectors below are eigenvectors for the matrix

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

For those that are, give the corresponding eigenvalue.

$$\vec{\mathbf{u}}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \quad \vec{\mathbf{u}}_2 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \quad \vec{\mathbf{u}}_3 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad \vec{\mathbf{u}}_4 = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}, \quad \vec{\mathbf{u}}_5 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$$

5. For each transformation described geometrically, give as many *independent* eigenvectors, and their corresponding eigenvalues, as you can. *Keep in mind that any vectors that become the zero vector under the transformation, with zero as an eigenvalue.*

   (a) The transformation that reflects every vector in $\mathbb{R}^2$ across the y-axis.

   (b) The transformation that projects every vector in $\mathbb{R}^2$ onto the y-axis.

   (c) The transformation that projects every vector in $\mathbb{R}^2$ onto the line $y = 3x$.

   (d) The transformation that reflects every vector in $\mathbb{R}^3$ across the yz-plane.

   (e) The transformation that rotates every vector in $\mathbb{R}^3$ 90 degrees around the z-axis.

   (f) The transformation that projects every vector in $\mathbb{R}^3$ onto the xz-plane.

   (g) The transformation that reflects every vector in $\mathbb{R}^3$ across the plane with equation $z = -y$. (**Hint:** Sketch a picture of the graph of this equation in the yz-plane.)

5.7 Finding Eigenvalues and Eigenvectors

Performance Criteria:

5. (j) Find the characteristic polynomial for a 2×2 or 3×3 matrix. Use it to find the eigenvalues of the matrix.
   (k) Give the eigenspace E_j corresponding to an eigenvalue λ_j of a matrix.
   (l) Determine the principal stresses and the orientation of the principal axes for a two-dimensional stress element.

So where are we now? We know what eigenvectors, eigenvalues and eigenspaces are, and we know how to determine whether a vector is an eigenvector of a matrix. There are two big questions at this point:

- Why do we care about eigenvalues and eigenvectors?
- If we are just given a square matrix A, how do we find its eigenvalues and eigenvectors?

We will not see the answer to the first question until the end of this section. First we'll address the second question.

Finding Eigenvalues

We begin by rearranging the eigenvalue/eigenvector equation $A\vec{\mathbf{x}} = \lambda\vec{\mathbf{x}}$ a little. First, we can subtract $\lambda\vec{\mathbf{x}}$ from both sides to get

$$A\vec{\mathbf{x}} - \lambda\vec{\mathbf{x}} = \vec{\mathbf{0}}\ .$$

Note that the right side of this equation must be the zero *vector*, because both $A\vec{\mathbf{x}}$ and $\lambda\vec{\mathbf{x}}$ are vectors. At this point we want to factor $\vec{\mathbf{x}}$ out of the left side, but if we do so carelessly we will get a factor of $A-\lambda$, which makes no sense because A is a matrix and λ is a scalar! Note, however, that we can replace $\vec{\mathbf{x}}$ with $I\vec{\mathbf{x}}$, thus we can replace $\lambda\vec{\mathbf{x}}$ with $(\lambda I)\vec{\mathbf{x}}$, allowing us to factor $\vec{\mathbf{x}}$ out:

$$\begin{aligned} A\vec{\mathbf{x}} - \lambda\vec{\mathbf{x}} &= \vec{\mathbf{0}} \\ A\vec{\mathbf{x}} - (\lambda I)\vec{\mathbf{x}} &= \vec{\mathbf{0}} \\ (A - \lambda I)\vec{\mathbf{x}} &= \vec{\mathbf{0}} \end{aligned}$$

Now $A - \lambda I$ is just a matrix - let's call it B for now. Any *nonzero* (by definition) vector $\vec{\mathbf{x}}$ that is a solution to $B\vec{\mathbf{x}} = \vec{\mathbf{0}}$ is an eigenvector for A. Clearly the zero vector is a solution to $B\vec{\mathbf{x}} = \vec{\mathbf{0}}$, and if B is invertible that will be the only solution. But since eigenvectors are nonzero vectors, A will have eigenvectors only if B is not invertible. Recall that one test for invertibility of a matrix is whether its determinant is nonzero. For B to not be invertible, then, its determinant must be zero. But B is $A - \lambda I$, so we want to find values of λ for which $\det(A - \lambda I) = 0$. (Note that the determinant of a matrix is a scalar, so the zero here is just the scalar zero.) We introduce a bit of special language that we use to discuss what is happening here:

<u>Definition 5.7.1</u>: **Characteristic Polynomial and Equation**

Taking λ to be an unknown, $\det(A-\lambda I)$ is a polynomial called the **characteristic polynomial** of A. The equation $\det(A - \lambda I) = 0$ is called the **characteristic equation** for A, and its solutions are the eigenvalues of A.

Before looking at a specific example, you would probably find it useful to go back and look at Examples 3.8(a), (b) and (c), and to recall the following.

Determinant of a 2×2 Matrix

The determinant of the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\det(A) = ad - bc$.

Determinant of a 3×3 Matrix

To find the determinant of a 3×3 matrix,

- Augment the matrix with its first two columns.
- Find the product down each of the three complete "downward diagonals" of the augmented matrix, and the product up each of the three "upward diagonals."
- Add the products from the downward diagonals and subtract each of the products from the upward diagonals. The result is the determinant.

Now we're ready to look at a specific example of how to find the eigenvalues of a matrix.

◇ **Example 5.7(a):** Find the eigenvalues of the matrix $A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$.

Solution: We need to find the characteristic polynomial $\det(A-\lambda I)$, then set it equal to zero and solve.

$$A - \lambda I = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} -4-\lambda & -6 \\ 3 & 5-\lambda \end{bmatrix}$$

$$\det(A-\lambda I) = (-4-\lambda)(5-\lambda) - (3)(-6)) = (-20 - \lambda + \lambda^2) + 18 = \lambda^2 - \lambda - 2$$

We now factor this and set it equal to zero to find the eigenvalues:

$$\lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1) = 0 \qquad \Longrightarrow \qquad \lambda = 2, -1$$

We use subscripts to distinguish the different eigenvalues: $\lambda_1 = 2,\ \lambda_2 = -1$.

Finding Eigenvectors

We now need to find the eigenvectors or, more generally, the eigenspaces, corresponding to each eigenvalue. We defined eigenspaces in the previous section, but here we will give a slightly different (but equivalent) definition.

Definition 5.7.2: **Eigenspace Corresponding to an Eigenvalue**

For a given eigenvalue λ_j of an $n \times n$ matrix A, the **eigenspace** E_j corresponding to λ_j is the set of all solutions to the equation

$$(A - \lambda_j I)\,\vec{\mathbf{x}} = \vec{\mathbf{0}}\ .$$

It is a subspace of $\mathbb{R}^n$.

Note that we indicate the correspondence of an eigenspace with an eigenvalue by subscripting them with the same number.

◇ **Example 5.7(b):** Find the eigenspace E_1 of the matrix $A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$ corresponding to the eigenvalue $\lambda_1 = 2$.

Solution: For $\lambda_1 = 2$ we have $A - \lambda I = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} -6 & -6 \\ 3 & 3 \end{bmatrix}$

The augmented matrix of the system $(A - \lambda I)\,\vec{\mathbf{x}} = \vec{\mathbf{0}}$ is then $\begin{bmatrix} -6 & -6 & 0 \\ 3 & 3 & 0 \end{bmatrix}$, which reduces to $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. The top row represents the equation $x_1 + x_2 = 0$ so any values of x_1 and x_2 that make this true will give us an eigenvector so, for example, we can take $\vec{\mathbf{x}}$ to be $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. The eigenspace corresponding to $\lambda_1 = 2$ can then be described by either of

$$E_1 = \left\{ t \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\} \qquad \text{or} \qquad E_1 = \operatorname{span}\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$$

It would be beneficial for the reader to repeat the above process for the second eigenvalue $\lambda_2 = -1$ and verify that

$$E_2 = \left\{ t \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}.$$

When first seen, the whole process for finding eigenvalues and eigenvectors can be a bit bewildering! Here is a summary of the process:

Finding Eigenvalues and Bases for Eigenspaces

The following procedure will give the eigenvalues and corresponding eigenspaces for a square matrix A.

1) Find $\det(A - \lambda I)$. This is the characteristic polynomial of A.

2) Set the characteristic polynomial equal to zero and solve for λ to get the eigenvalues.

3) For a given eigenvalue λ_i, solve the system $(A - \lambda_j I)\vec{\mathbf{x}} = \vec{\mathbf{0}}$. The set of solutions is the eigenspace corresponding to λ_j. The vector or vectors whose linear combinations make up the eigenspace are a basis for the eigenspace.

◇ **Example 5.7(c):** Give the characteristic polynomial of the matrix

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

and use it to determine the eigenvalues. Then, given that one of the eigenvalues is $\lambda_1 = 3$, give the corresponding eigenspace E_1.

Solution: First we see that

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

The characteristic polynomial is

$$\begin{aligned} \det(A - \lambda I) &= \det\left(\begin{bmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{bmatrix}\right) \\ &= (1-\lambda)(2-\lambda)(1-\lambda) - (1-\lambda) - (1-\lambda) \\ &= (1-\lambda)(2-\lambda)(1-\lambda) - 2(1-\lambda). \end{aligned}$$

Ordinarily we would just multiply everything out, combine like terms and solve by factoring using algebra methods or a computational tool. In this case, however, we can factor $(1-\lambda)$ out of both terms to get

$$\begin{aligned} \det(A - \lambda I) &= (1-\lambda)[(2-\lambda)(1-\lambda) - 2] \\ &= (1-\lambda)[(2 - 3\lambda + \lambda^2) - 2] \\ &= (1-\lambda)(\lambda^2 - 3\lambda) \\ &= \lambda(1-\lambda)(\lambda - 3). \end{aligned}$$

This last expression is the characteristic polynomial, in factored form, and we can see that the eigenvalues are 0, 1 and 3.

To find the eigenspace corresponding to $\lambda_1 = 3$ we solve the characteristic equation $(A - \lambda I)\vec{\mathbf{x}} = \vec{\mathbf{0}}$ for $\lambda = 3$. First we compute

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} - 3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{bmatrix} \end{aligned}$$

The augmented matrix for $(A - \lambda I)\vec{\mathbf{x}} = \vec{\mathbf{0}}$ is shown below and to the left, and its row reduced form is below and to the right.

$$\begin{bmatrix} -2 & -1 & 0 & 0 \\ -1 & -1 & -1 & 0 \\ 0 & -1 & -2 & 0 \end{bmatrix} \quad \overset{rref}{\Longrightarrow} \quad \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{1}$$

We see that x_3 is a free variable. Letting $x_3 = t$ and solving $x_2 + 2x_3 = 0$ gives us $x_2 = -2t$. Solving $x_1 - x_3 = 0$ gives us $x_1 = t$. The solution to $(A - \lambda I)\vec{\mathbf{x}} = \vec{\mathbf{0}}$ is then the set of vectors of the form

$$\begin{bmatrix} t \\ -2t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

and the eigenspace corresponding to $\lambda_1 = 3$ is

$$E_1 = \left\{ t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}.$$

We can vary the above process slightly as follows. After obtaining the row-reduced form (1) we get the equation $x_2 + 2x_3 = 0$, and we can see that $x_2 = -2$ and $x_3 = 1$ is a solution. For that choice of x_3 the first equation $x_1 - x_3 = 0$ gives us that $x_1 = 1$ as well. This gives us the single eigenvector $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$, and the corresponding eigenspace is all scalar multiples of that vector, as shown above.

An Application - Principal Stress

Figure 5.7(a) below and to the left shows a cantilevered beam embedded in a wall at its right end. There is a force acting downward on the left end of the beam, due to the weight of the beam and perhaps a load applied to the end of the beam. The small square **stress element** shown on the side of the beam toward us is an imaginary square extending back through to the far side of the beam. That element is subject to stresses of tension (indicated by the little arrows) and compression, as well as some **shear stress**. The element is blown up in Figure 5.7(b) and all of the stresses on it are shown. σ_x and σ_y are **normal stresses**, and τ_{xy} and τ_{yx} are **shear stresses**. The fact that the element is not rotating gives us that $\tau_{xy} = \tau_{yx}$.

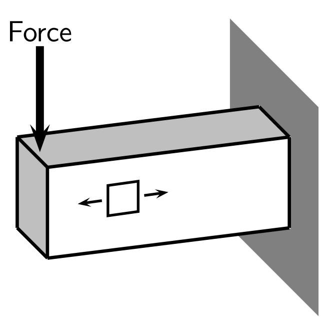


Figure 5.7(a)

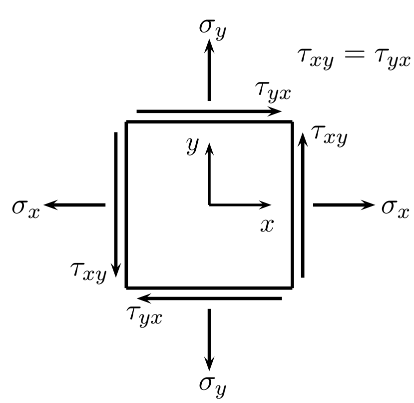


Figure 5.7(b)

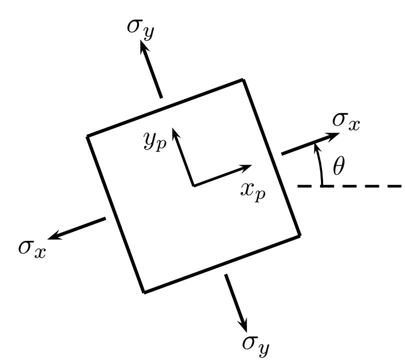


Figure 5.7(c)

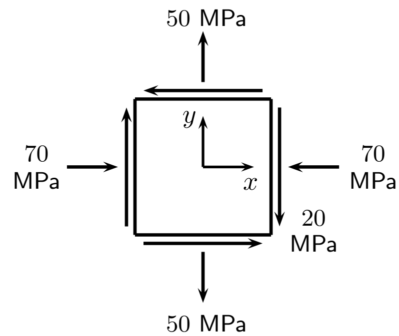
We can rotate the stress element about its center and the stresses will all change as we do that.

There is one angle of rotation we are particularly interested in. If we rotate by a particular angle θ all of the stress will be normal, with no shear stresses. This is depicted in Figure 5.7(c) above and to the right. The new axes x_p and y_p are called the **principal axes**. The two normal stresses for this orientation of the stress element are called **principal stresses**, and one of them is the greatest stress at the location of the stress element in the beam. The key to finding the principal stresses and their directions is eigenvalues and eigenvectors! We begin by setting up a **stress matrix**:

$$\begin{bmatrix} \sigma_x & \tau_{xy} \\ \tau_{yx} & \sigma_y \end{bmatrix},$$

keeping in mind that $\tau_{xy} = \tau_{yx}$. Stresses oriented in the directions of the arrows in Figure 5.7(b) are taken to be positive, any oriented opposite of those arrows are considered to be negative. We then find the eigenvalues and corresponding eigenvectors of the stress matrix. The eigenvalues give us the normal stresses in the directions of their corresponding eigenvectors. If we reorient the stress element to have the principal stresses as its normal stresses there will be no shear stresses. Let's see how this happens in practice.

◇ **Example 5.7(d):** The normal and shear stresses on a stress element are shown to the right. Determine the principal stresses and the angle of rotation (from the positive x-axis, as shown in Figure 5.7(c)). Sketch the stress element and indicate the stresses when the normal stresses are the principal stresses. Give values in decimal form, rounded to the nearest tenth.



Solution: Accounting for the orientations of the stresses, the stress matrix is $\begin{bmatrix} -70 & -20 \\ -20 & 50 \end{bmatrix}$.

The characteristic polynomial is $(-70-\lambda)(50-\lambda)-(-20)^2 = \lambda^2+20\lambda-3900$. Setting it equal to zero and solving with the quadratic formula gives us $\lambda = -73.2, 53.2$. The augmented matrices for the system $(A-\lambda I)\,\vec{\mathbf{x}} = \vec{\mathbf{0}}$ for each of these eigenvalues are

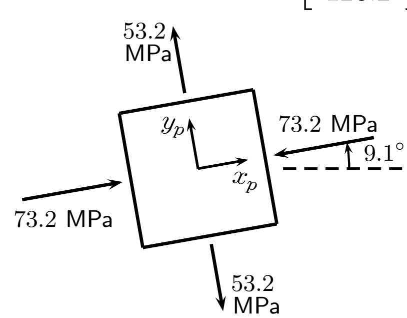
$$\left[\begin{array}{cc|c} 3.2 & -20 & 0 \\ -20 & 126.2 & 0 \end{array}\right] \quad \text{and} \quad \left[\begin{array}{cc|c} -126.2 & -20 & 0 \\ -20 & -3.2 & 0 \end{array}\right].$$

These will not row-reduce to give a second row of zeros because of needing to round the eigenvalues, so we assume that they reduce to

$$\left[\begin{array}{cc|c} 3.2 & -20 & 0 \\ 0 & 0 & 0 \end{array}\right] \quad \text{and} \quad \left[\begin{array}{cc|c} -126.2 & -20 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

An eigenvector for $\lambda_1 = -73.2$ is $\begin{bmatrix} 20 \\ 3.2 \end{bmatrix}$ and for $\lambda = 53.2$ we get the eigenvector $\begin{bmatrix} -20 \\ 126.2 \end{bmatrix}$.

We can verify that these vectors are essentially perpendicular by taking their dot product and seeing that it is very close to zero, and would be if we hadn't rounded our eigenvalues. The angle of the first vector with the x-axis is $\theta = \tan^{-1}\dfrac{3.2}{20} = 9.1^\circ$. (Draw a sketch of the vector to see how we get this.) The stress element reoriented along the principal axes, along with the principal stresses, is shown to the right.



When analyzing beams we generally encounter stress elements like the one in Example 5.7(d), with tension in one direction (the y- and y_p-directions in this case) and compression in the other (here in the x- and x_p-directions). When analyzing a sheet structure we can see tension in both normal directions, and something like soil underground can exhibit compression in both normal directions. Those cases will arise in the exercises.

Section 5.7 Exercises To Solutions

1. Use the method of Example 5.7(b) to find the eigenspace of $A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$ corresponding to $\lambda_2 = -1$.

2. Find the eigenvalues and corresponding eigenspaces for each of the following matrices. Answers are given in the back of the book, but check your answers yourself by multiplying the matrix times each basis eigenvector to make sure the result is the same as multiplying the eigenvector by the eigenvalue.

   (a) $\begin{bmatrix} 8 & 3 \\ 2 & 7 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$

3. For each of the following a matrix is given, and it's action as a transformation on vectors/points in $\mathbb{R}^2$ is described. Find the eigenvectors and eigenspaces, and make sure that your results make sense for the transformation described.

   (a) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, which reflects vectors across the line $y = x$.

   (b) $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$, which projects vectors onto the line $y = -x$.

   (c) $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, which reflects vectors across the y-axis.

4. Follow a process like the last half of Example 5.7(c) to find the eigenspaces E_2 and E_3 of

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

   corresponding to the eigenvalues $\lambda_2 = 1$ and $\lambda_3 = 0$.

5. Consider the matrix $A = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$.

   (a) Find characteristic polynomial by computing $\det(A - \lambda I)$. As in Example 5.7(c), you will initially have two terms that both have a factor of $1 - \lambda$ in them. *Do not expand (multiply out) these terms - instead, factor the common factor of $1 - \lambda$ out of both*, then combine and simplify the rest, as done in that example.

(b) Give the characteristic equation for matrix A, which is obtained by setting the characteristic polynomial equal to zero. Remember that you are doing this because the equation $A\vec{\mathbf{x}} = \lambda\vec{\mathbf{x}}$ will only have solutions $\vec{\mathbf{x}} \neq \vec{\mathbf{0}}$ if $\det(A - \lambda I) = 0$. Find the solutions (eigenvalues) of the equation by factoring the part that is not already factored.

(c) One of your eigenvalues should be one; let's refer to it as λ_1. Find a basis for the eigenspace E_1 corresponding to $\lambda = 1$ by solving the equation $(A - I)\vec{\mathbf{x}} = \vec{\mathbf{0}}$. ($(A - \lambda I)$ becomes $(A - I)$ because $\lambda_1 = 1$.) Conclude by giving the eigenspace E_1 using correct notation.

(d) Give the eigenspaces corresponding to the other two eigenvalues. *Make it clear which eigenspace is associated with which eigenvalue.*

(e) Check your answers by multiplying each eigenvector by the original matrix A to see if the result is the same as multiplying the eigenvector by the corresponding eigenvalue. In other words, if the eigenvector is $\vec{\mathbf{x}}$, check to see that $A\vec{\mathbf{x}} = \lambda\vec{\mathbf{x}}$.

6. (a) For $A = \begin{bmatrix} 1 & 2 & 1 \\ 6 & -1 & 0 \\ -1 & -2 & -1 \end{bmatrix}$, find characteristic polynomial. In this case you can't factor something out as in Example 5.7(c), so you must multiply everything out.

(b) Give the characteristic equation for matrix A, which is obtained by setting the characteristic polynomial equal to zero. You should be able to solve it by first factoring $-\lambda$ out, then factoring the remaining quadratic. Do this and give the eigenvalues.

(c) Give the eigenspaces corresponding to the eigenvalues. *Make it clear which eigenspace is associated with which eigenvalue.*

7. (a) Find characteristic polynomial for $A = \begin{bmatrix} -2 & -4 & 2 \\ -2 & 1 & 2 \\ 4 & 2 & 5 \end{bmatrix}$. You again can't factor something out, so just have multiply out (carefully!) and combine like terms.

(b) Use Wolfram Alpha or some other tool to factor the characteristic polynomial. (Just use x instead of λ when doing this.) Set the result equal to zero (giving the characteristic equation) and solve to get the eigenvalues.

(c) Give the eigenspaces corresponding to the eigenvalues, again making it clear which eigenspace is associated with which eigenvalue.

8. (a) Computing $\det(A - \lambda I)$ for $A = \begin{bmatrix} 0 & 0 & 2 \\ -3 & 1 & 6 \\ 0 & 0 & -1 \end{bmatrix}$ will directly result in a factored form of the characteristic polynomial of A. Give it.

(b) Give the eigenvalues.

(c) Find the eigenspaces corresponding to the eigenvalues by solving $(A - \lambda I)\vec{\mathbf{x}} = \vec{\mathbf{0}}$ for each eigenvalue. For one eigenvalue you will have two free variables, resulting in an eigenspace of dimension two (with two independent basis eigenvectors).

9. Find the eigenvalues and eigenspaces for $A = \begin{bmatrix} 7 & 0 & -3 \\ -9 & -2 & 3 \\ 18 & 0 & -8 \end{bmatrix}$, using a tool if necessary for factoring the characteristic polynomial. As with Exercise 7, there are only two eigenspaces, and one of them has dimension two.

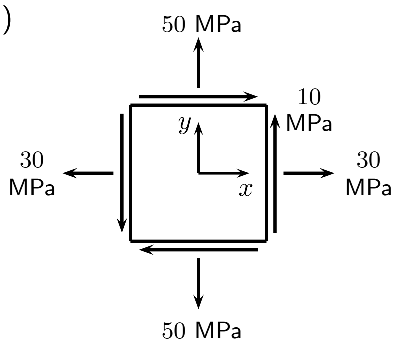
10. So far, every $n \times n$ matrix we have worked with has had n linearly independent eigenvectors, perhaps with fewer eigenvalues (Exercises 7 and 8). The matrix

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

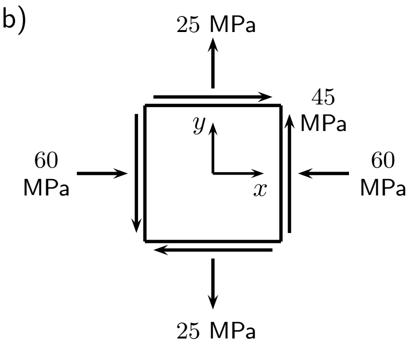
has only two eigenvalues and only two independent eigenvectors. Find the eigenvalues and eigenvectors.

11. For each of the following a stress element is given, along with its normal and shear stresses. Determine the principal stresses and their eigenvectors. Then determine the angle between the principal x_p-axis and the original x-axis, and whether the angle is positive or negative, following the standard convention that counterclockwise is positive and clockwise is negative. Sketch the rotated stress element, showing the angle of rotation of the axes and the principal stresses, indicating whether each is tension or compression by the direction of its arrow, as done in Example 5.7(d).

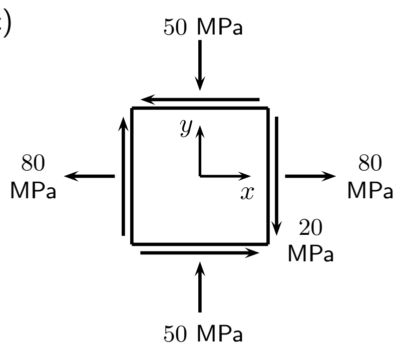
(a)



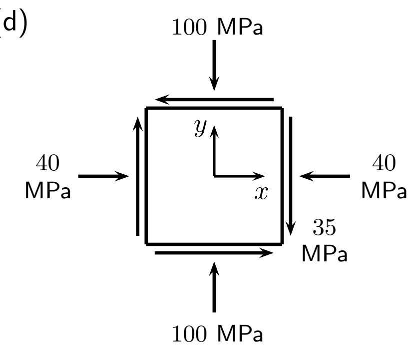
(b)



(c)



(d)



5.8 Diagonalization of Matrices

Performance Criterion:

5. (m) Diagonalize a matrix; know the forms of the matrices P and D from $P^{-1}AP = D$.

We begin with an example involving the matrix A from Examples 5.5(a) and (b).

◇ **Example 5.8(a):** For $A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$ and $P = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}$, find the product $P^{-1}AP$.

Solution: First we obtain $P^{-1} = \dfrac{1}{(-1)(1)-(1)(-2)} \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}$. Then

$$P^{-1}AP = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

We want to make note of a few things here:

- The columns of the matrix P are eigenvectors for A.
- The matrix $D = P^{-1}AP$ is a diagonal matrix.
- The diagonal entries of D are the eigenvalues of A, in the order of the corresponding eigenvectors in P.

For a square matrix A, the process of creating such a matrix D in this manner is called **diagonalization** of A. This cannot always be done, but often it can. (We will fret about exactly when it can be done later.) The point of the rest of this section is to see a use or two of this idea.

Before getting to the key application of this section we will consider the following. Suppose that we wish to find the kth power of a 2×2 matrix A with eigenvalues λ_1 and λ_2 and having corresponding eigenvectors that are the columns of P. Then solving $P^{-1}AP = D$ for A gives $A = PDP^{-1}$ and

$$\begin{aligned} A^k = (PDP^{-1})^k &= (PDP^{-1})(PDP^{-1})\cdots(PDP^{-1}) \\ &= PD(P^{-1}P)D(P^{-1}P)\cdots(P^{-1}P)DP^{-1} \\ &= PDDD\cdots DP^{-1} \\ &= PD^kP^{-1} \\ &= P\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^k P^{-1} \\ &= P\begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix} P^{-1} \end{aligned}$$

Therefore, once we have determined the eigenvalues and eigenvectors of A we can simply take each eigenvector to the kth power, then put the results in a diagonal matrix and multiply *once* by P on the left and P^{-1} on the right.

◇ **Example 5.8(b):** Diagonalize the matrix $A = \begin{bmatrix} 3 & 12 & -21 \\ -1 & -6 & 13 \\ 0 & -2 & 6 \end{bmatrix}$.

Solution: First we find the eigenvalues by solving $\det(A - \lambda I) = 0$:

$$\begin{aligned} \det \begin{bmatrix} 3-\lambda & 12 & -21 \\ -1 & -6-\lambda & 13 \\ 0 & -2 & 6-\lambda \end{bmatrix} &= (3-\lambda)(-6-\lambda)(6-\lambda) - 42 + 26(3-\lambda) + 12(6-\lambda) \\ &= (-18 + 3\lambda + \lambda^2)(6-\lambda) - 42 + 78 - 26\lambda + 72 - 12\lambda \\ &= -108 + 18\lambda + 18\lambda - 3\lambda^2 + 6\lambda^2 - \lambda^3 + 108 - 38\lambda \\ &= -\lambda^3 + 3\lambda^2 - 2\lambda \\ &= -\lambda(\lambda^2 - 3\lambda + 2) \\ &= -\lambda(\lambda - 2)(\lambda - 1) \end{aligned}$$

The eigenvalues of A are then $\lambda = 0, 1, 2$. We now find an eigenvector corresponding to $\lambda = 0$ by solving the system $(A - \lambda I)\,\vec{\mathbf{x}} = 0$. The augmented matrix and its row-reduced form are shown below:

$$\begin{bmatrix} 3 & 12 & -21 & 0 \\ -1 & -6 & 13 & 0 \\ 0 & -2 & 6 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 0 & 5 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Longrightarrow \begin{array}{l} \text{Let } x_3 = 1. \\ \text{Then } x_2 = 3 \\ \text{and } x_1 = -5 \end{array}$$

The eigenspace corresponding to the eigenvalue $\lambda = 0$ is then the span of the vector $\vec{\mathbf{v}}_1 = [-5, 3, 1]$. For $\lambda = 1$ we have

$$\begin{bmatrix} 2 & 12 & -21 & 0 \\ -1 & -7 & 13 & 0 \\ 0 & -2 & 5 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 0 & \frac{9}{2} & 0 \\ 0 & 1 & -\frac{5}{2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Longrightarrow \begin{array}{l} \text{Let } x_3 = 2. \\ \text{Then } x_2 = 5 \\ \text{and } x_1 = -9 \end{array}$$

The eigenspace corresponding to the eigenvalue $\lambda = 1$ is then the span of the vector $\vec{\mathbf{v}}_2 = [-9, 5, 2]$ (obtained by multiplying the solution vector by two in order to get a vector with integer components). Finally, for $\lambda = 2$ we have

$$\begin{bmatrix} 1 & 12 & -21 & 0 \\ -1 & -8 & 13 & 0 \\ 0 & -2 & 4 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Longrightarrow \begin{array}{l} \text{Let } x_3 = 1. \\ \text{Then } x_2 = 2 \\ \text{and } x_1 = -3 \end{array}$$

so the eigenspace corresponding to the eigenvalue $\lambda = 2$ is then the span of the vector $\vec{\mathbf{v}}_3 = [-3, 2, 1]$. The diagonalization of A is then $D = P^{-1}AP$, where

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad \text{and} \qquad P = \begin{bmatrix} -5 & -9 & -3 \\ 3 & 5 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Section 5.8 Exercises To Solutions

1. Consider matrix again the matrix $A = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$ from Exercise 2 of Section 11.2.

   (a) Let P be a matrix whose columns are eigenvectors for the matrix A. (The basis vectors for each of the three eigenspaces will do.) Give P and P^{-1}, using your calculator to find P^{-1}.

   (b) Find $P^{-1}AP$, using your calculator if you wish. The result should be a diagonal matrix with the eigenvalues on its diagonal. If it isn't, check your work from Exercise 4.

2. Now let $B = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$.

   (a) Find characteristic polynomial by computing $\det(B - \lambda I)$. If you expand along the second column you will obtain a characteristic polynomial that already has a factor of $2 - \lambda$.

   (b) Give the characteristic *equation* (make sure it has an equal sign!) for matrix B. Find the roots (eigenvalues) by factoring. *Note that in this case one of the eigenvalues is repeated.* This is not a problem.

   (c) Find and describe (as in Exercise 1(c)) the eigenspace corresponding to each eigenvalue. The repeated eigenvalue will have *TWO* eigenvectors, so that particular eigenspace has dimension two. State your results as sentences, and use set notation for the bases.

3. Repeat the process from Exercise 1 for the matrix B from Exercise 2.

5.9 Solving Systems of Differential Equations

Performance Criteria:

5. (n) Write a system of linear differential equations in matrix-vector form. Write the initial conditions in vector form.

   (0) Solve a system of two linear differential equations; solve an initial value problem for a system of two linear differential equations.

We now get to the centerpiece of this section. Recall that the solution to the initial value problem $x'(t) = kx(t),\ x(0) = C$ is $x(t) = Ce^{kt}$. Now let's consider the **system of two differential equations**

$$\begin{aligned} x_1' &= x_1 + 2x_2 \\ x_2' &= 3x_1 + 2x_2, \end{aligned}$$

where x_1 and x_2 are functions of t. Note that the two equations are **coupled**; the equation containing the derivative x_1' contains the function x_1 itself, but also contains x_2. The same sort of situation occurs with x_2'. The key to solving this system is to *uncouple* the two equations, and eigenvalues and eigenvectors will allow us to do that!

We will also add in the initial conditions $x_1(0) = 10,\ x_2(0) = 5$. If we let $\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ we can rewrite the system of equations and initial conditions as follows:

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \end{bmatrix}$$

which can be condensed to

$$\vec{\mathbf{x}}' = A\,\vec{\mathbf{x}}, \quad \vec{\mathbf{x}}\,(0) = \begin{bmatrix} 10 \\ 5 \end{bmatrix} \tag{1}$$

This is the matrix initial value problem that is completely analogous to $x'(t) = kx(t),\ x(0) = C$.

Before proceeding farther we note that the matrix A has eigenvectors $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ with corresponding eigenvalues $\lambda = 4$ and $\lambda = -1$. Thus, if $P = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$ we then have $P^{-1}AP = D = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$ and $A = PDP^{-1}$.

We can substitute this last expression for A into the vector differential equation in (1) to get $\vec{\mathbf{x}}' = PDP^{-1}\,\vec{\mathbf{x}}$. If we now multiply both sides on the left by P^{-1} we get $P^{-1}\,\vec{\mathbf{x}}' = DP^{-1}\,\vec{\mathbf{x}}$. We now let $\vec{\mathbf{y}} = P^{-1}\,\vec{\mathbf{x}}$; Since P^{-1} is simply a matrix of constants, we then have $\vec{\mathbf{y}}' = (P^{-1}\,\vec{\mathbf{x}})' = P^{-1}\,\vec{\mathbf{x}}'$ also. Making these two substitutions into $P^{-1}\,\vec{\mathbf{x}}' = DP^{-1}\,\vec{\mathbf{x}}$ gives us $\vec{\mathbf{y}}' = D\,\vec{\mathbf{y}}$. By the same substitution we also have $\vec{\mathbf{y}}\,(0) = P^{-1}\,\vec{\mathbf{x}}\,(0)$. We now have the new initial value problem

$$\vec{\mathbf{y}}' = D\,\vec{\mathbf{y}}, \quad \vec{\mathbf{y}}\,(0) = P^{-1}\,\vec{\mathbf{x}}\,(0). \tag{3}$$

Here the vector $\vec{\mathbf{y}}$ is simply the unknown vector $\vec{\mathbf{y}} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ and $\vec{\mathbf{y}}\,(0) = \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix}$ which can be determined by $\vec{\mathbf{y}}\,(0) = P^{-1}\,\vec{\mathbf{x}}\,(0)$. Because the coefficient matrix of the system in (3) is diagonal, the two differential equations can be uncoupled and solved to find $\vec{\mathbf{y}}$. Now since $\vec{\mathbf{y}} = P^{-1}\vec{\mathbf{x}}$ we also have $\vec{\mathbf{x}} = P\vec{\mathbf{y}}$, so after we find $\vec{\mathbf{y}}$ we can find $\vec{\mathbf{x}}$ by simply multiplying $\vec{\mathbf{y}}$ by P.

So now it is time for you to make all of this happen!

Section 5.9 Exercises To Solutions

1. Write the system $\vec{\mathbf{y}}' = D\vec{\mathbf{y}}$ in that form, then as two differential equations. Solve the differential equations. There will be two arbitrary constants; distinguish them by letting one be C_1 and the other C_2. solve the two equations to find $y_1(t)$ and $y_2(t)$.

2. Find P^{-1} and use it to find $\vec{\mathbf{y}}(0)$. Use $y_1(0)$ and $y_2(0)$ to find the constants in your two differential equations.

3. Use $\vec{\mathbf{x}} = P\vec{\mathbf{y}}$ to find x. Finish by giving the functions $x_1(t)$ and $x_2(t)$.

4. Check your final answer by doing the following. If your answer doesn't check, go back and find your error. I had to do that, so you might as well also!

   (a) Make sure that $x_1(0) = 10$ and $x_2(0) = 5$.

   (b) Put x_1 and x_2 into the equations (1) and make sure you get true statements.

B Solutions to Exercises

B.5 Chapter 5 Solutions

Section 5.1 Solutions Back to 5.1 Exercises

1. (a) $T(\vec{\mathbf{u}}) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $T(\vec{\mathbf{v}}) = \begin{bmatrix} -6 \\ 9 \end{bmatrix}$ (b) $T(\vec{\mathbf{u}}) = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$, $T(\vec{\mathbf{v}}) = \begin{bmatrix} -1 \\ -9 \\ -2 \end{bmatrix}$

   (c) $T(\vec{\mathbf{v}}) = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$, $T(\vec{\mathbf{w}}) = \begin{bmatrix} -1 \\ 2 \\ 7 \end{bmatrix}$ (d) $T(\vec{\mathbf{u}}) = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 2 \end{bmatrix}$

   (e) $T(\vec{\mathbf{u}}) = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$, $T(\vec{\mathbf{w}}) = \begin{bmatrix} 7 \\ -5 \\ 6 \end{bmatrix}$ (f) $T(\vec{\mathbf{v}}) = \begin{bmatrix} -9 \\ 12 \\ -1 \\ -5 \\ 1 \end{bmatrix}$

2. (a) no matrix (b) $A = \begin{bmatrix} 1 & 1 \\ 0 & 3 \\ -1 & 0 \end{bmatrix}$ (c) $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$

   (d) $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ (e) no matrix (f) $A = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$

3. (a) $T : \mathbb{R}^2 \to \mathbb{R}^2, \quad T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}$ (b) $T : \mathbb{R}^2 \to \mathbb{R}^2, \quad T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 \end{bmatrix}$

   (c) $T : \mathbb{R}^2 \to \mathbb{R}^2, \quad T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 3 \\ x_2 + 1 \end{bmatrix}$ (d) $T : \mathbb{R}^2 \to \mathbb{R}^2, \quad T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ -x_1 \end{bmatrix}$

   (e) $T : \mathbb{R}^2 \to \mathbb{R}^2, \quad T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$ (f) $T : \mathbb{R}^3 \to \mathbb{R}^3, \quad T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_2 \\ x_3 \end{bmatrix}$

   (g) $T : \mathbb{R}^3 \to \mathbb{R}^3, \quad T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \\ x_3 \end{bmatrix}$ (h) $T : \mathbb{R}^3 \to \mathbb{R}^3, \quad T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_3 \\ x_2 \\ x_1 \end{bmatrix}$

   (i) $T : \mathbb{R}^3 \to \mathbb{R}^3, \quad T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ x_3 \end{bmatrix}$ (j) $T : \mathbb{R}^3 \to \mathbb{R}^3, \quad T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ x_2 \\ 0 \end{bmatrix}$

(k) $T:\mathbb{R}^2 \to \mathbb{R}^3, \quad T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 2 \end{bmatrix}$

(l) $T:\mathbb{R}^3 \to \mathbb{R}^3, \quad T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - 1 \\ x_3 + 4 \end{bmatrix}$

4. $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix}$

5. $T:\mathbb{R}^3 \to \mathbb{R}^3, \quad T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ x_2 \\ 0 \end{bmatrix}$

Section 5.2 Solutions

Back to 5.2 Exercises

1. (a) The transformation is linear.

(b) $T\left(2\begin{bmatrix} 3 \\ 5 \end{bmatrix}\right) = T\begin{bmatrix} 6 \\ 10 \end{bmatrix} = \begin{bmatrix} 16 \\ 60 \end{bmatrix}$ and $2T\begin{bmatrix} 3 \\ 5 \end{bmatrix} = 2\begin{bmatrix} 8 \\ 15 \end{bmatrix} = \begin{bmatrix} 16 \\ 30 \end{bmatrix}$,

so $T\left(2\begin{bmatrix} 3 \\ 5 \end{bmatrix}\right) \neq 2T\begin{bmatrix} 3 \\ 5 \end{bmatrix}$ and T is not linear

(c) $T\left(-2\begin{bmatrix} 3 \\ -5 \end{bmatrix}\right) = T\begin{bmatrix} -6 \\ 10 \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \end{bmatrix}$ and $-2T\begin{bmatrix} 3 \\ -5 \end{bmatrix} = -2\begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} -6 \\ -10 \end{bmatrix}$

so $T\left(-2\begin{bmatrix} 3 \\ -5 \end{bmatrix}\right) \neq -2T\begin{bmatrix} 3 \\ -5 \end{bmatrix}$ and T is not linear.

(d) The transformation is linear.

2. (a) $T(\vec{\mathbf{u}} + \vec{\mathbf{v}}) = T\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right) = T\begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 + 2(u_2 + v_2) \\ 3(u_2 + v_2) - 5(u_1 + v_1) \\ u_1 + v_1 \end{bmatrix} =$

$$\begin{bmatrix} u_1 + 2u_2 \\ 3u_2 - 5u_1 \\ u_1 \end{bmatrix} + \begin{bmatrix} v_1 + 2v_2 \\ 3v_2 - 5v_1 \\ v_1 \end{bmatrix} = T\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + T\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = T(\vec{\mathbf{u}}) + T(\vec{\mathbf{v}})$$

(b) $T(c\,\vec{\mathbf{u}}) = T\left(c\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) = T\begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix} = \begin{bmatrix} cu_1 + 2(cu_2) \\ 3(cu_2) - 5(cu_1) \\ cu_1 \end{bmatrix} =$

$$\begin{bmatrix} c(u_1 + 2u_2) \\ c(3u_2 - 5u_1) \\ c(u_1) \end{bmatrix} = c\begin{bmatrix} u_1 + 2u_2 \\ 3u_2 - 5u_1 \\ u_1 \end{bmatrix} = cT\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = cT(\vec{\mathbf{u}})$$

3. (a) The transformation is linear:

(b) Not linear: $T\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}\right) = T\begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix} = \begin{bmatrix} 44 \\ 5 \end{bmatrix}$ and $T\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + T\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} =$

$\begin{bmatrix} 5 \\ 1 \end{bmatrix} + \begin{bmatrix} 26 \\ 4 \end{bmatrix} = \begin{bmatrix} 31 \\ 5 \end{bmatrix}$, so $T\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}\right) \neq T\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + T\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$

4. (a) The transformation is linear:

(b) Not linear. Note that $T\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, which violates Theorem 5.2.2.

5. (a) $T\begin{bmatrix} -5 \\ -2 \end{bmatrix} = \begin{bmatrix} -21 \\ 5 \end{bmatrix}$ (b) $T\begin{bmatrix} -6 \\ 3 \end{bmatrix} = \begin{bmatrix} 16 \\ -19 \end{bmatrix}$

(c) $T\begin{bmatrix} 2 \\ 7 \\ -1 \end{bmatrix} = \begin{bmatrix} 24 \\ -14 \end{bmatrix}$ (d) $T\begin{bmatrix} 11 \\ 3 \\ -5 \end{bmatrix} = \begin{bmatrix} 11 \\ 14 \\ 17 \end{bmatrix}$

Section 5.3 Solutions Back to 5.3 Exercises

1. (a) $[T] = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$ (b) $[T] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ (c) $[T] = \begin{bmatrix} 1 & 2 \\ -5 & 3 \\ 1 & 0 \end{bmatrix}$

(d) $[T] = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ (e) $[T] = \begin{bmatrix} 3 & 0 \\ 1 & -1 \end{bmatrix}$

Section 5.4 Solutions Back to 5.4 Exercises

1. (a) $S \circ R : \mathbb{R}^2 \to \mathbb{R}^3$, $(S \circ R)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1x_2 + x_1 - x_2 \\ x_1 - x_2 \\ x_1 - x_2 - 3x_1x_2 \end{bmatrix}$

(b) $T \circ S : \mathbb{R}^2 \to \mathbb{R}^2$, $(T \circ S)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 - 5 \\ -x_1 + 2x_2 + 2 \end{bmatrix}$

(c) $R \circ T : \mathbb{R}^3 \to \mathbb{R}^2$, $(R \circ T)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1x_2 + x_2x_3 - 5x_1 + 2x_2 - 5x_3 - 10 \\ -x_1 + x_2 - x_3 - 7 \end{bmatrix}$

(d) $S \circ T : \mathbb{R}^3 \to \mathbb{R}^3$, $(S \circ T)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + x_3 - 8 \\ x_1 + x_3 + 2 \\ x_1 - 3x_2 + x_3 + 17 \end{bmatrix}$

2. (a) $(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 + 2 \\ 2x_1 + 2x_2 \end{bmatrix}$, $(T \circ S)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_3 + 1 \\ x_1 + x_2 + x_3 - 1 \\ 3x_1 + x_2 \end{bmatrix}$

3. $(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 - 1 \\ x_1 + x_2 \\ x_1 + 1 \end{bmatrix}$, $S \circ T : \mathbb{R}^2 \to \mathbb{R}^3$

4. (a) $[S] = \begin{bmatrix} 1 & 1 \\ 2 & 0 \\ 0 & -3 \end{bmatrix}$, $[T] = \begin{bmatrix} 5 & -1 \\ 1 & 4 \end{bmatrix}$

(b) $(S\circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = S\begin{bmatrix} 5x_1 - x_2 \\ x_1 + 4x_2 \end{bmatrix} = \begin{bmatrix} 6x_1 + 3x_2 \\ 10x_1 - 2x_2 \\ -3x_1 - 12x_2 \end{bmatrix}$

(c) $[S\circ T] = \begin{bmatrix} 6 & 3 \\ 10 & -2 \\ -3 & -12 \end{bmatrix}$ (d) $[S][T] = \begin{bmatrix} 6 & 3 \\ 10 & -2 \\ -3 & -12 \end{bmatrix}$

(e) $[S\circ T] = [S][T]$

5. (a) $R\circ T$, $S\circ R$, $S\circ T$, $T\circ R$, $T\circ S$

(b) $(R\circ T)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 + 2x_3 \\ 2x_1 + 2x_2 + 2x_3 \end{bmatrix}$, $(S\circ T)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1x_2 + x_1x_3 + x_2x_3 + x_3^2 + x_1 + x_2 \\ x_1 + x_3 \end{bmatrix}$

(c) $[R] = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$, $[T] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$, $[R\circ T] = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$

$$[R][T] = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix} = [R\circ T]$$

Section 5.5 Solutions Back to 5.5 Exercises

1. (a) $[R_{-90^\circ}] = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $[T_{(3,-5)}] = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{bmatrix}$

(c) $[R_{(1,0)}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (d) $[T_{(1,2)}] = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

(e) $[R_{(1,-1)}] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (f) $[R_{\pi/3}] = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$

2. (a) $R_{50^\circ}\circ R_{(2,3)}$ (b) $R_{(2,3)}\circ R_{50^\circ}$

(c) $T_{(6,-2)}\circ R_{-25^\circ}\circ T_{(-6,2)}$ (d) $T_{(0,-3)}\circ R_{(1,1)}\circ T_{(0,3)}$

Section 5.6 Solutions Back to 5.6 Exercises

1. (a) $A\vec{\mathbf{x}}_1 = \begin{bmatrix} 9 \\ 18 \end{bmatrix}$, $A\vec{\mathbf{x}}_2 = \begin{bmatrix} 1 \\ -8 \end{bmatrix}$, $A\vec{\mathbf{x}}_3 = \begin{bmatrix} 6 \\ 18 \end{bmatrix}$, $A\vec{\mathbf{x}}_4 = \begin{bmatrix} -6 \\ -6 \end{bmatrix}$, $A\vec{\mathbf{x}}_5 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$

(b) $\vec{\mathbf{x}}_1$ is an eigenvector with eigenvalue $\lambda = 3$, $\vec{\mathbf{x}}_4$ and $\vec{\mathbf{x}}_5$ are eigenvectors with eigenvalue $\lambda = 2$.

(c) Every eigenvector corresponding to $\lambda = 3$ has the form $t\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

(d) Every eigenvector corresponding to $\lambda = 2$ has the form $t\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

2. (a) $\vec{\mathbf{u}}_2$ is an eigenvector with eigenvalue $\lambda = -5$, $\vec{\mathbf{u}}_3$ is an eigenvector with eigenvalue $\lambda = 1$

   (b) $\vec{\mathbf{v}}_1$ is an eigenvector with eigenvalue $\lambda = 5$, $\vec{\mathbf{v}}_4$ is an eigenvector with eigenvalue $\lambda = 0$

   (c) $\vec{\mathbf{w}}_1$ is an eigenvector with eigenvalue $\lambda = 1$, $\vec{\mathbf{w}}_2$ is an eigenvector with eigenvalue $\lambda = 4$, $\vec{\mathbf{w}}_4$ is an eigenvector with eigenvalue $\lambda = 2$

   (d) $\vec{\mathbf{u}}_1$ is an eigenvector with eigenvalue $\lambda = 1$, $\vec{\mathbf{u}}_3$ is an eigenvector with eigenvalue $\lambda = 0$, $\vec{\mathbf{u}}_4$ is an eigenvector with eigenvalue $\lambda = 3$

3. $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_5$ are eigenvectors with eigenvalue $\lambda = -1$, $\vec{\mathbf{v}}_3$ and $\vec{\mathbf{v}}_4$ are eigenvectors with eigenvalue $\lambda = 2$

4. $\vec{\mathbf{u}}_2$ and $\vec{\mathbf{u}}_4$ are independent eigenvectors with eigenvalue $\lambda = -1$, $\vec{\mathbf{u}}_5$ is an eigenvector with eigenvalue $\lambda = 8$.

5. For each of the following, any scalar multiples of the given vectors are also eigenvectors with the same respective eigenvalues.

   (a) $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = -1$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = 1$.

   (b) $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = 0$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = 1$.

   (c) $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = 0$, $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = 1$.

   (d) $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = -1$, any vector in the yz-plane is an eigenvector with with eigenvalue $\lambda = 1$. In particular, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ are independent eigenvectors with eigenvalues $\lambda = 1$.

   (e) $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ or any scalar multiple of it is an eigenvector with eigenvalue $\lambda = 1$. There are no other eigenvectors independent of that one.

   (f) $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = 0$, any vector in the xz-plane is an eigenvector with with eigenvalue $\lambda = 1$. In particular, $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ are independent eigenvectors with eigenvalues $\lambda = 1$.

   (g) $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue $\lambda = -1$, any vector in the plane $z = -y$ is an eigenvector with with eigenvalue $\lambda = 1$. In particular, $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ are independent eigenvectors with eigenvalues $\lambda = 1$.

Section 5.7 Solutions Back to 5.7 Exercises

1. $E_2 = \left\{ t \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right\}$

2. The basis eigenvector for any given eigenspace can be any scalar multiple of the vector given.

   (a) $\lambda_1 = 1,\ E_1 = \left\{ t \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\}, \qquad \lambda_2 = 5,\ E_2 = \left\{ t \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

   (b) $\lambda_1 = 1,\ E_1 = \left\{ t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}, \qquad \lambda_2 = 3,\ E_2 = \left\{ t \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

   (c) $\lambda_1 = 3,\ E_1 = \left\{ t \begin{bmatrix} -4 \\ 1 \end{bmatrix} \right\}, \qquad \lambda_2 = -2,\ E_2 = \left\{ t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

3. The basis eigenvector for any given eigenspace can be any scalar multiple of the vector given.

   (a) $\lambda_1 = 1,\ E_1 = \left\{ t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}, \qquad \lambda_2 = -1,\ E_2 = \left\{ t \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

   (b) $\lambda_1 = 1,\ E_1 = \left\{ t \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}, \qquad \lambda_2 = 0,\ E_2 = \left\{ t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

   (c) $\lambda_1 = 1,\ E_1 = \left\{ t \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \qquad \lambda_2 = -1,\ E_2 = \left\{ t \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$

4. $E_2 = \left\{ t \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$ and $E_3 = \left\{ t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

5. (a) The characteristic polynomial is $(1-\lambda)(\lambda^2 - 5\lambda + 6)$ or $(1-\lambda)(\lambda-2)(\lambda-3)$

   (b) $\lambda_1 = 1,\ \lambda_2 = 2,\ \lambda_3 = 3$

   (c), (d) $E_1 = \left\{ t \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\},\ E_2 = \left\{ t \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} \right\},\ E_3 = \left\{ t \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$

6. (a) The characteristic polynomial is $-\lambda^3 - \lambda^2 + 12\lambda$

   (b) The characteristic equation is $-\lambda^3 - \lambda^2 + 12\lambda = 0$, which can be factored to get $-\lambda(\lambda + 4)(\lambda - 3) = 0$, giving the eigenvalues $\lambda_1 = 0,\ \lambda_2 = -4,\ \lambda_3 = 3$

   (c) $E_1 = \left\{ t \begin{bmatrix} 1 \\ 6 \\ -13 \end{bmatrix} \right\},\ E_2 = \left\{ t \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right\},\ E_3 = \left\{ t \begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix} \right\}$

7. (a) The characteristic polynomial is $-\lambda^3 + 4\lambda^2 + 27\lambda - 90$

   (b) The factored characteristic equation is $-(\lambda-3)(\lambda+5)(\lambda+6) = 0$, giving the eigenvalues $\lambda_1 = 3,\ \lambda_2 = -5,\ \lambda_3 = -6$

   (c) $E_1 = \left\{ t \begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix} \right\},\ E_2 = \left\{ t \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right\},\ E_3 = \left\{ t \begin{bmatrix} 1 \\ 6 \\ 16 \end{bmatrix} \right\}$

8. (a) The characteristic polynomial is $-\lambda(1-\lambda)^2$

(b) $\lambda_1 = 0,\ \ \lambda_2 = 1$

(c) $E_1 = \left\{ t \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} \right\},\ \ E_2 = \left\{ s \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$

9. $\lambda_1 = 1,\ \ E_1 = \left\{ t \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \right\},\qquad \lambda_2 = -2,\ \ E_2 = \left\{ s \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} \right\}$

10. $\lambda_1 = 1,\ \ E_1 = \left\{ t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\},\qquad \lambda_2 = 2,\ \ E_2 = \left\{ t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

11. See all rotated stress elements below the other answers.

(a) $\lambda_1 = 54.1,\ \ \vec{\mathbf{x}}_1 = \begin{bmatrix} 40 \\ 24.1 \end{bmatrix},\ \ \lambda_2 = 25.9,\ \ \vec{\mathbf{x}}_2 = \begin{bmatrix} -40 \\ 4.1 \end{bmatrix},\ \ \theta = \tan^{-1}\dfrac{24.1}{40} = 31.1^\circ$

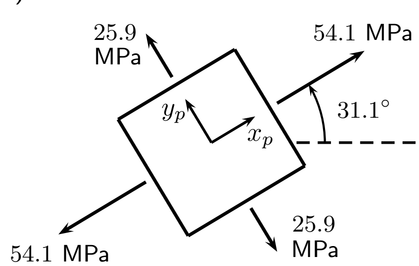
(b) $\lambda_1 = -79.4,\ \ \vec{\mathbf{x}}_1 = \begin{bmatrix} 45 \\ -19.4 \end{bmatrix},\ \ \lambda_2 = 44.4,\ \ \vec{\mathbf{x}}_2 = \begin{bmatrix} 45 \\ 104.4 \end{bmatrix},$

$\theta = \tan^{-1}\dfrac{-19.4}{45} = -23.3^\circ$

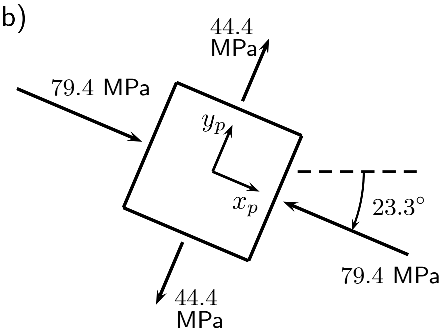
(c) $\lambda_1 = 83.0,\ \ \vec{\mathbf{x}}_1 = \begin{bmatrix} 20 \\ -3.0 \end{bmatrix},\ \ \lambda_2 = -53.0,\ \ \vec{\mathbf{x}}_2 = \begin{bmatrix} 20 \\ 133.0 \end{bmatrix},\ \ \theta = \tan^{-1}\dfrac{-3.0}{20} = -8.5^\circ$

(d) $\lambda_1 = -116.1,\ \ \vec{\mathbf{x}}_1 = \begin{bmatrix} 35 \\ 76.1 \end{bmatrix},\ \ \lambda_2 = -23.9,\ \ \vec{\mathbf{x}}_2 = \begin{bmatrix} -35 \\ 16.1 \end{bmatrix},\ \ \theta = \tan^{-1}\dfrac{76.1}{35} = 65.3^\circ$

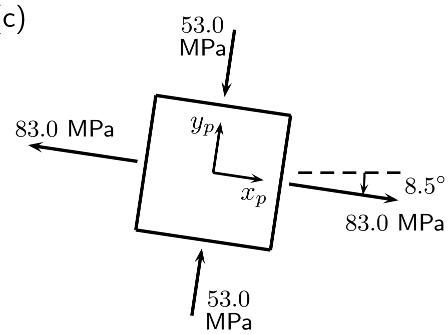
(a)



(b)



(c)



(d)