MAT224 Linear Algebra II
Lecture Notes

Tianyu Du

January 2018

Info.
Created: January. 9 2018
Last modified: April 3, 2019

Contents

1 Lecture1 Jan.9 2018 **3**
1.1 Vector spaces 3
1.2 Examples of vector spaces 4
1.3 Some properties of vector spaces 5

2 Lecture2 Jan.10 2018 **6**
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

3 Lecture3 Jan.16 2018 **9**
3.1 Linear Combination 9
3.2 Combination of subspaces 10

4 Lecture4 Jan.17 2018 **13**
4.1 Cont'd 13
4.2 Linear Independence 14

5 Lecture5 Jan.23 2018 . . . **15**
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

6 Lecture6 Jan.24 2018 . . . **20**
6.1 Basis and Dimension . . . 20

7 Lecture7 Jan.30. 2018 . . . **23**
7.1 Linear Transformations . . . 23
7.2 Properties of linear transformations . . . 24
7.3 Definitions . . . 26

8 Lecture8 Jan.31 2018 . . . **27**
8.1 Linear Transformations . . . 27
8.2 Applications of dimension theorem . . . 29

9 Lecture9 Feb.6 2018 . . . **30**
9.1 Applications of dimension theorem . . . 30
9.2 Isomorphisms . . . 33
9.3 Coordinates . . . 35

10 Lecture10 Feb.7 2018 . . . **36**
10.1 Matrix of linear transformation . . . 36

11 lecture11 Feb.13 2018 . . . **38**
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

12 Lecture12 Feb.14 2018 . . . **42**

13 Lecture13 Feb.27 2018 . . . **44**
13.1 Diagonalization . . . 44
13.2 How to find eigenvalues and eigenvectors of T . . . 44

14 Lecture14 Feb.28 2018 **47**

15 Lecture15 Mar.6 2018 **50**
15.1 Fields 50
15.2 Complex Numbers 51

16 Lecture16 Mar.7 2018 **52**
16.1 Vector space over a field 52
16.2 Complex vector space 52

17 Lecture 17 Mar.13 2018 **53**

18 Lecture 18 Mar.14 2018 **54**
18.1 Triangular form 54
18.2 Nilpotent transformation 55

19 Lecture 19 Mar.20 2018 **57**

20 Lecture 20 Mar.21 2018 **58**
20.1 Nilpotent Transformations 58
20.2 Canonical Forms for Transformations $T : V \to V$ with Single Eigenvalue λ 59
20.3 Graph(Computational aspect) 59

21 Lecture21 Mar.27 2018 **59**
21.1 Goal 59

22 Lecture22 Mar.28 2018 **64**
22.1 Examples on finding JCF. 64

1 Lecture1 Jan.9 2018

1.1 Vector spaces

Definition A real [1] **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: $\forall \vec{v}, \vec{y} \in V, \vec{x} + \vec{y} = \vec{y} + \vec{x}$

[1] 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: $\exists\ \vec{0} \in V s.t. \forall \vec{x} \in V, \vec{x} + \vec{0} = \vec{x}$
5. **AI** Additive Inverse: $\forall \vec{x} \in V, \exists - \vec{x} \in V s.t. \vec{x} + (-\vec{x}) = \vec{0}$
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.[1] usual component-wise addition and scalar multiplication.

Example 2 $\mathbb{M}_{m\times n}(\mathbb{R})$ 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 $\mathbb{P}^\star_n(\mathbb{R})$ 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: $(1 + x^n), (1 - x^n) \in \mathbb{P}^\star_n(\mathbb{R})$ but $(1 + x^n) + (1 - x^n) = 2 \notin \mathbb{P}^\star_n(\mathbb{R})$

[1] 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.

$$\text{Assume } \vec{0}, \vec{0^\star} \text{ are two zero vectors in } V$$
$$\text{WTS: } \vec{0} = \vec{0^\star}$$
$$\text{Since } \vec{0} \text{ is the zero vector, by Z } \vec{0^\star} + \vec{0} = \vec{0^\star}$$
$$\text{Similarly, } \vec{0} + \vec{0^\star} = \vec{0}$$
$$\text{Also, } \vec{0} + \vec{0^\star} = \vec{0^\star} + \vec{0} \text{ by commutative vector addition.}$$
$$\text{So, } \vec{0^\star} = \vec{0}$$

■

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.

$$\text{By property of number 0: } 0\vec{x} = (0+0)\vec{x}$$
$$\text{By DSA: } 0\vec{x} = 0\vec{x} + 0\vec{x}$$
$$\text{By AI, } \exists(-0\vec{x}) s.t.$$
$$0\vec{x} + (-0\vec{x}) = 0\vec{x} + 0\vec{x} + (-0\vec{x})$$
$$\text{By AA}$$
$$\implies 0\vec{x} = \vec{0}$$

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* <u>*additive inverse*</u> *with scalar multiplication with (-1)*)[1].
proof.

$$\begin{aligned}(-\vec{x}) &= (-\vec{x}) + \vec{0} \quad \text{By property of zero vector} \\ &= (-\vec{x}) + 0\vec{x} \quad \text{By property3} \\ &= (-\vec{x}) + (1 + (-1))\vec{x} \quad \text{By property of zero as real number} \\ &= (-\vec{x}) + 1\vec{x} + (-1)\vec{x} \\ &= \vec{0} + (-1)\vec{x} \\ &= (-1)\vec{x}\end{aligned}$$

■

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

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

[1] The scalar multiplication here is the one defined in vector space V.

proof.

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

■

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[1] 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 us closed under scalar multiplication. $\forall \vec{x} \in W, c \in \mathbb{R}, c\vec{x} \in W$.

Proof.

[1] 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 Clousure
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

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

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,

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

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.

$\blacksquare$

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

$$c_1\vec{v_1} + \cdots + c_n\vec{v_n}$$

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

$$c_1\vec{w_1} + \cdots + c_k\vec{w_k} \in W$$

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

Theorem Let V be a vector space, let $\vec{v_1}, \ldots, \vec{v_k} \in V$ then the set of all linear combination of $\vec{v_1}, \ldots, \vec{v_k}$

$$W = \{\sum_{i=1}^{k} c_i \vec{v_i} | c_i \in \mathbb{R} \forall\ i\}$$

is a subspace of V.
proof.

$$\begin{gathered} \text{Consider } \vec{0} \in W \\ \text{So, } W \neq \emptyset \\ \text{Let } c \in \mathbb{R}, \text{ Let } \vec{x} \in W \wedge \vec{y} \in W \\ \text{By definition of span, we have,} \\ \vec{x} = \sum_{i=1}^{k} a_i \vec{v_i}, \quad \vec{y} = \sum_{i=1}^{k} b_i \vec{v_i} \\ \text{Consider, } \vec{x} + c\vec{y} \\ \vec{x} + c\vec{y} = \sum_{i=1}^{k} a_i \vec{v_i} + c\sum_{i=1}^{k} b_i \vec{v_i} = \sum_{i=1}^{k} (a_i + cb_i)\vec{v_i} \in W \end{gathered}$$

■

Definition Let V be a vector space, $\vec{v_1}, \ldots, \vec{v_k} \in V$, **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.

$$\begin{gathered}\text{Falsify the statement by providing counter-example:}\\ \text{Consider,}\\ W_1 = \{(x_1, x_2) \mid x_1 \in \mathbb{R}, x_2 = 0\}\\ W_2 = \{(x_1, x_2) \mid x_2 \in \mathbb{R}, x_1 = 0\}\\ \begin{pmatrix}0\\1\end{pmatrix} \in W_1 \cup W_2 \quad \begin{pmatrix}1\\0\end{pmatrix} \in W_1 \cup W_2\\ \text{But, } \begin{pmatrix}0\\1\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix} \notin W_1 \cup W_2\end{gathered}$$

■

proof.

$$\begin{gathered}\text{Because } W_1 \text{ and } W_2 \text{ are both subspaces, so}\\ \vec{0} \in W_1 \cap W_2 \implies W_1 \cap W_2 \neq \emptyset\\ \text{Let } \vec{x}, \vec{y} \in W_1 \cap W_2, c \in \mathbb{R}\\ \text{Consider, } \vec{x} + c\vec{y}\\ \text{Sine } W_1, W_2 \text{ are subspaces,}\\ \vec{x} + c\vec{y} \in W_1 \wedge \vec{x} + c\vec{y} \in W_2\\ \implies \vec{x} + c\vec{y} \in W_1 \cap W_2\\ \text{So, } W_1 \cap W_2 \text{ is a subspace.}\end{gathered}$$

■

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 $\vec{y} = 0 \in W_2$, $\forall \vec{x} \in W_1,\ \vec{x} \in W_1 + W_2$. 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.

$$\text{Let } \vec{x_1}, \vec{x_2} \in W_1, \vec{y_1}, \vec{y_2} \in W_2$$
$$\text{By properties of subspaces,}$$
$$\forall c \in \mathbb{R}, \vec{x_1} + c\vec{x_1} \in W_1 \wedge \vec{y_2} + c\vec{y_2} \in W_2$$
$$\text{Consider, } \vec{x_1} + \vec{y_1} \in W_1 + W_2, \vec{x_2} + \vec{y_2} \in W_1 + W_2$$
$$(\vec{x_1} + \vec{y_1}) + c(\vec{x_2} + \vec{y_2})$$
$$= (\vec{x_1} + c\vec{x_2}) + (\vec{y_1} + c\vec{y_2}) \in W_1 + W_2$$

■

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.

$$\text{(Forward direction) Suppose } V = W_1 \oplus W_2$$
$$\text{WTS. } V = W_1 + W_2 \wedge W_1 \cap W_2 = \{\vec{0}\}$$
$$\text{Let } V = W_1 \oplus W_2$$
$$\implies \forall \vec{x} \in V, \text{ can be written uniquely as}$$
$$\vec{x} = \vec{w_1} + \vec{w_2},\ \vec{w_1} \in W_1,\ \vec{w_2} \in W_2$$
$$\implies V = W_1 + W_2 \text{ by definition of } \textit{sum}.$$
$$\text{Let } \vec{x} \in W_1 \cap W_2$$
$$\text{Decomposition, let } \vec{z} \in W_1 \cap W_2 \subseteq V$$
$$\vec{z} = \vec{z} + \vec{0},\ \vec{z} \in W_1, \vec{0} \in W_2$$
$$\vec{z} = \vec{0} + \vec{z},\ \vec{0} \in W_1, \vec{z} \in W_2$$
$$\text{Since decomposition is unique, } \vec{z} = \vec{0}$$
$$\text{So, } W_1 \cap W_2 = \{\vec{0}\}$$
$$\text{(Backward direction) Suppose } V = W_1 + W_2 \wedge W_1 \cap W_2 = \{\vec{0}\}$$
$$\text{WTS. } V = W_1 \oplus W_2$$
$$\text{Assume } \vec{x} = \vec{w_1} + \vec{w_2},\ \vec{w_1} \in W_1, \vec{w_2} \in W_2$$
$$\vec{x} = \vec{w_1'} + \vec{w_2'},\ \vec{w_1'} \in W_1, \vec{w_2'} \in W_2$$
$$\implies \vec{w_1} + \vec{w_2} = \vec{w_1'} + \vec{w_2'}$$
$$\implies \vec{w_1} - \vec{w_1'} = \vec{w_2'} - \vec{w_2}$$
$$\text{Where, by definition of subspace, } \vec{w_1} - \vec{w_1'} \in W_1 \wedge \vec{w_2'} - \vec{w_2} \in W_2$$
$$\text{So, } \vec{w_1} - \vec{w_1'} = \vec{w_2'} - \vec{w_2} \in W_1 \cap W_2$$
$$\text{Since } W_1 \cap W_2 = \{\vec{0}\}$$
$$\implies \vec{w_1} = \vec{w_1'} \wedge \vec{w_2} = \vec{w_2'}$$
$$\text{So the decomposition is unique.}$$

∎

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

$$span\{\vec{x_1}, \dots \vec{x_n}, \vec{x}\} = span\{\vec{x_1}, \dots \vec{x_n}\}$$

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

$$\text{let } \vec{x} \in span\{\vec{x}, \dots, \vec{x_n}\}$$
$$\vec{x} = \sum_{i=1}^{n} c_i \vec{x_i} \text{ for } c_i \in \mathbb{R}\ \forall i$$
$$\text{So, } span\{\vec{x_1}, \dots, \vec{x_n}, \vec{x}\} = \{\sum_{i=1}^{n} a_i \vec{x_i} + z\vec{x} \mid a_i, z \in \mathbb{R}\forall i\}$$
$$= \{\sum_{i=1}^{n} a_i \vec{x_i} + z \sum_{i=1}^{n} c_i \vec{x_i} \mid a_i, c_i \in \mathbb{R}\forall i\}$$
$$= \{\sum_{i=1}^{n} (a_i + zc_i)\vec{x_i} \mid a_i, c_i \in \mathbb{R}\forall i\}$$
$$\text{Let } d_i = a_i + zc_i \in \mathbb{R}$$
$$= \{\sum_{i=1}^{n} d_i \vec{x_i} \mid d_i \in \mathbb{R}\forall i\}$$
$$= span\{\vec{x_1}, \dots \vec{x_n}\}$$

■

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 $\{c_1, \cdots, c_n\}$ 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

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

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

$$\exists \vec{x} \in \{\vec{v_i}\}_{i=1}^n \ s.t. \ \vec{x_j} \in span\{\{\vec{v_i}\}_{i=1}^n \backslash \{\vec{x}\}\}$$

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

$$\forall \vec{x} \in \{\vec{v_i}\}_{i=1}^n, \ \vec{x_j} \notin span\{\{\vec{v_i}\}_{i=1}^n \backslash \{\vec{x}\}\}$$

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.

$$\sum_{i=1}^{k} c_1 \vec{x_k} = 0 (\star)$$

are $c_1 = \cdots = c_k = 0$
linearly dependent means <u>at least one</u> $c_i \neq 0$, $(\star)$ still holds.

5.1.1 Alternative definitions of linear independency

Definition(Alternative.1) $\{\vec{x_1} \dots \vec{x_k}\}$ is **linearly independent** <u>iff</u> none of them can be written as a linear combination of the remaining $k-1$ vectors.[1]

Definition(Alternative.2) $\{\vec{x_1} \dots \vec{x_k}\}$ is **linearly dependent** <u>iff</u> at least one of them can be written as a linear combination of the remaining $k-1$ vectors. [2]

5.2 Basis

Definition Let V be a vector space, a non-empty[3] set S of vectors from V is a **basis** for V if

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

[1] See theorem from the pervious lecture.
[2] See theorem from the pervious lecture.
[3] Specially, for an empty set, we define 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,

$$c_1\vec{x_1} + \cdots + c_k\vec{x_k} - (b_1\vec{x_1} + \cdots + d_k\vec{x_k}) = \vec{0}$$

$$\iff \sum_{i=1}^{k} (c_i - b_i)\vec{x_1} = \vec{0}$$

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: $V = span\{S\} \wedge S$ 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

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

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

■

Example

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

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 $\{(1, 0), (6, 3)\}$.

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

proof.

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

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

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

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

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 <u>linearly dependent</u>.

Fundamental Theorem (Alternative) If V is vector space spanned by m vectors, then any <u>linearly independent</u> 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.

$$\begin{gathered}\text{Since } S \text{ spans } V \text{ and } T \text{ is linearly independent}\\ \therefore l \le k\\ \text{(flip) Since } T \text{ spans } V \text{ and } S \text{ is linearly independent}\\ \therefore k \le l\\ \implies l \le k \wedge k \le l\\ \implies k = l\end{gathered}$$

■

Definition So we can define the **dimension** of V, as $dim(V)$ as the number vectors in <u>any</u> 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 <u>linearly dependent</u>.

Note $k \le n \not\Rightarrow S$ is linear dependent.

Theorem If W is subspace of vector space V, then

1. $dim(W) \le dim(V)$
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 $\vec{x} \notin span\{S\}$

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 cab be* ***reduced*** *to basis for V by removing the linearly dependent(redundant) vector in the spanning set, using* <u>*redundancy theorem*</u> *to get a linearly independent spanning set.*

2. *Linear independent set can be* ***enlarged*** *to a basis for V.*

proof.

<u>omitted.</u>

■

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 $1 \notin span\{p_1(x), p_2(x)\}$, 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 $x \notin span\{p_1(x), p_2(x)\}$

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.

$$\text{Let } \{\vec{v_i}\}_1^k \text{ be basis for } U \cap W$$
$$\implies dim(U \cap W) = k$$
$$\text{Since } \{\vec{v_i}\}_1^k \text{ is basis for } U \cap W \text{ then it's a linearly independent subset of U}$$
$$\text{So it could be enlarged ti basis for } U, \ \{\vec{v_1}, \dots, \vec{v_k}, \vec{y_1}, \dots, \vec{y_r}\}$$
$$\text{So } dim(U) = k + r$$
$$\text{We also could enlarge a basis for } W \ \{\vec{v_1}, \dots, \vec{v_k}, \vec{z_1}, \dots, \vec{z_s}\}$$
$$\implies dim(V) = k + s$$
$$\text{WTS. } \{\vec{v_1}, \dots, \vec{v_k}, \dots, \vec{y_1}, \dots, \vec{y_r}, \vec{z_1}, \dots, \vec{z_s}\} \text{ is a basis for } U + W$$
$$\text{(If we could show this) } dim(U + W) = k + r + s = (k + r) + (k + s) - k$$
$$= dim(U) + dim(W) - dim(U \cap W)$$
$$\text{Obviously, the above set spans } U + W$$
$$\text{WTS. } \{\vec{v_1}, \dots, \vec{v_k}, \dots, \vec{y_1}, \dots, \vec{y_r}, \vec{z_1}, \dots, \vec{z_s}\} \text{ is linearly independent}$$
$$\text{Consider } a_1\vec{v_1} + \cdots + a_k\vec{v_k} + b_1\vec{y_1} + \dots b_r\vec{y_r} + c_1\vec{z_1} + \cdots + c_s\vec{z_s} = \vec{0} \ (\star)$$
$$\text{From } (\star) \implies \sum(c_i\vec{z_i}) = -\sum(a_i\vec{v_i}) - \sum b_i\vec{y_i}$$
$$\implies \sum(c_i\vec{z_i}) \in U \wedge \sum(c_i\vec{z_i}) \in W$$
$$\iff \sum(c_i\vec{z_i}) \in U \cap W$$
$$\text{Since } \{\vec{v_i}\} \text{ is a basis for } U \cap W$$
$$\implies \sum(c_i\vec{z_i}) = \sum(d_i\vec{v_i})$$
$$\iff \sum(c_i\vec{z_i}) - \sum(d_i\vec{v_i}) = \vec{0} \in W$$
$$\implies c_i = d_i = 0 \text{ since } \{\vec{z_i}, \vec{v_i}\} \text{ is a basis}$$
$$\text{Rewrite } (\star)$$
$$\sum(a_i\vec{v_i}) + \sum b_i\vec{y_i} = 0 \in U$$
$$\implies a_i = b_i = 0 \text{ since } \{\vec{v_i}, \vec{y_i}\} \text{ is a basis for } U$$

■

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 shat $dim(U) = dim(W) = 2$, 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 $dim(U) = dim(V) = (n-1)$, proof:

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

7 Lecture7 Jan.30. 2018

7.1 Linear Transformations

Definition Let V, W be vector spaces, a function $T : V \to W$ is a **linear transformation**[1] if

1. $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})\ \forall \vec{x}, \vec{y} \in V$[2]
2. $T(c\vec{x}) = cT(\vec{x})\ \forall \vec{x} \in V,\ c \in \mathbb{R}$[3]

Linear transformation preserves vector additions ans saclar multiplications on vector spaces.

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

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

Linear transformations preserves linear combinations.

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

[1] In some textbooks, this is annotated as **linear mapping**.

[2] 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.

[3] 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 $T : \mathbb{R}^n \to \mathbb{R}^m$ defined as

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

is linear.

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

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

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

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

7.2 Properties of linear transformations

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

proof.

$$\text{Let}\{\vec{v_1}, \dots, \vec{v_k}\} \text{ be any basis for } V$$

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

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

$$\implies T(\vec{x}) = T(\sum_{i=1}^{k} c_i\vec{v_i})$$

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

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

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 $T : V \to W$ 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.

$$\text{Since } A \text{ is a subspace so it's non-empty, therefore } \exists T(\vec{x}),\ \vec{x} \in A$$
$$\text{So T(A)} \neq \emptyset$$
$$\text{Let } \vec{w_1}, \vec{w_2} \in T(A)$$
$$\implies \vec{w_1} = T(\vec{x_1}), \vec{w_2} = T(\vec{x_2}),\ \vec{x_1}, \vec{x_2} \in A$$
$$\implies \vec{w_1} + \vec{w_2} = T(\vec{x_1}) + T(\vec{x_2}) = T(\vec{x_1} + \vec{x_2}) \text{ since } T \text{ is linear.}$$
$$\text{Since } \vec{x_1} + \vec{x_2} \in A \text{ by the definition of subspaces.}$$
$$\implies \vec{w_1} + \vec{w_2} \in T(A)$$
$$\text{So } T(A) \text{ is closed under vector addition.}$$
$$\text{Let } \vec{w} \in T(A)$$
$$\implies \vec{w} = T(\vec{x}), \vec{x} \in A$$
$$\text{Let } c \in \mathbb{R}$$
$$\text{Consider } c\vec{w} = cT(\vec{x}) = T(c\vec{x})$$
$$\text{Since } c\vec{x} \in A$$
$$\text{So } c\vec{w} \in T(A)$$
$$\text{So } T(A) \text{ is closed under scalar multiplication.}$$

■

Property(derived from the definition) For all linear transformation $T : V \to W$, we have [1]

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

Property(iii) Let transformation $T : V \to W$ 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. [2]

[1] 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.

[2] The pre-image and inverse share the same notation, but in this case, transformation T is not necessarily invertible.

proof.

$$\text{Let } \vec{w_1}, \vec{w_2} \in T^{-1}(B)$$
$$\implies T(\vec{w_1}), T(\vec{w_2}) \in B$$
$$\implies aT(\vec{w_1}) + b(\vec{w_2}) \in B,\ \forall a, b \in \mathbb{R} \text{ since } B \text{ is a subspace.}$$
$$\implies T(a\vec{w_1} + b\vec{w_2}) \in B$$
$$\implies a\vec{w_1} + b\vec{w_2} \in T^{-1}(B)$$

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

So $T^{-1}(B)$ is a subspace.

■

7.3 Definitions

Let $T : V \to W$ 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 $T : P_3(\mathbb{R}) \to P_3(\mathbb{R})$ is <u>linear</u> defined by

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

find $Ker(T)$.

Theorem Let $T : V \to W$ be a linear transformation, let $\{\vec{v_1}, \ldots, \vec{v_k}\}$ be the <u>spanning set</u> of V [1], then $\{T(\vec{v_1}), \ldots, T(\vec{v_k})\}$ spans $Im(T)$

proof.

$$\text{Let } \vec{w} \in Im(T)$$
$$\text{Since } V = span\{\vec{v_1}, \ldots, \vec{v_k}\}$$
$$\text{For any } \vec{x} \in V \text{ can be written as}$$
$$\vec{x} = \sum_{i=1}^{k} c_i \vec{v_i}, \ c_i \in \mathbb{R}$$
$$\implies \vec{w} = T(\vec{x}) = T(\sum_{i=1}^{k} c_i \vec{v_i})$$
$$= \sum_{i=1}^{k} c_i T(\vec{v_i})$$
$$\text{as a linear combination of } \{T(\vec{v_1}), \ldots, T(\vec{v_k})\}$$
$$\text{So } Im(T) = span\{T(\vec{v_1}), \ldots, T(\vec{v_k})\}$$

■

8 Lecture8 Jan.31 2018

8.1 Linear Transformations

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

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

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.[2]. So by the <u>redundancy theorem</u> we could remove the last vector. There we have

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

[1] The set is only the spanning set of V, it's not necessarily to be a basis of V.
[2] Notice that the first three vectors is a basis of $P_2(\mathbb{R})$.

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

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

$$dim(V) = Nullity(T) + Rank(T)$$

Proof.

$$\text{Say } dim(V) = n$$
$$\text{Let } \{\vec{v_1}, \ldots, \vec{v_k}\} \text{ be a basis for } Ker(T)$$
$$\text{Since } Ker(T) \text{ is a subspace of } V, \text{ the set } \{\vec{v_i}\}_1^k \text{ is a subset of } V,$$
$$\text{It can be extended to a basis } \{\vec{v_i}\}_1^k \cup \{\vec{v_i}\}_{k+1}^n \text{ for V.}$$
$$\text{Claim: } \{T(\vec{v_{k+1}}), \ldots, T(\vec{v_n})\} \text{ is basis for } Im(T)$$
$$\text{If the claim is true, this prove the theorem since}$$
$$dim(Ker(T)) + dim(Im(T)) = k + n - k = n = dim(V)$$
$$\because T(\vec{v_i}) = \vec{0}, \ \forall i \in \mathbb{Z}_1^k$$
$$\text{and by the definition of kernel of linear transformation,}$$
$$\therefore \{T(\vec{v_i})\}_{k+1}^n \text{ spans } Im(T)$$
$$\text{Show if } \sum_{i=k+1}^{n} c_i T(\vec{v_i}) = \vec{0} \implies c_i = 0$$
$$\implies T(\sum_{i=k+1}^{n} c_i \vec{v_i}) = \vec{0}$$
$$\implies \sum_{i=k+1}^{n} c_i \vec{v_i} \in Ker(T)$$
$$\implies \sum_{i=k+1}^{n} c_i \vec{v_i} = \sum_{i=1}^{k} c_i \vec{v_i}$$
$$\implies c_1\vec{v_1} + \cdots + c_k\vec{v_k} - c_{k+1}\vec{v_{k+1}} - \cdots - c_n\vec{v_n} = \vec{0}$$
$$\text{Since } \{\vec{v_i}\}_i^n \text{ is a basis for } V.$$
$$\implies c_i = 0 \ \forall i$$

8.2 Applications of dimension theorem

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

Theorem Let transformation $T : V \to W$ 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 $T : P_2(\mathbb{R}) \to \mathbb{R}^2$ 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 $T : P_2(\mathbb{R}) \to \mathbb{R}^3$ defined by

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

Take $p(x) = a + bx + cx^2 \in P_2(\mathbb{R})$, $p(x) \in Ker(T)$ iff $T(p(x)) = \vec{0}$.
Let $p(x) \in Ker(T)$,
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 $dim(Ker(T)) = 0$. Therefore, T is **injective**.
By *dimension theorem*,

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

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

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

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

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

Theorem Let $T : V \to W$ 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 $\sum_i c_i v_i \in Ker(T)$. By definition of injective transformation, $\sum_i c_i v_i = \vec{0}$. 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 $T : V \to W$ 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 baed on image of a basis.

Proof.

$$\text{Let } \{\vec{v_i}\}_{i=1}^n \text{ be a basis of } V$$
$$\text{Consider } T(\vec{x}) = \vec{0}$$
$$\text{Since } \{\vec{v_i}\}_{i=1}^n \text{ is a basis}$$
$$\text{Let } \vec{x} = \sum c_i \vec{v_i}$$
$$\text{Assume } \vec{x} \in Ker(T)$$
$$T(\vec{x}) = \vec{0} \iff T(\sum c_i \vec{v_i}) = \vec{0}$$
$$\implies \sum c_i T(\vec{v_i}) = \vec{0}$$
$$\text{Since } \{T(\vec{v_i})\}_{i=1}^n \text{ are linearly independent.}$$
$$\implies c_i = 0$$
$$\text{Therefore } \vec{x} = \sum 0\vec{v_i} = \vec{0}$$
$$\text{Therefore } Ker(T) = \{\vec{0}\}$$
$$\text{Therefore } dim(Ker(T)) = 0$$
$$\implies \text{injective}$$

■

Theorem Let $T : V \to W$ be a linear transformation, [1]

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

[1] Consider the contrapositive predicates of this theorem.

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

Proof.

$$dim(V) = dim(Ker(T)) + dim(Im(T))$$
$$\because dim(Im(T)) \leq dim(W)$$
$$\therefore dim(V) \leq dim(Ker(T) + dim(W)$$
$$\implies dim(Ker(T)) \geq dim(V) - dim(W)$$
$$\implies dim(Ker(T)) > 0$$
$$\text{So } T \text{ could not be injective}$$
$$dim(V) = dim(Ker(T)) + dim(Im(T))$$
$$\because dim(Ker(T)) \geq 0$$
$$\therefore dim(V) \geq dim(Im(T))$$
$$\implies dim(Im(T)) < dim(W)$$
$$\text{So } T \text{ could not be surjective}$$

∎

Proof 2.

$$\text{Consider a transformation } T : V \to W \text{ is bijective.}$$
$$\text{By the contrapositive form of above theorem,}$$
$$\text{Injective} \implies dim(V) \leq dim(W)$$
$$\text{Surjective} \implies dim(V) \geq dim(W)$$
$$\text{Therefore bijective}$$
$$\implies dim(V) \leq dim(W) \wedge dim(V) \geq dim(W) \iff dim(V) = dim(W)$$
$$\text{Therefore bijective} \implies dim(V) = dim(W)$$
$$\text{So, take contrapositive, } dim(V) \neq dim(W) \implies \text{ not bijective.}$$

∎

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

Proof.

$$\text{By dimension theorem}$$
$$dim(V) = dim(Ker(T)) + dim(Im(T)) = dim(W)$$
$$\text{If injective } dim(Ker(T)) = 0$$
$$\implies dim(Im(T)) = dim(W)$$
$$\text{So surjective}$$
$$\text{If surjective } dim(Im(T)) = dim(W) = dim(V)$$
$$\implies dim(Ker(T)) = 0$$
$$\text{So injective}$$

■

9.2 Isomorphisms

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

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

Proof.

$$\rightarrow\ V, W \text{ isomorphic } \implies dim(V) = dim(W)$$
$$\text{Isomorphic means there exists a bijective transformation } T$$
$$\text{By dimension theorem } dim(V) = dim(Ker(T)) + dim(Im(T))$$
$$= 0 + dim(W)$$
$$\leftarrow\ dim(V) = dim(W) \implies V, W \text{ isomorphic}$$
$$\text{Equivalently, find a isomorphism(bijective) transformation}$$
$$\text{Let } \{\vec{v_i}\}_{i=1}^n \text{ be basis for } V$$
$$\text{Let } \{\vec{w_i}\}_{i=1}^n \text{ be basis for } W$$
$$\text{Claim } T : V \to W \text{ defined by}$$
$$T(\vec{v_i}) = \vec{w_i} \text{ is an isomorphism.}$$
$$\text{If } \vec{x} \in Ker(T) \subseteq V$$
$$\vec{x} = \sum c_i \vec{v_i}$$
$$\vec{0} = T(\vec{x})$$
$$= \sum c_i T(\vec{v_i})$$
$$= \sum (c_i \vec{w_i})$$
$$\implies c_i = 0 \text{ since } \vec{w_i} \text{ are basis.}$$
$$\implies \vec{x} = \vec{0}$$
$$\implies dim(Ker(T)) = 0$$
$$\implies \text{injective} \iff \text{surjective}$$
$$\text{Therefore } V \text{ and } W \text{ are isomorphic vector spaces.}$$

■

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

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

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

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

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

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

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 $T : V \to \mathbb{R}^n$

Procedure:

Let $\{\vec{v_i}\}_{i=1}^n$ be <u>any</u> 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 nay basis for V, $\forall \vec{x} \in V$ can be written uniquely as

$$\vec{x} = c_1\vec{v_1} + \cdots + c_n\vec{v_n}$$

then $c_1, \dots, c_n$ is called the <u>**coordinates** for $\vec{x}$ relative to basis α</u>, 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 α, α' are any two bases for V then *generally* $[\vec{x}]_\alpha \neq [\vec{x}]_{\alpha'}$ (except $\vec{0}$)[1].

10 Lecture10 Feb.7 2018

10.1 Matrix of linear transformation

Recall Let V be a vector space, let α 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 $\vec{x} \to [\vec{x}]_\alpha$ is an isomorphism that $V \to \mathbb{R}^n$.

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

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

So that

$$[T(\vec{x})]_\beta = [\sum c_i T(\vec{v_i})]_\beta = \sum c_i [T(\vec{v_i})]_\beta$$

$$= \begin{bmatrix} [T(\vec{v_1})]_\beta & \dots & [T(\vec{v_n})]_\beta \end{bmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

$\begin{bmatrix} [T(\vec{v_1})]_\beta & \dots & [T(\vec{v_n})]_\beta \end{bmatrix}$ is called the <u>the **matrix of** T w.r.t. bases α, β.</u> Denoted as $[T]_\alpha^\beta$, and by definition we have

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

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

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

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

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

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

$$\begin{gathered}
T(1-x) = x(1-x) = x - x^2 \\
x - x^2 = (-1)(1) + 2(1+x) + (-1)(1+x+x^2) + 0(1-x^3) \\
[T(1-x)]_\beta = (-1, 2, -1, 0) \\
T(1-x^2) = x - x^3 \\
[T(1-x^2)]_\beta = (-2, 1, 0, 1) \\
[T(x)] = x^2 \\
[T(x)]_\beta = (0, -1, 1, 0) \\
[T]_\alpha^\beta = \begin{bmatrix} -1 & -2 & 0 \\ 2 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}
\end{gathered}$$

Picture Let V, W be two vectors spaces, $\alpha = \{\vec{v_1}, \ldots, \vec{v_n}\}$ is a basis for V and $\beta = \{\vec{w_1}, \ldots, \vec{w_m}\}$ is a basis for W.

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

Remark

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

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

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

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

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

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

Solution

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

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

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

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

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

11 lecture11 Feb.13 2018

11.1 Algebra of Transformation

Theorem Let $T : V \to W$ be a linear transformation, where $\alpha = \{\vec{v_1}, \ldots, \vec{v_n}\}$ and $\beta = \{\vec{w_1}, \ldots, \vec{w_m}\}$ are bases for V, W respectively.

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

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

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

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

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

Theorem And, let α and β 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 $T : V \to W$ and $S : W \to U$ be two linear transformations, then the **composition** $ST : V \to U$ 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.

$$\begin{gathered} \text{Let } a, b \in \mathbb{R}, \ \vec{x}, \vec{y} \in V \\ ST(a\vec{x} + b\vec{y}) \\ = S(T(a\vec{x} + b\vec{y})) \\ = S(aT(\vec{x}) + bT(\vec{y})) \\ = a(ST(\vec{x})) + b(ST(\vec{y})) \end{gathered}$$

■

11.2 Matrix of composition of transformations

Consider $T : V \to W$ and $S : W \to U$ 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 $[ST]_\alpha^\gamma$.

$$\begin{gathered}\forall \vec{x} \in V, [ST]_\alpha^\gamma[\vec{x}]_\alpha \\ = [(ST)(\vec{x})]_\gamma \\ = [S(T(\vec{x}))]_\gamma \\ = [S]_\beta^\gamma[T(\vec{x})]_\beta \\ = [S]_\beta^\gamma[T]_\alpha^\beta[\vec{x}]_\alpha \\ \text{This holds true for all } \vec{x} \in V \\ \therefore [ST]_\alpha^\gamma = [S]_\beta^\gamma[T]_\alpha^\beta\end{gathered}$$

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

11.3 Inverse transformations

Theorem $T : V \to W$ is *isomorphism*[1] if and only if there exists function $S : W \to V$ such that

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

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

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

proof.($\leftarrow$) Assume $S : W \to V$ is the inverse of $T : V \to W$ 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. ■

[1] 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 $T : V \to W$ is an isomorphism then the inverse of T, $T^{-1} : W \to V$ is linear.[1]

Proof.

$$\begin{gathered}
\text{WTS } T^{-1}(a\vec{w_1} + b\vec{w_2}) = aT^{-1}(\vec{w_1}) + bT^{-1}(\vec{w_2}) \forall a, b \in \mathbb{R}, \forall \vec{w_1}, \vec{w_2} \in W \\
T^{-1}(\vec{w_1}) \text{ is the unique } \vec{x_1} \ s.t. \ T(\vec{x_1}) = \vec{w_1} \\
T^{-1}(\vec{w_2}) \text{ is the unique } \vec{x_2} \ s.t. \ T(\vec{x_2}) = \vec{w_2} \\
T^{-1}(a\vec{w_1} + b\vec{w_2}) \text{ is the unique } \vec{x} \ s.t. \ T(\vec{x}) = a\vec{w_1} + b\vec{w_2} \\
\because T(\vec{x}) = a\vec{w_1} + b\vec{w_2} \\
= aT(\vec{x_1}) + bT(\vec{x_2}) \\
= T(a\vec{x_1} + b\vec{x_2}) \\
\therefore \vec{x} = a\vec{x_1} + b\vec{x_2} \\
\text{Also } T(\vec{x}) = a\vec{w_1} + b\vec{w_2} \\
\therefore \vec{x} = T^{-1}(a\vec{w_1} + b\vec{w_2}) = a\vec{x_1} + b\vec{x_2} \\
= aT^{-1}(\vec{w_1}) + bT^{-1}(\vec{w_2})
\end{gathered}$$

■

Theorem $T : V \to W$ is isomorphism, then let α and β 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.

[1] Note: the conclusion could be changed into isomorphism.

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

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

Proof.

$$\begin{gathered} \text{Let } \vec{x} \in V \\ I(\vec{x}) = \vec{x} \\ [I(\vec{x})]_{\alpha'} = [\vec{x}]_{\alpha'} \\ [I]_{\alpha}^{\alpha'}[\vec{x}]_{\alpha} = [\vec{x}]_{\alpha'} \end{gathered}$$

■

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

Computation Let $\alpha = \{\vec{a_1}, \ldots, \vec{a_n}\}$, then[1]

$$[I]_{\alpha}^{\alpha'} = [[\vec{a_1}]_{\alpha'} \mid \cdots \mid [\vec{a_n}]_{\alpha'}]$$

12 Lecture12 Feb.14 2018

Recall Let α and β be bases for V and $I : V \to V$ 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 β be a basis for $P_2(\mathbb{R})$ and

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

find the basis β.
Solution <u>omitted</u>

[1]Construct column by column.

Theorem Suppose $T : V \to W$ is linear, α and α' are any two bases for V and β and β' are any two bases of W, then,

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

Proof.

$$\textbf{Recall } T = ITI$$
$$\text{Consider let } \vec{x} \in V$$
$$\begin{aligned} &[I]_{\beta}^{\beta'}[T]_{\alpha}^{\beta}[I]_{\alpha'}^{\alpha}[\vec{x}]_{\alpha'} \\ &= [I]_{\beta}^{\beta'}[T]_{\alpha}^{\beta}[\vec{x}]_{\alpha} \\ &= [I]_{\beta}^{\beta'}[T(\vec{x})]_{\beta} \\ &= [T(\vec{x})]_{\beta'} \\ &= [T]_{\beta'}^{\alpha'}[\vec{x}]_{\alpha'} \\ \implies & [T]_{\beta'}^{\alpha'} = [I]_{\beta}^{\beta'}[T]_{\alpha}^{\beta}[I]_{\alpha'}^{\alpha} \end{aligned}$$

■

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 <u>invertible</u> matrix P s.t.

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

Interpretation [1] 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 $T : V \to V$ is **diagonalizable** if and only $\exists$ a basis β for V s.t.

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

is diagonal.

Note Let $\beta = \{\vec{v_1}, \ldots, \vec{v_n}\}$ be a basis, $T : V \to V$ is diagonalizable if and only if $[T]_\beta^\beta$ is in form $\begin{bmatrix} \lambda_1 & ... & 0 \\ ... & ... & ... \\ 0 & ... & \lambda_n \end{bmatrix}$

Definition $T : V \to V$ 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 $\lambda \in \mathbb{R}$. λ is called the **eigenvalue** of T corresponding to vector $\vec{x}$.

Theorem Linear operator $T : V \to V$ 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, <u>in the same order</u>.

13.2 How to find eigenvalues and eigenvectors of T

Definition The **determinant** of T is defined as $det([T]_\alpha^\alpha)$ for <u>any</u> basis α for V.

Remark The determinant of linear operator T does <u>not</u> depends on the choice of basis of α for V, *since similar matrices have the same determinant.*

[1] 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.

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

■

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

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

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

Proof.

<u>By definition</u>

■

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

Note To find eigenvalues and eigenvectors of $T : V \to V$, choose any basis β for V, $\vec{x}$ is an eigenvector with corresponding eigenvalue λ if and only if $[\vec{x}]_\beta$ is an eigenvector of $[T]_\beta^\beta$ with corresponding eigenvalue λ.
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 β be a basis of V consisting of eigenvectors of T and α 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 λ_0 is an eigenvalue of linear operator $T : V \to V$, let $dim(E_{\lambda 0}) = k$, then $(\lambda - \lambda_0)^k$ divides $P_T(\lambda)$

Proof.

$$\text{Let}\{\vec{v_1}, \ldots, \vec{v_k}\} \text{ be basis for } E_{\lambda_0}$$
$$\text{Since } E_{\lambda_0} \subset V$$
$$\text{Let } dim(V) = n$$
$$\text{Extend basis of } E_{\lambda_0} \text{ to basis of } V.$$
$$\alpha = \{\vec{v_1}, \ldots, \vec{v_k}\} \cup \{\vec{v_{k+1}}, \ldots, \vec{v_n}\}$$
$$\text{Since } \vec{v_i} \in E_{\lambda 0},$$
$$\text{Therefore } T(\vec{v_i}) = \lambda_0 \vec{v_i}$$
$$[T]_\alpha^\alpha = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$$
$$\text{Where } A = diag(\lambda_0, \ldots, \lambda_0) \in \mathbb{M}_{k\times k}(\mathbb{R})$$
$$\text{And } B \in \mathbb{M}_{k\times n-k}(\mathbb{R}), D \in \mathbb{M}_{n-k\times n-k}(\mathbb{R})$$
$$P_T(\lambda) = det(A - \lambda I) * det(D - \lambda I)$$
$$= (\lambda_0 - \lambda)^k * det(D - \lambda I)$$
$$\text{Therefore}(\lambda - \lambda_0)^k \mid P_T(\lambda)$$

■

Definition The **multiplicity** of eigenvalue λ_0 is the number of times $(\lambda - \lambda_0)$ appears as a factor in $P_T(\lambda)$.

Note If eigenvalue λ has multiplicity m, the above theorem says

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

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

Theorem If $\lambda_1, \ldots, \lambda_k$ are *distinct* eigenvalues of $T : V \to V$ and $\alpha = \{\vec{x_1}, \ldots, \vec{x_k}\}$ are corresponding eigenvectors, then the set α is *linearly inde-*

pendent.

Proof.

Exercise

■

($\star$)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

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

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

Proof.

$$\leftarrow$$

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

$$\text{Consider } E_{\lambda_i}$$

$$\text{Take basis for } E_{\lambda_i},\ \text{note as } \{\vec{v}^i_1, \ldots, \vec{v}^i_{m_i}\}$$

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

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

All we need to show is linear independence.

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

$$\text{Consider } \sum_{j=1}^{m_i} c_{ij}\vec{v^i_j} \in E_{\lambda_i} = \vec{x_i}$$

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

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

Therefore, $\{\vec{x_{i1}}, \ldots, \vec{x_{ik}}\}$ is linearly independent

$$\text{So } \vec{x_i} = \vec{0}\ \forall i$$

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

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

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

Therefore T is diagonalizable.

$$\rightarrow$$

Suppose T is diagonalizable,

Since T is diagonalizable, then exists basis for V consisting of eigenvectors, say α

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

Where λ_1 takes first m_1 rows, λ_2 takes the next m_2 rows, etc.

$$P_T(\lambda) = det([T]^{\alpha}_{\alpha} - \lambda I)$$
$$= \prod_{i=1}^{k} (\lambda_i - \lambda)^{m_i}$$
$$\text{Since } 1 \le dim(E_{\lambda_i}) \le m_i \ \forall i$$
$$\implies dim(E_{\lambda_i}) = m_i \ \forall i$$

■

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

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 <u>ordered pair</u> of real numbers together with the following rules on basic operations.

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

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 $\overline{z} = a + i(-b)$, and if $z \neq 0$, then the **inverse** of z could be computed as

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

Definition A field F is **algebraically closed** is 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

$$\sum_{i=1}^{i} c_i \vec{x_i} = \vec{0},\ c_1, \dots, c_n \in F \implies c_1 = \dots = c_2 = 0 \in F$$

Definition **span**$\{\vec{x_1}, \dots, \vec{x_n}\}$ is defined as

$$\{\sum_{i=1}^{n} c_i \vec{x_i} | c_1, \dots, c_n \in F\}$$

Definition Consider V, W as two vector spaces over fields F then transformation $T : V(F) \to W(F)$ 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 $T : V \to V$ be a linear operator, and β is a basis for vector space V. Let W_i be the span of first i vectors in β, then $[T]_\beta^\beta$ is *upper-triangular* if and only if

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

Definition Let $T : V \to V$ 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 $T : V \to V$,[1]

1. V
2. $\{\vec{0}\}$
3. $Ker(T)$
4. $Im(T)$
5. $E_\lambda(T)$ for any eigenvalue λ of T [2]
6. $T : \mathbb{R}^3 \to \mathbb{R}^3$ 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 $T : V \to V$ be a linear operator, $\beta = \{\vec{x_1}, \ldots, \vec{x_k}\}$ 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 β, 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 $T : V \to V$ is said to be **triangularizable** if there exists a basis β for V such that $[T]_\beta^\beta$ is upper-triangular.

[1] Proofs are omitted.
[2] 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 λ_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 β.

18 Lecture 18 Mar.14 2018

18.1 Triangular form

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

Fact Any transformation $T : V \to V$ 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 $T : V \to V$

1. **Case 1** T has only eigenvalue 0 with multiplicity of $dim(V)$.
2. **Case 2** T has only eigenvalue λ with multiplicity of $dim(V)$. If T has only eigenvalue λ 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.*

[1] 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 $T : V \to V$ has only eigenvalue 0 if and only if $T^k = \mathbf{0}$ [1] for some $k \in \mathbb{Z}^+$.

Proof.

$$\leftarrow \text{Suppose } T^k = 0 \text{ for some } k \in \mathbb{Z}^+$$
$$\text{Let } \vec{x} \neq \vec{0} \text{ be an eigenvector for } T,$$
$$\text{And } \lambda \text{ is the corresponding eigenvalue,}$$
$$\text{Then } T(\vec{x}) = \lambda\vec{x}$$
$$\text{(Inductively) } T^k(\vec{x}) = \lambda^k\vec{x}$$
$$\text{Since } \vec{x} \neq \vec{0} \wedge T^k(\vec{x}) = \vec{0}$$
$$\implies \lambda^k = 0$$
$$\implies \lambda = 0$$

$\rightarrow$ 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 $T : V \to V$ is called **nilpotent** if

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

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

[1] The 0 here stands for zero transformation.

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

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

Example 2 Let $T : P_4(\mathbb{C}) \to P_4(\mathbb{C})$ 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 $\beta = \{T^{k-1}(\vec{x}), \ldots, T(\vec{x}), \vec{x}\}$ is linearly independent. And β is called a **cycle** of T generated by initial vector $\vec{x}$.

Proof.

$$\text{If } (\star) = c_{k-1}T^{k-1}(\vec{x}) + \cdots + c_1T(\vec{x}) + c_0\vec{x} = \vec{0}$$
$$\text{Apply } T^{k-1} \text{ on both sides of above equation,}$$
$$\text{That's } T^{k-1}(\star) = T^{k-1}(\vec{0}) = \vec{0}$$
$$\implies c_0T^{k-1}(\vec{0}) = \vec{0}$$
$$\implies c_0 = 0$$
$$\text{Recursively, } c_i = 0\ \forall i \in \mathbb{Z}_0^{k-1}$$
$$\text{Therefore } \beta \text{ is linearly independent.}$$

■

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

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

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,

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

Proof.

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

Therefore β is linearly independent by result from example 3

And β contains n vectors, so β is a basis for V.

■

19 Lecture 19 Mar.20 2018

Next Goal If $T : V \to V$ 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 $Ker(T)$
2. in $Ker(T^2) \backslash Ker(T)$
3. in $Ker(T^3) \backslash Ker(T^2)$
4. ...

Claim

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

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

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

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

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

$$T^{k-1}(\sum_{j=1}^{s} c_j \vec{w_j}) = \vec{0}$$

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

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

Since $\{\vec{w_1}, \dots, \vec{w_s}\}$ is a basis for W

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

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 $T : V \to V$ 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 $T : V \to V$ with Single Eigenvalue λ

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

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

and for the matrix of original transformation,

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

20.3 Graph(Computational aspect)

<u>omitted</u>

21 Lecture21 Mar.27 2018

21.1 Goal

Goal Prove for all $T : V \to V$ can decompose V into direct sum of two invariant subspaces s.t. on one subspace,[1] T has only single eigenvalue λ and on other no eigenvalue of T is λ.

Definition Let λ be an eigenvalue of $T : V \to V$, the **generalized eigenspace** corresponding to eigenvalue λ 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 $\vec{x}$.

[1]Transformation T restricted to this particular subspace.

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

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

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

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

Proof.

$$\begin{gathered}
\text{Let } \vec{v} \in K_\lambda \\
\text{i.e. } (T - \lambda I)^i(\vec{v}) = \vec{0},\ \forall i \geq k \\
\text{Consider } (T - \lambda I)^{k+1}(\vec{v}) = \vec{0} \\
\implies (T - \lambda I)^k(T - \lambda I)(\vec{0}) = \vec{0} \\
\implies (T - \lambda I)^k T(\vec{v}) - \lambda(T - \lambda I)^k(\vec{v}) = \vec{0} \\
\implies (T - \lambda I)^k T(\vec{v}) - \vec{0} = \vec{0} \\
\implies (T - \lambda I)^k T(\vec{v}) = \vec{0} \\
\text{We have shown that operator } (T - \lambda I)^k \text{ maps } T(\vec{v}) \text{ to } \vec{0} \\
\implies T(\vec{v}) \in K_\lambda
\end{gathered}$$

■

Note 3 The only eigenvalue of T on K_λ is λ. 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 $\mu = \lambda$

■

[1] 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.

$$\text{By dimension theorem,}$$
$$dim(V) = dim(Ker(T-\lambda I)^k) + dim(Im(T-\lambda I)^k)$$
$$\text{So to prove direct sum only need to show}$$
$$Ker(T-\lambda I)^k \cap Im(T-\lambda I)^k = \{\vec{0}\}$$
$$\text{Let } \vec{v} \in Ker(T-\lambda I)^k \cap Im(T-\lambda I)^k$$
$$\text{Since } \vec{v} \text{ is in the image, there exits } \vec{w} \in V$$
$$s.t.\ \vec{v} = (T-\lambda I)^k(\vec{w}) \in Ker(T-\lambda I)^k$$
$$\text{Therefore } (T-\lambda I)^k(\vec{v}) = (T-\lambda I)^k((T-\lambda I)^k(\vec{w}))$$
$$= (T-\lambda I)^{2k}(\vec{w}) = \vec{0} \text{ since } 2k > k$$
$$\Longrightarrow \vec{w} \in Ker(T-\lambda I)^{2k} = Ker(T-\lambda I)^k$$
$$\Longrightarrow \vec{v} = (T-\lambda I)^k(\vec{w}) = \vec{0}$$
$$\Longrightarrow Ker(T-\lambda I)^k \cap Im(T-\lambda I)^k = \{\vec{0}\}$$
$$\text{Therefore } V = Ker(T-\lambda I)^k \oplus Im(T-\lambda I)^k$$

■

Note 5 $T: V \to V$ is a linear operator and λ is an eigenvalue of T with multiplicity m, then
$$dim(K_\lambda) = m$$

In generally, the dimension of generalized eigenspace is equal to the multiplicity of λ

Proof.

$$\text{By Note 4, } V = Ker(T-\lambda I)^k \oplus Im(T-\lambda I)^k$$
$$\text{Let } \alpha, \beta \text{ be respective bases for } Ker(T-\lambda I)^k,\ Im(T-\lambda I)^k$$
$$\implies \gamma = \alpha \cup \beta \text{ is a basis for } V$$
$$\text{Let } Ker \text{ denote } Ker(T-\lambda I)^k$$
$$\text{Let } Im \text{ denote } Im(T-\lambda I)^k$$
$$[T]_\gamma^\gamma = \begin{bmatrix} [T|_{Ker}]_\alpha^\alpha & 0 \\ 0 & [T|_{Im}]_\beta^\beta \end{bmatrix}$$
$$\implies P_T(x) = P_{T|_{Ker}}(x) \times P_{T|_{Im}}(x)$$
$$\text{Since multiplicity of eigenvalue } \lambda \text{ is } m, \text{ factoring out,}$$
$$\implies P_T(x) = (x-\lambda)^m q(x),\ q(x) \neq 0$$
$$\text{Since } \lambda \text{ is the only eigenvalue for } T|_{Ker}$$
$$P_{T|_{Ker}}(x) = (x-\lambda)^l$$
$$\text{Now WTS } m = l$$
$$\text{For } T|_{Im}, \text{ it has no eigenvalue equals } \lambda$$
$$\text{Let } \vec{v} \in Im(T-\lambda I)^k \text{ and } T(\vec{v}) = \lambda\vec{v}$$
$$\vec{v} = (T-\lambda I)^k(\vec{w}) \text{ for some } \vec{w}$$
$$\implies T(\vec{v}) = T(T-\lambda I)^k(\vec{w}) = \lambda(T-\lambda I)^k(\vec{w})$$
$$\implies (T-\lambda I)^k(\vec{w}) \in E_\lambda \subset Ker(T-\lambda I)^k$$
$$\implies (T-\lambda I)^k(\vec{w}) \in Ker(T-\lambda I)^k \cap Im(T-\lambda I)^k = \{\vec{0}\}$$
$$\text{Contradict the fact that eigenvector cannot be zero vector,}$$
$$\text{Therefore } \lambda \text{ cannot be an eigenvalue of } T|_{Im}$$
$$\implies P_{T|_{Im}}(\lambda) \neq 0$$
$$\text{So } (x-\lambda)^m q(x) = (x-\lambda)^l P_{T|_{Im}}$$
$$\text{Where } q(x) \neq 0 \wedge P_{T|_{Im}}(x) \neq 0$$
$$\implies l = m$$

∎

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

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

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

Implication Let V be a vector space over $\mathbb{C}$, and $T : V \to V$ be a linear operator with <u>distinct</u> eigenvalues $\{\lambda_1, \dots, \lambda_l\}$ then

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

Proof(Sketch).

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

Apply induction on $dim(V)$

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

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

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

where B_{λ_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

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

Diagonal form is one of Jordan canonical form.

22 Lecture22 Mar.28 2018

22.1 Examples on finding JCF.

Example 1 Let $T : \mathbb{R}^4 \to \mathbb{R}^4$ 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 $T : \mathbb{R}^6 \to \mathbb{R}^6$ 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

■