Notes on MAT224: Linear Algebra II

Unversity of Toronto

Xinyue Li

Last Updated: April 3, 2023 (draft)

Contents

Chapter 1

Vector Spaces

Section 1. Vector Spaces

1.1. Definition: A (real) **vector space** is a set V (whose elements are called **vectors**) together with

(a) an operation called **vector addition**, which for each pair of vectors $x, y \in V$ produces another vector in V denoted $x + y$, and

(b) an operation called **multiplication** by a **scalar** (a real number), which for each vector $x \in V$, and each scalar $c \in \mathbb{R}$ produces another vector in V denoted $c \cdot x$.

1.2. Axiom: *These two operations must satisfy the following axioms:*

(1) ***associativity:*** $\forall x, y, z \in V, (x + y) + z = x + (y + z)$

(2) ***commutativity:*** $\forall x, y \in V, x + y = y + x$

(3) ***existence of*** $\vec{0}$***:*** $\exists \vec{0} \in V \ s.t \ \forall x \in V, \ \vec{0} + x = x + \vec{0} = x$

(4) ***inverse:*** $\forall x \in V, \exists -x \in V \ \ s.t \ x + (-x) = \vec{0}$ *('−' just a symbol)*

(5) ***distributivity:*** $\forall c \in \mathbb{R}, x, y \in V, c \cdot (x + y) = c \cdot x + c \cdot y$

(6) $\forall c, d \in \mathbb{R}, x \in V, (c + d) \cdot x = c \cdot x + d \cdot x$

(7) $\forall c, d \in \mathbb{R}, x \in V, (c \cdot d) \cdot x = c \cdot (d \cdot x)$

(8) $\forall x \in V, 1 \cdot x = x$

(1)$\forall x, y, z \in V, (x + y) + z = x + (y + z)$

Proof. Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{R}^2$.

$$\begin{aligned} (x+y)+z &= (\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}) + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\ &= \begin{pmatrix} (x_1 + y_1) + z_1 \\ (x_2 + y_2) + z_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 + z_1 \\ y_2 + z_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}) \\ &= x + (y + z) \end{aligned}$$

□

(3)$\exists\vec{0} \in V \ s.t \ \forall x \in V, \ \vec{0} + x = x + \vec{0} = x$

Proof. Take $\vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$.

$$\vec{0} + x = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 + x_1 \\ 0 + x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x$$

□

(5)$\forall c \in \mathbb{R}, x, y \in V, c \cdot (x + y) = c \cdot x + c \cdot y$

Proof. Let $c \in \mathbb{R}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$.

$$\begin{aligned} c \cdot (x + y) &= c \cdot (\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}) \\ &= c \cdot \begin{pmatrix} x_1 + y_1 \\ x_1 + y_2 \end{pmatrix} \\ &= \begin{pmatrix} c \cdot (x_1 + y_1) \\ c \cdot (x_2 + y_2) \end{pmatrix} \\ &= \begin{pmatrix} c \cdot x_1 + c \cdot y_1 \\ c \cdot x_2 + c \cdot y_2 \end{pmatrix} \\ &= \begin{pmatrix} c \cdot x_1 \\ c \cdot x_2 \end{pmatrix} + \begin{pmatrix} c \cdot y_1 \\ c \cdot y_2 \end{pmatrix} \\ &= c \cdot x + c \cdot y \end{aligned}$$

□

1.3. Example: $\mathbb{R}^2, \mathbb{R}^3, \mathbb{R}^n$

1.4. Example: $Mat_{2\times 2} := \{2 \times 2 \text{ matrix}\}$
Let $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in Mat_{2\times 2}$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a + a' & b + b' \\ c + c' & d + d' \end{pmatrix}$$

Let $r \in \mathbb{R}$.

$$r \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ra & rb \\ rc & rd \end{pmatrix}$$

(6)$\forall c, d \in \mathbb{R}, x \in V, (c+d)\cdot x = c\cdot x + d\cdot x$

Proof. Let $r, \delta \in \mathbb{R}$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in Mat_{2\times 2}$.

$$\begin{aligned}
(r+\delta)\cdot\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \begin{pmatrix} (r+\delta)\cdot a & (r+\delta)\cdot b \\ (r+\delta)\cdot c & (r+\delta)\cdot d \end{pmatrix} \\
&= \begin{pmatrix} ra+\delta a & rb+\delta b \\ rc+\delta c & rd+\delta d \end{pmatrix} \\
&= \begin{pmatrix} ra & rb \\ rc & rd \end{pmatrix} + \begin{pmatrix} \delta a & \delta b \\ \delta c & \delta d \end{pmatrix} \\
&= r\cdot\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \delta\cdot\begin{pmatrix} a & b \\ c & d \end{pmatrix}
\end{aligned}$$

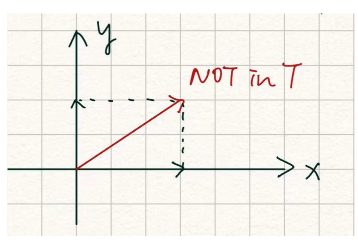
□

1.5. Example: $P_n(\mathbb{R}) = \{p(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_0 : a_n, a_{n-1}, ..., a_0 \in \mathbb{R}\}$
Let $p(x) = ax^2 + bx + c$ where $a, b, c \in \mathbb{R}$.Let $r \in \mathbb{R}$.

$$(ax^2 + bx + c) + (a'x^2 + b'x + c') = (a+a')x^2 + (b+b')x + (c+c')$$

$$r(ax^2 + bx + c) = (ra)x^2 + (rb)x + rc$$

1.6. Example (Counterexample): $T := x\text{-axis} \cup y\text{-axis} \in \mathbb{R}^2$
Q: Is$(T, +, \cdot)$ a vector space? A: No



1.7. Remark: In R^n there is clearly only one additive identity-the zero vector $(0, ..., 0) \in \mathbb{R}^n$. Moreover, each vector has only one additive inverse.

1.8. Proposition: *Let V be a vector space. Then*

a) The zero vector $\vec{0}$ is unique.

b) For all $x \in V, 0\cdot x = \vec{0}$

c) For each $x \in V$, the additive inverse $-x$ is unique.

d) For all $x \in V$ and all $c \in \mathbb{R}, (-c)\cdot x = -(c\cdot x)$

(a) The zero vector $\vec{0}$ is unique.

Proof. Suppose we had two vectors, $\vec{0}$ and $\vec{0}'$, both of which satisfy *Axiom* 3.
Then, $\vec{0}+\vec{0}'=\vec{0}$, since $\vec{0}'$ is an additive identity.
On the other hand, $\vec{0}+\vec{0}'=\vec{0}'+\vec{0}=\vec{0}'$, since addition is commutative and $\vec{0}$ is an additive identity.
Hence $\vec{0}=\vec{0}'$, or, in other words, there is only one additive identity in V. ∎

(b) For all $x \in V, 0 \cdot x = \vec{0}$

Proof. We have $0 \cdot x = (0+0)\cdot x = 0\cdot x + 0 \cdot x$, by *Axiom* 6.
Hence if we add the inverse of $0 \cdot x$ to both sides, we obtain $\vec{0} = 0 \cdot x$ ∎

(c) For each $x \in V$, the additive inverse $-x$ is unique.

Proof. Let $x \in V$, if $-x$ and $(-x)'$ are two inverses of x.
Then on one hand,by *Axioms* 1, 4, and 3,
we have $x+(-x)+(-x)' = (x+(-x))+(-x)' = \vec{0}+(-x)' = (-x)'$.
On the other hand, if we use *Axiom* 2 first before associating,
we have $x+(-x)+(-x)' = x+(-x)'+(-x) = (x+(-x)')+(-x) = \vec{0}+(-x) = -x$.
Hence, $-x = (-x)'$ and the additive inverse of x is unique. ∎

(d) For all $x \in V$ and all $c \in \mathbb{R}, (-c)\cdot x = -(c \cdot x)$.

Proof. We have $c \cdot x + (-c)\cdot x = (c+(-c))\cdot x = 0 \cdot x = \vec{0}$ by *Axiom* 6 and part b.
Hence $(-c)\cdot x$ also serves as an additive inverse for the vector $c \cdot x$.
By part c, therefore, we must have $(-c)\cdot x = -(c\cdot x)$ ∎

Section 2. Subspace

2.1. Example: Denote the set $V = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ is continuous}\}$ of functions by $C(\mathbb{R})$.

2.2. Lemma: *Let $f, g \in C(\mathbb{R})$, and let $c \in \mathbb{R}$. Then*

a) $f + g \in C(\mathbb{R})$

b) $cf \in C(\mathbb{R})$

(a) $f+g \in C(\mathbb{R})$

Proof. By the limit sum rule from calculus, for all $a \in \mathbb{R}$ we have

$$\lim_{x\to a}(f+g)(x) = \lim_{x\to a}(f(x)+g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$$

Since f and G are continuous, this last expression is equal to $f(a)+g(a) = (f+g)(a)$.
Hence $f+g$ is also continuous. ∎

(b) $cf \in C(\mathbb{R})$

Proof. By the limit product rule, we have

$$\lim_{x\to a}(cf)(x) = \lim_{x\to a} cf(x) = (\lim_{x\to a} c)\cdot(\lim_{x\to a} f(x)) = cf(a) = (cf)(a)$$

so cf is also continuous. □

2.3. Definition: Let V be a vector space and $W \subseteq V$ be a subset. If $(W, +, \cdot)$ itself is a vector space, then W is called **vector subspace** of V.

2.4. Example: $W := x - axis$ is a subset of $(\mathbb{R}^2, +, \cdot)$

$$\begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ 0 \end{pmatrix} \in W$$

$$c \cdot \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} c \cdot x_1 \\ c \cdot 0 \end{pmatrix} = \begin{pmatrix} cx \\ 0 \end{pmatrix} \in W$$

2.5. Theorem: Let V be a vector space, and let W be a nonempty subset of V. Then W is a vector subspace of V if and only if

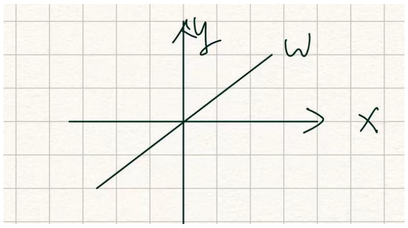
$$\forall w_1, w_2 \in W, \forall c \in \mathbb{R},\ \ c \cdot w_1 + w_2 \in W$$

2.6. Remark: By definition, a vector space must contain at least an additive identity element, hence the requirement that W be nonempty is certainly necessary.

Proof. $\rightarrow$: If W is subspace of V, because by the definition, a subspace W of V must be closed under vector sums and scalar multiples.
Then $\forall x \in W$, and $\forall c \in \mathbb{R}$, we have $c \cdot x \in W$.
And hence $\forall y \in W, c \cdot x + y \in W$ as well.

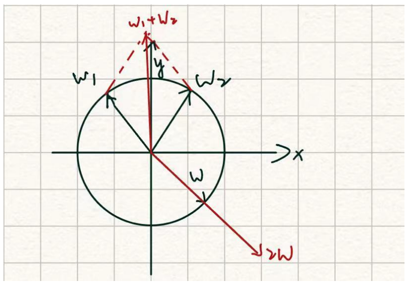
$\leftarrow$: Let W be any subset of V satisfying the condition of the theorem.
First, note that since $\forall x, y \in W, \forall c \in \mathbb{R}, c \cdot x + y \in W$.
We may specialize to the case $c = 1$.
Then we see that $1 \cdot x + y = x + y \in W$, so that W is closed under sums.
Next, let $x = y$ be any vector in W and $c = -1$.
Then $(-1) \cdot x + x = (-1 + 1) \cdot x = 0 \cdot x = \vec{0} \in W$.
Now .et x be any vector in W and let $y = 0$.
Then $c \cdot x + \vec{0} = c \cdot x \in W$.
So W is closed under scalar multiplication.
To see that these observations imply that W is a vector space,
note that the $Axioms$ $1, 2, 5$ through 8 are satisfied automatically for vectors in W.
Since they hold for all vectors in V.
$Axiom$ 3 is satisfied, since as we have seen $\vec{0} \in W$.
Finally, for each $x \in W$, by Proposition $(1.8d)$ $(-1) \cdot x = -x \in W$ as well.
Hence W is a vector space. □

2.7. Example: $W = \{(x, mx) \in \mathbb{R}^2 : x \in \mathbb{R}, m \neq 0\}$ is a subset of $(\mathbb{R}^2, +, \cdot)$



That is a vector subspace.

2.8. Example: $W = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$



That is not a vector subspace.

2.9. Example: $W := (x - axis) \cup (y - axis) = \{(0, 0)\}$ is a vector subspace.

2.10. Theorem: Let V be a vector space. Then the intersection of any collection of subspaces of V is a subspace of V.
Let $W_1, W_2, ..., W_k$ br vector subspaces of V. Then

$$W_1 \cap W_2 \cap ... \cap W_k$$

is a vector subspace of V.

Proof. Consider any collection of subspaces of V.
Note first that the intersection of the subspaces is nonempty.
Since it contains at least the zero vector from V.
Now, let x, y be any two vectors in the intersection of all the subspaces in the collection
($i.e.$ $x, y \in W$ for all W in the collection).
Since each W in the collection is a subspace of $V, c \cdot x + y \in W$.
Since this is true for all the W in the collection, $cx + y$ is in the intersection of all the subspaces in the collection.
Hence, the intersection is a subspace of V by Theorem (2.5).

□

2.11. Example: In $V = \mathbb{R}^3$ consider the subset

$$W = \{(x_1, x_2, x_3) \mid 4x_1 + 3x_2 - 2x_3 = 0 \text{ and } x_1 - x_3 = 0\}$$

Let $x, y \in W$ and $c \in \mathbb{R}$.
Then writing $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$, we have that the components of the vector

$$c \cdot x + y = (cx_1 + y_1,\ cx_2 + y_2,\ cx_3 + y_3)$$

satisfy the defining equations of the set W:

$$4(cx_1 + y_1) + 3(cx_2 + y_2) - 2(cx_3 + y_3) = c(4x_1 + 3x_2 - 2x_3) + (4y_1 + 3y_2 - 2y_3) = c0 + 0 = 0$$

Similarly.

$$(cx_1 + y_1) - (cx_3 + y_3) = c(x_1 - x_3) + (y_1 - y_3) = c0 + 0 = 0$$

Hence $c \cdot x + y \in W$. W is nonempty since the zero vector $(0, 0, 0)$ satisfies both equations.

2.12. Example: In $\mathbb{R}^n$, let $V = \{(x_1, ..., x_n) \in \mathbb{R}^n \mid a_1x_1 + ... + a_nx_n = 0,$ where $a_i \in \mathbb{R}$ for all $i\}$. Then V is a vector space, if we define the vector sum and scalar multiplication to be the same as the operations in the whole space $\mathbb{R}^n$.

2.13. Corollary: *Let $a_{ij}(1 \leq i \leq m, 1 \leq j \leq n)$ be any real numbers and let*

$$W = \{(x_1, ..., x_n) \in \mathbb{R}^n \mid a_{i1}x_1 + ... + a_{in}x_n = 0 \textit{ for all } i, 1 \leq i \leq m\}.$$

Then W is a subspace of R^n.

Proof. For each $i, 1 \leq i \leq m$, let $W = \{(x_1, ..., x_n) \mid a_{i1}x_1 + ... + a_{in}x_n = 0 \text{ for all } i, 1 \leq i \leq m\}$.
Then since W is precisely the set of solutions of the simutaneous system formed from the defining equations of all the W_i.
We have $W = W_1 \cap W_2 \cap ... \cap W_m$.
Each W_i is a subspace of $\mathbb{R}^n$ [see *Example*(2.12)], so W is also a subspace of $\mathbb{R}^n$. □

Section 3. Linear Combination

3.1. Definition: Let V be a vector space. A **linear combination** of vectors in V is any sum

$$a_1v_1 + a_2v_2 + ... + a_nv_n$$

where $a_1, a_2, ..., a_n \in \mathbb{R}$ and $v_1, v_2, ..., v_n \in V$

3.2. Definition: Let V be a vector space and S be a subset of V. Then **span** Span(S) of S in V is the set of all **linear combinations** of vectors in S.

$$\text{Span}(S) = \{a_1x_1 + a_2x_2 + ... + a_nx_n : x_1, x_2, ..., x_n \in S, a_1, a_2, ..., a_n \in \mathbb{R}\}$$

3.3. Example: Let V be any vector space and $S = \{0\}$.
Then $\text{Span}(S) = \{a \cdot 0 : a \in \mathbb{R}\} = \{0\}$.

3.4. Example: Let $V = \mathbb{R}^2$ and $S = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\}$.

Then $\text{Span}(S) = \{a \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} : a \in \mathbb{R}\} = \{\begin{pmatrix} a \\ 0 \end{pmatrix}\} =$ x-axis.

3.5. Definition: If $S = \emptyset$, we define $\text{Span}(S) = \{0\}$.

3.6. Definition: If W=Span(S), then we say that S spans W.

3.7. Example: Let $V = \mathbb{R}^2$ and $S = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$.

Then $\text{Span}(S) = \{a_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} : a_1, a_2 \in \mathbb{R}\} = \{\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\} = \mathbb{R}^2$.

Therefore, $\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$ spans $\mathbb{R}^2$.

3.8. Theorem: Let V be a vector space and let S be any subset of V. Then Span(S) is always a vector subspace of V.

Proof. We prove this by applying Theorem (2.5). Span(S) is non-empty by definition.
Furthermore, let $x, y \in \text{Span}(S)$, and let $c \in \mathbb{R}$.
Then we can write $x = a_1x_1 + ... + a_nx_n$, with $a_i \in \mathbb{R}$ and $x_i \in S$.
Similarly, we can write $y = b_1x'_1 + ... + bmx'_m$, with $b_i \in \mathbb{R}$ and $x'_i \in S$.
Then for any scalar c we have

$$\begin{aligned} cx + y &= c(a_1x_1 + ... + a_nx_n) + b_1x'_1 + ... + bmx'_m \\ &= ca_1x_1 + ... + ca_nx_n + b_1x'_1 + ... + bmx'_m \end{aligned}$$

Since this is also a linear combination of the vectors in the set S, we have that $cx + y \in \text{Span}(S)$.
Hence Span(S) is a subspace of V. □

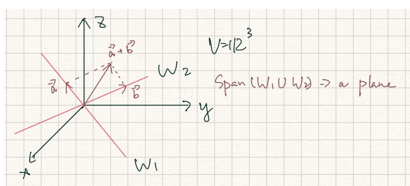
3.9. Definition: Let W_1, W_2 be subspaces of a vector space V. The **sum** $W_1 + W_2$ is the set

$$W_1 + W_2 = \{x \in V : x = x_1 + x_2 \text{ for some } x_1 \in W_1 \text{ and } x_2 \in W_2\}. = \text{Span}(W_1 \cup W_2)$$

3.10. Proposition: *Let V be vector space and S_1, S_2 be subsets of V. Let $W_1 = Span(S_1)$ and $W_2 = Span(S_2)$. Then $W_1 + W_2 = Span(S_1 \cup S_2)$.*

Proof. To see that $W_1 + W_2 \subseteq \text{Span}(S_1 \cup S_2)$. Let $v \in W_1 + W_2$.
Then $v = v_1 + v_2$, where $v_1 \in W_1$ and $v_2 \in W_2$.
Since $W_1 = \text{Span}(S_1)$, we can write $v_1 = a_1x_1 + ... + a_mx_m$, where each $x_1 \in S_1$ and each $a_i \in \mathbb{R}$.
Similarly, we can write $v_2 = b_1y_1 + ... + b_ny_n$, where each $y_i \in S_2$ and each $b_i \in \mathbb{R}$.
Hence we have $v = b_1y_1 + ... + b_ny_n + a_1x_1 + ... + a_mx_m$.
This is a linear combination of vectors that are either in S_1 or in S_2.
Hence $v \in \text{Span}(S_1 \cup S_2)$, and since this is true for all such v, $W_1 + W_2 \subseteq \text{Span}(S_1 \cup S_2)$.
Conversely, to see that $\text{Span}(S_1 \cup S_2) \subseteq W_1 + W_2$, we note that if $v \in \text{Span}(S_1 \cup S_2)$.
Then $v = c_1z_1 + ... + c_lz_l$, where each $z_k \in S_1 \cup S_2$ and each $c_k \in \mathbb{R}$.
Each z_k is in S_1 or in S_2, so by renaming the vectors and regrouping the terms.
We have $v = b_1y_1 + ... + b_ny_n + a_1x_1 + ... + a_mx_m$, where each $x_i \in S_1$ and $y_i \in S_2$.
Hence, by definition, we have written v as the sum of a vector in W_1 and a vector in W_2.
So $v \in W_1 + W_2$.
Since this is true for all $v \in \text{Span}(S_1 \cup S_2)$, we have $\text{Span}(S_1 \cup S_2) \subseteq W_1 + W_2$. □

3.11. Example: $W_1 \cup W_2$ is **NOT** a vector space but $\text{Span}(W_1 \cup W_2)$ is.



3.12. Theorem: Let W_1 and W_2 be subspaces of a vector space V. Then $W_1 + W_2$ is also a subspace of V.

Proof. It is clear that $W_1 + W_2$ is nonempty, since W_1 and W_2 are nonempty.
Let x, y be any two vectors in $W_1 + W_2$ and let $c \in \mathbb{R}$.
Since x and $y \in W_1 + W_2$.
We can write $x = x_1 + x_2, y = y_1 + y_2$, where $x_1, y_1 \in W_1$ and $x_2, y_2 \in W_2$.
Then we have

$$\begin{aligned} cx + y &= c(x_1 + x_2) + (y_1 + y_2) \\ &= (cx_1 + y_1) + (cx_2 + y_2) \end{aligned}$$

Since W_1 and W_2 are subspaces of V, we have $cx_1 + y_1 \in W_1$ and $cx_2 + y_2 \in W_2$.
Hence by the definition, $cx + y \in W_1 + W_2$.
By Theorem (2.5), $W_1 + W_2$ is a subspace of V. □

3.13. Remark: In general, if W_1 and W_2 are subspaces of V, then $W_1 \cup W_2$ will note be a subspaces of $\mathbb{R}^2$ given in Example (3.11).

3.14. Proposition: *Let W_1 and W_2 be subspaces of a vector space V, and let W be a subspace of V such that $W_1 \cup W_2 \subseteq W$. Then $W_1 + W_2 \subseteq W$.*

Proof. Let $v_1 \in W_1$ and $v_2 \in W_2$ be any vectors.
Since $v_1 \in W_1 \subset W_1 \cup W_2$, $v_1 \in W$ as well.
Similarly, $v_2 \in W$.
Hence, since W is a subspace of V, $v_1 + v_2 \in W$.
But this shows that every vector in $W_1 + W_2$ is contained in W. □

3.15. Example: Let $V = \mathbb{R}^3, W = xy - plane, W_1 = x - axis, W_2 = y - axis$.

$$W_1 + W_2 = \text{Span}(x - axis \cup y - axis) = xy - plane \subseteq W$$

3.16. Example: Let $V = Pol_{\leq 5} = \{a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 : a_0, ..., a_5 \in \mathbb{R}\}$,
$W = Pol_{\leq 4} = \{a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 : a_0, ..., a_4 \in \mathbb{R}\}$,
$W_1 = Pol_{\leq 1} = \{a_1x + a_0 : a_0, a_1 \in \mathbb{R}\}$ and $W_2 = \{a_2x^2 + a_1x + a_0 : a_0 + a_1 + a_2 = 0\}$.

$$\begin{aligned} W_1 + W_2 &= \text{Span}(W_1 \cup W_2) \\ &= \{c_1(a_1x + a_0) + c_2(b_2x^2 + b_1x + b_0) : b_0 + b_1 + b_2 = 0\} \end{aligned}$$

Given an arbitrary polynomial: $d_2x_2 + dx + d_0$ of degree ≤ 2, we have

$$d_2x^2 + d_1x + d_0 = 1 \cdot ((d_1 + d_2)x + d_0) + 1 \cdot (d_2x^2 + (-d_2)x + 0)$$

Hence $W_1 + W_2 = Pol_{\leq 2} \subset Pol_{\leq 4} = W$

Section 4. Linear Independence

4.1. Definition: Let V be a vector space and S be a subset of V.
A **linear dependence** among elements of S is a relationship of the form

$$a_1x_1 + a_2x_2 + ... + a_nx_n = 0$$

where $x_1, x_2, ..., x_n \in S$ and $a_1, a_2, ..., a_n \in \mathbb{R}$ are not all zero.
We say S is **linearly dependent** if there exists a linear dependence among elements of S.

4.2. Example: Let V be any vector space and $S = \{0\}$.
For any $a \in \mathbb{R}, a \neq 0$, we have $a \cdot 0 = 0$ which is a linear dependence.
Therefore, S is linearly dependent.

4.3. Example: Let $V = \mathbb{R}^2$ and $S = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$.

Assume $\exists a_1, a_2 \in \mathbb{R}$ $s.t$ $a_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Then

$$\begin{pmatrix} a_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad i.e \quad \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad i.e \quad \begin{cases} a_1 = 0 \\ a_2 = 0 \end{cases}$$

Hence, there is no linear dependence.
In other word, S is linearly independent.

4.4. Example: Let $V = \mathbb{R}^2$ and $S = \{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\}$.
We have

$$1 \cdot \begin{pmatrix}1\\0\end{pmatrix} + 1 \cdot \begin{pmatrix}0\\1\end{pmatrix} + (-1) \cdot \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

Hence, S is linearly dependent.

4.5. Example: Let $V = \mathbb{R}^2$ and $S = \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\2\end{pmatrix}, \begin{pmatrix}3\\2\end{pmatrix}\}$.
We need to solve

$$a_1 \begin{pmatrix}1\\1\end{pmatrix} + a_2 \begin{pmatrix}-1\\2\end{pmatrix} + a_3 \begin{pmatrix}3\\2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

where $a_1 . a_2, a_3 \in \mathbb{R}$.
Equivalent need to solve

$$\begin{cases} a_1 - a_2 + 3a_3 = 0 \\ a_1 + 2a_2 + 2a_3 = 0 \end{cases}$$

Take $a_1 = t \in \mathbb{R}$, need to solve

$$\begin{cases} a_2 - 3a_3 = t \\ 2a_2 + 2a_3 = -t \end{cases} \quad i.e \quad \begin{cases} a_2 = -\frac{1}{8}t \\ a_3 = -\frac{3}{8}t \end{cases}$$

Take t to be any non-zero real number, get a linear dependence.

4.6. Proposition:

(a) *Let S be a linearly dependent subset of a vector space V, and let S' be another subset of V that contains S. Then S' is also linearly dependent.*

(b) *Let S be a linearly independent subset of vector space V and let S' be another subset of V that is contained in S. Then S' is also linearly independent.*

Section 5. Bases and Dimension

5.1. Definition: A subset S of a vector space V is called a **basis** of V if

(1) $V = \text{Span}(S)$ and

(2) S is linearly independent.

5.2. Example: Let $V = \mathbb{R}^2$ and $S = \{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\}$.
Since $\text{Span}(S) = \mathbb{R}^2$ [see Example (3.7)] and S is linearly independent [see Example (4.3)].
Thus S is a basis of $\mathbb{R}^2$.

5.3. Example: Let $V = \mathbb{R}^2$ and $S = \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\2\end{pmatrix}\}$.

$$\begin{aligned}\text{Span}(S) &= \{a_1 \cdot \begin{pmatrix}1\\1\end{pmatrix} + a_2 \cdot \begin{pmatrix}-1\\2\end{pmatrix} : a_1, a_2 \in \mathbb{R}\}\\ &= \{\begin{pmatrix}a_1 - a_2\\a_1 + a_2\end{pmatrix} : a_1, a_2 \in \mathbb{R}\} = \mathbb{R}^2\end{aligned}$$

For all $\begin{pmatrix}c\\d\end{pmatrix}$, we want to solve $\begin{pmatrix}a_1 - a_2\\a_1 + a_2\end{pmatrix} = \begin{pmatrix}c\\d\end{pmatrix}$.

$$\begin{cases} a_1 - a_2 = c \\ a_1 + a_2 = d \end{cases} \implies \begin{cases} a_1 = \frac{1}{3}(d + 2c) \\ a_2 = \frac{1}{3}(d - c) \end{cases}$$

Want to solve $a_1 \cdot \begin{pmatrix}1\\1\end{pmatrix} + a_2 \cdot \begin{pmatrix}-1\\2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$.
This is the special case where $c = d = 0$. So must have $a_1 = a_2 = 0$.
Therefore, S is linearly independent.
Thus S is a basis of V.

5.4. Example: Let $V = Pol_{\leq 2}, S = \{1, 1 + x, 1 + x + x^2\}$.

$$\begin{aligned}\text{Span}(S) &= \{a_1 \cdot 1 + a_2 \cdot (1 + x) + a_3 \cdot (1 + x + x^2) : a_1, a_2, a_3 \in \mathbb{R}\}\\ &= \{(a_1 + a_2 + a_3) + (a_2 + a_3)x + a_3x^2 : a_1, a_2, a_3 \in \mathbb{R}\}\\ &\subseteq Pol_{\leq 2}\end{aligned}$$

For all $b_0 + b_1x_+b_2x^2 \in Pol_{\leq 2}$, need to solve

$$b_0 + b_1x + b_2x^2 = (a_1 + a_2 + a_3) + (a_2 + a_3)x + a_3x^2$$

$$\Leftrightarrow \begin{cases} a_1 + a_2 + a_3 = b_0 \\ a_2 + a_3 = b_1 \\ a_3 = b_2 \end{cases} \Leftrightarrow \begin{cases} a_1 = b_0 - b_1 \\ a_2 = b_1 - b_2 \\ a_3 = b_2 \end{cases}$$

Thus $\text{Span}(S) \supseteq V \implies \text{Span}(S) = V$.
Suppose $a_1 \cdot 1 + a_2 \cdot (1 + x) + a_3 \cdot (1 + x + x^2) = 0$.
(Note: " = " means the equality of polynomials)
This is the special case where $b_0 = b_1 = b_2 = 0$.
Get $a_1 = a_2 = a_3 = 0$. Hence S is linearly independent.
Thus S is a basis of V.

5.5. Theorem: Let V be a vector space, and let S be a nonempty subset of V.
Then S is a basis of V if and only if every vector $x \in V$ may be written uniquely as a linear combination of the vectors in S.

5.6. Theorem: Let V be a vector space that has a finite spanning set, and let S be a linearly independent subset of V. Then there exists a basis S' of V, with $S \subseteq S'$.

5.7. Remark: The content of the theorem is usually summarized by saying that every linearly independent set may be **extended to a basis** (by adjoining further vectors).

5.8. Lemma: *Let S be a linearly independent subset of V and let $x \in V$, but $x \notin S$.
Then $S \cup \{x\}$ is linearly independent if and only if $x \notin Span(S)$.*

Proof. Let $T = \{y_1, ..., y_n\}$ be a finite set that spans V, and let $S = \{x_1, ..., x_m\}$ be a linearly independent set in V.
We claim the following process will produce a basis of V.
First, by start by setting $S' = S$.
Then, for each $y_1 \in T$ in turn do the following:
If $S' \cup \{y_1\}$ is linearly independent, replace the current S' by $S' \cup \{y_1\}$.
Otherwise, leave S' unchanged. Then go on to the next y_i.
When the "loop" is completed, S' will be a basis of V.
To see why this works, note first that we are only including the y_i such that $S' \cup \{y_i\}$ is linearly independent at each stage.
Hence the final set S' will also be a linearly independent set.
Second, note that every $y_i \in T$ is in the span of the final set S'.
Since that set contains all the y_i that are adjoined to the original S.
On the other hand, by Lemma (5.8), each time the current $S' \cup \{y_i\}$ is not linearly independent, that $y_i \in \text{Span}(S')$ already.
Since T spans V, and every vector in T is in $\text{Span}(S')$, it follows that S' spans V as well.
Hence S' is a basis of V. □

5.9. Example: Let $V = \mathbb{R}^2, S = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\}$.
S is linearly independent, bu not a basis.
Take $S' = \{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\}$.
S' is a basis for $\mathbb{R}^2$. Also, $S \subseteq S'$.

5.10. Theorem: Let V be a vector space and let S be a spanning set for V, which has m elements. Then no linearly independent set in V can have more than m element.

5.11. Corollary: *Let V be a vector space and let S and S' be two bases of V, with m and m' elements, respectively. Then $m = m'$.*

Proof. Since S spans V and S' is linearly independent, by Theorem (5.10) we have that $m \geq m'$. On the other hand, since S' spans V and S is linearly independent, by Theorem (5.10) again, $m' \geq m$. It follows that $m = m'$. $\square$

5.12. Definition:

(a) If V is a vector space with some finite basis (possibly empty), we say V is **finite-dimensional**.

(b) Let V be a finite-dimensional vector space. The **dimension** of V, denoted $\dim(V)$, is the number of vectors in a (hence any) basis of V.

(c) If $V = \{0\}$, we define $\dim(V) = 0$.

5.13. Theorem: Let W be a subspace of a finite dimensional vector space V. Then $\dim(W) \leq \dim(V)$. Furthermore, $\dim(W) = \dim(V)$ if and only if $W = V$.

5.14. Example:
Let $V = Pol_{\leq 3}, W = Pol_{\leq 2}$.
$\dim(V) = 4 \leq 3 = \dim(W)$.

5.15. Corollary: *Let W be a subspace of $\mathbb{R}^n$ defined by a system of homogeneous linear equations. Then $\dim(W)$ is equal to the number of free variables in the corresponding echelon form system.*

5.16. Theorem: Let W_1 and W_2 be finite-dimensional subspaces of a vector space V. Then

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$$

5.17. Example: Let $V = \mathbb{R}^2, W_1 = W_2 = x - axis$.
$\dim(W_1 + W_2) = \dim(x - axis) = 1$
$\dim(W_1) + \dim(W_2) = 1 + 1 = 2$
$\dim(W_1 \cap W_2) = 1$

Chapter 2

Linear Transformation

Section 1. Linear Transformations

1.1. Definition: A function $T : V \to W$ is called a **linear transformation** or a **linear mapping** if it satisfies

(i) $T(u+v) = T(u) + T(v)$ for all u and $v \in V$

(ii) $T(av) = aT(v)$ for all $a \in \mathbb{R}$ and $v \in V$.

V is called the **domain** of T and W is called the **target** of T.

1.2. Example: Let $f : \mathbb{R}^2 \to \mathbb{R}^2$, $\begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} 2a-b \\ b \end{pmatrix}$.

$$\begin{aligned} f(\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix}) &= \begin{pmatrix} 2(a+c)-(b+d) \\ (b+d) \end{pmatrix} \\ &= \begin{pmatrix} 2a+2c-b-d \\ b+d \end{pmatrix} \\ &= \begin{pmatrix} 2a-b \\ b \end{pmatrix} + \begin{pmatrix} 2c-d \\ d \end{pmatrix} \\ &= f(\begin{pmatrix} a \\ b \end{pmatrix}) + f(\begin{pmatrix} c \\ d \end{pmatrix}) \end{aligned}$$

$$f(c \cdot \begin{pmatrix} a \\ b \end{pmatrix}) = f(\begin{pmatrix} ca \\ cb \end{pmatrix}) = \begin{pmatrix} 2(ca)-(cb) \\ cb \end{pmatrix} = \begin{pmatrix} c(2a-b) \\ cb \end{pmatrix} = c \cdot f(\begin{pmatrix} a \\ b \end{pmatrix})$$

1.3. Corollary: *If a function $T : V \to W$ is a linear transformation. We write 0_v for the zero vector in V and 0_w for the zero vector in W. Then,*

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

1.4. Proposition: *A function $T : V \to W$ is a linear transformation if and only if for all a and $b \in \mathbb{R}$ and all u and $v \in V$*

$$T(au+bv) = aT(u) + bT(v)$$

1.5. Corollary: *A function $T : V \to W$ is a linear transformation if and only if for all $a_1, ..., a_k \in \mathbb{R}$ and for all $v_1, ...v_k \in V$:*

$$T(\sum_{i=1}^{k} a_i v_i) = \sum_{i=1}^{k} a_i T(v_i)$$

1.6. Remark:

(i) **identity transformation**: $I : V \to V$

(ii) **zero transformation**: $T(v) = 0_w$

1.7. Example:
Let V be the vector space $C^x(\mathbb{R})$ (of functions $f : \mathbb{R} \to \mathbb{R}$ with derivatives of all orders.)
Let $D : C^x(\mathbb{R}) \to C^x(\mathbb{R})$ be the mapping that takes each function $F \in C^x(\mathbb{R})$ to its derivative function $D(f) = f' \in C^x(\mathbb{R})$
D is a linear transformation.

1.8. Example: Let V denote the vector space $C[a, b]$ of continuous functions on the closed interval $[a, b] \subset \mathbb{R}$, and let $W = \mathbb{R}$.
Define Int:$V \to W$ by the rule $\text{Int}(f) = \int_a^b f(x)dx \in \mathbb{R}$.
Int is a linear transformation.

1.9. Definition: Let $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$. The **linear product** of $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ is

$$< \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} > := x_1 y_1 + x_2 y_2$$

1.10. Definition: If the line segment is a vector $\vec{v}$, its length is denoted $\|\vec{v}\|$ such that

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2} = \sqrt{< v, v >}$$

1.11. Proposition: *If $\vec{a}$ and $\vec{b}$ are none-zero vectors in $\mathbb{R}^2$, the angle θ between $\vec{a}$ and $\vec{b}$ satisfies*

$$\cos(\theta) = \frac{< \vec{a}, \vec{b} >}{\|\vec{a}\|\|\vec{b}\|}$$

Proof. Use Cosine Law,

$$\begin{aligned} \|x\|^2 + \|y\|^2 - \|x - y\|^2 &= 2\|y\|\|x\| \cos\theta \\ < x, x > + < y, y > - < x - y, x - y > &= 2\|y\|\|x\| \cos\theta \end{aligned}$$

Since

$$\begin{aligned} < x - y, x - y > &= < \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \end{pmatrix}, \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \end{pmatrix} > \\ &= (x_1 - y_1)^2 + (x_2 - y_2)^2 \\ &= x_1^2 - 2x_1 y_1 + y_1^2 + x_2^2 - 2x_2 y_2 + y_2^2 \\ &= \|x\|^2 + \|y\|^2 - 2 < x, y > \end{aligned}$$

Thus,

$$R.H.S = 2 < x, y > = 2\|x\|\|y\|\cos\theta$$
$$\cos\theta = \frac{< x, y >}{\|x\|\|y\|}$$

□

1.12. Corollary: *If $\vec{a} \neq \vec{0}$ and $\vec{b} \neq \vec{0}$ are vectors in $\mathbb{R}^2$, $\vec{a} \perp \vec{b}$ if and only if $< \vec{a}, \vec{b} >= 0$.*

Proof. Since $\vec{a} \perp \vec{b}$ implies that $\theta = \pm\frac{\pi}{2} \Leftrightarrow \cos\theta = 0 \Leftrightarrow \frac{< a, b >}{\|a\|\|b\|} = 0$.
Since $\|a\|\|b\|$ is non-zero.
Thus $< a, b >= 0$. □

1.13. Example: Rotation through an angle θ. Let $V = W = \mathbb{R}^2$, and let θ be a fixed real number that represents an angle in radians. Define a function $R_\theta : V \to V$ by

$$R_\theta(v) = \mathbb{R}^2 \to \mathbb{R}^2, x \mapsto \text{ rotation of } x \text{ by } \varphi$$

(Note that a positive angle is measured in a counterclockwise manner, whereas a negative angle is measured in a clockwise manner.)
If $w = R_\theta(v)$, then the expression for w in terms of its length and the angle it makes with the first coordinate axis is

$$w = \|v\| \begin{pmatrix} \cos(\varphi + \theta) \\ \sin(\varphi + \theta) \end{pmatrix}$$

where φ is the angle v makes with the first coordinate axis.
Using the formulas for cosine and sine of a sum of angles, we obtain:

$$w = \begin{pmatrix} \cos(\varphi) \cdot \cos(\theta) - \sin(\varphi) \cdot \sin(\theta) \\ \cos(\varphi) \cdot \sin(\theta) + \sin(\varphi) \cdot \cos(\theta) \end{pmatrix} = \begin{pmatrix} v_1 \cos(\theta) - v_2 \sin(\theta) \\ v_1 \sin(\theta) + v_2 \cos(\theta) \end{pmatrix}$$

Using this algebraic expression for $R_\theta(v)$ we can easily check that R_θ is a linear transformation.
Let $a, b \in \mathbb{R}$ and $u = (u_1, u_2)$ and $v = (v_1, v_2) \in \mathbb{R}^2$, then

$$\begin{aligned}
R_\theta(au + bv) &= R_\theta(\begin{pmatrix} au_1 + bv_1 \\ au_2 + bv_2 \end{pmatrix}) \\
&= \begin{pmatrix} (au_1 + bv_1)\cos(\theta) - (au_2 + bv_2)\sin(\theta) \\ (au_1 + bv_1)\sin(\theta) + (au_2 + bv_2)\cos(\theta) \end{pmatrix} \\
&= \begin{pmatrix} a(u_1 \cos(\theta) - u_2 \sin(\theta)) + b(v_1 \cos(\theta) - v_2 \sin(\theta)) \\ a(u_1 \sin(\theta) + u_2 \cos(\theta)) + b(v_1 \sin(\theta) + v_2 \cos(\theta)) \end{pmatrix} \\
&= a \begin{pmatrix} u_1 \cos(\theta) - u_2 \sin(\theta) \\ u_1 \sin(\theta) + u_2 \cos(\theta) \end{pmatrix} + b \begin{pmatrix} v_1 \cos(\theta) - v_2 \sin(\theta) \\ v_1 \sin(\theta) + v_2 \cos(\theta) \end{pmatrix} \\
&= aR_\theta(u) + bR_\theta(v)
\end{aligned}$$

Therefore, R_θ is a linear transformation.

1.14. Example: Let v be a nonzero vector in $\mathbb{R}^2$ and L be the line containing v. Define $proj_L : \mathbb{R}^2 \to L$. Then,

$$\|proj_L(x)\| = |\|x\| \cdot \cos(\theta)| = \|x\| \cdot \frac{|<x, v>|}{\|x\|\|v\|}$$

Suppose $proj_L(x) = c \cdot v$. Then

$$\begin{aligned} \|proj_L(x)\| &= \|c \cdot v\| = \|c\|\|v\| \\ \|x\| \cdot \frac{|<x, v>|}{\|x\|\|v\|} &= |c| \cdot \|v\| \\ |c| &= \frac{|<x, v>|}{\|v\|^2} \\ c &= \pm \frac{|<x, v>|}{\|v\|^2} \end{aligned}$$

Therefore, $proj_L(x) = \dfrac{|<x, v>|}{\|v\|^2} \cdot v$. Write $v = (v_1, v_2), x = (x_1, x_2)$. Then,

$$\begin{aligned} proj_L(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}) &= \frac{x_1 v_1 + x_2 v_2}{\|v\|^2} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{x_1 v_1 + x_2 v_2}{\|v\|^2} v_1 \\ \frac{x_1 v_1 + x_2 v_2}{\|v\|^2} v_2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{v_1^2}{\|v\|^2} x_1 + \frac{v_1 v_2}{\|v\|^2} x_2 \\ \frac{v_1 v_2}{\|v\|^2} x_1 + \frac{v_1^2}{\|v\|^2} x_2 \end{pmatrix} \end{aligned}$$

1.15. Proposition: *If $f : V \to W$ is a linear transformation and V is finite dimensional with basis $\{v_1, ..., v_n\}$. Suppose that we know $f(v_1), ..., f(v_n)$, then we know f if $v = a_1 v_1 + ... + a_n v_n$ where $a_1, ..., a_n \in \mathbb{R}$, then $f(v) = a_1 f(v_1) + ... + a_n f(v_n)$.*

1.16. Example: Let $V = W = Pol_{\leq 2}$.
A basis for V is given by the polynomials $\{1, 1 + x, 1 + x + x^2\}$.
Define T on this basis by $T(1) = x, T(1 + x) = x^2, T(1 + x + x^2) = 1$.
If we insist that T be linear, this defines a linear transformation.
If $p(x) = a_2 x^2 + a_1 x + a_0$, then the equation

$$p(x) = (a_0 - a_1)1 + (a_1 - a_2)(1 + x) + a_2(1 + x + x^2)$$

express $p(x)$ in terms of the basis.
Therefore, we can get $T(p(x)) = (a_1 - a_2)x^2 + (a_0 - a_1)x + a_2$.

Section 2. Linear Transformations Between Finite Dimensional Vector Spaces

2.1. Proposition: *Let $T : V \to W$ be a linear transformation between the finite dimensional vector spaces V and W. If $\{v_1, ..., v_k\}$ is a basis for V and $\{w_1, ..., w_l\}$ is a basis for W, then $T : V \to W$ is uniquely determined by the $l \cdot k$ scalars used to express $T(v_j), j = 1, ..., k$, in terms of $w_1, ..., w_l$.*

2.2. Example: Let $V = W = \mathbb{R}^2$. Choose the standard basis $e_1 = (1, 0)$ and $e_2 = (0, 1)$ for both V and W. Define T by $T(e_1) = e_1 + e_2$ and $T(e_2) = 2e_1 - 2e_2$. The four scalars $a_{11} = 1, a_{21} = 1, a_{12} = 2, a_{22} = -2$ determine T.

2.3. Definition: Let $a_{ij}, 1 \leq i \leq l$ and $1 \leq j \leq k$ be $l \cdot k$ scalars. The matrix whose entries are the scalars a_{ij} is the rectangular array of l rows and k columns:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1k} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2k} \\ & \dots & & & \\ a_{l1} & a_{l2} & a_{l3} & \dots & a_{lk} \end{bmatrix}$$

Thus, the scalar a_{ij} is the entry in the ith row and the j th column of the array. A matrix with l rows and k columns will be called an $l \times k$ matrix.

2.4. Remark: If we begin with a linear transformation between finite-dimensional vector spaces V and W, the transformation is determined by the choice of bases in V and W and a set of $l \cdot k$ scalars, where $k = \dim(V)$ and $l = \dim(W)$.

2.5. Definition: Let $T : V \to W$ be a linear transformation between the finite dimensional vector spaces V and W.
Let $\alpha = \{v_1, ..., v_k\}$ and $\beta = \{w_1, ..., w_l\}$, respectively, be any bases for V and W.
Let $a_{ij}, 1 \leq i \leq l$ and $1 \leq j \leq k$ be the $l \cdot k$ scalars that determine T with respect to the bases α and β.
The matrix whose entries are the scalars$a_{ij}, 1 \leq i \leq l$ and $1 \leq j \leq k$, is called the **matrix of the linear transformation T with respect to the bases α for V and β for W.**
This matrix is denoted by $[T]_\alpha^\beta$.

2.6. Example: Let $f : \mathbb{R}^2 \to \mathbb{R}^2, x \mapsto x$. Let $\alpha = \{e_1, e_2\}$ be a basis of V.
We computed $[f]_\alpha^\alpha = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.
The matrices of (ex 2.2) is $\begin{bmatrix} 1 & 2 \\ 1 & -2 \end{bmatrix}$

2.7. Remark:
If $T : V \to V$ be the identity transformation of a finite-dimensional vector space to itself, $T = l$. When with respect to any choice of basis α for V, the matrix l is the $k \times k$ matrix with 1 in each diagonal position and 0 in each off-diagonal position.

2.8. Example: The matrix of a rotation. [See example (1.13)]
In this example $V = W = \mathbb{R}^2$. and we take both bases α and β to be the standard basis: $e_1(1,0), e_2(0,1)$. Let $T = R_\theta$ be rotation through an angle θ in the plane.
Then for an arbitrary vector $V = (v_1, v_2)$

$$R_\theta(v) = \begin{pmatrix} v_1 \cos(\theta) - v_2 \sin(\theta) \\ v_1 \sin(\theta) + v_2 \cos(\theta) \end{pmatrix}$$

Therefore, the matrix of R_θ is

$$[R_\theta]_\alpha^\alpha = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

2.9. Definition: Let A be an $l \times k$ matrix, and let X be a column vector with k entries, then the **product of the vector x by the matrix A** is defined to be the column vector with l entries:

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + ... + a_{1k}x_k \\ . \\ . \\ . \\ a_{l1}x_1 + a_{l2}x_2 + ... + a_{lk}x_k \end{bmatrix}$$

and is denoted by Ax. If we write out the entire matrix A and the vector x, this becomes

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1k} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2k} \\ & & . & & \\ & & . & & \\ & & . & & \\ a_{l1} & a_{l2} & a_{l3} & \dots & a_{lk} \end{bmatrix} \begin{bmatrix} x_1 \\ . \\ . \\ . \\ x_k \end{bmatrix}$$

2.10. Remark: The ith entry of the product $Ax, a_{i1}x_1 + ... + a_{ik}x_k$, can be thought of as the product of the ith row A, considered as a $1 \times k$ matrix, with the column vector x, using this same definition.

2.11. Remark: The project of a $1 \times k$ matrix, which we can think of as a row vector, and a column vector generalizes the notion of the dot product in the plane.
If x and $y \in \mathbb{R}^2, < x, y >= x_1y_1 + x_2y_2$.
If we write x as a matrix $[x_1, x_2]$, then $\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} =< x, y >$

2.12. Remark: If the number of columns of the matrix A is not equal to the number of entries in the column vector X, matrix multiplication Ax is not defined.

2.13. Proposition: *Let $T : V \to W$ be a linear transformation between vector spaces V of dimension k and W of dimension l. Let $\alpha = \{v_1, ..., v_k\}$ be a basis for V and $\beta = \{w_1, ..., w_l\}$ be a basis for W. Then for each $v \in V$*

$$[T(v)]_\beta = [T]^\beta_\alpha [v]_\alpha$$

Proof. Let $v = x_1v_1 + ... + x_kv_k \in V$. Then if $T(v_j) = a_{1j}w_1 + ... + a_{lj}w_l$

$$\begin{aligned} T(v) &= \sum_{j=1}^{k} x_j T(v_j) \\ &= \sum_{j=1}^{k} x_j (\sum_{i=1}^{l} a_{ij} w_i) \\ &= \sum_{i=1}^{l} (\sum_{j=1}^{k} x_j a_{ij}) w_i \end{aligned}$$

Thus, The ith coefficient of $T(v)$ in terms of β is $\displaystyle\sum_{j=1}^{k} x_j a_{ij}$

$$T(v)_\beta = \begin{bmatrix} \sum_{j=1}^{k} x_j a_{1j} \\ . \\ . \\ . \\ \sum_{j=1}^{k} x_j a_{lj} \end{bmatrix}$$

which is precisely $[T]^\beta_\alpha [v]_\alpha$ $\square$

2.14. Remark: If v_j is the jth member of the basis α of V

$$f(v_j) = a_1 j w_1 + ... + a_l j w_l$$

Thus,

$$[T(v_j)]_\alpha^\beta = [T]_\alpha^\beta [v_j]_\alpha = \begin{bmatrix} a_{11} & ... & a_{1j} & ... & a_{1k} \\ & & ... & & \\ a_{l1} & ... & a_{lj} & ... & a_{lk} \end{bmatrix} \begin{bmatrix} 0 \\ . \\ 0 \\ 1 \\ 0 \\ . \\ 0 \end{bmatrix} \text{(jth row)} = \begin{bmatrix} a_{1j} \\ . \\ . \\ . \\ a_{lj} \end{bmatrix}$$

which is te jth column of the matrix $[T]_\alpha^\beta$.

2.15. Proposition: *Let A be an $l \times k$ matrix and u and v be column vectors with k entries. Then for every pair of real numbers a and b*

$$A(au + bv) = aAu + bAv$$

2.16. Proposition: *Let $\alpha = \{v_1, ..., v_k\}$ be a basis for V and $\beta = \{w_1, ..., w_l\}$ be a basis for W, and let $v = x_1 v_1 + ... + x_k v_k \in V$.*

(i) If A is an $l \times k$ matrix, then the function

$$T(v) = w$$

where $[w]_\beta = A[v]_\alpha$ is a linear transformation.

(ii) If $A = [S]_\alpha^\beta$ is the matrix of a transformation $S : V \to W$, then the transformation T constructed from $[S]_\alpha^\beta$ is equal to S.

(iii) If T is the transformation of (i) constructed from A, then

$$[T]_\alpha^\beta = A$$

2.17. Proposition: *Let V and W be finite-dimensional vector spaces. Let α be a basis for V and β a basis for W. Then the assignment of a matrix to a linear transformation form V to W given by T goes to $[T]_\alpha^\beta$ is* ***injective and surjective***.

Section 3. Kernel and Image

3.1. Definition: The **kernel** of T, denoted $\mathrm{Ker}(T)$, is the subset of V consisting of all vectors $v \in V$ such that $T(v) = 0$

3.2. Proposition: *Let $T : V \to W$ be a linear transformation. $\ker(T)$ is a subspace of V.*

Proof. Since T is linear, for all u and $v \in \ker(T)$ and $\alpha \in \mathbb{R}$,
$T(u + av) = T(u) + aT(v) = 0 + a0 = 0$.
Therefore, $u + av \in \mathrm{Ker}(T)$. □

3.3. Example: Let $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$,

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Therefore, $\mathrm{Ker}(R_\theta) = \{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\}$

3.4. Example: Let $f : P_2 \to P_2$,

$$p(x) \mapsto xp'(x)$$

Therefore, $\mathrm{Ker}(f) = \{p(x) \in P_2 : xp'(x) = 0\} = \{p(x) \in P_2 : p'(x) = 0\} = \{\text{constant polynomial}\}$.

3.5. Proposition: *For all $x \in V$, we have $x \in \mathrm{Ker}(f) \Leftrightarrow [f]_\alpha^\beta [x]_\alpha = 0$*

3.6. Proposition: *Let $T : V \to W$ be a linear transformation of finite-dimensional vector spaces, and let α and β be bases for V and W, respectively.*
Then $x \in \mathrm{Ker}(T)$ if and only if the coordinate vector of x, $[x]_\alpha$, satisfies the system of equations

$$a_{11}x_1 + \ldots + a_{1k}x_k = 0$$
$$\ldots$$
$$a_{l1}x_1 + \ldots + a_{lk}x_k = 0$$

where the coefficients a_{ij} are the entries of the matrix $[T]_\alpha^\beta$.

3.7. Proposition: *Let V be a finite-dimensional vector, space, and let $\alpha = \{v_1, v_2, \ldots, v_k\}$ be a basis for V. Then the vectors $x_1, \ldots, x_m \in V$ are linearly independent if and only if their corresponding coordinate vectors $[x_1]\alpha, \ldots, [x_m]\alpha$ are linearly independent.*

Proof. Assume $x_1, \ldots, x_m$ are linearly independent and

$$x_i = a_{1i}v_1 + \ldots + a_{ki}v_k$$

If $b_1, ..., b_m$ is any m-tuple of scalars with

$$b_1[x_1]\alpha + ... + b_m[x_m]\alpha = b_1 \begin{bmatrix} a_{11} \\ . \\ . \\ . \\ a_{k1} \end{bmatrix} + ... + b_m \begin{bmatrix} a_{1m} \\ . \\ . \\ . \\ a_{km} \end{bmatrix} = \begin{bmatrix} 0 \\ . \\ . \\ . \\ 0 \end{bmatrix}$$

then equating each of the components to zero, we have $\sum_{i=1}^{m} b_i a_{ji} = 0$, for all $j, 1 \leq j \leq k$. Thus

$$(\sum_{i=1}^{m} b_i a_{1i})v_1 + ... + (\sum_{i=1}^{m} b_i a_{ki})v_k = 0$$

Rearranging the yields

$$b_1(\sum_{j=1}^{k} a_{j1}v_j) + ... + b_m(\sum_{j=1}^{k} a_{jm}v_j) = 0$$

or

$$b_1x_1 + ... + b_mx_m = 0$$

Therefore, $b_1 = b_2 = ... = b_m = 0$ and the m coordinate vectors are also linearly independent. $\square$

3.8. Example:
Let V be a vector space of dimension 4, and let W be a vector space of dimension 3.
Let $\alpha = \{v_1, ..., v_4\}$ be a basis for V, $\beta = \{w_1, ..., w_3\}$ be a basis for W.
Let T be the linear transformation such that

$$[T]_\alpha^\beta = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & -1 & 3 & 5 \end{bmatrix}$$

Let us find the dimension of Ker(T).
We must solve the system

$$\begin{aligned} x_1 + 0 \cdot x_2 + x_3 + 2x_4 &= 0 \\ 2x_1 + x_2 + 0 \cdot x_3 + x_4 &= 0 \\ x_1 - x_2 + 3x_3 + 5x_4 &= 0 \end{aligned}$$

The free variables are x_3 and x_4. Yields two solutions $(-1, 2, 1, 0)$ and $(-2, 3, 0, 1)$.
Therefore, $\dim(\text{Ker}(T)) = 2$ and a basis fore Ker(T) is $\{-v_1 + 2v_2 + v_3, -2v_1 + 3v_2 + v_4\}$.

3.9. Definition: The subset of W consisting of all vectors $w \in W$ for which there exists a $v \in V$ such that $T(v) = w$ is called the **image** of T and is denoted by Im(T).

3.10. Proposition: *Let $T : V \to W$ be a linear transformation. The image of T is a subspace of W.*

Proof. Let w_1 and $w_2 \in \text{Im}(T)$, and let $a \in \mathbb{R}$.
Since $w_1, w_2 \in \text{Im}(T)$, there exist vectors v_1 and $v_2 \in V$ with $T(v_1) = w_1$ and $T(v_2) = w_2$.
Then we have $aw_1 + w_2 = aT(v_1) + T(v_2)$, since T is linear.
Therefore, $aw_1 + w_2 \in \text{Im}(T)$ and $\text{Im}(T)$ is a subspace of W. □

3.11. Example: Let $f : P_2 \to P_2$,

$$p(x) \mapsto xp'(x)$$

$$f(a_0 + a_1 x + a_2 x^2) = x \cdot (a_1 + 2a_2 x) = a_1 x + 2a_2 x^2$$

Therefore,

$$\text{Im}(f) = \{a_1 x + 2a_2 x^2 : a_1, a_2 \in \mathbb{R}\} = \{b_1 x + b_2 x^2 : b_1, b_2 \in \mathbb{R}\} = \{P_2 \text{ whose constant term is } 0\}$$

3.12. Proposition: *If $\{v_1, ..., v_m\}$ is any set that spans V, then $\{T(v_1), ..., T(v_m)\}$ spans* $\text{Im}(T)$.

3.13. Remark: $[f(v_k)]_\beta = k$th column of $[f]_\alpha^\beta$

3.14. Corollary: *If $\alpha = \{v_1, ..., v_k\}$ is a basis for V and $\beta = \{w_1, ..., w_l\}$ is a basis for W. Then the vectors in W whose coordinate vectors (in terms of β) are the columns of $[T]_\alpha^\beta$ span* $\text{Im}(T)$.

3.15. Remark:
Q: How to find a basis for $\text{Im}(f)$?
A: Choose any maximal linearly independent subset of $\{f(v_1), ..., f(v_k)\}$.

3.16. Example: Let $f : P_2 \to P_2$,

$$p(x) \mapsto xp'(x)$$

Let $\alpha = \beta = \{1, x, x^2\} \implies f(1) = 0, f(x) = x, f(x^2) = x^2$.
$\{x, 2x^2\}$ is a maximal linearly independent subset of $\{f(1), f(x), f(x^2)\}$.
Hence $\{x, 2x^2\}$ is a basis for $\text{Im}(f)$

3.17. Theorem (Rank-Nullity): If V is a finite-dimensional vector space and $T : V \to W$ is a linear transformation, then

$$\dim(\text{Ker}(T)) + \dim(\text{Im}(T)) = \dim(V)$$

3.18. Remark: Another answer of (3.15):
Suppose $\{v_1, ..., v_k\}$ is a basis of V such that $\{v_1, .., v_r\}(r \leq k)$ is a basis for $\text{Ker}(f)$.
Then $\{f(v_{r+1}), ..., f(v_k)\}$ is a basis for $\text{Im}(f)$.

Section 4. Applications of the Dimension Theorem

4.1. Proposition: *A linear transformation $T : V \to W$ is* ***injective*** *if and only if*

$$\dim(\mathrm{Ker}(T)) = 0.$$

Proof. **Only if**: Assume that f is injective.
Since we know that $T(0) = 0$ for all linear mappings.
Thus, $\forall x \in V$, if $f(x) = 0 \implies x = 0$.
Thus, $\mathrm{Ker} = \{0\} \implies \dim(\mathrm{Ker}(T)) = 0$.
If: Assume $\mathrm{Ker}(f) = 0$.
If $f(x_1) = f(x_2)$, then $f(x_1 - x_2) = 0$.
i.e $x_1 - x_2 \in \mathrm{Ker}(f) = \{0\}$ i.e $x_1 - x_2 = 0$ i,e $x_1 = x_2$.
Therefore, f is injective. □

4.2. Corollary: *A linear mapping $T : V \to W$ on a finite-dimensional vector space V is* ***injective*** *if and only if*

$$\dim(\mathrm{Im}(T)) = \dim(V).$$

4.3. Corollary: *If V and W are finite dimensional, then a linear mapping $T : V \to W$ can* ***be injective*** *only if*

$$\dim(W) \geq \dim(V)$$

Expanation. $\mathrm{Im}(f)$ is a vector subspace of $W \implies \dim(\mathrm{Im}(f)) \leq \dim(W)$.
If $\dim(V) > \dim(W)$, then

$$\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = \dim(V) > \dim(W) \geq \dim(\mathrm{Im}(f))$$

Therefore, $\dim(\mathrm{Ker}(f)) > 0 \implies f$ is not injective. □

4.4. Proposition: *If W is finite-dimensional, then a linear mapping $T : V \to W$ is* ***surjective*** *if and only if*

$$\dim(\mathrm{Im}(T)) = \dim(W)$$

4.5. Corollary: *A linear mapping $T : V \to W$ can be* ***surjective*** *only if* $\dim(V) \geq \dim(W)$.

Explanation. Since $\dim(\mathrm{Im}(T)) \leq \dim(V)$.
If $\dim(V) < \dim(W)$, then $\dim(\mathrm{Im}(T)) < \dim(W)$, and hence, T is not surjective. □

4.6. Proposition: *Let* $\dim(V) = \dim(W)$. *A linear transformation* $T : V \to W$ *is injective if and only if it is surjective.*

Proof.
f is injective $\Leftrightarrow \dim(\mathrm{Ker}(f)) = 0$.
f is surjective $\Leftrightarrow \dim(\mathrm{Im}(f)) = \dim(W) = \dim(V)$.
By rank-nullity theorem, $\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = \dim(V)$.
$\implies \dim(\mathrm{Ker}(f)) = 0 \Leftrightarrow \dim(\mathrm{Im}(f)) = \dim(V)$. $\square$

4.7. Proposition: *Let* $T : V \to W$ *be a linear transformation, and let* $w \in \mathrm{Im}(T)$.
Let v_1 *be any fixed vector with* $T(v_1) = w$.
Then every vector $v_2 \in T^{-1}(\{w\})$ *can be written uniquely as* $v_2 = v_1 + u$, *where* $u \in \mathrm{Ker}(T)$.

Proof. If $T(v_2) = w$, we let $u = v_2 - v_1$.
Then, $T(u) = T(v_1 - v_2) = 0$.
We claim that this choice of u is unique.
Suppose that u' is another vector in $\mathrm{Ker}(T)$ with $v_2 = v_1 + u'$.
Then we have $v_1 + u = v_1 + u'$ which implies that $u = u'$. $\square$

4.8. Remark: If a different v_1 were used, the corresponding u' would change too.

4.9. Proposition: *Let* $T : V \to W$ *be a linear transformation.*
Let $y \in \mathrm{Im}(T)$ *and* $x \in T^{-1}(y)$.
Then $x' \in T^{-1}(y)$ *if and only if* $x - x' \in \mathrm{Ker}(f)$.

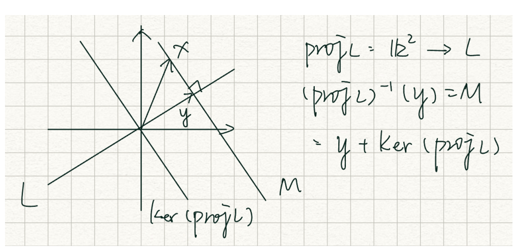
Explanation.
$x' \in T^{-1}(y) \implies f(x') = y \implies f(x') = f(x) \implies f(x' - x) = 0 \implies x' - x \in \mathrm{Ker}(T)$ $\square$

4.10. Corollary: *Let* $T : V \to W$ *be a linear transformation of finite-dimensional vector spaces. Let* $w \in W$. *Then there is a unique vector* $v \in V$ *such that* $T(v) = w$ *if and only if*

(i) $w \in \mathrm{Im}(T)$ *and*

(ii) $\dim(\mathrm{Ker}(T)) = 0$

4.11. Example:



4.12. Example: Let $f : \mathbb{R}^3 \to \mathbb{R}$ be linear transformation such that

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x + 2y$$

By observation, $\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \in f^{-1}(2)$ and $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in f^{-1}(2)$.

$\mathrm{Ker}(f) = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x + 2y = 0\}$. Therefore, $\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \in \mathrm{Ker}(f)$.

4.13. Proposition:

(i) The set of solutions of the system of linear equations $Ax = b$ is the subset $T^{-1}(\{b\})$ of $V = \mathbb{R}^n$.

(ii) The set of solutions of the system of linear equations $Ax = b$ is a subspace of V if and only of the system is homogeneous, in which case the set of solutions is $\mathrm{Ker}(T)$.

4.14. Corollary:

(i) The number of free variables in the homogeneous system $Ax = 0$ is equal to $\dim(\mathrm{Ker}(T))$.

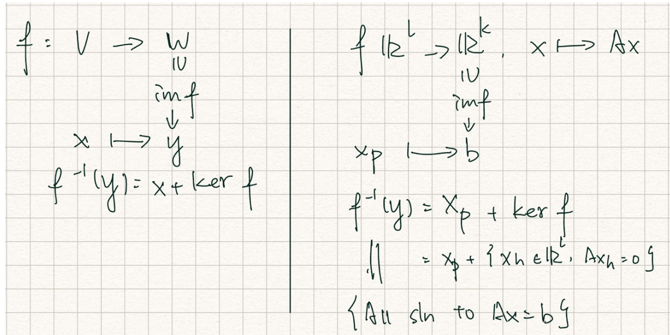
(ii) The number of basic variables of the system is equal to $\dim(\mathrm{Im}(T))$

4.15. Remark: Take a $k \times l$ matrix A.
Get a linear transformation $f : \mathbb{R}^l \to \mathbb{R}^k, x \mapsto Ax$.
Take $b \in \mathrm{Im}(f) \subseteq \mathbb{R}^k$.
Then $f^{-1}(b) = \{x \in \mathbb{R}^k : f(x) = b\} = \{x \in \mathbb{R}^k : Ax = b\}$

4.16. Definition: Given an inhomogeneous system of equations. $Ax = b$, any single vector x satisfying the system (necessarily $x \neq 0$) us caked a **particular solution** of the system of equations.
In other words, a particular solution to $Ax = b$ is a vector $x_p \in \mathbb{R}^l$ s.t $Ax_p = b$

4.17. Proposition: *Let x_p be a particular solution of the system $Ax = b$.
Then every solution to $Ax = b$ is of the form $x = x_p + x_h$, where $x_h \in \mathbb{R}^l$ satisfies $A(x_h) = 0$.
Furthermore, given x_p and x, there is a unique x_h such that $x = x_p + x_h$.*

Explanation.



4.18. Corollary: *The system $Ax = b$ has a unique solution if and only if $b \in \mathrm{Im}(T)$ and the* $\mathrm{Ker}(f) = 0$.

4.19. Note: Let S and $T : U \to V$ be linear transformations.
(i) If S and T are injective (surjective), then $S + T$ is **not** necessarily injective (surjective)
(ii) If S is injective (surjective) and $a \neq 0$, then aS must be injective (surjective).

Section 5. Composition of Linear Transformation

5.1. Proposition: *Let $S : U \to V$ and $T : V \to W$ be linear transformations, then TS is a linear transformation.*

Proof. Let a and $b \in \mathbb{R}$ and et u_1 and $u_2 \in U$.

$$\begin{aligned} TS(au_1 + bu_2) &= T(S(au_1 + bu_2)), \text{ by the definition of } TS \\ &= T(aS(u_1) + bS(u_2)), \text{ by the linearity of } S \\ &= aT(S(u_1)) + bT(S(u_2)), \text{ be the linearity of } T \\ &= aTS(u_1) + bTS(u_2) \end{aligned}$$

□

5.2. Proposition:

(i) ***Associativity**: Let $R : U \to V$, $S : V \to W$ and $T : W \to X$ be linear transformation of the vector spaces U, V, W and X as indicated. Then*

$$T(SR) = (TS)R$$

(ii) ***Distributivity**: Let $R : U \to V$, $S : U \to V$ and $T : V \to W$ be linear transformation of the vector spaces U, V and W as indicated. Then*

$$T(R + S) = TR + TS$$

(iii) ***Distributivity****: Let $R : U \to V$, $S : V \to W$ and $T : V \to W$ be linear transformation of the vector spaces U, V and W as indicated. Then*

$$(T + S)R = TR + SR$$

5.3. Proposition: *Let $S : U \to V$ and $T : V \to W$ be linear transformations. Then*

(i) $\mathrm{Ker}(S) \subseteq \mathrm{Ker}(TS)$

(ii) $\mathrm{Im}(TS) \subseteq \mathrm{Im}(T)$

5.4. Remark: These could be strict inclusions.
Let $T : \mathbb{R}^2 \to \mathbb{R}$ and $S : \mathbb{R} \to \{0\}$ be linear transformations, such that

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto x \mapsto 0$$

$\mathrm{Ker}(TS) = \mathbb{R}^2 \supset \mathrm{Ker}(T) = \{\begin{pmatrix} 0 \\ y \end{pmatrix} : y \in \mathbb{R}\}$

5.5. Corollary: *Let $S : U \to V$ and $T : V \to W$ be linear transformations of finite-dimensional vector spaces. Then*

(i) $\dim(\mathrm{Ker}(S)) \le \dim(\mathrm{Ker}(TS))$

(ii) $\dim(\mathrm{Im}(TS)) \le \dim(\mathrm{Im}(T))$

5.6. Proposition: *If $[S]_\alpha^\beta$ has entries a_{ij}, $i = 1, ..., n$ and $j = 1, ..., m$ and $[T]_\beta^\gamma$ has entries b_{kl}, $k = 1, ..., p$ and $l = 1, ..., n$, then the entries of $[TS]_\alpha^\gamma$ are $\sum_{l=1}^n b_{kl}a_{lj}$*

5.7. Definition: Let A be an $n \times m$ matrix and B a $p \times n$ matrix, then the matrix product BA is defines to be the $p \times m$ matrix whose entries are $\sum_{l=1}^n b_{kl}a_{lj}$ for $k = 1, ..., p$ and $j = 1, ..., m$.

5.8. Proposition: *Let $S : U \to V$ and $T : V \to W$ be linear transformations between finite-dimensional vector spaces. Let α, β and γ be bases for U, V, and W, respectively. Then*

$$[TS]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta$$

Explanation.
rth column of $[TS]_\alpha^\gamma$

$$\begin{aligned} TS(u_r) = T(S(u_r)) &= T(b_{1r}v_1 + ... + b_{nr}v_n) \\ &= b_{1r}T(v_1) + ... + b_{nr}T(v_n) \\ &= b_{1r}(a_{11}w_1 + ... + a_{p1}w_p) + ... + b_{nr}(a_{1n}w_1 + ... + a_{pn}) \\ &= (b_{1r}a_{11} + b_{2r}a_{12} ... + b_{nr}a_{1n})w_1 + ... + (b_{1r}a_{p1} + b_{2r}a_{p2} ... + b_{nr}a_{pn})w_p \end{aligned}$$

Hence, rth column of $[TS]_\alpha^\gamma = \begin{bmatrix} b_{1r}a_{11} + b_{2r}a_{12} ... + b_{nr}a_{1n} \\ ... \\ b_{1r}a_{p1} + b_{2r}a_{p2} ... + b_{nr}a_{pn} \end{bmatrix}$

5.9. Proposition:

(i) ***Associativity:*** *Let $A, B,$ and C be $m \times n$, $n \times p$ and $p \times r$ matrices. Then*

$$(AB)C = A(BC)$$

(ii) ***Distributivity:*** *Let A be an $m \times n$ matrix and B and C $n \times p$ matrices. Then*

$$A(B + C) = AB + AC$$

(iii) ***Distributivity:*** *Let A and B be an $m \times n$ matrix and C $n \times p$ matrices. Then*

$$(A + B)C = AC + BC$$

5.10. Remark: Let A be a $m \times n$ matrix and B be an $n \times p$.
But in general, BA does note make sense.
BA only makes sense when $m = p$. But even in this case, $BA \neq AB$.

5.11. Example: Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} = BA$$

Section 6. The Inverse of a Linear Transformation

6.1. Definition: Let $T : V \to W$ is a linear transformation that has an inverse transformation $S : W \to V$, we say that T is **invertible**, and we denote the inverse of T by T^{-1}.

6.2. Proposition: *If $T : V \to W$ is injective and surjective, then the inverse function $S : W \to V$ is a linear transformation.*

6.3. Proposition: *A linear transformation $T : V \to W$ has an inverse linear transformation S if and only if T is injective and surjective.*

6.4. Definition: If $T : V \to W$ is an invertible linear transformation, T is called an **isomorphism**, and we say V and W are **isomorphic vector spaces**.

6.5. Example:

(a) $Id_v : V \to V$

(b) $R_\theta : \mathbb{R}^2 \to R^2$

(c) $f : Mat_{2\times 2} \to \mathbb{R}^4$, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}$

(d) $f : P_2 \to \mathbb{R}^3$, $a_0 + a_1x + a_2x^2 \mapsto \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$

(e) **Nonex:** $f : P_2 \to P_1$, $p \mapsto p'$ (since f is not surjective)

6.6. Proposition: *Let V and W be finite-dimensional vector spaces, then there is an isomorphism $T : V \to W$ if and only if* $\dim(V) = \dim(W)$.

6.7. Remark (Setting):
V, W are finite dimensional vector spaces.
$\alpha : \{v_1, ..., v_k\}$ basis for V
$\beta : \{w_1, ..., w_l\}$ basis for W.
$f : V \to W$ invertible linear transformation

6.8. Remark: $f^{-1} : W \to V$, $f : V \to W$.
$[f \circ f^{-1}]^\beta_\beta = [f]^\beta_\alpha [f^{-1}]^\alpha_\beta$.
Observation: $[f \circ f^{-1}]^\beta_\beta = [I_w]^\beta_\beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
Hence $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [f]^\beta_\alpha [f^{-1}]^\alpha_\beta$.
Similarly, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = [f^{-1}]^\alpha_\beta [f]^\beta_\alpha$.

6.9. Definition: An $n \times n$ matrix A is called **invertible** if there exists an $n \times n$ matrix B so that $AB = BA = I$. B is called the **inverse** of A and is denoted by A^{-1}

6.10. Remark: The inverse matrix may not exists. ex: $A = (0)$.
The inverse matrix, if exists, is **unique**.
Suppose B and C are inverse of A.
Then $B = BI = B(AC) = (BA)C = IC = C$

6.11. Proposition: *Let $T : V \to W$ be an isomorphism of finite-dimensional vector spaces. Then for any choice of bases α for V and β for W*

$$[T^{-1}]_{\beta}^{\alpha} = ([T]_{\alpha}^{\beta})^{-1}$$

6.12. Remark:
Q: How to compute S^{-1}?
Consider $Ax = b$
Suppose A has an inverse. Then $A^{-1}(Ax) = A^{-1}b$.
i.e $(AA^{-1})x = A^{-1}b$ i.e $Ix = A^{-1}b \implies x = A^{-1}b$

Section 7. Change of Basis

7.1. Remark: $I : V \to V$ is an identity transformation from V to itself.

7.2. Proposition: *Let V be a finite-dimensional vector space, and let α and α' be bases for V. Let $v \in V$. Then the coordinate vector $[v]_{\alpha'}$ if v in the basis α' is related to the coordinate vector $[v]_{\alpha}$ of v in the basis α by*

$$[I]_{\alpha}^{\alpha'}[v]_{\alpha} = [v]_{\alpha'}$$

7.3. Definition: Let V be a finite-dimensional vector space, and let α and α' be bases for V. The matrix $[I_V]_{\alpha}^{\alpha'}$ is called the **change of basis matrix** from α to α'.

7.4. Remark: If $\dim(V) = n$, then $[I_V]_{\alpha}^{\alpha'}$ is of size $n \times n$.

7.5. Remark: I is an invertible linear mapping, and $I^{-1} = I$, so that $[I^{-1}]_{\alpha}^{\alpha'} = [I]_{\alpha}^{\alpha'}$. Hence, $([I]_{\alpha'}^{\alpha})^{-1} = [I]_{\alpha}^{\alpha'}$

7.6. Theorem: Let $T : V \to W$ be a linear transformation between finite-dimensional vector spaces V and W. Let $I_V : V \to V$ and $I_W : W \to W$ be the respective identity transformation of V and W. Let α and α' be two bases for V, and let β and β' be two bases for W. Then

$$[T]_{\alpha'}^{\beta'} = [I_W]_{\beta}^{\beta'} \cdot [T]_{\alpha}^{\beta} \cdot [I_V]_{\alpha'}^{\alpha}$$

7.7. Remark (Specialize): $W = V, \beta = \alpha, \beta' = \alpha'$.

$$[T]_{\alpha'}^{\alpha'} = ([I_V]_{\alpha'}^{\alpha})^{-1} \cdot [T]_{\alpha}^{\alpha} \cdot [I_V]_{\alpha'}^{\alpha}$$

7.8. Definition: Let A, B be $n \times n$ matrices, A and B are said to be **similar** if there is an invertible $n \times n$ matrix Q such that

$$B = Q^{-1}AQ$$

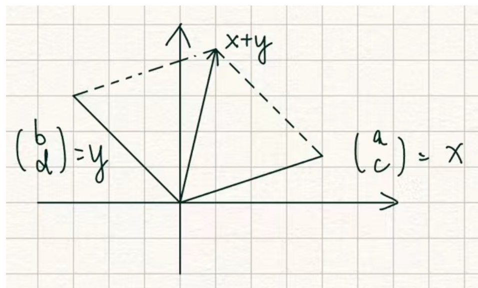
Chapter 3

The Determinant Function

Section 1. The Determinant as Area

1.1. Remark: Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation. Let $\alpha = \{v_1, v_2\}$ be a basis for $\mathbb{R}^2$.

$$[f]_\alpha^\alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$



Q: What is the area of this parallelogram?
A: $Area = \|x\|\|y\| \sin\theta$.
Recall: $< x, y >= \|x\|\|y\| \cos\theta$, $\cos^2\theta + \sin^2\theta = 1$.

$$\begin{aligned} Area &= \|x\|\|y\|\sqrt{1-\cos^2\theta} \text{positive, since } \theta \in [0, \pi] \\ &= \sqrt{\|x\|^2\|y\|^2 - \|x\|^2\|y\|^2\cos^2\theta} \\ &= \sqrt{\|x\|^2\|y\|^2 - (< x, y >)^2} \\ &= \sqrt{(a^2+c^2)(b^2+d^2) - (ab+cd)^2} \\ &= \sqrt{(ad-bc)^2} \\ &= \pm(ad-bc) \end{aligned}$$

1.2. Proposition: *The area of the parallelogram generated by* $\begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} b \\ d \end{pmatrix}$ *and* 0 *is* $\pm(ad-bc)$. *The area is zero if and only the* $\begin{pmatrix} a \\ c \end{pmatrix}$ *and* $\begin{pmatrix} b \\ d \end{pmatrix}$ *are linear dependent if and only if* $ad - bc = 0$.

1.3. Corollary: *Let* $V = \mathbb{R}^2$. $T : V \to V$ *is an isomorphism if and only the area of the parallelogram constructed previously is nonzero if and only if* $ad - bc \neq 0$.

1.4. Proposition: *The function* $Area(a_1, a_2)$ *has the following properties for* $a_1, a_2, a_1', a_2' \in \mathbb{R}^2$.

(i) $Area(ba_1 + ca_1', a_2) = b\ Area(a_1, a_2) + c\ Area(a_1', a_2)$ *for* $b, c \in \mathbb{R}$.

(ii) $Area(a_1, a_2) = \text{-}Area(a_2, a_1)$

(iii) $Area((1,0),(0,1)) = 1$

1.5. Remark:
If we fix the second argument in Area then we get a linear transformation $\mathbb{R}^2 \to \mathbb{R}$.
Similarly for the other argument.
A function $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ is linear in both arguments is called **multilinear**.
A function $f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ such that $f(x, y) = -f(y, x)$ for all $a, y \in \mathbb{R}^2$ is called **alternating**.

1.6. Proposition: *Let $B : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ be a multilinear and alternating function such that $B((1, 0), (0, 1)) = 1$. Then B is equal to the area function.*

1.7. Definition: The **determinant** of a 2×2 matrix A, denoted by $\det(A)$ or $\det(a_1, a_2)$, is the unique function of the rows of A satisfying

(i) $\det(ba_1 + ca_1', a_2) = b\det(a_1, a_2) + c\det(a_1', a_2)$ for $b, c \in \mathbb{R}$.

(ii) $\det(a_1, a_2) =$ -$\det(a_2, a_1)$

(iii) $\det(e_1, e_2) = 1$

When $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $\det(A) = ad - bc$

1.8. Proposition:

(i) A 2×2 matrix A is invertible if and only if $\det(A) \neq 0$.

(ii) If $T : V \to V$ is a linear transformation of a two-dimensional vector space V, then T is an isomorphism if and only if $\det[T]_\alpha^\alpha \neq 0$.

Section 2. The Determinant of An $n \times n$ Matrix

2.1. Definition: A function $f : \mathbb{R}^n \times \mathbb{R}^n \times ... \times \mathbb{R}^n \to \mathbb{R}$ of the rows of a matrix A is called **multilinear** if f is a linear function of each of its rows when the remaining rows are held fixed. That is, f is multilinear if for all b and $b' \in \mathbb{R}$

$$f(a_1, ..., ba_i + b'a_i', ..., a_n) = bf(a_1, ..., a_i, ..., a_n) + b'f(a_1, ..., a_i', ..., a_n).$$

2.2. Definition: A function f of the rows of a matrix A is said to be **alternating** if whenever any two rows of A are interchanged f changes sign, That is, for all $i \neq j, 1 \leq i, j \leq n$, we have

$$f(a_1, ..., a_i, .., a_j, ..., a_n) = -f(a_1, ..., a_j, ..., a_i, ..., a_n).$$

2.3. Lemma: *If f is an alternating real-valued function of the rows of an $n \times n$ matrix and two rows of the matrix A are identical, then* $f(A) = 0$

Proof. Assume $a_i = a_j$.
Then $f(A) = f(a_1, ..., a_i, .., a_j, ..., a_n) = -f(a_1, ..., a_j, ..., a_i, ..., a_n) = -f(A)$.
Therefore, $f(A) = 0$. $\square$

2.4. Example: $n = 2$. Take $f = \det$.

$$\det(\begin{bmatrix} a & a \\ c & c \end{bmatrix}) = 0$$

2.5. Definition: Let A be an $n \times n$ matrix with entries $a_{ij}, i, j = 1, ..., n$. The ijth **minor** of A is defined to be the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth column of A. The ijth minor is denoted by A_{ij}. Thus

$$A_{ij} = \begin{bmatrix} a_{11} & \dots & a_{1.j-1} & a_{1.j+1} & \dots & a_{1n} \\ & \vdots & & & \vdots & \\ a_{i-1.1} & \dots & a_{i-1.j-1} & a_{i-1.j+1} & \dots & a_{i-1.n} \\ a_{i+1.1} & \dots & a_{i+1.j-1} & a_{i+1.j+1} & \dots & a_{i+1.n} \\ & \vdots & & & \vdots & \\ a_{n1} & \dots & a_{n.j-1} & a_{n,j+1} & \dots & a_{nn} \end{bmatrix}$$

2.6. Example: $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$.

$$A_{11} = \begin{bmatrix} e & f \\ h & i \end{bmatrix}, A_{12} = \begin{bmatrix} d & f \\ g & i \end{bmatrix}, A_{13} = \begin{bmatrix} d & e \\ g & h \end{bmatrix}$$

2.7. Proposition: *Let A be a 3×3 matrix, and let f be an alternating multilinear function. Then*

$$f(A) = [a \det(A_{11}) - b \det(A_{12}) + c \det(A_{13})] f(I)$$

Proof. Expanding the first row of A in terms of the standard basis in $\mathbb{R}^3$ and using the multilinearity of f, we see that

$$f(A) = af(e_1, a_2, a_3) + bf(e_2, a_2, a_3) + cf(e_3, a_2, a_3)$$

Expanding a_2 in the same way, we obtain

$$f(e_1, a_2, a_3) = df(e_1, e_1, a_3) + ef(e_1, e_2, a_3) + ff(e_1.e_3.a_3) = eef(e_1, e_2, a_3) + ff(e_1.e_3.a_3)$$

applying Lemma2.3.Finally, expanding the third row yields

$$eif(e_1, e_2, e_3) + fhf(e_1.e_3.e_2)$$

The other terms are zero by Lemma2.3. Since f is alternating, we have $f(e_1, e_3, e_2) = -f(e_1, e_2, e_3)$, so the preceding expression equals to

$$\det(A_{11}) f(I).$$

Other two are similarly. □

2.8. Corollary: *There exists exactly one multilinear alternating function f of the rows of a 3×3 matrix such that $f(I) = 1$.*

2.9. Definition: The determinant function of a 3×3 matrix is the unique alternating multilinear function f with $f(I) = 1$. This function will denoted by $\det(A)$.

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

for $i = 1, 2, 3$.

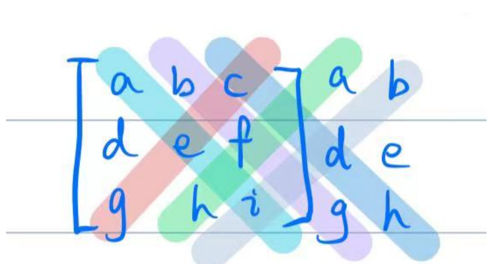
2.10. Example: $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

$$a \det(A_{11}) - b \det(A_{12}) = a \det[d] - b \det[c] = ad - bc = \det(A)$$

$$-c \det(A_{21}) + d \det(A_{22}) = -c \det[b] + d \det[a] = -cb + ad = \det(A)$$

2.11. Remark: $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

$$\det A = aei + bfg + cdh - ceg - bdi - afh$$



2.12. Remark:

$$f\left(\begin{pmatrix} a \\ d \\ g \end{pmatrix} + \begin{pmatrix} a' \\ d' \\ g' \end{pmatrix}, \begin{pmatrix} b \\ e \\ h \end{pmatrix}, \begin{pmatrix} c \\ f \\ i \end{pmatrix}\right)$$
$$= (a + a') \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \det\left(\begin{pmatrix} d \\ g \end{pmatrix} + \begin{pmatrix} d' \\ g' \end{pmatrix}, \begin{pmatrix} f \\ i \end{pmatrix}\right) + c \det\left(\begin{pmatrix} d \\ g \end{pmatrix} + \begin{pmatrix} d' \\ g' \end{pmatrix}, \begin{pmatrix} e \\ h \end{pmatrix}\right))$$
$$= a \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} + a' \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \begin{bmatrix} d & f \\ g & i \end{bmatrix} - b \begin{bmatrix} d' & f \\ g' & i \end{bmatrix} + c \begin{bmatrix} d & e \\ g & h \end{bmatrix} + c \begin{bmatrix} d' & e \\ g' & h \end{bmatrix}$$

$$= a \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} b \begin{bmatrix} d & f \\ g & i \end{bmatrix} c \begin{bmatrix} d & e \\ g & h \end{bmatrix} + a' \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \begin{bmatrix} d' & f \\ g' & i \end{bmatrix} + c \begin{bmatrix} d' & e \\ g' & h \end{bmatrix}$$
$$= \det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} + \det \begin{bmatrix} a' & b & c \\ d' & e & f \\ g' & h & i \end{bmatrix}$$

2.13. Remark:
$$\begin{aligned} f(\begin{pmatrix} b \\ e \\ h \end{pmatrix}, \begin{pmatrix} a \\ d \\ g \end{pmatrix}, \begin{pmatrix} c \\ f \\ i \end{pmatrix}) &= b \begin{bmatrix} d & f \\ g & i \end{bmatrix} - a \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} + \begin{bmatrix} d & e \\ g & h \end{bmatrix} \\ &= b \det(A_{12}) - a \det(A_{11}) + c(-\det(A_{13})) \\ &= -(a \det(A_{11}) - b \det(A_{12}) + c \det(A_{13})) \\ &= -f(A) \end{aligned}$$

2.14. Proposition: *Lett $g : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ be a multilinear alternating function. Then*
$$g(x_1, x_2, x_3) = f(x_1, x_2, x_3) g(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix})$$

2.15. Remark: There exists a unique multilinear alternating function g such that $g(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}) = 1$ and $g = f$ (Cor2.8)

2.16. Proposition: $\det(a_1, ..., a_n) = \det(a_1, ..., a_i + ba_j, ..., a_n)$ *where $a_i + ba_j$ is on the ith position.*

2.17. Proposition: *If an $n \times n$ matrix A is not invertible, then* $\det(A) = 0$.

2.18. Lemma: *If A is an $n \times n$ diagonal matrix, then* $\det(A) = a_{11}a_{22}...a_{nn}$.

2.19. Proposition: *If A is invertible, then* $\det(A) \neq 0$.

2.20. Theorem: Let A be an $n \times n$ matrix. A is invertible if and only if $\det(A) \neq 0$.

Section 3. Further Properties of The Determinant

3.1. Definition: Let A be an $n \times n$ matrix. The $n \times n$ matrix A' whose (i, j)-entry is $(-1)^{i+j}\det(A_{ji})$ is called the ***jith* cofactor** of A.

3.2. Proposition: $AA' = A'A = \det(A)I$

Proof. The (i, j)th entry of AA'
$= \sum_{k=1}^{n} a_{ik} \cdot (k, j)$th entry of A'
$= \sum_{k=1}^{n} a_{ik} \cdot (-1)^{k+j} \cdot \det(A_{jk})$
$= \sum_{k=1}^{n} (-1)^{k+j} \cdot a_{ik} \cdot \det(A_{jk})$
$= \det(A)I$ □

3.3. Corollary: *If A is an invertible $n \times n$ matrix, then A^{-1} us the matrix whose ijth entry is* $(-1)^{i+j}\det(A_{ji})/\det(A)$

3.4. Proposition: *For any fixed i, $1 \le i \le n$*

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

3.5. Proposition: *If A and B are $n \times n$ matrices, then*

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

(b) If A is invertible, then $\det(A^{-1}) = (\det(A))^{-1}$

3.6. Corollary: *If $T : V \to V$ is a linear transformation,* $\dim(V) = n$*, then*

$$\det([T]_{\alpha}^{\alpha}) = \det([T]_{\beta}^{\beta})$$

for all choices of bases α and β for V.

3.7. Definition: The **determinant** of a linear transformation $T : V \to V$ of a finite-dimensional vector space is the determinant of $\det([T]_{\alpha}^{\alpha})$ for any choice of α. We denote this by $\det(T)$

3.8. Proposition: *Let $S : V \to V$ and $T : V \to V$ be linear transformations of a finite-dimensional vector space, then*

(a) $\det(ST) = \det(S)\det(T)$ *and*

(b) if T is isomorphism, $\det(T^{-1}) = (\det(T))^{-1}$

3.9. Proposition: *The linear transformation f is invertible if and only if* $\det(f) \neq 0$.

3.10. Proposition (Cramer's rule): *Let A be an invertible $n \times n$ matrix. The solution x to the system of equations $Ax = b$ is the vector whose jth entry is the quotient*

$$\det(B_j)/\det(A)$$

where B_j is the matrix obtained from A by replacing the jth column of A by the vector b.

Explanation.
Since A is invertible, we have $A^{-1}(Ax) = A^{-1}b$ i.e., $x = A^{-1}b = \dfrac{1}{\det A} \cdot A' \cdot b$.
jth entry of $A'b = \sum_{k=1}^{n} (k,j)$th entry of $A' \cdot b_k = \sum_{k=1}^{n} (-1)^{j+k} \det(A_{jk}) b_k$.
Consider $B_j := \begin{bmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{bmatrix}$ where $\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$ is the jth column.
Expand $\det(B_j)$ along the jth column:
$\det(B_j) = \sum_{k=1}^{n} (-1)^{k+j} \cdot (k,j)$th entry of $B_j \cdot \det((k,j)$th minor of $B_j) = \sum_{k=1}^{n} (-1)^{j+k} \cdot b_k \det(A_{jk})$.
Therefore, the jth entry of $A'b$ is $\det(B_j)$.
Hence, $x_j = j$th entry of $(\frac{1}{\det(A)} A'b) = \frac{1}{\det(A)} \cdot j$th entry of $(A'b) = \frac{\det(B)}{\det(A)}$

3.11. Example: Solve $\begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
$A = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}, b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, B_1 = \begin{bmatrix} 0 & 1 \\ -1 & 3 \end{bmatrix}, B_2 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$.
$\det(A) = 7, \det(B_1) = -1, \det(B_2) = 2$.
Hence, $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -\frac{1}{7} \\ \frac{2}{7} \end{pmatrix}$

Chapter 4

Eigenvalues, Eigenvectors, Diagonalization, and the Spectral Theorem in $\mathbb{R}^n$

Section 1. Eigenvalues and Eigenvectors

1.1. Definition: Let $T : V \to V$ be a linear mapping.

(a) A vector $x \in V$ is called an **eigenvector** of T if $x \neq 0$ and there exists a scalar $\lambda \in \mathbb{R}$ such that $T(x) = \lambda x$.

(b) If x is an eigenvector of T and $T(x) = \lambda x$, the scalar λ is called the **eigenvalue** of T corresponding to x.

1.2. Example: $V = \mathbb{R}^2 \in L$: 1-dimensional subspace $Porj_L : V \to V$.
If $0 \neq x \in L$, then $Proj_L(x) = x = 1 \cdot x$.
If $0 \neq x$ and x is perpendicular to $L, Proj_L(x) = \vec{0} = 0 \cdot x$.

1.3. Proposition: *A vector x is an eigenvector of T with eigenvalue λ if and only if*

$$x \neq 0 \text{ and } \in \operatorname{Ker}(T - \lambda I).$$

1.4. Definition: Let $T : V \to V$ be a linear mapping, and let $\lambda \in \mathbb{R}$. The λ-**eigenspace** of T, denoted E_λ, is the set

$$E_\lambda = \{x \in V \mid T(X) = \lambda x\}$$

If λ is not an eigenvalue of T, then $E_\lambda = \{0\}$.

1.5. Proposition: *E_λ is a subspace of V for all λ.*

1.6. Example: $Proj_L : \mathbb{R}^2 \to \mathbb{R}^2$.

$$E_1 = \operatorname{Ker}(Proj_L - 1 \cdot I\mathbb{R}^2) = \{x \in \mathbb{R}^2 : Proj_L(x) = x\} = L$$

$$E_0 = \operatorname{Ker}(Proj_L - 0 \cdot I\mathbb{R}^2) = \{x \in \mathbb{R}^2 : Proj_L(x) = 0\} = L^2$$

If $\lambda \neq 0, 1, E_\lambda = \{0\}$.

1.7. Proposition:
Let $A \in M_{n\times n}(\mathbb{R})$. Then $\lambda \in \mathbb{R}$ is an eigenvalue of A if and only if $\det(A - \lambda I) = 0$.

Explanation. $\det(A - \lambda I_n) = 0$
$\Leftrightarrow A - \lambda I_n$ is not invertible
$\Leftrightarrow \operatorname{Ker}(A - \lambda I_n) \neq \{0\}$
$\Leftrightarrow E_\lambda \neq 0$
$\Leftrightarrow \lambda$ if an eigenvalue $\square$

1.8. Example: $Proj_L : \mathbb{R}^2 \to \mathbb{R}^2$.
$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \rightsquigarrow \det(A - \lambda I_2) = \det(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}) = \det(\begin{pmatrix} 1-\lambda & 0 \\ 0 & -\lambda \end{pmatrix})$
$= (1-\lambda)(-\lambda) = \lambda(\lambda - 1)$
$\Rightarrow \lambda_1 = 0, \lambda_2 = 1$

1.9. Definition: Let $A \in M_{n\times n}(\mathbb{R})$. The polynomial $\det(A - \lambda I)$ is called the **characteristic polynomial** of A.

1.10. Example: $n = 1, A = \begin{pmatrix} a \end{pmatrix}$
$P_A(\lambda) = \det(\begin{pmatrix} a \end{pmatrix} - \lambda \begin{pmatrix} 1 \end{pmatrix}) = a - \lambda$
$\Rightarrow$ degree 1 polynomial in λ.
$n = 2, A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$
$P_A(\lambda) = \det(\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}) = (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + (ad - bc)$
$\Rightarrow$ degree 2 polynomial in λ.

1.11. Proposition: *Similar matrices have equal characteristic polynomials.*

1.12. Remark: Let $f : V \to V$ be linear transformation and α, β be bases for V.
Since $[f]_\alpha^\alpha$ and $[f]_\beta^\beta$ are similar.
Thus, $P_{[f]_\alpha^\alpha}(\lambda) = P_{[f]_\beta^\beta}(\lambda)$ which is the characteristic polynomial of f.

1.13. Remark: The coefficient of λ is called the **trace** of the matrix A, and denoted by $\mathrm{Tr}(A)$.
In general, the trace is defined to be the sum of the diagonal entries.
For any $n \times n$ matrix A, the characteristic polynomial has the form

$$(-1)^n \lambda^n + (-1)^{n-1}\mathrm{Tr}(A)\lambda^{n-1} + c_{n-2}\lambda^{n-2} + ... + c_1\lambda + \det(A)$$

1.14. Corollary: *Let $A \in M_{n\times n}(\mathbb{R})$. Then A has no more than n distinct eigenvalues. In addition, if $\lambda_1, ..., \lambda_k$ are the distinct eigenvalues of A and λ_i is an m_i-fold root of the characteristic polynomial, then $m_1 + ... + m_k \leq n$.*

1.15. Theorem: Let $A \in M_{n\times n}(\mathbb{R})$, and let $p(t) = \det(A - tI)$ be its characteristic polynomial. Then $p(A) = 0$ (the $n \times n$ zero matrix)

Section 2. Diagonalizability

2.1. Definition: Let V be a finite-dimensional vector space, and let $T : V \to V$ be a linear mapping. T is said to be **diagonalizable** if there exists a basis of V, all of whose vectors are eigenvectors of T.

2.2. Proposition: *$T : V \to V$ is diagonalizable if and only if, for any basis α of V, the matrix $[T]_\alpha^\alpha$ is similar to a diagonal matrix.*

Proof. Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation.

Then $[A]_\alpha^\alpha = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}$ $\beta = \{e_1, ..., e_n\} \to A = [A]_\beta^\beta = [I]_\alpha^\beta [A]_\alpha^\alpha [I]_\beta^\alpha$

$\rightsquigarrow$ A is similar to $[A]_\alpha^\alpha$ $\square$

2.3. Example: Let the mapping of $V = \mathbb{R}^3$ defined by the matrix $A = \begin{bmatrix} 2 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$.

The characteristic polynomial of A is $(2-\lambda)^2(1-\lambda) \Rightarrow \lambda_1 = 1, \lambda_2 = \lambda_3 = 2$.

By calculating, $E_1 = \text{Span}\{\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}\}, E_2 = \text{Span}\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$.

If we form the change of basis matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$,

then it is easy to see that $Q^{-1}AQ = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, which diagonalizes A.

2.4. Example: Let the mapping of $V = \mathbb{R}^2$ defined by the matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

The characteristic polynomial of A is $\lambda^2 \Rightarrow \lambda_1 = \lambda_2 = 0$.

By calculating, $E_0 = \text{Span}\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\}$

$\rightsquigarrow \dim(E_0) = 1 < 2 = \dim(\mathbb{R}^2) \rightsquigarrow A$ is not diagonalizable

2.5. Proposition: *Let $x_i (1 \le i \le k)$ be eigenvectors of a linear mapping $T : V \to V$ corresponding to distinct eigenvalues λ_i. Then $\{x_1, ..., x_k\}$ is a linearly independent subset of V.*

2.6. Corollary: *For each $i(1 \leq i \leq k)$, let $\{x_{i.1}, ..., x_{i.n_i}\}$ be a linearly independent set of eigenvectors of T all with eigenvalue λ_i and suppose the λ_i are distinct. Then $S = \{x_{1.1}, ..., x_{1.n_1}\} \cup ... \cup \{x_{k.1}, ..., x_{k.n_k}\}$ is linearly independent.*

2.7. Corollary: *Let n be the dimension of V. Suppose that the linear transformation $f : V \to V$ has n distinct eigenvalues, then f is diagonalizable.*

2.8. Proposition: *Let λ be a root of the characteristic polynomial of f with multiplicity m, then*

$$1 \leq \dim(E_\lambda) \leq m$$

2.9. Theorem: Let $T : V \to V$ be a linear mapping on a finite-dimensional vector space V, and let $\lambda_1, ..., \lambda_k$ be its distinct eigenvalues. Let m_i be the multiplicity of λ_i as a root of the characteristic polynomial of T. Then T is diagonalizable if and only if

(i) $m_1 + ... + m_k = n = \dim(V)$, and

(ii) for each $i, \dim(E_{\lambda i}) = m_i$

2.10. Corollary: *A linear mapping $T : V \to V$ on a finite-dimensional space V is diagonlizable if and only if the sum of the multiplicities of the real eigenvalues is $n = \dim(V)$ and either*

(i) We have $\sum_{i=1}^{k} \dim(E_{\lambda_i}) = n$, or

(ii) We have $\sum_{i=1}^{k}(n - \dim(\mathrm{Im}(T - \lambda_i I))) = n$,

where the λ_i are the distinct eigenvalues of T.

Section 3. Geometry In $\mathbb{R}^n$

3.1. Definition: The **standard inner product** (or dot product) on $\mathbb{R}^n$ is the function $\langle -, - \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by the following rule:

$$\langle x, y \rangle = x_1 y_1 + ... + x_n y_n$$

if $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ in standard coordinates.

3.2. Proposition:

(a) $\forall c_1, c_2 \in \mathbb{R}, x_1, x_2 \in \mathbb{R}^n, \langle c_1 x_1 + c_2 x_2, y \rangle = c_1 \langle x_1, y \rangle + c_2 \langle x_2, y \rangle$

(b) $\forall x, y \in \mathbb{R}^n, \langle x, y \rangle = \langle y, x \rangle$

(c) $\forall x \in \mathbb{R}^n, \langle x, x \rangle \geq 0$ *and* $\langle x, x \rangle = 0$ *if and only if* $x = 0$.

3.3. Definition: Let V be a vector space. An **inner product** on V is a function $\langle -, - \rangle : V \times V \to \mathbb{R}$ which satisfies the properties above.

3.4. Definition:

(a) The **length** (or **norm**) of $x \in \mathbb{R}^n$ is the scalar

$$\|x\| = \sqrt{\langle x, x \rangle}$$

(b) x is called a **unit vector** if $\|x\| = 1$

3.5. Remark: If $x \neq \vec{0}$, then $\frac{x}{\|x\|}$ will be a unit vector

3.6. Proposition:

(a) ***The triangle inequality:*** $\forall x, y \in \mathbb{R}^n, \|x + y\| \leq \|x\| + \|y\|$.

(b) ***The Cauchy-Schwarz inequality:*** $\forall x, y \in \mathbb{R}^n, |\langle x, y \rangle| \leq \|x\| \cdot \|y\|$

3.7. Definition: The **angle, θ, between two nonzero vectors** $x, y \in \mathbb{R}^n$ is defined to be

$$\theta = \cos^{-1}\left(\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}\right)$$

3.8. Definition: Two vectors $x, y \in \mathbb{R}^n$ are said to be **orthogonal** (or perpendicular) if

$$\langle x, y \rangle = 0$$

3.9. Definition:

(a) A set of vectors $S \subset \mathbb{R}^n$ is said to be **orthogonal** if for every pair of vectors $x, y \in S$ with $x \neq y$, we have $\langle x, y \rangle = 0$.

(b) A set of vectors $S \subset \mathbb{R}^n$ is said to be **orthonormal** if S is orthogonal and, in addition, every vector in S is a unit vector.

3.10. Proposition: *If $x, y \in \mathbb{R}^n$ are orthogonal, nonzero vectors, then $\{x.y\}$ is linearly independent.*

Proof. Suppose $ax + by = 0$ for some scalars a, b. Then we have

$$\begin{aligned} \langle ax + by, x \rangle &= a\langle x, x \rangle + b\langle y, x \rangle \\ &= a\langle x, x \rangle \quad \text{since } \langle x, y \rangle = 0 \end{aligned}$$

On the other hand, $ax + by = 0$. So $\langle ax + by, x \rangle = \langle 0, x \rangle = 0$.
Since $x \neq 0, \langle x, x \rangle \neq 0$, so we must have $a = 0$.
Similarly, $b = 0$ as well. Hence, x and y are linearly independent. $\square$

Section 4. Orthogonal Projections And The Gram-Schmidt Process

4.1. Definition: The **orthogonal complement** of W, denoted $W^\perp$, is the set

$$W^\perp = \{v \in \mathbb{R}^n \mid \langle v, w \rangle = 0 \quad \text{for all } w \in W\}$$

4.2. Example:

(i) $W = \{0\} \Rightarrow \{0\}^\perp = \{x \in \mathbb{R}^n : \langle x, 0 \rangle = 0\} = \mathbb{R}^n$

(ii) $W = \mathbb{R}^n \Rightarrow (\mathbb{R}^n)^\perp = \{x \in \mathbb{R}^n : \langle x, y \rangle = 0\} \Rightarrow x = 0 \Rightarrow (\mathbb{R}^n)^\perp = \{0\}$

(iii) $W = \text{plane} \Rightarrow W^\perp = \text{line}$

(iv) $W = \text{line} \Rightarrow W^\perp = \text{plane}$

4.3. Proposition:

(a) For every subspace W of $\mathbb{R}^n$, $W^\perp$ is also a subspace of $\mathbb{R}^n$.

(b) $\dim(W) + \dim(W^\perp) = \dim(\mathbb{R}^n) = n$.

(c) For all subspaces W of $\mathbb{R}^n$, $W \cap W^\perp = \{0\}$.

(d) $\forall x \in \mathbb{R}^n, \exists x_1 \in W, x_2 \in W^\perp$ *s.t* $x_1 + x_2 = x \Rightarrow \mathbb{R}^n = W \oplus W^\perp$

4.4. Definition:
The map $P_W : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto x_1$ is called th **orthogonal projection** of $\mathbb{R}^n$ onto W.

4.5. Proposition:

(a) P_W is a linear mapping.

(b) $\mathrm{Im}(P_W) = W$ *and if $w \in W$, then $P_W(w) = w$*

(c) $\mathrm{Ker}(P_W) = W^\perp$

4.6. Definition: A basis $\{w_1, ..., w_k\}$ for W is **orthonormal** if

$$\langle w_i, w_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

4.7. Proposition: *Let $\{w_1, ..., w_k\}$ be an orthonormal basis for the subspace $W \subseteq \mathbb{R}^n$.*

(a) $\forall w \in W, w = \sum_{i=1}^k \langle w, w_i \rangle w_i$.

(b) $\forall x \in V, P_W(x) = \sum_{i=1}^k \langle x, w_i \rangle w_i$

4.8. Note (Gram-Schmidt Process):
$W \subseteq \mathbb{R}^n$: subspace, $\{w_1, ..., w_k\}$: basis for W.

(1) Put $w_1' := \frac{1}{\|w_1\|} \cdot w_1$

(2) Put $\widetilde{w_2} := w_2 - P_{span(w_1)}(w_2) = w_2 - \dfrac{\langle w_1, w_2 \rangle}{\langle w_1, w_1 \rangle} \cdot w_1$

$w_2' := \frac{1}{\|w_2\|} \cdot \widetilde{w_2}$

$\vdots$

(k) Put $\widetilde{w_k} := w_k - \dfrac{\langle w_1, w_k \rangle}{\langle w_1, w_1 \rangle} \cdot w_1 - \dfrac{\langle \widetilde{w_2}, w_k \rangle}{\langle \widetilde{w_2}, \widetilde{w_2} \rangle} \cdot \widetilde{w_2} - ... - \dfrac{\langle \widetilde{w_{k-1}}, w_k \rangle}{\langle \widetilde{w_{k-1}}, \widetilde{w_{k-1}} \rangle} \cdot \widetilde{w_{k-1}}$

$w_k' := \frac{1}{\|w_k\|} \cdot \widetilde{w_k}$

4.9. Theorem: Let W be a subspace of $\mathbb{R}^n$. Then there exists an orthonormal basis of W.

4.10. Note: $\{w_1', w_2', ..., w_k'\}$ is the orthonormal basis of W.

Section 5. Symmetric Matrices

5.1. Definition: A square matrix A is said to be **symmetric** if $A = A^T$, where A^T denotes the transpose of A produced by swapping the rows and columns of A.

5.2. Remark: By the definition of the transpose of a matrix, A is symmetric is and only if $a_{ij} = a_{ji}$ for all pairs i, j.

5.3. Proposition: *Let $A \in M_{n \times n}(\mathbb{R})$.*

(i) For all $x, y \in \mathbb{R}^n$, $\langle Ax, y \rangle = \langle x, A^T y \rangle$

(ii) A is symmetric if and only if $\langle Ax, y \rangle = \langle x, Ay \rangle$

5.4. Corollary: *Let V be any subspace of $\mathbb{R}^n$, let $T : V \to V$ be any linear mapping, and let $\alpha = \{x_1, ..., x_k\}$ be any orthonormal basis of V. Then $[T]_\alpha^\alpha$ is a symmetric matrix if and only if $\langle T(x), y \rangle = \langle x, T(y) \rangle$ for all vectors $x, y \in V$.*

5.5. Definition: Let V be a subspace of $\mathbb{R}^n$. A linear mapping $T : V \to V$ is said to be **symmetric** if if $\langle T(x), y \rangle = \langle x, T(y) \rangle$ for all vectors $x, y \in V$.

5.6. Theorem: Let $A \in M_{n\times n}(\mathbb{R})$ be a symmetric matrix. Then all the roots of the characteristic polynomial of A are real. In other words, the characteristic polynomial has n roots in $\mathbb{R}$ (counted with multiplicities).

5.7. Theorem: Let $A \in M_{n\times n}(\mathbb{R})$ be a symmetric matrix, let x_1 be an eigenvector of A with eigenvalue λ_1, and let x_3 be an eigenvector of A with eigenvalue λ_2, where $\lambda_1 \neq \lambda_2$. Then x_1 and x_2 are orthogonal vectors in $\mathbb{R}^n$.

Section 6. The Spectral Theorem

6.1. Theorem: Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a symmetric linear mapping. Then there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of T. In particular, T is diagonalizable.

6.2. Example: The mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ whose matrix with respect to the standard basis is

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

Since $A' = A$ and the standard basis is orthonormal, T is a symmetric mapping.
Computing, we find

$$\det(A - \lambda I) = -\lambda(\lambda - 2)^2$$

The eigenvalues of T are $\lambda = 0$ and $\lambda = 2$.
For $\lambda = 0$, we have $E_0 = \text{Span}\{(-1, 1, 0)\}$, whereas for $\lambda = 2$, $E_2 = \text{Span}\{(1, 1, 0), (0, 0, 1)\}$.
Since $\{(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)\}$ is an orthonormal basis of E_0.
And $\{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (0, 0, 1)\}$ is an orthonormal basis of E_2.
Hence, $\alpha = \{(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (0, 0, 1)\}$ is an orthonormal basis of $\mathbb{R}^3$.
With respect to this basis, $[T]_\alpha^\alpha = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is a diagonal matrix.

6.3. Theorem: Let $T : \mathbb{R}^n \to \mathbb{R}^n$ by a symmetric linear mapping, and let $\lambda_1, ..., \lambda_k$ be the distinct eigenvalues of T. Let P_i be the orthogonal projection of $\mathbb{R}^n$ onto the eigenspace E_{λ_i}. Then

(a) $T = \lambda_1 P_1 + ... + \lambda_k P_k$, and

(b) $I = P_1 + ... + P_k$

6.4. Remark: The quadratic terms can be interpreted as a matrix product

$$Ax_1^2 + 2Bx_1x_2 + Cx_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Chapter 5

Complex Numbers and Complex Vector Spaces

Section 1. Complex Numbers

1.1. Definition: The set of **complex numbers**, denoted $\mathbb{C}$, is the set of ordered pairs of real numbers (a, b) with the operations of addition and multiplication defined by

(i) For all (a, b) and $(c, d) \in \mathbb{C}$, the **sum** of (a, b) and (c, d) is the complex number defined by

$$(a, b) + (c, d) = (a + b, c + d)$$

(ii) and the **product** of (a, b) and (c, d) is the complex number defined by

$$(a, b)(c, d) = (ac - bd, ad + cb)$$

1.2. Remark: The subset of $\mathbb{C}$ consisting of those elements with second coordinate zero, $\{(a, 0) \mid a \in \mathbb{R}\}$, will be identified with the real numbers in the obvious way

$$a \in \mathbb{R} = (a, 0) \in \mathbb{C}$$

1.3. Definition: Let $z = a + bi \in \mathbb{C}$. The **real part of z**, denoted $\mathscr{R}(z)$, is the real number a. The **imaginary part of z**, denoted $\mathscr{I}(z)$, is the real number b. z is called a **real number** if $\Im(z) = 0$ and **purely imaginary** if $\mathscr{R}(z) = 0$.

1.4. Remark:

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

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

1.5. Definition: A **field** is a set F with two operations, defined on ordered pairs of elements of F, called addition and multiplication. Addition assigns to the pair x and $y \in F$ their **sum** and multiplication assigns to the pair x and $y \in F$ their **product**. These two operations must satisfy the following properties for all x, y, and $z \in F$.

(i) **Commutativity of addition:** $x + y = y + x$.

(ii) **Associativity of addition:** $(x + y) + z = x + (y + z)$

(iii) **Existence of an additive identity:** $\exists 0 \in F$, such that $x + 0 = x$.

(iv) **Existence of additive inverse:** $\forall x, \exists -x \in F$ such that $x + (-x) = 0$

(v) **Commutativity of multiplication:** $xy = yx$

(vi) **Associativity of multiplication:** $(xy)z = x(yz)$

(vii) **Distributivity:** $(x + y)z = xz + yz$ and $x(y + z) = xy + xz$

(viii) **Existence of a multiplicative identity:** $\exists 1 \in F$, such that $x \cdot 1 = x$.

(ix) **Existence of multiplicative inverses:** If $x \neq 0$, then there is an element $x^{-1} \in F$ such that $xx^{-1} = 1$

1.6. Proposition: *The set of complex numbers is a field with the operations of addition and scalar multiplication as defined previously.*

1.7. Proposition:

(i) The additive identity in a field is unique.

(ii) The additive inverse of an element of a field is unique.

(iii) The multiplicative identity of a field is unique.

(iv) The multiplicative inverse of a nonzero element of a field is unique.

1.8. Definition: The **absolute value** of the complex number $z = a + bi$ is the nonnegative real number $\sqrt{a^2 + b^2}$ and is denoted by $|z|$ or $r = |z|$. The **argument** of the complex number z is the angle θ of the polar coordinate representation of z.

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

1.9. Remark:

$$z_1 z_2 = r_1 r_2(\cos(\theta_1 + \theta_2) + i \cdot \sin(\theta_1 + \theta_2))$$

$$z^n = r^n(\cos(n\theta) + i \cdot \sin(n\theta))$$

where $r^n = r_0$ and $b\theta = \theta_0 + 2\pi k$ for k an integer

1.10. Remark: The n nth roots of the complex number $z_0 \neq 0$ are

$$r_0^{1/n}(\cos(\frac{\theta_0}{n} + \frac{2\pi k}{n}) + i\sin(\frac{\theta}{n} + \frac{2\pi k}{n}))$$

for $k = 0, 1, ..., n - 1$

1.11. Example: Solve $z^4 - 1 = 0$.
Suppose $z = r(\cos\theta + i\sin\theta), r > 0$.
Then $z^4 - 1 = 0$ is equivalent to

$$r^4(\cos(4\theta) + i\sin(4\theta)) = 1 = 1(\cos(2k\pi) + i\sin(2k\pi))$$

Equivalently,

$$\begin{cases} r^4 = 1 \\ 4\theta = 2k\pi, \quad k \in \mathbb{Z} \end{cases}$$

Hence,

$$\begin{cases} r = 1 \\ \theta = \frac{k}{2}\pi, \quad k \in \mathbb{Z} \end{cases}$$

Therefore, $z = \cos\theta + i\sin\theta$

$$k = 0 \to z = 1, k = 1 \to z = i, k = 2 \to z = -1, k = 3 \to z = -i$$

1.12. Definition: A field F is called **algebraically closed** if every polynomial $p(z) = a_n z^n + ... + a_1 z + a_0$ with $a_i \in F$ and $a_n \neq 0$ has n roots counted with multiplicity in F.

1.13. Theorem: $\mathbb{C}$ is algebraically closed and $\mathbb{C}$ is the smallest algebraically closed field containing $\mathbb{R}$.

1.14. Example (non-example):
$\mathbb{Q}$ is not algebraically closed.
Since $\pm\sqrt{2}$ are not rational. But $x^2 - 2$ is a polynomial with coefficients in $\mathbb{Q}$.
$\mathbb{R}$ is not algebraically closed.
The polynomial $x^2 + 1$ is a polynomial with coefficient in $\mathbb{R}$. But it does not have real roots.

Section 2. Vector Spaces Over A Field

2.1. Definition: A **vector space** over a field F is a set V (whose elements are called **vectors**) together with

(a) an operation called **vector addition**, which for each pair of vectors $x, y \in V$ produces a vector denoted $x + y \in V$, and

(b) an operation called **multiplication** by a **scalar** (a field element), which for each vector $x \in V$, and each scalar $c \in F$ produces a vector denoted $cx \in V$.

2.2. Axiom: *These two operations must satisfy the following axioms:*

(1) ***associativity:*** $\forall x, y, z \in V, (x + y) + z = x + (y + z)$

(2) ***commutativity:*** $\forall x, y \in V, x + y = y + x$

(3) ***existence of*** $\vec{0}$***:*** $\exists \vec{0} \in V$ *s.t* $\forall x \in V,\ \vec{0} + x = x + \vec{0} = x$

(4) ***inverse:*** $\forall x \in V, \exists -x \in V$ *s.t* $x + (-x) = \vec{0}$ *('$-$' just a symbol)*

(5) ***distributivity:*** $\forall c \in F, x, y \in V, c \cdot (x + y) = c \cdot x + c \cdot y$

(6) $\forall c, d \in F, x \in V, (c + d) \cdot x = c \cdot x + d \cdot x$

(7) $\forall c, d \in F, x \in V, (c \cdot d) \cdot x = c \cdot (d \cdot x)$

(8) $\forall x \in V, 1 \cdot x = x$

2.3. Example: We will find a solution to the following system of equations over $\mathbb{C}$.

$$(1 + i)x_1 + (3 + i)x_3 = 0$$
$$x_1 - ix_2 + (2 + i)x_3 = 0$$

We add $-(1+i)^{-}1$ times the first equations to the second.

$$\begin{aligned}(1+i)x_1 + (3+i)x_3 &= 0\\ -ix_2 + (2i)x_3 &= 0\end{aligned}$$

Multiply through to make the leading coefficients 1:

$$\begin{aligned}x_1 + (2-i)x_3 &= 0\\ x_2 - 2x_3 &= 0\end{aligned}$$

x_3 is the only free variable, so that setting $x_3 = 1$, we obtain the solution $(-2+i, 2, 1)$.

2.4. Definition: Let V and W be vector spaces over a field F. A linear transformation $T : V \to W$ is a function from V to W, which satisfies

$$T(au + bv) = aT(u) + bT(v)$$

for u and $v \in V$ and a and $b \in F$.

2.5. Definition: A subset W of V is a **subspace** if $(W, +, \cdot)$ is itself a vector space over F.

2.6. Theorem: A subset W of V is a subspace iff for all $x, y \in W$ and $c \in F$

$$c \cdot x + y \in W$$

Section 3. Geometry In A Complex Vector Space

3.1. Example: $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$
$P_A(\lambda) = \det(A - \lambda I_2) = \det \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} = \lambda^2 + 1.$
No real eigenvalues. A is not diagonalizable over $\mathbb{R}$.
Over $\mathbb{C}$, the root of $P_A(\lambda)$ are $\pm i$.
Hence, the complex matrix A has two distinct eigenvalues.
Therefore, A is diagonalizable over $\mathbb{C}$.

$$E_i = \operatorname{Ker}(A - iI_2) = \operatorname{Ker} \begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix} = \operatorname{Span}_{\mathbb{C}}(\begin{bmatrix} -i \\ 1 \end{bmatrix})$$

$$E_{-i} = \operatorname{Ker}(A + iI_2) = \operatorname{Ker} \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} = \operatorname{Span}_{\mathbb{C}}(\begin{bmatrix} i \\ 1 \end{bmatrix})$$

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

3.2. Definition: Let V be a complex vector space. A **Hermitian inner product** on V is a complex valued function on pairs of vectors in V, denoted by $\langle u, v\rangle \in \mathbb{C}$ for $u, v \in V$, which satisfies the following properties:

(a) For all $u, b, w \in V$ and $a, b \in \mathbb{C}$, $\langle au + bv, w\rangle = a\langle u, w\rangle + b\langle v, w\rangle$.

(b) For all $u, v \in V$, $\langle u, v\rangle = \overline{\langle v, u\rangle}$, and

(c) For all $v \in V$, $\langle v, v\rangle \geq 0$ and $\langle v, v\rangle = 0$ implies $v = 0$.

3.3. Remark: If $a + ib \in \mathbb{C}$, then $\overline{a + ib} := a - ib$.

3.4. Example: V= $\mathbb{C}^n$

$$\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle = x_1\overline{y_1} + ... + x_n\overline{y_n}$$

$$\begin{aligned}
\langle \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \rangle &= y_1\overline{x_1} + ... + y_n\overline{x_n} \\
&= \overline{\overline{y_1}}\ \overline{x_1} + ... + \overline{\overline{y_n}}\ \overline{x_n} \\
&= \overline{\overline{y_1}x_1} + ... + \overline{\overline{y_n}x_n} \\
&= \overline{\overline{y_1}x_1 + ... + \overline{y_n}x_n} \\
&= \overline{\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle}
\end{aligned}$$

3.5. Remark: Let V is a complex vector space and $\langle -, -\rangle$ a Hermitian inner product on V. Can define perpendicularity, length/norm of vector, orthogonal projection, Gram-Schmidt process.

3.6. Definition: Let V be a complex vector space with a Hermitian inner product. The **norm** or **length** of a vector $v \in V$ is $\|v\| = \langle v, v\rangle^{1/2}$. A set of nonzero vectors $v_1, ..., v_k \in V$ is called **orthogonal** if $\langle v_i, v_j\rangle = 0$ for $i \neq j$. If in addition $\langle v_i, v_i\rangle = 1$ for all i, the vectors are called **orthonormal**.

3.7. Definition: Let $A \in Mat_{2\times 2}(\mathbb{C})$. The **Hermitian transpose** A^* of A is the $n \times m$ matrix whose (i, j)-th entry is $\overline{a_{ji}}$

3.8. Example: $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^* = (\overline{x_1}, ..., \overline{x_n})$.

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}^* \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\overline{y_1}, ..., \overline{y_n}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\overline{y_1} + ... + x_n\overline{y_n} = \langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle$$

3.9. Definition: $T : V \to V$ is called **Hermitian** or **self-adjoint** if $T = T^*$.
An $n \times n$ complex matrix is called **Hermitian** or **self-adjoint** if $A = A^*$.

3.10. Theorem: If λ is an eigenvalue of the self-adjoint linear transformation T, then $\lambda \in \mathbb{R}$.

3.11. Proposition: *If u and v are eigenvectors, respectively, for the distinct eigenvalues λ and μ of $T : V \to V$, then u and v are **orthogonal**.*

3.12. Theorem: Let $T : V \to V$ be a self-adjoint transformation of a complex vector space V with Hermitian inner product. Then there is an orthonormal basis of V consisting of eigenvectors for T and T is diagonalizable.

3.13. Theorem: Let $T : V \to V$ be a self-adjoint transformation of a complex vector space V with Hermitian inner product. Let $\lambda_1, ..., \lambda_k \in \mathbb{R}$ be the distinct eigenvalues of T, and let P_i be the orthogonal projections of V onto the eigenspaces E_{λ_i}, then

(a) $T = \lambda_1 P_1 + ... + \lambda_k P_k$

(b) $I = P_1 + ... + P_k$

Chapter 6

Jordan Canonical Form

Section 1. Triangular Form

1.1. Definition: Let $f : V \to V$ be a linear transformation. A subspace W of V is **invariant** or **stable** under f if

$$f(W) \subseteq W$$

1.2. Example:

(a) $\{0\}$ and V itself are invariant under all linear mappings $T : V \to V$.

(b) $\mathrm{Ker}(T)$ and $\mathrm{Im}(T)$ are invariant subspaces as well.

(c) If λ is an eigenvalue of T, then the eigenspace E_k is invariant under T as well.
Since $\forall v \in E_\lambda, T(v) = \lambda v \in E_\lambda$.

1.3. Proposition: *Let $\beta = \{x_1, ..., x_n\}$ be a basis for V. Then $[f]_\beta^\beta$ is upper triangular if and only if each of the subspaces $W_i = Span(x_1, ..., x_i)$ is invariant under T.*

1.4. Remark: $\{0\} \subset W_1 \subset W_2 \subset ... \subset W_n = V$

1.5. Definition: A linear mapping $T : V \to V$ on a finite-dimensional vector space V is **triangularizable** if there exists a basis β such that $[T]_\beta^\beta$ is upper-triangular.

1.6. Corollary: $\{$*eigenvalues of* $f \mid w \subseteq$ *eigenvalues of* $f\}$

1.7. Theorem: Let V be a finite-dimensional vector space over a field F, and let $T : V \to V$ be a linear mapping. Then T is triangularizable if and only if the $P_f(\lambda)$ has $\dim(V)$ roots (counted with multiplicities) in the field F.

1.8. Corollary: *If $F = \mathbb{C}$, then f is triangularizable.*

Proof. Let $n = \dim(V)$.
By the theorem, enough to show that $P_f(\lambda)$ has n roots (counting multiplicity) in $\mathbb{C}$.
Recall that $P_f(\lambda)$ is a polynomial of degree n with coefficients in $\mathbb{C}$.
By the fundamental theorem of algebra, $P_f(\lambda)$ has n roots in $\mathbb{C}$. □

1.9. Lemma: *Let $T : V \to V$ be as in the theorem, and assume that $P_T(\lambda)$ has $n = \dim(V)$ roots in F. If $W \subsetneq V$ is an invariant subspace under T, then there exists a vector $x \neq 0$ in V such that $x \notin W$ and $W + Span(\{x\})$ is also invariant under T.*

1.10. Corollary: *If $T : V \to V$ is triangularizable, with eigenvalues λ_i with respective multiplicities m, then there exists a basis β for V such that $[T]_\beta^\beta$ is upper-triangular, and the diagonal entries of $[T]_\beta^\beta$ are $m_1 \lambda_1$'s, followed by $m_2 \lambda_2$'s, and so on.*

1.11. Theorem (Cayley-Hamilton): Let $T : V \to V$ be a linear mapping on a finite-dimensional vector space V, and let $p(t) = \det(T - tI)$ be its characteristic polynomial. Assume that $p(t)$ has $\dim(V)$ roots in the field F over which V is defined. Then $p(T) = 0$.

1.12. Remark: If

$$P_A(A) = (-1)^n A^n + ... + a_1 A + \det(A) I = 0$$

then

$$A^{-1} = \frac{-1}{\det(A)}((-1)^n A^{n-1} + ... + a_1 I)$$

1.13. Example: Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 1 \\ 0 & 1 & 3 \end{bmatrix}$.
We have $p(t) = -t^3 + 3t^2 + 6t - 16$.
Then,

$$\begin{aligned} A^{-1} &= \frac{1}{16}(-A^2 + 3A + 6I) \\ &= \frac{1}{16}\begin{bmatrix} 4 & 6 & -2 \\ 6 & -3 & 1 \\ -2 & 1 & 5 \end{bmatrix} \end{aligned}$$

Section 2. A Canonical Form For Nilpotent Mappings

2.1. Definition: $f : V \to V$ is **nipotent** if $f \circ ... \circ f = 0$ for some $k \in \mathbb{N}$.
A matrix $A \in M_{n\times n}(\mathbb{R})$ is said to be **nilpotent** if $A^k = 0$ for some integer $k \geq 1$.

2.2. Definition: Let A be an $n \times n$ nilpotent matrix and x be a nonzero vector in F^n.

(a) $k :=$ minimal natural number such that $A^k x = 0$.

(b) The set $\{x, Ax, ..., A^{k-1}x\}$ is called the **cycle** generated by x.

(c) The set $\text{Span}(\{x, Ax, ..., A^{k-1}x\})$ is called the **cycle subspace** generated by x and denoted $C(x)$.

(d) k is called the **length** of the cycle.

(e) x is called the **initial vector** of the cycle

2.3. Example: Let $A = \begin{bmatrix} 0 & 2 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ and $x = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}$.

We noticed that $A^4 = 0$ and $A^3 \neq 0$.

Thus, $Ax = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$, $A^2x = \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $A^3x = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

Hence, $\{A^2x, Ax, x\}$ is a cycle of length 3.

2.4. Remark: Different vectors may generate cycles of different lengths.

2.5. Proposition:

(a) $N^{k-1}(x)$ *is an eigenvector of* N *with eigenvalue* $\lambda = 0$.

(b) $C(x)$ *is an invariant subspace of* V *under* N.

(c) They cycle generated by $x \neq 0$ *is a linearly independent set. Hen* $\dim(C(x)) = k$*, the length of the cycle.*

2.6. Proposition: *Let* $\alpha_i = \{N^{k_i-1}(x_i), ..., x_i\}(1 \leq i \leq r)$ *be cycles of length* k_i*, respectively. If the set of eigenvectors* $\{N^{k_1-1}(x_1), ..., N^{k_r-1}(r_i)\}$ *is linearly independent, then* $\alpha_1 \cup ... \cup \alpha_r$ *is linear independent.*

2.7. Definition: We say that the cycles $\alpha_i = \{N^{k_i-1}(x_i), ..., x_i\}$ are **non-overlapping** cycles if $\alpha_1 \cup ... \cup \alpha_r$ is linearly independent.