

# 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 . . . . .	7
1.3.1	Toolbox Corollaries and Results . . . . .	10
1.4	Direct Sums . . . . .	11
<b>2</b>	<b>Matrices</b>	<b>11</b>
2.1	Linear Equations . . . . .	13
<b>3</b>	<b>Mappings</b>	<b>13</b>
3.1	Consequences of Properties of Linear Transformations . . . . .	15
3.2	Kernel . . . . .	16
3.2.1	Consequences of Kernel . . . . .	16
3.3	Compositions and Inverse Linear Mappings . . . . .	17
<b>4</b>	<b>Linear Maps and Matrices</b>	<b>18</b>
4.1	Bases, Matrices, and Linear Maps . . . . .	18
<b>5</b>	<b>Scalar Products and Orthogonality</b>	<b>21</b>
5.1	Scalar Products . . . . .	21
5.2	Orthonormal Basis . . . . .	25
5.3	Application to Linear Equations: Rank . . . . .	26
5.4	Scalar Products Under Complex Numbers . . . . .	27
5.5	General Orthogonal Bases . . . . .	27
5.5.1	Properties and Types of Scalar Products . . . . .	27
5.6	Quadratic Forms . . . . .	31
5.7	Sylvester's Theorem . . . . .	32
5.8	Riesz Representation . . . . .	34
<b>6</b>	<b>Operators</b>	<b>35</b>
6.1	Multilinear k-form . . . . .	35
6.1.1	Consequences of the Big Count . . . . .	39
6.1.2	Matrix Representation . . . . .	40
<b>7</b>	<b>Determinants</b>	<b>41</b>
7.1	Row Determinants . . . . .	42
<b>8</b>	<b>Symmetric, Hermitian, Unitary Operators</b>	<b>42</b>
<b>9</b>	<b>Eigenvectors and Eigenvalues</b>	<b>43</b>

Goals of this course are to discuss

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

# 1 Vector Space

## 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  $(a, b, c) + (d, e, f) = (a + d, b + e, c + f)$
- $\times$  : 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  $f + g := (f + g)(x)$
- $\times$  : 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  $a \in K, v \in W \implies aw \in W$  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

$$a_1 v_1 + \dots + a_n v_n \quad a_i \in K$$

**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  $a = a_1 v_1 + \dots + a_n v_n \in W$  and  $b = b_1 v_1 + \dots + b_n v_n \in W$

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  $a = a_1 v_1 + \dots + a_n v_n \in W$  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  $a_1 v_1 + \dots + a_n v_n = 0$ , 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  $v_i = a_1 v_1 + \dots + a_n v_n$  for some  $v_i \notin \{v_1, \dots, v_n\}$

Then we have

$$O = a_1 v_1 + \dots + a_n v_n + (-1)v_i$$

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

$$a_1 v_1 + \dots + a_n v_n = O$$

WLOG, assume  $a_1 \neq 0$ . Then  $v_1 a_1 + \dots + a_n v_n = a_1 v_1$

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

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

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. Notably,  $\emptyset$  is the basis for  $\{O\}$

**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  $\{u_1, \dots, u_r, b_1, \dots, b_m\}$  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  $\text{Mat}_{2 \times 2}(K)$ , has dimension 3

**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 an  $m \times 1$  matrix

**Definition - Transpose:** Given an  $m \times n$  matrix  $A$ , the **transpose**  ${}^t A$  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  ${}^t A$  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 the following system

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

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

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

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

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

**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  $\text{Image of } F = R$

**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:**  $T(a_1v_1 + \dots + a_nv_n) = a_1T(v_1) + \dots + a_nT(v_n)$

*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  $T(v) = w$

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

**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  $a_1v_1 + \dots + a_mv_m = O_V$

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

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 linear 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  $v \in V \implies T(v) \in \text{Im}(T) = a_1 w_1 + \dots + a_m w_m$

Let  $s = a_1 v_1 + \dots + a_m v_m \in S$ .

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

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

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

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

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  $A = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 4 & 2 \end{bmatrix}$

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

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

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

**Upshot:** 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  $\theta$

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

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

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

**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  $B'$

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

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

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

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

**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

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

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

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  $p(x) = 5 + 3x + 4x^2$ ,

$$D(p(x)) = 3 + 8x = 5(0, 0, 0) + 3(0, 1, 0) + 4(0, 2, 0)$$

**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

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

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

Furthermore, for any  $v \in V, v = x_1v_1 + \cdots + x_nv_n$ , 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  $T(v) = y_1w_1 + \cdots + y_mw_m$  (**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

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

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

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

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  $> 0$  and is continuous

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

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

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

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

$$W^\perp = \{v \in V \mid \forall w \in W, v \perp w\}$$

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

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

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

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

**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

Let  $w \in V$ . For any  $v \in V$  there exists  $c \in K$  such that  $v - cw \perp w$ . That is  $\langle v - cw, w \rangle = 0$

**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  $v - cw \perp w \implies \langle v - cw, w \rangle = 0$

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

- 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  $|\langle v, w \rangle| = 0 \leq \|v\| \|w\| = 0$

Otherwise, assume that  $w$  is a unit vector. Using the definition of projection, we have  $cw \perp v - cw$ . Thus we see

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

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

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

*Proof:*

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &= \|v\|^2 + 2\langle v, w \rangle + \|w\|^2 \\ &\leq \|v\|^2 + \underbrace{2\|v\| \|w\|}_{\text{by Schwartz}} + \|w\|^2 \\ &\leq (\|v\| + \|w\|)^2 \\ \implies \|v + w\| &\leq \|v\| + \|w\| \end{aligned}$$

**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  $a_1 w_1 + \dots + a_r w_r = O_V$ . Then we have

$$\langle w_i, a_1 w_1 + \dots + a_r w_r \rangle = \langle w_i, a_1 w_1 \rangle + \dots + \langle w_i, a_r w_r \rangle = 0 \quad \text{since each } w \text{ is pairwise orthogonal}$$

Thus  $\langle w_i, a_i w_i \rangle = 0 \implies a \langle w_i, w_i \rangle = 0 \implies a_i = 0$  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:**  $\left(v - \sum_{j=1}^r c_j w_j\right) \perp$  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  $w_i$

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

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  $W$

$$\|v - \text{proj}_W v\| \leq \|v - w\| \quad \text{for any arbitrary } w \in W$$

*Proof:* Choose any  $w \in W = \text{span}(\{w_1, \dots, w_r\})$ , then  $w = \sum_{i=1}^r a_i w_i$ . Then we have

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

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

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

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

## 5.2 Orthonormal Basis

**Definition - Orthonormal Basis:**  $\{w_1, \dots, w_r\} \subseteq W$  is an **orthonormal basis** if

1.  $\{w_1, \dots, w_r\}$  are pairwise orthogonal and none are zero
2.  $\|w_i\| = 1$  for  $i \in \{1, \dots, r\}$

**Corollary:** If  $\{w_1, \dots, w_r\}$  is orthonormal, then  $\forall w \in W, w = \sum_{i=1}^r \langle w, w_i \rangle w_i$

**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. Repeat

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

**Theorem:** Let  $V$  be a finite dimension  $R$  Vector Space with a positive definite scalar product. Then for any subspace  $W \subseteq V$

$$V = W \oplus W^\perp$$

*Proof:*

- Show that  $V = W + W^\perp$

Choose  $v \in V$  and let  $w^* = \text{proj}_W v \in W$ . Then  $v - w^* \in W^\perp$

$$\text{Thus } v = \underbrace{w^*}_{\in W} + \underbrace{(v - w^*)}_{\in W^\perp}$$

- Show that  $W \cap W^\perp = \{O\}$

Choose  $w \in W \cap W^\perp$

Since  $w \in W^\perp$ ,  $w$  is orthogonal to all vectors in  $W$

In particular,  $w \perp w \implies \langle w, w \rangle = 0 \implies w = O$  since the scalar product is positive definite

**Corollary:** If  $W \subseteq V$  is a subspace, then

$$\dim(V) = \dim(W) + \dim(W^\perp)$$

### 5.3 Application to Linear Equations: Rank

Let  $A$  be an  $m \times n$  matrix with entries in  $R$

- Let  $C_A \subseteq R^m$  be the span of column vectors of  $A$
- Let  $R_A \subseteq R^n$  be the span of row vectors of  $A$
- Let  $\text{Null}(A) = \{v \in R^n \mid Av = O\}$

Recall that any  $m \times n$  matrix  $A$  describes a linear transformation  $L_A : R^n \rightarrow R^m$  where  $L_A(v) = Av \in R^m$

Thus  $\text{Im}(L_A) = C_A$

Furthermore,  $\text{Ker}(L_A) = \{v \in R^n \mid Av = O\} = \text{Null}(A)$

Thus we have

$$\begin{aligned} \dim(R^n) &= \dim(\text{Im}(L_A)) + \dim(\text{Ker}(L_A)) \\ &= \dim(C_A) + \dim(\text{Null}(A)) \end{aligned}$$

Now consider using scalar products

Take  $v \in \text{Null}(A)$ . Thus  $Av = O$

$$\text{Thus } A_i \cdot v = 0 \iff A_i \perp v \iff v \perp \text{all } u \in R_A \implies v \in (R_A)^\perp$$

$$\text{Thus } \text{Null}(A) = \text{Ker}(A) = (R_A)^\perp$$

Thus  $R_A \subseteq R^n$  is a subspace of  $R^n$ .

Thus we have

$$\begin{aligned} \dim(R^n) &= \dim(R_A) + \dim((R_A)^\perp) \\ n &= \dim(R_A) + \dim(\text{Null}(A)) \end{aligned}$$

Thus we have  $\dim(R_A) = \dim(C_A)$

**Definition - Rank:** The **rank** of a matrix  $A$  is  $\dim(R_A) = \dim(C_A)$

## 5.4 Scalar Products Under Complex Numbers

We want a positive definite scalar product for  $C$

Take the **complex conjugate**

$$(a + bi)(a - bi) = a^2 + b^2$$

Then we see that

$$\|z\| = \sqrt{\langle z, \bar{z} \rangle} \in R$$

**Definition - Hermitian Inner Product:** For  $(y_1, \dots, y_n)$  and  $(z_1, \dots, z_n) \in C^n$ , define

$$\langle y, z \rangle = y_1 \bar{z}_1 + \dots + y_n \bar{z}_n$$

- **Note:** This is NOT a scalar product since  $\langle y, z \rangle \neq \langle z, y \rangle$

Now we list the properties of the Hermitian Inner Product

- $\langle w, v \rangle = \overline{\langle v, w \rangle}$
- $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$
- $\langle cv, w \rangle = c \langle v, w \rangle$  AND  $\langle v, cw \rangle = \bar{c} \langle v, w \rangle$

**Proposition:** The Hermitian Inner Product is positive definite

*Proof:* We look at

$$\langle v, v \rangle = x_1 \bar{x}_1 + \dots + x_n \bar{x}_n = \|x_1\|^2 + \dots + \|x_n\|^2 \in R$$

We see that  $\langle v, v \rangle \geq 0$ . If it happens that  $\langle v, v \rangle = 0 \implies x_1 = \dots = x_n = 0$

## 5.5 General Orthogonal Bases

### 5.5.1 Properties and Types of Scalar Products

Repeating a lot of what was stated above for clarity

A **scalar product** satisfies

1. Symmetry:  $\langle v, w \rangle = \overline{\langle w, v \rangle}$
2. Linear:  $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$
3. Scalar:  $\langle cv, w \rangle = c \langle v, w \rangle$  AND  $\langle v, cw \rangle = \bar{c} \langle v, w \rangle$

Types of scalar products (each progressively weaker):

- **Positive Definite:**  $\forall v \in V, \langle v, v \rangle \geq 0$  AND  $\langle v, v \rangle = 0 \implies v = 0$
- **Non-Degenerate:** For  $v \neq 0, \exists w \in V$  such that  $\langle v, w \rangle \neq 0$
- **Non-Trivial:**  $\exists v, w \in V$  such that  $\langle v, w \rangle \neq 0$

**Upshot:** positive definite  $\implies$  non-degenerate  $\implies$  non-trivial

We also consider **Trivial Scalar Products** where  $\forall v, w \in V$ , we have  $\langle v, w \rangle = 0$

For a positive definite  $\langle, \rangle$ , we proved that

1. Every finite dimensional Vector Space  $V$  has an orthonormal basis (**Gram Schmidt Process**)
2. For any subspace  $W \subseteq V$ , we have  $V = W \oplus W^\perp$  (**Projection**)

**Observation:** If  $\langle, \rangle$  is trivial, then any basis of  $V$  is orthogonal

**Lemma:** Suppose  $\langle v, v \rangle = 0$  for all  $v \in V$ , then  $\langle, \rangle$  is trivial

*Proof:* Choose any  $v, w \in V$ . Then we see

$$\langle v + w, v + w \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle$$

Thus we have

$$\langle v, w \rangle = \frac{1}{2}(\langle v + w, v + w \rangle - \langle v, v \rangle - \langle w, w \rangle) = 0$$

**Corollary:** If  $\langle v, v \rangle = 0$  for all  $v \in V$ , then any basis of  $V$  is orthogonal

*Proof:* Since  $\langle, \rangle$  is trivial (shown from the Lemma), by the observation above, any basis of  $V$  is orthogonal

**Theorem 1:** If  $\langle, \rangle$  is any scalar product on  $V$ , then  $V$  has an orthogonal basis

*Proof:* By Induction on  $n = \dim(V)$

Claim: If  $\langle, \rangle$  is any scalar product on any finite dimensional Vector Space  $V$  with  $\dim(V) \leq n$ , then  $V$  has an orthogonal basis

Base Case:  $n = 0 : \dim(V) \implies B = \{\}$  is a basis and is an orthogonal basis

Base Case:  $n = 1 : \dim(V) = 1 \implies \{v_1\}$  is an orthogonal basis for  $v_1 \in V, v_1 \neq 0$

IH: Assume the claim holds for  $\dim(V) = n - 1$

IS: Suppose  $\dim(V) = n$

- Case 1:  $\forall v \in V, \langle v, v \rangle = 0$ . Then by the preceding Lemma,  $\langle, \rangle$  is trivial and any basis for  $V$  is an orthogonal basis
- Case 2:  $\exists v_1 \in V$  such that  $\langle v_1, v_1 \rangle \neq 0$

Let  $V_1 = \text{span}(\{v_1\}) \subseteq V$  be a subspace. We show that  $V = V_1 \oplus V_1^\perp$

– Show that  $V = V_1 + V_1^\perp$

Choose  $v \in V$ . Since  $\langle v_1, v_1 \rangle \neq 0$  we can use projection:  $\text{proj}_{v_1} v = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 \in V_1$

Thus  $(v - \text{proj}_{v_1} v) \perp v_1 \implies (v - \text{proj}_{v_1} v) \in V_1^\perp$

Thus  $v = \underbrace{(\text{proj}_{v_1} v)}_{\in V_1} + \underbrace{(v - \text{proj}_{v_1} v)}_{\in V_1^\perp}$

– Show  $V_1 \cap V_1^\perp = \{O\}$

Choose  $v \in V_1 \cap V_1^\perp$

$v \in V_1^\perp$  and  $v \in V_1 \implies v \perp v \implies \langle v, v \rangle = 0$

However,  $v \in V_1 \implies v = dv_1 \implies 0 = \langle v, v \rangle = \langle dv_1, dv_1 \rangle = d^2 \underbrace{\langle v_1, v_1 \rangle}_{\neq 0}$

Thus we see that  $d = 0 \implies v = O$

Now we have  $\dim(V) = \dim(V_1) + \dim(V_1^\perp) \implies \dim(V_1^\perp) = n - 1$  which by IH has an orthogonal basis  $\{v_2, \dots, v_n\}$

Finally, since  $v_1 \perp v_i$  for  $2 \leq i \leq n$ , we see that  $\{v_1, v_2, \dots, v_n\}$  is a orthogonal basis for  $V$

**Definition - Dual Space:**  $K$ -Vector Space  $V^* = \mathcal{L}(V, K)$  where each element of  $V^*$  is a linear transformation  $\phi : V \rightarrow K$

- **Note:** For any  $w_1, \dots, w_n \in W$ , there is exactly one Linear Transformation  $T : V \rightarrow W$  such that  $T(v_i) = w_i$  for  $1 \leq i \leq n$

**Example:** Let  $B = \{v_1, \dots, v_n\}$  be a basis for  $V$  and take

$$\begin{aligned}\phi_1 : V &\rightarrow K & \phi_1(v) &= \phi_1(a_1v_1 + \dots + a_nv_n) = a_1 \\ \phi_2 : V &\rightarrow K & \phi_2(v) &= \phi_2(a_1v_1 + \dots + a_nv_n) = a_2 \\ && \dots &\end{aligned}$$

Thus we see that  $\phi_i(v_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$

Let  $B' = \{\phi_1, \dots, \phi_n\}$ . Then we see that  $B'$  is a basis for  $V^*$

- Show linear independence: Take  $a_i \in K$  such that  $\underbrace{O}_{O \text{ mapping}} = \underbrace{a_1\phi_1 + \dots + a_n\phi_n}_{\text{mapping}}$

This equality means that  $\forall w \in V$ , we have  $(a_1\phi_1 + \dots + a_n\phi_n)(w) = O(w)$

Now applying the transformation to  $v_1$ , we see that  $a_1\phi_1(v_1) = O(v_1) = 0 \implies a_1 = 0$

Similar logic shows that  $a_i = 0$  for  $1 \leq i \leq n$

- Show  $B'$  spans  $\mathcal{L}(V, K)$

Choose any  $T \in \mathcal{L}(V, K)$ . Then we see

$$T(v_1) = b_1 \in K, \dots, T(v_n) = b_n \in K$$

Now let  $\phi^* = b_1\phi_1 + \dots + b_n\phi_n$ . Clearly  $\phi \in \text{span}(B')$

We show that  $\phi^* = T$  (they need to agree on all input)

It suffices so show that  $\phi^*(v_j) = T(v_j)$  for  $v_j \in B$  since  $B$  is a basis of  $V$

Simple calculations show that  $\phi^*(v_j) = (b_1\phi_1 + \dots + b_n\phi_n)(v_j) = b_j = T(v_j)$

Thus  $T \in \text{span}(B)$

**Corollary:**  $\dim(V^*) = \dim(V) = n$  (so same size as basis)

**Corollary:**  $V$  is isomorphic to  $V^*$ . Namely, there exists a 1-1, onto linear transformation  $F : V \rightarrow V^*$  where

$$F(v_1) = \phi_1, \dots, F(v_n) = \phi_n$$

These  $\phi_i$  uniquely describe  $F$

Consider a subspace  $W \subseteq V$

**Definition - Annihilator:**  $\text{Ann}(W) = \{\phi \in V^* \mid \forall w \in W, \phi(w) = 0\}$ , so the set of linear transformations  $\phi$  in  $V^*$  such that  $W \subseteq \text{Ker}(\phi)$

**Annihilator Theorem:** For any  $W \subseteq V$

$$\dim(W) + \dim(\text{Ann}(W)) = \dim(V) = n$$

*Proof:* Choose a basis for  $W$ ,  $\{w_1, \dots, w_r\}$

Now extend it to a basis for  $V$ ,  $B = \{w_1, \dots, w_r, w_{r+1}, \dots, w_n\}$

Let  $B' = \{\phi_1, \dots, \phi_n\}$  be the dual basis of  $V^*$  corresponding to  $B$

We claim that  $\{\phi_{r+1}, \dots, \phi_n\}$  is a basis for  $\text{Ann}(W)$

- For any  $w \in W$ ,  $w = a_1 w_1 + \dots + a_r w_r$ , and  $j \geq r+1$ , we have that  $w_j = 0 \implies \phi_j(w) = 0 \implies \{\phi_{r+1}, \dots, \phi_n\} \subseteq \text{Ann}(W)$
- $\{\phi_{r+1}, \dots, \phi_n\}$  is linearly independent since  $B'$  is linearly independent
- To show that  $\text{span}(\{\phi_{r+1}, \dots, \phi_n\}) = \text{Ann}(W)$

Take  $T \in \text{Ann}(W) \implies T : V \rightarrow K$  is a linear transformation

Furthermore, we have  $T(w_1) = 0, \dots, T(w_r) = 0$

Since  $T \in B'$  (since  $B'$  is a basis for  $V^*$ ), we have that  $T = a_1 \phi_1 + \dots + a_r \phi_r + \dots + a_n \phi_n$

Now we see  $T(w_1) = (a_1 \phi_1 + \dots + a_n \phi_n)(w_1) = a_1 = 0$

Similarly, we see  $a_i = 0$  for  $1 \leq i \leq r$

Thus  $T = a_{r+1} \phi_{r+1} + \dots + a_n \phi_n \in \text{span}(\{\phi_{r+1}, \dots, \phi_n\})$

**Theorem 2:** If  $\langle, \rangle$  is non-degenerate, then for every subspace  $W \subseteq V$ , we have

$$V = W \oplus W^\perp$$

Now consider a  $\langle, \rangle$  non-degenerate

**Claim:**  $\forall v \in V$ , given a linear transformation  $L_v : V \rightarrow K$ , let  $L_v(w) = \langle v, w \rangle \in K$ , then  $F : V \rightarrow V^*$  where  $F(v) = L_v$  is an isomorphism

## 5.6 Quadratic Forms

**Definition - Symmetric Bilinear Form:** Another way of calling scalar products on a vector space  $V$

- **Symmetric** comes from  $\langle v, w \rangle = \langle w, v \rangle$
- **Bilinear** comes from  $\langle v, w_1 + w_2 \rangle = \langle v, w_1 \rangle + \langle v, w_2 \rangle$  and  $\langle v, cw \rangle = c \langle v, w \rangle = \langle cv, w \rangle$
- **Form** comes from the mapping  $(v, w) \rightarrow \langle v, w \rangle$ , often denoted as a function

$$g : V \times V \rightarrow K \quad g(v, w) = \langle v, w \rangle$$

**Definition - Quadratic Form:** Given a scalar product  $g = \langle, \rangle$ , the **quadratic form** determined by  $g$  is a function

$$f : V \rightarrow K \quad f(v) = g(v, v) = \langle v, v \rangle$$

**Example:** If  $V = K^n$  then  $f(X) = X \cdot X = x_1^2 + \dots + x_n^2$  is the quadratic form determined by regular dot product

In general, if  $V = K^n$  and  $C$  is a symmetric matrix, then the quadratic form is given by

$$F(X) = {}^t X C X = \sum_{i,j=1}^n c_{ij} x_i x_j$$

For a diagonal matrix  $C$ , this simplifies to

$$F(X) = c_1 x_1^2 + \cdots + c_n x_n^2$$

## 5.7 Sylvester's Theorem

Let  $V = R^2$  and let the form be represented by the symmetric matrix

$$C = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

Then the vectors

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

form an orthogonal basis using  $f(X) = \langle X, X \rangle = {}^t X C X$ . Indeed

$$\langle v_1, v_1 \rangle = -1 \quad \langle v_2, v_2 \rangle = 0$$

Now we generalize the situation above to arbitrary dimensions

Let  $\{v_1, \dots, v_n\}$  be an orthogonal basis of  $V$  and let

$$c_i = \langle v_i, v_i \rangle$$

After some renumbering of elements in our basis, we can assume that

$$\begin{aligned} c_1, \dots, c_r &> 0 \\ c_{r+1}, \dots, c_s &< 0 \\ c_{s+1}, \dots, c_n &= 0 \end{aligned}$$

We are interested in looking at the number of positive, negative, and zero terms among  $c_i = \langle v_i, v_i \rangle$  i.e. the numbers  $r$  and  $s$

Let  $X$  be the coordinate vector of an element of  $V$  with respect to our basis and let  $f$  be the quadratic form associated with our scalar product. Then

$$F(X) = c_1 x_1^2 + \cdots + c_r x_r^2 + \cdots + c_s x_s^2$$

Here we see  $r$  positive terms,  $s - r$  negative terms, and that  $n - s$  of the terms have disappeared

We can see this more clearly by normalizing the basis

**Definition - Orthonormal:** A basis  $\{v_1, \dots, v_n\}$  is **orthonormal** if for each  $i$  we have

$$\langle v_i, v_i \rangle = 1 \quad \text{or} \quad \langle v_i, v_i \rangle = -1 \quad \text{or} \quad \langle v_i, v_i \rangle = 0$$

If  $\{v_1, \dots, v_n\}$  is a orthogonal basis, we can always obtain an orthonormal basis by taking



- $c_i = 0 \implies v'_i = v_i$
- $c_i > 0 \implies v'_i = \frac{v_i}{\sqrt{c_i}}$
- $c_i < 0 \implies v'_i = \frac{v_i}{\sqrt{-c_i}}$

Then  $\{v'_1, \dots, v'_n\}$  is an orthonormal basis

Now suppose that  $X$  is the coordinate vector of an element of  $V$ . In terms of the orthonormal basis, we have

$$f(X) = x_1^2 + \dots + x_r^2 - x_{r+1}^2 - \dots - x_s^2$$

Thus we can clearly see the number of positive and negative terms

We now show that number of positive, negative, and zero terms don't depend on the orthonormal basis

**Theorem 8.1:** Let  $V$  be a finite dimensional vector space over  $R$  with a scalar product. Take the subspace  $V_0 \subseteq V, V_0 = \{v \in V \mid \forall w \in V, \langle v, w \rangle = 0\}$ . Then the number of integers  $i$  such that  $\langle v_i, v_i \rangle = 0$  is equal to the dimension of  $V_0$

*Proof:* Suppose  $\{v_1, \dots, v_n\}$  is ordered such that

$$\langle v_1, v_1 \rangle \neq 0, \dots, \langle v_s, v_s \rangle \neq 0 \quad \text{but } \langle v_i, v_i \rangle = 0 \quad \text{for } i > s$$

Since  $\{v_1, \dots, v_n\}$  is orthogonal, clearly  $v_{s+1}, \dots, v_n \in V_0$

Now we take  $v \in V_0$

$$v = x_1 v_1 + \dots + x_s v_s + \dots + x_n v_n$$

Taking the scalar product with any  $v_j$  for  $j \leq s$ , we get

$$0 \langle v, v_j \rangle = x_j \langle v_j, v_j \rangle \implies x_j = 0 \implies v \in \text{span}(\{v_{s+1}, \dots, v_n\})$$

Furthermore, since  $\{v_{s+1}, \dots, v_n\}$  is linearly independent, we have that  $\{v_{s+1}, \dots, v_n\}$  is a basis for  $V_0$

**Definition - Index of Nullity:** From the proof above, we call  $V_0$  the **index of nullity of the form**

- **Note:** Here form is non-degenerate if and only if the index of nullity = 0

**Sylvester's Theorem:** Let  $V$  be a finite dimensional vector space of  $R$ . Then there exists  $r \geq 0$  such that if  $\{v_1, \dots, v_n\}$  is a basis, then there are precisely  $r$  integers such that

$$\langle v_i, v_i \rangle > 0$$

*Proof* Let  $\{v_1, \dots, v_n\}$  and  $\{w_1, \dots, w_n\}$  be orthogonal bases for  $V$ . Arrange them such that

$$\begin{array}{ll} \langle v_i, v_i \rangle > 0 & 1 \leq i \leq r \\ \langle v_i, v_i \rangle < 0 & r+1 \leq i \leq s \\ \langle v_i, v_i \rangle = 0 & s+1 \leq i \leq n \\ \langle w_i, w_i \rangle > 0 & 1 \leq i \leq r' \\ \langle w_i, w_i \rangle < 0 & r'+1 \leq i \leq s' \\ \langle w_i, w_i \rangle = 0 & s'+1 \leq i \leq n \end{array}$$

We show that  $v_1, \dots, v_r, w_{r'+1}, \dots, w_n$  is linearly independent

Suppose that we have

$$x_1 v_1 + \dots + x_r v_r + y_{r'+1} w_{r'+1} + \dots + y_n w_n = 0 \implies x_1 v_1 + \dots + x_r v_r = -(y_{r'+1} w_{r'+1} + \dots + y_n w_n)$$

Let  $c_i = \langle v_i, v_i \rangle$  and  $d_i = \langle w_i, w_i \rangle$

Taking the scalar product of both sides with itself, we see that

$$c_1 x_1^2 + \dots + c_r x_r^2 = d_{r'+1} y_{r'+1}^2 + \dots + d_n y_n^2$$

Clearly the LHS  $\geq 0$  and the RHS  $\leq 0 \implies$  both sides are 0

Thus  $x_1 = \dots = x_r = 0 \implies y_{r'+1} = \dots = y_n = 0$  by linear independence

Finally, since  $\dim(V) = n$ , we see that  $r + n - r' \leq n \implies r \leq r'$

However, by symmetric we also get that  $r' \leq r$

Thus we must have that  $r = r'$

**Definition - Index of Positivity:** From Sylvester's Theorem, the integer  $r$  is called the **index of positivity**

## 5.8 Riesz Representation

Recall that  $P_2(R) = \{a_0 + a_1 x + a_2 x^2 \mid a_i \in R\}$

Also recall that if  $\langle, \rangle$  is non-degenerate, then  $L^* : V \rightarrow V^*$  is an isomorphism where

$$L^*(v) = L_v : V \rightarrow K \quad L_v(w) = \langle v, w \rangle$$

**Riesz Representation Theorem:** For any finite dimensional vector space  $V$  with a non-degenerate  $\langle, \rangle$ , for any linear function  $\phi : V \rightarrow K \in V^*$ , there exists a unique  $u \in V$  such that  $\phi = L_u$

*Proof:* Since  $L^* : V \rightarrow V^*$  is an isomorphism, we let  $u = (L^*)^{-1}(\phi)$

**Proposition:** There is a polynomial  $u(x) \in P_2(R)$  such that for all  $p(x) \in P_2(R)$

$$\int_0^1 p(x) u(x) dx = \int_{-\pi}^{\pi} p(x) \cos(x) dx$$

*Proof:* Clearly  $V = P_2(R)$  is finite dimensional and  $\langle f, g \rangle = \int_0^1 fg$  is non-degenerate

Let

$$\phi : P_2(R) \rightarrow R \quad \phi(p) = \int_{-\pi}^{\pi} p(x) \cos(x) dx$$

Now we use Riesz Representation Theorem to get  $u$  such that

$$\int_0^1 p(x) u(x) dx = \langle u, p \rangle = \int_{-\pi}^{\pi} p(x) \cos(x) dx$$

**Proposition:** There is a  $u(x) \in P_2(R)$  such that for all  $p(x) \in P_2(R)$  we have

$$\int_0^1 p(x)u(x) dx = P(0) = a_0$$

*Proof:* Let

$$\psi : P_2(R) \rightarrow R \quad \psi(a_0 + a_1x + a_2x^2) = a_0$$

Then apply Riesz Representation Theorem

## 6 Operators

**Definition - Operators:** Linear transformations  $T : V \rightarrow V$

**Definition**  $\mathcal{L}(V, V)$ : Set of all linear transformations  $T : V \rightarrow V$

- **Note:**  $\mathcal{L}(V, V)$  is a Vector Space

For the remainder of the course, we look at **operators** of  $V$

For every linear transformation  $T : V \rightarrow V$ , we have an  $n \times n$  matrix  $A$

However, there are many different  $n \times n$  matrices associated to the same transformation  $T$

In fact, for any basis  $B = \{v_1, \dots, v_n\}$ , we get a matrix  $M_{n \times n}(T)_B^B$

In particular, we study properties of  $n \times n$  matrices  $A$  that don't depend on the change of basis

### 6.1 Multilinear k-form

**Definition - Multilinear k-form:** A function  $\omega : \underbrace{V \times \dots \times V}_{k \text{ factors}} \rightarrow K$  such that for all  $1 \leq i \leq n$ , for all  $v_1, \dots, v_i, w_i, v_{i+1}, \dots, v_k$ , and  $a, b \in K$  we have

$$\omega(v_1, \dots, v_{i-1}, (av_i + bw_i), v_{i+1}, \dots, v_k) = a\omega(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_k) + b\omega(v_1, \dots, v_{i-1}, w_i, v_{i+1}, \dots, v_k)$$

**Upshot:** It's linear on each coordinate, provided that the other coordinates stay fixed

Let  $ML_k(V)$  be the set of all multilinear k-forms  $\omega : V^k \rightarrow K$

- **Note:**  $ML_k(V)$  is a  $K$ -Vector Space

**Consider:** What is a multilinear 1-form

$\omega : V \rightarrow K$  is a linear transformation. Thus  $\{\omega : V \rightarrow K\} = V^* = \text{dual space}$

**Consider:** What is a multilinear 2-form (**bilinear form**)

$\omega : V \times V \rightarrow K$  is linear in each coordinate

$ML_2(V)$  is the set of all bilinear forms on  $V$

- **Note:** Scalar Products  $\subseteq ML_2$

**Definition - Alternating:** A multilinear  $k$ -form  $\omega : V^k \rightarrow K$  is **alternating** if some  $v_i = v_j$  for  $i \neq j$  then

$$\omega(v_1, \dots, v_k) = 0$$

**Example:**  $\begin{vmatrix} 5 & 0 & 0 \\ 4 & 3 & 3 \\ 2 & 6 & 6 \end{vmatrix} = 0$

**Definition -  $\Lambda(V)$ :** All alternating multilinear  $k$ -forms

- **Note:**  $\Lambda(V)$  is a subspace of  $\text{ML}_k(V)$ 
  - In particular  $0 \in \Lambda(V) \subseteq \text{ML}_k(V)$ . This is the 0 mapping

**Consider:** For a fixed  $V$  with dimension  $n$ , what is  $\Lambda(V)$ ?

**Definition - Permutation:** 1-1, onto mapping  $\sigma : [n] \rightarrow [n]$

**Example:** For  $n = 4$ ,  $(1, 2, 3, 4) \rightarrow (2, 4, 1, 3)$  corresponds to  $\sigma(1) = 2, \sigma(2) = 4, \sigma(3) = 1, \sigma(4) = 3$

We can also compose permutations

Let  $\tau(1) = 1, \tau(2) = 3, \tau(3) = 3, \tau(4) = 4$ . Then

- $\tau \circ \sigma(1) = 3$
- $\tau \circ \sigma(2) = 4$
- $\tau \circ \sigma(3) = 1$
- $\tau \circ \sigma(4) = 2$

Furthermore, every permutation  $\sigma : [n] \rightarrow [n]$  has an inverse function  $\sigma^{-1}$ , satisfying  $\sigma^{-1}\sigma = \text{id}$

- $\sigma^{-1}(1) = 3$
- $\sigma^{-1}(2) = 1$
- $\sigma^{-1}(3) = 4$
- $\sigma^{-1}(4) = 2$

**Definition - Transposition:** A permutation  $\tau$  that swaps two entries and fixes everything else

- **Note** For a transposition  $\tau$ , we have that  $\tau^{-1} = \tau \implies \tau^2 = \text{id}$

Let  $S_n$  be the set of all permutations of  $[n]$

**Claim:**  $S_n$  has  $n!$  elements

*Proof:* on the homework

**Claim:** For all  $n \geq 1$ , every  $\sigma \in S_n$  can be written as a (possibly empty) product of transpositions

$$\sigma = \tau_r \circ \dots \circ \tau_1$$

*Proof by Induction:*

Base Case: For  $n = 1$ , we have  $S_1 \implies S_1 = \{\text{id}\}$  where  $\text{id}$  is the product of no transpositions

Base Case: For  $n = 2$ , we have  $S_2 \implies S_2 = \{\text{id}, \tau_{1,2}\}$  where  $\tau_{1,2}$  swaps 1, 2

IH: Suppose for an arbitrary  $n$ , every  $\sigma \in S_n$  can be written as a (possibly empty) product of transpositions

IS: Choose an arbitrary  $\sigma \in S_{n+1}$

- Case 1: Suppose  $\sigma(n+1) = n+1$ . Then we can look at the remaining elements  $[n]$ , which by IH, any  $\sigma \in S_n$  can be written as a product of transpositions
- Case 2: Suppose  $\sigma(n+1) = j$  for some  $j \leq n$ . Then let  $\tau$  be the transposition swapping  $j, n+1$ . Then  $\tau \in S_{n+1}$  and  $\tau\sigma(n+1) = n+1$

By using Case 1, we can write

$$\tau\sigma = \tau_r \circ \dots \circ \tau_1 \implies \tau\tau\sigma = \sigma = \tau(\tau_r \circ \dots \circ \tau_1)$$

**Definition -  $\epsilon$ :** Is a function  $\epsilon : S_k \rightarrow \{-1, +1\}$

$$\epsilon(\sigma) = \begin{cases} +1 & \sigma \text{ is even} \\ -1 & \sigma \text{ is odd} \end{cases}$$

- **Note:** Any  $\sigma \in S_k$  permutes  $\{x_1, \dots, x_k\} \rightarrow \{x_{\sigma(1)}, \dots, x_{\sigma(k)}\}$

**Notation:** For each  $\omega \in \Lambda_k(V)$  and each  $\sigma \in S_k$ , we let

$$(\sigma\omega)(x_1, \dots, x_k) = \omega(x_{\sigma(1)}, \dots, x_{\sigma(k)})$$

**Example:** For  $k = 3$ , suppose  $\sigma(x_1, x_2, x_3) = (x_3, x_1, x_2)$

Then for any  $(v_1, v_2, v_3) \in V^3$ , we have

$$(\sigma\omega)(v_1, v_2, v_3) = \omega(v_3, v_1, v_2)$$

**Theorem:** If  $\omega \in \Lambda(V)$  and  $\sigma \in S_k$ , then

$$(\sigma\omega) = \epsilon(\sigma)\omega$$

Meaning that for all  $(v_1, \dots, v_k) \in V^k$ , we have

$$(\sigma\omega)(v_1, \dots, v_k) = \omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = \epsilon(\sigma)\omega(v_1, \dots, v_k)$$

*Proof:* Since  $\sigma$  is a product of transpositions, it suffices to prove that when  $\sigma$  is a transposition  $\tau$  swapping  $i, j$

- **Note:**  $\epsilon(\tau) = -1$

We need to show that for all  $(v_1, \dots, v_k) \in V^k$ , we have

$$\omega(v_{\tau(1)}, \dots, v_{\tau(k)}) = -\omega(v_1, \dots, v_k)$$

Notation wise, let  $\bar{\omega}(x, y)$  denote

$$\omega(v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_{j-1}, y, v_{j+1}, v_k)$$

Note that  $\bar{\omega}(x+y, x+y) = 0$  since  $\omega$  is alternating

Thus we see that

$$\bar{\omega}(x+y, x+y) = \bar{\omega}(x, x) + \bar{\omega}(x, y) + \bar{\omega}(y, x) + \bar{\omega}(y, y) = 0$$

This shows that

$$\bar{\omega}(x, y) = \bar{\omega}(y, x) \implies \bar{\omega}(v_j, v_i) = -\bar{\omega}(v_i, v_j)$$

**Theorem:** Suppose  $\{v_1, \dots, v_k\} \subseteq V$  is linearly dependent. Then for all  $\omega \in \Lambda_k(V)$ , we have

$$\omega(v_1, \dots, v_k) = 0$$

*Proof:* Suppose that  $v_i$  is a linear combination of the other vectors in the basis

$$v_i = \sum_{j \neq i} a_j v_j$$

Then we see that

$$\omega(v_1, \dots, v_{i-1}, (\sum_{j \neq i} a_j v_j), v_{i+1}, \dots, v_k) = \sum_{j \neq i} a_j \underbrace{\omega(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_k)}_{\text{by multilinearity}} = 0$$

This last part follows since there are 2  $v_j$  and  $\omega$  is alternating

**Upshot:** Alternating multilinear k-forms preserve linear dependence

**The Big Count:** Suppose  $\dim(V) = n$  and  $V$  has a basis  $B = \{b_1, \dots, b_n\}$ . Take any  $\omega \in \Lambda_k(V)$ . Then for any  $(v_1, \dots, v_n) \in V^n$

$$\omega(v_1, \dots, v_n) = (\sum_j a_{1j} b_j, \dots, \sum_j a_{nj} b_j) = \underbrace{\sum_{1 \leq j_1, \dots, j_n \leq n} a_{1j_1}, \dots, a_{nj_n} \omega(b_{j_1}, \dots, b_{j_n})}_{n^n \text{ terms}}$$

- **Note:** This follows from  $v_i = \sum_j a_{ij} b_j$

However, the terms in the summation above are non-zero only when  $j_1, \dots, j_n$  are distinct

Thus the terms in the summation can be viewed as permutations  $\sigma : \{1, \dots, n\} \rightarrow \{j_1, \dots, j_n\}$

Thus the summation actually only involves  $n!$  terms

$$\sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \underbrace{\omega(b_{\sigma(1)}, \dots, b_{\sigma(n)})}_{(\sigma_\omega)(b_1, \dots, b_n)}$$

Finally, we see that this is equal to

$$\sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \epsilon(\sigma) \omega(b_1, \dots, b_n) = \omega(b_1, \dots, b_n) \underbrace{\sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \epsilon(\sigma)}_{\in K}$$

### 6.1.1 Consequences of the Big Count

Let  $\{b_1, \dots, b_n\}$  be a basis of  $V$

1. If  $\omega(b_1, \dots, b_n) = 0 \implies \omega = 0$ , the 0 mapping
2. If  $\omega(b_1, \dots, b_n) \neq 0$  for some basis, then  $\omega(c_1, \dots, c_n) \neq 0$  for any other basis of  $V$ ,  $\{c_1, \dots, c_n\}$
3. If  $w, w' \neq 0$  are 2 different elements in  $\Lambda_n(V)$ , then they are linearly dependent
  - This means that  $w' = cw$  for some  $c \in K \implies \dim(\Lambda_n(V)) \leq 1$

**Theorem:** If  $\dim(V) = n \geq 1$ , then  $\dim(\Lambda_n(V)) = 1$

- This means that there is some non-zero  $\omega \in \Lambda_n(V)$

*Proof by Induction on  $k \leq n$*

We will show that there is some non-zero  $\omega \in \Lambda_n(V)$

Base case  $k = 1$ . Recall that  $\text{ML}_1(V) = V^*$ , which has dimension  $\geq 1$

IH: Assume there exists a non-zero  $\omega \in \Lambda_k(V)$  with  $k < n$

IS: Show that there is a  $\hat{\omega} \in \Lambda_{k+1}(V)$  where  $\hat{\omega} \neq 0$

**TODO FINISH THIS PROOF**

Now take a linear transformation  $T : V \rightarrow V$  that induces another linear transformation  $T^* : \Lambda_n(V) \rightarrow \Lambda_n(V)$  defined by

$$T^*(\omega) : V^n \rightarrow K \quad T^*(\omega)(u_1, \dots, u_n) = \omega(T(u_1), \dots, T(u_n))$$

Clearly,  $T^* : \Lambda_n(V) \rightarrow \Lambda_n(V)$  is just scalar multiplication, meaning that there is some  $d \in K$  such that

$$\forall \omega \in \Lambda_n(V), T^*(\omega) = d\omega$$

**Definition - Determinant:** The **determinate** of  $T$  is exactly the  $d$  above. That is  $\det(T) = d \in K$

**Properties of  $\det(\mathbf{T})$ :**

1. Suppose  $T : V \rightarrow v$  is multiplication by  $a$ . That is  $T(v) = av$

$$\text{Then } T^* : \Lambda_n(V) \rightarrow \Lambda_n(V) \quad T^*(\omega)(u_1, \dots, u_n) = \omega(T(u_1), \dots, T(u_n)) = \omega(au_1, \dots, au_n) = a^n \omega(u_1, \dots, u_n)$$

$$\text{Thus } T^*(\omega) = a^n \omega \text{ for } \omega \in \Lambda_n(V)$$

$$\text{Here } \det(T) = a^n$$

**Special Cases:**

- $\text{id} : V \rightarrow V \quad \forall v \in V, \text{id}(v) = v \implies \det(\text{id}) = 1$
  - $\text{zero} : V \rightarrow V \quad \forall v \in V, \text{zero}(v) = 0 \implies \det(\text{zero}) = 0$
2. Take two linear transformations  $S, T : V \rightarrow V$

Then the composition  $S \circ T : V \rightarrow V$  is also a linear transformation

$$\text{We claim that } \det(S \circ T) = \det(S) \det(T)$$

For any  $\omega \in \Lambda_n(V)$ , we have that

$$\begin{aligned}
(S \circ T)^*(\omega)(u_1, \dots, u_n) &= \omega(S \circ T(u_1), \dots, S \circ T(u_n)) \\
&= \det(S)\omega(T(u_1), \dots, T(u_n)) \\
&= \det(S)\det(T)\omega(u_1, \dots, u_n) \\
\implies (S \circ T)^*(\omega) &= \det(T)\det(S)\omega
\end{aligned}$$

**Special Cases:**

- Suppose that  $T : V \rightarrow v$  is invertible, then clearly  $T^{-1} \circ T = \text{id} \implies \det(T^{-1} \circ T) = \det(\text{id}) = 1$

$$\text{Thus } \det(T^{-1})\det(T) = 1 \implies \det(T^{-1}) = \frac{1}{\det(T)}$$

Thus  $T$  is invertible if and only if  $\det(T) \neq 0$

**TFAE Theorem:**

1.  $T$  is an isomorphism
2.  $T$  is invertible
3.  $\text{rank}(T) = n$
4.  $\det(T) \neq 0$

*Proof:*  $1 \iff 2 \iff 3$  is shown by the previous proof

To show that 4 must hold, by the special case before, if any of 1, 2, 3 hold, then  $\det(T) \neq 0$

Now we show that if 1, 2, 3 fail, then  $\det(T) = 0$

Since 3 fails, we must have that  $\text{rank}(T) = \dim(\text{Im}(T)) < \dim(V) = n$

Now choose any  $\omega \in \Lambda_n(V)$  and choose any  $(u_1, \dots, u_n) \in V^n$

We see that

$$(T^*)(\omega)(u_1, \dots, u_n) = \omega(T(u_1), \dots, T(u_n))$$

Since  $\{T(u_1), \dots, T(u_n)\}$  are  $n$  vectors in  $\mathfrak{S}(T)$ , they must be linearly dependent

Thus since  $\omega$  respect linearly dependency, we see that

$$\omega(T(u_1), \dots, T(u_n)) = 0$$

Thus for any  $\omega \in \Lambda_n(V)$ , we must have

$$T^*(\omega) = 0 \implies \det(T) = 0$$

### 6.1.2 Matrix Representation

Now take  $A \in M_{n \times n}(K)$

We know that  $A$  encodes a linear transformation, namely  $T_A : K^n \rightarrow K^n$

Thus  $\det(A) = \det(T_A)$

**Consequences:**

1.  $\det(I_n) = 1$  since  $T_{I_n} = \text{id} : K^n \rightarrow K^n$  and  $\det(\text{id}) = 1$
2.  $\det(O) = 0$  since  $T_{\text{zero}} = \text{zero} : K^n \rightarrow K^n$  and  $\det(\text{zero}) = 0$



3. For  $A, B \in M_{n \times n}(K)$ ,  $\det(AB) = \det(A) \det(B)$

The linear transformation  $T_{AB}$  is described by the composition  $T_A \circ T_B \implies \det(T_{AB}) = \det(T_A) \det(T_B) = \det(A) \det(B)$

**TFAE Theorem:** For  $A \in M_{n \times n}(K)$ , the following are equivalent

1.  $T_A$  is an isomorphism
2.  $A$  is invertible
3.  $\text{rank}(A) = n$
4.  $\det(A) \neq 0$

Now we can compute  $\det(A)$  for  $A \in M_{n \times n}(K)$  by applying The Count Theorem

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

**Example:** For  $n = 2$  we have  $S_2 = \{\text{id}, \tau\}$  where  $\epsilon(\text{id}) = 1$  and  $\epsilon(\tau) = -1$

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \epsilon(\text{id}) a_{11} a_{22} - \epsilon(\tau) a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21}$$

**Note:** Since any linear transformation can be represented as a matrix  $A$ , we have that

$$\det(A) = \sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \det({}^t A)$$

## 7 Determinants

Determinants only make sense for square  $n \times n$  matrices. We define the **determinate** as

- $1 \times 1 \implies \det(a) = a$
- $2 \times 2 \implies \det : M_{2 \times 2}(K) \rightarrow K$  where  $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$
- $3 \times 3 \implies \det : M_{3 \times 3}(K) \rightarrow K$  where  $\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$

**Example:**  $\begin{vmatrix} 2 & 1t \\ 3 & 5t \end{vmatrix} = 2(5t) - 3(t) = 10t - 3t = 7t$

**Example:**  $\begin{vmatrix} a+a' & b \\ c+c' & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b \\ c' & d \end{vmatrix}$

- **Upshot:** Freezing a column gives us linearity with the other column

**Example:**  $\begin{vmatrix} b & a \\ d & c \end{vmatrix} = bc - ad = -1 \begin{vmatrix} a & b \\ c & d \end{vmatrix}$

- **Upshot:** Switching columns changes the sign of the determinant

**Example:**  $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$

**Example:**  $\begin{vmatrix} 5 & 1 & 2 \\ 3 & 2 & 0 \\ 4 & 1 & 3 \end{vmatrix} + \begin{vmatrix} 5 & 2 & 2 \\ 3 & -1 & 0 \\ 4 & 0 & 3 \end{vmatrix} = 11 - 25 = -14 = \begin{vmatrix} 5 & 3 & 2 \\ 3 & 1 & 0 \\ 4 & 1 & 3 \end{vmatrix}$

## 7.1 Row Determinants

We look at what exactly are row reductions and their impact on determinants

Suppose  $A^1, \dots, A^n$  are columns of  $A$ . Let  $B$  have the same columns, except two swapped columns

From the rules of determinants, we have that

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

Now consider replacing a column by itself plus some scalar multiple of another column

That is  $B = [A^1 + cA^2, A^2, \dots]$ . Then we see that

$$\det(B) = \det(A^1, A^2, \dots) + c \det(A^2, A^2, A^3, \dots) = \det(A)$$

Finally, since  $\det({}^t A) = \det(A)$ , these equalities work under row operations as well

## 8 Symmetric, Hermitian, Unitary Operators

**Definition - Operator:** A linear transformation  $T : V \rightarrow V$

Consider when  $V$  is a  $K$ -Vector Space and  $\langle, \rangle$  is a positive definite scalar product

Recall that

- $\|v\| = \sqrt{\langle v, v \rangle}$
- Gram-Schmidt process takes a basis  $B$  and produces an orthonormal basis
- $V^*$  is the set of linear transformation  $\phi : V \rightarrow R$  and that  $V \approx V^*$  under

$$L^* : V \rightarrow V^* \quad L^*(w) : V \rightarrow R \quad L^*(w)(v) = \langle v, w \rangle \forall v \in V$$

### Fundamental Fact

For any operator  $A : V \rightarrow V$ , there exists a unique operator  $B : V \rightarrow V$  such that for all  $v, w \in V$

$$\langle Av, w \rangle = \langle v, Bw \rangle$$

Here  $B$  is called the **transpose** of  $A$ , namely  $B = {}^t A$

Now given  $A$ , how do we find  $B$ ?

Take  $w \in V$  and let  $L_w^A : V \rightarrow R$  be defined by  $L_w^A(v) = \langle Av, w \rangle$

- It can be shown that  $L_w^A$  is a linear transformation. Thus  $L_w^A \in L^*$

Furthermore, since  $L^* : V \rightarrow V^*$  is an isomorphism, there exists a unique  $w' \in V$  such that

$$L^*(w') = L_w^A$$

- Importantly,  $L^*(w')$  is the same function as  $L_w^A$

But then we have that

$$\forall v \in V \quad L_w^A(v) = L^*(w')(v) \implies \langle Av, w \rangle = \langle v, w' \rangle$$

Now we define  $B : V \rightarrow V$  such that  $B(w) = w'$ . Thus we have

- **Note:** It can be shown that  $B$  is a linear transformation

$$\langle Av, w \rangle = \langle v, Bw \rangle$$

Furthermore we have that

$$\langle Av, w \rangle = \langle v, {}^tAw \rangle$$

**Definition - Symmetric:** An operator  $A : V \rightarrow V$  is **symmetric** if and only if any  $n \times n$  matrix representing  $A$  is a symmetric matrix

$${}^tA = A$$

**Definition - Unitary:** An operator  $A : V \rightarrow V$  is **unitary** if

$$\forall v, w \in V \quad \langle Av, Aw \rangle = \langle v, w \rangle$$

- **Note:** We say  $A$  is **norm-preserving** if for all  $v \in V$ ,  $\langle Av \rangle = \langle v \rangle$

**Proposition:**  $A$  Is unitary if and only if  $A$  is norm-preserving

*Proof:*  $\implies$  Assume that  $A$  is unitary and choose  $v \in V$ . Clearly

$$\|Av\|^2 = \langle Av, Av \rangle = \langle v, v \rangle = \|v\|^2$$

$\Leftarrow$  Assume  $A$  is norm-preserving and choose  $v, w \in V$ . Then

$$\langle v + w, v + w \rangle - \langle v - w, v - w \rangle = 4\langle v, w \rangle$$

Similarly, we have that

$$\langle A(v + w), A(v + w) \rangle - \langle A(v - w), A(v - w) \rangle = 4\langle Av, Aw \rangle$$

Thus we have that  $\|v + w\|^2 - \|v - w\|^2 = \|A(v + w)\|^2 - \|A(v - w)\|^2 \implies \langle v, w \rangle = \langle Av, Aw \rangle$

## 9 Eigenvectors and Eigenvalues

For the rest of the course, we consider a Finite Dimensional Vector Space  $V$  over  $R$  or  $C$  with a positive definite  $\langle, \rangle$  that is

- Scalar if we are dealing with  $R$
- Hermitian if we are dealing with  $C$

We study operators  $A : V \rightarrow V$

Some operators are easy to understand. For example consider  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$   $A = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$

Then we have  $A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$  and  $A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$

Thus  $A$  stretches  $x$  by 5 and stretches  $y$  by 3  $\implies A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5x \\ 3y \end{bmatrix}$

However, some operators are more difficult to understand. For example, consider  $B : \mathbb{R}^2 \rightarrow \mathbb{R}^2$   $B = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

Then we have  $B \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$  and  $B \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

Clearly using the standard basis, it's difficult to describe the operator

However, if we instead use the basis  $B = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ , we get

$$B \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad B \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Which is much easier to describe the operator

**Definition - Eigenvector:** Nonzero  $v \in V$  such that  $Av = \lambda v$  for some  $\lambda \in K$

**Definition - Eigenvalue:** Some  $\lambda \in K$  such that  $Av = \lambda v$  for some nonzero  $v \in V$

So for the examples above

- Eigenvalues for  $A$  are  $\lambda \in \{5, 3\}$  corresponding to  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$  and  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$
- Eigenvalues for  $B$  are  $\lambda \in \{4, 2\}$  corresponding to  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  and  $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

**Definition - Eigenspace:** For any  $\lambda$ , the **eigenspace** is

$$W_\lambda = \{v \in V \mid Av = \lambda v\}$$

- **Note:**  $W_\lambda$  is a subspace of  $V$

**Example:** Rotations. Fix an angle  $\theta$

Define the operator  $R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  (rotation counter-clockwise) by

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

- **Note:** There are no eigenvalues in  $\mathbb{R}$  and thus no eigenvectors
- **Note:** However, for  $v \in \mathbb{C}^2$ ,  $R_\theta(v) = e^{i\theta}v$

Now show how to determine if  $\lambda$  is an eigenvalue

By definition,  $\lambda$  is an eigenvalue if and only if there is some nonzero  $v \in V$  such that  $Av = \lambda v$

But this means that  $(\lambda I - A)v = 0$

Which means that  $v \in \text{Ker}(\lambda I - A)$

Which means that  $\text{Ker}(\lambda I - A) \neq \{O\}$

Which means that  $\det(\lambda I - A) = 0$

**Definition - Characteristic Polynomial:** For an operator  $A$ , we define the **characteristic polynomial**

$$P_A(t) = \det(tI - A)$$

- **Note:**  $\lambda$  is an eigenvalue of  $A$  if and only if  $P_A(\lambda) = 0$

**Example:** Consider  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$       $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$ . Determine the eigenvalues of  $A$

We look at

$$\det(\lambda I - A) = \lambda^2 - 4\lambda + 3 - 8 = 0 \implies t \in \{5, -1\}$$

Thus eigenvalues of  $A$  are  $\{5, -1\}$

**Note:** By the Fundamental Theorem of Algebra, every non-constant monic polynomial has a root in  $\mathbb{C}$ . Thus every  $A : V \rightarrow V$  has an eigenvalue (hence eigenvector) in  $\mathbb{C}$