

# MAT224 Linear Algebra II

## Lecture Notes

Tianyu Du

January 2018

### Info.

**Created:** January. 9 2018

**Last modified:** April 6, 2018

This work is licensed under a Creative Commons  
“Attribution-ShareAlike 3.0 Unported” license.



### Contents

<b>1</b>	<b>Lecture1 Jan.9 2018</b>	<b>3</b>
1.1	Vector spaces . . . . .	3
1.2	Examples of vector spaces . . . . .	4
1.3	Some properties of vector spaces . . . . .	5
<b>2</b>	<b>Lecture2 Jan.10 2018</b>	<b>6</b>
2.1	Some properties of vector spaces-Cont'd . . . . .	6
2.2	Subspaces . . . . .	7
2.3	Examples of subspaces . . . . .	8
2.4	Recall from MAT223 . . . . .	9
<b>3</b>	<b>Lecture3 Jan.16 2018</b>	<b>9</b>
3.1	Linear Combination . . . . .	9
3.2	Combination of subspaces . . . . .	10
<b>4</b>	<b>Lecture4 Jan.17 2018</b>	<b>13</b>
4.1	Cont'd . . . . .	13
4.2	Linear Independence . . . . .	14

<b>5</b>	<b>Lecture5 Jan.23 2018</b>	<b>15</b>
5.1	Linear independence, recall definitions . . . . .	15
5.1.1	Alternative definitions of linear independency . . . . .	15
5.2	Basis . . . . .	15
5.3	Dimensions . . . . .	18
5.3.1	Consequences of fundamental theorem . . . . .	19
5.3.2	Use dimension to prove facts about linearly (in)dependent sets and subspaces . . . . .	19
<b>6</b>	<b>Lecture6 Jan.24 2018</b>	<b>20</b>
6.1	Basis and Dimension . . . . .	20
<b>7</b>	<b>Lecture7 Jan.30. 2018</b>	<b>23</b>
7.1	Linear Transformations . . . . .	23
7.2	Properties of linear transformations . . . . .	24
7.3	Definitions . . . . .	26
<b>8</b>	<b>Lecture8 Jan.31 2018</b>	<b>27</b>
8.1	Linear Transformations . . . . .	27
8.2	Applications of dimension theorem . . . . .	29
<b>9</b>	<b>Lecture9 Feb.6 2018</b>	<b>30</b>
9.1	Applications of dimension theorem . . . . .	30
9.2	Isomorphisms . . . . .	33
9.3	Coordinates . . . . .	35
<b>10</b>	<b>Lecture10 Feb.7 2018</b>	<b>36</b>
10.1	Matrix of linear transformation . . . . .	36
<b>11</b>	<b>lecture11 Feb.13 2018</b>	<b>38</b>
11.1	Algebra of Transformation . . . . .	38
11.2	Matrix of composition of transformations . . . . .	39
11.3	Inverse transformations . . . . .	40
11.4	Change of basis . . . . .	41
<b>12</b>	<b>Lecture12 Feb.14 2018</b>	<b>42</b>
<b>13</b>	<b>Lecture13 Feb.27 2018</b>	<b>44</b>
13.1	Diagonalization . . . . .	44
13.2	How to find eigenvalues and eigenvectors of $T$ . . . . .	44

<b>14 Lecture14 Feb.28 2018</b>	<b>47</b>
<b>15 Lecture15 Mar.6 2018</b>	<b>50</b>
15.1 Fields . . . . .	50
15.2 Complex Numbers . . . . .	51
<b>16 Lecture16 Mar.7 2018</b>	<b>52</b>
16.1 Vector space over a field . . . . .	52
16.2 Complex vector space . . . . .	52
<b>17 Lecture 17 Mar.13 2018</b>	<b>53</b>
<b>18 Lecture 18 Mar.14 2018</b>	<b>54</b>
18.1 Triangular form . . . . .	54
18.2 Nilpotent transformation . . . . .	55
<b>19 Lecture 19 Mar.20 2018</b>	<b>57</b>
<b>20 Lecture 20 Mar.21 2018</b>	<b>58</b>
20.1 Nilpotent Transformations . . . . .	58
20.2 Canonical Forms for Transformations $T : V \rightarrow V$ with Single Eigenvalue $\lambda$ . . . . .	59
20.3 Graph(Computational aspect) . . . . .	59
<b>21 Lecture21 Mar.27 2018</b>	<b>59</b>
21.1 Goal . . . . .	59
<b>22 Lecture22 Mar.28 2018</b>	<b>64</b>
22.1 Examples on finding JCF. . . . .	64

## 1 Lecture1 Jan.9 2018

### 1.1 Vector spaces

**Definition** A real<sup>1</sup> **vector space** is a set  $V$  together with two vector operations vector addition and scalar multiplication such that

1. **AC** Additive Closure:  $\forall \vec{x}, \vec{y} \in V, \vec{x} + \vec{y} \in V$
2. **C** Commutative:  $\forall \vec{v}, \vec{y} \in V, \vec{x} + \vec{y} = \vec{y} + \vec{x}$

---

<sup>1</sup>A vector space is real if scalar which defines scalar multiplication is real.

3. **AA** Additive Associative:  $\forall \vec{x}, \vec{y}, \vec{z} \in V, (\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$
4. **Z** Zero Vector:  $\exists \vec{0} \in V \text{ s.t. } \forall \vec{x} \in V, \vec{x} + \vec{0} = \vec{x}$
5. **AI** Additive Inverse:  $\forall \vec{x} \in V, \exists -\vec{x} \in V \text{ s.t. } \vec{x} + (-\vec{x}) = \vec{0}$
6. **SC** Scalar Closure:  $\forall \vec{x}, c \in \mathbb{R}, c\vec{x} \in V$
7. **DVA** Distributive Vector Additions:  $\forall \vec{x}, \vec{y} \in V, c \in \mathbb{R}, c(\vec{x} + \vec{y}) = c\vec{x} + c\vec{y}$
8. **DSA** Distributive Scalar Additions:  $\forall \vec{x} \in V, c, d \in \mathbb{R}, (c+d)\vec{x} = c\vec{x} + d\vec{x}$
9. **SMA** Scalar Multiplication Associative:  $\forall \vec{x} \in V, c, d \in \mathbb{R}, (cd)\vec{x} = c(d\vec{x})$
10. **O** One:  $\forall \vec{x} \in V, 1\vec{x} = \vec{x}$

**Note** For  $V$  to be a vector space, need to know or be given operations of vector additions multiplication and check all 10 properties hold.

## 1.2 Examples of vector spaces

**Example 1**  $\mathbb{R}^n$  w.r.t.<sup>1</sup> usual component-wise addition and scalar multiplication.

**Example 2**  $M_{m \times n}(\mathbb{R})$  set of all  $m \times n$  matrices with real entry. w.r.t. usual entry-wise addition and scalar multiplication.

**Example 3**  $\mathbb{P}_n(\mathbb{R})$  set of polynomials with real coefficients, of degree less or equal to  $n$ , w.r.t. usual degree-wise polynomial addition and scalar multiplication.

**Note** If define  $\mathbb{P}_n^*(\mathbb{R})$  as set of all polynomials of degree exactly equal to  $n$  w.r.t. normal degree-wise multiplication and addition.

Then it is **NOT** a vector space.

**Explanation:**  $(1 + x^n), (1 - x^n) \in \mathbb{P}_n^*(\mathbb{R})$  but  $(1 + x^n) + (1 - x^n) = 2 \notin \mathbb{P}_n^*(\mathbb{R})$

---

<sup>1</sup>w.r.t. is the abbreviation of "with respect to".

**Example 4** Something unusual, define  $V$  as

$$V = \{(x_1, x_2) | x_1, x_2 \in \mathbb{R}\}$$

with vector addition

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1 + 1, x_2 + y_2 + 1)$$

and scalar multiplication

$$c(x_1, x_2) = (cx_1 + c - 1, cx_2 + c - 1)$$

This is a vector space.

### 1.3 Some properties of vector spaces

Suppose  $V$  is a vector space, then it has the following properties.

**Property 1** The zero vector is unique.

*proof.*

Assume  $\vec{0}, \vec{0}^*$  are two zero vectors in  $V$

$$\text{WTS: } \vec{0} = \vec{0}^*$$

Since  $\vec{0}$  is the zero vector, by Z  $\vec{0}^* + \vec{0} = \vec{0}^*$

$$\text{Similarly, } \vec{0} + \vec{0}^* = \vec{0}$$

Also,  $\vec{0} + \vec{0}^* = \vec{0}^* + \vec{0}$  by commutative vector addition.

$$\text{So, } \vec{0}^* = \vec{0}$$

■

**Property 2**  $\forall \vec{x} \in V$ , the additive inverse  $-\vec{x}$  is unique.

*proof.*

Exercise. (By Cancellation Law)

**Property 3**  $\forall \vec{x} \in V, 0\vec{x} = \vec{0}$ .

*proof.*

By property of number 0:  $0\vec{x} = (0 + 0)\vec{x}$

By DSA:  $0\vec{x} = 0\vec{x} + 0\vec{x}$

By AI,  $\exists(-0\vec{x})$  s.t.

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

By AA

$$\implies 0\vec{x} = \vec{0}$$

**Property 4**  $\forall c \in \mathbb{R}, c\vec{0} = \vec{0}$

*proof.*

$$c\vec{0} = c(\vec{0} + \vec{0}) = c\vec{0} + c\vec{0}$$

## 2 Lecture2 Jan.10 2018

### 2.1 Some properties of vector spaces-Cont'd

**Property 5** For a vector space  $V$ ,  $\forall \vec{x} \in V$ ,  $(-1)\vec{x} = (-\vec{x})$ . (we could use this property to find the additive inverse with scalar multiplication with  $(-1)$ )<sup>1</sup>.

*proof.*

$$\begin{aligned} (-\vec{x}) &= (-\vec{x}) + \vec{0} \quad \text{By property of zero vector} \\ &= (-\vec{x}) + 0\vec{x} \quad \text{By property 3} \\ &= (-\vec{x}) + (1 + (-1))\vec{x} \quad \text{By property of zero as real number} \\ &= (-\vec{x}) + 1\vec{x} + (-1)\vec{x} \\ &= \vec{0} + (-1)\vec{x} \\ &= (-1)\vec{x} \end{aligned}$$

■

**Property 6** For a vector space  $V$ , let  $\vec{x} \in V$  and  $c \in \mathbb{R}$ , then,

$$c\vec{x} = \vec{0} \implies c = 0 \vee \vec{x} = \vec{0}$$

<sup>1</sup>The scalar multiplication here is the one defined in vector space  $V$ .

*proof.*

$$\begin{aligned}
 &\text{if } c = 0 \implies \text{True} \\
 &\text{else } c^{-1}c\vec{x} = c^{-1}\vec{0} \\
 &\implies (c^{-1}c)\vec{x} = \vec{0} \\
 &\implies 1\vec{x} = \vec{0} \\
 &\implies \vec{x} = \vec{0} \\
 &\implies \text{True}
 \end{aligned}$$

■

## 2.2 Subspaces

**Loosely** A subspace is a space contained within a vector space.

**Definition** Let  $V$  be a vector space and  $W \subseteq V$ ,  $W$  is a **subspace** of  $V$  if  $W$  is itself a vector space w.r.t. operations of vector addition and scalar multiplication from  $V$ .

**Theorem** Let  $V$  be a vector space, and  $W \subseteq V$ ,  $W$  has the same<sup>1</sup> operations of vector addition and scalar multiplication as in  $V$ . Then,  $W$  is a subspace of  $V$  iff:

1.  $W$  is non-empty.  $W \neq \emptyset$ .
2.  $W$  is closed under addition.  $\forall \vec{x}, \vec{y} \in W, \vec{x} + \vec{y} \in W$ .
3.  $W$  is closed under scalar multiplication.  $\forall \vec{x} \in W, c \in \mathbb{R}, c\vec{x} \in W$ .

*Proof.*

---

<sup>1</sup>Other properties of vector spaces related to vector addition and scalar multiplication are immediately inherited from the parent vector space.

Forward:

If  $W$  is a subspace

$$\implies \vec{0} \in W$$

$$\implies W \neq \emptyset$$

Also, additive and scalar multiplication closures  $\implies (ii), (iii)$

Backward:

Let  $W \neq \emptyset \wedge (ii) \wedge (iii)$

WTS. 10 axioms in definition of vector space hold

$(ii) \implies$  Additive Closure

$(iii) \implies$  Scalar Multiplication Closure

Because  $W \subseteq V$ , and  $V$  is a vector space, so properties hold  $\forall \vec{w} \in W$ .

Additive inverse: by property 5 and scalar multiplication closure,

$$\forall \vec{x} \in W, -\vec{x} = (-1)\vec{x} \in W.$$

Also, existence of additive identity:  $(-\vec{x}) + \vec{x} = \vec{0} \in W$ .

## 2.3 Examples of subspaces

**Example 1** Let  $V = \mathbb{M}_{n \times n}(\mathbb{R})$ ,  $V$  is a subspace.

**Example 2** Define  $W$  as

$$W = \{A \in \mathbb{M}_{n \times n}(\mathbb{R}) \mid A \text{ is not symmetric}\}$$

*Explanation:* Let  $A_1 = \begin{bmatrix} 0 & -2 \\ -1 & 0 \end{bmatrix}$  and  $A_2 = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$   $A_1, A_2 \in W$  but

$$A_1 + A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \notin W.$$

Since there's no additive identity in set  $W$ , so  $W$  failed to be a vector space, therefore  $W$  is not a subspace.

**Example 3** Let  $V = \mathbb{P}_2(\mathbb{R})$ , is  $W$  defined as following,

$$W = \{p(x) \in V \mid p(1) = 0\}$$



a subspace of  $V$  ?

*proof.*

WTS: (i)

Let  $z(x) = 0$  or  $z(x) = x^2 - 1, \forall x \in \mathbb{R}$

$\implies W \neq \emptyset$

WTS: (ii)

Let  $p_1, p_2 \in W$ , which means  $p_1(1) = p_2(1) = 0$

$(p_1 + p_2)(1) = p_1(1) + p_2(1) = 0 + 0 = 0$

$\implies p_1 + p_2 \in W$

$\implies W$  is closed under addition.

WTS: (iii) Let  $p \in W$  and  $c \in \mathbb{R}$

$\implies p(1) = 0$

Since  $(c * p)(x) = c * p(x)$ , we have  $(c * p)(1) = c * p(1) = c * 0 = 0$

$\implies cp \in W$ .

So  $W$  is a subspace of  $V$ .



## 2.4 Recall from MAT223

Let  $A \in \mathbb{M}_{m \times n}(\mathbb{R})$ , then  $Nul(A)$  is a subspace of  $\mathbb{R}^n$  and  $Col(A)$  is a subspace of  $\mathbb{R}^m$ .

## 3 Lecture3 Jan.16 2018

### 3.1 Linear Combination

**Definition** Let  $V$  be a vector space,  $\vec{v}_1, \dots, \vec{v}_n \in V$ ,  $a_1, \dots, a_n \in \mathbb{R}$  the expression

$$c_1 \vec{v}_1 + \dots + c_n \vec{v}_n$$

is called a **linear combination** of  $\vec{v}_1, \dots, \vec{v}_n$ .

**Theorem** Let  $V$  be a vector space,  $W$  is a subspace of  $V$ ,  $\forall \vec{w}_1, \dots, \vec{w}_k \in W$ ,  $c_1, \dots, c_k \in \mathbb{R}$ , we have

$$c_1 \vec{w}_1 + \dots + c_k \vec{w}_k \in W$$

*Subspaces are closed under linear combinations, since subspaces are closed under scalar multiplication and vector addition.*

**Theorem** Let  $V$  be a vector space, let  $\vec{v}_1, \dots, \vec{v}_k \in V$  then the set of all linear combination of  $\vec{v}_1, \dots, \vec{v}_k$

$$W = \left\{ \sum_{i=1}^k c_i \vec{v}_i \mid c_i \in \mathbb{R} \forall i \right\}$$

is a subspace of  $V$ .

*proof.*

Consider  $\vec{0} \in W$

So,  $W \neq \emptyset$

Let  $c \in \mathbb{R}$ , Let  $\vec{x} \in W \wedge \vec{y} \in W$

By definition of span, we have,

$$\vec{x} = \sum_{i=1}^k a_i \vec{v}_i, \quad \vec{y} = \sum_{i=1}^k b_i \vec{v}_i$$

Consider,  $\vec{x} + c\vec{y}$

$$\vec{x} + c\vec{y} = \sum_{i=1}^k a_i \vec{v}_i + c \sum_{i=1}^k b_i \vec{v}_i = \sum_{i=1}^k (a_i + cb_i) \vec{v}_i \in W$$

■

**Definition** Let  $V$  be a vector space,  $\vec{v}_1, \dots, \vec{v}_k \in V$ , **span** of the set of vectors  $\{\vec{v}_i\}_{i=1}^k$  is defined as the collection of all possible linear combinations of  $\{\vec{v}_i\}_{i=1}^k$ . By pervious theorem, span is a subspace.

### 3.2 Combination of subspaces

**Definition** Let  $W_1, W_2$  be two sets, then the **union** of  $W_1, W_2$  is defined as:

$$W_1 \cup W_2 = \{\vec{w} \mid \vec{w} \in W_1 \vee \vec{w} \in W_2\}$$

the **intersection** of  $W_1, W_2$  is defined as:

$$W_1 \cap W_2 = \{\vec{w} \mid \vec{w} \in W_1 \wedge \vec{w} \in W_2\}$$

Now consider  $W_1, W_2$  to be two subspaces of vector space  $V$ , then we have,

1.  $W_1 \cup W_2$  is **not** a subspace.

2.  $W_1 \cap W_2$  is a subspace.

*proof.*

Falsify the statement by providing counter-example:

Consider,

$$W_1 = \{(x_1, x_2) \mid x_1 \in \mathbb{R}, x_2 = 0\}$$

$$W_2 = \{(x_1, x_2) \mid x_2 \in \mathbb{R}, x_1 = 0\}$$

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \in W_1 \cup W_2 \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in W_1 \cup W_2$$

$$\text{But, } \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin W_1 \cup W_2$$

■

*proof.*

Because  $W_1$  and  $W_2$  are both subspaces, so

$$\vec{0} \in W_1 \cap W_2 \implies W_1 \cap W_2 \neq \emptyset$$

$$\text{Let } \vec{x}, \vec{y} \in W_1 \cap W_2, c \in \mathbb{R}$$

$$\text{Consider, } \vec{x} + c\vec{y}$$

Sine  $W_1, W_2$  are subspaces,

$$\vec{x} + c\vec{y} \in W_1 \wedge \vec{x} + c\vec{y} \in W_2$$

$$\implies \vec{x} + c\vec{y} \in W_1 \cap W_2$$

So,  $W_1 \cap W_2$  is a subspace.

■

**Definition** Let  $W_1, W_2$  be subspaces of vector space  $V$ , define the **sum** of two subspaces as:

$$W_1 + W_2 = \{\vec{x} + \vec{y} \mid \vec{x} \in W_1 \wedge \vec{y} \in W_2\}$$

**Note** Let  $\vec{x} = \vec{0} \in W_1, \forall \vec{y} \in W_2, \vec{y} \in W_1 + W_2$  so that,  $W_2 \subseteq W_1 + W_2$ . Similarly, let  $\vec{y} = \vec{0} \in W_2, \forall \vec{x} \in W_1, \vec{x} \in W_1 + W_2$ . so that,  $W_1 \subseteq W_1 + W_2$ . So we have  $\forall \vec{v} \in W_1 \cap W_2, \vec{v} \in W_1 + W_2$ . So that,

$$W_1 \cap W_2 \subseteq W_1 + W_2$$

**Note**  $W_1 + W_2$  is a subspace of  $V$ .  
*proof.*

Let  $\vec{x}_1, \vec{x}_2 \in W_1, \vec{y}_1, \vec{y}_2 \in W_2$

By properties of subspaces,

$$\forall c \in \mathbb{R}, \vec{x}_1 + c\vec{x}_1 \in W_1 \wedge \vec{y}_2 + c\vec{y}_2 \in W_2$$

Consider,  $\vec{x}_1 + \vec{y}_1 \in W_1 + W_2, \vec{x}_2 + \vec{y}_2 \in W_1 + W_2$

$$\begin{aligned} & (\vec{x}_1 + \vec{y}_1) + c(\vec{x}_2 + \vec{y}_2) \\ &= (\vec{x}_1 + c\vec{x}_2) + (\vec{y}_1 + c\vec{y}_2) \in W_1 + W_2 \end{aligned}$$

■

**Definition(Unique Representation)** Let  $W_1, W_2$  be subspaces of vector space  $V$ , say  $V$  is **direct sum** of  $W_1$  and  $W_2$ , written as  $V = W_1 \oplus W_2$ , if every  $\vec{x} \in V$  can be written uniquely as  $\vec{x} = \vec{w}_1 + \vec{w}_2$  where  $\vec{w}_1 \in W_1$  and  $\vec{w}_2 \in W_2$ .

**Equivalently** Let  $W_1$  and  $W_2$  be subspaces of  $V$ ,  $V = W_1 \oplus W_2 \iff V = W_1 + W_2 \wedge W_1 \cap W_2 = \{\vec{0}\}$ .

## 4 Lecture4 Jan.17 2018

### 4.1 Cont'd

#### Cont'd Proof of Theorem

*proof.*

(Forward direction) Suppose  $V = W_1 \oplus W_2$

WTS.  $V = W_1 + W_2 \wedge W_1 \cap W_2 = \{\vec{0}\}$

Let  $V = W_1 \oplus W_2$

$\implies \forall \vec{x} \in V$ , can be written uniquely as

$\vec{x} = \vec{w}_1 + \vec{w}_2$ ,  $\vec{w}_1 \in W_1$ ,  $\vec{w}_2 \in W_2$

$\implies V = W_1 + W_2$  by definition of *sum*.

Let  $\vec{x} \in W_1 \cap W_2$

Decomposition, let  $\vec{z} \in W_1 \cap W_2 \subseteq V$

$\vec{z} = \vec{z} + \vec{0}$ ,  $\vec{z} \in W_1, \vec{0} \in W_2$

$\vec{z} = \vec{0} + \vec{z}$ ,  $\vec{0} \in W_1, \vec{z} \in W_2$

Since decomposition is unique,  $\vec{z} = \vec{0}$

So,  $W_1 \cap W_2 = \{\vec{0}\}$

(Backward direction) Suppose  $V = W_1 + W_2 \wedge W_1 \cap W_2 = \{\vec{0}\}$

WTS.  $V = W_1 \oplus W_2$

Assume  $\vec{x} = \vec{w}_1 + \vec{w}_2$ ,  $\vec{w}_1 \in W_1, \vec{w}_2 \in W_2$

$\vec{x} = \vec{w}'_1 + \vec{w}'_2$ ,  $\vec{w}'_1 \in W_1, \vec{w}'_2 \in W_2$

$\implies \vec{w}_1 + \vec{w}_2 = \vec{w}'_1 + \vec{w}'_2$

$\implies \vec{w}_1 - \vec{w}'_1 = \vec{w}'_2 - \vec{w}_2$

Where, by definition of subspace,  $\vec{w}_1 - \vec{w}'_1 \in W_1 \wedge \vec{w}'_2 - \vec{w}_2 \in W_2$

So,  $\vec{w}_1 - \vec{w}'_1 = \vec{w}'_2 - \vec{w}_2 \in W_1 \cap W_2$

Since  $W_1 \cap W_2 = \{\vec{0}\}$

$\implies \vec{w}_1 = \vec{w}'_1 \wedge \vec{w}_2 = \vec{w}'_2$

So the decomposition is unique.

■

## 4.2 Linear Independence

**Theorem (Redundancy theorem)** Let  $V$  be a vector space,  $\{\vec{x}_1, \dots, \vec{x}_n\}$ , let  $\vec{x} \in \{\vec{x}_1, \dots, \vec{x}_n\}$ , then

$$\text{span}\{\vec{x}_1, \dots, \vec{x}_n, \vec{x}\} = \text{span}\{\vec{x}_1, \dots, \vec{x}_n\}$$

we say  $\vec{x}$  is the **redundant** vector that contributes nothing to the span.  
proof.

$$\begin{aligned} \text{let } \vec{x} &\in \text{span}\{\vec{x}_1, \dots, \vec{x}_n\} \\ \vec{x} &= \sum_{i=1}^n c_i \vec{x}_i \text{ for } c_i \in \mathbb{R} \forall i \\ \text{So, } \text{span}\{\vec{x}_1, \dots, \vec{x}_n, \vec{x}\} &= \left\{ \sum_{i=1}^n a_i \vec{x}_i + z \vec{x} \mid a_i, z \in \mathbb{R} \forall i \right\} \\ &= \left\{ \sum_{i=1}^n a_i \vec{x}_i + z \sum_{i=1}^n c_i \vec{x}_i \mid a_i, c_i \in \mathbb{R} \forall i \right\} \\ &= \left\{ \sum_{i=1}^n (a_i + z c_i) \vec{x}_i \mid a_i, c_i \in \mathbb{R} \forall i \right\} \\ \text{Let } d_i &= a_i + z c_i \in \mathbb{R} \\ &= \left\{ \sum_{i=1}^n d_i \vec{x}_i \mid d_i \in \mathbb{R} \forall i \right\} \\ &= \text{span}\{\vec{x}_1, \dots, \vec{x}_n\} \end{aligned}$$

■

**Definition** Let  $V$  be a vector space, let  $\{\vec{x}_1, \dots, \vec{x}_n\} \in V$ , we say  $\{v_i\}_{i=1}^n$  is **linearly independent** if the only set of scalars  $\{c_1, \dots, c_n\}$  that satisfies,

$$\sum_{i=1}^n c_i \vec{x}_i = 0$$

is  $\{0, \dots, 0\}$ .

**Definition** In contrast, we say a set of vector, with size  $n$ , is **linearly dependent** if

$$\exists \vec{c} \neq \vec{0} \in \mathbb{R}^n, \text{ s.t. } \sum_{i=1}^n c_i \vec{v}_i = 0$$

**Theorem** Let  $V$  be a vector space,  $\{\vec{v}_i\}_{i=1}^n \in V$  is *linearly dependent* if and only if,

$$\exists \vec{x} \in \{\vec{v}_i\}_{i=1}^n \text{ s.t. } \vec{x} \in \text{span}\{\{\vec{v}_i\}_{i=1}^n \setminus \{\vec{x}\}\}$$

**Theorem** Let  $V$  be a vector space,  $\{\vec{v}_i\}_{i=1}^n \in V$  is *linearly independent* if and only if,

$$\forall \vec{x} \in \{\vec{v}_i\}_{i=1}^n, \vec{x} \notin \text{span}\{\{\vec{v}_i\}_{i=1}^n \setminus \{\vec{x}\}\}$$

## 5 Lecture5 Jan.23 2018

### 5.1 Linear independence, recall definitions

*Acknowledgement: special thanks to Frank Zhao.*

**Definition** Let  $\{\vec{x}_1, \dots, \vec{x}_k\}$  is **linearly independent** if only scalars  $c_1 \dots c_k$  s.t.

$$\sum_{i=1}^k c_i \vec{x}_i = 0(\star)$$

are  $c_1 = \dots = c_k = 0$

**linearly dependent** means at least one  $c_i \neq 0$ ,  $(\star)$  still holds.

#### 5.1.1 Alternative definitions of linear independency

**Definition(Alternative.1)**  $\{\vec{x}_1 \dots \vec{x}_k\}$  is **linearly independent** iff none of them can be written as a linear combination of the remaining  $k - 1$  vectors.<sup>1</sup>

**Definition(Alternative.2)**  $\{\vec{x}_1 \dots \vec{x}_k\}$  is **linearly dependent** iff at least one of them can be written as a linear combination of the remaining  $k - 1$  vectors.<sup>2</sup>

### 5.2 Basis

**Definition** Let  $V$  be a vector space, a non-empty<sup>3</sup> set  $S$  of vectors from  $V$  is a **basis** for  $V$  if

1.  $V = \text{span}\{S\}$

<sup>1</sup>See theorem from the pervious lecture.

<sup>2</sup>See theorem from the pervious lecture.

<sup>3</sup>Specially, for an empty set, we define  $\text{span } \emptyset = \{\vec{0}\}$

2.  $S$  is linearly independent.

**Theorem (characterization of basis)** A non-empty subset  $S = \{\vec{x}_i\}_{i=1}^n$  of vector space  $V$  is basis for  $V$  iff every  $\vec{x} \in V$  can be written uniquely as linear combination for vectors in  $S$ .



*proof.*

### Forwards

Suppose  $S$  is a basis for  $V$

So every  $\vec{x} \in V$  can be written as a linear combination of vectors in  $S$

To prove the uniqueness, assume two expressions of  $\vec{x} \in V$

$$\vec{x} = \begin{cases} c_1\vec{x}_1 + \cdots + c_k\vec{x}_k \\ b_1\vec{x}_1 + \cdots + d_k\vec{x}_k \end{cases}$$

Consider,

$$c_1\vec{x}_1 + \cdots + c_k\vec{x}_k - (b_1\vec{x}_1 + \cdots + d_k\vec{x}_k) = \vec{0}$$

$$\iff \sum_{i=1}^k (c_i - b_i)\vec{x}_i = \vec{0}$$

Since vectors in basis  $S$  are linear independent,

$$c_i = b_i \forall i \in \mathbb{Z} \cap [1, k]$$

So the representation is unique.

### Backwards

Suppose every  $\vec{x} \in V$  can be written uniquely as linear combination of vectors in  $S$ .

WTS:  $V = \text{span}\{S\} \wedge S$  is linearly independent

By the assumption, spanning set is shown.

All we need to show is linear independence.

Consider,

$$\sum_{i=1}^n c_i \vec{x}_i = \vec{0}$$

Also, we know

$$\sum_{i=1}^n 0\vec{x}_i = \vec{0}$$

By the uniqueness of representation

$$\text{We have identical expression } \sum_{i=1}^n c_i \vec{x}_i = \sum_{i=1}^n 0\vec{x}_i$$

$$\therefore c_i = 0 \forall i \in \mathbb{Z} \cap [1, n]$$

■

**Example**

$$\begin{aligned}
 V &= \{(x_1, x_2) \mid x_1, x_2 \in \mathbb{R}\} \\
 (x_1, x_2) + (y_1, y_2) &= (x_1 + y_1 + 1, x_2 + y_2 + 1) \\
 c(x_1, x_2) &= (cx_1 + c - 1, cx_2 + c - 1)
 \end{aligned}$$

Show that  $\{(1, 0), (6, 3)\}$  is a basis of  $V$ .

By theorem,  $\{(1, 0), (6, 3)\}$  is basis if every  $(a, b) \in V$  can be written uniquely as linear combination of  $\{(1, 0), (6, 3)\}$ .

$$\exists \text{ unique scalars } c_1, c_2 \in \mathbb{R} \text{ s.t. } c_1(1, 0) + c_2(6, 3) = (a, b)$$

*proof.*

By definition of scalar multiplication and vector addition in this space,

$$\begin{aligned}
 \text{Consider } (a, b) &= c_1(1, 0) + c_2(6, 3) = (2c_1 - 1, c_1 - 1) + (7c_2 - 1, 4c_2 - 1) \\
 &= (2c_1 + 7c_2 - 1, c_1 + 4c_2 - 1)
 \end{aligned}$$

Consider the coefficients of variables

$$\begin{cases} 2c_1 + 7c_2 - 1 = a \\ c_1 + 4c_2 - 1 = b \end{cases}$$

WTS, the above system of linear equations has unique solution for all  $a, b$

The system has a unique solution  $\forall a, b \in \mathbb{R}$

Since the coefficient matrix has rank 2

$$\text{rank}\left(\begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix}\right) = 2$$

Since obviously the columns are linearly independent.

■

**5.3 Dimensions**

**Definition** For a vector space  $V$ , the **dimension** of  $V$  is the minimum number of vectors required to span  $V$ .

**Fundamental Theorem** if  $V$  vector space is spanned by  $m$  vectors, then any set of more than  $m$  vectors from  $V$  must be linearly dependent.

**Fundamental Theorem (Alternative)** If  $V$  is vector space spanned by  $m$  vectors, then any linearly independent set in  $V$  must contain less or equal to  $m$  vectors.

### 5.3.1 Consequences of fundamental theorem

**Theorem** if  $S = \{\vec{v}_i\}_{i=1}^k$  and  $T = \{\vec{w}_i\}_{i=1}^l$  are two bases of vector space  $V$  then  $l = k$ . *Bases have the same size.*

*proof.*

Since  $S$  spans  $V$  and  $T$  is linearly independent

$$\therefore l \leq k$$

(flip) Since  $T$  spans  $V$  and  $S$  is linearly independent

$$\therefore k \leq l$$

$$\implies l \leq k \wedge k \leq l$$

$$\implies k = l$$

■

**Definition** So we can define the **dimension** of  $V$ , as  $\dim(V)$  as the number vectors in any basis for  $V$ . For special case  $V = \{\vec{0}\}$ ,  $\dim(V) = 0$ .

#### Example

- $\dim(\mathbb{R}^n) = n$
- $\dim(\mathbb{P}_n(\mathbb{R})) = n + 1$
- $\dim(\mathbb{M}_{m \times n}(\mathbb{R})) = m \times n$

### 5.3.2 Use dimension to prove facts about linearly (in)dependent sets and subspaces

**Theorem** If  $V$  is a vector space,  $\dim(V) = n$ ,  $S = \{\vec{x}_k\}_{i=1}^k$  is subset of  $V$ , if  $k > n$  then  $S$  is linearly dependent.

**Note**  $k \leq n \nRightarrow S$  is linear dependent.

**Theorem** If  $W$  is subspace of vector space  $V$ , then

1.  $\dim(W) \leq \dim(V)$
2.  $\dim(W) = \dim(V) \iff W = V$

*proof.*

(1) Suppose  $\dim(V) = n, \dim(W) = k$

WTS,  $k \leq n$

Any basis for  $W$  is a linearly independent set of  $k$  vectors from  $V$ .

Since  $V$  is spanned by  $n$  vectors, since  $\dim(V) = n$

By fundamental theorem,  $k \leq n$

$$\iff \dim(W) \leq \dim(V)$$

(2) By contradiction, assume  $\dim(V) = \dim(W) = n$  but  $V \neq W$

Then  $\exists \vec{x} \in V \wedge \vec{x} \notin W$

Take  $S$  as a basis of  $W$ , then  $\vec{x} \notin \text{span}\{S\}$

Then  $S \cup \vec{x}$  is linearly independent

$\implies S \cup \{\vec{x}\}$  is linearly independent in  $V$  containing  $n + 1$  vectors

This contradicts the assumption by fundamental theorem since  $\dim(V) = n$  so it could not contain more than  $n$  linearly independent vectors

■

## 6 Lecture6 Jan.24 2018

### 6.1 Basis and Dimension

**Theorem** Let  $V$  be a vector space,  $S$  is a spanning set of  $V$ , and  $I$  is a linearly independent subset of  $V$ , s.t.  $I \subseteq S$ , then  $\exists$  basis  $B$  for  $V$  s.t.  $I \subseteq B \subseteq S$ .

#### Explaining

1. Any spanning set for  $V$  can be **reduced** to basis for  $V$  by removing the linearly dependent(redundant) vector in the spanning set, using redundancy theorem to get a linearly independent spanning set.
2. Linear independent set can be **enlarged** to a basis for  $V$ .

*proof.*

omitted.

■

**Corollary** Let  $V$  be a vector space and  $\dim(V) = n$ , any set of  $n$  linearly independent vectors from  $V$  is a basis for  $V$ .

*proof.* If  $n$  linearly independent vectors did not span  $V$ , then could be enlarged to a basis of  $V$  by pervious theorem, but then have a basis containing more than  $n$  vectors from  $V$ , which is impossible by the fundamental theorem since we given the  $\dim(V) = n$ , proven by contradiction.

**Example** Let  $V = P_2(\mathbb{R})$ ,  $p_1(x) = 2 - 5x$ ,  $p_2(x) = 2 - 5x + 4x^2$ , find  $p_3 \in P_2(\mathbb{R})$  s.t.  $\{p_1(x), p_2(x), p_3(x)\}$  is basis for  $P_2(\mathbb{R})$

**Note** Since  $\dim(P_2(\mathbb{R})) = 3$  so any 3 linearly independent vectors from  $P_2(\mathbb{R})$  will be a basis for  $P_2(\mathbb{R})$ .

**Solutions** e.g. constant function  $p_3(x) = 1$ , since  $1 \notin \text{span}\{p_1(x), p_2(x)\}$ , so  $\{p_1(x), p_2(x), p_3(x)\}$  is a basis of  $P_2(\mathbb{R})$ . e.g.  $p_3(x) = x$ , since  $x \notin \text{span}\{p_1(x), p_2(x)\}$

**Theorem** Let  $U$  and  $W$  be subspaces of vector space  $V$ , then we have

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$$

*proof.*

Let  $\{\vec{v}_i\}_1^k$  be basis for  $U \cap W$

$$\implies \dim(U \cap W) = k$$

Since  $\{\vec{v}_i\}_1^k$  is basis for  $U \cap W$  then it's a linearly independent subset of  $U$

So it could be enlarged to basis for  $U$ ,  $\{\vec{v}_1, \dots, \vec{v}_k, \vec{y}_1, \dots, \vec{y}_r\}$

$$\text{So } \dim(U) = k + r$$

We also could enlarge a basis for  $W$   $\{\vec{v}_1, \dots, \vec{v}_k, \vec{z}_1, \dots, \vec{z}_s\}$

$$\implies \dim(W) = k + s$$

WTS.  $\{\vec{v}_1, \dots, \vec{v}_k, \dots, \vec{y}_1, \dots, \vec{y}_r, \vec{z}_1, \dots, \vec{z}_s\}$  is a basis for  $U + W$

$$\begin{aligned} \text{(If we could show this)} \quad \dim(U + W) &= k + r + s = (k + r) + (k + s) - k \\ &= \dim(U) + \dim(W) - \dim(U \cap W) \end{aligned}$$

Obviously, the above set spans  $U + W$

WTS.  $\{\vec{v}_1, \dots, \vec{v}_k, \dots, \vec{y}_1, \dots, \vec{y}_r, \vec{z}_1, \dots, \vec{z}_s\}$  is linearly independent

Consider  $a_1\vec{v}_1 + \dots + a_k\vec{v}_k + b_1\vec{y}_1 + \dots + b_r\vec{y}_r + c_1\vec{z}_1 + \dots + c_s\vec{z}_s = \vec{0}$  (\*)

$$\text{From (*)} \implies \sum (c_i\vec{z}_i) = -\sum (a_i\vec{v}_i) - \sum b_i\vec{y}_i$$

$$\implies \sum (c_i\vec{z}_i) \in U \wedge \sum (c_i\vec{z}_i) \in W$$

$$\iff \sum (c_i\vec{z}_i) \in U \cap W$$

Since  $\{\vec{v}_i\}$  is a basis for  $U \cap W$

$$\implies \sum (c_i\vec{z}_i) = \sum (d_i\vec{v}_i)$$

$$\iff \sum (c_i\vec{z}_i) - \sum (d_i\vec{v}_i) = \vec{0} \in W$$

$$\implies c_i = d_i = 0 \text{ since } \{\vec{z}_i, \vec{v}_i\} \text{ is a basis}$$

Rewrite (\*)

$$\sum (a_i\vec{v}_i) + \sum b_i\vec{y}_i = \vec{0} \in U$$

$$\implies a_i = b_i = 0 \text{ since } \{\vec{v}_i, \vec{y}_i\} \text{ is a basis for } U$$

■

**Corollary** For direct sum, since the intersection is  $\{\vec{0}\}$

$$\dim(U \oplus W) = \dim(U) + \dim(W)$$

**Example** Let  $U, W$  are subspaces of  $\mathbb{R}^3$  such that  $\dim(U) = \dim(W) = 2$ , why is  $U \cap W \neq \{\vec{0}\}$

**Solutions** Geometrically,  $U$  and  $W$  are planes through origin then the intersection would be a line through origin ( $U \neq W$ ) or a plane through origin ( $U = W$ ), so shown.

**Question**  $V$  is a vector space,  $\dim(V) = n$ ,  $U \neq W$  are subspaces of  $V$  but  $\dim(U) = \dim(W) = (n - 1)$ , proof:

1.  $V = U + W$
2.  $\dim(U \cap W) = (n - 2)$

## 7 Lecture7 Jan.30. 2018

### 7.1 Linear Transformations

**Definition** Let  $V, W$  be vector spaces, a function  $T : V \rightarrow W$  is a **linear transformation**<sup>1</sup> if

1.  $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y}) \quad \forall \vec{x}, \vec{y} \in V$
2.  $T(c\vec{x}) = cT(\vec{x}) \quad \forall \vec{x} \in V, c \in \mathbb{R}$

*Linear transformation preserves vector additions and scalar multiplications on vector spaces.*

**Theorem(Alternative definition)** Transformation  $T : V \rightarrow W$  is linear if and only if

$$T(c\vec{x} + d\vec{y}) = cT(\vec{x}) + dT(\vec{y}), \quad \forall \vec{x}, \vec{y} \in V, c, d \in \mathbb{R}$$

*Linear transformations preserves linear combinations.*

**Example** (form 223) Rotation through angle  $\theta$  about the origin in  $\mathbb{R}^2$ .

<sup>1</sup>In some textbooks, this is annotated as **linear mapping**.

<sup>2</sup>Notice that the vector additions on the left and right sides of the equation are defined in different vector spaces, in  $V$  and  $W$  respectively.

<sup>3</sup>Notice that the scalar multiplication on the left and right sides of the equation are defined in different vector spaces, in  $V$  and  $W$  respectively.

**Example** (from 223) Matrix transformation, let  $A \in M_{m \times n}(\mathbb{R})$ , transformation  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  defined as

$$T(\vec{x}) = A\vec{x}$$

is linear.

**Example** Derivative  $T : P_n(\mathbb{R}) \rightarrow P_{n-1}(\mathbb{R})$  defined by

$$T(\vec{p}(x)) = \vec{p}'(x)$$

**Example** Matrix transpose  $T : M_{m \times n}(\mathbb{R}) \rightarrow M_{n \times m}(\mathbb{R})$  defined by

$$T(A) = A^T$$

## 7.2 Properties of linear transformations

**Property(i)** Linear transformation  $T : V \rightarrow W$  are uniquely defined by their values on any basis for  $V$ .

*proof.*

Let  $\{\vec{v}_1, \dots, \vec{v}_k\}$  be any basis for  $V$

Every vector  $\vec{x} \in V$  can be uniquely written as some linear combination of the  $\{\vec{v}_i\}_{i=1}^k$

$$\vec{x} = \sum_{i=1}^k c_i \vec{v}_i, \quad c_i \in \mathbb{R}, \quad \text{and } c_i \text{ are uniquely determined } \forall \vec{x} \in V$$

$$\implies T(\vec{x}) = T\left(\sum_{i=1}^k c_i \vec{v}_i\right)$$

$$= \sum_{i=1}^k c_i T(\vec{v}_i) \text{ since the transformation } T \text{ is linear.}$$

Since  $c_i$ s are uniquely determined by  $\{\vec{v}_i\}_{i=1}^k$

so the value of  $T(\vec{x})$  is uniquely determined by its value on basis vectors  $\{\vec{v}_i\}_{i=1}^k$ . ■

**Property(ii)** Let  $T : V \rightarrow W$  be a linear transformation, let  $A$  be a subspace of vector space  $V$ , then the **image**  $T(A)$  defined as

$$T(A) = \{T(\vec{x}) \mid \vec{x} \in A\}$$

called the image of  $A$  under linear transformation  $T$  is a subspace of  $W$ .  
*Linear transformation maps subspaces of  $V$  to subspaces of  $W$ .*



*proof.*

Since  $A$  is a subspace so it's non-empty, therefore  $\exists T(\vec{x}), \vec{x} \in A$

So  $T(A) \neq \emptyset$

Let  $\vec{w}_1, \vec{w}_2 \in T(A)$

$\implies \vec{w}_1 = T(\vec{x}_1), \vec{w}_2 = T(\vec{x}_2), \vec{x}_1, \vec{x}_2 \in A$

$\implies \vec{w}_1 + \vec{w}_2 = T(\vec{x}_1) + T(\vec{x}_2) = T(\vec{x}_1 + \vec{x}_2)$  since  $T$  is linear.

Since  $\vec{x}_1 + \vec{x}_2 \in A$  by the definition of subspaces.

$\implies \vec{w}_1 + \vec{w}_2 \in T(A)$

So  $T(A)$  is closed under vector addition.

Let  $\vec{w} \in T(A)$

$\implies \vec{w} = T(\vec{x}), \vec{x} \in A$

Let  $c \in \mathbb{R}$

Consider  $c\vec{w} = cT(\vec{x}) = T(c\vec{x})$

Since  $c\vec{x} \in A$

So  $c\vec{w} \in T(A)$

So  $T(A)$  is closed under scalar multiplication. ■

**Property(derived from the definition)** For all linear transformation  $T : V \rightarrow W$ , we have <sup>1</sup>

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

**Property(iii)** Let transformation  $T : V \rightarrow W$  be linear, let  $B$  be a subspace of  $W$ , then its **pre-image** defined as

$$T^{-1}(B) = \{\vec{x} \in V \mid T(x) \in B\}$$

is a subspace of  $V$ . <sup>2</sup>

---

<sup>1</sup>In the equation, clearly, the zero vector on the left side of the equation is in space  $V$  and the zero vector on the right side is in space  $W$ .

<sup>2</sup>The pre-image and inverse share the same notation, but in this case, transformation  $T$  is not necessarily invertible.

*proof.*

$$\begin{aligned}
 & \text{Let } \vec{w}_1, \vec{w}_2 \in T^{-1}(B) \\
 & \implies T(\vec{w}_1), T(\vec{w}_2) \in B \\
 & \implies aT(\vec{w}_1) + bT(\vec{w}_2) \in B, \forall a, b \in \mathbb{R} \text{ since } B \text{ is a subspace.} \\
 & \implies T(a\vec{w}_1 + b\vec{w}_2) \in B \\
 & \implies a\vec{w}_1 + b\vec{w}_2 \in T^{-1}(B)
 \end{aligned}$$

So  $T^{-1}(B)$  is closed under both vector addition and scalar multiplication,

So  $T^{-1}(B)$  is a subspace. ■

### 7.3 Definitions

Let  $T : V \rightarrow W$  to be a linear transformation,

**Definition** the **Image** of transformation  $T$  is defined as

$$Im(T) = T(V) = \{T(\vec{x}) \mid \vec{x} \in V\}$$

**Definition** the **Rank** of transformation  $T$  is defined as

$$Rank(T) = dim(Im(T))$$

**Definition** the **Kernel** of transformation  $T$  is defined as

$$Ker(T) = T^{-1}(\{\vec{0}\}) = \{\vec{x} \in V \mid T(\vec{x}) = \vec{0}\}$$

**Definition** the **Nullity** of transformation  $T$  is defined as

$$Nullity(T) = dim(ker(T))$$

**Example**  $T : P_3(\mathbb{R}) \rightarrow P_3(\mathbb{R})$  is linear defined by

$$T(\vec{p}(x)) = \vec{p}(2x + 1) - 8\vec{p}(x)$$

find  $Ker(T)$ .

**Theorem** Let  $T : V \rightarrow W$  be a linear transformation, let  $\{\vec{v}_1, \dots, \vec{v}_k\}$  be the spanning set of  $V$ <sup>1</sup>, then  $\{T(\vec{v}_1), \dots, T(\vec{v}_k)\}$  spans  $Im(T)$

*proof.*

Let  $\vec{w} \in Im(T)$

Since  $V = span\{\vec{v}_1, \dots, \vec{v}_k\}$

For any  $\vec{x} \in V$  can be written as

$$\vec{x} = \sum_{i=1}^k c_i \vec{v}_i, \quad c_i \in \mathbb{R}$$

$$\implies \vec{w} = T(\vec{x}) = T\left(\sum_{i=1}^k c_i \vec{v}_i\right)$$

$$= \sum_{i=1}^k c_i T(\vec{v}_i)$$

as a linear combination of  $\{T(\vec{v}_1), \dots, T(\vec{v}_k)\}$

So  $Im(T) = span\{T(\vec{v}_1), \dots, T(\vec{v}_k)\}$

■

## 8 Lecture8 Jan.31 2018

### 8.1 Linear Transformations

**Example**  $T : P_3(\mathbb{R}) \rightarrow P_3(\mathbb{R})$

$$T(p(x)) = p(2x + 1) - 8p(x)$$

Find the image of  $T$ .

We know  $B = \{1, x, x^2, x^3\}$  is the standard basis for  $P_3(\mathbb{R})$ , consider the set  $P(B)$

$$P(B) = \{-7, 1 - 6x, 1 + 4x - 4x^2, 1 + 6x + 12x^2\}$$

spans  $Im(T)$ . Notice the first three vectors in the set is linearly independent, the last vector is clearly dependent to the pervious three.<sup>2</sup> So by the redundancy theorem we could remove the last vector. There we have

$$Im(T) = span\{-7, 1 - 6x, 1 + 4x - 4x^2\}$$

<sup>1</sup>The set is only the spanning set of  $V$ , it's not necessarily to be a basis of  $V$ .

<sup>2</sup>Notice that the first three vectors is a basis of  $P_2(\mathbb{R})$ .

as basis.

In this example, the dimension of  $\text{Ker}(T)$  is 1 and the dimension of  $\text{Im}(T)$  is 3, and dimension of  $P_3(\mathbb{R})$  is 4. We have,  $\dim(P_3(\mathbb{R})) = \text{Nullity}(T) + \text{Rank}(T)$

**Theorem(Dimension Theorem)** Let  $T : V \rightarrow W$  be a linear transformation,

$$\dim(V) = \text{Nullity}(T) + \text{Rank}(T)$$

*Proof.*

Say  $\dim(V) = n$

Let  $\{\vec{v}_1, \dots, \vec{v}_k\}$  be a basis for  $\text{Ker}(T)$

Since  $\text{Ker}(T)$  is a subspace of  $V$ , the set  $\{\vec{v}_i\}_1^k$  is a subset of  $V$ ,

It can be extended to a basis  $\{\vec{v}_i\}_1^k \cup \{\vec{v}_i\}_{k+1}^n$  for  $V$ .

Claim:  $\{T(\vec{v}_{k+1}), \dots, T(\vec{v}_n)\}$  is basis for  $\text{Im}(T)$

If the claim is true, this prove the theorem since

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

$$\because T(\vec{v}_i) = \vec{0}, \forall i \in \mathbb{Z}_1^k$$

and by the definition of kernel of linear transformation,

$$\therefore \{T(\vec{v}_i)\}_{k+1}^n \text{ spans } \text{Im}(T)$$

$$\text{Show if } \sum_{i=k+1}^n c_i T(\vec{v}_i) = \vec{0} \implies c_i = 0$$

$$\implies T\left(\sum_{i=k+1}^n c_i \vec{v}_i\right) = \vec{0}$$

$$\implies \sum_{i=k+1}^n c_i \vec{v}_i \in \text{Ker}(T)$$

$$\implies \sum_{i=k+1}^n c_i \vec{v}_i = \sum_{i=1}^k c_i \vec{v}_i$$

$$\implies c_1 \vec{v}_1 + \dots + c_k \vec{v}_k - c_{k+1} \vec{v}_{k+1} - \dots - c_n \vec{v}_n = \vec{0}$$

Since  $\{\vec{v}_i\}_i^n$  is a basis for  $V$ .

$$\implies c_i = 0 \forall i$$

## 8.2 Applications of dimension theorem

**Definition** A linear transformation  $T : V \rightarrow W$  is called **injective**(one-to-one) if and only if

$$T(\vec{v}_1) = T(\vec{v}_2) \implies \vec{v}_1 = \vec{v}_2$$

**Definition** A linear transformation  $T : V \rightarrow W$  is called **surjective**(onto) if and only if

$$Im(T) = W$$

*Every vector in  $W$  has a pre-image in  $V$ .*

**Definition** A linear transformation  $T : V \rightarrow W$  is called **bijective** if it's both injective and surjective.

**Theorem** Let transformation  $T : V \rightarrow W$  is linear,  $T$  is injective if and only if  $dim(Ker(T)) = 0$ .

*Proof.*

Exercise

■

**Theorem**  $T$  is surjective if and only if  $dim(Im(T)) = dim(W)$ .

**Example**  $T : P_2(\mathbb{R}) \rightarrow \mathbb{R}^2$  defined by

$$T(p(x)) = \begin{pmatrix} p(1) \\ p(2) \end{pmatrix}$$

is  $T$  injective? surjective?

Not injective but surjective.

**Solution**

$$Ker(T) = span\{(x-1)(x-2)\}$$

So  $T$  has nullity of 1 and since  $dim(P_2(\mathbb{R})) = 3$ , by the dimension theorem we have  $Rank(T) = 2$  and since  $Im(T)$  is a subspace of  $\mathbb{R}^2$  which has dimension of 2, we could conclude that  $Im(T) = \mathbb{R}^2$ .

## 9 Lecture9 Feb.6 2018

### 9.1 Applications of dimension theorem

**Example**  $T : P_2(\mathbb{R}) \rightarrow \mathbb{R}^3$  defined by

$$T(p(x)) = (p(1), p(2), p(3))$$

Take  $p(x) = a + bx + cx^2 \in P_2(\mathbb{R})$ ,  $p(x) \in \text{Ker}(T)$  iff  $T(p(x)) = \vec{0}$ .

Let  $p(x) \in \text{Ker}(T)$ ,

Obviously the only solution for the system

$$\begin{cases} a + b + c = 0 \\ a + 2b + 4c = 0 \\ a + 3b + 9c = 0 \end{cases}$$

is  $a = b = c = 0$ , i.e.  $p = \vec{0} \in P_2(\mathbb{R})$ . So  $\dim(\text{Ker}(T)) = 0$ . Therefore,  $T$  is **injective**.

By *dimension theorem*,

$$\dim(P_2(\mathbb{R})) = 3 = 0 + \dim(\text{Im}(T)) \implies \dim(\text{Im}(T)) = 3 = \dim(\mathbb{R}^3)$$

therefore  $T$  is **surjective**. Therefore,  $T$  is **bijective**.

**Question**  $T : P_n(\mathbb{R}) \rightarrow P_n(\mathbb{R})$

$$T(p(x)) = xp'(x)$$

**Solution** Not injective because any *constant function* in  $P_n(\mathbb{R})$  is mapped to  $\vec{0} \in P_n(\mathbb{R})$ , therefore  $\text{Ker}(T) \neq \{\vec{0}\}$ . Also not surjective by the dimension theorem.

**Theorem** Let  $T : V \rightarrow W$  be an injective linear transformation, if  $\{\vec{v}_i\}_{i=1}^k$  is linearly independent in  $V$ , then the set  $\{T(\vec{v}_i)\}_{i=1}^k$  is linearly independent in  $W$ . *Injective transformation maps linearly independent set to linear independent set.*

*Proof.* Consider  $\sum_i c_i T(\vec{v}_i) = \vec{0}$ , then we have  $T(\sum_i c_i \vec{v}_i) = \vec{0}$ , which implies  $\sum_i c_i \vec{v}_i \in \text{Ker}(T)$ . By definition of injective transformation,  $\sum_i c_i \vec{v}_i = \vec{0}$ . Since  $\{\vec{v}_i\}_{i=1}^k$  is linearly independent, so  $c_i = 0$ ,  $\forall i$ . Therefore  $\{T(\vec{v}_i)\}_{i=1}^k$  is linearly independent. ■

**Theorem** Let  $T : V \rightarrow W$  be a linearly transformation,  $\{\vec{v}_i\}_{i=1}^n$  is a basis for  $V$ , then if  $\{T(\vec{v}_i)\}_{i=1}^n$  is linear independent, then  $T$  is injective.  
*A criteria for  $T$  to be injective based on image of a basis.*

*Proof.*

Let  $\{\vec{v}_i\}_{i=1}^n$  be a basis of  $V$

Consider  $T(\vec{x}) = \vec{0}$

Since  $\{\vec{v}_i\}_{i=1}^n$  is a basis

Let  $\vec{x} = \sum c_i \vec{v}_i$

Assume  $\vec{x} \in \text{Ker}(T)$

$$T(\vec{x}) = \vec{0} \iff T\left(\sum c_i \vec{v}_i\right) = \vec{0}$$

$$\implies \sum c_i T(\vec{v}_i) = \vec{0}$$

Since  $\{T(\vec{v}_i)\}_{i=1}^n$  are linearly independent.

$$\implies c_i = 0$$

$$\text{Therefore } \vec{x} = \sum 0 \vec{v}_i = \vec{0}$$

$$\text{Therefore } \text{Ker}(T) = \{\vec{0}\}$$

$$\text{Therefore } \dim(\text{Ker}(T)) = 0$$

$$\implies \text{injective}$$

■

**Theorem** Let  $T : V \rightarrow W$  be a linear transformation,<sup>1</sup>

1. If  $\dim(V) > \dim(W)$ , then  $T$  cannot be injective.
2. If  $\dim(V) < \dim(W)$ , then  $T$  cannot be surjective.

---

<sup>1</sup>Consider the contrapositive predicates of this theorem.

**Lemma** For a linear transformation between spaces with different dimensions, it could not be bijective.

*Proof.*

$$\begin{aligned}
 \dim(V) &= \dim(\text{Ker}(T)) + \dim(\text{Im}(T)) \\
 &\because \dim(\text{Im}(T)) \leq \dim(W) \\
 \therefore \dim(V) &\leq \dim(\text{Ker}(T)) + \dim(W) \\
 \implies \dim(\text{Ker}(T)) &\geq \dim(V) - \dim(W) \\
 &\implies \dim(\text{Ker}(T)) > 0
 \end{aligned}$$

So  $T$  could not be injective

$$\begin{aligned}
 \dim(V) &= \dim(\text{Ker}(T)) + \dim(\text{Im}(T)) \\
 &\because \dim(\text{Ker}(T)) \geq 0 \\
 \therefore \dim(V) &\geq \dim(\text{Im}(T)) \\
 \implies \dim(\text{Im}(T)) &< \dim(W)
 \end{aligned}$$

So  $T$  could not be surjective

■

*Proof 2.*

Consider a transformation  $T : V \rightarrow W$  is bijective.

By the contrapositive form of above theorem,

$$\text{Injective} \implies \dim(V) \leq \dim(W)$$

$$\text{Surjective} \implies \dim(V) \geq \dim(W)$$

Therefore bijective

$$\implies \dim(V) \leq \dim(W) \wedge \dim(V) \geq \dim(W) \iff \dim(V) = \dim(W)$$

$$\text{Therefore bijective} \implies \dim(V) = \dim(W)$$

So, take contrapositive,  $\dim(V) \neq \dim(W) \implies$  not bijective.

■



**Theorem** (*Half is good enough*) Let  $T : V \rightarrow W$  is linear, and  $\dim(V) = \dim(W)$ . Then  $T$  is injective if and only if surjective.

*Proof.*

By dimension theorem

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

If injective  $\dim(\text{Ker}(T)) = 0$

$$\implies \dim(\text{Im}(T)) = \dim(W)$$

So surjective

If surjective  $\dim(\text{Im}(T)) = \dim(W) = \dim(V)$

$$\implies \dim(\text{Ker}(T)) = 0$$

So injective

■

## 9.2 Isomorphisms

**Definition** If  $T : V \rightarrow W$  is bijjective, we call  $T$  an **isomorphism**. If there exists an isomorphism  $T : V \rightarrow W$  say  $V$  and  $W$  are **isomorphic** vector spaces.

**Theorem**  $V, W$  are isomorphic iff  $\dim(V) = \dim(W)$ .

*Proof.*

$$\rightarrow V, W \text{ isomorphic} \implies \dim(V) = \dim(W)$$

Isomorphic means there exists a bijective transformation  $T$

$$\begin{aligned} \text{By dimension theorem } \dim(V) &= \dim(\text{Ker}(T)) + \dim(\text{Im}(T)) \\ &= 0 + \dim(W) \end{aligned}$$

$$\leftarrow \dim(V) = \dim(W) \implies V, W \text{ isomorphic}$$

Equivalently, find a isomorphism(bijective) transformation

Let  $\{\vec{v}_i\}_{i=1}^n$  be basis for  $V$

Let  $\{\vec{w}_i\}_{i=1}^n$  be basis for  $W$

Claim  $T : V \rightarrow W$  defined by

$T(\vec{v}_i) = \vec{w}_i$  is an isomorphism.

If  $\vec{x} \in \text{Ker}(T) \subseteq V$

$$\vec{x} = \sum c_i \vec{v}_i$$

$$\vec{0} = T(\vec{x})$$

$$= \sum c_i T(\vec{v}_i)$$

$$= \sum (c_i \vec{w}_i)$$

$$\implies c_i = 0 \text{ since } \vec{w}_i \text{ are basis.}$$

$$\implies \vec{x} = \vec{0}$$

$$\implies \dim(\text{Ker}(T)) = 0$$

$$\implies \text{injective} \iff \text{surjective}$$

Therefore  $V$  and  $W$  are isomorphic vector spaces. ■

**Note** if  $T : V \rightarrow W$  is an isomorphism, then  $T$  maps a basis for  $V$  to a basis for  $W$ .

**Example**  $T : P_2(\mathbb{R}) \rightarrow \mathbb{R}^3$ ,

$$T(p(x)) = (p(1), p(2), p(3))$$

is an isomorphism. And  $P_2(\mathbb{R})$  and  $\mathbb{R}^3$  are isomorphic.

**Example**  $T : P_2(\mathbb{R}) \rightarrow \mathbb{R}^3$ ,

$$T(1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad T(x) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad T(x^2) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

is an isomorphism.

**Example**  $M_{2 \times 2}(\mathbb{R})$ ,  $P_3(\mathbb{R})$  and  $\mathbb{R}^4$  are isomorphic.

**Theorem** Any  $n$ -dim vector space  $V$  is isomorphic to  $\mathbb{R}^n$ . What is an isomorphism  $T : V \rightarrow \mathbb{R}^n$

*Procedure:*

Let  $\{\vec{v}_i\}_{i=1}^n$  be any basis for  $V$

We know that  $\forall \vec{x} \in V$ ,

By property of basis,

$$\vec{x} = \sum c_i \vec{v}_i$$

Then transformation  $T$  defined by

$$T(\vec{x}) = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n \text{ is an isomorphism.}$$

### 9.3 Coordinates

**Definition** Let  $V$  be a vector space,  $\alpha = \{\vec{v}_1, \dots, \vec{v}_n\}$  be any basis for  $V$ ,  $\forall \vec{x} \in V$  can be written uniquely as

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

then  $c_1, \dots, c_n$  is called the **coordinates** for  $\vec{x}$  relative to basis  $\alpha$ , with notation

$$[\vec{x}]_\alpha = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \iff \vec{x} = \sum c_i \vec{v}_i$$

**Claim**  $[\vec{x} + c\vec{y}]_\alpha = [\vec{x}]_\alpha + c[\vec{y}]_\alpha \quad \forall \vec{x}, \vec{y} \in V, c \in \mathbb{R}.$

**Remark** if  $\alpha, \alpha'$  are any two bases for  $V$  then *generally*  $[\vec{x}]_\alpha \neq [\vec{x}]_{\alpha'}$  (except  $\vec{0}$ )<sup>1</sup>.

## 10 Lecture10 Feb.7 2018

### 10.1 Matrix of linear transformation

**Recall** Let  $V$  be a vector space, let  $\alpha$  be any basis for  $V$ .

$$\forall \vec{x} \in V, x = \sum c_i \vec{v}_i$$

$$[\vec{x}]_\alpha = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

So transformation  $\vec{x} \rightarrow [\vec{x}]_\alpha$  is an isomorphism that  $V \rightarrow \mathbb{R}^n$ .

**Definition** Let  $W$  be a vector space and let  $\beta = \{\vec{w}_i\}_{i=1}^m$  be any basis of  $W$ , let  $T : V \rightarrow W$  be a linear operator.

$$T(\vec{x}) = \sum c_i T(\vec{v}_i)$$

So that

$$\begin{aligned} [T(\vec{x})]_\beta &= [\sum c_i T(\vec{v}_i)]_\beta = \sum c_i [T(\vec{v}_i)]_\beta \\ &= \begin{bmatrix} [T(\vec{v}_1)]_\beta & \dots & [T(\vec{v}_n)]_\beta \end{bmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \end{aligned}$$

$\begin{bmatrix} [T(\vec{v}_1)]_\beta & \dots & [T(\vec{v}_n)]_\beta \end{bmatrix}$  is called the the **matrix of  $T$**  w.r.t. bases  $\alpha, \beta$ . Denoted as  $[T]_\alpha^\beta$ , and by definition we have

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

---

<sup>1</sup> $[\vec{0}]_\beta = \vec{0} \in \mathbb{R}^n, \forall \beta$  as basis for  $V$ .

**Example**  $T : P_2(\mathbb{R}) \rightarrow P_3(\mathbb{R})$

$$T(p(x)) = xp(x)$$

$$\alpha = \{1 - x, 1 - x^2, x\}, \beta = \{1, 1 + x, 1 + x + x^2, 1 - x^3\}$$

Find  $[T]_{\alpha}^{\beta}$ .

$$T(1 - x) = x(1 - x) = x - x^2$$

$$x - x^2 = (-1)(1) + 2(1 + x) + (-1)(1 + x + x^2) + 0(1 - x^3)$$

$$[T(1 - x)]_{\beta} = (-1, 2, -1, 0)$$

$$T(1 - x^2) = x - x^3$$

$$[T(1 - x^2)]_{\beta} = (-2, 1, 0, 1)$$

$$[T(x)]_{\beta} = x^2$$

$$[T(x)]_{\beta} = (0, -1, 1, 0)$$

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

**Picture** Let  $V, W$  be two vectors spaces,  $\alpha = \{\vec{v}_1, \dots, \vec{v}_n\}$  is a basis for  $V$  and  $\beta = \{\vec{w}_1, \dots, \vec{w}_m\}$  is a basis for  $W$ .

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \downarrow [\ ]_{\alpha} & & \downarrow [\ ]_{\beta} \\ \mathbb{R}^n & \xrightarrow{[T]_{\alpha}^{\beta}} & \mathbb{R}^m \end{array}$$

**Remark**

1.  $\vec{x} \in \text{Ker}(T) \iff T(\vec{x}) = \vec{0} \iff [T(x)]_{\beta} = [\vec{0}]_{\beta} \in \mathbb{R}^m \iff [T]_{\alpha}^{\beta}[\vec{x}]_{\alpha} = 0 \iff [\vec{x}]_{\alpha} \in \text{Ker}([T]_{\alpha}^{\beta})$
2.  $\vec{w} \in \text{Im}(T) \iff [\vec{w}]_{\beta} \in \text{Col}([T]_{\alpha}^{\beta})$

**Theorem Rank nullity for transformation matrix** Let  $T : V \rightarrow W$  be a linear operator and  $\dim(V) = n$ , then

$$\dim(\text{Ker}([T]_{\alpha}^{\beta})) + \dim(\text{Col}([T]_{\alpha}^{\beta})) = n$$

**Example**  $T : P_2(\mathbb{R}) \rightarrow M_{2 \times 2}(\mathbb{R})$

$$T(a + bx + c^2) = \begin{bmatrix} c & -c \\ a - c & a + c \end{bmatrix}$$

And given bases  $\alpha = \{x^2 - x, x - 1, x^2 + 1\}$  and  $\beta = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \right\}$

**Solution**

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

$$\text{Nul}([T]_{\alpha}^{\beta}) = \text{span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

$$\text{Nul}(T) = \text{span}\{2x\}$$

$$\text{Col}([T]_{\alpha}^{\beta}) = \text{span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} \right\}$$

$$\text{Col}(T) = \text{span}\left\{ \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ -1 & -1 \end{bmatrix} \right\}$$

## 11 lecture11 Feb.13 2018

### 11.1 Algebra of Transformation

**Theorem** Let  $T : V \rightarrow W$  be a linear transformation, where  $\alpha = \{\vec{v}_1, \dots, \vec{v}_n\}$  and  $\beta = \{\vec{w}_1, \dots, \vec{w}_m\}$  are bases for  $V, W$  respectively.

$$\vec{x} \in \text{Ker}(T) \iff [\vec{x}]_{\alpha} \in \text{Ker}([T]_{\alpha}^{\beta})$$

$$\vec{x} \in \text{Im}(T) \iff [\vec{x}]_{\beta} \in \text{Col}([T]_{\alpha}^{\beta})$$

**Definition**  $T_1, T_2 : V \rightarrow W$  are linear transformations, define **addition** and **scalar multiplication** of transformation as

$$(T_1 + T_2)(\vec{x}) = T_1(\vec{x}) + T_2(\vec{x}) \quad \forall \vec{x} \in V$$

$$(cT_1)(\vec{x}) = c(T_1(\vec{x})) \quad \forall \vec{x} \in V, \quad c \in \mathbb{R}$$

**Theorem** And, let  $\alpha$  and  $\beta$  be bases for  $V, W$  respectively, then,

$$[T_1]_{\alpha}^{\beta} + [T_2]_{\alpha}^{\beta} = [T_1 + T_2]_{\alpha}^{\beta}$$

$$c[T_1]_{\alpha}^{\beta} = [cT_1]_{\alpha}^{\beta}$$

**Definition** Let  $T : V \rightarrow W$  and  $S : W \rightarrow U$  be two linear transformations, then the **composition**  $ST : V \rightarrow U$  is defined as

$$(ST)(\vec{x}) = S(T(\vec{x})) \quad \forall \vec{x} \in V$$

**Remark** If  $S, T$  are *linear* then the composition  $ST$  is also *linear*.

*Proof.*

$$\text{Let } a, b \in \mathbb{R}, \quad \vec{x}, \vec{y} \in V$$

$$\begin{aligned} & ST(a\vec{x} + b\vec{y}) \\ &= S(T(a\vec{x} + b\vec{y})) \\ &= S(aT(\vec{x}) + bT(\vec{y})) \\ &= a(ST(\vec{x})) + b(ST(\vec{y})) \end{aligned}$$

■

## 11.2 Matrix of composition of transformations

Consider  $T : V \rightarrow W$  and  $S : W \rightarrow U$  as linear transformations, let  $\alpha, \beta, \gamma$  be bases of  $V, W, U$  respectively.

We know how to compute  $[T]_\alpha^\beta$  and  $[S]_\beta^\gamma$ . Now want to find  $[ST]_\alpha^\gamma$ .

$$\begin{aligned}
 \forall \vec{x} \in V, [ST]_\alpha^\gamma [\vec{x}]_\alpha &= [(ST)(\vec{x})]_\gamma \\
 &= [S(T(\vec{x}))]_\gamma \\
 &= [S]_\beta^\gamma [T(\vec{x})]_\beta \\
 &= [S]_\beta^\gamma [T]_\alpha^\beta [\vec{x}]_\alpha
 \end{aligned}$$

This holds true for all  $\vec{x} \in V$

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

**Conclusion** the matrix of  $ST$  = matrix of  $S \times$  matrix of  $T$ .

### 11.3 Inverse transformations

**Theorem**  $T : V \rightarrow W$  is *isomorphism*<sup>1</sup> if and only if there exists function  $S : W \rightarrow V$  such that

$$(ST)(\vec{v}) = \vec{v} \quad \forall \vec{v} \in V \quad \wedge \quad (TS)(\vec{w}) = \vec{w} \quad \forall \vec{w} \in W$$

**Definition** And the above-mentioned linear operator  $S$  is called the **inverse** of  $T$ , written as  $T^{-1}$ .

*proof.*( $\rightarrow$ )  $T$  is an isomorphism means every vector in  $W$  has an unique pre-image in  $V$  the function  $S : W \rightarrow V$  maps *every* vector in  $W$  to its *unique* pre-image in  $V$ , so  $S$  is the inverse of  $T$ .

*proof.*( $\leftarrow$ ) Assume  $S : W \rightarrow V$  is the inverse of  $T : V \rightarrow W$  then  $T(S(\vec{y})) = \vec{y}$ ,  $\forall \vec{y} \in V$ , this means  $T$  is surjective since every  $\vec{y} \in W$  has pre-image under  $T$ , which is  $S(\vec{y}) \in V$ . Now suppose  $T(\vec{x}_1) = T(\vec{x}_2)$ , apply transformation  $S$  on both sides of the equation,  $S(T(\vec{x}_1)) = S(T(\vec{x}_2))$  we have  $\vec{x}_1 = \vec{x}_2$ . This implies the transformation is injective. Therefore, transformation  $T$  is bijective, that's isomorphism. ■

<sup>1</sup>Recall that isomorphism is equivalent to bijective.



**Note**  $T^{-1}(\vec{y})$  is the unique vector  $\vec{x}$ , s.t.  $T(\vec{x}) = \vec{y}$ . That's

$$T(\vec{x}) = \vec{y} \iff T^{-1}(\vec{y}) = \vec{x}$$

**Theorem** If  $T : V \rightarrow W$  is an isomorphism then the inverse of  $T$ ,  $T^{-1} : W \rightarrow V$  is linear.<sup>1</sup>

*Proof.*

$$\text{WTS } T^{-1}(a\vec{w}_1 + b\vec{w}_2) = aT^{-1}(\vec{w}_1) + bT^{-1}(\vec{w}_2) \forall a, b \in \mathbb{R}, \forall \vec{w}_1, \vec{w}_2 \in W$$

$$T^{-1}(\vec{w}_1) \text{ is the unique } \vec{x}_1 \text{ s.t. } T(\vec{x}_1) = \vec{w}_1$$

$$T^{-1}(\vec{w}_2) \text{ is the unique } \vec{x}_2 \text{ s.t. } T(\vec{x}_2) = \vec{w}_2$$

$$T^{-1}(a\vec{w}_1 + b\vec{w}_2) \text{ is the unique } \vec{x} \text{ s.t. } T(\vec{x}) = a\vec{w}_1 + b\vec{w}_2$$

$$\because T(\vec{x}) = a\vec{w}_1 + b\vec{w}_2$$

$$= aT(\vec{x}_1) + bT(\vec{x}_2)$$

$$= T(a\vec{x}_1 + b\vec{x}_2)$$

$$\therefore \vec{x} = a\vec{x}_1 + b\vec{x}_2$$

$$\text{Also } T(\vec{x}) = a\vec{w}_1 + b\vec{w}_2$$

$$\therefore \vec{x} = T^{-1}(a\vec{w}_1 + b\vec{w}_2) = a\vec{x}_1 + b\vec{x}_2$$

$$= aT^{-1}(\vec{w}_1) + bT^{-1}(\vec{w}_2)$$

■

**Theorem**  $T : V \rightarrow W$  is isomorphism, then let  $\alpha$  and  $\beta$  are bases of  $V$  and  $W$  representing then  $[T]_{\alpha}^{\beta}$  is invertible, and

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

*Proof.*

omitted

## 11.4 Change of basis

*What's the effect of a change of basis on coordinate of a vector and matrix of transformation.*

<sup>1</sup>Note: the conclusion could be changed into isomorphism.

**Theorem** Let  $\alpha$  and  $\alpha'$  be two bases of  $V$ , and  $\vec{x} \in V$ , then

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

*Proof.*

Let  $\vec{x} \in V$

$$I(\vec{x}) = \vec{x}$$

$$[I(\vec{x})]_{\alpha'} = [\vec{x}]_{\alpha'}$$

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

■

**Definition** The above-mentioned  $[I]_{\alpha}^{\alpha'}$  is called the **change of basis matrix** from  $\alpha$  to  $\alpha'$ .

**Computation** Let  $\alpha = \{\vec{a}_1, \dots, \vec{a}_n\}$ , then<sup>1</sup>

$$[I]_{\alpha}^{\alpha'} = [[\vec{a}_1]_{\alpha'} \mid \dots \mid [\vec{a}_n]_{\alpha'}]$$

## 12 Lecture12 Feb.14 2018

**Recall** Let  $\alpha$  and  $\beta$  be bases for  $V$  and  $I : V \rightarrow V$  is the identity transformation, then

$$[I]_{\alpha}^{\beta} [\vec{x}]_{\alpha} = [\vec{x}]_{\beta}$$

Also,

$$[I]_{\beta}^{\alpha} [\vec{x}]_{\beta} = [\vec{x}]_{\alpha}$$

**Example** Let  $\alpha = \{x^2, 1+x, x+x^2\}$  and  $\beta$  be a basis for  $P_2(\mathbb{R})$  and

$$[I]_{\alpha}^{\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \text{ and } [p(\vec{x})]_{\beta} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

find the basis  $\beta$ .

**Solution** omitted

---

<sup>1</sup>Construct column by column.

**Theorem** Suppose  $T : V \rightarrow W$  is linear,  $\alpha$  and  $\alpha'$  are any two bases for  $V$  and  $\beta$  and  $\beta'$  are any two bases of  $W$ , then,

$$[T]_{\alpha'}^{\beta'} = [I]_{\beta}^{\beta'} [T]_{\alpha}^{\beta} [I]_{\alpha'}^{\alpha}$$

*Proof.*

**Recall**  $T = ITI$

Consider let  $\vec{x} \in V$

$$\begin{aligned} [I]_{\beta}^{\beta'} [T]_{\alpha}^{\beta} [I]_{\alpha'}^{\alpha} [\vec{x}]_{\alpha'} &= [I]_{\beta}^{\beta'} [T]_{\alpha}^{\beta} [\vec{x}]_{\alpha} \\ &= [I]_{\beta}^{\beta'} [T(\vec{x})]_{\beta} \\ &= [T(\vec{x})]_{\beta'} \\ &= [T]_{\beta'}^{\alpha'} [\vec{x}]_{\alpha'} \\ \implies [T]_{\beta'}^{\alpha'} &= [I]_{\beta}^{\beta'} [T]_{\alpha}^{\beta} [I]_{\alpha'}^{\alpha} \end{aligned}$$

■

Also,

$$[T]_{\alpha}^{\beta} = [I]_{\beta'}^{\beta} [T]_{\alpha'}^{\beta'} [I]_{\alpha}^{\alpha'}$$

**Special Case** Consider when  $V = W$ ,  $\alpha = \beta$  and  $\alpha' = \beta'$ . we have

$$[T]_{\alpha'}^{\alpha'} = [I]_{\alpha}^{\alpha'} [T]_{\alpha}^{\alpha} [I]_{\alpha'}^{\alpha}$$

where

$$([I]_{\alpha}^{\alpha'})^{-1} = [I]_{\alpha'}^{\alpha}$$

the equation becomes

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

and can be written in the form of

$$B = P^{-1}AP$$

**Definition** Two matrices  $A$  and  $B$  are **similar** if there exists an invertible matrix  $P$  s.t.

$$B = P^{-1}AP$$

**Interpretation** <sup>1</sup> Linear operators  $A$  and  $B$  are **similar** if and only if  $A$  and  $B$  representing the same transformation relative to different bases and  $P$  is the change of basis matrix.

## 13 Lecture13 Feb.27 2018

### 13.1 Diagonalization

**Definition** Consider a linear operator  $T : V \rightarrow V$  is **diagonalizable** if and only  $\exists$  a basis  $\beta$  for  $V$  s.t.

$$[T]_{\beta}^{\beta}$$

is diagonal.

**Note** Let  $\beta = \{\vec{v}_1, \dots, \vec{v}_n\}$  be a basis,  $T : V \rightarrow V$  is diagonalizable if and only if  $[T]_{\beta}^{\beta}$  is in form 
$$\begin{bmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_n \end{bmatrix}$$

**Definition**  $T : V \rightarrow V$  is a linear operator on  $V$ , a *non-zero* vector  $\vec{x} \in V$  is an **eigenvector** of  $T$  if and only if  $T(\vec{x}) = \lambda\vec{x}$  for some  $\lambda \in \mathbb{R}$ .  $\lambda$  is called the **eigenvalue** of  $T$  corresponding to vector  $\vec{x}$ .

**Theorem** Linear operator  $T : V \rightarrow V$  is *diagonalizable* if and only exist a basis of  $V$  consisting of *eigenvectors* of  $T$ . If  $T$  is diagonalizable, the diagonal entries of  $[T]_{\beta}^{\beta}$  are corresponding eigenvalues of  $T$ , in the same order.

### 13.2 How to find eigenvalues and eigenvectors of $T$

**Definition** The **determinant** of  $T$  is defined as  $\det([T]_{\alpha}^{\alpha})$  for any basis  $\alpha$  for  $V$ .

**Remark** The determinant of linear operator  $T$  does not depends on the choice of basis of  $\alpha$  for  $V$ , *since similar matrices have the same determinant*.

---

<sup>1</sup>Could be used as alternative definition for similarity between matrices.

**Theorem**  $\lambda \in \mathbb{R}$  is an eigenvalue of  $T$  if and only if

$$\det(T - \lambda I) = 0$$

*Proof.*

$$\begin{aligned}
 & \text{Let } \lambda \text{ be an eigenvalue of } T, \\
 & \text{Let } \alpha \text{ be any basis for } V, \\
 \iff & \exists \vec{x} \in V, \vec{x} \neq \vec{0}, \text{ s.t. } T(\vec{x}) = \lambda \vec{x} \\
 \iff & T(\vec{x}) - \lambda \vec{x} = \vec{0} \\
 \iff & (T - \lambda I)(\vec{x}) = \vec{0} \\
 \iff & \vec{x} \in \text{Ker}(T - \lambda I) \\
 \therefore & \text{Ker}(T - \lambda I) \neq \{\vec{0}\} \\
 \iff & (T - \lambda I) \text{ is not injective} \\
 \iff & [T - \lambda I]_{\alpha}^{\alpha} \text{ is not injective and not invertible} \\
 \iff & \det([T - \lambda I]_{\alpha}^{\alpha}) = \det(T - \lambda I) = 0
 \end{aligned}$$

■

**Definition**  $\det(T - \lambda I) = 0$  is called the **characteristic polynomial** of  $T$ , written as  $P_T(\lambda) := \det(T - \lambda I)$ , the degree of  $P_T(\lambda)$  is the dimension of  $V$ .

**Note**  $\lambda$  is an eigenvalue  $\iff \lambda$  is a root of  $P_T(\lambda)$ .

**Theorem**  $T : V \rightarrow V$  is a linear operator and  $\lambda$  is an eigenvalue of  $T$ ,  $\vec{x}$  is an eigenvector of  $T$  corresponding to eigenvalue  $\lambda$ , if and only if

$$\vec{x} \neq \vec{0} \wedge \vec{x} \in \text{Ker}(T - \lambda I)$$

*Proof.*

By definition

■

**Definition**  $\text{Ker}(T - \lambda I)$  is called the **eigenspace** of  $T$  corresponding to eigenvalue  $\lambda$ , noted as  $E_{\lambda}(T)$ , and it is a subspace of  $V$ .

**Note** To find eigenvalues and eigenvectors of  $T : V \rightarrow V$ , choose any basis  $\beta$  for  $V$ ,  $\vec{x}$  is an eigenvector with corresponding eigenvalue  $\lambda$  if and only if  $[\vec{x}]_\beta$  is an eigenvector of  $[T]_\beta^\beta$  with corresponding eigenvalue  $\lambda$ .  
That's

$$\begin{aligned} T(\vec{x}) &= \lambda \vec{x} \\ \implies [T(\vec{x})]_\beta &= [\lambda \vec{x}]_\beta \\ \iff [T]_\beta^\beta [\vec{x}]_\beta &= \lambda [\vec{x}]_\beta \end{aligned}$$

**Note** Consider diagonalization in MAT223,

$$D = P^{-1}AP$$

Let  $D$  and  $A$  representing the same linear operator  $[T]_V^V$  and let  $\beta$  be a basis of  $V$  consisting of eigenvectors of  $T$  and  $\alpha$  is another basis of  $V$ . Then, the above equation is

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

## 14 Lecture14 Feb.28 2018

**Theorem** Suppose  $\lambda_0$  is an eigenvalue of linear operator  $T : V \rightarrow V$ , let  $\dim(E_{\lambda_0}) = k$ , then  $(\lambda - \lambda_0)^k$  divides  $P_T(\lambda)$

*Proof.*

Let  $\{\vec{v}_1, \dots, \vec{v}_k\}$  be basis for  $E_{\lambda_0}$

Since  $E_{\lambda_0} \subset V$

Let  $\dim(V) = n$

Extend basis of  $E_{\lambda_0}$  to basis of  $V$ .

$\alpha = \{\vec{v}_1, \dots, \vec{v}_k\} \cup \{\vec{v}_{k+1}, \dots, \vec{v}_n\}$

Since  $\vec{v}_i \in E_{\lambda_0}$ ,

Therefore  $T(\vec{v}_i) = \lambda_0 \vec{v}_i$

$$[T]_{\alpha}^{\alpha} = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$

Where  $A = \text{diag}(\lambda_0, \dots, \lambda_0) \in \mathbb{M}_{k \times k}(\mathbb{R})$

And  $B \in \mathbb{M}_{k \times n-k}(\mathbb{R}), D \in \mathbb{M}_{n-k \times n-k}(\mathbb{R})$

$$P_T(\lambda) = \det(A - \lambda I) * \det(D - \lambda I)$$

$$= (\lambda_0 - \lambda)^k * \det(D - \lambda I)$$

$$\text{Therefore } (\lambda - \lambda_0)^k \mid P_T(\lambda)$$

■

**Definition** The **multiplicity** of eigenvalue  $\lambda_0$  is the number of times  $(\lambda - \lambda_0)$  appears as a factor in  $P_T(\lambda)$ .

**Note** If eigenvalue  $\lambda$  has multiplicity  $m$ , the above theorem says

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

if  $m = 1$ , then  $\dim(E_{\lambda}) = 1$ .

**Theorem** If  $\lambda_1, \dots, \lambda_k$  are *distinct* eigenvalues of  $T : V \rightarrow V$  and  $\alpha = \{\vec{x}_1, \dots, \vec{x}_k\}$  are corresponding eigenvectors, then the set  $\alpha$  is *linearly inde-*

pendent.

*Proof.*

Exercise



(★)**Theorem** Sufficient condition for diagonalizability Let  $\lambda_1, \dots, \lambda_k$  be distinct eigenvalues of  $T$ , suppose the characteristic polynomial is in form

$$P_T(\lambda) = \prod_{i=1}^k (\lambda - \lambda_i)^{m_i}$$

and  $T$  is diagonalizable if and only if

$$\dim(E_{\lambda_i}) = m_i, \quad \forall i$$

**Note** Also,  $\sum_{i=1}^k m_i = \dim(V) = n$



*Proof.*

←

Assume  $\dim(E_{\lambda_i}) = m_i \forall i$

Consider  $E_{\lambda_i}$

Take basis for  $E_{\lambda_i}$ , note as  $\{\vec{v}_1^i, \dots, \vec{v}_{m_i}^i\}$

Claim: the union of bases of  $E_{\lambda_i} \forall i$  gives a basis consisting of eigenvectors of  $T$ .

$$\text{Note } |\cup_{i=1}^k \{\vec{v}_1^i, \dots, \vec{v}_{m_i}^i\}| = \sum_{i=1}^k m_i = \dim(V)$$

All we need to show is linear independence.

$$\text{Consider } \sum_{i=1}^k \sum_{j=1}^{m_i} c_{ij} \vec{v}_j^i = \vec{0}(\star)$$

$$\text{Consider } \sum_{j=1}^{m_i} c_{ij} \vec{v}_j^i \in E_{\lambda_i} = \vec{x}_i$$

$$\text{So } (\star) \text{ becomes } \sum_{i=1}^k \vec{x}_i = \vec{0} \text{ where } \vec{x}_i \in E_{\lambda_i}, \forall i$$

Since  $\vec{x}_i$  is eigenvectors for  $T$ , corresponding to different eigenvalues,

Therefore,  $\{\vec{x}_{i_1}, \dots, \vec{x}_{i_k}\}$  is linearly independent

$$\text{So } \vec{x}_i = \vec{0} \forall i$$

$$\text{That's } \sum_{j=1}^{m_i} c_{ij} \vec{v}_j^i = \vec{x} = \vec{0} \forall i$$

$$\implies c_{ij} = 0 \forall i, j$$

Therefore linearly independent, so exists basis for  $V$  consisting of eigenvectors,

Therefore  $T$  is diagonalizable.

→

Suppose  $T$  is diagonalizable,

Since  $T$  is diagonalizable, then exists basis for  $V$  consisting of eigenvectors, say  $\alpha$

$$\text{Consider } [T]_{\alpha}^{\alpha} = \begin{bmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \lambda_1 & \dots & \dots & 0 \\ 0 & \dots & \lambda_2 & \ddots & 0 \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

Where  $\lambda_1$  takes first  $m_1$  rows,  $\lambda_2$  takes the next  $m_2$  rows, etc.

$$P_T(\lambda) = \det([T]_{\alpha}^{\alpha} - \lambda I)$$

$$= \prod_{i=1}^k (\lambda_i - \lambda)^{m_i}$$

$$\text{Since } 1 \leq \dim(E_{\lambda_i}) \leq m_i \quad \forall i$$

$$\implies \dim(E_{\lambda_i}) = m_i \quad \forall i$$

■

## 15 Lecture15 Mar.6 2018

### 15.1 Fields

**Definition** A **field** is a set  $F$  together with two operations, *addition* and *multiplication* that satisfies the following properties.

1.  $\forall x, y \in F, x + y = y + x$
2.  $\forall x, y, z \in F, (x + y) + z = x + (y + z)$
3. *Additive identity*  $\exists 0 \in F, \text{ s.t. } \forall x \in F, 0 + x = x$
4. *Additive inverse*  $\forall x \in F, \exists (-x) \in F \text{ s.t. } x + (-x) = 0$
5.  $\forall x, y \in F, xy = yx$
6.  $\forall x, y, z \in F, (xy)z = x(yz)$
7. *Multiplicative identity*  $\exists 1 \in F, \text{ s.t. } \forall x \in F, 1 \times x = x$
8. *Multiplicative inverse*  $\forall x \in F, x \neq 0, \exists x^{-1} \text{ s.t. } x \times x^{-1} = 1$

**Note** Every field has at least 2 elements: 0, the *additive identity* and 1, the *multiplicative identity*.

**Examples**

1.  $\mathbb{R}$  is a field.
2.  $\mathbb{Z}$  is not a field.
3.  $\mathbb{N}$  is not a field.
4.  $\mathbb{Q}$  is a field.
5. Irrational numbers is not a field.

**15.2 Complex Numbers**

**Definition** The set of **complex number**  $\mathbb{C}$  is the set of ordered pair of real numbers together with the following rules on basic operations.

1. Addition:  $(a, b) + (c, d) = (a + c, b + d)$
2. Multiplication:  $(a, b) \cdot (c, d) = (ac - bd, ad + bc)$

With set notation we define complex numbers as

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

\* altogether with operations of addition and multiplication defined above.

**Note** (*Connection to  $\mathbb{R}$* ) Any complex number with second component as 0,  $(a, 0)$  is identified as  $a \in \mathbb{R}$ , i.e.  $\mathbb{R} \subsetneq \mathbb{C}$

**Alt. notation**  $\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R} \wedge i^2 = -1\}$

**Definition** Let  $w, z \in \mathbb{C}$ , we define  $w$  **equals**  $z$  as,

$$w = z \iff (\Re(z) = \Re(w)) \wedge (\Im(z) = \Im(w))$$

**Definition** Let  $z = a + ib \in \mathbb{C}$  then the **conjugate** of  $z$  is  $\bar{z} = a + i(-b)$ , and if  $z \neq 0$ , then the **inverse** of  $z$  could be computed as

$$z^{-1} = \frac{\bar{z}}{z\bar{z}}$$

**Definition** A field  $F$  is **algebraically closed** if every polynomial of degree  $n$  in  $F$  has  $n$  roots in  $F$ . (Counting multiplicities)

**Examples**  $\mathbb{C}$  is algebraically closed and  $\mathbb{R}$  is not.

## 16 Lecture16 Mar.7 2018

### 16.1 Vector space over a field

**Definition** A vector space over field  $F$  is a set  $V$  together with two operations, addition and scalar multiplication s.t. [Very similar to those those defining properties for real vector space.]

### 16.2 Complex vector space

Complex vector space  $\mathbb{C}^n = \{(z_1, \dots, z_n) | z_1, \dots, z_n \in \mathbb{C}\}$  is a vector space over  $\mathbb{C}$ , with dimension  $n$  and standard basis  $\{\vec{e}_1, \dots, \vec{e}_n\}$

**Definition** Let  $F$  be a field, then

$$F^n = \{(x_1, \dots, x_n) | x_1, \dots, x_n \in F\}$$

and

$$\dim(F^n) = n$$

$F^n$  is a vector space over field  $F$  w.r.t. usual coordinate wise addition and scalar multiplication.

**Definition** Let  $V$  vector space over field  $F$ , then  $\{\vec{x}_1, \dots, \vec{x}_n\}$  is **linearly independent** if and only if

$$\sum_{i=1}^i c_i \vec{x}_i = \vec{0}, c_1, \dots, c_n \in F \implies c_1 = \dots = c_n = 0 \in F$$

**Definition**  $\text{span}\{\vec{x}_1, \dots, \vec{x}_n\}$  is defined as

$$\left\{ \sum_{i=1}^n c_i \vec{x}_i | c_1, \dots, c_n \in F \right\}$$

**Definition** Consider  $V, W$  as two vector spaces over fields  $F$  then transformation  $T : V(F) \rightarrow W(F)$  is **linear** if and only if

$$\forall \vec{v}_1, \vec{v}_2 \in V, c, d \in F, T(c\vec{v}_1 + d\vec{v}_2) = cT(\vec{v}_1) + dT(\vec{v}_2)$$

## 17 Lecture 17 Mar.13 2018

**Theorem** Let  $T : V \rightarrow V$  be a linear operator, and  $\beta$  is a basis for vector space  $V$ . Let  $W_i$  be the span of first  $i$  vectors in  $\beta$ , then  $[T]_\beta^\beta$  is *upper-triangular* if and only if

$$T(W_i) \subset W_i, \forall i$$

**Definition** Let  $T : V \rightarrow V$  be a linear operator, a subspace  $W$  of  $V$  is called **invariant** under  $T$  (*T-invariant*) if and only if

$$T(W) \subset W$$

**Examples** For linear operator  $T : V \rightarrow V$ ,<sup>1</sup>

1.  $V$
2.  $\{\vec{0}\}$
3.  $\text{Ker}(T)$
4.  $\text{Im}(T)$
5.  $E_\lambda(T)$  for any eigenvalue  $\lambda$  of  $T$ <sup>2</sup>
6.  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  defined as

$$T((x, y, z)) = (3x + 2y, y - z, 4x + 2y - z)$$

Then subspace of  $\mathbb{R}^3$ :  $W = \{(x, y, x) \mid x, y \in \mathbb{R}\}$  is  $T$ -invariant.

**Theorem** Let  $T : V \rightarrow V$  be a linear operator,  $\beta = \{\vec{x}_1, \dots, \vec{x}_k\}$  is a basis for  $V$ , then  $[T]_\beta^\beta$  is upper-triangular if and only if  $W_i$ , defined as the span of first  $i$  vectors in  $\beta$ , is  $T$ -invariant for all  $i \leq k$ .

**Note**  $\{\vec{0}\} \subset W_1 \subset W_2 \subset W_3 \cdots \subset W_k = V$

**Definition** Linear operator  $T : V \rightarrow V$  is said to be **triangularizable** if there exists a basis  $\beta$  for  $V$  such that  $[T]_\beta^\beta$  is upper-triangular.

---

<sup>1</sup>Proofs are omitted.

<sup>2</sup>As eigenspace is defined as kernel.

**Remark** (Consider property of determinant of triangular form matrix) If  $[T]_{\beta}^{\beta}$  is upper-triangular, the characteristic polynomial  $P_T(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n)$  where  $\lambda_i$  are entries on the main diagonal.

**Remark** Entries above main diagonal are **not** uniquely determined by  $T$ , it's also depends on the choice of basis  $\beta$ .

## 18 Lecture 18 Mar.14 2018

### 18.1 Triangular form

**Theorem** Let  $V$  be a vector space over field  $F$ , let  $T : V \rightarrow V$  be a linear operator, suppose the characteristic polynomial has  $\dim(V)$  roots in  $F$ ,<sup>1</sup> then there exists  $\beta$  as a basis of  $V$  such that  $[T]_{\beta}^{\beta}$  is upper-triangular.

**Fact** Any transformation  $T : V \rightarrow V$  whose eigenvalues all have multiplicity of 1, then  $T$  is diagonalizable. (Since there would be  $\dim(V)$  unique eigenvalues.)

**Contra-positive of above fact** non-diagonalizable  $\implies \exists \lambda_i$  with multiplicity greater than 1.

**Consider** *Break down the problem* Linear operator  $T : V \rightarrow V$

1. **Case 1**  $T$  has only eigenvalue 0 with multiplicity of  $\dim(V)$ .
2. **Case 2**  $T$  has only eigenvalue  $\lambda$  with multiplicity of  $\dim(V)$ . If  $T$  has only eigenvalue  $\lambda$  then  $S = (T - \lambda I)$  has eigenvalue 0 only, as in case 1.
3. **Case 3**  $T$  has multiple eigenvalues. *the direct sum of single eigenvalue case.*

---

<sup>1</sup>i.e. field  $F$  is algebraically closed, e.g.  $F = \mathbb{C}$

## 18.2 Nilpotent transformation

**Theorem** Let  $V$  be a vector space over  $\mathbb{C}$  and linear operator  $T : V \rightarrow V$  has only eigenvalue 0 if and only if  $T^k = \mathbf{0}$ <sup>1</sup> for some  $k \in \mathbb{Z}^+$ .

*Proof.*

$\leftarrow$  Suppose  $T^k = 0$  for some  $k \in \mathbb{Z}^+$

Let  $\vec{x} \neq \vec{0}$  be an eigenvector for  $T$ ,

And  $\lambda$  is the corresponding eigenvalue,

Then  $T(\vec{x}) = \lambda\vec{x}$

(Inductively)  $T^k(\vec{x}) = \lambda^k\vec{x}$

Since  $\vec{x} \neq \vec{0} \wedge T^k(\vec{x}) = \vec{0}$

$\implies \lambda^k = 0$

$\implies \lambda = 0$

$\rightarrow$  Suppose only eigenvalue of  $T$  is 0.

We know there exists basis for  $V$ ...

so the matrix of  $T$  relative to this basis is upper-triangular...

with 0 along diagonal.

And matrix of  $T^2$  relative to this basis has 0 on the super diagonal

And with every composition of additional  $T$ ,...

the zero diagonal is pushed up for at least one step higher.

Eventually, for the worst case we could guarantee  $T^{\dim(V)} = 0$

**Note:** the actual value of  $k$  might be smaller than  $\dim(V)$ ,...

and  $k$  is bounded above by  $\dim(V)$ .

As composition of zero transformations is zero,

There must exist  $k \leq \dim(V)$  s.t.  $T^k = 0$

■

**Definition** A linear operator  $T : V \rightarrow V$  is called **nilpotent** if

$$\exists k \in \mathbb{Z}^+ \text{ s.t. } T^k = 0$$

the *smallest* possible  $k$  that  $T^k = 0$  is called the **order/index** of  $T$ .

<sup>1</sup>The 0 here stands for zero transformation.

**Theorem** (Same as above theorem) A linear operator  $T : V \rightarrow V$  is nilpotent if and only if  $T$  has only eigenvalue 0.

**Example 1** Let  $T : P_n(\mathbb{C}) \rightarrow P_n(\mathbb{C})$  and  $T(p(x)) = p'(x)$ ,  $T$  is nilpotent with order  $n + 1$ .

**Example 2** Let  $T : P_4(\mathbb{C}) \rightarrow P_4(\mathbb{C})$  and  $T(p(x)) = p''(x) + p'''(x)$ ,  $T$  is nilpotent with order 3.

**Example 3/Theorem** If  $T^{k-1}(\vec{x}) \neq \vec{0}$  for non-zero  $\vec{x}$ , and  $T^k(\vec{x}) = \vec{0}$ , i.e.  $T$  is a nilpotent transformation with degree  $k$ . Then  $\beta = \{T^{k-1}(\vec{x}), \dots, T(\vec{x}), \vec{x}\}$  is linearly independent. And  $\beta$  is called a **cycle** of  $T$  generated by initial vector  $\vec{x}$ .

*Proof.*

$$\text{If } (\star) = c_{k-1}T^{k-1}(\vec{x}) + \dots + c_1T(\vec{x}) + c_0\vec{x} = \vec{0}$$

Apply  $T^{k-1}$  on both sides of above equation,

$$\text{That's } T^{k-1}(\star) = T^{k-1}(\vec{0}) = \vec{0}$$

$$\implies c_0T^{k-1}(\vec{0}) = \vec{0}$$

$$\implies c_0 = 0$$

$$\text{Recursively, } c_i = 0 \ \forall i \in \mathbb{Z}_0^{k-1}$$

Therefore  $\beta$  is linearly independent. ■

**Theorem** Let  $T : V \rightarrow V$  be a nilpotent with degree  $n = \dim(V)$ , then there exists  $\vec{x} \in V$  (*not necessarily unique*) such that

$$\beta = \{T^{n-1}(\vec{x}), \dots, T(\vec{x}), \vec{x}\}$$

is a basis for  $V$ . And  $[T]_\beta^\beta$  is upper-triangular with zero on main diagonal and one on super-diagonal, and zero elsewhere, like,

$$[T]_\beta^\beta = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$



*Proof.*

Since  $T^n = 0 \wedge T^{n-1} \neq 0$

Therefore  $\beta$  is linearly independent by result from example 3

And  $\beta$  contains  $n$  vectors, so  $\beta$  is a basis for  $V$ . ■

## 19 Lecture 19 Mar.20 2018

**Next Goal** If  $T : V \rightarrow V$  is nilpotent in order between 1 and  $\dim(V)$ , then the matrix of  $T$  relative to some basis is in the form of

$$\begin{pmatrix} J_{m_1} & 0 & \dots & 0 \\ 0 & J_{m_2} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & J_{m_k} \end{pmatrix}$$

where  $J_{m_i} \in \mathbb{M}_{m_i \times m_i}(F)$  in the form with ones on super-diagonal and zeros elsewhere.

**Essential procedures** Identify vectors

1. in  $\text{Ker}(T)$
2. in  $\text{Ker}(T^2) \setminus \text{Ker}(T)$
3. in  $\text{Ker}(T^3) \setminus \text{Ker}(T^2)$
4. ...

**Claim**

$$\{\vec{0}\} \subseteq \text{Ker}(T) \subseteq \text{Ker}(T^2) \subseteq \dots \subseteq \text{Ker}(T^k) = V$$

**Theorem** Let  $T : V \rightarrow V$  is nilpotent of order  $k$ , let  $W$  be a subspace of  $\text{Ker}(T^k)$  s.t.  $W \cap \text{Ker}(T^{k-1}) = \{\vec{0}\}$ , then

$$\dim(T^i(W)) = \dim(W), \forall i < k$$

*Proof.*

Let  $\{\vec{w}_1, \dots, \vec{w}_s\}$  be a basis for subspace  $W$

So  $\dim(W) = s$

Let  $i < k$ , know  $\{T^i(\vec{w}_1), \dots, T^i(\vec{w}_s)\}$  spans  $T^i(W)$

WTS linear independency, so that  $\{T^i(\vec{w}_j)\}$  is a basis for  $T^i(W)$

So that we could show they have same dimension by checking the sizes of their bases.

$$\text{Consider } \sum_{j=1}^s c_j T^i(\vec{w}_j) = \vec{0}$$

$$\text{That's } T^i\left(\sum_{j=1}^s c_j \vec{w}_j\right) = \vec{0}$$

Applying  $T^{k-i-1}$  on both side of above equation

$$T^{k-1}\left(\sum_{j=1}^s c_j \vec{w}_j\right) = \vec{0}$$

$$\text{So } \sum_{j=1}^s c_j \vec{w}_j \in W \cap \text{Ker}(T^{k-1})$$

$$\text{Therefore } \sum_{j=1}^s c_j \vec{w}_j = \vec{0} \in W$$

Since  $\{\vec{w}_1, \dots, \vec{w}_s\}$  is a basis for  $W$

So  $c_1 = c_2 = \dots = c_s = 0$

So  $\{\vec{w}_1, \dots, \vec{w}_s\}$  is a basis for  $T^i(W)$

So  $\dim(T^i(W)) = \dim(W) = s$

■

## 20 Lecture 20 Mar.21 2018

### 20.1 Nilpotent Transformations

**Goal** Show that every nilpotent  $T : V \rightarrow V$  can be brought into canonical form (in some basis).

**Theorem** Two nilpotent transformations are similar (i.e. they represents the same transformations relative to different bases) if and only if they have the same canonical form.

## 20.2 Canonical Forms for Transformations $T : V \rightarrow V$ with Single Eigenvalue $\lambda$

If linear operator  $T$  has only eigenvalue  $\lambda$  the linear operator  $(T - \lambda I)$  (*nilpotent*) has eigenvalue 0 only. Therefore, linear operator  $T : V \rightarrow V$  has only eigenvalue  $\lambda$  means operator  $(T - \lambda I)$  is nilpotent, so for some bases  $\beta$  of  $V$ ,  $[T - \lambda I]_{\beta}^{\beta}$  could be in canonical form.

$$[T - \lambda I]_{\beta}^{\beta} = J = \begin{pmatrix} J_{m_1} & 0 & 0 & 0 \\ 0 & J_{m_2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & J_{m_k} \end{pmatrix}$$

and for the matrix of original transformation,

$$[T]_{\beta}^{\beta} = J + \lambda I$$

## 20.3 Graph(Computational aspect)

omitted

## 21 Lecture21 Mar.27 2018

### 21.1 Goal

**Goal** Prove for all  $T : V \rightarrow V$  can decompose  $V$  into direct sum of two invariant subspaces s.t. on one subspace,<sup>1</sup>  $T$  has only single eigenvalue  $\lambda$  and on other no eigenvalue of  $T$  is  $\lambda$ .

**Definition** Let  $\lambda$  be an eigenvalue of  $T : V \rightarrow V$ , the **generalized eigenspace** corresponding to eigenvalue  $\lambda$  is

$$K_{\lambda} = \{\vec{x} \in V \mid (T - \lambda I)^i(\vec{x}) = \vec{0} \text{ for some } i \in \mathbb{Z}^+\}$$

*In the definition,  $i$  might be different for different  $\vec{x}$ .*

---

<sup>1</sup>Transformation  $T$  restricted to this particular subspace.

**Note 1**  $K_\lambda = \text{Ker}(T - \lambda I)^k$  for some  $k$ . Since

$$\{\vec{0}\} \subset \text{Ker}(T - \lambda I) \subset \text{Ker}(T - \lambda I)^2 \dots$$

the chain cannot grow forever must eventually stabilize.<sup>1</sup> That's, there exists a (smallest)  $k$  s.t.  $\text{Ker}(T - \lambda I)^k = \text{Ker}(T - \lambda I)^{k+1}$ , more generally the  $\text{Ker}(T - \lambda I)^l = \text{Ker}(T - \lambda I)^k$ ,  $\forall l > k$ .  $k$  is the degree where kernel gets stabilized.

**Note 2**  $K_\lambda$  is  $T$  invariant.  $\iff (\vec{v} \in K_\lambda \implies T(\vec{v}) \in K_\lambda)$

*Proof.*

$$\begin{aligned} &\text{Let } \vec{v} \in K_\lambda \\ &\text{i.e. } (T - \lambda I)^i(\vec{v}) = \vec{0}, \forall i \geq k \\ &\text{Consider } (T - \lambda I)^{k+1}(\vec{v}) = \vec{0} \\ &\implies (T - \lambda I)^k(T - \lambda I)(\vec{v}) = \vec{0} \\ &\implies (T - \lambda I)^k T(\vec{v}) - \lambda(T - \lambda I)^k(\vec{v}) = \vec{0} \\ &\implies (T - \lambda I)^k T(\vec{v}) - \vec{0} = \vec{0} \\ &\implies (T - \lambda I)^k T(\vec{v}) = \vec{0} \end{aligned}$$

We have shown that operator  $(T - \lambda I)^k$  maps  $T(\vec{v})$  to  $\vec{0}$   
 $\implies T(\vec{v}) \in K_\lambda$

■

**Note 3** The only eigenvalue of  $T$  on  $K_\lambda$  is  $\lambda$ . Equivalently,

$$T(\vec{v}) = \mu \vec{v} \implies \mu = \lambda$$

*Proof.*

Consider  $(T - \lambda I)^i(\vec{v}) = (\mu - \lambda)^i(\vec{v}) = \vec{0}$  by definition of generalized eigenspace.

Since  $\vec{v} \neq \vec{0}$  by definition of eigenvector.

So  $\mu = \lambda$

■

---

<sup>1</sup>As kernel is a subspace of  $V$ , its dimension could not exceed  $\dim(V)$ .

**Note 4**  $V = Ker(T - \lambda I)^k \oplus Im(T - \lambda I)^k$

*Check:  $Im(T - \lambda I)^k$  is  $T$ -invariant*

*Proof.*

By dimension theorem,

$$dim(V) = dim(Ker(T - \lambda I)^k) + dim(Im(T - \lambda I)^k)$$

So to prove direct sum only need to show

$$Ker(T - \lambda I)^k \cap Im(T - \lambda I)^k = \{\vec{0}\}$$

Let  $\vec{v} \in Ker(T - \lambda I)^k \cap Im(T - \lambda I)^k$

Since  $\vec{v}$  is in the image, there exists  $\vec{w} \in V$

$$s.t. \vec{v} = (T - \lambda I)^k(\vec{w}) \in Ker(T - \lambda I)^k$$

$$\text{Therefore } (T - \lambda I)^k(\vec{v}) = (T - \lambda I)^k((T - \lambda I)^k(\vec{w}))$$

$$= (T - \lambda I)^{2k}(\vec{w}) = \vec{0} \text{ since } 2k > k$$

$$\implies \vec{w} \in Ker(T - \lambda I)^{2k} = Ker(T - \lambda I)^k$$

$$\implies \vec{v} = (T - \lambda I)^k(\vec{w}) = \vec{0}$$

$$\implies Ker(T - \lambda I)^k \cap Im(T - \lambda I)^k = \{\vec{0}\}$$

$$\text{Therefore } V = Ker(T - \lambda I)^k \oplus Im(T - \lambda I)^k$$

■

**Note 5**  $T : V \rightarrow V$  is a linear operator and  $\lambda$  is an eigenvalue of  $T$  with multiplicity  $m$ , then

$$dim(E_\lambda) = m$$

*In generally, the dimension of generalized eigenspace is equal to the multiplicity of  $\lambda$*

*Proof.*

By Note 4,  $V = \text{Ker}(T - \lambda I)^k \oplus \text{Im}(T - \lambda I)^k$

Let  $\alpha, \beta$  be respective bases for  $\text{Ker}(T - \lambda I)^k, \text{Im}(T - \lambda I)^k$

$\implies \gamma = \alpha \cup \beta$  is a basis for  $V$

Let  $\text{Ker}$  denote  $\text{Ker}(T - \lambda I)^k$

Let  $\text{Im}$  denote  $\text{Im}(T - \lambda I)^k$

$$[T]_{\gamma}^{\gamma} = \begin{bmatrix} [T|_{\text{Ker}}]_{\alpha}^{\alpha} & 0 \\ 0 & [T|_{\text{Im}}]_{\beta}^{\beta} \end{bmatrix}$$

$$\implies P_T(x) = P_{T|_{\text{Ker}}}(x) \times P_{T|_{\text{Im}}}(x)$$

Since multiplicity of eigenvalue  $\lambda$  is  $m$ , factoring out,

$$\implies P_T(x) = (x - \lambda)^m q(x), \quad q(x) \neq 0$$

Since  $\lambda$  is the only eigenvalue for  $T|_{\text{Ker}}$

$$P_{T|_{\text{Ker}}}(x) = (x - \lambda)^l$$

Now WTS  $m = l$

For  $T|_{\text{Im}}$ , it has no eigenvalue equals  $\lambda$

Let  $\vec{v} \in \text{Im}(T - \lambda I)^k$  and  $T(\vec{v}) = \lambda \vec{v}$

$$\vec{v} = (T - \lambda I)^k(\vec{w}) \text{ for some } \vec{w}$$

$$\implies T(\vec{v}) = T(T - \lambda I)^k(\vec{w}) = \lambda(T - \lambda I)^k(\vec{w})$$

$$\implies (T - \lambda I)^k(\vec{w}) \in E_{\lambda} \subset \text{Ker}(T - \lambda I)^k$$

$$\implies (T - \lambda I)^k(\vec{w}) \in \text{Ker}(T - \lambda I)^k \cap \text{Im}(T - \lambda I)^k = \{\vec{0}\}$$

Contradict the fact that eigenvector cannot be zero vector,

Therefore  $\lambda$  cannot be an eigenvalue of  $T|_{\text{Im}}$

$$\implies P_{T|_{\text{Im}}}(\lambda) \neq 0$$

$$\text{So } (x - \lambda)^m q(x) = (x - \lambda)^l P_{T|_{\text{Im}}}$$

$$\text{Where } q(x) \neq 0 \wedge P_{T|_{\text{Im}}}(x) \neq 0$$

$$\implies l = m$$

■

**Goal / crucial idea**  $T : V \rightarrow V$  is a linear operator with  $\lambda$  as an eigenvalue with multiplicity  $m$ , then

$$V = \text{Ker}(T - \lambda I)^k \oplus \text{Im}(T - \lambda I)^k = K_\lambda \oplus \text{Im}(T - \lambda I)^k$$

and both  $\text{Ker}(T - \lambda I)^k$  and  $\text{Im}(T - \lambda I)^k$  are invariant under  $T$ , the only eigenvalue of  $T|_{\text{Ker}(T - \lambda I)^k}$  is  $\lambda$  and no eigenvalue of  $T|_{\text{Im}(T - \lambda I)^k}$  is equal to  $\lambda$ . Also  $\dim(K_\lambda) = m$ .

**Implication** Let  $V$  be a vector space over  $\mathbb{C}$ , and  $T : V \rightarrow V$  be a linear operator with distinct eigenvalues  $\{\lambda_i, \dots, \lambda_l\}$  then

$$V = \bigoplus_{i=1, \dots, l} K_{\lambda_i}$$

*Proof(Sketch).*

$$V = K_{\lambda_1} \oplus \text{Im}(T - \lambda_1 I)^k$$

Apply induction on  $\dim(V)$

Keep splitting, one-by-one, until there are no more eigenvalues left. ■

So  $V$  is a vector space over  $\mathbb{C}$ ,  $T : V \rightarrow V$  has matrix (in some basis)

$$\begin{pmatrix} B_{\lambda_1} & 0 & 0 & 0 \\ 0 & B_{\lambda_2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & B_{\lambda_l} \end{pmatrix}$$

where  $B_{\lambda_i}$  is a Jordan block. And the matrix is called **Jordan canonical form** of  $T$ , and is unique *up to ordering of Jordan blocks*.

**Theorem** Two matrices are similar (i.e. representing same transformation relative to different bases) if and only if they have same JCF.

**Note** If  $T$  is diagonalizable, then

$$V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_l}$$

*Diagonal form is one of Jordan canonical form.*

## 22 Lecture22 Mar.28 2018

### 22.1 Examples on finding JCF.

**Example 1** Let  $T : \mathbb{R}^4 \rightarrow \mathbb{R}^4$  be a linear transformation and  $T$  has matrix  $A$  relative to standard basis of  $\mathbb{R}^4$ ,

$$A = \begin{pmatrix} 2 & -2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Find Jordan canonical form for  $T$  and a canonical basis.

*Solution:*

Omitted

■

**Example 2** Let  $T : \mathbb{R}^6 \rightarrow \mathbb{R}^6$  has matrix

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

Find the Jordan Canonical Form of  $T$ .

*Solution:*

Omitted

■