MATH 544, SPRING 2022

CONTENTS

1. Introductory remarks about the course.

Here are some preliminary remarks about the course.

(1) My name is Professor Kustin. (My last name rhymes with "Justin".)
(2) Quiz 1 on Wednesday, January 19 is one of the assigned HW problems from 1–4.
(3) I have taught Math 544 many times from many textbooks. The book that bothers me the least is by Johnson-Riess-Arnold. The early editions were relatively inexpensive and had a minimal amount of junk. The later editions have become very expensive and have a great deal of junk. If you feel that you need a text book be sure to get a book that emphasizes vector spaces (rather than matrix calculations). If you find a book that discusses vector spaces, linear independence, and basis near the beginning, it will probably be okay. In particular, the various editions of the book by Johnson-Riess-Arnoldis are okay. It is not required that you have a book. I have posted a complete set of lecture notes and I will post Homework problems.
(4) There will be an Exam or Quiz essentially every other class. The exams and quizzes will be given at the end of class. When you finish your quiz or exam, **take a picture of your solution for your records** and give me your answers. I will send my comments back by way of e-mail. In general, I won't return papers.
(5) If you miss an exam or quiz or do poorly on an exam or quiz; don't worry about it. There will be plenty more chances for you to demonstrate competence. Be sure to learn how to do missed or wrong problem correctly. I'll surely ask about it again.
(6) I plan to post typed versions of the class lectures on my website. If you miss a class, the typed version of my lectures will be of some use to you. (They might also be of use if you attend class. Sometimes what I convey in class is not as articulate as what I type sitting at my desk.) I encourage you to form partnerships with your classmates and share notes that are taken from class. (Surely, there will be times that I think some something helpful to add as I am lecturing. In particular, my answers to student's questions will not make their way into the typed notes.) I do not mind if you take pictures of my lectures on the whiteboard.
(7) To make the class work, please do the following.
    (a) Master every lecture.
    (b) Do every homework problem.
    (c) Ask questions. There are many ways to ask questions: raise your hand in class, send me an e-mail, leave me a note (either in my office or at the front of the classroom), etc. Do whatever works for you.

   (d) Learn from your mistakes.
   (e) Don't give up.

(8) The course is a mix of calculation and theory. You are supposed to have had Math 300, which is the course about "How to prove mathematical statements". Traditionally, Math 544 and Math 574 are the first 500 level Math courses that students take. These two courses are a mix of theory and calculation; whereas, Math 546 and 554 are straight theory.
   (a) I believe that theory and calculation reinforce one another. As one makes calculations, one learns how and why things work. As one learns theory, one learns how to streamline calculations.
   (b) **Learn every definition.** Learn the official definition (because that is what one uses to prove things). Also learn the intuitive idea behind each concept (because we are human beings and not robots).
   (c) Be sure to look at the old exams. Notice the high percentage of the problems that are: "Prove", "True or False", "Define". I am especially fond of True or False questions. I believe that these are the questions that confront us most in life. Our co-worker says "blah blah blah" to us; we are left trying to figure out if this makes any sense. Fortunately, for elementary mathematical assertions one can usually produce a counterexample or a proof.

(9) There is a linear algebra course that does not have any theory and does not have any proofs and has a prerequisite of Math 142 (rather than Math 300). That course is called Math 344, "Applied Linear Algebra".

Linear algebra is the study of vector spaces. A <u>vector space</u> is a set of things that can be added and multiplied by scalars. Of course "addition" and "scalar multiplication" must satisfy certain rules. Eventually, I will tell you the official rules. For the time being, lets just say "the usual rules". For our present purposes, the most important rules are the closure rules. When you add two elements of a vector space, you get another element of the vector space. When you multiply an element of the vector space by a real number scalar, you get an element of the vector space.

Examples. (1) The set of directed line segments in the *xy*-plane forms a vector space. (In this situation two "vectors" are considered the same of they have the same length and direction. One can move a vector as long as one does not change its length or direction.) Such vectors can be added and multiplied by scalars. These vectors are studied in Math 241.

(2) The set of column vectors with 2 entries

$$\left\{ \begin{bmatrix} a \\ b \end{bmatrix} \middle| \, a \text{ and } b \text{ are real numbers} \right\}$$

forms a vector space that we call $\mathbb{R}^2$. One adds two vectors component-wise and one multiplies a vector by a scalar component-wise.

(3) The set of column vectors with n entries

$$\left\{ \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \middle| a_1, \dots, a_n \in \mathbb{R} \right\}$$

forms a vector space that we call $\mathbb{R}^n$. Again addition and multiplication take place component-wise.

(4) Consider a system of m homogeneous linear equations in n unknowns:

$$\begin{cases} c_{11}x_1 + \cdots + c_{1n}x_n = 0 \\ \qquad\vdots \\ c_{m1}x_1 + \cdots + c_{mn}x_n = 0. \end{cases} \tag{1.0.1}$$

(In this system of equations the c_{ij} are real numbers. They are called the coefficients of the equations. The x_j are the unknowns. The equations are homogeneous and linear because every (non-zero) term that appears has degree EXACTLY one in the unknowns.) The set of solutions of this set of equations is a vector space. Indeed, if

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

each are a solution of (1.0.1).[1] Then

$$\begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{bmatrix}$$

[1]This means that

$$\begin{cases} c_{11}a_1 + \cdots + c_{1n}a_n = 0 \\ \qquad\vdots \\ c_{m1}a_1 + \cdots + c_{mn}a_n = 0. \end{cases}$$

and

$$\begin{cases} c_{11}b_1 + \cdots + c_{1n}b_n = 0 \\ \qquad\vdots \\ c_{m1}b_1 + \cdots + c_{mn}b_n = 0. \end{cases}$$

is a solution of (1.0.1)[2] and

$$\begin{bmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{bmatrix}$$

is a solution of (1.0.1).[3] for each real number λ.

(5) The set of differentiable functions $f : \mathbb{R} \to \mathbb{R}$ forms a vector space. (This is the object studied in Math 141.)

(6) The set of solutions of a homogeneous linear differential equation, like

$$y' + y = 0,$$

forms a vector space. (This vector space is studied in Math 242 and the techniques you used to solve homogeneous linear differential equation are vector space techniques.)

(7) Fix $D(x)$ a polynomial with real coefficients. Consider the set of rational functions of the form

$$\frac{c_0 + \cdots + c_n x^n}{D(x)}$$

where n is arbitrary and $c_0, \ldots, c_n$ are arbitrary real numbers. This set forms a vector space. In math 142 you learned how to integrate every element in the set using the technique of Partial Fractions. The game of partial fractions is a vector space game.

---

[2]Plug

$$\begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{bmatrix}$$

into the top equation of (1.0.1), get

$$c_{11}(a_1 + b_1) + \cdots + c_{1n}(a_n + b_n).$$

Distribute and regroup to get

$$\Big(c_{11}(a_1) + \cdots + c_{1n}(a_n)\Big) + \Big(c_{11}(b_1) + \cdots + c_{1n}(b_n)\Big) = 0 + 0 = 0.$$

The same game works for each equation in (1.0.1), not just the top equation.

[3]Again, one plugs

$$\begin{bmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{bmatrix}$$

into the left side of each equation of (1.0.1) and then factors out the λ to see that

$$\begin{bmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{bmatrix}$$

actually works in each equation of (1.0.1).

2. Systems of linear equations.

Warm up. Here are three easy examples of systems of equations. It is easy to solve each of these systems and we get three different types of answers.

Example One. The system of linear equations

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

has exactly one solution; namely the point $(1,0)$. (Actually, we write the solution as the vector $\begin{bmatrix}1\\0\end{bmatrix}$.) The corresponding picture is two lines in the xy plane that meet at a point. (See the page of pictures that follows the three examples.)

Example Two. The system of linear equations

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

has an infinite number of solutions; namely, every point on the line $x+y=1$ is also on $2x+2y=2$. The corresponding picture is one line in the xy-plane. (See the page of pictures that follows the three examples.)
Typically we will attack Example Two by saying that the solution set of Example Two is the same as the solution set of

$$x+y=1. \tag{2.0.1}$$

The solution set of (2.0.1) is the same as the solution set of

$$\begin{cases} x=1-y\\ y=\phantom{1-{}}y. \end{cases} \tag{2.0.2}$$

The solution set of (2.0.2) is

$$\boxed{\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}+y\begin{bmatrix}-1\\1\end{bmatrix},\ \text{where } y \text{ is free to take any value.}}$$

Example Two and a half. Solve

$$\begin{cases} x_1+2x_2++4x_4=6\\ \phantom{x_1+2x_2+{}}x_3+5x_4=7. \end{cases} \tag{2.0.3}$$

Typically, we would say: the solution set of (2.0.3) is the same as the solution set of

$$\begin{cases} x_1=6-2x_2-4x_4\\ x_2=x_2\\ x_3=7-5x_4\\ x_4=x_4 \end{cases} \tag{2.0.4}$$

The solution set of (2.0.4) is

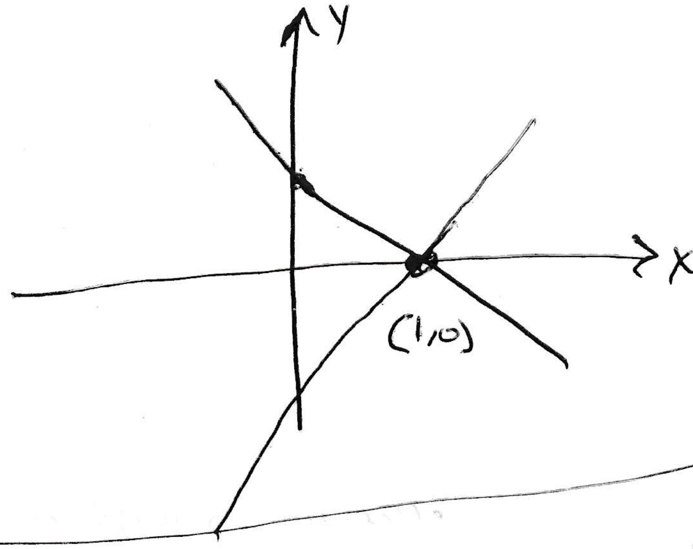
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 7 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -4 \\ 0 \\ -5 \\ 1 \end{bmatrix}, \text{ where } x_2 \text{ and } x_4 \text{ are free to take any value.}$$

Example Three. The system of equations

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

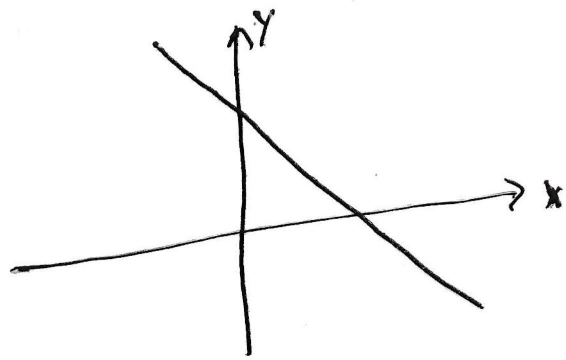
has no solution. The corresponding picture is two parallel lines in the xy-plane. (See the next page.)

(a) $\begin{cases} x + y = 1 \\ x - y = 1 \end{cases}$



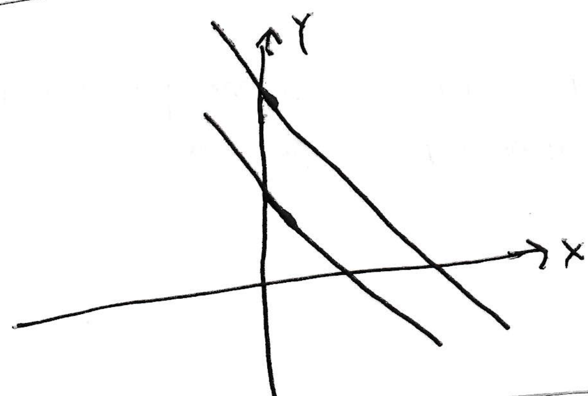
The solution is one point.

(b) $\begin{cases} x + y = 1 \\ 2x + 2y = 2 \end{cases}$



The solution is the entire line $x+y=1$, hence there are a infinite number of points in the solution

(c) $\begin{cases} x + y = 1 \\ x + y = 2 \end{cases}$



This system of equations has no solution. There are no points that are on both of these parallel lines

A system of linear equations has either

(a) a unique solution,

(b) infinitely many solutions, or

(c) no solution

To solve a system of linear equations we replace the given system of equations by an "easier" system of equations with the same solution step. We continue this process until we have a system of equations that is is "easy" as possible. Then we read off the answer.

The technique that I am about to show is called **Gaussian elimination** or **Gauss-Jordan elimination**. (It is how we solve systems of linear equations.)

It turns out that if we apply the procedure to

$$\begin{cases} x_1 + \ 3x_2 + \ 7x_3 = \ 28 \\ 2x_1 + \ 7x_2 + 16x_3 = \ 64 \\ 3x_1 + 11x_2 + 26x_3 = 103, \end{cases} \tag{2.0.5}$$

then we end up with

$$\begin{cases} x_1 + 0x_2 + 0x_3 = 1 \\ 0x_1 + \ x_2 + 0x_3 = 2 \\ 0x_1 + 0x_2 + \ x_3 = 3. \end{cases} \tag{2.0.6}$$

We read that the original system of equations has exactly one solution and that solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Projects 2.1. *We have three projects.*

(a) *What operations can we perform on a system of equations that yield a system of equations with the same solution set?*
(b) *When do we stop?*
(c) *How do we read off the answer?*

Operations on systems of linear equations which do not change the solution set. For the time being, lets call these the "elementary equation operations". (Usually, we will deal with this problem using matrices instead of equations. Our usual name for these operations is "elementary row operations".)

(1) We may exchange two equations.
(2) We may multiply an equation by a non-zero constant.
(3) We may add a multiple of one equation to another equation.

Please notice that these operations yield a system of linear equations with the exact same solution set as the original system of equations. Each of these elementary operations can be undone using another elementary operation.

We now apply elementary equation operations to (2.0.5) in order to obtain (2.0.6).

I like the $1x_1$ in the first equation. Let us use this $1x_1$ to get rid of the $2x_1$ in the second equation and the $3x_1$ in the third equation. In other words, we replace

Equation 2 by Equation 2 -2 times Equation 1, and we replace Equation 3 with Equation 3 -3 times Equation 1. We obtain

$$\begin{aligned} x_1 + 3x_2 + 7x_3 &= 28 \\ x_2 + 2x_3 &= \ \ 8 \\ 2x_2 + 5x_3 &= 19. \end{aligned}$$

I like the $1x_2$ in the second equation. Let us use this $1x_2$ to get rid of the $3x_2$ in the first equation and the $2x_2$ in the third equation. Replace Equation 1 by Equation 1 -3 times Equation 2. Replace Equation 3 by Equation 3 -2 times Equation 2. We obtain

$$\begin{aligned} x_1 \qquad + \ \ x_3 &= 4 \\ x_2 + 2x_3 &= 8 \\ x_3 &= 3. \end{aligned}$$

I like the $1x_3$ in the third equation. We use this $1x_3$ to get rid of x_3 in the first equation and $2x_3$ in the second equation. Replace Equation 1 by Equation 1 minus Equation 3. Replace Equation 2 by Equation 2 -2 Equation 3. We obtain

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

This is the system of equations (2.0.6). The unique solution of the system of equations (2.0.5) is

$$\boxed{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.} \tag{2.1.1}$$

Check Of course, (2.1.1) is the solution of (2.0.5) because

$$\begin{aligned} 1 + \ \ 3(2) + \ \ 7(3) &= \ \ 28 \\ 2 + \ \ 7(2) + 16(3) &= \ \ 64 \\ 3 + 11(2) + 26(3) &= 103. \end{aligned}$$

There is no particular reason to write the x's and the equal signs.

We write the system of equations (2.0.5) as an **augmented matrix**:

$$\left[\begin{array}{ccc|c} 1 & 3 & 7 & 28 \\ 2 & 7 & 16 & 64 \\ 3 & 11 & 26 & 103 \end{array}\right] \tag{2.1.2}$$

and then apply Elementary Row Operations (EROs).

(1) We may exchange two rows.
(2) We may multiply a row by a non-zero constant.
(3) We may add a multiple of one row to another row.

Please notice that if one applies any of these elementary row operations then the solution set of the system of equations that corresponds to the original matrix is exactly the same as the solution set that corresponds to the new system of

equations. Each of these elementary row operations can be undone by another elementary row operation.

If we apply EROs to (2.1.2) we obtain:
$R2 \mapsto R2 - 2R1$
$R3 \mapsto R3 - 3R1$

$$\left[\begin{array}{ccc|c} 1 & 3 & 7 & 28 \\ 0 & 1 & 2 & 8 \\ 0 & 2 & 5 & 19 \end{array}\right].$$

$R1 \mapsto R1 - 3R2$
$R3 \mapsto R3 - 2R2$

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 3 \end{array}\right].$$

$R1 \mapsto R1 - R3$
$R2 \mapsto R2 - 2R3.$

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

We reinterpret the most recent matrix to say the system of equations (2) has a unique solution namely

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Project 2.1.(b). On page 9 we identified three projects. We have carried out Project 2.1.(a). Now we do Project 2.1.(b).

We must answer the following question. Suppose you want to solve a system of equations. You have written the system of equations as an augmented matrix. You are applying Elementary Row Operations. When do you stop?

The answer: Stop when the matrix is in “row echelon form” or “reduced row echelon form”.

Roughly speaking, a matrix is in row echelon form when it is as much like an upper triangular matrix with one’s on the main diagonal as possible. A matrix is in reduced row echelon form when the matrix is as much like the identity matrix as possible.

Here are some examples. The matrices in the left column are in row echelon form. The matrices in the right column are in reduced row echelon form. The symbol $*$ can be any number.

row echelon form	reduced row echelon form
$\begin{pmatrix} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 1 & * & * \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & * & * \\ 0 & 1 & 0 & * & * \\ 0 & 0 & 1 & * & * \end{pmatrix}$
$\begin{pmatrix} 1 & * & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & * & 0 & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$
$\begin{pmatrix} 1 & * & * & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & * & * & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

We give a complete description of Row Echelon Form and Reduced Row Echelon Form.

Definition. A matrix is in <u>Row Echelon Form</u> (REF) if

(i) the left-most non-zero entry in each row is a one,
(ii) the leading one in a row k appears to the right of the leading one in row $k-1$, and
(iii) the rows which consist of all zeros are at the bottom.

A matrix is in <u>Reduced Row Echelon Form</u> (RREF) if the matrix is in Row Echelon Form and every entry above every leading one is also zero.

Remark. The word echelon conveys the notion of a hierarchy. Here is what the Merriam-Webster dictionary says:

(1) an arrangement of a body of troops with its units each somewhat to the left or right of the one in the rear like a series of steps;
(2) a formation of units or individuals resembling such an echelon geese flying in echelon;
(3) a flight formation in which each airplane flies at a certain elevation above or below and at a certain distance behind and to the right or left of the airplane ahead.

The dictionary also reminded me of the phrase "the upper echelons of management".

Remark. If you are actually solving a system of equations, we probably want to put the matrix in Reduced Row Echelon Form. If you only want to know how many solutions a system of equations has, you can stop with the matrix in Row Echelon Form (or maybe even a little less.)

Project 2.1.(c). On page 9 we identified three projects. We have carried out Projects 2.1.(a) and 2.1.(b). Now we do Project 2.1.(c).

The question is the following. Suppose that a matrix is in Reduced Row Echelon Form. What is the solution set of the corresponding system of equations? We will answer the question by exhibiting four examples.

Example One. The system of equations that corresponds to

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

has exactly one solution and that solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Example Two. The system of equations that corresponds to

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 4 \\ 0 & 1 & 0 & 2 & 5 \\ 0 & 0 & 1 & 3 & 6 \end{array}\right)$$

is

$$\begin{aligned} x_1 + x_4 &= 4 \\ x_2 + 2x_4 &= 5 \\ x_3 + 3x_4 &= 6 \end{aligned} \tag{2.1.3}$$

We rewrite these equations as

$$\begin{aligned} x_1 &= 4 - x_4 \\ x_2 &= 5 - 2x_4 \\ x_3 &= 6 - 3x_4 \\ x_4 &= x_4. \end{aligned}$$

The solution set of (2.1.3) is the set

$$\boxed{\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \;\middle|\; \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 6 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -2 \\ -3 \\ 1 \end{bmatrix} \text{ where } x_4 \text{ can be any real number.} \right\}}.$$

This is a line in four-space.

Example Three. The system of equations that corresponds to

$$\left(\begin{array}{cccc|c} 1 & 2 & 0 & 3 & 5 \\ 0 & 0 & 1 & 4 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

is

$$\begin{aligned} x_1 + 2x_2 + 3x_4 &= 5 \\ x_3 + 4x_4 &= 6 \end{aligned} \tag{2.1.4}$$

We rewrite these equations as

$$\begin{aligned} x_1 &= 5 - 2x_2 - 3x_4 \\ x_2 &= \phantom{5 -{}} x_2 \\ x_3 &= 6 \phantom{{}- 2x_2} - 4x_4 \\ x_4 &= x_4. \end{aligned}$$

The solution set of (2.1.4) is the set

$$\boxed{\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \;\middle|\; \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 6 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ -4 \\ 1 \end{bmatrix}, \begin{array}{l} \text{where } x_2 \text{ and } x_4 \text{ can be} \\ \text{any real numbers.} \end{array} \right\}.}$$

This is a plane in four space.

Example Four. The system of equations that corresponds to

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 5 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

is

$$\begin{aligned} x_1 + 2x_2 + 3x_3 \phantom{{}+ x_4} &= 5 \\ x_4 &= 6 \\ 0 &= 1. \end{aligned}$$

This system of equations $\boxed{\text{has no solution}}$ because 0 is not equal to 1.

3. MATRICES.

Definition 3.1. A matrix is a rectangular array of numbers. Two matrices are equal if they have the same shape and the corresponding entries are equal. In particular if A is a matrix with r rows, c columns, and a_{ij} in row i and column j and B is a matrix with p rows, q columns, and b_{ij} in row i and column j, then the matrices A and B are equal if $r = p$, $c = q$, and $a_{ij} = b_{ij}$ for all i and j.

One multiplies a matrix by a scalar component-wise. If two matrices have the same shape (that is they have the same number of rows and the same number of columns), then one adds these matrices component-wise:

$$2\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}=\begin{bmatrix}2&4&6\\8&10&12\end{bmatrix}$$

and

$$\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}+\begin{bmatrix}11&12&13\\14&15&16\end{bmatrix}=\begin{bmatrix}12&14&16\\18&20&22\end{bmatrix}.$$

"Matrix multiplication consists of many dot products." That is, if $A = (a_{ij})$ is[4] an $m\times n$ matrix[5] and $B = (b_{ij})$ is a $p\times q$ matrix, then AB makes sense if $n = p$; furthermore, if $n = p$, then AB is an $m\times q$ matrix[6] and the entry in row r and column c of AB is

$$a_{r1}b_{1c}+a_{r2}b_{2c}+\cdots+a_{rn}b_{nc}. \tag{3.1.1}$$

$$\begin{bmatrix}1&2&3\end{bmatrix}\begin{bmatrix}4\\5\\6\end{bmatrix}=1\cdot4+2\cdot5+3\cdot6=4+10+18=32.$$

$$\begin{bmatrix}1&2&3\\4&5&6\end{bmatrix}\begin{bmatrix}7\\8\\9\end{bmatrix}=\begin{bmatrix}1\cdot7+2\cdot8+3\cdot9\\4\cdot7+5\cdot8+6\cdot9\end{bmatrix}=\begin{bmatrix}50\\122\end{bmatrix}$$

Problem 3.2. Express

$$\begin{cases} x_1+\ 3x_2+\ 7x_3= 28\\ 2x_1+\ 7x_2+16x_3= 64\\ 3x_1+11x_2+26x_3=103,\end{cases}$$

in the form $Ax = b$ where A and b are matrices of constants and x is a matrix of unknowns.

---

[4]The notation $A = (a_{ij})$ is a matrix is intended to indicate that the entry of A in row i and column j is called a_{ij}.

[5]The phrase A is $m\times n$ matrix is intended to indicate that A has m rows and n columns.

[6]An $m\times n$ matrix times an $n\times q$ matrix is equal to an $m\times q$ matrix. It is as though the middle n annihilate one another.

Answer.

$$A = \begin{bmatrix} 1 & 3 & 7 \\ 2 & 7 & 16 \\ 3 & 11 & 26 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} 28 \\ 64 \\ 103 \end{bmatrix}.$$

Notice that some rules that hold for the multiplication of numbers also hold for the multiplication of matrices; for example:

- matrix multiplication distributes over addition; that is
$$A(B+C) = AB + AC,$$
and
- Matrix multiplication associates; that is,
$$A(BC) = (AB)C.$$

HOWEVER, matrix multiplication is not commutative:

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 6 \end{bmatrix}, \quad \text{but} \quad \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} \text{ does not make sense.}$$

Furthermore, even if both products are legal, they might not be equal:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 10 & 14 \end{bmatrix}, \quad \text{but} \quad \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 8 & 12 \\ 7 & 10 \end{bmatrix}.$$

Also, **it is possible for a matrix product AB to be zero with $A \neq 0$ and $B \neq 0$.** For example,

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} -2 & -8 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Definition 3.3. If $A = (a_{ij})$ is an $m \times n$ matrix then the transpose of A, denoted A^{T} is the $n \times m$ matrix with a_{ji} in row i and column j.

Example 3.4.

$$\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Observation 3.5. *If A and B are matrices and AB makes sense, then $B^{\mathrm{T}}A^{\mathrm{T}}$ makes sense and $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$*

Example 3.6.

$$\left(\begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 & 7 \\ 5 & 8 \\ 6 & 9 \end{bmatrix} \right)^{\mathrm{T}} = \begin{bmatrix} 4+10+18 & 7+16+27 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 32 & 48 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 32 \\ 50 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 7 \\ 5 & 8 \\ 6 & 9 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 4+10+18 \\ 7+16+27 \end{bmatrix} = \begin{bmatrix} 32 \\ 50 \end{bmatrix}.$$

We got the same answer both times, as expected.

Proof. Suppose A is an $m \times n$ matrix. The product AB makes sense; so B is an $n \times q$ matrix for some q. Observe that AB is an $m \times q$ matrix; $(AB)^{\mathrm{T}}$ is a $q \times m$ matrix; and $B^{\mathrm{T}}A^{\mathrm{T}}$ is the (attempted) product of a $q \times n$ matrix times an $n \times m$ matrix. This product makes sense and is equal to a $q \times m$ matrix. Thus, both products make sense and both products are equal to a $q \times m$ matrix.

Now we show that the entry of $(AB)^{\mathrm{T}}$ in row r and column c is equal to the entry of $B^{\mathrm{T}}A^{\mathrm{T}}$ in row r and column c.

Let

$$\begin{cases} a_{ij} \text{ be the entry of } A \text{ in row } i \text{ and column } j \text{ and let} \\ b_{ij} \text{ be the entry of } B \text{ in row } i \text{ and column } j. \end{cases} \tag{3.6.1}$$

Observe that

$$\begin{aligned} [(AB)^{\mathrm{T}}]_{rc} &= (AB)_{cr}, && \text{by Definition 3.3,} \\ &= \sum_{i=1}^{n} a_{ci}b_{ir}, && \text{by (3.1.1).} \end{aligned}$$

On the other hand,

$$\begin{aligned} [B^{\mathrm{T}}A^{\mathrm{T}}]_{rc} &= \sum_{i=1}^{n} (B^{\mathrm{T}})_{ri}(A^{\mathrm{T}})_{ic}, && \text{by (3.1.1),} \\ &= \sum_{i=1}^{n} (B)_{ir}(A)_{ci}, && \text{by Definition 3.3,} \\ &= \sum_{i=1}^{n} b_{ir}a_{ci}, && \text{by (3.6.1),} \\ &= \sum_{i=1}^{n} a_{ci}b_{ir}, && \text{because multiplication of numbers commutes.} \end{aligned}$$

Thus $[(AB)^{\mathrm{T}}]_{rc} = [B^{\mathrm{T}}A^{\mathrm{T}}]_{rc}$ for all r and c with $1 \le r \le m$ and $1 \le c \le q$, and

$$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}.$$

□

Definition 3.7. A matrix A is called symmetric if $A^{\mathrm{T}} = A$.

Example 3.8. The matrix

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

is a symmetric matrix.

We are particularly interested in matrices which consist of exactly one column. We call such matrices vectors or column vectors.

3.9. We call the set of column vectors with m entries $\mathbb{R}^m$.

If x is a column vector, then $\sqrt{x^{\mathrm{T}}x}$ is called the length of x.

For example, the vector $\begin{bmatrix}1\\2\end{bmatrix}$ has length equal to $\sqrt{5}$.

4. Linearly independent vectors and nonsingular matrices.

Recall from 3.9 that $\mathbb{R}^m$ is the set of column vectors with m entries.

There are three parts to this section:

A. Linearly independent vectors;
B. Nonsingular matrices; and
C. The connection between linearly independent vectors and nonsingular matrices.

The concepts "linearly independent vectors" and "nonsingular matrices" are the first theoretical concepts that we meet in this course. However, to answer the questions "Are these vectors linearly independent?" or "Is this matrix nonsingular?", one must determine the number of solutions for a particular system of linear equations.

4.A. **Linearly independent vectors.**

Definition 4.1. The vectors $v_1,\dots,v_p$ in $\mathbb{R}^m$ are <u>linearly independent</u> if the only numbers $c_1,\dots,c_p$ with $\sum_{i=1}^p c_iv_i = 0$ are $c_1 = c_2 = \dots = c_p = 0$. [7]

Question 4.2. *Is $v_1 = \begin{bmatrix}1\\2\end{bmatrix}$ linearly independent?*

Answer. Yes. If c_1 is a number and $c_1v_1 = 0$, then

$$c_1\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix};$$

hence c_1 times 1 equals 0 and c_1 times 2 equals 0. At any rate, if $c_1v_1 = 0$, then c_1 **has to be** the real number zero. The vector v_1 is linearly independent.

Question 4.3. *Are the vectors $v_1 = \begin{bmatrix}1\\2\end{bmatrix}$ and $v_2 = \begin{bmatrix}2\\4\end{bmatrix}$ linearly independent?*

Answer. No. There is a non-trivial linear combination[8] of v_1 and v_2 which is equal to zero; namely, $2v_1 - v_2 = 0$. I presume that this non-trivial linear combination leapt into your eyes. If so, that is great. If not, you say, "Hmm. I don't know. I'll have to make a calculation." Then you calculate all numbers c_1 and c_2 with

$$c_1v_1 + c_2v_2 = \begin{bmatrix}0\\0\end{bmatrix}$$

You solve the equations

$$\begin{aligned}1c_1 + 2c_2 &= 0\\2c_1 + 4c_2 &= 0\end{aligned}$$

---

[7] Please notice that 0 in the equation $\sum_{i=1}^p c_iv_i = 0$ is the zero vector in $\mathbb{R}^m$. This is a column with m entries and each of the entries is zero. On the other hand, the 0 in the expression $c_1 = c_2 = \dots = c_p = 0$ is the real number zero.

[8] If $v_1,\dots,v_p$ are vectors, then every vector of the form $c_1v_1 + \dots + c_pv_p$, where $c_1,\dots,c_p$ are numbers, is called a <u>linear combination</u> of $v_1,\dots,v_p$.

Probably, you convert the problem into an augmented matrix

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 2 & 4 & 0 \end{array}\right].$$

You apply Elementary Row Operations to put this matrix into Reduced Row Echelon Form (RREF). Then you read the answer. Apply

$$R2 \mapsto R2 - 2R1$$

to obtain

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

This matrix is in (RREF). We read the solution $c_1 = -2c_2$ and c_2 is free to take any value. In particular, if $c_2 = -1$, then $c_1 = 2$ and $2v_1 - 2v_2 = 0$ is a non-trivial linear combination of v_1 and v_2 that is equal to the zero vector.

Question 4.4. *Are the vectors* $v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ *and* $v_2 = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}$ *linearly independent?*

Answer. Yes. When we solve

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

the top row tells us that c_1 must be zero and the second row tells us that c_2 must be zero.

Question 4.5. *Are the vectors* $v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ *and* $v_2 = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}$ $v_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ *linearly independent?*

Answer. No. We see that $1v_1 + 0v_2 - v_3 = 0$ is a non-trivial linear combination on v_1, v_2, v_3 which is equal to zero.

Question 4.6. *Are the vectors* $v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ *and* $v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ *linearly independent?*

Answer. Yes. When we solve

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

the top row tells us that c_1 must be zero; the second row tells us that c_2 must be zero: and the third row tells us that c_3 must be zero.

Question 4.7. *Let a, b, c, d, e, f be numbers. Are the vectors $v_1 = \begin{bmatrix} a \\ b \end{bmatrix}$ and $v_2 = \begin{bmatrix} c \\ d \end{bmatrix}$ $v_3 = \begin{bmatrix} e \\ f \end{bmatrix}$ linearly independent?*

Answer. No. The system of equations $c_1v_2 + c_2v_2 + c_3v_3 = 0$ has at least one solution (namely $c_1 = c_2 = c_3 = 0$). We must determine if it has more than one solution. We solve the system of equations by applying Elementary Row Operations to the augmented matrix

$$\left[\begin{array}{ccc|c} a & c & e & 0 \\ b & d & f & 0 \end{array}\right]$$

Eventually, there will leading ones in some rows. There are two rows; so there will be at most two leading ones. There are three columns before the augmentation; so at least one of the columns before the augmentation will not have a leading one. This column corresponds to a variable that is **free to take any value**. Hence, the system of equations has more than one solution!

Theorem 4.8. *If $m < p$, then any collection of p vectors in $\mathbb{R}^m$ is linearly dependent.*

Proof. Let $v_1, \dots v_p$ be p vectors in $\mathbb{R}^m$. We find all numbers $c_1, \dots, c_p$ in $\mathbb{R}$ with

$$\sum_{i=1}^{p} c_i v_i = 0. \tag{4.8.1}$$

The system of equations (4.8.1) clearly has at least one solution; namely,

$$c_1 = c_2 = \cdots = c_p = 0.$$

We will show that (4.8.1) has more than one solution. We use Gaussian elimination to solve this system of equations. The original augmented matrix has m rows and p columns before the augmentation. The column after the augmentation consists entirely of zero all through the process. When the matrix is in reduced row echelon form, some of the rows have leading ones. Thus,

$$\text{the number of leading ones} \le \text{the number of rows} < \begin{cases} \text{the number of columns} \\ \text{before the augmentation.} \end{cases}$$

Some column before the augmentation does not have a leading one. This column corresponds to a variable that is free to take any value. We conclude that (4.8.1) has more than one solution; hence the vectors $v_1, \dots, v_p$ are linearly dependent. $\square$

Problem 4.9. Suppose v_1, v_2, and v_3 are three vectors in $\mathbb{R}^m$, for some m, with v_1, v_2 linearly independent, v_1, v_3 linearly independent, and v_2, v_3 linearly independent. Do the vectors v_1, v_2, v_3 have to be linearly independent? If the answer is yes, prove it. If the answer is no, give a counterexample.

Answer. Of course not. Take

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{in } \mathbb{R}^2.$$

It is easy to see that the vectors v_1, v_2 are linearly independent; the vectors v_1, v_3 are linearly independent; and the vectors v_2, v_3 are linearly independent. However, the vectors v_1, v_2, v_3 are linearly dependent by Theorem 4.8 or Question 4.7.

Problem 4.10. Suppose v_1, v_2, v_3, and v_4 are linearly independent vectors in $\mathbb{R}^m$, for some m. Do the vectors v_1, v_2, v_3 have to be linearly independent? If the answer is yes, prove it. If the answer is no, give a counterexample.

Answer. Yes, of course the vectors v_1, v_2, v_3 are linearly independent. Suppose c_1, c_2, c_3 are numbers with

$$c_1v_1 + c_2v_2 + c_3v_3 = 0.$$

Then c_1, c_2, c_3 and $c_4 = 0$ are numbers with

$$c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 = 0.$$

The vectors v_1, v_2, v_3, v_4 are linearly independent; hence

$$c_1 = c_2 = c_3 = c_4 = 0.$$

In particular, $c_1 = c_2 = c_3 = 0$ and v_1, v_2, v_3 are linearly independent.

Problem 4.11. Suppose v_1, v_2, and v_3 are linearly independent vectors in $\mathbb{R}^3$. Create the vectors

$$v_1^* = \begin{bmatrix} v_1 \\ 1 \end{bmatrix}, \quad v_2^* = \begin{bmatrix} v_2 \\ 1 \end{bmatrix}, \quad v_3^* = \begin{bmatrix} v_3 \\ 1 \end{bmatrix}$$

in $\mathbb{R}^4$. Do the vectors v_1^*, v_2^*, v_3^* have to be linearly independent? If the answer is yes, prove it. If the answer is no, give a counterexample.

Answer. Yes, v_1^*, v_2^*, v_3^* are linearly independent. Let c_1, c_2, c_3 be numbers with

$$c_1v_1^* + c_2v_2^* + c_3v_3^* = 0.$$

Look at the top three rows to see that

$$c_1v_1 + c_2v_2 + c_3v_3 = 0 \tag{4.11.1}$$

and look at the bottom row to see that $c_1 + c_2 + c_3 = 0$. At any rate, the vectors v_1, v_2, v_3 are linearly independent by hypothesis; hence, equation (4.11.1) yields that $c_1 = c_2 = c_3 = 0$.

Remark 4.12. Almost always, if you want to show that vectors $v_1, \dots, v_p$ are linearly independent, we start by saying "Suppose $c_1, \dots, c_p$ are numbers with

$$c_1v_1 + \cdots + c_pv_p = 0."$$

4.B. Nonsingular matrices.

Definition 4.13. The $n \times n$ matrix A is <u>nonsingular</u> if the only vector x with $Ax = 0$ is the zero vector.

Problems 4.14. (a) Is $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ non-singular?

(b) Is $B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 4 & 1 \end{bmatrix}$ non-singular?

(c) Is $C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ non-singular?

Answers. (a) No, $A \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 0$ and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ is not the zero vector.

(b) No, $B \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = 0$ and $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ is not the zero vector.

(c) Yes. Observe that $Cx = x$ for all x. If $Cx = 0$, then $x = 0$.

4.C. The Nonsingular matrix theorem, version 1.

I planned

- Problem 4.14.(a) to be essentially the same as Example 4.3;
- Problem 4.14.(b) to be essentially the same as Example 4.5; and
- Problem 4.14.(c) to be essentially the same as Example 4.6.

Indeed, if A is an $m \times n$ matrix, $A_{*,i}$ is the i^{th} column of A, and $c_1, \dots, c_n$ are numbers, then

$$A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1 A_{*,1} + \dots c_n A_{*,n}.$$

In the next result, we record the promised connection between "linearly independent" and "nonsingular". We also have a statement about solving systems of linear equations of the form

$$Ax = b,$$

where A is a nonsingular matrix.

Theorem 4.15. [The Nonsingular matrix theorem, version 1] *Let A be an $n \times n$ matrix. The following conditions are equivalent.*[9]

(a) *The matrix A is nonsingular.*
(b) *The columns of A are linearly independent.*
(c) *The system of equations $Ax = b$ has a unique solution for all $b \in \mathbb{R}^n$.*

[9] This means that if one of the conditions holds, then all of the conditions hold. If one of the conditions fails, then all of the conditions fail.

Proof. Let $A_{*,i}$ be the i^{th} column of A.

(a)⇒ (b) We assume that A is a nonsingular matrix. We want to show that the vectors $A_{*,1},\dots,A_{*,n}$ are linearly independent. We remember Remark 4.12 and we say, "Suppose $c_1,\dots,c_n$ are numbers with $c_1A_{*,1}+\cdots+c_nA_{*,n}=0$." Then

$$A\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}=0.$$

The matrix A is non-singular, hence

$$\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$$

is the zero vector and $c_1=\cdots=c_n=0$. We conclude that the vectors $A_{*,1},\dots,A_{*,n}$ are linearly independent.

(b)⇒ (a) We assume that the vectors $A_{*,1},\dots,A_{*,n}$ are linearly independent. We want to prove that the matrix A is non-singular. Suppose

$$\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$$

is a vector in $\mathbb{R}^n$ with

$$A\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}=0.$$

It follows that

$$c_1A_{*,1}+\cdots+c_nA_{*,n}=0.$$

But the vectors $A_{*,1},\dots,A_{*,n}$ are linearly independent by hypothesis; hence

$$c_1=\cdots=c_n=0.$$

Thus

$$\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$$

must be the zero vector and A is non-singular.

(c)⇒ (a) We are told that $Ax=b$ has a unique solution for all $b\in\mathbb{R}^n$. Well, 0 is a vector in $\mathbb{R}^n$. So $Ax=0$ has a unique solution. On the other hand $A0=0$. Thus, 0 must be the unique solution of $Ax=0$. We conclude that if $Ax=0$, then $x=0$. Thus, A is a nonsingular matrix.

(a)⇒ (c) Assume A is a nonsingular matrix $n\times n$ matrix. Let b be an arbitrary element of $\mathbb{R}^n$. We must prove that there exists a unique matrix $v\in\mathbb{R}^n$ with $Av=b$.

(Notice that we have two jobs. We must prove that v exists and we must prove that v is unique.)

Uniqueness. Suppose v_1 and v_2 both are vectors in $\mathbb{R}^n$ with $Av_1 = b$ and $Av_2 = b$. Then

$$\begin{aligned} A(v_1 - v_2) &= Av_1 - Av_2, \quad \text{because matrix multiplication distributes over addition} \\ &= b - b \\ &= 0 \end{aligned}$$

But the matrix A is nonsingular; hence $v_1 - v_2$ must be the zero vector and $v_1 = v_2$.

Existence. Consider the $n+1$ vectors

$$A_{*,1},\ \dots,\ A_{*,n},\ b \tag{4.15.1}$$

in $\mathbb{R}^n$, where $A_{*,i}$ represents column i of A. Apply Theorem 4.8 to see that the vectors of (4.15.1) are linearly dependent. Thus, there are numbers $c_1, \dots, c_{n+1}$,

$$\text{with some } c_i \text{ not zero,} \tag{4.15.2}$$

and

$$c_1 A_{*,1} + \cdots + c_n A_{*,n} + c_{n+1} b = 0. \tag{4.15.3}$$

We claim that c_{n+1} is not zero. (We will do a quick argument by contradiction to establish this claim.) Indeed, if $c_{n+1} = 0$, then

$$c_1 A_{*,1} + \cdots + c_n A_{*,n} = 0 \tag{4.15.4}$$

from (4.15.3). The matrix A is nonsingular and we already proved that (a)$\Rightarrow$ (b); so, the columns of A are linearly independent. It follows from (4.15.4) that

$$c_1 = \cdots = c_n = 0.$$

On the other hand, some c_i is not zero by our choice of c_i in (4.15.2). This is a contradiction. Thus, c_{n+1} is not zero.

Now that we have proven that c_{n+1} is not zero, we rewrite (4.15.3) as

$$\tfrac{-c_1}{c_{n+1}} A_{*,1} + \cdots + \tfrac{-c_n}{c_{n+1}} A_{*,n} = b.$$

Thus

$$A \begin{bmatrix} \frac{-c_1}{c_{n+1}} \\ \vdots \\ \frac{-c_n}{c_{n+1}} \end{bmatrix} = b$$

and we have proven that there does exist a vector v with $Av = b$. $\square$

5. Matrix inverses.

Definition 5.1. The identity matrix I is a square matrix with 1's on the main diagonal[10] and zeros elsewhere.

Fact 5.2. *Notice that if I is an identity matrix and M is a matrix, then $IM = M$ and $MI = M$, whenever the products are defined.*

Examples 5.3.

(a) The 2×2 identity matrix is
$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

(b) The 3×3 identity matrix is
$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(c)
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix},$$
$$\begin{bmatrix} 5 & 6 \\ 7 & 8 \\ 9 & 10 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 7 & 8 \\ 9 & 10 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \\ 9 & 10 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 7 & 8 \\ 9 & 10 \end{bmatrix}.$$

Definition 5.4. If A is an $n \times n$ matrix, then the $n \times n$ matrix B is called an inverse of A if
$$AB = I \quad \text{and} \quad BA = I.$$
One usually writes A^{-1} to denote an inverse of A.

Example 5.5. If $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, then $B = \frac{-1}{2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$ is an inverse of A because
$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ \frac{3}{2} & \frac{-1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
and
$$BA = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & \frac{-1}{2} \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Example 5.6. The matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ has no inverse because if $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is the inverse of A, then
$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

[10] The main diagonal is the diagonal from the upper left corner to the lower right corner.

In particular,

$$1a+2c=1$$
$$2a+4c=0$$

and this system of equations has no solution.

Definition 5.7. The $n\times n$ matrix A is called invertible if A has an inverse.

Possibly Examples 5.5 and 5.6 give you the idea that there is a connection between the concepts "nonsingular" and "invertible".

Theorem 5.8. [The Nonsingular matrix theorem, version 2] *Let A be an $n\times n$ matrix. The following conditions are equivalent.*

(a) *The matrix A is nonsingular.*
(b) *The columns of A are linearly independent.*
(c) *The system of equations $Ax=b$ has a unique solution for all $b\in R^n$.*
(d) *The matrix A is invertible.*

Proof. We already saw that conditions (a), (b), and (c) are equivalent. Now we show that these three conditions are equivalent to (d).

(d)$\Rightarrow$ (a) Assume that A is invertible. We want to prove A is nonsingular. Suppose v is a vector ad $Av=0$. Multiply both sides of the equation on the left by A^{-1} to obtain

$$\begin{aligned} A^{-1}Av &= A^{-1}0; && \text{hence}\\ Iv &= 0; && \text{hence}\\ v &= 0. \end{aligned}$$

We conclude that if $Av=0$, then $v=0$. In other words, we conclude that A is nonsingular.

(a)$\Rightarrow$ (d) Assume A is a nonsingular $n\times n$ matrix. We prove that A is an invertible matrix.

We saw in Theorem 4.15 that if A is a nonsingular $n\times n$ matrix then the system of equations $Ax=b$ has a unique solution, for every b in $\mathbb{R}^n$. Thus, for each i with $1\le i\le n$, there is a (unique) vector v_i with Av_i equal to the i^{th} column of the identity matrix. Let B be the matrix whose columns are $v_1,v_2,\dots,v_n$ (in that order). We have chosen B to have the property that

$$AB=I. \tag{5.8.1}$$

We still must show that $BA=I$. Multiply both sides of equation (5.8.1) on the right by A to obtain

$$ABA=A.$$

It follows that

$$A(BA-I)=0.$$

The matrix A is nonsingular by hypothesis. The matrix A times the i^{th} column of $BA-I$ is zero. Thus, the i^{th} column of $BA-I$ is zero for all i. Thus, $BA-I=0$ and the proof is complete. $\square$

Problems 5.9. Here are some easy questions about inverses.

(a) Suppose that A is an invertible matrix. Is the inverse of A unique?
(b) Suppose that A is an invertible matrix. Is A^{-1} invertible?
(c) Suppose A and B are invertible matrices of the same shape. Is AB invertible?
(d) Suppose that A is an invertible matrix. Is A^{T} invertible?
(e) How can one tell if the matrix
$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$
is invertible?

Answers.
(a) Yes. Suppose B and C both are inverses of A. Then

$$\begin{aligned} C &= IC && \text{because, this is the property of the identity matrix,} \\ &= (BA)C, && \text{because } B \text{ is an inverse of } A \\ &= B(AC), && \text{because matrix multiplication associates} \\ &= BI, && \text{because } C \text{ is an inverse of } A \\ &= B, && \text{because, this is the property of the identity matrix.} \end{aligned}$$

(b) Yes. If A^{-1} is the inverse of A then A and A^{-1} are square matrices of the same size, $A^{-1}A=I$ and $AA^{-1}=I$. These three statements are exactly what it takes to show that A is the inverse of A^{-1} is A.

(c) Yes. The inverse of AB is $B^{-1}A^{-1}$ because
$$(AB)(B^{-1}A^{-1})=I \quad \text{and} \quad (B^{-1}A^{-1})(AB)=I.$$

(d) Yes. The inverse of A^{T} is the transpose of A^{-1}. We use Observation 3.5 twice:
$$(A^{\mathrm{T}})(A^{-1})^{\mathrm{T}}=(A^{-1}A)^{\mathrm{T}}=I^{\mathrm{T}}=I$$
and
$$(A^{-1})^{\mathrm{T}}(A^{\mathrm{T}})=(AA^{-1})^{\mathrm{T}}=I^{\mathrm{T}}=I.$$

(e)[11] The matrix $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if $ad-bc\neq 0$.

[11]The expression $ad-bc$ is the determinant of A. It is true for all square matrices A that A is invertible if and only if the determinant of A is non-zero. We are not ready to prove the general statement, but we can easily prove the result for 2×2 matrices. The determinant of a matrix determines if the matrix is invertible.

($\Leftarrow$) Assume $ad - bc \neq 0$. We prove that A is invertible. Indeed, one easily check that the inverse of A is
$$\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$
($\Rightarrow$)[12] Assume $ad - bc = 0$. We prove that A is singular.[13]

If $ad - bc = 0$, then
$$A\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$
If any of the four numbers a, b, c, or d are non-zero, then we have exhibited a non-zero vector which is sent to zero by A; hence A is a singular matrix. On the other hand, if $a = b = c = d = 0$, then A is the zero matrix and every vector is sent to zero by A; hence A is singular in this case also.

Question 5.10. *In general, how does one find the inverse of a matrix?*

Answer. Apply the technique which was used in the proof of
Theorem 5.8 (a) $\Rightarrow$ (d).

Example 5.11. Let
$$A = \begin{bmatrix} 2 & 1 & 1 \\ 5 & 2 & 1 \\ 3 & 1 & 1 \end{bmatrix}.$$
Find the inverse of A. The inverse of A is
$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$
where
$$A\begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad A\begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad A\begin{bmatrix} b_{13} \\ b_{23} \\ b_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$
We have to solve three systems of equations, where each system has 3 equations in 3 unknowns. But the 3×3 matrix of coefficients is same for all three systems. We apply Elementary Row Operations to
$$\left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 5 & 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right].$$
If A is invertible, then eventually, we will obtain
$$\left[I \;\mid\; \text{something}\right].$$

[12]Actually, I am proving the contrapositive of ($\Rightarrow$). The contrapositive of $P \implies Q$ is not $Q \implies$ not P. A statement is equivalent to its contrapositive.

[13]If A is a square matrix, then singular if A is not nonsingular. We proved that a square matrix is invertible if and only if it is nonsingular. Hence, a matrix is not invertible if and only it it is singular

In this case “something” is the inverse of A. Replace $R1$ by $(1/2)R1$ and obtain

$$\left[\begin{array}{ccc|ccc} 1 & 1/2 & 1/2 & 1/2 & 0 & 0 \\ 5 & 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right].$$

Replace $R2$ by $R2-5R1$ and $R3$ by $R3-3R1$ to obtain

$$\left[\begin{array}{ccc|ccc} 1 & 1/2 & 1/2 & 1/2 & 0 & 0 \\ 0 & -1/2 & -3/2 & -5/2 & 1 & 0 \\ 0 & -1/2 & -1/2 & -3/2 & 0 & 1 \end{array}\right].$$

Replace $R1$ by $R1+R2$ and replace $R3$ by $R3-R2$ to obtain

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & -2 & 1 & 0 \\ 0 & -1/2 & -3/2 & -5/2 & 1 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right].$$

Replace $R2$ by $-2R2$ to obtain

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & -2 & 1 & 0 \\ 0 & 1 & 3 & 5 & -2 & 0 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right].$$

Replace $R1$ by $R1+R3$ and $R2$ by $R2-3R3$ to obtain

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & 2 & 1 & -3 \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right].$$

It appears that we calculated that

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

Lets check it.

$$AA^{-1} = \begin{bmatrix} 2 & 1 & 1 \\ 5 & 2 & 1 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 2 & 1 & -3 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} -2+2+1 & 1-1 & 2-3+1 \\ -5+4+1 & 2-1 & 5-6+1 \\ -3+2+1 & 1-1 & 3-3+1 \end{bmatrix} = I \checkmark$$

and

$$A^{-1}A = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 1 & -3 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 5 & 2 & 1 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -2+3 & -1+1 & -1+1 \\ 4+5-9 & 2+2-3 & 2+1-3 \\ 2-5+3 & 1-2+1 & 1-1+1 \end{bmatrix} = I \checkmark.$$

We conclude that

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

6. Introduction to vector spaces.

There are two parts to this section:

A. Some random examples of subspaces of $\mathbb{R}^n$ and
B. Some official examples of subspaces of $\mathbb{R}^n$.

6.A. Some random examples of subspaces of $\mathbb{R}^n$.

Definition 6.1. A subset V of $\mathbb{R}^n$ is called a vector space or a subspace of $\mathbb{R}^n$ if

(a) the zero vector is in V,
(b) V is closed under addition,[14]
(c) V is closed under scalar multiplication. [15]

Examples 6.2. Which of the following sets are vector spaces? Prove your answer.

(a) $V = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \,\middle|\, x_3 = 2x_1 + 14x_2 \right\}$

(b) $V = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \,\middle|\, x_3 = 2x_1 + 14 \right\}$

(c) $V = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \,\middle|\, \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$

(d) $V = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \,\middle|\, \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$

(e) $V = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \,\middle|\, x_1 = 0 \right\}$

(f) $V = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \,\middle|\, x_1 = 0 \text{ and } x_2 = 0 \right\}$

(g) $V = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \,\middle|\, x_1 = 0 \text{ or } x_2 = 0 \right\}$

(h) $V = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \,\middle|\, \sin x_1 = 0 \right\}$

Answers.

[14] The set V is closed under addition if $v_1, v_2 \in V$ implies $v_1 + v_2 \in V$.
[15] The set V is closed under scalar multiplication if $v \in V$ and $r \in \mathbb{R}$ implies $rv \in V$.

(a) Yes. The zero vector is in V. If $v_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $v_2 = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ are in V, then $x_3 = 2x_1 + 14x_2$ and $y_3 = 2y_1 + 14y_2$. It follows that

$$x_3 + y_3 = 2(x_1 + y_1) + 14(x_2 + y_2);$$

hence $v_1 + v_2$ is in V. Similarly, $rx_3 = 2(rx_1) + 14(rx_2)$; hence $rv_1 \in V$, for any real number r

(b) No. The zero vector is not in V

(c) Yes. The zero vector is in V. If $v_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $v_2 = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ are in V, then

$$2x_1 + 3x_2 + 4x_3 = 0 \quad \text{and} \quad 5x_1 + 6x_2 + 7x_3 = 0$$

and

$$2y_1 + 3y_2 + 4y_3 = 0 \quad \text{and} \quad 5y_1 + 6y_2 + 7y_3 = 0.$$

It follows that

$$2(x_1+y_1)+3(x_2+y_2)+4(x_3+y_3)=0 \quad \text{and} \quad 5(x_1+y_1)+6(x_2+y_2)+7(x_3+y_3)=0.$$

Thus, $v_1 + v_2$ is in V. Similarly, if r is any real number then

$$2(rx_1) + 3(rx_2) + 4(rx_3) = 0 \quad \text{and} \quad 5(rx_1) + 6(rx_2) + 7(rx_3) = 0;$$

hence rv_1 is also in V.

(d) No. The zero vector is not in V.

(e) Yes. The zero vector is in V. If $v_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $v_2 = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ are in V, then $x_1 = 0$ and $y_1 = 0$. Thus. $v_1 + v_2 = \begin{bmatrix} 0 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix}$ and this is in V. Similarly $rv_1 = \begin{bmatrix} 0 \\ rx_2 \\ rx_3 \end{bmatrix}$ and this is in V.

(f) Yes. The zero vector is in V. If $v_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $v_2 = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ are in V, then $x_1 = 0$, $x_2 = 0$, $y_1 = 0$, and $y_2 = 0$. Thus. $v_1 + v_2 = \begin{bmatrix} 0 \\ 0 \\ x_3 + y_3 \end{bmatrix}$ and this is in V. Similarly $rv_1 = \begin{bmatrix} 0 \\ 0 \\ rx_3 \end{bmatrix}$ and this is in V.

(g) No. This set is not closed under addition $v_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$ and $v_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$ are both in V, but $v_1 + v_2 \notin V$.

(h) No. This set is not closed under scalar multiplication: $v = \begin{bmatrix}\pi\\0\\0\end{bmatrix} \in V$, but $\frac{1}{2}v \notin V$.

6.B. **Some official examples of subspaces of $\mathbb{R}^n$: Span (as a noun), Null space, and Column space.** We have used the expression "linear combination" before; see footnote 8 on page 19. Recall that if $v_1, \dots, v_r$ and v are vectors in $\mathbb{R}^n$, then v is a linear combination of $v_1, \dots, v_r$ if there are real numbers $c_1, \dots, c_r$ with $v = c_1 v_1 + \dots + c_r v_r$.

Definition 6.3. If $v_1, \dots, v_r$ are elements of $\mathbb{R}^n$, then the span of $\{v_1, \dots, v_r\}$ is

$$\{v \in \mathbb{R}^n \mid v \text{ is a linear combination of } v_1, \dots, v_r\}.$$

Observation 6.4. *If $v_1, \dots, v_r$ are elements of $\mathbb{R}^n$, then the span of $\{v_1, \dots, v_r\}$ is a subspace of $\mathbb{R}^n$.*

Proof. Let V equal the span of $\{v_1, \dots, v_r\}$.

- We see that the zero vector is in V because $0 = 0v_1 + \dots + 0v_r$.
- We see that V is closed under addition. If v and w are in V, then
$$v = a_1 v_1 + \dots + a_r v_r \quad \text{and} \quad w = b_1 v_1 + \dots + b_r v_r$$
for some a_i and b_i in $\mathbb{R}$. It follows that
$$v + w = (a_1 + b_1)v_1 + \dots + (a_r + b_r)v_r,$$
which is in V.
- We see that V is closed under scalar multiplication. If $v \in V$ and c is a real number, then $v = a_1 v_1 + \dots + a_r v_r$ for some a_i in $\mathbb{R}$. It follows that $cv = ca_1 v_1 + \dots ca_r v_r$, which is in V.

$\square$

Questions 6.5.

(a) Is the span of $\left\{ \begin{bmatrix}1\\2\end{bmatrix} \right\}$ equal to all of $\mathbb{R}^2$?

(b) Is the span of $\left\{ \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\4\end{bmatrix} \right\}$ equal to all of $\mathbb{R}^2$?

(c) Is the span of $\left\{ \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix} \right\}$ equal to all of $\mathbb{R}^2$?

Answers.

(a) No, because $\begin{bmatrix}1\\0\end{bmatrix}$ is in $\mathbb{R}^2$ but not in the span of $\left\{ \begin{bmatrix}1\\2\end{bmatrix} \right\}$.

(b) No, because $\begin{bmatrix}1\\0\end{bmatrix}$ is in $\mathbb{R}^2$ but not in the span of $\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\}$.

(c) Yes, because $Ax = b$ has a solution for all b in $\mathbb{R}^2$, where $A = \begin{bmatrix}1 & 3\\2 & 4\end{bmatrix}$. To see this I calculate the determinant of A (which is $1\times 4 - 3\times 2 = -2 \neq 0$ and apply Problem 5.9.(e) and Theorem 5.8. (Of course, I could make calculations to reach the same conclusion.)

Definition 6.6. If A is an $m\times n$ matrix, then

$$\{v \in \mathbb{R}^n \mid Av = 0\}$$

is called the null space of A.

Example 6.7. The vector spaces (a), (c), (e) and (f) in Example 6.2 all are null spaces. Example (a) is the null space of $\begin{bmatrix}-2 & -14 & 1\end{bmatrix}$. Example (c) is the null space of $\begin{bmatrix}2 & 3 & 4\\5 & 6 & 7\end{bmatrix}$. Example (e) is the null space of $\begin{bmatrix}1 & 0 & 0\end{bmatrix}$. Example (f) is the null space of $\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\end{bmatrix}$.

Observation 6.8. *If A is a matrix, then the null space of A is a vector space.*

Proof. Let A have m rows and n columns. The zero vector from $\mathbb{R}^n$ is in the null space of A. If v_1 and v_2 are in the null space of A, then

$$A(v_1+v_2) = Av_1 + Av_2 = 0+0 = 0.$$

The first equality holds because matrix multiplication distributes over addition; the second equality holds because v_1 and v_2 are in the null space of A. We have shown that the null space of A is closed under addition. Notice that all three 0's in the most recent display are the zero vector in $\mathbb{R}^m$. If r is a real number, then

$$A(rv_1) = rA(v_1) = r \text{ times the zero vector} = 0.$$

We have shown that the null space of A is closed under scalar multiplication. $\square$

Observation 6.9. *Let A be an $m\times n$ matrix. Consider the following three subsets of $\mathbb{R}^m$:*

(a) *S_1 is the span of the columns of A,*
(b) *S_2 is $\{Av \mid v \in \mathbb{R}^n\}$, and*
(c) *S_3 is $\{b \in \mathbb{R}^m \mid$ there exists $v \in \mathbb{R}^n$ with $Av = b\}$.*

Then the three sets S_1, S_2, and S_3 are exactly equal.

Proof. Let $A_{*1}, \dots A_{*,n}$ be the columns of A.

$S_1 \subseteq S_2$: Let w be an element of S_1. Then w is a linear combination of $A_{*1}, \dots A_{*,n}$. Thus, there are numbers $c_1,\dots,c_n$ with

$$w = c_1A_{*,1} + \cdots + c_nA_{*,n}.$$

The most recent equation can be rewritten as

$$w = Av, \quad \text{where} \quad v = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

This shows that $w \in S_2$.

$S_2 \subseteq S_1$: Let $w \in S_2$. Then $w = Av$ for some $v \in \mathbb{R}^n$. If $v = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$, then $w = Av$ may be rewritten as

$$w = c_1 A_{*,1} + \cdots + c_n A_{*,n}.$$

The most recent equation exhibits w as a linear combination of the columns of A. Thus $w \in S_1$.

$S_2 \subseteq S_3$: Let w be an element of S_2. In this case $w = Av$ for some $v \in \mathbb{R}^n$. Thus, $w \in S_3$.

$S_3 \subseteq S_2$: Take an element b of S_3. According to the definition of S_3, there exists a vector $v \in \mathbb{R}^n$ with $b = Av$. Now we see that $b \in S_2$.

□

Definition 6.10. Let A be an $m \times n$ matrix. Any one of the three sets S_1, S_2, or S_3 of Observation 6.9 is called the column space of A.

Problems 6.11.

(a) What is the null space of $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$?
(b) What is the column space of $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$?
(c) What is the null space of the $n \times n$ identity matrix?
(d) What is the column space of the $n \times n$ identity matrix?
(e) Suppose A is a nonsingular $n \times n$ matrix. What is the null space of A?
(f) Suppose A is a nonsingular $n \times n$ matrix. What is the null space of A?
(g) Consider the assigned Homework problems 23–41, When possible describe the vector space under consideration as a null space or a column space.

Answers.

(a) **What is the null space of $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$?** The null space of $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$ is $\mathbb{R}^3$.
(b) **What is the column space of $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$?** The column space of $\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$ is $\{[0]\}$, which is a subset of $\mathbb{R}^1$.
(c) **What is the null space of the $n \times n$ identity matrix?**
The null space of the $n \times n$ identity matrix consists of the zero vector in $\mathbb{R}^n$.
(d) **What is the column space of the $n \times n$ identity matrix?** The column space of the $n \times n$ identity matrix is all of $\mathbb{R}^n$.
(e) **Suppose A is a nonsingular $n \times n$ matrix. What is the null space of A?** The null space of an $n \times n$ nonsingular matrix consists of the zero vector in $\mathbb{R}^n$.

(f) **Suppose A is a nonsingular $n \times n$ matrix. What is the column space of A?**
The column space of an $n \times n$ nonsingular matrix is all of $\mathbb{R}^n$.

(g) **Consider the assigned Homework problems 23–41, When possible describe the vector space under consideration as a null space or a column space.**

7. THE OFFICIAL DEFINITION OF A VECTOR SPACE AND MANY MORE EXAMPLES.

Definition 7.1. A vector space over $\mathbb{R}$ is a set V together with two operations called addition and scalar multiplication which satisfy the following properties.

(a) (closure under addition) If v_1 and v_2 are in V, then $v_1 + v_2$ is in V.
(b) (addition associates) If v_1, v_2, v_3 are in V, then
$$v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3.$$
(c) (additive identity element) There exists an element 0 in V with $0 + v = v$ for all $v \in V$.
(d) (additive inverses) If v is an element of V, then there is an element $-v$ of V with $v + (-v) = 0$.
(e) (addition commutes) If v_1 and v_2 are in V, then $v_1 + v_2 = v_2 + v_1$.
(f) (closure under scalar multiplication) If $r \in R$ and $v \in V$, then $rv \in V$.
(g) If $r \in R$ and v_1 and v_2 are in V, then
$$r(v_1 + v_2) = rv_1 + rv_2.$$
(h) If r_1 and r_2 are in $\mathbb{R}$ and $v \in V$, then
$$(r_1 + r_2)v = r_1 v + r_2 v.$$
(i) If r_1 and r_2 are in $\mathbb{R}$ and $v \in V$, then
$$(r_1 r_2)v = r_1(r_2 v).$$
(j) If 1 is the real number 1 and $v \in V$, then $1v = v$.

Remark 7.2. If $(V, +)$ is a set together with one operation which satisfies properties (a)-(e), then V is called an Abelian group. (Groups are the main object studied in Math 546.)

Remark 7.3. The rules do not state that the zero element of $\mathbb{R}$ times an arbitrary element of a vector space V is equal to the zero element of V. However, this statement follows quickly from the axioms. Indeed, if V is a vector space, $0_{\mathbb{R}}$ is the zero element of R, 0_V is the zero element of V, and $v \in V$, then
$$0_{\mathbb{R}} v = (0_{\mathbb{R}} + 0_{\mathbb{R}})v = 0_{\mathbb{R}} v + 0_{\mathbb{R}} v.$$
Add $-(0_{\mathbb{R}} v)$ to both sides; associate; use the property of additive inverse; and use the property of additive identity. Conclude that
$$0_V = 0_{\mathbb{R}} v.$$

Examples 7.4. (a) If n is a positive integer, then $\mathbb{R}^n$ is a vector space over $\mathbb{R}$,
(b) If m and n are positive integers, then the set of $m \times n$ matrices is a vector space over $\mathbb{R}$, denoted $\mathrm{Mat}_{m\times n}(\mathbb{R})$. In particular, the set of row vectors with n entries is a vector space for each positive integer n.

(c) If $a < b$ are real numbers, then the set of continuous functions from the closed interval $[a, b]$ to $\mathbb{R}$ is a vector space over $\mathbb{R}$, denoted $\mathscr{C}[a, b]$.
(d) The set of polynomials in one variable over $\mathbb{R}$ is a vector space over $\mathbb{R}$, denoted $\mathbb{R}[x]$.
(e) The set of polynomials of degree at most n in one variable over $\mathbb{R}$ is a vector space over $\mathbb{R}$, denoted $\mathscr{P}_n$.

Definition 7.5. If $W \subseteq V$ are vector spaces with the same addition and scalar multiplication, then W is a <u>subspace</u> of V.

Observation 7.6. *Let V be a vector space. If W is a non-empty subset of V which is closed under addition and scalar multiplication, then W is a subspace of V.*

Proof. All of the other axioms of vector space hold automatically in W because they hold in V.

In particular, The zero element of V is automatically an element of W. Indeed, we are promised that W has at least one element. Let w be an element of W. We are also promised that W is closed under scalar multiplication; so

$$0_{\mathbb{R}} w \in W.$$

Apply Remark 7.3 to conclude that $0_V \in W$. □

Example 7.7. If A is an $m \times n$ matrix, then the set of row vectors which can be written as a linear combination of the rows of A is a vector space, which we call the row space of A. Notice that if the matrix B is obtained from the matrix A by applying a sequence of Elementary Row Operations, then the matrices A and B have the same row space.

Examples 7.8. Consider the following subsets of $\mathscr{P}_3$. Which are subspaces of $\mathscr{P}_3$?
(a) $V_1 = \{f(x) \in \mathscr{P}_3 \mid f(0) = 0\}$,
(b) $V_2 = \{f(x) \in \mathscr{P}_3 \mid f(1) = 0\}$,
(c) $V_3 = \{f(x) \in \mathscr{P}_3 \mid f(0) = 1\}$,
(d) $V_4 = \{f(x) \in \mathscr{P}_3 \mid f(1) = 1\}$,
(e) $V_5 = \{f(x) \in \mathscr{P}_3 \mid f(1) = 0 \text{ and } f'(1) = 0\}$.

Answers. It is easy to see that V_1, V_2, and V_5 all are non-empty and closed under addition and scalar multiplication. Thus, V_1, V_2, and V_5 are subspaces of $\mathscr{P}_3$, On the other hand, the zero polynomial is not in V_3 or V_4; so neither of these sets is a subspace of $\mathscr{P}_3$. (We are using the proof of Observation 7.6. If W is a subspace of V, then the zero element of V must be in W.)

Examples 7.9. Consider the following subsets of $\mathrm{Mat}_{2\times 2}(\mathbb{R})$. Which are subspaces of $\mathrm{Mat}_{2\times 2}(\mathbb{R})$?
(a) $V_1 = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| \, a = 0 \right\}$,

(b) $V_2 = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| \, a+b=0 \right\}$,

(c) $V_3 = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| \, ab=0 \right\}$,

(d) $V_4 = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| \, a=1 \right\}$,

(e) $V_5 = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| \, b=c \right\}$.

Answers. It is easy to see that V_1, V_2, and V_5 all are non-empty and closed under addition and scalar multiplication. Thus, V_1, V_2, and V_5 are subspaces of $\mathrm{Mat}_{2\times 2}(\mathbb{R})$. The zero matrix is not in V_4, hence is not a subspace of $\mathrm{Mat}_{2\times 2}(\mathbb{R})$. (Once again, we use the proof of Observation 7.6. If W is a subspace of V, then the zero element of V must be in W.) The set V_3 is not closed under addition. For example, $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ are both in V_3, but $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not in V_3. Thus V_3 is not a vector space.

8. Bases for vector spaces

Definition 8.1. Let $v_1, \dots, v_p$ be elements of the vector space V.

(a) [16] The vectors $v_1, \dots, v_p$ are <u>linearly independent</u> if the only numbers $c_1, \dots, c_p$ with $\sum_{i=1}^{p} c_i v_i = 0$ are $c_1 = \cdots = c_p = 0$.

(b) [17] The vectors $v_1, \dots, v_p$ <u>span</u> V if every element of V is equal to a linear combination of $v_1, \dots, v_p$.

(c) The vectors $v_1, \dots, v_p$ are a <u>basis</u> for V if $v_1, \dots, v_p$ span V and are linearly independent.

Examples 8.2. (a) The vectors

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

form a basis for $\mathbb{R}^3$.

Indeed, the vectors v_1, v_2, v_3 are linearly independent. If c_1, c_2, c_3 are numbers with

$$c_1 v_1 + c_2 v_2 + c_3 v_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

then

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and $c_1 = 0$, $c_2 = 0$, and $c_3 = 0$.

Also, the vectors v_1, v_2, v_3 span $\mathbb{R}^3$. If

$$v = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

is an arbitrary element of $\mathbb{R}^3$, then v can be written as a linear combination of v_1, v_2, and v_3. Indeed,

$$v = av_1 + bv_2 + cv_3.$$

(b) The vectors

$$v_1 = \begin{bmatrix} 1 \\ 6 \\ 94 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 1 \\ 122 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

form a basis for $\mathbb{R}^3$.

[16] The concept of linearly independent vectors was introduced in Definition 4.1. We have changed nothing. We emphasize that the notion of linear independence makes sense in every vector space.

[17] We defined "span" as a noun in Definition 6.3. Here we define "span" as a verb.

Indeed, the vectors v_1, v_2, v_3 are linearly independent. If c_1, c_2, c_3 are numbers with

$$c_1 v_1 + c_2 v_2 + c_3 v_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

then

$$\begin{bmatrix} c_1 \\ 6c_1 + c_2 \\ 94c_1 + 122c_2 + c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and $c_1 = 0$ from the top row, $c_2 = 0$ from the second row, and $c_3 = 0$ from the third row.

Also, the vectors v_1, v_2, v_3 span $\mathbb{R}^3$. If

$$v = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

is an arbitrary element of $\mathbb{R}^3$, then v can be written as a linear combination of v_1, v_2, and v_3. Indeed,

$$v = av_1 + (b - 6a)v_2 + (c - 94a - 122(b - 6a))v_3.$$

(c) The vectors

$$v_1 = \begin{bmatrix} 2 \\ 5 \\ 3 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

form a basis for $\mathbb{R}^3$.

For this one I appeal to Example 5.11, where we proved that the matrix whose columns are v_1, v_2, v_3 is an invertible matrix. The columns of an invertible matrix are linearly independent (by, say the Nonsingular Matrix Theorem, Theorem 5.8). Thus, v_1, v_2, v_3 are linearly independent.

The vectors v_1, v_2, v_3 span $\mathbb{R}^3$. Indeed, if v is an arbitrary vector in $\mathbb{R}^3$, then

$$\begin{bmatrix} v_1 | v_2 | v_3 \end{bmatrix} x = v$$

has a (unique) solution by the Nonsingular Matrix Theorem (again). In other words, v can be written as a linear combination of v_1, v_2, v_3.

(d) The polynomials $1, x, x^2, x^3$ are a basis for $\mathscr{P}_3$.

These polynomials span $\mathscr{P}_3$ because every element of $\mathscr{P}_3$ is equal to $r_0 + r_1 x + r_2 x^2 + r_3 x^3$, for some $r_i \in \mathbb{R}$.

These polynomials are linearly independent because, if $p(x) = r_0 + r_1 x + r_2 x^2 + r_3 x^3$ is the zero polynomial, then all of the coefficients are zero.

(e) The polynomials $1, x-1, (x-1)^2, (x-1)^3$ are a basis for $\mathscr{P}_3$.

These polynomials span $\mathscr{P}_3$ because if $p(x)$ is an arbitrary element of $\mathscr{P}_3$, then

$$p(x) = p(1) + p'(1)(x-1) + \frac{p''(1)}{2}(x-1)^2 + \frac{p'''(1)}{3!}(x-1)^3.$$

(I have recorded the Taylor's series for $p(x)$ about $x = 1$. If you don't trust this argument you can create a grubbier argument that expresses an arbitrary element $p(x)$ from $\mathscr{P}_3$ in terms of 1, $x-1$, $(x-1)^2$, and $(x-1)^3$. Once you get your answer. Step back and look at it. You will see that you have produced the answer given by Taylor's Theorem.)

The polynomials $1, x-1, (x-1)^2, (x-1)^3$ are linearly independent. Suppose c_0, c_1, c_2, c_3 are real numbers with

$$c_0 + c_1(x-1) + c_2(x-1)^2 + c_3(x-1)^3 \tag{8.2.1}$$

equal to the zero polynomial. The coefficient of x^3 in the zero polynomial is zero. So the coefficient of x^3 in (8.2.1). The coefficient of x^3 in (8.2.1) is c_3. Thus, c_3 must be zero.

The coefficient of x^2 in the zero polynomial is zero. The coefficient of x^2 in (8.2.1) (now that $c_3 = 0$) is c_2. Thus, c_2 must also be zero. etc.

(f) The matrices $\begin{bmatrix}1&0\\0&0\end{bmatrix}, \begin{bmatrix}0&1\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\1&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix}$ is a basis for $\mathrm{Mat}_{2\times 2}(\mathbb{R})$.

8.A. **How to find bases for the row space, the null space, and the column space of a matrix.**

Observation 8.3. *Let A be a matrix. Apply Elementary Row Operations to A in order to obtain a matrix B which is in Reduced Row Echelon Form. Then the following statements hold.*

(a) *The non-zero rows of B are a basis for the row space of A.*
(b) *The matrices A and B have the same null space and it is easy to read a basis for the null space of B.*
(c) *The columns in A which correspond to the leading ones in B form a basis for the column space of A.*

Example 8.4. Let $A = \begin{bmatrix}1&2&3\\2&4&6\end{bmatrix}$. Find bases for the row space, null space, and column space of A. Replace R2 by R2-2R1 to obtain $B = \begin{bmatrix}1&2&3\\0&0&0\end{bmatrix}$. Observe that B is in Reduced Row Echelon Form.

The matrix B has one non-zero row. We conclude that

$$\boxed{\text{the vector } \begin{bmatrix}1&2&3\end{bmatrix} \text{ is a basis for the row space of } A.}$$

We read the null space of A from B. When one writes $Bx = 0$ one obtains the equation $x_1 + 2x_2 + 3x_3 = 0$, which means x_2 and x_3 are free to take any value and

$$x_1 = -2x_2 - 3x_3.$$

In other words,

$$\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$$

is in the null space of A if and only if

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

We see that

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

span the null space of A. It is also clear that these two vectors are linearly independent.

$$\boxed{\text{The vectors } \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \text{ form a basis for the null space of } A.}$$

The matrix B has one leading one. This leading one is in column 1. Thus, column 1 of A is a basis for the column space of A.

$$\boxed{\text{The vector } \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ is a basis for the column space of A.}}$$

Proof. (a) The row space of A is equal to the row space of B. It suffices to show that the non-zero rows of B form a basis for the row space B. It is clear that the non-zero rows of B span the row space of B. Each non-zero row of B contains a leading one. The entries in B above and below the leading one all are zero. Thus, the non-zero rows of B are linearly independent.

(b) It is clear that if the matrix M' is obtained from the matrix M by way of an elementary row operation, then M' and M have the same null space. (To see this, consider the elementary row operations one at a time. If I exchange two rows of M, then the new matrix has the same null space as the old matrix. If I multiply one row of M by a non-zero constant, then the new matrix has the same null space as the old matrix. If I add a non-zero multiple of one row to another row then the two matrices have the same null space.) Thus, A and B have the same null space. It is easy to read a basis of the null space from B; but that is not a statement that is provable. (It is a value judgment.) There is nothing to prove.

(c) It is fairly easy to make sense of (c); but fairly cumbersome to write down a proof. The point is that if x_i is a free variable, then there is an element in the null space of A which involves one copy of column i of A and the columns of A which correspond to leading variables. Thus, column i is in the span of the columns of A which correspond to leading variables. Make this observation for each column of A which corresponds to a free variable and conclude that the column space of A is spanned by the columns of A which corrpespond to leadings. Observe further, that the only element in the null space of A which has zero in the rows corresponding to leading ones is the element zero. $\square$

Example 8.5. Let

$$A = \begin{bmatrix} 1 & 4 & 5 & 1 & 8 \\ 1 & 4 & 5 & 2 & 10 \\ 3 & 12 & 15 & 4 & 26 \end{bmatrix}.$$

Find a basis for the row space, the null space, and the column space of A. Express each row of A in terms of the proposed basis for the row space of A. Express each column of A in terms of the proposed basis for the column space of A.

Apply elementary row operations to A in order to obtain the matrix

$$B = \begin{bmatrix} 1 & 4 & 5 & 0 & 6 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

which is in row echelon form.

Here are the row operations. Start with

$$\begin{bmatrix} 1 & 4 & 5 & 1 & 8 \\ 1 & 4 & 5 & 2 & 10 \\ 3 & 12 & 15 & 4 & 26 \end{bmatrix}.$$

Replace Row 2 with Row 2 minus Row 1. Replace Row 3 with Row 3 minus 3 Row 1 to get

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

Replace Row 1 with Row 1 minus Row 2 and replace Row 3 with Row 3 minus Row 2 to get

$$\begin{bmatrix} 1 & 4 & 5 & 0 & 6 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The matrices A and B have the same null space. We see from B that the null space of A is equal to

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} \middle| \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -5 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -6 \\ 0 \\ 0 \\ -2 \\ 1 \end{bmatrix}, \text{ where } x_2, x_3, \text{ and } x_5 \text{ are arbitrary elements in } \mathbb{R} \right\}.$$

(You are welcome to put an intermediate step here if you like.)

In particular,

$$\boxed{v_1 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} -5 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} -6 \\ 0 \\ 0 \\ -2 \\ 1 \end{bmatrix}}$$

is a basis for the null space of A. Furthermore, the fact that v_1 is in the null space of A tells us that $\boxed{A_{*2} = 4A_{*1}}$; the fact that v_2 is in the null space of A tells us that $\boxed{A_{*3} = 5A_{*1}}$; and the fact that v_3 is in the null space of A tells us that $\boxed{A_{*5} = 6A_{*1} + 2A_{*4}}$. We conclude that the column space of A is spanned by $\{A_{*1}, A_{*4}\}$. We wonder if the vectors A_{*1}, A_{*4} are linearly independent. That is the same as wondering if there are any non-zero vectors in the null space of A of the form

$$\begin{bmatrix} * \\ 0 \\ 0 \\ * \\ 0 \end{bmatrix}?$$

The answer is no. Every element in the null space of A has the form

$$c_1v_1 + c_2v_2 + c_3v_3 = \begin{bmatrix} -4c_1 - 5c_2 - 6c_3 \\ c_1 \\ c_2 \\ -2c_3 \\ c_3 \end{bmatrix}$$

for some numbers c_1, c_2, c_3. Of course, when we compare rows 2, 3, and 5 of

$$\begin{bmatrix} * \\ 0 \\ 0 \\ * \\ 0 \end{bmatrix} = \begin{bmatrix} -4c_1 - 5c_2 - 6c_3 \\ c_1 \\ c_2 \\ -2c_3 \\ c_3 \end{bmatrix},$$

we learn that $c_1 = c_2 = c_3 = 0$, hence both $*$'s are zero and A_{*1}, A_{*4} are linearly independent. We conclude that

$$\boxed{A_{*1}, A_{*4} \text{ is a basis for the column space of } A.}$$

The matrices A and B have the same row space. The vectors

$$\boxed{w_1 = \begin{bmatrix} 1 & 4 & 5 & 0 & 6 \end{bmatrix} \quad \text{and} \quad w_2 = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \end{bmatrix}}$$

form a basis for the row space of A. We see that

$$\boxed{A_{1*} = w_1 + w_2, \quad A_{2*} = w_1 + 2w_2, \quad \text{and} \quad A_{3*} = 3w_1 + 4w_2.}$$

9. Vector space dimension.

Definition 9.1. If the vector space V has a finite basis, then V is a finite dimensional vector space.

In this section we prove four theorems about finite dimensional vector spaces. (These theorems continue to hold for arbitrary vector spaces provided one uses the notion of "cardinality" in place of the notion of the "number of elements". Significantly more care is required in the general case.)

Theorem 9.2. *Every basis for the finite dimensional vector space V has the same number of vectors.*

We use the following Lemma twice to prove Theorem 9.2.

Lemma 9.3. *If $v_1, \dots, v_q$ span the vector space V and $u_1, \dots, u_p$ are elements of V with $q < p$, then $u_1, \dots, u_p$ are linearly dependent.*

The Lemma reminds us of the short/fat theorem (Theorem 4.8) which states that if $q < p$, then every collection of p vectors in $\mathbb{R}^q$ is linearly dependent. Indeed, we will prove the Lemma by maneuvering the given data into the hypotheses of Theorem 4.8. The vector space V might not be $\mathbb{R}^q$; but there are column vectors with q entries that are very relevant.

Proof. Write $u_j = \sum_{i=1}^q a_{ij} v_i$, with $a_{ij} \in \mathbb{R}$. The columns of

$$A = \begin{bmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{q1} & \dots & a_{qp} \end{bmatrix}$$

are p vectors in $\mathbb{R}^q$ with $q < p$. Theorem 4.8 guarantees that there exists a nonzero vector $c = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$ in $\mathbb{R}^p$ with $Ac = 0$. Observe that $\sum\limits_{j=1}^{p} c_j u_j = 0$. Indeed,

$$\sum_{j=1}^{p} c_j u_j = \sum_{j=1}^{p} c_j \left(\sum_{i=1}^{q} a_{ij} v_i \right) = \sum_{i=1}^{q} \left(\underbrace{\sum_{j=1}^{p} c_j a_{ij}}_{0} \right) v_i = 0.$$

We have identified numbers $c_1, \dots, c_p$, not all of which are zero, with $\sum\limits_{j=1}^{p} c_j u_j = 0$. Thus, we have shown that $u_1, \dots, u_p$ are linearly dependent. □

Proof of Theorem 9.2. The hypothesis guarantees that some basis of V is finite. Thus, Lemma 9.2 guarantees that every basis of V is finite. Let $v_1, \dots, v_n$ and $w_1, \dots, w_m$ be bases for V.

Apply Lemma 9.3 to the spanning set $v_1, \dots, v_n$ and the linearly independent set $w_1, \dots, w_m$ to see that $m \le n$. Now apply Lemma 9.3 to the spanning set $w_1, \dots, w_m$ and the linearly independent set $v_1, \dots, v_n$ to see that $m \le n$. Conclude that $m = n$. □

Definition 9.4. If V is a finite dimensional vector space, then the number of elements in a basis for V is called the <u>dimension</u> of V.

Theorem 9.5. *If V is a finite dimensional vector space, then every linearly independent subset in V is part of a basis for V.*

Proof. Let $d = \dim V$. By hypothesis, there exists a basis of V with d elements. Apply Lemma 9.3 to see that every subset of V with $d+1$ elements is linearly dependent. Conclude that

$$\begin{cases} \text{every linearly independent subset of } V \text{ with } d \text{ elements} \\ \text{is already a basis for } V. \end{cases} \tag{9.5.1}$$

Let $v_1, \dots, v_i$ be a linearly independent subset of V. If $v_1, \dots, v_i$ span V, then $v_1, \dots, v_i$ is a basis for V and the proof is complete. If $v_1, \dots, v_i$ do not span V, then there is an element v_{i+1} in V but not in the span of $v_1, \dots, v_i$. Observe $v_1, \dots, v_{i+1}$ is a linearly independent subset of V. Proceed in this manner. According to (9.5.1), the process stops once your list of linearly independent vectors contains d vectors. □

Theorem 9.6. *If V is a finite dimensional vector space, then every subset of V which spans V contains a basis for V.*

Proof. Let S be a set of elements of V which spans V.

We first show that some finite subset of S also spans V. Let $d = \dim V$. By hypothesis, there exists a basis of V with d elements. Each element of this basis is a finite linear combination of elements of S. A finite union of finite sets is finite. We have identified a finite subset S' of S that spans V.

The rest of the proof is easy. We throw elements of S' away, one at a time, until we have a basis for V. If S' is linearly independent, then S' is a basis for V. If S' is linearly dependent, then an element s_1 of S' can be written in terms of the other elements. Throw this away. We still have that $S' \setminus \{s_1\}$ spans V.

We continue in this manner until we obtain a subset S'' of S which spans V and has d elements. No proper subset of S'' can span V by Lemma 9.3 (because we have a linearly independent subset of V with d elements). Thus, S'' must both span V and be linearly independent. We have produced a basis for V which is a subset of the original set S. □

Theorem 9.7. *If M is a matrix, then the dimension of the column space of M plus the dimension of the null space of M is equal to the number of columns of M.*

Remarks. (a) This theorem is usually called the “Rank-Nullity Theorem” because the dimension of the column space of M is called the rank of M and the dimension of the null space of M is called the nullity of M.

(b) I think of this result as a conservation result (as in conservation of energy). One must account for all of the vector space dimension. Here is what I mean. Suppose M is an $m \times n$ matrix. Then multiplication by M is a function $\mathbb{R}^n \to \mathbb{R}^m$. The image of this map is the column space of M which is a subspace of $\mathbb{R}^m$. The stuff that was burned (that is sent to zero) is the null space of M, which is a subspace of $\mathbb{R}^n$. The theorem says that the dimension of the domain (that is n) is equal to the dimension of the image (that is $\operatorname{rank} M$) plus the dimension of the stuff that was burned (that is the nullity of M).

Proof. Let M be an $m \times n$ matrix. Let $v_1, \dots, v_s$ in $\mathbb{R}^n$ be a basis for the null space of M and let $w_1, \dots, w_t$ in $\mathbb{R}^n$ be vectors with $Mw_1, \dots, Mw_t$ are a basis for the column space of M.

We prove that $v_1, \dots v_s, w_1, \dots, w_t$ is a basis for $\mathbb{R}^n$.

First we show that the vectors $v_1, \dots, v_s, w_1 \dots w_t$ are linearly independent. Suppose

$$\sum_{i=1}^{s} a_i v_i + \sum_{j=1}^{t} b_j w_j = 0 \tag{9.7.1}$$

for numbers $a_1, \dots, a_t, b_1, \dots, b_s$. Multiply by M to see that $\sum_{j=1}^{t} b_j M w_j = 0$. The vectors $Mw_1, \dots, Mw_t$ are linearly independent; hence, each b_j is equal to zero. and (9.7.1) becomes

$$\sum_{i=1}^{s} a_i v_i = 0.$$

The vectors $v_1, \dots, v_s$ are linearly independent; thus, each a_i is also equal to zero.

Now we show that $v_1, \dots, v_s, w_1 \dots w_t$ span $\mathbb{R}^n$. Let u be an arbitrary element of $\mathbb{R}^n$. Observe that Mu is in the vector space spanned by $Mw_1, \dots, Mw_t$. Hence there are numbers $\beta_1 \dots, \beta_t$ with $Mu = \sum_{j=1}^{t} \beta_j M w_j$. Rewrite the last equation to see that $u - \sum_{j=1}^{t} \beta_j w_j$ is in the null space of M. It follows that there exist numbers $\alpha_1, \dots, \alpha_s$ with

$$u - \sum_{j=1}^{t} \beta_j w_j = \sum_{i=1}^{s} \alpha_i v_i.$$

□

Corollary 9.8. *If M is a matrix, then the row space of M and the column space of M have the same dimension.*

Proof. Apply elementary row operations to M to obtain a matrix B in row echelon form. The dimension of the row space of M is equal to the number of nonzero rows in B, which is the same as the number of leading ones in B. The dimension of the null space of M is equal to the number of columns of B without a leading one. Thus, the dimension of the row space of M plus the dimension of the null space of M is equal to the number of columns of M. On the other hand, the rank-nullity theorem shows that the dimension of the column space of M plus the dimension of the null space of M is equal to the number of columns of M. The stated result follows. $\square$

Problems 9.9. Let V be a vector space.

(a) Suppose $\dim V = n$ and $v_1, \ldots, v_n$ are linearly independent vectors in V. What else do you know for sure? Why?

(b) Suppose $\dim V = n$ and $v_1, \ldots, v_n$ span V. What else do you know for sure? Why?

(c) Suppose V_1 is a subspace of V with $V_1 \neq V$. Suppose further that $v_1, \ldots, v_{n-1}$ are elements of V_1 which are linearly independent. What else do you know for sure? Why?

(d) Let V be the vector space

$$V = \{p(x) \in \mathscr{P} | p(1) = 0 \text{ and } p'(1) = 0\}.$$

Find a basis for V. If possible, find your basis with out doing any hard work.

(e) Suppose $U \subseteq V$ are finite dimensional vector spaces and that $u_1, \ldots, u_r$ is a basis for U. Prove that there exist $v_1, \ldots, v_s$ in V so that $u_1, \ldots, u_r, v_1, \ldots, v_s$ is a basis for V.

10. Linear Transformations.

Definition 10.1. A function $T : V \to W$ from the vector space V to the vector space W is called a linear transformation if

(a) $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all v_1 and v_2 in V, and
(b) $T(rv) = rT(v)$ for all $r \in R$ and $v \in V$.

Examples 10.2. (a) If A is an $m \times n$ matrix, then the function $T : \mathbb{R}^n \to \mathbb{R}^m$, which is given by $T(v) = Av$, for all $v \in V$, is a linear transformation.

(b) Let $\mathscr{C}$ be the set of all functions from $\mathbb{R}$ to $\mathbb{R}$, $\mathscr{C}^1$ be the set of all differentiable functions from $\mathbb{R}$ to $\mathbb{R}$, and $T : \mathscr{C}^1 \to \mathscr{C}$ be differentiation. Then T is a linear transformation.

(c) Let $\mathscr{C}$ be the set of continuous function from the closed interval $[0,1]$ to $\mathbb{R}$. Then $T : \mathscr{C} \to \mathbb{R}$ given by $T(f) = \int_a^b f(x)dx$ is a linear transformation.

(d) Let V be a vector space of dimension n and $v_1, \dots, v_n$ be a basis for V. Call this basis $\mathscr{B}$. For each vector $v \in V$, let $[v]_{\mathscr{B}}$ be the column vector

$$\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

in $\mathbb{R}^n$ with $v = \sum_{i=1} c_i v_i$. (The vector $[v]_{\mathscr{B}}$ is called the coordinate vector of v with respect to the basis $\mathscr{B}$ of V.) The function $T : V \to \mathbb{R}^n$ which is given by $T(v) = [v]_{\mathscr{B}}$ is a linear transformation.

(e) Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Let A be the $m \times n$ matrix

$$A = \left[T(e_1) | T(e_2) | \cdots | T(e_n) \right],$$

where e_j is the element of $\mathbb{R}^n$ with 1 in row j and zero everywhere else. Then $T(v) = Av$ for[18] all v in $\mathbb{R}^n$.

(f) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the function which fixes the origin and rotates the xy-plane by ϕ radians counter-clock-wise. Find a matrix A with $Tv = AV$ for all $v \in \mathbb{R}^2$. Conclude that T is a linear transformation.

To do this problem we use polar coordinates

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix}$$

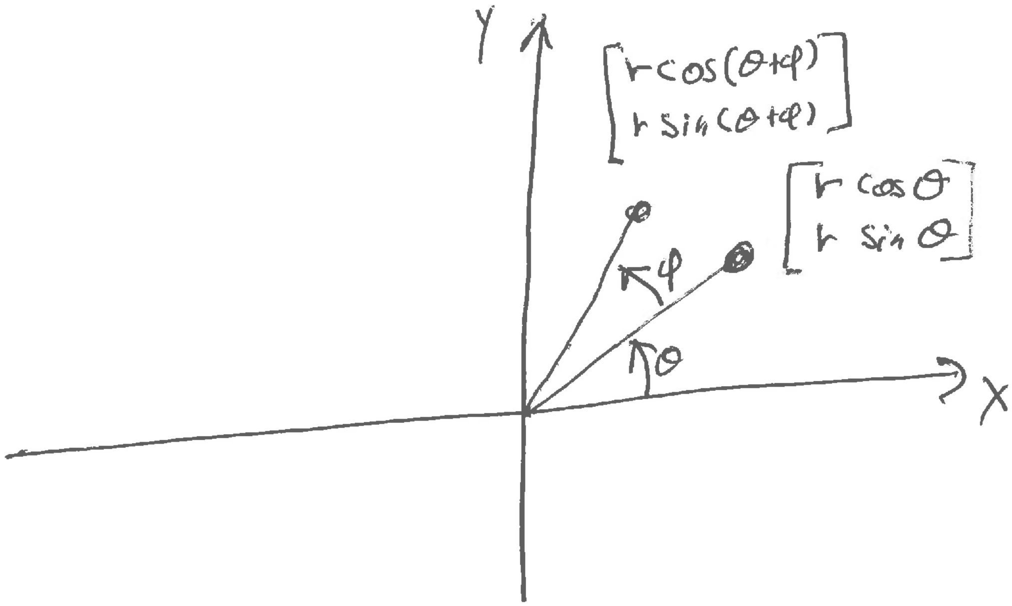
and

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix} \right) = \begin{bmatrix} r\cos(\theta+\phi) \\ r\sin(\theta+\phi) \end{bmatrix}.$$

I put a picture on the next page.

[18] One reads this equation as T **OF** v is equal to A **TIMES** v.

Rotation


Y
X
$\begin{bmatrix} r\cos(\theta+\varphi) \\ r\sin(\theta+\varphi) \end{bmatrix}$
$\begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix}$
φ
θ

It would be useful if you know

$$\begin{aligned}\cos(\theta+\phi) &= \cos\theta\cos\phi - \sin\theta\sin\phi \\ \sin(\theta+\phi) &= \sin\theta\cos\phi + \cos\theta\sin\phi\end{aligned} \tag{10.2.1}$$

Maybe you learned these things in High School; maybe you did not learn them until you took Differential Equations (Math 242 or Math 552). At any rate Euler proved that

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

We can see this using Taylor's series:

$$\begin{aligned}e^z &= 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \frac{z^5}{5!} + \dots, \\ \cos(z) &= 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots, \\ \sin(z) &= z - \frac{z^3}{3!} + \frac{z^5}{5!} + \dots,\end{aligned}$$

for all complex numbers z. It follows that

$$\begin{aligned}e^{i\theta} &= 1 + (i\theta) + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \dots \\ &= 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \dots + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \dots\right) \\ &= \cos\theta + i\sin\theta.\end{aligned}$$

At any rate,

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

Equate the real and imaginary components to conclude (10.2.1).

Now we see that

$$\begin{aligned}T\left(\begin{bmatrix}x\\y\end{bmatrix}\right) &= T\left(\begin{bmatrix}r\cos\theta\\r\sin\theta\end{bmatrix}\right) = \begin{bmatrix}r\cos(\theta+\phi)\\r\sin(\theta+\phi)\end{bmatrix} = \begin{bmatrix}r\cos\theta\cos\phi - r\sin\theta\sin\phi\\r\sin\theta\cos\phi + r\cos\theta\sin\phi\end{bmatrix} \\ &= \begin{bmatrix}x\cos\phi - y\sin\phi\\y\cos\phi + x\sin\phi\end{bmatrix} = \begin{bmatrix}\cos\phi & -\sin\phi\\\sin\phi & \cos\phi\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}.\end{aligned}$$

Thus T is a linear transformation as we claimed.

(g) Find a matrix A so that Av is the reflection of v across the x-axis for all v in $\mathbb{R}^2$.

This is easy. When $\begin{bmatrix} x \\ y \end{bmatrix}$ is reflected across the x-axis, the result is $\begin{bmatrix} x \\ -y \end{bmatrix}$. It follows that
$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

(h) Let ℓ be the line in $\mathbb{R}^2$ through the origin which makes the angle ϕ with the positive x-axis. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the function which fixes ℓ and reflects the xy-plane across ℓ. Find a matrix A with $Tv = AV$ for all $v \in \mathbb{R}^2$. Conclude that T is a linear transformation.

We find A in three steps.

- Rotate ℓ to become the x-axis.
- Reflect across the x-axis.
- Rotate the x-axis back to ℓ.

So,
$$\begin{aligned} A &= \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos(-\phi) & -\sin(-\phi) \\ \sin(-\phi) & \cos(-\phi) \end{bmatrix} \\ &= \begin{bmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\phi - \sin^2\phi & 2\cos\phi\sin\phi \\ 2\sin\phi\cos\phi & \sin^2\phi - \cos^2\phi \end{bmatrix} = \begin{bmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{bmatrix}. \end{aligned}$$

Once again, we conclude that T is a linear transformation as claimed.

11. Eigenvalues and Eigenvectors.

Definition 11.1. Let A be an $n \times n$ matrix. The number λ is an eigenvalue of A if there is a non-zero vector $v \in \mathbb{R}^n$ with $Av = \lambda v$. If λ is an eigenvalue of A, then every vector v (including the zero vector) with $Av = \lambda v$ is an eigenvector of A associated to the eigenvalue λ.

Example 11.2. Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$. Observe that

$$A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

It follows that 1 is an eigenvalue of A and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector of A associated to the eigenvalue 1. Observe also, that

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

It follows that 3 is an eigenvalue of A and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of A associated to the eigenvalue 3.

11.A. **How does one find the eigenvalues of the square matrix A.** Observe that

$$\begin{aligned} & \lambda \text{ is an eigenvalue of } A \\ \iff & Av = \lambda v \text{ for some non-zero vector } v \\ \iff & (A - \lambda I)v = 0 \text{ for some non-zero vector } v \\ \iff & A - \lambda I \text{ is a singular matrix} \\ \iff & \det(A - \lambda I) = 0 \end{aligned}$$

Actually, you might not know the last step. Recall that the determinant of the 2×2 matrix

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

is $ad - bc$. Furthermore, A is nonsingular if and only if $\det A \neq 0$. See 5.9.(e) on page 28. We will not have time to deal with determinants for bigger matrices. Nonetheless, if A is a square matrix, then there is a number associated to A, called $\det A$ and A is nonsingular if and only if $\det A \neq 0$.

At any rate,

True Statement 11.3. *The number λ is an eigenvalue of the square matrix A if and only if* $\det(A - \lambda I) = 0$.

Example 11.4. Find all eigenvalues of $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$.

Well, $\det(A-\lambda I) = \det\begin{bmatrix}1-\lambda & 2\\ 0 & 3-\lambda\end{bmatrix} = (1-\lambda)(3-\lambda)$. Thus, $\det(A-\lambda I)=0$ if only if $(1-\lambda)(3-\lambda)=0$. The eigenvalues of A are $\lambda=1$ and $\lambda=3$.

11.B. **Once you know that the number λ is an eigenvalue of the square matrix A, how do you find all of the eigenvectors of A associated to the eigenvalue λ?**

Remark 11.5. Notice, first of all, the the set of all eigenvectors of A associated to the eigenvalue λ is a **vector space**. Indeed this set is the null space of $A-\lambda I$.

Definition 11.6. If λ is an eigenvalue of the square matrix A, then the eigenspace of A associated to λ is the set of all eigenvectors of A associated to λ.

Example 11.7. Find all eigenvectors of the matrix $A=\begin{bmatrix}1 & 2\\ 0 & 3\end{bmatrix}$.

We saw in Example 11.4 that the eigenvalues of A are 1 and 3.

- We find the eigenspace of A associated to $\lambda=1$. That is, we find all vectors $x=\begin{bmatrix}x_1\\ x_2\end{bmatrix}$ such that
$$Ax = 1x.$$
In other words, we find all x with
$$(A-I)x = 0.$$
We find the null space of
$$A-I.$$
We find the null space of
$$\begin{bmatrix}0 & 2\\ 0 & 2\end{bmatrix}. \tag{11.7.1}$$
If you apply Elementary Row Operations (ERO) you obtain
$$\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}. \tag{11.7.2}$$
(Of course, the point of (EROs) is that the matrices (11.7.1) and (11.7.2) have the same null space. The null space of (11.7.2) is the vector space with basis[19]
$$v_1=\begin{bmatrix}1\\ 0\end{bmatrix}.$$

$$\boxed{\text{The eigenspace of } A \text{ associated to the eigenvalue } \lambda=1 \text{ is the vector space with basis } v_1=\begin{bmatrix}1\\ 0\end{bmatrix}.}$$

is half of our answer.

[19]This is not a big deal. We see that x_1 is a free variable and x_2 is required to be zero.

- We find the eigenspace of A associated to $\lambda = 3$. That is, we find all vectors $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ such that
$$Ax = 3x.$$
In other words, we find all x with
$$(A - 3I)x = 0.$$
We find the null space of
$$A - 3I.$$
We find the null space of
$$\begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix}. \tag{11.7.3}$$
If you apply Elementary Row Operations (ERO) you obtain
$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}. \tag{11.7.4}$$
(Of course, the point of (EROs) is that the matrices (11.7.3) and (11.7.4) have the same null space. The null space of (11.7.4) is the vector space with basis[20]
$$v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

$$\boxed{\text{The eigenspace of } A \text{ associated to the eigenvalue } \lambda = 3 \text{ is the vector space with basis } v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.}$$

is the other half of our answer.

[20]This is not a big deal. We see that x_2 is a free variable and x_1 is required to be x_2.

12. Diagonalization.

There are many procedures which are easy for numbers and easy for diagonal matrices, but hard for ordinary matrices.

In this discussion, let a be a number,

$$D = \begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 & \dots & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & \dots & d_n \end{bmatrix}$$

be a diagonal matrix, and A be an arbitrary matrix.

- raising to a power: a^m is easy and
$$D^m = \begin{bmatrix} d_1^m & 0 & 0 & \dots & 0 \\ 0 & d_2^m & 0 & \dots & 0 \\ 0 & 0 & d_3^m & \dots & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & \dots & d_n^m \end{bmatrix}$$
is easy; but A^m is a nuisance;
- taking the limit of the m^{th} power: $\lim_{m\to\infty} a^m$ is easy and
$$\lim_{m\to\infty} D^m = \begin{bmatrix} \lim_{m\to\infty} d_1^m & 0 & 0 & \dots & 0 \\ 0 & \lim_{m\to\infty} d_2^m & 0 & \dots & 0 \\ 0 & 0 & \lim_{m\to\infty} d_3^m & \dots & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & \dots & \lim_{m\to\infty} d_n^m \end{bmatrix}$$
is easy; but it is not at all clear how one computes (or if one can compute) $\lim_{m\to\infty} A^m$. (These are interesting problems if the entry of the matrix in row i column j represents the probability in some Markov process of moving from state i to state j, when one wants to learn the ultimate state of the process.)
- Linear Differential equations. It is easy to solve the linear differential equation $\frac{dy}{dx} = ay$. (The solution is $y = ce^{ax}$.) It is easy to solve the system of linear differential equations
$$\begin{bmatrix} \frac{dy_1}{dx} \\ \frac{dy_2}{dx} \\ \vdots \\ \frac{dy_n}{dx} \end{bmatrix} = D \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$
The solution is
$$y_1 = c_1 e^{d_1 x}, \dots, y_n = c_n e^{d_n x}.$$

But it is harder to solve

$$\begin{bmatrix} \frac{dy_1}{dx} \\ \frac{dy_2}{dx} \\ \vdots \\ \frac{dy_n}{dx} \end{bmatrix} = A \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

- taking roots: It is easy to find $\sqrt{a}$. It is easy to find a matrix B with $B^2 = D$; namely

$$B = \begin{bmatrix} \sqrt{d_1} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{d_2} & 0 & \dots & 0 \\ 0 & 0 & \sqrt{d_3} & \dots & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & \dots & \sqrt{d_n} \end{bmatrix}$$

  But it is not clear how one might find B with $B^2 = A$, when A is an arbitrary matrix.

Definition 12.1. The square matrices A and B are <u>similar</u> if there exists an invertible matrix P with $A = PBP^{-1}$.

Definition 12.2. The square matrix A is <u>diagonalizable</u> if A is similar to a diagonal matrix.

Theorem 12.3. *Let A be an $n \times n$ matrix. If $v_1, \dots, v_n$ are linearly independent vectors and each v_i is an eigenvector of A, then A is diagonalizable.*

Proof. Suppose $Av_i = \lambda_i v_i$ for all i. Observe that

$$A \left[v_1 | v_2 | \cdots | v_n \right] = \left[v_1 | v_2 | \cdots | v_n \right] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

Let D be the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

and P be the invertible[21] matrix $\left[v_1 | v_2 | \cdots | v_n \right]$. We have shown that $A = PDP^{-1}$. Thus, we have shown that A is diagonalizable. $\square$

Example 12.4. Let $A = \begin{bmatrix} 7 & 6 \\ -3 & -2 \end{bmatrix}$. Find a matrix B with $B^2 = A$.

We diagonalize A. Find the eigenvalues and eigenvectors of A (See for example Examples 11.4 and 11.7.)

[21] The columns of P are linearly independent; so P is invertible.

At any rate, the eigenvalues of A are 1 and 4. Furthermore,

$$A\begin{bmatrix}1\\-1\end{bmatrix}=1\begin{bmatrix}1\\-1\end{bmatrix}\quad\text{and}\quad A\begin{bmatrix}2\\-1\end{bmatrix}=4\begin{bmatrix}2\\-1\end{bmatrix}.$$

Let $D=\begin{bmatrix}1&0\\0&4\end{bmatrix}$ and $P=\begin{bmatrix}1&2\\-1&-1\end{bmatrix}$. We have shown that $AP=PD$; hence $A=PDP^{-1}$. Let $B=P\begin{bmatrix}1&0\\0&2\end{bmatrix}P^{-1}$. Observe that

$$B^2=\left(P\begin{bmatrix}1&0\\0&2\end{bmatrix}P^{-1}\right)\left(P\begin{bmatrix}1&0\\0&2\end{bmatrix}P^{-1}\right)=P\begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}1&0\\0&2\end{bmatrix}P^{-1}=P\begin{bmatrix}1&0\\0&4\end{bmatrix}P^{-1}=A.$$

Let us do the arithmetic in order to actually exhibit B. Well,[22]

$$B=P\begin{bmatrix}1&0\\0&2\end{bmatrix}P^{-1}=\begin{bmatrix}1&2\\-1&-1\end{bmatrix}\begin{bmatrix}1&0\\0&2\end{bmatrix}\begin{bmatrix}-1&-2\\1&1\end{bmatrix}=\begin{bmatrix}1&4\\-1&-2\end{bmatrix}\begin{bmatrix}-1&-2\\1&1\end{bmatrix}=\boxed{\begin{bmatrix}3&2\\-1&0\end{bmatrix}}.$$

We check

$$B^2=\begin{bmatrix}3&2\\-1&0\end{bmatrix}\begin{bmatrix}3&2\\-1&0\end{bmatrix}=\begin{bmatrix}7&6\\-3&-2\end{bmatrix}=A\checkmark.$$

[22] I used Example 5.9.(e) to find P^{-1}.

13. ORTHOGONAL SETS.

Definition 13.1. The set of vectors $u_1, \dots, u_p$ in $\mathbb{R}^n$ is an orthogonal set if $u_i^{\mathrm{T}} u_j = 0$, for $i \neq j$.

Examples 13.2.

- The vectors
$$\begin{bmatrix}1\\0\\0\end{bmatrix}, \quad \begin{bmatrix}0\\1\\0\end{bmatrix}, \quad \begin{bmatrix}0\\0\\1\end{bmatrix}$$
form an orthogonal set.
- The vectors
$$\begin{bmatrix}1\\1\\1\end{bmatrix}, \quad \begin{bmatrix}1\\-1\\0\end{bmatrix}, \quad \begin{bmatrix}1\\1\\-2\end{bmatrix}$$
form an orthogonal set.

This section has two ideas:

- Orthogonal sets are nice to have.
- How do we find orthogonal sets?

13.A. **Orthogonal sets are nice to have.**

(a) If $u_1, \dots, u_p$ is an orthogonal set of non-zero vectors in $\mathbb{R}^n$, then the vectors $u_1, \dots, u_p$ are linearly independent.

(b) Suppose $u_1, \dots, u_p$ is an orthogonal set of non-zero vectors in $\mathbb{R}^n$. Let U be the subspace of $\mathbb{R}^n$ spanned by $u_1, \dots, u_p$. If u is an arbitrary element of U, then it is easy to find the coefficients of u with respect to $u_1, \dots, u_p$.

The proof of (a). Suppose $c_1, \dots, c_p$ are real numbers and $\sum_{i=1}^{p} c_i u_i = 0$. Multiply by u_j^{T} to learn that

$$c_j(u_j^{\mathrm{T}} u_j) = 0,$$

for each j. Notice that c_j and $(u_j^{\mathrm{T}} u_j)$ are both real numbers. If the product of two real numbers is zero, then one of the numbers is zero. The real number $(u_j^{\mathrm{T}} u_j)$ is the sum of perfect squares of real numbers (and u_j has at least one non-zero entry by hypothesis). It follows that $c_j = 0$ for each j. $\square$

Justification of (b). Obviously, you could Elementary Row Operations to the matrix

$$[u_1|u_2|\dots|u_p|u]$$

in order to find numbers $c_1, \dots, c_p$ with

$$\sum_{i=1}^{p} c_i u_i = u.$$

But that takes a significant amount of effort.

There is a much easier way to find numbers $c_1,\dots,c_p$ with

$$\sum_{i=1}^{p} c_i u_i = u.$$

Merely multiply both sides of the equation by u_j^{T} to see that

$$c_j u_j^{\mathrm{T}} u_j = u_j^{\mathrm{T}} u;$$

hence,

$$c_j = \frac{u_j^{\mathrm{T}} u}{u_j^{\mathrm{T}} u_j},$$

for each j.

Example 13.3. Write $u = \begin{bmatrix}1\\2\\3\end{bmatrix}$ as a linear combination of

$$u_1 = \begin{bmatrix}1\\1\\1\end{bmatrix}, \quad u_2 = \begin{bmatrix}1\\-1\\0\end{bmatrix}, \quad u_3 = \begin{bmatrix}1\\1\\-2\end{bmatrix}$$

If $u = \sum_{i=1}^{3} c_i u_i$, then

$$u_1^{\mathrm{T}} u = c_1 u_1^{\mathrm{T}} u_1;$$

hence $2 = \frac{1+2+3}{3} = c_1$;

$$u_2^{\mathrm{T}} u = c_2 u_2^{\mathrm{T}} u_2;$$

hence $\frac{-1}{2} = \frac{1-2}{2} = c_2$; and $c_3 = \frac{3-6}{6} = \frac{-1}{2}$.

Check.

$$2\begin{bmatrix}1\\1\\1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\-1\\0\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\1\\-2\end{bmatrix} = \begin{bmatrix}2-1\\2\\2+1\end{bmatrix} = \begin{bmatrix}1\\2\\3\end{bmatrix} \checkmark.$$

13.B. **How do we find orthogonal sets?** One uses Gram-Schmidt Orthogonalization to find an orthogonal basis for a vector space.

Let $v_1,\dots,v_n$ be linearly independent vectors. We produce $u_1,\dots,u_n$, which is an orthogonal set and also is a basis for the span of $v_1,\dots,v_n$.

Define

$$\begin{aligned}
u_1 &= v_1 \\
u_2 &= v_2 - \frac{u_1^{\mathrm{T}} v_2}{u_1^{\mathrm{T}} u_1} u_1 && \text{Notice that } u_1^{\mathrm{T}} u_2 = 0. \\
u_3 &= v_3 - \frac{u_1^{\mathrm{T}} v_3}{u_1^{\mathrm{T}} u_1} u_1 - \frac{u_2^{\mathrm{T}} v_3}{u_2^{\mathrm{T}} u_2} u_2 && \text{Notice that } u_1^{\mathrm{T}} u_3 = u_2^{\mathrm{T}} u_3 = 0. \\
u_4 &= v_4 - \frac{u_1^{\mathrm{T}} v_4}{u_1^{\mathrm{T}} u_1} u_1 - \frac{u_2^{\mathrm{T}} v_4}{u_2^{\mathrm{T}} u_2} u_2 - \frac{u_3^{\mathrm{T}} v_4}{u_3^{\mathrm{T}} u_3} u_3 && \text{Notice that } u_1^{\mathrm{T}} u_4 = u_2^{\mathrm{T}} u_4 = u_3^{\mathrm{T}} u_4 = 0.
\end{aligned}$$

etc.

Example 13.4. Find an orthogonal basis for the null space of $A = \begin{bmatrix} 1 & 2 & 1 & 3 \end{bmatrix}$.

One basis for the null space of A is

$$v_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} -3 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Let $u_1 = v_1$. Let

$$u_2' = v_2 - \frac{u_1^{\mathrm{T}} v_2}{u_1^{\mathrm{T}} u_1} u_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{5} \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} -1 \\ -2 \\ 5 \\ 0 \end{bmatrix}.$$

Let $u_2 = 5u_2' = \begin{bmatrix} -1 \\ -2 \\ 5 \\ 0 \end{bmatrix}$. We verify that u_2 is in the null space of A and $u_2^{\mathrm{T}} u_1 = 0$. Let

$$u_3' = v_3 - \frac{u_1^{\mathrm{T}} v_3}{u_1^{\mathrm{T}} u_1} u_1 - \frac{u_2^{\mathrm{T}} v_3}{u_2^{\mathrm{T}} u_2} u_2 = \begin{bmatrix} -3 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{6}{5} \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{3}{30} \begin{bmatrix} -1 \\ -2 \\ 5 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} -3 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{6}{5} \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{10} \begin{bmatrix} -1 \\ -2 \\ 5 \\ 0 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} -5 \\ -10 \\ -5 \\ 10 \end{bmatrix} = \frac{5}{10} \begin{bmatrix} -1 \\ -2 \\ -1 \\ 2 \end{bmatrix}.$$

Let

$$u_3 = 10u_3' = \begin{bmatrix} -1 \\ -2 \\ -1 \\ 2 \end{bmatrix}.$$

Thus

$$\boxed{u_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad u_2 = \begin{bmatrix} -1 \\ -2 \\ 5 \\ 0 \end{bmatrix}, \quad u_3 = \begin{bmatrix} -1 \\ -2 \\ -1 \\ 2 \end{bmatrix}}$$

is an orthogonal basis for the null space of A. Be sure to verify that $Au_i = 0$ and $u_i^{\mathrm{T}} u_j = 0$ for $i \neq j$.