Chapter 1

What Happened to my Algebra and why is it Linear?

This class will likely be very different from the majority of math classes you have taken in your academic career. In your previous classes, there was likely an emphasis on computations and less so on theory/concepts; in this class we will put equal emphasis on computations and theory/concepts. That is to say: you will be expected to ably compute examples while, at the same time, demonstrate a working knowledge of the theory. I do not mention this to alarm or frighten you. Rather, it is important for you to know that this class may diverge from your usual notion of a math class, and that you persevere in the face of difficulty. With that being said you might be wondering:

Question: Why Take this class?

Answer: The cop-out answer is that this depends on what you want to use linear algebra for (I suggest you think about this as the class progresses), but as this may be your first introduction to the subject, it might be hard to answer this. In light of this, I will do my best to provide an answer that will encompass as many backgrounds as possible. You should study linear algebra because it is useful in many different areas of study, and, most importantly, you will strengthen your problem solving abilities and being comfortable with and overcoming mathematical/intellectual hurdles. I am here to help you when things get challenging; as such I encourage you to unitize office hours. In addition, I encourage you to collaborate with and befriend with your classmates. **I want this to be an enjoyable and worthwhile endeavor for each of you!**

With all of that said, lets answer the question posed by the title of the chapter: What is Linear Algebra?

There are many correct answers to what is linear algebra, each of which depends strongly on what you will use linear algebra for; for now, we will keep in mind a down to earth answer. However, I strongly encourage you to come back periodically and think about what linear algebra means to you as you learn more about this beautiful subject.

Answer: Linear algebra is the study of n-dimensional spaces, the functions between them, and how they fit inside each other. By an n-dimensional space, I mean a 0-dimensional space is a point, a 1-dimensional space is a line, a 2-dimensional is the plane, and so on.

Now, you might be wondering:

Question: How do I succeed in this class.

Answer: To succeed in this class, you should, at minimum, do the following:

1. Study the online lecture notes and/or textbook before and after each class.

2. Come to class with questions.
3. Start homework early! It will be challenging.
4. Come to office hours. **You do not need to have questions to come to office hours; it can be a space for you to do homework with others.**
5. Talk and study with your peers.
6. Don't give up when things get hard (I can't stress this enough)!
7. Most importantly, have fun!

With all of that said, lets embark on our journey into the fantastic world of linear algebra!

Chapter 2

Systems of Linear Equations and Matrices

2.1 Systems of Linear Equations

In the last chapter, we said that linear algebra is the study of n-dimensional spaces, the functions between them, and how they fit inside of each other. Our first step toward making sense of this is through systems of linear equations. First, a couple of definitions, and then we will look at why systems of linear equations are the right thing to look at if we want to study n-dimensional spaces.

Linear Equation

Definition 2.1. A linear equation in the variables $x_1, \ldots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b,$$

where all a_i and b are complex numbers.

Example 2.1. The following are examples of linear equations:

1. $x_1 + 2x_2 = \sqrt{2}$
2. $x_2 - 8x_3 - 2x_5 = 3x_1 + 1$
3. $21x_2 + 5x_4 = 0$

Example 2.2. The following **are not** examples of linear equations:

1. $x_1^2 - x_2 = 5$
2. $e^{x_1 - x_2} + x_3 = 2$
3. $x_1x_4 - x_2x_3 = 0$
4. $\frac{1}{x_1} = 2$
5. $\sin(x_1^2 - x_2^2) = \sqrt{2}$
6. $\sqrt{x_1} + 2x_2 = \sqrt{7}$

Linear equations are very good models of n-dimensional spaces. Indeed, the graph of a linear equation in 2 variables is a line (or 1-dimensional space), the graph of a linear equation in 3 variables is a plane (or 2-dimensional space), and so on. We will want to see how these equations (spaces) interact when they are in the same ambient space; that is: we want to figure out when they intersect. This leads us to the following definition.

System of Linear Equations

Definition 2.2. A system of linear equations is a collection of one or more linear equations involving the same variables.

Example 2.3. The following are examples of systems of linear equations:

1. $\begin{cases} 4x_1 + x_2 = 7 \\ 2x_1 - 2x_2 = \sqrt{3} \\ x_2 + \sqrt{14}x_2 = 0 \end{cases}$

2. $\begin{cases} 2x_1 + x_3 = 0 \\ 2x_1 - 2x_2 = \sqrt{3} \\ x_1 - x_4 = 0 \end{cases}$

Solution

Definition 2.3. A solution to a system of linear equations is a tuple $(s_1, \ldots, s_n)$ of real numbers that make each linear equation in the system a true statement.

Example 2.4. $(3, 2)$ is a solution for the following system of linear equations, but $(0, 1)$ is not a solution.

$$\begin{cases} x_1 - 2x_2 = -1 \\ -x_1 + 3x_2 = 3 \end{cases}$$

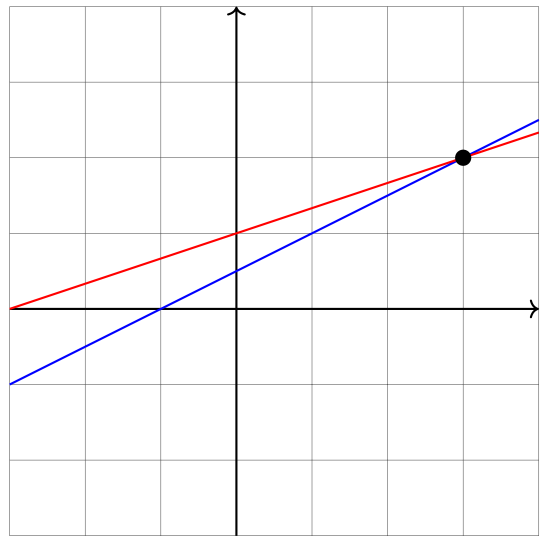
Solution Set

Definition 2.4. The set of all possible solutions to a system of linear equations is called the solution set. Two linear systems are said to be equivalent if they have the same solution sets.

Example 2.5. 1. The solution set to the system

$$\begin{cases} x_1 - 2x_2 = -1 \\ -x_1 + 3x_2 = 3 \end{cases}$$

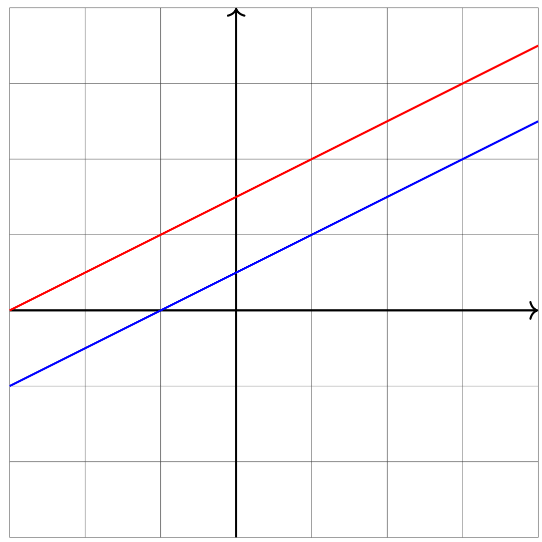
is the set {(3,2)}. Here's a geometric picture of what is going on:



2. The solution set to the system

$$\begin{cases} x_1 - 2x_2 = -1 \\ -x_1 + 2x_2 = 3 \end{cases}$$

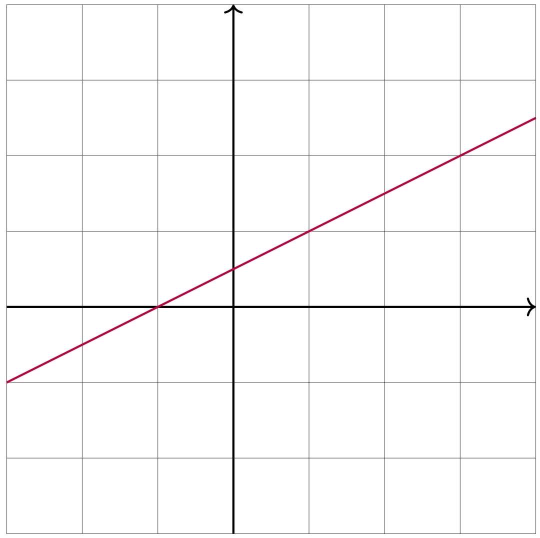
is the empty set, meaning there is no solution. Here is a geometric picture of what is going on:



3. The solution set to the system

$$\begin{cases} x_1 - 2x_2 = -1 \\ 2x_1 - 4x_2 = -2 \end{cases}$$

is the line carved out by $x_1 - 2x_2 = -1$. Here is a geometric picture of what is going on:



These three examples illustrate the following fact:

Theorem 2.1. The solution set to any system of linear equations in any number of variables is either

1. The empty set (i.e no solution).
2. One and only one point.
3. Infinitely many points.

Proof. Postponed for now. □

Consistent and Inconsistent Systems

Definition 2.5. If a system of linear equations has at least one solution, it is called indexconsistent consistent. If it has no solutions, it is called inconsistent.

Example 2.6. Using an online graphing convince yourself that Theorem 2.1 is true for systems of linear equations of three variables.

2.1.1 Matrix Notation

It is very convenient to encode information about a system of linear equations into a rectangular array called a **matrix**. There are two ways to do this, which we will demonstrate through an example.

Example 2.7. Consider the following system of linear equations

$$\begin{cases} 3x_1 + x_2 + 2x_3 = 0 \\ x_1 + x_3 = 2 \\ -2x_1 + 2x_2 = 7 \end{cases},$$

The **coefficient matrix** of the system of linear equations is made by arranging the coefficients of the system into the following matrix

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

We will find it useful to define the **augmented matrix** of a system of linear equations by also including the values on the right hand side of the equal signs:

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

Size of a Matrix

Definition 2.6. A matrix is an $m \times n$ matrix if it has m rows (horizontal) and n columns (vertical).

So, why should we look at matrices? It turns out that the augmented matrix of a system of linear equations is sufficient information to find the solution set. We will dive more into this soon. For now, we will talk about *elementary row operations* as these will very useful tools in our venture to find solution sets for any size system of linear equations!

Elementary Row Operations

Definition 2.7. The *elementary row operations* are the following:

1. Replacement: replace one row by the sum of itself and a multiple of another row.
2. Interchange: Switch the position of two rows.
3. : Multiply all entries in a row by a **nonzero** constant.

Row Equivalent Matrices

Definition 2.8. The *elementary row operations* are the following:
We say that two matrices are row equivalent if one can use elementary row operations to from one matrix to the other.

Here is why we consider these operations:

Theorem 2.2. Suppose a matrix, which we will call A, is the augmented matrix for a system of linear equations. If B is a matrix that is row equivalent to A (i.e we can use row operations to go from A to B), then the system of linear equations represented by B has the same solution set as the original system.

Proof. The proof is outlined as a bonus question on Homework 1. □

This is incredibly powerful! We can simplify an augmented matrix in such a way that we can determine whether or not a system is consistent, as we shall now see:

Example 2.8. Consider the system of linear equations

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ 5x_1 - 5x_3 = 10 \end{cases}.$$

The augmented matrix of this system is

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix}.$$

Since elementary row operations do not change the solution set to a system of linear equations, lets simplify the above matrix to something more tame.

$$\begin{aligned} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix} &\xrightarrow{-5R_1+R_3 \mapsto R_3} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 10 & -10 & 10 \end{bmatrix} \xrightarrow{\frac{1}{10}R_3 \mapsto R_3} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 1 & -1 & 1 \end{bmatrix} \\ &\xrightarrow{-2R_3+R_2 \mapsto R_2} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & -6 & 6 \\ 0 & 1 & -1 & 1 \end{bmatrix} \xrightarrow{\frac{1}{6}R_2 \mapsto R_2} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix} \end{aligned}$$

We may stop here (although we can go further, which we will do next time). From this new matrix, we can pick out a solution to the original system of linear equations. Indeed, row 2 tells us that $x_3 = -1$. Using this, row 3 tells us that $x_2 = 0$. Finally, row 1 tells us that $x_1 = 1$. Therefore, the original system is consistent with $(1, 0, -1)$ being a solution.

Helpful Tip: It is helpful to think of matrices and systems of linear equations as the same thing under different guises. What we mean by this is that if you see a matrix you should think about it as being the augmented matrix of some system of equations, and, on the other hand, when you see a system of linear equations, you should think of its augmented matrix. **As with many aspects of mathematics, different points of view of the same thing is a power that cannot be overestimated.**

2.2 More Row Reduction and the Echelon Forms

In the last section we introduced elementary row operations and saw how they can aid us in solving a system of linear equations. It seems reasonable, at this point, to ask

Question: Can we find the solution set to any system of linear equations by writing down its augmented matrix and performing row operations?

Answer: Yes!

Not only can we find the solution set to any system of linear equations using row operations, but there are *always* two forms we can row reduce the augmented matrix into that yields valuable information about our system of linear equations called **Row Echelon Form (REF)** and **Reduced Row Echelon Form (RREF)**. These forms and the information they posses will be the content of this section, and

The Echelon Forms

Definition 2.9.

1. A matrix is said to be in row Echelon form (REF), or simply Echelon form, if it satisfies the following properties:

   (a) All nonzero rows are above any rows consisting of all zeros.

   (b) Each leading entry **(i.e the left most nonzero entry)** of a row is in a column tot he right of the leading entry of the row above it.

   (c) All entries below a leading term are 0.

2. A matrix is said to be in row reduced Echelon form (RREF) if it satisfies the following conditions:

   (a) It satisfies all properties of being in REF.

   (b) The leading entry of each row is 1.

   (c) Each leading 1 is the only nonzero entry in its column.

If a matrix is in RREF, then it is in REF. However, if a matrix is in REF, then it may not be true that it is in RREF, as we shall see in examples 2.9 and 2.10.

Example 2.9. The following matrix is in REF but not in RREF:

$$\begin{bmatrix} 2 & 0 & 7 & 9 & 1 & 10 \\ 0 & 0 & 3 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

The following matrix is not in REF:

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

Example 2.10. The following matrix is in RREF:

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

The following matrix is in REF but not RREF:

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

Theorem 2.3. Given any matrix A, there exists a series of row operations that put a matrix into a REF. Similarly, there is a series of row operations that put a matrix into a RREF. In other words, A is row equivalent to REF and RREF matrices.

Proof. We will postpone this for now. Later, we will describe an algorithm that shows this. □

Uniqueness of RREF

Theorem 2.4. Given any matrix A, the RREF of A is **unique**. That is to say there is only one RREF we can row reduce A to.

Proof. We won't prove this; however, I encourage you to try! □

The following is a homework exercise, but as it is important, we will state it here.

A matrix A has a unique RREF by Theorem 2.2. However, a matrix A can have many different REF's. Construct an example of a matrix with multiple REF's.

Before we discuss an algorithm to row reduce a matrix to a REF or RREF, we will find it helpful to define a couple of terms. But first, a related meme:

Pivots

Definition 2.10. A **pivot position** in a matrix A (not necessarily in RREF) is a location in A that corresponds to a leading 1 in the RREF of A. A pivot column of A is a column of A that contains a **pivot column**.

This is, perhaps, a strange definition; so, lets do an example and identify the pivot positions and pivot columns of a matrix.

Example 2.11. We will find all of the pivot positions and pivot columns of the following matrix

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

Lets perform some row operations to put A into RREF:

$$\begin{bmatrix} 3 & 0 & 2 & 1 \\ 2 & 2 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 3 & 0 & 2 & 1 \end{bmatrix} \xrightarrow[-3R_1+R_3 \mapsto R_3]{-2R_1+R_2 \mapsto R_2} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 3 & 2 & 1 \end{bmatrix}$$

$$\xrightarrow{\frac{1}{4}R_2 \mapsto R_2} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 3 & 2 & 1 \end{bmatrix} \xrightarrow[R_2+R_1 \mapsto R_1]{-3R_2+R_3 \mapsto R_3} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix}$$

$$\xrightarrow{\frac{1}{2}R_3 \mapsto R_3} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

This is the RREF of A. Below are circled the leading ones, which correspond to the pivot positions of A. Also, highlighted, are the pivot columns of A.

As the above example shows **not all columns of a matrix will be pivot columns!**.

2.2.1 The Row Reduction Algorithm (Gaussian Elimination)

We will describe the algorithm of Gaussian Elimination, which yields a way to put a matrix into REF or RREF. Given a matrix A, we proceed in the following steps:

Gaussian Elimination

Algorithm 2.1.

1. Start with the left most nonzero column. Select a nonzero entry in the column to be a pivot. If necessary, interchange two rows so the pivot is at the top of the matrix.
2. Use row operations to get zeros in all positions below the pivot you found in step 1.
3. Cover the first row of the new matrix and apply 1-2 to the new matrix with the first row deleted. Keep doing this until you cannot.
4. Starting form the right most column, create zeros above each pivot.

Let's return to systems of linear equations! Given an augmented matrix, we may put it into REF or RREF, which yields information about the solution set of the system of linear equations. Before we see an example of this, we will take a brief detour through the world of solution sets. As we said in the previous section, there are systems of equations with infinitely many solutions; for these systems it would be very inefficient (and impossible) to enumerate all solutions by hand. To remedy this, we can describe all solutions using parameters (or free variables). To this end, we will find the next definition to be useful.

Free and Basic Variables

Definition 2.11. Suppose a matrix is A is the augmented matrix of a system of linear equations. The variables corresponding to pivot columns are called basic variables and the other variables are called free variables.

Example 2.12. We will find the solution set of the following system of linear equations by putting its augmented matrix into RREF and identifying pivot positions/columns:

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 0 \\ 4x_1 + 4x_3 + x_4 = 8 \\ 2x_1 + 6x_3 + 8x_4 = 4 \end{cases}.$$

Let's apply Gaussian elimination to put the augmented matrix into RREF:

$$\begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 4 & 0 & 4 & 1 & 8 \\ 2 & 0 & 6 & 8 & 4 \end{bmatrix} \xrightarrow[-2R_1+R_3\mapsto R_3]{-4R_1+R_2\mapsto R_2} \begin{bmatrix} 1 & 2 & 2 & 0 & 0 \\ 0 & -8 & -8 & 1 & 8 \\ 0 & -4 & 0 & 8 & 4 \end{bmatrix} \xrightarrow{\frac{-1}{4}R_3\mapsto R_3} \begin{bmatrix} 1 & 2 & 2 & 0 & 0 \\ 0 & -8 & -8 & 1 & 8 \\ 0 & 1 & 0 & -2 & -1 \end{bmatrix}$$

$$\xrightarrow{R_2\leftrightarrow R_3} \begin{bmatrix} 1 & 2 & 2 & 0 & 0 \\ 0 & 1 & 0 & -2 & -1 \\ 0 & -8 & -8 & 1 & 8 \end{bmatrix} \xrightarrow{8R_2+R_3\mapsto R_3} \begin{bmatrix} 1 & 2 & 2 & 0 & 0 \\ 0 & 1 & 0 & -2 & -1 \\ 0 & 0 & -8 & -15 & 0 \end{bmatrix}$$

$$\xrightarrow{\frac{-1}{8}R_3\mapsto R_3} \begin{bmatrix} 1 & 2 & 2 & 0 & 0 \\ 0 & 1 & 0 & -2 & -1 \\ 0 & 0 & 1 & \frac{15}{8} & 0 \end{bmatrix} \xrightarrow{-2R_3+R_1\mapsto R_1} \begin{bmatrix} 1 & 2 & 0 & \frac{-15}{4} & 0 \\ 0 & 1 & 0 & -2 & -1 \\ 0 & 0 & 1 & \frac{15}{8} & 0 \end{bmatrix}$$

$$\xrightarrow{-2R_2+R_1\mapsto R_1} \begin{bmatrix} 1 & 0 & 0 & \frac{1}{4} & 2 \\ 0 & 1 & 0 & -2 & -1 \\ 0 & 0 & 1 & \frac{15}{8} & 0 \end{bmatrix}$$

We then see that x_1, x_2, and x_3 are basic variables for our system, and x_4 is a free variable. From the RREF we found, we find that $x_3 = -\frac{15}{8}x_4$, $x_2 = 2x_2 - 1$, and $x_1 = -\frac{1}{4}x_4 + 2$. A sufficient way to write this is to say the following is a **parametric description** of the general solution:

$$\begin{cases} x_1 - \frac{1}{4}x_4 + 2 \\ x_2 = 2x_2 - 1 \\ x_3 = -\frac{15}{8}x_4 \\ x_4 \text{ is free} \end{cases}$$

Note: sometimes people replace free variables with letters like t or s; we won't do that here, but we might in the future. Here, we see that x_4 can be any number, and it will determine what x_1, x_2, and x_3 need to be to give a solution to the system (this is why we say that x_4 is a free variable). For example, set $x_4 = 0$. Then, $x_1 = 2$, $x_2 = -1$ and $x_3 = 0$. Thus, $(2, -1, 0, 0)$ is a solution to our original solution..

The number of free variables of a system is the dimension of the solution set. To make sense of this, we need to agree on a notion of dimension, which will come up later.

We will end this section with a beautiful theorem that determines whether or not a system is consistent just by analyzing the REF of a matrix!

Theorem 2.5. A system of linear equations is consistent if and only if the right most olumn of the augmented matrix is not a pivot column. In other words, the REF of the augmented matrix has *no row* of the form

$$\begin{bmatrix} 0 & \dots & 0 & b \end{bmatrix}$$

where b is any *nonzero* number.

2.3 Vectors in $\mathbb{R}^n$

In this section we will introduce the notion of vectors; however, for now, we will only focus on vectors in $\mathbb{R}^n$. As we will see later, there is a more abstract notion of vectors. We will find that vectors offer us a very convenient language to describe what is going on with systems of linear equations and the spaces that they carve out!

$\mathbb{R}^n$

Definition 2.12. The set $\mathbb{R}^n$ is define to be the set of of n-tuples of real numbers.

Example 2.13. Lately, we have been working with $\mathbb{R}^2$, which we visualize as the Cartesian plane. We've even thought about $\mathbb{R}^3$ a bit, which we visualize as 3 dimensional space. In general we may think of $\mathbb{R}^n$ as n-dimensional space. Though, I must caution you that later on n-dimensional space will encompass more than just $\mathbb{R}^n$; however, for now, we will abuse this terminology!

Vectors

Definition 2.13. A vector (or column vector) in $\mathbb{R}^n$ is a $n \times 1$ matrix of real numbers. **Caution:** later on vector will mean something that will encompass more than the elements of $\mathbb{R}^n$.

Example 2.14. The following is a vector in

$$\begin{bmatrix} 1 \\ 0 \\ 2 \\ \sqrt{6} \end{bmatrix}$$

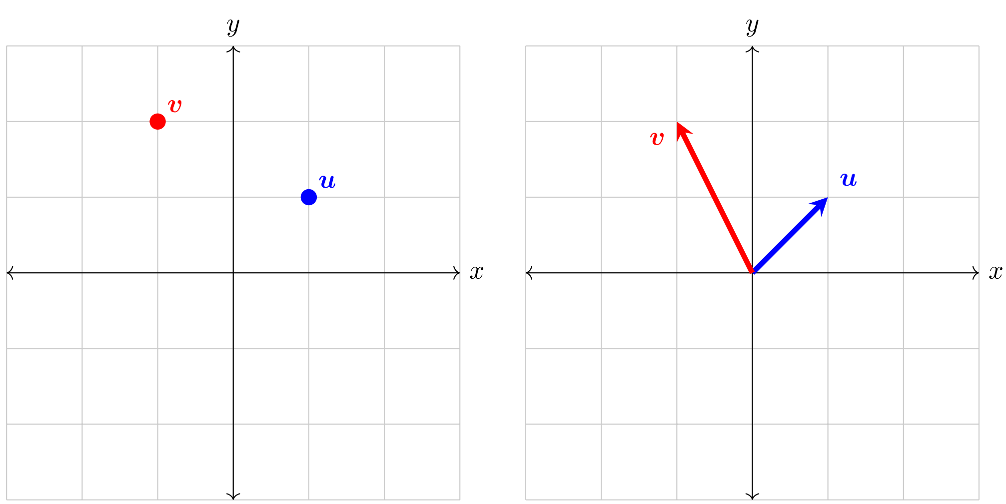
in $\mathbb{R}^4$

As we can see in the last example, sometimes it takes up a lot of space to write a vector as an $n \times 1$ matrix, so we will sometimes write a vector as an n-tuple. Precisely, we shall take

$$\begin{bmatrix} a_1 \\ \ddots \\ a_n \end{bmatrix} = (a_1, \dots, a_n)$$

Caution: $(a_1, \dots, a_n) \neq \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix}$. The latter is an $1 \times n$ matrix.

A vector $\mathbf{a} = (a_1, \dots, a_n)$ can be visualized in two ways: 1) They can be thought of as points $(a_1, \dots, a_n)$ in $\mathbb{R}^n$; 2) We can also think of the vector $\mathbf{a}$ as an arrow from the origin of $\mathbb{R}^n$ to the point $(a_1, \dots, a_n)$. We will often take the view of 2) when thinking about vectors. Below are two figures that illustrate these two viewpoints.



Visualizing vectors as arrows is extremely useful, as we shall see shortly.

2.3.1 Operations with Vectors

While vectors **are not numbers** (rather they are an ordered list of numbers), we can still perform some operations with them, like addition, subtraction, and scaling. Though we must caution ourselves!

Vector Addition and Scalar Multiplication

Definition 2.14. We add two vectors $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ component wise:

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, \ldots, a_n + b_n).$$

Suppose that $c \in \mathbb{R}$ (c is a real number). Then, $\mathbf{a}$ scaled by c is

$$c\mathbf{a} = (ca_1, \ldots, ca_n).$$

Proposition 2.1. Let $\mathbf{a}, \mathbf{b}$, and $\mathbf{c}$ be vectors in $\mathbb{R}^n$ and $c, d \in \mathbb{R}$. Then, the following hold:

a) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$.
b) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$.
c) $\mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}$.
d) $\mathbf{a} + (-\mathbf{a}) = -\mathbf{a} + \mathbf{a} = \mathbf{0}$.
e) $c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b} = (\mathbf{a} + \mathbf{b})c$.
f) $(c + d)\mathbf{a} = c\mathbf{a} + d\mathbf{a}$.
g) $c(d\mathbf{a}) = cd\mathbf{a}$.
h) $1\mathbf{a} = \mathbf{a}$.

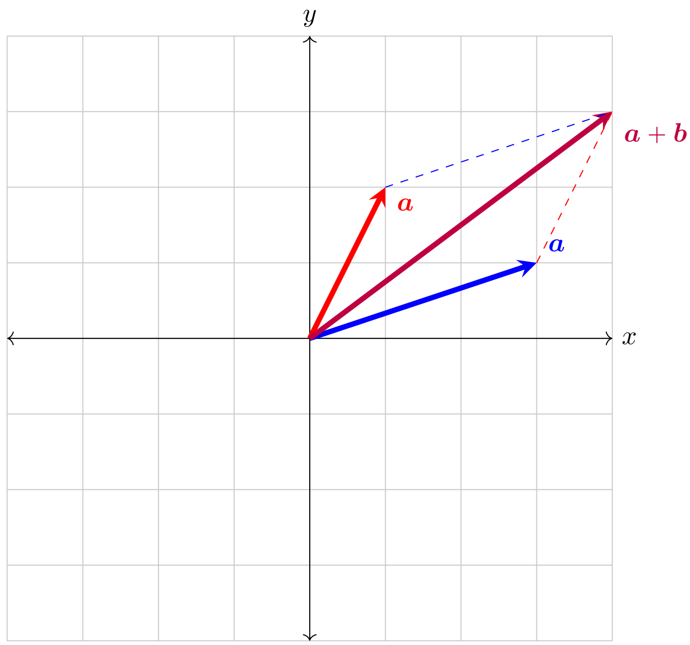
Example 2.15. We compute:

$$\frac{1}{2}\left(\begin{bmatrix} 2 \\ 4 \\ -6 \\ 0 \\ -8 \end{bmatrix} + \begin{bmatrix} 4 \\ -2 \\ 0 \\ 2 \\ -8 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 0 \\ -4 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ 0 \\ 1 \\ -4 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ -3 \\ 1 \\ -8 \end{bmatrix}.$$

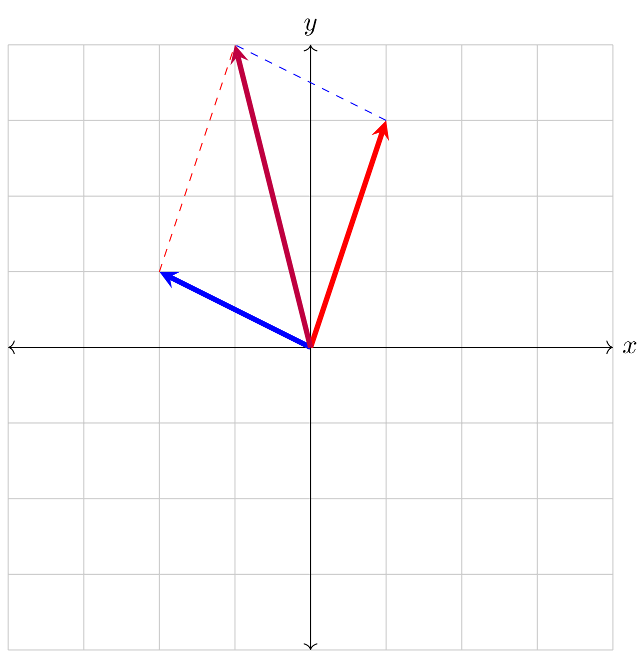
Note, there are a few different ways to go about computing this. I encourage you to try and find another way!

Parallelogram Law for Vectors

Proposition 2.2. The addition of two vectors **a** and **b** in $\mathbb{R}^2$ is the fourth vertex of the parallelogram whose other three vertices are $\mathbf{0}, \mathbf{a}$, and **b**.



Example 2.16. Below, we draw the vectors $\mathbf{a} = (-2, 1)$ and $\mathbf{b} = (1, 3)$, and then draw $\mathbf{a} + \mathbf{b}$ by using the Parallelogram Law for Vectors.



2.3.2 Linear Combinations and Spans

Linear Combination

Definition 2.15. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be vectors in $\mathbb{R}^n$ and $c_1, \ldots, c_n \in R$. Then, we call

$$c_1\mathbf{x}_1 + \ldots + c_n\mathbf{x}_n$$

a **linear combination** of the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with weights $c_1, \ldots, c_n \in R$

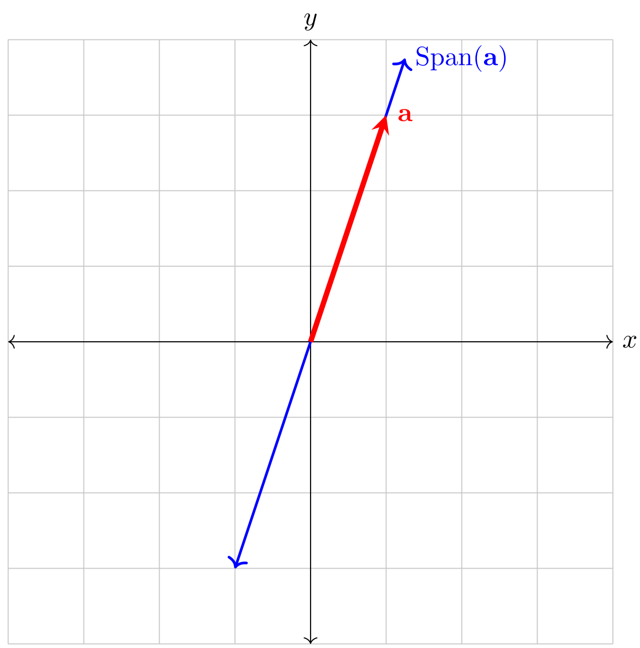
Often, we will want to consider all linear combinations of a set of vectors. To this end, we define the **span** of a set of vectors.

Span

Definition 2.16. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be vectors. We define the Span of $_{\mathbf{1}}, \ldots, _{\mathbf{n}}$ to be the set of all linear combinations of $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Often, we will denote this set by $\mathrm{Span}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$

Lets get a geometric idea of what spans of vectors look like.

Example 2.17. Consider the vector $\mathbf{a} = (1, 3)$ in $\mathbb{R}^2$. The span of $\mathbf{a}$ is the set of all linear combinations of $\mathbf{a}$; that is the set of all things of the form $c\mathbf{a}$ with $c \in \mathbb{R}$. Geometrically, this is the line through the orgin passing through $(1, 3)$, as depicted below.



Example 2.18. Consider the vectors $\mathbf{a} = (0, 1, 1)$ and $\mathbf{b} = (0, 0, 1)$ in $\mathbb{R}^3$. What do you think $\mathrm{Span}(\mathbf{a})$ looks like? What about $\mathrm{Span}(\mathbf{a}, \mathbf{b})$?

Using the same logic as found in the previous example, we can surmise that $\mathrm{Span}(\mathbf{a})$ is the line in $\mathbb{R}^3$ that passes through the origin passing through the point $(0, 1, 1)$.

Finding $\mathrm{Span}(\mathbf{a}, \mathbf{b})$ is slightly more challenging. It is the plane in $\mathbb{R}^3$ that contains the orgin, $\mathbf{a}$, and $\mathbf{b}$. Try and convince yourself of this!

Example 2.19. Let $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$. We will determine whether or not the vector $\mathbf{a} = \begin{bmatrix} -1 \\ -6 \\ -2 \end{bmatrix}$ is in the span of $\mathbf{x}_1$ and $\mathbf{x}_2$. Translating from math to English, can we find weights c_1 and c_2 such that

$$c_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -6 \\ -2 \end{bmatrix}.$$

Lets use our knowledge of addition of vectors to simplify the above equation to get:

$$\begin{bmatrix} c_1 + c_2 \\ 2c_1 + 4c_2 \\ c_2 \end{bmatrix} = \begin{bmatrix} -1 \\ -6 \\ -2 \end{bmatrix}.$$

Thus, $\begin{bmatrix} -1 \\ -6 \\ -2 \end{bmatrix}$ is a linear combination of $\mathbf{x}_1$ and $\mathbf{x}_n$ if and only if the system given by

$$\begin{cases} c_1 + c_2 = -1 \\ 2c_1 + 4c_2 = -6 \\ c_2 = -2 \end{cases}.$$

As if, magically, by design, we have talked about how to solve systems like this! There are a few different ways to go about it, but we will use Guassian Elimination to determine if the system is consistent or not. Skipping a few steps, of which I will leave to you to check, the REF of

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

is the matrix

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

Thus, by Theorem the above system is consistant and so $\mathbf{a} = \begin{bmatrix} -1 \\ -6 \\ -2 \end{bmatrix}$ is a linear combination of $\mathbf{x}_1$ and $\mathbf{x}_2$. Going a bit further, the REF we just found tells us that $c_1 = 1$ and $c_2 = -2$ will work (check this!)

The above example demonstrates the following theorem

Theorem 2.6. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$, and $\mathbf{b}$ be vectors. Then, the following are equivalent (i.e the following say the same thing):

1. $\mathbf{b}$ is in the span of $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$

2. there are $c_1, \ldots, c_n \in \mathbb{R}$ such that
$$c_1\mathbf{x}_1 + \ldots c_n\mathbf{x}_n = \mathbf{b}.$$

3. the system, whose augmented matrix is
$$\begin{bmatrix}\mathbf{x}_1 & \ldots & \mathbf{x}_n & \mathbf{b}\end{bmatrix},$$
is consistent.

Example 2.20. Using this theorem, we can see that the span of the vector $(1,0)$ and $(0,1)$ is all of $\mathbb{R}^2$! Here, we will notice that we usually think of the dimension of $\mathbb{R}^2$ as 2, and there are two vectors that span $\mathbb{R}^2$. Hmm... maybe dimension and spanning are related. We will make this observation formal soonish!

2.4 Matrix Equations

As we have seen, it is often useful to interpret systems of equations as information encoded into a matrix (i.e its augmented matrix). We will continue this theme of translating ideas into expressions involving matrices! Last time we talked about linear combinations of vectors in $\mathbb{R}^n$; it turns out that we can encode this infomration into something called a matrix equation. Before we do this, we should discuss matrix operations.

2.4.1 Matrix Equations

Matrix Addition

Definition 2.17. Let A and B be two matrices of the same size, say $m \times n$ with
$$\text{A=}\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & \ldots & \vdots \\ a_{mn} & \ldots & a_{mn} \end{bmatrix} \qquad \text{B=}\begin{bmatrix} b_{11} & \ldots & a_{1n} \\ b_{21} & \ldots & b_{2n} \\ \vdots & \ldots & \vdots \\ b_{m1} & \ldots & b_{mn} \end{bmatrix}.$$
We define the addition of A and B to be
$$A + B \begin{bmatrix} a_{11}+b_{11} & \ldots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & \ldots & a_{2n}+b_{2n} \\ \vdots & \ldots & \vdots \\ a_{1m}+b_{m1} & \ldots & a_{mn}+b_{mn} \end{bmatrix}.$$
We refer to this as component wise addition.

Note that in the definition of $A + B$, where A and B are matrices of the same size, it does not make a whole lot of sense to add two matrices of different sizes!

Now lets talk about how we can multiply two matrices. Unfortunately, we cannot just multiply any two matrices we wish; the two matrices we want to multiply must complement eachother in some way.

L

Definition 2.18. t A and B be two matrices of possibly different sizes. Let A have size $m \times n$ and B have size $j \times k$. Then,

1. if the number of columns of A is the number of rows of B (or $n = j$), we define

$$AB = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \dots & \vdots \\ a_{mn} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \dots & b_{1k} \\ b_{n1} & \dots & b_{2k} \\ \vdots & \dots & \vdots \\ b_{n1} & \dots & b_{jk} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \dots + a_{1n}b_{n1} & a_{11}b_{12} + a_{12}b_{22} + \dots + a_{1n}b_{n2} & \dots & a_{11}b_{1k} + a_{12}b_{2k} + \dots + a_{1n}b_{nk} \\ a_{21}b_{11} + a_{22}b_{21} + \dots + a_{2n}b_{n1} & a_{21}b_{12} + a_{22}b_{22} + \dots + a_{2n}b_{n2} & \dots & a_{21}b_{1k} + a_{22}b_{2k} + \dots + a_{2n}b_{nk} \\ \vdots & \vdots & \dots & \dots \\ a_{m1}b_{11} + a_{m2}b_{21} + \dots + a_{mn}b_{n1} & a_{m1}b_{12} + a_{m2}b_{22} + \dots + a_{mn}b_{n2} & \dots & a_{m1}b_{1k} + a_{m2}b_{2k} + \dots + a_{mn}b_{nk} \end{bmatrix}$$

2. if the number of columns of A is not the number of rows of B, then AB is not defined.

Using this definition of matrix multiplication, we can reframe the notion of linear combinations of vectors using matrix notation.

Proposition 2.3. Let A be an $m \times n$ matrix, with columns given by the $\mathbf{a}_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \dots, \mathbf{a}_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$. In other words

$$A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \dots & \vdots \\ a_{mn} & \dots & a_{mn} \end{bmatrix}.$$

Let $\mathbf{c}$ be any vector in $\mathbb{R}^n$ (note that this is the same n that occurs in the size of A). Then, the product of A and $\mathbf{c}$,

$$A\mathbf{c} = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1\mathbf{a}_1 + \dots + c_n\mathbf{a}_n$$

is the linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_n$ with weights $c_1, \dots, c_n$. **Warning:** $A\mathbf{c}$ makes sense, but $\mathbf{c}A$ does not! (why?)

Lets practice using this propositon!

Example 2.21. 1. $\begin{bmatrix} 2 & 7 & 8 & 11 \\ 1 & 8 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 6 \\ 0 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3\begin{bmatrix} 7 \\ 8 \end{bmatrix} + 6\begin{bmatrix} 8 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 11 \\ 0 \end{bmatrix}$

2. For any vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}+3$ in $\mathbb{R}^3$, we can write the linear combination $3\mathbf{x}_1 - 2\mathbf{x}_2 + 8\mathbf{x}_3$ as a matrix times a vector:

$$\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 8 \end{bmatrix} = 3\mathbf{x}_1 - 2\mathbf{x}_2 + 8\mathbf{x}_3.$$

All of these equations may seem confusing and daunting, but remember: we are using matrix multiplication to translate linear combinations to an equality involving matrices (and vice versa). It is two ways of writing the same thing, and being able to fluidly go back and forth between the two is very important!

The following is an important and useful fact:

Proposition 2.4. Let A and B be $m \times n$ matrices, let $\mathbf{a}$ and $\mathbf{b}$ be vectors in ^{n}, and let $c \in \mathbb{R}$. Then,

1. $A(\mathbf{a}+\mathbf{b}) = A\mathbf{a} + A\mathbf{b}$
2. $(A+B)(\mathbf{a}) = A\mathbf{a} + B\mathbf{b}$
3. $A(c\mathbf{a}) = (cA)\mathbf{a}$

2.4.2 Old results under new guises.

A constant theme in this class is rewriting a number of things in different ways; we've had a bit of practice with this already! So, for many sections in these notes, we will be rewriting old theorems in new terminology. We will start this practice by relating systems of equations to the language of matrix equations!

Theorem 2.7. Let $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix}$ be an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$. Then, the following are equivalent:

1. $\mathbf{x} = (x_1, \dots x_n)$ is a solution to the system of linear equations represented by the augmented matrix $\begin{bmatrix} \mathbf{a}_1 & \dots \mathbf{a}_n & \mathbf{b} \end{bmatrix}$
2. $A\mathbf{x} = \mathbf{b}$
3. $\mathbf{b}$ is a linear combination of the vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ with weights $x_1, \dots, x_n$. That is

$$x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n = \mathbf{b}.$$

The following corollary is a consequence of the above Theorem.

Corollary 2.1. Let $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{bmatrix}$. Then, the following are equivalent:

1. Every $\mathbf{b} \in \mathbb{R}^m$, there is a solution, $\mathbf{x}$, t0 the matrix equation $A\mathbf{x} = \mathbf{b}$.
2. Every $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the columns of A.
3. The columns of A span $\mathbb{R}^m$. In other words, $\text{Span}(\mathbf{a}_1, \dots, \mathbf{a}_n) = \mathbb{R}^m$.
4. A has a pivot position in every row.

Lets get a bit of practice using these useful facts!

Example 2.22. Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 1 \\ 0 & 1-2 & \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ be any vector in $\mathbb{R}^3$. We will determine whether or not $A\mathbf{x} = \mathbf{b}$ has a solutions (equivalently consistent) for all possible b_1, b_2, b_3.

First, lets row reduce the augmented matrix for $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} 1 & 2 & 0 & b_1 \\ 2 & 3 & 1 & b_2 \\ 0 & 1 & -2 & b_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 2 & 0 & b_1 \\ 0 & 1 & -2 & b_3 \\ 0 & 0 & 1 & \frac{-2b_1+b_2+b_3}{3} \end{bmatrix}.$$

From this, we see that A has a pivot position in each row. Hence, by our corollary above, $A\mathbf{x} = \mathbf{b}$ is consistent for any choice of $\mathbf{b}$. Note that we could have just row reduced A rather than the augmented matrix, but sometimes its useful to know what is happening to the b_i, since we can use them to solve for solutions.

2.5 Solution Sets and Applications / Worksheet 1

Homogeneous System

Definition 2.19. A system of equation is said to be homogeneous, if it can be written as $A\mathbf{x} = \mathbf{0}$. A homogeneous system always has a solution, namely $\mathbf{0}$, which we call the trivial solution. Any other solution, if it exists is called a nontrivial solution.

1. Consider the system

$$\begin{cases} 3x_1 + 5x_2 - 4x_3 = 0 \\ -3x_1 - 2x_2 - 4x_3 = 0 \\ 6x_1 + x_2 - 8x_3 - x_4 = 0 \end{cases}.$$

   (a) Does the system have a nontrivial solution?

(b) Find a parametric description of its solution set.

(c) Think of a way to rewrite your answer in (b) as a vector equation. Hint: Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ be a solution, and use your answer in (b) to find something this vector is equal to.

2. Fill in the blank: the homogeneous equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution if and only if the equation has at least one _________ variable. Hint: see problem 1.

Non-Homogeneous System

Definition 2.20. A system of equation is said to be non-homogeneous, if it can be written as $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} \neq \mathbf{0}$. In other words, a system of equations is said to be non-homogeneous if it is not homogeneous.

3. Come up with a non-homogeneous system of equations that does not have a solution. How is this different than homogeneous systems?

4. Consider the homogeneous system of equations

$$\begin{cases} 3x_1 + 5x_2 - 4x_3 = 1 \\ -3x_1 - 2x_2 - 4x_3 = -1 \\ 6x_1 + x_2 - 8x_3 - x_4 = 2 \end{cases} .$$

(a) Is the system consistent?

(b) Find a parametric description of its solution set.

(c) Think of a way to rewrite your answer in (b) as a vector equation. Hint: Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$ be a solution, and use your answer in (b) to find something this vector is equal to.

5. Compare your answers in 1c and 4c. What do you notice?

Problem 5 alludes to the following theorem:

S

Theorem 2.8. ppose $A\mathbf{x} = \mathbf{b}$ is *consistent* with a solution $\mathbf{p}$ (it can be any solution you want). Then any solution, $\mathbf{w}$, to $A\mathbf{x} = \mathbf{b}$ is given by

$$\mathbf{w} = \mathbf{v}_h + \mathbf{p},$$

where $\mathbf{v}_h$ is a solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$. **Warning:** the choice of $\mathbf{v}_h$ depends on $\mathbf{w}$.

6. This exercise will outline the proof of the above theorem. There are two main parts to the proof.

   (a) Suppose that $\mathbf{w} = \mathbf{v}_h t + \mathbf{p}$ with $t \in \mathbb{R}$. Show that $\mathbf{w}$ is a solution to $A\mathbf{x} = \mathbf{b}$.

   (b) We aren't done yet! We still need to show that every solution to $A\mathbf{x} = \mathbf{b}$ has the form $\mathbf{v}_h + \mathbf{p}$ for some solution, v_h, to the homogeneous system $A\mathbf{x} = \mathbf{0}$.

      i. Show that $\mathbf{w} - \mathbf{p}$ is a homogeneous solution to $A\mathbf{x} = \mathbf{0}$.

      ii. Set $\mathbf{v}_h := \mathbf{w} - \mathbf{p}$, and show that $\mathbf{w} = \mathbf{v}_h + \mathbf{p}$.

   (c) Briefly explain why parts (a) and (b) complete the proof of the theorem.

7. Boron Sulfide reacts with water to create boric acid and hydrogen sulfide gas. We will use linear algebra to balance the following chemical equation that illustrates this reaction:

$$x_1 B_2S_3 + x_2 H_2O \rightarrow x_3 H_3BO_3 + x_4 H_2S.$$

To do so, find whole numbers whole numbers x_1, x_2, x_3, and x_4 such that the total number of Boron (B), Sulfur (S), Hydrogen (H), and Oxygen (0) on the left matches the number on the right. Hint: Try and set up a system of linear equations.

Index