Linear Algebra Notes

Atticus Kuhn

May 11, 2023

Contents

1 March 20

Definition 1 *A **linear equation** of the variables $x_1, x_2, \ldots x_n$ is any equation of the form*

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

where $a_1, \ldots a_n$ are called the coefficents and b is called the constant.

Definition 2 *A **system of linear equation** is a set of 1 or more linear equations.*

Example 1

$$x + 2y = 3$$

is a linear system and

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

is a linear system. But

$$5\sqrt{x} - y = 1$$

is not a linear system.

Given a linear system, we want to find all substitutions which satisfy the linear system. A **solution set** is a set of n-tuples, all of which are valid solutions to a given system of linear equations. Our central problem is: given a linear system, find its solution set. Big problems

1. Existance Question: Do there exist any solutions? If there are, we call the system **consistent**. If there are no solutions, we call the solution **inconsistent**.

2. Uniqueness Question: How many solutions are there? If there is exactly 1 solution, we call that **unique**.

Theorem 1 *If a linear system has solutions, it either has 1 solution or infintely many solutions. A linear system cannot have 12 solutions.*

Example 2

$$x + 2y = 5, x + 4y = 6$$

has solution

$$x = -4, y = 9/2$$

This linear system corresponds to the matrix multiplication

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

We can turn this into an ***augmented matrix*** *to get*

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

We are allowed to perform row operations on the rows to get

$$\begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 9/2 \end{bmatrix} \implies \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 9/2 \end{bmatrix}$$

So $(x, y) = (-4, 9/2)$*. The solution is unique.*

Theorem 2 *We are allowed to do these elementary row operations*

Definition 3 *We say two systems of equations are* **equivalent** *if their solutions sets are equal.*

1. *replacement: replace one row with the sum of itself and another row*
2. *interchange: swap 2 rows*
3. *scaling: multiply all entries in a row by a constant*

The reason the elementary operations work is because they give an equivalent system, i.e. they do not change the solution set.

Convince yourself that each of the elementary operations do not change the solution set.

Example 3 *Let's solve this system*

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

The solution is

$$x = -3h/(k-3), y = \frac{(h+1)k - 3}{k - 3}, z = \frac{hk}{k - 3}$$

The system would have no solutions (inconsistent) iff $k = 3$ *and* $h \neq 0$*. The system has infinite solutions if* $k = 3$ *and* $h = 0$*. The system has 1 solution if* $k \neq 3$*.*

Definition 4 *A matrix is in* ***echelon form*** *iff*

1. *All nonzero rows are above any rows of all zeroes*

2. *each leading entry in a row is in a column to the right of the leading endtry of the row above it*

3. *All entries in a column below a leading entry are zeroes*

Definition 5 *A matrix is in **reduce row echelon form** iff*

1. *The leading entry in each nonzero row is* 1

2. *each leading* 1 *is the only nonzero entry in its column*

Example 4 *This matrix*

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

is in REF, but not in RREF.

Example 5 *This matrix*

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

is in REF, and it would be in RREF if we changed the 2 *to a* 1*.*

Theorem 3 *A given matrix has infinitely many equivalent REF, but only 1 RREF.*

Example 6 *Reduce this matrix*

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

Solution:

$$\begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

Definition 6 *If a matrix is in REF, the **pivots** are leading entries. The **pivot columns** are the columns which contain a pivot.*

Definition 7 *A **basic variable** is a variable which corresponds to a pivot column. A **free variable** is a variable which does not correspond to a pivot column.*

Example 7 *Given the matrix*

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

This matrix is in RREF. Variables x_1, x_2, x_4 are basic variables and variable x_3 is a free variable.

If the righmost column of A is a pivot column, then we cannot tell if $\mathbf{A}\vec{x} = \vec{b}$ is consistent or inconsistent.

Theorem 4 *If the rightmost column of* $\begin{bmatrix} \mathbf{A} & \vec{b} \end{bmatrix}$ *is a pivot column, then* $\mathbf{A}\vec{x} = \vec{b}$ *is not consistent.*

Theorem 5 *Even if you have a free variable, that doesn't guarantee infintely many solutions.*

If you have a free variable and the system is consistent, then the system has infinitely many solutions.

Example 8 *Given that the RREF is*

$$\begin{bmatrix} 1 & 2 & 3 & 0 & 0 & 6 \\ 0 & 0 & 0 & 1 & 0 & 7 \\ 0 & 0 & 0 & 0 & 1 & 8 \end{bmatrix}$$

Describe the solution set. It is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ 0 \\ 0 \\ 7 \\ 8 \end{bmatrix}$$

This solution set is infinitely big because it has free varaibles. This is called ***parametric vector form*** *of the solution set.*

2 March 22

Let's review what we learned last session about matricies and linear systems.

1. the RREF form of a matrix is unique (but the REF form is not).
2. the elementary row operations do not affect the solution set of a linear system.
3. pivot columns correspond to basic variables, non-pivot columns correpond to free variables.
4. every linear system has either 0, 1, or ∞ solutions.
5. an augmented matrix can be used to represent a linear system. It is the coefficent matrix concatenated with the constant vector.
6. We can express the solution set of a linear system in parametric vector form.
7. If a system is consistent and has a free variable, it has infinitely many variables (but even inconsistent systems may have free variables)
8. row operations do not change pivot columns.

Example 9 *Let*

$$p = (0, 1) \qquad q = (1, 1) \qquad r = (2, 2)$$

Prove there exists no line with all 3 points.

Prove there exists exactly one unique parabola with all 3 points
Solution:

The line has linear system

$$1 = 0a + b \qquad 1 = a + b \qquad 2 = 2a + b$$

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

So inconsistent.

For the parabola, it has the system

$$1 = 0a + 0b + c \qquad 1 = a + b + c \qquad 2 = 4a + 2b + c$$

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

i.e.

$$y = \frac{1}{2}x^2 - \frac{1}{2}x + 1$$

solution is unique because no free varaibles.

2.1 1.3 – Vector Equation

Definition 8 ***matrix multiplication by vector****: Given*

$$A = [\vec{a_1}, \vec{a_2}, \vec{a_3}, \ldots, \vec{a_n}]$$

and

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{bmatrix}$$

then

$$A\vec{v} = \vec{a_1}v_1 + \vec{a_2}v_2 + \cdots + \vec{a_n}v_n = \sum_{i=1}^{n} \vec{a_i}v_i$$

We call this a ***linear combination*** *of a_i. A linear combination is a sum of scalings of vectors, i.e. scale the vectors and then add them up.*

Example 10

$$\begin{bmatrix} 1 & 2 & 3 \\ 5 & 4 & 7 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a + 2b + 3c \\ 5a + 4b + 7c \end{bmatrix}$$

Definition 9 ***span***: *Given a set of vectors $S = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$, then span(S) is the set of all linear combinations of the vectors. In other words, it is the set of all vectors reachable by using only the vectors in S. In equational form,*

$$span(S) = \left\{ \sum_{i=1}^{n} c_i \vec{v}_i \;\middle|\; \forall c_i \in \mathbb{R} \right\}$$

Example 11 *In $\mathbb{R}^3$,*

$$span(\{\vec{v}\}) = \{k\vec{v} \quad | \quad k \in \mathbb{R}\}$$

This set is a line going through the origin (unliness $\vec{v} = \vec{0}$, then it's just the origin). The $\vec{0}$ vector is always in the span set.

$$span(\{\vec{u}, \vec{v}\})$$

this is a plane passing through the origin (unless the vectors are linearly dependent, then it's a line). It could be origin, or line thru origin, or plane thru origin.

$$span(\{\vec{u}, \vec{v}, \vec{w}\})$$

is either the origin, or a line thru the origin, or a plane thru the origin, or the entire space.

Example 12

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 1 & 0 \end{bmatrix}, \qquad v = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

Is $Ax = b$ consistent? This is equivalent to asking if b is a linear combination of the columns of A and it is equivalent to asking if b is an element of the spanning set of the columns of A.

This is not consistent because

$$\begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Thus, $b \notin span(A)$

Example 13

$$x + 2y + 3z = 4 \qquad 5x + 6y + 7z = 8$$

So

$$\begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

Theorem 6 *Asking whether the system $Ax = b$ is consistent is the same as asking if $b \in span(A)$*

3 March 23

3.1 1.4 – Matrix Equation

Example 14 *Let*

$$A = \begin{bmatrix} 1 & 4 & 5 \\ -3 & -1 & -14 \\ 2 & 8 & 10 \end{bmatrix}$$

Let $b \in \mathbb{R}^m$. Is $Ax = b$ consistent for any b?
Solution:
Not consistent because if you rref

$$\begin{pmatrix} 1 & 4 & 5 & b_1 \\ -3 & -11 & -14 & b_2 \\ 2 & 8 & 10 & b_3 \end{pmatrix}$$

you get the $-2b_1 + b_3 = 0$.

Now, find the set spanned by the column vectors of A.

(note that the spanning set is NOT $\mathbb{R}^3$, because the columns have a rank of 2). This is

$$span(A) = \left\{ \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \middle| -2b_1 + b_3 = 0 \right\} = \left\{ b_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + b_3 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} |\forall b_2, b_3 \in \mathbb{R} \right\} = span\left(\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \right\} \right)$$

You can do this is Sage with

```
A =  span(column_matrix([[1,4,5],[3,-11,-14], [2,8,10]]))
```

$$A = \mathrm{RowSpan}_Z \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix}$$

Theorem 7 *The following statements are logically equivalent: Let A be an $m \times n$ matrix and let b be an $m \times 1$ vector.*

1. *for all $b \in \mathbb{R}^m$, $Ax = b$ has a solution.*
2. *$span(A) = \mathbb{R}^m$*
3. *The columns of A span $\mathbb{R}^m$*
4. *A has 1 pivot position in every row.*
5. *b can be written as a linear combination of the columns of A.*
6. *$Ax = b$ is consistent for all $b \in \mathbb{R}^m$.*

Make sure you don't mess up the details. We want a pivot in every ROW of A, not every column of A.

Example 15 *Let's apply theorem 7.*

$$span\left(\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\} \right) \neq \mathbb{R}^2$$

because it doesn't have a pivot in the second row.

Example 16

$$span\left(\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}2\\1\end{bmatrix}\right\}\right)=\mathbb{R}^2$$

because there's a pivot in every row. Note that there is not a pivot in every column, but the theorem requires a pivot in every row.

Another mistake people make is thinking that the theorem applies if there is a pivot in every row of $[A \quad b]$. We are only concerned about a pivot in every row of A.

4 March 27

4.1 Homogenous System

Definition 10 *The linear system $Ax = b$ is called **homogenous** iff $b = \vec{0}$. A homogenous linear system is any system of the form $Ax = \vec{0}$.*

Theorem 8 *A homogenous linear system is always consistent because*

$$\mathbf{A}\vec{0} = \vec{0}$$

Definition 11 *We call $x = \vec{0}$ the **trivial solution** to the linear system $Ax = \vec{0}$.*

Theorem 9 *A homogenous linear system has infinite nontrivial solutions if and only if it contains free variables.*

Example 17 *consider this homogenous linear system*

$$\begin{bmatrix}1 & 2\\3 & 6\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

The solution set is

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

Example 18 *consider this homogenous linear system*

$$\begin{bmatrix}1 & 2\\3 & 6\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}4\\12\end{bmatrix}$$

Then

$$\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}-2y+4\\y\end{bmatrix}$$

This solution set cannot be written as a span because it doesn't contain $\vec{0}$.

The solution set of a homogenous system can always be expressed as the span of some vectors.

Example 19 *Here are 2 systems to consider.*

$$S_1 : x + 2y + 3z = 0$$

$$S_2 : x + 2y + 3z = 4$$

Compare and contrast the solution sets.
Solution:

The solution set for S_1 *is span* $\left\{ \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\}$. *The solution set for* S_2 *is*

$$\left\{ \begin{bmatrix} -2y - 3z + 4 \\ y \\ z \end{bmatrix} \middle| \forall y, z \in \mathbb{R} \right\}$$

Theorem 10 *The solution set to* $Ax = b$ *is parallel to the solution set of* $Ax = 0$. *You can think of the solution set of* $Ax = b$ *as being a constant vector translation from the solution set of* $Ax = 0$.

Suppose p *is any solution to the system* $\mathbf{A}\vec{x} = \vec{b}$. *Then, any other solution to* $\mathbf{A}\vec{x} = \vec{b}$ *can be written as* $\vec{p} + \vec{v}_h$, *where* $\vec{v}_h$ *is any solution to* $\mathbf{A}\vec{x} = \vec{0}$.

Warning: you can only think of the solution set of $\mathbf{A}\vec{x} = \vec{b}$ as a translation of the solution set of $\mathbb{A}\vec{x} = \vec{0}$ is the linear system is consistent. Otherwise, it's solution set is the empty set.

4.2 1.6 – Applications

There are a bunch of applications of matricies in the textbook.

Example 20 *Find the general traffic pattern in the freeway network shown in the figure*
Describe the general traffic pattern when the road x_1 *is closed.*
When $x_4 = 0$ *what is the minimum value of* x_1*?*
Solution:
For the solution, we use the general principle that (flow in) - (flow out) = 0.

$$A = \begin{bmatrix} 1 & 0 & -1 & -1 & 0 & 40 \\ -1 & -1 & 0 & 0 & 0 & -200 \\ 0 & 1 & 1 & 0 & -1 & 100 \\ 0 & 0 & 0 & 1 & 1 & 60 \end{bmatrix}$$

$$\begin{aligned} A : x_1 &= x_3 + x_4 + 40 \\ B : 200 &= x_1 + x_2 \\ C : x_2 + x_3 &= 100 + x_5 \\ D : x_4 + x_5 &= 60 \end{aligned}$$

This reduces to

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 1 & 100 \\ 0 & 1 & 1 & 0 & -1 & 100 \\ 0 & 0 & 0 & 1 & 1 & 60 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

So x_1, x_2, x_4 *are basic variables and* x_3, x_5 *are free variables. The solution set is*

$$\left\{ \begin{bmatrix} x_3 - x_5 + 100 \\ -x_3 + x_5 + 100 \\ x_3 \\ -x_5 + 60 \\ x_5 \end{bmatrix} | \forall x_3, x_5 \in \mathbb{R} \right\}$$

If $x_4 = 0$, then $x_5 = 60$, so we get the solution set

$$\left\{ \begin{bmatrix} x_3 + 40 \\ -x_3 + 160 \\ x_3 \\ 0 \\ 60 \end{bmatrix} | \forall x_3 \in [0, 160] \right\}$$

The minimum value of x_1 is 40.

4.3 1.7 – Linear Independence

Definition 12 *Given a set of vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3, \ldots$, we say that the set is **linearly independent** if and only if the only solution to*

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 + \cdots = \vec{0} \qquad c_i \in \mathbb{R}$$

is the trivial solution.

*We say the set is **linearly dependent** if and only if there is a solution to the equation*

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 + \cdots = \vec{0} \qquad c_i \in \mathbb{R}$$

which is not the trivial solution.

Theorem 11 *The columns of $\mathbb{A}$ are linearly independent if and only if the only solution to*

$$\mathbf{A}\vec{x} = \vec{0}$$

is the trivial solution $\vec{x} = \vec{0}$.

If one vector lies in the plane spanned by the other two, then the vectors are linearly dependent.

Theorem 12 *To find out if a set of vectors is linearly dependent, RREF their matrix.*

5 March 29

No class due to doctor's appointment

6 March 30

6.1 1.7 – Linearly Indepedence

Definition 13 *A **linearly independent set** is a set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots \vec{v}_n\}$ such that*

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots c_3\vec{v}_3 = 0 \iff c_1 = c_2 = c_3 = \cdots = 0$$

In other words, we say the columns of $\mathbf{A}$ are linearly independent iff

$$\mathbf{A}\vec{x} = \vec{0} \iff \vec{x} = \vec{0}$$

There are some special cases to easily determine if a set is linearly independent or dependent.

1. The empty set $\{\}$ is linearly independent.

2. The singleton set $\{\vec{v}\}$ is linearly independent if and only if $\vec{v} \neq \vec{0}$

3. The two-element set $\{\vec{v}_1, \vec{v}_2\}$ is linearly independent iff $\nexists k \in \mathbb{R}, \vec{v}_1 = k\vec{v}_2$.

4. Given a set of p vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ in $\mathbb{R}^n$ and $p > n$, then the set is linearly dependent.

5. If a set contains the $\vec{0}$ vector, then it is linearly dependent.

Practically, to find out if vectors are linearly independent, RREF the matrix. If the RREF has a pivot in every row, then the vectors are linearly independent. Otherwise, the vectors are linearly dependent.

Example 21 *Are the column vectors in this matrix linearly independent? Just RREF it.*

6.2 1.8 – Linear Transformations

Definition 14 *A* ***transformation*** *is a function $T : \mathbb{R}^n \to \mathbb{R}^m$. We call $\mathbb{R}^n$ the domain and we call $\mathbb{R}^m$ the codomain.*

Definition 15 *Given an $m \times n$ matrix $\mathbf{A}$, a* ***Matrix Transformation*** *is a linear transformation*

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

$$T(\vec{x}) = \mathbf{A}\vec{x}$$

We call $\mathbf{A}$ the ***standard matrix*** *of T.*

Example 22 *For example, given the matrix*

$$\mathbf{A} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$$

The matrix transformation of $\mathbf{A}$ is

$$T : \mathbb{R}^2 \to \mathbb{R}^3$$

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$

Definition 16 *Given a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, we say a transformation is a* ***linear transformation*** *iff it satisfies 2 axioms:*

1. *Vector addition:*

$$T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) \qquad \forall \vec{u}, \vec{v} \in \mathbb{R}^n$$

2. *Scaling:*

$$T(c\vec{v}) = cT(\vec{v}) \qquad \forall c \in \mathbb{R} \quad \forall \vec{v} \in \mathbb{R}^n$$

Theorem 13 *If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then*

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

because of the scaling rule.

By the contrapositive, if $T(\vec{0}) \neq \vec{0}$, then T is not a linear transformation.

Theorem 14 *All matrix transformations are also linear transformations. Given a linear transformation and a basis, there is a unique matrix transformation corresponding to the linear transformation.*

In other words, a matrix transformation is basically the same thing as a linear transformation

Theorem 15 *The derivative $\frac{d}{dx}$ from the set of differentiable functions to functions. To see why, it satisfies the axioms*

$$\frac{d}{dx}(f(x)+g(x)) = \frac{d}{dx}(f(x)) + \frac{d}{dx}(g(x))$$

$$\frac{d}{dx}(cf(x)) = c\frac{d}{dx}(f(x))$$

7 April 12

Theorem 16 *If $T : \mathbb{R}^n \to \mathbb{R}^p$ is a linear transformation, then there exists a matrix A called the* ***standard matrix*** *defined by*

$$A = [T(e_1), T(e_2), T(e_3), T(e_4), \dots, T(e_n)]$$

Where $e_1, e_2, \dots, e_n$ are the ***basis vectors****.*

The way to see this is that any arbitrary vector $\vec{v} \in \mathbb{R}^n$ can be written as a linear combination of the basis vectors as $\vec{v} = c_1\vec{e}_1 + \cdots c_n\vec{e}_n$.

Example 23 *The matrix of a vertical reflection in $\mathbb{R}^2$ is*

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

The matrix of a horizontal contraction (or expansion) by a factor of k is

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

The matrix of the projection onto the y axis is

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

The matrix of the horizonal shear by a factor of k is

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

The matrix of a θ counter-clockwise about the origin is

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Definition 17 *Let $f : X \to Y$. f is called **surjective** iff every element of Y is hit.*

*f is called **injective** iff $f(x_1) = f(x_2) \implies x_1 = x_2$.*

Surjective functions are also called "onto functions". Injective functions are also called "one-to-one functions"

Definition 18 *We say $f : X \to Y$ is **bijective** iff it is both surjective and injective.*

Theorem 17 *Let $T : \mathbb{R}^n \to \mathbb{R}^p$ be a linear transformation, and let T have standard matrix A.*

T is injective/one-to-one iff the equation $\boldsymbol{A}\vec{x} = \vec{0}$ only has the solution $\vec{x} = \vec{0}$. This means that A has a pivot in every column.

In other words, $T(\vec{x}) = \vec{0} \implies \vec{x} = 0$.

The following are equivalent

1. *T is an injective/one-to-one function*
2. *A has a pivot in every column*
3. *A has no free variables*
4. *$T(\vec{x}) = \vec{0} \implies \vec{x} = \vec{0}$*
5. *The columns of $\mathbf{A}$ are linearly independent.*

Theorem 18 *Let $T : \mathbb{R}^n \to \mathbb{R}^p$ be a linear transformation with matrix $\mathbf{A}$. The follow are equivalent*

1. *T is a surjection / onto function*
2. *A has a pivot in every row*
3. *The system $\mathbf{A}\vec{x} = \vec{b}$ is consistent for every $\vec{b} \in \mathbb{R}^p$*
4. *$span(A) = \mathbb{R}^p$.*

Example 24 *Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by*

$$T(\vec{x}) = \begin{bmatrix} 1 & -1 & 3 \\ -2 & 2 & -6 \end{bmatrix} \vec{x}$$

Find Range(T)?
Solution:
Range$(T) = span(\begin{bmatrix} 1 \\ -2 \end{bmatrix})$
Find the set of all pre-images of $\begin{bmatrix} 2 \\ -4 \end{bmatrix}$?
Solution:

$$S = \left\{ k \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + l \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} | \forall k, l \in \mathbb{R}^3 \right\} = span \left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right) + \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

Is T onto/surjective?
Solution:
no.
Is T one-to-one/injective?
Solution:
no.

8 April 13

Definition 19 *Given a matrix A, we call A^{-1} the* ***inverse*** *of A iff $AA^{-1} = I$ If A has an inverse, then A is called* ***invertable****.*

If A is invertiable, then A^{-1} is its unique inverse.
We only talk about inverses in the context of square matricies.
For a 1×1 matrix, the inverse is

$$a^{-1} = \frac{1}{a}$$

For a 2×2 matrix, the inverse is

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

Theorem 19 *If A is invertable, then the system*

$$\mathbf{A}\vec{x} = \vec{b}$$

has the unique solution

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

Example 25 *Solve this system*

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

Theorem 20

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

Theorem 21 *A matrix A is invertible if and only if it is row-equivalent (by elementary row operatins) to the identity matrix I.*

Theorem 22 ***Finding Inverse****: To find the inverse on the $n \times n$ matrix A, perform row operations $[A \quad I_n]$ to get to $[I_n \quad A]$ In other words*

$$[A \quad I_n] \sim [I_n \quad A]$$

Example 26

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

$$\sim \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{3}{2} & 1 & 0 \end{pmatrix}$$

Theorem 23 ***Invertible Matrix Theorem****: The following statements are all logically equivalent*

1. *A is an invertiable $n \times n$ matrix*
2. *$A \sim I_n$*
3. *A has n pivot positions*
4. *The only solution $\mathbf{A}\vec{x} = \vec{0}$ is the trivial solution $\vec{x} = \vec{0}$*
5. *The columns of A form a linearly inpendent set/basis of $\mathbb{R}^n$*
6. *The linear transformation $T(\vec{x}) = \mathbf{A}\vec{x}$ is one-to-one/injective*
7. *The linear transformation $T(\vec{x}) = \mathbf{A}\vec{x}$ is onto/surjective*
8. *The equation $\mathbf{A}\vec{x} = \vec{b}$ has exactly one solution $\vec{x} = \mathbf{A}^{-1}\vec{b}$ for all $\vec{b} \in \mathbb{R}^n$*

8.1 Chapter 3 – Determinants

Definition 20 *Given a matrix A, the **sub-matrix** A_{ij} is the matrix obtained by deleting row i and column j*

Theorem 24 ***Laplace Expansion**: The determinant can be calculated by*

$$|A| = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}|$$

*Some people call $(-1)^{i+j}|A_{ij}| = c_{ij}$ the **co-factor**.*

Note that Laplace expansion is very slow, it has time complexity $O(n!)$. If you want to calculate the determinant quickly, elementary row operations can calculate the determinant in $O(n^3)$.

Theorem 25 *If A is a triangular matrix, then $|A|$ is just the product of the diagonal entries.*

$$|A| = \prod_{i=1}^{n} a_{ii}$$

It is a common mistake that $A \sim B \implies |A| = |B|$, but this is false.

Theorem 26 *Each of the elementary row operations has an effect on the determinant*
***scaling**: multiplying a row by k multiplies the determinant by k*
***interchange**: swapping two rows multiplies the determinant by -1*
***replacement**: adding a row to another row doesn't change the determinant*

Theorem 27 ***Properties of determinant***

$$|AB| = |A||B|$$

$$|A^{-1}| = |A|^{-1}$$

$$|A^T| = |A|$$

9 April 17

Let's review determinant.

$$\begin{vmatrix} d-2g & e-2 & g-2i \\ 3a & 3b & 3c \\ -a+4g & -b+4h & -c+4i \end{vmatrix}$$

$$= -3 \cdot 4 \cdot \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

Theorem 28 ***Cramers Rule****: If A is an invertiable $n \times n$ matrix, then the solution to the linear system*

$$\mathbf{A}\vec{x} = \vec{b}$$

has entries

$$\vec{x}_i = \frac{\det \mathbf{A}_i(\vec{b})}{\det \mathbf{A}} \qquad \forall i \in \{1, 2, \ldots, n\}$$

Where

$$\mathbf{A}_i(\vec{b}) := [\vec{a}_1 \ldots \vec{b} \ldots \vec{a}_n]$$

Theorem 29 *The determinant is the (signed) volume or area of the parallepiped or parallelogram with edges given by the columns of* $\mathbf{A}$*.*

9.1 Practice Knowledge Check Chapter 1

1. Consider the following System

$$x + y + z = 2 \quad x - y + z = h \quad x - y + kz = 3$$

   (a) apply elementary row oeprations to write in echelon form.

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

   (b) Determine the values of k, h such that the system has 0, 1, or ∞ solutions? 0 solutions iff $k = 1 \wedge h \neq 3$. 1 solution iff $k \neq 1$, ∞ solutions iff $k = 1 \wedge k = 3$

   (c) When the system has ∞ solutions, find the general solution in parametric vector form

$$\begin{bmatrix} 5/2 \\ -1/2 \\ 0 \end{bmatrix} + span \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

   (d) describe the above set geometrically? A line passing through $\begin{bmatrix} 5/2 \\ -1/2 \\ 0 \end{bmatrix}$ and parallel to $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ in $\mathbb{R}^3$

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

   (a) Do the columns of A span $\mathbb{R}^3$? No

   (b) Find the span of the columns of A?

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

   (c) Are the columns of A linearly independent? No

   (d) Write a dependence relation among the columns of A?

$$\vec{a}_3 = -\frac{1}{2}\vec{a}_1 + \frac{1}{2}\vec{a}_2$$

   (e) Geometrically describe $span\,\{a_1, a_2, a_3\}$? A place in $\mathbb{R}^3$ passing through the origin $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, and $\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$.

3. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation that maps $(1,1,1) \mapsto (0,1)$, $(0,1,1) \mapsto (1,1)$, and $(0,1,1) \mapsto (1,-1)$

   (a) Is T a matrix transformation?
   Yes

   (b) Find the matrix of T?

$$T(\vec{x}) = \begin{bmatrix} -1 & 1 & 0 \\ 2 & 1 & -2 \end{bmatrix} \vec{x}$$

   (c) Find the image of (x, y, z)?

$$\begin{bmatrix} -x + y \\ 2x + y - 2z \end{bmatrix}$$

   (d) Prove that T is a linear transformation?

$$T(\vec{a} + \vec{b}) = T(\vec{a} + \vec{b})$$
$$T(c\vec{a}) = cT(\vec{a})$$

   (e) Find the pre-image of $(1, 4)$

$$S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 2/3 \\ 2/3 \\ 1 \end{bmatrix} z \,\middle|\, \forall z \in \mathbb{R} \right\}$$

(f) Is T one-to-one/injective?
no

(g) Is T onto/surjective?
yes

10 April 20

Definition 21 *A set V is called a* ***vector space*** *iff it satisfies the following axioms*

$$\begin{aligned} \vec{u}+\vec{v} &= \vec{v}+\vec{u} \\ (\vec{u}+\vec{v})+\vec{w} &= \vec{u}+(\vec{v}+\vec{w}) \\ \vec{v}+\vec{0} &= \vec{v} \\ \vec{u}+(-\vec{u}) &= 0 \\ c(\vec{u}+\vec{v}) &= c\vec{u}+c\vec{v} \\ (c+d)\vec{u} &= c\vec{u}+d\vec{u} \\ c(d\vec{u}) &= (cd)\vec{u} \\ 1\vec{u} &= \vec{u} \end{aligned}$$

Note that the empty set is *not* a vector space because it does not contain $\vec{0}$.
The **zero space** $\{\vec{0}\}$ is a vector space.

Example 27 *Examples of vector spaces*

1. *the zero space* $\{\vec{0}\}$.
2. $\mathbb{R}^n$ *for any integer* n.
3. $\mathbb{P}$, *the set of polynomials with real coefficients*
4. $\mathbb{P}_n$, *the set of polynomials with real coefficients with degree at most* n.
5. *signals:* $\mathbb{S}$ *the set of lists of real numbers that extend infinitely in both directions* $(\ldots, x_{-1}, x_0, x_1, \ldots)$
6. *function space: real valued functions from any set* X *to the real numbers* $\mathbb{R}^n$
7. *Any plane passing through the origin.*

11 April 24

Nothing happend.

12 April 26

Here are some more examples of vector spaces

1. The set of $m \times n$ matricies
2. The set of integrable functions from a set X to $\mathbb{R}$.

3. For a fixed $m \times n$ matrix B, the set $\{\mathbf{A}|\forall \mathbf{A} \in \mathbb{R}^{p\times m}, \mathbf{AB} = \mathbf{0}\}$ (because it is the kernel/nullspace of $\mathbf{B}$).

4. For a fixed $m \times n$ matrix B, the set $\{\mathbf{A}|\forall \mathbf{A} \in \mathbb{R}^{p\times m}, \mathbf{AB} = \mathbf{A}\}$ (because it is an invariant subspace)

Definition 22 *Given a vector space V, we say that a set W is called a **vector subspace** of V if and only if*

1. $W \subseteq V$
2. $\vec{0} \in W$
3. *addition is closed under W, meaning that $\forall \vec{u}, \vec{v} \in W \implies (\vec{u} + \vec{v}) \in W$*
4. *W is closed under scalar multiplication, meaning that $\forall c \in \mathbb{R} \quad \forall \vec{u} \in W \implies k\vec{u} \in W$.*

Example 28 *Here is a list of examples of vector subspaces*

1. *For any vector space V, the zero subspace $\{\vec{0}\}$ is a subspace of V*
2. $\left\{ \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} | \forall a, b \in \mathbb{R} \right\}$ *is a vector subspace of* $\mathbb{R}^3$
3. $\left\{ \begin{bmatrix} x \\ y \\ x+y \end{bmatrix} | \forall x, y \in \mathbb{R} \right\}$ *is a vector subspace of* $\mathbb{R}^3$

Theorem 30 *If W is a vector subspace of V, then W is automatically a vector space (you don't need to check the axioms).*

Theorem 31 *If V is a vector space and $S \subseteq V$, then $span(S)$ is a vector subspace of V.*

Definition 23 *Given a linear transformation $T : V \to W$, we define the **null space** as*

$$Nul(T) := \left\{ \vec{x} | \vec{x} \in V, T(\vec{x}) = \vec{0} \right\} = T^{-1}(\vec{0})$$

It is a theorem that $Nul(T)$ is a vector subspace of V. The null space is sometimes called the kernel of T.

Definition 24 *Given a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, we define the **column space** as*

$$Col(A) := span(A) = \{\mathbf{A}\vec{x} | \forall \vec{x} \in \mathbb{R}^n\} = \left\{ \vec{b} | \quad \exists \vec{x} \in \mathbb{R}^{m\times n}, \mathbf{A}\vec{x} = \vec{b} \right\}$$

*The column space is also called the **range** of the linear transformation $\vec{x} \mapsto \mathbf{A}\vec{x}$.*

It is a theorem that $Col(\mathbf{A})$ is a vector subspace of $\mathbb{R}^m$.

Example 29 *Let*

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

Find the column space and the null space.

$$Col(A) = span(\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \})$$

$$Nul(A) = span(\{ \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix} \})$$

13 April 27

Example 30 *Consider* $T : \mathbb{P}_2 \to \mathbb{R}^{2\times 3}$

$$T(p(t) = \begin{bmatrix} p(0) & p(0) & p(1) \\ p(1) & p(1) & p(0) \end{bmatrix}$$

Prove that T is a linear transformation.
T satisfies the linear transformation axioms. T preserves vector addition and preserves scalar multiplication.
Find $Ker(T)$.

$$Ker(T) = span\{\lambda x.x^2 - x\}$$

Is T a matrix transformation? According to YHPL, no.
Find Range(T).

$$Range(T) = span\left\{ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}, \right\}$$

Find all pre-images of $\begin{bmatrix} 2 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$*?*

$$T^{-1}(\begin{bmatrix} 2 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}) = \{\lambda x.ax^2 - (1+a)x + 2 | \forall a \in \mathbb{R}\}$$

Is T injective/one-to-one? No because nullspace is not trivial.
Is T surjective/onto? Impossible because the codomain is larger than the domain.

13.1 4.3 – Basis

Definition 25 *Let V be a vector space. We say that $B \subseteq V$ is called a* ***basis*** *of V if and only if*

1. *$span(B) = V$*

2. *B is linearly independent*

In other words, a basis is a smallest spanning set.

Note that a basis of V is not unique: V usually has infinitely many different bases. For example, $\{1, x, x^2\}$ is a basis of $\mathbb{P}_2$, but $\{1+x, x^2, x+x^2\}$ is also a basis of $\mathbb{P}_2$.

Example 31 *Here are some examples of a basis*

1. $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ *is a basis of* $\mathbb{R}^3$

2. $\{1, x, x^2\}$ *is a basis of* $\mathbb{P}_2$

Theorem 32 *If S spans V, then any superset of S also spans V.*

If S is a linearly independent set, then any subset of S is also linearly indepenent.

If S is a linearly dependent set, then any superset of S is also linearly depenent.

If B is a basis of V, then you cannot add or remove any vector from B and still have it be a basis of V.

14 May 1

Example 32 *Does the set $\{1+x, x^2, x+x^2\}$ span $\mathbb{P}_2$? Yes.*

To see why, we need to show that an abitrary vector, call it $a+bx+cx^2$ can be written as a linear combination of the basis. We want to show that $\exists k, l, m \in \mathbb{R}$

$$k(1+x) + lx^2 + m(x+x^2) = a + bx + cx^2 \implies a + bx + cx^2 = k + (k+m)x + (l+m)x^2$$

So we get the linear system

$$\begin{aligned} l + m &= a \\ k + m &= b \\ k &= c \end{aligned}$$

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

Be warned that the pivot columns of A form a basis of $Col(A)$, not the pivots in the reduced matrix.

14.1 4.4 – Coordinate Systems

Theorem 33 ***Unique representation theorem**: If B is a basis of V, then every vector $v \in V$ can be uniquely represented as a linear combination of B. In other words, there exist unique scalars $c_1, c_2, \dots c_n$ such that*

$$v = c_1\vec{b}_1 + c_2\vec{b}_2 + \cdots + c_n\vec{b}_n$$

There is one and only one set of scalars $c_1, \dots, c_n$ to satisfy the equation.

Definition 26 *With respect to a given basis B, we define the* ***coordinates*** *relative to B to a notation as*

$$[\vec{x}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}_B = c_1\vec{b}_1 + c_2\vec{b}_2 + \cdots + c_n\vec{b}_n$$

This notation is well-defined because of the unique representation theorem.

A vector space may have many different bases.

To change a coordinate vector from B_1 to B_2 there exists a matrix called the **change of basis** matrix which changes the basis.

Example 33 *Let $a = \langle 6, 5\rangle$ in $\begin{bmatrix} 6 \\ 5 \end{bmatrix}$ is the coordinate vector in the elementary basis. If we set our basis to $\{\langle 3, 1\rangle, \langle 0, 1\rangle\}$, then the coordinate vector is $\begin{bmatrix} 2 \\ 3 \end{bmatrix}_B$ because $\begin{bmatrix} 2 \\ 3 \end{bmatrix}_B = 2\langle 3, 1\rangle + 3\langle 0, 1\rangle = \langle 6, 5\rangle$*

In this case, the change-of-basis matrix $\begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix}$.

$$\begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Example 34 *Let $v = \mathbb{P}_2$ and let $B = \{1 + t, t^2, t + t^2\}$.*

$$[1 + 2t + 3t^2]_B = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

because

$$1 + 2t + 3t^2 = 1(1 + t) + 2(t^2) + 1(t + t^2)$$

Theorem 34

$$\mathbb{P}_n \cong \mathbb{R}^{n+1}$$

Because we haev the bijective linear mapping (called a vector-space isomorphism) one example is

$$T(a + bx + cx^2) \mapsto \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Example 35 *Let $B = \left\{ \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \right\}$. Find the coordinate vector of $\begin{bmatrix} 9 \\ 13 \\ 15 \end{bmatrix}$ in span(B).*

$$\begin{bmatrix} 9 \\ 13 \\ 15 \end{bmatrix}_B = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

Becuase

$$3 \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 13 \\ 15 \end{bmatrix}$$

Definition 27 *If B is a basis of V, then the* ***dimension*** *of V is the length/cardinality of B.*

No matter which basis we choose, each basis of V will have the same length, which is the basis of V.

Example 36 *The vector space $\mathbb{R}^n$ is n-dimensional because any basis of $\mathbb{R}^n$ has n elements.*

Example 37 *The trivial vector space $\{\vec{0}\}$ is 0-dimensional because the empty list $\{\}$ spans the vector space. $\vec{0}$ is the only 0-dimensional vector space.*

Theorem 35 *If V is a vector space of dimension p, then any linearly independent set of size p is automatically a basis of V.*

Theorem 36

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

$$dim(\mathbb{R}^n) = n$$

$$dim(\mathbb{P}_n) = n + 1$$

14.2 4.6 – Rank

Definition 28 *Given a linear transformation $T : V \to W$ we define the* ***rank*** *as*

$$rank(T) := dim(Range(T))$$

Theorem 37 *To calculate the rank of a matrix, the rank is the number of pivot columns, which is equal to the number of pivot rows.*

Definition 29 *Given a linear transformation $T : V \to W$, we define the* ***nullity*** *as*

$$dim(ker(T))$$

Theorem 38 ***rank-nullity theorem****: If $T : V \to W$ is a linear transformation, then the rank and the nullity are related by the following equation*

$$dim(range(T)) + dim(ker(T)) = dim(V)$$

To prove this, the rank is the number of pivot columns, and the nullity is the number of nonpivot columns, and $dim(V)$ is the total number of columns.

15 May 3

Definition 30 *If $A \in \mathbb{R}^{n\times n}$ is a matrix and $\vec{x} \in \mathbb{R}^n$ is a vector with $\vec{x} \neq \vec{0}$ and $\lambda \in \mathbb{R}$ is a scalar and*

$$\mathbf{A}\vec{x} = \lambda\vec{x}$$

then we say $\vec{x}$ is an ***eigenvector*** *of* **A** *and we say that λ is the* ***eigenvalue*** *of $\vec{x}$.*

Note that only square matricies have eigenvalues.

Theorem 39 *The folllowing are logically equivalent*

$$\begin{aligned} &\mathbf{A}\vec{x} = \lambda\vec{x} \\ \iff &(\mathbf{A} - \lambda I)\vec{x} = 0 \\ \iff &det(\mathbf{A} - \lambda I) = 0 \end{aligned}$$

Definition 31 *Given $\mathbf{A} \in \mathbb{R}^{n\times n}$, we define the* ***characteristic polynomial*** *of $\mathbf{A}$ as*

$$p(\lambda) = \det(\mathbf{A} - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the n eigenvalues of $\mathbf{A}$. The characteristic polynomial is an n-degree polynomial of λ.

Theorem 40 *If $\mathbf{A} \in \mathbb{R}^{n\times n}$, then $\mathbf{A}$ has exactly n eigenvalues (counting with multiplicity).*

Definition 32 *Given a fixed eigenvalue $\lambda \in \mathbb{R}$ of $\mathbf{A} \in \mathbb{R}^{n\times n}$, we define the* ***eigenspace*** *of λ as*

$$\{\vec{x} : \mathbf{A}\vec{x} = \lambda\vec{x}\}$$

Note that the eigenspace is a vector subspace of $\mathbb{R}^n$ because it is the nullspace of $(\mathbf{A} - \lambda I)\vec{x} = \vec{0}$. To find the eigenspace of λ, solve the homogenous linear equation $(\mathbf{A} - \lambda I)\vec{x} = \vec{0}$

16 May 4

Definition 33 *If $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$, we say A and B are* ***similar*** *if and only if there exists a matrix $P \in \mathbb{R}^{n\times n}$ such that*

$$PB = AP$$

P is called the change-of-basis matrix

Definition 34 *We say a matrix $A \in \mathbb{R}^{n\times n}$ is* ***diagonalizable*** *if it is similar to a diagonal matrix, i.e*

$$AP = PD$$

equivalently

$$A = PDP^{-1}$$

for some diagonal matrix D the diagonal entries of D are the eigenvalues of A.

The diagonal matrix is the ***eigenvector basis****.*

Theorem 41 *A is diagonalizable iff it has n linearly independent eigenvectors.*

Theorem 42 *two matricies are similar if and only if they have the same eigenvalues with the same multiplicties.*

Theorem 43 *If A has n distinct eigenvalues, then A is diagonalizable. (warning: the converse does not always hold.)*

Definition 35 *A* ***stochastic matrix*** *represents a markov chain with a matrix. Please note that the COLUMN vectors represent the probability vectors. The sum of the column vectors must be 1. a_{rc} represents the probability that we will go from state r to state c. The top represents now, the right represents next.*

The ***state vector*** *is a column vector which represnets the current probabilities.*

Theorem 44 *$\lambda = 1$ is an eigenvalue of all stochastic matricies. Use this fact to easily factor the characteristic polynomial of a stochastic matrix. This corresponds to the steady-state vector.*

Example 38 *Here is an example of a stochastic matrix*

$$A = \begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix}$$

The steady state vector is

$$\vec{x} = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$$

The eigenvalues are

$$\lambda = 1, \lambda = 0.7$$

The diagonalization is

$$\begin{bmatrix} .9 & .2 \\ .1 & .8 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & .7 \end{bmatrix} (1/3 \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix})$$

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

Theorem 45 *It doesn't matter the inital state vector after a long time, all the columns of M^{∞} are the same.*

Definition 36 *The **steady state** vector is the fixedpoint of the markov chain, i.e. the probabilities don't change. If $\mathbf{A}\vec{x} = \vec{x}$ then $\vec{x}$ is called the steady-state vector of $\mathbf{A}$.*

In other words, a steady state vector is an eigenvector for $\lambda = 1$.

To find the steady-state vector, solve the homogenous system

$$(\mathbf{M} - \mathbf{I})\vec{x} = 0$$

17 May 8

17.1 Chapter 6 – Least Squares

Even if the equation $\mathbf{A}\vec{x} = \vec{b}$ is inconsistent, we can still project $\vec{b}$ onto $span(\mathbf{A})$. We have that $\hat{b} = proj_{\mathbf{A}}(\vec{b})$.

Definition 37 *Given $\vec{u}, \vec{v} \in W$, the **inner product** is any function that satisfies the following axioms*

1. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
2. $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$
3. $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot) = \vec{u} \cdot (c\vec{v})$
4. $\vec{u} \cdot \vec{u} \geq 0$

The most important inner product is called the **dot product** and is defined as

$$\vec{u} \cdot \vec{v} = \sum_{i=1}^{n} u_i v_i$$

Definition 38 *The **norm** of a vector is*

$$\| \cdot \| : \mathbb{R}^n \to \mathbb{R}$$
$$\|\vec{u}\| := \sqrt{\vec{u} \cdot \vec{u}}$$

Definition 39 *The **distance** between $\vec{u}$ and $\vec{v}$ is*

$$dist(\vec{u}, \vec{v}) := \|\vec{u} - \vec{v}\|$$

Theorem 46 ***orthogonality**: Two vectors are said to be orthogonal if and only if their dot products are* 0

$$\vec{u} \perp \vec{v} \iff \vec{u} \cdot \vec{v} = 0$$

Theorem 47 ***pythagorean theorem**:*

$$\vec{u} \perp \vec{v} \iff \|\vec{u}\|^2 + \|\vec{v}\|^2 = \|\vec{u} + \vec{v}\|^2$$

Theorem 48 $\vec{0}$ *is orthogonal to every vector.*

Definition 40 ***perpendicular to vector space** Let W be a vector subspace of V and let $\vec{u} \in V$. We say $\vec{u}$ is perpendicular to W iff*

$$\vec{u} \perp W \iff \vec{u} \in W^{\perp} \iff \forall \vec{v} \in W, \vec{u} \perp \vec{v}$$

Definition 41 ***orthogonal complement**: Let W be a vector subspace of V. The orthogonal complement of W is*

$$W^{\perp} := \{\vec{u} : \forall \vec{v} \in W, \vec{u} \perp \vec{v}\}$$

Definition 42 *If $S \subseteq \mathbb{R}^n$, then S is called an **orthogonal subset** of $\mathbf{R}^n$ iff*

$$\forall i, j \in [n] \qquad S_i \cdot S_j = \delta_{ij}$$

Theorem 49 *An orthogonal subset S is linearly independent iff $\vec{0} \notin S$*

Theorem 50 *If $\{\vec{v}_1, \dots \vec{v}_n\}$ is an orthogonal basis of V, and $y \in V$, then*

$$\vec{y} = \sum_{i=1}^{n} c_i \vec{v}_i$$

where

$$c_i = \frac{\vec{y} \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i}$$

Definition 43 *A set $S \subset V$ is called an **orthogonal basis** of V iff S is a basis of V and S is an orthogonal subspace of V.*

Theorem 51

$$\forall \mathbf{A} \in \mathbb{R}^{r \times c} \qquad (Row(A))^{\perp} = Null(A)$$
$$\forall \mathbf{A} \in \mathbb{R}^{r \times c} \qquad Col(A)^{\perp} = Null(A^{\top})$$

Theorem 52 ***orthogonal decomposition**: If W is a vector subspace of $\mathbb{R}^n$, then every $y \in \mathbb{R}^n$ can be uniquely written as*

$$\vec{y} = \hat{y} + \vec{z}$$

Where $\hat{y} \in W$ and $Z \in W^{\perp}$

18 May 10

Example 39 *Here is an application of orthogonal projection in math.*

Find the shortest possible distance from $(2,1)$ *to the line passing through* $(0,0)$ *and* $(1,-3)$*?*

Let $\vec{p} = \langle 2,1 \rangle, \vec{q} = \langle 1,-3 \rangle$.

$$\hat{p} = proj_q(p) = \frac{p \cdot q}{q \cdot q} q = -\frac{1}{10}\langle 1,-3 \rangle$$

The orthogonal decomposition is

$$\vec{p} = \hat{p} + \vec{z}$$

so

$$\langle 2,1 \rangle = -\frac{1}{10}\langle 1,-3 \rangle + \frac{7}{10}\langle 3,1 \rangle$$

So the distance is

$$d = \|\vec{z}\| = \|\frac{7}{10}\langle 3,1 \rangle\| = \frac{7}{\sqrt{10}}$$

Definition 44 *We say* B *is an* ***orthonormal basis*** *if*

$$\forall i,j \qquad b_i \cdot b_j = \delta_{ij}$$

Basically all the vectors are unit vectors and all the vectors are pairwise orthogonal.

Definition 45 *We say that* M *is an* ***orthogonal matrix*** *iff*

$$M^T = M^{-1}$$

Another way of saying this is

$$M^T M = M M^T = I$$

Theorem 53 *If* M *is an orthogonal matrix, then* M *is always invertible.*

18.1 6.4 – Geometric Interpretation of Orthogonal Projection

Theorem 54 ***orthogonal decomposition****: If* W *is a vector subspace of* $\mathbb{R}^n$*, then every* $y \in \mathbb{R}^n$ *can be uniquely written as*

$$\vec{y} = \hat{y} + \vec{z}$$

Where $\hat{y} \in W$ *and* $z \in W^\perp$

And if W *has an orthogonal basis* $B = \{\vec{u}_1, \ldots, \vec{u}_p\}$*, the coordinates of* $\hat{y}$ *are given by*

$$\hat{y} = proj_W(\vec{y}) = proj_{\vec{u}_1}(\vec{y}) + \cdots + proj_{\vec{u}_p}(\vec{y}) = \sum_{i=1}^{p} proj_{\vec{u}_i}(\vec{y})$$

Warning: this only works for an orthogonal basis.

Theorem 55 ***closest approximation theorem****: Basically, an orthogonal projection is the closest vector to the vector.*

Let W *be a vector subspace of* $\mathbb{R}^n$ *and let* $\vec{y} \in \mathbb{R}^n$ *then*

$$\forall \vec{v} \in W \qquad \|\vec{y} - proj_W(\vec{y})\| \leq \|\vec{y} - \vec{v}\|$$

Example 40 *Let* $B = \left\{ \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} \right\}$ *be an orthogonal basis of some subspace. Find the orthogonal decomposition of* $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Solution:

$$\hat{y} = \frac{6}{10} \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 9/5 \\ 2 \\ 3/5 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 9/5 \\ 2 \\ 3/5 \end{bmatrix} + \begin{bmatrix} -4/5 \\ 0 \\ 12/5 \end{bmatrix}$$

Theorem 56 ***Gram-Schmidt Procedure****: Given a vector subspace* W *of* V *and Given a basis* $B = \{\vec{x}_1, \ldots, \vec{x}_n\}$ *of* W*, then the set*

$$\begin{aligned} \vec{v}_1 &= \vec{x}_1 \\ \vec{v}_2 &= \vec{x}_2 - proj_{\{\vec{v}_1\}}(\vec{x}_2) \\ \vec{v}_3 &= \vec{x}_3 - proj_{\{\vec{v}_1, \vec{v}_2\}}(\vec{x}_3) \\ &\vdots \\ \vec{v}_p &= \vec{x}_p - proj_{\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{p-1}\}}(\vec{x}_p) \end{aligned}$$

The set $V = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ *spans* W *and is orthogonal.*

19 May 11

19.1 6.5 – Least Squares

Here is a problem: If the system

$$\mathbf{A}\vec{x} = \vec{b}$$

has no solution, does it at least have a close solution

$$\mathbf{A}\hat{x} = \vec{b}$$

By projection,

$$\hat{b} = proj_{Col(\mathbf{A})}(\vec{x})$$

Definition 46 *If* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *and* $\vec{b} \in \mathbb{R}^m$*, the* ***least squares solution*** *of the linear system*

$$\mathbf{A}\vec{x} = \vec{b}$$

is the vector $\hat{b} \in \mathbb{R}^n$ *such that*

$$\forall \vec{y} \in \mathbb{R}^n \qquad \|\vec{b} - \mathbf{A}\hat{x}\| \leq \|\vec{b} - \mathbf{A}\vec{x}$$

Basically, even if the linear system is inconsitent, $\hat{x}$ *is the closest possible solution.*

Theorem 57 *By the closest approximation theorem from yesterday, the least squares solutions are all the solutions to*

$$\mathbf{A}\hat{x} = proj_{Col(\mathbf{A})}(\vec{b})$$

This method of calculating is very bad because you need an orthogonal basis.

Definition 47 *For* $\mathbf{A} \in \mathbb{R}^{m\times n}$*,*

$$\mathbf{A}^T\mathbf{A}$$

is called the ***normal matrix*** *of* $\mathbf{A}$*.*

If $\mathbf{A}\vec{x} = \vec{b}$ *is a linear system, then*

$$\mathbf{A}^T\mathbf{A}\vec{x} = \mathbf{A}^T\vec{x}$$

is called the ***normal equation****.*

Theorem 58 ***calculating least-squares****: The least squares solutions of*

$$\mathbf{A}\vec{x} = \vec{b}$$

are the same as the set of solutions to

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

Example 41 *Find the least squares solution to the linear system*

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

The normal equation is

$$\begin{pmatrix} 8 & 4 \\ 4 & 5 \end{pmatrix} \hat{x} = \begin{bmatrix} 8 \\ 8 \end{bmatrix}$$

Example 42 *Given the points*

$$(1,1), (2,1), (3,2)$$

find the least-squares regression line.

We want the least-squares solution to

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

So

$$y = \frac{1}{2}x + \frac{1}{3}$$