# Linear Algebra I

## Skills, Concepts and Applications

Gregg Waterman
Oregon Institute of Technology

# Contents


**1 Systems of Linear Equations** **1**
1.1 Linear Equations and Systems of Linear Equations . . . . . 2
1.2 Curve Fitting, Temperature Equilibrium and Electric Circuits . . . . . 7
1.3 Solving Systems of Linear Equations . . . . . 13
1.4 Solving With Matrices . . . . . 17
1.5 "When Things Go Wrong" . . . . . 25
1.6 Back to Applications . . . . . 31
1.7 Chapter 1 Exercises . . . . . 35

**2 Euclidean Space and Vectors** **40**
2.1 Euclidean Space . . . . . 41
2.2 Introduction to Vectors . . . . . 45
2.3 Addition and Scalar Multiplication of Vectors, Linear Combinations . . . . . 48
2.4 Linear Combination Form of a System . . . . . 54
2.5 Sets of Vectors . . . . . 58
2.6 Vector Equations of Lines and Planes . . . . . 62
2.7 Interpreting Solutions to Systems of Linear Equations . . . . . 68
2.8 The Dot Product of Vectors, Projections . . . . . 72

**3 Matrices and Vectors** **76**
3.1 Introduction to Matrices . . . . . 78
3.2 Matrix Times a Vector . . . . . 82
3.3 Actions of Matrices on Vectors: Transformations in $\mathbb{R}^2$ . . . . . 90
3.4 Multiplying Matrices . . . . . 94
3.5 Inverse Matrices . . . . . 102
3.6 Determinants and Systems of Equations . . . . . 108
3.7 Applications: Transformation Matrices, Graph Theory . . . . . 115

**4 Vector Spaces and Subspaces** **121**
4.1 Span of a Set of Vectors . . . . . 122
4.2 Closure of a Set Under an Operation . . . . . 128
4.3 Vector Spaces and Subspaces . . . . . 131
4.4 Column Space and Null Space of a Matrix . . . . . 138
4.5 Least Squares Solutions to Inconsistent Systems . . . . . 142
4.6 Linear Independence . . . . . 147
4.7 Bases of Subspaces, Dimension . . . . . 155
4.8 Bases for the Column Space and Null Space of a Matrix . . . . . 160
4.9 Solutions to Systems of Equations . . . . . 165
4.10 Chapter 4 Exercises . . . . . 168

**5 Linear Transformations** **170**
5.1 Transformations of Vectors . . . . . 171
5.2 Linear Transformations . . . . . 176
5.3 Linear Transformations and Matrices . . . . . 183
5.4 Compositions of Transformations . . . . . 186
5.5 Transformations and Homogeneous Coordinates . . . . . 190
5.6 An Introduction to Eigenvalues and Eigenvectors . . . . . 197
5.7 Finding Eigenvalues and Eigenvectors . . . . . 202
5.8 Diagonalization of Matrices . . . . . 211

5.9 Solving Systems of Differential Equations . . . . . . . . . . . . . . . . . . . . . . . . . 214

**A Index of Symbols** **217**

**B Solutions to Exercises** **219**

B.1 Chapter 1 Solutions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 219
B.2 Chapter 2 Solutions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 221
B.3 Chapter 3 Solutions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 229
B.4 Chapter 4 Solutions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 235
B.5 Chapter 5 Solutions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 241

# 1 Systems of Linear Equations

**Learning Outcome:**

1. Solve systems of linear equations using Gaussian elimination, use systems of linear equations to solve problems.

   **Performance Criteria:**

   (a) Determine whether an equation in $n$ unknowns is linear.

   (b) Set up a system of linear equations to find coefficients of a line or polynomial through a given set of points, or to model flow in a network or equilibrium temperatures in a solid object.

   (c) Determine whether an $n$-tuple is a solution to a linear equation or a system of linear equations.

   (d) Solve a system of two linear equations by the addition method.

   (e) Give the coefficient matrix and augmented matrix for a system of equations.

   (f) Determine whether a matrix is in row-echelon form. Perform, by hand, elementary row operations to reduce a matrix to row-echelon form.

   (g) Determine whether a matrix is in reduced row-echelon form. Use technology to reduce a matrix to reduced row-echelon form.

   (h) For a system of equations having a unique solution, determine the solution from either the row-echelon form or reduced row-echelon form of the augmented matrix for the system.

   (i) Use a calculator to solve a system of linear equations having a unique solution.

   (j) Given the row-echelon or reduced row-echelon form of an augmented matrix for a system of equations, determine the leading variables and free variables of the system.

   (k) Given the row-echelon or reduced row-echelon form for a system of equations:

   - Determine whether the system has a unique solution, and give the solution if it does.
   - If the system does not have a unique solution, determine whether it is inconsistent (no solution) or dependent (infinitely many solutions).
   - If the system is dependent, give the general form of a solution and give some particular solutions.

   (l) Use systems of equations to solve network problems.

## 1.1 Linear Equations and Systems of Linear Equations

**Performance Criteria:**

1. (a) Determine whether an equation in $n$ unknowns is linear.
   (b) Set up a system of linear equations to find coefficients of a line or polynomial through a given set of points, or to model flow in a network or equilibrium temperatures in a solid object.
   (c) Determine whether an $n$-tuple is a solution to a linear equation or a system of linear equations.

### Linear Equations and Their Solutions

It is natural to begin our study of linear algebra with the process of solving systems of linear equations, and applications of such systems.

Definition 1.1.1: A **linear equation** in $n$ unknowns is an equation that can be put in the form

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b, \tag{1}$$

where $a_1, a_2, ..., a_n$ and $b$ are *known* constants and $x_1, x_2, ... , x_n$ are unknown values. A **solution to a linear equation** is a collection of values for the unknowns that makes the equation true.

◇ **Example 1.1(a):** Which of the equations

$$y = -\tfrac{2}{3}x + 4 \qquad y = -16t^2 + 61t + 7 \qquad 5.3x + 7.2y + 1.4z = 16.9$$

$$a_{41}x_1 + a_{42}x_2 + \cdots + a_{4n}x_n = b_4,$$

where $a_{11}$, $a_{12}, ..., a_{1n}, b_1$ are all known numbers, are linear equations?

**Solution:** The first equation can be rewritten as $\frac{2}{3}x + y = 4$, so it is a linear equation. The second equation can be written $-16t^2 + 61t - y = -4$, but the $t^2$ prevents this from being a linear equation. The third and fourth equations are in exactly the form (1), so they are linear.

A few comments are in order:

- For the third equation above we see that

$$5.3(1) + 7.2(2) + 1.4(-2) = 16.9,$$

so $x = 1, y = 2$ and $z = -2$ is a solution to $5.3x + 7.2y + 1.4z = 16.9$. To save writing we usually write such a solution as $(1, 2, -2)$, a form you are likely familiar with.

- The equation $y = -\frac{2}{3}x + 4$ can also be rewritten as $2x + 3y = 12$ instead of $\frac{2}{3}x + y = 4$. An $(x, y)$ pair that is a solution to any one of the forms is also a solution to the other two (and any pair that is *NOT* a solution to any one of them will not be a solution to the other two either). We can multiply or divide both sides of a linear equation by a value in order to make it easier to work with, if we wish.

- Although you may have used $x$ and $y$, or $x$, $y$ and $z$ as the unknown quantities in the past, like in the third equation above, we will often use $x_1$, $x_2$, ..., $x_n$ instead. Thus the third equation could be written

$$5.3x_1 + 7.2x_2 + 1.4x_3 = 15.9,$$

  which is equivalent to the fourth equation with $a_{41} = 5.3$, $a_{42} = 7.2$, $a_{43} = 1.4$ and $b_4 = 15.9$. One obvious advantage to using the letter $a$ for all of the numbers is that we don't have to fret about what letters to use, and there is no danger of running out of letters! You will eventually see that there is also a very good mathematical reason for using just $x$ (or some other single letter), with subscripts denoting different values.

It is important that you easily recognize the form (1) from the definition of a linear equation. Soon we will be interested in similar equations, but of the form

$$number \cdot vector + number \cdot vector + \cdots + number \cdot vector = vector.$$

We now move on to the concept that forms the beginning of our study of linear algebra:

> Definition 1.1.2: A **system of linear equations** is a set of linear equations containing the same unknowns. (Not every equation needs to contain every unknown.) A **solution to a system of linear equations** is a collection of values for the unknowns that makes every equation of the system true.

◇ **Example 1.1(b):** Which of the following are systems of linear equations?

$$\begin{aligned} x + 3y - 2z &= -4 \\ 3x + 7y + z &= 4 \\ -2x + y + 7z &= 7 \end{aligned} \qquad \begin{aligned} x + y^2 &= 3 \\ x^2 + y^2 &= 5 \end{aligned} \qquad \begin{aligned} 4t_1 - t_2 - t_3 \phantom{- t_4} &= 108 \\ -t_1 + 4t_2 \phantom{- t_3} - t_4 &= 106 \\ -t_1 + 4t_3 - t_4 &= 94 \\ -t_2 - t_3 + 4t_4 &= 96 \end{aligned}$$

**Solution:** The first and third systems are systems of linear equations, the second is not. The second is a system of *nonlinear* equations. One can verify that $(3, -1, 2)$ is a solution to the first system of equations.

Here we note the following:

- The first system in the previous example is a system of three equations in three unknowns. We will spend a lot of time with such systems because they exhibit just about everything that we would like to see but are small enough to be manageable to work with. As noted before, we will often use $x_1$, $x_2$ and $x_3$ instead of $x$, $y$ and $z$ for the unknowns.

- The numbers that the unknowns are multiplied by are **coefficients** of the system. It is customary to get the coefficient/unknown terms on the left, and the numbers not multiplying an unknown on the right, as shown in the first and third (and second, for that matter) examples above. The numbers without unknowns are often referred to as the "right hand sides."

- One should note carefully the coefficients of the third system and how they are arranged, as shown to the right. Later we will put some brackets around such an array and call it a **matrix**. The fours are on what we will call the **diagonal** of the matrix. (It seems that there is another diagonal with zeros on it, but that diagonal, from lower left to upper right, has no real significance. We therefore make no special note of what is going on there.) In addition to noting the fours on the diagonal, we also need to make special note of the way that the zeros and negative ones are arranged symmetrically across the diagonal. That sort of pattern is commonly encountered in physical applications of systems of linear equations.

$$\begin{array}{rrrr} 4 & -1 & -1 & 0 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 4 & -1 \\ 0 & -1 & -1 & 4 \end{array}$$

- When discussing a system of linear equations in general, we often use the following notation, given for a system of $m$ equations in $n$ unknowns:

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array} \qquad (2)$$

  Here the $a_{ij}$s are the coefficients, with the first subscript of each giving the equation and the second subscript giving which unknown $x_1$, $x_2$, ... it is with.

Systems of equations arise naturally in many engineering applications (as well as applications in other areas like business). Some of the uses of systems of equations that we'll work with are

- analysis of electric circuits
- equilibrium distribution of heat in solid materials
- stress and strain in solid materials
- linear regression (least-squares approximation)

You will begin exploring some of these applications in the exercises for this section and the next. There are applications of other linear algebra concepts that we'll see later, such as

- robotics and computer graphics
- sports and internet search rankings
- air travel routing
- signal processing

**Section 1.1 Exercises** To Solutions

1. Which of the following equations are linear equations?

   (a) $4x^2 + 3y^2 = 5$ (b) $\dfrac{t_1 + t_2 + 83}{4} = t_3$ (c) $3x_1 - x_2 + 4x_3 = x_2$

   (d) $4.3 = 1.7m + b$ (e) $y = \sqrt{10 - x}$ (f) $\dfrac{5}{x} + \dfrac{2}{y} = 7$

2. Which of the systems of equations below are linear?

$$\begin{aligned} x_1 - x_2 + x_3 &= 3 \\ 2x_1 - x_2 + 4x_3 &= 7 \\ 3x_1 - 5x_2 - x_3 &= 7 \end{aligned} \qquad \begin{aligned} 3x + y - 2z &= -4 \\ 5x + 4z &= 3 \\ x - y + 2z &= 0 \end{aligned} \qquad \begin{aligned} x^2 - y &= 3 \\ x - y &= 1 \end{aligned}$$

3. (a) Determine which of the following are solutions to the first system of equations in the previous exercise: $(5, -2, 4)$, $(-2, -3, 2)$, $(7, 3, 1)$

   (b) Determine which of the following are solutions to the second system of equations in the previous exercise: $(3, -19, -3)$, $(-1, 3, 2)$, $(5, -2, 4)$

   (c) Determine which of the following are solutions to the third system of equations in the previous exercise: $(2, 1)$, $(3, 5)$, $(-1, -2)$

4. Consider the equation $y = ax^3 + bx^2 + cx + d$, representing a third degree polynomial.

   (a) Substitute the value $-2$ for $x$ and $5$ for $y$ into $y = ax^3 + bx^2 + cx + d$, and simplify the result. Is the resulting equation linear?

   (b) Substitute the values $a = 7$, $b = -2$, $c = -5$ and $d = 1$ into $y = ax^3 + bx^2 + cx + d$. Is the resulting equation linear?

5. It turns out that there is exactly one third degree polynomial with equation $y = ax^3 + bx^2 + cx + d$ whose graph goes through the four points

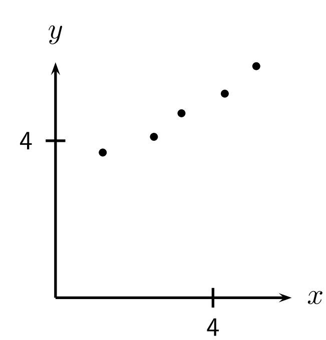
$$(-2, 5) \qquad (-1, 2) \qquad (1, 3) \qquad (3, 0)$$

   Substitute each of those pairs into the equation (one pair at a time) to obtain four equations in the four unknowns $a$, $b$, $c$ and $d$. Give your final system in the form (2). Once we know how to solve such a system we can determine the values of $a$, $b$, $c$ and $d$.

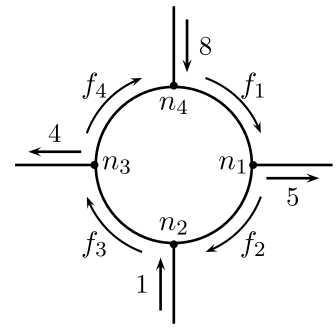
6. The graph to the right shows a plot of the five points with coordinates

$$(1.2, 3.7) \quad (2.5, 4.1) \quad (3.2, 4.7) \quad (4.3, 5.2) \quad (5.1, 5.9).$$

   You can see that there is no line containing them, but the points are arranged somewhat linearly. In many applications it is desirable to find the line that comes closest (in a sense to be described later) to passing through all of the points. Substitute each of the points individually into the equation $y = mx + b$ to obtain a system of five equations. (What are the unknowns?) Give the system in the form (2).



7. In many engineering applications we are interested in flow through a **network**. The flow could consist of water or air through pipes or ductwork (mechanical engineering), electrons through a circuit (or current, electrical engineering) or automobiles on a raodway (civil engineering). Such a network can be modeled by a set of **nodes** or **vertices** connected by arcs or line segments, typically called **edges**. The guiding principle for most such networks is simple: *the flow into any node must equal the flow out*. The network to the right represents a traffic circle. The numbers next to each of the paths leading into or out of the circle are the net flows in the directions of the arrows, in vehicles per minute, during the first part of lunch hour. The unknowns $f_1$, $f_2$, $f_3$ and $f_4$ represent the flows in the corresponding arcs of the traffic circle.




(a) The nodes of the network have been labeled $n_1$, $n_2$, $n_3$ and $n_4$. At node $n_1$, "flow in equals flow out" gives us $f_1 = f_2 + 5$. Rearranging this to get the unknowns on one side with $f_1$ positive, we get $f_1 - f_2 = 5$. Repeat at nodes $n_2$, $n_3$ and $n_4$, with the corresponding flows being positive in each case. That is, $f_2$ should be positive in the equation obtained at $n_2$, and so on. Give the resulting system of equations.

(b) Determine which of the following "4-tuples" (an $n$-**tuple** is an ordered "tuple" or collection of values, separated by commas) are solutions to the system that you obtained. *Note that this illustrates that a system can have more than one solution.*

$$(12, 7, 8, 4) \qquad (7, 2, 3, -1) \qquad (10, 5, 6, 2) \qquad (9, 4, 3, 1)$$

(c) Which of the 4-tuples that you found to be a solution to your system would cause a problem with the traffic circle? Explain.

(d) Suppose that $f_3 = 9$. Just by looking at the traffic circle and using the "flow in equals flow out, determine the other three flows.

## 1.2 Curve Fitting, Temperature Equilibrium and Electric Circuits

**Performance Criterion:**

1. (b) Set up a system of linear equations to find coefficients of a line or polynomial through a given set of points, or to model flow in a network or equilibrium temperatures in a solid object.

### Curve Fitting

Curve fitting refers to the process of finding a polynomial function of "minimal degree" whose graph contains some given points. We all know that any two distinct points (that is, points that are not the same) in $\mathbb{R}^2$ have exactly one line through them. In a previous course you should have learned how to find the equation of that line in the following manner. Suppose that we wish to find the equation of the line through the points $(2,3)$ and $(6,1)$. We know that the equation of a line looks like $y = mx + b$, where $m$ and $b$ are to be determined. $m$ is the slope, which can be found by $m = \dfrac{3-1}{2-6} = \dfrac{2}{-4} = -\frac{1}{2}$. Therefore the equation of our line looks like $y = -\frac{1}{2}x + b$. To find $b$ we simply substitute either of the given ordered pairs into our equation (the fact that both pairs lie on the line means that either pair is a solution to the equation) and solve for $b$: $3 = -\frac{1}{2}(2) + b \implies b = 4$. The equation of the line through $(2,3)$ and $(6,1)$ is then $y = -\frac{1}{2}x + 4$.

We will now solve the same problem in a different way. A student should understand that whenever a new approach to a familiar exercise is taken, there is something to be gained by it. Usually the new method is in some way more powerful, and allows the solving of additional problems. This will be the case with the following example, which uses a process you should have seen in a previous course, and that we will review in detail in the next section.

◇ **Example 1.2(a):** Find the equation of the line containing the points $(6,1)$ and $(2,3)$.

**Solution:** We are again trying to find the two constants $m$ and $b$ of the equation $y = mx + b$. Here we substitute the values of $x$ and $y$ from each of the two points into the equation $y = mx + b$ (separately, of course!) to get two equations in the two unknowns $m$ and $b$. The resulting system is then solved for $m$, then $b$.

$$\begin{aligned} 1 &= 6m + b \\ 3 &= 2m + b \end{aligned} \quad\Longrightarrow\quad \begin{aligned} 1 &= 6m + b \\ -(3 &= 2m + b) \\ \hline -2 &= 4m \\ -\tfrac{1}{2} &= m \end{aligned} \quad\Longrightarrow\quad \begin{aligned} 3 &= 2(-\tfrac{1}{2}) + b \\ 3 &= -1 + b \\ 4 &= b \end{aligned}$$

The equation of the line containing $(6,1)$ and $(2,3)$ is $y = -\frac{1}{2}x + 4$.

---

The process of solving systems of two linear equations in two unknowns will be covered in more detail in the next section.

The equation of a line is considered to be a first-degree polynomial, since the power of $x$ in $y = mx+b$ is one. Note that when we have two points in the $xy$-plane we can find a first-degree polynomial whose graph contains the points, and there is only one such line. Similarly, when given three points we can find a second-degree polynomial (quadratic polynomial) whose graph contains the three points. In general,

> THEOREM 1.2.1: Given $n$ points in the plane such that (a) no two of them have the same $x$-coordinate and (b) they are not collinear, we can find a *unique* polynomial function of degree $n-1$ whose graph contains the $n$ points.

Often in mathematics we are looking for some object (solution) and we wish to be certain that such an object exists. In addition, it is generally preferable that *only one* such object exists. We refer to the first of these wishes as "existence," and the second is "uniqueness." If we have, for example, four points meeting the two conditions of the above theorem, there would be infinitely many fourth degree polynomials whose graphs would contain them, and the same would be true for fifth degree, sixth degree, and so on. Additionally, a set of four points meeting the above conditions will likely *NOT* not have a polynomial of degree two whose graph passes through all of them. But the theorem guarantees us that there is one, and only one, third degree polynomial whose graph contains the four points. In Exercise 3 of the previous section you saw how to construct a system of linear equations whose solution gives us the coefficients of the third degree polynomial whose graph contains four given points. In Example 1.4(e) we'll see how to find such a polynomial, from start to finish.

**Temperature Equilibrium**

Consider the following hypothetical situation: We have a plate of metal that is perfectly insulated on both of its faces so that no heat can get in or out of the faces. Each point on the edge (which we will call the **boundary**), however, is held at a constant temperature (constant at that point, but possibly differing from point to point). The temperatures at points on the boundary affect the temperatures at interior points. If the plate is left alone for a long time ("infinitely long"), the temperature at each point in the interior of the plate will reach a constant temperature, called the "equilibrium temperature." This equilibrium temperature at any given interior point is a **weighted average** of the temperatures at all the boundary points, with temperatures at closer boundary points being weighted more heavily in the average than the temperatures at boundary points that are farther away.

The task of trying to determine those interior temperatures based on the edge temperatures is a famous problem of applied mathematics, called the **Dirichlet problem** (pronounced "dir-i-shlay"). Finding the exact solution involves methods beyond the scope of this course, but we will use systems of equations to solve the problem "numerically," which means to approximate the exact solution, usually by some non-calculus method. The key to solving the Dirichlet problem is the following:

> THEOREM 1.2.2: **Mean Value Property**
>
> The equilibrium temperature at any interior point $P$ is the average of the temperatures of all interior points on *any* circle centered at $P$.

We will solve what are called **discrete** versions of the Dirichlet problem, which means that we only know the temperatures at a finite number of points on the boundary of our metal plate, and we will only find the equilibrium temperatures at a finite number of the interior points. These finite points, both on the boundary and in the interior, are usually evenly spaced on a rectangular grid. Consider the plate shown in Figure 1.2(a) on the next page, with boundary temperatures known at the indicated points. We can then construct a square grid in the interior of the plate, as shown in Figure 1.2(b). The unknown temperatures at the **mesh points** of the grid are denoted by $t_1$, $t_2$, $t_3$ and $t_4$, as shown in Figure 1.2(b). By the mean value property, the temperature $t_1$ is the average of the temperatures at all points on the circle shown in Figure 1.2(c). Such an average is obtained by an integral, but in our case we will simply average the temperatures at the four boundary and mesh points that are on the circle. This gives us the equation

$$t_1 = \frac{61 + 68 + t_2 + t_3}{4}.$$

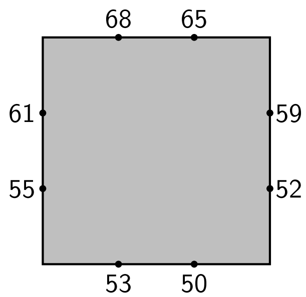



Figure 1.2(a)

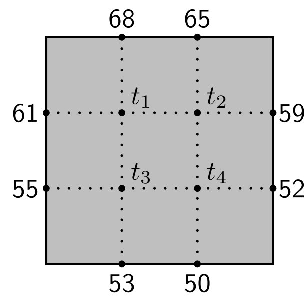



Figure 1.2(b)

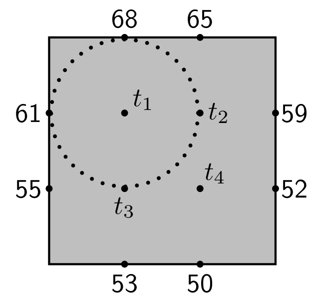
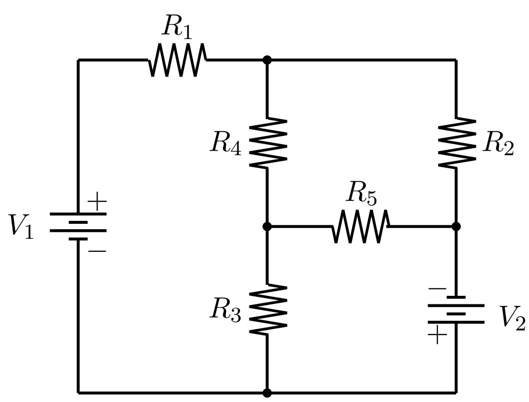


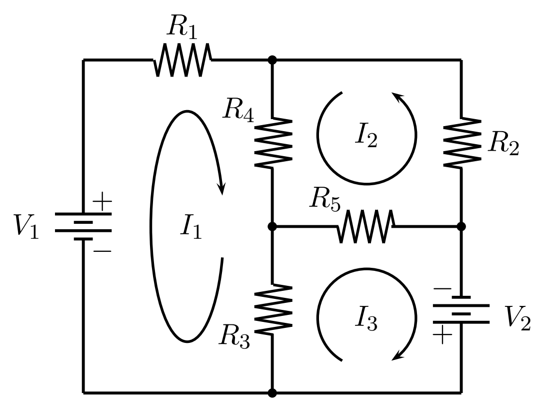

Figure 1.2(c)

We can then find three more such equations for circles centered at the mesh points with temperatures $t_2$, $t_3$ and $t_4$. If we then multiply both sides of each equation by four, combine the two known numerical values and get all of the unknowns on one side of each equation, we can obtain a system of linear equations in the standard form. You will do this in the exercises at the end of the section.

**Electric Circuits**

We could spend a great deal of time that we don't have on electric circuits, so here we'll just learn one method to get a system of linear equations modeling a circuit with constant **voltage sources** (batteries) and **resistors**. An example of such a circuit is shown below and to the left. The lines indicate wires that are connected at the dots, which are called **principal nodes**, and the portion of the circuit between two principle nodes we'll call a **branch**. A set of branches that can be placed end to end to get from one node back to itself is called a **loop**.




Each "zig-zag" is a resistor and the parallel long and short lines with $+$ and $-$ at either side are batteries. Each resistor has a characteristic called its **resistance**, which is measured in **ohms**. Similarly, each battery has a **voltage**, measured in ... **volts**! We will let $V_1$ and $V_2$ represent both the voltage sources and their voltages, and $R_1$ through $R_5$ will represent both the indicated resistors and their resistances.

The voltage sources cause something called **current** to flow in the circuit. Intuitively, we can think of the voltage sources as "pumps" pushing current through the wires, like pushing water through pipes. The resistors "resist" the flow of current. *Our objective is to find the current in each branch of the circuit.* To find the current in each branch we will proceed as follows:

1) Establish a clockwise or counterclockwise direction of current in each loop. If there is a voltage source in a loop, establish the current in the direction from the negative side to the positive side. If there is no voltage source, the current can be in either direction that you wish. (If you choose the "wrong" direction you will simply obtain a negative value for the current.) The diagram above and to the right shows currents established for each loop - we will use $I$ for current.

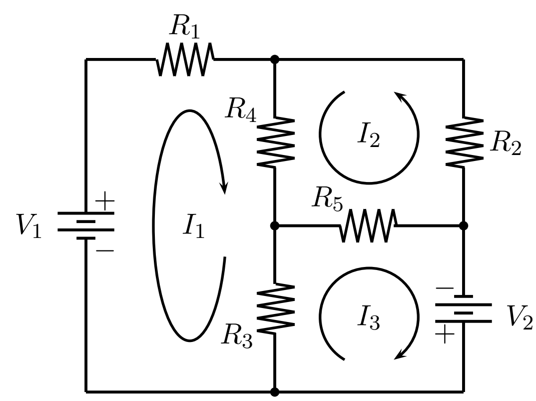
2) The "voltage drop" across each resistor is given by **Ohm's Law**, $V = IR$. **Kirchoff's Voltage Law** then tells us that the voltage supplied in a loop is equal to the sum of the voltage drops across each of the resistors in the loop. *When working in a loop and calculating the voltage drop across a resistor shared with another loop, the current used is the one for the loop under consideration plus or minus the current from the adjacent loop, depending on whether that current is going the same, or the opposite, direction as the current in the loop under consideration.* Write an equation for each loop based on Kirchoff's Voltage Law. If there is no voltage source in a loop, the voltage supplied is zero.

3) Get each equation in the form $aI_1 + bI_2 + cI_3 = V_k$, where $k$ is the loop the equation was obtained from.

4) Solve the system of equations.

In Section 1.4 we'll see how to solve such systems; for now we will only complete steps 1, 2 and 3 above.

◇ **Example 1.2(b):** Use the steps above to obtain a system of three equations that models the circuit shown below and to the right.




**Solution:** For the loop with current $I_1$ the voltage supplied is $V_1$. Going around the loop from the battery the voltage drops are

$$I_1R_1, \qquad (I_1 + I_2)R_4, \qquad (I_1 - I_3)R_3.$$

Kirchoff's Voltage Law then gives us

$$I_1R_1 + (I_1 + I_2)R_4 + (I_1 - I_3)R_3 = V_1.$$

The equations for the other two loops are

$$I_2R_2 + (I_2 + I_1)R_4 + (I_2 + I_3)R_5 = 0$$

and

$$(I_3 - I_1)R_3 + (I_3 + I_2)R_5 = V_2.$$

Putting each of our three equations in the form $aI_1 + bI_2 + cI_3 = V_k$ gives us

$$\begin{array}{rcrcrcl}
(R_1 + R_3 + R_4)I_1 & + & R_4I_2 & - & R_3I_3 & = & V_1 \\
R_4I_1 & + & (R_2 + R_4 + R_5)I_2 & + & R_5I_3 & = & 0 \\
-\ R_3I_1 & + & R_5I_2 & + & (R_3 + R_5)I_3 & = & V_2
\end{array}$$

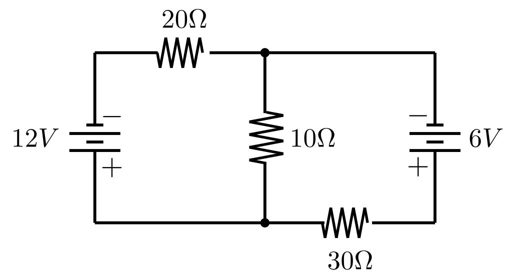
It is worth noting the array of coefficients of the three unknowns $I_1$, $I_2$ and $I_3$:

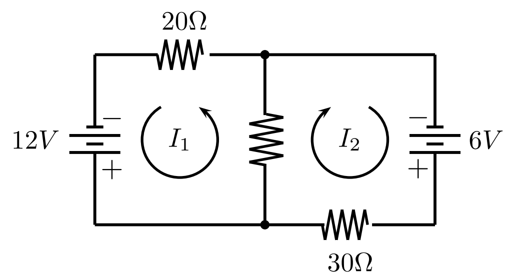
$$\begin{array}{ccc} R_1+R_3+R_4 & R_4 & -R_3 \\ R_4 & R_2+R_4+R_5 & R_5 \\ -R_3 & R_5 & R_3+R_5 \end{array}$$

Once again we see symmetry across the diagonal!

Let's do another example with numerical values for the voltage and resistances:

◇ **Example 1.2(c):** Find a system of equations that models the circuit below and to the right.




**Solution:** Here we establish the current $I_1$ in a counterclockwise direction in the left loop, and $I_2$ in a clockwise direction in the right loop, as shown in the lower picture to the right. For the left loop we get the equation

$$(I_1+I_2)10+20I_1=12$$

and, from the right loop,

$$30I_2+(I_2+I_1)10=6.$$

Distributing the resistances and regrouping gives us the system

$$\begin{aligned} 30I_1+10I_2 &= 12 \\ 10I_1+40I_2 &= 6 \end{aligned}$$

---

**Section 1.2 Exercises** To Solutions

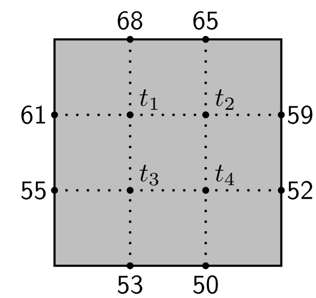
1. Consider the four points $(-1,3)$, $(1,5)$, $(2,4)$ and $(4,-1)$. By Theorem 1.2.1, there is a unique third degree polynomial of the form
$$y=a+bx+cx^2+dx^3 \qquad (1)$$
whose graph contains those four points.

    (a) Substitute the $x$ and $y$ values from the first ordered pair into (1) and rearrange the resulting equation so that it has all of the unknowns on the left and a number on the right, like all of the linear equations we have worked with so far.

    (b) Repeat (a) for the other 3 ordered pairs, and give the system of equations whose solution is the four coefficients $a$, $b$, $c$ and $d$.

2. Give a system of equations that can be solved to find the values of $a$, $b$ and $c$ for the quadratic polynomial $y=ax^2+bx+c$ whose graph is the parabola passing through the points $(-1,-4)$, $(1,1)$ and $(3,0)$.

3. To the right is a diagram of the metal plate described in the discussion of the mean value property for temperature equilibrium. In this exercise you will set up a system of equations whose solution gives the unknown temperatures $t_1$, $t_2$, $t_3$ and $t_4$ at the four interior points.



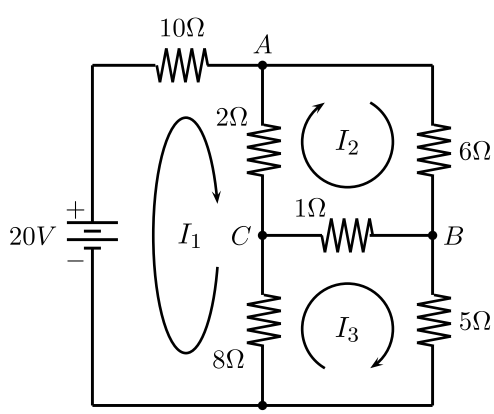

   (a) Applying the mean value property at the mesh point with temperature $t_1$ gave us the equation $t_1 = \dfrac{61+68+t_2+t_3}{4}$. Multiply both sides by four to eliminate the fraction, subtract $t_2$ and $t_3$ from both sides and add 61 and 68. You should end up with an equation of the form $at_1 + bt_2 + ct_3 + dt_4 = e$, where $d = 0$.

   (b) Follow a similar process for each of the other three interior mesh points in order to obtain a system of four equations in the four unknowns $t_1$, $t_2$, $t_3$ and $t_4$. In each equation one of $a$, $b$, $c$ or $d$ will be zero, so each equation will actually only contain three of $t_1$, $t_2$, $t_3$ and $t_4$.

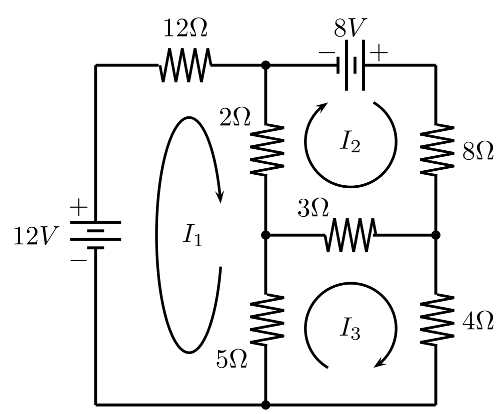
   (c) Although we can put the four equations in any order we want, arrange them so that the coefficients of four are along the diagonal of the left side, as we saw in third system of Example 1.1(b). Be sure that the remaining coefficients are symmetric about the diagonal. If they are not, find and correct your error.

4. (a) Give the system of equations modeling the circuit below and to the left.

   (b) Give the system of equations modeling the circuit below and to the right.




Exercise 4(a)




Exercise 4(b)

   (c) The solution to the circuit for Exercise 4(a) is $I_1 = 1.36$ amperes, $I_2 = 0.39$ amperes and $I_3 = 0.81$ amperes. Given this information, what is the current in the branch from point $A$ to point $C$? (Note that it is $I_1$ and $I_2$ combined, with the direction of each taken into account.) Does the current flow from $A$ to $C$, or from $C$ to $A$?

   (d) Again considering the circuit for Exercise 4(a) with the current values given above, what is the current in the branch from point $B$ to point $C$? Does the current flow from $B$ to $C$, or from $C$ to $B$?

## 1.3 Solving Systems of Linear Equations

**Performance Criteria:**

1. (d) Solve a system of two linear equations by the addition method.

Now that we know some applications of systems of equations, and how to set up systems for an application, it is time we learn how to solve a system. In this section we remember how to solve a system of two equations in two unknowns by the addition method, and extend the method to a system of three equations in three unknowns. Then in Section 1.4, we will introduce the method that we will use throughout the rest of the course.

Consider the system $\begin{aligned} x - 3y &= 6 \\ -2x + 5y &= -5 \end{aligned}$ of linear equations. In this case a solution to the system is an **ordered pair** $(x, y)$ that makes *both* equations true. In the past you should have learned two methods for solving such systems, the **addition method** and the **substitution method**. The method we want to focus on is the addition method. In this case we could multiply the first equation by two and add the resulting equation to the second. The result is $\begin{aligned} x - 3y &= 6 \\ -y &= 7 \end{aligned}$; from this we can see that $y = -7$. This value is then substituted into the first equation to get $x = -15$.

Sometimes we have to do something a little more complicated:

◇ **Example 1.3(a):** Solve the system $\begin{aligned} 2x - 4y &= 18 \\ 3x + 5y &= 5 \end{aligned}$ using the addition method.

**Solution:** Here we can eliminate $x$ by multiplying the first equation by 3 and the second by $-2$, then adding:

$$\begin{aligned} 2x - 4y &= 18 \\ 3x + 5y &= 5 \end{aligned} \quad \Longrightarrow \quad \begin{aligned} 6x - 12y &= 54 \\ -6x - 10y &= -10 \\ \hline -22y &= 44 \\ y &= -2 \end{aligned}$$

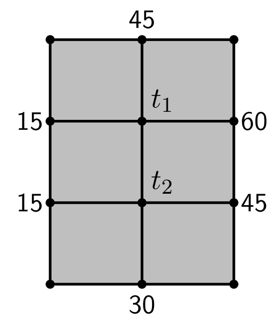
Now we can substitute this value of $y$ back into either equation to find $x$:

$$\begin{aligned} 2x - 4(-2) &= 18 \\ 2x + 8 &= 18 \\ 2x &= 10 \\ x &= 5 \end{aligned}$$

The solution to the system is then $x = 5$, $y = -2$, which we usually write as the ordered pair $(5, -2)$. It can be easily verified that this pair is a solution to both equations.

Let's now solve an applied problem that uses a system of two equations.

◇ **Example 1.3(b):** The temperatures (in degrees Fahrenheit) at six points on the edge of a rectangular plate are shown to the right. Assuming that the temperatures in the plate have reached equilibrium, find the interior temperatures $t_1$ and $t_2$ at their indicated "mesh points."




**Solution:** The discrete version of the mean value property tells us that the equilibrium temperature at any interior point of the mesh is the average of the four adjacent points. This gives us the two equations

$$t_1 = \frac{15+45+60+t_2}{4} \qquad \text{and} \qquad t_2 = \frac{15+t_1+45+30}{4}$$

If we multiply both sides of each equation by four, combine the constants and get the $t_1$ and $t_2$ terms on the left side we get the system of equations $\begin{aligned} 4t_1 - t_2 &= 120 \\ -t_1 + 4t_2 &= 90 \end{aligned}$. Multiplying the first equation by four and adding the result to the second gives us $15t_1 = 570$, from which we find that $t_1 = 38$. Substituting that into either equation and solving for $t_2$ gives $t_2 = 32$. These values can easily be shown to verify our discrete mean value property:

$$\frac{15+45+60+t_2}{4} = \frac{15+45+60+32}{4} = 38 = t_1\,,$$

$$\frac{15+t_1+45+30}{4} = \frac{15+38+45+30}{4} = 32 = t_2$$

---

**Geometric Interpretation**

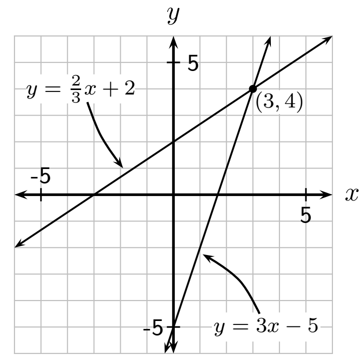
At the start of this section we saw that the system $\begin{aligned} x - 3y &= 6 \\ -2x + 5y &= -5 \end{aligned}$ has the solution $(-15,-7)$. You should be aware that if we graph the equation $x-3y=6$ we get a line. Technically speaking, what we have graphed is the **solution set**, the set of all pairs $(x,y)$ that make the equation true. *Any pair $(x,y)$ of numbers that makes the equation true is on the line, and the $(x,y)$ representing any point on the line will make the equation true.* If we plot the solution sets of both equations in the system $\begin{aligned} x - 3y &= 6 \\ -2x + 5y &= -5 \end{aligned}$ together in the coordinate plane we will get two lines. Since $(-15,-7)$ is a solution to both equations, *the two lines cross at the point with those coordinates!* We could use this idea to (somewhat inefficiently and possibly inaccurately) solve a system of two equations in two unknowns:

◇ **Example 1.3(c):** Solve the system $\begin{aligned} 2x - 3y &= -6 \\ 3x - y &= 5 \end{aligned}$ graphically.

**Solution:** We begin by solving each of the equations for $y$; this will give us the equations in $y = mx+b$ form, for easy graphing. The results are

$$y = \tfrac{2}{3}x + 2 \qquad \text{and} \qquad y = 3x - 5$$

If we graph these two equations on the same graph, we get the picture to the right. Note that the two lines cross at the point $(3,4)$, so the solution to the system of equations is $(3,4)$, or $x=3,\ y=4$.




It is possible that two lines in the standard two-dimensional plane might be parallel; in that case a system consisting of the two equations representing those lines will have no solution. It is also possible that two equations might actually represent the same line, in which case the system consisting of those two equations will have infinitely many solutions. Investigation of those two cases will lead us to more complex considerations that we will avoid for now.

### A System of Equations in Three Unknowns

The previous examples were two linear equations with two unknowns. Now we consider the following system of *three linear equations in three unknowns*.

$$\begin{aligned} x+3y-2z &= -4 \\ 3x+7y+z &= 4 \\ -2x+y+7z &= 7 \end{aligned} \tag{1}$$

We can use the addition method here as well; first we multiply the first equation by negative three and add it to the second. We then multiply the first equation by two and add it to the third. This eliminates the unknown $x$ from the second and third equations, giving the second system of equations shown below. We can then add $\frac{7}{2}$ times the second equation to the third to obtain a new third equation in which the unknown $y$ has been eliminated. This "final" system of equations is shown to the right below.

$$\begin{aligned} x+3y-2z &= -4 \\ 3x+7y+z &= 4 \\ -2x+y+7z &= 7 \end{aligned} \quad\Longrightarrow\quad \begin{aligned} x+3y-2z &= -4 \\ -2y+7z &= 16 \\ 7y+3z &= -1 \end{aligned} \quad\Longrightarrow\quad \begin{aligned} x+3y-2z &= -4 \\ -2y+7z &= 16 \\ \tfrac{55}{2}z &= 55 \end{aligned} \tag{2}$$

Above we have three different systems, each with three equations. The three systems are *equivalent*, meaning that a solution to any one of them is also a solution to the other two. The point of the above process is to obtain a system that is equivalent to the original but easier to solve. We can see that the solution to the last equation of the third system is $z=2$. That result is then substituted into the second equation in the last system to get $y=-1$. Finally, we substitute the values of $y$ and $z$ into the first equation to get $x=3$. The solution to the system is then the **ordered triple** $(3,-1,2)$. The process of finding the last unknown first, substituting it to find the next to last, and so on, is called **back substitution**. The word "back" here means that we find the last unknown (in the order they appear in the equations) first, then the next to last, and so on.

You might note that we could eliminate any of the three unknowns from any two equations, then use the addition method with those two to eliminate another variable. However, we will always follow a process that first uses the first equation to eliminate the first unknown from all equations but the first one itself. After that we use the second equation to eliminate the second unknown from all equations from the third on, and so on. One reason for this is that if we were to create a computer algorithm to solve systems, it would need a consistent method to proceed, and what we have done is as good as any.

What is the geometric interpretation of this? Since there are three unknowns, the appropriate geometric setting is three-dimensional space. *The solution set to any equation* $ax + by + cz = d$ *is a plane* in three-dimensional space, as long as not all of $a$, $b$ and $c$ are zero. Therefore, a solution to the system is a point that lies on each of the planes representing the solution sets of the three equations. For our example, then, the planes representing the three equations intersect at the point $(3, -1, 2)$.

In the study of linear algebra we will be defining new concepts and developing corresponding notation. We begin the development of notation with the following. The set of all real numbers is denoted by $\mathbb{R}$, and the set of all ordered pairs of real numbers is $\mathbb{R}^2$, spoken as "R-two." Geometrically, $\mathbb{R}^2$ is the familiar Cartesian coordinate plane. Similarly, the set of all ordered triples of real numbers is the three-dimensional space referred to as $\mathbb{R}^3$, "R-three."

All of the algebra that we will be doing using equations with two or three unknowns can easily be done with more unknowns. In general, when we are working with $n$ unknowns, we will get solutions that are $n$-**tuples** of numbers. Any such $n$-tuple represents a location in $n$-dimensional space, denoted $\mathbb{R}^n$. Note that a linear equation in two unknowns represents a line in $\mathbb{R}^2$, in the sense that the set of solutions to the equation forms a line. We consider a line to be a one-dimensional object, so the linear equation represents a one-dimensional object in two-dimensional space. The solution set to a linear equation in three unknowns is a plane in three-dimensional space. The plane itself is two-dimensional, so we have a two-dimensional "flat" object in three dimensional space.

Similarly, when we consider the solution set of a linear equation in $n$ unknowns, its solution set represents an $n-1$-dimensional "flat" object in $n$-dimensional space. When such an object has more than two dimensions, we usually call it a **hyperplane**. Although such objects can't be visualized, they certainly exist in a mathematical sense.

**Section 1.3 Exercises** To Solutions

1. Solve each of the following systems by the addition method.

(a) $\begin{aligned} 2x - 3y &= -7 \\ -2x + 5y &= 9 \end{aligned}$ (b) $\begin{aligned} 2x - 3y &= -6 \\ 3x - y &= 5 \end{aligned}$ (c) $\begin{aligned} 4x + y &= 14 \\ 2x + 3y &= 12 \end{aligned}$

(d) $\begin{aligned} 7x - 6y &= 13 \\ 6x - 5y &= 11 \end{aligned}$ (e) $\begin{aligned} 5x + 3y &= 7 \\ 3x - 5y &= -23 \end{aligned}$ (f) $\begin{aligned} 5x - 3y &= -11 \\ 7x + 6y &= -12 \end{aligned}$

2. Solve each of the following systems by graphing, as done in Example 1.3(c).

(a) $\begin{aligned} 3x - 4y &= 8 \\ x + 2y &= 6 \end{aligned}$ (b) $\begin{aligned} 4x - 3y &= 9 \\ x + 2y &= -6 \end{aligned}$ (c) $\begin{aligned} 5x + y &= 12 \\ 7x - 2y &= 10 \end{aligned}$

## 1.4 Solving With Matrices

**Performance Criteria:**

1. (e) Give the coefficient matrix and augmented matrix for a system of equations.

   (f) Determine whether a matrix is in row-echelon form. Perform, by hand, elementary row operations to reduce a matrix to row-echelon form.

   (g) Determine whether a matrix is in reduced row-echelon form. Use technology to reduce a matrix to reduced row-echelon form.

   (h) For a system of equations having a unique solution, determine the solution from either the row-echelon form or reduced row-echelon form of the augmented matrix for the system.

   (i) Use a calculator to solve a system of linear equations having a unique solution.

Note that when using the addition method for solving the system of three equations in three unknowns in the previous section, the symbols $x$, $y$ and $z$ and the equal signs are simply "placeholders" that are "along for the ride." To make the process cleaner we can simply arrange the constants $a$, $b$, $c$ and $d$ for each equation $ax + by + cz = d$ in an array form called a **matrix**, which is simply a table of values like

$$\begin{bmatrix} 1 & 3 & -2 & -4 \\ 3 & 7 & 1 & 4 \\ -2 & 1 & 7 & 7 \end{bmatrix}. \tag{1}$$

Each number in a matrix is called an **entry** of the matrix. Each horizontal line of numbers in a matrix is a **row** of the matrix, and each vertical line of numbers is a **column**. The **size** or **dimensions** of a matrix is (are) given by first telling the number of rows, then the number of columns, with the $\times$ symbol between them. The size of the above matrix is $3 \times 4$, which we say as "three by four."

Suppose that the above matrix came from the system of equations

$$\begin{aligned} x + 3y - 2z &= -4 \\ 3x + 7y + z &= 4 \\ -2x + y + 7z &= 7 \end{aligned}$$

When a matrix represents a system of equations, as (1) does, it is called the **augmented matrix** of the system. The matrix consisting of just the coefficients of $x$, $y$ and $z$ from each equation is called the **coefficient matrix**:

$$\begin{bmatrix} 1 & 3 & -2 \\ 3 & 7 & 1 \\ -2 & 1 & 7 \end{bmatrix}$$

We are not interested in the coefficient matrix at this time, but we will be later. The reason for the name "augmented matrix" will also be seen later.

Once we have the augmented matrix, we can perform a process called **row-reduction**, which is essentially what we did in the previous section, but we work with just the matrix rather than the system of equations. The following example shows how this is done for the above matrix.

◇ **Example 1.4(a):** Solve the system

$$\begin{aligned} x+3y-2z &= -4\\ 3x+7y+z &= 4\\ -2x+y+7z &= 7 \end{aligned}$$

from the previous section by row-reduction.

**Solution:** We begin with the augmented matrix for the system, shown below and to the left. We then add negative three times the first row to the second, and put the result in the second row. Then we add two times the first row to the third, and place the result in the third. Using the notation $R_n$ (not to be confused with $\mathbb{R}^n$!) to represent the $n$th row of the matrix, we can symbolize these two operations as shown in the middle below. The matrix to the right below is the result of those operations.

$$\begin{bmatrix} 1 & 3 & -2 & -4\\ 3 & 7 & 1 & 4\\ -2 & 1 & 7 & 7 \end{bmatrix} \quad \begin{array}{c} -3R_1+R_2\rightarrow R_2\\ \Longrightarrow\\ 2R_1+R_3\rightarrow R_3 \end{array} \quad \begin{bmatrix} 1 & 3 & -2 & -4\\ 0 & -2 & 7 & 16\\ 0 & 7 & 3 & -1 \end{bmatrix}$$

Next we finish with the following:

$$\begin{bmatrix} 1 & 3 & -2 & -4\\ 0 & -2 & 7 & 16\\ 0 & 7 & 3 & -1 \end{bmatrix} \quad \begin{array}{c} \frac{7}{2}R_2+R_3\rightarrow R_3\\ \Longrightarrow \end{array} \quad \begin{bmatrix} 1 & 3 & -2 & -4\\ 0 & -2 & 7 & 16\\ 0 & 0 & \frac{55}{2} & 55 \end{bmatrix}$$

The process just outlined is called **row reduction**. At this point we return to the equation form

$$\begin{aligned} x+3y-\ 2z &= -4\\ 0x-2y+\ 7z &= 16\\ 0x+0y+\tfrac{55}{2}z &= 55 \end{aligned}$$

and perform back-substitution. The last equation gives us that $z=2$. We can than substitute this value into the second equation to get $-2y+14=16$, resulting in $y=-1$. These values of $y$ and $z$ are substituted into the first equation which is then solved to get $x=3$. The solution to the system is then $(3,-1,2)$.

---

The final form of the matrix before we went back to equation form is something called **row-echelon** form. (The word "echelon" is pronounced "esh-el-on.") The first non-zero entry in each row is called a **leading entry**; in this case the leading entries are the numbers $1$, $-2$ and $\frac{55}{2}$. To be in row-echelon form means that

- any rows containing all zeros are at the bottom of the matrix and
- the leading entry in any row is to the right of any leading entries above it.

◇ **Example 1.4(b):** Which of the matrices below are in row-echelon form?

$$\begin{bmatrix} 1 & 3 & -2 & -4\\ 0 & 0 & 3 & -5\\ 0 & 7 & -10 & -1 \end{bmatrix} \quad \begin{bmatrix} 2 & 6 & -1 & 9 & 5\\ 0 & 0 & -8 & 1 & -3\\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \quad \begin{bmatrix} 7 & -12 & 5 & 0\\ 0 & 0 & 0 & 0\\ 0 & -5 & 1 & 8 \end{bmatrix}$$

**Solution:** The leading entries of the rows of the first matrix are $1$, $3$ and $7$. Because the leading entry of the third row $(7)$ is not to the right of the leading entry of the second row $(3)$, the first matrix *is not* in row-echelon form. In the third matrix, there is a row of zeros that is not at the bottom of the matrix, so it *is not* in row-echelon form. The second matrix *is* in row-echelon form.

---

Note that if we switch the second and third rows of the first and third matrices in the above example, which we are usually allowed to do, then both will then be in row-echelon form.

It is possible to continue with the matrix operations to obtain something called **reduced row-echelon form**, from which it is easier to find the values of the unknowns. The requirements for being in reduced row-echelon form are the same as for row-echelon form, with the addition of the following:

- All leading entries are ones.
- The entries above any leading entry are all zero *except perhaps in the last column*.

Obtaining reduced row-echelon form requires more matrix manipulations, and nothing is really gained by obtaining that form if you are doing this by hand. However, when using software or a calculator it is most convenient to obtain reduced row-echelon form. Here are two examples of matrices in reduced row-echelon form:

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -7 \\ 0 & 0 & 1 & 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 6 & 0 & 9 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

In the next section we will see how to interpret what the second matrix would be telling us if it came from a system of equations. The next example shows what the first matrix tells us.

◇ **Example 1.4(c):** Suppose that the matrix to the right is the result of row-reduction of the augmented matrix for a system of three equations in the unknowns $x_1$, $x_2$ and $x_3$. Determine the values of the unknowns.

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

**Solution:** When using row-reduction to solve a system we first create the augmented matrix for the system, then row-reduce it, and then we go back to equations. The equations we would return to for the above matrix are

$$\begin{aligned} 1x_1 + 0x_2 + 0x_3 &= 3 \\ 0x_1 + 1x_2 + 0x_3 &= -7 \\ 0x_1 + 0x_2 + 1x_3 &= 4 \end{aligned}$$

and from these we can easily see the solution: $x_1 = 3$, $x_2 = -7$ and $x_3 = 4$.

---

In practice, very large systems are solved by row-reduction. Many issues arise when doing this. For example, coefficients are often obtained from some sorts of measurements that give rounded values. At every step of row-reduction more rounding needs to take place, resulting in rounding errors. Additionally, matrices used in practice can have entries that cause introduction of other errors in the process of row-reduction. We could spend an entire course examining such concerns, but instead we'll focus on less numerically oriented aspects of linear algebra.

That said, let's look at one thing that can come up in the process of row-reduction, illustrated in the following example.

◇ **Example 1.4(d):** Row-reduce the matrix $\begin{bmatrix} 1 & 3 & -2 & -4 \\ 2 & 6 & -1 & -13 \\ -1 & 4 & -8 & 3 \end{bmatrix}$.

**Solution:** We begin by adding negative two times the first row to the second, and put the result in the second row. Then we add two times the first row to the third, and place the result in the third. Using the notation $R_n$ (not to be confused with $\mathbb{R}^n$!) to represent the $n$th row of the matrix, we can symbolize these two operations as shown in the middle below. The matrix to the right below is the result of those operations.

$$\begin{bmatrix} 1 & 3 & -2 & -4 \\ 2 & 6 & -1 & -13 \\ -1 & 4 & -8 & 3 \end{bmatrix} \quad \begin{array}{c} -2R_1 + R_2 \to R_2 \\ \Longrightarrow \\ R_1 + R_3 \to R_3 \end{array} \quad \begin{bmatrix} 1 & 3 & -2 & -4 \\ 0 & 0 & 3 & -5 \\ 0 & 7 & -10 & -1 \end{bmatrix}$$

We can see that the matrix would be in row-echelon form if we simply switched the second and third rows (which is equivalent to simply rearranging the order of our original equations), so that's what we do:

$$\begin{bmatrix} 1 & 3 & -2 & -4 \\ 0 & 0 & 3 & -5 \\ 0 & 7 & -10 & -1 \end{bmatrix} \quad \begin{array}{c} R_2 \longleftrightarrow R_3 \\ \Longrightarrow \end{array} \quad \begin{bmatrix} 1 & 3 & -2 & -4 \\ 0 & 7 & -10 & -1 \\ 0 & 0 & 3 & -5 \end{bmatrix}$$

---

The act of rearranging rows in a matrix is called **permuting** them. In general, a **permutation** of a set of objects is simply a rearrangement of them. When solving a system by row-reduction, permuting simply amounts to changing the order of the original equations, and doing so will not affect the solution to the system.

### Row Reduction Using Technology

There are three main technologies that can be used to get an augmented matrix into reduced row-echelon form:

- Most or all graphing calculators (and the TI-36X Pro, a non-graphing calculator) will perform row reduction via the *rref* function. Do a search to find an article or video on how to *rref* with your particular model of calculator.
- There are numerous matrix calculators that can be found online - there is a link to one at the class web page.
- Various mathematical software programs, like MATLAB, will perform row-reduction.

Now that we know how to solve systems of linear equations we can complete an application.

◇ **Example 1.4(e):** Find the equation of the third degree polynomial containing the points $(-1,-7)$, $(0,1)$, $(1,5)$ and $(2,11)$.

**Solution:** A general third degree polynomial has an equation of the form $y = ax^3+bx^2+cx+d$; our goal is to find values of $a$, $b$, $c$ and $d$ so that the given points all satisfy the equation. Since the values $x = -1$, $y = -7$ must make the general equation true, we have $-7 =$

$a(-1)^3 + b(-1)^2 + c(-1) + d = -a + b - c + d$. Doing this with all four given ordered pairs and "flipping" each equation gives us the system

$$\begin{aligned} -a + b - c + d &= -7 \\ d &= 1 \\ a + b + c + d &= 5 \\ 8a + 4b + 2c + d &= 11 \end{aligned}$$

If we enter the augmented matrix for this system in our calculators and *rref* we get

$$\begin{bmatrix} -1 & 1 & -1 & 1 & -7 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 5 \\ 8 & 4 & 2 & 1 & 11 \end{bmatrix} \quad \stackrel{rref}{\Longrightarrow} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

So $a = 1$, $b = -2$, $c = 5$, $d = 1$, and the desired polynomial equation is $y = x^3 - 2x^2 + 5x + 1$.

---

**Section 1.4 Exercises** To Solutions

1. Give the coefficient matrix and augmented matrix for the system of equations

$$\begin{aligned} x + y - 3z &= 1 \\ -3x + 2y - z &= 7 \\ 2x + y - 4z &= 0 \end{aligned}.$$

2. Determine which of the following matrices are in row-echelon form.

$$A = \begin{bmatrix} 3 & -7 & 5 & 0 & 2 & -4 \\ 0 & 0 & 0 & -2 & 5 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 5 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 3 & 5 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 0 & 4 & 4 \\ 0 & 1 & -3 & 2 \\ 6 & 1 & 3 & 5 \end{bmatrix} \qquad E = \begin{bmatrix} 1 & 3 & -5 & 10 & -7 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad F = \begin{bmatrix} 1 & 3 & 0 & 0 & -7 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

3. Determine which of the matrices in Exercise 2 are in reduced row-echelon form.

4. Perform the first two row operations for the augmented matrix from Exercise 1, to get zeros in the bottom two entries of the first column.

5. Fill in the blanks in the second matrix with the appropriate values after the first step of row-reduction. Fill in the long blanks with the row operations used.

(a) $$\begin{bmatrix} 1 & 5 & -7 & 3 \\ -5 & 3 & -1 & 0 \\ 4 & 0 & 8 & -1 \end{bmatrix} \quad \begin{matrix} \underline{\hspace{3cm}} \\ \Longrightarrow \\ \underline{\hspace{3cm}} \end{matrix} \quad \begin{bmatrix} \_\_ & \_\_ & \_\_ & \_\_ \\ 0 & \_\_ & \_\_ & \_\_ \\ 0 & \_\_ & \_\_ & \_\_ \end{bmatrix}$$

(b) $\begin{bmatrix} 2 & -8 & -1 & 5 \\ 0 & -2 & 0 & 0 \\ 0 & 6 & -5 & 2 \end{bmatrix} \quad \overset{\Longrightarrow}{\underline{\hspace{3cm}}} \quad \begin{bmatrix} \underline{\quad} & \underline{\quad} & \underline{\quad} & \underline{\quad} \\ 0 & \underline{\quad} & \underline{\quad} & \underline{\quad} \\ 0 & 0 & \underline{\quad} & \underline{\quad} \end{bmatrix}$

(c) $\begin{bmatrix} 1 & -2 & 4 & 1 \\ 0 & 3 & 5 & -2 \\ 0 & 2 & -8 & 1 \end{bmatrix} \quad \overset{\Longrightarrow}{\underline{\hspace{3cm}}} \quad \begin{bmatrix} \underline{\quad} & \underline{\quad} & \underline{\quad} & \underline{\quad} \\ 0 & \underline{\quad} & \underline{\quad} & \underline{\quad} \\ 0 & 0 & \underline{\quad} & \underline{\quad} \end{bmatrix}$

6. Find $x$, $y$ and $z$ for the system of equations that reduces to each of the matrices shown.

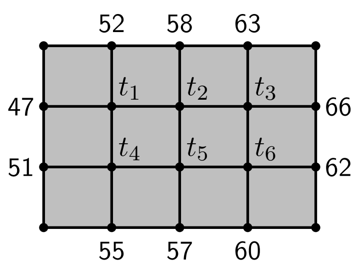
(a) $\begin{bmatrix} 1 & 6 & -2 & 7 \\ 0 & 8 & 1 & 0 \\ 0 & 0 & -2 & 8 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 6 & -2 & 7 \\ 0 & 2 & -5 & -13 \\ 0 & 0 & 3 & 3 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 0 & 7 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -4 & 8 \end{bmatrix}$

7. Use row operations (by hand) on an augmented matrix to solve each system of equations.

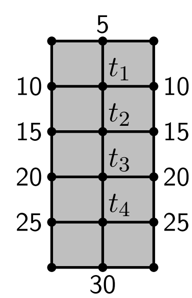
(a) $\begin{aligned} x - 2y - 3z &= -1 \\ 2x + y + z &= 6 \\ x + 3y - 2z &= 13 \end{aligned}$ (b) $\begin{aligned} -x - y + 2z &= 5 \\ 2x + 3y - z &= -3 \\ 5x - 2y + z &= -10 \end{aligned}$ (c) $\begin{aligned} x + 2y + 4z &= 7 \\ -x + y + 2z &= 5 \\ 2x + 3y + 3z &= 7 \end{aligned}$

8. Use the *rref* capability of your calculator to solve each of the systems from the previous exercise.

9. Temperatures at points along the edges of a rectangular plate are as shown below and to the left. Find the equilibrium temperature at each of the interior points, to the nearest tenth.




Exercise 9




Exercise 10

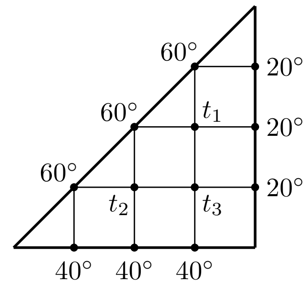
10. Consider the rectangular plate with boundary temperatures shown below and to the right of Exercise 9.

    (a) Intuitively, what do you think that the equilibrium temperatures $t_1$, $t_2$, $t_3$ and $t_4$ are?

    (b) Set up a system of equations and find the equilibrium temperatures. How was your intuition?

11. Look at your solutions *and the boundary temperatures* for Exercises 9 and 10. For each plate, look at where the maximum and minimum temperatures occur. What can we say in general about the locations of the maximum and minimum temperatures? Can you see how this is implied by the Mean Value Property?

12. For the diagram to the right, the mean value property still holds, even though the plate in this case is triangular. Find the interior equilibrium temperatures, rounded to the nearest tenth.



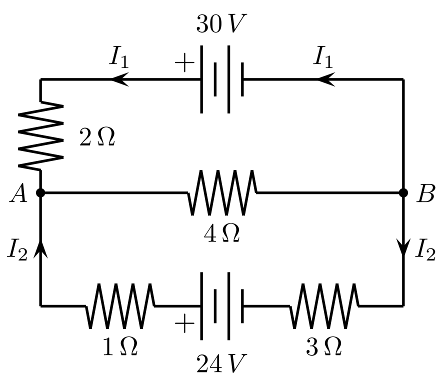

13. (a) Plot the points $(-4,0)$, $(-2,2)$, $(0,0)$, $(2,2)$ and $(3,0)$ neatly on an $xy$ grid. Sketch the graph of a polynomial function with the fewest number of turning points ("humps") possible that goes through all the points. What is the degree of the polynomial function?

    (b) Find a fourth degree polynomial $y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$ that goes through the given points.

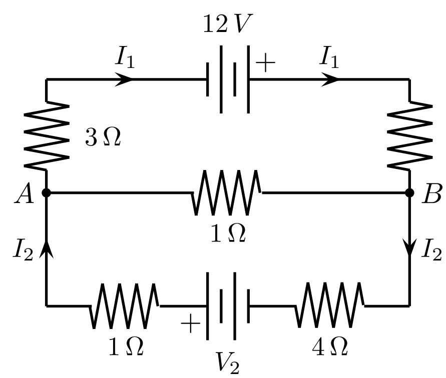
    (c) Graph your function from (b) on your calculator and sketch it, using a dashed line, on your graph from (a). Is the graph what you expected?

14. (a) Find the currents $I_1$ and $I_2$ in the circuit with the diagram shown below and to the left.

    (b) What is the value of the current through the 4 ohm resistor, and does it flow from $A$ to $B$, or from $B$ to $A$?




Exercise 14




Exercise 15

15. Consider the circuit shown to the right below Exercise 14.

    (a) Find the currents $I_1$, $I_2$ and $I_3$ when the voltage $V_2$ is 6 volts, rounded to the tenth's place.

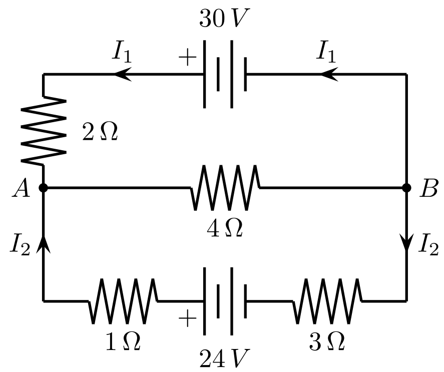
    (b) Does the current in the middle branch of the circuit flow from $A$ to $B$, or from $B$ to $A$?

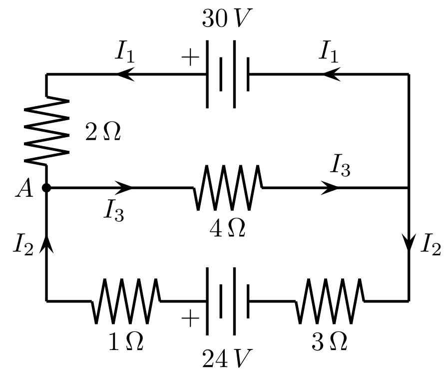
    (c) Find the currents $I_1$, $I_2$ and $I_3$ when the voltage $V_2$ is 24 volts, rounded to the tenth's place.

    (d) Does the current in the middle branch of the circuit flow from $A$ to $B$, or from $B$ to $A$?

    (e) Determine the voltage needed for $V_2$ in order that no current flows through the middle branch. (You might wish to row reduce by hand for this...)

16. In this exercise you will find the currents in the circuit from Exercise 14 (shown to the left below) by a slightly different manner. Rather than working with just the currents $I_1$ and $I_2$ and then adding them to find the current from $A$ to $B$, We will begin with another unknown current $I_3$, as shown in the diagram to the right below.




(a) Set up equations for both the upper and lower loops as before, but use $I_3$ as the current through the 4 ohm resistor, rather than $I_1 + I_2$. This will give you two equations with three unknowns in them, $I_1$, $I_2$ and $I_3$.

(b) We need one more equation, which we get as follows: The current into node $A$ must equal the current out. Use this to write an equation, then get all of $I_1$, $I_2$ and $I_3$ on one side.

(c) Solve your system to get the three currents.

17. The equation of a non-vertical plane in $\mathbb{R}^3$ can always be written in the form $z = a + bx + cy$, where $a$, $b$ and $c$ are constants and $(x, y, z)$ is any point on the plane. Use a method similar to the method for finding the equation of a polynomial through a given set of points to find the equation of the plane through the three points $P_1(-5, 0, 2)$, $P_2(4, 5, -1)$ and $P_3(2, 2, 2)$. Use your calculator's *rref* command to solve the system. Round $a$, $b$ and $c$ to the thousandth's place.

## 1.5 "When Things Go Wrong"

**Performance Criteria:**

1. (j) Given the row-echelon or reduced row-echelon form of an augmented matrix for a system of equations, determine the leading variables and free variables of the system.

   (k) Given the row-echelon or reduced row-echelon form for a system of equations:

   - Determine whether the system has a unique solution, and give the solution if it does.
   - If the system does not have a unique solution, determine whether it is inconsistent (no solution) or dependent (infinitely many solutions).
   - If the system is dependent, give the general form of a solution and give some particular solutions.

Consider the three systems of equations

$$\begin{aligned} x - 3y &= 6 \\ -2x + 5y &= -5 \end{aligned} \qquad \begin{aligned} x - 2y &= 3 \\ -2x + 4y &= 1 \end{aligned} \qquad \begin{aligned} x - 2y &= 3 \\ -2x + 4y &= -6 \end{aligned}$$

For the first system, if we multiply the first equation by 2 and add it to the second, we get $-y = 7$, so $y = -7$. This can be substituted into either equation to find $x = -15$, and the system is solved!

When attempting to solve the second and third systems, things do not "work out" in the same way. In both cases we would likely attempt to eliminate $x$ by multiplying the first equation by two and adding it to the second. For the second system this results in $0 = 7$ and for the third the result is $0 = 0$. So what is happening? Let's keep the unknown value $y$ in both equations: $0y = 7$ and $0y = 0$. There is no value of $y$ that can make $0y = 7$ true, so there is no solution to the second system of equations. We call a system of equations with no solution **inconsistent**.

The equation $0y = 0$ is true for *any* value of $y$, so $y$ can be anything in the third system of equations. Thus we will call $y$ a **free variable**, meaning it is free to have any value. *In this sort of situation we will assign another unknown, usually $t$, to represent the value of the free variable.* (If there is another free variable we usually use $s$ and $t$ for the two free variables.) Once we have assigned the value $t$ to $y$, we can substitute it into the first equation and solve for $x$ to get $x = 2t + 3$.

What all this means is that any ordered pair of the form $(2t + 3, t)$ will be a solution to the third system of equations above. For example, when $t = 0$ we get the ordered pair $(3, 0)$, when $t = -6$ we get $(-9, -6)$. You can verify that both of these are solutions, as are infinitely many other pairs. At this point you might note that we could have made $x$ the free variable, then solved for $y$ in terms of whatever variable we assigned to $x$. *It is standard convention, however, to start assigning free variables from the last variable, and you will be expected to follow that convention in this class.* A system like this, with infinitely many solutions, is called a **dependent** system.

The fundamental fact that should always be kept in mind is this.

**Solutions to a System of Equations**

Every system of linear equations has either

- one unique solution
- no solution (the system is inconsistent)
- infinitely many solutions (the system is dependent)

In the context of both linear algebra and differential equations, mathematicians are always concerned with "existence and uniqueness." What this means is that when attempting to solve a system of equations or a differential equation, one cares about

1) whether at least one solution exists and
2) if there is at least one solution, is there exactly one; that is, is the solution unique?

We'll now see if we can learn to recognize which of the above three situations is the case, based on the row-echelon or reduced row-echelon form of the augmented matrix of a system. If the three systems we have been discussing are put into augmented matrix form and row reduced we get

$$\begin{bmatrix} 1 & 0 & -15 \\ 0 & 1 & -7 \end{bmatrix} \qquad \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & -2 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

It should be clear that the first matrix gives us the unique solution to that system. The second line of the second matrix "translates" back to the equation $0x + 0y = 7$, which clearly cannot be true for any values of $x$ or $y$. So that system has no solution.

If the row reduced augmented matrix for a system has any row with entries all zeros EXCEPT the last one, the system has no solution. The system is said to be **inconsistent**.

We now consider the third row reduced matrix. The last line of it "translates" to $0x + 0y = 0$, which is true for *any* values of $x$ and $y$. That means we are free to choose the value of either one but, as discussed before, it is customary to let $y$ be the free variable. So we let $y = t$ and substitute that into the equation $x - 2y = 3$ represented by the first line of the reduced matrix. As before, that is solved for $x$ to get $x = 2t + 3$. The solutions to the system are then $x = 2t + 3$, $y = t$ for all values of $t$.

Of the three cases (1) exactly one solution, (2) no solution, (3) infinitely many solutions, the third case is the most challenging to interpret in most situations. As an introduction, let's consider the system shown below and to the left; its augmented matrix reduces to the form shown below and to the right.

$$\begin{aligned} x_1 - x_2 + x_3 &= 3 \\ 2x_1 - x_2 + 4x_3 &= 7 \\ 3x_1 - 5x_2 - x_3 &= 7 \end{aligned} \qquad \begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We now make the following definitions:

- The **leading variables** are the variables corresponding to the columns of the reduced matrix containing the first non-zero entries (always ones for reduced row-echelon form) in each row. For the above system the leading variables are $x_1$ and $x_2$.
- Any variables that are not leading variables are **free variables**, so $x_3$ is the free variable in the above system. This means it is free to take any value.

It is a bit difficult to explain how to solve systems with infinitely many solutions, and it is probably best seen by some examples. However, let me try to describe it. Start with the last variable and solve for it if it is a leading variable. If it is not, assign it a parameter, like $t$. If the next to last variable is a leading variable solve for it, either as a number or in terms of the parameter assigned to the last variable. Continue in this manner until all variables have been determined as numbers or in terms of parameters.

◇ **Example 1.5(a):** Solve the system
$$\begin{aligned} x_1 - x_2 + x_3 &= 3 \\ 2x_1 - x_2 + 4x_3 &= 7 \\ 3x_1 - 5x_2 - x_3 &= 7 \end{aligned}$$

**Solution:** The row-reduced form of the augmented matrix for this system is $\begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

In this case the leading variables are $x_1$ and $x_2$. Any variables that are not leading variables are free variables, so $x_3$ is the free variable in this case. If we let $x_3 = t$, the last non-zero row gives the equation $x_2 + 2t = 1$, so $x_2 = -2t + 1$. The first row gives the equation $x_1 + 3x_3 = 4$, so $x_1 = -3t + 4$ and the final solution to the system is

$$x_1 = -3t + 4, \qquad x_2 = -2t + 1, \qquad x_3 = t$$

We can also think of the solution as being any ordered triple of the form $(-3t + 4, -2t + 1, t)$.

---

◇ **Example 1.5(b):** A system of three equations in the four variables $x_1$, $x_2$, $x_3$ and $x_4$ gives the row-reduced matrix

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

Give the general solution to the system.

**Solution:** The leading variables are $x_1$, $x_2$ and $x_4$. Any variables that are not leading variables are the free variables, so $x_3$ is the free variable in this case. We can see that the last row gives us $x_4 = 4$. Letting $x_3 = t$, the second equation from the row-reduced matrix is $x_2 - 5t = 2$, so $x_2 = 5t + 2$. The first equation is $x_1 + 3t = -1$, giving $x_1 = -3t - 1$. The final solution to the system is then

$$x_1 = -3t - 1, \qquad x_2 = 5t + 2, \qquad x_3 = t, \qquad x_4 = 4,$$

or $(-3t - 1, 5t + 2, t, 4)$.

---

The solutions given in the previous two examples are called **general solutions**, because they tell us what any solution to the system looks like in the cases where there are infinitely many solutions. We can also produce some specific numbers that are solutions as well, which we will call **particular solutions**. These are obtained by simply letting any parameters take on whatever values we want.

◇ **Example 1.5(c):** Give three particular solutions to the system in Example 1.5(a).

**Solution:** If we take the easiest choice for $t$, zero, we get the particular solution $(4, 1, 0)$. Letting $t$ equal negative one and one gives us the particular solutions $(7, 3, -1)$ and $(1, -1, 1)$.

---

The following examples show a situation in which there are two free variables, and one in which there is no solution.

◇ **Example 1.5(d):** A system of equations in the four variables $x_1$, $x_2$, $x_3$ and $x_4$ that has the row-reduced matrix

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

Give the general solution and four particular solutions.

**Solution:** In this case the leading variables are $x_1$ and $x_3$, and the free variables are $x_2$ and $x_4$. We begin by letting $x_4 = t$; we have the equation $x_3 - 2t = 3$, giving us $x_3 = 2t + 3$. Since $x_2$ is a free variable, we call it something else. $t$ has already been used, so let's say $x_2 = s$. The first equation indicated by the row-reduced matrix is then $x_1 + 2s - t = 2$, giving us $x_1 = -2s + t + 2$. The solution to the corresponding system is

$$x_1 = -2s + t + 2, \qquad x_2 = s, \qquad x_3 = 2t + 3, \qquad x_4 = t$$

If we let $s = 0$ and $t = 0$ we get the solution $(2, 0, 3, 0)$, and if we let $s = 2$ and $t = -1$ we get $(-3, 2, 1, -1)$. Letting $s = 0$ and $t = 1$ gives the particular solution $(3, 0, 5, 1)$ and letting $s = 1$ and $t = 0$ gives the particular solution $(0, 1, 3, 0)$.

---

The values used for the parameters in Examples 1.5(c) and (d) were chosen arbitrarily; any values can be used for $s$ and $t$.

◇ **Example 1.5(e):** A system of equations in the four variables $x_1$, $x_2$, $x_3$ and $x_4$ has the row-reduced matrix

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

Solve the system.

**Solution:** Since the last row is equivalent to the equation $0x_1 + 0x_2 + 0x_3 + 0x_4 = 5$, which has no solution, the system itself has no solution.

---

**Section 1.5 Exercises** To Solutions

1. Consider the system of equations $\begin{aligned} 2x - 4y - z &= -4 \\ 4x - 8y - z &= -4 \\ -3x + 6y + z &= 4 \end{aligned}$.

   (a) Determine which of the following ordered triples are solutions to the system of equations:

   $$(6,3,4) \qquad (3,-1,4) \qquad (0,0,4) \qquad (-2,-1,4) \qquad (5,2,0) \qquad (2,1,4)$$

   Look for a pattern in the ordered triples that *ARE* solutions. Try to guess another solution, and test your guess by checking it in all three equations. How did you do?

   (b) When you tried to solve the system using your calculator, you should have gotten the reduced echelon matrix as

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

   Give the system of equations that this matrix represents. Which variable can you determine?

   (c) It is not possible to determine $y$, so we simply let it equal some arbitrary value, which we will call $t$. So at this point, $z = 4$ and $y = t$. Substitute these into the first equation and solve for $x$. Your answer will be in terms of $t$. Write the ordered triple solution to the system.

   **NOTE:** The system of equations you obtained in part (b) and solved in part (c) has infinitely many solutions, but we do know that every one of them has the form $(2t, t, 4)$. Note how this explains the results of part (a).

2. The reduced echelon form of the matrix for the system

   $$\begin{aligned} 3x - 2y + z &= -7 \\ 2x + y - 4z &= 0 \\ x + y - 3z &= 1 \end{aligned} \qquad \text{is} \qquad \begin{bmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

   (a) Give the free variable(s) and leading variable(s).

   (b) In this case, $z$ cannot be determined, so we let $z = t$. Now solve for $y$, in terms of $t$. Then solve for $x$ in terms of $t$.

   (c) Pick a specific value for $t$ and substitute it into your general form of a solution triple for the system. Check it by substituting it into all three equations in the original system.

   (d) Repeat (b) for a different value of $t$.

3. The reduced echelon forms of some systems are given below.

- If the system has a unique solution, give it. If the system has no solution, say so.
- If the system has infinitely many solutions, give the general solution in terms of parameters $s$, $t$, etc., then give two particular solutions.

(a) $\left[\begin{array}{rrrrr} 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & 2 & 0 & -5 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right]$

(b) $\left[\begin{array}{rrrrr} 1 & 3 & 0 & 0 & 5 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -4 \end{array}\right]$

(c) $\left[\begin{array}{rrrrr} 1 & -3 & 0 & 1 & -4 \\ 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

(d) $\left[\begin{array}{rrrr} 1 & 0 & 1 & 5 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 1 \end{array}\right]$

(e) $\left[\begin{array}{rrrrr} 1 & 0 & -2 & 1 & 6 \\ 0 & 1 & 3 & 5 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

(f) $\left[\begin{array}{rrrr} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{array}\right]$

(g) $\left[\begin{array}{rrrr} 1 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

(h) $\left[\begin{array}{rrrrr} 1 & 4 & -1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$

(i) $\left[\begin{array}{rrrrr} 1 & 5 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

4. Give four particular solutions from the general solution of Example 1.5(b), which had general solution $(-3t-1, 5t+2, t, 4)$.

5. For the systems whose augmented matrices row reduce to the forms shown below, do one of the following:

- If the system has a unique solution, give it. If the system has no solution, say so.
- If the system has infinitely many solutions, give the general solution in terms of parameters $s$, $t$, etc., then give two particular solutions.

$$\left[\begin{array}{rrr} 1 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{array}\right] \qquad \left[\begin{array}{rrrrr} 1 & 2 & -1 & 0 & 5 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

## 1.6 Back to Applications

**Performance Criterion:**

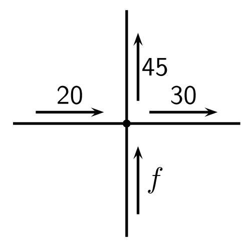
1. (l) Use systems of equations to solve network problems.

The concept of a network was introduced in Exercise 7 of Section 1.1. To review, a network is a set of junctions, which we'll call **nodes**, connected by what could be called pipes, or wires, but which we'll call **directed edges**. The word "directed" is used to mean that we'll assign a direction of flow to each edge. There will also be directed edges coming into or leaving the network. It is probably easiest to just think of a network of plumbing, with water coming in at perhaps several places, and leaving at several others.

Our study of networks will be based on one simple idea, known as **conservation of flow**:

At each node of a network, the flow into the node must equal the flow out.

◇ **Example 1.6(a):** A one-node network is shown to the right. Find the unknown flow $f$.

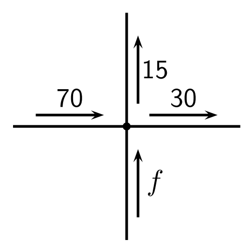



**Solution:** The flow in is $20 + f$ and the flow out is $45 + 30$, so we have

$$20 + f = 45 + 30.$$

Solving, we find that $f = 55$.

◇ **Example 1.6(b):** Another one-node network is shown to the right. Find the unknown flow $f$.



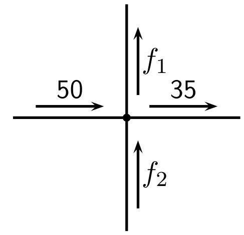

**Solution:** The flow in is $70 + f$ and the flow out is $15 + 30$, so we have

$$70 + f = 15 + 30.$$

Solving, we find that $f = -25$, so the flow at the arrow labeled $f$ is actually in the direction opposite to the arrow.

When setting up a network we must commit to a direction of flow for any edges in which the flow is unknown, but when solving the system we may find that the flow is in the opposite direction from the way the edge was directed initially, as we just saw. We may also have less information than we did in the previous two examples, as shown by the next example.

◇ **Example 1.6(c):** For the one-node network is shown to the right, find the unknown flow $f_1$ in terms of the flow $f_2$.




**Solution:** By conservation of flow,

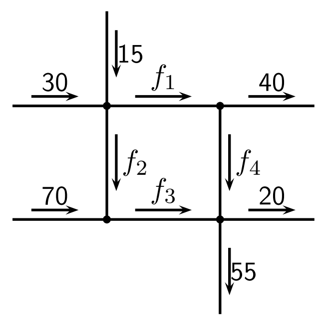
$$50 + f_2 = f_1 + 35.$$

Solving for $f_1$ gives us $f_1 = f_2 + 15$. Thus if $f_2$ was $10$, $f_1$ would be $25$ (look at the diagram and think about that), if $f_2$ was 45, $f_1$ would be $60$, and so on.

The previous example represents, in an applied setting, the idea of a **free variable**. In this example either variable can be taken as free, but if we know the value of one of them, we'll "automatically" know the value of the other. The way the example was worded, we were taking $f_2$ to be the free variable, with the value of $f_1$ then depending on the value of $f_2$.

The systems in these first three examples have been very simple; let's now look at a more complex system.

◇ **Example 1.6(d):** Determine the flows $f_1$, $f_2$, $f_3$ and $f_4$ in the network shown to the right.




**Solution:** Utilizing conservation of flow at each node, we get the equations

$$30 + 15 = f_1 + f_2, \qquad 70 + f_2 = f_3,$$

$$f_1 = 40 + f_4, \qquad f_3 + f_4 = 20 + 55$$

Rearranging these give us the system of equations shown below and to the left. The augmented matrix for this system reduces to the matrix shown below and to the right.

$$\begin{array}{rcrcrcrcr} f_1 & + & f_2 & & & & & = & 45 \\ & & f_2 & - & f_3 & & & = & -70 \\ f_1 & & & & & - & f_4 & = & 40 \\ & & & & f_3 & + & f_4 & = & 75 \end{array} \quad \Longrightarrow \quad \begin{bmatrix} 1 & 1 & 0 & 0 & 45 \\ 0 & 1 & -1 & 0 & -70 \\ 1 & 0 & 0 & -1 & 40 \\ 0 & 0 & 1 & 1 & 75 \end{bmatrix} \quad \Longrightarrow \quad \begin{bmatrix} 1 & 0 & 0 & -1 & 40 \\ 0 & 1 & 0 & 1 & 5 \\ 0 & 0 & 1 & 1 & 75 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

From this we can see that $f_4$ is a free variable, so lets say it has value $t$. The solution to the network is then

$$f_1 = 40 + t, \qquad f_2 = 5 - t, \qquad f_3 = 75 - t, \qquad f_4 = t,$$

where $t$ is the flow $f_4$.

### Underdetermined and Overdetermined Systems

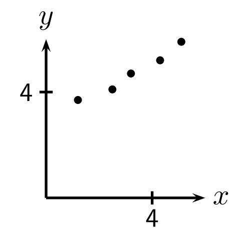
Let's think a bit more about this last example. Suppose that $f_4 = t = 0$. The equations given as the solution to the network then give us $f_1 = 40$, $f_2 = 5$, $f_3 = 75$. We can see this without even solving the system of equations. Looking at the node in the lower right, if $f_4 = 0$ one can easily see that $f_3$ must be 75 in order for the flow in to equal the flow out. Knowing $f_3$, we can go to the node in the lower left and see that $f_2 = 5$. Finally, $f_2 = 5$ gives is $f_1 = 40$. The information given originally was not sufficient to determine the values of the flows $f_1$, $f_2$, $f_3$ and $f_4$. In such a case, we sometimes say that the system is **undetermined**, meaning that there is too little information to guarantee a single solution. We just saw that with one more piece of information, the value of $f_4$, all of the remaining flows were then determined by that value.

Now consider the situation described in Exercise 6 of Section 1.1. Given the points

$$(1.2, 3.7) \qquad (2.5, 4.1) \qquad (3.2, 4.7) \qquad (4.3, 5.2) \qquad (5.1, 5.9)$$

we wish to find the equation $y = mx + b$ of a line containing them. Substituting each pair into $y = mx + b$ gives us the system shown below and to the left.

$$\begin{array}{rcl} 1.2m + b & = & 3.7 \\ 2.5m + b & = & 4.1 \\ 3.2m + b & = & 4.7 \\ 4.3m + b & = & 5.2 \\ 5.1m + b & = & 5.9 \end{array} \quad \Longrightarrow \quad \begin{bmatrix} 1.2 & 1 & 3.7 \\ 2.5 & 1 & 4.1 \\ 3.2 & 1 & 4.7 \\ 4.3 & 1 & 5.2 \\ 5.1 & 1 & 5.9 \end{bmatrix} \quad \overset{rref}{\Longrightarrow} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

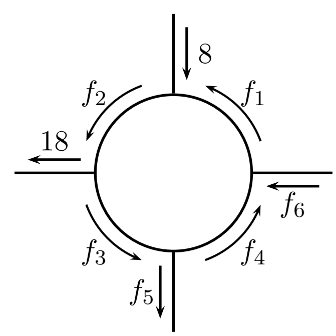



When we row-reduce the augmented matrix for this system, we get the last matrix above, indicating that the system has no solution. The five points are plotted above and to the right, and we can see that they are not on a line, which is why we were not able to solve the system. In this case the system is **overdetermined**, meaning that there is too much information to allow a solution to the system.

You might think "Well, why not just use less data, so that the resulting system has a solution?" Well the additional data gives us some redundancy that can give us better results *if we know how to deal with it.* The way out of this problem is a method called **least-squares**, which we'll see later. It is a method for dealing with systems that don't have solutions. What it allows us to do is obtain values that are in some sense the "closest" values there are to an actual solution. Again, more on this later.

## Section 1.6 Exercises

To Solutions




1. The network to the right represents a traffic circle. The numbers next to each of the paths leading into or out of the circle are the net flows in the directions of the arrows, in vehicles per minute, during the first part of lunch hour.

   (a) Suppose that $f_3 = 7$ and $f_5 = 4$. You should be able to work your way around the circle, eventually figuring out what each flow is. Do this.

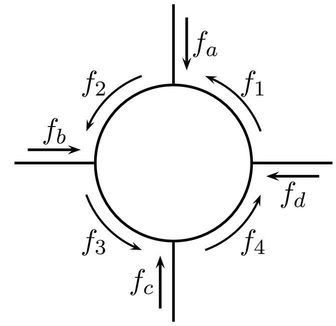
   (b) Now assume that the only flows you know are the ones shown in the diagram. When you set up a system of equations, based on the flows in and out of each junction, how many equations will you have? How many unknowns?

   (c) Go ahead and set up the system of equations. Give the augmented matrix and the reduced matrix (obtained with your calculator), and then give the general solution to the system.

   (d) Choose the value(s) of the parameter(s) that make $f_3 = 7$ and $f_5 = 4$, then find the resulting particular solution. *If your answers do not match what you got for part (a), go back and check your work for (c).*

   (e) What restriction(s) is(are) there on the parameter(s), in order that all flows go in the directions indicated. (Allow a flow of zero for each flow as well.)

2. For another traffic circle, a student uses the diagram shown to the right and obtains the flows given below, in vehicles per minute.




$$f_1 = t - 8, \qquad f_2 = t + 3, \qquad f_3 = t - 5, \qquad f_4 = t$$

(a) Determine the minimum value of $t$ that makes each of $f_1$ through $f_4$ zero or greater. Give the minimum allowable values for each flow, in the form $f_i \geq a$, assuming that no vehicles ever go the wrong way around a portion of the circle. Remember that setting a value for any flow determines all the other flows. You may neglect units.

(b) Give each of the flows $f_1$ through $f_4$ when the flow in the northeast quarter $(f_1)$ is 12 vehicles per minute. You may neglect units.

(c) Determine each of the flows $f_a$ through $f_d$, still for $f_1 = 12$. You should be able to do this based only on the four equations given. At least one of them will be negative, indicating that the corresponding arrow(s) should be reversed.

## 1.7 Chapter 1 Exercises

1. Consider the system of equations below and to the right. Solve the system by Gaussian elimination (get in row-echelon form, then perform back substitution), **by hand** (no calculator). Show all steps, including what operation was performed for each step. **Hint:** You may find it useful to put the equations in a different order before forming the augmented matrix.

$$\begin{aligned} 5x - y + 2z &= 17 \\ x + 3y - z &= -4 \\ 2x + 4y - 3z &= -9 \end{aligned}$$

2. Find the equation of the parabola through the points $(0,3)$, $(1,4)$ and $(3,18)$.

3. Consider the points $(1,5)$, $(2,2)$, $(4,3)$ and $(5,4)$.

   (a) What is the smallest degree polynomial whose graph will contain all of these points?

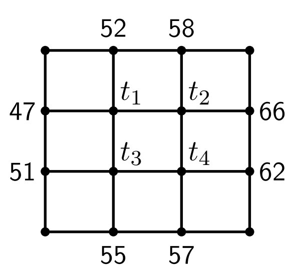
   (b) Find the polynomial whose graph contains all the points.

   (c) Check by graphing on your calculator.

4. Why would we not be able to find the equation of a line through $(0,6)$, $(2,3)$ and $(6,1)$? We will see later what this means in terms of systems of equations, and we will resolve the problem in a reasonable way.

5. Find the equation of the plane through the three points $P_1(4,1,-3)$, $P_2(0,-5,1)$ and $P_3(3,3,2)$.

6. (a) A student is attempting to find the equilibrium temperatures at points $t_1$, $t_2$, $t_3$ and $t_4$ on a plate with a grid and boundary temperatures shown below and to the left. They get $t_1 = 50.3$, $t_2 = 67.4$, $t_3 = 53.6$, $t_4 = 60.5$. Explain in one *complete sentence* why their answer must be incorrect, *without finding the solution*.




$$\left[\begin{array}{rrrrrrr} 4 & -1 & 0 & 0 & -1 & 0 & 103 \\ -1 & 4 & -1 & 0 & 0 & -1 & 92 \\ 0 & -1 & 4 & -1 & -1 & 0 & 110 \\ 0 & 0 & -1 & 4 & 0 & 0 & 98 \\ -1 & 0 & -1 & 0 & 4 & 0 & 105 \\ 0 & -1 & 0 & 0 & -1 & 4 & 107 \end{array}\right]$$

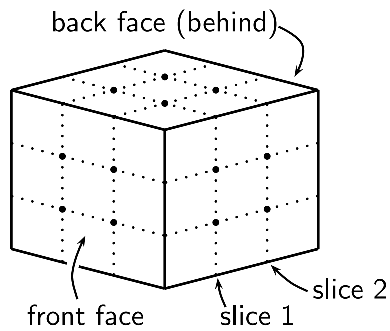
   (b) A different student is trying to solve another such problem, and their augmented matrix is shown above and to the right. How do we know that one of their equations is incorrect, *without setting up the equations ourselves?*

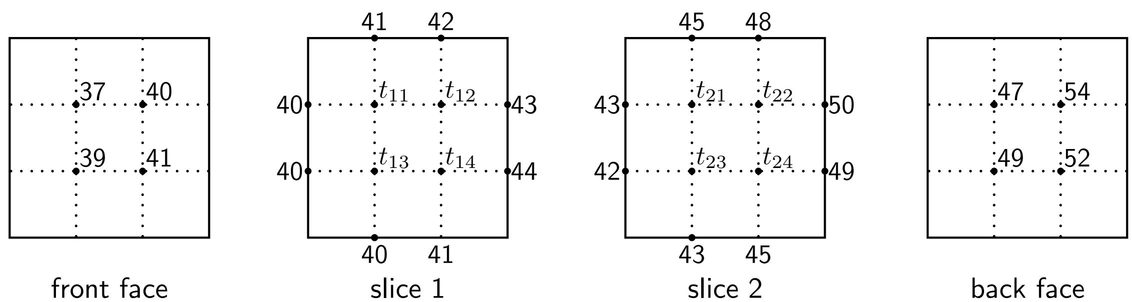
7. Suppose we are solving a system of three equations in the three unknowns $x_1$, $x_2$ and $x_3$, *with the unknowns showing up in the equations in that order*. It is possible to do row reduction in such a way as to obtain the matrix

$$\left[\begin{array}{rrrr} 1 & 0 & 0 & 5 \\ 3 & -2 & 0 & 7 \\ -1 & 5 & 2 & -3 \end{array}\right]$$

   Determine $x_1$, $x_2$ and $x_3$ *without row-reducing this matrix!* you should be able to simply set up equations and find values for the unknowns.

8. Given a cube of some solid material, it is possible to put a three-dimensional grid into the solid, in the same way that we put a two-dimensional grid on a rectangular plate. Given temperatures at all nodes on the exterior faces of the cube, we can find equilibrium temperatures at each interior node using a system of equations. Once again the key is the mean-value property. In this three dimensional case this property tells us that the equilibrium temperature at each interior node is equal to the average of all the temperatures at nodes of the grid that are immediately adjacent to the point in question. To the right I have shown a cube that has eight interior grid points. The word "slice" is used here to mean a cross section through the cube. The grids below show temperatures, known or unknown, at all nodes on the front face, each of the two slices, and the back face. Above and to the right I have "exploded" the cube to show the temperatures on the front and back faces, and the two slices. Of course each node on any slice is connected to the corresponding node on the adjacent slice or face.




(a) Using the Mean Value Property in three dimensions, the temperature at each interior point will *NOT* be the average of four temperatures, like it was on a plate. How many temperatures will be averaged in this case?

(b) Set up a system of equations to solve for the interior temperatures, and find each to the nearest tenth.

9. Do any of your observations from Exercise 7 change in the three dimensional case?

10. Do one of the following for each of the systems whose augmented matrices row reduce to the forms shown below. **Assume that the unknowns are $x_1$, $x_2$, ...**

- If the system has a unique solution, give it. If the system has no solution, say so.
- If the system has infinitely many solutions, give the general solution in terms of parameters $s$, $t$, etc., then give two particular solutions.

(a) $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & -1 & 0 & -4 \\ 0 & 1 & 2 & 0 & 5 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

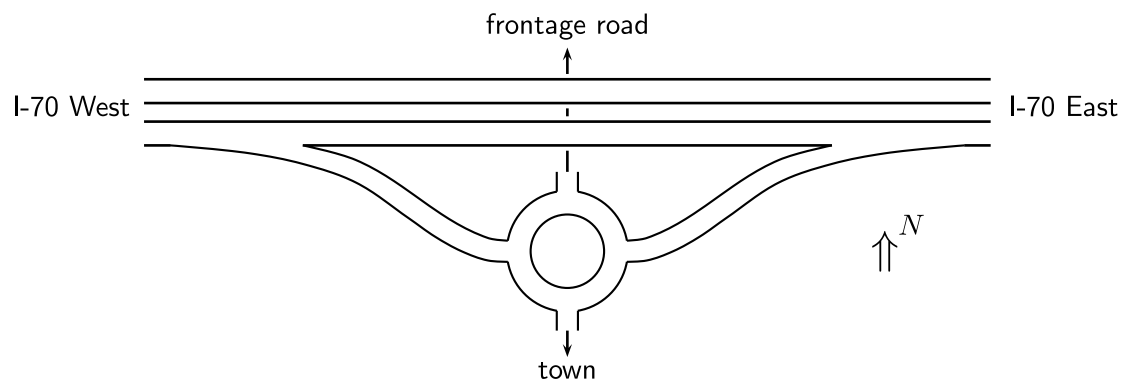
11. Consider the row-echelon augmented matrix $\begin{bmatrix} 1 & -1 & 3 & -2 & 4 \\ 0 & 0 & 1 & 2 & -5 \\ 0 & 0 & 0 & 2 & -8 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

(a) Give the general solution to the system of equations that this arose from.

(b) Give three specific solutions to the system.

(c) Change one entry in the matrix so that the system of equations would have no solution.

12. Vail, Colorado recently put in traffic "round-a-bouts" at all of its exits off Interstate 70. Each of these consists of a circle in which traffic is only allowed to flow counter-clockwise (do that all turns are right turns), and four points at which the circle can be entered or exited. See the diagram below.




It is known that at 7:30 AM the following is occurring:

- 22 vehicles per minute are entering the roundabout from the west. (These are the workers who cannot afford to live in Vail, and commute on I-70 from towns 30 and 40 miles west.)
- 4 vehicles per minute are exiting the roundabout to go east on I-70. (These are the tourists headed to the airport in Denver.)
- 7 vehicles per minute are exiting the roundabout toward town and 11 per minute are exiting toward the frontage road.

Solve the system and answer the following:

(a) What is the minimum number of cars per minute passing through the southeast quarter of the roundabout?

(b) If 18 vehicles per minute are passing through the southeast (SE) quarter of the roundabout per minute, how many are passing through each of the other quarters (NW, NE, SW)?

13. Consider the system
$$\begin{aligned} x_1 - x_2 - 4x_3 &= 6 \\ 5x_1 + x_2 - 2x_3 &= 18 \\ 2x_1 + 4x_2 + 10x_3 &= 0 \end{aligned}.$$

(a) Use your calculator or an online tool to reduce the matrix to reduced row echelon form. Write the system of two equations represented by the first two rows of the reduced matrix. (The last equation is of no use, so don't bother writing it.)

(b) The second equation contains $x_2$ and $x_3$. Suppose that $x_3 = 1$ and compute $x_2$ using that equation. Then use the values you have for $x_2$ and $x_3$ in the first equation to find $x_1$.

(c) Verify that the values you obtained in (b) are in fact a solution to the original system given.

(d) Now let $x_3 = 0$ and repeat the process from (b) to obtain another solution. Verify that solution as well.

(e) Let $x_3 = 2$ to find yet another solution.

(f) Because there is no equation allowing us to determine $x_3$, we say that it is a **free variable**. What we will usually do in situations like this is let $x_3$ equal some **parameter** (number) that we will denote by $t$. That is, we set $x_3 = t$, which is really just renaming it. Substitute $t$ into the second equation from (a) and solve for $x_2$ in terms of $t$. Then substitute that result into the first equation for $x_2$, along with $t$ for $x_3$, and solve for $x_1$ in terms of $t$. Summarize by giving each of $x_1$, $x_2$ and $x_3$ in terms of $t$, all in one place.

(g) Substitute the number one for $t$ into your answer to (f) and check to see that it matches what you got for (b). If doesn't, you've gone wrong somewhere - find the error and fix it.

14. Solve each of the following systems of equations that have solutions. Do/show the following things:

- Enter the augmented matrix for the system in your calculator.
- Get the row-reduced echelon form of the matrix using the *rref* command. **Write down the resulting matrix**.
- Write the system of equations that is equivalent to the row-reduced echelon matrix.
- Find the solutions, if there are any. *Use the letters that were used in the original system for the unknowns!* For those with multiple solutions, give them in terms of a parameter $t$ or, when necessary, two parameters $s$ and $t$.

(a)
$$\begin{aligned} x_1 - x_2 + 3x_3 &= -4 \\ -2x_1 + 3x_2 - 8x_3 &= 13 \\ 5x_1 - 3x_2 + 11x_3 &= -10 \end{aligned}$$

(b)
$$\begin{aligned} c_1 + 3c_2 + 5c_3 &= 3 \\ 2c_1 + 7c_2 + 9c_3 &= 5 \\ 2c_1 + 6c_2 + 11c_3 &= 7 \end{aligned}$$

(c)
$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= -1 \\ -7x_1 - 21x_2 + 14x_3 &= 7 \\ 2x_1 + 6x_2 - 4x_3 &= -2 \end{aligned}$$

(d)
$$\begin{aligned} c_1 - c_2 + 3c_3 &= -4 \\ -2c_1 + 3c_2 - 8c_3 &= 13 \\ 5c_1 - 3c_2 + 11c_3 &= 4 \end{aligned}$$

(e)
$$\begin{aligned} x - 3y + 7z &= 4 \\ 5x - 14y + 42z &= 29 \\ -2x + 5y - 20z &= -16 \end{aligned}$$

(f)
$$\begin{aligned} x + 3y \phantom{{}+z} &= 2 \\ 4x + 12y + z &= 1 \\ -x - 3y - 2z &= 12 \end{aligned}$$

- Give three *specific* solutions to the system from part (a) above.
- Give three *specific* solutions to the system from part (c) above.
- Solve the system from part (b) above by hand, showing all steps of the row reduction and indicating what you did at each step.

15. Give the reduced row echelon form of an augmented matrix for a system of four equations in four unknowns $x_1$, $x_2$, $x_3$ and $x_4$ for which

- $x_4 = 7$
- $x_2$ and only $x_2$, is a free variable

16. (Erdman) Consider the system $\begin{array}{rcl} x + ky &=& 1 \\ kx + \ y &=& 1 \end{array}$, where $k$ is some constant.

   (a) Set up the augmented matrix and use a row operation to get a zero in the lower left corner.

   (b) For what value or values of $k$ would the system have infinitely many solutions? What is the form of the general solution?

   (c) For what value or values of $k$ would the system have no solution?

   (d) For all remaining values of $k$ the system has a unique solution (that depends on the choice of $k$). What is the solution in that case? Your answer will contain the parameter $k$.

17. (Erdman) Consider the system $\begin{array}{rcl} x - \ y - 3z &=& 3 \\ 2x + \qquad\quad z &=& 0 \\ 2y + 7z &=& c \end{array}$, where $c$ is some constant.

   (a) Set up the augmented matrix and use a row operation to get a zero in the first entry of the second row.

   (b) Look at the second and third rows. For what value or values of $c$ can the system be solved? Give the solution if there is a unique solution. Give the general solution if there are infinitely many solutions.

18. In a previous exercise, you may have attempted to find the equation of a parabola through the three points $(-1, -6)$, $(0, -4)$ and $(1, -1)$. You set up a system to find values of $a$, $b$ and $c$ in the parabola equation $y = ax^2 + bx + c$. There was a unique solution, meaning that there is only one parabola containing those three points. Expect the following to not work out as "neatly."

   (a) Use a system of equations to find the equation of a parabola that goes through just the two points $(-1, -6)$ and $(1, -1)$. Explain your results.

   (b) Use a system of equations to find the equation $y = mx + b$ of a line through the four points $(1.3, 1.5)$, $(0.8, 0.4)$, $(2.6, 3.0)$ and $(2.0, 2.0)$.

   (c) Plot the four points from (b) on a neat and accurate graph, and use what you see to explain your answer to (b). **You should be able to give your explanation in one or two complete sentences.**

# 2 Euclidean Space and Vectors

**Outcome:**

2. Understand vectors and their algebra and geometry in $\mathbb{R}^n$. Understand the relationship of vectors to systems of equations.

   **Performance Criteria:**

   (a) Recognize the equation of a plane in $\mathbb{R}^3$ and determine where the plane intersects each of the three axes. Sketch a graph of the part of a plane in the first quadrant. Give the equation of a plane from a geometric description.

   (b) Find the vector from one point to another in $\mathbb{R}^n$. Find the length of a vector in $\mathbb{R}^n$.

   (c) Multiply vectors by scalars and add vectors, algebraically. Find linear combinations of vectors algebraically.

   (d) Illustrate the parallelogram method and tip-to-tail method for finding a linear combination of two vectors.

   (e) Find a linear combination of vectors equalling a given vector.

   (f) Give the linear combination form of a system of equations, give the system of linear equations equivalent to a given vector equation.

   (g) Sketch a picture illustrating the linear combination form of a system of equations of two equations in two unknowns.

   (h) Give an algebraic description of a set of a set of vectors that has been described geometrically, and vice-versa.

   (i) Determine whether a set of vectors is closed under vector addition; determine whether a set of vectors is closed under scalar multiplication. If it is, prove that it is; if it is not, give a counterexample.

   (j) Give the vector equation of a line through two points in $\mathbb{R}^2$ or $\mathbb{R}^3$ or the vector equation of a plane through three points in $\mathbb{R}^3$.

   (k) Write the solution to a system of equations in vector form and determine the geometric nature of the solution.

In the study of linear algebra we will be defining new concepts and developing corresponding notation. *The purpose of doing so is to develop more powerful machinery for investigating the concepts of interest.* We begin the development of notation with the following. The set of all real numbers is denoted by $\mathbb{R}$, and the set of all ordered pairs of real numbers is $\mathbb{R}^2$, spoken as "R-two." Geometrically, $\mathbb{R}^2$ is the familiar Cartesian coordinate plane. Similarly, the set of all ordered triples of real numbers is the three-dimensional space referred to as $\mathbb{R}^3$, "R-three." The set of all ordered ***n*-tuples** (lists of $n$ real numbers in a particular order) is denoted by $\mathbb{R}^n$. Although difficult or impossible to visualize physically, $\mathbb{R}^n$ can be thought of as $n$-dimensional space. All of the $\mathbb{R}^n$s are what are called **Euclidean space**.

## 2.1 Euclidean Space

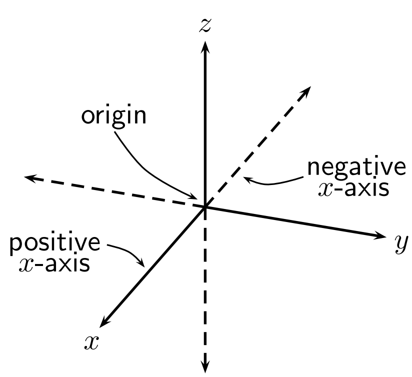
**Performance Criteria:**

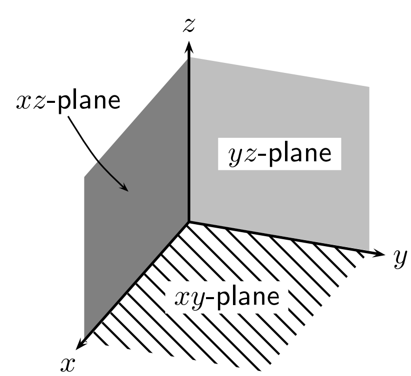
2. (a) Recognize the equation of a plane in $\mathbb{R}^3$ and determine where the plane intersects each of the three axes. Sketch a graph of the part of a plane in the first quadrant. Give the equation of a plane from a geometric description.

It is often taken for granted that everyone knows what we mean by the **real numbers**. To actually define the real numbers precisely is a bit of a chore and very technical. Suffice it to say that the real numbers include all numbers other than complex numbers (numbers containing $\sqrt{-1} = i$ or, for electrical engineers, $j$) that a scientist or engineer is likely to run into. The numbers $5$, $-31.2$, $\pi$, $\sqrt{2}$, $\frac{2}{7}$, and $e$ are all real numbers. We denote the set of all real numbers with the symbol $\mathbb{R}$, and the geometric representation of the real numbers is the familiar **real number line**, a horizontal line on which every real number has a place. This is possible because the real numbers are ordered: given any two real numbers, either they are equal to each other, one is less than the other, or vice-versa.

As mentioned previously, the set $\mathbb{R}^2$ is the set of all ordered pairs of real numbers. Geometrically, every such pair corresponds to a point in the **Cartesian plane**, which is the familiar $xy$-plane. $\mathbb{R}^3$ is the set of all ordered triples, each of which represents a point in three-dimensional space. We can continue on - $\mathbb{R}^4$ is the set of all ordered "4-tuples", and can be thought of geometrically as four dimensional space. Continuing further, an "$n$-tuple" is $n$ real numbers, in a specific order; each $n$-tuple can be thought of as representing a point in $n$-dimensional space. These spaces are sometimes called "two-space," "three-space" and "$n$-space" for short.

Two-space is fairly simple, with the only major features being the two axes and the four quadrants that the axes divide the space into. Three-space is a bit more complicated. Obviously there are three coordinate axes instead of two. In addition to those axes, there are also three coordinate planes as well, the $xy$-plane, the $xz$-plane and the $yz$-plane. Finally the three coordinate planes divide the space into eight **octants**. The pictures below illustrate the coordinate axes and planes. The first octant is the one we are looking into, where all three coordinates are positive. It is not important that we know the numbering of the other octants.




*Every* plane in $\mathbb{R}^3$ (we will be discussing only $\mathbb{R}^3$ for now) consists of a set of points that behave in an orderly mathematical manner, described here:

**Equation of a Plane in $\mathbb{R}^3$:** The ordered triples corresponding to all the points in a plane satisfy an equation of the form

$$ax + by + cz = d, \tag{1}$$

where $a$, $b$, $c$ and $d$ are constants, and not all of $a$, $b$ and $c$ are zero.

Note that equation (1) is a linear equation! The equation of any line in $\mathbb{R}^2$ can be written in the form $ax + by = c$, which is also a linear equation. A line is a one-dimensional object and a plane is a two-dimensional object, in a sense that we will see later. $ax + by = c$ then describes a one-dimensional "straight" object in the two-dimensional space $\mathbb{R}^2$, and (1) describes a two-dimensional "flat" object in $\mathbb{R}^3$. The corresponding equation in $\mathbb{R}^n$ is

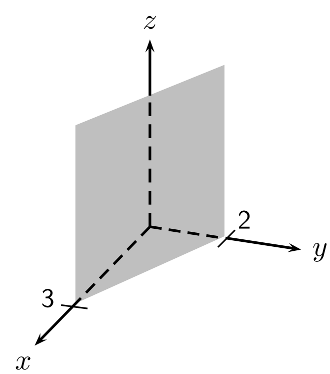
$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b, \tag{2}$$

where all of $a_1$, $a_2$, ..., $a_n$ and $b$ are numbers and $x_1$, $x_2$, ..., $x_n$ are unknowns. (2) describes an $n-1$-dimensional object in $n$-dimensional space. When $n > 3$ we often call such an object a **hyperplane**.

Going back to (1) as describing a plane in $\mathbb{R}^3$, the $xy$-plane in $\mathbb{R}^3$ is the plane containing the $x$ and $y$-axes. The only condition on points in that plane is that $z = 0$, so that is the equation of that plane. In that case the constants $a$, $b$ and $d$ are all zero, and $c = 1$. The plane $z = 5$ is a horizontal plane that is five units above the $xy$-plane, and $x = -2$ describes the vertical plane that is parallel to the $yz$-plane and passes through the $x$-axis at $x = -2$. What about an equation of the form $ax + by = c$ when we are in $\mathbb{R}^3$?

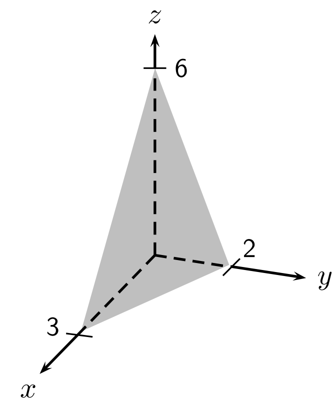
◇ **Example 2.1(a):** Graph the equation $2x + 3y = 6$ in the first octant. Indicate clearly where it intersects each of the coordinate axes, if it does.

**Solution:** Some points that satisfy the equation are $(3, 0, 0)$, $(6, -2, 5)$, and so on. Since $z$ is not included in the equation, there are no restrictions on $z$; it can take any value. If we were to fix $z$ at zero and plot all points that satisfy the equation, we would get a line in the $xy$-plane through the two points $(3, 0, 0)$ and $(0, 2, 0)$. These points are obtained by first letting $y$ and $z$ be zero, then by letting $x$ and $z$ be zero. Since $z$ can be anything, the set of points satisfying $2x + 3y = 6$ is a vertical plane intersecting the $xy$-plane in that line. The plane is shown to the right.




Next we'll see how we can sometimes graphically represent a portion of a plane in 3-space.

◇ **Example 2.1(b):** Graph the equation $2x + 3y + z = 6$ in the first octant. Indicate clearly where it intersects each of the coordinate axes, if it does.




**Solution:** The set of points satisfying this equation is also a plane, but $z$ is no longer free to take any value. The simplest way to "get a handle on" such a plane is to find where it intercepts the three axes. Note that every point on the $x$-axis has $y$- and $z$-coordinates of zero. So to find where the plane intersects the $x$-axis we put zero into the equation for $y$ and $z$, then solve for $x$, getting $x = 3$. The plane then intersects the $x$-axis at $(3, 0, 0)$. A similar process gives us that the plane intersects the $y$ and $z$ axes at $(0, 2, 0)$ and $(0, 0, 6)$. From this information we get that the graph of the plane is that shown in the drawing above and to the right.

---

Consider now a system of equations like

$$\begin{aligned} x + 3y - 2z &= -4 \\ 3x + 7y + z &= 4 \\ -2x + y + 7z &= 7 \end{aligned} ,$$

which has solution $(3, -1, 2)$. We now know that each of the three equations represents a plane in $\mathbb{R}^3$. The point $(3, -1, 2)$ is where the three planes intersect! This is completely analogous to the interpretation of the solution of a system of two linear equations in two unknowns as the point where the two lines representing the equations cross. This is the first of three interpretations we'll have for the solution to a system of equations.

The only other basic geometric fact we need about three-space is this:

**Distance Between Points:** The distance between two points $(x_1, y_1, z_1)$ and $(x_2, y_2, z_2)$ in $\mathbb{R}^3$ is given by

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

This is simply a three-dimensional version of the Pythagorean Theorem, and an equivalent formula is used in higher dimensions. Even though we cannot visualize the distance geometrically, this idea is both mathematically valid and useful in applications.

◇ **Example 2.1(c):** Find the distance in $\mathbb{R}^4$ between the points $(-4, 7, 1, -5)$ and $(13, 0, -6, 2)$.

**Solution:** Using the above formula with one more dimension we get

$$\begin{aligned} d &= \sqrt{(-4 - 13)^2 + (7 - 0)^2 + (1 - (-6))^2 + (-5 - 2)^2} \\ &= \sqrt{(-17)^2 + 7^2 + 7^2 + (-7)^2} \;=\; \sqrt{436} \;\approx\; 20.9 \end{aligned}$$

---

## Section 2.1 Exercises

To Solutions

1. Determine whether each of the equations given describes a plane in $\mathbb{R}^3$. If not, say so. If it does describe a plane, give the points where it intersects each axis. If it doesn't intersect an axis, say so.

   (a) $-2x - y + 3z = -6$ (b) $x + 3z = 6$ (c) $y = -6$

   (d) $x + 3z^2 = 12$ (e) $x - 2y + 3z = -6$

2. (a) Give an equation of the plane in $\mathbb{R}^3$ that intersects the $x$-axis at $2$, the $y$-axis at $5$, and the $z$-axis at $4$.

   (b) Give an equation of the plane in $\mathbb{R}^3$ that intersects the $x$-axis at $-3$, the $y$-axis at $1$, and the $z$-axis at $7$.

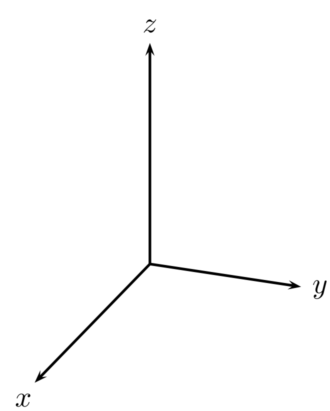
   (c) Give an equation of the plane in $\mathbb{R}^3$ that intersects the $x$-axis at $3$, the $z$-axis at $-2$, and does not intersect the $y$-axis.

   (d) Give an equation of the plane in $\mathbb{R}^3$ that intersects the $y$-axis at $4$ and does not intersect either of the other two axes.

   (e) Give an equation of the plane in $\mathbb{R}^3$ that intersects the $y$-axis at $-4$, the $z$-axis at $1$, and does not intersect the $x$-axis.

   (f) Give an equation of the plane in $\mathbb{R}^3$ that intersects the $z$-axis at $-2$ and does not intersect either of the other two axes.

3. Each of the following equations describes a plane in $\mathbb{R}^3$. For each, sketch the graph of the part of the plane in the first octant, in the manner done in Examples 2.1(a) and (b). Begin with a sketch of the positive parts of the three coordinate axes, as shown to the right.




   (a) $2x + 2y + 5z = 10$ (b) $2y + 3z = 6$

   (c) $x + 3y = 6$ (d) $2x + 4y + z = 8$

   (e) $z = 3$ (f) $3x + y + 2z = 6$

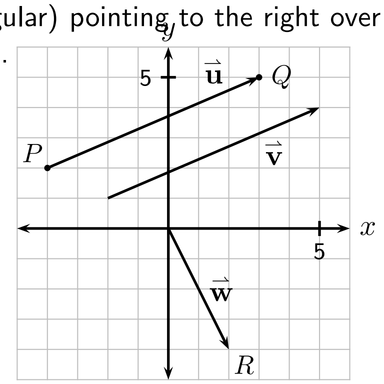
   (g) $2x + 4z = 8$ (h) $y = 5$ (i) $3x + y + 3z = 3$

## 2.2 Introduction to Vectors

**Performance Criteria:**

2. (b) Find the vector from one point to another in $\mathbb{R}^n$. Find the length of a vector in $\mathbb{R}^n$.

There are a number of different ways of thinking about **vectors**; it is likely that you think of them as "arrows" in two- or three-dimensional space, which is the first concept of a vector that most people have. Each such arrow has a **length** (sometimes called **norm** or **magnitude**) and a direction. Physically, then, vectors represent quantities having both amount and direction. Examples would be things like force or velocity. Quantities having only amount, like temperature or pressure, are referred to as **scalar** quantities. We will represent scalar quantities with lower case italicized letters like $a, b, c, ..., x, y, z$ and we'll represent vectors with lower case boldface letters with "harpoon" arrows over them, like $\vec{\mathbf{u}}, \vec{\mathbf{v}}, \vec{\mathbf{x}}$, and so on.. In many texts vectors are denoted just by lower case boldface letters like $\mathbf{u}$, $\mathbf{v}$, $\mathbf{x}$, and so on. When writing by hand we'll just put a small arrow (harpoon or regular) pointing to the right over any letter denoting a vector, without trying to make the letter boldface.




Consider a vector represented by an arrow in $\mathbb{R}^2$. We will call the end with the arrowhead the **tip** of the vector, and the other end we'll call the **tail**. (The more formal terminology is **terminal point** and **initial point**.) The picture to the right shows three vectors $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ in $\mathbb{R}^2$. It should be clear that a vector can be moved around in $\mathbb{R}^2$ in such a way that the direction and magnitude remain unchanged. Sometimes we say that two vectors related to each other this way are equivalent, but in this class we will simply say that they are the same vector. The vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are the same vector, just in different positions.

It is sometimes convenient to denote points with letters, and we use italicized capital letters for this. We commonly use $P$ (for point!) $Q$ and $R$, and the origin is denoted by $O$. (That's capital "oh," not zero.) Sometimes we follow the point immediately by its coordinates, like $P(-4, 2)$. The notation $\overrightarrow{PQ}$ denotes the vector that goes from point $P$ to point $Q$, which in this case is vector $\vec{\mathbf{u}}$. Any vector $\overrightarrow{OR}$ with its tail at the origin is said to be in **standard position**, and is called a **position vector**; $\vec{\mathbf{w}}$ above is an example of such a vector. Note that for any point in $\mathbb{R}^2$ (or in $\mathbb{R}^n$), there is a corresponding vector that goes from the origin to that point. *In linear algebra we think of a point and its position vector as interchangeable.* In the next section you will see the advantage of thinking of a point as a position vector.

We will describe vectors with numbers - in $\mathbb{R}^2$ we give a vector as two numbers, the first telling how far to the right (positive) or left (negative) one must go to get from the tail to the tip of the vector, and the second telling how far up (positive) or down (negative) from tail to tip. These numbers are generally arranged in one of two ways. The first way is like an ordered pair, but with "square brackets" instead of parentheses. The vector $\vec{\mathbf{u}}$ above is then $\vec{\mathbf{u}} = [7, 3]$, and $\vec{\mathbf{w}} = [2, -4]$. The second way to write a vector is as a **column vector**: $\vec{\mathbf{u}} = \begin{bmatrix} 7 \\ 3 \end{bmatrix}$. This is, in fact, the form we will use more often. Observe that the vector from one point to another is obtained by subtracting the corresponding coordinates of the first point from those of the second point. So the vector $\overrightarrow{PQ}$ from $P(-4, 2)$ to $Q(3, 5)$ is $\overrightarrow{PQ} = \begin{bmatrix} 3-(-4) \\ 5-2 \end{bmatrix} = \begin{bmatrix} 7 \\ 3 \end{bmatrix}$.

The two numbers quantifying a vector in $\mathbb{R}^2$ are called the **components** of the vector. We generally use the same letter to denote the components of a vector as the one used to name the vector, but we distinguish them with subscripts. Of course the components are scalars, so we use italic letters for them. So we would have $\vec{\mathbf{u}} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\vec{\mathbf{v}} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. The direction of a vector in $\mathbb{R}^2$ is given, in some sense, by the combination of the two components. The length is found using the Pythagorean theorem. For a vector $\vec{\mathbf{v}} = [v_1, v_2]$ we denote and define the length of the vector by $\| \vec{\mathbf{v}} \| = \sqrt{v_1^2 + v_2^2}$. Of course everything we have done so far applies to vectors in higher dimensions. A vector $\vec{\mathbf{x}}$ in $\mathbb{R}^n$ would be denoted by $\vec{\mathbf{x}} = [x_1, x_2, ..., x_n]$. This shows that, in some sense, a **vector** is just an ordered list of numbers, like an $n$-tuple but with differences you will see in the next section. The length of a vector in $\mathbb{R}^n$ is found as follows.

<u>DEFINITION 2.2.1</u>: **Magnitude of a Vector in $\mathbb{R}^n$**

The **magnitude**, or **length** of a vector $\vec{\mathbf{x}} = [x_1, x_2, ..., x_n]$ in $\mathbb{R}^n$ is given by

$$\| \vec{\mathbf{x}} \| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

◇ **Example 2.2(a):** Find the vector $\vec{\mathbf{x}} = \overrightarrow{PQ}$ in $\mathbb{R}^3$ from the point $P(5, -3, 7)$ to $Q(-2, 6, 1)$, and find the length of the vector.

**Solution:** The components of $\vec{\mathbf{x}}$ are obtained by simply subtracting each coordinate of $P$ from each coordinate of $Q$:

$$\vec{\mathbf{x}} = \overrightarrow{PQ} = \begin{bmatrix} -2-5 \\ 6-(-3) \\ 1-7 \end{bmatrix} = \begin{bmatrix} -7 \\ 9 \\ -6 \end{bmatrix}$$

The length of $\vec{\mathbf{x}}$ is

$$\| \vec{\mathbf{x}} \| = \sqrt{(-7)^2 + 9^2 + (-6)^2} = \sqrt{166} \approx 12.9$$

There will be times when we need a vector with length zero; this is the special vector we will call (surprise, surprise!) the **zero vector**. It is denoted by a boldface zero, $\vec{\mathbf{0}}$, to distinguish it from the scalar zero. This vector has no direction.

Let's finish with the following important note about how we will work with vectors in this class:

In this course, when working with vectors geometrically, we will almost always be thinking of them as position vectors. When working with vectors algebraically, we will always consider them to be column vectors.

**Section 2.2 Exercises** To Solutions

1. Find the magnitude of each vector - label each answer using correct notation.

$$\vec{\mathbf{u}} = \begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix}, \qquad \vec{\mathbf{x}} = \begin{bmatrix} -5 \\ 7 \end{bmatrix}, \qquad \vec{\mathbf{v}} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \qquad \vec{\mathbf{b}} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

2. For each pair of points $P$ and $Q$, find the vector $\overrightarrow{PQ}$ in the appropriate space. Then find the length of the vector.

   (a) $P(-4, 11, 7),\ Q(13, 5, -8)$

   (b) $P(-5, 1),\ Q(7, -2)$

   (c) $P(-3, 0, 6, 1),\ Q(7, -1, -1, 10)$

3. (a) The vector $\overrightarrow{PQ} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ in $\mathbb{R}^2$ has initial point $P(3, 5)$. What is the terminal point $Q$?

   (b) The vector $\overrightarrow{PQ} = [4, 0, -2, 1, 5]$ in $\mathbb{R}^5$ has initial point $P(-2, 7, 1, -8, 2)$. What is the terminal point $Q$?

   (c) The vector $\overrightarrow{PQ} = \begin{bmatrix} -4 \\ 6 \\ -1 \end{bmatrix}$ in $\mathbb{R}^3$ has terminal point $Q(5, -2, 4)$. What is the initial point $P$?

4. Consider the vector $\vec{\mathbf{u}} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$ in $\mathbb{R}^3$.

   (a) Find the magnitude $\| \vec{\mathbf{u}} \|$, labeling it as such.

   (b) Later we will define what we mean by multiplication or division of a vector by a scalar. Divide each component of $\vec{\mathbf{u}}$ by your answer to (a). The result is the vector $\dfrac{\vec{\mathbf{u}}}{\| \vec{\mathbf{u}} \|}$, so label it that way.

   (c) Find $\left\| \dfrac{\vec{\mathbf{u}}}{\| \vec{\mathbf{u}} \|} \right\|$, the magnitude of $\dfrac{\vec{\mathbf{u}}}{\| \vec{\mathbf{u}} \|}$ from (b).

5. Repeat Exercise 4 for the vector $\vec{\mathbf{v}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$ in $\mathbb{R}^2$. How does your final result compare with that of Exercise 4?

## 2.3 Addition and Scalar Multiplication of Vectors, Linear Combinations

**Performance Criteria:**

2. (c) Multiply vectors by scalars and add vectors, algebraically. Find linear combinations of vectors algebraically.

   (d) Illustrate the parallelogram method and tip-to-tail method for finding a linear combination of two vectors.

   (e) Find a linear combination of vectors equalling a given vector.

In the previous section a vector $\vec{\mathbf{x}} = [x_1, x_2, ..., x_n]$ in $n$ dimensions was starting to look suspiciously like an $n$-tuple $(x_1, x_2, ..., x_n)$ and we established a correspondence between any point and the position vector with its tip at that point. One might wonder why we bother then with vectors at all! The reason is that we can perform algebraic operations with vectors that make sense physically, while such operations make no sense with $n$-tuples. The two most basic things we can do with vectors are add two of them or multiply one by a scalar, and both are done component-wise:

DEFINITION 2.3.1: **Addition and Scalar Multiplication of Vectors**

Let $\vec{\mathbf{u}} = [u_1, u_2, ..., u_n]$ and $\vec{\mathbf{v}} = [v_1, v_2, ..., n_n]$, and let $c$ be a scalar. Then we define the vectors $\vec{\mathbf{u}} + \vec{\mathbf{v}}$ and $c\vec{\mathbf{u}}$ by

$$\vec{\mathbf{u}} + \vec{\mathbf{v}} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} \quad \text{and} \quad c\vec{\mathbf{u}} = c\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{bmatrix}$$

*Note that result of adding two vectors or multiplying a vector by a scalar is also a vector.* It clearly follows from these that we can get subtraction of vectors by first multiplying the second vector by the scalar $-1$, then adding the vectors. With just a little thought you will recognize that this is the same as just subtracting the corresponding components.

◇ **Example 2.3(a):** For $\vec{\mathbf{u}} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$ and $\vec{\mathbf{v}} = \begin{bmatrix} -4 \\ 9 \\ 6 \end{bmatrix}$, find $\vec{\mathbf{u}} + \vec{\mathbf{v}}$, $3\vec{\mathbf{u}}$ and $\vec{\mathbf{u}} - \vec{\mathbf{v}}$.

**Solution:**

$$\vec{\mathbf{u}} + \vec{\mathbf{v}} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} + \begin{bmatrix} -4 \\ 9 \\ 6 \end{bmatrix} = \begin{bmatrix} 5 + (-4) \\ -1 + 9 \\ 2 + 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \\ 8 \end{bmatrix}$$

$$3\vec{\mathbf{u}} = 3\begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3(5) \\ 3(-1) \\ 3(2) \end{bmatrix} = \begin{bmatrix} 15 \\ -3 \\ 6 \end{bmatrix}$$

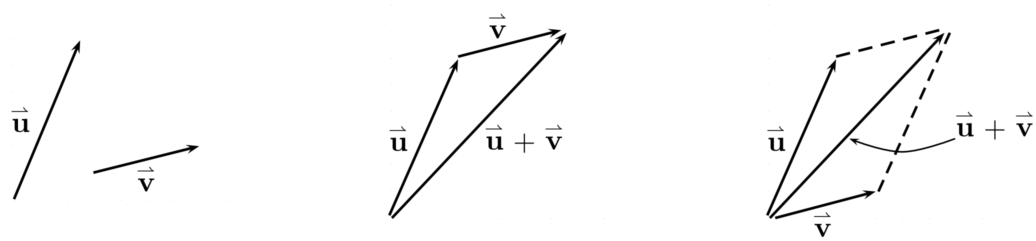
$$\vec{\mathbf{u}} - \vec{\mathbf{v}} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} - \begin{bmatrix} -4 \\ 9 \\ 6 \end{bmatrix} = \begin{bmatrix} 5-(-4) \\ -1-\ 9 \\ 2\ -\ 6 \end{bmatrix} = \begin{bmatrix} 9 \\ -10 \\ -4 \end{bmatrix}$$

---

Addition of vectors can be thought of geometrically in two ways, both of which are useful. The first way is what we will call the **tip-to-tail method**, and the second method is called the **parallelogram method**. You should become very familiar with both of these methods, as they each have their advantages; they are illustrated below.

◇ **Example 2.3(b):** Add the two vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ shown below and to the left, first by the tip-to-tail method, and second by the parallelogram method.

**Solution:** To add using the tip-to-tail method, move the second vector so that its tail is at the tip of the first. (Be sure that its length and direction remain the same!) The vector $\vec{\mathbf{u}} + \vec{\mathbf{v}}$ goes from the tail of $\vec{\mathbf{u}}$ to the tip of $\vec{\mathbf{v}}$. See in the middle below.




To add using the parallelogram method, put the vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ together at their tails (again being sure to preserve their lengths and directions). Draw a dashed line from the tip of $\vec{\mathbf{u}}$, parallel to $\vec{\mathbf{v}}$, and draw another dashed line from the tip of $\vec{\mathbf{v}}$, parallel to $\vec{\mathbf{u}}$. $\vec{\mathbf{u}} + \vec{\mathbf{v}}$ goes from the tails of $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ to the point where the two dashed lines cross. See to the right above. The reason for the name of this method is that the two vectors and the dashed lines create a parallelogram.

---

Each of these two methods has a natural physical interpretation. For the tip-to-tail method, imagine an object that gets *displaced* by the direction and amount shown by the vector $\vec{\mathbf{u}}$. Then suppose that it gets displaced by the direction and amount given by $\vec{\mathbf{v}}$ after that. Then the vector $\vec{\mathbf{u}} + \vec{\mathbf{v}}$ gives the *net* (total) displacement of the object. Now look at that picture for the parallelogram method, and imagine that there is an object at the tails of the two vectors. If we were then to have two forces acting on the object, one in the direction of $\vec{\mathbf{u}}$ and with an amount (magnitude) indicated by the length of $\vec{\mathbf{u}}$, and another with amount and direction indicated by $\vec{\mathbf{v}}$, then $\vec{\mathbf{u}} + \vec{\mathbf{v}}$ would represent the net force. (In a statics or physics course you might call this the **resultant** force.)

A very important concept in linear algebra is that of a **linear combination**. Let me say it again:

> Linear combinations are one of the most important concepts in linear algebra! You need to recognize them when you see them and learn how to create them. They will be central to almost everything that we will do from here on.

A linear combination of a set of vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, ..., \vec{\mathbf{v}}_n$ (note that the subscripts now distinguish different *vectors*, not the components of a single vector) is obtained when each of the vectors is multiplied by a scalar, and the resulting vectors are added up. So if $c_1, c_2, ..., c_n$ are the scalars that $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, ..., \vec{\mathbf{v}}_n$ are multiplied by, the resulting linear combination is the *single vector* $\vec{\mathbf{v}}$ given by

$$\vec{\mathbf{v}} = c_1 \vec{\mathbf{v}}_1 + c_2 \vec{\mathbf{v}}_2 + c_3 \vec{\mathbf{v}}_3 + \cdots + c_n \vec{\mathbf{v}}_n \, .$$

Emphasizing again the importance of this concept, let's provide a slightly more concise and formal definition:

> DEFINITION 2.3.2: **Linear Combination**
>
> A **linear combination** of the vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, ..., \vec{\mathbf{v}}_n$, all in $\mathbb{R}^n$, is any vector $\vec{\mathbf{v}}$ of the form
> $$\vec{\mathbf{v}} = c_1 \vec{\mathbf{v}}_1 + c_2 \vec{\mathbf{v}}_2 + c_3 \vec{\mathbf{v}}_3 + \cdots + c_n \vec{\mathbf{v}}_n,$$
> where $c_1, c_2, ..., c_n$ are scalars.

Note that when we create a linear combination of a set of vectors we are doing virtually everything possible algebraically with those vectors, which is just addition and scalar multiplication!

You have seen this idea before; every polynomial like $5x^3 - 7x^2 + \frac{1}{2}x - 1$ is a linear combination of $1, x, x^2, x^3, ....$ Those of you who have had a differential equations class have seen things like $\dfrac{d^2y}{dt^2} + 3\dfrac{dy}{dt} + 2y$, which is a linear combination of the second, first and "zeroth" derivatives of a function $y = y(t)$. Here is why linear combinations are so important: In many applications we seek to have a basic set of objects (vectors) from which all other objects can be built as linear combinations of objects from our basic set. A large part of our study will be centered around this idea. This may not make any sense to you now, but hopefully it will by the end of the course.

◇ **Example 2.3(c):** For the vectors $\vec{\mathbf{v}}_1 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} -4 \\ 9 \\ 6 \end{bmatrix}$ and $\vec{\mathbf{v}}_3 = \begin{bmatrix} 0 \\ 3 \\ 8 \end{bmatrix}$, give the linear combination $2\vec{\mathbf{v}}_1 - 3\vec{\mathbf{v}}_2 + \vec{\mathbf{v}}_3$ as one vector.

**Solution:**

$$\begin{aligned} 2\vec{\mathbf{v}}_1 - 3\vec{\mathbf{v}}_2 + \vec{\mathbf{v}}_3 &= 2\begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} - 3\begin{bmatrix} -4 \\ 9 \\ 6 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \\ 8 \end{bmatrix} \\ &= \begin{bmatrix} 10 \\ -2 \\ 4 \end{bmatrix} - \begin{bmatrix} -12 \\ 27 \\ 18 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \\ 8 \end{bmatrix} = \begin{bmatrix} -2 \\ -26 \\ 30 \end{bmatrix} \end{aligned}$$

◇ **Example 2.3(d):** For the same vectors $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$ and $\vec{\mathbf{v}}_3$ as in the previous exercise and scalars $c_1$, $c_2$ and $c_3$, give the linear combination $c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3$ as one vector.

**Solution:**

$$\begin{aligned} c_1 \vec{\mathbf{v}}_1 + c_2 \vec{\mathbf{v}}_2 + c_3 \vec{\mathbf{v}}_3 &= c_1 \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -4 \\ 9 \\ 6 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 3 \\ 8 \end{bmatrix} \\ &= \begin{bmatrix} 5c_1 \\ -1c_1 \\ 2c_1 \end{bmatrix} + \begin{bmatrix} -4c_2 \\ 9c_2 \\ 6c_2 \end{bmatrix} + \begin{bmatrix} 0c_3 \\ 3c_3 \\ 8c_3 \end{bmatrix} \\ &= \begin{bmatrix} 5c_1 - 4c_2 + 0c_3 \\ -1c_1 + 9c_2 + 3c_3 \\ 2c_1 + 6c_2 + 8c_3 \end{bmatrix} \end{aligned}$$

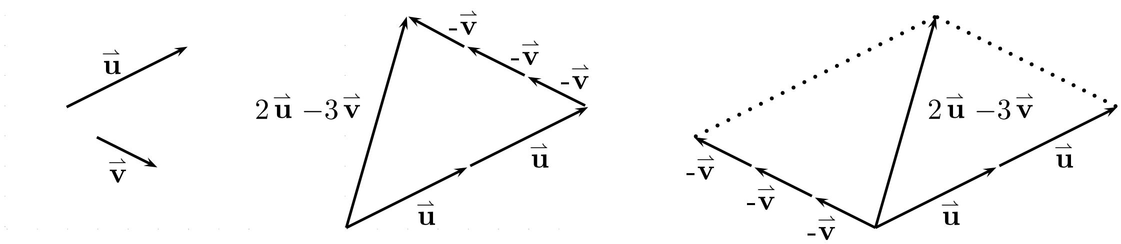
Note that the final result is a single vector with three components that look suspiciously like the left sides of a system of three equations in three unknowns!

---

In the previous two examples we found linear combinations algebraically - in the next example we find a linear combination geometrically.

◇ **Example 2.3(e):** In the space below and to the right, sketch the vector $2\vec{\mathbf{u}} - 3\vec{\mathbf{v}}$ for the vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ shown below and to the left.

**Solution:** In the center below the linear combination is obtained by the tip-to-tail method, and to the right below it is obtained by the parallelogram method.




---

The final example is probably the most important in this section.

◇ **Example 2.3(f):** Find a linear combination of the vectors $\vec{\mathbf{v}}_1 = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$ and $\vec{\mathbf{v}}_2 = \begin{bmatrix} 7 \\ -3 \end{bmatrix}$ that equals the vector $\vec{\mathbf{w}} = \begin{bmatrix} 1 \\ -14 \end{bmatrix}$.

**Solution:** We are looking for two scalars $c_1$ and $c_2$ such that $c_1 \vec{\mathbf{v}}_1 + c_2 \vec{\mathbf{v}}_2 = \vec{\mathbf{w}}$. By the method of Example 2.3(d) we have

$$c_1 \begin{bmatrix} 3 \\ -4 \end{bmatrix} + c_2 \begin{bmatrix} 7 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ -14 \end{bmatrix}$$

$$\begin{bmatrix} 3c_1 \\ -4c_1 \end{bmatrix} + \begin{bmatrix} 7c_2 \\ -3c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -14 \end{bmatrix}$$

$$\begin{bmatrix} 3c_1 + 7c_2 \\ -4c_1 - 3c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -14 \end{bmatrix}$$

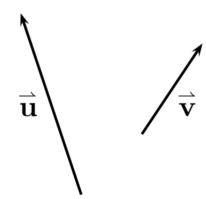
In the last line above we have two vectors that are equal. It should be intuitively obvious that this can only happen if the individual components of the two vectors are equal. This results in the system $\begin{aligned} 3c_1 + 7c_2 &= 1 \\ -4c_1 - 3c_2 &= -14 \end{aligned}$ of two equations in the two unknowns $c_1$ and $c_2$. Solving, we arrive at $c_1 = 5,\ c_2 = -2$. It is easily verified that these are correct:

$$5\begin{bmatrix} 3 \\ -4 \end{bmatrix} - 2\begin{bmatrix} 7 \\ -3 \end{bmatrix} = \begin{bmatrix} 15 \\ -20 \end{bmatrix} - \begin{bmatrix} 14 \\ -6 \end{bmatrix} = \begin{bmatrix} 1 \\ -14 \end{bmatrix}$$

We now conclude with an important observation. Suppose that we consider all possible linear combinations of a single vector $\vec{\mathbf{v}}$. That is then the set of all vectors of the form $c\,\vec{\mathbf{v}}$ for some scalar $c$, which is just all scalar multiples of $\vec{\mathbf{v}}$. At the risk of being redundant, the set of all linear combinations of a single vector is all scalar multiples of that vector.

**Section 2.3 Exercises** To Solutions

1. Illustrate the tip-to-tail and parallelogram methods for finding $\vec{\mathbf{w}} = -\vec{\mathbf{u}} + 2\vec{\mathbf{v}}$ for the two vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ shown to the right. Make it clear what portion of your diagram represents $\vec{\mathbf{w}}$ in each case.




2. For the vectors $\vec{\mathbf{v}}_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$, $\vec{\mathbf{v}}_3 = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$ and $\vec{\mathbf{v}}_4 = \begin{bmatrix} -8 \\ 1 \end{bmatrix}$, give the linear combination $5\vec{\mathbf{v}}_1 + 2\vec{\mathbf{v}}_2 - 7\vec{\mathbf{v}}_3 + \vec{\mathbf{v}}_4$ as one vector.

3. For the vectors $\vec{\mathbf{v}}_1 = \begin{bmatrix} -1 \\ 3 \\ -6 \end{bmatrix}$ and $\vec{\mathbf{v}}_2 = \begin{bmatrix} -8 \\ 1 \\ 4 \end{bmatrix}$, give the linear combination $c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2$ as one vector.

4. Give a linear combination of $\vec{\mathbf{u}} = \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}$ and $\vec{\mathbf{w}} = \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix}$ that equals $\begin{bmatrix} 17 \\ -4 \\ -9 \end{bmatrix}$.
Demonstrate that your answer is correct by filling in the blanks:

$$\underline{\quad}\begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} + \underline{\quad}\begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} + \underline{\quad}\begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix} = \begin{bmatrix} \underline{\quad} \\ \underline{\quad} \\ \underline{\quad} \end{bmatrix} + \begin{bmatrix} \underline{\quad} \\ \underline{\quad} \\ \underline{\quad} \end{bmatrix} + \begin{bmatrix} \underline{\quad} \\ \underline{\quad} \\ \underline{\quad} \end{bmatrix} = \begin{bmatrix} 17 \\ -4 \\ -9 \end{bmatrix}$$

5. For each of the following, find a linear combination of the vectors $\vec{\mathbf{u}}_1$, $\vec{\mathbf{u}}_2$,...,$\vec{\mathbf{u}}_n$ that equals $\vec{\mathbf{v}}$. Conclude by giving the actual linear combination, not just some scalars.

(a) $\vec{\mathbf{u}}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\vec{\mathbf{u}}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

(b) $\vec{\mathbf{u}}_1 = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 8 \\ -2 \end{bmatrix}$

(c) $\vec{\mathbf{u}}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} -4 \\ 1 \\ 2 \end{bmatrix}$, $\vec{\mathbf{u}}_3 = \begin{bmatrix} 5 \\ -3 \\ 2 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 6 \\ -18 \\ -7 \end{bmatrix}$

(d) $\vec{\mathbf{u}}_1 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{\mathbf{u}}_3 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 8 \\ -6 \end{bmatrix}$

(e) $\vec{\mathbf{u}}_1 = \begin{bmatrix} 7 \\ 1 \\ 3 \\ 0 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} -2 \\ 5 \\ 1 \\ -3 \end{bmatrix}$, $\vec{\mathbf{u}}_3 = \begin{bmatrix} 2 \\ 2 \\ -3 \\ 1 \end{bmatrix}$, $\vec{\mathbf{u}}_4 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 19 \\ 10 \\ -12 \\ 12 \end{bmatrix}$

(f) $\vec{\mathbf{u}}_1 = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} -1 \\ 1 \\ -4 \end{bmatrix}$, $\vec{\mathbf{u}}_3 = \begin{bmatrix} 3 \\ 1 \\ -8 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$

(g) $\vec{\mathbf{u}}_1 = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} -1 \\ 1 \\ -4 \end{bmatrix}$, $\vec{\mathbf{u}}_3 = \begin{bmatrix} 3 \\ 1 \\ -8 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} -1 \\ 3 \\ -14 \end{bmatrix}$

6. (a) Consider the vectors $\vec{\mathbf{u}}_1 = \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\vec{\mathbf{u}}_3 = \begin{bmatrix} -2 \\ 6 \\ 5 \end{bmatrix}$, $\vec{\mathbf{w}} = \begin{bmatrix} 11 \\ 5 \\ 8 \end{bmatrix}$.

   If possible, find scalars $a_1$, $a_2$ and $a_3$ such that $a_1\vec{\mathbf{u}}_1 + a_2\vec{\mathbf{u}}_2 + a_3\vec{\mathbf{u}}_3 = \vec{\mathbf{w}}$.

   (b) Consider the vectors $\vec{\mathbf{v}}_1 = \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\vec{\mathbf{v}}_3 = \begin{bmatrix} -7 \\ 2 \\ 5 \end{bmatrix}$, $\vec{\mathbf{w}} = \begin{bmatrix} 11 \\ 5 \\ 8 \end{bmatrix}$.

   If possible, find scalars $b_1$, $b_2$ and $b_3$ such that $b_1\vec{\mathbf{v}}_1 + b_2\vec{\mathbf{v}}_2 + b_3\vec{\mathbf{v}}_3 = \vec{\mathbf{w}}$.

   (c) To do each of parts (a) and (b) you should have solved a system of equations. Let $A$ be the coefficient matrix for the system in (a) and let $B$ be the coefficient matrix for the system in part (b). Use your calculator to find $det(A)$ and $\det(B)$, the determinants of matrices $A$ and $B$. You will probably find the command for the determinant in the same menu as *rref*.

## 2.4 Linear Combination Form of a System

**Performance Criterion:**

2. (f) Give the linear combination form of a system of equations, give the system of linear equations equivalent to a given vector equation.

   (g) Sketch a picture illustrating the linear combination form of a system of equations of two equations in two unknowns.

It should be clear that two vectors are equal if and only if their corresponding components are equal. That is,

$$\vec{\mathbf{u}} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \vec{\mathbf{v}} \qquad \text{if, and only if,} \qquad \begin{array}{c} u_1 = v_1, \\ u_2 = v_2, \\ \vdots \\ u_n = v_n \end{array}$$

The words "if, and only if" mean that the above works "both ways" in the following sense:

- If we have two vectors of length $n$ that are equal, then their corresponding entries are all equal, resulting in $n$ equations.
- If we have a set of $n$ equations, we can create a two vectors, one of whose components are all the left hand sides of the equations and the other whose components are all the right hand sides of the equations, and the two vectors created this way are equal.

Using the second bullet above, we can take the system of equations below and to the left and turn them into the single vector equation shown below and to the right:

$$\begin{array}{rcr} x_1 + 3x_2 - 2x_3 & = & -4 \\ 3x_1 + 7x_2 + x_3 & = & 4 \\ -2x_1 + x_2 + 7x_3 & = & 7 \end{array} \qquad \Longrightarrow \qquad \begin{bmatrix} x_1 + 3x_2 - 2x_3 \\ 3x_1 + 7x_2 + x_3 \\ -2x_1 + x_2 + 7x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \\ 7 \end{bmatrix}$$

We can take the vector on the left side of the equation and break it into three vectors to get

$$\begin{bmatrix} x_1 \\ 3x_1 \\ -2x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 7x_2 \\ x_2 \end{bmatrix} + \begin{bmatrix} -2x_3 \\ x_3 \\ 7x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \\ 7 \end{bmatrix}$$

and then we can factor the scalar unknown out of each vector to get the *vector* equation

$$x_1 \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 7 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \\ 7 \end{bmatrix} \tag{1}$$

Our previous geometric interpretation of solving

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= -4 \\ 3x_1 + 7x_2 + \ x_3 &= \ \ 4 \\ -2x_1 + \ x_2 + 7x_3 &= \ \ 7 \end{aligned}$$

was that we were looking for the point $(x_1, x_2, x_3)$ where the planes with equations $x_1+3x_2-2x_3 = -4$, $3x_1 + 7x_2 + \ x_3 = 4$ and $-2x_1 + \ x_2 + 7x_3 = 7$ intersect. (1) now gives us another interpretation - we are looking for the linear combination of the vectors

$$\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ 7 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix} \text{ that equals } \begin{bmatrix} -4 \\ 4 \\ 7 \end{bmatrix}$$

The form (1) of a system of equations is quite important, so we give it a definition:

<u>Definition 2.4.1</u> **Linear Combination Form of a System**

A system

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

of $m$ linear equations in $n$ unknowns can be written as a linear combination of vectors equalling another vector:

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

We will refer to this as the **linear combination form of the system of equations**.

Thus the system of equations below and to the left can be rewritten in the linear combination form shown below and to the right.

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= -4 \\ 3x_1 + 7x_2 + x_3 &= 4 \\ -2x_1 + x_2 + 7x_3 &= 7 \end{aligned} \qquad x_1 \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 7 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \\ 7 \end{bmatrix}$$

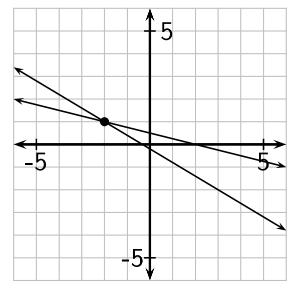
The question we originally asked for the system of linear equations was "Are there numbers $x_1$, $x_2$ and $x_3$ that make all three equations true?" Now we can see this is equivalent to a different question, "Is there a linear combination of the vectors $\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 7 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix}$ that equals the vector $\begin{bmatrix} -4 \\ 4 \\ 7 \end{bmatrix}$?"

◇ **Example 2.4(a):** Give the linear combination form of the system $\begin{array}{rcr} 3x_1 + 5x_2 & = & -1 \\ x_1 + 4x_2 & = & 2 \end{array}$ of linear equations.

**Solution:** The linear combination form of the system is $x_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

---

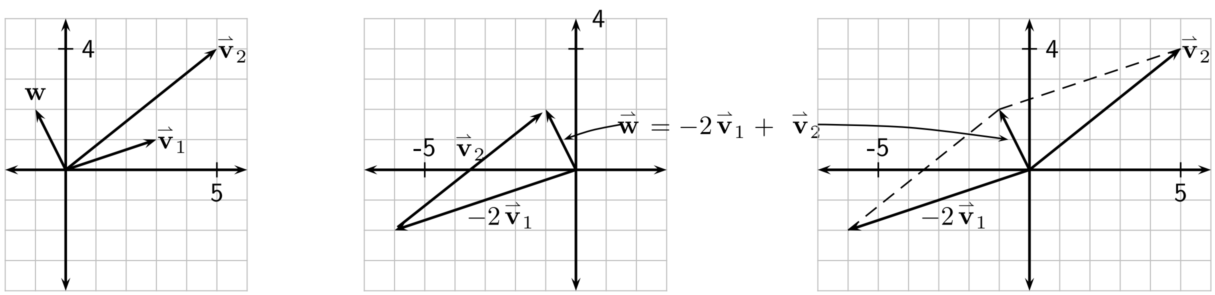
Let's consider the system from this last example a bit more. The goal is to solve the system of equations $\begin{array}{rcr} 3x_1 + 5x_2 & = & -1 \\ x_1 + 4x_2 & = & 2 \end{array}$. In the past our geometric interpretation has been this: The set of solutions to the first equation is a line in $\mathbb{R}^2$, and the set of solutions to the second equation is another line. The solution to the system happens to be $x_1 = -2,\ x_2 = 1$, and the point $(-2, 1)$ in $\mathbb{R}^2$ is the point where the two lines cross. This is shown in the picture to the right.




Now consider the linear combination form $x_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ of the system. Let $\vec{\mathbf{v}}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$ and $\vec{\mathbf{w}} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. These vectors are shown in the diagram to the left at the top of the next page. The solution $x_1 = -2,\ x_2 = 1$ to the system is the scalars that we can use for a linear combination of the vectors $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ to get the vector $\vec{\mathbf{w}}$. That is,

$$-2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

This is shown in the middle diagram below by the tip-to-tail method, and in the diagram below and to the right by the parallelogram method.




tip-to-tail method

parallelogram method

**Section 2.4 Exercises** To Solutions

1. Give the linear combination form of each system:

(a)
$$\begin{aligned} x + y - 3z &= 1 \\ -3x + 2y - z &= 7 \\ 2x + y - 4z &= 0 \end{aligned}$$

(b)
$$\begin{aligned} 5x_1 \qquad + x_3 &= -1 \\ 2x_2 + 3x_3 &= 0 \\ 2x_1 + x_2 - 4x_3 &= 2 \end{aligned}$$

(c)
$$\begin{aligned} b + 0.5m &= 8.1 \\ b + 1.0m &= 6.9 \\ b + 1.5m &= 6.2 \\ b + 2.0m &= 5.3 \\ b + 2.5m &= 4.5 \\ b + 3.0m &= 3.8 \\ b + 3.5m &= 3.0 \end{aligned}$$

(d)
$$\begin{aligned} x_1 - 4x_2 + x_3 + 2x_4 &= -1 \\ 3x_1 + 2x_2 - x_3 - 7x_4 &= 0 \\ -2x_1 + x_2 - 4x_3 + x_4 &= 2 \end{aligned}$$

2. Give the system of equations that is equivalent to each vector equation.

(a) $$x_1 \begin{bmatrix} -3 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

(b) $$x_1 \begin{bmatrix} 5 \\ 1 \\ -4 \\ -3 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 2 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 0 \\ 5 \\ 4 \end{bmatrix} + x_4 \begin{bmatrix} 7 \\ -4 \\ 6 \\ 7 \end{bmatrix} = \begin{bmatrix} -8 \\ 1 \\ 5 \\ 4 \end{bmatrix}$$

(c) $$x_1 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ -7 \\ 5 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ 3 \\ -4 \end{bmatrix}$$

3. The system of equations $\begin{aligned} 2x - 3y &= -6 \\ 3x - y &= 5 \end{aligned}$ has solution $x = 3$, $y = 4$. Write the system in linear combination form, then replace $x$ and $y$ with their values. Finally, sketch a picture illustrating the resulting vector equation. See the explanation after Example 2.4(a) if you have no idea what I am talking about.

## 2.5 Sets of Vectors

**Performance Criterion:**

2. (h) Give an algebraic description of a set of a set of vectors that has been described geometrically, and vice-versa.

   (i) Determine whether a set of vectors is closed under vector addition; determine whether a set of vectors is closed under scalar multiplication. If it is, prove that it is; if it is not, give a counterexample.

One of the most fundamental concepts of mathematics is that of **sets**, collections of objects called **elements**. You have probably encountered various sets of numbers, like the **whole numbers**

$$\{0, 1, 2, 3, ...\}$$

and the integers

$$\mathbb{Z} = \{..., -3, -2, -1, 0, 1, 2, 3, ...\}.$$

As shown above, when we describe a set by listing all or some of its elements, we enclose them with "curly brackets." Sets are usually named by upper case letters. The above two sets are **infinite sets**.

Another kind of infinite set is an interval of the number line, like all real numbers between 1 and 5, including 1 and 5. You have likely seen the interval notation $[1, 5]$ for such sets. This set is also infinite, but in a different sense than the whole numbers and integers. That difference is not of concern to us here, but some of you may encounter that idea again. Of course, there are also finite sets like $\{2, 4, 6, 8\}$.

There will be a time soon when we will be very interested in sets of vectors, both finite and infinite. For example we might be interested in the finite set

$$A = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\},$$

or the infinite set of all vectors of the form $\begin{bmatrix} a \\ 2a \end{bmatrix}$, where $a$ is any real number. Let's examine this set a bit more.

◇ **Example 2.5(a):** Let $B$ be the set of all vectors of the form $\begin{bmatrix} a \\ 2a \end{bmatrix}$, where $a$ is any real number. Are the vectors $\vec{\mathbf{u}} = \begin{bmatrix} 3 \\ 10 \end{bmatrix}$ and $\vec{\mathbf{v}} = \begin{bmatrix} -2 \\ -4 \end{bmatrix}$ in $B$?

**Solution:** Because $2(3) \neq 10$, $\vec{\mathbf{u}}$ is not in $B$. But $2(-2) = -4$, so $\vec{\mathbf{v}}$ is in $B$.

---

◇ **Example 2.5(b):** Let $C$ be the set of all vectors of the form $\begin{bmatrix} a \\ a+1 \end{bmatrix}$, where $a$ is any real number. Are the vectors $\vec{\mathbf{u}} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\vec{\mathbf{v}} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$ in $C$?

**Solution:** $3 + 1 = 4$ and $-2 + 1 = -1$, so both $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are in $C$.

---

In the future we will also be very interested in this question: Given an infinite set of vectors, is the sum of any two vectors a vector that is also in the set? When faced with such a question we should do one of two things:

- Give two *specific* vectors that are in the set, and show that their sum is not.
- Give two *arbitray* (general) vectors in the set and show that their sum is also in the set.

The following examples illustrate these.

◇ **Example 2.5(c):** Let $B$ be the set of all vectors of the form $\begin{bmatrix} a \\ 2a \end{bmatrix}$, where $a$ is any real number. Determine whether the sum of any two vectors in $B$ is also in $B$.

**Solution:** The vectors $\begin{bmatrix} 3 \\ 6 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ -4 \end{bmatrix}$ are both in $B$, and their sum $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is as well. We may have just gotten lucky, though, and maybe the sum of *any* two vectors in $B$ is not necessarily in $B$. Let's see if the sum of two *arbitrary* vectors in $B$ is in $B$. For any numbers $a$ and $b$, the vectors $\begin{bmatrix} a \\ 2a \end{bmatrix}$ and $\begin{bmatrix} b \\ 2b \end{bmatrix}$ are in $B$. We then compute their sum to get

$$\begin{bmatrix} a \\ 2a \end{bmatrix} + \begin{bmatrix} b \\ 2b \end{bmatrix} = \begin{bmatrix} a+b \\ 2a+2b \end{bmatrix} = \begin{bmatrix} a+b \\ 2(a+b) \end{bmatrix} = \begin{bmatrix} c \\ 2c \end{bmatrix},$$

where $c = a + b$. Therefore the sum of *any* two vectors in $B$ is a vector in $B$.

---

◇ **Example 2.5(d):** Let $C$ be the set of all vectors of the form $\begin{bmatrix} a \\ a+1 \end{bmatrix}$, where $a$ is any real number. Is the sum of any two vectors in $C$ also a vector in $C$?

**Solution:** The vectors $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ -1 \end{bmatrix}$ are both in $C$, as shown in Example 2.5(b). Their sum is the vector $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, which is not in $C$ because $1 + 1 \neq 3$. Thus the sum of any two vectors in $C$ is not necessarily another vector in $C$.

---

◇ **Example 2.5(e):** Let $D$ be the set of all vectors of the form $\begin{bmatrix} x \\ y \end{bmatrix}$, where $x \geq 0$ and $y \geq 0$. Is the sum of any two vectors in $D$ also a vector in $D$?

**Solution:** Suppose that $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ and $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ are both in $D$, so all of $x_1, y_1, x_2, y_2$ are greater than or equal to zero. Clearly $x_1 + x_2 \geq 0$ and $y_1 + y_2 \geq 0$, so their sum $\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}$ is in $D$.

---

Given a set of vectors, we are also interested in whether a scalar multiple of a vector in the set is in the set as well. In the next example we determine whether that is the case for the set $D$ from the previous example.

◇ **Example 2.5(f):** Let $D$ be the set of all vectors of the form $\begin{bmatrix} x \\ y \end{bmatrix}$, where $x \geq 0$ and $y \geq 0$. Is *any* scalar times any vector in $D$ also a vector in $D$?

**Solution:** Suppose that $\begin{bmatrix} x \\ y \end{bmatrix}$ is in $D$, so both $x \geq 0$ and $y \geq 0$. Is it possible that for some scalar $a$, the vector $a\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax \\ ay \end{bmatrix}$ is *NOT* in $D$? That would only be the case if either $ax < 0$ or $ay < 0$, which would happen if $a$ was negative and at least one of $x$ or $y$ was positive. To give a specific example, for the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $a = -3$, $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is in $D$ but $(-3)\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -3 \\ -6 \end{bmatrix}$ is not.

---

It is very important that we note the difference between the approaches of Examples 2.5(e) and (f). When trying to show that something is true in general about a set, we must demonstrate it for arbitrary elements in the set, as done in Example 2.5(e). When trying to show something is not true all that is needed is one specific example that fails to be true, which is called a **counterexample**. The vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and scalar $-3$ are a specific counterexample for Example 2.5(f). We *ALWAYS* use *specific* counterexamples to show that something is not true, but a general example to show that something *IS* true.

**Section 2.5 Exercises** To Solutions

Do each of the following for each of Exercises 1 - 10.

(a) Give several vectors in the set.

(b) Determine whether the set is closed under addition.

(c) Determine whether the set is closed under scalar multiplication.

1. $\mathcal{S} = \left\{ \begin{bmatrix} a \\ a^2 \end{bmatrix} \in \mathbb{R}^2 \right\}$

2. $\mathcal{S} = \left\{ \begin{bmatrix} a \\ b \\ a+b \end{bmatrix} \;\middle|\; a, b \in \mathbb{R} \right\}$

3. $\mathcal{S} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \mid xy \geq 0 \right\}$

4. $\mathcal{S} = \left\{ t\begin{bmatrix} 1 \\ -2 \end{bmatrix} \;\middle|\; t \in \mathbb{R} \right\}$

5. $\mathcal{S} = \left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ -2 \end{bmatrix} \;\middle|\; t \in \mathbb{R} \right\}$

6. $\mathcal{S} = \left\{ s\begin{bmatrix} 3 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ -2 \end{bmatrix} \;\middle|\; s, t \in \mathbb{R} \right\}$

7. $\mathcal{S} = \left\{ \begin{bmatrix} -4 \\ 2 \end{bmatrix} + t\begin{bmatrix} 2 \\ -1 \end{bmatrix} \;\middle|\; t \in \mathbb{R} \right\}$

8. $\mathcal{S} = \left\{ \begin{bmatrix} a \\ b \\ |a| \end{bmatrix} \in \mathbb{R}^3 \right\}$

9. $\mathcal{S} = \left\{ \begin{bmatrix} a \\ 2a \\ 3a \end{bmatrix} \;\middle|\; a \in \mathbb{R} \right\}$

10. $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \;\middle|\; s, t \in \mathbb{R} \right\}$

11. On separate graphs, plot the points corresponding to each of the sets of vectors (taken as position vectors) in Exercises 1 - 4 above.

12. (a) The sets in Exercises 5 and 7 look very similar, but one is closed under both addition and scalar multiplication, and the other is not. What do you notice about the vectors used to describe the set that *IS* closed under both addition and scalar multiplication?

    (b) Graph the sets from Exercises 5 and 7 on separate graphs. How are the two graphs alike? How are they different?

13. Determine what the set described in Exercise 10 is, geometrically.

14. You might find it surprising that the vectors $\begin{bmatrix} 17 \\ -22 \end{bmatrix}$ and $\begin{bmatrix} -10 \\ -10 \end{bmatrix}$ are in the set $\mathcal{S}$ described in Exercise 6 above. For each of those two vectors, determine the values of $s$ and $t$ that give them, to the nearest hundredth.

15. Determine what the set described in Exercise 6 is, geometrically.

## 2.6 Vector Equations of Lines and Planes

**Performance Criterion:**

2. (j) Give the vector equation of a line through two points in $\mathbb{R}^2$ or $\mathbb{R}^3$ or the vector equation of a plane through three points in $\mathbb{R}^3$.

The idea of a linear combination does more for us than just give another way to interpret a system of equations. The set of points in $\mathbb{R}^2$ satisfying an equation of the form $y = mx + b$ is a line; any such equation can be rearranged into the form $ax + by = c$. (The values of $b$ in the two equations are clearly not the same.) But if we add one more term to get $ax + by + cz = d$, with the $(x, y, z)$ representing the coordinates of a point in $\mathbb{R}^3$, *we get the equation of a plane, not a line!* In fact, we cannot represent a line in $\mathbb{R}^3$ with a single *scalar* equation. The object of this section is to show how we can represent lines, planes and higher dimensional objects called **hyperplanes** using linear combinations of vectors.

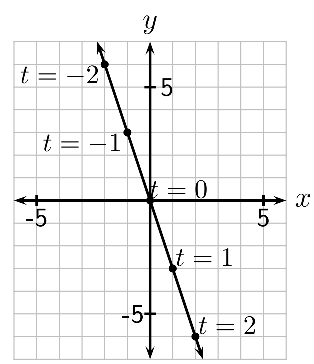
For the bulk of this course, we will think of most vectors as position vectors. (Remember, this means their tails are at the origin.) We will also think of each position vector as corresponding to the point at its tip, so the coordinates of the point will be the same as the components of the vector. Thus, for example, in $\mathbb{R}^2$ the vector $\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ corresponds to the ordered pair $(x_1, x_2) = (1, -3)$.

◇ **Example 2.6(a):** Graph the set of points corresponding to all vectors $\vec{\mathbf{x}}$ of the form $\vec{\mathbf{x}} = t\begin{bmatrix} 1 \\ -3 \end{bmatrix}$, where $t$ represents any real number.

**Solution:** We already know that when $t = 1$ the the vector $x$ corresponds to the point $(1, -3)$. We then let $t = -2, -1, 0, 2$ and determine the corresponding vectors $\vec{\mathbf{x}}$:

$$t = -2 \Rightarrow x = \begin{bmatrix} -2 \\ 6 \end{bmatrix}, \qquad t = -1 \Rightarrow x = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

$$t = 0 \Rightarrow x = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad t = 2 \Rightarrow x = \begin{bmatrix} 2 \\ -6 \end{bmatrix}$$




These vectors correspond to the points with ordered pairs $(-2, 6)$, $(-1, 3)$, $(0, 0)$ and $(2, -6)$, which lie on a line through the origin. If we were to continue plotting more such points for all possible values of $t$ we get the line shown above and to the right.

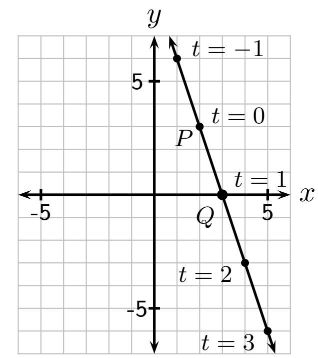
It should be clear from the above example that we could create a line through the origin in any direction by simply replacing the vector $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ with a vector in the direction of the desired line. The next example illustrates how we get a line not through the origin using vectors.

◇ **Example 2.6(b):** Graph the set of points corresponding to all vectors $\vec{\mathbf{x}}$ of the form $\vec{\mathbf{x}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, where $t$ represents any real number.

**Solution:** Performing the scalar multiplication by $t$ and adding the two vectors, we get

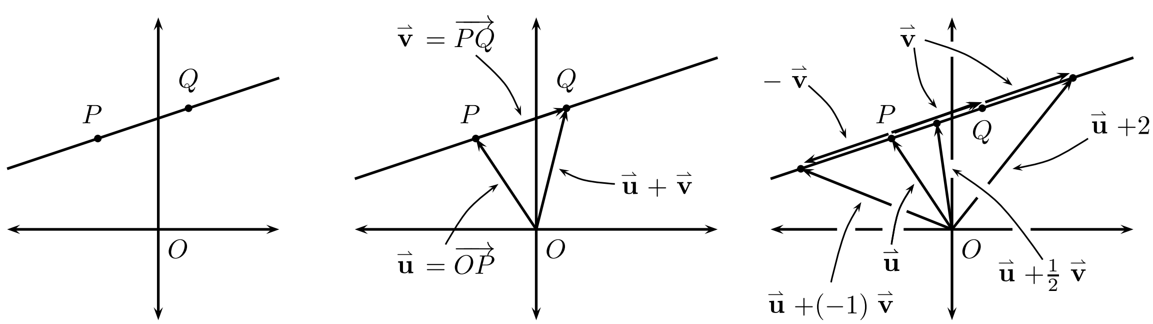
$$\vec{\mathbf{x}} = \begin{bmatrix} 2-3t \\ 3+\ t \end{bmatrix}.$$

These vectors then correspond to all points of the form $(2-3t, 3+t)$. When $t=0$ this is the point $(2,3)$ so our line clearly passes through that point. Plotting the points obtained when we let $t=-1$, $1$, $2$ and $3$, we see that we will get the line shown to the right.

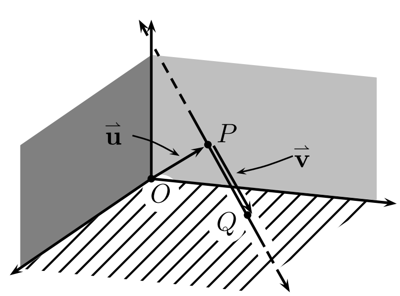



---

Now let's make two observations about the set of points represented by the set of all vectors $\vec{\mathbf{x}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, where $t$ again represents any real number. These vectors correspond to the ordered pairs of the form $(2-3t, 3+t)$. Plotting these results in the line through the point $(2,3)$ and in the direction of the vector $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$. This is not a coincidence. Consider the line shown below and to the left, containing the points $P$ and $Q$. If we let $\vec{\mathbf{u}} = \overrightarrow{OP}$ and $\vec{\mathbf{v}} = \overrightarrow{PQ}$, then the points $P$ and $Q$ correspond to the vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{u}} + \vec{\mathbf{v}}$ (in standard position, which you should assume we mean from here on), as shown in the second picture. From this you should be able to see that if we consider all the vectors $\vec{\mathbf{x}}$ defined by $\vec{\mathbf{x}} = \vec{\mathbf{u}} + t\,\vec{\mathbf{v}}$ as $t$ ranges over all real numbers, the resulting set of points is our line! This is shown in the third picture, where $t$ is given the values $-1$, $0$, $\frac{1}{2}$ and $2$.




This may seem like an overly complicated way to describe a line, but with a little thought you should see that the idea translates directly to three (and more!) dimensions, as shown in the picture to the right. This is all summarized by the following.



**Lines in $\mathbb{R}^2$ and $\mathbb{R}^3$**

The **vector equation of a line** through two points $P$ and $Q$ in $\mathbb{R}^2$ and $\mathbb{R}^3$ (and even higher dimensions) is

$$\vec{\mathbf{x}} = \overrightarrow{OP} + t\,\overrightarrow{PQ}.$$

By this we mean that the line consists of all the points corresponding to the position vectors $\vec{\mathbf{x}}$ as $t$ varies over all real numbers. The vector $\overrightarrow{PQ}$ is called the **direction vector** of the line.

◇ **Example 2.6(c):** Give the vector equation of the line in $\mathbb{R}^2$ through the points $P(-4,1)$ and $Q(5,3)$.

**Solution:** We need two vectors, one from the origin out to the line, and one in the direction of the line. For the first we will use $\overrightarrow{OP}$, and for the second we will use $\overrightarrow{PQ} = [9,2]$. We then have

$$\vec{\mathbf{x}} = \overrightarrow{OP} + t\overrightarrow{PQ} = \begin{bmatrix} -4 \\ 1 \end{bmatrix} + t\begin{bmatrix} 9 \\ 2 \end{bmatrix},$$

where $\vec{\mathbf{x}} = [x_1, x_2]$ is the position vector corresponding to any point $(x_1, x_2)$ on the line.

◇ **Example 2.6(d):** Give a vector equation of the line in $\mathbb{R}^3$ through the points $(-5,1,2)$ and $(4,6,-3)$.

**Solution:** Letting $P$ be the point $(-5,1,2)$ and $Q$ be the point $(4,6,-3)$, we get $\overrightarrow{PQ} = \langle 9,5,-5\rangle$. The vector equation of the line is then

$$\vec{\mathbf{x}} = \overrightarrow{OP} + t\overrightarrow{PQ} = \begin{bmatrix} -5 \\ 1 \\ 2 \end{bmatrix} + t\begin{bmatrix} 9 \\ 5 \\ -5 \end{bmatrix},$$

where $\vec{\mathbf{x}} = [x_1, x_2, x_3]$ is the position vector corresponding to any point $(x_1, x_2, x_3)$ on the line.

*The vector equation of a line is not unique!* The first vector can be any point on the line, so it could be the vector $\overrightarrow{OQ} = [4,6,-3]$ instead of $[-5,1,2]$, for example. The second vector is simply a direction vector, so can be any scalar multiple of $\overrightarrow{PQ} = [9,5,-5]$, including $\overrightarrow{QP} = [-9,-5,5]$.

The same general idea can be used to describe a plane in $\mathbb{R}^3$. Before seeing how that works, let's define something and look at a situation in $\mathbb{R}^2$. We say that two vectors are **parallel** if one is a scalar multiple of the other. Now suppose that $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ are two nonzero vectors in $\mathbb{R}^2$ that are *not* parallel, as shown in Figure 2.6(a) on the next page, and let $P$ be the randomly chosen point in $\mathbb{R}^2$ shown in the same picture. Figure 2.6(b) shows that a linear combination of $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ can be formed that gives us a vector $s\vec{\mathbf{v}} + t\vec{\mathbf{w}}$ corresponding to the point $P$. In this case the scalar $s$ is positive and less than one, and $t$ is positive and greater than one. Figures 2.6(c) and 2.6(d) show the same thing for another point $Q$, with $s$ being negative and $t$ positive in that case. It should now be clear that *any* point in $\mathbb{R}^2$ can be obtained in this manner.

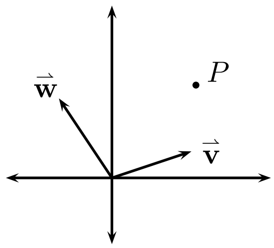



Figure 2.6(a)

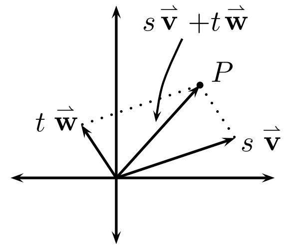



Figure 2.6(b)

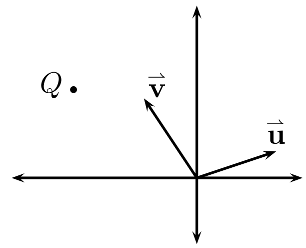



Figure 2.6(c)

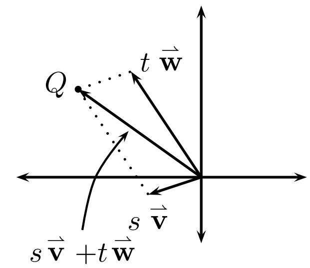
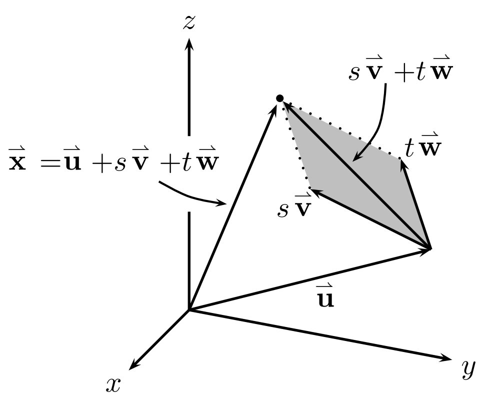
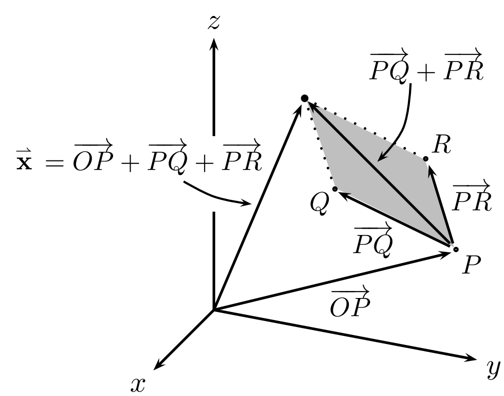



Figure 2.6(d)

Now let $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ be three vectors in $\mathbb{R}^3$, and consider the vector $\vec{\mathbf{x}} = \vec{\mathbf{u}} + s\vec{\mathbf{v}} + t\vec{\mathbf{w}}$, where $s$ and $t$ are scalars that are allowed to take all real numbers as values. The vectors $s\vec{\mathbf{v}} + t\vec{\mathbf{w}}$ all lie in the plane containing $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$. Adding $\vec{\mathbf{u}}$ "moves the plane off the origin" to where it passes through the tip of $\vec{\mathbf{u}}$ (again, in standard position). This is probably best visualized by thinking of adding $s\,\vec{\mathbf{v}}$ and $t\,\vec{\mathbf{w}}$ with the parallelogram method, then adding the result to $\vec{\mathbf{u}}$ with the tip-to-tail method. I have attempted to illustrate this to the left at the top of the next page, with the gray parallelogram being part of the plane created by all the points corresponding to the vectors $\vec{\mathbf{x}}$.




The same diagram above and to the right shows how all of the previous discussion relates to the plane through three points $P$, $Q$ and $R$ in $\mathbb{R}^3$. This leads us to the description of a plane in $\mathbb{R}^3$ given at the top of the next page.

**Planes in $\mathbb{R}^3$**

The **vector equation of a plane** through three points $P$, $Q$ and $R$ in $\mathbb{R}^3$ (or higher dimensions) is

$$\vec{\mathbf{x}} = \overrightarrow{OP} + s\,\overrightarrow{PQ} + t\,\overrightarrow{PR}.$$

By this we mean that the plane consists of all the points corresponding to the position vectors $\vec{\mathbf{x}}$ as $s$ and $t$ vary over all real numbers.

◇ **Example 2.6(e):** Give a vector equation of the plane in $\mathbb{R}^3$ through the points $(2,-1,3)$, $(-5,1,2)$ and $(4,6,-3)$. What values of $s$ and $t$ give the point $R$?

**Solution:** Letting $P$ be the point $(2,-1,3)$, $Q$ be $(-5,1,2)$ and $R$ be $(4,6,-3)$, we get

$\overrightarrow{PQ} = [-7, 2, -1]$ and $\overrightarrow{PR} = [2, 7, -6]$. The vector equation of the plane is then

$$\vec{\mathbf{x}} = \overrightarrow{OP} + s\overrightarrow{PQ} + t\overrightarrow{PR} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} + s \begin{bmatrix} -7 \\ 2 \\ -1 \end{bmatrix} + t \begin{bmatrix} 2 \\ 7 \\ -6 \end{bmatrix},$$

where $\vec{\mathbf{x}} = [x_1, x_2, x_3]$ is the position vector corresponding to any point $(x_1, x_2, x_3)$ on the plane. It should be clear that there are other possibilities for this. The first vector in the equation could be any of the three position vectors for $P$, $Q$ or $R$. The other two vectors could be any two vectors from one of the points to another.

The vector corresponding to point $R$ is $\overrightarrow{OR}$, which is equal to $\vec{\mathbf{x}} = \overrightarrow{OP} + \overrightarrow{PR}$ (think about that), so $s = 0$ and $t = 1$.

---

We now summarize all of the ideas from this section.

**Lines in $\mathbb{R}^2$ and $\mathbb{R}^3$, Planes in $\mathbb{R}^3$**

Let $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ be vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ with $\vec{\mathbf{v}} \neq \mathbf{0}$. Then the set of points corresponding to the vector $\vec{\mathbf{x}} = \vec{\mathbf{u}} + t\,\vec{\mathbf{v}}$ as $t$ ranges over all real numbers is a line through the point corresponding to $\vec{\mathbf{u}}$ and in the direction of $\vec{\mathbf{v}}$. (So if $\vec{\mathbf{u}} = \vec{\mathbf{0}}$ the line passes through the origin.)

Let $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ be vectors $\mathbb{R}^3$, with $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ being nonzero and not parallel. (That is, not scalar multiples of each other.) Then the set of points corresponding to the vector $\vec{\mathbf{x}} = \vec{\mathbf{u}} + s\,\vec{\mathbf{v}} + t\,\vec{\mathbf{w}}$ as $s$ and $t$ range over all real numbers is a plane through the point corresponding to $\vec{\mathbf{u}}$ and containing the vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$. (If $\vec{\mathbf{u}} = \vec{\mathbf{0}}$ the plane passes through the origin.)

## Section 2.6 Exercises

To Solutions

1. For each of the following, give the vector equation of the line or plane described.

   (a) The line in $\mathbb{R}^2$ through the points $P(3, -1)$ and $Q(6, 0)$.

   (b) The plane in $\mathbb{R}^3$ through the points $P(3, -1, 4)$, $Q(2, 6, 0)$ and $R(-1, 0, 3)$.

   (c) The line in $\mathbb{R}^3$ through the two points $P(3, -1, 4)$ and $Q(2, 6, 0)$.

   (d) The line in $\mathbb{R}^2$ through the points $(-1, 4)$ and $(2, 5)$.

   (e) The line in $\mathbb{R}^2$ through $(-4, 3)$ and the origin. (**Hint:** Use the method of Example 2.6(c), taking $P$ to be $(0, 0)$ and $Q$ to be $(-4, 3)$.)

   (f) The plane in $\mathbb{R}^3$ through $(-5, 1, 3)$, $(2, 0, 4)$ and $(1, -2, 3)$.

   (g) The line in $\mathbb{R}^3$ through $(2, 0, -1)$ and $(1, 1, 3)$.

   (h) The plane in $\mathbb{R}^3$ through $(-4, 5, 1)$, $(2, 2, 2)$ and the origin.

   (i) The line in $\mathbb{R}^3$ through $(1, 2, 3)$ and the origin.

2. Each of the lines or planes in the previous exercise constitutes a set of points in $\mathbb{R}^2$ or $\mathbb{R}^3$. Which are sets closed under addition and scalar multiplication?

3. Give the equation of the three-dimensional hyperplane in $\mathbb{R}^4$ containing the points $(-2,1,1,5)$, $(3,7,-5,2)$, $(4,0,5,-2)$ and $(3,-5,4,7)$.

4. Consider the two points $P(5,-1,4)$ and $Q(1,1,2)$ in $\mathbb{R}^3$.

   (a) Give the vector equation of the line for which the value $t=0$ for the parameter gives the point $Q$ and $t=1$ gives the point $P$.

   (b) Give the vector equation of the line for which the value $t=0$ for the parameter gives the point $Q$ and $t=-1$ gives the point $P$.

   (c) Give the vector equation of the line for which the value $t=0$ for the parameter gives the point $P$ and $t=2$ gives the point $Q$.

5. Consider the two points $P(-3,1)$ and $Q(2,5)$ in $\mathbb{R}^2$.

   (a) Give the "next" point with integer coordinates as one goes from $P$ to $Q$.

   (b) Give the vector equation for the line for which the parameter value $t=0$ gives the point $P$ and $t=1$ gives $Q$.

   (c) What value of $t$ in your equation from (b) gives the point you found in part (a)?

   (d) Give the "next" point with integer coordinates as one goes from $Q$ to $P$. What value of the parameter $t$ in your equation from (b) gives this point?

6. Consider the three points $P(1,6,-2)$, $Q(-5,7,4)$ and $R(3,0,-1)$ in $\mathbb{R}^3$. Give the vector equation of plane in the form $\vec{\mathbf{x}} = \vec{\mathbf{u}} + s\vec{\mathbf{v}} + t\vec{\mathbf{w}}$ for which $s=0$ and $t=0$ gives the point $Q$, $s=0$ and $t=1$ gives $R$, and $s=1, t=0$ gives $P$.

7. Give the "next" three points on the line in $\mathbb{R}^3$ containing $P(5,-1,4)$ and $Q(1,1,2)$, traveling in the direction from $P$ to $Q$.

8. Find another point in the plane containing $P_1(-2,1,5)$, $P_2(3,2,1)$ and $P_3(4,-2,-3)$. **Show clearly how you do it**. (**Hint:** Find and use the vector equation of the plane.)

9. "Usually" a vector equation of the form $\vec{\mathbf{x}} = \vec{\mathbf{p}} + s\vec{\mathbf{u}} + t\vec{\mathbf{v}}$ gives the equation of a plane in $\mathbf{R}^3$. Answer the following first allowing any of $\vec{\mathbf{p}}$, $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ to be the zero vector, then give answers assuming that none of them are zero.

   (a) Under what conditions on $\vec{\mathbf{p}}$ and/or $\vec{\mathbf{u}}$ and/or $\vec{\mathbf{v}}$ would this be the equation of a line?

   (b) Under what conditions on $\vec{\mathbf{p}}$ and/or $\vec{\mathbf{u}}$ and/or $\vec{\mathbf{v}}$ would this be the equation of a plane through the origin?

## 2.7 Interpreting Solutions to Systems of Linear Equations

**Performance Criterion:**

2. (k) Write the solution to a system of equations in vector form and determine the geometric nature of the solution.

We begin this section by considering the following two systems of equations.

$$\begin{aligned} 3x_1 - 3x_2 + 3x_3 &= 9 \\ 2x_1 - x_2 + 4x_3 &= 7 \\ 3x_1 - 5x_2 - x_3 &= -3 \end{aligned} \qquad\qquad \begin{aligned} x_1 - x_2 + 2x_3 &= 1 \\ -3x_1 + 3x_2 - 6x_3 &= -3 \\ 2x_1 - 2x_2 + 4x_3 &= 2 \end{aligned}$$

The augmented matrices for these two systems reduce to the following matrices, respectively.

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

Let's look at the first system. $x_3$ is a free variable, and $x_1$ and $x_2$ are leading variables. The general solution is $x_1 = -3t + 4$, $x_2 = -2t + 1$, $x_3 = t$. Algebraically, $x_1$, $x_2$ and $x_3$ are just numbers, but we can think of $(x_1, x_2, x_3)$ as a point in $\mathbb{R}^3$. The corresponding position vector is

$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4-3t \\ 1-2t \\ t \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -3t \\ -2t \\ t \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -3 \\ -2 \\ 1 \end{bmatrix}$$

We will call this the **vector form of the solution** to the system of equations. The beauty of expressing the solutions to a system of equations in vector form is that we can see what the set of all solutions looks like. In this case, the set of solutions is the set of all points in $\mathbb{R}^3$ on the line through $(4, 1, 0)$ and with direction vector $[-3, -2, 1]$.

◇ **Example 2.7(a):** The general solution to the second system of equations is $x_1 = 1 + s - 2t$, $x_2 = s$, $x_3 = t$. Express the solution in vector form and determine the geometric nature of the solution set in $\mathbb{R}^3$.

**Solution:** A process like the one just carried out leads to the general solution with position vector

$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

(Check to make sure that you understand how this was arrived at.) Here the set of solutions is the set of all points in $\mathbb{R}^3$ on the plane through $(1, 0, 0)$ with direction vectors $[1, 1, 0]$ and $[-2, 0, 1]$.

Now recall that the three equations from this last example,

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

represent three planes in $\mathbb{R}^3$, and when we solve the system we are looking for all points in $\mathbb{R}^3$ that are solutions to all three equations. Our results tell us that the set of solution points in this case is itself a plane, which can only happen if all three equations represent the same plane. If you look at them carefully you can see that the second and third equations are multiples of the first, so the points satisfying them also satisfy the first equation.

◇ **Example 2.7(b):** The general solution to the second system of equations is $x_1 = 1 + s - 2t$, $x_2 = s$, $x_3 = t$. Express the solution in vector form and determine the geometric nature of the solution set in $\mathbb{R}^3$.

**Solution:** A process like the one just carried out leads to the general solution with position vector

$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

(Check to make sure that you understand how this was arrived at.) Here the set of solutions is the set of all points in $\mathbb{R}^3$ on the plane through $(1, 0, 0)$ with direction vectors $[1, 1, 0]$ and $[-2, 0, 1]$.

---

◇ **Example 2.7(c):** Give the vector form of the solution to the system

$$\begin{aligned} 3x_2 - \phantom{1}6x_3 - 4x_4 - 3x_5 &= -5 \\ x_1 - 3x_2 + 10x_3 + 4x_4 + 4x_5 &= \phantom{-}2 \\ 2x_1 - 6x_2 + 20x_3 + 2x_4 + 8x_5 &= -8 \end{aligned}$$

**Solution:** The augmented matrix of the system reduces to

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

We can see that $x_3$ and $x_5$ are free variables, and we can also see that $x_4 = 2$. Letting $x_5 = t$ and $x_3 = s$, $x_2 = 1 + 2s + t$ and $x_1 = -3 - 4s - t$. Therefore

$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} -4 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

---

How do we interpret this result geometrically? The set of points $(x_1, x_2, x_3, x_4, x_5)$ represents a two-dimensional plane in five-dimensional space. We could also have ended up with one, three or four dimensional "plane", often called a **hyperplane**, in five-dimensional space.

**Section 2.7 Exercises** To Solutions

1. For each of the following, a student correctly finds the given the general solution $(x_1, x_2, x_3)$ to a system of three equations in three unknowns. Give the vector form of the solution, then tell whether the set of all particular solutions is a point, line or plane.

   (a) $x_1 = s - t + 5, \quad x_2 = s, \quad x_3 = t$ (b) $x_1 = 2t + 5, \quad x_2 = t, \quad x_3 = -1$

   (c) $x_1 = s - 2t + 5, \quad x_2 = s, \quad x_3 = t$

2. In each of the following, the vector form of the solution to a system of linear equations is given. Give the dimension of the solution, and the dimension of the space it is in. For example you might answer "three-dimensional plane in five-dimensional space."

   (a)
$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -1 \\ 4 \end{bmatrix} + r\begin{bmatrix} 3 \\ 7 \\ 1 \\ -4 \end{bmatrix} + s\begin{bmatrix} 6 \\ 3 \\ -10 \\ 8 \end{bmatrix} + t\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

   (b)
$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 2 \\ 0 \end{bmatrix} + s\begin{bmatrix} -4 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

   (c)
$$\vec{\mathbf{x}} = \begin{bmatrix} 1 \\ 0 \\ -3 \\ 1 \\ 0 \\ 2 \\ 0 \end{bmatrix} + t_1\begin{bmatrix} 5 \\ -3 \\ -4 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t_2\begin{bmatrix} 1 \\ 5 \\ -1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + t_3\begin{bmatrix} 2 \\ -5 \\ 1 \\ 0 \\ 3 \\ 9 \\ 3 \end{bmatrix} + t_4\begin{bmatrix} 5 \\ 7 \\ -4 \\ -2 \\ 1 \\ -4 \\ 5 \end{bmatrix} + t_5\begin{bmatrix} 3 \\ 1 \\ -1 \\ 6 \\ 10 \\ -4 \\ 1 \end{bmatrix}$$

   (d)
$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 4 \\ -7 \\ 3 \\ 6 \\ -1 \end{bmatrix} + t\begin{bmatrix} 2 \\ 0 \\ -8 \\ -1 \\ 4 \end{bmatrix}$$

(e)

$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 3 \\ 7 \\ 1 \\ -4 \end{bmatrix} + s \begin{bmatrix} 6 \\ 3 \\ -10 \\ 8 \end{bmatrix}$$

(f)

$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = t_1 \begin{bmatrix} 1 \\ 5 \\ -1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + t_2 \begin{bmatrix} 2 \\ -5 \\ 1 \\ 0 \\ 3 \\ 9 \\ 3 \end{bmatrix} + t_3 \begin{bmatrix} 5 \\ 7 \\ -4 \\ -2 \\ 1 \\ -4 \\ 5 \end{bmatrix} + t_4 \begin{bmatrix} 3 \\ 1 \\ -1 \\ 6 \\ 10 \\ -4 \\ 1 \end{bmatrix}$$

## 2.8 The Dot Product of Vectors, Projections

**Performance Criteria:**

3. (g) Find the dot product of two vectors, determine the length of a single vector.
   (h) Determine whether two vectors are orthogonal (perpendicular).
   (i) Find the projection of one vector onto another, graphically or algebraically.

### The Dot Product and Orthogonality

There are two ways to "multiply" vectors, both of which you have likely seen before. One is called the **cross product**, and only applies to vectors in $\mathbb{R}^3$. It is quite useful and meaningful in certain physical situations, but it will be of no use to us here. More useful is the other method, called the **dot product**, which is valid in all dimensions.

DEFINITION 2.8.1: **Dot Product**

Let $\vec{\mathbf{u}}= [u_1, u_2, ..., u_n]$ and $\vec{\mathbf{v}}= [v_1, v_2, ..., v_n]$. The **dot product** of $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$, denoted by $\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}$, is given by

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}= u_1v_1 + u_2v_2 + u_3v_3 + \cdots + u_nv_n$$

The dot product is useful for a variety of things. Recall that the length of a vector $\vec{\mathbf{v}}= [v_1, v_2, ..., v_n]$ is given by $\| \vec{\mathbf{v}} \| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} = \sqrt{\vec{\mathbf{v}} \cdot \vec{\mathbf{v}}}$. Note also that $v_1^2 + v_2^2 + \cdots + v_n^2 = \vec{\mathbf{v}} \cdot \vec{\mathbf{v}}$, which implies that $\| \vec{\mathbf{v}} \| = \sqrt{\vec{\mathbf{v}} \cdot \vec{\mathbf{v}}}$. Perhaps the most important thing about the dot product is that the dot product of two vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ is zero if, and only if, the two vectors are perpendicular. In general, we make the following definition.

DEFINITION 2.8.2: **Orthogonal Vectors**

Two vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ in $\mathbb{R}^n$ are said to be **orthogonal** if, and only if, $\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}= 0$.

◇ **Example 2.8(a):** For the three vectors $\mathbf{u} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix}$, find $\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}$, $\vec{\mathbf{u}} \cdot \vec{\mathbf{w}}$ and $\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$. Are any of the vectors orthogonal to each other?

**Solution:** We find that

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}= (5)(-1) + (-1)(3) + (2)(4) = -5 + (-3) + 8 = 0,$$

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{w}} = (5)(2) + (-1)(-1) + (2)(-3) = 10 + 1 + (-6) = 5,$$

$$\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = (-1)(2) + (3)(-1) + (4)(-3) = -2 + (-3) + (-12) = -17$$

From the first computation we can see that $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are orthogonal.

---

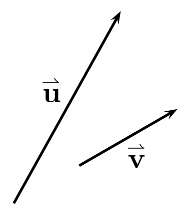
**Projections**

Given two vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$, we can create a new vector $\vec{\mathbf{w}}$ called the **projection of $\vec{\mathbf{u}}$ onto $\vec{\mathbf{v}}$**, denoted by $\text{proj}_{\mathbf{v}}\, \vec{\mathbf{u}}$. This is a very useful idea, in many ways. Geometrically, we can find $\text{proj}_{\mathbf{v}}\, \vec{\mathbf{u}}$ as follows:

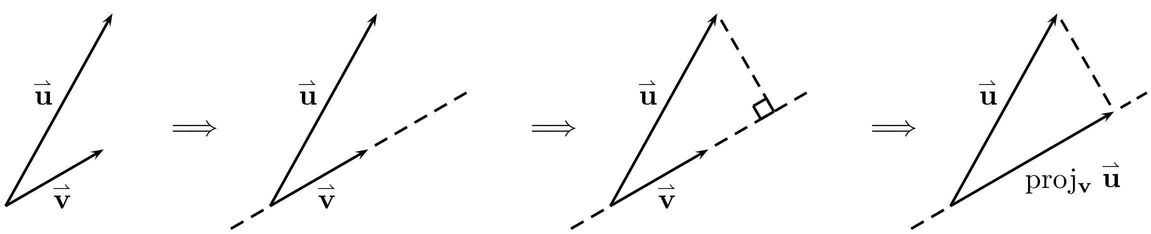
- Bring $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ together tail-to-tail.
- Sketch in the line containing $\vec{\mathbf{v}}$, as a dashed line.
- Sketch in a dashed line segment from the tip of $\vec{\mathbf{u}}$ to the dashed line containing $\vec{\mathbf{v}}$, *perpendicular to that line.*
- Draw the vector $\text{proj}_{\mathbf{v}}\, \vec{\mathbf{u}}$ from the point at the tails of $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ to the point where the dashed line segment meets $\vec{\mathbf{v}}$ or the dashed line containing $\vec{\mathbf{v}}$.

Note that $\text{proj}_{\mathbf{v}}\, \vec{\mathbf{u}}$ is parallel to $\vec{\mathbf{v}}$; if we were to find $\text{proj}_{\mathbf{u}}\, \vec{\mathbf{v}}$ instead, the result would be parallel to $\vec{\mathbf{u}}$ in that case. The above steps are illustrated in the following example.

◇ **Example 2.8(b):** For the vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ shown to the right, find the projection $\text{proj}_{\mathbf{v}}\, \vec{\mathbf{u}}$.



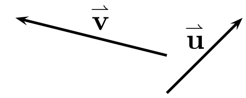

**Solution:** Following the above steps we get



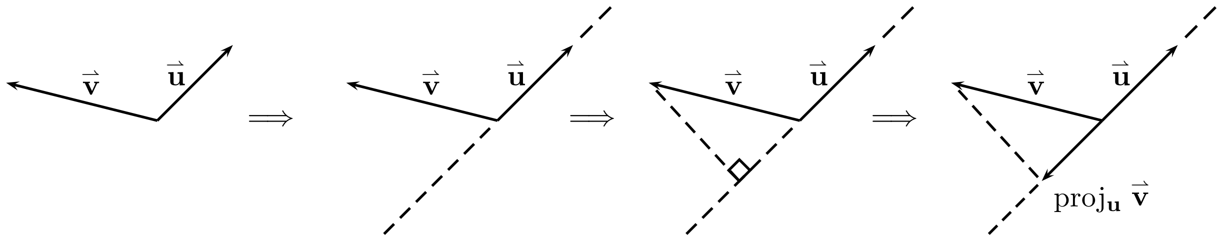

---

Projections are a bit less intuitive when the angle between the two vectors is obtuse, as seen in the next example.

⋄ **Example 2.8(c):** For the vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ shown to the right, find the projection $\text{proj}_{\mathbf{u}}\ \vec{\mathbf{v}}$.




**Solution:** We follow the steps again, noting that this time we are projecting $\vec{\mathbf{v}}$ onto $\vec{\mathbf{u}}$:




Here we see that $\text{proj}_{\text{u}}\ \vec{\mathbf{v}}$ is in the direction opposite $\vec{\mathbf{u}}$.

---

We will also want to know how to find projections algebraically:

> DEFINITION 2.8.3: **The Projection of One Vector on Another**
>
> For two vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$, the vector $\text{proj}_{\text{v}}\ \vec{\mathbf{u}}$ is given by
>
> $$\text{proj}_{\text{v}}\ \vec{\mathbf{u}} = \frac{\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}}{\vec{\mathbf{v}} \cdot \vec{\mathbf{v}}}\ \vec{\mathbf{v}}$$

Note that since both $\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}$ and $\vec{\mathbf{v}} \cdot \vec{\mathbf{v}}$ are scalars, so is $\dfrac{\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}}{\vec{\mathbf{v}} \cdot \vec{\mathbf{v}}}$. That scalar is then multiplied times $\vec{\mathbf{v}}$, resulting in a vector parallel $\vec{\mathbf{v}}$. if the scalar is positive the projection is in the direction of $\vec{\mathbf{v}}$, as shown in Example 2.8(b); when the scalar is negative the projection is in the direction opposite the vector being projected onto, as shown in Example 2.8(c).

⋄ **Example 2.8(d):** For the vectors $\mathbf{u} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix}$, find $\text{proj}_{\mathbf{u}}\ \vec{\mathbf{v}}$.

Note that here we are projecting $\vec{\mathbf{v}}$ onto $\vec{\mathbf{u}}$. first we find

$\vec{\mathbf{v}} \cdot \vec{\mathbf{u}} = (2)(5) + (-1)(-1) + (-3)(2) = 5$ and $\vec{\mathbf{u}} \cdot \vec{\mathbf{u}} = 5^2 + (-1)^2 + 2^2 = 30$

Then

$$\text{proj}_{\mathbf{u}}\ \vec{\mathbf{v}} = \frac{\vec{\mathbf{v}} \cdot \vec{\mathbf{u}}}{\vec{\mathbf{u}} \cdot \vec{\mathbf{u}}}\ \vec{\mathbf{u}} = \frac{5}{30}\begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{5}{6} \\ -\frac{1}{6} \\ \frac{1}{3} \end{bmatrix}$$

---

As stated before, the idea of projection is extremely important in mathematics, and arises in situations that do not appear to have anything to do with geometry and vectors as we are thinking of them now. You will see a clever geometric use of vectors in one of the exercises.

**Section 2.8 Exercises** To Solutions

1. Consider the vectors $\vec{\mathbf{v}} = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$ and $\vec{\mathbf{b}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

   (a) Draw a *neat and accurate* graph of $\vec{\mathbf{v}}$ and $\vec{\mathbf{b}}$, with their tails at the origin, labeling each.

   (b) Use the formula to find $\text{proj}_{\mathbf{b}}\,\vec{\mathbf{v}}$, with its components rounded to the nearest tenth.

   (c) Add $\text{proj}_{\mathbf{b}}\,\vec{\mathbf{v}}$ to your graph. Does it look correct?

2. For each pair of vectors **v** and **b** below, do each of the following

   i) Sketch **v** and **b** with the same initial point.

   ii) Find $\text{proj}_{\mathbf{b}}\vec{\mathbf{v}}$ algebraically, using the formula for projections.

   iii) On the same diagram, sketch the $\text{proj}_{\mathbf{b}}\vec{\mathbf{v}}$ you obtained in part (ii). If it does not look the way it should, find your error.

   iv) Find $\text{proj}_{\mathbf{b}}\vec{\mathbf{v}}$, and sketch it as a new sketch. Compare with your previous sketch.

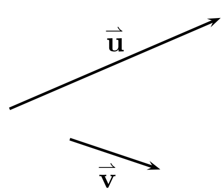
   (a) $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$

   (b) $\mathbf{v} = \begin{bmatrix} -5 \\ 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

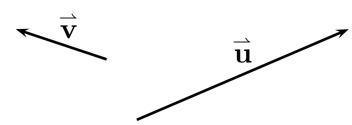
   (c) $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -2 \\ -4 \end{bmatrix}$

3. For each pair of vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$, sketch $\text{proj}_{\mathbf{v}}\,\vec{\mathbf{u}}$. Indicate any right angles with the standard symbol.

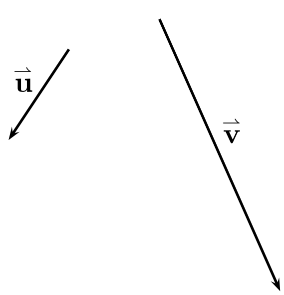
(a)



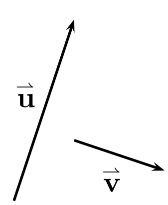

(b)




(c)




(d)




($\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are orthogonal)

# 3 Matrices and Vectors

**Outcome:**

3. Understand matrices, their algebra, and their action on vectors. Use matrices to solve problems. Understand the algebra of matrices.

   **Performance Criteria:**

   (a) Give the dimensions of a matrix. Identify a given entry, row or column of a matrix.

   (b) Identify matrices as square, upper triangular, lower triangular, symmetric, diagonal. Give the transpose of a given matrix; know the notation for the transpose of a matrix.

   (c) Add or subtract matrices when possible. Multiply a matrix by a scalar.

   (d) Multiply a matrix times a vector, give the linear combination form of a matrix times a vector.

   (e) Give a specific matrix that is multiplied times an arbitrary vector to obtain a given resulting vector.

   (f) Express a system of equations as a coefficient matrix times a vector equalling another vector.

   (g) For a given matrix $A$ and vector $\vec{\mathbf{b}}$, find a vector $\vec{\mathbf{x}}$ for which $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$,

   (h) Determine whether a matrix is a projection matrix, reflection matrix or rotation matrix, or none of these, by its action on a few vectors.

   (i) Determine whether a vector is an eigenvector of a matrix. If it is, give the corresponding eigenvalue.

   (j) Know when two matrices can be multiplied, and know that matrix multiplication is not necessarily commutative. Multiply two matrices "by hand."

   (k) Determine whether two matrices are inverses without finding the inverse of either.

   (l) Find the inverse of a 2×2 matrix using the formula. Find the inverse of a matrix using the Gauss-Jordan method. Describe the Gauss-Jordan method for finding the inverse of a matrix.

   (m) Solve a system of equations using an inverse matrix. Describe how to use an inverse matrix to solve a system of equations.

   (n) Find the determinant of a 2×2 or $3 \times 3$ matrix by hand. Use a calculator to find the determinant of an $n \times n$ matrix.

   (o) Use the determinant to determine whether a system of equations has a unique solution.

   (p) Determine whether a homogeneous system has more than one solution.

Continued on the next page.

**Outcome:**

3. Understand matrices, their algebra, and their action on vectors. Use matrices to solve problems. Understand the algebra of matrices.

   **Performance Criteria:**

   (q) Use Cramer's rule to solve a system of equations.

   (r) Give the geometric or algebraic representations of the inverse or square of a rotation. Demonstrate that the geometric and algebraic versions are the same

   (s) Give the incidence matrix of a graph or digraph. Given the incidence matrix of a graph or digraph, identify the vertices and edges using correct notation, and draw the graph.

   (t) Determine the number of $k$-paths from one vertex of a graph to another. Solve problems using incidence matrices.

## 3.1 Introduction to Matrices

**Performance Criteria:**

3. (a) Give the dimensions of a matrix. Identify a given entry, row or column of a matrix.
   (b) Identify matrices as square, upper triangular, lower triangular, symmetric, diagonal. Give the transpose of a given matrix; know the notation for the transpose of a matrix.
   (c) Add or subtract matrices when possible. Multiply a matrix by a scalar.

A **matrix** is simply an array of numbers arranged in rows and columns. Here are some examples:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} -5 & 1 \\ 0 & 4 \\ 2 & -3 \end{bmatrix}, \qquad D = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}, \qquad L = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 5 & -2 & 1 \end{bmatrix}$$

*We will always denote matrices with italicized capital letters.* There should be no need to define the **rows** and **columns** of a matrix. The number of rows and number of columns of a matrix are called its **dimensions**. The second matrix above, $B$, has dimensions $3 \times 2$, which we read as "three by two." The numbers in a matrix are called its **entries**. Each entry of a matrix is identified by its row, then column. For example, the $(3, 2)$ entry of $L$ is the entry in the 3rd row and second column, $-2$. In general, we will define the $(i, j)$th entry of a matrix to be the entry in the $i$th row and $j$th column.

There are a few special kinds of matrices that we will run into regularly:

- A matrix with the same number of rows and columns is called a **square matrix**. Matrices $A$, $D$ and $L$ above are square matrices.

- The entries that are in the same number row and column of a square matrix are called the **diagonal entries** of the matrix. For example, the diagonal entries of $A$ are 1 and 3. All the diagonal entries taken together are called the **diagonal** of the matrix. (This *ALWAYS* refers to only the diagonal from upper left to lower right.)

- A square matrix with zeros "above" the diagonal is called a **lower triangular matrix**; $L$ is an example of a lower triangular matrix. Similarly, an **upper triangular matrix** is one whose entries below the diagonal are all zeros. (Note that the words "lower" and "upper" refer to the triangular parts of the matrices where the entries are *NOT* zero.)

- A square matrix all of whose entries above *AND* below the diagonal are zero is called a **diagonal matrix**. $D$ is an example of a diagonal matrix. *Any diagonal matrix is also both upper and lower triangular.*

- A diagonal matrix with only ones on the diagonal is called "the" **identity matrix**. We use the word "the" because in a given size there is only one identity matrix. We will soon see why it is called the "identity."

- Given a matrix, we can create a new matrix whose rows are the columns of the original matrix. (This is equivalent to the columns of the new matrix being the rows of the original.) The new matrix is called the **transpose** of the original. The transposes of the matrices $B$ and $L$ above are denoted by $B^T$ and $L^T$. They are the matrices

$$B^T = \begin{bmatrix} -5 & 0 & 2 \\ 1 & 4 & -3 \end{bmatrix} \qquad L^T = \begin{bmatrix} 1 & -3 & 5 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

If a matrix $a$ is $m \times n$, then $A^T$ is $n \times m$. Note that when a matrix is square, its transpose is obtained by "flipping the matrix over its diagonal."

- Notice that $A^T = A$. Such a matrix is called a **symmetric matrix**. One way of thinking of such a matrix is that the entries across the diagonal from each other are equal. Matrix $D$ is also symmetric, as is the matrix
$$\begin{bmatrix} 1 & 5 & 0 & -2 \\ 5 & -4 & 7 & 3 \\ 0 & 7 & 0 & -6 \\ -2 & 3 & -6 & -3 \end{bmatrix}$$

When discussing an arbitrary matrix $A$ with dimensions $m \times n$ we refer to each entry as $a$, but with a double subscript with each to indicate its position in the matrix. The first number in the subscript indicates the row of the entry and the second indicates the column of that entry:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1k} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & & & \vdots & & \vdots \\ a_{j1} & a_{j2} & a_{j3} & \cdots & a_{jk} & \cdots & a_{jn} \\ \vdots & \vdots & & & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mk} & \cdots & a_{mn} \end{bmatrix}$$

Under the right conditions it is possible to add, subtract and multiply two matrices. We'll save multiplication for a little, but we have the following:

DEFINITION 3.1.1: **Adding and Subtracting Matrices**

When two matrices have the same dimensions, they are added or subtracted by adding or subtracting their corresponding entries.

◇ **Example 3.1(a):** Determine which of the matrices below can be added, and add those that can be.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} -5 & 1 \\ 0 & 4 \\ 2 & -3 \end{bmatrix}, \qquad C = \begin{bmatrix} -7 & 4 \\ 1 & 5 \end{bmatrix}$$

**Solution:** $B$ cannot be added to either $A$ or $C$, but

$$A + C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} -7 & 4 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 3 & 8 \end{bmatrix}$$

It should be clear that $A$ and $C$ could be subtracted, and that $A+C=C+A$ but $A-C \neq C-A$.

> DEFINITION 3.1.2: **Scalar Times a Matrix**
>
> The result of a scalar $c$ times a matrix $A$ is the matrix each of whose entries are $c$ times the corresponding entry of $A$.

◇ **Example 3.1(b):** For the matrix $A = \begin{bmatrix} 3 & 1 & -1 \\ -2 & 8 & 5 \\ 6 & -4 & -3 \end{bmatrix}$, find $3A$.

$$3A = 3\begin{bmatrix} 3 & 1 & -1 \\ -2 & 8 & 5 \\ 6 & -4 & -3 \end{bmatrix} = \begin{bmatrix} 9 & 3 & -3 \\ -6 & 24 & 15 \\ 18 & -12 & -9 \end{bmatrix}$$

---

**Section 3.1 Exercises** To Solutions

1. (a) Give the dimensions of matrices $A$, $B$ and $C$ in Exercise 3 below.

   (b) Give the entries $b_{31}$ and $c_{23}$ of the matrices $B$ and $C$ in Exercise 3 below.

2. Give the names of the matrices at the top of the next page that are

   (a) square (b) symmetric (c) diagonal

   (d) upper triangular (e) lower triangular

$$A = \begin{bmatrix} -4 & 0 & 0 \\ 2 & 6 & 0 \\ -1 & 5 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 1 & -4 & 3 \\ -4 & 5 & 2 \end{bmatrix} \quad C = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

$$D = \begin{bmatrix} 1 & -4 \\ -4 & 5 \end{bmatrix} \quad E = \begin{bmatrix} 2 & 5 & -3 & 1 \\ 0 & 4 & 0 & 2 \\ 0 & 0 & -1 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix} \quad F = \begin{bmatrix} 2 & 7 \\ -1 & 4 \\ 6 & 5 \end{bmatrix}$$

$$G = \begin{bmatrix} 3 & 4 & -5 & 0 \\ 4 & 1 & 7 & 0 \\ -5 & 7 & -2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \quad H = \begin{bmatrix} 3 & 1 \\ 4 & 5 \end{bmatrix} \quad J = \begin{bmatrix} 3 & 2 & -1 \\ 2 & 1 & 3 \\ -1 & 3 & -5 \end{bmatrix}$$

3. Give examples of each of the following types of matrices.

   (a) lower triangular (b) diagonal

   (c) symmetric (d) identity

   (e) upper triangular but not diagonal (f) symmetric but without any zero entries

   (g) symmetric but not diagonal (h) diagonal but not a multiple of an identity

4. Give the transpose of each matrix. Use the correct notation to denote the transpose.

$$A = \begin{bmatrix} 1 & 0 & 5 \\ -3 & 1 & -2 \\ 4 & 7 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 \\ -3 & 1 \\ 4 & 7 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & -1 & 3 \\ -3 & 1 & 2 & 0 \\ 4 & 7 & 0 & -2 \end{bmatrix} \qquad D = \begin{bmatrix} 1 & -3 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}$$

5. Give all possible sums and differences of matrices from Exercise 4.

6. Consider the matrix below and to the left.

   (a) What are some ways we could describe the matrix? Give all you can think of.

   (b) Below and to the right the dotted lines indicate how the matrix can be broken into four **blocks**, each of which is a $3 \times 3$ matrix. Give all ways you can think of to describe the matrix consisting of only the block in the upper left.

   (c) Give all ways you can think of to describe the matrix consisting of the block in the lower right.

   (d) Give all ways you can think of to describe the matrices in the upper right and lower left. There is one more word to describe them that was not given in the section. What do you think it is?

$$\begin{bmatrix} \frac{1}{E} & -\frac{\nu}{E} & -\frac{\nu}{E} & 0 & 0 & 0 \\ -\frac{\nu}{E} & \frac{1}{E} & -\frac{\nu}{E} & 0 & 0 & 0 \\ -\frac{\nu}{E} & -\frac{\nu}{E} & \frac{1}{E} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{G} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{G} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G} \end{bmatrix} \qquad \left[\begin{array}{ccc:ccc} \frac{1}{E} & -\frac{\nu}{E} & -\frac{\nu}{E} & 0 & 0 & 0 \\ -\frac{\nu}{E} & \frac{1}{E} & -\frac{\nu}{E} & 0 & 0 & 0 \\ -\frac{\nu}{E} & -\frac{\nu}{E} & \frac{1}{E} & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & \frac{1}{G} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{G} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G} \end{array}\right]$$

7. (a) For the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, find the matrix $B = A + A^T$.

   (b) What kind of matrix is $B$?

## 3.2 Matrix Times a Vector

**Performance Criteria:**

3. (d) Multiply a matrix times a vector, give the linear combination form of a matrix times a vector.
   (e) Give a specific matrix that is multiplied times an arbitrary vector to obtain a given resulting vector.
   (f) Express a system of equations as a coefficient matrix times a vector equalling another vector.
   (g) For a given matrix $A$ and vector $\vec{\mathbf{b}}$, find a vector $\vec{\mathbf{x}}$ for which $A\,\vec{\mathbf{x}} = \vec{\mathbf{b}}$.

### Matrix Times a Vector

Multiplying a matrix times a vector is in some sense *THE* foundational operation of linear algebra. When we study algebra, trigonometry and calculus, what we are really interested in how functions act on numbers. *In linear algebra we are interested in how matrices act on vectors*. This action is usually referred to as multiplication of a matrix times a vector, but we will be well served by remembering that it is truly an *action* of a matrix on a vector.

Before getting into how to do this, we need to devise a useful notation. Consider the matrix

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

Each column of $A$, taken by itself, is a vector. We'll refer to the first column as the vector $\vec{\mathbf{a}}_{*1}$, with the asterisk * indicating that the row index will range through all values, and the $1$ indicating that the values all come out of column one. Of course $\vec{\mathbf{a}}_{*2}$ denotes the second column, and so on. Similarly, $\vec{\mathbf{a}}_{1*}$ will denote the first row, $\vec{\mathbf{a}}_{2*}$ the second row, etc. Technically speaking, the rows are not vectors, but we'll call them **row vectors** and we'll call the columns **column vectors**. If we use just the word *vector*, we will mean a column vector.

◇ **Example 3.2(a):** Give $\vec{\mathbf{a}}_{2*}$ and $\vec{\mathbf{a}}_{*3}$ for the matrix $A = \begin{bmatrix} -5 & 3 & 4 & -1 \\ 7 & 5 & 2 & 4 \\ 2 & -1 & -6 & 0 \end{bmatrix}$

$$\vec{\mathbf{a}}_{2*} = \begin{bmatrix} 7 & 5 & 2 & 4 \end{bmatrix} \qquad \text{and} \qquad \vec{\mathbf{a}}_{*3} = \begin{bmatrix} 4 \\ 2 \\ -6 \end{bmatrix}$$

A row vector can be multiplied times a column vector (*only in that order*) if they have the same number of components. To do this, we simply multiply each entry of the row vector times each entry of the column vector and add all the results. *The result of a row vector times a column vector is then a single number.* This is shown in the following example.

◇ **Example 3.2(b):** Multiply $\vec{\mathbf{u}} = \begin{bmatrix} 5 & -3 & 2 \end{bmatrix}$ times $\vec{\mathbf{v}} = \begin{bmatrix} -4 \\ 2 \\ 1 \end{bmatrix}$.

$\vec{\mathbf{u}}\vec{\mathbf{v}} = (5)(-4) + (-3)(2) + (2)(1) = -20 + (-6) + 2 = -24.$ Note that if $\vec{\mathbf{u}}$ was a true (column) vector, this would be $\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}$.

---

Technically, the result $-24$ should be the $1 \times 1$ matrix $[-24]$, but we are really only interested in the above operation as it relates to taking a matrix times a vector. Once we are able to do that, there is no need to do the above operation except as a part of the process of taking a matrix times a vector. We now define how to multiply a matrix times a vector.

**Definition 3.2.1: Matrix Times a Vector**

An $m \times n$ matrix $A$ can be multiplied times a vector $\vec{\mathbf{x}}$ with $n$ components. The result is a vector with $m$ components, the $i$th component being the product of the $i$th row of $A$ with $\vec{\mathbf{x}}$, as shown below.

$$A\,\vec{\mathbf{x}} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{a}}_{1*}\vec{\mathbf{x}} \\ \vec{\mathbf{a}}_{2*}\vec{\mathbf{x}} \\ \vdots \\ \vec{\mathbf{a}}_{m*}\vec{\mathbf{x}} \end{bmatrix}$$

◇ **Example 3.2(c):** Multiply $\begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \\ 1 & -6 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -7 \end{bmatrix}$.

$$\begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \\ 1 & -6 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ -7 \end{bmatrix} = \begin{bmatrix} (3)(2) + (0)(1) + (-1)(-7) \\ (-5)(2) + (2)(1) + (4)(-7) \\ (1)(2) + (-6)(1) + (0)(-7) \end{bmatrix} = \begin{bmatrix} 13 \\ -36 \\ -4 \end{bmatrix}$$

---

There is no need for the matrix multiplying a vector to be square, but when it is not, the resulting vector is not the same length as the original vector:

◇ **Example 3.2(d):** Find $A\vec{\mathbf{x}}$ for $A = \begin{bmatrix} 7 & -4 & 2 \\ -1 & 0 & 6 \end{bmatrix}$ and $\vec{\mathbf{x}} = \begin{bmatrix} 3 \\ -5 \\ 1 \end{bmatrix}$.

$$A\vec{\mathbf{x}} = \begin{bmatrix} 7 & -4 & 2 \\ -1 & 0 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -5 \\ 1 \end{bmatrix} = \begin{bmatrix} (7)(3) + (-4)(-5) + (2)(1) \\ (-1)(3) + (0)(-5) + (6)(1) \end{bmatrix} = \begin{bmatrix} 43 \\ 3 \end{bmatrix}$$

---

We sometimes describe what we see in the above example by saying that the matrix $A$ *transforms* the vector $\vec{\mathbf{x}} = \begin{bmatrix} 3 \\ -5 \\ 1 \end{bmatrix}$ into the vector $\begin{bmatrix} 43 \\ 3 \end{bmatrix}$. This emphasizes the fact that we are interested in a matrix times a vector as more than just an algebraic computation.

Note that, if a matrix $A$ is $1 \times n$ and $\vec{\mathbf{x}}$ is a vector with $n$ components, the multiplication $\vec{\mathbf{x}}A$ can be performed, but such multiplication do not arise naturally in applications. We will *always* multiply a matrix times a vector, in that order, so we always write $A\vec{\mathbf{x}}$ rather than $\vec{\mathbf{x}}A$.

### Matrix Times a Vector Form of a System

We now see an example that has important implications.

◇ **Example 3.2(e):** Multiply $A = \begin{bmatrix} 1 & -1 & 2 \\ -3 & 4 & -2 \\ 2 & 1 & 5 \end{bmatrix}$ times the vector $\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

$$A\vec{\mathbf{x}} = \begin{bmatrix} 1 & -1 & 2 \\ -3 & 4 & -2 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 + 2x_3 \\ -3x_1 + 4x_2 - 2x_3 \\ 2x_1 + x_2 + 5x_3 \end{bmatrix}$$

Consider now the system shown below and to the left:

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 5 \\ -3x_1 + 4x_2 - 2x_3 &= -1 \\ 2x_1 + x_2 + 5x_3 &= 2 \end{aligned} \qquad \begin{bmatrix} x_1 - x_2 + 2x_3 \\ -3x_1 + 4x_2 - 2x_3 \\ 2x_1 + x_2 + 5x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$$

Because two vectors are equal if their corresponding components are equal, we can rewrite the system in the form shown above and to the right and, as a result of Example 3.2(e), this leads to

$$\begin{bmatrix} 1 & -1 & 2 \\ -3 & 4 & -2 \\ 2 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}.$$

This is what we will refer to as the **matrix times a vector form** of a system of equations.

**Definition 3.2.2 Matrix Times a Vector Form of a System**

A system

$$\begin{array}{rcl} a_{11}x_1 + \ a_{12}x_2 + \cdots + \ a_{1n}x_n & = & b_1 \\ a_{21}x_1 + \ a_{22}x_2 + \cdots + \ a_{2n}x_n & = & b_2 \\ & \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \end{array}$$

of $m$ linear equations in $n$ unknowns can be written as a matrix times a vector equalling another vector:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

We will refer to this as the **matrix times a vector form of the system of equations**, and we express it compactly as $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$.

◇ **Example 3.2(f):** Give the matrix times vector form of the system $\begin{array}{rcr} 3x_1 + 5x_2 & = & -1 \\ x_1 + 4x_2 & = & 2 \end{array}$.

**Solution:** The matrix times vector form of the system is $\begin{bmatrix} 3 & 5 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$.

---

Now we can solve the "inverse problem" of a matrix times a vector:

◇ **Example 3.2(g):** Let $A = \begin{bmatrix} 3 & 5 \\ 1 & 4 \end{bmatrix}$ and $\vec{\mathbf{b}} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$. Find a vector $\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ for which $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$.

**Solution:** We are trying to solve the matrix-vector equation $\begin{bmatrix} 3 & 5 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$. This is the matrix times a vector form of the system

$$\begin{array}{rcr} 3x_1 + \ x_2 & = & 4 \\ x_1 + 4x_2 & = & -1 \end{array},$$

which we can solve by the addition method or row-reduction to obtain $\vec{\mathbf{x}} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

---

**Linearity of the Action of a Matrix on a Vector**

Multiplication of vectors by matrices has the following important properties, which are easily verified.

> THEOREM 3.2.3
>
> Let $A$ and $B$ be matrices, $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ be vectors, and $c$ be any scalar. Assuming that all the indicated operations below are defined (possible), then
>
> (a) $A(\vec{\mathbf{x}}+\vec{\mathbf{y}}) = A\vec{\mathbf{x}} + A\vec{\mathbf{y}}$  (b) $A(c\vec{\mathbf{x}}) = c(A\vec{\mathbf{x}})$
>
> (c) $(A+B)\vec{\mathbf{x}} = A\vec{\mathbf{x}} + B\vec{\mathbf{x}}$

We now come to a very important idea that depends on the first two properties of Theorem 3.2.3. When we act on a mathematical object with another object, the object doing the "acting on" is often called an **operator**. Some operators you are familiar with are the derivative operator and the antiderivative operator (indefinite integral), which act on functions to create other functions. Note that the derivative operator has the following two properties, for any functions $f$ and $g$ and real number $c$:

$$\frac{d}{dx}(f+g) = \frac{df}{dx} + \frac{dg}{dx}, \qquad \frac{d}{dx}(cf) = c\frac{df}{dx}$$

These are the same as the first two properties above for multiplication of a vector by a matrix. A matrix can be thought of as an operator that operates on vectors (through multiplication). The first two properties of multiplication of a vector by a matrix, as well as the corresponding properties of the derivative, are called the **linearity properties**. Both the derivative operator and the matrix multiplication operator are then called **linear operators**. This is why this subject is called *linear* algebra!

**Linear Combination Form of a Matrix Times a Vector**

There is another way to compute a matrix times a vector. It is not as efficient to do by hand as the method implied by Definition 3.2.1, but it will be very important conceptually quite soon. Using our earlier definition of a matrix $A$ times a vector $\vec{\mathbf{x}}$, we see that

$$A\vec{\mathbf{x}} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \begin{bmatrix} a_{21}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{m2}x_2 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix}$$

$$= x_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2\begin{bmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Let's think about what the above shows. It gives us the result below, which is illustrated in Examples 3.2(h) and (i).

**Linear Combination Form of a Matrix Times a Vector**

The product of a matrix $A$ and a vector $\vec{\mathbf{x}}$ is a linear combination of the columns of $A$, with the scalars being the corresponding components of $\vec{\mathbf{x}}$.

◇ **Example 3.2(h):** Give the linear combination form of $\begin{bmatrix} 7 & -4 & 2 \\ -1 & 0 & 6 \end{bmatrix}\begin{bmatrix} 3 \\ -5 \\ 1 \end{bmatrix}$.

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

◇ **Example 3.2(i):** Give the linear combination form of $\begin{bmatrix} 1 & 3 & -2 \\ 3 & 7 & 1 \\ -2 & 1 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

$$\begin{bmatrix} 1 & 3 & -2 \\ 3 & 7 & 1 \\ -2 & 1 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ 7 \\ 1 \end{bmatrix} + x_3\begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix}$$

**Section 3.2 Exercises** To Solutions

1. Multiply $\begin{bmatrix} 1 & 0 & -1 \\ -3 & 1 & 2 \\ 4 & 7 & -2 \end{bmatrix}\begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 3 & -4 & 0 \\ -1 & 5 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 5 \\ 2 \end{bmatrix}$ **by hand**.

2. Multiply each of the following by hand, when possible. Some are on the next page.

(a) $\begin{bmatrix} 3 & -1 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} -4 \\ 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & -5 & 2 \\ 6 & 3 & -4 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & -5 & 2 \\ 6 & 3 & -4 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 6 \\ -5 & 3 \\ 2 & -4 \end{bmatrix}\begin{bmatrix} 3 \\ 5 \end{bmatrix}$

(e) $\begin{bmatrix} 7 & -2 & 0 & 4 \\ 1 & 5 & 3 & -3 \\ 2 & 1 & -1 & 5 \\ -3 & 7 & 2 & 1 \end{bmatrix}\begin{bmatrix} -4 \\ 1 \\ 3 \\ 2 \end{bmatrix}$

(f) $\begin{bmatrix} 1 & 0 & -5 \\ 2 & 2 & 3 \\ -4 & 7 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$

3. Give each of the products from Exercise 2 in linear combination form.

4. Give the product $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ as

   (a) a single vector (b) a linear combination of vectors

5. (a) Find a matrix $A$ such that $A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 - 5x_2 \\ x_1 + x_2 \end{bmatrix}$.

   (b) Find a matrix $B$ such that $B\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 3x_2 \\ 2x_1 - x_2 \\ 5x_1 + 4x_2 \end{bmatrix}$.

   (c) Find a matrix $C$ such that $C\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 4x_2 - x_3 \\ -5x_1 + x_2 + 2x_3 \\ x_1 + 3x_2 \end{bmatrix}$.

   (d) Find a matrix $D$ such that $D\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_3 \\ x_1 - x_2 - x_3 \end{bmatrix}$.

6. Give the matrix times a vector form of each system:

   (a)
$$\begin{aligned} x + y - 3z &= 1 \\ -3x + 2y - z &= 7 \\ 2x + y - 4z &= 0 \end{aligned}$$

   (b)
$$\begin{aligned} 5x_1 \phantom{+ x_2} + x_3 &= -1 \\ 2x_2 + 3x_3 &= 0 \\ 2x_1 + x_2 - 4x_3 &= 2 \end{aligned}$$

   (c)
$$\begin{aligned} b + 0.5m &= 8.1 \\ b + 1.0m &= 6.9 \\ b + 1.5m &= 6.2 \\ b + 2.0m &= 5.3 \\ b + 2.5m &= 4.5 \\ b + 3.0m &= 3.8 \\ b + 3.5m &= 3.0 \end{aligned}$$

   (d)
$$\begin{aligned} x_1 - 4x_2 + x_3 + 2x_4 &= -1 \\ 3x_1 + 2x_2 - x_3 - 7x_4 &= 0 \\ -2x_1 + x_2 - 4x_3 + x_4 &= 2 \end{aligned}$$

7. For each part of this exercise, find a vector $\vec{\mathbf{x}}$ for which $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ for the $A$ and $\vec{\mathbf{b}}$ given.

   (a) $A = \begin{bmatrix} 3 & -1 & 5 \\ 2 & 0 & 2 \\ -1 & 4 & -3 \end{bmatrix}, \quad \vec{\mathbf{b}} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$

   (b) $A = \begin{bmatrix} 1 & -3 \\ 5 & -2 \end{bmatrix}, \quad \vec{\mathbf{b}} = \begin{bmatrix} 24 \\ 29 \end{bmatrix}$

   (c) $A = \begin{bmatrix} 1 & 2.1 \\ 1 & 2.7 \\ 1 & 3.2 \\ 1 & 3.9 \end{bmatrix}, \quad \vec{\mathbf{b}} = \begin{bmatrix} 2.54 \\ 2.78 \\ 2.98 \\ 3.26 \end{bmatrix}$

   (d) $A = \begin{bmatrix} 1 & 3 & 0 & 0 \\ 3 & 1 & 2 & 0 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & 3 & 1 \end{bmatrix}, \quad \vec{\mathbf{b}} = \begin{bmatrix} 13 \\ 19 \\ 11 \\ 7 \end{bmatrix}$

8. The relationship between stress and strain in a small cube of solid material can be expressed by the matrix equation

$$\begin{bmatrix} \frac{1}{E} & -\frac{\nu}{E} & -\frac{\nu}{E} & 0 & 0 & 0 \\ -\frac{\nu}{E} & \frac{1}{E} & -\frac{\nu}{E} & 0 & 0 & 0 \\ -\frac{\nu}{E} & -\frac{\nu}{E} & \frac{1}{E} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{G} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{G} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{G} \end{bmatrix} \begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \tau_{xy} \\ \tau_{yz} \\ \tau_{zx} \end{bmatrix} = \begin{bmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ \gamma_{xy} \\ \gamma_{yz} \\ \gamma_{zx} \end{bmatrix}$$

where
$E$ is elastic modulus
$G$ is shear modulus
$\nu$ is Poisson's ratio
$\epsilon$ is normal strain
$\gamma$ is shear strain
$\sigma$ is normal stress
$\tau$ is shear stress

and the subscripts indicate faces of the cube and directions of forces. The vector being multiplied is the stress vector, and the right hand side is the strain vector.

(a) Give the relationship expressed by the product of the second row of the matrix times the stress vector.

(b) Give an expression for the $\epsilon_{zz}$ strain in terms of the three stresses $\sigma_{xx}$, $\sigma_{yy}$, $\sigma_{zz}$ and the parameters $E$ and $\nu$.

(c) Give the relationship between $\tau_{zx}$ and $\gamma_{zx}$.

(d) The coefficient matrix is what we call a **block matrix** made up of blocks, each of which is a smaller matrix. In this case there are four $3 \times 3$ matrices, two of which are zero matrices. Each of the nonzero blocks is multiplied by only a portion of the stress vector and gives only a portion of the strain vector. Write the two matrix equations for the nonzero blocks.

## 3.3 Actions of Matrices on Vectors: Transformations in $\mathbb{R}^2$

**Performance Criteria:**

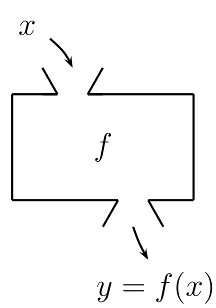
3. (h) Determine whether a matrix is a projection matrix, reflection matrix or rotation matrix, or none of these, by its action on a few vectors.

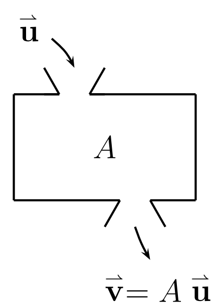
   (i) Determine whether a vector is an eigenvector of a matrix. If it is, give the corresponding eigenvalue.

Most of the mathematics that you have studied revolves around the idea of a function, which is simply a rule that assigns to every number in the *domain* of the function another number. There need not be any logic to how this is done, but usually there is, and that logic is seen in the fact that the function is given as some sort of specific relationship, like

$$f(x) = x^2 - 5x + 2\,, \qquad y = \sqrt{x+3}\,, \qquad g(t) = 3\sin(2t-5)\,, \qquad h(x) = \ln x.$$

(The last two of these are sort of redundant, in the sense that they are built out of other functions, sine and the natural logarithm.) For those who are visually or mechanically inclined, a function can also be thought of as a "machine," commonly named $f$, whose input is a number $x$ and output is some other number $y$. We write $y = f(x)$ and say "$y$ equals $f$ of $x$" to indicate that $y$ is the result of the function $f$ acting on $x$. This is shown in the picture below and to the left.



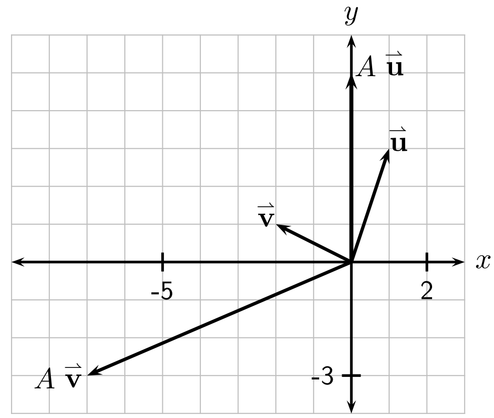

When we multiply a vector by a matrix, the result is another vector - *this is essentially the same idea as a function, but with vectors playing the role of numbers and a matrix taking the place of the function*. This is shown in the picture above and to the right. We should really think of a matrix times a vector as the matrix acting on the vector to create another vector. We sometimes say that the matrix $A$ *transforms* the original vector $\vec{\mathbf{u}}$ to the new one $\vec{\mathbf{v}} = A\,\vec{\mathbf{u}}$. This happens by the purely computational means that you learned in the previous section.

One powerful tool in the study of any function is its graphical representation. Unfortunately, it is difficult to graphically represent the action of a matrix on all vectors - we instead must picture the action of a matrix on vectors one or two vectors at a time. Even that is difficult with vectors in $\mathbb{R}^3$ and impossible in higher dimensions. Let's see how we do it in $\mathbb{R}^2$. Suppose that we have the matrix and vectors

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

for which

$$A\,\vec{\mathbf{u}} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}, \qquad A\,\vec{\mathbf{v}} = \begin{bmatrix} -7 \\ -3 \end{bmatrix}.$$




The graph to the right shows the vectors $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, $A\,\vec{\mathbf{u}}$ and $A\,\vec{\mathbf{v}}$, and we can see the result of $A$ acting on $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$. The two attributes of any vector are of course magnitude and direction, and we can see that $A$ altered the magnitude and the direction of both vectors, but apparently not in any special way.

In general, when a matrix acts on a vector the resulting vector will have a different magnitude and direction than the original vector, with the change in magnitude and direction being different for most vectors. There are a few notable exceptions to this:

- The matrix that acts on a vector without actually changing it at all is called the **identity matrix**. Clearly, then, when the identity matrix acts on a vector, neither the direction or magnitude is changed.

- A matrix that rotates every vector in $\mathbb{R}^2$ through a fixed angle $\theta$ is called a **rotation matrix**. In this case the direction changes, but not the magnitude. (Of course the direction doesn't change if $\theta = 0°$ and, in some sense, if $\theta = 180°$. In the second case, even though the direction is opposite, the resulting vector is still just a scalar multiple of the original.)

- For most matrices there are certain vectors, called **eigenvectors** whose directions don't change (other than perhaps reversing) when acted on by by the matrix under consideration. In those cases, the effect of multiplying such a vector by the matrix is the same as multiplying the vector by a scalar. This has very useful applications.

In the following exercises you will see rotation matrices and eigenvectors, along with some other matrices that do interesting things to vectors geometrically.

**Section 3.2 Exercises**

For the following exercises you will be multiplying each of several vectors by a given matrix and trying to see what the matrix does to the vectors. This can be pretty tedious by hand, so I would suggest that you use the UCSMP Polygon Plotter that you can link to from the class web page (or find with a web search for "UCSMP polygon plotter"). You will need to enter each vector as a position vector from the origin, and then transform it by the transformation matrix you are working with. Here's how you do all that:

- Under "Enter New" you are asked "How Many Points". Enter 2, meaning you are creating a polygon with only two vertices (a line segment!).

- Below that you are to describe your polygon as a "matrix." The first column should be zeros, and the second column should be the components of your vector. (The first column is the coordinates of the initial point, and the second column is the coordinates of the terminal point.)

- Once you put the values in for the vector, click "Enter" and you should see your vector, in red.

- Enter your transformation matrix. For the ones with with entries containing roots, use their decimal representations rounded to the thousandth's place. **Fractions may be entered as they are**.
- Once you have entered the transformation matrix, click "Transform!" You will see the result of the transformation as a vector in black.
- If you now look below "Select Polygon:" you will see your vector under "Preimage: AB," and the result of the transformation as "Image: A'B'."
- To multiply another vector by the same transformation, enter the new vector and click "Enter" again, followed by "Transform!"
- To repeat all this with a different transformation matrix, click "Clear Grid" at the bottom and start over.

1. Let $A = \begin{bmatrix} 3 & -1 \\ -4 & 0 \end{bmatrix}$, $\vec{\mathbf{u}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\vec{\mathbf{w}} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$.

   (a) Find $A\,\vec{\mathbf{u}}$, $A\,\vec{\mathbf{v}}$ and $A\,\vec{\mathbf{w}}$ **by hand**. Check your answers with the polygon plotter.

   (b) Plot and label $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, $\vec{\mathbf{w}}$, $A\,\vec{\mathbf{u}}$, $A\,\vec{\mathbf{v}}$ and $A\,\vec{\mathbf{w}}$ on one $\mathbb{R}^2$ coordinate grid. Label each vector by putting its name near its tip.

   You should not see any special relationship between the vectors $\vec{\mathbf{x}}$ and $A\,\vec{\mathbf{x}}$ (where $\vec{\mathbf{x}}$ is to represent any vector) here.

2. Let $B = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$, $\vec{\mathbf{u}} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, $\vec{\mathbf{w}} = \begin{bmatrix} 3 \\ 4.5 \end{bmatrix}$.

   (a) Plot and label $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, $\vec{\mathbf{w}}$, $B\,\vec{\mathbf{u}}$, $B\,\vec{\mathbf{v}}$ and $B\,\vec{\mathbf{w}}$ on one $\mathbb{R}^2$ coordinate grid. Label each vector by putting its name near its tip.

   (b) You should be able to see that $B$ does not seem to change the length of a vector. To verify this, find $\|\,\vec{\mathbf{w}}\,\|$ and $\|\,B\,\vec{\mathbf{w}}\,\|$ to the nearest hundredth.

   (c) What does the matrix $B$ seem to do to every vector? **Think about the two attributes of any vector, direction and magnitude.**

   (d) The entries of $B$ should look familiar to you. What is special about $\frac{1}{2}$ and $\frac{\sqrt{3}}{2}$?

3. Let $C = \begin{bmatrix} \frac{16}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{9}{25} \end{bmatrix}$, $\vec{\mathbf{u}} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, $\vec{\mathbf{w}} = \begin{bmatrix} 4.5 \\ 6 \end{bmatrix}$.

   (a) Plot and label $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, $\vec{\mathbf{w}}$, $C\,\vec{\mathbf{u}}$, $C\,\vec{\mathbf{v}}$ and $C\,\vec{\mathbf{w}}$ on one $\mathbb{R}^2$ coordinate grid.

   (b) What does the matrix $C$ seem to do to every vector? (Does the magnitude change? Does the direction change?)

(c) Try $C$ times $\begin{bmatrix} -4 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ -4 \end{bmatrix}$. Hmmm...

(d) Can you see the role of the entries of the matrix here?

4. Let $D = \begin{bmatrix} \frac{7}{25} & \frac{24}{25} \\ \frac{24}{25} & -\frac{7}{25} \end{bmatrix}$, $\vec{\mathbf{u}} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, $\vec{\mathbf{w}} = \begin{bmatrix} 4.5 \\ 6 \end{bmatrix}$.

(a) Plot and label $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, $\vec{\mathbf{w}}$, $D\,\vec{\mathbf{u}}$, $D\,\vec{\mathbf{v}}$ and $D\,\vec{\mathbf{w}}$ on one $\mathbb{R}^2$ coordinate grid.

(b) What does the matrix $D$ seem to do to every vector? (Does the magnitude change? Does the direction change?)

(c) Try $D$ times $\begin{bmatrix} -4 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ -4 \end{bmatrix}$. Hmmm...

(d) Can you see the role of the entries of the matrix here?

5. Again let $A = \begin{bmatrix} 3 & -1 \\ -4 & 0 \end{bmatrix}$, (see Exercise 1) but let

$\vec{\mathbf{u}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\vec{\mathbf{w}} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$.

(a) Plot and label $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, $\vec{\mathbf{w}}$, $A\,\vec{\mathbf{u}}$, $A\,\vec{\mathbf{v}}$ and $A\,\vec{\mathbf{w}}$ on one $\mathbb{R}^2$ coordinate grid.

(b) For one of the vectors, there should be no apparent relationship between the vector and the result when it is multiplied by the matrix. Discuss what happened to the direction and magnitude of each of the other two vectors when the matrix acted on it.

(c) Pick one of your two vectors for which something special happened and multiply it by three, and multiply the result by $A$; what is the effect of multiplying by $A$ in this case?

(d) Pick the other special vector, multiply it by five, then by $A$. What effect does multiplying by $A$ have on the vector?

## 3.4 Multiplying Matrices

**Performance Criteria:**

3. (j) Know when two matrices can be multiplied, and know that matrix multiplication is not necessarily commutative. Multiply two matrices "by hand."

When two matrices have appropriate sizes they can be multiplied by a process you are about to see. Although the most reliable way to multiply two matrices and get the correct result is with a calculator or computer software, *it is very important that you get quite comfortable with the way that matrices are multiplied*. That will allow you to better understand certain conceptual things you will encounter later.

The process of multiplying two matrices is a bit clumsy to describe, but I'll do my best here. First I will try to describe it informally, then I'll formalize it with a definition based on some special notation. To multiply two matrices we just multiply each row of the first with each column of the second as we did when multiplying a matrix times a vector, with the results becoming the elements of the second matrix. Here is an informal description of the process:

(1) Multiply the first row of the first matrix times the first column of the second. The result is the (1,1) entry (first row, first column) of the product matrix.

(2) Multiply the first row of the first matrix times the second column of the second. The result is the (1,2) entry (first row, second column) of the product matrix.

(3) Continue multiplying the first row of the first matrix times each column of the second to fill out the first row of the product matrix, stopping after multiplying the first row of the first matrix times the last column of the second matrix.

(4) Begin filling out the second row of the product matrix by multiplying the *second* row of the first matrix times the first column of the second matrix to get the (2,1) entry (second row, first column) of the product matrix.

(5) Continue multiplying the second row of the first matrix times each column of the second until the second row of the product matrix is filled out.

(6) Continue multiplying each row of the first matrix times each column of the second until the last row of the first has been multiplied times the last column of the second, at which point the product matrix will be complete.

Note that for this to work the number of columns of the first matrix must be equal the number of rows of the second matrix. Let's look at an example.

◇ **Example 3.4(a):** For $A = \begin{bmatrix} -5 & 1 \\ 0 & 4 \\ 2 & -3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 7 & 3 \end{bmatrix}$, find the product $AB$.

Video Example

**Solution:**

$$AB = \begin{bmatrix} -5 & 1 \\ 0 & 4 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 7 & 3 \end{bmatrix} = \begin{bmatrix} -5(1)+1(7) & -5(2)+1(3) \\ 0(1)+4(7) & 0(2)+4(3) \\ 2(1)+(-3)(7) & 2(2)+(-3)(3) \end{bmatrix} = \begin{bmatrix} 2 & -7 \\ 28 & 12 \\ -19 & -5 \end{bmatrix}$$

In order to make a formal definition of matrix multiplication, we need to remember the special notation from Section 3.2 for rows and columns of a matrix. Given a matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

we refer to, for example the third row as $\vec{\mathbf{a}}_{3*}$. Here the first subscript $3$ indicates that we are considering the third row, and the $*$ indicates that we are taking the elements from the third row *in all columns*. Therefore $\vec{\mathbf{a}}_{3*}$ refers to a $1\times n$ matrix. Similarly, $\vec{\mathbf{a}}_{*2}$ is the vector that is the second column of $A$. So we have

$$\vec{\mathbf{a}}_{3*} = \begin{bmatrix} a_{31} & a_{32} & \cdots & a_{3n} \end{bmatrix} \qquad\qquad \vec{\mathbf{a}}_{*2} = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{2m} \end{bmatrix}$$

A $1\times n$ matrix like $\vec{\mathbf{a}}_{3*}$ can be thought of like a vector; in fact, we sometimes call such a matrix a **row vector**. Note that the transpose of such a vector is a column vector. We then define a product like product $\vec{\mathbf{a}}_{i*}\vec{\mathbf{b}}_{*j}$ by

$$\vec{\mathbf{a}}_{i*}\vec{\mathbf{b}}_{*j} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

This is the basis for the following formal definition of the product of two matrices.

Definition 3.4.1: **Matrix Multiplication**

Let $A$ be an $m \times n$ matrix whose *rows* are the vectors $\vec{\mathbf{a}}_{1*}, \vec{\mathbf{a}}_{2*}, ..., \vec{\mathbf{a}}_{m*}$ and let $B$ be an $n \times p$ matrix whose *columns* are the vectors $\vec{\mathbf{b}}_{*1}, \vec{\mathbf{b}}_{*2}, ..., \vec{\mathbf{b}}_{*p}$. Then $AB$ is the $m \times p$ matrix

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}$$

$$= \begin{bmatrix} \vec{\mathbf{a}}_{1*} \\ \vec{\mathbf{a}}_{2*} \\ \vdots \\ \vec{\mathbf{a}}_{m*} \end{bmatrix} \begin{bmatrix} \vec{\mathbf{b}}_{*1} & \vec{\mathbf{b}}_{*2} & \vec{\mathbf{b}}_{*3} & \cdots & \vec{\mathbf{b}}_{*p} \end{bmatrix}$$

$$= \begin{bmatrix} \vec{\mathbf{a}}_{1*}\vec{\mathbf{b}}_{*1} & \vec{\mathbf{a}}_{1*}\vec{\mathbf{b}}_{*2} & \cdots & \vec{\mathbf{a}}_{1*}\vec{\mathbf{b}}_{*p} \\ \vec{\mathbf{a}}_{2*}\vec{\mathbf{b}}_{*1} & \vec{\mathbf{a}}_{2*}\vec{\mathbf{b}}_{*2} & \cdots & \vec{\mathbf{a}}_{2*}\vec{\mathbf{b}}_{*p} \\ \vdots & \vdots & \ddots & \vdots \\ \vec{\mathbf{a}}_{m*}\vec{\mathbf{b}}_{*1} & \vec{\mathbf{a}}_{m*}\vec{\mathbf{b}}_{*2} & \cdots & \vec{\mathbf{a}}_{m*}\vec{\mathbf{b}}_{*p} \end{bmatrix}$$

For the above computation to be possible, products in the last matrix.This implies that the number of columns of $A$ must equal the number of rows of $B$.

◇ **Example 3.4(b):** For $C = \begin{bmatrix} -5 & 1 & -2 \\ 7 & 0 & 4 \\ 2 & -3 & 6 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -7 & 0 \\ 5 & 2 & 3 \end{bmatrix}$, find $CD$ and $DC$.

**Solution:**

$$CD = \begin{bmatrix} -5 & 1 & -2 \\ 7 & 0 & 4 \\ 2 & -3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ -3 & -7 & 0 \\ 5 & 2 & 3 \end{bmatrix} = \begin{bmatrix} -5-3-10 & -10-7-4 & 5+0-6 \\ 7+0+20 & 14+0+8 & -7+0+12 \\ 2+9+30 & 4+21+12 & -2+0+18 \end{bmatrix}$$

$$= \begin{bmatrix} -18 & -21 & -1 \\ 27 & 22 & 5 \\ 41 & 37 & 16 \end{bmatrix}$$

$$DC = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -7 & 0 \\ 5 & 2 & 3 \end{bmatrix} \begin{bmatrix} -5 & 1 & -2 \\ 7 & 0 & 4 \\ 2 & -3 & 6 \end{bmatrix} = \begin{bmatrix} 7 & 4 & 0 \\ -34 & -3 & -22 \\ -5 & -4 & 16 \end{bmatrix}$$

We want to notice in the last example that $CD \neq DC$! This illustrates something very important:

**Matrix multiplication is not necessarily commutative! That is, given two matrices $A$ and $B$, it is not necessarily true that $AB = BA$. It is possible, but is not "usually" the case. In fact, one of $AB$ and $BA$ might exist and the other not.**

This is not just a curiosity; the above fact will have important implications in how certain computations are done. The next example, along with Example 3.4(a), shows that one of the two products might exist and the other not.

◇ **Example 3.4(c):** For the same matrices $A = \begin{bmatrix} -5 & 1 \\ 0 & 4 \\ 2 & -3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 7 & 3 \end{bmatrix}$ from Example 3.4(a), find the product $BA$.

**Solution:** When we try to multiply $BA = \begin{bmatrix} 1 & 2 \\ 7 & 3 \end{bmatrix} \begin{bmatrix} -5 & 1 \\ 0 & 4 \\ 2 & -3 \end{bmatrix}$ it is not even possible. We can't find the dot product of a row of $B$ with a column of $A$ because, as vectors, they don't have the same number of components. Therefore the product $BA$ does not exist.

---

◇ **Example 3.4(d):** For $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 \\ 7 & 3 \end{bmatrix}$, find the products $I_2B$, $BI_2$, $CB$ and $BC$.

**Solution:**

$$I_2B = BI_2 = \begin{bmatrix} 1 & 2 \\ 7 & 3 \end{bmatrix}, \qquad CB = BC = \begin{bmatrix} 3 & 6 \\ 21 & 9 \end{bmatrix}$$

---

The notation $I_2$ here means the $2 \times 2$ **identity matrix**. Note that when it is multiplied by another matrix $A$ *on either side* the result is just the matrix $A$.

Let's take a minute to think a bit more about the idea of an "identity." In the real numbers we say zero is the **additive identity** because adding it to any real number $a$ does not change the value of the number:

$$a + 0 = 0 + a = a$$

Similarly, the number one is the multiplicative identity:

$$a \times 1 = 1 \times a = a$$

Here the symbol $\times$ is just multiplication of real numbers. When we talk about an identity matrix, we are talking about a *multiplicative* identity, like the number one. There is no confusion because even though there are matrices that could be considered to be additive identities, they are not useful, so we don't consider them. When the size of the identity matrix is clear from the context, which is almost always the case, we omit the subscript and just write $I$.

There are many other special and/or interesting things that can happen when multiplying two matrices. Here's an example that shows that we can take powers of a matrix if it is a square matrix.

◇ **Example 3.4(e):** For the matrix $A = \begin{bmatrix} 3 & 1 & -1 \\ -2 & 8 & 5 \\ 6 & -4 & -3 \end{bmatrix}$, find $A^2$ and $A^3$.

**Solution:**

$$A^2 = AA = \begin{bmatrix} 3 & 1 & -1 \\ -2 & 8 & 5 \\ 6 & -4 & -3 \end{bmatrix}\begin{bmatrix} 3 & 1 & -1 \\ -2 & 8 & 5 \\ 6 & -4 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 15 & 5 \\ 8 & 42 & 27 \\ 8 & -14 & -17 \end{bmatrix}$$

$$A^3 = AA^2 = \begin{bmatrix} 3 & 1 & -1 \\ -2 & 8 & 5 \\ 6 & -4 & -3 \end{bmatrix}\begin{bmatrix} 1 & 15 & 5 \\ 8 & 42 & 27 \\ 8 & -14 & -17 \end{bmatrix} = \begin{bmatrix} 3 & 101 & 59 \\ 102 & 236 & 121 \\ -50 & -36 & - \end{bmatrix}$$

---

In Section 3.1 we saw that a scalar times a matrix is defined by multiplying each entry of the matrix by the scalar. With a little thought the following should be clear:

> THEOREM 3.4.2
>
> Let $A$ and $B$ be matrices for which the product $AB$ is defined, and let $c$ be any scalar. Then
> $$c(AB) = (cA)B = A(cB)$$

Note this carefully - when multiplying a product of two matrices by a scalar, we can instead multiply *one or the other, but NOT BOTH* of the two matrices by the scaler, then multiply the result with the remaining matrix.

Although one can do a great deal of study of matrices themselves, linear algebra is primarily concerned with the action of matrices on vectors. The following simple result is extremely important conceptually:

> THEOREM 3.4.3
>
> Let $A$ and $B$ be matrices and $\vec{\mathbf{x}}$ a vector. Assuming that all the indicated operations below are defined (possible), then
> $$(AB)\vec{\mathbf{x}} = A(B\vec{\mathbf{x}})$$

The following illustrates the difference between $(AB)\vec{\mathbf{x}}$ and $A(B\vec{\mathbf{x}})$ from a computational standpoint.

◇ **Example 3.4(g):** For the matrices $A = \begin{bmatrix} 1 & -1 \\ -2 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & -3 \\ 7 & 0 \end{bmatrix}$ and the vector $\vec{\mathbf{x}} = \begin{bmatrix} 3 \\ -6 \end{bmatrix}$, find $(AB)\vec{\mathbf{x}}$ and $A(B\vec{\mathbf{x}})$.

**Solution:** Be sure to note the difference between the how the two calculations are performed, along with the fact that the resluts are the same:

$$(AB)\vec{\mathbf{x}} = \left( \begin{bmatrix} 1 & -1 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 7 & 0 \end{bmatrix} \right) \begin{bmatrix} 3 \\ -6 \end{bmatrix} = \begin{bmatrix} -3 & -3 \\ 27 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -6 \end{bmatrix} = \begin{bmatrix} 9 \\ 45 \end{bmatrix}$$

$$A(B\vec{\mathbf{x}}) = \begin{bmatrix} 1 & -1 \\ -2 & 5 \end{bmatrix} \left( \begin{bmatrix} 4 & -3 \\ 7 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ -6 \end{bmatrix} \right) = \begin{bmatrix} 1 & -1 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} 30 \\ 21 \end{bmatrix} = \begin{bmatrix} 9 \\ 45 \end{bmatrix}$$

---

Let's now continue the analogy between functions and multiplication of a vector by a matrix. Consider the functions

$$f(x) = x^2 \qquad \text{and} \qquad g(x) = 2x - 1.$$

Suppose that we wanted to apply $g$ to the number three, and then apply $f$ to the result. We show this symbolically by

$$f\big[g(3)\big] = f\big[2(3) - 1\big] = f[5] = 25.$$

We can form a new function called the **composition** of $f$ and $g$, $f \circ g$. This function is *defined* for any value of $x$ by

$$(f \circ g)(x) = f\big[g(x)\big].$$

In the case of our particular $f$ and $g$ the composition is

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

and

$$(f \circ g)(3) = 4(3)^2 - 4(3) + 1 = 36 - 12 + 1 = 25,$$

showing that $f \circ g$ does indeed act on the number three just as $f$ and $g$ did in sequence. Let's reiterate - $f \circ g$ is a *single* new function that is equivalent to performing $g$ followed by $f$, *in that order*. If we were to perform the two functions on the number three, but in the opposite order, we would get

$$g\big[f(3)\big] = g\big[3^2\big] = g[9] = 2(9) - 1 = 17.$$

We also see that

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

so *the functions* $f \circ g$ *and* $g \circ f$ *are not the same!*

Now Theorem 3.4.3 tells us that, for two matrices $A$ and $B$ and any vector $\vec{\mathbf{x}}$,

$$(AB)\vec{\mathbf{x}} = A(B\vec{\mathbf{x}})$$

when all of the operations are defined. Here $A$ and $B$ can be thought of like the functions $f$ and $g$ above, except they act on vectors rather than numbers. The product $AB$ is like the composition $f \circ g$ - it is a single matrix whose action $\vec{\mathbf{x}}$ is equivalent to $B$ acting on $\vec{\mathbf{x}}$, and then $A$ acting on the result $B\vec{\mathbf{x}}$. The fact that $CD$ is not usually equal to $DC$ for two matrices $C$ and $D$ for which both $CD$ and $DC$ exist (see Example 3.4(b)) is analogous to the fact that the two compositions $f \circ g$ and $g \circ f$ of two functions $f$ and $g$ are not generally the same.

**Section 3.4 Exercises** To Solutions

1. Multiply (a) $\begin{bmatrix} 2 & -1 \\ -3 & 4 \end{bmatrix}\begin{bmatrix} 4 & 1 \\ 5 & -1 \end{bmatrix}$ and (b) $\begin{bmatrix} 1 & 2 & -1 \\ -3 & 4 & 1 \\ 2 & -1 & 3 \end{bmatrix}\begin{bmatrix} -2 & 4 & 1 \\ 3 & -5 & -1 \\ 0 & 1 & 2 \end{bmatrix}$ **by hand**.

2. For the following matrices, there are *THIRTEEN* multiplications possible, *including squaring* some of the matrices. Find and do as many of them do as many of them as you can. When writing your answers, tell which matrices you multiplied to get any particular answer. For example, it *IS* possible to multiply $A$ times $B$ (how about $B$ times $A$?), and you would then write

$$AB = \begin{bmatrix} -10 & 0 & 25 \\ -14 & 21 & -4 \end{bmatrix}$$

to give your answer. Now you have twelve left to find and do.

$$A = \begin{bmatrix} 0 & 5 \\ -3 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & -7 & 3 \\ -2 & 0 & 5 \end{bmatrix} \qquad C = \begin{bmatrix} -5 \\ 4 \\ -7 \end{bmatrix}$$

$$D = \begin{bmatrix} 6 & 0 & 3 \\ -5 & 4 & 2 \\ 1 & 1 & 0 \end{bmatrix} \qquad E = \begin{bmatrix} 5 & -1 & 2 \end{bmatrix} \qquad F = \begin{bmatrix} 2 & -1 \\ 6 & 9 \end{bmatrix}$$

3. Fill in the blanks: $\begin{bmatrix} -5 & 1 & 3 \\ 2 & 4 & 0 \\ 1 & -1 & -6 \end{bmatrix}\begin{bmatrix} 6 & 0 & -1 \\ -5 & 7 & 2 \\ -4 & 1 & 3 \end{bmatrix} = \begin{bmatrix} * & * & * \\ \_\_\_ & * & * \\ * & \_\_\_ & * \end{bmatrix}$

4. Suppose that $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & & \\ a_{31} & & \ddots & \\ \vdots & & & \end{bmatrix}$ is a $5 \times 5$ matrix. Write an expression for the third row, second column entry of $A^2$.

5. In the previous section you found that the matrix $P = \begin{bmatrix} \frac{16}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{9}{25} \end{bmatrix}$ projects every vector in $\mathbb{R}^2$ onto the line through the origin and the point $(4, 3)$. If we wish to calculate $P^2$ we can apply Theorem 3.4.2 to factor $\frac{1}{25}$ out of each copy of the matrix, multiply the resulting matrices with integer entries, then multiply the product of the two $\frac{1}{25}$ scalars back in at the end. Do this using your calculator for multiplying numbers, but not the actual matrix multiplication, and reduce the entries when you are done. The result may surprise you a bit!

6. Let $A = \begin{bmatrix} -5 & 1 \\ 0 & 4 \\ 2 & -3 \end{bmatrix}$.

(a) Give $A^T$, the transpose of $A$.

(b) Find $A^TA$ and $AA^T$. Are they the same (equal)?

(c) Your answers to (b) are special in two ways. What are they? (What I'm looking for here is two of the special types of matrices described in Section 3.1.)

## 3.5 Inverse Matrices

**Performance Criteria:**

3. (k) Determine whether two matrices are inverses without finding the inverse of either.
   (l) Find the inverse of a 2×2 matrix using the formula. Find the inverse of a matrix using the Gauss-Jordan method. Describe the Gauss-Jordan method for finding the inverse of a matrix.
   (m) Solve a system of equations using an inverse matrix. Describe how to use an inverse matrix to solve a system of equations.

### Inverse Matrices

Let's begin with an example!

◇ **Example 3.5(a):** Find $AC$ and $CA$ for $A=\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}$ and $C=\begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix}$.

**Solution:**

$$AC=\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad CA=\begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We see that $AC=CA=I_2$!

Now let's remember that the identity matrix is like the number one for multiplication of numbers. Note that, for example, $\frac{1}{5}\cdot 5=5\cdot\frac{1}{5}=1$. This is exactly what we are seeing in the above example. We say the numbers $5$ and $\frac{1}{5}$ are **multiplicative inverses**, and we say that the matrices $A$ and $C$ above are inverses of each other.

<u>Definition 3.5.1</u> **Inverse Matrices**

Suppose that for matrices $A$ and $B$ we have $AB=BA=I$, with the size of the identity being the same in both cases. Then we say that $A$ and $B$ are **inverse matrices**.

Notationally we write $B=A^{-1}$ or $A=B^{-1}$, and we will say that $A$ and $B$ are **invertible**. Note that in order for us to be able to do both multiplications $AB$ and $BA$, both matrices must be square and of the same dimensions. It also turns out that that to test two square matrices to see if they are inverses we only need to multiply them in one order:

<u>Theorem 3.5.2</u> **Test for Inverse Matrices**

To test two *square* matrices $A$ and $B$ to see if they are inverses, compute $AB$. If it is the identity, then the matrices are inverses.

Here are a few notes about inverse matrices:

- Not every square matrix has an inverse, but "many" do. If a matrix does have an inverse, it is said to be **invertible**.
- The inverse of a matrix is unique, meaning there is only one.
- Matrix multiplication *IS* commutative for inverse matrices.

Two questions that should be occurring to you now are

1) How do we know whether a particular matrix has an inverse?

2) If a matrix does have an inverse, how do we find it?

There are a number of ways to answer the first question; here is one:

<u>THEOREM 3.5.3</u> **Test for Invertibility of a Matrix**

A square matrix $A$ is invertible if, and only if, *rref* $(A) = I$.

**Finding Inverse Matrices**

Here is the answer to the second question above in the case of a $2 \times 2$ matrix:

<u>THEOREM 3.5.4</u> **Inverse of a $2 \times 2$ Matrix**

The inverse of a $2 \times 2$ matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $A^{-1} = \dfrac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

◇ **Example 3.5(b):** Find the inverse of $A = \begin{bmatrix} -2 & 7 \\ 1 & -5 \end{bmatrix}$.

**Solution:**

$$A^{-1} = \frac{1}{(-2)(-5) - (1)(7)}\begin{bmatrix} -5 & -7 \\ -1 & -2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -5 & -7 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} -\frac{5}{3} & -\frac{7}{3} \\ -\frac{1}{3} & -\frac{2}{3} \end{bmatrix}$$

Before showing how to find the inverse of a larger matrix we need to go over the idea of **augmenting** a matrix with a vector or another matrix. To augment a matrix $A$ with a matrix $B$, both matrices must have the same number of rows. A new matrix, denoted $[\,A\,|\,B\,]$ is formed as follows: the first row of $[A|B]$ is the first row of $A$ followed by the first row of $B$, and every other row in $[A|B]$ is formed the same way.

◇ **Example 3.5(c):** Let $A = \begin{bmatrix} -5 & 1 & -2 \\ 7 & 0 & 4 \\ 2 & -3 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 9 & 1 \\ -1 & 8 \\ -6 & -3 \end{bmatrix}$ and $\vec{\mathbf{x}} = \begin{bmatrix} -7 \\ 10 \\ 4 \end{bmatrix}$.

Give the augmented matrices $[\,A \mid \vec{\mathbf{x}}\,]$ and $[A \,|\, B]$.

**Solution:** $[\,A \mid \vec{\mathbf{x}}\,] = \begin{bmatrix} -5 & 1 & -2 & -7 \\ 7 & 0 & 4 & 10 \\ 2 & -3 & 6 & 4 \end{bmatrix}$, $\quad [A \,|\, B] = \begin{bmatrix} -5 & 1 & -2 & 9 & 1 \\ 7 & 0 & 4 & -1 & 8 \\ 2 & -3 & 6 & 6 & -3 \end{bmatrix}$

---

**Gauss-Jordan Method for Finding Inverse Matrices**

Let $A$ be an $n \times n$ invertible matrix and $I_n$ be the $n \times n$ identity matrix. Form the augmented matrix $[A \mid I_n]$ and find *rref* $([A \mid I_n]) = [I_n \mid B]$. (The result of row-reduction will have this form.) Then $B = A^{-1}$.

◇ **Example 3.5(d):** Find the inverse of $A = \begin{bmatrix} 2 & 3 & 0 \\ 1 & -2 & 1 \\ 2 & 0 & 1 \end{bmatrix}$, if it exists.

**Solution:** We begin by augmenting with the $3 \times 3$ identity: $[A \mid I_3] = \begin{bmatrix} 2 & 3 & 0 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$.

Row reducing then gives $\begin{bmatrix} 1 & 0 & 0 & 2 & 3 & -3 \\ 0 & 1 & 0 & -1 & -2 & 2 \\ 0 & 0 & 1 & -4 & -6 & 7 \end{bmatrix}$, so $A^{-1} = \begin{bmatrix} 2 & 3 & -3 \\ -1 & -2 & 2 \\ -4 & -6 & 7 \end{bmatrix}$.

---

The above example is a bit unusual; the inverse of a randomly generated matrix will usually contain fractions.

◇ **Example 3.5(e):** Find the inverse of $B = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 2 & -1 \\ 0 & 2 & -2 \end{bmatrix}$, if it exists.

**Solution:** We compute

$$[B \mid I_n] = \begin{bmatrix} 1 & -1 & 2 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 2 & -2 & 0 & 0 & 1 \end{bmatrix} \overset{rref}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & -1 & \frac{3}{2} \end{bmatrix}.$$

Because the left side of the reduced matrix is not the identity, the matrix $B$ is not invertible.

---

◇ **Example 3.5(f):** Find a matrix $B$ such that $AB = C$, where $A = \begin{bmatrix} -3 & 1 \\ 2 & -1 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & -3 \\ -2 & 3 \end{bmatrix}$.

**Solution:** Note that if we multiply both sides of $AB = C$ on the left by $A^{-1}$ we get $A^{-1}AB = A^{-1}C$. But $A^{-1}AB = IB = B$, so we have

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

---

**Inverse Matrices and Systems of Equations**

Let's consider a simple algebraic equation of the form $ax = b$, where $a$ and $b$ are just constants. If we multiply both sides on the left by $\frac{1}{a}$, the multiplicative inverse of $a$, we get $x = \frac{1}{a} \cdot b$. for example,

$$\begin{aligned} 3x &= 5 \\ \tfrac{1}{3}(3x) &= \tfrac{1}{3} \cdot 5 \\ \left(\tfrac{1}{3} \cdot 3\right) x &= \tfrac{5}{3} \\ 1x &= \tfrac{5}{3} \\ x &= \tfrac{5}{3} \end{aligned}$$

The following shows how an inverse matrix can be used to solve a system of equations by exactly the same idea:

$$\begin{aligned} A\vec{\mathbf{x}} &= \vec{\mathbf{b}} \\ A^{-1}(A\vec{\mathbf{x}}) &= A^{-1}\vec{\mathbf{b}} \\ (A^{-1}A)\vec{\mathbf{x}} &= A^{-1}\vec{\mathbf{b}} \\ I\vec{\mathbf{x}} &= A^{-1}\vec{\mathbf{b}} \\ \vec{\mathbf{x}} &= A^{-1}\vec{\mathbf{b}} \end{aligned}$$

*Note that this only "works" if $A$ is invertible!* The upshot of all this is that when $A$ is invertible the solution to the system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is given by $\vec{\mathbf{x}} = A^{-1}\vec{\mathbf{b}}$. The above sequence of steps shows the details of why this is. Although this may seem more straightforward than row reduction, it is more costly in terms of computer time than row reduction or $LU$-factorization and can lead to poor results. Therefore it is not used in practice.

◇ **Example 3.5(g):** Solve the system of equations $\begin{aligned} 5x_1 + 4x_2 &= 25 \\ -2x_1 - 2x_2 &= -12 \end{aligned}$ using an inverse matrix, showing all steps given above.

**Solution:** The matrix form of the system is $\begin{bmatrix} 5 & 4 \\ -2 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 25 \\ -12 \end{bmatrix}$, and $A^{-1} = -\frac{1}{2}\begin{bmatrix} -2 & -4 \\ 2 & 5 \end{bmatrix}$. $A^{-1}$ can now be used to solve the system:

$$\begin{bmatrix} 5 & 4 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 25 \\ -12 \end{bmatrix}$$

$$-\frac{1}{2}\begin{bmatrix} -2 & -4 \\ 2 & 5 \end{bmatrix} \left( \begin{bmatrix} 5 & 4 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = -\frac{1}{2}\begin{bmatrix} -2 & -4 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 25 \\ -12 \end{bmatrix}$$

$$\left( -\frac{1}{2}\begin{bmatrix} -2 & -4 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ -2 & -2 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -\frac{1}{2}\begin{bmatrix} -2 \\ -10 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

The solution to the system is $(1, 5)$.

---

**Section 3.5 Exercises** To Solutions

1. Determine whether $A = \begin{bmatrix} 2 & 5 \\ 3 & 8 \end{bmatrix}$ and $C = \begin{bmatrix} 8 & -4 \\ -3 & 2 \end{bmatrix}$ are inverses, *without actually finding the inverse of either*. Show clearly how you do this.

2. Consider the matrices $A = \begin{bmatrix} 3 & 0 & -1 \\ -1 & -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 5 & 3 \end{bmatrix}$. Find $AB$, then give two reasons why $A$ and $B$ are not inverses.

3. Consider the matrix $\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$.

   (a) Apply row reduction ( *"by hand"*) to $[\,A \mid I_2\,]$ until you obtain $[\,I_2 \mid B\,]$. That is, find the *reduced* row-echelon form of $[\,A \mid I_2\,]$.

   (b) Find $AB$ and $BA$.

   (c) What does this illustrate?

4. Assume that you have a system of equations $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ for some invertible matrix $A$. Show how the inverse matrix is used to solve the system, showing all steps in the process clearly. Check your answer against what is shown at the bottom of page 111.

5. Consider the system of equations $\begin{aligned} 2x_1 - 3x_2 &= 4 \\ 4x_1 + 5x_2 &= 3 \end{aligned}$.

(a) Write the system in matrix times a vector form $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$.

(b) Apply the formula in Theorem 3.5.4 to obtain the inverse matrix $A^{-1}$. **Show a step or two in how you do this.**

(c) **Demonstrate** that your answer to (b) really is the inverse of $A$.

(d) Use the inverse matrix to solve the system. **Show ALL steps outlined in Example 3.5(g), and give your answer in exact form.**

(e) Apply row reduction (**"by hand"**) to $[\,A \mid I_2\,]$ until you obtain $[\,I_2 \mid B\,]$. That is, find the *reduced* row-echelon form of $[\,A \mid I_2\,]$. What do you notice about $B$?

6. Consider the system of equations $\begin{aligned} 5x + 7y &= -1 \\ 2x + 3y &= 4 \end{aligned}$

(a) Write the system in $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ form.

(b) Use Theorem 3.5.4 to find $A^{-1}$.

(c) Give the matrix that is to be row reduced to find $A^{-1}$ by the Gauss-Jordan method. Then give the reduced row-echelon form obtained using your calculator.

(d) Repeat *EVERY* step of the process for solving $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ using the inverse matrix.

## 3.6 Determinants and Systems of Equations

**Performance Criterion:**

3. (n) Find the determinant of a 2×2 or $3 \times 3$ matrix by hand. Use a calculator to find the determinant of an $n \times n$ matrix.
   (o) Use the determinant to determine whether a system of equations has a unique solution.
   (p) Determine whether a homogeneous system has more than one solution.
   (q) Use Cramer's rule to solve a system of equations.

Associated with every *square* matrix is a scalar that is called the **determinant** of the matrix, and determinants have numerous conceptual and practical uses. For a square matrix $A$, the determinant is denoted by $\det(A)$. This notation implies that the determinant is a function that takes a matrix returns a scalar. The determinant of the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ written as $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ or $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

There is a simple formula for finding the determinant of a $2 \times 2$ matrix:

DEFINITION 3.6.1: **Determinant of a $2 \times 2$ Matrix**

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

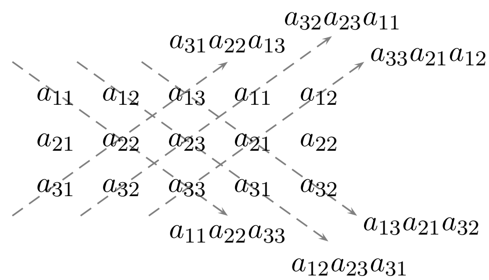
◇ **Example 3.6(a):** Find the determinant of $A = \begin{bmatrix} 5 & 4 \\ -2 & -2 \end{bmatrix}$

$$\det(A) = (5)(-2) - (-2)(4) = -10 + 8 = -2$$

There is a fairly involved method of breaking the determinant of a larger matrix down to where it is a linear combination of determinants of $2 \times 2$ matrices, but we will not go into that here. It is called the **cofactor expansion** of the determinant, and can be found in any other linear algebra book, or online. Of course your calculator will find determinants of matrices whose entries are numbers, as will online matrix calculators and various software like *MATLAB*.

Later we will need to be able to find determinants of matrices containing an unknown parameter, and it will be necessary to find determinants of $3 \times 3$ matrices. For that reason, we now show a relatively simple method for finding the determinant of a $3 \times 3$ matrix. (This will not look simple here, but it is once you are familiar with it.) *This method only works for $3 \times 3$ matrices.* To perform this method we begin by augmenting the matrix with its own first two columns, as shown below.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \Longrightarrow \quad \begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix}$$

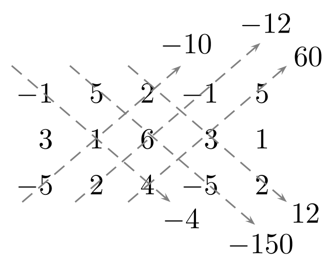



We get the determinant by adding up each of the results of the downward multiplications and then subtracting each of the results of the upward multiplications. This is shown below.

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

◇ **Example 3.6(b):** Find the determinant of $A = \begin{bmatrix} -1 & 5 & 2 \\ 3 & 1 & 6 \\ -5 & 2 & 4 \end{bmatrix}$. Video Example

$$\det \begin{bmatrix} -1 & 5 & 2 \\ 3 & 1 & 6 \\ -5 & 2 & 4 \end{bmatrix} \Longrightarrow$$




$$\det(A) = (-4) + (-150) + 12 - (-10) - (-12) - 60 = -4 - 150 + 12 + 10 + 12 - 60 = -180$$

In the future we will need to compute determinants like the following.

◇ **Example 3.6(c):** Find the determinant of $B = \begin{bmatrix} 1-\lambda & 0 & 3 \\ 1 & -1-\lambda & 2 \\ -1 & 1 & -2-\lambda \end{bmatrix}$.

$$\begin{aligned} \det(B) &= (1-\lambda)(-1-\lambda)(-2-\lambda) + (0)(2)(-1) + (3)(1)(1) \\ &\qquad - (-1)(-1-\lambda)(3) - (1)(2)(1-\lambda) - (-2-\lambda)(1)(0) \\ &= (-1+\lambda^2)(-2-\lambda) + 3 - 3 - 3\lambda - 2 + 2\lambda \\ &= 2 + \lambda - 2\lambda^2 - \lambda^3 - \lambda - 2 \\ &= -\lambda^3 - 2\lambda^2 \end{aligned}$$

Here is why we care about determinants right now:

THEOREM 3.6.2: **Determinants and Invertibility, Systems**

Let $A$ be a square matrix.

(a) $A$ is invertible if, and only if, $\det(A) \neq 0$.

(b) The system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a unique solution if, and only if, $A$ is invertible.

(c) If $A$ is not invertible, the system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ will have either no solution or infinitely many solutions.

Recall that when things are "nice" the system $A\,\vec{\mathbf{x}}=\vec{\mathbf{b}}$ can be solved as follows:

$$\begin{aligned} A\vec{\mathbf{x}} &= \vec{\mathbf{b}} \\ A^{-1}(A\vec{\mathbf{x}}) &= A^{-1}\vec{\mathbf{b}} \\ (A^{-1}A)\vec{\mathbf{x}} &= A^{-1}\vec{\mathbf{b}} \\ I\vec{\mathbf{x}} &= A^{-1}\vec{\mathbf{b}} \\ \vec{\mathbf{x}} &= A^{-1}\vec{\mathbf{b}} \end{aligned}$$

In this case the system will have the unique solution $\vec{\mathbf{x}} = A^{-1}\vec{\mathbf{b}}$. (When we say unique, we mean only one.) *If $A$ is not invertible, the above process cannot be carried out, and the system will not have a single unique solution*. In that case there will either be no solution or infinitely many solutions.

We previously discussed the fact that the above computation is analogous to the following one involving simple numbers and an unknown number $x$:

$$\begin{aligned} 3x &= 5 \\ \tfrac{1}{3}(3x) &= \tfrac{1}{3}\cdot 5 \\ \left(\tfrac{1}{3}\cdot 3\right)x &= \tfrac{5}{3} \\ 1x &= \tfrac{5}{3} \\ x &= \tfrac{5}{3} \end{aligned}$$

Now let's consider the following two equations, of the same form $ax = b$ but for which $a = 0$:

$$0x = 5 \qquad\qquad 0x = 0$$

We first recognize that we can't do as before and multiply both sides of each by $\frac{1}{0}$, since that is undefined. The first equation has no solution, since there is no number $x$ that can be multiplied by zero and result in five! In the second case, every number is a solution, so the system has infinitely many solutions. *These equations are analogous to $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ when $\det(A) = 0$. The one difference is that $A\,\vec{\mathbf{x}}=\vec{\mathbf{b}}$ can have infinitely many solutions even when $\vec{\mathbf{b}}$ is NOT the zero vector.*

Homogenous systems are important and will come up in a couple places in the future, but there is not a whole lot that can be said about them! A **homogeneous system** is one of the form $A\,\vec{\mathbf{x}}= \mathbf{0}$. With a tiny bit of thought this should be clear: *Every homogenous system has at least one solution - the zero vector!* Given the Theorem 3.6.2, if $A$ is invertible (so $\det(A) \neq 0$), that is the only solution. If $A$ is not invertible there will be infinitely many solutions, the zero vector being just one of them.

**Cramer's Rule**

Cramer's rule is a method for finding solutions to systems of equations. It is not generally used for solving large systems with numerical solutions, but it is used sometimes for solving smaller systems containing an unknown parameter.

Let's consider the system of equations $\begin{array}{rcl} 5x+3y &=& 1 \\ 4x+2y &=& 2 \end{array}$. To solve for $x$ we can multiply the first equation by $2$ and the second equation by $-3$ to obtain

$$\begin{array}{rcr} 10x+6y &=& 2 \\ -12x-6y &=& -6 \end{array}.$$

We then add the two equations to obtain the equation $-2x=-4$, so $x=2$. Note that we could instead multiply the first equation by $2$, the second by $3$, and then *subtract* the two equations to get

$$\begin{array}{rcl} 5x+3y=1 & \Longrightarrow & 10x+6y=2 \\ 4x+2y=2 & \Longrightarrow & \underline{12x+6y=6} \\ & & \quad -2x=-4, \end{array}$$

solving the resulting equation to get $x=2$. Using this process on a general system of two equations we get

$$\begin{array}{rcl} ax+by=e & \Longrightarrow & adx+bdy=ed \\ cx+dy=f & \Longrightarrow & \underline{bcx+bdy=bf} \\ & & adx-bcx=ed-bf \end{array}$$

$$\Longrightarrow \quad (ad-bc)x=ed-bf \quad \Longrightarrow \quad x=\frac{ed-bf}{ad-bc} \tag{1}$$

We can also multiply the top equation by $c$ and the bottom equation by $a$ and then subtract the top equation from the bottom one to get

$$\begin{array}{rcl} ax+by=e & \Longrightarrow & acx+bcy=ce \\ cx+dy=f & \Longrightarrow & \underline{acx+ady=af} \\ & & ady-bcy=af-ce \end{array}$$

$$\Longrightarrow \quad (ad-bc)y=af-ce \quad \Longrightarrow \quad y=\frac{af-ce}{ad-bc} \tag{2}$$

Note the matrix form of the system of equations: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=\begin{bmatrix} e \\ f \end{bmatrix}$. If we look carefully at the expressions for $x$ and $y$ that were obtained in (1) and (2) above, we see that they both have the same denominator $ad-bc$, the determinant of the coefficient matrix! How about the numerators of the expressions for $x$ and $y$? We can see that they are the determinants

$$\begin{vmatrix} e & b \\ f & d \end{vmatrix} \qquad \text{and} \qquad \begin{vmatrix} a & e \\ c & f \end{vmatrix},$$

respectively. This leads us to

**Theorem 3.6.3: Cramer's Rule**

The system of two equations in two unknowns with standard and matrix forms

$$\begin{aligned} ax + by &= e \\ cx + dy &= f \end{aligned} \qquad \text{and} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} e \\ f \end{bmatrix}.$$

has solution given by

$$x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \qquad \text{and} \qquad y = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}.$$

*when the determinant of the coefficient matrix is not zero.*

We reiterate: the denominators of both of these fractions are the determinant of the coefficient matrix. The numerator for finding $x$ is the determinant of the matrix obtained when the coefficient matrix has its first column (the coefficients of $x$) replaced with the numbers to the right of the equal signs. Let's see an example:

◇ **Example 3.6(d):** Use Cramer's rule to solve the system of equations $\begin{aligned} 5x + 3y &= 1 \\ 4x + 2y &= 2 \end{aligned}$.

**Solution:** Cramer's Rule gives us

$$x = \frac{\begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix}}{\begin{vmatrix} 5 & 3 \\ 4 & 2 \end{vmatrix}} = \frac{1 \cdot 2 - 2 \cdot 3}{5 \cdot 2 - 4 \cdot 3} = \frac{-4}{-2} = 2 \qquad y = \frac{\begin{vmatrix} 5 & 1 \\ 4 & 1 \end{vmatrix}}{\begin{vmatrix} 5 & 3 \\ 4 & 2 \end{vmatrix}} = \frac{5 \cdot 2 - 4 \cdot 1}{5 \cdot 2 - 4 \cdot 3} = \frac{6}{-2} = -3$$

---

We conclude by noting that the use of Cramer's rule is not restricted to systems of two equations in two unknowns:

**Theorem 3.6.4: Cramer's Rule for More Unknowns**

Any system of $n$ equations in $n$ unknowns whose coefficient matrix has nonzero determinant can be solved in the same manner as above. That is, the value of each unknown is obtained by replacing the column of the coefficient matrix corresponding to that unknown with the right hand side vector, then dividing the determinant of the resulting matrix by the determinant of the coefficient matrix.

**Section 3.6 Exercises** To Solutions

1. Find the determinant of each matrix *by hand*, giving your answer in fraction form.

(a) $A = \begin{bmatrix} 3 & 5 \\ 1 & 3 \end{bmatrix}$ (b) $B = \begin{bmatrix} 2 & 2 \\ -1 & 4 \end{bmatrix}$ (c) $C = \begin{bmatrix} 2 & -1 \\ 2 & 3 \end{bmatrix}$

(d) $A = \begin{bmatrix} -1 & -2 \\ 3 & 1 \end{bmatrix}$ (e) $B = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$ (f) $C = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$

(e) $A = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$ (f) $B = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix}$

2. (a) $\begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix}$ (b) $\begin{bmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{bmatrix}$

(c) $\begin{bmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{bmatrix}$ (d) $\begin{bmatrix} 3-\lambda & 2 & 4 \\ 2 & 0-\lambda & 2 \\ 4 & -2 & 3-\lambda \end{bmatrix}$

3. Explain/show how to use the determinant to determine whether
$$\begin{aligned} x + 3y - 3z &= -5 \\ 2x - \ y + \ z &= -3 \\ -6x + 3y - 3z &= 4 \end{aligned}$$
has a unique solution. **You may use your calculator for finding determinants - be sure to conclude by saying whether or not this particular system has a solution!**

4. Suppose that you hope to solve a system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ of $n$ equations in $n$ unknowns.

   (a) If the determinant of $A$ is zero, what does it tell you about the nature of the solution? (By "the nature of the solution" I mean no solution, a unique solution or infinitely many solutions.)

   (b) If the determinant of $A$ is *NOT* zero, what does it tell you about the nature of the solution?

5. Suppose that you hope to solve a system $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$ of $n$ equations in $n$ unknowns.

   (a) If the determinant of $A$ is zero, what does it tell you about the nature of the solution? (By "the nature of the solution" I mean no solution, a unique solution or infinitely many solutions.)

   (b) If the determinant of $A$ is *NOT* zero, what does it tell you about the nature of the solution?

6. Use Cramer's Rule to solve each of the following systems of equations. Check your answers by solving with *rref*.

(a) $$\begin{aligned} -2x + 5y &= 13 \\ 4x + 7y &= 25 \end{aligned}$$

(b) $$\begin{aligned} 1x - 3y &= -17 \\ -2x + 5y &= 29 \end{aligned}$$

(c) $$\begin{aligned} 8x - 3y &= 32 \\ 4x - 5y &= 16 \end{aligned}$$

(d) $$\begin{aligned} 1x + 7y &= 48 \\ 4y &= 28 \end{aligned}$$

## 3.7 Applications: Transformation Matrices, Graph Theory

**Performance Criteria:**

3. (r) Give the geometric or algebraic representations of the inverse or square of a rotation. Demonstrate that the geometric and algebraic versions are the same

   (s) Give the incidence matrix of a graph or digraph. Given the incidence matrix of a graph or digraph, identify the vertices and edges using correct notation, and draw the graph.

   (t) Determine the number of $k$-paths from one vertex of a graph to another. Solve problems using incidence matrices.

### Rotation, Projection and Reflection Matrices

In the Section 3.3 Exercises you encountered matrices that rotated every vector in $\mathbb{R}^2$ thirty degrees counterclockwise, projected every vector in $\mathbb{R}^2$ onto the line $y = \frac{3}{4}x$, and reflected every vector in $\mathbb{R}^2$ across the line $y = \frac{3}{4}x$. Here are the general formulas for rotation, projection and reflection matrices in $\mathbb{R}^2$:

**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 $\theta$.

**Projection Matrix in $\mathbb{R}^2$**

For the matrix $B = \begin{bmatrix} \dfrac{a^2}{a^2+b^2} & \dfrac{ab}{a^2+b^2} \\ \dfrac{ab}{a^2+b^2} & \dfrac{b^2}{a^2+b^2} \end{bmatrix}$ and any position vector $\vec{\mathbf{x}}$ in $\mathbb{R}^2$, the

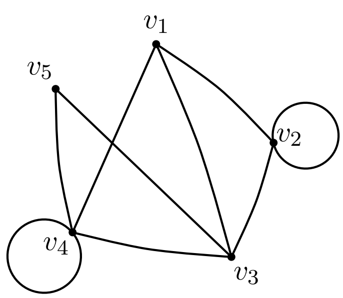
product $B\,\vec{\mathbf{x}}$ is the vector resulting when $\vec{\mathbf{x}}$ is projected onto the line containing the origin and the point $(a, b)$.

**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 $(a,b)$.

## Graphs and Digraphs

A **graph** is a set of dots, called **vertices**, connected by segments of lines or curves, called **edges**. An example is shown to the left below. We will usually label each of the vertices with a subscripted $v$, as shown. Note that a vertex can be connected to itself, as shown by the circles at $v_2$ and $v_4$. We can then create a matrix, called an **incidence matrix** to show which pairs of vertices are connected (and which are not). The $(i,j)$ and $(j,i)$ entries of the matrix are one if $v_i$ and $v_j$ are connected by a single edge and a zero if they are not. If $i=j$ the entry is a one if that vertex is connected to itself by an edge, and zero if it is not. You should be able to see this by comparing the graph and and corresponding incidence matrix below. *Note that the incidence matrix is symmetric; that is the case for all incidence matrices of graphs.*




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

Even though there is no edge from vertex one to vertex five for the graph shown, we can get from vertex one to vertex five via vertex three or vertex four. We call such a "route" a **path**, and we denote the paths by the sequence of vertices, like $v_1v_3v_5$ or $v_1v_4v_5$. These paths, in particular, are called **2-paths**, since they consist of two edges. There are other paths from vertex one to vertex five, like the 3-path $v_1v_2v_3v_5$ and the 4-path $v_1v_2v_2v_3v_5$.

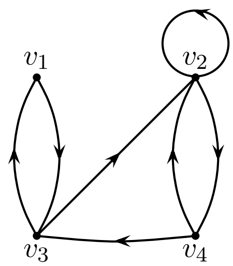
◇ **Example 3.7(a):** Give two more 3-paths from vertex $v_1$ to vertex $v_5$.

**Solution:** $v_1v_3v_4v_5$ is a fairly obvious 3-path from $v_1$ to vertex $v_5$. Less obvious is $v_1v_4v_4v_5$, which travels the edge from $v_4$ to itself.

---

We should note that when forming a path we are allowed to travel the same edge multiple times, including reversing the direction. Thus, for example, $v_1v_4v_5v_4v_5$ is a 4-path from $v_1$ to $v_5$. It is often desired to find all $n$-paths from one vertex to another, and it can be difficult to determine when all of them have been found. You will find in the exercises a clever way to determine how many $n$-paths there are from one vertex to another, so we know how many we are looking for. This doesn't necessarily make it any easier to find them, but we can know whether we have them all or not.

In some cases we want the edges of a graph to be "one-way." We indicate this by placing an arrow on each edge, indicating the direction it goes. *We will not put two arrowheads on one edge; if we can travel both ways between two vertices, we will show that by drawing TWO edges between them.* Such a graph is called a **directed graph**, or **digraph** for short. Below is a digraph and its incidence matrix. The $(i, j)$ entry of the incidence matrix is one only if there is a directed edge from $v_i$ to $v_j$. Of course *the incidence matrix for a digraph need not be symmetric*, since there may be an edge going one way between two vertices but not the other way. Digraphs have incidence matrices as well. Below is a digraph and its incidence matrix.




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

Both the graph and the digraph above are what we call **connected graphs**, meaning that every two vertices are connected by some path (but not necessarily an edge). A graph that is not connected will appear to be two or more separate graphs. All graphs that we will consider will be connected; we will leave further discussion/investigation of graphs and incidence matrices to the exercises.

## Section 3.7 Exercises

To Solutions

1. Here we investigate projections.

   (a) Sketch a set of coordinate axes in $\mathbb{R}^2$, with an additional line passing through the origin. Let $P$ be the matrix that projects all vectors onto the line.

   (b) Sketch a position vector $\vec{\mathbf{u}}$ that is on the line. What can you say about the vector $P\vec{\mathbf{u}}$ in this case?

   (c) Sketch a position vector $\vec{\mathbf{v}}$ that is perpendicular to the line. What is $P\vec{\mathbf{v}}$?

   (d) Sketch a third vector $\vec{\mathbf{w}}$ that is not on the line or perpendicular to it. Draw the vector $P\vec{\mathbf{w}}$. Now suppose that we applied $P$ to $P\vec{\mathbf{w}}$ to get $P(P\vec{\mathbf{w}}) = P^2\vec{\mathbf{w}}$. How does that result compare with $P\vec{\mathbf{w}}$? What does this tell us about $P^2$?

   (e) Discuss $P^3$, $P^4$, ..., $P^n$.

   (f) Suppose that we know $P\vec{\mathbf{x}}$ for some unknown vector $\vec{\mathbf{x}}$. Can we determine $\vec{\mathbf{x}}$? What does this tell us about $P^{-1}$?

2. And now we investigate reflections.

   (a) Sketch another set of coordinate axes in $\mathbb{R}^2$, with an additional line passing through the origin. Let $C$ be the matrix that reflects all vectors across the line.

   (b) Sketch a position vector $\vec{\mathbf{u}}$ that is on the line. What can you say about the vector $C\vec{\mathbf{u}}$ in this case?

(c) Sketch a position vector $\vec{\mathbf{v}}$ that is perpendicular to the line. What is $C\vec{\mathbf{v}}$?

(d) Sketch a third vector $\vec{\mathbf{w}}$ that is not on the line or perpendicular to it. Draw the vector $C\vec{\mathbf{w}}$. Now suppose that we applied $C$ to $C\vec{\mathbf{w}}$ to get $C(C\vec{\mathbf{w}}) = C^2\vec{\mathbf{w}}$. What does this tell us about $C^2$?

(e) Discuss $C^3$, $C^4$, ..., $C^n$.

(f) Suppose that we know $C\vec{\mathbf{x}}$ for some unknown vector $\vec{\mathbf{x}}$. Can we determine $\vec{\mathbf{x}}$? What does this tell us about $C^{-1}$?

**Some Trigonometric Identities**

$$\sin^2\theta + \cos^2\theta = 1 \qquad \cos(-\theta) = \cos\theta \qquad \sin(-\theta) = -\sin\theta$$

$$\sin(2\theta) = 2\sin\theta\cos\theta \qquad \cos(2\theta) = \cos^2\theta - \sin^2\theta$$

3. Consider the general rotation matrix $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

   (a) Suppose that we were to apply $A$ to a vector $\vec{\mathbf{x}}$, then apply $A$ again, to the result. Thinking only geometrically (don't do any calculations), give a single matrix $B$ that should have the same effect.

   (b) Find the matrix $A^2$ algebraically, by multiplying $A$ by itself.

   (c) Use some of the trigonometric facts above to continue your calculations from part (b) until you arrive at matrix $B$. This of course shows that that $A^2 = B$.

4. Consider again the general rotation matrix $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

   (a) Give a matrix $C$ that should "undo" what $A$ does. Do this thinking only geometrically.

   (b) Find the matrix $A^{-1}$ algebraically, using the formula for the inverse of a $2\times 2$ matrix..

   (c) Use some of the trigonometric facts above to show that $C = A^{-1}$. Do this by starting with $C$, then modifying it a step at a time to get to $A^{-1}$.

   (d) Give the transpose matrix $A^T$. It should look familiar - tell how.

5. Let $R_\theta$ be the matrix that rotates all vectors counter-clockwise by the angle $\theta$.

   (a) $R_\theta^2$ is equal to $R_\phi$ for what angle $\phi$, in terms of $\theta$?

   (b) $R_\theta^{-1}$ is equal to $R_\phi$ for what angle $\phi$, in terms of $\theta$?

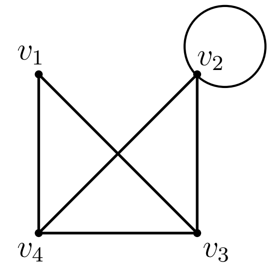
   (c) Give an angle $\theta$ for which $R_\theta^3 = I$.

   (d) Give an angle $\theta$ for which $R_\theta^{-1} = R_\theta$. That is, what rotation is its own inverse?

6. Use the graph to the right for the following.

   (a) Give all 2-paths from $v_3$ to $v_4$.

   (b) Give all three paths from $v_1$ to $v_4$. Don't forget that you can follow the same edge more than once, including in opposite directions.

   (c) Give the incidence matrix $A$ for the graph, and give the additional matrices $A^2$ and $A^3$.

   (d) Look at the $(3,4)$ and $(4,3)$ entries of $A^2$. How do they relate to the number of 2-paths from $v_3$ to $v_4$?

   (e) Look at the $(1,4)$ and $(4,1)$ entries of $A^3$. How do they relate to the number of 3-paths from $v_1$ to $v_4$? (You won't see the connection here if you didn't find all of the 3-paths from $v_1$ to $v_4$. You should have found five.)

   (f) Based on what you observed in parts (d) and (e), how many 3-paths from $v_2$ to $v_3$ are there? Give all 3-paths from $v_2$ to $v_3$. In this case, figure that you can follow the same edge more than once, including in opposite directions, and that you can travel the edge from $v_2$ to itself.

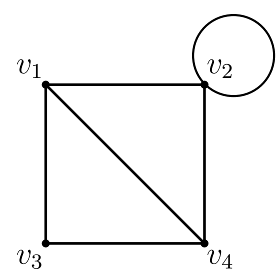
7. (a) Draw the graph with the incidence matrix $A$ shown below, with vertices labeled $v_1$, $v_2$, ...

   (b) Draw the directed graph with the incidence matrix $B$ shown below.

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

8. Use the graph to the right for the following.

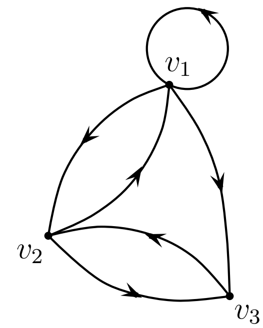
   (a) Give the incidence matrix for the graph; call it $A$. Find and give $A^3$ also.

   (b) How many 3-paths from $v_1$ to $v_4$ do you expect? Give all of them by listing the vertices the path goes through, *in order and including* $v_2$ *and* $v_4$, as done in class. (For example, then, a path from $v_4$ to $v_3$, through $v_2$ would be denoted $v_4v_2v_3$.)

   (c) How many 3-paths are there from $v_4$ to $v_1$? What characteristic of the matrix $A^3$ relates your answer to the number of 3-paths from $v_1$ to $v_4$?

9. Use the directed graph to the right for the following.




(a) Give the incidence matrix; again, call it $A$.

(b) The number of n-paths from vertex $i$ to vertex $j$ is given by the $(i, j)$ entry of $A^n$. How many 4-paths are there from $v_1$ to $v_3$? Show how you get your answer.

(c) Give all 4-paths from $v_2$ to $v_3$.

(d) Give all 3-paths from $v_3$ to $v_2$. Is it the same as the number from $v_2$ to $v_3$?

# 4 Vector Spaces and Subspaces

**Outcome:**

4. Understand subspaces of $\mathbb{R}^n$. Understand bases of vector spaces and subspaces. Find a least squares solution to an inconsistent system of equations.

   **Performance Criteria:**

   (a) Describe the span of a set of vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ as a line or plane containing a given set of points.

   (b) Determine whether a vector **w** is in the span of a set $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, ..., \vec{\mathbf{v}}_k\}$ of vectors. If it is, write $\mathbf{w}$ as a linear combination of $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, ..., \vec{\mathbf{v}}_k$.

   (c) Determine whether a set is closed under an operation. If it is, prove that it is; if it is not, give a counterexample.

   (d) Determine whether a subset of $\mathbb{R}^n$ is a subspace. If so, prove it; if not, give an appropriate counterexample.

   (e) Determine whether a vector is in the column space or null space of a matrix, based only on the definitions of those spaces.

   (f) Find the least-squares approximation to the solution of an inconsistent system of equations. Solve a problem using least-squares approximation.

   (g) Give the least squares error and least squares error vector for a least squares approximation to a solution to a system of equations.

   (h) Determine whether a set $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, ..., \vec{\mathbf{v}}_k$ of vectors is a linearly independent or linearly dependent. If the vectors are linearly dependent, (1) give a non-trivial linear combination of them that equals the zero vector, (2) give any one as a linear combination of the others when possible.

   (i) Determine whether a given set of vectors is a basis for a given subspace. Give a basis and the dimension of a subspace.

   (j) Find the dimensions of, and bases for, the column space and null space of a given matrix.

   (k) Given the dimension of the column space and/or null space of the coefficient matrix for a system of equations, say as much as you can about how many solutions the system has.

   (l) Determine, from given information about the coefficient matrix $A$ and vector $\vec{\mathbf{b}}$ of a system $A\vec{\mathbf{x}}=\vec{\mathbf{b}}$, whether the system has any solutions and, if it does, whether there is more than one solution.

A very important concept in linear algebra is that all vectors of interest in a given situation can be constructed out of a small set of vectors, using linear combinations. That is the key idea that we will explore in this chapter. This will seem to take us farther from some of the more concrete ideas that we have used in applications, but these ideas have huge value in a practical sense as well.

## 4.1 Span of a Set of Vectors

**Performance Criteria:**

4. (a) Describe the span of a set of vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ as a line or plane containing a given set of points.

   (b) Determine whether a vector $\mathbf{w}$ is in the span of a set $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, ..., \vec{\mathbf{v}}_k\}$ of vectors. If it is, write $\mathbf{w}$ as a linear combination of $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, ..., \vec{\mathbf{v}}_k$.

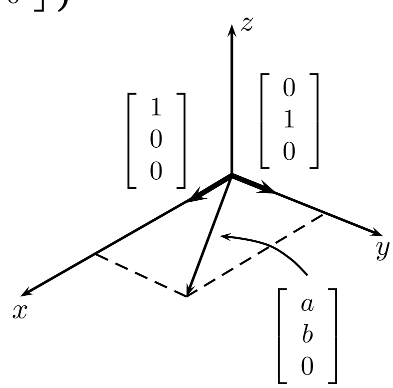
DEFINITION 4.1.1: The **span of a set** $\mathcal{S}$ **of vectors**, denoted $\text{span}(\mathcal{S})$ is the set of all linear combinations of those vectors.

◇ **Example 4.1(a):** Describe the span of the set $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ in $\mathbb{R}^3$.

**Solution:** Note that *ANY* vector with a zero third component can be written as a linear combination of these two vectors:

$$\begin{bmatrix} a \\ b \\ 0 \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

All the vectors with $x_3 = 0$ (or $z = 0$) are the $xy$ plane in $\mathbb{R}^3$, so the span of this set is the $xy$ plane. Geometrically we can see the same thing in the picture to the right.




◇ **Example 4.1(b):** Describe $\text{span}\left( \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \right)$.

**Solution:** By definition, the span of this set is all vectors $\vec{\mathbf{v}}$ of the form

$$\vec{\mathbf{v}} = c_1 \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix},$$

which, because the two vectors are not scalar multiples of each other, we recognize as being a plane through the origin. It should be clear that all vectors created by such a linear combination will have a third component of zero, so the particular plane that is the span of the two vectors is the $xy$-plane. Algebraically we see that any vector $[a, b, 0]$ in the $xy$-plane can be created by

$$\left(\frac{a-3b}{7}\right) \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + \left(\frac{2a+b}{7}\right) \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{a-3b}{7} \\ \frac{-2a+6b}{7} \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{6a+3b}{7} \\ \frac{2a+b}{7} \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{7a}{7} \\ \frac{7b}{7} \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \\ 0 \end{bmatrix}$$

You might wonder how one would determine the scalars $\dfrac{a-3b}{7}$ and $\dfrac{2a+b}{7}$. You will see how this is done in the exercises!

At this point we should make a comment and a couple observations:

- First, some language: we can say that the span of the two vectors in Example 4.1(b) is the $xy$-plane, but we also say that the two vectors span the $xy$-plane. That is, the word span can be either a noun or a verb, depending on how it is used.
- Note that in the two examples above we considered two different sets of two vectors, but in each case the span was the same. This illustrates that *different sets of vectors can have the same span.*
- Consider also the fact that if we were to include in either of the two sets additional vectors that are also in the $xy$-plane, it would not change the span. However, if we were to add another vector not in the $xy$-plane, the span would increase to all of $\mathbb{R}^3$.
- In either of the preceding examples, removing either of the two given vectors would reduce the span to a linear combination of a single vector, which is a line rather than a plane. But in some cases, removing a vector from a set does not change its span.
- The last two bullet items tell us that *adding or removing vectors from a set of vectors may or may not change its span.* This is a somewhat undesirable situation that we will remedy in Section 4.7.
- It may be obvious, but it is worth emphasizing that (in this course) we will consider spans of finite (and usually rather small) sets of vectors, but the span of a finite set $\mathcal{S}$ always contains infinitely many vectors (unless $\mathcal{S}$ consists of only the zero vector).

It is often of interest to know whether a particular vector is in the span of a certain set of vectors. The next examples show how we do this.

◇ **Example 4.1(c):** Is $\vec{\mathbf{v}} = \begin{bmatrix} 3 \\ -2 \\ -4 \\ 1 \end{bmatrix}$ in the span of $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -3 \\ 1 \end{bmatrix} \right\}$?

**Solution:** The question is, "can we find scalars $c_1$, $c_2$ and $c_3$ such that

$$c_1 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} + c_3 \begin{bmatrix} 2 \\ 0 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ -4 \\ 1 \end{bmatrix} \text{?"} \qquad (1)$$

We should recognize this as the linear combination form of the system of equations below and to the left. The augmented matrix for the system row reduces to the matrix below and to the right.

$$\begin{aligned} c_1 + c_2 + 2c_3 &= 3 \\ 2c_1 - c_2 \qquad &= -2 \\ 3c_1 + c_2 - 3c_3 &= -4 \\ 4c_1 - c_2 + \ \ c_3 &= 1 \end{aligned} \qquad\qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This tells us that the system above and to the left has no solution, so there are no scalars $c_1$, $c_2$ and $c_3$ for which equation (1) holds. Thus $\vec{\mathbf{v}}$ is not in the span of $\mathcal{S}$.

◇ **Example 4.1(d):** Is $\vec{\mathbf{v}} = \begin{bmatrix} 19 \\ 10 \\ -1 \end{bmatrix}$ in $\text{span}(\mathcal{S})$, where $\mathcal{S} = \left\{ \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 7 \\ -4 \end{bmatrix} \right\}$?

**Solution:** Here we are trying to find scalars $c_1$, $c_2$ and $c_3$ such that

$$c_1 \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 7 \\ -4 \end{bmatrix} = \begin{bmatrix} 19 \\ 10 \\ -1 \end{bmatrix} \tag{2}$$

This is the linear combination form of the system of equations below and to the left, whose augmented matrix row reduces to the matrix below and to the right.

$$\begin{array}{rcl} 3c_1 - 5c_2 + \phantom{4}c_3 & = & 19 \\ -c_1 \phantom{- 5c_2} + 7c_3 & = & 10 \\ 2c_1 + \phantom{5}c_2 - 4c_3 & = & -1 \end{array} \qquad \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

This tells us that (2) holds for $c_1 = 4$, $c_2 = -1$ and $c_3 = 2$, so $\vec{\mathbf{v}}$ is in $\text{span}(\mathcal{S})$.

Sometimes, with a little thought, no computations are necessary to answer such questions, as the next examples show.

◇ **Example 4.1(e):** Is $\vec{\mathbf{v}} = \begin{bmatrix} -4 \\ 2 \\ 5 \end{bmatrix}$ in the span of $\mathcal{S} = \left\{ \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix} \right\}$?

**Solution:** One can see that any linear combination of the two vectors in $S$ will have zero as its second component:

$$c_1 \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3c_1 \\ 0 \\ 2c_1 \end{bmatrix} + \begin{bmatrix} -5c_2 \\ 0 \\ 1c_2 \end{bmatrix} = \begin{bmatrix} -3c_1 - 5c_2 \\ 0 \\ 2c_1 + c_2 \end{bmatrix}$$

Since the second component of $\vec{\mathbf{v}}$ is not zero, $\vec{\mathbf{v}}$ is not in the span of the set $\mathcal{S}$.

◇ **Example 4.1(f):** Is $\vec{\mathbf{v}} = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$ in $\text{span}\left( \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right)$?

**Solution:** Here we can see that if we multiply the three vectors in $\mathcal{S}$ by 4, 7 and $-1$, respectively, and add them, the result will be $\vec{\mathbf{v}}$:

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

Therefore $\vec{\mathbf{v}}$ is in $\text{span}\left( \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right)$.

Sometimes we will be given an infinite set of vectors, and we'll ask whether a particular finite set of vectors *spans the infinite set*. By this we are asking whether the span of the finite set is the infinite set. For example, we might ask whether the vector $\vec{\mathbf{v}}=\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ spans $\mathbb{R}^2$. Because the span of the single vector $\vec{\mathbf{v}}$ is just a line, $\vec{\mathbf{v}}$ does not span $\mathbb{R}^2$. With the knowledge we have at this point, it can sometimes be difficult to tell whether a finite set of vectors spans a particular infinite set. In Sections 4.6 and 4.7 we will encounter some concepts that will give us a means for making such a judgement a bit easier.

We conclude with a few more observations. With a little thought, the following can be seen to be true. (Assume all vectors are non-zero.)

- The span of a single vector is all scalar multiples of that vector. In $\mathbb{R}^2$ or $\mathbb{R}^3$ the span of a single vector is a line through the origin.
- The span of a set of two non-parallel vectors in $\mathbb{R}^2$ is all of $\mathbb{R}^2$. In $\mathbb{R}^3$ it is a plane through the origin.
- The span of three vectors in $\mathbb{R}^3$ that do not lie in the same plane is all of $\mathbb{R}^3$.

**Section 4.1 Exercises** To Solutions

1. Describe the span of each set of vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ by telling what it is geometrically and, if it is a standard set like one of the coordinate axes or planes, specifically what it is. If it is a line that is not one of the axes, give two points on the line. If it is a plane that is not one of the coordinate planes, give three points on the plane.

   (a) The vector $\begin{bmatrix} 5 \\ 0 \end{bmatrix}$ in $\mathbb{R}^2$.

   (b) The set of vectors $\left\{ \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \right\}$ in $\mathbb{R}^3$.

   (c) The vectors $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$ in $\mathbb{R}^2$.

   (d) The set $\left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$ in $\mathbb{R}^3$.

   (e) The vectors $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$ in $\mathbb{R}^3$.

2. For each of the following, determine whether the vector $\vec{\mathbf{w}}$ is in the span of the set $S$. If it is, write it as a linear combination of the vectors in $S$.

   (a) $\vec{\mathbf{w}} = \begin{bmatrix} -2 \\ 5 \\ 4 \end{bmatrix}, \qquad S = \left\{ \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 7 \end{bmatrix}, \begin{bmatrix} 3 \\ -11 \\ -1 \end{bmatrix} \right\}$

(b) $\vec{\mathbf{w}} = \begin{bmatrix} -5 \\ -23 \\ 12 \\ 8 \end{bmatrix}, \qquad S = \left\{ \begin{bmatrix} 1 \\ -4 \\ -3 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 6 \\ -4 \\ 5 \end{bmatrix} \right\}$

(c) $\vec{\mathbf{w}} = \begin{bmatrix} 8 \\ 38 \\ -14 \\ 11 \end{bmatrix}, \qquad S = \left\{ \begin{bmatrix} 1 \\ -4 \\ -3 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 6 \\ -4 \\ 5 \end{bmatrix} \right\}$

(d) $\vec{\mathbf{w}} = \begin{bmatrix} 3 \\ 7 \\ -4 \end{bmatrix}, \qquad S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

Say we have a set $\mathcal{S}$ of vectors in $\mathbb{R}^n$, and we consider $\text{span}(\mathcal{S})$. Then suppose that we include an additional new vector $\vec{\mathbf{v}}$ in $\mathcal{S}$, to obtain a new set we'll call $\mathcal{S}'$. If $\vec{\mathbf{v}}$ is in $\text{span}(\mathcal{S})$, then $\text{span}(\mathcal{S}') = \text{span}(\mathcal{S})$. If $\vec{\mathbf{v}}$ is not in $\text{span}(\mathcal{S})$, then $\text{span}(\mathcal{S}')$ contains more vectors than $\text{span}\ (\mathcal{S})$ and the two spans are not the same. This concept will be useful for parts of the following exercise.

3. In each of the following, two sets $\mathcal{S}_1$ and $\mathcal{S}_2$ are given. In each case, determine whether or not $\text{span}(\mathcal{S}_1)$ and $\text{span}(\mathcal{S}_2)$ are equal. You should be able to do (a) - (d) without any computations other than some simple mental arithmetic.

   (a) $S_1 = \left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 7 \end{bmatrix} \right\}, \qquad S_2 = \left\{ \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \begin{bmatrix} -6 \\ 2 \end{bmatrix} \right\}$

   (b) $S_1 = \left\{ \begin{bmatrix} 1 \\ -5 \\ 2 \end{bmatrix} \right\}, \qquad S_2 = \left\{ \begin{bmatrix} 2 \\ 0 \\ 7 \end{bmatrix} \right\}$

   (c) $S_1 = \left\{ \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix} \right\}, \qquad S_2 = \left\{ \begin{bmatrix} -9 \\ 3 \\ -12 \end{bmatrix} \right\}$

   (d) $S_1 = \left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 7 \end{bmatrix} \right\}, \qquad S_2 = \left\{ \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \end{bmatrix} \right\}$

   (e) $S_1 = \left\{ \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 7 \end{bmatrix} \right\}, \qquad S_2 = \left\{ \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ -4 \\ 1 \end{bmatrix} \right\}$

   (f) $S_1 = \left\{ \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 7 \end{bmatrix} \right\}, \qquad S_2 = \left\{ \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$

4. (a) For the following sets $\mathcal{S}_1$ and $\mathcal{S}_2$, check to see if each vector in $\mathcal{S}_2$ is in $\text{span}(\mathcal{S}_1)$.

$$\mathcal{S}_1 = \left\{ \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ 6 \\ -1 \end{bmatrix} \right\}, \qquad \mathcal{S}_2 = \left\{ \begin{bmatrix} 15 \\ -6 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -6 \\ 0 \end{bmatrix} \right\}$$

Are $\text{span}(\mathcal{S}_1)$ and $\text{span}(\mathcal{S}_2)$ equal?

(b) Give an example of two sets $\mathcal{S}_1$ and $\mathcal{S}_2$, with the same number of vectors in each, for which every vector in $\mathcal{S}_2$ is in $\text{span}(\mathcal{S}_1)$, but $\text{span}(\mathcal{S}_1) \neq \text{span}(\mathcal{S}_2)$.

(c) How could we test two sets of vectors to see if their spans are the same?

5. This exercise will be important in the next section and those that follow. Read both parts of the exercise before attempting part (a) - the two parts are "inverses" of each other, in a sense.

(a) Letting $\mathcal{S} = \left\{ \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$, what does any vector in $\text{span}(\mathcal{S})$ look like, *as a single vector*?

(b) Give a set $\mathcal{S}$ consisting of two *specific* vectors in $\mathbb{R}^3$ whose span consists of all vectors of the form

$\begin{bmatrix} a+b \\ 2a \\ 3b \end{bmatrix}$, where $a$ and $b$ are any real numbers.

6. Consider the sets

$$\mathcal{S}_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \qquad \mathcal{S}_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$$\mathcal{S}_3 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \qquad \mathcal{S}_4 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

$$\mathcal{S}_5 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

(a) For each of the given sets, try to find a linear combination of the vectors in the set that equals the vector $\vec{\mathbf{v}} = \begin{bmatrix} -3 \\ 2 \\ 5 \end{bmatrix}$. When you can find such a linear combination, write your results down, assuming the names of the vectors in each set are $\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2, \vec{\mathbf{u}}_3, ....$

(b) In the cases where you *can* find a linear combination equaling $\vec{\mathbf{v}}$, is the linear combination unique? That is, is there only one linear combination equalling the vector $\vec{\mathbf{v}}$?

## 4.2 Closure of a Set Under an Operation

**Performance Criteria:**

4. (c) Determine whether a set is closed under an operation. If it is, prove that it is; if it is not, give a counterexample.

Consider the set $\{0, 1, 2, 3, ...\}$, which are called the whole numbers. Notice that if we add or multiply any two whole numbers the result is also a whole number, but if we try subtracting two such numbers it is possible to get a number that is not in the set. We say that the whole numbers are **closed under addition and multiplication**, but the set of whole numbers is not closed under subtraction. If we enlarge our set to be the integers $\{..., -3, -2, -1, 0, 1, 2, 3, ...\}$ we get a set that is closed under addition, subtraction and multiplication. These operations we are considering are called **binary operations** because they take two elements of the set and create a single new element. An operation that takes just one element of the set and gives another (possibly the same) element of the set is called a **unary operation**. An example would be absolute value; note that the set of integers is closed under absolute value.

<u>DEFINITION 4.2.1</u>: **Closed Under an Operation**

A set $\mathcal{S}$ is said to be closed under a binary operation $*$ if for every $s$ and $t$ in $\mathcal{S}$, $s * t$ is in $\mathcal{S}$. $\mathcal{S}$ is closed under a unary operation $\langle\ \rangle$ if for every $s$ in $\mathcal{S}$, $\langle s \rangle$ is in $\mathcal{S}$.

Notice that the term "closed," as defined here, only makes sense in the context of a set with an operation. Notice also that *it is the set that is closed*, not the operation. The operation is important as well; as we have seen, a given set can be closed under one operation but not another.

When considering closure of a set $\mathcal{S}$ under a binary operation $*$, our considerations are as follows:

- We first wish to determine whether we think $\mathcal{S}$ *IS* closed under $*$.
- If we do think that $\mathcal{S}$ is closed under $*$, we then need to prove that it is. To do this, we need to take two general, or *arbitrary* elements $x$ and $y$ of $\mathcal{S}$ and show that $x * y$ is in $\mathcal{S}$.
- If we think that $\mathcal{S}$ is not closed under $*$, we need to take two *specific* elements $x$ and $y$ of $\mathcal{S}$ and show that $x * y$ is not in $\mathcal{S}$.

◇ **Example 4.2(a):** The odd integers are the numbers $..., -5, -3, -1, 1, 3, 5, ....$ Are the odd integers closed under addition? Multiplication?

**Solution:** We see that $3 + 5 = 8$. Because $3$ and $5$ are both odd but their sum isn't, the odd integers are not closed under addition. Let's try multiplying some odds:

$$3 \times 5 = 15 \qquad -7 \times 9 = -63 \qquad -1 \times -7 = 7$$

Based on these three examples, it appears that the odd integers are perhaps closed under multiplication. Let's attempt to prove it. First we observe that any number of the form $2n + 1$, where $n$ is any integer, is odd. (This is in fact the definition of an odd integer.) So if we have two *possibly different* odd integers, we can write them as $2m+1$ and $2n+1$, where $m$ and $n$ are not necessarily the same integers. Their product is

$$(2m+1)(2n+1) = 4mn + 2m + 2n + 1 = 2(2mn + m + n) + 1.$$

Because the integers are closed under multiplication and addition, $2mn+m+n$ is an integer and the product of $2m+1$ and $2n+1$ is of the form two times an integer, plus one, so it is odd as well. Therefore the odd integers are closed under multiplication.

---

Closure of a set under an operation is a fairly general concept; let's narrow our focus to what is important to us in linear algebra.

◇ **Example 4.2(b):** Prove that the span of a set $\mathcal{S} = \{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3, ..., \vec{\mathbf{v}}_k\}$ in $\mathbb{R}^n$ is closed under addition and scalar multiplication.

**Solution:** Suppose that $\vec{\mathbf{u}}$ and $\vec{\mathbf{w}}$ are in span($\mathcal{S}$). Then there are scalars $c_1, c_2, c_3, ..., c_k$ and $d_1, d_2, d_3, ..., d_k$ such that

$$\vec{\mathbf{u}} = c_1 \vec{\mathbf{v}}_1 + c_2 \vec{\mathbf{v}}_2 + c_3 \vec{\mathbf{v}}_3 + \cdots + c_k \vec{\mathbf{v}}_k \qquad \text{and} \qquad \vec{\mathbf{w}} = d_1 \vec{\mathbf{v}}_1 + d_2 \vec{\mathbf{v}}_2 + d_3 \vec{\mathbf{v}}_3 + \cdots + d_k \vec{\mathbf{v}}_k .$$

Therefore

$$\begin{aligned} \vec{\mathbf{u}} + \vec{\mathbf{w}} &= (c_1 \vec{\mathbf{v}}_1 + c_2 \vec{\mathbf{v}}_2 + c_3 \vec{\mathbf{v}}_3 + \cdots + c_k \vec{\mathbf{v}}_k) + (d_1 \vec{\mathbf{v}}_1 + d_2 \vec{\mathbf{v}}_2 + d_3 \vec{\mathbf{v}}_3 + \cdots + d_k \vec{\mathbf{v}}_k) \\ &= (c_1 + d_1) \vec{\mathbf{v}}_1 + (c_2 + d_2) \vec{\mathbf{v}}_2 + (c_3 + d_3) \vec{\mathbf{v}}_3 + \cdots + (c_k + d_k) \vec{\mathbf{v}}_k \end{aligned}$$

This last expression is a linear combination of the vectors in $\mathcal{S}$, so it is in span($\mathcal{S}$). Therefore span($\mathcal{S}$) is closed under addition. Now suppose that $\vec{\mathbf{u}}$ is as above and $a$ is any scalar. Then

$$\begin{aligned} a \vec{\mathbf{u}} &= a(c_1 \vec{\mathbf{v}}_1 + c_2 \vec{\mathbf{v}}_2 + c_3 \vec{\mathbf{v}}_3 + \cdots + c_k \vec{\mathbf{v}}_k) \\ &= (ac_1) \vec{\mathbf{v}}_1 + (ac_2) \vec{\mathbf{v}}_2 + (ac_3) \vec{\mathbf{v}}_3 + \cdots + (ac_k) \vec{\mathbf{v}}_k \end{aligned}$$

which is also a linear combination of the vectors in $\mathcal{S}$, so it is also in span($\mathcal{S}$). Thus span($\mathcal{S}$) is closed under multiplication by scalars.

---

The result of the above example is that

> THEOREM 4.2.2: The span of a set $\mathcal{S}$ of vectors is closed under vector addition and scalar multiplication.

This seemingly simple observation is the beginning of one of the most important stories in the subject of linear algebra. The remainder of this chapter will fill out the rest of that story.

**Section 4.2 Exercises** To Solutions

1. Determine whether each of sets described is closed under addition and scalar multiplication. For those that are not closed, give a *specific* example demonstrating that.

   (a) $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} \right\}$

   (b) $\mathcal{S} = \left\{ t \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} \right\}$, where $t$ ranges over all real numbers

   (c) $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} + t \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix} \right\}$, where $t$ ranges over all real numbers

   (d) $\mathcal{S} = \left\{ s \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} + t \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix} \right\}$, where $s$ and $t$ range over all real numbers

   (e) $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \\ 2 \end{bmatrix} \right\}$

2. Determine whether each of the following sets (considered as sets of position vectors) is closed under addition and scalar multiplication. then

   - for those that are not, give a specific counterexample demonstrating that,
   - for those that are, give a set of *specific* vectors that span the given set.

   (a) The line with equation $y = 2x$ in $\mathbb{R}^2$.

   (b) The $yz$-plane in $\mathbb{R}^3$.

   (c) The line with equation $y = 2x + 1$ in $\mathbb{R}^2$.

   (d) $\mathcal{S} = \left\{ s \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} + t \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \right\}$, where $s$ and $t$ range over all real numbers

   (e) The set of all vectors of the form $\begin{bmatrix} a \\ b \\ a+b \end{bmatrix}$, where $a$ and $b$ can be any real number.

   (f) The set of all vectors of the form $\begin{bmatrix} a \\ b \\ ab \end{bmatrix}$, where $a$ and $b$ can be any real number.

## 4.3 Vector Spaces and Subspaces

**Performance Criterion:**

4. (d) Determine whether a subset of $\mathbb{R}^n$ is a subspace. If so, prove it; if not, give an appropriate counterexample.

### Vector Spaces

The term "space" in math simply means a set of objects with some additional special properties. There are metric spaces, function spaces, topological spaces, Banach spaces, and more. The vectors that we have been dealing with make up the **vector spaces** called $\mathbb{R}^2$, $\mathbb{R}^3$ and, for larger values, $\mathbb{R}^n$. In general, a vector space is simply a collection of objects called vectors, a set of scalars, and two operations that satisfy certain properties.

DEFINITION 4.3.1: **Vector Space**

A **vector space** is a set $V$ of objects called **vectors** and a set of scalars (usually the real numbers $\mathbb{R}$), with the operations of vector addition and scalar multiplication, for which the following properties hold for all $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}$, $\vec{\mathbf{w}}$ in $V$ and scalars $c$ and $d$.

1. $\vec{\mathbf{u}}+\vec{\mathbf{v}}$ is in $V$
2. $\vec{\mathbf{u}}+\vec{\mathbf{v}}=\vec{\mathbf{v}}+\vec{\mathbf{u}}$
3. $(\vec{\mathbf{u}}+\vec{\mathbf{v}})+\vec{\mathbf{w}}=\vec{\mathbf{u}}+(\vec{\mathbf{v}}+\vec{\mathbf{w}})$
4. There exists a vector $\vec{\mathbf{0}}$ in $V$ such that $\vec{\mathbf{u}}+\vec{\mathbf{0}}=\vec{\mathbf{u}}$. This vector is called the **zero vector**.
5. For every $\vec{\mathbf{u}}$ in $V$ there exists a vector $-\vec{\mathbf{u}}$ in $V$ such that $\vec{\mathbf{u}}+(-\vec{\mathbf{u}})=\vec{\mathbf{0}}$.
6. $c\,\vec{\mathbf{u}}$ is in $V$.
7. $c\,(\vec{\mathbf{u}}+\vec{\mathbf{v}})=c\vec{\mathbf{u}}+c\vec{\mathbf{v}}$
8. $(c+d)\vec{\mathbf{u}}=c\vec{\mathbf{u}}+d\vec{\mathbf{u}}$
9. $c\,(d\vec{\mathbf{u}})=(cd)\,\vec{\mathbf{u}}$
10. $1\vec{\mathbf{u}}=\vec{\mathbf{u}}$

Note that items 1 and 6 of the above definition say that a vector space is closed under addition and scalar multiplication.

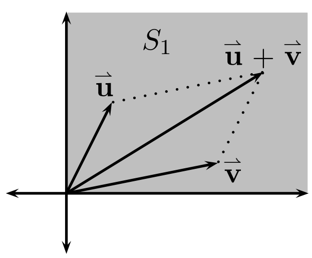
When working with vector spaces, we will be very interested in certain subsets of those vector spaces that are the span of a set of vectors. As you proceed, recall Example 4.2(b), where we showed that *the span of a set of vectors is closed under addition and scalar multiplication.*

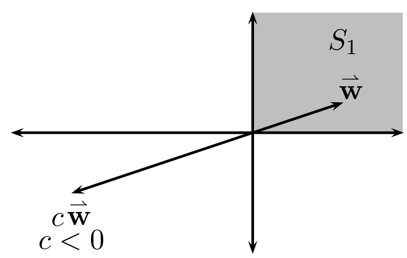
### Subspaces of Vector Spaces

As you should know by now, the two main operations with vectors are multiplication by scalars and addition of vectors. (Note that these two combined give us linear combinations, the foundation of almost everything we've done.) A given vector space can have all sorts of subsets; consider the following subsets of $\mathbb{R}^2$.

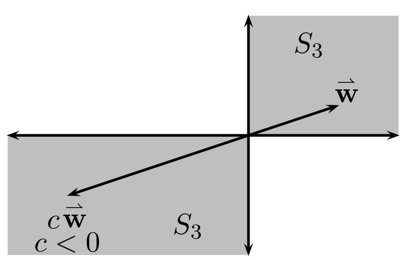
- The set $S_1$ consisting of the first quadrant and the nonnegative parts of the two axes, or all vectors of the form $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ such that $x_1 \geq 0$ and $x_2 \geq 0$.
- The set $S_2$ consisting of the line containing the vector $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$. Algebraically this is all vectors of the form $t\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ where $t$ ranges over all real numbers.
- The set $S_3$ consisting of the first and third quadrants and both axes. This can be described as the set of all vectors $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ with $x_1 x_2 \geq 0$.

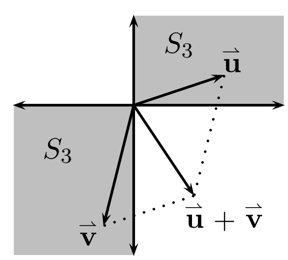
Our current concern is whether these subsets of $\mathbb{R}^2$ are closed under addition and scalar multiplication. With a bit of thought you should see that $S_1$ is closed under addition, but not scalar multiplication when the scalar is negative:




In some sense we can solve the problem of not being closed under scalar multiplication by including the third quadrant as well to get $S_3$, but then the set isn't closed under addition:




Finally, the set $S_2$ is closed under both addition and scalar multiplication. That is a bit messy to show with a diagram, but consider the following. $S_2$ is the span of the single vector $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$, and we showed in the last section that the span of any set of vectors is closed under addition and scalar multiplication.

It turns out that when working with vector spaces the only subsets of any real interest are the ones that are closed under both addition and scalar multiplication. We give such subsets a name:

**Definition 4.3.2: Subspace of $\mathbb{R}^n$**

A subset $S$ of $\mathbb{R}^n$ is called a **subspace** of $\mathbb{R}^n$ if for every scalar $c$ and any vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ in $S$, $c\vec{\mathbf{u}}$ and $\vec{\mathbf{u}}+\vec{\mathbf{v}}$ are also in $S$. That is, $S$ is closed under scalar multiplication and addition.

You will be asked whether certain subsets of $\mathbb{R}^2$, $\mathbb{R}^3$ or $\mathbb{R}^n$ are subspaces, and it is your job to back your answers up with some reasoning. This is done as follows:

- When a subset *IS* a subspace a general proof is required. That is, we must show that the set is closed under scalar multiplication and addition, for *ALL* scalars and *ALL* vectors. We may have to do this outright, but if it is clear that the set of vectors is the span of some set of vectors, then we know from the argument presented in Example 4.2(b) that the set is closed under addition and scalar multiplication, so it is a subspace.
- When a subset *IS NOT* a subspace, we demonstrate that fact with a *SPECIFIC* example. Such an example is called a **counterexample**. Notice that all we need to do to show that a subset is not a subspace is to show that either it is not closed under scalar multiplication *OR* it is not closed under scalar multiplication vector addition. If either is the case, then the set in question is not a subspace. *Even if both are the case, we need only show one.*

The following examples illustrate these things.

◇ **Example 4.3(a):** Show that the set $S_1$ consisting of all vectors of the form $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ such that $x_1 \geq 0$ and $x_2 \geq 0$ is not a subspace of $\mathbb{R}^2$.

**Solution:** As mentioned before, this set is not closed under multiplication by negative scalars, so we just need to give a specific example of this. Let $\vec{\mathbf{u}}=\begin{bmatrix} 3 \\ 5 \end{bmatrix}$ and $c=-2$. Clearly $\vec{\mathbf{u}}$ is in $S_1$ and $c\vec{\mathbf{u}}=(-2)\begin{bmatrix} 3 \\ 5 \end{bmatrix}=\begin{bmatrix} -6 \\ -10 \end{bmatrix}$, which is not in $S_1$. Therefore $S_1$ is not closed under scalar multiplication so it is not a subspace of $\mathbb{R}^2$.

---

◇ **Example 4.3(b):** Show that the set $S_2$ consisting of all vectors of the form $t\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$, where $t$ ranges over all real numbers, is a subspace of $\mathbb{R}^3$.

**Solution:** Let $c$ be any scalar and let $\vec{\mathbf{u}}=s\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$ and $\vec{\mathbf{v}}=t\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$. Then $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are both in $S_2$ and

$$c\,\vec{\mathbf{u}}=c\left(s\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}\right)=(cs)\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \vec{\mathbf{u}}+\vec{\mathbf{v}}=s\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}+t\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}=(s+t)\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}.$$

Because $cs$ and $s+t$ are scalars, we can see that both $c\vec{\mathbf{u}}$ and $\vec{\mathbf{u}}+\vec{\mathbf{v}}$ are in $S_2$, so $S_2$ is a subspace of $\mathbb{R}^3$.

---

This last example demonstrates the general method for showing that a set of vectors is closed under addition and scalar multiplication. That said, the given subspace could have been shown to be a subspace by simply observing that it is the span of the set consisting of the single vector $\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$. From Theorem 4.2.2 we know that the span of any set of vectors is closed under addition and scalar multiplication, so a span of a set is a subspace. Therefore the set described in the above example is a subspace of $\mathbb{R}^2$. Let's formalize all that:

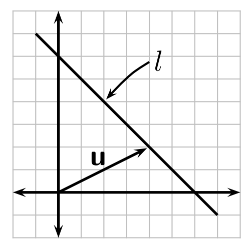
COROLLARY 4.3.3: **The Span of a Set is a Subspace**

The span of a set $\mathcal{S}$ in $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$.

In Example 4.3(d) you will see how we can use the above to prove that a set is a subspace. First we look at another example.

◇ **Example 4.3(c):** Determine whether the subset $S$ of $\mathbb{R}^3$ consisting of all vectors of the form $\vec{\mathbf{x}} = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + t \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix}$ is a subspace. If it is, prove it. If it is not, provide a counterexample.

**Solution:** We recognize this as a line in $\mathbb{R}^3$ passing through the point $(2, 5, -1)$, and it is not hard to show that the line does not pass through the origin. Remember that what we mean by the line is really all position vectors (so with tails at the origin) whose tips are on the line. Considering a similar situation in $\mathbb{R}^2$, we see that $\vec{\mathbf{u}}$ is such a vector for the line $l$ shown. It should be clear that if we multiply $\vec{\mathbf{u}}$ by any scalar other than one, the resulting vector's tip will not lie on the line. Thus we would guess that the set $S$, even though it is in $\mathbb{R}^3$, is probably not closed under scalar multiplication.




Now let's prove that it isn't. To do this we first let $\vec{\mathbf{u}} = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + 1 \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$, which is in $S$. Let $c = 2$, so $2\vec{\mathbf{u}} = \begin{bmatrix} 12 \\ 8 \\ 4 \end{bmatrix}$. We need to show that this vector is not in $S$. If it were, there would have to be a scalar $t$ such that $\begin{bmatrix} 12 \\ 8 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + t \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix}$. Subtracting $\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$ from both sides we get $\begin{bmatrix} 10 \\ 3 \\ 5 \end{bmatrix} = t \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix}$. We can see that the value of $t$ that would be needed to give the correct second component would be $-3$, but this would result in a third component of $-9$, which is not correct. Thus there is no such $t$ and the vector $\begin{bmatrix} 12 \\ 8 \\ 4 \end{bmatrix}$ is not in $S$. Thus $S$ is not closed under scalar multiplication, so it is not a subspace of $\mathbb{R}^3$.

This example may seem to contradict Corollary 4.3.3. The reason it does not is because the vectors of the form $\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + t\begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix}$ do not include *ALL* linear combinations of the vectors $\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix}$.

We should compare the results of Examples 4.3(b) and 4.3(c). Note that both are lines in their respective $\mathbb{R}^n$'s, but the line in 4.3(b) passes through the origin, and the one in 4.3(c) does not. *It is no coincidence that the set in 4.3(b) is a subspace and the set in 4.3(c) is not.* If a set $S$ is a subspace, being closed under scalar multiplication means that zero times any vector in the subspace must also be in the subspace. But zero times a vector is the zero vector $\mathbf{0}$. Therefore

> <u>THEOREM 4.3.4</u>: **Subspaces Contain the Zero Vector**
>
> If a subset $S$ of $\mathbb{R}^n$ is a subspace, then the zero vector of $\mathbb{R}^n$ is in $S$.

This type of a statement is called a **conditional statement**. Related to any conditional statement are two other statements called the **converse** and **contrapositive** of the conditional statement. In this case we have

- **Converse:** If the zero vector is in a subset $S$ of $\mathbb{R}^n$, then $S$ is a subspace.
- **Contrapositive:** If the zero vector is *not* in a subset $S$ of $\mathbb{R}^n$, then $S$ is *not* a subspace.

When a conditional statement is true, its converse may or may not be true. In this case the converse is not true. This is easily seen in Example 4.3(a), where the set contains the zero vector but is not a subspace. However, when a conditional statement is true, its contrapositive is true as well. Therefore the second statement above is the most useful of the three statements, since it gives us a quick way to rule out a set as a subspace. Let's recognize it formally.

> <u>COROLLARY 4.3.5</u>: **Test For Not a Subspace**
>
> If a subset $S$ of $\mathbb{R}^n$ does not the zero vector of $\mathbb{R}^n$, then it is not a subspace of $\mathbb{R}^n$.

In Example 4.3(c) this would have saved us the trouble of providing a counterexample, although we'd still need to convincingly show that the zero vector is not in the set.

◇ **Example 4.3(d):** Determine whether the set of all vectors of the form $\vec{\mathbf{x}} = \begin{bmatrix} a \\ a+b \\ b \\ a-b \end{bmatrix}$, for some real numbers $a$ and $b$, is a subspace of $\mathbb{R}^4$.

**Solution:** We first note that a vector $\vec{\mathbf{x}}$ of the given form will be the zero vector if $a = b = 0$. We cannot then use the above result to rule out the possibility that the given set is a subspace, but neither do we yet know it *IS* a subspace. But we observe that

$$\vec{\mathbf{x}} = \begin{bmatrix} a \\ a+b \\ b \\ a-b \end{bmatrix} = \begin{bmatrix} a \\ a \\ 0 \\ a \end{bmatrix} + \begin{bmatrix} 0 \\ b \\ b \\ -b \end{bmatrix} = a\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}.$$

Thus, by Corollary 4.3.3, the set of vectors under consideration is the span of $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix} \right\}$, so it is a subspace of $\mathbb{R}^4$.

---

We conclude this section with an example that gives us the "largest" and "smallest" subspaces of $\mathbb{R}^n$.

◇ **Example 4.3(e):** Of course a scalar times any vector in $\mathbb{R}^n$ is also in $\mathbb{R}^n$, and the sum of any two vectors in $\mathbb{R}^n$ is in $\mathbb{R}^n$, so $\mathbb{R}^n$ is a subspace of itself. Also, the zero vector by itself is a subspace of $\mathbb{R}^n$ as well, often called the **trivial subspace**.

---

At this point we have seen a variety of subspaces, and some sets that are not subspaces as well. Now supose that we have two vectors in $\mathbb{R}^3$ that are not scalar multiples of each other. We know that the span of the two vectors is a plane through the origin in $\mathbb{R}^3$. Note that we could impose a coordinate system on any plane to make it essentially $\mathbb{R}^2$, so we can think of this particular variety of subspace as just being a copy of $\mathbb{R}^2$ "sitting inside" $\mathbb{R}^3$. This illustrates what is in fact a general principle: *any subspace of* $\mathbb{R}^n$ *is essentially a copy of* $\mathbb{R}^m$, *for some* $m \le n$, *sitting inside* $\mathbb{R}^n$ *with its origin at the origin of* $\mathbb{R}^n$. More formally we have the following:

**Subspaces of** $\mathbb{R}^n$

- The only non-trivial subspaces of $\mathbb{R}^2$ are lines through the origin and all of $\mathbb{R}^2$.
- The only non-trivial subspaces of $\mathbb{R}^3$ are lines through the origin, planes through the origin, and all of $\mathbb{R}^3$.
- The only non-trivial subspaces of $\mathbb{R}^n$ are hyperplanes (including lines) through the origin and all of $\mathbb{R}^n$.

**Section 4.3 Exercises** To Solutions

1. For each of the following sub*sets* of $\mathbb{R}^3$, think of each point as a position vector; each set then becomes a set of vectors rather than points. For each,

   - determine whether the set is a sub*space* and
   - if it is *NOT* a subspace, give a reason why it isn't by doing one of the following:
     - ◇ stating that the set does not contain the zero vector
     - ◇ giving a vector that is in the set and a scalar multiple that isn't (show that it isn't)
     - ◇ giving two vectors that in the set and showing that their sum is not in the set

   (a) All points on the horizontal plane at $z = 3$.

   (b) All points on the $xz$-plane.

   (c) All points on the line containing $\vec{\mathbf{u}} = [-3, 1, 4]$.

   (d) All points on the lines containing $\vec{\mathbf{u}} = [-3, 1, 4]$ and $\vec{\mathbf{v}} = [5, 0, 2]$.

   (e) All points for which $x \geq 0, y \geq 0$ and $z \geq 0$.

   (f) All points $\vec{\mathbf{x}}$ given by $\vec{\mathbf{x}} = \vec{\mathbf{w}} + s\,\vec{\mathbf{u}} + t\,\vec{\mathbf{v}}$, where $\vec{\mathbf{w}} = [1, 1, 1]$ and $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are as in (d).

   (g) All points $\vec{\mathbf{x}}$ given by $\vec{\mathbf{x}} = s\,\vec{\mathbf{u}} + t\,\vec{\mathbf{v}}$, where $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are as in (d).

   (h) The vector **0**.

   (i) All of $\mathbb{R}^3$.

2. Consider the vectors $\vec{\mathbf{u}} = \begin{bmatrix} 8 \\ -2 \\ 4 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 7 \\ 0 \\ 1 \end{bmatrix}$ and $\vec{\mathbf{w}} = \begin{bmatrix} -16 \\ 4 \\ -8 \end{bmatrix}$.

   (a) Is the set of all vectors $\vec{\mathbf{x}} = \vec{\mathbf{u}} + t\,\vec{\mathbf{v}}$, where $t$ ranges over all real numbers, a subspace of $\mathbb{R}^3$? If not, tell why not.

   (b) Is the set of all vectors $\vec{\mathbf{x}} = \vec{\mathbf{u}} + t\,\vec{\mathbf{w}}$, where $t$ ranges over all real numbers, a subspace of $\mathbb{R}^3$? If not, tell why not.

## 4.4 Column Space and Null Space of a Matrix

**Performance Criteria:**

4. (e) Determine whether a vector is in the column space or null space of a matrix, based only on the definitions of those spaces.

In this section we will define two important subspaces associated with a matrix $A$, its **column space** and its **null space**.

DEFINITION 4.4.1: **Column Space of a Matrix**

The **column space** of an $m \times n$ matrix $A$ is the span of the columns of $A$. It is a subspace of $\mathbb{R}^m$ and we denote it by col$(A)$.

◇ **Example 4.4(a):** Determine whether $\vec{\mathbf{u}} = \begin{bmatrix} 3 \\ 3 \\ 8 \end{bmatrix}$ and $\vec{\mathbf{v}} = \begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix}$ are in the column space of $A = \begin{bmatrix} 2 & 5 & 1 \\ -1 & -7 & -5 \\ 3 & 4 & -2 \end{bmatrix}$.

**Solution:** We need to solve the two vector equations of the form

$$c_1 \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} 5 \\ -7 \\ 4 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ -5 \\ -2 \end{bmatrix} = \vec{\mathbf{b}}, \tag{1}$$

with $\vec{\mathbf{b}}$ first being $\vec{\mathbf{u}}$, then $\vec{\mathbf{v}}$. The respective reduced row-echelon forms of the augmented matrices corresponding to the two systems are

$$\begin{bmatrix} 1 & 0 & -2 & 4 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Therefore we can find scalars $c_1$, $c_2$ and $c_3$ for which (1) holds when $\vec{\mathbf{b}} = \vec{\mathbf{u}}$, but not when $\vec{\mathbf{b}} = \vec{\mathbf{v}}$. From this we deduce that $\vec{\mathbf{u}}$ is in col$(A)$, but $\vec{\mathbf{v}}$ is not.

---

Recall that the system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ of $m$ linear equations in $n$ unknowns can be written in linear combination form:

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Note that the left side of this equation is simply a linear combination of the columns of $A$, with the scalars being the components of $\vec{\mathbf{x}}$. The system will have a solution if, and only if, $\vec{\mathbf{b}}$ can be written as a linear combination of the columns of $A$. Stated another way, we have the following:

> THEOREM 4.4.2: A system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a solution (meaning *at least* one solution) if, and only if, $\vec{\mathbf{b}}$ is in the column space of $A$.

Let's consider now only the case where $m = n$, so we have $n$ linear equations in $n$ unknowns. We have the following facts:

- If col$(A)$ is all of $\mathbb{R}^n$, then $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ will have a solution for any vector $\vec{\mathbf{b}}$. What's more, *the solution will be unique*.
- If col$(A)$ is a proper subspace of $\mathbb{R}^n$ (that is, it is not all of $\mathbb{R}^n$), then the equation $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ will have a solution if, and only if, $\vec{\mathbf{b}}$ is in col$(A)$. If $\vec{\mathbf{b}}$ is in col$(A)$ the system will have infinitely many solutions.

Next we define the **null space** of a matrix.

> DEFINITION 4.4.3: **Null Space of a Matrix**
>
> The **null space** of an $m \times n$ matrix $A$ is the set of all solutions to $A\vec{\mathbf{x}} = \mathbf{0}$. It is a subspace of $\mathbb{R}^n$ and is denoted by null$(A)$.

◇ **Example 4.4(b):** Determine whether $\vec{\mathbf{u}} = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}$ and $\vec{\mathbf{v}} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ are in the null space of $A = \begin{bmatrix} 2 & 5 & 1 \\ -1 & -7 & -5 \\ 3 & 4 & -2 \end{bmatrix}$.

**Solution:** A vector $\vec{\mathbf{x}}$ is in the null space of a matrix $A$ if $A\vec{\mathbf{x}} = \mathbf{0}$. We see that

$$A\vec{\mathbf{u}} = \begin{bmatrix} 2 & 5 & 1 \\ -1 & -7 & -5 \\ 3 & 4 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ -21 \\ 11 \end{bmatrix} \quad \text{and} \quad A\vec{\mathbf{v}} = \begin{bmatrix} 2 & 5 & 1 \\ -1 & -7 & -5 \\ 3 & 4 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

so $\vec{\mathbf{v}}$ is in null$(A)$ and $\vec{\mathbf{u}}$ is not.

Again considering the case where $m = n$, we have the following fact about the null space:

- If null$(A)$ is just the zero vector, $A$ is invertible and $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a unique solution for any vector $\vec{\mathbf{b}}$.

We conclude by pointing out the important fact that for an $m \times n$ matrix $A$, the null space of $A$ is a subspace of $\mathbb{R}^n$ and the column space of $A$ is a subspace of $\mathbb{R}^m$.

**Section 4.4 Exercises** To Solutions

1. For each of the following, a matrix $A$ and a vector $\vec{\mathbf{u}}$ are given. in each case, determine whether $\vec{\mathbf{u}}$ is in the column space of $A$. **You should be able to do all of these by inspection, meaning without doing any computations other than some mental arithmetic**.

   (a) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \vec{\mathbf{u}} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$

   (b) $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \vec{\mathbf{u}} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$

   (c) $A = \begin{bmatrix} 2 & -3 \\ -5 & 1 \end{bmatrix}, \quad \vec{\mathbf{u}} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$

   (d) $A = \begin{bmatrix} 2 & -3 \\ -4 & 6 \end{bmatrix}, \quad \vec{\mathbf{u}} = \begin{bmatrix} 8 \\ -16 \end{bmatrix}$

   (e) $A = \begin{bmatrix} 2 & -3 \\ -4 & 6 \end{bmatrix}, \quad \vec{\mathbf{u}} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$

   (f) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \vec{\mathbf{u}} = \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix}$

2. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & -2 \\ -1 & -4 & 6 \end{bmatrix}, \quad \vec{\mathbf{u}}_1 = \begin{bmatrix} 2 \\ 9 \\ -17 \end{bmatrix}, \quad \vec{\mathbf{u}}_2 = \begin{bmatrix} 3 \\ 15 \\ 2 \end{bmatrix}$.

   (a) Remember that the column space of $A$ is simply the set of all vectors that are linear combinations of the columns of $A$. Determine whether the vector $\vec{\mathbf{u}}_1$ is in the column space of $A$ by determining whether $\vec{\mathbf{u}}_1$ is a linear combination of the columns of $A$. Give the vector equation that you are trying to solve, and your row reduced augmented matrix. **Be sure to tell whether $\vec{\mathbf{u}}_1$ is in the column space of $A$ or not! Do this with a brief sentence.**

   (b) If $\vec{\mathbf{u}}_1$ *IS* in the column space of $A$, give a *specific* linear combination of the columns of $A$ that equals $\vec{\mathbf{u}}_1$.

   (c) Repeat parts (a) and (b) for the vector $\vec{\mathbf{u}}_2$.

3. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & -2 \\ -1 & -4 & 6 \end{bmatrix}, \quad \vec{\mathbf{v}}_1 = \begin{bmatrix} 8 \\ -8 \\ -4 \end{bmatrix}, \quad \vec{\mathbf{v}}_2 = \begin{bmatrix} 5 \\ 0 \\ -7 \end{bmatrix}$. The null space of $A$ is just all the vectors $\vec{\mathbf{x}}$ for which $A\vec{\mathbf{x}} = \mathbf{0}$, and it is denoted by null($A$). This means that to check to see if a vector $\vec{\mathbf{x}}$ is in the null space we need only to compute $A\vec{\mathbf{x}}$ and see if it is the zero vector. Use this method to determine whether either of the vectors $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ is in null($A$). Give your answer as a brief sentence.

4. Let $A = \begin{bmatrix} 3 & -2 & 1 \\ 1 & 3 & 4 \\ -1 & -4 & -5 \end{bmatrix}$.

   (a) Is $\vec{\mathbf{u}}_1 = \begin{bmatrix} 5 \\ -2 \\ 3 \end{bmatrix}$ in $\mathrm{col}(A)$? Justify your answer.

   (b) Is $\vec{\mathbf{u}}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in $\mathrm{col}(A)$? Justify your answer.

   (c) Does the system $A\,\vec{\mathbf{x}} = \vec{\mathbf{b}}$ have a solution for all vectors $\vec{\mathbf{b}}$ in $\mathbb{R}^3$? Explain.

   (d) Give a non-zero vector $\vec{\mathbf{v}}_1$ that is in the null space of $A$.

   (e) Is $A$ invertible? Explain.

5. (a) Give a $2 \times 2$ matrix $A$ with no zero entries whose column space is all of $\mathbb{R}^2$.

   (b) Give a $2 \times 2$ matrix $B$ with no zero entries whose column space is *NOT* all of $\mathbb{R}^2$.

   (c) For either of your matrices $A$ and $B$ from parts (a) and (b) that you can, find a nonzero vector in the null space of the matrix.

6. For non-square matrices, the column space and null space contain vectors in different $\mathbb{R}^n$s. Consider the matrix

$$A = \begin{bmatrix} 2 & -3 & 1 \\ -5 & 1 & 4 \end{bmatrix}.$$

   (a) Vectors in the column space are in what $\mathbb{R}^n$?

   (b) Vectors in the null space are in what $\mathbb{R}^n$?

   (c) Can you find a vector in the $\mathbb{R}^n$ of the column space that *IS NOT* in the column space of $A$? If so, give one. If not, explain.

   (d) Can you find a nonzero vector in the $\mathbb{R}^n$ of the null space that *IS* in the null space of $A$? If so, give one. If not, explain.

## 4.5 Least Squares Solutions to Inconsistent Systems

**Performance Criterion:**

4. (f) Find the least-squares approximation to the solution of an inconsistent system of equations. Solve a problem using least-squares approximation.
   (g) Give the least squares error and least squares error vector for a least squares approximation to a solution to a system of equations.

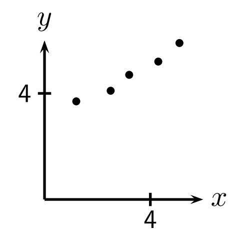
Let's begin with a motivating example.

◇ **Example 4.5(a):** Find the equation of the line containing the five points with coordinates

$$(1.2, 3.7) \quad (2.5, 4.1) \quad (3.2, 4.7) \quad (4.3, 5.2) \quad (5.1, 5.9).$$

**Solution:** We substitute each pair of points into the equation $y = mx + b$ of a line to get the system shown below and to the left.

$$\begin{aligned} 1.2m + b &= 3.7 \\ 2.5m + b &= 4.1 \\ 3.2m + b &= 4.7 \\ 4.3m + b &= 5.2 \\ 5.1m + b &= 5.9 \end{aligned} \quad \Longrightarrow \quad \begin{bmatrix} 1.2 & 1 & 3.7 \\ 2.5 & 1 & 4.1 \\ 3.2 & 1 & 4.7 \\ 4.3 & 1 & 5.2 \\ 5.1 & 1 & 5.9 \end{bmatrix} \stackrel{rref}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$




When then row-reduce the augmented matrix for this system in order to solve the system. When we do this, we get the last matrix above as our row-reduced matrix, indicating that the system has no solution. The five points are plotted above and to the right, and we can see that they are not on a line, which is why we were not able to solve the system. You may remember that such a system is sometimes referred to as **overdetermined**, meaning that there is too much information to allow a solution to the system. We conclude that there is no line through all of the given points.

---

Recall that an inconsistent system is one for which there is no solution. Often we wish to solve an inconsistent system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$, and it is just not acceptable to have no solution. In those cases we can find some vector (whose components are the values we are trying to find when attempting to solve the system) that is "closer to being a solution" than all other vectors. The theory behind this process is part of the second term of this course, but we now have enough knowledge to find such a vector in a "cookbook" manner.

Suppose that we have a system of equations $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$. Pause for a moment to reflect on what we know and what we are trying to find when solving such a system: We have a system of linear equations, and the entries of $A$ are the coefficients of all the equations. The vector $\vec{\mathbf{b}}$ is the vector whose components are the right sides of all the equations, and the vector $\vec{\mathbf{x}}$ is the vector whose components are the unknown values of the variables we are trying to find. So we know $A$ and $\vec{\mathbf{b}}$ and we are trying to find $\vec{\mathbf{x}}$. If $A$ is invertible, the solution vector $\vec{\mathbf{x}}$ is given by $\vec{\mathbf{x}}= A^{-1}\vec{\mathbf{b}}$. If $A$ is not invertible there will be no solution vector $\vec{\mathbf{x}}$, but we can usually find a vector $\overline{\mathbf{x}}$ (usually spoken as "ex-bar") that comes "closest" to being a solution. Here is the formula telling us how to find that $\overline{\mathbf{x}}$:

**THEOREM 4.5.1: The Least Squares Theorem:** Let $A$ be an $m \times n$ matrix and let $\vec{\mathbf{b}}$ be in $\mathbb{R}^m$. If $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a **least squares solution** $\overline{\mathbf{x}}$, it is given by

$$\overline{\mathbf{x}} = (A^T A)^{-1} A^T \vec{\mathbf{b}}$$

◇ **Example 4.5(b):** Find the least squares solution to

$$\begin{aligned} 1.3x_1 + 0.6x_2 &= 3.3 \\ 4.7x_1 + 1.5x_2 &= 13.5 \\ 3.1x_1 + 5.2x_2 &= -0.1 \end{aligned}.$$

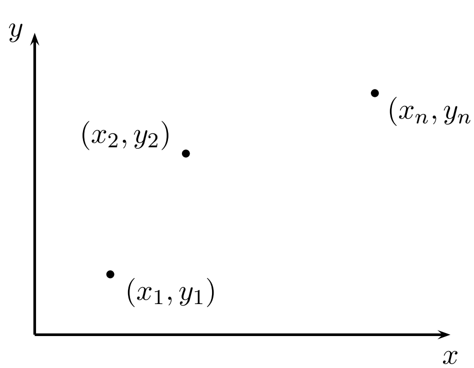
**Solution:** First we note that if we try to solve by row reduction we get no solution; this is an overdetermined system because there are more equations than unknowns. The matrix $A$ and vector $\vec{\mathbf{b}}$ are

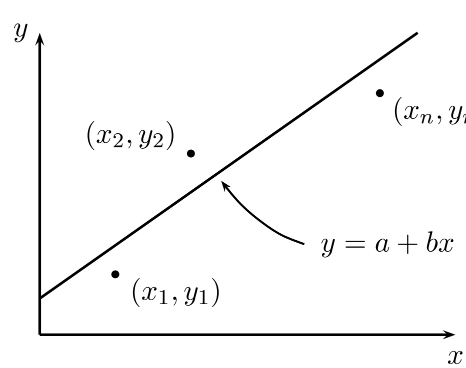
$$A = \begin{bmatrix} 1.3 & 0.6 \\ 4.7 & 1.5 \\ 3.1 & 5.2 \end{bmatrix}, \qquad \vec{\mathbf{b}} = \begin{bmatrix} 3.3 \\ 13.5 \\ -0.1 \end{bmatrix}$$

Using a calculator or *MATLAB*, we get

$$\overline{\mathbf{x}} = (A^T A)^{-1} A^T \vec{\mathbf{b}} = \begin{bmatrix} 3.5526 \\ -2.1374 \end{bmatrix}$$

Example 4.5(a) is a classic example of when we want to do something like this. We have a bunch of $(x, y)$ data pairs from some experiment, and when we graph all the pairs they describe a trend. We then want to find a simple function $y = f(x)$ that best models that data. In some cases that function might be a line, in other cases maybe it is a parabola, and in yet other cases it might be an exponential function. Let's try to make the connection between this and linear algebra. Suppose that we have the data points $(x_1, y_1)$, $(x_2, y_2)$, ..., $(x_n, y_n)$, and when we graph these points they arrange themselves in roughly a line, as shown to the left below. We then want to find an equation of the form $a + bx = y$ (note that this is just the familiar $y = mx + b$ rearranged and with different letters for the slope and $y$-intercept) such that $a + bx_i \approx y_i$ for $i = 1, 2, ..., n$, as shown to the right below.




If we substitute each data pair into $a + bx = y$ we get a system of equations which can be thought of in several different ways. Remember that all the $x_i$ and $y_i$ are known values - the unknowns are

$a$ and $b$.

$$\begin{array}{c} a + x_1 b = y_1 \\ a + x_2 b = y_2 \\ \vdots \\ a + x_n b = y_n \end{array} \quad \Longleftrightarrow \quad \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \Longleftrightarrow \quad A\vec{\mathbf{x}} = \vec{\mathbf{b}}$$

Above we first see the system that results from putting each of the $(x_i, y_i)$ pairs into the equation $a + bx = y$. After that we see the $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ form of the system. We must be careful of the notation here. $A$ is the matrix whose columns are a vector in $\mathbb{R}^n$ consisting of all ones and a vector whose components are the $x_i$ values. It would be logical to call this last vector $\vec{\mathbf{x}}$, but instead $\vec{\mathbf{x}}$ is the vector $\begin{bmatrix} a \\ b \end{bmatrix}$. $\vec{\mathbf{b}}$ is the column vector whose components are the $y_i$ values. Our task, as described by this interpretation, is to *find a vector* $\vec{\mathbf{x}}$ *in* $\mathbb{R}^2$ *that* $A$ *transforms into the vector* $\vec{\mathbf{b}}$ *in* $\mathbb{R}^n$. Even if such a vector did exist, it couldn't be given as $\vec{\mathbf{x}} = A^{-1}\vec{\mathbf{b}}$ because $A$ is not square, so can't be invertible. However, it is likely no such vector exists, but we *CAN* find the least-squares vector $\overline{\mathbf{x}} = \begin{bmatrix} a \\ b \end{bmatrix} = (A^T A)^{-1} A^T \vec{\mathbf{b}}$. When we do, its components $a$ and $b$ are the intercept and slope of our line.

◇ **Example 4.5(c):** Find the the least squares solution to the problem from Example 4.5(a).

**Solution:** Recall that we obtained the system shown below and to the left. The matrix form of the system is shown to the right of the system.

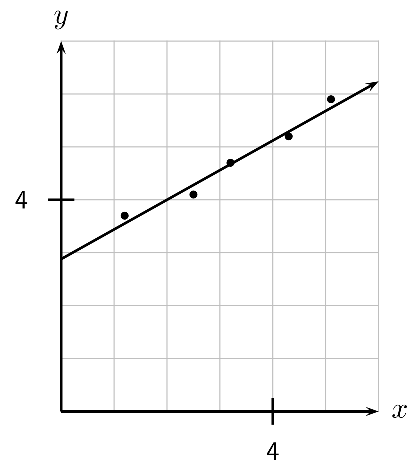
$$\begin{array}{rcl} 1.2m + b & = & 3.7 \\ 2.5m + b & = & 4.1 \\ 3.2m + b & = & 4.7 \\ 4.3m + b & = & 5.2 \\ 5.1m + b & = & 5.9 \end{array} \quad \Longrightarrow \quad \begin{bmatrix} 1.2 & 1 \\ 2.5 & 1 \\ 3.2 & 1 \\ 4.3 & 1 \\ 5.1 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 3.7 \\ 4.1 \\ 4.7 \\ 5.2 \\ 5.9 \end{bmatrix}$$

From this we determine that $A = \begin{bmatrix} 1.2 & 1 \\ 2.5 & 1 \\ 3.2 & 1 \\ 4.3 & 1 \\ 5.1 & 1 \end{bmatrix}$, $\vec{\mathbf{x}} = \begin{bmatrix} m \\ b \end{bmatrix}$, $\vec{\mathbf{b}} = \begin{bmatrix} 3.7 \\ 4.1 \\ 4.7 \\ 5.2 \\ 5.9 \end{bmatrix}$. Note the difference between the $y$-intercept $b$ of the line we are looking for and the vector $\vec{\mathbf{b}}$! We now apply the least squares theorem to obtain $\bar{x} = (A^T A)^{-1} A^T \vec{\mathbf{b}} = \begin{bmatrix} 0.56 \\ 2.88 \end{bmatrix}$. The equation of the line that is "closest" to containing the points

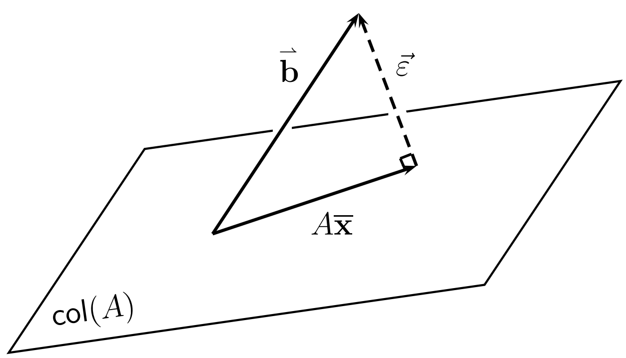
$$(1.2, 3.7) \quad (2.5, 4.1) \quad (3.2, 4.7) \quad (4.3, 5.2) \quad (5.1, 5.9).$$

is then $y = 0.56x + 2.88$.

To the right we see a plot of the five points and the line $y = 0.56x + 2.88$, showing that the line does a good job of coming close to going through all of the points.




Theoretically, here is what is happening when use least squares to solve an inconsistent system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$. The fact that the system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has no solution means that $\vec{\mathbf{b}}$ is not in the column space of $A$. The least squares solution to $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is simply the vector $\overline{\mathbf{x}}$ for which $A\overline{\mathbf{x}}$ is the closest vector to $\vec{\mathbf{b}}$ that is still in $\mathrm{col}(A)$ - it is the projection of $\vec{\mathbf{b}}$ onto the column space of $A$. This is shown simplistically below, for the situation where the column space is a plane in $\mathbb{R}^3$.




To recap a bit, suppose we have a system of equations $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ where there is no vector $\vec{\mathbf{x}}$ for which $A\vec{\mathbf{x}}$ equals $\vec{\mathbf{b}}$. What the least squares approximation allows us to do is to find a vector $\overline{\mathbf{x}}$ for which $A\overline{\mathbf{x}}$ is as "close" to $\vec{\mathbf{b}}$ as possible. We generally determine "closeness" of two objects by finding the difference between them. Because both $A\overline{\mathbf{x}}$ and $\vec{\mathbf{b}}$ are both vectors with the same number of components, we can subtract them to get a vector $\vec{\varepsilon}$ that we will call the **error vector**, shown above. The **least squares error** is then the magnitude of this vector:

<u>Definition 4.5.2</u>: If $\overline{\mathbf{x}}$ is the least-squares solution to the system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$, the **least squares error vector** is

$$\vec{\varepsilon} = \vec{\mathbf{b}} - A\overline{\mathbf{x}}$$

and the **least squares error** is the magnitude of $\vec{\varepsilon}$.

◇ **Example 4.5(b):** Find the least squares error vector and least squares error vector for the solution obtained in Example 4.5(a).

**Solution:** The least squares error vector is

$$\vec{\varepsilon} = \vec{\mathbf{b}} - A\overline{\mathbf{x}} = \begin{bmatrix} 3.3 \\ 13.5 \\ -0.1 \end{bmatrix} - \begin{bmatrix} 1.3 & 0.6 \\ 4.7 & 1.5 \\ 3.1 & 5.2 \end{bmatrix} \begin{bmatrix} 3.5526 \\ -2.1374 \end{bmatrix} = \begin{bmatrix} -0.0359 \\ 0.0089 \\ 0.0016 \end{bmatrix}$$

The least squares error is $\|\vec{\varepsilon}\| = 0.0370.$

---

## Section 4.5 Exercises

To Solutions

1. Consider the points $(1,8)$, $(2,7)$, $(3,5)$, $(4,2)$.

   (a) Find the least squares approximating line $y = a + bx$ for the points. Give the system of equations to be solved (in any form), the matrix $A$ and vector $\vec{\mathbf{b}}$, the solution vector $\vec{\mathbf{x}}$ and the equation of the line.

   (b) Plot the points and your line using something like *Desmos*. (To plot the points with *Desmos*, just list them exactly as they are given above.) Does the line seem to do a good job of coming close to all of the points?

   (c) Find the least squares approximating parabola $y = a + bx + cx^2$ for the points. Give the system of equations to be solved (in any form), the matrix $A$ and vector $\vec{\mathbf{b}}$, the solution vector $\vec{\mathbf{x}}$ and the equation of the parabola.

   (d) Plot the points and your parabola using something like *Desmos*. Does the parabola seem to do a good job of coming close to all of the points?

2. We know that through any three *non-collinear* points in $\mathbb{R}^3$ there is exactly one plane. When we have more than three such points there is likely not a plane containing all the points, because this is an overdetermined situation. But we can use least-squares to determine the plane that best approximates a plane through all of the points. If the plane is not vertical, its equation can be written

$$z = a + bx + cy,$$

   where each of $a$, $b$ and $c$ are constants. Find the equation of the plane closest to containing the points

$$(1.1, 0.2, 4.7), \qquad (4.3, 6.5, 5.0), \qquad (3.2, 5.1, 4.5), \qquad (7.4, 2.8, 13.8),$$

   rounding each of $a$, $b$ and $c$ to the hundredth's place.

3. Consider the system $\begin{aligned} 2x_1 + 3x_2 &= 26 \\ -x_1 + 5x_2 &= 13 \end{aligned}$.

   (a) Find the least-squared solution to the system.

   (b) Solve the system by some method you have used for consistent systems (those that have "ordinary" solutions).

   (c) How do your solutions from (a) and (b) compare?

## 4.6 Linear Independence

> **Performance Criterion:**
>
> 4. (h) Determine whether a set $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, ..., \vec{\mathbf{v}}_k$ of vectors is a linearly independent or linearly dependent. If the vectors are linearly dependent, (1) give a non-trivial linear combination of them that equals the zero vector, (2) give any one as a linear combination of the others, when possible.

In Exercise 6 of Section 4.1 we considered several sets of vectors in $\mathbb{R}^3$, including the sets

$$\mathcal{S}_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}, \qquad \mathcal{S}_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\},$$

$$\mathcal{S}_5 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

It should be clear that set $\mathcal{S}_1$ does not span $\mathbb{R}^3$, and that any vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ in $\mathbb{R}^3$ is in the span of $\mathcal{S}_2$:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Because $\mathcal{S}_2$ spans $\mathbb{R}^3$, including the additional vector as done in $\mathcal{S}_5$ does not increase the span, but of course it doesn't decrease it either. Therefore we can construct any vector in $\mathbb{R}^3$ as a linear combination of iether vectors from $\mathcal{S}_2$ or vectors from $\mathcal{S}_5$. We will find that $\mathcal{S}_5$ is undesirable for getting vectors in $\mathbb{R}^3$ for the following reason:

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

and

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

This shows that the vector $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ can be constructed as a linear combination of vectors in $\mathcal{S}_5$ in more than one way. In fact, there are infinitely many different linear combinations of the vectors in $\mathcal{S}_5$ that equal *any* vector in $\mathbb{R}^3$. On the other hand, it should be clear that, given any vector in $\mathbb{R}^3$, there is *only one way* to obtain that vector as a linear combination of vectors in $\mathcal{S}_2$. So even though both sets $\mathcal{S}_2$ and $\mathcal{S}_5$ span $\mathbb{R}^3$, We will find that a set like $\mathcal{S}_2$, with enough vectors to span, but not "too many," is more desirable.

Let's take a look at what is going on from a different perspective. Suppose that we are trying to create a set $\mathcal{S}$ of vectors that spans $\mathbb{R}^3$. We might begin with one vector, say $\vec{\mathbf{u}}_1 = \begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix}$, in $\mathcal{S}$.

We know by now that the span of this single vector is all scalar multiples of it, which is a line in $\mathbb{R}^3$. If we wish to increase the span, we would add another vector to $\mathcal{S}$. If we were to add a vector like $\begin{bmatrix} 6 \\ -2 \\ -4 \end{bmatrix}$ to $\mathcal{S}$, we would not increase the span, because this new vector is a scalar multiple of $\vec{\mathbf{u}}_1$, so it is on the line we already have and would contribute nothing new to the span of $\mathcal{S}$. To increase the span, we need to add to $\mathcal{S}$ a second vector $\vec{\mathbf{u}}_2$ that is not a scalar multiple of the vector $\vec{\mathbf{u}}_1$ that we already have. It should be clear that the vector $\vec{\mathbf{u}}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is not a scalar multiple of $\vec{\mathbf{u}}_1$, so adding it to $\mathcal{S}$ would increase it's span.

The span of $\mathcal{S} = \{\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2\}$ is a plane. When $\mathcal{S}$ included only a single vector, it was relatively easy to determine a second vector that, when added to $\mathcal{S}$, would increase it's span. Now we wish to add a third vector to $\mathcal{S}$ to further increase its span. Geometrically it is clear that we need a third vector that is *not in the plane spanned by* $\{\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2\}$. Probabilistically, just about any vector in $\mathbb{R}^3$ would do, but what we would like to do here is create an algebraic condition that needs to be met by a third vector so that adding it to $\mathcal{S}$ will increase the span of $\mathcal{S}$.

Let's begin with what we *DON'T* want: we don't want the new vector to be in the plane spanned by $\{\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2\}$. Now every vector $\vec{\mathbf{v}}$ in that plane is of the form $\vec{\mathbf{v}} = c_1\vec{\mathbf{u}}_1 + c_2\vec{\mathbf{u}}_2$ for some scalars $c_1$ and $c_2$. We say the vector $\vec{\mathbf{v}}$ created this way is "dependent" on $\vec{\mathbf{u}}_1$ and $\vec{\mathbf{u}}_2$, and that is what causes it to not be helpful in increasing the span of a set that already contains those two vectors. Assuming that neither of $c_1$ and $c_2$ is zero, we could also write

$$\vec{\mathbf{u}}_1 = \frac{c_2}{c_1}\vec{\mathbf{u}}_2 - \frac{1}{c_1}\vec{\mathbf{v}} \qquad \text{and} \qquad \vec{\mathbf{u}}_2 = \frac{c_1}{c_2}\vec{\mathbf{u}}_1 - \frac{1}{c_2}\vec{\mathbf{v}},$$

showing that $\vec{\mathbf{u}}_1$ is "dependent" on $\vec{\mathbf{u}}_2$ and $\vec{\mathbf{v}}$, and $\vec{\mathbf{u}}_2$ is "dependent" on $\vec{\mathbf{u}}_1$ and $\vec{\mathbf{v}}$. So whatever "dependent" means (we'll define it more formally soon), all three vectors are dependent on each other. We can create another equation that is equivalent to all three of the ones given so far, and that does not "favor" any particular one of the three vectors:

$$c_1\vec{\mathbf{u}}_1 + c_2\vec{\mathbf{u}}_2 + c_3\vec{\mathbf{v}} = \vec{\mathbf{0}},$$

where $c_3 = -1$.

Of course, if we want a third vector $\vec{\mathbf{u}}_3$ to add to $\{\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2\}$ to increase its span, we would not want to choose $\vec{\mathbf{u}}_3 = \vec{\mathbf{v}}$; instead we would want a third vector that is "independent" of the two we already have. Based on what we have been doing, we would suspect that we would want

$$c_1\vec{\mathbf{u}}_1 + c_2\vec{\mathbf{u}}_2 + c_3\vec{\mathbf{u}}_3 \neq \vec{\mathbf{0}}\,. \tag{1}$$

Of course even if $\vec{\mathbf{u}}_3$ was not in the plane spanned by $\vec{\mathbf{u}}_1$ and $\vec{\mathbf{u}}_2$, (1) would be true if $c_1 = c_2 = c_3 = 0$, but we want that to be the only choice of scalars that makes (1) true.

We now make the following definition, based on our discussion:

> DEFINITION 4.6.1: **Linear Dependence and Independence**
>
> A set $\mathcal{S} = \{\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2, ..., \vec{\mathbf{u}}_k\}$ of vectors is **linearly dependent** if there exist scalars $c_1, c_2, ..., c_k$, *not all equal to zero* such that
>
> $$c_1\vec{\mathbf{u}}_1 + c_2\vec{\mathbf{u}}_2 + \cdots + c_k\vec{\mathbf{u}}_k = \mathbf{0}. \tag{2}$$
>
> If (2) only holds for $c_1 = c_2 = \cdots = c_k = 0$ the set $\mathcal{S}$ is **linearly independent**.

We can state linear dependence (independence) in either of two ways. We can say that the set is linearly dependent, or the vectors are linearly dependent. Either way is acceptable. Often we will get lazy and leave off the "linear" of linear dependence or linear independence. This does no harm, as there is no other kind of dependence/independence that we will be interested in.

◇ **Example 4.6(a):** Determine whether the vectors $\begin{bmatrix} -1 \\ -7 \\ 3 \\ 11 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -3 \\ 2 \\ 5 \end{bmatrix}$ and $\begin{bmatrix} 7 \\ -1 \\ 4 \\ 3 \end{bmatrix}$ are linearly dependent, or linearly independent. If they are dependent, give a non-trivial linear combination of them that equals the zero vector. (Non-trivial means that not all of the scalars are zero!)

**Solution:** To make such a determination we always begin with the vector equation from the definition:

$$c_1 \begin{bmatrix} -1 \\ -7 \\ 3 \\ 11 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -3 \\ 2 \\ 5 \end{bmatrix} + c_3 \begin{bmatrix} 7 \\ -1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{3}$$

We recognize this as the linear combination form of a system of equations that has the augmented matrix shown below and to the left, which reduces to the matrix shown below and to the right.

$$\begin{bmatrix} -1 & 1 & 7 & 0 \\ -7 & -3 & -1 & 0 \\ 3 & 2 & 4 & 0 \\ 11 & 5 & 3 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

From this we see that there are infinitely many solutions, so there are certainly values of $c_1$, $c_2$ and $c_3$, not all zero, that make (4) true, so the set of vectors is linearly dependent. To find a non-trivial linear combination of the vectors that equals the zero vector we let the free variable $c_3$ *be any value other than zero*. (You should try letting it be zero to see what happens.) If we take $c_3$ to be one, then $c_2 = -5$ and $c_1 = 2$. Then

$$2 \begin{bmatrix} -1 \\ -7 \\ 3 \\ 11 \end{bmatrix} - 5 \begin{bmatrix} 1 \\ -3 \\ 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 7 \\ -1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -14 \\ 6 \\ 22 \end{bmatrix} + \begin{bmatrix} -5 \\ 15 \\ -10 \\ -25 \end{bmatrix} + \begin{bmatrix} 7 \\ -1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

---

◇ **Example 4.6(b):** Determine whether the vectors $\begin{bmatrix} -3 \\ -1 \\ 2 \end{bmatrix}$, $\begin{bmatrix} 4 \\ 7 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 5 \\ -1 \end{bmatrix}$ are linearly dependent, or linearly independent. If they are dependent, give a non-trivial linear combination of them that equals the zero vector. (Non-trivial means that not all of the scalars are zero!)

**Solution:** To make such a determination we always begin with the vector equation from the definition:

$$c_1 \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 4 \\ 7 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} -2 \\ 5 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

We recognize this as the linear combination form of a system of equations that has the augmented matrix shown below and to the left, which reduces to the matrix shown below and to the right.

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

We see that the only solution to the system is $c_1 = c_2 = c_3 = 0$, so the vectors are linearly independent.

---

A comment is in order at this point. The system $c_1 \vec{\mathbf{u}}_1 + c_2 \vec{\mathbf{u}}_2 + \cdots + c_k \vec{\mathbf{u}}_k = \mathbf{0}$ is homogeneous, so it will always have at least the zero vector as a solution. It is precisely *when the only solution is the zero vector that the vectors are linearly independent*.

Now suppose we have a set $\{\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2, ..., \vec{\mathbf{u}}_k\}$ of linearly dependent vectors in $\mathbb{R}^n$. By definitions, there are scalars $c_1, c_2, ..., c_k$, not all equal to zero, such that

$$c_1 \vec{\mathbf{u}}_1 + c_2 \vec{\mathbf{u}}_2 + \cdots + c_k \vec{\mathbf{u}}_k = \vec{\mathbf{0}}$$

Let $c_j$, for some $j$ between $1$ and $k$, be one of the non-zero scalars. (By definition there has to be at least one such scalar.) Then we can do the following:

$$\begin{aligned} c_1 \vec{\mathbf{u}}_1 + c_2 \vec{\mathbf{u}}_2 + \cdots + c_j \vec{\mathbf{u}}_j + \cdots + c_k \vec{\mathbf{u}}_k &= \vec{\mathbf{0}} \\ c_j \vec{\mathbf{u}}_j &= -c_1 \vec{\mathbf{u}}_1 - c_2 \vec{\mathbf{u}}_2 - \cdots - c_k \vec{\mathbf{u}}_k \\ \vec{\mathbf{u}}_j &= -\frac{c_1}{c_j} \vec{\mathbf{u}}_1 - \frac{c_2}{c_j} \vec{\mathbf{u}}_2 - \cdots - \frac{c_k}{c_j} \vec{\mathbf{u}}_k \\ \vec{\mathbf{u}}_j &= d_1 \vec{\mathbf{u}}_1 + d_2 \vec{\mathbf{u}}_2 + \cdots + d_k \vec{\mathbf{u}}_k \end{aligned}$$

This, along with the fact that the above computation can be reversed, gives us the following:

> Theorem 4.6.2: A set $S = \{\vec{\mathbf{u}}_1, \vec{\mathbf{u}}_2, ..., \vec{\mathbf{u}}_k\}$ is linearly dependent if, and only if, at least one of these vectors can be written as a linear combination of the remaining vectors.

The importance of this, which we'll reiterate again later, is that *if we have a set of linearly dependent vectors with a certain span, we can eliminate at least one vector from our original set without reducing the span of the set*. If, on the other hand, we have a set of linearly *independent* vectors, eliminating any vector from the set will reduce the span of the set.

◇ **Example 4.6(c):** In Example 4.6(a) we determined that the vectors $\begin{bmatrix} -1 \\ -7 \\ 3 \\ 11 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -3 \\ 2 \\ 5 \end{bmatrix}$ and $\begin{bmatrix} 7 \\ -1 \\ 4 \\ 3 \end{bmatrix}$ are linearly dependent. Give one of them as a linear combination of the others.

**Solution:** In Example 4.6(a) we found that the vector equation

$$c_1 \begin{bmatrix} -1 \\ -7 \\ 3 \\ 11 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -3 \\ 2 \\ 5 \end{bmatrix} + c_3 \begin{bmatrix} 7 \\ -1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

had infinitely many solutions, one of which was

$$2 \begin{bmatrix} -1 \\ -7 \\ 3 \\ 11 \end{bmatrix} - 5 \begin{bmatrix} 1 \\ -3 \\ 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 7 \\ -1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{4}$$

We can easily solve that equation for the third vector to get

$$\begin{bmatrix} 7 \\ -1 \\ 4 \\ 3 \end{bmatrix} = -2 \begin{bmatrix} -1 \\ -7 \\ 3 \\ 11 \end{bmatrix} + 5 \begin{bmatrix} 1 \\ -3 \\ 2 \\ 5 \end{bmatrix}$$

Note that we could also have solved (4) for either the first or second vectors. Solving for the first would give us

$$\begin{bmatrix} -1 \\ -7 \\ 3 \\ 11 \end{bmatrix} = \frac{5}{2} \begin{bmatrix} 1 \\ -3 \\ 2 \\ 5 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 7 \\ -1 \\ 4 \\ 3 \end{bmatrix},$$

and we could solve for the second in a similar manner.

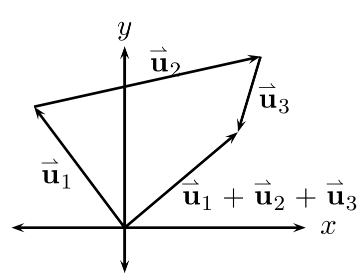
---

◇ **Example 4.6(d):** Determine whether the vectors $\begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix}$ are linearly dependent.

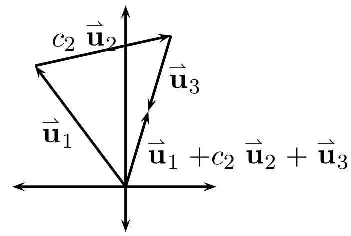
**Solution:** We can easily observe that $\begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix}$, so the first vector is a linear combination of the other two. By Theorem 4.6.2, the three vectors are linearly dependent.

---

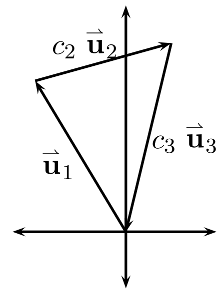
We now consider three vectors $\vec{\mathbf{u}}_1$, $\vec{\mathbf{u}}_2$ and $\vec{\mathbf{u}}_3$ in $\mathbb{R}^2$ whose sum is not the zero vector, and for which no two of the vectors are parallel. I have arranged these to show the tip-to-tail sum in the top diagram to the right; clearly their sum is not the zero vector.



At this point if we were to multiply $\vec{\mathbf{u}}_2$ by some scalar $c_2$ less than one we could shorten it to the point that after adding it to $\vec{\mathbf{u}}_1$ the tip of $c_2\,\vec{\mathbf{u}}_2$ would be in such a position as to line up $\vec{\mathbf{u}}_3$ with the origin. This is shown in the bottom diagram to the right.




Finally, we could then multiply $\vec{\mathbf{u}}_3$ by a scalar $c_3$ greater than one to lengthen it to the point of putting its tip at the origin. We would then have $\vec{\mathbf{u}}_1 + c_2\,\vec{\mathbf{u}}_2 + c_3\,\vec{\mathbf{u}}_3 = \mathbf{0}$. You should play around with a few pictures to convince yourself that this can always be done with three vectors in $\mathbb{R}^2$, as long as none of them are parallel (scalar multiples of each other). This shows us that *any three vectors in* $\mathbb{R}^2$ *are always linearly dependent*. In fact, we can say even more:




$\vec{\mathbf{u}}_1 + c_2\,\vec{\mathbf{u}}_2 + c_3\,\vec{\mathbf{u}}_3 = \mathbf{0}$

> THEOREM 4.6.3: Any set of more than $n$ vectors in $\mathbb{R}^n$ must be linearly dependent.

◇ **Example 4.6(e):** Determine whether the vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 5 \\ 5 \end{bmatrix}$.

**Solution:** This is a set of three vectors in $\mathbb{R}^2$ ($\mathbb{R}^n$ with $n = 2$), so by the above theorem they must be dependent.

Note that Theorem 4.6.3 *doesn't* tell us that a set of $n$ or fewer vectors in $\mathbb{R}^n$ must be linearly independent!

◇ **Example 4.6(f):** Determine whether the vectors $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$ are independent.

**Solution:** Note that this is a set of two vectors in $\mathbb{R}^n$ with $n = 3$. The above theorem is therefore not helpful. However, we can see that the second vector is a linear combination of the first (two times it), so by Theorem 4.6.2 the vectors are linearly dependent.

**Section 4.6 Exercises** To Solutions

1. Determine whether each set of vectors is linearly independent or linearly dependent without applying the definition. That is, use Theorems 4.6.2 and 4.6.3. In each case, explain your reasoning.

(a) $S = \left\{ \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ 7 \\ 0 \end{bmatrix} \right\}$,

(b) $S = \left\{ \begin{bmatrix} 4 \\ 1 \end{bmatrix} \right\}$

(c) $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

(d) $S = \left\{ \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 6 \\ 9 \end{bmatrix} \right\}$

(e) $S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 7 \\ 7 \end{bmatrix} \right\}$

2. Consider the vectors $\vec{\mathbf{u}}_1 = \begin{bmatrix} -5 \\ 9 \\ 4 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} 5 \\ 0 \\ 6 \end{bmatrix}$, and $\vec{\mathbf{u}}_3 = \begin{bmatrix} 5 \\ 9 \\ 16 \end{bmatrix}$.

   (a) Give the *VECTOR* equation that we must consider in order to determine whether the three vectors are linearly independent.

   (b) Your equation has one solution for sure. What is it? What does it mean (in terms of linear dependence or independence) if that is the *ONLY* solution?

   (c) Write your equation from (a) as a system of linear equations. Then give the augmented matrix for the system.

   (d) Does the system have more solutions than the one you gave in (b)? If so, find one of them. (By "one" I mean one ordered triple of three numbers.)

   (e) Find each of the three vectors as a linear combination of the other two.

3. Show that the vectors $\vec{\mathbf{u}} = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 5 \\ 1 \\ -6 \end{bmatrix}$ and $\vec{\mathbf{w}} = \begin{bmatrix} -4 \\ 4 \\ 9 \end{bmatrix}$ are linearly dependent. Then give one of the vectors as a linear combination of the others.

4. For the following, use the vectors $\vec{\mathbf{v}}_1 = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$, $\vec{\mathbf{v}}_3 = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$.

   (a) Determine whether $\vec{\mathbf{u}} = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}$ and $\vec{\mathbf{w}} = \begin{bmatrix} 0 \\ 5 \\ -5 \end{bmatrix}$ are in $\text{span}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3)$.

(b) Show that the vectors $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$ and $\vec{\mathbf{v}}_3$ are linearly dependent by the definition of linearly dependent. In other words, produce scalars $c_1$, $c_2$ and $c_3$ and demonstrate that they and the vectors satisfy the equation given in the definition.

(c) Since the vectors are linearly dependent, at least one the vectors can be expressed as a linear combination of the other two. Express $\vec{\mathbf{v}}_1$ as a linear combination of $\vec{\mathbf{v}}_2$ and $\vec{\mathbf{v}}_3$.

## 4.7 Bases of Subspaces, Dimension

**Performance Criterion:**

4. (i) Determine whether a given set of vectors is a basis for a given subspace. Give a basis and the dimension of a subspace.

We have seen that the span of any set of vectors in $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$. In a sense, the vectors whose span is being considered are the "building blocks" of the subspace. That is, every vector in the subspace is some linear combination of those vectors. Now, recall that if a set of vectors is linearly dependent, one of the vectors can be represented as a linear combination of the others. So if we are considering the span of a set of dependent vectors, we can throw out the one that can be obtained as a linear combination without affecting the span of the set of vectors.

So given a subspace, it is desirable to find what we might call a *minimal spanning set*, the smallest set of vectors whose linear combinations gives the entire subspace. Such a set is called a **basis**.

DEFINITION 4.7.1: **Basis of a Subspace**

For a subspace $S$, a **basis** is a set $\mathcal{B}$ of vectors such that

- the span of $\mathcal{B}$ is $S$,
- the vectors in $\mathcal{B}$ are linearly independent

The plural of basis is *bases* (pronounced "base-eez"). With a little thought, you should believe that *every subspace has infinitely many bases*. (This is a tiny lie - the trivial subspace consisting of just the zero has no basis vectors, which is a funny consequence of logic.)

◇ **Example 4.7(a):** Is the set $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ a basis for $\mathbb{R}^3$?

**Solution:** Clearly for any vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ in $\mathbb{R}^3$ we have $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, so the

span of $\mathcal{B}$ is all of $\mathbb{R}^3$. The augmented matrix for testing for linear independence is simply the identity augmented with the zero vector, giving only the solution where all the scalars are zero, so the vectors are linearly independent. Therefore the set $\mathcal{B}$ is a basis for $\mathbb{R}^3$.

The basis just given is called the **standard basis** for $\mathbb{R}^3$, and its vectors are often denoted by $\vec{\mathbf{e}}_1$, $\vec{\mathbf{e}}_2$ and $\vec{\mathbf{e}}_3$. There is a standard basis for every $\mathbb{R}^n$, and $\vec{\mathbf{e}}_1$ is always the vector whose first component is one and all others are zero, $\vec{\mathbf{e}}_2$ is the vector whose second component is one and all others are zero, and so on.

◇ **Example 4.7(b):** Let $S_1 = \left\{ \begin{bmatrix} -3 \\ 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ 7 \\ -4 \end{bmatrix} \right\}$. In the previous section we saw that span$(S_1)$ is a subspace of $\mathbb{R}^3$. Is $S_1$ a basis for span$(S_1)$?

**Solution:** Clearly $S_1$ meets the first condition for being a basis and, since we can see that neither of these vectors is a scalar multiple of the other, they are linearly independent. Therefore they are a basis for span$(S_1)$.

---

◇ **Example 4.7(c):** Let $S_2 = \left\{ \begin{bmatrix} -3 \\ 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ 7 \\ -4 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \\ -9 \end{bmatrix} \right\}$. Span$(S_2)$ is a subspace of $\mathbb{R}^3$; is $S_2$ a basis for span$(S_2)$?

**Solution:** Once again this set meets the first condition of being a subspace. We can also see that $(-1)\begin{bmatrix} -3 \\ 1 \\ 5 \end{bmatrix} + \begin{bmatrix} 2 \\ 7 \\ -4 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ -9 \end{bmatrix}$, so the set $S_2$ is linearly dependent. Therefore it is NOT a basis for span$(S)$.

---

◇ **Example 4.7(d):** The $yz$-plane in $\mathbb{R}^3$ is a subspace. Give a basis for this subspace.

**Solution:** We know that a set of two linearly independent vectors will span a plane, so we simply need two vectors in the $yz$-plane that are not scalar multiples of each other. The simplest choices are the two vectors $\vec{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ and $\vec{\mathbf{e}}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, so they are a basis for the $yz$-plane.

---

Considering this last example, it is not hard to show that the set $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ is also a basis for the $yz$-plane, and there are many other sets that are bases for that plane as well. All of those sets will contain two vectors, illustrating the fact that *every basis of a subspace has the same number of vectors*. This allows us to make the following definition:

> DEFINITION 4.7.2: **Dimension of a Subspace**
>
> The **dimension** of a subspace is the number of elements in a basis for that subspace.

Looking back at Examples 4.7(a), (b) and (d), we then see that $\mathbb{R}^3$ has dimension three, and span$(S_1)$ has dimension two, and the $yz$-plane in $\mathbb{R}^3$ has dimension two.

Although its importance may not be obvious to you at this point, here's why we care about a basis rather than any set that spans a subspace:

THEOREM 4.7.3: Any vector in a subspace $S$ with basis $\mathcal{B}$ is represented by one, and only one, linear combination of vectors in $\mathcal{B}$.

◇ **Example 4.7(e):** In Example 4.7(d) we determined that the set of all vectors of the form $\vec{\mathbf{x}} = \begin{bmatrix} a \\ a+b \\ b \\ a-b \end{bmatrix}$, for some real numbers $a$ and $b$, is a subspace of $\mathbb{R}^4$. Give a basis for that subspace.

**Solution:** The key computation in Example 4.7(d) was

$$\vec{\mathbf{x}} = \begin{bmatrix} a \\ a+b \\ b \\ a-b \end{bmatrix} = \begin{bmatrix} a \\ a \\ 0 \\ a \end{bmatrix} + \begin{bmatrix} 0 \\ b \\ b \\ -b \end{bmatrix} = a\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}.$$

The set of vectors under consideration is spanned by $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix} \right\}$, and we can see that those two vectors are linearly independent (because they aren't scalar multiples of each other, which is sufficient for independence when considering just two vectors). Therefore they form a basis for the subspace of vectors of the given form.

---

## Section 4.7 Exercises

To Solutions

1. Which of the following is a basis for $\mathbb{R}^2$? For those that aren't, tell why not.

$$S_1 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \qquad S_2 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\} \qquad S_3 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$

2. Which if the following is a basis for $\mathbb{R}^3$? For those that aren't, tell why not.

$$S_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \qquad S_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

$$S_3 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} \right\} \qquad S_4 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\}$$

3. Which of the following is a basis for the line in $\mathbb{R}^2$ containing the vector $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$? For those that aren't, tell why not.

$$S_1 = \left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 2 \end{bmatrix} \right\} \qquad S_2 = \left\{ \begin{bmatrix} 6 \\ 2 \end{bmatrix} \right\} \qquad S_3 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$

4. Which of the following is a basis for the $yz$-plane in $\mathbb{R}^3$? For those that aren't, tell why not.

$$S_1 = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \right\} \qquad S_2 = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

$$S_3 = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\} \qquad S_4 = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

5. For each of the following sub*sets* of $\mathbb{R}^3$, think of each point as a position vector; each set then becomes a set of vectors rather than points. For each,

- determine whether the set is a sub*space* and
- if it is *NOT* a subspace, give a reason why it isn't by doing one of the following:
  - ◇ stating that the set does not contain the zero vector
  - ◇ giving a vector that is in the set and a scalar multiple that isn't (show that it isn't)
  - ◇ giving two vectors that in the set and showing that their sum is not in the set
- if it *IS* a subspace, give a basis for the subspace.

(a) All points on the horizontal plane at $z = 3$.

(b) All points on the $xz$-plane.

(c) All points on the line containing $\vec{\mathbf{u}} = [-3, 1, 4]$.

(d) All points on the lines containing $\vec{\mathbf{u}} = [-3, 1, 4]$ and $\vec{\mathbf{v}} = [5, 0, 2]$.

(e) All points for which $x \geq 0$, $y \geq 0$ and $z \geq 0$.

(f) All points $\vec{\mathbf{x}}$ given by $\vec{\mathbf{x}} = \vec{\mathbf{w}} + s\vec{\mathbf{u}} + t\vec{\mathbf{v}}$, where $\vec{\mathbf{w}} = [1, 1, 1]$ and $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are as in (d).

(g) All points $\vec{\mathbf{x}}$ given by $\vec{\mathbf{x}} = s\vec{\mathbf{u}} + t\vec{\mathbf{v}}$, where $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are as in (d).

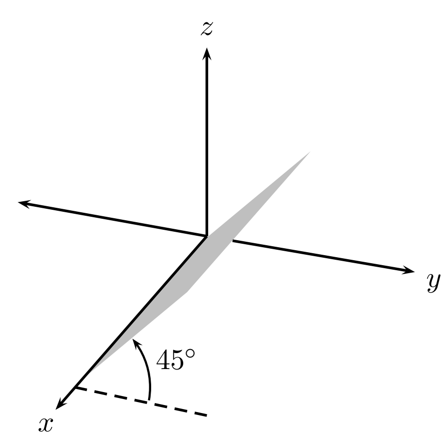
(h) The vector **0**.

(i) All of $\mathbb{R}^3$.

6. Determine whether each of the following is a subspace. If not, give an appropriate counterexample; if so, give a basis for the subspace.

(a) The subset of $\mathbb{R}^2$ consisting of all vectors on or to the right of the $y$-axis.

(b) The subset of $\mathbb{R}^3$ consisting of all vectors in a plane containing the $x$-axis and at a 45 degree angle to the $xy$-plane. See diagram to the right.



7. The $xy$-plane is a subspace of $\mathbb{R}^3$.

   (a) Give a set of at least two vectors in the $xy$-plane that is not a basis for that subspace, and tell why it isn't a basis.

   (b) Give a different set of at least two vectors in the $xy$-plane that is not a basis for that subspace *for a different reason*, and tell why it isn't a basis.

## 4.8 Bases for the Column Space and Null Space of a Matrix

**Performance Criteria:**

4. (j) Find the dimensions of, and bases for, the column space and null space of a given matrix.

   (k) Given the dimension of the column space and/or null space of the coefficient matrix for a system of equations, say as much as you can about how many solutions the system has.

In a previous section you learned about two special subspaces related to a matrix $A$, the column space of $A$ and the null space of $A$. Remember the importance of those two spaces:

A system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a solution if, and only if, $\vec{\mathbf{b}}$ is in the column space of $A$.

If the null space of a square matrix $A$ is just the zero vector, $A$ is invertible and $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a unique solution for any vector $\vec{\mathbf{b}}$.

We would now like to be able to find bases for the column space and null space of a given vector $A$. The following describes how to do this:

Theorem 4.8.1: **Bases for Null Space and Column Space**

- A basis for the column space of a matrix $A$ is the columns of $A$ corresponding to columns of $rref(A)$ that contain leading ones.
- The solution to $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$ (which can be easily obtained from $rref(A)$ by augmenting it with a column of zeros) will be an arbitrary linear combination of vectors. Those vectors form a basis for null($A$).

◇ **Example 4.8(a):** Find bases for the null space and column space of the matrix

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

**Solution:** The reduced row-echelon form of $A$ is shown below and to the left. We can see that the first through third columns contain leading ones, so a basis for the column space of $A$ is the set shown below and to the right.

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -7 \\ 0 & 0 & 1 & 4 \end{bmatrix} \qquad \left\{ \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ 7 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix} \right\}$$

If we were to augment $A$ with a column of zeros to represent the system $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$ and row reduce we'd get the matrix shown above and to the left but with an additional column of zeros on the right. We'd then have $x_4$ as a free variable $t$, with $x_1 = -3t$, $x_2 = 7t$ and $x_3 = -4t$. The solution to $A\vec{\mathbf{x}} = \mathbf{0}$ is any scalar multiple of $\begin{bmatrix} -3 \\ 7 \\ -4 \\ 1 \end{bmatrix}$, so that vector is a basis for the null space of $A$.

---

◇ **Example 4.8(b):** Find a basis for the null space and column space of the matrix

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 3 & 7 & 1 \\ -2 & 1 & 7 \end{bmatrix}.$$

**Solution:** The reduced row-echelon form of this matrix is the identity, so a basis for the column space consists of all the columns of $A$. If we augment $A$ with the zero vector and row reduce we get a solution of the zero vector, so the null space is just the zero vector (which is of course a basis for itself).

---

We should note in the last example that the column space is all of $\mathbb{R}^3$, so for any vector $\vec{\mathbf{b}}$ in $\mathbb{R}^3$ there is a vector $\vec{\mathbf{x}}$ for which $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$. Thus $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a solution for every choice of $\vec{\mathbf{b}}$.

There is an important distinction to be made between a subspace and a basis for a subspace:

- Other than the trivial subspace consisting of the zero vector, *a subspace is an infinite set of vectors*.
- A basis for a subspace *is a finite set of vectors*. In fact a basis consists of relatively few vectors; the basis for any subspace of $\mathbb{R}^n$ contains at most $n$ vectors (and it only contains $n$ vectors if the subspace is all of $\mathbb{R}^n$).

To illustrate, consider the matrix $A = \begin{bmatrix} 1 & 3 & -2 & -4 \\ 3 & 7 & 1 & 4 \\ -2 & 1 & 7 & 7 \end{bmatrix}$ from Example 4.8(a). The set $\left\{ \begin{bmatrix} -3 \\ 7 \\ -4 \\ 1 \end{bmatrix} \right\}$ is a basis for the null space of $A$, whereas the set $\left\{ t \begin{bmatrix} -3 \\ 7 \\ -4 \\ 1 \end{bmatrix} \right\}$ *IS* the null space of $A$.

We finish this section with a couple definitions and a major theorem of linear algebra. The importance of these will be seen in the next section.

DEFINITION 4.8.2: **Rank and Nullity of a Matrix**

- The **rank** of a matrix $A$, denoted rank$(A)$, is the dimension of its column space.
- The **nullity** of a matrix $A$, denoted nullity$(A)$, is the dimension of its null space.

THEOREM 4.8.3: **The Rank Theorem**

For an $m \times n$ matrix $A$, rank$(A)$ + nullity$(A) = n$.

**Section 4.8 Exercises** To Solutions

1. Consider the matrix $A = \begin{bmatrix} 1 & 1 & -2 \\ -3 & -3 & 6 \\ 2 & 2 & -4 \end{bmatrix}$, which has row-reduced form $\begin{bmatrix} 1 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

   In this exercise you will see how to find a basis for the null space of $A$. All this means is that *you are looking for a "minimal" set of vectors whose span (all possible linear combinations of them) give all the vectors* $\vec{\mathbf{x}}$ *for which* $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$.

   (a) Give the augmented matrix for the system of equations $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$, then give its row reduced form.

   (b) There are two free variables, $x_3$ and $x_2$. Let $x_3 = t$ and $x_2 = s$, then find $x_1$ (in terms of $s$ and $t$). Give the vector $\vec{\mathbf{x}}$, in terms of $s$ and $t$.

   (c) Write $\vec{\mathbf{x}}$ as the sum of two vectors, one containing only the parameter $s$ and the other containing only the parameter $t$. Then factor $s$ out of the first vector and $t$ out of the second vector. You now have $\vec{\mathbf{x}}$ as all linear combinations of two vectors.

   (d) Those two vectors are linearly independent, since neither of them is a scalar multiple of the other, so both are essential in the linear combination you found in (c). They then form a basis for the null space of $A$. Write this out as a full sentence, "A basis for ...". *A basis is technically a set of vectors, so use the set brackets* $\{\ \}$ *appropriately.*

2. Consider the matrix $A = \begin{bmatrix} 1 & -1 & 5 \\ 3 & 1 & 11 \\ 2 & 5 & 3 \end{bmatrix}$

   (a) Solve the system $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$. You should get infinitely many solutions containing one or more parameters. Give the general solution, in terms of the parameters. **Give all values in exact form**.

   (b) If you didn't already, you should be able to give the general solution as a linear combination of vectors, with the scalars multiplying them being the parameter(s). Do this.

   (c) The vector or vectors you see in (c) is (are) a basis for the null space of $A$. Give the basis.

3. Consider the matrix $A$ from the previous exercise.

   (a) What is the nullity of $A$?

   (b) From the Rank Theorem, what is the rank of $A$?

   (c) When doing part (a) of the previous exercise you should have obtained the row reduced form of the matrix $A$ (of course you augmented it). A basis for the column space of $A$ is the columns of $A$ (*NOT* the columns of the row reduced form of $A$!) corresponding to the leading variables in the row reduced form of $A$. Give the basis for the column space of $A$.

   (d) Does your result from part (c) agree with the rank that you determined in (b)? If not, find what's wrong and correct it.

4. Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & -2 \\ -1 & -4 & 6 \end{bmatrix}$, $\vec{\mathbf{u}}_1 = \begin{bmatrix} 2 \\ 9 \\ -17 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} 3 \\ 15 \\ 2 \end{bmatrix}$

   (a) Determine whether each of $\vec{\mathbf{u}}_1$ and $\vec{\mathbf{u}}_2$ is in the column space of $A$.

   (b) Find a basis for col$(A)$. **Give your answer with a brief sentence, and indicate that the basis is a set of vectors.**

   (c) One of the vectors $\vec{\mathbf{u}}_1$ and $\vec{\mathbf{u}}_2$ *IS* in the column space of $A$. Give a linear combination of the *basis vectors* that equals that vector.

   (d) What is the rank of $A$?

5. Again let $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & -2 \\ -1 & -4 & 6 \end{bmatrix}$, and let $\vec{\mathbf{v}}_1 = \begin{bmatrix} 8 \\ -8 \\ -4 \end{bmatrix}$, $\vec{\mathbf{v}}_2 = \begin{bmatrix} 5 \\ 0 \\ -7 \end{bmatrix}$.

   (a) Determine whether each of the vectors $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ is in null$(A)$. Give your answer as a brief sentence.

   (b) Determine a basis for null$(A)$, giving your answer in a brief sentence.

   (c) Referring to your answer from part (d) of the previous exercise, what is the nullity of $A$?

   (d) Give the linear combinations of the basis vectors of the null space for either of the vectors $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ that are in the null space.

6. Each of the following matrices is the row-reduced matrix of a given matrix $A$. For each, give $\text{rank}(A)$ and $\text{nullity}(A)$.

   (a) $\begin{bmatrix} 1 & 0 & 6 & -1 \\ 0 & 1 & -5 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

   (b) $\begin{bmatrix} 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 & -1 & 0 & -4 \\ 0 & 1 & 2 & 0 & 5 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

(f) $\begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 0 & 0 & 1 & -3 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

(g) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

## 4.9 Solutions to Systems of Equations

**Performance Criterion:**

4. (l) Determine, from given information about the coefficient matrix $A$ and vector $\vec{\mathbf{b}}$ of a system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$, whether the system has any solutions and, if it does, whether there is more than one solution.

You may have found the last section to be a bit overwhelming, and you are probably wondering why we bother with all of the definitions in that section. The reason is that those ideas form tools and language for discussing whether a system of equations

(a) has a solution (meaning at least one) and

(b) if it does have a solution, is there only one.

Item (a) above is what mathematicians often refer to as the *existence* question, and item (b) is the *uniqueness* question. Concerns with "existence and uniqueness" of solutions is not restricted to linear algebra; it is a big deal in the study of differential equations as well.

Consider a system of equations $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$. We saw previously that the product $A\vec{\mathbf{x}}$ is the linear combination of the columns of $A$ with the components of $\vec{\mathbf{x}}$ as the scalars of the linear combination. This means that the system will only have a solution if $\vec{\mathbf{b}}$ is a linear combination of the columns of $A$. But all of the linear combinations of the columns of $A$ is just the span of those columns - the column space! the conclusion of this is as follows:

A system of equations $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a solution (meaning *at least* one solution) if, and only if, $\vec{\mathbf{b}}$ is in the column space of $A$.

Let's look at some consequences of this.

◇ **Example 4.9(a):** Let $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ represent a system of five equations in five unknowns, and suppose that rank$(A) = 3$. Does the system have (for certain) a solution?

**Solution:** Since the system has five equations in five unknowns, $\vec{\mathbf{b}}$ is in $\mathbb{R}^5$. Because $\text{rank}(A) = 3$, the column space of $A$ only has dimension three, so it is not all of $\mathbb{R}^5$ (which of course has dimension five). Therefore $\vec{\mathbf{b}}$ may or may not be in the column space of $A$, and we can't say for certain that the system has a solution.

---

◇ **Example 4.9(b):** Let $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ represent a system of three equations in five unknowns, and suppose that rank$(A) = 3$. Does the system have (for certain) a solution?

**Solution:** Because there are three equations and five unknowns, $A$ is $3 \times 5$ and the columns of $A$ are in $\mathbb{R}^3$. Because $\text{rank}(A) = 3$, the column space must then be all of $\mathbb{R}^3$. Therefore $\vec{\mathbf{b}}$ will be in the column space of $A$ and the system has at least one solution.

---

Now suppose that we have a system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ and a vector $\vec{\mathbf{x}}$ that *IS* a solution to the system. Suppose also that $\text{nullity}(A) \neq 0$. Then there is some $\vec{\mathbf{y}} \neq \vec{\mathbf{0}}$ such that $A\vec{\mathbf{y}} = \vec{\mathbf{0}}$ and

$$A(\vec{\mathbf{x}} + \vec{\mathbf{y}}) = A\vec{\mathbf{x}} + A\vec{\mathbf{y}} = \vec{\mathbf{b}} + \vec{\mathbf{0}} = \vec{\mathbf{b}}\,.$$

This shows that both $\vec{\mathbf{x}}$ and $\vec{\mathbf{x}} + \vec{\mathbf{y}}$ are solutions to the system, so the system does not have a unique solution. The thing that allows this to happen is the fact that the null space of $A$ contains more than just the zero vector. This illustrates the following:

> A system of equations $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ can have a unique solution only if the nullity of $A$ is zero (that is, the null space contains only the zero vector).

Note that this says nothing about whether a system has a solution to begin with; it simply says that if there is a solution and the nullity is zero, then that solution is unique.

◇ **Example 4.9(c):** Consider again a system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ of three equations in five unknowns, with $\text{rank}(A) = 3$, as in Example 4.9(b). We saw in that example that the system has at least one solution - is there a unique solution?

**Solution:** We note first of all that $A$ is $3 \times 5$, so the $n$ of the Rank Theorem is five. We know that $\text{rank}(A)$ is three so, by the Rank Theorem, the nullity is then two. Thus the null space contains more than just the zero vector, so the system does not have a unique solution.

---

◇ **Example 4.9(d):** Suppose we have a system $A\vec{\mathbf{x}} = \vec{\mathbf{0}}$, with $\text{nullity}(A) = 2$. Does the system have a solution and, if it does, is it unique?

**Solution:** Because the system is homogeneous, it has at least one solution, the $zero$ vector. But the null space contains more than just the zero vector, so the system has more than one solution, so there is not a unique solution.

---

## Section 4.9 Exercises

To Solutions

1. Let $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ be a system of equations, with $A$ an $m \times n$ matrix where $m = n$ unless specified otherwise. For each situation below, determine whether the system *COULD* have

   (*i*) no solution   (*ii*) exactly one solution   (*iii*) infinitely many solutions

   **Give all possibilities for each.**

   (a) $\det(A) = 0$   (b) $\det(A) \neq 0$   (c) $\vec{\mathbf{b}} = \vec{\mathbf{0}}$

   (d) $\vec{\mathbf{b}} = \vec{\mathbf{0}}$, $A$ invertible   (e) $m < n$   (f) $m > n$

   (g) columns of $A$ linearly independent   (h) columns of $A$ linearly dependent

(i) $\vec{\mathbf{b}} = \vec{\mathbf{0}}$, columns of $A$ linearly independent

(j) $\vec{\mathbf{b}} = \vec{\mathbf{0}}$, columns of $A$ linearly dependent

2. Consider the system of equations
$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 4 \\ 2x_1 + x_2 - 4x_3 &= 3 \\ -3x_1 + 4x_2 - x_3 &= -2 \end{aligned}.$$

   (a) Give the $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ form of the system.

   (b) Find a basis for the column space of $A$.

   (c) What is $\operatorname{rank}(A)$? Is the system guaranteed to have a solution?

   (d) What is $\operatorname{nullity}(A)$? If the system has a solution, is it unique? That is, is there only one solution?

   (e) Find a basis for the null space of $A$.

   (f) Solve the system if possible. If it isn't, say so.

3. Now consider the system of equations
$$\begin{aligned} x_1 \qquad + 2x_3 &= -5 \\ -2x_1 + 5x_2 \qquad &= 11 \\ 2x_1 + 5x_2 + 8x_3 &= -7 \end{aligned}.$$
Repeat the steps of the previous exercise.

4. Compare the system
$$\begin{aligned} x_1 \qquad + 2x_3 &= -1 \\ -2x_1 + 5x_2 \qquad &= -1 \\ 2x_1 + 5x_2 + 8x_3 &= -5 \end{aligned}$$
with the one in Exercise 3.

   (a) Note that the system has the same $A$ as that from Exercise 3, so the column space, null space rank and nullity are the same as they were in that exercise. Is the vector $\vec{\mathbf{b}}$ of this exercise in $\operatorname{col}(A)$?

   (b) Solve the system.

## 4.10 Chapter 4 Exercises

1. (a) Give a set of three non-zero vectors in $\mathbb{R}^3$ whose span is a line.

   (b) Suppose that you have a set of two non-zero vectors in $\mathbb{R}^3$ that are not scalar multiples of each other. What is their span? How can you create a new vector that is not a scalar multiple of either of the other two vectors but, when added to the set, does not increase the span?

   (c) How many vectors need to be in a set for it to have a chance of spanning all of $\mathbb{R}^3$?

2. Give a set of nonzero vectors $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ in $\mathbb{R}^2$ that **DOES NOT** span $\mathbb{R}^2$. Then give a third vector $\vec{\mathbf{v}}_3$ so that all three vectors **DO** span $\mathbb{R}^2$.

3. Give a set of three vectors, with no one being a scalar multiple of just one other, that span the $xy$-plane in $\mathbb{R}^3$.

4. The things in a set are called *elements*. The union of two sets $A$ and $B$ is a new set $C$ consisting of every element of $A$ along with every element of $B$ and nothing else. (If something is an element of both $A$ and $B$, it is only included in $C$ once.) Every subspace of an $\mathbb{R}^n$ is a subset of that $\mathbb{R}^n$ that possesses some additional special properties. Show that the union of two subspaces is not generally a subspace by giving a specific $\mathbb{R}^n$ and two specific subspaces, then showing that the union is not a subspace.

5. Suppose that the column space of a $3 \times 3$ matrix $A$ has dimension two. What does this tell us about the nature of the solutions to a system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$? Show that the vectors $\vec{\mathbf{u}}_1 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, $\vec{\mathbf{u}}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{\mathbf{u}}_3 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are linearly dependent. Then give $\vec{\mathbf{u}}_2$ as a linear combination of $\vec{\mathbf{u}}_1$ and $\vec{\mathbf{u}}_3$.

6. (a) Give three non-zero linearly dependent vectors in $\mathbb{R}^3$ for which removing any one of the three leaves a linearly independent set.

   (b) Give three non-zero linearly dependent vectors in $\mathbb{R}^3$ for which removing one vector leaves a linearly independent set but removing a different one (of the original three) leaves a linearly dependent set.

7. Consider the vectors $\vec{\mathbf{u}} = \begin{bmatrix} 8 \\ -2 \\ 4 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 7 \\ 0 \\ 1 \end{bmatrix}$ and $\vec{\mathbf{w}} = \begin{bmatrix} -16 \\ 4 \\ -8 \end{bmatrix}$.

   (a) Is the set of all vectors $\vec{\mathbf{x}} = \vec{\mathbf{u}} + t\,\vec{\mathbf{v}}$, where $t$ ranges over all real numbers, a subspace of $\mathbb{R}^3$? If so, give a basis for the subspace; if not, tell why not.

   (b) Is the set of all vectors $\vec{\mathbf{x}} = \vec{\mathbf{u}} + t\,\vec{\mathbf{w}}$, where $t$ ranges over all real numbers, a subspace of $\mathbb{R}^3$? If so, give a basis for the subspace; if not, tell why not.

(c) Is the set of all vectors $\vec{\mathbf{x}} = \vec{\mathbf{u}} + s\,\vec{\mathbf{v}} + t\,\vec{\mathbf{w}}$, where $t$ ranges over all real numbers, a subspace of $\mathbb{R}^3$? If so, give a basis for the subspace; if not, tell why not.

8. Consider the matrix $A = \begin{bmatrix} 2 & 2 & 2 \\ -2 & 5 & 2 \\ 8 & 1 & 4 \end{bmatrix}$.

   (a) Find a basis for row($A$), the row space of $A$. What is the dimension of row($A$)?

   (b) Find a basis for col($A$), the column space of $A$. What is the dimension of col($A$)?

9. Give bases for the null and column spaces of the matrix $A = \begin{bmatrix} -6 & 3 & 30 \\ 2 & -1 & -10 \\ -4 & 2 & 20 \end{bmatrix}$.

10. (a) Give a $3 \times 3$ matrix $B$ for which the column space has dimension one. (**Hint:** What kind of subspace of $\mathbb{R}^3$ has dimension one?)

    (b) Find a basis for the column space of $B$.

    (c) What should the dimension of the null space of $B$ be?

    (d) Find a basis for the null space of $B$.

# 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 $2 \times 2$ or $3 \times 3$ matrix. Use it to find the eigenvalues of the matrix.

(k) Give the eigenspace $E_j$ corresponding to an eigenvalue $\lambda_j$ 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,

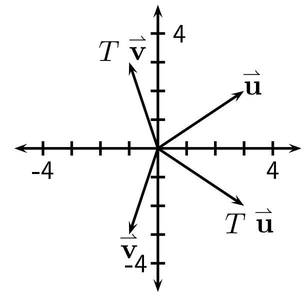
$$\frac{d}{dx}(x^3 - 5x^2 + 2x - 1) = 3x^2 - 10x + 2,$$

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 $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ that reflects vectors across the $x$-axis.

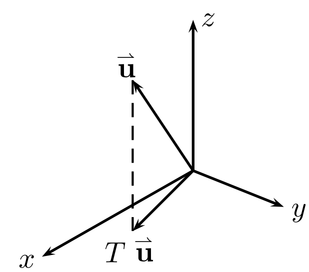
**Solution:** First we might wish to draw a picture to see what such a transformation does to a vector. To the right we see the vectors $\vec{\mathbf{u}}=[3,2]$ and $\vec{\mathbf{v}}=[-1,-3]$, and their transformations $T\,\vec{\mathbf{u}}=[3,-2]$ and $T\,\vec{\mathbf{v}}=[-1,3]$. From these we see that 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 $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ 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 $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ 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, $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$. In general, we can use any $m \times n$ matrix $A$ to define a transformation $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ 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) $T : \mathbb{R}^2 \to \mathbb{R}^2$, $\quad T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1x_2 \\ x_2^2 \end{bmatrix}$, $\quad \vec{\mathbf{u}} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\vec{\mathbf{v}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

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

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

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

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

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

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.

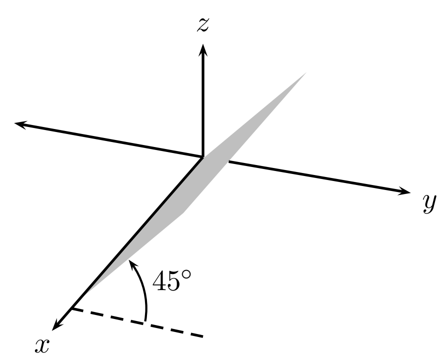
   (i) The transformation that projects every point in $\mathbb{R}^3$ across the $xz$-plane.

   (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 $z = 1$.*

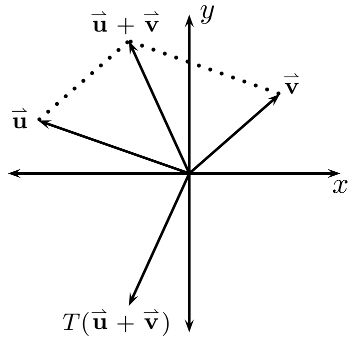
## 5.2 Linear Transformations

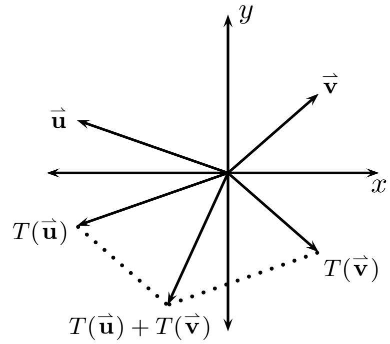
**Performance Criteria:**

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:

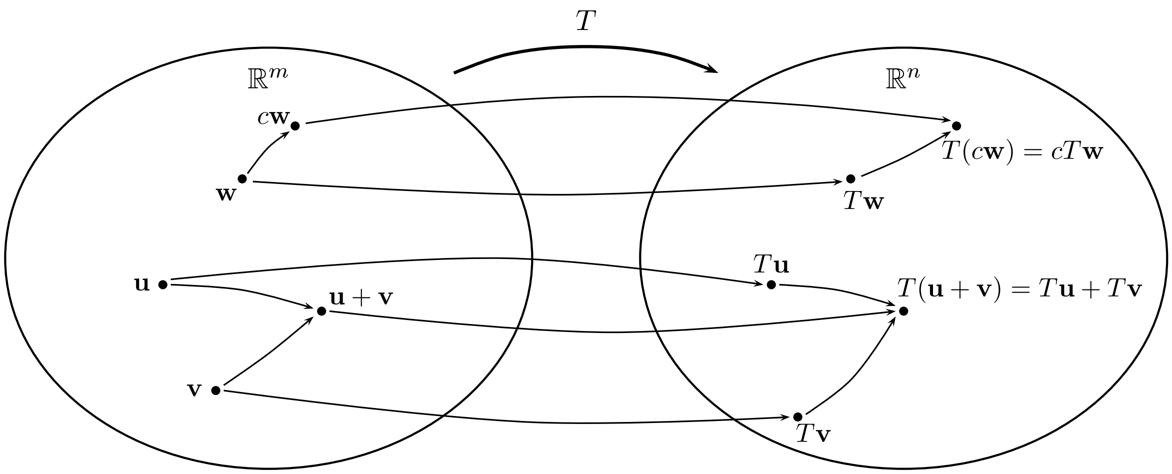
<u>DEFINITION 5.2.1</u>: **Linear Transformation**

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

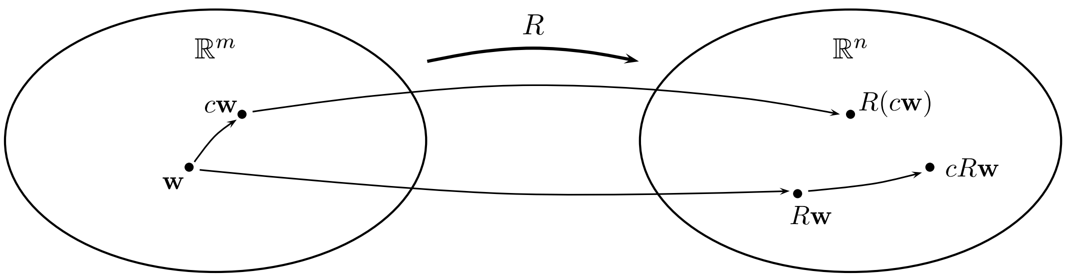
1) $T(\vec{\mathbf{u}}+\vec{\mathbf{v}}) = T(\vec{\mathbf{u}})+T(\vec{\mathbf{v}})$ (**additivity**) and

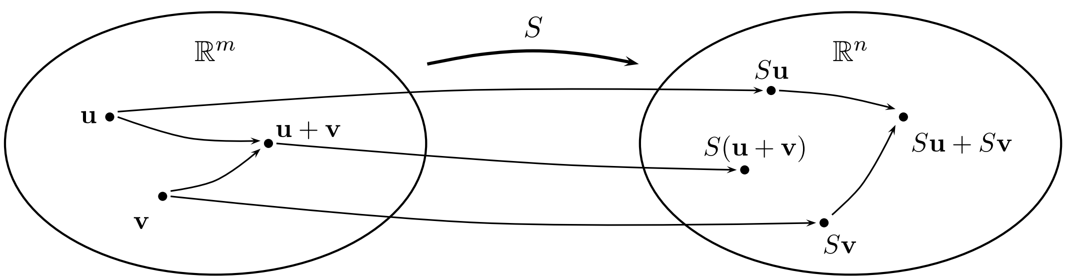
2) $T(c\,\vec{\mathbf{u}}) = c\,T(\vec{\mathbf{u}})$ (**homogeneity**).

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.

◇ **Example 5.2(b):** Is $T:\mathbb{R}^2\rightarrow\mathbb{R}^3$ 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.

◇ **Example 5.2(c):** Determine whether $T:\mathbb{R}^3\rightarrow\mathbb{R}^2$ 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:

> THEOREM 5.2.3: 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

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

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 $T : \mathbb{R}^2 \to \mathbb{R}^2$ 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 $T : \mathbb{R}^2 \to \mathbb{R}^3$ 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

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

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 $\vec{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$,

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

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

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

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

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

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 $AB$.

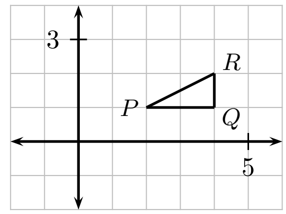
**Solution:** $AB = \begin{bmatrix} 3 & -1 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} 1 & 7 & 3 \\ 4 & -2 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 23 & 6 \\ 22 & 4 & 21 \end{bmatrix}$.

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.

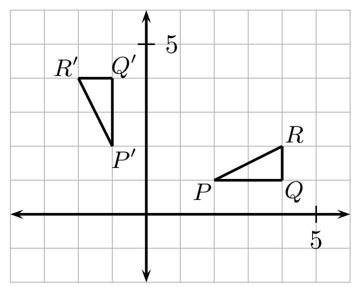



**Solution:** We can represent the triangle $\triangle PQR$ by the matrix $[PQR] = \begin{bmatrix} 2 & 4 & 4 \\ 1 & 1 & 2 \end{bmatrix}$.

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 $\triangle P'Q'R'$.

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

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.

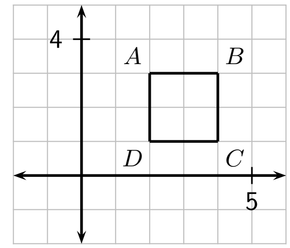
**Solution:** In this case the homogeneous translation matrix is $A=\begin{bmatrix}1&0&-3\\0&1&1\\0&0&1\end{bmatrix}$ and the homogeneous form of $\vec{\mathbf{x}}$ is $\vec{\mathbf{x}}_h=\begin{bmatrix}5\\2\\1\end{bmatrix}$. Multiplying them gives us

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

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 $\theta$.

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

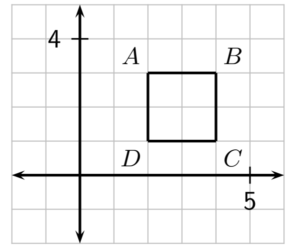
**Solution:** Note that the transformation $T$ shifts every point three units to the left and two units down, so $T^{-1}$ must shift every point three units to the right and two units up. We can determine the homogeneous matrices of these by using the form demonstrated in Example 5.5(e), and the rotation 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.

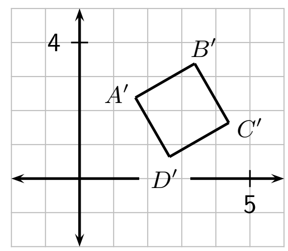



**Solution:** We can represent the square $ABCD$ with a homogenous matrix, shown below and to the left. Each column gives the coordinates of a vertex, with an extry component of one to represent the point in homogeneous coordinates. Converting the entries of the transformation matrix from Example 5.5(g) 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 $\theta$ around the origin we would simply compute the new coordinates $(w, z)$ by

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

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.

- $R_\theta$ will be a rotation of $\theta$, with the rotation being copunterclockwise if $\theta$ is positive, and clockwise if $\theta$ 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 $3 \times 3$ 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) $R_{-90^\circ}$  (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 $(6, -2)$.

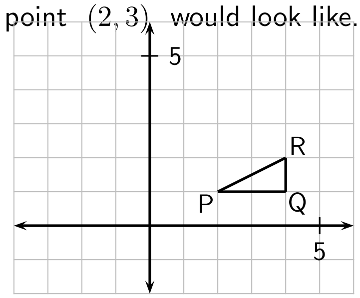
   (d) A reflection across the line $y = x - 3$. (**Hint:** Translate so that the line goes through the origin, reflect, translate back.)

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

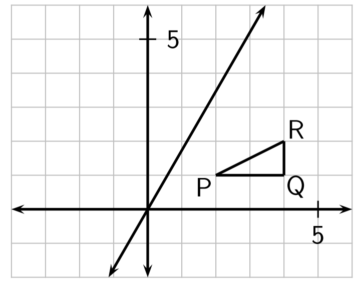
   (b) Consider the triangle $\triangle PQR$ shown to the right below. Sketch what you think the result $\triangle P'Q'R'$ of the rotation $90^\circ$ counterclockwise about the point $(2, 3)$ would look like.

   (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 $\triangle P'Q'R'$.

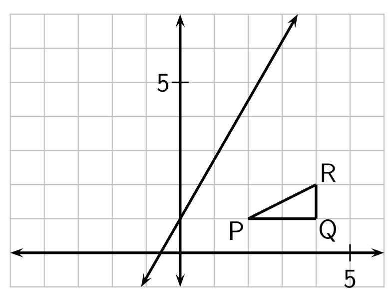
   (c) Draw $\triangle P'Q'R'$ on the graph to the right. If it doesn't look like the reflection of $\triangle PQR$ across the line, find your error and correct it.

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 $C = \begin{bmatrix} \frac{a^2-b^2}{a^2+b^2} & \frac{2ab}{a^2+b^2} \\ \frac{2ab}{a^2+b^2} & \frac{b^2-a^2}{a^2+b^2} \end{bmatrix}$ 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:**

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

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 $\lambda$ 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 $\lambda$.

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 $\lambda$ 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 $2 \times 2$ 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 $\lambda$ to represent eigenvalues. Now suppose that $\vec{\mathbf{x}}$ is an eigenvector of an $n \times n$ matrix $A$, with corresponding eigenvalue $\lambda$, 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 $\lambda$. 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 $\lambda$. $\vec{\mathbf{x}}$, 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.

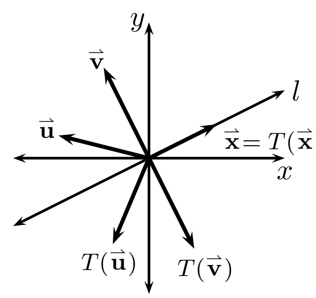
Definition 5.6.3: **Eigenspace Corresponding to an Eigenvalue**

For a given eigenvalue $\lambda_j$ of an $n\times n$ matrix $A$, the **eigenspace** $E_j$ corresponding to $\lambda$ is the set of all eigenvectors corresponding to $\lambda_j$. 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 $\lambda$. 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 $2 \times 2$ or $3 \times 3$ matrix. Use it to find the eigenvalues of the matrix.
   (k) Give the eigenspace $E_j$ corresponding to an eigenvalue $\lambda_j$ 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 $\lambda$ 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 $\lambda$ 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:

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

Taking $\lambda$ 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 $2 \times 2$ Matrix**

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

**Determinant of a $3 \times 3$ Matrix**

To find the determinant of a $3 \times 3$ 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 $\lambda_j$ of an $n \times n$ matrix $A$, the **eigenspace** $E_j$ corresponding to $\lambda_j$ 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

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

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 $\lambda$ to get the eigenvalues.

3) For a given eigenvalue $\lambda_i$, solve the system $(A - \lambda_j I)\vec{\mathbf{x}} = \vec{\mathbf{0}}$. The set of solutions is the eigenspace corresponding to $\lambda_j$. 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.

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

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. $\sigma_x$ and $\sigma_y$ are **normal stresses**, and $\tau_{xy}$ and $\tau_{yx}$ are **shear stresses**. The fact that the element is not rotating gives us that $\tau_{xy} = \tau_{yx}$.

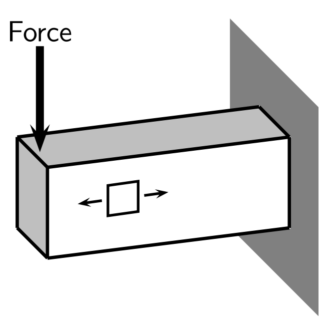



Figure 5.7(a)

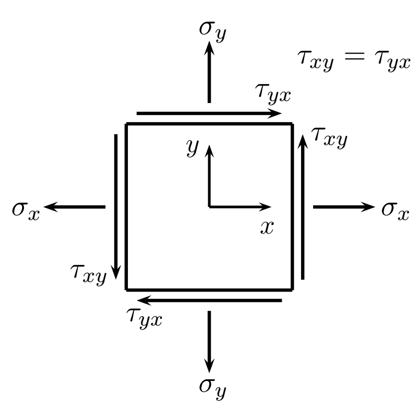



Figure 5.7(b)

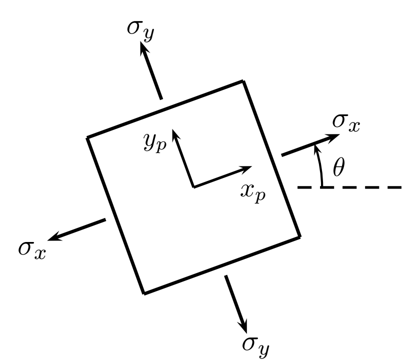



Figure 5.7(c)

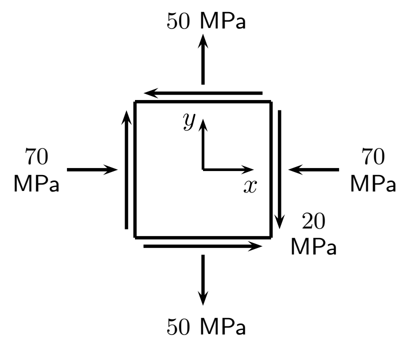
We can rotate the stress element about its center and the stresses will all change as we do that.

There is one angle of rotation we are particularly interested in. If we rotate by a particular angle $\theta$ 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

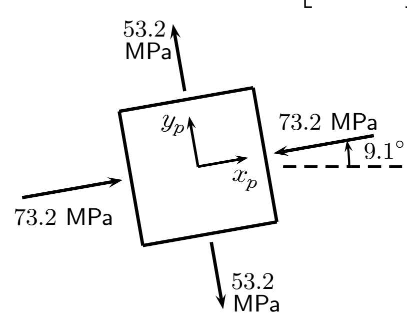
$$\begin{bmatrix} 3.2 & -20 & 0 \\ -20 & 126.2 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -126.2 & -20 & 0 \\ -20 & -3.2 & 0 \end{bmatrix}.$$

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

$$\begin{bmatrix} 3.2 & -20 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -126.2 & -20 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

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 $\theta = \tan^{-1}\dfrac{3.2}{20} = 9.1°$. (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 $\lambda_1$. 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 $\lambda$ 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.

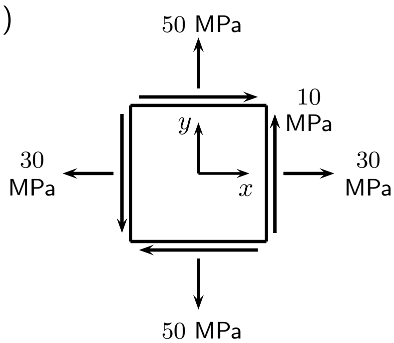
10. So far, every $n \times n$ matrix we have worked with has had $n$ linearly independent eigenvectors, perhaps with fewer eigenvalues (Exercises 7 and 8). The matrix

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

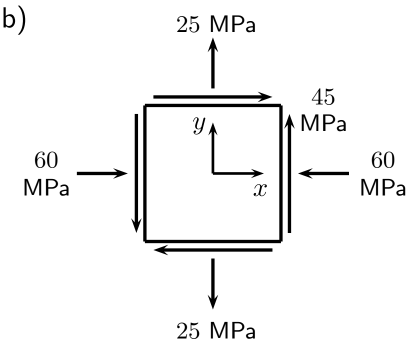
has only two eigenvalues and only two independent eigenvectors. Find the eigenvalues and eigenvectors.

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).

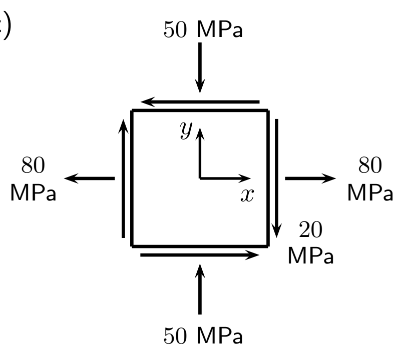
(a)



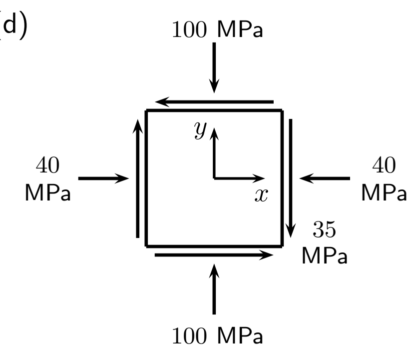

(b)




(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 $k$th power of a $2\times 2$ matrix $A$ with eigenvalues $\lambda_1$ and $\lambda_2$ 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 $k$th 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.

# A Index of Symbols


| | | |
|---|---|---|
| $\mathbb{R}$ | real number | 7 |
| *rref* | reduced row echelon form | 10 |
| $\mathbb{R}^2$, $\mathbb{R}^3$ | two and three dimensional Euclidean space | 33, 34 |
| $a, b, c, x, y, z$ | scalars | 37 |
| $\vec{\mathbf{u}}$, $\vec{\mathbf{v}}\vec{\mathbf{w}}, \vec{\mathbf{x}}$ | vectors | 37 |
| $\overrightarrow{OP}$ | position vector from the origin to point $P$ | 37 |
| $\overrightarrow{PQ}$ | vector from point $P$ to point $Q$ | 37 |
| $\Vert \vec{\mathbf{v}} \Vert$ | magnitude (or length) of vector $\vec{\mathbf{v}}$ | 38 |
| $\vec{\mathbf{u}} \cdot \vec{\mathbf{v}}$ | dot product of $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ | 56 |
| $\text{proj}_{\mathbf{v}}\, \vec{\mathbf{u}}$ | projection of $\vec{\mathbf{u}}$ onto $\vec{\mathbf{v}}$ | 57, 58 |
| $A$, $B$ | matrices | 64 |
| $A^T$ | transpose of a matrix | 64 |
| $A(i, j)$ | $i$th row, $j$th column of matrix $A$ | 64 |
| $I_n$ | $n \times n$ identity matrix | 64 |
| $m \times n$ | dimensions of a matrix | 64, 65 |
| $a_{ij}$ | $i, j$th entry of matrix $A$ | 65 |
| $\vec{\mathbf{a}}_{i*}$, $\vec{\mathbf{a}}_{*j}$ | $i$th row and $j$th column of matrix $A$ | 66 |
| $A\,\vec{\mathbf{u}}$ | matrix $A$ times vector $u$ | 66 |
| $AB$ | product of matrices $A$ and $B$ | 77 |
| $A^{-1}$ | inverse of the matrix $A$ | 81 |
| $L$ | lower triangular matrix | 91 |
| $U$ | upper triangular matrix | 91 |
| $\det(A)$ | determinant of matrix $A$ | 95 |
| $\mathcal{S}$ | a finite set of vectors | 104 |
| $\text{col}(A)$ | column space of $A$ | 116 |
| $\text{null}(A)$ | null space of $A$ | 117 |
| $\bar{\vec{\mathbf{x}}}$ | least squares solution | 118 |
| $\vec{\mathbf{e}}$ | error vector for least squares approximation | 120 |
| $\Vert \vec{\mathbf{e}} \Vert$ | error for least squares approximation | 120 |
| $\vec{\mathbf{e}}_j$ | $j$th standard bassi vector | 132 |
| $\mathcal{B}$ | basis | 132 |
| $T$, $S$ | transformations | 146 |
| $[T]$ | matrix for the linear transformation $T$ | 152 |
| $S \circ T$ | composition of transformations $S$ and $T$ | 154, 155 |
| $\lambda$, $\lambda_j$ | eigenvalues | 160 |
| $E_j$ | eigenspace corresponding to eigenvalue $\lambda_j$ | 164 |
| $P^{-1}AP$ | diagonalization of matrix $A$ | 166 |

# B Solutions to Exercises

## B.1 Chapter 1 Solutions

**Section 1.1 Solutions** Back to 1.1 Exercises

1. (b), (c) and (d) are linear equations

2. The first and second systems are linear, the third is not.

3. (a) $(-2,-3,2)$ and $(7,3,1)$ are solutions, $(5,-2,4)$ is not.
   (b) Only $(-1,3,2)$ is a solution.
   (c) $(2,1)$ and $(-1,-2)$ are solutions, $(3,5)$ is not.

4. (a) $5=-8a+4b-2c+d$, this is a linear equation
   (b) $y=7x^3-2x^2-5x+1$, this is not a linear equation

5.
$$\begin{aligned} -8a+4b-2c+d &= 5\\ -a+b-c+d &= 2\\ a+b+c+d &= 3\\ 27a+9b+3c+d &= 0 \end{aligned}$$

6.
$$\begin{aligned} 1.2m+b &= 3.7\\ 2.5m+b &= 4.1\\ 3.2m+b &= 4.7\\ 4.3m+b &= 5.2\\ 5.1m+b &= 5.9 \end{aligned}$$

7. (a)
$$\begin{aligned} f_1-f_2 &= 5\\ f_2-f_3 &= -1\\ f_3-f_4 &= 4\\ -f_1+f_4 &= -8 \end{aligned}$$

7. (b) $(12,7,8,4)$, $(7,2,3,-1)$ and $(10,5,6,2)$ are solutions, $(9,4,3,1)$ is not
   (c) The solution $(7,2,3,-1)$ indicates a net flow of one vehicle *in the wrong direction* between nodes $n_3$ and $n_4$.
   (d) $f_4=5$, $f_1=13$ and $f_2=8$

**Section 1.2 Solutions** Back to 1.2 Exercises

1. (a) $a-b+c-d=3$

   (b)
$$\begin{aligned} a+b+c+d &= 5\\ a+2b+4c+8d &= 4\\ a+4b+16c+64d &= -1 \end{aligned}$$

2.
$$\begin{aligned} a-b+c &= -4\\ a+b+c &= 1\\ 9a+3b+c &= 0 \end{aligned}$$

3. (a) $4t_1-t_2-t_3=129$

   (b)
$$\begin{aligned} t_2 &= \frac{t_1+65+59+t4}{4} &\implies\quad -t_1+4t_2-t_4 &= 129\\ t_3 &= \frac{t_1+t_4+53+55}{4} &\implies\quad -t_1+4t_3-t_4 &= 108\\ t_4 &= \frac{t_2+52+50+t_3}{4} &\implies\quad -t_2-t_3+4t_4 &= 102 \end{aligned}$$

4. (a)
$$\begin{aligned} 10I_1+2(I_1-I_2)+8(I_1-I_3) &= 20 &\implies\quad 20I_1-2I_2-8I_3 &= 20\\ 6I_2+1(I_2-I_3)+2(I_2-I_1) &= 0 &\implies\quad -2I_1+9I_2-I_3 &= 0\\ 5I_3+8(I_3-I_2)+1(I_3-I_2) &= 0 &\implies\quad -8I_1-I_2+14I_3 &= 0 \end{aligned}$$

(b) 
$$\begin{aligned} 12I_1 + 2(I_1 - I_2) + 5(I_1 - I_3) &= 12 & \implies && 19I_1 - 2I_2 - 5I_3 &= 12 \\ 8I_2 + 3(I_2 - I_3) + 2(I_2 - I_1) &= 8 & \implies && -2I_1 + 13I_2 - 3I_3 &= 8 \\ 4I_3 + 5(I_3 - I_2) + 3(I_3 - I_2) &= 0 & \implies && -5I_1 - 3I_2 + 12I_3 &= 0 \end{aligned}$$

(c) The current from $A$ to $C$ is $I_1 - I_2$ because $I_1$ is from $A$ to $C$ and $I_2$ is from $C$ to $A$. Thus the current from $A$ to $C$ is $1.36 - 0.39 = 0.97$ amperes. The current flows from $A$ to $C$ because this value is positive.

(d) The current from $B$ to $C$ is $I_2 - I_3$ because $I_2$ is from $B$ to $C$ and $I_3$ is from $C$ to $B$. Thus the current from $B$ to $C$ is $0.39 - 0.81 = -0.42$ amperes. The current flows from $C$ to $B$ because this value is negative.

**Section 1.3 Solutions** Back to 1.3 Exercises

1. (a) $(-2, 1)$ (b) $(3, 4)$ (c) $(3, 2)$
   (d) $(1, -1)$ (e) $(-1, 4)$ (f) $(-2, \frac{1}{3})$

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

**Section 1.4 Solutions** Back to 1.4 Exercises

1. The coefficient matrix is $\begin{bmatrix} 1 & 1 & -3 \\ -3 & 2 & -1 \\ 2 & 1 & -4 \end{bmatrix}$ and the augmented matrix is $\begin{bmatrix} 1 & 1 & -3 & 1 \\ -3 & 2 & -1 & 7 \\ 2 & 1 & -4 & 0 \end{bmatrix}$

2. All the matrices but $D$ are in row-echelon form.

3. Matrices $B$ and $F$ are in reduced row-echelon form.

4. $$\begin{bmatrix} 1 & 1 & -3 & 1 \\ -3 & 2 & -1 & 7 \\ 2 & 1 & -4 & 0 \end{bmatrix} \quad \begin{array}{c} 3R_1 + R_2 \to R_2 \\ \Longrightarrow \\ -2R_1 + R_3 \to R_3 \end{array} \quad \begin{bmatrix} 1 & 1 & -3 & 1 \\ 0 & 5 & -10 & 10 \\ 0 & -1 & 2 & -2 \end{bmatrix}$$

5. (a) $$\begin{bmatrix} 1 & 5 & -7 & 3 \\ -5 & 3 & -1 & 0 \\ 4 & 0 & 8 & -1 \end{bmatrix} \quad \begin{array}{c} 5R_1 + R_2 \to R_2 \\ \Longrightarrow \\ -4R_1 + R_3 \to R_3 \end{array} \quad \begin{bmatrix} 1 & 5 & -7 & 3 \\ 0 & 28 & -36 & 15 \\ 0 & -20 & 36 & -13 \end{bmatrix}$$

   (b) $$\begin{bmatrix} 2 & -8 & -1 & 5 \\ 0 & -2 & 0 & 0 \\ 0 & 6 & -5 & 2 \end{bmatrix} \quad \begin{array}{c} \Longrightarrow \\ 3R_2 + R_3 \to R_3 \end{array} \quad \begin{bmatrix} 2 & -8 & -1 & 5 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -5 & 2 \end{bmatrix}$$

   (c) $$\begin{bmatrix} 1 & -2 & 4 & 1 \\ 0 & 3 & 5 & -2 \\ 0 & 2 & -8 & 1 \end{bmatrix} \quad \begin{array}{c} \Longrightarrow \\ -\frac{2}{3}R_2 + R_3 \to R_3 \end{array} \quad \begin{bmatrix} 1 & -2 & 4 & 1 \\ 0 & 3 & 5 & -2 \\ 0 & 0 & -\frac{34}{3} & \frac{7}{3} \end{bmatrix}$$

6. (a) $(-4, \frac{1}{2}, -4)$ (b) $(33, -4, 1)$ (c) $(7, 0, -2)$

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

8. Same as Exercise 7.

9. $t_1 = 52.6°$, $t_2 = 57.3°$, $t_3 = 61.6°$, $t_4 = 53.9°$, $t_5 = 57.1°$, $t_6 = 60.2°$

12. $t_1 = 44.6°$, $t_2 = 49.6°$, $t_3 = 38.6°$

13. (a) fourth degree (b) $y = \frac{2}{5}x + \frac{23}{30}x^2 - \frac{1}{10}x^3 - \frac{1}{15}x^4$ or $y = 0.4x + 0.77x^2 - 0.1x^3 - 0.07x^4$

14. (a) $I_1 = 4.5$ amperes, $I_2 = 0.75$ amperes (b) $I_3 = 5.25$ amperes

17. $z = 3.765 + 0.353x - 1.235y$

**Section 1.5 Solutions** Back to 1.5 Exercises

1. (a) $(6,3,4)$, $(0,0,4)$, $(-2,-1,4)$, $(2,1,4)$
   (b) $x - 2y = 0$, $z = 4$, we can determine $z$ (c) $(2t,t,4)$

2. (a) $z$ is a free variable, $x$ and $y$ are leading variables
   (b) $y = 2t + 2$, $x = t - 1$

3. (a) $(4+t, -5-2t, t, 3)$, $(4,-5,0,3)$, $(5,-7,1,3)$, $(6,-9,2,3)$, $(3,-3,-1,3), ...$
   (b) $(5-3t, t, 1, -4)$, $(5,0,1,-4)$, $(2,1,1,-4)$, $(-1,2,1,-4)$, $(8,-1,1,-4), ...$
   (c) $(-4+3s-t, s, 5+2t, t)$, $(-4,0,5,0)$, $(-1,1,5,0)$, $(-5,0,7,1)$, $(-2,1,7,1), ...$
   (d) no solution
   (e) $(6+2s-t, -3-3s-5t, s, t)$, $(6,-3,0,0)$, $(6,-3,0,0)$, $(8,-6,1,0)$, $(5,-8,0,1)$, $(7,-11,1,1), ...$
   (f) $(-1,2,0)$
   (g) $(1-2s+t, s, t)$, $(1,0,0)$, $(-1,1,0)$, $(2,0,1)$, $(0,1,1), ...$
   (h) no solution
   (i) $(2-5t, t, 1, -4)$, $(2,0,1,-4)$, $(-3,1,1,-4)$, $(-8,2,1,-4)$, $(7,-1,1,-4), ...$

4. $t = -2$: $(5,-8,-2,4)$, $t = -1$: $(2,-3,-1,4)$, $t = 0$: $(-1,2,0,4)$, $t = 1$: $(-4,7,1,4)$, $t = 2$: $(-7,12,2,4)$

5. There is no solution to the system with the first reduced matrix given. The system with the second reduced matrix has general solution $(-2s+t+5, s, t, -4)$ and some particular solutions of

$$s = t = 0 : (5,0,0,-4), \qquad s = 1,\ t = 0 : (3,1,0,-4), \qquad s = 0,\ t = 1 : (6,0,1,-4)$$

## B.2 Chapter 2 Solutions

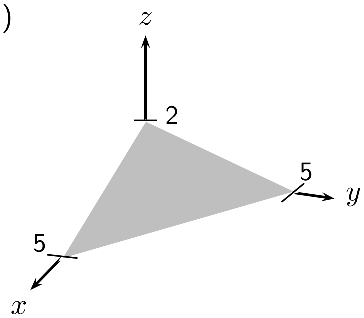
**Section 2.1 Solutions** Back to 2.1 Exercises

1. (a) Plane, intersects the $x$, $y$ and $z$-axes at $(3,0,0)$, $(0,6,0)$ and $(0,0,-2)$, respectively.
   (b) Plane, intersects the $x$ and $z$-axes at $(6,0,0)$ and $(0,0,2)$, respectively. Does not intersect the $y$-axis.
   (c) Plane, intersects only the $y$-axis, at $(0,-6,0)$.
   (d) Not a plane.
   (e) Plane, intersects the $x$, $y$ and $z$-axes at $(-6,0,0)$, $(0,3,0)$ and $(0,0,-2)$, respectively.

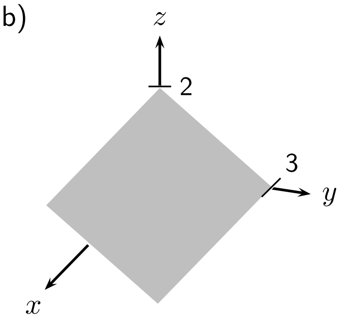
2. (a) $10x + 4y + 5z = 20$ (b) $-7x + 21y + 3z = 21$

(c) $2x - 3z = 6$ (d) $y = 4$ (e) $-y + 4z = 4$ (f) $z = -2$

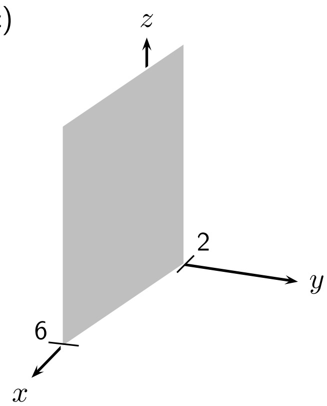
3. (a)



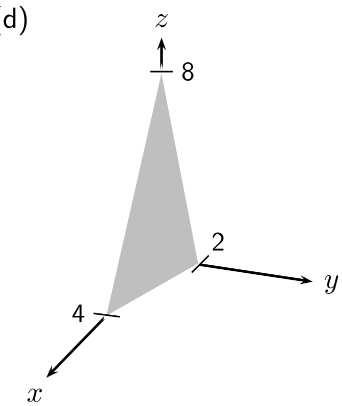

(b)



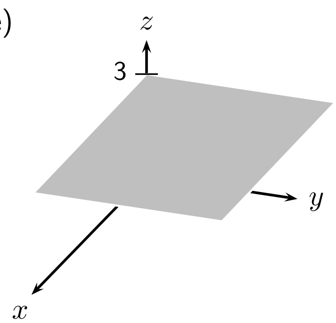

(c)



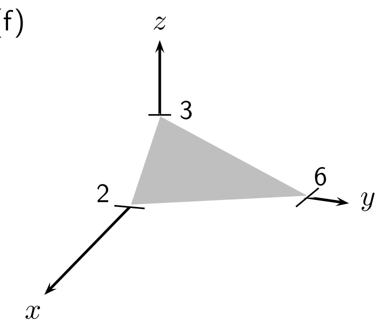

(d)



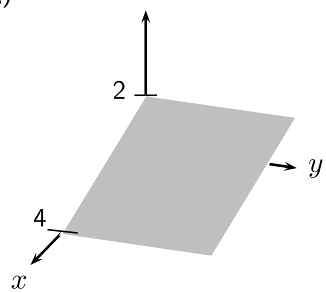

(e)



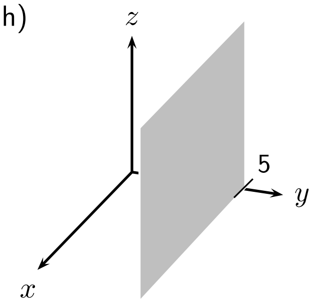

(f)



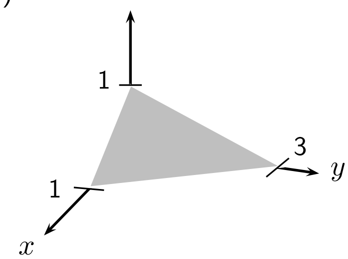
(g)




(h)




(i)




**Section 2.2 Solutions**

Back to 2.2 Exercises

1. $\| \vec{\mathbf{u}} \| = \sqrt{14}$ , $\quad \| \vec{\mathbf{x}} \| = \sqrt{74}$ , $\quad \| \vec{\mathbf{v}} \| = \sqrt{30}$ , $\quad \| \vec{\mathbf{b}} \| = \sqrt{b_1^2 + b_2^2 + b_3^2 + b_4^2}$

2. (a) $\overrightarrow{PQ} = \begin{bmatrix} 17 \\ -6 \\ -15 \end{bmatrix}$, $\|\overrightarrow{PQ}\| = \sqrt{550} = 23.4$

   (b) $\overrightarrow{PQ} = \begin{bmatrix} 12 \\ -3 \end{bmatrix}$, $\|\overrightarrow{PQ}\| = \sqrt{153} = 12.4$

   (c) $\overrightarrow{PQ} = \begin{bmatrix} 10 \\ -1 \\ -7 \\ 9 \end{bmatrix}$, $\|\overrightarrow{PQ}\| = \sqrt{231} = 15.2$

3. (a) $Q(5,4)$ (b) $Q(6,7,-1,-7,7)$ (c) $P(9,-8,5)$

4. (a) $\| \vec{\mathbf{u}} \| = 3$ (b) $\dfrac{\vec{\mathbf{u}}}{\| \vec{\mathbf{u}} \|} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$ (c) $\left\| \dfrac{\vec{\mathbf{u}}}{\| \vec{\mathbf{u}} \|} \right\| = 1$

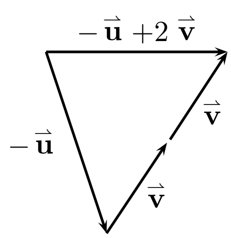
5. (a) $\| \vec{\mathbf{v}} \| = 5$ (b) $\dfrac{\vec{\mathbf{v}}}{\| \vec{\mathbf{v}} \|} = \begin{bmatrix} \frac{4}{5} \\ -\frac{3}{5} \end{bmatrix}$ (c) $\left\| \dfrac{\vec{\mathbf{v}}}{\| \vec{\mathbf{v}} \|} \right\| = 1$

   For any vector $\vec{\mathbf{v}}$, the magnitude of $\dfrac{\vec{\mathbf{v}}}{\| \vec{\mathbf{v}} \|}$ is always one.

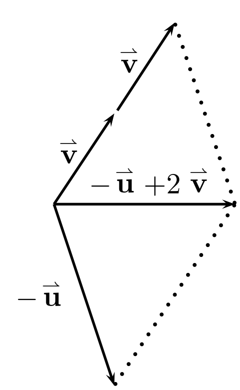
**Section 2.3 Solutions**

Back to 2.3 Exercises

1. Tip-to-tail:




Parallelogram:




2. $\begin{bmatrix} -45 \\ 30 \end{bmatrix}$

3. $\begin{bmatrix} -c_1 - 8c_2 \\ 3c_1 + c_2 \\ -6c_1 + 4c_2 \end{bmatrix}$

5. (a) $0\vec{\mathbf{u}}_1 + 2\vec{\mathbf{u}}_2 + 1\vec{\mathbf{u}}_3 = \vec{\mathbf{v}}$ (b) $2\vec{\mathbf{u}}_1 - 3\vec{\mathbf{u}}_2 = \vec{\mathbf{v}}$

   (c) $-5\vec{\mathbf{u}}_1 + \vec{\mathbf{u}}_2 + 3\vec{\mathbf{u}}_3 = \vec{\mathbf{v}}$

   (d) Any vector of the form $(-\frac{1}{2}t + \frac{7}{2})\vec{\mathbf{u}}_1 + (\frac{1}{2}t - \frac{5}{2})\vec{\mathbf{u}}_2 + t\vec{\mathbf{u}}_3$ equals $\vec{\mathbf{v}}$

   (e) $2\vec{\mathbf{u}}_1 - \vec{\mathbf{u}}_2 + 4\vec{\mathbf{u}}_3 - 5\vec{\mathbf{u}}_4 = \vec{\mathbf{v}}$

   (f) There is no linear combination of $\vec{\mathbf{u}}_1$, $\vec{\mathbf{u}}_2$ and $\vec{\mathbf{u}}_3$ equalling $\vec{\mathbf{v}}$.

   (g) Any vector of the form $(1-2t)\vec{\mathbf{u}}_1 + (4-3t)\vec{\mathbf{u}}_2 + t\vec{\mathbf{u}}_3$ equals $\vec{\mathbf{v}}$

6. (a) $a_1 = \frac{3}{2}, \quad a_2 = 4, \quad a_3 = -\frac{1}{2}$ (b) There are no such $b_1$, $b_2$ and $b_3$

   (c) $\det(A) = -42, \quad \det(B) = 0$

**Section 2.4 Solutions**

Back to 2.4 Exercises

1. (a) $$x\begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} + y\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + z\begin{bmatrix} -3 \\ -1 \\ -4 \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \\ 0 \end{bmatrix}$$

   (b) $$x_1\begin{bmatrix} 5 \\ 0 \\ 2 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} + x_3\begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$$

   (c) $$b\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + m\begin{bmatrix} 0.5 \\ 1.0 \\ 1.5 \\ 2.0 \\ 2.5 \\ 3.0 \\ 3.5 \end{bmatrix} = \begin{bmatrix} 8.1 \\ 6.9 \\ 6.2 \\ 5.3 \\ 4.5 \\ 3.8 \\ 3.0 \end{bmatrix}$$

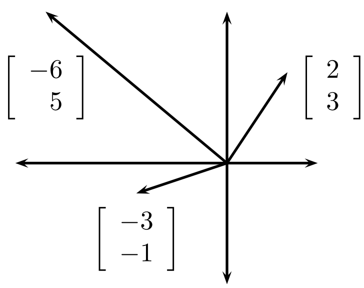
   (d) $$x_1\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + x_2\begin{bmatrix} -4 \\ 2 \\ 1 \end{bmatrix} + x_3\begin{bmatrix} 1 \\ -1 \\ -4 \end{bmatrix} + x_4\begin{bmatrix} 2 \\ -7 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$$

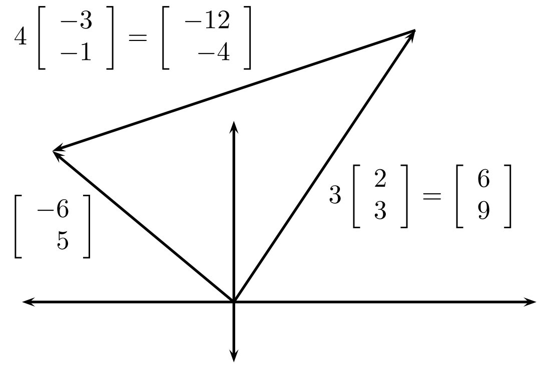
2. (a) $$\begin{aligned} -3x_1 + x_2 &= 5 \\ x_1 + x_2 &= -2 \end{aligned}$$

   (b) $$\begin{aligned} 5x_1 - 3x_2 + 2x_3 + 7x_4 &= -8 \\ x_1 + 2x_2 \phantom{+ 5x_3} - 4x_4 &= 1 \\ -4x_1 + x_2 + 5x_3 + 6x_4 &= 5 \\ -3x_1 - x_2 + 4x_3 + 7x_4 &= 4 \end{aligned}$$

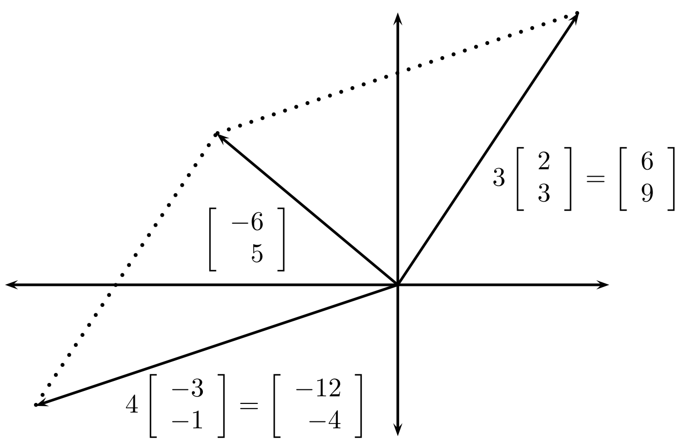
   (c) $$\begin{aligned} 3x_1 + 4x_2 + x_3 &= -5 \\ 2x_1 - 7x_2 + x_3 &= 3 \\ x_1 + 5x_2 + x_3 &= -4 \end{aligned}$$

3. $x\begin{bmatrix} 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} -3 \\ -1 \end{bmatrix} = \begin{bmatrix} -6 \\ 5 \end{bmatrix} \quad \Longrightarrow \quad 3\begin{bmatrix} 2 \\ 3 \end{bmatrix} + 4\begin{bmatrix} -3 \\ -1 \end{bmatrix} = \begin{bmatrix} -6 \\ 5 \end{bmatrix}$




tip-to-tail method




parallelogram method

**Section 2.5 Solutions** Back to 2.5 Exercises

1. (a) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 9 \end{bmatrix}, \begin{bmatrix} -2 \\ 4 \end{bmatrix}, \begin{bmatrix} \pi \\ \pi^2 \end{bmatrix}, \ \cdots$

   (b) The set is not closed under addition. (c) $\mathcal{S}$ is not closed under scalar multiplication.

2. (a) $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ -7 \\ -2 \end{bmatrix}, \begin{bmatrix} -3 \\ -1 \\ -4 \end{bmatrix}, \begin{bmatrix} \pi \\ e \\ \pi + e \end{bmatrix}, \ \cdots$

   (b) The set is closed under addition. (c) The set is closed under scalar multiplication.

3. (a) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -2 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ -5 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \ \cdots$

   (b) $\mathcal{S}$ is not closed under addition. (c) $\mathcal{S}$ is closed under scalar multiplication.

4. (a) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -4 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \end{bmatrix}, \begin{bmatrix} -3 \\ 6 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \ \cdots$

(b) $\mathcal{S}$ is closed under scalar multiplication. (c) The set is closed under addition.

5. (a) $\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ -3 \end{bmatrix}, \begin{bmatrix} 6 \\ -5 \end{bmatrix}, \begin{bmatrix} 0 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \dots$

   (b) $\mathcal{S}$ is not closed under addition. (c) $\mathcal{S}$ is not closed under scalar multiplication.

6. (a) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ -3 \end{bmatrix}, \begin{bmatrix} 17 \\ -22 \end{bmatrix}, \begin{bmatrix} -10 \\ -10 \end{bmatrix}, \dots$

   (b) The set is closed under addition. (c) $\mathcal{S}$ is closed under scalar multiplication.

7. (a) $\begin{bmatrix} -4 \\ -2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -6 \\ 3 \end{bmatrix}, \dots$

   (b) The set is closed under addition. (c) $\mathcal{S}$ is closed under scalar multiplication.

8. (a) $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -8 \\ 0 \\ 8 \end{bmatrix}, \dots$

   (b) The set is not closed under addition. (c) The set is not closed under scalar multiplication.

9. (a) $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \\ -6 \end{bmatrix}, \begin{bmatrix} \pi \\ 2\pi \\ 3\pi \end{bmatrix}, \dots$

   (b) $\mathcal{S}$ is closed under addition. (c) $\mathcal{S}$ is closed under scalar multiplication.

10. (a) $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 200 \\ -300 \end{bmatrix}, \begin{bmatrix} 1 \\ \pi \\ e \end{bmatrix}, \begin{bmatrix} 1 \\ -7 \\ 1 \end{bmatrix}, \dots$

    (b) The set is not closed under addition. (c) The set is not closed under scalar multiplication.

**Section 2.6 Solutions** Back to 2.6 Exercises

1. Each of the following is just one possibility - each line or plane has more than one equation possible.

   (a) $\vec{\mathbf{x}} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} + t \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

   (b) $\vec{\mathbf{x}} = \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix} + s \begin{bmatrix} -1 \\ 7 \\ -4 \end{bmatrix} + t \begin{bmatrix} -4 \\ 1 \\ -1 \end{bmatrix}$

   (c) $\vec{\mathbf{x}} = \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix} + t \begin{bmatrix} -1 \\ 7 \\ -4 \end{bmatrix}$

   (d) $\vec{\mathbf{x}} = \begin{bmatrix} -1 \\ 4 \end{bmatrix} + t \begin{bmatrix} 3 \\ -1 \end{bmatrix}$

(e) $\vec{\mathbf{x}} = t\begin{bmatrix} -4 \\ 3 \end{bmatrix}$ (f) $\vec{\mathbf{x}} = \begin{bmatrix} -5 \\ 1 \\ 3 \end{bmatrix} + s\begin{bmatrix} 7 \\ -1 \\ 1 \end{bmatrix} + t\begin{bmatrix} 6 \\ -3 \\ -1 \end{bmatrix}$

(g) $\vec{\mathbf{x}} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} + t\begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix}$ (h) $\vec{\mathbf{x}} = s\begin{bmatrix} -4 \\ 5 \\ 1 \end{bmatrix} + t\begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$

(i) $\vec{\mathbf{x}} = t\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

2. The sets described in parts (e), (h) and (i) are closed under addition and scalar multiplication.

3. $\vec{\mathbf{x}} = \begin{bmatrix} -2 \\ 1 \\ 1 \\ 5 \end{bmatrix} + r\begin{bmatrix} 5 \\ 6 \\ -6 \\ -3 \end{bmatrix} + s\begin{bmatrix} 6 \\ -1 \\ 4 \\ -7 \end{bmatrix} + t\begin{bmatrix} 5 \\ -6 \\ 3 \\ 2 \end{bmatrix}$

4. (a) $\vec{\mathbf{x}} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + t\begin{bmatrix} 4 \\ -2 \\ 2 \end{bmatrix}$ (b) $\vec{\mathbf{x}} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + t\begin{bmatrix} -4 \\ 2 \\ -2 \end{bmatrix}$

(c) $\vec{\mathbf{x}} = \begin{bmatrix} 5 \\ -1 \\ 4 \end{bmatrix} + t\begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$

5. (a) $(7, 9)$ (b) $\vec{\mathbf{x}} = \begin{bmatrix} -3 \\ 1 \end{bmatrix} + t\begin{bmatrix} 5 \\ 4 \end{bmatrix}$ (c) $t = 2$ (d) $(-8, -3),\ t = -1$

6. $\vec{\mathbf{x}} = \begin{bmatrix} -5 \\ 7 \\ 4 \end{bmatrix} + s\begin{bmatrix} 6 \\ -1 \\ -6 \end{bmatrix} + t\begin{bmatrix} 8 \\ -7 \\ -5 \end{bmatrix}$ 7. $(-3, 3, 0),\ (-7, 5, -2),\ (-11, 7, -4)$

8. The vector equation of the plane is $\vec{\mathbf{x}} = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} + s\begin{bmatrix} 5 \\ 1 \\ -4 \end{bmatrix} + t\begin{bmatrix} 6 \\ -3 \\ -8 \end{bmatrix}$. Letting $s = 1$ and $t = 1$ gives the point $(9, -1, -7)$. (This is just one possibility - we can get other points by choosing other values of $s$ and $t$.)

9. (a) If either, but not both, of $\vec{\mathbf{u}}$ or $\vec{\mathbf{v}}$ are the zero vector, then $\vec{\mathbf{x}} = \vec{\mathbf{p}} + s\vec{\mathbf{u}} + t\vec{\mathbf{v}}$ will be the equation of a line. If $\vec{\mathbf{u}} \neq \vec{\mathbf{0}}$ and $\vec{\mathbf{v}} \neq \vec{\mathbf{0}}$ are scalar multiples of each other, then $\vec{\mathbf{x}} = \vec{\mathbf{p}} + s\vec{\mathbf{u}} + t\vec{\mathbf{v}}$ will be the equation of a line.

(b) If $\vec{\mathbf{p}}$ is the zero vector, then $\vec{\mathbf{x}} = \vec{\mathbf{p}} + s\vec{\mathbf{u}} + t\vec{\mathbf{v}}$ will be the equation of a plane through the origin. If $\vec{\mathbf{p}} \neq \vec{\mathbf{0}}$ is a scalar multiple of either $\vec{\mathbf{u}} \neq \vec{\mathbf{0}}$ or $\vec{\mathbf{v}} \neq \vec{\mathbf{0}}$ and $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are *NOT* scalar multiples of each other, then $\vec{\mathbf{x}} = \vec{\mathbf{p}} + s\vec{\mathbf{u}} + t\vec{\mathbf{v}}$ will be the equation of a plane through the origin.

## Section 2.7 Solutions

Back to 2.7 Exercises

1. (a) $\vec{\mathbf{x}} = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ (b) $\vec{\mathbf{x}} = \begin{bmatrix} 5 \\ 0 \\ -1 \end{bmatrix} + t\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

   (c) $\vec{\mathbf{x}} = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$

2. (a) three-dimensional plane in four-dimensional space
   (b) three-dimensional plane in five-dimensional space
   (c) five-dimensional plane in seven-dimensional space
   (d) one-dimensional plane (line) in five-dimensional space
   (e) two-dimensional plane in four-dimensional space
   (f) four-dimensional plane in seven-dimensional space

## Section 2.8 Solutions

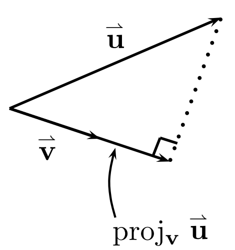
Back to 2.8 Exercises

2. (a) $\text{proj}_{\mathbf{b}}\,\vec{\mathbf{v}} = \dfrac{15-2}{25+4}\begin{bmatrix} 5 \\ -2 \end{bmatrix} = \dfrac{13}{29}\begin{bmatrix} 5 \\ -2 \end{bmatrix} = \begin{bmatrix} \frac{65}{29} \\ -\frac{26}{29} \end{bmatrix}$

   (b) $\text{proj}_{\mathbf{b}}\,\vec{\mathbf{v}} = \dfrac{10-0}{4+1}\begin{bmatrix} -2 \\ 1 \end{bmatrix} = 2\begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$

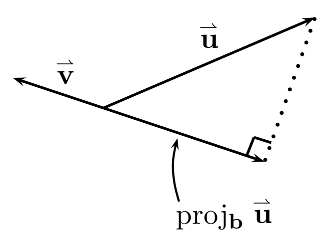
   (c) $\text{proj}_{\mathbf{b}}\,\vec{\mathbf{v}} = \dfrac{-6-4}{16+4}\begin{bmatrix} -2 \\ -4 \end{bmatrix} = -\dfrac{1}{2}\begin{bmatrix} -2 \\ -4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

3. (a)

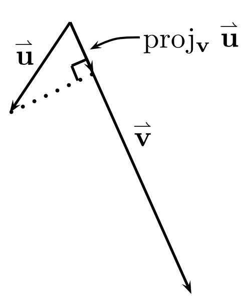



(tail is where $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ meet)

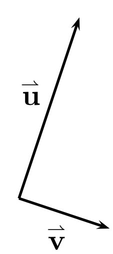
(b)




(c)




(d)




because $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are orthogonal, $\text{proj}_{\mathbf{v}}\,\vec{\mathbf{u}}$ is the zero vector

## B.3 Chapter 3 Solutions

**Section 3.1 Solutions**

Back to 3.1 Exercises

1. (a) $A$ is $3 \times 3$, $B$ is $3 \times 2$, $C$ is $3 \times 4$

   (b) $b_{31} = 4$, $c_{23} = 2$

2. (a) all but $B$, $F$ (b) $C$, $D$, $G$, $J$ (c) $C$ (d) $C$, $E$ (e) $A$, $C$

3. The following are *possible* answers:

   (a) $\begin{bmatrix} 2 & 0 & 0 \\ -5 & 1 & 0 \\ 7 & 3 & -8 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -8 \end{bmatrix}$ (c) $\begin{bmatrix} 2 & -5 & 7 \\ -5 & 1 & 3 \\ 7 & 3 & -8 \end{bmatrix}$

   (d) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (e) $\begin{bmatrix} 2 & -1 & 5 \\ 0 & 4 & 3 \\ 0 & 0 & -8 \end{bmatrix}$ (f) see (c)

   (g) see (c) (h) see (b)

4. $A^T = \begin{bmatrix} 1 & -3 & 4 \\ 0 & 1 & 7 \\ 5 & -2 & 0 \end{bmatrix}$ $B^T = \begin{bmatrix} 1 & -3 & 4 \\ 0 & 1 & 7 \end{bmatrix}$ $C^T = \begin{bmatrix} 1 & -3 & 4 \\ 0 & 1 & 7 \\ -1 & 2 & 0 \\ 3 & 0 & -2 \end{bmatrix}$

5. $B + D = \begin{bmatrix} 2 & -3 \\ -2 & 3 \\ 5 & 11 \end{bmatrix}$, $B - D = \begin{bmatrix} 0 & 3 \\ -4 & -1 \\ 3 & 3 \end{bmatrix}$, $D - B = \begin{bmatrix} 0 & -3 \\ 4 & 1 \\ -3 & -3 \end{bmatrix}$

6. (a) The matrix is square and symmetric. (b) The matrix is square and symmetric.

   (c) The matrix is square, symmetric, diagonal and both upper and lower triangular.

   (d) The matrices are square, symmetric, diagonal and both upper and lower triangular. They are also called **zero matrices**.

7. (a) $B = \begin{bmatrix} 2 & 5 \\ 5 & 8 \end{bmatrix}$ (b) $B$ is a symmetric matrix.

**Section 3.2 Solutions**

Back to 3.2 Exercises

1. $\begin{bmatrix} 3 \\ -12 \\ 9 \end{bmatrix}$, $\begin{bmatrix} -23 \\ 28 \end{bmatrix}$

2. (a) $\begin{bmatrix} -13 \\ -18 \end{bmatrix}$ (b) $\begin{bmatrix} -5 \\ 18 \end{bmatrix}$ (c) not possible

(d) $\begin{bmatrix} -3 \\ 1 \\ -28 \end{bmatrix}$ (e) $\begin{bmatrix} -22 \\ 4 \\ 0 \\ 27 \end{bmatrix}$ (f) $\begin{bmatrix} 33 \\ 0 \\ -14 \end{bmatrix}$

3. (a) $-4\begin{bmatrix} 3 \\ 5 \end{bmatrix} + 1\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ (b) $1\begin{bmatrix} 1 \\ 6 \end{bmatrix} + 0\begin{bmatrix} -5 \\ 3 \end{bmatrix} - 3\begin{bmatrix} 2 \\ -4 \end{bmatrix}$

(c) not possible (d) $3\begin{bmatrix} 1 \\ -5 \\ 2 \end{bmatrix} + 5\begin{bmatrix} 6 \\ 3 \\ -4 \end{bmatrix}$

(e) $-4\begin{bmatrix} 7 \\ 1 \\ 2 \\ -3 \end{bmatrix} + 1\begin{bmatrix} -2 \\ 5 \\ 1 \\ 7 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 3 \\ -1 \\ 2 \end{bmatrix} + 2\begin{bmatrix} 4 \\ -3 \\ 5 \\ 1 \end{bmatrix}$

(f) $2\begin{bmatrix} 1 \\ 2 \\ -4 \end{bmatrix} - 3\begin{bmatrix} 0 \\ 2 \\ 7 \end{bmatrix} + 1\begin{bmatrix} -5 \\ 3 \\ 1 \end{bmatrix}$

4. (a) $\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{bmatrix}$ (b) $x_1\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + x_2\begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} + x_3\begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}$

5. (a) $A = \begin{bmatrix} 3 & -5 \\ 1 & 1 \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & 3 \\ 2 & -1 \\ 5 & 4 \end{bmatrix}$

(c) $C = \begin{bmatrix} 2 & 4 & -1 \\ -5 & 1 & 2 \\ 1 & 3 & 0 \end{bmatrix}$ (d) $D = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix}$

6. (a) $\begin{bmatrix} 1 & 1 & 3 \\ -3 & 2 & -1 \\ 2 & 1 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \\ 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 & 3 \\ -3 & 2 & -1 \\ 2 & 1 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \\ 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0.5 \\ 1 & 0.5 \\ 1 & 0.5 \\ 1 & 0.5 \\ 1 & 0.5 \\ 1 & 0.5 \\ 1 & 0.5 \end{bmatrix}\begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} 8.1 \\ 6.9 \\ 6.2 \\ 5.3 \\ 4.5 \\ 3.8 \\ 3.0 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & -4 & 1 & 2 \\ 3 & 2 & -1 & -7 \\ -2 & 1 & -4 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$

7. (a) $\vec{\mathbf{x}} = \begin{bmatrix} -5 \\ 2 \\ 4 \end{bmatrix}$ (b) $\vec{\mathbf{x}} = \begin{bmatrix} 3 \\ -7 \end{bmatrix}$ (c) $\vec{\mathbf{x}} = \begin{bmatrix} 1.7 \\ 0.4 \end{bmatrix}$ (d) $\vec{\mathbf{x}} = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$

8. (a) $-\frac{\nu\sigma_{xx}}{E} + \frac{\sigma_{yy}}{E} - \frac{\nu\sigma_{zz}}{E} = \epsilon_{yy}$ (b) $-\frac{\nu\sigma_{xx}}{E} - \frac{\nu\sigma_{yy}}{E} + \frac{\sigma_{zz}}{E} = \epsilon_{zz}$ (c) $\frac{\tau_{zx}}{\gamma_{zx}}$

(d) $$\begin{bmatrix} \frac{1}{E} & -\frac{\nu}{E} & -\frac{\nu}{E} \\ -\frac{\nu}{E} & \frac{1}{E} & -\frac{\nu}{E} \\ -\frac{\nu}{E} & -\frac{\nu}{E} & \frac{1}{E} \end{bmatrix} \begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \end{bmatrix} = \begin{bmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{G} & 0 & 0 \\ 0 & \frac{1}{G} & 0 \\ 0 & 0 & \frac{1}{G} \end{bmatrix} \begin{bmatrix} \tau_{xy} \\ \tau_{yz} \\ \tau_{zx} \end{bmatrix} = \begin{bmatrix} \gamma_{xy} \\ \gamma_{yz} \\ \gamma_{zx} \end{bmatrix}$$

**Section 3.4 Solutions** Back to 3.4 Exercises

1. (a) $\begin{bmatrix} 3 & 3 \\ 8 & -7 \end{bmatrix}$ (b) $\begin{bmatrix} 4 & -7 & -3 \\ 18 & -31 & -5 \\ -7 & 16 & 9 \end{bmatrix}$

2. $A^2 = \begin{bmatrix} -15 & 5 \\ -3 & -14 \end{bmatrix}$ $AF = \begin{bmatrix} 30 & 45 \\ 0 & 12 \end{bmatrix}$ $BC = \begin{bmatrix} -69 \\ -25 \end{bmatrix}$

$BD = \begin{bmatrix} 62 & -25 & -2 \\ -121 & 80 & 36 \end{bmatrix}$ $CE = \begin{bmatrix} -25 & 5 & -10 \\ 20 & -4 & 8 \\ -35 & 7 & -14 \end{bmatrix}$ $DC = \begin{bmatrix} -51 \\ 27 \\ -1 \end{bmatrix}$

$D^2 = \begin{bmatrix} 39 & 3 & 18 \\ -48 & 18 & -7 \\ 1 & 4 & 5 \end{bmatrix}$ $EC = \begin{bmatrix} -43 \end{bmatrix}$ $ED = \begin{bmatrix} 37 & -2 & 13 \end{bmatrix}$

$FA = \begin{bmatrix} 3 & 9 \\ -27 & 39 \end{bmatrix}$ $FB = \begin{bmatrix} 10 & -35 & 10 \\ 6 & 147 & -18 \end{bmatrix}$ $F^2 = \begin{bmatrix} -2 & -11 \\ 66 & 75 \end{bmatrix}$

3. The $(2,1)$ entry is $-8$ and the $(3,2)$ entry is $-13$.

4. $a_{31}a_{12} + a_{32}a_{22} + a_{33}a_{32} + a_{34}a_{42} + a_{35}a_{52}$

5. $$P^2 = \begin{bmatrix} \frac{16}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{9}{25} \end{bmatrix} \begin{bmatrix} \frac{16}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{9}{25} \end{bmatrix} = \left(\frac{1}{25}\right)\left(\frac{1}{25}\right) \begin{bmatrix} 16 & 12 \\ 12 & 9 \end{bmatrix} \begin{bmatrix} 16 & 12 \\ 12 & 9 \end{bmatrix} =$$

$$\left(\frac{1}{625}\right) \begin{bmatrix} 400 & 300 \\ 300 & 225 \end{bmatrix} = \begin{bmatrix} \frac{16}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{9}{25} \end{bmatrix}$$

6. (a) $A^T = \begin{bmatrix} -5 & 0 & 2 \\ 1 & 4 & -3 \end{bmatrix}$ (b) $A^TA = \begin{bmatrix} 29 & -11 \\ -11 & 26 \end{bmatrix}$, $AA^T = \begin{bmatrix} 26 & 4 & -13 \\ 4 & 16 & -12 \\ -13 & -12 & 13 \end{bmatrix}$

(c) $A^TA$ and $AA^T$ are both square, symmetric matrices

**Section 3.5 Solutions** Back to 3.5 Exercises

1. Because $\begin{bmatrix} 2 & 5 \\ 3 & 8 \end{bmatrix}\begin{bmatrix} 8 & -4 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 4 \end{bmatrix} \neq I_2$, the matrices are not inverses.

2. $AB = I_2$ but $A$ and $B$ are not inverses because (1) neither is square and (2) $BA \neq I$

3. (a) $[I_2 \mid B] = \begin{bmatrix} 1 & 0 & -\frac{5}{2} & \frac{3}{2} \\ 0 & 1 & 2 & -1 \end{bmatrix}$

(b) $AB = BA = I_2$ (c) $B$ is the inverse of $A$

5. (a) $\begin{bmatrix} 2 & -3 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ (b) $A^{-1} = \dfrac{1}{10-(-12)}\begin{bmatrix} 5 & 3 \\ -4 & 2 \end{bmatrix} = \frac{1}{22}\begin{bmatrix} 5 & 3 \\ -4 & 2 \end{bmatrix}$

(c) $\frac{1}{22}\begin{bmatrix} 5 & 3 \\ -4 & 2 \end{bmatrix}\begin{bmatrix} 2 & -3 \\ 4 & 5 \end{bmatrix} = \frac{1}{22}\begin{bmatrix} 22 & 0 \\ 0 & 22 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(d)

$$\begin{aligned}
\begin{bmatrix} 2 & -3 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 4 \\ 3 \end{bmatrix} \\
\frac{1}{22}\begin{bmatrix} 5 & 3 \\ -4 & 2 \end{bmatrix}\left(\begin{bmatrix} 2 & -3 \\ 4 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= \frac{1}{22}\begin{bmatrix} 5 & 3 \\ -4 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \end{bmatrix} \\
\left(\frac{1}{22}\begin{bmatrix} 5 & 3 \\ -4 & 2 \end{bmatrix}\begin{bmatrix} 2 & -3 \\ 4 & 5 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \frac{1}{22}\begin{bmatrix} 5 & 3 \\ -4 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \end{bmatrix} \\
\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \frac{1}{22}\begin{bmatrix} 29 \\ -10 \end{bmatrix} \\
\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} \frac{29}{22} \\ -\frac{10}{22} \end{bmatrix}
\end{aligned}$$

(e)

$$\begin{bmatrix} 2 & -3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{bmatrix}$$
$$\begin{bmatrix} 2 & -3 & 1 & 0 \\ 0 & 11 & -2 & 1 \end{bmatrix}$$
$$\begin{bmatrix} 2 & -3 & 1 & 0 \\ 0 & 1 & -\frac{2}{11} & \frac{1}{11} \end{bmatrix}$$
$$\begin{bmatrix} 2 & 0 & \frac{5}{11} & \frac{3}{11} \\ 0 & 1 & -\frac{2}{11} & \frac{1}{11} \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 & \frac{5}{22} & \frac{3}{22} \\ 0 & 1 & -\frac{2}{11} & \frac{1}{11} \end{bmatrix}$$

6. (a) $\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$ (b) $A^{-1} = \dfrac{1}{(5)(3)-(2)(7)}\begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix} = \begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix}$

(c) $\begin{bmatrix} 5 & 7 & 1 & 0 \\ 2 & 3 & 0 & 1 \end{bmatrix} \overset{rref}{\Longrightarrow} \begin{bmatrix} 1 & 0 & 3 & -7 \\ 0 & 1 & -2 & 5 \end{bmatrix}$, so $A^{-1} = \begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix}$

(d)

$$\begin{aligned}
\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} -1 \\ 4 \end{bmatrix} \\
\begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix}\left(\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}\right) &= \begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} -1 \\ 4 \end{bmatrix} \\
\left(\begin{bmatrix} 3 & -7 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}\right)\begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} -31 \\ 22 \end{bmatrix} \\
\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} -31 \\ 22 \end{bmatrix} \\
\begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} -31 \\ 22 \end{bmatrix}
\end{aligned}$$

## Section 3.6 Solutions

Back to 3.6 Exercises

1. (a) $\det(A) = 4$ (b) $\det(B) = 10$ (c) $\det(C) = 8$

   (d) $\det(A) = 5$ (e) $\det(B) = 1$ (f) $\det(C) = 0$

   (e) $\det(A) = 2$ (f) $\det(B) = -1$

2. (a) $\lambda^2 - 5\lambda$ (b) $\lambda^2 - 4\lambda + 3$

   (c) $-\lambda^3 + 4\lambda^2 - 3\lambda$ (d) $-\lambda^3 + 6\lambda^2 + 15\lambda + 8$

3. The determinant of the coefficient matrix is zero, so the system *DOES NOT* have a unique solution.

4. (a) If the determinant of $A$ is zero, then the system has no solution or infinitely many solutions.
   (b) If the determinant of $A$ is not zero, then the system has a unique solution.

5. (a) If the determinant of $A$ is zero, then the system has infinitely many solutions. (It can't have no solutions, because $\vec{\mathbf{x}} = \mathbf{0}$ is a solution.
   (b) If the determinant of $A$ is not zero, then the system has the unique solution $\vec{\mathbf{x}} = \mathbf{0}$.

## Section 3.7 Solutions

Back to 3.7 Exercises

1. (b) $P\vec{\mathbf{u}} = \vec{\mathbf{u}}$ (c) $P\vec{\mathbf{v}} = \vec{\mathbf{0}}$ (d) $P(P\vec{\mathbf{w}}) = P^2\vec{\mathbf{w}} = P\vec{\mathbf{w}}$, so $P^2 = P$
   (e) $P^n = P$ for $n = 1, 2, 3, 4, \dots$ This says that once we have projected a vector, any additional projecting of the result is just what we got on the first projection.
   (f) (Infinitely) many different vectors project to the same vector, so if we know the result of a projection we cannot determine the original vector. Therefore $P^{-1}$ does not exist for a projection matrix $P$.

2. (b) $C\vec{\mathbf{u}} = \vec{\mathbf{u}}$ (c) $C\vec{\mathbf{v}} = -\vec{\mathbf{v}}$ (d) $C(C\vec{\mathbf{w}}) = C^2\vec{\mathbf{w}} = \vec{\mathbf{w}}$, so $C^2 = I$
   (e) $C^n = I$ if $n$ is even, and $C^n = C$ if $n$ is odd.
   (f) Yes, we can determine $\vec{\mathbf{x}}$ by simply applying $C$ to $C\vec{\mathbf{x}}$. $C$ is invertible and, in fact, $C$ is its own inverse!

5. (a) $R_\theta^2 = R_{2\theta}$ (b) $R_\theta^{-1} = R_{-\theta}$ (c) $R_{120°}^3 = R_{360°} = I$
   (d) $R_{180°}^{-1} = R_{180°}$ This holds for any multiple of $180°$.

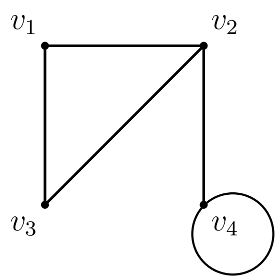
6. (a) $v_3v_2v_4,\ v_3v_1v_4$ (b) $v_1v_3v_1v_4,\ v_1v_3v_2v_4,\ v_1v_4v_1v_4,\ v_1v_4v_2v_4,\ v_1v_4v_3v_4$

   (c) $A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$ $A^2 = \begin{bmatrix} 2 & 2 & 1 & 1 \\ 2 & 3 & 2 & 2 \\ 1 & 2 & 3 & 2 \\ 1 & 2 & 2 & 3 \end{bmatrix}$ $A^3 = \begin{bmatrix} 2 & 4 & 5 & 5 \\ 4 & 7 & 7 & 7 \\ 5 & 7 & 5 & 6 \\ 5 & 7 & 6 & 5 \end{bmatrix}$

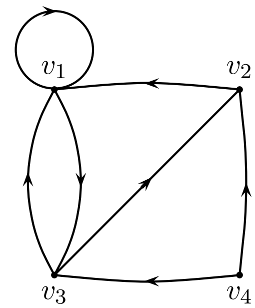
   (d) The $(3,4)$ and $(4,3)$ entries of $A^2$ are both two, the number of 2-paths from $v_3$ to $v_4$.
   (e) The $(1,4)$ and $(4,1)$ entries of $A^3$ are both five, the number of 3-paths from $v_1$ to $v_4$.

(e) There are seven 3-paths from $v_2$ to $v_3$. They are $v_1v_2v_2v_3$, $v_2v_2v_4v_3$, $v_2v_3v_1v_3$, $v_2v_3v_2v_3$, $v_2v_3v_4v_3$, $v_2v_4v_1v_3$, $v_2v_4v_2v_3$

7. (a)




(b)




8. (a)

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \qquad A^3 = \begin{bmatrix} 5 & 7 & 5 & 6 \\ 7 & 7 & 4 & 7 \\ 5 & 4 & 2 & 5 \\ 6 & 7 & 5 & 5 \end{bmatrix}$$

(b) We expect six 3-paths from $v_1$ to $v_4$. they are

$v_1v_2v_1v_4$
$v_1v_2v_2v_4$
$v_1v_4v_1v_4$
$v_1v_4v_2v_4$
$v_1v_4v_3v_4$
$v_1v_3v_1v_4$

(c) There are six 3-paths from $v_4$ to $v_1$. This is indicated by the fact that the matrix $A^3$ is symmetric.

9. (a)

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

(b) There are eight 4-paths from $v_1$ to $v_3$, as indicated by the $(1, 3)$ entry of

$$A^4 = \begin{bmatrix} 8 & 8 & 8 \\ 5 & 6 & 5 \\ 3 & 2 & 3 \end{bmatrix}$$

(c) $v_2v_3v_2v_1v_3$
$v_2v_1v_2v_1v_3$
$v_2v_1v_1v_1v_3$
$v_2v_1v_1v_2v_3$
$v_2v_1v_3v_2v_3$

(d) $v_3v_2v_3v_2$
$v_3v_2v_1v_2$
No, there are two 3-paths from $v_3$ to $v_2$, but three 3-paths from $v_2$ to $v_3$.

## B.4 Chapter 4 Solutions

**Section 4.1 Solutions** Back to 4.1 Exercises

1. (a) The span of the set is the $x$-axis.

   (b) The span of the set is the $xz$-plane, or the plane $y = 0$.

   (c) The span of the set is all of $\mathbb{R}^2$.

   (d) The span of the set is the origin.

   (e) The span of the set is a line through the origin and the point $(1, 2, 3)$.

2. (a) $\vec{\mathbf{w}}$ is not in the span of $S$. (b) $\vec{\mathbf{w}}$ is not in the span of $S$.

   (c) $\begin{bmatrix} 8 \\ 38 \\ -14 \\ 11 \end{bmatrix} = -2 \begin{bmatrix} 1 \\ -4 \\ -3 \\ 7 \end{bmatrix} + 5 \begin{bmatrix} 2 \\ 6 \\ -4 \\ 5 \end{bmatrix}$ (d) $\begin{bmatrix} 3 \\ 7 \\ -4 \end{bmatrix} = -4 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 11 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - 4 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

3. (a) Because the vectors in $\mathcal{S}_1$ are not scalar multiples of each other, they span all of $\mathbb{R}^2$. The vectors in $\mathcal{S}_2$ are scalar multiples of each other, so their span is just a line in $\mathbb{R}^2$, and the spans of the two sets are not the same.

   (b) The span of each set is a line, but the lines spanned are not the same, so the spans of the two sets are different.

   (c) The two sets each span the same line. We can tell this because the vector in $\mathcal{S}_2$ is $-3$ times the vector in $\mathcal{S}_1$.

   (d) The two vectors in each set are not scalar multiples of each other, so both sets span all of $\mathbb{R}^2$ and their spans are equal.

   (e) We see that $\mathcal{S}_2$ is $\mathcal{S}_1$ with the additional vector $\begin{bmatrix} 2 \\ -4 \\ 1 \end{bmatrix}$ included. We can find that

   $$\begin{bmatrix} 2 \\ -4 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 3 \\ -1 \\ 4 \end{bmatrix} - \begin{bmatrix} 4 \\ 2 \\ 7 \end{bmatrix},$$

   so the third vector in $\mathcal{S}_2$ is in the span of $\mathcal{S}_1$, so the spans of the two sets are the same.

   (f) This time $\mathcal{S}_2$ is $\mathcal{S}_1$ with the additional vector $\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ included. We can find that this vector is not in the span of $\mathcal{S}_1$, so the spans of the two sets are not the same in this case.

**Section 4.2 Solutions** Back to 4.2 Exercises

1. (a) $\mathcal{S}$ is not closed under either addition or scalar multiplication.

   (b) $\mathcal{S}$ is closed under both addition and scalar multiplication.

   (c) $\mathcal{S}$ is not closed under either addition or scalar multiplication.

   (d) $\mathcal{S}$ is closed under both addition and scalar multiplication.

   (e) $\mathcal{S}$ is not closed under either addition or scalar multiplication.

2. (a) The set is closed under both addition and scalar multiplication, and is spanned by $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$.

(b) The set is closed under both addition and scalar multiplication, and is spanned by $\mathcal{S} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

(c) The set is not closed under either addition or scalar multiplication.

(d) The set is closed under both addition and scalar multiplication, and is spanned by $\mathcal{S} = \left\{ \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix} \right\}$.

(e) The set is closed under both addition and scalar multiplication, and is spanned by $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$.

(f) The set is not closed under either addition or scalar multiplication.

**Section 4.3 Solutions**

Back to 4.3 Exercises

1. (a) Not a subspace, doesn't contain the zero vector.

(b) Subspace. (c) Subspace.

(d) Not a subspace, the vector $\begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix} + \begin{bmatrix} 5 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 6 \end{bmatrix}$ is not on either line because it is not a scalar multiple of either vector.

(e) Not a subspace, the vector $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is in the set, but $-2\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ -6 \end{bmatrix}$ is not.

(f) The set is a plane not containing the zero vector, so it is not a subspace.

(g) This is a plane containing the origin, so it is a subspace.

(h) The vector $\mathbf{0}$ is a subspace. (i) Subspace.

2. (a) The set is not a subspace because it does not contain the zero vector. We can tell this because $\vec{\mathbf{u}}$ and $\vec{\mathbf{v}}$ are not scalar multiples of each other.

(b) The set is a subspace, and either of the vectors $\vec{\mathbf{u}}$ or $\vec{\mathbf{w}}$ by itself is a basis, as is any scalar multiple of either of them.

**Section 4.4 Solutions** Back to 4.4 Exercises

1. (a) $\vec{\mathbf{u}}$ is in the column space of $A$.
   (b) $\vec{\mathbf{u}}$ is not in the column space of $A$.
   (c) $\vec{\mathbf{u}}$ is in the column space of $A$.
   (d) $\vec{\mathbf{u}}$ is in the column space of $A$.
   (e) $\vec{\mathbf{u}}$ is not in the column space of $A$.
   (f) $\vec{\mathbf{u}}$ is in the column space of $A$.

2. (a) $c_1\begin{bmatrix}1\\2\\-1\end{bmatrix}+c_2\begin{bmatrix}1\\3\\-4\end{bmatrix}+c_3\begin{bmatrix}0\\-2\\6\end{bmatrix}=\begin{bmatrix}2\\-9\\17\end{bmatrix}$, $\quad\begin{bmatrix}1&1&0&2\\2&3&-2&-9\\-1&-4&6&17\end{bmatrix}$

   $\vec{\mathbf{u}}_1$ is in the column space.

   (b) $-3\begin{bmatrix}1\\2\\-1\end{bmatrix}+5\begin{bmatrix}1\\3\\-4\end{bmatrix}=\begin{bmatrix}2\\-9\\17\end{bmatrix}$.

   (c) $c_1\begin{bmatrix}1\\2\\-1\end{bmatrix}+c_2\begin{bmatrix}1\\3\\-4\end{bmatrix}+c_3\begin{bmatrix}0\\-2\\6\end{bmatrix}=\begin{bmatrix}3\\15\\2\end{bmatrix}$, $\quad\begin{bmatrix}1&1&0&3\\2&3&-2&15\\-1&-4&6&2\end{bmatrix}$

   $\vec{\mathbf{u}}_2$ is not in the column space.

3. $\vec{\mathbf{v}}_1$ is in null($A$) and $\vec{\mathbf{v}}_2$ is not.

4. (a) $\vec{\mathbf{u}}_1$ is in $\text{col}(A)$. For example $\begin{bmatrix}3\\1\\-1\end{bmatrix}-\begin{bmatrix}-2\\3\\-4\end{bmatrix}=\begin{bmatrix}5\\-2\\3\end{bmatrix}$.

   (b) $\vec{\mathbf{u}}_2$ is not in $\text{col}(A)$ because there is no linear combination of the columns of $A$ that equals $\vec{\mathbf{u}}_2$.

   (c) We know that $A\vec{\mathbf{x}}=\vec{\mathbf{u}}_2$ has no solution, because of Theorem 4.4.2.

   (d) There are many possibilities - one example is $\begin{bmatrix}1\\1\\-1\end{bmatrix}$.

   (e) $A$ is not invertible. If it were, there would be a solution to $A\vec{\mathbf{x}}=\vec{\mathbf{b}}$ for all $\vec{\mathbf{b}}$ in $\mathbb{R}^3$, which we know is not the case.

6. (a) $\mathbb{R}^2$ (b) $\mathbb{R}^3$

   (c) It is clear that the first two columns of $A$ span all of $\mathbb{R}^2$, so there are no vectors in $\mathbb{R}^2$ not in the column space of $A$.

   (d) $\begin{bmatrix}1\\1\\1\end{bmatrix}$ is in the null space of $A$. There are, of course, other vectors in $\mathbb{R}^3$ that are in $\text{null}(A)$.

**Section 4.5 Solutions** Back to 4.5 Exercises

1. (a) $y = 10.5 - 2x$
   (b) The line seems to do a pretty good job of coming close to all of the points.
   (c) $y = 8 + 0.5x - 0.5x^2$
   (d) The parabola goes through all of the points.

2. $z = 2.95 + 1.79x - 0.85y$

**Section 4.6 Solutions** Back to 4.6 Exercises

1. (a) The set is linearly independent because the vectors are not scalar multiples of each other.
   (b) A set of one vector $\vec{\mathbf{v}}$ is linearly independent because the only solution to $c\,\vec{\mathbf{v}} = \vec{\mathbf{0}}$ is $c = 0$.
   (c) The set is linearly dependent because the second vector is the sum of the first and third.
   (d) The set is linearly dependent because the vectors are scalar multiples of each other. We can see this because both are scalar multiples of $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$.
   (e) The set is linearly dependent because it consists of more than three vectors in $\mathbb{R}^3$.

2. (a) $c_1\,\vec{\mathbf{u}}_1 + c_2\,\vec{\mathbf{u}}_2 + c_3\,\vec{\mathbf{u}}_3 = \vec{\mathbf{0}}$.
   (b) The zero vector (in other words, $c_1 = c_2 = c_3 = 0$) is a solution. If it is the *ONLY* solution, then the vectors are linearly independent.
   (c) $$\begin{aligned} -5c_1 + 5c_2 + \ 5c_3 &= 0 \\ 9c_1 + 0c_2 + \ 9c_3 &= 0 \\ 4c_1 + 6c_2 + 16c_3 &= 0 \end{aligned} \quad \Longrightarrow \quad \begin{bmatrix} -5 & 5 & 5 & 0 \\ 9 & 0 & 9 & 0 \\ 4 & 6 & 16 & 0 \end{bmatrix}$$
   (d) $c_1 = -1,\ c_2 = -2,\ c_3 = 1$ OR $c_1 = 1,\ c_2 = 2,\ c_3 = -1$ OR any scalar multiple of these.
   (e) $\vec{\mathbf{u}}_1 = -2\,\vec{\mathbf{u}}_2 + \vec{\mathbf{u}}_3,\ \ \vec{\mathbf{u}}_2 = -\frac{1}{2}\,\vec{\mathbf{u}}_1 + \frac{1}{2}\,\vec{\mathbf{u}}_3,\ \ \vec{\mathbf{u}}_3 = \vec{\mathbf{u}}_1 + 2\,\vec{\mathbf{u}}_2$

3. Solving $c_1 \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix} + c_2 \begin{bmatrix} 5 \\ 1 \\ -6 \end{bmatrix} + c_3 \begin{bmatrix} -4 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ gives the general solution $c_1 = -\frac{3}{2}t,\ c_2 = \frac{1}{2}t,\ c_3 = t$. Therefore $\vec{\mathbf{w}} = \frac{3}{2}\,\vec{\mathbf{u}} - \frac{1}{2}\,\vec{\mathbf{v}}$. (Also $\vec{\mathbf{u}} = \frac{1}{3}\,\vec{\mathbf{v}} + \frac{2}{3}\,\vec{\mathbf{w}}$ and $\vec{\mathbf{v}} = 3\,\vec{\mathbf{u}} - 2\,\vec{\mathbf{w}}$.)

4. (a) $\vec{\mathbf{u}}$ is not in $\text{span}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3)$, but $\vec{\mathbf{w}}$ is.
   (b) $-6\,\vec{\mathbf{v}}_1 + \vec{\mathbf{v}}_2 + 5\,\vec{\mathbf{v}}_3 = \vec{\mathbf{0}}$.
   (c) $\vec{\mathbf{v}}_1 = \frac{1}{6}\,\vec{\mathbf{v}}_2 + \frac{5}{6}\,\vec{\mathbf{v}}_3$.

## Section 4.7 Solutions

Back to 4.7 Exercises

1. $\mathcal{S}_2$ is a basis for $\mathbb{R}^2$. $\mathcal{S}_1$ is not a basis because it is not linearly independent, $\mathcal{S}_3$ is not a basis because it doesn't span $\mathbb{R}^2$.

2. $\mathcal{S}_3$ is not a basis because the vectors are not independent. The otehr sets are all bases for $\mathbb{R}^3$.

3. $\mathcal{S}_2$ is a basis for the line in $\mathbb{R}^2$ containing the given vector. $\mathcal{S}_1$ is not a basis because the set is not independent, $\mathcal{S}_3$ is not a basis because it does not span the space because it isn't in it (it's not on the line)!

4. $\mathcal{S}_1$ is a basis for the $yz$-plane. $\mathcal{S}_2$ and $\mathcal{S}_4$ are not bases for the $yz$-plane because they doesn't span the space, and $\mathcal{S}_3$ isn't a basis because it is not a linearly independent set.

5. (a) Not a subspace, doesn't contain the zero vector.

   (b) Subspace, a basis is $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$  (c) Subspace, $\mathbf{u}$ is a basis.

   (d) Not a subspace, the vector $\begin{bmatrix}-3\\1\\4\end{bmatrix}+\begin{bmatrix}5\\0\\2\end{bmatrix}=\begin{bmatrix}2\\1\\6\end{bmatrix}$ is not on either line because it is not a scalar multiple of either vector.

   (e) Not a subspace, the vector $\begin{bmatrix}1\\2\\3\end{bmatrix}$ is in the set, but $-2\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}-2\\-4\\-6\end{bmatrix}$ is not.

   (f) It can be shown that $\mathbf{w}$, $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, so the set is a plane not containing the zero vector, so it is not a subspace.

   (g) This is a plane containing the origin, so it is a subspace. The set $\{\vec{\mathbf{u}},\vec{\mathbf{v}}\}$ is a basis.

   (h) The vector $\mathbf{0}$ is a subspace.  (i) All of $\mathbb{R}^3$ is a subspace.

6. (a) Not a subspace. $\begin{bmatrix}2\\3\end{bmatrix}$ is in the set, but $-1\begin{bmatrix}2\\3\end{bmatrix}=\begin{bmatrix}-2\\-3\end{bmatrix}$.

   (b) Subspace. A basis would be $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$.

## Section 4.8 Solutions

Back to 4.8 Exercises

1. (a) These are simply the matrices given in the exercise, each augmented with a column of zeros.

   (b) $\vec{\mathbf{x}}=\begin{bmatrix}-s+2t\\s\\t\end{bmatrix}$  (c) $\vec{\mathbf{x}}=s\begin{bmatrix}-1\\1\\0\end{bmatrix}+t\begin{bmatrix}2\\0\\1\end{bmatrix}$

   (d) A basis for the null space of $A$ is $\left\{\begin{bmatrix}-1\\1\\0\end{bmatrix},\begin{bmatrix}2\\0\\1\end{bmatrix}\right\}$.

2. (a) $x_1=-4t$, $x_2=t$ and $x_3=t$.  (b) $\vec{\mathbf{x}}=t\begin{bmatrix}-4\\1\\1\end{bmatrix}$  (c) $\begin{bmatrix}-4\\1\\1\end{bmatrix}$

3. (a) $\text{nullity}(A) = 1$ (b) $\text{rank}(A) = 2$

(b) A basis for the column space of $A$ is $\left\{ \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} \right\}$.

(c) The fact that the basis for the column space has two vectors agrees with the rank.

4. (a) $\vec{\mathbf{u}}_1$ is in the column space and $\vec{\mathbf{u}}_2$ is not.

(b) A basis for col($A$) is $\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix} \right\}$.

(c) $\vec{\mathbf{u}}_1 = -3 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + 5 \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$.

(d) $\text{rank}(A) = 2$

5. (a) $\vec{\mathbf{v}}_1$ is in null($A$) and $\vec{\mathbf{v}}_2$ is not. (b) A basis for null($A$) is $\left\{ \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\}$.

(b) $\text{nullity}(A) = 1$ (d) $\vec{\mathbf{v}}_1 = -2 \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$.

6. (a) $\text{rank}(A) = 2, \quad \text{nullity}(A) = 2$ (b) $\text{rank}(A) = 4, \quad \text{nullity}(A) = 1$

(c) $\text{rank}(A) = 3, \quad \text{nullity}(A) = 2$ (d) $\text{rank}(A) = 3, \quad \text{nullity}(A) = 1$

(e) $\text{rank}(A) = 2, \quad \text{nullity}(A) = 1$ (f) $\text{rank}(A) = 2, \quad \text{nullity}(A) = 3$

(g) $\text{rank}(A) = 3, \quad \text{nullity}(A) = 0$

## Section 4.9 Solutions

Back to 4.9 Exercises

1. (a) $(i)$, $(iii)$ (b) $(ii)$ (c) $(ii)$, $(iii)$ (d) $(ii)$

(e) $(i)$, $(iii)$ (f) $(i)$, $(ii)$, $(iii)$ (g) $(ii)$ (h) $(i)$, $(iii)$

(i) $(ii)$ (j) $(iii)$

2. (a) $\begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & -4 \\ -3 & 4 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ -2 \end{bmatrix}$

(b) $\mathcal{B}_{\text{col}(A)} = \left\{ \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ -1 \end{bmatrix} \right\}$

(c) $\text{rank}(A) = 3$, so the column space spans $\mathbb{R}^3$ and so $\vec{\mathbf{b}}$ must be in the column space, no matter what it is. The system therefore has a solution.

(d) $\text{nullity}(A) = 0$ by the Rank Theorem. We already know that the system has a solution, and now we know it is unique.

(e) $\mathcal{B}_{\text{null}(A)} = \emptyset$, the empty set.

(f) The solution to the system is $\vec{\mathbf{x}} = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$

3. (a) $\begin{bmatrix} 1 & 0 & 2 \\ -2 & 5 & 0 \\ 2 & 5 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5 \\ 11 \\ -7 \end{bmatrix}$

(b) $\mathcal{B}_{\text{col}(A)} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ 5 \end{bmatrix} \right\}$

(c) $\text{rank}(A) = 2$, so the column space does not span $\mathbb{R}^3$. The system may or may not have a solution, depending on whther $\vec{\mathbf{b}}$ is in $\text{col}(A)$.

(d) $\text{nullity}(A) = 1$ by the Rank Theorem. If the solution has a solution, it is not unique.

(e) $\mathcal{B}_{\text{null}(A)} = \left\{ \begin{bmatrix} -2 \\ -\frac{4}{5} \\ 1 \end{bmatrix} \right\}$, or any set containing just one scalar multiple of that vector.

(f) The system has no solution.

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

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

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

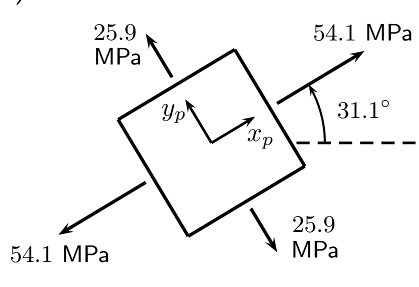
(b) $\lambda_1 = -79.4, \;\; \vec{\mathbf{x}}_1 = \begin{bmatrix} 45 \\ -19.4 \end{bmatrix}, \;\; \lambda_2 = 44.4, \;\; \vec{\mathbf{x}}_2 = \begin{bmatrix} 45 \\ 104.4 \end{bmatrix},$

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

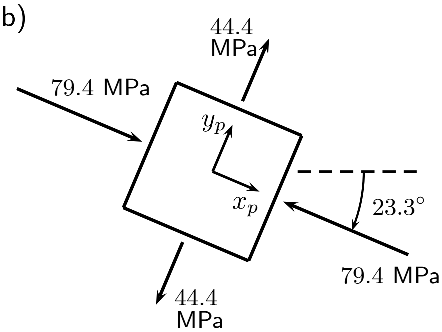
(c) $\lambda_1 = 83.0, \;\; \vec{\mathbf{x}}_1 = \begin{bmatrix} 20 \\ -3.0 \end{bmatrix}, \;\; \lambda_2 = -53.0, \;\; \vec{\mathbf{x}}_2 = \begin{bmatrix} 20 \\ 133.0 \end{bmatrix}, \;\; \theta = \tan^{-1}\dfrac{-3.0}{20} = -8.5^\circ$

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

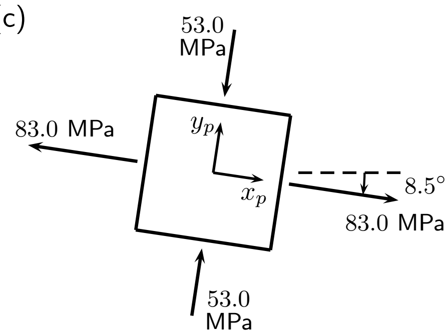
(a)




(b)




(c)




(d)