MATH 4651

Contents

Chapter 1

Course Introduction

Professor Information: Valia Gazaki. Office Kerchof 221. The textbook is needed for HW assignments. Most of the weeks the HW comes from the book.

Homework will be kind of long, if you do the homework well then the exams will be similar and simpler. There are 2 Midterms. Each is 25%. The Final Exam is the remaining 35%. The first Midterm will be in class the second will be take-home. The Final Exam will be TBD. The timing will be a 3-4 hour commitment spread out over 24 Hours. The HW will be due Thursdays at 11:00 PM. On Gradescope. Possible dates for midterms. Created a Piazza in the past. Students used the piazza in semesters past.

There will be one quiz at some point before the midterm before the midterm that should be used for your own benefit.

Proof based substantial course. You need to know what you are doing.

Very natural transition from Survey of Algebra.

Office Hours. Tuesdays from 11:00-12:00 AM additionally We. from 3:30 - 5:00.

Follow the book mostly.

Chapter 2

Course Material

2.1 Introduction to Fields

The main example of a field that we all know and love is the real numbers.

Definition 2.1.1: Definition of a Field

A field is a set equipped with two operations $+, \cdot$. A field has the following 10 properties:

1. FA1) Closure under addition: For all $x, y \in F$, the sum of $x + y \in F$.
2. FA2) Closure under multiplication: For all $x, y \in F$, the product $x \cdot y \in F$.
3. FA3) Commutativity of addition: For all $x, y \in F$ $x + y = y + x$.
4. FA4) Associativity of addition: For all $x, y, z \in F x + (y + z) = (x + y) + z$.
5. FA5) Additive identity: There exists $0 \in F$ such that for all $x \in F$, we have $0 + x = x$.
6. FA6) Additive inverses: For all $x \in F$ there exists $(-x) \in F$ such that $x + (-x) = 0$.
7. FA7) Commutativity of multiplication: For all $x, y \in F$, $x \cdot y = y \cdot x$.
8. FA8) Associativity of multiplication: For all $x, y, z \in F$, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.
9. FA9) Multiplicative identity: there exists $1 \in F$ such that for all $x \in F$, $1 \cdot x = x$.
10. FA10) Multiplicative inverses: for all $x \in F$ such that $x \neq 0$, there exists $x^{-1} \in F$ such that $x \cdot x^{-1} = 1$.
11. FA11) Distributivity of multiplication over addition: For all $x, y, z \in F$, we have $x \cdot (y + z) = x \cdot y + x \cdot z$.
12. FA12) Distinct Additive and Multiplicative identities: $1 \neq 0$.

2.2 Common Examples of Fields

Example 2.2.1 (Examples of Fields)

1. The rational numbers $\mathbb{Q}$
2. The real numbers $\mathbb{R}$
3. The complex numbers $\mathbb{C}$
4. 3×3 Matrices $M_3(\mathbb{R})$ with real entries in $\mathbb{R}$.

Definition 2.2.1: Definition of a Vector Space

Fix F to be a field with the properties as defined above. A vector space V over a field F is defined as a non-empty set V with a binary operation $(+)$ and a binary function $(\cdot)$ that satisfy the axioms below. In this context, the elements in V are called vectors and the elements in F are called scalars.

1. Associativity of Vector addition: for all $x, y, z \in V$, we have $x + (y + z) = (x + y) + z$.
2. Commutativity of Vector addition: for all $x, y \in V$, $x + y = y + x$.
3. Existence of Additive element in V: There exists $0_v \in V$ such that for all $x \in V$: $x + 0_v = x$.
4. Existence of Additive Inverses in V: for all $x \in V$, there is a corresponding (and unique) $y \in V$ such that $x + y = 0_v$. We represent y as $-x$.
5. Compatibility of Scalar Multiplication with Field Multiplication: for $a, b \in F$ and $v \in V$ we have $a(bv) = (ab)v$.
6. Identity element of scalar multiplication: $1v = v$, where $1v$ represents the multiplicative identity in F.
7. Distributivity of Scalar multiplication with respect to vector addition, for $a \in F$ and $x, y \in V$: $a(x + y) = ax + ay$
8. Distributivity of Scalar multiplication with respect to field addition, let $a, b \in F$ and $v \in V$: $(a+b)v = av + bv$

2.3 Common Examples of Vector Spaces

Example 2.3.1 (F^n)

Define a vector space V to be F^n. $F^n = \{(a_1, a_2, a_3, \dots, a_n) : a_i \in F\}$. We use coordinate wise operations. When adding, we add a_1 to a_1 in both vectors. Scalar multiplication is defined intuitively with: $c(a_1, a_2, \dots, a_n) = (ca_1, ca_2, ca_3, \dots, ca_n)$.

Example 2.3.2 (Matrices)

$$V = M_{n\times m}(F) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}$$

Example 2.3.3 (Polynomials)

The set of all polynomials with real coefficients is a real vector space, with the usual operations of addition of polynomials and multiplication of polynomials by scalars (in which all of the coefficients of polynomial is multiplied by the same real number).
$V = F[x]$ the book uses $P(F)$ $V = \{p(x) = a_n x^n + \dots + a_1 x_1 + a_0\}$, where $n \geqslant 0, a_i \in F$ We add functions and multiply operations.

Definition 2.3.1: Definition of degree

let $f(x) \in F[x], f(x) = a_m x^n + ... + a_1 x + a_0$. Then the degree of $f(x)$, $deg(f) = max\{n \geqslant 0 s.t. a_n \neq 0\}$.

Example 2.3.4 (Example)

$p(x) = 3x^3 - 2x + 1$, the corresponding degree is 3. If $f(x) = 0$, then by convention $f(x)$ has no degree.

There are some properties of degree.

① A zero degree polynomial is not defined.

② The degree of a polynomial is one if it is a constant.

③ The degree of a polynomial times a scalar is the degree of the polynomial unless the scalar is zero.

④ When adding two polynomials the resulting degree of the polynomial is the max of the degree of the two polynomials, except if both polynomials are equal to zero.

⑤ When multiplying two polynomials the resulting degree of the polynomial is the max of the largest degree in each polynomial or the sum of the two largest polynomials.

Example 2.3.5 (Continuous Real Functions)

The set $C(D, \mathbb{R})$ of all continuous real-valued functions defined over a given subset of D of the real numbers is a real vector space: if $x \longrightarrow f(x)$ and $x \longrightarrow g(x)$ are continuous functions on D then so are $x \longrightarrow f(x)+g(x)$ and $x \longrightarrow cf(x)$ for all real numbers c. Moreover, these operations of addition and scalar multiplication follow the axioms of vector spaces.

2.4 Non-Examples:

Example 2.4.1 (Addition Changed)

Let $V = \{(a, b) \in \mathbb{C}^2\}$ but with vector addition defined as $(a_1 + b_1, a_2 - b_2)$. V is not a vector space because addition is not commutative. This is because $(5, 3) + (2, 5) = (7, -2) \neq (7, 2) = (2, 5) + (5, 3)$

Example 2.4.2 (Multiplication Changed)

$T = \{(a, b) \in \mathbb{Q}^2\}$ with normal addition defined as $(a, b) + (x, y) = (a + x, b + y)$. But with multiplication defined as follows: $c(a_1, a_2) = (ca_1, 0)$. This breaks existence of a multiplicative identity. This is for the example of $c = 1, (4, 4) = (4, 0) \neq (4, 4)$.

2.5 Proof that additive inverses are unique

Theorem 2.5.1 Additive inverses are unique. Multiplicative inverses are unique.

Proof: Prove that additive identities are unique:

For some $v \in V$, assume that we have two unique additive identities for v. That is $a + v = 0 = b + v$.

$$a = 0 + a$$
$$a = (b + v) + a$$

$$a = b + (v + a)$$
$$a = b + 0$$
$$a = b$$

😄

Theorem 2.5.2 Cancellation Law for Vector Space

For $a, b, c, d \in V$: $a + v = b + v$
$a = b$

Proof: For $x, y, z \in V$ we are given $x + z = y + z$, and we want to show that $x = y$. Our proof proceeeds directly. Additionally, let $z + w = 0$. x

$$x + 0_v$$
$$x + (z + w)$$
$$(x + z) + w$$
$$(x + z) + w = (y + z) + w$$
$$x + (z + w) = y + (z + w)$$
$$x + 0 = y + 0$$
$$x = y$$

😄

Theorem 2.5.3 Proving a Vector Space

Let V be a vector space. Then the following are true:

1. $0 \cdot v = 0_v$ for all v
2. $-a(v) = -(av)$ for all v
3. $a \cdot 0_v = 0_v$ for all $a \in F$

Proof: Proving One Above

$$0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v$$

Apply the cancellation law. $0 \cdot v =$ This means that $0 \cdot v$ is the zero vector.

😄

Proof: Proving Two Above

The left hand side is the additive inverse of a scalar $a \in F$ times a vector $v \in V$, the right hand side is the additive inverse of a scalar times a vector. If we can show that the two are additive inverses of the same thing, then the uniqueness of additive inverses will guarantee $-a(v) = -(av)$ for all $v \in V$ and $a \in F$.

The following is true: $av + (-a)v = 0_v$

$$(a + -a)v = 0_v$$
$$(0)v = 0_v$$

Therefore, $av + (-a)v = 0_v$. Then $(-a)v = -(av)$. By subtraction.

😄

Definition 2.5.1: Definition of a subspace

Let V be a vector space over F. Let $W \subseteq V$ and write $W \leqslant V$ if W is itself a vector space over F with addition and scalar multiplication inherited by the ones of V. This means that for $v, w \in W, w + v \in V$.

2.6 Verify something is a subspace

There are two ways to verify something is a subspace. The first option is to verify all 8 axioms of a vector space, and then find a set with the same operations that is larger, (a corresponding V). There is however a more optimal way to proving subspaces.

2.7 How to show $W \leqslant V$?

How to verify if something is a subspace. We need to show that first the subspace is a vector space. Verify that $(v1) - (v8)$ are true The optimal way to do this is :

Theorem 2.7.1 Optimal Way to check subspace

Let V be a vector space over F and $W \subseteq V$, Then: $W \leqslant V$ if and only if the following conditions are met:

1. The zero vector of V is an element of W. $0_v \in W$.
2. $\forall x, y \in W \to x + y \in W$. Simply, W is closed under addition.
3. $\forall x \in W$ and $c \in F$, it follows that $c \cdot x \in W$. This means that W is closed under scalar multiplication.

Proof: Backward direction. If something satisfies these three then it is a vector space

Suppose that 1,2,3 hold. Verify $W \times W \to W$, $(x, y) \to x + y$. So satisfies $V1$.
$x + y$ is addition in V, because we know that V is a vector space which means that it satisfies $V1$. $V2, V3$ are similar. We will verify $V4$.
We want every element of W to have an additive inverse. Let $w \in W$, by property 3 $(-1) \cdot w \in W$. From Theorem 2 this is equal to $-W$. This gives us $-w \in W$. Which means that the additive inverse of w lies in W.

Prove the other direction. Note that some axioms are left as practice.

Suppose $W \leqslant V$ Axioms two and three are automatic by the definition of being a subspace. The only thing that remains to show is that $0_v \in W$. We use a similar argument. We know that W is a vector space, so it must have a 0_v itself. We will show that it must be the same one as the one in V.
Since W is a vector space, it satisfies $(V4) : \exists 0_v \in W$ such that $0_w + x = x$ for all $x \in W$. We will show that $0_w = 0_v$.
$0_w = 0_w + 0_w$. This is because $0_w \in W$ and 0_w is the zero vector of W.
Similarly, $0w \in V$ gives $0_w = 0_w + 0_v$. This uses V4 for v. Now this gives us the cancellation law. It gives us $0_w + 0_v = 0_w + 0_v$.

😃

Question 1: Do we need the first property of the theorem to show that one thing is a subspace

Why do we need to show that the two subspaces have the same identity vector. We can get the first property from the second property unless the set is empty. The empty set otherwise would be a vector space. But the vector space must have at least one element in it. We do not allow $W =$ but this does not satisfy $V4$.

Question 2: Practice Question

Show that these three conditions in the theorem can be replaced by $W \leqslant V \leftrightarrow$:

1. $W \neq \emptyset$
2. For every $x, y \in W$ and $c, d \in F \to c \cdot x + d \cdot y \in W$

Example 2.7.1 ($V = f(\mathbb{R}, \mathbb{R})$)

$W = \{f \in V$ such that $f(1) = 0\}$. The zero function $f : \mathbb{R}, f(x) = 0 \forall x \in W$.
Let $f_1, g_1 \in W \to f(1) = g(1) = 0$. Then, $(f + g)(1) = f(1) + g(1) = 0 \to f + g \in W$ and W is closed under $+$. Let $c \in R$ be a scalar and $f \in W$, Then $(c \cdot f)(1) = c \cdot (f(1)) = c \cdot 0 = 0 \to c \cdot f \in W$ and W is closed under scalar multiplication.

Example 2.7.2 (Example one Prime)

$U = \{f \in F(\mathbb{R}, \mathbb{R}) s.t. f(1) = 1\}$ Not even the zero function is in this subspace.

Example 2.7.3 (square Matricies)

$V = M_{n \times m}(F)$
$W =$ Upper triangular Matricies. 0 Below the diagonal. $W \leqslant M_{m \times n}(F)$

Example 2.7.4 (Polynomials)

$V = f[x]$
$W = \{p(x) \in V s.t. \deg p = 3\}$. The zero is not even in here. Not closed under addition. Generally this is not closed under addition. $x^3 + (-x^3) \notin W$.

We can fix this one with $P_3(F) = \{p(x) \in F[x]$ such that $\deg p \leqslant 3\} \bigcup \{0\}$.

Theorem 2.7.2 Avoid doing work

Let V be a $F - v.s.p$ Let $\{W_i\}_{i \in I}$ be a collection of possibly infinite of subspaces then the intersection of w_i is again a subspace. $\bigcap_{i \in I} W_i$.

Proof: Verfiy 0 is in the intersection. Each w_i is a subspace so it must be the case that 0 is in all of their intersections. We know $0 \in W_i$ for all $i \in I$, because W_i is a subspace of V This implies that it is in the intersection becuase it lies in all of them.

For addition we want to show that we still have closure under addition in the intersection.
Let $x, y \in \bigcap_{i \in I} W_i$ we want to show that $x + y \in \bigcap_{i \in I} W_i$. $x, y \in \bigcap_{i \in I} W_i \to x, y \in W_i$ for all i Since W_i is a subspace $\to x + y \in W_i$ for all $i \to x + y \in \bigcap_{i \in I} W_i$.

The multiplication is similar

😊

Example 2.7.5 (Solution to sets to linear system of equations)

$V = F^2 = \{(a, b) : a, b \in F\}$
. Define $W_{a,b} = \{(x, y) \in F^2 s.t. ax + by = 0\}$ We want to show that $W_{a,b} \leqslant F^2$

Proof: $(0, 0)$ is in $W_{a,b}$ since $a \cdot 0 + b \cdot 0 = 0$. For two, let $(x_1, y_1), (x_2, y_2) \in W_{a,b}$. We want to show that $(x_1, y_1) + (x_2, y_2) \in W_{a,b}$. We know that $(x_1, y_1), (x_2, y_2) \in W_{a,b} \to ax_1 + by_1 = 0, ax_2 + by_2 = 0$. We want to show that $(x_1 + x_2, y_1 + y_2) \in W_{a,b}$.

$$a(x_1 + x_2) + b(y_1 + y_2) = ax_1 + ax_2 + by_1 + by_2 = (ax_1 + by_1) + (ax_2 + by_2) = 0 + 0$$

😊

Example 2.7.6 (Example 8)

Let $U = \{(x, y) \in F^2 s.t. 3x + 2y = 0$ and $5x - 11y = 0\}$ $U = W_{3,2} \bigcap W_{5,11}$ we know these are both subspaces so that it a subspace.

2.8 Linear Combinations and Span

Definition 2.8.1: Linear Combination

Let V be a $F - VSP$, let $S \subseteq V$, $S \neq \emptyset$ We say that a vector $v \in V$ is a linear combination of vectors in S if there are finitely many vectors $v_1, ... v_n \in S$ and scalars $a_1, ..., a_n \in F$ such that $v = a_1 v_1 + ... v_n$

Chapter 3

Tuesday September 3rd

Definition 3.0.1: Linear Combination

Let V be a vector space over a field F. Let $S \subseteq V$ with $S \neq \emptyset$. We say that a vector $v \in V$ is a finite linear combination of vectors in S, if $\exists n \in \mathbb{N}$ and $v_1, \ldots, v_n \in S$ and scalars $a_1, \ldots, a_n \in F$ such that $v = a_1 v_1 + \ldots + a_n v_n$.

Note:-

Take the field that makes sense otherwise the problem will clearly state that this is not the case.

Example 3.0.1 (Example 1)

Take V to be the space of all polynomials $V = \mathbb{R}[x]$. $S = \{h_1(x) = x^3 - 2x^2 - 5x - 3, h_2(x) = 3x^3 - 5x^2 - 4x - 9\}$. Check if the polynomials $f(x) = 2x^3 - 2x^2 + 12x - 6$ and $g(x) = 3x^3 - 2x^2 + 7x + 8$ are linear combinations from S.
We need to find scalars that make this work, the vectors are given to us. This should lead us to a linear system of equations. The question becomes, does their exist $a, b \in \mathbb{R}$ such that $f(x) = a \cdot h_1(x) + b \cdot h_2(x)$.
This is equivalent to asking $2x^3 - 2x^2 + 12x - 6 = a(x^3 - 2x^2 - 5x - 3) + b(3x^3 - 5x^2 - 4x - 9)$
The LHS $= x^3(a + 3b) + x^2(-2a - 5b) + x(-5a - 4b) + (-3a - 9b)$. Which is equivalent to:
$2 = 9 + 3b$
$-2 = -2a - 5b$
$12 = -5a - 4b$
$-6 = -3a - 9b$
Solve for a, b. We get $a = -4$ and $b = 2$. In this case $f(x)$ is a linear combination from elements of S. In this case we get something that doesn't have a solution. This means that $g(x)$ is not a linear combination of elements from S.

Example 3.0.2 (Example 2)

Lets say that $V = \mathbb{F}(\mathbb{R}, \mathbb{R})$. Lets say that $S = \{1, x, x^2, \ldots, x^n\} = \{x^n : n \geqslant 0\}$. Let's take $f(x)$ to be the absolute value function: $f(x) = |x|$.
Question is $f(a)$ a linear combination from elements in S?
Note that $g(x) = e^x$ is not a linear combination because it is not a finite linear combination from S, although we know that the taylor expansion for $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$

Proof: Suppose for contradiction that e^x is a linear combination from S. This implies that e^x is a polynomial of deg $n \geqslant 0$. Then if we take the $n + 1$ derivative of e^x should be zero. However we know that the derivative of e^x is itself. This is our contradiction.

☺

Proof: Similarly we can apply the same logic to the absolute value function. We know that the absolute value function is a linear combination of all of the powers of x if it is differentiable everywhere. However, we know that the absolute value funciton is not differentiable everywhere, therefore, it must be the case that it is not a linear combination of elements from S.

😊

Note:-

In order to say an element is a linear combination we should be able to say that there are a finite number of components in the linear combination.

Definition 3.0.2: Span

Let V be a vector space over F, $S \subseteq V, S \neq \emptyset$. The span of S denoted $\text{span}\,(S)$ = the set of all finite linear combinations of vectors in S.
$\text{span}\,(S) = \{v \in V \text{ such that } \exists n \geqslant 1, v_1, \ldots, v_n \in S, a_1, \ldots, a_n \in F \text{ such that } v = a_1v_1 + \ldots + a_nv_n\}$
The set of all vectors that are linear combinations of S.

Note:-

Convention

By convention, we allow the span of the empty set to be the zero vector.
$\text{span}(0) = \{0\}$.

Example 3.0.3 (Example 3)

Let $V = \mathbb{C}^3$, $S = \{(i,0,0),(2i,1,0)\}$. Our claim is that $\text{span}(S) = \{(x,y,0) : x,y \in \mathbb{C}\}$.

Proof: We begin by showing that the sets are sets of one another. We have one set equal to another set let $U = \{(x,y,0) : x,y \in \mathbb{C}\}$. We want to show that $\text{span}\,(S) \subseteq U$ and that $U \subseteq \text{span}(S)$.
$\text{span}(S) \subseteq U$: we start with an arbitrary element in the span of S. Let V be an element in $\text{span}(S)$.
This means that $v = (x,y,z) \in \mathbb{C}^3$, such that there exists scalars $a, b \in \mathbb{C}$ with
$(x,y,z) = a(i,0,0) + b(2i,1,0)$. This iplies
$(x,y,z) = (ai + 2bi, b, 0)$
This means that $x = ai + 2bi$
$y = b$
$z = 0$
This implies that $v \in U$.
The other direction: $U \subseteq \text{span}(S)$, we will start for an element and look for scalars. In the first example we don't have a specific vector.
Let $v \in U$, by definition $v = (x,y,z)$ for some $x, y \in \mathbb{C}$.
We look for $a, b \in \mathbb{C}$ such that $(x,y,0) = a(i,0,0) + b(2i,1,0)$. The only difference now is that we are looking for these a, b values. It will be exactly as we had before:
$x = ai + 2bi$
$y = b$
Note that we need to solve this for a, b in terms of x, y.
The second relation gives us that $b = y$, which we can plug into the first equation. We can solve this for a. We get that a is equal to $\frac{x-2yi}{i}$. Since we have a solution, we know that $v \in \text{span}(S)$.

😊

Theorem 3.0.1 Theorem 1 From class today

Let V be a vector space, $S \subseteq V$. Then the following are true:

1. $\text{span}(S)$ is a subspace of V.

② $\operatorname{span}(S)$ is the smallest subspace of V that contains S. This means that if $W' \leqslant V$ such that $S \subseteq W'$, then $\operatorname{span}(S) \subseteq W'$.

Proof: Case 1) If $S = \emptyset$, then $\operatorname{span}(S) = \{0\}$. And this is a subspace by definition. Case 2): Suppose that $S \neq \emptyset$, we need to show that $0_v \in \operatorname{span}(S)$: Since $S \neq \emptyset \longrightarrow \exists v_1 \in S$. Then it follows that $0 \cdot v_1 \in \operatorname{span}(S)$. We showed last class that this is equal to the 0_v.
To show that $\operatorname{span}(S)$ is closed under addition:

Note:-

We want to show that if v_1, w are in the span, then we want to show that $v + w \in \operatorname{span}(S)$. $v \in \operatorname{span}(S)$ means that we can write it in the form $a_1 v_1 + \ldots + a_n v_n$ for some $n \in \mathbb{N}$, some $v_1, \ldots, v_n \in S$, and some $a_1, \ldots, a_n \in F$.
$w \in \operatorname{span}(S) \longrightarrow \exists m \in \mathbb{N}, v_1', \ldots, v_m' \in S$, and $b_1, \ldots, b_n \in F$ such that $w = b_1 v_1' + \ldots + b_m v_m'$.
By enlarging the sets of vectors by adding some coefficients $= 0$ we may assume that $\exists n \in \mathbb{N}, v_1, \ldots, v_n \in S$, and $a_1, \ldots, a_n \in F$, such that $v = a_1 v_1 + \ldots + a_n v_n$ and $w = b_1 v_1 + \ldots + b_n v_n$.
If there is any vector in one of them that doesn't appear in the other one we can put a zero to make them of the same degree. We continue with our proof of closure under addition. Let $v, w \in \operatorname{span}(S)$, we want to show that $v + w \in \operatorname{span}(S)$. We may assume (by possibly adding 0 coefficients), that $\exists n \geqslant 1, v_1, \ldots, v_n \in S, a_1, \ldots, a_n, b_1, \ldots, b_n \in F$ such that $v = a_1 v_1 + \ldots + a_n v_n$ and $w = b_1 v_1 + \ldots + b_n v_n$. Such that when adding we get $v + w = (a_1 + b_1) v_1 + \ldots + (a_n + b_n) v_n$. Where the elements within the parentheses are elements are elements within F call it $c_i \in F$. Then we have $v + w = c_1 v_1 + \ldots + c_n v_n \in \operatorname{span}(S)$.
Show that $\operatorname{span}(S)$ is closed under scalar multiplication. This concludes our proof of part 1.

😄

Proof of the second Case:

Proof: Let $W' \leqslant V$ such that $S \subseteq W'$, we want to show that the $\operatorname{span}(S) \leqslant W'$. This one is simpler. Let V be an element of the span(S) then by definition this means $\exists n \geqslant 1$, $v_1, \ldots, v_n \in S, a_1, \ldots, a_n \in F$ such that $v = a_1 v_1 + \ldots + a_n v_n$.
$S \subseteq W'$ and $v_1, \ldots, v_n \in S$ implies the fact that:
$v_1, \ldots, v_n \in W'$. Since W' is a subspace of V, W' is closed under scalar multiplication $\longrightarrow a_1 v_1, \ldots, a_n v_n \in W'$ and $a_1 v_1 + \ldots + a_n v_n \in W' \longrightarrow \operatorname{span}(S) \subseteq W'$.

😄

Note:-

Span of S

The span of S isn't contained in S because S isn't a subspace. If S is a subspace then the span of S is equal to S by minimality.

Note:-

HW 2 Note

$\operatorname{span}(S) = \bigcap W$, $W \leqslant V$, $S \subseteq W$.

Example 3.0.4 (For Practice)
Lets Take $P_2(\mathbb{R}) = \{p(x) \in \mathbb{R}[x],\ \text{degree } p(x) \leqslant 2\} \bigcup \{0\}$. Show if $S = \{1 + x + x^2, 3 - x, 5 + x\}$, then the $\operatorname{span}(S) = p_2(\mathbb{R})$
Do this problem.

3.1 Summary

Starting with $S \subseteq V$ we created $W = \operatorname{span}(S) \leqslant V$. We will say that the set S generates W or spans the subspace W. We can ask the following questions now:

1. If $W \leqslant V$, can find $S \subseteq W$ that generates it?
   Yes take $S = W$.

2. Can we find a minimal generating set for W?
   (S would be minimal if the following is true)

   (a) $\operatorname{span}(S) = W$

   (b) If $S' \subseteq S$ and $\operatorname{span}(S) = W$, then $S' = S$.

   Equivalently S is minimal if $\operatorname{span}(S) = W$, and if $T \subset S$, then T does not span W

3. Given a generating set $S \subseteq W$ such that $\operatorname{span}(S) = W$, can we extract a mimimal one from it?

4. If S and S' are 2 minimal generating sets, for W, do they have the same size? Do they have the same number of elements?

We will provide a positive answer to all of these questions. This is what the next couple of classes will be on.

Definition 3.1.1: Linear Dependence / Independence

Let V be a vector space over F, and $S \subseteq V$.

1. S is called linearly dependent if $\exists$ distinct vectors $v_1, \ldots, v_n \in V$ and scalars $a_1, \ldots, a_n \in S$ not all 0, such that $a_1 v_1 + \ldots + a_n v_n = 0_v$.
2. S is linearly independent if S is not linearly dependent. This means given any vectors $v_1, \ldots, v_n \in S$, the only solution to the equation $a_1 \cdot v_1 + \ldots + a_n \cdot v_n = 0_v$ with $a_i \in F$, is $a_1 = a_2 = \ldots = a_n = 0$. The trivial solution.

Chapter 4

Thursday September 5th

Professor Gazaki will be leaving next week Wednesday.

4.1 Review from last time

Definition 4.1.1: Span

V is a vector space over a field F and $S \subseteq V$, $S \neq \emptyset$. $\text{span}(S) = \{v \in V : \exists n \geqslant 1, v_1, \ldots, v_n \in s \text{ and } a_1, \ldots, a_n \in F \text{ such that } v = a_1v_1 + \ldots + a_nv_n\}$

Definition 4.1.2: Linearly dependent

We say that S is linearly dependent if $\exists v_1, \ldots, v_n \in S$, $a_1, \ldots, a_n \in F$, not all 0, such that $a_1v_1 + \ldots + a_nv_n = 0_v$. Similarly we say that S is linearly indepenedent if whenever we have $a_1v_1 + \ldots + a_nv_n = 0_v$ with $a_1 = \ldots = a_n = 0$.

Example 4.1.1 (Example 1)

Let $S = \{0_v\}$. We claim then S is linearly dependent. Any scalar in the field times the 0_v is 0.

Proof: $1 \cdot 0_v = 0_v$ and $1 \neq 0$.

☺

Note:-

General Rule

More generally if $S \subseteq V$ with $0_v \in S \longrightarrow S$ is linearly dependent.

Proof by contradiction that a subset with a non zero vector is linearly independent.

Example 4.1.2 (Example 3)

$V = M_{2\times2}(\mathbb{R})$ and S the following subset Set of three matrices. We need to explore whether S is linearly independent. Suppose $a, b, c \in \mathbb{R}$ such that a scalar times all three of the matrices is equal to zero. We can then create a system of equations. If all three of the valeus are zero then our set is linearly independent.

Example 4.1.3 (Proposition 1)

Suppose that $S = \{v_1, \ldots, v_n\} \subseteq V$. Then S is linearly dependent $\leftrightarrow \exists v_i \in S$ that can be written as a linear combination of $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$.

Proof: Forward direction. Suppose that S is linearly dependent. This means by definition that

$\exists a_1, \ldots, a_n \in F$ not all equal to zero such that $a_1 \cdot v_1 + \ldots + a_n v_n = 0_v$.

At least one of these $a_i \neq 0$. This will be the one that we use. Without loss of generality assume that $a_1 \neq 0$. Now we can solve for v_1:

$$a_1 v_1 = -a_2 v_2 - \ldots - a_n v_n$$
$$v_1 = -\frac{a_2}{a_1} v_2 - \ldots - \frac{a_n}{a_1} v_n$$

Then $v_1 \in \text{ span}\{v_1, \ldots, v_n\}$. 😊

In the other direction we assume that one of the v_i can be written as a combination of the others.

Proof: Assume $\exists i \in \{1, \ldots, n\}$ and $a_1, \ldots, a_n$ such that $v_i = a_1 v_1 + \ldots + a_{i-1} v_{i-1} + a_{i+1} v_{i+1} + a_n v_n$. $a_1 v_1 + \ldots + a_{i-1} v_{i-1} + (-1) v_i + a_{i+1} v_{i+1} + \ldots + a_n v_n = 0$. This is a non trivial solution to our first equation because the scalar $a_i = -1 \neq 0$. 😊

Theorem 4.1.1

Let V be a vector space over a field F, and lets say that $S \subseteq T \subseteq V$. Then:

1. If T is independent then S is linearly independent.
2. If S is dependent then T is dependent.

Example 4.1.4 (Proposition 2)

Let $S = \{v_1, \ldots, v_n\} \subseteq V$ be linearly independent. Let $v \in V \setminus S$. Then $S \bigcup \{v\}$ is linearly dependent $\leftrightarrow v$ is in the span S.

Proof: The converse follows from proposition 1. We don't need to prove it.

Assume that $S \bigcup \{v\}$ is linearly independent. By definition this means that we can find $\exists a_1, \ldots, a_n, a_{n+1} \in F$ not all 0 such that $a_1 v_1 + \ldots + a_n v_n + a_{n+1} v = 0$. It is enough to show that $a_{n+1} \neq 0$ becuase we can solve for v. Suppose for contradiction that $a_{n+1} = 0$. Then $a_1 v_1 + \ldots + a_n v_n = 0_v$ with not all $a_1, \ldots, a_n$ equal to zero. But this contradicts S as linearly independent.

😊

Example 4.1.5 (Proposition 3)

If $S \subseteq T \subseteq V$, then span$(S)$ $\leqslant$span(T)

Example 4.1.6 (Proposition 4)

Let $S \subseteq V$ and v $\in$span(S), then span$(S \bigcup \{v\})$ is the span(S).

Example 4.1.7 (Proposition 5)

Let $S, T \subseteq V$: $S = \{v_1, \ldots, v_n\}, T = \{u_1, \ldots, u_n\}$, if for each $i \in \{1, \ldots, n\}$. $v_i \in$ span(T) and for each $j \in \{1, \ldots, m\}$, $u_j \in$ span$(S) \longrightarrow$ span$(S) =$ span(T)

Example 4.1.8 (6)

$S = \{v_1, v_2\} \subseteq V$, then S is linearly dependent $\leftrightarrow$ one of the vectors is amultiple of the other

Example 4.1.9 ((7) Important example)

$V = \mathbb{C}^2$, $S = \{(2,i),(2+2i,i-1)\}$. We can do two different things. Consider V as an $\mathbb{R}$ vector space. This means scalars are real numbers. We will show that S is $\mathbb{R}$ linearly independent.

Proof: Let $a,b \in \mathbb{R}$ such that $a(2,i)+b(2+2i,i-1)=(0,0)$

$$\text{This is equivalent to } (2a+2b+2bi, ai-b) = (0,0)$$
$$\text{This is equivalent to } 2a+2b+2bi = 0+0i$$
$$ai-b = 0+0i$$
$$\text{Because } a,b \text{ are real numbers we have } 2a+2b=0$$
$$2b=0$$
$$a=0$$
$$b=0$$
$$\text{In particular } a=b=0 \text{ is the only solution.}$$

This implies that S is $\mathbb{R}$ linearly independent. ☺

Part (B): Consider V as a $\mathbb{C}$ vector space now. Suppose that $a,b \in \mathbb{C}$ such that $a(2,i)+b(2+2i,i-1) = (0,0) \leftrightarrow$
$2a+2b+2bi=0$
$ai-b+bi=0$
this has non trivial solutions. Take $b=1$, $a=1+i$. This means that S is $\mathbb{C}$ linearly dependent.

Definition 4.1.3

A subset $S \subseteq V$ which is a vector space over F is called a basis for V if:

1. The $\text{span}(S)=V$
2. S is linearly independent

Note:-

Every finite dimensional vector space has a basis.

Example 4.1.10 (Basis)

Recall that a basis is a minimal generating set. If $W = \text{span}(S)$ then S is a minimal generating set for W if for every $S' \subsetneq S$ $\text{span}(S') \neq W$.

Theorem 4.1.2

Suppose that $W \leqslant V$ and that $W = \text{span}(S)$ for some $S \subseteq V$, then S is minimal if and only if S is linearly independent.

Proof: Forward direction Suppose that S is minimal but for contradiction suppose S is linearly dependent.

By one of the first things we proved today. We get that there exists some vector in S that can be written as a linear combination of other vectors $v_1,\dots,v_n \in S$. $\exists v \in S$.
We can extract v from our set and it will still span the entire thing. This means we can remove it. We know we can remove it because v is a linear combination of other vectors in our set. Let $S_0 = S \setminus \{v\}$, then $\text{span}(S_0) = \text{span}(S)$. But this contradicts the mminimality of S.

Suppose that S is independent. Assume towards contradiction that S is not minimal.
This means we there exists some $S_0 \subsetneq S$ such that $\text{span}(S_0)=W$. Since S_0 is strictly contained in S, this means there exists at least one $v \in S \setminus S_0$ since $S \subseteq W$,$v \in W \to v \in \text{span}(S_0)$. This means that v is a linear

combination from s_0 and this contradicts linear independence. Go in the other direction. Suppose now that S is independent 😄

Note:-

Conclusion

If $\mathbb{B}$ is a basis for V then B is a minimal generating set for V.

Example 4.1.11 (Standard Bases)

The standard bases:

1. $V = F^n$, $e_1, e_2, \dots : (1, 0, \dots, 0), (0, 1, 0, \dots, 0)$ etc.
2. Take Polynomials up to degree n. Powers of $X : 1, x, x^2, \dots, x^n$
3. Matrices: $V = M_{n\times m}(F)$. Standard Bases. 1 in one entry and 0 everywhere else.

Example 4.1.12 (Basis Example 2)

Takes an example of three matrices, show that B is a basis for the subspace W of all the upper triangular matrices. Three matrices given with entries in the upper corners. Some matrices are missing values in the upper diagonal.
Show linearly independent and spans everything.

Chapter 5

Tuesday September 10th

5.1 Announcements

Thursday will be taught virtually. Live Lecture. Use Piazza. Open Piazza Notifications.

Note:-

Span($S_1 \bigcap S_2$) Relates to the intersection of either span

5.2 Review

Definition 5.2.1: Basis

V be a vector space over a field F. A subset $S \subseteq V$ is a basis for V if:

1. $\text{span}(S) = V$
2. S is linearly independent

5.3 Goals To Show Today

1. If Someone gives us a finite subset of V,such that the span of S is the entire V, then S contains a basis. Meaning we can extract the basis.
2. If we start with $S \subseteq V$, linearly independent then we can extend S into a basis for V.
3. Any two bases for V have the same number of elements

Theorem 5.3.1

Let V be a vector space over F. Let $u_1, \ldots, u_n \in V$ be distinct vectors. Then $B = \{u_1, \ldots, u_n\}$ is a basis for $V \leftrightarrow$ each $v \in V$ can be written uniquely in the form $v = a_1u_1 + \ldots + a_nu_n$ for scalars $a_1, \ldots, a_n \in F$.

Note:-

Similar to the last problem in the first hw

Proof: Statement A: $B = \{u_1, \ldots, u_n\}$ is a basis. This implies that the span$B = V$. Meaning that every $v \in V$ is a linear combination of $u_1, \ldots, u_n$.Additionally, B is linearly independent.
Statement B: Each $v \in V$ has an expression $V = a_1u_1 + \ldots + a_nu_n$ with the $a_i \in F$, and two this expression is unique for each $v \in V$.

Note:-

Each Statement contains two pieces of information

Forward Direction($\rightarrow$): If we know that $B$is a basis then each $v \in V$ has an expression as $v = a_i u_1 + \ldots + a_n u_n$, for some $a_1, \ldots, a_n \in F$.

We want to show uniqueness. We start by assuming we have two expressions and we show they are the same. Suppose that a vector $v \in V$ has 2 expressions:

$v = a_1 u_1 + \ldots + a_n u_n$ and $v = b_1 u_1 + \ldots + b_n u_n$ for $a_i, b_i \in F$. We want to show that $a_i = b_i$.

If we subtract equation one from equation 2. we get:

$(a_1 - b_1)u_1 + \ldots + (a_n - b_n)u_n = 0$. Because we know that $B = \{u_1, \ldots, u_n\}$ is a basis and linearly independent we know that $a_i - b_i = 0$.

Backward Direction($\leftarrow$): We assume Statement B and want to show statement A. For the same reason, we only need to show that uniqueness implies that B is linearly independent.

Note:-

It is clear that span$(B) = V$

Remains to show that B is linearly independent. We always start with a linear combination equal to 0. Let $a_1, \ldots, a_n \in F$ such that $a_1 u_1 + \ldots + a_n u_n = 0$. We want to show that $a_1 = a_2 = \ldots + a_n = 0$.

We have: $a_1 u_1 + \ldots + a_n u_n = 0$, but also $0u_1 + \ldots + 0u_n = 0$. We have two expressions, and by uniqueness of expression, we know that $a_i = 0$ for $i \in \mathbb{N}$.

😊

Theorem 5.3.2

Let V be a vector space over F, $S \subseteq V$ a finite subset of V such that span(S)= V, then S contains a basis.

Proof: Special Case: Suppose first that S doesn't contain any non zero vectors. This assumption implies that $S = \emptyset$ or $S = \{0_v\}$. By assumption span$(S) = V$ implieas that $V = \{0\}$. Then $B = \emptyset$ is a basis contained in S.

Next suppose that S contains at least one non zero vector: This implies that for at least one non zero vector $\{v_1\} \subseteq S$, that v_1 is linearly independent. We can continue in an inductive manner:

Case 1: Suppose all other vectors in S, are linear combinations of v_1. Then $\{v_1\}$ is a basis and we are done.

Case 2: Suppose $\exists v_2 \in S$ not a multiple of v_1, then $\{v_1, v_2\}$ is linearly independent.

This process continues, until it finishes, we know that it finishes because S is a finite subset of V.

😊

Note:-

Convention

For $V = \{0_v\}$, $S = \emptyset$ is the basis.

Example 5.3.1 (Example 1)

Let $V = \mathbb{R}^3$ and $S = \{(1, -2, 3), (0, 1, 1), (2, -9, 1), (1, 1, 11)\}$.
There is a preliminary step left as practice to show that span$(S) = \mathbb{R}^3$.
Step 2: we are going to extract a basis from S. Start with the first non zero vector.

$S_1 = \{(1, -2, 3)\}$ and this is linearly independent because v_1 is not the zero vector.

Check if $(1, -2, 3), (0, 1, 1)$ are independent. They are.
Let $S_2 = \{(1, 02, -3), (0, 1, 1)\}$

Now check if the third vector is in the span of the first two.

Check if $(2, -9, 1) \in \text{span}\{(1, -2, 3), (0, 1, 1)\}$. Look for $a, b \in \mathbb{R}$ such that $(2, -9, 1) = a(1, -2m3) + b(0, 1, 1)$

There is a solution:

$a = 2$
$-2a_b = -9$

$3a + b = 1$

$a = 2, b = 5$. This shows that this vector is indeed in the span and so it shouldn't be included in the basis.

Check if the last vector is in the span of the first two that are in S_2 :

$(1, 1, 11) = a(1, -2, 3) + b(0, 1, 1)$. It is not which means that we have found our basis. This implies that:

$B = \{(1, -2, 3), (0, 1, 1), (1, 1, 11)\}$

Note:-

Question about the proof

How do we know that the linearly independent set spans the vector space. You know the original subset spans the set. So by getting rid of the redundancies we know that we are done.

Note:-

There could be multiple bases

We could have started with a different first vector.

Note:-

Make the assumption that we start with the finite basis

Generally there are other vector spaces that do not meet a finite basis.

Example 5.3.2 (Polynomials)

$V = F[x]$, has an infinite basis. To show that a basis exists we use Zorm's Lemma

Theorem 5.3.3 Replacement Theorem

Let V be a vector space over F and suppose that V has a finite generating set defined as $H = \{u_1, \ldots, u_n\}$ with n elements. Suppose $S = \{v_1, \ldots, v_m\}$ is a linearly independent subset of V. Then:

1. $m \leqslant n$
2. $\exists T \subseteq H$ containing exactly $n - m$ vectors such that $T \bigcup S$ spans V.

Note:-

The basis has less elements than tany set of linearly independent vectors

Note:-

Replacement theorem because we can create a new basis by replacing vectors in S

Note:-

This proof will be by induction on m (the number of elements in the set S)

Proof: Base Case: Suppose $m = 0$. This means that $S = \emptyset$. Then take $T = H$ and we are done. Clearly $0 \leqslant n$. And step two is satisfied if we let $T = H$. We know that the span of H is the entire space V.

Note:-

(my answer)Induction Hypothesis: For some $k \geqslant 0$ supassume that for a finite generating set H and linearly independent set S we have that the theorem is true.

Induction Hypothesis: Suppose that the theorem is true for all linearly independents subsets $S_0 \subseteq V$ with exactly $m \geqslant 0$ elements. Meaning whenever $S_0 \subseteq V$, is a linearly independent subset of V, with exactly $m \geqslant$ elements, we get that $m \leqslant n$ and $\exists T_0 \subseteq H$ with exactly $n - m$ elements such that $S_0 \bigcup T_0$ spans V.

Induction Step: Let $S \subseteq V$ be a linearly independent subset with $m+1$ elements. We want to show that $m+1 \leqslant n$ and that $\exists T \subseteq H$ with $n-m-1$ elements such that $S \bigcup T$ spans V.
Let $S = \{v_1, ..., v_n + 1\}$
We have $\{v_1, ..., v_n\} \bigcup \{v_{n+1}\}$. Set $S_0 = \{v_1, ..., v_n\}$, then S_0 is a subset of S and S is independent, by HW2 we also know that S_0 is also independent, we also know that S_0 has m vectors. We can use the induciton hypothesis for S_0. This implies that $m \leqslant n$ and second that $\exists T_0 \subseteq H$ with $n-m$ elements such that $S_0 \bigcup T_0$ spans V.
Say $T_0 = \{u_1, ..., u_{n-m}\}$. Then $S_0 \bigcup T_0 = \{v_1, ..., v_m, u_1, ...u_{n-m}\}$. Now we need to introduce a new vector in a sensitive manner. We have that the span$(S_0 \bigcup T_0) = V$. This implies that v_{m+1} is a linear combination of $V_1, ..., v_m, u_1, ..., u_{n-m}$.
$v_{m+1} = a_1 v_1 + ... + a_m v_m + b_1 u_1 + ... + b_{n-m} u_{n-m}$ for some $a_1, a_m, b_1, b_{n-m} \in F$.
Claim 1 $m+1 \leqslant n$: these are integers, so it is enough to show that $m < n$. Proceed by contradiction. Suppose that $m = n$. That will imply that $T_0 = \emptyset$, has no elements. $V_{m+1} \in$span(S_0) implies that S is linearly dependent which is a contradiction. For the same reason, one of the b_i has to equal zero by the same argument.
Claim 2: $\exists \in \{1, ..., n-m\}$ such that $b_i \neq 0$. For the same reason, if all of the $b_i = 0$ then we would have $v_{m+1} \in$span(S_0). which is a contradiction.
Without loss of generality, assume that $b_1 \neq 0$, then our equation for v_{m+1}, allows us to solve for u_1.

$$S_0 = \{v_1, ..., v_m\}$$
$$S = \{v_1, ..., v_m, v_{m+1}\}$$
$$T_0 = \{u_1, ..., u_{n-m}\}$$
$$T = \{u_2, ..., u_{n-m}\}$$

Take $T = \{u_2, ..., u_{n-m}\}$ has $n-m-1$ vectors and span$(S \bigcup T)$ =span$(S_0 \bigcup T_0) = V$.

😄

Note:-

Review for the Replacement Theorem

Take everything as a given, and apply induction on m and let S be any S.

Start with a linearly independent set that has one more vector than the hypothesis allows. Extract the first n vectors and call S_0. Then apply the hypothesis. Because we know that S_0 spans V we can write it as a linear combination of S_0 and $T_0, T_0 = \{u_1, ..., u_{n-m}\}$. Because weknow the span is everything we can create our linear combination equation. Then we need to verify the elements in our list.

Chapter 6

Thursday September 12th

6.1 Last Time

Theorem 6.1.1 Replacement Theorem

Let V be a vector space over F suppose that V has a generating set H with n vectors. Let $S \subseteq V$. Linearly independent with m vectors.

1. Then $m \leqslant n$
2. $\exists T \subseteq H$ with exactly $n - m$ elements such that $S \bigcup T$ spans V.

6.2 Application of the Remainder Theorem

1. Let V be a vector space and suppose that V has a finite generating set. Then we can conclude:
   (a) V has a finite basis
   (b) Any two bases of V have the same number of vectors

Proof: Last time we showed that if H spans V, then we can extract a basis from H.
This implies that V has a finite basis. The interesting part is 2 from above.
Let B be a basis for V with n vectors. The previous step guarantees that such a B exists.
We want to show that any other basis has also n vectors.
Follow our noses and let Γ be another basis for V with m vectors.
One the one hand we have B is a basis, therefore, B spans V. Γ is a basis, implies that Γ is linearly independent.
We will apply the first part of the replacement theorem with $H = B$ and $S = \Gamma$.
Part 1 of the Replacement Theorem implies that $m \leqslant n$.
Similarly now, Γ spans V (as a basis) and B is linearly independent. Therefore again, the replacement theorem gives us $n \leqslant m$.
Now we are done, we have $m = n$.

😄

Definition 6.2.1: Dimension

Let V be a vector space over F having a finite basis B with n vectors. Then we say that V has dimension n, and we write that $\dim V = n$.

Note:-

$\dim V$ is the size of any basis

All bases have the same size.

6.3 Examples

Example 6.3.1 (Example 1)
$V = F^n$ implies that $\dim V = n$. This is because the standard basis has exactly n vectors.

Standard Basis: $\{(1, 0, ..., 0), ..., (0, 0, ..., 1)\}$ has n vectors.

Example 6.3.2 (Example 2)
What is the standard $V = M_{n\times m}(F) \longrightarrow \dim V = n \cdot m$

Example 6.3.3 (Example 3)
$V = P_n(F) = B_{st} = \{1, x, x^2, ..., x^n\}$.

Example 6.3.4 (Example 4)
Let $V = \mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$ considered as an $\mathbb{R}$ vector Space.

Question 3: What would be considered the basis for $\mathbb{C}$ viewed as an $\mathbb{R}$ vector space

Solution: $B_{st} = \{1, i\} \longrightarrow dim_{\mathbb{R}}\mathbb{C} = 2$.

While the $dim_{\mathbb{C}}\mathbb{C} = 1$.

Example 6.3.5 (Example 4')
What would be the $dim_{\mathbb{R}}M_{n\times m}(\mathbb{C}) = 2nm$. Whenever we take complex vector spaces and consider them as $\mathbb{R}$ vector spaces the dimension is multiplied by 2.

Note:-
The $\mathbb{R}$ vector spaces

$\mathbb{R}[x], \mathbb{F}(\mathbb{R}, \mathbb{R}), seq(\mathbb{R})$ are infinite dimensional, meaning they don't admit a finite basis.

6.4 Another Application of Replacement Theorem

Example 6.4.1 (Example 5)
Let V be a vector space over F of $dimn$. Let $S = \{v_1, ..., v_n\} \subseteq V$. Then the following are equivalent:

1. $span(S) = V$
2. S is linearly independent
3. S is a basis

Proof: We will show that $a \longrightarrow b \longrightarrow c$. This is sufficient to show that they are all equivalent.

$(c) \longrightarrow (a)$:, we don't need to do any work this is because if S is a basis of V then by definition we know that S spans V.

$(a) \longrightarrow (b)$: We assume that $span(S) = V$. We also know: By the statement of the theorem, that $dim(V) = n$, which is equal to the size of S.

Since $span(S) = V$, we get that S must contain a basis B for V (Result from last time). But $dim(V) = n \longrightarrow B$ has size n and $B \subseteq S$, implies that $B = S$ and S is linearly independent by the definition of B.
$(b) \longrightarrow (c)$: Suppose that S is linearly independent. Let B be a basis for V. Since the $dim(V) = n \longrightarrow B$ has size n. We will do the Replacement Theorem again.
We have S is linearly independent and B spans V. The second part of replacement theorem gives us that $\exists T \subseteq B$ and $S \bigcup T$ spans V and T has exactly size of B minus size of S is equal to 0 vectors which implies that $T \neq \emptyset$. Therefore S spans V.

😄

6.5 Effect of this Proof

Example 6.5.1 (Eample 7)

Let $S = \{\begin{bmatrix}1 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}-1 & 0\\1 & 0\end{bmatrix}, \begin{bmatrix}2 & 0\\3 & 1\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\}$.
To show that S is a basis for V, all we need to do is show that all of the elements in S are independent. This is because V is $M_{2\times 2}(F)$.

Since $dim(V) = 4 =$ size S, it is enough to show that S is linearly independent.

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

Adding co-ordinate wise:

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

This produces the following augmented matrix:

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

Which shows that $a = b = c = d = 0$.

Example 6.5.2 (Example 7)

$V = P_3(\mathbb{R})$, $S = \{1 + x^2, x^3 - 5x, 2\}$, S is not a basis since $dim(V) = 4$. Since S contains 3 polynomials of different degree, S is linearly independent.

By the replacement theorem we can take the standard basis:
$B_{st} = \{1, x, x^2, x^3\}$

Now we have that S is independent with 3 vectos and B spans V, so the replacement theorem guarantees that we $\exists T \subseteq B$ with 1 vector such that $S \bigcup T$ is a basis.

Note:-
Often more than one will work

Sometimes one will not work. We will show why x^2 will not work either:

Actually: $1, x^2$ are in the span(S).

$x^2 = 1 + x^2 - \frac{1}{2} \cdot 2 \in$span$(S)$.

We can pick two of the other elements in T. This implies that we can use $B_1 = \{S \bigcup \{x\}\}$ or $B_2 = \{S \bigcup \{x^3\}\}$.

Example 6.5.3 (Example 8)

Let $W = \{A \in M_{n\times m}(\mathbb{F}) : A = A^t\}$ of symmetric $n \times n$ matrices.

Question 4

We want to find the dimension of W.

The most obvious thing to do is to find a standard basis, then the number of elements in the standard basis is the number.

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

$$B_{st} = \{\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \dots & 0 \\ 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}\} \bigcup \{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}\} \bigcup ... \bigcup \{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}\}$$

of the diagonals, and we union with all of the symmetric values where there is a one on both sides of $e_{ij} = e_{j,i}$.

We get $n + \frac{n^2-n}{2}$. Total of n^2 and we can subtract the middle diagonal, and we divide by two because we count two at a time.

Note:-

If $V = \{$ upper triangular matrices$\}$ then the $dimV = \frac{n(n+1)}{2}$.

6.6 Application 3

Example 6.6.1 (Example 9)

Let V be a vector space of $dim(V) = n$. Then every linearly independent subset of V can be extended to a basis.

Proof: Let B be any basis of V. By the Replacement Theorem, we can get $S \bigcup T$ to be a basis for some $T \subseteq V$.

☺

6.7 Application 4

Example 6.7.1 (Example 10)
Let V, $dim(V) = n$. Let W be a subspace of V, $W \leqslant V$: then

1. W is finite dimensional
2. Any Basis of W can be extended to a bais of V

Note:-
Proof is left as practice

Similar to previous proofs.

Note:-
This is an important ending point of chapter 1

6.8 Chapter 2: Linear Transformations

Definition 6.8.1

Let V, W be two vector spaces over the **same** field F. A linear transformation from V to W:

$T : V \longrightarrow W$ is a function such that:

1. $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$
2. $T(c \cdot v) = c \cdot T(v)$, for all $c \in F$, $v \in V$.

6.9 Properties of Linear Transformations

1. $T(0_v) = 0_w$

Proof: $T(0_v) = T(0_v + 0_v) = T(0_v) + T(0_v)$
$T(0v) = 2T(0_v) \longrightarrow T(0_v) = 0_w$

😀

2. T is linear if and only if $T(c \cdot v_1 + d \cdot v_2) = c \cdot T(v_1) + d \cdot T(v_2)$ for all $c, d \in F$ and $v_1, v_2 \in V$
3. $T(v_1 - v_2) = T(v_1) - T(v_2)$

Example 6.9.1 (Example 1)
Let V, W be arbitrary vector spaces over F. We can take $T : V \longrightarrow W$, where $v \longrightarrow 0_w$ for $v \in V$. This is the zero linear transformation.

Example 6.9.2 (Example 2)
Take $T : V \longrightarrow V$, to define T as the identity function.

Example 6.9.3 (Example 3)
Take $T : V \longrightarrow V$ for $T(v) = c \cdot v$ for some fixed scalar $c \in F$

Example 6.9.4 (Example 4)
Take $T : \mathbb{R}^3 \to \mathbb{R}^2$, where $T(x, y, z) = (3x - y, 5z - 11x + 2y)$

Check that T is linear: Let $v_1 = (x_1, y_1, z_1)$, and $v_2 = (x_2, y_2, z_2) \in \mathbb{R}^3$.

We want to show that $T(v_1 + v_2) = T(v_1) + T(v_2)$

$v_1 + v_2 = (x_1 + x_2, y_1 + y_2 + z_1 + z_2) \sim$
$T(v_1 + v_2) = (3(x_1 + x_2) - (y_1 + y_2), 5(z_1 + z_2) - 11(x_1 + x_2) + 2(y_1 + y_2))$

And this equals $(3x_1 - y_1, 5z_1 - 11x_1 + 2y_1) + (3x_2 - y_2, 5z_2 - 11x_2 + 2y_2)$. Which is equal to $T(v_1) + T(v_2)$.

Example 6.9.5 (Example 5)
Let $T : M_{n\times n}(F) \longrightarrow M_{n\times n}(F)$ sending the matrix $A \longrightarrow A^t$.

Hopefully we saw from HW1: $(A + B)^t = A^t + B^t$

And $(c \cdot A)^t = c \cdot A^t$. Both of these facts imply that T is linear.

Example 6.9.6 (Example 6)
$V = \{f : \mathbb{R} \to \mathbb{R}, f$ is infinitely differentiable $\}$. This means that f has all possible derivatives.

Then we can take $T : V \longrightarrow V, T(F) = F'$. Calculus pretty much tells us that this is linear.

$(f + g)' = f' + g'$
$(c \cdot f)' = c \cdot f'$

Example 6.9.7 (Example 7)
Let $V = \{f : \mathbb{R} \to \mathbb{R}$ continuous $\}$. Let $a, b \in \mathbb{R}$ where $a < b$, $T : V \longrightarrow \mathbb{R}$: $f \longrightarrow \int_a^b f(t)dt$

Again Linear Calculus tells us that $\int_a^b (c \cdot f + d \cdot g)dt = c \cdot \int_a^b f dt = d \cdot \int_a^b g dt$

6.10 Rotations in $\mathbb{R}^2$

Example 6.10.1 (Example 8)
Let $V = \mathbb{R}^2$, let $\theta \in (0, \pi)$ an angle.

A vector in $\mathbb{R}^2$ can be represented in the natural sense.

Let θ be some angle. Now let's say we rotate our vector in some counter clockwise direction by adding theta to it. See Drawing.

$T_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$:

$T_\theta(x, y) = (0, 0)$ if $(x, y) = (0, 0)$.
Otherwise: $T_\theta(x, y) =$counter clockwise rotation by θ.

Chapter 7

Tuesday September 17th

7.1 Midterm 1

Tuesday October 1st or October 8th. No Hw for the week. Go back and look at the HW. Focus more on conceptual problems.

Note:-

Every Finite dimensional space has a basis.

Definition 7.1.1: Linear Transformations

Linear Transformations: $T : V \longrightarrow W$ is linear by definition if $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$.

Additionally, $T(cv) = cT(v)$ for all $c \in F$ and $v \in V$.

Combine this for $T(cv_1 + dv_2) = cT(v_1) + dT(v_2)$

Definition 7.1.2: Rotations in $\mathbb{R}^2$

$V = \mathbb{R}^2$, $\Theta \in (0, \pi)$
$T_\Theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$:

$T_\Theta(x, y) = 0, 0$ if $(x, y) = (0, 0)$, otherwise counter clockwise rotation by Θ.

Example 7.1.1 (Verify that T is a linear function)
From the picture it is not completely clear. Let's find an explicit formula for T_Θ.

Start with a vector $v = (x, y)$ and we will be using polar co-ordinates. $v = (x, y) = r = \sqrt{x^2 + y^2}$ and $x = rcos\omega$ and $y = rsin\omega$. Now we want to explore what happens to $T_\Theta(x, y)$.

$T_\Theta(x, y) = (rcos(\omega + \theta), rsin(\omega + \theta)) = r(cos\omega cos\theta - rsin\omega rsin\theta, rsin\omega cos\theta + rcos\omega sin\theta$
$cos\theta rcos\omega - sin\theta rsin\omega, cos\theta rsin\omega + sin\theta rcos\omega$

$= (cos(\theta)x - sin(\theta)y, cos(\theta)y + sin(\theta)x) = T_\theta(x, y)$

Check linear combination becomes easy.

Definition 7.1.3: Kernel

Let $T : V \longrightarrow W$ of all of the linear transformations of T.

The kernel of T is $ker(T) = \{v \in V : T(v) = 0_w\} = T^{-1}(\{0_w\})$.

Definition 7.1.4: Image

The image of T is $Im(T) = T(V) = \{w \in W$ such that $\exists v \in V$ with $T(v) = w\}$.

Note:-

The $ker(T) \leqslant V$ and $Im(T) \leqslant W$. This is very immediate and follows from the definitions. The kernel of T is a subspace of the domain and the image of T is the subspace of the codomain.

Theorem 7.1.1

Let $T : V \longrightarrow W$ Linear transformation. Let $S = \{v_1, \ldots, v_n\} \subseteq V$ such that $span(S) = V$. Then the set $\{T(v_1), \ldots, T(v_n)\}$ spans $I_m(T)$.

Proof: We want to show that $span\{T(v_1), \ldots, T(v_n)\} = Im(T)$ this is double inclusion. Show the first direction. Let $w \in span\{T(v_1), \ldots, T(v_n)\}$, This measn that $\exists c_1, \ldots, c_n \in F$ such that $w = c_1T(v_1) + \ldots + c_nT(v_n)$. Because T is linear we have: $T(c_1v_1 + \ldots + c_nv_n)$ set $v = c_1v_1 + \ldots + c_nv_n \in V$. This implies that $w = T(v) \in Im(T)$. The other direction, let $w \in Im(T)$, this implies that there exists a v such that $w = T(v)$. Since $span\{v_1, \ldots, v_n\} = V$ $\exists a_1, \ldots, a_n \in F$ such that $v = a_1v_1 + \ldots + a_nv_n$. This implies that $w = T(v) = T(a_1v_1 + \ldots + a_nv_n)$. By definition of linearity $a_1T(v_1) + \ldots + a_nT(v_n)$. This means that $w \in span\{T(v_1), \ldots, T(v_n)\}$.

😊

Example 7.1.2 (Example 1)

$T : P_2(\mathbb{R}) \longrightarrow M_{2\times2}(\mathbb{R})$

$f(x) \longrightarrow \begin{bmatrix} f(1) - f(2) & 0 \\ 0f(0) & \end{bmatrix}$.

Compute $ker(T)$ and $Im(T)$ and find the bases for them.

$Ker(T) = \{f(x) \in P_2(\mathbb{R})$ such that $\begin{bmatrix} f(1) - f(0) & 0 \\ 0 & f(0) \end{bmatrix}\}$ is zero. We need $f(1) - f(2) = 0$ and $f(0) = 0$.

Note:-

We need to write down polynomials of degree of utmost 2.

We can write $f(x) \in P_2(\mathbb{R})$ in the form:
$f(x) = ax^2 + bx + c$ for some scalars $a, b, c \in \mathbb{R}$. we see that $f \in ker(T) \leftrightarrow f(0) = 0 \wedge f(1) = f(2)$

This implies that $f(0) \longrightarrow c = 0$ and $a + b + c = 4a + 2b + c$. This implies that $c = 0$ and $b = -3a$.

This tells us that $ker(T) = \{f(x) = ax^2 - 3ax, a \in \mathbb{R}\}$.
If $f(x) \in Ker(T) \longrightarrow f(x) = a(x^2 - 3x)$ for some $a \in \mathbb{R}$. Then this implies that $ker(T) \subseteq span\{x^2 - 3x\}$. But since $x^2 - 3x \in ker(T)$ then we can see that $ker(T) = span\{x^2 - 3x\}$.

This implies that $B = \{x^2 - 3x\}$ is a basis for $ker(T)$ that is also linearly independent.

Example 7.1.3 (The image for the same problem)

We can start with the standard basis for the polynomials, $B_{st} = \{1, x, x^2\}$. Theorem 1 tells us that $\{T(1), T(x), T(x^2)\}$ spans the image.

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

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

$T(x^2) = \begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}$.

Note that the first two are independent. But the third is a linear combination of the second.

$\{T(1) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, T(x) = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}\}$ is a basis for $Im(T)$.

Note:-

$dimKer(T) = 1$ and $dimIm(T) = 2$ and $dim(P_2(\mathbb{R}) = 3$. This proof we will start with on Thursday.

Note:-

Let $T : V \to W$ be linear. Then T is one to one as a function if and only if $ker(T) = \{0_v\}$.

Proof: Suppose T is 1:1 as a function. Let $v \in ker(T) \to T(v) = 0_v$ but because T is linear $T(v) = T(0_v)$ and because T is one to one, we get that $v = 0$ and we are done:
$ker(T) = \{0_v\}$.
For the other direction, suppose that $ker(T) = \{0_v\}$ and we want to show that T is 1:1. Let $v_1, v_2 \in V$ such that $T(v_1) = T(v_2) \to T(v_1) - T(v_2) = 0 \to T(v_1 - v_2) = 0 \to v_1 - v_2 \in ker(T) = \{0_v\} \to v_1 = v_2$.

😊

Theorem 7.1.2 Dimension Theorem

Let $T : V \to W$ be linear. Suppose that $dimV = n \to dimV = n = dimKer(T) + dimIm(T)$.

The proof we be similar to some of this week's HW.

Proof: $dimV = n$, $Ker(T) \leqslant V \to$every basis of $ker(T)$ can be extended to a basis of V.
Set $m = dimKer(T) \to m \leqslant n$.
Let $B_0 = \{v_1, ..., v_m\}$ be a basis of $kerT$ from the above we can extend it for a basis of B_0.
$B = \{v_1, ..., v_m, v_{m+1}, ..., v_n\}$. Our goal is to show that $dimIm(T) = n - m$.
Theorem one of the day implies that $\{T(v_1), ..., T(v_m), T(v_{m+1}), ..., T(v_n)\}$ spans$Im(T) \in Ker(T)$. Since $T(v_1), ..., T(v_m)$ are elements of the kernel. Which implies they are equal to zero, we don't need them and so $\{T(v_{n+1}, ..., T(v_n)\}$ is enough to span $Im(T)$.
If we know that these guys are linearly independent then we are done. Remains to show that $T(v_{n+1}), ..., T(v_n)$ are linearly independent because then B_1 which is the new basis without 0 elements would be a basis for the image.
Let $a_{m+1}, ..., a_m \in F$ such that $a_{m+1}T(v_{m+1}) + ... + a_nT(v_n) = 0$. We want to show that $a_i = 0$.
Because T is linear we can rewrite this as $T(a_{m+1}v_{n+1} + ... + a_nv_n) = 0$. We can say for sure that the vector for sure inside the argument is in the kernel of T. Since b_0 is a basis for $ker(T)$ we can find $a_1, ..., a_m \in F$ such that $a_{m+1}V_{m+1} + ... + a_nv_n = a_1v_1 + ... + a_mv_m$. Now we can put everything to one side. And we know that this equals zero. Because B is a basis we know that all of the coefficients have to be zero. This implies that $dimIm(T) = n - m$.

😊

Definition 7.1.5

Let $T : V \longrightarrow W$ be linear. We say that:

1. T is injective if $ker(T) = 0$ or equivalently if T is one to one.
2. We say that T is surjective if the image of $T = W$
3. Or bijective, or there exists an isomorphism if it is both.

Note:-

Let V, W be vector spaces over F with the same dimension: $dimV = dimW$. Let $T : V \longrightarrow W$ be linear. Then the following are equivalent:

1. T is surjective
2. T is injective
3. T is an isomorphism

Note:-

If we have a subspace of a vector space: $W \leqslant V$ if $dimW = dimV$, then $W = V$.

This is because any basis for W can be extended for a basis for V.

Chapter 8

Thursday September 19th

Note:-

Midterm will be on October 1st

Theorem 8.0.1 Dimension Theorem

Let V, W be vector spaces over F and $T : V \longrightarrow W$ a linear transformation. Then the conclusion is that $dimV = dimKer(T) + dimIm(T)$.

Note:-

Suppose that $dimV = dimW = n$, $T : V \longrightarrow W$, then the following are equivalent:

1. T is injective (i.e.) 1:1
2. T is surjective meaning onto
3. T is an isomorphism

Example 8.0.1 (Example 1)

Take $T : P_2(\mathbb{R}) \longrightarrow P_3(\mathbb{R})$ where $f(x) \longrightarrow 5f'(x) + \int_0^x 2f(t)dt$. Compute the kernel, and the image and find basis for them.

Solution: If we start with a basis of the domain, and apply T to it, it will always span the image.

Let $B_{st} = \{1, x, x^2\}$ be the standard basis for $P_2(\mathbb{R})$.

1. $T(1) = 2x$
2. $T(2) = 5 + x^2$
3. $T(3) = 10x + \frac{2}{3}x^3$

Claim is that we can immediately write down $Im(T), Ker(T)$ and find bases for them.

We know from last time that the image will actually be the span$(T(1), T(x), T(x^2))$.

These are three polynomials of different degrees, so they are linearly independent.

This gives us $B = \{2x, 5 + x^2, 10x + \frac{2}{3}x^3\}$ is a basis for $Im(T)$.

We know that $Ker(T) = \{0\}$ because by the dimension theorem we get that the dimension of the kernel has to be zero. $DimKer(T) = dimP_2(\mathbb{R}) - dimIm(T)$.

Note:-

Note that if the kernel is just 0, as in the example above, we know that the function is injective.

Theorem 8.0.2 Construction of Linear Maps

Let V, W be vector spaces over F. Let $B = \{v_1, ..., v_n\}$ be a basis for V.

Let $w_1, ..., w_n$ be arbitrary vectors in W. Then, $\exists$ unique linear transformation $T : V \longrightarrow W$ such that $T(v_i) = w_i$ for all $i = 1, ..., n$.

Note:-

This is a very powerful theorem.

Proof: We need to prove existence and uniqueness. We will start with existence.

We want a linear transformation $T : V \longrightarrow W$ that sends $v_i \longrightarrow w_i$, for all $i = 1, ..., n$.
Let $v \in V$, we want to define $T(v)$. Because B is a basis: $B = \{v_1, ..., v_n\}$, the vector V can be written uniqueley as a linear combination among the $v_i's$ in B for unique scalars $a_1, ..., a_n \in F$.
Define $T(v) = a_1w_1 + ... + a_nw_n$. We need to verify that this satifies what we want, that it is linear and that indeed each v_i goes to w_1.
Verify T satisfies these desired properties:First $T(v_i) = w_i$ for $i = 1, ..., n$. v_i has a unique expression as $v_i = 0v_1 + ... + 1v_1 + ... + 0v_n$, this implies that $T(v_i) = 1w_i = w_i$. As desired.
T is linear: $v_1, v_2 \in V$, and $c, d \in F$. We want to show that $T(c \cdot u_1 + du_2) = cT(u_1) + dT(u_2)$.
Write u_1 uniquely as a linear combination of the u_i and we do the same for u_2:
v_1, v_2 have unique expressions, as $u_1 = a_1v_1 + ... + a_nv_n$ and $v_2 = b_1v_1 + ... + b_nv_n$ for unique scalars $a_i, b_i \in F$.
This implies that $cu_1 + du_2 = c(a_1 + db_1)v_1 + ... + (ca_n + db_n)v_n$.
Since $\{v_1, ..., v_n\}$ is a basis for V, the previous line gives the unique expression of $cu_1 + du_2$ as a linear combination of $v_1, ..., v_n$. This implies that $T(cu_1 + du_2) =$
$c(a_1 + db_1)w_1 + ... + (ca_n + db_n)w_1$
$cT(u_1) + dT(u_2)$.
Thus T exists.

☺

Now we show uniqueness.

Proof: Suppose that we have another linear transformation $S : V \longrightarrow W$ is a linear transformation with $S(v_i) = w_i$ for all $i = 1, ..., n$. We want to show that $S = T$. Meaning, we want to show that $S(v) = T(v) \forall v \in V$.
Let $v \in V \longrightarrow v$ has a unique expression as $v = a_1v_1 + ... + a_nv_n$ with $a_i \in F$.
Then if we apply S, $S(v) = S(a_1v_1 + ... + a_nv_n)$, because S is assumed to be linear we can do $a_1S(v_1) + ... + a_nS(v_n)$ by assumption $a_1w_1 + ... + a_nw_n$ and this is exactly how we defined $T(v)$. Thus the two must be the same and they are unique.

☺

Example 8.0.2 (Example 2)

Take $V = \mathbb{C}^3$ and $W = M_{2\times2}(\mathbb{C})$. We can fix the standard basis $B = \{e_1, e_2, e_3\}$ as a standard basis for C^3.

Given any 3 matrices A_1, A_2, A_3 in $M_{2\times2}(\mathbb{C})$ we can find a linear transformation $T : C^3 \longrightarrow M_{2\times2}(\mathbb{C})$ that sends $e_1 \longrightarrow A_1$ $e_2 \longrightarrow A_2$ and $e_3 \longrightarrow A_3$.

Example 8.0.3 (True False)

1. Suppose that $S = \{v_1, ..., v_n\}$ spans V, $w_1, ..., w_n$ arbitrary vectors in W. Then $\exists T : v \longrightarrow W$ for $v_1 \longrightarrow w_i$ for $i = 1, ..., n$. ***Solution:*** We can't create unique representations for the vectors in v because they are not linearly independent.

Counter Example: Let $V = \mathbb{R}^2, W = \mathbb{R}^3$. Let $S = \{(1,0),(0,1),(1,1)\}$ and let $W_1 = (1,0,0), W_2 = (0,1,0), W_3 = (0,0,1)$. Then we know that there is not a Linear transformation that does this.

2. Suppose $S = \{v_1, ..., v_n\} \subseteq V$ is linearly independent $w_1, ..., w_n \in W \longrightarrow \exists T : V \longrightarrow W$, linear such that $T(v_i) = w_i, i = 1, ..., n$.
Solution: You can add vectors to the basis and corresponding vectors to the new set. We will lose uniquess in the theorem.

Note:-
If S is not a basis then T will not be unique.

Example 8.0.4 (Application)

Let V, W be vector spaces over F. Assume that they have the same dimension: $dim(V) = dim(W)$. Then $\exists T : V \longrightarrow W$ which is an isomorphism.

Proof: Consider Bases $B_v = \{v_1, ..., v_n\}$ and $B_w = \{w_1, ..., w_n\}$. By the theorem there exists a unique linear transformation T sending each B_i to each V_i.

By construction the w_i are in the image of T. The image is a subspace, of W therefore we get that span$\{w_1, ..., w_n\}$ to be inside the $Im(T)$. But this is a basis for W, which implies that the span is everything. This gives us that the image of T is W and so T is surjective.

Dimension theorem tells us that we have an isomorphism because the two have the same dimension.

😄

8.1 The Matrix Representation of a Linear Transformation:

Definition 8.1.1

Let V be a finite dimensional vector space. An ordered basis for V is $B = \{v_1, ..., v_n\}$ with prescribed order, meaning that $\{v_1, v_2, ..., v_n\}$.

Note:-
$B_1 = \{v_1, v_2, ..., v_n\}$ and we take $B_2 = \{v_2, v_1, ..., v_n\}$, these are different.

Now if $B = \{v_1, ..., v_n\}$ ordered basis for V, and we have $v \in V$, then v has a unique expression of the basis vectors: $a_1v_1 + ... + a_nv_n$

Definition 8.1.2

The co-ordinate vector of v with respect to the basis B is $[v]_b = (a_1, ..., a_n)$ column vector in F^n.

Example 8.1.1

Take $V = \mathbb{R}^3$ and Take $B_1 = \{(1,0,0),(0,1,0),(0,0,1)\}$
$B_2 = \{(0,0,1),(1,0,0),(0,1,0)\}$
$B_3 = \{(2,1,0),(-1,0,1),(0,0,1)\}$.

Given a $v \in \mathbb{R}^3$ we want to find $[v]_{b_1}$, $[v]_{b_2}$ and $[v]_{b_3}$

We start with an arbitrary vector x, y, z and we need to work with the basis.

We look for $a, b, c \in \mathbb{R}$ such that $(x, y, z) = a(2,1,0) + b(-1,0,1) + c(0,0,1)$. This gives us:

$2a - b = x$
$a = y$
$b + c = z$

We need to solve for a, b, c in terms of x, y, z.

$b = 2y - x$,$a = y$,$c = z + x - 2y$.
$[v]_b = (x, y, z)$, $[v]_{b_2} = (z, x, y)$, $[v]_{b_3} = (y, 2y - x, z + x - 2y)$.

Note:-

$T : V \longrightarrow W$ linear. $B_v = \{v_1, ..., v_n\}$, $\Gamma = \{w_1, ..., w_n\}$ ordered basis for v_1, w.

If I take $T(v_1) = a_{11}w_1 + a_{21}w_2 + ... + a_{mn}$.

This is why we write $T(v_i) = a_{ij}w_1 + ... + a_{mj}v_m$.

This way $[T]_b^\Gamma$ where the first column is given by $T(v_1)$.

Note:-

$[T]_b^\Gamma$ is a matrix and we call it the matrix representation of T with respect to the basis B, Γ.

Chapter 9

Tuesday September 24th

Note:-
Midterm Next Tuesday

Note:-
Material is up to today

Definition 9.0.1: Notation

Given vector space V and $B = \{v_1, ..., v_n\}$ a basis for V, for $v \in V$ we write $[v]_b = (a_1, ..., a_n)$ as a column vector.

Definition 9.0.2

Column vector such that $v = a_1v_1 + ... + a_nv_n$
$T : V \longrightarrow W$ linear transformation
$B = \{v_1, ..., v_n\}$ as an ordered basis
$\Gamma = \{w_1, ..., w_m\}$ as an ordered basis for V, W respectively.

$[T]_B^\Gamma = ([T(v_1)]_\Gamma, ..., [T(v_n)]_\Gamma)$ where each [] represents the ith column.

Definition 9.0.3

If $T(v_i) = \Sigma_{j=1}^m a_{ij}w_j$ for $i = 1, ..., n$

Then $[T]_B^\Gamma = \begin{bmatrix} a_{11} & ... & a_{n1} \\ \vdots & ... & \vdots \\ a_m & ... & a_{nm} \end{bmatrix}$

Example 9.0.1
$T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$
$T(a, b) = (3a - 5b + b + 8a, 2a - b)$
Compute $[T]_{B_{st}}^{B_{2st}}$ where $B_{st} = \{e_1, e_2\}$ and $B_{2st} = \{e_1, e_2, e_3\}$.

$T(1, 0) = (3, 8, 2) = 3e_1 + 8e_2 + 2e_3$
$T(0, 1) = (-5, 1, -1) = -5e_1 + e_2 - 1e_3$

$$[T]_{B_{st}}^{B_{st}} = \begin{bmatrix} 3 & 5 \\ 8 & 1 \\ 2 & -1 \end{bmatrix}$$

Note:-

Get as many columns as the dimension of V.

Note:-

Notation:

Example 9.0.2

$T : V \longrightarrow V$ and just one ordered basis for V, just write $[T]_B$ instead of $[T]_B^B$

Generally:

$T : (V, B) \longrightarrow (W, \Gamma)$

Note:-

Different choices of ordered basis will give different matrix representations.

Example 9.0.3

Consider the same transformation from example 1:

But take now $B' = \{(1,1), (1,0)\}$ basis for $\mathbb{R}^2$

$\Gamma = \{e_1, e_2, e_3\}$. The first column will stay the same but the second column changes:

$T(1,0) = 3e_1 + 8e_2 + 2e_3$

$T(1,1) = (-2, 9, 1) = -2e_1 + 9e_2 + e_3$

This implies that $[T]_{B'}^{\Gamma} = \begin{bmatrix} 3 & -2 \\ 8 & 9 \\ 2 & 1 \end{bmatrix} \neq [T]_B^{\Gamma}$

Example 9.0.4 (Special Case)

$T : F^n \longrightarrow F^n$, and we take the standard basis for both:

B, Γ =the standard basis for F^n and F^n.

$A = [T]_B^{\Gamma} = (a_{i,j} \in M_{n\times m}(F)$, then $T(v) = A \cdot v$ where this is now matrix multiplication.

Meaning $T(v) = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \dots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ if $v = (x_1, \dots, x_n)$.

Theorem 9.0.1

Let $T : V \longrightarrow W$ be linear transformation B, Γ ordered bassi for V, W respectively. Then: for all $v \in V$, To compute $T(v)$ it is enough to tell me how to compute in the terms of $[T(v)]_\Gamma = [T]_B^\Gamma \cdot [v]_B$.

Note that we have on the left $m \times n$ matrix and on the right $n \times 1$ matrix. Thus it is well defined.

Note:-

As a result of this theorem. We get that $[T]_B^\Gamma$ completely determines the linear transformation T.

Proof: Write $B = \{v_1, \ldots, v_n\}$ and $\Gamma = \{w_1, \ldots, w_n\}$.
Let $v \in V \longrightarrow v$ has unique expression as $c_1 v_1 + \ldots + c_n v + n$ for unique scalars $c_1, \ldots, c_n \in F$.
This implies that:

$$[v]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \text{ By definition.}$$

Because $T(v)$ is linear, $c_1 T(v_1) + \ldots + c_n T(v_n)$ which means that for each $i = 1, \ldots, n$ let
$T(v_i) = a_{1i} w_1 + \ldots + a_{mi} w_m$. In other words:
$T(v_i) = \sum_{j=1}^m a_{ji} w_j$ for unique scalars $a_{ij} \in F$.
Now back in terms of $T(v) = \Sigma_{i=1}^n c_i \Sigma_{j=1}^m a_{ij} w_j$
But from here we can compute:
First we need to exchange the summations, something times w_1 plus something times w_m.
These coefficients will give me the let hand side of our equation with matrix multiplication.

$$= \Sigma_{j=1}^m \Sigma_{i=1}^m c_i a_{ji} w_j$$

$$[T]_\Gamma = \begin{bmatrix} \Sigma_{i=1}^n c_i a_i j \\ \vdots \\ \Sigma_{i=1}^n c_i a_{mi} \end{bmatrix}$$

Note:-

Right hand Side now:

$$= [T]_B^\Gamma \cdot [V]_B$$

$$= ([T(v_1)]_\Gamma \ldots [T(v_n)]_\Gamma) \cdot \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

😆

Example 9.0.5

Determine the linear Transformation $T : \mathbb{R}^3 \longrightarrow P_2(\mathbb{R})$ such that $[T]_B^\Gamma = \begin{bmatrix} 1 & -3 & 4 \\ 0 & 2 & 1 \\ 2 & 0 & 0 \end{bmatrix}$. Where $B = \{(1,1,0), (1,1,1), (0,1,0)\}$ and $\Gamma = \{1, 1+x, 1+x^2\}$.

Question 5

What would we need to do in order to compute a formula, $T(a,b,c)$.

Solution: Solution is the problem

1. Step 1: for arbitrary $(x,y,z) \in \mathbb{R}^3$compute $[v]_B$.

   Look for $a, b, c \in \mathbb{R}$ such that $(x,y,z) = a \cdot (1,1,0) + b \cdot (1,1,1) + c \cdot (0,1,0)$. Create a system and show this for a, b, c in terms of x, y, z.

$$(x, y, z) = (a + b, a + b + c, b)$$
$$x = a + b$$
$$y = a + b + c$$
$$z = b$$

This implies that $a = x - z$,$c = y - 2z - x$, and that $b = z$.

$$\text{This implies that } [v]_B = \begin{bmatrix} x - z \\ z \\ y - 2z - x \end{bmatrix}$$

2. One way is to compute $[T(v)]_\Gamma = \begin{bmatrix} 1 & -3 & 4 \\ 0 & 2 & 1 \\ 2 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x - z \\ z \\ y - 2z - x \end{bmatrix}$ but this is extra work now because I need to compute the basis with respect to w.

3. The fastest way is to compute T with the standard basis of $\mathbb{R}^3$. Compute $T(e_1), T(e_2), T(e_3)$. Doing this is sufficient.

$$[T(e_1)]_\Gamma = \begin{bmatrix} 1 & -3 & 4 \\ 0 & 2 & 1 \\ 2 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \\ 2 \end{bmatrix}$$

This implies that $T(e_1) = -3 \cdot 1 + (-1) \cdot (1 + x) + 2 \cdot (1 + x^2)$

$$= T(e_1) = 2x^2 - x - 2$$

Same computation for e_2

$$[T(e_2)]_\Gamma = \begin{bmatrix} 1 & -3 & 4 \\ 0 & 2 & 1 \\ 2 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix} \text{ which implies that } T(e_2) = 4w_1 + w_2$$

Which means that $T(e_2) = x + 5$

Similarly $T(e_3) = -2x^2 - 14$

All together then $T(a, b, c) = aT(e_1) + bT(e_2) + cT(e_3)$
$T(a, b, c) = x^2(2a - 2c) + x(-a + b) + 5b - 2a - 14c$

Note:-

Recap of the steps

1. Given $[T]_B^\Gamma$ to compute $T(v)$:

2. Compute the column vector with respect to the basis B: Compute $[v]_B$

3. Compute $[e_i]_B$ for $B_{st} = \{e_1, \ldots, e_n\}$ which is the standard basis for V.

4. Step 3 is to compute $T(e_i)$ using the formula:
$[T(e_i)]_\Gamma = [T]_B^\Gamma \cdot [e_i]_B$ These $T(e_i)$ determine T.

Note:-

1. Let $T : V \to W$
The zero transformation meaning every v goes to 0_w. Then no matter the choice of basis I take, we will always get the 0 matrix:

   $[T]_B^\Gamma = 0$ matrix

2. Let B be an ordered basis for V, then if we take $1v : V \longrightarrow V$, the identity transformation where $v \longrightarrow v$. Then the matrix with respect to the basis will be the idenity matrix:
   $[1v]_B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

3. $[1_v]_B = I_n$ = identity $n \times n$ matrix, with $n = \dim V$

   But consider $(V, B) \longrightarrow (V, \Gamma)$, if $B \neq \Gamma$ then $[1v]_B^\Gamma \neq I_n$.

Question 6

What can you tell me about this matrix?

Solution: I can go backwards, perhaps go Γ to B. But $P = [1_v]_B^\Gamma$ the guess is that P is invertible $n \times n$ matrix. With the inverse $P^{-1} = [1v]_\Gamma^B$.

Note:-

We will call this a change of basis matrix.

Note:-

Suppose we are given $T_1, T_2 : V \longrightarrow W$, both lineary transformations.

Then the claim is that $T_1 + T_2$ is also linear.

Proof: Let $v_1, v_2 \in V$ and $c, d \in F$: Then $(T_1 + T_2)(cv_1 + dv_2) = T_1(cv_1 + dv_2) + T_2(cv_1 + dv_2)$
$= cT_1(v_1) + dT_1(v_1) + cT_2(v_1) + dT_2(v_2)$
$= c(T_1(v_1) + T_2(v_1)) + d(T_1(v_2) + T_2(v_2))$
$= c(T_1 + T_2)(v_1) + d(T_1 + T_2))v_2$
Thus $T_1 + T_2$ is linear.

☺

Note:-

Similarly if $T : V \longrightarrow W$ is linear then $c \in F$ a scalar then $cT : V \longrightarrow W$ is linear. $v \longrightarrow c \cdot T(v)$.

Definition 9.0.4: Given F vector spaces V, W let $L(v, w) = \{T : V \longrightarrow W\}$ be the set of all linear transformations from V to W.

Question 7

What can I say about this L by summarizing these two remarks?

Solution: **This implies that L should be a vector space over the same field F.**

Note:-

Let $B = \{v_1, ..., v_n\}$ and $\Gamma = \{w_1, ..., w_n\}$, ordered bases for V, W respectively. Let T_1, T_2 $V \longrightarrow W$ be a linear transformation and $c, d \in F$. Then we have:
$[cT_1 + dT_2]_B^\Gamma = c[T_1]_B^\Gamma + d[T_2]_B^\Gamma$.

Chapter 10

Tuesday October 8th

10.1 Overview of Chapter 3

Note:-

x, b are vectors in this section

Given a system of linear equations: $a_{11}x_1 + a_{12}x_2 + \ldots + a_m x_n = b1$ $a_{m1}x_1 + \ldots + a_{mn}x_n = bm$
Where the set of $x_1, \ldots, x_n$ to be the unknowns.

Definition 10.1.1: Homogeneous

$A = (a_{ij}) \in M_{m\times n}(F)$, let $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$. $Ax = b$ sepcial case $b = 0_v$ the system is homogeneous

Note:-

The main method to solve this is to use row reductions, reduced row echelon form

10.2 Immediate Deductions

Suppose we have $Ax = 0$, then $L_A : F^n \longrightarrow F^m$, $v \longrightarrow Av$, then the set of solutions is precisely $ker(L_A)$.

1. A Homogeneous System always has at least one solution, namely the 0 solution.
2. Suppose $m < n$. Then the dimension theorem gives us $dim(ker(L_A)) = n - dimIM(L_A)$. This is at least $n - m > 0$. This implies that the homogeneous system has infinitely many solutions.
3. Proposition: Let $b \in F^n$ suppose that $Ax = b$ has at least one solution, $s = \begin{bmatrix} s_1 \\ \vdots \\ s_n \end{bmatrix}$, then all other solutions are of the following form: $\{s + x : Ax = 0\}$.
4. Let $Ax = b$ be a linear system. Then if A is invertible, then there exists a unique solution. The unique solution is $x = A^{-1}b$
5. If $m < n$ then the homogeneous system has infinitely many solutions, The inhomogeneous $Ax = b$ either has no solutions, or infinitely many.
6. If $m = n$ and A^{-1} exists, then there is one unique solution.
7. If $m = n$ but A^{-1} does not exist then simir to $m < n$ case

10.3 Row and Column Operations

Note:-

Row Reductions

Theorem 10.3.1

Let $Ax = b$, be a linear system. With $A \in M_{m\times n}(F)$, let C be an invertible $m \times m$ matrix, then the system $Ax = b$ is equivalent to $CAx = Cb$.

Note:-

The idea is that each row reduction is given left multiplication by an invertible matrix.

10.4 Types of Operations

Fix $A \in M_{m\times n}(F)$

Note:-

For $r \in F$, $i \neq j$, $i, j \in \{1, \ldots, m\}$, we want to add $r \cdot j^{th}$ row to i^{th} row. This is achieved by the linear transformation:

$\mathbb{E}_{ij}(r) : M_{m\times n}(F) \longrightarrow M_{m\times n}(F)$, $A \longrightarrow E_{ij}(r) \cdot A$
Where $E_{ij}(r) = (a_{kl})$ where $a_{kl} = \{1, k = l, r \text{ if } k = 1, l = j0, \text{ else}\}$

10.5 Operation 2 is flipping 2 rows

Note:-

Flipping rows i and j, $F_{ij} = M_{m\times n}(F) \longrightarrow M_{m\times n}(F)$, where $A \longrightarrow F_{ij}A$ where F_{ij} is the matrix obtained from the identity by flipping the i and j row. Note that F_{ij} is invertible because the rows and the columns are linearly independent.

10.6 Multiply a row by a non zero scalar

Note:-

$\mathbb{E}_i(r)$: $M_{m\times n}(F) \longrightarrow M_{m\times n}(F)$, $A \longrightarrow C_i(r)A$ where $C_i r$ is the diagonal matrix with only r on ii

10.7 Column Operations

Note:-

Column operations are given by right multiplication of the elements we created for row operations. $E_{ij}(r), F_{ij}, C_i(r)$

Note:-

Can't use them when solving systems, because the system will change.

Note:-

We can use either row or column operations to compute the rank of a matrix.

Definition 10.7.1: Rank of a Matrix

The rank of a matrix A is:
$rk(A) = dimIm(L_A)$ dimension of the image of left multiplication by A.

Theorem 10.7.1 Why both row and col operations work

Let $A \in M_{m\times n}(F)$, $P \in M_{m\times m}(F)$, $Q \in M_{n\times n}(F)$ with p, q invertible. Then $rk(A) = rk(PA) = rk(AQ) = rk(PAQ)$

Proof: Want to show $rk(A) = rk(PA)$, $rk(A) = dim(Im(L_A))$ we want to show that it is the $dim(Im(L_{PA}))$, and last week we showed that this is the same as $dim(Im(L_P L_A))$, because P is invertibel L_P is an isomorphism so the rank doesn't change.

😊

10.8 Test Review

1. True. $T : F \to F$ linear, We want to show that $T(v) = cv$ for some $c \in F$

   Proof: $T(v) = T(v \cdot 1) = vT(1)$

   😊

2. False. Take all of them and show they are independent, argue that $dim(W_1 + W_2) = dim(W_1) + dim(W_2) - dim(W_1 \cap W_2)$. This implies that $dim(W_1 \cap W_2) \geqslant 1$ which implies that it is not a direct sum.
3. Skip 1.3
4. $\exists T : C^4 \to C^2$ such that $kerT = \{(x_1, x_2, x_3, x_4) \in C^4 : s.t. | x_1 = 3x_2, x_2 = x_3 = x_4\}$. The main thing to notice is that the kernel of T is spanned by one vector $\{(3, 1, 1, 1)\}$. If Such T existed then $dimKer(T) = 1$ $Dim(C^4) = 1 + dimIm(T)$ but this is impossible becuase the latter is at most 2.

1. $T : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}^2$.

   First compute the kernel and not the image. The kernel might give us all of the information we need. There is no way that we can guess what the kernel is. Must compute.

   Note:-

   $2a + b - c = 0$
   $d - 7a + 6b + 11c = 0$

   Immediately solve for $c = 2a + b$ the second one gives us $d = 7a - 6b - 11c$
   $c = -2a - b$, $d = -15a - 17b$, because of this we can compute a basis for the kernel. The kernel $T = \{(\begin{bmatrix} a & b \\ 2a + b & -15a - 17b \end{bmatrix})\}$ Where a, b are free variables. This is the span of $\{\begin{bmatrix} 1 & 0 \\ 2 & -15 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & -17 \end{bmatrix}\}$, this gives us that $dimKerT = 2$, because T is surjective, we can take the standard basis for the image.

2. Part B was harder. We want to construct a new linear map $F : M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ such that $kerF = ImF = KerT$. Start with $kerT = \text{span}\{\begin{bmatrix} 1 & 0 \\ 2 & -15 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & -17 \end{bmatrix}\}$ we need to take these two vectors and extend it to a basis for the entire vector space. We only need to extend the basis of $kerT$ to a basis of $M_{2\times 2}(\mathbb{R})\}$. Send $e_3 \to v_1$ and then $e_4 \to v_2$.

1. $W_1 = \{p(x) \in V : p(0) = p'(0) = 0\}$ and $W_2 = \text{span}\{1 + x, 2x^2 + x - 1\}$

   $W_1 = \{p(x) = ax^3 + bx^2 : a, b \in \mathbb{R}\}$ is an easier description.
   We only need to show one of the intersection is zero or that they add to V. By the Dimension theorem.

Chapter 11

Tuesday October 8th

11.1 Overview of Chapter 3

Note:-

x, b are vectors in this section

Given a system of linear equations: $a_{11}x_1 + a_{12}x_2 + \ldots + a_m x_n = b1$ $a_{m1}x_1 + \ldots + a_{mn}x_n = bm$
Where the set of $x_1, \ldots, x_n$ to be the unknowns.

Definition 11.1.1: Homogeneous

$A = (a_{ij}) \in M_{m\times n}(F)$, let $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$. $Ax = b$ sepcial case $b = 0_v$ the system is homogeneous

Note:-

The main method to solve this is to use row reductions, reduced row echelon form

11.2 Immediate Deductions

Suppose we have $Ax = 0$, then $L_A : F^n \longrightarrow F^m$, $v \longrightarrow Av$, then the set of solutions is precisely $ker(L_A)$.

1. A Homogeneous System always has at least one solution, namely the 0 solution.
2. Suppose $m < n$. Then the dimension theorem gives us $dim(ker(L_A)) = n - dimIM(L_A)$. This is at least $n - m > 0$. This implies that the homogeneous system has infinitely many solutions.
3. Proposition: Let $b \in F^n$ suppose that $Ax = b$ has at least one solution, $s = \begin{bmatrix} s_1 \\ \vdots \\ s_n \end{bmatrix}$, then all other solutions are of the following form: $\{s + x : Ax = 0\}$.
4. Let $Ax = b$ be a linear system. Then if A is invertible, then there exists a unique solution. The unique solution is $x = A^{-1}b$
5. If $m < n$ then the homogeneous system has infinitely many solutions, The inhomogeneous $Ax = b$ either has no solutions, or infinitely many.
6. If $m = n$ and A^{-1} exists, then there is one unique solution.
7. If $m = n$ but A^{-1} does not exist then simir to $m < n$ case

11.3 Row and Column Operations

Note:-
Row Reductions

Theorem 11.3.1
Let $Ax = b$, be a linear system. With $A \in M_{m\times n}(F)$, let C be an invertible $m \times m$ matrix, then the system $Ax = b$ is equivalent to $CAx = Cb$.

Note:-
The idea is that each row reduction is given left multiplication by an invertible matrix.

11.4 Types of Operations

Fix $A \in M_{m\times n}(F)$

Note:-
For $r \in F$, $i \neq j$, $i, j \in \{1, ..., m\}$, we want to add $r \cdot j^{th}$ row to i^{th} row. This is achieved by the linear transformation:

$\mathbb{E}_{ij}(r) : M_{m\times n}(F) \longrightarrow M_{m\times n}(F)$, $A \longrightarrow E_{ij}(r) \cdot A$
Where $E_{ij}(r) = (a_{kl})$ where $a_{kl} = \{1, k = l, r \text{ if } k = 1, l = j0, \text{ else}\}$

11.5 Operation 2 is flipping 2 rows

Note:-
Flipping rows i and j, $F_{ij} = M_{m\times n}(F) \longrightarrow M_{m\times n}(F)$, where $A \longrightarrow F_{ij}A$ where F_{ij} is the matrix obtained from the identity by flipping the i and j row. Note that F_{ij} is invertible because the rows and the columns are linearly independent.

11.6 Multiply a row by a non zero scalar

Note:-
$\mathbb{E}_i(r)$: $M_{m\times n}(F) \longrightarrow M_{m\times n}(F)$, $A \longrightarrow C_i(r)A$ where $C_i r$ is the diagonal matrix with only r on ii

11.7 Column Operations

Note:-
Column operations are given by right multiplication of the elements we created for row operations. $E_{ij}(r), F_{ij}, C_i(r)$

Note:-
Can't use them when solving systems, because the system will change.

Note:-
We can use either row or column operations to compute the rank of a matrix.

Definition 11.7.1: Rank of a Matrix

The rank of a matrix A is:
$rk(A) = dimIm(L_A)$ dimension of the image of left multiplication by A.

Theorem 11.7.1 Why both row and col operations work

Let $A \in M_{m\times n}(F)$, $P \in M_{m\times m}(F)$, $Q \in M_{n\times n}(F)$ with p, q invertible. Then $rk(A) = rk(PA) = rk(AQ) = rk(PAQ)$

Proof: Want to show $rk(A) = rk(PA)$, $rk(A) = dim(Im(L_A))$ we want to show that it is the $dim(Im(L_{PA}))$, and last week we showed that this is the same as $dim(Im(L_P L_A))$, because P is invertibel L_P is an isomorphism so the rank doesn't change.

😊

11.8 Test Review

1. True. $T : F \to F$ linear, We want to show that $T(v) = cv$ for some $c \in F$

Proof: $T(v) = T(v \cdot 1) = vT(1)$

😊

2. False. Take all of them and show they are independent, argue that $dim(W_1 + W_2) = dim(W_1) + dim(W_2) - dim(W_1 \cap W_2)$. This implies that $dim(W_1 \cap W_2) \geqslant 1$ which implies that it is not a direct sum.
3. Skip 1.3
4. $\exists T : C^4 \to C^2$ such that $kerT = \{(x_1, x_2, x_3, x_4) \in C^4 : s.t. | x_1 = 3x_2, x_2 = x_3 = x_4\}$. The main thing to notice is that the kernel of T is spanned by one vector $\{(3, 1, 1, 1)\}$. If Such T existed then $dimKer(T) = 1$ $Dim(C^4) = 1 + dimIm(T)$ but this is impossible becuase the latter is at most 2.

1. $T : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}^2$.

   First compute the kernel and not the image. The kernel might give us all of the information we need. There is no way that we can guess what the kernel is. Must compute.

   Note:-

   $2a + b - c = 0$
   $d - 7a + 6b + 11c = 0$

   Immediately solve for $c = 2a + b$ the second one gives us $d = 7a - 6b - 11c$
   $c = -2a - b$, $d = -15a - 17b$, because of this we can compute a basis for the kernel. The kernel $T = \{(\begin{bmatrix} a & b \\ 2a + b & -15a - 17b \end{bmatrix})\}$ Where a, b are free variables. This is the span of $\{\begin{bmatrix} 1 & 0 \\ 2 & -15 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & -17 \end{bmatrix}\}$, this gives us that $dimKerT = 2$, because T is surjective, we can take the standard basis for the image.

2. Part B was harder. We want to construct a new linear map $F : M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ such that $kerF = ImF = KerT$. Start with $kerT = \text{span}\{\begin{bmatrix} 1 & 0 \\ 2 & -15 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & -17 \end{bmatrix}\}$ we need to take these two vectors and extend it to a basis for the entire vector space. We only need to extend the basis of $kerT$ to a basis of $M_{2\times 2}(\mathbb{R})\}$. Send $e_3 \to v_1$ and then $e_4 \to v_2$.

1. $W_1 = \{p(x) \in V : p(0) = p'(0) = 0\}$ and W_2 =span$\{1 + x, 2x^2 + x - 1\}$

   $W_1 = \{p(x) = ax^3 + bx^2 : a, b \in \mathbb{R}\}$ is an easier description.
   We only need to show one of the intersection is zero or that they add to V. By the Dimension theorem.

Chapter 12

Thursday October 10th

12.1 Determinants

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(F)$

Definition 12.1.1: Determinant

The determinant of A or $\det(A)$ or $|A|$ is $ad - bc$

Example 12.1.1

$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 0 \\ 3 & 2 \end{bmatrix}$ note that $A + B = \begin{bmatrix} 0 & 2 \\ 6 & 6 \end{bmatrix}$. We have $\det(A) = \det(B)$ which is -2 while $\det(A + B) = -12$ this implies that det is not a linear funciton.

Note:-

But the determinant operation satisfies a different form of linearity.

Theorem 12.1.1

The function $\det : M_{2\times 2} \to F$. Is linear on each row if the other row is held fixed.

Note:-

For u, v, w row vectors in F^2, then the theorem says that $\det \begin{bmatrix} u + cv \\ w \end{bmatrix} = \det \begin{bmatrix} u \\ w \end{bmatrix} + c \det \begin{bmatrix} v \\ w \end{bmatrix}$ similarly in the other direction.

$\det \begin{bmatrix} u \\ w + cv \end{bmatrix} = \det uw + c \det uv$

Proof: Write $u = (a_1, a_2), v = (b_1, b_2), w = (d_1, d_2)$

$$\begin{bmatrix} u + cv \\ w \end{bmatrix} = \begin{bmatrix} a_1 + cb_1 & a_2 + cb_2 \\ d_1 & d_2 \end{bmatrix}$$

The RHS is: $= \begin{bmatrix} a_1 & a_2 \\ d_1 & d_2 \end{bmatrix} + c \begin{bmatrix} b_1 & b_2 \\ d_1 & d_2 \end{bmatrix}$

😄

Theorem 12.1.2

Let $A \in M_{2\times 2}(F)$ then $\det A \neq 0 \leftrightarrow A$ is invertible. In this case $A^{-1} = \frac{1}{\det(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Proof: $(\rightarrow)$ Suppose $\det(A) \neq 0$. Let $B = \frac{1}{\det(A)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. A quick verification yields $AB = BA = I \rightarrow B = A^{-1}$

$(\leftarrow)$ Prove Contrapositive: $\det(A) = 0 \rightarrow A$ not invertible. $\det(A) = ad - bc = 0$ if $a = b = c = d = 0 \rightarrow A = 0$ and $A \rightarrow$ not invertible. Withough loss of generality assume that $d \neq 0$. Then we can solve for a where $a = \frac{bc}{d}$. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \frac{bc}{d} & b \\ c & d \end{bmatrix}$ Then $u = \begin{bmatrix} \frac{bc}{d} \\ c \end{bmatrix}$ and the second column $v = \begin{bmatrix} b \\ d \end{bmatrix}$ this implies that $v = \frac{c}{d}u \rightarrow u, v$ are linearly independent. This implies that the $rank(A) < 2$ which implies that left multiplication is not an isomorphism. A is not invertible.

😃

Note:-

The det function is linear on each column if the other columns are held fixed.

12.2 Summary of Determinants

1. det is linear on each row or respectively column if the other row (respective column)is held fixed.
2. det of the identity matrix is 1.
3. $\det \begin{bmatrix} a & b \\ a & b \end{bmatrix} = 0$

Note:-

The key point is that properties 1,2,3completely determine the function.

If $\delta : M_{2\times 2}(F) \rightarrow F$ satisfying 1,2,3 above, then $\delta(A) = \det(A)$.

Example 12.2.1 (Application of uniqueness of det)

The area of a parallaleagram in $\mathbb{R}^2$. Start with $B = \{u, v\}$ as a basis in $\mathbb{R}^2$. $u = (a_1, a_2)$ and $v = (b_1, b_2)$. These are not parallel. We can form the parallelagram by the two. We will say that the area is $A\begin{bmatrix} u \\ v \end{bmatrix}$.

Theorem 12.2.1

$A\begin{bmatrix} u \\ v \end{bmatrix} = |\det(\begin{bmatrix} u \\ v \end{bmatrix})|$. Determinant has a nice geometric meaning. It computes something explicit.

Example 12.2.2 (Continued)

Preparation: We will prove instead the following. $\det\begin{bmatrix} u \\ v \end{bmatrix} = o\begin{bmatrix} u \\ v \end{bmatrix} A\begin{bmatrix} u \\ v \end{bmatrix}$. Where $o\begin{bmatrix} u \\ v \end{bmatrix}$ is what we call the orientation of the basis, or $o\begin{bmatrix} u \\ v \end{bmatrix} = \frac{\det\begin{bmatrix} u \\ v \end{bmatrix}}{|\det\begin{bmatrix} u \\ v \end{bmatrix}}|$

Note:-

Why the name orientation? We will show in Homework 7 that ($o \begin{bmatrix} u \\ v \end{bmatrix} = \pm 1$), u can be rotated in counter clockwise direction through an angle $\theta \in (0, \pi)$ to coincide with the direction of v

Note:-

Our strategy is to show that $o \begin{bmatrix} u \\ v \end{bmatrix} a \begin{bmatrix} u \\ v \end{bmatrix}$ satisfies properties one and three.

Proof: Set $\delta \begin{bmatrix} u \\ v \end{bmatrix} = o \begin{bmatrix} u \\ v \end{bmatrix} A \begin{bmatrix} u \\ v \end{bmatrix}$. Step 1 will be to extend δ to a function $\delta : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}$.

Let $u, v \in \mathbb{R}^2$ if u, v independent then define $\delta \begin{bmatrix} u \\ v \end{bmatrix} = 0 \begin{bmatrix} u \\ v \end{bmatrix} A \begin{bmatrix} u \\ v \end{bmatrix}$ if u, v independent define $\delta \begin{bmatrix} u \\ v \end{bmatrix} = 0$.

Step 2 is to verify that δ satisfies properties 2,3. $\delta \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = o \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1 \cdot 1$

Property 3 is clear by definition of δ

Step 3: Show $\delta \begin{bmatrix} u + c \cdot v \\ w \end{bmatrix} = \delta \begin{bmatrix} u \\ w \end{bmatrix} + c\delta \begin{bmatrix} v \\ w \end{bmatrix} \forall v, w, u \in \mathbb{R}^2, \forall c \in \mathbb{R}$.

Step 3.1: Prove linearity for scalars in the following sense: If you have $A = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$ then we want to show that

$$\delta \begin{bmatrix} a_1 & a_2 \\ cb_1 & cb_2 \end{bmatrix} = c\delta \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} = \delta \begin{bmatrix} ca_1 & ca_2 \\ b_1 & b_2 \end{bmatrix}$$

Proof: If $(a_1, a_2), (b_1, b_2)$ are dependent, then so are $(a_1, a_2), (cb_1, cb_2)$ and $(ca_1, ca_2), (b_1, b_2)$ implies that all sides are equal to zero.
Case 2 is if $c = 0$, this implies that these are all 0 for the same reason.
Case 3: $u = (a_1, a_2), v = (b_1, b_2)$ are independent and $c \neq 0$.
$A \begin{bmatrix} u \\ v \end{bmatrix} = h|v|$, $\begin{bmatrix} u \\ cv \end{bmatrix} = |c|h|v|$ together we get that $A \begin{bmatrix} u \\ cv \end{bmatrix} = |c| \begin{bmatrix} u \\ v \end{bmatrix}$

$O \begin{bmatrix} u \\ cv \end{bmatrix} = \dfrac{\det \begin{bmatrix} u \\ cv \end{bmatrix}}{|\det \begin{bmatrix} u \\ v \end{bmatrix}|} = \dfrac{c \det \begin{bmatrix} u \\ v \end{bmatrix}}{|c| |\det \begin{bmatrix} u \\ v \end{bmatrix}|}$ This gives us what we wanted if you carry out the calculations. Step 3.2:

Show $\delta \begin{bmatrix} u \\ u + w \end{bmatrix} = \delta \begin{bmatrix} u \\ w \end{bmatrix}$ for all $u, w \in \mathbb{R}$.
If u, w are linearly independent, then $u, w, +w$ are independent. Suppose u, w are indpendent 😃

😃

12.3 Thursday October 17th

Note:-

Review Permutations and their sign over the weekend

12.3.1 Reminders

Last time we discussed determinants for 2×2 Matrices. We showed that det satisifes:

1. det is linear on each row if the other rows are held fixed.
2. If I have a matrix that has two identical rows or two identical columns, then the det = 0.
3. det of $I_n = 1$.

4. $A \in M_{2\times 2}(F) \leftrightarrow \det(A) \neq 0$

12.3.2 Today we will do this for $n \times n$ Matrices

Note:-

$M_n(F)$ are $n \times n$ matrices

Example 12.3.1

$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 0 & -1 \\ 3 & 3 & 0 \end{bmatrix} = 1\begin{bmatrix} 0 & -1 \\ 3 & 0 \end{bmatrix} - 4\begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} + (-7)\begin{bmatrix} 2 & 0 \\ 3 & 3 \end{bmatrix} = x$$

Note:-

Notation. Let $A = (a_{ij}) \in M_n(F), n \geqslant 2$. For $i, j \in \{1, \dots, n\}$ denoted by $A_{ij} = (n-1)\times(n-1)$ matrix obtained from $\tilde{A}$ by deleting the i^{th} row and j^{th} column.

Thus $A_{11} = \begin{bmatrix} 0 & -1 \\ 3 & 0 \end{bmatrix}$

Definition 12.3.1: Recursive definition of Determinant Using Co-factor Expansion along the first row

Let $A = (a_{ij}) \in M_n(F), n \geqslant 2$. Then we define $\det(A) = \Sigma_{j=1}^{n}(-1)^{1+j} a_{ij} \det(\tilde{A}_{1j})$

Theorem 12.3.1 Theorem 1 Linearity of the Determinant

The function $\det : M_n(F) \to F$ is linear on each row if all the other rows are held fixed.

Meaning $A = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ u_i =row vector. Then this means $\det \begin{bmatrix} u_1 \\ \vdots \\ u_1 + c \cdot w \\ \vdots \\ u_n \end{bmatrix} = \det \begin{bmatrix} u_1 \\ \vdots \\ u_i \\ u_n \end{bmatrix} + c \det \begin{bmatrix} u_1 \\ \vdots \\ w \\ u_n \end{bmatrix}$ $\forall c \in F, w \in F^n$

Proof: Proceed by induction on $n \geqslant 2$. The base case is fine, we have already proved this. ($n = 2$.) Let $n \geqslant 3$. Suppose that the statement is true for $(n-1)(n-1)$ matrices.

Let $A = \begin{bmatrix} a_1 & \dots & a_m \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix}$ Suppose that the i^{th} row can be composed as follows:

$(a_{i1}, \dots, a_{in}) = b_{i1} + cd_{i1}, \dots, b_{in} + cd_{in})$ for some $c \in F$, $(b_{i1}, \dots, b_{in}), (d_{i1}, \dots, d_{in}) \in F^n$

Set $B = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ b_{i1} & \dots & b_{in} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix}$ and set $D = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ d_{i1} & \dots & d_{in} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix}$

Case 1: $i = 1$ implies that $B(b_{ij}), D = (d_{ij})$ observe that $a_{1j} = b_{1j} + cd_{1j}$ but A, B, C have rows $2-n$ the same $\tilde{A}_{1j} = \tilde{B}_{1j} = \tilde{D}_{ij}$ for all $j = 1, \dots, n$.

With all of this in mind we can just compute:

$\det(A) = \Sigma_{j=1}^{n}(-1)^{1+j} a_{1j} \det(\tilde{A}_{1j}) = \Sigma_{j=1}^{n}(-1)^{1+j}(b_{1j} + cd_{1j} \cdot \det(\tilde{A}_{1j})$

$= \Sigma_{j=1}^{n}(-1)^{1+j} \cdot b_{1j} \det(\tilde{B}_{1j}) + c\Sigma_{j=1}^{n}(-1)^{1+j} d_{1j} \det(\tilde{D}_{1j}) = \det(B) + c\det(D)$. For Case 2: $i > 1$. In this case, A, B, D

have the same first row. The $(i-1)$ row of $\tilde{B}_{1j} + c(i-1)$row of $\tilde{D}_{1j}$ These are $(n-1)\times(n-1)$ matrices which means that we can apply the induction hypothesis. This means that $\det(\tilde{A}_{1j}) = \det(\tilde{B}_{1j}) + c\cdot\det(\tilde{D}_{1j})\forall j = 1,\dots,n$ $\Sigma_{j=1}(-1)^n a_{ij} + c\cdot\Sigma_{j=1}^n(-1)^{1+j}a_{1j}\cdot\det(\tilde{D}_{1j})$ but either side is equal to $\det(\tilde{B}_{1j})$ and $\det(\tilde{D}_{1j})$.

Note:-

This proves linearity of a determinant

😊

Theorem 12.3.2 Theorem 2

Suppose $A\in M_n(F)$ has 2 identical rows, then $\det(A) = 0$.

Proof: $A = (a_{ij})\in M_n(F)$ has two identical rows. We want to show that $\det(A) = 0$. Again we will proceed by induciton on $n\geqslant 2$. Base case we already know. Let $n\geqslant 3$. Assume the statement is true for $(n-1)\times(n-1)$ matrices. A has two identical rows, say $u_i = u_j$ for some $i\neq j$. Since $n\geqslant 3$, then there is at least one more row u_k with $k\neq i,j$. We will compute the determinant using u_k, expand along the row k using these formulas here.

Compute $\det(A)$ using expansion along row k:

$\det(A) = \Sigma_{l=1}^n(-1)^{k+l}a_{kl}\det(\tilde{A}_{kl})$, because $k\neq i,j$ each of the $\tilde{A}_{kl}$ has two identical rows and $\tilde{A}_{kl}\in M_{n-1}(F)$ allows us to apply the induction hypothesis for $\tilde{A}_{kl}\longrightarrow\det(\tilde{A}_{kl}) = 0\forall l = 1,\dots,n\longrightarrow\det(A) = 0$.

😊

Theorem 12.3.3 Det Using cofactor expansion along row i

Let $A = (a_{ij})\in M_n(F)$, then $\det(A) = \Sigma_{j=1}^n(-1)^{i+j}a_{ij}\det(\tilde{A}_{ij})$.

Proof: Proof sketch. First prove Theorem 3 (current theorem) in the following special case. When the i_{th} row of A is $e_k = (0,\dots,1,\dots,0)$ for some $k\in\{1,\dots,n\}$. Note in this case $\det(A) = \Sigma_{j=1}^n(-1)^{i+j}a_{ij}\det(\tilde{A}_{ij})$, in the special case:

$$\begin{array}{cc} & \\ & \\ i & \\ & \\ & \end{array}\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ 0 & 1 & 0 \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} = (-1)^{i+k}\det(\tilde{A}_{i+k}).$$

Special Case. Implies theorem 3. Let $A = \begin{bmatrix} u_1 \\ \vdots \\ u_i \\ \vdots \\ u_n \end{bmatrix}$, $u_i = \Sigma_{j=1}^n a_{ij}e_j$ then the $\det(A) = \det(\begin{bmatrix} u_1 \\ \vdots \\ \Sigma_{j=1}a_{ij}e_j \\ \vdots \\ u_n \end{bmatrix})$ and by linearity we have $\Sigma_{j=1}^n a_{ij}\det(\begin{bmatrix} u_1 \\ \vdots \\ e_j \\ \vdots \\ u_n \end{bmatrix})$

Where here I can apply the special case: $\Sigma_{j=1}^n a_{ij}(-1)^{i+j}\det(\tilde{A}_{ij})$ this is the same for A and i $\begin{bmatrix} u_1 \\ \vdots \\ e_j \\ \vdots \\ u_n \end{bmatrix}$

😊

Note:-

Let $A \in M_n(F), B \in M_n(F)$ is obtained by flipping 2 rows of A, then $\det(B) = -\det(A)$

.

Proof: Let $A = \begin{bmatrix} u_1 \\ \vdots \\ u_i \\ u_j \\ \vdots \\ u_n \end{bmatrix}$ and $B = \begin{bmatrix} u_1 \\ \vdots \\ u_j \\ u_i \\ \vdots \\ u_n \end{bmatrix}$. Let $C = \begin{bmatrix} u_1 \\ \vdots \\ u_i + u_j \\ u_i + u_j \\ \vdots \\ u_n \end{bmatrix}$. Theorem 2 tells us that $\det(C) = 0$ Using linearity we can write that $0 = \det(C) = \det(A) + \det(B)$. Note that we cancelled out two copies of rows having the same two elements in either case.

😄

12.3.3 How row operations affect determinant

1. Add $r \cdot j^{th}$ row to i^{th} row of A
   $B = E_{ij}(r) \cdot A$, claim that this operation doesn't affect the determinant.

   Proof: $A = \begin{bmatrix} u_1 \\ \vdots \\ u_i \\ \vdots \\ u_j \\ \vdots \\ u_n \end{bmatrix}, B = \begin{bmatrix} u_1 \\ \vdots \\ u_i + ru_j \\ \vdots \\ u_j \\ \vdots \\ u_n \end{bmatrix}$

   By linearity we can decompose this into $\det(\begin{bmatrix} u_1 \\ \vdots \\ u_i \\ \vdots \\ u_n \end{bmatrix})$

   😄

2. Flip i, j rows of A. $B = F_{ij}A \longrightarrow \det(B) = -\det(A)$ By the corrolary.
3. Multiply a row by a constant,

12.4 Tuesday October 22

Note:-

Finish Determinants today; need some stuff for the HW.

12.4.1 Reminders

Note:-

$A = (a_{ij}) \in M_n(F)$

Note:-

$\det(A) = \Sigma_{j=1}^{m} (-1)^{1+j} a_{1j} \tilde{A}_{1j}$
Or any other row

Note:-

det : $M_n(F) \to F$ is linear on each row if all other rows are held fixed.

Note:-

$\det(A) = 0$ if A has 2 identical rows

Note:-

$\det(I_n) = 1$

Note:-

1,2 above imply that $\det(E_{ij}(r)A) = \det(A) \to$ first row operation doesn't change determinant.

Note:-

$\det(F_{ij}A) = -\det(A)$

Note:-

$\det(C(r)A) = r\det(A)$, where $C(r) = \begin{bmatrix} 1 & \dots & 0 \\ 0 & r & 0 \\ 0 & \dots & 0 \end{bmatrix}$

Theorem 12.4.1

$A \in M_n(F)$ is invertible $\leftrightarrow \det(A) \neq 0$.

Theorem 12.4.2

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

We will prove the first direction of theorem 1. Then Proving 2 gives us the second half of Theorem 2.

Proof: If A is not invertible $\to \det(A) = 0$. This is the contrapositive of Theorem 2.
A not invertible means that $L_A : F^n \to F^n$ is not an isomorphism. Because the dimensions are the same, this means that L_A is not surjective. In other words $rk(A) < n$. This implies that the rows of A are linearly dependent.

$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ where $u_i = i^{th}$ row of A. This implies that $\exists i \in \{1, \dots, n\}$ such that u_i is a linear combination of the others.
Without loss of generality assume that $i = 1$. This is okay because we can apply flips if necessary which changes the sign of the determinant, but we are shoiwng that the determinant is 0 so this doesn't make any difference.

$u_1 = c_2u_2 + \dots + c_nu_n$ for some $c_2, \dots, c_n \in F$.

$$\det(\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}) = \det(\begin{bmatrix} c_2u_2 + \dots + c_nu_n \\ u_2 \\ \vdots \\ u_n \end{bmatrix}).$$

Through row operations we get that (adding negative copies of c_i from each row):

$$\det(A) = \det(E_{12}(-C_2)E_{13}(-C_3)\dots E_{1n}(-C_n)A)$$

But if the first row is zero then we are done.

This concludes the first part.

😄

Definition 12.4.1

A matrix $E \in M_n(F)$ is called elementary if $E = E_{ij}(r)$ or $E = F_{ij}$ or $E = C_i(r)$

Note:-

Below we prove $\det(AB) = \det(A)\det(B)$

Proof: Compute $\det(E)$, E elementary. $E = E_{ij}(r)$ then we know that $\det(E_{ij}(r)I_n)$ is equal to the determinant of $\det(E_{ij}(r))$ but this is also equal to the determinant of the identity because $E_{ij}(r)$ does not change the determinant. This gives us that $\det(E_{ij}(r)) = 1$. Similarly we get $\det(F_{ij}) = -1$ and $E = C_i(r) = \det(C_i(r)) = r$ We want to prove that $\det(AB) = \det(A)\det(B)$. First note that $\det(EA) = \det(E)\det(A)$ The proof is similar and uses the computation of $\det(E)$.
Next Key proposition. Let $A \in M_n(F)$ be invertible. Then A is a product of of elementary matrices. Meaning $A = E_1 ... E_l$ for some $E_1, ..., E_l$ elementary.

Note:-

First perform row operations to bring A to row echelon form. Because row operations are given by left multiplication of elementary matrices, this means $\exists P_1, ..., P_r$ elementary matrices such that $P_1 ... P_r A$ is in row echelon form. Now we perform column operations. These are right multiplications by elementary matrices. To bring $P_1 ... PrA$ to $\begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$ This means that $\exists Q_1, ..., Q_m$ elementary such that $P_1 ... P_r A Q_1 ... Q_m = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$, because A is invertible, the $rk(A) = n$. This means that doing these operations gives us the identity matrix.

Proof of Theorem 2. $\det(AB) = \det(A)\det(B)$ Case 1. Suppose that A is not invertible. Then by the first proposition we proved, we get that $\det(A) = 0$. This means that the right hand side is 0. Thus we want to show that $\det(AB) = 0$. It is enough to show that AB is not invertible.
Suppose for contradiction that AB is invertible. This implies that the $rk(AB) = n$ which implies that L_{AB} is an isomorphism. Recall that $L_{AB} = L_A \circ L_B$. This is implies that L_A is surjective which implies that $rk(A) = n \longrightarrow A$ is invertible which is a contradiction.
Case 2: Suppose A is invertible. We get that $A = E_1, ..., E_l$ is a product of elementary matrices. Now we compute each one. $\det(A) = \det(E_1 ... E_l) \longrightarrow \det(E_1(E_2 ... E_n))$ By the special case: $\det(E_1)\det(E_2 ... E_n)$. Continue from here inductively to get $\det(E_1) ... \det(E_l)$.
We play the same game for B
Proposition Let $A \in M_n(F)$ If A invertible then $\det(A) \neq 0$ and $\det(A^{-1}) = \frac{1}{\det(A)}$
Then finally $AA^{-1} = I_n \longrightarrow \det(AA^{-1}) = 1$ by Theorem 2. $\det(A)\det(A^{-1}) = 1 \longrightarrow \det(A) \neq 0$ and $\det(A^{-1}) = \frac{1}{\det(A)}$.

☺

Note:-

Let $A \in M_n(F)$ hten $\det(A) = \det(A^t)$.

Proof: Case 1 suppose that $\det(A) = 0 \longrightarrow rk(A) < n$ and $rk(A^t) < n$. This implies that A^t is not invertible which implies that $\det(A^t) = 0$.
Case 2: Suppose that $\det(A) \neq 0 \longrightarrow A = E_1 ... E_r$ product of determinant matrices $\longrightarrow A^t = E_r^t ... E_1^T$. As an easy verification $\det(A^t) = \det(E_r^t) ... \det(E_1^t)$. Easy computation $\det(E) = \det(E^T)$.

☺

Note:-

$\det(A) = \det(A^t)$ would follow if we could compute $\det(A)$ using expansions along the columns instead of rows.

Theorem 12.4.3

Suppose we are given a function $\delta : M_n(F) \longrightarrow F$ function that satisfies the following:

1. δ is linear on each row if all other rows are held fixed.
2. $\delta(A) = 0$ if A has 2 identical rows
3. $\delta(I_n) = 1$.

Then we can conclude that $\delta(A) = \det(A)$. Mean det is the unique funciton satisfying 1-3 above.

1. Show $\delta(E) = \det(E)$ for all E elementary.
   This follows similarly to det: $\delta(E_{ij}(r)A) = \delta(A)$ etc.

2. Show $\delta(EA) = \delta(E)\delta(A)$ We can do that by computing $\delta(E)$.

3. Show $\delta(A) = \det(A)$. Case 1: If $rk(A) < n \to RHS = 0 \to \delta(A) = 0$ by the same proof. Case 2: If $rk(A) = n \to A = E_1 ... E_l$. Proceeding as we did before.

Example 12.4.1 (Application: Another description of the Determinant)

Preparation: From 3354. For $n \geqslant 1$ we have $S_n = \{\sigma : 1, ..., n\} \to \{1, 2, ..., n\}$ The set of all 1-1 functions from zero to itself. We learn in survey of Algebra that this thing becomes a group with multiplication given by composition. Meaning $\sigma\tau = \sigma \circ \tau$

Theorem 12.4.4 Leibniz Formula for det

Let $A = (a_{ij}) \in M_n(F)$ then $\det(A) = \Sigma_{\sigma \in S_n} \operatorname{sign}(\sigma) \pi_{i=1}^{n} a_{i\sigma(i)}$

Note:-

This idea follows relatively easy by uniqueness.

Example 12.4.2 (Application of Determinant)

Cramer's Rule. Solving Linear Systems.

Theorem 12.4.5 Cramer's Rule

Suppose we are given a linear System. $Ax = b$ a linear system with $A \in M_n(F)$ meaning the system is in n equations and n variables. Suppose that A is invertible, or that $\det(A) \neq 0$. Then the system has unique solution $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ with the following formula $x_k = \frac{\det(M_k)}{\det(A)}$ where M_k is the matrix obtained from A by replacing its k_{th} column by B.

12.5 Chapter 5 Diagonalization

Definition 12.5.1

Let V be an F vector space. A linear transformation T from $V \to V$ is called diagonalizable if $\exists B = \{v_1, ..., v_n\}$ ordered basis for B such that $[T]_B$ is diagonal $\begin{bmatrix} d_1 & ... & 0 & \\ 0 & d_2 & 0 & ... \\ d_n & & & \end{bmatrix}$

Example 12.5.1

$T : R^2 \to R^2$ The reflection along $y = 2x$. $B = \{(1, 2), (-2, 1)\}$, with $[T]_B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ with this particular example T is diagonalizable.

Definition 12.5.2

A matrix $A \in M_n(F)$ is called diagonalizable if $L_a : F^n \to F^n$ is diagonalizable as a linear transformation.

Note:-

Recall $[L_a]_{B_{st}} = A$ of the standard Basis of F^n. A is diagonalizable $\leftrightarrow \exists B$ ordered basis of F^n such that $[L_a]_B$ is diagonal.

Let P be the change of basis matrix from B to B_{st}. This is always the easy one to form. Then the i^{th} column of P is v_i.

Note:-

Recall that $A = [L_a]_{B_{st}} = [I_{F^n}]_B^{B_{st}}[L_a]_B[I_{F^n}]_{B_{st}}^B$

Note:-

If $A \in M_n(F)$ is diagonalizable then $A = P^{-1}DP$ for some D, diagonal P invertible.

Note:-

A is diagonalizable $\leftrightarrow$ A is similar to a diagonal matrix

12.5.1 Why Diagonalizability is a nice property

Note:-

If A is diagonalizable, then we can compute its powers very easily.

$$\text{If } A = PDP^{-1} \text{ then } A^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}, \text{ more generally, } A^n = PD^nP^{-1} = P \begin{bmatrix} d_1^m \\ \vdots \\ d_n^m \end{bmatrix} P^{-1}$$

Example 12.5.2 (Non-Example)

Fix $\theta \in (0, \pi)$. Let $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be the counterclockwise rotation by θ. Claim is that T_θ is not diagonalizable.

Suppose $T : V \to V$ linear. Suppose T is diagonalizable. By definition this means $\exists B = \{v_1, ..., v_n\}$ ordered basis of V such that $[T]_B = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}$ This means that $T(v_1) = \lambda_1 v_1 + 0v_2 + ... + 0v_n$ and so on. This implies that transforming a non zero vector takes it parallel to itself in terms of the basis.

This shows that T_θ is not diagonalizable because for every $v \neq 0, v \in \mathbb{R}^2$. Accordingly, $T_\theta(v), v$ are not parallel.

Note:-

The rest of the class will be to find concrete criterai for diagonalizability. What if we remove criteria? Plan for the rest of class.

Note:-

Suppose $\exists B = \{v_1, ..., v_n\}$ ordered basis of V such that $T(v_i) = \lambda_i v_i$ for some $\lambda_i \in F$ for each $i = 1, ..., n$.

Note:-

Below is shorthand for diagonal matrix

This implies that $[T]_B = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}$

Definition 12.5.3: EigenVector

Let $T : V \longrightarrow V$ linear. A nonzero vector $v \in V$ is called an eigenvector of T if $\exists \lambda \in F$ such that $T(v) = \lambda v$.

Definition 12.5.4: Eigen Value

The scalar λ is called an eigen value of T corresponding to the eigen vector v.

Definition 12.5.5: EigenVector Matrix

Let $A \in M_n(F)$ A nonzero $v \in F^n$ is called an eigenvector of A if $Av = \lambda v$ for some $\lambda \in F$.

Note:-

$T : V \longrightarrow V$ is diagonalizable $\leftrightarrow$ V has a basis B consisting of eigenvectors.

$A \in M_n(F)$ is diagonalizable $\leftrightarrow \exists B$ basis of F^n consisting of eigenvectors for A.

Note:-

If I have 1 eigenvector, I have infinitely many. If T has 1 eigenvector $v \neq 0$ then any scalar multiple of it is also an eigenvector.

T is diagonalizable $\leftrightarrow$ we can find n linearly independent eigenvectors where n is the dimension of V.

Note:-

The notion of diagonalizability is Tied to finite dimensional vector spaces, but the notion of eigen value or eigen vectors is not.

Example 12.5.3

Let $V = L^\infty(\mathbb{R}) = \{f : \mathbb{R} \longrightarrow \mathbb{R}$ infinitely differentiable functions$\}$, Take $T : V \longrightarrow V$ and $f \longrightarrow f'$. Let $\lambda \in \mathbb{R}$ be an arbitrary scalar with $f(t) = e^{\lambda t}$ is an eigen vector with a corresponding eigenvalue λ.

Note:-

If f is an eigenvector with eigenvalue λ, then $f'(t) = \lambda f(t)$. This is a first order linear ODE: Solving for $f(t) = ce^{\lambda t}, c \in \mathbb{R}$.

12.5.2 How to find eigenvectors / eigenvalues

For a matrix $A \in M_n(F)$: Suppose $\lambda \in F$ is an eigenvalue for A and $v \in F^n, v \neq 0$ a corresponding eigenvector. By definition this means $Av = \lambda v \longrightarrow (A - \lambda I_n)v = 0$. Observe that $(A - \lambda I_n) \in M_n(F)$ this implies that for B the linear system, $Bv = 0$ has nonzero solutions. This implies that B is not invertible: $A - \lambda I_n$ is not invertible. This implies that $\det(A - \lambda I) = 0$. Thus if $\lambda \in F$ is an eigenvalue of $A \longrightarrow \det(A - \lambda I) = 0$.

Note:-

Converse. Let $\lambda \in F$ such that $\det(A - \lambda I) = 0 \longrightarrow A - \lambda I$ is not invertible, this not invertible implies that the system $(A - \lambda I)x = 0$ has at least one nonzero solution, $v \neq 0$. Justify $A - \lambda I$ not invertible implies that $L_{A-\lambda I} : F^n \longrightarrow F^n$ is not an isomorphism. Because the dimensions are the same we know that the $ker(L_{A-\lambda I} \neq 0$. This implies $\exists v \neq 0, v \in ker(L_{A-\lambda I})$.

$\lambda \in F$ is an eigenvalue of $A \leftrightarrow \det(A - \lambda I) = 0$.

Definition 12.5.6

Let $A \in M_n(F)$ the characteristic polynomial of A is $P_A(t) = \det(A - tI)$.

Note:-

$\lambda \in F$ is an eigenvalue of A if and only if $P_A(\lambda) = 0$ if and only if λ is a root of $P_a(t)$

Example 12.5.4

$A = \begin{bmatrix} 1 & 4 \\ 4 & 1 \end{bmatrix}$ $P_a(T) = |\begin{bmatrix} 1-t & 1 \\ 4 & 1-t \end{bmatrix}| = (1-t)^2 - 4 = (-t-1)(3-t)$ this gives us A has 2 eigenvalues $\lambda_1 = -1, \lambda_2 = 3$. Find Eigen Vectors:

For $\lambda_1 = -1$: Look for vectors $\begin{bmatrix} x \\ y \end{bmatrix} \in F^2$ such that $A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix} \leftrightarrow \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix}$.

$x + y = -x$ and $4x + y = -y$. Toegether this is $y = -2x$. Let $x = 1, y = -2$, then $v_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ is an eigenvector corresponding to λ_1.

For $\lambda_2 = 3$ look for $v = \begin{bmatrix} x \\ y \end{bmatrix}$ such that $Av = 3v \leftrightarrow \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ 3y \end{bmatrix}$. Solving this system gives us:

$x + y = 3x$ and $4x + y = 3y$ together we have $y = 2x$: Take $v_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ to be eigenvectors corresponding to λ_2.

Now we have two eigenvectors $B = \{\begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}\}$ which means that this is a basis because it has the correct amount of vectors and the two vectors are clearly independent. Since we found a basis consisting of eigen vectors we can conclude that A is diagonalizable. This diagonal matrix is $A = P \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} P^{-1}$. Then $A \sim \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$ $[L_A]_B = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$.

Example 12.5.5

$T_\theta : \mathbb{R}^2 \to \mathbb{R}^2 \to A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, $\theta \in (0, \pi)$, $P_A(t) = |\begin{bmatrix} \cos\theta - t & -\sin\theta \\ \sin\theta & \cos\theta - t \end{bmatrix}| = t^2 - 2\cos\theta t + 1$ then $b^2 - 4ac = 4(\cos^2\theta - 1)$ if $\theta \in (0, \pi)$. Because $4(\cos^2\theta - 1) < 0$ we cannot factor it in the real numbers. But if we consider $A_\theta \in M_2(\mathbb{C})$ T becomes diagonalizable.

Example 12.5.6

$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ not diagonalizable. Also $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ doesn't produce enough eigen values. If $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ were diagonalizable it would be similar to the identity. But everything that is similar to the identity is the identity itself. But this matrix is not the identity.

12.6 Tuesday October 29th

12.6.1 Reminders

1. $A \in M_n(F)$ is diagonalizable $\leftrightarrow$ A is similar to a diagonal Matrix $\leftrightarrow \exists P \in M_n(F)$ invertible and $D \begin{bmatrix} \lambda_1 & \cdots & 0 \\ 0 & \cdots & 0 \\ 0 & \cdots & y_n \end{bmatrix}$ such that $A = PDP^{-1} \leftrightarrow F^n$ has a basis $\{v_1, ..., v_n\}$ consisting of eigen vectors for $A = Av_i\lambda_i v_i$
2. $\lambda \in F$ is an eigenvalue of $A \leftrightarrow \exists v \neq 0$ vector such that $Av = \lambda v$ similarly for linear $T : V \to V$, $T(v) = \lambda v$
3. $\lambda \in F$ is an eigenvalue of $A \in M_n(f) \leftrightarrow \det(A - \lambda I) = 0 \leftrightarrow \lambda$ is a root of $P_A(t) = \det(A - tI) =$ characteristic polynomial of A

Let $T : V \longrightarrow V$ linear, we want to define $P_T(t)$ = characteristic polynomial of T. Which matrix representation are we looking for?

Definition 12.6.1

Let $T : V \longrightarrow V$ be linear. The characteristic polynomial of T is $P_T(t) = \det([T]_B - tI)$ where $B = \{v_1, \dots, v_n\}$ is some fixed ordered basis for V.

Question 8

We need to show that this is well defined. Meaning, it is independent of the choice of ordered basis B.

Note:-

Let $A, B \in M_n(F)$ be similar. Then $P_A(t) = P_B(t)$.

Proof: A, B similar $\longrightarrow \exists P \in M_n(F)$ invertible such that $A = PBP^{-1}$. We need to compute $A - tI$, we take the determinant of this guy to get the characteristic polynomial, $A - tI = PBP^{-1} - tI =$
$PBP^{-1} - PtIP^{-1}$, because tI commutes with everything.
$P(B - tI)P^{-1}$.
We can conclude that if A, B are similar then $A - tI, B - tI$ are also similar. In the HW we proved that they also have the same determinant, that $\det(A - tI) = \det(B - tI)$ which are exactly the characteristic of $P_A(t), P_B(t)$, respectively.

😊

Note:-

The characteristic polynomial $P_T(t)$ of a linear transformation $T : V \longrightarrow V$ is independent of choice.

Proof: If B_1, B_2 are 2 ordered bases for $V \longrightarrow [T]_{B_1}, [T]_{B_2}$ are similar. By previous lectures we get that $\longrightarrow \det([T]_{B_1} - tI) = \det([T]_{B_2} - tI)$

😊

12.6.2 Summary of similalr matrices

If A, B are similar then A, B have the same

1. Trace
2. det
3. Characteristic Polynomial.

Note:-

However none of these go backwards. We will see a quick example as to why not.

Example 12.6.1

$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, Trace A = Trace $B = 2$, $\det(A) = \det(B) = 1$, $P_A(t) = (1 - t^2) = P_B(t)$ But $A \nsim B$.

Suppose $A \sim B$, then $\exists P$ invertible such that $B = P \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} P^{-1}$ but this of course equals I so it must be the case that this is not true.

Note:-

If someone gives two diagnoalizable matrices, then the two matrices are similar if and only if they are similar to the same diagonal matrix.

Example 12.6.2

$T : P_2(\mathbb{R}) \longrightarrow P_2(\mathbb{R})$, $f(x) \longrightarrow f(x) + (x+1)f'(x)$.

Question 9: Is T diagonalizable

Solution: Reminder this means: Does there exist a basis B of $P_2(\mathbb{R})$ such that $[T]_B$ is diagonal.

Let $B_{st} = \{1, x, x^2\}$. $T(1) = 1, T(x) = x + x + 1, T(x^2) + 3x^2 + 2x$

$A = [T]_{B_{st}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$. Now we explore if A is diagonalizable.

Step 1: Find Eigenvalues: $P_A(t) = \begin{bmatrix} 1-t & 1 & 0 \\ 0 & 2-t & 2 \\ 0 & 0 & 3-t \end{bmatrix}$

Which equals $(1-t)(2-t)(3-t)$, where $\lambda_1 = 1, \lambda_2 = 2\lambda_3 = 3$.

Compute EigenVectors:

$\lambda_1 = 1$ Look for $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$ such that $A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$\lambda_2 = 2$ look for $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$ such that $A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2 \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

$\lambda_3 = 3$ look for $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$ such that $A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 3 \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

$v_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

Conclusion so far: The vectors $\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\}$ form a basis for $\mathbb{R}^3$. A is diagonalizable. $A = P \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix})P^{-1}$, where $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$, Give a basis for $P_2(\mathbb{R})$ consisting of eigenvectors for T.

$P_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$
$1 \longrightarrow e_1$
$x \longrightarrow e_2$
$x^2 \longrightarrow e_3$

Transfer B_0 to a basis B of $P_2(\mathbb{R})$. $B = \{1, 1+x, 1+2x+x^2\}$ is a basis consisting of eigenvectors for T.

12.6.3 Important Observations

1. $A = (a_{ij}) \in M_n(F)$, then $P_A(t) = \begin{bmatrix} a_{11}-t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-t & \cdots & a_{2n} \\ \vdots & \cdots & \cdots & \cdots \\ a_{n_1} & a_{n_2} & \cdots & a_{nn}-t \end{bmatrix}$

2. $P_A(t)$ is a polynomial of degree exactly n (Do an induction proof.)

3. The leading term of $P_A(t)$ is $(-1)^n$ and the constant term is $\det(A)$

1. $\rightarrow P_A(t)$ has at most n roots implies that A has at most n eigenvalues.

2. 2). Suppose $A \in M_n(\mathbb{C})$ by Fundamental Theorem of Algebra we get that $P_A(t)$ has at least 1 root. This implies A has at least 1 eigen value and a linearly independent eigenvector. This is not always true over the reals.

3. Let $\lambda \in F$ be an eigenvalue of A. Set $E_\lambda = \{v \in F^n$ such that $Av = \lambda v\}$we call this the eigenspace of A corresponding to lambda.

   Claim is that E_λ is a subspace of F^n

4. Let $T : V \longrightarrow V$ linear, $\lambda \in F$ an eigenvalue then E_λ is a T invariant subspace.

Definition 12.6.2

If we are given $T : V \longrightarrow V$, a subspace $W \leqslant V$ is called T invariant if $T(w) \subseteq W$.

Proof: $E_\lambda = \ker(L_{A-\lambda I}$, similarly for $T : V \longrightarrow V$ linear.

😊

Theorem 12.6.1

Let $T : V \longrightarrow V$ linear transformation. Let $\lambda_1, ..., \lambda_k$ be distinct eigen values of T. For each $i \in \{1, ..., k\}$, let $S_i \subseteq E_{\lambda_i}$ be a linearly independent subset. Then if $S = S_1 \cup S_2 \cup ... \cup S_k$ is linearly independent.

Note:-

The S_i's are pairwise disjoint.

Proof: Let $v \in S_i \cap S_j$ for $i \neq j$. Then $v \in S_i \subseteq E_{\lambda i} \longrightarrow T(v) = \lambda i v$
$v \in S_j \subseteq E_{lambda_j} \longrightarrow T(v) = \lambda_j v$ together these to imply that $\lambda_i(v) = \lambda_j(v) = (\lambda_i - \lambda_j)v = 0$. This implies that $v = 0$.
This shows that $E_{\lambda i} \cap E_{\lambda i} = \{0\}$ since S_i, S are independent and $v \neq 0, S_i \cap S_j \neq 0$.

😊

Proof: By induction on $k \geqslant 1$. Base case $K = 1$ then $S = S_1$. S_1 is assumed to be independent.
Induction Hypothesis: Suppose the statement is true for $k-1, k \geqslant 2$. We want to show that $S = S_1 \cup ... \cup S_k$ is linearly independent.
For each $i = 1, ..., k$ write $S_i = \{v_{i1}, v_{i2}, \cdots, v_{in}\}$ linearly independent subset of $E_{\lambda i}$. We want to show that $S = S_1 U \cdots U S_k$ is linearly independent. Note that by the induciton hypothesis, the set $S_1 \cup S_2 \cup \cdots \cup S_{k-1}$ is linearly independent. Let a_{ij} be scalars such that $\Sigma_{i=1}^k \Sigma_{j=1}^{n_i} a_{ij} v_{ij} = 0$. This is the linear combination of what is above.
$0 = \Sigma_{i=1}^k \Sigma j = 1_i^n a_{ij} v_{ij} =$
$0 = \Sigma_{i=1}^k - 1\Sigma_{j=1}^{n_i} a_{ij} v_{ij} + \Sigma_{j=1}^{n_k} a_{kj} v_{kj}$ The right hand side is in $E_{\lambda k}$, applying the linear transformation. $T - \lambda_k I$ to the right hand side in $E_{\lambda k}$.
$(T - \lambda_k I)(0) = (T - \lambda_k I)(RHS)$
Which equals: $(T - \lambda_k I)(\Sigma_{i=1}^{k-1} \Sigma_{j=1}^{n_i} a_{ij} v_{ij}) + (T - \lambda_k I)$

$$(T_{\lambda_k} I)(\Sigma_{i=1}^{k-1}\Sigma_{j=1}^{n_i} a_i v_{ij} = 0$$
$$\Sigma_{i=1}^{k-1}\Sigma_{j=1}^{n_i} a_{ij} T(v_{ij}) - \Sigma_{i=1}^{k-1}\Sigma_{j=1}^{n_i} a_{ij}\lambda_k v_{ij} = 0$$
$$\Sigma_{i=1}^{k-1}\Sigma_{j=1}^{n_i} a_{ij}(\lambda_i - \lambda_k)v_{ij} = 0.$$

Linear combinations of vectors in $S_1 \cup S_2 \cup \ldots \cup S_{k-1}$ are independent implies that $a_{ij}(\lambda_i - \lambda_k) =$ and so $\forall 1 \leqslant i \leqslant k-1$, $1 \leqslant j \leqslant n_i$ because eigen values are distinct, $\lambda_i \neq \lambda_k$ for all $i < k$. $a_{ij} = 0$ for all $1 \leqslant i \leqslant k-1$, $1 \leqslant j \leqslant n_i$

Put this back to 1 gives us $\Sigma_{j=1}^{n_k} a_{jk} v_{kj} = 0$ $S_k = \{v_{k1}, \ldots, v_{k_n}\}$ is linearly indepdnent by assumption.

☺

12.7 Thursday October 31st

Note:-

Midterm open on Gradescope next Friday morning. The gradescope submission will close saturday afternoon.

Note:-

Material: up to today.

12.7.1 Reminders

1. $T : V \to V$ linear. $\lambda \in F$, an eigen value of T.

2. $E_\lambda = \{v \in V \text{ such that } T(v) = \lambda v\}$ = Eigen space of T corresponding to λ

3. Theorem From Last Time:

   Theorem 12.7.1

   Let $T : V \to V$ linear $\lambda_1, \ldots, \lambda_n$ distinct eigen values of T for each $i = 1, \ldots, k$ Let S_i a linear independent subset of $E_{\lambda i}$ then $S = S_1 \cup S_2 \cup \ldots \cup S_n$ is linearly independent.

Note:-

If $\lambda_1, \ldots, \lambda_k$ are all the distinct eigenvalues of T. And B_i = a basis for E_{λ_i} then $S = B_1 \cup \ldots \cup B_k$ is linearly independent.

Note:-

If $\dim V = n$ and T has n distinct eigeenvalues $\longrightarrow T$ is diagonalizable.

Proof: For each λ_i the dimension of $\dim E_{\lambda_i} \geqslant 1$. This is because λ_i is an eigen value. Let B_i = a basis of E_{λ_i}. Then $B = B_1 \cup B_2 \cup \ldots \cup B_n$ is a linearly independent subset of V, with at least n elements.
Since $\dim V = n \longrightarrow B$ has exactly n elements and B is a basis of V, which implies by definition that T is diagonalizable.

☺

Note:-

Be careful. The converse is not true.

Question 10: T diagonalizable does not imply T has n distinct eigen values

Solution: Counter example for the converse. Take $n \geqslant 2$ and I_n. By definition I_n is diagonal. In this case B_{st} is a basis of eigen vectors.

Note:-

$W_1, W_2 \leqslant V, W_1 + W_2$ is direct $\leftrightarrow W_1 \cap W_2 = 0$

Definition 12.7.1

Let $W_1, ..., W_k \leqslant V$ we say that $W_1 + W_2 + ... + W_k$ is direct if for each $i \in \{1, ..., k\}$, if we take $W_i \cap (\Sigma_{j \neq i} W_j) = \{0\}$

Let $W_1, ..., W_k \leqslant V$ for each $i \in \{1, ..., k\}$ let B_i be a basis of W_i. If: $B_i \cap B_j = \emptyset \forall i \neq j$ and $S = B_1 \cup B_2 \cup ... \cup B_k$ is linearly independent, then the sum is direct: $W_1 \oplus W_2$.

Proof: We want to show $\forall i \in \{1, ..., k\}, W_i \cap (\Sigma_{j \neq i} W_j) = \{0\}$.
Write $B_i = \{v_{i1}, v_{i2}, ..., v_{ini}\}$ for each $i = 1, ..., k$.
Let $v \in W_i \cap (\Sigma_{j \neq i} W_j)$.
$v \in W_i \rightarrow v = a_{i1} v_i + ... + a_{ini} v_{ini}$, for some scalars $a_{i1}, ..., a_{ini} \in F$.
$v \in \Sigma_{j \neq i} W_j \rightarrow$we can write $v = W_1 + ... + W_{i-1} + W_{2i+1} + ... + W_k$. With $w_j \in W_j$.

$$W_1 = a_{11} v_{11} + ... + a_{1n1} v_{1n1}$$
$$\vdots$$
$$W_k = a_{k1} v_{k1} + \cdots + a_{knk} v_{knk} \;\; v = \Sigma_{j \neq 1} \Sigma_{l=1}^{n_j} a_{jl} v_{jl}$$

Becuase we can subtract two equations above to get zero, we get that linear combination of the vectors in $B_1 \cup ... \cup B_k = 0$ are linearly independent. Since $B_1 \cup ... \cup B_k$ are linaerly independent scalars $= 0 \rightarrow v = 0$. As an exercise prove the converse.

Note:-

Converse of this is true.

😊

Note:-

Let $T : V \rightarrow V$ linear, $\lambda_1, ..., \lambda_k$ be distinct eigenvalues of T. Then $E_{\lambda_1} \oplus E_{\lambda_2} \oplus ... \oplus E_{\lambda_k}$ is direct.

Conclusion: $T : V \rightarrow V$ is diagonalizable $\leftrightarrow \exists$ basis B of V consisting of eigenvectors of $T \leftrightarrow V = E_{\lambda_i} \oplus ... \oplus E_{\lambda_k}$. Note that each eigen space is a T invariant subsapce of V.

Note:-

1. If $A \in M_n(F)$ has n distinct eigenvalues, then immediately A is diagonalizable. Meaning if $P_A(t) = (\lambda_1 - t)...(\lambda_n - t)$ with $\lambda_i \in F$ pairwise distinct, then A is diagonalizable.
2. Suppose $P_A(t) = (\lambda - t)^n$. Then A diagonalizable if and only if A is diagonal.

Proof: Since λ is the only eigenvalue, if A is diagonalizable then $A \sim \begin{bmatrix} \lambda & \cdots & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \lambda I$
If $A = P \lambda I \cdot P^{-1} \rightarrow$
$A = \lambda I$ must be diagonal itself.

😊

Definition 12.7.2

A polynomial $f(t) \in F[t]$ splits completely over F if $f(t)$ can be factored into linear terms in $F[t]$. Meaning $f(t) = c(t - a_1)...(t - a_n)$ with a_i not necessarily distinct.

Note:-

Remark towards verifying diagonalizable

Let $A \in M_n(F)$ if $P_A(t)$ does not split completely over $F \rightarrow A$ not diagonalizable.

Example 12.7.1

$A \in M_4(\mathbb{R})$ with $P_A(t) = (1-t)(5-t)(t^2+2)$.

$P_A(t)$ does not split completely over $\mathbb{R} \longrightarrow A$ is not diagonalizable in $M_4(\mathbb{R})$. But if we consider $A \in M_4(\mathbb{C}) \longrightarrow A$ is diagonalizable because $P_A(t) = (1-t)(5-t)(t+\sqrt{2}i)(t-\sqrt{2}i)$, with $\lambda_1 = 1, \lambda_2 = 5, \lambda_3 = \sqrt{2}i, \lambda_4 = -\sqrt{2}i$
P_A has 4 distinct roots in $\mathbb{C} \longrightarrow A$ is diagonalizable in $M_4(\mathbb{C})$.

Proof: We will prove the contrapositive. A diagonalizable $\longrightarrow P_A(t)$ splits completely in $F[t]$.
A diagonalizable means $\exists P$ invertible and D diagonal matrix D with lambdas along the diagonal such that
$A = PDP^{-1} \longrightarrow A - tI = P(T-tI)P^{-1}$
Which means that $A - tI$ and $D - tI$ are similar and have the same determinant.
$P_A(t) = \det(A-tI) = \det(D-tI)$ but $D - tI$ is diagonalizable so we know that it equals:
$(\lambda_1 - t)...(\lambda_n - t)$.

☺

Note:-
Outside of these three cases we have to compute eigen spaces.

Note:-
If $P_A(t)$ splits over F and has repeated roots, unless we are in case 2, we have to compute the eigen spaces.

From now on we will assume that $A \in M_n(\mathbb{C})$, everything splits completely, $\longrightarrow P_A(t)$ always splits.

Suppose $P_A(t) = (\lambda_1 - t)_1^k...(\lambda_n - t)^{k_r}$ with $k_i \geqslant 1$. $\lambda_1, ..., \lambda_i$ the distinct eigen values of A.

Definition 12.7.3

Let λ be an eigen value of A. The algebraic multiplicity of λ is the largest positive integer k such that $(\lambda - t)^k$ is a factor of $P_A(t)$.

The geometric multiplicity of $\lambda = \dim E_\lambda$.

Question 11: Do we have a guess for how these are related

Solution: The geometric multiplicity is at most the algebraic multiplicity.

Theorem 12.7.2

$T : V \longrightarrow V$. Let $m = \dim E_\lambda$ be the geometric multiplicity of λ = an eigen value of T. Let k = its algebraic multiplicity then $1 \leqslant m \leqslant k$.

Proof: $m \geqslant 1$ is clear because λ eigenvalues $\longrightarrow \exists v \neq 0$ eigenvector $\longrightarrow \dim E_\lambda \geqslant 1$.
Let $\{v_1, ..., v_n\}$ be a basis for E_λ extend it to a basis of V. $B = \{v_1, ..., v_m, v_{m+1}, ..., v_n\} \longrightarrow [T]_B = \begin{bmatrix} diag & * \\ 0 & B \end{bmatrix}$.
This is enough because $P_A(t) = \det(A - tI) = \det(\begin{bmatrix} \lambda - t & 0 \\ 0 & \lambda - t \end{bmatrix} \cdot \det(B - tI)$

☺

12.8 Thursday November 7th

Definition 12.8.1: Diagonalizability

$T : V \to V$ linear. T is Diagonalizabile if and only if $\exists B = \{v_1, .., v_n\}$ ordered basis consistin of eigenvectors for T. $\leftrightarrow \exists B$ ordered basis such that $[T]_B$ = a diagonal matrix of eigen values.

$\leftrightarrow V = E_{\lambda i} \oplus ... \oplus E_{\lambda r}$ where $\lambda_1, ..., \lambda_r$ are the distinct eigen values of T.

Note:-

In general $E_{\lambda i} \oplus ... \oplus E_{\lambda r}$ is always direct.

Note:-

T is diagonalizabile $\longrightarrow P_T(t)$ splits completely in $F[t]$.

Note:-

Caution: The converse is not true.

Counter example for the converse. $A = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$ Two distinct eigen values. Consider the Characteristic Polynomial. $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$: $P_A(t) = (1-t)^2$. If A diagonalizable then $A \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ because $\lambda = 1$ is the only eigenvalue of A. But $P \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} P^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $A \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Note:-

From now on assume that $F = \mathbb{C}$ so every characteristic polynomial splits.

Definition 12.8.2: Algebraic Multiplicity

$P_A(t) = (\lambda_1 - t)^{k_1} \cdots (\lambda_r - t)^{k_r}$
k_i is the largest positive integer such that $(\lambda_i - t)_i^k$ is a factor of $P_A(t)$.

Definition 12.8.3: Geometric Multiplicity

$m_i = \dim E_{\lambda_i}$.

Theorem 12.8.1

Let $T : V \to V$ linear. $\lambda \in F$ and eigenvalue. Let $m = \dim E_\lambda$, k = algebraic multiuplicity of λ. Then $1 \leqslant m \leqslant k$.

Proof: Since λ is an eigenvalue $\longrightarrow \dim E_\lambda \geqslant 1$. This implies that $m \geqslant 1$. It remains to show $m \leqslant k$. . It remains to show $m \leqslant k$. Let $B_\lambda = \{v_1, ..., v_m\}$ be a basis for E_λ extend it to a basis for $B = \{v_1, ..., v_m, v_{m+1}, ..., v_n\}$ of V.
Because $v_1, ..., v_m \in E_\lambda$ we know that $T(v_1) = \lambda v_1 = \lambda v_1 + 0v_2 + ... + 0v_n$
All the way to $T(v_m) = \lambda v_m = 0v_1 + ... + \lambda v_m + 0v_m + 1 + 0 + ...$
At the end of the day we get $[T]_B = \begin{bmatrix} \lambda I_m & c \\ 0 & B \end{bmatrix}$ for $B = M_{n-m}(F)$.

$$\longrightarrow P_T(t) = \begin{bmatrix} \lambda - tI_m & c \\ 0 & B - tI_m \end{bmatrix}$$

$\longrightarrow P_T(t) = (\lambda - t)^m P_B(t)$ remember that $m = \dim E_\lambda$
This implies that $(\lambda - t)^m$ is a factor of $P_T(t)$ and that $1 \leqslant m \leqslant k$ because k is the largest possible factor in the characteristic polynomial.

☺

Theorem 12.8.2

Let $T : V \longrightarrow V$ linear transformation suppose that $P_T(t)$ splits completely. Let $P_T(t) = (\lambda_1 - t)^{k_1} ... (\lambda_r - t)^{k_r}$. For each $i \in \{1, ..., r\}$ let $m_i = \dim E_{\lambda i}$. Then T is diagonalizable $\leftrightarrow m_i = k_i$ for $i = 1, ..., r$.

Proof: $(\longrightarrow)$ suppose that T is diagonalizable. $\longrightarrow \exists B$ a basis for V consisting of eigen vectors for T.
For each $i = 1, ..., r$ let $B_i = B \cap E_{\lambda i}$.
B_i is a linearly independent subset of the eigenspace. This implies that $\#B_i \leqslant \dim E_{\lambda i} \longrightarrow \#B \leqslant m_i$.
If we add all of these inequaliteis together we get that $\#B_1 + ... + \#B_r \leqslant m_1 + ... + m_r$ On the left we have the sum equal to $\#B = \dim V = n$. On the other hand, by our previous theorem we know that $\Sigma_{i=1}^r m_i \leqslant \Sigma_{i=1}^r k_i$. The sum of the k_i's is n. This is because n is the degree of the polynomila. This is because the degree of the $P_t(T)$ is the sum of the k_i's. This implies that all the inequalities must be equalities: $\#B_i \leqslant m_i \leqslant k_i$ must be equalities.
$m_i = k_i = \#B_i$ for all $i = 1, ..., r$. We also deduce that in fact $B_i = B \cap E_{\lambda i}$ is a basis of $E_\lambda i$.
$(\longleftarrow)$ Suppose $k_i = m_i$ for $i = 1, ..., r$.
For each $i = 1, ..., r$ Let B_i =a basis of $E_\lambda 2$
From a theorem of previous lectures we get that $B = B_1 \cup B_2 \cup ... \cup B_r$ is linearly independent subset of V.
Also showed that $B_i \cap B_j = \emptyset$ for $i \neq j$.
We want to show that B is a basis of V because then B will be a basis consisting of eigen vectors of T. So T will be diagonalizable.

😊

Big idea from diagonalizability:

1. $T : V \longrightarrow V$ linear $\lambda \in F$ an eigenvalue
2. E_λ is a T invariant subspace meaning, $T(E_\lambda) \subseteq E_\lambda$
3. If $B_\lambda = \{v_1, ..., v_m\}$ a basis for E_λ then $B = \{v_1, ..., v_m, v_{m+1}, ..., v_n\}$ extends to a basis of V.
4. $[T]_B = \begin{bmatrix} \lambda diag & C \\ 0 & B \end{bmatrix}$

12.8.1 Generalize this:

In general if $W \leqslant V$ a $T-$ invariant subspace of V we can proceed similarly $B_w = \{v_1, ..., v_m\}$ basis of W we can exten this to a basis of V: $B = \{v_1, ..., v_m, v_{m+1, ..., v_n}\}$ Then $[T]_B = \begin{bmatrix} [T_w]_{B_w} & C \\ 0 & B \end{bmatrix}$

Where T_w = the restriction of T to W. $T_W = T_W : W \longrightarrow W$. This implies that $P_T(t) = P_{T_w}(t) \cdot g(t)$.

12.8.2 Take-aways

1. If $W \leqslant V$ is a T invariant subspace of V then $P_T(t) = P_{T_w}(t) \cdot g(t)$ for some $g(t) \in F[t]$. Meaning $P_{T_w}(t)$ is a factor of $P_T(t)$.
2. Suppose that T is not diagonalizable. Goal will be to generalize the notion of an eigenvector $\longrightarrow$ generalized eigenvectors. This will gives us $K_{\lambda i}$ = generalized eigenspaces. The big goal will be to show tha t$V = K_{\lambda_i} \oplus ... \oplus K_{\lambda r}$.

Note:-

Preparation

From HW8 we have Given $A \in M_n(F)$, and $f(t) = a_n t^n + ... + a_1 t + a_0$, polynomial we can evaluate f at A: $F(A) = a_n A^n + ... + a_1 A + a_0 I$ if $T : V \longrightarrow V$ is linear we can evaluate f at T, $f(T) = a_n T^n + ... + a_1 T + a_0 I_v$ $I_v : V \longrightarrow V$, $v \longrightarrow v$.

Theorem 12.8.3 Cayley-Hamilton Theorem

1. Let $A \in M_n(F)$ then A satisfies its characteristic polynomial.

Meaning $P_A(A) = 0$.

2. Let $T : V \longrightarrow V$ linear then $P_T(T) = 0$. This means $P_T(T)(v) = 0 \forall v \in V$.

Example 12.8.1 (Verify Cayley Hamilton)

$$A = P \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} P^{-1}$$
$$\text{Set } D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$
$$P_A(t) = P_D(t) = (\lambda_1 - t)(\lambda_2 - t)$$
$$\text{What we want to show is that } (\lambda_1 I - A)(\lambda_2 I - A) = 0.$$
$$(\lambda_1 I - D)(\lambda_2 I - D) = \begin{bmatrix} 0 & 0 \\ 0 & \lambda_1 - \lambda_1 \end{bmatrix} \begin{bmatrix} \lambda_1 - \lambda_2 & 0 \\ 0 & 0 \end{bmatrix}$$
$$A = PDP^{-1}$$
$$P_A(A) = P_D(A) =$$
$$= P \cdot P_D(D) P^{-1}$$

12.9 Reminders

Redo p6 if you didnt do it well from the midterm.

Note:-

Final will be take home. Before the actual exam date so we can finish earlier. Towards the end of the first week after classes end.

Theorem 12.9.1 Cayley-Hamilton Theorem

If $A \in M_n(\mathbb{F})$, then $P_A(A) = 0$. For linear transformations if $T : V \longrightarrow V$ linear, then $P_T(T) = 0$. Proof for $P_T(T)$ below. $P_T(T) : V \longrightarrow V$ means that $P_T(T)(v) = 0, \quad \forall v \in V$.

12.9.1 Preparation

If $T : V \longrightarrow V$ is linear and we have $W \leqslant V$ a T invariant subspace, set $T_W : W \longrightarrow W$ the restriction of T to W. Then $P_{T_W}(t)$ is a factor of $P_T(t)$.

This means that $P_T(T) = P_{T_W}(t) g(t)$ for some $g(t) \in F[t]$.

Proof: Strategy of Cayley-Hamilton Theorem. We want to show $P_T(T)(v) = 0, \quad \forall v \in V$. Fix $T : V \longrightarrow V$ linear.

1. If $v = 0 \longrightarrow$ clearly true becuase $P_T(T)$ is a linear transformation.

2. Fix $v \neq 0$. We do a proof for each individual v. For this v we will find a specific T invariant subspace denoted $W_v \longrightarrow P_T(t) = P_{T_{W_v}}(t) g(t) \longrightarrow$
$P_T(T) = P_{T_{W_V}}(T) g(T) =$
$g(T) P_{T_{W_V}} \longrightarrow$
$P_T(T)(v) = g(T)(P_{T_{W_V}}(T)(v))$ it is enough to show that $P_{T_{W_V}}(T)(V) = 0$.

From now on: $T : V \longrightarrow V$ linear, and $v \in V, v \neq 0$ fixed. Start first with a definition.

Definition 12.9.1: T cyclic subspace of V corresponding to v

$W_v = \text{ span}\{v, T(v), T^2(v), \dots, T^n(v)\}$

Note:-

$W \leqslant V$ and $W = \text{span}\{v_1, ..., v_k\}$ then W is Tinvariant $\leftrightarrow$ each $T(v_i) \in W$ for $i = 1, ..., k$.

This implies that W_v is $T-$invariant since $T(T^i(v)) = T^{i+1}(v) \in W_v \quad \forall i \geqslant 0$.
$W_v \leqslant V$, if $\dim V = n$ then $\dim W_v = k$ for some $1 \leqslant k \leqslant n$. This is because $v \in W_v, v \neq 0$.
From now on, we assume that $\dim W_v = k$.
Suppose $\dim W_v = k$. Then the claim is that the set $\{v, T(v), ..., T^{k-1}(v)\}$ is a basis for W_v.

Proof: $W_v = \text{span}\{v, t(v), ..., T^n(v)\}$.
Let $1 \leqslant j \leqslant k$ Be the largest integer such that the vectors $v, T(v), ..., T^{j-1}(v)$ are linearly inedpendent. We can find a j such that the first j vectors are linearly independent. We want to show that $j = k$. We cannot have more than k of these guys.
It is enough to show that $\text{span}\{v, T(v), ..., T^{j-1}(v)\} = W_v$.
Set $U = \text{span}\{v, T(v), ..., T^{j-1}(v)\}$. We want to show that $U = W_v$. U is definitely contained in W_v. That much is clear. All we have to show is that every higher power of $T(v)$ is still in U. This amounts to showing that $T^n(v) \in U \ \forall n \geqslant j$.
We will start with T^j and all of the rest will be exactly the same. By definition of j we get that the vectors $\{v, T(v), ..., T^{j-1}(v), T^j(v)\}$ are dependent. In fact $\exists a_0, a_1, ..., a_{n-1} \in F$ such that $T^j(v) + a_{j-1}T^{j-1}(v) + ... + a_1T(v) + a_ov = 0$. Now we can solve for $T^j(v)$ it is in the $\text{span}\{v, T(v), ..., T^{j-1}(v)\} = U$.
Now we can write $T^j(v) = -a_{j-1}T^{j-1}(v)...a_1T(v) - a_0v \rightarrow$
$T^{j+1}(v) = -a_{j-1}T^j(v)...a_1T^2(v) - a_0T(v) \in \text{span}\{v, T(v), ..., T^{j-1}(v)\}$.
$\rightarrow T^{j+1}(v) \in \text{span}\{v, T(v), ..., T^{j-1}(v)\}$ Similarly, $T^n(v) \in U, \forall n \geqslant j$.

😊

Note:-

Conclusion so Far:
$T : V \rightarrow V$ linear and $v \neq 0$. $W_v = \text{span}\{v, T(v), ..., T^{k-1}(v)\}$ with $k = \dim W_v$ and $T^k(v) \in W_v \rightarrow \exists a_0, a_1, ..., a_{k-1} \in F$ such that $T^k(v) + a_{k-1}T^{k-1}(v) + ... + a_1T(v) + a_0v = 0$.

The characteristic polynomial of T_{W_V} is $P_{t_{W_V}} = (-1)^k(a_0 + a_1t + ... + a_{k-1}t^{k-1} + t^k)$. Proposition 2 implies cayley hamilton because we want to show $P_{T_{W_V}}(T)(V) = 0$. Which follows from $T^k(v) + a_{k-1}T^{k-1}(v) + ... + a_1T(v)$ where $\forall v = 0$.

Proof: $T_{W_V} : W_V \rightarrow W_V$. Take the ordered basis $\beta = \{v, T(v), ..., T^{k-1}(v)\}$ of W. Now we compute $[T_{W_V}]_\beta$

$$T(v_1) = T(v) = v_2$$
$$T(v_1) = 0v_1 + 1v_2 + 0 + ...0v_k$$
$$T(v_2) = T(T(v_1) = T^2(v) = v_3$$
$$\cdots \ T(v_{k-1}) = v_k = 0v_1 + ... + 0v_{k-1} + 1v_k$$
$$T(v_k) = T(T^{k-1}(v)) = T^k(v) = -a_0v - a_1T(v) - ... - a_{k-1}T^{k-1}(v)$$
$$= -a_0v_1 - a_1v_2 - ... - a_{k-1}v_k$$

$$[T_{W_V}]_\beta = \begin{bmatrix} 0 & 0 & ... & -a_0 & & \\ 1 & 0 & ... & 0 & -a_1 & \\ 0 & 1 & 0 & ... & 0 & \vdots \end{bmatrix}$$ You get the vandermonde determinant in the characteristic polynomial!

$$P_{T_W}(T) = (-1)^k(t^k + a_{k-1}t^{k-1} + ... + a_1t + a_0)$$

😊

😊

Example 12.9.1 (Application)

Let $A = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$ We can compute the characteristic: $P_A(t) = (1-t)(2-t)(3-t)$ By Cayley Hamilton we get that $P_A(A) = 0$. This implies that $-A^3 + 6A^2 - 11A + 6I = 0$. Get this by factoring out the roots. We know that A is not invertible because 0 is not an eigenvalue. Then we can multiply by A^{-1} and get $-A^2 + 6A - 11I + 6A^{-1} = 0$ we can solve here for A^{-1}.

Chapter 13

Chapter 7 Jordan Canonical Form (JCF)

13.0.1 Reminders

1. $A \in M_n(\mathbb{F})$ is diagonalizable if and only if $F^n = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda r}$ where $\lambda_1, \ldots, \lambda_r$ are distinct eigenvalues of A. If and only if $\dim E_{\lambda i}$ = algebraic multiplicity of λ_i for $i = 1, \ldots, r$. One more equivalence. $\leftrightarrow v \in \mathbb{F}^n$ has unique expression as $v = v_1 + \ldots + v_r$ with $v_i \in E_{\lambda i}$.

Note:-

If $T : V \longrightarrow V$ linear, and v and eigenvector with corresponding eigenvalue $\lambda \in F$ then $(T - \lambda I)(v) = 0$.

Definition 13.0.1

Let $T : V \longrightarrow V$ linear $\lambda \in \mathbb{F}$ a scalar. A vector $v \in V, v \neq 0$ is called a generalized eigenvector of T with respect to λ if $(T - \lambda I)^p(v) = 0$ for some $p \geqslant 1$.

1. If v is a regular eigenvector of T with eigenvalue λ, then v is a generalized eigenvector. With $p = 1$.

2. We can have more generalized eigenvectors than regular eigenvectors.

Example 13.0.1

$A = \begin{bmatrix} \lambda & 2 & 3 \\ 0 & \lambda & 4 \\ 0 & 0 & \lambda \end{bmatrix}$ we know this is not diagonalizable since A has unique eigenvalue λ and it is not diagonal.

$$P_A(t) = (\lambda - t)^3$$

By Cayley-Hamilton $\longrightarrow$

$P_A(A) = 0 \longrightarrow (A - \lambda I)^3 = 0$ which implies that every $v \in F^3$ is a gerneralized eigenvector of A. In fact $p = 3$ works for every $v \in F^3$.

Note:-

Suppose $v \in V$ for $v \neq 0$ is a generalized eigenvector of T with respect to some $\lambda \in F$. Then λ must be an eigenvalue.

Proof: $\exists p \geqslant 1$ such that $(T - \lambda I)^p(v) = 0$. Now we take cases:

1. If $p = 1$ then we are done v is a regular eigenvector, therefore λ is an eigenvalue.

2. Suppose $p > 1$ and it is the smallest possible positive interger such that $(T - \lambda I)^p(v) = 0$. Because it is the smallest possible, this means that $(T - \lambda I)^{p-1}(v) \neq 0$ by minimality of p. Set $w = (T - \lambda I)^{p-1}(v) \neq 0$ Then $T(-\lambda I)(w) =$
$(T - \lambda I)((T - \lambda I)^{p-1}(v)) = (T - \lambda I)^p(v) = 0$ implies that w is a regular eigenvector of T.

☺

13.0.2 Reminders

1. If $A \in M_n(\mathbb{F})$ then $P_A(A) = 0$
2. If $T : V \longrightarrow V$ linear then $P_T(T) = 0$

Definition 13.0.2

Let $T : V \longrightarrow V$ linear. A vector $v \in V, v \neq 0$ is a generalized eigenvector of T with respect to $\lambda \in F$ if $\exists p \geqslant$ such that $(T - \lambda I)^p(v) = 0$.

Note:-

Takeaways

1. λ must be an eigenvalue of T
2. Suppose $p \geqslant 1$ is the smallest possible $(T - \lambda I)^p(v) = 0$ This implies that $(T - \lambda I)^{p-1})(v) = w \neq 0$ for w a regular eigenvector. Then $(T - \lambda I)^{p-1}(v) \in E_\lambda$

Definition 13.0.3

K_λ for λ an eigenvalue of T. We define $K_\lambda = \{v \in V$ such that v is a generalized eigenvector corresponding to $\lambda\}$

$= \{v \in V s.t \exists p \geqslant 1 s.t (T - \lambda I)^p(v) = 0\}$

Let this be the generalized eigenspace of T correpsonding to λ

Theorem 13.0.1

K_λ is a subspace of V

Proof: K_λ is closed under scalar multiplication. Let $v \in K_\lambda, c \in \mathbb{F}$ and we want to show that $c \cdot v \in K_\lambda$.

$$v \in K_\lambda \longrightarrow \exists p \geqslant s.t.(T - \lambda I)^p(v) = 0$$

Apply the same thing to (cv)

$$(T - \lambda I)^p(cv) = c(T - \lambda I)^p(v) = c \cdot 0.$$

Note:-

Closure under addition below here.

Let $v, w \in K_\lambda$ and we want to show $v + w \in K_\lambda$. $v \in K_\lambda \longrightarrow \exists p \geqslant 1$ such that $(T - \lambda I)^p(v) = 0$ and $w \in K_\lambda \longrightarrow \exists s \geqslant 1$ such that $(T - \lambda I)^s(w) = 0$.

If we have a p that works then any larger p works. This is because $(T - \lambda I)^p(v) = 0$ for some $p \geqslant 1$ then

$$(T - \lambda I)^r(v) = 0 \quad \forall r \geqslant p$$

$$(T - \lambda I)^r(v) =$$

$$(T - \lambda I)^{r-p}(T - \lambda I)^p(v) =$$

$= (T - \lambda I)^{r-p}((T - \lambda I)^p)(v)) = 0$ because the right hand side is equal to 0.

Set $r = \max\{s, p\}$.

Then $(T - \lambda I)^r(v) = (T - \lambda I)^r(w) = 0$ This implies that $(t - \lambda I)^r(v + w) = 0$.

☺

Example 13.0.2

Let $A = \begin{bmatrix} \lambda & 2 & 3 \\ 0 & \lambda & 4 \\ 0 & 0 & \lambda \end{bmatrix}$

$P_A(T) = (\lambda - t)^3$ by Cayley Hamilton we know that $P_A(A) = 0 \to (A - \lambda I)^3 = 0$. This means $\forall v \in F^3, \quad (A - \lambda I)^3(v) = 0$. This implies that $F^3 = K_\lambda$

1. Compute E_λ
2. Look for $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in F^3$ such that $(A - \lambda I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \leftrightarrow \begin{cases} 2y + 3z = 0 \\ 4z = 0 \end{cases} \quad \to y = z = 0$

   This implies that $E_\lambda = \{\begin{bmatrix} x \\ y \\ z \end{bmatrix} : x \in F\} = \text{ span}\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\}$

3. We only have one eigenvector. Step 2 is to look for a generalized eigenvector of order 2. Meaning we seek $v_2 \in F^3$ such that $(A - \lambda v_2)^2(v_2) = 0$, but $(A - \lambda I)(v_2) \neq 0$. We call this generalized eigenvector of order 2 if 2 is the smallest guy that sends to 0.

   Find $v_2 \in F^3$ such that $(A - \lambda I)(v_2) = v_1$. This is something that we can solve by hand.

4. Look for $v_2 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in F^3$ such that $\begin{bmatrix} 0 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

   $\begin{cases} 2y + 3z = 1 \\ 4z = 0 \end{cases} \quad \to y = \frac{1}{2}$ and $z = 0$.

   If we take $v_2 = \begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \end{bmatrix}$ satisfies what we want.

   So far we have $v_1 \in E_\lambda$ and a v_2 such that $(T - \lambda I)(v_2) = v_1$.

5. Find $v_3 \in F^3$ such that $(A - \lambda I)^3(v_3) = 0$, but $(A - \lambda I)^2(v_3) \neq 0$. As before the easy way to do this is to look for:

   $v_3 = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ such that $(A - \lambda I)(v_3) = v_2$. As before this can be solved:

   $\begin{bmatrix} 0 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{3}{16} \\ \frac{1}{8} \end{bmatrix} = v_3$ satisfies this.

   Note:-

   Now let's see what we have achieved.

   $\beta = \{v_1, v_2, v_3\} = \{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -\frac{3}{16} \\ \frac{1}{8} \end{bmatrix}\}$ Becuase these are clearly independent β is clearly a basis.

   We know that $Av_1 = \lambda v_1 = \lambda v_1 + 0v_2 + 0v_3$.

   The way we constructed v_2 : is $(A - \lambda I)(v_2) = v_1 \to Av_2 = v_1 + \lambda v_2$

   Similarly for v_3 we have $(A - \lambda I)(v_3) = v_2 \to Av_3 = v_2 + \lambda v_3$.

This gives us $A = P \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$ for $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{3}{16} \\ 0 & 0 & \frac{1}{8} \end{bmatrix}$, thus $A \sim \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$

Note:-

Standard method going forward below

Example 13.0.3

Let $B = \begin{bmatrix} \lambda & 2 & 4 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$

$P_B(t) = (\lambda - t)^3 \longrightarrow F^3 = K_\lambda$ by Cayley Hamilton.

1. Step 1 is to always compute the eigenspace. Look for $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in F^3$ such that $(B - \lambda I) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \leftrightarrow$

   $\begin{cases} 2y + 4z = 0 \\ y = -2z \end{cases}$

   Then $E_\lambda = \{ \begin{bmatrix} x \\ -2z \\ z \end{bmatrix} : x, z \in F\} = \text{ span}\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} \}$

2. Choose a generalized eigenvector of order 2. Need $V = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ such that $(B - \lambda I)^2(v) = 0$ but $v \notin E_\lambda$.

   Becuase if it is not in E_λ then it is not in order 2.

   $(B - \lambda I)^2(v) = 0$

   $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} v = 0$. This means that every $v \in F^3$ satisfies that $(B - \lambda I)^2 v = 0$. Then we can take the simplest vector in F^3 that is not in E_λ take $v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

3. Now we begin from here: $v_2 = (B - \lambda I)v_3 = \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}$. So far we have built two vectors v_2, v_3 for v_2 an eigenvector and v_3 lies above v_2 in the same way as before. $Bv_3 = v_2 + \lambda v_3$.

   Take $\beta = \{v_1, v_2, v_3\} = \{ \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \}$ this gives us that:

   $B \sim \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$ We observe that this is different than the one that we got before.

4. $K_\lambda = \{v \in V : (T - \lambda I)^p(v) = 0 \text{ for some } p \geqslant 1\}$

Let $v \in K_\lambda v \neq 0$, we say v is a generalized eigenvector of order $p \geqslant 1$ if $(T - \lambda I)^p(v) = 0$ and $(T - \lambda I)^{p-1}(v) \neq 0$. This implies that for $w = (T - \lambda I)^{p-1}(v) \in E_\lambda$.

Definition 13.0.4

$\{(T - \lambda I)^{p-1}(v), (T - \lambda I)^{p-2}(v), \ldots, (T - \lambda I)(v), v\}$ is called a cycle of generalized eigenvectors of T corresponding to λ.

The first guy is always an eigenvector: $(T - \lambda I)^{p-1}(v)$ is called the initial vector of the cycle. And v is called the end vector of the cycle. We say that the cycle has length p. This is clear because we have exactly p vectors in our set.

Theorem 13.0.2

The cycle/ set $\{(T - \lambda I)^{p-1}(v), \ldots, (T - \lambda I)(v), v\}$ is always linearly inedpendent.

Note:-

Assuming the above proposition / theorem. Given $v \in K_\lambda$ of order p we get a linearly indpendent subset of size p which we can extend to a basis. $B = \{(T - \lambda I)^{p-1}(v), \ldots, (T - \lambda I)(v), v\}$, provide an ordering to the first p of these, meaning v is u_p.

What happens when we try and compute the matrix representation with respect to B? We get a block which isλ on the main diagonal and 1s on the off diagonal and 0 below, and then the rest.

$$U_1 \in E_\lambda \longrightarrow T(u_1) = \lambda u_1$$
$$U_2 \in (T - \lambda I)^{p-2}(v) \longrightarrow (T - \lambda I)(u_2) = u_1 \longrightarrow T(u_2) = \lambda u_2 + u_1.$$
$$T(u_p) = \lambda u_p + u_{p-1}$$

This matrix is special and we will treat it as such its name is:

Definition 13.0.5

Let $\lambda \in F$ and $p \geqslant 1$ we will define $J(\lambda, p) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$ we will call this the Jordan λ block of size p. This is of course in $M_p(\mathbb{F})$.

What we will show is that every $A \in M_n(\mathbb{C})$ is similar to a matrix in what we will call Jordan Canonical Form.

The number of jordan blocks and the sizes will be completely determine the similarity class. Because of the $\dim E_\lambda$ $j - \lambda$ blocks.

13.1 Tuesday November 19th

13.2 Reminders

1. $T : V \longrightarrow V$ linear, and $\lambda \in F$ an eigen value.
2. $K_\lambda = \{v \in V$ such that $(T - \lambda I)^p(v) = 0\}$ for some $p \geqslant 1$. This is what we call the generalized eigen space, correpsponding to λ.
3. Starting with $v \in K_\lambda$ we take $p \geqslant 1$ smallest possible such that $(T - \lambda I)^p(v) = 0$, and from this we make a cycle of length p which is the following:
$\{(T - \lambda I)^{p-1}(v), (T - \lambda I)^{p-2}(v), \ldots., (T - \lambda I)(v), (v)\}$. Note that $(T - \lambda I)^{p-1}(v) \in E_\lambda$

Theorem 13.2.1

In the above situation the cycle $\{(T-\lambda I)^{p-1}(v), (T-\lambda I)^{p-2}(v), ..., (T-\lambda I)(v), (v)\}$ is linearly independent.

Proof: Let $c_0, c_1, ..., c_{p-1} \in F$ such that $c_0 V + c_1(T-\lambda I)(v) + ... + c_{p-1}(T-\lambda I)^{p-1}(v) = 0$. We want to show that $c_0 = c_1 = ... = c_{p-1}$. We apply the linear transformation $(T-\lambda I)^{p-1}$ to the above equation that we have. This gives us:

$$c_0(T-\lambda I)^{p-1}(v) + c_1(T-\lambda I)^{p-1}((T-\lambda I)(v)) + ...$$

In the end we get:

$$c_0(T-\lambda I)^{p-1}(v) + c_1(T-\lambda I)^{p}(v) + ... + c_{p-1}(T-\lambda I)^{2p-L}(v)$$

What is nice is that everything except the first term is zero. This is becasue

$$(T-\lambda I)^p(v) = 0 \longrightarrow (T-\lambda I)^r(v) = 0, \forall r \geqslant p.$$

By doing this we get $c_0(T-\lambda I)^{p-1}(v) = 0$. We know however that c_0 must be equal to zero because the other factors are not zero by definition.

Put this back into equation one above, and we get:

$c_1(T-\lambda I)(v) + ... + c_{p-1}(T-\lambda I)^{p-1}(v) = 0$. From here we now apply $(T-\lambda I)^{p-2}$ gives $c_1 = 0$. We continue in this fashion semi inductively.

😄

Note:-

Recall $W =$ span $\{(T-\lambda I)^{p-1}(v), ..., (T-\lambda I)(v), v\}$. We will show that W is T invariant. We also know that this span is a basis for B, call it B_w.

If we take $[T_w]_{B_w} = J(\lambda, i, p)$. This is what we call Jordan λ block of size p.

Theorem 13.2.2

$T : V \longrightarrow V$ linear, $\lambda \in F$ an eigenvalue. Then all of the following are true:

1. K_λ is T invariant.
2. Let $\mu \neq \lambda$ be any other eigenvalue of T. Then $K_\lambda \cap K_\mu = \{0\}$
3. For any scalar $\mu \in F$ not necesarrily an eigenvalue:

   If we take $T - \mu I|_{k_\lambda}$ is an isomorphism.

Note:-

If $W \leqslant V$ is T invariant, then W is actually $F(T)$-invariant $\forall f(t) \in F[t]$.

Example 13.2.1

$f(t) = t - \alpha$ for $\alpha \in F$. Then $f(T) = T - \alpha I_v$

$w \in W \longrightarrow f(T)(w) = T(w) - \alpha w \in W$. This argument extends to for all polynomials.

Proof: Proof of (a). We already proved that generalized eigenspaces are subspaces. All we must verify now is that these subspaces are T invariant.

Already know $K_\lambda \leqslant V$. W.T.S $T-$ invariant.

Let $v \in K_\lambda \longrightarrow \exists p \geqslant 1$ such that $(T-\lambda I)^p(v) = 0$.

We want to show that $T(v) \in K_\lambda$. This follows from the following: If we do:

$(T-\lambda I)^p(v) = (T-\lambda I) \circ T(v)$. Since these are polynomial expansions of T they commute. This allows us to commute them:

$$= T \circ (T-\lambda I)^p(v) = T((T-\lambda I)^p(v))$$
$$= T(0) = 0 \longrightarrow T(v) \in K_\lambda.$$

😄

(b) is the main one as (c) follows from (b).

Proof: Let $v \in K_\lambda \cap K_\mu$ suppose for contradiction that $v \neq 0$.
Then we can find $p, q \geqslant 1$ smalles possible such that $(T - \lambda I)^p(v) = 0$ and $(T - \mu I)^q(v) = 0$.
Because $(T - \lambda I)^p(v) = 0$ and $(T - \mu I)^q(v) = 0$. This gives us $(T - \lambda I)^{p-1}(v) = U \neq 0$ and $(T - \lambda I)^{q-1}(v) = W \neq 0$.
$U \in E_\lambda$. We calim that $U \in E_\lambda \cap K_\mu$. Same is true for the other counterpart. This follows from $v \in K_\mu \longleftarrow f(T)-$invariant. This invariance implies that $(T - \lambda I)^{p-1}(v) \longleftarrow K_\mu$
Given that $U \in E_\lambda \cap K_\mu \longrightarrow U \in E_\lambda$. Because $U \in E_\lambda \longrightarrow T(U) = \lambda U$.
Additionally this gives us $(T - UI)^r(v) = (\lambda - u)^r(u) \neq 0$. $\forall \lambda \geqslant 0$. But $u \in K_u$ which is a contradiction.
We used from the homework the fact that if v is an eigenvector corresponding to λ for $T \longrightarrow v$ for $f(T)$ corresponding to $f(\lambda)$. We applied this for $f(t) = (t - \mu)^r$.
The contradiction that u is supposed to be in K_μ but all of the powers end up giving me something that is non zero. This is our contradiction.

😊

Note:-

u is a polynomial expression of v since $v \in K_\mu \longrightarrow u \in K_\mu$.

Proof: First recall that $\forall \mu \in F$, K_λ is $T - \mu I$ invariant. This is because of the polynomial representation also being in $K\mu$?
Because of this if we restrict:
$T - \mu I|_{k_\lambda} : k_\lambda \longrightarrow k_\lambda$. We want to show that for every $\mu \neq \lambda$, that $k_\lambda \longrightarrow k_\lambda$ is an isomorphism.
We have a linear transformation:
$T_\mu I|_{K_\lambda} : K_\lambda \longrightarrow K_\lambda$. Since the dimension and the target are the same it is sufficient to show that $T_\mu I|_{K_\lambda}$ is injective. This follows from the dimension theorem. There are two cases to consider:

1. If μ is not an eigenvalue, μ not an eigenvalue implies that $T - \mu I$ is invertible. This implies that its restrection: $T_\mu I|_{k_\lambda}$ is injective. This automatic by definition because that is what the μ guys do.
2. Suppose $\mu \neq \lambda$ is an eigenvalue. Let $v \in \ker(T - \mu I|_{k_\lambda})$.
$v \in \ker(T - \mu I|k_\lambda$ means that $(T - \mu I)(v) = 0 \longrightarrow v \in E_\mu$. But also $v \in K_\lambda$, this means that $\ker(T - \mu I|_{k_\lambda} \subseteq E_m u \cap K_\lambda = \{0\}$. By proof of (b) we can conclude because we showed that this is equal to $\{0\}$.

😊

Theorem 13.2.3

$T : V \longrightarrow V$ linear. Let $P_T(t) = (\lambda_1 - t)^{k_1} ... (\lambda r - t)^{k_r}$. For $\lambda_1, ..., \lambda_k$ the distinct eigenvalues.

1. $V = K_{\lambda_1} \oplus K_{\lambda_2} \oplus ... \oplus K_{\lambda_r}$
2. $\dim K_{\lambda_i} = K_i$ for all $i = 1, ..., r$.

Note:-

Follows immediately from Cayley Hamilton if there is only one eigenvalue.

Proof: First we will show that for each $i = 1, ..., r$ $\dim K_{\lambda_i} \leqslant k_i$.
Next we will show that $K_{\lambda_i} = \ker((T - \lambda_i I)^{k_i})$.
If $P_A(t) = (\lambda - t)^n$ it follows by Cayley Hamilton that $P_A(A) = 0$. Which means that $K_\lambda = V$.

Note:-

Now we begin our proof of the proposition.

We know that K_{λ_i} is T invariant from Theorem 1. This means that if we take $P_{T_{k_{\lambda_i}}}(t)$ is a factor of $P_T(t)$. We discussed this last week.
Our $P_T(t) = (\lambda_1 - t)^{k_1} ... (\lambda_r - t)^{k_r}$. Theorem 1 above part (c) implies that the only eigenvalue of $T_{k_{\lambda_i}}$ is λ. This is because $\forall \mu \neq \lambda$, $T - \mu I|_{k_\lambda}$ is an isomorphism. This implies that μ is not an eigenvalue of $T|_{k_{\lambda_i}}$.
We know that the characteristic polynomial of this fourth subscript transformation must be a factor of $P_T(t)$.

K_{λ_i} is T invariant this gives me that $P_{T_{k_{\lambda_i}}}$. Since it is a factor of the original characteristic polynomial then it's power cannot exceed the algebraic multiplicity of the original.

$T_{K_{\lambda_i}} : K_{\lambda_i} \to K_{\lambda_i}$ This implies that $\dim K_{\lambda_i} = l_i$ This implies $\dim K_{\lambda_i} \leqslant k_i$. This proves Proposition one.

Note:-

Now we prove proposition 2: $K_{\lambda_i} = \ker((T - \lambda_i I)^{k_i})$

$P_{T_{\lambda_i}}(t) = (\lambda_i - t)^{l_i}, l_i \leqslant k_i$. Cayley Hamilton Applied to $T_{k_{\lambda_i}} \to P_{T_{K_{\lambda_i}}}(T_{k_{\lambda_i}}) = 0 \to (\lambda_i I - T_{k_{\lambda_i}})^{l_i} = 0$ because $l_i \leqslant k_i \to (\lambda_i I - T_{k_{\lambda_i}})^{k_i} = 0 \to (\lambda I = T_{k_{\lambda_i}})^{k_i}(v) = 0$ for all $v \in K_{\lambda_i}$. But now we are done because

$$K_{\lambda_i} = \ker)(T_{k_{\lambda_i}} - \lambda_i I)^k$$

😆

Note:-

Some reductions.

1. Theorem 2 above has two parts, if we can show 1 then part two follows immediately from proposition two leading up to the proof for theorem 2.
   If we prove $V = K_{\lambda_1} \oplus ... \oplus K_{\lambda_r}$ then $\dim K_{\lambda_i} = k_i$ which is the algebraic multiplicity:

Note:-

This follows from. $\dim k_{\lambda_i} = l_i \leqslant k_i$. If $V = K_{\lambda_i} \oplus ... \oplus K_{\lambda_r}$. This implies that $\dim V = n = l_1 + l_2 + ... l_4$. By the proposition this is less than or equal to $k_1 + ... + k_2 + ... + k_r$.

This implies that each of the inequaliteis of $l_i \leqslant k_i$ must be an equality.

Note:-

Another reduction: Theorem 2 part 1 is equivalent to proving the following:

Theorem 13.2.4 Theorem 2 Prime

Every $v \in V$ has unique expression as $v = v_1 + v_2 + ... + v_r$ with each $v_i \in K_{\lambda_i}$.

Note:-

Preparation for the next class.

Let $f_1(t), ..., f_r(t) \in F[t]$ be relatively prime. This means they don't share any factors other than the identity and common factors.

Relatively prime means that if $p(t) \in F[t]$ is a common factor of $f_1, ..., f_r$ then $p(t) = c$ is a constant.

The fact that we will be using is that if $f_1(t), ..., f_r(t) \in F[t]$ are relatively prime, then $\exists q_1(t), ..., q_r(t) \in F[t]$ generally not unique, with the following property: $f_1(t)q_1(t) + ... + f_r(t)q_r(t) = 1$.

13.3 Thursday November 21st

13.3.1 Reminders

1. $T : V \to V$ linear. With a characteristic polynomial that splits. $P_T(t) = (\lambda_1 - t)_1^k ... (\lambda_r - t)^{k_r}$ with distinct λ_i.

2. We proved a proposition last class with three statements. (a) For each $i = 1, ..., r$ $\dim K_{\lambda_i} \leqslant k_i$ and (b) $k_{\lambda_i} = \ker((T - \lambda_i I)^{k_i})$ and lastly (c) if $\mu \neq \lambda_i$ is any scalar $\to T_\mu I : K_{\lambda_i} \to K_{\lambda_i}$ is an isomorphism.

Theorem 13.3.1

$V = K_{\lambda_i} \oplus ... \oplus K_{\lambda_r}$ and the $\dim K_{\lambda_i} = k_i =$ the algebraic multiplicity of λ_i.

We discussed that it is actually enough to prove the following:

Theorem 13.3.2

Every $v \in V$ has unique expression as $v = v_1 + ... + v_r$ with $v_i \in K_{\lambda_i}$.

Proof: We will use the following fact. (Probaby prove later today). If we have $f_1(t), ..., f_r(t) \in F[t]$, relatively prime polynomials, then $\exists q_1(t), ..., q_r(t) \in F[t]$ such that $f_1(t)q_1(t) + ... + f_r(t)q_r(t) = 1$.

Note:-

We start the proof below beginning with existence.

Existence of Expression: We need to come up with the $v's$. For each $j = 1, ..., r$ I will consider the following polynomial:

$f_j(t) = \prod_{i \neq j}(\lambda_i - t)^{k_i}$, in other words take the characteristic and forget the $(\lambda_j - t)^{k_j}$ factor.

Note:-

We claim that these polynomials $f_1(t), ..., f_r(t)$ are relatively prime.

Suppose not that $f_1(t), ..., f_r(t)$ are not relatively prime. Take $f_1(T) = \prod_{j=2}^{r}(\lambda_j - t)^{k_j}$. If $f_1(t)$ has a common non constant factor with $f_2, ..., f_r$, then $f_1(t)$ will have a linear common factor with the others. This linear factor has to be of the form $(\lambda_j - t)$ for some $j \geqslant 2$. But $f_j(t)$ does not have $(\lambda_j - t)$ as a factor. Explained for f_1 to keep the notation simple.

Note:-

Continue with the proof.

By the Abstract Algebra fact above, This implies that $\exists q_1(t), ..., q_r(t) \in F[t]$ such that $f_1(t)q_1(t) + ... + f_r(t)q_r(t) = 1$. Emphasize the fact that the q_i are not unique.

This implies $\longrightarrow * = f_1(T)q_1(T) + ... + f_r(T)q_r(T) = I$. The equality of linear transformations $V \longrightarrow V$.

Let $v \in V$. Evaluating $*$ at v gives:

$f_1(T)(q_1(T)(v)) + ... + f_r(T)(q_r(T)(v)) = V$, and now for each $j = 1, ..., r$ set $v_j = f_j(T)(q_j(T)(v))$. Using this notation we get an expression that v is $v_1 + v_2 + ... + v_r$.

Note:-

So far we have set up the problem. Now we need to verify that each one of the v_i map to the corresponding K_{λ_i}.

Note:-

Claim that each $v_j \in K_{\lambda_j}$ for $j = 1, ..., r$. Reminder that $v_j = f_j(T)(q_j(T)(v))$. And $f_j(T) = \prod_{i \neq j}(\lambda_i - t)^{k_i}$.

Notice that $(\lambda_j - t)^{k_j} f_j(t) = P_T(t)$.

Note:-

Now we will do Cayley Hamilton as we have been doing the last couple of lectures.

Cayley Hamilton tells me that $P_T(T) = 0 \rightarrow P_T(T)(w) = 0$Here we apply $w = q_j(T)(v)$. This holds $\forall w \in V$. This implies that $(\lambda_j I - T)^{k_j} \circ f_j(T) = 0 \rightarrow (\lambda_j I_v - T)^{kj}(f_j(T)(w)) = 0$ for all $w \in V$.

The above relation implies that $\lambda_j I_V - T)^{k_j}(v_j) = 0$. This implies that $v_j \in K_{\lambda_j}$ for $j = 1, ..., r$.

Note:-

This concludes existence.

Suppose that I have $v = v_1 + v_2 + ... + v_r$ and $v = w_1 + ... + w_r$ are two expressions of v. With $v_i, w_i \in K_{\lambda_i}$. We want to show that each $v_i = w_i$. As per usual this is equivalent to showing that if:

$$v_1 + v_2 + ... + v_r = 0 \text{ with } v_i \in K_{\lambda_i} \text{ then all of the } v_i = 0 \text{ for } i = 1, 2, ..., r$$

Given $v_1 + v_2 + ... + v_r = 0$ and each $v_i \in K_{\lambda_i}$. I will use our polynomial. Fix an index $j \in \{1, ..., r\}$. We want to show that $v_j = 0$.

I will apply the linear transformation $f_j(T) = \prod_{i \neq j}(\lambda_i I - T)^{k_i}$ to $v_1 + v_2 + ... + v_r = 0$.

Because T is linear we get:

$$f_j(T)(v_1) + ... + f_j(T)(v_j) + ... + f_r(T)(v_r) = 0$$

Note:-

Claim for each $i \neq j$, for each $f_j(T)(v_i) = 0$.

Thus we will rewrite this as $\prod_{s \neq i,j}(\lambda_s I - T)^{ks}(\lambda I - t)^{k_i}(v_i i)$. Because $((\lambda_i I - T)^{k_i}(v_i)) = 0$ We know that $v_i \in K_{\lambda_i} = \ker((T - \lambda_I)^{k_i})$.

We are left with $f_j(T)(v_j) = 0$.We want to show that $v_j = 0$. $f_j(T)(v_j) = \prod_{i \neq j}(\lambda_i I - t)^{k_i}(v_j)$. By proposition part c, each $\lambda I - T : K_{\lambda_j} \to k_{\lambda_j}$ is an isomorphism for $i \neq j$. Composition of isomorphisms implies that $f_j(T) : K_{\lambda_j} \to K_{\lambda_j}$ is an isomorphism. $f_j(T)(v_j) = 0 \to v_j = 0$ and so we are done.

😃

13.3.2 Proof of fact for 2 relatively prime polynomials

Proof: 2 Relatively prime polynomials $f(t), g(t) \in F[t]$. I start with 2 relatively prime polynomials and I want to show that there exist polynomials $\alpha, \beta \in F[t]$ such that $f(t)\alpha(t) + g(t)\beta(t) = 1$.

Consider the following set: $S = \{p(t), \in F[t] \text{ such that } p(t) = f(t)s(t) + g(t)u(t)\}$ with the constraint that p is non zero. This is the set of all polynomials of the form that we want to show above.

Convince yourself that this is non empty. We can take s, u to be very large degrees. Definitely non empty.

There exists an element in this $p(t) \in S$ of the smallest possible degree.

By assumption all of my polynomials are non zero. By dividing with the leading coefficient, we may assume that $p(t)$ is monic, this means that $p(x)$ has a leading coefficeint of 1. The claim that we want to show is that $p(t) = 1$.

Since f, g are relatively prime, it suffices to show that $p(t)$ is a common factor of $f(t)g(t)$.

We will prove that $p(t)$ is a factor of $f(t)$ and the statement is exactly symmetric.

We will do long division of polynomials. (Polynomial division of $f(t)$ by $p(t)$.) Kill the leading coefficients one by one until we get a remainder that is smaller and then we stop.

$\to \exists q(t), r(t) \in F[t]$ such that $f(t) = q(t)p(t) + r(t)$ where $r(t) = 0$ or the degree of $r(t)$ is strictly smaller than the degree of p.

We want to show that $r(t) = 0$. We will go by contradiction. Suppose not that $f(t) = q(t)p(t) + r(t)$ and that degree $r(t)$ <degree$p(t)$. Now however, we want to use the main property of $p(t)$, which is the fact that it is in S with the smallest degree.

$$r(t) = f(t) - q(t)p(t)$$

$$p(t) \in S \to \exists \alpha(t), \beta(t) \in F[t] \text{ such that } p(t) = \alpha(t)f(t) + \beta(t)g(t)$$

This gives me:

$$r(t) = f(t) - q(t)(\alpha(t)f(t) + \beta(t)g(t))$$

$r(t) = f(t)(1 - q(t)\alpha(t)) + g(t)(-q(t) + \beta(t))$. This implies that $r(t) \in S$. If we let the factors to be some new polynomial.

But because $p(t)$ is a factor it contradicts the minamilty of $p(t) \in S$ and so our assumption must be false. It must be the case that $r(t) = 0$.

The argument for $g(t)$ is exactly symmetric.

😃

Theorem 13.3.3

$T : V \to V$ linear, with $P_T(t) = (\lambda_1 - t)^{k_1} ... (\lambda_r - t)^{k_r}$, then $V = K_{\lambda_1} \oplus ... \oplus K_{\lambda_r}$. Each K_{λ_i} is T invariant.

Note:-

The theorem implies that if B_i is an ordered basis for K_{λ_i} for each $i = 1, ..., r$ then if we take $B = B_1 \cup ... \cup B_r$ is a basis for V. Note disjoint union. Because of the T invariance. If we compute $[T]_B = \begin{bmatrix} [T_{k_{\lambda_1}}]_{B_2} & 0 \\ 0 & [T_{K_{\lambda_2}}]_{B_2} \end{bmatrix}$

such that we have a block diagonal.

We want to show that for each K_{λ_i} we can find a basis $B_i = \Gamma_1 \cup ... \cup \Gamma_s$. Disjoint union where each of the Γ_i is a cycle of generalized eigenvectors corresponding to λ_i.

Note:-

If $T : V \to V$ and $v \in K_\lambda$ of length $p \geqslant 1$, then $(a)\Gamma = \{(T - \lambda I)^{p-1}(v), ..., (T - \lambda I)(v), v\}$ cycle of length p. Then Γ is linearly independent.

Note:-

We also showed that if we extend Γ to a basis B of $V \to [T]_B = \begin{bmatrix} J(\lambda, P) & B \\ 0 & C \end{bmatrix}$

Our goal above the previous two notes will imply the existence of a JCF. The first step next time will be to fix a λ eigenvalue if $\Gamma = \Gamma_1 \cup ... \cup \Gamma_s$ where each Γ_i is a cycle as before, if the different eigenvalues in the cycle are independent then the entire guys are all independent.

13.4 Thursday November 21st

13.4.1 Reminders

1. $T : V \to V$ linear. With a characteristic polynomial that splits. $P_T(t) = (\lambda_1 - t)^k_1 ... (\lambda_r - t)^{k_r}$ with distinct λ_i.
2. We proved a proposition last class with three statements. (a) For each $i = 1, ..., r$ $\dim K_{\lambda_i} \leqslant k_i$ and (b) $k_{\lambda_i} = \ker((T - \lambda_i I)^{k_i})$ and lastly (c) if $\mu \neq \lambda_i$ is any scalar $\to T_\mu I : K_{\lambda_i} \to K_{\lambda_i}$ is an isomorphism.

Theorem 13.4.1

$V = K_{\lambda_i} \oplus ... \oplus K_{\lambda_r}$ and the $\dim K_{\lambda_i} = k_i =$ the algebraic multiplicity of λ_i.

We discussed that it is actually enough to prove the following:

Theorem 13.4.2

Every $v \in V$ has unique expression as $v = v_1 + ... + v_r$ with $v_i \in K_{\lambda_i}$.

Proof: We will use the following fact. (Probaby prove later today). If we have $f_1(t), ..., f_r(t) \in F[t]$, relatively prime polynomials, then $\exists q_1(t), ..., q_r(t) \in F[t]$ such that $f_1(t)q_1(t) + ... + f_r(t)q_r(t) = 1$.

Note:-

We start the proof below beginning with existence.

Existence of Expression: We need to come up with the $v's$. For each $j = 1, ..., r$ I will consider the following polynomial:
$f_j(t) = \prod_{i \neq j}(\lambda_i - t)^{k_i}$, in other words take the characteristic and forget the $(\lambda_j - t)^{k_j}$ factor.

Note:-

We claim that these polynomials $f_1(t), ..., f_r(t)$ are relatively prime.

Suppose not that $f_1(t), ..., f_r(t)$ are not relatively prime. Take $f_1(T) = \prod_{j=2}^{r}(\lambda_j - t)^{k_j}$. If $f_1(t)$ has a common non constant factor with $f_2, ..., f_r$, then $f_1(t)$ will have a linear common factor with the others. This linear factor has to be of the form $(\lambda_j - t)$ for some $j \geqslant 2$. But $f_j(t)$ does not have $(\lambda_j - t)$ as a factor. Explained for f_1 to keep the notation simple.

Note:-

Continue with the proof.

By the Abstract Algebra fact above, This implies that $\exists q_1(t), \ldots, q_r(t) \in F[t]$ such that $f_1(t)q_1(t) + \ldots + f_r(t)q_r(t) = 1$. Emphasize the fact that the q_i are not unique.
This implies $\rightarrow * = f_1(T)q_1(T) + \ldots + f_r(T)q_r(T) = I$. The equality of linear transformations $V \rightarrow V$.
Let $v \in V$. Evaluating $*$ at v gives:
$f_1(T)(q_1(T)(v)) + \ldots + f_r(T)(q_r(T)(v)) = V$, and now for each $j = 1, \ldots, r$ set $v_j = f_j(T)(q_j(T)(v))$. Using this notation we get an expression that v is $v_1 + v_2 + \ldots + v_r$.

Note:-

So far we have set up the problem. Now we need to verify that each one of the v_i map to the corresponding K_{λ_i}.

Note:-

Claim that each $v_j \in K_{\lambda_j}$ for $j = 1, \ldots, r$. Reminder that $v_j = f_j(T)(q_j(T)(v))$. And $f_j(T) = \prod_{i \neq j}(\lambda_i - t)^{k_i}$.

Notice that $(\lambda_j - t)^{k_j} f_j(t) = P_T(t)$.

Note:-

Now we will do Cayley Hamilton as we have been doing the last couple of lectures.

Cayley Hamilton tells me that $P_T(T) = 0 \rightarrow P_T(T)(w) = 0$Here we apply $w = q_j(T)(v)$. This holds $\forall w \in V$. This implies that $(\lambda_j I - T)^{k_j} \circ f_j(T) = 0 \rightarrow (\lambda_j I_v - T)^{kj}(f_j(T)(w)) = 0$ for all $w \in V$.
The above relation implies that $\lambda_j I_V - T)^{k_j}(v_j) = 0$. This implies that $v_j \in K_{\lambda_j}$ for $j = 1, \ldots, r$.

Note:-

This concludes existence.

Suppose that I have $v = v_1 + v_2 + \ldots + v_r$ and $v = w_1 + \ldots + w_r$ are two expressions of v. With $v_i, w_i \in K_{\lambda_i}$. We want to show that each $v_i = w_i$. As per usual this is equivalent to showing that if:
$v_1 + v_2 + \ldots + v_r = 0$ with $v_i \in K_{\lambda_i}$ then all of the $v_i = 0$ for $i = 1, 2, \ldots, r$
Given $v_1 + v_2 + \ldots + v_r = 0$ and each $v_i \in K_{\lambda_i}$. I will use our polynomial. Fix an index $j \in \{1, \ldots, r\}$. We want to show that $v_j = 0$.
I will apply the linear transformation $f_j(T) = \prod_{i \neq j}(\lambda_i I - T)^{k_i}$ to $v_1 + v_2 + \ldots + v_r = 0$.
Because T is linear we get:
$f_j(T)(v_1) + \ldots + f_j(T)(v_j) + \ldots + f_r(T)(v_r) = 0$

Note:-

Claim for each $i \neq j$, for each $f_j(T)(v_i) = 0$.

Thus we will rewrite this as $\prod_{s \neq i,j}(\lambda_s I - T)^{ks}(\lambda I - t)^{k_i}(v_i i)$. Because $((\lambda_i I - T)^{k_i}(v_i)) = 0$ We know that $v_i \in K_{\lambda_i} = \ker((T - \lambda_I)^{k_i})$.
We are left with $f_j(T)(v_j) = 0$.We want to show that $v_j = 0$. $f_j(T)(v_j) = \prod_{i \neq j}(\lambda_i I - t)^{k_i}(v_j)$. By proposition part c, each $\lambda I - T : K_{\lambda_j} \rightarrow k_{\lambda_j}$ is an isomorphism for $i \neq j$. Composition of isomorphisms implies that $f_j(T) : K_{\lambda_j} \rightarrow K_{\lambda_j}$ is an isomorphism. $f_j(T)(v_j) = 0 \rightarrow v_j = 0$ and so we are done.

😃

13.4.2 Proof of fact for 2 relatively prime polynomials

Proof: 2 Relatively prime polynomials $f(t), g(t) \in F[t]$. I start with 2 relatively prime polynomials and I want to show that there exist polynomials $\alpha, \beta \in F[t]$ such that $f(t)\alpha(t) + g(t)\beta(t) = 1$.
Consider the following set: $S = \{p(t), \in F[t]$ such that $p(t) = f(t)s(t) + g(t)u(t)\}$ with the constraint that p is non zero. This is the set of all polynomials of the form that we want to show above.
Convince yourself that this is non empty. We can take s, u to be very large degrees. Definitely non empty.
There exists an element in this $p(t) \in S$ of the smallest possible degree.
By assumption all of my polynomials are non zero. By dividing with the leading coefficient, we may assume that $p(t)$ is monic, this means that $p(x)$ has a leading coefficeint of 1. The claim that we want to show is that $p(t) = 1$.
Since f, g are relatively prime, it suffices to show that $p(t)$ is a common factor of $f(t)g(t)$.

We will prove that $p(t)$ is a factor of $f(t)$ and the statement is exactly symmetric.
We will do long division of polynomials. (Polynomial division of $f(t)$ by $p(t)$.) Kill the leading coefficients one by one until we get a remainder that is smaller and then we stop.
$\rightarrow \exists q(t), r(t) \in F[t]$ such that $f(t) = q(t)p(t) + r(t)$ where $r(t) = 0$ or the degree of $r(t)$ is strictly smaller than the degree of p.
We want to show that $r(t) = 0$. We will go by contradiction. Suppose not that $f(t) = q(t)p(t) + r(t)$ and that degree $r(t)$ <degree$p(t)$. Now however, we want to use the main property of $p(t)$, which is the fact that it is in S with the smallest degree.
$r(t) = f(t) - q(t)p(t)$
$p(t) \in S \rightarrow \exists \alpha(t), \beta(t) \in F[t]$ such that $p(t) = \alpha(t)f(t) + \beta(t)g(t)$
This gives me:
$r(t) = f(t) - q(t)(\alpha(t)f(t) + \beta(t)g(t))$
$r(t) = f(t)(1 - q(t)\alpha(t)) + g(t)(-q(t) + \beta(t))$. This implies that $r(t) \in S$. If we let the factors to be some new polynomial.
But because $p(t)$ is a factor it contradicts the minamilty of $p(t) \in S$ and so our assumption must be false. It must be the case that $r(t) = 0$.
The argument for $g(t)$ is exactly symmetric.

😄

Theorem 13.4.3

$T : V \rightarrow V$ linear, with $P_T(t) = (\lambda_1 - t)^{k_1} ... (\lambda_r - t)^{k_r}$, then $V = K_{\lambda_1} \oplus ... \oplus K_{\lambda_r}$. Each K_{λ_i} is T invariant.

Note:-

The theorem implies that if B_i is an ordered basis for K_{λ_i} for each $i = 1, ..., r$ then if we take $B = B_1 \cup ... \cup B_r$ is a basis for V. Note disjoint union. Because of the T invariance. If we compute $[T]_B = \begin{bmatrix} [T_{k_{\lambda_1}}]_{B_2} & 0 \\ 0 & [T_{K_{\lambda_2}}]_{B_2} \end{bmatrix}$ such that we have a block diagonal.

We want to show that for each K_{λ_i} we can find a basis $B_i = \Gamma_1 \cup ... \cup \Gamma_s$. Disjoint union where each of the Γ_i is a cycle of generalized eigenvectors corresponding to λ_i.

Note:-

If $T : V \rightarrow V$ and $v \in K_\lambda$ of length $p \geqslant 1$, then $(a)\Gamma = \{(T - \lambda I)^{p-1}(v), ..., (T - \lambda I)(v), v\}$ cycle of length p. Then Γ is linearly independent.

Note:-

We also showed that if we extend Γ to a basis B of $V \rightarrow [T]_B = \begin{bmatrix} J(\lambda, P) & B \\ 0 & C \end{bmatrix}$

Our goal above the previous two notes will imply the existence of a JCF. The first step next time will be to fix a λ eigenvalue if $\Gamma = \Gamma_1 \cup ... \cup \Gamma_s$ where each Γ_i is a cycle as before, if the different eigenvalues in the cycle are independent then the entire guys are all independent.

13.5 Tuesday November 26th

Note:-

Given any matrix we can find a basis such that when we change basis we can get into JCF

13.5.1 Reminders

1. $T : V \rightarrow V$ a linear map. Characteristic polynomial splits completely: $P_T(t) = (\lambda_1 - t)^{k_1} ... (\lambda_r - t)^{k_r}$
2. We have a theorem that showed that $V = K_{\lambda_1} \oplus ... \oplus K_{\lambda_r}$
3. If $\Gamma = \{(T - \lambda I)^{p-1}(v), ..., (T - \lambda I)(v), v\}$ is a cycle of length p of generalized eigenvectors of T corresponding to λ then $\Gamma$is linearly inedpendent.

4. As a reminder in the Γ set, the inital vector is always an eigenvector.

Note:-
Today we will prove two theorems. We will fix each K_λ

Theorem 13.5.1
Let λ be an eigenvalue of T. Let $\Gamma_1, \ldots, \Gamma_s$ be cycles of generalized eigenvectors of T corresponding to λ. All of these cycles correspond to the same λ. For each $i = 1, \ldots, s$ write $\Gamma_i = \{u_{i1}, u_{i1}, \ldots, u_{ipi}\}$, all corresponding to E_λ. Suppose the set $\{U_{11}, U_{21}, \ldots, U_{s1}\}$ of initial vectors forms a linearly independent subset of E_λ. Then the conclusion is:

1. Γ_i is pariwise disjoint from all other Γ_i for $i \neq i$
2. $\Gamma = \Gamma_1 \cup \ldots \cup \Gamma_s$ is linearly independent.

Note:-
Unravel the theorem a bit. Now we have a fixed eigenvalue λ, we are working just with K_λ.

We have $\Gamma_1 = \{u_{11}, u_{12}, \ldots, u_{1p1}\} = \{(T - \lambda I)^{p-1}(v_1), \ldots, (T - \lambda I)(v_1), (v_1)\}$ and we can do all of the following Γ_s in the same manner:

$\Gamma_s = \{u_{s1}, u_{st}, \ldots, u_{sps}\} = \{(T - \lambda I)^{ps-1}(v_s), \ldots, (T - \lambda I)(v_s), (v_s)\}$. The statement assumes that the first value in our cycle is an eigenvalue and as a result the disjoint union of all of the $\Gamma's$ are linearly independdent.

Proof: Showing pairwise disjoint requires contradiction. Suppose for contradiction that for some $i \neq j$ we have $\Gamma_i \cap \Gamma_j \neq \emptyset$. Let $w \in \Gamma_i \cap \Gamma_j$. This means that $\exists 0 \leqslant l \leqslant p_{i-1}$ and $0 \leqslant m \leqslant p_{j-1}$ such that

$$w = (T - \lambda I)^l(v_i) = (T - \lambda I)^m(v_j).$$

Note:-
We take two cases.

1. Suppose that $p_i - 1 - l \neq p_j - 1 - m$. Without loss of generality suppose that LHS > RHS, then we will apply $(T - \lambda I)^{p_i - 1 - l}$ to both sides of relation one above our note. This gives us:
   $(T - \lambda I)^{p_i - 1 - l + l}(v_i) = (T - \lambda I)^{p-1}(v_i) = (T - \lambda I)^r(v_j)$ where $r = p_i - 1 - l + m > p_j - 1 - m + m$.
   Now the relation has become:
   $(T - \lambda I)^{p_i - 1}(v) = (T - \lambda I)^r(v_j) = 0$ on the right hand side we get zero because r is bigger than $p_j - 1$ we know that it goes to zero because it is bigger than the cycle, but the left hand side is not equal to zero. Contradiction.
2. Case 2. Suppose $p_i - 1 - l = p_j - l - m$. The same process as before gives us $u_{i1} = u_{j1}$, which contradicts the linear independence $\{u_{i1}, u_{j1}\}$.

😊

Note:-
Now we prove b

Proof: $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \ldots \cup \Gamma_s$ and we want to show that Γ is independent. For $m = \#\Gamma$ we will do induction on $m \geqslant 1$.

Base case: $m = 1$. $\Gamma = \Gamma_1$. Done.

Induction Hypothesis. Suppose that $m > 1$ and we have $\Gamma' = \Gamma'_1 \cup \ldots \cup \Gamma'_t$ another disjoint union of cycles corresponding to λ such that the following:

1. $\#\Gamma' < m$
2. The initial vectors form a linearly independent subset of E_λ. Then Γ' is independent.

Note:-

This is a strong induction proof. We need all of the previous dominos to fall not just the previous dominos.

Now we want to show that Γ is independent, $\#\Gamma = m$. We will do the following. Set $W = \text{span}\Gamma$. Then it suffices to show that $W = \dim W = m$.
This is now our new goal.

Note:-

We know that each $\text{span}(\Gamma_i)$ is T invariant subspace. This implies that W is T invariant, and therefore will be $f(T)$ invariant for every polynomial $f \in \mathbb{F}[t]$.

Because of this we can consider the linear transformation $F = (T - \lambda I)|_w : W \longrightarrow W$. Because W is T invariant this will map W to W. W is the span of all of the Γ_i.

Note:-

We make some observations about this f;

1. W has a generating set of Γ_i. How does f act on the generators?

   $f(v_i) = (T - \lambda I)(v_i)$

   Where $f(v_i) \longrightarrow f(u_i p_i)$
   $(T - \lambda I)(v_i)$.

   We send each element in our generator to the next value until we get to the last value which we send to 0. This is because the last value is in the eigenspace $E_\lambda . F(U_{i1}) = 0$.

Note:-

For $i = 1, \dots, s$ we get $f(u_{i1}) = 0$. This means that $\{u_{11}, u_{21}, \dots, u_{s1}\} \subseteq \ker F$. By assumption this is linearly independent. This implies that the dimension of the kernel of F is at least s.

Note:-

The second conclusion is that if we forget the last column, everything else is in the image of f.

For each $i = 1, \dots, s$ set $\Gamma_i' = \{(T - \lambda I)^{p_i - 1}(v_i), \dots, (T - \lambda I)(v_i)\} = \Gamma - \{v_i\}$. The other relations tell us that $\Gamma_1' \cup \dots \cup \Gamma_s' \subseteq \text{im}(F)$.

Each Γ_i' is a cycle of length $p_i - 1$ and the initial vector of Γ_i' is the initial vector of Γ_i.

This is the guy by part (a) if we set $\Gamma' = \Gamma_1' \cup \Gamma_2' \cup \dots \cup \Gamma_s'$ is a disjoint union of cycles because the first elements are linearly independent.

We also note that span $\leqslant \text{Im}(F)$.
The size of $\Gamma' = \#\Gamma - s$ because we removed one vector from each of the s cycles in Γ. In particular the size of $\Gamma' < m$ and the induction hypothesis applies. This implies that Γ' is independent.
The dimension theorem for f gives me that the dimension of W is the dimension of the kernel of f plus the dimension of the image.
$\dim W = \dim\ker W + \dim \text{Im}(F)$. We know however that $\dim\ker \geqslant S$ and $\dim \text{Im}(F) \leqslant m - s$, which implies that $\dim W \geqslant m$. And we knew that $\dim W \leqslant m \longrightarrow \dim W = m$.

😃

Theorem 13.5.2 Existence of JCF

Fix an eigenvalue λ of $T : V \longrightarrow V$. Then K_λ has an ordered basis of the form $\Gamma = \Gamma_1 \cup \dots \cup \Gamma_s$ like we had in Theorem 1. Consisting of cycles of generalized eigenvectors.

Proof: By induction of $\dim K_\lambda = n \geqslant 1$. If $n = 1$ clear.

Induction Hypothesis: Let $F : W \to W$, λ an eigenvalue of F such that $\dim K_\lambda(f) < n$, then $K_\lambda(f)$ has a basis Γ of the desired form.

The key will be to construct a subspace of K_λ of strictly smaller dimension and then restrcit T to W, Then that will be our dimenion.

Take the linear transformation $T - \lambda I|_{k_\lambda} : K_\lambda \to K_\lambda$,

The kernel of this is E_λ. In particular the kernel is non zero. This implies that since T has a kernel, the transformation T won't be surjective either. We can set $W = \mathrm{Im}(T - \lambda I|K_\lambda)$. This means that $\dim W < n$. Which is the dimension of K_λ. We claim that W is T-nvariant.

Note:-

Take $w \in W$, $w = \mathrm{im}(T - \lambda I|_{k_\lambda}$. for some $v \in K_\lambda$. So if I apply T. I get $T(w) = (T - \lambda I)(T(v))$, and $v \in K_\lambda$, K_λ is T invariant because of this $T(w) \in W$ as desired.

Take F towards the induction hypothesis to be $T_w : W \to W$. For this linear transformation F W applies. $K_\lambda(F) = W$ this implies that $\dim K_\lambda(f)) < n$ which means that the induction hypothesis applies. This means that $W = \mathrm{Im}(T - \lambda I|_{k_\lambda})$ has a basis $\Gamma = \Gamma_1 \cup \ldots \cup \Gamma_q$ a basis of disjoint cycles of generalized eigenvectors of T_w corresponding to λ by the induction of hypothesis. We can drop the w for T_w because T is restricted to W.

$$\Gamma_i = \{(T - \lambda I)^{p_i - 1}(w_i), \ldots, (T - \lambda I)(w_i), (w_i)\}, \text{ for some } w_i \in W \text{ and } i = 1, \ldots, q.$$

$$w_i \in W = \mathrm{Im}(T - \lambda I|_{k_\lambda} \to w_i = (T - \lambda I)(v_i) \text{ for some } v_i \in K_\lambda.$$

$$\Gamma_1 = \{(T - \lambda I)^{p_i}(v_i), .., (T - \lambda I)(v_1)\}$$

$$\vdots$$

$$\Gamma_q = \{(T - \lambda I)^{p_q}(v_q), \ldots, (T - \lambda I)(v_q)\}$$

For each of these we will append the end vector. We will still have a cycle, and the size of the cycle will go up by 1. For $i = 1, \ldots, q$ extend Γ_i to $\tilde{\Gamma_i} = \Gamma_i \cup \{v_i\} = \{(T - \lambda I)^{pi}(v_i), \ldots, (T - \lambda I)(v_i), v_i\}$ and now this becomes a cycle of length $p_i + 1$.

Notice something important. All we did really was add something extra to all of the Γ_i. For $i = 1, \ldots, q$. The initial vectors stayed the same. In particular they are independent because the set of all initial vectors is independent.

$\{u_1, \ldots, u_q\}$ which are the initial vectors, this is a linearly independent subset of E_λ. If it is a basis I stop. If not I need to extend this to a basis for E_λ. $\{u_1, \ldots, u_q, u_{q+1,\ldots,u_s}\}$ some basis for E_λ. For each $q + 1 \leqslant j \leqslant s$ let $\tilde{\Gamma_j}$ be just $\{u_j\}$ which is a cycle of generalized eigenvectors of length 1.

Now we take $\Gamma = \tilde{\Gamma_1} \cup \ldots \cup \tilde{\Gamma_q} \cup \{u_{q+1}\} \cup \ldots \cup \{u_{q+s}\}$.

The final claim is that Γ is a basis for K_λ. Theorem 1 gives us that Γ is linearly independent. This is because the initial vectors form a basis of E_λ. The rest of this is just the counting of the dimension of W. By the dimension theorem. $\dim W = \dim K_\lambda - s\ldots$ claim follows by counting.

😃

13.6 Tuesday December 03

13.6.1 Sequence of Theorems

Theorem 13.6.1

$T : V \to V$ linear, main assumption is that the characteristic polynomial splits, i.e. $P_T(t) = (\lambda_i - t)^k \ldots (\lambda_r - t)^k_r$ splits.

$V = K_{\lambda_1} \oplus \ldots \oplus K_{\lambda_r}$ and dimension of $K_{\lambda_i} = k_i$ =algebraic multiplicity of λ_i.

Theorem 13.6.2

Fix an eigenvalue λ of T. Suppose we have $\Gamma_1, \ldots, \Gamma_s$ disjoint cycles of generalized eigenvectors with respect to $\Gamma_1 = \{(T - \lambda I)^{p_1}(v), \ldots, (T - \lambda I)(v_1), v_1\}$, with each Γ_i of this form all the way to Γ_s. Assume that the initial vectors are linearly independent. Assume that these initial vectors form a linearly independent subset of E_λ. Then the conclusion is that $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \ldots \cup \Gamma_s$ is linearly independent.

Theorem 13.6.3

Fix an eigenvalue λ of Γ, then K_λ has an ordered basis of the form $\Gamma = \Gamma_1 \cup \ldots \cup \Gamma_s$, consisting of disjoint cycles of generalized eigenvectors corresponding to λ.

Theorem 1 and Theorem 3 implies that T has a Jordan Basis. This means that if I order the eigenvalues in some way, for each K_λ I can build a basis of the form $\Gamma = \Gamma_1 \cup \ldots \cup \Gamma_s$. Strining these bases together we can get $[T]_B = \begin{bmatrix} J(\lambda_i, p_i) & 0 \\ 0 & J(\lambda_r, p_r) \end{bmatrix}$

Theorem 13.6.4

Given $T : V \longrightarrow V$ linear with $P_T(t)$ that splits then $\exists B$ an ordered basis of V such that $[T]_B$ is in the JCF

Theorem 13.6.5

$A \in M_n(F)$ with $P_A(t)$ that splits. Then A is similar to a matrix in JCF.

Note:-

Not diagonalizable over the real numbers mean there is not a JCF for the real numbers. It is all about the characteristic polynomials splitting.

Question 12

How unique is the JCF? What blocks appear? Number of blocks and the sizes.

For each eigenvalue λ, we have that the number of Jordan λ blocks in the JCF(A) is precisely the dimension of the eigenspace E_λ.

Theorem 13.6.6

For a fixed eigenvalue λ, of T or of A, we have: Let r_j = # Jordan λ blocks of size $\geqslant j$. Then $r_j = \operatorname{rank}((T-\lambda I)^{j-1}) - \operatorname{rank}((T-\lambda I)^j)$, this is still a uniquely determined number.

13.6.2 Conclusions

Given $A \in M_n(F)$, then

1. The λ's that contribute Jordan blocks are precisely the eigenvalues which are exactly the roots of $P_A(t)$.
2. For a fixed eigenvalue λ, we have dimension E_λ = # Jordan λ blocks.
3. The sizes for the Jordan-λ blocks are determined by The formula in 0.1.5

Example 13.6.1 (Build a JCF)

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

Step 1 is to compute the characteristic polynomial. $P_A(t)$.

$$P_A(t) = (3-t)(2-t^3)$$

Immediately, $\lambda_1 = 3$ and $\lambda_2 = 2$. We know that dimension of $K_3 = 1 \longrightarrow \dim E_3 = 1$ and the dimension of $E_3 = 3$ by theorem 1. To compute E_3:

$$\text{Look for } \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \mathbb{C}^4 \text{ such that } (A-3I)\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

This gives us:

$$\begin{cases} w - x - y = 0 \\ -z = 0 \\ y - 2z = 0 \\ y = 0 \end{cases} \quad \text{Which gives us:}$$

$y = z = 0$ and $x = w$:

$$E_3 = \{\begin{bmatrix} x \\ 0 \\ 0 \\ x \end{bmatrix} : x \in \mathbb{R}\} = \text{span}\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}\}$$

Now we compute E_2

Look for $\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \mathbb{C}^4$ such that:

$$(A - 2I)\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \leftrightarrow \begin{cases} y = z = w \\ x \text{ is free} \end{cases}$$

This gives us $E_2 = \{\begin{bmatrix} x \\ y \\ y \\ y \end{bmatrix} : x, y \in \mathbb{C}^4\} = \text{span}\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0, 1, 1, 1 \end{bmatrix}\}$, This tells us that there exists exactly 2 jordan 2 blocks. Because $A \in M_4(\mathbb{C})$ and $\dim K_3 = 1$, we are left with $2 + 1 = 3$. Meaning we must have exactly 1 two block of size 1 and 2 block of size 2.

$$\text{JCF}(A) = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Dimension of $E_2 = 2$ tells me I have $\Gamma_1 \cup \Gamma_2$ but dimension of $K_3 = \#(\Gamma_1 \cup \Gamma_2)$

The last step to find the jordan basis:

We need to find $\Gamma_1 \cup \Gamma_2$. We need $E_2 = \text{span}\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}\}$. Compute K_2 explicitly. Let $v = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \mathbb{C}^4$ such that $(A - 2I)^2(v) = 0$, in general we only have to go up to the cubed which is the geometric multiplicity, but here we go up to 2.

$$(A - 2I)^2(v) = 0 \leftrightarrow A = \begin{bmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$2y + z + w = 0 \leftrightarrow w = 2y - z$$

$$K_w = \{\begin{bmatrix} x \\ y \\ z \\ 2y - z \end{bmatrix} : x, y, z \in \mathbb{C}^4\}$$

So it is the $\text{span}\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}\}$, pick a $v \in K_2 - E_2$, Take $v = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 2 \end{bmatrix}$, Then compute $(A - 2I)(v)$ this is supposed to be in E_2, $= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, this way we will build a 2 cycle where we put our eigenvector first and then the generalized cycle:

$\{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\\2\end{bmatrix}\}$, take $v_3 \in E_2$ to be linearly independent from these guys:

$$\{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\\2\end{bmatrix}, \begin{bmatrix}1\\0\\0\\0\end{bmatrix}\}, \text{ this is a basis of } K_2.$$

$$\text{At the very end take } B = B_{K_2} \cup \{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}\} \text{ is a Jordan Basis for } A.$$

Example 13.6.2

If for example dim $E_{\lambda_2} = E_{\lambda_1} = 1$, we can ompute $K_{\lambda_2} = \{(A - \lambda_@ I)^2 v = 0\}$. Step one is to find v such that $(A - \lambda_2 I)^3(v) = 0$ but the square is not.

Note:-

JCF of A is unique up to reordering the Jordan Blocks. As a result:

Theorem 13.6.7

Two matrices A B $\leftrightarrow$JCF(A) =JCF(B).

Example 13.6.3 (2)

Find a # of similarity classes of $A \in M_8(\mathbb{C})$ such that $P_A(t) = (1-t)^3(3-t)^5$. How many non similar JCF's can I possibly have.

Solution: $\lambda_1 = 1$ and $\lambda_2 = 3$. dim$K_1 = 3$ and dim$k_2 = 5$, we need to find all of the possible partitions of 3 and all of the possible partitions of 5. $1+1+1, 1+2, 3$. Similarly for $5 : 1+1+1+1+1, 1+3+2, 1+4, 5, 2+2+1$, we get seven different partitions. Anyone of the first three can be combined with anyone of the second 3.

Definition 13.6.1: Minimal Polynomial

Fix $T : V \longrightarrow V$ linear. Cayley Hamilton gives us that $P_T(T) = 0$. The characteristic is a polynomial of degree n where n is the dimension. Consider the set $S = \{n \geqslant 1$ such that $\exists f(t) \in F[t]$ such that $f(T) = 0\}$. Cayley Hamilton gives us that $S \neq \emptyset$, because the characteristic polynomial satisfies. S has a smallest element and some polynomial needs to give us this degree and this is what we call the minimal polynomial.

Definition 13.6.2

A polynomial $M_T(t)$ is called a minimal polynomial for T, if:

1. $\mu_T(T) = 0$
2. μ_T is of smallest possible degree
3. μ_T is monic

13.7 Thursday December 05

Note:-

Minimal Polynomial about diagonalizability

Definition 13.7.1: Minimal Polynomial

$T : V \to V$ linear. $\mu(t) \in F[t]$ is a minimal polynomial for T if:

1. degree $\mu \geqslant 1$
2. $\mu(T) = 0$
3. μ has the smallest possible degree
4. $\mu(t)$ is monic
5. $\mu(t) = t^m + a_1 t^{m-1} + ... + a_n t^0$

Note:-

Our goal will be to show that $\mu(t)$ is unique.

Theorem 13.7.1

Let $f(t) \in F[t]$ be an arbitrary polynomial such that $f(T) = 0$. Then the conclusion is that $\mu(t)$ is a factor of $f(T)$.

Proof: We will do polynomial division of $f(t)$ with $\mu(t)$. Do long division of $f(t)$ by $\mu(t)$. This gives us that $\exists q(t), r(t) \in F[t]$ such that $f(t) = \mu(t)q(t) + r(t)$. and $r(t) = 0$ or that degree $r(t) <$ degree $\mu(t)$. It is enough to show that $r(t) = 0$.

Suppose towards contradiction. That $r(t) \neq 0$. This implies that the degree of $r(t) <$ degree of $\mu(t)$. Evaluate $f(t) = \mu(T)q(T) + r(T)$ But this gives us that $0 = 0 + r(T)$, which contradictions the minimality of the degree of μ 😄

Theorem 13.7.2 $\mu_T(t)$ is unique

Proof: Let $\mu(t)$ and $\mu'(t)$ both satisfying properties 1-4 above. $\mu(t)$ is a factor of $\mu'(t)$ implies that $\exists q(t) \in F[t]$ such that $\mu'(t) = \mu(t)q(t)$ and the degree of $q(t)$ is bigger or equal to zero, and the other way around is true: $\exists B(t) \in F[t]$ such that $\mu(t) = \mu'(t)B(t)$. The only possibilities is if these guys are constants. In particular this implies that $\mu'(t) = c(\mu(t)$ for some $c \in F$. But because they are both monic, $c = 1$.

😄

Theorem 13.7.3

$\mu_T(t)$ is a factor of the characteristic polynomial $P_T(t)$.

If $P_T(t) = (\lambda_1 - t)^{k_1} \cdots (\lambda_1 - t)^{k_r}$ then $\mu_T(t) = (t - \lambda_1)^{k_1} \cdots (t - \lambda_r)^{k_r}$ for some $0 \leqslant l_i \leqslant k_i$ for $i = 1, ..., r$.

Note:-

We will show that $(t - \lambda_1)(t - \lambda_2)...(t - \lambda_r)$ always is a factor of $\mu_T(t)$.

Example 13.7.1

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

$P_A(t) = (1-t)^3$, but $\mu_A(t) = 1-t$

Example 13.7.2

$J = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$ $\mu_A(t) = (t-\lambda)^3$, $(t-\lambda_i)$ to the largest size of the jordan block.

To see if it is minimal check $(A-\lambda I)^n = 0$?

Theorem 13.7.4

A Scalar $\lambda \in F$ is a root of $\mu_T(t) \leftrightarrow \lambda$ is a roof of $P_T(t) \leftrightarrow \lambda$ is an eigenvalue.

Proof: $(\rightarrow)$ Suppose λ is a root of μ, then $\exists q(t) \in F[t]$ such that $P_T(t) = \mu_T(t)q(t)$. So if we evaluate at $\lambda \rightarrow P_T(\lambda) = \mu_T(\lambda)q(\lambda) = 0$, this implies that λ is an eigenvalue.
$(\leftarrow)$ Let λ be an eigenvalue of T, let $v \in V$ be a corresponding eigenvector.
$\mu_T(T) = 0 \rightarrow \mu_T(T)(w) = 0, \forall w \in V$. In particular $\mu_T(T)(v) = 0$.
Recall that if v is an eigenvector, then $T(v) = \lambda v$. Which implies that $f(T)(v) = f(\lambda)V, \forall f(t) \in F[t]$. Evaulated at T implies $\mu_T(T)(v) = \mu_T(\lambda)v = 0 \rightarrow \mu_T(\lambda) = 0$.

😃

Theorem 13.7.5 Criterion for Diagonalizability

$T : V \rightarrow V$ is diagonalizable $\leftrightarrow \mu_T(t)$ is a product of distinct linear factors, meaning $\mu_T(t)$ is the smallest possible.

Proof: Suppose that T is diagonalizable. We want to show that $\mu_T(t) = (t-\lambda)\cdots(t-\lambda_r)$. It suffices to show that $f(t) = (t-\lambda_1)\cdots(t-\lambda_r)$, it suffices to show that $f(T) = 0$.
T diagonalizable means that $\exists B = \{v_1, ..., v_n\}$ basis of V consisting of eigenvectors. We want to show that a certain linear transformation is zero. It is enough to show that it is zero for each of the basis elemenets.
$f(T)(v_i) = 0$ for $i = 1, ..., n$. Let v_i have corresponding eigenvalue λ_j.
$f(T) = (T-\lambda_1 I)\cdots(T-\lambda_r I)$. The left sides commute.
$f(T)(v_i) = g(T)((T-\lambda I)(v_i))$ is equal to zero. On the non right side.
$(\leftarrow)$ Now suppose that $\mu_T(t) = (t-\lambda_1)\cdots(t-\lambda_r)$, and we want to show that T is diagonalizable.
This means that we need to show that there exists a basis of eigenvectors.
We will do induction on n = dimension of V. Of course if V is one dimensional, then everything is diagonalizable. $V = E_\lambda$, so diagonalizable.
Suppose that $n \geqslant 1$, and suppose that if $T : W \rightarrow W$ with dimension $W < n$, and μ_T = product of distinct linear factors $\rightarrow T$ is diagonalizable.
$T : V \rightarrow V$, dimension of $V = n$. Take W =image$(T-\lambda_r I)$.
λ_r eigenvalue of $T \rightarrow E_{\lambda r}$ =ker$(T-\lambda_r I) \neq 0$ which implies that $T-\lambda_r I$ is not invertible which implies not surjective. This implies that dimension of W is strictly smaller than n.
Easy check: W is T invariant. Because of this We can restrict $T_W : W \rightarrow W$ from a space now of strictly smaller dimension.
Consider $f(t) = (t-\lambda_1)\cdots(t-\lambda_r)$, the claim is that $f(T_w) = 0$.
Let $w \in W$ =Im$(T-\lambda_r I) \rightarrow w = (T-\lambda_r I)(v)$ for some $v \in V$.
When I apply this polynomial $f(T)(w) = (T-\lambda_1 I)\cdots(T-\lambda_r)(v) = \mu_T(T)(v) = 0$.
Our conclusion is the following that $\mu_{T_W}(t)$ is a factor of $(t-\lambda_1)\cdots(t-\lambda_{r-1}) \rightarrow$ the induction hypothesis applies for $T_W : W \rightarrow W$. This gives us that T_W is diagonalizable. We did this extra work to also imply that λ_r is not an eigenvalue of T_W. This is because if it were, it should appear as a factor in the minimal polynomial of T_W but it doesn't. This gives us that $W \cap E_{\lambda_r} = \{0\}$, because if it were more λ_r would be an eigenvalue of T_W.
In particular, $W \oplus E_{\lambda_r}$ is direct. By the dimension theorem, this direct sum must equal V. Now we are done because we can append our two bases of eigenvectors.

😃

Example 13.7.3

Let $A \in M_{2024}(\mathbb{C})$ such that $A^4 = I$. Then A is diagonalizable. No matter how it looks like. ***Solution:*** $A^4 - I = 0$. This gives me the information that $\mu_A(t)$ divides $t^4 - 1$. For $A \in M_{2024}(\mathbb{R})$ we get possible $\mu_A(t)$:

$$t-1, t+1, t^2-1, t^2+1, (t-1)(t^2+1), (t+1)(t^2+1).$$

Question 13: 1B

$T : V \to V$ $W \leqslant V$ T-invariant subspace. Assume that P_{T_W} splits. The question is, does $P_T(t)$ split. ***Solution:*** No. Look for $A \in M_3(\mathbb{R})$ with $P_A(t) = (t^2+1)(t-3)$. We want to construct a matrix with. 3 is an eigenvalue of A and dimension of $E_3 = 1$. $E_3 = \text{span}\{v\}$ and this E_3 will be my W. The characteristic of $T_w = 3 - t$, W is one dimensional it is the span of a single vector this means that $T_W : \text{span}\{v\} \to \text{span}\{v\}$. $[T_W]_{\{v\}}$