# 115B – Linear Algebra

University of California, Los Angeles

Duc Vu

Spring 2021

This is math 115B – Linear Algebra which is the second course of the undergrad linear algebra at UCLA – continuation of 115A(H). Similar to 115AH, this class is instructed by Professor Elman, and we meet weekly on MWF from 2:00 pm to 2:50 pm. There is no official textbook used for the class. You can find the previous linear algebra notes (115AH) with other course notes through my github. Any error in this note is my responsibility and please email me if you happen to notice it.

## Contents

# List of Theorems



# List of Definitions

# §1 Lec 1: Mar 29, 2021

## §1.1 Vector Spaces

Notation: if $\star : A \times B \to B$ is a map (= function) write $a \star b$ for $\star(a, b)$, e.g., $+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ where $\mathbb{Z}$ = the integer.

**Definition 1.1** (Field) — A set $F$ is called a FIELD under

- Addition: $+ : F \times F \to F$
- Multiplication: $\cdot : F \times F \to F$

if $\forall a, b, c \in F$, we have

A1) $(a+b)+c = a+(b+c)$

A2) $\exists\, 0 \in F \ni a + 0 = a = 0 + a$

A3) A2) holds and $\exists x \in F \ni a + x = 0 = x + a$

A4) $a + b = b + a$

M1) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

M2) A2) holds and $\exists 1 \neq 0 \in F$ s.t. $a \cdot 1 = a = 1 \cdot a$ ( 1 is unique and written 1 or $1_F$)

M3) M2) holds and $\forall 0 \neq x \in F \quad \exists y \in F \ni xy = 1 = yx$ ($y$ is seen to be unique and written $x^{-1}$)

M4) $x \cdot y = y \cdot x$

D1) $a \cdot (b + c) = a \cdot b + a \cdot c$

D2) $(a + b) \cdot c = a \cdot c + b \cdot c$

**Example 1.2**

$\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are fields as is

$$\mathbb{F}_2 := \{0, 1\} \text{ with } + : \text{ given by}$$

| + | 0 | 1 |
|---|---|---|
| 0 | 0 | 1 |
| 1 | 1 | 0 |

| • | 0 | 1 |
|---|---|---|
| 0 | 0 | 0 |
| 1 | 0 | 1 |

**Fact 1.1.** Let $p > 0$ be a prime number in $\mathbb{Z}$. Then $\exists$ a field $F_{p^n}$ having $p^n$ elements write $|F_{p^n}| = p^n \quad \forall n \in \mathbb{Z}^+$.

**Definition 1.3** (Ring) — Let $R$ be a set with

- $+ : R \times R \to R$
- $\cdot : R \times R \to R$

satisfying A1) – A4), M1), M2), D1), D2), then $R$ is called a RING.
A ring is called

i) a commutative ring if it also satisfies M4).

ii) an (integral) domain if it is a commutative ring and satisfies

$$\text{M 3')}\; a \cdot b = 0 \implies a = 0 \text{ or } b = 0$$

($0 = \{0\}$ is also called a ring – the only ring with $1 = 0$)

**Example 1.4** (Proof left as exercises)

1. $\mathbb{Z}$ is a domain and not a field.
2. Any field is a domain.
3. Let $F$ be a field
$$F[t] \coloneqq \{\text{polys coeffs in } F\}$$
with usual $+, \cdot$ of polys, is a domain but not a field. So if $f \in F[t]$
$$f = a_0 + a_1 t + \ldots + a_n t^n$$
where $a_0, \ldots, a_n \in F$.
4. $\mathbb{Q} \coloneqq \left\{ \frac{a}{b} | a, b \in \mathbb{Z}, b \neq 0 \right\} < \mathbb{C}$ ($<$ means $\subset$ and $\neq$) with usual $+, \cdot$ of fractions. (when does $\frac{a}{b} = \frac{c}{d}$?)
5. If $F$ is a field
$$F(t) \coloneqq \left\{ \frac{f}{g} | f, g \in F[t],\, g \neq 0 \right\} \text{ (rational function)}$$
with usual $+, \cdot$ of fractions is a field.

**Example 1.5** (Cont'd from above)

6. $\mathbb{Q}[\sqrt{-1}] \coloneqq \left\{\alpha + \beta\sqrt{-1} \in \mathbb{C} | \alpha, \beta \in \mathbb{Q}\right\} < \mathbb{C}$. Then $\mathbb{Q}[\sqrt{-1}]$ is a field and

$$\begin{aligned}\mathbb{Q}(\sqrt{-1}) &\coloneqq \left\{\frac{a}{b} | a, b \in \mathbb{Q}[\sqrt{-1}],\, b \neq 0\right\} \\ &= \mathbb{Q}[\sqrt{-1}] \\ &= \left\{\frac{a}{b} | a, b \in \mathbb{Z}[\sqrt{-1}],\, b \neq 0\right\}\end{aligned}$$

where $\mathbb{Z}[\sqrt{-1}] \coloneqq \left\{\alpha + \beta\sqrt{-1} \in \mathbb{C},\, \alpha, \beta \in \mathbb{Z}\right\} < \mathbb{C}$. How to show this? – rationalize ($\mathbb{Z}[\sqrt{-1}]$ is a domain not a field, $F[t] < F(t)$ if $F$ is a field so we have to be careful).

7. $F$ a field

$$\mathbb{M}_n F \coloneqq \{n \times n \text{ matrices entries in } F\}$$

is a ring under $+, \cdot$ of matrices.

$$1_{\mathbb{M}_n F} = I_n = n \times n \text{ identity matrix } \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

$$0_{\mathbb{M}_n F} = 0 = 0_n = n \times n \text{ zero matrix } \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

is <u>not</u> commutative if $n > 1$.

In the same way, if $R$ is a ring we have

$$\mathbb{M}_n R = \{n \times n \text{ matrices entries in } R\}$$

e.g., if $R$ is a field $\mathbb{M}_n F[t]$.

8. Let $\emptyset \neq I \subset \mathbb{R}$ be a subset, e.g., $[\alpha, \beta], \alpha < \beta \in \mathbb{R}$. Then

$$C(I) = \{f : I \to \mathbb{R} | f \text{ continuous}\}$$

is a commutative ring and not a domain where

$$\begin{aligned}(f \dotplus g)(x) &\coloneqq f(x) + g(x) \\ 0(x) &= 0 \\ 1(x) &= x\end{aligned}$$

for all $x \in I$.

<u>Notation</u>: Unless stated otherwise $F$ is always a field.

**Definition 1.6** (Vector Space) — Let $F$ be a field, $V$ a set. Then $V$ is called a VECTOR SPACE OVER $F$ write $V$ is a vector space over $F$ under

- $+ : V \times V \to V$ – Addition
- $\cdot : F \times V \to V$ – Scalar multiplication

if $\forall x, y, z \in V \quad \forall \alpha, \beta \in F$.

1. $(x + y) + z = x + (y + z)$
2. $\exists 0 \in V \ni x + 0 = x = 0 + x$ (0 is seen to be unique and written 0 or $0_V$)
3. 2) holds and $\exists v \in V \ni x + v = 0 = v + x$ ($v$ is seen to be unique and written $-x$)
4. $x + y = y + x$
5. $1_F \cdot x = x$.
6. $(\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)$
7. $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$
8. $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$

**Remark 1.7.** The usual properties we learned in 115A hold for $V$ a vector space over $F$, e.g., $0_F V = 0_V$, general association law,...

# §2 Lec 2: Mar 31, 2021

## §2.1 Vector Spaces (Cont'd)

**Example 2.1**

The following are vector space over $F$

1. $F^{m\times n} := \{m \times n \text{ matrices entries in } F\}$, usual $+$, scalar multiplication, i.e., if $A \in F^{m\times n}$, let $A_{ij} = ij^{\text{th}}$ entry of $A$. If $A, B \in F^{m\times n}$, then

$$(A+B)_{ij} := A_{ij} + B_{ij}$$
$$(\alpha A)_{ij} := \alpha A_{ij} \quad \forall \alpha \in F$$

   i.e., component-wise operations.

2. $F^n = F^{1\times n} := \{(\alpha_1, \dots, \alpha_n) \,|\alpha_i \in F\}$

3. Let $V$ be a vector space over $F$, $\emptyset \neq S$ a set. Define

$$\mathcal{F}cn(S, V) := \{f : S \to V |\, f \text{ a fcn}\}$$

   Then $\mathcal{F}cn(S,V)$ is a vector space over $F$ $\forall f, g \in \mathcal{F}cn(S,V)$, $\forall \alpha \in F$. For all $x \in S$,

$$f + g : x \mapsto f(x) + g(x)$$
$$\alpha f : x \mapsto \alpha f(x)$$

   i.e.

$$(f+g)(x) = f(x) + g(x)$$
$$(\alpha f)(x) = \alpha f(x)$$

   with 0 by $0(x) = 0_V$ $\forall x \in S$.

4. Let $R$ be a ring under $+, \cdot, F$ a field $\ni F \subseteq R$ with $+, \cdot$ on $F$ induced by $+, \cdot$ on $R$ and $0_F = 0_R$, $1_F = 1_R$, i.e.

$$\underbrace{+}_{\text{on } R}\Big|_{\underbrace{F\times F}_{\text{restrict dom}}} : F \times F \to F \text{ and } \underbrace{\cdot}_{\text{on } R}\Big|_{\underbrace{F\times F}_{\text{restrict dom}}} : F \times F \to F$$

   i.e. closed under the restriction of $+, \cdot$ on $R$ to $F$ and also with $0_F = 0_R$ and $1_F = 1_R$ (we call $F$ a <u>subring</u> of $R$). Then $R$ is a vector space over $F$ by restriction of scalar multiplication, i.e., same $+$ on $R$ but scalar multiplication

$$\cdot\big|_{F\times R} : F \times R \to R$$

   e.g., $\mathbb{R} \subseteq \mathbb{C}$ and $F \subseteq F[t]$.

**Example 2.2** (Cont'd from above)

*Note*: $\mathbb{C}$ is a vector space over $\mathbb{R}$ by the above but as a vector space over $\mathbb{C}$ is different.

5. In 4) if $R$ is also a field (so $F \subseteq R$ is a subfield). Let $V$ be a vector space over $R$. Then $V$ is also a vector space over $F$ by restriction of scalars, e.g., $M_n\mathbb{C}$ is a vector space over $\mathbb{C}$ so is a vector space over $\mathbb{R}$ so is a vector space over $\mathbb{Q}$.

## §2.2 Subspaces

**Definition 2.3** (Subspace) — Let $V$ be a vector space under $+, \cdot, \emptyset \neq W \subseteq V$ a subset. We call $W$ a subspace of $V$ if $\forall w_1, w_2 \in W$, $\forall \alpha \in F$,

$$\alpha w_1,\ w_1 + w_2 \in W$$

with $0_W = 0_V$ is a vector space over $F$ under $+|_{W\times W}$ and $\cdot|_{F\times W}$ i.e., closed under the operation on $V$.

**Theorem 2.4**

Let $V$ be a vector space over $F$, $\emptyset \neq W \subseteq V$ a subset. Then $W$ is a subspace of $V$ iff $\forall \alpha \in F$, $\forall w_1, w_2 \in W$, $\alpha w_1 + w_2 \in W$.

**Example 2.5**

1. Let $\emptyset \neq I \subseteq \mathbb{R}$, $C(I)$ the commutative ring of continuous function $f : I \to \mathbb{R}$. Then $C(I)$ is a vector space over $\mathbb{R}$ and a subspace of $\mathcal{F}cn(I, \mathbb{R})$.
2. $F[t]$ is a vector space over $F$ and $n \geq 0$ in $\mathbb{Z}$.

$$F[t]_n := \{f \mid f \in F[t],\ f = 0 \text{ or } \deg f \leq d\}$$

is a subspace of $F[t]$ (it is not a ring).

Attached is a review of theorems about vector spaces from math 115A.

## §2.3 Direct Sums

**Problem 2.1.** Can you break down an object into simpler pieces? If yes can you do it uniquely?

**Example 2.6**

Let $n > 1$ in $\mathbb{Z}$. Then $n$ is a product of primes unique up to order.

**Example 2.7**

Let $V$ be a finite dimensional inner product space over $\mathbb{R}$ (or $\mathbb{C}$) and $T : V \to V$ a hermitian (=self adjoint) operator. Then $\exists$ an ON basis for $V$ consisting of eigenvectors for $T$. In particular, $T$ is diagonalizable. This means

$$V = E_T(\lambda_1) \perp \ldots \perp E_T(\lambda_r) \tag{*}$$

$E_T(\lambda_i) := \{v \in V \,|\, Tv = \lambda_i v\} \neq 0$ eigenspace of $\lambda_i$; $\lambda_1, \ldots, \lambda_r$ the distinct eigenvalues of $T$. So

$$T\big|_{E_T(\lambda_i)} : E_T(\lambda_i) \to E_T(\lambda_i)$$

i.e., $E_T(\lambda_i)$ is T-invariant and

$$T\big|_{E_T(\lambda_i)} = \lambda_i 1_{E_T(\lambda_i)}$$

and (*) is unique up to order.

Goal: Generalize this to $V$ any finite dimensional vector space over $F$, any $F$, and $T : V \to V$ linear. We have many problems to overcome in order to get a meaningful result, e.g.,

**Problem 2.2.**
1. $V$ may not be an inner product space.
2. $F \neq \mathbb{R}$ or $\mathbb{C}$ is possible.
3. $F \nsubseteq$ is possible, so cannot even define an inner product.
4. $V$ may not have any eigenvalues for $T : V \to V$.
5. If we prove an existence theorem, we may not have a uniqueness one.

We shall show: given $V$ a finite dimensional vector space over $F$ and $T : V \to V$ a linear operator. Then $V$ breaks up uniquely up to order into small $T$-invariant subspace that we shall show are completely determined by polys in $F[t]$ arising from $T$. Motivation: Generalize the concept of linear independence, Spectral Theorem Decomposition, to see how pieces are put together (if possible).

**Definition 2.8** (Span) — Let $V$ be a vector space over $F$, $W_i \subseteq V$, $i \in I$ – may not be finite, subspaces. Let

$$\sum_{i\in I} W_i = \sum_{i \in I} W_i := \left\{ v \in V | \exists w_i \in W_i,\, i \in I,\text{ almost all } w_i = 0 \ni v = \sum_{i\in I} w_i \right\}$$

when almost all zero means only finitely many $w_i \neq 0$. Warning: In a vector space/F we can only take finite linear combination of vectors. So

$$\sum_{i\in I} W_i = \operatorname{Span}\left(\bigcup_{i\in I} W_i\right) = \left\{ \text{finite linear combos of vectors in } \bigcup_{i\in I} W_i \right\}$$

e.g., if $I$ is finite, i.e., $|I| < \infty$, say $I = \{1, \ldots, n\}$ then

$$\sum_{i\in I} W_i = W_1 + \ldots + W_n := \{w_1 + \ldots + w_n | \, w_i \in W_i \, \forall i \in I\}$$

**Definition 2.9** (Direct Sum) — Let $V$ be a vector space over $F$, $W_i \subseteq V$, $i \in I$, subspace. Let $W \subseteq V$ be a subspace. We say that $W$ is the (internal) direct sum of the $W_i$, $i \in I$ write $W = \bigoplus_{i\in I} W_i$ if

$$\forall w \in W \, \exists! \, w_i \in W_i \text{ almost all } 0 \ni \; w = \sum_{i\in I} w_i$$

e.g., if $I = \{1, \ldots, n\}$, then

$$w \in W_1 \oplus \ldots \oplus W_n \text{ means } \exists! \, w_i \in W_i \ni \; w = w_1 + \ldots + w_n$$

Warning: It may not exist.

# §3 Lec 3: Apr 2, 2021

## §3.1 Direct Sums (Cont'd)

**Definition 3.1** (Independent Subspace) — Let $V$ be a vector space over $F, W_i \subseteq V$, $i \in I$ subspaces. We say the $W_i$, $i \in I$, are independent if whenever $w_i \in W_i$, $i \in I$, almost all $w_i = 0$, satisfy $\sum w_i = 0$, then $w_i = 0 \, \forall i \in I$.

**Theorem 3.2**

Let $V$ be a vector space over $F$, $W_i \subseteq V$, $i \in I$ subspaces, $W \subseteq V$ a subspace. Then the following are equivalent:

1. $W = \bigoplus_{i \in I} W_i$
2. $W = \sum_{i \in I} W_i$ and $\forall i$
$$W_i \cap \sum_{j \in I \setminus \{i\}} W_j = 0 := \{0\}$$
3. $W = \sum_{i \in I} W_i$ and the $W_i$, $i \in I$, are independent.

*Proof.* 1) $\implies$ 2) Suppose $W = \bigoplus_{i \in I} W_i$. Certainly, $W = \sum_{i \in I} W_i$. Fix $i$ and suppose that

$$\exists x \in W_i \cap \sum_{j \in I \setminus \{i\}} W_j$$

By definition, $\exists w_i \in W_i$, $w_j \in W_j$, $j \in I \setminus \{i\}$ almost all 0 satisfying

$$w_i = x = \sum_{j \neq i} w_j$$

So

$$0_V = 0_W = w_i - \sum_{j \neq i} w_j$$

But

$$0_W = \sum_I 0_{W_k} \qquad 0_{W_k} = 0_V \, \forall k \in I$$

By uniqueness of 1), $w_i = 0$ so $x = 0$.
2) $\implies$ 3) Let $w_i \in W_i$, $i \in I$, almost all zero satisfy

$$\sum_{i \in I} w_i = 0$$

Suppose that $w_k \neq 0$. Then

$$w_k = -\sum_{i \in I \setminus \{k\}} w_i \in W_k \cap \sum_{i \neq k} w_i = 0,$$

a contradiction. So $w_i = 0 \, \forall i$
3) $\implies$ 1) Suppose $v \in \sum_{i \in I} W_i$ and $\exists w_i, w_i' \in W_i$, $i \in I$, almost all 0 $\ni$

$$\sum_{i \in I} w_i = v = \sum_{i \in I} w_i'$$

Then $\sum_{i\in I}(w_i - w_i') = 0$, $w_i - w_i' \in W_i\, \forall i$. So

$$w_i - w_i' = 0, \text{ i.e., } w_i = w_i' \quad \forall i$$

and the $w_i's$ are unique. □

Warning: 2) DOES NOT SAY $W_i \cap W_j = 0$ if $i \neq j$. This is too weak. It says $W_i \cap \sum_{j\neq i} W_j = 0$.

**Corollary 3.3**

Let $V$ be a vector space over $F, W_i \subseteq V$, $i \in I$ subspaces. Suppose $I = I_1 \cup I_2$ with $I_1 \cap I_2 = \emptyset$ and $V = \bigoplus_{i\in I} W_i$. Set

$$W_{I_1} = \bigoplus_{i\in I_1} W_i \quad \text{and} \quad W_{I_2} = \bigoplus_{j\in I_2} W_j$$

Then

$$V = W_{I_1} \oplus W_{I_2}$$

*Proof.* Left as exercise – Homework. □

Notation: Let $V$ be a vector space over $F$, $v \in V$. Set

$$Fv := \{\alpha v | \alpha \in F\} = \text{Span}(v)$$

if $v \neq 0$, then $Fv$ is the line containing $v$, i.e., $Fv$ is the one dimensional vector space over $F$ with basis $\{v\}$.

**Example 3.4**

Let $V$ be a vector space over $F$.

1. If $\emptyset \neq S \subseteq V$ is a subset, then
$$\sum_{v\in S} Fv = \text{Span}(S)$$
the span of $S$. So
$$\text{Span } S = \{\text{all finite linear combos of vectors in } S\}$$

2. If $\emptyset \neq S$ is linearly indep. (i.e. meaning every finite nonempty subset of $S$ is linearly indep.), then
$$\text{Span}(S) = \bigoplus_{s\in S} Fs$$

**Example 3.5** (Cont'd from above) 3. If $S$ is a basis for $V$, then $V = \bigoplus_{s\in S} Fs$.

4. If $\exists$ a finite set $S \subseteq V \ni V = \text{Span}(S)$, then $V = \sum_{s\in S} Fs$ and $\exists$ a subset $\mathscr{B} \subseteq S$ that is a basis for $V$, i.e., $V$ is a finite dimensional vector space over $F$ and $\dim V = \dim_F V = |\mathscr{B}|$ is indep. of basis $\mathscr{B}$ for $V$.

5. Let $V$ be a vector space over $F, W_1, W_2 \subseteq V$ finite dimensional subspaces. Then $W_1 + W_2$, $W_1 \cap W_2$ are finite dimensional vector space over $F$ and
$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$$
So
$$W_1 + W_2 = W_1 \oplus W_2 \iff W_1 \cap W_2 = \emptyset$$
Warning: be very careful if you wish to generalize this.

**Definition 3.6** (Complementary Subspace) — Let $V$ be a finite dimensional vector space over $F, W \subseteq V$ a subspace if
$$V = W \oplus W', \quad W' \subseteq V \text{ a subspace}$$
We call $W'$ a complementary subspace of $W$ in $V$.

**Example 3.7**
Let $\mathscr{B}_0$ be a basis of $W$. Extend $\mathscr{B}_0$ to a basis $\mathscr{B}$ for $V$ (even works if $V$ is not finite dimensional). Then
$$W' = \bigoplus_{\mathscr{B}\setminus\mathscr{B}_0} Fv \text{ is a complement of } W \text{ in } V$$
*Note*: $W'$ is not the unique complement of $W$ in $V$ – counter-example?

Consequences: Let $V$ be a finite dimensional vector space over $F, W_1, \ldots, W_n \subseteq V$ subspaces, $W_i \neq 0\,\forall i$. Then the following are equivalent

1. $V = W_1 \oplus \ldots \oplus W_n$.
2. If $\mathscr{B}_i$ is a basis (resp., ordered basis) for $W_i\,\forall i$, then $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_n$ is a basis (resp. ordered) – with obvious order – for $V$.

*Proof.* Left as exercise (good one)! □

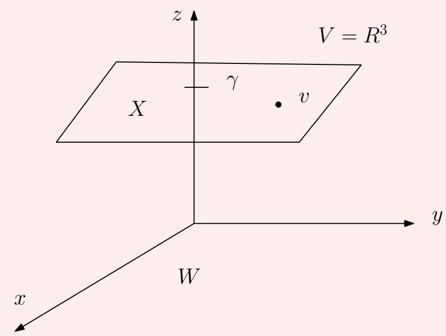
Notation: Let $V$ be a vector space over $F$, $\mathscr{B}$ a basis for $V$, $x \in V$. Then, $\exists! \alpha_v \in F$, $v \in \mathscr{B}$, almost all $\alpha_v = 0$ (i.e., all but finitely many) s.t. $x = \sum_{\mathscr{B}} \alpha_v v$. Given $x \in V$,
$$x = \sum_{v\in\mathscr{B}} \alpha_v v$$
to mean $\alpha_v$ is the unique complement of $x$ on $v$ and hence $\alpha_v = 0$ for almost all $v \in \mathscr{B}$.

## §3.2 Quotient Spaces

Idea: Given a surjective map $f : X \to Y$ and "nice", can we use properties of $Y$ to obtain properties of $X$?

**Example 3.8**

Let $V = \mathbb{R}^3$, $W = X - Y$ plane. Let $X =$ plane parallel to $W$ intersecting the z-axis at $\gamma$.




So

$$\begin{aligned} X &= \{(\alpha, \beta, \gamma) | \alpha, \beta \in \mathbb{R}\} \\ &= \{(\alpha, \beta, 0) + (0, 0, \gamma) | \alpha, \beta \in \mathbb{R}\} \\ &= W + \gamma \underbrace{e_3}_{(0,0,1)} \end{aligned}$$

<u>*Note*</u>: $X$ is a vector space over $\mathbb{R}$ $\iff$ $\gamma = 0$ $\iff$ $W = X$ (need $0_V$). Let $v \in X$. So $v = (x, y, \gamma)$ some $x, y \in \mathbb{R}$. So

$$\begin{aligned} W + v &:= \left\{ \underbrace{(\alpha, \beta, 0)}_{\text{arbitrary}} + \underbrace{(x, y, \gamma)}_{\text{fixed}} \,|\, \alpha, \beta \in \mathbb{R} \right\} \\ &= \{(\alpha + x, \beta + y, \gamma) \,|\, \alpha, \beta \in \mathbb{R}\} \\ &= W + \gamma_{e_3} \end{aligned}$$

It follows if $v, v' \in V$, then

$$W + v = W + v' \implies v - v' \in W$$

Conversely, if $v, v' \in V$ with $X = W + v$, then

$$v' \in X \implies v' = w + v \text{ some } w \in W$$

hence

$$v' - v \in W$$

So for arbitrary $v, v' \in V$, we have the <u>conclusion</u> $W + v = W + v' \iff v - v' \in W$. We can also write $W + v$ as $v + W$.

# §4 Lec 4: Apr 5, 2021

## §4.1 Quotient Spaces (Cont'd)

Recall from the last example of the last lecture, we have

$$V = \bigcup_{v \in V} W + v$$

If $v, v' \in V$, then

$$0 \neq v'' \in (W + v) \cap (W + v')$$

means

$$W + v - W + v'' = W + v'$$

This means either $W + v = W + v'$ or $W + v \cap W + v' = \emptyset$, i.e., planes parallel to the xy-plane partition $V$ into a disjoint unions of planes.

Let

$$S := \{W + v | \, v \in V\}$$

the set of these planes. We make $S$ into a vector space over $\mathbb{R}$ as follows: $\forall v, v' \in V$, $\forall \alpha \in \mathbb{R}$ define

$$\begin{aligned}(W + v) + (W + v') &:= W + (v + v') \\ \alpha \cdot (W + v) &:= W + \alpha v\end{aligned}$$

We must check these two operations are well-defined and we set

$$0_S := W$$

Then $(W + v) + W = W + v = W + (W + v)$ make $S$ into a vector space over $\mathbb{R}$.
If $v \in V$ let $\gamma_v^1 =$ the $k^{\text{th}}$ component of $v$. Define

$$S \to \{(0, 0, \gamma) | \, \gamma \in \mathbb{R}\} \to \mathbb{R}$$

by

$$W + v \mapsto (0, 0, \gamma_v) \mapsto \gamma$$

both maps are bijection and, in fact, linear isomorphism. So

$$S \cong \{(0, 0, \gamma) | \, \gamma \in \mathbb{R}\} \cong \mathbb{R}$$

<u>Note</u>: $\dim V = 3$, $\dim W = 2$, $\dim S = 1$ and we also have a linear transformation

$$V \to S \text{ by } (\alpha, \beta, \gamma) \mapsto W + \gamma_{e_3}$$

a surjection.
We can now generalize this.
<u>Construction</u>: Let $V$ be a vector space over $F$, $W \subseteq V$ a subspace. Define $\equiv \mod W$ called <u>congruent mod $W$ on $V$</u> as follows: if $x, y \in V$, then

$$x \equiv y \mod W \iff x - y \in W \iff \exists w \in W \ni x = w + y$$

Then, for all $x, y, z \in V$, $\equiv \mod W$ satisfies

1. $x \equiv x \mod W$
2. $x \equiv y \mod W \implies y \equiv x \mod W$
3. $x \equiv y \mod W$ and $y \equiv z \mod W \implies x \equiv z \mod W$

We can conclude that $\equiv \mod W$ is an equivalence relation on $V$.

Notation: For $x \in V$, $W \subseteq V$, let

$$\overline{x} \coloneqq \{y \in V | \, y \equiv x \mod W\}$$

We can also write $\overline{x}$ as $[x]_W$ if $W$ is not understood. Also, $\overline{x} \subseteq V$ is a subset and not an element of $V$ called a coset of $V$ by $W$. We have

$$\begin{aligned}\overline{x} &= \{y \in V | \, y \equiv x \mod W\} \\ &= \{y \in V | \, y = w + x \text{ for some } w \in W\} \\ &= \{w + x | \, w \in W\} = W + x = x + W\end{aligned}$$

**Example 4.1**

$\overline{0}_V = W + 0_V = W$.

Note: $W + x$ translates every element of $W$ by $x$. By 2), 3) of $\equiv \mod W$, we have check

$$y \in \overline{x} = W + x \iff x \in \overline{y} = W + y$$

and

$$x \equiv y \mod W \iff \overline{x} = \overline{y} \iff W + x = W + y$$

and check

$$\overline{x} \cap \overline{y} = \emptyset \iff (W + x) \cap (W + y) = \emptyset \iff x \not\equiv y \mod W$$

This means the $W + x$ partition $V$, i.e.,

$$V = \bigcup_V (W + x) \text{ with } (W + x) \cap (W + y) = \emptyset \text{ if } \overline{x} = (W + x) \neq (W + y) = \overline{y}$$

Let

$$\overline{V} \coloneqq V/W \coloneqq \{\overline{x} | \, x \in V\} = \{W + x | \, x \in V\}$$

a collection of subsets of $V$.

# §5 Lec 5: Apr 7, 2021

## §5.1 Quotient Spaces (Cont'd)

Suppose we have $W \subseteq V$ a subspace. For $x, y, z, v \in V$

$$\begin{aligned} x &\equiv y \mod W \\ z &\equiv v \mod W \end{aligned} \tag{+}$$

Then

$$(x+z)-(y+v) = \underbrace{(x-y)}_{\in W} + \underbrace{(z-v)}_{\in W} \in W$$

So

$$x+z \mod y+v \mod W$$

and if $\alpha \in F$

$$\alpha x - \alpha y = \alpha(x-y) \in W \quad \forall x, y \in V$$

So

$$\alpha x \equiv \alpha y \mod W$$

Therefore, $\overline{V} = V/W$. If (+) holds, then for all $x, y, z, v \in V$ and $\alpha \in F$, we have

$$\begin{aligned} \overline{x+z} = \overline{y+v} \in \overline{V} \\ \overline{\alpha x} = \overline{\alpha y} \in \overline{V} \end{aligned}$$

Notice $\overline{V} = V/W$ satisfies all the axioms of a vector space with $0_{\overline{V}} = \overline{0_V} = \{y \in V | \, y \equiv 0 \mod W\} = W + 0_V = W$.

We call $\overline{V} = V/W$ the Quotient Space of $V$ by $W$.
We also have a map

$$- : V \to \overline{V} = V/W \text{ by } x \mapsto \overline{x} = W + x$$

which satisfies

$$\alpha v + v' \overset{-}{\mapsto} \overline{\alpha u + v'} = \alpha\overline{v} + \overline{v'}$$

for all $v, v' \in V$ and $\alpha \in F$. Then

$$\begin{aligned} \dim V &= \dim \ker^{-} \\ \dim V &= \dim W + \dim V/W \\ \dim V/W &= \dim V - \dim W \end{aligned}$$

which is called the <u>codimension</u> of $W$ in $V$.

**Proposition 5.1**
Let $V$ be a vector space over $F$, $W \subseteq V$ a subspace, $\overline{V} = V/W$. Let $\mathscr{B}_0$ be a basis for $W$ and

$$\mathscr{B}_1 = \{v_i | \, i \in I, \, v_i - v_j \notin W \text{ if } i \neq j\}$$

where $\overline{v_i} \neq \overline{v_j}$ if $i \neq j$ or $w + v_i \neq w + v_j$ if $i \neq j$.

Let

$$\mathscr{C} = \{\overline{v_i} = W + v_i | \, i \in I, \, v_i \in \mathscr{B}_1\}$$

If $\mathscr{C}$ is a basis for $\overline{V} = V/W$, then $\mathscr{B}_0 \cup \mathscr{B}_1$ is a basis for $V$ (compare with the proof of the Dimension Theorem).

*Proof.* Hw 2 # 3. □

## §5.2 Linear Transformation

A review of linear of linear transformation can be found here.
Now, we consider

$$GL_nF := \{A \in \mathbb{M}_nF |\ \det A \neq 0\}$$

The elements in $GL_nF$ in the ring $\mathbb{M}_nF$ are those having a multiplicative inverse. If $R$ is a commutative ring, determinants are still as before but

$$\begin{aligned} GL_nR &:= \{A \in \mathbb{M}_nR |\ \det A \text{ is a unit in } R\} \\ &= \left\{A \in \mathbb{M}_nR |\ A^{-1} \text{ exists}\right\} \end{aligned}$$

**Example 5.2**

Let $V$ be a vector space over $F, W \subseteq V$ a subspace. Recall

$$\overline{V} = V/W = \{\overline{v} = W + v |\ v \in V\}$$

a vector space over $F$ s.t. for all $v_1, v_2 \in F$ and $\alpha \in F$

$$\begin{aligned} 0_{\overline{V}} &= \overline{0_V} = W \\ \overline{v_1} + \overline{v_2} &= \overline{v_1 + v_2} \\ \alpha\overline{v_1} &= \overline{\alpha v_1} \end{aligned}$$

Then

$$- : V \to V/W = \overline{V} \text{ by } v \mapsto \overline{v} = W + v$$

is an epimorphism with $\ker^- = W$.

Recall from 115A(H) that the most important theorem about linear transformation is Universal Property of Vector Spaces. As a result, we can deduce the following corollary

**Corollary 5.3**

Let $V, W$ be vector space over $F$ with bases $\mathscr{B}, \mathscr{C}$ respectively. Suppose there exists a bijection $f : \mathscr{B} \to \mathscr{C}$, i.e., $|\mathscr{B}| = |\mathscr{C}|$. Then $V \cong W$.

*Proof.* There exists a unique $T : V \to W \ni T\big|_{\mathscr{B}} = f$. $T$ is monic by the Monomorphism Theorem ($T$ takes linearly indep. sets to linearly indep. sets iff it's monic) and is onto as $W = \text{Span}(\mathscr{C}) = \text{Span}\left(f(\mathscr{B})\right)$. □

# §6 Lec 6: Apr 9, 2021

## §6.1 Linear Transformation (Cont'd)

**Theorem 6.1**

Let $T: V \to W$ be linear. Then $\exists X \subseteq V$ a subspace s.t.

$$V = \ker T \oplus X \text{ with } X \cong \text{im } T$$

*Proof.* Let $\mathscr{B}_0$ be a basis for $\ker T$. Extend $\mathscr{B}_0$ to a basis $\mathscr{B}$ for $V$ by the Extension Theorem. Let $\mathscr{B}_1 = \mathscr{B} \setminus \mathscr{B}_0$, so $\mathscr{B} = \mathscr{B}_0 \vee \mathscr{B}_1$ ($\mathscr{B} = \mathscr{B}_0 \cup \mathscr{B}_1$ and $\mathscr{B}_0 \cap \mathscr{B}_1 = \emptyset$) and let

$$X = \bigoplus_{\mathscr{B}_1} Fv$$

As $\ker T = \bigoplus_{\mathscr{B}_0} Fv$, we have

$$V = \ker T \oplus X$$

and we have to show

$$X \cong \text{im } T$$

**Claim 6.1.** $Tv$, $v \in \mathscr{B}_1$ are linearly indep.

In particular, $Tv \neq Tv'$ if $v, v' \in \mathscr{B}_1$ and $v \neq v'$. Suppose

$$\sum_{v \in \mathscr{B}} \alpha_v Tv = 0_W, \quad \alpha_v \in F \text{ almost all } \alpha_v = 0$$

Then

$$0_W = T\left(\sum_{v \in \mathscr{B}_1} \alpha_v v\right), \quad \text{i.e. } \sum_{\mathscr{B}_1} \alpha_v v \in \ker T$$

Hence

$$\sum_{\mathscr{B}_1} \alpha_v v = \sum_{\mathscr{B}_0} \beta_v v \in \ker T \text{ almost all } \beta_v \in F = 0$$

As $\sum_{\mathscr{B}_1} \alpha_v v - \sum_{\mathscr{B}_0} \beta_v v = 0$ and $\mathscr{B} = \mathscr{B}_0 \cup \mathscr{B}_1$ is linearly indep., $\alpha_v = 0 \, \forall v$. This proves the above claim.

Let $\mathscr{C} = \{Tv | \, v \in \mathscr{B}_1\}$. By the claim

$$\mathscr{B}_1 \to \mathscr{C} \text{ by } v \mapsto Tv \text{ is } 1-1$$

and onto as $\mathscr{C}$ is linearly indep. Lastly, we must show $\mathscr{C}$ spans im $T$. Let $w \in \text{im } T$. Then $\exists x \in V \ni Tx = w$. Then

$$\begin{aligned} w = Tx &= T\left(\sum_{\mathscr{B}_0} \alpha_v v\right) + T\left(\sum_{\mathscr{B}_1} \alpha_v v\right) \\ &= \sum_{\mathscr{B}_0} \alpha_v Tv + \sum_{\mathscr{B}_1} \alpha_v Tv = \sum_{\mathscr{B}_1} \alpha_v Tv \end{aligned}$$

lies in span $\mathscr{C}$ as needed. □

**Remark 6.2.** Note that the proof is essentially the same as the proof of the Dimension Theorem.

**Corollary 6.3** (Dimension Theorem)
If $V$ is a finite dimensional vector space over $F$, $T : V \to W$ linear then

$$\dim V = \dim \ker T + \dim \ \text{im}\ T$$

**Corollary 6.4**
If $V$ is a finite dimensional vector space over $F$, $W \subseteq V$ a subspace, then

$$\dim V = \dim W + \dim V/W$$

*Proof.* $- : V \to V/W$ by $v \mapsto \overline{v} = W + v$ is an epi. ☐

Important Construction: Set

$$\begin{gathered} T : V \to Z \text{ be linear} \\ W = \ker T \\ \overline{V} = V/W \\ - : V \to V/W \text{ by } v \mapsto \overline{v} = W + v \text{ linear} \end{gathered}$$

$\forall x, y \in V$ we have

$$\overline{x} = \overline{y} \in \overline{V} \iff x \equiv y \mod W \iff x - y \in W \iff T(x-y) = 0_Z$$

i.e., when $W = \ker T$

$$\overline{x} = \overline{y} \iff Tx = Ty \tag{*}$$

This means

$$\overline{T} : \overline{V} \to Z \text{ defined by } W + v = \overline{v} \mapsto Tv$$

is well-defined, i.e., via function, since if $\overline{x} = \overline{y}$, then $\overline{T}(\overline{x}) := Tx = Ty =: \overline{T}(\overline{y})$. From (*),

$$\overline{x} = \overline{y} \iff \overline{T}(\overline{x}) = T(x) = T(y) =: \overline{T}(\overline{y})$$

so

$$\overline{T} : \overline{V} \to Z \text{ is also injective}$$

As $\overline{T}$ is linear, let $\alpha \in F$, $x, y \in V$, then

$$\begin{aligned} \overline{T}\left(\alpha\overline{x} + \overline{y}\right) &= \overline{T}\left(\overline{\alpha x + y}\right) = T(\alpha x + y) \\ &= \alpha T x + T y = \alpha \overline{T}(\overline{x}) + \overline{T}(\overline{y}) \end{aligned}$$

as needed. Therefore,

$$\overline{T} : \overline{V} \to Z \text{ by } \overline{x} \mapsto T(x)$$

is a monomorphism, so induces an isomorphism onto im $\overline{T}$ and we recall im $\overline{T} =$ im $T$, so

$$\overline{V} \cong \text{im}\ \overline{T} = \text{im}\ T$$

and we have a commutative diagram

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "-"'] & Z \\ V/\ker T = \overline{V} \arrow[ur, "\overline{T}"'] & \end{tikzcd}$$

This can also be written as

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, "-"'] & Z \\
V/\ker T = \overline{V} \arrow[r, "\overline{T}"] & im\, T \arrow[u, bend right, "\text{inclusion map}"']
\end{tikzcd}$$

<u>Consequence</u>: Any linear transformation $T : V \to Z$ induces an isomorphism

$$\overline{T} : V/\ker T \to \text{im } T \text{ by } \overline{v} = \ker T + v \mapsto Tv$$

This is called the First Isomorphism Theorem. We also have

$$V = \ker T \oplus X \text{ with } X \subseteq V \text{ and } X \cong \text{im } T \cong V/\ker T$$

This means that all images of linear transformations from $V$ are determined, up to isomorphism, by $V$ and its subspaces. It also means, if $V$ is a finite dimensional vector space over $F$, we can try prove things by induction.

## §6.2 Projections

<u>Motivation</u>: Let $m < n$ in $\mathbb{Z}^+$ and

$$\pi : \mathbb{R}^n \to \mathbb{R}^n \text{ by } (\alpha_1, \dots, \alpha_n) \mapsto (\alpha_1, \dots, \alpha_n, 0, \dots, 0)$$

a linear operator onto $\bigoplus_{i=1}^m \Gamma e_i$ where $e_i = \left(0, \dots, \underbrace{1}_{i^m}, \dots, 0\right)$.

**Definition 6.5** (T-invariant) — Let $T : V \to V$ be linear, $W \subseteq V$ a subspace. We say $W$ is $T$-invariant if $T(W) \subseteq V$ if this is the case, then the restriction $T\big|_W$ of $T$ can be viewed as a linear operator

$$T\big|_W : W \to W$$

**Example 6.6**

Let $T : V \to V$ be linear.

1. $\ker T$ and im $T$ are $T$-invariant.
2. Let $\lambda \in F$ be an eigenvalue of $T$, i.e., $\exists 0 \neq v \in V \ni Tv = \lambda v$, then any subspace of the eigenspace
$$E_T(\lambda) := \{v \in V \,|\, Tv = \lambda v\}$$
is $T$-invariant as $T\big|_{E_T(\lambda)} = \lambda 1_{E_T(\lambda)}$

**Remark 6.7.** Let $V$ be a finite dimensional vector space over $F$, $T: V \to V$ linear. Suppose that

$$V = W_1 \oplus \ldots \oplus W_n$$

with each $W_i$ $T$-invariant, $i = 1, \ldots, n$ and $\mathscr{B}_i$ an ordered basis for $W_i$, $i = 1, \ldots, n$. Let $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_n$ be a basis of $V$ ordered in the obvious way.
Then the matrix representation of $T$ in the $\mathscr{B}$ basis is

$$[T]_{\mathscr{B}} = \begin{pmatrix} \left[T\big|_{W_1}\right]_{\mathscr{B}_1} & & 0 \\ & \ddots & \\ 0 & & \left[T\big|_{W_n}\right]_{\mathscr{B}_n} \end{pmatrix}$$

**Example 6.8**

Suppose that $T: V \to V$ is diagonalizable, i.e., there exists a basis $\mathscr{B}$ of eigenvectors of $T$ for $V$. Then, $T: V \to V$,

$$V = \bigoplus E_T(\lambda_i)$$

each $E_T(\lambda_i)$ is $T$-invariant.

$$T\big|_{E_T(\lambda_i)} = \lambda_i 1_{E_T(\lambda_i)}$$

<u>Goal</u>: Let $V$ be a finite dimensional vector space over $F$, $n = \dim V$, $T: V \to V$ linear. Then $\exists W_1, \ldots, W_m \subseteq V$ all $T$-invariant subspaces with $m = m(T)$ with each $W_i$ being as small as possible with $V = W_1 \oplus \ldots \oplus W_m$. This is the theory of canonical forms.
<u>Recall</u>: If $V$ is a finite dimensional vector space over $F$, $T: V \to V$ linear, $\mathscr{B}$ an ordered basis for $V$, then the matrix representation $[T]_{\mathscr{B}}$ is only unique up to <u>similarity</u>, i.e., if $\mathscr{C}$ is an another ordered basis

$$[T]_{\mathscr{C}} = P\,[T]_{\mathscr{B}}\,P^{-1}$$

where $P = [1_V]_{\mathscr{B},\mathscr{C}} \in GL_n F$, the change of basis matrix $\mathscr{B} \to \mathscr{C}$.

**Definition 6.9** (Projection) — Let $V$ be a vector space over $F$, $P: V \to V$ linear. We call $P$ a <u>projection</u> if $P^2 = P \circ P = P$.

**Example 6.10**

1. $P = 0_V$ or $1_V: V \to V$, $V$ is a vector space over $F$.
2. An orthogonal projection in 115A.
3. If $P$ is a projection, so is $1_V - P$.

If $T: V \to V$ is linear, then

$$V = \ker T \oplus X \text{ with } X \cong \operatorname{im} T$$

**Lemma 6.11**

Let $P : V \to V$ be a projection. Then

$$V = \ker P \oplus \text{im } P$$

Moreover, if $v \in \text{im } P$, then

$$Pv = v$$

i.e.

$$P\big|_{\text{im } P} : \text{im } P \to \text{im } P \text{ is } 1_{\text{im } P}$$

In particular, if $V$ is a finite dimensional vector space over $F, \mathscr{B}_1$ an ordered basis for $\ker P$, $\mathscr{B}_2$ an ordered basis for im $P$, then $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2$ is an ordered basis for $V$ and

$$[P]_{\mathscr{B}} = \begin{pmatrix} [0]_{\mathscr{B}_1} & 0 \\ 0 & [1_{\text{im } P}]_{\mathscr{B}_2} \end{pmatrix} = \begin{pmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

*Proof.* Let $v \in V$, then $v - Pv \in \ker P$, since

$$P(v - Pv) = Pv - P^2 v = Pv - Pv = 0$$

Hence

$$v = (v - Pv) + Pv \in \ker P + \text{im } P$$

$\ker P \cap \text{im } P = 0$ and $P\big|_{\text{im } P} = 1_{\text{im } P}$. Let $v \in$ im P. By definition, $Pw = v$ for some $w \in V$. Therefore,

$$Pv = PPw = Pw = v$$

Hence

$$P\big|_{\text{im } P} = 1_{\text{im } P}$$

If $v \in \ker P \cap \text{im } P$, then

$$v = Pv = 0$$

□

# §7 Lec 7: Apr 12, 2021

## §7.1 Projection (Cont'd)

**Lemma 7.1**

Let $V$ be a vector space over $F$, $W, X \subseteq V$ subspaces. Suppose

$$V = W \oplus X$$

Then $\exists!\, P : V \to V$ a projection satisfying

$$\begin{aligned} W &= \ker P \\ X &= \text{im } P \end{aligned} \tag{*}$$

We say such a $P$ is the projection along $W$ onto $X$.

*Proof.* Existence: Let $v \in V$. Then

$$\exists! w \in W,\, x \in X \ni v = w + x$$

Define

$$P : V \to V \text{ by } v \mapsto x$$

To show $P^2 = P$ , we suppose $v \in V$ satisfies $v = w + x$, for unique $w \in W$, $x \in X$. Then

check $P$ is linear and well defined

$$Pv = Pw + Px = Px = 1_X x = x$$

so

$$P^2 v = Px = x = Pv \qquad \forall v \in V$$

hence $P^2 = P$.
Uniqueness: Any $P$ satisfying (*) takes a basis for $W$ to 0 and fix a basis of $X$. Therefore, $P$ is unique by the UPVS. □

> **Remark 7.2.** Compare the above to the case that $V$ is an inner product space over $F$, $W \subseteq V$ is a finite dimensional subspace and $P : V \to V$ by $v \mapsto v_W$, the orthogonal projection of $P$ onto $W$.

**Proposition 7.3**

Let $V$ be a vector space over $F$, $W, X \subseteq V$ subspaces s.t. $V = W \oplus X$, $P : V \to V$ the projection along $W$ onto $X$, and $T : V \to V$ linear. Then the following are equivalent:

1. $W$ and $X$ are both $T-$invariant.
2. $PT = TP$.

*Proof.* 2) $\implies$ 1) : $W$ is $T$-invariant: We have $W = \ker P$, so if $w \in W$, $Pw = 0$. Hence

$$PTw = TPw = T0 = 0$$

$Tw \in \ker P = W$ so $W$ is $T$-invariant.
$X$ is $T$-invariant, $X = \text{im } P$, $P|_X = 1_X$. So if $x \in X$

$$Tx = TPx = PTx \in \text{im } P = X$$

So $X$ is $T$-invariant.
1) $\implies$ 2) Let $v \in V$. Then $\exists!\, w \in W$, $x \in X$ s.t.

$$v = w + x$$

As $P|_X = 1_X$ and $P|_W = 0$, so $Pv = Px$. By 1), $W$ and $X$ are $T$-invariant, so

$$\begin{aligned} PTv &= PT(w+x) = PTw + PTx \\ &= 0 + Tx = TPx = TPw + TPx = TPv \end{aligned}$$

for all $v \in V$ and $PT = TP$. □

**Remark 7.4.** One can easily generalize from the case

$$V = W_1 \oplus W_2$$

that we did to the case

$$V = W_1 \oplus \ldots \oplus W_n$$

by induction on $n$ as

$$V = W_i \oplus \left( W_1 \oplus \ldots \oplus \underbrace{\hat{W_i}}_{\text{omit}} \oplus \ldots \oplus W_n \right)$$

<u>Construction</u>: Let

$$V = W_1 \oplus \ldots \oplus W_n$$

as above. Define

$$P_{W_i} : V \to V$$

to be the projection along $W_1 \oplus \ldots \oplus \hat{W_i} \oplus \ldots \oplus W_n$, i.e.

$$\ker P_{W_i} = W_1 \oplus \ldots \oplus \hat{W_i} \oplus \ldots \oplus W_n$$

and onto $W_i = \operatorname{im} P_{W_i}$ as in the above Proposition. Then we have

a) Each $P_{W_i}$ is linear (and a projection).

b) $\ker P_{W_i} = W_1 \oplus \ldots \oplus \hat{W_i} \oplus \ldots \oplus W_n$.

c) $W_i$ is $P_{W_i}$-invariant and $P_{W_i}\big|_{W_i} = 1_{W_i}$. In particular, $\operatorname{im} P_{W_i} = W_i$.

d) $P_{W_i} P_{W_j} = \delta_{ij} P_{W_i}$ where

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

e) $1_V = P_{W_1} + \ldots + P_{W_n}$.

Moreover, if $T : V \to V$ is linear and each $W_i$ is $T$-invariant, then

$$TP_{W_i} = P_{W_i} T, \quad i = 1, \ldots, n$$

Hence

$$\begin{aligned} T = T1_V = T\left(P_{W_1} + \ldots P_{W_n}\right) &= TP_{W_1} + \ldots + TP_{W_n} \\ &= P_{W_1} T + \ldots + P_{W_n} T \end{aligned}$$

i.e., $1_V T = T 1_V$. This implies

$$T\big|_{W_i} : W_i \to W_i$$

is given by

$$T\big|_{W_i} = TP_{W_i}\big|_{W_i}$$

or $T$ is determined by what it does to each $W_i$.

**Remark 7.5.** Compare this to the case that $T$ is diagonalizable and the $W_i$ are the eigenspaces.

**Question 7.1.** Let $V$ be a real or complex finite dimensional inner product space, $T : V \to V$ hermitian. What can you replace $\oplus$ by? What if $V$ is a complex finite dimensional inner product space and $T : V \to V$ is normal.

**Exercise 7.1.** Suppose $V$ is a vector space over $F$, $P_1, \ldots, P_n : V \to V$ linear and satisfy

i) $P_i - P_j = \delta_{ij} P_i$, $i = 1, \ldots, n$

ii) $1_V = P_1 + \ldots + P_n$

iii) $W_i = \text{im } P_i$, $i = 1, \ldots, n$

Then

$$\begin{aligned} V &= W_1 \oplus \ldots \oplus W_n \\ P_i &= P_{W_i} \qquad i = 1, \ldots, n \end{aligned}$$

## §7.2 Dual Spaces

**Question 7.2.** Let $V = \mathbb{R}^3$, $v \in V$. What is the first question that we should ask about $v$?

Motivation/Construction: Let $V$ be a vector space over $F$, $\mathscr{B}$ a basis for $V$. Fix $v_0 \in \mathscr{B}$. By the UPVS, $\exists! \, f_{v_0} : V \to F$ linear satisfying

$$f_{vv_0}(v) = \begin{cases} 1 & \text{if } v_0 = v \\ 0 & \text{if } v_0 \neq v \end{cases} = \delta_{v,v_0} \quad \forall v \in \mathscr{B}$$

**Example 7.6**

Let $\mathscr{E}_n = \{e_1, \ldots, e_n\}$ be the standard basis for $\mathbb{R}^n$ and in the above $e_1 = v_0 \ldots$ Then

$$f_{e_1} : \mathbb{R}^n \to \mathbb{R} \text{ satisfies}$$

If $v = (\alpha_1, \ldots, \alpha_n)$ in $\mathbb{R}^n$

$$v = \sum_{i=1}^n \alpha_i e_i$$

so

$$\begin{aligned} f_{e_1}(v) &= f_{e_1}\left(\sum_{i=1}^n \alpha_i e_i\right) \\ &= \sum_{i=1}^n \alpha_i f_{e_1}(e_i) = \sum_{i=1}^n \alpha_i \delta_{ii} = \alpha_1 \end{aligned}$$

this first coordinate of $v$.

Notation: If $A \subseteq B$ are sets, we write $A < B$ if $A \neq B$.
As $v_0 \neq 0$,

$$0 < \text{im } f_{v_0} \subseteq F \text{ is a subspace}$$

Notice $\dim_F F = 1$, so $\dim \text{im } f_{v_0} \leq \dim F = 1$ and

$$\dim \text{im } f_{v_0} = 1, \quad \text{i.e. im } f_0 = F$$

So $f_{v_0} : V \to F$ is a surjective linear transformation. Since this is true for all $v_0 \in \mathscr{B}$, for each $v \in \mathscr{B}$, $\exists!\, f_v : V \to F$ s.t.

$$f_v(v') = \delta_{v,v'} = \begin{cases} 1 & \text{if } v = v' \\ 0 & \text{if } v \neq v' \end{cases} \quad \forall v' \in \mathscr{B}$$

Now suppose that $x \in V$, then

$$\exists!\, \alpha_v \in F,\, v \in \mathscr{B}, \text{ almost all 0 s.t. } x = \sum_{\mathscr{B}} \alpha_v v$$

Hence

$$\begin{aligned} f_{v_0}(x) &= f_{v_0}\left( \sum_{v \in \mathscr{B}} \alpha_v v \right) = \sum_{\mathscr{B}} \alpha_v f_{v_0}(v) \\ &= \sum_{\mathscr{B}} \alpha_v \delta_{v,v_0} = \alpha_{v_0} \end{aligned}$$

**Example 7.7**

$\mathscr{B} = \mathscr{E}_n$ standard basis for $\mathbb{R}^n$

$$f_{e_i}(e_j) = \delta_{e_i,e_j} = \delta_{i,j} = \begin{cases} 1 & \text{if } e_i = e_j \\ 0 & \text{if } e_i \neq e_j \end{cases}$$

Then if $v = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n = V$. Then

$$f_{e_i}(v) = f_{e_i}(\alpha_1, \dots, \alpha_n) = \alpha_i$$

So we observe in the above that if $x \in V$, then

$$x = \sum_{\mathscr{B}} f_v(x) v$$

We call $f_v$ the coordinate function on $v$ relative to $\mathscr{B}$.

**Example 7.8**

Let $V$ be a finite dimensional inner product space over $\mathbb{R}$, $\mathscr{B} = \{v_1, \dots, v_n\}$ an orthonormal basis. Then if $x = \sum_{\mathscr{B}} \alpha_i v_i$, then

$$\alpha_i = \langle x, v_i \rangle$$

Take

$$\begin{aligned} \langle x, v_i \rangle &= \langle \sum \alpha_j v_j, v_i \rangle = \sum \alpha_j \langle v_j, v_i \rangle \\ &= \sum \alpha_j \delta_{ij} \|v_i\|^2 = \sum \alpha_j \delta_{ij} = \alpha_i \end{aligned}$$

i.e. the linear map

$$f_{v_i} := \langle , v_i \rangle : V \to \mathbb{R} \text{ by } x \mapsto \langle x, v_i \rangle$$

is the coordinate function on vectors relative to $\mathscr{B}$.

**Definition 7.9** (Dual Space) — Let $V$ be a vector space over $F$. A linear transformation $f : V \to F$ is called a linear functional. Set

$$V^* := L(V,F) := \{f : V \to F | f \text{ is linear}\}$$

is called the dual space of $V$.

**Proposition 7.10**

Let $V, W$ be a vector space over $F$. Then

$$L(V,W) := \{T : V \to W | T \text{ linear}\}$$

is a vector space over $F$. Moreover, if $V, W$ are finite dimensional vector spaces over $F$

$$\dim L(V,W) = \dim V \dim W$$

In particular, if $V$ is a finite dimensional vector space over $F$, then so is $V^*$ and

$$\dim V = \dim V^*$$

so

$$V \cong V^*$$

*Proof.* 115A. □

**Example 7.11**

Let $V$ be a vector space over $F$. Then the following are linear functionals

1. $0 : V \to F$
2. Let $0 \neq v_0 \in V$ then $\{v_0\}$ is a basis for $Fv_0$. Therefore, $\{v_0\}$ extends to a basis $\mathscr{B}$ for $V$. Let $fv_0 \in V^*$ be the coordinate function for $V$ on $v_0$ relative to $\mathscr{B}$. Then $fv_0 \in \mathscr{B}^* := \{fv | v \in \mathscr{B}\}$.

# §8 Lec 8: Apr 14, 2021

## §8.1 Dual Spaces (Cont'd)

**Example 8.1** (Cont'd from Lec 7) 3. trace: $\mathbb{M}_n F \to F$ by

$$A \mapsto \sum_{i=1}^{n} A_{ii}$$

4. $\alpha < \beta \in \mathbb{R}$, then

$$I : C[\alpha, \beta] \to \mathbb{R} \text{ by } f \mapsto \int_{\alpha}^{\beta} f$$

5. Fix $\gamma \in [\alpha, \beta]$, $\alpha < \beta \in \mathbb{R}$. Then the evaluation map at $\gamma$

$$e_\gamma : C[\alpha, \beta] \to \mathbb{R} \text{ by } f \mapsto f(\gamma)$$

**Lemma 8.2**
Let $V$ be a vector space over $F$, $\mathscr{B}$ a basis for $V$,

$$\mathscr{B}^* := \{fv_0 : V \to F | \text{ coordinate function on } v_0 \text{ relative to } \mathscr{B}\}$$

so

$$fv_0(v) = \delta_{v_0,v} \qquad \forall v \in \mathscr{B}$$

the set of coordinate functions relative to $\mathscr{B}$. Then $\mathscr{B}^* \subseteq V^*$ is linearly indep.

*Proof.* Suppose

$$0 = 0_{V^*} = \sum_{v \in \mathscr{B}} \beta v f v, \quad \beta v \in F \text{ almost all } 0$$

We need to show $\beta v = 0 \, \forall v \in \mathscr{B}$. Evaluation at $v_0 \in \mathscr{B}$ yields

$$\begin{aligned} 0 = 0_{V^*}(v_0) &= \left( \sum_{\mathscr{B}} \beta v f v \right) (v_0) = \sum \beta v f v(v_0) \\ &= \sum_{\mathscr{B}} \beta v f_{v,v_0} = \beta v_0 \end{aligned}$$

So $\beta v = 0 \, \forall v \in \mathscr{B}$ and the lemma follows. □

**Corollary 8.3**
Let $V$ be a vector space over $F$ with basis $\mathscr{B}$. Then the linear transformation

$$D_{\mathscr{B}} : V \to V^* \text{ induced by } \mathscr{B} \to \mathscr{B}^* \text{ by } v \mapsto fv$$

is a monomorphism.
In particular, if $V$ is a finite dimensional vector space over $F$, then $\mathscr{B}^*$ is a basis for $V^*$ and

$$D_{\mathscr{B}} : V \to V^* \text{ is an isomorphism}$$

*Proof.* By the Monomorphism Theorem, $D_{\mathscr{B}}$ is monic in view of he lemma if $V$ is a finite dimensional vectors space over $F$, then

$$\dim V = \dim V^*$$

so $V \cong V^*$ by the Isomorphism Theorem. □

**Remark 8.4.**

1. If $V = \mathbb{R}_f^\infty := \{(\alpha_1, \alpha_2, \ldots) \,|\alpha_i \in \mathbb{R} \text{ almost all } 0\}$, then by HW1 # 4,

$$D\mathscr{E}_\infty : V \to V^* \text{ is not an isomorphism}$$

2. $D_{\mathscr{B}} : V \to V^*$ in the corollary depends on $\mathscr{B}$. There exists no monomorphism $V \to V^*$ that does not depend on a choice of basis. However, there exists a "nice" monomorphism, i.e., defined independent of basis.

$$L : V \to (V^*)^* =: V^{**}$$

$V^{**}$ is called the double dual of $V$. We now construct it.

**Lemma 8.5**

Let $V$ be a vector space over $F$, $v \in V$. Then

$$L_v : V^* \to F \text{ by } f \mapsto L_v(f) := f(v)$$

the evaluation map at $v$ is linear, i.e.

$$L_v \in V^{**}$$

*Proof.* For all $f, g \in V^*$, $\alpha \in F$

$$L_v(\alpha f + g) = (\alpha f + g)(v) = \alpha f(v) + g(v) = \alpha L_v f + L_v g$$

□

**Theorem 8.6**

The "natural" map

$$L : V \to V^{**} \text{ by } v \mapsto L(v) := L_v$$

is a monomorphism.

*Proof.* $L$ is linear: Let $v, w \in V$, $\alpha \in F$. Then for all $f \in V^*$, as $V^{**} = (V^*)^*$

$$\begin{aligned} L(\alpha v + w)(f) &= L_{\alpha v + w}(f) = f(\alpha v + w) \\ &= \alpha f(v) + f(w) = \alpha L_v f + L_w f = (\alpha L_v + L_w)(f) \\ &= (\alpha L(v) + L(w))(f) \end{aligned}$$

So

$$L(\alpha v + w) = \alpha L(v) + L(w)$$

$L$ is monic. Suppose $v \neq 0$. To show $L_v = L(v) \neq 0$. By example 2,

$$\exists 0 \neq f \in V^* \ni f(v) \neq 0$$

So

$$L_v f = f(v) \neq 0$$

so $L_v = L(v) \neq 0$ and $L$ is monic. □

**Corollary 8.7**
If $V$ is a finite dimensional vector space over $F$, then $L : V \to V^{**}$ is a natural isomorphism.

*Proof.* $\dim V = \dim V^* = \dim V^{**}$ and the Isomorphism Theorem. □

Identification: Let $V$ be a finite dimensional vector space over $F$. Then $\forall v, w \in V$

1. $v = w \iff L_v = L_w$
2. $\forall f \in V^*\, f(v) = f(w) \iff L_v f = L_w f$

Moreover, if $W$ is also a finite dimensional vector space over $F$, then if $T : V \to W$ is linear, $\exists! \tilde{T} : V^{**} \to W^{**}$ linear and if $\tilde{T} : V^{**} \to W^{**}$ $\exists! T : V \to W$ linear. In other words, $V$ and $V^{**}$ can be identified by

$$v \leftrightarrow L_v$$

because

$$L_v(f) = f(v) \qquad \forall v \in V \quad \forall f \in V^*$$

Construction: Let $V$ be a finite dimensional vector space over $F$ with basis $\mathscr{B} = \{v_1, \dots, v_n\}$. Then

$$\mathscr{B}^* := \{f_1, \dots, f_n\}$$

defined by

$$f_i(v_j) = \delta_{ij} \qquad \forall i, j$$

i.e., $f_i$ is the coordinate function on $v_i$ relative to $\mathscr{B}$. Since

$$L_{v_i}(f_j) = f_j(v_i) = \delta_{ij} \qquad \forall i, j$$

$L_{v_i} \in V^{**}$

$$\mathscr{B}^{**} := \{L_{v_1}, \dots, L_{v_n}\}$$

is the dual basis of $\mathscr{B}^*$ for $V^{**}$. So we have if $x = \sum_{i=1}^n \alpha_i v_i \in V$, $g = \sum_{i=1}^n \beta_i f_i \in V^*$.

$$x = \sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n f_i(x) v_i$$
$$g = \sum_{i=1}^n \beta_i f_i = \sum_{i=1}^n L_{v_i}(g) f_i = \sum_{i=1}^n g(v_i) f_i$$

i.e.

$$x = \sum_{i=1}^n f_i(x) v_i \qquad \forall x \in V$$
$$g = \sum_{i=1}^n g(v_i) f_i \qquad \forall g \in V^*$$

Motivation: Let $V$ be an inner product space over $\mathbb{R}$, $\emptyset \neq S \subseteq V$ a subset. What is $S^\perp$?
*Note*: $\forall v \in V$, $\langle , v\rangle : V \to \mathbb{R}$ by $x \mapsto \langle x, v\rangle$ is a linear functional. To generalize this to an arbitrary vector space over $F$, we define the following.

**Definition 8.8** (Annihilator) — Let $V$ be a vector space over $F$, $\emptyset \neq S \subseteq V$ a subset. Define the annihilator of $S$ to be

$$\begin{aligned} S^\circ &:= \{f \in V^* \mid f(x) = 0\, \forall x \in S\} \\ &= \{f \in V^* \mid f|_S = 0\} \subseteq V^* \end{aligned}$$

**Remark 8.9.** Many people write $\langle v, f\rangle$ for $f(v)$ in the above even though $f \notin v$.

# §9 Lec 9: Apr 16, 2021

## §9.1 Dual Spaces (Cont'd)

**Lemma 9.1**

Let $V$ be a vector space over $F$, $\emptyset \neq S \subseteq V$ a subset. Then

1. $S^\circ \subseteq V^*$ is a subspace.
2. If $V$ is a finite dimensional vector space over $F$ and we identify $V$ as $V^{**}$ (by $v \leftrightarrow L_v$), then $S \subseteq S^{\circ\circ} := (S^\circ)^\circ$.

*Proof.* 1. For all $f, g \in S^\circ$, $\alpha \in F$, we have

$$(\alpha f + g)(x) = \alpha f(x) + g(x) = 0 \qquad \forall x \in S$$

Hence $\alpha f + g \in S^\circ$ and $S^\circ \subseteq V^*$ is a subspace.

2. Let $x \in S$. Then $\forall f \in S^\circ$, we have

$$0 = f(x) = L_x f, \quad \text{so } L_x \in (S^\circ)^\circ = S^{\circ\circ}$$

□

**Theorem 9.2**

Let $V$ be a finite dimensional vector space over $F$, $S \subseteq V$ a subspace. Then

$$\dim V = \dim S + \dim S^\circ$$

*Proof.* Let $\mathscr{B}_0 = \{v_1, \ldots, v_k\}$ be a basis for $S$. Extend this to

$$\begin{aligned} \mathscr{B} &= \{v_1, \ldots, v_n\} \text{ a basis for } V \\ \mathscr{B}_0 &= \{f_1, \ldots, f_n\} \text{ the dual basis of } \mathscr{B} \end{aligned}$$

**Claim 9.1.** $\mathscr{C} := \{f_{k+1}, \ldots, f_n\}$ is a basis for $S^\circ$.

If we show this, the theorem follows. Let $f \in S^\circ$. Then

$$\begin{aligned} f &= \sum_{i=1}^n L_{v_i}(f) f_i = \sum_{i=1}^n f(v_i) f_i \\ &= \sum_{i=1}^k f(v_i) f_i + \sum_{i=k+1}^n f(v_i) f_i = \sum_{i=k+1}^n f(v_i) f_i \end{aligned}$$

lies in span $\mathscr{C}$ so $\mathscr{C}$ spans. As $\mathscr{C} \subseteq \mathscr{B}^*$ which is linearly indep., so is $\mathscr{C}$. This proves the claim. □

**Corollary 9.3**

Let $V$ be a finite dimensional vector space over $F$, $S \subseteq V$ a subspace. Then $S = S^{\circ\circ}$.

*Proof.* As $S \subseteq S^{\circ\circ}$, it suffices to show $\dim S = \dim S^{\circ\circ}$. By the theorem, we have

$$\begin{aligned} \dim V &= \dim S + \dim S^{\circ} \\ \dim V^* &= \dim S^{\circ} + \dim S^{\circ\circ} \end{aligned}$$

where $\dim V = \dim V^*$. So $\dim S = \dim S^{\circ\circ}$. □

**Remark 9.4.** If $V$ is an inner product space over $\mathbb{R}$, compare all this to $\emptyset \neq S \subseteq V$ a subset and $S^{\perp}, S^{\perp\perp}$.

## §9.2 The Transpose

Construction: Fix $T : V \to W$ linear. For every $S : W \to X$, we have a composition

$$S \circ T : V \to X \text{ is linear}$$

So $T :\to W$ linear induces a map

$$T^{\star} : L(W, X) \to L(V, X)$$

by

$$S \mapsto S \circ T$$

**Proposition 9.5**

Let $V, W, X$ be vector spaces over $F$, $T : V \to W$ linear. Then

$$T^{\star} : L(W, X) \to L(V, X)$$

is linear.

*Proof.* Let $S_1, S_2 \in L(W, X)$, $\alpha \in F$. Then

$$\begin{aligned} T^{\star}(\alpha S_1 + S_2) &= (\alpha S_1 + S_2) \circ T \\ &= \alpha S_1 \circ T + S_2 \circ T = \alpha T^{\star} S_1 + T^{\star} S_2 \end{aligned}$$

□

**Corollary 9.6**

Let $T : V \to W$ be linear. Then

$$T^* : W^* \to V^* \text{ by } f \mapsto f \circ T$$

is linear.

*Proof.* Let $X = F$ in the proposition. □

**Definition 9.7** (Transpose) — Let $T : V \to W$ be linear. The linear map $T^{\star} : W^* \to V^*$ in the corollary is called the transpose of $T$ and denoted by $T^{\top}$.

Note: The transpose "turns thing around"

$$V \xrightarrow{T} W$$
$$V^* \xleftarrow{T^\top} W^*$$

**Lemma 9.8**

Let $T : V \to W$ be linear. Then

$$\ker T^\top = (\text{im T})^\circ \in W^*$$

*Proof.* $g \in \ker T^\top \iff T^\top g = 0 \iff (T^\top g)(v) = 0\ \forall v \in V \iff (g \circ T)(v) = 0\ \forall v \in V \iff g(Tv) = 0\ \forall v \in V \iff g \in (\text{im } T)^\circ$. □

**Theorem 9.9**

Let $V, W$ be finite dimensional vector space over $F$, $T : V \to W$ linear. Then

$$\dim \text{im } T = \dim \text{im } T^\top$$

*Proof.* Consider:

$$\dim W^* = \dim \ker T^\top + \dim \text{im } T^\top$$
$$\dim W = \dim \text{im } T + \dim(\text{im } T)^\circ$$

Notice that $\dim W^* = \dim W$. By the lemma, $\dim \text{im } T = \dim \text{im } T^\top$. □

Computation: Let $V, W$ be finite dimensional vector space over $F$.

$$\mathscr{B},\ \mathscr{B}^* \text{ ordered dual bases for } V,\ V^*$$
$$\mathscr{C},\ \mathscr{C}^* \text{ ordered dual bases for } W,\ W^*$$

Suppose

$$\mathscr{B} = \{v_1, \dots, v_n\}, \quad \mathscr{B}^* = \{f_1, \dots, f_n\}$$
$$f_i(v_j) = \delta_{ij} \quad \forall i, j$$

So

$$\mathscr{C} = \{w_1, \dots, w_n\}, \quad \mathscr{C}^* = \{g_1, \dots, g_n\}$$
$$g_i(w_j) = \delta_{ij} \quad \forall i, j$$

Let

$$A = [T]_{\mathscr{B},\mathscr{C}}, \qquad B = [T^\top]_{\mathscr{C}^*,\mathscr{B}^*}$$

be the matrix representation of $T, T^\top$ in the ordered bases $\mathscr{B}, \mathscr{C}$ and $\mathscr{C}^*, \mathscr{C}^*$ respectively. By definition of $A$ and $B$, we have

$$Tv_k = \sum_{i=1}^m A_{ik} w_i \qquad k = 1, \dots, n$$

$$T^\top g_j = \sum_{i=1}^n B_{ij} f_i \qquad j = 1, \dots, m$$

So

$$B_{kj} = A_{jk} \qquad \forall j, k$$

So we just proved...

**Theorem 9.10**

Let $V, W$ be finite dimensional vector space over $F$, $T : V \to W$ linear, $\mathscr{B}, \mathscr{B}^*$ ordered dual bases for $V, V^*$ and $\mathscr{C}, \mathscr{C}^*$ ordered dual bases for $W, W^*$. Then

$$[T^\top]_{\mathscr{C}^*, \mathscr{B}^*} = \left([T]_{\mathscr{B}, \mathscr{C}}\right)^\top$$

**Definition 9.11** (Row/Column Rank) — Let $A \in F^{m\times n}$. The row (column) rank of $A$ is the dimension of the span of the rows (columns) of $A$.

We know if $A \in F^{m\times n}$, we can view

$$A : F^{n\times 1} \to F^{m\times 1} \text{ by } v \mapsto A \cdot v$$

a linear transformation and the matrix representation of $A$ is

$$A = [A]_{\mathscr{E}_{n,1}, \mathscr{E}_{m,1}}$$

where $\mathscr{E}_{n,1}$, $\mathscr{E}_{m,1}$ are the standard bases for $F^{n\times 1}$ and $F^{m\times 1}$ respectively.

**Corollary 9.12**

Let $A \in F^{m\times n}$. Then

$$\text{row rank } A = \text{column rank } A$$

and we call this common number the rank of $A$.

## §9.3 Polynomials

**Definition 9.13** (Polynomial Division) — Let $f, g \in F[t]$, $f \neq 0$. We say that $f$ divides $g \in F[t]$ write $f|g$ if $\exists h \in F[t]$ s.t. $g = fh$, i.e. $g$ is multiple of $f$, e.g. $t + 1|t^2 - 1$.

**Lemma 9.14**

If $f|g$ and $f|h$ in $F[t]$, then $f|gk + hl$ in $F[t]$ for all $k, l \in F[t]$.

*Proof.* By definition,

$$g = fg_1, \quad h = fh_1, \quad g_1, h_1 \in F[t]$$

So

$$gk + hl = fg_1k + fh_1l = f(g_1k + h_1l)$$

in $F[t]$. □

> **Remark 9.15.** If $f|g \in F[t]$ and $0 \neq a \in F$, then $af|g$ and $f|ag$.

**Definition 9.16** (Polynomial Degree and Leading Coefficient) — Let

$$0 \neq f = at^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0 \in F[t]$$

with $a, a_0, \ldots, a_{n-1} \in F$ and $a \neq 0$. We call $n$ the degree of $f$ write $\deg f = n$ and $a$ the leading coefficient of $F$ write lead $f = a$. If $a = 1$, we say $f$ is monic.

We can define the degree of $0 \in F[t]$ to be the symbol $-\infty$ or just do not define it at all.

**Remark 9.17.** Let $f, g \in F[t] \setminus \{0\}$. Then

$$\operatorname{lead}(fg) = \operatorname{lead}(f) \cdot \operatorname{lead}(g) \neq 0 \in F$$

So

$$\deg(fg) = \deg f + \deg g$$

# §10 Lec 10: Apr 19, 2021

## §10.1 Polynomials (Cont'd)

Division Algorithm: Let $0 \neq f \in F[t]$, $g \in F[t]$. Then

$$\exists!\ q,\ r \in F[t]$$

satisfying

$$g = fq + r \quad \text{with} \quad r = 0 \quad \text{or} \quad \deg r < \deg f$$

**Definition 10.1** (Greatest Common Divisor) — Let $f, g \in F[t] \setminus \{0\}$. We say $d$ in $F[t]$ is a gcd (greatest common divisor) of $f, g$ if

i) $d$ is monic.

ii) $d|f$ and $d|g$ in $F[t]$.

iii) if $e|f$ and $e|g$ in $F[t]$, then $e|d$ in $F[t]$.

**Remark 10.2.** If a gcd of $f, g$ exists, then it is unique.

**Remark 10.3.** If $d = 1$ is a gcd of $f, g \in F[t]$, we say that $f, g$ are relatively bear.

**Remark 10.4.** Compare the above with analogous in $\mathbb{Z}$.

**Theorem 10.5**

Let $f, g \in F[t] \setminus \{0\}$. Then a gcd of $f, g$ exists and is unique write gcd(f,g) for the gcd of $f, g$. Moreover, we have an equation

$$d = fk + gl \in F[t] \text{ for some } k, l \in F[t] \qquad (\star)$$

*Proof.* The existence and $(\star)$ follow from the Euclidean Algorithm. Let $f, g \in F[t] \setminus \{0\}$. Then iteration of the Division Algorithm produces equations in $F[t]$, if $f + g \in F[t]$,

$$\begin{aligned}
g &= q_1 f + r_1 \qquad \deg r_1 < \deg f \\
f &= q_2 r_1 + r_2 \qquad \deg r_2 < \deg r_1 \\
&\vdots \\
r_{n-3} &= q_{n-1} r_{n-2} + r_{n-1} \qquad \deg r_{n-1} < \deg r_{n-2} \\
r_{n-2} &= q_n r_{n-1} + r_n \qquad \deg r_{n-1} < \deg r_n \\
r_{n-1} &= q_{n+1} + r_n
\end{aligned}$$

where $r_n$ is the remainder of least degree ($r_n \neq 0$).
This must stop in $\leq \deg f$ steps. Plugging from the bottom up and using the lemma shows

$$r_n = fk + gl \in F[t]$$

and if $e|r_1 \to e|r_2 \to \ldots \to e|r_n$ then $(\text{lead } r_n)^{-1} r_n$ is the gcd of $f$ and $g$ in $F[t]$ if $a = \text{lead } f$

$$a^{-1} r_n = a^{-1} fk + a^{-1} gl$$

□

**Definition 10.6** (Irreducible Polynomial) — $f \in F[t] \setminus F$ is called <u>irreducible</u> if there does not exist $g, h \in F[t] \ni f = gh$ with $\deg g, \deg h < \deg f$. Equivalently, if

$$f = gh \in F[t], \quad \text{then } 0 \neq g \in F \text{ or } 0 \neq h \in F$$

**Example 10.7**

If $f \in F[t]$, $\deg f = 1$, then $f$ is irreducible.

**Remark 10.8.** If $f, g \in F[t] \setminus F$ with $f$ irreducible, then either $f$ and $g$ are relatively prime or $f|g$ since only $a, af, 0 \neq a \in F$ can divide $f$.

**Lemma 10.9** (Euclid)

Let $f \in F[t]$ be irreducible and $f|gh$ in $F[t]$. Then $f|g$ or $f|h$.

*Proof.* Suppose $f \times g$ where $\times$ means does not divide. Then $f$ and $g$ are relatively prime. By the Euclidean Algorithm, there exists an equation

$$1 = fk + gl \in F[t]$$

Hence

$$h = fhk + ghl \in F[t]$$

As $f|fhk$ and $f|ghl$ in $F[t]$, $f|h$ by the lemma. $\square$

**Remark 10.10.** In $\mathbb{Z}$ the analog of an irreducible element is called a prime element.

**Remark 10.11.** Euclid's lemma is the key idea. The "correct" generalization of "prime" is the conclusion of Euclid's lemma. This generalization is profound as, in general, there is difference between the two conditions "irreducible" and "prime", although not for $\mathbb{Z}$ or $F[t]$.

We know that any positive integer is a product of positive primes unique up to order $n$. If we allow $n < 0$ such is unique up to $\pm 1$.

**Theorem 10.12** (Fundamental Theorem of Arithmetic (Polynomial Case))

Let $g \in F[t] \setminus F$. Then there exists uniquely $a \in F$, $r \in \mathbb{Z}^+$, $p_1, \dots, p_r \in F[t]$ distinct monic irreducible polynomial, $e_1, \dots, e_r \in \mathbb{Z}^+$ s.t. we have a factorization

$$g = ap_1^{e_1} \dots p_r^{e_r}$$

unique up to order.

*Proof.* (Sketch) <u>Existence</u>: We induct on $n = \deg g \geq 1$. If $g$ is irreducible, $a$, $(\text{lead } g)^{-1} g$, $a = \text{lead } g$ work. If $g$ is reducible,

$$g = fh \in F[t], \qquad 1 < \deg f, \quad \deg h < \deg g$$

By induction, $f, h$ have factorization hence we're done as $g = fh$.

Uniqueness: We induct on $n = \deg g \geq 1$. If

$$ap_1^{e_1} \dots p_r^{e_r} = g = bq_1^{f_1} \dots q_s^{f_s}$$

with $p_i, q_i$ monic irreducible, $a, b \in F$, $e_i, f_j \in \mathbb{Z}^+$ for all $i, j$, $\deg q_1 \geq 1$, so $\deg q_1 \times a$. By Euclid's lemma

$$q_i | p_j \text{ for some } j$$

Changing notation, we may assume that $j = 1$. As $p_1$ is irreducible $p_1 = q_1$ and by $(M3')$

$$g_0 \coloneqq ap_1^{e_1 - 1} p_2^{e_2} \dots p_r^{e_r} = bq_1^{f_1 - 1} q_2^{f_2} \dots q_s^{f_s}$$

As $\deg g_0 < \deg g$, induction yields

$$r = s,\ e_1 - 1 = f_1 - 1,\ e_i = f_i,\ i > 1,\ a = b = \text{lead } g_0,\ p_i = q_i\ \forall i,\ e_i = f_i\ \forall i$$

□

**Remark 10.13.** Applying the Euclidean Algorithm is relatively fast to compute, (for $f|g$ takes $\leq \deg f$ steps to get a gcd). Factoring into the irreducible is not.

# §11 Lec 11: Apr 21, 2021

## §11.1 Minimal Polynomials

We use the following theorem from 115A, Matrix Theory Theorem.

**Remark 11.1.** Let $T : V \to V$ be linear. If $f = a_n t^n + \ldots + a_1 t + a_0 \in F[t]$, we can plug $T$ in for $t$ to get

$$f(T) = a_n T^n + \ldots + a_1 T + a_0 1_V \in L(V,V)$$

More precisely

$$e_T : F[t] \to L(V,V) \text{ by } t \mapsto T$$

i.e. $f = \sum a_i t^i \mapsto f(T) = \sum a_i T^i$ is a ring homomorphism. Since we have

$$T^n = T \underbrace{\circ \ldots \circ}_{n} T, \qquad n \geq 0$$

Can we use the remark if $V$ is a finite dimensional vector space over $F$?

**Lemma 11.2**

Let $V$ be a finite dimensional vector space over $F$, $f, g, h \in F[t]$, $\mathscr{B}$ an ordered basis for $V$, $T : V \to V$ linear. Then

1. $[g(T)]_{\mathscr{B}} = g\left([T]_{\mathscr{B}}\right)$
2. If $f = gh \in F[t]$, then
$$f(T) = g(T)h(T)$$

*Proof.*
- By MTT, if $g = \sum_{i=0}^{n} a_i t^i \in F[t]$, then

$$\begin{aligned} [g(T)]_{\mathscr{B}} &= \left[\sum_{i=0}^{n} a_i T^i\right]_{\mathscr{B}} = \sum_{i=0}^{n} a_i \left[T^i\right]_{\mathscr{B}} \\ &= \sum a_i \left[T\right]_{\mathscr{B}}^i = g\left([T]_{\mathscr{B}}\right) \end{aligned}$$

- Left as exercise.

□

**Lemma 11.3**

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. Then $\exists q \in F[t] \setminus \{0\} \ni q(T) = 0$ and if $a = \text{lead } q$, then $q_0 := a^{-1} q$ is moinc and satisfies $q_0(T) = 0$

$$q \in \ker e_T := \{f \in F[t] | \, f(T) = 0\}$$

*Proof.* Let $n = \dim V$. By MTT

$$\dim L(V,V) = \dim \mathbb{M}_n F = n^2 < \infty$$

So

$$1_V, T, T^2, \ldots, T^{n^2} \in L(V,V)$$

are linearly dependent. So $\exists a_0, \dots, a_{n^2} \in F$ not all 0 s.t.

$$\sum_{i=0}^{n^2} a_i T^i = 0$$

Then $q = \sum_{i=0}^{n^2} a_i t^i$ works. $\square$

**Theorem 11.4**

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. Then $\exists! 0 \neq q_T \in F[t]$ monic called the minimal polynomial of $T$ having the following properties:

1. $q_T(T) = 0$
2. If $g \in F[t]$ satisfies $g(T) = 0$, then $q_T | g \in F[t]$. In particular, if $0 \neq g \in F[t]$ satisfies $g(T) = 0$, then $\deg g \geq \deg q_T$ and if $\deg g = \deg q_T$, then $g = (\text{lead } g) q_T$

*Proof.* By the lemma, $\exists 0 \neq q \in F[t]$ monic s.t. $q(T) = 0$. Among all such $q$, choose one with $\deg q$ minimal.

**Claim 11.1.** $q$ works.

Let $g \neq 0$ in $F[t]$ satisfy $g(T) = 0$. To show $q|g \in F[t]$. Write $g = qh + r$ in $F[t]$ with $r = 0$ or $\deg r < \deg q$. Then

$$0 = g(T) = q(T)hh(T) + r(T) = r(T)$$

If $r \neq 0$, then $r_0 = (\text{lead } r)^{-1} r$ is a monic poly satisfying $r_0(T) = 0$, $\deg r_0 < \deg q$, contradicting the minimality of $\deg q$. So $r_0 = 0$ and $q|g \in F[t]$. If $q'$ also satisfies 1) and 2), then

$$q|q' \text{ and } q'|q \in F[t] \text{ both monic so } q = q'$$

The last statement follows as if

$$h, g \in F[t], \quad g|h,\ h \neq 0, \text{ then } \deg h \geq \deg q$$ $\square$

**Corollary 11.5**

Let $V$ be a finite dimensional vector space over $F$, $\mathscr{B}$ an ordered basis for $V_1$ and $T : V \to V$ linear. Then

$$q_T = q_{[T]_{\mathscr{B}}}$$

In particular, if $A, B \in \mathbb{M}_n F$ are similar write $A \sim B$. Then

$$q_A = q_B$$

*Proof.* $q_T = q_{[T]_{\mathscr{B}}}$ by MTT and the first lemma. $\square$

Note:By the theorem, if $V$ is a finite dimensional vector space over $F$ $g \in F[t]$ $g \neq 0$, and $\deg g < \deg q_T$, then $q(T) \neq 0$.
Goal: Let $V$ be a finite dimensional vector space over $F$, $\mathscr{B}$ an ordered basis of $V$, $T : V \to V$ linear. Call

$$tI - [T]_{\mathscr{B}} \text{ the characteristics matrix of } T \text{ relative to } \mathscr{B}$$

Recall the characteristics polynomial $f_T$ of $T$ is defined to be

$$f_T := f_{[T]_{\mathscr{B}}} = \det\left(tI - [T]_{\mathscr{B}}\right) \in F[t]$$

We want to show $f_T$ satisfies the

**Theorem 11.6** (Cayley-Hamilton)
If $V$ is a finite dimensional vector space over $F$, $T : V \to V$ linear, then

$$q_T | f_T, \quad \text{hence } f_T(T) = 0$$

In particular, $\deg q_T \leq \deg f_T$.

**Remark 11.7.**
1. There exists a determinant proof of this – essentially Cramer's rule.
2. A priori we only know $\deg q_T \leq n^2$, where $n = \dim V$.
3. $f_T$ is independent of $\mathscr{B}$ depends on properties of $\det : \mathbb{M}_n F[t] \to F[t]$

$$\begin{aligned}\det(tI - A) &= \det\left(P(tI - A)P^{-1}\right) \\ &= \det\left(tI - PAP^{-1}\right)\end{aligned}$$

for each $P \in GL_n F$

**Proposition 11.8**
Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. Then $q_T$ and $f_T$ have the same roots in $F$, the eigenvalues of $T$.

*Proof.* Let $\lambda$ be a root of $q_T$. To show $\lambda$ is an eigenvalue of $T$, i.e., a root of $f_T$. As $\lambda$ is a root of $q_T$, using the Division Algorithm that

$$q_T = (t - \lambda)h \in F[t]$$

So

$$0 = q_T(T) = (T - \lambda 1_V)h(T)$$

As

$$0 \leq \deg h < \deg q_T, \quad \text{we have } h(T) \neq 0$$

Since $h(T) \neq 0 \exists 0 \neq v \in V$ s.t.

$$w = h(T)v \neq 0$$

Then

$$0 = q_T(T)v = (T - \lambda 1_V)\, h(T)v = (T - \lambda 1_V)\, w$$

So $0 \neq w \in E_T(\lambda)$ and $\lambda$ is an eigenvalue of $T$.
Conversely, suppose $\lambda$ is a root of $f_T$ so an eigenvalue of $T$. Let $0 \neq v \in E_T(\lambda)$. Then $t - \lambda \in F[t]$ satisfies $(T - \lambda)w = 0$ for all $w \in Fv$, i.e. it is the minimal poly of $T|_{Fv} : Fv \to Fv$. But $q_T(T) = 0$ on $V$ so $t - \lambda | q_T$ by the definition that $t - \lambda$ is the minimal poly of $T|_{Fv}$. $\square$

## §11.2 Algebraic Aside

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. Te minimality poly $q_T$ of $T$ is algebraically more interesting than $f_T$. Recall we have a ring homomorphism

$$e_T : F[t] \to L(V, V)$$

given by

$$\sum a_i t^i \mapsto \sum a_i T^i$$

so $e_T$ is not only a linear transformation but a ring homomorphism, i.e., it also follows that

$$(fg)(T) = f(T)g(T) \qquad \forall f, g \in F[t]$$

We know that

$$\dim_F F[t] = \infty$$

which has $\{1, t, \ldots, t^n, \ldots\}$ is a basis for $F[t]$ and

$$\dim_F L(V,V) = (\dim V)^2 < \infty$$

by MTT. So

$$0 < \ker e_T \coloneqq \{f \in F[t] | e_T f = f(T) = 0\}$$

is a vector space over $F$ and a subspace of $F[t]$. This induces a linear transformation

$$\overline{e_T} : V/\ker e_T \to \text{im } e_T = F[T]$$

which is an isomorphism. If $\overline{V} = V/\ker T$, we have

$$\begin{aligned} \overline{e_T}\left(\overline{\sum a_i t^i}\right) &= \overline{e_T\left(\sum a_i t^i\right)} = \sum \overline{a_i}\overline{T}^i \\ &= \sum a_i \overline{T}^i = \sum a_i T^i \end{aligned}$$

Check that $\overline{e}_T$ is also a ring isomorphism onto im $e_T$. By definition, if $f(T) = 0$, $f \in F[t]$, then

$$q_T | f \in F[t]$$

It follows that

$$\ker e_T = \{q_t g | g \in F[t]\} \subseteq F[t]$$

called an <u>ideal</u> in the ring $F[t]$.
The first isomorphism of rings gives rise to $\ker e_T$ whit quotient isomorphic to $F[t] \subseteq L(V,V)$. So we are at a higher level of algebra. Then this allows us to view $F[t]$ as acting on $V$, i.e. there exists a map

$$F[t] \times V \to V \tag{*}$$

by

$$\begin{gathered} f \cdot v \coloneqq f(T)v \\ q_T(T) = 0 \end{gathered}$$

This turns $V$ into what is called an $F[t]$-module, i.e., $V$ via (*) satisfies the axioms of a vector space over $F$ but the scalars $F[t]$ are now a ring rather than only a field.

# §12 Lec 12: Apr 23, 2021

## §12.1 Triangularizability

**Proposition 12.1**

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear, $W \subseteq V$ a $T$-invariant subspace. Then $T$ induces a linear transformation

$$\overline{T} : V/W \to V/W \text{ by } \overline{T}(\overline{v}) := \overline{T(v)}$$

where $\overline{v} = W + v$, $\overline{V} = V/W$ and

$$q_{\overline{T}} | q_T \in F[t]$$

*Proof.* By the hw, we need only to prove that

$$q_{\overline{T}} | q_T \in F[t]$$

But also by the hw,

$$q_T(\overline{T}) = \overline{q_T(T)}$$

As $q_T(T) = 0$,

$$0 = \overline{q_T(T)} = q_T(\overline{T})$$

so

$$q_{\overline{T}} | q_T$$

by the defining property of $q_{\overline{T}}$. □

**Definition 12.2** (Triangularizability) — Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. We say $T$ is <u>triangularizable</u> if $\exists$ an ordered basis $\mathscr{B}$ for $V$ s.t. $A = [T]_{\mathscr{B}}$ satisfies $A_{ij} = 0 \; \forall i < j$, i.e.

$$A = \begin{pmatrix} * & & 0 \\ & \ddots & \\ * & & * \end{pmatrix} \text{ is lower triangular} \tag{*}$$

<u>*Note*</u>: If $\mathscr{B} = \{v_1, \ldots, v_n\}$ in (*) and $\mathscr{C} = \{v_n, v_{n-1}, \ldots, v_1\}$, then

$$[T]_{\mathscr{C}} = \begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix} \text{ is upper triangular}$$

Hence, by Change of Basis Theorem,

$$[T]_{\mathscr{B}} \sim [T]_{\mathscr{C}}$$

**Remark 12.3.** Suppose $V$ is a finite dimensional vector space over $F$, $\dim V = n$, $T : V \to V$ linear, $\mathscr{B}$ an ordered basis for $V$, $A = [T]_{\mathscr{B}}$ is triangular (upper or lower). Then

$$f_T = (t - A_{11}) \ldots (t - A_{nn}) \in F[t]$$

and $A_{11}, \ldots, A_{nn}$ are all the eigenvalues of $T$ (not necessarily distinct) and hence roots of $q_T$.

**Definition 12.4** (Splits) — We say $g \in F[t] \setminus F$ splits in $F[t]$ if $g$ is a product of linear polys in $F[t]$, i.e.,

$$g = (\text{lead } g)(t-\alpha_1)\dots(t-\alpha_n) \in F[t]$$

**Example 12.5**

If $V$ is a finite dimensional vector space over $F$, $T : V \to V$ linear and $T$ is triangularizable, then $f_T$ splits in $F[t]$.

*Note*: $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{M}_2\mathbb{R}$ is not triangularizable as it has no eigenvalues.

**Theorem 12.6**

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. Then $T$ is triangularizable if and only if $q_T$ splits in $F[t]$.

*Proof.* " $\implies$ " We induct on $n = \dim V$.
$n = 1$ : It's obvious.
$n > 1$ : We proceed by induction: let $\lambda$ be a root of $q_T$ in $F$ ($q_T$ splits in $F[t]$). Then $\lambda$ is a root of $q_T$ hence an eigenvalue of $T$. Let $0 \neq v_n \in E_T(\lambda)$, so $W = Fv_n$ is $T$-invariant. By the Proposition, $T$ induces a linear map

$$\overline{T} : V/W \to V/W \text{ by } \overline{v} \mapsto \overline{T(v)}$$

and

$$q_{\overline{T}} | q_T \in F[t]$$

We also know that

$$W = \ker(- : V \to V/W) \text{ by } v \mapsto \overline{v}$$

and

$$\dim V/W = \dim V - \dim W = n-1$$

as $- : v \to \overline{v}$ is epic. Since $q_T$ splits in $F[t]$ and $q_{\overline{T}} | q_T$ in $F[t]$, $q_{\overline{T}}$ also splits in $F[t]$ by Fundamental Theorem of Algebra. Thus, by induction,

$$\exists v_1, \dots, v_{n-i} \in V \ni \mathscr{C} = \{\overline{v_1}, \dots, \overline{v}_{n-1}\}$$

is an ordered basis for $\overline{V} = V/W$ with $A = [\overline{T}]_{\mathscr{C}}$ is lower triangular, i.e., $A_{ij} = 0$ if $i < j \leq n-1$. Thus

$$\overline{T}\overline{v}_j = \sum_{i=j}^{n-1} A_{ij}\overline{v}_i, \qquad 1 \leq j \leq n-1$$

hence

$$0 = \overline{T}\overline{v}_j - \sum_{i=j}^{n-1} A_{ij}\overline{v}_i = \overline{Tv_j - \sum_{i=j}^{n-1} A_{ij}v_i}$$

$1 \leq j \leq n-1$ in $\overline{V} = V/W$. Therefore,

$$Tv_j - \sum_{i=j}^{n-1} A_{ij}v_i \in \ker^- = W = Fv_n$$

by definition as $W = \ker^- : V \to V/W$.

In particular, $\exists A_{nj} \in F$, $1 \le j \le n-1$ satisfying

$$Tv_j - \sum_{i=j}^{n-1} A_{ij} v_i = A_{nj} v_n$$

So

$$Tv_j = \sum_{i=j}^{n} A_{ij} v_n \qquad 1 \le j \le n-1$$

By choice, $A_{ij} = 0$, $i < j \le n-1$ and

$$Tv_n = \lambda v_n$$

By hw 2 # 3, $\mathscr{B} = \{v_1, \dots, v_n\}$ is an ordered basis for $V$ and

$$[T]_{\mathscr{B}} = \begin{pmatrix} & [T]_{\mathscr{C}} & 0 \\ & & \vdots \\ & & 0 \\ A_{n1} \dots A_{n,n-1} & & \lambda \end{pmatrix}$$

which is lower triangular, as needed. " $\implies$ " Let $\mathscr{B} = \{v_1, \dots, v_n\}$ be an ordered basis for $V$. $A = [T]_{\mathscr{B}}$ is lower triangular. Then

$$f_T = \prod_{i=1}^{n} (t - A_{ii}) \text{ splits in } F[t]$$

$A_{11}, \dots, A_{nn}$ are the (not necessarily distinct) eigenvalues of $T$ and hence roots of $q_T$.
Let $\lambda_i = A_{ii}$, $i = 1, \dots, n$. We have

$$\begin{aligned} Tv_j &= \sum_{i=1}^{n} A_{ij} v_i = \lambda_j v_j + \sum_{i=j+1}^{n} A_{ij} v_i, \qquad 1 \le j \le n-1 \\ Tv_n &= \lambda_n v_n \end{aligned}$$

So

$$(T - \lambda_j 1_V) v_j = \sum_{i=j+1}^{n} A_{ij} v_i \in \operatorname{Span}(v_{j+1}, \dots, v_n) \quad \forall 1 \le j \le n-1 \tag{*}$$

Now

$$(T - \lambda_n 1_V) v_n = 0$$

So

$$(T - \lambda_n 1_V) v_{n-1} \in \operatorname{Span}(v_n) \text{ by } (*)$$

This implies

$$(T - \lambda_n 1_V)(T - \lambda_{n-1} 1_V) v_{n-1} = 0$$

By induction, we may assume that

$$(T - \lambda_n 1_V) \dots (T - \lambda_j 1_V) v_j = 0$$

So by (*),

$$(T - \lambda_n 1_V) \dots (T - \lambda_j 1_V)(T - \lambda_{j-1} 1_V) v_{j-1} = 0$$

Therefore,

$$f_T(T) v_i = (T - \lambda_n 1_V) \dots (T - \lambda_i 1_V) v_i = 0$$

for $i = 1, \dots, n$. As $\mathscr{B}$ is a basis for $V$, $f_T(T) = 0$. Thus $q_T | f_T \in F[t]$. In particular, $q_T$ splits in $F[t]$. $\square$

**Corollary 12.7**

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ a triangularizable linear operator. Then

$$q_T | f_T \in F[t]$$

In particular,

$$f_T(T) = 0$$

**Definition 12.8** (Algebraically Closed) — A field $F$ is called algebraically closed if every $f \in F[t] \setminus F$ splits in $F[t]$. Equivalently, $f \in F[t] \setminus F$ has a root in $F$.

**Corollary 12.9** (Cayley-Hamilton – Special Case)

Let $F$ be algebraically closed, $V$ a finite dimensional vector space over $F$, $T : V \to V$ linear. Then

1. $T$ is triangularizable.
2. $q_T | f_T$
3. $f_T(T) = 0$

**Theorem 12.10** (Fundamental Theorem of Algebra)

(FTA) $\mathbb{C}$ is algebraically closed.

*Proof.* It's assumed (proven in 132 – Complex Analysis or 110C – Algebra). □

# §13 Lec 13: Apr 26, 2021

## §13.1 Triangularizability (Cont'd)

**Remark 13.1.** Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear, $\mathscr{B}$ an ordered basis for $V$, $A = [T]_{\mathscr{B}}$. So $q_A = q_T$ and $f_A = f_T$.

Let $n = \dim V$. Given a field $F$, $\exists \tilde{F}$ an algebraically closed field satisfying $F \subseteq \tilde{F}$ is a subfield. Then

$$A \in \mathbb{M}_n F \subseteq \mathbb{M}_n \tilde{F}$$

So by the corollary,

$$f_A(A)v = 0 \qquad \forall v \in \tilde{F}^{n\times 1}$$

where we view $A : \tilde{F}^{n\times 1} \to \tilde{F}^{n\times 1}$ linear. Then

$$f_A(A)v = 0 \qquad \forall v \in F^{n\times 1} \subseteq \tilde{F}^{n\times 1}$$

viewing

$$A : F^{n\times 1} \to F^{n\times 1} \text{ linear}$$

Thus,

$$f_A(A) = 0$$

Hence $f_T(T) = 0$ and $q_T = q_A | f_A = f_T$. So $q_T | f_T$ in $F[t]$. Thus, if we knew such an $\tilde{F}$ exists in general, we would have proven the Cayley-Hamilton Theorem in general, i.e., if $V$ is a finite dimensional vector space over $F$ and $T : V \to V$ linear, then

$$q_T | f_T \in F[t]$$
$$f_T(T) = 0$$

This is, in fact, true (and proven in Math 110C). Of course, assuming FTA, this proves Cayley-Hamilton for all fields $F \subseteq \mathbb{C}$.

**Remark 13.2.** The symmetric matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{M}_2\mathbb{F}_2 \text{ and } \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \in \mathbb{M}_2\mathbb{F}_5$$

are both triangularizable, but not diagonalizable.

## §13.2 Primary Decomposition

Algebraic Motivation: Let $f \in F[t] \setminus F$ be monic. Write

$$f = p_1^{e_1} \dots p_r^{e_r},\ p_1, \dots, p_r \text{ distinct monic}$$

irreducible polys in $F[t]$, $e_i > 0\, \forall i$. Set

$$q = \frac{f}{p_i^{e_i}} = p_1^{e_1} \dots p_i^{e_i} \dots p_r^{e_r}$$

Then $p_i, q_i$ are relatively prime so there exists an equation

$$1 = p_i^{e_i} k_i + q_i g_i \in F[t], \quad i = 1, \dots, n \tag{*}$$

if we plug a linear operator $T : V \to V$ into (*), we get

$$1_V = p_i^{e_i}(T) k_1(T) + q_i(T) g_i(T) \qquad \forall i$$

Linear Algebra Motivation: Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. Suppose

$$V = W_1 \oplus W_2, \quad W_1, W_2 \subseteq V \text{ subspaces}$$

with $W_1, W_2$ both $T$-invariant.
Let $\mathscr{B}_i$ be an ordered basis for $W_i$, $i = 1, 2$ and $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2$ an ordered basis for $V$. Then

$$[T]_{\mathscr{B}} = \begin{pmatrix} [T|_{W_1}]_{\mathscr{B}_1} & 0 \\ 0 & [T|_{W_2}]_{\mathscr{B}_2} \end{pmatrix}$$

Let $P_{W_i} : V \to V$ be the projection onto $W_i$ along $W_j$, $j \neq i$. Then we know

$$\begin{aligned} 1_V &= P_{W_1} + P_{W_2} \\ P_{W_i} P_{W_j} &= \delta_{ij} P_{W_j} \\ P_{W_i} T &= T P_{W_i}, \quad i = 1, 2 \\ T &= T P_{W_1} + T P_{W_2} = T|_{W_1} + T|_{W_2} \end{aligned}$$

By hw 4 # 6

$$q_T = \operatorname{lcm}\left(q_T|_{W_1}, q_T|_{W_2}\right)$$

This easily extends to more blocks.

**Lemma 13.3**

Let $f \in F[t]$, $T : V \to V$ linear. Then $\ker f(T)$ is $T$-invariant.

*Proof.* If $v \in \ker f(T)$, to show $Tv \in \ker f(T)$. But

$$f(T)\, Tv = T f(T) v = 0$$

so this is immediate. □

**Lemma 13.4**

Let $g, h \in F[t] \setminus F$ be relatively prime. Set $f = gh \in F[t]$. Suppose $T : V \to V$ is linear and $f(T) = 0$. Then

$$\ker g(T) \text{ and } \ker h(T) \text{ are } T\text{-invariant}$$

subspaces of $V$ and

$$V = \ker g(T) \oplus \ker h(T) \tag{+}$$

*Proof.* By the lemma we just proved, we need only show (+). Since $g, h$ are relatively prime, there exists equation

$$1 = gk + hl \in F[t]$$

Hence

$$1_V = g(T)k(T) = h(T)l(T)$$

as linear operators on $V$ i.e. $\forall v \in V$

$$v = g(T)k(T)v + h(T)l(T)v \tag{*}$$

Since $f(T) = 0$ we have

$$0 = f(T)k(T)v = h(T)g(T)k(T)v$$

Therefore,

$$g(T)k(T)v \in \ker h(T)$$

and

$$0 = f(T)l(T)v = g(T)h(T)l(T)v$$

so

$$h(T)l(T)v \in \ker g(T)$$

It follows by (*), $\forall v \in V$

$$v = g(T)k(T)v + h(T)l(T)v \in \ker h(T) + \ker g(T)$$

where

$$V = \ker g(T) + \ker h(T)$$

By (*), if $v \in \ker g(T) \cap \ker h(T)$, then

$$v = g(T)k(T)v + h(T)l(T)v = 0$$

Hence

$$V = \ker g(T) \oplus \ker h(T)$$

as needed. $\square$

# §14 Lec 14: Apr 28, 2021

## §14.1 Primary Decomposition (Cont'd)

**Proposition 14.1**

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear, $g, h \in F[t] \setminus F$ monic and relatively prime. Suppose that

$$q_T = gh \in F[t]$$

Then $\ker g(T)$ and $\ker h(T)$ are $T$-invariant.

$$V = \ker g(T) \oplus \ker h(T)$$

and

$$g = q_T\big|_{\ker g(T)} \text{ and } h = q_T\big|_{\ker h(T)}$$

*Proof.* By the last lemma in last lecture, we need only prove the last statement. By definition, we have

$$g(T)\big|_{\ker g(T)} = 0 \text{ and } h(T)\big|_{\ker h(T)} = 0$$

So by definition,

$$q_T\big|_{\ker q(T)}|g \text{ and } q_T\big|_{\ker h(T)}|h \in F[t]$$

As $g$ and $h$ are relatively prime, by the FTA, so are

$$q_T\big|_{\ker g(T)} \text{ and } q_T\big|_{\ker h(T)}$$

Therefore, we have

$$\begin{aligned} f &:= \operatorname{lcm}\left(q_T\big|_{\ker g(T)}, q_T\big|_{\ker h(T)}\right) \\ &= q_T\big|_{\ker q(T) q_T}\big|_{\ker h(T)} \end{aligned}$$

Since

$$\begin{gathered} V = \ker g(T) \oplus \ker h(T) \\ f(T)v = 0 \quad \forall v \in V \end{gathered}$$

Hence

$$q_T | f \in F[t]$$

By (+) and FTA

$$f | gh = q_T$$

As both $f$ and $q_T$ are monic,

$$f = q_T$$

Applying FTA again, we conclude that

$$g = q_T\big|_{\ker g(T)} \text{ and } h = q_T\big|_{\ker h(T)} \qquad \square$$

We now generalize the proposition to an important result that decomposes a finite dimensional vector space over $F$ relative to a linear operator $T : V \to V$.

**Theorem 14.2** (Primary Decomposition)
Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear, and $q_T = p_1^{e_1} \dots p_r^{e_r}$, with $p_1, \dots, p_r$ distinct monic irreducible polys in $F[t]$, $e_1, \dots, e_r \in \mathbb{Z}^+$. Then there exists a direct sum decomposition of $V$ into subspaces $W_1, \dots, W_r$

$$V = W_1 \oplus \dots \oplus W_r \tag{*}$$

satisfying all of the following:

i) Each $W_i$ is $T$-invariant, $i = 1, \dots, r$

ii) $q_T|_{W_i} = p_i^{e_i}$, $i = 1, \dots, r$

iii) $q_T = \prod_{i=1}^r p_i^{e_i} = \prod_{i=1}^r q_{T|_{W_i}}$

iv) If $\mathscr{B}_i$ is an ordered basis for $W_i$, $i = 1, \dots, r$, $\mathscr{B} = \mathscr{B}_1 \cup \dots \cup \mathscr{B}_r$ is an ordered basis for $V$ with

$$[T]_{\mathscr{B}} = \begin{pmatrix} [T|_{W_1}]_{\mathscr{B}_1} & & 0 \\ & \ddots & \\ 0 & & [T|_{W_r}] \end{pmatrix}$$

Moreover, any direct sum decomposition (*) of $V$ satisfying $i), ii), iii)$ is uniquely determined by $T$ and the $p_1, \dots, p_r$ up to order. If in addition, this is the case, then

$$W_i = \ker p_i^{e_i}(T) \qquad i = 1, \dots, r$$

*Proof.* We induct on $r$.

- $r = 1$ is immediate
- $r > 1$ By TFA, $p_1^{e_1}$ and $g = p_2^{e_2} \dots p_r^{e_r}$ are relatively prime, so by the Proposition

  $$V = W_1 \oplus V_1$$

  where

  $$\begin{aligned} W_1 &= \ker p_1^{e_1}(T) \text{ and } W_1 \text{ is } T\text{-invariant} \\ V_1 &= \ker g(T) \text{ and } V_1 \text{ is } T\text{-invariant} \\ q_{T|_{W_1}} &= p_1^{e_1} q_{T|_{V_1}} \qquad\qquad = p_2^{e_2} \dots p_r^{e_r} \end{aligned}$$

  Let

  $$T_1 = T|_{V_1} : V_1 \to V_1$$

  By induction on $r$, we may assume all of the following:

  $$\begin{aligned} V_1 &= W_2 \oplus \dots \oplus W_r \\ W_i &= \ker p_i^{e_i}(T_1) \text{ and is } T_1\text{-invariant} \\ q_{T_1|_{W_i}} &= p_i^{e_i} \text{ for } i = 2, \dots, r \end{aligned}$$

  <u>Note</u>:

  $$\ker p_i^{e_i}(T_1) \cap \sum_{\substack{j=2 \\ j \neq i}}^r \ker p_j(T_1) = 0 \quad \forall i > 0$$

  **Claim 14.1.** Let $2 \le i \le r$. Then

  $$\ker p_i^{e_i}(T) = \ker p_i^{e_i}(T_1)$$

Let $v \in \ker p_i^{e_i}(T)$, $i > 1$. So

$$p_i^{e_i}(T)v = 0$$

Hence

$$0 = \prod_{j=2}^{r} p_j^{e_j}(T)v = g(T)v,$$

i.e.,

$$v \in \ker g(T) = V_1$$

So

$$Tv = T|_{V_1}v = T_1 v$$

and

$$0 = p_i^{e_i}(T)v = p_i^{e_i}(T_1)v$$

as needed.

Let $v \in \ker p_i^{e_i}(T_1), i > 1$. By definition, $v \in V_1$, so

$$\begin{aligned} 0 &= p_i^{e_i}(T_1)v = p_i^{e_i}\left(T|_{V_1}\right)v \\ &= p_i^{e_i}(T)|_{V_1}v = p_i^{e_i}(T)v \end{aligned}$$

This proves the claim.

The existence of (*), $i), ii), iii)$ nad $W_i = \ker p_i^{e_i}(T)$, $i = 1, \ldots, r$, now follow. Moreover, $i)$ and (*) yield $iv)$.

Uniqueness: Suppose that

$$V = W_1 \oplus \ldots \oplus W_r$$

satisfies i), ii), iii). If we show

$$W_i = \ker p_i^{e_i}(T), \qquad i = 1, \ldots, r$$

the result will follow. It suffices to do the case $i = 1$. Let

$$\begin{aligned} V_1 &= W_2 \oplus \ldots \oplus W_r \\ V &= W_1 \oplus V_1 \end{aligned}$$

As each $W_i$ is $T$-invariant and $V_1$ is $T$-invariant. As before

$$p_1^{e_1} \text{ and } g = p_2^{e_2} \ldots p_r^{e_r}$$

and relatively prime by FTA. So by hw 4 # 6

$$q_T = \operatorname{lcm}\left(q_{T|_{V_1}}, q_{T|_{V_1}}\right)$$

It follows that

$$q_{T|_{V_1}} = p_2^{e_2} \ldots p_r^{e_r} = g$$

Moreover, we have an equation

$$1 = p_1^{e_1}k + gl \in F[t]$$

So

$$1_V = p_1^{e_1}(T)k(T) + g(T)l(T) \tag{+}$$

**Claim 14.2.** $W_1 = \ker p_1^{e_1}(T)$ and hence we are done.

Since

$$q_{T|_{W_1}} = p_1^{e_1}$$

We have

$$p_1^{e_1}(T)v = 0 \qquad \forall v \in W_1$$

Hence

$$W_1 \subseteq \ker p_1^{e_1}(T)$$

To finish, we must know

$$\ker p_1^{e_1}(T) \subseteq W_1$$

Let

$$v \in \ker p_1^{e_1}(T) \subseteq V = W_1 \oplus V_1$$

So $\exists! w_1 \in W_1$, $v_1 \in V_1$ s.t.

$$v = w_1 + v_1$$

Since $W_1 \subseteq \ker p_1^{e_1}(T)$,

$$p_1^{e_1}(T)W_1 = 0$$

By assumption, $p_1^{e_1}(T)v = 0$, so

$$p_1^{e_1}(T)v_1 = 0$$

As $V_1 = W_2 \oplus \ldots \oplus W_r$

$$p_i^{e_i} = q_{T|_{W_i}}, \qquad i = 2, \ldots, r \text{ by (ii)}$$

We have

$$p_2^{e_2}(T)\ldots p_r^{e_r}(T)v_1 = 0$$

Hence by (+)

$$v_1 = 1_V v_1 = p_1^{e_1}(T)k(T)v_1 + p_2^{e_2}(T)\ldots p_r^{e_r}(T)l(T)v_1 = 0$$

Therefore,

$$v = w_1 + v_1 = w_1 \in W_1$$

and it follows that $\ker p_1^{e_1}(T) \subseteq W_1$ as needed. $\square$

Recall: Let $V$ be a finite dimensional vector space over $F$, $T: V \to V$ linear is called diagonalizable if there exists an ordered basis $\mathscr{B}$ for $V$ consisting of eigenvectors of $T$. By hw 2 # 2, this is equivalent to

$$V = \bigoplus_{\lambda} E_T(\lambda)$$

# §15 Lec 15: Apr 30, 2021

## §15.1 Primary Decomposition (Cont'd)

Recall: Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear is called diagonalizable if there exists an ordered basis $\mathscr{B}$ for $V$ consisting of eigenvectors of $T$. By hw 2 # 2, this is equivalent to

$$V = \bigoplus_{\lambda} E_T(\lambda)$$

**Theorem 15.1**

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. Then $T$ is diagonalizable iff $q_T$ splits in $F[t]$ and has no repeated roots in $F$. If this is the case, then

$$q_T = \prod_{i=1}^{r} (t - \lambda_i), \quad \lambda_1, \ldots, \lambda_r \text{ the distinct roots of } q_T$$

*Proof.* " $\Longleftarrow$ " $q_T = \prod_{i=1}^{r} (t - \lambda_i)$, $\lambda_1, \ldots, \lambda_r$ the distinct roots of $q_T$. Let $V_i = \ker(T - \lambda_i 1_V) = E_T(\lambda_i)$, $i = 1, \ldots, r$. Then by the Primary Decomposition Theorem,

$$V = V_1 \oplus \ldots \oplus V_r$$

SO $T$ is diagonalizable.
" $\Longrightarrow$ " Let $\mathscr{B} = \{v_1, \ldots, v_n\}$ be an ordered basis for $V$ consisting of eigenvectors of $T$ with $\lambda_i$ the eigenvalue of $v_i$ and ordered s.t.

$$\lambda_1, \ldots, \lambda_r \text{ are the distinct eigenvalues of } T$$

For each $j$, $1 \le j \le n$, we have

$$(T - \lambda_i 1_V)\, v_j = T v_j - \lambda_i v_j = (\lambda_j - \lambda_i)\, v_j, \quad j = 1, \ldots, n$$

So

$$\prod_{i=1}^{r} (T - \lambda_i 1_V)\, v_j = 0 \quad \text{for } j = 1, \ldots, n$$

i.e.,

$$\prod_{i=1}^{r} (T - \lambda_i 1_V) \text{ vanishes on a basis for } V$$

hence vanishes on all of $V$. It follows that

$$q_T \Big| \prod_{i=1}^{r} (t - \lambda_i) \in F[t]$$

In particular, $q_T$ splits in $F[t]$ and has no multiple roots in $F$ by FTA. As every eigenvalue of $T$ is a root of $f_T$, we have

$$t - \lambda_i | q_T, \qquad i = 1, \ldots, r$$

using $f_T$ and $q_T$ have the same roots. Therefore,

$$q_T = \prod_{i=1}^{r} (t - \lambda_i) \in F[t] \qquad \square$$

## §15.2 Jordan Blocks

**Definition 15.2** (Jordan Block Matrix) — $J \in \mathbb{M}_n F$ is called a Jordan block matrix of eigenvalue $\lambda$ of size $n$ if

$$J = J_n(\lambda) := \begin{pmatrix} \lambda & & & 0 \\ 1 & \lambda & & \\ & 1 & & \\ & & \ddots & \lambda \\ 0 & & & 1 \end{pmatrix} \in \mathbb{M}_n F$$

*Note*: $f_{J_n}(\lambda) = \det(tI - J_n(\lambda)) = (t-\lambda)^n \in F[t]$, so splits with just one root of multiplicity.

**Definition 15.3** (Nilpotent) — $T : V \to V$ linear is called nilpotent if $q_T = t^m$, some $m$, i.e., $\exists M \in \mathbb{Z}^+ \ni T^M = 0$.

**Example 15.4**

$J = J_n(0)$ is nilpotent and has $q_J = t^m$ for some $m$. In fact, $q_J = t^n$ – why?
In fact, let $A \in \mathbb{M}_n F$, $A : F^{n\times 1} \to F^{n\times 1}$ linear with $A \sim N$ with

$$N = J_n(\lambda_- \lambda I_n = J_n(0)$$

Then as $N$ is nilpotent and

$$A = PNP^{-1}, \quad \text{some } P \in GL_n F,$$

we have

$$A^n = \left(PNP^{-1}\right)^n = PNP^{-1}PNP^{-1}\dots PNP^{-1} = PN^nP^{-1} = 0$$

So $A$ is nilpotent. Now $N$ is nilpotent.
If $\mathscr{S} = \{e_1, \dots, e_n\}$ is the standard basis for $F^{n\times 1}$

$$\begin{aligned} Ne_i &= e_{i+1}, \quad i \le n-1 \\ Ne_n &= 0 \\ N^2 e_i &= N - Ne_i = e_{i+2}, \quad i \le n-2 \\ &\vdots \end{aligned}$$

**Example 15.5** (Cont'd from above)

In any case, we have

$$\left.\begin{aligned} \dim \operatorname{im} N^r &= n - r \\ \dim \ker N^r &= r \end{aligned}\right\} \text{if } r \le n$$

$$\left.\begin{aligned} \dim \operatorname{im} N^r &= 0 \\ \dim \operatorname{im} \ker N^r &= n \end{aligned}\right\} \text{if } r > n$$

**Lemma 15.6**

Let $J = J_n(\lambda) \in \mathbb{M}_n F$. Then

1. $\lambda$ is the only eigenvalue of $J$.
2. $\dim E_J(\lambda) = 1$
3. $t_J = q_J = (t-\lambda)^n$
4. $f_J(J) = 0$

*Proof.* Let

$$N = J - \lambda I \in \mathbb{M}_n F$$

the characteristics matrix of $J$

$$N^{n-1} = \begin{pmatrix} 0 & & \dots & 0 \\ \vdots & & & \vdots \\ 0 & & & 0 \\ 1 & 0 & \dots & 0 \end{pmatrix} \in \mathbb{M}_n F$$

is not the zero matrix, but

$$N^n = 0$$

So

$$q_T | (t-\lambda)^n \text{ and } q_J \not| (t-\lambda)^{n-1}$$

It follows that $q_J = (t-\lambda)^n = f_J$. This shows 3) and 4). By the computation,

$$\dim \ker N = 1$$

and

$$\ker N = E_T(\lambda)$$

This gives 2) as $f_T = (t-\lambda)^n$, 1) is clear. □

**Remark 15.7.** $J_n(\lambda)$ has only a line as an eigenspace, so among triangulariazable operator away from being diagonalizable when $n \geq 1$.

**Proposition 15.8**

Let $A \in \mathbb{M}_n F$ be triangularizable. Suppose $f_A = (t-\lambda)^n$ for some $\lambda \in F$. Then $A$ is diagonalizable iff $q_A = (t-\lambda)$ iff $A = \lambda I$.

*Proof.* If $q_A = t - \lambda$, then $A = \lambda I$ as

$$F^{n\times 1} = \ker (A - \lambda I)$$

The converse is immediate. □

Computation: Let $V$ be a finite dimensional vector space over $F$, $\dim V = n$, $T : V \to V$ linear. Suppose there exists $\mathscr{B} = \{v_1, \dots, v_n\}$ an ordered basis for $V$ satisfying

$$[T]_{\mathscr{B}} = J_n(\lambda)$$

Then by definition

$$\begin{aligned}
Tv_1 &= \lambda v_1 + v_2 \quad \text{i.e. } (T - \lambda 1_V)v_1 = v_2 \\
Tv_2 &= \lambda v_2 + v_3 \quad \text{i.e. } (T - \lambda 1_V)v_2 = v_3 \\
&\vdots \\
Tv_{n-1} &= \lambda v_{n-1} + v_n \quad \text{i.e. } (T - \lambda 1_V)\, v_{n-1} = v_n \\
Tv_n &= \lambda v_n
\end{aligned} \tag{+}$$

So

$$E_\lambda(\lambda) = Fv_n$$

$v_1, \ldots, v_{n-1}$ are not eigenvectors, but do satisfy

$$\begin{aligned}
(T - \lambda 1_V)v_i &= v_{i+1} \qquad i = 1, \ldots, n-1 \\
(T - \lambda 1_V)^{n-i} v_i &= v_n \qquad \text{, an eigenvector}
\end{aligned}$$

So we can compute $v_1, \ldots, v_{n-1}$ from the eigenvalue $v_n$.

# §16 Lec 16: May 3, 2021

## §16.1 Jordan Blocks (Cont'd)

**Definition 16.1** (Sequence of Generalized Eigenvectors) — Let $T : V \to V$ be linear, $0 \neq v_n \in E_T(\lambda)$. We say $v_1, \ldots, v_n$ is an (ordered) sequence of generalized eigenvectors of eigenvalue $\lambda$ of length $n$ if (+) above holds, i.e.,

$$\begin{aligned}(T - \lambda 1_V)v_i &= v_{i+1}, \qquad i = 1, \ldots, n-1\\ (T - \lambda 1_V)v_n &= 0\end{aligned}$$

We let

$$\begin{aligned}g_n(\lambda) = g_n(v_n, \lambda) &:= \{v_1, \ldots, v_n\}\\ &= \left\{v_1, (T - \lambda 1_V)^{n-1} v_1\right\}\end{aligned}$$

be an ordered sequence of generalized eigenvectors for $T$ of length $n$ relative to $\lambda$.

*Note*: We should really write

$$g_n(v_n, \lambda, v_1, \ldots, v_{n-1})$$

**Lemma 16.2**

Let $V$ be a vector space over $F$, $T : V \to V$ linear, $0 \neq v_n \in E_T(\lambda)$, $v_1, \ldots, v_n$ an ordered sequence of generalized eigenvectors of $T$ of length $n$, $g_n(\lambda) = \{v_1, \ldots, v_n\}$. Then

1. $g_n(\lambda)$ is linearly independent.
2. If $V$ is a finite dimensional vector space over $F$, $\dim V = n$, then
   i) $g_n(\lambda)$ is an ordered basis for $V$
   ii) $[T]_{g_n(\lambda)} = J_n(\lambda)$

*Proof.* 1. We have seen that $(*)$ implies

$$\begin{aligned}(T - \lambda 1_V)^{n-i} v_i &= v_n \qquad i < n\\ (T - \lambda 1 - V)v_n &= 0\end{aligned}$$

So

$$(T - \lambda 1_V)^k v_i = 0 \qquad \forall k > n - i$$

Suppose

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0, \qquad \alpha_i \in F \text{ not all } 0$$

Choose the least $k$ s.t. $\alpha_k \neq 0$. Then

$$0 = (T - \lambda 1_V)^{n-k} (\alpha_k v_k + \ldots + \alpha_n v_n) = \alpha_k v_n$$

As $v_n \neq 0$, $\alpha_k = 0$, a contradiction.

So 1) follows and 1) $\to$ 2). $\square$

**Definition 16.3** (Jordan Canonical Form) — $A \in \mathbb{M}_n F$ is called a matrix in Jordan canonical form (JCF) if $A$ has the block form

$$A = \begin{pmatrix} J_{r_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{r_m}(\lambda_m) \end{pmatrix}$$

$\lambda_1, \ldots, \lambda_m$ not necessarily distinct.

**Definition 16.4** (Jordan Basis) — Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. An ordered basis $\mathscr{B}$ for $V$ is called a Jordan basis (if it exists) for $V$ relative to $T$ if $\mathscr{B}$ is the union

$$g_{r_1}(v_{1,r_1}, \lambda_1) \cup \ldots \cup g_{r_m}(v_{m,r_m}, \lambda_m) \tag{$\star$}$$

where $g_{r_j}(v_{j,r_j}, \lambda_j)$ is an ordered sequence of generalized eigenvectors of $T$ relative to $\lambda_j$ ending at eigenvector $v_{j,r_j}$. The $\lambda_1, \ldots, \lambda_m$ need not be distinct.

**Proposition 16.5**

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. Then $V$ has a Jordan basis relative to $T \iff T$ has a matrix representation in Jordan canonical form (JCF).

*Proof.* Let $w_i = g_{r_i}(v_{i,r_i}, \lambda_i)$ in ($\star$). The only thing to show is: $W_i$ is $T$-invariant, but this follows from our computation. ☐

<u>Conclusion</u>: Let $T : V \to V$ be linear with $V$ having a Jordan basis relative to $T$. Gathering all the Jordan blocks with the same eigenvalues together and ordering these into increasing size, we can write such a Jordan basis as follows:

$$\lambda_1, \ldots, \lambda_m \text{ the distinct eigenvalues of } T$$

$$\begin{aligned} \mathscr{B} = & \, g_{r_{11}}(v_{11}, \lambda_1) \cup \ldots \cup g_{r_{1,n_1}}(v_{1,n_1}, \lambda_1) \\ & \cup g_{r_{21}}(v_{21}, \lambda_2) \cup \ldots \cup g_{r_{2,n_2}}(v_{2,n_2}, \lambda_2) \\ & \vdots \\ & \cup g_{r_m,1}(v_{m,1}, \lambda_m) \cup \ldots \cup g_{r_m,n_m}(v_{m,r_m}, \lambda_m) \end{aligned}$$

with

$$r_{i1} \le r_{i2} \le \ldots \le r_i n_i, \qquad 1 \le i \le m$$

e.g.

$$[T]_{\mathscr{B}} = \begin{pmatrix} 1 & 0 & & & & & \\ 0 & 1 & & & & & \\ & 1 & 0 & & & & \\ & 1 & 1 & & & & \\ & & 0 & 2 & 0 & 0 & \\ & & & 1 & 2 & 0 & \\ & & & 0 & 1 & 2 & \end{pmatrix} = \begin{pmatrix} J_1(1) & & & \\ & J_1(1) & & \\ & & J_2(1) & \\ & & & J_3(2) \end{pmatrix}$$

Let

$$W_{ij} = \text{Span } g_{r_i,j}(v_{ij}, \lambda_i) \qquad \forall i, j$$

These are all $T$-invariant. We have

$$f_T = \prod_{i,j} (t-\lambda_i)^{r_{ij}}$$

and

$$\begin{aligned} q_T &= \prod_i \operatorname{lcm}\left((t-\lambda_i)^{r_{ij}} | j = 1,\dots,n_i\right) \\ &= \prod_i (t-\lambda_i)^{r_{in_i}} \end{aligned}$$

So

$$q_T | f_T \text{ and } f_T(T) = 0$$

Also

$$q_{T|W_{ij}} = f_{T|W_{ij}} = (t-\lambda_i)^{r_{ij}}$$

for all $1 \le j \le n_j$, $1 \le i \le m$. There are called the elementary divisors of $T$

$$V = W_{11} \oplus \dots \oplus W_{1,n_1} \oplus \dots \oplus W_{m1} \oplus \dots \oplus W_{mn_m}$$

Now let $P_{ij}$ be the projection onto $W_{ij}$ along

$$W_{11} \oplus \dots \oplus \underbrace{\widehat{W_{ij}}}_{\text{omit}} \oplus \dots \oplus W_{m,n_m}$$

Then

$$P_{ij}P_{kl} = \delta_{ik}\delta_{jl}P_{jl} = \begin{cases} P_{jl} \text{ if } i = k \text{ and } j = l \\ 0 \text{ otherwise} \end{cases}$$
$$1_V = P_{11} + \dots + P_{mn_m}$$
$$TP_{ij} = P_{ij}T$$
$$T = TP_{11} + \dots + TP_{mn_m} = T\big|_{W_{11}} + \dots + T\big|_{W_{mn_m}}$$

Abusing notation

$$\lambda_1,\dots,\lambda_m \text{ are the distinct eigenvalues of } T$$

Let

$$W_i = W_{i1} \oplus \dots \oplus W_{in_i} \quad i = 1,\dots,m$$

As $r_{i1} \le \dots \le r_{in_i}$,

$$\begin{aligned} (T-\lambda_i 1_V)^{r_{in_i}}\big|_{W_{ij}} &= 0, \qquad 1 \le j \le n_i \\ (T-\lambda_i 1_V)^{r_i n_i - 1}\big|_{W_{ij}} &\ne 0 \end{aligned}$$

showing

$$q_T | W_i = (t-\lambda_i)^{r_{in_i}}$$

So

$$V = W_1 \oplus \dots \oplus W_m$$

is the unique primary decomposition of $V$ relative to $T$.
*Note*: The Jordan canonical form of $T$ above is completely determined by the elementary divisors of $T$.

## §16.2 Jordan Canonical Form

**Theorem 16.6**

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear. Suppose that $q_T$ splits in $F[t]$. Then there exists a Jordan basis $\mathscr{B}$ for $V$ relative to $T$. Moreover, $[T]_{\mathscr{B}}$ is unique up to the order of the Jordan blocks. In addition, all such matrix representations are similar.

*Proof.* Reduction 1: We may assume that

$$q_T = (t-\lambda)^r$$

Suppose that

$$q_T = (t-\lambda_1)^{r_1} \dots (t-\lambda_m)^{r_m} \in F[t]$$

$\lambda_1, \dots, \lambda_m$ distinct. Set

$$W_i = \ker \left(T - \lambda_i 1_V\right)^{r_i}, \quad i = 1, \dots, m$$

By the Primary Decomposition Theorem,

$$V = W_1 \oplus \dots \oplus W_m$$

$W_i$ is $T$-invariant, $i = 1, \dots, n$

$$q_{T|_{W_i}} = (t-\lambda_i)^{r_i}, \qquad i = 1, \dots, m$$

So we need only find a Jordan basis for each $W_i$. □

# §17 Lec 17: May 5, 2021

## §17.1 Jordan Canonical Form (Cont'd)

*Proof.* (Cont'd from Lec 16) Reduction 2: We may assume that $q_T = t^r$, i.e., $\lambda = 0$. Suppose that we have proven the case for $\lambda = 0$. Let $S = T - \lambda 1_V$, $T$ as in Reduction 1. Then

$$S^r = (T - \lambda 1_V)^r = 0 \text{ and } S^{r-1} = (T - \lambda 1_V)^{r-1} \neq 0$$

Therefore,

$$q_S = t^r$$

if $\mathscr{B}$ is a Jordan basis for $V$ relative to $S$, then

$$[S]_{\mathscr{B}} = [T]_{\mathscr{B}} - \lambda I$$

is a JCF with diagonal entries 0. Hence

$$[T]_{\mathscr{B}} = [S]_{\mathscr{B}} + \lambda I$$

is a JCF with diagonal entries $\lambda$ and $\mathscr{B}$ is also a Jordan basis for $V$ relative to $T$. Reduction 2 now follows easily. We turn to
Existence: We have reduced to the case

$$q_T = t^r, \quad \text{i.e.,} \quad T^r = 0, \quad T^{r-1} \neq 0$$

In particular, $T$ is nilpotent. We induct on $\dim V$.

- $\dim V = 1$ is immediate.
- $\dim V > 1$: $T$ is singular, so $0 < \ker T$, as $\lambda = 0$ is an eigenvalue. Since $V$ is a finite dimensional vector space over $F$, by the Dimension Theorem, $T$ is not onto, i.e.,

  $$\text{im } T < V$$

  As im $T$ is $T$-invariant, we can (and do) view

  $$T\big|_{\text{im } T} : \text{ im } T \to \text{ im } T \text{ linear}$$

  As $T^r = 0$, certainly $\left(T|_{\text{im } T}\right)^r = 0$, so

  $$T\big|_{\text{im } T} \text{ is also nilpotent}$$

  and

  $$q_T\big|_{\text{im } T} | q_T \in F[t]$$

  since

  $$q_T\left(T\big|_{\text{im } T}\right) = 0 = q_T(T)$$

  So $q_{T|_{\text{im } T}}$ splits in $F[t]$ and

  $$q_{T\big|_{\text{im } T}} = t^s, \quad \text{for some } s \leq r$$

  by FTA. By induction on $\dim V$, im $T$ has a Jordan basis relative to $T|_{\text{im } T}$. So

  $$\text{im } T = W_1 \oplus \ldots \oplus W_m, \text{ some } m$$

  with each $W_i$ being $T|_{\text{im } T}-$ (hence $T-$) invariant and $W_i$ has a basis of an ordered sequence of generalized eigenvectors for $T|_{W_i}$, hence for $T|_{\text{im } T}$ and $T$,

  $$g_{r_i}(0) = \left\{w_i, Tw_i, \ldots, T^{r_i - 1} w_i\right\}, \quad r_i \geq 1$$

Thus we have

$$T^{r_i} w_i = 0, \qquad i = 1, \ldots, m$$
$$q_{T\big|_{W_i}} = t^{r_i}, \qquad i = 1, \ldots, m$$

Since $w_i \in W_i \subseteq \text{ im } T$,

$$\exists v_i \in V \ni T v_i = w_i, \quad i = 1, \ldots, m$$

So we also have

$$T^{r_i+1} v_i = T^{r_i} T v_i = T^{r_i} w_i = 0$$

and

$$T^{r_i} v_i = T^{r_i - 1} T v_i = T^{r_i - 1} w_i \neq 0$$

Therefore, $v_i, T v_i, \ldots, T^{r_i} v_i$ is an ordered sequence of generalized eigenvalues for $T$ in $V$, and, in particular, linearly independent. For each $i = 1, \ldots, m$, let

$$V_i = \text{Span}\left\{v_i, T v_i, \ldots, T^{r_i} v_i\right\}$$

So

$$\begin{aligned} V_i &= \left\{ \sum_{j=0}^{r_i} \alpha_j T^j v_i | \alpha_j \in F \right\} \\ &= \{f(T) v_i | f \in F[t],\ f = 0 \text{ or } \deg f \leq r_i\} \\ &= F[T]_{V_i} \end{aligned}$$

Since each $V_i$ is spanned by an ordered sequence of generalized eigenvectors for $T$, each $V_i$ is $T$-invariant, $i = 1, \ldots, m$.

*<u>Note</u>*: If $f \in F[t]$ and $f(T) w_i = 0$, then $f(T) = 0$ in $W_i$ and similarly if $f \in F[t]$ and $f(T) v_i = 0$, then $f(T) = 0$ on $V_i$ as $f(T) w_i = 0$ implies

$$0 = T^j f(T) w_i = f(T) T^j w_i = 0 \quad \forall i$$

Set

$$V' = V_1 + \ldots + V_m$$

Each $V_i$ is $T$-invariant, so $V'$ is $T$-invariant.

**Claim 17.1.** $V' = V_1 \oplus \ldots \oplus V_m$

In particular,

$$\mathscr{B}_0 = \{v_1, T v_1, \ldots, T^{r_i} v_1, \ldots, v_m, T v_m, \ldots, T^{r_m} v_m\}$$

is a basis for $V'$. □

# §18 Lec 18: May 7, 2021

## §18.1 Jordan Canonical Form (Cont'd)

*Proof.* (Cont'd) Suppose $u_i \in V_i$, $i = 1, \ldots, m$ satisfies

$$u_1 + \ldots + u_m = 0 \tag{1}$$

To show $u_i = 0$, $i = 1, \ldots, m$. As $u_i \in V_i$, $\exists f_i \in F[t] \ni$

$$u_i = f_i(T)v_i$$

where we let $f_i = 0$ if $u_i = 0$. So (1) becomes

$$f_i(T)v_1 + \ldots + f_m(T)v_m = 0 \tag{2}$$

Since $Tf(T) = f(T)T$ $\forall f \in F[t]$ and

$$w_i = Tv_i \qquad i = 1, \ldots, m$$

taking $T$ of (2) yields

$$f_1(T)w_1 + \ldots + f_m(T)w_m = 0$$

As the $T$-invariant $W_i$ satisfying

$$W_1 + \ldots + W_m = W_1 \oplus \ldots \oplus W_m \tag{*}$$

We have

$$f_i(T)w_i = 0, \qquad i = 1, \ldots, m$$

Hence

$$f_i(T) = 0 \text{ on } W_i, \qquad i = 1, \ldots, m$$

Thus

$$t^{r_i} = q_{T\big|_{W_i}} \big| f_i \in F[t], \quad i = 1, \ldots, m$$

In particular, since $r_i \geq 1$ $\forall i$, we can write

$$f_i = tg_i \in F[t], \qquad i = 1, \ldots, m$$
$$\deg g_i < \deg f_i, \qquad i = 1, \ldots, m \text{ if } f_i \neq 0$$

Since

$$f_i(T) = Tg_i(T) = g_i(T)T$$

and

$$w_i = Tv_i, \qquad i = 1, \ldots, m$$

(2) now becomes

$$g_1(T)w_1 + \ldots + g_m(T)w_m = 0 \tag{3}$$

Since each $W_i$ is $T$-invariant, by (*)

$$g_i(T)w_i = 0, \quad \text{hence } g_i(T) = 0 \text{ on } W_i$$

for $i = 1, \ldots, m$ by the definition of $W_i$. Therefore, for each $i$, $i = 1, \ldots, m$

$$t^{r_i} = q_{T\big|_{W_i}} \big| g_i \in F[t]$$

In particular, we can write

$$g_i = t^{r_i}h_i \in F[t], \qquad i = 1, \ldots, m$$

So

$$f_i = t^{r_i+1}h_i \in F[t], \qquad i = 1, \ldots, m$$

Thus we have

$$u_i = f_i(T)v_i = h_i(T)T^{r_i+1}v_i = 0, \quad i = 1, \ldots, m$$

This establishes claim 1. As

$$w_i = Tv_i \in W_i, \qquad i = 1, \ldots, m$$

We have

$$\begin{aligned} TV' &= TV_1 \oplus \ldots \oplus TV_m \\ &= W_1 \oplus \ldots \oplus W_m = TV \end{aligned} \tag{$\star$}$$

since each $W_i$, $V_i$ is $T$-invariant and

$$TV_i = W_i, \qquad i = 1, \ldots, m$$

Therefore,

$$T\big|_{V'} = T\big|_{V_1} + \ldots + T\big|_{V_m}$$

**Claim 18.1.** $V = \ker T + V'$

Let $v \in V$. Since

$$TV' = TV$$

by $(\star)$, we have $\forall v \in V$

$$\exists v' \in V' \ni Tv' = Tv,$$

so

$$v - v' \in \ker T$$

and

$$v = v' + w \text{ some } w \in \ker T$$

i.e.

$$v \in V' + \ker T$$

as needed.
Now by construction, we have a Jordan basis $\mathscr{B}_0$ for the $T$-invariant subspace $V'$ relative to $T|_{V'}$. Let

$$\mathscr{C} = \{u_1, \ldots, u_k\} \text{ be a basis for } \ker T = E_T(0)$$

Modifying the Toss In Theorem, we get a basis for $V$ as follows. If $u_1 \notin \text{Span } \mathscr{B}_0$, let $\mathscr{B}_1 = \mathscr{B}_0 \cup \{u_1\}$. Otherwise, let $\mathscr{B}_1 = \mathscr{B}_0$. If $u_2 \notin \text{Span } \mathscr{B}_1$, let $\mathscr{B}_2 = \mathscr{B}_1 \cup \{u_2\}$. Otherwise, let $\mathscr{B}_2 = \mathscr{B}_1$. In either case, $\mathscr{B}_2$ is a linearly independent set. Continuing in this way, since $\mathscr{B}_0 \cup \mathscr{C}$ spans $V$, we get a spanning set of $V$

$$\mathscr{B} = \mathscr{B}_0 \cup \{u_{j_1}, \ldots, u_{j_r}\} \subseteq V$$

with

$$T_{u_{j_i}} = 0$$

for some $u_{j_i}$ constructed above, $1 \leq i \leq s$.
Using claim 1, we have

$$\begin{aligned} V &= V' \oplus \text{Span } \{u_{j_1}, \ldots, u_{j_s}\} \\ &= V_1 \oplus \ldots \oplus V_m \oplus Fu_{j_1} \oplus \ldots \oplus Fu_{j_s} \end{aligned}$$

and $[T]_{\mathscr{B}}$ is in Jordan canonical form. This proves existence.

<u>*Note*</u>: $Fu_{j_i}$ are the $g_1(u_{j_i}, 0)$ and the $u_{j_i}$ are eigenvectors that cannot be extended to $g_i(v_i, 0)$ of longer length.
<u>Uniqueness</u>: By reduction 1) and 2), we have

$$q_T = t^r, \quad T^r = 0, \quad T^{r-1} \neq 0$$

Let $\mathscr{C}$ be an ordered basis for $V$. Then by MTT

$$m_j = \dim \operatorname{im} T^j = \text{rank } \left[T^j\right]_{\mathscr{C}} = \text{rank } [T]^j_{\mathscr{C}} \tag{*}$$

Let $\mathscr{B}$ be any Jordan basis for $V$ relative to $T$, say

$$[T]_{\mathscr{B}} = \begin{pmatrix} J_{r_1}(0) & & 0 \\ & \ddots & \\ 0 & & J_{r_m}(0) \end{pmatrix}$$

the corresponding Jordan canonical form. Prior computation showed for each $i$, $1 \le i \le m$,

$$\begin{cases} \text{rank } J^j_{r_i}(0) = r_i - j \\ \dim \ker J^j_{r_i}(0) = j \end{cases} \quad \text{if } j < r_i$$

and

$$\begin{cases} \text{rank } J^j_{r_i}(0) = 0 \\ \dim \ker J^j_{r_i}(0) = r_i \end{cases} \quad \text{if } j \ge r_i$$

Clearly, for each $i$,

$$[T]^j_{\mathscr{B}} = \begin{pmatrix} J^j_{r_1}(0) & & \\ & \ddots & \\ & & J^j_{r_m}(0) \end{pmatrix}$$

as $[T]_{\mathscr{B}}$ is in block form. So by (*),

$$m_j = \text{rank } [T]^j_{\mathscr{B}} = \sum_{i=1}^m \text{rank } J^j_{r_i}(0)$$

It follows that we have

$$\begin{aligned} m_{j-1} - m_j &= \text{rank } [T]^{j-1}_{\mathscr{B}} - \text{rank } [T]^j_{\mathscr{B}} \\ &= \#\text{ of } l \times l \text{ Jordan blocks } J_l(0) \text{ in } (+) \text{ with } l \ge j \end{aligned}$$

We also have, in the same way,

$$\begin{aligned} m_j - m_{j+1} &= \text{rank } [T]^j_{\mathscr{B}} - \text{rank } [T]^{j+1}_{\mathscr{B}} \\ &= \#\text{ of } l \times l \text{ Jordan blocks } J_l(0) \text{ in } (+) \text{ with } l \ge j+1 \end{aligned}$$

Consequently, there are precisely

$$(m_{j-1} - m_j) - (m_j - m_{j+1}) = m_{j-1} - 2m_j + m_{j+1}$$

which equals the number of $l \times l$ Jordan blocks $J_l(0)$ in (+) with $l = j$. This number is independent of $\mathscr{B}$ as it is

$$\text{rank } T^{j-1} - 2 \text{ rank } T^j + \text{rank } T^{j+1}$$

Thus, $[T]_{\mathscr{B}}$ is unique up to order of the Jordan blocks. This proves uniqueness.
If $\mathscr{B}'$ is another Jordan basis, then

$$[T]_{\mathscr{B}'} \sim [T]_{\mathscr{B}}$$

by the Change of Basis Theorem. This finishes the proof **(phewww... such a long proof!)** $\square$

**Corollary 18.1**

Let $A \in \mathbb{M}_n F$. If $q_A \in F[t]$ splits in $F[t]$, then $A$ is similar to a matrix in JCF unique up to the order of the Jordan blocks.

**Corollary 18.2**

Let $F$ be an algebraically closed field, e.g., $F = \mathbb{C}$. Then every $A \in \mathbb{M}_n F$ is similar to a matrix in JCF unique up to the order of the Jordan blocks and for every $V$, a finite dimensional vector space over $F$, and $T : V \to V$ linear, $V$ has a Jordan basis relative to $T$. Moreover, the Jordan blocks of $[T]_{\mathscr{B}}$ are completely determined by the elementary divisors (minimal polys) that correspond to the Jordan blocks.

**Theorem 18.3**

Let $F$ be an algebraically closed field, e.g., $F = \mathbb{C}$, $A, B \in \mathbb{M}_n F$. Then, the following are equivalent

1. $A \sim B$
2. $A$ and $B$ have the same JCF (up to block order)
3. $A$ and $B$ have the same elementary divisors counted with multiplicities.

**Corollary 18.4**

Let $F$ be an algebraically closed field. Then $A \sim A^\top$.

*Proof.* For any $B \in \mathbb{M}_n F$, $q_B = q_{B^\top}$. □

## §18.2 Companion Matrix

**Definition 18.5** (Companion Matrix) — Let $g = t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0 \in F[t]$, $n \geq 1$. The matrix

$$C(g) := \begin{pmatrix} 0 & 0 & \ldots & 0 & \text{———} & a_0 \\ 1 & 0 & & 0 & \text{———} & a_1 \\ 0 & 1 & & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots \\ & & & 0 & \text{———} & a_{n-2} \\ 0 & 0 & \ldots & 1 & \text{———} & a_{n-1} \end{pmatrix}$$

is called the companion matrix of $g$.

**Example 18.6**

$C(t^n) = J_n(0)$.

<u>Note</u>: If $f, g \in F[t]$ are monic, then

$$f = g \iff C(f) = C(g)$$

**Lemma 18.7**

Let $g \in F[t] \setminus F$ be moinc. Then

$$f_{C(g)} = g$$

*Proof.* Let $g = t^n + a_{n-1}t^{n-1} + \ldots + a_0 \in F[t] \setminus F$. We induct on $n$, using properties about determinants.

- $n = 1$ is immediate
- $n > 1$ Expanding on the determinant

$$f_{C(g)} = \det(tI - C(g)) = \det \begin{pmatrix} t & 0 & \ldots & 0 & a_0 \\ -1 & t & & \vdots & \\ 0 & -1 & & \vdots & \\ \vdots & 0 & & \vdots & \\ 0 & \ldots & \ldots & -1 & t + a_{n-1} \end{pmatrix}$$

  along the top row and induction yields

$$t\left(t^{n-1} + a_{n-1}t^{n-2} + \ldots + a_1\right) + (-1)^{n-1}a_0(-1)^{n-1} = g$$

□

**Lemma 18.8**

Let $g \in F[t] \setminus F$ be monic. Then

$$q_{C(g)} = f_{C(g)} = g$$

In particular,

$$f_{C(g)}\left(C(g)\right) = 0$$

# §19 Lec 19: May 10, 2021

## §19.1 Companion Matrix (Cont'd)

**Remark 19.1.** If $C$ is a companion matrix in $\mathbb{M}_n F$, viewing

$$C : F^{n\times 1} \to F^{n\times 1} \text{ linear,}$$

then

$$\mathscr{B} = \left\{e_1, Ce_1, \ldots, C^{n-1}e_1\right\}$$

is a basis for $F^{n\times 1}$ and

$$\begin{aligned} F^{n\times 1} &= \left\{\sum_{i=0}^{n-1} \alpha_i C^i e_i | \alpha_i \in F\right\} \\ &= F[C]e_1 := \{f(C)e_1 | f \in F[t]\} \end{aligned}$$

**Definition 19.2** (T-Cyclic) — Let $V$ be a vector space over $F$, $T : V \to V$ linear. We say $v \in V$ is a $T$-cyclic vector for $V$ and $V$ is $T$-cyclic if

$$V = \text{Span}\,\{v, Tv, \ldots, T^n v, \ldots\} = F[T]v$$

**Warning**: Let $T : V \to V$ be linear. It is rare that $V$ is $T$-cyclic. However, if $v \in V$, then $F[t]v \subseteq V$ is a $T$-invariant subspace and $F[T]v$ is $T$-cyclic. So $T$-cyclic subspace generalize the notion of a line in $V$.

**Proposition 19.3**

Let $V$ be a finite dimensional vector space over $F$, $n = \dim V$, $T : V \to V$ linear. Suppose there exists a $T$-cyclic vector $v$ for $V$, i.e., $V = F[T]v$. Then all of the following are true

i) $\mathscr{B} = \left\{v, Tv, \ldots, T^{n-1}v\right\}$ is an ordered basis for $V$

ii) $[T]_{\mathscr{B}} = C(f_T)$

iii) $f_T = q_T$

*Proof.* i) As $\dim V = n$, the set $\{v, Tv, \ldots, T^n v\}$ must be linearly independent. Let $j \le n$ be the first positive integer s.t.

$$T^j v \in \text{Span}\left\{v, Tv, \ldots, T^{j-1}v\right\}$$

say

$$T^j v = \alpha_{j-1}T^{j-1}v + \alpha_{j-2}T^{j-2}v + \ldots + \alpha_1 Tv + \alpha_0 v \tag{*}$$

for $\alpha_0, \ldots, \alpha_{j-1} \in F$. Take $T$ of (*), to get

$$T^{j+1}v = \alpha_{j-1}T^j v + \alpha_{j-2}T^{j-1}v + \ldots + \alpha_1 T^2 v + \alpha_0 Tv$$

which lies in $\text{Span}(v, Tv, \ldots, T^{j-1}v)$ by (*). Iterating this process shows

$$T^N v \in \text{Span}\left\{v, Tv, \ldots, T^{j-1}v\right\} \qquad \forall N \ge j$$

It follows that

$$v = F[T]v = \text{Span}\left\{v, Tv, \ldots, T^{j-1}v\right\}$$

So

$$n = \dim V \leq j, \qquad \text{hence } n = j$$

This proves $i$).

ii) The computation proving $i$) shows

$$\mathscr{B} = \left\{v, Tv, \ldots, T^{n-1}v\right\}$$

is an ordered basis for $V$. As

$$\begin{aligned}[T]_{\mathscr{B}} &= \left([Tv]_{\mathscr{B}} \quad \left[T^2v\right]_{\mathscr{B}} \quad \ldots \quad \left[T^{n-2}v\right]_{\mathscr{B}} \quad [T^nv]_{\mathscr{B}}\right) \\ &= \begin{pmatrix} 0 & 0 & & 0 & * \\ 1 & 0 & & \vdots & * \\ 0 & 1 & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & * \end{pmatrix}\end{aligned}$$

it is a companion matrix, hence must be $C(f_T)$ and by the lemma, we have proven $ii$).

iii) $f_T = f_{[T]_{\mathscr{B}}} = q_{[T]_{\mathscr{B}}} = q_T$ as $[T]_{\mathscr{B}} = C(f_T)$. □

**Example 19.4**

Let $V$ be a finite dimensional vector space over $F$, $\dim V = n$, $T : V \to V$ linear s.t. there exists an ordered basis $\mathscr{B}$ with

$$[T]_{\mathscr{B}} = J_n(\lambda)$$

Set $S = T - \lambda 1_V : V \to V$ linear. Then $\exists v \in V \ni$

$$\mathscr{B} = \left\{v, Sv, \ldots, S^{n-1}v\right\}$$

So $v$ s an $S$-cyclic vector and

$$V = F[S]v$$

**Fact 19.1.** If $A \in \mathbb{M}_r F[t]$, $C \in \mathbb{M}_s F[t]$, $B \in F[t]^{r\times s}$, then

$$\det \begin{pmatrix} A & B \\ O & C \end{pmatrix} = \det A \det C$$

where

$$\det D = \sum \operatorname{sgn} \sigma D_{1\sigma(1)} \ldots D_{n\sigma(n)}$$

## §19.2 Smith Normal Form

We say that $A \in F[t]^{m\times n}$ is in Smith Normal Form (SNF) if $A$ is the zero matrix or if $A$ is the matrix of the form

$$\begin{pmatrix} q_1 & 0 & \ldots & & & & \\ 0 & q_2 & & & & & \\ \vdots & & \ddots & & & & \\ & & & q_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & & & & \end{pmatrix}$$

with $q_1|q_2|q_3|\ldots|q_r$ in $F[t]$ and all monic, i.e., there exists a positive integer $r$ satisfying $r \le \min(m,n)$ and $q_1|q_2|q_3|\ldots|q_r$ monic in $F[t]$ s.t. $A_{ii} = q_i$ for $1 \le i \le r$ and $A_{ij} = 0$ otherwise.
We generalize Gaussian elimination, i.e., row (and column) reduction for matrices with entries in $F$ to matrices with entries in $F[t]$. The only difference arises because most elements of $F[t]$ do not have multiplicative inverses.
Let $A \in \mathbb{M}_n(F[t])$. We say that $A$ is an elementary matrix of

i) Type I: if there exists $\lambda \in F[t]$ and $l \neq k$ s.t.

$$A_{ij} = \begin{cases} 1 & \text{if } i = j \\ \lambda & \text{if } (i,j) = (k,l) \\ 0 & \text{otherwise} \end{cases}$$

ii) Type II: If there exists $k \neq l$ s.t.

$$A_{ij} = \begin{cases} 1 & \text{if } i = j \neq l \text{ or } i = j \neq k \\ 0 & \text{if } i = j = l \text{ or } i = j = k \\ 1 & \text{if } (k,l) = (i,j) \text{ or } (k,l) = (j,i) \\ 0 & \text{otherwise} \end{cases}$$

iii) Type III: If there exists a $0 \neq u \in F$ and $l$ s.t.

$$A_{ij} = \begin{cases} 1 & \text{if } i = j \neq l \\ u & \text{if } i = j = l \\ 0 & \text{otherwise} \end{cases}$$

**Remark 19.5.** Let $A \in F[t]^{m\times n}$. Multiplying $A$ on the left (respectively right) by a suitable size elementary matrix of

a) Type I is equivalent to adding a multiple of a row (respectively column) of $A$ to another row (respectively column) of $A$.
b) Type II is equivalent to interchanging two rows (respectively columns) of $A$.
c) Type III is equivalent to multiplying a row (respectively column) of $A$ by an element in $F[t]$ having a multiplicative inverse.

**Remark 19.6.**
1. All elementary matrices are invertible.
2. The definition of elementary matrices of Types I and II is exactly the same as that given when defined over a field.
3. The elementary matrices of Type III have a restriction. The $u$'s appearing in the definition are precisely the elements in $F[t]$ having a multiplicative inverse. The reason for this is so that the elementary matrices of Type III are invertible.

Let

$$GL_n\left(F[t]\right) \coloneqq \{A|A \text{ is invertible}\}$$

**Warning**: A matrix in $\mathbb{M}_n(F[t])$ having $\det(A) \neq 0$ may no longer be invertible, i.e., have an inverse. What is true is that $GL_n(F[t]) = \{A|0 \neq \det(A) \in F\}$, equivalently $GL_n(F[t])$ consists of those matrices whose determinant have a multiplicative inverse in $F[t]$.

**Definition 19.7** (Equivalent Matrix) — Let $A, B \in F[t]^{m\times n}$. We say that $A$ is equivalent to $B$ and write $A \approx B$ if there exist matrices $P \in GL_m(F[t])$ and $Q \in GL_n(F[t])$ s.t. $B = PAQ$.

**Theorem 19.8**

Let $A \in F[t]^{m\times n}$. Then $A$ is equivalent to a matrix in Smith Normal Form. Moreover, there exist matrices $P \in GL_m(F[t])$ and $Q \in GL_n(F[t])$, each a product of matrices of Type I, Type II, Type III, s.t. $PAQ$ is in SNF.

**Remark 19.9.** The SNF derived by this algorithm is, in fact, unique. In particular, the monic polynomials $q_1|q_2|q_3|\dots|q_r$ arising in the SNF of a matrix $A$ are unique and are called the **invariant factor** of $A$. This is proven using results about determinant.

# §20 Lec 20: May 12, 2021

## §20.1 Rational Canonical Form

If $A, B \in F[t]^{m\times n}$ then $A \approx B$ if and only if they have the same SNF if and only if they have the same invariant factors. So what good is the NSF relative to linear operators on finite dimensional vector spaces?

Let $A, B \in \mathbb{M}_n(F)$. Then $A \sim B$ if and only if $tI - A \approx tI - B$ in $\mathbb{M}_n(F[t])$ and this is completely determined by the SNF hence the invariant factors of $tI - A$ and $tI - B$. Now the SNF of $tI - A$ may have some of its invariant factors 1, and we shall drop these.
Let $V$ be a finite dimensional vector space over $F$ with $\mathscr{B}$ an ordered basis. Let $T : V \to V$ be a linear operator. If one computes the SNF of $tI - [T]_{\mathscr{B}}$, it will have the form

$$\begin{pmatrix} 1 & 0 & & \dots & \dots & 0 \\ 0 & 1 & & & & 0 \\ \vdots & & \ddots & & & \vdots \\ & & & q_1 & & \\ & & & & q_2 & \\ \vdots & & & & & \ddots & \vdots \\ 0 & & & \dots & \dots & q_r \end{pmatrix}$$

with $q_1|q_1|\dots|q_r$ are all the monic polynomials in $F[t] \setminus F$. These are called the **invariant factors** of $T$. They are uniquely determined by $T$. The main theorem is that there exists an ordered basis $\mathscr{B}$ for $V$ s.t.

$$[T]_{\mathscr{B}} = \begin{pmatrix} C(q_1) & 0 & \dots & 0 \\ 0 & C(q_2) & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & C(q_r) \end{pmatrix}$$

and this matrix representation is unique. This is called the rational canonical form or **RCF** of $T$. Moreover, the minimal polynomial $q_t$ of $T$ is $q_r$. The algorithm computes this as well as all invariant factors of $T$. The characteristic polynomial $f_T$ of $T$ is the product of $q_1 \dots q_r$. This works over any field $F$, even if $q_T$ does not split. The basis $\mathscr{B}$ gives a decomposition of $V$ into $T$-invariant subspaces $V = W_1 \oplus \dots \oplus W_r$ where $f_{T|W_i} = q_{T|W_i} = q_i$ and if $\dim(W_i) = n_i$ then $\mathscr{B}_i = \left\{v_i, Tv_i, \dots, T^{n_i - 1}v_i\right\}$ is a basis for $W_i$.
Let $V$ be a finite dimensional vector space over $F$ with $\mathscr{B}$ an ordered basis. Let $T : V \to V$ be a linear operator. Suppose that $q_T$ splits over $F$. Then we know that there exists a Jordan canonical form of $T$.

**Question 20.1.** How do we compute it?

We use the Smith Normal Form of $tI - [T]_{\mathscr{B}}$ to compute the invariant factors $q_1|q_1|\dots|q_r$ of $T$ just as one does to compute the **RCF** of $T$. We then factor each $q_i$. Suppose this factorization is

$$q_i = (t - \lambda_1)^{r_1} \dots (t - \lambda_m)^{r_m}$$

in $F[t]$ with $\lambda_1, \dots, \lambda_m$ distinct. Note that $q_{i+1}$ has this as a factor so it has the form

$$q_{i+1} = (t - \lambda_1)^{s_1} \dots (t - \lambda_m)^{s_m} \dots (t - \lambda_{m+k})^{s_{m+k}}$$

with $s_i \ge r_i$ for each $1 \le i \le m$ and $m + 1, \dots, m + k \ge 0$ with $\lambda_1, \dots, \lambda_{m+k}$ distinct. Then the totality of all the $(t - \lambda_i)^{r_j}$, including repetition if they occur in different $q_i$'s give all the elementary divisors of $T$. So to get the JCF of $T$ we take for each $q_i$ as factored above the block matrix

$$\begin{pmatrix} J_{r_1}(\lambda_1) & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & J_{r_m}(\lambda_m) \end{pmatrix}$$

and replace $C(q_i)$ by it in the RCF, i.e., we take all the Jordan blocks $J_r(\lambda)$ associated to each and every factor of the form $(t-\lambda)^r$ in each and every invariant factor $q_i$ determined by the SNF and form a matrix out of all such blocks. This is the JCF which is unique only up to block order.
Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear, $v \in V$. Then as before, if $v \in V$

$$F[t]v = \{f(T)v | f \in F[t]\} \subseteq V$$

the $T$-cyclic subspace of $V$ generated by $v$ and satisfies

$$n_v := \dim F[T]v \leq \dim V$$

and has ordered basis

$$\mathscr{B}_v := \left\{v, Tv, \ldots, T^{n_v - 1}v\right\}$$

As $F[T]v$ is $T$-invariant,

$$\left[T|_{F[T]v}\right]_{\mathscr{B}v} = C\left(f_{T|_{F[T]v}}\right)$$

and

$$q_{T|_{F[T]v}} = f_{T|_{F[T]v}}$$

We want to show that $V$ can be decomposed as a direct sum of $T$-cyclic subspaces of $V$. The SNF of the characteristic matrix

$$tI - [T]_{\mathscr{C}}$$

$\mathscr{C}$ is an ordered basis for $V$, which gives rise to invariants of $T$

$$q_1 | \ldots | q_r \in F[t] \tag{*}$$

$q_1 \neq 1$, $q_i$ monic for all $i$.
<u>*Note*</u>: The SNF of (+) has no 0's on the diagonal $asf_T \neq 0$. We want to show there exists an ordered basis $\mathscr{B}$ for $V$ with all the following properties

i) $V = W_1 \oplus \ldots \oplus W_r$, $n_i = \dim W_i$, $i = 1, \ldots, r$

ii) $W_i$ is $T$-invariant, $i = 1, \ldots, r$

iii) $W_i = F[T]v_i$ are $T$-cyclic, $W_i = \ker q_{T|_{W_i}}(T|_{W_i})$

iv) $q_i = q_{T|_{W_i}} = f_{T|_{W_i}}$, $i = 1, \ldots, r$ with $q_i$ as in (*)

v) $q_T = q_r$

vi) $f_T = q_1 \ldots q_r = q_{T|_{W_1}} \cdots q_{T|_{W_r}}$

vii) $\mathscr{B}_{v_i} = \left\{v_i, Tv_i, \ldots, T^{n_i - 1}v_i\right\}$ is an ordered basis for $W_i$, $i = 1, \ldots, r$

viii) $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_r$ is an ordered basis for $V$ satisfying

$$[T]_{\mathscr{B}} = \begin{pmatrix} C(q_1) & & 0 \\ & \ddots & \\ 0 & & C(q_r) \end{pmatrix}$$

called the rational canonical form of $T$ and it is unique.

The uniqueness follows from the uniqueness of SNF. From the definition of equivalent matrix, we have the following remark

**Remark 20.1.** If $A \in \mathbb{M}_n F[t]$ is in SNF, then

$$A \in GL_n F[t] \iff A = I$$

since

$$\begin{pmatrix} q_1 & & & & & 0 \\ & \ddots & & & & \\ & & q_r & & & \\ & & & 0 & & \\ & & & & & \\ 0 & & & & & \ddots \end{pmatrix}$$

means $0 \dots 0 \cdot q_1 \dots q_r \in F \setminus \{0\}$ if there are any 0's on the diagonal, which is inseparable.

**Lemma 20.2**

Let $g \in F[t] \setminus F$ be monic of degree $n$. Then

$$It - C(q) \approx \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & q \end{pmatrix}$$

**Corollary 20.3**

Let $V$ be a finite dimensional vector space over $F$, $T : V \to V$ linear $q_1 | \dots | q_r$ the invariants of $T$ in $F[t]$. Then

$$tI - \begin{pmatrix} C(q_1) & & 0 \\ & \ddots & \\ 0 & & C(q_r) \end{pmatrix}$$

where $\dim V = \sum_{i=1}^{r} \deg q_i$

Certainly, if there exists an ordered basis $\mathscr{B}$ for $V$ a finite dimensional vector space over $F$, $T : V \to V$ linear s.t. $[T]_{\mathscr{B}}$ is in RCF, then everything in goal falls out. So by the above, the goal will follow if we prove the following

**Theorem 20.4**

Let $A_0, B_0 \in \mathbb{M}_n F$, $A = tI - A_0$ and $B = tI - B_0$ in $\mathbb{M}_n F[t]$, the corresponding characteristic matrices. Then the following are equivalent

i) $A_0 \sim B_0$ (i.e. $A_0$ and $B_0$ are similar)

ii) $A \sim B$ (i.e., $A$ and $B$ are equivalent)

iii) $A$ and $B$ have the same SNF.

We need two preliminary lemmas.

**Lemma 20.5**

Let $A \approx B$ in $\mathbb{M}_n F[t]$. Then $\exists P, Q \in GL_m F[t]$ each products of elementary matrices s.t. $A = PBQ$.

*Proof.* $P \in GL_n F[t]$ iff its SNF $= I$ which we get using elementary matrices. □

For the second lemma, we need the "division algorithm" by "linear polys" in $\mathbb{M}_n F[t]$. If we were in $F[t]$, we know if $f, g \in F[t]$, $f \neq 0$,

$$g = fq + r \in F[t] \text{ with } r = 0 \text{ or } \deg r < \deg f$$

So if $f = t - a$, $r \in F$, i.e., $r = g(a)$ by plugging in $a$ into (*). But for matrices,

$$AQ + R \neq QA + R$$

but the same argument to get (*) for polys, will give a right and left remainder.
Notation: Let $A_i \in \mathbb{M}_r F$, $i = 0, \ldots, n$ and let

$$A_n t^n + A_{n-1} t^{n-1} + \ldots + A_0$$

denote

$$A_n(t^n I) + \ldots + A_0 I \in \mathbb{M}_n F[t]$$

So if

$$A = (\alpha_{ij})$$

then

$$At^n = (\alpha_{ij} t^n)$$

i.e., two matrix polynomials are the same iff all their corresponding entries are equal, i.e.,

$$(\mathbb{M}_n F)[t] = \mathbb{M}_r \left(F[t]\right)$$

**Lemma 20.6**

Let $A_0 \in \mathbb{M}_n F$, $A = tI - A_0 \in \mathbb{M}_n F[t]$ and

$$0 \neq P = P(t) \in \mathbb{M}_n F[t]$$

Then there exist matrices $M, N \in \mathbb{M}_n F[t]$ and $R, S \in \mathbb{M}_n F$ satisfying

i) $P = AM + R$

ii) $P = NA + S$

# §21 Lec 21: May 14, 2021

## §21.1 Rational Canonical Form (Cont'd)

Recall from last lecture,

**Lemma 21.1**
Let $A_0 \in \mathbb{M}_n F$, $A = tI - A_0 \in \mathbb{M}_n F[t]$ and

$$0 \neq P = P(t) \in \mathbb{M}_n F[t]$$

Then there exist matrices $M, N \in \mathbb{M}_n F[t]$ and $R, S \in \mathbb{M}_n F$ satisfying

i) $P = AM + R$

ii) $P = NA + S$

*Proof.* i) Let

$$m = \max_{l,k} \deg P_{lk}, \quad P_{lk} \neq 0$$

and $\forall i, j$ let

$$\alpha_{ij} = \begin{cases} \text{lead } P_{ij} \text{ if } \deg P_{ij} = m \\ 0 \quad \text{if } P_{ij} = 0 \text{ or } \deg P_{ij} < m \end{cases}$$

So

$$P_{ij} = \alpha_{ij} t^m + \text{ lower terms in } t \in F[t]$$

Let $\alpha_{ij} \in \mathbb{M}_n F$ and let

$$P_{m-1} = (\alpha_{ij}) t^{m-1} = \left(\alpha_{ij} t^{m-1}\right)$$

Every entry in

$$\begin{aligned} AP_{m-1} &= (tI - A_0)\,(\alpha_{ij}) t^{m-1} \\ &= (\alpha_{ij}) t^m - A_0 (\alpha_{ij}) t^{m-1} \end{aligned}$$

has deg $= m$ or is zero and the $t^m$-coefficient of $(AP_{m-1})_{ij}$ is $\alpha_{ij}$. Thus, $P - AP_{m-1}$ has polynomial entries of degree at most $m-1$ (or $= 0$). Apply the same argument to $P - AP_{m-1}$ (replacing $m$ by $m-1$ in (*)) to produce a matrix $P_{m-2}$ in $\mathbb{M}_n F[t]$ s.t. all the polynomial entries in $(P - AP_{m-1}) - AP_{m-2}$ have degree at most $m-2$ (or $= 0$). Continuing this way, we construct matrices $P_{m-3}, \ldots, P_0$ satisfying if

$$M := P_{m-1} + P_{m-2} + \ldots + P_0$$

then

$$R := P - AM$$

has only constant entries, i.e., $R \in \mathbb{M}_n F$. So

$$P = AM + R$$

as needed.

ii) This can be proven in an analogous way. □

**Theorem 21.2**

Let $A_0, B_0 \in \mathbb{M}_n F$, $A = tI - A_0$, $B = tI - B_0$ in $\mathbb{M}_n F[t]$. Then

$$A \approx B \in \mathbb{M}_n F[t] \iff A_0 \sim B_0 \in \mathbb{M}_n F$$

*Proof.* " $\Longleftarrow$ " If

$$B_0 = PA_0P^{-1}, \quad P \in GL_n F,$$

then

$$P\left(tI - A_0\right)P^{-1} = PtP^{-1} - PA_0P^{-1} = tI - B_0 = B$$

So $B = PAP^{-1}$ and $B \approx A$.
" $\implies$ " Suppose there exist $P_1, Q_1 \in GL_n F[t]$, hence each a product of elementary matrices by Lemma 20.5, satisfying

$$B = tB - B_0 = P_1 A Q_1 = P_1\left(tI - A_0\right)Q_1$$

Applying Lemma 21.1, we can write

i) $P_1 = BP_2 + R$, $P_2 \in \mathbb{M}_n F[t]$, $R \in \mathbb{M}_n F$

ii) $Q_1 = Q_2 B + S$, $Q_2 \in \mathbb{M}_n F[t]$, $S \in \mathbb{M}_n F$

Since $B = P_1 A Q_1$, $P_1, Q_1 \in GL_n F[t]$, we also have

iii) $P_1 A = BQ^{-1}$

iv) $AQ_1 = P_1^{-1} B$

Thus, we have

$$\begin{aligned} B = P_1 A Q_1 &\overset{i)}{=} (BP_2 + R)AQ_1 = BP_2AQ_1 + RAQ_1 \\ &\overset{iv)}{=} BP_2P_1^{-1}B + RAQ_1 \overset{ii)}{=} BP_2P_1^{-1}B + RA(Q_2B + S) \\ &= BP_2P_1^{-1}B + RAQ_2B + RAS \end{aligned}$$

i.e., we have

v) $B = BP_2P_1^{-1}B + RAQ_2B + RAS$

By i)

$$R = P_1 - BP_2$$

Plugging this into $RAQ_2B$, yields

$$\begin{aligned} RAQ_2B &\overset{i)}{=} (P_1 - BP_2)AQ_2B = P_1AQ_2B - BP_2AQ_2B \\ &\overset{iii)}{=} BQ_1^{-1}Q_2B - BP_2AQ_2B = B\left[Q_1^{-1}Q_2 - P_2AQ_2\right]B \end{aligned}$$

i.e.

vi) $RAQ_2B = B\left[Q_1^{-1}Q_2 - P_2AQ_2\right]B$

Plug vi) into v) to get

$$\begin{aligned} B &\overset{v)}{=} BP_2P_1^{-1}B + RAQ_2B + RAS \\ &\overset{vi)}{=} BP_2P_1^{-1}B + B\left[Q_1^{-1}Q_2 - P_2AQ_2\right]B + RAS \\ &= B\left[P_2P_1^{-1} + Q_1^{-1}Q_2 - P_2AQ_2\right]B + RAS \end{aligned}$$

Let

$$T = P_2P_1^{-1} + Q_1^{-1}Q_2 - P_2AQ_2$$

Then

vii) $B = BTB + RAS \in \mathbb{M}_n F[t]$

We next look at the degree of the poly entries of these matrices.

viii) Every entry of $B = tI - B_0$ is zero or has deg $\leq 1$ and every entry of $RAS = R(tI - A_0)S$ has is zero or has deg $\leq 1$.

**Question 21.1.** What about $BTB$?

Let $T = T_m t^m + T_{m-1}t^{m-1} + \ldots + T_0$ with $T_0, \ldots, T_m \in \mathbb{M}_n F$. Then

$$\begin{aligned} BTB &= (tI - B_0)\left(T_m t^m + T_{m-1}t^{m-1} + \ldots + T_0\right)(tI - B_0) \\ &= T_m t^{m+2} + \text{ lower terms in } t \end{aligned}$$

Comparing coefficients of the matrix of polys $BTB = B - RAS$ using vii), viii) shows

$$T_m = 0$$

Hence

$$T = 0$$

So vii) becomes

$$\begin{aligned} tI - B_0 = B = BTB + RAS = RAS &= R(tI - A_0)S \\ &= RST + RA_0S \end{aligned} \tag{*}$$

comparing coefficients of the poly matrices in (*) shows

$$\begin{aligned} I &= RS \\ B_0 &= RA_0S \end{aligned}$$

i.e., $B_0 = RA_0S = RA_0R^{-1}$. □

**Theorem 21.3**

Let $A_0, B_0 \in \mathbb{M}_n F$, $A = tI - A_0$, $B = tI - B_0$ in $\mathbb{M}_n F[t]$. Then the following are equivalent

i) $A_0 \sim B_0$

ii) $A \approx B$

iii) $A$ and $B$ have the same SNF.

iv) $A_0$ and $B_0$ have the same invariant factors.

In particular, if $V$ is a finite dimensional vector space over $F$, $T : V \to V$ linear, $q_1 | \ldots | q_r$ the invariants of $T$, then

$$\begin{aligned} V &= \ker q_1(T) \oplus \ldots \oplus \ker q_n(T) \\ q_r &= q_T \\ f_T &= q_1 \ldots q_r \end{aligned}$$

<u>*Note*</u>: If $q_i = \prod_{j=1}^{r}(t - \lambda_i)^{e_j}$ is an invariant factor, then

$$C(q_i) \sim \begin{pmatrix} J_{e_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{e_r}(\lambda_r) \end{pmatrix}$$

**Corollary 21.4**

Let $A, B \in \mathbb{M}_n F$, $F \subseteq K$ a subfield. Then $A \sim B$ in $\mathbb{M}_n F$ iff $A \sim B$ in $\mathbb{M}_n K$.

# §22 Lec 22: May 17, 2021

## §22.1 Inner Product Spaces

Notation: $\overline{-}: \mathbb{C} \to \mathbb{C}$ by $\alpha + \beta\sqrt{-1} \mapsto \alpha - \beta\sqrt{-1}$ $\forall \alpha, \beta \in \mathbb{R}$ is called the **complex conjugation**. If $F \subseteq \mathbb{C}$, set

$$\overline{F} := \{\overline{\alpha} | \, \alpha \in F\}$$

is a field, e.g., $\overline{F} = F$ if $F \subseteq \mathbb{R}$.

**Definition 22.1** (Inner Product Space) — Let $F \subseteq \mathbb{C}$ satisfy $F = \overline{F}$, $V$ a vector space over $F$. Then $V$ is called an inner product space over $F$ relative to

$$\langle , \rangle = \langle , \rangle_V : V \times V \to F$$

satisfies

1. $p_v : V \to F$ by $p_v(w) := \langle w, v\rangle$ is linear for all $v \in V$, i.e., $p_v \in V^*$
2. $\langle v, w\rangle = \overline{\langle v, w\rangle}$ for all $v, w \in V$
3. $\|v\|^2 := \langle v, v\rangle \in \mathbb{R} \cap F$ for all $v \in V$ and $\|v\|^2 \geq 0$ in $\mathbb{R}$ and $= 0$ iff $v = 0$ (*)

Let $V$ be an inner product space over $F$. Then,

1. If $v \in V$ satisfies $\langle w, v\rangle = 0$ for all $w \in V$, then $v = 0$.
2. Let $v_1, v_2 \in V \setminus \{0\}$,
$$w = \frac{\langle v_2, v_1\rangle}{\|v_1\|^2} v_1$$
is called the orthogonal projection of $v_2$ on $v_1$ and $v = v_2 - w$ is orthogonal to $w$, i.e. $\langle v, w\rangle = 0$, write $v \perp w$.

**Definition 22.2** (Sesquilinear Map) — A map $f : V \to W$ of inner product space over $F$ is called sesquilinear if $v_1, v_2 \in V$, $\alpha \in F$

$$f(v_1 + \alpha v_2) = f(v_1) + \overline{\alpha} f(v_2)$$

Let $V^\dagger := \{f : V \to F | \, f \text{ sesquilinear}\}$ a vector space over $F$.

**Example 22.3**
If $F \subseteq \mathbb{R}$, then any sesquilinear map is linear and $V^\dagger = V^*$.

**Remark 22.4.** Let $V$ be an inner product space over $F$.

1. $p : V \to V^*$ by $v \mapsto p_v$ is sesquilinear.

$$\begin{aligned} p(\alpha v_1 + v_2)(w) &= \langle w, \alpha v_1 + v_2 \rangle \\ &= \overline{\alpha}\langle w, v_2 \rangle + \langle w, v_1 \rangle = \overline{\alpha} p(v_1) + p(v_2) \end{aligned}$$

   for all $\alpha \in F$, $v_1, v_2, w \in V$. Also, we can deduce that $p$ is an injection and if $V$ is finite dimensional, then $p$ is a bijection.

2. If $v \in V$, let $\lambda_v : V \to F$ by $w \mapsto \langle v, w \rangle$, i.e., $\lambda_v(w) = \langle v, w \rangle$. Then $\lambda_v$ is sesquilinear. Moreover,

$$\lambda : V \to V^\dagger \text{ by } v \mapsto \lambda v$$

   is linear. As $\langle v, w \rangle = 0$ for all $w \to v = 0$, $\lambda$ is injective hence monic. If $V$ is finite dimensional then $\lambda$ is an isomorphism.

3. If $f : V \to W$ is sesquilinear, it is called a sesquilinear isomorphism if it is bijective and $f^{-1}$ is sesquilinear. Then $f$ is a sesquilinear isomorphism iff $f$ is bijective.

Let $V$ be an inner product space over $F$.

1. If $v \in V$, $\|v\| := \sqrt{\|v\|^2} \geq 0$ is called the length of $v$.

2. Length and $\angle$ make sense in $V$ by the <u>Cauchy – Schwarz</u> inequality

$$|\langle v, w \rangle| \leq \|v\| \|w\| \quad \forall v, w \in V$$

   and $V$ is a metric space by distances from $v, w := d(v, w) := \|v - w\|$ as the triangle inequality

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

   holds for all $v, w \in W$.

3. <u>Gram – Schmidt</u>: If $W \subseteq V$ is a finite dimensional subspaces, then $\exists$ an orthogonal basis for $W$

$$\mathscr{B} = \{w_1, \ldots, w_n\}, \quad \text{i.e. } \langle w_i, w_j \rangle = 0 \text{ if } i \neq j$$

   and if $\|w_i\| \in F$ $\forall i$, then $\exists$ an orthonormal basis

$$\mathscr{C} = \left\{ \frac{w_1}{\|w_1\|}, \ldots, \frac{w_n}{\|w_n\|} \right\}$$

4. In 3), if $v \in V$ let $\mathscr{B} = \{w_1, \ldots, w_n\}$ be an orthogonal basis for $W$. Set

$$v_w := \sum_{i=1}^n \frac{\langle v, w_i \rangle}{\|w_i\|^2} w_i = \sum_{i=1}^n \langle v, \frac{w_i}{\|w_i\|^2} \rangle w_i$$

   Then, the $w_i$-coordinate of $v_w$ is $\frac{\langle v, w_i \rangle}{\|w_i\|^2} \in F$. Hence

$$f_i = p_{\frac{w_i}{\|w_i\|^2}} : V \to F$$

   is the corresponding coordinate function, so $\mathscr{B}^* = \{f_1, \ldots, f_n\}$ is the dual basis of $\mathscr{B}$.

5. Let $\emptyset \neq S \subseteq V$ be a subset. The orthogonal complement $S^\perp$ of $S$ is defined by

$$S^\perp := \{x \in V |\, x \perp s \ \forall s \in S\} \subseteq V$$

   a subspace.

*Note*: The sesquilinear map

$$p : V \to V^* \text{ by } v \mapsto p_v$$

induces an injective sesquilinear map

$$p\big|_{S^\perp} : S^\perp \to S^\circ$$

and we have

$$S \subseteq S^{\perp\perp} := (S^\perp)^\perp$$

If $S$ is a subspace, $S \cap S^\perp = 0$ so

$$S + S^\perp = S \oplus S^\perp$$

write

$$S + S^\perp = S \perp S^\perp$$

called an **orthogonal direct sum** and if $V$ is finite dimensional then

$$S = S^{\perp\perp}$$

e.g., if $v \in V$, then

$$\ker p_v = (Fv)^\perp$$

so

$$V = Fv \perp (Fv)^\perp$$

More generally, we have the following crucial result.

**Theorem 22.5** (Orthogonal Decomposition)
Let $V$ be an inner product space over $F$, $S \subseteq V$ a finite dimensional subspace. Then

$$V = S \perp S^\perp$$

i.e., if $v \in V$

$$\exists! s \in S,\ s^\perp \in S^\perp \ni v = s + s^\perp$$

In particular, $s = v_S$. If $V$ is finite dimensional, then

$$\dim V = \dim S + \dim S^\perp$$

**Theorem 22.6** (Best Approximation)
Let $V$ be an inner product space over $F$, $S \subseteq V$ a finite dimensional subspace, $v \in V$. Then $v_S \in S$is the best approximation to $v$ in $S$, i.e., for all $s \in S$

$$\|v - v_S\| \le \|v - s\| \text{ with equality iff } s = v_S$$

**Remark 22.7.** More generally, if $V$ is an inner product space over $F$,

$$V = W_1 \oplus \ldots \oplus W_n$$

with

$$w_i \perp w_j \qquad \forall w_i \in W_i,\ w_j \in W_j,\ i \neq j$$

We call $V$ an orthogonal direct sum or orthogonal decomposition of $V$.

By the Orthogonal Decomposition Theorem,

$$V = W_i \perp W_i^\perp$$

and

$$W_i^\perp = W_1 \perp \ldots \underbrace{\hat{W}_i}_{\text{omit}} \perp \ldots \perp W_n$$

Let $P_i : V \to V$ be the projection along

$$W_i^\perp = W_1 \perp \ldots \perp \hat{W}_i \perp \ldots \perp W_n$$

onto $W_i$. Then we have

$$\begin{aligned} \ker P_i &= W_i^\perp \\ \text{im } P_i &= W_i \\ P_i P_j &= \delta_{ij} P_j \qquad \forall i, j \\ 1_V &= P_1 + \ldots + P_n \end{aligned}$$

The $P_i$ are called **orthogonal projections**. As $W_i \subseteq V$ is finite dimensional in the above,

$$P_i(v) = v_{W_i}$$

So

$$v = v_{W_1} + \ldots + v_{W_n}$$

is a unique decomposition of $v$ relative to (*).

**Definition 22.8** (Adjoint) — Let $V, W$ be inner product spaces over $F$, $T : V \to W$ linear. A linear transformation $T^* : W \to V$ is called the adjoint of $T$ if

$$\langle Tv, w \rangle_W = \langle v, T^* w \rangle_V \qquad \forall v \in V \ \forall w \in W$$

**Theorem 22.9**

Let $V, W$ be finite dimensional inner product space over $F$, $T : V \to W$ linear. Then the adjoint $T^* : W \to V$ exists.

# §23 Lec 23: May 19, 2021

## §23.1 Inner Product Spaces (Cont'd)

**Corollary 23.1**

Let $V, W$ be finite dimensional inner product space over $F$, $T : V \to W$ linear. Then

$$T = T^{**} := (T^*)^*$$

and

$$\langle T^*w, v\rangle_V = \langle w, Tv\rangle_W \qquad \forall w \in W \; \forall v \in V$$

*Proof.* We have

$$\begin{aligned}\langle Tv, w\rangle_W &= \langle v, T^*w\rangle_V = \overline{\langle T^*w, v\rangle_V} \\ &= \overline{\langle w, T^{**}v\rangle_W} = \langle T^{**}v, w\rangle_W\end{aligned}$$

which completes the proof. □

**Definition 23.2** (Isometry) — Let $V, W$ be inner product space over $F$, $T : V \to W$ linear. Then $T$ is called an isometry (or isomorphism of inner product space over $F$) if

1. $T$ is an isomorphism of vector space over $F$
2. $T$ preserves inner products, i.e.,

$$\langle Tv, Tv'\rangle_W = \langle v, v'\rangle_V \qquad \forall v, v' \in V$$

**Remark 23.3.** Let $T : V \to W$ linear of inner product space over $F$. If $T$ preserves inner products, then $T$ is monic.

$$Tv = 0 \iff \|Tv\| = 0 \iff \langle Tv, Tv\rangle = 0 \iff \langle v, v\rangle = 0$$

**Theorem 23.4**

Let $V, W$ be finite dimensional inner product space over $F$ with $\dim V = \dim W$ and $T : V \to W$ linear. Then the following are equivalent

1. $T$ preserves inner product.
2. $T$ is an isometry.
3. If $\mathscr{B} = \{v_1, \ldots, v_n\}$ is an orthogonal basis for $V$, then $\mathscr{C} = \{Tv_1, \ldots, Tv_n\}$ is an orthogonal basis for $W$ and

$$\|Tv_i\| = \|v_i\| \qquad i = 1, \ldots, n$$

4. $\exists$ an orthogonal basis $\mathscr{B} = \{v_1, \ldots, v_n\}$ for $V$ s.t. $\mathscr{C} = \{Tv_1, \ldots, Tv_n\}$ is an orthogonal basis for $W$ with $\|Tv_i\| = \|v_i\|$ $i = 1, \ldots, n$.

*Proof.* 1) $\implies$ 2) $T$ is monic by the remark above, so an isomorphism by the Isomorphism theorem.
2) $\implies$ 3) By the Isomorphism theorem, $\mathscr{C}$ is a basis for $W$ and $\mathscr{C}$ is orthogonal with $\|v_i\| = \|Tv_i\|$ for all $i$.
3) $\implies$ 4) is immediate.
4) $\implies$ 1) By the Isomorphism theorem, $T$ is an isomorphism of vector space over $F$. If $x, y \in V$, let $x = \sum_{i=1}^n \alpha_i v_i$, $y = \sum_{i=1}^n \beta_i v_i$, then

$$\begin{aligned}\langle x, y\rangle &= \sum_{i,j} \alpha_i \overline{\beta_j} \langle v_i, v_j\rangle = \sum_{i,j} \alpha_i \overline{\beta_j} \delta_{ij} \|v_i\|^2 \\ &= \sum_{i,j} \alpha_i \overline{\beta_j} \delta_{ij} \|Tv_i\|^2 = \sum_{i,j} \alpha_i \overline{\beta_j} \delta_{ij} \langle Tv_i, Tv_j\rangle \\ &= \langle Tx, Ty\rangle\end{aligned}$$

□

**Corollary 23.5**

Let $V, W$ be finite dimensional inner product space over $F$ both having orthonormal basis. Then $V$ is isometric to $W$ if and only if $\dim V = \dim W$.

*Proof.* Apply UPVS and the theorem above. □

**Theorem 23.6**

Let $V, W$ be inner product space over $F$, $T : V \to W$ linear. Then $T$ preserves inner products iff $T$ preserves lengths, i.e., $\|Tv\|_W = \|v\|_V$ for all $v \in V$.

*Proof.* " $\implies$ " The result is immediate.
" $\impliedby$ " Let $x, y \in V$ and

$$\begin{aligned}\langle x, y\rangle_V &= \alpha + \beta\sqrt{-1} \\ \langle Tx, Ty\rangle_W &= \gamma + \delta\sqrt{-1}\end{aligned}$$

for $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. We notice that

$$2\alpha = 2\gamma \implies \alpha = \gamma$$

So we are done if $F \subseteq \mathbb{R}$. Suppose $F \nsubseteq \mathbb{R}$, then there exists $0 \neq \mu \in \mathbb{R}$ s.t. $\mu\sqrt{-1} \in F$. Then

$$\begin{aligned}\langle x, \sqrt{-1}\mu y\rangle_V &= -\sqrt{-1}\mu\langle x, y\rangle_V = -\mu\sqrt{-1}\alpha + \beta\mu \\ \langle Tx, \sqrt{-1}\mu Ty\rangle_W &= -\sqrt{-1}\mu\langle Tx, Ty\rangle_W = -\mu\sqrt{-1}\gamma + \delta\mu\end{aligned}$$

Analogous to (*),

$$\beta\mu = \delta\mu, \quad \text{so } \beta = \delta$$

Hence $\langle x, y\rangle_V = \langle Tx, Ty\rangle_W$. □

# §24 Lec 24: May 21, 2021

## §24.1 Inner Product Spaces (Cont'd)

**Definition 24.1** (Unitary Operator) — Let $V$ be an inner product space over $F$, $T : V \to V$ linear. We call $T$ a unitary operator if $T$ is an isometry. If $F \subseteq \mathbb{R}$, such a $T$ is called an orthogonal operator.

**Proposition 24.2**

Let $V$ be an inner product space over $F$, $T : V \to V$ linear. Suppose that $T^*$ exists. Then, $T$ is an isometry if and only if $T^* = T^{-1}$, i.e., $TT^* = 1_V = T^*T$.

*Proof.* " $\implies$ " As $T$ is an isomorphism of vector space over $F$, $T^{-1} : V \to V$ exists and is linear. As $T$ preserves inner products, for all $x, y \in V$

$$\langle Tx, y\rangle = \langle Tx, 1_V y\rangle = \langle Tx, TT^{-1}y\rangle = \langle x, T^{-1}y\rangle$$

It follows that $T^* = T^{-1}$ by uniqueness.
" $\impliedby$ " As $T^*T = 1_V = TT^*$, $T$ is invertible with $T^{-1} = T^*$, so $T$ is an isomorphism. Since

$$\langle Tx, Ty\rangle = \langle x, T^*Ty\rangle = \langle x, y\rangle$$

for all $x, y \in V$. $T$ preserves inner products. □

**Remark 24.3.** Let $V$ be a finite dimensional inner product space over $F$, $T : V \to V$ linear.

1. $T$ is monic iff $T$ is epic iff $T$ is an iso of vector space over $F$.
2. $T$ is unitary $\iff$ $T^*T = 1_V$ $\iff$ $TT^* = 1_V$
3. $T$ is unitary $\iff$ $T^*$ is unitary as $T^{**} = T$

**Definition 24.4** (Unitary Matrix) — Let $F \subseteq \mathbb{C}$, $\overline{F} = F$. We say $A \in \mathbb{M}_n F$ is unitary if $A^*A = I$. Equivalently, $AA^* = I$. Let

$$U_n F \coloneqq \{A \in GL_n F |\, AA^* = I\}$$

If $F \subseteq \mathbb{R}$, we say $A \in \mathbb{M}_n F$ is orthogonal if $A^\top A = I$. Equivalently, $AA^\top = I$. Let

$$O_n F \coloneqq \left\{A \in GL_n F |\, AA^\top = I\right\}$$

**Remark 24.5.**

1. Let $F \subseteq \mathbb{C}$, $F = \overline{F}$, $F^{n\times 1}$, $F^{1\times n}$ inner product space over $F$ via the dot product. If $A \in \mathbb{M}_n F$, then
$$A = [A]_{s_{n,1}} : F^{n\times 1} \to F^{n\times 1}$$
linear and $s_{n,1}$ the ordered standard basis. Then $A$ is unitary iff
   i) The columns of $A$ form an ordered orthonormal basis for $F^{n\times 1}$
   ii) The rows of $A$ form an ordered orthonormal basis for $F^{1\times n}$
2. If $T : V \to V$ is linear, $V$ an inner product space over $F$ with $\dim V = n$, $\mathscr{B}, \mathscr{C}$ ordered orthonormal bases for $V$, then $T$ is unitary iff $[T]_{\mathscr{B},\mathscr{C}}$ is unitary.

## §24.2 Spectral Theory

**Lemma 24.6**

Let $V$ be an inner product space over $F$, $T : V \to V$ linear, $W \subseteq V$ a subspace. Suppose that $T^*$ exists. Then the following is true: If $W$ is $T$-invariant, then $W^\perp$ is $T^*$-invariant.

*Proof.* Let $v \in W^\perp$, $w \in W$, then

$$\langle w, T^*v\rangle = \langle Tw, v\rangle = 0$$

□

**Lemma 24.7**

Let $V$ be a finite dimensional inner product space over $F$, $T : V \to V$ linear. Then the following is true: If $\lambda$ is an eigenvalue of $T$, then $\overline{\lambda}$ is an eigenvalue of $T^*$.

*Proof.* Let $S = T - \lambda 1_V : V \to V$ linear. Then

$$S^* = T^* - \overline{\lambda} 1_V : V \to V \text{ linear}$$

Then $\forall w \in V$,

$$0 = \langle 0, w\rangle = \langle Sv, w\rangle = \langle v, S^*w\rangle$$

Hence $v \perp \text{ im } S^*$ and $v \notin \text{ im } S^*$ as $v \neq 0$. By the Dimension Theorem,

$$0 < \ker S^*, \quad E_{T^*}(\overline{\lambda}) \neq 0$$

□

**Theorem 24.8** (Schur)

Let $V$ be a finite dimensional inner product space over $F$ with $F = \mathbb{R}$ or $\mathbb{C}$ and $T : V \to V$ linear. Suppose that $f_T$ splits in $F[t]$. Then, there exists an ordered orthonormal basis $\mathscr{B}$ for $V$ s.t. $[T]_{\mathscr{B}}$ is upper triangular.

*Proof.* We induct on $n = \dim V$.
$n = 1$ is immediate.
$n > 1$. By the 2nd lemma, $\exists \overline{\lambda} \in F$ and $0 \neq v \in E_{T^*}(\overline{\lambda})$. By the Orthogonal Decomposition Theorem,

$$V = Fv \perp (Fv)^\perp$$

and

$$\dim (Fv)^\perp = \dim V - \dim Fv = n - 1$$

$Fv$ is $T^*$-invariant, hence $(Fv)^\perp$ is $T^{**} = T$-invariant. Let $\mathscr{C}_0$ be an ordered basis for $(Fv)^\perp$. Then $\mathscr{C} = \mathscr{C}_0 \cup \{v_0\}$ is an ordered basis for $V$ and we have

$$[T]_{\mathscr{C}} = \begin{pmatrix} \left[T\big|_{(Fv)^\perp}\right]_{\mathscr{C}_0} & * \\ & * \\ & \vdots \\ & * \\ 0 & [Tv_0]_{\mathscr{C}} \end{pmatrix}$$

By expansion,

$$f_{T\big|_{(Fv)^\perp}} \Big| f_T \in F[t]$$

hence $f_{T\big|_{(Fv)^\perp}} \in F[t]$ splits. By induction, there exists an orthonormal basis $\mathscr{B}_0 = \{v_1, \ldots, v_{n-1}\}$ for $(Fv)^\perp$ s.t. $\left[T\big|_{(Fv)^\perp}\right]_{\mathscr{B}_0}$ is upper triangular. Then $\mathscr{B} = \mathscr{B}_0 \cup \left\{\frac{v}{\|v\|}\right\}$ is an orthonormal basis for $V$ s.t. $[T]_{\mathscr{B}}$ is upper triangular. $\square$

# §25 Lec 25: May 24, 2021

## §25.1 Spectral Theory (Cont'd)

**Definition 25.1** (Hermitian(Self-Adjoint)) — Let $V$ be an inner product space over $F$, $T : V \to V$ linear. Suppose that $T^*$ exists. We say that $T$ is normal

$$TT^* = T^*T$$

and is Hermitian if $T = T^*$, i.e.

$$\langle Tv, w\rangle = \langle v, Tw\rangle \qquad \forall v, w \in V$$

<u>*Note*</u>: If $T$ is Hermitian, $T^*$ exists automatically and $T$ is normal.

**Lemma 25.2**

Let $V$ be an inner product space over $F$, $\lambda \in F$, $0 \neq v \in V$, $T : V \to V$ a normal operator. Then

$$v \in E_T(\lambda) \iff v \in E_{T^*}(\overline{\lambda})$$

*Proof.* Let $S = T - \lambda 1_V$, then $S^* = T^* - \overline{\lambda} 1_V$. It follows that

$$SS^* = S^*S, \quad \text{i.e.} \quad S \text{ is normal}$$

Then

$$\begin{aligned}\|Sv\|^2 &= \langle Sv, Sv\rangle = \langle v, S^*Sv\rangle \\ &= \langle v, SS^*v\rangle = \langle S^*v, S^*v\rangle \\ &= \|S^*v\|^2\end{aligned}$$

So

$$v \in E_T(\lambda) \iff Sv = 0 \iff S^*v = 0 \iff v \in E_{T^*}(\overline{\lambda})$$

□

**Corollary 25.3**

Let $V$ be an inner product space over $F$, $T : V \to V$ normal, $\lambda \neq \mu$ eigenvalue of $T$. Then, $E_T(\lambda)$ and $E_T(\mu)$ are orthogonal. In particular,

$$\sum_\lambda E_T(\lambda) = \frac{1}{\lambda} E_T(\lambda)$$

*Proof.* Let $0 \neq v \in E_T(\lambda)$, $0 \neq w \in E_T(\mu)$. Then by the lemma, $w \in E_{T^*}(\overline{\mu})$ and

$$\begin{aligned}\lambda\langle v, w\rangle &= \langle \lambda v, w\rangle = \langle Tv, w\rangle = \langle v, T^*w\rangle \\ &= \langle v, \overline{\mu} w\rangle = \mu\langle v, w\rangle\end{aligned}$$

As $\lambda \neq \mu$, we obtain $\langle v, w\rangle = 0$. □

**Proposition 25.4**

Let $V$ be a finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to V$ linear, $\mathscr{B}$ an ordered orthonormal basis for $V$ s.t. $[T]_{\mathscr{B}}$ is upper triangular. Then, $T$ is normal if and only if $[T]_{\mathscr{B}}$ is diagonal.

*Proof.* " $\Longleftarrow$ " If

$$[T]_{\mathscr{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

then

$$[T^*]_{\mathscr{B}} = [T]^*_{\mathscr{B}} = \begin{pmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{pmatrix}$$

So

$$\begin{aligned} [TT^*]_{\mathscr{B}} = [T]_{\mathscr{B}} [T^*]_{\mathscr{B}} &= \begin{pmatrix} |\lambda_1|^2 & & 0 \\ & \ddots & \\ 0 & & |\lambda_n|^2 \end{pmatrix} \\ &= [T^*]_{\mathscr{B}} [T]_{\mathscr{B}} \\ &= [T^*T]_{\mathscr{B}} \end{aligned}$$

Hence, $TT^* = T^*T$ by the Matrix Theory Theorem.
" $\Longrightarrow$ " Let $\mathscr{B} = \{v_1, \dots, v_n\}$ be an orthonormal basis for $V$ s.t. $A = [T]_{\mathscr{B}}$ is upper triangular. By the lemma,

$$Tv_1 = A_{11} v_1 \quad \text{and} \quad T^* v_1 = \overline{A_{11}} v_1$$

By definition,

$$T^* v_1 = \sum_{i=1}^n (A^*)_{i1} v_i = \sum_{i=1}^n \overline{A_{1i}} v_i$$

So

$$\overline{A_{1i}} = 0 \qquad \forall i > 1$$

Hence,

$$A_{1i} = 0 \qquad \forall i > 1$$

In particular,

$$A_{12} = 0$$

By the lemma,

$$Tv_2 = A_{22} v_2, \quad \text{hence} \quad T^* v_2 = \overline{A_{22}} v_2$$

The same argument shows $\overline{A_{2i}} = 0$, $i \neq 2$, i.e.,

$$A_{2i} = 0, \quad i \neq 2$$

Continuing this process, we conclude $A$ is diagonal. □

**Theorem 25.5** (Spectral Theorem for Normal Operators)

Let $V$ be a finite dimensional inner product space over $\mathbb{C}$, $T : V \to V$ linear. Then $T$ is normal if and only if there exists an orthonormal basis $\mathscr{B}$ for $V$ consisting of eigenvectors of $T$. In particular, if $T$ is normal, then $T$ is diagonalizable.

*Proof.* This follows immediately by Schur's theorem, FTA, and the above proposition. □

**Remark 25.6.** Let $V$ be a finite dimensional inner product space over $\mathbb{R}$, $T : V \to V$ linear. Suppose that $f_T \in \mathbb{R}[t]$ splits. Then $T$ is normal iff $\exists$ an orthonormal basis $\mathscr{B}$ for $V$ consisting of eigenvectors for $T$.
By Schur's theorem, $T$ is triangularizable via an orthonormal basis for $V$. The same result follows by the proposition in the case $F = \mathbb{R}$.

**Spectral Decomposition and Resolution for Normal Operators**:

Let $V$ be a finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to V$ linear s.t. $f_T$ splits. So $T$ is normal. Let $\lambda_1, \ldots, \lambda_r$ be all the distinct eigenvalues of $T$ in $F$, $\mathscr{C}$ an orthonormal basis for $V$. We know

$$v \in E_T(\lambda_i) \iff v \in E_{T^*}(\overline{\lambda_i}) \qquad \forall i \tag{+}$$

Let $P_i : V \to V$ be the orthogonal projection along $E_T(\lambda_i)^\perp$ for $i = 1, \ldots, r$ omit at $i^{\text{th}}$ onto $E_T(\lambda_i)$.
By (+), $P_i : V \to V$ is also the orthogonal projection along $E_{T^*}(\overline{\lambda_i})^\perp$ onto $E_{T^*}(\overline{\lambda_i})$.
This is a unique decomposition

$$\begin{gathered}
P_{E_T(\lambda_i)} = P_i = P_{E_T^*}(\overline{\lambda_i}) \qquad \forall i \\
TP_i = P_i T \quad \text{and} \quad T^* P_i = P_i T^* \qquad \forall i \\
1_V = P_1 + \ldots + P_r \\
P_i P_j = \delta_{ij} P_i \qquad \forall i \\
T = \lambda_1 P_1 + \ldots + \lambda_r P_r \\
T^* = \overline{\lambda_1} P_1 + \ldots + \overline{\lambda_r} P_r
\end{gathered}$$

Let $\mathscr{B}_i$ be an ordered orthonormal basis for $E_T(\lambda_i)$, so $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_r$ is an ordered orthonormal basis for $V$ with $[T]_{\mathscr{B}}$ and $[T^*]_{\mathscr{B}}$ is diagonal.
Let $\mathscr{Q} = [1_V]_{\mathscr{B},\mathscr{C}}$. Then $\mathscr{Q}$ is unitary as it takes an orthonormal basis to an orthonormal basis, hence

$$\begin{gathered}
\mathscr{Q}^{-1} = \mathscr{Q}^* \\
[T]_{\mathscr{B}} = \mathscr{Q}^* [T]_{\mathscr{C}} \mathscr{Q} \\
[T^*]_{\mathscr{B}} = \mathscr{Q}^* [T^*]_{\mathscr{C}} \mathscr{Q}
\end{gathered}$$

**Theorem 25.7**

Let $V$ be a finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to V$ linear with $f_T \in F[t]$ splits. Then, $T$ is normal if and only if $\exists g \in F[t]$ s.t. $T^* = g(T)$.

# §26 Lec 26: May 26, 2021

## §26.1 Spectral Theory (Cont'd)

> **Remark 26.1.** A rotation $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ by $\angle\theta$, $0 < \theta < 2\pi$, $\theta \neq \pi$ has no eigenvalues, but is normal (with $\mathbb{R}^2$ an inner product space over $\mathbb{R}$ via the dot product) as it is unitary.

**Lemma 26.2**

Let $V$ be an inner product space over $F$, $T : V \to V$ hermitian. If $\lambda$ is an eigenvalue of $T$, then $\lambda \in F \cap \mathbb{R}$.

*Proof.* Let $0 \neq v \in E_T(\lambda)$. Then

$$\begin{aligned}\lambda\|v\|^2 &= \lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Tv, v\rangle \\ &= \langle v, T^*v\rangle = \langle v, Tv\rangle = \langle v, \lambda v\rangle \\ &= \overline{\lambda}\langle v, v\rangle = \overline{\lambda}\|v\|^2\end{aligned}$$

As $\|v\| \neq 0$, $\lambda = \overline{\lambda}$, so it's real. □

**Lemma 26.3**

Let $V$ be a finite dimensional inner product space over $F$ with $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to V$ hermitian. Then $f_T \in F[t]$ splits in $F[t]$.

*Proof.* By previous result, we can assume that $F = \mathbb{R}$. Let $\mathscr{B}$ be an orthonormal basis for $V$. Then

$$A := [T]_{\mathscr{B}} = [T^*]_{\mathscr{B}} = [T]_{\mathscr{B}}^* = A^*$$

in $\mathbb{M}_n\mathbb{R} \subseteq \mathbb{M}_n\mathbb{C}$, $n = \dim V$. As

$$A : \mathbb{C}^{n\times 1} \to \mathbb{C}^{n\times 1} \text{ is Hermitian}$$

$f_A$ splits with real roots by Lemma 26.2. (and FTA), i.e.,

$$f_A = \prod (t - \lambda_i) \in \mathbb{C}[t], \quad \lambda_i \in \mathbb{R} \quad \forall i$$

So $f_T = f_A = \prod(t - \lambda_i) \in \mathbb{R}[t]$ splits. □

**Theorem 26.4** (Spectral Theorem for Hermitian Operators)

Let $V$ be a finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to V$ hermitian. Then, there exists an orthonormal basis for $V$ of eigenvectors of $T$ and all all eigenvalues are real.

*Proof.* If $F = \mathbb{C}$, the result follows by Lemma 26.2 as $T$ is normal. So we may assume $F = \mathbb{R}$. As $f_T \in \mathbb{R}[t]$ splits by Lemma 26.3, there exists an orthonormal basis $\mathscr{B}$ for $V$ s.t. $[T]_{\mathscr{B}}$ is upper triangular by Schur's Theorem. As $T$ is normal, it is diagonalizable. The result follows by Lemma 26.2. □

## §26.2 Hermitian Addendum

**Theorem 26.5**

If $0 \neq V$ is a finite dimensional inner product space over $\mathbb{R}$, $T : V \to V$ hermitian, then $T$ has an eigenvalue.

The proof in Axler's book is very nice, and he does not use determinant theory. He uses the following arguments

1. If $V$ is a finite dimensional vector space over $F$, $T : V \to V$ linear, then there exists $q \in F[t]$ monic s.t. $q(T) = 0$
2. If $0 \neq q \in \mathbb{R}[t]$, then there exists a factorization
$$q = \beta(t-\lambda_1)^{e_1} \dots (t-\lambda_r)^{e_r} q_1^{f_1} \dots q_s^{f_s}$$
in $\mathbb{R}[t]$ with $q_i$ monic irreducible quadratic polynomials in $\mathbb{R}[t]$.

This follows by the FTA.

**Lemma 26.6**

Let $q = t^2 + bt + c$ in $\mathbb{R}[t]$, $b^2 < 4c$, i.e., $q$ is an irreducible monic quadratic polynomial in $\mathbb{R}[t]$. If $V$ is a finite dimensional inner product space over $\mathbb{R}$ and $T : V \to V$ is Hermitian, then $q(T)$ is an isomorphism.

*Proof.* It suffices to show $q(T)$ is a monomorphism by the Isomorphism Theorem. So it suffices to show if $0 \neq v \in V$, then $q(T)v \neq 0$. We have

$$\begin{aligned}
\langle q(T)v, v\rangle &= \langle T^2 v, v\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle \\
&= \langle Tv, Tv\rangle + b\langle Tv, v\rangle + c\langle v, v\rangle \\
&= \|Tv\|^2 + b\langle Tv, v\rangle + c\|v\|^2 \\
&\geq \|Tv\|^2 - |b|\|Tv\|\|v\| + c\|v\|^2 \\
&= \left(\|Tv\| - \frac{|b|\|v\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|v\|^2 > 0
\end{aligned}$$

So $q(T)v \neq 0$. □

*Proof.* (of Theorem) Let $q \in \mathbb{R}[t]$ in 2) satisfy $q(T) = 0$. So

$$0 = q(T) = (T - \lambda_1 1_V)^{e_1} \dots (T - \lambda_r 1_V)^{e_r} q_1(T)^{f_1} \dots q_s(T)^{f_s}$$

As all the $q_i(T)$ are isomorphism, at least one of the $(T - \lambda_i 1_V)$ is not injective, i.e., $\lambda_i$ is an eigenvalue. □

# §27 Lec 27: May 28, 2021

## §27.1 Positive (Semi-)Definite Operators

Let $V$ be a finite dimensional inner product space over $F$, where $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to V$ hermitian, $\mathscr{B} = \{v_1, \dots, v_n\}$ an orthonormal basis of eigenvectors of $T$, i.e.,

$$Tv_i = \lambda_i v_i, \qquad i = 1, \dots, n$$

So $\lambda_i \in \mathbb{R}$, $i = 1, \dots, n$. Suppose $v \in V$. Then

$$v = \sum_{i=1}^{n} \alpha_i v_i, \qquad \alpha_i \in F \ \forall i$$

and

$$\begin{aligned}
\langle Tv, v\rangle &= \langle \sum_{i=1}^{n} T(\alpha_i v_i), \sum_{j=1}^{n} \alpha_j v_j \rangle \\
&= \langle \sum_{i=1}^{n} \lambda_i \alpha_i v_i, \sum_{j=1}^{n} \alpha_j v_j \rangle \\
&= \sum_{i,j=1}^{n} \lambda_i \alpha_i \overline{\alpha_j} \langle v_i, v_j \rangle \qquad (*) \\
&= \sum_{i,j=1}^{n} \lambda_i \alpha_i \overline{\alpha_j} \delta_{ij} \\
&= \sum_{i=1}^{n} \lambda_i |\alpha_i|^2
\end{aligned}$$

**Definition 27.1** (Positive/Negative (Semi-) Definite) — Let $V$ be a finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to V$ hermitian. We say that $T$ is positive or positive definite if

$$\langle Tv, v \rangle > 0 \qquad \forall 0 \neq v \in V$$

and positive semi-definite if

$$\langle Tv, v \rangle \geq 0 \qquad \forall 0 \neq v \in V$$

We can define $T$ as negative (semi-) definite similarly.

It follows from (*) that we have

**Proposition 27.2**

Let $V$ be a finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to V$ hermitian. Then $T$ is positive semi-definite (respectively positive) if and only if all eigenvalues of $T$ are non-negative (respectively positive).

**Question 27.1.** What does this say about the $2^{\text{nd}}$ derivative test for $C^2$ function, $f : S \to \mathbb{R}$ at a critical point in the interior of $S$?

**Theorem 27.3**

Let $V$ be a finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to V$ hermitian. Then $T$ is non-negative (respectively positive) iff $\exists S : V \to V$ non-negative s.t.

$$T = S^2$$

i.e., $T$ has a square root (respectively, and $S$ is invertible).

*Proof.* " $\implies$ " Let $\mathscr{B} = \{v_1, \ldots, v_n\}$ be an ordered orthonormal basis for $V$ of eigenvectors of $T$

$$Tv_i = \lambda_i v_i, \qquad \lambda_i \geq 0 \in \mathbb{R},\ \ i = 1, \ldots, n$$

Then $\exists \mu_i \in \mathbb{R}$, $\mu_i \geq 0$ s.t. $\lambda_i = \mu_i^2$, $i = 1, \ldots, n$. Let

$$B = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} = \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}$$

So

$$B^2 = [T]_{\mathscr{B}}$$

By MTT, $\exists S : V \to V$ linear s.t. $[S]_{\mathscr{B}} = B$. So

$$[T]_{\mathscr{B}} = B^2 = [S]_{\mathscr{B}}^2 = \left[S^2\right]_{\mathscr{B}}$$

Hence $T = S^2$ by MTT. As $\mathscr{B}$ is orthonormal, $\mu_i \in \mathbb{R}$ for all $i$

$$[S^*]_{\mathscr{B}} = [S]_{\mathscr{B}}^* = B^* = B = [S]_{\mathscr{B}}$$

Thus, $S = S^*$ by MTT; so hermitian if $\lambda_i > 0 \forall i$, $\det B \neq 0$, so $B \in GL_n F$.
" $\impliedby$ " Let $\mathscr{B}$ be an ordered orthonormal basis for $V$ of eigenvectors for $S$. Then

$$[S]_{\mathscr{B}} = \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}, \quad \mu_i \geq 0 \in \mathbb{R} \text{ and}$$

$$[T]_{\mathscr{B}} = \left[S^2\right]_{\mathscr{B}} = \begin{pmatrix} \mu_1^2 & & 0 \\ & \ddots & \\ 0 & & \mu_n^2 \end{pmatrix}$$

is diagonal. Therefore, $\mathscr{B}$ is also an orthonormal basis for $V$ of eigenvectors of $T$. As $\mu_i^2 \geq 0$ ($> 0$ if $S$ is invertible), $T$ is non-negative (respectively positive if $S$ is invertible). $\square$

**Theorem 27.4**

Let $V$ be a finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$ and $T : V \to V$ hermitian. Suppose that $T$ is non-negative. Then $T$ has a unique square root $S$, i.e., $S : V \to V$ non-negative s.t. $S^2 = T$.

*Proof.* Let $S^2 = T$, $S : V \to V$ non-negative. The Spectral Theorem gives unique orthogonal decompositions

$$\begin{aligned} V &= E_T(\lambda_1) \perp \ldots \perp E_T(\lambda_r) \\ T &= \lambda_1 P_{\lambda_1} + \ldots + \lambda_r P_{\lambda_r} \\ P_{\lambda_i} P_{\lambda_j} &= \delta_{ij} P_{\lambda_i} P_{\lambda_j}, \qquad \forall i, j \\ 1_V &= P_{\lambda_1} + \ldots + P_{\lambda_r} \end{aligned}$$

and we also have

$$\begin{aligned} V &= E_S(\mu_1) \perp \ldots \perp E_S(\mu_s), \qquad \mu_i \geq 0, \;\; i = 1, \ldots, s \\ S &= \mu_1 P_{\mu_1} + \ldots + \mu_s P_{\mu_s} \\ P_{\mu_i} P_{\mu_j} &= \delta_{ij} P_{\mu_i}, \qquad \forall i, j \\ 1_V &= P_{\mu_1} + \ldots + P_{\mu_s} \end{aligned}$$

In particular,

$$\begin{aligned} S^2 &= (\mu_1 P_{\mu_1} + \ldots + \mu_s P_{\mu_s})(\mu_1 P_{\mu_1} + \ldots + \mu_s P_{\mu_s}) \\ &= \mu_1^2 P_{\mu_1} + \ldots + \mu_s^2 P_{\mu_s} \end{aligned}$$

As $T = S^2$,

$$\mu_1^2 P_{\mu_1} + \ldots + \mu_s^2 P_{\mu_s} = \lambda_1 P_{\lambda_1} + \ldots + \lambda_r \mu_r$$

So by uniqueness, we must have $s = r$ and changing the order if necessary

$$\mu_i^2 = \lambda_i, \qquad P_{\mu_i} = P_{\lambda_i}, \quad \forall i$$

□

**Lemma 27.5**

Let $V, W$ be finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to W$ linear. Then $T^*T : V \to V$ is hermitian and non-negative.

> **Remark 27.6.** If in the definition of positive operator, etc, we omit $V$ being finite dimensional but assume $T^*$ exists, then we would still have $T^*T$ hermitian.

*Proof.* Let $x, y \in V$. Then

$$\langle x, (T^*T)^* y \rangle_V = \langle T^*Tx, y \rangle_V = \langle Tx, Ty \rangle_W = \langle x, T^*Ty \rangle_V$$

Since this is true for all $x, y$

$$(T^*T)^* = (T^*T^{**})^* = T^*T$$

is hermitian, hence has real eigenvalues. Let $\lambda$ be an eigenvalue of $T^*T$, $0 \neq v \in V$ s.t. $T^*Tv = \lambda v$. Then

$$\begin{aligned} \lambda \|v\|_V^2 = \lambda \langle v, v \rangle_V = \langle \lambda v, v \rangle_V &= \langle T^*Tv, v \rangle_V \\ = \langle Tv, Tv \rangle_W &= \|Tv\|_W^2 \geq 0 \end{aligned}$$

So

$$\lambda = \frac{\|Tv\|_W^2}{\|v\|_V^2} \geq 0$$

as $\|v\|_V^2 \neq 0$. □

**Corollary 27.7**

Let $V$ be a finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$, $T : V \to V$ linear. Then $T$ is non-negative (respectively positive) iff $\exists S : V \to V$ linear (respectively an isomorphism) s.t. $T = S^*S$.

*Proof.* Use the theorem and lemma presented above. □

<u>Notation</u>:

- $F = \mathbb{R}$ or $\mathbb{C}$, $A \in F^{m\times n}$
- $A^{(i)}$ = the $i^{\text{th}}$ column of $A$
- $A = \left[A^{(1)} \; \ldots \; A^{(m)}\right]$
- $\langle,\rangle$ = the dot product on $F^N$ for any $N \geq 1$
- $U_N(F) = \left\{U \in GL_N F \middle| U^* = U^{-1}\right\}$

**Definition 27.8** (Pseudodiagonal) — Let $D \in F^{m\times n}$. We call $D$ pseudodiagonal if $D_{ij} = 0$ $\forall i \neq j$, i.e., only $D_{ii}$ can have non-zero entries.

**Theorem 27.9** (Singular Value)
Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in F^{m\times n}$. Then $\exists U \in U_n(F)$, $X \in U_m(F)$ s.t.

$$X^*AU = D = \begin{pmatrix} \mu_1 & & & & & 0 \\ & \ddots & & & & \\ & & \mu_r & & & \\ & & & 0 & & \\ 0 & & & & & \ddots \end{pmatrix} \in F^{m\times n}$$

is a pseudodiagonal matrix satisfying

$$\mu_1 \geq \ldots \geq \mu_r > 0$$

and

$$r = \operatorname{rank}(A)$$

*Proof.* By the lemma, $A^*A \in \mathbb{M}_n F$ is hermitian and has non-negative eigenvalues. Let $\lambda_1, \ldots, \lambda_r$ be the positive eigenvalues ordered s.t.

$$\lambda_1 \geq \ldots \geq \lambda_r > 0$$

By the Spectral Theorem for Hermitian Operators, $\exists U \in U_n F$ s.t.

$$(AU)^*(AU) = U^*A^*AU = \begin{pmatrix} \lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix}$$

in $\mathbb{M}_n F$. Let $C = AU \in F^{m\times n}$. Then

$$C^*C = (AU)^*(AU) \in \mathbb{M}_n F$$

Write

$$\lambda_i = \mu_i^2, \quad \mu_i > 0, \qquad 1 \leq i \leq r$$

So

$$\mu_1 \geq \ldots \geq \mu_r > 0$$

Set

$$B = \begin{pmatrix} \mu_1 & & & & & 0 \\ & \ddots & & & & \\ & & \mu_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} \in \mathbb{M}_n F$$

If $i > r_1$ let $\lambda_i = 0$. Then, we have

$$\begin{aligned} \lambda_i \delta_{ij} = (C^* C)_{ij} &= \sum_l (C^*)_{il} C_{lj} = \sum_l \overline{C_{li}} C_{lj} \\ &= \sum_l C_{lj} \overline{C_{li}} = \langle C^{(j)}, C^{(i)} \rangle \end{aligned}$$

Hence

$$C = \begin{bmatrix} C^{(1)} & \ldots & C^{(r)} & 0 & \ldots & 0 \end{bmatrix}$$

We continue with the proof in the next lecture. $\square$

# §28 Lec 28: Jun 2, 2021

## §28.1 Positive (Semi-)Definite Operators (Cont'd)

*Proof.* (Cont'd) Recall, we have proven so far

$$C = \begin{bmatrix} C^{(1)} & \dots & C^{(r)} & 0 & \dots & 0 \end{bmatrix}$$

and thus $\{C^{(1)}, \dots, C^{(r)}\}$ is an orthogonal set in $F^{m\times 1}$. As $C^{(i)} \neq 0$, $i = 1, \dots, r$, $C^{(1)}, \dots, C^{(r)}$ are linearly independent. In particular,

$$\text{rank } C = r$$

We also have

$$\|C^{(i)}\|^2 = \langle C^{(i)}, C^{(i)} \rangle = \lambda_i = \mu_i^2$$

for $i = 1, \dots, m$. As $U$ is invertible,

$$\text{rank } A = \text{rank } AU = \text{rank } C = r$$

So rank $A = r$ as needed.
Now let

$$X^{(i)} := \frac{1}{\mu_i} C^{(i)}, \quad i = 1, \dots, r$$

Then $\{X^{(1)}, \dots, X^{(r)}\}$ is an orthonormal set. Extend this to an orthonormal basis $\mathscr{B} = \{X^{(1)}, \dots, X^{(m)}\}$. Then

$$X = \begin{bmatrix} X^{(1)} & \dots & X^{(m)} \end{bmatrix} = [1_{F^{m\times 1}}]_{\mathscr{S}_{m,1}, \mathscr{B}}$$

Since both $\mathscr{S}_{m,1}$ and $\mathscr{B}$ are orthonormal bases, $X \in U_m(F)$. Let $D$ be the pseudo-diagonal matrix

$$D := \begin{pmatrix} \mu_1 & & & & 0 \\ & \ddots & & & \\ & & \mu_r & & \\ & & & 0 & \\ 0 & & & & \ddots \end{pmatrix} \in F^{m\times n}$$

as in the statement of the theorem. Then

$$\begin{aligned} XD &= \begin{bmatrix} X^{(1)} & \dots & X^{(m)} \end{bmatrix} \begin{pmatrix} \mu_1 & & & & \\ & \ddots & & & \\ & & \mu_r & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix} \\ &= \begin{bmatrix} \mu_1 X^{(1)} & \dots & \mu_r X^{(r)} & 0 & \dots & 0 \end{bmatrix} \\ &= C = AU \end{aligned}$$

Hence

$$X^* AU = D$$

as needed. □

**Definition 28.1** (Singular Value Decomposition) — Let $A \in F^{m\times n}$, $F = \mathbb{R}$ or $\mathbb{C}$.

$$A = XDU^*, \quad U \in U_n F, \quad X \in U_m F$$

$$D = \begin{pmatrix} \mu_1 & & & & 0 \\ & \ddots & & & \\ & & \mu_r & & \\ & & & 0 & \\ 0 & & & & \ddots \end{pmatrix} \in F^{m\times n} \tag{*}$$

$$\mu_1 \geq \ldots \geq \mu_r > 0 \in \mathbb{R}$$

Then (*) is called a singular value decomposition (SVD) of $A$, $\mu_1, \ldots, \mu_r$ are the singular values of $A$, $D$ is the pseudo-diagonal matrix of $A$.

*Note*: Let $A = XDU^*$ be an SVD of $A$. Then

1. The singular values of $A$ are the (positive) square roots of the positive eigenvalues of $A^*A$.
2. The columns of $X$ form an orthonormal basis for $F^{m\times 1}$ of eigenvectors of $AA^*$.
3. The columns of $U$ form an orthonormal basis for $F^{n\times 1}$ of eigenvectors of $A^*A$.

**Corollary 28.2**

The singular values of $A \in F^{m\times n}$, $F = \mathbb{R}$ or $\mathbb{C}$ are unique (including multiplicity) up to order.

*Proof.* Let $A = XDU^*$ be a SVD of $A$, $X \in U_m F$, $U \in U_n F$. Then

$$A^*A = (XDU^*)^*(XDU^*) = UD^*X^*XDU^* = UD^*DU^*$$

as $X^*X = I$. So

$$A^*A \sim D^*D = \begin{pmatrix} \alpha_{11}^2 & & \\ & \ddots & \\ & & \ddots \end{pmatrix} \in \mathbb{M}_n F$$

have the same eigenvalues $\alpha_{11}^2, \ldots,$ as $A^*A$. □

**Remark 28.3.** An SVD of $A \in F^{m\times n}$, $F = \mathbb{R}$ or $\mathbb{C}$ may not be unique.

**Corollary 28.4**

The singular values of $A \in F^{m\times n}$, $F = \mathbb{R}$ or $\mathbb{C}$ are the same as the singular values of $A^* \in F^{n\times m}$.

*Proof.* $(XDU^*)^* = UD^*X^*$ and $D, D^*$ have the same non-zero diagonal eigenvalues. □

The abstract version of the singular value theorem is

**Theorem 28.5** (Singluar Value - Linear Transformation Form)
Let $F = \mathbb{R}$ or $\mathbb{C}$, $V$ a finite dimensional inner product space over $F$ and $T : V \to W$ linear of rank $r$. Then there exists orthonormal basis

$$\mathscr{B} = \{v_1, \dots, v_n\} \text{ for } V$$
$$\mathscr{C} = \{w_1, \dots, w_m\} \text{ for } W$$
$$\mu_1 \geq \ldots \geq \mu_r > 0 \in \mathbb{R}$$

satisfying

$$Tv_i = \begin{cases} \mu_i w_i, & i = 1, \dots, r \\ 0, & i > r \end{cases}$$

Conversely, suppose the above conditions are all satisfied. Then $v_i$ is an eigenvector for $T^*T$ with eigenvalue $\mu_i^2$ for $i = 1, \dots, r$ and eigenvalue $0$ for $i = r+1, \dots, n$. In particular, $\mu_1, \dots, \mu_r$ are uniquely determined.

*Proof.* Left as exercise. □

**Remark 28.6.** So we see for an arbitrary linear transformation $T : V \to W$ of finite dimensional inner product space over $F$, $F = \mathbb{R}$ or $\mathbb{C}$, singular values can be viewed as a substitute for eigenvalues.

When $F = \mathbb{R}$ or $\mathbb{C}$ and $A \in \mathbb{M}_n F$, we get a generalization of the polar representation of eigenvalues $z \in \mathbb{C}$ where $z = re^{\sqrt{-1}\theta}$.

**Theorem 28.7** (Polar Decomposition)
Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in \mathbb{M}_n F$. Then there exists $\tilde{U} \in U_n F$, $N \in \mathbb{M}_n F$ hermitian with all its eigenvalues real and non-negative satisfying

$$A = \tilde{U} N$$

here $N \leftrightarrow r$, $\tilde{U} \leftrightarrow e^{\sqrt{-1}\theta}$ for $n = 1$.

*Proof.* In the singular value theorem, we have $m = n$. Let $A = XDU^*$ be an SVD, $X, U \in U_n F$. We have $D = D^*$ is hermitian with non-negative eigenvalues. So

$$A = XDU^* = X(U^*U)DU^* = (XU^*)(UDU^*)$$

Since

$$(XU^*)^*(XU^*) = UX^*XU^* = UU^* = I$$

$XU^* \in U_n F$ also. Let $\tilde{U} = XU^* \in U_n F$, $N = UDU^*$ which completes the proof. □

## §28.2 Least Squares

We give an application of SVD

**Problem 28.1.** Let $F = \mathbb{R}$ or $\mathbb{C}$, $V$ a finite dimensional inner product space over $F$, $W \subseteq V$ a subspace. Let

$$P_W : V \to V \text{ by } v \mapsto v_W$$

be the orthogonal projection of $V$ onto $W$. By the Approximation Theorem, $v_W$ is the best approximation of $v \in V$ onto $W$. Now let $X$ be another finite dimensional inner product space over $F$ and $T : X \to V$ linear with $W = T(X) = \operatorname{im} T$. Let $v \in V$ and $x \in X$. We call

i) $x$ a best approximation to $v$ via $T$ if

$$Tx = v_W = P_W(v)$$

ii) $x$ an optimal approximation to $v$ via $T$ if it is a best approximation to $v$ via $T$ and $\|x\|$ is minimal among all best approximation to $v$ via $T$.

Find an optimal approximation.

<u>**Solution**</u>:

$$\langle x, T^* y \rangle_X = \langle Tx, y \rangle_V,$$

we have

$$W^\perp = (\text{im } T)^\perp = \ker T^*$$

Since

$$v - v_W \in W^\perp = (\text{im } T)^\perp \qquad \text{(by the OR Decomposition Theorem)}$$

and

$$T^* v = T^* v_W$$

So if $x$ is a best approximation of $v$ via $T$, then

$$T^* T x = T^* v \tag{*}$$

i.e., $x$ is also a solution to $T^* T x = T^* v$. Conversely, if (*) holds, then

$$Tx - v \in \ker T^* = (\text{im } T)^\perp = W^\perp$$

In particular,

$$\begin{aligned} v_W = P_W v &= P_W \left( Tx - (Tx - v) \right) \\ &= P_W(Tx) - P_W(Tx - v) \\ &= Tx + 0 = Tx \end{aligned}$$

<u>Conclusion</u>: $x$ is a best approximation to $v$ via $T$ if and only if $T^* T x = T^* v$.

**Claim 28.1.** Suppose that $T$ is monic. Then

$$T^* T : X \to X \text{ is an isomorphism}$$

and

$$P_W = T \left( T^* T \right)^{-1} T^* : V \to V \tag{+}$$

Suppose that $x \in X$ satisfies $T^* T x = 0$. Then

$$0 = \langle T^* T x, x \rangle_X = \langle Tx, Tx \rangle_V = \|Tx\|_V^2 \tag{$\star$}$$

Therefore, $Tx = 0$. But $T$ is monic, so $x = 0$. Hence $T^* T : V \to V$ is monic hence an isomorphism. We now show (+) holds.
Let $v \in V$. Since $T^* T$ is an isomorphism, there exists $x \in X$ s.t.

$$T^* T x = T^* v \tag{$\star\star$}$$

and

$$\begin{aligned} T(T^* T)^{-1} T^* v &= T(T^* T)^{-1} T^* T x \\ &= Tx = v_W = P_W(v) \end{aligned}$$

showing (+). This proves the claim and also shows that the $x$ in ($\star\star$) is a best approximation to $v$ via $T$.

# §29 Lec 29: Jun 4, 2021

## §29.1 Least Squares (Cont'd)

**Claim 29.1.** Let $v \in V$. Then $\exists! x \in X$ an optimal approximation to $v$ via $T$. Moreover, this $x$ is characterized by

$$P_Y(x) = 0 \text{ where } Y = \ker T^*T$$

Let $x, x'$ be two best approximation to $v$ via $T$. Then,

$$T^*Tx = T^*v = T^*Tx'$$

Therefore,

$$x - x' \in \ker T^*T =: Y$$

It follows if $x$ is a best approximation to $v$ via $T$, then any other is of the form $x + y$, $y \in Y$. We also have for such $x + y$

$$P_Y(x+y) = P_Y(x) + P_Y(y) = P_Y(x) + y$$

Let $x'' = x - P_Y(x)$. Then

$$P_Y(x'') = P_Y(x) - P_Y^2(x) = 0, \quad \text{i.e., } x'' \perp Y$$

So

$$\|x'' + y\|^2 = \|x''\|^2 + \|y\|^2 \geq \|x''\|^2 \qquad \forall y \in Y$$

by the Pythagorean Theorem. Hence, $x'' = P_{Y^\perp}(x)$ is the unique optimal approximation. This proves the claim above.
Let $A = T : F^{n\times 1} \to F^{m\times 1}$, $A \in F^{m\times n}$, $v \in F^{m\times 1}$ with $F = \mathbb{R}$ or $\mathbb{C}$. Let

$$A = XDU^*, \quad D = \begin{pmatrix} \mu_1 & & & & \\ & \ddots & & & \\ & & \mu_r & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix} \in F^{m\times n}$$

and

$$\mu_1 \geq \ldots \geq \mu_r > 0 \in \mathbb{R}$$

be an SVD. Let's define

$$D^\dagger := \begin{pmatrix} \mu_1^{-1} & & & & \\ & \ddots & & & \\ & & \mu_r^{-1} & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix} \in F^{n\times m}$$

Then

$$A^\dagger := UD^\dagger X^* \in F^{n\times m}$$

is called the Moore-Penrose generalized pseudoinverse of $A$. Then the following are true

i) $\operatorname{rank}(A) = \operatorname{rank}(A^\dagger)$

ii) $A^\top v$ is an optimal approximation in $F^{n\times 1}$ to $v$ via $A$ and is unique.

iii) If $\text{rank}(A) = n$, then
$$A^\dagger = (A^*A)^{-1}A^*$$

*Proof.* i) $\text{rank}(A) = \text{rank}(D) = \text{rank}(D^\dagger) = \text{rank}(A^\dagger)$ as $X, U$ are invertible.

ii) **<u>Case 1:</u>** $A = D$, i.e., $X, U$ are the appropriate identity matrices. Let $W = \text{im } A$, $U = \ker D^\dagger D$, $W = \text{span}\left\{e_i \in \mathscr{S}_{m,1} | D_{ii} \neq 0\right\}$
If $v \in F^{m\times 1}$, then
$$v_W = P_W(v) = DD^\dagger v = D\left(D^\dagger v\right)$$
So $D^\dagger v$ is a best approximation to $v$ relative to $D$. As
$$U = \ker D^\dagger D = \text{Span}\left\{e_j \in \mathscr{S}_{n,1} | D_{jj} = 0\right\}$$
and we have
$$D^\dagger v \in \text{Span}\left\{e_j \in \mathscr{S}_{n,1} | D_{jj} \neq 0\right\} = Y^\perp,$$
and $P_Y\left(D^\dagger v\right) = 0$.
$$D^\dagger v \text{ is optimal approximation to } v \text{ relative to } D$$

**<u>Case 2:</u>** $A = XDU^*$ in general. $X, U$ are unitary, so they preserve dot products, so $z$ is an optimal approximation to $v$ relative to $A = AUU^*$ if and only if $U^*z$ is an optimal approximation to $v$ relative to $AU$ (*). We also have
$$\begin{aligned}\|Az - v\| &= \|XDU^*z - v\| = \|X^*\left(XDU^*z - v\right)\| \\ &= \|DU^*z - X^*v\|\end{aligned}$$
So (*) is true iff $U^*z$ is an optimal approximation to $X^*v$ relative to $D$. By case 1, $D^\dagger X^*v$ is an optimal approximation to $X^*v$ relative to $D$. As $A^\dagger = UD^\dagger X^*$
$$D\left(D^\dagger X^*v\right) \overset{\text{SVD}}{=} \left(X^*AU\right)\left(D^\dagger X^*v\right) = X^*A\left(A^\dagger v\right)$$
Therefore, $A^\dagger v$ is the optimal approximation to $X^*v$ relative to $X^*A$. Thus, as $X^*$ is an isometry, $A^\dagger v$ is the optimal approximation to $v$ relative to $A$.

iii) This follows as in (ii) for if $\text{rank}(A) = n$, then $(A^*A)^{-1}A^*v$ is the unique optimal best approximation to $Az = v$. □

<u>Warning:</u> In general, $(AB)^\dagger \neq B^\dagger A^\dagger$.
Let $A \in F^{m\times n}$, $F = \mathbb{R}$ or $\mathbb{C}$. Solve
$$AX = B \text{ for } X \in F^{n\times 1}$$
for $X \in F^{n\times 1}$. As $A$ can be inconsistent, we want an optimal approximation to a solution.

**Example 29.1**

Let $F = \mathbb{R}$ or $\mathbb{C}$. Given data $(x_1, y_1), \ldots, (x_n, y_n)$ in $F^2$, find the best line relative to this data, i.e., find

$$y = \lambda x + b, \quad \lambda = \text{ slope}$$

Let

$$A = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}, \quad X = \begin{pmatrix} \lambda \\ b \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

To solve $AX = Y$, we want the optimal solution

$$\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} \lambda \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Let $W = \operatorname{im} A$. To find the optimal approximation to $AX = Y_W$, $X = A^\dagger Y$ works. But $\operatorname{rank}(A) = 2$ is most probable

$$X = (A^* A)^{-1} A^* Y$$

## §29.2 Rayleigh Quotient

Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in \mathbb{M}_n F$. The euclidean norm of $A$ is defined by

$$\|A\| := \max_{0 \neq v \in F^{n \times 1}} \frac{\|Av\|}{\|v\|}$$

If $A \in \mathbb{M}_n F$ is hermitian, then the Rayleigh Quotient of $A$

$$R(v) = R_A(v) : F^{n \times 1} \setminus \{0\} \to \mathbb{R}$$

is defined by

$$R(v) := \frac{\langle Av, v \rangle}{\|v\|^2}$$

Rayleigh quotients are used to approximate eigenvalues of hermitian $A \in \mathbb{M}_n F$.

**Theorem 29.2**

Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in \mathbb{M}_n F$ hermitian. Then,

i) $\max_{v \neq 0} R(v)$ is the largest eigenvalue of $A$.

ii) $\min_{v \neq 0} R(v)$ is the smallest eigenvalue of $A$.

*Proof.* By the Spectral Theorem, $\exists$ an orthonormal basis $\{v_1, \ldots, v_n\}$ of eigenvectors for $A$ with $Av_i = \lambda v_i$, $i = 1, \ldots, n$. We may assume

$$\lambda_1 \geq \ldots \geq \lambda_n \in \mathbb{R}$$

i) Let $v \in F^{n \times 1}$ and $v = \sum_{i=1}^n \alpha_i v_i$, $\alpha_i \in F$, $i = 1, \ldots, n$. Then

$$\begin{aligned} R(v) = \frac{\langle Av, v \rangle}{\|v\|^2} &= \langle \sum_{i=1}^n \alpha_i \lambda_i v_i, \sum_{j=1}^n \alpha_j v_j \rangle / \|v\|^2 \\ &= \frac{\sum_{i,j=1}^n \lambda_i \alpha_i \overline{\alpha_j} \delta_{ij} \langle v_i, v_j \rangle}{\|v\|^2} = \frac{\sum_{i=1}^n \lambda_i - |\alpha_i|^2}{\|v\|^2} \end{aligned}$$

By the Pythagorean Theorem

$$\sum_{i=1}^{n} |\alpha_i|^2 = \|v\|^2$$

So

$$R(v) \le \frac{\sum_{i=1}^{n} \lambda_1 |\alpha_i|^2}{\|v\|^2} = \frac{\lambda_1 \|v\|^2}{\|v\|^2} = \lambda_1$$

ii) Prove similarly. $\square$

**Corollary 29.3**

Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in \mathbb{M}_n F$. Then $\|A\| < \infty$. Moreover, if $\mu$ is the largest singular value of $A$, then

$$\|A\| = \mu$$

*Proof.* Consider:

$$0 \le \frac{\|Av\|^2}{\|v\|^2} = \frac{\langle Av, Av\rangle}{\|v\|^2} = \frac{\langle A^* Av, v\rangle}{\|v\|^2}$$

for all $v \neq 0$. Since $A^*A$ is non-negative, the result follows. $\square$

We know that the singular value of $A \in F^{m\times n}$ are the same as for $A^* \in F^{n\times m}$ if $F = \mathbb{R}$ or $\mathbb{C}$. Therefore,

**Corollary 29.4**

Let $A \in GL_n F, F = \mathbb{R}$ or $\mathbb{C}$, $\mu$ the smallest singular value of $A$. Then

$$\|A^{-1}\| = \frac{1}{\sqrt{\mu}}$$

*Proof.* If $B \in GL_n F$ has an eigenvalue $\lambda \neq 0$, $0 \neq v \in E_B(\lambda)$, then

$$Bv = \lambda v, \quad \text{so } \frac{1}{\lambda} v = B^{-1} v$$

Hence if

$$\mu_1 \ge \ldots \ge \mu_n > 0$$

are the singular values of $A$,

$$\mu_n \ge \ldots \ge \mu_1 > 0$$

are the singular values of $A^{-1}$ as $(A^{-1})^* A^{-1} = (AA^*)^{-1}$. $\square$

# §30 Additional Materials: Jun 04, 2021

## §30.1 Conditional Number

Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in GL_n F$, $b \neq 0$ in $F^{n\times 1}$. Suppose $Ax = b$.

**Problem 30.1.** What happens if we modify $x$ a bit, i.e., by $\delta x \in F^{n\times 1}$. Then we get a new equation

$$A(x + \delta x) = b + \delta b, \qquad \delta b \in F^{n\times 1}$$

and we would like to understand the variance in $b$.

Since $A$ is linear,

$$A(x + \delta x) = b + A(\delta x)$$

i.e.

$$A(\delta x) = \delta b \text{ or } \delta x = A^{-1}(\delta b)$$

and we know, therefore, that

$$\begin{aligned} \|b\| &= \|Ax\| \leq \|A\| \cdot \|x\| \\ \|\delta\| &= \|A^{-1}(\delta b)\| = \|A^{-1}\| \cdot \|\delta b\| \end{aligned}$$

Therefore,

$$\begin{aligned} & \frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|} \text{ as } \|x\| \neq 0 \qquad (b \neq 0) \\ \implies & \frac{\|\delta x\|}{\|x\|} \leq \frac{\|A^{-1}\|\|\delta b\|}{1} \cdot \frac{\|A\|}{\|b\|} = \|A\|\|A^{-1}\| \frac{\|\delta b\|}{\|b\|} \end{aligned}$$

Similarly,

$$\frac{1}{\|A\|\|A^{-1}\|} \frac{\|\delta b\|}{\|b\|} \leq \frac{\|\delta x\|}{\|x\|}$$

We call the number $\|A\|\|A^{-1}\|$ the Conditional Number of $A$ and denote it $\text{cond}(A)$.

**Theorem 30.1**

Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in GL_n F$, $b \neq 0$ in $F^{n\times 1}$. Then

1. $\frac{1}{\text{cond}(A)} \frac{\|\delta b\|}{\|b\|} \leq \frac{\|\delta x\|}{\|x\|} \leq \text{cond}(A) \frac{\|\delta b\|}{\|b\|}$
2. Let $\mu_1 \geq \ldots \geq \mu_r > 0$ be the singular values of $A$. Then

$$\text{cond}(A) = \frac{\mu_1}{\mu_n}$$

*Proof.* 1. from the computation above.

2. follows over computation on the Rayleigh function.

□

**Remark 30.2.** From the theorem,

1. If $\text{cond}(A)$ is close to one, then a small relative error in $b$ forces a small relative error in $x$.
2. If $\text{cond}(A)$ is large, even a small relative error in $x$ may cause a relatively large error in $b$.

> **Remark 30.3.** If there is an error $SA$ of $A$, things would get more complicated. For example, $A+\delta A$ may no longer be invertible.

There exist conditions that can control this. For example, if $A+SA \in GL_nF$, $F=\mathbb{R}$ or $\mathbb{C}$, it is true that

$$\frac{\|\delta x\|}{\|x+\delta x\|} \leq \operatorname{cond}(A)\frac{\|\delta A\|}{\|A\|}$$

One almost never computes $\operatorname{cond}(A)$, as error arises trying to compute it as we need to compute the singular values. However, in some cases, remarkable estimates can be found.

## §30.2 Mini-Max

Let $F=\mathbb{R}$ or $\mathbb{C}$, $A \in \mathbb{M}_nF$. We want a method to compute its eigenvalues if $A$ is hermitian. Since $A$ is hermitian, by the Spectral Theorem,

$$U^*AU = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad U \in U_nF$$

where $A=[A]_{\mathscr{S}_{n,1}}$.
$\mathscr{B}=\{v_1,\ldots,v_n\}$ is an ordered orthonormal basis of eigenvectors for $V=F^{n\times 1}$ satisfying

$$Av_i=\lambda_i v_i$$

So

$$v_i = \text{the } i^{\text{th}} \text{ column of } U^*$$

We let the order be s.t.

$$\lambda_1 \geq \ldots \geq \lambda_n$$

As $(Fv_1)^{\perp}$ is $A$-invariant, $A\big|_{(Fv_1)^{\perp}}$ has maximum eigenvalue $\lambda_2$ obtained from $v_2$, i.e.,

$$\max_{x\in(Fv_1)^{\perp}} R_A(x)=\lambda_{n-1}$$

is obtained from $x=v_2$. The constraint is

$$\langle x, v_1\rangle = 0$$

We can obtain $\lambda_{n-1}$ without knowing $v_1$ or $\lambda_1$. Let $x\in V$ be constrained by $\langle x,z\rangle=0$, some $z\neq 0$. Let $y=U^*x$. Then $\langle x,z\rangle=0$ is equivalent to $\langle y,w\rangle=0$ where $w=Uz$. Computation shows the Rayleigh quotient $R_U$ for $U$ satisfies

$$\max_{\substack{y\\ \langle y,w\rangle=0}} R_U(y)\leq \lambda_n$$

$$\max_{\substack{y\\ \langle y,w\rangle=0}} R_U(y)\geq \lambda_{n-1}$$

So

$$\min_{w\neq 0}\max_{\substack{y\\ \langle y,w\rangle=0}} R_U(y)\geq \lambda_{n-1}$$

gives an upper and lower bound for $R_U(y)$. Let

$$y=\begin{pmatrix} y_1\\ \vdots \\ y_n\end{pmatrix}, \quad \tilde{y}=\begin{pmatrix} y_1\\ y_2\\ 0\\ \vdots\\ 0\end{pmatrix}$$

with $\langle \tilde{y}, w \rangle = 0$. In addition, computation shows,

$$R_U(\tilde{y}) = \lambda_2$$

Let $w = e_1$. Then

$$\max_{\substack{y \\ \langle y, e_1 \rangle}} R_U(y) = \lambda_2$$

So

$$\min_{w \neq 0} \max_{\substack{y \\ \langle y, w \rangle = 0}} R_U(y) = \lambda_2$$

and

$$\min_{w_1, w_2 \neq 0} \max_{\substack{y \\ \langle y, w_1 \rangle = 0 \\ \langle y, w_2 \rangle = 0}} R_U(y) = \lambda_3$$

Proceed inductively.

**Theorem 30.4** (Minimax Principle)
Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in \mathbb{M}_n F$ hermitian with eigenvalues

$$\lambda_1 \geq \ldots \geq \lambda_n$$

Then

$$\min_{z_1, \ldots, z_k \neq 0} \max_{\substack{\langle x, z_1 \rangle = 0 \\ \vdots \\ \langle x, z_k \rangle = 0}} R_A(x) = \lambda_k$$

**Remark 30.5.** The minimax principle is also formulated by

$$\min_{V_j} \max_{x \in V_j} R_A(x) = \lambda_{n-j}, \qquad j = 1, \ldots, n$$

where $V_j$ denotes an arbitrary subspace of $\dim j$.

## §30.3 Uniqueness of Smith Normal Form

Consult Professor Elman's notes.