Elementary Linear Algebra [MA351] Notes

DeVon Herr

Summer 2019

These are my course notes for Elementary Linear Algebra, MA35100, taught by Beichuan Deng, based on chapters 1-5 of *Linear Algebra: Ideas and Applications* by Richard C. Penny.

Contents

1 Day 1: June 10

Syllabus stuff!

1.1 What is Linear Algebra?

There's two words here, linear and algebra. So, what is an algebra? We've learned this pretty well:

- "Numbers"
- Real numbers
- Algorithms $(+, -, \times, /)$

We will learn these algorithms again, but no longer with simply numbers. Here, we will learn these under the context of vectors and matrices. This will be the "linear" part. In linear algebra, we will learn both scalar and vector multiplication.

Another important term is a set, such as $\mathbb{R}$ and $\mathbb{C}$. In linear algebra, we will learn sets. We will learn what the set $M(m, n)$ is; these are for matrices. $\mathbb{R}^n$ is for vectors.

We will also learn what a space is. Here, we will hand-wave and define it as a space with some additional structure.

As an example, we can think of our backpack as a set:

$$\text{Backpack} = \{\text{computer}, \text{binder}, \text{pencil}, \dots\}$$

or a folder as one too, but this folder being specifically for MA351:

$$\text{Folder}_{\text{MA351}} = \{\dots\}$$

where all elements in this set has the additional property of being related to MA351.

1.1.1 What is *linear?*

These are things that are "straight" such as lines and planes. We know the standard equations of these as $y = ax + b$ and $ax + by + cz = d$. Both are linear functions, the first in x and y, the latter in x, y and z. In this course, we will learn about the *algebraic* definition of what linear really means. What is a linear function? What is a linear transformation? Etc.

1.2 The vector space of $m \times n$ matrices

Our elements are matrices.

We begin with the classic introduction of a table. Here, let this be a grading report of some kind:

	Homework	Exam 1	Exam 2	Final
Student A	95	85	78	92
Student B	87	88	89	90
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

If we look solely at the numbers, we can consider this a matrix. An important notion is the size of a matrix. Here, we represent the size by writing as the number of rows by the number of columns. If there were thirty eight students, then this matrix is size 38×4. Importantly when converting from a table to a matrix, we leave out the labels as rows or columns. We are solely counting the *entries*.

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

This is a matrix, and is 2×2.

$$\begin{bmatrix} 1 & 3 & 8 \\ 4 & & \end{bmatrix}$$

This is not, as there are numbers missing. Compare this to

$$\begin{bmatrix} 1 & 3 & 8 \\ 4 & 0 & 0 \end{bmatrix}$$

which is. This is a 2×3 matrix.

Now, we use $M(m, n)$ to mean the set of all matrices with size $m \times n$. So the first table would belong to the set $M(38, 4)$. The first matrix would belong in $M(2, 2)$ and so on.

Some special matrices exist such as column vectors like

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

which belongs in $M(3, 1)$. Importantly, this is also in $\mathbb{R}^n$, which is the set of all n-dimensional column vectors. This vector is also in $\mathbb{R}^3$. For posterity, entries must be real numbers.

We may also deal with something like $\mathbb{C}^3$, which would be three entires which are complex numbers. For instance,

$$\begin{bmatrix} 1+i \\ 1-i \\ 2+3i \end{bmatrix}$$

is not in $\mathbb{R}^n$ but is in $\mathbb{C}^n$. Most of the time, we will working in $\mathbb{R}^n$.

Row vectors also exist too. For example,

$$\begin{bmatrix} 1 & 2 & 4 & -5 \end{bmatrix}$$

is in $M(1, 4)$ as a four dimensional row vector.

1.3 Algorithms and Operations on Matrices

1.3.1 Addition

Addition of two matrices is well defined, but the two matrices must be of the same size. Two matrices of size 5×3 can be added, but we can't add, say, a 5×4 to a 3×4.

For notation, If we represent a matrix with a capital letter, (such as A) then we refer to its entries *in general* with the subscripts ij (such as A_{ij}) but sometimes with its lowercase letter (such as a_{ij}). With this notation, i refers to the row and j refers to the column. Thus, we can write for a given $m \times n$ matrix,

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

Then, the sum of two matrices $A + B$ is defined as $(a_{ij} + b_{ij})$. Thus,

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}.$$

For example

$$\begin{bmatrix} 1 & 3 & 4 \\ 2 & -1 & 5 \end{bmatrix} + \begin{bmatrix} 3 & 3 & 0 \\ 2 & 2 & -8 \end{bmatrix} = \begin{bmatrix} 4 & 6 & 4 \\ 4 & 1 & -3 \end{bmatrix}$$

note imporantly, if we sum two $x \times y$ matrices, the result is still $x \times y$. The sizes do not change. As another example,

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

is not defined.

1.3.2 Scalar multiplication

This is defined by $c \cdot A$, where c is some real number and A is a matrix. Here, $c \cdot A$ is defined as A with all of it's entires multiplied by c, which gives us $c \cdot A = (c \cdot a_{ij})$. As an example,

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

Importantly, $0 \cdot A$ does not give 0, but *the zero matrix.* So

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

Scalar multiplication by negative numbers behave as expected, so we have both that

$$A - B = A + (-1)B$$

and

$$A - B = (a_{ij} - b_{ij})$$

An example is not needed here.

1.4 On Systems of Linear Equations

An extremely common reason for why people study linear algebra is to solve systems of linear equations. Something such as $ax = b$ is fairly trivial to solve. Something like $ax + by = c$ has infinitely many solutions, but not if we add in another equation like so:

$$\begin{aligned} ax + by &= c \\ cx + dy &= e \end{aligned}$$

These are also pretty easy to solve with algebraic methods (substitution and elimination etc).

So solving a 2×2 is pretty easy, but what about solving a 10×10? If we try substitution, we are almost dead in the water. This is *horrendously* difficult to do. Elimination is a bit better, but still complicated. Which variable do we solve for when?

The key is to use a matrix to solve them.

2 Day 2: June 11

2.1 What is a Linear Equation and What is a System of them?

What is a linear equation? We say that an equation is linear if and only it is in the form

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

Then naturally, a system of linear equations is just a a set of them:

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

Here, we have n variables and m equations. These are considered linear because all variables x_n exist solely in first order expressions, such as $2x$ or $3y$. Something like x^2 or $\sin(x)$ is not linear.

Naturally, we care about finding solutions to the system, that is, finding all possible n-tuples

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

that satisfy all equations.

We've seen before that something like $5x = 10$ is easy to solve – just divide both sides by 5. For a system of two linear equations (of two variables), we can use substitution and elimination.

2.2 Discerning the Number of Solutions

But before solving, it may be worth asking if the system is even solvable in the first place. And how many, if at all?

An intersection of two lines in the xy-plane is the classical representation of the solution of a 2×2 system, where the intsersection represents a point which lies on both lines, so it naturally satisfies both equations. In this case, the solution is unique. But is this the only possibility? No. The two lines (technically, the two equations) could completely overlap and thus be the same. Thus all points on the line are solutions to the system, and there are infinitely many solutions. There are also systems where the two lines don't intersect (in $\mathbb{R}^2$, they would be parallel), and there are no solutions to the system.

So,

There are three different situations:

- There exists a unique solution.
- There exists infinitely many solutions.
- There do not exist any solutions.

This generalizes to arbitrary equations/dimensions.

For instance, in three dimensions, the system

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= b_1 \\ a_{21}x + a_{22}y + a_{23}z &= b_2 \\ a_{31}x + a_{32}y + a_{33}z &= b_3 \end{aligned}$$

represents three planes. If there exists a unique solution, it is the single point where all three planes intersect.

Things get a bit more complicated with overlap, but not by much. If two planes completely overlap but intersect with a third, the intersection line represents infinitely many solutions. If all three planes overlap, there are infinitely many solutions that lie on the plane.

Of course, if not all of the three planes intersect, then there exist no solutions to the system.

In higher dimensions, the equations are now hyperplanes (planes in multiple dimensions).

2.3 Solving a Linear System

Note that when we did elimination, for instance of the following system:

$$\begin{aligned} 2x - 3y &= -1 \\ 2x + 3y &= 5 \end{aligned}$$

when we add the two equations, we have that $4x = 4$. But we didn't have to add the two equations directly. We could just add the numbers. You didn't physically write out $(2x - 3y) + (2x + 3y) = -1 + 5$, did you? No. We just summed the coefficients.

Instead, we can reprsent this as a matrix:

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

This is a 2×3 matrix. For posterity, the number of rows is the number of equations, and the number of columns is the number of variables $+1$. When we applied elimination by adding the two equations, our matrix became

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

The bottom row reprsents the equation $4x + 0y = 4$ or simply just $4x = 4$. Then to solve for y, we want something of the form $0x + Cy = b$, so we get rid of the 2 in the first row by subtracting one half of the second row:

$$\begin{bmatrix} 0 & -3 & -3 \\ 4 & 0 & 4 \end{bmatrix}$$

Then to actually find x, we can divide the second row by 4. The same principle applies to the first row to give us

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

which gives us that $y = 1$ and $x = 1$.

This process is known as Gaussian elimination. This is the process of:

$$\underbrace{\begin{bmatrix} 2 & -3 & -1 \\ 2 & 3 & 5 \end{bmatrix}}_{\text{Augmented matrix}} \Rightarrow \underbrace{\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}}_{\text{Solution matrix}}$$

2.4 Special Matrices

We can also describe the coefficient matrix which is obviously just the coefficients. In this example, it would be

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

which is a 2×2 matrix. The number of columns is the number of variables and the number of rows is the number of equations.

As an example, the following system

$$\begin{aligned} x + 2y + z &= 1 \\ 3x + y + 4z &= 0 \\ 2x + 2y + 3z &= 2 \end{aligned}$$

would have an augmented matrix size of 3×4:

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

We can also go backwards. The following augmented matrix

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

has three equations and two variables and is reprsented by the system

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

As a pathological case, the following augmented matrix

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

yields the equation $0 = 2$, an obvious contradiction. Thus, the system the matrix represents has no solution.

A matrix such as

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

isn't as nice as a solution matrix, but is still quite nice. Once we have that $z = -1$, we can find that $y + 1(-1) = 2$ which gives us $y = 3$, and then get that $x + 2(3) + 3(-1) = 1$ to find that $x = -2$.

Identity Matrices

A special type of matrix is an identity matrix. For instance, the 3×3 identity matrix is

$$I_{\mathbb{R}^3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which is the "coefficient part" of a solution matrix of a system of three equations in three variables. Higher degree versions exist, which are square matrices with 1s on the down-right diagonal and all other entries zero.

A "worse" version of an identity matrix is

Diagonal Matrices

These are matrices where all non-diagonal entries must be zero. For posterity, diagonal entries *can* be zero. For example,

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

is a diagonal matrix.

An "even *worse*" version of a diagonal matrix is an upper-triangular matrix:

Upper-triangular matrices

These are matrices where everything *except* the diagonal entry and *above* must be zero. For example,

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

is an upper-triangular matrix.

We can also define the lower-triangular matrix similarly. I won't bother.

2.5 Gaussian Elimination Revisited

The goal of Gaussian elimination is to transform the coefficient matrix part of an augmented matrix into an identity matrix. This is done by first transforming it into an upper-triangular matrix, before reducing it to an identity matrix.

We can't just rewrite the numbers in a matrix. That would be redefining the system, which is not what we want to do (find solutions). Here are things we *can* do:

- Swap the order of two rows.
- Add a multiple of a row onto another row.
- Multiply a row by a non-zero constant.

These are known as elementary row operations. It's pretty clear that the first and third don't change the system.

3 Day 3: 2019 June 12

We will talk more about Gaussian elimination.

3.1 On Gaussian Elimination

Review of Yesterday

The goal of Gaussian elimination is to transform the coefficient matrix part of an augmented matrix into an identity matrix. This is done by first transforming it into an upper-triangular matrix, before reducing it to an identity matrix.
We can't just rewrite the numbers in a matrix. That would be redefining the system, which is not what we want to do (find solutions). Here are things we *can* do:

- Swap the order of two rows.
- Add a multiple of a row onto another row.
- Multiply a row by a non-zero constant.

These are known as elementary row operations. It's pretty clear that the first and third don't change the system.

As an example, take the following system of linear equations

$$\begin{aligned} x + 2y + z &= 1 \\ 3x + y + 4z &= 0 \\ 2x + 2y + 3z &= 2 \end{aligned}$$

which is represented by the following augmented matrix:

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

Before we get solve the actual system, we can discuss some theorems about matrix representations.

Definition of Row Equivalence of Matrices

We say two matrices A and B are row equivalent if and only if there is a sequence of elementary row operations that transforms A into B.

So the two matrices

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

are row equivalent.

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

are also row equivalent, as the second row is the first row minus three times the second. This isn't trivial to notice, so I wouldn't bother (I didn't see it until it was pointed out, either.)

The notion of row equivalence is useful as it yields the following theorem:

"Theorem 1.3"

Suppose that A and B are row equivalent. Then the system having A as an augmented matrix has the same solution set as the system having B as its augmented matrix.

This should be pretty straightforward: if two matrices are row equivalent, then they should have the same solutions under a system of linear equations. Again, this is because all elementary row operations maintain the solution set of a matrix, and two matrices are row equivalent if and only they can be obtained by means of elementary row operations. This theorem technically validates the method of Gaussian elimination.

3.2 Echelon Form

The best situation is that Gaussian elimination correctly turns the coefficient matrix into an upper-triangular matrix, but sometimes this is not achievable. If our entire system is simply

$$x - 2y + 3z = 4$$

we get the augmented matrix of

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

which cannot be modified to get an upper-triangular matrix.

So we can generalize the definition of an upper-triangular matrix to echelon form:

"Definition 1.11:" Echelon Form

A matrix A is in echelon form if and only if it satisfies the following conditions:

- The first nonzero entry in any nonzero row occurs to the right of the first such entry in the row directly above it.
- All zero rows are grouped together at the bottom

For example, let's look at

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

This is not echelon form, as the second row's first nonzero entry is not to the right of the one directly above it, as the first entry of the first row is in the same column as the first entry of the second row.

As another example,

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

satisfies the condition. For instance, in the second row, the first nonzero entry, 2, is to the right (column wise) of the first nonzero entry in the one above it, 1. The same principle applies to the third row.

Here's another:

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

this is also in echelon form. 2 is to the right of 1, and 3 is to the right of 2.

Here's yet another:

$$\begin{bmatrix} 0 & 3 & 1 & 7 \\ 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This fails to be in echelon form, as 1 is to the *left* of 3 column wise.

Here's yet *another:*

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

this is not in echelon form, as there is a zero row that is not at the bottom.

It should be pretty straightforward to note any upper triangular matrix is in echelon form.

With echelon form, we can generalize Gaussian elimination to now look like this:

$$A \longrightarrow \text{Echelon form} \longrightarrow I$$

"Theorem 1.4"

Every matrix A is row equivalent to a matrix in echelon form.

With this, we can always guarantee the first step of Gaussian elimination. Let's go back to the following example:

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

We want to get this to echelon form. The first row is fine, but the second and third rows both have a nonzero entry as their first entry that is not to the right of the first entry of the first row column wise. The method is to subtract a multiple of the first equation such that the second rows first entry becomes zero. We can do this by having $A_2 = A_2 - 3A_1$ and $A_3 = A_3 - 2A_1$. This gives us

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

Now our second row is okay, but the third row is not. So we do $A_3 = 5A_3 - 2A2$:

$$\begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & -5 & 1 & -3 \\ 0 & 0 & 3 & 6 \end{bmatrix}$$

Now the matrix is in echelon form.

From echelon form, we can see if a solution set exists or not. Recall that if you see a row such as

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

and a is not zero, then the solution set does not exist. If we do not see a row like this, there must exist a solution.

3.3 Row Reduced Echelon Form

We now define a special kind of echelon form:

"Definition 1.13:" Row Reduced Echelon Form

A matrix A is in row reduced echelon form if and only if

- A is in echelon form.
- All of the pivot entries are 1.
- All the entries above the pivots are 0.

A pivot entry is the first nonzero entry in a nonzero form.

Note that the identity matrix is in row reduced echelon form.
The following matrix is in row reduced echelon form:

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

We can then generalize Gaussian Elimination to be:

$$A \longrightarrow \text{Echelon form} \longrightarrow \text{Row Reduced Echelon Form}$$

For example, take the echelon form matrix

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

Then, to make all the pivot entries 1, we simply divide all rows by their pivot:

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

Then to ensure all entries above the pivots are 0, we subtract multiples of rows below them, from the lowest row to the top. So we do $A_1 = A_1 - A_3$ and $A_2 = A_2 - 1/5A_3$. This gives

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

Then we do $A_1 = A_1 - 2A_1$ which gives

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

which gives the solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix}$$

Sometimes we will want to arrange rows before applying some row reduction. For example,

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

doesn't work. The first row has a zero, which can't be made into a 1. So we can just move it somewhere else. Gaussian elimination takes care of the rest.

Generally speaking, you want the first rows pivot to be a 1 so you don't run into fractions.

As another example, consider the matrix

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

After a series of row reduction, you get

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

The bottom most row has that $0 = 1$, so we can conclude that there is no solution.

4 Day 4: 2019 June 13

Previously, we discussed the cases where Gaussian elimination yields a single unique solution or no solutions. Today we will discuss another case.

4.1 Infinitely Many Solutions

Consider the equation $x - 2y = 4$. If we graph the line as $y = 1/2x - 2$, all points on the line are solutions. Though subtle, we the equation does not give us a solution. In reality, it's more like

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} s \\ \frac{1}{2}s - 2 \end{bmatrix}$$

Here, we let s be an arbitrary constant. We can actually factor into constant and variable terms like so:

$$\begin{bmatrix} x \\ y \end{bmatrix} = s \begin{bmatrix} 1 \\ \frac{1}{2} \end{bmatrix} + \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

There is an important geometric interpretation to this. The vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is, well, a position vector on $\mathbb{R}^2$, so if we are able to multiply it by a constant s, we can arbitrarily stretch or compress and even flip directions. The set of all possiblities of $s\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ forms a line. The constant vector $\begin{bmatrix} 0 \\ -2 \end{bmatrix}$ is just some constant vector. Since we have that

$$\begin{bmatrix} x \\ y \end{bmatrix} = s\begin{bmatrix} 1 \\ \frac{1}{2} \end{bmatrix} + \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

the geometric interepretation of our vector sum is actually the line created by the variable vector shifted by the constant vector. For posterity, all values s yields a different solution.

What about in 3 variables? Consider if we have the equation $2x - y + 3z = 1$. Then we have that $x = (1 + y - 3z)/2$. Importantly, both y and z are independent and free. We can then interepret the solution set, using s and t as constants as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{2} + \frac{s}{2} - \frac{3}{2}t \\ s \\ t \end{bmatrix}$$

which we can then factor by constants and variables to yield

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -\frac{3}{2} \\ 0 \\ 1 \end{bmatrix}$$

Importantly, the first vector is constant, but the latter two are variable. We know from multivariate calculus that the expression $2x - y + 3z = 1$ forms a plane in $\mathbb{R}^3$, but we can derive this from linear algebra. Note that we have two free independent vectors, which actually is a subspace of $\mathbb{R}^2$ which spans a plane. Then, the full vector expression is simply the plane spanned by the two vectors shifted by the constant vector.

Let's consider the system

$$\begin{aligned} x + y + 2z + w &= 1 \\ 3x - 4y + z + w &= 2 \\ 4x - 3y + 3z + 2w &= -1 \\ 5x - 2y + 5z + 32 &= 0 \end{aligned}$$

which simplfies to

$$\begin{bmatrix} 1 & 1 & 2 & 1 & 1 \\ 0 & -7 & -5 & -2 & -1 \\ 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

which fails to have a solution, since the third row states that $0 = -4$.

But what if we can vary the constants? That is, given

$$\begin{aligned} x + y + 2z + w &= a \\ 3x - 4y + z + w &= b \\ 4x - 3y + 3z + 2w &= c \\ 5x - 2y + 5z + 3 &= d \end{aligned}$$

can we ensure *a* solution exists? The problem, again, is caused by the third and fourth rows which state that $0 = 4$. So if we have something like

$$\begin{bmatrix} 1 & 1 & 2 & 1 & a \\ 3 & -4 & 1 & 1 & b \\ 4 & -3 & 3 & 2 & 0 \\ 5 & -2 & 5 & 3 & 0 \end{bmatrix}$$

we are good.

Let's try solving

$$\begin{bmatrix} 1 & 1 & 2 & 1 & a \\ 3 & -4 & 1 & 1 & b \\ 4 & -3 & 3 & 2 & c \\ 5 & -2 & 5 & 3 & d \end{bmatrix}$$

After row reduction, we have

$$\begin{bmatrix} 1 & 1 & 2 & 1 & a \\ 0 & -7 & -5 & -2 & b - 3a \\ 0 & 0 & 0 & 0 & c - b - a \\ 0 & 0 & 0 & 0 & d - b - 2a \end{bmatrix}$$

which has that $c - b - a = 0$ and $d - b - 2a = 0$. Note that in our original situation, this was not true, so there were no solutions. If we take these conditions and put it into an augmented matrix, we have that

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

The first column corresponds to a, the second b and so on. We use this to introduce a related concept to pivot entries:

Pivot Variables and Free Variables

A variable that has a pivot entry in its corresponding column is said to be a pivot variable but a free variable otherwise.

To further solve, we aim to solve for the pivot variables in terms of the free variables.

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} -s + t \\ 2s - t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

Note that this vector sum expresses all constant vectors that yields a solvable system.

Let's consider the system

$$\begin{bmatrix} 3 & 4 & 5 & 3 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 1 & 0 \\ 1 & 2 & 3 & 1 & 1 \\ 3 & 5 & 7 & 3 & 2 \end{bmatrix}$$

which is a bitch to simplify if you don't make the very first pivot entry 1 by switching rows. If you don't, you end up having to divide the first row by 3 which is nasty. Gaussian elimination gives us

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

Transforming this into equation form, we have that $x - z + w = -1$ and $y + 2z = 1$. Solving for x and y by letting z and w be constants, we have that $x = -z + w - 1$ and $y = 1 - 2z$. This gives us

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -s + t - 1 \\ 1 - 2s \\ s \\ t \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

It's worth noting that different expressions exist for solution sets. This is because for a given plane, there are infinitely many vectors. Thus, there are infinitely many combinations of two vectors which span the plane.

4.2 Vector Spaces

We define a vector space as a set of vectors with some additional properties. We'll call it V. This set is closed under addition and scalar multiplication.

In total, we have 10 properties. They are as follows:

Given three vectors $X, Y, Z \in V$

- $X + Y \in V$
- $X + Y = Y + X$
- $X + (Y + Z) = (X + Y) + Z$
- There exists some element (a zero vector) such that $X + \text{Zero vector} = X$
- For any X, there exists a not necessarily unique element $-X \in V$ such that $X + (-X) = 0$

These are also known as group axioms.

For any scalars k and l,

- $kX \in V$
- $k(lX) = klX$
- $k(X + Y) = kX + kY$
- $(k + l)X = kX + lX$
- $1X = X$ and $\vec{0}X = 0$.

5 Day 5: 2019 June 14

We discussed what a vector space is. To review,

Review

We define a vector space as a set of vectors with some additional properties. We'll call it V. This set is closed under addition and scalar multiplication. In total, we have 10 properties. They are as follows:

Given three vectors $X, Y, Z \in V$

- $X + Y \in V$
- $X + Y = Y + X$
- $X + (Y + Z) = (X + Y) + Z$
- There exists some element (a zero vector) such that $X + \text{Zero vector} = X$
- For any X, there exists a not necessarily unique element $-X \in V$ such that $X + (-X) = 0$

These are also known as group axioms.

For any scalars k and l,

- $kX \in V$
- $k(lX) = klX$
- $k(X + Y) = kX + kY$
- $(k + l)X = kX + lX$
- $1X = X$ and $\vec{0}X = 0$.

5.1 Examples of Vector Spaces

It's worth noting that $M(m, n)$ is a vector space. I don't know why this is being covered, but we can verify this for $M(2, 2)$.

- The group is indeed closed under addition. $\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} + \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & d_1 + d_2 \end{bmatrix}$
- Matrix addition is commutative
- Matrix addition is associative
- The additive identity exists, where $\vec{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

- An additive inverse exists for all elements, where $A + (-A) = \vec{0}$.
- Scalar multplication is well defined. $c\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ca & cb \\ cc & cd \end{bmatrix}$
- Scalar multiplication is associative.
- Scalar multiplication is distributive – commutative with respect to a scalar and the sum of two matrices.
- Scalar multiplication is distributive – commutative with respect to the sum of two scalars and a matrix.
- A multiplicative identity exists.

Some more:

- $\mathbb{R}^2$ is a vector space.
- Interestingly, the set of *only* $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is a vector space as well.
- The set of only $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ fails to be a vector space, as it is not closed under addition.
- The set $S = \{\alpha \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \alpha \in \mathbb{R}\}$ is a vector space. Addition is trivially commutative and associative as this space uses vector addition. The same applies for scalar multiplication being associative and distributive in both regards. All that's necessary is to prove that it is closed under addition and scalar multiplication.
- The set $S = \{\alpha \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \alpha \in \mathbb{R}^+\}$ is not a vector space as it lacks a zero vector, and thus, an additive inverse.

Here's a more interesting one:

$$P_n(a, b) = \{\text{All of the polynomials defined on } (a, b) \text{ whose degree} \leq N\}$$

So $P_2(a, b)$ is the set of all constant, linear and quadratic functions. This is indeed a vector space, we can add polynomials together and get another one, apply scalar multiplication to get anothe one, and so on.

5.2 Subspaces

A subset is pretty obvious. In fact, I can't even define it with words. The sets $S = \{\alpha \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \alpha \in \mathbb{R}\}$ and $S = \{\alpha \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \alpha \in \mathbb{R}^+\}$ are both subsets of $\mathbb{R}^2$ but only the former is a subspace.

A subspace is a bit more involved:

"Theorem (1.3)" Subspace

Firstly, a subspace must first be a subset. A subset W of a vector is a susbspace if and only if the following hold:

- If X and Y are elements of W, then $X + Y \in W$
- For any $X \in W$ and a scalar c, $cW \in X$.
- $0 \in X$

That is, a subset is a subspace if and only if it is a vector space itself.

Note that $\mathbb{R}^2$ is *not* a subset of $\mathbb{R}^3$. The number of entries has to be consistent, otherwise there's no reasonable way to map the first onto the second. Of course we can just say "add a z" component, but then there has to be a choice of *where* (first, second or third entry) and at that point it's insufficient. There is no canonic embedding.

5.3 Linear Combinations

"Definition 1.3" Linear Combinations

Let $S = \{A_1, \ldots, A_n\}$ be a set of elements in the vector space V. An element C is a linear ocmbiation of A_i if and only if there exist scalars b_i such that

$$C = b_1 A_1 + \ldots b_n A_n$$

For instance, if $S = \{A_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$, then $\begin{bmatrix} 2 \\ -3 \end{bmatrix}$ is a linear combination, as $C = 2A_1 + (-3A_2)$. If $C = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, C is not a linear combination.

Here's a harder example: Let $S = \{A_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, A_2 = [1, 1, 0], A_3 = \begin{bmatrix} 1 & 0 & 2 \end{bmatrix}\}$ and $C = \begin{bmatrix} -1 & 2 & 4 \end{bmatrix}$. To determine if C is a linear combination, we see if we can find constants a b and c such that

$$\begin{bmatrix} -1 & 2 & 4 \end{bmatrix} = a \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} + b\,[1, 1, 0] + c \begin{bmatrix} 1 & 0 & 2 \end{bmatrix}$$

which when multiplied out and solved for constants:

$$\begin{aligned} -1 &= a + b + c \\ 2 &= 0a + b + 0c \\ 4 &= 0a + 0b + 2c \end{aligned}$$

which we can then set up as an augmented matrix and solve

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

Since the solution exists, C is indeed a linear combination.

5.4 Linear Dependence of Sets

There exists an application of linear combinations:

Linear Dependence of Sets

Let S be a vector space. S is linearly dependent if at least one of A_i is a linear combination of the other elements of the other elements in S. S is said to be linearly independent otherwise.

We also define the set of solely the zero vector as linearly independent.

The set $S = \{\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}\}$ is linearly dependent. There's no way to write any of the vectors as a linear combination of the others. However, we have to check *every single* vector. If a single vector isn't a linear combination of the other vectors, it does not rule out the possibility of the another vector being a linear combination. Here, we can simply not that for each vector, it has a nonzero entry where all other vectors have said entry 0.

For instance, the set $S = \{\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 2 & -3 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}\}$ is linearly dependent as the third vector can be written as a linear combination of the other vectors *even though one of them cannot.*

For posterity, if the set S has solely two vectors, we can verify just by checking if the two vectors are a scale factor of eachother.

6 Day 6: 2019 June 17

We discussed linear combinations and linear dependence.

6.1 Span

"Definition 1.5:" Span

Let $S = \{A_1, A_2, \ldots, A_k\}$ be a set of elements in a vector space. Then span$S$ (verbally, the span of S)is that set of all elements of the form

$$B = c_1 A_1 + c_2 A_2 + \ldots + c_k A_k$$

where c_i are scalars.

So the span of a set of vectors is the set of all possible linear combinations of the vectors. We can also write that

$$\text{span}(s) = \{c_1 A_1 + \ldots + c_k A_k\} \quad c_1, \ldots c_k \in \mathbb{R}$$

For example, if $S = \begin{bmatrix}1\\0\end{bmatrix}$, what is span $\left(\begin{bmatrix}1\\0\end{bmatrix}\right)$? By definition, we have that the span is equal to $c\begin{bmatrix}1\\0\end{bmatrix}$, which is a straight line. More specifically, it is the x-axis.

Similarly, span $\left(\begin{bmatrix}1\\0\end{bmatrix}\right)$ is the y-axis.

More intrestingly, span $\left(\begin{bmatrix}0\\0\end{bmatrix}\right)$ is the origin. Algebraically, it is the set $\{\begin{bmatrix}0\\0\end{bmatrix}\}$. **Most of the time a span of a set of vectors contains infinitely many vectors**, but this is a special case.

span $\left(\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}2\\0\end{bmatrix}\right)$ is still just the x-axis. Note that even though we have another vector, it still doesn't affect the span since the two vectors are on the same line.

What if the two vectors aren't on the same line? span $\left(\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right)$ is $\mathbb{R}^2$. To see why, we apply the definition of span to get

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

to which we can pick arbitrary reals c_1 and c_2 to get all possible tuples $\begin{bmatrix}x\\y\end{bmatrix}$ of $\mathbb{R}^2$.

For a harder example, find span $\left(\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}\right)$. Interestingly enough, the span is also $\mathbb{R}^2$. Geometrically, we can construct a linear combination by drawing an arbitrary vector, and using some parallelograms.

6.1.1 Color

One potential application/analogy with linear dependence and linear combinations is that of mixing color. A set of only yellow and orange would be linear independent, since you can't get yellow from orange and vice versa. But something like a set of red, orange, yellow would not be lienarly dependent, since you can mix red and yellow to make orange. What if we add green? Still dependent, since there exists at least one vector that is linearly dependent (orange is still there). It doesn't matter if green is linearly independent. The set of green, red and yellow is linearly independent.

The span of a pallette of colors would then be all possible color combinationst hat can be yielded from mixing the colors.

6.2 Test for Independence

Recall that previously the set $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is linearly dependent. This was obvious to see, but what about something like $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} \right\}$? We first proceed by checking

$$\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = a \begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix} + b \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix}$$

which is not possibe, as the first entry is 1, whereas the first entry for the other two matrices are zero, so it is not possible. For the second vector, we check

$$\begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix} = a \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} + b \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix}$$

which has that $a = 0$ in order to match the first entry. For the third vector, we have that

$$\begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} = a \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix}$$

which gives that $b = 0$ in order to match the second entry. However, this yields a contradiction, as if a and b are zero, we end up with the zero vector. Thus, the set is linearly dependent.

In general, if we have an upper or lower triangular matrix, the vectors are linearly independent. We can take this further and say that if a single vector has a zero entry, but all vectors have nonzero entries, the matrix is linearly independent. Stair-case like structures (instead of upper/lower triangular matrices) are also linearly independent.

This was not fun to do. To make this faster, we make use of something known as the the test for independence. This is an equivalent though much faster method.

Note that when we check for linear combinations, we check if the following systems, well, hold:

$$\begin{aligned} A_1 &= a_1^1 A_2 + a_2^1 A_2 + \ldots + a_k^1 A_k \\ A_2 &= a_1^2 A_2 + a_2^2 A_3 + \ldots + a_k^2 A_k \\ \vdots &= \vdots \end{aligned}$$

We can rearrange to get

$$\begin{aligned} 0 &= -A_1 + a_1^1 A_2 + a_2^1 A_2 + \ldots + a_k^1 A_k \\ 0 &= -A_2 + a_1^2 A_2 + a_2^2 A_3 + \ldots + a_k^2 A_k \\ 0 &= \vdots \end{aligned}$$

Summing, then dividing by $n-1$ (don't worry about it, it's not important to understand this), we have that

$$0 = a_1 A_1 + a_2 A_2 + \ldots + a_k A_k$$

which has the trivial solution $a_i = 0$. But if we have other solutions, say we have

$$0 = 2A_2 - 3A_3$$

we would have that $A_2 = 3/2A_3$. We would then say that these two are linearly dependent. This applies for any equation/any group of vectors.

If we have a nonzero solution, the set is linearly dependent. If the only solution is the zero vector, the set is linearly independent. As a sidenote, systems where the constant vector is the zero vector are homogeneous systems.

As another example, let's reconsider the original problem which then asks us to find constants a b and c such that

$$a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This yields the augmented matrix

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

which, as the identity matrix, yields that $a, b, c = 0$, concluding that the set is indeed linearly independent.

What about the set $S = \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \right\}$? Translating into row reduced form, we have

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

which tells us that c is free, and a and b are linearly dependent. From this, we have that $a + c = 0$ and $b + c = 0$. Solving, we have that $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -s \\ -s \\ s \end{bmatrix}$ which yields the vector $s\begin{bmatrix}-1\\-1\\1\end{bmatrix}$. Letting $s=1$, we have that $\begin{bmatrix}a\\b\\c\end{bmatrix}=\begin{bmatrix}-1\\-1\\1\end{bmatrix}$ which gives us that $-A_1-A_2+A_3=0$.

Tomorrow, we will discuss dimension, an integer measurement that refers to how "large" a vector space is. This is how we can conclude $\mathbb{R}^3>\mathbb{R}^1$ in some sense.

7 Day 7: June 18

We will continue or discussions on vector spaces.

7.1 Dimension of a Vector Space

"Definition 2.4:" Dimension of a Vector Space

The dimension of a vector space V is the smallest number of elements necessary to span V. That is, V has dimension n if and only if there exists a set $\{A_1,A_2,\ldots,A_n\}$ whose span equals V, while no set of $n-1$ or fewer elements of V spans V.

This is usually a non-negative integer $\{0,1,2,\ldots\}$. We denote this by saying $\dim V$.

Geometrically $\dim\left(\mathbb{R}^2\right)=2$. This is because $\mathbb{R}^2$ is a plane and we only need two vectors. This is not possible with any fewer, as a single vector cannot span $\mathbb{R}^2$, as a single vector spans a line. Trivially, zero vectors doesn't work either.

The same logic works for $\mathbb{R}^3$, $\dim\left(\mathbb{R}^3\right)=3$.

"Definition 2.2:" Basis

A basis for a vector space V is a linearly independent set

$$\{A_1,A_2,\ldots A_n\}$$

of V that spans V.

For example, the standard basis of $\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}$ is the basis of $\mathbb{R}^2$. The same idea applies for $\mathbb{R}^3$ with its own basis.

7.2 Vector Spaces Revisited

Our old definition was abstract. Most vector spaces are infinite.

As an analogy, consider preperation to paint. We want to bring the fewest number of paints while still enabling us to paint whatever we want. For instance, if we take red, blue and yellow, we can get any color we want.

The same principle here applies for a vector space, we want the know the fewest number of elements such that it spans our given vector space. We can obviously take *more* elements then we need and still span a given vector space, so for a given basis of a vector space, we have that

$$\dim V \leq \text{The number of elements in the basis.}$$

"Proposition 2.1"

Let $\{A_1, A_2, \ldots, A_k\}$ be a set of elements in V such that A_k is a linear combination of $\{A_1, A_2, \ldots, A_{k-1}\}$. Then $\operatorname{span}(\{A_1, A_2, \ldots, A_k\}) = \operatorname{span}(\{A_1, A_2, \ldots, A_{k-1}\})$.

For example, if we use the standard basis for $\mathbb{R}^2$ then take some random element in $\mathbb{R}^2$ and add that to the basis, then this extra element doesn't do anything.

To be even more concrete, we can do

$$\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \left[2, \tfrac{1}{2}\right]\right\}$$

and find that the span is given by

$$C = a\begin{bmatrix}1\\0\end{bmatrix} + b\begin{bmatrix}0\\1\end{bmatrix} + c\left[2, \tfrac{1}{2}\right]$$

or $C = aA_1 + bA_2 + cA_3$. Since $A_3 = 2A_1 + 1/2A_2$, we can just replace the span of with

$$c = aA_1 + bA_2 + c(2A_1 + 1/2A_2)$$

which then simplifies to $(a + 2c)A_1 + (b + 1/2c)A_2$, so we can just take a tuple of the three vectors and replace it with only two.

The principle is that we can remove vectors which are linearly dependent on the rest and we don't change the span. If we have a set of four linearly independent vectors and remove one, there isn't a way to get it back from the other three. Whereas if we have a linearly dependent set, we can remove a vector and still get it "back" by using our original vectors.

We then have that

$$\dim V = \text{Smallest number of elements in a given basis.}$$

In linear algebra, dimension is finite. Additionally, $\dim M(m,n) = m \times n$. For a very fast demonstration, consider this example of $M(2,2)$:

$$\begin{bmatrix}a & b\\c & d\end{bmatrix} = a\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + b\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + c\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + d\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$$

To generalize, this is because is $M(m,n)$ has $m \times n$ entries, so we can have $m \times n$ degrees of freedom.

Let $P_2(a,b,c)$ be the set of all polynomials ax^2+bx+c. This has dimension 3, because we can think of the span as a linear combination of x^2, bx and c. As a result the basis is $\{x^2, x, 1\}$.

Checking if a given basis is linearly independent is easy (we have the test for linear dependence), but it's not so easy to check if it actually spans the vector space we want. For this, we have the following:

> "Theorem 2.6"
>
> Suppose that the vector space V can be spanned by n elements. Then any set containing more than n elements must be linearly dependent. That is, if $S = \{A_1, A_2, \ldots, A_n\}$ is a subset of V where $m > n$, then S is linearly dependent.

We don't have to apply the test of linear dependence if we can just test with this first.

> "Theorem 2.7"
>
> Let V be a vector space, and $\dim(V) = n$. Then all bases of V have exactly n elements. In particular, the dimension of V is the number of elements in any basis of V.

8 Day 8: June 19

We will further our discussion on dimension. Last time, we defined what dimension and basis were.

> "Theorem 2.6"
>
> Suppose that the vector space V can be spanned by n elements. Then any set containing more than n elements must be linearly dependent. That is, if $S = \{A_1, A_2, \ldots, A_n\}$ is a subset of V where $m > n$, then S is linearly dependent.

> "Theorem 2.7"
>
> Let V be a vector space, and $\dim(V) = n$. Then all bases of V have exactly n elements. In particular, the dimension of V is the number of elements in any basis of V.

8.1 Proving Sets as Bases

We will discuss another theorem:

"Theorem 2.8"

Let V be an n-dimensional vector space. Then:

- Any set of n elements that spans V must be linearly independent and thus is a basis.
- Any linearly independent set of n elements of V must span V and thus is a basis.

Recall that $\{A_1, \dots A_m\}$ is a basis of V if and only if:

- The span of the set is V
- All A_i are linearly independent.

However, checking these conditions are not straightforward. For the second, we have the test for independence. But for the first, we are currently out of luck. We have to check if a linear combination of the vectors can give us any arbitrary vector in V, but this actually introduces more variables. This means checking if a set of vectors is a basis of a vector space can be difficult.

For instance, $V = \mathbb{R}^2$ and our vectors are

$$\begin{bmatrix} 4 \\ -3 \end{bmatrix}, \begin{bmatrix} 10 \\ 2 \end{bmatrix}$$

we can apply the test for linear independence by solving

$$xA_1 + yA_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and seeing if there exists solutions that are not $x, y = 0$.

For the latter, we have to check if

$$aA_1 + bA_2 = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

and see if a solution always exists for arbitrary choice α and β.

There's a trick, however. We have a failsafe that so long as the dimension of the vector space is the number of vectors, we only need to check one other condition. Going back to our previous example, since the dimension of $\mathbb{R}^2$ is trivially 2 and we have two vectors, we do indeed have this "new" condition satisfied. Then all we have to do is to check if

$$xA_1 + yA_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

has a non-zero solution. This is because having a n dimensional vector space with *one* of the conditions implies the other. In general, we only have to show two of the three conditions are met for a set of vectors of V to be a basis:

- The dimension of V and the number of vectors are the same
- The set of vectors is linearly independent
- The span of set of vectors is V

Example

Let $X_1 = \begin{bmatrix}1\\2\\1\end{bmatrix}, X_2 = \begin{bmatrix}2\\0\\2\end{bmatrix}, X_3 = \begin{bmatrix}1\\1\\1\end{bmatrix}$. Prove that this set of vectors is a basis of $\mathbb{R}^3$

It is very trivial that the dimension of $\mathbb{R}^3$ is 3 and that we have three vectors, so we have one down. All that's left is to prove one of the other conditions. I think showing that it is linearly independent is slightly easier, so we will proceed to find the homogenous solution

$$xX_1 + yX_2 + zX_3 = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

By Gaussian elimination, we have that

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

which does admit a non-unique solution. Continuing to get to row reduced echelon form, we have that

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

which implies that z is a free variable and x and y are pivot variables. Thus, we can let z be some arbitrary constant s and get that

$$\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}-1/2s\\-1/4s\\s\end{bmatrix} = s\begin{bmatrix}-1/2\\-1/4\\1\end{bmatrix}$$

So this fails to be a basis.

Example

Let $X_1 = \begin{bmatrix}1\\2\\1\end{bmatrix}, X_2 = \begin{bmatrix}2\\0\\2\end{bmatrix}, X_3 = \begin{bmatrix}1\\1\\2\end{bmatrix}$. Prove that this set of vectors is a basis of $\mathbb{R}^3$

By the same logic, we have the number of vectors matching the dimensions, so we now seek to find if its linearly dependent or not. By Gaussian elimination, we have that

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

which has that $x, y, z = 0$, which means that they are indeed linearly independent, so the set does indeed form a vector space.

In the homework, there are generally two kinds of problems:

- Find a vector space V
- Find a basis of V

but these are kind of the same thing. This is because a vector space is defined as the span of a basis, so either way we need to find a basis at some point.

The following two statements are essentially equivalent:

$$\text{The basis is } \{A_1, A_2, A_3\}$$

and

$$\mathbb{R}^3 = \operatorname{span}(A_2, A_2, A - 3)$$

note however the word span. It makes a difference!

8.2 Special Kinds of Vector Spaces

"Definition 1.17:" Multiplication of a Matrix by a Vector

Let $A = [A_1, A_2, \ldots A_n]$ be an $m \times n$ matrix, wehre A_i are column vectors of A and let $X \in \mathbb{R}^n$. We define the product of A with X by $AX := x_1A_1 + X_2A_2 + \ldots + x_nA_n$ where

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

It's important that the number of columns of A matches the number of entries in X otherwise some products don't get well defined. The product

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

is not well defined but

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 1 \begin{bmatrix} 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 10 \end{bmatrix}$$

For posterity, one can note that the product of an $m \times n$ matrix with a $n \times 1$ vector yields a $m \times 1$ vector.

Transpose

Let

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

Then we define the tranpose of A, A^t, pronounced A-transpose, as the matrix where all rows and columns are exchanged:

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

With this in mind, we can change something like

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

So we can repackage a system of linear equations into one line: $AX = b$, where A is the coefficient matrix, X is the variable vector, and b is the constant vector. Then, a homogenous system of linear equations is defined as $AX = 0$.

"Definition 1.18:" Nullspace

The nullspace for a matrix A is the solution set to the homogenous system $AX = 0$.

If we have that $A \in M(m, n)$ and $X \in \text{Null}(A)$, $X \in \mathbb{R}^n$ for the homogeneous system of linear equation to be well defined. We have that the nullspace is a subset of n-dimensional space, so we may be interested if it is also a subspace. The answer is yes.

We also have that $A \cdot (cX) = c \cdot (AX)$.

9 Day 9: June 20

We will continue talking about the nullspace.

For any matrix $A \in M(m, n)$, we can always find its nullspace. It is the subspace of R^n. We showed that it was closed under addition and scalar multiplication and that it contained the zero vector. It's also not too hard to check that the nullspace is a subset of $\mathbb{R}^n$. From all of this, we can conclude that it's a subspace of $\mathbb{R}^n$.

Again, null(A) is the set of elements $X \in \mathbb{R}^n$ such that $AX = 0$.

9.1 Nullspace and Columnspace as Vector Spaces

We can also think of S as the set of elements $X \in \mathbb{R}^n$ such that $AX \neq 0$. This is a subset of $\mathbb{R}^n$, but it is *not* a subspace. It fails all the conditions. Obviously, it fails to contain a zero element, as zero multiplied by a matrix yields zero, which negates its existence in the set. What about closure under addition and scalar multiplication? Then $A \cdot (X + Y) = B$, where B is an element in the set. This seems fine. The set is not closed under scalar multiplication, as $A \cdot (cX) = b$ where b is in the set, but this gives us $c\dot{A}X = c \cdot b$, which isn't okay for $c = 0$.

We can think of this as a line. If it goes through the origin, we can say that it's a vector space, represented by all scalar multiples of a given vector. But if it doesn't go through the origin, we can't represent it like this (there's a shift), so it fails to be a vector space.

9.2 Finding the Nullspace

Solve for $AX = 0$.

Find the nullspace of

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

We find the nullspace by solving

$$AX = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which gives us the augmented matrix of

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

Note that the augmented matrix is just the matrix A with a zero column vector tacked on. Now we can solve for it! We apply Gauss-Jordan elimination and get

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

which is consistent so far. Continuing, we have

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

which tells us that x and y are pivot variables and z is a free variable, so we can solve for x, y in terms of z. The system reads that $x - z = 0$ and $y + z = 0$, which gives us that $x = z$ and $y = -z$. Letting $z = s$, we have that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} s \\ -s \\ s \end{bmatrix} = s \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

We can then say the nullspace is the set of all scalar multiples of the vector $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}^t$, or even the span of the vector.

> What is the nullspace of the identity matrix?

Solely the zero element. The augmented matrix gives us

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

which has a solution if and only if we have that $x, y, z = 0$.

> What is the nullsapce of
> $$\begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 7 \\ 0 & 0 & 2 \end{bmatrix}$$

The augmented matrix gives

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

which further simplifies to

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

which has only the zero vector.

Find the nullspace of

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

This one's a bit weird. Note from the definition that the nullspace is the set of all vectors where multiplying the matrix by the vector yields zero, to which it would be any vector. Since the set of all vectors with three entires is $\mathbb{R}^3$, the rest follows.

Let A be a zero matrix with size $M(m,n)$. What is it's nullsapce?

Recall that matrix multiplication is only defined of a matrix with size $m \times n$ if the vector is $n \times 1$. Since any vector multiplied by the zero vector is zero, the nullspace is $\mathbb{R}^n$.

Find the nullspace of

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

By applying Gaussian elimination to the augmented matrix, we have that

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

Here, we have four variables x, y, z, w, where x and y are pivot variables and z and w are free variables. We can then solve for x and y by getting $x = -2z - 3w$ and $y = -z + w$. Since z and w are free, we can assign them arbitrary constant variables s and t. This gives us

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -2s - 3t \\ -s + t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

As the set is all linear combinations of these two vectors, the nullspace is a plane in $\mathbb{R}^2$.

"Theorem 1.12"

Let $A \in M(m,n)$. The sanning vectors for the solution set of the system $AX = 0$ span the nullspace of A and are linearly independent.

Note that this then implies that the spanning vectors for the solution set is then the basis of the nullspace.

As an aside, the dimension of the nullspace of a matrix A is given by $\text{null}(A)$.

9.3 Column Space

"Definition 1.15"

The column space of a matrix is given by the span of the columns of the matrix.

What is the column space of

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

$\mathbb{R}^3$.

The colum space of a given matrix of size $m \times n$ is a subset of $\mathbb{R}^m$.

Say we wish to find a basis of the column space of a given matrix. First, we define what a pivot column is:

Pivot Column

A pivot column is a column vector of the original matrix that, when row reduced, contains a pivot entry.

The basis of the column space is simply the set of the pivot columns of A.

$$\operatorname{span}\left(\left\{\begin{bmatrix} \frac{-10}{3} & -\frac{1}{3} & 1 \end{bmatrix}\right\}\right)$$

10 Day 10: June 21

We will continue our topics on column space.

10.1 Column Space, Again

"Definition 2.3:" Pivot Columns

Let $A = [A_1, A_2, \ldots, A_N]$ be a matrix where A_i are the columns of A. We say that A_i is a pivot column for A if x_i is a pivot variable for the system $AX = B$.

This means we have the logical procedure of

$$\text{Pivot Entry} \Rightarrow \text{Pivot Variable} \Rightarrow \text{Pivot Column}.$$

"Theorem 2.4"

The pivot columns of a matrix A form a basis for the column space of A.

Note that in order to find pivot entries we use Gaussian elimination to get a row reduced echeleon form matrix, so in reality the process is more like

$$\text{RRE} \Rightarrow \text{Pivot Entry} \Rightarrow \text{Pivot Variable} \Rightarrow \text{Pivot Column}.$$

What is the column space of

$$A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 2 & -4 & 4 \\ 1 & 3 & 0 & 4 \end{bmatrix}?$$

Trivially, the column space is given by the span of the four column vectors, but since we have four vectors in three variables, we have that the set is then linearly dependent. What if we want the basis of the column space?

We proceed by applying Gaussian elimination. Reducing, we have that

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

which tells us that the first and second columns contain pivot entires, so that our basis is the set

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \right\}$$

For posterity, the columns in the row reduced matrix do not span the column space. Be careful! This also tells us that any other column vectors in our matrix can be expressed as a linear combination of these two vectors.

10.2 Row Space and Rank

"Definition 2.5:" Row Space

The row space of an $m \times n$ matrix A is spanned by the rows of A. It is denoted by Row Space(A).

This also forms a subspace of $M(1, n)$.

Let A and B be row equivalent matrices. Then A and B have the same row space.

This is because elementary row operations doesn't affect the row space.

Theorem 2.12

The nonzero rows of any echelon form of a matrix A form a basis for the row space of A.

$$A \Rightarrow EF(A) \Rightarrow \text{Non-zero Rows}$$

Note that in finding either column space or row space, we begin by applying row reduction.

"Definition 2.6:" Rank

The rank of A is the dimension of the row space. It is computable as the number of nonzero rows in any echelon form of the matrix A.

Rank Theorem

For any matrix A, the row space and column space have the same dimension. This dimension is the rank of the matrix.

A fast loose proof is that if a row is nonzero, it must have a pivot entry. Thus, we expect the number of pivot entries to be equal to the number of nonzero rows. Thus, the dimension of the column space matches the dimension of the row space. To be clear, however, even though they have the same dimension, they are not necessarily the same set.

For instance, consider the matrix

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

The row space is given by the first and second rows, which together span the xz-plane in $\mathbb{R}^3$. However, the column space is equal to the span of the first and third column, which is just the xy-plane in $\mathbb{R}^2$.

"Theorem 2.14"

The rank of a matrix A is equal to the transpose matrix A^t.

"Theorem 2.17:" Rank-nullity theorem

Let $A \in M(m, n)$. Then

$$\text{Rank}(A) + \text{Null}(A) = n$$

11 Day 11: June 24

Recall that we describe the rank of a given matrix A as

$$\text{rank}(A) = \dim(\text{row space}(A))$$

which, by the rank theorem, that

$$\text{rank}(A) = \dim(\text{column space}(A))$$

and by the rank-nullity (recall the nullity, $\text{null}(A) = \dim(\text{nullspace}(A))$ is the dimension of the nullspace) theorem, if $A \in M(m, n)$,

$$\text{rank}(A) + \text{null}(A) = n.$$

which, for posterity, is the number of columns.

11.1 Quick Proof of Rank-Nullity

Recall that the rank of a matrix A can be computed by the number of pivot entries. For each pivot entry, there is one pivot variable.

Going back to the rank-nullity theorem, the right hand side is n. In a given $m \times n$ matrix, this is the number of variables. We also know that there are two kinds of variables:

- Pivot variables
- Free variables

We know that the rank of A is the number of pivot variables. So all that remains is to show that the nullity of A is equal to the number of free variables to prove the theorem.

Let's find the nullity of

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 4 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

which becomes

$$\begin{bmatrix} 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & 2 & 4 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1 & -3 & 0 \\ 0 & 1 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This implies we have two pivot variables, say x and y, but two free variables, say z and w. From this matrix, we have that $x - z + w = 0$ and $y + 2z + 4w = 0$, which then implies

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

which yields the solution vector of

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} s + 3t \\ -2s - 4t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 3 \\ -4 \\ 0 \\ 1 \end{bmatrix}$$

The basis of the nullspace of A is given by those two vectors, so the nullity of A is 2. Since for every free variable we get another linearly independent vector in the basis, we have that the nullity of the matrix is equal to the number of free variables.

This completes the "proof."

As an aside, now that we know that the nullity of A is 2 and that there are 4 columns (we don't count the right hand side), we have that the rank of A is $4 - 2 = 2$. We can verify this by noting that the row reduced echelon form has only two non-zero rows, so that the rank is 2.

11.2 "Backwards" Linear Algebra

If we have a matrix, we can derive a lot of properties about its corresponding linear system. A question, is then, if we know said properties (rank, nullity, etc), what can we actually say about the linear system?

"Theorem 1.10:" The Translation Theorem

Let A be an $m \times n$ matrix and let T be any particular solution to $AX = b$. Then the general solution is

$$X = T + W$$

where $W \in \text{nullspace}(A)$.

Here, W is a vector in the nullspace of A. If T is a solution, then $T + W$ is a solution as well.

This is easy enough to check. If T is a solution, then we know $AT = b$. But what about $A(T + W)$? By the distributive law, we have $AT + AW$. Since $AW = 0$ and $b + 0 = 0$, we have that $A(T + W) = b$.

So how can we use this? Suppose we have the following linear system

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 4 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

which tells us that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \underbrace{\begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix}}_{\text{A constant vector.}} + \underbrace{s \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}}_{\text{A vector in the nullspace}}$$

Note that the vector T is a solution, and that $AW = 0$. Since we have an arbitrary scalar s, we have that the nullspace is given by the span of the vector W. The sum of these two vectors is precisely the form of the translation theorem.

So when is a solution unique? Recall, that any solution is given by the form $X = T + W$. So if W is a non-zero vector, we can arbitrary translate/shift a given solution T to yield infinitely many solutions. But if the nullspace contains solely the zero vector, or $W = \vec{0}$, $T + W = T$, so there is only a single solution. As an aside, a vector space having solely the zero vector is equivalent to the dimension being zero. From this, if the nullity of A is zero, then T cannot be shifted, so there is only a single, unique solution.

> "Theorem 2.18"
>
> Let $A \in M(m, n)$ where A is the coefficient matrix. There is at most one solution to $AX = b$ for all $b \in \mathbb{R}^m$ if and only if the rank of A is n.

Here, n is the number of columns.

If the rank of the matrix is n, then the nullity of the matrix is 0. Then there must exist only a single unique solution to the matrix.

Why at most 1? As an example, we have that the augmented matrix is given by

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

which yields the inconsistent $0 = 2$, so there is no solution. But the rank of the coefficient matrix

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

is 1.

This theorem tells us that if a solution exists, we can determine whether or not it's unique. However, this theorem doesn't comment on if a solution exists. To this, we have another tool:

> "Theorem 2.16"
>
> The linear system $AX = b$ is solvable for all $b \in \mathbb{R}^m$ if and only if the rank of A is equal to m.

Here, m is the number of rows.

Let's begin with AX. We know that this is equal to $x_1A_1 + x_2A_2 + \ldots + x_nA_n$, where each of A_i is a column vector, or $A_i \in \mathbb{R}^m$. Thus, AX is a linear combination of the column vectors A_i. We then have that b is equal to that linera combination, so b must thus be in the span of the A_is $\text{span}\{A_1, A_2, \ldots, A_n\}$ if we want the solution to exist. So if b is not in the span, the solution does not exist.

So what is $\text{span}\{A_1, A_2, \ldots, A_n\}$? This is the column space of the matrix. We want this to work for all b, so we need the column space to be $\mathbb{R}^m$ so that our column vectors to be spanning. We don't care if they're linearly dependent–so long as they span $\mathbb{R}^m$, it's ok. So thus we require that the rank of A be equal to m.

11.3 Bounding Rank by Size

If $A \in M(3,5)$, the rank of A is either 0, 1, 2 or 3. The rank of A is defined as the number of non-zero rows, and we only have three rows. But we can also have all rows be zero, in which case we have zero non-zero rows. Thus,

$$0 \leq \text{rank}(A) \leq 3$$

What about if $A \in M(5,3)$? Again, the rank is either 0, 1, 2 or 3. Here, we appeal to the fact that the rank is the number of pivot entries, and we can only have one per column. Since we have three, we can either have no pivot variables to all three columns having a pivot varaible.

From this exercise, we have that

$$0 \leq \text{rank}(A) \leq \min(m,n).$$

Let $A \in M(10,9)$. Is it possible that $AX = b$ is solvable for all $b \in \mathbb{R}^{10}$?

No. The rank is less than or equal to 9, which fails the condition of needing to be 10 to have a solution for all $b \in \mathbb{R}^{10}$. So when $m > n$, or where the number of equations is greater than the number of variables, it is not guaranteed to have a solution.

If $m < n$, we have that if the solution exists, it is non-unique. The rank is at best m, which is not n, so the nullity is at least 1, so we have infinitely many solutions.

11.4 Summary

"Definition 2.8"

An $m \times n$ matrix A is said to be nonsingular if the solution to $AX = b$ both exists and is unique for all $b \in \mathbb{R}^m$.

It's worth noting that it says *all* $b \in \mathbb{R}^m$, because there can exist *some* b that work, but not all.

If we want the existence for all $b \in \mathbb{R}^m$, we need the rank of A to be equal to the number of rows, m. And to guarantee uniqueness, we need the rank of A to be equal to the number of columns, n. Putting these two together,

"Theorem 2.19"

A is nonsingular if and only if the rank of the matrix rank(A) is equal to the number of rows m and the number of columns n.

12 Day 12: June 25

We discussed two things last class:

"Theorem 2.16"

The linear system $AX = b$ is solvable for all $b \in \mathbb{R}^m$ if and only if the rank of A is equal to m.

The linear system $AX = b$, if a solution X exists, is unique if and only if the rank of A is equal to n.

"Theorem 2.19"

A is nonsingular (a unique solution always exists) if and only if the rank of the matrix $\text{rank}(A)$ is equal to the number of rows m and the number of columns n.

"Theorem 2.20"

Let $A \in M(n, n)$ (so A is a square matrix). Then the following statements are equivalent to each other:

- The nullspace of A is just the zero vector/just the origin/is $\{\vec{0}\}$
- For each $b \in \mathbb{R}^n$, the system $AX = b$ has at most one solution. (This comes from theorem 2.18)
- The system $AX = b$ has a solution for all $b \in \mathbb{R}^n$. (This comes from theorem 2.16)
- A is non-singular.

If A is non-singular, we know that the rank of A is equal to n, the number of rows *and* the number of columns. The other statements follow naturally. The first statement is also a consequence of the rank-nullity theorem.

12.1 Review of Chapters 1-2

- Solving a linear system.
    - Coefficient and augmented matrices.
    - Gaussian elimination. We aim to transform $A \Rightarrow \text{EF}(A) \Rightarrow$ Solution!?.
        - If at any point, we have that a row of zeros equals a nonzero, we have an inconsistency.
        - A unique solution exists if we have the identity matrix as our coefficient matrix.

        * Infinitely many solutions exist if we have the coefficient matrix as a row-reduced echelon form. We can then find the general solution by solving for the pivot variables in terms of the free variables.

- Test for independence. This is achieved by solving the homogenous system $AX = \vec{0}$ or $x_1 A_1 + \ldots x_n A_n = \vec{0}$.
- Theoretical stuff
    - Vector spaces – a set of vectors closed under addition and scalar multiplication.
    - Bases are a finite set of vectors which generates the (often) infinitely many vectors in a vector space. They must be linearly independent, spanning (their span *is* the vector space) and their dimension is the number of elements. Only two of these are necessary to prove that a set is a basis.
    - Examples of vector spaces are $M(m, n)$ and $\mathbb{R}^n$.
    - Examples of subspaces are nullspaces, column spaces, and row space, which can all be derived from the matrix.
    - Rank theorem (rank of the column space equals the rank of the row space). The spaces are different, but they have the same dimension.
    - Rank-nullity theorem. $\text{rank}(A) + \text{null}(A) = n$ where n is the number of columns.
    - Rank tells us information about the number of solutions. If the matrix is "full rank", there exists only one solution for all constant vectors.

12.2 Every T/F Question

1.1 True.

1.2 False. Not true in general, we can get either a linearly independent or a linearly dependent subset.

1.3 False. Not true in general, a set can contain a set of linearly independent vectors, but still have a vector which is a linear combination of those linearly independent vectors.

1.4 True.

1.5 False. Not true in general, only one vector has to be a linear combination of the others, but not all of the others.

1.6 True. This is true by definition, but also because any linear combination of the zero vector (with all other vectors zero) is the zero vector.

1.7 False. We have a vector which is linearly dependent on the others, so the set is linearly dependent.

1.8 False. It is possible that X is linearly independent from the rest, and that some other vector is linearly dependent.

1.13 False. A plane is two dimensional, so it's ok.

1.14 True. We can have 0, 1 or infinitely many, but no other possibilities.

1.15 False. By the rank-nullity theorem and the translation theorem, our solution space is two dimensional and is thus a plane.

1.16 False. Two parallel planes is the fastest counterexample.

1.17 False. Infinitely many solutions is still an option.

1.18 True. We can't have more non-zero rows than there are rows.

1.19
- a True. If the first system has a unique solution, the coefficient matrix must be row equivalent to an identity matrix. Since the second system has the same coefficient matrix, we can end up in the same situation and get a single solution.
- b False. Adding one to an entry of the constant vector shifts the plane, which can cause the planes to no longer intersect.
- c False. The exact same process from above applies, albeit in the opposite direction.

1.20 False. Not true in general, we can get an identity matrix with a constant column, but we can also get an inconsistent row as well. It has at most one solution.

1.21 False. If there are infinitely many solutions, the nullity must be at least one. But if there are three non-zero rows, the rank is 3. This contradicts the rank-nullity theorem.

1.22 "do it urself lol"

1.23 True. Add the third row to the fourth.

1.24 True. We get the identity matrix, which is consistent as a coefficient matrix.

1.25 True. The rank-nullity theorem is at least one.

1.26 False. The rank can be 3, which would imply the nullity is 0, which means the nullspace is solely the zero vector.

1.27 True. Rewriting this as two equations then applying subtraction implies the vector belongs to the nullspace.

1.28 True. Decomposition into iterated vector addition and scalar multiplication implies that it is part of the nullspace.

1.29 True. With some algebra and simplifications, this is actually the span of only one vector, so it's the solution of a homogenous system.

1.30 False. By the same logic as 1.29, the vector is a solution to a homogeonus system.

1.31 False. If we set up an augmented matrix and apply row reduction and then check for the conditions where we have a consistent system, our solutions lie on the span of a single vector, which is a line.

1.32 True. If we explicitly write out the span as the set of all possible linear combinations, the first span has three degrees of freedom, but the second only has two.

1.33 False. You can't get a zero element. Also, using any two vectors makes it very clear it's not closed under addition. This also isn't closed under scalar multiplication.

1.34 True.

1.35 False. This fails to be closed under scalar multiplication.

1.36 False. The subspace containing just the zero vector is indeed a subspace. Note, however, if we have any non-zero vector, we have infinitely many elements by scalar multiplication.

2.1 False. Rank can be at best 5, so only 5 linearly independent columns.

2.2
- a False. We only have two pivot entries, so we only have two pivot columns, which form a basis of the column space, so the other columns necessarily are linear combinations of those two pivot columns.
- b True. It's not a pivot column, so it must exist as a linear combination.
- c True.
- d True. Inspsection is sufficient.
- e False. Inspection is sufficient.

2.3 True. An $n \times n$ matrix row reduces to a identity matrix if their rows are linearly independent. So both will be the identity matrix.

2.4 True. If there are infinitely many solutions, the nullity of the matrix must be non-zero, so the rank of the matrix is not the number of rows or columns. Thus we have fewer linearly independent columns than we have columns, so we have at least one linearly dependent column.

2.10 False. By inspection, they are linearly dependent, and has dimension two, which is insufficient to span $\mathbb{R}^3$.

2.11 True. Although they are not a basis (since we have more vectors than we have dimension), they do indeed span $\mathbb{R}^2$ since we can find two linearly independent vectors.

2.12 False. We can get a plane that isn't one of the axial planes, and it can't be spanned by the axes. In another sense, we only have three total planes possible. Importantly, these three vectors do indeed span three dimensional space, but we cannot pick two and get any plane in three dimensional space.

2.13 False. We have four three dimensional vectors.

2.14 True. If the nullity is four, then the rank of the matrix must be zero, so it would be a zero matrix. This is a contradiction.

2.15 False. These four vectors must be linearly independent.

2.16 False. We don't know for sure if the set of four vectors has dimension 4 or not. It's possible that we removed a linearly independent vector.

2.17 True. If five vectors span a four dimensional space one is linearly dependent, then we just need to pick four linearly independent vectors to get a basis.

2.18 True.

2.19 False. We only have four vectors to span a four dimensional space, so they're all necessary.

2.20
 a True. We have that the nullity is three, so the rank of the matrix is 2 by the rank nullity theorem.
 b False. This is only possible if the rank is equal to three, but it's not.
 c False. This is only possible if the rank is equal to five, but it's not.

2.21 True. If there always exists a solution, the rank of the matrix is m. The transpose of the matrix is $n \times m$. Thus we have uniqueness for a solution if it exists.

2.22 True. The same reasoning from 2.21 applies here.

2.23 False. The same reasoning from 2.21 applies here.

2.25 True.

2.26 True. If the column space is a line, the rank is 1, so the dimension of the row space is also 1, so there is only one linearly independent row.

2.27 True. The rank-nullity theorem puts a restriction on the number of columns but not the number of rows.

13 Day 14: June 27

We begin chapter 3, the study of

13.1 On Linear Transformations

We've discussed what it means to be linear before, but we will discuss it again too. We can consider a transformation as a special kind of function.

We've worked with cartesian coordinates, but we've also worked with polar coordinates. With the formulas $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1} y/x$, we can actually transform from cartesian to polar. However, this transformation isn't linear. We again will discuss what that means later, but we will first start with calculations.

13.2 Matrix Multiplication

In chapters 1-2, we've looked at multiplying a matrix A with a vector X to get AX. Can we then multiply two *matrices*? Can we compute AB if A and B are both matrices?

We can consider the second matrix, B, as a set of column vectors $[B_1, \ldots, B_n]$, so $AB = A \cdot [B_1, \ldots B_p]$. Since we know how to multiply matrices with column vectors, we can do a sort of distributive property and get

$$A \cdot B = [AB_1, \ldots AB_p] \, .$$

AB_1 is only defined if $B \in \mathbb{R}^n$, otherwise matrix multiplcation here will not be defined. This follows for all the other column vectors. For posterity, $AB_1 \in \mathbb{R}^m$.

Since each AB_i is $m \times 1$ and we have p of them, our matrix is of size $m \times p$.

If $A \in M(m, n)$, matrix multiplication AB is well defined if and only if $B \in M(n, p)$. It is defined as

$$AB = [AB_1, \ldots AB_p] \, .$$

Compute

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

This is indeed well defined, as A is 2×3, and B is 3×2.

Computing, we have that the product is

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

The product is 2×2 since we have 2 columns in B and 2 rows in A

Compute

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

This is not well defined, as we have 3 columns in A but 2 rows in B.

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

This is indeed well defined, as we have two columns in A and 2 rows in B. Computing, we have that
$$\begin{bmatrix} 2 & 5 & -1 \\ 0 & 3 & 3 \end{bmatrix}.$$

Commutativity of Matrix Multiplication

In general, matrix multiplication is not commutative. For two matrices A and B,
$$AB \neq BA.$$

Compute
$$\begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 2 & -5 \end{bmatrix}.$$
and
$$\begin{bmatrix} -1 & 1 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -3 \end{bmatrix}.$$

These are both well defined since both matrices are 2×2. Computing, we have that
$$\begin{bmatrix} 3 & -9 \\ -6 & 15 \end{bmatrix} \qquad \begin{bmatrix} 1 & -5 \\ 2 & 19 \end{bmatrix}$$

Note that the identity matrix multiplied by an arbitrary square matrix is just the square matrix. We actually have commutativity here.

13.3 Some Properties of Matrix Multiplication

We can rewrite A as a set of row vectors, and we have
$$AB = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \cdot B$$

Since A is $m \times n$, we have that each row vector is $1 \times n$. B is $n \times p$, so the product of each row with B is $1 \times p$. Since we have m total rows, each being p long, we have that the matrix is $m \times p$.

Either perspective is correct in computing the matrix product.

Proposition 3.2

In general, for the product of two matrices A and B,

In addition,

- Matrix multiplication is not commutative.
- Matrix multiplication is associative. $ABC = A(BC) = (AB)C$. There's a proof, but I won't bother with it.
- Matrix multiplcation is distributive over matrix addition. That is, $A(B + C) = AB + AC$ If $A \in M(m, n)$ and $B, C \in M(n, p)$.
- Matrix addition is distributive over matrix multiplication. That is, $(B + C)A = BA + CA$ if $A \in M(n, p)$ and $B, C \in M(m, n)$.
- $AB^t = B^t \cdot A^t$.
- $cAB = A(cB)$. In general, we can introduce a scalar at any point in the equation and it will keep its value.

What happens to the ranks if we multiply two matrices?

"Theorem 3.3:" Rank of Products

$$\operatorname{rank}(AB) \leq \min(\operatorname{rank}(A), \operatorname{rank}(B)).$$

If A and B are non-singular, then AB is also non-singular with rank n.

13.4 Inverse Matrices

When we studied the real numbers, we had four binary operators, addition, subtraction, multiplication and division. Note that these come in pairs as subtraction is the inverse of addition and likewise.

So we have defined addition, subtraction and multiplication for matrices. Can we define an inverse to multiplication? Given a matrix A, what is the matrix equivalent to $1/A$? Once we have a reciprocal, we can define an inverse to multiplcation.

Taking $1/A$, what is "1" in a set of matrices? It has two properties:

- It is a multiplicative identity: $1 \cdot a = a \cdot 1$.
- $\frac{1}{a} \cdot a = 1$.

Do we have one of these? Yes. This is the identity matrix.

14 Day 15: June 28

Test!

15 Day 16: July 1

15.1 Review on Matrix Multiplication and the Identity Matrix

Recall we discussed matrix multiplication prior. The first matrix A must be $A \in M(m, n)$ and $B \in M(n, p)$ for AB to be well defined. We also discussed a very important matrix, the identity matrix I, which is also a square matrix. It has the very important proeprty where $A \cdot I = A$ and $I \cdot A = A$.

Take care, however. For example, if $A \in M(2, 3)$, then $A \cdot I$ is only well defined if and only if I is a 3×3 identity matrix. But if we wanted to compute $I \cdot A$, then I must be 2×2 for the product to be well defined. They both yield the same product, though.

15.2 On Inverse Matrices

We discussed this as a means to include "division" for matrices. For the real numbers, we have the multiplcative inverse $1/a$, which is only well-defined if and only if a is not zero.

Now, for matrices, we are looking for I/A where A is some matrix. We should also expect $A \cdot I/A = A$. This motivates the following definition:

"Definition 3.8:" Inverse matrices

An $n \times n$ matrix A is invertible if there is a matrix B such that $AB = I = BA$ where I is the identity matrix. We can then write $B = A^{-1}$ and refer to B as the inverse of A or "A-inverse".

As a result, we should expect both AA^{-1} and $A^{-1}A$ to yield the identity matrix.

So what kinds of matrices are invertible? And why do require A to be $n \times n$? Why does it have to be square? Well, the identity matrix is a square matrix, so if A is not a square matrix, we run into issues. If A is 5×4 and we wish to multiply it by a matrix to yield an identity matrix, we get that B is a 4×5 matrix for a 5×5 identity matrix.. But then BA yields a 4×4 identity matrix. A perhaps more important reason is

$$\operatorname{rank}(AB) \leq \min(\operatorname{rank}(A), \operatorname{rank}(B)).$$

where in this context, AB is the identity matrix, which gives us

$$\operatorname{rank}(I) = n \leq \min(n, n).$$

This only holds if all of their ranks are n, so A and B are non-singular. This motivates the following theorem:

"Theorem 3.9:" Nonsingular Matrices are Invertible

Let A be an $n \times n$ matrix. A is invertible if and only if A is nonsingular.

A is inveritble if and only if A has rank n.

The process for inspecting if a given matrix is invertible is thus as follows (this is in lieu of using examples):

- Check if the matrix is $n \times n$
- Check if the rank is n.

"Theorem 3.10:" Inverse Matrices are Unique

Let A be an $n \times n$ matrix. If either of the following statements hold, then A is invertible and $B = A^{-1}$:

- There is a matrix B such that $BA = I$ (there exists a left inverse).
- There is a matrix B such that $AB = I$ (there exists a right inverse).

There only exists a single unique inverse matrix–the left and right inverse are both the inverse.

We make this seemingly pedantic point simply because $AB \neq BA$. This is useful because we only have to show one side and the other is implied.

15.3 Finding Inverse Matrices

The simplest case if we are trying to find the inverse of an identity matrix. The inverse of it is itself.

What about a diagonal matrix such as

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

We are looking for a matrix such that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Multiplying through, we have that

$$\begin{bmatrix} -3a & 2b \\ -3c & 2d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

which tells us that $a = -1/3$ and $d = 1/2$, so our inverse matrix is

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

We can generalize this:

> The inverse of a diagonal matrix is the matrix with all of its diagonal entries reciprocated.

> Find the inverse matrix of
> $$\begin{bmatrix} 1 & 2 \\ 3 & -2 \end{bmatrix}.$$

Following the same procedure, we end up with

$$\begin{aligned} a + 3b &= 1 \\ 2a + 2b &= 0 \\ c + 3d &= 0 \\ 2c + 2d &= 1. \end{aligned}$$

which we can solve. Things get out of control easily, though. For a given $n \times n$ matrix, we have to juggle n^2 variable. So we wish to find a more efficient method.

15.4 A Better Method for Inverting Matrices

15.4.1 Row Swaps as Matrix Multiplication

First we need to talk about Gaussian Elimination. Note that if A is nonsingular, it gives us I. Is this process somehow related to matrix multiplication? Yes.

We can interchange two rows. This is equivalent to multiplying our matrix with a matrix on its left. Consider the operation

$$\begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} \Rightarrow \begin{bmatrix} A_2 \\ A_1 \\ A_3 \end{bmatrix}$$

but we can also accomplish this by multiplying PA. What is P? It would be

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

Where did this come from? Well we have to start with a 3×3 matrix. Since we switched the first and second rows, we can just switch the first and second rows of the identoity matrix.

Let A be $n \times n$. We can find a matrix P such that PA corresponds to switching the x and yth rows of A by taking the $n \times n$ identity matrix and switching the x and yth rows of it.

Note that P corresponds to switching the x and yth row. If we do it twice, we go back to the original matrix. Thus $PPA = A$. Note that by identity matrix properties, $PP = I$ if $PPA = A$, so $P^{-1} = P$.

What about if P corresponds to a matrix that can only be yielded by a *sequence* of exchange of rows? Let P_1 be the first exchange, then P_2 the second. Then our desired matrix is given by P_2P_1A. The product of multiple matrices for elementary row opeartions stacks *leftward*.

15.4.2 Stripping Rows as Matrix Multiplication

If we want to add cA_j to the row A_i. Then P is given by where the constant c is in the i-th row and j-th column.

So if we wanted to find P such that corresponds to

$$\begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} \Rightarrow \begin{bmatrix} A_1 \\ A_2 + 2A_1 \\ A_3 \end{bmatrix}$$

This would be given by

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

15.4.3 Scaling Rows as Matrix Multiplication

This one's straightroward. If we want to scale the ith row by c, we simply take the identity matrix, and replace the ith diagonal entry by c.

15.5 Gaussian Elimination, Again

Gaussian elimination can then be represented as a sequence of left matrix multiplications.

16 Day 17: July 2

We will continue our discussion on inverse matrices.

Nonsingular Matrices are Invertible

If A is nonsingular, then A is invertible and A^{-1} exists.

Again, we know that $A^{-1}A = I$, where $I \in M(n, n)$. Last time, we introduced the idea of writing elementary operations as a left multiplcation of a matrix P.

Swapping two rows corresponds to switching said rows of the identity matrix. Adding a row by a constant of another row is accomplished by adding a constant in the column of the row we are adding and in the row we are adding to.

Let
$$A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & 1 \\ -2 & 1 & 5 \end{bmatrix}$$
write the elementary row operations to reduce it as a sequence of left-matrix multiplication.

We first begin by subtracting two times the first row from the second row to give us
$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & -5 \\ -2 & 1 & 5 \end{bmatrix}.$$
which corresponds to
$$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Continuing, we look to add two times the first row to the third row, which yields
$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & -5 \\ 0 & 1 & 11 \end{bmatrix}.$$
which can also be expressed as
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & -5 \\ -2 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & -5 \\ 0 & 1 & 11 \end{bmatrix}.$$

Then we want to add negative one times the second row to the third row, which we can then represent as
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & -5 \\ 0 & 1 & 11 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & -5 \\ 0 & 0 & 6 \end{bmatrix}.$$

Then to make the last entry in the third row 1, we divide by 6 which is accomplished like so:
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & -5 \\ 0 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & -5 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then we want to add five times the third row to the second

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

Then we want to subtract three times the third row to the fisrt

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

Then to complete, we change the sign of the second row.

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

I haven't labeled them, but we've done a total of 7 matrix multiplications to reduce our original matrix. Importantly, note that as we multiply our matrices, they come from *previous steps.* So our first product is given by P_1A, so the second one is given by P_2 multiplied by the previous product, which is thus P_2P_1A. Generalizing to the very end, where we got the identity matrix, we get the horrendously long equivalent of Gaussian elimination:

$$P_7P_6P_5P_4P_3P_2P_1A = I.$$

Because matrix multiplication is *associative*, we can multiply all the Ps together, call it B I guess, and we get $BA = I$, which is precisely the definition of an inverse: $B = A^{-1}$ so

$$A^{-1} = P_7P_6P_5P_4P_3P_2P_1.$$

This still sucks, since computing seven products is awful.

16.1 Adjoining the Identity Matrix to Invert

Instead of repeatedly reducing A to I by doing PA, we can create a new matrix by connecting A next to I, which yields a $n \times 2n$ matrix. We can represent this new matrix as

$$\begin{bmatrix} A & I \end{bmatrix}.$$

Applying the first elementary row operation then gives us

$$P_1 \begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} P_1A & P_1 \end{bmatrix}.$$

Generalizing, we end up with

$$P_7 \begin{bmatrix} P_6P_5P_4P_3P_2P_1A & P_6P_5P_4P_3P_2P_1 \end{bmatrix} = \begin{bmatrix} I & A^{-1} \end{bmatrix}.$$

The idea is to do apply elementary row operations to the left hand side and apply Gauassian elimination. We still have to apply the operations to the entire matrix, importantly. THe right half will, step by step, transform into the identity matrix.

This was an extremely long demonstration of this:

> To compute the inverse matrix of A, adjoin the identity matrix I to the right of A to get a $n \times 2n$ matrix, and apply elementary row operations as if you were trying to reduce the left half. When done, the right half will be the inverse matrix.

I'm not going to typeset examples here, sorry.

16.2 Applications of Inverse Matrices

We can actually work with "division" now, where $A/B = AB^{-1}$. For this to be well defined, B must be non-singular. This is supposedly equivalent to the condition that we can't divide by 0.

More importantly, however, if we have $AX = b$ where there exists a solution X for all b, we can solve for X by computing

$$A^{-1}AX = A^{-1}b$$

which simplifes to

$$IX = A^{-1}b$$

and then

$$X = A^{-1}b.$$

However this isn't actually that useful. We first have to find A^{-1} and then apply it to b, which is two steps instead of just applying elementary row reduction to $AX = b$. On the contrary, this is useful when solving a *series* of system of linear equations, which apparently exist. So if we are to solve

$$\begin{aligned} AX_1 &= b_1 \\ AX_2 &= b_2 \\ &\vdots \\ AX_k &= b_k \end{aligned}$$

this can be done by multiplying all sides by A^{-1}. Then we only have to do elementary row opeartions once, and just carry out k matrix multiplications rather than doing elementary row opeartions k times.

16.3 Properties of the Inverse

We know that $(AB)^t = B^t A^t$. There is a similar idea for $(AB)^{-1}$, where the inverse is $B^{-1}A^{-1}$. We can think of it as first undoing the multiplication by B then undoing the multiplication by A.

16.4 LU Decmoposition Sneak Peek

In the "real world", we don't actually solve matrices by hand. We usually just let a computer do it. However, even though computers are faster than humans in computing, there's still a barrier where it becomes infeasible. In general, the big-Oh of Gaussian elimination is $O(n^3)$, but there exists a technique to bring it down to $O(n^2)$

17 Day 18: July 3

17.1 Compuational Costs for Gaussian Elimination

Gaussian elimination is $O\left(n^3\right)$. The logic is because Gaussian elimination works as follows:

- We apply elementary row opeartions for each pivot entry. For a given $n \times n$ matrix the worst case is n pivot entries.
- For each row below the pivot entry, we have to subtract away a multiple of the previous row. This is thus done $n(n-1)/2$ times. This addition/subtraction process grows with respect to row size so it's $O(n)$.
- Backsubstitution exists, which takes $n(n-1)/2$ operations.

With some extremely shitty multiplication, we can then conclude that Gaussian elimination is $O\left(n^3\right)$.

However, this is is the worst case. We have some better options for matrices, too.

- The computational cost of applying Gaussian elimination to an identity matrix is 0. It takes n checks but we don't need to actually calculate anything.
- For a diagonal matrix, the cost is n. We just do n divisions.

What about for a lower triangular matrix? The advantage is that if we have a lower matrix, the cost of slicing away rows is constant. The first row has one entry, and the second entry has two. Subtracting the first entry away from all the lower rows brings them down an entry, so the process of slicing away rows doesn't grow with size, just the number of slices we have to do. Thus this is only $O\left(n^2\right)$. For what it's worth, the logic here also applies to upper triangular matrices.

17.2 LU Decomposition

If we can take a given system $AX = b$ and factor A into a product of a lower l and upper triangular matix u, we can then rewrite this system into

$$LU(X) = b.$$

Denote the product of U and X to be Y I guess, and this then becomes

$$LY = b.$$

which we can then solve in $O\left(n^2\right)$. Since matrix multiplication is $O\left(n^2\right)$, this entire process is $O\left(n^2\right)$, an order better than regular Gaussian elimination.

When we looked ofr the inverse, we proceeded by solving

$$\begin{bmatrix} A & I \end{bmatrix}$$

to find A^{-1}. But this isn't what we're looking for. Instead, we stop when the left hand side is in echelon form, then the right hand side will then be in upper triangular form. The idea is that once we stop, we take the right hand side, which is now in upper triangular form, and then apply the process of finding its inverse.

LU Decomposition

Apply the same process of finding an inverse matrix for a matrix A, but then stop when the left side matrix becomes echelon form. Denote the left side matrix at this point U and the right side matrix as P. Find the inverse matrix of P. Call the inverse of P L. Then we have that $A = LU$.

The idea is to turn A into an upper triangular matrix, U and find its associated P_1, then take P_1 and apply row reduction to it to get P_2. Then we have that $A = P_2U$.

17.3 Applying LU Decomposition to Solve a System

Say we have that

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

and we are looking for where $AX = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Proceed by solving $LY = b$, which involves adjoining the constant vector to our lower triangular matrix.

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

Then solve $UX = Y$, which we do by adjoining the constant vector from the previous step to our upper triangular matrix. Once solved, we are finished.

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

17.4 Restrictions on LU Decomposition

Not all matrices can be decomposed as such.

Type-2 Matrices

We say that a given $m \times n$ matrix is type-2 if and only if A may be brought into echelon form using a sequence of only type-2 elementary row operations.

Here, a type-2 elementary row operation is adding multiples of rows to other matrices.

This is of some importance because if we had to switch rows, we switch rows of P which renders it no longer lower-triangular.

The inverse of an upper-triangular matrix is an upper triangular matrix and vice-versa for a lower-triangular matrix.

"Theorem 3.12:"

Let A be an $m \times n$ type-2 matrix. Then there is an $m \times m$ lower triangular unipotent matrix L and an $m \times n$ echeleon matrix U such that $A = LU$.

We define a unipotent matrix as a lower triangular matrices with it's rightmost entry in each row to be 1.

This directly implies that if a matrix is not type-2, we cannot apply LU decomposition on it.

If A is nonsingular, then U is upper-triangular.

17.5 PLU Decomposition for non-type 2 Matrices

We will discuss this tomorrow for how to attack said matrices.

18 Day 19: July 4

Happy independence day! No class today.

19 Day 20: July 5

I took notes for this on paper. I'll get to adding this later.

20 Day 21: July 8

20.1 Exam Review

I missed one question due to misreading a question and missed one question by forgetting that the test for linear (in)dependence has that the solution

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

doesn't show dependence as it is a trivial solution.

20.2 Textbook Errata

The textbook claims that the nullspace of a $m \times n$ matrix A is a subspace of $\mathbb{R}^m$, but is actually a subspace of $\mathbb{R}^n$. This makes sense since the number of columns is the number of variables, so our vectors in the subspace have n entries. The column space of A is a subspace of $\mathbb{R}^m$. This is because the number of entries per given vector is based on the number of rows.

20.3 Matrix and Linear Transformations

Let T be a transformation

$$T : V \to W$$

where V and W are vector spaces. We say T is linear if it satisfies the two conditions:

- Linearity of addition/additive property – $T(X+Y) = T(X)+T(Y)$
- Distributive of scalar multiplication/scalar property – $T(c \cdot X) = c \cdot T(X)$.

We don't have a specific expression (something like $f(x) = x^2$) so far but just a list of properties. So, for an example, are matrix transofrmation linear? Going through the properties:

- $T_A : \mathbb{R}^n \to \mathbb{R}^m$, so it is indeed a map from and to vector spaces.
- $T_A(X + Y) = A(X + Y) = AX + AY = T_A(X) + T_A(Y)$ through the distributive law, so we do have the additive property.

- $T_A(cX) = A(cX) = c(AX) = cT_A(x)$ since vector spaces are closed through scalar multiplcation, so we have the scalar property.

This proves that all matrix transformations are linear transformatins. But what about the other direction? Are all linear transformation matrix transformations? The answer is no. We have some restrictions on domain and range for matrix transformations ($\mathbb{R}^n \to \mathbb{R}^m$), but we can have linear transformations for arbitrary vector spaces. That being said, all linear transformations which match these domain and range restrictios are indeed matrix transformations.

"Theorem 3.1:" The Matrix Representation Theorem

Let
$$T : \mathbb{R}^n \to \mathbb{R}^m$$
be any linear transformation. Then there exists a unique matrix A such that
$$T(X) = A \cdot X$$
for any and all $X \in \mathbb{R}^n$.

20.4 Finding Matrix Transformations

"Definiton 3.4:" The Identity Transformation

This transforms anything to itself.
Let V be a vector space. The identity transformation V is the transformation
$$I_V : V \to V$$
defined by $I_V(X) = X$ for all $X \in V$.

As a simple consequence, the domain and range are the same. I won't prove that it's a linear transformation either. The matrix which corresponds to this transformation is the $n \times n$ identity matrix, so we have that

$$I_V(X) = A \cdot X = X.$$

Previously we discussed the concept of an image $T(S)$. Consider the identity transformation $I : \mathbb{R}^2 \to \mathbb{R}^2$. On the domain, we have the vectors $\begin{bmatrix}1 & 2\end{bmatrix}^t$ and $\begin{bmatrix}2 & 1\end{bmatrix}^t$. So what is $I(S)$? the set of these vectors *after* the transformation? It's the same set.

What about if S is some vertical line segment? This corresponds to all vectors stemming from the origin whose heads end on the line segment. $I(S)$ will still be that segment.

Okay, not what about if S is the unit square? Okay you get this gist of it.

21 Day 22: July 9

We previously discussed that for a given linear transformation, if the domain is $\mathbb{R}^n$ and the range is $\mathbb{R}^m$, then the linear transformation is a matrix transformation. This is not true in general, however. We also discussed the identity transformation $I_V(X)$ where V denotes the vector space which is simultaneously the domain and the range for the identity matrix. If V is $\mathbb{R}^n$, then the matrix transformation which corresponds to this is multiplication by the $n \times n$ identity matrix.

21.1 More Examples of Linear Transformations

The 3B1B Trick

Given n basis vectors $x_1, x_2, \dots, x_n$, the matrix A which corresponds to the linear transformation

$$\begin{aligned} x_1 &\to x_1' \\ x_2 &\to x_2' \\ &\vdots \\ x_n &\to x_n' \end{aligned}$$

is given by

$$A = \begin{bmatrix} x_1' & x_2' & \dots & x_n' \end{bmatrix}.$$

For a more complete treatment, refer to this video.

The majority following subsubsections (except rotations) are admittedly not very important if you internalize this.

21.1.1 Dialation

Suppose we have the unit circle in $\mathbb{R}^2$. Its center is the origin with radius 1. Suppose we want to get an ellipse from it, where the major axis is the y-axis from -2 to 2 and the minor axis is the x-axis from -1 to 1. This is accomplished by stretching the unit circle on the y-axis, which we do by taking all points on the unit circle and doubling its y-coordinate.

We have this procedure for other stretches and squashes. These linear transformations are given by some diagonal matrix

$$A = \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}.$$

For instance, if we wanted to find the matrix which corresponds to stretching the unit circle on the x-axis three times, we would look for a matrix A such that

$$A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ y \end{bmatrix}.$$

We can do matrix multiplication and find it explicitly, but the 3B1B trick tells us it should be

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

We also have higher dimensional analogues. I won't bother.

21.1.2 Reflections

Suppose we have a right triangle with vertices at the origin, $(1, 0)$ and $(1, 1)$. A reflection would be if we flipped it onto the other side across the y-axis, now giving us the origin, $(-1, 0)$ and $(-1, 1)$. It should be straightforward to see that we can obtain this reflection by taking all points on the original triangle and multiplying their x-coordinate by negative one. This will obviously be given by

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

21.1.3 Rotations

What if we want to rotate the unit square $45°$ using the origin as a pivot?

Instead of using $\begin{bmatrix} x & y \end{bmatrix}^t$ as our vector, what if we use polar coordinates? This should feel normals, as polar makes circular things easier. Thus an arbitrary point on the unit square will have the coordinates

$$\begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \end{bmatrix}.$$

If we want to find the matrix that rotates our vector by α, we are looking for a matrix such that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \end{bmatrix} = \begin{bmatrix} r\cos(\theta + \alpha) \\ r\sin(\theta + \alpha) \end{bmatrix}$$

I really don't want to follow the proof since it's not very enlightening and uses more trigonometric identities than I'm comfortable with, so here's a rough outline instead:

$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Going off of the 3B1B trick, the first column should be fairly straightfoward: the vector $\begin{bmatrix} 1 & 0 \end{bmatrix}^t$ θ "starts" at 0 radians, so rotating it by θ should be given by the x and y coordinates of the vector on the unit circle at angle θ, which naturally come from $\cos(\theta)$ and $\sin(\theta)$. For the second column, which corresponds to the transformed y basis vector, rotating by θ naturally makes the x coordinate *smaller* and starts at 0, so it must be given by $-\sin(\theta)$. Likewise, the y component starts at 1 but decreases, which is $\cos(\theta)$.

Since we have identified a matrix, this is indeed a linear transformation.

21.1.4 Shears

This is the linear algebra equivalent to skewing an object in MS Paint.

22 Day 23: July 10

We went over the following linear transformations:

- Identity transformation – Multiplication by the corresponding identity matrix
- Dialation – Multiplication by a diagonal matrix with all entries greater than zero.
- Reflection – Multiplication by a diagonal matrix with all entries either positive or negative 1.
- Rotation – Multiplication by the matrix $\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$. This is only true in two dimensions though. In general, we have that a rotation matrix A has the property $A \cdot A^T = I$. Any matrix which satifies this property is known as an orthogonal matrix.
- Shear – Multiplication by a triangular matrix.

22.1 Observing Images

Suppose we have the unit square as a set of vectors S and we have the matrix transformation

$$T_A(X) = \begin{bmatrix} \cos(90) & -\sin(90) \\ \sin(90) & \cos(90) \end{bmatrix}.$$

What is $T(S)$?

Since $T_A(X)$ is a counter clockwise rotation by 90 degrees, the image $T(S)$ is given by the unit square rotated into quadrant two.

What about if S is the unit circle?

Circles are infinitely rotationally invariant (very rotationally symmetric), so this is the same thing.

If S is the unit square that touches 1 and -1 on every axis and

$$T_A(X) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

what is $T(S)$?

This corresponds to the basis vector for x being stretched by a factor of 2 and the basis vector for y being stereteched by 3. This gives us a rectangle that is taller than it is wide.

22.2 Composition of Matrix Transformations

We know how to get a circle to an ellipse, but what if we want to get a rotated ellipse (such that the major/minor axes aren't the coordinate axes)? The principle is to identity a matrix which first stretches the circle into the ellipse and then a matrix which rotates it.

Composition of Real Functions

Say we have $f(x) = 2x + 1$ and $g(x) = \sin(x)$. Then $f(g(x)) = 2\sin(x) + 1$ and $g(f(x)) = \sin(2x + 1)$. Importantly, composition is generally not commutative.

Back to Matrices

If we have two matrix multiplications T_A and T_B, we can proceed by similarly defining $(T_A \circ T_B)(X) = T_A(T_B(X))$. What if $T_A(X) = AX$ and $T_B(X) = BX$? We then have that

$$(T_A \circ T_B)(X) = T_A \cdot (BX) = (A \cdot B)X.$$

This follows from the associative law.

This also implies that

$$(T_A \circ T_B) = AB$$

which indeed proves that composition of matrix transformations are themselves matrix transformations and are thus also linear transformations.

"Definition 3.5:" Composition of Matrix Transformations

Let U, V and W be sets and define the following transformations T_A and T_B as follows:

$$T_A : U \to V$$

$$T_B : V \to W$$

Then the composoite of T_B with T_A is the transformation

$$U \to W.$$

This implies that the domain of the composition is given by the first set U and the range of the composition is the final set V.

Find the matrix transformation that strtches a circle to an ellipse with its major axis as the y-axis with a factor of two then rotates it clockwise 45 degrees.

We attack this by using composition. We look for the entire matrix T as a composition of two other matices $T_B \circ T_A$. The first matrix transformation, T_A

is just a dialation by a factor of 2 on the y-axis, which is

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

The second matrix transformation, T_B is a rotation by 45 degrees clockwise, which is given by

$$B = \begin{bmatrix} \cos(-45) & -\sin(45) \\ \sin(-45) & \cos(-45) \end{bmatrix}.$$

By definition, we have that $T(X) = (B \cdot A)X$. I'm not going to bother showing all of the steps here, but note that we can factor out the $\sqrt{2}/2$ in the first matrix. Carrying everything out we have that

$$T(X) = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}.$$

22.3 Orthogonal Matrices

An orthogonal matrix is defined as having the property $A^T \cdot A = I$. This also implies that $A^T = A^{-1}$. This also means that

$$T_{A^T} \circ T_A = T_I.$$

We have that this is true for all rotation matrices since the transpose A^T corresponds to the rotation matrix rotated in the opposite direction. Their composition, which corresponds to rotation a certain number of degrees then rotating it back, is then obviously the identity transofrmation.

In 2 and 3 dimensions (so 2×2 and 3×3 matrices), all orthogonal matrices necessarily correspond to a rotation matrix.

22.4 Inverting Matrix Transformations

"Definition 3.7:" Inverse Linear Transformations

Let

$$T : V \to W$$

be a linear transformation ebtween two vector spaces V and W. We say T is invertible if for all $Y \in W$ there is a unique solution X to the equation $T(X) = Y$. In this case, the inverse transformation $T^{-1} : W \to V$ is dfined by $T^{-1}(Y) = X$ if and only if $T(X) = Y$.

It should be fairly straightforward that $(T^{-1} \circ T(X)) = I(X) = X$.

"Proposition 3.4:" Inverse goes both ways

For a given linear transformation corresponding to a matrix transformation, the linear transformation is invertible if and only if the matrix is invertible.

22.4.1 Properties of Invertible Linear Transformations

- If A is linear, A^{-1} is linear.
- If A is inveritble A^{-1} is also invertible.
- If T_A, T_B are linear transformations, then their composition $T_A \circ T_B$ is also linear.

23 Day 24: July 11

23.1 T/F Section 3.1

3.1 If we know that $A \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 3 \end{bmatrix}$ and likewise for the second equation, we can expand and solve for entries. Then we check if it's possible. Supposedly this is false.

3.2 True. Geometrically, linear transformations can't create more/new lines nor can they stretch them.

3.3 False. Shears.

3.4 False. Projections.

3.5 False. This is a trick question, because the question did not specify *linear* transformations. Also multiplication by the zero matrix collapses $\mathbb{R}^2$ to the origin.

3.6 True. Two linearly independent vectors can form a basis of $\mathbb{R}^2$, so we are able to create any vector in $\mathbb{R}^2$ from those two and use linear transformation properties to compute them.

23.2 T/F Section 3.2

3.12 False. This is not true in general.

3.13 False. Just check sizes.

3.14 True.

3.15 True. Since matrix multiplication is writing down entries of dot products with rows and columns, if one column of B is zero, then every row multiplied by said column is zero, so we have an entire column in the product which is zero.

3.16 True. The rationale from 3.15 applies.

3.17 False. Counterexample is if A is a zero matrix and B is a filled matrix. Note that if A is nonsingular, then B must be the zero matrix. This is because if A is nonsingular (and thus has nullspace of solely the zero element), then any product of A with B that yields zero implies that that column of B belongs in the nullspace and is thus the zero element.

3.18 True. If the columns are linearly dependent, then we have some non-trivial X such that $BX = 0$. So what about AB? Note that if $BX = 0$, then $ABX = 0$, as $A0 = 0$.

3.19 True. Apply the rank theorem with the logic of 3.18

3.20 False. We can have two dimensional collapses. The only thing we know is that all columns of B are in the nullspace of A.

23.3 T/F Section 3.3

3.21 False. Just check if the rank is 4. It's not, since the fourth row is the sum of the first three.

3.22 False. The rank of A and B both cannot be larger than 3. So the rank of the products also can't be larger than 3. Since the resultant matrix is 4×4 and the rank is 3, it is not invertible.

3.23 False. If A is inveritble or nonsingular, then the rank of AB is equal to the rank of B.

3.24 False.

3.25 True. If AB is invertible, it is full rank. Since A and B are both square, they are also full rank and thus inveritble.

3.26 True. Invertible matrices are commutative.

3.27 True. We factor to get $A(A + 7I) = I$, which implies that $(A + 7I) = A^{-1}$.

3.28 True. If something is invertible, it is row equivalent to the identity matrix. Follow this train to B.

3.29 True. Any element in the nullspace of B is in the nullspace of AB, since $BX = 0$ and $A0 = 0$. A doesn't contribute since it's invertible. In general, if A is nonsingular, AB has the same spaces of B.

3.30 False. Multiply A on both sides, apply the associative law, and you get $AB = BA$, which isn't true in general.

24 Day 25: July 12

24.1 Coordinates

We've seen coordinates before:

- Cartesian
- Polar

Generally when working in $\mathbb{R}^n$ we use the first. The cartesian coordinate system has what's known as a "standard basis:"

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

However, this isn't the only option. So long as we have three linearly independent elements from $\mathbb{R}^3$, we can get a basis.

Supposer we have an arbitrary vector

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

which we can represent as

$$\begin{bmatrix}a\\b\\c\end{bmatrix} = a\begin{bmatrix}1\\0\\0\end{bmatrix} + b\begin{bmatrix}0\\1\\0\end{bmatrix} + c\begin{bmatrix}0\\0\\1\end{bmatrix}.$$

Here, we have that a is the x-coordinate and so on.

So we can say in the standard basis, its coordinate vector in the standard basis X' is given by

$$X' = \begin{bmatrix}a\\b\\c\end{bmatrix}$$

which is the same thing. But what if we don't have the standard basis? What if our basis is something like

$$B = \left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\}?$$

Then for a vector $x = \begin{bmatrix}a\\b\\c\end{bmatrix}$ we may have something like

$$X' = \begin{bmatrix}a-b-c\\b-c\\c\end{bmatrix}$$

and $X' \neq X$.

"Definition 3.10:" Ordered basis

This is just some basis where there's an ordering, like

$$B = \{x_1, \ldots, x_n\}$$

where there is an ordering (so swapping elements here produces a new ordered basis).

Given that

$$X = a_1 X_1 + \ldots + a_n X_n$$

where a_i are scalars, we further define

$$X' = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

and say that X' is the B coordinate vector of X.

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

and

$$X = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

find X'.

By inspection,

$$X' = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

and

$$X = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$$

find X'.

By inspection,

$$X' = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}.$$

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

and

$$X = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$$

find X'.

By inspection

$$X' = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}.$$

This demonstrates that changing the order of the elements in the basis can change the coordinate vector.

General Method

When given some vector X and some n-th dimensional basis $B = \{B_1, \ldots B_n\}$, we can find the X' coordinate in B by solving

$$a_1 B_1 + \ldots a_n B_n = X$$

which becomes the matrix

$$\begin{bmatrix} B_1 & \ldots & B_n & X \end{bmatrix}.$$

which is solvable by whatever.

We call the matrix of all the basis vectors

$$P_B = \begin{bmatrix} B_1 & \ldots & B_n \end{bmatrix}$$

the point matrix.

We call it the point matrix since it point matrix multiplied by original vector yields the point vector.

Point Matrices are Invertible

This is because point matrices stem from bases, which are a set of linearly independent vectors which span, which implies full-rank and so they are invertible.

Starting with

$$P_B \cdot X' = X$$

and multiplying both sides by P_B^{-1}, we get

$$X' = P_B^{-1} X.$$

we can also denote P_B^{-1} as the bluecoordinate matrix C_B, as multiplying the original vector by this matrix yields the coordinate vector $C_B \cdot X = X'$.

Why study coordinates? Because sometimes we might have a general vector which doesn't exist in $\mathbb{R}^n$, but we can find its coordinate in $\mathbb{R}^n$, which is a lot easier to study.

24.2 Isomorphisms

"Definition 3.12:" Isomorphisms are Invertible Linear Transofmrations

Let V and W be vector spaces. Then an isomorphism of V onto W is an inveritble linear transformation

$$L : V \to W.$$

In this case, we say V is isomorphic with W.

Some examples:

- Trivially, V is isomorphic with itself. Here, the linear transformation is the identity transformation.
- $X \to X^t$ is an isomorphism. The linear transformation is just transposition.
- Let $P_N(\mathbb{R})$ be the set of all single variable polynomials whose degree is less than or equal to N. P_N is isomorphic to $\mathbb{R}^{N+1}$ since we need $N+1$ total variables to describe all the coefficients. In this case, we have that (for example where N is 3) $L : P_3(\mathbb{R}) \to \mathbb{R}^4$ so

$$L(ax^3 + bx^2 + cx + d) = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

since for each cubic we can find a unique vector and vice versa, it is an isomorphism.

Then, to show that it's linear, we just need to see if

$$L(P(x) + Q(X) = L(P(X)) + L(Q(X))$$

and

$$L(e \cdot P(X)) = e \cdot L(P(X)).$$

We can do all the epxansion, but I don't really want to.

25 Day 26: July 13

We discussed what an isomorphism is:

We say there exists an isomorphism between two vector spaces V and W if there exists an *invertible* linear transformation

$$L : V \to W.$$

We also discussed an isomorphism between the space of all polynomials with maximal degree N $P_N(\mathbb{R})$ to $\mathbb{R}^{N+1}$ by the transformation

$$L\left(a_n x^n + \ldots + a_0\right) = \begin{bmatrix} a_n \\ \vdots \\ a_0 \end{bmatrix}.$$

Also worth mentioning that this is *not* a matrix transformation, since we can't really multiply a polynomial by a matrix here. We then showed that this an isomorphism as

- It is closed with respect to addition and scalar multipliation
- It is one-to-one and is thus invertible.

25.1 Applications of Isomorphisms

"Proposition 3.9"

Suppose that that W is isomorphic with an n-dimensional vector space V. Then

$$\dim(W) = n.$$

So $\dim(P_N(\mathbb{R})) = N + 1$.

This proposition doesn't give us any information about V, though. We know a lot of information about $V = \mathbb{R}^n$, but what if this isn't the vector space we work with?

"Definition 3.11:" Point and coordinate transformations

Let

$$\mathcal{B} = \{x_1, \ldots, x_n\}$$

be an ordered basis for a vector space W. The point transformation with respect to $\mathcal{B}$ is the transformation

$$T_{\mathcal{B}}^{P} : \mathbb{R}^n \to W$$

defined by

$$T_{\mathcal{B}}^{P}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = x_1 X_1 + \ldots x_n X_n.$$

The coordinate transformation is the mapping

$$T_{\mathcal{B}}^{C} := \left(T_{\mathcal{B}}^{P}\right)^{-1} : W \to \mathbb{R}^n$$

that transformats the element $X \in W$ into its coordinate vector $X' \in \mathbb{R}^n$ with respect to $\mathcal{B}$.

This tells us that for any and all vector spaces W, we have the coordinate transformation to $\mathbb{R}^n$ and the point transformation backwards.

any n-dimensional vector space W is isomorphic to $\mathbb{R}^n$.

Suppose we want to transform some vector space V to W through some transformation L. This can be hard, espicially if they are different dimensions.

$$\begin{tikzcd} V \arrow[r, "L"] \arrow[d, "T^C_{\mathcal{B}_V}"'] & W \\ \mathbb{R}^n \arrow[r, "M"'] & \mathbb{R}^m \arrow[u, "T^P_{\mathcal{B}_W}"'] \end{tikzcd}$$

We proceed by doing the following:

- Convert the vector space V to $\mathbb{R}^n$ through a coordinate transformation $T_{\mathcal{B}_V}^{C}$.
- Convert $\mathbb{R}^n$ to $\mathbb{R}^m$ through a matrix transformation M.
- Convert the vector space $\mathbb{R}^m$ to W through a point transformation $T_{\mathcal{B}_W}^{P}$.

If we let the first element X be in V and the second element Y be its corresponding element in W we have this very gruesome expression

$$T_{\mathcal{B}_W}^{P}\left(M \cdot T_{\mathcal{B}_V}^{C}(X)\right) = Y.$$

This is further encapsulated in

"Theorem 3.14:"

Let V and W be finite dimensional vector spaces where $\dim(V) = n$ and $\dim(W) = m$. Let $L : V \to W$ be a linear transformation. Let $\mathcal{B}_V$ and $\mathcal{B}_W$ be ordered bases for V ad W respectively. THen there exists a unique $m \times n$ matrix M such that

$$L(X)' = M \cdot X'$$

$$Y' = M \cdot X'.$$

This means we can reprsent the transformation L as a single matrix multiplication by a special matrix M.

"Theorem 3.14"

Let $A \in M(m, n)$ and let $\mathcal{B}$ and $\bar{\mathcal{B}}$ be ordered bases for $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Then the matrix of

$$T_A : \mathbb{R}^n \to \mathbb{R}^m$$

with respect to the bases $\mathcal{B}$ and $\bar{\mathcal{B}}$ is

$$M = C_{\bar{\mathcal{B}}} \cdot AP_{\mathcal{B}}.$$

which proves the matrix M in the diagram below

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "A"] \arrow[d, "\mathcal{B}"'] & \mathbb{R}^m \\
\mathbb{R}^n \arrow[r, "M"'] & \mathbb{R}^m \arrow[u, "\bar{\mathcal{B}}"']
\end{tikzcd}$$

always exists.

Suppose we have

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 7 \\ -11 \end{bmatrix} \right\},$$

$$\bar{\mathcal{B}} = \left\{ \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \end{bmatrix} \right\}$$

with

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

so

$$T_A : \mathbb{R}^3 \to \mathbb{R}^2.$$

Find M.

Note that
$$M = C_{\bar{\mathcal{B}}} \cdot A \cdot P_{\mathcal{B}},$$
so we need to find everything.
$$P_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 7 \\ 0 & 0 & -11 \end{bmatrix}.$$
$$C_{\bar{\mathcal{B}}} = \begin{bmatrix} 2 & 1 \\ -3 & 4 \end{bmatrix}^{-1} = \frac{1}{11}\begin{bmatrix} 4 & -1 \\ 3 & 2 \end{bmatrix}.$$
Then we just multiply everything out.

Note that the matrix transforamation determined by A is the transformation given in the standard bases, but M is the transformation given by the bases $\mathcal{B}$ and $\bar{\mathcal{B}}$.

Linear transformations correspond to a changing of basis.

25.2 Preview of Determinants

We define the determinant to be a function on square matrices that assigns it a real number, which tells us some neat things.

26 Day 27: July 16

We define the determinant as a function such that takes in an arbitrary square matrix and yields a real number:
$$D : M(n, n) \to \mathbb{R}.$$

26.1 Determinants

For a 1×1, it's itself, since 1×1 matrices are already real numbers, so we don't have anything to do.

For a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we define it as
$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$
So if $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, then $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 4 - 6 = -2$.

If $A = \begin{bmatrix} 2 & 4 \\ 3 & 6 \end{bmatrix}$, then $\det(A) = 12 - 12 = 0$.

What about the 3×3 determinant? We do it recursively.

The determinant for a 3×3 matrix is a little more involved and requires a bit of recursion. This generalizes to arbitrary dimensions.

The procedure is as follows. Take the first entry in the top row, and then delete all entries on the same row and column of it. Multiply the first entry with the determinant of the remaining entries not including the first entry. Then, repeat this process with the second entry in the first row but negative. Flip signs each time.

$$\begin{aligned}\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} &= a_1 \begin{vmatrix} \square & \square & \square \\ \square & b_2 & c_2 \\ \square & b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} \square & \square & \square \\ a_2 & \square & c_2 \\ a_3 & \square & c_3 \end{vmatrix} + c_1 \begin{vmatrix} \square & \square & \square \\ a_2 & b_2 & \square \\ a_3 & b_3 & \square \end{vmatrix} \\ &= a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}.\end{aligned}$$

There exists an alternative way:

- Adjoin a copy of the matrix horizontally to the matrix.
- Draw down-right diagonal lines (these must be through 3 entries). Sum the products on these lines.
- Draw down-left diagonal lines. Sum the products on these lines.
- Subtract the second from the first.

Compute

$$\begin{vmatrix} 2 & 0 & 1 \\ 6 & 1 & 7 \\ 4 & 5 & 9 \end{vmatrix}.$$

This gives us $(18 + 0 + 30) - (70 + 0 + 4) = 48 - 74 = -26$.

Determinants are extremely slow to compute and are $O(n!)$. So why do we define it using such a complex recursive process? The answer is because its deifned in geoemtrically. In 2 dimensions, the determinant of a matrix represents the area; in 3, the volume; in 4, the 4-dimensional hypervolume.

For a given $n \times n$ matrix, the determinant is defined as the n-dimensional volume spanned by the n vectors in n-dimensional space by their parallelepepied.

26.2 Determinant Properties

"Theorem 4.1:" Row interchange property

Let $A \in M(n, n)$. Suppose B is obtained by inerchanging two rows of A. Then

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

"Theorem 4.3:" Row scalar property

Suppose the $n \times n$ matrix B is obtained from A by multiplying each element in the ith row by some scalar c. Then

$$\det(B) = c \cdot \det(A).$$

"Theorem 4.6:" Row additive property

Let U, V and A_i be $1 \times n$ row vectors, where $i = 2, 3, \ldots, n$. Then

$$\det \begin{bmatrix} U + V \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} U \\ \vdots \\ A_n \end{bmatrix} + \begin{bmatrix} V \\ \vdots \\ A_n \end{bmatrix}.$$

$$\det(a + b)$$

does not have a closed form expression in general, and it is **NOT**

$$\det(a) + \det(b).$$

27 The rest of the course

I starting taking notes in OneNote at this point. Please beware! THe nice formatting and explanations go dramatically downhill.

Friday, July 19, 2019 10:59 AM

$$A^{-1}_{ij} = \frac{(-1)^{i+j} \det(A_{ij})}{\det(A)}$$

$\det(A)$ AFTER DELETING ROW i AND COL j.

SHIT'S MAD EASY FOR 2x2.

$$A^{-1}_{11} = \frac{A_{22}}{\det(A)} \qquad A^{-1}_{12} = \frac{-A_{21}}{\det(A)}$$

$$A^{-1}_{21} = \frac{-A_{12}}{\det(A)} \qquad A^{-1}_{22} = \frac{A_{11}}{\det(A)}.$$

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} A_{22} & -A_{21} \\ -A_{12} & A_{11} \end{bmatrix}$$

$$\boxed{\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \frac{1}{AD-BC} \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}}$$

IF THE MATRIX IS NOT INVERTIBLE, THIS IS NOT WELL-DEFINED.

3x3 IS "NOT TOO BAD."

Monday, July 22, 2019 11:03 AM

X IS A λ-EIGENVECTOR FOR A IF

$$AX = \lambda X.$$

X: NONZERO, $X \in \mathbb{R}^n$.

λ: SCALAR, $\lambda \in \mathbb{R}$.

A: MATRIX, $A \in M(m,n)$

MOST VECTORS AREN'T E-VECTORS.

TO FIND,

$$\det(A - \lambda I) = 0.$$

$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $\lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$, SO

$$A - \lambda I = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix}$$

$$\underbrace{\det(A - \lambda I) = 0}_{\text{"CHARACTERISTIC POLYNOMIAL"}} \Rightarrow (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0.$$

$$\underbrace{\lambda^2 - (a_{11} + a_{22})\lambda - (a_{11}a_{22} - a_{12}a_{21}) = 0}_{\text{WE CAN SOLVE FOR } \lambda}$$

(1) $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$, FIND E-VALUES / E-VECTORS

$$A - \lambda I = \begin{bmatrix} 3 - \lambda & 2 \\ 2 & 3 - \lambda \end{bmatrix}$$

$$\det(A - \lambda I) = 0 \Leftrightarrow (3 - \lambda)(3 - \lambda) - 2 \times 2 = 0$$

$$\lambda^2 - 6\lambda + 9 - 4 = 0$$

$$\lambda^2 - 6\lambda + 5 = 0$$

$$(\lambda - 5)(\lambda - 1) = 0$$

$\boxed{\lambda = 1, 5.}$ ← THESE MAKE $A - \lambda I$ SINGULAR.

E-VECTORS?

COME BACK TO

$(A - \lambda I)X = 0$ FOR AN E-VECTOR X. $X \in \text{Nullspace}(A)$

$\lambda = 1 \Rightarrow A - 1I = \begin{bmatrix} 3-1 & 2 \\ 2 & 3-1 \end{bmatrix}$
$= \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$

FIND NULLSPACE

$\left[\begin{array}{cc|c} 2 & 2 & 0 \\ 2 & 2 & 0 \end{array}\right] \Rightarrow \begin{array}{ccc} \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \\ x_1 \quad x_2 \quad \vec{0} \end{array} \Rightarrow x_1 + x_2 = 0 \Rightarrow x_1 = -x_2$

$\Rightarrow X = \begin{bmatrix} -s \\ s \end{bmatrix} = s \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$\lambda = 5 \Rightarrow A - 5I = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$

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

$\Rightarrow x_1 - x_2 = 0 \Rightarrow X = \begin{bmatrix} s \\ s \end{bmatrix} = s \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$x_1 = x_2$

$\therefore$ 1-EIGENSPACE: $\text{span}\left(\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}\right)$

5-EIGENSPACE: $\text{span}\left(\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}\right)$

(2) $A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix}$

$$A-\lambda I \quad \begin{bmatrix} 3-\lambda & 1 & 0 \\ 0 & 1-\lambda & 0 \\ 4 & 2 & 1-\lambda \end{bmatrix}$$

$$\begin{aligned} \det(A-\lambda I) &= (3-\lambda)(1-\lambda)(1-\lambda) + (1\cdot 0 \cdot 4) + (0\cdot 0\cdot 2) \\ &\quad - ((3-\lambda)\cdot 0\cdot 2) - ((1\cdot 0\cdot (1-\lambda)) - (0\cdot (1-\lambda)\cdot 4) \\ &= (3-\lambda)(1-\lambda)(1-\lambda) \end{aligned}$$

$\therefore \lambda = 1, 3$

$\lambda = 3 \quad A-3I = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 4 & 2 & -2 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & -1 & 0 \end{bmatrix} \Rightarrow$ $x_2 = 0$, $2x_1 - x_3 = 0 \Rightarrow x_1 = \frac{1}{2}x_3$

$\Rightarrow X = \begin{bmatrix} \frac{1}{2}s \\ 0 \\ s \end{bmatrix} = s\begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}$

$\lambda = 1$

$A-I \quad \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 4 & 2 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ $2x_1 + x_2 = 0 \Rightarrow x_1 = -\frac{1}{2}x_2$

x_3 IS FREE.

$\Rightarrow X = s\begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

THE MULTIPLICITY OF AN E-VALUE λ_0 IS THE NUMBER OF TIMES $(\lambda - \lambda_0)$ APPEARS IN THE CHARACTERISTIC POLYNOMIAL.

MATRIX SIZE (n) = NUMBER OF EIGENVALUES + MULTIPLICITY.
$\geq$ NUMBER OF LIN-IN E-VECS.

(3) $A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 5 \end{bmatrix}$

$$A-\lambda I = \begin{bmatrix} 3-\lambda & 1 & 1 \\ 0 & 3-\lambda & 1 \\ 0 & 0 & 5-\lambda \end{bmatrix}$$

$$\det(A-\lambda I) = (3-\lambda)^2(5-\lambda)$$

Tuesday, July 23, 2019 10:57 AM

THEOREM 5.1

LET $A \in M(n,n)$

THEN λ IS AN EIGENVALUE FOR A IFF

$$p(\lambda) := \det(A - I\lambda) = 0$$

THE FUNCTION $p(\lambda)$ IS A POLYNOMIAL WITH DEGREE n WITH LEADING COEFFICIENT $(-1)^n$ AND IS CALLED THE CHARACTERISTIC POLYNOMIAL FOR A.

WE FIND E-Vs BY SOLVING $\det(A - I\lambda) = 0$.
ALL SOLUTIONS λ ARE EIGENVALUES.

TO FIND EIGENVECTORS, SOLVE FOR $(A - I\lambda)X = 0$ FOR EACH λ.

(2)

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

INCL. MULTIPLICITY

$\underline{n=3}$, HAS 3 EIGENVALUES AND $\underline{3}$ <u>REPRESENTATIVE EIGENVECTORS</u>

$\therefore$ DIAGONALIZABLE

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 5 \end{bmatrix}$$

$\underline{n=3}$

3 EIGENVALUES

2 REP E-VECS

3 EIGENVALUES
2 REP E-VECS.
∴ DEFICIENT.

THEOREM 5.2
LET $p(\lambda)$ BE THE CHARACTERISTIC POLYNOMIAL FOR $A \in M(n,n)$ and LET $\underbrace{\lambda_1, \ldots, \lambda_k}_{\text{DISTINCT.}}$ BE THE ROOTS OF $p(\lambda)$. THEN
THE DIMENSION OF THE λ_i-EIGENSPACE IS AT MOST THE MULTIPLICITY (m_i) OF λ_i AS A ROOT OF $p(\lambda)$.

A IS DEFICIENT IF

FOR ANY i, $\overbrace{\#\text{ OF } \lambda_i\text{-EIGENVECTORS}}^{d_i < m_i} < m_i$

IF WE HAVE, FOR A PARTICULAR EIGENVALUE λ_i FEWER EIGENVECTORS THAN THE MULTIPLICITY OF λ_i, THE MATRIX IS DEFICIENT.

λ_i-EIGENSPACE $:= \{x_1, \ldots, x_{d_i}\}$

$$1 \le d_i \le m_i$$

WE ALWAYS HAVE 1 EIGEN VECTOR

NUMBER OF EIGENVECTORS,

MULTIPLICITY OF λ_i.

(1)
LET $A \in M(n,n)$. A HAS n DISTINCT EIGENVALUES. THEN A IS DIAGONALIZABLE
PROOF IF IT HAS n DISTINCT THE MULTIPLICITIES

THEN A IS DIAGONALIZABLE

PROOF. IF IT HAS n DISTINCT, THE MULTIPLICITIES ARE ALL 1, SO $1 \le d_i \le 1$ FOR EACH EIGENVALUE. AS $d_i = m$ FOR ALL λ_i, IT IS DIAGONALIZABLE.

(2)

A IS SINGULAR IFF $\lambda = 0$ IS AN EIGENVALUE

A IS NONSINGULAR IFF $\lambda = 0$ IS NOT AN EIGENVALUE

EXAMPLE 3

$$p(\lambda) = \lambda^2 (1-\lambda)(3-\lambda)^3$$

$\lambda_1 = 0 \quad m_1 = 2$

$\lambda_2 = 1 \quad m_2 = 1$

$\lambda_3 = 3 \quad m_3 = 3$

(1) A IS NOT INVERTIBLE AS 0 IS AN EIGENVALUE.

(2) Null(A)?

$\boxed{\text{NOTE} \quad \text{NULLSPACE} := 0\text{-EIGENSPACE}}$

SO WE HAVE

$$1 \le \text{null}(A) \le 2$$

$\text{null}(A) = \{1, 2\}$

(3) Rank(A)?

$n = 6$, AS WE HAVE 6 ROOTS (WITH MULTIPLICITY).

null = 1 => Rank = $\boxed{5}$
= 2 $\quad$ 4

(3) THE CONSTANT TERM OF $p(\lambda) = \det(A)$.

THIS PROVES THE SINGULARITY/INVERTIBILITY CONDITION.

- IF λ^k IS A FACTOR, THERE EXISTS NO CONSTANT TERM
- THUS $\det(A) = 0$
- YADA YADA

APPLICATIONS

COMPUTING $A^n X$, LIKE $A^{100} X$ IS HARD
BUT IF $AX = \lambda X$, THEN

$$A^2 X = \lambda^2 X$$
$$A^n X = \lambda^n X.$$

BUT WHAT IF X ISN'T AN EIGENVECTOR?
WRITE X AS A LINEAR COMBINATION OF EIGENVECTORS.

(1) $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$ $\quad \lambda_1 = 5,\ X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
$\quad \lambda_2 = 1,\ X_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

FIND $A^{100} \begin{bmatrix} 4 \\ -2 \end{bmatrix}$.

REWRITE $\begin{bmatrix} 4 \\ -2 \end{bmatrix}$ AS LINEAR COMBO OF X_1, X_2.

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

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

$$\begin{bmatrix} \\ -2 \end{bmatrix} = 1\begin{bmatrix} \\ 1 \end{bmatrix} + 3\begin{bmatrix} \\ -1 \end{bmatrix}$$

$$\begin{aligned}
A^{100}\begin{bmatrix} 4 \\ -2 \end{bmatrix} &= A^{100}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 1 \\ -1 \end{bmatrix} \right) \\
&= \left(A^{100}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) + \left(A^{100}\, 3\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) \\
&= 5^{100}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3 \cdot A^{100}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \\
&= 5^{100}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3 \cdot 1^{100}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} 5^{100} + 3 \\ 5^{100} - 3 \end{bmatrix}
\end{aligned}$$

Wednesday, July 24, 2019 11:02 AM

APPLICATIONS PART 2.

(1) $A_1 = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix}$ $A_2 = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 5 \end{bmatrix}$

$\lambda_1 = 1$

$m_1 = 2$

$X_1 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$

$X_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

$\lambda_2 = 3$

$m_2 = 1$

$X_3 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$

$\{X_1, X_2, X_3\}$ IS LINEARLY INDEPENDENT.

THEOREM 5.5

LET $A \in M(n,n)$ AND LET $Q_1, \ldots, Q_k$ BE REPRESENTATIVE EIGENVECTORS FOR EIGENVALUES λ_i. THEN

$$\{Q_1, \ldots, Q_k\}$$

IS LINEARLY INDEPENDENT.

$\{X_1, X_2, \ldots, X_k\}$ IS LINEARLY INDEPENDENT AND THERE ARE n OF THEM, THEN A IS DIAGONALIZABLE AND

$$\{X_1, \ldots, X_n\}$$

FORMS AN EIGEN BASIS FOR $\mathbb{R}^n$ AND IS

IF WE HAVE AN n-DIMENSIONAL EIGEN BASIS, WE CAN EASILY

IF WE HAVE AN n-DIMENSIONAL EIGEN BASIS, WE CAN EASILY COMPUTE $A^n \cdot b$.

(2)

$$B = \begin{bmatrix} 2 \\ 2 \\ 8 \end{bmatrix}$$

$A^n B$?

FIRST, REWRITE B AS A LIN-COM OF E-VECS.

$$\begin{bmatrix} 2 \\ 2 \\ 8 \end{bmatrix} = a\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + c\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

$$= 1\ x_1 + 2x_2 + 3x_3$$

$$\begin{aligned}
A^{20}B &= A^{20}(x_1 + 2x_2 + 3x_3) \\
&= A^{20}x_1 + A^{20}2x_2 + A^{20}3x_3 \\
&= A^{20}x_1 + 2A^{20}x_2 + 3A^{20}x_3 \\
&= \lambda_1^{20}x_1 + 2\lambda_2^{20}x_2 + 3\lambda_3^{20}x_3 \\
&= 1^{20} \cdot \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + 2 \cdot 1^{20}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + 3 \cdot 3^{20}\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \\
&= \begin{bmatrix} 3^{21} - 1 \\ 2 \\ 2 + 2 \cdot 3^{21} \end{bmatrix}
\end{aligned}$$

thicc

IF A IS DEFICIENT, IT'S NOT GUARANTEED WE CAN DO THIS FOR ARBITRARY B.

IF A IS DIAGONALIZABLE

$$M = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

IF THERE EXISTS AN INVERTIBLE P SUCH THAT

$$A = PMP^{-1}$$

THEN A AND M ARE SIMILAR.

IF A IS DIAGONALIZABLE, IT IS SIMILAR TO A DIAGONAL MAP.

T/F POG

Tuesday, July 23, 2019 9:11 PM

3.31 Let A be an $m \times n$ matrix and let M be the matrix of T_A with respect to bases $\mathcal{B}$ of $\mathbb{R}^m$ and $\overline{\mathcal{B}}$ of $\mathbb{R}^n$. Then rank A = rank M. [*Hint:* Consider formula (3.36).]

3.32 Let $\mathcal{B}$ and $\overline{\mathcal{B}}$ be ordered bases for $\mathbb{R}^n$. Then the matrix of the identity transformation of $\mathbb{R}^n$ into itself with respect to $\mathcal{B}$ and $\overline{\mathcal{B}}$ is the $n \times n$ identity matrix I.

3.33 Let $\mathcal{B}$ and $\overline{\mathcal{B}}$ be ordered bases for $\mathbb{R}^n$ where $\mathcal{B} = \overline{\mathcal{B}}$. Then the matrix of the identity transformation of $\mathbb{R}^n$ into itself with respect to $\mathcal{B}$ and $\overline{\mathcal{B}}$ is the $n \times n$ identity matrix I.

3.34 Let A be an $m \times n$ matrix and let M be the matrix of T_A with respect to bases $\mathcal{B}$ and $\overline{\mathcal{B}}$. Then the dimensions of the column spaces of A and M are equal.

3.35 Let A be an $m \times n$ matrix and let M be the matrix of T_A with respect to bases $\mathcal{B}$ and $\overline{\mathcal{B}}$. Then the the column spaces of A and M are equal.

3.31 True.

$$M = C_{\overline{B}} A P_B$$

INVERTIBLE, SO MULT BY THEM DOES NOT CHANGE RANK

$\therefore$ Rank M = Rank A

3.32 False. Not necessarily true, as

$$M = C_{\overline{B}} A P_B$$

IDENTITY MATRIX

$$= P_{\overline{B}}^{-1} A P_B$$

Is only true if B bar is B.

3.33 True, as this satisfies the condition above.

3.34 Let A be an $m \times n$ matrix and let M be the matrix of T_A with respect to bases $\mathcal{B}$ and $\overline{\mathcal{B}}$. Then the dimensions of the column spaces of A and M are equal.

True. Dimension of the column space is the rank, and they have the same rank by 3.31

3.35 Let A be an $m \times n$ matrix and let M be the matrix of T_A with respect to bases B and $\overline{B}$. Then the the column spaces of A and M are equal.

False. Even though the dimensions are equal, they are not the same spaces, as A = C_B bar M P_B, and so unless multiplying by C_B bar and M_B does nothing which they generally will not, the two matrices won't have the same space.

3.36 Let A be an $m \times n$ matrix and let M be the matrix of T_A with respect to bases B and $\overline{B}$. Then the dimensions of the row spaces of A and M are equal.

True, as the dimension of the row space is the rank, and we know that the ranks are equal.

3.37 Let A be an $m \times n$ matrix and let M be the matrix of T_A with respect to bases B and $\overline{B}$. Then the row spaces of A and M are equal.

False, by the logic of 3.35

3.38 Let A be an $m \times n$ matrix and let M be the matrix of T_A with respect to bases B and $\overline{B}$. Then the dimensions of the nullspaces of A and M are equal.

True, they have the same rank, so they have the same nullity by the rank-nullity theorem.

3.39 Let A be an $m \times n$ matrix and let M be the matrix of T_A with respect to bases B and $\overline{B}$. Then A and M have the same nullspace.

False, by the logic of 3.35

4.1 The following matrices have the same determinant. [*Hint:* Factor some scalars out from the determinants.]

$$\begin{bmatrix} 2 & 4 & 2 & 6 \\ 3 & 3 & 27 & 33 \\ 2 & 1 & 5 & 2 \\ 6 & 1 & -3 & 3 \end{bmatrix} \qquad \begin{bmatrix} 3 & 6 & 3 & 9 \\ 2 & 1 & 5 & 2 \\ 6 & 1 & -3 & 3 \\ 2 & 2 & 18 & 22 \end{bmatrix}$$

True. Just spam some determinant operations.

4.2 Let A be a 3×3 matrix. Then $\det(5A) = 5\det(A)$.

False. If you multiply ONE ROW by 5, the determinant is scaled by 5, but 5A scales 3 rows.

As an aside, just consider the determinant of an nxn matrix analogous to

$f(x) = x^n$

So det (5x) here would be 5^3 x^n. And also f(x + y) isn't just x^n +y^n.

4.3 Let A and B be 3×3 matrices. Then $\det(A + B) = \det(A) + \det(B)$.

False. Consider the above reasoning, but this also isn't a real formula.

4.4 The following statement is true:

$$\begin{vmatrix} 2 & 4 & 2 & 6 \\ 3 & 3 & 27 & 33 \\ 2 & 1 & 5 & 2 \\ 6 & 1 & -3 & 3 \end{vmatrix} = 2 \begin{vmatrix} 2 & 4 & 2 & 6 \\ 1 & 1 & 2 & 0 \\ 2 & 1 & 5 & 2 \\ 6 & 1 & -3 & 3 \end{vmatrix} + \begin{vmatrix} 2 & 4 & 2 & 6 \\ 1 & 1 & 23 & 33 \\ 2 & 1 & 5 & 2 \\ 6 & 1 & -3 & 3 \end{vmatrix}$$

Spam theorems.

4.5 The following statement is true:

$$\begin{vmatrix} 2 & 4 & 2 & 6 \\ 3 & 3 & 7 & 8 \\ 2 & 1 & 5 & 2 \\ 6 & 1 & -3 & 3 \end{vmatrix} = 4 \begin{vmatrix} 3 & 7 & 8 \\ 2 & 5 & 2 \\ 6 & -3 & 3 \end{vmatrix} - 3 \begin{vmatrix} 2 & 2 & 6 \\ 2 & 5 & 2 \\ 6 & -3 & 3 \end{vmatrix} + \begin{vmatrix} 2 & 2 & 6 \\ 3 & 7 & 8 \\ 6 & -3 & 3 \end{vmatrix} - \begin{vmatrix} 2 & 2 & 6 \\ 3 & 7 & 8 \\ 2 & 5 & 2 \end{vmatrix}$$

False. You also spam theorems here, but it's a bit more subtle. Note that det(A^t) = det(A), switch two rows, then transpose it back. You'll be off by a sign.

4.6 The following matrices have the same determinant:

$$\begin{bmatrix} 1753 & 0 & 0 & 0 \\ 27 & 33 & 0 & 0 \\ 13 & 911 & 1411 & 0 \\ -15 & 44 & 32 & 1001 \end{bmatrix}, \quad \begin{bmatrix} 1753 & 27 & 13 & -15 \\ 0 & 33 & 911 & 44 \\ 0 & 0 & 1411 & 32 \\ 0 & 0 & 0 & 1001 \end{bmatrix}$$

True. If they're triangular, you can just multiply it across their diagonals.

4.7 The following matrices have the same determinant:

$$\begin{bmatrix} 1753 & 0 & 0 & 0 \\ 27 & 33 & 0 & 0 \\ 13 & 911 & 1411 & 0 \\ -15 & 44 & 32 & 1001 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 & 1753 \\ 0 & 0 & 33 & 27 \\ 0 & 1411 & 911 & 13 \\ 1001 & 32 & 44 & -15 \end{bmatrix}$$

True. This is a bit more subtle. The second matrix actually isn't a triangular matrix (you need either the upper right or bottom left to be zero entries). You can transform the second matrix INTO a triangular matrix by exchanging rows into an upper triangular matrix. This takes two swaps, so the sign doesn't change.

4.8 Given that none of the determinants are 0, the following matrices have the same determinant:

(c) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 4 & 4 \\ -1 & 2 & -1 \end{bmatrix}, \quad 4\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ -1 & 2 & -1 \end{bmatrix}$

(d) $\begin{vmatrix} 7 & -3 & 4 & -2 \\ 13 & 6 & 5 & 11 \\ 1 & 2 & 1 & 0 \\ 17 & 16 & 5 & 6 \end{vmatrix}, \quad \begin{vmatrix} 7 & -3 & 4 & -2 \\ 1 & 2 & 1 & 0 \\ 10 & 0 & 2 & 11 \\ 17 & 16 & 5 & 6 \end{vmatrix}$

(e) $\begin{bmatrix} 2 & 3 & -5 & 3 \\ 7 & 1 & 2 & 3 \\ 1 & 4 & 4 & 1 \\ 3 & 2 & 4 & 5 \end{bmatrix}, \quad \begin{bmatrix} 2 & 7 & 1 & 3 \\ 3 & 1 & 4 & 2 \\ -5 & 2 & 4 & 4 \\ 3 & 3 & 1 & 5 \end{bmatrix}$

False. This is a trick question. The second matrix is 4 times the entire matrix, not just times the second row. If we replaced the brackets with determinant vertical lines, they would then be equal.

False. The third row in the second matrix is a linear combination, but there's also a row swap, so no.

True. This is just the transpose.

4.9 For all $n \times n$ matrices A and B, $\det(AB) = \det(BA)$.

True. Even though AB != BA in general, det(AB) = det A * det B and the rest is pretty fast to prove.

4.10 Suppose that $\det(A + I) = 3$, and $\det(A - I) = 5$. Then

$$\det(A^2 - I) = 20$$

False. det AB = det a * det b, so this should be 15.

4.11 Suppose that $\det A = 2$, $\det(A + I) = 3$, and $\det(A + 2I) = 5$. Then

$$\det(A^4 + 3A^3 + 2A^2) = 48$$

$\det(A^2(A^2+3A+2I))$

$\det(A^2(A+2I)(A+I))$

$\det A \cdot \det A \cdot \det A+2I \cdot \det A+I$

$2 \cdot 2 \cdot 3 \cdot 5$

$\neq 48.$

As an aside most $(A^2-B^2) \neq (A+B)(A-B).$
Factorizations across two
Matrices generally don't work because they rely on middle terms cancelling which isn't

Matrices generally don't work because they rely on middle terms cancelling which isn't always guaranteed due to anticommutativity.

4.12 In the following system, you may assume that the determinant of the coefficient matrix is not zero. Then, $z = 0$. [*Hint:* Compute the sum of the first two columns of the coefficient matrix.]

$$\begin{aligned} 2x + 3y - 5z + 3w &= 5 \\ 7x + y + 2z + 3w &= 8 \\ x + 4y + 4z + w &= 5 \\ 3x + 2y + 4z + 5w &= 5 \end{aligned}$$

4.13 If the following matrix is invertible, then the $(1, 4)$ entry of A^{-1} is zero. [*Hint:* Compute the sum of the first two rows of A.]

$$A = \begin{bmatrix} 1 & 2 & 1 & 13 \\ 2 & 3 & -5 & -1 \\ 0 & 5 & -4 & 12 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Use Cramer's rule.

5.1 If A is an $n \times n$ matrix that has zero for an eigenvalue, then A cannot be invertible.

True. If we have a zero eigenvalue, it cannot be invertible.

5.2 There is no 3×3 matrix A with $p_A(\lambda) = (\lambda - 2)^2(\lambda - 3)$.

False. The leading term of a characteristic polynomial must be -1 to the power of the degree (since factors are of the form k - lambda). This is a trick question since it's of the form lambda - k.

5.3 If A is a 3×3 matrix with $p_A(\lambda) = (2 - \lambda)^2(3 - \lambda)$ then:

(a) There is a nonzero vector X such that $AX = 2X$.

(b) Then there is at most one nonzero vector X such that $AX = 3X$.

(c) There are two linearly independent vectors X_1 and X_2 such that $AX_i = 2X_i$.

A is true. We always have at least one representative eigenvector.
B is false. Scalar multiples of representative eigenvectors are also eigenvectors. There is only one linearly independent X for the condition, though.
C is false. The matrix can be deficient.

5.4 The sum of two eigenvectors is an eigenvector.

False. This is only true if the two eigenvectors belong to the same eigenspace.

5.5 If X is an eigenvector for A with eigenvalue 3, then $2X$ is an eigenvector for A with eigenvalue 6.

False.

with eigenvalue 6.

False.

$$A(2x) = 2 \cdot Ax$$
$$= 2 \cdot 3x$$
$$= 6x$$
$$= 3(2x)$$

$\lambda = 3$,

5.6 If X is an eigenvector for an $n \times n$ matrix A, then X is also an eigenvector for $2A$.

True. Note though the eigenvalue changes.

$$(2A)X = 2 \cdot AX$$
$$= 2 \cdot \lambda X$$
$$= 2\lambda \cdot X$$

EIGENVALUE FOR $2A$.

5.7 If 3 is an eigenvalue for A, then 9 is an eigenvalue for A^2.

True.

$$AX = 3X$$
$$A^2X = A \cdot AX$$
$$= A \cdot 3x$$
$$= 3 \cdot 3x$$
$$= 9x$$

EIGENVALUE FOR A^2.

5.8 Suppose that A is a 3×3 matrix with 2 and 3 as its only eigenvalues. Then A is deficient.

False. One of these can have multiplicity 2 and then have 2 eigenvectors with the other having multiplicity 1 and have 1 eigenvector. Also consider

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

Which is pretty obviously a diagonal matrix.

5.9 The only nondeficient 3×3 matrix that has 1 as its only eigenvalue is the identity matrix. [*Hint:* What is the dimension of the $\lambda = 1$ eigenspace in this context.]

True.

5.10 Suppose that $p_A(\lambda) = -\lambda^3(\lambda - 2)(\lambda + 3)^2$. Then, the nullspace of A is at most two-dimensional.

False. The nullspace is the zero eigenspace, and it can be up to 3 here.

5.11 There is a 3×3 matrix with eigenvalues 1, 2, 3, and 4.

False. We can't have more eigenvalues (including multiplicity) than we have dimension of the matrix.

5.12 Suppose that A is a 3×3 matrix with eigenvalues 2, 3, and 4. Then $\det A = 9$.

False. We have that for a given matrix that is diagonalizable,

$$\det(A) = \det(PDP^{-1})$$
$$= \det(D)$$

WHERE $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_3 \end{bmatrix}$

$\therefore \prod \lambda = \det(A)$.

ALTERNATIVELY, THE CONSTANT OF $p_A(\lambda) = \det(A)$.

FOR A DEFICIENT MATRIX A, WE HAVE

FOR A DEFICIENT MATRIX A, WE HAVE

$$A = PJP^{-1}$$

WHERE J IS AN UPPER TRIANGULAR MATRIX, WHERE

$$\underbrace{\begin{bmatrix} \lambda_1 & \partial_1 & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}}_{\text{JORDAN FORM}}$$

5.12 Suppose that A is a 3×3 matrix with eigenvalues 2, 3, and 4. Then $\det A = 9$.

5.13 Suppose that A is a 3×3 diagonalizable matrix such that A^2 has eigenvalues 1, 4, and 9. Then A has eigenvalues 1, 2, and 3.

False. Don't forget that A can have eigen values NEGATIVE one, two and three.

5.14 There are at least two 2×2 matrices, with eigenvectors $Q_1 = [1, 2]^t$ and $Q_2 = [-3, 2]^t$ corresponding to the respective eigenvalues $\lambda_1 = -2$ and $\lambda_2 = 0$.

5.15 There is no matrix A that has the vectors $X_1 = [1, 1, 1]^t$, $X_2 = [1, 0, 1]^t$, and $X_3 = [2, 1, 2]^t$ as eigenvectors corresponding, respectively, to the eigenvalues 1, 2, and 3.

True. X_3 = X_1 + X_2, so it can't have its own eigenvalue. Eigenvectors for different eigenvalues must be linearly independent. So the matrix doesn't exist.

5.16 There is no matrix A that has the vectors $X_1 = [1, 1, 1]^t$, $X_2 = [1, 0, 1]^t$, and $X_3 = [2, 1, 3]^t$ as eigenvectors corresponding, respectively, to the eigenvalues 1, 2, and 3.

False. They're linearly independent, so it's ok. To find such a matrix A, just solve

$$\underbrace{\begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 1 & & \\ & 2 & \\ & & 3 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix}^{-1}}_{P^{-1}}$$

EIGENVALUES

CHAPTER 4 REVIEW

Friday, July 26, 2019 11:31 AM

1. DEF $\det(A)$

$$a\left|\quad\right| + b\left|\quad\right| + c\left|\quad\right|$$

2. PROPERTIES

ROW ADDITIVE

SCALAR

INTERCHANG

3. ADVANCED

LOWER/UPPER = PRODUCT OF DIAGONALS

$$\det(A^t) = \det(A)$$

$$\det(A) = 0 \text{ IFF } A \text{ IS SINGULAR}$$

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

$$\det(A) = \det(P)\det(L)\det(U)$$

4. APPLICATIONS

CRAMER'S RULE TO SOLVE $AX=b$

FINDING A^{-1}

CHAPTER 5 REVIEW

Friday, July 26, 2019 11:40 AM

DEFINITION $AX = \lambda X$

FIND λ BY $P_A(\lambda) = 0$

$\det(A - I\lambda) = 0.$

$\#\lambda$ (WITH MULT) $= n$.

$$\underbrace{1 \le X_\lambda \le \lambda \text{ (MULT)}}$$

DIAGONALIZABLE AT $=$,

DEFICIENT AT $<$

$A = PDP^{-1}$

$$\begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_n \end{bmatrix}$$

APPLICATION

$A^k b$.

SIMILARITY

A IS SIMILAR TO M IF THEY HAVE THE SAME EIGENVALUES.

HAVE THE SAME EIGENVALUES.

3.1, 3.4-5, 4.1-3, 5.1-2.