MATH 19620 Final Study Guide

Anthony Yoon

May 2024

Disclaimer

This will be a study guide for MATH 19620, Linear Algebra (Math 19620). Please refer to given notes for the most accurate representation of what is to be expected in class. There will not that many **applied** approaches, as I believe that that will be very easy to learn from the applied approach from the **fundamental** approaches in mathematics. Also this is a very **verbose** guide, so if it is intimidating, then I'm really sorry. If you want to see the giant list of invertible matrix properties, see the very back of this study guide.

1 What is Linear Algebra?

If I where to breakdown what Linear Algebra is, it is the study of studying lines, planes, vectors, and other things using systems of linear equations. Keep this definition in mind as we go through this document.

1.1 So what is a system of linear equations?

Fundamentally, a linear equation contains only "linear" terms. What this means is that all of the variables have exponents that equal **1**. So a great example of this is the equation (and the main "style" of the equations we will be studying):

$$\mathbf{a + b - c = 0}$$

The following equation $a^2 + b + c = 2$ is an example of an **nonlinear** equation. Logically, a system of linear equations is multiple of these linear equations that are all linked to each other. So the following is a system of linear equation:

$$
\begin{aligned}
a + b - c &= 0 \\
a + 2b + c &= 9 \\
a + 3b + 2c &= 15
\end{aligned}
$$

System 1

Up till this, we have learned the classic methods of solving these linear equations, where we could have

1. Added two of the three equations together
2. Subtracted two of the three equations together
3. Add a scalar multiple of an equation to another
4. Multiplied an equation by a value (or more formally, multiplied an equation by a scalar value)

Eventually, once you performed all operations, you would end up the following solution: $a = -10, b = 9, c = -1$. However, in linear algebra, we can represent this as a **matrix**.

1.2 So what is a matrix?

Think of a matrices as a thing that stores values in a fixed manner. For CS people, you can think of this as an array (or list for you python people). Semantically, these are referred to by **Row by Column** (So a matrix with 3 rows and 4 columns is a 3 by 4 matrix). So for system 1, the matrix that contains the equations can be represented as:

$$
\begin{bmatrix}
1 & 1 & -1 & 0 \\
1 & 2 & -1 & 9 \\
1 & 3 & 2 & 15
\end{bmatrix}
$$

Where it is evident that the respective coefficients are stored within the matrix.

Matrix operations

Notice in the "normal" operations, only the coefficients where actually manipulated. Thus, we can do something similar with the "rows" of this matrix. These are called **row operations**, meaning that you can.

1. multiply a row by a scalar multiple (analogous with multiplying an equation with a scalar multiple)

2. switch a row with another (analogous with switching two of the equations around)

3. adding/subtracting a row to another (analogous with adding two equations with each other)

4. add a row multiplied by a scalar multiple with another.

The end goal of using these operations is to be able to **row reduce** a matrix down to see solutions. But how do we know that these things are solutions within matrix form? We can first look at what the system of equations looks like.
Let the solution to a 3 equation and 3 variables system of equations be $a = x, b = y, andc = z$. if we were to rewrite these to the original system of equations, we can represent them as

$$\begin{aligned} a + \mathit{0}\,b + \mathit{0}\,c = x \\ \mathit{0}\,a + b + \mathit{0}\,c = y \\ \mathit{0}\,a + \mathit{0}\,b + c = z \end{aligned}$$

So if we were to represent this in matrix form, we could represent the following system as the matrix

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

So therefore, we are interested in the diagonal of the matrix being all ones and every other entry being 0s. This form is very important in linear algebra and is called **Reduced Row Echelon Form or RREF**. So for the original equation (see system 1), the following operations can be performed where $\mathbf{R_2} + \mathbf{R_1}$ means row 2 + row 1 is the new row one. (Basically The rightmost row is being the row that is being replaced.[1]

$$\begin{bmatrix}1&1&-1&0\\1&2&-1&9\\1&3&2&15\end{bmatrix} \xrightarrow{\mathbf{-R_1+R_2\to R_2}} \begin{bmatrix}1&1&-1&0\\0&1&0&9\\1&3&2&15\end{bmatrix} \xrightarrow{\mathbf{-R_1+R_3\to R_3}} \begin{bmatrix}1&1&-1&0\\0&1&0&9\\0&2&3&15\end{bmatrix}$$

$$\xrightarrow{\mathbf{-2R_2+R_3\to R_3}} \begin{bmatrix}1&1&-1&0\\0&1&0&9\\0&0&3&-3\end{bmatrix} \xrightarrow{\mathbf{1/3R_3\to R_3}} \begin{bmatrix}1&1&-1&0\\0&1&0&9\\0&0&1&-1\end{bmatrix} \xrightarrow[\mathbf{R_3+R_1\to R_1}]{\mathbf{-R_2+R_1\to R_1}}$$

$$\begin{bmatrix}1&0&0&-10\\0&1&0&9\\0&0&1&-1\end{bmatrix}$$

These matrices that represent these system of equations can be written as an **augmented matrix**. It is shown below.

$$\left[\begin{array}{ccc|c}1&-1&1&0\\1&2&-1&9\\1&3&2&15\end{array}\right]$$

The coefficients of the equations can be put into an **coefficient matrix**, which is shown below.

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

2 RREF and all related topics

Reduced Row Echelon Form is very important in the study of linear algebra. It enables us to quickly determine the solutions and characteristics of a system of linear equations. The characteristics of RREF systems are

- all rows have 0s at the bottom
- each row will have 0s that will end with a singular 1. (note: the 1st row generally always start with a 1.
- the leading one (which is known as a pivot) is always to the right of another leading one.
- all entries above and below a leading one are 0

[1] I prefer the notation $\mathbf{R_2} + \mathbf{R_1} \to \mathbf{R_1}$, but some professors prefer the above notation. Consider both interchangeable

A leading one is considered a **pivot point**. It can be best visualized here.

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

The number of pivot points that a matrix has can be written as the **Rank** of the matrix. So formally, if we are given the matrix A, then $\mathbf{Rank(A)} = \mathbf{3}$.
So given these characteristics, we can understand a lot about the solutions of a matrix. There exists 3 possibilities of the solution of any system of matrices:

- No Solution
- One Solution
- Infinitely many solutions

Now that we know these solutions, then there are matrix representations we can rely on to demonstrate these properties. For example:

- If there are no solutions, then within the matrice, there will be an inherent contradiction. Typically, this will be seen as [0 0 0 ... c] for some constant c. This is because that it implies that 0 = a constant value, which is a contradiction/
- One solution is just the matrix form described above.
- Infinitely many solutions is in the form $\left[\begin{array}{ccc|c} 1 & a & b & e \\ 0 & 1 & c & f \\ 0 & 0 & 0 & 0 \end{array}\right]$, where $a, b, c, e, f \in \mathbb{R}$.
  A common characteristic of these systems are having a row of 0s on the bottom row and having a rank $<$ than the column.

Note, if you were to solve an equation with infinitely many solutions, the values on the right of the pivot point are what are refereed to as **free variables**. If you were to solve these out, you would solve for equations in the form:

$$\text{variable in the pivot point} = \text{constant} + \text{free variables}$$

From there, you set each of these free variables to a different variable (each free variable gets it's different variable) and **clearly state that the variable** (in this case, let's say s) is $s \in \mathbb{R}$ So given what we know about the solutions, there must be an easy way to find these things. We can do this using rank.

2.1 Using rank to determine solutions

If a linear system has m rows and i columns, then the following points hold:

- if r = m, then there is exactly one solution
- if r = n, then there cannot be exactly infinitely many solutions
- if the system has exactly solution, then n ≤ m
- if r = m = n, thn there is exactly one solution
- rank + free variables = n

2.2 Matrix Multiplication and Addition

2.2.1 Matrix Addition

Let A = $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and B = $\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ Matrix addition is defined for 2 by 2 matrices is

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

You can extend this definition to bigger matrices. Let

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

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

Therefore, matrix addition can be defined as

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b22 & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

Similarly, this can be applied to be scalar multiplication.

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

2.2.2 Vector Multiplication

A vector is a matrix with only one column. Thus, a vector with n entries is called a vector in $\mathbb{R}^n$. This can be seen in:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

Let A be a matrix with columns $\vec{v_1}, \vec{v_2}, \vec{v_3}, \dots, \vec{v_n}$. Then we can define the operation $A\vec{x}$ as

$$A\vec{x} = x_1\vec{v_1} + x_2\vec{v_2} + \dots + x_n\vec{v_n}$$

The above is an example of a **Linear Combination**

3 Linear Transformation

A linear transformation is a function that is defined as $f : \mathbb{R}^n \to \mathbb{R}^m$ where an unique vector $\vec{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$ maps to $\vec{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$ You can represent a linear transformation as $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$. A good rule of thumb is that if the linear transformation representative matrix has any non-linear terms or adding of any constants, then it is not a linear transformation.

3.1 Properties of Linear Transformation

A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if,

1. $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$

2. $T(k\vec{v}) = kT(\vec{v})$

One thing to note that for a transformation to be a linear transformation. The following must hold.

$$T(\vec{0}) = \vec{0}$$

3.2 Proving that a linear transformation is indeed a linear transformation

A function can be a linear transformation or it may not be one. Such, here's what to do in each case.

- **Not a linear transformation**: Provide a counter example and state what property it is violating. Here are some red flags you can look for:
  1. There are constants being added inside the subspace.
  2. The variables are non-linear
- **Is a Linear Transformation**: Either find the matrix associated with the linear transformation or prove the properties hold using matrices that hold variables that can take any value.

4 Linear Transformations and their compositions

4.1 Geometric Linear Transformations

You can represent vector transformations as linear transformations, which have geometric effects.

1. Scaling by a factor of k: So this means $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$
2. Reflecting in the the X-axis: $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$
3. Reflecting in the Y axis: $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$
4. Rotation θ°counterclockwise[2]: $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

[2] Plug in 90°for the 90°rotation

5. Horizontal Shear: $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

6. Vertical Shear: $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

7. Orthogonal Projection onto the X axis: $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

8. Orthogonal Projection onto the Y axis: $T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

From here, you can "chain" these transformations together. When we "chain" these together, we just multiply the matrices together from the rules above. For example, the chain of operation **Reflecting in the Y axis followed by a horizontal shear by a factor of 2** can be represented as the following:

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

This chaining of matrix operations is called composition of linear functions. However, now we can attach a more precise definition to Matrix Multplication.

4.2 Matrix Multiplication in depth.

[3]If A is an m x n matrix and B is as $B = \begin{bmatrix} | & | & \dots & | \\ v_1 & v_2 & \dots & v_p \\ | & | & \dots & | \end{bmatrix}$, then

$$AB = \begin{bmatrix} | & | & \dots & | \\ Av_1 & Av_2 & \dots & Av_p \\ | & | & \dots & | \end{bmatrix}$$

4.3 Properties of Matrix Multiplication

- Not Commutative, so it is not necessarily true that AB = BA
- Is Associative and Distributive, meaning: A(BC) = (AB)C and A(B+C) = AB + AC
- if $k \in \mathbb{R}$, then A(kB) = k(AB)

[3]The vertical lines show that each column is indicated by the line

5 Inverse Matrix

Before we cover what an inverse matrix is, it may be worthwhile to explore what an identity matrix is. An identity matrix is just a **square RREF matrix that has rank that equals the columns**. So an example of that is:

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

A crucial property of the identity matrix is that **Any matrix multiplied by the identity matrix will just be itself. I.E** $AI_n = A = I_nA$ A linear transformation is invertible if and only if for any $\vec{y} \in \mathbb{R}$, there is an unique $\vec{x} \in \mathbb{R}$ such that $T(\vec{y}) = \vec{x}$ Following from this, the official definition of what a invertible matrix is that a matrix A is invertible if the corresponding linear transformation T is invertible. We define A^{-1} to be the matrix for T^{-1}.
One important property of the invertible matrix is that

$$A^{-1}A = AA^{-1} = I_n$$

This is one of the ***MANY*** properties of invertible matrices we should know. So, at this point in time, this is properties that we know of. [4]

- T is invertible
- A is invertible
- For any $\vec{y} \in \mathbb{R}^n$ the system $A\vec{x} = \vec{y}$ has a unique solution.
- The system $A\vec{x} = \vec{y}$ has an unique solution $\vec{x} = \vec{0}$
- $rank(A) = n$
- $rref(A) = I_n$

One cool thing is that every invertible matrix has it's own unique inverse. The proof can be seen below
Proof. Suppose, for contradiction, that there is a matrix A with two inverses. So there are matrices B and C such that

$$AB = BA = I_n$$

[4]There will be a massive list of these properties in the back of this document, just listing all of them here would not make any sense whatsoever.

and

$$AC = CA = I_n.$$

Then

$$\begin{aligned}
C &= CI_n && \text{by the property of the identity matrix} \\
&= C(AB) && \text{since we assumed } AB = I_n \\
&= (CA)B && \text{by associativity of matrix multiplication} \\
&= I_nB && \text{since we assumed } CA = I_n \\
&= B && \text{by the property of the identity matrix.}
\end{aligned}$$

In general, $AB \neq BA$ even if A and B are square matrices, however, if the product is the identity matrix, it turns out that in fact they will be equal. [5]

5.1 So how do we find invertible matrices?

It is easier to show how it's done rather than explaining. So, consider a 2×2 matrix A given by:

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

To find A^{-1}, we perform row operations on the augmented matrix $(A|I)$, where I is the 2×2 identity matrix.

$$\left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right] \xrightarrow{\mathbf{R_1}/a \to \mathbf{R_1}}$$

$$\left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ c & d & 0 & 1 \end{array}\right] \xrightarrow{-cR_1 + R_2 \to R_2}$$

$$\left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & d - \frac{bc}{a} & -\frac{c}{a} & 1 \end{array}\right] \xrightarrow{R_2/(d - \frac{bc}{a}) \to R_2}$$

$$\left[\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{1}{d - \frac{bc}{a}} \end{array}\right] \xrightarrow{-\frac{b}{a}R_2 + R_1} R_1$$

$$\left[\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right]$$

Basically, we have 2 matrices, one on the right and one on the left, and we row reduce the left matrix while still manipulating the right matrix so that when we finish all operations, then we get the identity matrix on the left and the resultant inverse

[5] Normally, I don't put proofs in review guides, but this one's pretty cool.

matrix on the right. The following is an example for a 3 by 3 system.

$$\left[\begin{array}{ccc|ccc} 5 & 6 & -1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 1 & 2 & 4 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\mathbf{R_1 \div 5}}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1.2 & -0.2 & 0.2 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 1 & 2 & 4 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\substack{\mathbf{-R_1+R_2\to R_2} \\ \mathbf{-R_1+R_3\to R_3}}}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1.2 & -0.2 & 0.2 & 0 & 0 \\ 0 & 0.8 & 3.2 & -0.2 & 1 & 0 \\ 0 & 0.8 & 4.2 & -0.2 & 0 & 1 \end{array}\right] \xrightarrow{\mathbf{R_2 \div 0.8}}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1.2 & -0.2 & 0.2 & 0 & 0 \\ 0 & 1 & 4 & -0.25 & 1.25 & 0 \\ 0 & 0.8 & 4.2 & -0.2 & 0 & 1 \end{array}\right] \xrightarrow{\mathbf{-0.8R_2+R_3\to R_3}}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1.2 & -0.2 & 0.2 & 0 & 0 \\ 0 & 1 & 4 & -0.25 & 1.25 & 0 \\ 0 & 0 & 0.8 & 0.0 & -1 & 1 \end{array}\right] \xrightarrow{\mathbf{R_3 \div 0.8}}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1.2 & -0.2 & 0.2 & 0 & 0 \\ 0 & 1 & 4 & -0.25 & 1.25 & 0 \\ 0 & 0 & 1 & 0.0 & -1.25 & 1.25 \end{array}\right] \xrightarrow{\substack{\mathbf{-4R_3+R_2\to R_2} \\ \mathbf{0.2R_3+R_1\to R_1}}}$$

$$\left[\begin{array}{ccc|ccc} 1 & 1.2 & 0 & 0.2 & -0.25 & 0.25 \\ 0 & 1 & 0 & -0.25 & 6.25 & -5 \\ 0 & 0 & 1 & 0.0 & -1.25 & 1.25 \end{array}\right] \xrightarrow{\mathbf{-1.2R_2+R_1\to R_1}}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0.5 & -8 & 6.25 \\ 0 & 1 & 0 & -0.25 & 6.25 & -5 \\ 0 & 0 & 1 & 0.0 & -1.25 & 1.25 \end{array}\right]$$

Notice in the 2 by 2 determinant, there is notion of $ad - bc$. This is called the **determinant** of a 2 by 2 matrix, and this behavior can be expanded to higher dimension very easily. But for now, understand this for the 2×2 matrix. The form we derived from the $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ matrix is very important, as it allows quick access to the inverse of a 2 by 2 matrix.
Notice if $ad - bc = 0$, then the whole matrix falls apart, which is a hint a very important property of invertible matrices: **If the determinant of a matrix is 0, then it is uninvertible.** One thing to note is that you can use the idea of invertible matrices to solve systems of linear equations. Let $A\vec{x} = b$, where A is a matrix and b and x are vectors. Then the following holds.

$$\vec{x} = A^{-1}A\vec{x} = A^{-1}\vec{b}$$

6 Image and Kernel

Before we introduce what the image and kernel are, we must first discuss what a *span* is. A span is a set that contains vectors that can be used to create other vectors. Defined rigorously, *a span is the set of all linear combination of vectors.* However, the vectors in the span must be "unique". This means that you cannot create the vector using any sort of *linear combination* of the vectors in the set.

6.1 But what is a linear combination?

We defined what a linear combination is informally above. But, rigorously it is defined as

Let $c_1, c_2, c_3, c_4, \ldots, c_n \in \mathbb{R}$ [6] and let $\vec{v}_1, \vec{v}_2, \vec{v}_3, \ldots, \vec{v}_n$ be vectors. Then a linear combination is follows:

$$c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 + \ldots c_5v_5$$

So if you can express a vector as in a span as a linear combination of vectors in a set, then that vector is redundant and thus can be **removed** from the span. Some vocab to know:

- if $c_1 = c_2 = c_3 = \cdots = c_n = 0$, that is the *trivial solution* and is how you know something is linearly independent
- if the solution is not the trivial solution, then the it is linearly dependent.

6.2 Image and Kernel defined

An image is the set of all possible outcomes of a function. A kernel is the set of all inputs that can lead to the output of $\vec{0}$.[7]

The image of a linear transformation can be found by taking the span of the columns of the matrix representative of the linear transformation. The kernel is a bit more nuanced. Out of the 3 possible outcomes of number of solutions, we are only interested in the infinitely many solutions, as no solutions and one solution are trivial. And the trick to solving these questions is to RREF the matrix representative of the linear transformation. More specifically:

1. Find matrix representative of the linear transformation

[6] all the cs are constants

[7] A set is just a way of storing things in math. Think: array or list in cs

2. RREF the matrix

3. If you are finding the image of the linear transformation, the columns containing the pivot points in the RREF matrix are the same columns within the original matrix that will be put into the span of the image.

4. If you are finding the kernel of the linear transformation, create an augmented matrix where the right side of the matrix is the 0 vector (right of the line) and the left side is the matrix.

An example of finding the kernel is as follows. Given the matrix:

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

If we were to set this up as an augmented matrix to solve for the kernel:

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

The RREF form of this matrix is:

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

Thus if we were to represent this equation form.

$$x + \frac{4}{3}z = 0$$
$$y + \frac{1}{3}z = 0$$

which is equivalent to

$$x = -\frac{4}{3}z$$
$$y = -\frac{1}{3}z$$

And if we let $z = t \in \mathbb{R}$ Then, the following is the kernel of transformation.

$$\ker(T) = \left\{ \begin{bmatrix} -\frac{4}{3}t \\ -\frac{1}{3}t \\ t \end{bmatrix} \middle| t \in \mathbb{R} \right\} = \operatorname{span}\left\{ \begin{bmatrix} -\frac{4}{3} \\ -\frac{1}{3} \\ 1 \end{bmatrix} \right\}.$$

From these, we can now add 2 properties to invertible matrices.

1. $ker(T) = \vec{O}$

2. $im(T) = \mathbb{R}^n$

7 Subspace and Bases

A subspace is just a subset of a vector space. Or if one wants to be more precise: $W \subseteq \mathbb{R}$. But for W to be a subspace, the following theorem must hold.

Theorem

$W \subseteq \mathbb{R}$ if and only if

1. W is closed under addition
2. W is closed under scalar multiplication
3. W holds $\vec{0}$

Mathematically, this means that if we have given that $\vec{v}, \vec{y} \in W$

1. $\vec{v} + \vec{y} \in W$
2. Given a $k \in \mathbb{R}$, $k\vec{v} \in W$

Note that that these are very similar to those of linear transformations. To prove these, there are 2 cases [8].

- **Not a subspace**: Typically, these will have a couple of red flags.
  1. There are constants being added inside the subspace.
  2. The variables are non-linear.

  if that is the case, all you have to do is **provide a counter example to the given and state which property of subspace it is violating**

- **Is a subspace**: Prove using general terms. Meaning that you set every entry in the matrix as a variable that can take on any value and show that the properties hold.

7.1 Geometric Interpretation

We know of two possible geometric interpretations of subspaces:

- In $\mathbb{R}^2$, possible subspaces are $\{\mathbf{0}\}$, lines through the origin, and $\mathbb{R}^2$.
- In $\mathbb{R}^3$, possible subspaces are $\{\mathbf{0}\}$, lines through the origin, and planes through the origin.

[8] These are patterns I've observed after just a lot of practice

7.2 Basis

A basis is a span that is linearly independent. That's it really it.

To prove that a span is linearly independent, we set each vector in the basis as a column of matrice, and the RREF of the matrix should be:

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \dots & 1 & 0 \end{array}\right]$$

You can apply this to concept to geometric problems. If you are solving the basis of a plane, they are very similar to the kernel problems. There are some to know.[9]

- A basis in $\mathbb{R}^n$ must contain n vectors

7.3 Dimension

Dimension is just the number of vectors in a span or basis. This important in the following theorem.

Rank Nullity Theorem

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then

$$\dim(\text{im}(T)) + \dim(\ker(T)) = n.$$

Equivalently,

$$\text{rank}(T) + \text{nullity}(T) = n.$$

We can use this to "generate" matrices given information about rank and kernel.

8 Using the Rank Nullity Theorem

Here is how to use the Rank Nullity Theorem to generate matrices. [10]

1. **Calculate the missing value**: If we are given *dim(rank)* or vice versa, use the Rank Nullity Theorem to find this value.

[9] I'll add more if necessary, but I think this is the only one you need to know.
[10] This is how I personally like to do them, but there are other ways of doing it

2. **Fill in the pivot points**: Using the value of *dim(rank)*, set the first *dim(rank)* columns as pivot points, from left to right.

3. **Fill in the Free Variables**: Once we fill in the pivot points, fill in the remaining spots with *free variables* in the remaining columns.

4. **Double Check**: Make sure the matrix you generated has the properties.

9 Coordinates

If we are given a basis, say B. The coordinates of vector $\vec{a}$ is just a vector that contains the coefficients of the linear combination of the vectors in B that will create $\vec{a}$. Or more formally:
Consider a basis $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ of a subspace of $\mathbb{R}^n$. Then there are unique scalars $c_1, \dots, c_k$ such that

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k.$$

We call the scalars $c_1, c_2, \dots, c_k$ the B-coordinates of $\mathbf{x}$, and the vector

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

is the B-coordinate vector of $\mathbf{x}$, denoted by $[\mathbf{x}]_B$.
Note: The first vector's coefficient is the first entry in the matrix.

9.1 How to solve these problems

1. **Identify the Basis**: Make sure that you have copied down the matrix down correctly on your work, otherwise you will fail.

2. **Find the coefficients**: If you can see how the vectors in the basis can from your target vector, just write it down. If not, solve the matrix problem:

   $$\mathbf{A}\vec{x} = \text{target vector}$$

   where $\mathbf{A}$ is the matrix created by using the vectors in the basis as columns and $\vec{x}$ represent coefficients in the linear combination.

3. **Double Check**: Double Check that your coordinates actually work.

9.2 Properties of these coordinates

These properties are taken directly from Professor Ziesler's notes, so full credit to the Professor.

9.2.1 Theorem 1

Given a subspace W of $\mathbb{R}^n$, with basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, and given a vector $\mathbf{w} \in W$, $[\mathbf{w}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ is the unique solution to the equation

$$\begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_k \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{w}.$$

Thus, if we define $S = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_k \\ | & | & \cdots & | \end{bmatrix}$, then $S[\mathbf{w}]_B = \mathbf{w}$ and also $S^{-1}\mathbf{w} = [\mathbf{w}]_B$. (If this equation has no solution then $\mathbf{w}$ is not in W.)

9.2.2 Theorem 2

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for $\mathbb{R}^n$. Then there is a unique matrix B such that $[T(\mathbf{x})]_B = B[\mathbf{x}]_B$, with

$$B = \begin{bmatrix} | & | & \cdots & | \\ [T(\mathbf{v}_1)]_B & [T(\mathbf{v}_2)]_B & \cdots & [T(\mathbf{v}_n)]_B \\ | & | & \cdots & | \end{bmatrix}$$

For Theorem 2, here is a step by step process of how to solve this problem.

1. Plug in v_1 into the function and find the coordinates of that
2. Do that over and over again until you can fill in the B vector

10 Change of Basis

Let A be the matrix and B be the matrix above (I.E the matrix in respect to the basis of interest) and S be the matrix such that $S = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_k \\ | & | & \cdots & | \end{bmatrix}$, where the

$\mathbf{v_n}$ are vectors in the basis of interest. Then the following holds.

$$B = S^{-1}AS \ and A = SBS^{-1}$$

The following is a little chart that demonstrates the operations we can do to get the appropriate things.

$$\begin{array}{ccc} [\mathbf{x}]_B & \xrightarrow{B} & [T(\mathbf{x})]_B \\ \downarrow S & & \uparrow S^{-1} \\ \mathbf{x} & \xrightarrow{A} & T(\mathbf{x}) \end{array}$$

The arrow indicates the multiplication required.

11 Review of the vector properties

Let $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$.

1. The dot product of $\mathbf{v}$ and $\mathbf{w}$ is given by

$$\mathbf{v} \cdot \mathbf{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n.$$

2. Vectors $\mathbf{v}$ and $\mathbf{w}$ are orthogonal if $\mathbf{v} \cdot \mathbf{w} = 0$.

3. The length (magnitude/norm) of $\mathbf{v}$ is given by

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

4. A vector is an unit vector if and only if it has a length of 1

5. $\vec{a}(\vec{b} \cdot \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$

11.1 Orthonormal Vectors

Vectors $\vec{u_1}, \vec{u_2}, \vec{u_3}, \ldots, \vec{u_n}$ in $\mathbb{R}^n$ if

- All vectors are unit vectors
- All vectors are orthogonal

Note that a list of orthonormal vectors are linearly independent. So, if these vectors form a span, that span is automatically a basis. And this kind of matrix is a special kind of matrix called the **Orthonormal Basis**, and hence referred as an ONB. Because of these be.

1. A list of n linearly independent vectors in a n-dimensional space$subspace is a basis.
2. A list of n that spans a n-dimensional space$subspace is a basis
3. Any list of orthonormal vectors is linearly independent. It then follows that n orthonormal vectors in an n-dimensional space are automatically a basis.
4. If $B = \{\vec{u_1}, \vec{u_2}, \vec{u_3}, \ldots, \vec{u_k}\}$ is an ONB for a subspace W and $\vec{w}$ is in W then the coefficient vector for $\vec{w}$ for respect to B is given by

$$[\vec{w}]_B = \begin{bmatrix} \vec{u_1} \cdot \vec{w} \\ \vec{u_2} \cdot \vec{w} \\ \vec{u_3} \cdot \vec{w} \\ \vdots \\ \vec{u_k} \cdot \vec{w} \end{bmatrix}$$

12 Orthogonal Projections, Gram-Schmidt Process, Orthogonal Matrix

Some definitions to get out of the way

Definition

A vector is orthogonal to a subspace W if it is orthogonal to all vectors in W. We can define $W^\perp$ as the subspace of all orthogonal vectors. A subsequent theorem that follows is:

Theorem 1

Given any vector $\vec{v} \in \mathbb{R}^n$ and a subspace W of $\mathbb{R}^n$, we can write $\vec{v}$ as the sum of a vector $\vec{v}^{//}$ in W with a vector $\vec{v}^\perp$ orthogonal to W. We have

$$\vec{v} = \vec{v}^{//} \ + \vec{v}^{\perp]}$$

Definition

We can call $\vec{v}^{//}$ the projection of $\vec{v}$ on W, where W has an ONB = $\{\vec{u_1}, \vec{u_2}, \dots, \vec{u_k}\}$ and write it as

$$proj_w \vec{u} = (\vec{u_1} \cdot \vec{v})\vec{u_1} + (\vec{u_2} \cdot \vec{v})\vec{u_2} + \cdots + (\vec{u_k} \cdot \vec{v})\vec{u_k}$$

One key thing to note is that for you to use the projection, the thing underneath the projection (AKA the subspace) **must be an ONB**.

12.1 Gram-Schimdt Orthonormalization

This is an algorithm that will take a basis of a subspace W of $\mathbb{R}^n$ and uses it to create an ONB.
Let the basis be B = $\{\vec{u_1}, \vec{u_2}, \dots, \vec{u_k}\}$. Follow these steps.

1. Take the very first vector and let that become a unit vector. So $\vec{u_1} = \frac{\vec{v_1}}{\|\vec{v_1}\|}$
2. Let $W_1 = span\{\vec{u_1}\}$
   - Find $proj_{W_1}\vec{v_2} = (\vec{u_1} \cdot \vec{v_2})\vec{u_1}$
   - $\vec{v_2} - proj_{W_1}\vec{v_2}$ is orthogonal to W_1. So we have to normalize to get an unit vector and define $u_2 = \frac{\vec{v_2} - proj_{W_1}\vec{v_2}}{\|\vec{v_2} - proj_{W_1}\vec{v_2}\|}$
3. Let $W_2 = span\{\vec{u_1}, \vec{u_2}\}$.
4. Redo step 2 with W_2 and $\vec{v_3}$ and keep iterating until you get enough vectors for a basis.

12.2 Orthogonal Matrix and Transformations

A linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is called an **orthogonal transformation** if T preserves the length of vectors in $\mathbb{R}^n$. I.E, for all $\vec{x} \in \mathbb{R}^n$,

$$\|T(\vec{x})\| = |\|\vec{x}\|$$

An $n \times n$ matrix A is an **orthogonal matrix** if and only if it is the matrix representative of an orthogonal transformation.

Another property is that if given T as an orthogonal transformation, then for all $\vec{v}, \vec{w} \in \mathbb{R}^n$,

$$T(\vec{v}) \cdot T(\vec{w}) = \vec{v} \cdot \vec{w}$$

Another characterization of orthogonal transformations is: $T : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation if and only if $\{T(\vec{e_1}), T(\vec{e_2}), \ldots, T(\vec{e_n})\}$ is an **<u>ONB</u>** of $\mathbb{R}^n$. In other words, A is an $n \times n$ orthogonal matrix if and only if the columns of A are an ONB of $\mathbb{R}^n$. To simplify this meaning even more: **If the columns of the matrix are orthonormal, then the matrix is orthogonal.**

13 More orthogonal matrix properties

We have to define what a transpose is first before we continue. A **transpose** of an $m \times n$ matrix A $= \begin{bmatrix} | & | & \ldots & | \\ \vec{v_1} & \vec{v_2} & \ldots & \vec{v_n} \\ | & | & \ldots & | \end{bmatrix}$ is the matrix. The transpose is defined by:

$$A^T = \begin{bmatrix} - & \vec{v_1} & - \\ - & \vec{v_2} & - \\ \vdots & \vdots & \vdots \\ - & \vec{v_n} & - \end{bmatrix}$$

A comprehensive list of characterizations of orthogonal matrix for a T is an orthogonal linear transformation:

1. A is an orthogonal matrix
2. $\|A\vec{x}\| = \|\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$
3. The columns are an orthonormal to each other
4. $A^T A = I_n$
5. $A^-1 = A^T$
6. $(A\vec{x}) \cdot (A\vec{y}) = \vec{x} \cdot \vec{y}$

Properties of Transpose that may be useful:

1. **Transpose of a sum**:
$$(A+B)^T = A^T + B^T$$
2. **Transpose of a scalar multiple**:
$$(kA)^T = kA^T$$

3. **Transpose of a product**:

$$(AB)^T = B^T A^T$$

4. **Transpose of an inverse** (for all invertible matrices A):

$$(A^T)^{-1} = (A^{-1})^T$$

5. **Rank equality**:

$$\text{rank}(A) = \text{rank}(A^T)$$

13.1 Least Square Solution

For future reference, I will refer to Least Square Solution as LSQ. The whole premise of this is to solve the solution. Given $A\vec{x} = \vec{b}$, find the solution to

$$A\vec{x^*} = proj_{im(A)}\vec{b}$$

And doing so minimizes the quantity:

$$\|\vec{b} - A\vec{x^*}\|^2$$

. However, this gets really tedious, so we can just use this:

$$A^T A\vec{x^*} = A^T\vec{b}$$

However, not everything in life is perfectly modeled by a linear solution. In cases in equations in the form:

$$y = ae^{bx}$$

We can manipulate this equation using logs to get the respective equation.:

$$ln(y) = c_1 + c_2 x$$

Where $c_1 = ln(a)$ and $c_2 = ln(b)$. Similarly,, if the equation is in the form:

$$y = ax^b$$

We can rewrite the equation to be in the form:

$$ln(y) = c_1 + bln(x)$$

Where $c_1 = ln(a)$.

14 Determinants

Going way, way back. Remember that the if given matrix A = $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant or Det(A) = ad - bc. However, let A = $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \vec{v} & \vec{w} \end{bmatrix}$. Then det(A) = ad - bc = $\|\vec{v}\| \cdot \|\vec{w}\| sin\theta$, where θ is the angle from v and w.

Imagine two vectors being drawn in a 2 dimensional plane. You can "draw" a parallelogram in between the two. The area would be just be $\mid Det(A) \mid$ where the columns are the vectors in any order. [11]

Another theorem that may be worth noting is for any parallelogram Ω and T is a linear transformation and A is the associated matrix:

$$\text{area of } T(\Omega) = \mid detA \mid \text{(area of } \Omega)$$

This applies to higher dimensions. Just replace area with volume and you're good to go.

14.1 How to find determinants, Part 1

Let A be an $n \times n$ matrix. Define A_{ij} to be the $(n-1) \times (n-1)$ matrix obtained by removing the i-th row and j-th column of A. We call these matrices *minors*. From here, we also need to see what an Laplace expansion is. The following matrix is very important.

$$\begin{bmatrix} + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Select a row, and then go down the row and eliminate columns and get minors. From there in the row, the first value in the column you eliminated, that is your coefficient times the proper + or - sign according to the Laplace expansion. From there, calculate the determinant of the minors, multiply by the coefficient and add to get your determinants. [12]

[11] The reasoning why order does not matter will be covered later.

[12] This is surprisingly hard to explain and frankly I do not have it in me right now to type up examples, but you know how to do them .

14.2 How to find determinants part 2

If a matrix is in diagonal/triangle form, then it is very easy to calculate the determinant. **Just multiply the diagonals**. See below for the different types of matrices.

$$\text{Upper Triangular Matrix: } \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$$

Figure 1: An example of an upper triangular matrix.

$$\text{Lower Triangular Matrix: } \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Figure 2: An example of a lower triangular matrix.

$$\text{Diagonal Matrix: } \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$$

Figure 3: An example of a diagonal matrix.

Since we see how easy it is to create a diagonal matrix, we can row reduce (or column reduce for that matter) into a more favorable matrix. We can do row/column operations to reduce the matrix into a RREF matrix.[13]

- If we multiply a row/column of A by a scalar k, the determinant also gets multiplied by a factor of k. Note: when you are actually **solving** these matrices, the determinant in the end will be multiplied by a $\frac{1}{k}$ rather than k, as this statement would be detB = detA

- If we swap a row/column around, then the determinant also gets multiplied by a factor of -1.

[13] I'll put in why we can do this later

- If we were to add a row multiplied by a scalar multiple to another, nothing changes.

When solving these problems, please keep track **all determinant** values on the side. Trust me, this makes life a lot easier.

15 How to find determinants part 3

Note, that if there are 2 rows or columns that are all 0s, then the determinant is 0. Here are some properties of determinants to know of:

- $det(I_n) = 1$
- $det(AB) = det(A)det(B)$
- $det(A^{-1}) = \frac{1}{det(A)}$
- if $det(A) = 0$ then, A is not invertible.

Some of my own notes about finding determinants:

- If you can see that a matrix is really RREFable[14], then go through the row operations.
- If you see a bunch of zeros in a matrix, then it may be worthwhile to go through the Laplace expansion, as many things will get that neutralized.
- And my own personal opinion, when in doubt for anything 3×3 and below, use the Laplace expansion. Anything above, default to the row reduction form.

16 Eigenvectors and Eigenvalues

The root word *Eigen* means *own* in German[15]. So deriving from the root word, **An eigenvector is a vector that when multiplied by a matrix, will return itself times a scalar multiple, denoted as** λ. From there, we can create a **diagonal matrix**, where the literally the diagonals are the eigenvalues. Thus, a matrix is that can have a diagonal matrix is called a **diagonalizable** matrix.

An **eigenbasis** is a basis that consists of all possible eigen**vectors**. For any given matrix, we can have **up to n values** and many eigenvectors, so please keep that in mind. If we want a matrix to be diagonalizable, then there should be **n eigenvalues**

[14] we should make this a word

[15] Source: Google Translate

16.1 Eigenvectors from a geometric perspective

From a geometric perspective, the transformation should one that just returns the vector multiplied by some value. Basically, the only way the vector can change direction is if the vector is rotated 180°. But once we get to higher dimensional spaces is when that notion may change. Some notes from me:

- If we are in 2d, normally the only way for the transformation to have any eigenvectors or eigenvalues is to be some sort of scaling by a factor. If we are rotating by any quantity that is not 180°multiplied by some scalar multiple, then usually there are no eigenvectors or values with this transformation
- If we are in 2d, and reflecting over a line, the transformation is generally diagnolizable. The 2 vectors that would work would be
  - vector that lies on the line $\rightarrow$ will return itself
  - vector that is perpendicular to the line, which would indicate that the vector returned would just be 180 degree rotation of the original (Or in less wordy terms, multiplied by -1)
- If we are in 3d, think carefully about the rotation. We are rotating over a certain axis, then usually there only exists 1 possible eigenvector. Other than that, think real carefully

16.2 Properties of Eigenvalues and Eigenvectors

Theorems

- If T is an orthogonal transformation with standard matrix A then the only possible eigenvalues for A are 1 and -1.
- A is invertible if, and only if, 0 is not an eigenvalue for A

16.3 Finding Eigenvalues

When finding eigenvalues, we are interested in solving the following:
If A is a $n \times n$ matrice with eigenvalue λ if and only if

$$det(A - I_n\lambda) = 0$$

So when we solve a matrix in this for, we are interested in the form of matrix:

$$\begin{bmatrix} a_{11}-\lambda & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & a_{13} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33}-\lambda & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn}-\lambda \end{bmatrix}$$

And from there solve the resulting polynomial once we find the determinant. And because of that **triangular matrices have their eigenvalues along the diagonal**.

The function $det(A - I_n\lambda)$ is called the characteristic polynomial, with the roots being the eigenvalues. However, with this, we introduce **algebraic multiplicity**[16], basically the number of times the eigenvalue appears which is usually denoted by the exponent in the factored state.

Note: If n is odd, there exists at least one eigenvalue and we can have at most N eigenvalues.

Some notes about solving these problems:

- This is just solving a polynomial problem, so do not over complicate things
- As tempting as it may be, **do not skip steps here**. What I mean is write out every step in the matrix without fail, and things will become significantly easier. **the step of writing the resultant matrix of $A - I_n$ is very crucial.**

More definitions of eigenvectors and eigenvalues

Definition: A matrix A is diagnolizable if and only if there is an eigenbasis of A. Moreover, the diagonal matrix is given by SAS^{-1}, where S is the matrix whose columns are the vectors of the eigenbasis. The entries on the diagonal are the eigenvalues corresponding to the eigenvectors in the basis.

Definition: Suppose that λ is an eigenvalue of an $n \times n$ matrix A. Then the *eigenspace* associated with λ is $E_\lambda = Ker(A - \lambda I_n)$.

Definition: The **Geometric Multiplicity** is $dim(ker(A - \lambda I_n))$

[16] Please do not get confused with geometric multiplicity, which appears much later.

Definition: Eigenvectors corresponding to distinct eigenvalues are always linearly independent. It follows that if A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagnolizable. [17]

Definition: An $n \times n$ matrix A is diagonalizable if the sum of the geometric multiplicities of its eigenvalues is n.

So what does this mean in English? If we were to create an eigenspace for some λ, which is just the kernel of $A - \lambda I_n$, we would get a span with some amount of vectors. We can then place those vectors into an eigenbasis. Thus, the diagonal matrix is just the eigenvalues in the diagonal. **Note**: that if we do not have enough vectors in all eigenspaces, we cannot create an eigenbasis.

16.4 Appying the $A = SBS^{-1}$ rule

A = original matrix, B = diagonal matrix, S = eigenbasis basically.

16.5 Trace

The trace of an $n \times n$ matrix, written as tr(A), is the sum of the diagonal in the diagonal matrix or the diagonal of the original matrix. In other words, **the sum of the eigenvalues or the diagonal values of the original matrix.** With this, we can have the following ideas.

$$tr(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n \ and \ det(A) = \lambda_1 \lambda_2 \ldots \lambda_n$$

16.6 More properties of similar matrices

Suppose that A and B are similar $n \times n$ matrices. Then

(a) A and B have the same characteristic polynomial.

(b) A and B have the same eigenvalues, with the same algebraic multiplicities. However, the eigenvectors need not be the same.

(c) $\det A = \det B$ and $\operatorname{tr}(A) = \operatorname{tr}(B)$.

(d) $\operatorname{rank}(A) = \operatorname{rank}(B)$ and $\operatorname{nullity}(A) = \operatorname{nullity}(B)$.

(e) $A^n = SB^nS^{-1}$, for all natural numbers n.

[17] I prefer to think about as an eigenvalue corresponds to an eigenvector

17 Discrete Dynamical Systems

This is just a way of seeing how transition matrices perform over long periods of time. Refer to the idea that $A^n = SB^nS^{-1}$ applied to Eigenbasis and Eigenvalues. In this case, We are interested in the same form, but both sides being multiplied by $\vec{x}$, where A is equal to the transition matrix. And how do we actually solve these problems?

1. Identify the transition matrix and distribution vector.
2. Find the appropriate eigenbasis for the transition matrix.
3. From there, build S, S^{-1} and B
4. Solve equation SB^nS^{-1} while preserving the n. Note that the only the diagonal entries will get an exponent of n.
5. From there, use the original definition of matrix multiplication (Scalar Value times column of matrix) to find the SB^nS^{-1} in separate entities
6. Observe the behavior of the matrix that contains the n exponent as $\lim_{n\to\infty} n$. Usually this results in a part of the equation disappearing or exploding in value.

If you were to set an distribution vector as $\begin{bmatrix} a \\ b \end{bmatrix}$ or similar quantity, we can enable the analysis of proportions and what initial distributions of populations would cause an explosion in value or decrease. Note that we were to take any shortcuts in this process, we would miss any insight we would derive from the final product, so it is mandatory we do not skip any steps here.

18 Long list of matrix properties

If we let A be the $n \times n$ matrix corresponding to T, then this tells us that the following are all equivalent:[18]

1. T is invertible
2. For any $\mathbf{y} \in \mathbb{R}^n$, the system $A\mathbf{x} = \mathbf{y}$ (or equivalently, $T(\mathbf{x}) = \mathbf{y}$) has a unique solution.
3. The system $A\mathbf{x} = \mathbf{0}$ (or equivalently, $T(\mathbf{x}) = \mathbf{0}$) has the unique solution $\mathbf{x} = \mathbf{0}$.
4. A is invertible
5. $\mathrm{rref}(A) = I_n$
6. $\mathrm{rank}(A) = n$
7. $\mathrm{nullity}(A) = 0$
8. $\mathrm{im}(A) = \mathbb{R}^n$
9. $\ker(A) = \{\mathbf{0}\}$
10. $\det(A) \neq 0$
11. A has n linearly independent columns
12. The column vectors of A span $\mathbb{R}^n$
13. The column vectors of A form a basis for $\mathbb{R}^n$
14. A has n linearly independent rows
15. The row vectors of A span $\mathbb{R}^n$
16. The row vectors of A form a basis for $\mathbb{R}^n$
17. A does not have 0 as an eigenvalue

[18] Full credit to Professor Ziesler for compiling this list.