MATH 349 Notes

Vladislav Taranchuk

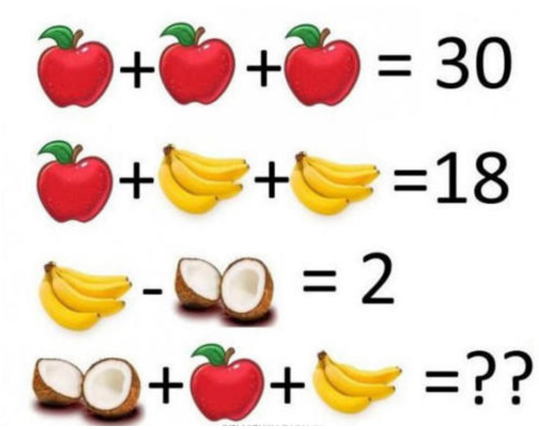
November 18, 2022

Contents

1 Lecture - 08/31/2022

1.1 What is Linear Algebra?

Linear algebra is a fascinating mathematical topic that has applications to just about every scientific field. In some sense, linear algebra is easy, I mean, you have been working with linear equations since 6th grade. You've been plotting lines of the form $y = mx + b$ and solving even *systems of linear equations* without referring to them as such. Linear algebra is so embedded into our daily lives, that some of you have been solving linear algebra problems while scrolling through memes! I'm sure many of you will find this familiar:



Why don't we take all the fun out of it by formalizing what the question above is asking. Let's denote: Apple $= x$, Bananas $= y$, and Coconut $= z$. Now we reduce it to

$$\begin{aligned} 3x &= 30 \\ x + 2y &= 18 \\ y - z &= 2 \\ x + y + z &= ?? \end{aligned}$$

Solution. This particular problem is set up to be fairly straight forward. First we solve for x using the first equation, then substitute our answer into the second equation to find y, likewise for z. Finally, we can solve for $x + y + z$. So $3x = 30$ means $x = 10$. Then $x + 2y = 18$ means $10 + 2y = 18$ so $y = 4$. Finally $y - z = 2$ means $4 - z = 2$ so $z = 2$. So then $x + y + z = 10 + 4 + 2 = 16$.

Of course, solving problems like this is just the tip of the iceberg. Linear algebra is a very deep and rich field, and one which could be explored with some training in how to prove results. Math 349 does not require any proof based mathematics as a pre-requisite, so you will not be expected to regurgitate difficult proofs, though we may encounter some as we make our way through the course. Before we can get started, let us cover come preliminaries regarding notation and certain mathematical notions that will make our lives much easier if we learn to use them effectively.

1.2 Sets

We begin by defining the notion of a set. It is a rather ambiguous notion, yet mathematics is fundamentally built from the ground up using sets. There is a whole branch of mathematics called Set Theory which studies this exact idea. Set Theory aims to formalize the language of mathematics in order to ensure that there are no logical contradictions lurking within the foundation of mathematics.

Definition. A **set** is just a collection of *distinct* things that we wish to group together. The "things" themselves are referred to as *elements* of the set. To denote a set we, enclose the elements of the set with $\{\}$, curly braces.

Examples of Sets:

- { apple, dog, 6, car, John } has elements: 'apple', 'dog', '6', 'car', 'John'.
- The set of people in this class.

To denote that a element belongs to a set, we use the symbol $\in$, and $\notin$ if an element is *not* in a set.
Example 1: Let $A = \{2, 3, 7, 11\}$. Determine if the following are true or false:

1. $3 \in A$ - *True*
2. $4 \in A$ - *False*
3. $12 \notin A$ - *True*
4. $\{2, 3\} \in A$ - *False*

The last one is false in particular because even though the elements 2 and 3 are in the set, $\{2, 3\}$ itself is not an element.

Example 2: We now introduce some common notation for sets that will be frequently used.

- $\emptyset = \{\}$ the empty set.
- $[n] = \{1, 2, \ldots, n\}$
- $\mathbb{N} = \{1, 2, 3, \ldots\}$ is the set of all natural numbers.
- $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ is the set of all integers.
- $\mathbb{Q} = \{\frac{m}{n} : m, n \in \mathbb{Z}\}$
- $\mathbb{R} =$ the set of all real numbers.

Let us consider something a little more difficult, but something we will encounter more often. Make sure you understand why the following is true.

- $\{x \in \mathbb{R} : x^2 - 1 = 0\} = \{-1, 1\}$
- $\{(x, y) \in \mathbb{R}^2 : y = 3x + 2\} = \{(x, 3x + 2) : x \in \mathbb{R}\}$ is the set of all points on the line $y = 3x + 2$.

1.3 Linear Systems

Let us introduce some terms that we will be using regularly.

Definition. A linear equation in n variables $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where the coefficients $a_1, a_2, \ldots, a_n$ and constant term b are specific numbers in $\mathbb{R}$.

Example 3: The following are linear equations:

- $5x + \sqrt{2}y + \pi z = \ln 2$
- $3x_1 + 0.1x_2 = -\frac{1}{1000}x_3$

Definition. A solution to a linear equation $a_1x_1+a_2x_2+\cdots+a_nx_n = b$ is a vector $(s_1, \ldots, s_n)$ such that when we substitute $(x_1, \ldots, x_n) = (s_1, \ldots, s_n)$ (i.e. $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$) into our linear equation, the equality holds.

Example 4: Find one solution to $2x_1 - x_2 = 2$. Can you find more than one?

Definition. A system of linear equations is a finite collection of linear equations. A solution of a system is a vector that is a solution of each linear equation. The solution set of a system is the set of ALL solutions to the system. (This is where set notation will come in handy for us).

Example 5: Find a solution to the system:

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

Solution. There are several ways we can approach this problem. By the end of today, we will develop a systematic method for tackling these sorts of problems. Notice that $x+y = x-y$. This implies that $2y = 0$ so $y = 0$. Now if $y = 0$, then $x = 7$, so $(x, y) = (7, 0)$ is the unique solution.

Example 6: Find a solution to the system:

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

Solution. Let us try solving this one using substitution. Meaning, $x + 2y = 1$ implies $x = 1-2y$. Now it is also the case that $-2x-4y = 0$. So we may substitute for x, to obtain

$$0 = -2x - 4y = -2(1 - 2y) + 4y = -2 - 4y + 4y = -2.$$

So -2 = 0! But how could that be? As it turns out, not every linear system has a solution

Definition. We say that a system of linear equations with coefficients in $\mathbb{R}$ is either

- a consistent system (has at least one solution) or
- an inconsistent system (has no solutions)

For more examples, look at pages 59 and 60 in your textbook.

1.4 Solving Linear Systems

So what is this systematic method for solving linear systems of equations? Let us do some small examples before making an observation which should help us solve larger systems.

Example 7: Find the *solution set* of the equation $x - y = 2$.

Solution. Here we have two variables, so a solution set will be of the form $\{(s_1, s_2) : s_1 - s_2 = 2\}$. But, we want to write the solution so that we use as the least amount of variables to describe it. We observe that $x - y = 2$ means that $x = 2 + y$ or $y = x - 2$. Now, we can parameterize the solution using the established relationship between x and y. We can set $x = s$, and therefore $y = s - 2$, and so a solution is of the form $(x, y) = (s, s - 2)$ for any real number s. So the solution set is

$$\{(s, s - 2) : s \in \mathbb{R}\}.$$

Example 8: Find the solution set of the following linear system:

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

Solution. To solve this system, we can perform a method known as back substitution. We know that $z = 1$, and so we *substitute* z into $y - z = 2$ to obtain that $y = 3$. Finally, $x + y + z = 3$ means that $x = -1$. So $(-1, 3, 1)$ is the *unique* solution, and we do not need to introduce any new parameters to describe the full solution set, since the solution set has only one element, namely $(-1, 3, 1)$.

Given a linear system, we can summarize the system in what is called augmented matrix form. We collect the coefficients by each variable, as well as the constant term and build a matrix. It is best described using some examples:

Example 9: The augmented matrix form of the linear system

$$\begin{aligned} x + y + z &= 3 \\ y - z &= 2 \\ z &= 1 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

The augmented form of a matrix is particularly handy as we will see in the next section. To understand why, we introduce the notion of equivalent systems of linear equations

Definition. We say that two linear systems are equivalent if they're solution sets are equal.

Example: Prove that the following two linear systems are equivalent.

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

Once in augmented matrix form, there are ways to manipulate a matrix and change the linear system to an equivalent one, but somewhat simpler looking. In particular, there are three elementary row operations that can be done on an augmented matrix that preserve the solution set of the linear system. You can:

1. Add a scalar multiple of one row to another,
2. swap any two rows,
3. scale a row by any number.

I will leave it to you to think about why these operations do not affect the solution set of a linear system. The point of doing row operations is to eliminate variables and obtain a form of a matrix on which we can solve using back substitution. We have a name for the form of such a matrix.

Definition. We say that a matrix is in row echelon form if:

1. Any rows of zeroes are at the very bottom.
2. In each row, the first non-zero entry from the left, called the leading entry of that row appears in a column to the left of the leading entries in rows above it.

Exercise: Which of the following matrices are in row echelon form?

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

Example: Find the solution set of the following linear system:

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

Solution. We will cover this next class.

2 Lecture - 09/07/2022

Today we will cover material from Sections 2.2 in the book. Section 2.2 aims to introduce us to common techniques used for solving systems of linear equations. So far, we have learned that what we should try is turn a system into its corresponding augmented matrix form, and then reduce the augmented matrix to row echelon form using elementary row operations. Let's finish the example we started last time.

Example 1: Find the solution set of the following linear system:

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

Solution. First we convert the system to augmented matrix form, and then perform row operations on it.

$$\left[\begin{array}{ccc|c} 1 & 1 & -4 & -2 \\ 3 & -1 & 2 & 4 \\ -2 & 1 & 3 & 2 \end{array}\right] \xrightarrow{R_2 - 3R_1} \left[\begin{array}{ccc|c} 1 & 1 & -4 & -2 \\ 0 & -4 & 14 & 10 \\ -2 & 1 & 3 & 2 \end{array}\right] \xrightarrow{R_3 + 2R_1} \left[\begin{array}{ccc|c} 1 & 1 & -4 & -2 \\ 0 & -4 & 14 & 10 \\ 0 & 3 & -5 & -2 \end{array}\right]$$

$$\xrightarrow{R_3 + \frac{3}{4}R_2} \left[\begin{array}{ccc|c} 1 & 1 & -4 & -2 \\ 0 & -4 & 14 & 10 \\ 0 & 0 & \frac{11}{2} & \frac{11}{2} \end{array}\right]$$

Which implies that the original linear equation is equivalent to

$$\begin{aligned} x_1 + x_2 - 4x_3 &= -2 \\ -4x_2 + 14x_3 &= 10 \\ \tfrac{11}{2}x_3 &= \tfrac{11}{2} \end{aligned}$$

So then $\frac{11}{2}x_3 = \frac{11}{2} \implies x_3 = 1$. Then $-4x_2 + 14x_3 = 10 \implies x_2 = 1$, and finally, $x_1 + x_2 - 4x_3 = -2 \implies x_1 = 1$. Which means the system has a unique solution since we do not need to parameterize any variables.

2.1 Introduction to Proofs

Definition. We say that two matrices A and B are row equivalent if there is a sequence of elementary row operations that can turn A into B, or equivalently, B into A.

Theorem 2.1.1. *Matrices A and B are row equivalent if and only if they can be reduced to the same row echelon form.*

Before we proceed with a proof, let us mention a few things about logical statements and proving things. The theorem can be understood to have two logical statements, in particular they are:

1. Matrices A and B are row equivalent, **AND**

2. Matrices A and B can be reduced to the same row echelon form.

The theorem has the term *if and only if*, (you will see it abbreviated as *iff* often times in the future). "If and only if" means that we need to prove that the two statements above *imply* each other. When a statement P implies a statement Q, we write $P \implies Q$. When P is true if and only if Q is true, then we write $P \iff Q$.

In essence, the theorem above says the following two statements:

1. ($\implies$) If matrices A and B are row equivalent, then they can be reduced to the same row echelon form, **AND**

2. ($\impliedby$) If matrices A and B can be reduced to the same row echelon form, then they are row equivalent.

We will now prove the theorem.

Proof. ($\implies$): We prove the first statement. If A and B are row equivalent, then there is a sequence of elementary row operations R such that $A \to B$. After applying these row operations, we obtain the matrix B. But then we can reduce B into row echelon form using a sequence of elementary row operations. So both A and B can be reduced to the same row echelon form using some sequence of row operations.

($\impliedby$): Now suppose that A and B can both be turned into the same row echelon form R via some sequence of elementary row operations. But then we can undo the operations done to B to go from $A \to R \to B$ and vice versa. So A and B are row equivalent. □

2.2 Gaussian Elimination (with back substitution)

We introduced the notion of equivalent systems last class. Note then, given two augmented matrices that represent some linear system. If the two matrices are row equivalent, then they are equivalent in the sense that they have the same solution set. **So the goal of solving systems now is to reduce systems into row echelon form using elementary row operations, because the solution set does not change!**

The particular method we have already been suing is called Gaussian elimination, named after the mathematician Carl Friedrich Gauss. To summarize, Gaussian elimination requires us to follow these three steps. (Technically, Gaussian elimination is just step 2, and step 3 is called back-substitution.)

1. Write the linear system in augemented matrix form.

2. Use elementary row operations to reduce the augmented matrix to row echelon form.

3. Using back substitution, solve the system.

Example: Let us do Example 2.11 in the book, at the bottom of page 70.

After reducing a system to row echelon form, we say that columns with leading entries correspond to leading variable and columns that have no leading entries correspond to free variables. The number of free variables, will be precisely the number of distinct parameters in our solution. In the example above, there were two columns without leading entries, and two distinct parameters s and t.

3 Lecture - 9/12/2022

3.1 Gauss-Jordan Elimination

There is another method that we can sue to solve linear systems that is in some sense a continuation of Gaussian elimination. Gaussian elimination has us reduce a matrix into row echelon form, while Gauss-Jordan elimination will take this idea even further.

Definition. A matrix is said to be in reduced row echelon form the following properties are all true:

1. The matrix is in row echelon form.
2. The leading entry in any non-zero row is called a 1.
3. Any column with a leading 1 has all other entries as 0.

So how does one turn a matrix into reduced row echelon form? We summarize the answer to this question in following, called Gauss-Jordan Elimination:

1. Write linear system as augmented matrix.
2. Reduce the matrix into row echelon form.
3. Divide each row with a leading entry by the leading entry to obtain a leading 1.
4. Use each row with a leading 1 to subtract a multiple of it from the rows above, so that the only numbers in the same row as the leading 1, are all zero.

Example: Let's do the example we did at the beginning of last class. Use Gauss-Jordan Elimination to find the solution set of the following linear system.

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

Solution. Recall that, we already converted this system to row echelon form, so let us begin there. We performed a sequence of elementary row operations to obtain

$$\left[\begin{array}{ccc|c} 1 & 1 & -4 & -2 \\ 3 & -1 & 2 & 4 \\ -2 & 1 & 3 & 2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & -4 & -2 \\ 0 & -4 & 14 & 10 \\ 0 & 0 & \frac{11}{2} & \frac{11}{2} \end{array}\right].$$

Observe that this system has 3 leading variables and no free variables according to the definitions we gave. Now continuing on we still need to complete steps 3 and 4 of the Gauss-Jordan elimination process. First, we do step 3, to obtain:

$$\left[\begin{array}{ccc|c} 1 & 1 & -4 & -2 \\ 0 & -4 & 14 & 10 \\ 0 & 0 & \frac{11}{2} & \frac{11}{2} \end{array}\right] \xrightarrow{-\frac{1}{4}R_2, \quad \frac{2}{11}R_3} \left[\begin{array}{ccc|c} 1 & 1 & -4 & -2 \\ 0 & 1 & -\frac{7}{2} & -\frac{5}{2} \\ 0 & 0 & 1 & 1 \end{array}\right]$$

Proceeding to step 4, we now eliminate starting from the bottom, going up.

$$\left[\begin{array}{ccc|c} 1 & 1 & -4 & -2 \\ 0 & 1 & -\frac{7}{2} & -\frac{5}{2} \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{R_2 + \frac{7}{2}R_3} \left[\begin{array}{ccc|c} 1 & 1 & -4 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow{R_1 + 4R_3} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\xrightarrow{R_1 - R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

Which states exactly what we found before, $x_1 = 1, x_2 = 1, x_3 = 1$.

Definition. A linear system is called homogeneous if the constant term for each linear equation is 0. That is, each linear equation is equal to 0. Otherwise, the system is called, inhomogeneous.

Example: Find the solution set of the following linear system.

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

Solution. Let us convert this system to augmented matrix form and solve using Gauss-Jordan elimination. We have

$$\left[\begin{array}{cccc|c} 1 & 1 & 4 & 2 & 0 \\ 1 & 1 & 2 & -2 & 0 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cccc|c} 1 & 1 & 4 & 2 & 0 \\ 0 & 0 & -2 & -4 & 0 \end{array}\right]$$

$$\xrightarrow{-\frac{1}{2}R_2} \left[\begin{array}{cccc|c} 1 & 1 & 4 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 \end{array}\right] \xrightarrow{R_1 - 4R_2} \left[\begin{array}{cccc|c} 1 & 1 & 0 & -6 & 0 \\ 0 & 0 & 1 & 2 & 0 \end{array}\right]$$

Notice that columns 1 and 3 have leading 1's, but columns 2 and 4, do not. So this system has 2 free variables. We can solve for x_1 and x_3 in terms of the free variables x_2 and x_4.

$$\left.\begin{aligned} x_1 + x_2 + 4x_3 + 2x_4 &= 0 \\ x_1 + x_2 + 2x_3 - 2x_4 &= 0 \end{aligned}\right\} \iff \begin{aligned} x_1 + x_2 \qquad - 6x_4 &= 0 \\ x_3 + 2x_4 &= 0 \end{aligned}$$

So we must have that $x_3 = -2x_4$ and $x_1 = -x_2 + 6x_4$. Let us now parameterize our solution. Set $x_2 = s$ and $x_4 = t$, then the solutions are of the form

$$\begin{bmatrix} -s + 6t \\ s \\ -2t \\ t \end{bmatrix} = \begin{bmatrix} -s \\ s \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 6t \\ 0 \\ -2t \\ t \end{bmatrix} = s \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 6 \\ 0 \\ -2 \\ 1 \end{bmatrix}.$$

Let $\overline{v} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and $\overline{w} = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 1 \end{bmatrix}$ and the solution set to the system is

$$\{s\overline{v} + t\overline{w} : s, t \in \mathbb{R}\}.$$

Important: A few things worth brining attention to:

1. Row Echelon form means to use elementary row operations to reduce a system to an upper triangular form. That is any entries below the diagonal of a matrix, are 0.
2. Homogeneous systems always have at least one solution, namely the zero vector.
3. A linear systems has either: no solution, one solution, or infinitely many solutions. If a system is consistent and has a free variable, then it has infinitely many solutions. If it is consistent with no free variables, then it has one solution.

3.2 Linear Combinations and Span

Definition. Let $S = \{v_1, \dots, v_k\} \subset \mathbb{R}^n$ be a finite set of vectors. A linear combination of vectors in S is

$$v = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k$$

for some $a_1, \dots, a_k \in \mathbb{R}$.

Definition. The span of a set of vectors $S = \{v_1, \dots, v_k\} \subset \mathbb{R}^n$, is the set of all linear combinations of these vectors, denoted

$$\text{span}(S) = \{a_1 v_1 + a_2 v_2 + \cdots + a_k v_k : a_1, \dots, a_k \in \mathbb{R}\}.$$

If $\text{span}(S) = \mathbb{R}^n$, then S is called a spanning set for $\mathbb{R}^n$.

Example: Is the vector $u = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in the span of the vectors $v = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ and $w = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$?

Solution. The way these problems are approached are always the same way. We suppose that it is in the span, and then attempt to find the correct coefficients that solve our problems. We must see if there exist numbers $x_1, x_2 \in \mathbb{R}$ such that

$$u = x_1 v + x_2 w.$$

In other words, we wish to know if there is a solution to

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

Does this look familiar? It should! We are just asking to see if a particular system of linear equations has a solution! So let's turn it into augmented matrix form, and reduce it. The augmented matrix form is given by

$$\left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{array}\right] \xrightarrow{R_3 + R_1} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{array}\right] \xrightarrow{R_3 - R_2} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{array}\right]$$

This system is inconsistent, so u is not in the span of v and w.

Remark: The approach for solving these kinds of problems is always the same. When asked to determine if a vector b is in $\text{span}(\{v_1, \dots, v_k\})$, define the matrix A to have as columns the vectors $v_1, \dots, v_k$, and attempt to solve the augmented matrix system $[A|b]$. If the system is consistent, then b is in the span, otherwise, b is not in the span.

4 Lecture 9/14/2022

4.1 Linear Dependence and Independence

Today's lecture will focus on Section 2.3 in the book and some potential applications of what we have learned so far.

Important: Suppose we have a $m \times n$ matrix A (just a regular matrix, not an augmented matrix), with $n > m$. Here, and everywhere else, m is the number of rows and n is the number of columns. So a matrix with more columns, than rows. Note that when we reduce A to row echelon form, each row has at most one leading entry. But that also means that each column has at most one leading entry. So, **if there are more columns, than rows, you are guaranteed to have a free variable!** However, this does not mean that the system is necessarily consistent. But it does mean that if the system is consistent, then it must have infinitely many solutions.

Theorem 4.1.1. *Let A be an $m \times n$ matrix with $n > m$, then $[A|0]$ has infinitely many solutions, particularly, the solution set contains more than just the zero vector.*

Proof. We mentioned that if A is $m \times n$ with $n > m$, then the row echelon form always has more columns than leading entries, so it has free variables. We also know that $[A|0]$ is always consistent. Therefore, by the comment above, $[A|0]$ must have infinitely many solutions. □

Notation: Denote $\mathbb{R}^n$ by

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} : x_1, \dots, x_n \in \mathbb{R} \right\}.$$

We say that a set S is a subset of $\mathbb{R}^n$, denoted $S \subset \mathbb{R}^n$, if each vector $v \in S$ we have $v \in \mathbb{R}^n$.

Definition. The set of vectors $\{v_1, \dots, v_k\} \subset \mathbb{R}^n$ is said to be linearly dependent, if there exists scalars (real numbers), $c_1, \dots, c_k$, that are not all zero, such that

$$c_1 v_1 + \cdots + c_k v_k = 0.$$

A set that is not linearly independent is called linearly independent.

Theorem 4.1.2. *The vectors $v_1, \dots, v_k \in \mathbb{R}^n$ are linearly dependent if and only if at least one of the vectors can be expressed as a linear combination of the others.*

Proof. Since the theorem says if and only if, then each statement should imply the other.

($\implies$) : Suppose that $v_1, \dots, v_k$ are linearly dependent. Then there exists scalars $c_1, \dots, c_k$ such that

$$c_1 v_1 + \cdots + c_k v_k = 0.$$

We can always rearrange the order of the vectors, so that the non-zero scalar appears as c_1. So suppose $c_1 \neq 0$, then

$$c_2 v_2 + \cdots c_k v_k = -c_1 v_1 \implies v_1 = -\frac{c_2}{c_1} v_2 - \cdots - \frac{c_k}{c_1} v_k.$$

So v_1 can be written as a linear combination of the other vectors.

($\Longleftarrow$): Now suppose that one of the vectors, without loss of generality, we can assume it is v_1, can be written as a linear combination of the other vectors. Then, there exists scalars $c_2, \ldots, c_k$ such that

$$c_2v_2 + \cdots c_kv_k = v_1 \implies -v_1 + c_2v_2 + \cdots c_kv_k = 0.$$

Thus, the vectors $v_1, \ldots, v_k$ are linearly dependent. $\square$

Example: Is the set $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ linearly independent?

Solution. Let v_1 be the first vector in the set, v_2, the second, and v_3 the third. To answer the question, we need to understand whether or not there exist scalars c_1, c_2, c_3 such that $c_1v_1 + c_2v_2 + c_3v_3 = 0$. But that's just asking us to solve a homogeneous system of linear equations. Starting with the augmented matrix form, we have

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] \xrightarrow{R_3 - R_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right]$$

Now we can either perform back substitution or continue on with Gauss-Jordan Elimination. I will continue with Gauss-Jordan elimination.

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right] \xrightarrow{(1/2)R_3} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \xrightarrow{R_2 + R_3} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \xrightarrow{R_1 - R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

The above tells us that the only solution is $c_1 = c_2 = c_3 = 0$, and therefore the vectors are *linearly independent*.

So we see that the question of linear independence also comes back to solving a linear system of equations. We summarize this observation in a theorem.

Theorem 4.1.3. *Let $v_1, \ldots, v_n$ be column vectors in $\mathbb{R}^m$ and let A be the $n \times m$ (m rows, n columns) matrix $[v_1v_2 \ldots v_n]$. Then $v_1, \ldots, v_n$ are linearly dependent if and only if the homogeneous system $[A|0]$ has a nontrivial solution.*

Corollary. *Let $v_1, \ldots, v_n$ be column vectors in $\mathbb{R}^m$ and suppose $n > m$, then $v_1, \ldots, v_n$ are linearly dependent.*

Proof. We have that $v_1, \ldots, v_n$ are linearly dependent if and only if $[A|0]$ has a non-trivial solution. Since $n > m$, then by Theorem 4.0.1 in the notes, $[A|0]$ has infinitely many solutions, and so has a non-trivial solution. Thus $v_1, \ldots, v_n$ are linearly dependent according to Theorem 4.0.3 above. $\square$

Remark: Putting together the ideas of span and linear independence, we observe that a set of linearly independent vectors allows us to have a representation of an infinite space using a finite number of vectors and considering it's span. When the finite set of vectors contains linearly *dependent* vectors, then it contains redundant information, so linear *independence* guarantees us unique solutions to systems, if the system is consistent.

4.2 Applications

Here we will briefly touch on a application of linear algebra to real world problems. The material will come from section 2.4. You will **NOT** be tested on application type problems, but it is important to recognize the use of linear algebra outside the classroom.

Example: Suppose that a company produces desks, tables, and chairs. The resources it requires is

- Lumber per foot
- Screw per pound
- Labor per hour

Furthermore, suppose that the company requires the following amount of each resource to build each product

- Desk: 20 feet of lumber, 2 pounds of screws, 6 hours labor.
- Chair: 6 feet of lumber, 1 pound of screw, 4 hours labor.
- Table: 15 feet of lumber, 3 pounds of screws, 3 hours labor.

If in the next week, the company has 227 feet of lumber, 35 pounds of screws, and 87 hours of labor available, how many tables, desks, and chairs can they make?

Solution. We need to construct a system of linear equations that we will then go ahead and proceed to solve. Let x_1 denote the number of desks made, x_2 the number of chairs made, and x_3 the number of tables. Then we have

$$\begin{aligned} 20x_1 + 6x_2 + 15x_3 &= 227 \text{ (Feet of lumber)} \\ 2x_1 + x_2 + 3x_3 &= 35 \text{ (Pounds of screws)} \\ 6x_1 + 4x_2 + 3x_3 &= 87 \text{ (Hours of labor)} \end{aligned} \iff \left[\begin{array}{ccc|c} 20 & 6 & 15 & 227 \\ 2 & 1 & 3 & 35 \\ 6 & 4 & 3 & 87 \end{array}\right]$$

Now we can proceed to reducing the system. Sometimes it's easier to start the problem by swapping some rows first. For example, it would be rather ugly computation to use the 20 in the top left corner to eliminate the 2 and 6 below. But if the 2 were in the corner, this would be easier. So I begin by swappings rows 1 and 2, and then swap rows 2 and 3. So we have

$$\left[\begin{array}{ccc|c} 2 & 1 & 3 & 35 \\ 6 & 4 & 3 & 87 \\ 20 & 6 & 15 & 227 \end{array}\right] \xrightarrow{R_2 - 3R_1} \left[\begin{array}{ccc|c} 2 & 1 & 3 & 35 \\ 0 & 1 & -6 & -18 \\ 20 & 6 & 15 & 227 \end{array}\right] \xrightarrow{R_3 - 10R_1} \left[\begin{array}{ccc|c} 2 & 1 & 3 & 35 \\ 0 & 1 & -6 & -18 \\ 0 & -4 & -15 & -123 \end{array}\right]$$

$$\xrightarrow{R_3 + 4R_2} \left[\begin{array}{ccc|c} 2 & 1 & 3 & 35 \\ 0 & 1 & -6 & -18 \\ 0 & 0 & -39 & -195 \end{array}\right] \implies \begin{aligned} 2x_1 + x_2 + 3x_3 &= 35 \\ x_2 - 6x_3 &= -18 \\ -39x_3 &= -195 \end{aligned} \xrightarrow{\text{back substitution}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 12 \\ 5 \end{bmatrix}$$

So the company can make 4 desks, 12 chairs, and 5 tables with the given resources.

5 Lecture 9/19/2022

Today we start on Chapter 3 in the textbook, and should cover section 3.1 and 3.2, which are all about matrices and doing algebra with matrices.

Definition. The set $\mathbb{R}^{m\times n}$ is the set of all $m \times n$ matrices. So to say A is an $m \times n$ matrix, we will often just write $A \in \mathbb{R}^{m\times n}$.

Definition. If $A \in \mathbb{R}^{m\times n}$, then we denote it as $A = [a_{ij}]$, where a_{ij} is the entry in the i^{th} row and j^{th} column, that is:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

5.1 Matrix Addition

Just in the same way that we have been dealing with vectors, one can add matrices, entry wise. **You can only add matrices that have the same corresponding dimensions, meaning they are both $m \times n$.** We can also scale matrices by a scalar, which is equivalent to multiplying every entry in the matrix by that scalar. Putting this together we have that if $A, B \in \mathbb{R}^{m\times n}$ and $A = [a_{ij}]$ and $B = [b_{ij}]$, and $c \in \mathbb{R}$. Then if $D = A + cB$, then

$$D = [d_{ij}] = [a_{ij} + cb_{ij}].$$

Example: Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$. Then

$$A - 2B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 2\begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} -8 & -6 \\ -4 & -2 \end{bmatrix} = \begin{bmatrix} -7 & -4 \\ -1 & 2 \end{bmatrix}$$

Theorem 5.1.1. *Let $A, B, c \in \mathbb{R}^{m\times n}$, and $c, d \in \mathbb{R}$, then the following properties of matrix addition hold.*

1. $A + B = B + A$, *(Commutativity)*
2. $(A + B) + C = A + (B + C)$, *(Associativity)*
3. $c(A + B) = cA + cB$, *(Distributivity)*
4. $(c + d)A = cA + dA$, *(Distributivity)*

5.2 Dot Products and Matrix Multiplication

Matrix multiplication may be new to you, and a little less intuitive than one might think, but it is defined in a way that allows us to talk about matrices in many different ways, including thinking of matrix multiplication as composition of function (more on that later in the semester).

In order to define matrix multiplication, we first define the dot product of two *vectors*.

Definition. Let $v, w \in \mathbb{R}^n$ be vectors. Then we define the dot product of v and w as

$$v \cdot w = v_1 w_1 + v_2 w_2 + \cdots v_n w_n.$$

Informally speaking, when taking the product of two matrices A and B, the ij^{th} entry of AB is given by the dot product of the i^{th} row of A with the j^{th} column of B. Note that dot products can only be done when the two vectors have the same number of coordinates. Therefore, when multiplying two matrices, the number of entries in a row of A should equal the number of entries in a column of B. **Equivalently, in order to multiply matrices A and B in the manner AB, A should have the same number of columns as B does rows.** So if $A \in \mathbb{R}^{m\times n}$, then B must be in $\mathbb{R}^{n\times d}$.

Definition. Let $A = [a_{ij}] \in \mathbb{R}^{m\times n}$ and $B \in [a_{ij}] \in \mathbb{R}^{n\times d}$, then if $C = AB$, we have that

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

And $C \in \mathbb{R}^{m\times d}$

Example: Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -4 \\ -2 & 3 \end{bmatrix}$, and $C = \begin{bmatrix} 3 & 2 & 0 \\ -1 & -1 & 1 \end{bmatrix}$. Find AB, AC, BA, BC, CA, CB.

Solution. t We have $A, B \in \mathbb{R}^{2\times 2}$, so the number of rows columns in A equals the number of rows in B, so we *can* multiply them, thus

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & -4 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} 1(1)+2(-2) & 1(-4)+2(3) \\ 3(1)+4(-2) & 3(-4)+4(3) \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ -5 & 0 \end{bmatrix}$$

Here $A \in \mathbb{R}^{2\times 2}$ and $C \in \mathbb{R}^{2\times 3}$. Again the number of columns in A equals the number of rows in C, so we can multiply them, thus

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

$$BA = \begin{bmatrix} 1 & -4 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} -11 & -14 \\ 7 & 8 \end{bmatrix}$$

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

Now, notice that it is not possible to compute CA, nor CB, since C has three columns, but both A and B have only two rows.

Theorem 5.2.1. *Let A, B, C be matrices have corresponding dimensions so that the following matrix multiplication is allowed. Then the following properties of matrix multiplication hold.*

1. $(AB)C = A(BC)$, *(Associativity)*

2. $C(A+B) = CA + CB$, *(Distributivity)*

3. $(A+B)C = AC + BC$, *(Distributivity)*

But, in general, it is NOT true that $AB = BA$. That is, in general, matrix multiplication is NOT commutative, as the example above demonstrates.

6 Lecture 9/21/2022

6.1 Matrix-Vector Multiplication

A vector $v \in \mathbb{R}^n$ written as a column vector can be considered as an $n \times 1$ matrix. So matrix-vector multiplication can be deduced from the definition of matrix multiplication in general. Let $A \in \mathbb{R}^{m\times n}$, $A = [a_{ij}]$, and $v \in \mathbb{R}^n$. Then the multiplication Av makes sense, and Av becomes a $m \times 1$ vector. In particular, we have

$$Av = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix}$$

In particular, observe that Av is just a linear combination of the columns of A with v_i as the coefficients.

$$\begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix} = v_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + v_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + v_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

A big part of this class, will be determining when $Ax = b$ has a solution, for some matrix A and corresponding vectors x and b. The key observation of solving $Ax = b$ is that this is equivalent to solving a linear system, as we have been doing.

Example: Let $A = \begin{bmatrix} 3 & 2 & 0 \\ -1 & -1 & 1 \\ 2 & -1 & 5 \end{bmatrix}$ and $v = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$. Find Av.

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

To see the connection between solving linear systems, and matrix multiplication, consider the following linear system.

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

Note that if we solve this system, the solution will precisely be the vector $v = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$.

Why? Well, consider $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. What is Ax?

$$Ax = \begin{bmatrix} 3 & 2 & 0 \\ -1 & -1 & 1 \\ 2 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3x_1 + 2x_2 \\ -x_1 - x_2 + x_3 \\ 2x_1 - x_2 + 5x_3 \end{bmatrix}$$

Which is precisely the system of equations we started with. So, when given a system of linear equations to solve, in the form

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

Then solving this linear system, is equivalent to finding all vectors x such that

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

has a solution.

6.2 Inverse of a Matrix

This section corresponds to section 3.3 in the book. We will focus on square matrices in this section, and many of the future sections.

Definition. The identity matrix $I \in \mathbb{R}^{n\times n}$, is a matrix such that for any other matrix $A \in \mathbb{R}^{n\times n}$,

$$IA = A = AI.$$

In particular,

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Definition. If $A \in \mathbb{R}^{n\times n}$, an inverse of A, is a matrix B such that

$$AB = I = BA.$$

If A has an inverse, then A is called invertible.

Theorem 6.2.1. *Let $A \in \mathbb{R}^{n\times n}$ be invertible, then the inverse of A is unique and is denoted A^{-1}.*

Proof. Let B and C both be matrices such that $AB = I = BA$ and $AC = I = CA$. If the inverse is unique, then we need to show that $B = C$. Note that

$$AB = I \implies C = CI = C(AB) = (CA)B = IB = B.$$

□

Theorem 6.2.2. *If $A \in \mathbb{R}^{n\times n}$ is invertible, then $Ax = b$ always has a unique solution, namely $x = A^{-1}b$.*

Proof. All we need to show is that $x = A^{-1}b$ s a solution, and that there are no others. First we check that it is a solution.

$$Ax = A(A^{-1})b = (AA^{-1})b = Ib = b.$$

So it is a solution. Now suppose that y is also a solution to $Ax = b$. Then $Ay = b$, but

$$Ay = b \implies A^{-1}(Ay) = A^{-1}b \implies y = A^{-1}b,$$

so the solution to $Ax = b$ is unique. □

But how does one compute the inverse of a matrix? Well, it is not so dissimilar to what we have already been doing. Instead of augmenting a matrix A with a vector to solve for it, we will augment a matrix A with the identity matrix I on the right hand side, like $[A|I]$, and then reduce the left hand side to Reduced Row Echelon Form (RREF). What will appear on the right hand side once we have reduced A completely (using Gauss-Jordan Elimination) will be the inverse.

Example: Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. A is invertible, find A^{-1}.

Solution. As mentioned, we augment A with I and reduce A to RREF.

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right] \xrightarrow{R_2 - 3R_1} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right]$$

$$\xrightarrow{R_1 + R_2} \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{array}\right] \xrightarrow{-(1/2)R_2} \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & 1/2 \end{array}\right]$$

Now that A has been reduced to RREF, the matrix appearing on the right hand side willbe the inverse of A, that is

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

It is no coincidence that the RREF form of A is I. This is in fact true for any invertible matrix.

Theorem 6.2.3. *Let $A \in \mathbb{R}^{n\times n}$ be invertible. Then the reduced row echelon form of A is the identity matrix I.*

Proof. Recall that A is invertible implies that $Ax = b$ has a unique solution. Then that means $Ax = 0$ has a unique solution, just the 0 solution. So A in RREF cannot have free variables, meaning there must be a leading 1 in each row. Since A is a square, then A has a leading 1 in each column as well. But, RREF means above and below each leading 1, we should have zeroes. This is precisely the identity matrix I. □

7 Lecture 9/26/2022

7.1 Inverses Continued

Theorem 7.1.1. *If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then A is invertible if and only if $ad - bc \neq 0$, in which case*

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Theorem 7.1.2. *If A is invertible, then*

1. $(A^{-1})^{-1} = A$.
2. $(cA)^{-1} = \frac{1}{c}A^{-1}$, *for any* $c \in \mathbb{R}$.
3. *If B is also invertible of the same size as A, then*
$$(AB)^{-1} = B^{-1}A^{-1}.$$
4. $(A^n)^{-1} = (A^{-1})^n$, *for any positive integer* n.

For some proofs, see page 168 in the book.

Examples: Let $A = \begin{bmatrix} 2 & -2 \\ 3 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -5 \\ 0 & 1 \end{bmatrix}$. Find

1. A^{-1}
2. B^{-1}
3. $(AB)^{-1}$

Solution. To find the inverses of A and B, we can just use the formula given.

$$A^{-1} = \frac{1}{2(2)-(3)(-2)}\begin{bmatrix} 2 & 2 \\ -3 & 2 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 2 & 2 \\ -3 & 2 \end{bmatrix}$$

and

$$B^{-1} = \frac{1}{1(1)-0}\begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}$$

Finally,

$$(AB)^{-1} = B^{-1}A^{-1} = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}\left(\frac{1}{10}\begin{bmatrix} 2 & 2 \\ -3 & 2 \end{bmatrix}\right) = \frac{1}{10}\begin{bmatrix} -13 & 12 \\ -3 & 2 \end{bmatrix}$$

7.2 Elementary Matrix

Last time, we performed an example of how to find the inverse of a matrix. We will now attempt to understand the theory of why this method works. We have discussed elementary row operations before, but we have not mentioned, is that perform an elementary row operation on a matrix A is the same as multiplying the matrix A by a matrix E called an *elementary matrix*.

Definition. An Elementary Matrix is any matrix that can be obtained by performing an elementary row operation on the identity matrix.

Examples: Which of the following are elementary matrices. For each elementary matrix, what row operation was performed?

1. $E_1 = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
2. $E_2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$
3. $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Solution.

1. E_1 is an elementary matrix. It can be obtained by performing $R_1 - 4R_3$ to I.
2. E_2 is not an elementary matrix. Because there is no single row operation that can be done to I to obtain E_2. Instead one would need to do $R_1 R_3$, and then $R_2 \leftrightarrow R_3$, which is two operations.
3. E_3 is an elementary matrix that can be obtained from I by performing $15R_2$. Multiplying row 2 by 15.

Theorem 7.2.1. *Let E be the elementary matrix obtained by performing an elementary row operation on the identity matrix $I_n \in \mathbb{R}^{n\times n}$. If the same row operation is performed on $A \in \mathbb{R}^{n\times m}$, then the resultant matrix is equal to EA.*

Examples: Write down the elementary matrix E needed to take the given matrix A to the matrix A' so that $EA = A'$.

1. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \to A' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

2. $A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 5 & 0 \\ 0 & 1 & -2 \end{bmatrix} \to A' = \begin{bmatrix} 0 & -13 & 1 \\ 1 & 5 & 0 \\ 0 & 1 & -2 \end{bmatrix}$

Solution.

1. To go from A to A', one needs to perform $R_3 + 4R_1$. Equivalently, we must multiply by $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$, so that $EA = A'$.

2. Here, we need to perform $R_1 - 2R_2$ on A. Equivalenty, we can multiply A on the left by $E = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Theorem 7.2.2. *Let E be an elementary matrix obtained by performing some row operation on I, then E^{-1} exists, and is also an elementary matrix obtained from I by undoing the operation performed by E.*

1. *If E swaps row i and row j, then E^{-1} swaps rows i and j again. That is $E^{-1} = E$.*

2. *If E performs $R_i + cR_j$, then E^{-1} performs $R_i - cR_j$.*

3. *If E performs cR_i, then E^{-1} performs $\frac{1}{c}R_i$.*

Recall that when reducing an invertible matrix to RREF, then it's RREF form is precisely I, the identity matrix. We have also just shown that each elementary row operation can be represented as matrix multiplication. So suppose that A is an invertible matrix, and we perform the operations given by elementary matrices $E_1, \ldots, E_k$ to A to bring it to RREF. Beginning with E_1, then E_2, etc. But this means that the product

$$E_k \cdots E_2 E_1 A = I. \implies A^{-1} = E_k \cdots E_1.$$

This is precisely the idea that is mimicked by the method we introduced for finding the inverse of a matrix.

Example: Let $A = \begin{bmatrix} 1 & -3 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{bmatrix}$. Write down the product of elementary matrices with A such that $E_k \cdots E_1 A = I$. Then using matrix multiplication, determine A^{-1}.

Solution. Let us perform row operations until we have reduced the matrix to RREF, then, if we kept track of the row operations, we will be able to build the matrices. We have

$$\begin{bmatrix} 1 & -3 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{bmatrix} \xrightarrow{(1/2)R_3} \begin{bmatrix} 1 & -3 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_2 + 2R_3} \begin{bmatrix} 1 & -3 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$\xrightarrow{R_1 - R_3} \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 + 3R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Now we can represent each row operation with an elementary matrix.

1. $(1/2)R_3 \longrightarrow E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$.

2. $R_2 + 2R_3 \longrightarrow E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

3. $R_1 - R_3 \longrightarrow E_3 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

4. $R_1 + 3R_2 \longrightarrow E_4 = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

So then

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

and

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

8 Lecture 9/28/2022

Theorem 8.0.1 (The Fundamental Theorem of Invertible Matrices: Part 1)**.** *Let $A \in \mathbb{R}^{n\times n}$, then the following are equivalent:*

1. *A is invertible.*
2. *$Ax = b$ has a unique solution for each $b \in \mathbb{R}^n$.*
3. *$Ax = 0$ has only the trivial solution.*
4. *The columns of A are linearly independent.*
5. *The reduced row echelon form of A is the identity $I \in \mathbb{R}^{n\times n}$.*
6. *A is the product of elementary matrices.*

The bits here that are new that are worth bringing attention to are parts 3 and 4 in the above theorem. We will prove that in fact statement 3 implies statements 2 and that statement 3 is in fact equivalent to statement 4.

END OF EXAM 1 MATERIAL

8.1 Subspaces

We will now continue on to Section 3.5 in the book. We will finally get to define terms you may have heard already, words like *basis, dimension, vector space, subspace, and rank.* The set $\mathbb{R}^n$ is called a vector space over $\mathbb{R}$.

Definition. A subspace of $\mathbb{R}^n$ is a subset S of $\mathbb{R}^n$ satisfying the following conditions:

1. $0 \in S$, (the zero vector)
2. If $u, v \in S$, then $u + v \in S$ (closed under addition)
3. If $u \in S$ and $c \in \mathbb{R}$, then $cu \in S$ (closed under scalar multiplication)

We have already encountered many sets that are subspaces of $\mathbb{R}^n$. The following theorem illuminates this connection for us.

Theorem 8.1.1. *Let $v_1, \ldots, v_k \in \mathbb{R}^n$, then $S = span(v_1 \ldots, v_k)$ is a subspace of $\mathbb{R}^n$.*

Proof. All we need to do is verify that axioms of a subspace hold for the set S.

1. $0 \in S$ since the linear combination $0 = 0v_1 + \cdots 0v_k \in S$.
2. If $u, w \in S$, then u, w are linear combinations of vector $v_1, \ldots, v_k$. So $u = \alpha_1 v_1 + \cdots \alpha_k v_k$ and $w = \beta_1 v_1 + \cdots \beta_k v_k$, and therefore $u + w = (\alpha_1 + \beta_1)v_1 + \cdots + (\alpha_k + \beta_k)v_k$ is also a linear combination, thus $u + w \in S$.
3. Finally, if $u \in S$, and $c \in \mathbb{R}$, then if $u = \alpha_1 v_1 = \cdots + \alpha_k v_k$, then $cu = c\alpha_1 v_1 + \cdots c\alpha_k v_k$, which again is a linear combination of $v_1, \ldots, v_k$ so $cu \in S$

Therefore span$(v_1, \ldots, v_k)$ is a subspace of $\mathbb{R}^n$. □

Example: Let $S \subset \mathbb{R}^3$ be the set of vectors satisfying $x_1 = x_2 + x_3$, the $S = \{(x_2+x_3, x_2, x_3) : x_2, x_3 \in \mathbb{R}\}$. Is S a subspace of $\mathbb{R}^3$.

Solution. Yes, S is a subspace of $\mathbb{R}^3$, and we can verify it.

1. Is $0 \in S$? Yes, since if $x_2 = x_3 = 0$, then the correpsonding vector in S is $(0+0, 0, 0) = (0, 0, 0)$.
2. Let $(x_2 + x_3, x_2, x_3), (y_2 + y_3, y_2, y_3) \in S$, then
$$(x_2 + x_3, x_2, x_3) + (y_2 + y_3, y_2, y_3) = ((x_2 + y_2) + (x_3 + y_3), x_2 + y_2, x_3 + y_3) \in S.$$
3. Finally, we have if $c \in \mathbb{R}$ and $(x_2 + x_3, x_2, x_3) \in S$, then
$$c(x_2 + x_3, x_2, x_3) = (cx_2 + cx_3, cx_2, cx_3) \in S.$$

There are also some subspaces that are important to know that are related to a matrix. Before we introduce them, let us first prove that one of them is in fact a subspace.

Theorem 8.1.2. *Let $A \in \mathbb{R}^{m\times n}$. Define null$(A) = \{v \in \mathbb{R}^n : Av = 0\}$. Then null$(A)$ is a subspace of $\mathbb{R}^n$.*

Proof. We just need to check that the three axioms of a subspace hold. Note that null(A) is in fact a subset of $\mathbb{R}^n$, since $A \in \mathbb{R}^{m\times n}$, then the product Av is only allowed when $v \in \mathbb{R}^n$.

1. $A0 = 0$, therefore, $0 \in$ null(A):
2. Let $u, v \in$ null(A), then $A(u + v) = Au + Av = 0 + 0 = 0$, therefore $u + v \in$ null(A).
3. Let $u \in$ null(A) and $c \in \mathbb{R}$, then $A(cu) = c(Au) = c0 = 0$, so $cu \in$ null(A).

Thus null(A) is a subspace of $\mathbb{R}^n$. □

Definition. Let $A \in \mathbb{R}^{m\times n}$.

1. The row space of A is the subspace row(A) of $\mathbb{R}^n$ spanned by the rows of A.
2. The column space of A is the subspace col(A) of $\mathbb{R}^m$ spanned by the columns of A.
3. The null space of A is the subspace null(A) of $\mathbb{R}^n$, defined by null$(A) = \{v \in \mathbb{R}^n : Av = 0\}$.

Recall that if $A \in \mathbb{R}^{m\times n}$, then Ax is a vector that is a linear combination of the columns of A. So it is also true that col$(A) = \{Av : v \in \mathbb{R}^n\}$.

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

Solution. Note that we have already solved this exact type of problem before. All we need to do is find the solution set to the system $[A|0]$. The solution set is precisely null(A). Note that the solution set is obtained from $x_1 + x_2 = 0$, so $x_1 = -x_2$. Let $s = x_2$, so then

$$\text{null}(A) = \left\{ s \begin{bmatrix} -1 \\ 1 \end{bmatrix} : s \in \mathbb{R} \right\}$$

8.2 Basis and Dimension

We now come to fundamental results that allow us to classify subspaces according to "size". The subspaces of a space are infinite there is a notion in which we can compare their size, and this will relate to a property called dimension. Before we can talk about the dimension of a space, we must first introduce the notions that allow us to quantify something like dimension.

Definition. A basis of a subspace $S \subset \mathbb{R}^n$ is a collection of vectors $\{u_1, \ldots, u_k\} \subset S$ such that:

1. the vectors $u_1, \ldots, u_k$ span S.
2. the vectors $u_1, \ldots, u_k$ are linearly independent.

Example: Do the following sets form a basis for $\mathbb{R}^3$?

- $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$
- $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$
- $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Solution.

The last example gives us some insight into something. A basis for $\mathbb{R}^3$ should have precisely 3 vectors, no more, no less. In this way, we can say that $\mathbb{R}^3$ *has dimension 3*. This is the way in which we can quantify the *size* of a vector space.

9 Lecture 10/2/2022

9.1 Bases and Dimension

Theorem 9.1.1. *Let $S \subset \mathbb{R}^n$ be a subspace. Then any two bases for S must have the same size.*

The proof I will follow is the exact one offered in the book, see Theorem 3.2.3, page 202.

Definition. If $S \subset \mathbb{R}^n$ is a subspace, then the number of vectors in a basis for S is called the <u>dimension</u> of S.

Example: Let $A = \begin{bmatrix} 1 & 1 & 2 & 0 \\ -1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 2 \\ 1 & 3 & 7 & 3 \end{bmatrix}$. Find a basis and determine the dimension of each of the following spaces.

1. row(A)
2. col(A)
3. null(A).

Solution. First we solve part (1). In essence, this question is again asking us to solve a problem that we already know how to solve. When we perform row operations, we eliminate any linear dependence between the rows, so we should turn the matrix into REF form.

$$A = \begin{bmatrix} 1 & 1 & 2 & 0 \\ -1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 2 \\ 1 & 3 & 7 & 3 \end{bmatrix} \xrightarrow{R_2 + R_1 \text{ and } R_4 - R_1} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 5 & 3 \end{bmatrix}$$

$$\xrightarrow{R_3 - R_2 \text{ and } R_4 - 2R_2} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \xrightarrow{R_4 + R_3} \begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = R$$

The matrix is now in REF form, so a basis for the row space is just the first three rows, either of the reduced matrix. Note that we did *not* need to reduce to RREF to find a basis. So the vectors $\{(1,1,2,0),(0,1,3,1),(00,-1,1)\}$ form a basis for the row space. So the row space has dimension 3.

To answer part (2), we will reduce the matrix one step further, though we will see later that this is not necessary. We will first multiply row 3 by -1, and start there

$$\begin{bmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 - 3R_3 \text{ and } R_1 - 2R_3} \begin{bmatrix} 1 & 1 & 0 & 2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_1 - R_2} \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Note that row operations, do **NOT** preserve the column space. However, we know that row operations do preserve any linear independence between the columns. This is precisely idea we take advantage of when we solve linear systems. So if columns were linearly dependent before row reduction, they must stay that way after row reduction. Looking at the columns in the RREF matrix we obtained, we see that columns 1 and 2 are linearly independent, and columns 3 and 4 can be written as linear combinations of columns 1 and 2. Therefore, a basis for the column space is the first two columns in the *original* matrix A. Again, the column space of the columns in the RREF matrix is different from the columns in the original matrix! However, since linear dependence is preserved, we can safely assume that the columns in which we have leading entries after row reduction must have been linearly independent to start with. So a basis for the column space of A is

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \\ 7 \end{bmatrix} \right\}.$$

So the column space has dimension 3.

Finally to answer part (3), this is precisely what we have done before. We solve the homogeneous system. Considering the RREF form then, we see that

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

So $x_4 = s$ are our free variables, and we have that the solution set is

$$\left\{ \begin{bmatrix} 2s \\ -4s \\ s \\ s \end{bmatrix} : s \in \mathbb{R} \right\} = \left\{ s \begin{bmatrix} 2 \\ -4 \\ 1 \\ 1 \end{bmatrix} : s \in \mathbb{R} \right\}.$$

So a basis for the null space, is a basis for the solution set, which is given precisely by the vectors corresponding to each free variable above. So a basis for null(A) is

$$\left\{ \begin{bmatrix} 2 \\ -4 \\ 1 \\ 1 \end{bmatrix} \right\}$$

So the null space has dimension 1.

10 Lecture - 10/10/2022

10.1 Rank

To summarize, how to find a basis for each of the spaces, we perform the following steps. Let $A \in \mathbb{R}^{m\times n}$.

1. Reduce A to REF, call the reduced matrix R. The non-zero rows in the REF form of A form a basis for the row space of A.

2. The columns of R which contain a leading term correspond to leading variables. Then, the columns in A which correspond to the leading variables form a basis for the row space.

3. To find a basis for null(A), we solve the homogeneous system $[A|0]$. We then parameterize our solution using our free variables. For each variable, we obtain a corresponding solution vector. Taking one such vector for each free variable we obtain a basis for null(A).

Theorem 10.1.1. *The row space and column space of a matrix A have the same dimension.*

Proof. The dimension of the row space of a matrix corresponds to the number of non-zero rows remaining after reducing the matrix to REF or RREF form. This is equal to the number of leading entries, or leading 1's in the matrix. But as we have seen, row operations preserve linear dependence of columns. How?

Recall that performing row operations does not change the solution set to our system. In particular, the homogeneous system $[A|0]$ has the same solution set as $[R|0]$ where R is the RREF form matrix of A. But homogeneous systems capture linear dependence of the columns, so since solution sets are preserved, the so i linear dependence, and thus, linear independence.

S number of vectors in the basis for the column space is also equal to the number of leading 1's. Thus the dimensions of the both the row space and columns space are the same. □

Definition. The <u>rank</u> of a matrix A is the dimension of it's column space (or row space, equivalently), denoted rank(A). The dimension of the null space of A is called the nullity of A and is denoted nullity(A).

Theorem 10.1.2 (The Rank Theorem)**.** *Let $A \in \mathbb{R}^{m\times n}$, then*

$$rank(A) + nullity(A) = n$$

Proof. Take A and reduce it to RREF. A has n columns. Denote the number of leading ones by r, it equals the dimension of the column space. The remaining columns correspond to free variables. Each free variable contributes one basis vector to the null space, as we have seen, so the dimension of the null space is $n - r$. Thus

$$\text{rank}(A) + \text{nullity}(A) = r + (n - r) = n.$$

□

Example: Find the rank and nullity of $A = \begin{bmatrix} 1 & 2 \\ -2 & -3 \\ 2 & -1 \end{bmatrix}$.

Solution. First we reduce the matrix to REF. We have

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

So A has 2 leading ones, therefore the rank of A is 2. By the rank-nullity theorem, nullity of A is 2 - 2 = 0.

We now come to the next version of the fundamental theorem of invertible matrices.

Theorem 10.1.3 (The fundamental theorem of invertible matrices: part 2)**.** *Let $A \in \mathbb{R}^{n\times n}$ be a square matrix. The following statements are equivalent:*

1. *A is invertible.*
2. *$Ax = b$ has a unique solution for each $b \in \mathbb{R}^n$.*
3. *$Ax = 0$ has only the trivial solution, $x = 0$.*
4. *The RREF form of A is I_n, the identity matrix.*
5. *A is the product of elementary matrices.*
6. *$\text{rank}(A) = n$*
7. *$\text{nullity}(A) = 0$.*
8. *The column vectors and row vectors each are linearly independent.*
9. *The column vectors and row vectors each form a basis for $\mathbb{R}^n$.*
10. *The column vectors and row vectors each span $\mathbb{R}^n$.*

10.2 Orthogonality

We now move ahead to section 5.1 in the book.

Definition. The norm of a vector $v = (v_1, v_2 \ldots, v_n) \in \mathbb{R}^n$ is defined as

$$||v|| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + \cdots v_n^2}.$$

In particular, $||v|| > 0$ if $v \neq 0$.

Definition. Let $u, v \in \mathbb{R}^n$. We say that u is orthogonal to v if $u \cdot v = 0$.

In terms of geometry, this means that the vectors are perpendicular. Consider for example the vectors

$$\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\-1\end{bmatrix} \qquad \begin{bmatrix}1\\2\end{bmatrix} \cdot \begin{bmatrix}2\\-1\end{bmatrix} = 2(1) + (-1)(2) = 0.$$

If we draw these vectors in $\mathbb{R}^2$, we see that they are in fact perpendicular.

Definition. A set of vectors $\{v_1, \ldots, v_k\}$ in $\mathbb{R}^n$ is called an orthogonal set of $v_i \cdot v_j = 0$ for all pairs $i \neq jj$. That is, vectors v_i and v_j are orthogonal for any $i \neq j$.

Example: The standard basis vectors e_1, e_2, e_3 form an orthogonal set.

Theorem 10.2.1. *If $\{v_1, \ldots, v_k\}$ is an orthogonal set of non-zero vectors in $\mathbb{R}^n$, then these vectors are linearly independent.*

Proof. Suppose that

$$c_1 v_1 + \cdots c_k v_k = 0.$$

We wish to show that $c_i = 0$ for all i. Consider the dot product

$$v_i \cdot (c_1 v_1 + \cdots c_k v_k) = v_i \cdot 0 = 0.$$

Equivalently, we can distribute the v_i into the equation, and recalling that $v_i \cdot v_j = 0$ when $i \neq j$, we have

$$0 = c_1(v_i \cdot v_1) + \cdots + c_k(v_1 \cdot v_k) = c_i(v_i \cdot v_i)$$

Now $v_i \cdot v_i = ||v_i||^2 > 0$ since v_i is a non-zero vector. Thus $c_i = 0$. Since v_i was arbitrary, this implies that $c_i = 0$ for all $1 \leq i \leq k$. So the vectors $v_1, \ldots, v_k$ are linearly independent. □

Definition. An orthogonal basis for a subspace S is a orthogonal set that forms a basis for S.

Example: Find a set of vectors that form an orthogonal basis for the subspace $S \subset \mathbb{R}^3$ defined by $[x, y, z] \in S$ if $x + y + z = 0$.

Solution. If $x + y + z = 0$, then we can write $x = -y - z$, to see that vectors in S are of the form

$$\begin{bmatrix}-y-z\\y\\z\end{bmatrix} = y\begin{bmatrix}-1\\1\\0\end{bmatrix} + z\begin{bmatrix}-1\\0\\1\end{bmatrix}.$$

So the two vectors above form a basis for the space, but we can see that they are not orthogonal since $[-1,1,0]\cdot[-1,0,1] = 1 \neq 0$. Let's create an orthogonal basis as so. First, we choose any vector in S to start wit, let's take one of the basis vectors from the basis associated with y and z above. I will use the vector $[-1,1,0]$.

Now, we wish to find a vector in S that is also orthogonal to the vector $[-1,1,1,0]$. We know two is enough because any basis for S should have two elements. We consider these two statements now to obtain a system of linear equations that will allow us to find the vector we need.

$$[x,y,z] \in S \implies x+y+z=0$$

$$[x,y,z]\cdot[-1,1,0] = 0 \implies -x+y=0.$$

So we need to solve

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

Solving this system gives us the set $\{s[-1,-1,2] : s \in \mathbb{R}\}$. So in particular,

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \right\}$$

is an orthogonal basis for the subspace S.

11 Lecture - 10/12/2022

11.1 Orthogonality Continued

Today's lecture will continue section 5.1.

Theorem 11.1.1. *Let $\{v_1,\ldots,v_k\}$ be an orthogonal basis for a subspace S of $\mathbb{R}^n$ and let w be an vector in S. Then the unique scalars $c_1,\ldots,c_k$ such that*

$$w = c_1v_1 + \cdots + c_kv_k$$

are given by

$$c_i = \frac{w\cdot v_i}{v_i\cdot v_i} = \frac{w\cdot v_i}{||v_i||^2}.$$

Proof. We note that since $\{v_1,\ldots,v_k\}$ is a basis for S, then there is in fact a unique representation of each $w \in S$ as a linear combination of $v_1,\ldots,v_k$. Now, let

$$w = c_1v_1 + \cdots + c_kv_k$$

be the unique representation of w. Then observe that

$$v_i\cdot w = v_i(c_1v_1 + \cdots + .c_kv_k) = c_1(v_i\cdot v_1) + \cdots + c_k(v_i\cdot v_k) = c_i(v_i\cdot v_i) \implies c_i = \frac{v_i\cdot w}{v_i\cdot v_i}.$$

□

The importance of the theorem is that it can SIGNIFICANTLY simplify our work to solve linear systems. If we know that we are working with an orthogonal set of vectors, computing the solution can be reduced to performing a dot product. We can completely avoid having to perform any sort of Gaussian Elimination.

Example: The following is an orthonormal basis for $\mathbb{R}^3$,

$$\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}.$$

Find the coefficients c_1, c_2, c_3 so that

$$c_1 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 7 \\ -2 \end{bmatrix}$$

Definition. A set of vectors in $\mathbb{R}^n$ is an orthonormal set if it is orthogonal, and each vector has norm 1. An orthonormal basis for a subspace $S \subset \mathbb{R}^n$ is a basis for S that is an orthonormal set.

In particular, if $S \subset \mathbb{R}^n$ is a subspace and $\{v_1, \ldots, v_k\} \subset S$ is an orthonormal basis for S. Then we have that the unique representation of any $w \in S$ as a linear combination of the orthonormal basis vectors is

$$w = (w \cdot v_1)v_1 + (w \cdot v_2)v_2 + \cdots + (w \cdot v_k)v_k.$$

Definition. Let $A \in \mathbb{R}^{n \times m}$. The transpose of A is the matrix $A^T \in \mathbb{R}^{m \times n}$ such that the ij^{th} entry in A is the ji^{th} entry in A^T.

Examples: Let $A = \begin{bmatrix} 1 & 2 \\ -2 & -3 \\ 2 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$. Then

$$A^T = \begin{bmatrix} 1 & -2 & 2 \\ 2 & -3 & -1 \end{bmatrix} \qquad B^T = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$$

Lemma 11.1.1. *Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times \ell}$, then $(AB)^T = B^T A^T$.*

When working with orthonormal bases, we get more than just an easy way of representing vectors as linear combinations quickly. We also obtain an easy way of finding inverses of matrices.

Theorem 11.1.2. *Let $Q \in \mathbb{R}^{m \times n}$ so that the columns of Q form an orthonormal set of vectors. Then $Q^T Q = I_n$.*

Proof. Note that since the columns of Q are orthogonal, then they are linearly independent, and so $n \leq m$. Let $Q = [v_1 | \cdots | v_n]$, where v_i is the column vector of column i in Q. Then observe that Q^T, now has as rows, the columns of Q. Therefore, when we perform $Q^T Q$, the entry ij in $Q^T Q$ is given by $v_j \cdot v_i = 0$ if $i \neq 0$ and $v_j \cdot v_i = 1$ if $i = j$. Therefore, $Q^T Q = I_n$. □

Definition. Let $Q \in \mathbb{R}^{n\times n}$ be a square matrix, with an orthonormal set of column vectors. Then Q is called an orthogonal matrix.

Example: Consider $A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$. Show that A is an orthogonal matrix. What is A^{-1}?

Solution. We just need to see that the columns of Q form an orthonormal set of vectors in $\mathbb{R}^2$. We have

$$\begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix} = 1 \qquad \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix} = 1 \qquad \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix} \cdot \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix} = 0$$

So Q is an orthogonal matrix.

Theorem 11.1.3. *Let Q be an orthogonal matrix. Then $Q^{-1} = Q^T$.*

12 Lecture - 10/17/2022

12.1 Determinants

We now make our way back to Chapter 4, and we begin with 4.1, determinants. Many of you have heard about them before, and maybe even used or computed them before. But what do they represent? Determinants can be positive or negative. Forgetting about this for the moment? What does the absolute value of the determinant represent? As it turns out, the determinant of an $n \times n$ matrix tells you about how much the space $\mathbb{R}^n$ is stretched. In particular, the absolute value of the determinant of will tell you how by how much the matrix changes the n-dimensional volume. Let's consider two 2×2 examples that will demonstrate how exactly we can measure this.

Video on Determinants

The key property of the determinant is that it tells us how much multiplication by A stretches or collapses a space.

12.2 Determinants of small matrices

Many of you may recognize the 2×2 determinant. We will defined it here, and then show how to generalize finding the determinant of an $n \times n$ matrix.

Definition. Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. The determinant of A is denoted $\det(A)$ and is given by

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Examples: Let $A = \begin{bmatrix} 3 & 3 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$. Let's check that the intuitive approach we took to determinants coincides numerically with the theory we are now introducing.

$$\det\left(\begin{bmatrix} 3 & 3 \\ 0 & 2 \end{bmatrix}\right) = 3 \cdot 2 - 3 \cdot 0 = 6, \qquad \det\left(\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}\right) = 1 \cdot 2 - 2 \cdot 1 = 0.$$

The determinant of a 3×3 matrix $A = (a_{ij})$, is given in terms of 2×2 determinants of submatrices of A in the following way:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

This formula is called the cofactor expansion of A across the first row. Note that in particular, these submatrices come from A in the following method

$$\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Example: Let's consider the 3×3 we did at the beginning of class. Let $A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, then

$$\det(A) = \begin{vmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} = 1 \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} - 2 \begin{vmatrix} 1 & 1 \\ 0 & 0 \end{vmatrix} - 1 \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} = 0.$$

As it turns out, we can actually perform cofactor expansion across *any row or column* of A to find the determinant, but we need to be careful. Note that in the cofactor expansion across the first row, we have coefficients of the 2×2 matrices as $a_{11}, -a_{12}, a_{13}$. Note that it is $-a_{12}$, not just a_{12}. So, the way we know whether we should put a $+$ or a $-$ when performing cofactor expansion across any row or column, is as follows.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \longrightarrow \begin{vmatrix} a_{11}^{+} & a_{12}^{-} & a_{13}^{+} \\ a_{21}^{-} & a_{22}^{+} & a_{23}^{-} \\ a_{31}^{+} & a_{32}^{-} & a_{33}^{+} \end{vmatrix}.$$

This extends to $n \times n$ determinants. The top left entry is consider a $+$, and then going right to left, and up to down, we alternate $+$ and $-$.

Example: Find $\begin{vmatrix} 1 & 0 & 2 \\ 3 & -1 & 2 \\ 0 & 0 & 5 \end{vmatrix}$ by performing a cofactor expansion a cross column 2.

Solution. We begin as follows:

$$\begin{vmatrix} 1^{+} & 0^{-} & 2^{+} \\ 3^{-} & -1^{+} & 2^{-} \\ 0^{+} & 0^{-} & 5^{+} \end{vmatrix} = -0 \begin{vmatrix} -1 & 2 \\ 0 & 5 \end{vmatrix} + (-1) \begin{vmatrix} 1 & 2 \\ 0 & 5 \end{vmatrix} - 0 \begin{vmatrix} 3 & -1 \\ 0 & 0 \end{vmatrix} = 0(-5) + (-1)(5) - 0(0) = -5.$$

The thing to note there is that it is useful to perform cofactor expansion across any row or column with the most number of zeroes.

12.3 Properties of Determinants

Let us introduce some notation regarding the cofactor expansion of a matrix A. Let $A \in \mathbb{R}^{n\times n}$, then denote A_{ij}, to be the submatrix of A that excludes row i and column j. For example, if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \implies A_{11} = \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}, \quad A_{13} = \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}$$

So that now we can describe $\det(A) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$. In general, the determinant of any $A \in \mathbb{R}^{n\times n}$ computed by cofactor expansion across row i (or column j) is given by

$$\det(A) = \underbrace{\sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij})}_{\text{across row } i} = \underbrace{\sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij})}_{\text{across column } j}$$

1. The first and most important property, is that if $A \in \mathbb{R}^{n\times n}$, then $\det(A) = 0$ if and only if A *is not* invertible. In other words, $\det(A) \neq 0$ if and only if A *is* invertible. *Why?*

   Well, let's consider the meaning of $\det(A)$ as we introduced it. If $\det(A) = 0$, then it sends the n-dimensional cube with one corner at the origin and volume 1, to some shape that has strictly less than n dimensions, so that it's n-dimensional volume is 0. (Think of 2×2 matrices mapping the unit square to just a line, a line is flat, and so has area = 0, while the square has area 1.) In other words, the standard basis vectors $\{e_1, \ldots, e_n\}$ which are linearly independent, are mapped to a smaller dimensional space, and so $\{Ae_1, \ldots, Ae_n\}$ must be linearly dependent, meaning A is not invertible.

2. If A is a triangular matrix (either upper or lower triangular), then the determinant of A is just the product of the diagonal entries. Let's quickly verify this with a 4×4 example. In this example, we will perform cofactor expansion across column 1.

$$\begin{vmatrix} 1^{+} & 2^{-} & -3^{+} & 0^{-} \\ 0^{-} & 5^{+} & -1^{-} & 2^{+} \\ 0^{+} & 0^{-} & -1^{+} & 3^{-} \\ 0^{-} & 0^{+} & 0^{-} & 2^{+} \end{vmatrix} = 1\begin{vmatrix} 5 & -1 & 2 \\ 0 & -1 & 3 \\ 0 & 0 & 2 \end{vmatrix} - 0\begin{vmatrix} 0 & -1 & 2 \\ 0 & -1 & 3 \\ 0 & 0 & 2 \end{vmatrix}$$

$$+0\begin{vmatrix} 0 & 5 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 2 \end{vmatrix} - 0\begin{vmatrix} 0 & 5 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{vmatrix} = 1\begin{vmatrix} 5 & -1 & 2 \\ 0 & -1 & 3 \\ 0 & 0 & 2 \end{vmatrix}$$

   So we see that all that is left over is the 3×3 determinant of the bottom right corner, if we continue this, we will obtain that

$$\begin{vmatrix} 1 & 2 & -3 & 0 \\ 0 & 5 & -1 & 2 \\ 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 2 \end{vmatrix} = 1\begin{vmatrix} 5 & -1 & 2 \\ 0 & -1 & 3 \\ 0 & 0 & 2 \end{vmatrix} = 1(5)\begin{vmatrix} -1 & 3 \\ 0 & 2 \end{vmatrix} = 1(5)(-1)2 = -10.$$

13 Lecture - 10/19/2022

13.1 Properties of Determinants Continued..

We will continue here with the properties of determinants. We recall the first two properties above. Let $A \in \mathbb{R}^{n\times n}$, be a square matrix.

1. The determinant, $\det(A)$ gives us information about how multiplying $\mathbb{R}^n$ by A stretches or collapses $\mathbb{R}^n$. In particular:

2. If the rows or columns of A are linearly dependent, then $\det(A) = 0$.

3. If A is triangular (upper or lower), then $\det(A) = a_{11}a_{22}\cdots a_{nn}$, the product of the diagonal entries of A.

4. If A, B, C are matrices that are identical except with a differing row i, such that row i of matrix C is the sum of row i in A and row i in B, then $\det(C) = \det(A) + \det(B)$.

   Example: Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix}$, and $C = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}$. Then

   $$\det(A) + \det(B) = 3 + (-7) = -4 = \det(C).$$

   Proof. Note, we said that all rows in A, B, C are identical except row i, for which we have that

   $$c_{ij} = a_{ij} + b_{ij}$$

   for any $1 \le j \le n$. Then, using our general cofactor expansion formula, let us expand across row i, to find $\det(C)$.

   $$\begin{aligned}\det(C) &= \sum_{j=1}^{n}(-1)^{i+j}(a_{ij} + b_{ij})\det(A_{ij}) \\ &= \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}) + \sum_{j=1}^{n}(-1)^{i+j}b_{ij}\det(A_{ij}) = \det(A) + \det(B)\end{aligned}$$

   □

   The same exact result applies to columns.

5. If C is obtained from A by adding a multiple of row k to row i, then $\det(C) = \det(A)$.

   Example: A whole class of examples can be considered from triangular matrices. Consider $A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{bmatrix}$ and B obtained from A by adding row 2 to row 1. Then

   $$\det(B) = \begin{vmatrix} 2 & 2 & 4 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{vmatrix} = 6 = \begin{vmatrix} 2 & -1 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{vmatrix} = \det(A).$$

   In particular, such an operation cannot change the diagonal entries, so if the new matrix again is triangular, then it is easy to verify the claim.

Proof. Let A be the matrix as given and B be identical to A, except, replace row i with c times row k of A. So in particular, B has rows that are linearly dependent as row i is a multiple of row k. Therefore $\det(B) = 0$. By applying property 4 above, we have

$$\begin{aligned}\det(C) &= \sum_{j=1}^{n}(-1)^{i+j}(a_{ij} + ca_{kj})\det(A_{ij}) \\ &= \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}) + \sum_{j=1}^{n}(-1)^{i+j}ca_{kj}\det(A_{ij}) = \det(A) + \underbrace{\det(B)}_{0} = \det(A)\end{aligned}$$

□

The same holds if we do this for columns.

6. If C is obtained from A by multiplying row i by $k \in \mathbb{R}$, then $\det(C) = k\det(A)$.

   Example: Let $A = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$. Then $\det(C) = 6 = 2\det(A)$

   Proof. Perform cofactor expansion across row i to see that

$$\det(A) = \sum_{j=1}^{n}(-1)^{i+j}ka_{ij}\det(A_{ij}) = k\left(\sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij})\right) = k\det(A)$$

   □

   The same holds if we do this for columns.

7. If B is obtained from A by swapping rows i and j, then $\det(B) = -\det(A)$.

   Proof. We have shown how the elementary row operations defined by multiplying a row by a scalar, and adding a multiple of one row to another, affects the determinant. With this knowledge, we will show that in fact, swappping two rows, can be done via a sequence of the other two elementary row operations.

   Let the rows of A be $[R_1, \dots, R_i, \dots, R_j, \dots R_n]$. Then we may summarize the determinant properties above as

$$\begin{aligned}\det([R_1, \dots, cR_i, \dots, R_n]) &= c\det([R_1, \dots, R_i, \dots, R_n]) \\ \det([R_1, \dots, R_i + cR_k, \dots, R_n]) &= \det([R_1, \dots, R_i, \dots, R_n]) \ \ (i \neq k)\end{aligned}$$

   We will now perform a series of row operations to get from A to B without performing the swap operation.

$$\begin{aligned}\det(A) &= \det([R_1, \dots, R_i, \dots, R_j, \dots, R_n]) \\ &= \det([R_1, \dots, R_i, \dots, R_j - R_i, \dots, R_n]) \\ &= \det([R_1, \dots, R_i + (R_j - R_i), \dots, R_j - R_i, \dots, R_n]) \\ &= \det([R_1, \dots, R_j, \dots, R_j - R_i, \dots, R_n]) \\ &= \det([R_1, \dots, R_i, \dots, -R_i, \dots, R_n]) \\ &= -\det([R_1, \dots, R_j, \dots, R_i, \dots, R_n]) = -\det(B)\end{aligned}$$

□

The key point of discussing these properties is that they inform us that we can find the determinant of matrices more quickly if we first perform some row operations on them. This is the way a computer would approach this problem.

13.2 Determinants and Elementary Matrices

We immediately obtain the following corollary as a result of proving the properties above.

Corollary. *Let $E \in \mathbb{R}^{n\times n}$ be an elementary matrix. Then*

1. $\det(I_n) = 1$.
2. *If E is obtained from I by adding row i to row j, then* $\det(E) = 1$.
3. *If E is obtained from I by swapping two rows, then* $\det(E) = -1$.
4. *If E is obtained from I by multiplying row i by $c \in \mathbb{R}$,* $\det(E) = c$.

Recalling that multiplying a matrix A by an elementary matrix, is equivalent to performing the corresponding operation on the matrix. Therefore, the corollary and properties above together yield the following lemma.

Lemma 13.2.1. *Let $A \in \mathbb{R}^{n\times n}$ and E be an elementary matrix. Then*

$$\det(EA) = \det(E)\det(A).$$

Theorem 13.2.1. *A square matrix $A \in \mathbb{R}^{n\times n}$ is invertible if and only if* $\det(A) \neq 0$.

Proof. Assume that A is invertible, then A can be written as a product of elementary matrices, so $A = E_k \dots E_1$. Then $\det(A) = \det(E_k \cdots E_1) = \det(E_k)\cdots\det(E_1) \neq 0$, since $\det(E_i) \neq 0$ for any elementary matrix.

Now suppose that $\det(A) \neq 0$. Then reduce A to row echelon form R. We know that this can be done via a series of row operations. Thus $R = E_\ell \cdots E_1 A$. Thus

$$\det(R) = \det(E_k)\cdots\det(E_1)\det(A).$$

Now, $\det(E_i) \neq 0$, so if $\det(A) \neq 0$, then $\det(R) \neq 0$. But since R is upper triangular, it implies that the diagonal entries of R cannot be 0. Therefore, the RREF form of R is I, meaning it is invertible, which means that A is invertible, since RREF form of R is the RREF form of A. □

14 Lecture - 10/24/22

14.1 Important Properties of Determinants

Theorem 14.1.1. *A square matrix* $A \in \mathbb{R}^{n\times n}$ *is invertible if and only if* $\det(A) \neq 0$.

Theorem 14.1.2. *Let* $A, B \in \mathbb{R}^{n\times n}$*, then*

$$\det(AB) = \det(A)\det(B).$$

Proof. We will consider two cases. Either $\det(A) = 0$, in which case the REF form of A has a row of zeroes, or $\det(A) \neq 0$.

If $\det(A) = 0$, then A the REF form of A, call it the matrix R, has a bottom row of zeroes, this implies that RB will have a bottom row of zeroes. So then $A = E_k \cdots E_1 R$, thus

$$\det(AB) = \det(E_k \cdots E_1 RB) = \det(E_k)\cdots\det(E_1)\det(RB) = 0 = \det(A)\det(B).$$

If $\det(A) \neq 0$, then A is invertible, so $A = E_k \cdots E_1$, is a product of elementary matrices. Thus

$$\det(AB) = \det(E_k \cdots E_1 B) = \det(E_k)\cdots\det(E_1)\det(B) = \det(E_k \cdots E_1)\det(B) = \det(A)\det(B).$$

□

Theorem 14.1.3. *Let* $A \in \mathbb{R}^{n\times n}$*, then* $\det(kA) = k^n \det(A)$

Corollary. *Let* $A \in \mathbb{R}^{n\times n}$ *be invertible, then*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Proof. We have

$$1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1}) \implies \det(A^{-1}) = \frac{1}{\det(A)}.$$

□

Theorem 14.1.4. *Let* $A \in \mathbb{R}^{n\times n}$*, then* $\det(A^T) = \det(A)$.

14.2 Eigenvalues and Eigenvectors

We now come to one of the most important topics in Linear Algebra, eigenvalues and eigenvectors. In essence, the question of the eigenvalues and eigenvectors of a matrix ultimately seeks to classify matrices by certain properties, like how much does multiplying by A stretch the space, and in which direction? Particularly, our goal will be to understand the following: Given $A \in \mathbb{R}^{n\times n}$, for which values λ does there exists a non-zero vector v such that

$$Av = \lambda v.$$

For example, the identity matrix satisfies

$$Iv = 1 \cdot v$$

for any $v \in \mathbb{R}^n$. Notice that if for some square matrix there exists a vector v and $\lambda \in \mathbb{R}$ such that

$$Av = \lambda V \implies Av - \lambda v = 0 \implies (A - \lambda I)v = 0.$$

Now, if v is not the zero vector, then there are many ways to interpret this statement. For example, it would mean all the following things:

- $v \in \text{null}(A - \lambda I)$.
- $A - \lambda I$ is *not* invertible
- $\det(A - \lambda I) = 0$.
- etc....

Definition. Let $A \in \mathbb{R}^{n \times n}$. An eigenvalue of A is a real number $\lambda \in \mathbb{R}$ such that there exists a non-zero vector v such that $Av = \lambda v$. If λ is an eigenvalue of A, then any non-zero vector v with $Av = \lambda v$ is called an eigenvector of A.

Practically speaking, the way we find eigenvalues is through the following theorem.

Theorem 14.2.1. *Let $A \in \mathbb{R}^{n \times n}$, then λ is an eigenvalue of A if and only if* $\det(A - \lambda I) = 0$

Proof. Suppose that $\det(A - \lambda I) = 0$. We have shown that this is true, if and only if $A - \lambda I$ is not invertible, which happens if and only if there exists a non-zero solution vector v to $(A - \lambda I)v = 0$, implying that λ is an eigenvalue. □

Definition. The polynomial in λ, $\det(A - \lambda I)$ is called the characteristic polynomial of A.

Example: Let $A = \begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix}$. Find the characteristic polynomial and the eigenvalues of A.

Solution. We have that

$$\det(A - \lambda I) = \begin{vmatrix} 2 - \lambda & -3 \\ 1 & -2 - \lambda \end{vmatrix} = (2 - \lambda)(-2 - \lambda) - (1)(-3) = \lambda^2 - 4 + 3 = \lambda^2 - 1.$$

Now

$$\lambda^2 - 1 = 0 \implies (\lambda - 1)(\lambda + 1) = 0 \implies \lambda = \pm 1$$

are the eigenvalues of A.

Now, how does one find the eigenvectors of A? And more specifically, if λ is an eigenvalue of A, then how do we find the eigenvectors corresponding to the eigenvalue λ. Well, we must substitute in the precise value of λ into the matrix $A - \lambda I$, and then solve $[A - \lambda I|0]$.

Definition. Let A be a square matrix with λ as an eigenvalue. Then $\text{null}(A - \lambda I)$ is called the eigenspace of A corresponding to λ, denoted E_λ.

Example: Let $A = \begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix}$. Find the eigenvlaues and corresponding eigenspaces of A.

Solution. We already found the $\lambda = \pm 1$ are the eigenvalues of A So to find the corresponding eigenspaces, we now solve for $\text{null}(A - I)$ and $\text{null}(A + I)$. Note

$$A - I = \begin{bmatrix} 1 & -3 \\ 1 & -3 \end{bmatrix} \qquad A + I = \begin{bmatrix} 3 & -3 \\ 1 & -1 \end{bmatrix}.$$

Now,

$$\left[\begin{array}{cc|c} 1 & -3 & 0 \\ 1 & -3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & -3 & 0 \\ 0 & 0 & 0 \end{array}\right] \implies x_1 = 3x_2.$$

So the null space is one dimensional, and namely, we have

$$\text{null}(A - I) - \left\{ s \begin{bmatrix} 3 \\ 1 \end{bmatrix} : s \in \mathbb{R} \right\}.$$

Any vector of the form $s \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is an eigenvector of A with corresponding eigenvalue 1. Similarly, we find that

$$\text{null}(A + I) = \left\{ s \begin{bmatrix} 1 \\ 1 \end{bmatrix} : s \in \mathbb{R} \right\}.$$

So any vector of the for $s \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of A with corresponding eigenvalue $-1.=$

So why is this useful or interesting? Well, let's see how this can speed up matrix multiplication against a vector. In the 2×2 case, this may not see impressive, but just think about having matrices with thousands of rows and columns, and what this would mean for computational speed and efficiency.

Let's take the same 2×2 we had in the example above. Note that any two eigenvectors of A, each coming from the different eigenspaces will form a basis for $\mathbb{R}^2$. In particular, we can see that any vector in $\mathbb{R}^2$ has the following break down,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \left(\frac{x - y}{2} \right) \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \left(\frac{3y - x}{2} \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Now, this implies that

$$A \begin{bmatrix} x \\ y \end{bmatrix} = A \left(\frac{x - y}{2} \right) \begin{bmatrix} 3 \\ 1 \end{bmatrix} + A \left(\frac{3y - x}{2} \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \left(\frac{x - y}{2} \right) \left(A \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right) + \left(\frac{3y - x}{2} \right) \left(A \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$$

$$= \left(\frac{x - y}{2} \right) \left(1 \cdot \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right) + \left(\frac{3y - x}{2} \right) \left(-1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = \left(\frac{x - y}{2} \right) \begin{bmatrix} 3 \\ 1 \end{bmatrix} - \left(\frac{3y - x}{2} \right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2x - 3y \\ x - 2y \end{bmatrix}$$

15 Lecture 10/26/2022

15.1 Eigenvalues and Eigenvectors Continued

Let's do an example to refresh our memories about eigenvalues.

Example: Find the eigenvalues of $A = \begin{bmatrix} 1 & -5 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 2 \end{bmatrix}$.

Solution. We have

$$\det(A - \lambda I) - \begin{vmatrix} 1-\lambda & -5 & 2 \\ 0 & 2-\lambda & -1 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda)(2-\lambda)(1-\lambda).$$

Therefore, solutions to $\det(A - I) = 0$ are $\lambda = 1, 2$. So the eigenvalues of A are 1 and 2 where 1 appears two time, and 2 appears one time.

We notice that what made computing the iegenvlaues easy in this case is that A was upper triangular. In particular, if A is upper triangular, then $A - \lambda I$ is triangular, meaning that the eigenvalues of A are just the entries down the diagonal of A. This proves the following theorem.

Theorem 15.1.1. *Let A be a triangular matrix. Then the eigenvalues of A are precisely just the diagonal entries of A.*

Let us now go on to find the eigenspaces of the same matrix A associated with each eigenvalue. Again, A has eigenvalues 1 and 2. So now we compute $\text{null}(A - I)$ and $\text{null}(A - 2I)$. Working with $A - I$, we reduce it to the following using row operations.

$$A-I = \begin{bmatrix} 0 & -5 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 0 & -5 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \implies \text{null}(A-I) = \left\{ x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} : x \in \mathbb{R} \right\}$$

So we have determined that the eigenspace of A corresponding to eigenvalue 1 is as given and is precisely 1 dimensional. Now, working with $A - 2I$, we have

$$A - 2I = \begin{bmatrix} -1 & -5 & 2 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 5 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \implies \text{null}(A - 2I) = \left\{ x \begin{bmatrix} -5 \\ 1 \\ 0 \end{bmatrix} : x \in \mathbb{R} \right\}.$$

So we have determined that the eigenspace of A corresponding to eigenvalue 2 is as given and is precisely 1 dimensional.

Definition. Let $A \in \mathbb{R}^{n \times n}$ and λ_1 be an eigenvalue of A. Then the multiplicity of the root $(\lambda - \lambda_1)$ in the equation $\det(A - \lambda I) = 0$ is called the algebraic multiplicity of λ_1. The dimension of the null space of $A - \lambda_1 I$ is called the geometric multiplicity of λ_1.

Note that the above example tells us that the geometric multiplicity and algebraic multiplicity of an eigenvalue of some matrix A need note be the same. In particular, in the example above, the eigenvalue 2 has algebraic multiplicity two but geometric multiplicity one.

Theorem 15.1.2. *Let $A \in \mathbb{R}^{n\times n}$, then A is invertible if and only if 0 is not an eigenvalue of A.*

Proof. If 0 is an eigenvalue, then this implies that there exists a non-zero vector solving $Av = 0v = 0$, which implies that the null space of A is non-empty which can only happen if and only if A is not invertible. Thus, A is invertible if and only if 0 is *not* an eigenvalue of A. □

Theorem 15.1.3. *Let A be a square matrix with eigenvalue λ corresponding eigenvector v. Then:*

1. *For any positive integer n, λ^n is an eigenvalue of A^n with corresponding eigenvector v.*
2. *If A is invertible, then $1/\lambda$ is an eigenvalue of A^{-1} with corresponding eigenvector v.*

16 Lecture - 10/31/2022

Theorem 16.0.1. *Let $A \in \mathbb{R}^{n\times n}$ have distinct eigenvalues $\lambda_1, \ldots, \lambda_m$, with corresponding eigenvectors $v_1, \ldots, v_m$. Then $v_1, \ldots, v_m$ are linearly independent.*

Proof. Suppose that $v_1, \ldots, v_m$ are not linearly independent, then, order the eigenpairs such that $\lambda_{i+1} > \lambda_i$. Now suppose that we find the the first k vectors are linearly indpendent, but adding the $k+1$t vector creates a linear dependence, so that

$$v_{k+1} = a_1v_1 + \cdots + a_kv_k \implies Av_{k+1} = A(a_1v_1 + \cdots a_kv_k).$$

Now, we have that

$$Av_{k+1} = \lambda_{k+1}v_{k+1} = a_1\lambda v_1 + \cdots a_k\lambda_k v_k.$$

On the other hand, since $v_{k+1} = a_1v_1 + \cdots a_kv_k$, we have that

$$\lambda_{k+1}a_1v_1 + \cdots + \lambda_{k+1}a_kv_k = \lambda_1a_1v_1 + \cdots \lambda_ka_kv_k \implies (\lambda_{k+1} - \lambda_1)a_1v_1 + \cdots + (\lambda_{k+1} - \lambda_k)a_kv_k = 0.$$

Since $v_1, \cdots, v_k$ were linearly independent to begin with, then we need each $a_i(\lambda_{k+1} - \lambda_i) = 0$. Now, since $v_{k+1} \neq 0$, then at least one of the $a_i \neq 0$, implying that for this index i, $\lambda_{k+1} - \lambda_i = 0$. But this is impossible since the eigenvalues are all distinct! So these vectors must be linearly independent, so we are done. □

Theorem 16.0.2. *Let $A \in \mathbb{R}^{n\times n}$ and $\lambda_1, \ldots, \lambda_m$ be eigenvalues of A with corresponding eigenvectors $v_1, \ldots, v_m$. If x can be written as a linear combination of these eigenvectors, $x = a_1v_1 + \cdots a_mv_m$, then*

$$A^kx = \lambda_1^ka_1v_1 + \cdots \lambda_m^ka_mv_m.$$

Example: Let $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Let us find the eigenvalues and their corresponding eigenspaces.

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = -\lambda \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} - \begin{vmatrix} 1 & 1 \\ 1 & -\lambda \end{vmatrix} + \begin{vmatrix} 1 & -\lambda \\ 1 & 1 \end{vmatrix}$$

$$= -\lambda(\lambda^2 - 1) - (-\lambda - 1) + (1 + \lambda) = -(\lambda - 2)(\lambda + 1)^2$$

So the eigenvalues are $\lambda = -1, 2$. We see that the algebraic multiplicity of $\lambda = -1$ is two, and algebraic multiplicity of $\lambda = 2$ is one. Let's move on to find the eigenspaces.

Beginning with $\lambda = -1$, we have

$$[A + I|0] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \implies x_1 + x_2 + x_3 = 0.$$

Here x_2 and x_3 are free variables, so we parameterize and get $x_2 = s$, $x_3 = t$ and $x_1 = -s-t$, and so

$$\mathrm{null}(A + I) - \left\{ s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} : s, t \in \mathbb{R} \right\}.$$

This space is 2 dimensional, and so the geometric multiplicity of $\lambda = -1$ is two.

Now for $\lambda = 2$, we have

$$[A - 2I|0] = \left[\begin{array}{ccc|c} -2 & 1 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & 1 & -2 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 1 & -2 & 1 & 0 \\ -2 & 1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & 3 & -3 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 1 & -2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \implies x_1 = x_3 \text{ and } x_2 = x_3.$$

Here x_3 can be the free variable, and so we set $x_3 = s$. Therefore

$$\mathrm{null}(A - 2I) = \left\{ s \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} : s \in \mathbb{R} \right\}.$$

So this space is one dimensional, meaning the geometric multiplicity of $\lambda = 2$, is one.

A basis for $\mathrm{null}(A + I)$, and $\mathrm{null}(A - 2I)$ respectively, is

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

We have proven a theorem that says that the eigenvectors corresponding to different eigenvalues are linearly independent. So that means, altogether, the three vectors are linearly independent, and so they span $\mathbb{R}^3$, since $\mathbb{R}^3$ is three dimensional. So we can write any vector a linear combination of these three vectors. In particular, we can solve

$$\left[\begin{array}{ccc|c} -1 & -1 & 1 & x \\ 1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right]$$

to find that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \left(\frac{2y-x-z}{3}\right)\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + \left(\frac{2z-x-y}{3}\right)\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \left(\frac{x+y+z}{3}\right)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Now using what we know about the eigenvalues and eigenvectors of A, we get that

$$\begin{aligned} A\begin{bmatrix} x \\ y \\ z \end{bmatrix} =& A\left(\frac{2y-x-z}{3}\right)\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + A\left(\frac{2z-x-y}{3}\right)\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + A\left(\frac{x+y+z}{3}\right)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ =&(-1)\left(\frac{2y-x-z}{3}\right)\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + (-1)\left(\frac{2z-x-y}{3}\right)\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + (2)\left(\frac{x+y+z}{3}\right)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} y+z \\ x+z \\ x+y \end{bmatrix} \end{aligned}$$

While this could have been easy to compute by hand, using the other theorem, we can also compute any power of A applied to any vector, very quickly, for example

$$\begin{aligned} A^7\begin{bmatrix} x \\ y \\ z \end{bmatrix} =& A^7\left(\frac{2y-x-z}{3}\right)\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + A^7\left(\frac{2z-x-y}{3}\right)\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + A^7\left(\frac{x+y+z}{3}\right)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ =&(-1)^7\left(\frac{2y-x-z}{3}\right)\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + (-1)^7\left(\frac{2z-x-y}{3}\right)\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + (2)^7\left(\frac{x+y+z}{3}\right)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ =&(-1)\left(\frac{2y-x-z}{3}\right)\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + (-1)\left(\frac{2z-x-y}{3}\right)\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + 128\left(\frac{x+y+z}{3}\right)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\ =&\begin{bmatrix} 42x+43y+43z \\ 43x+42y+43z \\ 43x+43y+42z \end{bmatrix} \end{aligned}$$

17 Lecture 11/02/2022

17.1 Similar Matrices

Suppose we have two matrices $A, B \in \mathbb{R}^{n\times n}$ such that they have the same eigenvalues (with the same multiplicity), and furthermore, there is a basis $\{v_1, \dots, v_n\}$ of eigenvectors of A for $\mathbb{R}^n$ and a basis $\{u_1, \dots, u_n\}$ of eigenvectors of B. Suppose also that for each i, we have

$$Av_i = \lambda_i v_i \qquad Bu_i = \lambda_i u_i,$$

that is the eigenvalue λ_i of A corresponding to v_i, is the same as the eigenvalue of B corresponding to u_i. Finally, let's also consider the matrix

$$C = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

So C is diagonal. Notice that there is something "similar" about how A, B, C multiply against vectors in $\mathbb{R}^n$. For example, $A(a_1v_1 + \cdots a_nv_n) = \lambda_1a_1v_1 + \cdots + \lambda_na_nv_n$, but it is also true that $B(a_1u_1 + \cdots a_nu_n) = \lambda_1a_1u_1 + \cdots + \lambda_na_nu_n$. Even though the vectors $a_1v_1 + \cdots a_nv_n$ is not necessarily equal to $a_1u_1 + \cdots + a_nu_n$, the way A and B multiply against the respective vectors is essentially the same. But notice, this is also true for $C(a_1e_1 + \cdots + a_ne_n) = \lambda_1a_1e_1 + \cdots + \lambda_na_ne_n$, where $\{e_1, \dots, e_n\}$ is the standard ordered basis, and can be seen to be a set of eigenvectors for C. This similar action is what we wish to understand and make more concrete.

Definition. Two matrices $A, B \in \mathbb{R}^{n\times n}$ are said to be **similar** if there exists an invertible matrix P such that $B = P^{-1}AP$. If A is similar to B, then we write $A \sim B$.

Example: Recall the matrix from last class $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. We determined that A had eigenvalues $\lambda = -1, 2$.

Let

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

Then observe that

$$P^{-1}AP = \frac{1}{3}\begin{bmatrix} -1 & 2 & -1 \\ -1 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & -2 & 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -1 & 2 & -1 \\ -1 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 2 \\ -1 & -1 & 2 \\ 0 & 2 & 2 \end{bmatrix}$$

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

So in this case, the matrix A is *similar* to the diagonal matrix with the eigenvalues of A down the diagonal. This is no coincidence, and there are particular conditions under which we know that this will necessarily happen.

Remark: Observe that an equivalent statement about A being similar to A is if $AP = PB$.

Theorem 17.1.1. *Let $A, B, C \in \mathbb{R}^{n\times n}$, then*

1. $A \sim A$.
2. $A \sim B$ *if and only if* $B \sim A$.
3. $A \sim B$ *and* $B \sim C$ *implies that* $A \sim C$.

Theorem 17.1.2. *Let $A, B \in \mathbb{R}^{n\times n}$ with $A \sim B$. Then*

1. $\det(A) = \det(B)$.
2. *A is invertible if and only if B is invertible.*
3. *A and B have the same rank.*
4. *A and B have the same characteristic polynomial.*
5. *A and B have the same eigenvalues.*
6. $A^m \sim B^m$, *for all* $m \geq 0$.
7. *If A is invertible, then $A^m \sim B^m$ for all integer m.*

Proof. We will only prove some of these. Suppose that $A \sim B$, then there exists an invertible matrix P such that $B = P^{-1}AP$.

1. Then

$$\det(B) = \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P) = \frac{1}{\det(P)}\det(A)\det(P) = \det(A).$$

2. From this, part 2 falls out immediately since we know that a matrix is invertible if and only if it's determinant is not 0.

3. If A and B are invertible, then we are done. If they are not invertible, then rank(A) = rank(B) if and only if nullity(A) = nullity(B). The nullity is the just the dimension of the null space. Note that if $S = \{v : Av = 0\}$ is the null space of A, then the null space of B is $P^{-1}S = \{P^{-1}v : Av = 0\}$. Since P^{-1} is invertible, then the dimension of the space does not change.

4. The characteristic polynomial is $\det(A - \lambda I)$ and $\det(B - \lambda I)$. Note that

$$\det(B - \lambda I) = \det(P^{-1}AP - \lambda(P^{-1}P)) = \det(P^{-1})\det(A - \lambda I)\det(P) = \det(A - \lambda I).$$

□

So we have that an easy necessary condition is to check if two matrices are similar using the theorem above.

Example: Let $A \begin{bmatrix} 6 & 3 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 3 \\ 0 & 4 \end{bmatrix}$. Show that A and B are not similar.

Solution. Note that $\det(A) = 12 = \det(B)$, so we must proceed. Note that the eigenvalues of A are 2 and 6, while the eigenvlaues of B are 3 and 4. Therefore the matrices are not similar.

17.2 Diagonalization

When matrices are similar to diagonal matrices, then it can really simplify any work done with such matrices. In this subsection, we will try and understand when it is that a matrix is similar to a diagonal matrix. Note that if A is similar to a diagonal matrix D, then the diagonal entries necessary of D are necessarily the eigenvalues of A, each eigenvalue appearing with the same number as its algebraic multipicity.

Definition. A matrix $A \in \mathbb{R}^{n\times n}$ is diagonalizable if there is a diagonal matrix D such that A is similar to D. That is, there exists an invertible matrix P such that $P^{-1}AP = D$.

Theorem 17.2.1. *Let $A \in \mathbb{R}^{n\times n}$. Then A is diagonalizable if and only if A has n linearly independent eigenvectors.*

More precisely, there exists an invertible matrix P and diagonal matrix D such that $P^{-1}AP = D$ if and only if the columns of P are linearly independent eigenvectors of A and the diagonal entries of D are the eigenvalues of A corresponding to the eigenvectors. That is, D_{ii} is the eigenvalue which corresponds to the column i in P.

Proof. Let A be an $n \times n$ matrix with n linearly independent eigenvectors, $p_1, \ldots, p_n$. Let $P = [p_1|\ldots|p_n]$ be matrix with the linearly independent eigenvectors of A as columns. Let D be a diagonal matrix such that $D_{ii} = \lambda_i$ is the eigenvalue of A such that $Ap_i = \lambda_i p_i$.

We mentioned that an equivalent way of stating similarity, is that $AP = PD$. Let us prove that this is the case. Observe that

$$AP = A[p_1|\ldots|p_n] = [Ap_1|\ldots|Ap_n] = [\lambda_1 p_1|\ldots|\lambda_n p_n] = [p_1|\ldots|p_n] \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} = PD$$

This proves that if A has n linearly independent eigenvectors, then A is diagonalizable. Now we need to show that if A is diagonalizable, then in fact A has n linearly independent eigenvectors and the diagonal entries of D are the eigenvalues which correspond to the columns of P which are eigenvectors of A.

Suppose $P^{-1}AP = D \implies AP = PD$. Let's consider what happens one columns at a time. Take the column 1 in P, call it p_1. Then $Ap_1 = p_1\lambda_1$, but this just means that p_1 is an eigenvectors of A with eigenvalue λ_1. This holds true for any column p_i, so that

$$Ap_i = \lambda_i p_i.$$

Since P is invertible, then the columns are linearly independent. □