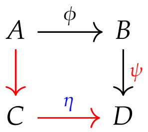
Math 354
Notes

Charlie Cruz

Contents



Information about the class: Math 354 Honors Linear Algebra.
Professor Varilly-Alvarado (Dr. V.) with his email being varilly@rice.edu.
Office: HBH 222
Office Hours: Monday 3:30 – 5:00 pm, F 3-4PM (to be confirmed)
We will be using the book *Linear Algebra Done Right* by Sheldon Axler.

Notation	Definition
$\mathbb{N}$	The set of natural numbers = $\{1, 2, 3, 4, \ldots\}$
$\mathbb{Z}$	The set of integers = $\{\ldots, -2, -1, 0, 1, 2 \ldots\}$
$\mathbb{Q}$	The set of rational numbers = $\{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\}$
$\mathbb{R}$	The set of real numbers
$\mathbb{C}$	The set of complex numbers = $\{a + bi \mid a, b \in \mathbb{R}, i^2 = -1\}$
$\in$	The symbol $\in$ means "is an element of" or "belongs to"
$\notin$	The symbol $\notin$ means "is not an element of" or "does not belong to"
$\subset$	The symbol $\subset$ means "is a subset of"
$\subseteq$	The symbol $\subseteq$ means "is a subset of or equal to"
$\cap$	The symbol $\cap$ means "intersection of"
$\cup$	The symbol $\cup$ means "union of"
$\setminus$	The symbol $\setminus$ means "set difference of"
$\emptyset$	The symbol $\emptyset$ means "the empty set"
$\forall$	The symbol $\forall$ means "for all"
$\exists$	The symbol $\exists$ means "there exists"
$\mid$	The symbol $\mid$ in $\{a \mid a \in \mathbb{R}\}$ means "such that"
$\implies$	The symbol $\implies$ means "implies"
$\iff$	The symbol $\iff$ means "if and only if"
$\vec{a}$	The symbol $\vec{a}$ means "the vector a"

Mathematical Induction: Set of Natural Numbers

(a) $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$ Natural Numbers

(b) $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ Integers

Mathematical Induction is a technique of proof that allows you to verify statements indexed by $\mathbb{N}$ or a subset of $\mathbb{Z}$.

Example 0.0.1
For all $n \in \mathbb{N}$, it is true that:

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$$

e.g.: $n = q \quad 1 = \frac{1 \cdot 2}{2}$, and so on for $n = 2, 3, \ldots$.

Induction: $P(n)$ is a statement depending on $n \in \mathbb{N}$.
e.g. "$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$"
Now suppose that

(i) $P(1)$ is true (base case)

(ii) If $P(k)$ is true for some $k \in \mathbb{N}$, then $P(k+1)$ is true (inductive step)

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Note:-

Think about this as a domino effect.

Example: Let $P(n) : 1 + 2 + \ldots + n = \frac{n(n+1)}{2}$.

(i) $P(1)$ is true because $1 = \frac{1(1+1)}{2}$.

(ii) Suppose $P(k)$ is true for some $k \in \mathbb{N}$. Then

$$\begin{aligned} 1 + 2 + \ldots + k + (k+1) &= \frac{k(k+1)}{2} + (k+1) \\ &= \frac{k(k+1) + 2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2} \\ &= \frac{(k+1)((k+1)+1)}{2} \\ &= \frac{n(n+1)}{2} \text{ where } n = k+1 \end{aligned}$$

Therefore, $P(k+1)$ is true.

😆

Note:-

Baby Version Let $A \subseteq \mathbb{N}$ be a subset. Suppose that

(i) $1 \in A$

(ii) If $k \in A$, then $k + 1 \in A$

Then $A = \mathbb{N}$
Baby version $\implies$ PMI (let $A \, \{n \in \mathbb{N} \colon p(n) \text{ is true}\}$)

Example 0.0.2 (Same proof in two different ways)
Let $p(n)$: $\frac{1}{1*2} + \frac{1}{2*3} + \frac{1}{n(n+1)} = \frac{n}{n+1}$ for all $n \in \mathbb{N}$.

Proof 1.0: For $n \in \mathbb{N}$, let $p(n)$ be the statement:

$$\frac{1}{1*2} + \frac{1}{2*3} + \frac{1}{n(n+1)} = \frac{n}{n+1}.$$

We show that $p(n)$ is true for all $n \in \mathbb{N}$ by induction.

Base Case: $p(1)$ is true because $\frac{1}{1*2} = \frac{1}{2} = \frac{1}{1+1}$.

Inductive Step: Assume that $p(k)$ is true for some $k \in \mathbb{N}$. Then

$$\frac{1}{1*2} + \frac{1}{2*3} + \ldots + \frac{1}{k(k+1)} = \frac{k}{k+1} \quad (1)$$

We want to deduce that $p(k+1)$ is true i.e.

$$\frac{1}{1*2} + \frac{1}{2*3} + \ldots + \frac{1}{k(k+1)} + \frac{1}{(k+1)(k+2)} = \frac{k+1}{k+2} \quad (2)$$

To do this, add $\frac{1}{(k+1)(k+2)}$ to both sides of (1).

$$\frac{1}{1*2}+\frac{1}{2*3}+\ldots+\frac{1}{k(k+1)}+\frac{1}{(k+1)(k+2)}=\frac{k}{k+1}+\frac{1}{(k+1)(k+2)}.$$

The RHS of this equation is:

$$\frac{k}{k+1}+\frac{1}{(k+1)(k+2)}=\frac{k(k+2)+1}{(k+1)(k+2)}=\frac{k^2+2k+1}{(k+1)(k+2)}=\frac{(k+1)^2}{(k+1)(k+2)}=\frac{k+1}{k+2}$$

This shows that $p(k+1)$ true, completing the induction step.
Therefore, $p(n)$ is true for all $n \in \mathbb{N}$. 😊

Proof 2.0: We prove the statement by induction on n.
The base case, when $n = 1$, is true because $\frac{1}{1*2} = \frac{1}{2} = \frac{1}{1+1}$.
For the inductive step, assume the claim is the true for some $k \in \mathbb{N}$.

$$\frac{1}{1*2}+\frac{1}{2*3}+\ldots+\frac{1}{k(k+1)}=\frac{k}{k+1}$$

Add $\frac{1}{(k+1)(k+2)}$ to both sides of the equation to obtain

$$\begin{aligned}\frac{1}{1*2}+\frac{1}{2*3}+\ldots+\frac{1}{k(k+1)}+\frac{1}{(k+1)(k+2)} &= \frac{k}{k+1}+\frac{1}{(k+1)(k+2)}\\ &= \frac{k(k+2)+1}{(k+1)(k+2)}\\ &= \frac{k^2+2k+1}{(k+1)(k+2)}\\ &= \frac{(k+1)^2}{(k+1)(k+2)}\\ &= \frac{k+1}{k+2}\end{aligned}$$

Then, the claim is true for $k+1$, completing the inductive step.
We deduce that the claim is true for all $n \in \mathbb{N}$ by induction. 😊

Chapter 1

Vector Spaces

1.1 Fields

Definition 1.1.1: Fields or "sets of scalars"

We have a lot of experience with this, in fact:

(i) $\mathbb{Q} = \left\{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\right\}$ Rational Numbers

(ii) $\mathbb{R}$, the real numbers: $\mathbb{Q} \subseteq \mathbb{R}$ e.g. $\sqrt{2} \notin \mathbb{Q}$

(iii) $\mathbb{C}$, the complex numbers: $i^2 + 1 = 0$

A field is a set F with two operations $+$ and $\cdot$ such that:

(i) Associativity: for all $a, b, c \in F$, $a + (b + c) = (a + b) + c$ and $a(bc) = (ab)c$

(ii) Commutativity: for all $a, b \in F$, $a + b = b + a$ and $ab = ba$

(iii) Identity: there are elements 0 and 1 in $\mathbb{F}$ such that for all $a \in F$, we have: $a + 0 = a$ and $a \cdot 1 = a$

(iv) Inverses: for all $a \in F$, there is an element b such that $a+b = 0$. We write $b = -a$ and $x-y: = x+(-y)$. Additionally, for any $a \neq 0$ in F, there is an element $b \neq 0$ such that $ab = 1$. We write $b = a^{-1}$.

(v) Distributivity: for all $a, b, c \in F$, we have $a(b + c) = ab + ac$

(vi) $0 \neq 1$

Example 1.1.1

$F_2 = \{0, 1\}$ is a field.
If we write the addition table:

+	0	1
0	0	1
1	1	0

Now for the multiplication table:

$\cdot$	0	1
0	0	0
1	0	1

Example 1.1.2

Let's define $\mathbb{F} = \{\triangle, \square, \circ\}$
With the addition and multiplication tables as followed:

$+$	$\triangle$	$\square$	$\circ$
$\triangle$	$\triangle$	$\square$	$\circ$
$\square$	$\square$	$\circ$	$\triangle$
$\circ$	$\circ$	$\triangle$	$\square$

$\cdot$	$\triangle$	$\square$	$\circ$
$\triangle$	$\triangle$	$\triangle$	$\triangle$
$\square$	$\triangle$	$\square$	$\circ$
$\circ$	$\triangle$	$\circ$	$\square$

$\mathbb{F}$ is a field.
With $o_{\mathbb{F}} = \triangle$ and $1_{\mathbb{F}} = \square$.
Now, "2_F" $= 1_{\mathbb{F}} + 1_{\mathbb{F}} = \square + \square = \circ$.
Let's also define, $-\square = \circ$.

Theorem 1.1.1 The additive identity of a field $\mathbb{F}$ is unique.

Proof: Suppose $0_F, 0'_F$ are additive identities of $\mathbb{F}$.
Then:

$$0_F \overbrace{=}^{\text{Because } O_F\text{` is an additive identify}} 0_F + 0'_F \underbrace{=}_{\text{Beacuse } O_f \text{ is an additive identity}} 0'_F$$

Thus, the additive identity is unique.

😊

Theorem 1.1.2
Let $\mathbb{F}$ be a field and let $a \in \mathbb{F}$. Then $a \cdot O_F = 0_F$.

Note:-
Don't think of zero is nothing, think of its meaning and how important it is to a field

Proof: Let $-a \cdot 0_F$ be the additive inverse of $a \cdot 0_F$.
Then:

$$\begin{aligned}
0_F + 0_F &= 0_F \quad \text{as } 0_F \text{ is an additive identity} \\
a \cdot (0_F + 0_F) &= a \cdot 0_F \\
a \cdot 0_F + a \cdot 0_F &= a \cdot 0_F \quad \text{by distributivity} \\
(a \cdot 0_F + a \cdot 0_F) + (-a \cdot 0_F) &= a \cdot 0_F + (-a \cdot 0_F) \\
(a \cdot 0_F + a \cdot 0_F) + -a \cdot 0_F &= 0_F \quad \text{by additive inverse} \\
a \cdot 0_F + (a \cdot 0_F + -a \cdot 0_F) &= 0_F \quad \text{by associativity} \\
a \cdot 0_F + 0_F &= 0_F \quad \text{by additive inverse} \\
a \cdot 0_F &= 0_F \quad \text{by additive identity}
\end{aligned}$$

😊

Theorem 1.1.3
Let $\mathbb{F}$ be a field and let $a \in \mathbb{F}$. Then $-(-a) = a$.

Known: Additive inverse are unique.

$$(-a) + a = 0_F$$

This says that a is an additive inverse of $-a$.
Additive inverses are unique, so a must be the additive inverse of $-(-a)$.

Note:-

you can try this at home:

$$(-1_F) \cdot (a) = -a$$

Where -1 is the additive inverse of 1, and $-a$ is the additive inverse of a.
Hint: $(-1) + 1 = 0_f$

Building $\mathbb{C}$ out of $\mathbb{R}$: A complex number is an ordered pair (a, b) of real numbers.

$$\mathbb{C} = \{(a, b) \colon a, b \in \mathbb{R}\}$$

(i) Addition: $(a, b) + (c, d) \colon = (a + c, b + d)$, where $(a, b) = z_1$ and $(c, d) = z_2$. As such, we use $\mathbb{R}$ addition to define $\mathbb{C}$ addition.

(ii) Multiplication: $(a, b) \cdot (c, d) \colon = (ac - bd, ad + bc)$, where $(a, b) = z_1$ and $(c, d) = z_2$

Note:-

You might want to think of (a, b) as $a + bi$, where $i^2 = -1$. As such:

$$(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i$$

(iii) Additive identity: $(0_{\mathbb{R}}, 0_{\mathbb{R}}) = 0_{\mathbb{C}}$.

(iv) Multiplicative identity: $(1_{\mathbb{R}}, 0_{\mathbb{R}}) = 1_{\mathbb{C}}$.

Note:-

We can check that $i \colon = (0_{\mathbb{R}}, 1_{\mathbb{R}})$.
Now, $(0_{\mathbb{R}}, 1_{\mathbb{R}}) \underbrace{\cdot}_{\mathbb{C}} (0_{\mathbb{R}}, 1_{\mathbb{R}}) = (-1_{\mathbb{R}}, 0_{\mathbb{R}}) = -(1_{\mathbb{R}}, 0_{\mathbb{R}}) = -1_{\mathbb{C}}$.

Definition 1.1.2: Lists T tuples

Let F be a field (think: $F_2, F_3, \mathbb{Q}, \mathbb{R}, \mathbb{C}$).
Then we will have a list of length n that they are ordered $x_1, \dots, x_n, x_i \in \mathbb{F}$.
Remember order matters! Note $(2,3) \neq (3,2)$
Let's define: $\mathbb{F}^n = \{(x_1, \dots, x_n) \colon x_i \in \mathbb{F}, i = 1, \dots, n\}$. For instance $\mathbb{R}^2 = \{(x,y) \colon x, y \in \mathbb{R}\}$
Sometimes we write $\underline{x}$ or $\vec{x} \in \mathbb{F}^n$ for $(x_1, \dots, x_n)$.
F^n is the archetype of a "finite-dimensional vector space".
This means that the following properties hold:

(i) $\vec{x} + \vec{y} = (x_1, \dots, x_n) + (y_1, \dots, y_n) \coloneqq (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$. Note that we are using the addition of $\mathbb{F}$ to define the addition of $\mathbb{F}^n$.

(ii) Addition has a neutral element: $\vec{0} = (0_{FF}, \dots, 0_{FF})$, n times. Thus:

$$\begin{aligned} \vec{x} + \vec{0} &= (x_1, \dots, x_n) + (0_{FF}, \dots, 0_{FF}) \\ &= (x_1 + 0_{FF}, \dots, x_n + 0_{FF}) \\ &= (x_1, \dots, x_n) \\ &= \vec{x} \end{aligned}$$

There are additive inverses: if $\vec{x} = (x_1, \dots, x_n)$ then setting $-\vec{x} = (-x_1, \dots, -x_n)$ we get

$$\vec{x} + (-\vec{x}) = \vec{0}$$

(iii) Elements of $\mathbb{F}^n$ can be scaled by elements of $\mathbb{F}$. If $\vec{x} = (x_1, \dots, x_n)$ and $\lambda \in \mathbb{F}$, then $\lambda \cdot \vec{x} = (\lambda \cdot x_1, \dots, \lambda \cdot x_n)$.

Note:-

Warning! In general, elements of $\mathbb{F}^n$ cannot be multiplied with each other unless we define a multiplication operation on $\mathbb{F}^n$.

1.2 Vectors Spaces

Definition 1.2.1: Vector Space in general

Let F be a field, where $F = (F, +_F, \cdot_F)$.
A vector space over $\mathbb{F}$ is a set V together with two operations:
Define Addition of Vectors as

$$+ : V \times V \mapsto V, (u, v) \mapsto u + v$$

And scalar multiplication as

$$\cdot : \mathbb{F} \times V \mapsto V, (\lambda, v) \mapsto \lambda \cdot v$$

These operations satisfy the following properties:

(i) Commutativity: $u + v = v + u$ for all $u, v \in V$

(ii) Associativity of addition: $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$. Also:

$$(\lambda_1^{\mathbb{F}} \cdot \lambda_2^{\mathbb{F}}) \cdot_V v = \lambda_1^{\mathbb{F}} \cdot_V (\lambda_2^{\mathbb{F}} \cdot_V v) \quad \text{for all } \lambda_1, \lambda_2 \in \mathbb{F} \text{ and } v \in V$$

(iii) Additive identity: there is a vector $0_V \in V$ such that $0_V + v = v$ for all $v \in V$.

(iv) Additive inverse: for every $v \in V$, there is a vector $-v \in V$ such that $v + (-v) = 0_V$.

(v) Scalar Multiplicative identity: $1_{\mathbb{F}} \cdot v = v$ for all $v \in V$.

(vi) Distributivity: $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$ for all $\lambda \in \mathbb{F}$ and $u, v \in V$. Also, $(\lambda_1^{\mathbb{F}} + \lambda_2^{\mathbb{F}}) \cdot v = \lambda_1^{\mathbb{F}} \cdot v + \lambda_2^{\mathbb{F}} \cdot v$ for all $\lambda_1, \lambda_2 \in \mathbb{F}$ and $v \in V$.

Note:-

If $\mathbb{F} = \mathbb{R}$, call V a real vector space.
If $\mathbb{F} = \mathbb{C}$, call V a complex vector space.

Summary: To specify a vector space, we need 4 pieces of data:

(a) V - the set of vectors

(b) $\mathbb{F}$ - the "numbers that we can scale by"

(c) $+_V$ - addition of vectors

(d) $\cdot_V$ - scalar multiplication

For a while, we will write $(V, \mathbb{F}, +, \cdot)$ for all this data.
In fact, $(V, \mathbb{F}, +, \cdot) = (V, (\mathbb{F}, +_{\mathbb{F}}, \cdot_{\mathbb{F}}), +_V, \cdot_V)$.
For instance. Take a field $\mathbb{F}$, $(\mathbb{F}^n, \mathbb{F}, +, \cdot)$.
Now given $\vec{x}, \vec{y} \in \mathbb{F}^n, \lambda \in \mathbb{F}$. We can define:

$$\begin{aligned}
\vec{x} + \vec{y} &= (x_1, \ldots, x_n) + (y_1, \ldots, y_n) \\
&= (x_1 + y_1, \ldots, x_n + y_n) \\
\lambda \cdot \vec{x} &= \lambda \cdot (x_1, \ldots, x_n) \\
&= (\lambda \cdot x_1, \ldots, \lambda \cdot x_n)
\end{aligned}$$

Example 1.2.1

(i) (a) $(V, \mathbb{F}, +, \cdot) = (\mathbb{R}^2, \mathbb{R}, +, \cdot)$ is a real vector space.

(b) $(x_1, y_1) +_{\mathbb{R}^2} (x_2, y_2) = (x_1 +_{\mathbb{R}} x_2, y_1 +_{\mathbb{R}} y_2)$.

(c) $\lambda \in \mathbb{R}$, $\lambda \cdot^{\mathbb{R}^2} (x, y) = (\lambda \cdot^{\mathbb{R}} x, \lambda \cdot^{\mathbb{R}} y)$.

(ii) (a) $(V, \mathbb{F}, +, \cdot) = (\mathbb{C}^2, \mathbb{C}, +_{\mathbb{C}^2}, \cdot_{\mathbb{C}^2})$ is a complex vector space.

(b) $z_1, z_2 +_{\mathbb{C}^2} (Z_1', Z_2') = (z_1 +_{\mathbb{C}} z_1', z_2 +_{\mathbb{C}} z_2')$.

(c) $\lambda \in \mathbb{C}$, $\lambda \cdot^{\mathbb{C}^2} (z_1, z_2) = (\lambda \cdot^{\mathbb{C}} z_1, \lambda \cdot^{\mathbb{C}} z_2)$.

(iii) (a) $(V, \mathbb{F}, +, \cdot) = (\mathbb{C}^2, \mathbb{R}, +_{\mathbb{R}^2}, \cdot_{\mathbb{R}^2})$ is a complex vector space, but we are using real numbers to scale.

(b) Addition is the same as (ii), but $z_1, z_2 \in \mathbb{C}^2$ i.e. $a + bi$

(c) Scalar multiplication: $\lambda \in \mathbb{R}, \lambda \cdot^{\mathbb{R}^2} (z_1, z_2) = (\lambda \cdot^{\mathbb{R}} z_1, \lambda \cdot^{\mathbb{R}} z_2)$.

(iv) $(V, \mathbb{F}, +, \cdot) = (\mathbb{F}^n, \mathbb{F}, +, \cdot)$ is a vector space.

(v) Let F be a field, and S be a set. Let $V = \mathbb{F}^s := \{\text{functions } f : S \mapsto \mathbb{F}\}$.

(i) Addition: $f, g \in V$, $f : S \mapsto F$, and $g : S \mapsto F$. Then $(f + g) : S \mapsto F$ is defined by $(f + g)(s) = f(s) + g(s)$ for all $s \in S$. Or $s \mapsto f(s) + g(s)$

(ii) Scalar Multiplication: $\lambda \in F, f \in V$, $f : S \mapsto F$. We need to show that $\lambda\mathbb{F} \in V$ i.e. $\lambda\mathbb{F} : S \mapsto F$, where $s \longrightarrow \lambda f(s)$. Also $(\lambda f)(s) = \lambda f(s)$ for all $s \in S$.

(iii) Additive identity: $\vec{0}_V : S \longrightarrow \mathbb{F}$, $s \longrightarrow 0_F$.
Check: $(f + \vec{0}_V)(s) = f(s) + \vec{0}_V(s) = f(s) + 0_F = f(s)$ for all $s \in S$.

Now, we want to talk about its relationship to $\mathbb{F}^n$. Take $S = \{1, 2, \ldots, n\}$

Now, let $V = \mathbb{F}^{\{1,\ldots,n\}} = \{\text{functions } f : \{1, \ldots, n\} \mapsto \mathbb{F}\}$.

We can create a function $F^{\{1,\ldots,n\}} \mapsto \mathbb{F}^n$ by:

$$f : \{1, \ldots, n\} \mapsto \mathbb{F} \longrightarrow (f(1), \ldots, f(n))$$

Note:-
This is a bijection!

Example 1.2.2

Let $S = [0, 1]$ and $F = \mathbb{R}$.
Now set $V = \mathbb{F}^s = \mathbb{R}^{[0,1]} = \{\text{functions } f : [0, 1] \longrightarrow \mathbb{R}\}$.
Another Example.
Let $\mathbb{F} = \mathbb{R}$.
Now $V = \{\text{polynomials of degree } \leq 19 \text{ with coefficients in } \mathbb{R}\}$.
Now, $+_V =$ usual addition of polynomials and $\cdot_V =$ usual scalar multiplication of polynomials.
For instance,

$$x^{19} + x + 1 \in V$$
$$-x^{19} + x^{17} - x^2 \in V$$
$$9 * (x^2 + 2) = 9x^2 + 18 \in V$$

Note:-
Sometimes we denote that the degree of 0 (the zero polynomial) is $-\infty$.

First properties of vector spaces: Let V be a vector space over a field $\mathbb{F}$.

(i) Additive identities are unique. Suppose $\vec{0}, \vec{0}' \in V$ are additive identities. Then $\vec{0} = \vec{0} + \vec{0}' = \vec{0}'$.

(ii) Additive inverses are unique:
Say w, w' are additive inverse of $v \in V$.
$w = w + \vec{0} = w + (v + w') = (w + v) + w' = \vec{0} + w' = w'$.

(iii) $0 \cdot v = \vec{0}, \forall v \in V$.

$$\begin{aligned}
0_F &= 0_F + 0_F \\
&\implies 0_F \cdot v = (0_F + 0_F) \cdot v \\
&\implies 0_F \cdot v = 0_F \cdot v + 0_F \cdot v \\
0_F \cdot v + (-0_F \cdot v) &= (0_F \cdot v + 0_F \cdot v) + (-0_F \cdot v) \\
\vec{0} &= 0_F \cdot v + (-0_F \cdot v + 0_F \cdot v) \\
\vec{0} &= 0_F \cdot v + \vec{0} \\
\vec{0} &= 0_F \cdot v
\end{aligned}$$

1.3 Subspaces

Definition 1.3.1: Subspaces

Let $(V, \mathbb{F}, +, \cdot)$ be a vector space. A subset $U \subseteq V$ is a subspace if $(U, \mathbb{F}, +_{u \in U}, *_{u \in U})$ is a vector space in its own right.

Example 1.3.1
Let $V = \mathbb{R}^3$ and $\mathbb{F} = \mathbb{R}$.
$$U = \{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\} \nsubseteq \mathbb{R}^3 = V$$

1.34 in book / conditions for a subspace: To check that $U \subseteq V$ is a subspace, it is enough to check:

(i) $\vec{0} \in U$

(ii) U is closed under addition: if $u, v \in U$ then $u + v \in U$

(iii) U is closed under scalar multiplication: if $u \in U$ and $\lambda \in \mathbb{F}$ then $\lambda \cdot u \in U$

Reason: These three conditions ensure that U has an additive identity vector, and that addition and scalar multipliation makes sense in U.
The remaining axioms for U to be a vector space are inherited from V.

Example 1.3.2
Let's check associativity of addition: Let $u, v, w \in U$.
But we know that $u, v, w \in V$ as $U \subseteq V$, so $u + (v + w) = (u + v) + w (\star)$ in V.
Since U is closed under addition, $u + v \in U$.
Again, since $u + v \in U$, and $w \in U$, we know that $(u + v) + w \in U$.
Likewise, $u + (v + w) \in U$. This means that $(\star)$ is also true in U.
Ditto for the other axioms. Thus, we would be proving the same thing twice.

Example 1.3.3 (Charlie add the graphs)
$V = \mathbb{R}^2$

(i) $U = \{(a, a,) : a \geqslant 0\}$ is not closed under scalar multiplication.

(ii) $U = \{(a, a) : a \in \mathbb{R}\} \cup \{(-a, a) : a \in \mathbb{R}\}$ is not closed under addition.

(iii) $U = \{(a, a + a) : a \in \mathbb{R}\}$ does not contain the additive identity of $\mathbb{R}^2$

Example 1.3.4
Let $\mathbb{F} = \mathbb{R}$ and $V = \mathbb{R}^{(0,3)} = \{\text{functions } f : (0, 3) \to \mathbb{R}\}$.
Let $U \subseteq V$ be the subset $\{\text{functions } f : (0, 3) \to \mathbb{R} \mid f \text{ is differentiable and } f'(2) = 0\}$.

Proof: Let's check that $U \subseteq V$ is a subspace:

(i) Show that $\vec{0}_V \in U$: $\vec{0}_V$ is $\vec{0}_V : (0, 3) \to \mathbb{R}, x \mapsto 0_{\mathbb{R}}$.
$\vec{0}_V$ is differentiable and $\vec{0}_V{}'(2) = 0$.

(ii) Show that U is closed under addition: Let $f, g \in U$. We need to show that $f + g \in U$.
This means that both $f : (0, 3) \to \mathbb{R}$ and $g : (0, 3) \to \mathbb{R}$ are differentiable, and that $(f + g)'(2) = 0$.
Then $f + g : (0, 3) \to \mathbb{R}$ is differentiable as both f and g are differentiable.
Moreover, $(f + g)'(2) = f'(2) + g'(2) = 0 + 0 = 0$.
Thus, $f + g \in U$.

(iii) Show that U is closed under scalar multiplication: Let $f \in U$ and $\lambda \in \mathbb{R}$. We need to show that $\lambda \cdot f \in U$.
This means that $f : (0, 3) \to \mathbb{R}$ is differentiable and that $(\lambda \cdot f)'(2) = 0$.
Then $\lambda \cdot f : (0, 3) \to \mathbb{R}$ is differentiable as f is differentiable.
Moreover, $(\lambda \cdot f)'(2) = \lambda \cdot f'(2) = \lambda \cdot 0 = 0$.
Thus, $\lambda \cdot f \in U$.

All three conditions are satisfied, so $U \subseteq V$ is a subspace. ☺

Definition 1.3.2: Sums of Subsets

Let $(V, \mathbb{F}, +, \cdot)$ be a vector space over a field $\mathbb{F}$. Let $U, W \subseteq V$ be subsets.
Let $U_1, \ldots, U_m$ be subsets of V.
Where

$$U_1, \ldots, U_m = \{V_1 + V_2 + \ldots + V_m : V_i \in U_i \text{ for all } i = 1, \ldots, m\}$$

Example 1.3.5
Our field will be $\mathbb{F} = \mathbb{R}$, vectors space will be $V = \mathbb{R}^3$.
Let $U_1 = \{(x, 0, 0) : x \in \mathbb{R}\}, U_2 = \{(0, y, 0) : y \in \mathbb{R}\}$.
Let $U_1 + U_2 = \{V_1 + V_2 : V_1 \in U_1, V_2 \in U_2\}$.
This means that this is equal to $\{(x, 0, 0) + (0, y, 0) : x, y \in \mathbb{R}\} = \{(x, y, 0) : x, y \in \mathbb{R}\}$

Theorem 1.3.1
If $U_1, \ldots, U_m$ are subspaces of V, then $U_1 + \ldots + U_m$ is the smallest subspace of V containing $U_1, \ldots, U_m$.
Have to prove that:

(i) $U_1 + \ldots + U_m$ is a subspace of V (not just a subset).

(ii) $U_1 \subseteq U_1 + \ldots, U_m, U_2 \subseteq U_1 + \ldots + U_m, \ldots, U_m \subseteq U_1 + \ldots + U_m$.

(iii) $U_1 + \ldots + U_m$ is the smallest subspace of V containing $U_1, \ldots, U_m$.

Proof: We are given that each U_i is a subspace, meaning that $\vec{0} \in U_i$, so

$$\vec{0} = \vec{0}_{\in u_1} + \ldots + \vec{0}_{\in U_m} \in U_1 + \ldots + U_m$$

Thus, we have shown that the additive identity is in $U_1 + \ldots + U_m$.
Now, we want to show that this sum is closed under addition.
Let $\vec{v}, \vec{w} \in U_1 + \ldots + U_m$. We need to show that $\vec{v} + \vec{w} \in U_1 + \ldots + U_m$.
Then, $\vec{V} = \vec{V}_1 + \ldots + \vec{V}_m, \vec{W} = \vec{w}_1 + \ldots + \vec{W}_m$
As such

$$\begin{aligned} \vec{w} + \vec{v} &= (\vec{w}_1 + \ldots + \vec{w}_m) + (\vec{v}_1 + \ldots + \vec{v}_m) \\ &= (\vec{w}_1 + \vec{v}_1) + \ldots + (\vec{w}_m + \vec{v}_m) \\ &\in U_1 + \ldots + U_m \end{aligned}$$

Since each U_i is closed under addition.
Now, we want to show that $U_1 + \ldots + U_m$ is closed under scalar multiplication.
Let $\lambda \in \mathbb{F}$ and $\vec{v} \in U_1 + \ldots + U_m$. We need to show that $\lambda \cdot \vec{v} \in U_1 + \ldots + U_m$.
Then,

$$\begin{aligned} \lambda * \vec{V} &= \lambda * (\vec{v}_1 + \ldots + \vec{v}_m) \\ &= (\lambda * \vec{v}_1) + \ldots + (\lambda * \vec{v}_m) \\ &\in U_1 + \ldots + U_m \end{aligned}$$

Since each U_i is closed under scalar multiplication.
Thus, $U_1 + \ldots + U_m$ is a subspace of V.
Now, now we need to show that each U_i is contained in $U_1 + \ldots + U_m$.
Let $u \in U_i$, we want to show that $u \in U_1 + \ldots + U_m$.
Then we can set $\vec{0}_{u_1} + \ldots + \vec{0}_{u_{i-1}} + u + \vec{0}_{u_{i+1}} + \ldots + \vec{0}_{u_m} \in U_1 + \ldots + U_m$.
Obviously, if we set the rest of the vectors to be $\vec{0}$, then we get $u \in U_1 + \ldots + U_m$.
Finally, we want to prove that $U_1 + \ldots + U_m$ is the smallest subspace of V containing $U_1, \ldots, U_m$.
Let X be a subspace of V such that $U_i \in X$ for all $i = 1, \ldots, m$.
We want to show that $U_1 + \ldots + U_m \subseteq X$.
Let $\vec{v} \in U_1 + \ldots + U_m$, so $\vec{v} = \vec{v}_1 + \ldots + \vec{v}_m$ where $\vec{v}_i \in U_i$ for all $i = 1, \ldots, m$.
Since $U_i \subseteq X$ for all $i = 1, \ldots, m$, we know that $\vec{v}_i \in X$ for all $i = 1, \ldots, m$.
Thus, $\vec{v} = \vec{v}_1 + \ldots + \vec{v}_m \in X$ since X is closed under vector addition. 😊

Definition 1.3.3: Direct sum

Let $U_1 + \ldots + U_M$ is a direct sum if
for each $\vec{v} \in U_1 + \ldots U_m$, there is exactly one way to write $\vec{v} = \vec{u_1} + \ldots + u_{\vec{m}}$ with $\vec{u_i} \in U_i$.

Example 1.3.6

Let $U_1 = \{(x, y, 0) \colon x, y \in \mathbb{R}\}$, $U_2 = \{(0, 0, z) \colon z \in \mathbb{R}\}$.

Claim: Then $U_1 + U_2$ is a direct sum.
Let's prove our claim

Proof: Let $\vec{v} \in U_1 + U_2$. We know $\vec{v} = \vec{u_1} + \vec{u_2}$ for some $\vec{u_1} \in U_1, \vec{u_2} \in U_2$.
We want to show: if $\vec{v} = \vec{u_1} + \vec{u_2}$ for any $\vec{u_1} \in U_1, u_2 \in U_2$
then $\vec{u_1} = \vec{u_1}‘$ and $\vec{u_2} = \vec{u_2}‘$.
Now we know that $\vec{u_1} = (x, y, 0)$ and $\vec{u_2} = (0, 0, z)$, and
the same for our primes (i.e. $\vec{u_1}‘ = (x‘, y‘, 0)$ and $\vec{u_2}‘ = (0, 0, z‘)$) for some $x, y, z, x‘, y‘, z‘ \in \mathbb{R}$.

$$\begin{aligned} \vec{v} &= \vec{u_1} + \vec{u_2} = \vec{u_1}‘ + \vec{u_2}‘ \\ &(x, y, 0) + (0, 0, z) = (x‘, y‘, 0) + (0, 0, z‘) \\ &\implies (x, y, z) \qquad = (x‘, y‘, z‘) \\ &\implies \vec{u_1} = (x, y, 0) = (x‘, y‘, 0) = \vec{u_1}‘ \\ &\implies \vec{u_2} = (0, 0, z) = (0, 0, z‘) = \vec{u_2}‘ \end{aligned}$$

😄

Non-example: Let $U_3 = \{(0, y, y) \colon y \in \mathbb{R}\}$.

Claim: Then $U_1 + U_2 + U_3$ is not a direct sum.

Proof: Thus, one way to write the zero vector is as follows:

$$\vec{0} = \vec{0}_{\in u_1} + \vec{0}_{\in u_2} + \vec{0}_{\in u_3} = (0, -1, 0)_{\in u_1} + (0, 0, -1)_{\in u_2} + (0, 1, 1)_{\in u_3}$$

😄

Theorem 1.3.2

$U_1 + \ldots + U_m$ is a direct sum if and only if

$$\vec{0} \text{can be written uniquely as } \vec{0} = \vec{0}_{\in u_1} + \ldots + \vec{0}_{\in u_m}$$

We need to prove it both ways.

Proof of $\implies$ ***:*** If $U_1 + \ldots + U_m$ is direct, then every vector in $\vec{v} \in U_1 + \ldots + U_m$ can be written uniquely as a sum of vectors from $U_1, \ldots, U_m$.
In particular, if $\vec{v} = \vec{0}$, ten we can only write $\vec{0}$ in one way as $\vec{0} = \vec{0}_{\in u_1} + \ldots + \vec{0}_{\in u_m}$.
And we done.

😄

Proof of $\impliedby$ ***:*** Suppose $\vec{0}$ can only be written in one way as

$$\vec{0} = \vec{0}_{\in u_1} + \ldots + \vec{0}_{\in u_m}, u_i \in U_i$$

Let $\vec{v} \in U_1 + \ldots + U_m$ be arbitrary, and suppose

$$\vec{v} = \vec{u_1} + \ldots + \vec{u_m} = \vec{u_1}‘ + \ldots + \vec{u_m}‘$$

We want to show that $\vec{u_i} = \vec{u_i}‘$ for all $i = 1, \ldots, m$.

Then,

$$\vec{0} = \vec{v} - \vec{v} = (\vec{u_1} + \ldots + \vec{u_m}) - (\vec{u_1}` + \ldots + \vec{u_m}`)$$
$$\implies \vec{0} = (\vec{u_1} - \vec{u_1}`)_{\in u_1} + \ldots + (\vec{u_m} - \vec{u_m}`)_{\in u_m}$$
$$\implies \vec{u_1} - \vec{u_1}` = \vec{0}_{\in u_1}, \ldots, \vec{u_m} - \vec{u_m}` = \vec{0}_{\in u_m}$$

And we are done.

😄

Thus, we have shown that $U_1 + \ldots + U_m$ is a direct sum if and only if $\vec{0} = \vec{0}_{\in u_1} + \ldots + \vec{0}_{\in u_m}$ is the only way to write $\vec{0}$ as a sum of vectors from $U_1, \ldots, U_m$.

Alternative proof that $U_1 + U_2$ (from our example) is direct using criterion from our prvious theorem.

Alternative Proof: Let $U_1 = \{(x, y, 0) : x, y \in \mathbb{R}\}$ and $U_2 = \{(0, 0, z) : z \in \mathbb{R}\}$.
If $\vec{0} = \vec{u_1} + \vec{u_2}$ for some $\vec{u_1} \in U_1, \vec{u_2} \in U_2$, then

$$\implies \vec{0} = \vec{u_1} + \vec{u_2} = (x, y, 0) + (0, 0, z) = (x, y, z)$$
$$\implies x = y = z = 0$$
$$\implies \vec{u_1} = (x, y, 0) = (0, 0, 0), \vec{u_2} = (0, 0, z) = (0, 0, 0)$$

😄

Theorem 1.3.3
If U_1, U_2 are subspaces of of a vector space V, then

$$(U_1 + U_2 \text{ is direct}) \iff (U_1 \cap U_2 = \left\{\vec{0}\right\})$$

Proof of $\implies$: Suppose $U_1 + U_2$ is direct, then we want to show that $U_1 \cap U_2 = \left\{\vec{0}\right\}$.
In other words we need to prove subset inclusion in both directions.

$\subseteq$ **:** We have $\{0\} \subseteq U_1 \cap U_2$ since $\vec{0} \in U_1$ and $\vec{0} \in U_2$.

$\supseteq$ **:** Let $\vec{v} \in U_1 \cap U_2$. We want to show that $\vec{v} = \vec{0}$.

$$\implies \vec{v} \in U_1 \text{ and } \vec{v} \in U_2$$
$$\implies -\vec{v} \in U_1 \text{ and } -\vec{v} \in U_2 \text{ as they are closed under scalar multiplication}$$
$$\implies \vec{0} = \vec{v} + (-\vec{v}) \in U_1 + U_2 \text{ by our previous theorem}$$
$$\implies \vec{v} = \vec{0}, -\vec{v} = \vec{0}$$

Thus, $U_1 \cap U_2 = \left\{\vec{0}\right\}$.

😄

Proof of $\impliedby$: Suppose $U_1 \cap U_2 = \left\{\vec{0}\right\}$, then we want to show that $U_1 + U_2$ is direct.
Suppose $\vec{0} = \vec{u_1} + \vec{u_2}$ for some $\vec{u_1} \in U_1, \vec{u_2} \in U_2$.
We want to show that $\vec{u_1} = \vec{u_2} = \vec{0}$.

$$
\begin{aligned}
0 = \vec{u_1} + \vec{u_2} \implies \vec{u_1} = -\vec{u_2} \implies \vec{u_1} \in U_1 \text{ and } \vec{u_1} \in U_2 \\
\text{so } \vec{u_1} \in U_1 \cap U_2 = \left\{\vec{0}\right\} \\
\implies \vec{u_1} = \vec{0} \implies \vec{u_2} = -\vec{u_1} = -\vec{0} = \vec{0}
\end{aligned}
$$

By our previous theorem, $U_1 + U_2$ is direct because we can only write $\vec{0}$ in one way as a sum of vectors from U_1 and U_2. 😄

Thus, we have shown that $U_1 + U_2$ is direct if and only if $U_1 \cap U_2 = \left\{\vec{0}\right\}$.

Third proof: Let $\vec{v} = U_1 \cap U_2$, $\vec{v} = (x, y, 0) = (0, 0, z)$.

Then $x = y = z = 0$, so $\vec{v} = \vec{0}$.

This means that $U_1 \cap U_2 = \left\{\vec{0}\right\}$. 😄

Chapter 2

Finite-Dimensional Vector Spaces

2.1 Span and linear independence

Definition 2.1.1

Let $(V, F, +, \cdot)$ be a vector space.
A linear combination of $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m} \in V$ is a vector of the form:

$$\alpha \cdot \vec{v_1} + \alpha \cdot \vec{v_2} + \cdots + \alpha \cdot \vec{v_m} \text{ for some } \alpha_1, \alpha_2, \ldots, \alpha_m \in F$$

Example 2.1.1

Let $V = \mathbb{R}$, and our field being $\mathbb{F} = \mathbb{R}$.

$$6 \cdot (2, 1, -3) + 5 \cdot (1, -2, 4) = (17, -4, -2)$$

So $17, -4, 2$ is a linear combination of $(2, 1, -3)$ and $(1, -2, 4)$.

Definition 2.1.2

The span of $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m} \in V$ is the set of all linear combinations of $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}$.

$$\text{span}(\vec{v_1}, \vec{v_2}, \ldots, \vec{v_m}) = \left\{\alpha_1 \cdot \vec{v_1} + \alpha_2 \cdot \vec{v_2} + \cdots + \alpha_m \cdot \vec{v_m} \mid \alpha_1, \alpha_2, \ldots, \alpha_m \in F\right\}$$

Note:-

We have a few convention: $\text{span}\,() \coloneqq \left\{\vec{0}_v\right\}$.

Proposition 2.1.1

The span $(v_1, \ldots, v_m)$ is the smallest subspace of V that contains $v_1, \ldots, v_m$.

Proof: We have to show three things in 1.34.

(a) We know that $\vec{0}_v = 0_{\mathbb{F}} \cdot \vec{v_1} + 0_{\mathbb{F}} \cdot \vec{v_2} + \cdots + 0_{\mathbb{F}} \cdot \vec{v_m} \in \text{span}\left(\vec{v_1}, \ldots, \vec{v_m}\right)$. Thus, we are done

(b) Closed under addition $+_v$:

$$\underbrace{(a_1 \cdot \vec{v_1} + \ldots + a_m \cdot \vec{v_m})}_{\in \text{span}(\vec{v_1}, \ldots, \vec{v_m})} + \underbrace{(b_1 \cdot \vec{v_1} + \ldots + b_m \cdot \vec{v_m})}_{\in \text{span}(\vec{v_1}, \ldots, \vec{v_m})} = \underbrace{(a_1 + b_1) \cdot \vec{v_1} + \ldots + (a_m + b_m) \cdot \vec{v_m}}_{\in \text{span}(\vec{v_1}, \ldots, \vec{v_m})}$$

(c) Now closed under scalar multiplication:

$$\underbrace{\lambda}_{\in \mathbb{F}} \underbrace{\cdot}_{\text{in } V} \underbrace{(a_1 \cdot \vec{v_1} + \ldots + a_m \cdot \vec{v_m})}_{\in \text{span}(\vec{v_1},\ldots,\vec{v_m})} = \underbrace{\lambda \cdot a_1 \cdot \vec{v_1} + \ldots + \lambda \cdot a_m \cdot \vec{v_m}}_{\in \text{span}(\vec{v_1},\ldots,\vec{v_m})}$$

Now we have to show that this span contains $\vec{v_1}, \ldots, \vec{v_m}$:
In other words,

$$\vec{v_2} = 0_{\mathbb{F}} \cdot \vec{v_1} + 1_{\mathbb{F}} \cdot \vec{v_2} + 0_{\mathbb{F}} \cdot \vec{v_3} + \ldots + 0_{\mathbb{F}} \cdot \vec{v_m} \in \text{span}\left(\vec{v_1}, \ldots, \vec{v_m}\right)$$

Now, we must show it is the smallest.

Note:-
Draw some pics charlie

Suppose that $U \subseteq V$ is a subspace that contains $\vec{v_1}, \ldots, \vec{v_m}$.
Must show that span $\left(\vec{v_1}, \ldots, \vec{v_m}\right) \subseteq U$.
Let $v \in$ span $\left(\vec{v_1}, \ldots, \vec{v_m}\right)$, and is arbitrary.
We want to show that $v \in U$
We know some some things:

1. $v = a_1 \cdot \vec{v_1} + \ldots + a_m \cdot \vec{v_m}$ for some $a_1, \ldots, a_m \in \mathbb{F}$
2. $\vec{v_1}, \ldots, \vec{v_m} \in U$. Since $v_i \in U$, then $a_i \cdot \vec{v_i} \in U$ for all $i = 1, \ldots, m$.
   This is because U is a subspace, and is closed under scalar multiplication.
   But then $a_1 \cdot \vec{v_1} + \ldots + a_m \cdot \vec{v_m} \in U$ since U is closed under addition.

Therefore $v \in U$, and we are done. 😊

Special Situation: If span $(v_1, \ldots, v_m) = V$, we say that $v_1, \ldots, v_m$ spans V.

Example 2.1.2

Let $V = \mathbb{R}^3$, and the field $\mathbb{F} = \mathbb{R}$.
Then span $((1,0,0),(0,1,0),(0,0,1)) = \mathbb{R}^3$.

Proof: Let $(a,b,c) \in \mathbb{R}^3$ be arbitrary.
Then, $(a,b,c) = a \cdot (1,0,0) + b \cdot (0,1,0) + c \cdot (0,0,1)$.

😊

Definition 2.1.3

We say that V is finite-dimensional if V can be spanned by a finite list $v_1, v_2, \ldots, v_m$.

Example 2.1.3

$P_m(F) = \{\text{Polys of degree } \leq m \text{ with coefficients in } F\}$
And we claim that this is spanned by $1, x, x^2, \ldots, x^m$.
Because any $p(x) \in P_m(F)$ has the form $a_m \cdot x^m + \ldots + a_1 \cdot x + a_0$ for some $a_0, \ldots, a_m \in F$.

Proposition 2.1.2

$P(F) = \{\text{Polys with coefficients in } F\}$ is not finite-dimensional.

Proof: We procced by contradiction.
Suppose, for a contradiction, that $P(F)$ is finite-dimensional.

Then, there exists a finite list $p_1(x), \ldots, p_m(x)$ that spans $P(F)$.
In other words, span $(p_1(x), \ldots, p_m(x)) = P(F)$.
Let $n = max\,(\deg(p_1(x)), \ldots, \deg(p_m(x)))$.
Then, $deg(a_1 \cdot p_1(x) + \ldots + a_m \cdot p_m(x)) \leq n$ for all $a_1, \ldots, a_m \in F$.
So the degree of every element of span $(p_1(x), \ldots, p_m(x))$ is at most n.
Hence, $1_{\mathbb{F}} \cdot X^{n+1} \notin$ span $(p_1(x), \ldots, p_m(x))$.
This means that span $(p_1(x), \ldots, p_m(x)) \subsetneq P(F)$.
This is absurd!
So our assumption that $P(F)$ is finite-dimensional is false.

😊

Definition 2.1.4

Linear (In)depence.
Let $(V, \mathbb{F}, +, \cdot)$ be a vector space.
A list $\vec{v_1}, \ldots, \vec{v_m} \in V$ is linearly independent if the only way to write

$$\vec{0}_v = \alpha_1 \cdot \vec{v_1} + \ldots + \alpha_m \cdot \vec{v_m}, \alpha_1, \ldots, \alpha_m \in \mathbb{F}$$

Is to take $a_1 = \ldots = a_m = 0_{\mathbb{F}}$, otherwise it is linearly dependent.

Example 2.1.4

We want to show that $(1,0,0), (0,1,0)0, (0,0,1)$ are linearly independent in $\mathbb{R}^3 = V$.
because if

$$\vec{0}_{\mathbb{R}^3} = (0,0,0) = a_1 \cdot (1,0,0) + a_2 \cdot (0,1,0) + a_3 \cdot (0,0,1)$$

Then, $(0,0,0) = (a_1, a_2, a_3)$, so $a_1 = a_2 = a_3 = 0$.
Now suppose that $\vec{v_1}, \ldots, \vec{v_m} \in V$ is linearly independent and $v \in$ span $\left(\vec{v_1}, \ldots, \vec{v_n}\right)$.
This means: $\vec{v} = a_1 v_1 + \ldots + a_m v_m$ for some $a_1, \ldots, a_m \in \mathbb{F}$.
Now, suppose that $V = b_1 v_1 + \ldots + b_m v_m$ for some $b_1, \ldots, b_m \in \mathbb{F}$ as well
Now, let's subtract:

$$\vec{0}_V = v - v = (a_1 - b_1)v_1 + \ldots + (a_m - b_m)v_m$$

Since $\vec{v_1}, \ldots, \vec{v_m}$ is linearly independent (L.I, we must have $a_i - b_i = 0$ for all $i = 1, \ldots, m$.
This implies that $a_i = b_i$ for all $i = 1, \ldots, m$.
Thus, there is exactly one way to write V as a linear combination of $\vec{v_1}, \ldots, \vec{v_m}$

Key result: Let $(V, \mathbb{F}, +, \cdot)$ be a finite-dimensional vector space.
Then the length of any-list of Linear Independence vectors is at most the length of any list of spanning vectors.

Example 2.1.5

We want to show that $(1,0,0), (0,1,0), (0,0,1)$ spans $\mathbb{R}^3$.
This implies that the list $(2,-1,\pi), (\sqrt{3}, -7, e), (\sqrt{19}, -1, 7), (0, -5, \sqrt{2} + \sqrt{3})$ is not linearly independent.
Since the length of the first list is 3, and the length of the second list is 4.
Thus, this list cannot be linearly independent.

Lemma 2.1.1 Linear Dependence Lemma (LDL)

We want to prove this, but let's do some prep work first.

Prep work: Say $\vec{v_1}, \ldots, \vec{v_n} \in V$ is linearly dependent. Then there is a $j \in \{1, \ldots, m\}$ such that

(i) $v_j \in \text{span}\left(\vec{v_1}, \ldots, \vec{v_j - 1}\right)$

(ii) $\text{span}\left(\vec{v_1}, \ldots, \vec{v_m}\right) = \text{span}\left(v_1, \ldots, \hat{v_j}, \ldots, v_m\right)$, where $\hat{v_j}$ means that we remove v_j from the list.

Now, let's prove this.

Proof: Since $\vec{v_1}, \ldots, \vec{v_m}$ are linearly dependent, there are $a_1, \ldots, a_m$ not all zero such that

$$\vec{0}_v = a_1 \cdot \vec{v_1} + \ldots + a_m \cdot \vec{v_m}, a_1, \ldots, a_m \in \mathbb{F}$$

(i) Let $j = \max\{i \mid a_i \neq 0\}$, so that $a_1 v_1 + \ldots + a_j v_j = \vec{0}_v$ and $a_j \neq 0$.

$$\Longrightarrow v_j = -\frac{1}{a_j}\left(a_1 v_1 + \ldots + a_{j-1} v_{j-1}\right) = \left(-\frac{a_1}{a_j}\right) v_1 + \ldots + \left(-\frac{a_{j-1}}{a_j}\right) v_{j-1}$$
$$\Longrightarrow v_j \in \text{span}\left(v_1, \ldots, v_{j-1}\right)$$

(ii) $\text{span}\left(v_1, \ldots, \hat{v_j}, \ldots, v_m\right) \subseteq \text{span}\left(v_1, \ldots, v_m\right)$.

Note:-

We have to do the one above as well.

Now, we want to show the other direction as well.

$$\text{span}\left(\vec{v_1}, \ldots, \vec{v_m}\right) \subseteq \text{span}\left(v_1, \ldots, \hat{v_j}, \ldots, v_m\right)$$

Let $v \in \text{span}\left(\vec{v_1}, \ldots, \vec{v_m}\right)$.
Then, $v = b_1 v_1 + \ldots + b_m v_m$ for some $b_1, \ldots, b_m \in \mathbb{F}$.

$$\Longrightarrow \quad v = b_1 v_1 + \ldots + b_j\left[\left(-\frac{a_1}{a_j}\right) v_1 + \ldots + \left(-\frac{a_{j-1}}{a_j}\right) v_{j-1}\right] + b_{j+1} v_{j+1} + \ldots + b_m v_m \quad \text{where } v_j \text{ (from (i))}$$
$$\Longrightarrow \quad v \in \text{span}\left(v_1, \ldots, \hat{v_j}, \ldots, v_m\right)$$

Thus, $\text{span}\left(\vec{v_1}, \ldots, \vec{v_m}\right) = \text{span}\left(v_1, \ldots, \hat{v_j}, \ldots, v_m\right)$.

😄

Proof key result: Let $\vec{v_1}, \ldots, \vec{v_m} \in V$ be a linearly independence list.
Let $\vec{u_1}, \ldots, \vec{u_n} \in v$ be a spanning list, $V = \text{span}\left(\vec{u_1}, \ldots, \vec{u_n}\right)$.
We need to show that $m \leq n$.

Step 1:

$$v_1 \in \text{span}\left(\vec{u_1}, \ldots, \vec{u_n}\right) \overbrace{\Longrightarrow}^{\text{PSET4}} \vec{v_1}, \vec{u_1}, \ldots, \vec{u_n} \text{ is linearly dependent}$$

With the linear independence lemma, we know there exits $\vec{u_{j_1}}$ such that

$$\vec{u_{j_1}} \in \text{span}\left(\vec{v_1}, \vec{u_1}, \ldots, \vec{u_{j_1} - 1}\right)$$

And

$$\text{span}\left(\vec{v_1}, \vec{u_1}, \ldots, \vec{u_n}\right) = \text{span}\left(\vec{v_1}, \vec{u_1}, \ldots, \hat{\vec{u_{j_1}}}, \ldots, \vec{u_n}\right)$$

Note:-

NB means nota bene, which means note well.
Notice that v_1 is not plucked out from our list when we apply LDL.
If it were, then LDL would say $v_1 \in \operatorname{span}() = \left\{\vec{0}_v\right\}$.
This implies that $v_1 = \vec{0}_v$,
But $\vec{v_1}, \ldots, \vec{v_m}$ is linearly independent.
As $\vec{0}_v = 1_{\mathbb{F}} \cdot \vec{v_1} + 0_{\mathbb{F}} \cdot \vec{v_2} + \ldots + 0_{\mathbb{F}} \cdot \vec{v_m}$ is the only way to write $\vec{0}_v$ as a linear combination of $\vec{v_1}, \ldots, \vec{v_m}$.
Thus, $v_1 \neq \vec{0}_v$.

Step 2: $v_2 \in \operatorname{span}\left(\vec{u_1}, \ldots, \vec{u_n}\right) = \operatorname{span}\left(\vec{v_1}, \vec{u_1}, \ldots, \vec{u_n}\right) = \operatorname{span}\left(\vec{v_1}, \vec{u_1}, \ldots, \hat{\vec{u_{j_1}}}, \ldots, \vec{u_n}\right)$

Again, with the result in PSET4, we know that

$$\vec{v_1}, \vec{v_2}, \vec{u_1}, \ldots, \hat{\vec{u_{j_1}}}, \ldots, \vec{u_n} \quad \text{is linearly dependent}$$

With the linear independence lemma, we know there exits $\vec{u_{j_2}}$ such that

$$\operatorname{span}\left(\vec{v_1}, \vec{u_1}, \ldots, \hat{\vec{u_{j_1}}}, \ldots, \vec{u_n}\right) = \operatorname{span}\left(\vec{v_1}, \vec{v_2}, \vec{u_1}, \ldots, \hat{\vec{u_{j_1}}}, \ldots, \hat{\vec{u_{j_2}}}, \ldots, \vec{u_n}\right)$$

After m steps: Our list is $\vec{v_1}, \ldots, \vec{v_m}$, some u's implies that $m \leq n$

Thus, we have shown that $m \leq n$.

😄

2.2 Basis

Definition 2.2.1

Let $(V, \mathbb{F}, +, \cdot)$ be a vector space.
A basis for V is a list $\vec{v_1}, \ldots, \vec{v_n}$ that spans V.
(i.e., $V = \text{span}\left(\vec{v_1}, \ldots, \vec{v_n}\right)$) and is linearly independent.

Example 2.2.1

(i) Let $V = \mathbb{F}^n$ (think $V = \mathbb{R}^n$ or $\mathbb{C}^n$)

We can define the standard basis for $\mathbb{F}^n$ as:

$$\begin{aligned} v_1 &= (1, 0, \ldots, 0) \\ v_2 &= (0, 1, \ldots, 0) \\ &\vdots \\ v_n &= (0, 0, \ldots, 1) \end{aligned}$$

e.g., $V = \mathbb{R}^3 = \text{span}\left((1,0,0), (0,1,0), (0,0,1)\right)$. This list is linearly independent.

(ii) $V = \mathbb{R}^2$ The list $(1,2), (2,3)$ is a basis.

Linearly Independence: If

$$\begin{aligned} a_1(1,2) + a_2(2,3) &= (0,0)_{\vec{0}_{\mathbb{R}^2}} \\ \implies (a_1 + 2a_2, 2a_1 + 3a_2) &= (0,0) \\ \implies a_1 = a_2 &= 0 \end{aligned}$$

(iii) $V = P_m(\mathbb{R})$

Thus, the list $1, x, x^2, \ldots, x^m$ is a basis for V

Proposition 2.2.1

$\vec{v_1}, \ldots, \vec{v_n} \in V$ is a bais for V if and only if every $\vec{v} \in V$ can be written uniquely as a linear combination of $\vec{v_1}, \ldots, \vec{v_n}$.

Proof of $\implies$: Say $\vec{v_1}, \ldots, \vec{v_n} \in V$ is a basis.
Let $\vec{v} \in V$. Since $V = \text{span}\left(\vec{v_1}, \ldots, \vec{v_n}\right)$,
we know that $\vec{v} = a_1\vec{v_1} + \ldots + a_n\vec{v_n}$ for some $a_1, \ldots, a_n \in \mathbb{F}$.
Since $\vec{v_1}, \ldots, \vec{v_n}$ are linearly independent, we know this representation is unique. 😊

Proof of $\impliedby$: Suppose that every $\vec{v} \in V$ can be written uniquely as $\vec{v} = a_1\vec{v_1} + \ldots + a_n\vec{v_n}$ for some $a_1, \ldots, a_n \in \mathbb{F}$.
Then $\vec{v} \in \text{span}\left(\vec{v_1}, \ldots, \vec{v_n}\right)$, so $V \subseteq \text{span}\left(\vec{v_1}, \ldots, \vec{v_n}\right)$.
By the definition of span, we know that $\text{span}\left(\vec{v_1}, \ldots, \vec{v_n}\right) \subseteq V$.
Thus, $V = \text{span}\left(\vec{v_1}, \ldots, \vec{v_n}\right)$.
Next, let $\vec{v} = \vec{0}_v$.
We know that $\vec{0}_v = a_1\vec{v_1} + \ldots + a_n\vec{v_n}$ for unique $a_1, \ldots, a_n \in \mathbb{F}$.
On the other hand (OTOH): taking $a_1 = \ldots = a_n = 0$ works!
Therefore, the only way to write $\vec{0}_v$ is a linearly combination of $\vec{v_1}, \ldots, \vec{v_n}$

is to take $a_1 = \ldots = a_n = 0_{\mathbb{F}}$.
The definition implies that $\vec{v_1}, \ldots, \vec{v_n}$ is linearly independent. ☺

Thus, we have shown that $\vec{v_1}, \ldots, \vec{v_n} \in V$ is a basis for V if and only if every $\vec{v} \in V$ can be written uniquely as a linear combination of $\vec{v_1}, \ldots, \vec{v_n}$.

Theorem 2.2.1

Let $(V, \mathbb{F}, +, \cdot)$ be a finite-dimensional vector space (fdvs).
Then every spanning list for V can be trimmed to a basis.

Proof: Say that $\vec{v_1}, \ldots, \vec{v_n}$ is a strong list for V.

Algorithm 1: Trimming

$B = \left\{\vec{v_1}, \ldots, \vec{v_n}\right\}$; /* Note that B has no order. */
for $j = 1, \ldots, n$ **do**
  if $v_j \in \text{span}\left(\left\{\vec{v_1}, \ldots, \vec{v_{j-1}}\right\} \cap B\right)$ **then**
    Delete v_j from B;
end

When the loop is finished, the set B gives rise to a basis (any order). ☺

Example 2.2.2

$V = \mathbb{R}^3$.
Let $v_1 = (1, 0, 0)$, $v_2 = (1, 1, 1)$, $v_3 = (0, 1, 1)$, and $v_4 = (0, 0, 1)$.
Let $B = \left\{\vec{v_1}, \vec{v_2}, \vec{v_3}, \vec{v_4}\right\}$

Step 1: Is $v_1 \in \text{span}(\emptyset \cap B) = \text{span}() = \left\{\vec{0}_v\right\}$?
NO. Leave B alone.

Step 2: Is $v_2 \in \text{span}(\{v_1\} \cap B) = \text{span}(v_1)$?
Does $v_2 = a_1 \cdot v_1$.
No!
Leave B alone.

Step 3: Is $v_3 \in \text{span}(\{v_1, v_2\} \cap B) = \text{span}(v_1, v_2)$?
Does $v_3 = a_1 \cdot v_1 + a_2 \cdot v_2$?
Yes!

$$v_3 = -v_1 + v_2$$

New $B = \{v_1, v_2, v_4\}$

Step 4: Is $v_4 \in \text{span}(\{v_1, v_2, v_3\} \cap B) = \text{span}(v_1, v_2)$?
Does $v_4 = a_1 \cdot v_1 + a_2 \cdot v_2$?
No!
Leave B alone.
Thus, $B = \{v_1, v_2, v_4\}$ is a basis for V through trimming.

Corollary 2.2.1

Any linearly independence list $\vec{v_1}, \ldots, \vec{v_m}$ on V can be extended to a basis.

Proof: Let $\vec{u_1}, \ldots, \vec{u_n}$ be any basis for V .
Trim the enlarged list $\vec{v_1}, \ldots, \vec{v_m}, \vec{u_1}, \ldots, \vec{u_n}$.

No $\vec{v}_i$ is deleted during trimming (LDL). 😊

Semi-simplicity: Let $(V, \mathbb{F}, +, \cdot)$ be a finite-dimensional vector space.
Let $U \subseteq V$ be a subspace.
Then, there is a subspace $W \subseteq V$ (not necessarily unique) such that $V = U \oplus W$.

Idea: Let $\vec{u_1}, \ldots, \vec{u_n}$ be a basis for U.
Complete to a spanning list of V.
$\vec{u_1}, \ldots, \vec{u_n}, \vec{w_1}, \ldots, \vec{w_m}$.
The space $W = \text{span}\left(\vec{w_1}, \ldots, \vec{w_n}\right)$ works!

Claim: U itself is finite-dimensional.
Assume claim: Let $\vec{u_1}, \ldots, \vec{u_n}$ be a basis for U .
This implies that $\vec{u_1}, \ldots, \vec{u_n}$ is linearly independent in U , but also in V .
Now, extend to a basis of V: $\vec{u_1}, \ldots, \vec{u_n}, \vec{w_1}, \ldots, \vec{w_m}$.
Take $W = \text{span}\left(\vec{w_1}, \ldots, \vec{w_m}\right)$.
We want

(i) $U + W \supseteq V$, the other direciton is trivial.

(ii) $U \cap W = \left\{\vec{0}_v\right\}$

Ok, let's start.

(i) Let $v \in V$. Since $V = \text{span}\left(\vec{u_1}, \ldots, \vec{u_n}, \vec{w_1}, \ldots, \vec{w_m}\right)$, we know:

$$v = \underbrace{a_1\vec{u_1} + \ldots + a_n\vec{u_n}}_{\in U, a_i \in \mathbb{F}} + \underbrace{b_1\vec{w_1} + \ldots + b_m\vec{w_m}}_{\in W, b_i \in \mathbb{F}} = U + W$$

As such, $V = U + W$

(ii) Let $v \in U \cap W$.

$$v = a_1\vec{u_1} + \ldots + a_n\vec{u_n} \quad (vinU = \text{span}\left(\vec{u_1}, \ldots, \vec{u_n}\right))$$
$$v = b_1\vec{w_1} + \ldots + b_m\vec{w_m} \quad (vinW = \text{span}\left(\vec{w_1}, \ldots, \vec{w_m}\right))$$

Now, let's substract

$$\vec{0}_v = v - v = a_1\vec{u_1} + \ldots + a_n\vec{u_n} - b_1\vec{w_1} - \ldots - b_m\vec{w_m}$$

Since u's and w's are linearly independent in V, this forces a's and b's to be all $0_{\mathbb{F}}$
This implies that $v = \vec{0}_v$
Thus, $U \cap W \subseteq \left\{\vec{0}_v\right\}$.

Thus, $U \cap W = \left\{\vec{0}_v\right\}$ and $U + W = V$.
Therefore, $V = U \oplus W$.

proof of claim: If $U = \left\{\vec{0}_v\right\}$ then we are done!
This is because $U = \text{span}()$
Otherwise, there is a $\vec{v_1} \neq \vec{0}_v$ in U .
If $U = \text{span}\left(\vec{v_1}\right)$, then we are done.
This is because U is finite-dimensional.
Otherwise, there is a $\vec{v_2} \in U$ such that $\vec{v_2} \notin \text{span}\left(\vec{v_1}\right)$.

This implies that (v_1, v_2) is a linearly independent list in U .
Which means that the list is also linearly independent in V .
If $U = \text{span}\left(\vec{v}_1, \vec{v}_2\right)$, then we are done.
Otherwise there is a $\vec{v}_3 \in U$ such that $\vec{v}_3 \notin \text{span}\left(\vec{v}_1, \vec{v}_2\right)$.
This implies that (v_1, v_2, v_3) is a linearly independent list in U .
Which means that the list is also linearly independent in V .
This process terminates:
V is finite dimensional, which implies $V = \text{span}\left(\vec{x}_1, \ldots, \vec{x}_p\right)$
At step m we produce a linearly independent list $\vec{v}_1, \ldots, \vec{v}_m$ of V .
The key result we have proved in class: $m \leq p$. 😆

2.3 Dimension

Theorem 2.3.1

Any two bases of a finite-dimensional vector space V have the same length.

Proof: Say $\vec{v}_1, \ldots, \vec{v}_m$ and $\vec{u}_1, \ldots, \vec{u}_n$ are bases for V .
Let $\vec{v}_1, \ldots, \vec{v}_m$ are linearly independent in V .
Let $\vec{u}_1, \ldots, \vec{u}_n$ span V .
By the key result $m \leq n$. Reverse roles to get $n \leq m$.
Thus, $m = n$.
The length of any basis for V is called the dimension of V.
😆

Example 2.3.1

(i) $V = \mathbb{R}^n$ standard basis $\vec{e}_1, \ldots, \vec{e}_n$.

These vectors look like $(0, \ldots, 0, 1, 0, \ldots, 0)$ where the 1 is in the ith position for each $i = 1, \ldots, n$.

This implies that the dimension of $\mathbb{R}^n$ is n.

(ii) $P_m(\mathbb{R})$ has basis $1, x, x^2, \ldots, x^m$.

This implies that the dimension of $P_m(\mathbb{R})$ is $m + 1$.

Properties: (i) If $U \subseteq V$ is a subspace, then $\dim U \leq \dim V$.

Say V is finite-dimensional, which implies that U is finite-dimensional.

A basis $\vec{u}_1, \ldots, \vec{u}_n$ for U is a linearly independent in V .

This means we can extend a basis $\vec{u}_1, \ldots, \vec{u}_n, \vec{w}_1, \ldots, \vec{w}_m$ of V.

Thus, $\dim U = n \leq n + m = \dim V$

(ii) Say that $\dim V = n$, and $\vec{v}_1, \ldots, \vec{v}_n$ is a linearly independent list in V,

Then $\vec{v}_1, \ldots, \vec{v}_n$ spans V .

Proof: Extend $\vec{v}_1, \ldots, \vec{v}_n$ to a basis of V.

Result is a basis for V. This basis has length $\dim V = n$.

This means the extension process didn't add new vectors.

Which means that $\vec{v}_1, \ldots, \vec{v}_n$ is already a basis.

Thus, $\vec{v}_1, \ldots, \vec{v}_n$ spans V . 😆

(iii) Say that $\dim V = n$ and that $\vec{v}_1 \ldots \vec{v}_n$ spans V.

Then $\vec{v}_1, \ldots, \vec{v}_n$ is a linearly independent list.

Note:-
Do this as an exercise

.

Example 2.3.2

Take $V = \{p(x) \in P_3(\mathbb{R}) \colon p`(5) = 0\} \subseteq P_3(\mathbb{R})$
We know that $P_3(\mathbb{R})$ is 4-dimensional, with a basis $1, x, x^2, x^3$.

Claim: $\dim V < 4$ and that V is 3-dimensional.

Proof: Since $V \subseteq P_3(\mathbb{R})$, we know that V is finite-dimensional i.e, $\dim V < 4$.
We just need to rule out that $\dim V = 4$.
Suppose that $\dim V = 4$.
Then $1, x, x^2, x^3$ is a basis for V.
Then $V \subset P_3(\mathbb{R})$ both have dimension 4.
Let p_1, p_2, p_3, p_4 be a basis for V
Then p_1, p_2, p_3, p_4 are linearly independent in $P_3(\mathbb{R})$.
But, the $\dim P_3(\mathbb{R}) = 4$, so p_1, p_2, p_3, p_4 also spans $P_3(\mathbb{R})$.
This means that $V = P_3(\mathbb{R})$. This is as they both span $(p_1, \ldots, p_4)$.
Let $p(x) = x$.
Then $p`(5) = 1$
Thus, $p(x) \notin V$.
Therefore, $V \neq P_3(\mathbb{R})$.

😄

Definition 2.3.1: Dimension of a sum

Let $U_1, U_2 \subseteq V$ be finite-dimensional subspaces.
Then $\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$.

Proof: Let $\vec{u}_1, \ldots, \vec{u}_n$ be a basis for $U_1 \cap U_2$.
Then, we can extend the basis in two ways:

(i) a basis $\vec{u}_1, \ldots, \vec{u}_n, \vec{v}_1, \ldots, \vec{v}_m$ for U_1

(ii) a basis $\vec{u}_1, \ldots, \vec{u}_n, \vec{w}_1, \ldots, \vec{w}_p$ for U_2

Claim: Let $\vec{u}_1, \ldots, \vec{u}_n, \vec{v}_1, \ldots, \vec{v}_m, \vec{w}_1, \ldots, \vec{w}_p$ is a basis for $U_1 + U_2$.
Assume claim true for now.

$$\begin{aligned}\dim(U_1 + U_2) &= n + m + p \quad \text{our claim and defintion of dim} \\ &= (n+m) + (n+p) - n \quad \text{algebra} \\ &= \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2) \quad \text{defintion of dim}\end{aligned}$$

For the claim we need to prove:

Proof of span: This is left for us. 😃

Proof of linear independence: Suppose there are scalars $a_1, \ldots, a_n, b_1, \ldots, b_m, c_1, \ldots, c_p \in \mathbb{F}$ such that:

$$a_1\vec{u}_1 + \ldots + a_n\vec{u}_n + b_1\vec{v}_1 + \ldots + b_m\vec{v}_m + c_1\vec{w}_1 + \ldots + c_p\vec{w}_p = \vec{0}_v$$

We want to show that $a_1 = \ldots = a_n = b_1 = \ldots = b_m = c_1 = \ldots = c_p = 0_{\mathbb{F}}$.
Let's introduce sum notation:

$$\underbrace{\sum_{i=1}^{n} a_i\vec{u}_i + \sum_{j=1}^{m} b_j\vec{v}_j}_{\in U_1} = \underbrace{-\sum_{k=1}^{p} c_k\vec{w}_k}_{\in U_1}$$

This shows that $\sum_{k=1}^{p} c_k\vec{w}_k \in U_1 \cap U_2 = \operatorname{span}\left(\vec{u}_1, \ldots, \vec{u}_n\right)$.
The u's are basis for $U_1 \cap U_2$, so there are scalars $d_1, \ldots, d_n \in \mathbb{F}$ such that:

$$\sum_{k=1}^{p} c_k\vec{w}_k = \sum_{i=1}^{n} d_i\vec{u}_i$$

This impplies that $c_1\vec{w}_1 + \ldots + c_p\vec{w}_p - d_1\vec{u}_1 - \ldots - d_n\vec{u}_n = \vec{0}_v$.
u's and w's are a basis for U_2.
This implies that they are linearly independent and $c_1 = \ldots = c_p = d_1 = \ldots = d_n = 0_{\mathbb{F}}$.
This shows that $\sum_{i=1}^{n} a_1\vec{u}_i + \sum_{j=1}^{m} b_j\vec{v}_j = \vec{0}_v$.
Next:
u's and v's are a absis for U_1.
This implies that they are linear independence.
Which implies that $a_1 = \ldots = a_n = b_1 = \ldots = b_m = 0_{\mathbb{F}}$.
Thus, we have proven this basis is linearly independent.

😃

Thus, we have proven the claim.
Thus, we proven the theorem. 😃

Chapter 3

Linear Transformations

3.1 Linear Maps

Definition 3.1.1: Linear Maps

Let V, W be vector spaces over the same field $F(= \mathbb{R} V \mathbb{C})$.
Meaning that $V = (V, \mathbb{F}, +_V, \cdot_V)$ and $W = (W, \mathbb{F}, +_W, \cdot_W)$.
A linear map: $T\colon V \longrightarrow W$ is a function such that:

(i) $T(u +_v v) = T(u) +_w T(v)$ for all $u, v \in V$

(ii) $T(\lambda \cdot_v v) = \lambda \cdot_w T(v)$ for all $v \in V$ and $\lambda \in \mathbb{F}$

In other words, they preserve the vector space structure.

Note:-

Observation: $T(\vec{0_v}) = \vec{0_w}$.
Reason:

$$T(\vec{0_v}) = T(\vec{0}_v +_v \vec{0}_v) = T(\vec{0}_v) +_w T(\vec{0}_v)$$

Adding $-T(\vec{0}_v)$ to both sides, we get:

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

Example 3.1.1

We will be showing a lot of examples today!

(i) Zero map:

$$\begin{aligned} 0\colon V &\longrightarrow W \\ v &\longmapsto 0_w \end{aligned}$$

(ii) Identity map:

$$\begin{aligned} \mathrm{id}_V\colon V &\longrightarrow V \\ v &\longmapsto v \end{aligned}$$

Note:-

Notation $\mathscr{L}(V, W) = \{T\colon V \longrightarrow W \mid T \text{ is linear}\}$.

(iii) Differentiation map: $D \in \mathscr{L}(P(\mathbb{R}), P(\mathbb{R}))$

$$\begin{aligned} D\colon P(\mathbb{R}) &\to P(\mathbb{R}) \\ p(x) &\mapsto p\prime(x) \end{aligned}$$

Let's check!

Linear:

(a) $D(p(x) + q(x)) = (p(x) + q(x))\prime = p\prime(x) + q\prime(x) = D(p(x)) + D(q(x))$

(b) $D(\lambda p(x)) = (\lambda p(x))\prime = \lambda p\prime(x) = \lambda D(p(x))$

(iv) Integration: $I \in \mathscr{L}(P(\mathbb{R}), P(\mathbb{R}))$

$$\begin{aligned} I\colon P(\mathbb{R}) &\to P(\mathbb{R}) \\ p(x) &\mapsto \int_0^1 p(x)dx \end{aligned}$$

Let's check!

Linear:

(a) $I(p(x) +_{P(\mathbb{R})} q(x)) = \int_0^1 (p(x) + q(x))dx = \int_0^1 p(x)dx + \int_0^1 q(x)dx = I(p(x)) +_{\mathbb{R}} I(q(x))$

(b) $I(\lambda \cdot_{P(\mathbb{R})} p(x)) = \int_0^1 (\lambda p(x))dx = \lambda \int_0^1 p(x)dx = \lambda I(p(x))$

(v) Shift: $S \in \mathscr{L}(\mathbb{F}^\infty, \mathbb{F}^\infty)$

$$\begin{aligned} S\colon \mathbb{F}^\infty &\to \mathbb{F}^\infty \\ (x_1, x_2, \ldots) &\mapsto (0, x_1, x_2, \ldots) \end{aligned}$$

(vi) $T\colon \mathbb{R}^3 \to \mathbb{R}^2$

$$\begin{aligned} T\colon \mathbb{R}^3 &\to \mathbb{R}^2 \\ (x_1, x_2, x_3) &\mapsto (5x + 7y - z, 2x - y) \end{aligned}$$

Properties: Remember our notation:

$$\mathscr{L}(V, W) = \{T\colon V \to W \mid T \text{ is linear}\}$$

The set $\mathscr{L}(V, W)$ can be given the structure of a vector space over $\mathbb{F}$.

(i) Addition: Let $T, S \in \mathscr{L}(V, W)$. Where $T\colon V \to W$ and $S\colon V \to W$.

$$\begin{aligned} (T + S)\colon V &\to W \\ v &\mapsto T(v) +_W S(v) \end{aligned}$$

if and only if $(S + T)(v) = S(v) +_W T(v)$ for all $v \in V$.

(ii) Multiplication: Let $T \in \mathscr{L}(V, W)$ and $\lambda \in \mathbb{F}$. With $T\colon V \to W$.

$$\begin{aligned} (\lambda T)\colon V &\to W \\ v &\mapsto \lambda \cdot_W T(v) \end{aligned}$$

If and only if $(\lambda T)(v) = \lambda \cdot_W T(v)$ for all $v \in V$.

(iii) Bonus structure!

We can also multiply linearly maps using function composition:

$$U \overbrace{\longrightarrow}^{T} V \overbrace{\longrightarrow}^{S} W$$

Thus, we can define $(S \cdot T)(u) = S(T(u))$.

Propositions of composition:

(a) Associativity: $U \overbrace{\longrightarrow}^{T_1} V \overbrace{\longrightarrow}^{T_2} W \overbrace{\longrightarrow}^{T_3} X$

$$T_3 \cdot (T_2 \cdot T_1) = (T_3 \cdot T_2) \cdot T_1$$

(b) Identities: $T\colon V \longrightarrow W$

$$\begin{aligned} \mathrm{id}_V \colon V &\longrightarrow V \\ v &\longmapsto v \\ \mathrm{id}_W \colon W &\longrightarrow W \\ w &\longmapsto w \end{aligned}$$

Thus, $\mathrm{id}_W \cdot T = T = T \cdot \mathrm{id}_V$.

(c) Distributivity: $S_1, S_2 \colon V \longrightarrow W$ and $T\colon W \longrightarrow X$

$$T \cdot (S_1 + S_2) = T \cdot S_1 + T \cdot S_2$$

Important: Say V is a finite dimensional vector space over $\mathbb{F}_1$, and $\vec{v}_1, \dots, \vec{v}_n$ is a basis for V . Then a linear map $T\colon V \longrightarrow W$ is determined by the values $T(\vec{v}_1), \dots, T(\vec{v}_n)$.

Reason: Let $\vec{v} \in V = \operatorname{span}\left(\vec{v}_1, \dots, \vec{v}_n\right)$.
This implies that $\vec{v} = \lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n$ for some and unique $\lambda_1, \dots, \lambda_n \in \mathbb{F}$.
Then:

$$\begin{aligned} T(\vec{v}) &= T(\lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n) \\ &= T(\lambda_1\vec{v}_1) + \cdots + T(\lambda_n\vec{v}_n) \\ &= \lambda_1 T(\vec{v}_1) + \cdots + \lambda_n T(\vec{v}_n) \end{aligned}$$

Theorem 3.1.1 Axler 3.5

Now suppose that $\vec{w}_1, \dots, \vec{w}_n \in W$, not necessarily a basis.
Then there is exactly one linear map $T\colon V \longrightarrow W$ mapping
the basis $\vec{v}_1, \dots, \vec{v}_n$ to the vectors $\vec{w}_1, \dots, \vec{w}_n$ respectively.
Meaning that $T(\vec{v}_i) = \vec{w}_i$ for all $i = 1, \dots, n$.
Again: $\vec{v} \in V, \vec{v} = \lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n$.
Then:

$$T(\vec{v}) = T(\lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n) = \lambda_1 T(\vec{v}_1) + \cdots + \lambda_n T(\vec{v}_n) = \lambda_1\vec{w}_1 + \cdots + \lambda_n\vec{w}_n$$

3.2 Null spaces and Ranges

Definition 3.2.1: Kernels or null spaces

Let $T\colon V \longrightarrow W$ be a linear map.
The kernel (null spaces) of T is $\ker T := \left\{\vec{v} \in V : T(\vec{v}) = \vec{0}_W\right\}$

Note:-

The image on our canvas page is this definition.

Note:-

We know that $T(\vec{0}_V) = \vec{0}_W$, so $\vec{0}_V \in \ker T$.

Example 3.2.1

(a) $\ker(0) = V$

$$\begin{aligned} 0\colon V &\longrightarrow W \\ v &\longmapsto 0_W \end{aligned}$$

(b) $\ker(\mathrm{id}_V) = \left\{\vec{0}_V\right\}$

$$\begin{aligned} \mathrm{id}_V\colon V &\longrightarrow V \\ v &\longmapsto v \end{aligned}$$

(c)

$$\begin{aligned} D\colon P(\mathbb{R}) &\longrightarrow P(\mathbb{R}) \\ p(x) &\longmapsto p\prime(x) \end{aligned}$$

Then:

$$\begin{aligned} \ker D &= \{p(x) \in P(\mathbb{R})\colon p\prime(x) = 0\} \\ &= \{p(x) \in P(\mathbb{R})\colon p(x) = a_0\} \\ &= \{a_0\colon a_0 \in \mathbb{R}\} \\ &= \mathbb{R} \end{aligned}$$

(d) Shift

$$\begin{aligned} S\colon \mathbb{F}^\infty &\longrightarrow \mathbb{F}^\infty \\ (x_1, x_2, \ldots) &\longmapsto (x_2, x_3, \ldots) \end{aligned}$$

Then:

$$\begin{aligned} \ker S &= \left\{(x_1, x_2, \ldots) \in \mathbb{F}^\infty \mid (x_2, x_3, \ldots) = \vec{0}_{\mathbb{F}^\infty}\right\} \\ &= \{(x_1, 0, 0, \ldots) \in \mathbb{F}^\infty \mid x_1 \in \mathbb{F}\} \end{aligned}$$

Proposition 3.2.1

In general, $\ker T$ is a subspace of V.

Proof: Let $T\colon V \longrightarrow W$ be a linear map.
Now we want to check 1.34:

(i) $T(\vec{0}_V) \in \ker T$ as $T(\vec{0}_V) = \vec{0}_W$.

(ii) Closed under addition: Let $\vec{u}, \vec{v} \in \ker T \subseteq V$.
We want to show that $\vec{u} +_V \vec{v} \in \ker T$.

$$\begin{aligned} T(\vec{u} +_V \vec{v}) &= T(\vec{u}) +_W T(\vec{v}) \\ &= \vec{0}_W +_W \vec{0}_W \\ &= \vec{0}_W \end{aligned}$$

Thus, $\vec{u} +_V \vec{v} \in \ker T$.

(iii) Closed under scalar multiplication: Let $\vec{u} \in \ker T$ and $\lambda \in \mathbb{F}$.
We want to show that $\lambda \cdot_V \vec{u} \in \ker T$.

$$\begin{aligned} T(\lambda \cdot_V \vec{u}) &= \lambda \cdot_W T(\vec{u}) \\ &= \lambda \cdot_W \vec{0}_W \\ &= \vec{0}_W \end{aligned}$$

Thus, $\lambda \cdot_V \vec{u} \in \ker T$.

Therefore, $\ker T$ is a subspace of V.

😊

Definition 3.2.2: Injective

A linear map is injective if:

$$\underbrace{T(\vec{u})}_{\text{equal outputs}} = T(\vec{v}) \quad \underbrace{\implies}_{\text{must come from}} \quad \underbrace{u = v}_{\text{equal inputs}}$$

The cont appositive:

$$\underbrace{u \neq v}_{\text{unequal inputs}} \implies \underbrace{T(\vec{u}) \neq T(\vec{v})}_{\text{unequal outputs}}$$

Proposition 3.2.2

Let $T\colon V \longrightarrow W$ be a linear map.
Then T is injective if and only if $\ker T = \left\{\vec{0}_V\right\}$.

Proof of $\implies$ ***:*** Assume $T\colon V \longrightarrow W$ is injective.
We know that $\vec{0}_V \in \ker T$.
We want to show that $\ker T \subseteq \left\{\vec{0}_V\right\}$.

Let $\vec{v} \in \ker T$.
Then $T(\vec{v}) = T(\vec{0}_V + \vec{0}_V) = T(\vec{0}_V) + T(\vec{0}_V) = \vec{0}_W + \vec{0}_W = \vec{0}_W$. 😄

Proof of $\Longleftarrow$ **:** We are given that $\ker T = \left\{\vec{0}_V\right\}$.
We want to show that T is injective.
Suppose $T(\vec{u}) = T(\vec{v})$.
Then $T(\vec{u}) - T(\vec{v}) = \vec{0}_W$.
By linearity, $T(\vec{u} - \vec{v}) = \vec{0}_W$.
Thus, $\vec{u} - \vec{v} \in \ker T$.
This means that $\vec{u} - \vec{v} = \vec{0}_V$.
Therefore, $\vec{u} - \vec{v} = \vec{0}_V \implies \vec{u} = \vec{v}$.

😄

As we have proven both directions, we have proven the proposition.

Definition 3.2.3: Images

Let $T \in \mathscr{L}(V, W)$. Then the image of T is $Im(T) = \{w \in W \mid w = T(v) \text{ for some } v \in V\}$.
Also denoted as Range(T).
It is a subspace of W (Axle 3.19)

Example 3.2.2

(i) $Im(0) = \left\{\vec{0}_W\right\}$

$$\begin{aligned} 0\colon V &\longrightarrow W \\ v &\longmapsto \vec{0}_W \end{aligned}$$

(ii) $Im(\mathrm{id}_V) = V$

$$\begin{aligned} \mathrm{id}_V\colon V &\longrightarrow V \\ v &\longmapsto v \end{aligned}$$

(iii) $Im(D) = P(\mathbb{R})$

$$\begin{aligned} D\colon P(\mathbb{R}) &\longrightarrow P(\mathbb{R}) \\ p(x) &\longmapsto p\prime(x) \end{aligned}$$

(iv) An example of polynomials with $m = 5$

$$\begin{aligned} D\colon P_5(\mathbb{R}) &\longrightarrow P_5(\mathbb{R}) \\ &\text{Note: } x^5 \notin Im(D_5) \end{aligned}$$

Definition 3.2.4: Surjective

A map $T: V \longrightarrow W$ is surjective if
for any $w \in W$ there is a $v \in V$ such that $T(v) = w$.
i.e., T is surjective if (and only if) $Im(T) = W$.

Theorem 3.2.1 Rank-nullity Theorem (Fundamental Theorem of linear Maps)
Let V be a finite dimensional vector space over $\mathbb{F}$ and $T: V \longrightarrow W$ be a linear map.
Then $Im(T)$ is a finite dimensional vector space, and

$$\dim V = \dim \ker T + \dim Im(T)$$

Proof: Let V be a finite dimensional vector space over $\mathbb{F}$ and $\ker T \subseteq V$ be a subspace.
This means that $\ker T$ is finite dimensional.
Let $\vec{u_1}, \ldots, \vec{u_m}$ be a basis for $\ker T$.
Which means that $\vec{u_1}, \ldots, \vec{u_n}$ is linearly independent in $\ker T$.
Therefore, it also linearly independent in V.
We can extend this list to a full basis $\vec{u_1}, \ldots, \vec{u_n}, \vec{v_1}, \ldots, \vec{v_m}$ for V.
Then $\dim V = n + m$, and $\dim \ker T = n$

Claim: $T(\vec{v_1}), \ldots, T(\vec{v_m})$ is a basis for $Im(T)$.
Thus, if the claim is true, then $Im(T)$ is finite dimensional and $\dim Im(T) = m$.
Thus, $\dim V = \dim \ker T + \dim Im(T)$.

Proof of claim: We need to show that $T(\vec{v_1}), \ldots, T(\vec{v_m})$ is linearly independent in $Im(T)$ and spans $Im(T)$.

(i) $Im(T) = \text{span}\left(T(\vec{v_1}) \ldots, T(\vec{v_m})\right)$: $\supseteq$ definition of span

(ii) We want to prove $\subseteq$.
Let $w \in Im(T)$.
Then there is a $v \in V$ such that $T(v) = w$.
We know that $v = a_1\vec{u_1} + \cdots + a_n\vec{u_n} + b_1\vec{v_1} + \cdots + b_m\vec{v_m}$ for some $a_1, \ldots, a_n, b_1, \ldots, b_m \in \mathbb{F}$.
Then:

$$\begin{aligned} T(v) = T(a_1\vec{u_1} + \cdots + a_n\vec{u_n} + b_1\vec{v_1} + \cdots + b_m\vec{v_m}) &= T(a_1\vec{v_1}) + \cdots + T(a_n\vec{v_n}) + T(b_1\vec{v_1}) + \cdots + T(b_m\vec{v_m}) \\ &= a_1T(\vec{v_1}) + \cdots + a_nT(\vec{v_n}) + b_1T(\vec{v_1}) + \cdots + b_mT(\vec{v_m}) \\ &\text{we know that } T(\vec{u_1}) = \cdots = T(\vec{u_n}) = \vec{0}_W \\ &= b_1T(\vec{v_1}) + \cdots + b_mT(\vec{v_m}) \\ &\in \text{span}\left(T(\vec{v_1}), \ldots, T(\vec{v_m})\right) \end{aligned}$$

Thus, this shows that $Im(T) \subseteq \text{span}\left(T(\vec{v_1}), \ldots, T(\vec{v_m})\right)$.

(iii) $T(\vec{v_1}), \ldots, T(\vec{v_m})$ are linearly independent in $Im(T)$:
Suppose that $c_1T(\vec{v_1}) + \cdots + c_mT(\vec{v_m}) = \vec{0}_W$ for some $c_1, \ldots, c_m \in \mathbb{F}$.
Thus, $T(c_1\vec{v_1}) + \cdots + T(c_m\vec{v_m}) = \vec{0}_W$.
Then $T(c_1\vec{v_1} + \cdots + c_m\vec{v_m}) = \vec{0}_W$.
Hence, $c_1\vec{v_1} + \cdots + c_m\vec{v_m} \in \ker T = \text{span}\left(\vec{u_1}, \ldots, \vec{u_n}\right)$.
Thus, $c_1\vec{v_1} + \cdots + c_m\vec{v_m} = d_1\vec{u_1} + \cdots + d_n\vec{u_n}$ for some $d_1, \ldots, d_n \in \mathbb{F}$.
Then $d_1\vec{u_1} + \cdots + d_n\vec{u_n} - c_1\vec{v_1} - \cdots - c_m\vec{v_m} = \vec{0}_V$.

Since $\vec{u}_1, \ldots, \vec{u}_n, \vec{v}_1, \ldots, \vec{v}_m$ are linearly independent in V,

it follows that $d_1 = \cdots = d_n = c_1 = \cdots = c_m = 0$.

Thus, $T(\vec{v}_1), \ldots, T(\vec{v}_m)$ are linearly independent in $Im(T)$.

As we have shown that $T(\vec{v}_1), \ldots, T(\vec{v}_m)$ are linearly independent in $Im(T)$ and span $Im(T)$, we have proven the claim.

😄

Thus, we have proven the theorem. 😄

Application: Suppose we have a system of linear equations:
Variables $x_1, \ldots, x_n$, $a_{i,j} \in \mathbb{R}$
Then we can write this as a matrix equation:

$$\begin{bmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n = 0_{\mathbb{R}} \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n = 0_{\mathbb{R}} \end{bmatrix}$$

Thus, there are m equations.

One solution: Let $x_1 = \cdots = x_n = 0_{\mathbb{R}}$.
Then the system is satisfied.
But are there others?

Rephrase: Let's rephrase this in terms of linear maps:

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

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{bmatrix}$$

We can check T is linear!
Thus, $x_1 = \cdots = x_n = 0$ is $\vec{0}_{\mathbb{R}} \in \ker T$.

Rank-nullity: By the theorem, we know that $\underbrace{\dim \mathbb{R}^n}_{n} = \dim \ker T + \underbrace{\dim Im(T)}_{\leq m}$.

Thus, $n \leq \dim \ker T + m$.
As such, $\dim \ker T \geq n - m$
Suppose that $n - m > 0$ (more variables than equations).
Therefore, $\dim \ker T > 0$.
Meaning that there are non-zero solutions to the system of equations.

Note:-

Is $\ker T = \left\{\vec{0}_{\mathbb{R}^n}\right\} = \left\{\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}\right\}$?

Or is there something else?

Theorem 3.2.2

Let V, W be a finite dimensional vector space over $\mathbb{F}$ and $\dim V > \dim W$.

Then any linear map $T : V \longrightarrow W$ is not injective, i.e., $\ker T \neq \left\{\vec{0}_V\right\}$.

Proof:

$$\begin{aligned}\dim\ker T &= \dim V - \dim Im(T) \quad \text{By Rank-nullity}\\ &\geq \dim V - \dim W \quad \text{Since } Im(T) \subseteq W \implies \dim Im(T) \leq \dim W\\ &> 0 \quad \text{by hypothesis}\end{aligned}$$

Thus, $\ker T \neq \left\{\vec{0}_V\right\}$. 😄

Note:-

Going back to systems of linear equations:

$$\begin{aligned}\text{Theorem} &\implies \text{if } n > m \text{ then } T\colon \mathbb{R}^n \to \mathbb{R}^m \text{ is not injective}\\ &\implies \ker T \neq \left\{\vec{0}_{\mathbb{R}^n}\right\}\\ &\implies \text{there are non-zero solutions to the system of equations}\end{aligned}$$

Look at Axler 3.24 and 3.27 for more information.

3.3 Matrix of a linear map

Definition 3.3.1: Matrix of a linear map

Let V, W be finite dimensional vector spaces over $\mathbb{F}$, and $T \in \mathscr{L}(V, W)$.
Choose basis:

$$\vec{v}_1, \ldots, \vec{v}_n \text{ for } V$$
$$\vec{w}_1, \ldots, \vec{w}_m \text{ for } W$$

Now, we can write:

$$T(\vec{v}_1) \in W = \text{span}\left(\vec{w}_1, \ldots, \vec{w}_m\right) \implies T(\vec{v}_1) = a_{1,1}\vec{w}_1 + \cdots + a_{m,1}\vec{w}_m, a_{i,1} \in \mathbb{F}$$
$$\vdots$$
$$T(\vec{v}_n) \in W = \text{span}\left(\vec{w}_1, \ldots, \vec{w}_m\right) \implies T(\vec{v}_n) = a_{1,n}\vec{w}_1 + \cdots + a_{m,n}\vec{w}_m, a_{i,n} \in \mathbb{F}$$

Recall: A linear map is determined by what it does to a basis.
This implies that the array of coefficients in $\mathbb{F}$ determines T:

$$\begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}$$

This is called the matrix of T, with respect to the bases $\vec{v}_1, \ldots, \vec{v}_n$ and $\vec{w}_1, \ldots, \vec{w}_m$.
Where, the above is an $m \times n$ matrix, where m is the number of rows and n is the number of columns.

Note:-

Notation:

$$\mathcal{M}(T, (\vec{v}_1, \ldots, \vec{v}_n), (\vec{w}_1, \ldots, \vec{w}_m)) \text{ or}$$
$$\mathcal{M}(T)$$

Example 3.3.1

Let $T: \mathbb{R}^2 \to \mathbb{R}^3$ be a linear map.
With $(x, y) \mapsto (x + 3y, 2x + 5y, 7x + 9y)$.
Choose standard bases:

$$\underbrace{(1,0)}_{V_1} \text{ and } \underbrace{(0,1)}_{V_2} \text{ for } \mathbb{R}^2$$
$$\underbrace{(1,0,0)}_{W_1}, \underbrace{(0,1,0)}_{W_2}, \underbrace{(0,0,1)}_{W_3} \text{ for } \mathbb{R}^3$$

Then, we can write:

$$T(v_1) = T((1,0)) = (1,2,7) = 1 \cdot_{\mathbb{R}} w_1 + 2 \cdot_{\mathbb{R}} w_2 + 7 \cdot_{\mathbb{R}} w_3$$
$$T(v_2) = T((0,1)) = (3,5,9) = 3 \cdot_{\mathbb{R}} w_1 + 5 \cdot_{\mathbb{R}} w_2 + 9 \cdot_{\mathbb{R}} w_3$$

Thus, $\mathcal{M}(T) = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{bmatrix}$.

Example 3.3.2

Differentiation:

$$\begin{aligned} D &\in \mathscr{L}(P_3(\mathbb{R}), P_2(\mathbb{R})) \\ D(p(x)) &= p`(x) \end{aligned}$$

Check bases:

$$\begin{aligned} \underbrace{1, x, x^2, x^3}_{V_1, V_2, V_3, V_4} &\text{ for } P_3(\mathbb{R}) \\ \underbrace{1, x, x^2}_{W_1, W_2, W_3} &\text{ for } P_2(\mathbb{R}) \end{aligned}$$

Then:

$$\begin{aligned} D(v_1) = D(1) = 0 &= 0 \cdot_{\mathbb{R}} 1 + 0 \cdot_{\mathbb{R}} x + 0 \cdot_{\mathbb{R}} x^2 \\ D(v_2) = D(x) = 1 &= 1 \cdot_{\mathbb{R}} 1 + 0 \cdot_{\mathbb{R}} x + 0 \cdot_{\mathbb{R}} x^2 \\ D(v_3) = D(x^2) = 2x &= 0 \cdot_{\mathbb{R}} 1 + 2 \cdot_{\mathbb{R}} x + 0 \cdot_{\mathbb{R}} x^2 \\ D(v_4) = D(x^3) = 3x^2 &= 0 \cdot_{\mathbb{R}} 1 + 0 \cdot_{\mathbb{R}} x + 3 \cdot_{\mathbb{R}} x^2 \end{aligned}$$

Thus,

$$\mathcal{M}(D) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Addition of Matrices: Let V, W be finite dimensional vector spaces over $\mathbb{F}$.

If $S, T \in \mathscr{L}(V, W)$, then define $S + T \in \mathscr{L}(V, W)$ by:

$$(S + T)(v) \coloneqq S(v) + T(v)$$

What is the matrix of $S + T$?

Choose bases $\vec{v}_1, \ldots, \vec{v}_n$ for V and $\vec{w}_1, \ldots, \vec{w}_m$ for W.

Then:

$$\begin{aligned} T(v_k) &= a_{1,k}\vec{w}_1 + \cdots + a_{m,k}\vec{w}_m \quad 1 \leq k \leq n \\ S(v_k) &= b_{1,k}\vec{w}_1 + \cdots + b_{m,k}\vec{w}_m \quad 1 \leq k \leq n \end{aligned}$$

Thus,

$$\mathcal{M}(T, \{v\text{'s}\}, \{w\text{'s}\}) = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}$$

And,

$$\mathcal{M}(S, \{v\text{'s}\}, \{w\text{'s}\}) = \begin{bmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{bmatrix}$$

Therefore:

$$\begin{aligned} (S+T)(v_k) &= S(v_k) + T(v_k) \\ &= (b_{1,k}\vec{w}_1 + \cdots + b_{m,k}\vec{w}_m) + (a_{1,k}\vec{w}_1 + \cdots + a_{m,k}\vec{w}_m) \\ &= (a_{1,k} + b_{1,k})\vec{w}_1 + \cdots + (a_{m,k} + b_{m,k})\vec{w}_m \end{aligned}$$

Thus,

$$\mathcal{M}(S+T, \{v\text{'s}\}, \{w\text{'s}\}) = \begin{bmatrix} a_{1,1} + b_{1,1} & \cdots & a_{1,n} + b_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} + b_{m,1} & \cdots & a_{m,n} + b_{m,n} \end{bmatrix}$$

So we define addition of matrices so that:

$$\mathcal{M}(S) \underbrace{+}_{\text{we defined this!}} \mathcal{M}(T) := \mathcal{M}(S +_{\mathscr{L}(V,W)} T)$$

Scalar Multiplication: Let $T \in \mathscr{L}(V, W)$ with bases $\vec{v}_1, \ldots, \vec{v}_n$ for V and $\vec{w}_1, \ldots, \vec{w}_m$ for W.
Remember that $M(T) = M(T, (\vec{v}_1, \ldots, \vec{v}_n), (\vec{w}_1, \ldots, \vec{w}_m))$.
This looks like:

$$\begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}$$

i.e,. $T(\vec{v}_k) = a_{1,k}\vec{w}_1 + \cdots + a_{m,k}\vec{w}_m$.
Then, for $\lambda \in \mathbb{F}$, we define $\lambda T \in \mathscr{L}(V, W)$ by:

$$\lambda \cdot M(T) := M(\lambda T)$$

We compute:

$$\begin{aligned} (\lambda \cdot T)(\vec{v}_k) &:= \lambda \cdot T(\vec{v}_k) \\ &= \lambda \cdot (a_{1,k}\vec{w}_1 + \cdots + a_{m,k}\vec{w}_m) \\ &\implies M(\lambda \cdot T) = \begin{bmatrix} \lambda a_{1,1} & \cdots & \lambda a_{1,n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m,1} & \cdots & \lambda a_{m,n} \end{bmatrix} \end{aligned}$$

In other words:

$$\lambda \cdot \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} = \begin{bmatrix} \lambda\lambda a_{1,1} & \cdots & \lambda a_{1,n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m,1} & \cdots & \lambda a_{m,n} \end{bmatrix}$$

Notational shift: Let $F^{m,n} := \{m \times n \text{ matrices with entries in } \mathbb{F}\}$
Having addition + scalar multiplication implies that $F^{m,n}$ is a vector space over $\mathbb{F}$.
Soon: $F^{m,n} \cong \mathscr{L}(\mathbb{F}^n, \mathbb{F}^m)$.

Composition of maps:

$$U \xrightarrow{S} V \xrightarrow{T} Z$$

Now pick bases: $\vec{u}_1, \ldots, \vec{u}_p$ for U, $\vec{v}_1, \ldots, \vec{v}_n$ for V, and $\vec{w}_1, \ldots, \vec{w}_m$ for W.
Let $j = 1, \ldots, p$ and $k = 1, \ldots, n$.
Then:

$$S(\vec{u}_j) = b_{1,j}\vec{v}_1 + \cdots + b_{n,j}\vec{v}_n$$
$$T(\vec{v}_k) = a_{1,k}\vec{w}_1 + \cdots + a_{m,k}\vec{w}_m$$

Now remember, $M(S, (\vec{u}_1, \ldots, \vec{u}_p), (\vec{v}_1, \ldots, \vec{v}_n)) = M(S)$

$$M(S) = \begin{bmatrix} b_{1,1} & \cdots & b_{1,p} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,p} \end{bmatrix}$$

And $M(T, (\vec{v}_1, \ldots, \vec{v}_n), (\vec{w}_1, \ldots, \vec{w}_m)) = M(T)$

$$M(T) = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}$$

Now, let's define $M(T) \cdot M(S) \coloneqq M(T \circ S)$
What is $M(T \circ S)$?
We know that $T \circ S \in \mathscr{L}(U, W)$ with bases $\vec{u}_1, \ldots, \vec{u}_p$ for U and $\vec{w}_1, \ldots, \vec{w}_m$ for W.
Let's look how the j^{th} column of $M(T \circ S)$ is determined by $M(S)$ and $M(T)$.

$$\begin{aligned}
(T \circ S)(\vec{u}_j) &= T(S(\vec{u}_j)) \\
&= T(b_{1,j}\vec{v}_1 + \cdots + b_{n,j}\vec{v}_n) \\
&= b_{1,j}T(\vec{v}_1) + \cdots + b_{n,j}T(\vec{v}_n) \quad \text{by linearity} \\
&= b_{1,j}(a_{1,1}\vec{w}_1 + \cdots + a_{m,1}\vec{w}_m) + \cdots + b_{n,j}(a_{1,n}\vec{w}_1 + \cdots + a_{m,n}\vec{w}_m) \\
&= (a_{1,1} \cdot b_{1,j} + \cdots + a_{1,n} \cdot b_{n,j})\vec{w}_1 + \cdots + (a_{m,1} \cdot b_{1,j} + \cdots + a_{m,n} \cdot b_{n,j})\vec{w}_m
\end{aligned}$$

All told:

$$(T \circ S)(\vec{u}_j) = \left(\sum_{k=1}^{n} a_{1,k} \cdot b_{k,j}\right)\vec{w}_1 + \left(\sum_{k=1}^{n} a_{2,k} \cdot b_{k,j}\right)\vec{w}_2 + \cdots + \left(\sum_{k=1}^{n} a_{m,k} \cdot b_{k,j}\right)\vec{w}_m$$

So for the j^{th} column of $M(T \circ S)$, we have:

$$M(T \circ S, (\vec{u}_1, \ldots, \vec{u}_p), (\vec{w}_1, \ldots, \vec{w}_m)) = \begin{bmatrix} \sum_{k=1}^{n} a_{1,k} \cdot b_{k,j} \\ \sum_{k=1}^{n} a_{2,k} \cdot b_{k,j} \\ \vdots \\ \sum_{k=1}^{n} a_{m,k} \cdot b_{k,j} \end{bmatrix}$$

Thus, the ij-th entry of $M(T \circ S)$ is $\sum_{k=1}^{n} a_{i,k} \cdot b_{k,j}$.
So the matrix multiplication looks like:

$$\underbrace{\begin{bmatrix} a_{i,j} \end{bmatrix}}_{m \times n \text{ matrix}} \cdot \underbrace{\begin{bmatrix} b_{i,j} \end{bmatrix}}_{n \times p \text{ matrix}} = \underbrace{\begin{bmatrix} \sum_{k=1}^{n} a_{i,k} \cdot b_{k,j} \end{bmatrix}}_{m \times p \text{ matrix}}$$

Theorem 3.3.1 Matrix multiplication is associative

Let $A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$, where $a_{i,j} \in \mathbb{R}$.

$$T_A : \mathbb{R}^n \to \mathbb{R}^m$$
$$e_k = \underbrace{(0, \ldots, 1, \ldots, 0)}_{k^{\text{th}} \text{ place}} \mapsto (a_{1,k}, a_{2,k}, \ldots, a_{m,k})$$
$$A = M(T_A, \text{standard basis of } \mathbb{R}^2, \text{standard basis of } \mathbb{R}^m)$$

Let A, B, C be matrices with $m \times n, n \times p, p \times r$ dimensions respectively with entries in $\mathbb{R}$.
Then:

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C$$

Proof: Let

$$\begin{aligned} T_A &: \mathbb{R}^m \to \mathbb{R}^n \\ A &= M(T_A) \\ T_B &: \mathbb{R}^n \to \mathbb{R}^p \\ B &= M(T_B) \\ T_C &: \mathbb{R}^p \to \mathbb{R}^r \\ C &= M(T_C) \end{aligned}$$

Then:

$$\begin{aligned} A \cdot (B \cdot C) &= M(T_A) \cdot (M(T_B) \cdot M(T_C)) \\ &= M(T_A) \cdot M(T_B \circ T_C) \\ &= M(T_A \circ (T_B \circ T_C)) \\ &= M((T_A \circ T_B) \circ T_C) \\ &= M(T_A \circ T_B) \cdot M(T_C) \\ &= (M(T_A) \cdot M(T_B)) \cdot M(T_C) \\ &= (A \cdot B) \cdot C \end{aligned}$$

☺

3.4 Invertible Linear Maps

Definition 3.4.1: Invertible

T is invertible if there is a $S \in \mathscr{L}(W,V)$ such that $T \circ S = Id_W$ and $S \circ T = Id_V$.
Then we declare $S :=$ the inverse of T, and write $S = T^{-1}$.
Inverses if they exist are unique.

Reason: Say $S_1, S_2 \in \mathscr{L}(W,V)$ are inverses for $T \in \mathscr{L}(V,W)$.
Then:

$$\begin{aligned} S_1 = S_1 \circ Id_W &\underbrace{=}_{\text{since } S_2 \text{ is an inverse}} S_1 \circ (T \circ S_2) \\ &= (S_1 \circ T) \circ S_2 \underbrace{=}_{\text{since } S_1 \text{ is an inverse}} Id_V \circ S_2 \\ &= S_2 \end{aligned}$$

Theorem 3.4.1

Let $T \in \mathscr{L}(V,W)$ is invertible if and only if T is bijective (injective and surjective).

Proof of $\implies$: Say that $T \in \mathscr{L}(V,W)$ is invertible. Let $T^{-1}: W \to V$ be the inverse.

(i) T is injective: Suppose that for some $u, v \in V$, we have $T(u) = T(v)$.
Then:

$$u = T^{-1}(T(u)) = T^{-1}(T(v)) = v$$

(ii) T is surjective: Let $w \in W$ be arbitrary.

$$w = T(T^{-1}(w)) \implies w \in Im(T)$$

So $W \subseteq Im(T)$.

Thus, T is bijective.

😊

Proof of $\impliedby$: Say that $T \in \mathscr{L}(V,W)$ is bijective i.e., T is injective and surjective.
Let's construct an inverse:

$$S: W \to V$$
$$w \text{ the unique } v \in V \text{ such that } T(v) = w$$

Thus, the existence of v is guaranteed by surjectivity.
And the uniqueness of v is guaranteed by injectivity.

Check: We have three things to check:

(i) $T \circ S = Id_W$ i.e., $T(S(w)) = w$ for all $w \in W$.
Then $T(S(w)) = T(v)$ where $v \in V$ is the unique vector such that $T(v) = w$.
Thus, $T(S(w)) = w$.

(ii) $S \circ T = Id_v$. We want $S(T(v)) = v$ for all $v \in V$.

$$\begin{aligned} T(S(T(v))) &= (T \circ (S \circ T))(v) \quad T \text{ injective} \\ &= ((T \circ S) \circ T)(v) \quad \implies S(T(v)) = v \\ &= (T \circ S)(T(v)) \\ &= Id_W(T(v)) \\ &= T(v) \end{aligned}$$

(iii) We need to check that S is linear.

(a) Additivity:
One on hand we have:

$$\begin{aligned} T(S(w_1) + S(w_2)) &= T(S(w_1)) + T(S(w_2)) \quad T \text{ is linear} \\ &= Id_W(w_1) + Id_W(w_2) \\ &= w_1 + w_2 \end{aligned}$$

On the other hand:

$$\begin{aligned} T(S(w_1) + S(w_2)) &= (T \circ S)(w_1 + w_2) \\ &= Id_W(w_1 + w_2) \\ &= w_1 + w_2 \end{aligned}$$

As T is injective, we know that $S(w_1) + S(w_2) = S(w_1 + w_2)$.

(b) Homogeneity: So on one hand we have:

$$\begin{aligned} T(\lambda \cdot S(w)) &= \lambda \cdot T(S(w)) \\ &= \lambda \cdot (T \circ S)(w) \\ &= \lambda \cdot Id_W(w) \\ &= \lambda \cdot w \end{aligned}$$

On the other hand:

$$\begin{aligned} T(S(\lambda \cdot w)) &= (T \circ S)(\lambda \cdot w) \\ &= Id_W(\lambda \cdot w) \\ &= \lambda \cdot w \end{aligned}$$

Since T is injective, we know that $S(\lambda \cdot w) = \lambda \cdot S(w)$.

😄

As we have proven both directions, we have proven the theorem.

Definition 3.4.2

An invertible linear map $T \in \mathscr{L}(V, W)$ is called an isomorphism between V and W.
Notation: $V \cong W$.

Proposition 3.4.1

Say V, W are finite dimensional vector spaces over $\mathbb{F}$, and $V \cong W$.
Then $\dim V = \dim W$.

Proof: If $V \cong W$, then there is an invertible linear map $T : V \longrightarrow W$.
By the rank-nullity theorem, we know that:

$$\begin{aligned}
\dim V &= \dim \ker T + \dim Im(T) \\
&\quad \text{Since } T \text{ is invertible, we know that } \dim \ker T = 0 \\
&= 0 + \dim Im(T) \\
&\quad \text{Since } T \text{ is surjective, we know that } \dim Im(T) = \dim W \\
&= 0 + \dim W \\
&= \dim W
\end{aligned}$$

😄

Converse is also true (Axler 3.5): If V, W are finite dimensional vector spaces over $\mathbb{F}$ and $\dim V = \dim W$, then $V \cong W$.

Proof: Let $\vec{v}_1, \ldots, \vec{v}_n$ be a basis for V.
And let $\vec{w}_1, \ldots, \vec{w}_n$ be a basis for W.
Define a linear map $T : V \longrightarrow W$ by setting $T(\vec{v}_i) = \vec{w}_i, 1 \leq i \leq n$.

T ***is surjective:*** Let $\vec{w} \in W$ be arbitrary.
Then:

$$\begin{aligned}
\vec{w} &= a_1\vec{w}_1 + \cdots + a_n\vec{w}_n, a_i \in \mathbb{F} \\
&= a_1T(\vec{v}_1) + \cdots + a_nT(\vec{v}_n) \\
&= T(a_1\vec{v}_1) + \cdots + T(a_n\vec{v}_n) \\
&= T(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n)
\end{aligned}$$

This implies that $w \in ImT$, so $W \subseteq ImT$.
Thus, T is surjective.

T ***is injective:*** By rank-nullity, we know that:

$$\begin{aligned}
\dim V &= \dim \ker T + \dim Im(T) \\
&= \ \text{since } T \text{ is surjective, we know that } \dim Im(T) = \dim W \\
&\implies \dim \ker T = 0 \quad \text{since } \dim V = \dim W \\
&\implies \ker T = \left\{\vec{0}_V\right\}
\end{aligned}$$

Thus, T is injective.
Thus, we have shown that T is bijective, and thus T is an isomorphism.

😄

Example 3.4.1

We know that $P_3(\mathbb{C})$ and $\mathbb{C}^4$ are isomorphic.

Proof gives us:

$$\begin{aligned} T: 1 &\mapsto (1,0,0,0) \\ x &\mapsto (0,1,0,0) \\ x^2 &\mapsto (0,0,1,0) \\ x^3 &\mapsto (0,0,0,1) \end{aligned}$$

Under this map, for some $a_0, a_1, a_2, a_3 \in \mathbb{C}$, we have:

$$\begin{aligned} T(a_0 + a_1 x + a_2 x^2 + a_3 x^3) &= a_0 T(1) + a_1 T(x) + a_2 T(x^2) + a_3 T(x^3) \\ &= a_0 \cdot (1,0,0,0) + a_1 \cdot (0,1,0,0) + a_2 \cdot (0,0,1,0) + a_3 \cdot (0,0,0,1) \\ &= (a_0, a_1, a_2, a_3) \end{aligned}$$

Example 3.4.2

Let V, W be finite dimensional vector spaces over $\mathbb{F}$.
Choose bases $\vec{v}_1, \ldots, \vec{v}_n$ for V and $\vec{w}_1, \ldots, \vec{w}_m$ for W.
Let's define:

$$\begin{aligned} M : \mathscr{L}(V,W) &\longrightarrow F^{m,n} \\ T &\mapsto M(T, (\vec{v}_1, \ldots, \vec{v}_n), (\vec{w}_1, \ldots, \vec{w}_m)) \end{aligned}$$

Now, recall that M is linear since:

$$\begin{aligned} M(T +_{\mathscr{L}(V,W)} S) &= M(T) +_{\mathbb{F}^{m,n}} M(S) \\ M(\lambda \cdot_{\mathscr{L}(V,W)} T) &= \lambda \cdot_{\mathbb{F}^{m,n}} M(T) \end{aligned}$$

Now, by Axler 3.60, M is an isomorphism.
By PSET 6, $\dim F^{m,n} = mn$.
This implies that $\dim \mathscr{L}(V,W) = \dim V \cdot \dim W$.

Definition 3.4.3: Endomorphisms (Linear opeartions)

A linear map: $T: V \longrightarrow V$ is called an endomorphism or a linear operation of V.

Notation: $\mathscr{L}(V) := \mathscr{L}(V,V)$

Example 3.4.3

Here are some examples:

(i)

$$\begin{aligned} T : P(\mathbb{R}) &\longrightarrow P(\mathbb{R}) \\ p(x) &\mapsto x^2 p(x) \end{aligned}$$

Note, that this map is injective but not surjective.

(ii)

$$S : \mathbb{C}^\infty \to \mathbb{C}^\infty$$
$$(x_1, x_2, x_3, \ldots) \mapsto (x_2, x_3, \ldots)$$

Note, that this map is surjective but not injective.

Theorem 3.4.2

Let V be a finite dimensional vector space over $\mathbb{F}$.
Let $T \in \mathscr{L}(V)$.
Then the following are equivalent:

(i) T is injective

(ii) T is surjective

(iii) T is invertible

We are going to prove $(i) \implies (ii) \implies (iii) \implies (i)$.

Proof of $(iii) \implies (i)$ ***:*** We have already proven this in class.

☺

Proof of $(i) \implies (ii)$ ***:*** Assume that T is injective.
Then, we know that $\ker T = \left\{\vec{0}_V\right\}$
By rank-nullity, we know that $\dim V = \dim \ker T + \dim Im(T)$.
Thus, $\dim V = 0 + \dim Im(T)$, so $\dim V = \dim Im(T)$.
Since $T \in \mathscr{L}(V)$, we know that $Im(T) \subseteq V$.
By Axler 2.C.1, we know that $Im(T) = V$.
Thus, T is surjective. ☺

Proof of $(ii) \implies (iii)$ ***:*** Now assume that T is surjective.
Then $Im(T) = V$
By rank-nullity, we know:

$$\begin{aligned} \dim V &= \dim \ker T + \dim Im(T) \\ &= \dim \ker T + \dim V \\ &\implies \dim \ker T = 0 \\ &\implies \ker T = \left\{\vec{0}_V\right\} \\ &\implies T \text{ is injective} \implies T \text{ is bijective} \implies T \text{ is invertible} \end{aligned}$$

Thus, T is invertible as desired.

☺

As we have proven all three directions, we have proven the theorem.

Corollary 3.4.1

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

$$T \text{ is invertible} \iff T \text{ is injective} \iff T \text{ is surjective}$$

Question 1

Show that, given $q(x) \in P(\mathbb{R})$, there exists another polynomial $p(x)$ such that:

$$q(x) = \left[\left(x^2 + 2x + 3\right) \cdot p(x)\right]''$$

Solution: First, make everything finite-dimensional. Say $q(x)$ has degree m.
Now let's define:

$$\begin{aligned} T \colon P_m(\mathbb{R}) &\to P_m(\mathbb{R}) \\ p(x) &\mapsto \left[\left(x^2 + 2x + 3\right) \cdot p(x)\right]'' \end{aligned}$$

Exercise: Show that T is linear.
We want to show that T is surjective.

Claim: T is injective

Proof of claim: The kernel consists of $p(x)$ such that $\left[\left(x^2 + 2x + 3\right) \cdot p(x)\right]'' = 0$.
Thus, it must have the form $[ax + b]'$.
Thus, we need $\left(x^2 + 2x + 3\right) \cdot p(x)$ to have the form $ax + b$.

$$\deg((x^2 + 2x + 3) \cdot p(x)) \geqslant 2 \text{ as long as } p(x) \neq 0$$
$$\deg(ax + b) \leq 1$$

Thus, the only way for this to be true is if $\ker T = \left\{0_{P_m(\mathbb{R})}\right\}$.
This implies that T is injective.
Then, by the previous theorem, we know that if T is injective, then T is surjective.
Thus, given $q(x) \in P(\mathbb{R})$, there exists another polynomial $p(x)$ such that $T(p(x)) = q(x)$.
Therefore, $\left[\left(x^2 + 2x + 3\right) \cdot p(x)\right]'' = q(x)$.

😊

Linear Maps as Matrix multiplication: Let V be a finite dimensional vector space over $\mathbb{F}$.
Let $\vec{v}_1, \ldots, \vec{v}_n$ be a basis for V.
Now for any $v \in V$, we can write for some scalars $c_1, \ldots, c_n \in \mathbb{F}$:

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

Let's define:

$$M(v) := \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Example 3.4.4

Let $V = P_3(\mathbb{R})$ with basis $1, x, x^2, x^3$.
Then,

$$v = 2 - 7x + 5x^3 = 2 \cdot 1 - 7 \cdot x + 0 \cdot x^2 + 5 \cdot x^3$$

Or in other words:

$$M(v) = \begin{bmatrix} 2 \\ -7 \\ 0 \\ 5 \end{bmatrix}$$

Note: $M(v_0 + w_0) = M(v_0) + M(w_0)$ and $M(\lambda v) = \lambda M(v)$.
Say that $T \in \mathscr{L}(V, W)$.
Let $\vec{w_1}, \ldots, \vec{w_m}$ be a basis for W.
Then, for any $v \in V$, we can write:

$$M(T(u)) = M(T) \cdot M(u)$$

In other words, linear maps act like matrix multiplication.
We can say:

$$M(T, (\vec{v_1}, \ldots, \vec{v_n}), (\vec{w_1}, \ldots, \vec{w_m})) = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix}$$

Then:

$$\begin{aligned} T(v) &= T(c_1\vec{v_1} + \cdots + c_n\vec{v_n}) \\ &= c_1T(\vec{v_1}) + \cdots + c_nT(\vec{v_n}) \end{aligned}$$

Which implies $M(T(v)) = c_1M(T(\vec{v_1})) + \cdots + c_nM(T(\vec{v_n}))$.
On the other hand, we have:

$$T(v_k) = a_{1,k}\vec{w_1} + \cdots + a_{m,k}\vec{w_m}$$

Now, $M(T(v_k)$ is the k^{th} column of $M(T)$.

$$\begin{bmatrix} a_{1,k} \\ \vdots \\ a_{m,k} \end{bmatrix}$$

Thus, we have:

$$M(T(v)) = \begin{bmatrix} c_1a_{1,1} \\ \vdots \\ c_1a_{m,1} \end{bmatrix} + \cdots + \begin{bmatrix} c_na_{1,n} \\ \vdots \\ c_na_{m,n} \end{bmatrix} = \begin{bmatrix} c_1a_{1,1} + \cdots + c_na_{1,n} \\ \vdots \\ c_1a_{m,1} + \cdots + c_na_{m,n} \end{bmatrix} = M(T) \cdot M(v)$$

Row Reduction *I* ***over*** $\mathbb{F}$ ***:*** System of m linear equations with n unknowns: $x_1, \ldots, x_n$

$$\begin{bmatrix} a_{1,1}x_1 + \ldots + a_{1,n}x_n = b_1 \\ \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n = b_m \end{bmatrix}, \quad a_{i,j}, b_k \in \mathbb{F}$$

Which can be written as a Matrix:

$$\underbrace{\begin{bmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{bmatrix}}_{A \in \mathbb{F}^{m,n}} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{\vec{x} \in \mathbb{F}^{n,1}} = \underbrace{\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}}_{\vec{b} \in \mathbb{F}^{m,1}}$$

Then we can define:

$$\begin{aligned} T_A : \mathbb{F}^n &\mapsto \mathbb{F}^m \quad \text{linear map} \\ \vec{x} &\mapsto A\vec{x} = \vec{b} \end{aligned}$$

Question is: $Is \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \text{image}(T_A)$?

Row operations are used on the augmented matrix:

$$[A \mid B] = \left[\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_m \end{array}\right] \in \mathbb{F}^{m,n+1}$$

to simplify the original systems of equations.

Need elementary matrices to express row operations: $E \in \mathbb{F}^{m,m}$

Thus, we get three types:

(i) Where $a \in \mathbb{F}$ is in position i, j

$$E = \begin{bmatrix} 1 & & \\ & \ddots & a \\ & & 1 \end{bmatrix} \text{ or } \begin{bmatrix} 1 & & \\ & \ddots & \\ & a & 1 \end{bmatrix}$$

Which means $E \cdot A$: modify A by adding $a \cdot (\text{row } j)$ to row i.

(ii) Given:

$$a_{i,i} \mapsto 0 \qquad a_{i,j} \mapsto 1$$
$$a_{j,j} \mapsto 0 \qquad a_{j,i} \mapsto 1$$

Then:

$$E = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

Thus, $E \cdot A$: modify A by exchanging rows i and j.

(iii) Given: $a_{i,i} \mapsto c \in \mathbb{F}, c \neq 0$

Thus,

$$E = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$

Meaning, $E \cdot A$: modify A by multiplying row i by c.

Example 3.4.5

(i)

$$\underbrace{\begin{bmatrix}1 & 7 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}}_{E_{(i)}}\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix} = \begin{bmatrix}29 & 37 & 45\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}$$

(ii)

$$\underbrace{\begin{bmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{bmatrix}}_{E_{(ii)}}\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix} = \begin{bmatrix}7 & 8 & 9\\4 & 5 & 6\\1 & 2 & 3\end{bmatrix}$$

(iii)

$$E = \underbrace{\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 3\end{bmatrix}}_{E_{(iii)}}\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix} = \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\21 & 24 & 27\end{bmatrix}$$

Lenma 3.4.1

Elementary matrices are invertible:
if E is an elementary matrix, then there exists a matrix E^{-1} such that $E \cdot E^{-1} = E^{-1} \cdot E = I$.

Proof: By EXAMPLES LOl:

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

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

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

😄

Upshot: Elementary row operations $\overset{\text{1-1}}{\Longleftrightarrow}$ Elementary matrices.

Example 3.4.6

$$
\begin{aligned}
A = \begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 1 & 1 & 2 & 6 & 10 \\ 1 & 2 & 5 & 2 & 7 \end{bmatrix} &\xrightarrow{-R_1+R_2 \mapsto R_2} \begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 5 & 5 \\ 1 & 2 & 5 & 2 & 7 \end{bmatrix} \\
&\xrightarrow{-R_1+R_3 \mapsto R_3} \begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 3 & 1 & 2 \end{bmatrix} \\
&\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 5 & 5 \end{bmatrix} \\
&\xrightarrow{\frac{1}{5}R_3 \mapsto R_3} \begin{bmatrix} 1 & 1 & 2 & 1 & 5 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \\
&\xrightarrow{-R_2+R_1 \mapsto R_1} \begin{bmatrix} 1 & 0 & -1 & 0 & 3 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \\
&\xrightarrow{-R_3+R_2 \mapsto R_2} \begin{bmatrix} 1 & 0 & -1 & 0 & 3 \\ 0 & 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} := A'
\end{aligned}
$$

In other words:

$$
A' = \begin{bmatrix} 1 & 0 & 0 & & & \\ 0 & 1 & -1 & 0 & 0 & 1 \end{bmatrix} \cdot \underbrace{\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{-R_2+R_1 \mapsto R_1}
$$

Note:-

I didn't finish the above but therey are equal

Solving systems of linear equations:

$$
\underbrace{A}_{\mathbb{F}^{m,m}} \cdot \underbrace{\vec{x}}_{\mathbb{F}^n} = \underbrace{B_{\mathbb{F}^m}}
$$

Meaning that the augmented matrix $M = [A \mid B]$

$$
\begin{aligned}
M' &= \underbrace{E_k \cdot \ldots \cdot E_1}_{\text{elementary matrices } (m \times m)} \cdot M \\
&= [A' \mid B'] = \left[\begin{array}{cccc|c} a'_{1,1} & a'_{1,2} & \ldots & a'_{1,n} & b'_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a'_{m,1} & a'_{m,2} & \ldots & a'_{m,n} & b'_m \end{array}\right]
\end{aligned}
$$

Important: $\star \left\{\vec{x} \in \mathbb{F}^n \mid A \cdot \vec{x} = B\right\} = \left\{\vec{x} \in \mathbb{F}^n \mid A' \cdot \vec{x} = B'\right\}$
Meaning, the solutions to our original system of equations are the same as the solutions to our modified system of equations.

Proof: Let $P = E_k \cdot \ldots \cdot E_1$ is invertible.
Where, $P^{-1} = E_1^{-1} \cdot \ldots \cdot E_k^{-1}$
And $I = P^{-1} \cdot P = E_1^{-1} \cdot \ldots \cdot E_k^{-1} \cdot E_k \cdot \ldots \cdot E_1$
Say $\vec{x} \in$ LHS of $\star$:

Where $M' = P \cdot M = \left[\underbrace{P * A}_{A'} \mid \underbrace{P * B}_{B'}\right]$

$$\begin{aligned} A \cdot \vec{x} &= B \\ P \cdot A \cdot \vec{x} &= P \cdot B \\ A' \cdot \vec{x} &= B' \\ &\implies \vec{x} \in \text{RHS}. \end{aligned}$$

Use P^{-1} to show the other direction. 😊

(Reduced) Row-Echelon form: Notation: $M \in \mathbb{F}^{m,n}$, write M_i for the ith row of M.

Definition 3.4.4

$M \in \mathbb{F}^{m,n}$ is in (reduced) row-echelon form if:

(i) If $M_i = (0, \ldots, 0)$ then $M_j = (0, \ldots, 0)$ for all $j > i$.

(ii) If $M_i \neq (0, \ldots, 0)$, then the left most nonzero entry is a 1 (pivot).

(iii) If $M_{i+1} \neq (0, \ldots, 0)$ as well, then the pivot in M_{i+1} is to the right of the pivot in M_i.

(iv) The entries above and below a pivot are 0.

Example 3.4.7

Think $\mathbb{F} = \mathbb{Q}, \mathbb{R}$, or $\mathbb{C}$.

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

Theorem 3.4.3

Let $M \in \mathbb{F}^{m,n}$. There is a sequence of elementary row operations, $E_k, \ldots, E_1$, such that $M' = E_k \cdot \ldots \cdot E_1 \cdot M$ is in row-echelon form.
M' is unique.

Solving systems of linear equations using Row-Echelon matrices : Sat $A \cdot \vec{x} = B \implies M = [A \mid B]$

Suppose that the row-echelon form of M is:

$$M' = [A' \mid B'] = \left[\begin{array}{cccc|c} 1 & 6 & 0 & 1 & 3 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

This would imply that:

$$\begin{aligned} A' \cdot \vec{x} &= B' \\ x_1 + 6x_2 + x_4 &= 0 \\ x_3 + 2x_4 &= 0 \\ 0 &= 1 \end{aligned}$$

Thus, there are no solutions.

If instead we had:

$$M' = [A' \mid B'] = \left[\begin{array}{cccc|c} 1 & 6 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Thus would imply that:

$$\begin{aligned} x_1 + 6x_2 + x_4 =& 1 \\ x_3 + 2x_4 =& 3 \\ 0 =& 0 \end{aligned}$$

Thus, we have solutions!
Let $x_2 = a$ and $x_4 = b$ be constants. Solve for pivot variables:

$$\begin{aligned} & x_1 = 1 - 6a - b \\ & x_3 = 3 - 2b \\ \implies & \vec{x} = (x_1, x_2, x_3, x_4) = (1 - 6a - b, a, 3 - 2b, b) \end{aligned}$$

In general:
Let $M' = [A' \mid B']$ be in row-echelon form.

(i) $A' \cdot \vec{x} = B'$ has no solutions B contains a pivot.

(ii) If B' has no pivot:

  (a) Give the non=pivotal variables constant values.

  (b) Solve for pivot variables.

Lenma 3.4.2

Let $\vec{x_s}$ be a solution to $T(\vec{x}) = \vec{b}$.
Where T is a linear map that maps $\vec{x} \in \mathbb{R}^n$ to $\vec{b} \in \mathbb{R}^m$ by a matrix A: $T(\vec{x}) = A \cdot \vec{x}$.
Then, if there are other solutions, $\vec{x_\star}$, to $T(\vec{x}) = \vec{b}$,
Then there exists an $\vec{x_k} \in \ker T$ such that every other solution is given by:

$$\vec{x_\star} = \vec{x_s} + \vec{x_k}$$

Proof: We know $T(\vec{x_s}) = \vec{b}$ and $T(\vec{x_\star}) = \vec{b}$ as they are solutions to $T(\vec{x}) = \vec{b}$. Consider, $T(\vec{x_\star} - \vec{x_s})$.
Then, by the fact that T is a linear map, the following is true:

$$T(\vec{x_\star} - \vec{x_s}) = T(\vec{x_\star}) - T(\vec{x_s}) = \vec{b} - \vec{b} = \vec{0}$$

.
Therefore, $\vec{x_\star} - \vec{x_s} \in \ker T$.
Thus, there exists $\vec{x_k} \in \ker T$ such that $\vec{x_k} = \vec{x_\star} - \vec{x_s}$.
By definition, $T(\vec{x_k}) = 0$, meaning it is a solution to $T(\vec{x}) = \vec{0}$.
Thus, every other solution to $T(\vec{x}) = \vec{b}$ is given by a solution, $\vec{x_s}$ plus a solution to $T(\vec{x}) = \vec{0}$, $\vec{x_k}$. 😊

Connection to linear maps: Let $A \in \mathbb{R}^{m,n}$.
Then:

$$\begin{aligned} T_A \colon \mathbb{R}^n &\to \mathbb{R}^m \\ \vec{x} &\mapsto A \cdot \vec{x} \end{aligned}$$

Remember:

$$
\begin{aligned}
e_1, \ldots, e_n &\text{ standard basis for } \mathbb{R}^n \\
f_1, \ldots, f_m &\text{ standard basis for } \mathbb{R}^m \\
f_1 &= \underbrace{(1, 0, \ldots, 0)}_{m}
\end{aligned}
$$

We can write:
$A = M(T_A, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$

$$
\begin{aligned}
\ker T_A &= \left\{ \vec{x} \in \mathbb{R}^n \mid A \cdot \vec{x} = \vec{0}_{\mathbb{R}^m} \right\} \\
&= \left\{ \vec{x} \in \mathbb{R}^n \mid A' \cdot \vec{x} = \vec{0}_{\mathbb{R}^m} \right\} \text{ where } A' \text{ is in row-echelon form} \\
&= \ker T_{A'}
\end{aligned}
$$

Example 3.4.8
Let:

$$
A' = \begin{bmatrix} 1 & 6 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{3,4}
$$

This implies:

$$
\begin{aligned}
A' \cdot \vec{x} &= \vec{0}_{\mathbb{R}^3} \\
x_1 + 6x_2 + x_4 &= 0 \\
x_3 + 2x_4 &= 0 \\
0 &= 0
\end{aligned}
$$

Non-pivot variables x_2 and x_4 are free variables.
Say $x_2 = a$ and $x_4 = b$ are constants.
Then solve for pivot variables:

$$
\begin{aligned}
x_1 &= -6a - b \\
x_3 &= -2b
\end{aligned}
$$

Solutions:

$$
(x_1, x_2, x_3, x_4) = (-6a - b, a, -2b, b) = a(-6, 1, 0, 0) + b(-1, 0, -2, 1)
$$

So $\ker T_{A'} = \ker T_A = \{(-6a - b, a, -2b, b) \mid a, b \in \mathbb{R}\} \subseteq \mathbb{R}^4$.

$$
\begin{aligned}
a = 1, b = 0 &\implies (-6, 1, 0, 0) \\
a = 0, b = 1 &\implies (-1, 0, -2, 1)
\end{aligned}
$$

Thus:

$$
\begin{aligned}
T_A &= \operatorname{span}((-6, 1, 0, 0), (-1, 0, -2, 1)) \\
&= \operatorname{span}(-6e_1 + e_2, -e_1 - 2e_3 + e_4).
\end{aligned}
$$

Images: Given $T_A : \mathbb{R}^n \to \mathbb{R}^m$.
Compute a basis $\vec{v}_1, \ldots, \vec{v}_r$ for the kernel.
Let $i_1, \ldots, i_{n-r}$be the indices of the pivot columns of A'

Claim: $\vec{v}_1, \ldots, \vec{v}_n, e_i, \ldots, e_{n-r}$ is a basis for $\mathbb{R}^n$ (see notes).
Assume claim.
Proof of rank-nullity shows that $T(e_i), \ldots T(e_{n-r})$ is a basis for $Im(T_A)$.

Example 3.4.9

$$A = \begin{bmatrix} 1 & 1 & 2 & 1 \\ 1 & 1 & 2 & 6 \\ 1 & 2 & 5 & 2 \end{bmatrix}$$

This implies: $T_A : \mathbb{R}^4 \to \mathbb{R}^3$.
Row-echelon form of A is:

$$A' = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, i_1 = 1, i_2 = 2, i_3 = 4$$

$$\begin{aligned} Im_T &= \operatorname{span}(T_A(e_1), T_A(e_2), T_A(e_4)) \\ &= \operatorname{span}(1 \cdot f_1 + 1 \cdot f_2 + 1 \cdot f_3, 1 \cdot f_1 + 1 \cdot f_2 + 2 \cdot f_3, 1 \cdot f_1 + 6 \cdot f_2 + 2 \cdot f_3) \end{aligned}$$

Definition 3.4.5: Elementary matrices + invertibility

$A \in \mathbb{F}^{m,n}$ is invertible if there is a $B \in \mathbb{F}^{n,m}$ such that:

$$A \cdot B = B \cdot A = I_n = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$$

Notation: $B = A^{-1}$.

Note:-

$A \in \mathbb{F}^{n,n}$ implies:

$$\begin{aligned} T_A : \mathbb{F}^n &\to \mathbb{F}^n \\ \vec{x} &\mapsto A \cdot \vec{x} \end{aligned}$$

(a) A is invertible $\iff$ T_A is an isomorphism.
In this case: $(T_A)^{-1} = T_{A^{-1}}$.

(b) Elementary matrices are invertible.

Theorem 3.4.4

Let $A \in \mathbb{F}^{n,n}$. The following are equivalent (TFAE):

1. The reduced row-echelon form of A is I_n.
2. $A = E_k \cdot \ldots \cdot E_1$ where $E_1, \ldots, E_k$ are elementary matrices.
3. A is invertible.

Proof of $1 \implies 2$**:** Let $I_n = A' = E_k \cdot \ldots \cdot E_1 \cdot A$.
Since elementary matrices are invertible: $(E_k \cdot \ldots \cdot E_1)^{-1} = E_1^{-1} \cdot \ldots \cdot E_k^{-1}$.
Then, $A = E_1^{-1} \cdot \ldots \cdot E_k^{-1}$.
But E_i^{-1} is elementary for $1 \leq i \leq k$ are also elementary matrices.
Thus, A is a product of elementary matrices. 😊

Proof of $2 \implies 3$ **:** If $A = E_1 \cdot \ldots \cdot E_k$, then $A^{-1} = E_k^{-1} \cdot \ldots \cdot E_1^{-1}$.
Since:

$$E_1 \cdot \ldots \cdot E_k = A$$
$$E_k^{-1} \cdot \ldots \cdot E_1^{-1} = B = A^{-1}$$

Thus:

$$A \cdot B = E_1 \cdot \ldots \cdot E_k \cdot E_k^{-1} \cdot \ldots \cdot E_1^{-1} = I_n$$

Thus, A is invertible. 😊

Proof of $3 \implies 1$ **:** Assume A is invertible.
Let $A' = E_k \cdot \ldots \cdot E_1 \cdot A$ be the row-echelon form of A.
Either $A' = I_n$ or the bottom row of A' is $(0, \ldots, 0)$.
If the bottom row of A' has all zeros, then:

$$T_{A'} : \mathbb{F}^n \longrightarrow \mathbb{F}^n \text{ is not surjective.}$$

Thus, $T_{A'}$ is not an isomorphism.
Meaning, A' is not invertible.
Therefore, A is not invertible. 😊

Consequence: If A is invertible, then row-reduce it to reduced row-echelon form.
Then, $I_n = E_k \cdot \ldots \cdot E_1 \cdot A$.
Where $E_k, \ldots, E_1 = A^{-1}$.

Note:-

Notice that we started to talk about determinants after section 3.C.
I've moved this to chapter 10 to correspond with the textbook.
Click here to go to the determinants section: Determinants

3.5 Products and quotients of Vector Spaces

Definition 3.5.1

Let $V_1, \ldots, V_m$ be vector spaces over $\mathbb{F}$.
The the product of $V_1, \ldots, V_m$ is:

$$V_1 \times \ldots \times V_m = \{(v_1, \ldots, v_m) \mid v_i \in V_i \text{ for } 1 \leq i \leq m\}$$

I.e., think of this in terms of a cartesian product.

Example 3.5.1

Elements of $\mathbb{R}^2 \times \mathbb{R}^3$ look like:

$$((3, 5), (1, 0, -7.2)) \in \mathbb{R}^2 \times \mathbb{R}^3$$

Example 3.5.2
Vectors in $P_2(\mathbb{R}) \times \mathbb{R}$ looks like:

$$(-3 + x - x^2, (2, 7))$$

Definition 3.5.2

Let's define vector addition + scalar multiplication on $V_1 \times \ldots \times V_m$.
They are defined component-wise:

$$(v_1, \ldots, v_m) + (w_1, \ldots, w_m) = (v_1 + w_1, \ldots, v_m + w_m)$$
$$\lambda \cdot (v_1, \ldots, v_m) = (\lambda \cdot v_1, \ldots, \lambda \cdot v_m)$$

Thus, the product of $V_1, \ldots, V_m$ is a vector space over $\mathbb{F}$.

Proposition 3.5.1
If $V_1, \ldots, V_m$ are finite dimensional over $\mathbb{F}$, then so is $V_1 \times \ldots \times V_m$.
In fact, the dimension of $V_1 \times \ldots \times V_m$ is:

$$\dim(V_1 \times \ldots \times V_m) = \dim(V_1) + \ldots + \dim(V_m)$$

Sketch of Proof: Say V_i has basis $\{v_{i,1}, \ldots, v_{i,m}\}$ for $1 \leq i \leq m$.
Then $V_1 \times \ldots \times V_m$ has basis:

$$\{(v_{1,1}, 0, \ldots, 0), \ldots, (v_{1,m}, 0, \ldots, 0), (0, v_{2,1}, 0, \ldots, 0), \ldots, (0, v_{m,m})\}$$

😆

Example 3.5.3
$P_2(\mathbb{R}) \times \mathbb{R}^2$:

(i) $P_2(\mathbb{R})$ has basis $\left\{1, x, x^2\right\}$.

(ii) $\mathbb{R}^2$ has basis $\{(1, 0), (0, 1)\}$.

Which means that $P_2(\mathbb{R}) \times \mathbb{R}^2$ has basis:

$$\left\{(1, (0, 0)), (x, (0, 0)), (x^2, (0, 0)), (0, (1, 0)), (0, (0, 1))\right\}$$

Connection between products and direct sums: Let $U_1, \ldots, U_m$ be subspaces of V over $\mathbb{F}$.
Let's define:

$$\Gamma \colon U_1 \times \ldots \times U_m \to U_1 + \ldots + U_m$$

So, $\Gamma(u_1, \ldots, u_m) = u_1 + \ldots + u_m$.
Is Γ a linear map?

Proof of linear map: Vector addition:

$$\begin{aligned}\Gamma((v_1, \ldots, v_m) + (u_1, \ldots, u_m)) &= \Gamma((v_1 + u_1, \ldots, v_m + u_m)) \\ &= (v_1 + u_1) + \ldots + (v_m + u_m) \\ &= (v_1 + \ldots + v_m) + (u_1 + \ldots + u_m) \\ &= \Gamma((v_1, \ldots, v_m)) + \Gamma((u_1, \ldots, u_m))\end{aligned}$$

Thus, it is additive.
Now, we check for homogeneity:

$$\begin{aligned}\Gamma(\lambda \cdot (v_1, \ldots, v_m)) &= \Gamma((\lambda \cdot v_1, \ldots, \lambda \cdot v_m)) \\ &= \lambda \cdot v_1 + \ldots + \lambda \cdot v_m \\ &= \lambda \cdot (v_1 + \ldots + v_m) \\ &= \lambda \cdot \Gamma((v_1, \ldots, v_m))\end{aligned}$$

Thus, it is homogeneous.
Therefore, Γ is a linear map as desired. 😄

Moreover:

(i) Γ is surjective:

if $u_1 + \ldots + u_m \in U_1 + \ldots + U_m$, then $\Gamma((u_1, \ldots, u_m)) = u_1 + \ldots + u_m$.

(ii) Γ is injective $\iff \ker(\Gamma) = \{0\}$

If $\Gamma((u_1, \ldots, u_m)) = 0$, then $u_1 + \ldots + u_m = 0$.

That means the only way to write 0 as a sum of vectors in $U_1, \ldots, U_m$ is if $u_1 = \ldots = u_m = 0$.

Or, if the sum is a direct sum: $U_1 \oplus \ldots \oplus U_m$.

Rank nullity: We know that:

$$\dim(U_1 \times \ldots \times U_m) = \dim Im(\Gamma) + \dim \ker(\Gamma)$$

Since we know that Γ is surjective, $\dim Im(\Gamma) = \dim(U_1 + \ldots + U_m)$.
Furthermore, we know that Γ is injective $\iff \ker(\Gamma) = \{0\}$.
Meaning that $\dim(U_1 \times \ldots \times U_m) = \dim(U_1 + \ldots + U_m)$
Which means that:

$$\dim(U_1) + \ldots + \dim(U_m) = \dim(U_1 \oplus \ldots \oplus U_m)$$

Thus, we have:
If $U_1, \ldots, U_m$ are finite subspaces of V over $\mathbb{F}$, then:

$$U_1 + \ldots + U_m = U_1 \oplus \ldots \oplus U_m \iff \dim(U_1 + \ldots + U_m) = \dim(U_1) + \ldots + \dim(U_m)$$

Definition 3.5.3

Let V be a vector space over $\mathbb{F}$.
With $U \subseteq V$ is a subspace.
With $v \in V$:

$$v + U = \{v + u \mid u \in U\} \text{ "afine subset parallel to } U \text{ "}$$

In other words, $v + U$ is an affine subset parallel to U.

Example 3.5.4
$V = \mathbb{R}^2, U = \{(x, 2x) \mid x \in \mathbb{R}\}$.
Then $v + U$ is the set of all lines parallel to U.
Let $v_1 = (3, 1)$ and $v_2 = (4, 3)$.

$$
\begin{aligned}
v + U &= \{(3,1) + (x, 2x) \mid x \in \mathbb{R}\} \\
&= \{(3 + x, 1 + 2x) \mid x \in \mathbb{R}\} \\
&= \{(4 + x, 3 + 2x) \mid x \in \mathbb{R}\}
\end{aligned}
$$

So even though $v_1 \neq v_2$ but $v_1 + U = v_2 + U$.

Lenma 3.5.1

(i) $v_1 + U = v_2 + U$

(ii) $v_2 - v_1 \in U$

(iii) $(v_1 + U) \cap (v_2 + U) \neq \emptyset$

Proof of $ii \implies i$ ***:*** Let $v \in v_1 + U$.
So $v = v_1 + u$ for some $u \in U$.

$$
\begin{aligned}
v = v_1 + u &= v_2 - v_2 - v_1 + u \\
&= v_2 + (v_1 - v_2) + u \in v_2 + U
\end{aligned}
$$

Similarly, $v_2 + U \subseteq v_1 + U$.

😄

Last time we proved: $(ii) \implies (i)$ and $(i) \implies (iii)$. Clear

Proof of $(iii) \implies (ii)$ ***:*** Take $w \in (v_1 + U) \cap (v_2 + U)$.
Then $w = v_1 + u_1$, $w = v_2 + u_2$ for some $u_1, u_2 \in U$.

$$
\begin{aligned}
\vec{0}_V &= (v_1 + u_1) - (v_2 + u_2) \\
&= (v_1 - v_2) + (u_1 - u_2) \\
\implies v_2 - v_1 &= u_1 - u_2 \in U
\end{aligned}
$$

Thus, we have shown that $v_2 - v_1 \in U$.

😄

Example 3.5.5 (Quotient space)
We have $V \setminus U := \{v + U : v \in V\}$.
Set of affine parallel subsets to U.
E.g. Let $V = \mathbb{R}^2, U = \{(x, 2x) : x \in \mathbb{R}\}$.
With $V \setminus U$ = the set of all lines parallel to U.
Then that means that it is the set of all lines with slope 2.

Note:-
An element of $\mathbb{R}^2 \setminus U$ is a whole line parallel to U.

This means that $V \setminus U$ is an $\mathbb{F}-$vector space!
Let's check addition:

$$
(v + U) +_{V \setminus U} (w + U) = (v +_V w) + U
$$

Scaler multiplication

$$\lambda \cdot (v + U) \coloneqq \lambda \cdot v + U$$

Have to check that, e.g, addition is well defined:
Say $v_1 + U = v_2 + U$ and $w_1 + U = w_2 + U$.
Then we need to show that:

$$\begin{aligned}(v_1 + U) + (w_1 + U) &\overbrace{=}^{?} (v_2 + U) + (w_2 + U)\\ (v_1 + w_1) + U &\overbrace{=}^{?} (v_2 + w_2) + U\\ &\iff (v_1 + w_1) - (v_2 + w_2) \in U\\ &\iff \underbrace{(v_1 - v_2)}_{\in U} + \underbrace{(w_1 - w_2)}_{\in U} \in U\end{aligned}$$

Which means that $v_1 - v_2 \in U$ and $w_1 - w_2 \in U$.
We also know that scalar multiplication is well defined:
Say $v_1 + U = v_2 + U$.

$$\begin{aligned}&\implies v_1 - v_2 \in U\\ &\implies \lambda \cdot (v_1 - v_2) \in U\\ &\implies \lambda \cdot v_1 - \lambda \cdot v_2 \in U\\ &\implies \lambda \cdot v_1 + U = \lambda \cdot v_2 + U\\ &\implies \lambda \cdot (v_1 + U) = \lambda \cdot (v_2 + U)\end{aligned}$$

Let's give an example:

Example 3.5.6

Let:

$$\begin{aligned}\pi : V \setminus U\\ v \mapsto v + U\end{aligned}$$

Check that π is linear.
Say V is finite dimensional, then so is U.
By rank-nullity, $\dim V = \dim \ker \pi + \dim Im(\pi)$.
π is surjective, which means that $\dim Im(\pi) = \dim V \setminus U$.
But $Im(\pi)$ is finite dimensional, which means that $V \setminus U$ is also finite dimensional.
Let's prove it.

$$\begin{aligned}\ker \pi &= \left\{v \in V : \pi(v) = \left\{\vec{0}_{V \setminus U}\right\}\right\}\\ &= \left\{v \in V : \pi(v) = \vec{0}_v + U\right\}\\ &= \{v \in V : v + U = U\}\\ &= \{v \in V : v \in U\}\\ &= V \setminus U\end{aligned}$$

The result follows, probably.

Theorem 3.5.1 1st isomorphism theorem

Let $T \in \mathscr{L}(V, W)$ with $Im(T) \subseteq W$, $\ker T \subseteq V$.
Which means that $V \longrightarrow V \setminus \ker T$.
Let's define the following:

$$\begin{aligned}\widetilde{T} : V \setminus \ker T &\longrightarrow Im(T) \\ v + \ker T &\mapsto T(v)\end{aligned}$$

Claims: We have the following claims:

(i) $\widetilde{T}$ is well defined.

If $v_1 + \ker T = v_2 + \ker T$, then we want to show that $\widetilde{T}(v_1 + \ker T) = \widetilde{T}(v_2 + \ker T)$.

We have:

$$\begin{aligned} v_1 - v_2 &\in \ker T \\ T(v_1 - v_2) &= \vec{0_W} \\ T(v_1) - T(v_2) \end{aligned}$$

Thus, it is well defined.

(ii) $\widetilde{T}$ is linear.

e.g., $\widetilde{T}((v + \ker T) + (w + \ker T)) = \widetilde{T}((v + w) + \ker T)$.

This is equal to $T(v + w) = T(v) + T(w)$

Meaning that $\widetilde{T}(v + \ker T) + \widetilde{T}(w + \ker T)$.

We leave homogeneity as an exercise.

(iii) $\widetilde{T}$ is injective!

Say $\widetilde{T}(v + \ker T) = \vec{0_W}$

This means that $T(v) = \vec{0_W}$.

Which implies that $v \in \ker T$.

Hence, $\vec{0} + \ker T = v + \ker T$.

This is $\vec{0_{V \setminus \ker T}}$

(iv) $\widetilde{T}$ is surjective!

Let $w \in Im(T)$.

Then $w = T(v)$ for some $v \in V$.

Which means that $w = \widetilde{T}(v + \ker T)$.

Thus, $\widetilde{T} \in \mathscr{L}(V \setminus \ker T, Im(T))$ is an isomorphism of $\mathbb{F}$ vector spaces.
i.e.,

$$V \setminus \ker T \cong Im(T)$$

Chapter 4

Polynomials

Definition 4.0.1

Let $z = a + bi$ where $a, b \in \mathbb{R}$. Then:

(i) The real part of z is a, denoted $\Re(z)$ or $\text{Re}(z)$.

(ii) The imaginary part of z is b, denoted $\Im(z)$ or $\text{Im}(z)$.

Hence, $z = \Re(z) + i\Im(z)$.

Definition 4.0.2

Let $z \in \mathbb{C}$, then
The complex conjugate of z is $\overline{z} = \Re(z) - \Im(z)i$.
The absolute value of z is $|z| = \sqrt{\Re(z)^2 + \Im(z)^2}$.

Properties of Complex numbers: Let $w, z \in \mathbb{C}$, where:

$$\begin{aligned} z &= a + bi \\ w &= c + di \\ \overline{z} &= a - bi \\ \overline{w} &= c - di \end{aligned}$$

(i) Sum of z and $\overline{z}$: $z + \overline{z} = 2\Re(z)$

Proof:

$$\begin{aligned} z + \overline{z} &= (a + bi) + (a - bi) \\ &= 2a \\ &= 2\Re(z) \end{aligned}$$

😄

(ii) Difference of z and $\overline{z}$: $z - \overline{z} = 2i\Im(z)$

Proof:

$$\begin{aligned} z - \overline{z} &= (a + bi) - (a - bi) \\ &= 2bi \\ &= 2\Im(z)i \end{aligned}$$

😄

(iii) Product of z and $\overline{z}$: $z\overline{z} = |z|^2$

Proof:

$$\begin{aligned} z\overline{z} &= (a+bi)(a-bi) \\ &= a^2 - abi + abi - b^2 i^2 \\ &= a^2 + b^2 \\ &= |z|^2 \end{aligned}$$

😄

(iv) Additivity of complex conjugate: $\overline{w+z} = \overline{w} + \overline{z}$

Proof:

$$\begin{aligned} \overline{z} + \overline{w} &= (a-bi) + (c-di) \\ &= (a+c) - (b+d)i \\ &= \overline{w+z} \end{aligned}$$

😄

(v) Multiplicativity of complex conjugate: $\overline{wz} = \overline{w} \cdot \overline{z}$

Proof:

$$\begin{aligned} \overline{w} \cdot \overline{z} &= (c-di)(a-bi) \\ &= ac - adi - bci - bdi^2 \\ &= ac - adi - bci + bd \\ &= (ac+bd) - (ad+bc)i \\ &= \overline{wz} \end{aligned}$$

😄

(vi) Conjugate of a conjugate: $\overline{\overline{z}} = z$

Proof:

$$\begin{aligned} \overline{\overline{z}} &= \overline{a-bi} \\ &= a+bi \\ &= z \end{aligned}$$

😄

(vii) Real and imaginary parts are bounded by $|z|$:

Proof:

$$\begin{aligned} |z|^2 &= z\overline{z} \\ &= (a+bi)(a-bi) \\ &= a^2 + b^2 \\ |z|^2 &\geqslant a^2 \\ |z|^2 &\geqslant b^2 \\ |z| &\geqslant a \\ |z| &\geqslant b \end{aligned}$$

😄

(viii) Absolute value of the complex conjugate: $|\overline{z}| = |z|$

Proof:

$$
\begin{aligned}
|\overline{z}| &= |a - bi| \\
&= \sqrt{a^2 + b^2} \\
&= |z|
\end{aligned}
$$

😄

(ix) Multiplicativity of absolute value: $|wz| = |w||z|$

Proof:

$$
\begin{aligned}
|wz|^2 &= (wz)(\overline{wz}) \\
|wz| &= \sqrt{(wz)(\overline{wz})} \\
&= \sqrt{(w\overline{w})(z\overline{z})} \\
&= \sqrt{w\overline{w}}\sqrt{z\overline{z}} \\
&= |w||z|
\end{aligned}
$$

😄

(x) Triangle Equality: $|w + z| \leqslant |w| + |z|$

Proof:

$$
\begin{aligned}
|w + z|^2 &= (w + z)(\overline{w} + \overline{z}) \\
&= w\overline{w} + w\overline{z} + z\overline{w} + z\overline{z} \\
&= |w|^2 + w\overline{z} + \overline{\overline{w}z} + |z|^2 \\
&= |w|^2 + |z|^2 + 2\Re(w\overline{z}) \\
&\leqslant |w|^2 + |z|^2 + 2\,|w\overline{z}| \\
&\leqslant |w|^2 + |z|^2 + 2\,|w|\,|\overline{z}| \\
&= (|w| + |z|)^2
\end{aligned}
$$

😄

Definition 4.0.3

Geometric interpretation of complex numbers:
Let $w, z \in \mathbb{C}$, $\theta, \phi \in \mathbb{R}$.
Let's write $z = |z|(\cos(\theta) + i\sin(\theta))$,
And $w = |w|(\cos(\phi) + i\sin(\phi))$.
Then:

$$zw = |z||w|(\cos(\theta + \phi) + i\sin(\theta + \phi))$$

Proof: Let's use trig identities:

$$\begin{aligned} zw &= (r(\cos(\theta) + i\sin(\theta)))(s(\cos(\phi) + i\sin(\phi))) \\ &= rs(\cos(\theta)\cos(\phi) - \sin(\theta)\sin(\phi) + i(\cos(\theta)\sin(\phi) + \sin(\theta)\cos(\phi))) \\ &= rs(\cos(\theta + \phi) + i\sin(\theta + \phi)) \end{aligned}$$

We used the following trig identities:

$$\begin{aligned} \cos(\alpha + \beta) &= \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \\ \sin(\alpha + \beta) &= \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta) \end{aligned}$$

😊

Theorem 4.0.1

Let $a_0, \ldots, a_m \in \mathbb{F}$. If:

$$a_0 + a_1 x + \ldots + a_m x^m = 0$$

For every $x \in \mathbb{F}$, then $a_0 = \ldots = a_m = 0$.

Proof: Assume the contrapositive. Let our polynomial be given by

$$p(x) = a_0 + a_1 x + \cdots + a_m x^m$$

If this polynomial is not the zero function, then there exists some coefficient $a_k \neq 0$.
Without loss of generality, let's assume that a_m is that coefficient.
We want to show that there exists some value $x = z$ for which the polynomial does not evaluate to zero.
Specifically, we'll show that the term $a_m z^m$ will dominate all other terms for a sufficiently large z, such that the polynomial cannot evaluate to zero.
To do this, let's choose z such that

$$z > \frac{\sum_{j=0}^{m-1} |a_j|}{|a_m|}$$

Given this choice of z, the magnitude of the term $a_m z^m$ will exceed the combined magnitudes of all the other terms:

$$|a_m z^m| > |a_0| + |a_1 z| + \cdots + |a_{m-1} z^{m-1}|$$

Now, when we evaluate $p(z)$:

$$p(z) = a_0 + a_1 z + \cdots + a_{m-1} z^{m-1} + a_m z^m$$

Given our choice of z, it's clear that $p(z) \neq 0$.
This completes the proof by contrapositive.
Thus, if a polynomial is the zero function, all of its coefficients must be zero.

😊

Question 2

Fix a real number c.

(a) Show that if p has degree $n > 0$, then there is some monomial q such that $p-(x-c)q$ is a polynomial of degree less than n. (A monomial is a polynomial that has only one non-zero term.)

(b) Suppose that p is a polynomial with a root at $x =$ i.e., $p(c) = 0$. Show that $(x - c)$ is a factor of p (that is, there is some polynomial r such that $p = (x - c)r$).

Proof of a: Given polynomial p with degree $n > 0$, and the form $p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$, for $a_i \in \mathbb{R}$.
Let's fix c, now, we want to show that there is some monomial q
such that $p - (x - c)q$ is a polynomial of degree less than n.
Let's proceed by induction on $n \in \mathbb{N}$,

Base Case: Let $n = 1$, which means that $p(x)$ has a degree of 1.

$$p(x) = a_0 + a_1x$$

Clearly, we can pick $q = a_1$ (as it is a monomial).
Moreover, if we solve for $p - (x - c)q$, we get

$$\begin{aligned} p - (x - c)q &= (a_0 + a_1x) - (x - c)(a_1) \\ &= a_0 + a_1x - a_1x + a_1c \\ &= a_0 + a_1c \end{aligned}$$

Notice, that $a_0 + a_1c$ is a constant polynomial, meaning that its degree is 0, which is less than 1.
Hence, the base case holds.

Inductive Step: Assume the statement holds for all polynomials p with degree less than n.
Thus, for all $k < n, k \in \mathbb{N}$, we have a monomial q s.t. $p - (x - c)q$ is a polynomial of degree less than k.
Now, we want to show that the statement holds for n.
Let's consider a polynomial p with degree n, then we can write:

$$p(x) = a_nx^n + p_{n-1}(x), \text{ where } p_{n-1}(x) \text{ is a polynomial of degree less than } n$$

By our inductive hypothesis, we know that there is some monomial $q_{n-1}(x)$
such that $p_{n-1} - (x - c)q_{n-1}$ is a polynomial of degree less than $n - 1$.
Combining this information, let's pick $q(x) = a_nx^{n-1}$. Clearly, $q(x)$ is a monomial.
Thus, we have:

$$\begin{aligned} p - (x - c)q &= (a_nx^n + p_{n-1}(x)) - (x - c)(a_nx^{n-1}) \\ &= a_nx^n + p_{n-1}(x) - a_nx^n + a_ncx^{n-1} \quad \text{leading term cancels out} \\ &= p_{n-1}(x) + a_ncx^{n-1} \end{aligned}$$

Notice that the degree of $p_{n-1}(x) + a_ncx^{n-1}$ is less than n.
This holds as the degree of $p_{n-1}(x)$ is less than $n - 1$ and a_ncx^{n-1} is a term of degree $n - 1$.
Which means the polynomial $p - (x - c)q$ is a polynomial of degree less than n.
Therefore, the inductive step holds.
Thus, by the principle of mathematical induction, we have shown that if p has degree $n > 0$,
then there is some monomial q such that $p - (x - c)q$ is a polynomial of degree less than n for all $n \in \mathbb{N}$. 😊

Proof of b: Assume that p is a polynomial with a root at $x = c$, i.e., $p(c) = 0$.
We want to show that $(x - c)$ is a factor of p, i.e., there is some polynomial r such that $p = (x - c)r$.
Let's proceed by induction on $n \in \mathbb{N}$ for the degree of p.
Note that if $n = 0$, then it is trivially true that $p = (x - c)r$.

Base Case: If $n = 1$, then $p(x) = a_0 + a_1 x$.
As $p(c) = 0$, this implies that $a_0 + a_1 c = 0$ and $a_0 = -a_1 c$.
Hence, $p(x) = a_1(x - c)$ and $(x - c)$ is a factor of p.
Notice that $r = a_1$, so the base case holds.

Inductive Step: Assume that the statement holds for some $k \in \mathbb{N}$, then
for any polynomial p of degree k with $p(c) = 0$, $(x - c)$ is a factor of p.
Now, let's consider a polynomial p of degree $k + 1$.
By part (a), there exists a monomial q such that

$$p - (x - c)q \text{ is a polynomial of degree less than } k + 1$$

Hence, we can write p as:

$$p(x) = (x - c)q(x) + s(x)$$

Where $s(x)$ is the difference of the two polynomials with degree less than $k + 1$.
Now substituting $x = c$ into $p(x)$, we get:

$$\begin{aligned} p(c) &= (c - c)q(c) + s(c) \\ &= 0 + s(c) \\ &= 0 \\ \implies s(c) &= 0 \end{aligned}$$

Thus, by our inductive hypothesis, we know that $(x - c)$ is a factor of $s(x)$.
Which means we can write:

$$s(x) = (x - c)t(x)$$

Substituting this into our original equation, we get:

$$\begin{aligned} p(x) &= (x - c)q(x) + s(x) \\ &= (x - c)q(x) + (x - c)t(x) \\ &= (x - c)(q(x) + t(x)) \end{aligned}$$

Thus, $(x - c)$ is a factor of p with some polynomial $r = q(x) + t(x)$.
Completing the inductive step.

Thus, through the principle of mathematical induction,
we have shown that if p is a polynomial with a root at $x = c$, i.e., $p(c) = 0$,
then $(x - c)$ is a factor of p, i.e., there is some polynomial r such that $p = (x - c)r$. 😊

Chapter 5

Eigenvalues, Eigenvectors, and Invariant Subspaces

5.1 Invariant subspaces (5.A + 5.B)

Note:-

Goal: understand the building blocks / internal structure of $T \in \mathscr{L}(V)$, especially when V is finite-dimensional.
Idea: Maybe $V = \bigoplus_{i=1}^{m} U_i$
Restrict attention to $T|_{U_i}: U_i \to V$.

Definition 5.1.1

Let $U \subseteq V$ is an invariant subspace under T if

$$u \in U \implies T(u) \in U$$

in other words, if $Im(T|_U) \subseteq U$,
or $T|_U: U \to U$, i.e., $T|_u \in \mathscr{L}(U)$ where $T: V \to V$.

Example 5.1.1

What does a 1 dimensional invariant subspace under T look like?
$U = \text{span}(v)$. Then $T(v) \in U$, so $T(v) = \lambda v$ for some $\lambda \in \mathbb{F}$.
Conversely, if $v \neq \vec{0}_v$ and $T(v) = \lambda v$ for some $\lambda \in \mathbb{F}$,
then $U = \text{span}(v)$ is 1-dimensional invariant subspace under T.
We call λ an eigenvalue of T.
If $v \neq \vec{0}_v$, then v is an eigenvector for the eigenvalue λ.

Proposition 5.1.1

Suppose that V is a finite-dimensional vector space over $\mathbb{F}$ and $T \in \mathscr{L}(V)$.
Then the following are equivalent:

(a) $\lambda \in \mathbb{F}$ is an eigenvalue of T.

(b) $T - \lambda Id$ is not injective.

(c) $T - \lambda Id$ is not surjective.

(d) $T - \lambda Id$ is not invertible.

Where $T - \lambda Id \in \mathscr{L}(V)$:

$$
\begin{aligned}
(T - \lambda Id)(v) &= T(v) - \lambda Id(v) \\
&= T(v) - \lambda v
\end{aligned}
$$

And given $T, S \in \mathscr{L}(V, W)$, $(T + S)(v) = T(v) + S(v)$ for all $v \in V$.
Thus, $T, \lambda Id \in \mathscr{L}(V, V)$, so $T - \lambda Id \in \mathscr{L}(V, V)$.

Proof of $1 \iff 2$ ***:*** I didn't get this :sob: 😃

Before we continue, let's prove a claim:

Claim 5.1.1

Eigenvectors corresponding to distinct eigenvalues are linearly independent.
Let $T \in \mathscr{L}(V)$, and let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of T,
with eigenvectors $v_1, \ldots, v_m$ respectively.
Then $v_1, \ldots, v_m$ are linearly independent.

Proof: Suppose for contradiction that $v_1, \ldots, v_m$ are linearly dependent.
Then by the linear dependence lemma, there exists a (smallest) $k \in \{1, \ldots, m\}$ such that

$$v_k \in \operatorname{span}(v_1, \ldots, v_{k-1}) \text{ (so } v_1, \ldots, v_{k-1} \text{ are linearly independent)}$$

This implies that $v_k = a_1 v_1 + \ldots + a_{k-1} v_{k-1}$ for some $a_1, \ldots, a_{k-1} \in \mathbb{F}$.
Now apply T:

$$
\begin{aligned}
T(v_k) &= T(a_1 v_1 + \ldots + a_{k-1} v_{k-1}) \\
&= a_1 T(v_1) + \ldots + a_{k-1} T(v_{k-1}) \\
&= a_1 \lambda_1 v_1 + \ldots + a_{k-1} \lambda_{k-1} v_{k-1} \\
\lambda_k \cdot v_k &= a_1 \lambda_1 v_1 + \ldots + a_{k-1} \lambda_{k-1} v_{k-1}
\end{aligned}
$$

Now take: $v_k \cdot (a_1 v_1 + \ldots + a_{k-1} v_{k-1}) - (a_1 \lambda_1 v_1 + \ldots + a_{k-1} \lambda_{k-1} v_{k-1})$:

$$\vec{0}_v = a_1(\lambda_k - \lambda_1) v_1 + \ldots + a_{k-1}(\lambda_k - \lambda_{k-1}) v_{k-1}$$

Since $v_1, \ldots, v_{k-1}$ are linearly independent, this implies that:

$$a_1(\lambda_k - \lambda_1) = \ldots = a_{k-1}(\lambda_k - \lambda_{k-1}) = 0$$

Since we are given that $\lambda_1, \ldots, \lambda_m$ are distinct, we have that $\lambda_k - \lambda_i \neq 0$ for all $i \in \{1, \ldots, k-1\}$.
This means that $a_1 = \ldots = a_{k-1} = 0_{\mathbb{F}}$.
Thus, $v_k = \vec{0}_v$ thus v_k is not an eigenvector.
Which is a contradiction!

😃

Thus, the claim holds.

Last Time: Let V be a finite-dimensional vector space over $\mathbb{F}$ and $T \in \mathscr{L}(V)$.
i.e., $T : v \longrightarrow V$ is linear.
WE know that $V \cong \mathbb{F}^n$ for some $n \in \mathbb{N}$.
Think $T : \mathbb{F}^n \rightarrow \mathbb{F}^n$.
Let $A = \mathcal{M}(T, (e_1, \ldots, e_n))$, where $e_1, \ldots, e_n$ is the standard basis for V.
We defined the characteristic polynomial of T to be:

$$\det(A - xI_n) \in P_n(\mathbb{F})$$

We showed $\lambda \in \mathbb{F}$ is an eigenvalue for $T \iff \det(A - \lambda I_n) = 0$.

The first part shows that there exists $v \neq \vec{0}_v, T(v) = \lambda \cdot v$

Theorem 5.1.1

Let $v \neq \left\{\vec{0}_v\right\}$ be a finite-dimensional vector space over $\mathbb{C}$.
Let $T \in \mathscr{L}(V)$.
Then T has at least one eigenvalue.

Proof: Let $n = \dim V$. Note $n \geqslant 1$, since $v \neq \left\{\vec{0}_v\right\}$.
Then $\det(A - xI_n)$ is a polynomial of degree n with complex coefficients.
By the fundamental theorem of algebra (proved in Math 427),
a non-constant polynomial with complex coefficients has a root in $\mathbb{C}$.
Thus, there exists $\lambda \in \mathbb{C}$ such that $\det(A - \lambda I_n) = 0$. 😊

Example 5.1.2

Let

$$A = \begin{pmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix}$$

Let:

$$\det(A - x \cdot I_3) = \det\left(\begin{pmatrix} 1 & 4 & 5 \\ 0 & 2 & 6 \\ 0 & 0 & 3 \end{pmatrix} - \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix}\right)$$

Thus, we get the following:

$$\det\left(\begin{pmatrix} 1-x & 4 & 5 \\ 0 & 2-x & 6 \\ 0 & 0 & 3-x \end{pmatrix}\right) = (1-x)(2-x)(3-x)$$

Roots of characteristic polynomial are $x = 1, 2,$ or3.

Change of basis: Does the characteristic polynomial depend on A, or does it depend only on T?

$$T : \mathbb{F}^n \to \mathbb{F}^n$$

We can have:

$$\begin{aligned} A &= \mathcal{M}(T, (e_1, \ldots, e_n)) \\ A' &= \mathcal{M}(T, (f_1, \ldots, f_n)) f_j && = a_{1,j}e_1 + \ldots + a_{n,j}e_n \end{aligned}$$

Which means our polynomial looks like:

$$P = \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \ldots & a_{n,n} \end{pmatrix}$$

Where we get a new basis in terms of old basis.
To get from A to A':

$$A' = \underbrace{P^{-1}}_{\text{converts e's to f's}} \cdot \underbrace{A}_{\text{apply T WRT e's}} \cdot \underbrace{P}_{\text{converts f's to e's}}$$

What does this mean for our characteristic polynomial?: We have that:

$$P^{-1}(A - xI_n)P = P^{-1}AP - \underbrace{P^{-1}xI_nP}_{x \cdot I_n} = P^{-1}AP - xI_n$$

Thus, we can write our characteristic polynomial with respect to f's as:

$$\begin{aligned}
\det(A' - xI_n) &= \det(P^{-1}AP - xI_n) \\
&= \det(P^{-1}AP - xP^{-1}I_nP) \\
&= \det(P^{-1}(A - xI_n)P) \\
&= \det(p^{-1}) \cdot \det(A - xI_n) \cdot \det(P) \\
&= \frac{1}{\det(P)} \cdot \det(A - xI_n) \cdot \det(P) \\
&= \det(A - xI_n) \text{ which is our characteristic polynomial with respect to E's}
\end{aligned}$$

Our next foal is to find basis of V such that $\mathcal{M}T$ has many zeros!
Makes computing determinants + eigenvalues easier.

Definition 5.1.2

$\mathcal{M}(T)$ is upper triangular if every entry below the diagonal is 0.
For instance:

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

Proposition 5.1.2

Let $T \in \mathscr{L}(V)$. V is a finite-dimensional vector space over $\mathbb{F}$.
Let $\vec{v_1}, \ldots, \vec{v_n}$ be a basis for V.
The following are equivalent:

1. $\mathcal{M}(T, (v_1, \ldots, v_n))$ is upper triangular.
2. $T(v_j) \in \operatorname{span}(v_1, \ldots, v_j)$ for all $j \in \{1, \ldots, n\}$.
3. For all $j \in \{1, \ldots, n\}$, $\operatorname{span}(v_1, \ldots, v_j)$ is invariant subspace for T.

This means that $T(\operatorname{span}(v_1, \ldots, v_j)) \subseteq \operatorname{span}(v_1, \ldots, v_j)$.

Proof of $1 \iff 2$ ***:*** Definition of $\mathcal{M}(T)$. 😊

Proof of $2 \implies 3$ ***:*** Definition of invariant subspace. 😊

Proof of $2 \implies 3$ ***:*** Fix $j \geq 1$.
Then:

$$\begin{aligned}
T(v_1) &\in \operatorname{span}(v_1) \subseteq \operatorname{span}(v_1, \ldots, v_j) \\
T(v_2) &\in \operatorname{span}(v_1, v_2) \subseteq \operatorname{span}(v_1, \ldots, v_j) \\
&\vdots \\
T(v_j) &\in \operatorname{span}(v_1, \ldots, v_j)
\end{aligned}$$

Since T is linear.
Then this implies that $T(a_1v_1 + \ldots + a_jv_j) = a_1T(v_1) + \ldots + a_jT(v_j) \in \operatorname{span}(v_1, \ldots, v_j)$.
Hence, $T(\operatorname{span}(v_1, \ldots, v_j)) \subseteq \operatorname{span}(v_1, \ldots, v_j)$. 😊

Theorem 5.1.2

Let V be a finite-dimensional vector space over $\mathbb{C}$ and $T \in \mathscr{L}(V)$.
Then there exists a basis $\vec{v}_1, \ldots, \vec{v}_n$ of V such that $\mathcal{M}(T, (v_1, \ldots, v_n))$ is upper triangular (UT).
For this, we need the above proposition.

Proof: Let's proceed on induction on $n = \dim V$.

Base Claim: Let $n = 1$, then clearly every 1×1 matrix is upper triangular.

Inductive Step: Assume that the statement holds for all $S \in \mathscr{L}(W)$ with $\dim W < \dim V$.
Let $\lambda \in \mathbb{C}$ be an eigenvalue for T. This means it exists such that $v \neq \vec{0}_v$.
Now consider $(T - \lambda \cdot Id : V \longrightarrow V)$.
Set $W = Im(T - \lambda \cdot Id) \subseteq V$.

Claim: W is an invariant subspace under T.

Proof of claim: Let $w \in W$. Then:

$$\begin{aligned} T(w) &= T(w) - \lambda \cdot w + \lambda \cdot w \\ &= \underbrace{(T - \lambda \cdot Id)(w)}_{\in W \text{ by definition}} + \underbrace{\lambda \cdot w}_{\in W} \\ &\in W \text{ since } W \text{ is closed under addition} \end{aligned}$$

😊

Since λ is an eigenvalue, we know that $T - \lambda \cdot Id$ is not surjective.
Which implies that $W \subsetneq V$, thus $\dim W < \dim V$.
With the claim, we can write:

$$T\mid_W : W \longrightarrow W$$

This implies that $T\mid_W \in \mathscr{L}(W)$.
By our inductive hypothesis, there exists a basis $\vec{w}_1, \ldots, \vec{w}_m$ of W such that $\mathcal{M}(T\mid_W, (w_1, \ldots, w_m))$ is upper triangular.
By our proposition, $T(w_j) \in \operatorname{span}(w_1, \ldots, w_j)$ for some j.
Extend to a basis for V: $\vec{w}_1, \ldots, \vec{w}_m, \vec{v}_1, \ldots, \vec{v_{n-m}}$.
Then, for $k = 1, \ldots, n - m$, we have:

$$\begin{aligned} T(v_k) &= T(v_k) - \lambda \cdot v_k + \lambda \cdot v_k \\ &= \underbrace{(T - \lambda \cdot Id)(v_k)}_{\in W = \operatorname{span}(\vec{w}_1, \ldots, \vec{w}_m)} + \lambda \cdot v_k \\ &\in \operatorname{span}(w_1, \ldots, w_m, v_k) \subseteq \operatorname{span}(w_1, \ldots, w_m, v_1, \ldots, v_k) \end{aligned}$$

By the proposition, $\mathcal{M}(T, (v_1, \ldots, v_n))$ is upper triangular.
Thus, through the principle of strong mathematical induction, the statement holds for all finite-dimensional vector spaces over $\mathbb{C}$. 😊

Claim 5.1.2 Upper triangular + invertibility

Let V be a finite-dimensional vector space over $\mathbb{F}$ and $T \in \mathscr{L}(V)$.
Now suppose there exists a basis for V such that $\mathcal{M}(T)$ is UT (e.g., $\mathbb{F} = \mathbb{C}$).
Then T is invertible $\iff$ all diagonal entries of $\mathcal{M}(T)$ are non-zero.

Proof: By hypothesis, there exists a basis $\vec{v}_1, \ldots, \vec{v}_n$ of V such that:

$$\mathcal{M}(T,(v_1,\ldots,v_n)) = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Now, we proofed with the biconditional.

⟸ : Suppose $\lambda_i \neq 0$ for all $i \in \{1,\ldots,n\}$.
Then $T(v_1) = \lambda_1 v_1$.
Which means that if $\lambda_1 \neq 0 \implies$:

$$v_1 = \frac{1}{\lambda_1}T(v_1) = T\left(\frac{1}{\lambda_1}v_1\right)$$

Which means that $v_1 \in Im(T)$.
$T(v_2) = a_{1,2}v_1 + \lambda_2 v_2$.
If $\lambda_2 \neq 0$, then:

$$v_2 = \frac{1}{\lambda_2}T(v_2) - \frac{a_{1,2}}{\lambda_2}v_1 = \underbrace{\underbrace{T\left(\frac{1}{\lambda_2}v_2\right)}_{\in Im(T)} - \frac{a_{1,2}}{\lambda_2}\underbrace{v_1}_{\in Im(T)}}_{\in Im(T) \text{ as it is a subspace}}$$

Now, induct on n, to show $v_n \in Im(T)$.
This means that $\operatorname{span}(v_1,\ldots,v_n) \subseteq Im(T)$.
Which means that $V \subseteq Im(T) \subseteq V$.
Thus, T is surjective.
Which means that T is invertible since we are working in a finite dimensional vector space.

Proof: Suppose the converse i.e., $\exists j \in \{1,\ldots,n\}$ such that $\lambda_j = 0$.
Then $T(v_j) = a_{1,j}v_1 + \ldots + a_{j-1,j}v_{j-1} + \lambda_j v_j$.
Notice that the last term is 0, so $T(v_j) \in \operatorname{span}(v_1,\ldots,v_{j-1})$.
Thus, $T(\operatorname{span}(v_1,\ldots,v_j)) \subseteq \operatorname{span}(v_1,\ldots,v_{j-1})$.
But the latter is dim j and the latter is dim $j-1$.
Which means that $T|_{\operatorname{span}(v_1,\ldots,v_j)}$ is not surjective.
As:

$$T|_{\operatorname{span}(v_1,\ldots,v_j)} : \operatorname{span}(v_1,\ldots,v_j) \longrightarrow \operatorname{span}(v_1,\ldots,v_{j-1})$$

Which means that T is not surjective, and thus not invertible.
Hence, the converse holds.

😄

😄

Theorem 5.1.3
If $\mathcal{M}(T)$ is UT then the eigenvalues of T are the diagonal entries of $\mathcal{M}(T)$.

Proof: Say:

$$\mathcal{M}(T,(v_1,\ldots,v_n)) = \begin{pmatrix} \lambda_1 & & \star \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Let $\lambda \in \mathbb{F}$, then $\mathcal{M}(T - \lambda \cdot Id)$ is UT.

$$\mathcal{M}(T - \lambda \cdot Id, (v_1, \ldots, v_n)) = \begin{pmatrix} \lambda_1 - \lambda & & \star \\ & \ddots & \\ 0 & & \lambda_n - \lambda \end{pmatrix}$$

Hence:

$$\begin{aligned} \lambda \in \mathbb{F} \text{ is an eigenvalue for } T &\iff T - \lambda \cdot Id \text{ not invertible} \\ &\iff \mathcal{M}(T - \lambda \cdot Id) \text{ not invertible} \\ &\iff \lambda_i - \lambda = 0 \text{ for some } i \end{aligned}$$

😄

5.2 Eigenspaces

Definition 5.2.1: Eigenspaces

Let V be a finite-dimensional vector space over $\mathbb{F}$ and $T \in \mathscr{L}(V)$.
With $\lambda \in \mathbb{F}$.
Then the eigenspace corresponding to λ is:

$$E(\lambda, T) := \ker(T - \lambda \cdot Id)$$

As:

$$\begin{aligned} (T - \lambda \cdot Id)(v) &= 0 \\ T(v) - (\lambda \cdot Id)(v) &= 0 \\ T(v) &= \lambda \cdot v \end{aligned}$$

i.e., this is the set of eigenvectors corresponding to λ together with $\vec{0}_v$.

Note:-

λ is an eigenvalue for $T \iff E(\lambda, T) \neq \left\{\vec{0}_v\right\}$.

Proposition 5.2.1

Let V be a finite-dimensional vector space over $\mathbb{F}$ and $T \in \mathscr{L}(V)$.
Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues.
Then $E(\lambda_1, T), \ldots, E(\lambda_m, T) \subseteq V$ is a direct sum.
Moreover:

$$\dim E(\lambda_1, T) + \ldots + \dim E(\lambda_m, T) \leqslant \dim V$$

Proof: Let $u_i \in E(\lambda_i, T)$ for $i = 1, \ldots, m$.
Suppose that $u_1 + \ldots + u_m = \vec{0}_v$.
Recall that eigenvectors for distinct eigenvalues are linearly independent.
Which implies that $u_i = \vec{0}_v$ for all i.
Thus, $E(\lambda_1, T), \ldots, E(\lambda_m, T)$ is a direct sum.
Thus:

$$\dim E(\lambda_1, T) + \ldots + \dim E(\lambda_m, T) = \dim \left(E(\lambda_1, T) + \ldots + E(\lambda_m, T)\right) \leqslant \dim V$$

😄

Definition 5.2.2

Diagonal matrix: $A \in \mathbb{F}^{n,n}$:

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Then, $T \in \mathscr{L}(V)$ is diagonalizable if there is a basis $v_1, \ldots, v_n$ of V such that:
$\mathcal{M}(T, (v_1, \ldots, v_n))$ is diagonal.

Example 5.2.1

$T : \mathbb{R}^3 \to \mathbb{R}^3$ with $\mathcal{M}(T, (e_1, e_2, e_3)$ is:

$$\begin{pmatrix} 5 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

Then $T(x, y, z) = (5x, 8y, 8z)$.
With $T(e_1) = 5e_1, T(e_2) = 8e_2, T(e_3) = 8e_3$.
Then $E(5, T) = \text{span}\,(e_1)$, $E(8, T) = \text{span}\,(e_2, e_3)$, and $E(0, T) = \mathbb{R}^3$.
Thus:

$$\dim E(5, T) + \dim E(8, T) \leq \dim \mathbb{R}^3$$

Example 5.2.2

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ with $(x, y) \mapsto (41x + 7y, -20x + 74y$
Use standard basis: $e_1 = (1, 0), e_2 = (0, 1)$

$$\begin{aligned} T(e_1) &= (41, -20) = 41e_1 - 20e_2 \\ T(e_2) &= (7, 74) = 7e_1 + 74e_2 \end{aligned}$$

Thus:

$$\mathcal{M}(T, (e_1, e_2)) = \begin{pmatrix} 41 & 7 \\ -20 & 74 \end{pmatrix}$$

Now try $v_1 = (1, 4), v_2 = (7, 5)$.
We claim that v_1, v_2 is a basis for $\mathbb{R}^2$.

$$\begin{aligned} T(v_1) &= (69, 276) = 69v_1 + 0v_2 \\ T(v_2) &= (322, 230) = 0v_1 + 46v_2 \end{aligned}$$

Thus, we get:

$$\mathcal{M}(T, (v_1, v_2)) = \begin{pmatrix} 69 & 0 \\ 0 & 46 \end{pmatrix}$$

So T is diagonalizable.

$$\begin{aligned} E(69, T) &\supseteq \text{span}\,(v_1) \\ E(46, T) &\supseteq \text{span}\,(v_2) \end{aligned}$$

Hence:

$$1 + 1 \leq \dim E(69, T) + \dim E(46, T) \leq \dim \mathbb{R}^2 = 2$$

Which implies that $E(69, T) = \text{span}\,(v_1)$ and $E(46, T) = \text{span}\,(v_2)$.

Theorem 5.2.1

Let V be a finite-dimensional vector space over $\mathbb{F}$ and $T \in \mathscr{L}(V)$.
Let $\lambda_1, \ldots, \lambda_m$ be a complete list of the distinct eigenvalues of T.
The following are equivalent:

1. T is diagonalizable.
2. V has a basis consisting of eigenvectors of T.
3. $V = E(\lambda_1, T) \oplus \ldots \oplus E(\lambda_m, T)$.
4. $\dim V = \dim E(\lambda_1, T) + \ldots + \dim E(\lambda_m, T)$.

Proof: We want to show:

$$\begin{aligned} 1 &\iff 2 \\ 2 &\implies 3 \\ 3 &\implies 4 \\ 4 &\implies 2 \end{aligned}$$

Let's start:

Proof of $1 \iff 2$ ***:*** This is trivial.
If $\mathcal{M}(T, (v_1, \ldots, v_n))$ is diagonal, then

$$T(v_i) = \mu_i v_i \text{ for some } \mu_i \in \mathbb{F}, i = 1, \ldots, n$$

Proof of $2 \implies 3$ ***:*** Say V has a basis consisting of eigenvectors of T.
This means that all $v \in V$ are linear combinations of eigenvectors.
Which means that $V \subseteq E(\lambda_1, T) + \ldots + E(\lambda_m, T) \subseteq V$.
Thus, $V = E(\lambda_1, T) + \ldots + E(\lambda_m, T)$.

Proof of $3 \implies 4$ ***:*** We showed this in $3.E$, where we showed that the sum of direct sums is less than or equal to the dimension of the vector space.

Proof of $4 \implies 2$ ***:*** Choose bases for each $E(\lambda_i, T)$ for $i = 1, \ldots, m$.
Concatenate to get a list $v_1, \ldots, v_n$ of V

Claim: $v_1, \ldots, v_n$ is a basis for V.

Proof of claim: We need to show span and linear independence.
Linearly independence:
Suppose that $a_1 v_1 + \ldots + a_n v_n = \vec{0}_v$.
In other words, $\sum_{k=1}^{n} a_k v_k = \vec{0}_v$.
Reorganize as $u_1 + \ldots + u_m = \vec{0}_v$ where $u_i \in E(\lambda_i, T)$.
By taking $u_i = \sum_{k \in K_i} a_k v_k$,
where $K_i = \{k \mid v_k \in E(\lambda_i, T)\}$.
Note: $u_i \in E(\lambda_i, T)$.
The u_i are either 0, or eigenvectors for distinct eigenvalues.
Such eigenvectors would be LI, as otherwise it would contradict $u_1 + \ldots + u_m = \vec{0}_v$.

Hence, $u_i = \vec{0_v}$ for $i = 1, \ldots, m$.
But $u_i = \sum_{k \in K_i} a_k v_k$.
Since these v_k's are LI (they are all in $E(\lambda_i, T)$).
Which implies that $a_k = 0$ for $k \in K_i$ and $i = 1, \ldots, m$.
Thus, $v_1, \ldots, v_n$ is linearly independent.
Now, our condition says that $\dim V = \dim E(\lambda_1, T) + \ldots + \dim E(\lambda_m, T)$.
Let's denote this as n, which is the dimension of V.
Hence, it's a linearly independent list with an appropriate dimension, which implies that it is a basis for V.

😊

Thus, the implication holds.
Hence, we have shown all the implications; thus the statement holds.

😊

Example 5.2.3

Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ with $T(w, z) = (z, 0)$.
Standard basis e_1, e_2, then $\mathcal{M}(T, (e_1, e_2)) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

$$\begin{aligned} T(e_1) &= (0, 0) = 0e_1 + 0e_2 \\ T(e_2) &= (1, 0) = e_1 + 0e_2 \end{aligned}$$

We know the eigenvalues are 0 and 0. What is $E(0, T)$?

$$\begin{aligned} E(0, T) &= \left\{ v \in \mathbb{C}^2 \mid T(v) = 0 \right\} \\ &= \left\{ (w, z) \in \mathbb{C}^2 \mid (z, 0) = (0, 0) \right\} \\ &= \operatorname{span}\left((1, 0)\right) \\ \implies & \dim E(0, T) = 1 \end{aligned}$$

Thus, you will never be able to find a basis of eigenvectors for T that makes $\mathcal{M}(T)$ diagonal.
Since $2 = \dim \mathbb{C}^2 \neq \dim E(0, T) = 1$, we conclude that T is not diagonalizable.

Chapter 6

Inner Product Spaces

6.1 Inner product spaces

Definition 6.1.1: Inner Products

Let V be a vector space over $\mathbb{F}$, with $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$.
The inner product:

$$\begin{aligned} \langle \, , \, \rangle : V \times V &\to \mathbb{F} \\ (v, w) &\mapsto \langle v, w \rangle \end{aligned}$$

Such that:

1. $\langle v, v \rangle \in \mathbb{R}$ and $\langle v, v \rangle \geq 0$ for all $v \in V$
2. $\langle v, v \rangle = 0 \iff v = \vec{0}_v$
3. $\langle u + w, v \rangle = \langle u, v \rangle + \langle w, v \rangle$ for all $u, w, v \in V$
4. $\langle \lambda \cdot_v v, w \rangle = \lambda \cdot_{\mathbb{F}} \langle v, w \rangle$ for all $\lambda \in \mathbb{F}$ and $v, w \in V$
5. $\langle u, v \rangle = \overline{\langle v, u \rangle}$ for all $u, v \in V$

Example 6.1.1

(i) Let $V = \mathbb{R}^n$ with $\mathbb{F} = \mathbb{R}$ and $\langle \, , \, \rangle =$ the dot product.
Then,

$$\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle \coloneqq x_1 y_1 + \ldots + x_n y_n$$

And:

$$\langle (x_1, \ldots, x_n), (x_1, \ldots, x_n) \rangle = x_1^2 + \ldots + x_n^2 = 0 \iff (x_1, \ldots, x_n) = (0, \ldots, 0)$$

(ii) Let $V = \mathbb{C}^n$ with $\mathbb{F} = \mathbb{C}$ and $\langle \, , \, \rangle =$ the dot product.
Then,

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle \coloneqq z_1 \overline{w_1} + \ldots + z_n \overline{w_n}$$

And:

$$\langle (z_1, \ldots, z_n), (z_1, \ldots, z_n) \rangle = z_1\overline{z_1} + \ldots + z_n\overline{z_n} = |z_1|^2 + \ldots + |z_n|^2 \geq 0$$

(iii) $V = P(\mathbb{R})$ with $\mathbb{F} = \mathbb{R}$ and $\langle\, , \,\rangle$ = the integral.

Then,

$$\langle p, q \rangle \coloneqq \int_0^\infty p(x)q(x) \cdot e_{-x} dx$$

(iv) $V = \{f : [-1, 1] \to \mathbb{R} \mid f \text{ is continuous}\}$ with $\mathbb{F} = \mathbb{R}$ and $\langle\, , \,\rangle$ = the integral.

Then,

$$\langle f, g \rangle \coloneqq \int_{-1}^1 f(x)g(x) dx$$

Definition 6.1.2: Inner product space

Vector space with an inner product is called an inner product space.
Consequences of axoims:

(i) Fix $u \in V$. Define:

$$\begin{aligned} T_u : V &\to \mathbb{F} \\ v &\mapsto \langle v, u \rangle \end{aligned}$$

Then T_u is a linear map.

Additivity:

$$T_u(v + w) = \langle v + w, u \rangle = \langle v, u \rangle + \langle w, u \rangle = T_u(v) + T_u(w)$$

Homogeneity:

$$T_u(\lambda \cdot_v v) = \langle \lambda \cdot_v v, u \rangle = \lambda \cdot_{\mathbb{F}} \langle v, u \rangle = \lambda \cdot_{\mathbb{F}} T_u(v)$$

(ii) $\left\langle \vec{0}_v, v \right\rangle = 0_{\mathbb{F}} : \left\langle \vec{0}_v, v \right\rangle = T_v(\vec{0}_v) = 0_{\mathbb{F}}$

(iii) $(\langle v, u + w \rangle = \langle v, u \rangle + \langle v, w \rangle)$:

Reasoning:

$$\begin{aligned} \langle v, u + w \rangle &= \overline{\langle u + w, v \rangle} \\ &= \overline{\langle u, v \rangle + \langle w, v \rangle} \\ &= \overline{\langle u, v \rangle} + \overline{\langle w, v \rangle} \\ &= \overline{\overline{\langle v, u \rangle}} + \overline{\overline{\langle v, w \rangle}} \\ &= \langle v, u \rangle + \langle v, w \rangle \end{aligned}$$

(iv) $\left\langle v, \vec{0}_v \right\rangle = 0_{\mathbb{F}} : \left\langle v, \vec{0}_v \right\rangle = \overline{\left\langle \vec{0}_v, v \right\rangle} = \overline{0_{\mathbb{F}}} = 0_{\mathbb{F}}$

(v) $\langle v, \lambda \cdot w \rangle = \overline{\lambda} \cdot \langle v, w \rangle$:

Reason:

$$\begin{aligned} \langle v, \lambda \cdot w \rangle &= \overline{\langle \lambda \cdot w, v \rangle} \\ &= \overline{\lambda \cdot \langle w, v \rangle} \\ &= \overline{\lambda} \cdot \overline{\langle w, v \rangle} \\ &= \overline{\lambda} \cdot \langle v, w \rangle \end{aligned}$$

Definition 6.1.3

Inner products is approx "size of a vector".
Norm:

$$|v| := \sqrt{\langle v, v \rangle}$$

Example 6.1.2

1. $V = \mathbb{R}^n, \mathbb{F} = \mathbb{R}, \langle\,,\,\rangle =$ dot product.

$$\|(x_1, \ldots, x_n)\| = \sqrt{x_1^2 + \ldots + x_n^2}$$

2. $V = \{f : [-1,1] \to \mathbb{R} \mid f \text{ is continuous}\}, \mathbb{F} = \mathbb{R}, \langle f(x), g(x)\rangle = \int_{-1}^{1} f(x)g(x)dx$

$$\|f\| = \sqrt{\int_{-1}^{1} f(x)^2 dx}$$

Meaning $\|f - g_n\| \to 0$ as $n \to \infty$.

For norms:

$$\begin{aligned}\|v\| = 0_{\mathbb{F}} &\iff \sqrt{\langle v, v\rangle} = 0_{\mathbb{F}} \\ &\iff \langle v, v\rangle = 0_{\mathbb{F}} \\ &\iff v = \vec{0}_v\end{aligned}$$

Definition 6.1.4: orthogonal

Two vectors u, v are orthogonal if $\langle u, v\rangle = 0_{\mathbb{F}}$.

Example 6.1.3

Let $V = \mathbb{R}^2$:

$$\begin{aligned}\langle (x_1, y_1), (x_2, y_2)\rangle = 0 &\iff x_1x_2 + y_1y_2 = 0 \\ &\iff \frac{x_2}{y_2} = -\frac{y_1}{x_1} \\ &\iff (x_1, y_1) \text{ and } (x_2, y_2) \text{ are perpendicular}\end{aligned}$$

Theorem 6.1.1 Pythagoras

Suppose $u, v \in V$ are orthogonal. Then:

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

Proof:

$$\begin{aligned}\|u + v\|^2 &= \langle u + v, u + v\rangle \\ &= \langle u, u + v\rangle + \langle v, u + v\rangle \\ &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &= \langle u, u\rangle + 0_{\mathbb{F}} + 0_{\mathbb{F}} + \langle v, v\rangle \\ &= \|u\|^2 + \|v\|^2\end{aligned}$$

😊

Loose End:

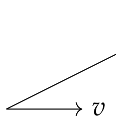
$$\|\lambda v\| = |\lambda| \cdot \|v\|$$

As both are non-negative.

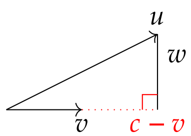
Proof:

$$\begin{aligned} \|\lambda v\| &= \sqrt{\langle \lambda v, \lambda v\rangle} \\ &= \sqrt{\lambda \cdot \overline{\lambda} \cdot \langle v, v\rangle} \\ &= \sqrt{|\lambda|^2 \cdot \langle v, v\rangle} \\ &= |\lambda| \cdot \sqrt{\langle v, v\rangle} \\ &= |\lambda| \cdot \|v\| \end{aligned}$$

Orthonormal Decompositions: Given:



Find $c \in \mathbb{F}, w \in V$ such that we complete the triangle, i.e. $u = c \cdot v + w$.



Want:

$$\begin{aligned} \langle w, v\rangle &\iff \langle u - c \cdot v, v\rangle = 0 \\ &\iff \langle u, v\rangle - c \cdot \langle v, v\rangle = 0 \\ &\iff \langle u, v\rangle - c \cdot \|v\|^2 = 0 \\ &\implies c = \frac{\langle u, v\rangle}{\|v\|^2} \end{aligned}$$

Note:-

$$w = u - cv = u - \frac{\langle u, v\rangle}{\|v\|^2} \cdot v$$

Where v and w are orthogonal.

Theorem 6.1.2 Cauchy-Schwarz inequality

Suppose that V is an inner product space and $u, v \in V$. Then:

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$$

With equality if and only if u and v are linearly dependent.

Proof: If $v = \vec{0}_v$, then both sides are $0_{\mathbb{F}}$.

Note:-

In this case equality holds and u, v are linearly dependent.

If $v \neq \vec{0}_v$, then v and $w := u - \frac{\langle u, v\rangle}{\|v\|^2} \cdot v$ are orthogonal.
Which means so are $\alpha \cdot v$ and w for any $\alpha \in \mathbb{F}$.
Recall that Pythagoras:

$$\|w + \alpha \cdot v\|^2 = \|w\|^2 + \|\alpha v\|^2 = \|w\|^2 + |\alpha|^2 \cdot \|v\|^2$$

Pick $\alpha = \frac{\langle u,v\rangle}{\|v\|^2}$, so that $w + \alpha v = u$
This implies that:

$$\|u\|^2 = \|w\|^2 + \left|\frac{\langle u,v\rangle}{\|v\|^2}\right|^2 \cdot \|v\|^2$$

Where the right term of the sum is:

$$\frac{|\langle u,v\rangle|^2}{\|v\|^4} \cdot \|v\|^2$$

Which is:

$$\|v\|^2 = \|w\|^2 + \frac{|\langle u,v\rangle|^2}{\|v\|^2} \geqslant \frac{|\langle u,v\rangle|^2}{\|v\|^2}$$

So we get:

$$\|v\|^2 \geqslant \frac{|\langle u,v\rangle|^2}{\|v\|^2} \implies \|u\| \cdot \|v\| \geqslant |\langle u,v\rangle|$$

Notice that equality holds if and only if:

$$\begin{aligned}
\|w\| = 0 &\iff w = \vec{0}_v \\
&\iff u = \frac{\langle u,v\rangle}{\|v\|^2} \cdot v = 0 \\
&\iff u = \frac{\langle u,v\rangle}{\|v\|^2} \cdot v \iff u, v \text{ are linearly dependent}
\end{aligned}$$

Example 6.1.4

1. Let $V = \mathbb{R}^n, \mathbb{F} = \mathbb{R}, \langle\, ,\, \rangle =$ dot product.

$$\vec{x} = (x_1, \ldots, x_n)$$
$$\vec{y} = (y_1, \ldots, y_n)$$

C.S. tell us:

$$\left|\langle \vec{x}, \vec{y}\rangle\right|^2 \leqslant \left\|\vec{x}\right\|^2 \cdot \left\|\vec{y}\right\|^2$$
$$(x_1y_1 + \ldots + x_ny_n)^2 \leqslant (x_1^2 + \ldots + x_n^2) \cdot (y_1^2 + \ldots + y_n^2)$$

2. Let $V = \{f : [-1,1] \to \mathbb{R} \mid f \text{ is continuous}\}, \mathbb{F} = \mathbb{R}, \langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$.
By C.S., we know that:

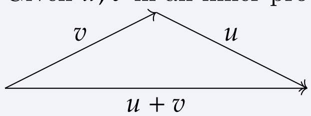
$$|\langle f, g\rangle|^2 \leqslant \|f\|^2 \cdot \|g\|^2$$

Thus:

$$\left(\int_{-1}^{1} f(x)g(x)dx\right)^2 \leqslant \left(\int_{-1}^{1} f(x)^2dx\right)\cdot\left(\int_{-1}^{1} g(x)^2dx\right)$$

Theorem 6.1.3 Triangle Inequality

Given u, v in an inner product space V, we have:



The triangle inequality states that:

$$\|u+v\| \leqslant \|u\| + \|v\|$$

Proof:

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \\
&= \|u\|^2 + \langle u,v\rangle + \overline{\langle u,v\rangle} + \|v\|^2 \\
&= \|u\|^2 + 2\cdot \operatorname{Re}\langle u,v\rangle + \|v\|^2 \text{ as } (a+bi)+(a-bi)=2a \\
&\leq \|u\|^2 + 2\cdot |\langle u,v\rangle| + \|v\|^2 \text{ by } \star \\
&\leq \|u\|^2 + 2\cdot \|u\|\cdot\|v\| + \|v\|^2 \text{ by C.S.} \\
&= (\|u\|+\|v\|)^2
\end{aligned}$$

Thus, we get:

$$\begin{aligned}
\|u+v\|^2 &\leqslant (\|u\|+\|v\|)^2 \\
\|u+v\| &\leqslant \|u\|+\|v\|
\end{aligned}$$

So why is $\star$ true?

$$\begin{aligned}
\langle u,v\rangle = a+bi &\implies |\langle u,v\rangle| = \sqrt{a^2+b^2} \\
\operatorname{Re}\langle u,v\rangle &= a
\end{aligned}$$

But why is $a \leqslant \sqrt{a^2+b^2}$?
True since:

$$\begin{aligned}
a &\leq |a| \\
&= \sqrt{a^2} \\
&\leq \sqrt{a^2+b^2}
\end{aligned}$$

Which means the triangle inequality holds. 😄

Example 6.1.5

Let $V = \mathbb{R}^n, \vec{x} = (x_1, \dots, x_n), \vec{y} = (y_1, \dots, y_n)$.
We have:

$$\left\|\vec{x} + \vec{y}\right\|^2 = \left\|\vec{x}\right\|^2 + 2 \cdot \left\langle \vec{x}, \vec{y} \right\rangle + \left\|\vec{y}\right\|^2$$

Thus:

$$\begin{aligned}\sum_{i=1}^{n}(x_i + y_i)^2 &\leqslant \left(\sqrt{\sum_{i=1}^{n} x_i^2} + \sqrt{\sum_{i=1}^{n} y_i^2}\right)^2 \\ &= \sum_{i=1}^{n} x_i^2 + 2 \cdot \sqrt{\sum_{i=1}^{n} x_i^2} \cdot \sqrt{\sum_{i=1}^{n} y_i^2} + \sum_{i=1}^{n} y_i^2\end{aligned}$$

6.2 Orthonormal bases

Definition 6.2.1

Let V be an inner product space over $\mathbb{F}$.
Let $e_1, \dots, e_n$ be a list of vectors in V.
Then we say $\{e_1, \dots, e_n\}$ is an orthogonal list if:

$$\left\langle e_i, e_j \right\rangle = \delta_{ij} := \begin{cases} 1_{\mathbb{F}} & i = j \\ 0_{\mathbb{F}} & i \neq j \end{cases}$$

E.g.

$$\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) \in \mathbb{R}^3$$

We normalize the vectors to get a length of 1.
If an orthogonal list is also a basis, then the following holds:

$$\begin{aligned}\|a_1e_1 + \dots + a_ne_n\|^2 &= \langle a_1e_1 + \dots + a_ne_n, a_1e_1 + \dots + a_ne_n \rangle \\ &= \langle a_1e_1, a_1e_1 \rangle + \dots + \langle a_ne_n, a_ne_n \rangle \\ &= a_1 \cdot \overline{a_1} \langle e_1, e_1 \rangle + \dots + a_n \cdot \overline{a_n} \langle e_n, e_n \rangle \\ &= |a_1|^2 + \dots + |a_n|^2\end{aligned}$$

A list of orgthogonal vectors is linearly independent, but might not span.

Definition 6.2.2

V is an inner product space over $\mathbb{R}$ or $\mathbb{C}$.
Then we say $\{V_1, \ldots, V_n\}$ is an orthonormal if :

$$\langle V_i, V_j \rangle = \begin{cases} 1_{\mathbb{F}} & i = j \\ 0_{\mathbb{F}} & i \neq j \end{cases}$$

Claim 6.2.1

Suppose $\{V_1, \ldots, V_n\}$ is orthonormal, then $\{V_1, \ldots, V_n\}$ is linearly independent.

Proof: Suppose there are some scalars $c_1, \ldots, c_n \in \mathbb{F}$ such that:

$$c_1 v_1 + \ldots c_n v_n = 0$$

Then it follows:

$$\langle c_1 v_1 + \ldots c_n v_n, c_1 v_1 + \ldots c_n v_n \rangle = \|c_1 v_1 + \ldots c_n v_n\|^2 = 0$$

Which means:

$$|c_1|^2 + \ldots + |c_n|^2 = 0 \implies |c_1|^2 = \ldots = |c_n|^2 = 0$$

Thus,

$$c_1 = \ldots = c_n = 0_{\mathbb{F}}$$

😄

: Suppose $\{e_1, \ldots, e_n\}$ is an orthonormal basis for V. and let $v \in V$. Then:

Algorithm 2: Gram-Schmidt Process

Input: $\vec{v_1}, \ldots, \vec{v_n} \in V$. Linearly independent set.
Output: $e_1, \ldots, e_n \in V$ orthonormal basis and $\text{span}(e_1, \ldots, e_n) = \text{span}(v_1, \ldots, v_n)$

/* We want $\langle e_1, e_2 \rangle = 0_{\mathbb{F}}$ */

$\vec{e_1} = \frac{\vec{v_1}}{\|\vec{v_1}\|}$;

$\vec{e_2} = \frac{\vec{v_2} - \langle \vec{v_2}, \vec{e_1} \rangle \cdot \vec{e_1}}{\|\vec{v_2} - \langle \vec{v_2}, \vec{e_1} \rangle \cdot \vec{e_1}\|}$;

$\vec{e_j} = \frac{\vec{v_j} - \langle \vec{v_j}, \vec{e_1} \rangle \cdot \vec{e_1} - \ldots - \langle \vec{v_j}, \vec{e_{j-1}} \rangle \cdot e_{j-1}}{\|\vec{v_j} - \langle \vec{v_j}, \vec{e_1} \rangle \cdot \vec{e_1} - \ldots - \langle \vec{v_j}, \vec{e_{j-1}} \rangle \cdot \vec{e_{j-1}}\|}$;

Example 6.2.1

Let $V = \mathcal{P}_2(\mathbb{R}), \langle p, q \rangle = \int_{-1}^{1} p(x)q(x)dx$.
Start with $\vec{v_1} = 1, \vec{v_2} = x, \vec{v_3} = x^2$.

1.

$$\vec{e_1} = \frac{\vec{v_1}}{\|\vec{v_1}\|} = \int_{-1}^{1} 1^2 dx = \sqrt{2}$$
$$= \frac{1}{\sqrt{2}}$$

2.

$$v_2 - \langle v_2, e_1 \rangle \cdot e_1 = x - \int_{-1}^{1} x \cdot \frac{1}{\sqrt{2}} dx \cdot \frac{1}{\sqrt{2}}$$

notice that the integral is 0 as x is odd

$$= x$$

Remember that:

$$\vec{e_2} = \frac{\vec{v_2} - \left\langle \vec{v_2}, \vec{e_1} \right\rangle \cdot \vec{e_1}}{\left\| \vec{v_2} - \left\langle \vec{v_2}, \vec{e_1} \right\rangle \cdot \vec{e_1} \right\|} = \frac{x}{\|x\|}$$

$$\|x\|^2 = \langle x, x \rangle = \int_{-1}^{1} x^2 dx = \frac{2}{3}$$

$$\implies \|x\| = \sqrt{\frac{2}{3}} \implies \vec{e_2} = \frac{x}{\sqrt{\frac{2}{3}}}$$

3.

$$v_3 - \langle v_3, e_1 \rangle \cdot e_1 - \langle v_3, e_2 \rangle \cdot e_2 = x^2 - \int_{-1}^{1} x^2 \cdot \frac{1}{\sqrt{2}} dx \cdot \frac{1}{\sqrt{2}} - \int_{-1}^{1} x^2 \cdot \frac{x}{\sqrt{\frac{2}{3}}} dx \cdot \frac{x}{\sqrt{\frac{2}{3}}}$$

notice that the integral is 0 as the right side is odd

$$= x_2 - \frac{1}{3}$$

Hence,

$$\left\| x^2 - \frac{1}{3} \right\| = \sqrt{\int_{-1}^{1} (x^2 - \frac{1}{3})^2 dx} = \sqrt{\frac{8}{45}} \implies \vec{e_3} = \frac{x^2 - \frac{1}{3}}{\sqrt{\frac{8}{45}}}$$

This week: (i) Inner product spaces

(ii) Some properties:

$$u = 0 \iff \langle v, u \rangle = 0 \text{ for all } v \in V$$

In particular:

$$u = u'$$
$$\iff u - u' = 0$$
$$\iff \forall v \in V, \langle v, u - u' \rangle = 0$$

Goal: Study linear operators between inner product spaces

Definition 6.2.3

A linear functional on V is a linear map $V \xrightarrow{\phi} \mathbb{F}$.
i.e., $\phi \in \mathscr{L}(V, \mathbb{F})$

Theorem 6.2.1 Riesz Representation Theorem (RRT)

Let V be a finite-dimensional vector space over $\mathbb{F}$ and $\phi \in \mathscr{L}(V, \mathbb{F})$.
Then there exists a unique $u \in V$ such that:

$$\phi(v) = \langle v, u\rangle \text{ for all } v \in V$$

Proof of part 1: Find u.

$$\phi(v) = \phi(\langle v, e_1\rangle e_1 + \ldots + \langle v, e_n\rangle e_n)$$

For some orthonormal basis $\{e_1, \ldots, e_n\}$ of V.
This means we get:

$$\begin{aligned}
&= \langle v, e_1\rangle \phi(e_1) + \ldots + \langle v, e_n\rangle \phi(e_n) \\
&= \left\langle v, \overline{\phi(e_1)}e_1\right\rangle + \ldots + \left\langle v, \overline{\phi(e_n)}e_n\right\rangle \\
&= \left\langle v, \overline{\phi(e_1)}e_1 + \ldots + \overline{\phi(e_n)}e_n\right\rangle
\end{aligned}$$

Which is u!
Thus,

$$\phi(v) = \langle v, u\rangle \text{ for all } v \in V$$

☺

Uniqueness:

$$\phi(v) = \langle v, u\rangle = \langle v, u'\rangle \text{ for all } v \in V$$

Show $u = u' \iff$ show $\langle v, u - u'\rangle = 0$ for all $v \in V$.

$$\begin{aligned}
\langle v, u - u'\rangle &= \langle v, u\rangle - \langle v, u'\rangle \\
&= \phi(v) - \phi(v) \\
&= 0
\end{aligned}$$

Thus, $u = u'$. ☺

Note:-
Because of uniqueness the u in the proof cannot / doesn't depend on the choice of basis.

Example 6.2.2

Let $\mathcal{P}_2(\mathbb{R})$ with $\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)dx$.
This has an orthonormal basis:

$$e_1 = \sqrt{\frac{1}{2}}, e_2 = \sqrt{\frac{3}{2}}x, e_3 = \sqrt{\frac{45}{8}}(x^2 - \frac{1}{3})$$

Let $\phi \in \mathscr{L}(\mathcal{P}_2(\mathbb{R}), \mathbb{R})$ be defined by:

$$\phi(p) = \int_{-1}^{1} p(x)\cos(\pi x)dx \in \mathscr{L}(\mathcal{P}_2(\mathbb{R}), \mathbb{R})$$

Note:-

We have:

$$\langle p, \cos(\pi x)\rangle = \phi(p)$$

but $\cos(\pi x) \notin \mathcal{P}_2(\mathbb{R})$.

Thus, by using RRT,

$$\phi(p) = \langle p, u\rangle$$

Where $u = \overline{\phi(e_1)}e_1 + \overline{\phi(e_2)}e_2 + \overline{\phi(e_3)}e_3$.
Notice that the second term is

$$\overline{\phi(e_2)}e_2 = \overline{\int_{-1}^{1} x\cos(\pi x)dx} \cdot \sqrt{\frac{3}{2}}x$$

Computing gives us:

$$u = -\frac{45}{2\pi^2}(x^2 - \frac{1}{3})$$

Now, let $(V, \langle\ ,\ \rangle_V), (W, \langle\ ,\ \rangle_W)$ be inner product spaces over $\mathbb{F}$.
Let $T \in \mathscr{L}(V, W)$.
For each $w \in W$, create: $\phi_w \in \mathscr{L}(V, \mathbb{F})$ by:

$$\phi_w(v) = \langle T(v), w\rangle_W$$

By RRT, for all $w \in W$, there exists a unique $u_w \in V$ such that:

$$\phi_w(v) = \langle v, u_w\rangle_V$$

Now, notice:

$$\langle v, u_w\rangle_V = \langle T(v), w\rangle_W$$

By uniqueness of u_w, we can define:

$$T^* : W \to V, w \mapsto u_w := T^*(w)$$

Definition 6.2.4

The adjoint of a linear map $T : V \to W$ between inner product spaces is the linear map $T^* : W \to V$ characterized by:

$$\langle T(v), w\rangle_W = \langle v, T^*(w)\rangle_V$$

Example 6.2.3

Let $\mathbb{R}^3, \mathbb{R}^2$ with the standard inner product i.e., dot product.

$$T : \mathbb{R}^3 \to \mathbb{R}^2, \quad T(x_1, x_2, x_3) = (x_1 + x_2, 2x_2 + x_3)$$

What is $T^* : \mathbb{R}^2 \to \mathbb{R}^3$?

$$
\begin{aligned}
\langle T(x_1, x_2, x_3), (y_1, y_2)\rangle_{\mathbb{R}^2} &= \langle (x_1 + x_2, 2x_2 + x_3), (y_1, y_2)\rangle_{\mathbb{R}^2} \\
&= \langle (x_1, x_2, x_3), T^*(y_1, y_2)\rangle_{\mathbb{R}^3} \\
&= (x_1 + x_2)y_1 + (2x_2 + x_3)y_2 \\
&= x_1y_1 + x_2y_1 + 2x_2y_2 + x_3y_2 \qquad = \langle (x_1, x_2, x_3), (y_1, y_1 + 2y_2, y_2)\rangle_{\mathbb{R}^3}
\end{aligned}
$$

Thus, $T^*(y_1, y_2) = (y_1, y_1 + 2y_2, y_2)$.

Note:-

Is T^* is linear?
Adjoins are linear:
If $T : V \longrightarrow W$ is linear, then $T^* : W \longrightarrow V$ is linear.

Chapter 7

Operators on Inner Product Spaces

Chapter 8

Operators on Complex Vector Spaces

Chapter 9

Operators on Real Vector Spaces

Chapter 10

Determinants and Traces

10.1 Determinants and Permutations

Definition 10.1.1: Determinants

$\det : \mathbb{F}^{n,n} \implies \mathbb{F}$

(a) If $n = 1$, then $\det(a) = a$.

(b) If $n = 2$, then $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$.

(c) If $n \geq 3$, then we need a recursive definition.
If $A \in \mathbb{F}^{n,n}$, then the ij-th minor of A is $A_{i,j}$.
Where $A_{i,j}$ means you take A and delete the ith row and jth column.

Example 10.1.1
Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

.
Then:

$$A_{2,1} = \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix}$$

Thus, given $A \in \mathbb{F}^{n,n}$, define its determinant as:

$$\det A = \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \cdot \det A_{i,1}$$

Example 10.1.2

Let:

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 2 \\ 0 & 5 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Thus,

$$\begin{aligned} \det A &= a_{1,1} \cdot \det A_{1,1} - a_{2,1} \cdot \det A_{2,1} + a_{3,1} \cdot \det A_{3,1} \\ &= 1 \cdot \det \begin{pmatrix} 1 & 2 \\ 5 & 1 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 0 & 3 \\ 5 & 1 \end{pmatrix} + 0 \cdot \det \begin{pmatrix} 0 & 3 \\ 1 & 2 \end{pmatrix} \\ &= 1 \cdot (-9) - 2 \cdot (-15) + 0 \cdot (-3) \\ &= 21 \end{aligned}$$

Theorem 10.1.1 Det 1

There exists a unique function $\delta : \mathbb{F}^{n,n} \longrightarrow \mathbb{F}$ with the following properties:

1. $\delta(I_n) = 1$
2. δ is row-linear.
3. If A has two identical rows, then $\delta(A) = 0$.

Point: we will show that $\det = \delta$.

Row-linear: This means that:

$$\delta\left(\begin{pmatrix} 1 & 2 & 3 \\ 4\lambda + 2\mu & 5\lambda + 5\mu & 6\lambda + 8\mu \\ 7 & 8 & 9 \end{pmatrix}\right) = \lambda \cdot \delta \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \mu \cdot \delta \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 7 & 8 & 9 \end{pmatrix}$$

Assume the previous theorem is true for now.
What is the value of δ on elementary matrices?

Theorem 10.1.2 Det 2

E elementary matrix. Then:

$$\delta(E \cdot A) = \begin{cases} \delta(A) & \text{if } E \text{ is type (i)} \\ -\delta(A) & \text{if } E \text{ is type (ii)} \\ c \cdot \delta(A) & \text{if } E \text{ is type (iii)} \end{cases}$$

S is determined on elementary matrices.

Corollary 10.1.1 Related to thm 2

$$\delta(E) = \begin{cases} +1 & \text{if } E \text{ is type (i)} \\ -1 & \text{if } E \text{ is type (ii)} \\ c & \text{if } E \text{ is type (iii)} \end{cases}$$

Proof: Take $A = I_n$ in theorem 2. 😁

Proof to det 2: For E of type (iii) this is jut row-linearity of δ.

Let A_i be the ith row of A.

$$\delta\left(\begin{pmatrix}1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1\end{pmatrix}\cdot\begin{pmatrix}-- & A_1 & -- \\ -- & A_2 & -- \\ & \vdots & \\ -- & A_n & --\end{pmatrix}\right) = \delta\begin{pmatrix}-- & A_1 & -- \\ & \vdots & \\ -- & cA_i & -- \\ & \vdots & \\ -- & A_n & --\end{pmatrix} = c\cdot\delta\begin{pmatrix}-- & A_1 & -- \\ & \vdots & \\ -- & A_i & -- \\ & \vdots & \\ -- & A_n & --\end{pmatrix} = c\cdot\delta(A)\begin{pmatrix}1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1\end{pmatrix}$$

Since we did not require $c \neq 0$, then this is true for all $c \in \mathbb{F}$.

Thus, $\delta(E \cdot A) = c \cdot \delta(A)$ for all $c \in \mathbb{F}$.

If a row contains only zeros, then $\delta(A) = 0$.

For types (i) and (ii), we do the special case when E acts on consecutive rows.

(Type: i):

$$E \cdot A = \begin{pmatrix}1 & & & & \\ & \ddots & & a_{i,j} & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1\end{pmatrix}\cdot\begin{pmatrix}-- & A_1 & -- \\ -- & A_2 & -- \\ & \vdots & \\ -- & A_i & -- \\ -- & A_{i+1} & -- \\ & \vdots & \\ -- & A_n & --\end{pmatrix} = \begin{pmatrix}-- & A_1 & -- \\ -- & A_2 & -- \\ & \vdots & \\ -- & a_{i,j}A_i + A_j & -- \\ -- & A_{i+1} & -- \\ & \vdots & \\ -- & A_n & --\end{pmatrix}$$

Special case, $j = i + 1$

$$\delta(E \cdot A) = \delta\begin{pmatrix}-- & A_1 & -- \\ -- & A_2 & -- \\ & \vdots & \\ -- & A_{i+1} & -- \\ -- & aA_i + A_{i+1} & -- \\ & \vdots & \\ -- & A_n & --\end{pmatrix} = a \cdot \delta\begin{pmatrix}-- & A_1 & -- \\ -- & A_2 & -- \\ & \vdots & \\ -- & A_i & -- \\ -- & A_i & -- \\ & \vdots & \\ -- & A_n & --\end{pmatrix} + \delta\begin{pmatrix}-- & A_1 & -- \\ -- & A_2 & -- \\ & \vdots & \\ -- & A_i & -- \\ -- & A_{i+1} & -- \\ & \vdots & \\ -- & A_n & --\end{pmatrix}$$

But, the first matrix's determinant is 0 since it has two identical rows.

Thus, $\delta(E \cdot A) = a \cdot 0 + \delta(A) = \delta(A)$.

(Type: ii): Swap rows.

Again, this assumes theorem 1.

Let's assume that we swap row i with row $i + 1$

$$\delta(A) = \delta\begin{pmatrix}-- & A_1 & -- \\ & \vdots & \\ -- & A_i & -- \\ -- & A_{i+1} & -- \\ & \vdots & \\ -- & A_n & --\end{pmatrix}$$

by part one:

$$= \delta \begin{pmatrix} -- & A_1 & -- \\ & \vdots & \\ -- & A_i - A_{i+1} & -- \\ -- & A_{i+1} & -- \\ & \vdots & \\ -- & A_n & -- \end{pmatrix}$$

by part one again:

$$= \delta \begin{pmatrix} -- & A_1 & -- \\ & \vdots & \\ -- & A_i - A_{i+1} & -- \\ -- & A_{i+1} + (A_i - A_{i+1}) & -- \\ & \vdots & \\ -- & A_n & -- \end{pmatrix}$$

$$= \delta \begin{pmatrix} -- & A_1 & -- \\ & \vdots & \\ -- & A_i - A_{i+1} & -- \\ -- & A_i & -- \\ & \vdots & \\ -- & A_n & -- \end{pmatrix}$$

by row linearity:

$$= \delta \begin{pmatrix} -- & A_1 & -- \\ & \vdots & \\ -- & A_i & -- \\ -- & A_i & -- \\ & \vdots & \\ -- & A_n & -- \end{pmatrix} - \delta \begin{pmatrix} -- & A_1 & -- \\ & \vdots & \\ -- & A_{i+1} & -- \\ -- & A_i & -- \\ & \vdots & \\ -- & A_n & -- \end{pmatrix}$$

but for the first matrix is zero since it has two identical rows:

$$= -\delta \begin{pmatrix} -- & A_1 & -- \\ & \vdots & \\ -- & A_{i+1} & -- \\ -- & A_i & -- \\ & \vdots & \\ -- & A_n & -- \end{pmatrix}$$

$$= -\delta(E \cdot A)$$

Thus, $\delta(A) = -\delta(E \cdot A)$, which implies

$$\delta(E \cdot A) = -\delta(A)$$

In general, for part 2, we want to swap i with row j.
Assume $i < j$.

(i) We bubble down row i to row j indices $j - 1$ exchanges.

Thus, row i is in the right place.

But Row j is row in Row $j - 1$.

(ii) Bubble up row j (in position $j - 1$ right now) to row i.

This involves $(j - 1) - i$ exchanges.

Which means that the total number of exchanges is:

$$j - 1 + (j-1) - i = 2(j-i) - 1$$

Which is odd!

This means that $\delta(E \cdot A) = (-1)^{2(j-i)-1} \cdot \delta(A) = -\delta(A)$.

Note:-

We can also do this for part 1.
Do this an exercise.

As such, we have proven theorem 2. 😄

Corollary 10.1.2

$\delta(S \cdot B) = \delta(A) \cdot \delta(B)$ for any $A, B \in \mathbb{F}^{n,n}$.
We know that $\delta(E) \cdot \delta(A) = \delta(E \cdot A)$ if E is an elementary matrix.
Let $A' = E_k \cdots E_1 \cdot A$ be a reduced row echelon form of A.
Then either:

(i) $A' = I_n$ or

(ii) the last row of A' is all zeros. (could be more than the last row)

Then:

(i) If $A' = I_n$,

$$\begin{aligned} A' = I_n &\implies A = E_1^{-1} \cdots E_k^{-1} \cdot I_n \\ &\implies \delta(A) = \delta(E_1^{-1}) \cdots \delta(E_k^{-1}) \end{aligned}$$

On the other hand:

$$\begin{aligned} \delta(A \cdot B) &= \delta(E_1^{-1}) \cdots \lambda(E_k^{-1}) \cdot \delta(B) \\ &= \delta(A) \cdot \delta(B) \end{aligned}$$

(ii) If A' has a row of zeros, then:

$\delta(A') = 0$, which implies that $\delta(A) = 0$.

Since $\delta(A') = \delta(E_k) \cdots \delta(E_1) \cdot \delta(A)$.

Where $\delta(E_i) \neq 0$ for all i.

This implies that $\delta(A) = 0$.

And exercise: $\delta(A \cdot B) = 0$ as well.

Proof of det 1: ***Proof of uniqueness:*** Suppose there are functions $\delta : \mathbb{F}^{n,n} \to \mathbb{F}$ and $\delta' : \mathbb{F}^{n,n} \to \mathbb{F}$.
Each satisfying the three desired properties.
Let $A \in \mathbb{F}^{n,n}$ such that $A' = E_k \cdots E_1 \cdot A$ is a reduced row echelon form of A.
Either we get $A' = I_n$ or its last row is all zeros.
In either case, $\delta(A') = \delta'(A') = 1$.
Or $\delta(A') = \delta'(A') = 0$.
That means that $\delta(E_k \cdots E_1 \cdot A) = \delta'(E_k \cdots E_1 \cdot A)$ in either case.
Thus,

$$\delta(E_k) \cdots \delta(E_1) \cdot \delta(A) = \delta'(E_k) \cdots \delta'(E_1) \cdot \delta'(A)$$

But by theorem 2, we get $\delta(E_i) = \delta'(E_i)$.
Which means that $\delta(A) = \delta'(A)$ for all $A \in \mathbb{F}^{n,n}$. 😄

Proof of existence: We'll show $\det : \mathbb{F}^{n,n} \to \mathbb{F}$ satisfies the three properties to be δ.
Let's proceed by induction on $n \in \mathbb{N}$

Base Case: Let n be 1.
Then $\det : \mathbb{F}^{1,1} \to \mathbb{F}$ is defined by $\det(a) = a$.
Thus, $\det(I_1) = 1$.
Now, for row linear:

$$\det(\lambda a + \mu b) = \lambda \cdot \det(a) + \mu \cdot \det(b)$$

Part 3 is trivial since there is only one row.

Inductive Step: Assume that $\det : \mathbb{F}^{n-1,n-1} \to \mathbb{F}$ satisfies the three properties.
We show the $n \times n$ case!
We need to show the three properties.

(i) I_n.

$$\begin{aligned}\delta(I_n) &= \delta\begin{pmatrix}1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1\end{pmatrix}\\ &= 1 \cdot \delta(I_{n-1})\\ &= 1 \cdot 1 \quad \text{by inductive hypothesis}\\ &= 1\end{aligned}$$

(ii) Let $A, B, D \in \mathbb{F}^{n,n}$ be identical matrices except for row k.
Where $D_k = \lambda A_k + \mu B_k$.
We want to show that $\delta(D) = \lambda \cdot \delta(A) + \mu \cdot \delta(B)$.

Claim 10.1.1

$d_{i,1} \cdot \det(D_{i,1}) = \lambda \cdot a_{i,1} \cdot \det(A_{i,1}) + \mu \cdot b_{i,1} \cdot \det(B_{i,1})$ is true for all $i \in \{1, \ldots, n\}$.
If claim is true then we can:

(a) Multiply equation by $(-1)^{i+1}$

(b) Add from $i = 1$ to n to get $\delta(D) = \lambda \cdot \delta(A) + \mu \cdot \delta(B)$.

Proof of claim: We have two cases:

Case (i) $i = k$, then The minors $A_{k,1}, B_{k,1}$ and $D_{k,1}$ are equal.
I.e., the kth row of A, B, D is deleted.
Which means:
Claim is true $\iff d_{i,1} = \lambda \cdot a_{i,1} + \mu \cdot b_{i,1}$.
This is true by our construction of D.

Case (ii) $i \neq k$, then
A', B', D' rows with $n-1$ entries after deleting the kth row.
Then $D'_k = \lambda \cdot A'_k + \mu \cdot B'_k$.

All other rows of A', B', D' are equal.
Thus, by inductive hypothesis:

$$\det D_{i,1} = \lambda \cdot \det A_{i,1} + \mu \cdot \det B_{i,1}$$

But also if $i \neq k$, $a_{i,1} = b_{i,1} = d_{i,1}$.
Thus,

$$d_{i,1} \cdot \det D_{i,1} = \lambda \cdot a_{i,1} \cdot \det A_{i,1} + \mu \cdot b_{i,1} \cdot \det B_{i,1}$$

Thus, the claim is true in this case as well.

😃

(iii) Moved a bit ahead in these notes, you can see the final part of the proof.

😃

😃

Note:-

On Mondays' class we showed that:

(a) δ is unique, if it exists

(b) $\det : \mathbb{F}^{m,n} \to \mathbb{F}$ such that: $A \mapsto \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \cdot \det A_{i,1}$ is row-linear and $\det I_n = 1$.
We showed this by induction on n.

Proof of Det 1.3: Let's proceed by induction on n.
Suppose rows k and $k+1$ of A are equal.
Then if $i \neq k$ or $k+1$, then
$(n-1) \times (n-1)$ minor $A_{i,1}$ has two consecutive equal rows.
By inductive hypothesis, $\det A_{i,1} = 0$. Then:

$$\det(A) = (-1)^{k+1} \cdot a_{k,1} \cdot \det A_{k,1} + (-1)^{k+2} \cdot a_{k+1,1} \cdot \det A_{k+1,1}$$

Since $A_k = A_{k+1}$ have $a_{k,1} = a_{k+1,1}$ and $A_{k,1} = A_{k+1,1}$
This implies that:

$$\det A = (-1)^{k+1} \left[a_{k,1} \cdot \det A_{k,1} + (-1) \cdot a_{k,1} \cdot \det A_{k,1} \right] = 0$$

Thus, $\det A = 0$.
Therefore, by the principle of mathematical induction, $\det A = 0$ for all A with two identical rows. 😃

Corollary 10.1.3

These are given free by the theorem of det 1:

(a) $\det(A \cdot B) = \det(A) \cdot \det(B)$

(b) $\det(A) = 0$ if A has a row of zeros.

(c) $\det(A) = 0$ if $A_j = \lambda \cdot A_i$ for some $i \neq j$ and $\lambda \in \mathbb{F}$.

Other formulas: (a) General column expansion:

This lands among the jth column:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{i,j} \cdot \det(A_{i,j})$$

(b) General row expansion:

This expands along the ith row:

$$\det(A) = \sum_{j=1}^{n}(-1)^{i+j} \cdot a_{i,j} \cdot \det(A_{i,j})$$

Definition 10.1.2: Permutations

A permutation of S is a bijection $\sigma : S \xrightarrow{\sim} S$.
e.g. $S = \{1, 2, 3, 4, 5\}$.

S	1	2	3	4	5
$\sigma(S)$	3	5	1	4	2

Then:

$$S_n \coloneqq \left\{\text{permutations } \sigma : \{1, \ldots, n\} \xrightarrow{\sim} \{1, \ldots, n\}\right\}$$

Notice that this is the symmetric group on n elements.
We can see that size is:

$$|S_n| = n!$$

Can compare permutations:

$$\tau : \{1, \ldots, n\} \xrightarrow{\sim} \{1, \ldots, n\}$$
$$\sigma : \{1, \ldots, n\} \xrightarrow{\sim} \{1, \ldots, n\}$$

Then $\tau \circ \sigma$ is also a bijection (”group law”).

Cycle Notation: Take the explicit σ above.
Given: $1 \mapsto 3 \mapsto 4 \mapsto 1$
draw a 3-cycle
And $2 \mapsto 5 \mapsto 2$
draw a 2-cycle
We can write:

$$\begin{aligned}\sigma &= (1, 3, 4)(2, 5) \text{ this is cycle notation.} \\ &= (52)(341) \text{ cycle notation is not unique}\end{aligned}$$

Example:

S	1	2	3	4
$\sigma(S)$	4	1	3	2

Thus:

$$\begin{aligned}\sigma &= (142)(3) \\ &= (142)\end{aligned}$$

Where (3) is a fixed point.
Thus, we can notice the composition in cycle notation:

$$
\begin{aligned}
\sigma &= (134)(25) \\
\tau &= (1452) \\
\tau \circ \sigma &= \underbrace{[(1452)]}_{\text{then this}} \circ \underbrace{[(134)(25)]}_{\text{first do this}} \\
&= (135)(2)(4) \\
&= (135) \\
(\tau \circ \sigma)(1) &= \tau(\sigma(1)) = \tau(3) = 5
\end{aligned}
$$

Question 3

Problem 5. The trace of a square matrix A is the sum of its diagonal entries:

$$\operatorname{tr}(A) := a_{11} + a_{22} + \cdots + a_{nn}$$

Show that
(a) $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$;
(b) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$;
(c) if B is invertible, then $\operatorname{tr}(A) = \operatorname{tr}\left(BAB^{-1}\right)$.

***Proof of** a* **:** Let A, B be two matrices size $n \times n$ with entries in $\mathbb{F}$.
Let $C = A + B$, which means that it looks like:

$$C = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \cdots & a_{nn} + b_{nn} \end{pmatrix}$$

Let's take the trace of C:

$$
\begin{aligned}
tr(C) &= \sum_{i=1}^{n} c_{ii} \\
&= (a_{1,1} + b_{11}) + \ldots + (a_{n,n} + b_{n,n}) \\
&= a_{11} + \ldots + a_{n,n} + b_{11} + \ldots + b_{n,n}
\end{aligned}
$$

Now let's take the trace of A and B separately:

$$
\begin{aligned}
tr(A) + tr(b) &= \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} \\
&= a_{11} + \ldots + a_{n,n} + b_{11} + \ldots + b_{n,n}
\end{aligned}
$$

As both sides are equal, we have shown that $tr(A + B) = tr(A) + tr(B)$. 😃

***Proof of** b* **:** Let A, B be two matrices size $n \times n$ with entries in $\mathbb{F}$.
If they are not of the same size, then we cannot multiply them.
So, let's assume they are both matrices of size $n \times n$.

Note:-

Don't worry, I have a program that generates these matrices for me.

Let $C = AB$, which means that it looks like:

$$C = \begin{pmatrix} \sum_{k=1}^{n} a_{1k}b_{k1} & \sum_{k=1}^{n} a_{1k}b_{k2} & \cdots & \sum_{k=1}^{n} a_{1k}b_{kn} \\ \sum_{k=1}^{n} a_{2k}b_{k1} & \sum_{k=1}^{n} a_{2k}b_{k2} & \cdots & \sum_{k=1}^{n} a_{2k}b_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} a_{nk}b_{k1} & \sum_{k=1}^{n} a_{nk}b_{k2} & \cdots & \sum_{k=1}^{n} a_{nk}b_{kn} \end{pmatrix}$$

$$C_{i,j} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

Let $D = BA$, which means that it looks like:

$$D = \begin{pmatrix} \sum_{k=1}^{n} b_{1k}a_{k1} & \sum_{k=1}^{n} b_{1k}a_{k2} & \cdots & \sum_{k=1}^{n} b_{1k}a_{kn} \\ \sum_{k=1}^{n} b_{2k}a_{k1} & \sum_{k=1}^{n} b_{2k}a_{k2} & \cdots & \sum_{k=1}^{n} b_{2k}a_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{n} b_{nk}a_{k1} & \sum_{k=1}^{n} b_{nk}a_{k2} & \cdots & \sum_{k=1}^{n} b_{nk}a_{kn} \end{pmatrix}$$

$$D_{i,j} = \sum_{k=1}^{n} b_{ik}a_{kj}$$

Let's take the trace of C and show it is equal to the trace of D:

$$\begin{aligned} tr(C) &= \sum_{i=1}^{n} c_{ii} \\ &= \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik}b_{ki} \\ &= \sum_{i=1}^{n} (a_{i1}b_{1i} + \ldots + a_{in}b_{ni}) \\ &= \sum_{i=1}^{n} (b_{1i}a_{i1} + b_{2i}a_{i2} + \ldots + b_{ni}a_{in}) \text{ by commutativity of } \cdot \text{ in } \mathbb{F} \\ &= (b_{11}a_{11} + \ldots + b_{n1}a_{1n}) + \ldots + (b_{1n}a_{n1} + \ldots + b_{nn}a_{nn}) \\ &= (b_{11}a_{11} + b_{12}a_{21} + \ldots + b_{1n}a_{n1}) + \ldots + (b_{n1}a_{1n} + \ldots + b_{nn}a_{nn}) \\ &= \sum_{i=1}^{n} (b_{i1}a_{1i} + b_{i2}a_{2i} + \ldots + b_{in}a_{ni}) \\ &= \sum_{i=1}^{n} \sum_{k=1}^{n} b_{ik}a_{ki} \\ &= \sum_{i=1}^{n} d_{ii} \\ &= tr(D) \end{aligned}$$

Thus, we have shown that $tr(AB) = tr(BA)$. 😄

Proof of c: This follows directly from part b.

Let A, B be two matrices size $n \times n$ with entries in $\mathbb{F}$.
Remember that we prove in part b, that $tr(AB) = tr(BA)$. Thus:

$$tr(BAB^{-1}) = tr(AB^{-1}B) = tr(AI), \text{ as } BB^{-1} = I$$

Remember that multiplying a matrix by the identity matrix does not change the matrix.
Thus,

$$tr(BAB^{-1}) = tr(AI) = tr(A)$$

This means that $tr(A) = tr(BAB^{-1})$, as desired. 😄

Consider the group:

$$\Gamma = \left\langle \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & -\sqrt{2} \\ \sqrt{2}+1 & \sqrt{2}+1 \end{bmatrix} \right\rangle$$