Math 305: Linear Algebra Lecture Notes

Created by: Professors Amanda Harsy and Michael Smith

July 12, 2024

About the Course

Thanks for taking Linear Algebra with me! It is one of my favorite classes to teach. You may be asking yourself (or have asked yourself), "What is Linear Algebra, and why do I have to take this class?"

"I believe that linear algebra is ***the*** *most important subject in college mathematics. Isaac Newton would not agree! But he isn't taking mathematics in the 21^{st} century."*

-Gilbert Strang

I, and many other mathematicians, definitely agree with Strang in that Linear Algebra has emerged as a giant in mathematics -especially now that we have computers to do what used to be extremely tedious work. It has such a vast set of applications and touches almost every mathematical subject. Linear algebra provides "essential preparation for advanced work in the sciences, statistics, and computing. Linear algebra also introduces students to discrete mathematics, algorithmic thinking, a modicum of abstraction, moderate sophistication in notation, and simple proofs. Linear algebra helps students develop facility with visualization, see connections among mathematical areas, and appreciate the power of abstract thinking."[1]

Linear algebra is basically the study of multivariate linear systems and transformations and, in my opinion, is at its core trying to solve Two Fundamental Problems:

1. Solving $Ax = b$, and
2. Diagonalizing a matrix A (AKA Eigenvalue Problems).

The first problem relates to exploiting linear methods to solve complex and dynamical systems and situations. Basically, it is using one of the best problem-solving techniques mathematicians use, what I like to call the "Wouldn't it be nice if..." approach. That is, real life is a mess and we often have to deal with really complex functions (if we are even lucky enough to have a function at all!) which are difficult to manage. So instead we use linear functions (lines and planes) to approximate or model the complex, real-world situation, which is much easier.

The second problem relates to simplifying our system so that we can more easily solve or approximate systems. The fact that some systems don't have solutions leads directly into the mathematical field of Numerical Analysis and we will dive into some basic numerical analysis in this course. Because so many situations in life can be modeled linearly, Linear Algebra shows up in many topics including (but not exhaustively) "Markov chains, graph theory, correlation coefficients, cryptology, interpolation, long-term weather prediction, the Fibonacci sequence, difference equations, systems of linear differential equations, network analysis, linear least squares, graph theory, Leslie population models, the power method of approximating the dominant eigenvalue, linear programming, computer graphics, coding theory, spectral decomposition, principal component analysis, discrete and continuous

[1] Schumacher, etc. 2015 *2015 CUPM Curriculum Guide to Majors in the Mathematical Sciences* **37,39**

dynamical systems, iterative solutions of linear systems, image processing, and traffic flow."[1]

Another reason you are taking this course (which perhaps is why anyone is required to take a mathematics class) is to learn how to think abstractly through problem solving.

I hope you will enjoy this semester and learn a lot! Some students struggle at first with Linear Algebra because it is one of the first math courses students take which starts to really exploit abstract notation and thinking. Hang in there! It just takes time to digest. Please make use of my office hours and plan to work hard in this class. My classes have a high work load (as all math classes usually do!), so make sure you **stay on top of your assignments and get help early**. Remember you can also email me questions if you can't make my office hours or make an appointment outside of office hours for help. When I am at Lewis, I usually keep the door open and feel free to pop in at any time. If I have something especially pressing, I may ask you to come back at a different time, but in general, I am usually available. The Practice Problems for Exams are at the end of this course packet, and I have most likely handed out separate packets for the HW Assignments and Labs as well.

We have worked hard to create this course packet for you, but it is still a work in progress. Please be understanding of the typos we have not caught, and politely bring them to my attention so I can fix them for the next time I teach this course. I look forward to meeting you and guiding you through the wonderful course that is Linear Algebra.

Cheers,
Professor Smith, Professor Stratton, and Dr. Harsy

Acknowledgments: No math teacher is who they are without a little help. I would like to thank two professors in particular that have had an impact on the way I teach and approach problems: Dr. Amanda Harsy of our own Lewis University (who also teaches a section of this course!) and Dr. Nathan Krislock of Northern Illinois University. I also want to thank Dr. Heather Moon, Dr. Marie Snipes, Dr. Tom Asaki, Dr. Tim Chartier, Dr. Scott Kaschner, the members from both the IOLA and MathVote projects for sharing some of their resources and labs from their own courses. *And finally, we would like to thank you and all the other students for making this job worthwhile and for all the suggestions and encouragement you have given me over the years to improve.*

Contents

1 Solving Systems of Linear Equations

1.1 MA 307 –ICE 0 -Motivation for Linear Algebra

1. Think of a number between 1 and 20. Double the number and then add 10. Next divide by 2 and then subtract your original number. What number did you get?

2. What is linear algebra used for?

In algebra class, we learned how to solve systems of equations. Let's review some of these!

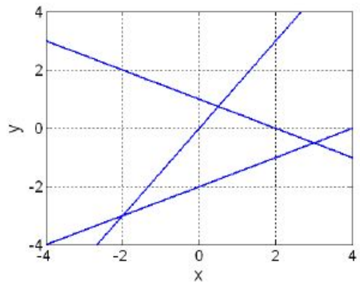
4. A system of linear equations could not have exactly ________ solutions.
(a) 0
(b) 1
(c) 2
(d) infinite
(e) All of these are possible numbers of solutions to a system of linear equations.

5. What is the solution to the following system of equations?
$2x + y = 3$
$3x - y = 7$

(a) $x = 4$ and $y = -5$
(b)$x = 4$ and $y = 5$
(c) $x = 2$ and $y = -1$
(d) There are an infinite number of solutions to this system.
(e) There are no solutions to this system.

6. Create a set of linear equations (called a linear system) that has no solutions.

7. We have a system of three linear equations with two unknowns, as plotted in the graph below. How many solutions does this system have?



(a) 0
(b) 1
(c) 2
(d) 3
(e) Infinite

8. Set up and solve the the following problem: "There are three classes of grain, of which three bundles of the first class, two of the second, and one of the third make 39 measures. Two of the first, three of the second, and one of the third make 34 measures. And one of the first, two of the second, and three of the third make 26 measures. How many measures of grain are contained in one bundle of each class?"

-*Jiuzhang Suanshu*, a Chinese manuscript from about 200 BC. [1]

[1] *When Life is Linear* pg 3.

1.2 Linear Systems and Matrix Notation

During your Ice sheet, we reviewed solving simple systems of equations. This is very nice when we have only a few variables, but it quickly gets tedious when we have systems with even more variables. Up until now, our method for solving these systems has been a little ad-hoc. In this section we will discuss an algorithm which will help us solve larger systems of equations. Now, this algorithm may seem weird and tedious, and is a rather uninspiring way to start this course, but it has allowed us to program computer to do this work for us. And after HW 1, you will be able to use Octave, Matlab, Sage, or a calculator to do this for you. Let's get started...

Definition 1.1. *A **linear combination** of variables (unknowns) $x_1, ..., x_n$ is an expression in the form:*

where the a_i's $(a_1, a_2, ..., a_n)$ are real numbers which are called the combination's ________________

Examples:

Non-Examples:

Definition 1.2. *We can turn a linear combination into a **linear equation** by adding "$= b_i$":*

*Then we can create a **system** of linear equations:*

$$\begin{gathered}
a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + ... + a_{1,n}x_n = b_1 \\
a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + ... + a_{2,n}x_n = b_2 \\
\vdots \\
a_{n,1}x_1 + a_{n,2}x_2 + a_{n,3}x_3 + ... + a_{n,n}x_n = b_n
\end{gathered}$$

Definition 1.3. *A solution to a linear system is an "n-tuple"* ________________ *which solves **each** equation in the system.*

$$
\begin{gathered}
a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 + ... + a_{1,n}x_n = b_1 \\
a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 + ... + a_{2,n}x_n = b_2 \\
\vdots \\
a_{n,1}x_1 + a_{n,2}x_2 + a_{n,3}x_3 + ... + a_{n,n}x_n = b_n
\end{gathered}
$$

We would like to be able to record a linear system compactly. We can do this by using Matrices:

The Coefficient Matrix:

$$A = (a_{ij}) = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} \end{bmatrix}$$

The Augmented Matrix:

$$\big(A \mid b\big) = \left(\begin{array}{ccccc|c} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} & b_n \end{array}\right)$$

Example 1.1. *Write the following system as an augmented matrix:*
$x_1 - 3x_2 + 5x_3 - 2x_4 = 0$
$x_2 - 3x_3 = 2$
$9x_3 - 4 = 0$
$x_4 = 1$

Example 1.2. *a) Convert the following augmented matrix back into a system of equations:*

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

b) When you solve this system, you get $x = -1, y = 2, z = -1$. Is this an equivalent form of the system of equations above?

c) Write this system in augmented form.

Wouldn't it have been nice if we were given the latter system to solve rather than the initial system? It turns out we have a name for the "nicest form" of an augmented matrix.

That is we want an augmented matrix that has the maximum number of zeros, but doesn't change the original system's solution. It turns out there is a an algorithm we can use to convert a system into this "nice form." But first, let's describe what we mean by "nice" a little more formally.

1.3 Echelon Forms

Definition 1.4. *The **leading term** of a row is the leftmost nonzero term in that row. If a row has all zeros, it has no leading term.*

Definition 1.5. *A Matrix/system is in **echelon form** if*

1. *Every leading term is in a column to the left of the leading term of the row below it.*
2. *Any zero rows are at the bottom of the matrix.*

Examples:

$$\begin{bmatrix} 1 & 2 & 3 & 5 \\ 0 & 4 & 1 & -2 \\ 0 & 0 & 3 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 0 & 5 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \left(\begin{array}{cc|c} 0 & 1 & 5 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right), \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \left(\begin{array}{ccc|c} 0 & 1 & 5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Example 1.3. *Which of the following matrices are in echelon form?*

$$\left(\begin{array}{cc|c} 0 & 1 & 5 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right) \quad \begin{bmatrix} 1 & 0 & 0 & 5 \\ 2 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 2 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 5 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Definition 1.6. *In an echelon form system, a* ***pivot position*** *is an entry that corresponds to a leading term in the echelon form of A. Variables that correspond to a pivot position are called* ***pivot variables****.*

Definition 1.7. *In an echelon form system, the variables that are not leading are called __________ variables.*

Example 1.4. *Which variable(s) are pivots? Which variable(s) are free for the following matrix representation with variables x_1, x_2, x_3, x_4?*

$$\left(\begin{array}{cccc|c} 2 & 4 & 0 & 4 & 0 \\ 0 & 9 & 18 & 3 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right).$$

Definition 1.8. *A Matrix/system is in* ***reduced row echelon form*** *if*

1. *It is in echelon form.*
2. *All pivot positions contain a 1.*
3. *The only nonzero term in a pivot column is in pivot position.*

Examples:

$$\begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 1 & 2 \end{bmatrix}, \qquad \left(\begin{array}{ccccc|c} 0 & 1 & 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & -6 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right), \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Example 1.5. *Which of the following matrices are in Reduced Row Echelon form?*

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

1.4 Gauss's Method for Solving Systems of Linear Equations

It would be nice if we were given a system whose matrix form was in echelon form or even better, reduced row echelon form (RREF) (sometimes this form is called row reduced echelon form). In algebra courses, you often solve a system of 2 (maybe 3) equations using substitution or elimination. Unfortunately, this can get tricky when we have large systems though. Luckily we have an algorithm to help us called **Gaussian Elimination.**

Gaussian Elimination/ Gauss's Method: If a linear system is changed to another by one of the operations below, then the two systems have the same set of solutions (that is, they are equivalent systems).

1. **Swapping:** Swap one row (equation) with another.

2. **Rescaling:** Multiply one row (equation) by a nonzero constant.

3. **Row Combination:** Replace one row (equation) by the sum of itself and a multiple of another.

Definition 1.9. *The 3 operations above are called* ***elementary row operations****. Note that they are reversible.*

Basic Strategy:
Replace one system of equations with an *equivalent* system that is **easier** to solve.

In the end, we want a matrix that is as close as it can be to looking like: $\left(\begin{array}{cccccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & * \\ 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & * \\ 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & * \end{array}\right)$

Why is this form helpful?

Tie in with Echelon Forms:
Notice there seems to be two "steps" in Gaussian Elimination:

Step 1: Get matrix in "**Upper Triangular Form:** $\left(\begin{array}{cccccccc|c} 1 & * & * & * & * & * & \dots & * & * \\ 0 & 1 & * & * & * & * & \dots & * & * \\ 0 & 0 & 1 & * & * & * & \dots & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & * \end{array}\right)$

Note this can also be something like: $\left(\begin{array}{cccccccc|c} 1 & * & * & * & * & * & \dots & * & * \\ 0 & 1 & * & * & * & * & \dots & * & * \\ 0 & 0 & 1 & * & * & * & \dots & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & * \end{array}\right)$

Step 2: Back substitute to get the matrix in Row Reduced Reduced Echelon Form.

Note each matrix's Row Reduced Echelon Form is unique!

Example 1.6. *Solve the linear system represented by the augmented matrix using Gaussian Elimination:*

$$\left(\begin{array}{ccc|c} 6 & 3 & -2 & -4 \\ 0 & 2 & -6 & -8 \\ 1 & 0 & 2 & 3 \end{array}\right)$$

What could you be asking yourself about linear systems?

1.

2.

We can use Guassian Elimination to identify how many solutions we will have in a system.

Example 1.7. *Convert the following augmented matrices back into a system of equations:*

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

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

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

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

Theorem 1.1. *A system of equations can have __________ solutions, exactly __________ solutions, or __________ solutions.*

Definition 1.10. *A system is called **consistent** if it has __________ or __________ solutions. A system is called **inconsistent** if it has __________ solutions.*

Example 1.8. *Solve the system of equations using Gaussian Elimination.*

$$\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 4 \\ x_1 & + & x_2 & + & x_3 & = & 0 \end{array}$$

Example 1.9. *Solve the system of equations using Gaussian Elimination.*

$$\begin{array}{rcrcrcr} 2x & + & 2y & + & 2z & = & 4 \\ -x & + & y & - & 2z & = & -2 \\ x & + & y & + & z & = & 2 \end{array}$$

1.5 ICE 1: Solving Systems of Equations

1. What is the solution to the following system of equations? (Note this is called a homogeneous system since the right-hand-sides of each equation is 0.)

   $x + 2y + \ \ z = 0$

   $x + 3y - 2z = 0$

   Archer has reduced this system into Row Reduced Echelon Form (RREF): $\left(\begin{array}{ccc|c} 1 & 0 & 7 & 0 \\ 0 & 1 & -3 & 0 \end{array}\right)$

   (a) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \{(-7z, 3z, z) | z \in \mathbb{R}\}$

   (b) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \{(7z, -3z, z) | z \in \mathbb{R}\}$

   (c) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \{(-7z, 3z, 0) | z \in \mathbb{R}\}$

   (d) None of the above.

   (e) More than one of the above.

   Bonus question: How many vectors are in this solution set?

2. What is the solution to the following system of equations? (Note this system is called a non-homogeneous system.)

   $x + 2y + \ \ z = 3$

   $x + 3y - 2z = 4$

   Eva has reduced this system into RREF: $\left(\begin{array}{ccc|c} 1 & 0 & 7 & 1 \\ 0 & 1 & -3 & 1 \end{array}\right)$

   (a) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \{(7 + z, 3 + z, z) | z \in \mathbb{R}\}$

   (b) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \{(-1 - 7z, -1 + 3z, z) | z \in \mathbb{R}\}$

   (c) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \{(1 - 7z, 1 + 3z, z) | z \in \mathbb{R}\}$

   (d) None of the above.

   (e) More than one of the above.

Bonus question: What is the relationship between the previous two solution sets?

3. What is the solution to the following system of equations (called a non-homogeneous system)?

$x + 2y + z = -2$

$x + 3y - 2z = 1$

Archer has reduced this system into RREF: $\left(\begin{array}{ccc|c} 1 & 0 & 7 & -8 \\ 0 & 1 & -3 & 3 \end{array}\right)$

(a) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \{(-8 - 7z, 3 + 3z, z) | z \in \mathbb{R}\}$

(b) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \{(8 - 7z, -3 + 3z, z) | z \in \mathbb{R}\}$

(c) $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \{(-8 + 7z, 3 - 3z, -z) | z \in \mathbb{R}\}$

(d) None of the above.

(e) More than one of the above.

4. What is the value of a so that the linear system represented by the following matrix would have infinitely many solutions?

$\left(\begin{array}{cc|c} 2 & 6 & 8 \\ 1 & a & 4 \end{array}\right)$

(a) a = 0

(b) a = 2

(c) a = 3

(d) This is not possible.

(e) More than one of the above

5. Solve the system represented by the augmented matrix $\begin{bmatrix} 1 & 7 & 3 & -4 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}$. [Hint: Eva says that you can answer this without doing any work!]

6. Rewrite the linear system represented by the augmented matrix below in Row Reduced Echelon Form to find the solution set: $\begin{bmatrix} 1 & -1 & 0 & 0 & -4 \\ 0 & 1 & -3 & 0 & -7 \\ 0 & 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 1 & -3 \end{bmatrix}$.

7. Let matrix R be the reduced row echelon form of matrix A. True or False The solutions to $Rx = 0$ are the same as the solutions to $Ax = 0$.

   (a) True, and I am very confident
   (b) True, but I am not very confident
   (c) False, but I am not very confident
   (d) False, and I am very confident

8. Let matrix R be the reduced row echelon form of matrix A. True or False: The solutions to $Rx = b$ are the same as the solutions to $Ax = b$.

   (a) True, and I am very confident
   (b) True, but I am not very confident
   (c) False, but I am not very confident
   (d) False, and I am very confident

9. Solve the system from *Jiuzhang Suanshu*.

   $3x + 2y + z = 39$
   $2x + 3y + z = 34$
   $x + 2y + 3z = 26$

2 Vector Spaces

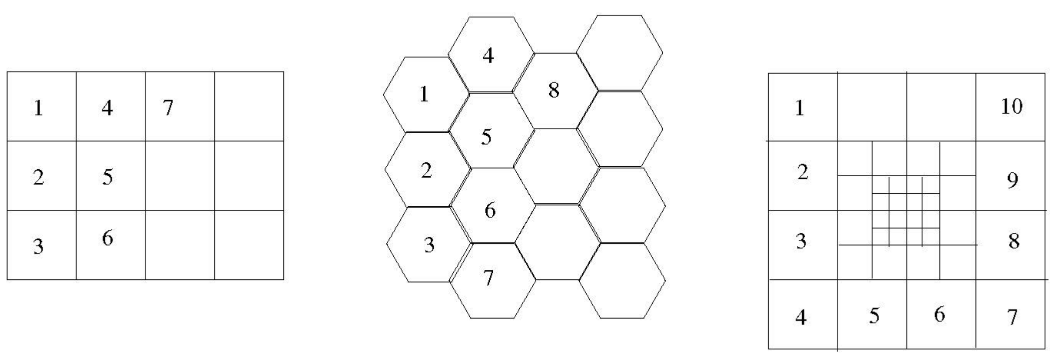
Today we are going to introduce the "*Playground*" that Linear Algebra exists in along with a set of usual suspects that play in this arena. If we were doing a more theoretical linear algebra, we would do lots of proofs in this section. For our purposes, treat this mainly as a "Who's Who" and learning about what two operations we can use in our "playground." Seriously though, there is a lot of nice theory here which provides the underpinning for all we do later this semester.

2.1 Review of Lab 1

In Lab 1, we noticed that if you add two images, you get a new image. We also saw that if you multiply an image by a scalar, you get a new image. We also saw in Lab that (rectangular pixelated) images can be represented as a rectangular array of values or equivalently as a rectangular array of grayscale patches. This is a very natural idea especially since the advent of digital photography. It is tempting to consider an image (or image data) as a matrix – after all, it certainly looks like one. Recall, we defined an image in the following way:

Definition 2.1. *An* **image** *is a finite ordered list of real values with an associated geometric array description.*

We also defined pixel-wise addition and scalar multiplication of images.
Three examples of arrays along with an index system specifying the order of patches can be seen in Figure below. Each patch would also have a numerical value indicating the brightness of the patch (not shown). The first is a regular pixel array commonly used for digital photography. The second is a hexagon pattern which also nicely tiles a plane. The third is a square pixel set with enhanced resolution toward the center of the field of interest. The key point here is that only the first example can be written as a matrix, but all satisfy the definition of image. We found that these operations had a place in a real-life example and when applying these operations it is very important that our result is still an image in the same configuration. It turns out that these operations are important enough that we give a name to sets that have these operations with some pretty standard properties.



2.2 Vectors and Vector Spaces

The word vector may be familiar for many students who have taken Vector Calculus and/or Physics. In these courses, there is a very specific type of vector used, vectors in $\mathbb{R}^m$. That is, the word vector may bring to mind something that looks like $\langle a, b\rangle$, $\langle a, b, c\rangle$, or $\langle a_1, a_2, \ldots, a_n\rangle$. Maybe you've even seen things like any of the following

$$(a, b), \quad (a, b, c), \quad (a_1, a_2, \ldots, a_n), \quad \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

called vectors. In this section, we first define the type of set we want to consider when talking about linear algebra concepts. Then, we discuss the elements of these sets called **vectors.**

Definition 2.2. *A set V with a set of scalars and operations $+$ and scalar multiplication $\cdot$ is called a* **vector space** *if the following ten properties hold. Elements of V are called* **vectors**. *Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ be vectors and α, β be scalars (constants).*

1. V is closed under addition $+$: $\mathbf{u} + \mathbf{v} \in V$ (Adding two vectors gives a vector in the set.)

2. V is closed under scalar multiplication $\alpha \cdot \mathbf{u} \in V$. (Multiplying a vector by a scalar gives a vector in the set.)

   Question: Does the set of odd real numbers satisfy the properties above? How about the set of even real numbers?

3. Addition is commutative: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

4. Addition is associative: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.

5. Scalar multiplication is associative: $\alpha \cdot (\beta \cdot \mathbf{v}) = (\alpha \cdot \beta)\mathbf{v}$.

6. Scalar multiplication distributes over vector addition: $\alpha \cdot (\mathbf{u} + \mathbf{v}) = \alpha \cdot \mathbf{u} + \alpha \cdot \mathbf{v}$.

7. Vector multiplication distributes over scalar addition: $(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \beta \cdot \mathbf{v}$.

8. V contains the $\mathbf{0}$ vector, where $\mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = v$.
   Note: $\mathbf{0}$ is the additive identity.

9. V has additive inverses $-\mathbf{v}$: $\mathbf{v} + -\mathbf{v} = \mathbf{0}$.

10. The scalar set has an identity element 1 for scalar multiplication: $1 \cdot \mathbf{v} = \mathbf{v}$ for all $\mathbf{v} \in V$.

Careful: It is important to note that the identity element for scalar multiplication need not be the number 1 and the zero vector need not be (and in general is not) the number 0 or a vector of 0's.

Notice also that elements of a vector space are called **vectors**. These need not look like the vectors presented above.
We now present some examples of vector space and the corresponding vectors arguments.

2.2.1 Common Vector Space Example 1: $\mathbb{R}^n$:

$\mathbb{R}^n$: the set of points or vectors in n-dimensional space is a vector space with scalars taken from the set of real numbers. Here, we recognize that addition of two vectors in $\mathbb{R}^n$ is component-wise. Here the vectors are just like the vectors we knew before Linear Algebra.
Examples: $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\begin{pmatrix} \pi \\ \sqrt{2} \end{pmatrix}$, and many more.

Example 2.1. $\mathbb{R}$, *the set of real numbers is a vector space with scalars taken from the set of real numbers. We define addition and multiplication as usual.*

Example 2.2. *We will show that* $\mathbb{R}^2$ *is a vector space and recognize how the same proofs generalize. Let* $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \mathbb{R}^2$ *be vectors and* α, β *be scalars.*

- *V is closed under addition $+$:* $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix} \in \mathbb{R}^2.$

- *V is closed under scalar multiplication $\cdot$:* $\alpha \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \end{pmatrix} \in \mathbb{R}^2.$

- *Addition is commutative:*
  $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix} = \begin{pmatrix} v_1 + u_1 \\ v_2 + u_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$

- *Addition is associative:*
  $\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right) + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} (u_1 + v_1) + w_1 \\ (u_2 + v_2) + w_2 \end{pmatrix} =$
  $\begin{pmatrix} u_1 + (v_1 + w_1) \\ u_2 + (v_2 + w_2) \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \left(\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\right).$

- *Scalar multiplication is associative:*
  $\alpha \cdot \left(\beta \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right) = \alpha \begin{pmatrix} \beta v_1 \\ \beta v_2 \end{pmatrix} = \begin{pmatrix} \alpha(\beta v_1) \\ \alpha(\beta v_2) \end{pmatrix} = \begin{pmatrix} (\alpha\beta) v_1 \\ (\alpha\beta) v_2 \end{pmatrix} = (\alpha\beta) \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$

- *Scalar multiplication distributes over vector addition:* $\alpha \cdot \left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right) = \alpha \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix} = \begin{pmatrix} \alpha(u_1 + v_1) \\ \alpha(u_2 + v_2) \end{pmatrix} = \begin{pmatrix} \alpha u_1 + \alpha v_1 \\ \alpha u_2 + \alpha v_2 \end{pmatrix} = \begin{pmatrix} \alpha u_1 \\ \alpha u_2 \end{pmatrix} + \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \end{pmatrix} = \alpha \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \alpha \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

- *Scalar multiplication distributes over scalar addition:* $(\alpha+\beta) \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} (\alpha+\beta)v_1 \\ (\alpha+\beta)v_2 \end{pmatrix} = \begin{pmatrix} \alpha v_1 + \beta v_1 \\ \alpha v_2 + \beta v_2 \end{pmatrix} = \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \end{pmatrix} + \begin{pmatrix} \beta v_1 \\ \beta v_2 \end{pmatrix} = \alpha \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \beta \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

- *V contains the 0 vector, where* $\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 + v_1 \\ 0 + v_2 \end{pmatrix} = \begin{pmatrix} v_1 + 0 \\ v_2 + 0 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

- *V has additive inverses $-v$:* $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} -v_1 \\ -v_2 \end{pmatrix} = \begin{pmatrix} v_1 + (-v_1) \\ v_2 + (-v_2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

- *The scalar set has an identity element 1 for scalar multiplication:* $1 \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1(v_1) \\ 1(v_2) \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$.

Key: If I want to show a set is **not** a vector space, I just need to show that ________ of the 10 properties fails.

Example 2.3. *Which property of vector spaces is **not** true for the following subset of $\mathbb{R}^2$?*

$$V = \{ \begin{bmatrix} -2 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \}$$

(a) Closure under vector addition
(b) Existence of an additive identity
(c) Existence of an additive inverse for each vector
(d) None of the above

Question: What would I have to add to the set to make it a vector space?

2.2.2 Common Vector Space Example 2: Matrices $\mathcal{M}_{m\times n}$:

Example 2.4. $\mathcal{M}_{2\times 3} = \left\{ \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \mid a,b,c,d,e,f \in \mathbb{R} \right\}$ *is a vector space when addition and scalar multiplication are defined as usual with Matrix operations and scalars are taken from* $\mathbb{R}$.

Closure under $+$: $\begin{pmatrix} a_1 & b_1 & c_1 \\ d_1 & e_1 & f_1 \end{pmatrix} + \begin{pmatrix} a_2 & b_2 & c_2 \\ d_2 & e_2 & f_2 \end{pmatrix} =$

Closure under $\cdot$: $\alpha \cdot \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} =$

Because matrix properties have the same structure as vectors in $\mathbb{R}^n$, all Vector Properties for Vector Spaces hold.

Additive Identity:

Additive Inverse for arbitrary vector $\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$:

Scalar Identity:

2.2.3 Common Vector Space Example 3: Polynomial Spaces $\mathcal{P}_n$:

Example 2.5. $\mathcal{P}_n = \{a_0 + a_1x + a_2x^2 + a_3x^3 + ... + a_nx^n \mid a_0, a_1, a_2, ..., a_n \in \mathbb{R}\}$ *is a vector space with scalars taken from* $\mathbb{R}$ *and addition and scalar multiplication defined in the standard way for polynomials.*

For example, $\mathcal{P}_2 = \{ax^2 + bx + c \mid a,b,c \in \mathbb{R}\}$

Example vectors in $\mathcal{P}_2$:

Is $x^3 \in \mathcal{P}_2$?

Is $x^3 \in \mathcal{P}_3$?

Why are Polynomial Spaces vector spaces? (We will show $\mathcal{P}_2$ is a vector space)

Closure under $+$:

Closure under $\cdot$:

Vector Properties: Yes because we can add polynomials component-wise, we can represent $ax^2 + bx + c$ as a vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ so all vector properties hold.

Additive Identity:

Additive Inverse for arbitrary vector $ax^2 + bx + c$:

Scalar Identity:

2.2.4 Common Vector Space Example 4: Function Spaces $\mathcal{F}$:

$\mathcal{F} = \{f : \mathbb{R} \to \mathbb{R}\}$: the set of all functions whose domain is $\mathbb{R}$ and whose range is a subset of $\mathbb{R}$. $\mathcal{F}$ is a vector space with scalars taken from $\mathbb{R}$. We can define addition and scalar multiplication in the standard way. Note: Here, the vectors are **functions**.

Examples vectors in this Vector Space:

How we define our two operations: **Addition**: $f + g = (f + g)(x) = f(x) + g(x)$
Scalar Multiplication: $\alpha f = (\alpha f)(x) = \alpha \cdot f(x)$
Check it is a Vector Space: Let $f, g, h \in \mathcal{F}$ and $\alpha, \beta \in \mathbb{R}$ then

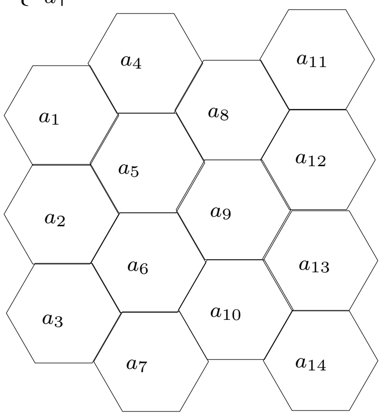
- $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$. Based on the definition of addition, $f + g : \mathbb{R} \to \mathbb{R}$. So $\mathcal{F}$ is closed over addition.
- Similarly, $\mathcal{F}$ is closed under scalar multiplication.
- Addition is commutative: $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$. So, $f + g = g + f$.
- Addition is associative: $((f + g) + h)(x) = (f + g)(x) + h(x) = (f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) = f(x) + (g + h)(x) = (f + (g + h))(x)$. So $(f + g) + h = f + (g + h)$.
- Scalar multiplication is associative: $(\alpha \cdot (\beta \cdot f))(x) = (\alpha \cdot (\beta f(x))) = (\alpha\beta) f(x) = ((\alpha\beta) \cdot f)(x)$. So $\alpha \cdot (\beta \cdot f) = (\alpha\beta) \cdot f$.
- Scalar multiplication distributes over vector addition: $(\alpha \cdot (f + g))(x) = \alpha \cdot (f + g)(x) = \alpha \cdot (f(x) + g(x)) = \alpha \cdot f(x) + \alpha \cdot g(x) = (\alpha \cdot f + \alpha \cdot g)(x)$. So $\alpha(f + g) = \alpha f + \alpha g$.
- Scalar multiplication distributes over scalar addition: $((\alpha + \beta) \cdot f)(x) = (\alpha + \beta) \cdot f(x) = \alpha \cdot f(x) + \beta \cdot f(x) = (\alpha \cdot f + \beta \cdot f)(x)$. So, $(\alpha + \beta) \cdot f = \alpha \cdot f + \beta \cdot f$.
- $\mathcal{F}$ contains the constant function defined by $z(x) = 0$ for every $x \in \mathbb{R}$. And, $(z + f)(x) = z(x) + f(x) = 0 + f(x) = f(x) = f(x) + 0 = f(x) + z(x) = (f + z)(x)$. That is, $z + f = f + z = f$. So, the 0 vector is in $\mathcal{F}$.
- $\mathcal{F}$ has additive inverses $-f$ defined to as $(-f)(x) = -f(x)$ and $(f + (-f))(x) = f(x) + (-f(x)) = 0 = z(x)$, where z is defined in part 2.2.4. So, $f + (-f) = z$.
- The real number 1 satisfies: $(1 \cdot f)(x) = 1 \cdot f(x) = f(x)$. So, $1 \cdot f = f$.

2.2.5 Common Vector Space Example 5: Space of Greyscale Images:

Notice that the set of images of a with a specified geometric arrangement is a vector space with scalars taken from $\mathbb{R}$. That is, if we consider the set

$V = \{I_a |\ I$ is of the form below and $a_1, a_2, \ldots a_{14} \in \mathbb{R}\}$.

$I_a =$



We know this is a vector space since, by definition of addition and scalar multiplication on images, we see that both closure properties hold. Notice that there is a corresponding real number to each pixel (or voxel). Because addition and scalar multiplication are taken pixel-wise (or voxel-wise), we can see that these 10 properties hold within each pixel (or voxel). So, we know all 10 properties hold (just like they did in the last example2.2).

Note: It turns out that given any geometric configuration our definitions of image and operations on images guarantee that the space of images with the chosen configuration is a vector space. The vectors are then images.

Example 2.6. *Is the following space* $V = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$ *a Vector Space?*

(a) Yes, and I am very confident
(b) Yes, but I am not very confident
(c) No, but I am not very confident
(d) No, and I am very confident

2.2.6 Common Vector Space Example 6: Homogeneous System of Equations:

Example 2.7. *Determine whether or not V is a vector space. Verify your answer.*

$$V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, x + y + z = 1, 2x + 2y + 2z = 2, \textit{ and } - x - y - z = -1 \right\}.$$

Notice when we change the system to a homogeneous system we get a vector space.

Example 2.8. *Determine whether or not V is a vector space.*

$$V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,\middle|\, x + y + z = 0, 2x + 2y + 2z = 0, \textit{ and } - x - y - z = 0 \right\}.$$

2.2.7 Vector Space Example 7: Heat States in Diffusions and other models:

A manufacturing company uses a process called diffusion welding to adjoin several smaller rods into a single longer rod. The diffusion welding process leaves the final rod heated to various temperatures along the rod with the ends of the rod having the same temperature. Every $a$cm along the rod, a machine records the temperature difference from the temperature at the ends to get an array of temperatures called a *heat state*.

1. Plot the heat state given below (let the horizontal axis represent distance from the left end of the rod and the vertical axis represent the temperature difference from the ends).

$$u = (0, 1, 13, 14, 12, 5, -2, -11, -3, 1, 10, 11, 9, 7, 0)$$

2. How long is the rod represented by u, the above heat state, if $a = 1$cm?

3. Give another example of a heat state for the same rod, sampled in the same locations.

4. Notice that the set of all heat states, for this rod, is a vector space.

5. What do the vectors in this vector space look like? That is, what is the key property (or what are the key properties) that make these vectors stand out?

2.2.8 Vector Space Example 8: Galois Field 2: $GF(2)^n$:

See HW 4.

2.3 Span and Subspace Motivation

Span: It is now natural to ask the question: What images can be created as linear combinations of a given set of images? We say that image z is in the **span** of a set of images $I = \{x, y, ...\}$ if z can be written as a linear combination of images in I. We write $z \in$Span(I). We also say that Span(I) is the set of all linear combinations of the images in I.

Example 2.9. *Image 3 from Lab 1 was in the span of Images A, B and C since Image 3= a*Image A+b*Image B+c*Image C. But Image 4 is not in Span(Image A, Image B, Image C) since we could not write it as a linear combination of Images A, B and C.*

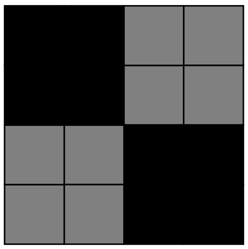


Image A

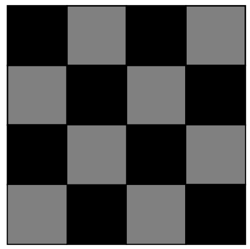


Image B

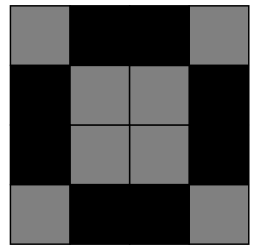


Image C

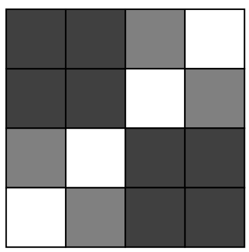


Image 3

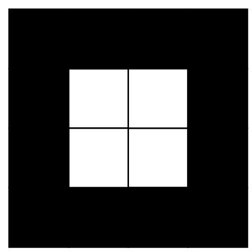


Image 4

Is Image 3$\in$ span {Image A, Image B, Image C}?

Subspace: Because Span(I) is determined by the same operations as closure of vector spaces, we might ask if Span(I) is a vector space of images. Indeed it is, and it is also a subset of the set of all images (of the same geometry). We say that W is a subspace of V if it is nonempty and is closed under image addition and scalar multiplication.

Notation: $x \in V$ means x is an element of V.

Example: If $V = \mathbb{R}$, then $3 \in V$.

$U \subseteq V$ or $U \subset V$ means U is a subset of V.

Example: If $U = \{$ even numbers $\}$, then $U \subseteq \mathbb{R}$.

2.4 Span

In this class, we will use the word *Span* as **both** a __________ and __________. Let us begin with the definition of the *noun* span.

Definition 2.3. *(n.) Let V be a vector space and let $X = \{v_1, v_2, \ldots, v_n\} \subset V$. Then the* **span** *of the set X is the set of all linear combinations of the elements of X. That is,*

$$\text{span}\, X = \text{span}\, \{v_1, v_2, \ldots, v_n\} = \{\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n |\ \alpha_1, \alpha_2, \ldots, \alpha_n \text{ are scalars}\}.$$

Why do we care about "Span"?

- Allows us to encode a Vector Space *compactly* into its building blocks (uses bases -coming soon!)
- Helps us to represent the set of possible outputs of a Vector Space.
- Allows us to determine if subsets of vectors produce the same set of outputs.
- Helps us determine if a vector is in a Vector Space.

Method: How to show a vector $\mathbf{u}$ is in the span $\{v_1, v_2, ...v_n\}$:
Try to find scalars/coefficients α_i such that $\mathbf{u} = \alpha_1 \mathbf{v_1} + \alpha_2 \mathbf{v_2} + \alpha_3 \mathbf{v_3} + \cdots + \alpha_n \mathbf{v_n}$.

Example 2.10. *Let $v_1 = 3x+4, v_2 = 2x+1, v_3 = x^2+2,$ and $v_4 = x^2$. Is $v_1 \in \text{span}\, \{v_2, v_3, v_4\}$?*

Example 2.11. *Find the span of the two polynomials* x *and* 1 *in* $\mathcal{P}_1$*, denoted span* $\{x, 1\}$*. Recall* $\mathcal{P}_1 = \{ax + b|\ a, b \in \mathbb{R}\}$*. Also determine span* $\{x, 2\}$*. Is span* $\{x, 2\}$ $=$ *span* $\{x, 1\}$*?*

Note: This example is interesting because it shows two different ways to write the same set as a span.

2.4.1 ICE 2a: Span: Part 1

Eva and Archer are going on a trip for the first time. Dr. Harsy wants to help them on their journey so she gives them two gifts. Specifically, she gives them two forms of transportation: a hover board and a magic carpet. Dr. Harsy informs them that both the hover board and the magic carpet have restrictions in how they operate:

We denote the restriction on the hover board's movement by the vector $\begin{bmatrix}3\\1\end{bmatrix}$.

By this we mean that if the hover board traveled "forward" for one hour, it would move along a "diagonal" path that would result in a displacement of 3 miles East and 1 mile North of its starting location.

We denote the restriction on the magic carpet's movement by the vector $\begin{bmatrix}1\\2\end{bmatrix}$.

By this we mean that if the magic carpet traveled "forward" for one hour, it would move along a "diagonal" path that would result in a displacement of 1 mile East and 2 miles North of its starting location.

Scenario One: The Maiden Voyage
Eva and Archer's first adventure is to go visit their cousin Mako who lives in a cabin that is 107 miles East and 64 miles North of their home.

Task 1a:
Investigate whether or not Eva and Archer can use the hover board and the magic carpet to get to Mako's cabin. If so, how? If it is not possible to get to the cabin with these modes of transportation, why is that the case?

Task 1b: What is another "linear algebra" way to ask the question in Task 1?

Task 2a: Suppose the magic carpet is acting up so Eva and Archer can only use their hoverboard. Can they get to Mako's cabin now? Why or why not?

Task 2b: Since they can't use their magic carpet, Eva and Archer need to borrow a magic carpet from their friend's Habañero (Dr. Haven's cat) to get to Mako's cabin. She has many different carpets which movement restrictions given by $c_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $c_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$, $c_3 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, and $c_4 = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$. Archer doesn't want to do much work calculating this answer. Eva says that Archer could probably answer this in his head. Which carpet should he borrow? Check your answer if you have time.

2.5 Spanning Sets

Definition 2.4. *(v.) We say that the set of vectors* $\{v_1, v_2, \ldots, v_n\}$ ***spans*** *a set* X *if* $X = span\ \{v_1, v_2, \ldots, v_n\}$. *In this case, we call the set* $\{v_1, v_2, \ldots, v_n\}$ **a spanning set** *of* X.

Example 2.12. *True or False: Spanning Sets are unique.*

(a) True, and I am very confident
(b) True, but I am not very confident
(c) False, but I am not very confident
(d) False, and I am very confident

Method: How to show a subset of vectors $\{v_1, v_2, ...v_n\}$ spans (is a spanning set for) a Vector Space W:

This means we want to show we can produce any vector $\mathbf{w}$ in W using a linear combination of vectors in the set $\{v_1, v_2, ...v_n\} \subseteq W$.
The problem is that W often has infinitely many vectors in it. How can we check them all?

Solution: Pick an arbitrary vector in W and see if we can write the arbitrary vector as a linear combination of $v_1, v_2, ...v_n$ Try to find scalars/coefficients (α_i's) such that
$\mathbf{u} = \alpha_1 \mathbf{v_1} + \alpha_2 \mathbf{v_2} + \alpha_3 \mathbf{v_3} + \cdots + \alpha_n \mathbf{v_n}$.

What is an arbitrary vector representative for $\mathbb{R}^3$?

What is an arbitrary vector representative for $\mathcal{P}_3$?

Note spanning sets need to be a subset of X. For example $D = \{x, 1, x^2\}$ cannot be a spanning set for $\mathcal{P}_1 = \{ax + b | a, b \in \mathbb{R}\}$. Why?

Example 2.13. *Which of the following sets are spanning sets for* $\mathcal{P}_1 = \{ax + b | a, b \in \mathbb{R}\}$*:*

$A = \{200, 1\}$

$B = \{x\}$

$C = \{x, x+1\}$

$D = \{x+1, 2x+2\}$

$E = \{x, 1, 3x+4\}$

Example 2.14. *Show that* $\mathbb{R}^2$ *is spanned by both* $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$ *and* $\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \end{pmatrix}\right\}$.

Notice that in each of these examples, we found that the span of a set turned out to be a vector space. It turns out that this is always true.
Because of this, we now have a new way to determine if a set is a vectors space -check if it can be written as a span.

Method: How to write a set of vectors $\{v_1, v_2, ...v_m\}$ as a span (if you can):

Ultimate Goal: Be able to write the set $S = \{v_1, v_2, ...v_m\}$ in the form
$\{\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n | \ \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}\} = \text{span } \{v_1, v_2, ..., v_n\}$
*Note there are NO restrictions on the constants!

For example, $\mathcal{P}_1 = \{ax + b | a, b \in \mathbb{R}\} = \text{span } \{1, x\}$.

But $S = \{v_1 + av_2 | a \in \mathbb{R}\}$ has the restriction that the constant in front of v_1 must be 1. Thus, $S = \{av_1 + bv_2 | a = 1, b \in \mathbb{R}\}$. So is S a vector space?

Example 2.15. *Which of the following sets are vector spaces?*

$T = \{5 + av_1 | a \in \mathbb{R}\}$

$V = \{a(v_1 + v_2) | a \in \mathbb{R}\}$

Example 2.16. *Show that* $\{a_0 + a_1 x | a_0 + a_1 = 0\}$ *is a vector space by writing it as a span of vectors.*

Here, we summarize the terminology:

We say $\{v_1, v_2, \ldots, v_n\}$ **spans a set** V if V is the span of $\{v_1, v_2, \ldots, v_n\}$.

We say $\{v_1, v_2, \ldots, v_n\}$ is a **spanning set** for V if V is the span of $\{v_1, v_2, \ldots, v_n\}$.

We say V is **spanned by** $\{v_1, v_2, \ldots, v_n\}$ if V is the span of $\{v_1, v_2, \ldots, v_n\}$.

Finally, all of these mean that if $v \in V$ then there are scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ so that $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots \alpha_n v_n$.

Example 2.17. *How do you describe the span of the vectors* $(1, 0, 0)$*,* $(0, 1, 0)$*,* $(0, 0, 1)$*?*
(a) A point
(b) A line segment
(c) A line
(d) $\mathbb{R}^2$
(e) $\mathbb{R}^3$

Example 2.18. *How do you describe the set of all linear combinations of the vectors* $(1, 0, 0)$*,* $(0, 1, 0)$*?*
(a) A point
(b) A line segment
(c) A line
(d) $\mathbb{R}^2$
(e) $\mathbb{R}^3$

Example 2.19. *How do you describe the set of all linear combinations of the vector* $(1, 0, 0)$*?*
(a) A point
(b) A line segment
(c) A line
(d) $\mathbb{R}^2$
(e) $\mathbb{R}^3$

2.5.1 ICE 2b: Span: Part 2

Scenario two: Getting Back Home

Suppose Eva and Archer are now in a three-dimensional world for the carpet ride problem, and they have three models of transportation: $v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $v_2 = \begin{bmatrix} 6 \\ 3 \\ 8 \end{bmatrix}$, $v_3 = \begin{bmatrix} 4 \\ 1 \\ 6 \end{bmatrix}$.

Task 1:[2] Eva and Archer are only allowed to use each mode of transportation **once** (in the forward or backward direction) for a fixed amount of time (c_1 on v_1, c_2 on v_2, c_3 on v_3). Find the amounts of time on each mode of transportation (c_1, c_2, and c_3, respectively) needed to go on a journey that starts **and ends at home** OR explain why it is not possible to do so.

Hint 1: We don't know where Eva and Archer's house is (and Dr. Harsy is not sure she wants to share this with her class), so what could we use to represent the location of their house?

Hint 2: Set up an equation you want to solve with these vectors. And see if you can find values for c_1, c_2, and c_3.

[2] 2 sided!

Bonus Task 1: What equation do we really solve in Task 1?

Task 2: Is there more than one way to make a journey that meets the requirements described above? (In other words, are there different combinations of times you can spend on the modes of transportation so that you can get back home?) If so, how? (use the same vectors from the first page.[3])

[3] $v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 6 \\ 3 \\ 8 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 4 \\ 1 \\ 6 \end{bmatrix}.$

2.6 Subspaces

Many times, we work in subsets of vector spaces.
Definition: Let V be a vector space. If $W \subseteq V$ is a ***subspace*** of V if it satisfies the following three properties.

1. __________ $\in W$
   (We don't allow W to be empty and so we require that it contains the $\mathbf{0}$ vector.)
2. Let $\mathbf{u} \in W$ and α be a scalar. Then __________ $\in W$.
   The Closure Property under scalar multiplication.
3. Let $\mathbf{u},\mathbf{v} \in W$ Then __________ $\in W$.
   The Closure Property under vector addition.

Note: You can combine 2 & 3 by checking whether the space is *closed under linear combinations:*
Let $u, v \in W$ and α, β be scalars. Then $\alpha u + \beta v \in W$.

1. $\mathbf{0} \in W$
   (We don't allow W to be empty and so we require that it contains the 0 vector.)
2. If $u, v \in W$ and α, β are scalars, then __________ $+$ __________ $\in W$

Heuristic Verification: Why Only These 2 Properties Are Necessary To Check: Notice that if W is a subspace of V, then W is a vector space as well. The Vector Properties are *inherited* from V since V is like a parent set to W. The scalar 1 still exists in the scalar set also. The only issue that may happen comes with the Closure Properties. This means that we need only show closure under addition and scalar multiplication and that $\mathbf{0} \in W$. $\square$

Theorem: Let V be a vector space and let $v_1, v_2, \ldots, v_n \in V$. Then span $\{v_1, v_2, \ldots, v_n\}$ is a subspace of V.

Method: How to show a subset of vectors, W, of a Vector Space V is a subspace: There are 2 methods:

First Method:
Show that both
1) $\mathbf{0} \in W$ and
2) Pick arbitrary vector $\mathbf{u} \in W$ & scalar α and show that $\alpha\mathbf{u} \in W$.
3) Pick arbitrary vectors $\mathbf{v},\mathbf{u} \in W$ and show that $\mathbf{u} + \mathbf{v} \in W$.

Second Method:
Write the set as a span.
That is, show that
$W = \{a_1v_1 + a_2v_2 + ... + a_nv_n | a_i \in \mathbb{R}\}$
$\Rightarrow W = \text{span}\ \{v_1, v_2, ..., v_n\}$

*Note that in both methods, the subset must follow the normal rules of the bigger space.

Example 2.20. *Let* $W = \{(x, y) \in \mathbb{R}^2 | -2x + 7y = 9, x - 4y = -2\}$.
Is W *a subspace of* $\mathbb{R}^2$*?*

Note: A line in $\mathbb{R}^2$ that does NOT go through the origin is not a subspace of $\mathbb{R}^2$. Similarly a hyper-plane in $\mathbb{R}^n$ that doesn't contain the zero vector is not a subspace.

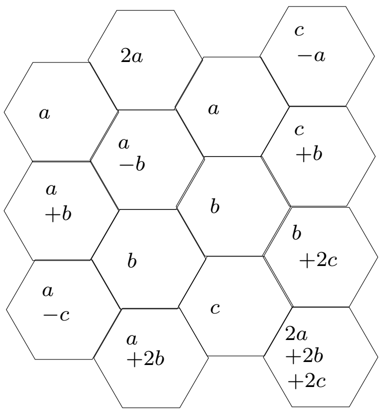
Example 2.21. *Is* $\mathbb{R}^2$ *a subspace of* $\mathbb{R}^3$*?*

Example 2.22. *Is* $W = \{a \begin{bmatrix} 2 \\ 1 \end{bmatrix} |a \in \mathbb{R}\}$ *a subspace of* $\mathbb{R}^2$*?*

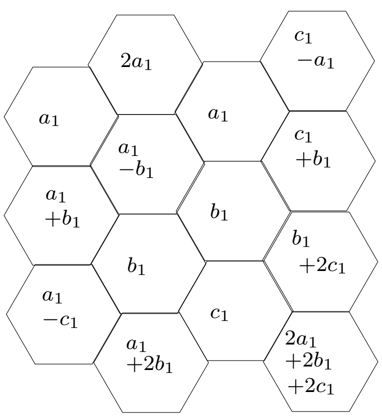
Example 2.23. *Let* $V = \{(x, y, z) \in \mathbb{R}^3 |\ x + y + z = 0, 3x + 3y + 3z = 0\}$. *Write* V *as a span to show that* V *is a subspace of* $\mathbb{R}^3$.

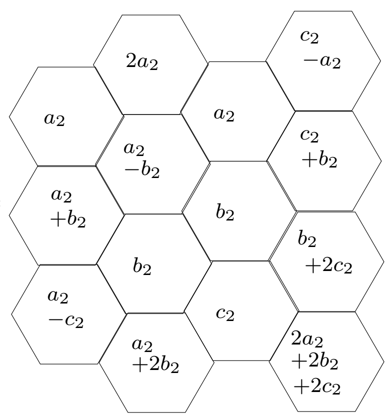
Reading Example: Please read on your own time.

Example 2.24. *Consider the set of images $V = \{I \mid I$ is of the form below and $a, b, c \in \mathbb{R}\}$.*

$I =$ 

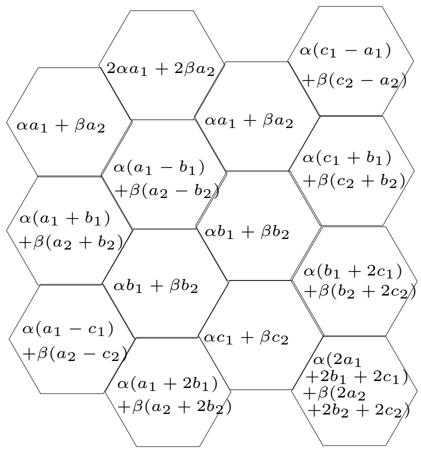
We can show that V is a subspace of images with the same geometric configuration. We showed above that the set of these images is a vector space, so we need only show the two subspace properties. First, notice that the 0 image is the image with $a = b = c = 0$ and this image is in V. Now, we need to show that linear combinations are still in the set V. Let $\alpha, \beta \in \mathbb{R}$, be scalars and let $I_1, I_2 \in V$, then there are real numbers $a_1, b_1, c_1, a_2, b_2,$ and c_2 so that

$I_1 =$ 

$I_2 =$ 

Notice that $\alpha I_1 + \beta I_2$ is also in V since... *(see image on next page)*

$\alpha I_1 + \beta I_2 =$



Notice that performing the operations inside each pixel shows that we can write $\alpha I_1 + \beta I_2$ in the same form as I above. That is, $\alpha I_1 + \beta I_2 \in V$. Thus V is a subspace of images that are laid out in the geometric form above. Notice, this means that V is in itself a vector space.

2.6.1 ICE 2c: Span: Part 3 -Subspaces

Which of these subsets are subspaces of $\mathcal{M}_{2\times 2}$? For each one that is a subspace, write the set as a span. For each that is not, show the condition that fails.[4]!

$$A = \left\{ \left(\begin{array}{cc} a & 0 \\ 0 & b \end{array} \right) \middle| \; a, b \in \mathbb{R} \right\}$$

$$B = \left\{ \left(\begin{array}{cc} a & 0 \\ c & b \end{array} \right) \middle| \; a - 2b = 0, c \in \mathbb{R} \right\}$$

[4]Two Sided

$$C = \left\{ \left(\begin{array}{cc} a & 0 \\ c & 0 \end{array} \right) \middle| \; a + c = 1 \right\}$$

$$D = \left\{ \left(\begin{array}{cc} a & 1 \\ 0 & b \end{array} \right) \middle| \; a, b \in \mathbb{R} \right\}$$

2.7 Linear Independence

In the Radiography and Tomography Lab #1, you were asked to write one image using arithmetic operations on three others. You found that this was possible sometimes and not others. It becomes a useful tool to know whether or not this is possible.

Definition 2.5. *Let V be a vector space. We say that the set $\{v_1, v_2, \ldots, v_n\} \subset V$, is* **linearly independent** *if no element is in the span of the others. This means, we* ***can't*** *write this vector as a linear combination of the other vectors in the set.*

Check for Linear Independence: When $\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$ is true only when $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0.$, then the vectors $v_1, v_2, \ldots, v_n$ **are linearly independent**.
This check is called the **linear dependence relation.**

Definition 2.6. *If $\{v_1, v_2, \ldots, v_n\}$ is not linearly independent, then we say that it is* **linearly dependent***. Note this means that we can write one element as a linear combination of the others.*

Here is the process to check whether or not $\{v_1, v_2, ...v_n\}$ is linear independent:
Step 1: Set an arbitrary linear combination of v_i's equal to 0:
$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = 0$
Step 2: See if you can find values for the α_i's that are not all 0. If you can, the set is dependent. If $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$ is the *only* possible values for the α's, then the set is linear independent.

Careful: Of course, one solution for the linear dependence relation will always be $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$, but this tells us nothing about the linear dependence of the set. That is, just because $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$ works, it doesn't mean the set is linear independent. You must show that it is the __________ solution!

Demonstration: If the set $\{v_1, v_2\}$ is linearly dependent, then we can find α_1 and α_2, BOTH not zero, such that $\alpha_1 v_1 + \alpha_2 v_2 = 0$.

Suppose α_1 is not zero.
Set $\alpha_1 v_1 + \alpha_2 v_2 = 0$.
Since $\alpha_1 \neq 0$, we can solve for v_1:
$v_1 = -\frac{\alpha_2}{\alpha_1} v_2$.
This shows that, v_1 is a scalar multiple of v_2
which means v_1 can be built from v_2
aka v_1 is *dependent* of v_2.

Example 2.25. *The set* $\{0, v_1, v_2, \ldots, v_n\}$ *is always linearly dependent.*

Thus, it if we are considering the linear dependence a two element set, we need only check whether one can be written as a scalar multiple of the other.

Example 2.26. *Determine whether* $\{x^2+x, x^2, 1\} \subseteq \mathcal{P}_2$ *is linearly dependent or independent.*

Notice that whenever we are determining whether a set is linearly independent or dependent, we always start with the linear dependence relation and determine whether or not there is only one set of scalars that (when they are all zero) to make the linear dependence relation true.

Example 2.27. *Determine the linear dependence of the set* $\left\{ \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \right\}$.

(To See Row Reduction)[5]

5

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 3 & 1 & 2 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 0 \end{array}\right) \xrightarrow[R_3=-r_1+r_3, R_4=-r_1+r_4]{R_2=-3r_1+r_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & -1 & 0 \end{array}\right)$$

$$\xrightarrow[R_4=-r_2+r_4]{R_2=\frac{1}{-2}r_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{R_1=-r_2+r_1} \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Example 2.28. *Let* $v_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, *and* $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^3$. *Can we write* $Span(\{v_1, v_2, v_3\})$ *as a span of 2 vectors? 1 vector?*

2.8 Basis

Idea: Describe a vector space in terms a much smaller subset. This set can be considered building blocks which we can use to build any vector from the vector space.
Recall, a vector space can be written as a **span of vectors**. Sometimes when we compute the span of a set of vectors, we see that it can be written as a span of a smaller set of vectors. This means that the vector space can be described with the smaller set of vectors. This happens when the larger set of vectors is linearly dependent.

Main Goal: Find a spanning set that is big enough to describe *all* of the elements of our vector space, but not so big that there's repetition.

Definition 2.7. *Given a vector space V, we call $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ the* **basis** *of V if and only if $\mathcal{B}$ satisfies the following conditions:*

1. *span* $(\mathcal{B}) = $ __________ *and*
2. *$\mathcal{B}$ is linearly independent.*

Think of it this way, we are like Goldie Locks, we want our set to be
1) Big Enough so
2) Not to Big that we have

Example 2.29. *The standard basis for $\mathbb{R}^3$ is* $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Notation: Because the standard basis is used often, we introduce notation for each of the vectors. We let e_1, e_2, and e_3 denote the three vectors in $\mathcal{S}$, where

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

In general, for $\mathbb{R}^n$, the standard basis is $\{e_1, e_2, \ldots, e_n\}$, where e_i is the $n \times 1$ vector array with zeros in every entry except the ith entry which contains a one.

Example 2.30. *Show that another basis for* $\mathbb{R}^3$ *is* $\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Example 2.31. *The standard basis for* $\mathcal{P}_2$ *is* $\mathcal{S} = \{x^2, x, 1\}$. *Show* $\mathcal{S}$ *is a basis for* $\mathcal{P}_2$. *Recall,* $\mathcal{P}_2 = \{ax^2 + bx + c \cdot 1 | a, b, c \in \mathbb{R}\} = \text{span}\,\{x^2, x, 1\}$.

Example 2.32. $\mathcal{B} = \{1, x + x^2, x^2\}$ *is also a basis for* $\mathcal{P}_2$. *In Exercise2.26, we showed that* $\mathcal{B}$ *is linearly independent. So, we need only show that span* $\mathcal{B} = \mathcal{P}_2$. *again, we need only show that* $\mathcal{P}_2 \subseteq \text{span}\,\mathcal{B}$ *since it is clear that span* $\mathcal{B} \subseteq \mathcal{P}_2$.

Notice in the last four examples, we see two examples where the bases of a vector space had the same number of elements. That is, both bases of $\mathbb{R}^3$ had the same number of elements and both bases for $\mathcal{P}_2$ had the same number of elements. One should wonder if this is always true. The answer is given in the next theorem.

Theorem 2.1. *Let V be a vector space with bases $\mathcal{B}_1 = \{v_1, v_2, \ldots, v_n\}$ and $\mathcal{B}_2 = \{u_1, u_2, \ldots, u_m\}$. Then the number of elements, n in $\mathcal{B}_1$ is the same as the number of elements, m in $\mathcal{B}_2$.*

That is, $n =$ __________.

Proof. Suppose both $\mathcal{B}_1$ and $\mathcal{B}_2$ are bases for V. We show that this is true by assuming it is not true and showing that this is an impossible scenario. That is, we will assume that $m \neq n$ and find a reason that this cannot be true. Suppose $m > n$ (a very similar argument can be made if we assumed $n > m$). Since both $\mathcal{B}_2$is a subset of V, we know that there exist $\alpha_{i,j}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$ so that

$$\begin{aligned}
u_1 =& \alpha_{1,1}v_1 + \alpha_{1,2}v_2 + \ldots + \alpha_{1,n}v_n \\
u_2 =& \alpha_{2,1}v_1 + \alpha_{2,2}v_2 + \ldots + \alpha_{2,n}v_n \\
& \vdots \\
u_m =& \alpha_{m,1}v_1 + \alpha_{m,2}v_2 + \ldots + \alpha_{m,n}v_n.
\end{aligned}$$

We want to show that $\mathcal{B}$ cannot be linearly independent (which would be impossible if it is a basis). Let

$$\beta_1 u_1 + \beta_2 u_2 + \ldots + \beta_m u_m = 0.$$

We will then find $\beta_1, \beta_2, \ldots, \beta_m$. Notice that if we replace $u_1, u_2, \ldots, u_n$ with the linear combinations above, we can rearrange to get

$$\begin{aligned}
&(\beta_1\alpha_{1,1} + \beta_2\alpha 2, 1 + \ldots + \beta_m\alpha_{m,1})v_1 \\
+&(\beta_1\alpha_{1,2} + \beta_2\alpha 2, 2 + \ldots + \beta_m\alpha_{m,2})v_2 \\
&\vdots \\
+&(\beta_1\alpha_{1,n} + \beta_2\alpha 2, n + \ldots + \beta_m\alpha_{m,n})v_n = 0.
\end{aligned}$$

Since $\mathcal{B}_1$ is a basis, we get that the coefficients of $v_1, v_2, \ldots, v_n$ are all zero. That is

$$\beta_1\alpha_{1,1} + \beta_2\alpha 2, 1 + \ldots + \beta_m\alpha_{m,1}) = 0 \quad (1)$$

$$\beta_1\alpha_{1,2} + \beta_2\alpha 2, 2 + \ldots + \beta_m\alpha_{m,2} = 0 \quad (2)$$

$$\vdots \quad (3)$$

$$\beta_1\alpha_{1,n} + \beta_2\alpha 2, n + \ldots + \beta_m\alpha_{m,n} = 0. \quad (4)$$

We know that this system has a solution because it is homogeneous. But, because there are more scalars $\beta_1, \beta_2, \ldots, \beta_m$ (that we are solving for) than there are equations, this system must have infinitely many solutions. This means that $\mathcal{B}_2$ cannot be linearly independent and so it cannot be a basis. Thus, the only way both $\mathcal{B}_1$ and $\mathcal{B}_2$ can be bases of the same vector space is if they have the same number of elements. $\square$

Example 2.33. *Is* $\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\5\\1\end{bmatrix}, \begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}1\\5\\2\end{bmatrix}\}$ *and basis for* $\mathbb{R}^3$*?*

a) Yes, and I am very confident!
b) Yes, but I am not very confident.
c) No, and I am very confident!
d) No, but I am not very confident.

Because the number of elements in a basis is unique to the vector space, we can give it a name.

Definition 2.8. *Given a vector space V whose basis $\mathcal{B}$ has n elements, we say that the* **dimension** *of V is n, the number of elements in $\mathcal{B}$ ($\dim V = n$.)*
Sometimes we say that V is n-dimensional.

Example 2.34. *We can see that the last 4 examples show that both $\mathbb{R}^3$ and $\mathcal{P}_2$ are*

a) 1 dimensional
b) 2 dimensional
c) 3 dimensional
d) 4 dimensional
e) 5 dimensional

Note: What we see is that, in order to find the dimension of a vector space, we need to find a basis and count the elements in the basis.

This can also help us immediately check whether or not a set of vectors is a basis.

Since $P_4 = \{ax^4 + bx^3 + cx^2 + dx + e | a, b, c, d, e \in \mathbb{R}\} = \text{span}\,\{x^4, x^3, x^2, x, 1\}$, it's dimension is __________ so we know any set that doesn't have exactly __________ vectors will not be a basis.

Bases are not unique, but often we want to the easiest basis (set of building blocks.)
What would be the easiest (standard) basis for $\mathcal{M}_{2\times 2} = \{\begin{bmatrix}a & b\\c & d\end{bmatrix} | a, b, c, d \in \mathbb{R}\}$?

Example 2.35. *Let* $V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \; x+y+z=0, 2x+y-4z=0, 3x+2y-3z=0 \right\}$. *Find the dimension of* V*. First, we need to rewrite* V *as a span.*

$$\begin{array}{rcrcrcl} x & + & y & + & z & = & 0 \\ 2x & + & y & - & 4z & = & 0 \\ 3x & + & 2y & - & 3z & = & 0 \end{array} \xrightarrow{\textit{use matrix}} \left(\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 2 & 1 & -4 & 0 \\ 3 & 2 & -3 & 0 \end{array}\right) \xrightarrow[R_3=-3r_1+r_3]{R_2=-2r_1+r_2} \left(\begin{array}{rrr|r} 1 & 1 & 1 & 0 \\ 0 & -1 & -6 & 0 \\ 0 & -1 & -6 & 0 \end{array}\right)$$

$$\xrightarrow[R_3=-r_2+r_3, R_2=-r_2]{R_1=r_2+r_1} \begin{pmatrix} 1 & 0 & -5 & 0 \\ 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow$$

2.8.1 ICE 3: Basis

2 sides!

1. True or False: Each vector space has a unique basis.
   (a) True, and I am super confident!
   (b) True, but I am not very confident.
   (c) False, and I am super confident!
   (d) False, but I am not very confident.

2. Which of the following describes a basis for a subspace V ?
   (a) A basis is a linearly independent spanning set for V .
   (b) A basis is a minimal spanning set for V .
   (c) A basis is a largest possible set of linearly independent vectors in V .
   (d) All of the above
   (e) Some of the above

3. Which of the following sets of vectors **forms a basis** for $\mathbb{R}^3$?

   i. $\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}$ ii. $\begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

   iii. $\begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 6 \\ 2 \\ -2 \end{bmatrix}$ iv. $\begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 8 \\ 12 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$

   (a) i, ii, iii, and iv
   (b) ii, iii, and iv only
   (c) ii and iii only
   (d) iii and iv only
   (e) ii only

4. Which of the following sets of vectors **spans** $\mathbb{R}^3$?

i. $\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}$ ii. $\begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$

iii. $\begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 6 \\ 2 \\ -2 \end{bmatrix}$ iv. $\begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 8 \\ 12 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$

(a) i, ii, iii, and iv
(b) ii, iii, and iv only
(c) ii and iii only
(d) iii and iv only
(e) ii only

5. Which of the following describes the subspace of $\mathbb{R}^3$ spanned by the vectors from : $\left\{ \begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 8 \\ 12 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix} \right\}$? (note this is set iv from the previous problem.)

(a) A line
(b) A plane
(c) $\mathbb{R}^2$
(d) All of $\mathbb{R}^3$
(e) Both (b) and (c)

3 Least Squares

Sometimes when we are problem solving, there is no solution to a system of equations. When this happens, we need to find an approximation for this system. There are a variety of techniques to do this, but one of the most common is using the method of *Least Squares.*

First we need to introduce some definitions.

3.1 Norm, Distance, Transpose, and Inner Products

Definition 3.1. *The **transpose** of A, A^T is a matrix whose columns are formed from the corresponding rows of A.*

That is, the rows of A become the columns of A^T.

Example 3.1. *Given a matrix* $A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$, $A^T =$

So if A is an $n \times m$ matrix, A^T is a __________.

Octave Code: In octave A^T can be written as A'

Definition 3.2. *A is a **symmetric matrix** if $A^T = A$.*

Definition 3.3. *The **inner product** or **dot product** between two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, denoted $\mathbf{u} \cdot \mathbf{v}$, is the product $\mathbf{u}^T \cdot \mathbf{v}$*

Let $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} =$

Example 3.2. *Determine $\mathbf{u} \cdot \mathbf{v}$, $\mathbf{v} \cdot \mathbf{u}$, and $\mathbf{v} \cdot \mathbf{v}$ for* $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$ *and* $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$

Definition 3.4. *The **length** or **norm** of a vector* $\boldsymbol{v} \in \mathbb{R}^n$ *is denoted as* $||\mathbf{v}||$ *and defined as* $||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$

Definition 3.5. *Given* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$*, the **distance between u and v**, denoted* $dist(\mathbf{u}, \mathbf{v})$ *is the length of the vector* $\mathbf{u} - \mathbf{v} = ||\mathbf{u} - \mathbf{v}||$.

Note: There are many different ways to define distances/norms. In this class, the norm we use is called the standard Euclidean Norm.

Example 3.3. *Find the distance between* $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$ *and* $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$

3.2 Method of Least Squares

So if we want to the best approximation for solving $A\mathbf{x} = \mathbf{b}$, we want to minimize the distance between $A\mathbf{x}$ and $\mathbf{b}$. That is, our goal is to minimize:

Definition 3.6. *Given* $A \in \mathcal{M}_{n\times m}$ *and* $\boldsymbol{b} \in \mathbb{R}^n$*, a **least-squares solution** of* $A\mathbf{x} = \mathbf{b}$ *is a* $\hat{\mathbf{x}} \in \mathbb{R}^m$ *such that* $||\mathbf{b} - A\hat{\mathbf{x}}|| \leq ||\mathbf{b} - A\mathbf{x}||$ *for all* $x \in \mathbb{R}^m$.

Theorem 3.1. *The set of least-square solutions of* $A\mathbf{x} = \mathbf{b}$ *coincides with the nonempty set of solutions to the normal equation:* $A^T A\mathbf{x} = A^T\mathbf{b}$.

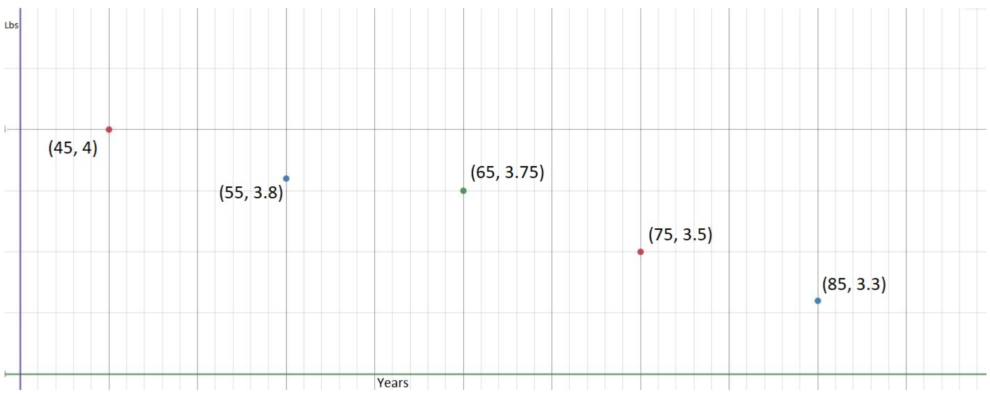
Proof. Omitted (We don't have the theory yet!).

Example 3.4. *Find a least-squares solution for* $A\mathbf{x} = \mathbf{b}$ *where* $A = \begin{bmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{bmatrix}$ *and* $\mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$.

3.3 Linear Modeling

One of the main problems that scientists, mathematicians, and statisticians face is trying to analyze and understand relationships from a data set. In calculus classes, we are often given functions to analyze and use. Unfortunately, often we have to create or build our own function or formula to describe our data. The good news is that linear algebra and the method of least squares can help!

Example 3.5. *Suppose we collect data on age versus brain mass (in the table below) and wanted to use this to help us hypothesize what happens with brain mass after 45 years.*
a) What would you hypothesize?



b) Suppose we hypothesize that brain mass decreases linearly with age. This means we want to use the function $f(x) = cx + a$ to model this. Our goal now is to find values for a and c which best fit our data. Find values for c and a so we can find a ***best-fit line*** *for our data and use it to predict the brain size of an 90 year old.*

Age (yrs)	**Brain Mass (lbs)**
45	4
55	3.8
65	3.75
75	3.5
85	3.3

We can generalize this. If we want to find the best fit n-degree polynomial to a data set $\{(x_1, y_1,), (x_2, y_2,), (x_3, y_3,), ...(x_m, y_m,)\}$, we could set up the following system of equations:

$$\begin{array}{llllll}
a_0 + & a_1 x_1 & +a_2(x_1)^2 & +a_3(x_1)^3 + ... + & a_n(x_1)^n = & y_1 \\
a_0 + & a_1 x_2 & +a_2(x_2)^2 & +a_3(x_2)^3 + ... + & a_n(x_2)^n = & y_2 \\
a_0 + & a_1 x_3 & +a_2(x_3)^2 & +a_3(x_3)^3 + ... + & a_n(x_3)^n = & y_3 \\
\vdots\ \vdots & \vdots & \vdots & \vdots\ \vdots\ \ \vdots\ \vdots & \vdots\ \vdots & \vdots \\
a_0 + & a_1 x_m & +a_2(x_m)^2 & +a_3(x_m)^2 + ... + & a_n(x_m)^n = & y_m
\end{array}$$

$$\begin{bmatrix} 1 & x_1 & (x_1)^2 & (x_1)^3 & ... & (x_1)^n \\ 1 & x_2 & (x_2)^2 & (x_2)^3 & ... & (x_2)^n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & (x_m)^2 & (x_m)^3 & ... & (x_m)^n \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{bmatrix}$$

Example 3.6. *Set up the matrix system we would solve if we wanted to find the best fit quadratic function for the following dataset.*

X	**Y**
0	1
1	3
2	6
3	12

We can also extend this theory to functions that are not polynomials. In fact, many functions are often approximated by linear combinations of sine and cosine. Example 3.7 gives us an example in which our function is a trigonometric polynomial. (Note: these types of functions show up in something called Fourier Series which is beyond the scope of this course, but you may see in later math and physics courses!)

Example 3.7. *A certain experiment produces the data* $(0, 7.9), (\frac{\pi}{2}, 5.4), (\pi, -0.9)$. *Find values for A and B which describe the model that produces a least squares fit of the points by a function of the form*
$f(x) = A\cos x + B\sin x$.

4 Linear Transformations

4.1 Motivation for Linear Transformations

Terminology: A *Transformation* is just another word for ________. This function will transform one space to another space. In Calculus 3, this occurred when we used a change of variables. For example, we can transform rectangular coordinates to polar through a transformation. In Figure 1 we see that this change of variables transformation takes round things and makes them rectangular.[6]

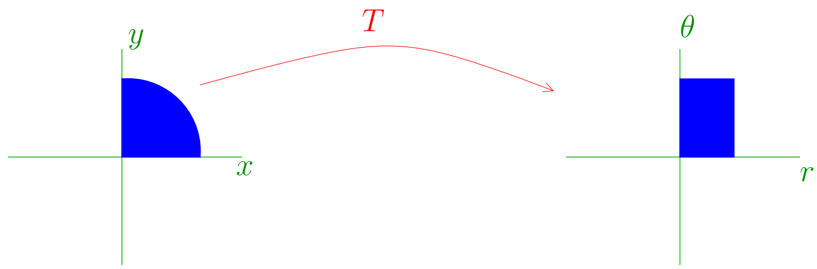


Figure 1:

Another type of transformation is when we project from one space to a lower dimensional space. For example, we can project a points in $\mathbb{R}^2$ to their component along the y axis as in Figure 2.

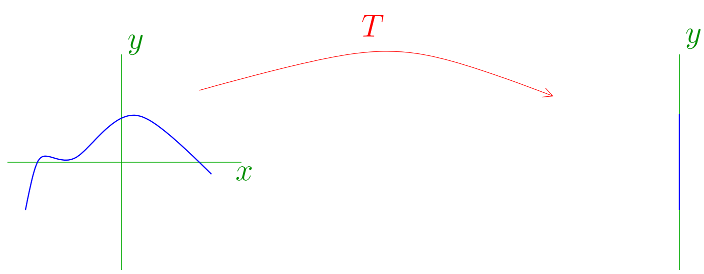


Figure 2:

In Lab #2, we found that when we apply a radiographic transformation to a linear combination, $\alpha u + \beta v$, we get a linear combination of radiographs, $\alpha T u + \beta T v$ out. This property is useful because we may wonder what is an object made of. If we know part of what is being radiographed (and what the corresponding radiograph should be), we can subtract that away to be able to view the radiograph of the part we don't know.

[6]Some Images from Heather Moon and Beezer's *A First Course in Linear Algebra*

Example 4.1. *Suppose we expect an object to look like what we see on the left in Figure 3 but the truth is actually like we see on the right in Figure 3.*

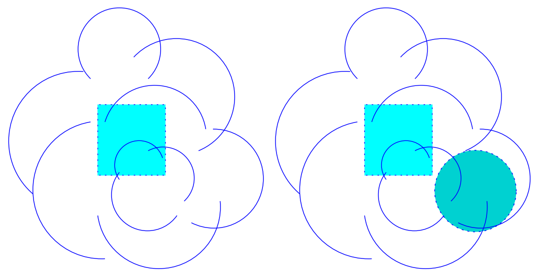


Figure 3:

We then expect that the radiograph that comes from an object, x_{exp} like we see on the left in Figure 3 to be a certain radiograph, call it b_{exp}. But, we take a radiograph of the actual object, x, and we get the radiograph b. Now, we know for sure that the expected parts are in the object. We can then remove b_{exp} from the radiograph so that we can dig deeper into what else is present. Thus, we want to know about the radiograph, b_{unexp}, of the unexpected object, call this x_{unexp}, that is present. So we compute this radiograph like this:

Another reason this might be helpful comes in possible changes in an object.

Example 4.2. *Suppose you radiograph an object as in Figure 3 and find that the radiograph is b_{exp}, but weeks later, you radiograph the same object (or so you think) and you get a radiograph that is* 1.3 *times the radiograph b_{exp}. This could mean that the object now looks more like one of the objects we see in Figure 4.*

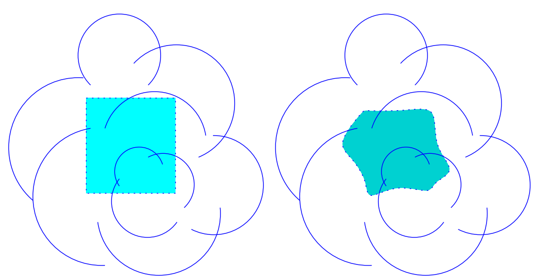


Figure 4:

Again, we can see that the density is proportionally larger, possibly meaning the object grew.

Being able to perform the type of investigation above with our radiographic transformation is very useful. It turns out that this property is useful beyond the application of radiography, so useful that we define transformations with this property.

4.2 Introduction to Linear Transformations

Definition 4.1. *Let V and W be vector spaces and let T be a function that maps vectors in V to vectors in W. We say that $T : V \to W$ is a **linear transformation** if both of the following are satisfied:*
1) For any arbitrary vectors $u_1, u_2 \in V$, $T(u_1 + u_2) = \underline{\hspace{3cm}}$
2) For any scalar α and vector $u \in V$, $T(\alpha u) = \underline{\hspace{3cm}}$

Note: You can combine these two steps into a 1-Step Check by showing that
$T(\alpha u_1 + \beta u_2) = \underline{\hspace{6cm}}$.

Terminology: We can also call T a **linear function** or a **homomorphism**.

Method: To check if $T : V \to W$ is a linear transformation, take 2 arbitrary elements in V, and an arbitrary scalar and check that
1) $T(u_1 + u_2) = T(u_1) + T(u_2)$, and
2) $T(\alpha u_1) = \alpha T(u_1)$

Note: The radiographic transformation is an example of a linear transformation.

$$\begin{tikzcd}
\mathbf{u}_1, \mathbf{u}_2 \arrow[r, "T"] \arrow[d, "+"] & T(\mathbf{u}_1), T(\mathbf{u}_2) \arrow[d, "+"] \\
\mathbf{u}_1 + \mathbf{u}_2 \arrow[r, "T"] & T(\mathbf{u}_1 + \mathbf{u}_2) = T(\mathbf{u}_1) + T(\mathbf{u}_2)
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathbf{u} \arrow[r, "T"] \arrow[d, "\alpha"] & T(\mathbf{u}) \arrow[d, "\alpha"] \\
\alpha\mathbf{u} \arrow[r, "T"] & T(\alpha\mathbf{u}) = \alpha T(\mathbf{u})
\end{tikzcd}$$

Example 4.3. *Consider the map $T : \mathbb{R}^3 \to W$ where $T \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a + b + c \end{pmatrix}$.*

a) Determine $T \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

b) Determine whether T is a linear transformation.

Example 4.4. *Is the map $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = 2x + 1$ a linear transformation?*

(a) Yes, and I am very confident
(b) Yes, but I am not very confident
(c) No, but I am not very confident
(d) No, and I am very confident

Example 4.5. *$f(x) = mx + b$ is a linear transformation only if what is true:*

(a) $m = 0$ and $b = 0$
(b) $m = 0$ and b can be anything
(c) $b = 0$ and m can be anything
(d) m and b can be anything!
(e) I like cats!

Theorem 4.1. *Let V and W be vector spaces. If $T : V \to W$ is a linear transformation, then $T(0) =$ __________.*

Proof. Let V and W be vector spaces and let $T : V \to W$ be a linear transformation. Notice that $0 \in V$ and $0 \in W$. (Note also that these two vectors called 0 need not be the same vector.) We also know that, for any scalar α, $T(0) = T(\alpha 0) = \alpha T(0)$. We can use this equation to solve for $T(0)$ and we get that either $\alpha = 1$ or $T(0) = 0$. Since we know that α can be any scalar, $T(0) = 0$ must be true. $\square$

Quick Check for Linear Transformations:
We can determine whether or not $T(0) =$ __________. **If not, T cannot be linear.**
If yes, you still have to check that $T(u_1 + u_2) = T(u_1) + T(u_2)$ and $T(\alpha u) = \alpha T(u)$.

Example 4.6. *Determine which of the following is a linear transformation.*

a) Define $f : \mathbb{R}^3 \to \mathbb{R}^2$ by $f(v) = Mv + x$, where $M = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix}$ and $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

b) Define $\mathcal{T} : \mathcal{P}_2 \to \mathcal{P}_1$, where by $\mathcal{T}(ax^2 + bx + c) = 2ax + b$.

Example 4.7. *Define* $T(v) = Av$, *where* $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$. *Then* $T(v)$.... *[Hint: let* $v = \begin{bmatrix} x \\ y \end{bmatrix}$*]*

(a) reflects v about the x-axis.
(b) reflects v about the y-axis.
(c) rotates v clockwise $\frac{\pi}{2}$ radians about the origin.
(d) rotates v counterclockwise $\frac{\pi}{2}$ radians about the origin.
(e) None of the above

4.2.1 Motivation for Matrix Forms of Linear Transformations:

Functions are relatively easy for us to understand, but it would be nice to be able to encode them into a matrix form so computers can understand and use them better. Matrices tend to be a good way to store information in a computer. They are also, at times, easier to work with (as they were when we solved systems of equations). Coding a linear transformation based on the formula can at times be very tedious. So, we want to be able to use matrices as a tool for linear transformations as well. Let's look at an example that suggests that this might be a good idea.

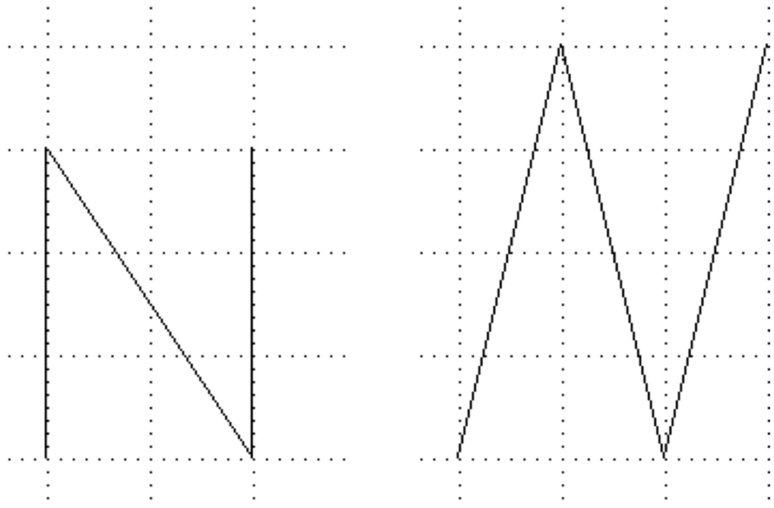
Example: Define $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $Tx = Mx$, where M is a 3×2 matrix. T is a linear transformation. We can show this using properties of matrix multiply. Let $x, y \in \mathbb{R}^2$ and let α, β be scalars. Then

$$\begin{aligned} T(\alpha x + \beta y) &= M(\alpha x + \beta y) \\ &= M\alpha x + M\beta y \\ &= \alpha Mx + \beta My \\ &= \alpha Tx + \beta Ty. \end{aligned}$$

This example shows us that matrix multiply defines a linear transformation. So, what if we can define all linear transformations with a matrix multiply? That would be really great!
Issue: We can't just multiply a vector, in say $\mathcal{P}_2$, by a matrix. What would that mean?

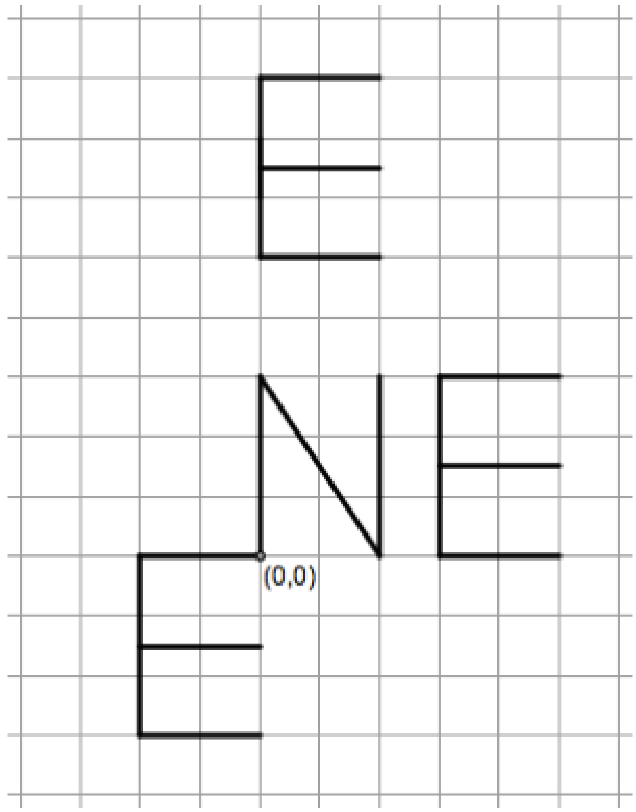
Key: Represent it using a *Coordinate Vector*
(coming soon after this ICE sheet!)

4.2.2 ICE 4 -Linear Transformations



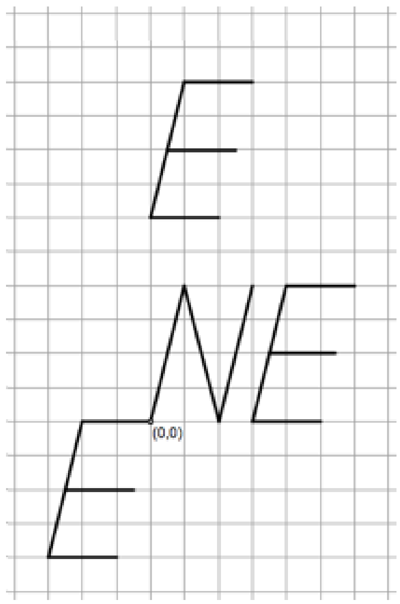
1) Suppose "N" on the left is written in regular 12-point font. Find a matrix A that will transform N into that letter on the right, which is written in 'italics' in 16-point font.

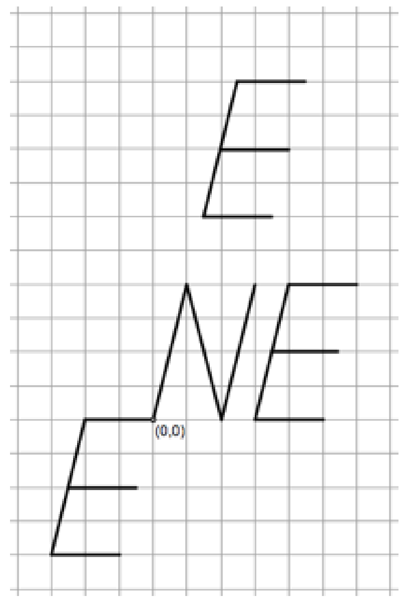
$A =$



2) After class, Eva, Archer and their kitty friends Gilbert and Mako are wondering how letters placed in other locations in the plane would be transformed under $A = \begin{bmatrix} 1 & 1/3 \\ 0 & 4/3 \end{bmatrix}$. If an "E" is placed around the "N," the kitties argued over four different possible results for the transformed E's on the next page. Which choice below is correct, and why? If none of the four options are correct, what would the correct option be, and why?

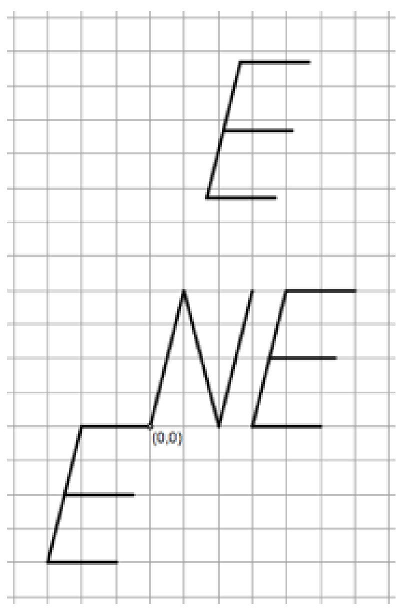
Choice A

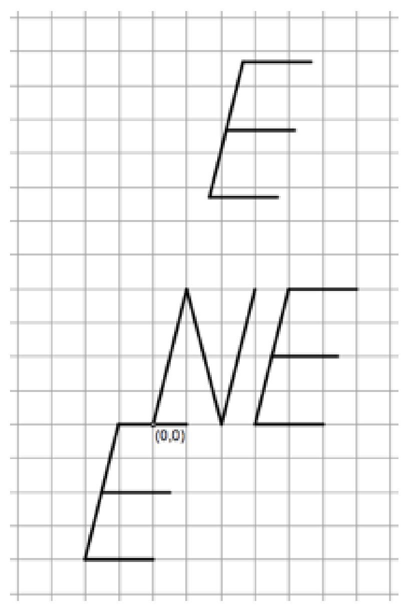




Choice B

Choice C

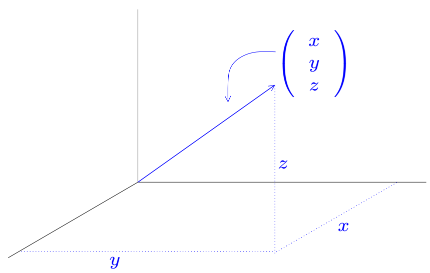




Choice D

5 Coordinate Vectors

In our typical 3D space, we talk about vectors that look like $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$. And we see that the coordinates are x, y, and z respectively. What this means is that the vector points from the origin to a point that is x units horizontally, y units vertically, and z units up from the origin (as in the Figure below. The truth is that when we say that, we are assuming that we are working with the *standard basis for* $\mathbb{R}^3$. (This makes sense because, it is the basis that we usually think about, x-axis perpendicular to the y-axis, forming a horizontal plane and the z axis perpendicular to this plane.)



So, using the standard basis $\mathcal{E} = \{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\}$, we can write

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Notice that the coordinates are the scalars in this linear combination. That's what we mean by coordinates. Notice the coordinates of the vector $v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ in the standard basis are __________, __________, & __________ respectively.

Now, let us consider a different basis for $\mathbb{R}^3$, $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. Let's find the coordinates for $v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ in this new basis. These are found by finding the scalars α_1, α_2, and α_3 so that $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Going through the motions to solve for α_1, α_2, and α_3, we find that $\alpha_1 = 2, \alpha_2 = -1$, and $\alpha_3 = 2$.

Thus, we can represent the vector v in coordinates according to the basis $\mathcal{B}$ as

$$[v]_{\mathcal{B}} = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}.$$

This vector is called the $\mathcal{B}-$*coordinate vector of* v *or the coordinate vector relative to* $\mathcal{B}$ *and is denoted as* $[v]_{\mathcal{B}}$. Notice that we indicate that we are looking for the coordinates in terms of the basis B by using the notation, $[v]_{\mathcal{B}}$ by using the notation, $[v]_{\mathcal{B}}$. This means if we are given a vector space V and bases $\mathcal{B}_1 = \{v_1, v_2, \ldots, v_n\}$ and $\mathcal{B}_2 = \{u_1, u_2, \ldots, u_n\}$, then if $w = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$, we have $[w]_{\mathcal{B}_1} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$.

But, if $w = \beta_1 u_1 + \beta_2 u_2 + \ldots + \beta_n u_n$, we have $[w]_{\mathcal{B}_2} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$.

In our example from $\mathbb{R}^3$, notice $[v]_{\mathcal{B}} = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$
but with respective to the standard basis, $[v]_{\mathcal{E}_3} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

Important: Notice also that coordinate vectors look like vectors in $\mathbb{R}^n$ for some n.

Let's do another example from a vector space that isn't already in vector form:

Suppose we have $v = 3 + x - x^2 \in \mathcal{P}_2$. Given the standard basis for $\mathcal{P}_2, \mathcal{B} = \{1, x, x^2\}$, we can think of $v = 3 + x - x^2 = 3(1) + 1(x) - 1(x^2)$. We define $3, 1, -1$ as the *$\mathcal{B}$-coordinates of v*. Now we can represent $3 + x - x^2$ as a vector in $\mathbb{R}^3$:

Example 5.1. *What is the coordinate vector of $3 + x - x^2$ relative to the standard basis of $\mathcal{P}_3$, $\mathcal{S} = \{1, x, x^2, x^3\}$? Does order matter?*

Key Idea: I like to think of the coordinate vector of v relative to a basis as the "instruction manual" for how to put the basis elements together to create v.

Example 5.2. *Let $V = \{ax^2 + (b)x + (c) \in \mathcal{P}_2 |\ a + b - 2c = 0\}$. One basis for V is $\mathcal{B} = \{-x^2 + x, 2x^2 + 1\}$ (check this on your own).*

a) Since $v = 3x^2 - 3x \in V$, we can write v as a coordinate vector with respect to $\mathcal{B}$.

b) Now suppose we know the coordinates for a vector w are $[w]_{\mathcal{B}} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. We can use this to determine what w actually is in V.

Let's now consider representations of a vector when we view the vector space in terms of two different bases.

Example 5.3. *Let* $V = \left\{ \begin{pmatrix} a & b-a \\ a+b & a+2b \end{pmatrix} \middle| \; a, b \in \mathbb{R} \right\}$. *Archer has found one base for* V*:* $\mathcal{B}_1 = \{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}\}$. *Eva has used the basis to construct another basis for* V*:* $\mathcal{B}_2 = \{\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 1 & -2 \\ 0 & -1 \end{bmatrix}\}$.

a) Find a vector such that $[w]_{\mathcal{B}_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

b) Write find the coordinates of the vector your found in terms of Eva's basis.(That is find the instruction manual to build it from Eva's basis.)

Reading Example:

Example 5.4. *Let* $V = \{ax^2 + (b-a)x + (a+b) | a, b \in \mathbb{R}\}$. *Find a basis for* V *and find the coordinate vector for* $v = 3x^2 + x + 7 \in V$ *with respect to this basis.*

$$V = \{a(x^2 - x + 1) + b(x+1) |\ a, b \in \mathbb{R}\} = \text{span}\{x^2 - x + 1, x + 1\}.$$

So a basis for V is

$$\mathcal{B} = \{v_1 = x^2 - x + 1, v_2 = x + 1\}.$$

This means that $\dim V = 2$ and so vectors in V can be represented by vectors in $\mathbb{R}^2$. Notice that $v = 3x^2 + x + 7 \in V$. You actually should be able to check this ($a = 3, b = 4$). We can write v in terms of the basis of V as $v = 3v_1 + 4v_2$. We can check this as follows

$$3v_1 + 4v_2 = 3(x^2 - x + 1) + 4(x + 1) = 3x^2 + x + 7.$$

Thus, the coordinate vector for v is $[v]_{\mathcal{B}} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

6 Matrix Forms of Linear Transformations:

Now that we have discussed coordinate vectors in more detail, let's go back to the reason why we went down that bunny trail.

Recall, we wanted to be able to define a linear transformations with a matrix, but the issue is that we can't just multiply a vector, in say $\mathcal{P}_2$, by a matrix.

The Key is to use Coordinate Vectors!

Given a basis for V, we are able to represent any vector $v \in V$ as a coordinate vector in $\mathbb{R}^n$, where $n = \dim V$. Suppose $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ is a basis for V, then we find the coordinate vector $[v]_{\mathcal{B}}$ by finding the scalars, α_i, that make the linear combination $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$ and we get

$$[v]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n.$$

After our next section, we will talk a little bit more about the theory behind why we construct the matrix of a linear transformation (and how we create a linear transformation that changes basis representation) later.

Our goal is to create a matrix that does what we want: $[T(v)]_{\mathcal{B}_W} = M[v]_{\mathcal{B}_V}$.

So given a Linear Transformation $T : V \to W$, we first need to convert vectors in V into vectors in $\mathbb{R}^n$. To do this we create the coordinate vector for each basis element in the standard basis. We then take the image of each basis element from V, $T(v)$ and write it as a coordinate vector in $\mathcal{B}_W$. We do this one basis element at a time and construct a matrix so that

$$\begin{aligned} M[v_1]_{\mathcal{B}_V} &= [T(v_1)]_{\mathcal{B}_W} \\ M[v_2]_{\mathcal{B}_V} &= [T(v_2)]_{\mathcal{B}_W} \\ &\vdots \\ M[v_n]_{\mathcal{B}_V} &= [T(v_n)]_{\mathcal{B}_W}. \end{aligned}$$

Notice that

$$[v_1]_{\mathcal{B}_V} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ and } [v_2]_{\mathcal{B}_V} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \text{ etc....}$$

So, $M[v_1]_{\mathcal{B}_V}$ gives the first column of M, $M[v_2]_{\mathcal{B}_V}$ gives the second column of M, $\ldots$, and $M[v_n]_{\mathcal{B}_V}$ gives the nth column of M. Thus, we have that

$$M = \begin{pmatrix} | & | & & | \\ [T(v_1)]_{\mathcal{B}_W} & [T(v_2)]_{\mathcal{B}_W} & \ldots & [T(v_n)]_{\mathcal{B}_W} \\ | & | & & | \end{pmatrix}$$

Steps for finding the Matrix Form of a Linear Transformation:

To find a matrix form for the linear transformation $T : V \to W$:
Example: $T : \mathcal{P}_2 \to \mathcal{M}_{2\times 2}$

Step 1: Find bases for the domain (in this example, this is V) and co-domain (in this example, this is W).

Ex: $V = \mathcal{P}_2$ so $\mathcal{B}_V =$ ____________________ and $W = \mathcal{M}_{2\times 2}$ so $\mathcal{B}_W =$ ____________________

Step 2: For each element of the basis of the domain (in this example, this is V), $\mathcal{B}_V$, find its image after apply T to it.

Ex: $T(v_i)$ for $\mathcal{B}_V$:

Step 3: Write the image as a coordinate vector with respect to the basis for co-Domain (in this example, this is W), $\mathcal{B}_W$.

Ex: $[T(v_i)]_{\mathcal{B}_W}$: $[T(1)]_{\mathcal{B}_W}$, $[T(x)]_{\mathcal{B}_W}$, $[T(x^2)]_{\mathcal{B}_W}$

Step 4: Create M using the coordinate vectors you created in Step 3: $M = [T(v_1)_{\mathcal{B}_W}, T(v_2)_{\mathcal{B}_W}, ..., T(v_n)_{\mathcal{B}_W},$

Ex: Our matrix would be:

Useful, but optional step: Check that your matrix works to see if you get the same thing as you would with the original transformation.

Note if you go from an n dimensional space to an m dimensional space, your matrix will be an $m \times n$ matrix.

Example 6.1. *Find a matrix form for the linear transformation* $T : \mathcal{P}_1 \to \mathcal{P}_2$ *where* $T(ax+b) = ax^2 + (a+b)x - b$.

First of all, since we are going from a domain of $V = \mathcal{P}_1$ which is a __________ dimensional vector space to a codomain of $W = \mathcal{P}_2$ which is a __________ dimensional vector space. The matrix we create should have the dimensions __________ $\times$ __________.

Step 1: Find bases for V and W.
$\mathcal{B}_V =$

$\mathcal{B}_W$

Step 2: For each element of the basis of V, $\mathcal{B}_V$, find its image after apply T to it.

Step 3: Write the image as a coordinate vector with respect to the basis for W, $\mathcal{B}_W$.

Step 4: Create M using the coordinate vectors you created in Step 3: $M =$

Check that your matrix works to see if you get the same thing as $T(8x+4)$

Example 6.2. *Let's try to find a matrix representation for the linear transformations in Exercise 4.2. That is, define a matrix representation for the linear transformation:* $T : \mathbb{R}^3 \to \mathbb{R}$ *is defined as follows.* $T\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a+b+c \end{pmatrix}$.

Example 6.3. *The linear transformation* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *where* $T(x, y) = (x + 2y, x - 2y)$, *can be written as a matrix transformation* $T(x, y) = A\begin{bmatrix} x \\ y \end{bmatrix}$ *where*

(a) $A = \begin{bmatrix} x & 2y \\ x & -2y \end{bmatrix}$
(b) $A = \begin{bmatrix} 1 & 2 \\ 1 & -2 \end{bmatrix}$
(c) $A = \begin{bmatrix} 1 & -2 \\ 1 & 2 \end{bmatrix}$
(d) It can't be written in matrix form.
(e) Archer is my favorite.

Example 6.4. *The linear transformation* $T(x, y) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$, *can be written as*

(a) $T(x, y) = (x, y)$
(b) $T(x, y) = (y, x)$
(c) $T(x, y) = (-x, y)$
(d) $T(x, y) = (-y, x)$
(e) None of the above

Things get a little stranger when you look at subspaces. When you have this type of situation, you must first find a basis for the subspace of V.

Example 6.5. *Find a matrix representation for the Linear Transformation* $h : V \to \mathcal{P}_1$*, where* $V = \left\{ \left(\begin{array}{ccc} a & b & c \\ 0 & b-c & 2a \end{array} \right) \middle| \; a, b, c \in \mathbb{R} \right\} \subseteq \mathcal{M}_{2\times 3}$ *and*

$$h \left(\begin{array}{ccc} a & b & c \\ 0 & b-c & 2a \end{array} \right) = ax + c.$$

6.0.1 Radiographic Connection

When working with radiographic transformations, we found a matrix using tomomap. But, our objects weren't vectors in $\mathbb{R}^N$ that could be multiplied by the matrix we found. Let's use the above information to explore, through an example, what was really happening. Let $V = \mathcal{I}_{2\times 2}$, the space of 2×2 objects. Let T be the radiographic transformation with 6 views having 2 pixels each. This means that the codomain is the set of radiographs with 12 pixels. To figure out the matrix M representing this radiographic transformation, we first change the objects in V to coordinate vectors in $\mathbb{R}^4$ via the isomorphism T_1. So T_1 is defined by:

$$\begin{array}{ccc} V & \xrightarrow{\quad T_1 \quad} & \mathbb{R}^4 \\ \begin{array}{|c|c|}\hline x_1 & x_3 \\ \hline x_2 & x_4 \\ \hline \end{array} & \longmapsto & \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \end{array}$$

After performing the matrix multiply, we will change from coordinate vectors in $\mathbb{R}^{12}$ back to radiographs via T_2. So T_2 is defined by by:

$$\begin{array}{ccc} \mathbb{R}^{12} & \xrightarrow{\quad T_2 \quad} & W \\ \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \\ b_8 \\ b_9 \\ b_{10} \\ b_{11} \\ b_{12} \end{bmatrix} & \longmapsto & \begin{array}{c|c} b_1 & b_2 \\ \hline b_3 & b_4 \\ \hline b_5 & b_6 \\ \hline b_7 & b_8 \\ \hline b_9 & b_{10} \\ \hline b_{11} & b_{12} \\ \hline \end{array} \end{array}$$

Our radiographic transformation is then represented by the matrix M (which we called T in the labs). M will be a 12×4 matrix determined by the radiographic set up. We've computed M several times in previous labs, but the real mathematics was all a bit hand-wavy and so now we see that really, what we have is that T maps from V to W by taking a side excursion through coordinate spaces and doing a matrix multiply.

Note: Typically, to simplify notation, we write $T(v) = Mv$ when we actually mean $[T(v)]_{\mathcal{B}_W} = M[v]_{\mathcal{B}_V}$. This is understandable by a mathematician because we recognize that when two spaces are isomorphic, they are "essentially" the same set. The only difference is that the vectors look different. In this class, we will maintain the notation discussed in this section being aware that this is just to get used to the ideas before relaxing our notation.

6.0.2 ICE 5 -Matrix Linear Transformations

1. Let's try to find a matrix representation for the linear transformation in Exercise 4.3.That is, Define a matrix representation for the linear transformation: $\mathcal{T} : \mathcal{P}_2 \to \mathcal{P}_1$, where $\mathcal{T}(ax^2 + bx + c) = 2ax + b$.

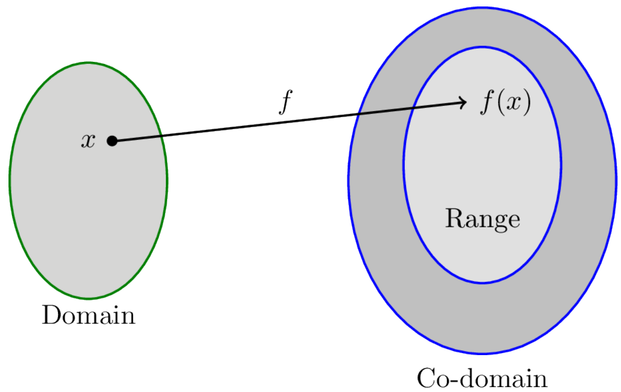
2. Suppose $T : \mathcal{P}_2 \to \mathcal{P}_1$ where $T(ax^2 + bx + c) = (a - 2b)x + a - c$. Find a matrix representation for this linear transformation.

3. Slightly more challenging example: Let $V = \{ax^2 + (b-a)x + (a+b) |\ a, b \in \mathbb{R}\}$ and let $W = \mathcal{M}_{2\times 2}$. Consider the transformation $T : V \to W$ defined by $T(ax^2 + (b-a)x + (a+b)) = \begin{pmatrix} a & b-a \\ a+b & a+2b \end{pmatrix}$.
   Construct the Matrix Representation of T. Remember you will need to find bases for V and W!
   Note: Archer thinks the basis for V is $\{\begin{bmatrix}1 & -1 & 1\end{bmatrix}, \begin{bmatrix}0 & 1 & 1\end{bmatrix}, \}$

7 Properties of Transformations

7.1 Injective and Surjective Transformations

Recall, a linear transformation is just a function with special properties. [7]



In Lab # 3, we will see several properties of linear transformations that are useful to recognize. We will see that it is possible for two objects to produce the same radiograph. This can be an issue in the case of brain radiography. We would like to know if it would be possible for an abnormal brain to produce the same radiograph as a normal brain. We also will see that it was possible to have radiographs that could not be produce from any object. This becomes important in being able to recognize noise or other corruption in a given radiograph. Again, it turns out that these properties are not only important in radiography. There are many other scenarios (some application based and some theoretical) where we need to know if a transformation has these properties. So, let's define them.

Due to the press for time in this course and the "applied" focus for the course, we will not dwell as much on proving transformations have these properties. Later we will have a more computational way to check these properties. I will provide some of the arguments though for completeness and for reference.

Definition 7.1. *Let V and W be vector spaces. We say that the transformation $T : V \to W$ is* ***injective*** *if the following property is true:*
Whenever $T(u) = T(v)$ it must be true that __________ = __________.

A transformation that is injective is also said to be <u>one-to-one</u> (1-1) or an <u>injection</u>.

Idea:

[7] Image used with permission from https://www.probabilitycourse.com

Process to show Injectivity: Start with two arbitrary elements in the range of T and suppose they are equal, that is suppose $T(u) = T(v)$ and work backwards to see if $u = v$. In other words, can we find two possible input values that map to the same output? If yes, T is NOT injective.

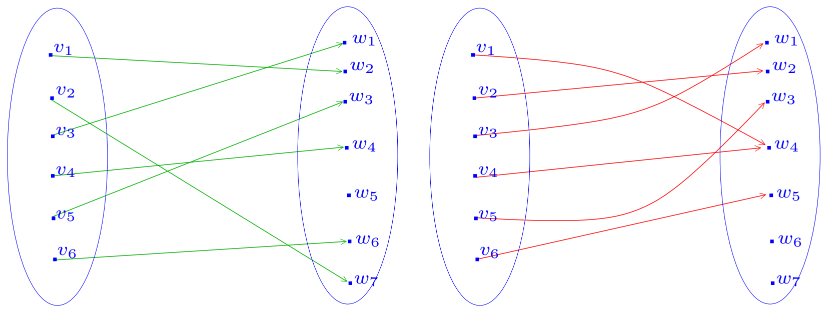


Figure 5: Notice, we see that the transformation represented on the left is 1-1, but the transformation represented on the right is not because both v_1 and v_2 map to w_4, but $v_1 \neq v_2$.

Example 7.1. *Let* $T : \mathbb{R}^2 \to \mathcal{M}_{3\times 2}$ *be the linear transformation defined by* $T\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a+b \\ 0 & -a \end{pmatrix}$. *Determine whether it is one-to-one.*

Proof. To show: T is injective/(one-to one)
Assume

$$T\begin{pmatrix} a \\ b \end{pmatrix} = T\begin{pmatrix} c \\ d \end{pmatrix}$$

Then

$$\begin{pmatrix} a & -b \\ b & a+b \\ 0 & -a \end{pmatrix} = \begin{pmatrix} c & -d \\ d & c+d \\ 0 & -c \end{pmatrix}$$

Matching up entries, gives us $a = c, -b = -d, b = d, a + b = c + d, 0 = 0$, and $-a = -c$. Thus,

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}.$$

So, by definition, T is one-to-one. □

Example 7.2. *Let $T : \mathcal{P}_2 \to \mathbb{R}^2$ be defined by $T(ax^2 + bx + c) = \begin{pmatrix} a-b \\ b+c \end{pmatrix}$. Notice that T is a linear transformation. Is T an injection?*

Proof. T is not injective, notice that both $x + 4$ and $5 - x^2$ map to $\begin{bmatrix} -1 \\ 5 \end{bmatrix}$,
that is $T(x + 4) = T(5 - x^2) = \begin{bmatrix} -1 \\ 5 \end{bmatrix}$. □

But how could we determine this? Here is another way.
Suppose $T(ax^2+bx+c) = T(ex^2+fx+g)$, this implies (after apply T) that $\begin{bmatrix} a-b \\ b+c \end{bmatrix} = \begin{bmatrix} e-f \\ f+g \end{bmatrix}$.

This means (after matching the entries) that $\begin{cases} a-b = e-f \\ b+c = f+g \end{cases}$

We can set up the augmented equation to solve this: $\left(\begin{array}{ccc|c} 1 & -1 & 0 & e-f \\ 0 & 1 & 1 & f+g \end{array}\right)$

Notice we have a free variable which means we have the freedom to play a bit. At this point, we should be suspicious that this transformation is NOT one-to-one, but we need to find an output that has 2 different inputs that can map to it. So for example, we can pick any values for e, f, and g like $e = 1, f = 2, g = 3$ to help us find one of the two vectors which can show that we break the injectivity requirement.

Now we need to solve: $\begin{cases} a-b = e-f \\ b+c = f+g \end{cases} \Rightarrow \begin{cases} a-b = 1-2 \\ b+c = 2+3 \end{cases} \Rightarrow \begin{cases} a-b = -1 \\ b+c = 5 \end{cases}$

So now we have this system to solve:
$\left(\begin{array}{ccc|c} 1 & -1 & 0 & -1 \\ 0 & 1 & 1 & 5 \end{array}\right)$
Solving this gives us
$a = -1 + b$
$b =$ free baby!
$c = 5 - b$.

So we can pick two different values for b. In the example above we picked $b = 1$ which means $a = 0, c = 4$ which gives the vector $x + 4$. We then picked $b = 0$ which means $a = -1, c = 5$ which gives us the vector $-x^2 + 5$.

Definition 7.2. *Let V and W be vector spaces. We say that the transformation $T : V \to W$ is **surjective** if every element in W is mapped to. That is, if $w \in W$, then there exists a $v \in V$ so that* __________ = __________.

A transformation that is surjective is also said to be onto or a surjection.

Idea: Dr. Harsy is going to do her absolute worst to present you with the most obscure element of the codomain. Are you certain that this function will be able to map to this element no matter what Dr. Harsy presents you with?

Process to show surjectivity: Pick an arbitrary element in co-domain, find an element that maps to it.

Example 7.3. *Consider the function $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^2$. Note that this is neither injective nor onto (and isn't a linear transformation either). Why?*

Example 7.4. *Determine which of the maps below are surjective and which maps are injective.*

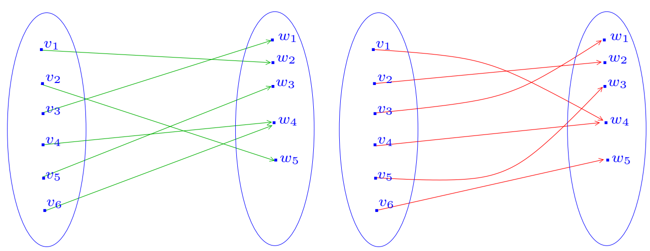


Figure 6:

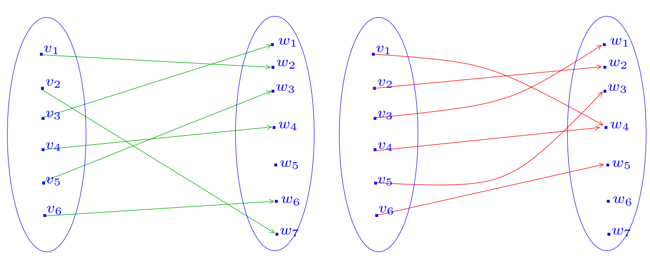


Figure 7:

Example 7.5. *Let $T : \mathbb{R}^2 \to \mathcal{M}_{3\times 2}$ be defined by $T\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a+b \\ 0 & -a \end{pmatrix}$. We already showed this linear transformation was one-to-one in Ex 7.1. Is it onto?*

Proof. Notice that,

$$w = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} \in \mathcal{M}_{3\times 2}.$$

But, there is no $v \in \mathbb{R}^2$ so that $T(v) = w$ because every output has a 0 in the $(3,1)$ entry. □

Example 7.6. *Let $T : \mathbb{R}^2 \to \mathcal{P}_1$ be defined by $T\begin{pmatrix} a \\ b \end{pmatrix} = 2ax + b$. Is this map 1-to-1 or onto?*

Proof. **T is 1-to-1:**
Suppose $T\begin{pmatrix} a \\ b \end{pmatrix} = T\begin{pmatrix} c \\ d \end{pmatrix}$. Then implies $2ax + b = 2cx + d$. Matching up like terms gives us that $\begin{cases} 2a = 2c \\ b = d \end{cases}$ Solving this, gives us $a = c$ and $b = d$. That is $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}$. Therefore T is one-to-one.

T is also onto:
We want to pick an arbitrary representative of a vector in the codomain, $\mathcal{P}_1$, say $ax + b$ and we want to find a vector in the domain, $\mathbb{R}^2$, $\begin{bmatrix} c \\ d \end{bmatrix}$ such that $T(\begin{bmatrix} c \\ d \end{bmatrix}) = ax + b$. Let $w = ax + b \in \mathcal{P}_1$, and Notice also that $v = \begin{bmatrix} \frac{a}{2} \\ b \end{bmatrix} \in \mathbb{R}^2$ maps to $ax + b$. Thus $T(v) = w$. Therefore, T is onto. □

7.1.1 ICE 5 -Injective and Surjective Linear Transformations

For the next 4 questions, try to construct the following transformations. If you can, justify or prove that your transformation has the property you claim it has or give an argument why it is impossible.

1. If possible, create a transformation that maps from $\mathbb{R}^3$ to $\mathcal{P}_1$ that is surjective.

2. If possible, create a transformation that maps from $\mathbb{R}^3$ to $\mathcal{P}_1$ that is injective.

3. If possible, create a transformation that maps from $\mathcal{P}_1$ to $\mathbb{R}^3$ that is injective.

4. If possible, create a transformation that maps from $\mathcal{P}_1$ to $\mathbb{R}^3$ that is surjective.

5. What do you notice? Given $T : V \to W$
In order for T to be injective, the dim V needs to be ________ than dim W.
In order for T to be surjective, the dim V needs to be ________ than dim W.

7.2 Bijections and Isomorphisms

We see above that sometimes a linear transformation can be both injective and surjective. In this subsection we discuss this special type of linear transformation.

Definition 7.3. *We say that a linear transformation, $T : V \to W$, is* ***bijective*** *if T is both injective and surjective. We call a bijective transformation a* ***bijection*** *or an* ***isomorphism****.*

Definition 7.4. *Let V and W be vector spaces. If there exists a bijection mapping between V and W, then we say that V is* ***isomorphic*** *to W and we write $V \cong W$.*

Example 7.7. *Notice that 7.5 we found that T was a bijection. This means that $\mathcal{P}_1 \cong \mathbb{R}^2$. Notice also that $\dim \mathcal{P}_1 = \dim \mathbb{R}^2$. This is not a coincidence.*

Lemma 7.1. *Let V and W be vector spaces. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ be a basis for V. $T : V \to W$ is an injective linear transformation if and only if $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a linearly independent set in W.*

Proof. As with every proof about linear dependence/independence, we will assume the following linear dependence relation is true. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be scalars so that

$$\alpha_1 T(v_1) + \alpha_2 T(v_2) + \ldots + \alpha_n T(v_n) = 0.$$

Then because T is linear, we know that

$$T(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) = 0.$$

But, we also know that $T(0) = 0$. That means that

$$T(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) = T(0).$$

And, since T is 1-1, we know that (by definition)

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0.$$

Finally, since $\mathcal{B}$ is a basis for V, $\mathcal{B}$ is linearly independent. Thus,

$$\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0.$$

Thus, $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is linearly independent.
Now suppose that T is linear, let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ be a basis for V, and suppose $\{T(v_1), T(v_2), \ldots, T(v_n)\} \subset W$ is linearly independent. Suppose $u, v \in V$ so that $T(u) = T(v)$. So, $T(u - v) = 0$. Since $u, v \in V$, there are scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ and $\beta_1, \beta_2, \ldots, \beta_n$ so that

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n \quad \text{and} \quad v = \beta_1 v_1 + \beta_2 v_2 + \ldots + \beta_n v_n.$$

Thus
$$T\left((\alpha_1-\beta_1)v_1+(\alpha_2-\beta_2)v_2+\ldots+(\alpha_n-\beta_n)v_n\right)=0.$$
This leads us to the linear dependence relation
$$(\alpha_1-\beta_1)T(v_1)+(\alpha_2-\beta_2)T(v_2)+\ldots+(\alpha_n-\beta_n)T(v_n)=0.$$
Since $\{T(v_1),T(v_2),\ldots,T(v_n)\}$ is linearly independent, we know that
$$\alpha_1-\beta_1=\alpha_2-\beta_2=\ldots=\alpha_n-\beta_n=0.$$
That is, $u=v$. So, T is injective. $\square$

Note: Notice that Lemma 7.1 tells us that if V is n-dimensional then basis elements of V map to basis elements of an n-dimensional subspace of W. In particular, if $\dim W=n$ also, then we see a basis of V maps to a basis of W. This is very useful and leads to the following useful result.

Theorem 7.1. *Given (finite dimensional) vector spaces V and W, then $V\cong W$ if and only if $\dim V=\dim W$.*

Proof. Suppose $\dim V=\dim W$. Suppose also that a basis for V is $\mathcal{B}_V=\{v_1,v_2,\ldots,v_n\}$ and a basis for W is $\mathcal{B}_W=\{w_1,w_2,\ldots,w_n\}$. Then we can define $T:V\to W$ to be the linear transformation so that
$$T(v_1)=w_1,T(v_2)=w_2,\ldots,T(v_n)=w_n.$$
We will show that T is an isomorphism. Now, we know that if $w\in W$, then $w=\alpha_1w_1+\alpha_2w_2+\ldots+\alpha_nw_n$ for some scalars $\alpha_1,\alpha_2,\ldots,\alpha_n$. We also know that $v=\alpha_1v_1+\alpha_2v_2+\ldots+\alpha_nv_n\in V$. Since T is linear, we can see that
$$\begin{aligned}T(v)=&T(\alpha_1v_1+\alpha_2v_2+\ldots+\alpha_nv_n)\\=&\alpha_1T(v_1)+\alpha_2T(v_2)+\ldots+\alpha_nT(v_n)\\=&\alpha_1w_1+\alpha_2w_2+\ldots+\alpha_nw_n=w.\end{aligned}$$
Thus, T is onto. Now, suppose that $T(v)=T(u)$ where $v=\alpha_1v_1+\alpha_2v_2+\ldots+\alpha_nv_n$ and $u=\beta_1v_1+\beta_2v_2+\ldots+\beta_nv_n$ are vectors in V. Then we have
$$\begin{aligned}T(\alpha_1v_1+\alpha_2v_2+\ldots+\alpha_nv_n)=&T(\beta_1v_1+\beta_2v_2+\ldots+\beta_nv_n)\\\alpha_1T(v_1)+\alpha_2T(v_2)+\ldots+\alpha_nT(v_n)=&\beta_1T(v_1)+\beta_2T(v_2)+\ldots+\beta_nT(v_n)\\\alpha_1w_1+\alpha_2w_2+\ldots+\alpha_nw_n=&\beta_1w_1+\beta_2w_2+\ldots+\beta_nw_n\\(\alpha_1-\beta_1)w_1+(\alpha_2-\beta_2)w_2+\ldots+(\alpha_n-\beta_n)w_n=&0.\end{aligned}$$

Notice that this last equation is a linear dependence relation for the basis $\mathcal{B}_W$. Since $\mathcal{B}_W$ is linearly independent, we know that

$$\begin{aligned}\alpha_1 - \beta_1 &= 0\\ \alpha_2 - \beta_2 &= 0\\ &\vdots\\ \alpha_n - \beta_n &= 0.\end{aligned}$$

That is to say $u = v$. Thus, T is injective. And, therefore, since T is both injective and surjective, T is an isomorphism. Now, since there is an isomorphism between V and W, we know that $V \cong W$.
Now we will prove the other direction. That is, we will show that if $V \cong W$ then $\dim V = \dim W$. First, let us assume that $V \cong W$. This means that there is an isomorphism, $T : V \to W$, mapping between V and W.
Suppose, for the sake of contradiction, that $\dim V \neq \dim W$. Without loss of generality, assume $\dim V > \dim W$. (We can make this assumption because we can just switch V's and W's in the following argument and argue for the case when $\dim V < \dim W$.) Let $\mathcal{B}_V = \{v_1, v_2, \ldots, v_n\}$ be a basis for V and $\mathcal{B}_W = \{w_1, w_2, \ldots, w_m\}$ be a basis for W. Then $m < n$. We will show that this cannot be true. Lemma 7.1 tells us that since T is one-to-one, the basis $\mathcal{B}_V$ maps to a linearly independent set $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ with n elements. But by Theorem 2.1, we know that this is not possible. Thus, our assumption that $n > m$ cannot be true. Again, the argument that tells us that $n > m$ also cannot be true is very similar with V's and W's switched. Thus, $n = m$. That is, $\dim V = \dim W$. $\square$

Careful: If we allow the dimension of a vector space to be infinite, then this is not necessarily true. In this class, we restrict to discussions of finite dimensional vector spaces only.

This theorem gives us a tool for creating isomorphisms, when they exist. It also tells us that isomorphisms exist between two vector spaces so long as they have the same ________.

Example 7.8. *Suppose we have a linear transformation $T : P_2 \to \mathbb{R}^2$. Using dimensional analysis, which of the following properties could T have: One-to-One, Onto, and/or Isomorphism?*

Example 7.9. *Suppose we have a linear transformation matrix, A, what dimensions will A need to have for it to possibly represent a bijective linear transformation?*

Example 7.10. *Let $V = \mathcal{M}_{2\times 3}$ and $W = \mathcal{P}_5$. We know that $V \cong W$ because both are* 6*-dimensional vector spaces. Indeed, a basis for V is*

$$\mathcal{B}_V = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

and a basis for W is $\mathcal{B}_W = \{1, x, x^2, x^3, x^4, x^5\}$. Create a bijection T that maps V to W.

Using the Theorem 7.1 we define T as follows

$$T\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \underline{\hspace{3cm}}, T\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \underline{\hspace{3cm}}, T\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \underline{\hspace{3cm}}$$
$$T\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \underline{\hspace{3cm}}, T\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \underline{\hspace{3cm}}, T\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \underline{\hspace{3cm}}.$$

Notice that if we have any vector $v \in V$, we can find where T maps it to in W. Since $v \in V$, we know there are scalars, a, b, c, d, e, f so that

$$v = a\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + e\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + f\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

That is,

$$v = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

Thus, since T is linear,

$$T(v) = a(1) + b(x) + c(x^2) + d(x^3) + e(x^4) + f(x^5) = a + bx + cx^2 + dx^3 + ex^4 + fx^5.$$

8 Transformation Spaces

In Lab # 3, we saw more properties of linear transformations that were useful to recognize other than injectivity and surjectivity. Some radiographic transformations cannot "see" certain objects or we say that those objects are "invisible" to the transformation. In this application, we really want to know if something might be present yet invisible to the transformation we are using. In the case of brain scans, it would be most unhelpful if we cannot see certain abnormalities because they are invisible to the radiographic setup. If something we want to look for in a brain scan is invisible to our current setup, we can adjust the setup to "see" the object we are looking for. Say, for example, we know that the object on the right in Figure 8 is invisible to our radiographic transformation. Then when it is present along with what we expect (Figure from Example 1 from our Linear Tranformation notes), we get the same radiographic transformation b_{exp} and we might go along our merry way, not knowing that something unexpected is present in our object so that instead of what is seen in our Example 1 figure, we actually have what is on the left in Figure 8.

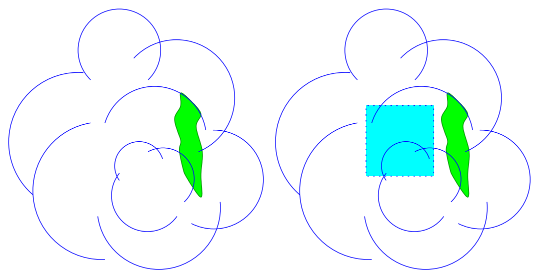


Figure 8:

8.1 Nullspace

Wanting to know which objects are "invisible" to a transformation extends beyond the application of Radiography and Tomography. So, we define the space of all "invisible" objects below.

Definition 8.1. *The* ***nullspace*** *of a linear transformation, $T : V \to W$, is the subset of V that map to $0 \in W$. That is, null(T) =*

Sometimes this is called the *Kernel* of T, or *ker(T)*.

Definition 8.2. *We say that the* ***nullity*** *of a linear transformation, T, is the dimension of the subspace null(T).*

To find the nullity: Find a basis for the nullspace (a spanning set that is linearly independent) and count the number of vectors.

Example 8.1. *Let* $V = \{ax^2 + 2bx + b |\ a, b \in \mathbb{R}\} \subseteq \mathcal{P}_2$. *Define* $\mathcal{F} : V \to \mathcal{P}_1$ *by* $\mathcal{F}(ax^2 + 2bx + b) = 2ax + 2b$. *Find the nullspace of* $\mathcal{F}$.

In Example 8.1, the nullity is __________ because there are no elements in the basis of null$\mathcal{F}$ and we say that $\mathcal{F}$ has a __________ nullspace, or that the nullspace of $\mathcal{F}$ is trivial.

Example 8.2. *Define* $f : \mathcal{M}_{2\times 2} \to \mathbb{R}^4$ *by* $f\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \\ b+a \\ b \\ c \end{pmatrix}$. *Find null*$(f)$.

In Example 8.2 the nullity is __________ because there is __________ element in the basis for the nullspace.
Also notice that there is more than one element of the nullspace.
That means, since $f\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = f\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, but $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$, f is not __________.

Note: The above examples are indeed examples of linear transformations (See Linear Transformation Notes).

Comment: The name "nullspace" seems to imply that this set is a __________ __________. In fact, we have discussed basis and treated it as if it is a vector space in Lab 3 and other examples. The next theorem justifies this treatment.

Theorem 8.1. *Given vector spaces V and W and a linear transformation $T : V \to W$, the nullspace null(T) is a __________ of V.*

Proof. By definition of null(T), we know that null$(T) \subseteq V$. We also know that the zero vector always maps to 0. Thus $0 \in$ null(T). Now, let α and β be scalars and let $u, v \in$ null(T). Then $T(u) = 0$ and $T(v) = 0$. Thus, since T is linear, we have

$$T(\alpha u + \beta v) = \alpha T(u) + \beta T(v) = \alpha \cdot 0 + \beta \cdot 0 = 0.$$

So, $\alpha u + \beta v \in$ null(T). Therefore, null(T) is a subspace of V. □

8.1.1 Nullspace of a Matrix

Recall any linear transformation has a matrix representation. So the Matrix Spaces are defined analogously to linear transformations spaces. And Matrix Reduction makes the calculations easier.

Definition 8.3. *The* ***nullspace****, null(M), of an $m \times n$ matrix M is the nullspace of the corresponding transformation, T, where $T(x) = Mx$ for all $x \in \mathbb{R}^n$.*
That is, null(M) =

Example 8.3. *Given the matrix $M = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ -1 & 0 & 2 \end{pmatrix}$, Determine null$(M)$.*[8]*:*

[8]

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & -1 & 0 \\ -1 & 0 & 2 & 0 \end{pmatrix} \xrightarrow[R_3=r1+r_3]{R_2=-2r_1+r_2} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & -3 & 0 \\ 0 & 1 & 3 & 0 \end{pmatrix} \xrightarrow[R_3=r2+r_3]{R_1=r_2+r_1, R_2=-r_2} \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

8.2 Range Spaces

When considering a transformation, we want to know to which vectors we are allowed to apply the transformation. In the case of a Radiographic transformation, we wonder what is the shape and size of objects/images that the particular radiographic transformation uses. This was all part of our radiographic setup. As with most functions, this set is called the __________. In linear algebra, we consider only sets that are **vector spaces**. So, it is often referred to as the __________ space. There is also a space to which all of the vectors in the domain space map. This space is defined next.

Definition 8.4. *We say that the* ***codomain*** *of a linear transformation, $T : V \to W$, is the vector space to which we map. That is, the codomain of $T : V \to W$ is __________.*

In Examples 8.1 and 8.2 the codomains are $\mathcal{P}_1$ and $\mathbb{R}^4$, respectively. The codomain tends to be much less interesting for a given problem than the set of all thing mapped to. Often not all the vectors in the codomain are mapped to. If they were, then we would say that T is __________.

Definition 8.5. *We say that the* ***range space*** *of a linear transformation, $T : V \to W$, is the subspace of the codomain W that contains all of the outputs from V under the transformation T. That is, ran(T) =*

Definition 8.6. *We say that the* ***rank*** *of a linear transformation, T, is the dimension of the subspace ran(T).*

To find the rank: Find a basis for the range space of T (a spanning set that is linearly independent) and count the number of vectors.

Theorem 8.2. *Let V and W be vector spaces and let $T : V \to W$ be a linear transformation. Then ran(T) is a __________ of W.*

So the range is a __________ space.

Proof. By definition of ran(T), we know that null$(T) \subseteq W$. Since V is a vector space, $\mathbf{0} \in V$. The zero vector always maps to 0, thus $T(\mathbf{0}) = \mathbf{0} \in W$. Thus $0 \in$ ran(T). Now, let α and β be scalars and let $u, w \in$ ran(T). This means there exists a $v_1, v_2 \in V$ such that $T(v_1) = u$ and $T(v_2) = w$ Thus, since T is linear, we have

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2) = \alpha \cdot u + \beta \cdot w.$$

So, there exists a $v \in V$, $v = \alpha v_1 + \beta v_2$, such that $T(v) = \alpha u + \beta w$. Therefore $\alpha u + \beta w \in$ ran(T) and thus, ran(T) is a subspace of W. □

Example 8.4. *Define $\mathcal{F} : V \to \mathcal{P}_1$, where $V = \{ax^2 + \boldsymbol{2bx} + b \mid a, b \in \mathbb{R}\} \subseteq \mathcal{P}_2$. by $\mathcal{F}(ax^2 + \boldsymbol{2bx} + b) = 2ax + 2b$. Show that the range of $\mathcal{F}$ is $\mathcal{P}_1$.*

Since $\text{ran}(\mathcal{F}) = \mathcal{P}_1$, the rank of $\mathcal{F}$ is ___________.
And in fact, this means $\mathcal{F}$ is an ___________ map!
Note: As we have seen the codomain need not be the range. This is only true if our map is onto.

Example 8.5. *Define $f : \mathcal{M}_{2\times 2} \to \mathbb{R}^4$ by $f\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \\ b+a \\ b \\ c \end{pmatrix}$. Find ran$(f)$. What is the rank of f?*

Note: In the example above, the codomain is ___________ and $\text{ran}(f) \neq$ ___________. That means there are elements in $\mathbb{R}^4$ that are not mapped to through f. That is, f is not ___________.

8.2.1 Column/Range Space for Matrix

Definition 8.7. *The* ***range space, ran***(M) *for an $m \times n$ matrix is the set ran$(M) =$*
This is often called the ***Column Space*** *of M, denoted Col(M).*

Note: To find Col(M), we want to find all $b \in \mathbb{R}^m$ so that there is an $v \in \mathbb{R}^n$ so that $Mv = b$.

Process: To find a basis for the column space of a matrix M:
1) Use elementary row operations to write M in RREF.
2) Identify the columns that have pivots.
3) The basis for Col(M) will be the span of the columns of the **original** M associated with the pivot columns.

Example 8.6. *Given the matrix* $M = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ -1 & 0 & 2 \end{pmatrix}$, *Determine ran$(M)$. We have Row reduced M below:*

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & -1 \\ -1 & 0 & 2 \end{pmatrix} \xrightarrow[R_3=r1+r_3]{R_2=-2r_1+r_2} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -3 \\ 0 & 1 & 3 \end{pmatrix} \xrightarrow[R_3=r2+r_3]{R_1=r_2+r_1, R_2=-r_2} \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Aside: An alternate way to solve Example 8.6, is to row reduce the augmented matrix below:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 2 & 1 & -1 & b \\ -1 & 0 & 2 & c \end{array}\right) \xrightarrow[R_3=r_1+r_3]{R_2=-2r_1+r_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & a \\ 0 & -1 & -3 & -2a+b \\ 0 & 1 & 3 & a+c \end{array}\right) \xrightarrow[R_3=r_2+r_3]{R_1=r_2+r_1, R2=-r_2} \left(\begin{array}{ccc|c} 1 & 0 & -2 & -a+b \\ 0 & 1 & 3 & 2a-b \\ 0 & 0 & 0 & -a+b+c \end{array}\right).$$

We then look at the requirements for this system to be consistent and use this to find our basis. Note that our basis spans the same set that we found in Ex 8.6.

8.3 Injectivity and Surjectivity Revisited

Let's consider this discussion again from the point of view of radiography. We saw that some transformations had the property where two objects could give the same radiograph. This particular radiographic transformation would not be __________. Furthermore, we found that if two objects produce the same radiograph, that there difference would then be invisible. That is, the difference is in the __________ of the radiographic transformation. Thus if there is an object that is invisible to the radiographic transformation, any scalar multiple of it will also be invisible. This means that two different objects are invisible, producing the same radiograph, 0. Therefore, the radiographic transformation would not be injective.

Theorem 8.3. *A linear transformation, $T : V \to W$, is injective if and only if __________.*

Proof. Let $T : V \to W$ be a linear transformation. First we will prove that if T is injective then $\text{null}(T) = \{0\}$. Since T is an linear transformation, $T(0) = 0$. Because T is an injection, only one vector can map to 0. Thus $\text{null}(T) = \{0\}$ and there is nothing else that can be in this set.

On the other hand, suppose $\text{null}(T) = \{0\}$. We want to show T is an injection and we will prove this by showing that the contrapositive statement is true. That is, we will prove if T is not injective then $\text{null}(T) \neq \{0\}$. Suppose T is not injective. This means there exists at least 2 vectors, say u and v both in V such that $u \neq v$, but $T(u) = T(v)$. Thus $T(u) - T(v) = 0$. By linear transformation properties, $0 = T(u) - T(v) = T(u - v)$. Thus $(u - v) \in \text{null}(T)$ and therefore $\text{null}(T) \neq \{0\}$. Since we have show that the contrapositive statement is true, the original statement is also true and thus, $\text{null}(T) = \{0\} \implies$ T is injective.

Therefore T is injective $\iff$ $\text{null}(T) = \{0\}$ □

New process to determine Injectivity: We now have a computational way to determine if a linear transformation is injective. Just find the Nullspace (or nullity)!

Recall, also, that we found that there were radiographs that could not be produced from an object given a certain radiographic transformation. This means that there is a radiograph in the codomain that is not mapped to from the domain. If this happens, the radiographic transformation is not __________.

Theorem 8.4. *Let V and W be vector spaces and let $T : V \to W$ be a linear transformation. T is surjective if and only if $\text{ran}(T) =$ __________.*

Proof. Suppose $\text{ran}(T) = W$ then, by definition of $\text{ran}(T)$ if $w \in W$, there is a $v \in V$ so that $f(v) = w$. Thus T is onto. Now, if T is onto, then for all $w \in W$ there is a $v \in V$ so that $T(v) = w$. That means that $W \subseteq \text{ran}(T)$. But, by definition of T and $\text{ran}(T)$, we already know that $\text{ran}(T) \subseteq W$. Thus, $\text{ran}(T) = W$. □

Corollary 8.1. *$T : V \to ran(T)$ is a bijection if and only if $null(T) = \{0\}$.*

Proof. Note since the codomain is equal to the range for this transformation, T is onto.

Suppose T is one-to-one and suppose that $u \in \text{null}(T)$. Then $T(u) = 0$. But, $T(0) = 0$. So, since T is 1-1, we know that $u = 0$. Thus, $\text{null}(T) = \{0\}$.

Now, suppose $\text{null}(T) = \{0\}$. We want to show that T is 1-1. Notice that if $u, v \in V$ satisfy

$$T(u) = T(v)$$

then

$$T(u) - T(v) = 0.$$

But since T is linear this gives us that

$$T(u - v) = 0.$$

Thus, $u - v \in \text{null}(T)$. But $\text{null}(T) = \{0\}$. Thus, $u - v = 0$. That is, $u - v$. So, T is 1-1. □

These theorems give us tools to check injectivity and surjectivity. Let's see how...

Example 8.7. *Define $\mathcal{F} : V \to \mathcal{P}_1$, where $V = \{ax^2 + \boldsymbol{2bx} + b |\ a, b \in \mathbb{R}\} \subseteq \mathcal{P}_2$. by $\mathcal{F}(ax^2 + \boldsymbol{2bx} + b) = 2ax + 2b$. Show that $V \cong \mathcal{P}_1$ using Example 8.1 and 8.4.*

Example 8.8. *Define* $f : \mathcal{M}_{2\times 2} \to \mathbb{R}^4$ *by* $f\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \\ b+a \\ b \\ c \end{pmatrix}$. *Use Example 8.2 and 8.5 to determine whether this map is injective or surjective.*

8.4 The Rank-Nullity Theorem

In each of the last examples of the previous section, we saw that the following theorem holds:

Theorem 8.5. ***The Rank Nullity Theorem:*** *Let* V *and* W *be a vector spaces and let* $T : V \to W$ *be a linear transformation. Then the following is true*

$$\dim V = \underline{\hspace{5cm}} + \underline{\hspace{5cm}}.$$

Recall, $\dim V$= the number of vectors in a basis for V.

Proof. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$.
First, we consider the case when $\text{ran}(T) = \{0\}$. Then, $\text{ran}(T)$ has no basis, so $\text{rank}(T) = 0$. We also know that if $v \in V$ then $T(v) = 0$. Thus, $\mathcal{B}$ is a basis for $\text{null}(T)$ and $\text{nullity}(T) = n$. Thus, $\text{rank}(T)+\text{nullity}(T) = n$.
Next, we consider the case when $\text{null}(T) = \{0\}$. In this case, $\text{null}(T)$ has no basis so $\text{nullity}(T) = 0$. Now, we refer to Lemma 7.1 and Theorem 8.3. We then know that $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is linearly independent and we also know that span $\{T(v_1), T(v_2), \ldots, T(v_n)\} = \text{ran}(T)$. Thus, $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a basis for $\text{ran}(T)$ and $\text{rank}(T) = n$. Thus, $\text{rank}(T)+\text{nullity}(T) = n$.
Finally, we consider the case where $\text{rank}(T) = m$ and $\text{nullity}(T) = k$. Let

$$\mathcal{B}_N = \{v_1, v_2, \ldots, v_k\}$$

be a basis for null(T). And let

$$\mathcal{B}_R = \{T(v_{k+1}), T(v_{k+2}), \ldots, T(v_{k+m})\}$$

be a basis for ran(T). Notice that we know that none of the elements of $\mathcal{B}$ are zero (for otherwise this set would not be linearly independent). So, we know that none of $T(v_1), T(v_2), \ldots, T(v_k) \in \mathcal{B}_R$. We also know that span $\{T(v_1), T(v_2), \ldots, T(v_n)\} = \text{ran}(T)$ so there must be m linearly independent vectors in $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ that form a basis for ran(T). So, our choice of vectors for the basis of $\mathcal{B}_R$ makes sense.
Our goal is to show that $m + k = n$. That is, we need to show that $\mathcal{B} = \{v_1, v_2, \ldots, v_{k+m}\}$. We know that $\{v_1, v_2, \ldots, v_{k+m}\} \subseteq \mathcal{B}$. We need to show that if $v \in \mathcal{B}$ then $v \in \{v_1, v_2, \ldots, v_{k+m}\}$.
Suppose $v \in \mathcal{B}$. Then $T(v) \in \text{ran}(T)$. Suppose $v \notin \text{null}(T)$ (for otherwise, $v \in \mathcal{B}_N$.) Then there are scalars $\alpha_1, \alpha_2, \ldots, \alpha_m$ so that

$$T(v) = \alpha_1 T(v_{k+1}) + \alpha_2 T(v_{k+2}) + \ldots + \alpha_m T(v_{k+m}).$$

So

$$\alpha_1 T(v_{k+1}) + \alpha_2 T(v_{k+2}) + \ldots + \alpha_m T(v_{k+m}) - T(v) = 0.$$

Using that T is linear, we get

$$T(\alpha_1 v_{k+1} + \alpha_2 v_{k+2} + \ldots + \alpha_m v_{k+m} - v) = 0.$$

Thus

$$\alpha_1 v_{k+1} + \alpha_2 v_{k+2} + \ldots + \alpha_m v_{k+m} - v \in \text{null}(T).$$

So either

$$\alpha_1 v_{k+1} + \alpha_2 v_{k+2} + \ldots + \alpha_m v_{k+m} - v = 0$$

or

$$\alpha_1 v_{k+1} + \alpha_2 v_{k+2} + \ldots + \alpha_m v_{k+m} - v \in \text{span } \mathcal{B}_N.$$

If

$$\alpha_1 v_{k+1} + \alpha_2 v_{k+2} + \ldots + \alpha_m v_{k+m} - v = 0$$

Then $v \in \text{span } \mathcal{B}_T$, but this is only true if $v \in \mathcal{B}_T$ because $\mathcal{B}$ is linearly independent and $v \in \mathcal{B}$ and all the elements of $\{v_{k+1}, v_{k+2}, \ldots, v_{k+m}\}$ are also in $\mathcal{B}$. Now, if

$$\alpha_1 v_{k+1} + \alpha_2 v_{k+2} + \ldots + \alpha_m v_{k+m} - v \in \text{span } \mathcal{B}_N$$

then $v \in \mathcal{B}_N$ again because $v \in \mathcal{B}$ and so are all the elements of $\mathcal{B}_N$.
Thus, $v \in \{v_{k+1}, v_{k+2}, \ldots, v_{k+m}\}$. So, we have that $\mathcal{B} = \{v_{k+1}, v_{k+2}, \ldots, v_{k+m}\}$. Thus $k + m = n$. That is, nullity(T)+rank(T) $= n$. $\square$

Important: A quick result of this theorem says that we can separate the basis of V into those that map to 0 and those to a basis of ran(T). More specifically, we have the following result.

Corollary 8.2. *Let V and W be vector spaces and let $T : V \to W$ be a linear transformation. If $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ is a basis for V, then $\{T(v_1), T(v_2), \ldots, T(v_n)\} = \mathcal{B}_r$ where $\mathcal{B}_r$ is a basis of ran(T).*

Proof. We can see in the proof of Theorem 8.5 that $\mathcal{B}$ was split into two sets $\{v_1, v_2, \ldots, v_k\}$ (a basis for null(T)) and $\{v_{k+1}, v_{k+2}, \ldots, v_n\}$, (where $\{T(v_{k+1}), T(v_{k+2}), \ldots, T(v_n)\}$ is a basis for ran(T)). $\square$

The Rank Nullity Theorem is useful in determining rank and nullity, along with proving results about subspaces.

Example 8.9. *Given a linear transformation $T : \mathcal{M}_{2\times 5} \to \mathcal{P}_4$.*
How do we know that T cannot be one-to-one?[9]

.

[9] Notice, we didn't know anything about T except for the spaces from and to which it maps.

Reading Example:

Example 8.10. *Define* $g : V \to \mathbb{R}^3$, *where* $V = \mathcal{P}_1$ *by* $g(ax+b) = \begin{pmatrix} a \\ b \\ a+b \end{pmatrix}$. *Show that* g *is injective and find the range space.*

Notice that $\text{null}(g) = \{ax+b \mid a, b \in \mathbb{R}, g(ax+b) = 0\}$
$= \{ax+b \mid a, b \in \mathbb{R}, \begin{pmatrix} a \\ b \\ a+b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\} = \{ax+b \mid a = 0, b = 0\} = \{0\}$.
Thus, g is injective.
Now we find the range space.
$\text{ran}(g) = \{g(ax+b) \mid a, b \in \mathbb{R}\} = \left\{ \begin{pmatrix} a \\ b \\ a+b \end{pmatrix} \middle| a, b \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$
Notice that since $\text{rank}(g) = 2$ and $\dim \mathbb{R}^3 = 3$, $\mathbb{R}^3 \neq \text{ran}(g)$ and thus g is not onto.
Notice also that $\dim V = 2$, $\text{nullity}(g) = 0$, and $\text{rank}(g) = 2$.

8.4.1 ICE 7: Column Space, Row Space, Nullspace

Let $A = \begin{bmatrix} 1 & 0 & 2 & 3 & 1 \\ 0 & 2 & -1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 1 \\ 3 & -1 & 0 & 2 & 2 \end{bmatrix}$. The Row Reduced Echelon form of A is $\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

Note: A is the matrix representation of some linear transformation.

1. A maps from a __________ dimensional space to a __________ dimensional space.
2. True or False: The column space of A is a subspace of $\mathbb{R}^4$.
   (a) True and I am very confident.
   (b) True, but I am not very confident.
   (c) False, but I am not very confident.
   (d) False and I am very confident.
3. True or False: The nullspace of A is a subspace of $\mathbb{R}^4$.
   (a) True and I am very confident.
   (b) True, but I am not very confident.
   (c) False, but I am not very confident.
   (d) False and I am very confident.
4. What is the dimension of the column space of A?
   (a) 0
   (b) 1
   (c) 2
   (d) 3
   (e) 4
5. What is the dimension of the nullspace of A?
   (a) 0
   (b) 1
   (c) 2
   (d) 3
   (e) 4
6. What is the rank of the linear transformation represented by the matrix A?

   (a) 0
   (b) 1
   (c) 2
   (d) 3
   (e) 4

7. Which columns would form a basis for the column space of A?

where $A = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 2 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 3 & -1 & 0 & 2 \end{bmatrix}$ and the Row Reduced Echelon form of A is $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

(a) All four
(b) The first three
(c) Any three
(d) Any two

8. The column space of a matrix A is the set of vectors that can be created by taking all linear combinations of the columns of A. Is the vector $b = \begin{bmatrix} -4 \\ 12 \end{bmatrix}$ in the column space of the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$?

(a) Yes, since we can find a vector x so that $Ax = b$.
(b) Yes, since $-2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 6 \end{bmatrix} = \begin{bmatrix} -4 \\ 12 \end{bmatrix}$
(c) No, because there is no vector x so that $Ax = b$.
(d) No, because we can't find α and β such that $\alpha \begin{bmatrix} 1 \\ 3 \end{bmatrix} + \beta \begin{bmatrix} 2 \\ 6 \end{bmatrix} = \begin{bmatrix} -4 \\ 12 \end{bmatrix}$
e) More than one of the above.

9. The row space of a matrix A is the set of vectors that can be created by taking all linear combinations of the rows of A. Which of the following vectors is in the row space of the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$?

(a) $x = \begin{bmatrix} -2 & 4 \end{bmatrix}$
(b) $x = \begin{bmatrix} 4 & 8 \end{bmatrix}$
(c) $x = \begin{bmatrix} 0 & 0 \end{bmatrix}$
(d) None of the above
(e) More than one of the above

10. The column space of the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ is

(a) the set of all linear combinations of the columns of A.
(b) a line in $\mathbb{R}^2$.
(c) the set of all multiples of the vector $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$
(d) All of the above
(e) None of the above

9 Inverse Transformations

Example 9.1. *If $T : V \to W$ and M is the matrix representation of T . What can you say about dimension of M if T is injective?*

a) M has the same number of rows and columns
b) The number of rows of M is greater than or equal to the number of columns
c) The number of rows of M is less than or equal to the number of columns
d) Nothing, but I love Linear Algebra!

Example 9.2. *If $T : V \to W$ and M is the matrix representation of T . What can you say about dimension of M if T is onto?*

a) M has the same number of rows and columns
b) The number of rows of M is greater than or equal to the number of columns
c) The number of rows of M is less than or equal to the number of columns
d) Nothing, but I love Linear Algebra!

9.1 Connection to Tomography

In this class, we've actually talked about an application: Radiography/Tomography. Well, we haven't talked about the Tomography part. We will get there. First note that nobody ever really computes the radiograph. This is done using a machine that sends x-rays through something. But what people want to be able to do is to figure out what the object being radiographed looks like. This is the idea behind Tomography. So, we do need to be able to find the object that was produced by a certain transformation.
Suppose we know that $T : \mathbb{R}^n \to \mathbb{R}^m$ is the transformation given by $T(v) = Mv$ where $M \in \mathcal{M}_{m\times n}$. Suppose also that we know that the vector $w \in \mathbb{R}^m$ is obtained by Tv for some v, but we don't know which v. How would we find it? Basically we want a way to work backwards!

9.2 The Inverse Matrix, when $m = n$

Recall, an inverse function of T, is a function such that $T \circ T^{-1}(x) = x$. As we learned in Calculus, it can be difficult to construct inverse functions and sometimes they don't even exist! But when we have a bijective transformation, $T : V \to W$, we <u>can</u> construct an inverse transformation $T^{-1} : W \to V$.
Sometimes it can be difficult to construct these inverse linear transformations, but this isn't too bad when we have a matrix transformation.

Definition 9.1. *Let M be an $n \times n$ square matrix. We say that M has an* ***<u>inverse</u>*** *(and call it M^{-1}) if and only if*

$$MM^{-1} = M^{-1}M = I_{n\times n},$$

where $I_{n\times n}$ is the square identity matrix. If M has an inverse, we say that it is ***<u>invertible</u>***

Algorithm to find the inverse of A.
Row reduce the augmented matrix: $[A|I]$, if A is row equivalent to I, then row reducing gives us $[I|A^{-1}]$. Otherwise, A does not have an inverse.

We can find the inverse of a matrix, A, in Octave by typing:

Example 9.3. *Find M^{-1} when $M = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$.*[10]

Example 9.4. *Find the inverse of $M = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$*
Again, we begin by reducing the augmented matrix $\left(\begin{array}{c|cc} M & e_1 & e_2 \end{array}\right)$ as follows:

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{array}\right).$$

This cannot be reduced further. From the reduced echelon form, we see that there is no vector v so that $Mv = e_1$ nor is there a vector v so that $Mv = e_2$. That is, there is no matrix M^{-1} so that $MM^{-1} = I_{2\times 2}$, the identity matrix. So, M does not have an inverse.

Lemma 9.1. *The inverse of $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$*

Proof. Notice $\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad-bc}\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ □

10 $\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right) \to \left(\begin{array}{cc|cc} 1 & 0 & 1 & -3 \\ 0 & 1 & 0 & 1 \end{array}\right)$

Now, let's look at a bigger example.

Example 9.5. *Find the inverse of* $M = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 2 & 1 \\ 2 & -1 & 1 \end{pmatrix}$.

No matter the size of M, we always start by reducing the augmented matrix $\left(\begin{array}{c|cc} M & e_1 & e_2 \end{array} \right)$.

$$\left(\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 2 & -1 & 1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & -3 & 5 & -2 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & -5 & 2 & -1 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & 14 & -5 & 3 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & -5 & 2 & -1 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & 1 & -5/6 & 1/2 & 1/6 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3/14 & 1/14 & 5/14 \\ 0 & 1 & 0 & 1/14 & -5/14 & -3/14 \\ 0 & 0 & 1 & -5/14 & 3/14 & 1/14 \end{array}\right).$$

Using the above understanding, we see that

$$M \begin{pmatrix} 3/14 & 1/14 & 5/14 \\ 1/14 & -5/14 & -3/14 \\ -5/14 & 3/14 & 1/14 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus, we know that the inverse is $M^{-1} = \begin{pmatrix} 3/14 & 1/14 & 5/14 \\ 1/14 & -5/14 & -3/14 \\ -5/14 & 3/14 & 1/14 \end{pmatrix}$.

Theorem 9.1. ***Properties of Invertible Matrices***

1. If A is invertible, then A^{-1} is invertible and $(A^{-1})^{-1} =$ __________
2. If A is invertible, then A^T is invertible and $(A^T)^{-1} =$ __________
3. If A and B are $n \times n$ invertible matrices, then $A \cdot B$ is invertible and $(AB)^{-1} =$ __________

Proof. (1) Notice $A^{-1}A = AA^{-1} = I$ so $(A^{-1})^{-1} = A$.
(2) By transpose properties, $(A^{-1})^T A^T = A^T(A^{-1})^T = (A(A^{-1}))^T = I^T = I$.
(3) Using the associative property of matrix multiplication, $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A(I)A^{-1} = AA^{-1} = I$ and $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}(I)B = B^{-1}B = I$. □

.

Theorem 9.2. *If A is an $n \times n$ invertible matrix, then for each $b \in \Re^n$, the equation $Ax = b$ has the unique solution $x = A^{-1}b$.*

Proof. Suppose A is an $n \times n$ invertible matrix and let $b \in \mathbb{R}^n$ and consider $A^{-1}b \in \mathbb{R}^n$. Now $A(A^{-1}b) = b$ so a solution does exist. We now want to show this solution is unique. To do this we will let u be an arbitrary solution to $Ax = b$ and show that $u = A^{-1}b$. So if $Au = b$ then $A^{-1}Au = A^{-1}b \Rightarrow u = AA^{-1}b$. Thus $Ax = b$ has the unique solution $x = A^{-1}b$. □

The Invertible Matrix Theorem:
Suppose A is a square $n \times n$. The following statements are either all true or all false:

1. A is invertible
2. A has n pivot positions
3. The Nullspace of A is trivial (The equation $Ax = 0$ has only the trivial solution.)
4. The columns of A form a linearly independent set.
5. The columns of A span $\mathbb{R}^n$
6. The linear transformation represented by A is injective.
7. The linear transformation represented by A is surjective.
8. $AX = b$ has a unique solution
9. A^T is an invertible matrix.
10. There is an $n \times n$ matrix C such that $CA = I$
11. There is an $n \times n$ matrix D such that $AD = I$

Proof: Omitted. Some of these properties we have already proved. Please see suggested textbook if interested in some of the details.

10 The Determinant of a Matrix

The determinant of a square matrix A measures the n dimensional volume of the parallelepiped that is generated by the column vectors of A. In this section, we discuss how to calculate a determinant of a matrix and discuss some information that we can obtain by finding the determinant. *First, note that this section is not a comprehensive discussion of the determinant of a matrix.* There are geometric interpretations that we will not discuss here. Instead this section will present the computation and tough on why we would ever consider finding the determinant. At first, this computation seems lengthy and maybe even more work than its worth. But, we will use the determinant to help us decide the outcome when solving systems of equations or when solving matrix equations. That is, it is often nice for us to know a hint about the solution before we begin the journey through matrix reduction (especially if we have to do these by hand and if they are big). With that, we begin by discussing the properties that we want the determinant of a matrix to have. These properties are all related to matrix reduction steps.

The Properties of the Determinant

Let α be a scalar. We want the determinant of an $n \times n$ matrix, A, to satisfy the following properties:

- $\det(\alpha A) = \alpha^n \det(A)$.
- $\det(A^T) = \det(A)$.
- If B is obtained by performing the row operation, $R_k = \alpha r_j + r_k$ on A, then $\det(B) = \det(A)$.
- If B is obtained by performing the row operation, $R_k = \alpha r_k$ on A, then $\det(B) = \alpha \det(A)$.
- If B is obtained by performing the row operation, $R_k = r_j$ and $R_j = r_k$ on A, then $\det(B) = -1 \cdot \det(A)$.
- If A is in echelon form, then $\det(A)$ is the product of the diagonal elements.

We can use these properties to find the determinant of a matrix by keeping track of the determinant as we perform row operations on the matrix. Let us try an example. We will find the determinant of $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ -1 & 1 & -2 \end{pmatrix}$. Our goal is to reduce A to echelon form all the while keeping track of how the determinant changes.

$$\underset{\det(A)}{\begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ -1 & 1 & -2 \end{pmatrix}} \underset{R_3=r_1+r_3}{\xrightarrow{R_2=-2r_1+r_2}} \underset{\det(A)}{\begin{pmatrix} 1 & 1 & 1 \\ 0 & -3 & 1 \\ 0 & 2 & -1 \end{pmatrix}} \xrightarrow{R_2=r_2+r_3} \underset{\det(A)}{\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 2 & -1 \end{pmatrix}}$$

$$\underset{R_3=2r_2+r_3}{\xrightarrow{R_2=-r_2}} \underset{-\det(A)}{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}} \underset{R_3=-r_3}{\longrightarrow} \underset{\det(A)}{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} = .$$

We will try one more example before giving another method for finding the determinant. We will find the determinant of $A = \begin{pmatrix} 2 & 2 & 2 \\ 1 & 0 & 1 \\ -2 & 2 & -4 \end{pmatrix}$. Again, we will reduce A to echelon form all the while keeping track of how the determinant changes.

$$\underset{\det(A)}{\begin{pmatrix} 2 & 2 & 2 \\ 1 & 0 & 1 \\ -2 & 2 & -4 \end{pmatrix}} \xrightarrow{R_1=\frac{1}{2}r_1} \underset{\frac{1}{2}\det(A)}{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ -2 & 2 & -4 \end{pmatrix}} \xrightarrow[R_3=2r_1+r_3]{R_2=-r_1+r_2} \underset{\frac{1}{2}\det(A)}{\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 4 & -2 \end{pmatrix}}$$

$$\xrightarrow[R_3=4r_2+r_3]{R_2=-r_2} \underset{-\frac{1}{2}\det(A)}{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}} \xrightarrow[R_3=-\frac{1}{2}r_3]{} \underset{\frac{1}{4}\det A}{\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}.$$

Clearly, there has to be another method because, well, I said that we would want to know the determinant before going through all of those steps. Another method for finding the determinant of a matrix is the method called **cofactor expansion**.
Now, if the matrix M is 2×2, it is much easier to compute the determinant:

Determinant of a 2 by 2 Matrix:
Let $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then the determinant is

Notation: $\det M$ is also denoted as $|M|$.

Example 10.1. *What is the determinant of* $\begin{bmatrix} 5 & 4 \\ 1 & 3 \end{bmatrix}$

(a) 4
(b) 11
(c) 15
(d) 19

If M is a bigger matrix, then there's more to do here. Here, we write out the formula given by this method.

Definition 10.1. ***Co-factor Expansion Method to Compute and Define the Determinant:*** *Let $A = (a_{i,j})$ be an $n \times n$ matrix and choose any j so that $1 \leq j \leq n$, then $|A| = \sum_{i=1}^{n}(-1)^{i+j}a_{i,j}|M_{i,j}|$, where $M_{i,j}$ is the sub-matrix of A where the ith row and jth column has been removed.*

Note, we can also expand about a column so that we choose an i so that $1 \leq i \leq n$, then $|A| = \sum_{i=j}^{n}(-1)^{i+j}a_{i,j}|M_{i,j}|$, Notice that if n is large, this process is iterative until the sub-matrices are 2×2. Here are some examples showing what this formula looks like.

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} =$$

Example 2: Determine $|A|$ for $A = \begin{pmatrix} 2 & 2 & 2 \\ 1 & 0 & 1 \\ -2 & 2 & -4 \end{pmatrix}$.

Example 10.2. *Let* $A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{pmatrix}$, *then*

$$|A| = a_{1,1}\begin{vmatrix} a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,2} & a_{4,3} & a_{4,4} \end{vmatrix} - a_{1,2}\begin{vmatrix} a_{2,1} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,3} & a_{4,4} \end{vmatrix} + a_{1,3}\begin{vmatrix} a_{2,1} & a_{2,2} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,4} \end{vmatrix} - a_{1,4}\begin{vmatrix} a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \\ a_{4,1} & a_{4,2} & a_{4,3} \end{vmatrix}.$$

Example 10.3. *What is the determinant of* $\begin{bmatrix} 5 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}$*?*

(a) 0
(b) 15
(c) 24
(d) 26

Example 10.4. *What is the determinant of* $\begin{bmatrix} 5 & 2 & -1 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{bmatrix}$*?*

(a) 0
(b) 6
(c) 15
(d) 22

The determinant of a triangular matrix is

Example 10.5. *What is the determinant of* $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$*?*

(a) 0
(b) 9
(c) 15
(d) none of the above

The determinant of a diagonal matrix is

Now we can add a few more lines to our Big Theorem:

10.1 The Invertible Matrix Theorem

Theorem 10.1. ***The Invertible Matrix Theorem*** *Suppose A is a square $n \times n$. The following statements are either all true or all false:*

1. A is invertible
2. A has n pivot positions
3. The Nullspace of A is trivial (The equation $Ax = 0$ has only the trivial solution.)
4. The columns of A form a linearly independent set.
5. The columns of A span $\mathbb{R}^n$
6. The linear transformation represented by A is injective.
7. The linear transformation represented by A is surjective.
8. $AX = b$ has a unique solution
9. A^T is an invertible matrix.
10. There is an $n \times n$ matrix C such that $CA = I$
11. There is an $n \times n$ matrix D such that $AD = I$
12. $\det A \neq$ __________

Proof: Omitted. Some of these properties we have already proved. Please see suggested textbook if interested in some of the details.

11 Eigenvalues and Eigenvectors

Recall, Linear algebra is basically the study of multivariate linear systems and transformations and, in my opinion, is at its core trying to solve Two Fundamental Problems:

1. Solving $Ax = b$, and
2. Diagonalizing a matrix A (AKA Eigenvalue Problems).

The first problem relates to exploiting linear methods to solve complex and dynamical systems and situations. Basically, it is using one of the best problem-solving techniques mathematicians use, what I like to call the "Wouldn't it be nice if..." approach. That is, real life is a mess and we often have to deal with really complex functions (if we are even lucky enough to have a function at all!) which are difficult to manage. So instead we use linear functions (lines and planes) to approximate or model the complex, real-world situation, which is much easier.

The second problem relates to simplifying our system so that we can more easily solve or approximate systems. The fact that some systems don't have solutions leads directly into the mathematical field of Numerical Analysis and we will dive into some basic numerical analysis in this course. Because so many situations in life can be modeled linearly, Linear Algebra shows up in many topics including (but not exhaustively) "Markov chains, graph theory, correlation coefficients, cryptology, interpolation, long-term weather prediction, the Fibonacci sequence, difference equations, systems of linear differential equations, network analysis, linear least squares, graph theory, Leslie population models, the power method of approximating the dominant eigenvalue, linear programming, computer graphics, coding theory, spectral decomposition, principal component analysis, discrete and continuous dynamical systems, iterative solutions of linear systems, image processing, and traffic flow."[1]

11.1 Eigenvalue and Eigenvectors Motivation

This next section of the class is primarily focused on the second fundamental problem we like to use Linear Algebra to solve. Before we begin, we first discuss some motivation for why we like and seek to find eigenvalues, eigenvectors, and eigenbases.

.

126

11.1.1 ICE 8: Motivation for Eigenvalues and Vectors

1. Compute the product $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

a) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$ c) $\begin{bmatrix} 3 & 3 \end{bmatrix}$ d) $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ e) None of the above.

2. Compute the product $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}^2 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

a) $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$ b) $\begin{bmatrix} 6 \\ 6 \end{bmatrix}$ c) $\begin{bmatrix} 9 \\ 9 \end{bmatrix}$ d) $\begin{bmatrix} 12 \\ 12 \end{bmatrix}$ e) None of the above.

3. Compute the product $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}^4 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

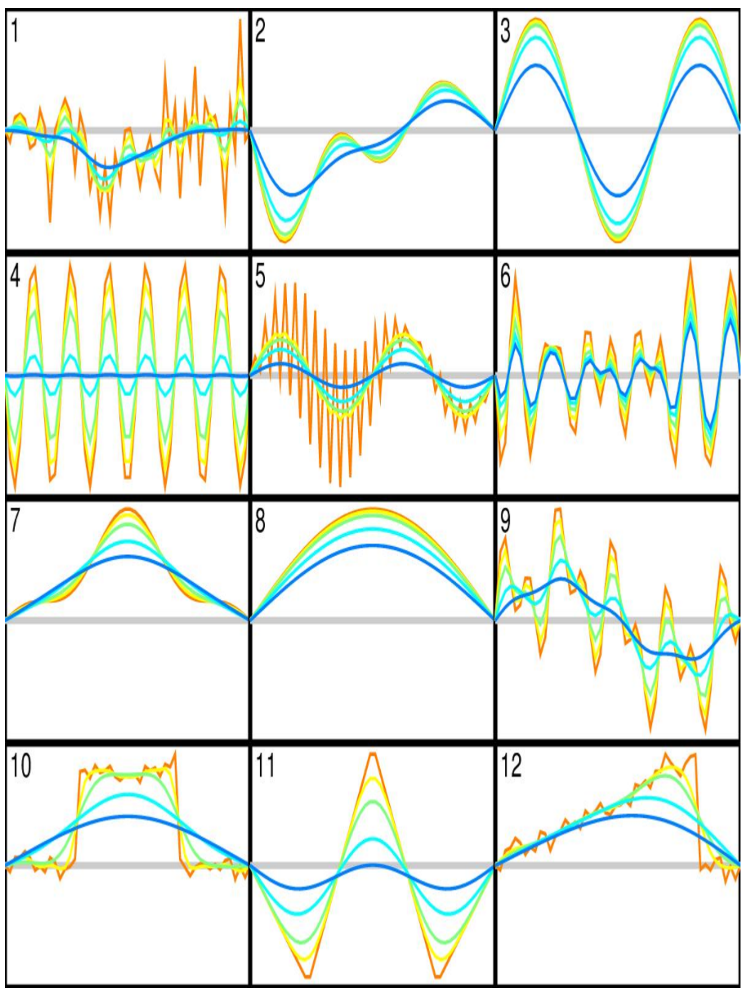
a) $\begin{bmatrix} 27 \\ 27 \end{bmatrix}$ b) $\begin{bmatrix} 81 \\ 81 \end{bmatrix}$ c) $\begin{bmatrix} 243 \\ 243 \end{bmatrix}$ d) $\begin{bmatrix} 729 \\ 729 \end{bmatrix}$ e) None of the above.

4. For any integer n, what will this product be? $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}^n \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

a) $\begin{bmatrix} 3n \\ 3n \end{bmatrix}$ b) $3^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ c) $n^3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ d) $3^n \begin{bmatrix} n \\ n \end{bmatrix}$ e) $\begin{bmatrix} 3 \\ 3 \end{bmatrix}^n$

Wouldn't it be nice if...

5. In the figure below there are 12 different heat state evolution scenarios. The initial heat states are in orange and their corresponding diffusions vary in time based on color from orange (initial) to dark blue (final). Looking specifically at the diffusion behavior for each, create groups that represent similar behavior. Briefly list the criteria you used to group them. **Circle 3 that seem to have the easiest, most predictable diffusion patterns.**



11.2 Eigenvalue and Eigenvectors Motivation Continued

In our motivating application example, we discovered that multiplying a matrix A multiple times can often cause us to have very complex matrix. So when we want to discover how a state, like in heat diffusion, changes after time, we want to find $u(t + n\Delta t) = E^n u(t)$. Wouldn't it be nice if we had special vectors which had the property $Eu = \lambda u$ where λ is a constant/scalar? If this were the case then $u(t + n\Delta t) = E^n u(t) = \lambda^n u(t)$. This will mean that our heat states will only change in amplitude in the heat diffusion (cooling) on a rod. This means that when we apply the diffusion operator to one of these heat states, we get a result that is a scalar multiple of the original heat state. Mathematically, this means we want to find vectors v such that:

$$Ev = \lambda v, \tag{5}$$

for some scalar λ. In other words, these vectors, v, satisfy the matrix equation

$$(E - \lambda I)v = 0. \tag{6}$$

Notice this is a *homogeneous* equation so we know that this equation has a solution. This means that either there is a unique solution (only the trivial solution) or infinitely many solutions. If we begin with a zero heat state (all temperatures are the same everywhere along the rod) then the diffusion is very boring (my opinion, I know) because nothing happens. It would be nice to find a nonzero vector satisfying the matrix Equation (6) because it gets us closer to the possibility of having a basis of these vectors. By the invertible matrix theorem, we know that this equation has a nonzero solution as long $\det(E - \lambda I) =$ __________. Recall, we are going to be very happy to find these special vectors because they help us with predictive modeling (stay tuned for the next few sections!)

11.3 Eigenvalue and Eigenvectors Definition

This property that the special vectors in a heat diffusion have is a very desirable property elsewhere. So, in Linear Algebra, we give these vectors a name.

Definition 11.1. *Let V be a vector space. Given a linear operator(transformation) $L : V \to V$, with corresponding square matrix, and a nonzero vector $v \in V$. If $Lv = \lambda v$ for some scalar λ, then we say v is an* ***eigenvector*** *of L with* ***eigenvalue*** *λ.*

As with the heat states, we see that eigenvectors (with positive eigenvalues) of a linear operator only change amplitude when the operator is applied to the vector. This makes repetitive applications of a linear operator to its eigenvalues very **simple**.

Note: Since every linear operator has an associate matrix, we will treat operators/transformations $L : V \to V$ as a matrix for the rest of this section.

Process: To find the eigenvalues and eigenvectors of a matrix, we need only solve the equation $\det(L - \lambda I) = 0$. That is, the scalars λ so that $\det(L - \lambda I) = 0$.

Definition 11.2. *We call the equation* $\det(L-\lambda I) = 0$. *the* ***characteristic equation*** *and* $\det(L-\lambda I)$ *the* ***characteristic polynomial*** *of L.*

Example 11.1. *Find the eigenvalues of the matrix* $M = \begin{pmatrix} 5 & -3 \\ 6 & -4 \end{pmatrix}$.

Example 11.2. *Find the eigenvectors associated with the eigenvalues found in Ex 11.1 (-1 and 2) of* $M = \begin{pmatrix} 5 & -3 \\ 6 & -4 \end{pmatrix}$.

Definition 11.3. *Given an eigenvalue* λ *for matrix* A*, we sometimes call the nullspace of* $A - \lambda I$ *the* __________*space corresponding to* λ*. A basis for an eigenspace is called an* __________*basis*

Example 11.3. *What is the eigenspace for* $\lambda = 2$ *for* $M = \begin{pmatrix} 5 & -3 \\ 6 & -4 \end{pmatrix}$ *from Ex 11.1?*

Example 11.4. *Find the eigenvalues for* $M = \begin{bmatrix} 5 & 1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

Theorem 11.1. *The eigenvalues of a triangular or diagonal matrix are the entries of its main diagonal.*

Proof. We will prove this for a triangular matrix since a diagonal matrix is a triangular matrix. Without loss of generality, assume A is an upper triangular matrix. Therefore

$$A - \lambda I = \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} & a_{14} & a_{15} & \dots & a_{1n} \\ 0 & a_{22} - \lambda & a_{23} & a_{24} & \dots & \dots & a_{2n} \\ 0 & 0 & a_{33} - \lambda & a_{34} & \dots & \dots & a_{3n} \\ & & \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & a_{nn} - \lambda \end{bmatrix}.$$

Since the determinant of a triangular matrix is the product of the diagonal entries. Solving $\det(A - \lambda I) = 0$ gives the characteristic polynomial $(a_{11} - \lambda)(a_{22} - \lambda) \dots (a_{nn} - \lambda) = 0$ And thus the roots of this equation (the eigenvalues) are $a_{11}, a_{22}, \dots, a_{nn}$. □

Lemma 11.1. *If 0 is an eigenvalue of A, then A is **not** invertible!*

Proof. If 0 is an eigenvalue of A, then for $\lambda = 0$, $0 = \det(A - \lambda I) = \det(A - 0) = \det(A)$. If $\det(A) = 0$, then A is not invertible by The Invertible Matrix Theorem (Theorem 10.1).

□

Example 11.5. *Find an eigenbasis for the eigenspace associated with the eigenvalue 0 for* $M = \begin{bmatrix} 5 & 1 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

11.4 ICE 9: Eigenvalue and Eigenvectors

Two Sided.

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

   (a) Find the eigenvalues for A.

   (b) Find the eigenspaces associated to each eigenvalue.

2. Consider $A = \begin{bmatrix} -2 & 4 \\ 3 & -1 \end{bmatrix}$[11]

   (a) Find the eigenvalues for A.

   (b) Find the eigenspaces associated to each eigenvalue.

[11] Note if $A = \begin{bmatrix} -2 & 4 \\ 3 & 1 \end{bmatrix}$ we have irrational evalues, which are ok, but messier.

11.5 Eigenbasis

Recall, when we are searching for eigenvectors and eigenvalues of an $n \times n$ matrix M, we are really considering the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ defined by $T(v) = Mv$. Then the eigenvectors are vectors in the domain of T that are scaled (by the eigenvalue) when we apply the linear transformation T to them. That is, v is an eigenvector with corresponding eigenvalue $\lambda \in \mathbb{R}$ if $v \in \mathbb{R}^n$ and $T(v) = \lambda v$.

So far, we have agreed that it would be nice if we could find a set of "simple vectors" that formed a basis for the space of heat states.

That is, this will allow us to predict long term behavior of a system in a rather simple way. But we can only do this if we have a basis consisting of eigenvectors. If we can find such a basis, we call this basis a special name, we say it is an "eigenbasis". That is, an *eigenbasis* is just a basis for $\mathbb{R}^n$ made up of eigenvectors. We define an eigenbasis more formally here. In this section, we want to explore when it is the case that we have an eigenbasis. Let's remind ourselves why we want such a basis with an example.

Application to the Heat Diffusion Operator

In the case of heat states, we recognize that if $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ is a basis of these special vectors so that $Ev_i = \lambda_i v_i$ and u_0 is our initial heat state, we can write u in coordinates according to $\mathcal{B}$. That is, there are scalars $\alpha_1, \alpha_2, \ldots \alpha_m$ so that

$$u_0 = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_m v_m.$$

Then, when we apply the diffusion operator to find the heat state, u_1 a short time later, we get

$$\begin{aligned} u_1 =& Eu_0 = E(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_m v_m) \\ =& \alpha_1 E v_1 + \alpha_2 E v_2 + \ldots + \alpha_m E v_m \\ =& \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \ldots + \alpha_m \lambda_m v_m. \end{aligned}$$

So, if we want to find u_k for some time step k, far into the future, we get

$$\begin{aligned} u_k =& E^k u_0 = \alpha_1 E^k v_1 + \alpha_2 E^k v_2 + \ldots + \alpha_m E^k v_m \\ =& \alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 + \ldots + \alpha_m \lambda_m^k v_m. \end{aligned}$$

With this we can predict long term behavior of the diffusion. (You will get a chance to do this in the exercises later.)

Definition 11.4. *Given an $n \times n$ matrix M. If M has n linearly independent eigenvectors, $v_1, v_2, \ldots, v_n$ then $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ is called an* ***eigenbasis*** *of M for $\mathbb{R}^n$.*

First let's recall the following lemma.

Lemma 11.2. *If the dimension of V is n, then any set of n linear independent vectors of V is a basis for V.*

Proof. If the dimension of V is n, then a set of n linear independent vectors will be a minimum spanning set for V and thus is a basis. $\square$

Goal: We want to construct an eigenbasis for M by finding a large enough set of linearly independent eigenvectors. So that the number of eigenvectors is equal to the dimension of M.

Example 11.6. *Determine if $M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$ has an eigenbasis for $\mathbb{R}^3$.*

So we want to find __________, linearly independent eigenvectors for M.

We begin by finding the __________ and __________.
That is, we want to know for which nonzero vectors v and scalars λ does $Mv = \lambda v$.
Eigenvalues =

$\lambda_1 = 2$ In this case we are solving the matrix equation $(M - 2I)v = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0$. Row reducing gives us $\left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$.

$\lambda_2 = 3$ Now, in this case we are solving the matrix equation $(M-3I)v = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0.$

Row reducing gives us $\left(\begin{array}{ccc|c} -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$

To try to find an eigenbasis, take the union of each basis for each eigenspace. In this case our set is:

So can we find an eigenbasis of M for $\mathbb{R}^3$?

Comment: Notice that in each example above, the set made by taking the union of each basis for each eigenspace is a linearly independent set. This theorem justifies why this union will be linearly independent.

Lemma 11.3. *Let M be a matrix and let v be an eigenvector with eigenvalue λ. Then for any scalar c, cv is an eigenvector with eigenvalue λ.*

Proof. Given a matrix M with eigenvalue λ with eigenvector v, then any vector in the eigenspace associated with λ will be an eigenvector for M. Thus any scalar multiple of v will also be in the eigenspace and thus will be an eigenvector associated with λ. □

Theorem 11.2. *Let M be a matrix and let v_1 and v_2 be eigenvectors with eigenvalues λ_1 and λ_2 respectively. If $\lambda_1 \neq \lambda_2$ then $\{v_1, v_2\}$ is linearly independent.*

Proof. Suppose v_1 and v_2 are nonzero eigenvectors with eigenvalues λ_1 and λ_2 respectively. Suppose also that $\lambda_1 \neq \lambda_2$. Then $Mv_1 = \lambda_1 v_1$ and $Mv_2 = \lambda_2 v_2$. By Lemma 11.3, αv_1 and βv_2 are eigenvectors with eigenvalues λ_1 and λ_2 respectively. Let's look at two cases 1.) $\lambda_1 = 0$ and $\lambda_2 \neq 0$ (the case when $\lambda_2 = 0$ and $\lambda_1 \neq 0$ is proved similarly so we won't prove it) and 2.) $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$:

Case 1: If $\lambda_1 = 0$ then by definition, $v_1 \in \text{null}M$. But since $\lambda_2 \neq 0$, $v_2 \notin \text{null}(M)$.
Suppose that $\alpha v_1 + \beta v_2 = 0$. Then $M(\alpha v_1 + \beta v_2) = 0$.
But this means that $\beta M v_2 = 0$. So $\beta = 0$. But then $\alpha v_1 = 0$ and so $\alpha = 0$ and $\{v_1, v_2\}$ is linearly independent.

Case 2: If $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$ and $\alpha v_1 + \beta v_2 = 0$.
Then $M(\alpha v_1 + \beta v_2) = 0$ tells us that $\alpha\lambda_1 v_1 = -\beta\lambda_2 v_2$.
Then $\alpha v_1 = \frac{\lambda_2}{\lambda_1}\beta v_2$. Thus, $\frac{\lambda_2}{\lambda_1}\beta v_2$ is an eigenvector with eigenvalue λ_1 then, $\frac{\lambda_2}{\lambda_1}\beta v_2$ is an eigenvector with eigenvalue λ_2. So

$$M\frac{\lambda_2}{\lambda_1}\beta v_2 = \lambda_2\beta v_2 \quad \text{and} \quad M\frac{\lambda_2}{\lambda_1}\beta v_2 = \frac{\lambda_2^2}{\lambda_1}\beta v_2.$$

So,

$$\lambda_2\beta v_2 = \frac{\lambda_2^2}{\lambda_1}\beta v_2 \quad \text{and} \quad (\lambda_1 - \lambda_2)\beta v_2 = 0.$$

Since, $\lambda_2 \neq \lambda_1$ and $v_2 \neq 0$, we see that $\beta = 0$. Thus,

$$M(\alpha v_1 + \beta v_2) = 0$$

implies $\alpha\lambda_1 v_1 = 0$. So, $\alpha = 0$. Therefore $\{v_1, v_2\}$ is linearly independent.

$\square$

12 Diagonalization

If we have a basis made up of eigenvectors, life is good. Well, that's maybe a bit over reaching. What we found was that if $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ is an eigenbasis for $\mathbb{R}^n$ corresponding to the diffusion matrix E then we can write any initial heat state vector $v \in \mathbb{R}^n$ as $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$. Suppose these eigenvectors have eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. Then with this decomposition into eigenvectors, we can find the heat state at any later time (say k time steps later) by multiplying the initial heat state by E^k. This became an easy computation with the above decomposition because it gives us, using the linearity of matrix multiplication,
$E^k v = E^k(\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n) = \alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 + \ldots + \alpha_n \lambda_n^k v_n$. We can then apply our knowledge of limits from Calculus here to find the long term behavior. That is, the long term behavior is $\lim\limits_{k\to\infty} \alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 + \ldots + \alpha_n \lambda_n^k v_n$. Depending on the eigenvalues. We also discussed how it would be nice to use a change of basis to transform E so that it acts like a diagonal matrix. Recall diagonal matrices are easy to raise to higher powers.

For example if $D = \begin{bmatrix} a & 0 & 0 & ... \\ 0 & b & 0 & ... \\ 0 & 0 & c & ... \\ 0 & 0 & 0 & ... \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$, $D^k =$

There is a formal method for this, called Diagonalization.

Definition 12.1. *Matrices A and B are **similar** if there exists an invertible matrix P such that $A =$ __________*

Definition 12.2. *An $n \times n$ (square) matrix A is **diagonalizable** if there exists $n \times n$ matrices D and P where D is diagonal and P is invertible such that $A =$ __________.*
That is, A is similar to a diagonal matrix.

Why is diagonalization so useful?

Suppose A is diagonalizable, that is, $A = PDP^{-1}$. Calculate A^3.

Example 12.1. *Suppose* $A = PDP^{-1}$ *where* $D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ *and* $P = \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}$. *Calculate* A^3.

Theorem 12.1. *An* $n \times n$ *matrix* A *is diagonalizable if and only if* A *has* n *linearly independent eigenvectors. In fact,* $A = PDP^{-1}$*, where* D *is a diagonal matrix, if and only if the columns of* P *are* n *linearly independent eigenvectors of* A *and the diagonal entries of* D *are the eigenvalues of* A *that correspond respectively to the eigenvectors of* P.

Note: In other words, A is diagonalizable if and only if there are enough eigenvectors to form a basis of $\mathbb{R}^n$. If such a basis exists, we call it an *eigenvector basis* or eigenbasis.

Proof. It is helpful for notation to first notice that if P is any $n \times n$ matrix with columns $v_1, v_2, ..., v_n$ and if D is a diagonal matrix with diagonal entries $\lambda_1, \lambda_2, ..., \lambda_n$, then

$$AP = A \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} = \begin{bmatrix} Av_1 & Av_2 & \dots & Av_n \end{bmatrix} \tag{7}$$

and

$$PD = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \dots & \lambda_n v_n \end{bmatrix} \tag{8}$$

First suppose A is diagonalizable. So there exists invertible P and diagonal D such that $A = PDP^{-1} \Rightarrow AP = PD$ Thus by 7 and 8,

$$\begin{bmatrix} Av_1 & Av_2 & \dots & Av_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \dots & \lambda_n v_n \end{bmatrix}$$

Setting the columns equal to each other gives us that $Av_i = \lambda_i v_i$ for $i = 1, ..., n$. Since P is invertible, $\{v_1, v_2, ..., v_n\}$ are linearly independent and non-zero by The Invertible Matrix Theorem (Thm 10.1). Furthermore, $\lambda_1, \lambda_2, ..., \lambda_n$ are eigenvalues with associated eigenvectors $v_1, v_2, ..., v_n$. Therefore A has n linearly independent eigenvectors (and thus has an eigenbasis -woo!).

On the other hand, suppose A has n linearly independent eigenvectors, $v_1, v_2, ..., v_n$. Define

$$P := \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}.$$

And put

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

Then

$$AP = \begin{bmatrix} Av_1 & Av_2 & \dots & Av_n \end{bmatrix}$$

and

$$PD = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \dots & \lambda_n v_n \end{bmatrix}.$$

By definition of eigenvalues and eigenvectors for A, $AP = DP$ and since the columns of P are the linearly independent eigenvectors, by the Invertible Matrix Theorem (Thm 10.1), P is invertible. Therefore $A = PDP^{-1}$ and thus is diagonalizable. $\square$

Process to Diagonalize a Matrix, A, If Possible

1. **Find the eigenvalues of the matrix.**
   Solve $\det(A - \lambda I) = 0$. That is, solve $|A - \lambda I| = 0$.

2. **Find a spanning set of linearly independent eigenvectors of the matrix.**
   For each e-value, λ_i, find a basis for $\text{Null}(A - \lambda_i I)$.
   Take the union of all bases for all of the eigenspaces.

3. **Determine if you have enough eigenvectors from Step 2, that you can span $\mathbb{R}^n$.**
   Note you can only proceed if there are the same number of eigenvectors as the dimension of A (n) if there are not enough eigenvectors, the matrix is NOT diagonalizable).

4. **Construct P from the eigenvectors from Step 2.**

5. **Construct D from the corresponding eigenvalues (order matters!).**

6. **Check to make sure P and D work by checking if $A = PDP^{-1}$, but it is easier to just check if $AP =$**

Example 12.2. *Diagonalize* $M = \begin{pmatrix} 5 & -3 \\ 6 & -4 \end{pmatrix}$ *if possible. Note in Ex. 11.3, we found that* M *has eigenvalues 2 and -1 with associated eigenspaces spanned by* $\mathcal{E}_2 = \mathit{span}\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$ *and* $\mathcal{E}_{-1} = \mathit{span}\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\}$.

Example 12.3. *Diagonalize* $M = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$ *from Example 11.6 if possible.*

Example 12.4. *Diagonalize* $A = \begin{bmatrix} 4 & 0 & -2 \\ 2 & 5 & 4 \\ 0 & 0 & 5 \end{bmatrix}$, *if possible. One eigenvalue is 4 and Eva has found that the E-space associated with* $\lambda = 4$ *to be* $\mathcal{E}_4$ =*span*$\{\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}\}$ [12]

[12] $\text{null}(A - 4I) = \text{null}(\begin{bmatrix} 0 & 0 & -1 \\ 2 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}) \to$ we solve $\begin{bmatrix} 0 & 0 & -1 & :0 \\ 2 & 1 & 4 & :0 \\ 0 & 0 & 1 & :0 \end{bmatrix} \to \begin{bmatrix} 1 & \frac{1}{2} & 0 & :0 \\ 0 & 0 & 1 & :0 \\ 0 & 0 & 0 & :0 \end{bmatrix}$

So the Eigenspace associated with 4 is $= \{\begin{bmatrix} x \\ y \\ z \end{bmatrix} | x = -\frac{y}{2}, z = 0\} = \{\begin{bmatrix} -\frac{y}{2} \\ y \\ 0 \end{bmatrix} | y \in \mathbb{R}\}$ =span $\{\begin{bmatrix} \frac{-1}{2} \\ 1 \\ 0 \end{bmatrix}\}$ =span $\{\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}\}$

12.1 ICE 10 -Diagonalization

1. Eva and Archer want to Diagonalize a **2 by 2** matrix A. Dr. Harsy has computed the following eigenvalues and associated eigenvectors for A: $\lambda_1 = 3$ with associated eigenvector $\{\begin{bmatrix}1\\1\end{bmatrix}\}$ and $\lambda_2 = 16$ with associated eigenvector $\begin{bmatrix}1\\-1\end{bmatrix}$. Which of the following would work for the diagonalization of A:

   a) $P = \begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$, $D = \begin{bmatrix}16 & 0\\0 & 3\end{bmatrix}$

   b) $P = \begin{bmatrix}1 & 1\\-1 & 1\end{bmatrix}$, $D = \begin{bmatrix}16 & 0\\0 & 3\end{bmatrix}$

   c) $P = \begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$, $D = \begin{bmatrix}3 & 0\\0 & 16\end{bmatrix}$

   d) $P = \begin{bmatrix}1 & 1\\-1 & 1\end{bmatrix}$, $D = \begin{bmatrix}3 & 0\\0 & 16\end{bmatrix}$

   e) More than one of the above.

2. Now Eva and Archer want to Diagonalize a **3 by 3** matrix A. Dr. Harsy has computed the following eigenvalues and associated eigenvectors for A: $\lambda_1 = 3$ with associated eigenbasis $\{\begin{bmatrix}1\\1\\0\end{bmatrix}\}$ and $\lambda_2 = 16$ with associated eigenbasis $\{\begin{bmatrix}0\\1\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\}$. Which of the following would work for the diagonalization of A:

   a) $P = \begin{bmatrix}1 & 1 & 0\\1 & 1 & 1\\0 & 1 & -1\end{bmatrix}$, $D = \begin{bmatrix}3 & 0 & 0\\0 & 3 & 0\\0 & 0 & 16\end{bmatrix}$

   b) $P = \begin{bmatrix}1 & 1 & 0\\1 & 1 & 1\\0 & 1 & -1\end{bmatrix}$, $D = \begin{bmatrix}3 & 0 & 0\\0 & 3 & 0\\0 & 0 & 16\end{bmatrix}$

   c) $P = \begin{bmatrix}1 & 1 & 0\\1 & 1 & 1\\0 & 1 & -1\end{bmatrix}$, $D = \begin{bmatrix}16 & 0 & 0\\0 & 16 & 0\\0 & 0 & 3\end{bmatrix}$

   d) $P = \begin{bmatrix}1 & 1 & 0\\1 & 1 & 1\\0 & 1 & -1\end{bmatrix}$, $D = \begin{bmatrix}3 & 0 & 0\\0 & 16 & 0\\0 & 0 & 16\end{bmatrix}$

   e) More than one of the above.

3. Now Eva and Archer want to Diagonalize a **3 by 3** matrix A. Dr. Harsy has computed the following eigenvalues and associated eigenvectors for A: $\lambda_1 = 3$ with associated eigenbasis $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$ and $\lambda_2 = 16$ with associated eigenbasis $\left\{ \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$. Which of the following would work for the diagonalization of A:

a) $P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$, $D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 16 \end{bmatrix}$

b) $P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$, $D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 16 \end{bmatrix}$

c) $P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$, $D = \begin{bmatrix} 16 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

d) More than one of the above.

e) A is not diagonalizable.

4. Suppose $A = PDP^{-1}$, write an expression for A^8 involving P and D.

13 Markov Chains

It turns out, we already kinda know about Markov Chains [13]

In our Heat Equation Discussion about the motivation for eigenvectors, we discussed how we could easily predict long term behavior of our heat state by calculating $E^k v_0$, where E is our diffusion matrix and v_0 is our initial heat state. This can get really difficult to calculate, especially if E is a large matrix. On the other hand, if we had a basis made up of eigenvectors, life is easier. When we have such a basis, our matrix E is diagonalizable. That is we can write $E = PDP^{-1}$. Then $E^k v_0 = PD^kP^{-1}v_0$. We can then also apply our knowledge of limits from Calculus here to find the long term behavior. That is, the long term behavior is

$$\lim_{k\to\infty} E^k v_0 = \lim_{k\to\infty} PD^kP^{-1}v_0 =$$

$$\lim_{k\to\infty} \alpha_1 E^k v_1 + \alpha_2 E^k v_2 + \ldots + \alpha_m E^k v_m = \lim_{k\to\infty} \alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 + \ldots + \alpha_m \lambda_m^k v_m$$

We see that this limit really depends on the size of the eigenvalues.

We can see how this equation allows us to find different heat states for a given rod. That is, $u_2 = Eu_1, u_3 = Eu_2 = E^2u_1, u_4 = Eu_3 = E^3u_1$, etc. This process of transitioning between "states" of a system over discrete time steps has many applications including:

- Movement of people between regions
- States of weather
- Movement between positions on a board games
- Your score in betting games like Blackjack
- Movements between webpages
- Positions of the runners on base and the number of outs in baseball
- Population changes of an ecosystem
- Subscription changes for amenities

Example 13.1. *We will divide the class into 3 groups, A,B, & C.*

- *1/3 of group A goes to group B and 1/3 of group A goes to group C.*
- *1/4 of group B goes to group A and 1/4 of group B goes to C.*
- *1/2 of group C goes to group B.*

What happens after a few state transitions?

[13] Some Examples from *When Life is Linear* and http://www.slideshare.net/leingang/lesson-11-markov-chains

13.1 Introduction to Markov Chains

Definition 13.1. *A **Markov chain** or **Markov Process** is a process in which the probability of the system being in a particular state at a given time period depends only on its state at previous time period.*
Note: *There is NO PAST in a Markov Chain, just the here and now!*

Common questions about a Markov Process:

- What is the long term behavior of the process?
- Is there a long-term stability to the process?
- What is the probability of transitions from state to state over multiple observations
- Can we hit an equilibrium state?

Example 13.2. *Suppose on any given class day you wake up and decide whether to come to class. If you went to class the time before, you're 70% likely to go today, and if you skipped the last class, you're 80% likely to go today. Some questions you may ask include:*
If I go to class on Tuesday, how likely am I to go to class on Thursday? Next Thursday?
Assuming the class is infinitely long (oh my!), what portion of the class will I approximately attend?

To answer these questions, we first will set up a Transition Matrix or Stochastic Matrix.

Definition 13.2. *Suppose a system has n possible states. The **Transition or Stochastic Matrix** for this system is given by $T = (t_{ij})$ where for each i and j, t_{ij} is the probability of switching from state j to state i.*

What is our transition matrix for Ex 13.2: Skipping Class?
If you went to class the time before, you're 70% likely to go today, and if you skipped the last class, you're 80% likely to go today.

What is our transition matrix for Ex. 13.1: Markov Dance?
Recall, 1/3 of group A goes to group B and 1/3 of group A goes to group C. 1/4 of group B goes to group A and 1/4 of group B goes to C. 1/2 of group C goes to group B.

What properties do these matrices have?

Properties of Stochastic Matrices:

- All entries are __________
- The columns add up to __________

Example 13.3. *Is our Diffusion Matrix from our Heat Lab a transition matrix? Recall* $0 \leq \delta < 1$

$$E = \begin{bmatrix} 1-2\delta & \delta & 0 & 0 & \dots \\ \delta & 1-2\delta & \delta & 0 & \dots \\ 0 & \delta & 1-2\delta & \delta & \\ \vdots & \vdots & \vdots & & \ddots \end{bmatrix}$$

Definition 13.3. *A Transition Matrix (for a corresponding Markov chain) is called* ***regular*** *if for some* $k \geq 1$, T^k *has all strictly positive entries. Note that this means there is a positive probability of eventually moving from every state to every state.*

Example: $T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is a ____________________ matrix but not __________.

Is A a Transition Matrix? Is it Regular?

$A = \begin{bmatrix} .5 & .5 & .5 \\ .5 & 0 & .5 \\ 0 & .5 & 0 \end{bmatrix}$, $A^2 = \begin{bmatrix} .5 & .5 & .5 \\ .25 & .5 & .25 \\ .25 & 0 & .25 \end{bmatrix}$, $A^3 = \begin{bmatrix} .5 & .5 & .5 \\ .375 & .25 & .375 \\ .125 & .25 & .125 \end{bmatrix}$.

Definition 13.4. *The* ***state vector*** *for a Markov chain with* n *possible state at a given time step* k *is denoted* $x^k = \begin{bmatrix} p_1^k \\ p_2^k \\ p_3^k \\ \vdots \\ p_n^k \end{bmatrix}$
where p_i^k *is the probability that the system is in state* i *at time-step* k*. These vectors are also called* __________ *vectors since they add up to 1 and are nonnegative.*

In order to find our state vectors for our two examples, we need more information.
For our skipping class example, we need to decide what is our initial vector state. What would be a good choice for this?

Find the initial state for our Markov Activity if we started with 20 students in A, 10 in B, and 10 in C.

Theorem 13.1. *If T is the transition matrix for a Markov chain, the then the state vector for the $(k+1)$th time step, denoted __________, can be determined from __________ by*

Proof: Omitted. See supplementary textbook.

Example 13.4. *Suppose you go to class on Tuesday, what's the probability that you will go to the next class? What about the class 3 days from now?*

Connection: Remember, if we want to calculate T^k for large k's (Lay's Linear Algebra Book says for $k > 30$), we may want to compute each x^k from the previous state rather than calculate T^k.

Even better, it would be nice if we can ____________________ T!

13.2 Steady State Vectors

As we mentioned before, it may be useful to determine the end behavior of a system. In other words, we want to determine $\lim_{k\to\infty} T^k x^0$.

Theorem 13.2. *If T is an $n \times n$ regular stochastic matrix, the T has a unique vector $\mathbf{u}$. That is if x^0 is the initial state, the $\lim_{k\to\infty} T^k x^0 =$ __________. That is, our Markov chain of states $\{x^i\}$ converges to $\mathbf{u}$.*

Proof: Omitted. See supplementary textbook.

Definition 13.5. *The vector $\mathbf{u}$ described above is called a steady-state vector.*

What does this mean?
$x^{i+1} = Tx^i$ so for large i...

Theorem 13.3. *The steady-state vector $\mathbf{u}$ is the unique probability vector satisfying the matrix equation:*

That is, **u** is an ____________________ for the associated ____________________ = __________!!

Proof: Omitted. See supplementary textbook.

So how can we find a steady-state vector?

Example 13.5. *Find the steady-state vector for the Skipping Ex. 13.2. Recall,* $T = \begin{bmatrix} 0.7 & 0.8 \\ 0.3 & 0.2 \end{bmatrix}$

13.3 Application to Google's Page Rank Model

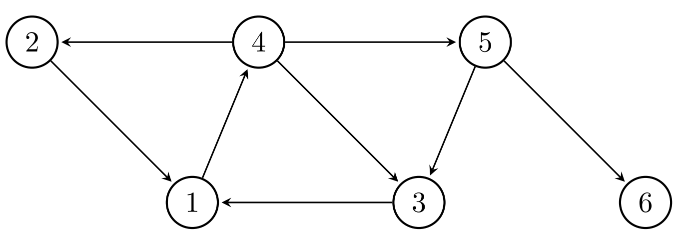
There are billions of web pages in the interwebs. How can we possibly determine the quality of a page? Google's Page Rank Model was developed by Google founders Larry Page and Sergey Brin in order to solve this problem. They determined the popularity of a web page by modeling Internet activity. That is, the frequency of time spent of time spent on a page yields that page's PageRank. So what is this model? Does Google have spies (probably cats) tracking your every move on the Internet?

Like every model, we start with some assumptions.
PageRank Model Assumptions:

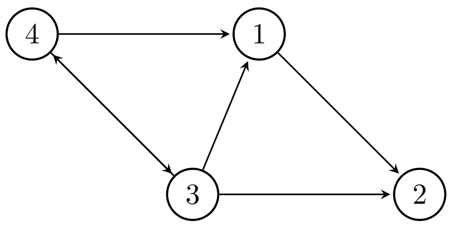
- Google treats everyone as a random surfer and assumes you will randomly choose links to follow.
- If there are links on the page you currently are on:
  - There is a 85% chance you will follow a hyperlink on a page that you are currently on.
  - There is a 15% chance you will jump to any web page in the network (with uniform probability),
- If there are no links on the web page you are currently on (this page is called a "dangling node"), you are equally likely to jump anywhere on the Internet.
  note you can go right back to your current page which is necessary so we can have a regular transition matrix

Example system (on your HW!):

Which page is the dangling node?



Example 13.6. *Below is a directed graph representing six web pages representing links to other pages. Create the Transition Matrix for this Markov process.*



b) Notice T is regular, so we can find a unique steady state vector. Let's find this vector.

13.4 Ice 11 -Markov Chains

1. A small, isolated town has two grocery stores, Archer's Market and Eva's Shoppe. While some customers are completely loyal to one store or another, there is another group of customers who change their shopping habits each month. Of the shoppers who favor Archer's Market one month, only 70% will still shop there the following month, while Eva's Shoppe retains 78% of its customer base each month. Everyone in the town shops at one of the two stores, and no one from out of town ever shops at either store. If Archer's Market currently has 2500 customers and Eva's Shoppe has 1900 customers, how many customers will Archer's Market have next month?

   (a) 418
   (b) 1750
   (c) 2168
   (d) 3080

2. Referring to the scenario in the previous question, what will the product $\begin{bmatrix} 0.70 & 0.22 \\ 0.30 & 0.78 \end{bmatrix} \cdot \begin{bmatrix} 2500 \\ 1900 \end{bmatrix}$ tell us?

   (a) This product will tell us the percentage of customers that will switch from one store to the other store next month.

   (b) This product will tell us the number of customers who will shop at each store next month.

   (c) This product will tell us the total number of customers who switched stores this month.

   (d) This product doesn't have any meaning.

3. Continuing the scenario from the previous questions, what does the (2, 1)-entry of the matrix $\begin{bmatrix} 0.70 & 0.22 \\ 0.30 & 0.78 \end{bmatrix}^3$ represent?

(a) This represents the probability that a customer will switch from Archer's Market to Eva's Shoppe between months 3 and 4.
(b) This represents the probability that a customer will switch from Eva's Shoppe to Archer's Market between months 3 and 4.
(c) This represents the probability that a customer who currently shops at Archer's Market will be shopping at Eva's Shoppe three months from now.
(d) This represents the probability that a customer who currently shops at Eva's Shoppe will be shopping at Archer's Market three months from now.

4. Suppose we found that the steady-state (equilibrium) vector for this shop example is $\begin{bmatrix} \frac{11}{26} \\ \frac{15}{26} \end{bmatrix}$. What does the steady-state vector mean in the context of Archer's Market and Eva's Shoppe?

(a) In the long-run, the probability of staying at Archer's Market will be 11/26 and the probability of switching to Eva's Shoppe will be 15/26.
(b) In the long-run, the probability of shopping at Archer's Market will be 11/26 and the probability of shopping at Eva's Shoppe will be 15/26.
(c) In the long-run, Archer's Market will approach 11/26 of the market share, while Eva's Shoppe will approach 15/26 of the market share.
(d) I like cats!

5. Verify that $\begin{bmatrix} \frac{11}{26} \\ \frac{15}{26} \end{bmatrix}$ is a steady state vector for this Markov Process.

14 Orthogonality

So far, we have talked about how lovely life is when we have a matrix (often representing a dynamical system) that is diagonalizable. Unfortunately, we don't always have this property. Sometimes we have matrices that don't have a complete and spanning eigenbasis and often we don't have square matrices. Luckily we have a way to deal with these less than ideal situations. This requires us to look at the matrix A^TA (which will be square) and create something called an orthogonal basis. So first let's talk about orthogonal vectors. Note, we are talking about this mainly because we will discuss a process which will allow us to find an orthogonal basis.

14.1 Orthogonal Vectors

Recall, when we discussed the Massey Method for sports ranking, we defined the following two terms:

Definition 14.1. *The **inner product** or **dot product** between two vectors* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$*, denoted* $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \cdot \mathbf{v} = \mathbf{v}^T \cdot \mathbf{u} = u_1 \cdot v_1 + u_2 \cdot v_2 + u_3 \cdot v_3 + ... + u_n \cdot v_n$

Definition 14.2. *The **length** or **norm** of a vector* $\boldsymbol{v} \in \mathbb{R}^n$ *is denoted* $||\mathbf{v}|| = \sqrt{v_1^2 + v_2^2 + ...v_n^2}$

We now have a few more definitions which are necessary for

Definition 14.3. *Vectors* $\boldsymbol{u}$ *and* $\boldsymbol{v} \in \mathbb{R}^n$ *are **orthogonal** if* $\mathbf{u} \cdot \mathbf{v} =$ __________

Definition 14.4. *Given a set of vectors* $S = \{u_1, u_2, u_3, ..., u_m\}$*, We say S is an orthogonal set if* $\mathbf{u_i} \cdot \mathbf{u_j} = 0$ *<u>for every</u>* i, j *<u>in the set</u>.*

Example 14.1. *a) Show that* $S = \{\begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}\}$ *is an orthogonal set.*

Solution: *We must check each of these 3 vectors pairwise which means we have 3 dot products to check.*
$\mathbf{u_1} \cdot \mathbf{u_2} = (-2)(1) + (1)(-1) + (-1)(-3) = -2 - 1 + 3 = 0.$
$\mathbf{u_2} \cdot \mathbf{u_3} = (1)(4) + (-1)(7) + (-3)(-1) = -4 - 7 + 3 = 0.$
$\mathbf{u_1} \cdot \mathbf{u_3} = (-2)(4) + (1)(-7) + (-1)(-1) = -8 + 7 + 1 = 0.$

b) What if we add $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$*?*

Solution: *We need to check the dot product of this new vector with each of the 3 vectors in S.*
$\mathbf{u_1} \cdot \mathbf{v} = (-2)(0) + (1)(2) + (-1)(2) = 0 + 2 - 2 = 0.$
$\mathbf{u_2} \cdot \mathbf{v} = (1)(0) + (-1)(2) + (-3)(2) = 0 - 2 - 6 = 0 - 8 \neq 0.$
Since $\mathbf{u_2} \cdot \mathbf{v} \neq 0$*, we know that* $S = \{\begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}\}$ *is not an orthogonal set and we do not need to check* $\mathbf{u_3} \cdot \mathbf{v}$*.*

Example 14.2. *Which of the following sets of vectors is not an orthogonal set?*

(a) (1, 1, 1), (1, 0,-1)
(b) (2, 3), (-6, 4)
(c) (3, 0, 0, 2), (0, 1, 0, 1)
(d) (0, 2, 0), (-1, 0, 3)
(e) $(\cos\theta, \sin\theta)$, $(\sin\theta, -\cos\theta)$

Theorem 14.1. *Suppose* $\boldsymbol{u}$ *and* $\boldsymbol{v}\in \mathbb{R}^n$*, then* $\mathbf{u}\cdot\mathbf{v} = 0$ *if and only if* $||\mathbf{u}+\mathbf{v}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2$.

Proof. Let $\mathbf{u}$ and $\mathbf{v}$ be orthogonal vectors, so by definition $\mathbf{u}\cdot\mathbf{v} = 0$. Then $||u+v||^2 = ||\mathbf{u}||^2 + 2\mathbf{u}\cdot\mathbf{v} + ||\mathbf{v}||^2 = ||\mathbf{u}||^2 + 0 + ||\mathbf{v}||^2$. □

Let's verify this theorem for $\begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}$

(Recall, we showed that these two matrices were orthogonal in Ex 14.1.)

$\mathbf{u}+\mathbf{v} = \begin{bmatrix} -1 \\ 0 \\ -4 \end{bmatrix}$,

$||\mathbf{u}+\mathbf{v}||^2 = (\sqrt{(-1)^2 + 0^2 + (-4)^2})^2 = \sqrt{17}^2 = 17$

$||\mathbf{u}||^2 = \sqrt{2^2 + 1^2 + 1^2}^2 = 6$ and $||\mathbf{v}||^2 = \sqrt{1^2 + 1^2 + 3^2}^2 = 11$

So $||\mathbf{u}||^2 + ||\mathbf{v}||^2 = 6 + 11 + 17$

Note: In general, $||\mathbf{u}+\mathbf{v}||^2 \neq ||\mathbf{u}||^2 + ||\mathbf{v}||^2$.

14.2 Orthogonal Spaces

Definition 14.5. *Let V be a vector space. We say a vector* $\boldsymbol{u}$ *is* ***orthogonal*** *to V (the whole space of V) if* $\mathbf{u}\cdot\vec{v} = 0$ *for every vector $v \in V$.*

Definition 14.6. *The set of all such vectors u that are orthogonal to V is called the* ***orthogonal complement of V*** *and is denoted by $V^\perp$.*

Note: $V^\perp$ is a vector space. You could check the closure property (if $\mathbf{v},\mathbf{u} \in V^\perp$, then $\alpha\mathbf{v} \in V^\perp$ and $\mathbf{v}+\mathbf{u} \in V^\perp$).

Theorem 14.2. *Let $\mathcal{B} = \{v_1, v_2, ...v_n\}$ be a basis for V then $u \in V^\perp$ (that is, $\mathbf{u}$ is orthogonal to every vector in V) if and only if $u \cdot v_i = 0$ for all $v_i \in \mathcal{B}$.*

Heuristic Proof: If a vector is orthogonal to all of the "building blocks" (basis vectors), then it is orthogonal to everything you can build/span with those vectors. Feel free to look at our suggested textbook for more details.

Example 14.3. *Suppose V =span$\{\begin{bmatrix}1\\1\\0\\2\end{bmatrix}, \begin{bmatrix}1\\1\\1\\-1\end{bmatrix}, \begin{bmatrix}2\\0\\1\\1\end{bmatrix}\}$. How can we find a basis for $V^\perp$?*

Solution: Let $\mathbf{v_1} = \begin{bmatrix}1\\1\\0\\2\end{bmatrix}, \mathbf{v_2} = \begin{bmatrix}1\\1\\1\\-1\end{bmatrix}, \mathbf{v_3} = \begin{bmatrix}2\\0\\1\\1\end{bmatrix}$. We want to find a linearly independent spanning set for the set of all vectors which are orthogonal to V. We don't want to check every vector individually, so instead we will find all vectors $\mathbf{u} = \begin{bmatrix}u_1\\u_2\\u_3\\u_4\end{bmatrix}$ such that

$$\mathbf{u} \cdot \mathbf{v_1} = 0 \Rightarrow u_1(1) + u_2(1) + u_3(0) + u_4(2) = 0 \quad (9)$$
$$\mathbf{u} \cdot \mathbf{v_2} = 0 \Rightarrow u_1(1) + u_2(1) + u_3(1) + u_4(-1) = 0 \quad (10)$$
$$\mathbf{u} \cdot \mathbf{v_3} = 0 \Rightarrow u_1(2) + u_2(0) + u_3(1) + u_4(1) = 0 \quad (11)$$
$$(12)$$

We can solve this system of equations by setting up the augmented matrix: $\left(\begin{array}{cccc|c} 1 & 1 & 0 & 2 & 0 \\ 1 & 1 & 1 & -1 & 0 \\ 2 & 0 & 1 & 1 & 0 \end{array}\right)$.

Note: This system is just $\left(A^T \mid 0\right)$ where $A = \begin{bmatrix}v_1 & v_2 & v_3\end{bmatrix} = \begin{bmatrix}1 & 1 & 2\\1 & 1 & 0\\0 & 1 & 1\\2 & -1 & 1\end{bmatrix}$.

Solving this system gives us: $\left(\begin{array}{cccc|c} 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0 \end{array}\right)$ which gives us the solution set
$\mathbf{u_1} = -2\mathbf{u_4}$, $\mathbf{u_2} = 0$, $\mathbf{u_3} = -3\mathbf{u_4}$, $\mathbf{u_4} = \mathbf{u_4}$.

Thus $V^\perp$ =span$\{\begin{bmatrix}-2\\0\\-3\\1\end{bmatrix}\}$. Since this vectors space is spanned by 1 vector, it is automatically linearly independent and thus is a basis.

Example 14.4. *Suppose V =span$\{\begin{bmatrix}-2\\1\\-1\end{bmatrix}, \begin{bmatrix}1\\-1\\-3\end{bmatrix}, \begin{bmatrix}4\\7\\-1\end{bmatrix}\}$. Find a basis for $V^{\perp}$.*

Solution: We can solve this by setting up the augmented matrix: $\left(\begin{array}{ccc|c}-2 & 1 & -1 & 0\\ 1 & -1 & -3 & 0\\ 4 & 7 & -1 & 0\end{array}\right)$ Row reducing gets us $\left(\begin{array}{ccc|c}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{array}\right)$ which gives the solution $\mathbf{u_1} = 0$, $\mathbf{u_2} = 0$, $\mathbf{u_3} = 0$ and thus $V^{\perp} = \{\mathbf{0}\}$.

14.3 Orthogonal Bases

Theorem 14.3. *A set of nonzero orthogonal vectors is a linearly independent set of vectors.*

Proof. Suppose $\{\mathbf{u_1}, \mathbf{u_2}, ..., \mathbf{u_n}\}$ is a set up nonzero orthogonal vectors. If $\mathbf{0} = a_1\mathbf{u_1} + a_2\mathbf{u_2} + ... + a_n\mathbf{u_n}$, then dotting both sides with $\mathbf{u_1}$ gives us
$0 = \mathbf{0} \cdot \mathbf{u_1} = a_1\mathbf{u_1} \cdot \mathbf{u_1} + a_2\mathbf{u_2} \cdot \mathbf{u_1} + ... + a_n\mathbf{u_n} \cdot \mathbf{u_1}$. Since $\mathbf{u_1}$ is orthogonal to $\mathbf{u_i}$ for $i = 2, ..., n$ $\mathbf{u_i} \cdot \mathbf{u_1} = 0$ for $i = 2, ..., n$. Thus, $0 = a_1\mathbf{u_1} \cdot \mathbf{u_1} + 0 + ... + 0$. Since $\mathbf{u_1} \neq \mathbf{0}$, the only solution to this equation is if a_1. We can continue to do this same method for $\mathbf{u_i}$ for $i = 2, ..., n$ to show $a_1 = a_2 = ... = a_n = 0$. Therefore our set is linearly independent.

□

Definition 14.7. *A basis made up of orthogonal vectors is called an* ***orthogonal basis.***

Why is this type of basis useful? Orthogonal sets are automatically linearly independent and also because we will have a nice way to write any arbitrary vector as a linear combination of our orthogonal basis. Orthogonal bases are also very stable.

Definition 14.8. *A set of vectors $\{v_1, v_2, v_3, ..., v_n\}$ is* ***orthonormal*** *if the set is orthogonal and $||v_i|| = 1$ for $i = 1, ..n$.*

Normalizing Process: Given an orthogonal set $\{v_1, v_2, v_3, ..., v_n\}$, we can normalize these vectors by dividing each vector by its norm.

Theorem 14.4. *Suppose V has an orthogonal basis $\{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}\}$, then any vector, $\mathbf{u} \in V$ can be written as $\mathbf{u} = a_1\mathbf{v_1} + a_2\mathbf{v_2} + ... + a_n\mathbf{v_n}$ where*
$a_i = \dfrac{\mathbf{v_i} \cdot \mathbf{u}}{||\mathbf{v_i}||^2} = \dfrac{\mathbf{v_i} \cdot \mathbf{u}}{\mathbf{v_i} \cdot \mathbf{v_i}}$

Proof. Let $\mathbf{u} \in V$ and suppose V has an orthogonal basis $\{\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v_n}\}$. Then $\exists a_1, a_2, ..., a_n$ such that $\mathbf{u} = a_1\mathbf{v_1} + a_2\mathbf{v_2} + ... + a_n\mathbf{v_n}$. Using a method similar to Thm 14.3, dotting both sides by nonzero $\mathbf{v_1}$ gives us $\mathbf{u} \cdot \mathbf{v_1} = a_1\mathbf{v_1} \cdot \mathbf{v_1}$. Since $\mathbf{v_1} \cdot \mathbf{v_1} \neq 0$, we can solve for a_1. Thus $a_1 = \dfrac{\mathbf{v_1} \cdot \mathbf{u}}{\mathbf{v_1} \cdot \mathbf{v_1}}$.
Similarly we can solve for a_i for $i = 2, .., n$ and determine that $a_i = \dfrac{\mathbf{v_i} \cdot \mathbf{u}}{\mathbf{v_i} \cdot \mathbf{v_i}}$. □

Theorem 14.5. *Given a subspace $W \subseteq V$, $V = W + W^{\perp}$. This means any $\mathbf{v} \in V$ can be written as $\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$, where $\mathbf{w} \in W$, $\mathbf{w}^{\perp} \in W^{\perp}$.*

Note: Theorem 14.5 makes our answer in Ex 14.4 make sense since these vectors are from Ex 14.1 and we already know they are orthogonal in $\mathbb{R}^3$. Thus $\mathbb{R}^3 = V + V^{\perp}$. Since V is 3 dimensional as is $\mathbb{R}^3$, $V^{\perp}$ must be the trivial vector space.

Proof. Let V be a vector space with $W \subseteq V$. We will first show that $W \cap W^{\perp} = \{0\}$: Suppose there is a vector $\mathbf{u} \in W$ and $W^{\perp}$, then $\mathbf{u} \cdot \mathbf{u} = 0$ which means $\mathbf{u} = \mathbf{0}$.
Next we show that any $\mathbf{v} \in V$ can be written as $\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$, where $\mathbf{w} \in W, \mathbf{w}^{\perp} \in W^{\perp}$. Let $\{o_1, o_2, ..., o_n\}$ be an orthonormal basis for W. Let $\mathbf{v} \in V, \mathbf{w} \in W$. By Theorem 14.4, we can write $\mathbf{w} = (\mathbf{v}\cdot\mathbf{o}_1)\mathbf{o}_1 + (\mathbf{v}\cdot\mathbf{o}_2)\mathbf{o}_2 + ... + (\mathbf{v}\cdot\mathbf{o}_n)\mathbf{o}_n$. Let $\mathbf{u} = \mathbf{v} - \mathbf{w}$. We then check that $\mathbf{u} = \mathbf{v} - \mathbf{w} \in W^{\perp}$ We can do this by checking that $\mathbf{v} - \mathbf{w}$ is orthogonal to our orthonormal basis for W. Indeed since $\mathbf{o}_i \cdot \mathbf{o}_j = 0$ for any $i \neq j$ and $\mathbf{o}_i \cdot \mathbf{o}_i = 1$. $\mathbf{o}_1 \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{o}_1 \cdot \mathbf{v} - \mathbf{o}_1 \cdot \mathbf{v} \cdot \mathbf{o}_1)\mathbf{o}_1 - \mathbf{o}_1 \cdot (\mathbf{v} \cdot \mathbf{o}_2)\mathbf{o}_2 - ... - \mathbf{o}_1 \cdot (\mathbf{v} \cdot \mathbf{o}_n)\mathbf{o}_n = \mathbf{o}_1 \cdot \mathbf{v} - \mathbf{o}_1 \cdot (\mathbf{v} \cdot \mathbf{o}_1)\mathbf{o}_1 = \mathbf{v} - \mathbf{v}$ We can similarly check the other dot products to show $\mathbf{u} \in W^{\perp}$. ☐

Example 14.5. [14] *Consider the set from Example 14.1:* $\{\begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}\}$.

a) How do we know this is a basis for $\mathbb{R}^3$?

Solution: *We have already shown that this is an orthogonal set of 3 vectors which means by Theorem 14.3, they are linearly independent. Since $\mathbb{R}^3$ is 3 dimensional, we know this gives us a basis for $\mathbb{R}^3$.*

b) Write $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}$ *as a linear combination of this basis.*

Solution:
One way to do this is to set $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix} = \alpha_1 \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix} + \alpha_3 \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$ and solve this system for $\alpha_1, \alpha_2, \alpha_3$. But because we have an orthogonal basis we can also use Theorem 14.4 which gives an algorithm to solve for $\alpha_1, \alpha_2, \alpha_3$ (which is often less prone to error!).
That is,

$$\alpha_1 = \frac{\mathbf{u} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \tag{13}$$

$$\alpha_2 = \frac{\mathbf{u} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \tag{14}$$

$$\alpha_1 = \frac{\mathbf{u} \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3} \tag{15}$$

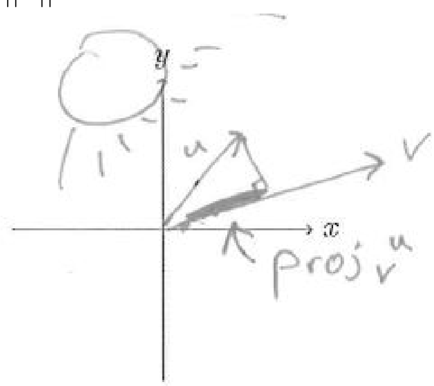
[14] Ex from Holt's Text

Thus $\alpha_1 = \dfrac{\mathbf{u} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} = \dfrac{3(-2) + (-1)(1) + 5(-1)}{(-2)^2 + 1^2 + (-1)^2} = \dfrac{-12}{6} = -2.$
Similarly, $\alpha_2 = \frac{-11}{11} = -1$ and $\alpha_3 = \frac{0}{66} = 0$.
So $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix} = -2 \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix} - 1 \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix} + 0 \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$.

14.4 Projections

In this section, we will consider a particular type of linear transformation, a projection. Before we generalize this concept to subspaces, lets recall how we defined projections in Calculus III or Physics.

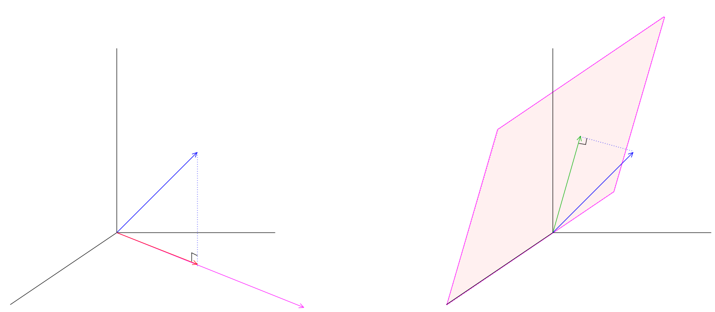
Definition 14.9. *Let* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$*, with* $\mathbf{v} \neq 0$*. The projection of* $\boldsymbol{u}$ *onto* $\boldsymbol{v}$ *is given by* $proj_{\mathbf{v}}\mathbf{u} = \dfrac{\mathbf{u}\cdot\mathbf{v}}{||\mathbf{v}||^2}\mathbf{v} = (\dfrac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}})\mathbf{v}$.



Note: This may remind you of the coefficients from our previous section which allowed us to write any vector as a linear combination of an orthogonal basis.

We would like to generalize this for a subspace. Typically, we use the notation π_V to denote the linear transformation that projects vectors to a space V. Let's first start with an example in $\mathbb{R}^3$ to understand the idea.

Example 14.6. *Suppose we have a vector* $\mathbf{u} \in \mathbb{R}^3$ *and we want to project* $\mathbf{u}$ *onto another vector* $\mathbf{v} \in \{(x, y, 0)|x, y \in \mathbb{R}\}$ *(the x-y plane). In reality, we are projecting onto the line that is parallel to v:* $L = \{\alpha v : \alpha \in \mathbb{R}\}$*. That is, we want to apply the transformation* $\pi_L : \mathbb{R}^3 \to L$*. Then we are looking for the vector* $proj_v(u)$ *shown in the figure on the left below. Notationally, we write* $\pi_L(u) = proj_v(u)$*.*



Now, suppose we want to project u onto the vector space spanned by two vectors, say $V =$ span $\{(1, 1, 1), (1, 0, 0)\}$. Then, we want to apply $\pi_V : \mathbb{R}^3 \to V$ to get proj$_V(u)$ shown in the figure on the right in the figure above.

In terms of the Linear Algebra language that we have been using, to find the projection of a vector u onto a space $V = $ span $\{v_1, v_2, \ldots, v_n\}$, we first find the part (or component) of u that is orthogonal to the basis elements of V (thus, orthogonal to all vectors in V). We call this component $u_\perp$. Then, the projection, $\text{proj}_V(u)$ is the vector that is left over after we subtract the orthogonal part: $\text{proj}_V(u) = u - u_\perp$.

Definition 14.10. *Let V be a nontrivial subspace with an orthogonal basis* $\{v_1, v_2, ...v_n\}$ *then the projection of u onto V is given by* $proj_V(u) = \dfrac{u \cdot v_1}{||v_1||^2}v_1 + \dfrac{u \cdot v_2}{||v_2||^2}v_2 + ... + \dfrac{u \cdot v_n}{||v_1 n||^2}v_n$.

Note that this is just writing **u** as a linear combination of the orthogonal basis of V. See Previous Notes.

Theoretical Note: $\text{proj}_V(u)$ does not depend on the choice of orthogonal basis for V.

14.5 The Gram-Schmidt Process for finding an Orthogonal Basis

Now we will discuss a method for finding an orthogonal basis for an arbitrary Vector Space.

Gram-Schmidt Process: Given a basis $\{b_1, b_2, ...b_n\}$ for a vector space V, we can create an orthogonal basis for V given by $\{v_1, v_2, ...v_n\}$ by the following process:

$v_1 = b_1$

$v_2 = b_2 - \text{proj}_{v_2} b_2$

$v_3 = b_3 - \text{proj}_{v_1} b_3 - \text{proj}_{v_2} b_3$

$v_4 = b_4 - \text{proj}_{v_1} b_4 - \text{proj}_{v_2} b_4 - \text{proj}_{v_3} b_4$

$\vdots$

$v_n = b_n - \text{proj}_{v_1} b_n - \text{proj}_{v_2} b_n - \text{proj}_{v_3} b_n - \cdots - \text{proj}_{v_{n-1}} b_n$

Gram-Schmidt Process -Alternate Representation:

$v_1 = b_1$

$v_2 = b_2 - \dfrac{b_2 \cdot v_1}{v_1 \cdot v_1} v_1$

$v_3 = b_3 - \dfrac{b_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \dfrac{b_3 \cdot v_2}{v_2 \cdot v_2} v_2$

$v_4 = b_4 - \dfrac{b_4 \cdot v_1}{v_1 \cdot v_1} v_1 - \dfrac{b_4 \cdot v_2}{v_2 \cdot v_2} v_2 - \dfrac{b_4 \cdot v_3}{v_3 \cdot v_3} v_3$

$\vdots$

$v_n = b_n - \dfrac{b_n \cdot v_1}{v_1 \cdot v_1} v_1 - \dfrac{b_n \cdot v_2}{v_2 \cdot v_2} v_2 - \dfrac{b_n \cdot v_3}{v_3 \cdot v_3} v_3 - \cdots - \dfrac{b_n \cdot v_{n-1}}{v_{n-1} \cdot v_{n-1}} v_{n-1}$

Computational Comment: The Gram-Schmidt process can suffer from significant round-off error. As we compute the orthogonal vectors, some dot $v_i \cdot v_j$ produces may not be close to zero when $|i-j|$ is large. This will cause a loss of orthogonality. There is a modified version of Gram-Schmidt which requires more operations but is more numerically stable. Givens Rotations or Householder Reflections are two such methods, but they are beyond the scope of this course.

Example 14.7. *Find an orthogonal basis for* $V = span\ \{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 6 \end{bmatrix} \}$
Note Eva has found $\mathbf{v_1} \cdot \mathbf{v_1} = 2, \mathbf{v_2} \cdot \mathbf{v_2} = 24, \mathbf{b_2} \cdot \mathbf{v_1} = -2, \mathbf{b_3} \cdot \mathbf{v_1} = 2, \mathbf{b_3} \cdot \mathbf{v_2} = 24.$

Recall: A set of vectors $\{v_1, v_2, v_3, ..., v_n\}$ is **orthonormal** if the set is orthogonal and $||v_i|| = 1$ for $i = 1, ..n.$ and given an orthogonal set $\{v_1, v_2, v_3, ..., v_n\}$, we can normalize these vectors by dividing each vector by its norm.

Note: When completing the Gram-Schmidt Process by hand, it is easier to normalize each orthogonal vector as you go!

Octave/MatLab: Matlab has a build in function that will orthogonalize the columns of a basis for V. Create a matrix A of its basis elements. Then use $[Q, R] = qr(A)$. Q gives you a matrix with orthonormal columns. You should know how to do this process by hand though.

Example 14.8. *Normalize your orthogonal basis from Ex 14.7:* $\{\begin{bmatrix}-1\\0\\1\end{bmatrix}, \begin{bmatrix}2\\4\\2\end{bmatrix}, \begin{bmatrix}3\\-3\\3\end{bmatrix}\}$

Eva has given us the norms of the vectors:
$||\mathbf{v_1}|| = \sqrt{\mathbf{v_1} \cdot \mathbf{v_1}} = \sqrt{(-1)^2 + 0^2 + 1^2} = \sqrt{2}$,
$||\mathbf{v_2}|| = \sqrt{2^2 + 4^2 + 2^2} = 2\sqrt{6}$,
$||\mathbf{v_3}|| = \sqrt{9^2 + 9^2 + 9^2} = 3\sqrt{3}$.

Definition 14.11. *A is an* ***orthogonal matrix*** *if its columns create an ortho**normal** set.*

Theorem 14.6. *If A is an $n \times n$ orthogonal matrix, then $A^{-1} = A^T$.*

You can show this on your own by showing that $A^{-1} = A^T$ for $A = \begin{bmatrix}\frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ 0 & \frac{2}{\sqrt{6}} & \frac{-1}{\sqrt{3}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\end{bmatrix}$

Proof. (of Theorem 14.6) Let $A = a_{ij}$ be an $n \times n$ orthogonal matrix, Then $A^T = a_{ji}$. Then

$$AA^T = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} a_{21} & a_{21} & a_{31} & \dots & a_{n1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{nn} \end{bmatrix} =$$

$$\begin{bmatrix} a_{11} \cdot a_{11} & a_{12} \cdot a_{21} & a_{13} \cdot a_{31} & \dots & a_{1n} \cdot a_{n1} \\ a_{21} \cdot a_{12} & a_{22} \cdot a_{22} & a_{23} \cdot a_{32} & \dots & a_{2n} \cdot a_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} \cdot a_{1n} & a_{n2} \cdot a_{2n} & a_{n3} \cdot a_{3n} & \dots & a_{nn} \cdot a_{nn} \end{bmatrix}$$

Since the columns are orthogonal and normal, $a_{ij} \cdot a_{ji} = 0$ for $i \neq j$ and $a_{ij} \cdot \cdot a_{ji} = 1$ for $i = j$. Thus $AA^T = I$. Similarly we can show that $A^T A = I$. Therefore $A^{-1} = A^T$. $\square$

14.6 Column Rank

Definition 14.12. *If all of the columns of a matrix A are linearly independent (are a basis for ran(A)), we say A has* ***full column rank.***

Example 14.9. $A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ *has full column rank.*

Definition 14.13. *If all of the rows of a matrix A are linearly independent, we say A has* ***full row rank.***

Note we can determine the column rank of A, by putting it in echelon form and counting the number of leading 1's in each column.

Example 14.10. $B = \begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$ *has full row rank, but NOT full column rank.*

Definition 14.14. *If all of the rows AND columns of a matrix A are linearly independent, we say A has* ***full rank.***

Property: If $n \times m$ matrix A has full column rank, what is the relationship between n and m?
a) $n \geq m$
b) $n \leq m$
c) $n = m$
d) I love math!

Property: If $n \times m$ matrix A has full row rank, what is the relationship between n and m?
a) $n \geq m$
b) $n \leq m$
c) $n = m$
d) I don't want this class to ever end!

Property: If $n \times m$ matrix A has full rank, what is the relationship between n and m?
a) $n \geq m$
b) $n \leq m$
c) $n = m$
d) I love cats!

14.7 QR Factorization

In linear algebra it is often useful to represent matrices in another form. We have already discussed the value of *Diagonalizing* a matrix, when it is possible. We will not talk about all of our different factorizations of matrices, but we will end our time in linear algebra discussing both the QR and SVD factorizations of matrices

QR Factorization: Let A be an $n \times m$ matrix with linearly independent columns (aka full column rank), then A can be factorized as $A = QR$ where Q is an $n \times m$ orthogonal matrix (has ortho**normal** columns) and R is an $m \times m$ upper triangular matrix with positive diagonal entries.

Note: Matrix A does not need to be square, but since A has to have full column rank, we need the number of columns of A to be ________ than the number of rows of A.

QR Process:

1. Find an orthonormal basis for the columns of A. (using Gram-Schmidt)

2. Find R by computing $R = Q^T A$ (works because Q is orthogonal so $Q^{-1} = Q^T$).

Example 14.11. *Calculate the QR factorization for* $A = \begin{bmatrix} -1 & 3 & 4 \\ 0 & 4 & 1 \\ 1 & 1 & 6 \end{bmatrix}$.

14.8 Usefulness of QR Factorization

Diving into Numerical Analysis:
Linear Algebra Has Two Fundamental Problems:

1. Solving $Ax = b$
2. Diagonalizing a matrix A (finding eigenvalues)

Problems: Theoretically perfect algorithms can be very numerically unstable. Errors occur and sometimes are unavoidable and can compound especially for real life applications when we have large matrices.

Solution: Orthogonal matrices are the best for numerical stability so we want to use them if we can.

Methods for Solving $Ax = b$

- Gaussian Elimination/LU Decomposition Advantages:
  - Works for any matrix
  - Finds all solutions when they exist
  - Easy to program (rref!)
  - Fast

  Disadvantages:
  - Can be unstable
  - Can't be used to approximate solutions when systems have no solution (aka least squares)
- QR Decomposition (using Gram-Schmidt): $Ax = b$ can be solved by $QRx = b \to Rx = Q^T b$ which is easier to solve.
  Advantages:
  - Works for any matrix
  - Finds all solutions when they exist
  - Easy to program
  - $Rx = Q^T b$ is well conditioned
  - Finds approximate solutions

  Disadvantages:
  - Slower than LU
  - More complicated

15 Singular Value Decomposition

15.1 Spectral Theorem

We have already discussed the value of diagonalizing a matrix. Recall, A is diagonalizable if there exists matrices P and D such that $A =$ __________. Unfortunately we can't always diagonalize a matrix. When can't we diagonalize a matrix?

Fortunately, we have a useful factorization called *Singular Value Decomposition* which will work for all matrices. Singular values appear in many linear algebra applications especially applications like clustering (eigenvectors help with this too!), statistics, signal processing, and other applications that involve big data. If we had an Applied Linear Algebra II course (oh boy!) we would go into more examples of these applications, but for now we will just discuss a few of them in a lab. Before we discuss these applications, we will first talk about a very nice theorem which is invoked in SVD.

Symmetric Matrices and Orthogonal Matrices have many nice properties. For example, last class we learned that if P is orthogonal then $P^{-1} =$ __________

Theorem 15.1. *If A is symmetric, then the eigenvectors associated with distinct eigenvalues are* ***orthogonal****.*

Pf. Omitted. See supplementary text.

Why is this nice?

Theorem 15.2. *A^TA has nonegative, real, eigenvalues. (Remember A^TA is symmetric!)*

Definition 15.1. *A square matrix A is* ***orthogonally diagonalizable*** *if there exists an orthogonal matrix P and diagonal matrix D such that $A = PDP^{-1}$.*

Note: This is the same idea as diagonalization, just normalize your eigenvectors (since by the previous theorem, they are orthogonal!).

Theorem 15.3. ***The Spectral Theorem:*** *A matrix is orthogonally diagonalizable if and only if A is symmetric.*

Pf. Omitted. The complete proof is rather difficult and is not included here. Often undergraduate linear algebra textbooks do not include the complete proof since half of the proof is easier to prove using complex numbers.

This means if A is symmetric then, A is ________________ ________________!

Example 15.1. *Orthogonally diagonalize A^TA for $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$. Eva found that the eigenspace associated with $\lambda = 1$ is $\{\begin{bmatrix} \frac{-1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}\}$.*

15.2 Singular Value Decomposition

Idea Behind SVD: For any $m \times n$ matrix A, A^TA is symmetric which means it is orthogonally diagonalizable! From Theorem 15.2 on the previous page, we know the eigenvalues for A^TA are nonegative. This allows us to order our eigenvalues in decreasing order:

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ... \geq \lambda_n \geq 0$$

Definition 15.2. *The* ***singular values of A*** *are the square roots of the eigenvalues of A^TA and are denoted $\{\sigma_1, \sigma_2, ...\sigma_n\}$.*

That is $\sigma_i = \sqrt{\lambda_i}$.

Note: The singular values of A are the lengths of the vectors Av_1, Av_2,.. Av_n where the v_i's are eigenvectors.

Theorem 15.4. *Let $\{v_1, v_2, ...v_n\}$ be an orthonormal basis of A^TA arranged in order of eigenvalues: $\lambda_1 \geq \lambda_2 \geq ...\lambda_n$. And suppose A has r nonzero singular values. Then $\{Av_1, Av_2, ...Av_r\}$ is an orthogonal basis for the range/column space of A and the rank of A is r.*

Pf. Omitted. Please see supplementary text for details.

Theorem 15.5. ***Singular Value Decomposition:*** *Suppose A is an $n \times m$ matrix with rank r. Then there exists an $n \times m$ matrix Σ (see below) for which the diagonal entries in D are the first r singular values of A $\sigma_1 \geq \sigma_2 \geq \sigma_r > 0$ and an $n \times n$ orthogonal matrix U and $m \times m$ orthogonal matrix V such that $A = U\Sigma V^T$*

Σ is a matrix that is the *same* size as A. We have two cases for Σ:

First Let D be the diagonal matrix with ordered singular values:

$$D = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \sigma_r \end{bmatrix}$$

Case 1: $n \geq m$: (A is a tall matrix)

$$\Sigma = \begin{bmatrix} D \\ 0 \end{bmatrix}$$

Case 2: $n \leq m$: (A is a wide matrix)

$$\Sigma = \begin{bmatrix} D & 0 \end{bmatrix}$$

Note: For case 2: we can sometimes make life easier for ourselves by considering $B = A^T$. Then B is a matrix with Case 1: and if $B = U\Sigma V^T$ then $A = B^T = (U\Sigma V^T)^T =$

Octave Code for SVD: [U,S,V] = svd(B)

Theorem 15.6. *Every $n \times m$ matrix A has a singular value decomposition.*

Heuristic Proof: Follow the construction from Thm 15.5 by completing the steps below.

Steps for Singular Value Decomposition:

1. Orthogonally Diagonalize A^TA to find V (V will be our P when $A^TA = PDP^T$).

2. Find Σ using the singular values of A.

3. Find U in 2 Steps

   (a) Find the first entries by $u_i = \dfrac{1}{\sigma_i} Av_i$
   Why? Our goals is to find U so that $A = U\Sigma V^T$ aka $AV = U\Sigma$. The ith column of AV is Av_i and the ith column of $U\Sigma$ is $\sigma_i u_i$. So we need $Av_i = \sigma_i u_i$. Thus
   $u_i = \frac{1}{\sigma_i} Av_i$

   (b) Fill out U by extending to an orthonormal basis of R^n. Basically you can do this by adding the normalized vectors from null(A^T). Here is why: The first vectors you find using $u_i = \dfrac{1}{\sigma_i} Av_i$, actually give you a basis for the range space of A (aka the column space). The column space is a subspace of $\mathbb{R}^3$ and remember our theorem which states $V = W \bigoplus W^\perp$ for any subspace W. So in our case, $\mathbb{R}^3 = (\text{ran}(A)) \bigoplus (\text{ran}(A))^\perp$. It turns out $(\text{ran}(A))^\perp = \text{null}(A^T)$ (proof omitted).

Example 15.2. *Find the singular value decomposition (SVD) for* $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Recall, in Ex 15.1 we found $A^TA = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \cdot \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$.

Example 15.3. *Find the SVD for* $A = \begin{bmatrix} -1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix}$

If you can use technology, load up Octave type [U,S,V]=svd(A). To do this by hand, do the following:
Step 1: Orthogonally Diagonalize $A^T A$ to find V (V will be our P when $A^T A = PDP^T$)
Note $A^T A = \begin{bmatrix} -1 & 2 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 & 2 \\ 3 & 5 & 2 \\ 2 & 1 & 1 \end{bmatrix}$. We could *TOTALLY Do this*, but if we diagonalize $B = A^T$, we will get a 2 by 2 matrix which is easier. Let's do this.
Let $B = A^T$.
Then $B^T B = AA^T = \begin{bmatrix} -1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 2 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 9 \end{bmatrix} = M$
We first find the eigenvalues of M. Because it is a diagonal matrix, we know the eigenvalues are $\lambda_1 = 9, \lambda_2 = 2$ (Note that I ordered them from largest to smallest.)

Next we find the eigenspaces:

$\mathcal{E}_9 = \text{null}(M - 9I) = \text{null}(\begin{bmatrix} -7 & 0 \\ 0 & 0 \end{bmatrix}) \Rightarrow \left(\begin{array}{cc|c} -7 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)$. So $x = 0, y = y$. So $\mathcal{E}_9 = \text{span}\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$
We need to normalize this vector, but it already has length 1 so we don't have to do anything else.

$\mathcal{E}_2 = \text{null}(M - 2I) = \text{null}(\begin{bmatrix} 0 & 0 \\ 0 & 7 \end{bmatrix}) \Rightarrow \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 7 & 0 \end{array}\right)$. So $x = x, y = 0$. So $\mathcal{E}_2 = \text{span}\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\}$
We also need to normalize this vector, but it already has length 1 (neat!).

Thus, $B^T B = PDP^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 9 & 0 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (Note in this example $P^T = P$).
Thus the V for B is $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. So $V^T = P^T = P$.

Step 2: Find Σ using the singular values of A.

Since B is a 3×2 matrix, Σ for B is a 3×2: $\Sigma = \begin{bmatrix} 3 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}$. Remember we put the singular values for the B along the diagonals (the square roots of the eigenvalues.)

We are almost done! Go to the next page!

Step 3: Find U in 2 Steps:
1) Find the first entries by $u_i = \dfrac{1}{\sigma_i}Av_i$.
2) Then Fill out U by extending to an orthonormal basis of R^n by finding the nullspace of A^T.

Since we are using B, we use $u_i = \dfrac{1}{\sigma_i}Bv_i$:

$u_1 = \frac{1}{3}\begin{bmatrix}-1 & 2\\ 1 & 2\\ 0 & 1\end{bmatrix} \cdot \begin{bmatrix}0\\1\end{bmatrix} = \frac{1}{3}\begin{bmatrix}2\\2\\1\end{bmatrix} = \begin{bmatrix}\frac{2}{3}\\ \frac{2}{3}\\ \frac{1}{3}\end{bmatrix}$.

$u_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}-1 & 2\\ 1 & 2\\ 0 & 1\end{bmatrix} \cdot \begin{bmatrix}1\\0\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\\0\end{bmatrix} = \begin{bmatrix}\frac{-1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\\ 0\end{bmatrix}$.

So we have found the first two entries of U for B. We need to fill out U by finding vectors in null(B^T) = null($\begin{bmatrix}-1 & 1 & 0\\ 2 & 2 & 1\end{bmatrix}$)$\Rightarrow \left(\begin{array}{ccc|c}-1 & 1 & 0 & 0\\ 2 & 2 & 1 & 0\end{array}\right) \Rightarrow \left(\begin{array}{ccc|c}1 & 0 & \frac{1}{4} & 0\\ 0 & 1 & \frac{1}{4} & 0\end{array}\right)$ So $x = \frac{-z}{4}, y = \frac{-z}{4}, z = z$. So the nullspace is $\begin{bmatrix}-1\\-1\\4\end{bmatrix}$ We need to normalize this vector:

$|| \begin{bmatrix}-1\\-1\\4\end{bmatrix} || = \sqrt{(-1)^{+}(-1)^2 + 4^2} = \sqrt{18}$, so our 3rd vector is $\begin{bmatrix}\frac{-1}{\sqrt{18}}\\ \frac{-1}{\sqrt{18}}\\ \frac{4}{\sqrt{18}}\end{bmatrix}$.

Hurray we have completed U! $U = \begin{bmatrix}\frac{2}{3} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{18}}\\ \frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{18}}\\ \frac{1}{3} & 0 & \frac{4}{\sqrt{18}}\end{bmatrix}$

Thus $B = U\Sigma V^T = \begin{bmatrix}\frac{2}{3} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{18}}\\ \frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{18}}\\ \frac{1}{3} & 0 & \frac{4}{\sqrt{18}}\end{bmatrix} \begin{bmatrix}3 & 0\\ 0 & \sqrt{2}\\ 0 & 0\end{bmatrix} \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$

Last Step because we were using $B = A^T$:
If $B = U\Sigma V^T$, and $B^T = A$, then $A = B^T = (U\Sigma V^T)^T = V\Sigma^T U^T$.

So $A = V\Sigma^T U^T = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \begin{bmatrix}3 & 0 & 0\\ 0 & \sqrt{2} & 0\end{bmatrix} \begin{bmatrix}\frac{2}{3} & \frac{2}{3} & \frac{1}{3}\\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ \frac{-1}{\sqrt{18}} & \frac{-1}{\sqrt{18}} & \frac{4}{\sqrt{18}}\end{bmatrix}$

16 Practice Problems and Review for Exams

The following pages are practice problems and main concepts you should master on each exam.

These problems are meant to help you practice for the exam and are often harder than what you will see on the exam. You should also look over all ICE sheets, Homework problems, and Lecture Notes. Make sure you can do all of the problems in these sets.

Disclaimer: The following lists are topics that you should be familiar with, and these are problems that you should be able to solve. This list may not be complete. You are responsible for everything that we have covered thus far in this course.

We will post the solutions to these practice exams on our blackboard site. If you find any mistakes with my solutions, please let us know right away and feel free to email at any hour.

Good Luck,
Dr. H and Professor Smith

MA 307 –Review for Exam 1

Remember to look over all Homework problems, Daily Assignments, and Lecture Notes (all are posted). Make sure you can do all of those problems. You can also practice with the problems listed as part of this working review. Disclaimer: The following is a list of topics that you should be familiar with, and a list of problems that you should be able to solve. This list may not be complete. You are responsible for everything that we have covered thus far in this course.

Main Concepts for Exam 1

Mastery Concept 1: Gaussian Elimination
Mastery Concept 2: Understanding solutions to systems of equations
Mastery Concept 3: Vector Spaces and Subspaces
Mastery Concept 4: Span
Mastery Concept 5: Basis/Linear Independence
Mastery Concept 6: Least Squares

1. **Concept 1: Gaussian Elimination**
   (a) Know how to use Gaussian Elimination to solve systems of linear equations.
      i. Know how to write solutions sets of systems
      ii. Be able to create a **Row Reduced Echelon Form** Matrix using Row Operations
      iii. Be able to do Gaussian Elimination by hand and also by using technology like Octave/Matlab or your calculator

2. **Concept 2: Solutions to systems of equations**
   (a) Know how to write solution sets for row reduced echelon form.
   (b) Know how to write solution sets as spans or parametric forms using free variables.
   (c) Know how to recognize a consistent system or inconsistent system.

3. **Concept 3: Vector Spaces**
   (a) Know the 10 properties that set need to satisfy in order to be defined as a vector space -remember these spaces only have 2 operations -addition and scalar multiplication!
   (b) Know how to recognize when a set is not a vector space (usually you should check the closure and identity properties)
   (c) Be able to list examples of Vectors spaces and understand why they are vector spaces. Some examples include, but are not limited to:
   $\mathbb{R}^n$, $GF(2)^n$, $\mathcal{M}_{n\times m}$, $\mathcal{P}_n$, and The space of greyscale images, a spanning set, the solution space for a homeogeneous system, the range space of a transformation, the nullspace of a transformation,(the last two may show up on Exam 2)...

4. **Concept 3: Subspaces**

   (a) Know how to test whether or not a subset of a vector space is a subspace (that is, the set acts like a self-contained vector space.)

   (b) How to show a **subset**, W, is a subspace:
   Method 1:

      i. Check if $\mathbf{0} \in W$ (that is make sure the additive identity is in W.)

      ii. Check that W is closed under linear combinations. That is let $\mathbf{u}, \mathbf{v}$ be arbitrary vectors in W and make sure $\alpha\mathbf{u} + \beta\mathbf{v} \in W$ for arbitrary scalars α, β.

   Method 2:

      i. Show that the subspace can be written as a span of vectors.

5. **Concept 4: Span**

   (a) Know the definition of *a span of vectors*:
   span$(v_1, v_2, ...v_n)$=$\{a_1v_1 + a_2v_2 + ... + a_nv_n | a_i \in \mathbb{R}\}$
   (the set of all linear combinations of $v_1, ...v_n$.)

   (b) Be able to write a vector space as a span of vectors. Another way to say this is to find a spanning set for a vector space.

      i. Process: get a set in the form:$\{a_1v_1+a_2v_2+...+a_nv_n|a_i \in \mathbb{R}\}$ so that there are no restrictions on the scalar coefficients. Often this can be done by creating an augmented matrix of the requirements of the set and using free variables to help find your coefficients.

      ii. See HW, Spanning Notes

   (c) Know the definition of *a spanning set*:
   $\{v_1, v_2, ...v_n\}$ is a spanning set for V if span$(v_1, v_2, ...v_n) = V$.

   (d) Be able to determine whether or not a particular vector is in the span of a set of vectors.

      i. Process: Find coefficients a_i such that $\vec{u} = a_1v_1 + a_2v_2 + ... + a_nv_n$, then $\vec{u} \in$span$(v_1, v_2, ...v_n)$

      ii. See HW Spanning Notes

   (e) Be able to show whether or not a set of vectors spans a whole Vector Space (that is, be able to show a set is a spanning set for a vector space.)

      i. Process: Pick an arbitrary vector from the Vector Space and try to find scalars a_i such that the vector can be written as a linear combination of the set of vectors.

6. **Concept 5: Linear Independence**

   (a) Know the definition of linear independence and dependence

(b) Know how to determine whether or not a set of vectors is linear independent or dependent. Here is the process to check whether or not $\{v_1, v_2, ...v_n\}$:
Step 1: Set an arbitrary linear combination of v_i's equal to 0:
$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n = 0$
Step 2: See if you can find values for the α_i's that are not all 0. If you can, the set is dependent. If $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$ is the *only* possible values for the α's, then the set is linear independent.

7. **Concept 5: Basis**

(a) Know the standard basis for $\mathbb{R}^n$, $\mathcal{P}_n$, and $\mathcal{M}_{n\times m}$.

(b) Know the requires/definition of a basis: A set $\{v_1, v_2, ...v_n\}$ is a basis for V if both of the following are true:

i. Span($\{v_1, v_2, ...v_n\}$)=V

ii. $\{v_1, v_2, ...v_n\}$ is linear independent

(c) Know how to find a basis for a set

8. **Concept 6: Least Squares**

- Be able to find an approximate solution to an inconsistent solution by solving the normal equation $A^T \cdot Ax = A^T b$.

Practice Problems:

1. Give the solution set of each system.

   (a)
   $$\begin{aligned} 3x + 2y + z &= 1 \\ x - y + z &= 2 \\ 5x + 5y + z &= 0 \end{aligned}$$

   (b)
   $$\begin{aligned} x + y - 2z &= 0 \\ x - y &= -3 \\ 3x - y - 2z &= -6 \\ 2y - 2z &= 3 \end{aligned}$$

   (c)
   $$\begin{aligned} 2x - y - z + w &= 4 \\ x + y + z &= -1 \end{aligned}$$

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

2. Verify that each is a vector space by checking the conditions.

   (a) The collection of 2×2 matrices with 0's in the upper right and lower left entries.

   $$\{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \mid a, b \in \mathbb{R}\}$$

   (b) The collection $\{a_0 + a_1x + a_3x^3 \mid a_0, a_1, a_3 \in \mathbb{R}\}$ of cubic polynomials with no quadratic term.

3. Determine if each set is linearly independent (in the set's natural vector space).

   (a) $\{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\} \subseteq \mathbb{R}^3$

   (b) $\{(1\ \ 3\ \ 1), (-1\ \ 4\ \ 3), (-1\ \ 11\ \ 7)\} \subseteq \mathcal{M}_{1\times 3}$

   (c) $\{\begin{pmatrix} 5 & 4 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 4 \end{pmatrix}\} \subseteq \mathcal{M}_{2\times 2}$

4. Is the vector in the span of the set?

   $$\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \qquad \{\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\}$$

5. Find a basis for each space. Verify that it is a basis.

   (a) The subspace $M = \{a + bx + cx^2 + dx^3 \mid a - 2b + c - d = 0\}$ of $\mathcal{P}_3$.

(b) This subspace of $\mathcal{M}_{2\times 2}$.

$$W = \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a - c = 0 \}$$

6. Give two different bases for $\mathbb{R}^3$. Verify that each is a basis.

7. Find a basis for, and the dimension of, each space.

   (a)

$$\{ \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \in \mathbb{R}^4 \mid x - w + z = 0 \}$$

   (b) the set of 5×5 matrices whose only nonzero entries are on the diagonal (e.g., in entry $1, 1$ and $2, 2$, etc.)

   (c) $\{ a_0 + a_1 x + a_2 x^2 + a_3 x^3 \mid a_0 + a_1 = 0 \text{ and } a_2 - 2a_3 = 0 \} \subseteq \mathcal{P}_3$

8. Give a basis for the span of each set, in the natural vector space.

   (a) $\{ \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 12 \\ 6 \end{pmatrix} \} \subseteq \mathbb{R}^3$

   (b) $\{ x + x^2, 2 - 2x, 7, 4 + 3x + 2x^2 \} \subseteq \mathcal{P}_2$

9. Find a least squares solution of $\begin{bmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{bmatrix} x = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$.

10. Find a least squares solution of $\begin{bmatrix} 1 & -2 \\ -1 & 2 \\ 0 & 3 \\ 2 & 5 \end{bmatrix} x = \begin{bmatrix} 3 \\ 1 \\ -4 \\ 2 \end{bmatrix}$.

MA 307 –Review for Exam 2

Remember to look over all Homework problems, Daily Assignments, and Lecture Notes (all are posted). Make sure you can do all of those problems. You can also practice with the problems listed as part of this working review. Disclaimer: The following is a list of topics that you should be familiar with, and a list of problems that you should be able to solve. This list may not be complete. You are responsible for everything that we have covered thus far in this course.

Main Concepts for Exam 2

Mastery Concept 1: Gaussian Elimination
Mastery Concept 2: Understanding solutions to systems of equations
Mastery Concept 3: Vector Spaces and Subspaces
Mastery Concept 4: Span
Mastery Concept 5: Basis/Linear Independence
Mastery Concept 6: Least Squares
Mastery Concept 7: Linear Transformations/Functions
Mastery Concept 8: Matrix representation for a linear transformation
Mastery Concept 9: Injective and Surjective Linear Transformations
Mastery Concept 10: Rank Nullity Theorem
Mastery Concept 11: Matrix Spaces: Nullspace, Column Space, Nullity, Rank
Mastery Concept 12: Determinants

New Concepts

Concept 7: Linear Transformation/Functions

1. Know the definition of a Linear Transformation:
   $T : U \to V$ is a linear transformation if $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$ for any arbitrary linear combination $\alpha u + \beta v \in U$.

2. Know how to test whether or not a Function/Mapping is a Linear Transformation:
   Process:

   (a) Method 1: Show that $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$ for any arbitrary linear combination.

   (b) Method 2: Show that both of the following hold:

      i. $T(u + v) = T(u) + T(v)$

      ii. $T(\alpha u) = \alpha T(u)$

   (c) Note: If $T(0) \neq 0$, T is NOT a linear transformation. But if $T(0) = 0$, you must use Method 1 or 2 to show it is a linear transformation.

3. Know how to find images of linear transformations.(Given T and x, what is $T(x)$?)

4. Know how to find a vector whose image is produced by a linear transformation. (Given T and b, what x gives $T(x) = b$?)

Concept 8: Matrix Representation of a Linear Transformation

- Know that every linear transformation can be represented in a matrix form.
- Process for finding the matrix form for a linear transformation (aka the change of basis matrix) for $T : V \to W$

  Step 1: Find the basis elements for the domain (V), codomain (W), and ran(T) (the range space of T)
  Step 2: Map the basis elements of V to the basis elements of ran(T)
  Step 3: Find the coordinate vectors or our ran(T) elements with respect to the basis for the codomain (W). That is write our ran(T) basis elements as a linear combination of the basis elements of the codomain, then list the coefficients into a column vector
  Step 4: Construct your matrix representation by using the coordinate vectors as the column vectors for M.

- You should know how to check whether your matrix representation does the same thing as the transformation. Do this by taking the a domain element and applying T to the element. Write the image of this element as a coordinate vector of the codomain. check to see if this coordinate vector gives you the same thing as writing your domain element as a coordinate vector and applying your matrix representation to it. Then check to see if your vector output matches the coordinate vector you got earlier.

Concept 9: Injective and Surjective Linear Transformations

1. Know the definition of an injective linear transformation: if $T(u) = T(v)$ then $u = v$.
2. Know the other terms for injective maps: one to one, injection
3. Know how to determine if a function is injective. There are 2 Methods:
   Method 1: Using the definition of injective functions:
   (a) Suppose two arbitrary values in the range are equal $T(u) = T(v)$.
   (b) Use the properties of the vector space and the definition of the function to show that the input values are equal ($u = v$.)

   Method 2: Use the Nullspace
   (a) Find the Nullspace(T), if Null(T)=$\{0\}$, T is injective.
4. Know the definition of a surjective linear transformation: For every value in the codomain, there is a value in the domain that maps to it. That is if $T : V \to W$, then T is a surjection if for every $w \in W$, there exists a $v \in V$ such that $T(v) = w$.
5. Know the other terms for surjective maps: onto, surjection

6. Know how to determine if a function is surjective. There are 2 Methods Method 1: Use the definition of a surjective map:

   (a) Pick an arbitrary value in the codomain.

   (b) See if there is an input value that will map to it.

   (c) Note often if the map has some restrictions to its output values, the map is not surjective -exploit these restrictions to show there is no input that can map to this output.

   Method 2: Use the range space

   (a) if the rank(T)=dim(codomain), T is surjective

7. Know the definition of a bijection or isomorphism

8. Know how to create an isomorphism between vector spaces (key: Map basis elements to basis elements.)

9. Know the definition of isomorphic vector spaces.

10. Nullspace and Range Space and Rank Nullity Theorem

    (a) Know the definition of the nullspace of a linear transformation, $T : V \to W$: $\text{null}(T) = \{v \in V |\ T(v) = 0\}$.

    (b) Know the definition of the nullity of a linear transformation, T: the dimension of $\text{null}(T)$.

    (c) Know that $\text{null}(T)$ is a subspace of V.

    (d) Know how to use the nullspace to determine whether or not a transformation is injective

    (e) Know how to find the nullspace of a transformation.

11. Range Space

    (a) Know the definition of the nullspace of a linear transformation, $T : V \to W$, $\text{ran}(T) = \{T(v) |\ v \in V\}$.

    (b) Know the definition of the rank of a linear transformation, T: the dimension of $\text{ran}(T)$.

    (c) Know that ran(T) is a subspace of W.

    (d) Know how to use the rank and range space to determine whether or not a transformation is surjective

    (e) Know how to find the range space of a transformation.

Concept 10: The Rank Nullity Theorem

1. Know what the domain and codomain of a linear transformation
2. Know the Rank Nullity Theorem: Given a linear transformation $T : V \to W$, dim(V)= Rank+Nullity
3. Know how to use this theorem to determine whether or not transformations can be injective or surjective
4. Be able to construct injective and surjective maps if they exist.

Concept 11: Matrix Spaces: Nullspace and Range/Column Space

- Know the definition of the nullspace for an $m \times n$ matrix M, nulls$(M) = \{v \in \mathbb{R}^n | Mv = 0\}$
- If M is an $m \times n$ matrix, the Nullspace is a subspace of $\mathbb{R}^n$
- Know the definition of the range space for an $m \times n$ matrix M, ran(M)=$\{Mx \in \mathbb{R}^m | \; x \in \mathbb{R}^n\}$
- If M is an $m \times n$ matrix, the range space is a subspace of $\mathbb{R}^m$
- Know that the *nullity* of a matrix to be the dimension of the nullspace and the *rank* of a matrix to be the dimension of the range space. Furthermore the same ideas for injective and surjective linear transformations work for Matrix Spaces.
- Know how to find the nullspace and range space for a matrix.
- Know that a shortcut for finding the range space or column space of a matrix: Find the echelon form of the matrix and pick the original columns that correspond to the columns with leading ones.
- Know that the number of columns with leading 1's of a matrix's echelon form is the same as the rank for that matrix.
- Know that the number of columns without leading 1's of a matrix's echelon form is the same as the nullity for that matrix.

Concept 12: Determinants

- Know how to find the determinant of a 2 by 2 matrix: $\left| \begin{array}{cc} a & b \\ c & d \end{array} \right| = ad - bc$
- Know how to efficiently do cofactor expansion to determine the determinant of a matrix: $|A| = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} |M_{i,j}|$,

For example: Let $A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{pmatrix}$, then

$$|A| = a_{1,1}\begin{vmatrix} a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,2} & a_{4,3} & a_{4,4} \end{vmatrix} - a_{1,2}\begin{vmatrix} a_{2,1} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,3} & a_{4,4} \end{vmatrix} + a_{1,3}\begin{vmatrix} a_{2,1} & a_{2,2} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,4} \end{vmatrix} - a_{1,4}\begin{vmatrix} a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \\ a_{4,1} & a_{4,2} & a_{4,3} \end{vmatrix}.$$

- Know that we can expand across any row or column when completing cofactor expansion.

Other Related Concepts;

1. **Inverse Matrices**

   - Know when a Matrix will have an Inverse (AKA the Invertible Matrix Theorem:) Suppose A is a square $n \times n$. The following statements are either all true or all false:
     (a) A is invertible
     (b) A has n pivot positions
     (c) The Nullspace of A is trivial (The equation $Ax = 0$ has only the trivial solution.)
     (d) The columns of A form a linearly independent set.
     (e) The columns of A span $\mathbb{R}^n$
     (f) The linear transformation represented by A is injective.
     (g) The linear transformation represented by A is surjective.
     (h) $AX = b$ has a unique solution
     (i) A^T is an invertible matrix.
     (j) There is an $n \times n$ matrix C such that $CA = I$
     (k) There is an $n \times n$ matrix D such that $AD = I$
     (l) $\det A \neq 0$

2. **Coordinate Vectors**

   - If we think of our basis elements as the building blocks for a vector space, our coordinate vectors represent the manual (or instructions) on how to put our basis elements together to get a particular vector.

   - Given a basis for V, we are able to represent any vector $v \in V$ as a coordinate vector in $\mathbb{R}^n$, where $n = \dim V$. Suppose $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ is a basis for V, then we find the coordinate vector $[v]_{\mathcal{B}}$ by finding the scalars, α_i, that make the linear combination $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$ and we get $[v]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n$.

Practice Problems:

1. For each map below, describe the range space and find the rank of the map.

   (a) $h\colon \mathcal{P}_3 \to \mathbb{R}^2$ given by
$$ax^2 + bx + c \mapsto \begin{pmatrix} a+b \\ a+c \end{pmatrix}$$

   (b) $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by
$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ x-y \\ 3y \end{pmatrix}$$

2. Verify that this map is an bijection (isomorphism): $h\colon \mathbb{R}^4 \to \mathcal{M}_{2\times 2}$ given by
$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \mapsto \begin{pmatrix} c & a+d \\ b & d \end{pmatrix}$$

3. Find the determinant of the following matrices.

   (a) $\begin{vmatrix} 1 & 4 \\ 2 & 8 \end{vmatrix}$

   (b) $\begin{vmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 6 & 4 & 1 \end{vmatrix}$

   (c) $\det \begin{bmatrix} 1 & 0 & -1 \\ 3 & 1 & 1 \\ -1 & 0 & 3 \end{bmatrix}$

4. Find the coordinate vectors for the standard basis for $\mathbb{R}^3$ in terms of the basis: $\mathcal{B} = \{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \}$

5. Use the echelon version of $A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 3 & 1 & -1 \\ 7 & 11 & 4 & -3 \end{bmatrix}$ to determine the column rank, row rank, rank, and nullity for A.

6. Assume that each matrix represents a map $h : \mathbb{R}^m \to \mathbb{R}^n$ with respect to the standard bases. In each case, (i) state m and n (ii) find rangespace (h) and rank of h (iii) find nullspace of nullity, and (iv) state whether the map is onto and whether it is one-to-one.

   (a) $\begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}$

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

(c) $\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$

7. If T is a linear transformation from $V \to W$ with rank$= 4$ and nullity$= 2$, what can we say about the dimension of V? of W?

8. If T is a linear transformation from $\mathcal{P}_4 \to \mathcal{M}_{2\times 3}$ can T be injective? If so, construct an example of such a possible linear transformation.

9. If T is a linear transformation from $\mathcal{P}_4 \to \mathcal{M}_{2\times 3}$ can T be surjective? If so, construct an example of such a possible linear transformation.

10. If T is a linear transformation from $\mathcal{P}_4 \to \mathcal{M}_{2\times 3}$ can T be bijective? If so, construct an example of such a possible linear transformation.

11. If T is an isomorphic linear transformation from $V \to W$. What can we say about the dimensions of V and W?

12. Find the nullspace, nullity, range/column space and rank for the following matrices:

    (a) $A = \begin{bmatrix} 1 & 0 & -4 & -3 \\ -2 & 1 & 13 & 5 \\ 0 & 1 & 5 & -1 \end{bmatrix}$ rref $\to \begin{bmatrix} 1 & 0 & -4 & -3 \\ 0 & 1 & 5 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

    (b) $B = \begin{bmatrix} 1 & -2 & 5 \\ 2 & 4 & 1 \\ -4 & 0 & 2 \\ 1 & -2 & 0 \\ 3 & 1 & 1 \end{bmatrix}$ rref $\to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

13. Verify that each map is a linear transformation.

    (a) $h\colon \mathcal{P}_3 \to \mathbb{R}^2$ given by

    $$ax^2 + bx + c \mapsto \begin{pmatrix} a+b \\ a+c \end{pmatrix}$$

    (b) $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by

    $$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ x-y \\ 3y \end{pmatrix}$$

14. Consider the linear transformation $T : \mathbb{R}^2 \to P_1$ where $T(\begin{bmatrix} a \\ b \end{bmatrix}) = bx - a$ Use these bases for the spaces: basis for domain: $\mathcal{B}_V$ = standard basis and basis for codomain:

$\mathcal{B}_{\mathcal{W}} = \{x, 1\}$. Represent this linear transformation (find the change of basis matrix:) Check that this matrix does the same thing as our transformation for $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

15. Consider the linear transformation $T : P_2 \to P_1$ where $T(ax^2 + bx + c) = 2ax + b$

    Use these bases for the spaces. Use these bases for the spaces: basis for domain: $\mathcal{B}_V =$ standard basis for P_2 basis for codomain: $\mathcal{B}_W =$ standard basis for P_1 Represent this linear transformation (find the change of basis matrix:) Check that this matrix does the same thing as our transformation for $x^2 - x + 1$.

MA 307 –Review for Exam 3

Remember to look over all Homework problems, Daily Assignments, and Lecture Notes (all are posted). Make sure you can do all of those problems. You can also practice with the problems listed as part of this working review. Disclaimer: The following is a list of topics that you should be familiar with, and a list of problems that you should be able to solve. This list may not be complete. You are responsible for everything that we have covered thus far in this course.

Main Concepts for Exam 3

Mastery Concept 1: Gaussian Elimination
Mastery Concept 2: Understanding solutions to systems of equations
Mastery Concept 3: Vector Spaces and Subspaces
Mastery Concept 4: Span
Mastery Concept 5: Basis/Linear Independence
Mastery Concept 6: Least Squares
Mastery Concept 7: Linear Transformations/Functions
Mastery Concept 8: Matrix representation for a linear transformation
Mastery Concept 9: Injective and Surjective Linear Transformations
Mastery Concept 10: Rank Nullity Theorem
Mastery Concept 11: Matrix Spaces: Nullspace, Column Space, Nullity, Rank
Mastery Concept 12: Determinants
Mastery Concept 13: Eigenvalues and Eigenvectors
Mastery Concept 14: Diagonalization
Mastery Concept 15: Markov Chains
Mastery Concept 16: Gram-Schmidt and Orthogonality

New Concepts
Mastery Concept 13: Eigenvalues and Eigenvectors

- Know the definition of eigenvectors, eigenvalues, and eigenspaces:
- Know the how to find the eigenvalues of a matrix A: find the scalars λ so that $\det(A - \lambda I) = 0$.
- Know that the characteristic equation is $\det(A - \lambda I) = 0$
- Know the characteristic polynomial is $\det(A - \lambda I)$.
- Know how to find the eigenvalues of a diagonal or triangular matrix (they are just the diagonal entries!)
- Know how to find an eigenvector for an eigenvalue, λ: Solve $(A - \lambda I)v = 0$ aka: find the nullspace of $(A - \lambda I)$
- Know how to find an eigenspace associated with eigenvalue λ: Solve $(A - \lambda I)v = 0$ aka: find the nullspace of $(A - \lambda I)$. Write the space as a span.

Mastery Concept 14: Diagonalization

- Know the definition of an eigenbasis for $\mathbb{R}^n$: a set of n linearly independent eigenvectors: $v_1, v_2, \ldots, v_n$
- Know how to determine whether or not a matrix is diagonalizable (and orthogonally diagonalizable)
- Know that two matrices, A and B, are similar if there exists an invertible matrix P such that $A = PBP^{-1}$
- Know the definition of a diagonalizable matrix: A matrix that is similar to a diagonal matrix. In other words there exists diagonal D and invertible P such that $A = PDP^{-1}$
- Know the **Process to Diagonalize a Matrix, If Possible**
  1. Find the eigenvalues of the matrix.
  2. Find linearly independent eigenvectors of the matrix.
  3. Determine if you have enough eigenvectors from Step 2, that you can span $\mathbb{R}^n$. Note you can only proceed if there are the same number of eigenvectors as the dimension of A (n) if there are not enough eigenvectors, the matrix is NOT diagonalizable).
  4. Construct P from the eigenvectors from Step 2.
  5. Construct D from the corresponding eigenvalues (order matters -make sure they match the eigenvectors!).
  6. Optional: Check to make sure P and D work by checking if $A = PDP^{-1}$, but it is easier to just check if $AP = PD$.
- Know the definition of a symmetric matrix: $A^T = A$
- Know that symmetric matrices can be orthogonally diagonalizable: This means $A = PDP^T$ where P is orthogonal.
- Know the process for orthogonally diagonalizing A^TA or any symmetric matrix (Same as diagonalization only normalize your eigenvectors.)

Mastery Concept 15: Markov Chains

- Know the definition of a Markov Chain/Process: A process in which the probability of the system being in a particular state at a given time period **depends only on its state at previous time period**.
- Know how to construct a transition for a Markov process and be able to use it to find future states.
- Know the properties of a transition matrix: All columns add up to 1 and every entry is between 0 and 1.

- Know what a steady state vector is for a Markov process: A vector such that $Tv = v$ where T is the transition matrix.
- Know how to find a steady state vector: Solve $(T - I_n)v = 0$ (find a normalized eigenvector for the eigenvalue 1.)

Mastery Concept 16: Gram Schmidt and Orthogonality

- Know the definition of Orthogonal matrices, a matrix in which all its columns are orthonormal.
- Know properties of orthogonal matrices like: $A^{-1} = A^T$
- Know the definition of the projection of u onto v: $\text{proj}_v u = \frac{u \cdot v}{||v||^2} v$
- Know the definition of an orthogonal set of matrices and how to check if two vectors are orthogonal (check $u \cdot v = 0$)
- Know how to determine the length of a vector: $||v|| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + ... + v_n^2}$
- Know the definition of an orthonormal set of vectors: a set which is orthogonal and all vectors have length 1.
- Know that a set of nonzero orthogonal vectors is a linearly independent.
- Know the definition of an orthogonal basis: A basis made up of orthogonal vectors
- Be able to use the Gram Schmidt process to create an orthogonal basis for a span of vectors.
- **Gram-Schmidt Process:** Given a basis $\{b_1, b_2, ...b_n\}$ for a vector space V, we can create an orthogonal basis for V given by $\{v_1, v_2, ...v_n\}$ by the following process:
  $v_1 = b_1$

  $$v_2 = b_2 - \frac{b_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

  $$v_3 = b_3 - \frac{b_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{b_3 \cdot v_2}{v_2 \cdot v_2} v_2$$

  $$v_4 = b_4 - \frac{b_4 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{b_4 \cdot v_2}{v_2 \cdot v_2} v_2 - \frac{b_4 \cdot v_3}{v_3 \cdot v_3} v_3$$

  $$\vdots$$

  $$v_n = b_n - \frac{b_n \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{b_n \cdot v_2}{v_2 \cdot v_2} v_2 - \frac{b_n \cdot v_3}{v_3 \cdot v_3} v_3 - \cdots - \frac{b_n \cdot v_{n-1}}{v_{n-1} \cdot v_{n-1}} v_{n-1}$$

- **Know how to write a matrix into its QR Decomposition:**
  1. Find an orthonormal basis for the columns of A. (using Gram-Schmidt)

2. Find R by computing $R = Q^T A$ (works because Q is orthogonal so $Q^{-1} = Q^T$).

Other Topics -not Mastery

Singular Value Decomposition

- Know some of the applications of SVD
- Know the difference between numerical rank and true rank
- **Steps for Singular Value Decomposition:**
  Step 1: Orthogonally Diagonalize $A^T A$ to find V (V will be our P when $A^T A = PDP^T$).
  Step 2: Find Σ using the singular values of A.
  Step 3: Find U in 2 Steps
  Step 3a: Find the first entries by $u_i = \dfrac{1}{\sigma_i} A v_i$
  Step 3b: Add the normalized vectors from null(A^T)

Column Rank and Left and Right Inverses

- Know the **The Invertible Matrix Theorem: *Restated*** Suppose A is a square $n \times n$. Then the following are equivalent (if one is true, they all are true)
  1. A is invertible
  2. $AX = b$ has a unique solution
  3. $\det A \neq 0$.
  4. A^T is an invertible matrix.
  5. A has full rank/ the columns of A are linearly independent
  6. $\text{nul}(A) = 0$
- Know how to determine if a matrix has full column rank: Every column of the echelon form of the matrix has a leading one.
- Know how to determine if a matrix has full rank: Every column AND row of the echelon form of the matrix has a leading one.
- Know how to determine if a matrix has full column rank: Every column of the echelon form of the matrix has a leading one.
- Know if A has full column rank, $(A^T A)^{-1} A^T$ is a left inverse of A.
- Know if A has full row rank, $A^T (AA^T)^{-1}$ is a right inverse of A.
- Know that while inverses of matrices are unique, but left and right inverses are not unique.
- Know that the set of least-square solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the nonempty set of solutions to the normal equation: $A^T A\mathbf{x} = A^T\mathbf{b}$.

Practice Problems:

1. Orthogonally diagonalize $A = \begin{bmatrix} 2 & 1 \\ 2 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ if you can. Can you do this for A^TA?

2. Find the singular value decomposition for $\begin{bmatrix} 2 & 1 \\ 2 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

3. Orthogonally diagonalize $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

4. Find the singular value decomposition for $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

5. Find an orthogonal basis for S=span$\{ \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \\ -4 \end{bmatrix} \}$

6. Find the Q matrix in the QR Decomposition of $\begin{bmatrix} 1 & 1 & 5 \\ -2 & -4 & 0 \\ 2 & 0 & -4 \end{bmatrix}$

7. Find the eigenvalues and eigenspaces for each associated eigenvalue for the following matrices:

   (a) $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$

   (b) $\begin{bmatrix} 7 & -8 \\ 4 & -5 \end{bmatrix}$

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

8. Diagonalize the following matrices if possible:

   (a) $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$

   (b) $\begin{bmatrix} 7 & -8 \\ 4 & -5 \end{bmatrix}$

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

9. In a linear algebra class of 100 students, on any given day, some are in class and the rest are absent. It is known that if a student is in class today, there is an 85% chance that he/she will be in class tomorrow, and if the student is absent today, there is a 60% chance that he/she will be absent tomorrow. Suppose today there are 76 kids in class.

   (a) Find the transition matrix for this scenario.

   (b) Predict the number of students in class five days from now. And predict the percentage of the class that will be absent.

   (c) Find the steady state vector.