

# MATH405: Linear Algebra

Michael Li

## Contents

<b>1</b>	<b>Vector Space</b>	<b>2</b>
1.1	Definitions . . . . .	2
1.2	Basis . . . . .	5
1.3	Dimension . . . . .	6
1.3.1	Toolbox Corollaries and Results . . . . .	9
1.4	Direct Sums . . . . .	9
<b>2</b>	<b>Matrices</b>	<b>10</b>
2.1	Linear Equations . . . . .	11
<b>3</b>	<b>Mappings</b>	<b>12</b>
3.1	Consequences of Properties of Linear Transformations . . . . .	13
3.2	Kernel . . . . .	14
3.2.1	Consequences of Kernel . . . . .	15
3.3	Compositions and Inverse Linear Mappings . . . . .	15
<b>4</b>	<b>Linear Maps and Matrices</b>	<b>16</b>
4.1	Bases, Matrices, and Linear Maps . . . . .	16
<b>5</b>	<b>Scalar Products and Orthogonality</b>	<b>19</b>
5.1	Scalar Products . . . . .	19
5.2	Orthonormal Basis . . . . .	22

# 1 Vector Space

Goals of this course is to discuss

- Vector spaces
- Linear transformations between vector spaces
- Other operations on vector spaces

## 1.1 Definitions

**Definition - Field:** A set of numbers containing 0, 1 that can be added, subtracted, multiplied, and divided (except cannot divide by 0) that satisfy the following **Field Axioms**

1.  $a, b \in K \implies a + b, ab \in K$
2.  $+, \times$  are commutative so  $a + b = b + a$  and  $ab = ba$
3.  $+, \times$  are associative so  $(a + b) + c = a + (b + c)$  and  $a(bc) = (ab)c$
4. Distributive Law:  $a(b + c) = ab + ac$
5. Additive Identity:  $a + 0 = 0 + a = a$
6. Multiplicative Identity:  $a \cdot 1 = 1 \cdot a = a$
7. Additive Inverse:  $\forall a \in K, \exists b$  such that  $a + b = 0$ , namely  $b = -a$  which is unique
8. Multiplicative Inverse:  $\forall a \in K, \exists b$  such that  $ab = 1$ , name  $b = 1/a$  which is unique

**Example:**  $R, Q$  are fields.  $Z$  is not a field since there is no multiplicative inverse of 2

**Example:**  $C = \{a + bi \mid a, b \in R\}$ , where  $i = \sqrt{-1}$  is a field under

- $+$  :  $(a + bi) + (c + di) = (a + c) + (b + d)i$
- $\times$  :  $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$

**Example:**  $F_2 = \{0, 1\}$  is a field under

- $+$  : where
$$0 + 0 = 0$$
$$0 + 1 = 1 + 0 = 1$$
$$1 + 1 = 0$$
- $\times$  : where
$$0 \cdot 0 = 0$$
$$0 \cdot 1 = 1 \cdot 0 = 0$$
$$1 \cdot 1 = 1$$

**Example:** For a prime  $p$ , let  $F_p = \{0, \dots, p-1\}$ . Then  $F_p$  is a field under

- $+$  :  $a + b \pmod{p}$
- $\times$  :  $ab \pmod{p}$

**Definition - Vector Space:** For an arbitrary field  $K$ , a  $K$ -vector space is a set  $V$ , with a distinguished element  $O$ , such that any 2 elements in  $V$  can be added and scalar multiplied by  $c \in K$

- $u, v \in V \implies u + v \in V$
- $c \in K, u \in V \implies cu \in V$

Satisfying the following properties

1. Commutative Addition:  $u + v = v + u$
2. Associative Addition:  $(u + v) + w = u + (v + w)$

3. Additive Identity:  $u + O = u$
4. Additive Inverse:  $\forall u \in V, \exists v \in V$  such that  $u + v = O$ , namely  $v = -u$  which is unique
5. Distributive Laws:  $\forall a, b \in K, a(u + v) = au + av$  and  $(a + b)u = au + bu$
6. Commutative Scalar Multiplication:  $(ab)u = a(bu)$
7. Multiplicative Identity:  $1 \cdot u = u$

**Example:**  $R^3$  is an  $R$ -vector space defined by the operations

$$R^3 = \{(x, y, z) \mid x, y, z \in R\}$$

- $+$  : add componentwise so  $(a, b, c) + (d, e, f) = (a + d, b + e, c + f)$
- Scalar  $\times$  : for  $r \in R, r(a, b, c) = (ra, rb, rc)$
- Additive Identity is  $O = (0, 0, 0)$

**Example:** For any field  $K, K^2$  is a  $K$ -vector space defined by the operations

$$K^2 = \{(x, y) \mid x, y \in K\}$$

- $+$  : add componentwise so  $(a, b) + (c, d) = (a + c, b + d)$
- Scalar  $\times$  : for  $k \in K, k(a, b) = (ka, kb)$
- Additive Identity is  $O = (0, 0)$

**Example:**  $R$  is an  $R$ -vector space since clearly the necessary properties hold

**Example**  $R$  is a  $Q$ -vector space since clearly the necessary properties hold

- Notably, for  $q \in Q$  and  $r \in R$ , we have  $qr \in R$ . Thus scalar multiplication is closed

**Example:** For any field  $K$ , the set  $\{O\}$  is a  $K$ -vector space

**Example:** Let  $X$  be any non-empty set and let  $\mathcal{F}(X)$  be the set of all functions  $f : X \rightarrow R$ . Then  $\mathcal{F}$  is an  $R$ -vector space under the operations

- $+$  : for  $f, g \in \mathcal{F}(X)$ , define  $f + g := (f + g)(x)$
- Scalar  $\times$  : let  $r \in R$ , then define  $rf := r(f(x))$
- Additive Identity is  $O = f(x) = 0$ , the function that takes any  $x$  to 0

**Example:** Take  $X = N$  and let  $F(X) = \{ \text{all functions } f : N \rightarrow R \}$  is a vector space

- **Note:**  $f : N \rightarrow R$  is a sequence  $(a_0, \dots, a_n)$  where  $a_n = f(n)$

**Lemma 1 - Cancellation:** For  $u, v, w \in V$  and if  $u + v = w + v$ , then  $u = w$

*Proof:*  $v \in V$  has an additive inverse, namely  $-v$ . Thus we have

$$u + v - v = w + v - v \implies u = w$$

**Lemma 2 - Unique Additive Inverse:** For all  $v \in V$ , there is a unique additive inverse, namely  $-v$

*Proof:* Suppose  $u, w$  are both additive inverses of  $v$ . Then we have

$$v + u = v + w \implies u = w$$

**Lemma 3 - 0 Times a Vector:** For all  $v \in V$ ,  $0v = O$

*Proof:*  $v = 1v = (0 + 1)v = 0v + 1v = 0v + v \implies 0v = O$

**Lemma 4 -  $(-1)v$  is the Additive Inverse:** For all  $v \in V$ ,  $(-1)v$  is the unique additive inverse of  $v$

*Proof:*  $(-1)v + v = (-1 + 1)v = 0v = O$ . Thus  $(-1)v$  is the additive inverse of  $v$ , which is unique by Lemma 2

**Definition - Subspace:** For a  $K$ -vector space  $V$  and a non-empty subset  $W \subseteq V$ ,  $W$  is a **subspace** if it satisfies

- $w_1, w_2 \in W \implies w_1 + w_2 \in W$
- $\forall a \in K, w \in W \implies aw \in W$
- $O \in W$

**Theorem 1:** Every subspace of a  $K$ -vector space is a  $K$ -vector space

*Proof:* We need to show that  $W \subseteq V$  satisfies all the necessary properties of a vector space

1. Verify  $O \in W$

Since  $W$  is non-empty and closed under scalar multiplication, take  $0w = O \in W$  by Lemma 3

2.  $u, v \in W \implies u + v \in W$  and  $a \in K, v \in W \implies av \in W$  by definition of subspace

3. Every  $w \in W$  has an additive inverse, namely  $-w$

Since  $W$  is closed under scalar multiplication,  $(-1)w = -w \in W$  by Lemma 4

4. Other conditions (associative addition, commutative addition, etc.) hold because  $u, v, w \in W \implies u, v, w \in V$

For example, choose  $u, v \in W$ , then  $u + v = v + u$ , since  $u, v \in V$ . Thus commutative addition is satisfied

**Example:** Take  $(5, 3, 2) \in R^3$ . Then let  $W = \{r(5, 3, 2) \mid r \in R\}$

Then  $W$  is an  $R$ -vector space. We prove this by showing that  $W$  is a subspace of  $R^3$

- $+$  : Choose 2 arbitrary elements of  $W$ ,  $r(5, 3, 2)$  and  $s(5, 3, 2)$  for  $r, s \in R$

Then  $r(5, 3, 2) + s(5, 3, 2) = (r + s)(5, 3, 2) \in W$

- $\times$  : Choose  $r(5, 3, 2) \in W$  and take  $s \in R$

Then  $s(r(5, 3, 2)) = (sr)(5, 3, 2) \in W$

**Example:** Let  $U = \{(x, y, z) \in R^3 \mid 2x + 3y = 0\}$ . We show that  $U$  is a vector space by showing it's a subspace of  $R^3$

- $+$  : Take  $(x_1, y_1, z_1)$  and  $(x_2, y_2, z_2) \in U \implies 2x_1 + 3y_1 = 0$  and  $2x_2 + 3y_2 = 0$

Then  $2(x_1 + x_2) + 3(y_1 + y_2) = 0$

Thus  $(x_1 + x_2, y_1 + y_2, z_1 + z_2) \in U$

- $\times$  : Let  $(x, y, z) \in U$  and  $r \in R$

Then  $2x + 3y = 0 \implies r(2x + 3y) = 2rx + 3ry = 0$

Thus  $r(x, y, z) \in U$

**Example:** Consider  $\sin(x), \cos(x) \in \mathcal{F}(R)$  and let  $W = \{a \sin(x) + b \cos(x) \mid a, b \in R\}$ . Then  $W$  is a subspace of  $\mathcal{F}(R)$

- $+$  : Take  $a_1 \sin(x) + b_1 \cos(x)$  and  $a_2 \sin(x) + b_2 \cos(x) \in W$ . Then  $(a_1 + a_2) \sin(x) + (b_1 + b_2) \cos(x) \in W$
- $\times$  : Take  $r \in R$ . Then  $r(a \sin(x) + b \cos(x)) = (ra) \sin(x) + (rb) \cos(x) \in W$

## 1.2 Basis

**Definition - Linear Combination:** For vectors  $\{v_1, \dots, v_n\} \subseteq V$ , a **linear combination** of  $\{v_1, \dots, v_n\}$  is a vector of the form

$$a_1v_1 + \dots + a_nv_n \quad a_i \in K$$

**Definition - Span:**  $\text{span}(\{v_1, \dots, v_n\}) = \{ \text{all linear combinations of } \{v_1, \dots, v_n\} \}$

**Proposition 1:**  $W = \text{span}(\{v_1, \dots, v_n\})$  is a subspace of  $V$  and thus is itself a  $K$ -Vector Space

*Proof:* We show that  $W$  satisfies the necessary criteria to be a subspace of  $V$

- $+$  : Let  $a = a_1v_1 + \dots + a_nv_n \in W$  and  $b = b_1v_1 + \dots + b_nv_n \in W$

Then  $a + b = (a_1 + b_1)v_1 + \dots + (a_n + b_n)v_n \in W$

Thus  $W$  is closed under addition

- Scalar  $\times$  : Let  $a = a_1v_1 + \dots + a_nv_n \in W$  and let  $c \in K$

Then  $ca = (ca_1)v_1 + \dots + (ca_n)v_n \in W$

Thus  $W$  is closed under scalar multiplication

**Example:** Take  $(5, 3, 1)$  and  $(4, 0, -2) \in R^3$

$\text{span}(\{(5, 3, 1), (4, 0, -2)\})$  is a plane in  $R^3$  passing through  $(0, 0, 0)$

**Example:** Take  $(5, 3, 1)$  and  $(10, 6, 2) \in R^3$

$\text{span}(\{(5, 3, 1), (10, 6, 2)\})$  is a line in  $R^3$  passing through  $(0, 0, 0)$

- **Note:**  $(10, 6, 2) = 2(5, 3, 1)$ . Thus  $\text{span}(\{(5, 3, 1), (10, 6, 2)\}) = a_1(5, 3, 1) + a_2(10, 6, 2) = (a_1 + 2a_2)(5, 3, 1)$

**Definition - Linearly Independent:**  $\{v_1, \dots, v_n\}$  is **linearly independent** if whenever  $a_1v_1 + \dots + a_nv_n = 0$ , then  $a_1 = \dots = a_n = 0$

- Otherwise  $\{v_1, \dots, v_n\}$  is **linearly dependent**

**Proposition 2:**  $\{v_1, \dots, v_n\}$  is linearly independent if and only if no  $v_i$  is a linearly combination of the other  $n - 1$  vectors

*Proof:*  $\implies$  Assume  $\{v_1, \dots, v_n\}$  is linearly independent

BWOC, assume some  $v_i = a_1v_1 + \dots + a_nv_n$  for some  $v_i \notin \{v_1, \dots, v_n\}$

Then we have

$$O = a_1v_1 + \dots + a_nv_n + (-1)v_i$$

Since  $v_i$  is a linear combination of  $\{v_1, \dots, v_n\}$ , the above equation shows that  $\{v_1, \dots, v_n\}$  is linearly dependent. Contradiction

Thus  $v_i$  cannot be written as a linear combination of the other vectors

$\Leftarrow$  Assume by way of contraposition that  $\{v_1, \dots, v_n\}$  is not linearly independent

Thus choose  $a_1, \dots, a_n \in K$ , not all 0 such that

$$a_1v_1 + \dots + a_nv_n = O$$

WLOG, assume  $a_1 \neq 0$ . Then  $v_2a_2 + \dots + a_nv_n = -a_1v_1$

Since  $a_1 \neq 0$  and  $K$  is a field, we have

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

Thus we have shown that  $v_1$  is a linear combination of the other  $n - 1$  vectors

**Corollary 3:**  $\{v_1, \dots, v_n\}$  is linearly independent if and only if for each  $i$ ,  $v_i \notin \text{span}(\{v_1, \dots, v_n\} \setminus \{v_i\})$

*Proof:* This follows from the previous proposition

**Definition - Spans:** Let  $W$  be a  $K$ -Vector Space and  $\{v_1, \dots, v_n\} \subseteq W$ . If  $\text{span}(\{v_1, \dots, v_n\}) = W$ , then  $\{v_1, \dots, v_n\}$  **spans**  $W$ , so every  $w \in W$  is a linear combination of  $\{v_1, \dots, v_n\}$

**Definition - Basis:**  $\{v_1, \dots, v_n\}$  is a **basis** of  $W$  if it spans  $W$  and is linearly independent

**Example:**  $\{(5, 3, 1), (4, 0, -2)\}$  is a basis for  $\text{span}(\{(5, 3, 1), (4, 0, -2)\})$

**Example:**  $\{(5, 3, 1), (10, 6, 2)\}$  is not a basis for  $\text{span}(\{(5, 3, 1), (10, 6, 2)\})$  since it is not linearly independent

**Proposition 4:** Let  $\{v_1, \dots, v_n\}$  be a basis for  $W$  and let  $w \in W$  be arbitrary. Then  $w$  can be written uniquely as

$$w = a_1v_1 + \dots + a_nv_n \quad a_i \in K$$

*Proof:* Since  $\{v_1, \dots, v_n\}$  spans  $W$ , every  $w \in W$  is a linear combination of  $\{v_1, \dots, v_n\}$

For uniqueness, suppose

$$w = a_1v_1 + \dots + a_nv_n = b_1v_1 + \dots + b_nv_n$$

Then we have

$$0 = (b_1 - a_1)v_1 + \dots + (b_n - a_n)v_n$$

Since  $\{v_1, \dots, v_n\}$  is linearly independent, we must have  $b_i - a_i = 0$ , and thus  $b_i = a_i$  for each  $i$

Thus each  $w \in W$  can be written uniquely as a linear combination of  $\{v_1, \dots, v_n\}$

**Example:** Let  $W = \text{span}(\{\sin(x), \cos(x)\}) = \{a \sin(x) + b \cos(x) \mid a, b \in R\}$

We know that  $W$  is an  $R$ -Vector Space

$\{\sin(x), \cos(x)\}$  is linearly independent. Otherwise  $\sin(x) = r \cos(x)$  for all  $x \in X$  and some  $r \in R$ . However, this cannot hold for when  $x = \pi/2$  since  $\sin(\pi/2) = 1 \neq r \cos(\pi/2) = r \cdot 0$

### 1.3 Dimension

Let  $\{v_1, \dots, v_n\} \subseteq V$  and let  $W = \text{span}(\{v_1, \dots, v_n\})$

Now let  $X = \{w_1, \dots, w_m\} \subseteq W$ . Then there are 2 desirable properties of  $X$

- **X is Big:**  $X$  spans  $W$  if  $\text{span}(X) = W$ , i.e. all  $w \in W$  is a linear combination of elements from  $X$
- **X is Small:**  $X$  is linearly independent, i.e. no element in  $X$  is a linear combination of the remaining elements

**Note:** the empty set  $\emptyset$  is linearly independent since no element in  $\emptyset$  is a linear combination of the others. More notably,  $\emptyset$  is a basis for  $\{0\}$

**Shrinking Lemma:** Let  $X = \{w_1, \dots, w_m\} \subseteq W$  and spans  $W$  but  $X$  is not linearly independent. Then  $X \setminus \{w_i\}$  still spans  $W$  for some  $w_i \in X$

*Proof:* Since  $X$  is not linearly independent, we know that some  $w_i$  is a linear combination of elements in  $X \setminus \{w_i\}$ . Suppose

$$w_i = a_1w_1 + \dots + a_mw_m \quad \text{without } w_i \text{ occurring}$$

Then take arbitrary  $u \in W$  where

$$u = b_1w_1 + \dots + b_mw_m$$

Replacing  $w_i$  above with the previous equation, we see that  $u$  is a linear combination of  $X \setminus \{w_i\}$

Thus  $X \setminus \{w_i\} = \text{span}(W)$

**Shrinking Theorem:** Let  $X = \{w_1, \dots, w_m\}$  span  $W$ . Then for some subset  $Y \subseteq X$  is a basis of  $W$

*Proof:*

Case 0: If  $X$  is linearly independent, then  $X$  is a basis by definition

Otherwise, apply the shrinking lemma to get  $X_1 = X \setminus \{w_i\}$ , which spans  $W$

Case 1: If  $X_1$  is linearly independent, then  $X_1$  is a basis

...

Since  $X$  is finite (it has  $m$  elements), we will stop eventually. Either

- Some  $X_i$  is linearly independent. Thus  $X_i$  is a basis for  $W$
- Otherwise if we hit case m:  $X_m = \emptyset$ , which is linearly independent, and thus  $X_m$  spans  $W = \{O\}$

**Corollary:** If  $W = \text{span}(\{v_1, \dots, v_n\})$ , then some subset of  $\{v_1, \dots, v_n\}$  is a basis

- **Note:** In particular,  $W$  has to have a basis

**Enlarging Lemma:** Suppose  $X = \{w_1, \dots, w_m\} \subseteq W$  and is linearly independent but doesn't span  $W$ . Then for any  $w \in W \setminus \text{span}(X)$ ,  $X \cup \{w\}$  is still linearly independent

*Proof:* Suppose  $a_1 w_1 + \dots + a_m w_m + b w = O$ . We show that  $a_1 = \dots = a_m = b = 0$

Suppose BWOC,  $b \neq 0$ , then we can solve for  $w$

$$w = \frac{-a_1}{b} w_1 + \dots + \frac{-a_m}{b} w_m$$

Which means that  $w \in \text{span}(X)$ . Contradiction

Thus  $b = 0$ . This gives

$$a_1 w_1 + \dots + a_m w_m + 0w = O$$

Since  $X = \{w_1, \dots, w_m\}$  is linearly independent, we also have  $a_1 = \dots = a_m = 0$

Thus  $X \cup \{w\}$  is linearly independent

**Main Question:** does the enlarging process above terminate? After some steps, do we get a set  $\{w_1, \dots, w_m\}$  that spans  $W$ ?

**Exchanging Lemma:** Let  $X = \{v_1, \dots, v_n\}$  be any basis for  $W$ . Choose any  $w \in W$  but  $w \notin \text{span}(\{v_k, \dots, v_n\})$ . Then  $\exists v_i, i < k$ , such that  $Y = (X \setminus \{v_i\}) \cup \{w\}$  is still a basis

- **Note:** If  $k > n$ , then  $\{v_k, \dots, v_n\} = \emptyset$

*Proof:* First we show that  $\text{span}(Y) = W$ . Since  $X$  spans  $W$ , we can write

$$w = a_1 v_1 + \dots + a_n v_n \implies v_1 = \frac{1}{a_1} w + \frac{-a_2}{a_1} v_2 + \dots + \frac{-a_n}{a_1} v_n$$

Since  $w \notin \text{span}(\{v_k, \dots, v_n\})$ , we must have  $a_i \neq 0$  for some  $i < k$

WLOG, let  $a_1 \neq 0$ . We show that  $Y$  spans  $W$

Since  $X$  spans  $W$ , for arbitrary  $u \in W$ , we have

$$u = d_1 v_1 + \dots + d_n v_n$$

Replacing  $v_1$  above with the previous equation, we see that  $u$  is a linear combination of elements of  $Y$  and thus  $u \in \text{span}(Y)$

Thus  $\text{span}(Y) = W$

Next we show that  $Y$  is linearly independent

Suppose we have

$$cw + b_2v_2 + b_nv_n = O$$

We show that  $c = b_2 = \dots = b_n = 0$

- If  $c = 0 \implies b_2 = \dots = b_n = 0$  since  $\{b_2, \dots, b_n\}$  is linearly independent
- Otherwise suppose  $c \neq 0$ , then we can solve for  $w$

$$w = \frac{-b_2}{c}v_2 + \dots + \frac{-b_n}{c}v_n \implies v_1 = \frac{1}{a_1}\left(\frac{-b_2}{c}v_2 + \dots + \frac{-b_n}{c}v_n\right) + \frac{-a_1}{a_1}v_2 + \dots + \frac{-a_m}{a_1}v_m$$

Thus  $v_1$  is a linear combination of  $\{v_2, \dots, v_n\}$ . Contradiction since we said  $X$  was linearly independent. Thus  $c = 0$

**Theorem:** Let  $X = \{v_1, \dots, v_n\}$  be a basis for  $W$ , and let  $\{w_1, \dots, w_m\} \subseteq W$  be linearly independent. Then  $m \leq n$

*Proof:* If  $m < n$ , we are done

Now assume  $m \geq n$ , we show that  $m = n$

Since  $\{w_1, \dots, w_m\}$  is linearly independent, we have that  $w_1 \neq O = \text{span}(\emptyset)$

Now apply the Exchanging Lemma to the basis  $X$ , with  $k > n$  and  $w_1$ . Then  $\exists v_i$  such that  $X_1 = (X \setminus \{v_i\}) \cup \{w_1\}$  is a basis

After reindexing, we see that  $X_1$  has  $n - 1$  vectors from  $X$  and 1 vector from  $w_1$

Now take  $k = n$ . Since  $\{w_1, \dots, w_m\}$  is linearly independent,  $w_2 \notin \text{span}(\{w_1\})$

Thus applying the Exchanging Lemma again, there exists  $j < k = n$  such that  $X_2 = (X_1 \setminus \{v_j\}) \cup \{w_2\}$  is a basis

Reindexing again, we get that  $X_2 = \{v_1, \dots, v_{n-2}, w_1, w_2\}$  is a basis

After  $n$  steps,  $X_n$  has no elements from  $X$  and  $X_n = \{w_1, \dots, w_n\}$  is a basis

Furthermore, we see that  $w_m \in \text{span}(\{w_1, \dots, w_n\})$ , contradicting that  $\{w_1, \dots, w_m\}$  is linearly independent

Thus  $m = n$

**Corollary:** If  $W$  is any  $K$ -vector space and some basis of  $W$  has  $n$  elements, then every basis of  $W$  has  $n$  elements

**Definition - Finite Dimensional:** Let  $W$  be a  $K$ -vector space. Then  $W$  is **finite dimensional** if some basis for  $W$  is finite

**Definition - Dimension:** Number of elements in any basis for a vector space  $W$

**Corollary:** Suppose  $\dim(W) = n$  and  $X = \{w_1, \dots, w_n\}$  are any  $n$ -vectors

1. If  $X$  spans  $W$ , then  $X$  is a basis for  $W$
2. If  $X$  is linearly independent, then  $X$  is a basis for  $W$

*Proof:*

1. By Shrinking Theorem, there exists a basis  $Y \subseteq X$

However,  $|Y| < n$  contradicts that  $\dim(W) = n$

Thus  $Y = X$ , i.e.  $X$  is a basis

2. By Enlarging Lemma, we can expand  $X$  to a basis  $Y$

However,  $|Y| > n$  contradicts that  $\dim(W) = n$

Thus  $Y = X$ , i.e.  $X$  is a basis



### 1.3.1 Toolbox Corollaries and Results

The following are useful corollaries that can be used to prove additional interesting results

Let  $V$  be a  $K$ -Vector Space with  $\dim(V) = n$ , i.e.  $V$  has some basis with  $n$  elements

1. Every basis for  $V$  has  $n$  elements
2. If  $X \supseteq V$  and  $\text{span}(X) = V$ , then  $X$  has at least  $n$  elements and some subset  $Y \subseteq X$  is a basis for  $V$
3. If  $Z \subseteq V$  is linearly independent, then  $Z$  has at most  $n$  elements and  $Z$  can be extended to a basis  $Y \supseteq Z$  for  $V$

**Example:** Let  $V = R^3$ . Since  $\dim(V) = 3$ ,  $V$  has a basis with 3 elements

- Consider the **Standard Basis:**  $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$

Suppose  $X = \{v_1, v_2, v_3\} \subseteq V$  for arbitrary vectors

- If  $\text{span}(X) = V$  then  $X$  is a basis
- If  $X$  is linearly independent, since  $|X| = 3$ ,  $X$  is a basis for  $V$

**Example:** Describe all subspaces  $W \subseteq R^3$

**Note:** Since  $\dim(V) = 3$ , we must have  $\dim(W) \leq \dim(V) = 3$

- Case 0:  $\dim(W) = 0$

Clearly  $W = \{O\}$

- Case 1:  $\dim(W) = 1$

$W$  is a line going through  $(0, 0, 0)$

Thus a basis for  $W$  will be  $\{w\}$  for any nonzero  $w \in W$

- Case 2:  $\dim(W) = 2$

$W$  is a plane containing  $(0, 0, 0)$

Thus a basis for  $W$  will be any 2 element set  $\{w_1, w_2\} \subseteq W$  such that

- Neither element is  $O$
- $w_2$  is not a scalar multiple of  $w_1$

- Case 3:  $\dim(W) = 3$

Only possibility is  $W = V = R^3$

**Examples:** Consider subspaces of  $\mathcal{F}(R)$  and look at small subspaces

- $W = \text{span}(\{e^x\}) = \{re^x \mid r \in R\}$

This can be thought of as a 1-dimensional subspace of  $\mathcal{F}(R)$

- $V = \text{span}(\{\sin(x), \cos(x)\}) = \{a \sin(x) + b \cos(x) \mid a, b \in R\}$

Clearly  $\dim(V) = 2$

Consider  $f(x) = \sin(x)$   $g(x) = \cos(x)$   $h(x) = 3 \sin(x) - 2 \cos(x)$

Since  $h = 3f + (-2)g$ ,  $\{f, g, h\}$  is not linearly independent

Thus  $\text{span}(\{f, g, h\}) = \text{span}(\{f, g\})$

## 1.4 Direct Sums

Let  $V$  be a  $K$ -Vector Space with  $\dim(V) = n$ . Let  $W \subseteq V$  be a subspace of  $V$ . Then  $\dim(W) \leq n$

Now choose another subspace  $U \subseteq V$

**Note:**  $W \cap U \neq \emptyset$  since both must contain  $O$

Thus the smallest we can make  $W \cap U$  is  $\{O\}$

Furthermore, it can be shown that both  $U \cap W$  and  $U + W$  are both subspaces of  $V$

**Definition - Direct Sum:**  $U \oplus W$  is called a **direct sum** if

- $U \oplus W = U + W$
- $U \cap W = \{O\}$

We often look at cases where  $V = U \oplus W$

**Example:** Consider  $R^3$  and let  $W$  be any plane containing  $(0, 0, 0)$

If  $U$  is any line through  $(0, 0, 0)$  such that  $U \not\subseteq W$ , then  $R^3 = W \oplus U$

**Theorem:** Let  $V$  be a  $K$ -Vector Space with  $\dim(V) = n$ . Let  $W \subseteq V$  be any subspace of  $V$ . Then there exists a subspace  $U \subseteq V$  such that

$$V = U \oplus W$$

*Proof:* Choose any basis  $Z = \{w_1, \dots, w_m\}$  of  $W$  (we know that  $m \leq n$ )

Now extend  $Z$  to  $Y = Z \cup \{u_1, \dots, u_r\}$ , which is a basis for  $V$

Let  $U = \text{span}(\{u_1, \dots, u_r\})$ . Then  $U$  is a subspace of  $V$  and  $\{u_1, \dots, u_r\}$  is a basis for  $U$

- Show that  $U \cap W = \{O\}$

Choose  $v \in U \cap W$

Then we have  $v = a_1 u_1 + \dots + a_r u_r = b_1 w_1 + \dots + b_m w_m$

Since  $Y$  is a basis for  $V$ , then  $\{u_1, \dots, u_r, w_1, \dots, w_m\}$  is linearly independent

Thus  $v - v = a_1 u_1 + \dots + a_r u_r - b_1 w_1 - \dots - b_m w_m = O \implies a_1 = \dots = a_r = b_1 = \dots = b_m = 0$

Thus  $v = O$

- Show that  $V = U + W$

Choose any  $v \in V$

Since  $Y$  is a basis for  $V$

$$v = \underbrace{a_1 u_1 + \dots + a_r u_r}_{u \in U} + \underbrace{b_1 w_1 + \dots + b_m w_m}_{w \in W}$$

Thus  $v = u + w \implies V = U + W$

## 2 Matrices

**Definition -  $m \times n$  Matrix:** Entries  $\in K$  of the form

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

**Example:**  $A = \begin{bmatrix} 4 & 0 & 2 \\ -1 & 3 & 6 \end{bmatrix}$  is a  $2 \times 3$  matrix with entries  $\in Q$

**Note:** Any  $2 \times 3$  matrices can be added together componentwise or multiplied by a scalar, resulting in a  $2 \times 3$  matrix

- Here the additive identity is  $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

- Here the additive inverse of  $A$  (from previous example) is  $-A = \begin{bmatrix} -4 & 0 & -2 \\ 1 & -3 & -6 \end{bmatrix}$

Thus  $\text{Mat}_{2 \times 3}(K)$ , the set of all  $2 \times 3$  matrices with entries in  $K$  is a  $K$ -Vector Space

Here the basis is  $B = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}$

- Clearly spans since any  $2 \times 3$  matrix  $\begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \end{bmatrix}$  can be written as a linear combination of elements in  $B$
- Clearly  $B$  is linearly independent since the only way to write  $O$  is to take each scalar  $a_i = 0$

Thus  $\dim(\text{Mat}_{2 \times 3}(K)) = 6$

**Upshot:** We can generalize the discussion above to show that  $\text{Mat}_{m \times n}(K)$  is a  $K$ -Vector Space of  $\dim = m \times n$

**Example:**  $\left\{ \begin{bmatrix} a & b \\ b & d \end{bmatrix} \right\}$ , **symmetric  $2 \times 2$  matrices**, is a subspace of  $\text{Mat}_{2 \times 2}(K)$ , which has dimension 4

**Non-Example:**  $\text{Mat}(K)$  is NOT a Vector Space since addition between  $2 \times 2$  and  $3 \times 3$  matrices is not defined

**Notation:**  $A_i = (a_{i1}, \dots, a_{in})$ , the  $i$ th row vector, is a  $1 \times n$  matrix

**Notation:**  $A^j = (a_{1j}, \dots, a_{mj})$ , the  $j$ th column vector, is a  $m \times 1$  matrix

**Definition - Transpose:** Given an  $m \times n$  matrix  $A$ , the **transpose**  ${}^tA$  is an  $n \times m$  matrix that swaps the rows and columns, and vice versa

- **Note:** If  $A$  is a square  $n \times n$  matrix, then  ${}^tA$  is also a square  $n \times n$  matrix

**Example:**  ${}^t \begin{bmatrix} 4 & 0 & 3 \\ -1 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ 0 & 3 \\ 2 & 6 \end{bmatrix}$

**Definition - Matrix Multiplication:** An  $m \times n$  matrix  $A$  can multiply with an  $n \times k$  matrix  $B$  where

$$C_{il} = \sum_{d=1}^n a_{id} b_{d,l}$$

- **Note:** If  $A, B$  are both  $n \times n$  matrices, then  $AB$  is an  $n \times n$  matrix
- **Upshot:** Square matrices are closed under transposition and matrix multiplication

**Example:**  $\begin{bmatrix} 2 & 3 & 4 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 22 \\ 4 \end{bmatrix}$

## 2.1 Linear Equations

Consider

$$\begin{aligned} 5x_1 + 3x_2 - 6x_3 &= 8 \\ x_1 - 2x_2 + x_3 &= 4 \end{aligned}$$

We can represent this using

$$A = \begin{bmatrix} 5 & 3 & -6 \\ 1 & -2 & 1 \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad B = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \implies AX = B$$

### 3 Mappings

**Definition - Function:** Mapping between 2 sets  $D, R$  such that for each  $x \in D$ , there exists a unique  $y \in R$  such that  $f(x) = y$

$$F : D \rightarrow R$$

- **Note:**  $D$  here is the **domain** of  $F$  and  $R$  is the **range** of  $F$

**Definition - Image:**  $F(D) = \{F(x) \mid x \in D\} \subseteq R$

**Example:**  $F : R \rightarrow R \quad F(x) = x^2$

- $\text{Domain}(F) = \text{Range}(F) = R$
- Image of  $F = \{y \in R \mid y \geq 0\} = [0, \infty)$

**Example:**  $G[0, \infty) \rightarrow R \quad G(x) = \sqrt{x}$

- Image of  $G = [0, \infty)$

**Example:**  $\mathcal{F} = \text{all functions } F : \mathbb{R} \rightarrow \mathbb{R}$

Let  $S$  be all “infinitely” differentiable functions

Let  $\frac{d}{dx} : S \rightarrow S$  where  $\frac{d}{dx}(f) = f'$

Thus  $\frac{d}{dx}$  is a function

**Example:**  $t : \text{Mat}_{2 \times 3}(K) \rightarrow \text{Mat}_{3 \times 2}(K)$

Then  $t(A) = {}^t A$  is a function

**Definition - Onto:** A function  $F : D \rightarrow R$  is **onto** if Image of  $F = R$

**Definition - 1-1:** A function  $F : D \rightarrow R$  is **1-1** if different elements from  $D$  get mapped to different elements of  $R$

$$F(d) = F(e) \implies d = e$$

**Definition - Bijection:** A function that is both onto and 1-1

**Definition - Inverse Function:** If  $F : D \rightarrow R$  is a bijection, there exists an inverse function  $F^{-1} : R \rightarrow D$  such that

$$\forall r, \in R, F(F^{-1}(r)) = r$$

$$\forall d, \in D, F^{-1}(F(d)) = d$$

**Definition - Linear Transformation:** For fixed  $K$ -Vector Spaces  $V, W$ , a **linear transformation**  $T : V \rightarrow W$  is a function satisfying

1.  $\forall v_1, v_2 \in W, T(v_1 + v_2) = T(v_1) + T(v_2)$
2.  $\forall c \in K, v \in W, T(cv) = cT(v)$

**Examples**

1.  $F : R \rightarrow R, F(x) = x^2$ 
  - Not onto since  $x^2$  cannot be negative

- Not 1-1 since  $1^2 = (-1)^2 = 1$
  - Not a linear transformation since  $(1 + 2)^2 = 9 \neq 1^2 + 2^2$
2.  $F : [0, \infty) \rightarrow R, F(x) = \sqrt{x}$
- Not onto since  $x^2$  cannot be negative
  - 1-1 since  $\sqrt{x} = \sqrt{y} \implies x = y$
  - Not a linear transformation since  $[0, \infty)$  isn't a Vector Space
3. Let  $S$  be the set of all infinite differentiable functions. Consider  $\frac{d}{dx} : S \rightarrow S$  where  $\frac{d}{dx}(f) = f'$
- Onto by the Fundamental Theorem of Calculus
  - Not 1-1 since  $f$  and  $f + 5$  share the same derivative
  - Is a linear transformation by addition and scalar multiplication properties of derivatives
4. Let  $C$  be the set of continuous functions on  $[0, 1]$ . Consider  $I : C \rightarrow R, I(f) = \int_0^1 f(t) dt$
- Onto since we can generate any value of  $R$  by taking the integral of the constant function
  - Not 1-1 since the definite integral of 2 functions could yield the same result
  - Is a linear transformation by additional and scalar multiplication properties of integrals
5.  $I^* : G \rightarrow C, I^*(f) = \int_0^x f(t) dt$
- Not onto since not all functions of  $f(0) = 0$
  - 1-1 since indefinite integral yields a unique function
  - Is a linear transformation by additional and scalar multiplication properties of integrals
6. Fix  $(4, 0, 2)$  and consider  $T_{(4,0,2)} : R^3 \rightarrow R^3, T_{(4,0,2)}((x, y, z)) = (x + 4, y, z + 2)$
- Clearly onto
  - Clearly 1-1
  - Not a linear transformation since  $T_{(4,0,2)}((0, 0, 0) + (1, 1, 1)) = (5, 0, 3) \neq T_{(4,0,2)}((0, 0, 0)) + T_{(4,0,2)}((1, 1, 1))$
7.  $E_\pi : R^3 \rightarrow R^3, E_\pi((x, y, z)) = (\pi x, \pi y, \pi z)$
- Clearly onto
  - Clearly 1-1
  - Is a linear transformation since  $E_\pi((a, b, c) + (d, e, f)) = (\pi(a + d), \pi(b + e), \pi(c + f)) = E_\pi((a, b, c)) + E_\pi((d, e, f))$

### 3.1 Consequences of Properties of Linear Transformations

**Proposition:** For any linear transformation  $T : V \rightarrow W$ , we have that

$$T(O_V) = O_W$$

*Proof:* Let  $w = T(O_V)$

Since  $O_V = 0 * O_V$ , we have that

$$T(O_V) = T(0 * O_V) = 0 * T(O_V) = 0 * w = O_W$$

**Proposition:**  $T(a_1 v_1 + \dots + a_n v_n) = a_1 T(v_1) + \dots + a_n T(v_n)$

*Proof:* Follows from linearly properties of linear transformations

- **Note:** If  $x = \{v_1, \dots, v_n\}$  is a basis for  $V$  and if  $w_1, \dots, w_n$  are arbitrary vectors in  $W$ , then there is a unique linear transformation  $T : V \rightarrow W$  such that

$$T(v_1) = w_1, \dots, T(v_n) = w_n$$

**Lemma:**  $\text{Im}(T)$  is a subspace of  $W$

*Proof:* We show the necessary conditions for a subspace

- $+$  :  $w_1, w_2 \in \text{Im}(T) \implies \exists v_1, v_2 \in V$  such that  $T(v_1) = w_1$  and  $T(v_2) = w_2$

$$\text{Then } w_1 + w_2 = T(v_1) + T(v_2) = T(\underbrace{v_1 + v_2}_{\in V}) \in \text{Im}(T)$$

- $\times : w \in \text{Im}(T) \implies \exists v \in V$  such that  $T(v) = w$

Then for  $c \in K$ , we have  $cw = c(Tv) = T(\underbrace{cv}_{\in V}) \in \text{Im}(T)$

**Definition - Pull Back:** Suppose  $Y = \{w_1, \dots, w_m\} \subseteq \text{Im}(T)$ . Then a **pull-back** is any set  $\{v_1, \dots, v_m\} \subseteq V$  such that

$$T(v_1) = w_1, \dots, T(v_m) = w_m$$

**Lemma:** If  $\{w_1, \dots, w_m\}$  is linearly independent in  $\text{Im}(T)$  (or in  $W$ ), then any pull back  $\{v_1, \dots, v_m\} \subseteq V$  is linearly independent in  $V$

*Proof:* Let  $a_1v_1 + \dots + a_mv_m = O_V$

Thus  $T(a_1v_1 + \dots + a_mv_m) = O_V = a_1T(v_1) + \dots + a_mT(v_m) = O_W$

Since  $\{w_1, \dots, w_m\}$  is linearly independent, we have  $a_1 = \dots = a_m = 0$  as desired

**Pull Back Property:** Suppose  $\{w_1, \dots, w_m\}$  is a basis for  $\text{Im}(T)$ , and let  $\{v_1, \dots, v_m\} \subseteq V$  be any pull back. Furthermore, let  $S = \text{span}(\{v_1, \dots, v_m\}) \subseteq V$  be a subspace. Then  $\{v_1, \dots, v_m\}$  is a basis for  $S$

*Proof:* By the previous lemma,  $\{v_1, \dots, v_m\}$  is linearly independent

Furthermore,  $\{v_1, \dots, v_m\}$  spans  $S$  by definition

**Corollary:** If  $T : V \rightarrow W$  is any linear transformation and if  $\dim(V) = n$ , then  $\dim(\text{Im}(T)) \leq n$

*Proof:* BWOC, suppose  $\dim(\text{Im}(T)) > n$ , thus we can create a set of  $n + 1$  linearly independent elements in  $\text{Im}(T)$ .

By the Pull Back Property, this pulls back to  $n + 1$  linearly independent elements in  $V$ . Contradiction since  $n + 1 > n = \dim(V)$

**Note:**  $T : V \rightarrow W$  where  $T(v) = \{O_W\}$  is a linearly transformation with  $\dim(\text{Im}(T)) = 0$ , regardless of the value of  $\dim(V)$

## 3.2 Kernel

**Definition - Kernel:** For  $T : V \rightarrow W$ , the **kernel**  $\text{Ker}(T) = \{v \in V \mid T(v) = O_W\}$

**Proposition:**  $\text{Ker}(T)$  is a subspace of  $V$

*Proof:* Clearly  $O_V \in \text{Ker}(T)$

- $+$  : For  $v_1, v_2 \in \text{Ker}(T)$ , we see that  $T(v_1 + v_2) = T(v_1) + T(v_2) = O_W + O_W = O_W$ . Thus  $v_1 + v_2 \in \text{Ker}(T)$
- $\times$  : For  $c \in K$  and  $v \in \text{Ker}(T)$ , we see that  $T(cv) = cT(v) = O_W$ . Thus  $cv \in \text{Ker}(T)$

**Proposition:** Let  $T : V \rightarrow W$  be any linear transformation. For any basis  $B = \{w_1, \dots, w_m\} \subseteq \text{Im}(T)$  and for any pullback  $\{v_1, \dots, v_m\} \subseteq V$ , we have

$$V = \text{Ker}(T) \oplus S \quad S = \text{span}(\{v_1, \dots, v_m\})$$

*Proof:* We need to show  $V = \text{Ker}(T) + S$  and  $\text{Ker}(T) \cap S = \{O_V\}$

- Take arbitrary  $v \in V \implies T(v) \in \text{Im}(T) = a_1w_1 + \dots + a_mw_m$

Let  $s = a_1v_1 + \dots + a_mv_m \in S$ .

Then  $T(s) = T(v) \implies T(v - s) = T(v) - T(s) = O_W \implies v - s \in \text{Ker}(T)$

Let  $u = v - s \in \text{Ker}(T)$

Thus clearly  $v = u + s$  for  $u \in \text{Ker}(T)$  and  $s \in S$

- Clearly  $O_V \in \text{Ker}(T) \cap S$  since both are subspaces of  $V$

Take any arbitrary  $v \in \text{Ker}(T) \cap S$

$$v \in S \implies v = b_1 v_1 + \cdots + b_m v_m \implies T(v) = b_1 w_1 + \cdots + b_m w_m$$

Since  $v \in \text{Ker}(T)$ , we have that  $T(v) = O_W \implies b_1 = \cdots = b_m = 0$  since  $\{w_1, \dots, w_m\}$  is linearly independent

$$\text{Thus we have } v = 0v_1 + \cdots + 0v_m = O_V \implies \text{Ker}(T) \cap S = \{O_V\}$$

Thus we have shown the necessary properties for  $V = \text{Ker}(T) \oplus S$

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

*Proof:* Choose a basis  $B = \{w_1, \dots, w_m\}$  for  $\text{Im}(T)$  and a pullback  $\{v_1, \dots, v_m\}$

Let  $S = \text{span}(\{v_1, \dots, v_m\})$

Since  $V = \text{Ker}(T) \oplus S$ , we have  $\dim(\text{Ker}(T)) + \dim(S) = \dim(\text{Ker}(T)) + \dim(\text{Im}(T)) = \dim(V)$

### 3.2.1 Consequences of Kernel

**Corollary 1:** For linear  $T : R^3 \rightarrow R^4$ ,  $T$  is NOT onto

*Proof:*  $\dim(\text{Im}(T)) \leq \dim(R^3) = 3 < 4 \implies \text{Im}(T) \neq R^4 \implies T$  is NOT onto

**Corollary 2:** For linear  $T : R^4 \rightarrow R^3$ ,  $T$  is NOT 1-1

*Proof:*  $\dim(\text{Ker}(T)) + \underbrace{\dim(\text{Im}(T))}_{\leq 3} = \dim(R^4) = 4 \implies \dim(\text{Ker}(T)) \geq 1$

Thus  $\text{Ker}(T)$  has something non-zero mapped to  $O_W \implies T$  is NOT 1-1

**Definition - Isomorphism:**  $T : V \rightarrow W$  such that  $T$  is linear transformation and a bijection

**Corollary 3:**  $\dim(V) = \dim(W)$  and  $T : V \rightarrow W$  is a linear transformation and 1-1  $\implies T$  is an isomorphism (i.e.  $T$  is onto)

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

But we know that  $\dim(\text{Ker}(T)) = 0 \implies \dim(\text{Im}(T)) = \dim(V) = \dim(W)$

Furthermore  $\text{Im}(T)$  is a subspace of  $W$  and  $\dim(\text{Im}(T)) = \dim(W) \implies T$  is onto

**Corollary 4:**  $\dim(V) = \dim(W)$  and  $T : V \rightarrow W$  is a linear transformation and onto  $\implies T$  is an isomorphism (i.e.  $T$  is 1-1)

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

But we know that  $\dim(\text{Im}(T)) = \dim(V) \implies \dim(\text{Ker}(T)) = 0$

### 3.3 Compositions and Inverse Linear Mappings

Consider Vector Spaces  $U, V, W$  and linear transformations  $T : U \rightarrow V$  and  $S : V \rightarrow W$

**Proposition:**  $S \circ T : V \rightarrow W$  is a linear transformation

*Proof:*

- $+$  : For  $u_1, u_2 \in U$  we have that

$$\begin{aligned}
S \circ T(u_1 + u_2) &= S(T(u_1 + u_2)) \\
&= S(T(u_1) + T(u_2)) \\
&= S(T(u_1)) + S(T(u_2)) \\
&= S \circ T(u_1) + S \circ T(u_2)
\end{aligned}$$

- $\times$  : For  $u \in U$  and  $c \in K$

$$\begin{aligned}
S \circ T(cu) &= S(T(cu)) \\
&= S(cT(u)) \\
&= cS(T(u)) \\
&= cS \circ T(u)
\end{aligned}$$

Thus  $S \circ T : V \rightarrow W$  is a linear transformation

**Definition - Inverse Mapping:**  $T^{-1} : W \rightarrow V$  where  $T^{-1}(w) =$  the unique  $v \in V$  such that  $T(v) = w$

**Proposition:**  $T^{-1} : W \rightarrow V$  is a linear transformation (and thus an isomorphism)

*Proof:*

- $+$  : Take  $w_1, w_2 \in W$  such that  $T(v_1) = w_1$  and  $T(v_2) = w_2$  for  $v_1, v_2 \in V$ . Then we see that

$$T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2$$

However, by definition of inverse mapping,  $v_1 + v_2$  is the unique element such that  $T(v_1 + v_2) = w_1 + w_2$

Thus by definition of  $T^{-1}$ , we have that  $T^{-1}(w_1 + w_2) = v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2)$

- $\times$  : Similar

## 4 Linear Maps and Matrices

**Definition -  $L_A$ :** For a  $m \times n$  matrix  $A$ ,  $L_A$  determines a linear transformation from  $R^n \rightarrow R^m$

**Example:** Consider  $L_A : R^3 \rightarrow R^2$  where  $A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 2 \end{bmatrix}$

$$\text{Then we see that } A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x + 3y + z \\ 4y + 2z \end{bmatrix}$$

It can be clearly shown that  $L_A$  is a linear transformation (follows from logic of dot products)

### 4.1 Bases, Matrices, and Linear Maps

For a given transformation  $T : V \rightarrow W$ , the matrix of  $T$  with respect to the standard basis is given by

$$A = (T(E_1), \dots, T(E_n))$$

**Example:**  $T : R^2 \rightarrow R^3$   $T(x, y) = (5x + y, x - y, x)$

$$T(E_1) = (5, 1, 1) \quad T(E_2) = (1, -1, 0)$$

$$\text{Thus we see that } A = \begin{bmatrix} 5 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix}$$



- $T({}^t(3, 2)) = \begin{bmatrix} 5 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = {}^t(17, 1, 3)$

**Example:**  $T : R^2 \rightarrow R^2$  where we stretch the  $x$ -coordinate by 2

$$T({}^t(1, 0)) = {}^t(2, 0) \quad T({}^t(0, 1)) = {}^t(0, 1)$$

Thus we see that  $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

**Example:**  $S \circ T : R^2 \rightarrow R^2$  where we first stretch by  $x$  by 3 then stretch  $y$  by 3

$$T({}^t(1, 0)) = {}^t(2, 0) \quad T({}^t(0, 1)) = {}^t(0, 3)$$

Thus we see that  $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

- **Note:** we see that applying functions just corresponds to matrix multiplication  $\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

**Example:** Fix  $\theta \in R$ , then rotate by  $\theta$

$$R_\theta({}^t(1, 0)) = {}^t(\cos(\theta), \sin(\theta)) \quad R_\theta({}^t(0, 1)) = {}^t(-\sin(\theta), \cos(\theta))$$

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

Thus given any  ${}^t(x, y) \in R^2$ , we see that  $T_\theta({}^t(x, y)) = \begin{bmatrix} x \cos(\theta) - y \sin(\theta) \\ x \sin(\theta) + y \cos(\theta) \end{bmatrix}$

**Example:** Stretch  $x$  by 2, rotate by  $\pi/4$ , and stretch  $y$  by 3

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

**Note:** Given  $T : K^n \rightarrow K^m$ , the matrix  $A$  for  $T$  depends on our choosing of bases for  $K^n$  and  $K^m$

**Example:**  $T : R^2 \rightarrow R^3$   $T(x, y) = (5x + y, x - y, x)$

Let  $B = \{\underbrace{(1, 4)}_{v_1}, \underbrace{(3, 0)}_{v_2}\}$  be a basis for  $R^2$  and  $B' = \{\underbrace{(3, 0, 0)}_{w_1}, \underbrace{(0, 5, 0)}_{w_2}, \underbrace{(0, 0, 1)}_{w_3}\}$  be a basis for  $R^3$

We can define a matrix of  $T$  with respect to  $B$  and  $B'$

$$M_{B'}^B(T) = \left( \underbrace{T(v_1) \quad T(v_2)}_{\text{in terms of } w_1, w_2, w_3} \right)$$

$$T(v_1) = T(1, 4) = (9, -3, 1) = 3w_1 - \frac{3}{5}w_2 + w_3$$

$$T(v_2) = T(3, 0) = (15, 3, 3) = 5w_1 + \frac{3}{5}w_2 + 3w_3$$

Thus we see that  $M_{B'}^B(T) = \begin{bmatrix} 3 & 5 \\ -3/5 & 3/5 \\ 1 & 3 \end{bmatrix}$

**Upshot:** Any vector, written in  $B$  coordinates, when multiplied by this matrix, yields an answer in  $B'$  coordinates. Thus for  $v = av_1 + bv_2$ , we have

$$T(v) = (3a + 5b)w_1 + (-3/5a + 3/5b)w_2 + (a + 3b)w_3$$

- As a sanity check, for  $v = (5, 8) \in R^2$ 
  - Normal Transformation:  $T(v) = (33, -3, 5)$

- Linear Map: writing  $v$  in terms of  $v_1, v_2$ , we get  $(5, 8) = a(1, 4) + b(3, 0) \implies (5, 8) = 2(1, 4) + 2(3, 0)$   
Thus we have

$$M_{B'}^B(T) \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ -3/5 & 3/5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 \\ -3/5 \\ 5 \end{bmatrix} \implies 11(3, 0, 0) - 3/5(0, 5, 0) + 5(0, 0, 1) = (33, -3, 5)$$

**Example:** Consider  $P_n = \{a_0 + a_1x + \dots + a_nx^n \mid a_i \in R\}$

It's easily verifiable that  $P_n$  is a subspace of  $\mathcal{F}(R)$ . Furthermore, the basis for  $P_n$  is  $\{1, x, \dots, x^n\} \implies \dim(P_n) = n + 1$

Let  $D : P_2 \rightarrow P_2$  be the derivative

$$D(a_0 + a_1x + a_2x^2) = a_1 + 2a_2x$$

Easily verifiable that  $D$  is a linear transformation. Consider what is the matrix of  $D$  with respect to  $B = \{1, x, x^2\}$ ?

$$A = [D(1) \quad D(x) \quad D(x^2)] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus we see that for  $p(x) = 5 + 3x + 4x^2$ ,  $SD(p(x)) = 3 + 8x = 5(0, 0, 0) + 3(0, 1, 0) + 4(0, 2, 0)$

**Upshot:** For a linear transformation  $T : V \rightarrow W$ , with  $\dim(V) = n$  and  $\dim(W) = m$ , if  $B = \{v_1, \dots, v_n\}$  and  $B' = \{w_1, \dots, w_m\}$  are bases for  $V, W$ , then

$$M_{B'}^B(T) = [T(v_1) \quad T(v_2) \quad \dots \quad T(v_n)]$$

is a  $m \times n$  matrix with column vectors containing coefficients of  $T(v_1)$  WRT  $B'$

Furthermore, for any  $v \in V, v = x_1v_1 + \dots + x_nv_n$ , we have

$$M_{B'}^B(T) \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

Thus  $T(v) = y_1w_1 + \dots + y_mw_m$  (**Note** coordinate is WRT to  $B'$ )

**Definition - Change of Basis:** Let  $B = \{v_1, \dots, v_n\}$  and  $B' = \{w_1, \dots, w_n\}$  be basis for the same vector space  $V$ , and let  $T : V \rightarrow V$  be the identity mapping. Then

$$M_{B'}^B(\text{id}) = \underbrace{[\text{id}(v_1) \quad \text{id}(v_2) \quad \dots \quad \text{id}(v_n)]}_{\text{WRT } B'}$$

is the **Change of Basis** matrix for  $V$

**Example:** Let  $V = P_1 = \{a_0 + a_1x \mid a_i \in R\}$  and let  $B = \{1, x\}$  and  $B' = \{3 + x, 5 + 2x\}$ , which are both bases for  $V$

$$1 = a(3 + x) + b(5 + 2x) \implies a = 2, b = -1 \implies 1 = 2(3 + x) - (5 + 2x)$$

$$x = c(3 + x) + d(5 + 2x) \implies c = -5, d = 3 \implies x = -5(3 + x) + 3(5 + 2x)$$

$$M_{B'}^B(\text{id}) = \underbrace{\begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}}_{\text{WRT } B'}$$

Furthermore, consider

$$M_B^{B'}(\text{id}) = \underbrace{[\text{id}(w_1) \quad \text{id}(w_2)]}_{\text{WRT } B} = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$$

Finally, we see that  $M_B^{B'}(M_{B'}^B(\text{id})) = \text{id}$

Thus the inverse of  $M_{B'}^B$  is  $M_B^{B'}$

## 5 Scalar Products and Orthogonality

### 5.1 Scalar Products

**Definition - Scalar Product::** For a Vector Space  $V$ , we define  $\langle, \rangle: V \times V \rightarrow K$

- **Example:** Think of dot products in  $R^n \times R^n \rightarrow R$

#### Properties of Scalar Products

1.  $\langle v, w \rangle = \langle w, v \rangle$
2.  $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$
3.  $\langle v, cw \rangle = c \langle v, w \rangle \quad \langle cv, w \rangle = c \langle v, w \rangle$

#### Consequences of Properties

- $\forall v_1, v_2, w \in V, \langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$

*Proof:* Follows from applying properties 1 and 2

- $\forall v \in V, \langle v, O_v \rangle = 0 = \langle O_v, v \rangle$

*Proof:* For any  $w \in V$ , we have  $\langle v, O_V \rangle = \langle v, 0w \rangle = 0 = \langle v, w \rangle$

**Definition - Non-Degenerate:** Scalar product that satisfies  $\forall v \neq 0, \exists w \in V$  such that  $\langle v, w \rangle \neq 0$

**Example:**  $\mathcal{F}([0, 1])$ , all functions  $f: [0, 1] \rightarrow R$

Let  $C([0, 1])$  be the set of all continuous functions  $f: [0, 1] \rightarrow R$ , which is clearly an  $R$  subspace

Now define  $\langle f, g \rangle = \int_0^1 f(x)g(x) dx$ . We claim that this is a scalar product

*Proof:*

- $\int_0^1 f(x)g(x) dx = \int_0^1 g(x)f(x) dx$  so property 1 holds
- $\int_0^1 f(x)(g_1(x) + g_2(x)) dx = \int_0^1 f(x)g_1(x) dx + \int_0^1 f(x)g_2(x) dx$  so property 2 holds
- $\int_0^1 f(x)cg(x) dx = c \int_0^1 f(x)g(x) dx$  so property 3 holds

We also claim that  $\langle f, g \rangle$  is non-degenerate since for  $f \neq 0$ , we have  $\langle f, f \rangle = \int_0^1 f(x)^2$ , which is always  $\geq 0$  and is continuous

**Example:**  $f(x) = 2x + 3 \quad g(x) = x^2$

$$\langle 2x + 3, x^2 \rangle = \int_0^1 (2x + 3)x^2 dx = 3/2$$

**Definition - Orthogonal:** Elements  $v, w \in V$  are **orthogonal**, denote  $v \perp w$ , if  $\langle v, w \rangle = 0$

**Definition - Orthogonal Complement:** Suppose  $W \subseteq V$  is a subspace, then the **orthogonal complement** of  $W$  is

$$W^\perp = \{v \in V \mid v \perp w\} \quad \text{for } w \in W$$

- **Note:**  $W^\perp \subseteq V$  is a subspace

**Definition - Positive Definite:** Scalar product that satisfies  $\forall v \neq O, \langle v, v \rangle \geq 0$ . Otherwise  $\langle v, v \rangle = 0 \implies v = O$

**Definition - Length:**  $\|v\| = \sqrt{\langle v, v \rangle}$

- Length between  $v$  and  $w$ :  $\|v - w\|$
- $\|c\| = |c|\|v\|$
- $\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle = \|v\|^2 - 2\langle v, w \rangle + \|w\|^2$
- $v \perp w \implies \langle v, w \rangle = 0 \implies \|v + w\|^2 = \|v - w\|^2 = \|v\|^2 + \|w\|^2$

**Pythagoras Theorem:** For  $v \perp w$ ,

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2$$

*Proof:*

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \\ &= \|v\|^2 + \|w\|^2 \end{aligned}$$

**Parallelogram Law:** For any  $v, w \in V$ , we have

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2$$

*Proof:* Follows from the definition/properties of length

**Definition - Unit Vector:**  $v \in V$  such that  $\|v\| = 1$

- If  $v \neq O$ , then  $(\frac{1}{\|v\|})v$  is a unit vector

**Definition - Projection:**  $\text{proj}_w v$  represents  $v$  as a scalar multiple of  $w$  where  $\text{proj}_w v = (\frac{\langle v, w \rangle}{\langle w, w \rangle})w$

- Definition comes from creating a right triangle where  $v - cw \perp cw \implies \langle v - cw, cw \rangle = 0$

$$\text{Thus we have } \langle v, cw \rangle - \langle cw, cw \rangle = c \langle v, w \rangle - c^2 \langle w, w \rangle \implies c = \frac{\langle v, w \rangle}{\langle w, w \rangle}$$

- Special case where  $\langle w, w \rangle = 1 \implies \text{proj}_w v = \langle v, w \rangle w$

**Schwartz Inequality:** For any  $v, w \in V$  we have

$$|\langle v, w \rangle| \leq \|v\|\|w\|$$

*Proof:* If  $w = O$ , then  $|\langle v, w \rangle| \leq 0$

Otherwise, using the definition of projection, we have  $cw \perp v - cw$ . Thus we see

$$\begin{aligned} \|v\|^2 &= \|v - cw\|^2 + \|cw\|^2 \\ \implies \|cw\|^2 &\leq \|v\|^2 \\ c^2 \|w\|^2 &\leq \|v\|^2 \\ \frac{\langle v, w \rangle^2}{\langle w, w \rangle^2} \|w\|^2 &\leq \|v\|^2 \\ \implies \langle v, w \rangle^2 &\leq \|v\|^2 \|w\|^2 \end{aligned}$$

**Triangle Inequality:** For  $v, w \in V$ , we have

$$\|v + w\| \leq \|v\| + \|w\|$$

*Proof:*

$$\begin{aligned}
\|v + w\|^2 &= \langle v + w, v + w \rangle \\
&= \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \\
&\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 \\
&\leq (\|v\| + \|w\|)^2 \\
\Rightarrow \|v + w\| &\leq \|v\| + \|w\|
\end{aligned}$$

**Proposition:** Suppose  $\{w_1, \dots, w_r\} \subseteq V$  is pairwise orthogonal and assume that each  $w_i \neq O$ . Then  $\{w_1, \dots, w_r\}$  is linearly independent

*Proof:* Let  $a_1w_1 + \dots + a_rw_r = O_V$ . Then we have

$$\langle w_i, a_1w_1 + \dots + a_rw_r \rangle = \langle w_i, a_1w_1 \rangle + \dots + \langle w_i, a_rw_r \rangle = 0 \quad \text{since each } w \text{ is pairwise orthogonal}$$

Thus  $\langle w_i, a_iw_i \rangle = 0 \Rightarrow a \langle w_i, w_i \rangle = 0 \Rightarrow a_i = 0$  since  $\langle w_i, w_i \rangle > 0$  since positive definite

Let  $W = \text{span}(\{w_1, \dots, w_r\}) \subseteq V$ . Then clearly  $\dim(W) = r$

Now take  $v \in V$  and define  $\text{proj}_W v = \sum_{i=1}^r c_i w_i$  where  $c_i w_i = \text{proj}_{w_i} v$

Clearly  $\text{proj}_W v \in W$

**Proposition:**  $\left(v - \sum_{i=1}^r c_i w_i\right) \perp$  each  $w_i$

*Proof:* Fix  $i$ , then

$$\sum_{j=1}^r c_j w_j = c_i w_i + \sum_{j \neq i} c_j w_j$$

Now take

$$v - \sum_{j=1}^r c_j w_j = (v - c_i w_i) - \sum_{j \neq i} c_j w_j$$

and take the inner product with  $w_i$

$$\begin{aligned}
\underbrace{\langle w_i, v - c_i w_i \rangle}_{0 \text{ b/c of projection}} &= \langle w_i, \underbrace{\sum_{j \neq i} c_j w_j}_{0 \text{ b/c orthogonal}} \rangle
\end{aligned}$$

Thus we have  $w_i \perp v - \sum_{j=1}^r c_j w_j$

**Corollary:**  $(v - \sum_{j=1}^r c_j w_j) \perp$  every  $w \in W$

*Proof:* Since each  $w_i$  in the basis is orthogonal to  $v - \sum_{j=1}^r c_j w_j$ , we must have

$$\langle w, v - \sum_{j=1}^r c_j w_j \rangle = 0$$

**Corollary:**  $(v - \sum_{j=1}^r c_j w_j) \in W^\perp$

*Proof:* Follows from the previous corollary

**Geometric Interpretation:** For any  $v \in V$ ,  $\text{proj}_W v$  is the closest point to  $v$  in  $W$

$$\|v - \text{proj}_W v\| \leq \|v - w\|$$

*Proof:* Choose any  $w \in W = \text{span}(\{v_1, \dots, v_r\})$ , then  $w = \sum_{i=1}^r a_i v_i$ . Then we have

$$\begin{aligned} \|v - w\|^2 &= \left\| v - \sum_{i=1}^r a_i v_i \right\|^2 \\ &= \left\| v - \underbrace{\sum_{i=1}^r c_i v_i}_{\perp W} + \underbrace{\sum_{i=1}^r (c_i a_i) v_i}_{\in W} \right\|^2 \\ &= \left\| v - \sum_{i=1}^r c_i v_i \right\|^2 + \left\| \sum_{i=1}^r (c_i - a_i) v_i \right\|^2 \quad \text{by Pythagoras} \end{aligned}$$

$$\text{Thus } \|v - w\|^2 \geq \left\| v - \sum_{i=1}^r c_i v_i \right\|^2 \implies \|v - w\| \geq \left\| v - \sum_{i=1}^r c_i v_i \right\|$$

**Corollary:** Suppose  $w \in W$ , then  $\text{proj}_W w$  is the element of  $W$  closest to  $w$

$$\text{But we have } w = \sum_{i=1}^r c_i v_i \implies c_i = \frac{\langle w, v_i \rangle}{\|v_i\|^2}$$

## 5.2 Orthonormal Basis

**Definition - Orthonormal Basis:**  $\{v_1, \dots, v_r\} \subseteq W$  is an **orthonormal basis** if

1.  $\{v_1, \dots, v_r\}$  are pairwise orthogonal and non are zero
2.  $\|v_i\| = 1$

**Corollary:** If  $\{v_1, \dots, v_r\}$  is orthonormal, then  $\forall w \in W, w = \sum_{i=1}^r \langle w, v_i \rangle v_i$

**Gram-Schmidt Process:** Turn any basis  $B = \{v_1, \dots, v_n\}$  into an orthonormal basis  $B' = \{u_1, \dots, u_n\}$

1. Given  $v_1$ , let  $u_1 = \frac{1}{\|v_1\|} v_1$ . Then we have  $\text{span}(\{u_1\}) = \text{span}(\{v_1\})$
2. Let  $p_2 = v_2 - \text{proj}_{u_1} v_2 = v_2 - \langle v_2, u_1 \rangle u_1$   
Now let  $u_2 = \frac{1}{\|p_2\|} p_2$ . Then  $\text{span}(\{u_1, u_2\}) = \text{span}(\{v_1, v_2\})$
3. Let  $p_3 = v_3 - \text{proj}_{\text{span}(\{u_1, u_2\})} v_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2$   
Now let  $u_3 = \frac{1}{\|p_3\|} p_3$ . Then  $\text{span}(\{u_1, u_2, u_3\}) = \text{span}(\{v_1, v_2, v_3\})$
4. ...

**Upshot:** Any finite  $R$  Vector Space  $V$  with a positive definite inner product has an orthonormal basis