LINEAR ALGEBRA WEEKS 1-6

Luke Rawlings

November 5, 2025

Contents

1 THINGS TO KNOW

1.1 WHY IS THIS COURSE DIFFERENT?

Welcome to a new level of mathematical work. You are no longer in a course focused solely on computational work, and you have reached a level of study that invites you to evolve and become more intellectually curious, inventive, creative, careful, and thoughtful, not simply precise. Being precise characterizes mathematics, as usual, but answering the question "*why does this work?*" makes a mathematician.

This course requires you to think and to communicate, in both writing and speaking, beyond the level of computational work, challenges you to make connections among concepts, and encourages you to understand that mathematics is about the art of generalization.

If you were to return to any of the mathematics courses which you have already studied, you might consider studying them with more critical thought. Here are a few random questions from your prior mathematical work to think about, which I hope might steer you in the direction of more profound thought regarding *this* course:

(a) Why are the Hindu-Arabic numerals 1, 2, 3, ... written in the way they are written? How did these numbers come to take their form and shape?

(a) Do you know the origin stories of any of the symbols you use, such as the "is equal to" symbol as used in $2+3=5$, or "the square root" symbol as used in $\sqrt{4}=2$, or any other symbol you have met in mathematics?

(c) Think about the operation of addition and how you have generalized this concept: Addition is repeated counting. Multiplication is repeated addition: $2+2+2 = 3 \cdot 2$. Exponentiation is repeated multiplication: $2 \cdot 2 \cdot 2 = 2^3$. Why stop there? Tetration is repeated exponentiation: $2^{2^2} = {}^{3}2$. You can keep going, too: there is pentation, hexation, heptation, octation, enneation, decation, vigintation, trigintation, centation, docentation, myriation, circulation... I'm not kidding. Is this not fun to know? These are called **hyper-operators** in mathematics.

(d) Speaking of exponentiation, I assume all of you are comfortable with seeing $2^2, 2^3$, and 2^4 and understanding what these powers of a number mean. So why on earth is $2^{-1} = \frac{1}{2}, 2^{-2} = \frac{1}{4}$, and $2^{-3} = \frac{1}{8}$?

(e) Why is 0!=1?

(f) You began learning calculus as a child: Calculus without limits is called *arithmetic*. arithmetic with limits is called *calculus*. What do I mean by this?

(g) Speaking of calculus, integration is addition with limits, and differentiation is subtraction with limits. Are you aware that you learn introductory calculus kind of backwards in that, historically speaking, integration came way before differentiation? What is so special about Newton and Leibniz?

(h If a line that touches a curve in one and only one point is called *tangent* and a line that cuts through a circle in two points is called *secant*, why are these words used in trigonometry as function names like $\tan\theta$ and $\sec\theta$?

(i) Linear algebra will generalize two statements you made as a child: Sentences such as "*Two points determine a line*" and "*Three non-collinear points determine a plane*" turn into "*One vector determines a line*" and "*Two linearly independent vectors determine a plane.*" The big generalization of these concepts is summed up in a single sentence: *A set of n linearly independent vectors determines an n-dimensional subspace.*

There are a lot more things that could be added to this list, all of which enhance a deeper thinking of the subject under which such questions could arise. These often delight students and bring a greater humanity to the subject of mathematics. Why is it that we leave out these stories when we teach and learn mathematics? (That's a question for any instructor of mathematics, but also one for any of you who might happen to make the courageous decision to one day teach it yourself!)

I dare you to abandon being obedient, only after thinking about when it is necessary to do so, and attempt to establish within yourself the art of play. Ask why when you can. I dare you to seek joy, or some sense of joy, as you explore this subject. Education, in my opinion, can sometimes chase the joy and fun out of learning mathematics: Our exam culture, with its system of rewards and punishments, ruins the party. I invite you to consider the idea that to truly learn and master anything, the art of play must be present.

1.2 WHAT ARE SCALARS?

The term *scalars* refers to numbers, and these numbers are initially limited to *real numbers*. Some of you have probably encountered *complex numbers*, and these are also scalars in some settings. Scalars will be denoted by lowercase, italicized letters and have the following properties:

A: To every pair of scalars a and b, there corresponds another scalar $a+b$ called the **sum** of a and b such that

1: $a+b=b+a$,

2: $a+(b+c)=(a+b)+c$,

3: there is a unique scalar called *zero*, denoted by 0 such that $a+0=a$ for every scalar, and

4: to every scalar a there corresponds a unique scalar $-a$ such that $a+(-a) = 0$.

B: To every pair of scalars a and b, there corresponds another scalar ab called the **product** of a and b such that

1: $ab = ba$,

2: $a(bc) = (ab)c$,

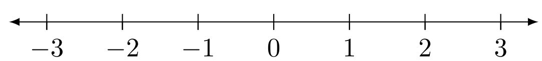
3: there exists a unique scalar 1 called *one* such that $1a = a$ for every scalar a, and

4: to every scalar that is not zero, there corresponds a unique scalar a^{-1} such that $aa^{-1} = 1$.

C: Multiplication is *distributive* with respect to addition: $a(b+c) = ab+ac$.

Given a set of scalars, if addition and multiplication are defined in such a way that all of the conditions above are satisfied, then that set is called a **field**. This is why you often hear someone mention they have a *field of scalars* in mathematics.

The set of all real numbers is denoted by $\mathbb{R}$. This is the field of scalars for this introductory course, meaning that anytime we refer to quantities known as scalars, they are real numbers. The geometric model of $\mathbb{R}$ is the real number line you developed long ago in elementary school: a line drawn horizontally, with a point known as the origin. To the left of the origin are the negative real numbers. Located to the right of the origin are the positive real numbers.



Hopefully you recognize the properties of a field as ones which you have encountered in a beginning algebra course: Both operations are commutative, associative, have an identity element, and an inverse element, and multiplication is distributive with respect to addition.

1.3 WHAT ARE VECTORS?

The term *vectors* refers to things that can be added together. Most of you have used *Euclidean vectors* in physics and mathematics: these are the directed line segments with a magnitude (length) and direction. These vectors indicate a displacement: something is carried from one location to another location. These are not the only kinds of vectors: Polynomials are vectors. Matrices are vectors. Examples abound! Vectors are elements of a space known as a **vector space**. This is the main concept of the course, one in which we do not refine until after several introductory chapters. Vectors are denoted by bold non-italicized letters. A **vector space** is a set V of elements called vectors satisfying the following axioms:

1: To every pair of vectors $\mathbf{x}$ and $\mathbf{y}$ in V, the sum $\mathbf{x}+\mathbf{y}$ is also a vector in V,

2: $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$,

3: $\mathbf{x}+(\mathbf{y}+\mathbf{z})=(\mathbf{x}+\mathbf{y})+\mathbf{z}$,

4: there is a unique vector called the *zero vector* $\mathbf{0}$ such that $\mathbf{x}+\mathbf{0}=\mathbf{x}$ for all vectors $\mathbf{x}\in V$,

5: to every vector $\mathbf{x}$ there corresponds a vector $-\mathbf{x}$ such that $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$,

6: for any scalar a and any vector $\mathbf{y}$ in V, the product $a\mathbf{y}$ is also a vector in V,

7: $a(b\mathbf{x})=(ab)\mathbf{x}$ for any scalars a and b,

8: $1\mathbf{x}=\mathbf{x}$,

9: $a(\mathbf{x}+\mathbf{y})=a\mathbf{x}+a\mathbf{y}$, and

10: $(a+b)\mathbf{x}=a\mathbf{x}+b\mathbf{x}$.

This course will focus on **Euclidean vector spaces**, ones in which you can imagine vectors to be directed line segments from one point to another. The field of scalars $\mathbb{R}$ is itself a vector space, but we will treat it separate from the vector spaces we will study. You are no doubt familiar with the xy-plane, also known as $\mathbb{R}^2$, formed by two real lines (axes), one orthogonal to the other, both meeting at the origin. $\mathbb{R}^2$ is the set of all vectors with two components. Ordered pairs (x, y) have coordinates x and y. Ordered pairs establish a location. The ordered pair $(2, 3)$ indicates a location at which you arrive if you move two units in the positive x-direction and three units in the positive y-direction. The vector from the origin $(0, 0)$ to the location $(2, 3)$ can be thought of as the directed line segment drawn from the origin to $(2, 3)$. If you took calculus, you distinguish between location and displacement in your written work: The location is $(2, 3)$ and the vector is $\langle 2, 3\rangle$. In this class, the two are the same thing; that is, we refer to (x, y) as a vector and do not use the calculus and physics notation. Here are the definitions of Euclidean vector spaces:

$\mathbb{R}^2$: the set of all vectors with two real number components (x_1, x_2).

$\mathbb{R}^3$: the set of all vectors with three real number components (x_1, x_2, x_3).

$\mathbb{R}^4$: the set of all vectors with four real number components (x_1, x_2, x_3, x_4).

... etc....

$\mathbb{R}^n$: the set of all vectors with n real number components $(x_1, x_2, ..., x_n)$.

VECTOR SPACE	DESCRIPTION
$\mathbb{P}_1$	all polynomials of the form $a_0 + a_1x$
$\mathbb{P}_2$	all polynomials of the form $a_0 + a_1x + a_2x^2$
$\mathbb{P}_3$	all polynomials of the form $a_0 + a_1x + a_2x^2 + a_3x^3$
$\mathbb{M}_{2\times 2}$	all matrices with two rows and two columns.

Table 1: A Few Non-Euclidean Vector Spaces

When n is finite, you have a finite vector space, but it is possible to have a vector space with an infinite number of components. This is denoted by $\mathbb{R}^\infty$ and seems confusing to think about at first, but here is something that helps: if you studied sequences and series, it was often the case that you had an infinite sequence. A sequence is an ordered list of numbers. The catch: **A sequence is a vector!**

Vectors are defined by two operations: **vector addition** and **scalar multiplication**:

DEFINITION 1: Let $\mathbf{u} = (u_1, u_2, ..., u_n)$ and $\mathbf{v} = (v_1, v_2, ..., v_n)$ be two vectors in $\mathbb{R}^n$ and let k be any scalar. **Vector addition** is

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, ..., u_n + v_n),$$

while **scalar multiplication** is

$$k\mathbf{v} = (kv_1, kv_2, ..., kv_n).$$

It is this definition that truly characterizes a vector. **Table 1.1** lists some examples of vectors from vector spaces that you may not have ever thought of as vectors; however, note the familiar mathematical moves you make and how these moves relate to this first definition in the following examples:

(1) In $\mathbb{R}^2$, $(1,3) + (2,4) = (1+2, 3+4) = (3,7)$ and $4(1,2) = (4\cdot 1, 4\cdot 2) = (4,8)$. If you have taken calculus and physics, you often see and denote an addition like this using the basis vector notation: If $\mathbf{v} = \mathbf{i} + 3\mathbf{j}$ and $\mathbf{w} = 2\mathbf{i} + 4\mathbf{j}$, then $\mathbf{v} + \mathbf{w} = 3\mathbf{i} + 7\mathbf{j}$. If $\mathbf{u} = \mathbf{i} + 2\mathbf{j}$, then $4\mathbf{u} = 4\mathbf{i} + 8\mathbf{j}$.

(2) In $\mathbb{R}^3$, $(1,5,-1) + (2,-7,3) = (3,-2,2)$ and $-3(1,2,-3) = (-3,-6,9)$.

(3) In $\mathbb{R}^4$, $(1,-1,0,3)+(3,2,1,7) = (4,1,1,10)$ and $5(1,2,3,2) = (5,10,15,10)$.

(4) The polynomials $1 + 3x$ and $2 + 4x$ are from the vector space $\mathbb{P}_1$. Notice that $(1+3x) + (2+4x) = (1+2) + (3x+4x) = 3 + 7x$ and $4(1+2x) = 4+8x$. In beginning algebra, you say you "add like terms" or "add similar terms" when adding polynomials. When multiplying a polynomial by a real number, you use the distributive property. This is vector addition and scalar multiplication. Note the same mathematics is executed in (1) above.

(5) Add $1+5x-x^2$ to $2-7x+3x^2$ and find $-3(1+2x-3x^2)$. Again, look at the mathematics done in (2) and compare the computations to the ones here. What is the difference?

(6) Matrices are rectangular arrays of numbers that you will see arise naturally from solving systems of equations. You will see them originate from the work we will do beginning in the next several chapters. The vector space of matrices $\mathbb{M}_{2\times 2}$ are all matrices with two rows and two columns, ones of the following form:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

In many books, there is a special notation for each element that appears in a matrix, a subscript notation ij where the first number i denotes the row the element appears and the second number j denotes the column in which the element appears. In this text, square brackets will be used to denote matrices, and when the matrix itself is referred to, an italicized capital letter is used to denote the matrix. Here are two matrices A and B using these notations:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ and } B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Matrices are vectors because you can add them and scalar multiply them. If k is any scalar (any real number), then vector addition and scalar multiplication for matrices are as follows:

$$A + B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{bmatrix}$$
$$kA = \begin{bmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{bmatrix}$$

1.4 HOW DO YOU WRITE AND TYPE SCALARS AND VECTORS?

In courses where students are using vector quantities for the first time, I advise using proper notation to distinguish between scalars and vectors. It is astonishing to see a growing number of students who dismiss this, ignore it, and could care less. What this often results in, especially in calculus, is complete confusion within problems over *location* and *displacement*. While mathematics can be a challenging subject to master, one of the reasons why students find it confusing at times is because little attention is given to how they communicate what they do in a problem. Accurate answers are important to reach in many problems you encounter in mathematics, which is a goal, of course; however, what many of you will find makes you valuable in a number of future jobs in the mathematical sciences is what a lot of people look for and treasure: GOOD COMMUNICATION SKILLS. In recent years, I have seen a ***lack of attention to the detail***

involved in writing mathematical notation. WRITING USING PROPER NOTATION THAT IS CLEAR AND EASY TO READ IS A CHOICE YOU MAKE AS A STUDENT OF THIS SUBJECT. I cannot express to you how many times I have sat down with a student who is lost in the subject, only to realize that once the student pays attention to how it is written, they begin to understand it much better. If you become confused in your work, consider the thought and care you put into your written communication of the mathematics you write. Have a high bar for yourself. Consider adding this question when you sit down to do homework: *What kind of experience will another human being have who reads my work?* Write as if you are communicating to someone else who is at your level or just beneath your level of mathematics, so your writing teaches them the mathematics. Write in such a way that you do not cause that reader of your work to do work in order to understand you. This can be challenging in this subject, and it is worth making an effort to be as clear as you can in your written communication. It will only benefit you and most likely help you understand the subject better. An answer to a problem is often an essay using the notation which acts as a kind of shorthand for how you speak it. You will find it important to mentally distinguish between vector quantities and scalar quantities. And if you write mathematics or speak it, I believe you should have others in mind who must read what you write or hear what you speak. Clarity goes a very long way, and you will either be appreciated for it, or you will frustrate and confuse others.

(1) Vectors, when typed in this text, are denoted by bold, non-italicized, lowercase letters:

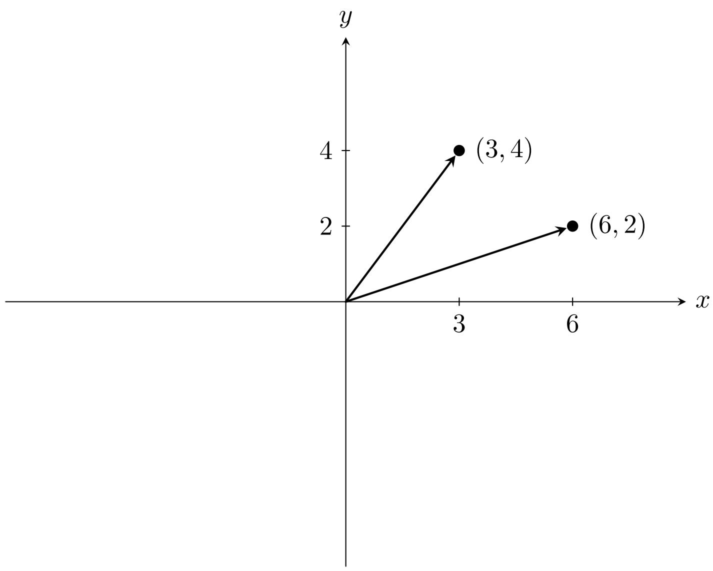
$$\mathbf{v}$$

(2) Vectors, when handwritten by your instructor in the lecture, are denoted by non-italicized lowercase letters with an arrow above them:

$$\vec{\mathrm{v}}$$

(3) When scalars are represented by letters, they are represented by italicized lowercase letters when both typed and handwritten: $k\mathbf{v}$ when typed, or $k\vec{\mathrm{v}}$ when handwritten, where k is the scalar. (You can *see* it is a scalar.)

(4) In this class, Euclidean vectors will be written in a number of different ways. Most of the time vectors will be written primarily as matrices. You will hear "column vector" spoken often. Below are two vectors in $\mathbb{R}^2$.



Both are drawn in what is called **standard position**, meaning from the origin to whatever point they each would carry a particle if the initial point was the origin. If you studied calculus, you denoted these vectors, most likely as $3\mathbf{i}+4\mathbf{j}$ and $6\mathbf{i}+2\mathbf{j}$, where $\mathbf{i}$ and $\mathbf{j}$ are the unit vectors in the positive x and y directions; that is, $3\mathbf{i}+4\mathbf{j}$ means move three units to the right and four units up. You also denoted these vectors as $\langle 3,4\rangle$ and $\langle 6,2\rangle$. This latter notation distinguishes these vectors from the locations $(3,4)$ and $(6,2)$. In *this* class, all of these notations are considered synonymous and we do not use the notation $\langle 3,4\rangle$. In fact, we consider $(3,4)$ to be the **space-saver notation** for a vector because we will write vectors in matrix form as one-column matrices. In other words, vectors will be written in the first two of the four forms you see below: For the vector from the origin to the point $(3,4)$, we have the column vector and the space-saver notation for that column vector, which equates to the second two forms with which many of you are familiar from calculus.

$$\begin{bmatrix}3\\4\end{bmatrix} = (3,4) = \langle 3,4\rangle = 3\mathbf{i}+4\mathbf{j}$$

To be clear about the names of each notation, the above vector, written four times from left to right, is first written as a single column matrix, then in the space-saver notation, then in the component vector notation, and finally in the standard basis vector notation.

It will probably be useful for you to think of all vectors in Euclidean vector spaces as column vectors, based on the mathematics you have yet to learn in this class, which involves how these vectors are seen and used in matrices. Below are column vector notations for some general vector $\mathbf{x}$ in some of the Euclidean vector spaces:

If $\mathbf{x}$ is from $\mathbb{R}^2$, then it is of the form $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

If $\mathbf{x}$ is from $\mathbb{R}^3$, then it is of the form $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

If $\mathbf{x}$ is from $\mathbb{R}^4$, then it is of the form $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$.

In general, if $\mathbf{x}$ is from $\mathbb{R}^n$, then it is of the form $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$.

(5) It is nice to have some kind of an outline or a "big picture" that kind of summarizes everything you will be learning in a course. This is sometimes challenging to do in mathematics courses, but this has helped some students in the past with clarity, especially when things get confusing. You all know how to solve equations of the form $ax = b$, where a and b are known real numbers and x is the unknown variable. For example, $3x = 6$ has the unique solution $x = 2$. It is fairly easy to come to terms with the three possible solutions for all equations of the form $ax = b$: There is one and only one solution if a has an inverse a^{-1}: The solution is $x = a^{-1}b$. There are an infinite number of solutions if both $a = 0$ and $b = 0$. There is no solution if $a = 0$ and $b \neq 0$. This is a first part of understanding the nature of all equations of the form $ax = b$: finding all values of x that satisfy the equation. A second part of understanding equations of this form is to look at the left side ax as a **function** $f(x) = ax$. As you will see in the last section of this chapter, all functions of the form $f(x) = ax$ are **linear transformations**. Keep this simple idea in mind, pretending that there is a course called BEGINNING BEGINNING LINEAR ALGEBRA out there somewhere that is broken up into these two parts:

BEGINNING BEGINNING LINEAR ALGEBRA

1. PART 1: All solutions to $ax = b$
2. PART 2: All functions $f(x) = ax$

Indeed, this may provide many of you with a simple idea for how to get back on track if you happen to become overwhelmed or confused. Any introductory course in linear algebra can be broken up into two similar parts:

INTRODUCTORY LINEAR ALGEBRA (MATH 208)

1. PART 1: All solutions to $A\mathbf{x} = \mathbf{b}$
2. PART 2: All functions $T(\mathbf{x}) = A\mathbf{x}$

What you most likely do not yet know, however, is that $A\mathbf{x} = \mathbf{b}$ is a matrix equation that represents a system of linear equations, and that $T(\mathbf{x}) = A\mathbf{x}$ represents what are called linear transformations.

1.5 WHAT IS FUNCTION, DOMAIN, CODOMAIN, IMAGE, RANGE?

The concept of *function* is so central to understanding mathematics beyond computational work. A function has a set of inputs, and produces a set of outputs with the following condition: *For each input, there is one and only one output.* Functions you have seen are denoted by a name, followed by a set of parentheses inside of which is a symbol representing any input, followed by an "is equal to" symbol, and followed by a symbol representing any output or the rule for how the inputs are transformed:

$$f(x) = y \tag{1}$$
$$f(x) = x^2 \tag{2}$$
$$f(x,y) = \sqrt{x-3y} \tag{3}$$
$$f(x_1,x_2,x_3,x_4) = (x_2+2x_4, 6x_1-x_2+x_3) \tag{4}$$

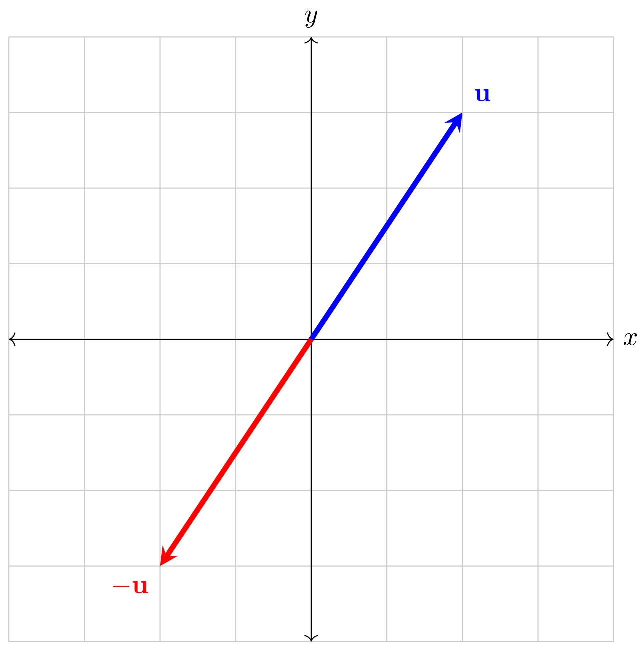
The **domain** of a function is the set of all inputs. The **range** of a function is the set of all outputs. The **codomain** of a function is the space in which its outputs are located. For the function $f(x) = x^2$, its domain is $\mathbb{R}$, its codomain is $\mathbb{R}$, and its range is all non-negative real numbers, sometimes more formally written in the set-builder notation as $\{x \in \mathbb{R} : x \geq 0\}$ or the interval notation $[0, \infty)$. The **image** of an element is the element to which it is mapped, so 4 is the image of 2 since $f(2) = 4$, 9 is the image of 3 since $f(3) = 9$, etc... Later in the course, we refer to functions as *mappings,* or *transformations*, as in f maps 2 to 4, or f transforms 3 to 9. The domain and the codomain are often presented as important spaces when a function is introduced. This function $f(x) = x^2$ maps the set of real numbers into the set of real numbers, and this is often written $f : \mathbb{R} \to \mathbb{R}$. This is a class of functions with which most of you are familiar, real-valued functions of one variable. Those of you who have studied mutlivariable functions are familiar with the example in (1.3), a function mapping points in the plane to real numbers. What is its domain? range? What is the image of $(2, -1)$? What is its codomain? The function in (1.4) is a function we will study in this course: it is a linear transformation. Later in the course, we will introduce a linear transformation like this in a formal way: Let $f : \mathbb{R}^4 \to \mathbb{R}^2$ be given by $f(x_1, x_2, x_3, x_4) = (x_2 + 2x_4, 6x_1 - x_2 + x_3)$. The domain of this function is $\mathbb{R}^4$. The codomain is $\mathbb{R}^2$. These are two important spaces to identify for linear transformations. We have quite a lot of material to cover before we study linear transformations, but it certainly does not hurt to let you know the stronger meaning of the word *linear* based on the functions you all know well from your study of algebra and beginning calculus, which is the focus of the next section.

1.6 WHAT FUNCTIONS $f(x)$ ARE LINEAR?

In this class, you must abandon a careless use of the the word *linear* and recognize that it is used in two ways in mathematics: you often refer to functions whose graphs are lines in the plane as linear functions, but this is a *geometric* use of the word. This class introduces an *algebraic* use of the word. The function $f(x) = 2x + 6$ is not linear. The function $f(x) = 2x$ is linear. Here is the definition of the word when the class of functions is restricted to real-valued functions of a single variable: A function $f(x)$ is **linear** if it satisfies two properties: For any two points x_1 and x_2 in the domain of f, and for any scalar k, **(a)** $f(x_1 + x_2) = f(x_1) + f(x_2)$ and **(b)** $f(kx_1) = kf(x_1)$. Property **(a)** is known as *additivity* and property **(b)** is known as *homogeneity*. These are usually spelled out together in other courses as *the principle of superposition*, which is $f(k_1x_1 + k_2x_2) = k_1f(x_1) + k_2f(x_2)$. You have all taken calculus and are familiar with the operations of differentiation and integration. Notice that these are operations that obey the principle of superposition. Are they linear operations? What do you think? (Yes, they are...) In the case of functions of one variable, all linear functions are of the form $f(x) = ax$, where a is any real number. The functions $f(x) = ax$ and $g(y) = by$ are linear functions. Combining them together means adding them, so why not call it a linear combination?

1.7 WHAT ARE LINEAR COMBINATIONS?

A linear combination of x and y is $ax + by$, where a and b are scalars. A linear combination of $\mathbf{u}$ and $\mathbf{v}$ is $a\mathbf{u} + b\mathbf{v}$. A linear combination of ω and ϵ is $a\omega + b\epsilon$. The ten systems you solved in problems 1-10, the opening exercises in this text, are systems involving linear combinations of x and y or linear combinations of x, y, and z, so we refer to them as systems of linear equations.



y
$\mathbf{u}$
x
$-\mathbf{u}$

2 THE ACTUAL SYLLABUS

Below is a description of how we will move through the course. Included in this description are the topics of study (in bold) with a few comments.

- **GAUSS-JORDAN ELIMINATION**
- **ROW OPERATIONS**
- **REDUCED ROW ECHELON FORM (RREF)**
- MATRIX MULTIPLICATION (DOT PRODUCT & OTHER OPERATIONS)
- **THE INVERSE OF A MATRIX**
- CONNECTIONS BETWEEN ROW OPERATIONS AND MATRIX MULTIPLICATION
- **MORE ABOUT MATRIX MULTIPLICATION**
- **LINEAR COMBINATIONS & SUBSPACES**
- **LINEAR INDEPENDENCE/DEPENDENCE**
- **SOLUTIONS TO** $A\mathbf{x} = \mathbf{b}$
- **BASIS & DIMENSION**
- **LINEAR TRANSFORMATIONS**
- **THE ROW & COLUMN SPACES OF A MATRIX**
- **VISUALIZING MATRICES**
- **THE FOUR FUNDAMENTAL SUBSPACES OF A MATRIX**
- **FINDING BASES**
- **RANK & NULLITY**
- **THE FUNDAMENTAL THEOREM OF LINEAR ALGEBRA**
- **THE METHOD OF LEAST SQUARES**
- **EIGENSPACES**
- **CHANGE OF BASIS**
- **ABSTRACT VECTOR SPACES**
- **DIAGONALIZATION** if time permits...

3 GAUSS-JORDAN ELIMINATION

3.1 HOMEWORK 1

Before any instruction, here is the very first homework assignment. Solve the following ten systems using *any* method you wish.

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

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

3.
$$\begin{aligned} x + 2y &= 3 \\ x - y &= -3 \end{aligned}$$

4.
$$\begin{aligned} x - y &= 2 \\ x + y &= 4 \end{aligned}$$

5.
$$\begin{aligned} x + y &= 6 \\ x - y &= 2 \end{aligned}$$

6.
$$\begin{aligned} x + 3y &= 2 \\ 2x + 6y &= 1 \end{aligned}$$

7.
$$\begin{aligned} x - 2y &= -5 \\ -3x + 6y &= 15 \end{aligned}$$

8.
$$\begin{aligned} x - y + z &= 1 \\ x - z &= 0 \\ x - y + 2z &= 2 \end{aligned}$$

9.

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

10.

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

The study of solving systems of linear equations is where we begin the mathematical work in our course. It is really the birthplace of the subject of linear algebra. Though you solved the first ten problems using *any method* you were taught prior to this course, you will now learn to solve systems by a strict algorithmic approach, similar to how any computer might be programmed to solve a system: *Gauss Jordan Elimination.* Understanding this method is simply a very systematic way of eliminating the unknown variables to arrive at a solution to the system, if it has one: *For any system of k equations and n unknowns* $x_1, x_2, ..., x_n$, eliminate x_1 from all other equations but the first, solving for this variable, then eliminate x_2 from all other equations but the second, solving the second equation for this variable, then eliminate x_3 from all other equations but the third, solving for this varible,... I hope that you understand the pattern and that it makes perfect sense. This is the single skill to practice on the following systems in this next set of exercises. Repeat what your instructor did in class, solving each system by the method described above, recording the moves that you make. Obtain a description of the solution to the system.

3.2 HOMEWORK 2

11.

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

12.

$$\begin{aligned} x_1 - x_2 &= 1 \\ x_1 + x_2 &= 3 \end{aligned}$$

13.

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

14.

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

15.

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

16.

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

17.

$$\begin{aligned} x_1 - x_2 + x_3 &= 1 \\ x_1 - x_3 &= 0 \\ x_1 - x_2 + 2x_3 &= 2 \end{aligned}$$

18.

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

3.3 The Process of Elimination

Of all the methods learned to solve systems, it is *elimination* to which we devote close attention. A majority of you, when solving the first ten problems on your own, most likely used some form of elimination, which involves altering one or both equations, then adding them so that one of the variables disappears. We will refine this process, and learn a solution method known as *Gauss-Jordan elimination*. This gives birth to matrices, and many of the important ideas of linear algebra. It is crucial that you pay close attention to the content in the course, working to understand it as best as you can. We take the following system as our first example of the process of solution by elimination, the first one you solved in the course:

Example 1:

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

Whatever solution method is used, the outcome will be the same: a value is found for x and a value is found for y. If we go about our business solving the problem, and keeping things orderly, we can write that the solution to the system is

$$\begin{aligned} x &= -4 \\ y &= 2 \end{aligned}$$

This corresponds to the point of intersection $(-4, 2)$ of the two lines in the system of equations. Here is the important thing to grasp: *From the system as it was given to the result above, we wish to use elimination to rewrite the original system to somehow change it to the form above.* Let me show you how this is done a first time. This is a meticulous process, but important to understand. Before beginning a detailed description of the solution method, one step at a time, notice there are four equations above with something in common: *each is a sentence containing a **linear combination** of x and y on the left side of the equation, and a single constant term on the right side.* Additionally, when the system is reinterpreted using the column perspective, it is a linear combination of two vectors. If we were to restate the specific linear combinations at the start, we have this:

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

At the end, we have:

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

In other words, we wish to work with the method of elimination in such a way as to produce this ending set of linear combinations. This brings an important emphasis on the way you should see the columns in a special way.

STEP 1: Acknowledge the system as it is given:

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

STEP 2: We want a 1 by the x in the first equation, first column. We have that. Beneath it, we want a zero in front of the x; that is, we want to eliminate x in the second equation. Therefore, *we must use the first equation to change the second*. We do this mentally by **multiplying the first equation by -1, and adding it to the second equation**. This action produces the following result:

$$\begin{aligned} 1x + 3y &= 2 \\ 0x + 1y &= 2 \end{aligned}$$

STEP 3: We want a 1 by the y in the second equation, second column. We have that. We want a zero in front of the y in the first equation, second column. We don't have that, so we use the second equation to change the first one; that is, we eliminate y in the first equation by **multiplying the second equation by -3 and adding it to the first**, which produces the following result:

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

Notice the algorithmic nature here: we eliminated variables in a strategic way, something we will soon be articulating in more detail. What we actually have done is sometimes surprising to know: matrix multiplication! Most likely, that means absolutely nothing to you unless you have had experience multiplying matrices. What on earth is matrix multiplication? We have some ground to cover. Let's first of all witness the birth of a matrix, nothing other than a rectangular array of numbers.

3.4 Matrices and Augmented Matrices

One of the most important observations you can make regarding this process of elimination is this: *The only elements involved in the mathematics performed were the coefficients and constants that appear in the system.* In fact, the process involved two specific steps spelled out in words as the action taken between steps that changed these numbers from one step to the next. If we were to look at the process again, using the same system as before, but writing it in a special way in which we view the coefficients and the constants as being in certain *columns* and form the following array of numbers, we can say that we have a system of two linear equations and two unknowns x and y, and it is represented by the following **matrix** with two rows and three columns, with a vertical bar added (not necessary) that is helpful to remind you of the system (it separates the coefficients of variables from the numbers to the right of the equals to symbol.

$$\left[\begin{array}{cc|c} 1 & 3 & 2 \\ 1 & 4 & 4 \end{array}\right]$$

A **matrix** is a rectangular array of numbers. The plural of *matrix* is **matrices**. A matrix is described by what we call its **size**, which is specified by the numbers of rows and columns it has in the following way: An $m \times n$ matrix is a matrix with m rows and n columns. The matrix we just produced from the system of two linear equations is a 2×3 matrix. Observe the same process of solution we used, which involved two operations:

Multiply the first equation by -1, and add it to the second equation.

Multiply the second equation by -3 and add it to the first equation.

Example 2: Solve the system in example 2 by using the augmented matrix and the row operations. (This is how you will be solving systems in your work.) Replace the words "the first equation" with "row 1" and "the second equation" with "row 2" and then use R1 for "row 1" and R2 for "row 2." These two operations can be described symbolically: (-1)R1+R2 and (-3)R2+R1, or just $-$R1+R2 and $-$3R2+R1. Furthermore, the process itself can be written out in the following way:

$$\left[\begin{array}{cc|c} 1 & 3 & 2 \\ 1 & 4 & 4 \end{array}\right] \xrightarrow{-R1+R2} \left[\begin{array}{cc|c} 1 & 3 & 2 \\ 0 & 1 & 2 \end{array}\right] \xrightarrow{-3R2+R1} \left[\begin{array}{cc|c} 1 & 0 & -4 \\ 0 & 1 & 2 \end{array}\right]$$

Written this way, you can interpret the steps detailed on the previous page concisely, as well as understand the intermediate actions taken to justify the changes. It is crucial that you pick up an important skill as quickly as possible: the ability to read a matrix when you are aware that the matrix represents a system of linear equations. We will learn more technical details about matrices as mathematical objects later, but systems of equations is where things originate. The 2×3 matrix representing the original system is called an **augmented matrix**.

Example 3: Write the augmented matrix representing the system in homwork problem 11.

Solution: You should be able to "see" that it is rather straight-forward to write

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 2 & -1 & 1 & 2 \end{array}\right]$$

Keep in mind that the bar may or may not be included (See problems 23-26.)

3.5 HOMEWORK 3

19. Write the augmented matrix for each system in the first homework assignment. Give the size of each matrix.

20. Give an example of a 3×3 matrix, selecting at random any numbers you wish. Then write the system of linear equations this matrix could represent.

21. **True** or **False**: A $m \times n$ augmented matrix represents a system of m equations and $n-1$ unknown variables.

22. Study carefully the solution to the system given in the three steps on page 6. Between steps 1 and 2, notice that the first equation was *used* to change the second equation; that is, the first equation *did not change* between steps 1 and 2. Notice that, between steps 2 and 3, the second equation was *used* in order to change the first equation. Take the system below and follow the instructions given to make changes to the equations, rewriting the result after each action is taken. (Avoid writing any matrices.)

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

Operation 1: Multiply the first equation by -2 and add it to the second equation in order to change the second equation.

Operation 2: Divide the second equation by -3.

Operation 3: Multiply the second equation by -1 and add it to the first equation in order to change the first equation.

23. Check the following description (in matrix form) of the work in the previous problem. Does it represent the four systems of equations you wrote as your answer?

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & -1 & 3 \end{bmatrix} \xrightarrow{-2R1+R2} \begin{bmatrix} 1 & 1 & 3 \\ 0 & -3 & -3 \end{bmatrix} \xrightarrow{R2/(-3)} \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow{-R2+R1} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$

24. Observe the matrix notation in the previous problem in order to appreciate what information is conveyed (especially how the operations are described above the arrows). Translate the following into a system of equations, and explain how the system was changed one step at a time in order to produce this information. At the end of your explanation, state the solution to the system.

$$\begin{bmatrix} 1 & 2 & 8 \\ 3 & 2 & 12 \end{bmatrix} \xrightarrow{-3R1+R2} \begin{bmatrix} 1 & 2 & 8 \\ 0 & -4 & -12 \end{bmatrix} \xrightarrow{R2/(-4)} \begin{bmatrix} 1 & 2 & 8 \\ 0 & 1 & 3 \end{bmatrix} \xrightarrow{-2R2+R1} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{bmatrix}$$

25. How would you interpret the matrix $\begin{bmatrix} 1 & 0 & 7 \\ 0 & 0 & 3 \end{bmatrix}$?

26. How would you interpret the matrix $\begin{bmatrix} 1 & 0 & 7 \\ 0 & 0 & 0 \end{bmatrix}$?

27. Use the descriptions of what to do between each step to replace all of the question marks with the correct numbers. (Do not worry about interpreting this work yet, but know that we will return to the results of this problem in the near future.)

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{-2R1+R2} \begin{bmatrix} 1 & 0 \\ ? & ? \end{bmatrix} \xrightarrow{R2/(-3)} \begin{bmatrix} 1 & 0 \\ ? & ? \end{bmatrix} \xrightarrow{-R2+R1} \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}$$

28. Repeat the same kind of work for the following:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{-3R1+R2} \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix} \xrightarrow{R2/(-4)} \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix} \xrightarrow{-2R2+R1} \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}$$

3.6 Clarification of Notation

Besides being described by its size, a matrix is often notated by an italicized capital letter. Below are several matrices that you will use in the next five problems.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix}, D = \begin{bmatrix} 3 & 2 & 2 \\ 1 & 0 & 4 \\ 2 & 2 & 0 \\ 0 & 1 & 7 \end{bmatrix}$$

29. Give the sizes of each matrix.

30. What system of linear equations does C represent?

31. What system of linear equations does D represent?

32. Use the following operations to change the matrix A. How do you interpret A as a system of linear equations and the matrix that results from these operations? **Operation 1:** $-4R1 + R2$; **Operation 2:** $R2/(-3)$; **Operation 3:** $-2R2 + R1$

33. Perform the same three operations on B.

3.7 Gauss-Jordan Elimination: Identifying Free Variables

From this point forward, we will abandon the use of x, y and z and use subscript notation that you will likely find convenient. All of the systems of linear equations that appear in this chapter are solved using Gauss-Jordan elimination, which means every time we are introduced a system, and are asked to find its set of solutions, we will be writing the system as an augmented matrix and employing elementary row operations to obtain what is known as the **reduced row echelon form** or RREF of the matrix. The big skill to master is the ability to take the matrix in this final form and write the solution set of the system. The important things are pointed out throughout the chapter through a succession of examples.

Example 4: Solve the system

$$\begin{aligned} 2x_1 - 5x_2 &= 8 \\ 3x_1 + 9x_2 &= -12 \end{aligned}$$

solution:

$$\begin{bmatrix} 2 & -5 & 8 \\ 3 & 9 & -12 \end{bmatrix} \xrightarrow[R1 \leftrightarrow R2]{R2/3} \begin{bmatrix} 1 & 3 & -4 \\ 2 & -5 & 8 \end{bmatrix} \xrightarrow{-2R1+R2} \begin{bmatrix} 1 & 3 & -4 \\ 0 & -11 & 16 \end{bmatrix}$$
$$\xrightarrow{R2/(-11)} \begin{bmatrix} 1 & 3 & -4 \\ 0 & 1 & \frac{-16}{11} \end{bmatrix} \xrightarrow{-3R2+R1} \begin{bmatrix} 1 & 0 & \frac{4}{11} \\ 0 & 1 & \frac{-16}{11} \end{bmatrix}$$

We had not yet met an example in which switching the equations mattered. That translates into one of the elementary row operations here. I try avoiding fractions wherever possible, so I observed that by dividing $R2$ by 3, I would obtain a 1 without introducing fractions if I, next, switched the two rows. Why is this permissible? Switching two rows is indicated by the symbol $\leftrightarrow$. Following this, it is the usual exercise of row reducing the augmented matrix, obtaining the solution $x_1 = \frac{4}{11}$ and $x_2 = \frac{-16}{11}$. Sometimes you cannot avoid fractions. In this section, we are going to view the solution sets as **column vectors**, the brackets indicating that the vector is a 2×1 matrix. Here, we let

$$\mathbf{x} = (x_1, x_2),$$

which, as a column vector, is written this way:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

The solution to the system is $\mathbf{x} = \begin{bmatrix} \frac{4}{11} \\ -\frac{16}{11} \end{bmatrix}$.

Example 5: Solve the system:

$$\begin{aligned} x_1 + 2x_2 + \ \ x_3 &= 1 \\ x_1 + \ \ x_2 + \ \ x_3 &= 0 \\ 2x_1 + \ \ x_2 + 2x_3 &= 1 \end{aligned}$$

solution not fully shown: I'll let you do the work here own your own: Write the augmented matrix for the system in example 5. Use Gauss-Jordan elimination to show that the matrix, writing down the operations performed, is $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

This system is *inconsistent*, meaning there is no solution. The third row reads $0 = 1$.

Example 6: Suppose row operations produced the following matrix (rref):

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

What is the solution set for the system of linear equations represented by this matrix?

solution: When you do not obtain a leading 1 where you expect it to show up, this means that the variable associated with the column with no leading 1 is a **free variable**, meaning the variable can be any real number. The third column in the above matrix does not have a leading 1, and so x_3 is the free variable, so let $x_3 = a$, where a is any real number. The above matrix can be translated literally as

$$\begin{aligned} 1x_1 + 0x_2 + 1x_3 &= 0 \\ 0x_1 + 1x_2 + 0x_3 &= 0 \\ 0x_1 + 0x_2 + 0x_3 &= 0 \end{aligned}$$

but with $x_3 = a$, we can solve the first equation for x_1, the second equation for x_2, and form the vector $\mathbf{x}$:

$$\begin{aligned} x_1 &= -a \\ x_2 &= 0 \\ x_3 &= a \end{aligned}$$

Thus the solution to the system is the set of all vectors of the form $(-a, 0, a)$. By scalar multiplication, $(-a, 0, a) = a(-1, 0, 1)$, so the solution to the system is the set of all scalar multiples of the vector $(-1, 0, 1)$, which is a line through the origin in $\mathbb{R}^3$. Of course, I am using the space-saver notation for this set of vectors. As a column vector, $\mathbf{x} = \begin{bmatrix} -a \\ 0 \\ a \end{bmatrix}$.

Systems do not always have to have the same number of equations as unknowns, as the next few examples show.

Example 7: Solve the system:

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

solution: $\begin{bmatrix} 1 & 2 & 3 & 1 \\ 1 & 3 & 2 & 1 \end{bmatrix} \xrightarrow{-R1+R2} \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix} \xrightarrow{-2R2+R1} \begin{bmatrix} 1 & 0 & 5 & 1 \\ 0 & 1 & -1 & 0 \end{bmatrix}$

There is no leading 1 for column 3 in this case, so x_3 is a free variable here. We read the following from this last matrix:

$$\begin{aligned} x_1 + 5x_3 &= 1 \\ x_2 - \quad x_3 &= 0 \\ x_3 &= x_3 \end{aligned}$$

Solving for the variables, letting $x_3 = a$, where a is any real number, we have that the solution to this system is the set of all vectors of the form $(1-5a, a, a)$. By vector addition, $(1-5a, a, a) = (1,0,0) + (-5a, a, a)$, and by scalar multiplication, $(1,0,0) + (-5a, a, a) = (1,0,0) + a(-5,1,1)$. So the solution set consists of all vectors on the line in $\mathbb{R}^3$ through the point $(1,0,0)$ and parallel to the vector $(-5,1,1)$. (If you are a student who has studied multi-variable calculus, you should recognize the vector equation of a line here.)

Example 8: Solve the system:

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

solution not fully shown: As in example 5, I'll let you do the work here on your own: Write the augmented matrix that represents the system, and then use row operations to show that the reduced row echelon form is $\begin{bmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

The third column has no leading 1, and so x_3 is a free variable. Let $x_3 = a$ and translate the matrix to read $x_1 + 2a = -1, x_2 - a = 2, x_3 = a$, and $x_4 = 4$. Solving for each variable, we conclude that the solution set is the set of all vectors of the form $(-2a-1, a+2, a, 4)$.[1]

Example 9: Solve the system

$$\begin{aligned} 2x_1 - 5x_2 - \quad x_3 + \quad 4x_4 &= 0 \\ 3x_1 - 5x_2 - 9x_3 + 11x_4 &= 0 \\ 3x_1 - 7x_2 - 3x_3 + \quad 7x_4 &= 0 \end{aligned}$$

solution:

$\begin{bmatrix} 2 & -5 & -1 & 4 & 0 \\ 3 & -5 & -9 & 11 & 0 \\ 3 & -7 & -3 & 7 & 0 \end{bmatrix}$ is the augmented matrix for this system of linear equations.

[1]This is a line in $\mathbb{R}^4$ passing through the point $(-1,2,0,4)$ and parallel to the vector $(-2,1,1,0)$.

It is worth noting the constants zero in each equation. Such a system is known as **homogeneous**. Homogeneous systems of linear equations are an important part of understanding the nature of the solution set of any system. (We explore this in the future.) I thought I would use this example to show you my favorite matrix calculator, free online and always available. (Your graphing calculator is also capable of doing this as well.) Go to www.desmos.com and, under *math tools* click *matrix calculator*. Figure out how to create a matrix of three rows and five columns, and then use the *rref* button to see that the reduced row echelon form of the above matrix is:

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

Columns 3 and 4 have no leading 1's, so for the system here, there are *two* free variables x_3 and x_4. Let $x_3 = a$ and $x_4 = b$, where a and b are any real numbers. Written directly from the above matrix, we have:

$$\begin{aligned} x_1 - 8x_3 + 7x_4 &= 0 \\ x_2 - 3x_3 + 2x_4 &= 0 \\ x_3 &= a \\ x_4 &= b \end{aligned}$$

Solving for each variable, we see that $x_1 = 8a - 7b$, $x_2 = 3a - 2b$, $x_3 = a$ and $x_4 = b$. The solution to the system is the set of all vectors of the form $(8a - 7b, 3a - 2b, a, b)$. You might be wondering why attention is paid to these solution sets when we utilize vector addition to separate elements, and scalar multiplication to factor. We are setting ourselves up to study important spaces. Solutions to homogeneous equations are always line-like things that contain the origin. These are known as *subspaces*, the kinds of spaces we encounter everywhere we work in linear algebra. They are all possible linear combinations of a set of vectors.

$$\begin{aligned} (8a - 7b, 3a - 2b, a, b) &= (8a, 3a, a, 0) + (-7b, -2b, 0, b) \\ &= a(8, 3, 1, 0) + b(-7, -2, 0, 1) \end{aligned}$$

The solution set to this system is the set of all linear combinations of the vectors $(8, 3, 1, 0)$ and $(-7, -2, 0, 1)$, which is a plane in $\mathbb{R}^4$ that passes through the origin. Before we begin discussing these special spaces, we must understand that we have been doing matrix multiplication in an interesting way. There is much more to see about solving systems. Be sure you practice this section of problems so well that you are comfortable and polished with all of the concepts established thus far.

Solve the following systems of equations, identifying any free variables, and state the solution set *as was shown in class in column vector form*. Show the row operations in your work and use a matrix calculator to check that you have correctly reduced each matrix.

3.8 HOMEWORK 4

34.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 14 \\ 2x_1 + 5x_2 + 8x_3 &= 36 \\ x_1 - x_2 &= -4 \end{aligned}$$

35.

$$\begin{aligned} x_1 - x_2 - x_3 &= -1 \\ -2x_1 + 6x_2 + 10x_3 &= 14 \\ 2x_1 + x_2 + 6x_3 &= 9 \end{aligned}$$

36.

$$\begin{aligned} x_1 + 2x_2 + 4x_3 &= 15 \\ 2x_1 + 4x_2 + 9x_3 &= 33 \\ x_1 + 3x_2 + 5x_3 &= 20 \end{aligned}$$

37.

$$\begin{aligned} x_1 + x_2 + x_3 &= 7 \\ 2x_1 + 3x_2 + x_3 &= 18 \\ -x_1 + x_2 - 3x_3 &= 1 \end{aligned}$$

38.

$$\begin{aligned} 3x_1 + 6x_2 - 3x_3 &= 6 \\ -2x_1 - 4x_2 - 3x_3 &= -1 \\ 3x_1 + 6x_2 - 2x_3 &= 10 \end{aligned}$$

39.

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

40.

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 7 \\ + 2x_2 + 2x_3 &= 11 \\ 2x_1 + 4x_2 + 3x_3 &= 18 \end{aligned}$$

41.

$$\begin{aligned} x_1 + x_2 - 3x_3 &= 10 \\ -3x_1 - 2x_2 + 4x_3 &= -24 \end{aligned}$$

42.

$$\begin{aligned} x_1 + 2x_2 - x_3 - x_4 &= 0 \\ x_1 + 2x_2 + x_4 &= 4 \\ -x_1 + 2x_2 + 2x_3 + 4x_4 &= 5 \end{aligned}$$

43.

$$\begin{aligned} x_2 + 2x_3 &= 7 \\ x_1 - 2x_2 - 6x_3 &= -18 \\ -x_1 - x_2 - 2x_3 &= -5 \\ 2x_1 - 5x_2 - 15x_3 &= -46 \end{aligned}$$

44.

$$\begin{aligned} x_1 + x_2 &= 2 \\ 2x_1 + 3x_2 &= 3 \\ x_1 + 3x_2 &= 0 \\ x_1 + 2x_2 &= 1 \end{aligned}$$

45. Show that the set of all vectors of the form (x, y) is the set of all linear combinations of $(1, 0)$ and $(0, 1)$.

4 SPEAKING LINEAR ALGEBRA, Part 1

4.1 the Set of all Possible lineAr combiNations of a set of vectors

"Linear algebra" translates as "line-like relationships." Think back to when you were in elementary school. At some point in time, you heard "two distinct points determine a line" and "three non-collinear points determine a plane." These are true statements if you are in the plane or in three-dimensional space, the two main geometric spaces made relatively comfortable and familiar by your mathematical training. The root of understanding the geometric aspect of this course is in those two claims; however, we use vectors to restate those two claims, which happen to be true in any dimension. Moreover, this course *generalizes* the statements:

(1) Two points that are not the same point will determine one line: the one and only one line containing both points.

(2) Three points that are not on the same line will determine a plane: the one and only one plane containing all three points.

(3) Four points that are not on the same plane will determine a (whatever-we-call-the-next-line-like-thing-one-dimension-higher)... Our brain breaks here due to our experience as humans, but here is how we cope with this: a line is easily conceptualized by imagining a point slide infinitely in one direction (and its opposite). The next line-like thing one dimension up is a plane, and it can be imagined as a square stretching infinitely in two directions (each having an opposite direction). The next line-like thing can be imagined as a cube extending indefinitely in three directions, each with an opposite direction. And so on and so on... This will be how we define what we mean by *dimension*, which we will cover later, but it ought to feel kind of natural to you to quickly accept that a point has no dimension, a line has dimension 1, a plane has dimension 2, the line-like thing formed from a cube has dimension 3, etc...

It is my belief that those who come to a clear understanding of this subject are those who understand the contents of this particular section, and the geometric aspects that characterize linear things. Two different points determine a vector, one from one of the points to the other. Three noncollinear points determine two vectors not pointing in the same direction. Take these two geometric situations: One nonzero vector determines a line. Two nonzero vectors not pointing in the same direction determine a plane. From here, we generalize these statements. Notice how "neat" the number works out. A line is one-dimensional. A plane is two-dimensional. Then we have a line-like thing that is three-dimensional. And on and on... Here is how we will say the above statements (1), (2), (3), etc... in our class:

(1) A single nonzero vector spans a line.

(2) Two linearly independent vectors span a plane.

(3) Three linearly independent vectors span a (whatever-we-call-the-next-line-like-thing-one-dimension-higher).

(4) Four linearly independent vectors span a (whatever-we-call-the-next-line-like-thing-one-dimension-higher-than-the-previous-line-like-thing).

This section takes care of the word SPAN as it is used in linear algebra. Your job as a student in linear algebra is to do the following two things: (a) Learn the vocabulary you are taught by understanding the definitions and using the new language in both writing and speaking. (b) Connect this information to the work we do with systems of linear equations. Students that do not understand the subject do not successfully spend the necessary time to clarify and connect. Our exam culture in education, I believe, is partly to blame for this, but your brain can do far more than you might believe is possible. In order to truly master higher mathematics, many students must agree with me that you have to abandon old habits such as learning a pattern in solving problems just so you can pass an exam, or cramming information at the last minute just so you can pass an exam, etc... Old keys often do not unlock new doors. The task now is to fully understand the sentences above, but I introduce clarity in small doses. Let's take care of SPAN.

How is it possible to take a single vector and create a line from it? A single vector certainly can be visualized as being on the line it determines, but "determines a line" *mathematically* means that every point on the line can somehow be accounted for, so how is a line determined by a vector? The answer is *scalar multiplication*. Let k be any scalar (any real number). Suppose you sketch the vector $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ in $\mathbb{R}^2$, the xy-plane, but instead of sketching the vector, you only want to sketch the point at which its arrow lands when drawn in standard position, the point $(3, 2)$. Plot that point. Now consider all possible scalar multiples of this vector, denoted by $k\mathbf{v}$, sketching all points. What do you ultimately see on your sketch? You ought to see that the points start *to cover* the line $y = \frac{2x}{3}$. How and why? And, by the way, since one of the scalars is $k = 0$, the origin $(0, 0)$ is on the line, so the set of all possible scalar multiples of any nonzero vector $\mathbf{v}$ will cover a line through the origin.

Imagine two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ not pointing in the same direction. Let k_1 and k_2 represent any scalars (any real numbers). The set of all possible linear combinations of these two vectors is written

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2.$$

If we were to sketch all points where the sum of these vectors in all these combinations land, we would cover a plane. One of the possibilities is that both scalars are zero, so the origin is included. Thus, all linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$ will cover a plane through the origin.

If we used the dictionary backwards to search for a word whose meaning is "to cover" one of the words is SPAN. In this class, it is used as both a noun and a verb. It is defined initially as a noun:

SPAN: The SPAN of a set of vectors is the set of all possible linear combinations of those vectors.

(1.) The span of one nonzero vector $\mathbf{v}$ is the set of all vectors $k\mathbf{v}$, which is a line through the origin of whatever space the vector $\mathbf{v}$ is located. All scalar multiples of a vector in any Euclidean vector space will be a line through the origin in that vector space. Used as a verb, we say: The vector $\mathbf{v}$ spans a line.

(2.) The span of two nonzero vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is the set of all linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$, which is either a line through the origin, or a plane through the origin. It is a line if the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are both located on the same line (parallel), and is a plane through the origin if $\mathbf{v}_1$ and $\mathbf{v}_2$ are not on the same line. Used as a verb, we say: The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ span a line through the origin, or a plane through the origin.

(3.) To generalize these concepts, consider this fact: The span of a set of vectors is a line-like thing through an origin.

QUESTION FOR YOU: ***If we run out of names to call each line-like thing in whatever dimension we are in, would it not be smart to INVENT a name for all of these kinds of line-like things?*** The answer to this question from mathematicians is YES, and the word used, which we will define at the end of the week, is SUBSPACE. For now, however, we need to return to systems of linear equations and connect this information to the work we are currently doing with systems of equations. I start this process of connection by assigning the reading of pages 29-this page, along with working the problems in the next section for HOMEWORK 5:

4.2 HOMEWORK 5

Solve the following systems using Gauss-Jordan elimination. Express your solutions in column vector form.

46.

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

47.

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

48.

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

49.

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

50.

$$\begin{aligned} x_1 + 4x_2 - 5x_3 - x_4 &= 0 \\ x_1 + 2x_2 - 3x_3 - x_4 &= 0 \\ x_1 + 3x_2 - 4x_3 - x_4 &= 0 \end{aligned}$$

51. I should have, in class, gone over how to think about - and work out the solution to - problem 45. If this problem caused you any unease or confusion, welcome to higher mathematics. This course is your entrance and part of the training ground for higher level mathematical study. Problem 45 leads naturally to the very next section you read carefully, word for word, and worked to understand what was written there, even if you had to get together with your peers and discuss. I celebrate that. Many of you come from calculus courses in which you have worked with Euclidean vectors in the plane and in three-dimensional space. What you might understand in this class as the column vector

$$\mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

in the xy-plane, in $\mathbb{R}^2$, is written in the component vector notation $\langle 2, 3 \rangle$, or the $\mathbf{i}$, $\mathbf{j}$ **basis vector notation** $2\mathbf{i} + 3\mathbf{j}$. While the component notation is easy to understand, let's make sure we are clear about the basis vector notation. The concept of *basis* is at the most serious level of linear algebra theory, and we will cover it in a couple of weeks. Remember that, in our class, we are allowed to write the vector as $(2, 3)$, thinking of that notation as being a space-saver notation; instead of using columns, you're thinking of $(2,3)$ as a column vector, first row 2, second row 3. From problem 45, the set of all vectors of the form (x,y) represent the set of all Euclidean vectors in $\mathbb{R}^2$. By vector addition in the first step, and by scalar multiplication in the second,

$$\begin{aligned}(x,y) &= (x,0)+(0,y)\\ &= x(1,0)+y(0,1).\end{aligned}$$

The vectors $(1,0)$ and $(0,1)$ are unit vectors (vectors that have a length of one unit) that point in the positive x and positive y directions. Discuss how this relates to the content you read super carefully six times because you loved it so much section 4.1. Discuss why you would think a good name for vectors like these two would be called "basis" vectors. Why is that a decent logical word to associate with them?

52. All solutions to the systems in 46-50 are, geometrically speaking, line-like things through origins. All of the systems have something in common. What is it?

5 REDUCED ROW ECHELON FORM

5.1 Solving by Elimination (A Second Example)

Perhaps the following example will shed more light on the strategy used to solve the system of equations in the previous chapter. (This is the system from problem 9.)

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

The solution to this system is

$$\begin{aligned} x &= 1 \\ y &= 1 \\ z &= 2 \end{aligned}$$

Remember the goal here: The goal is to perform steps in which we eliminate variables so that we obtain the form of the answer, which in the strictest of forms can be written in the following way:

$$\begin{aligned} 1x + 0y + 0z &= 1 \\ 0x + 1y + 0z &= 1 \\ 0x + 0y + 1z &= 2 \end{aligned}$$

This is a special form from which you can easily read the solution to the system. We will go through some meticulous steps before introducing the matrix language and notation for what is known as *Gauss-Jordan elimination*. It is important to understand what you are doing with the equations in the system in order to dispel any strange mysteries that arise when all of this work is spoken and written in matrix form. In order to reach our goal, we take the original system of equations, work from the left to the right, column by column, obtaining a 1 where we want it, and zeros everywhere else above and below that 1. Follow carefully:

STEP 1: Acknowledging the system in the form it is given, be sure everything is lined up properly, column-wise. Starting with the first equation, we want $1x$ here, and no x variables in the other two equations. We have the x, so we use this first equation to eliminate x in the other two. This requires two operations, both of which can be executed in a single step:

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

Using the first equation, **Operation 1:** multiply it by -2 and add it to the second, and **Operation 2:** multiply it by -1 and add it to the third.
The first two operations produce the following:

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

STEP 2: Next, we want a $1y$ in the second equation. We don't have that, but we can see how easy it is to "divide" by negative three:[2]**Operation 3:** Multiply the second equation by $-1/3$:

$$\begin{aligned} 1x + 1y - 1z &= 0 \\ 0x + 1y - 1z &= -1 \\ 0x + 1y + 0z &= 1 \end{aligned}$$

STEP 3: Next, we use the second equation to eliminate y from the other two. **Operation 4:** Multiply it by -1 and add it to the first, and **Operation 5:** multiply it by -1 and add it to the third:

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

STEP 4: Finally, we want a $1z$ in third equation, and we see we already have it. So, using the third equation to eliminate z from the other two (though here, z is already gone from the first equation, so only one operation remains: Using the third equation, **Operation 6:** add it to the second to obtain the result:

$$\begin{aligned} 1x + 0y + 0z &= 1 \\ 0x + 1y + 0z &= 1 \\ 0x + 0y + 1z &= 2 \end{aligned}$$

Moving from left to right, taking one column at a time, establishing a coefficient of 1 where we need it, and then *using it* to get coefficients of 0 in the other places, we see we are eliminating the variables x, y, and z, in that order. This produces the ideal set of equations that give the solution to the equation. When we do the same work in matrix form, we use different language, sometimes to the point of forgetting what the matrices represent! It is worth repeating several times

[2] Another option would be to switch the position of the second and third equations, a legal move which does not disturb anything.

throughout the course: the fundamental problem of linear algebra is solving a system of linear equations. The above procedure is known as **Gauss-Jordan elimination**, but it is almost always carried out by way of augmented matrices. The next sections of this chapter emphasize the new vocabulary you must learn, while restricting the use of this method to n equations and n unknown variables.

5.2 Row Operations

Each of the operations carried out in the previous examples are operations on the equations, and since we are really only dealing with the coefficients and constants, rewriting this procedure using matrices is convenient. The equations become rows, and so we refer to these operations as **row operations**. A question arises: *What are all possible types of row operations?* To answer this question, consider all of the moves you are allowed to perform with a set of equations comprising a system. It is permissible to...

(1) multiply any equation by a non-zero real number and add this to another equation, (such as Operations 1, 2, 4, 5, and 6 in the previous example),

(2) multiply an equation by a nonzero constant (such as Operation 3 in the previous example), and

(3) interchange two equations.

These moves are all legal mathematical moves that do not upset the system in terms of the solutions it possesses. No solutions are added or altered to the solution set, and these operations produce a succession of increasingly simpler systems. These three kinds of moves translate to what we call **elementary row operations**. Since the rows of an augmented matrix correspond to the equations, we can describe an elementary row operation as being an operation in which you

(1) add a constant multiple of one row to another,

(2) multiply a row by a nonzero constant, and

(3) interchange any two rows.

Gauss-Jordan elimination is the process by which you reach the ideal simple system from which you directly read the solution to the system. We borrow the word *echelon* from the French to describe the staircase formation of 1's we see on the main diagonal, and the matrix that is produced from all of these operations is known as a matrix in **reduced row echelon form**. This is how the work is shown in solving a system of linear equations by Gauss-Jordan elimination: the system is written as an augmented matrix, and then altered carefully by row operations, which you are required to notate:

Example 10: Solve the system:

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

solution: Begin with the augmented matrix, and proceed with identifying row operations between each of the matrices:

$$\begin{bmatrix} 1 & 1 & -1 & 0 \\ 2 & -1 & 1 & 3 \\ 1 & 2 & -1 & 1 \end{bmatrix} \xrightarrow[-R1+R3]{-2R1+R2} \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & -3 & 3 & 3 \\ 0 & 1 & 0 & 1 \end{bmatrix} \xrightarrow{R2/(-3)} \begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$
$$\xrightarrow[-R2+R3]{-R2+R1} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \xrightarrow{R3+R2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

The solution to the system can be read from the last matrix. For now, we can write $x = 1, y = 1, z = 2$ or $(1, 1, 2)$. (We will soon be writing these as vectors in column form.)

5.3 Reduced Row Echelon Form

Think carefully about the form of the augmented matrix that is reached if a system of n equations and n unknowns has a solution: There are 1's along the main diagonal of the matrix, with 0's above and below them. In the last column, there is a column of constants. Consider the next two examples (from the first homework assignment):

Example 11: Solve the system

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

.

solution:

$$\begin{bmatrix} 1 & 3 & 2 \\ 2 & 6 & 1 \end{bmatrix} \xrightarrow{-2R1+R2} \begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & -3 \end{bmatrix}$$

After a single row operation, you ought to pause and *interpret* the augmented matrix: what system of equations does it describe?

$$\begin{aligned} x + 3y &= 2 \\ 0x + 0y &= -3 \end{aligned}$$

The second equation makes no sense mathematically: $0 = -3$. This means the system has no solution, otherwise known as *inconsistent*.

Example 12: Solve the system

$$\begin{aligned} x - 2y &= -5 \\ -3x + 6y &= 15 \end{aligned}$$

solution:

$$\begin{bmatrix} 1 & -2 & -5 \\ -3 & 6 & 15 \end{bmatrix} \xrightarrow{3R1+R2} \begin{bmatrix} 1 & -2 & -5 \\ 0 & 0 & 0 \end{bmatrix}$$

After a single operation, note the 0's in the bottom row. Unlike the previous example, this augmented matrix *does* make sense in that, mathematically speaking, $0 = 0$. As you already know, there are an infinite number of solutions to this system, but writing as your answer "*infinite number of solutions*" does not describe *what the solutions are*.

In both of the previous examples, we were unable to reach the ideal form of an augmented matrix in which successive 1's appear to the right of the 1 in the row above it. Because the row operations produce simpler versions of the system, we say that the form we are approaching is some kind of *reduced* form, characterized by the staircase appearance of 1's along the main diagonal. The ultimate form we wish to reach is known as **reduced row echelon form**, which will be abbreviated as **rref** throughout this text.

DEFINITION: A matrix is said to be in **reduced row echelon form** if it has the following four properties:

(1) If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. We call this a **leading one**.

(2) If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.

(3) In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.

(4) Each column that contains a leading 1 has zeros everywhere else in that column.[3]

Another popular name for a leading 1 is **pivot**. This is due to the fact that, in the elimination process, going from left to right, one column at a time, a leading 1 is created and then used to annihilate the other numbers in the column, as if the 1 is strapped to a chair with some kind of number gun, and chair pivots around to face up and then face down, while the 1 makes everything 0 above it and below it. It will be important to be able to identify in which columns there are no pivots (no leading 1's). The columns without pivots located in them tell you that the variable whose coefficients appear initially in the column is a **free variable**, a detail we review in the next chapter.

[3] A matrix satisfying the first three properties is said to be in *row echelon form*.

5.4 HOMEWORK 6

53. Is the matrix $\begin{bmatrix} 1 & -2 & -5 \\ 0 & 0 & 0 \end{bmatrix}$ in reduced row echelon form? Why or why not?

54. Is the matrix $\begin{bmatrix} 1 & 3 & 2 \\ 0 & 0 & -3 \end{bmatrix}$ in reduced row echelon form? Why or why not?

55. Solve the system by writing the augmented matrix of the system in **rref**.

$$\begin{aligned} x + 4y &= 2 \\ -x + 5y &= 7 \end{aligned}$$

Use elementary row operations to obtain the reduced row echelon form of each matrix.

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

57. $\begin{bmatrix} 1 & 1 & 3 \\ 2 & 3 & 7 \\ 1 & 4 & 6 \end{bmatrix}$

58. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 5 \end{bmatrix}$

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

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

61. $\begin{bmatrix} 2 & 4 & 6 & 6 \\ 2 & 5 & 7 & 1 \\ 3 & 6 & 9 & 3 \end{bmatrix}$

62. $\begin{bmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix}$

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

64. $\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$

65. $\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & -1 \\ 2 & 6 & 0 & 8 & 4 & 18 & 6 \end{bmatrix}$

66. $\begin{bmatrix} 1 & 2 & 3 & 1 \\ 3 & 5 & 8 & 2 \\ 0 & 4 & 4 & 4 \\ 1 & 1 & 2 & 0 \end{bmatrix}$

67. $\begin{bmatrix} 1 & 2 & 4 & 1 & 1 \\ 3 & 1 & 7 & -2 & 3 \\ 2 & 2 & 6 & 0 & 2 \\ 1 & 3 & 5 & 2 & 1 \end{bmatrix}$

68. Interpret the matrix in problem 66 as one representing a system of equations. Write the system of equations. Then, based on the work you did in problem 66, write the solution to the system in column vector form.

69. Repeat the same thing for the matrix in problem 67.

70. Rewrite the systems in the previous two problems in the form $A\mathbf{x} = \mathbf{b}$ form; that is, rewrite these systems as a matrix equation.

6 SPEAKING LINEAR ALGEBRA, Part 2

6.1 Span

Let $S = \{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}\}$ be a set of n vectors. The span of this set of vectors is the set of all linear combinations of the vectors $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}$. This can be notated by

$$\text{span}\{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}\} = k_1\mathbf{v_1} + k_2\mathbf{v_2} + \cdots k_n\mathbf{v_n},$$

where each k_i is any real number. If we know the vector space that can be represented by all combinations of these vectors, then we say the vectors span that vector space.

Example 13. Look again at problem 45. *Every* vector in $\mathbb{R}^2$ can be represented by (x, y), where x and y are any real numbers. Show that the vectors $\mathbf{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ span $\mathbb{R}^2$.
Solution: Apply the definition by using vector addition and scalar multiplication, as you did in 45. Reread problem 51. That work shows that the vectors span $\mathbb{R}^2$. (*This will be a class example.*)
Other examples will be given in class, but given the one example above along with the definition, what would your written work be if you wrote solutions to the following problems in the following homework assignment? (To save space, column vectors are written in the "space-saver notation" but you are welcome to write them as column vectors in your work.

6.2 HOMEWORK 7

71. Determine the span of $(2, 2)$.

72. The vectors $\mathbf{i}$ and $\mathbf{k}$ in $\mathbb{R}^3$ span what?

73. Determine the span of the vectors $(1, -1)$ and $(-3, 3)$.

74. Determine the span of the vectors $(1, -1)$ and $(2, 3)$.

75. Determine the span of the vectors $(1, 1, 0)$ and $(0, 0, 1)$.

76. Determine the span of the vectors $(1, 0, 2, 0)$ and $(0, 1, 2, 1)$.

77. Determine the span of the set $\{(1, 0), (0, 1), (1, 1), (-1, 1)\}$.

78. The set of all vectors in $\mathbb{R}^3$ of the form $(a, b, a + b)$ is spanned by what vectors?

79. The set of all vectors in $\mathbb{R}^3$ of the form (a, a, b) is spanned by what vectors?

80. The set of all vectors in $\mathbb{R}^3$ of the form (a, b, a) is spanned by what vectors?

81. The set of all vectors in $\mathbb{R}^3$ of the form $(a+b, c, a+b)$ is spanned by what vectors?

For the problems 82-88, use the matrix here:

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

82. State the size of the matrix.

83. Interpret the rows as vectors from a Euclidean vector space. Identify the vector space and describe the span of those row vectors.

84. Interpret the columns as vectors from a Euclidean vector space. Identify the vector space and describe the span of those column vectors.

85. Write A in RREF.

86. Repeat problem 83 but using the rref of A. Is your answer here the same as it was in 83?

87. Repeat problem 84 but using the rref of A. Is your answer here the same as it was in 84?

88. If A represents a system of linear equations, what is the system and what is its solution?

89. What vector in $\mathbb{R}^2$ spans the line $y = 3x$?

90. What vectors in $\mathbb{R}^3$ span the plane $z = 2x - 3y$?

6.3 Linear Dependence and Linear Independence

The following two definitions are found in most modern texts on linear algebra:

DEFINITION: A set of vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ is said to be **linearly dependent** if there exist scalars $c_1, c_2, ..., c_n$, *not all zero*, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}.$$

DEFINITION: A set of vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ is said to be **linearly independent** if the only solution to

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

is $c_1 = c_2 = \cdots = c_n = 0$; that is, if the only linear combination of the set of vectors that adds to the zero vector is the one in which all of the scalars are zero.

To help you understand why the concept is defined in the way it is, we can use an example and language with which you ought to be familiar. Suppose someone gives you the equation $z = w + 2x - y$, where each variable represents a number. This equation is written in such a way that it makes sense to say that the value of z *depends* on the values of w, x, and y. The equation, however, can be written $w = z - 2x + y$, and we can say that the value of w depends on the values of z, x, and y. Of course, we can see that each of the variables can easily be isolated on one side, which means that the value of each variable depends on the values of the others:

$$\begin{aligned} z &= w + 2x - y \\ w &= z - 2x + y \\ x &= \frac{1}{2}z - \frac{1}{2}w + \frac{1}{2}y \\ y &= z - w - 2x \end{aligned}$$

What is special about this equation? (*Hint: Think of a sentence or four in which you use the words "linear" and "combination."*) When solved for one of the variables, we see that each variable can be written as a linear combination of the other. In such a situation, mathematicians do this: Write the equation in a form in which all of the variables are located on one side of the equation, with zero obtained on the right:

$$w + 2x - y - z = 0$$

Identifying the scalars in blue and the vectors in red, if each variable is dependent on the other by way of a linear combination, then there is a linear combination of the vectors equal to zero:

$$\color{blue}{1}\color{red}{w} \color{black}{+} \color{blue}{2}\color{red}{x} \color{black}{+} \color{blue}{(-1)}\color{red}{y} \color{black}{+} \color{blue}{(-1)}\color{red}{z} \color{black}{=} \color{red}{0}$$

This is what is at the heart of a precise definition of the concept of **linear dependence**, which some books define before linear independence, so think about this for a moment: if each vector in a set of vectors can be written as a linear combination of the other vectors, then there exists scalars that cannot all be zero that you can find to write a linear combination of the vectors that is equal to the zero vector. If *none* of the vectors can be written as a linear combination of the others, then *the only linear combination of the vectors that is equal to the zero vector is one in which all the scalars are zero.* This is the reason why the definition of linear independence is worded the way it is. Here is a more direct definition, but two mentioned prior to this one are the ones that contain the mathematical work a student often has to do to prove that a set is dependent or independent.

DEFINITION: If $S = \{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}\}$ is a set of two or more vectors in a vector space V, then S is said to be a **linearly independent set** if no vector in S can be expressed as a linear combination of the other vectors in the set. A set that is not linearly independent is said to be **linearly dependent**. If S contains one and only one vector, we will agree that the set is linearly independent if and only if that vector is nonzero.

A common problem across disciplines is to take a set of vectors and determine whether one of them can be written as a linear combination of one or more of the others. If one is found, your set is a linearly dependent set; that is, it is not linearly independent.

Example 29: Suppose $S = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4}\}$, and $\mathbf{v_2} = \mathbf{v_1} - 3\mathbf{v_3}$, and there is no relationship between $\mathbf{v_4}$ and any of the other three variables. How do you justify that the set S is a linearly dependent set? Is there a linear combination of all of the vectors in the set equal to the zero vector?

solution: You justify it by using the definition for linear independence. Since one vector can be written as a linear combination of two of the other vectors, then you automatically know, by definition, that S is not linearly independent.

The relationship that was given allows us to write $\mathbf{v_1} - \mathbf{v_2} - 3\mathbf{v_3} = \mathbf{0}$. In order to include the other vector in the set, the scalar zero is appended to the vector:

$$1\mathbf{v_1} + (-1)\mathbf{v_2} + (-3)\mathbf{v_3} + 0\mathbf{v_4} = \mathbf{0}$$

The purpose of this chapter is to, of course, learn these new definitions, which means you want to practice problems in which you might derive some meaning as to *why these concepts might be important.* By far, the important skill to practice is determining whether or not a set of vectors is linearly independent. This ought to come together soon, once we finish this phase of the course in which you are learning the language of linear algebra. The first phase is learning a general method for solving any system of linear equations, introducing matrices in order to have the matrix equation that represents the system $A\mathbf{x} = \mathbf{b}$. The second phase of the course is to learn the vocabulary of linear algebra (and where we, as instructors, can repeatedly point out that matrices are the things to which we eventually return). The third phase of the course is to learn to look at matrices and see them for what they really are: **linear transformations**.

You are often asked to determine whether or not a set of vectors is linearly independent. This means you might have to examine whether or not one vector in the set is a scalar multiple of another in the set, or whether or not the vector is a linear combination of two or more vectors in the set. This takes us right back to the beginning of our course: *solving systems but taking the column perspective!* If you did not pick up on this important detail, let me remind you of it using the following three examples:

Example 30: What is the column perspective of the system below?

$$\begin{aligned} 5x_1 + 3x_2 &= 1 \\ 6x_1 + 2x_2 &= -2 \\ -x_1 + x_2 &= 3 \end{aligned}$$

solution: This was the birthplace of matrices! The column perspective reads the given system in columns, and interprets the unknowns as scalars: Reinterpreted, this system is

$$x_1 \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}.$$

Is it possible to find scalars x_1 and x_2 so that the vectors $(5, 6, -1)$ and $(3, 2, 1)$ can be scaled in such a way that, when scaled and added together, result in the vector $(1, -2, 3)$? That is one way to restate the problem. Another way is to notice that the left side of the above equation is a linear combination of the vectors $(5, 6, -1)$ and $(3, 2, 1)$ equal to the vector $(1, -2, 3)$, so the system can be reworded in the exact same style as the next example is worded, so read the next example!

Example 31: Is $\mathbf{v_1} = (1, -2, 3)$ a linear combination of $\mathbf{v_2} = (5, 6, -1)$ and $\mathbf{v_3} = (3, 2, 1)$?

explanation and solution: This question directly translates to this more mathematical question: Are there scalars a and b that can be found such that $\mathbf{v_1} = a\mathbf{v_2} + b\mathbf{v_3}$? This is the column perspective of a system of linear equations. Notice how this translates to what you already know:

$$\begin{aligned} (1, -2, 3) &= a(5, 6, -1) + b(3, 2, 1) \\ (1, -2, 3) &= (5a, 6a, -a) + (3b, 2b, b) \\ (1, -2, 3) &= (5a + 3b, 6a + 2b, -a + b) \end{aligned}$$

The last step above (third step) is the system in example 30, as

$$\begin{aligned} 1 &= 5a + 3b \\ -2 &= 6a + 2b \\ 3 &= -a + b \end{aligned}$$

Rewritten to where we solve by Gauss-Jordan elimination (view the system in example 30), we have

$$\begin{bmatrix} 5 & 3 & 1 \\ 6 & 2 & -2 \\ -1 & 1 & 3 \end{bmatrix} \xrightarrow[R1 \leftrightarrow R3]{-R3, \frac{1}{2}R2} \begin{bmatrix} 1 & -1 & -3 \\ 3 & 1 & -1 \\ 5 & 3 & 1 \end{bmatrix} \tag{5}$$

$$\xrightarrow[-5R1+R3]{-3R1+R2} \begin{bmatrix} 1 & -1 & -3 \\ 0 & 4 & 8 \\ 0 & 8 & 16 \end{bmatrix} \tag{6}$$

$$\xrightarrow[\frac{1}{8}R3]{\frac{1}{4}R2} \begin{bmatrix} 1 & -1 & -3 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \tag{7}$$

$$\xrightarrow[-R2+R3]{R2+R1} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \tag{8}$$

In a step above, three row operations were carried out. Notice the original augmented matrix is such that if you multiply the third row by -1, you end up with a 1 where you want it if you then switch the first and third rows. At the same time, the second row is divisible by 2, so why not mutliply it by one-half while you can. The remaining steps read as usual, with the last matrix showing $a = -1$ and $b = 2$. Be careful to state a meaningful answer here! YES, the vector can be written as a linear combination of the other two: *If YES is your answer, then you had better state the combination!*

$$(1, -2, 3) = -1(5, 6, -1) + 2(3, 2, 1)$$

Example 32: Determine whether or not the set $S = \{(1, -2, 3), (5, 6, -1), (3, 2, 1)\}$ is linearly independent.

solution: The solution is really the previous example! And if you look closely enough at what you can do when this question is asked, what is a way of working the problem? Notice that the question must be asked: *is it possible to write one vector in the set as a linear combination of the others?* Since we already know the answer is yes, then S is linearly dependent.

Example 33: Determine whether or not the set $W = \{(4, -3), (-2, \frac{3}{2})\}$ is linearly independent.

solution: Since one vector is a scalar multiple of another, $(4, -3) = -2(-2, \frac{3}{2})$, then W is linearly dependent. *Always try and spot whether or not two vectors in a set are scalar multiples of one another. This is often the quickest way to note that a set of vectors is linearly dependent.*

Example 34: Find the reduced row echelon form of the matrix $B = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 7 \end{bmatrix}$. Determine whether or not the column vectors of B are linearly independent.

solution: $\begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 7 \end{bmatrix} \xrightarrow{-R1+R2} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \end{bmatrix} \xrightarrow{-2R2+R1} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 3 \end{bmatrix}$

By observation, no column vectors are scalar multiples of one another. Often students assume the set is linearly independent, but when it comes to a set of three or more vectors, you must ask whether one vector can be a linear combination of all, or some, of the other vectors in the set. In this case, since three vectors comprise the columns of B, then ask whether $(1,1)$ is a linear combination of $(2,3)$ and $(4,7)$.[4] If so, then

$$\begin{aligned} (1,1) &= a(2,3) + b(4,7) \\ &= (2a, 3a) + (4b, 7b) \\ &= (2a + 4b, 3a + 7b), \end{aligned}$$

which means that you must solve the system:

$$\begin{aligned} 2a + 4b &= 1 \\ 3a + 7b &= 1 \end{aligned}$$

Solving by Gauss-Jordan elimination

$$\begin{bmatrix} 2 & 4 & 1 \\ 3 & 7 & 1 \end{bmatrix} \xrightarrow{\frac{1}{2}R1} \begin{bmatrix} 1 & 2 & \frac{1}{2} \\ 3 & 7 & 1 \end{bmatrix} \xrightarrow{-3R1+R2} \begin{bmatrix} 1 & 2 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \end{bmatrix} \xrightarrow{-2R2+R1} \begin{bmatrix} 1 & 0 & \frac{3}{2} \\ 0 & 1 & -\frac{1}{2} \end{bmatrix},$$

the system has solution $a = \frac{3}{2}$ and $b = -\frac{1}{2}$, which means *yes*, the vector $(1,1)$ is a linear combination of the other two vectors: $(1,1) = \frac{3}{2}(2,3) - \frac{1}{2}(4,7)$. Notice the connection between the two problems here in that it was not necessary to do the work to conclude that the column vectors of B are linearly dependent: That work was done when we wrote the matrix in reduced row echelon form. When B is written in reduced row echelon form, one interpretation of that matrix is that it might possibly represent a system of linear equations whose solution, when interpreted in the column perspective, reveals that the column vectors are related through a linear combination $-2(1,1) + 3(2,3) = (4,7)$. Thus, the columns are linearly dependent.

For a moment, reconsider a factorization of B by looking back at example 20 in Chapter 8, and apply that process to obtain the CR-factorization of B. When examining the columns of B, one by one, and understanding the reduced form interpretation, the first column is $(1,1)$, and when you move to the second column, it is not a scalar multiple of the first, so these two are the columns of C:

$$B = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} & & \\ & & \end{bmatrix}$$

[4] Or ask whether the second column is a linear combination of the first and third, or whether the third column is a linear combination of the first and second...

The matrix R is found by understanding the column perspective: viewing the three columns of B as linear combinations of the columns of C: The first column of B is $1(1,1) + 0(2,3)$:

$$B = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & & \\ 0 & & \end{bmatrix}$$

The second column of B is $0(1,1) + 1(2,3)$:

$$B = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & \\ 0 & 1 & \end{bmatrix}$$

The third column of B is $-2(1,1) + 3(2,3)$:

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

Has anyone caught on to what this CR-factorization might be about? The matrix on the left, C, is a matrix whose column vectors are linearly independent. What is R? (Don't worry; we will return to this repeatedly!)

Example 35: Determine whether or not the columns of $M = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 2 & 0 & 6 \end{bmatrix}$ are linearly independent.

solution: If you suspect that any column vector *cannot* be written as a linear combination of the other two, then you are probably thinking there is a strong possibility that the vectors are linearly independent: **If a set of vectors is linearly independent, then the only linear combination of the vectors which add to the zero vector is the one in which all the scalars are zero.** In other words, to show that a set is linearly independent, use row operations to show that $c_1 = c_2 = c_3 = 0$; that is, if you set up the system of equations associated with the question posed, then show that the only solution to the system is the zero vector, which means the vectors are linearly independent:

$$c_1 \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{9}$$

$$(c_1 + c_2 + c_3, 2c_1 + 3c_2 + 5c_3, 2c_1 + 6c_3) = (0,0,0). \tag{10}$$

Notice, in (12.9), this sentence can be written after thinking *whether or not there exists a linear combination of the vectors equal to the zero vector.* This leads directly to (12.10), which is the system below:

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \\ 2c_1 + 3c_2 + 5c_3 &= 0 \\ 2c_1 + 6c_3 &= 0 \end{aligned}$$

I leave it to you, the student, to show that $\begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 3 & 5 & 0 \\ 2 & 0 & 6 & 0 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, which means that the only solution to the system is $c_1 = 0, c_2 = 0$, and $c_3 = 0$, and so the columns of M are linearly independent.

6.4 HOMEWORK 8

91. Determine whether or not the rows of M in example 35 are linearly independent.

92. Is the set $S = \{(2, -6), (-3, 9)\}$ linearly independent?

93. Determine whether or not the set $\{(2, 1), (1, 3), (1, 1)\}$ linearly independent.

94. Is the set $\{(1, 2, 1), (0, 0, 0), (0, 1, -1)\}$ linearly independent?

95. Show that the set $\{(1, 1), (-1, 1)\}$ is linearly independent and spans $\mathbb{R}^2$.

96. Let A be a square matrix. If its columns are linearly independent, then what is A in reduced row echelon form?

97. Let A be a square matrix. If its rows are linearly independent, then what is A in reduced row echelon form?

98. Write the system of linear equations represented by the augmented matrix $\begin{bmatrix} 1 & 1 & 2 & 1 \\ 2 & 1 & 3 & 1 \\ 1 & 2 & 3 & 2 \end{bmatrix}$. Write the matrix equation $A\mathbf{x} = \mathbf{b}$.

99. Write the matrix in rref, and, from there, write the solution to the system in column vector form.

100. If $\mathbf{v} = \mathbf{u} + \mathbf{w}$, prove that the set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent.

7 THE ROW AND COLUMN SPACES OF A MATRIX

The concepts of *span* and *linear independence/dependence* allow us to briefly introduce two important spaces associated with every matrix: the **row space** of a matrix, denoted by row(A) and the **column space** of a matrix, denoted by col(A). Later on in the course, we find these two spaces have a lot to do with the big picture of linear algebra. The purpose of this section is to become familiar with the two definitions and to further explore the concepts just learned, giving some meaning to the CR-factorization of a matrix and the rref of a matrix.

7.1 The Row Space of A

DEFINITION: Let A be an $m \times n$ matrix. The rows of A can be thought of as vectors in $\mathbb{R}^n$, and the row space of A, denoted by row(A) is the span of these row vectors; that is, the row(A) is the set of all linear combinations of the rows of A.

Example 36: Determine the row space of the matrix $A = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 5 & 1 \end{bmatrix}$.

solution: Always note the size of a matrix first. Here, A is a 2×3 matrix. The rows of A are vectors in $\mathbb{R}^3$. The span of the rows of A can be directly written by applying the definition of span:

$$\text{row}(A) = a(1, 2, 1) + b(4, 5, 1),$$

where a and b are any scalars. We could easily leave this here, but we want a much better description of this space, and there are two better descriptions. The above linear combination is an algebraic description that can be improved, and, when possible, a geometric description is ideal. By inspection, we can see that the two vectors are linearly independent, so the row space is a plane through the origin in $\mathbb{R}^3$. Two linearly independent vectors determine a plane. That's a good geometric description. What about a better algebraic description? Using scalar multiplication and vector addition, we can do the following:

$$\begin{aligned} a(1,2,1) + b(4,5,1) &= (a, 2a, a) + (4b, 5b, b) \\ &= (a + 4b, 2a + 5b, a + b) \end{aligned}$$

This allows us to give the description as the set of all vectors in $\mathbb{R}^3$ of the form $(a + 4b, 2a + 5b, a + b)$. But, if you were to describe this set of vectors this way, chances are the description is still not good. Look closely at what you have *algebraically*: Notice that if you chose values for a and b and noticed the vectors that you end up with as a result, you can probably conclude in words what all of the vectors in the set are like: For example, if $a = 1$ and $b = 1$, then you have $(5, 7, 2)$. If $a = 3$ and $b = 1$, then you have $(7, 11, 4)$. If $a = 1$ and $b = 2$, then you have $(9, 12, 3)$. What do each of these vectors have in common?

Look closely as it is something you can spot in the matrix: No matter what the first two components are, *the third component is always the second component minus the first component.* This would mean a far better description of this set of vectors would be "the set of vectors in $\mathbb{R}^3$ of the form $(a, b, b-a)$." THAT algebraic description is far superior to the first one. In fact, you can see that $(2a+5b)-(a+4b) = a+b$. How, then, do you arrive at the best algebraic description?? Take the matrix and rewrite it in reduced row echelon form, or, if you are picking up on the CR-factorization, described in detail in the previous section, factor the matrix this way. Either route you take illuminates the strength and power of row reduction. This matrix in reduced row echelon form is

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

Furthermore, the span of the rows of the matrix in rref is

$$a(1,0,-1)+b(0,1,1) = (a,0,-a)+(0,b,b) = (a,b,b-a).$$

This gets to the heart of a serious issue in the work force when mathematics is used, and how valuable mathematics can be: you want to be efficient, clear, and save time and energy. Notice what kind of power linear algebra gives you here! *Students, WAKE UP from the spell that our institutions of education have cast upon you with our exam culture!* You can make a serious living if you are catching on to what this course might give you if you are able to pry yourself from an old approach to mathematics and adopt a new approach, perhaps one more critical and intellectually curious! That, I know, is much easier said than done.

In your work, consider giving both types of descriptions. A good answer to this kind of problem in writing is to say the row(A) is the set of all vectors in $\mathbb{R}^3$ of the form $(a, b, b-a)$, a plane through the origin.[5]

7.2 The Column Space of A

DEFINITION: Let A be an $m \times n$ matrix. The columns of A can be thought of as vectors in $\mathbb{R}^m$, and the column space of A, denoted by col(A) is the span of these column vectors; that is, the col(A) is the set of all linear combinations of the columns of A.

Example 37: Determine the column space of the matrix $A = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 5 & 1 \end{bmatrix}$.

[5]To those of you taking calculus, this would be the set of all points (x, y, z) where $z = y - x$, which means the equation of the plane is $x - y + z = 0$, a plane passing through the origin in $\mathbb{R}^3$ with normal vector $\mathbf{n} = \langle 1, -1, 1 \rangle$. This connection ought to be made!

not a good solution: The answer that lacks critical thought but would be an accurate application of the definition would be to say that the column space of A is

$$\text{col}(A) = \text{span}\{\begin{bmatrix}1\\4\end{bmatrix}, \begin{bmatrix}2\\5\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\} = a(1,4) + b(2,5) + c(1,1),$$

where $a, b,$ and c are any scalars. This is a terrible answer. Why? The vectors are linearly dependent, which means you do not need all three! Furthermore, the span of the columns is all of $\mathbb{R}^2$. Why? How? Just as we did with the row space, we want to refine our answer here. If you can recognize that the span of the columns is all of $\mathbb{R}^2$ then you can immediately write that down. To show that this is the case, we need to do something similar like we did with the row space. But we only have row operations, and something we will show later is that row operations do not change the row space, but they almost always change the column space. So we introduce the idea of the **transpose of a matrix**, where we can perform row operations on that matrix to reduce a matrix that yields the best description of the column space.

DEFINITION: Let A be an $m \times n$ matrix. The **transpose** of A, denoted by A^T, is the $n \times m$ matrix formed from A where the first row of A^T is the first column of A, the second row of A^T is the second column of A, etc... In other words, switch the rows and columns. (Later, we will state some specific properties of matrices, including transpose properties, but we motivate its existence here in this section.

The transpose of A in example 36 is $A^T = \begin{bmatrix}1 & 4\\2 & 5\\1 & 1\end{bmatrix}$. When written in rref, A^T is $\begin{bmatrix}1 & 0\\0 & 1\\0 & 0\end{bmatrix}$, and the span of the rows of this matrix in rref is the column space of A. Examples 36 and 37 can be summed up in the following:

$$\text{row}(A) = \text{vectors of the form } (a, b, b-a) \text{ and } \text{col}(A) = \mathbb{R}^2,$$

since the span of the rows of the rref of A^T is $\mathbb{R}^2$, another plane through the origin.

7.3 HOMEWORK 9

Determine the row and column spaces of each matrix. To summarize this section, to find the best algebraic descriptions of these spaces, take both A and A^T and write these two matrices in rref. The row space of A is the span of the rows of the rref of A, and the column space of A is the span of the rows of the rref of A^T. We will do a few more examples in the lecture to clarify some other subtle details. Go back to problems 46-50, and think of writing each of those systems in $A\mathbf{x} = \mathbf{b}$ form. Select the matrix A from each for problems 101-105.

101. The matrix A in the system in problem 46.

102. The matrix A in the system in problem 47.

103. The matrix A in the system in problem 48.

104. The matrix A in the system in problem 49.

105. The matrix A in the system in problem 50.

106. $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

107. $A = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}$

108. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

109. $A = \begin{bmatrix} 1 & 1 & 2 & 0 \\ 3 & 7 & 10 & -4 \end{bmatrix}$

110. Create a 4×3 matrix with no zero entries that has linearly dependent column vectors. Determine the row and column spaces of this matrix.

The following problems are intended to prepare you for the first exam.

111. Use Gauss-Jordan elimination to solve the system, writing your solution in column vector form as was shown repeatedly in the lectures.

$$\begin{aligned} x_1 + x_2 + 5x_3 &= 1 \\ x_1 - x_2 - x_3 &= 1 \\ 2x_1 + x_2 + 7x_3 &= 2 \end{aligned}$$

112. Write the matrix equation $A\mathbf{x} = \mathbf{b}$ of the system in the previous problem.

113. Write the system in problem 111 as a linear combination of vectors.

114. Determine the row and column spaces of the matrix A from problem 112.

115. Give an example of a set of three linearly independent vectors in $\mathbb{R}^4$, and precisely define the meaning of this concept.

116. Is $(4, 5)$ a linear combination of $(1, -2)$ and $(-2, 2)$? Prove or disprove your answer.

117. Give an example of a set of nonzero, distinct linearly dependent vectors in $\mathbb{R}^3$ and then find the span of this set of vectors.

118. The set of all vectors of the form $(a, b, a + b, a - b)$ is spanned by what set of vectors?

7.4 EXTRA PROBLEMS TO PRACTICE FOR FUN

119. Is $(-1, 4, 11, 1)$ a linear combination of $(1, 2, 3, 1), (1, 1, 0, 2)$, and $(-1, 0, 1, 1)$? Prove or disprove your answer.

120. Explain why all four vectors mentioned in the previous problem are linearly independent or linearly dependent by first defining the two concepts, and then justifying your explanation.

121. Let $M = \begin{bmatrix} 1 & 2 & 7 & -3 \\ 3 & -1 & 7 & 5 \\ 2 & 1 & 8 & 0 \\ -1 & 4 & 5 & -9 \end{bmatrix}$. Find the reduced row echelon form of M. Be sure to show all row operations in your work.

122. Determine the row space of M in the previous problem.

123. Explain whether or not the columns of M in problem 121 are linearly independent or linearly dependent. Give a solid justification for your answer by demonstrating that you know the definitions of these terms.

124. Let W be the set of all vectors in $\mathbb{R}^4$ of the form $(a, b, 2a, 2b)$. Find a set of vectors that span W.

125. If the matrix M in problem 121 represents an augmented matrix for a system of linear equations, write the system of equations.

126. Find the solution to this system of linear equations in 125.

127. Referring to the matrix M in problem 121, find the solution to the homogeneous system $M\mathbf{x} = \mathbf{0}$ by first writing the system of linear equations.

128. Write the reduced row echelon form of $C = \begin{bmatrix} 1 & 3 & 4 & 1 & 0 \\ 1 & 2 & 3 & 0 & 1 \end{bmatrix}$.

129. Determine the row space of the matrix C in the previous problem.

130. Solve the system by viewing the work you did in 128:

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

131. Solve the system by Gauss-Jordan elimination, showing the row operations clearly, and then writing your solution in column vector form.

$$\begin{aligned} x_1 \; -_2 \; + x_3 &= 3 \\ 2x_1 - 3x_2 + x_3 &= 5 \end{aligned}$$

132. For the system in 131, write the matrix equation for the system $A\mathbf{x} = \mathbf{b}$. Identify the matrix A, which you will carry into the last several problems in this section.

133. Give both an algebraic and geometric description of the row space of A.

134. Give both an algebraic and geometric description of the column space of A.

135. Here is a definition: The **null space** of a matrix A, denoted by null(A), is the set of all vectors $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$; in other words, the null space of A is the set of all solutions to the homogeneous system $A\mathbf{x} = \mathbf{0}$. In more practical terms, it is all of the vectors that "get sent to the zero vector" when A "acts on the vectors." This language has real world meaning: If a matrix is used to re-organize data, the vectors in its null space will likely be the cause of any loss of that data. We will soon think of a matrix as being a special kind of transformation: a *linear* transformation, one that acts upon a vector. Linear transformations are at the heart and soul of this course. Find the null space of the matrix A and then look at the solution to the system in 131. What do you see?

136. Write the transpose of A (as you might have done in 134, perhaps?). Find the null space of the transpose, denoted by null(A^T).

137. Make one more observation about the matrix A: What is its size? Look at the vectors in each of the four spaces: row space, column space, null space of A, and null space of A^T. What can you say about the size of A and each of these four spaces?

8 INVERTIBLE SQUARE MATRICES

A traditional course in linear algebra focuses heavily on the theory of **square matrices**, which are matrices of size $n \times n$; that is, matrices that have the same number of rows as columns. These correspond with systems of n equations and n variables. If A is an $n \times n$ matrix, it might be **invertible**. In this section, we explore finding the inverse of a matrix, if it exists, make a connection between the row operations in Gauss-Jordan elimination and matrix multiplication, and study the algebra of solving $A\mathbf{x} = \mathbf{b}$.

When the word *singular* is used in mathematics, it is describing something that is strange, unusual, peculiar, striking, often deviating from the usual or expected. The point $x = 0$ is referred to as a *singularity* for the function $f(x) = \frac{1}{x}$ because the function is defined for all other values of x with the exception of $x = 0$. The center of a black hole, a point of infinite density and gravity where no object can escape, is referred to as a *singularity* in Albert Einstein's 1915 *Theory of General Relativity.* You are initially taught methods to solve systems of equations and experience the best case scenario of that method working well, producing the ideal result, which means that the system, when written in matrix form, contained a matrix A that was invertible. That feels like the usual. It is somewhat rare to obtain "no solution" or "infinite number of solutions" when using a method to solve the system, which means that the matrix equation of the system contains a matrix A that is a bit peculiar, or *singular.* This is the way to think of the word as it is used in the definition of an invertible matrix.

DEFINITION: Let A be a square matrix of order n, which is a fancy way of saying let A be an $n \times n$ matrix, also known as a square matrix. If there exists an $n \times n$ matrix B such that $AB = BA = I_n$, then A is said to be an **invertible** matrix or a **nonsingular** matrix, and B is the **inverse** of A. In this event, we write $B = A^{-1}$. In other words, if A is invertible, we denote its inverse by A^{-1}. If no such inverse exists, then A is said to be **singular**.

8.1 How to Find the Inverse

In matrix form, a system of linear equations can be interpreted as $A\mathbf{x} = \mathbf{b}$, where the goal is to find the unknown $\mathbf{x}$ such that when it is multiplied by A on the left, it results in $\mathbf{b}$. The quickest way to obtain that, in general, by hand, involves employing elementary row operations. The same logic applies to the question of finding an inverse. According to the definition, the inverse is an unknown matrix X with the property that $AX = I$. It makes sense to create an augmented matrix and use row operations to find X.

Example 38: Assume that $A = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix}$ is invertible. Find A^{-1}.

solution: Since the definition requires that $AA^{-1} = I$, then, in matrix form,

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

Form the augmented matrix and, as you did with solving systems, use elementary row operations to write it in reduced row echelon form:

$$\begin{bmatrix} 1 & 3 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{bmatrix} \xrightarrow{-R1+R2} \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{bmatrix} \xrightarrow{-3R2+R1} \begin{bmatrix} 1 & 0 & 4 & -3 \\ 0 & 1 & -1 & 1 \end{bmatrix}$$

The matrix at the end shows that $A^{-1} = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix}$. *I strongly encourage you to start paying close attention to **definitions** in this course.* If we check this answer and honor the definition, we would check and make sure that $AA^{-1} = I$ and $A^{-1}A = I$:

$$AA^{-1} = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4-3 & -3+3 \\ 4-4 & -3+4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and

$$A^{-1}A = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 4-3 & 12-12 \\ 1+1 & -3+4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Example 39: Solve the system by finding the inverse of its coefficients matrix, if it exists:

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 11 \\ 2x_1 - x_2 + 3x_3 &= 7 \\ 7x_1 - 3x_2 - 2x_3 &= 2 \end{aligned}$$

solution: As a matrix equation $A\mathbf{x} = \mathbf{b}$, where $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \\ 7 & -3 & -2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 11 \\ 7 \\ 2 \end{bmatrix}$, the solution to the system will be $\mathbf{x} = A^{-1}\mathbf{b}$. We set up the augmented matrix to find the inverse of A, thinking $AA^{-1} = I$:

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

Next, we apply row operations to obtain the reduced row echelon form of this augmented matrix. This one is a nice exercise in arithmetic:

$$\begin{bmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 7 & -3 & -2 & 0 & 0 & 1 \end{bmatrix} \xrightarrow[-7R1+R3]{-2R1+R2} \begin{bmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & -5 & 5 & -2 & 1 & 0 \\ 0 & -17 & 5 & -7 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[-\frac{1}{17}R3]{-\frac{1}{5}R2} \begin{bmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & \frac{2}{5} & -\frac{1}{5} & 0 \\ 0 & 1 & -\frac{5}{17} & \frac{7}{17} & 0 & -\frac{1}{17} \end{bmatrix}$$

$$\xrightarrow[-R2+R3]{-2R2+R1} \begin{bmatrix} 1 & 0 & 1 & \frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 1 & -1 & \frac{2}{5} & -\frac{1}{5} & 0 \\ 0 & 0 & \frac{12}{17} & \frac{1}{85} & \frac{1}{5} & -\frac{1}{17} \end{bmatrix}$$

$$\xrightarrow{\frac{17}{12}R2} \begin{bmatrix} 1 & 0 & 1 & \frac{1}{5} & \frac{2}{5} & 0 \\ 0 & 1 & -1 & \frac{2}{5} & -\frac{1}{5} & 0 \\ 0 & 0 & 1 & \frac{1}{60} & \frac{17}{60} & -\frac{1}{12} \end{bmatrix}$$

$$\xrightarrow[R3+R2]{-R3+R1} \begin{bmatrix} 1 & 0 & 0 & \frac{11}{60} & \frac{7}{60} & \frac{1}{12} \\ 0 & 1 & 0 & \frac{5}{12} & \frac{1}{12} & -\frac{1}{12} \\ 0 & 0 & 1 & \frac{1}{60} & \frac{17}{60} & -\frac{1}{12} \end{bmatrix}$$

The solution to the system is found by acknowledging

$$A^{-1} = \begin{bmatrix} \frac{11}{60} & \frac{7}{60} & \frac{1}{12} \\ \frac{5}{12} & \frac{1}{12} & -\frac{1}{12} \\ \frac{1}{60} & \frac{17}{60} & -\frac{1}{12} \end{bmatrix},$$

and the solution to the system is obtained by multiplying this inverse by $\mathbf{b}$. I'll leave it to you to arrive at the solution $x_1 = 3, x_2 = 5$, and $x_3 = 2$, often written in the space-saver notation as $(3, 5, 2)$, or as the column vector $\mathbf{x} = \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix}$.

Though the arithmetic in the previous example is busy, it is worth doing a few of these kinds of problems and then know when it is time to turn to using technology. I chose this example because I thought the inverse of A was worth playing with for a brief moment: Note the denominator of each fraction is either 12 or 60. If we rewrite the matrix such that every denominator is the same, then we can write

$$A^{-1} = \begin{bmatrix} \frac{11}{60} & \frac{7}{60} & \frac{5}{60} \\ \frac{25}{60} & \frac{5}{60} & -\frac{5}{60} \\ \frac{1}{60} & \frac{17}{60} & -\frac{5}{60} \end{bmatrix},$$

and we can factor out the scalar 1/60 from the matrix, and then write

$$A^{-1} = \frac{1}{60}\begin{bmatrix} 11 & 7 & 5 \\ 25 & 5 & -5 \\ 1 & 17 & -5 \end{bmatrix}.$$

The number 60 is a special number related to the original matrix A, a future topic. This is an example to which we will return.

I recall doing work like in the previous example as a student, questioning the use of this method for solving the system. While it is important to appreciate solving a system by Gauss-Jordan elimination, or converting it to a matrix equation, obtaining an inverse to multiply on both sides of that equation, thereby solving it, I noticed that the work done by hand is the same for the two different methods: had we solved the previous system by elimination, we would have performed the same elementary row operations. My instructor at the time convinced me that obtaining the inverse was sometimes an important task. I recall asking ”why can't we do both at the same time?” and, every time I had to solve a system of equations, I did my work like I show in the next example:

Example 40: Solve the system:

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

solution by a former smartass:

$$\begin{bmatrix} 1 & 2 & 3 & 1 & 0 \\ 1 & -1 & -3 & 0 & 1 \end{bmatrix} \xrightarrow{-R1+R2} \begin{bmatrix} 1 & 2 & 3 & 1 & 0 \\ 0 & -3 & -6 & -1 & 1 \end{bmatrix} \tag{11}$$

$$\xrightarrow{-\frac{1}{3}R2} \begin{bmatrix} 1 & 2 & 3 & 1 & 0 \\ 0 & 1 & 2 & \frac{1}{3} & -\frac{1}{3} \end{bmatrix} \tag{12}$$

$$\xrightarrow{-2R2+R1} \begin{bmatrix} 1 & 0 & -1 & \frac{1}{3} & \frac{2}{3} \\ 0 & 1 & 2 & \frac{1}{3} & -\frac{1}{3} \end{bmatrix} \tag{13}$$

This work shows that the solution to the system is $\mathbf{x} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, and the inverse of $A = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$ is $A^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{bmatrix}$. If you want to confirm the solution is $(-1, 2)$, then write the matrix equation of the system, and find the product $A^{-1}\mathbf{b}$.[6]

[6] In my later years now as a professor of mathematics, I sometimes ask students to solve by elimination, and then find the inverse of the coefficients matrix, and check the solution they found by demonstrating the solution to the equivalent matrix equation.

8.2 A Matrix that Integrates a Function

Let's return to an example of a matrix that differentiates a function:

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

differentiates all functions of the form $ae^{2x}\cos 3x + be^{2x}\sin 3x$. Suppose we put all functions of this form inside a set called $\mathbb{F}$. All functions in $\mathbb{F}$ can be represented by a vector in the form $\begin{bmatrix} a \\ b \end{bmatrix}$. These functions are vectors because you can add them and scalar multiply them. When the matrix D operates on all vectors of the form (a, b), where $a \in \mathbb{R}$ and $b \in \mathbb{R}$, the result is also a vector in the same family of functions. This is a property[7] that distinguishes important operators: If $\mathbf{x} \in \mathbb{F}$ then $D\mathbf{x} \in \mathbb{F}$; that is the derivative of any constant times $e^{2x}\cos 3x$ plus any constant times $e^{2x}\sin 3x$ is always a constant times $e^{2x}\cos 3x$ plus a constant times $e^{2x}\sin 3x$:

$$D\mathbf{x} = \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2a + 3b \\ -3a + 2b \end{bmatrix}$$

8.3 HOMEWORK 10

138. Differentiate the function as you did in a first course in calculus to show that the derivative of $ae^{2x}\cos 3x + be^{2x}\sin 3x$ is $(2a+3b)e^{2x}\cos 3x + (-3a+2b)e^{2x}\sin 3x$.

139. Find the inverse of D using the method shown in this chapter. (The method demonstrated in the examples in this chapter is often called the *Matrix Inversion Algorithm* or *Inversion Algorithm*.)

140. Recall the torture in integrating functions of this form. Use integration by parts to evaluate $\int e^{2x}\cos 3x\,dx$.

141. The function you integrated in the previous problem was $f(x) = e^{2x}\cos 3x$. This can be represented by the vector $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Find $D^{-1}\mathbf{v}$. Does this agree with your answer in 143? What does this inverse matrix *not* do that it ought to do?

Use the following matrices for exercises 142-149.

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

142. Find A^{-1}.

[7] This is called **closure**, a property we will define and use soon; As an example of something with which you should be familiar, the set of real numbers is *closed under addition* because adding two real numbers results in a real number.

143. Find B^{-1}.

144. Let $C = AB$. What is C?

145. Find C^{-1}.

146. Find the matrix product $A^{-1}B^{-1}$

147. Find the matrix product $B^{-1}A^{-1}$.

148. Let $D = BA$. Find D.

149. Find D^{-1}.

150. **True** or **False**: $\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -1 \\ 3 & -1 & 4 \end{bmatrix}$ is a singular matrix.

151. **True** or **False**: $\begin{bmatrix} 2 & -1 & 3 \\ 1 & 0 & -4 \\ 1 & -1 & 7 \end{bmatrix}$ is a singular matrix.

Solve the following systems in 152-154 by Gauss-Jordan elimination. Then check the solution to the system by converting it to a matrix equation, finding the inverse of the coefficients matrix, and using it to justify the solution.

152.

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

153.

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

154.

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

155. Derive a formula *on your own* to find the inverse of any 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. What condition, according to the formula you derived, must be met in order for an inverse to exist?

156. Test your formula out on the matrices A and B that precede problem 142.

9 THE ALGEBRA OF MATRICES

Here is a new notation to take care of representing any matrix. Let A be an $m \times n$ matrix, a matrix with m rows and n columns. The notation a_{ij} represents an element of the matrix A in the ith row and jth column. Here are a couple of examples:

Example 41: Let $A = \begin{bmatrix} 2 & 1 & 4 & 7 \\ 3 & 5 & 9 & 6 \end{bmatrix}$ and let $B = \begin{bmatrix} 0 & 2 \\ 1 & 3 \\ 9 & 5 \end{bmatrix}$. Identify the following elements: $a_{13}, a_{22}, b_{11},$ and b_{31}.

solution: a_{13} is the element of A in the first row, third column of A, so $a_{13} = 4$. Similarly, $a_{22} = 5, b_{11} = 0,$ and $b_{31} = 9$.

The following will be used to represent $m \times n$ matrices A and B:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \text{ and } B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

9.1 Matrices are Vectors

Matrices are vectors, meaning matrices of the same size can be added and scalar multiplied; that is, if A and B are both $m \times n$ matrices (matrices of the same size), and if k is any scalar, then elements in the same spots get added together, and all elements of a matrix are multiplied by k:

$$A + B = \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

$$kA = \begin{bmatrix} ka_{11} & \cdots & ka_{1n} \\ \vdots & \ddots & \vdots \\ ka_{m1} & \cdots & ka_{mn} \end{bmatrix}$$

The next section, in my opinion, contains far too much information that we never used; however, in mathematics, you have to at least state these properties. In all honesty, matrix multiplication is the important operation.

9.2 Properties of Matrix Addition and Multiplication

Let $A, B,$ and C be matrices of the same size. Let m and n be any scalars. Let O be the $m \times n$ matrix called the **zero matrix**, where all its entries are zero. The following are properties of matrices are established:

1. Matrix addition is commutative: $A + B = B + A$
2. Matrix addition is associative: $(A + B) + C = A + (B + C)$
3. $A + O = A$
4. $m(A + B) = mA + mB$
5. $(m + n)A = mA + nA$
6. $m(nA) = (mn)A$
7. Matrix multiplication is associative: $(AB)C = A(BC)$
8. $A(B + C) = AB + AC$
9. $(B + C)A = BA + CA$
10. $m(AB) = (mA)B = A(mB)$
11. $I_m A = A = AI_n$, where I_j is the $j \times j$ identity matrix.
12. $(A^T)^T = A$
13. $(A + B)^T$
14. $(mA)^T = mA^T$
15. $(AB)^T = B^T A^T$
16. $(A^{-1})^{-1} = A$
17. $(AB)^{-1} = B^{-1}A^{-1}$
18. $(A^T)^{-1} = (A^{-1})^T$

The above properties are actually *theorems* that can be proved, but you can do that in another course. We will prove *some* of these as we journey through different topics in the course. Keep in mind that A^T is the **transpose** and not "A to the power of T." Similarly, A^{-1} is the **inverse** of A not "A to the power of -1." It is matrix multiplication that I would like to bring into focus carefully: What is vital to understand in our course is what is *not* true about matrices that is true about real numbers and the algebra with which you are familiar. I emphasize this in both words and symbols, with examples in the lecture that have meaning not only to the following, but make sense when you see meaning associated with matrix multiplication:

- Matrix Multiplication is NOT commutative: $AB \neq BA$.
- If $AB = AC$, then you CANNOT conclude that $B = C$.
- $AB = O$ DOES NOT imply that $A = O$ or $B = O$.

Example 42: Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and let $B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. PROOF-WRITING ADVICE: *Whenever you are asked to prove that something is **not equal to** something else, all you have to do is present a single example. In other words, if you are ever asked to prove that matrix multiplication is not commutative, this is one of many examples you could give.* Show that $AB \neq BA$.

solution: I will leave it to the student to show that $AB = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and that $BA = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

9.3 The Dot Product

Here is the general mathematical definition of the dot product of any two vectors in $\mathbb{R}^n$. (In the lecture, we can discuss the origin of the dot product, if there is interest.)

DEFINITION: Let $\mathbf{u} = (u_1, u_2, ..., u_n)$ and $\mathbf{v} = (v_1, v_2, ..., v_n)$ be two vectors from the same space $\mathbb{R}^n$. The **dot product**, also known as the **inner product** or the **scalar product** of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u} \cdot \mathbf{v}$, is defined as the real number

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

Example 43: Find the dot product of the vectors $(1, 3)$ and $(-4, 2)$ in $\mathbb{R}^2$.
solution: By definition,

$$\begin{aligned} (1,3) \cdot (-4,2) &= (1)(-4) + (3)(2) \\ &= -4 + 6 \\ &= 2. \end{aligned}$$

Example 44: Find the dot product of the vectors $(1, 4)$ and $(-4, 2)$ in $\mathbb{R}^2$.
solution: By definition,

$$\begin{aligned} (1,4) \cdot (-4,2) &= (1)(-4) + (4)(2) \\ &= -4 + 8 \\ &= 4. \end{aligned}$$

The two dot products in the above examples is what defines the multiplication in the matrix equation

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

Example 45: Let $\mathbf{u} = (2, 1, -1, 2), \mathbf{v} = (0, 1, -2, -2)$. Find $\mathbf{u} \cdot \mathbf{v}$.
solution: By definition,

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= (2, 1, -1, 2) \cdot (0, 1, -2, -2) \\ &= (2)(0) + (1)(1) + (-1)(-2) + (2)(-2) \\ &= 0 + 1 + 2 - 4 \\ &= -1. \end{aligned}$$

9.4 Matrix Multiplication

Using the matrices that arise from the previous examples, coupled with the definition of the dot product, we see that matrix product $A\mathbf{x}$ is nothing more than the dot products of the vectors shown in the two dot product examples, where the matrix A contains components of vectors in its rows, while the matrix on the right is really a vector with the same number of components, but is written in column form, making matrix multiplication the art of computing the dot products of vectors. I have sometimes said to classes, when introducing a matrix product AB, "take the dot product of a row of A with a column of B." You can see how these dot products are carried out in the previous examples:

$$\begin{bmatrix} \color{red}{1} & \color{red}{3} \\ 1 & 4 \end{bmatrix} \begin{bmatrix} \color{red}{-4} \\ \color{red}{2} \end{bmatrix} = \begin{bmatrix} \color{red}{2} \\ \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ \color{blue}{1} & \color{blue}{4} \end{bmatrix} \begin{bmatrix} \color{blue}{-4} \\ \color{blue}{2} \end{bmatrix} = \begin{bmatrix} \color{red}{2} \\ \color{blue}{4} \end{bmatrix}$$

If this is how matrix multiplication is to be executed, then it must be the case that *the number of columns of the matrix on the left must equal the number of rows of the matrix on the right*. Why? The dot product between two vectors is defined only for vectors with the same number of components. This gives us a *matrix notation* for the dot product of two vectors. For example. the two dot products in the matrix product above can be written in vector form or matrix form:

$$(1, 3) \cdot (-4, 2) = 2$$

can be written

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

But... don't forget the column perspective of the system of linear equations! That perspective gives us another way to view matrix multiplication. All of this starts with this system of equations:

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

Remember that the column perspective means that the system can be interpreted as a linear combination of two vectors, where x_1 and x_2, the unknown variables, are the scalars needed to be found so that

$$x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

Notice that this means we can take the matrix product we were looking at, which was

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

and we can carry out the product by understanding that the rightmost matrix tells us what to do with the columns of the leftmost matrix; that is, the above product can be interpreted as "multiply the first column of A by -4, and add that to twice the second column of A." This gives a stronger meaning to the fact that, in order to multiply two matrices such as AB, the number of columns of A must equal the number of rows of B.

$$\begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \end{bmatrix} = -4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ -4 \end{bmatrix} + \begin{bmatrix} 6 \\ 8 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

I think knowing and using *both* are extremely useful. There are actually *five* different ways to compute a matrix product, and each method has its own use. At the moment, the two methods discussed serve us well in our work with systems of linear equations.

DEFINITION: If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then the **matrix product** AB is the $m \times n$ matrix whose entries are determined as follows: To find the entry in row i and column j of AB, single out row i from the matrix A and column j from the matrix B and, considering these as vectors, compute the dot product of row i and column j.

9.5 HOMEWORK 11

Use these matrices to study the example and work exercises 157-163.

$$A = \begin{bmatrix} 1 & -3 \\ 1 & -2 \end{bmatrix}, B = \begin{bmatrix} -6 \\ -2 \end{bmatrix}, C = \begin{bmatrix} -3 \\ 7 \end{bmatrix}, D = \begin{bmatrix} -2 & 3 \\ -1 & 1 \end{bmatrix}, E = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, F = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

Example 46: Find AF.
solution:

$$AF = \begin{bmatrix} 1 & -3 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + (-3)(0) & 1 \cdot 3 + (-3)(1) \\ 1 \cdot 1 + (-2)(0) & 1 \cdot 3 + (-2)(1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

157. AB

158. DA

159. EA

160. DC

161. FE

162. EF

163. FEA

164. Solve the following system by first writing the system as an augmented matrix, followed by Gauss-Jordan elimination. Be sure to record the row operations between changing matrices. There ought to be only two elementary row operations performed.

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

165. You might not be aware of the fact that as you solved the system in the previous problem, you did matrix multiplication. Observe how the augmented matrix changed after the first row operation. Which of the exercises in 157-163 do you see something similar?

166. Write the system as a *matrix equation* $A\mathbf{x} = \mathbf{b}$.

167. Take the matrix equation in 166, $A\mathbf{x} = \mathbf{b}$, and multiply both sides of the equation *on the left* by the matrix D. In other words, find $DA\mathbf{x} = D\mathbf{b}$. What is the result of doing this? (Convert the matrix equation back to a system of equations.)

For the last three exercises, use $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 0 & 2 \end{bmatrix}$.

168. Find A^{-1}, if it exists.

169. Write A^T and find $A^T A$ and AA^T.

170. Determine the row and column spaces of A.

The exercises in Homework 11 above focus on matrix multiplication. As you can see from the long list of properties of matrices, we have much more to explore with those properties in terms of the transpose and the inverse. We will cover more of those as time passes, but I figured I would put all of these properties in one place. Each of them are theorems, which means we can prove them. A later section in the course dives into the art of proof, an important ability for any future mathematician.

10 THE NULL SPACES OF A MATRIX A

Let A be an $m \times n$ matrix, a matrix with m rows and n columns. If A is multiplied on the right by a vector $\mathbf{x}$, denoted by $A\mathbf{x}$, then $\mathbf{x}$ must be a vector from $\mathbb{R}^n$, meaning $\mathbf{x}$ must have n components. The resulting product is a vector from $\mathbb{R}^m$. You should reason why these are the outcomes immediately before proceeding to the next section. The following picture might help, as the sizes of each matrix in the product is written as subscripts:

$$A_{m\times n}\mathbf{x}_{n\times 1} = \mathbf{b}_{m\times 1}$$

10.1 null(A)

DEFINITION: The null space of A is the set of all vectors $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$; that is, if A is an $\times n$ matrix, the null space is every vector in $\mathbb{R}^n$ that, when multiplied on the right of A results in the zero vector in $\mathbb{R}^m$. The null space of A is denoted by null(A).

Example 46: Find the null space of $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 7 & 11 \end{bmatrix}$.

Solution: Solve the system $A\mathbf{x} = \mathbf{0}$. Since A has four columns, then there are four unknowns. The system, if you wish to write it as a system of linear equations, would be the following:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + \ \ 4x_4 &= 0 \\ 4x_1 + 3x_2 + 7x_3 + 11x_4 &= 0 \end{aligned}$$

The system is solved by row reducing the augmented matrix

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 0 \\ 4 & 3 & 7 & 11 & 0 \end{bmatrix}.$$

But most people row reduce A and know how to write the solution to the system, since the rightmost column of zeros will never change. Nevertheless, row-reducing the augmented matrix gives us the following:

$$\begin{aligned} \begin{bmatrix} 1 & 2 & 3 & 4 & 0 \\ 4 & 3 & 7 & 11 & 0 \end{bmatrix} &\xrightarrow{-4R_1+R_2} \begin{bmatrix} 1 & 2 & 3 & 4 & 0 \\ 0 & -5 & -5 & -5 & 0 \end{bmatrix} \\ &\xrightarrow{-\frac{1}{5}R_2} \begin{bmatrix} 1 & 2 & 3 & 4 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix} \\ &\xrightarrow{-2R_2+R_1} \begin{bmatrix} 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix} \end{aligned}$$

That last matrix reads

$$
\begin{aligned}
x_1 + x_3 + 2x_4 &= 0 \\
x_2 + x_3 + x_4 &= 0
\end{aligned}
$$

and the free variables can be identified as x_3 and x_4. Solving for all four variables in terms of these free variables gives us:

$$
\begin{aligned}
x_1 &= -x_3 - 2x_4 \\
x_2 &= -x_3 - x_4 \\
x_3 &= x_3 \\
x_4 &= x_4
\end{aligned}
$$

This solution ought to be written in the column vector form as usual. I will write it as part of the answer to the question:

$$\text{The null}(A) \ = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \text{ where } x_3 \text{ and } x_4 \text{ are any scalars.}$$

NOTE: The null space is the span of two linearly independent vectors. The null space is all linear combinations of these two vectors. The null space is a plane through the origin in $\mathbb{R}^4$. The vectors $(-1, -1, 1, 0)$ and $(-2, -1, 0, 1)$ span this plane, which is the set of all vectors in $\mathbb{R}^4$ of the form $(-a - 2b, -a - b, a, b)$.

10.2 null(A^T)

The null space of the transpose of an $m \times n$ matrix A is, by definition, the set of all vectors $\mathbf{y}$ in $\mathbb{R}^m$ such that $A^T\mathbf{y} = \mathbf{0}$, the zero vector in $\mathbb{R}^n$. It is denoted by null(A^T). The null space of the transpose of a matrix is the **left null space of** A, and this next example will illustrate why it has that name.

Example 47: Let A be the 4×3 matrix $\begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 7 \\ 1 & 1 & 1 \\ 3 & 4 & 6 \end{bmatrix}$. Determine the null($A$) and null ($A^T$).

Solution process (full solution not shown): The process of doing this really requires writing both matrices in rref, but you don't want to forget the *definition* of these null spaces. They are all solutions to homogeneous systems for which A and its transpose represent in terms of the coefficients of unknowns. While you may learn quickly how to get these null spaces, take a moment to write the systems of linear equations you are actually solving, just to make sure you keep to the definition:

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 7 \\ 1 & 1 & 1 \\ 3 & 4 & 6 \end{bmatrix} \text{ and its transpose is } A^T = \begin{bmatrix} 1 & 1 & 1 & 3 \\ 2 & 3 & 1 & 4 \\ 4 & 7 & 1 & 6 \end{bmatrix}.$$

The two systems you are finding solutions to are these:

$$\begin{aligned} x_1 + 2x_2 + 4x_3 &= 0 \\ x_1 + 3x_2 + 7x_3 &= 0 \\ x_1 + \ \ x_2 + \ \ x_3 &= 0 \\ 3x_1 + 4x_2 + 6x_3 &= 0 \end{aligned} \qquad \begin{aligned} y_1 + \ \ y_2 + y_3 + 3y_4 &= 0 \\ 2y_1 + 3y_2 + y_3 + 4y_4 &= 0 \\ 4y_1 + 7y_2 + y_3 + 6y_4 &= 0 \end{aligned}$$

I will leave it to the student to find these null spaces, but state these answers here:

$$\text{null}(A) = x_3 \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} \text{ and null}(A^T) = y_3 \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \end{bmatrix} + y_4 \begin{bmatrix} -5 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

It is worth taking the time out to think about the size of the matrix A: 4×3. Notice what vectors in $\mathbb{R}^3$ are in the null space of A, and that if we take any of those vectors $\mathbf{x}$ and multiply $A\mathbf{x}$, we will get the zero vector in $\mathbb{R}^4$. Here is the reason why the null space of the transpose is really the **left null space of** A: Suppose we grab one vector from the null space of the transpose: if we choose $y_3 = y_4 = 1$, we obtain the vector $\mathbf{y} = (-7, 3, 1, 1)$ from the null space of A^T. A property of the transpose is the following: $(AB)^T = B^T A^T$. The vector $\mathbf{y}$ is the column vector shown in the product here: $A^T \mathbf{y} = \mathbf{0}$:

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

If we transpose that matrix equation, then using the properties of transposes, we get the following:

$$\begin{aligned}(A^T\mathbf{y})^T &= \mathbf{0}^T\\ \mathbf{y}^T A &= \mathbf{0}^T\end{aligned}$$

which makes sense if we rewrite the matrix product and check the sizes to make sure it makes sense:

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

The product on the previous page reads as a 3×4 multiplied by a 4×1 equals a 3×1, which makes sense. And the product above reads as a 1×4 multiplied by a 4×3 equals a 1×3, which makes sense. The sizes work and, as you can see above, a vector being multiplied on the left side of A. All vectors in the null space of the transpose of A are the vectors that result in the zero vector when multiplied on the left of A; hence, the name *left null space of A*.

10.3 Revisiting Span and Linear Independence

10.3.1 How to Read and Use Definitions

I restate the following definitions for a strong reason: to emphasize their use in the homework. *You cannot simply ignore definitions in mathematics and hope for the best when answering test questions that require them!* Definitions often literally tell you what to do in proof writing, a major skill you develop in higher mathematics. Many students fail out of math programs simply because they do not learn definitions, nor care to use them. Please read the following carefully and model your work after the work shown on the next few pages. This is a first course in mathematics where many of you find that you have to abandon an old approach to learning mathematics, and you have to put much more thought into the subject, not merely figure out a pattern to a problem so you can get the answer.

DEFINITION: A **linear combination** of two vectors $\mathbf{u}$ and $\mathbf{v}$ is an expression of the form

$$k_1\mathbf{u} + k_2\mathbf{v},$$

where k_1 and k_2 are real numbers (scalars).

A **linear combination** of three vectors $\mathbf{v}_1, \mathbf{v}_2,$ and $\mathbf{v}_3$ is an expression of the form

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3,$$

where $k_1, k_2,$ and k_3 are real numbers (scalars).

A **linear combination** of n number of vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ is an expression of the form

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots k_n\mathbf{v}_n,$$

where each k_i is a real number (scalar).

DEFINITION: The **span** of a set of vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ is the set of all possible linear combinations of the vectors.

DEFINITION: Let V be a vector space. If all vectors in V can be written as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$, then these vectors span V. If the vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$, span V, then all vectors in V can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$.

DEFINITION: The vectors $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ are **linearly dependent** if there exist scalars $c_1, c_2, ..., c_n$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}.$$

The vectors are **linearly independent** if the only solution to that linear combination is $c_1 = c_2 = \cdots = c_n = 0$.

Study the following. These are frequently asked questions on all exams in any course in linear algebra. Answering them requires paying attention to the definitions, which tell you exactly what to do.

1. The vector space $\mathbb{R}^3$ is the set of all vectors of the form (x, y, z), where $x, y,$ and z are any scalars, and

   $$\begin{aligned}(x, y, z) &= (x, 0, 0) + (0, y, 0) + (0, 0, z) \text{ (by vector addition)} \\ &= x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1) \text{ (by scalar multiplication)},\end{aligned}$$

   ,
   which shows that all vectors in $\mathbb{R}^3$ can be written as a linear combination of the vectors $(1, 0, 0), (0, 1, 0),$ and $(0, 0, 1)$. Therefore, the vectors $(1, 0, 0), (0, 1, 0),$ and $(0, 0, 1)$ span $\mathbb{R}^3$.

2. Let W be all vectors in $\mathbb{R}^3$ of the form (a, b, a), where a and b are any scalars. By vector addition and scalar multiplication, $(a, b, a) = (a, 0, a) + (0, b, 0) = a(1, 0, 1) + b(0, 1, 0)$, so all vectors in W can be written as linear combinations of $(1, 0, 1)$ and $(0, 1, 0)$. The vectors $(1, 0, 1)$ and $(0, 1, 0)$ span W.

3. The span of $\mathbf{u}$ and $\mathbf{v}$ is all linear combinations of the form $a\mathbf{u} + b\mathbf{v}$, where a and b are any scalars.

4. The span of $(1, 1)$ and $(-3, -3)$ is the set of all linear combinations of the form $a(1, 1) + b(-3, -3)$, but it is easy to spot that $(-3, -3) = -3(1, 1)$, which means $1(-3, -3) + 3(1, 1) = (0, 0)$, so the vectors are linearly dependent. In this case of two linearly dependent vectors span a line through the origin, so we only need one vector since all scalar multiples of one non-zero vector spans a line. The span of these two vectors is the line containing all vectors of the form (x, x), where x is any scalar.

5. How is it possible that the two vectors $(2, 3)$ and $(-1, 1)$ span $\mathbb{R}^2$? The Euclidean vector space $\mathbb{R}^2$ is the set of all vectors of the form (x, y), where x and y are any scalars. (*This is where people fall apart at first... there is always a way:* ***what does the definition say?***) Definitions tell you precisely what to do in mathematics. Vectors span V if every vector in V can be written as a linear combination of the vectors. Simply put things where they belong in that definition, read it, and it tells you what to do: If $(2, 3)$ and $(-1, 1)$ span $\mathbb{R}^2$, then every vector in $\mathbb{R}^2$ can be written as a linear combination of these two vectors: That would mean $(x, y) = a(2, 3) + b(-1, 1)$ and you can solve the system for a and b to obtain the way that every vector could be written this way:

$$\begin{aligned} 2a - b &= x \\ 3a + b &= y \end{aligned}$$

The solution is $a = \frac{x+y}{5}$ and $b = \frac{2y-3x}{5}$. For example, if you have the point $(-1, 6)$, then it is not too hard to find that $(-1, 6) = 1(2, 3) + 3(-1, 1)$. But the fact that you are able to write that all vectors of the form $(x, y) = (\frac{x+y}{5})(2, 3) + (\frac{2y-3x}{5})(-1, 1)$, then you have shown that all vectors in $\mathbb{R}^2$ can be written as a linear combination of $(2, 3)$ and $(-1, 1)$, so the vectors span $\mathbb{R}^2$.

6. Let $\mathbf{u}_1 = \mathbf{v} - \mathbf{w}, \mathbf{u}_2 = 2\mathbf{v}$, and $\mathbf{u}_3 = \mathbf{w} - 4\mathbf{v}$, all vectors from a vector space V. Are the vectors $\mathbf{u}_1, \mathbf{u}_2$, and $\mathbf{u}_3$ linearly independent or linearly dependent? **What does the definition say?** What possible scalars a, b, c can you find such that

$$a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{u}_3 = \mathbf{0}?$$

This is a homework problem, but you should be able to find that it is possible to find scalars that are not all zero such that the linear combination of the vectors is the zero vector. This can be strongly supported by the mathematics that leads to

$$1\mathbf{u}_1 + \frac{3}{2}\mathbf{u}_2 + 1\mathbf{u}_3 = \mathbf{0}.$$

and this means that the vectors are linearly dependent.

10.4 HOMEWORK 10/21:

171. Let $\mathbf{u}_1 = \mathbf{v} - \mathbf{w}, \mathbf{u}_2 = 2\mathbf{v}$, and $\mathbf{u}_3 = \mathbf{w} - 4\mathbf{v}$, all vectors from a vector space V. Are the vectors $\mathbf{u}_1, \mathbf{u}_2$, and $\mathbf{u}_3$ linearly independent or linearly dependent?

172. The dot product of two vectors $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ is denoted and defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3.$$

When the dot product is zero, then the vectors are perpendicular to one another. Let W be the set of all vectors in $\mathbb{R}^3$ that are perpendicular to the vector $(-1, 1, 1)$. Think geometrically what you would expect to see, perhaps the place where you could see all vectors perpendicular to this given vector, then find a set of vectors that you can show span W. (We will use this as a class example, but only start it so you finish successfully.)

173. What vectors span the set of all vectors in $\mathbb{R}^3$ whose first and third components are equal?

174. Are the vectors $(1, 1, 2), (3, 0, 1)$, and $(5, 2, 5)$ linearly independent or dependent?

175. Show that the span of $(1, 1, 2), (3, 0, 1)$, and $(5, 2, 5)$ is the same as the span of $(1, 1, 2)$ and $(3, 0, 1)$.

176. The span of $(3, 1)$ is ?

177. If the zero vector is included in a set of vectors, why is the set linearly dependent?

178. Determine the null space of A and the null space of A^T in example 47.

179. Determine the row space of A in example 47.

180. Determine the column space of A in example 47.

Use this matrix for problems 181-183: $A = \begin{bmatrix} 1 & 2 & 6 \\ 2 & 2 & 4 \\ 3 & 3 & 6 \end{bmatrix}$

181. Determine the null space of A.

182. Determine the null space of A^T.

183. Determine the row and column spaces of A.

Use the matrix $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ for 184 and 185:

184. Determine both null spaces of A.

185. Determine the row and column spaces of A.

186. If a matrix A is 4×7, to what Euclidean vector space $\mathbb{R}^n$ does its null space vectors belong? Explain.

187. Determine the null spaces of the 2×3 matrix whose entries are all 1.

188. Determine the null space and the row space of the 2×2 matrix whose entries are all 1. Sketch these two spaces on graph paper.

189. What is the null space of the matrix $\begin{bmatrix} 1 & 2 \\ 4 & 8 \end{bmatrix}$?

190. Find the null spaces of the matrix $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$.

11 SUBSPACES

11.1 Three Examples to Keep Close By

Here is a matrix: $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

The null space of this matrix is all $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$, and since A is 2×2, we know that $\mathbf{x}$ is from $\mathbb{R}^2$, the set of all Euclidean vectors with two components. Finding the null space of A, by definition, means solving the system

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_1 + x_2 &= 0 \end{aligned}$$

To most of you, this means writing the *augmented* matrix

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

in reduced row echelon form (by a single operation, in this case):

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

The solution, in column vector form, is

$$\mathbf{x} = x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

otherwise known as *all scalar multiples of a single vector*, a line through the origin in $\mathbb{R}^2$ that has equation $y = -x$. The row space of A is the span of the rows of A, but it is easy here to see that the row vectors of A are linearly dependent. (Why?) The row space is the span of $(1, 1)$ which is a line through the origin in $\mathbb{R}^2$, the line $y = x$.

Here, the transpose of A happens to be the same as A; that is, $A^T = A$, so it is not all that difficult to say that the null space of A^T is the same as the null space for A, and its row space is the column space of A. So... null space, null space of transose, row space, column space... *These are all line-like things passing through the origin* because they are always all linear combinations of a set of vectors, meaning they are the span of a set of vectors. Line-like things that pass through origins are known as **subspaces of a vector space**, not to be confused with the word *subsets*.

Let $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Its transpose is $A^T = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}$.

A is 2×3 and we know the null space of A will be the set of all vectors $\mathbf{x} \in \mathbb{R}^3$ such that $A\mathbf{x} = \mathbf{0}$.

A^T is 3×2 and we know the null space of A^T will be the set of all vectors $\mathbf{y} \in \mathbb{R}^2$ such that $A^T\mathbf{y} = \mathbf{0}$.

The two null spaces are found by solving the systems:

$$\begin{aligned} 0x_1 + 1x_2 + 0x_3 &= 0 \\ 0x_1 + 0x_2 + 2x_3 &= 0 \end{aligned}$$

$$\begin{aligned} 0x_1 + 0x_2 &= 0 \\ 1x_1 + 0x_2 &= 0 \\ 0x_1 + 1x_2 &= 0 \end{aligned}$$

The null(A) is the span of $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and null(A^T) contains one element only: $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

The row space of A is the span of the nonzero rows of the reduced row echelon form of A, so the row space is the span of $(0, 1, 0)$ and $(0, 0, 1)$.

The row space of A^T is the span of the nonzero rows of the rref of A^T, so row(A^T) is all of $\mathbb{R}^2$. The row space of the transpose of A is the column space of A.

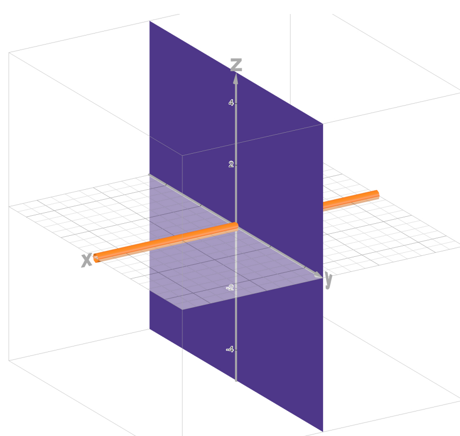


Figure 1: null(A)(x-axis), row(A)(yz-plane)

One more example: Let $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 3 & 3 \\ 3 & 3 & 3 \end{bmatrix}$. Its transpose is $A^T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 3 \\ 2 & 3 & 3 \end{bmatrix}$.

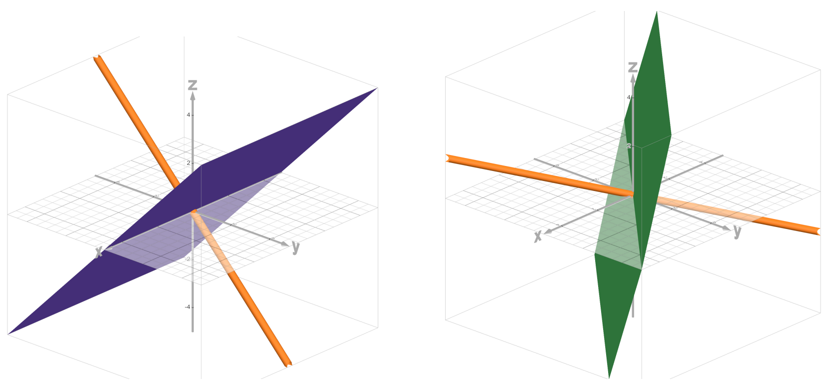
The rref of A is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. The rref of A^T is $\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$.

The null space of A is the span of $(0, -1, 1)$. The null space of A^T is the span of $(3, -3, 1)$.

The row space of A is the span of $(1, 0, 0)$ and $(0, 1, 1)$, which is the set of all vectors in $\mathbb{R}^3$ that have the same y and z components; that is, the set of all vectors of the form (a, b, b), and this is a plane passing through the origin. Its equation is $-y + z = 0$.

The row space of A^T is the span of $(1, 0, -3)$ and $(0, 1, 3)$, the set of all vectors in $\mathbb{R}^3$ of the form $(a, b, 3b - 3a)$. Its equation is $3x - 3y + z = 0$.

If you are thinking deeply about this material, you might have noticed something about these spaces. They come in pairs that literally *compliment* each other. Why? They are *complements*. (Note the change in spelling.)



(a) row(A) and null (A) (b) col(A) and null(A^T)

Figure 2: Pictures from the matrix A

Something extraordinary is going on here in the two pictures in Figure 2. The most important content in an entire course on introductory linear algebra can be found in the pictures in Figure 2. There is also a connection between (a) the picture on the left and (b) the picture on the right. What we will be doing in the next two weeks is completing this picture. All of the major results in linear algebra can be summarized in this picture! We will return often to the matrices so that we can visualize its pieces like this. Notice that all of these spaces are line-like things through the origin... They are all *subspaces*.

11.2 Closure

Add any two real numbers and you get another real number. Multiply a real number by a scalar (another real number) and you get a real number. These operations do not produce any kinds of numbers that are different than real numbers; that is, addition and multiplication of real numbers produce real numbers. No results will be *outside* of the set of all real numbers. In mathematics, we say that the set of real numbers is **closed under addition** and **closed under multiplication**. This is the property of **closure**:

Let $\mathbf{u}$ and $\mathbf{v}$ be any two vectors from a vector space V. If $\mathbf{u}+\mathbf{v} \in V$,, then V is **closed under vector addition**. If k is any scalar, and $\mathbf{u}$ is any vector in V, and if $k\mathbf{u} \in V$, then V is **closed under scalar multiplication**. In order to define a line-like thing in a vector space that passes through the origin of that vector space (like all lines and planes through $(0, 0, 0)$ in $\mathbb{R}^3$), you have to know that the line-like thing (1) contains the zero vector, and (2) is closed under both vector addition and scalar multiplication.

11.3 A Subspace of a Vector Space

Here is the technical mathematical definition of a line-like thing containing an origin.

DEFINITION: Let S be a nonempty subset of a vector space V, written $S \subset V$ or $S \subseteq V$. S is called a **subspace of** V if $\mathbf{0}_V \in S$, where $\mathbf{0}_V$ is the zero vector in V. and, for any two vectors $\mathbf{u}, \mathbf{v} \in S$ and any scalar k, S is closed under vector addition, which means $\mathbf{u} + \mathbf{v} \in S$, and S is closed under scalar multiplication, which means $k\mathbf{u} \in S$.

To show a subset S is a subspace of V, you must show three things. You must show that...

(1) ...the zero vector is in S,

(2) if you add any two vectors in S, the result is also a vector in S, and

(3) if you multiply any vector in S by any scalar, the result is also a vector in S.

This definition starts us on the trajectory of a theory we will develop that leads us to the "big picture of linear algebra" that involves any $m \times n$ matrix. That theory begins with **THEOREM 1** on the next page. For now, two examples proving that some set is a subset of a vector space:

Example 51: Let U be the set of all vectors in $\mathbb{R}^2$ of the form $(a, 2a)$. Show that U is a subspace of $\mathbb{R}^2$.

solution: Before writing the proof, which is a piece of writing, more than the typical mathematics problem, be clear that all the vectors in U are also vectors in $\mathbb{R}^2$: Vectors of the form $(a, 2a)$ in which the second component is twice the first component are also contained in $\mathbb{R}^2$, so $U \subseteq \mathbb{R}^2$. *Then you must do some preliminary thinking or brief notetaking where you jot down what you need in a formal explanation.* **Be clear about the three things you need to do in order to show** U is a subspace of $\mathbb{R}^2$: (1) Show the zero vector is in S by explaining its use or showing it is possible. Then introduce any two vectors in S and any scalar k. Show that the sum of the two vectors and any scalar multiple of a vector in S produces a result also in S. Here is what a formal proof would look like and how I am expecting you to write sentences as you would in any English course:

PROOF: Let U be the set of all vectors of the form $(a, 2a)$. If $a = 0$, then $(0, 0)$ is a vector in U. Let $\mathbf{v} = (v, 2v)$ and let $\mathbf{w} = (w, 2w)$ be any two vectors in U, and let k be any scalar.

$$\begin{aligned}\mathbf{v} + \mathbf{w} &= (v, 2v) + (w, 2w)\\ &= (v + w, 2v + 2w)\\ &= (v + w, 2(v + w)),\end{aligned}$$

so U is closed under vector addition.

$$\begin{aligned}k\mathbf{u} &= k(u, 2u)\\ &= (ku, 2(ku)),\end{aligned}$$

so U is closed under scalar multiplication. U is a subspace of $\mathbb{R}^2$. $\square$

Example 52: Let W be the set of all vectors in $\mathbb{R}^3$ of the form (a, a, b), where $a, b \in \mathbb{R}$. Prove, or disprove, that W is a subspace of $\mathbb{R}^3$.

solution: If all vectors in W are of the form (a, a, b), then it is possible that $a = 0$ and $b = 0$, so $(0, 0, 0) \in W$. Let $\mathbf{u} = (u_1, u_1, u_2), \mathbf{v} = (v_1, v_1, v_2) \in W$, and let k be any scalar.

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u_1, u_1, u_2) + (v_1, v_1, v_2)\\ &= (u_1 + v_1, u_1 + v_1, u_2 + v_2),\end{aligned}$$

which results in a vector with the first two components equal, so W is closed under vector addition. And

$$\begin{aligned}k\mathbf{u} &= k(u_1, u_1, u_2)\\ &= (ku_1, ku_1, ku_2),\end{aligned}$$

a vector with equal first two components, so W is closed under scalar multiplication. W is a subspace of $\mathbb{R}^3$. $\square$

Example 53: Let S be the set of all vectors of the form $(a, 2a, a^2)$. where a is any real number. Show that S is *not* a subspace of $\mathbb{R}^3$.

solution: If $\mathbf{u} = (u, 2u, u^2)$ and $\mathbf{v} = (v, 2v, v^2)$ are any two vectors in S,

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u, 2u, u^2) + (v, 2v, v^2) \\ &= (u + v, 2u + 2v, u^2 + v^2) \\ &= (u + v, 2(u + v), u^2 + v^2).\end{aligned}$$

If S is closed under vector addition, then the third component of the sum would have to be $u^2 + 2uv + v^2$, which it is not; therefore, S is not a subspace of $\mathbb{R}^3$ because it is not closed under addition. $\square$

11.4 The Big Theory, Part 1: Theorems 1 and 2

THEOREM 1: THE ROW SPACE AND NULL SPACE OF ANY $m \times n$ MATRIX ARE SUBSPACES OF $\mathbb{R}^n$.

PROOF: Let A be any $m \times n$ matrix. The null space of A is the set of all vectors $\mathbf{x}$ in $\mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{0}$. The zero vector is always a solution to this equation, so the zero vector is in the null space of A. Let $\mathbf{u}$ and $\mathbf{v}$ be any two vectors in the null space of A and let k be any scalar. Then $A\mathbf{u} = \mathbf{0}$ and $A\mathbf{v} = \mathbf{0}$. Are the vectors $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ also in the null space?

$$\begin{aligned}A(\mathbf{u} + \mathbf{v}) &= A\mathbf{u} + A\mathbf{v} \\ &= \mathbf{0} + \mathbf{0} \\ &= \mathbf{0}\end{aligned} \qquad\qquad \begin{aligned}A(k\mathbf{u}) &= k(A\mathbf{u}) \\ &= k(\mathbf{0}) \\ &= \mathbf{0}\end{aligned}$$

Since $A(\mathbf{u} + \mathbf{v}) = \mathbf{0}$ and $A(k\mathbf{u}) = \mathbf{0}$, then $\mathbf{u} + \mathbf{v} \in$ null(A) and $k\mathbf{u} \in$ null(A) (the null space of A is closed under vector addition and scalar multiplication). Therefore, the null(A) is a subspace of $\mathbb{R}^n$.

Now consider the row space of A. Since A is $m \times n$, then it has m row vectors, each with n components, so all vectors in the row space are vectors in $\mathbb{R}^n$. The row space of a matrix is the span of the m rows, so any vector in the row space can be written as a linear combination of the rows of A. Let $\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_m$ be the m row vectors of A. Let $\mathbf{u}$ and $\mathbf{v}$ be any vectors in the row space of A and let k be any scalar. Then $\mathbf{u} = a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + \cdots + a_m\mathbf{r}_m$ and $\mathbf{v} = b_1\mathbf{r}_1 + b_2\mathbf{r}_2 + \cdots + b_m\mathbf{r}_m$. If $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ can be written as a linear combination of the row vectors of A, then they are in the row space of A:

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + \cdots + a_m\mathbf{r}_m) + (b_1\mathbf{r}_1 + b_2\mathbf{r}_2 + \cdots + b_m\mathbf{r}_m) \\ &= (a_1 + b_1)\mathbf{r}_1 + (a_2 + b_2)\mathbf{r}_2 + \cdots + (a_m + b_m)\mathbf{r}, \text{ so } \mathbf{u} + \mathbf{v} \in \text{ row}(A). \\ k\mathbf{u} &= k(a_1\mathbf{r}_1 + a_2\mathbf{r}_2 + \cdots + a_m\mathbf{r}_m) \\ &= ka_1\mathbf{r}_1 + ka_2\mathbf{r}_2 + \cdots + ka_m\mathbf{r}_m \\ &= (ka_1)\mathbf{r}_1 + (ka_2)\mathbf{r}_2 + \cdots + (ka_m)\mathbf{r}_m, \text{ so } k\mathbf{u} \in \text{ row}(A).\end{aligned}$$

Thus the row space and null space of any $m \times n$ matrix A are each subspaces of $\mathbb{R}^n$.$\square$

THEOREM 2: THE COLUMN SPACE AND LEFT NULL SPACE OF ANY $m \times n$ MATRIX ARE SUBSPACES OF $\mathbb{R}^m$.

PROOF: Let A be any $m \times n$ matrix. A^T, the transpose of A, is an $n \times m$ matrix, meaning there are n row vectors of A^T, each with m components, so the row vectors of A^T are vectors from $\mathbb{R}^m$. Theorem 1 allows us state that the row space of A^T is a subspace of $\mathbb{R}^m$ and the null space of A^T is a subspace of $\mathbb{R}^m$. The row space of A^T is the column space of A. The null space of A^T is the left null space of A; therefore, the column space of A and the left null space of A are subspaces of $\mathbb{R}^m$.$\square$

Please note that the proof of Theorem 1 utilizes the definition of subspace (Theorem 1 shows the three conditions are satisfied for things to be a subspace) and the proof of Theorem 2 applies the result from Theorem 1 to another matrix related to A. The proofs of these two theorems are what are called *direct* proofs in mathematics, the easiest to understand since the logic follows from what is given. If you are headed towards a degree in mathematics, I highly recommend reading and writing these proofs (see if you can write them yourself once you pick up on the pattern in how they are written).

COMMENTS ON HOW TO UNDERSTAND THE COURSE: Any course in linear algebra is about line-like relationships. The theory we are now beginning will culminate in "the big picture of linear algebra" that includes the set of all matrices. Matrices are the mathematical objects we use in linear algebra because they are *linear transformations*, which depict these line-like relationships. You will learn this in the next several chapters. If you are interested in a degree in mathematics, it is the theoretical framework of this section that needs your sharp attention because this is the true *mathematics* of the subject: it explains how everything is connected. *What do these first two theorems say in a "big picture" kind of way?* They essentially say that any $m \times n$ matrix involves two Euclidean vector spaces: $\mathbb{R}^m$ and $\mathbb{R}^n$. For weeks now, we have been solving systems of linear equations $A\mathbf{x} = \mathbf{b}$. If A is $m \times n$, then $\mathbf{x}$ must be a vector in $\mathbb{R}^n$, so $A\mathbf{x}$ is a matrix product: an $m \times n$ matrix A, multiplied by an $n \times 1$ matrix $\mathbf{x}$, which will result in an $m \times 1$ matrix $A\mathbf{x}$, or $\mathbf{b}$. In other words, the matrix A can be thought of as transforming vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$. So we have two important Euclidean spaces related by a matrix, and each of those spaces can be split up into two subspaces. These subspaces (the row, column, and two null spaces of a matrix A) are known as the **four fundamental subspaces of** A. When we learn that A is a linear transformation, which means it is a function, we will learn that the input space is $\mathbb{R}^n$ and the output space is $\mathbb{R}^m$, and the four fundamental subspaces help us understand the linear transformation. If you view Figure 2 on page 79, the input space is pictured to the left, and the matrix A inputs vectors from that space to produce vectors in the output space, pictured on the right. We will literally draw the big picture of linear algebra from this type of picture.

Here is how the following homework assignment is structured: The first part of the assignment consists of problems built for your exam 2 preparation. (In the lecture, we summarized the procedures for these problems.) The second part of this homework includes the proofs on subspaces. The exercises in the second part can be completed *after* the second exam.

11.5 HOMEWORK 10/23 "DO THINGS IN THREE's"

PART I: EXAM 2 PREPARATION

Solve each system by Gauss-Jordan elimination, writing your solutions in column vector form.

191.

$$\begin{aligned} x_1 - 3x_2 &= -5 \\ -2x_1 + 6x_2 &= 10 \end{aligned}$$

192.

$$\begin{aligned} x_1 \qquad - 2x_3 &= 2 \\ 3x_1 + x_2 - 2x_3 &= 9 \end{aligned}$$

193.

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

For each A, determine the null(A), row(A), null(A^T), and col(A).

194. $A = \begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix}$ **195.** $A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \end{bmatrix}$ **196.** $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 4 & 5 & 1 \end{bmatrix}$

Quickly review how to multiply matrices. Use the following matrices to find the products in problems 197-199.

$$B = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & -2 & 2 \\ 3 & 1 & -2 & 9 \end{bmatrix}, D = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ -4 & 3 & 1 \end{bmatrix}, E = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 4 & 5 & 1 \end{bmatrix}$$

197. BC **198.** DE **199.** CC^T

Use the definitions of linear dependence/independence and span for the following problems.

200. Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be nonzero distinct vectors from a vector space V. Are the vectors $\mathbf{u}, \mathbf{u}+\mathbf{v}$, and $\mathbf{u}+\mathbf{v}+\mathbf{w}$ linearly independent? Prove your claim.

201. Determine whether the vectors in the set are linearly independent or dependent: $\{(1,2,2,2),(3,4,4,4),(0,1,1,1)\}$

202. Is the set $\{(1,1,1,-1),(-2,2,2,2),(3,-1,-1,-3)\}$ linearly dependent or linearly independent?

203. Let U be the set of all vectors in $\mathbb{R}^4$ whose fourth component is the sum of the first three. Find a set of vectors that span U.

204. Let V be the set of all vectors in $\mathbb{R}^3$ that are perpendicular to the vector $(2,-4,1)$. Determine a set of vectors that span V.

205. Once on an exam, I collected all of the different answers to the following problem: The line $y = 2x$ in $\mathbb{R}^2$ is spanned by what set of vectors? Four answers in the collection are listed below. Which answers are incorrect?

(A) The vector $(1,2)$ spans the line.
(B) The vector $(\frac{1}{2},1)$ spans the line.
(C) The vectors $(1,2),(2,4),$ and $(3,6)$ span the line.
(D) The vectors $(-1,-2)$ and $(1,2)$ span the line.

Answer the following:

206. The column space of a matrix A is the span of the columns of A, so if any vector $\mathbf{v}$ is in the column space of A, then $\mathbf{v}$ can be written as a linear combination of the columns of A. Is the vector $\begin{bmatrix}3\\1\\9\end{bmatrix}$ in the column space of $\begin{bmatrix}1&2&1\\0&1&1\\4&5&1\end{bmatrix}$?

207. If A is a 4×3 matrix, then how many components does any vector in its null space have? To what Euclidean vector space do the vectors in the column space of A^T belong?

208. Is it possible to write $(2,9)$ as a linear combination of $(1,3),(0,1)$, and $(-2,-2)$?

PART II: HOMEWORK AFTER EXAM 2 **Use the definition of *subspace* to write proofs for the following. Use examples 51, 52, 53, and the proofs of Theorem 1 and Theorem 2 as a guide and standard for what you produce in writing. Use complete sentences.**

209. Let P be the set of all vectors in $\mathbb{R}^3$ that are perpendicular to $(-3,-2,1)$. Prove, or disprove, that P is a subspace of $\mathbb{R}^3$.

210. Let W be the set of all vectors in $\mathbb{R}^4$ whose first and last coordinates are zero. Prove, or disprove, that W is a subspace of $\mathbb{R}^4$.

211. Let S be the set of all vectors in $\mathbb{R}^3$ of the form (a, b, ab). Prove, or disprove, that S is a subspace of $\mathbb{R}^3$.

212. Let V be a vector space. Let $Z = \{\mathbf{0}_V\}$, the set containing one vector, the zero vector of V. Prove that Z is a subspace of V.

213. Let H be the set of all vectors in $\mathbb{R}^3$ whose dot product with $(1,2,3)$ is 3. Is H a subspace of $\mathbb{R}^3$?

214. Let P be the set of all vectors in $\mathbb{R}^2$ of the form (x, x^2). Prove, or disprove, that P is a subspace of $\mathbb{R}^2$.

215. Let V be a vector space and let $\mathbf{u}$ be a nonzero vector in V. Let Q be the set of all vectors in V that are orthogonal to $\mathbf{u}$. Prove that Q is a subspace of V.

12 BASIS AND DIMENSION

If you have been a physics student or have taken multi-variable calculus, then you are no doubt familiar with the use of Euclidean vectors. A popular notation for vectors in $\mathbb{R}^3$ is what is known as the *basis vector notation* for any vector $\mathbf{v} = (a, b, c)$ in this space. The vector $\mathbf{v}$ is often written as $a\mathbf{i}+b\mathbf{j}+c\mathbf{k}$, where the vectors $\mathbf{i}, \mathbf{j},$ and $\mathbf{k}$ are the three orthogonal unit vectors pointing in the x, y, and z directions. These vectors form what is called the **standard basis vectors** for $\mathbb{R}^3$. Every vector in $\mathbb{R}^3$ can be written in terms of these three linearly independent vectors. A *basis* for a vector space is an extremely important concept. It would be an impossible task to send, one by one or in groups, every vector in a vector space; however, it is possible to send someone a set of vectors and, from that, they will be able to receive the vector space in its entirety. This is done by way of a basis, a special set of vectors that now becomes one of the most important topics in our subject. Prior to defining basis, I would like to have a brief conversation in the first section about *why* you ought to care about basis vectors. If you are inventing an operation on an entire vector space, then all you have to do is know exactly what happens with the basis vectors under that operation. Once that is established, you know what happens with every vector in the vector space. The conversation dispels some mystery surrounding the cross product, a strange product of two vectors that produces a third vector perpendicular to the two in the product. If you have been exposed to the cross product of two vectors, then it might be a good read. If not, I would advise skipping 12.1, and going to 12.2, where we define a basis B of a vector space V.

12.1 Making Sense of The Cross Product (optional)

When students learn what the cross product is in a physics or mathematics course, the instructor usually makes a choice to teach a class a way to find it that involves forming a "fake" determinant, and a set of instructions that contain several mysterious things that must be done in order to produce a correct vector. Usually, it is done this way: Let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ be any two nonzero vectors in $\mathbf{R}^3$. The **cross product** of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u} \times \mathbf{v}$ is a multiplication defined on two vectors in $\mathbb{R}^3$ that produces a third vector perpendicular to both $\mathbf{u}$ and $\mathbf{v}$. A convenient way to find the product is to form the determinant in which the first row contains the basis vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, the second row contains the components of the vector mentioned first, and the third row contains the components of the vector mentioned second; that is,

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

From here, you find the value of the determinant by a process known as *expansion of the first row by minor determinants*, a strange method taught without any explanation as to its origins: You take the first row, starting with $\mathbf{i}$, writing it as a thing to "multiply" by a determinant found by knocking out the row and the column in which $\mathbf{i}$ it is located, and forming the 2×2 determinant by the entries that are not knocked out, a determinant whose value (a real number) is the first component of the resulting product. The same thing is done with $\mathbf{j}$ and $\mathbf{k}$, but a negative sign is placed in front of the $\mathbf{j}$ vector for some unknown reason. The entire procedure looks like this:

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \mathbf{i} \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} - \mathbf{j} \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} + \mathbf{k} \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}$$

All of the above determinants result in the following cross product:

$$\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2, \quad u_3v_1 - u_1v_3, \quad u_1v_2 - u_2v_1), \text{ or}$$
$$\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2)\mathbf{i} + (u_3v_1 - u_1v_3)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}$$

The point of this section is to highlight how the basis vectors are a vital piece of knowing how the cross product works, and how the basis vectors play a significant role in mathematics in establishing specific facts about all vectors in the vector space for which they are a basis. Here is one explanation of how the cross product is developed, all from simply defining what happens with these three standard basis vectors for $\mathbb{R}^3$:

The cross product is motivated by a problem: how is it possible to find a vector orthogonal to both vectors in the product? In more mathematical terms, how can we be sure that $\mathbf{u} \times \mathbf{v}$ is perpendicular to $\mathbf{u}$ and also perpendicular to $\mathbf{v}$? What immediately helps that is to understand that the three vectors are all perpendicular to each other, and they are unit vectors (vectors that have a length of one unit). If you know anything about the xy-coordinate system and the xyz-coordinate system, you ought to know that both are *right-handed* systems. (The use of your right hand helps make sense, for example, why counterclockwise is considered "positive" orientation or direction, as on a circle: curl your fingers of your right hand in that direction of a curve and your thumb is pointing in the up direction, indicating positive direction along the curve.)
The three basis vectors are used to develop the product $\mathbf{u} \times \mathbf{v}$, which should produce only a single result, not two or three or more answers, in order to be *well-defined*. Suppose we want to develop a product that produces a third vector orthogonal to both vectors in the product. It makes sense that we would want to somehow take into account the magnitudes of these vectors since that is what you memorized in grade school, like $2 \times 3 = 6$ (the product of the magnitudes of these numbers is 6). The vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ each have a magnitude of 1, so that's nice to have. But if you think about two non-parallel vectors in standard position, two *linearly independent* vectors, which determine a plane, you can probably visualize the scenario of having a choice of two vectors. We want only one, and this is where we use the right hand: imagine putting the heel of your right hand on the vector mentioned first in the product, $\mathbf{u}$, and then curling your fingers in the direction of $\mathbf{v}$, the vector mentioned second in $\mathbf{u} \times \mathbf{v}$, where

your thumb is pointing. If you were to do all possible products of these three unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ with the right hand rule in mind, you can think about the products of their magnitudes $(1 \cdot 1 = 1)$, and the vector that results according to where your thumb is pointing. For example, $\mathbf{i} \times \mathbf{j}$ would result in the vector $\mathbf{k}$ by the "right-hand rule" you hear often in a physics course. You would establish the following results:

- $\mathbf{i} \times \mathbf{j} = \mathbf{k}$
- $\mathbf{j} \times \mathbf{k} = \mathbf{i}$
- $\mathbf{k} \times \mathbf{i} = \mathbf{j}$
- $\mathbf{j} \times \mathbf{i} = -\mathbf{k}$
- $\mathbf{k} \times \mathbf{j} = -\mathbf{i}$
- $\mathbf{i} \times \mathbf{k} = -\mathbf{j}$
- $\mathbf{i} \times \mathbf{i} = \mathbf{0}$
- $\mathbf{j} \times \mathbf{j} = \mathbf{0}$
- $\mathbf{k} \times \mathbf{k} = \mathbf{0}$

After this, you define any vector crossed with itself is not defined, and these are the third column results above: any vector "crossed" with itself or a vector parallel to it results in the zero vector. You can now do the regular multiplication of vectors after establishing the above results:

$$
\begin{aligned}
\mathbf{u} \times \mathbf{v} &= (u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}) \times (v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}) \\
&= (u_1v_1)\mathbf{i} \times \mathbf{i} + (u_1v_2)\mathbf{i} \times \mathbf{j} + (u_1v_3)\mathbf{i} \times \mathbf{k} \\
&\quad + (u_2v_1)\mathbf{j} \times \mathbf{i} + (u_2v_2)\mathbf{j} \times \mathbf{j} + (u_2v_3)\mathbf{j} \times \mathbf{k} \\
&\qquad + (u_3v_1)\mathbf{k} \times \mathbf{i} + (u_3v_2)\mathbf{k} \times \mathbf{j} + (u_3v_3)\mathbf{k} \times \mathbf{k} \\
&= (u_1v_2)(\mathbf{k}) + (u_1v_3)(-\mathbf{j}) \\
&\quad + (u_2v_1)(-\mathbf{k}) + (u_3v_1)(\mathbf{j}) \\
&\qquad + (u_3v_2)(-\mathbf{i}) + (u_2v_3)(\mathbf{i}) \\
&= (u_2v_3 - u_3v_2)\mathbf{i} + (u_3v_1 - u_1v_3)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}
\end{aligned}
$$

That's quite the exercise in algebra, as it is similar to multiplying two trinomials together, distributing one term at a time, and applying all of the above identities established, but it is a good exercise in which you see the result from the previous page, only this result makes sense, and also gives someone a reason why they might invent that strange determinant form and show how convenient it is to use it to find the cross product. Anytime you see a difference of products like the ones that appear in the three components, the form $ad - bc$, you can express this as a 2×2 *determinant* $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, so you can play with the forms above to figure out how someone would use a 3×3 to expand by minors. (You would see why the negative appears in the middle component.) You would not want to find it this algebraic way every time you had to find the cross product of two vectors. Each difference of products in the three components can, of course, be described in determinant form, but we have not covered determinants yet, and yes, that is something we will cover in the course. We could return and connect the above to the determinant form, but the point has been made here: *If you* *know what happens with basis vectors of a vector space, then you know what happens with ALL vectors in the vector space!* Time to return to our focus: what is a basis?

12.2 Basis of a Vector Space

If you are a painter and go to the art store and want to buy a minimum number of colors that you could use to mix to create any color you wanted, you would purchase only five colors: red, yellow, blue, white, and black. Think about this: *any* other color can be made from some combination of these five. In linear algebra speak, these five colors *span* the set of all colors. Think, too, about why each of these five colors might be special: white, for example, cannot be made from any other color in the set; that is, no amount of red, yellow, blue, and black would make white paint. No amount of yellow, white, blue, and black would make red paint. Each of these colors are *independent* of each other. And since they span the set of all colors, they form a *basis* for the set of all colors, since all colors are a certain combination of these five. This is the section in which all of the new vocabulary comes together in one conversation. The span of a set of vectors is all linear combinations of those vectors. The vectors $(2,3)$ and $(-1,4)$ are often written in terms of the standard basis vectors of $\mathbb{R}^2$ as $2\mathbf{i}+3\mathbf{j}$ and $-\mathbf{i}+4\mathbf{j}$. Every vector in $\mathbb{R}^2$ is of the form (x,y) and can also be written as $x\mathbf{i}+y\mathbf{j}$. You only need these two vectors $\mathbf{i}=(1,0)$ and $\mathbf{j}=(0,1)$ to get every vector in the plane. These two vectors are also linearly independent.

DEFINITION: Let B represent a set of vectors in a vector space V. B is a **basis for** V if the vectors in the set B span V and are linearly independent. The number of vectors in a basis for a vector space V is called the **dimension of** V.

A basis for a vector space is known as a *minimum spanning set*, the smallest number of vectors needed to write every vector in the vector space. Examples: Lines through the origin are spanned by a single vector: the line $y=\frac{2}{3}x$ is spanned by $(1,\frac{2}{3})$. It is also spanned by the vector $(3,2)$. Single vectors are automatically independent. $(3,2)$ is independent and spans the line, so a basis for the line $y=\frac{2}{3}x$ is $\{(3,2)\}$. The same can be said of the vector $(1,\frac{2}{3})$ as it could be a basis for the line. The set $S=\{(3,2),(\frac{3}{2},1)\}$ also spans the line $y=\frac{2}{3}x$, but S is not linearly independent since one vector is a scalar multiple of the other vector, so S is not a basis. The dimension of the line is 1 since there is only one vector in its basis.

The **standard basis** of any Euclidean vector space is a set of unit vectors, any two of which are orthogonal to each other. In linear algebra, we change the notation for $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ in $\mathbb{R}^2$ and $\mathbb{R}^3$ and beyond to $\mathbf{e_1},\mathbf{e}_2,$ and $\mathbf{e}_3$ because we run out of letters of the alphabet for higher-dimensional vectors in $\mathbb{R}^n$:

$$(2,3) = 2\mathbf{e}_1 + 3\mathbf{e}_2, \text{ where } \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \text{ and } \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

$$(1,3,-2) = \mathbf{e}_1 + 3\mathbf{e}_2 - 2\mathbf{e}_3, \text{ where } \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \text{ and } \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \text{ and } \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The **standard basis for** $\mathbb{R}^n$ is the set of n vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, ..., \mathbf{e}_n\}$, where each $\mathbf{e}_i$ is a unit vector with the number 1 in the ith component of the vector, and zeros in all the other components. This notation will come up in some of the theory we discuss. The dimension of $\mathbb{R}^n$ is n, of course. Notice how dimension is defined naturally here, and makes sense. Lines are spanned by a single vector, and lines have dimension 1, planes have dimension 2 and are spanned by a minimum of two independent vectors.

Always reason why a set is a basis for a vector space by showing the set is linearly independent and spans the space! Pay tribute to the definition of basis at all times, and keep in mind it is a minimum spanning set. For the most part, it will be your job to *find a basis* for a vector space or a subspace of a vector space.

Example 54: Let W be the set of all vectors in $\mathbb{R}^3$ whose last component is the sum of the first two components; that is, all vectors of the form $(a, b, a+b)$. Show that W is a subspace of $\mathbb{R}^3$ and then find a basis for W. State the dimension of W and give a geometric description of this subspace.

solution: Let $\mathbf{u} = (u_1, u_2, u_1 + u_2), \mathbf{v} = (v_1, v_2, v_1 + v_2) \in W$, and let k be any scalar.

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (u_1 + v_1, u_2 + v_2, u_1 + u_2 + v_1 + v_2) \\ &= (u_1 + v_1, u_2 + v_2, (u_1 + v_1) + (u_2 + v_2)), \end{aligned}$$

so W is closed under vector addition.

$$\begin{aligned} k\mathbf{u} &= (ku_1, ku_2, k(u_1 + u_2)) \\ &= (ku_1, ku_2, ku_1 + ku_2), \end{aligned}$$

so W is closed under scalar multiplication. W is a subspace of $\mathbb{R}^3$. Its vectors are of the form

$$\begin{aligned} (a, b, a+b) &= (a, 0, a) + (0, b, b) \\ &= a(1,0,1) + b(0,1,1), \end{aligned}$$

which shows that the vectors $(1,0,1)$ and $(0,1,1)$ span W. If these vectors are linearly independent, then since they also span W, they form a basis for W. To show the vectors are linearly independent, show that the only solution to $k_1(1,0,1) + k_2(0,1,1) = (0,0,0)$ is $k_1 = k_2 = 0$: This is left to the student to show! Therefore, a basis for W is the set $B = \{(1, 0, 1), (0, 1, 1)\}$. The dimension of W is 2, so W is a plane.

Example 55: Let V be the set of all vectors in $\mathbb{R}^2$ with a second component that is twice the first component. Prove that V is a one-dimensional subspace of $\mathbb{R}^2$.

solution: The set V consists of all vectors of the form $(a, 2a)$. First, show that this is a subspace of $\mathbb{R}^2$. Let $\mathbf{u} = (u, 2u)$ and $\mathbf{v} = (v, 2v)$ be any vectors in V, and let k be sny scalar.

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u.2u) + (v, 2v) \\ &= (u + v, 2u + 2v) \\ &= (u + v, 2(u + v)),\end{aligned}$$

and

$$\begin{aligned}k\mathbf{u} &= k(u, 2u) \\ &= (ku, 2ku),\end{aligned}$$

so V is closed under addition and scalar multiplication, since both adding and scalar multiplying yield vectors whose second component is twice the first component. V is a subspace of $\mathbb{R}^2$. Also, $(a, 2a) = a(1, 2)$, which means the vector $(1, 2)$ spans V, and, because it is a single vector, it is linearly independent. A basis for V is $\{(1, 2)\}$, which has dimension 1.

Example 56: Show that the set $\{(1, 1), (0, 1)\}$ is a basis for $\mathbb{R}^2$.

solution: All vectors in $\mathbb{R}^2$ are of the form (a, b), where $a, b \in \mathbb{R}$. To show the vectors span $\mathbb{R}^2$, we must show that all vectors of the form (a, b) can be written in terms of the two alleged basis vectors. The mathematical way to do this is to solve for the variables c_1 and c_2 in the equation $(a, b) = c_1(1, 1) + c_2(0, 1)$. Note this means that

$$\begin{aligned}a &= c_1 \\ b &= c_1 + c_2,\end{aligned}$$

so $c_1 = a$ and $c_2 = b - a$. Thus, any vector in $\mathbb{R}^2$ $(a, b) = a(1, 1) + (b - a)(0, 1)$. This shows that the vectors span $\mathbb{R}^2$. Since one vector is not a scalar mutliple of the other, then the set $\{(1, 1), (0, 1)\}$ is linearly independent, making it a basis for $\mathbb{R}^2$.

12.3 Homework 10/28

216. Let N be the set of all vectors in $\mathbb{R}^3$ of the form $(a, -a, -a)$, where a is any scalar. Show that N is a subspace of $\mathbb{R}^3$ and then find a basis for N. Give the dimension of N.

217. Let W be the set of all vectors in $\mathbb{R}^3$ orthogonal (perpendicular) to the vector $(-1, 1, 1)$. Determine a basis for W and give its dimension. You do not have to prove it is a subspace of $\mathbb{R}^3$, but you do need to do three very specific things: find an algebraic description of W, use addition and scalar multiplication to show that all vectors in W can be written in terms of a certain set of vectors (that action you take will show the vectors you find will span W), then show that the vectors are linearly independent.

218. Let U be the subspace of $\mathbb{R}^4$ spanned by $(1, 2, 1, 2), (3, 6, 4, 7),$ and $(2, 4, 1, 3)$. Find a basis for U, showing it is a basis.

219. Let W be the subspace of $\mathbb{R}^3$ spanned by the vectors $(1, 3, 2), (2, 6, 4), (1, 4, 1),$ and $(2, 7, 3)$. Find a basis for W, showing it is a basis.

220. Find a basis for the row space of a 3×2 matrix whose entries are all 3.

221. Using the same matrix in the previous problem, find a basis for its null space, a basis for its column space, and a basis for its left null space.

Use the following items for the remaining problems in this homework assignment: the matrix A for problems 222-223, and 229, the matrix B for 224-225, and 230, and the system for 227-228.

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 0 & 3 & 3 \\ 4 & 1 & 5 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 & 3 \\ 0 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}, \qquad \begin{aligned} x_1 - x_2 + x_3 + x_4 &= 1 \\ x_1 \phantom{{}+x_2} + 2x_3 + 3x_4 &= 1 \\ x_1 + x_2 + x_3 + 3x_4 &= 1 \end{aligned}$$

222. Determine bases for the row(A) and null(A).

223. Determine bases for the col(A) and the left null space of A.

224. Determine bases for the row(B) and null(B).

225. Determine bases for the col(B) and the left null space of B.

226. Solve the system above by Gauss Jordan elimination and write the solution in column vector form.

227. Write the system in the previous problem as a matrix equation $A\mathbf{x} = \mathbf{b}$. Determine bases for the null spaces, row and column spaces of this matrix A. State the dimension of each.

228. Note that the matrix equation you wrote in the previous problem can be understood in the following way: $A\mathbf{x} = \mathbf{b}$ is an equation showing that a matrix A *transforms* a vector $\mathbf{x}$ to a vector $\mathbf{b}$. When read this way, the matrix is taking action on a vector by multiplication and can be studied as a function of the form $T(\mathbf{x}) = A\mathbf{x}$. Find $T(0, -1, -1, 1)$.

229. We are about to learn that every matrix you meet defines a special function called a *linear transformation*. Take the matrix A from above, multiply it by any appropriate nonzero vector of your choice. Then take the vector you chose and use it to find $T(x_1, x_2, x_3) = (x_1 + 3x_2 + 4x_3, 3x_2, 3x_3, 4x_1 + x_2 + 5x_3)$. Compare the results.

230. Let $T(x, y, z) = (3z, 2y + 3z, x + 2y + 3z)$. Find $T(1, 2, 3)$. Then multiply B from above by this vector (write the vector $\mathbf{v} = (1, 2, 3)$ in column form and find $B\mathbf{v}$). Compare results.

13 LINEAR TRANSFORMATIONS

13.1 All Matrices Determine Linear Transformations

Let A be an $m \times n$ matrix. Then for any vector $\mathbf{x}$, the product $A\mathbf{x}$ is defined when $\mathbf{x} \in \mathbb{R}^n$. If $A\mathbf{x} = \mathbf{b}$, then $\mathbf{b} \in \mathbf{R}^m$. An $m \times n$ matrix A defines a function from $\mathbb{R}^n$ into $\mathbb{R}^m$ because it transforms vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$ by the multiplication just described. The linear transformation is $T(\mathbf{x}) = A\mathbf{x}$.

13.2 Clarification of Function Terms and Notation

The concept of a function is crucial to any work in mathematics at the college level and beyond. All persons present in this course are certainly familiar with *real-valued functions of a single variable*, functions like $f(x) = x^2, f(x) = mx + b, f(x) = \frac{x}{3x-1}$, etc... Some of you have studied functions of two variables, or three, or more, and some of you have studied *vector-valued functions of one or more variables*. A function is a rule that assigns to each input one and only one output. Think $f(\textit{input}) = \textit{output}$. A function *changes* an input, or *transforms* an input. A function is also known as a *transformation* since it does such a thing. Another common word for a function is the word *mapping*. The set of all inputs of a function is called the *domain* of the function. The set of all outputs of a function is called the *range* of the function. In linear algebra, we use the words *transformation* and *mapping* (the verb form of the word: a function maps a space onto another space, for example.) While most of you ought to be comfortable with these words, many of you have not used the word **codomain**. For example, the domain of $f(x) = x^2$ is $\mathbb{R}$. The range of $f(x) = x^2$ is a subset of $\mathbb{R}$, usually given in interval notation: $[0, \infty)$. The codomain of $f(x) = x^2$ is $\mathbb{R}$. The codomain of a function is nothing other than the space in which the outputs can be found. It is the space to which the inputs are mapped. There is a special notation you will see often in the course that specifies the function by its name, its domain, and its codomain:

$$f : \mathbb{R} \to \mathbb{R}$$

This notation is read "f maps the set of real numbers to the set of real numbers." These are, for us, vector spaces. If f maps the vector space V to the vector space W, we write $f : V \to W$. The following examples ought to illuminate these details, so be cautiously observant as you study them.

Example 57: Let $f(x, y) = 2xy + y^2$. This is a function familiar to those of you who have studied multi-variable differential calculus. Function notation reveals in a clear way in what vector space inputs are coming from, and to what vector space they are mapped. Here, $f : \mathbb{R}^2 \to \mathbb{R}$. The function maps vectors in the plane to real numbers. The domain is $\mathbb{R}^2$. The co-domain is $\mathbb{R}$. The range is $\mathbb{R}$. This is a real-valued function that maps the vector space $\mathbb{R}^2$, a plane, to the vector space $\mathbb{R}$, a line.

Example 58: Let $f(x, y) = (3x + y, x - y)$. When you read function notation, you must be able to *see* from what space the inputs and outputs are. Functions have a name: this one is *f*. Following the name is notation indicating the inputs: (x, y) represents familiar ordered pairs one finds in $\mathbb{R}^2$. Following the "is equals to" symbol is the rule for what to do with those inputs. Here, $(3x + y, x - y)$ is a thing with two components; thus, $\mathbb{R}^2$ is the vector space to which the inputs are mapped. This function maps vectors in $\mathbb{R}^2$ (domain) to vectors in $\mathbb{R}^2$ (codomain); that is, $f : \mathbb{R}^2 \to \mathbb{R}^2$. It is common to see this function written

$$f(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 3x + y \\ x - y \end{bmatrix}.$$

Example 59: Let $f(x_1, x_2, x_3) = (2x_1 - x_3, 3x_1 - x_2 + x_3)$. This is a function that maps vectors in $\mathbb{R}^3$ to vectors in $\mathbb{R}^2$; that is, $f : \mathbb{R}^3 \to \mathbb{R}^2$. The domain is $\mathbb{R}^3$. The codomain is $\mathbb{R}^2$. Know that we can replace the word *function* with *transformation*, as the next example demonstrates.

Example 60: Let $T(x_1, x_2) = (2x_2 - x_1, 2x_1 - x_2, x_1 + x_2, x_1)$. T is a transformation that maps vectors in $\mathbb{R}^2$ to vectors in $\mathbb{R}^4$; that is, $T : \mathbb{R}^2 \to \mathbb{R}^4$. The domain is $\mathbb{R}^2$. The codomain is $\mathbb{R}^4$.

Examples 58-60 have something in common: they are linear transformations.

13.3 Definition of a Linear Transformations

DEFINITION: A **linear transformation** is a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ satisfying

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \tag{14}$$
$$T(k\mathbf{u}) = kT(\mathbf{u}) \tag{15}$$

for all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and all scalars k.

Example 61: Prove that the function in example 58 is a linear transformation.

solution: $f(x, y) = (3x + y, x - y)$ is a function $f : \mathbb{R}^2 \to \mathbb{R}^2$. Let $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ be any vectors in the domain $\mathbb{R}^2$, and let k be any scalar.

$$\begin{aligned} f(\mathbf{u} + \mathbf{v}) &= f((u_1, u_2) + (v_1, v_2)) \\ &= f(u_1 + v_1, u_2 + v_2) \\ &= (3(u_1 + v_1) + (u_2 + v_2), (u_1 + v_1) - (u_2 + v_2)) \\ &= (3u_1 + 3v_1 + u_2 + v_2, u_1 + v_1 - u_2 - v_2) \\ &= (3u_1 + u_2 + 3v_1 + v_2, u_1 - u_2 + v_1 - v_2) \\ &= (3u_1 + u_2, u_1 - u_2) + (3v_1 + v_2, v_1 - v_2) \\ &= f(u_1, u_2) + f(v_1, v_2) \\ &= f(\mathbf{u}) + f(\mathbf{v}), \end{aligned}$$

and

$$\begin{aligned} f(k\mathbf{u}) &= f(k(u_1,u_2)) \\ &= f(ku_1,ku_2) \\ &= (3(ku_1)+ku_2, ku_1-ku_2) \\ &= (k(3u_1+u_2), k(u_1-u_2)) \\ &= k(3u_1+u_2, u_1-u_2) \\ &= kf(u_1,u_2) \\ &= kf(\mathbf{u}). \end{aligned}$$

Therefore, f is a linear transformation. □

Example 62: Prove that the function in example 57 is *not* a linear transformation.

solution: $f(x,y) = 2xy + y^2$ is a transformation in which $f : \mathbb{R}^2 \to \mathbb{R}$. Let $\mathbf{u} = (u_1,u_2)$ and $\mathbf{v} = (v_1,v_2)$ be any vectors in the domain $\mathbb{R}^2$.

$$\begin{aligned} f(\mathbf{u}+\mathbf{v}) &= f((u_1,u_2)+(v_1,v_2)) \\ io &= f(u_1+v_1, u_2+v_2) \\ &= 2(u_1+v_1)(u_2+v_2) + (u_2+v_2)^2 \\ &= 2(u_1u_2+u_1v_2+u_2v_1+v_1v_2) + u_2^2 + 2u_2v_2 + v_2^2 \\ &= 2u_1u_2 + 2u_1v_2 + 2u_2v_1 + 2v_1v_2 + u_2^2 + 2u_2v_2 + v_2^2 \\ &= 2u_1u_2 + u_2^2 + 2v_1v_2 + v_2^2 + 2u_1v_2 + 2u_2v_1 + 2u_2v_2 \\ &= f(\mathbf{u}) + f(\mathbf{v}) + 2u_1v_2 + 2u_2v_1 + 2u_2v_2 \\ &\neq f(\mathbf{u}) + f(\mathbf{v}), \end{aligned}$$

so f does not preserve addition; therefore, f is not a linear transformation. □

The fancier way to define a linear transformation is to say that it is a function (or mapping) that preserves both addition and scalar multiplication. Preserving addition is (14) in the definition. Preserving scalar multiplication is (15). Note the extreme detail from one step to the next in the two examples that follow the definition.[8] Also note that when you prove that a function is not linear, it is sufficient to show that only one of the two properties does not hold, since both are required in order for something to be called a linear transformation.[9] You ought to feel a sense of connection with the two operations that define vectors, and with what subspaces are.

[8] *Advice*: Work exercises 231 and 232 with this in mind, and mock the work you see in the examples.

[9] *Advice*: Work exercise 233 to note an alternative way to show a transformation is not linear, and consider which of the two properties might be less time consuming.

13.4 Homework 10/29

231. Show that the function in example 59 is a linear transformation.

232. Show that the function in example 60 is a linear transformation.

233. In example 62, it was shown that the function $f(x, y) = 2xy + y^2$ was not linear by showing that f does not preserve addition. We could have shown it was not linear by demonstrating it does not preserve scalar multiplication. Does this work?

234. If $f(x_1, x_2, x_3) = (2x_1 - x_3, 3x_1 - x_2 + x_3)$, find $f(0, 0, 0)$.

235. If $T(x_1, x_2) = (2x_2 - x_1, 2x_1 - x_2, x_1 + x_2, x_1)$, find $T(0, 0)$.

236. Assume that $T : V \to W$ is a linear transformation. Let $\mathbf{0}_V$ represent the zero vector in V and let $\mathbf{0}_W$ represent the zero vector in W. Prove that $T(\mathbf{0}_V) = \mathbf{0}_W$.

237. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by rotating any vector in the domain by $\frac{\pi}{2}$ radians. Prove that T is a linear transformation.

238. Let $T(x_1, x_2) = x_1 - x_2$. Identify the domain, codomain, and prove, or disprove, that T is a linear transformation.

239. Let $T(x, y, z) = (z, x + y)$. Identify the domain, codomain, and prove, or disprove, that T is a linear transformation.

240. Let $T(\mathbf{x}) = \mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^n$. Prove, or disprove, that T is linear.

241. Let $T(\mathbf{x}) = \mathbf{0}$, where $\mathbf{x}, \mathbf{0} \in \mathbb{R}^n$. Prove, or disprove, that T is linear.

The following problems prepare you for the next exam.

242. Solve the system by Gauss Jordan elimination.

$$\begin{aligned} x_1 - x_2 + x_3 - x_4 &= 0 \\ x_1 + x_2 + x_3 + x_4 &= 0 \\ x_1 - x_2 - x_4 - x_4 &= 0 \end{aligned}$$

243. Prove that the solution you just found to the system is a subspace of $\mathbb{R}^4$.

244. Find a basis for the solution you just proved was a subspace of $\mathbb{R}^4$, explaining why it is a basis and using the linear algebra vocabulary properly.

245. Write the system in 242 as a matrix equation $A\mathbf{x} = \mathbf{b}$, identifying the matrix A. Find bases for the null(A), the row(A), the left null space of A, and the col(A).

246. The matrix $\begin{bmatrix} 1 & 3 & 0 & -1 & 2 \\ 0 & 0 & 1 & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ is an augmented matrix representing a system of equations $A\mathbf{x} = \mathbf{b}$. The matrix is in reduced row echelon form. How many equations and unknowns are in the system, and what is the solution to the system?

247. Is the solution to the system a subspace of a vector space? Explain.

248. Re-read the first sentence in problem 246. What is the reduced row echelon form of the matrix A in the system? Find bases for the row space of A and null space of A.

249. Let $A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 0 \end{bmatrix}$. Determine the null space of A by giving its basis and dimension.

250. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be given by $T(x_1, x_2, x_3) = (x_1 + 3x_3, \quad 2x_1 + x_2)$. What is the domain and codomain of T? Find all vectors (x_1, x_2, x_3) in the domain of T such that $T(x_1, x_2, x_3) = (0, 0)$.

14 THE STANDARD MATRIX OF A LINEAR TRANSFORMATION

This section focuses on three things: determining the **standard matrix** of a linear transformation T, and two more function terms for T: the **kernel** of T, and the **range** of T.

14.1 Finding the Standard Matrix A of T

Every matrix defines a linear transformation; that is, you can think of any $m \times n$ matrix A as being a linear transformation from $\mathbb{R}^n \to \mathbb{R}^m$ because $T(\mathbf{x}) = A\mathbf{x}$ is the linear transformation. A wonderful exam question is to prove that if A is any $m \times n$ matrix, then $T(\mathbf{x}) = A\mathbf{x}$ is a linear transformation. When a linear transformation between finite vector spaces is given in the space-saver notation, such as $T(x, y, z) = (x+z, z-2y)$, then it is a matter of interpretation to extract the standard matrix from this transformation. Remember that the matrix A houses the coefficients of the linear combinations of unknown variables. In this example, we can read that $T : \mathbb{R}^3 \to \mathbb{R}^2$, so T is taking vectors with three components to those with two, and the matrix A will be conveying linear combinations of those three components. We can clearly see that there are two linear combinations of $x, y,$ and z. Written in column form:

$$T(x, y, z) = (x + z, z - 2y) = \begin{bmatrix} x + z \\ z - 2y \end{bmatrix} = \begin{bmatrix} 1x + 0y + 1z \\ 0x - 2y + 1z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Notice the transformation reads as two linear combinations of three things, so you can expect A to be a 2×3 matrix, which is consistent with the fact that $T : \mathbb{R}^3 \to \mathbb{R}^2$. This matrix is known as the *standard matrix* of the linear transformation. Further examples:

Example 63: Determine the standard matrices $A_1, A_2,$ and A_3 of the linear transformations $T_1(x_1, x_2) = (x_2 - 2x_1, x_2, x_1, x_1 + x_2)$, $T_2(x_1, x_2, x_3, x_4) = (x_2 + x_4, 3x_1 + 2x_4, 3x_3)$, and $T_3(x, y) = (3y, x - y, x + y, 2x - 5y)$.

solutions: Though this should be a skill that does not require "work shown" as you should be able to identify the size of A and immediately write it down from viewing these functions, here is the breakdown of how you can "see" the matrix from the linear combinations:

$$T_1(x_1, x_2) = (x_2 - 2x_1, x_2, x_1, x_1 + x_2) = \begin{bmatrix} -2x_1 + 1x_2 \\ 0x_1 + 1x_2 \\ 1x_1 + 0x_2 \\ 1x_1 + 1x_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$T_2(x_1,x_2,x_3,x_4) = (x_2+x_4, 3x_1+2x_4, 3x_3) = \begin{bmatrix} 0x_1+1x_2+0x_3+1x_4 \\ 3x_1+0x_2+0x_3+2x_4 \\ 0x_1+0x_2+3x_3+0x_4 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 1 & 0 & 1 \\ 3 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

$$T_3(x,y) = (3y, x-y, x+y, 2x-5y) = \begin{bmatrix} 0x+3y \\ 1x-1y \\ 1x+1y \\ 2x-5y \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ 1 & -1 \\ 1 & 1 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

The matrices are $A_1 = \begin{bmatrix} -2 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 3 & 0 & 0 & 2 \\ 0 & 0 & 3 & 0 \end{bmatrix}$, and $A_3 = \begin{bmatrix} 0 & 3 \\ 1 & -1 \\ 1 & 1 \\ 2 & -5 \end{bmatrix}$.
It is wise to note the domain and codomain of each transformation and see the correspondence of those spaces to the sizes of each matrix. For example, $T_1 : \mathbb{R}^2 \to \mathbb{R}^4$, which means A_1 is a 4×2 matrix.

14.2 The Kernel and Range of T

DEFINITION: The **kernel** of a transformation $T : U \to V$ is the set of all $\mathbf{x} \in U$ such that $T(\mathbf{x}) = \mathbf{0}_V$, where $\mathbf{0}_V$ is the zero vector in V. The kernel of T is denoted by $\ker(T)$. In other words, the kernel of the transformations are all of the input vectors that get mapped to the zero vector in the output space. In other words, the null space of the standard matrix A for T is the kernel of T; If $T(\mathbf{x}) = A\mathbf{x}$, then $\text{null}(A) = \ker(T)$.

It is often difficult to find the range of any given function, but for linear transformations, the range of the transformation is easily found. How so? First, the definition of *range* is, of course, the following:

DEFINITION: The **range** of a linear transformation $T : U \to V$ is the set of all vectors $\mathbf{y} \in V$ such that $T(\mathbf{x}) = \mathbf{y}$ where $\mathbf{x} \in U$; that is, the range of T is the set of all output vectors. The range of T is denoted by $\text{range}(T)$.

It is a surprising fact that the range of a linear transformation is the column space of the standard matrix for the linear transformation, something we establish in the theory that concludes this chapter. There is one more word to take care of prior to this theory: the word *image*.

If T maps a vector $\mathbf{v}$ to a vector $\mathbf{w}$; that is, if $T(\mathbf{v}) = \mathbf{w}$, then $\mathbf{w}$ is called the **image** of $\mathbf{v}$. The word *image* is used interchangeably with the word *range*. The images of all the vectors in the domain *is* the range, so sometimes people refer to the image of the vector space V is the range of the function. Think simple: if $f(x) = x^2$, then 9 is the image of 3, or the image of 3 is 9. Notice that for this function, the domain is $\mathbb{R}$, the codomain is $\mathbb{R}$, but the range is *not* $\mathbb{R}$. The range is a subset of $\mathbb{R}$, all real numbers greater than, or equal to, zero, and this set is the image of $\mathbb{R}$.

14.3 The Big Theory, Part 2: Theorems 3 and 4

The columns of a matrix tell you a lot about what the matrix does to a vector space in terms of a linear transformation. The standard basis for $\mathbb{R}^2$ is $\{(1,0),(0,1)\}$. The standard basis for $\mathbb{R}^3$ is the set $\{(1,0,0),(0,1,0),(0,0,1)\}$. The standard basis for $\mathbb{R}^4$ is... You can see how cumbersome it is to write the basis vectors for Euclidean vector spaces, so there is a notation mathematicians have invented to reduce the agony of writing a vector whose components are all zero except for the single component that is 1. They use **e** for "Euclidean" and a subscript indicating which component is 1. For example, $\mathbf{e_1}$ and $\mathbf{e_2}$ can describe the basis vectors of $\mathbb{R}^2$, meaning $\mathbf{e_1} = (1,0)$ and $\mathbf{e_2} = (0,1)$. For $\mathbb{R}^3$, the standard basis vectors are $\mathbf{e_1} = (1,0,0)$, $\mathbf{e_2} = (0,1,0)$, and $\mathbf{e_3} = (0,0,1)$. The basis for any Euclidean space $\mathbb{R}^n$ is written $\{\mathbf{e_1}, \mathbf{e_2}, ..., \mathbf{e_n}\}$.

What if T was a linear transformation mapping $\mathbb{R}^3$ to $\mathbb{R}^2$? Then the standard matrix A of T is a 2×3 matrix. Suppose $A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$, and so $T(x,y,z) = (ax+by+cz, dx+ey+fz)$. Justify each of the changes below:

$$T(x,y,z) = (ax+by+cz, dx+ey+fz) \tag{16}$$

$$T((x,0,0),(0,y,0),(0,0,z)) = (ax,dx)+(by,ey)+(cz,fz) \tag{17}$$

$$T(x(1,0,0)+y(0,1,0)+z(0,0,1)) = x(a,d)+y(b,e)+z(c,f) \tag{18}$$

$$xT(1,0,0)+yT(0,1,0)+zT(0,0,1) = x(a,d)+y(b,e)+z(c,f). \tag{19}$$

The standard basis vectors of $\mathbb{R}^3$ are $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$, and the final step shows that **the columns of the standard matrix are the vectors in the codomain to which the standard basis vectors are mapped.** Rewriting that last step in column vector notation gives one another reason why that notation is preferred if matrices are the objects used to house the data involved in a linear transformation:

$$xT(\begin{bmatrix}1\\0\\0\end{bmatrix}) + yT(\begin{bmatrix}0\\1\\0\end{bmatrix}) + zT(\begin{bmatrix}0\\0\\1\end{bmatrix}) = x\begin{bmatrix}a\\d\end{bmatrix} + y\begin{bmatrix}b\\e\end{bmatrix} + z\begin{bmatrix}c\\f\end{bmatrix}$$

Example 64: Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be linear, where $T(x_1,x_2,x_3) = (2x_1 + x_2, 3x_2 - x_3)$. Find the standard matrix of T and use it to find $T(0,1,-1)$.

solution: Perhaps you have arrived at a comfortable place in which you can look at the function and write its standard matrix quickly:

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

You see that $T(1,0,0) = (2,0)$, $T(0,1,0) = (1,3)$, and $T(0,0,1) = (0,-1)$, simply by observing the columns of A. Writing $(0,1,-1)$ in terms of the standard basis vectors, followed by using linearity,

$$\begin{aligned} (0,1,-1) &= 1(0,1,0) - 1(0,0,1) \\ T(0,1,-1) &= T(1(0,1,0) - 1(0,0,1)) \\ T(0,1,-1) &= T(0,1,0) - T(0,0,1) \\ T(0,1,-1) &= (1,3) - (0,-1) \\ T(0,1,-1) &= (1,4). \end{aligned}$$

THEOREM 3: The columns of a matrix are the vectors to which basis vectors are mapped. More precisely, an $m \times n$ matrix A determines a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$. The standard basis of $\mathbb{R}^n$ is $\{\mathbf{e_1}, \mathbf{e_2}, ..., \mathbf{e_n}\}$. If $A = \begin{bmatrix} \mathbf{c_1} & \mathbf{c_2} & \cdots & \mathbf{c_n} \end{bmatrix}$, where $\mathbf{c_i}$ is the ith column vector, then $T(\mathbf{e_i}) = \mathbf{c_i}$.

Before the proof, let's be clear on all of this notation. In example 64s, the standard matrix of the linear transformation was

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

If we made use of the notation introduced in this section, then $\mathbf{e_1}, \mathbf{e_2}$, and $\mathbf{e_3}$ are the standard basis vectors of the domain $\mathbb{R}^3$, best thought of as column vectors. The column vectors of A are $\mathbf{c_1} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $\mathbf{c_2} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, and $\mathbf{c_3} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$. I will leave it to the student to finish connecting this example to the theorem. Here is the proof:

PROOF: Let $\mathbf{x}$ be any vector in $\mathbb{R}^n$. In column vector form,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Given the standard basis of $\mathbb{R}^n$, we know that $\mathbf{x}$ can be written in terms of these vectors:

$$\mathbf{x} = x_1\mathbf{e_1} + x_2\mathbf{e_2} + \cdots + x_n\mathbf{e_n}$$

For any vector $\mathbf{x}$, $T(\mathbf{x}) = A\mathbf{x}$, and

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e_1} + x_2\mathbf{e_2} + \cdots + x_n\mathbf{e_n}) && (20) \\ &= x_1T(\mathbf{e_1}) + x_2T(\mathbf{e_2}) + \cdots + x_nT(\mathbf{e_n}), && (21) \end{aligned}$$

since a linear transformation preserves linear combinations. We also have that

$$\begin{aligned}
T(\mathbf{x}) &= A\mathbf{x} && (22)\\
&= \begin{bmatrix}\mathbf{c_1} & \mathbf{c_2} & \cdots & \mathbf{c_n}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix} && (23)\\
&= \mathbf{c_1}x_1 + \mathbf{c_2}x_2 + \cdots + \mathbf{c_n}x_n && (24)\\
&= x_1\mathbf{c_1} + x_2\mathbf{c_2} + \cdots + x_n\mathbf{c_n}. && (25)
\end{aligned}$$

By (21) and ,

$$x_1T(\mathbf{e_1}) + x_2T(\mathbf{e_2}) + \cdots + x_nT(\mathbf{e_n}),$$

so $T(\mathbf{e_1}) = \mathbf{c_1}, T(\mathbf{e_2}) = \mathbf{c_2}$, ..., and $T(\mathbf{e_n}) = \mathbf{c_n}$.□

THEOREM 4: Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation defined by $T(\mathbf{v}) = A\mathbf{v}$. The range of T is the span of the column vectors of A; that is, the range of T is the column space of A.

PROOF: Let $\mathbf{w}$ be any vector in the range of T, which means there exists a vector $\mathbf{v}$ in the domain of T such that $T(\mathbf{v}) = \mathbf{w}$; that is, $\mathbf{w}$ is the image of $\mathbf{v}$. Since $\mathbf{v} \in \mathbb{R}^n$, then it can be written in terms of the standard basis vectors of $\mathbb{R}^n$, so that $T(\mathbf{v}) = \mathbf{w}$ can be written as

$$T(c_1\mathbf{e_1} + c_2\mathbf{e_2} + \cdots + c_n\mathbf{e_n}) = \mathbf{w}.$$

The linearity of T allows us to write this as

$$c_1T(\mathbf{e_1}) + c_2T(\mathbf{e_2}) + \cdots + c_nT(\mathbf{e_n}) = \mathbf{w},$$

where we see that the vector in the range $\mathbf{w}$ is written as a linear combination of the images of the basis vectors of the domain; that is, $\mathbf{w}$ is a linear combination of the column vectors of the matrix A. By definition of *span*, the range of T is the span of the column vectors of A. □

14.4 HOMEWORK

Use the following items for the homework exercises:

$$T_1(x,y,z) = (x+y-z, 3z), \qquad T_2(x_1,x_2,x_3) = (x_3 - x_1, x_1 + x_2)$$

$$T_3(x_1,x_2) = (x_2, x_2, x_1, x_1 + x_2), \qquad T_4(x_1,x_2) = (x_1 + x_2, 4x_2 - 8x_1)$$

Determine the standard matrix for the linear transformation:

251. T_1

252. T_2

253. T_3

254. T_4

Determine the kernel and range of the linear transformation:

255. T_1

256. T_2

257. T_3

258. T_4

259. Describe the domain and codomain for T_1 and state the standard basis for the domain. Where do the basis vectors get mapped?

260. Repeat the same exercise for T_2.

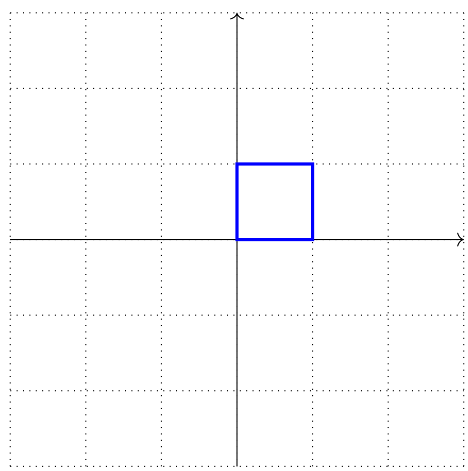
15 VISUALIZING MATRICES: Problems 261-280

In this chapter, the homework exercises are spread out within the exposition to emphasize the importance of reading through the examples and taking some time to do the sketching required.

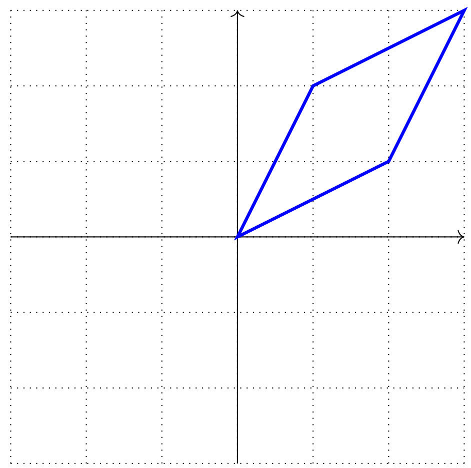
15.1 The Unit Square

In $\mathbb{R}^2$, the square with vertices (0,0), (1,0), (1,1), and (0,1) will be referred to as the **unit square**. When the adjective *unit* describes a mathematical object, it usually means that object has some measure of 1 associated with it: *unit vectors* are vectors of length 1, the *unit circle* $x^2 + y^2 = 1$ has a radius of 1, and, for our purposes here, the *unit square* has an area of 1. Euclidean vector spaces are measured by a quantity associated with their dimensions: $\mathbb{R}$, the real line, can be cut into pieces in which *length* is a measure; the plane, $\mathbb{R}^2$, *area*; $\mathbb{R}^3$, *volume*, etc... There is something about having spaces with a flatness associated with them, along with orthogonality (perpendicularity).

You often hear that two nonparallel vectors in the plane (in standard position) determine a parallelogram, so we can say that the unit square is determined by the basis vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.



The matrix $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ defines transformation $T(x, y) = (2x + y, x + 2y)$, and we know by looking at the columns that $T(1, 0) = (2, 1)$ and $T(0, 1) = (1, 2)$. If we sketch the two column vectors in the xy-plane, we see that these vectors determine the parallelogram with vertices the origin, (1,2), (3,3), and (2,1):



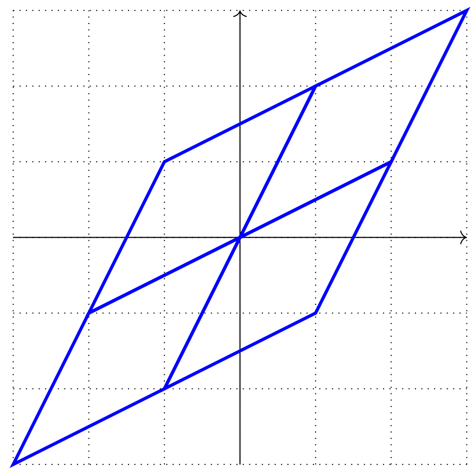
If you know to what vectors the basis vectors are mapped, *then you know how the entire space is altered*!

261. Find $T(1,1)$ by the formula $T(x,y) = (2x+y, x+2y)$.

262. Repeat finding $T(1,1)$ by using linearity, meaning $(1,1) = (1,0) + (0,1)$ and apply the transformation using only the knowledge given by the columns of the matrix.

263. What you see in the two figures preceding these first few exercises is how unit area is altered. This means *all of the unit squares are altered in the same way*. There are 36 unit squares in the first figure (dotted lines). Determine whether or not the four unit squares that surround the origin will be transformed as the picture below depicts.

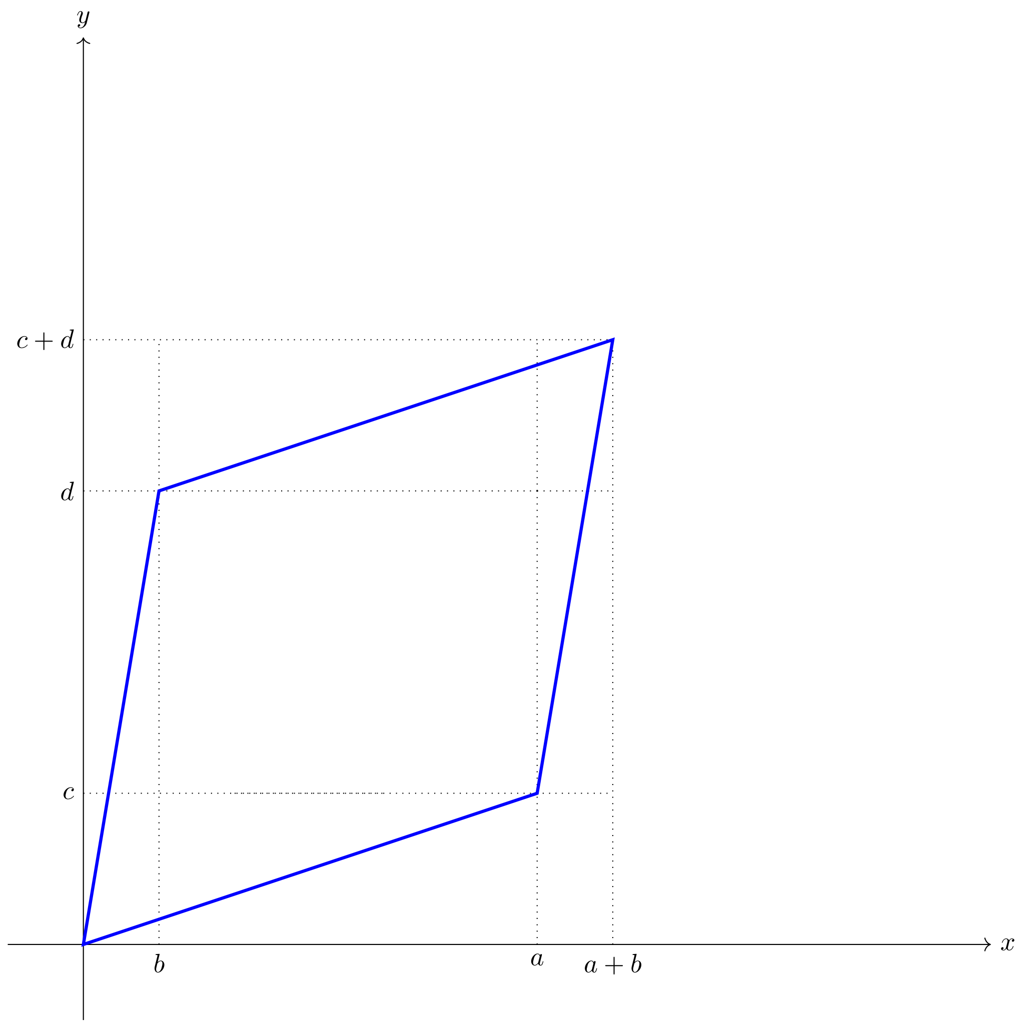
264. Notice in the figure above, that the parallelogram sits inside the area of a larger rectangle (that happens to be a square). Find the area of the parallelogram using grade school mathematics: Think "area of whole thing minus area of part will equal area of parallelogram."



If you use quadrille graph paper, ideally you imagine all the squares you see that make the graph paper what it is as unit squares. It represents the Euclidean vector space $\mathbb{R}^2$, the xy-plane, and it is built by the standard basis vectors. You can think of the matrix as transforming the coordinate system for the space, creating a new kind of graph paper in which the two major axes point diagonally and are not perpendicular to each other.

15.2 An Origin Story

There is a joke I repeatedly tell students, when the occasion arises: *Did you know that if you clean a vacuum cleaner, then you are a vacuum cleaner?* Another thought that I hope makes you chuckle, even if it is a little laugh, is this: *The other day, I went to the supermarket and bought a trash can that I put inside a bag. After I got home, I took the trash can out of the bag, and then put the bag inside the trash can.* Let's determine something called the determinant.



$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Assume we are given a matrix A that transforms the unit square in a similar manner as the matrix in the previous discussion. We can see that the matrix maps the vector $(1, 0)$ to the vector (a, c) and maps the vector $(0, 1)$ to the vector (b, d). The unit square is transformed to a parallelogram.

265. Find the area of the parallelogram using the same method used in 264: whole area minus part equals parallelogram area.

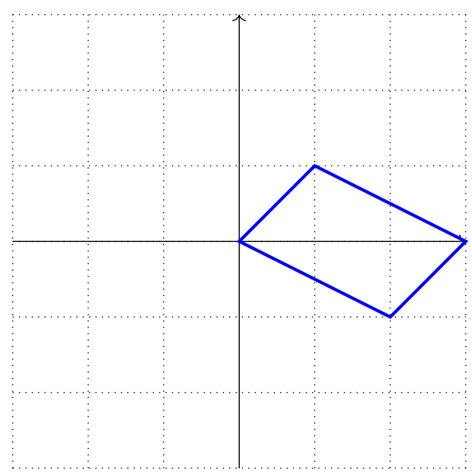
The area you found required some algebra, of course: The rectangle of area $(a+b)(c+d)$ minus the area of six other basic shapes (two triangles, each with area $\frac{ac}{2}$, two triangles, each with area $\frac{bd}{2}$, and two rectangles, each with area bc) results in the area of the parallelogram:

$$\begin{aligned}(a+b)(c+d) - ac - bd - 2bc &= ac + ad + bc + bd - ac - bd - 2bc \\ &= ad - bc\end{aligned}$$

Many of you were taught how to find what is called the determinant of a 2×2 matrix: Subtract the products of its diagonal elements: product of the main diagonal elements a and b subtracted by the product of the elements on the other diagonal b and c, a quick way to obtain the number $ad - bc$. How many of you were taught where the determinant comes from?! We will officially define the *determinant* of a matrix soon. Such a thing ought to *determine* something.

Example 66: Sketch the unit square in the xy-plane. Then sketch the image of the unit square when the matrix $M = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$ transforms the plane, and find the area of this image. Give the formula for the linear transformation using $T_M : \mathbb{R}^2 \to \mathbb{R}^2$.

partially written solution: I'll let you, the student, sketch the unit square. The columns of the matrix inform us that the image of the basis vector $(1, 0)$ is $(2, -1)$, and the image of the basis vector $(0, 1)$ is $(1, 1)$. Simply sketch these two vectors in the plane, and then sketch the parallelogram they determine. Your parallelogram should have vertices the origin, $(1, 1), (3, 0)$, and $(2, -1)$. Based on the formula we computed, the area of this parallelogram is $2(1) - (-1)(1) = 2 + 1 = 3$, so unit area is tripled under the matrix transformation.



The formula for the transformation $T_M : \mathbb{R}^2 \to \mathbb{R}^2$ is $T_M(x, y) = (2x + y, y - x)$.

Example 66:[10] Using the same matrix M as in the previous example, find the image of the vector $(1, 3)$, and sketch both vectors in the same coordinate plane.

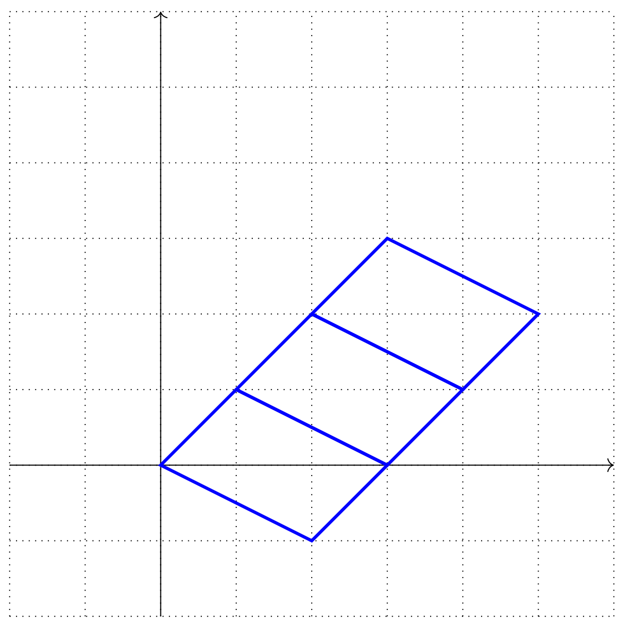
[10]The discussion in this example is *important*.

solution and some things to notice: The image of a vector $\mathbf{x}$ can be found by finding $T_M(\mathbf{x})$ or by multiplying the matrix by the vector: $M\mathbf{x}$:

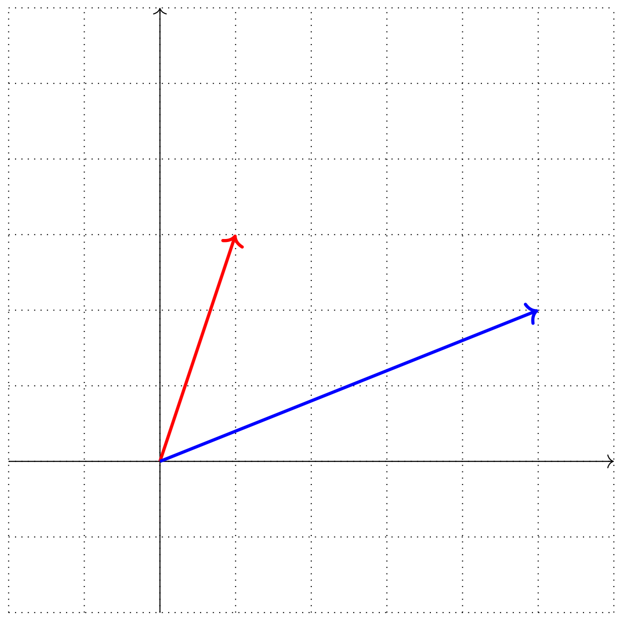
$$T_M(1,3) = (2(1)+3, 3-1) = (5,2)$$

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

Notice something here: The parallelogram in the previous example can serve as the basic building block of the plane in the same way that the unit square makes up graph paper. All scalar multiples of $(1,0)$ form the x-axis, and all scalar multiples of $(0,1)$ form the y-axis; You can say that the matrix M transforms the x-axis into a new axis parallel to the longer side of the parallelogram, while the y-axis becomes an axis parallel to the shorter side. To find $T_M(1,3)$, from the origin, go "right-ish one" long length on the parallelogram and then "up-ish three" short lengths, and notice that you land on $(5,2)$:



The example asks us to sketch the vector and its image in the same plane, which opens up another discussion that is vital to understanding some of the big things in the course.



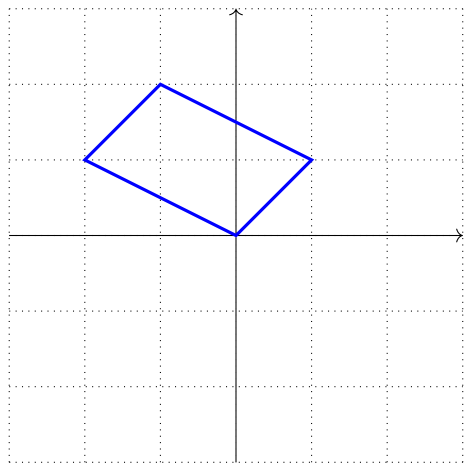
The red vector is transformed to the blue vector by the matrix. It ought to make sense to say that the matrix does two things in the transformation: it *rotates* the red vector, and then *stretches* the vector. In fact, one of the biggest, and most powerful, results in linear algebra is the fact that any matrix can be written as a product of matrices that rotate and stretch. There are two questions about what matrices do to vectors in a vector space, and these questions are central to the subject: *What vectors remain pointing in the same direction after a matrix transforms the space? If there are vectors that are perpendicular to each other before the transformation, are there vectors perpendicular to each other after the transformation?* We will come back to these questions as they lead to some of the most important mathematical tools in machine learning and data analysis.

Example 67: Sketch the unit square in the xy-plane. Then sketch the image of the unit square when the matrix $P = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ acts on the plane, and find the area of this image. Give the formula for the linear transformation $T_P : \mathbb{R}^2 \to \mathbb{R}^2$. Find the kernel of the transformation, which is the null space of the matrix P, and find the range of the transformation.

partial solution: I will leave it up to the student to show the work details here, and to realize that the image of unit square is a *line segment* from the origin to the point $(3, 6)$; in other words, unit area gets squashed. Notice that this agrees with the area formula we computed, that $1(4) - 2(2) = 0$. The formula for the linear transformation is $T_P(x, y) = (x + 2y, 2x + 4y)$. The null space of the matrix (the kernel of the transformation) is the set of all vectors of the form $(-2a, a)$, the line $x = -2y$. The range of the transformation is the span of the column vectors of P, which is all scalar multiples of $(1, 2)$, otherwise known as the line $y = 2x$. The matrix takes the plane and squashes it to a line.

Example 68: Let $Q = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}$. Sketch the image of the unit square under the matrix transformation. What is the area of this image? Write the formula T_Q for this transformation.

solution: Again, the column vectors of Q are the images of the standard basis vectors, and the image of the unit square is the following parallelogram:



It appears as though our area formula collapses here, since $(-2)(1)-1(1) = -3$. It is *not* because the image of the unit square appears in the second quadrant. The reason for this is due to the fact that the unit square was "flipped" somehow. The area is actually 3, the magnitude of the number computed by our formula. We will get into this later, but that is what a "negative determinant" means geometrically. $T_Q(x,y) = (y-2x, x+y)$.[11]

We are using $\mathbb{R}^2$ to give you exposure to what matrices do to a space, as it is easy to sketch things and visualize. Square matrices are, therefore, the big focus in an introductory course. As we continue, we will see some more matrix transformations. The last two sections of this chapter give you general matrix information for transformations of the plane.

15.3 Rotation Operators

Matrices that rotate vectors along arcs of circles centered at the origin are called *rotation operators.* Your background in trigonometry gives you the power to find the standard matrix for the rotation operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ that moves sets of points counterclockwise about the origin by θ, a positive angle.

266. Sketch the unit square in the plane. *In the same plane*, sketch the image of the unit square under the matrix transformation $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

[11] Note that $T(1,3) = (1,4)$. How would you find the image of the vector in this case using the parallelogram and the idea that it forms a new kind of graph paper?

267. Using the same matrix, sketch the vector $(2, 3)$ and its image under the matrix transformation.

268. Sketch the image of the unit square under the transformation by $\begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$.

269. Show that the standard matrix for rotating all points in the plane counterclockwise by an angle θ is the following matrix, which we will give the name R_θ:

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

270. Find the matrix that defines a rotation of the plane about the origin through each of the following angles. Determine the image of $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ under each transformation: **(a)** π. **(b)** $\frac{\pi}{6}$, and **(c)** $-\frac{\pi}{3}$.

271. Find the equation of the image of the ellipse $9x^2 + 4y^2 = 36$ under a rotation through an angle of $\frac{\pi}{2}$.

272. Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. What does this matrix do to vectors in $\mathbb{R}^2$? (Your job here is to determine an intelligent way to answer this with justification.)

15.4 Dilation and Contraction

Use the following matrices for the next four problems:

$$A = \begin{bmatrix} -7 & 5 \\ -3 & 4 \\ -1 & 9 \end{bmatrix}, B = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, C = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}, D = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

273. Find AB and AC.

274. Find BC and CB.

275. Find BD and CD.

276. What do the matrices B and C do to unit area in $\mathbb{R}^2$?

Consider the transformation $T(\begin{bmatrix} x \\ y \end{bmatrix}) = k \begin{bmatrix} x \\ y \end{bmatrix}$, one that maps every point in the plane k times as far from the origin. If $k > 1$, points are moved further away from the origin and the transformation is known as a dilation. If $0 < k < 1$, points are moved closer to the origin and the transformation is known as a contraction. The standard matrix of the transformation is a *diagonal matrix*:

$$\begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$$

Most of us use the word "stretched by a factor of k" rather than *dilation* or *contraction*.

277. Determine a single matrix that defines both rotation about the origin through an angle θ and a dilation of factor k.

278. Let T_1 be a rotation and let T_2 be a reflection in the x-axis. Are these transformations commutative? Discuss this geometrically first, and then algebraically.

279. Find and sketch the image of the unit square under the transformation $T(\mathbf{u}) = A\mathbf{u} + \mathbf{v}$ if A is a diagonal matrix with $k = 2$ and $\mathbf{v} = (4, 4)$.

280. Prove that $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x, y, z) = (0, y, 0)$ is linear. This transformation is known as a *projection*. Explain why this term is appropriate.

16 Answers to Selected Problems

Please note that the popular way to write solutions to systems is to write them as column vectors: For example, in problem 1, you will learn to write this solution as $\mathbf{x} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$.

1. $x = -4, y = 2$
2. $x = -1, y = -1$
3. $x = -1, y = 2$
4. $x = 3, y = 1$
5. $x = 4, y = 2$
6. no solution
7. Prior to this class, if you studied systems beyond two variables, you most likely wrote "infinitely many solutions," but in this class you want to understand exactly what you mean by that: here, the solutions set is the set of all vectors of the form $(2k-5, k)$, where k is any scalar. We will write this in a special form:
$$\mathbf{x} = \begin{bmatrix} -5 \\ 0 \end{bmatrix} + k \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$
8. $x = 1, y = 1, z = 1$
9. $x = 1, y = 1, z = 2$
10. the set of all vectors of the form $(1, k, k)$, where k is any scalar, which, when we learn to write this in column vector form will be
$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + k \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \text{ where } k \text{ is any scalar.}$$
11. $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + k \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, where x_3 is any scalar (free variable).
12. $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$
13. $\mathbf{x} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$
14. $\mathbf{x} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

15. $\mathbf{x} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$

16. no solution

17. $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

18. $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$

Please keep in mind that these first 18 problems were your first look at systems where you had not yet developed the current perspective you now have (if you have been properly working in the course so far, meaning completing the homework and reading this document carefully).

Answers not provided for HOMEWORK 3: problems 19-33.

HOMEWORK 4

34. $\mathbf{x} = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}$

35. $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

36. $\mathbf{x} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}$

37. $\mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, where x_3 is any scalar (free variable).

38. no solution (last row of matrix in rref reads 0=1)

39. no solution (last row of matrix in rref reads 0=1)

40. $\begin{bmatrix} 0 \\ \frac{3}{2} \\ 4 \end{bmatrix}$

41. $\mathbf{x} = \begin{bmatrix} 4 \\ 6 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix}$, where x_3 is any scalar (free variable).

42. $\mathbf{x} = \begin{bmatrix} 3.5 \\ 0.25 \\ 4 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -0.5 \\ -0.25 \\ -2 \\ 1 \end{bmatrix}$, where x_4 is any scalar (free variable).

43. Note here that the augmented matrix for the system is a 4×4 matrix and, after you perform all of the operations, that matrix reduces to a matrix where you have ones across the main diagonal and zeros everywhere else. In *this* system, which has three variables but four equations, the last row reveals the system has no solution. How so?

44. $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$

45. In a two-step process, using vector addition then scalar multiplication, you write $(x, y) = (x, 0) + (0, y) = x(1, 0) + y(0, 1)$, and from that result, you have shown that every vector in the plane can be written as a linear combination of $(1, 0)$ and $(0, 1)$. This shows that these two vectors span $\mathbb{R}^2$.

HOMEWORK 5

46. $x_3 \begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix}$, a line through the origin in $\mathbb{R}^3$

47. $x_4 \begin{bmatrix} -0.5 \\ -0.25 \\ -2 \\ 1 \end{bmatrix}$, a line through the origin in $\mathbb{R}^4$

48. $x_4 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, a line through the origin in $\mathbb{R}^4$

49. $x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, a plane through the origin in $\mathbb{R}^4$

50. $x_3 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, a plane through the origin in $\mathbb{R}^4$

51. I do a sneaky thing in this book throughout the term: occasionally reveal some great secret about an upcoming exam problem or give away something, or make a comment with the intention of reminding the class

about something important or draw awareness to a detail; that is, I like to see who is actually reading this book. This particular "problem" is clarification of the vector notation being used.

52. All of these systems are of the form $A\mathbf{x} = \mathbf{0}$ and all of the solutions are the span of a set of vectors; that is, all of the solutions are all possible linear combinations of a set of vectors; that is, all of the solutions are line-like things through origins in the respective Euclidean vector space they are located.

HOMEWORK 6 Problems 53-67 can be checked on your own. Use a calculator or the desmos matrix calculator I showed you in class to enter the matrix given to then find its rref... The rref of the matrix on the calculator should match what you got after doing the row operations. BE SURE YOU ARE WRITING DOWN ALL ROW OPERATIONS IN YOUR WORK. These are crucial to have written down as you rewrite the matrix. On exams, not showing the row operations for a matrix you are writing in rref consitutes a problem for which you will receive zero credit.

68. The matrix in 66, If it represents a system, the matrix in 66 represents a system of four equations and three unknowns. Your rref of this matrix can be read carefully to give the solution to that system. The system is

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 3x_1 + 5x_2 + 8x_3 &= 2 \\ 4x_2 + 4x_3 &= 4 \\ x_1 + x_2 + 2x_3 &= 0 \end{aligned}$$

and the solution to the system is $\mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$.

HOMEWORK 7

71. The line $y = x$ in $\mathbb{R}^2$ or "all vectors in $\mathbb{R}^2$ of the form (a, a), where a is any real number." or "all vectors in $\mathbb{R}^2$ that have the same two components" or....

72. The xz-plane or all vectors in $\mathbb{R}^3$ of the form $(a, 0, b)$ where a and b are any real numbers.

73. the same answer as the one in problem 71

74. $\mathbb{R}^2$

75. all vectors in $\mathbb{R}^3$ of the form (a, a, b) where a and b are any real numbers; i.e., all vectors in $\mathbb{R}^3$ whose first two components are equal

76. all vectors in $\mathbb{R}^4$ of the form $(a, b, 2a + 2b, b)$, where a and b are any real numbers

77. $\mathbb{R}^2$

78. $(1, 0, 1)$ and $(0, 1, 1)$

79. $(1, 1, 0)$ and $(0, 0, 1)$

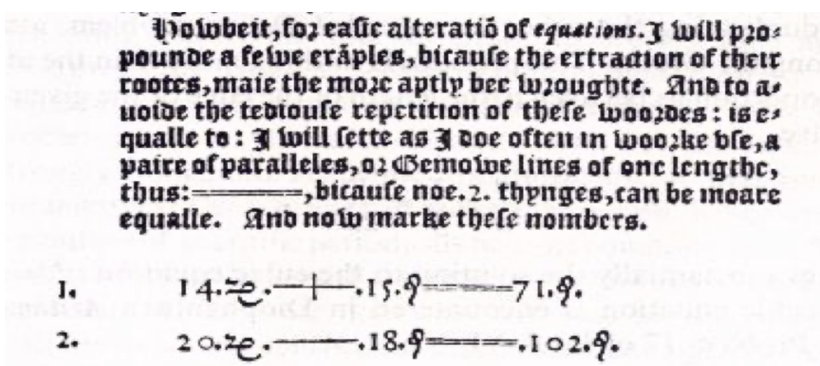
80. $(1, 0, 1)$ and $(0, 1, 0)$

81. Same answer as 80, but *remind me to talk about these two problems 80 and 81 in class!*

82. 2×3

89. $(1, 3)$ is one of the vectors that span the line

90. $z = 2x - 3y$ is an equation that can be read as the set of all points (ordered triples) of the form $(x, y, 2x - 3y)$, where x and y are any real numbers. Note that vector addition and scalar multiplication allow you to write $(x, y, 2x - 3y) = (x, 0, 2x) + (0, y, -3y) = x(1, 0, 2) + y(0, 1, -3)$, which means that the plane is spanned by the vectors $(1, 0, 2)$ and $(0, 1, -3)$.



Howbeit, for easie alteratiō of *equations*. I will pro-
pounde a fewe exāples, bicause the extraction of their
rootes, maie the more aptly bee wroughte. And to a-
uoide the tediouse repetition of these woordes : is e-
qualle to : I will sette as I doe often in woorke use, a
paire of paralleles, or Gemowe lines of one lengthe,
thus: ═══════, bicause noe .2. thynges, can be moare
equalle. And now marke these nombers.

1. 14.ᵹ.—+—.15.ϑ═══════71.ϑ.
2. 20.ᵹ.———.18.ϑ═══════.102.ϑ.

Figure 3: Robert Recorde's Reason for Inventing "is equal to" Symbol