

# Math 110, Spring 2019

## Contents

<b>1</b>	<b>Vector Space</b>	<b>3</b>
1.1	Vector Space over a field and subspace . . . . .	3
1.1.1	Lists (and vector spaces of lists) . . . . .	3
1.2	Subspaces . . . . .	4
1.3	Direct Sum . . . . .	8
<b>2</b>	<b>Finite Dimensional Vector Spaces</b>	<b>10</b>
2.1	Linear Dependence and Independence . . . . .	10
2.2	Bases and Dimension . . . . .	13
2.2.1	Dimension . . . . .	14
2.2.2	Dimension of a Sum . . . . .	15
<b>3</b>	<b>Linear Maps</b>	<b>17</b>
3.1	Linear Maps as Vector Space . . . . .	17
3.2	Null Space and Range . . . . .	18
3.3	Matrix Notation . . . . .	20
3.4	Matrix Representation . . . . .	21
3.5	Invertibility and Isomorphism . . . . .	22
3.5.1	Linear Operators . . . . .	24
3.6	Duality . . . . .	25
3.6.1	Dual Maps . . . . .	26
3.6.2	Matrix Representation of the dual map . . . . .	29
<b>4</b>	<b>Polynomials</b>	<b>31</b>
4.1	Axler's Recap on Polynomial . . . . .	31
4.2	Zero of polynomials and their algebraic manifestations . . . . .	31
<b>5</b>	<b>Eigenvalues, Eigenvectors, and Invariant Subspaces</b>	<b>34</b>
5.1	Invariant Subspaces . . . . .	34
5.1.1	Restriction Operators . . . . .	35
5.2	Eigenvectors and Upper-Triangular Matrices . . . . .	35
5.2.1	Polynomials in T . . . . .	35
5.3	Eigenspaces and Diagonal Matrices . . . . .	39
<b>6</b>	<b>Inner Product Spaces</b>	<b>41</b>
6.1	Inner Product and Norms . . . . .	41
6.2	Orthogonality . . . . .	45
6.3	Orthogonality and Orthogonal Projections . . . . .	47
<b>7</b>	<b>Operators on Inner Product Spaces</b>	<b>50</b>
7.1	Self-Adjoint and Normal Operators . . . . .	50
7.1.1	Matrix representation . . . . .	52

---

7.2	Spectral Theorem . . . . .	54
7.2.1	Over Complex Vector Space . . . . .	54
7.2.2	Over Real Vector Space . . . . .	55

# 1 Vector Space

## 1.1 Vector Space over a field and subspace

Recall that  $(\mathbb{F}, +, \cdot)$  or  $\langle \mathbb{F}, +, \cdot \rangle$ , where  $\mathbb{F}$  is a set, and  $+, \cdot$  are binary operations. We know that  $(\mathbb{F}, +)$  and  $(\mathbb{F} \setminus \{0\}, \cdot)$  and  $+, \cdot$  satisfy distributivity.

### Definition 1.1

$V$  is a vector space over a field  $\mathbb{F}$  if  $V$  is equipped with vector addition  $+: V \times V \rightarrow V$  and scalar multiplication  $\cdot: \mathbb{F} \times V \rightarrow V$ .

### 1.1.1 Lists (and vector spaces of lists)

#### Example 1.2

$\mathbb{R}^n, \mathbb{C}^n$ , or generally  $\mathbb{F}^n$ .

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_i \in \mathbb{R} \ \forall \ j = 1, 2, \dots, n\}$$

$$\mathbb{F}^n = \{(x_1, x_2, \dots, x_n) : x_i \in \mathbb{F} \ \forall \ j = 1, 2, \dots, n\}$$

We claim that  $\mathbb{F}^n$  is a vector space over  $\mathbb{F}$  provided  $\mathbb{F}$  is a field. We can define addition and scalar multiplication as

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) := (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

$$\alpha \cdot (x_1, x_2, \dots, x_n) = (\alpha \cdot x_1, \alpha \cdot x_2, \dots, \alpha \cdot x_n) \quad \alpha \cdot x_i \in \mathbb{F}$$

What rules / axioms should we impose?

- Commutativity  
 $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v} \ \forall \ \mathbf{v}, \mathbf{w} \in V.$
- Associativity  
 $(\mathbf{v} + \mathbf{w}) + \mathbf{u} = \mathbf{v} + (\mathbf{w} + \mathbf{u}) \ \forall \ \mathbf{v}, \mathbf{w}, \mathbf{u} \in V.$
- Additive Identity  
 $\exists \ \mathbf{0} \in V : \mathbf{v} + \mathbf{0} = \mathbf{v} + \mathbf{0} = \mathbf{v}$
- Additive Inverse  $\forall \mathbf{v} \in V \ \exists \ \mathbf{w} \in V : \mathbf{v} + \mathbf{w} = \mathbf{0}.$
- (Mixed) Scalar Multiplication Rules  
 $1 \cdot \mathbf{v} \in \mathbf{v} \quad \forall \mathbf{v} \in V$
- Distributivity:  

$$(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \beta \cdot \mathbf{v} \quad \forall \ a, b \in \mathbb{F} \quad \forall \ \mathbf{v} \in V$$

$$\alpha \cdot (\mathbf{v} + \mathbf{w}) = \alpha \cdot \mathbf{v} + \alpha \cdot \mathbf{w} \quad \forall \ a \in \mathbb{F} \quad \forall \ \mathbf{v}, \mathbf{w} \in V$$

Now we can check that these rules hold in  $\mathbb{F}^2$ :

$$\begin{aligned}(0, 0, \dots, 0) + (x_1, x_2, \dots, x_n) &= (x_1, x_2, \dots, x_n) \\ (x_1, x_2, \dots, x_n) + (-x_1, -x_2, \dots, -x_n) &= (0, 0, \dots, 0)\end{aligned}$$

**Basic Observation 0** is unique

*Proof.* Suppose  $\mathbf{0}_1$  and  $\mathbf{0}_2$  are both identity element with respect to  $+$ :

$$\mathbf{0}_1 = \mathbf{0}_1 + \mathbf{0}_2 + \mathbf{0}_2$$

A contradiction. ■

Additive inverse are unique, i.e., if  $\mathbf{v} + \mathbf{w} = \mathbf{0}$  and  $\mathbf{v} + \mathbf{u} = \mathbf{0}$ , then  $\mathbf{u} = \mathbf{w}$ .

*Proof.* Suppose  $\mathbf{v} + \mathbf{w} = \mathbf{0}$  and  $\mathbf{v} + \mathbf{u} = \mathbf{0}$ , then

$$\mathbf{w} = \mathbf{w} + \mathbf{0} = \mathbf{w} + (\mathbf{v} + \mathbf{u}) = (\mathbf{w} + \mathbf{v}) + \mathbf{u} = \mathbf{0} + \mathbf{u} = \mathbf{u}$$

A contradiction. ■

## Additive Inverse

$$\begin{aligned}(-1) \cdot \mathbf{v} + \mathbf{v} &= (-1) \cdot \mathbf{v} + 1 \cdot \mathbf{v} \\ &= ((-1) + 1) \cdot \mathbf{v} \\ &= 0 \cdot \mathbf{v} \\ 0 \cdot \mathbf{v} + 0 \cdot \mathbf{v} &= (0 + 0) \cdot \mathbf{v} \\ &= 0\mathbf{v} \implies \boxed{0 \cdot \mathbf{v} = \mathbf{0}}\end{aligned}$$

Additive inverse  $\implies 0 \cdot \mathbf{v} = \mathbf{0}$  on both sides.

## 1.2 Subspaces

### Definition 1.3

$V$  is a vector space over a field  $\mathbb{F}$ , Let  $W \subseteq V$ .

$W$  is called a subspace of  $V$  if  $W$  equipped with the same operations  $+, \cdot$  inherited from  $V$  is still a vector space.

**Is it enough for  $W$  to be just a subset of  $V$ ?**

Suppose  $V = \mathbb{R}^3$  is a vector space over  $\mathbb{R}$ . Let  $W := \{(1, 1, 1)\}$ , the additive inverse doesn't exist. Note that  $W$  is not closed in addition and scalar multiplication.

$$W := \{(x, 0, 0) : x_1 \in \mathbb{R}\}$$

**Why is  $\mathbf{0}$  in every subspace?**

We know that a vector space is a *non empty* set, and  $W$  is closed under multiplication, so since  $0 \in \mathbb{F}$ , therefore  $0 \cdot \mathbf{v} = \mathbf{0} \in W$ .

**Remark 1.4**

If  $\mathbf{v} + \mathbf{w} = \mathbf{v}$ , for some  $\mathbf{v} \in V$ , then  $\mathbf{w} = \mathbf{0}$ .

*Proof.* Suppose  $\mathbf{v} + \mathbf{w} = \mathbf{v}$ , then  $-\mathbf{v} + \mathbf{v} + \mathbf{w} = -\mathbf{v} + \mathbf{v} \implies \mathbf{0} + \mathbf{w} = \mathbf{0} \implies \mathbf{w} = \mathbf{0}$  ■

We recall that  $\mathbb{F}^n$  is defined as

$$\mathbb{F}^n = \{(x_1, x_2, \dots, x_n) : x_i \in \mathbb{F}\}$$

We can define  $\mathbb{F}^S$  for  $S$  being a set as  $\mathbb{F}^S = \left\{ f : \underbrace{S}_{\text{no structure needed}} \rightarrow \underbrace{\mathbb{F}}_{\text{field}} \right\}$

We can define addition and multiplication as

$$(f + g)(s) := f(s) + g(s) \quad \forall s \in S$$

$$(\lambda \cdot f)(s) := \lambda \cdot f(s) \in \mathbb{F}$$

Suppose  $S = \{1, 2, 3\}$ , what is  $\mathbb{F}^S$  or  $\mathbb{R}^S$ ? We can thought of  $\mathbb{R}^S$  as  $\mathbb{R}^3$ ..... why?

**Remark 1.5**

We can conclude  $\mathbb{F}^S \cong \mathbb{F}^{|S|}$ , where  $|S|$  is the cardinality of  $S$ . If  $S$  is finite.

What is  $\mathbb{R}^{\mathbb{N}}$ ? ← the set of all of all real sequences.

**Remark 1.6**

In the book we uses  $\mathbb{R}^{\infty}$ , we can conclude that

$$\mathbb{R}^{\infty} \cong \mathbb{R}^{\mathbb{N}}$$

We say that  $W$  is a subspace of  $\mathbb{R}^{\infty}$  with  $+, \cdot$ .

$$W := \left\{ s : \lim_{n \rightarrow \infty} s(n) = 0 \right\}$$

*Proof.* We can see that if  $\lim_{n \rightarrow \infty} s(n) = 0$  and  $\lim_{n \rightarrow \infty} t(n) = 0$ , then  $\lim_{n \rightarrow \infty} (s + t)(n) = 0$ .

If  $\lambda \in \mathbb{R}$ , then  $\lim_{n \rightarrow \infty} s(n) = 0 \implies \lim_{n \rightarrow \infty} (\lambda \cdot s)(n) = 0$ .

The zero sequence is in  $W$  so  $\mathbf{0} \in V$ . Therefore  $W$  is a subspace of  $V$ . ■

**Theorem 1.7**

$W$  is a subspace of  $V$  iff  $W$  is closed under addition, multiplication by scalar multiplication by scalars, and  $\mathbf{0} \in V$ .

Since the operation is inherent from vector space  $V$ , we do not need to verify the other property since they all for all  $V$  and  $W$  is a subspace of  $V$ .

**How do we form new subspaces from existing ones?**

**Theorem 1.8**

Suppose  $W_1, W_2$  are subspaces of  $V$ , then  $W_1 \cap W_2$  is a subspace of  $V$ .

*Proof.* We know that  $W_1, W_2$  are subspaces of  $V$ , therefore  $\mathbf{0} \in W_1$  and  $\mathbf{0} \in W_2$ , then  $\mathbf{0} \in W_1 \cap W_2$ .

Suppose  $\mathbf{v}, \mathbf{u} \in W_1 \cap W_2$ , we know that  $\mathbf{v}, \mathbf{u} \in W_1$  and  $\mathbf{v}, \mathbf{u} \in W_2$ . Since  $W_1, W_2$  is a subspace, therefore  $\mathbf{u} + \mathbf{v} \in W_1$  and  $\mathbf{u} + \mathbf{v} \in W_2 \implies \mathbf{u} + \mathbf{v} \in W_1 \cap W_2$ , therefore  $W_1 \cap W_2$  is closed under vector addition.

Suppose  $\alpha \in \mathbb{F}$  and  $\mathbf{v} \in W_1 \cap W_2$ . We know that  $\alpha \cdot \mathbf{v} \in W_1$  and  $\alpha \cdot \mathbf{v} \in W_2$  since they are both subspaces of  $V$ . Therefore we conclude  $\alpha \cdot \mathbf{v} \in W_1 \cap W_2$ , therefore  $W_1 \cap W_2$  is closed under multiplication.

Therefore  $W_1 \cap W_2$  is a subspace of  $V$ . ■

**Proposition 1.9**

The union of two subspaces of  $V$  are generally not a subspace of  $V$

*Proof.* We can see that  $\text{span}\{\mathbf{e}_1\}$  and  $\text{span}\{\mathbf{e}_2\}$  is not a subspace of  $\mathbb{R}^2$  as  $(1, 1) \notin W_1 \cup W_2$  ■

**Theorem 1.10**

Union of two subspaces of  $V$  is a subspace of  $V$  if and only if one of the subspaces is contained in the other.

*Proof.* The proof is left as an exercise. ■

**Theorem 1.11**

$W_1 + W_2$  is a subspace of  $V$ .

*Proof.* (identity)  $\mathbf{0} \in W_1$  and  $\mathbf{0} \in W_2 \implies \mathbf{0} + \mathbf{0} = \mathbf{0} \in W_1 + W_2$ .

(closure under addition) Suppose  $\mathbf{w}_1 + \mathbf{w}_2 \in W_1 + W_2$  and  $\tilde{\mathbf{w}}_1 + \tilde{\mathbf{w}}_2 \in W_1 + W_2$ . We compute  $(\mathbf{w}_1 + \mathbf{w}_2) + (\tilde{\mathbf{w}}_1 + \tilde{\mathbf{w}}_2) = \underbrace{(\mathbf{w}_1 + \tilde{\mathbf{w}}_1)}_{\in W_1} + \underbrace{(\mathbf{w}_2 + \tilde{\mathbf{w}}_2)}_{\in W_2} \implies (\mathbf{w}_1 + \mathbf{w}_2) + (\tilde{\mathbf{w}}_1 + \tilde{\mathbf{w}}_2) \in W_1 + W_2$ .

(closure under scalar multiplication) Suppose  $\mathbf{w}_1 + \mathbf{w}_2 \in W_1 + W_2$ , and  $\lambda \in \mathbb{F}$ , we compute  $\lambda \cdot (\mathbf{w}_1 + \mathbf{w}_2) = \underbrace{(\lambda \cdot \mathbf{w}_1)}_{\in W_1} + \underbrace{(\lambda \cdot \mathbf{w}_2)}_{\in W_2} \implies \lambda \cdot (\mathbf{w}_1 + \mathbf{w}_2) \in W_1 + W_2$  ■

**Remark 1.12**

$W_1 + W_2 + \dots + W_n$  is the smallest subspace containing  $W_1, W_2, \dots, W_n$ .

If  $\tilde{V}$  is a subspace of  $V$  containing  $W_j \forall j$ , since  $\tilde{V}$  is closed under  $+$ ,  $\mathbf{w}_1 + \mathbf{w}_2 + \dots + \mathbf{w}_n \in \tilde{V}$

**Example 1.13**

Suppose  $V = \mathbb{R}^3$ . Let

$$W_1 = \text{span}\{\mathbf{e}_1, \mathbf{e}_2\}, W_2 = \text{span}\{(0, 1, 1)\}, W_3 = \text{span}\{(x, y, z) : x + y + z = 0\}$$

What is  $W_1 + W_2 + W_3$ ?

Note that  $(0, 0, 1) = \underbrace{(0, \frac{1}{2}, \frac{1}{2})}_{\in W_2} + \underbrace{(0, -\frac{1}{2}, \frac{1}{2})}_{\in W_3}$ . We also know that  $(1, 0, 0) \in W_1$  and  $(0, 1, 0) \in W_2$ , therefore  $W_1 + W_2 + W_3 = \mathbb{R}^3$

**Discussion****Definition 1.14**

A vector space, is often denoted as  $(\underbrace{\mathbb{F}}_{\text{scalars}}, \underbrace{V}_{\text{vectors}}, \cdot : \underbrace{\mathbb{F} \rightarrow B}_{\text{scaling}})$

**Example 1.15**

$(\mathbb{R}, \mathbb{R}^n, \cdot : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n)$  is a vector space.

**Example 1.16**

$(\mathbb{R}, \mathbb{R}, \cdot : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R})$  is also a vector space.

**Notion of a field**

Suppose  $F = \{0, 1, 2, 3\}$ . Can  $F$  be a field?

**Definition 1.17**

A subset  $W$  of the vector space  $V$  is a subspace of  $V$  if it satisfy the following:

- 1)  $\mathbf{0} \in W$
- 2)  $+: W \times W \rightarrow W \subseteq V$  (closure under addition)
- 3)  $\cdot : \mathbb{F} \times W \rightarrow W \subseteq V$  (closure under scalar multiplication)

**Example 1.18**

Can we find a subset  $W$  of  $V$  such that  $W$  satisfy property 1), 2) but not 3)?  
Suppose  $W = \{(x, 0) : x \in \mathbb{Z}\}$  the proof is trivial and is left as an exercise.

**Example 1.19**

The set of functions  $\{f : (0, \infty) \rightarrow \mathbb{R}\} = \mathbb{R}^{(0, \infty)}$  is a vector space. We claim that  $W$  is a subspace of  $V$ .

$$W = \{f : (0, \infty) \rightarrow \mathbb{R} : f'(1) = 0\}$$

*Proof.* We begin by verifying the three properties

- 1) The zero function is in  $W$
- 2) Suppose  $f, g \in W$ , then  $(f + g)'(1) = f'(1) + g'(1) = 0 + 0 = 0 \implies f(x) + g(x) \in W$
- 3) Suppose  $f \in W$  and  $\lambda \in \mathbb{F}$ , then  $\lambda \cdot f'(1) = \lambda \cdot 0 = 0 \implies \lambda \cdot f(x) \in W$

Therefore  $W$  is a subspace of  $V$ . ■

**1.3 Direct Sum****Definition 1.20**

Let  $(\mathbb{F}, V, \cdot : \mathbb{F} \times V \rightarrow V)$  be a vector space. Given that  $U_1, U_2, \dots, U_n \subseteq V$  are subspaces of  $V$ , we can define the sum of the subspaces as

$$U_1 + U_2 + \dots + U_n = \{\mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_n : \mathbf{u}_i \in U_i\}$$

*Proof.* 1) We can see that  $\mathbf{0} \in U_i \forall i$ , and  $\mathbf{0} + \mathbf{0} + \dots + \mathbf{0} = \mathbf{0}$

- 2) Suppose  $\mathbf{x} = \sum_{i=1}^n \mathbf{x}_i \in U_i$  and  $\mathbf{y} = \sum_{j=1}^n \mathbf{y}_j \in U_j$ , we can see that  $\mathbf{x} + \mathbf{y} = \sum_{k=1}^n \mathbf{x}_k + \mathbf{y}_k \in U_k$ , therefore it's closed under addition.
- 3) Suppose  $\lambda \in \mathbb{F}$  and  $\mathbf{x} = \sum_{i=1}^n \mathbf{x}_i \in U_i$ , we compute  $\lambda \cdot \mathbf{x} = \lambda \cdot \sum_{i=1}^n \mathbf{x}_i \in U_i$ , therefore it's closed under scalar multiplication. ■

**Definition 1.21**

We say that  $U_1 + U_2 + \dots + U_n$  is a direct sum, denoted as  $U_1 \oplus U_2 \oplus \dots \oplus U_n$  if for every  $\mathbf{v} \in U_1 + U_2 + \dots + U_n$ ,  $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_n$  has a unique representation.

**Remark 1.22**

How best to check  $U_1 + U_2 + \dots + U_n$  is a direct sum?

Check that  $U_i \cap U_j = \{\mathbf{0}\}$ . We will go over in depth later.

What about  $W_1 + W_2 + \dots + W_n$  being a direct sum?



**Theorem 1.23**

The sum of subspaces  $W_1, W_2, \dots, W_n$ :

$$W_1 + W_2 + \cdots + W_n$$

is a direct sum iff  $\mathbf{0}$  can be written in only **one way** as a sum

$$\mathbf{w}_1 + \mathbf{w}_2 + \cdots + \mathbf{w}_n = \mathbf{0}$$

namely  $\mathbf{0} + \mathbf{0} + \cdots + \mathbf{0} = \mathbf{0}$ .

**Remark 1.24**

If  $W_1 \cap W_2 \neq \{\mathbf{0}\}$ ,  $W_1 \cap W_3 \neq \{\mathbf{0}\}$ ,  $W_2 \cap W_3 \neq \{\mathbf{0}\}$ , it is not possible for  $W_1, W_2, W_3$  to be a direct sum. However, the opposite of the proposition is not sufficient for being a direct sum as demonstrated in Remark 1.25.

**Remark 1.25**

$W_1 \cap W_2 = \{\mathbf{0}\}$ ,  $W_1 \cap W_3 = \{\mathbf{0}\}$ ,  $W_2 \cap W_3 = \{\mathbf{0}\}$  and  $W_1 + W_2 + W_3$  being not a direct sum is possible. For example, consider  $\mathbb{R}_2$ , for line  $x = y$ ,  $y = 0$  and  $x = 0$ , we can see that they only have the trivial intersection but they are not a direct sum. (credit: Catherine)

## 2 Finite Dimensional Vector Spaces

### 2.1 Linear Dependence and Independence

#### Definition 2.1

We will work with lists of vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ , then the span of  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  can be defined as

$$\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) := \{\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k\} \forall \alpha_i \in \mathbb{F}$$

If the list happens to cover the entire vector space  $V$ , we call the list a spanning list of  $V$ .

#### Definition 2.2

$V$  is finite dimensional if  $V$  is a span of finitely many vectors.

#### Remark 2.3

$V$  is not finite dimensional is logically equivalent to  $V$  is infinite dimensional.

#### Example 2.4

Consider the vector spaces:  $\mathcal{P}(x) := \{\alpha_0 + \alpha_1 x + \dots + \alpha_k x^k : \alpha_j \in \mathbb{F} \text{ for some } k\}$ . We can see that  $\mathcal{P}(x) \subseteq \mathbb{F}^{\mathbb{F}}$ , and  $\mathcal{P}(x)$  is infinite dimensional.

#### Definition 2.5

We can define the degree of a polynomial, denoted as  $\deg(f(x))$ , is the highest power of  $x$  whose coefficient ( $\alpha_k$ ) is nonzero. The zero function  $f(x) = 0$  has  $-\infty$  degree.

#### Example 2.6

$\mathcal{P}(x)$  is infinite dimensional.

*Proof.* Suppose  $\mathcal{P}(x) = \text{span}(f_1, f_2, \dots, f_k)$ , where  $f_j$  is polynomials, for all  $j$ . Let

$$D := \max \{\deg(f_1), \deg(f_2), \dots, \deg(f_k)\}$$

Suppose  $f(x) = x^{D+1} \in \mathcal{P}(x)$  however,  $x \notin \text{span}(f_1, f_2, \dots, f_k)$ . Since  $f(x)$  is not a linear combination of  $f_1, f_2, \dots, f_k$ . A contradiction, therefore  $\mathcal{P}(x)$  is an infinite dimensional vector space. ■

**Definition 2.7**

$V$  has dimension  $k$  over  $\mathbb{F}$  if you can find vectors  $v_1, v_2, \dots, v_k$  such that

$$\forall \mathbf{v} \in V : \mathbf{v} = \sum f_i \mathbf{v}_i \text{ uniquely}$$

**Definition 2.8**

$\mathcal{P}_d(x) :=$  all polynomials in  $g(x)$  of degree  $\leq d$ .

Note that  $\{1, x, x^2, \dots, x^d\}$  is a spanning list for  $\mathcal{P}_d(x)$

**Definition 2.9**

A list  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in V$  is called **linearly independent** if the equation

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k = 0 \implies \alpha_1, \alpha_2, \dots, \alpha_k = 0$$

**Definition 2.10**

A list  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in V$  is called **linearly dependent** if it is not independent.

**Digression on Logic**

Logic:  $A \implies B$  is equivalent to  $\neg A \vee B$ . Then we know that

$$\neg(A \implies B) \iff (\neg(\neg A \vee B)) \iff A \wedge \neg B$$

**Definition 2.11** (The better definition)

A list  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in V$  is called **linearly dependent** if for equation

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k = 0$$

has a nontrivial solution such that  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \neq 0$

**Example 2.12**

Is  $\{\}$  linearly independent?

By definition, it is linearly independent, because it is not linearly dependent. A set  $S$  is linearly dependent if there exists a finite set of vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  and corresponding scalars  $\alpha_1, \alpha_2, \dots, \alpha_n$  such that there exists at least one  $\alpha_i \neq 0$  so that

$$\sum_{i=1}^n \alpha_i \mathbf{v}_i = \mathbf{0}$$

since  $\alpha_i$  doesn't exist, we know that  $\{\}$  is linearly independent.

**Example 2.13**

Is  $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$  linearly independent on  $\mathbb{R}^3$ ?

$$\begin{aligned} \alpha_1(1, 0, 0) + \alpha_2(0, 1, 0) + \alpha_3(0, 0, 1) &= (0, 0, 0) \\ \implies (\alpha_1, \alpha_2, \alpha_3) &= (0, 0, 0) \\ \implies \alpha_1 = \alpha_2 = \alpha_3 &= 0 \end{aligned}$$

**Remark 2.14**

We can remove vectors from a linearly independent list can still remain independent, however, we cannot guarantee the result if we are still adding vectors; In mathematical terms, any sublist of the list is linearly independent, since  $\{\}$  is a sublist of any list, therefore its linearly independent.

**Lemma 2.15 (Linear Dependence Lemma)**

Suppose  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_h$  is linearly independent. Then there exists  $j$  between 1 and  $k$  such that

- $\mathbf{v}_j \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}\}$
- $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_h\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_h\}$

*Proof.* If  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  is a linearly dependent list, there are coefficients  $\alpha_1, \alpha_2, \dots, \alpha_k$  not all 0, such that

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0}$$

Take  $j$  such that  $\alpha_j$  is the largest index with  $\alpha_j \neq 0$ . Then  $\alpha_{j+1} = \alpha_{j+2} = \dots = \alpha_k = 0$  and

$$\mathbf{v}_j = \frac{-1}{\alpha_j} (\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_{j-1} \mathbf{v}_{j-1})$$

hence  $\mathbf{v}_j \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}\}$ . ■

**Lemma 2.16** (Very Important, a.k.a. Magic Lemma)

The length of the independent list  $\leq$  length of any spanning list.

*Proof.* Say  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$  is linearly independent say  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  is spanning. Then we want to establish that  $m \leq n$ .

- Step 1. Take the list  $\mathbf{u}_1, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ . It is linearly independent since  $\mathbf{u}_1 \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ . By the linear dependence lemma, there is a  $j$  such that  $\mathbf{v}_j$  can be removed (noted that  $\mathbf{u}_1$  cannot be subject to removal since  $\mathbf{u}_1$  comes from a linearly independent list). Consider the new list  $\{\mathbf{u}_1, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}, \dots, \mathbf{v}_m\}$ .
- Step 2. We can continue this process by bringing  $\mathbf{u}_2, \mathbf{u}_3, \dots, \mathbf{u}_n$ , we know that  $\mathbf{u}_i$  since they are linearly independent.

Note that this process preserves linear span of the whole list.

We know that this list contains all the  $\mathbf{u}_i$  (plus possibly some remaining  $\mathbf{v}_j$ ) and the length of the list is always  $n$ . So  $\boxed{m \leq n}$ . ■

## 2.2 Bases and Dimension

**Definition 2.17**

A basis a linearly independent spanning list.

**Theorem 2.18**

Any two basis in a finite dimensional space have the same number of vectors.

**Remark 2.19**

The span of  $\{\}$  is the zero vector.

**Theorem 2.20**

Suppose  $V$  is a finite dimensional vector space. Let  $W$  be a subspace of  $V$ , then  $W$  is finite dimensional.

*Proof.*  $V$  is finite dimensional means that  $V$  is spanned by some  $k$  vectors. Consider  $W$ . If  $W = \{\mathbf{0}\}$ , then  $w$  is spanned by the empty list  $\mathbf{0}$ . If  $W \neq \mathbf{0}$ , there exists  $\mathbf{w}_1 \in W$  such that  $W = \text{span}\{\mathbf{w}_1\}$ , done. Otherwise take  $\mathbf{w}_2 \in W \setminus \text{span}\{\mathbf{w}_1\}$ . Repeat this algorithm until it terminates. Now we want to show that this algorithm will terminate at  $\mathbf{w}_k$ , we know that  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_j$  is linearly independent by construction and the linearly dependence lemma. By remark 2.16, we know that the length of any such list will not exceed length  $k$ , therefore we know the algorithm will terminate in finite steps. This implies that  $W$  is finitely spanned, or  $W$  is finite dimension. ■

### 2.2.1 Dimension

#### Definition 2.21

Dimension of a vector space  $V$  is the cardinality of any basis in a finite dimensional space.

#### Proposition 2.22 (Criterion for a Basis)

$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  is a basis for  $V$  if and only if any  $v \in V$  can be uniquely written as a linear combination

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_n \mathbf{v}_n$$

*Proof.* We know that “can be written as linear combination” is logically equivalent a  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$  is a spanning list for  $V$ . “uniqueness” is logically equivalent as linear independence. Suppose

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2 + \dots + \beta_k \mathbf{v}_k$$

Not all  $\alpha_j = \beta_j$ . Then  $(\alpha_1 - \beta_1)\mathbf{v}_1 + (\alpha_2 - \beta_2)\mathbf{v}_2 + \dots + (\alpha_k - \beta_k)\mathbf{v}_k = \mathbf{0}$  is a nontrivial linear combination of  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  and vice versa. ■

#### Theorem 2.23

Any spanning set for a finite dimensional space can be shrink down to a basis.

*Proof.* Trivial by the linear dependence lemma. ■

#### Example 2.24

Consider  $\mathcal{P}_2(x)$  is spanned by  $\{x^2, (x-1)^2, (x-3)^2, (x-3)^2\}$ , we can see that this can be thinned down to  $\{x^2, (x-1)^2, (x-2)^2\}$ .

#### Corollary 2.25

Any linearly independent list in a finite dimensional space can be enlarged to a basis.

*Proof.* Add a spanning list at the back of our given list, then do removal for the linearly independent lemma. ■

#### Theorem 2.26

Suppose  $V$  is finite dimensional and  $W$  is a subspace, then there is a subspace  $U$  such that  $V = W \oplus U$ .

*Proof.* We already know by proceeding stuff  $W$  is finite dimensional and its dimension does not exceed that of  $V$ . Take any basis of  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$  of  $W$ . It's linearly independent so can be enlarged to a basis for  $V$ . Suppose the resulting basis is  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k, \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_l$ . Take  $U = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_l)$ . Then  $W + U = V$  and  $W \cap U = \{0\}$ . ■

### Remark 2.27

$\text{span}(\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_n)$  is a subspace of  $V$  by construction.

### Example 2.28

Consider  $\mathcal{P}(x)$ . We define  $W$  as

$$W := \{f \in \mathcal{P}_3(x) : f'(5) = 0\}$$

A basis for  $W$  can be taken as  $\{1, (x-5)^2, (x-5)^3\}$ .

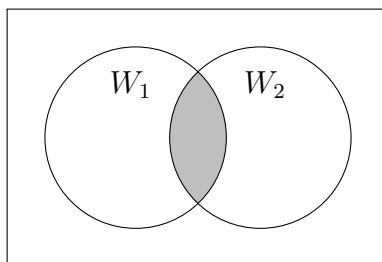
Now consider

$$\tilde{W} := \{f \in \mathcal{P}_3(x) : f''(5) = 0\}$$

. A basis for  $W$  can be taken as  $\{1, (x-5)^2, (x-5)^3\}$ .

## 2.2.2 Dimension of a Sum

### Principal of Inclusion for subspaces



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

Suppose  $W_1 \cap W_2$  forms a basis  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$ . We can extend the basis to

$$\mathbf{w}_1^{(1)}, \mathbf{w}_2^{(1)}, \dots, \mathbf{w}_l^{(1)}, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$$

is a basis for  $W_1$ . Similarly, we can extend the basis to

$$\mathbf{w}_1^{(2)}, \mathbf{w}_2^{(2)}, \dots, \mathbf{w}_m^{(2)}, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$$

is a basis for  $W_2$ .

We want to establish that

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2) = l + m + k$$

We want to prove that

$$\mathbf{w}_1^{(1)}, \mathbf{w}_2^{(1)}, \dots, \mathbf{w}_l^{(1)}, \mathbf{w}_1^{(2)}, \mathbf{w}_2^{(2)}, \dots, \mathbf{w}_m^{(2)}, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k$$

is a basis for  $W_1 + W_2$ . We can see that

$$\text{span}(\mathbf{w}_1^{(1)}, \mathbf{w}_2^{(1)}, \dots, \mathbf{w}_l^{(1)}, \mathbf{w}_1^{(2)}, \mathbf{w}_2^{(2)}, \dots, \mathbf{w}_m^{(2)}, \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k) \supseteq U_1, U_2$$

Hence  $\text{span}(\dots) \supseteq U_1 + U_2$ . Suppose the equation

$$\alpha_1 \mathbf{w}_1^{(1)} + \alpha_2 \mathbf{w}_2^{(1)} + \dots + \alpha_l \mathbf{w}_l^{(1)} + \beta_1 \mathbf{w}_1^{(2)} + \beta_2 \mathbf{w}_2^{(2)} + \dots + \beta_m \mathbf{w}_m^{(2)} + \gamma_1 \mathbf{w}_1 + \gamma_2 \mathbf{w}_2 + \dots + \gamma_k \mathbf{w}_k = \mathbf{0}$$

Manipulate the equation and we can see

$$\underbrace{(\alpha_1 \mathbf{w}_1^{(1)} + \alpha_2 \mathbf{w}_2^{(1)} + \dots + \alpha_l \mathbf{w}_l^{(1)} + \beta_1 \mathbf{w}_1^{(2)} + \beta_2 \mathbf{w}_2^{(2)} + \dots + \beta_m \mathbf{w}_m^{(2)})}_{\in W_1} = - \underbrace{(\gamma_1 \mathbf{w}_1 + \gamma_2 \mathbf{w}_2 + \dots + \gamma_k \mathbf{w}_k)}_{\in W_2 \setminus W_1}$$

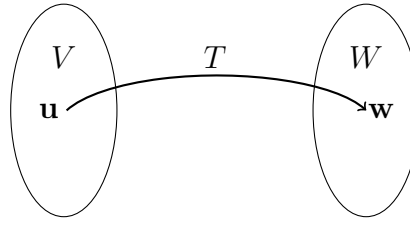
Since they belongs to different sets, clearly they cannot span each other. Therefore

$$\alpha_1 = \alpha_2 = \dots = \alpha_l = \beta_1 = \beta_2 = \dots = \beta_m = \gamma_1 = \gamma_2 = \dots = \gamma_k = 0$$

Hence the list of vectors is also linearly independent.



### 3 Linear Maps



#### 3.1 Linear Maps as Vector Space

Suppose  $V$  and  $W$  are two linear spaces over  $\mathbb{F}$ .  $T$  is a function with domain  $V$  and codomain  $W$ .  $T$  is called linear iff

1.  $T(\mathbf{u}_1 + \mathbf{u}_2) = T(\mathbf{u}_1) + T(\mathbf{u}_2)$
2.  $T(\lambda \mathbf{u}) = \lambda \cdot T(\mathbf{u})$

$\forall \mathbf{v}_1, \mathbf{v}_2 \in V$  and  $\forall \mathbf{v}_1 \mathbf{v}_2 \in V, \forall \lambda \in \mathbb{F}$ .

##### Example 3.1

Let  $V = \mathbb{R}^3, W = \mathbb{R}^4$ . Define  $T$  as  $(x_1, x_2, x_3) \mapsto (x_1, 0, 0, 0)$

##### Example 3.2

$T : \mathcal{P}(x) \rightarrow \mathcal{P}(x)$ , where  $f(x) \mapsto \int_{10}^x f(x) dx$  is a linear map.

##### Definition 3.3

$\mathcal{L}\{V, W\}$  denotes the set of all linear maps from  $V$  to  $W$ . Note that  $\mathcal{L}\{V, W\}$  with  $+$  and  $\cdot$  becomes a vector space over  $\mathbb{F}$ . This requires the additions of functions and multiplications of linear maps by scalars (from  $\mathbb{F}$ ). Given  $T_1, T_2 \in \mathcal{L}(V, W)$  we define addition as  $(T_1 + T_2)(\mathbf{u}) := T_1(\mathbf{u}) + T_2(\mathbf{u})$ , multiplication as  $(\lambda T)(\mathbf{u}) := \lambda \cdot T(\mathbf{u})$ .

##### Theorem 3.4

In finite vector space  $V, W$ , let  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$  be a basis for  $V$ , let  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$  be any vectors in  $W$ . Then there exist a unique linear map  $T \in \mathcal{L}\{V, W\}$  such that  $T(\mathbf{u}_j) = \mathbf{w}_j \forall j$ .

*Proof.* Any vector in  $V$  has a unique representation  $\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_n \mathbf{u}_n = \mathbf{u}$ .

Define  $T(\mathbf{u}) := \underbrace{\alpha_1 T(\mathbf{u})_1 + \alpha_2 T(\mathbf{u})_2 + \dots + \alpha_n T(\mathbf{u})_n}_{\in W}$  This makes  $T$  a linear map from  $V$  to  $W$ .

Indeed if  $\lambda \in \mathbb{F}$ , then  $T(\lambda \mathbf{u}) = T(\sum_{j=1}^n \lambda \alpha_j \mathbf{u}_j) = \lambda \sum_{j=1}^n \alpha_j \mathbf{w}_j$ . Suppose  $\tilde{T}(\mathbf{u}_j) = \mathbf{w}_j$  for all  $j$ , then  $T = \tilde{T}$  as a map function by linearity and basis. ■

## 3.2 Null Space and Range

### Theorem 3.5

Let  $\text{Null}(T) := \{\mathbf{u} \in V : T(\mathbf{u}) = 0\}$ .  $\text{Null}(T)$  is a subspace of  $V$ .

### Theorem 3.6

Let  $\text{Range}(T) := \{\mathbf{w} \in W : T(\mathbf{u}) = \mathbf{w}\}$ .  $\text{Range}(T)$  is a subspace of  $W$ .

*Proof.* The proof is trivial and is left as an exercise for the reader. ■

### Example 3.7

Let  $T : f \rightarrow f', V := \mathcal{P}(x), W := \mathcal{P}(x)$ .  $\text{Null}(T) = \mathcal{P}_0(x)$ ,  $\text{Range}(T) = \mathcal{P}_2(x)$ .  
 Let  $T : f \rightarrow f'', V := \mathcal{P}(x), W := \mathcal{P}(x)$ .  $\text{Null}(T) = \mathcal{P}_1(x)$ ,  $\text{Range}(T) = \mathcal{P}_1(x)$ .

### Example 3.8

Find a basis of  $\mathcal{L}(V, W)$  given bases  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$  and  $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$  of  $V$  and  $W$ . The basis consists of  $m \times n$  vectors as follows:

$$T_{11} = T(\mathbf{u}_1) = \mathbf{w}_1, T(\mathbf{u}_2) = \mathbf{0}, T(\mathbf{u}_3) = \mathbf{0}, \dots, T(\mathbf{u}_m) = \mathbf{0}$$

$$T_{12} = T(\mathbf{u}_1) = \mathbf{w}_2, T(\mathbf{u}_2) = \mathbf{0}, T(\mathbf{u}_3) = \mathbf{0}, \dots, T(\mathbf{u}_m) = \mathbf{0}$$

...

$$T_{mn} = T(\mathbf{u}_1) = \mathbf{0}, T(\mathbf{u}_2) = \mathbf{0}, T(\mathbf{u}_3) = \mathbf{0}, \dots, T(\mathbf{u}_m) = \mathbf{w}_n$$

### Example 3.9

Let  $\mathcal{U} = \{f : \mathbb{R} \rightarrow \mathbb{R} : f(x) = f(1-x) \ \forall x\}$ .

1. Show that  $\mathcal{U}$  is a subspace of  $f : \mathbb{R} \rightarrow \mathbb{R}$ .
2. Find a complement.

$$\mathcal{W} = \{g : \mathbb{R} \rightarrow \mathbb{R} : g(x) = -g(1-x) \ \forall x\}$$

*Proof.* We can see that the zero function  $f(x) = 0$  satisfies the requirement since  $0 = 0$  for all values of  $x$ .

Suppose  $f(x), g(x) \in \mathcal{U}$ , then we compute

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \\ &= f(1-x) + g(1-x) \\ &= (f+g)(1-x) \end{aligned}$$

Therefore we can see that  $\mathcal{U}$  is closed under addition.

Suppose  $f(x) \in \mathcal{U}$ ,  $\lambda \in \mathbb{R}$ , then we compute

$$\begin{aligned} (\lambda \cdot f)(x) &= \lambda \cdot f(x) \\ &= \lambda \cdot f(1-x) \\ &= (\lambda \cdot f)(1-x) \end{aligned}$$

Therefore we can see that  $\mathcal{U}$  is closed addition. Hence  $\mathcal{U}$  is a vector space. ■

*Proof.* The proof for subspace is similar to part (i) and is omitted here.

We now want to show that  $\mathcal{U} + \mathcal{W} = \mathbb{R}^{\mathbb{R}}$ . We can see that for  $f(x) \in \mathbb{R}^{\mathbb{R}}$ , we can rewrite  $f(x)$  as

$$f(x) = \frac{f(x) + f(1-x)}{2} + \frac{f(x) - f(1-x)}{2}$$

Clearly  $\frac{f(x) + f(1-x)}{2} \in \mathcal{U}$  and  $\frac{f(x) - f(1-x)}{2} \in \mathcal{W}$ . For uniqueness, suppose that a nonzero  $h(x) \in \mathcal{U} \cap \mathcal{W}$ , therefore  $h(x) = h(1-x) = -h(1-x)$ , and the only solution is  $f(x) = 0$ , a contradiction, therefore  $\mathcal{U} \cap \mathcal{W} = \{0\}$ . Hence  $\boxed{\mathbb{R}^{\mathbb{R}} = \mathcal{U} \oplus \mathcal{W}}$  ■

### Theorem 3.10 (Rank-Nullity Theorem also known as the Fundamental Theorem of Linear Maps)

Let  $V, W$  be finite dimensional vector spaces, let  $T \in \mathcal{L}(V, W)$ . Then

$$\dim V = \dim \text{Null } T + \dim \text{Range } T$$

*Proof.* Let  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$  to be the basis for the basis for Null  $T$ . By the linear independent list extension theory, this list can be extended to a basis of  $V$ . Say  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_l$  is asujc an extension to a basis of  $V$ . We can see that  $\dim V = k + l$ . We want to show that  $\text{Range } T = l$ . Consider  $T\mathbf{v}_1, T\mathbf{v}_2, \dots, T\mathbf{v}_l$ . We want to show that  $T\mathbf{v}_1, T\mathbf{v}_2, \dots, T\mathbf{v}_l$  is basis for Range  $T$ . Notice that  $\mathbf{v} \in V$  can be written as a linear combination of  $\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_k\mathbf{u}_k + \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2 + \dots + \beta_l\mathbf{v}_l$ . Then we compute

$$\begin{aligned} T\mathbf{v} &= \alpha_1 T\mathbf{u}_1 + \alpha_2 T\mathbf{u}_2 + \dots + \alpha_k T\mathbf{u}_k + \beta_1 T\mathbf{v}_1 + \beta_2 T\mathbf{v}_2 + \dots + \beta_l T\mathbf{v}_l \\ &= \beta_1 T\mathbf{v}_1 + \beta_2 T\mathbf{v}_2 + \dots + \beta_l T\mathbf{v}_l \end{aligned}$$

hence  $T\mathbf{v} \in \text{span}(T\mathbf{v}_1, T\mathbf{v}_2, \dots, T\mathbf{v}_l)$ .

Suppose  $\beta_1 T\mathbf{v}_1 + \beta_2 T\mathbf{v}_2 + \dots + \beta_l T\mathbf{v}_l = 0$ . Then  $\beta_1 T\mathbf{v}_1 + \beta_2 T\mathbf{v}_2 + \dots + \beta_l T\mathbf{v}_l \in \text{Null } T$ . So

$$\beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2 + \dots + \beta_l \mathbf{v}_l = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_k \mathbf{u}_k$$

for some  $\alpha_1, \alpha_2, \dots, \alpha_k$  since  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$  form a basis for Null  $T$ .

But  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_l, \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$  form a basis for  $V$ , all of the coefficient has to be 0. Therefore  $T\mathbf{v}_1, T\mathbf{v}_2, \dots, T\mathbf{v}_l$  is indeed a basis for Range  $T$ . ■

**Example 3.11** (Direct consequences of the Theorem)

Suppose  $\dim W < \dim V$  (both finite), and  $T \in \mathcal{L}(V, W)$ . Then  $T$  cannot be injective.

*Proof.*  $T$  is injective implies that  $\text{Null } T = \{\mathbf{0}\}$ . So  $\dim V = 0 + \dim \text{Range } T \leq \dim W < \dim V$ , a contradiction. ■

**Example 3.12** (Direct consequences of the Theorem)

Suppose  $\dim W > \dim V$  (both finite), and  $T \in \mathcal{L}(V, W)$ . Then  $T$  cannot be surjective.

*Proof.*  $T$  is surjective implies that  $\text{Range } T = W$ . So  $\dim V = \dim \text{Null } T + \dim \text{Range } T \geq \dim W > \dim V$ , a contradiction. ■

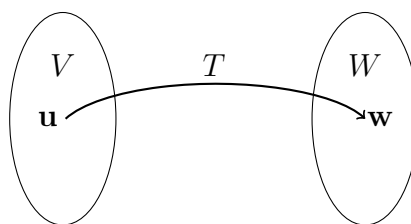
**Example 3.13** (Fun Question)

Suppose that  $p \in \mathcal{P}(\mathbb{R})$ , prove that  $\exists q \in \mathcal{P}(\mathbb{R})$  such that  $5q'' + 3q' = p$ .

[This exercise can be done without linear algebra, but it's more fun to do it using linear algebra.]

*Proof.* Let  $d = \deg p$ . Define linear transformation  $T : \mathcal{P}_{d+1}(\mathbb{R}) \rightarrow \mathcal{P}_d(\mathbb{R})$  as  $T : q \rightarrow 5q'' + 3q'$ . We can see that  $\dim \text{Null } T = 1$ , by the rank nullity theorem, know that  $T$  must be surjective as  $\dim \mathcal{P}_{d+1}(\mathbb{R}) = \dim \text{Null } T + \dim \text{Range } T = 1 + \dim \text{Range } T \implies \dim \text{Range } T = \dim \mathcal{P}_d(\mathbb{R})$ . ■

### 3.3 Matrix Notation



Recall this diagram, we want to understand  $T$  “correctly”. Pick a basis  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  for  $V$  and  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$  for  $W$ . We can see that  $\dim V = n$  and  $\dim W = m$ . We can define  $T$  as

$$Tv_j = A_{1,j}\mathbf{w}_1 + A_{2,j}\mathbf{w}_2 + \dots + A_{m,j}\mathbf{w}_m$$

Notice that  $A$  has the following form

$$\begin{aligned} \mathbf{w}_1 &\rightarrow \left[ A_{1,1} & A_{1,2} & \cdots & A_{1,n} \right] \\ \mathbf{w}_2 &\rightarrow \left[ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \right] \\ \vdots & \\ \mathbf{w}_m &\rightarrow \left[ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \right] \end{aligned}$$

This is called the matrix representation of  $T$ .

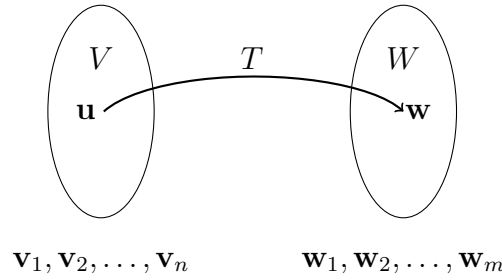
**Example 3.14**

Let  $D : V \rightarrow W$  be defined as  $D := p \rightarrow p'$ . Let  $V := \text{span}(1, \cos x, \sin x, \cos 2x, \sin 2x) = W$ . We can see that

$$\begin{array}{lcl} 1 \mapsto 0 \\ \cos x \mapsto -\sin x \\ \sin x \mapsto \cos x \\ \cos 2x \mapsto -2\sin 2x \\ \sin 2x \mapsto \cos 2x \end{array} \quad A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 2 & 0 \end{bmatrix}$$

**3.4 Matrix Representation**

Recall that if  $T$  is a linear transformation



$$T\mathbf{v}_u = \sum_{k=1}^m A_{i,k} \mathbf{w}_k$$

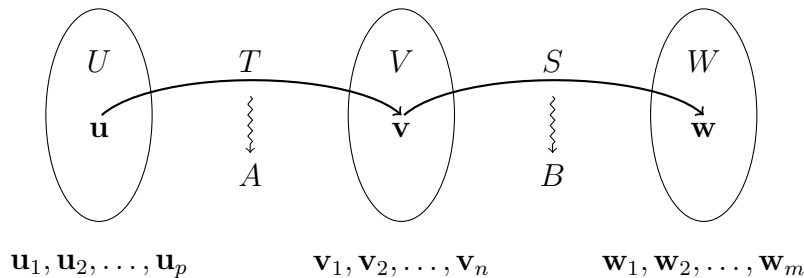
Note that Matrix  $A = [A_{i,k}]$  has  $m$  rows  $n$  columns.

$$\begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} \end{bmatrix}$$

Suppose  $v = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ ,  $T\mathbf{v} = c_1T\mathbf{v}_1 + c_2T\mathbf{v}_2 + \cdots + c_nT\mathbf{v}_n$ .

$$T\mathbf{v} = c_1 \sum_{i=1}^m A_{i,1} \mathbf{w}_1 + c_2 \sum_{i=1}^m A_{i,2} \mathbf{w}_2 + \cdots + c_n \sum_{i=1}^m A_{i,n} \mathbf{w}_n = \sum_{i=1}^m \left( \sum_{j=1}^n A_{i,j} c_j \right) \mathbf{w}_j$$

Notice that the operation is the equivalent as the matrix-vector multiplication.



$$ST\mathbf{u}_k = S(T\mathbf{u}_k) = S\left(\sum_{j=1}^n A_{j,k} \mathbf{v}_j\right) = \sum_{j=1}^n A_{j,k} (S\mathbf{v}_j) = \sum_{j=1}^n A_{j,k} \sum_{i=1}^m B_{i,j} \mathbf{w}_i = \sum_{i=1}^m \left( \sum_{j=1}^n B_{i,j} A_{j,k} \right) \mathbf{w}_i$$

Use name  $\mathcal{M}(S) := B$ ,  $\mathcal{M}(T) := A$ ,  $\mathcal{M}(ST) = BA = \mathcal{M}(S) \cdot \mathcal{M}(T)$ . So matrix representation multiply as matrices to produce a composition map or product.

**Remark 3.15** (Book Keeping)

$A_{*,j}$  denotes the  $j$ th column of  $A$ .

$A_{i,*}$  denotes the  $i$ th row of  $A$ .

Notice that  $\mathcal{M}$  is a linear map,  $\mathcal{L}(V, W) \xrightarrow{\mathcal{M}} \mathbb{F}^{m,n}$ .

**Proposition 3.16**

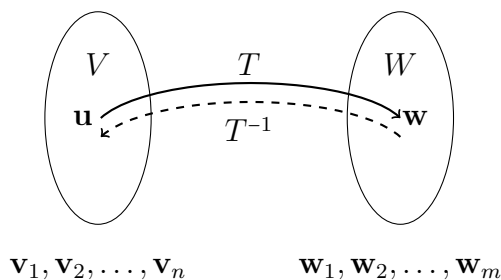
$\mathcal{M}$  is a linear map.

**Proposition 3.17**

$\mathbb{F}^{m,n}$  has a basis.

*Proof.* Consider  $E_{i,j}$ , the matrix consists of all zeros with the exception of 1 in position  $(i, j)$ . This can be done for all  $i = 1, 2, \dots, m$ ,  $j = 1, 2, \dots, n$ . Also notice that  $\dim \mathbb{F}^{m,n} = m \cdot n$  ■

### 3.5 Invertibility and Isomorphism



**Definition 3.18**

$T \in \mathcal{L}(V, W)$  is invertible provided that there exists a mapping  $T^{-1}$  from  $W$  to  $V$  (not necessarily linear) such that

$$T^{-1} \circ T = \mathbb{I}_V$$

$$T \circ T^{-1} = \mathbb{I}_W$$

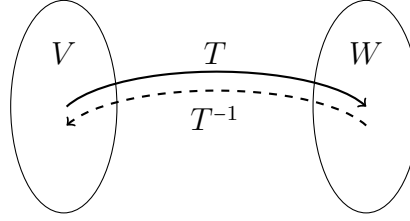
Where  $\mathbb{I}_V, \mathbb{I}_W$  is the identity map on  $V$  and  $W$ .

**Theorem 3.19**

$T$  is invertible if and only if  $T$  is both injective and surjective.

*Proof.* Suppose  $T$  is invertible, then  $T(T^{-1}\mathbf{w}) = \mathbf{w} \ \forall \mathbf{w} \in W$ , so  $\text{Range } T = W$ . Also we know that  $T^{-1}(T\mathbf{v}) = \mathbf{v}$ . Suppose  $T\mathbf{v}_1 = T\mathbf{v}_2$ , apply the left inverse and we have  $T^{-1}(T\mathbf{v}_1) = T^{-1}(T\mathbf{v}_2) \implies \mathbf{v}_1 = \mathbf{v}_2$ . Hence  $T$  is injective. Therefore  $T$  is bijective.

Now suppose  $T$  is bijective. We want to construct  $T^{-1}$



We need to take  $\mathbf{w} \in W$ , there is a  $\mathbf{v} \in V$  such that  $T\mathbf{v} = \mathbf{w}$  and such  $\mathbf{v}$  is unique since  $T$  is injective. We declare  $T^{-1}\mathbf{w}$  to be  $\mathbf{v}$ . So  $T^{-1} \circ T = \mathbb{I}_V$ . We compute

$$(T \circ T^{-1})\mathbf{w} = T(T^{-1}\mathbf{w}) = T\mathbf{v} = \mathbf{w} \ \forall \mathbf{w} \in W$$

So  $T \circ T^{-1} = \mathbb{I}_W$  ■

### Definition 3.20

If  $V, W$  are vector spaces, such that there exists an invertible linear map  $T \in \mathcal{L}(V, W)$  then  $V, W$  are isomorphic.

### Remark 3.21

Before we proceed, we want to check that  $T^{-1}$  is a linear map when  $T \in \mathcal{L}(V, W)$  and  $T^{-1}$  exists.

*Proof.* Take  $\mathbf{w}_1, \mathbf{w}_2 \in W, \lambda \in \mathbb{F}$ . We compute  $T^{-1}(\lambda\mathbf{w}_1 + \mathbf{w}_2)$ . We know that  $\mathbf{w}_1 = T\mathbf{v}_1$  and  $\mathbf{w}_2 = T\mathbf{v}_2$ . Then we know that  $T(\lambda\mathbf{v}_1 + \mathbf{v}_2) = \lambda T\mathbf{v}_1 + T\mathbf{v}_2 = \lambda\mathbf{w}_1 + \mathbf{w}_2$ . Substitute this into  $T^{-1}$  and we get

$$T^{-1}(\lambda\mathbf{w}_1 + \mathbf{w}_2) = T^{-1} \circ T(\lambda\mathbf{v}_1 + \mathbf{v}_2) = \mathbb{I}(\lambda\mathbf{v}_1 + \mathbf{v}_2) = \lambda\mathbf{v}_1 + \mathbf{v}_2 = \lambda T^{-1}\mathbf{w}_1 + T^{-1}\mathbf{w}_2$$

Hence  $T^{-1}$  is linear. ■

### Corollary 3.22

$\mathcal{M}$  is actually a bijection between  $\mathcal{L}(V, W)$  and  $\mathbb{F}^{m,n}$ , therefore  $\mathcal{L}(V, W)$  is isomorphic to  $\mathbb{F}^{m,n}$ .

### Theorem 3.23

Suppose  $T \in \mathcal{L}(V, W)$  is linear and invertible, and let  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  be a basis for  $V$ . Then  $T\mathbf{v}_1, T\mathbf{v}_2, \dots, T\mathbf{v}_m$  is a basis for  $W$ .

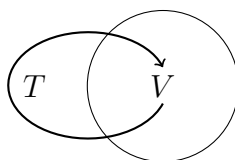
*Proof.* Suppose  $\alpha_1 T\mathbf{v}_1 + \alpha_2 T\mathbf{v}_2 + \cdots + \alpha_n T\mathbf{v}_n = 0$ . Then  $T(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n) = 0$ . Since  $T$  is injective, this implies  $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n = 0$ . Therefore  $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$  since  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  is a basis. Take  $\mathbf{w} \in W$ , then there exists a unique  $\mathbf{v} \in V$  such that  $T\mathbf{v} = \mathbf{w}$ , and  $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n$  for some  $\alpha_1, \alpha_2, \dots, \alpha_n$ , so  $T\mathbf{v} = \mathbf{w} = \alpha_1 T\mathbf{v}_1 + \alpha_2 T\mathbf{v}_2 + \cdots + \alpha_n T\mathbf{v}_n$ , hence span. ■

### Corollary 3.24

dim is invariant under isomorphism.

### 3.5.1 Linear Operators

We are dealing with a specific case where  $\mathcal{L}(V, W)$  is replaced by  $\mathcal{L}(V, V)$ .



Recall that  $\dim V = \dim \text{Null } T + \dim \text{Range } T$ . This gives a better test for invertability if  $W = V$ .

### Theorem 3.25

Let  $T \in \mathcal{L}(V, V)$ . If  $V$  is finite dimensional vector space, then the following are equivalent:

- (a)  $T$  is injective.
- (b)  $T$  is surjective.
- (c)  $T$  is invertible.

*Proof.*

(a)  $\implies$  (c). Trivial by definition.

(b)  $\implies$  (c). Suppose  $T$  is injective  $\xrightarrow{T \text{ being linear}} \text{Null } T = \{0\} \iff \dim \text{Null } T = 0$ . Therefore

$$\dim V = \dim \text{Null } T + \dim \text{Range } T = \dim \text{Range } T = 0$$

So  $\dim V = \dim \text{Range } T$ , hence  $\text{Range } T = V$ , so  $T$  is surjective.

(c)  $\implies$  (a).

$$\dim V = \dim \text{Null } T + \dim \text{Range } T = \dim \text{Null } T + \dim V \implies \dim \text{Null } T = 0 \implies \text{Null } T = \{0\}$$

So  $T$  is injective.  $T$  is already known to be surjective, so  $T$  is bijective, or  $T$  is invertable. ■



**Example 3.26**

The theorem does not hold for infinite dimension vector spaces, for example:

The differentiation map  $T : f(x) \mapsto f'(x)$  is surjective but not invertible over  $\mathcal{P}(\mathbb{R})$ . From calculus we know that for every  $f(x) \in \mathcal{P}(\mathbb{R})$ , there exists  $g(x) \in \mathcal{P}(\mathbb{R})$  such that  $g'(x) = f(x)$ , however, we can see that  $\text{Null } T \neq \{0\}$  as  $1 \in \text{Null } T$ . Hence  $T$  is not injective.

The integration map  $T : f(x) \mapsto \int_0^x f(t)dt$  is injective but not invertible over  $\mathcal{P}(\mathbb{R})$ . We can see that  $\text{Null } T = \{0\}$ , hence  $T$  is injective, however,  $1 \notin \text{Range } T$ .

**3.6 Duality****Definition 3.27**

Given a vector space  $V$ , we define its dual space as  $V' = \mathcal{L}(V, \mathbb{F})$ .

**Remark 3.28**

Objects in  $\mathcal{L}(V, \mathbb{F})$  are also called linear functional.

**Example 3.29**

Linear Functional on  $\mathbb{R}^3$ :  $(x_1, x_2, x_3) \mapsto \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 \quad \forall \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ .

**Definition 3.30**

Suppose  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  is a basis of  $V$ . The list of linear functionals  $\varphi_1, \varphi_2, \dots, \varphi_n \in V'$  such that

$$\varphi_i(\mathbf{v}_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

We claim that  $\varphi_1, \varphi_2, \dots, \varphi_n$  is the dual basis of  $V'$ .

**Lemma 3.31**

A dual basis is a basis of  $V'$ .

*Proof.* Suppose there exists  $\alpha_1, \alpha_2, \dots, \alpha_n$  such that

$$\alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \dots + \alpha_n \varphi_n = 0$$

We compute for  $v_1$

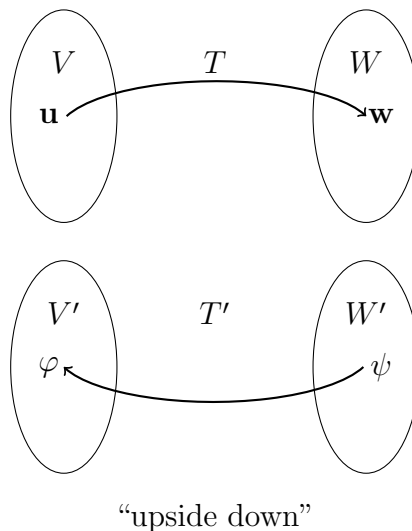
$$(\alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \dots + \alpha_n \varphi_n)(v_1) = \alpha_1 \cdot 1 + 0 \implies \alpha_1 = 0$$

Similarly, if we plug in an arbitrary  $v_j$

$$(\alpha_1\varphi_1 + \alpha_2\varphi_2 + \cdots + \alpha_n\varphi_n)\mathbf{v}_j \implies v_j = 0$$

This shows that  $\varphi_1, \varphi_2, \dots, \varphi_n$  is linear independent, and the count is  $n$ . So  $\varphi_1, \varphi_2, \dots, \varphi_n$  form a basis for  $V'$ . ■

### 3.6.1 Dual Maps



#### Definition 3.32

Given  $T \in \mathcal{L}(V, W)$ , we define  $T' : \varphi \mapsto \varphi \circ T$ .  $\phi \in W'$  i.e.  $\varphi \in \mathcal{L}(W, \mathbb{F})$ . Notice that  $\varphi \circ T \in \mathcal{L}(V, \mathbb{F})$ .

#### Example 3.33

Define  $V, W, \varphi, T$  as  $\varphi : f \mapsto \int_0^1 f(t)dt$ ,  $V = \mathcal{P}_3(\mathbb{R})$ ,  $W = \mathcal{P}_2(\mathbb{R})$ ,  $T : f \mapsto f'$ . What is  $\varphi \circ T$ ?

$$(\varphi \circ T)(f) = \varphi(T(f)) = \varphi(f') = \int_0^1 f'(t)dt = f(1) - f(0)$$

**Remark 3.34** (Algebraic Property of dual maps)

$$1. (S + T)' = S' + T'$$

*Proof.* For any  $S, T \in \mathcal{L}(V, W), S', T' \in \mathcal{L}(W', V')$ , we compute

$$(S + T)'(\varphi) = \varphi \circ (S + T) = \varphi \circ S + \varphi \circ T = S'(\varphi) + T'(\varphi)$$

■

$$2. (\lambda S)' = \lambda S'$$

*Proof.*

$$(\lambda S)'(\varphi) = \varphi \circ (\lambda S) = \lambda(\varphi \circ S) = \lambda \cdot S'(\varphi)$$

■

$$3. (ST)' = T'S'$$

*Proof.* Sanity check:  $S \in \mathcal{L}(V, W), T \in \mathcal{L}(U, V)$

$$(ST)' : \varphi \mapsto \varphi \circ (ST) = (\varphi \circ S) \circ T = T'(\varphi \circ S) = T'(S'(\varphi)) = T' \circ S'$$

■

**Definition 3.35** (Annihilators)

Let a set  $S$  to be a subset of a vector space  $V$ . We can define  $S^0$  as

$$S^0 := \{\varphi \in V' : \varphi(v) = 0 \ \forall v \in S\}$$

**Example 3.36**

Consider  $\mathbb{R}^3$ , let  $S := \{(1, 0, 0), (1, 1, 0)\}$ . We know that any  $\varphi \in \mathbb{R}^3$  will have the form of  $(x_1, x_2, x_3) \mapsto a_1x_1 + a_2x_2 + a_3x_3$  for some constant  $a_1, a_2, a_3$ . Plug in  $(1, 0, 0)$  and  $(1, 1, 0)$  and we get

$$a_1 \cdot 1 + a_2 \cdot 0 + a_3 \cdot 0 = 0 \implies a_1 = 0$$

$$a_1 \cdot 1 + a_2 \cdot 1 + a_3 \cdot 0 = 0 \implies a_2 = 0$$

We can see that  $S^0 = \{\varphi(x_1, x_2, x_3) = a_3x_3 \text{ for some } a_3 \in \mathbb{R}\}$  and forms a subspace for  $\mathbb{R}^3$ .

**Lemma 3.37**

Regardless of the nature of  $S$ ,  $S^0$  is always a subspace.

*Proof.* 1. The zero functional is clearly in  $S^0$ .

2. Suppose  $\varphi \in S^0$ , take  $\lambda \in \mathbb{F}$ , then  $(\lambda\varphi)(\mathbf{v}) = \lambda \cdot \varphi(\mathbf{v}) = 0$  for all  $\mathbf{v} \in S$ . So  $\lambda\varphi \in S$ .

3. Suppose  $\varphi, \psi \in S^0$ , then  $(\varphi + \psi)(\mathbf{v}) = \varphi(\mathbf{v}) + \psi(\mathbf{v}) = 0$  for all  $\mathbf{v} \in S$ . Therefore  $\varphi + \psi \in S^0$ . ■

### Theorem 3.38

Suppose  $S = U$ , where  $U$  is a subspace of  $V$ , then

$$\dim U + \dim U^0 = \dim V$$

*Proof.* Consider the inclusion map:

$$i : U \rightarrow V : \mathbf{u} \rightarrow \mathbf{u} \quad \forall \mathbf{u} \in U$$

Take a look at the dual of  $U$ :  $i' \in \mathcal{L}(V', U')$ . Apply Rank-Nullity to the dual map and we can see that  $\dim V' = \dim \text{Null } i' + \dim \text{Range } i'$ . We also know that

$$\text{Null } i' = \{\varphi \in V' : \varphi \circ i = 0\}$$

Notice that  $\varphi \circ i = 0$  as a functional implies

$$(\varphi \circ i)(\mathbf{u}) = 0 \implies \varphi(i(\mathbf{u})) = 0 \implies \varphi(\mathbf{u}) = 0 \quad \forall \mathbf{u} \in U$$

Therefore we can see that  $\text{Range } i' = \{\varphi \circ i : \varphi \in V'\} = U'$  since any linear functional on  $U$  extends to  $V$ . I have a clever proof for this but it does not fit in the margin of the page and is left as an exercise for the reader. ■

### Theorem 3.39

Let  $V, W$  be finite dimensional vector space and let  $T \in \mathcal{L}(V, W)$ . Then

$$(a) \quad \text{Null } T' = (\text{Range } T)^0$$

$$(b) \quad \dim \text{Null } T' = \dim \text{Null } T + \dim W - \dim V$$

*Proof.*

$$(a) \quad \varphi \in \text{Null } T' \iff \varphi \circ T = 0 \iff (\varphi \circ T)(\mathbf{v}) = 0 \quad \forall \mathbf{v} \in V \iff \varphi(T\mathbf{v}) = 0 \quad \forall \mathbf{v} \in V \\ \iff \varphi \in (\text{Range } T)^0$$

$$(b) \quad \dim \text{Null } T' = \dim (\text{Range } T)^0 = \dim W - \dim \text{Range } T = \dim W - (\dim V - \dim \text{Null } T) = \\ \dim W - \dim V + \dim \text{Null } T$$

■

**Corollary 3.40**

$T'$  is injective if and only if  $T$  is surjective.

**Theorem 3.41**

Suppose  $V$  and  $W$  are finite dimensional and  $T \in \mathcal{L}(V, W)$ , then

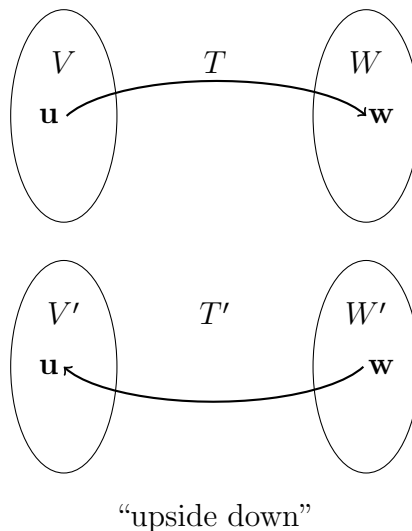
- (a)  $\dim \text{Range } T' = \dim \text{Range } T$
- (b)  $\text{Range } T' = (\text{Null } T)^0$

*Proof.*

- (a)  $\dim \text{Range } T' = \dim W' - \dim \text{Null } T' = \dim W - (\dim W - \dim V + \dim \text{Null } T) = \dim V - \dim \text{Null } T = \dim \text{Range } T$
- (b)  $\psi \in \text{Range } T' \iff \exists \varphi : \varphi \circ T = \psi \iff \varphi \circ T(\mathbf{v}) = \psi(\mathbf{v}) \ \forall \mathbf{v} \in V \iff \varphi(T\mathbf{v}) = \psi(\mathbf{v}) \ \forall \mathbf{v} \in V.$   
 So  $T\mathbf{v} = 0 \implies \psi(\mathbf{v}) = 0$ . This shows  $\text{Range } T' \subseteq (\text{Null } T)^0$ . But  $\dim \text{Range } T' = \dim \text{Range } T = \dim V - \dim \text{Null } T = \dim(\text{Null } T)^0$ . Hence  $\text{Range } T' = (\text{Null } T)^0$ . ■

**3.6.2 Matrix Representation of the dual map**

Recall that



We also recall that  $T' : \varphi \mapsto \varphi \circ T$ .

**Question 3.42**

How do we get  $\mathcal{M}(T')$  given  $\mathcal{M}(T)$ ?

**Answer 3.43.** For  $\mathcal{M}(T)$ , we need 2 bases  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  for  $V$  and  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$  for  $W$ . Take  $\varphi_1, \varphi_2, \dots, \varphi_n$ , a dual basis to  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ , it is a basis for  $V'$ .

Take  $\psi_1, \psi_2, \dots, \psi_m$ , a dual basis to  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ , it is a basis for  $W'$ .

Given  $\mathcal{M}(T)$ , we want to construct / understand  $\mathcal{M}(T')$  with regard to the basis  $\psi_1, \psi_2, \dots, \psi_m$  of  $W'$  and  $\varphi_1, \varphi_2, \dots, \varphi_n$  of  $V'$ .

We know  $\mathcal{M}(T)$  has  $m$  rows  $n$  columns, and  $\mathcal{M}(T')$  has  $n$  rows and  $m$  columns.

Suppose  $\mathcal{M}(T) = A, \mathcal{M}(T') = C$ , we then know

$$T\mathbf{v}_j = \sum_{i=1}^m A_{i,j} \mathbf{w}_i \quad \forall j = 1, 2, \dots, n, \quad T'\psi_l = \sum_{l=1}^n C_{l,k} \varphi_l \quad \forall l = 1, 2, \dots, m$$

$$T'\psi_k = \psi_k \circ T \implies (\psi_k \circ T)(\mathbf{v}_j) = \psi_k(T\mathbf{v}_j) = \psi_k\left(\sum_{i=1}^m A_{i,j} \mathbf{w}_i\right) = \sum_{i=1}^m A_{i,j} \psi_k(\mathbf{w}_i) = \sum_{i=1}^m A_{i,j} \delta_{ki} = A_{k,j}$$

$$(T'\psi_k)(\mathbf{v}_j) = \left(\sum_{l=1}^n C_{l,k} \varphi_l\right)(\mathbf{v}_j) = \sum_{l=1}^n C_{l,k} \varphi_l(\mathbf{v}_j) = \sum_{l=1}^n C_{l,k} \delta_{lj} = C_{j,k}$$

Notice that  $A_{k,j} = C_{j,k} \quad \forall j, k$ .

**Conjecture 3.44.** So we obtained that

$$\mathcal{M}(T') = \mathcal{M}(T)^T$$

provided that the basis of  $V'$  and  $W'$  are chosen to be the dual to the bases of  $V$  and  $W$ , respectively.

**Example 3.45**

Let  $T: p \mapsto p'$  for  $V = \mathcal{P}_3(\mathbb{R})$  with basis  $1, x, x^2, x^3$ , and  $W = \mathcal{P}_2(\mathbb{R})$  with  $1, x, x^2$ .

We can see that the dual basis for  $1, x, x^2, x^3$  is

$$\varphi_0: p \mapsto p(0), \varphi_1: p \mapsto p'(0), \varphi_2: p \mapsto \frac{p''(0)}{2}, \varphi_3: p \mapsto \frac{p'''(0)}{3!}$$

Dual basis for  $1, x, x^2$  is

$$\psi_0: p \mapsto p(0), \psi_1: p \mapsto p'(0), \psi_2: p \mapsto \frac{p''(0)}{2},$$

Notice that

$$\mathcal{M}(T) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \text{ and } \mathcal{M}(T') = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

## 4 Polynomials

Recall we call consider polynomials over  $\mathbb{C}$  or  $\mathbb{R}$ .

### Theorem 4.1

For any  $z_1, z_2 \in \mathbb{C}$ , we define  $|z| = \sqrt{a^2 + b^2}$ , we know that

1.  $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$
2.  $|z_1 \cdot z_2| \leq |z_1| + |z_2|$

*Proof.* Left as an exercise. ■

### 4.1 Axler's Recap on Polynomial

#### Theorem 4.2

Suppose  $p(x) \in \mathcal{P}(\mathbb{F})$ , is identically zero. Then all of its coefficient must be 0.

*Proof.* If  $p(x) = a_0 + a_1x + \cdots + a_nx^n$ , then  $a_j = \frac{p^{(j)}(0)}{j!}$ , If  $p(x) \equiv 0$ , then  $p^{(j)}(x) = 0$ , so  $a_j = \frac{0}{j!} = 0 \forall j$  ■

#### Corollary 4.3

Suppose  $p(x) \equiv q(x)$  for  $p, q \in \mathcal{P}(\mathbb{F})$ , then all coefficients of  $p$  are the same as all coefficients of  $q$ .

### 4.2 Zero of polynomials and their algebraic manifestations

#### Algorithm 4.4 (Euclidean Algorithm for polynomials)

Given  $p(x), s(x)$ , without the loss of generality,  $\deg p(x) > \deg s(x)$ , otherwise it's boring; we can always find  $q(x), r(x)$  such that  $p(x) = s(x)q(x) + r(x)$ , where  $\deg r(x) < \deg s(x)$ .

#### Corollary 4.5

$p(a) = 0 \iff p(x) = (x - a)q(x)$  for some  $a \in \mathbb{F}$ .

*Proof.* If  $p(a) = (x - a)q(x)$ , then  $p(a) = 0 \cdot q(a) = 0$ .

Conversely, suppose  $p(a) = 0$ , by division algorithm we have  $p(x) = (x - a)q(x) + r(x)$ , where  $\deg r \leq \deg(x - a)$ , therefore  $r(x) = c$  for some  $c \in \mathbb{F}$ . Plug in  $a$  and we get  $(a - a)q(a) + c = 0 \implies 0 + c = 0 \implies c = 0$ . Therefore  $p(x) = (x - a)q(x)$ . ■

**Theorem 4.6**

Let  $p(x)$  be a nonzero polynomial with coefficients in  $\mathbb{F}$  have degree  $n$ . Then  $p$  has at most  $n$  zeros in  $\mathbb{F}$ .

*Proof.*

*Base case:*  $\deg p = 1$ , i.e.  $p(x) = a_1x + a_0$  for some  $a_1 \in \mathbb{F}^\times, a_0 \in \mathbb{F}$ . Then  $p\left(\frac{-b}{a}\right) = 0$ , so  $p$  has exactly one zero.

*Inductive Hypothesis:* Suppose the statement is true for all polynomials for all polynomials of degree less than  $m$ .

*Inductive Step:* Take  $p(x)$  to be a degree  $m$  polynomial. If  $p$  has no zeros in  $\mathbb{F}$ , we are done. If  $p$  has a zero, by corollary we have  $p(x) = (x - a)q(x)$ , where  $\deg q = m - 1$ . So the inductive hypothesis applies and  $q$  at most  $n - 1$  distinct zeros in  $\mathbb{F}$ . ■

**Theorem 4.7 (Fundamental Theorem of Algebra)**

Every nonconstant polynomial with complex coefficients has a zero.

*Proof with “Black Box” from Complex Analysis.*

Assume  $\deg p \geq 1$ . Assume that  $p(a) \neq 0 \forall a \in \mathbb{C}$ . Consider the function  $\frac{1}{p(x)}$  is well-defined  $\forall x \in \mathbb{C}$  and is analytic in  $\mathbb{C}$ , more over  $\lim_{|z| \rightarrow \infty} \frac{1}{p(z)} = 0$ . We know that

$$\begin{aligned} p(x) &= a_0 + a_1x + \cdots + a_nx^n \\ &= x^n \left( \frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \cdots + a_n \right) \\ \frac{1}{p(x)} &= \frac{1}{x^n \left( \frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \cdots + a_n \right)} \end{aligned}$$

As  $|x| \rightarrow \infty$ ,  $\frac{1}{x^n} \rightarrow 0$ . Since  $\left| \frac{1}{x^n} \right| = \frac{1}{|x|^n} \rightarrow 0$ . But  $\frac{a_0}{x^n} + \frac{a_1}{x^{n-1}} + \cdots + a_n \rightarrow a_n \neq 0$ . Hence  $\frac{1}{p(x)} \rightarrow 0$  as  $|x| \rightarrow \infty$ .

By Louisville’s theorem, any analytic function with this property has to be constant. But  $\frac{1}{p(x)}$  is non-constant, so  $p$  must have at least 1 zero in  $\mathbb{C}$ . ■

**Corollary 4.8**

Any polynomial  $p(x)$  with coefficients in  $\mathbb{C}$  factors as follows

$$p(x) = c(x - a_1)(x - a_2) \cdots (x - a_m), c \neq 0$$

*Proof.* By Induction it’s clear for degree 1 and if  $\deg p = m$  then factor  $p(x) = (x - a)q(x)$  and repeat the process for  $q$ . ■

**Question 4.9**

What happens over  $\mathbb{R}$ ?



**Theorem 4.10**

If  $p(x)$  has coefficient in  $\mathbb{R}$ , and  $c \in \mathbb{C}$  is a zero of  $p$ , then  $\bar{c}$  is also a zero of  $p$ .

*Proof.*  $p(c) = 0$  means

$$a_0 + a_1c + a_2c^2 + \cdots + a_nc^n = 0$$

We then can see

$$\overline{a_0 + a_1c + a_2c^2 + \cdots + a_nc^n} = \overline{0} = 0$$

$$\overline{a_0} + \overline{a_1c} + \overline{a_2c^2} + \cdots + \overline{a_nc^n} = 0$$

$$a_0 + a_1\bar{c} + a_2\bar{c}^2 + \cdots + a_n\bar{c}^n = 0$$

Hence  $p(\bar{c}) = 0$  as well. ■

So over  $\mathbb{C}$ , a polynomial with real coefficient factors as follows

$$p(x) = (x - a_1)(x - a_2) \cdots (x - a_n)(x - \lambda_1)(x - \bar{\lambda}_1)(x - \lambda_2)(x - \bar{\lambda}_2) \cdots (x - \lambda_m)(x - \bar{\lambda}_m)$$

For some  $c \in \mathbb{R}, a_1, a_2, \dots, a_n \in \mathbb{R}, \lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{C}$ .

To translate this into a factorization over  $\mathbb{R}$ , we can see that  $x^2 - (\lambda + \bar{\lambda})x + |\lambda|^2$ . These are quadratic with  $\Delta < 0$ . Indeed,

$$(\lambda + \bar{\lambda})^2 - 4|\lambda|^2 = \lambda^2 - 2|\lambda|^2 + \bar{\lambda}^2 = 2\operatorname{Re}\lambda^2 - 2|\lambda|^2$$

Notice that  $\operatorname{Re}\lambda^2 \leq |\lambda|^2$  and  $\operatorname{Re}\lambda^2 = |\lambda|^2$  iff  $\lambda \in \mathbb{R}$ , therefore  $\Delta < 0$ .

**Question 4.11**

Why do we study polynomials?

**Answer 4.12.**

1. We will form polynomials in linear operators
2. We will associate special polynomials with linear operators.

**Remark 4.13**

An operator has the same co-domain as its domain.

## 5 Eigenvalues, Eigenvectors, and Invariant Subspaces

### 5.1 Invariant Subspaces

#### Definition 5.1

Let  $T \in \mathcal{L}(V, V)$  on a vector space  $V \neq \{0\}$ . A subspace  $U \subseteq V$  is called an invariant subspace is invariant under  $T$  if  $T\mathbf{u} \in U \ \forall \mathbf{u} \in U$ .

#### Example 5.2

For any  $T \in \mathcal{L}(V, V)$ , the following subspaces are invariant:

1.  $\{0\}$
2.  $V$
3.  $\text{Null } T = \{\mathbf{v} \in V : T\mathbf{v} = 0\}$   
If  $T\mathbf{v} \in \text{Null } T$ , then  $T\mathbf{v} = 0 \in \text{Null } T$ .
4.  $\text{Range } T = \{\mathbf{w} \in W : \mathbf{w} = T\mathbf{v} \text{ for some } \mathbf{v} \in V\}$   
So  $T\mathbf{w} \in \text{Range } T$ .

#### Question 5.3

What are 1-dimensional invariant subspaces?

**Answer 5.4.** Then  $U = \text{span}(\mathbf{u})$  for some  $\mathbf{u} \neq 0$ . Invariant means  $T\mathbf{u} = \lambda\mathbf{u}$  for some  $\lambda \in \mathbb{F}$ , where  $\mathbf{u}$  is the eigenvector of  $T$  and  $\lambda$  is the eigenvalues.

#### Remark 5.5

$\mathbf{u} \neq 0$  if  $\mathbf{u}$  is a eigenvector is  $T$ .  $\lambda = 0$  is possible.

#### Proposition 5.6

Let  $T$  be a linear operator in  $V$ , then the following are equivalent

1.  $\lambda$  is a eigenvalue of  $T$ .
2.  $T - \lambda\mathbb{I}$  is not invertible.
3.  $T - \lambda\mathbb{I}$  is not injective.
4.  $T - \lambda\mathbb{I}$  is not surjective.

We have already proven that statement 2, 3, 4 are logically equivalent.

**Theorem 5.7**

Suppose  $v_1, v_2, \dots, v_m$  are eigenvectors of  $T \in \mathcal{L}(V)$  corresponding to distinct eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_m$  will be linearly independent.

*Proof.* Suppose  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  are linearly independent. By linear dependence lemma, we find a the minimum index  $k \leq m$  such that  $\mathbf{v}_k \in \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1})$ . i.e.

$$\mathbf{v}_k = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_{k-1} \mathbf{v}_{k-1} \quad (1)$$

Apply linear transformation on both sides

$$T\mathbf{v}_k = T\alpha_1 \mathbf{v}_1 + \alpha_2 T\mathbf{v}_2 + \dots + \alpha_{k-1} T\mathbf{v}_{k-1} \quad (2)$$

$$\lambda \mathbf{v}_k = \alpha_1 \lambda_1 \mathbf{v}_1 + \alpha_2 \lambda_2 \mathbf{v}_2 + \dots + \alpha_{k-1} \lambda_{k-1} \mathbf{v}_{k-1} \quad (3)$$

We multiply by equation 1 by  $\lambda_m$  and subtract by from 3 and we get

$$0 = \alpha_1(\lambda_1 - \lambda_k)\mathbf{v}_1 + \alpha_2(\lambda_2 - \lambda_k)\mathbf{v}_2 + \dots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{v}_{k-1}$$

A contradiction since  $k$  is not the minimum index with the property chosen above. Therefore the list  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  must be linearly independent. ■

**Corollary 5.8**

An operator  $T \in \mathcal{L}(V)$  has at most  $\boxed{\dim V}$  distinct eigenvalues.

**5.1.1 Restriction Operators****Definition 5.9**

Suppose  $T \in \mathcal{L}(V)$  and  $U$  is a subspace of  $V$  invariant under  $T$ . Then the restriction operator  $T|_U \in \mathcal{L}(U)$  is defined by  $T|_U(\mathbf{u}) = T\mathbf{u}$  for all  $\mathbf{u} \in U$ .

**5.2 Eigenvectors and Upper-Triangular Matrices****5.2.1 Polynomials in T****Definition 5.10**

Suppose  $T \in \mathcal{L}(V)$ , then  $T^k$  is defined as

$$T^k := \underbrace{k \circ k \circ \dots \circ k}_{k \text{ times}}$$

Notice that  $T^0 = \mathbb{I}, T^1 = T$ .

**Definition 5.11**

If  $p(x) = a_0 + a_1x + \cdots + a_nx^n$ , then we can define  $p(T)$  as  $a_0\mathbb{I} + a_1T + a_2T^2 + \cdots + a_nT^n$ .

**Example 5.12**

Let  $V := \mathcal{P}(\mathbb{R})$ ,  $S : p \mapsto 3p'' + 2p' + p$ ,  $D : p \mapsto p'$ . We can see that  $S$  can be expressed as  $S = D^0 + 2D + 3D^2$ . Therefore

$$\mathcal{M}(S) = 3\mathcal{M}^2(D) + 2\mathcal{M}(D) + \mathcal{M}(\mathbb{I})$$

we need to have to take the same basis for inputs and output when forming  $\mathcal{M}(\cdot)$ . Let's use our favorite basis  $1, x, x^2, x^3$ . We then can see

$$\mathcal{M}(D) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \mathcal{M}(S) = \begin{bmatrix} 1 & 2 & 6 & 0 \\ 0 & 1 & 4 & 18 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

**Question 5.13**

What is the best matrix representation for an operator?

**Question 5.14**

What information about eigenvalues/eigenvectors can be read off from a matrix representation?

**Theorem 5.15**

Suppose  $T \in \mathcal{L}(V)$  and  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  is a basis of  $V$ . Then the following are logically equivalent:

1.  $\mathcal{M}(T)$  is upper triangular.
2.  $T\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j) \quad \forall j = 1, 2, \dots, n$ .
3.  $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j)$  is invariant under  $T \quad \forall j = 1, 2, \dots, n$ .

*Proof.* 1)  $\implies$  2)

$$\begin{bmatrix} * & * & * & * & \cdots & * \\ & * & * & * & \cdots & * \\ & & * & * & \cdots & * \\ & & & * & \cdots & * \\ & & & & \ddots & \vdots \\ & & & & & * \end{bmatrix}$$

We can see that 2) holds true by inspection.

2)  $\implies$  3) Consider  $T\mathbf{v}_h$  for  $h \leq j$ , by 2) we have  $T\mathbf{v}_k \in \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_h) \subseteq \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j)$ . So  $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j)$  is invariant under  $T$ .

3)  $\implies$  2) Consider  $T\mathbf{v}_j$ , by 3) it is a linear combination of  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j$  because  $T\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j)$  so  $\mathcal{M}(T)(i, j) = 0$  if  $i > j$ . ■

### Question 5.16

What about conditions for lower-triangular matrices?

### Lemma 5.17

Over  $\mathbb{C}$ , every linear operator has at least one eigenvalue.

*Proof.* Take  $\mathbf{v} \in V \setminus \{0\}$ , and consider the list  $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \dots, T^n\mathbf{v}$  where  $n = \dim V$ . There is a nontrivial linear combination of these vectors which is 0. Suppose the equation

$$a_0\mathbf{v} + a_1T\mathbf{v} + a_2T^2\mathbf{v} + \cdots + a_nT^n\mathbf{v} = 0$$

i.e.  $p(T)v = 0$  for nonconstant  $p(x) := a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ . By the fundamental theorem of algebra  $p$  splits into linear factors over  $\mathbb{C}$ .

$$p(x) = c(x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_m)$$

for some  $m \leq n$ . Therefore

$$p(T)v = c(T - \lambda_1\mathbb{I})(T - \lambda_2\mathbb{I}) \cdots (T - \lambda_m\mathbb{I})$$

Therefore at least one of these factors is not injective. This shows that  $T$  has at least 1 eigenvalue. ■

### Theorem 5.18

For any  $T \in \mathcal{L}(V)$ ,  $V$  is finite dimensional vector space over  $\mathbb{C}$ , there exists its matrix representation  $\mathcal{M}(T)$  which is upper-triangular.

*Proof.* We can induct on the dimension of  $V$ . *Base Step.*  $n = 1$  is trivially true.

*Inductive Hypothesis.* Suppose Theorem holds for all vector spaces of dimension less than  $\dim V$ .

*Inductive Step.* Consider  $\lambda \in \mathbb{C}$  an eigenvalue of  $T$  by lemma. We can define

$$U := \text{Range } (T - \lambda \mathbb{I})$$

$U$  is a subspace of  $V$ . By the characterization of eigenvalues,  $T - \lambda \mathbb{I}$  is not surjective, hence  $\text{Range } T - \lambda \mathbb{I} \neq V$ , hence  $\dim \text{Range } (T - \lambda \mathbb{I}) < \dim V$ . We want to show that  $U$  is invariant under  $T$ . Suppose  $\mathbf{v} \in U$ , then

$$T\mathbf{v} = \underbrace{(T - \lambda \mathbb{I})\mathbf{v}}_{\in U} + \underbrace{\lambda \mathbf{v}}_{\in U}$$

therefore we know that  $U$  is invariant under  $T$ . Consider

$$T|_U \in \mathcal{L}(U) : (T|_U)(\mathbf{v}) := T\mathbf{v} \forall \mathbf{v} \in U$$

If  $U \neq \{0\}$ , then there is a basis  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$  of  $U$  ( $m < n$ ) such that the matrix representation of  $T|_U$  with respect to  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$  is upper triangular by the inductive hypothesis. Extend  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$  to a basis of  $V$ ,  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ . We compute

$$T\mathbf{v}_j = \underbrace{(T - \lambda \mathbb{I})\mathbf{v}_j}_{\in U = \text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m)} + \lambda \mathbf{v}_j$$

We also know that  $T\mathbf{u}_i \in \text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{i-1})$ . We can see the matrix representation and hence we are done

$$m \left[ \begin{array}{cccc|ccc} * & * & \cdots & * & * & * & * \\ 0 & * & \cdots & * & * & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & * & * & * \\ \hline 0 & 0 & \cdots & 0 & \lambda & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \lambda & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \lambda \end{array} \right]$$

■

### Question 5.19

What about eigenvalues of an upper-triangular matrix?

### Theorem 5.20

An upper triangular matrix is invertible if and only if all its diagonal entries are nonzero.

*Proof.* Suppose all diagonal entries are nonzero. Prove surjectivity.

$$\begin{aligned}
 T\mathbf{v}_1 &= A_{1,1}v_1, A_{1,1} \neq 0 \implies \mathbf{v}_1 \in \text{Range } T \\
 T\mathbf{v}_2 &= A_{1,2}\mathbf{v}_1 + A_{2,2}\mathbf{v}_2, A_{2,2} \neq 0 \implies \mathbf{v}_2 \in \text{Range } T \\
 &\vdots \implies \\
 T\mathbf{v}_n &= A_{1,n}\mathbf{v}_1 + A_{2,n}\mathbf{v}_2 + \cdots + A_{n,n}\mathbf{v}_n \neq 0 \implies \mathbf{v}_n \in \text{Range } T
 \end{aligned}$$

Therefore  $\text{Range } T = V$ , so  $T$  is surjective, hence  $T$  is invertible. Suppose at least one diagonal entry is 0 we want to show that  $T$  is not invertible. Say  $A_{j,j} = 0$  for some  $j$  and upper triangular matrix  $A$ . If  $j = 1$ , then  $v_1 \in T$ , hence  $T$  is not invertible, and we are done. If  $j > 1$ , consider  $U := \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j)$ .  $T$  maps  $U$  to  $\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1})$ . This shows  $T|_U$  is not surjective, then we know that  $T|_U$  is not injective and there exists  $\mathbf{u} \in U$  such that  $\mathbf{u} \in T|_U \implies \mathbf{u} \in T$ . Therefore  $T$  is not injective. Hence  $T$  is not invertible. ■

### Corollary 5.21

An upper triangular matrix / operator in upper triangular form has the diagonal elements / entries as its eigenvalues.

### Example 5.22

The matrix

$$A = \begin{bmatrix} 5 & * & * & * & * \\ 0 & 9 & * & * & * \\ 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 8 & * \\ 0 & 0 & 0 & 0 & 10 \end{bmatrix}$$

has eigenvalue 1, 5, 9, 8, 10.

### Example 5.23

$T : \mathcal{P}_n(\mathbb{R}) \rightarrow \mathcal{P}_n(\mathbb{R}) : p \mapsto 3p'' - 5p' + 7p$  has eigenvalues 3, -5, 7.

## 5.3 Eigenspaces and Diagonal Matrices

### Definition 5.24

Suppose  $T \in \mathcal{L}(V)$  and  $\lambda \in \mathbb{F}$ . The eigenspace of  $T$  corresponding to  $\lambda$ , denoted as  $E(\lambda, T)$  is defined as

$$E(\lambda, T) := \{\mathbf{v} \in V : T\mathbf{v} = \lambda\mathbf{v}\} = \text{Null}(T - \lambda I)$$

**Definition 5.25**

An operator  $T \in \mathcal{L}(V)$  is called diagonalizable if the operator has a diagonal matrix with respect to some basis of  $V$ .

**Theorem 5.26**

For  $T \in \mathcal{L}(V)$ , where  $V$  is a finite dimensional vector space, then the following are equivalent

1.  $\mathcal{M}(T)$  is a diagonal matrix.
2. the corresponding basis for  $V$  consists of eigenvalue of  $T$ .
3.  $V = U_1 \oplus U_2 \oplus \cdots \oplus U_n$  where  $\dim U_j = 1$  and  $U_j$  is invariant under  $T$  for all  $j$ .
4.  $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ , where  $T|_{W_l} = \lambda_l \mathbb{I}$  for all  $l$  and  $W_l$  is invariant under  $T$ .
5.  $\dim V = \dim W_1 + \dim W_2 + \cdots + \dim W_k$ , where  $W_e = \ker(T - \lambda_e \mathbb{I})$ .

*Proof.* Refer to Axler Page 157. ■



## 6 Inner Product Spaces

### Motivation

#### Definition 6.1

In  $\mathbb{R}^n$ , the dot product of  $\mathbf{x}$  and  $\mathbf{y}$  is defined by

$$\mathbf{x} \cdot \mathbf{y} := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

for  $\mathbf{x} = (x_1, x_2, \dots, x_n), \mathbf{y} = (y_1, y_2, \dots, y_n)$ .

### 6.1 Inner Product and Norms

#### Settings

$V$  is a vector space over  $\mathbb{F}$ , we can define the following mapping  $\langle *, * \rangle : V \times V \rightarrow \mathbb{F}$ .

#### Definition 6.2

$\langle \cdot, \cdot \rangle$  is called an inner product if it satisfying the following rules:

1. (additivity in the first slot)  $\langle \mathbf{v} + \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{v} + \mathbf{w} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle, \forall \mathbf{v}, \mathbf{u}, \mathbf{w} \in V$
2. (homogeneity in the first slot)  $\langle \lambda \mathbf{v}, \mathbf{w} \rangle = \lambda \langle \mathbf{v} + \mathbf{w} \rangle, \forall \mathbf{v}, \mathbf{u}, \mathbf{w} \in V, \lambda \in \mathbb{F}$
3. (conjugate symmetry)  $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}, \forall \mathbf{v}, \mathbf{w} \in V$
4. (positivity)  $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0, \forall \mathbf{v} \in V$
5. (definiteness)  $\langle \mathbf{v}, \mathbf{v} \rangle = 0$  iff  $\mathbf{v} = \mathbf{0}$ .

#### Question 6.3

What about linearity in the second slot?

**Answer 6.4.** We can compute

$$\begin{aligned} \langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle &= \overline{\langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle} = \overline{\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle} = \overline{\langle \mathbf{u}, \mathbf{v} \rangle} + \overline{\langle \mathbf{w}, \mathbf{v} \rangle} = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle \\ \langle \mathbf{v}, \lambda \mathbf{u} \rangle &= \overline{\langle \lambda \mathbf{u}, \mathbf{v} \rangle} = \overline{\lambda \langle \mathbf{u}, \mathbf{v} \rangle} = \overline{\lambda} \overline{\langle \mathbf{u}, \mathbf{v} \rangle} = \overline{\lambda} \langle \mathbf{v}, \mathbf{u} \rangle \end{aligned}$$

Not quite. ☹

#### Remark 6.5

If  $\mathbf{v} \in V$  is fixed then the function  $\langle *, \mathbf{v} \rangle : \mathbf{u} \mapsto \langle \mathbf{u}, \mathbf{v} \rangle$  is a function functional.

**Example 6.6**

On  $\mathbb{R}^n$ , we could use any function of the type

$$c_1x_1y_1 + c_2x_2y_2 + \cdots + c_nx_ny_n$$

where all  $c_j \in \mathbb{R}^+$ .

**Remark 6.7** (Generalization to  $\mathbb{C}^n$ )

The inner product of this form of the standard product to  $\mathbb{C}^n$  can be defined as

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1\bar{y}_1 + x_2\bar{y}_2 + \cdots + x_n\bar{y}_n$$

**Remark 6.8** (Generalization to any function space)

$$\langle f, g \rangle := \int_D f(t)\overline{g(t)}dt$$

or generally

$$\langle f, g \rangle := \int_D f(t)\overline{g(t)}w(t)dt$$

where  $w(t)$  is the positive weight function. e.g. if  $V = \mathcal{P}(\mathbb{R})$ , or  $V = \mathcal{P}(\mathbb{C})$ , then

$$\langle f, g \rangle := \int_0^\infty f(t)\overline{g(t)}e^{-t}dt$$

**Definition 6.9**

For  $\mathbf{v} \in V$ , the (Euclidean) Norm is defined as

$$\|\mathbf{v}\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

**Theorem 6.10**

(Properties of Norms)

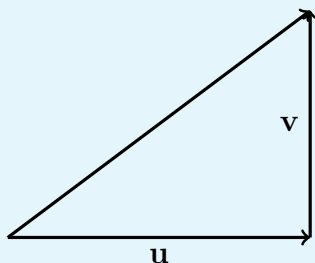
1.  $\|\lambda\mathbf{v}\| = |\lambda| \|\mathbf{v}\| \quad \forall \mathbf{v} \in V, \forall \lambda \in \mathbb{F}$
2.  $\|\mathbf{v}\| > 0$  for all  $\mathbf{v} \in V$
3.  $\|\mathbf{v}\| = 0$  if and only if  $\mathbf{v} = \mathbf{0}$

**Definition 6.11**

An inner product space is a vector space  $V$  along with an inner product on  $V$ .

**Definition 6.12**

For  $\mathbf{u}, \mathbf{v} \in V$ , we say  $\mathbf{u}$  and  $\mathbf{v}$  is orthogonal if  $\langle \mathbf{u}, \mathbf{v} \rangle = 0$

**Theorem 6.13** (Pythagorean Theorem)

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \iff \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle$$

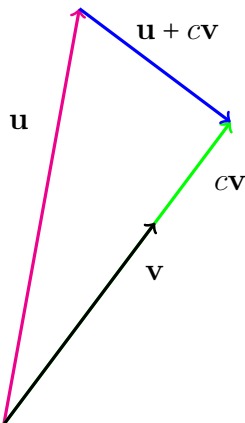
*Proof.* We compute

$$\langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 0 + 0 + \langle \mathbf{v}, \mathbf{v} \rangle = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

■

**Obeservation**

Given  $\mathbf{u}, \mathbf{v} \in V$  such that  $\mathbf{v} \neq 0$ , we want to modify  $\mathbf{u}$  such that  $\mathbf{u} + c\mathbf{v}$  is orthogonal to  $\mathbf{v}$ . We know that  $\langle \mathbf{v} + c\mathbf{v}, \mathbf{v} \rangle = 0$ , solve for  $c$  gives  $c = \frac{-\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}$ .



An orthogonal decomposition

**Theorem 6.14** (Cauchy-Schwarz Inequality)

For any  $u, v \in V$  where  $V$  is a inner product space, the following holds

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|$$

*Proof.* Given  $\mathbf{u}, \mathbf{v} \in V$ , we can assume without the loss of generality that  $\mathbf{v} \neq 0$ . So we can consider vectors  $\mathbf{u} + c\mathbf{v}$  and  $\mathbf{v}$  that are orthogonal for the choice that

$$c := \frac{-\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}$$

By Pythagorean theorem,  $\|\mathbf{u} + c\mathbf{v}\|^2 + \|c\mathbf{v}\|^2 = \|\mathbf{u}\|^2$ . But  $\|c\mathbf{v}\|^2 = |c|^2\|\mathbf{v}\|^2$  and recall

$$c = \frac{-\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}, \text{ so } c^2 = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\langle \mathbf{v}, \mathbf{v} \rangle^2} = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^4}, \text{ therefore } \|c\mathbf{v}\|^2 = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^4} \|\mathbf{v}\|^2 = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^2}$$

So by dropping  $\|\mathbf{u} + c\mathbf{v}\|^2 > 0$ , we obtain  $\|c\mathbf{v}\|^2 \leq \|\mathbf{u}\|^2$ , i.e.,

$$\frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^2} \leq \|\mathbf{u}\|^2 \implies |\langle \mathbf{u}, \mathbf{v} \rangle|^2 \leq \|\mathbf{u}\|^2 \cdot \|\mathbf{v}\|^2 \implies |\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|$$

■

### Theorem 6.15 (Triangle Inequality)

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

*Proof.* We have

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + \langle \mathbf{u}, \mathbf{v} \rangle + \overline{\langle \mathbf{u}, \mathbf{v} \rangle} = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\operatorname{Re}\langle \mathbf{u}, \mathbf{v} \rangle \\ &\leq \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\|\mathbf{u}\| \|\mathbf{v}\| \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned}$$

■

### Theorem 6.16 (Alternative Version of Triangle Inequality)

$$||\mathbf{u}\| - \|\mathbf{v}\|| \leq \|\mathbf{u} - \mathbf{v}\|$$

*Proof.* Notice that

$$\|\mathbf{u}\| - \|\mathbf{v}\| \leq \|\mathbf{u} - \mathbf{v}\| \iff \|\mathbf{u}\| \leq \|\mathbf{u} - \mathbf{v}\| + \|\mathbf{v}\|$$

Which is the triangle inequality. Swapping out  $\mathbf{u}$  and  $\mathbf{v}$  gives us

$$\|\mathbf{v}\| - \|\mathbf{u}\| \leq \|\mathbf{v} - \mathbf{u}\| \iff \|\mathbf{v}\| \leq \|\mathbf{v} - \mathbf{u}\| + \|\mathbf{u}\|$$

Combining these equations gives us

$$||\mathbf{u}\| - \|\mathbf{v}\|| \leq \|\mathbf{u} - \mathbf{v}\|$$

■

**Fact 6.17** (Fun inequalities)

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2)$$

**6.2 Orthogonality****Definition 6.18**

A list  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  in  $V$  is called orthonormal if

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

**Lemma 6.19**

Any list of orthonormal vectors is necessarily linearly independent.

*Proof.* Suppose  $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0}$ . We can compute on the standard inner product

$$\langle \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k, \mathbf{v}_1 \rangle = \alpha_1 \langle \mathbf{v}_1, \mathbf{v}_1 \rangle + \alpha_2 \langle \mathbf{v}_2, \mathbf{v}_1 \rangle + \dots + \alpha_k \langle \mathbf{v}_k, \mathbf{v}_1 \rangle \implies \alpha_1 = 0$$

$$\langle \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k, \mathbf{v}_2 \rangle = \alpha_1 \langle \mathbf{v}_1, \mathbf{v}_2 \rangle + \alpha_2 \langle \mathbf{v}_2, \mathbf{v}_2 \rangle + \dots + \alpha_k \langle \mathbf{v}_k, \mathbf{v}_2 \rangle \implies \alpha_2 = 0$$

$$\vdots$$

$$\langle \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k, \mathbf{v}_k \rangle = \alpha_1 \langle \mathbf{v}_1, \mathbf{v}_k \rangle + \alpha_2 \langle \mathbf{v}_2, \mathbf{v}_k \rangle + \dots + \alpha_k \langle \mathbf{v}_k, \mathbf{v}_k \rangle \implies \alpha_k = 0$$

Hence  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  is linearly independent. ■

**Question 6.20**

What is nice about orthonormal basis?

**Answer 6.21.** If  $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$  is an orthonormal basis, then an arbitrary vector can be written as

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{v}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \dots + \langle \mathbf{v}, \mathbf{v}_n \rangle \mathbf{v}_n$$

Furthermore, we can conclude the following theorem:

**Theorem 6.22** (Generalized Pythagorean Theorem)

$$\|\mathbf{v}\|^2 = \sum_{j=1}^n |\langle \mathbf{v}, \mathbf{v}_j \rangle|^2$$

**Algorithm 6.23** (Gram-Schmidt Algorithm)

**Input:** Any  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$  that is linearly independent.

**Output:**  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m$  such that  $\mathbf{e}_j \in \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j)$  for all  $j \leq m$ .

*Process.*

$$\begin{aligned}\mathbf{e}_1 &= \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \\ \mathbf{e}_2 &= \frac{\mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{e}_1 \rangle \mathbf{e}_1}{\|\mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{e}_1 \rangle \mathbf{e}_1\|} \\ \mathbf{e}_3 &= \frac{\mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{e}_1 \rangle \mathbf{e}_1 - \langle \mathbf{v}_3, \mathbf{e}_2 \rangle \mathbf{e}_2}{\|\mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{e}_1 \rangle \mathbf{e}_1 - \langle \mathbf{v}_3, \mathbf{e}_2 \rangle \mathbf{e}_2\|} \\ &\vdots \\ \mathbf{e}_n &= \frac{\mathbf{v}_n - \langle \mathbf{v}_n, \mathbf{e}_1 \rangle \mathbf{e}_1 - \langle \mathbf{v}_n, \mathbf{e}_2 \rangle \mathbf{e}_2 \cdots - \langle \mathbf{v}_n, \mathbf{e}_{n-1} \rangle \mathbf{e}_{n-1}}{\|\mathbf{v}_n - \langle \mathbf{v}_n, \mathbf{e}_1 \rangle \mathbf{e}_1 - \langle \mathbf{v}_n, \mathbf{e}_2 \rangle \mathbf{e}_2 \cdots - \langle \mathbf{v}_n, \mathbf{e}_{n-1} \rangle \mathbf{e}_{n-1}\|}\end{aligned}$$

■

**Proposition 6.24**

Every finite inner product vector space has a orthonormal basis.

*Proof.* Suppose  $V$  is a finite dimensional vector space. Let  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  be a basis for  $V$ . We then apply Gram-Schmidt Algorithm to the basis to obtain a orthonormal basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ . ■

**Remark 6.25** (Projection orthogonal with the respect to inner product)

Given a subspace  $U$  of  $V$  for finite dimensional vector space  $V$ , there is a projector  $P_U$  that project all vectors in  $V$  on  $U$  orthogonally.

**Remark 6.26** (Relations between inner product and linear functionals)

Suppose  $V$  is finite dimensional vector space. Given any  $\mathbf{u} \in V$ , then function  $\langle \cdot, \mathbf{u} \rangle$  is a linear functional (i.e. an element of  $V' = \mathcal{L}(V, \mathbb{F})$ )

**Theorem 6.27** (Riesz Representation Theorem)

For any  $\varphi \in V'$  there exists a *unique*  $\mathbf{u} \in V$  such that  $\langle \mathbf{v}, \mathbf{u} \rangle = \varphi(\mathbf{v})$  for all  $\mathbf{v} \in V$ .

**Proof. Existence**

Take an orthonormal basis  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  of  $V$ . Any  $\mathbf{v}$  can be written as a linear combination of the basis (To preserve linearity we want to put  $\mathbf{v}$  into the first slot)

$$\begin{aligned}\mathbf{v} &= \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2 + \dots + \langle \mathbf{v}, \mathbf{e}_n \rangle \mathbf{e}_n \\ \varphi(\mathbf{v}) &= \langle \mathbf{v}, \mathbf{e}_1 \rangle \varphi(\mathbf{e}_1) + \langle \mathbf{v}, \mathbf{e}_2 \rangle \varphi(\mathbf{e}_2) + \dots + \langle \mathbf{v}, \mathbf{e}_n \rangle \varphi(\mathbf{e}_n) \\ &= \langle \mathbf{v}, \overline{\varphi(\mathbf{e}_1)} \mathbf{e}_1 + \overline{\varphi(\mathbf{e}_2)} \mathbf{e}_2 + \dots + \overline{\varphi(\mathbf{e}_n)} \mathbf{e}_n \rangle \\ &= \mathbf{u}\end{aligned}$$

**Uniqueness**

Suppose there are  $\mathbf{u}_1, \mathbf{u}_2$  such that  $\langle \mathbf{v}, \mathbf{u}_1 \rangle = \langle \mathbf{v}, \mathbf{u}_2 \rangle$  for all  $\mathbf{v} \in V$ . Take  $\mathbf{v} = \mathbf{u}_1 - \mathbf{u}_2$ . We can see that

$$\langle \mathbf{v}_1, \mathbf{u}_1 \rangle + \langle \mathbf{v}, \mathbf{v}_2 \rangle \iff \langle \mathbf{v}, \mathbf{u}_1 - \mathbf{u}_2 \rangle$$

Plug in  $\mathbf{v}$  and we get

$$\langle \mathbf{u}_1 - \mathbf{u}_2, \mathbf{u}_1 - \mathbf{u}_2 \rangle = \|\mathbf{u}_1 - \mathbf{u}_2\|^2 = 0$$

This means that  $\mathbf{u}_1 - \mathbf{u}_2 = 0$ , or  $\mathbf{u}_1 = \mathbf{u}_2$ , hence such  $\mathbf{u}$  is unique. ■

**Example 6.28**

Let  $V := \mathcal{P}(\mathbb{C})$  and let  $\varphi(p) := \int_{-1}^1 p(t) \sin t \cdot dt$ . Find a representation in  $V$  with respect to

$$\langle f, g \rangle = \int_{-1}^1 f(t) \overline{g(t)} dt$$

Suppose  $p(t) = a_0 + a_1 t + a_2 t^2$ .

**6.3 Orthogonality and Orthogonal Projections****Definition 6.29**

Given an inner product space  $V$  and its subset  $U$  of  $V$  we can define

$$U^\perp := \{ \mathbf{v} \in V : \langle \mathbf{u}, \mathbf{v} \rangle = 0 \ \forall \mathbf{u} \in U \}$$

**Theorem 6.30**

Basic facts about orthogonal complement

- (a) If  $U$  is a subset of  $V$ , then  $U^\perp$  is a subspace of  $V$ .
- (b)  $\{\mathbf{0}\}^\perp = V$
- (c)  $V^\perp = \{\mathbf{0}\}$
- (d)  $U \cap U^\perp \subseteq \{\mathbf{0}\}$
- (e) If  $U \subseteq W$  then  $U^\perp \supseteq W^\perp$

*Proof.*

- (a) Clearly  $\mathbf{0} \in U^\perp$  as  $\langle \mathbf{v}, \mathbf{0} \rangle = 0 \ \forall \mathbf{v} \in V$ . Take  $\mathbf{v}_1, \mathbf{v}_2 \in U^\perp$  and  $\lambda \in \mathbb{F}$ , then  $\langle \mathbf{v}_1 + \lambda \mathbf{v}_2, \mathbf{u} \rangle = \langle \mathbf{v}_1, \mathbf{u} \rangle + \lambda \langle \mathbf{v}_2, \mathbf{u} \rangle = 0 \ \forall \mathbf{u} \in U$ .
- (b) Trivial by part (c).
- (c) Trivial by part (b).
- (d) Suppose  $\mathbf{v} \in U \cap U^\perp$ , then  $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \implies \mathbf{v} = \mathbf{0}$ .
- (e) Suppose  $\mathbf{v} \in W^\perp$ , then  $\langle \mathbf{v}, \mathbf{w} \rangle = 0 \ \forall \mathbf{w} \in W$ . Since  $U \subseteq W$ ,  $\langle \mathbf{v}, \mathbf{u} \rangle = 0 \ \forall \mathbf{u} \in U$ , hence  $U^\perp \supseteq W^\perp$ .

■

**Theorem 6.31**

If  $V$  is a finite dimensional inner product space and  $U$  is a subspace of  $V$ , then

$$U \oplus U^\perp = V$$

*Proof.* We already know that the sum is direct by  $U \cap U^\perp = \{\mathbf{0}\}$ . By Gram Schmidt we can construct an orthonormal basis of  $U$  and extend it to a normal orthonormal basis of  $V$  we have  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ . We claim that  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  is a basis of  $U^\perp$  as  $\langle \mathbf{u}_i, \mathbf{v}_j \rangle = 0$  for all  $i$  in  $1, 2, \dots, k$  and  $j$  in  $1, 2, \dots, n$ . Hence  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in U^\perp$ . On the other hand,

$$U^\perp \subseteq \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \dots + \alpha_k \mathbf{u}_k + \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2 + \dots + \beta_n \mathbf{v}_n$$

must satisfy  $\alpha_j = \langle \mathbf{v}_i, \mathbf{u}_j \rangle = 0$  for all  $i$  in  $1, 2, \dots, k$ . Therefore we have  $U^\perp = \text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$ . Hence  $U \oplus U^\perp = V$ . ■

**Theorem 6.32**

Suppose  $U$  is a finite dimensional subspace of  $V$ . Then

$$U = (U^\perp)^\perp$$



**Definition 6.33**

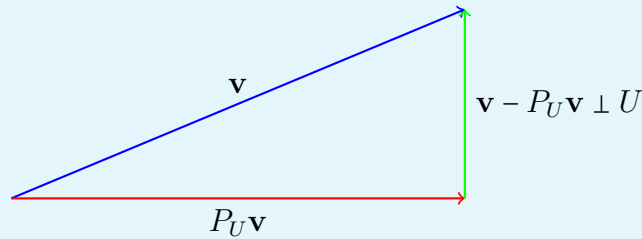
Suppose  $U$  is a finite-dimensional subspace of  $V$ . The orthogonal projection of  $V$  onto  $U$  is the operator  $P_U \in \mathcal{L}(V)$  defined as follows: For  $\mathbf{v} \in V$ , write  $\mathbf{v} = \mathbf{u} + \mathbf{w}$ , where  $\mathbf{u} \in U$  and  $\mathbf{w} \in U^\perp$ . Then  $P_U \mathbf{v} = \mathbf{u}$ .

**Theorem 6.34** (Properties of the orthogonal projection  $P_U$ )

Suppose  $U$  is finite dimensional subspace of  $V$  and  $\mathbf{v} \in V$ . Then

- (a)  $P_U \in \mathcal{L}(V)$
- (b)  $\text{Range } P_U = U$
- (c)  $\text{Null } P_U = U^\perp$
- (d)  $P_U^2 = P_U$
- (e)  $\text{Range } (\mathbb{I} - P_U) = U^\perp$
- (f)  $(\mathbb{I} - P_U)^2 = (\mathbb{I} - P_U)$
- (g)  $\|P_U \mathbf{v}\| \leq \|\mathbf{v}\|$
- (h) For every orthonormal basis  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  of  $U$ , we have

$$P_U \mathbf{v} = \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{v}, \mathbf{e}_2 \rangle \mathbf{e}_2 + \dots + \langle \mathbf{v}, \mathbf{e}_n \rangle \mathbf{e}_n$$



The red vector is the projection of the blue vector  $\mathbf{v}$  onto subspace  $U$ .

## 7 Operators on Inner Product Spaces

### 7.1 Self-Adjoint and Normal Operators

#### Definition 7.1

Suppose  $T \in L(V, W)$ , we define  $T^*$  by this formula

$$\langle T\mathbf{v}, \mathbf{w} \rangle_W = \langle \mathbf{v}, T^*\mathbf{w} \rangle_V$$

We can think of  $\mathbf{w}$  as fixed, we note that  $\langle T^*, \mathbf{w} \rangle$  is a linear function; hence it has a representer by Riesz, so we are entitled to call it  $T^*\mathbf{w}$  such that

$$\langle T\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, T^*\mathbf{w} \rangle$$

Hence we have  $T^* \in \mathcal{L}(W, V)$ . We want to verify this property. Consider  $T^*(\mathbf{w}_1 + \lambda\mathbf{w}_2)$ . We compute

$$\begin{aligned} \langle \mathbf{v}, T^*(\mathbf{w}_1 + \lambda\mathbf{w}_2) \rangle &= \langle T\mathbf{v}, \mathbf{w}_1 + \lambda\mathbf{w}_2 \rangle \\ &= \langle T\mathbf{v}, \mathbf{w}_1 \rangle + \bar{\lambda} \langle T\mathbf{v}, \mathbf{w}_2 \rangle \\ &= \langle \mathbf{v}, T^*\mathbf{w}_1 \rangle + \bar{\lambda} \langle \mathbf{v}, T^*\mathbf{w}_2 \rangle \\ &= \langle \mathbf{v}, T^*\mathbf{w}_1 \rangle + \langle \mathbf{v}, \lambda T^*\mathbf{w}_2 \rangle \\ &= \langle \mathbf{v}, T^*\mathbf{w}_1 + \lambda T^*\mathbf{w}_2 \rangle \forall \mathbf{v} \in V, \mathbf{w}_1, \mathbf{w}_2 \in W, \lambda \in \mathbb{F} \end{aligned}$$

So  $T^*(\mathbf{w}_1 + \lambda\mathbf{w}_2) = T^*\mathbf{w}_1 + \lambda T^*\mathbf{w}_2$ .

#### Theorem 7.2 (Properties of the adjoint)

Let  $T \in \mathcal{L}(V, W)$ , we have

1.  $(S + T)^* = S^* + T^*$
2.  $(\lambda T)^* = \bar{\lambda} T^*$
3.  $(S \cdot T)^* = T^* S^*$
4.  $(\lambda T)^* = \bar{\lambda} T^*$
5.  $\mathbb{I}^* = \mathbb{I}$

*Proof.* Refer to book page 206 ■

**Theorem 7.3** (Null space and range of adjoint)

Let  $T \in \mathcal{L}(V, W)$ , then

1.  $\text{Null } T^* = (\text{Range } T)^\perp$
2.  $\text{Range } T^* = (\text{Null } T)^\perp$
3.  $\text{Null } T = (\text{Range } T^*)^\perp$
4.  $\text{Range } T = (\text{Null } T^*)^\perp$

*Proof.* Let  $\mathbf{w} \in W$ . Then

$$\begin{aligned}
 \mathbf{w} \in \text{Null } T^* &\iff T^* \mathbf{w} = 0 \\
 &\iff \langle \mathbf{v}, T^* \mathbf{w} \rangle = 0 \quad \forall \mathbf{v} \in V \\
 &\iff \langle T \mathbf{v}, \mathbf{w} \rangle = 0 \quad \forall \mathbf{v} \in V \\
 &\iff \mathbf{w} \in (\text{Range } T)^\perp
 \end{aligned}$$

(b),(c),(d) follows by a similar logic and is left as an exercise. ■

**Definition 7.4**

Let  $T \in \mathcal{L}(V)$ .  $T$  is called self-adjoint if  $T^* = T$ .

**Definition 7.5**

$T$  is called normal if  $TT^* = T^*T$ .

**Example 7.6**

Suppose  $T \in \mathcal{L}(V)$ , we can define  $T : \mathbf{v} \mapsto \langle \mathbf{v}, \mathbf{x} \rangle \mathbf{y}$  for some fixed  $\mathbf{x}, \mathbf{y}$  in  $V$ . Compute  $T^*$ . We compute

$$\begin{aligned}
 \langle \mathbf{v}, T^* \mathbf{w} \rangle &= \langle T \mathbf{v}, \mathbf{w} \rangle \\
 &= \langle \langle \mathbf{v}, \mathbf{x} \rangle \mathbf{y}, \mathbf{w} \rangle \\
 &= \langle \mathbf{v}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{w} \rangle \\
 &= \langle \mathbf{v}, \overline{\langle \mathbf{y}, \mathbf{w} \rangle} \mathbf{x} \rangle \\
 &= \langle \mathbf{v}, \langle \mathbf{w}, \mathbf{y} \rangle \mathbf{x} \rangle
 \end{aligned}$$

hence we can conclude  $T^* \mathbf{w} = \langle \mathbf{w}, \mathbf{y} \rangle \mathbf{x}$  for all  $\mathbf{w} \in W$ .

### 7.1.1 Matrix representation

Suppose  $T \in \mathcal{L}(V, W)$ , where  $V, W$  are finite-dimensional vector spaces. Let  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  be an orthonormal basis for  $V$  and  $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m$  be a orthonormal basis for  $W$ . We can see that  $\mathcal{M}(T)$  is obtained through

$$T\mathbf{e}_j = \langle T\mathbf{e}_j, \mathbf{f}_1 \rangle \mathbf{f}_1 + \langle T\mathbf{e}_j, \mathbf{f}_2 \rangle \mathbf{f}_2 + \dots + \langle T\mathbf{e}_j, \mathbf{f}_m \rangle \mathbf{f}_m$$

$$T^*\mathbf{f}_k = \langle T^*\mathbf{f}_k, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle T^*\mathbf{f}_k, \mathbf{e}_2 \rangle \mathbf{e}_2 + \dots + \langle T^*\mathbf{f}_k, \mathbf{e}_n \rangle \mathbf{e}_n$$

Then

$$\mathcal{M}(T^*)(l, k) = \langle T^*\mathbf{f}_k, \mathbf{e}_l \rangle \implies \mathcal{M}(T^*)(j, i) = \langle T^*\mathbf{f}_i, \mathbf{e}_j \rangle$$

Therefore we have  $\mathcal{M}(T^*) = \overline{\mathcal{M}(T)}^T$

#### Remark 7.7

The above statement only holds if the basis for  $V$  and  $W$  are orthonormal.

#### Remark 7.8

If an operator  $T$  is self-adjoint, then  $T$  is normal, but not the converse.

#### Proposition 7.9

The eigenvalue of any self-adjoint operator is real.

*Proof.* Suppose self-adjoint  $T \in \mathcal{L}(V)$ . and  $\lambda$  is an eigenvalue of  $T$  and let  $\mathbf{v}$  be the eigenvector correspond to  $\lambda$ . We compute

$$\lambda \langle \mathbf{v}, \mathbf{v} \rangle = \langle \lambda \mathbf{v}, \mathbf{v} \rangle = \langle T\mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v}, T\mathbf{v} \rangle = \langle \mathbf{v}, \lambda \mathbf{v} \rangle = \bar{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle$$

We can see that  $\lambda = \bar{\lambda} \implies \lambda \in \mathbb{R}$ . ■

#### Question 7.10

Suppose  $\langle T\mathbf{v}, \mathbf{v} \rangle = 0 \forall \mathbf{v} \in V$ . Does the statement implies  $T$  is the zero map?

**Answer 7.11** (Surprisingly). Yes over  $\mathbb{C}$  and no over  $\mathbb{R}$ .

*Proof.* Suppose  $\mathbb{F} = \mathbb{C}$ , the following holds.

$$\langle T(\mathbf{v} + \mathbf{w}), \mathbf{v} + \mathbf{w} \rangle = \langle T\mathbf{v}, \mathbf{v} \rangle + \langle T\mathbf{v}, \mathbf{w} \rangle + \langle T\mathbf{w}, \mathbf{v} \rangle + \langle T\mathbf{w}, \mathbf{w} \rangle$$

$$\langle T(\mathbf{v} - \mathbf{w}), \mathbf{v} + \mathbf{w} \rangle = \langle T\mathbf{v}, \mathbf{v} \rangle - \langle T\mathbf{v}, \mathbf{w} \rangle - \langle T\mathbf{w}, \mathbf{v} \rangle + \langle T\mathbf{w}, \mathbf{w} \rangle$$

Subtrac the first equation by the second we have

$$\langle T(\mathbf{v} + \mathbf{w}), (\mathbf{v} + \mathbf{w}) \rangle - \langle T(\mathbf{v} - \mathbf{w}), (\mathbf{v} + \mathbf{w}) \rangle = \boxed{2\langle T\mathbf{v}, \mathbf{w} \rangle + 2\langle T\mathbf{w}, \mathbf{v} \rangle}$$

We also compute

$$\langle T(\mathbf{v} + i\mathbf{w}), (\mathbf{v} + i\mathbf{w}) \rangle - \langle T(\mathbf{v} - i\mathbf{w}), (\mathbf{v} - i\mathbf{w}) \rangle = \boxed{2i(\langle T\mathbf{w}, \mathbf{v} \rangle - \langle T\mathbf{v}, \mathbf{w} \rangle)}$$

Take the two boxed equation and divide the second one by  $i$  then subtract from first gives us

$$4\langle T\mathbf{v}, \mathbf{w} \rangle = 0$$

Suppose  $\mathbb{F} = \mathbb{R}$ . Consider  $\mathbb{R}^2$ . Take  $T\mathbf{v}$  and rotate  $\pi/2$  gives us  $T(x_1, x_2) := (-x_2, x_1)$ . We can see that  $\langle T\mathbf{v}, \mathbf{v} \rangle = 0 \forall \mathbf{v}$  but  $T \neq 0$ . However, if  $T$  is self-adjoint then  $T$  is 0. ■

### Remark 7.12

Suppose  $\mathbb{F} = \mathbb{R}$  and  $T = T^*$ . We have

$$4\langle T\mathbf{v}, \mathbf{w} \rangle = \langle T(\mathbf{v} + \mathbf{w}), (\mathbf{v} + \mathbf{w}) \rangle - \langle T(\mathbf{v} - \mathbf{w}), (\mathbf{v} - \mathbf{w}) \rangle$$

Hence  $T = 0$ .

### Corollary 7.13

$\langle T\mathbf{v}, \mathbf{v} \rangle \in \mathbb{R}$  in a complex space is equivalent to  $T$  being self adjoint.

*Proof.*

$$\langle T\mathbf{v}, \mathbf{v} \rangle \in \mathbb{R} \iff \langle T\mathbf{v}, \mathbf{v} \rangle = \langle T^*\mathbf{v}, \mathbf{v} \rangle \implies \langle (T - T^*)\mathbf{v}, \mathbf{v} \rangle = 0 \implies T - T^* = 0$$

We can see that  $T = T^*$ . Hence  $T$  is self-adjoint. ■

### Theorem 7.14

$T$  is normal if and only if  $\|T\mathbf{v}\| = \|T^*\mathbf{v}\| \quad \forall \mathbf{v} \in V$ .

*Proof.*

$$\|T\mathbf{v}\| = \|T^*\mathbf{v}\| \implies \langle T\mathbf{v}, T\mathbf{v} \rangle = \langle T^*\mathbf{v}, T^*\mathbf{v} \rangle \implies \langle T^*T\mathbf{v}, \mathbf{v} \rangle = \langle TT^*\mathbf{v}, \mathbf{v} \rangle$$

Hence  $T$  is normal since  $TT^* = T^*T$ . ■

### Theorem 7.15

Say  $\lambda, \mathbf{v}$  is an eigenpair of a normal operator  $T$ , then

$$\|(T - \lambda\mathbb{I})\mathbf{v}\| = \|(T^* - \bar{\lambda}\mathbb{I})\mathbf{v}\|$$

## 7.2 Spectral Theorem

### 7.2.1 Over Complex Vector Space

#### Theorem 7.16 (Spectral Theorem over Complex Vector Space)

Suppose  $T \in \mathcal{L}(V)$  where  $V$  is finite dimension vector space and  $\mathbb{F} = \mathbb{C}$  and  $T$  is normal. Then  $V$  has an orthonormal basis of eigenvectors of  $T$ , and vice versa, if  $T$  has a diagonal representation with respect to some orthonormal basis, then  $T$  is normal.

*Proof.* Suppose  $T$  has a diagonal matrix representation with respect to some orthonormal basis. i.e.

$$\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \quad \mathcal{M}(T^*) = \begin{bmatrix} \bar{\lambda}_1 & & & 0 \\ & \bar{\lambda}_2 & & \\ & & \ddots & \\ 0 & & & \bar{\lambda}_n \end{bmatrix}$$

Since any two diagonal matrices commute, we can see that

$$\mathcal{M}(T)\mathcal{M}(T^*) = \mathcal{M}(T^*)\mathcal{M}(T) = \begin{bmatrix} |\lambda_1|^2 & & & 0 \\ & |\lambda_2|^2 & & \\ & & \ddots & \\ 0 & & & |\lambda_n|^2 \end{bmatrix}$$

We have  $TT^* = T^*T$ , hence  $T$  is normal.

Suppose  $T$  is normal. By Schur's Theorem, there exists an orthonormal basis such that

$$\mathcal{M}(T) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \implies \mathcal{M}(T^*) = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & \cdots & \bar{a}_{1n} \\ 0 & \bar{a}_{22} & \cdots & \bar{a}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \bar{a}_{nn} \end{bmatrix}$$

Recall that  $\|T\mathbf{v}\| = \|T^*\mathbf{v}\| \quad \forall \mathbf{v} \in V$ . Call this this orthonormal basis  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ . We have  $T\mathbf{e}_1 = a_{11}\mathbf{e}_1$ , so  $\|T\mathbf{e}_1\| = |a_{11}|$ , we then compute

$$\begin{aligned} T^*\mathbf{e}_1 &= \bar{a}_{11}\mathbf{e}_1 + \bar{a}_{12}\mathbf{e}_2 + \cdots + \bar{a}_{1n}\mathbf{e}_n \\ \|T^*\mathbf{e}_1\| &= \sqrt{|a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2} \end{aligned}$$

Since  $\|T\mathbf{e}_1\| = \|T^*\mathbf{e}_1\|$ , we get  $|a_{12}| = |a_{13}| = \cdots = |a_{1n}| = 0$ .

Using a similar logic, we have  $\|T\mathbf{e}_j\| = \|T^*\mathbf{e}_j\|$  implies  $|a_{jj+1}| = |a_{jj+2}| = \cdots = |a_{jn}| = 0$ . Hence  $T$  is diagonal ■

#### Remark 7.17

So actually the schur form of a normal operator is necessarily diagonal.

### 7.2.2 Over Real Vector Space

#### Lemma 7.18

Suppose  $T \in \mathcal{L}(V)$  is self-adjoint and  $\beta, \gamma \in \mathbb{R}$  such that  $b\beta^2 - 4\gamma$  then

$$T^2 + \beta T + \gamma I$$

is invertible.

*Proof.* Consider nonzero  $\mathbf{v} \in V$ . We can factor

$$\begin{aligned} \langle (T^2 + \beta T + \gamma I)\mathbf{v}, \mathbf{v} \rangle &= \langle T^2\mathbf{v}, \mathbf{v} \rangle + \langle \beta T\mathbf{v}, \mathbf{v} \rangle + \gamma \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \langle T\mathbf{v}, T\mathbf{v} \rangle + \beta \langle T\mathbf{v}, \mathbf{v} \rangle + \gamma \|\mathbf{v}\|^2 \\ &\geq \|T\mathbf{v}\|^2 - |\beta| \cdot \|T\mathbf{v}\| \cdot \|\mathbf{v}\| + \gamma \|\mathbf{v}\|^2 \\ &= \left( \|T\mathbf{v}\| - \frac{|\beta| \cdot \|\mathbf{v}\|}{2} \right)^2 + \left( \gamma - \frac{\beta^2}{4} \right) \|\mathbf{v}\|^2 \\ &> 0 \end{aligned}$$

hence we can see that  $\text{Null}(T^2 + \beta T + \gamma I) = \{0\}$ . Hence it's injective. Since  $T^2 + \beta T + \gamma I \in \mathcal{L}(V)$ , we know  $(T^2 + \beta T + \gamma I)$  is invertible. ■

#### Theorem 7.19

$T$  has a eigenvalue if  $T$  is self-adjoint in any vector space.

*Proof.* Assume  $\dim V = n$ . Consider any  $\mathbf{v} \in V$ . Then  $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \dots, T^n\mathbf{v}$  are linearly dependent (I.e, there exist  $a_1, a_2, \dots, a_n \in \mathbb{R}$  such that

$$a_0\mathbf{v} + a_1T\mathbf{v} + \dots + a_nT^n\mathbf{v} = 0$$

Consider  $f(x) = a_0x + a_1x + \dots + a_nx^n$ . We know from chapter 4 we can factor  $f(x)$  as

$$f(x) = c(x^2 + \beta_1x + \gamma_1) \cdots (x^2 + \beta_mx + \gamma_m)(x - \lambda_1) \cdots (x - \lambda_n)$$

where all coefficients are real and  $\beta_i^2 - 4\gamma_i < 0$ . By lemma we know that the quadratic term is invertible, then we can simply factor them out. Therefore we have

$$0 = (T - \lambda_1 I) \cdots (T - \lambda_m I)\mathbf{v}$$

Hence one of the  $(T - \lambda_j I)$  is not injective. Hence  $T$  has an eigenvalue. ■

#### Theorem 7.20

Suppose  $T \in \mathcal{L}(V)$ , where  $V$  is a finite dimensional vector space and  $\mathbb{F} = \mathbb{R}$  and  $T$  is self-adjoint. Then  $T$  has a diagonal matrix representation with some orthonormal basis for  $V$ . And conversely, if  $T$  has a diagonal matrix representation with respect to some orthonormal basis, then  $T = T^*$ .

*Proof.* Suppose  $T$  has a diagonal matrix representation with some respect to some orthonormal basis. i.e.

$$\mathcal{M}(T) = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \quad \mathcal{M}(T^*) = \begin{bmatrix} \bar{\lambda}_1 & & & 0 \\ & \bar{\lambda}_2 & & \\ & & \ddots & \\ 0 & & & \bar{\lambda}_n \end{bmatrix}$$

We know  $\mathcal{M}(T) = \mathcal{M}(T^*)$  since  $\lambda = \bar{\lambda}$  in reals. Hence  $T$  is self-adjoint. Conversely, suppose  $T = T^*$ . We just found out  $T$  has at least one eigenvalue eigenvector pair. Say  $T\mathbf{u} = \lambda\mathbf{u}$ . Without the loss of generality  $\|\mathbf{u}\| = 1$ . If  $\mathbf{w} \perp \mathbf{u}$ , then  $\langle T\mathbf{u}, \mathbf{w} \rangle = 0 = \langle \mathbf{u}, T\mathbf{w} \rangle$ . So  $T\mathbf{w} \perp \mathbf{u}$ . Notice that  $T|_{\text{span}(\mathbf{u})^\perp}$  is still self-adjoint.

$$\langle T|_{\text{span}(\mathbf{u})^\perp} \mathbf{w}_1, \mathbf{w}_2 \rangle = \langle T\mathbf{w}_1, \mathbf{w}_2 \rangle = \langle \mathbf{w}_1, T\mathbf{w}_2 \rangle = \langle \mathbf{w}_1, T|_{\text{span}(\mathbf{u})^\perp} \mathbf{w}_2 \rangle \quad \forall \mathbf{w}_1, \mathbf{w}_2 \in \text{span}(\mathbf{u})^\perp$$

Hence  $\mathcal{M}(T) = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ 0 & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$  Now the problem is reduce to that for  $T|_{\text{span}(\mathbf{u})^\perp}$ , which has a

dimensional of  $\dim V - 1$ . By induction we can build a orthonormal basis of  $V$  which consists of eigenvectors. ■



Last updated: April 24, 2019