Math 115A(H)B – (Honors) Linear Algebra

University of California, Los Angeles

Duc Vu

Instructed by Richard Elman

Fall 2020 – Spring 2021

Contents

List of Theorems

List of Definitions

I

About the notes

This is math 115AH & 115B – Undergraduate (Honors) Linear Algebra sequence at UCLA. We meet weekly on MWF from 2:00pm – 2:50pm for lectures. There are two textbooks for the classes, *Linear Algebra* by *Hoffman & Kunze* used in 115AH and *Linear Algebra* by *Friedberg, Incel & Spence* which is optional for 115B. Keep in mind that there are a total of 57 official lectures; the first 28 are for 115AH, and the rest of them is from 115B with a few extra lectures provided by Professor Elman. Thus, the lecture number would be adjusted accordingly for each class. In addition, there are some overlaps in the definition and theorem listed above since a few materials covered in 115AH are supposed to be taught in 115B. All the typos/errors in the notes are my responsibility, and please let me know through my email if you spot any of them. Additional details with regard to note taking in live lecture and other course notes can also be found at my blog site.

II

115AH Lectures

§1 Lec 1: Oct 2, 2020

Remark 1.1. To know a definition, theorem, lemma, proposition, corollary,etc., you must

1. Know its precise statement and what it means without any mistake
2. Know explicit example of the statement and specific examples that do not satisfy it
3. Know consequences of the statement
4. Know how to compute using the statement
5. At least have an idea why you need the hypotheses – e.g., know counter-examples,. . .
6. Know the proof of the statement
7. Know the important (key) steps of in the proof, separate from the formal part of the proof – i.e., the main idea(s) of the proof

THIS IS NOT EASY AND TAKES TIME – EVEN WHEN YOU THINK THAT YOU HAVE MASTERED THINGS.

§1.1 Field

What are the properties of the REAL NUMBERS?

$$\mathbb{R} := \{x | x \text{ is a real no.}\}$$

– at least algebraically?
There are two FUNCTIONS (or MAPS)

- $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ called ADDITION write $a + b := +(a, b)$
- $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ called MULTIPLICATION write $a \cdot b := \cdot(a, b)$

that satisfy certain rule e.g., associativity, commutativity,. . .

Definition 1.2 (Field) — A set F is called a FIELD if there are two functions

- Addition: $+ : F \times F \to F$, write $a + b := +(a, b)$
- Multiplication: $\cdot : F \times F \to F$, write $a \cdot b := \cdot(a, b)$

satisfying the following AXIOMS(A: addition, M: multiplication, D: distributive)

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

A2 $\exists$ an element $0 \in F \ni a + 0 = a = 0 + a$ Existence of a Zero

A3 $\forall x \in F \exists y \in F \ni x + y = 0 = y + x$ Existence of an Additive Inverse

A4 $a + b = b + a$ Commutativity

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

M2 (A2) holds and $\exists$ an element $\in F$ with $1 \neq 0 \ni a \cdot 1 = a = 1 \cdot a$ Existence of a One

M3 (M2) holds and $\forall 0 \neq x \in F \ \ \exists y \in F \ni xy = 1 = yx$ Existence of a Multiplicative Inverse

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

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

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

<u>Comments</u>: Let F be a field, $a, b \in F$. Then the following are true

1. $F \neq \emptyset$ (F at least has 2 elements)
2. 0 and 1 are unique
3. If $a + b = 0$, then b is unique write b as $-a$:
   if $a + b = a + c$, then

$$\begin{aligned} b &= b + 0 \\ &= b + (a + c) \\ &= (b + a) + c \\ &= (a + b) + c \\ &= 0 + c \\ &= c \end{aligned}$$

4. if $a + b = a + c$, then $b = c$
5. if $a \neq 0$ and $ab = 1 = ba$, then b is unique write a^{-1} for b.
6. $0 \cdot a = 0 \forall a \in F$

$$0 \cdot a + 0 \cdot a = (0 + 0) \cdot a = 0 \cdot a = 0 \cdot a + 0$$

   so $0 \cdot a = 0$ by 3.
7. if $a \cdot b = 0$, then $a = 0$ or $b = 0$. If $a \neq 0$, then $0 = a^{-1}(ab) = (a^{-1}a)b = 1b = b$

8. if $a \cdot b = a \cdot c,\ a \neq 0$, then $b = c$
9. $(-a)(-b) = ab$
10. $-(-a) = a$
11. if $a \neq 0$, then $a^{-1} \neq 0$ and $(a^{-1})^{-1} = a$

Example 1.3

$$\mathbb{Q} := \left\{ \frac{a}{b} | a, b \in \mathbb{Z}, b \neq 0 \right\}$$
$$\mathbb{R} := \text{set of real no.}$$
$$\mathbb{C} := \{a + bi | a, b \in \mathbb{R}\} \text{ with}$$

$$(a + b\sqrt{-1} + (c + d\sqrt{-1}) = (a + c) + (b + d)\sqrt{-1}$$
$$(a + b\sqrt{-1}) \cdot (c + d\sqrt{-1}) = (ac - bd) + (ad + bc)\sqrt{-1}$$

$\forall a, b, c, d \in \mathbb{R}$
Under usual $+, \cdot$ of C

$$\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

are all field and we say $\mathbb{Q}$ is a subfield of $\mathbb{R}$, $\mathbb{Q}, \mathbb{R}$ subfield of $\mathbb{C}$, i.e., they have the same $+, \cdot, 0, 1$.
$\mathbb{Z}$ is not a field as $\nexists n \in \mathbb{Z} \ni 2n = 1$, so $\mathbb{Z}$ do not satisfy (M3).

Note:To show something is FALSE, we need only one COUNTER-EXAMPLE. To show something is TRUE, one needs to show true for all elements – not just example.

§2 Lec 2: Oct 5, 2020

§2.1 Field(Cont'd)

Note: $\mathbb{Z}$ does satisfy the weaker properly if $a, b \in \mathbb{Z}$ then
(M3') if $ab = 0$ in $\mathbb{Z}$, then $a = 0$ or $b = 0$ and all other axioms except M3 hold

1. Let $F = \{0, 1\}$, $0 \neq 1$. Define $+, \cdot$ by following table Then F is a field.

Table 0.1.: ADDITION

+	0	1
0	0	1
1	1	0

Table 0.2.: MULTIPLICATION

·	0	1
0	0	0
1	0	1

2. $\exists$ fields with n elements for

$$n = 2, 3, 4, 5, 7, 8, 9, 11, 13, 16, 17, 19, \ldots$$

[conjecture?]

3. Let F be a field

$$F[t] := \{(\text{formal polynomial in one variable}\}$$

with t, given by

$$(a_0 + a_1 t + a_2 t^2 + \ldots) + (b_0 + b_1 t + b_2 t^2 + \ldots) := (a_0 + a_1) + (a_1 + b_1)t + (a_2 + b_2)t^2 + \ldots$$
$$(a_0 + a_1 t + a_2 t^2 + \ldots) \cdot (b_0 + b_1 t + b_2 t^2 + \ldots) := a_0 b_0 + (a_0 b_1 + a_1 b_0)t + \ldots$$

Note: $f, g \in F[t]$ are EQUAL iff they have the same COEFFICIENTS(coeffs) for each t^i (if t^i does not occur we assume its coeff is 0.) $F[t]$ is not a field but satisfy all axioms except (M3) but it does satisfy (M3') (compare $\mathbb{Z}$). Let

$$F(t) := \left\{ \frac{f}{g} | f, g \in F[t], g \neq 0 \right\} \quad \text{with}$$

- $\frac{f}{g} = \frac{h}{k}$ if $fk = gh$
- $\frac{f}{g} + \frac{h}{k} := \frac{fk+gh}{gk} \quad \forall f, g, h, k \in F[t]$
- $\frac{f}{g} \cdot \frac{h}{k} := \frac{fh}{gk} \quad g \neq 0, \quad k \neq 0$

is a field, the FIELD of RATIONAL POLYS over F.

Note:the 0 in $F[t]$ is $\frac{0}{f}$, $f \neq 0$, and 1 in $F[t]$ is $\frac{f}{f}$, $f \neq 0$.

4. let F be a field.

$$M_nF := \{A | A \text{an} n \times n \text{matrix entries in} F\}$$

usual $+, \cdot$ of matrices, i.e. for $A, B \in M_nF$, let

$$A_{ij} := ij^{\text{th}} \text{ entry of A, etc}$$

Then

$$(A+B)_{ij} := A_{ij} + B_{ij}$$
$$(AB)_{ij} := C_{ij} := \sum_{k=1}^{n} A_{ik}B_{kj} \quad \forall i, j$$

<u>*Note*</u>: $A = B$ iff $A_{ij} = B_{ij} \;\; \forall i, j$.

If $n = 1$, then

F and M_1F and the "same" so M_1F is a field. If $n > 1$ then M_nF is not a field nor does it satisfy (M3), (M4), (M3'). It does satisfy other axioms with

$$I = I_n := \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}, \quad 0 = 0_n := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

§2.2 Vector Space

$\mathbb{R}^2 := \{(x, y) | x, y \in \mathbb{R}\} = \mathbb{R} \times \mathbb{R}$ Vector in $\mathbb{R}^2$ are added as above and if $v \in \mathbb{R}^2$ is a vector,

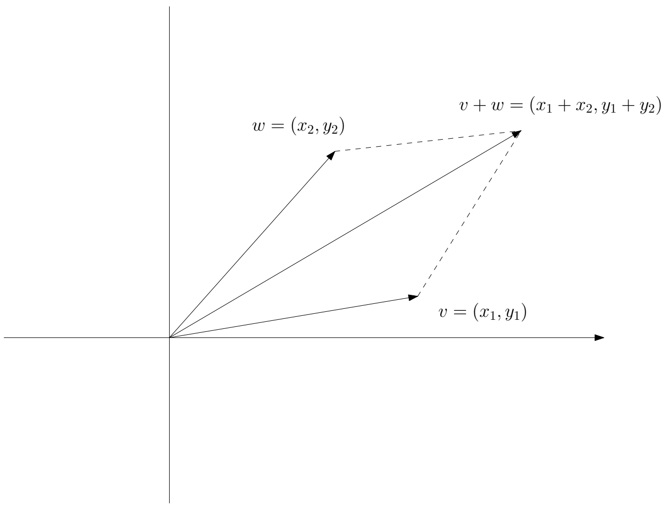


Figure 0.1.: Geometry in $\mathbb{R}^2$

αv makes sense $\forall \alpha \in F$ by $\alpha(x, y) = (\alpha x, \alpha y)$ called SCALAR MULTIPLICATION. For $+$, scalar mult and $(0, 0)$ is the ZERO VECTOR satisfying various axioms. e.g., assoc, comm, "distributive law…". To abstractify this

Definition 2.1 (Vector Space) — V is a vector space over F, via $+, \cdot$ or $(V, +, \cdot)$ is a vector space over F where

$$\begin{array}{cc} + : V \times V \to V & \cdot : F \times V \to V \\ \text{Addition} & \text{Scalar Multiplication} \\ \text{write:} v + w := +(v, w) & \text{write:} \alpha \cdot v := \cdot(\alpha, v) \ \text{ or } \ \alpha v \end{array}$$

if the following axioms are satisfied

$$\forall v, v_1, v_2, v_3 \in V, \quad \forall \alpha, \beta \in F$$

1. $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$
2. $\exists$ an element $0 \in V \ni \quad v + 0 = v = 0 + v$
3. (2) holds and the element $(-1)v$ in V satisfies
$$v + (-1)v = 0 = (-1)v + v$$
or (2) holds and $\forall v \in V \exists w \in V \ni \ v + w = 0 = w + v$
4. $v_1 + v_2 = v_2 + v_1$
5. $1 \cdot v = v$
6. $(\alpha \cdot \beta) \cdot v = \alpha(\beta \cdot v)$
7. $(\alpha + \beta)v = \alpha v + \beta v$
8. $\alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2$

Elements of V are called vector, elements of F scalars .

Comments: V: a vector space over F

1. The zero of F is unique and is a scalar. The zero of V is unique and is a vector. They are different (unless $V = F$) even if we write 0 for both – should write $0_F, 0_V$ for the zero of F, V respectively.
2. if $v, w \in V, \alpha \in F$ then
$$\begin{array}{cc} \alpha v + w & \text{makes sense} \\ v\alpha, vw & \text{do not make sense} \end{array}$$
3. We usually write

   vector using Roman letter

   scalar using Greek letter

   exception things like $(x_1, \dots, x_n) \in \mathbb{R}^n, x_i \in \mathbb{R} \forall i$
4. $+ : V \times V \to V$ says
$$\text{if } \ v, w \in V, \quad \text{then } \ v + w \in V$$
write $v, w \in V \underbrace{\to}_{\text{implies}} v + w \in V$. We say V is CLOSED under +

5. $\cdot : F \times V \to V$ says $\alpha \in F, v \in V \to \alpha v \in V$. We say V is CLOSED under SCALAR MULTIPLICATION.

Example 2.2

F a field, e.g., $\mathbb{R}$ or $\mathbb{C}$

1. F is a vector space over F with $+, \cdot$ of a field, i.e., the field operation are the vector space operation with $0_F = 0_V$.

2. $F^n := \{\alpha_1, \ldots, \alpha_n\} \, | \alpha_i \in F \forall i$ is a vector space over F under COMPONENT-WISE OPERATION and
$$0_{F^n} := (0, \ldots, 0)$$
Even have
$$F^{\infty}_{\text{finite}} = \{(\alpha_1, \ldots, \alpha_n, \ldots\} \, |\alpha_i \in F \forall i \text{ with only FINITELY MANY } \alpha_i \neq 0$$

3. Let $\alpha < \beta$ in $\mathbb{R}$
$$I = [\alpha, \beta]\,, \quad (\alpha, \beta), \quad [\alpha, \beta), \quad (\alpha, \beta]$$
including $(\alpha = -\infty, \beta = \infty)$. Let fxn $I := \{f : I \to \mathbb{R} | f \text{ a fxn}\}$ called the SET of REAL VALUE FXNS on I.

   Define $+, \cdot$ as follows: $\forall f, g \in \text{Fxn } I$,
$$\begin{aligned} f + g \quad & \text{by } (f+g)(x) := f(x) + g(x) \\ \alpha f \quad & \text{by } (\alpha f)(x) := \alpha f(x) \quad \forall \alpha \in \mathbb{R} \end{aligned}$$
and 0 by $0(\alpha) = 0 \forall \alpha \in F$. Then Fxn I is a vector space over $\mathbb{R}$.

§3 Lec 3: Oct 7, 2020

§3.1 Vector Space(Cont'd)

Example 3.1

F is a field, e.g. $\mathbb{R}$ or $\mathbb{C}$

1. F is a vector space over F with $+, \cdot$ of a field, i.e. the field operation are the vector space operation with $0_F = 0_V$.

2. $F^n := \{(\alpha_1, \ldots, \alpha_n) | \alpha_i \in F \forall i\}$ is a vector space over F under COMPONENTWISE OPERATIONS

$$(\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) := (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$$
$$\beta(\alpha_1, \ldots, \alpha_n) := (\beta\alpha_1, \ldots, \beta\alpha_n)$$

with $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in F$ and $0_{F^n} := (0, \ldots, 0)$.

Even have:

$$F^\infty = F^\infty_{\text{this}} : \{(\alpha_1, \ldots, \alpha_n, \ldots) | \alpha_i \in F \forall i \text{ with only FINITELY MANY } \alpha_i \neq 0\}$$

3. Let $\alpha < \beta$ in $\mathbb{R}$

$$I = [\alpha, \beta], \quad (\alpha, \beta), \quad [\alpha, \beta), \quad (\alpha, \beta]$$

(including $\alpha = -\infty, \beta = \infty$. Let function $I := \{f : I \to \mathbb{R} | f \text{ a function}\}$

Define $+, \cdot$ as follows: $\forall f, g \in$ Fxn I,

$$f + g \quad \text{by} \quad (f + g)(x) := f(x) + g(x)$$
$$\alpha f \quad \text{by} \quad (\alpha f)(x) := \alpha f(x) \quad \forall \alpha \in \mathbb{R}$$

and 0 by $0(\alpha) = 0 \forall \alpha \in F$. Then Fxn I is a vector space over $\mathbb{R}$.

Using this, we get subsets which are also vector space over $\mathbb{R}$ with same $+, \cdot, 0$.

- $C(I) := \{f \in \text{ fxn } I | f \text{ continuous on } I\}$
- Diff $(I) := \{f \in \text{ fxn } I | f \text{ differentiable on } I\}$
- $C^n(I) := \{f \in \text{ fxn } I | f(n) \text{ the} n^{\text{th}} \text{ derivative of } f \text{ and f exists on I and is cont on I}\}$
- $C^\infty(I) := \{f \in \text{ fxn } I | f(n) \text{ exists} \forall n \geq 0 \text{ on I and is cont}\}$
- $C^\omega(I) := \{f \in \text{ fxn } I | \text{ f converges to its Taylor Series}\}$
  (in a neighborhood of every $x \in I$ – be careful at boundary points)
- Int $(I) := \{f \in \text{ fxn } I | f \text{ is integrable on } I\}$

4. $F[t]$ the set of polys, coeffs in F old $+, \cdot$ with scalar mult

$$\alpha(\alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n) := \alpha\alpha_0 + \alpha\alpha_1 t + \ldots + \alpha\alpha_n t^n$$

5. $\underbrace{F[t]_n}_{\text{truncating} F[t]} := \{0 \in F[t]\} \cup \{f \in F[t] | \text{deg} f \leq n\}$ (not closed under $\cdot$ of polys)
where deg f = the highest power of t occurring non-trivially in f if $f \neq 0$ is a vector space over F with $+$, scalar mult,0.

Example 3.2 1. $F^{m\times n} :=$ set of $m \times n$ matrices entries in F where $A \in F^{m\times n}$, $\quad A_{ij} = ij^{\text{th}}$ entry of A

$$(A+B)_{ij} := A_{ij} + B_{ij} \in F \qquad \forall A, B \in F^{m\times n}$$
$$(\alpha A)_{ij} := \alpha A_{ij} \in F \qquad \forall \alpha \in F$$
$$0 = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} \text{(m rows and n columns)}$$

COMPONENTWISE OPERATION! Then $F^{m\times n}$ is a vector space over F, e.g. $M_n F$ is a vector space over F.

Example to GENERALIZE

Let V be a vector space over F, $\emptyset \neq S$ a set. Set $W := \{f : S \to V | f \text{ a map}\}$. Define $+, \cdot$ on W by

$$f+g \quad (f+g)(s) := f(s) + g(s) \in V$$
$$\alpha f \quad (\alpha f)(s) := \alpha(f(s)) \in V$$
$$0_W \quad 0(s) = 0_V \qquad \text{ZERO FUNCTION}$$

$\forall f, g \in W; \alpha \in F; s \in S$. Then W is a vector space over F.(of componentwise operation)

2. Let $F \subset K$ be a fields under $+, \cdot$ on K. Same 0,1, i.e. F is a SUBFIELD of k e.g. $\mathbb{R} \subset \mathbb{C}$. Then K is a vector space over F by RESTRICTION of SCALARS.

   i.e., $+ = +$ on K. With scalar mult, $F \times K \to K$ by

$$\underbrace{\alpha v}_{\text{in K as a vector space over } F} = \underbrace{\alpha v}_{\text{in K as a field}} \qquad \forall \alpha \in F \quad \forall v \in V$$

   e.g. $\mathbb{R}$ is a vector space over $\mathbb{Q}$ by $\frac{m}{n} r = \frac{mr}{n}$, $\quad m, n \in \mathbb{Z}, n \neq 0, r \in \mathbb{R}$. More generally, let V be a vector space over K, $F \subset K$ subfield, then it is a vector space over F by RESTRICTION of SCALARS.

$$\cdot|_{F\times V} : F \times V \to V$$

   e.g., K^n is a vector space over F (e.g. $\mathbb{C}^n$ is a vector space over $\mathbb{R}$).

Properties of Vector Space: Let V be a vector space over F. Then $\forall \alpha, \beta \in F$, $\quad \forall v, w \in V$, we have

1. The zero vector is unique write 0 or 0_V.
2. $(-1)v$ is the unique vector $w \ni w + v = 0 = v + w$ write $-v$.
3. $0 \cdot v = 0$
4. $\alpha \cdot 0 = 0$
5. $(-\alpha)v = -(\alpha v) = \alpha(-v)$
6. if $\alpha v = 0$, then either $\alpha = 0$ or $v = 0$

7. if $\alpha v = \alpha w$, $\alpha \neq 0$, then $v = w$
8. if $\alpha v = \beta v$, $v \neq 0$, then $\alpha = \beta$
9. $-(v+w) = (-v) + (-w) = -v - w$
10. can ignore parentheses in $+$

§3.2 Subspace

Definition 3.3 (Subspace) — Let V be a vector space over F, $W \subset V$ a subset. We say W is a subspace of V if W is a vector space over F with the operation $+, \cdot$ on V, i.e., $(V, +, \cdot)$ is a vector space over F, via $+ : V \times V \to V$ and $\cdot : F \times V \to V$ then W is a vector space over F via

- $+ = +/_{W \times W} : W \to W$: restrict the domain to $W \times W$
- $\cdot = \cdot|_{F \times W} : F \times W \to W$: restrict the domain to $F \times W$

  i.e. W is closed under $+, \cdot$ from V, $\forall^{w_1}_{w_2} \in W \quad \forall \alpha \in F, \quad w_1 + w_2 \in W$ and $\alpha w_1 \in W$ and $0_W = 0_V$.

Theorem 3.4 (Subspace)
Let V be a vector space over F, $\emptyset \neq W \subset V$ a subset. Then the following are equivalent:

1. W is a subspace for V
2. W is closed under $+$ and scalar mult from V
3. $\forall w_1, w_2 \in W$, $\forall \alpha \in F$, $\alpha w_1 + w_2 \in W$

Proof. Some of the implication are essentially ??
1) $\to$ 2) : by def. W is a subspace of V under $+, \cdot$ on V (and satisfies the axioms of a vector space over F) as $0_V = 0_W$.
2) $\to$ 1) claim: $0_V \in W$ and $0_W = 0_V$: As $\emptyset \neq W \exists w \in W$
By 2)$(-1)w \in W$, hence $0_V = w + (-w) \in W$. Since $0_V + w' = w' = w' + 0_V$ in V $\forall w' \in W$, the claim follows. The other axioms hold for elements of V hence for $W \subset V$.
2) $\to$ 3) : let $\alpha \in F$, $w_1, w_2 \in W$. As 2) holds, $\alpha w_1 \in W$ hence also $\alpha w_1 + w_2 \in W$
3) $\to$ 2) Let $\alpha \in F$, $w_1, w_2 \in W$. As above and 3)

$$0_V = w_1 + (-w_1) \in W \quad \text{and } 0_V = 0_W$$

Therefore,

$$w_1 + w_2 = 1 \cdot w_1 + w_2 \in W \quad \text{and } \alpha w_1 + \alpha w_1 + 0_V \in W$$

by 3). □

<u>*Note*</u>:Usually 3) is the easiest condition to check. WARNING: must subsets of a vector space over F are NOT subspace.

Example 3.5

V a vector space over F.

1. $0 \coloneqq \{0_V\}$ and V are subspace of V
2. Let $I \subset \mathbb{R}$ be an interval (not a point) then

$$C^\omega(I) < C^\infty(I) < \ldots < C^n(I) < \ldots < C'(I)$$
$$< \text{Diff I} < C(I) < \text{Int I} < \text{Fxn I}$$

   are subspaces of the vector space containing then... where we write

$$A < B \quad \text{if} \quad A \subset B \quad \text{and } A \neq B$$

3. Let F be afield, e.g $\mathbb{R}$. Then $F = F[t]_0 < F[t]_1 < \ldots < F[t_n] < \ldots < F[t]$ are vector space over F each a subspace of the vector space over F containing it.
4. If $W_1 \subset W_2 \subset V$, W_1, W_2 subspace of V,then $W_1 \subset W_2$ is a subspaces.
5. If $W_1 \subset W_2$ is a subspace and $W_2 \subset V$ is a subspace, then $W_1 \subset V$ is a subspace.
6. Let $W \coloneqq \{(0, \alpha_1, \ldots, \alpha_n | \alpha_i \in F, \quad 2 \leq i \leq n\} \subset F^n$ is a subspace, but $\{(1, \alpha_2, \ldots, \alpha_n | \alpha_i \in F, \quad 2 \leq i \leq n\}$ is not. Why?
7. Every line or plane through the origin in $\mathbb{R}^3$ is a subspace.

§4 Lec 4: Oct 9, 2020

§4.1 Span & Subspace

Definition 4.1 (Linear Combination) — Let V be a vector space over F, $v_1, \dots, v_n \in V$ we say $v \in V$ is a LINEAR COMBINATION of $v_1, \dots, v_n$ if $\exists \alpha_1, \dots, \alpha_n \in F \ni v = \alpha v_1 + \dots + \alpha_n v_n$.

Let

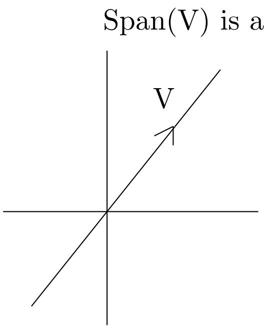
$$\text{Span}(v_1, \dots, v_n) := \{ \text{ all linear combos of} v_1, \dots, v_n \}$$

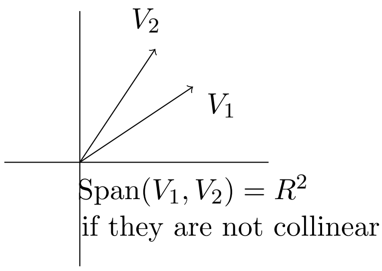
Let $v_1, \dots, v_n \in V$. Then

$$\text{Span}(v_1, \dots, v_n) = \left\{ \sum_{i=1}^{n} \alpha_i v_i | \alpha_1, \dots, \alpha_n \in F \right\}$$

is a subspace of V (by the Subspace Theorem) called the SPAN of $v_1, \dots, v_n$. It is the (unique) smallest subsapce of V containing $v_1, \dots, v_n$.

i.e., if $W \subset V$ is a subspace and $v_1, \dots, v_n \in W$ then $\text{Span}(v_1, \dots, v_n) \subset W$. We also let Span $\emptyset := \{0_V\} = 0$, the smallest vector space containing no vectors.





Question: If we view $\mathbb{C}$ as a vector space over $\mathbb{R}$, then $\mathbb{R}$ is a subspace of $\mathbb{C}$, but if we view $\mathbb{C}$ is a vector space over $\mathbb{C}$, then $\mathbb{R}$ is not a subspace of $\mathbb{C}$ (why? What's going on?) – not closed under operation(s).

Definition 4.2 (Span) — Let V be a vector space over $F, \emptyset \neq S \subset V$ a subset. Then, Span S := the set of all FINITE linear combos of vectors in S. i.e., if $V \in$ Span S, then

$$\exists v_1, \dots, v_n \in S, \quad \alpha_1, \dots, \alpha_n \in F \ni v = \alpha_1 v_1 + \dots + \alpha_n v_n$$

Span $S \subset V$ is a subspace. What is Span V?

Example 4.3 1. Let $V = \mathbb{R}^3$.

$$\text{Span}(i+j, i-j, k) = \text{Span}V = \text{Span}(i, j, i+j, k) = \text{Span}(i+j, i-j, k+i)$$

2. Define
$$\text{Symm}_n F := \left\{A \in M_n F | A = A^\top\right\}$$
Recall: $A^\top$ is the transpose of A, i.e.,
$$(A^\top)_{ij} := A_{ji} \quad \forall i, j$$
is a subspace of $M_n F$

3. $$V = \left\{\begin{pmatrix} a & c+di \\ c-di & b \end{pmatrix} | a, b, c, d \in \mathbb{R}\right\} \subset M_2 C$$
is NOT a subspace as a vector space over $\mathbb{C}$,eg,
$$i\begin{pmatrix} a & c+di \\ c-di & b \end{pmatrix} = \begin{pmatrix} ai & -d+ci \\ d+ci & bi \end{pmatrix}$$
does not lie in V if either $a \neq 0$ or $b \neq 0$ (cannot be imaginary). Also V is not a subspace of $M_2\mathbb{R}$ as a vector space over $\mathbb{R}$ as $V \not\subset M_2\mathbb{R}$. $V \subset M_2\mathbb{C}$ is a subspace as a vector space over $\mathbb{R}$.

4. (Important computational example) Fix $A \in F^{m\times n}$. Let
$$\ker A := \left\{x \in F^{n\times 1} | Ax = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \text{ in } F^{m\times 1}\right\}$$
called the KERNEL or NULL SPACE of A. Ker $A \subset F^{n\times 1}$ is a subspace and it is the SOLUTION SPACE of the system of m linear equations in n unknowns. – which we can compute by Gaussian elimination.

5. Let $W_i \subset V_i, i \in \underbrace{I}_{\text{indexing set}}$ be subspaces. Then $\bigcap_I W = \bigcap_{i\in I} W_i :=$ $\{x \in V | x \in W_i \quad \forall i \in I\}$ is a subspaces of V (why?)

6. In general, if $W_1, W_2 \subset V$ are subspaces, $W_1 \cup W_2$ is NOT a subspace.
e.g., Span(i) $\cup$ Span(j) $= \{(x, 0) | x \in \mathbb{R}\} \cup \{(0, y) | y \in \mathbb{R}\}$ is not a subspace
$$(x, y) = (x, 0) + (0, y) \notin \quad \text{Span(i)} \cup \text{Span(j)}$$
if $x \neq 0$ and $y \neq 0$

Definition 4.4 (Subspace & Span) — Let $W_1, W_2 \subset V$ be subspaces. Define

$$\begin{aligned} W_1 + W_2 &:= \{w_1 + w_2 | w_1 \in W_1, w_2 \in W_2\} \\ &= \text{Span}(W_1 \cup W_2) \end{aligned}$$

So $w_1 + w_2 \subset V$ is a subspace and the smallest subsapce of V containing W_1 and W_2.

More generally, if $W_i \in V$ is a subspace $\forall i \in I$ let

$$\sum_I W_i = \sum_{i \in I} W_i := +W_i := \text{Span}(\bigcup_I W_i)$$

the smallest subspace of V containing $W_i \forall i \in I$. What do elements in $\sum_I W_i$ look like? Determine the span of vector $v_1, \dots, v_n$ in $\mathbb{R}^n$

Suppose $v_i = (a_{i_1}, \dots, a_{ni}, \; i = 1, \dots, n.$ To determine when $w \in \mathbb{R}^n$ lies in $\text{Span}(u_1, \dots, u_n\ )$ i.e., if $w = (b_1, \dots, b_n) \in \mathbb{R}^n$ when does

$$w = \alpha_1 v_1 + \ldots + \alpha_n v_n, \qquad \alpha_1, \dots, \alpha_n \in \mathbb{R}$$

What v_i is an $n \times 1$ column matrix $\begin{pmatrix} \alpha_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}$

$$A = (a_{ij}), \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

view w as $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. To solve

$$Ax = B, \qquad X = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

is equivalent to finding all the $n \times 1$ matrices B (actually $B^\top$) s.t.

$$Ax = B$$

when the columns of A are the $v_i (v_i^\top)$.
Note: If $m = n$ an A is invertible then all B work.

§4.2 Linear Independence

We know that $\mathbb{R}^n$ is an n-dimensional vector space over $\mathbb{R}$. Since we need n coordinates (axes) to describe all vector in $\mathbb{R}^n$ but no fewer will do.
We want something like the following:

Let V be a vector space over F with $V \neq \emptyset$. Can we find distinct vectors $v_1 \dots, v_n \in V$, some n with following properties

1. $V = \text{Span}(v_1, \dots, v_n\ )$
2. No v_i is a linear combos of $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ (i.e. we need them all)

Then we want to call V an n-DIMENSIONAL VECTOR SPACE OVER F.

Lemma 4.5

Let V be a vector space over F, $n > 1$. Suppose $v_1, \dots, v_n$ are distinct. Then (2) is equivalent to

$$\text{If } \alpha_1 v_1 + \ldots + \alpha_n v_n = \beta_1 v_1 + \ldots + \beta_n v_n, \quad \alpha_i, \beta_i \in F \forall i, j$$

i.e. the "coordinates" are unique.

Proof. $(- >)$ If not, relabelling the $v_i's$, we may assume that $\alpha_1 \neq \beta_2$ in(*), then

$$(\alpha_1 - \beta_1)v_1 = \sum_{i=2}^{n} (\beta_i - \alpha_i)v_i$$

As $\alpha_1 - \beta_1 \neq 0$ in F, a field, $(\alpha_1 - \beta_1)^{-1}$ exists, so

$$v_1 = \sum_{i=2}^{n} (\alpha_1 - \beta_1)^{-1}(\beta_i - \alpha_i)v_i \in \text{ Span}(v_1, \dots, v_n)$$

a contradiction.
$(< -)$ Relabelling, we may assume that

$$v_1 = \alpha_2 v_2 + \ldots + \alpha_n v_n, \quad \text{some } \alpha_i \in F$$

Then,

$$1 \cdot v_1 + 0v_2 + \ldots + 0v_n = v_1 = 0 \cdot v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$$

so $1 = 0$, a contradiction. □

Remark 4.6. The case $n = 1$ is special because there are two possibilities
Case 1: $v \neq 0$: then $\alpha v = \beta v \to \alpha = \beta$
Case 2: $v = 0$: then $\alpha v = \beta v \forall \alpha, \beta \in F$

So the only time the above lemma is false is when $n = 1$ and $v = 0$. We do not want to say this, so we use another definition.

§5 Lec 5: Oct 12, 2020

§5.1 Linear Independence(Cont'd)

Definition 5.1 (Linear Independence & Dependence) — Let V be a vector space over F, $v_1, \ldots, v_n$ in V all distinct. We say $\{v_1, \ldots, v_n\}$ is LINEARLY DEPENDENT if $\exists \alpha_1, \ldots, \alpha_n \in F$ not all zero $\ni$

$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$$

and $\{v_1, \ldots, v_n\}$ is LINEARLY INDEPENDENT if it is NOT linearly dependent, i.e., if for any eqn

$$0 = \alpha v_1 + \ldots + \alpha_n v_n, \quad \alpha_1, \ldots, \alpha_n \in F,$$

then $\alpha_i = 0 \forall i$, i.e., the only linear comb of $v_1, \ldots, v_n$ – the zero vector is the TRIVIAL linear combo (we shall also say that distinct $v_1, \ldots, v_n$ are linearly independent if $\{v_1, \ldots, v_n\}$ is. More generally, a set $\emptyset \neq S \subset V$ is called LINEARLY DEPENDENT if for some FINITE subset (of distinct elements of S) of S is linearly dependent and it is called LINEARLY INDEPENDENT if every FINITE subset of S (of distinct elements) is linearly independent.
We say $v_i, i \in F$, all distinct are LINEARLY INDEPENDENT if $\{v_i\}_{i \in I}$ is linearly independent and $v_i \neq v_j \forall i, j \in I, i \neq j$.

Remark 5.2. Let V be a vector space over F, $\emptyset \neq S \subset V$ a subset

1. If $0 \in S$, then S is linearly dependent as $l \cdot 0 = 0$
2. distinct: $v_1, \ldots, v_n$ in V are linearly independent iff
   - no $v_i = 0$
   - $\alpha_1 v_1 + \ldots + \alpha_n v_n = \beta_1 v_1 + \ldots + \beta_n v_n, \quad \alpha_i, \beta_i \in F$ implies $\alpha_i = \beta_i \forall i$

   Note: v, v are linearly dependent if we allow repetitions – and $\{v, v\} = \{v\}$.

For homework, make sure to show this:
Suppose $v_1, \ldots, v_n$ are distinct, $n > 2$, no $v_i = 0$. Suppose no v_i is a scalar multiple of another v_j, $j \neq i$. It does not follow that $v_1, \ldots, v_n$ are linearly independent (in general).

Example 5.3 (counter-example)

$$(1,0), (0,1), (1,1) \quad \text{in} \quad V = \mathbb{R}^2$$

$(1,0), (0,1)$ are linearly indep. but not $(1,0), (0,1)$, and $(1,1)$.

Remark 5.4. Let $\emptyset \neq T \subset S$ be a subset. If T is linearly dependent, so is S. Then the contraposition is also true: if S is linearly indep., so is T.

More remarks:

1. Let $0 \neq v \in V$. Then $\{v\}$ is linearly independent and

$$Fv := \text{Span}(v)$$

is called a LINE in V:

$$\alpha v = 0 \to \alpha = 0$$

2. $u, v, w \in V \setminus \{0\}$ and $v \notin \text{Span}(w\ )$ (equivalently, $w \notin \text{Span}(v\ )$, then $\{v, w\}$ is linearly indep. and span(v, w) is called a PLANE in V.

3. $(1,1), (-2,-2)$ are linearly dep. in $\mathbb{R}^2$.

4. $(1,1), (2,-2)$ are linearly indep. in $\mathbb{R}^2$ (show coefficients are equal to each other and to 0).

5. More generally,

$$v_i = (a_{i_1}, \ldots, a_{i_n}) \quad \text{in } \mathbb{R}^n, \quad i = 1, \ldots, m \text{ (distinct)}$$

Then

$$\exists \alpha_1, \ldots, \alpha_m \in \mathbb{R} \text{ not all } 0 \;\ni\; \alpha_1 v_1 + \ldots + \alpha_m v_m = 0$$

iff $v_1, \ldots, v_m$ are linearly dep – iff $\exists \alpha_1, \ldots, \alpha_m \in \mathbb{R}$ not all 0 s.t.

$$\alpha_1(a_{11}, \ldots, a_{1m}) + \ldots + \alpha_m(a_{m1}, \ldots, a_{mn}) = 0$$

iff the matrix

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

with rows v_i row reduced to echelon form with a zero row. Also,

$$B = A^\top = \begin{pmatrix} a_{11} & \ldots & a_{m1} \\ \vdots & & \\ a_{1m} & & a_{mn} \end{pmatrix}$$

i.e., write the vectors v_i as columns then

$$\underbrace{B}_{n\times m}\underbrace{X}_{m\times 1} = 0$$

has a NON-TRIVIAL solution, i.e.,

$$\ker B \neq 0$$

where

$$\ker B := \left\{X \in F^{m\times 1} | BX = 0\right\}$$

the kernel of B.

6. Let $f_1, \ldots, f_n \in C^{n-1}(I), \quad I = (\alpha, \beta), \alpha < \beta$ in $\mathbb{R}$ and

$$\alpha_1 f_1 + \ldots + \alpha_n f_n = \underbrace{0}_{\text{the zero func}}$$

i.e., $(\alpha_1 f_1 + \ldots + \alpha_n f_n)(x) = 0 \;\; \forall x \in (\alpha, \beta)$. Taking the derivatives $(n-1)$ times and put them in matrix form, we have

$$\begin{pmatrix} f_1 & \ldots & f_n \\ f_1' & \ldots & f_n' \\ \vdots & \ldots & \vdots \\ f_1^{n-1} & \ldots & f_n^{n-1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

In particular, the Wronskian of $f_1, \ldots, f_n$ is not the zero func, i.e., $\exists x \in (\alpha, \beta) \ni W(f_1, \ldots, f_n)(x) \neq 0$. This means that the matrix above is invertible for some $x \in (\alpha, \beta)$. Then, $\alpha_1 = 0, \ldots, \alpha_n = 0$ by Cramer's rule – only the trivial soln.

Conclusion: $W(f_1, \ldots, f_n) \neq 0 \to \{f_1, \ldots, f_n\}$ is linearly indep.

WARNING: the converse is false.

Example 5.5 (of the conclusion)

Let $\alpha < \beta$ in $\mathbb{R}$.

1. $\sin x, \cos x$ are linearly indep. on (α, β).
2. We need some (sub) defns for this example.

For $x \in \mathbb{R}$, define the map

$$e_x : \mathbb{R}[t] \to \mathbb{R} \text{ by}$$

$g = \sum a_i t^i \mapsto g(x) \coloneqq \sum a_i x^i$ called EVALUATION at x.

We call a map $f : \mathbb{R} \to \mathbb{R}$ (or some $f : I \to \mathbb{R}(I \subset \mathbb{R})$) a POLYNOMIAL FUNCTION if

$$\exists P_f = \sum_{i=1}^{n} a_i t^i \in \mathbb{R}[t]$$

and

$$f(x) = e_x P_f = P_f(x) = \sum_{i=1}^{n} a_i x^i \quad \forall x \in \mathbb{R}$$

i.e., the function arising from a (formal) polynomial by evaluation at each x. We let

$$\mathbb{R}[x] \coloneqq \{f : \mathbb{R} \to \mathbb{R} | f \text{ a poly fcn } \}$$

Note:Polynomial fcns are defined on all of $\mathbb{R}$. $\mathbb{R}[x]$ is a vector space over $\mathbb{R}$.

Warning: if we replace $\mathbb{R}$ by F, $F[t]$ may be "very different" from $F[x]$, e.g., let $F = \{0, 1\}$. Then

$$t, t^2 \in F[t], \quad t \neq t^2 \text{ but } P_t = P_{t^2}$$

Now we can give our example using Wronskians

$$\{1, x, \ldots, x^n\}$$

is linearly indep. on (α, β) assuming $\alpha < \beta$.

HOMEWORK: Let $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ be distinct, then

$$e^{\alpha_1 t}, \ldots, e^{\alpha_n t}$$

are linearly indep. on (α, β). THINK OVER IT!

Theorem 5.6 (Toss In)

Let V be a vector space over F, $\emptyset \neq S \subset V$ a linearly indep. subset. Suppose that $v \in V \setminus \text{ Span } S$. Then $S \cup \{v\}$ is linearly indep.

Proof. Suppose this is false which is $S \cup \{v\}$ is linearly dep. Then $\exists v_1, \ldots, v_n \in S$ and $\alpha, \alpha_1, \ldots, \alpha_n \in F$ some n not all zero s.t.

$$\alpha v + \alpha_1 v_1 + \ldots + \alpha_n v_n = 0$$

Case 1: $\alpha = 0$
Then $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$ not all $\alpha_1, \ldots, \alpha_n$ zero so $\{v_1, \ldots, v_n\}$ is linearly dep., a contradiction.
Case 2: $\alpha \neq 0$
Then α^{-1} exists.

$$v = -\alpha^{-1}\alpha_1 v_1 - \ldots - \alpha^{-1}\alpha_n v_n$$

is a linear combo of $v_1, \ldots, v_n$, i.e., $v \in \text{Span}\ (v_1, \ldots, v_n)$ – a contradiction. Therefore, $S \cup \{v\}$ is linearly indep. $\square$

Corollary 5.7
Let V be a vector space over F and $v_1, \ldots, v_n \in V$ linearly indep. if

$$\text{Span}(v_1, \ldots, v_n) < V$$

then $\exists v_{n+1} \in V \ni v_1, \ldots, v_n, v_{n+1}$ are linearly indep. and

$$\text{Span}(v_1, \ldots, v_n) < \text{Span}(v_1, \ldots, v_{n+1}) \subset V$$

Question 5.1. Why can't we get a linearly indep. set spanning any vector space over F using this theorem?

Ans: Certainly we may not get a finite set. We shall only be interested in the case, much of the time, when such a finite linearly indep. set spans our vector space over F.

Example 5.8
$(1, 3, 1) \in \mathbb{R}^3$ is linearly indep. but $\text{Span}\ (1, 3, 1) < \mathbb{R}^3$.
$(1, 1, 0) \notin \text{Span}\ (1, 3, 1)$ so $(1, 3, 1), (1, 1, 0)$ are linearly indep. Similarly for $(0, 0, 1)$.
$\mathbb{R}^3 = \text{Span}((1, 3, 1), (1, 1, 0), (0, 0, 1))$

§6 Lec 6: Oct 14, 2020

§6.1 Bases

Definition 6.1 (Basis) — Let $\emptyset \neq V$ be a vector space over F. A BASIS B for V is a linearly indep. set in V and spans V. i.e.,

1. $V = \text{Span } B$.

2. B is linearly indep.

We say V is a FINITE DIMENSIONAL VECTOR SPACE OVER F if there exists B for V with finitely many elements, i.e., $|B| < \infty$.

Notation: If $V = 0$, we say V is a finite dimensional vector sapce over F of DIMENSION ZERO.
Goal: To show if V is finite dimensional vector space over F with bases B and b then $|B| = |b| < \infty$. This common integer is called the DIMENSION of V.

Example 6.2

Let V be a vector space over F, $S \subset V$ a linearly indep. set. Then S is a basis for Span S.
Warning: S is not a subspace just a subset.

Definition 6.3 (Ordered Basis) — If V is a finite dimensional vector space over F with a basis $B = \{v_1, \ldots, v_n\}$ we called it an ORDERED BASIS if the given order of $v_1, \ldots, v_n$ is to be used, i.e., the i^{th} vector in B is the i^{th} in the written list, e.g., $\{v_1, v_2, v_4, v_3, \ldots\}$ then v_4 is the 3^{rd} element in the ordered list if we want B to be ordered in this way.

Theorem 6.4 (Coordinate)

Let V be a finite dimensional vector space over F with basis $B = \{v_1, \ldots, v_n\}$ and $v \in V$. Then $\exists! \alpha_1, \ldots, \alpha_n \in F \ni v = \alpha_1 v_1 + \ldots + \alpha_n v_n$. We call $\alpha_1, \ldots, \alpha_n$ the COORDINATE of v relative to the basis B and call α_i the i^{th} coordinate relative to B.

Proof. Existence: By defn, $V = \text{Span } B$, so if $v \in V$

$$\exists \alpha_1, \ldots, \alpha_n \in F \ni v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

Uniqueness: Let $v \in V$ and suppose that $\alpha_1 v_1 + \ldots + \alpha_n v_n = \beta_1 v_1 + \ldots + \beta_n v_n$, for some $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in F$. Then

$$(\alpha_1 - \beta_1) v_1 + \ldots + (\alpha_n - \beta_n) v_n = 0$$

Since B is linearly indep,

$$\alpha_i = \beta_i = 0 \quad \text{for } i = 1, \ldots, n$$

□

Question 6.1. Does the above theorem hold if the basis B is not necessarily finite? If so prove it!

Exercise 6.1. Let V be a vector space over F, $v_1, \ldots, v_n \in V$ then

$$\text{Span}(v_1, \ldots, v_n) = \text{Span}(v_2, \ldots, v_n) \iff v_1 \in \text{Span}(v_2, \ldots, v_n)$$

Make sure to PROVE THIS

Note:For induction, you CAN'T assume n in the induction hypothesis is special in any way except it is greater than 1. Also, you can start induction at $n = 0$,i.e., show $P(0)$ true (or at any $n \in \mathbb{Z}$).

Theorem 6.5 (Toss Out)

Let V be a vector space over F. If V can be spanned by finitely many vector then V is a finite dimensional vector space over F. More precisely, if

$$V = \text{Span}(v_1, \ldots, v_n)$$

then a subset of $\{v_1, \ldots, v_n\}$ is a basis for V.

Proof. If $V = 0$, there is nothing to prove. So we may assume that $V \neq 0$. Suppose that $V = \text{Span}(v_1, \ldots, v_n)$. We can use induction on n and show a subset of $\{v_1, \ldots, v_n\}$ is a basis.

- $n = 1 : V = \text{Span}(v_1) \neq 0$ as $V \neq 0$, so $v_1 \neq 0$. Hence $\{v_1\}$ is linearly indep and it is the basis.
- Assume $V = \text{Span}(w_1, \ldots, w_n\ )$ – the induction hypothesis – to be true. Then a subset of $w_1, \ldots, w_n$ is a basis for V. Now suppose that $v = \text{Span}(v_1, \ldots, v_{n+1})$. To show a subset of $\{v_1, \ldots, v_{n+1}\}$ is a basis for V, we need to show if $\{v_1, \ldots, v_{n+1}\}$ is linearly indep., then it is a basis for V and it spans V and we are done. So let us assume that $\{v_1, \ldots, v_{n+1}\}$ is linearly dep. Hence,

  $$\exists \alpha_1, \ldots, \alpha_{n+1} \in F \text{ not all zero } \ni$$

  $$\alpha_1 v_1 + \ldots + \alpha_{n+1} v_{n+1} = 0$$

  Assume $\alpha_{n+1} \neq 0$, then

  $$v_{n+1} = -\alpha_{n+1}^{-1} \alpha_1 v_1 - \ldots - \alpha_{n+1}^{-1} \alpha_n v_n$$

  lies in $\text{Span}(v_1, \ldots, v_n)$. By the Exercise above,

  $$V = \text{Span}(v_1, \ldots, v_{n+1}) = \text{Span}(v_1, \ldots, v_n)$$

  By the induction hypo, a subset of $\{v_1, \ldots, v_n\}$ is a basis for V.

□

Example 6.6 1. Let $e_i = \{(0, \dots, 0, 1, 0, \dots)\} \in F^n$

$$s = s_n := \{e_1, \dots, e_n\} \subset F^n$$

If $v \in F^n$, then

$$v = (\alpha_1, \dots, \alpha_n) = \alpha_1 e_1 + \dots + \alpha_n e_n$$

since $\alpha_i \in F$, so $F^n = \text{Span } s$. If $0 = \alpha_1 e_1 + \dots + \alpha_n e_n = (\alpha_1, \dots, \alpha_n) = (0, \dots, 0)$, then $\alpha_i = 0 \forall i$. So s is linearly indep. Hence s is a basis for F^n called the standard basis. More generally, let

$$e_{ij} \in F^{m \times n} \text{ be the } m \times n \text{ matrix with all entries 0 except in the ith place.}$$

Then $s_{mn} := \{e_{ij} | 1 \le i \le m, 1 \le j \le n\}$ is a basis for $F^{m\times n}$ called the STANDARD BASIS for $F^{m\times n}$ – same proof – everything is done componentwise.

2. $V = F[t] := \{$ polys in t, coeffs in F.$\}$ ($F = \mathbb{R}$). Let $f \in V$. Then, there exists $n \ge 0$ in $\mathbb{Z}$ and $\alpha_0, \dots, \alpha_n$ in F s.t.

$$f = \alpha_0 + \alpha_1 t + \dots + \alpha_n t^n$$

So $B = \{t^n | n \ge 0\} = \left\{1, t, t^2, \dots\right\}$ spans V and by defn if

$$\alpha_0 + \alpha_1 t + \dots + \alpha_n t^n = \underbrace{0}_{\text{zero poly}}$$

then $\alpha_i = 0$ for all i so B is linearly indep. Hence B is a basis for $F[t]$. B is not a finite set. We shall see that $F[t]$ is not a finite dimensional vector space over F.

How?

3. $F[t]_n := \{f \in F[t] | f = 0 \text{or } \deg f \le n\} \subset F[t]$ is spanned by $\left\{1, t, t^2, \dots, t^n\right\}$. It is a subset of linearly indep. set. $\left\{1, t, t^2, \dots\right\} = \{t^n | n \ge 0\}$ so also linearly indep. and therefore a basis.

4. $\left\{1, \sqrt{-1}\right\}$ is a basis for $\mathbb{C}$ as a vector space over $\mathbb{R}$. $\{1\}$ is a basis for C as a vector space over $\mathbb{C}$(indeed, if F is a field, F is a vector space over F and if $0 \ne \alpha \in F$, then α^{-1} exists and $x = x\alpha^{-1}\alpha \in \text{Span } F$ so $\{\alpha\}$ is a basis. e.g., $\{\pi\}$ is a basis for $\mathbb{R}$ as a vector space over $\mathbb{R}$).

5. $\left\{e^{-x}, e^{3x}\right\}$ is a basis for

$$V := \left\{f \in \mathbb{C}^2(-\infty, \infty) | f'' - 2f' - 3f = 0\right\}$$

a vector space over $\mathbb{R}$.

6. Given $v_1, \dots, v_n \in F^n$, you know how to find $W = \text{Span}(v_1, \dots, v_n$). *Note*:If $m > n$ then rows reducing $A^\top$ must lead to a zero row so $v_1, \dots, v_m$ cannot be linearly indep. If $m = n$ we can see if

$$\det A^\top = 0 \quad (\text{or det A} = 0)$$

then linearly dep. And if

$$\det A^\top \ne 0 \quad (\text{or det A } \ne 0)$$

then linearly indep.

§7 Lec 7: Oct 16, 2020

§7.1 Replacement Theorem

Theorem 7.1 (Replacement)
Let V be a vector space over F, $\{v_1, \ldots, v_n\}$ a basis for V. Suppose that $v \in V$ satisfies

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n, \qquad \alpha_1, \ldots, \alpha_n \in F, \alpha_i \neq 0$$

Then

$$\{v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n\}$$

is also a basis for V.

Proof. Changing notation, we may assume $\alpha_1 \neq 0$. To show $\{v_1, v_2, \ldots, v_n\}$ is a basis for V, we have to show $\{v, v_2, \ldots, v_n\}$ spans V. Since

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n, \quad \alpha_1 \neq 0$$

α_1^{-1} exists, so

$$v_1 = \alpha_1^{-1} v - \alpha_1^{-1}\alpha_2 v_2 - \ldots - \alpha_1^{-1}\alpha_n v_n$$

lies in $\operatorname{Span}(v, v_2, \ldots, v_n\ )$. By Exercise ...,

$$V = \operatorname{Span}(v, v_1, \ldots, v_n) = \operatorname{Span}(v, v_2, \ldots, v_n)$$

So $\{v, v_2, \ldots, v_n\}$ spans V. Thus, $\{v, v_2, \ldots, v_n\}$ is linearly indep.
Suppose $\exists \beta_1, \beta_2, \ldots, \beta_n \in F$ not all $0 \ni$

$$\beta v + \beta_2 v_2 + \ldots + \beta_n v_n = 0$$

Case 1: $\beta = 0$
Then $\beta_2 v_2 + \ldots + \beta_n v_n = 0$ not all $\beta_i = 0$. So $\{v_2, \ldots, v_n\}$ is linearly dep., a contradiction.
Case 2: $\beta \neq 0$, so β^{-1} exists.
Then using (*), we see

$$v = 0 \cdot v_1 - \beta^{-1}\beta_2 v_2 - \ldots - \beta^{-1}\beta_n v_n = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

As $\{v_2, \ldots, v_n\}$ is a basis, by the Coordinate Theorem, we have

$$\alpha_1 = 0 \quad \text{and } \alpha_1 = \beta^{-1}\beta_i$$

a contradiction. □

Question 7.1. In the Replacement Theorem, do we need the basis to be finite?

Ans: I think it can be infinite ...

§7.2 Main Theorem

Theorem 7.2 (Main)

Suppose V is a vector space over F with $V = \text{Span}(v_1, \dots, v_n)$. Then any linearly indep. subset of V has at most n elements.

Proof. We know that a subset of $B = \{v_1, \dots, v_n\}$ is a basis for V by Toss Out Theorem. So we may assume B is a basis for V. It suffices to show any linearly indep. set in V has at most $|B| = n$ elements where B is a basis. Let $\{w_1, \dots, w_m\} \subset V$ be linearly indep. where no $w_i = 0$. To show $m \leq n$, the idea is to use Toss In and Toss out in conjunction with the Replacement Theorem.

Claim 7.1. After changing notation, if necessary, for each $k \leq n$

$$\{w_1, \dots, w_k, v_{k+1}, \dots, v_n\}$$

is a basis for V.

Suppose we have shown the above claim for $k = n$. Apply the claim to $k = n$ if $m > k$, then $\{w_1, \dots, w_{n+1}\}$ is linearly dep., a contradiction as $\{w_1, \dots, w_n\}$ is a basis. Thus, we prove the claim for $m \leq n$ as needed. We prove it by induction on k. BY the argument above, we may assume $k \leq n$.

- $k = 1$: As $w_1 \in$ Span $B =$ Span $(v_1, \dots, v_n)$ and $w_1 \neq 0$, $\exists \alpha_1, \dots, \alpha_n \in F$ not all 0 $\ni$
$$w_1 = \alpha_1 v_1 + \ldots + \alpha_n v_n$$
Changing notation, we may assume $\alpha_1 \neq 0$. By the Replacement Theorem,
$$\{w_1, v_2, \dots, v_n\} \text{ is a basis for } V$$
- Assume the claim hold for $k (k < n)$.
- We must show the claim holds for $k + 1$,
$$\{w_1, \dots, w_k, v_{k+1}, \dots, v_n\} \text{ is a basis for } V$$
We can write
$$0 \neq w_{k+1} = \beta_1 w_1 + \ldots + \beta_k w_k + \alpha_{k+1} v_{k+1} + \ldots + \alpha_n v_n$$
for some (new) $\beta_1, \dots, \beta_k, \alpha_{k+1}, \dots, \alpha_n \in F$ not all 0
Case 1: $\alpha_{k+1} = \alpha_{k+2} = \ldots = \alpha_n = 0$
Then $w_{k+1} \in \text{Span}(w_1, \dots, w_k)$, hence $\{w_1, \dots, w_{k+1}\}$ is linearly dep., a contradiction.
Case 2: $\exists i \ni \alpha_i \neq 0$:
Changing notation, we may assume $\alpha_{k+1} \neq 0$. By the Replacement Theorem
$$\{w_1, \dots, w_{k+1}, v_{k+2}, \dots, v_n\}$$
is a basis for V. This completes the induction step thus prove the claim and establish the theorem. □

§7.3 A Glance at Dimension

Corollary 7.3

Let V be a finite dimensional vector space over F, B_1, B_2 two bases for V. Then $|B_1| = |B_2| < \infty$. We call $|B_1|$ the dimension of V, write $\dim V = \dim_F V = |B_1|$ (dropping F if F is clear).

Proof. By defn of finite dimensional vector space over F, $\exists$ a basis b for V with $|b| < \infty$. By the Main Theorem, $|B| \leq |b|$, if B is a basis for V, so B is finite. Again by the Main Theorem, $|b| \leq |B|$ if B is a basis for V, so $|b| = |B|$ for any basis B of V. $\square$

The corollary above says $\dim V$ is well-defined for all finite dimensional vector space over F, i.e., "dim" : {finite dimensional vector space over $F \to \mathbb{Z}^+ \cup \{0\}$} is a function.
Warning: F makes a difference.

Example 7.4

$$\begin{aligned} &\dim_{\mathbb{C}} \mathbb{C} = 1 && \text{basis } \{1\} \\ &\dim_{\mathbb{R}} \mathbb{C} = 2 && \text{basis } \{1, \sqrt{-1}\} \\ &\dim_{\mathbb{Q}} \mathbb{C} = ? \end{aligned}$$

Corollary 7.5

$\dim_F F^n = n$.

Corollary 7.6

$\dim_F F^{m \times n} = mn$.

Corollary 7.7

$\dim_F F[t]_n = 1 + n$.

Note: If V is a finite dimensional vector space over F with bases B, then the Replacement Theorem allows us to find many other bases.

Corollary 7.8

Let V be a finite dimensional vector space over F, $n = \dim V$, $\emptyset \neq S \subset V$ a subset. Then

- If $|S| > n$, then S is linearly dep.
- If $|S| < n$, then Span $S < V$.

Proof.

- First bullet point: The Main Theorem says:

  A maximal linearly indep. set in V is a basis and can have at most n elements by Toss In Theorem.

- Second bullet point: By Toss Out Theorem, we can assume that S is linearly indep., so it cannot be a basis by Corollary ?.

□

Question 7.2. What is $\dim_{\mathbb{R}} M_n(\mathbb{C})$?

§8 Lec 8: Oct 19, 2020

§8.1 Extension and Counting Theorem

Theorem 8.1 (Extension)
Let V be a finite dimensional vector space over F, $W \subset V$ a subspace. Then every linearly independent subset S in W is finite and part of a basis for W which is a finite dimensional vector space over F.

Proof. Any linearly indep. set in W is linearly indep. subset S in V so $|S| \leq \dim V < \infty$ by the Main Theorem. In particular,

$$\dim \text{Span} S \leq \dim V$$

if $W = \text{Span } S$, we are done.
If not, $\exists w_1 \in W \setminus \text{Span } S$, and hence $S_1 = S \cup \{w_1\}$ is linearly indep. by Toss In Theorem and

$$|S_1| = |S \cup \{w_1\}| = |S| + 1 \leq \dim V$$

if Span $S_1 < W$, then $\exists w_2 \in W \setminus \text{Span } S_1$, so $S_2 = S \cup \{w_1, w_2\} \subset W$ is linearly indep., hence

$$|S_2| = |S| + 2 \leq \dim V$$

Continuing in this manner, we must stop when $n \leq \dim V - \dim \text{Span } S$ as $\dim V < \infty$. So S is a part of a basis for W and W is a finite dimensional vector space over F. □

Think about the proof for this

Corollary 8.2
Let V be a finite dimensional vector space over F. Then any linearly indep. set in V can be EXTENDED to a basis for V, i.e., is part of a basis for V. We often call this special case the Extension Theorem.

Corollary 8.3
Let V be a finite dimensional vector space over F, $W \subset V$ a subspace. Then W is a finite dimensional vector space over F and $\dim W \leq \dim V$ with equality iff $W = V$.

Proof. Left as exercise. □

Theorem 8.4 (Counting)
Let V be a finite dimensional vector space over F, $W_1, W_2 \subset V$ subspaces. Suppose that both W_1 and W_2 are finite dimensional vector space over F. Then

1. $W_1 \cap W_2$ is a finite dimensional vector space over F.
2. $W_1 + W_2$ is a finite dimensional vector space over F.
3. $\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2)$.

Proof. 1. $W_1 \cap W_2 \subset W_i, i = 1, 2$, so it is a finite dimensional vector space over F by corollary 8.2.

2. Let B_i be a basis for W_i, $i = 1, 2, \ldots$. Then $W_1 + W_2 = \text{Span } (B_1 \cup B_2)$ and

$$|B_1 \cup B_2| \leq |B_1| + |B_2| < \infty$$

So $W_1 + W_2$ is a finite dimensional vector space over F by Toss Out.

3. Let $B = \{v_1, \ldots, v_n\}$ be a basis for $W_1 \cap W_2$. Extend B to a basis

$$b_1 = \{v_1, \ldots, v_n, y_1, \ldots, y_r\} \text{ for } W_1$$
$$b_2 = \{v_1, \ldots, v_n, z_1, \ldots, z_s\} \text{ for } W_2$$

using the Extension Theorem.

Claim 8.1. $b_1 \cup b_2 = \{v_1, \ldots, v_n, y_1, \ldots, y_r, z_1, \ldots, z_s\}$ is a basis for $W_1 + W_2$ and has $n + r + s$ elements. So if we show the claim, the result will follow.

Certainly,

$$\text{Span}(b_1 \cup b_2) = \text{Span } b_1 + \text{ Span } b_2 = W_1 + W_2$$

So we need only to show $b_1 \cup b_2$ is linearly indep. Suppose this is false. Then

$$0 = \alpha_1 v_1 + \ldots + \alpha_n v_n + \beta_1 y_1 + \ldots + \beta_r y_r + \gamma_1 z_1 + \ldots + \gamma_s z_s \qquad (*)$$

for some $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n, \gamma_1, \ldots, \gamma_s$ in F not all zero.

Case 1: All the $\gamma_i = 0$. Since b_1 is linearly indep., this is a contradiction.

Case 2: Some $\gamma_i \neq 0$.

Changing notation, we may assume $\gamma_1 \neq 0$. Since b_2 is a basis, (*) leads to an equation

$$0 \neq z = \gamma_1 z_1 + \ldots + \gamma_s z_s = -\alpha_1 v_1 - \ldots - \alpha_n v_n - \beta_1 y_1 - \ldots - \beta_r y_r$$

Therefore, $0 \neq z$ lies in Span $b_2 \cap$Span $b_1 = W_2 \cap W_1$. So we can write $zi \in W_1 \cap W_2$ using basis B as

$$0 \neq z = \delta_1 v_1 + \ldots + \delta_n v_n \quad \text{some } \delta_1, \ldots, \delta_n \in F$$

Thus $W_2 = \text{Span } b_2$, we have

$$\delta_1 v_1 + \ldots + \delta_n v_n - 0 z_1 + \ldots + 0 z_s = z = 0 v_1 + \ldots + 0 v_n + \gamma_1 z_1 + \ldots + \gamma_s z_s$$

By the Coordinate Theorem, $\gamma_1 = 0$, a contradiction.

□

Corollary 8.5

Let V be a vector space over F, $W_1, W_2 \subset V$ finite dimensional subspaces of V. Then

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

iff

$$W_1 \cap W_2 = \emptyset$$

In this case, we write $W_1 + W_2 = W_1 \oplus W_2$ called the DIRECT SUM.

§8.2 Linear Transformation

In mathematics, whenever you have a collection of objects, one studies maps between them that preserves any special properties of the objects in the collection and tries to see what information can be gained from such maps.

Definition 8.6 (Linear Transformation) — Let V, W be a vector space over F. A map $T : V \to W$ is called a Linear Transformation, write $T : V \to W$ is linear if $\forall v_1, v_2 \in V, \forall \alpha \in F$

- $T(v_1 + v_2) = T(v_1) + T(v_2)$.
- $T(\alpha v_1) = \alpha T(v_1)$.
- $T(0_V) = 0_W$.

Notation: We write Tv for $T(v)$.

Remark 8.7. Let V, W be a vector space over F, $T : V \to W$ a map.

1. If T satisfies 1) and 2), then it satisfies 3):
$$0_W + T(0_V) = T(0_V) = T(0_V + 0_V) = T(0_V) + T(0_V)$$
so $0_W = T(0_V)$.
2. T is linear iff $T(\alpha v_1 + v_2) = \alpha T v_1 + T v_2 \quad \forall v_1, v_2 \in V, \forall \alpha \in F$.
3. If T is linear, $\alpha_1, \dots, \alpha_n \in F, v_1, \dots, v_n \in V$, then
$$T\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i T v_i$$
We leave a proof of 2) and 3) as exercises.

Example 8.8

Let V, W be a vector space over F. The followings are linear transformations

1. $0_{V,W} : V \to W$ by $v \mapsto 0_W$.
2. $V = W$, $1_V : V \to V$ by $v \mapsto v$.
   A linear transformation $T : V \to V$ is called a Linear Operator.
3. If $\emptyset \neq Z \subset W$ is a subset, then we have a map

$$\text{inc} : Z \to W$$

   given by $z \mapsto z$ called the Inclusion Map. Then, Z is a subspace of V iff inc: $Z \hookrightarrow W$ is linear.
   Note: inc $= \underbrace{1_W\Big|_Z}_{\text{Restriction map}}$.

   This is the Subspace Theorem.

4. $T : F^n \to F^{n-1}$ by $(\alpha_1, \dots, \alpha_n) \mapsto (\alpha_1, \dots, \overbrace{i}^{\text{omit}}, \dots, \alpha_n$ for a fixed i.
5. $T : F^n \to F$ by $(\alpha_1, \dots, \alpha_n) \mapsto \alpha_i$ for a fixed i.
6. $T : \mathbb{R}^{n-1} \to \mathbb{R}^n$ by $(\alpha_1, \dots, \alpha_{n-1} \mapsto (\alpha_1, \dots, \alpha_{i-1}, 0, \alpha_i, \dots, \alpha_n)$ for fixed i.
7. $T : \mathbb{R} \to \mathbb{R}^n$ by $\alpha \mapsto (0, 0, \dots, \alpha, 0, \dots, 0)$ for fixed i.
8. If $\alpha < \beta$ in $\mathbb{R}$, $D : C'(\alpha, \beta) \to C(\alpha, \beta)$ by $f \mapsto f'$.
9. If $\alpha < \beta$ in $\mathbb{R}$, Int: $C(\alpha, \beta) \to C'(\alpha, \beta)$ by $f \mapsto \int f$ where $\int f$ is the antiderivative – constant of integration 0.
10. Fix $\alpha \in F$, then $\lambda\alpha : V \to V$ by $v \mapsto \alpha v$. Left translation by α.
11. Let $A \in F^{m \times n}$. Define

$$T : F^{n \times 1} \to F^{m \times 1} \quad \text{by } T \cdot X = A \cdot X$$
$$\text{i.e. } \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \mapsto A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

   Matrices can be viewed as linear transformation. We should see the converse is true IF V is a finite dimensional vector space over F. It is not true in general.

§9 Lec 9: Oct 21, 2020

§9.1 Kernel, Image, and Dimension Theorem

Definition 9.1 (Kernel(Nullspace)) — Let V, W be a vector space over F, $T : V \to W$ linear set

$$N(T) = \ker T := \{v \in V | Tv = 0_W\}$$

called the nullspace or kernel of T.

Definition 9.2 (Range(Image)) — Let V, W be a vector space over F, $T : V \to W$ linear set

$$\begin{aligned} \text{im } T = T(V) &:= \{w \in W | \exists v \in V \ni Tv = w\} \\ &= \{Tv | v \in V\} \end{aligned}$$

called the range or image of T.

Proposition 9.3

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

1. $\ker T \subset V$ is a subspace.
2. $imT \subset W$ is a subspace.

Proof. Left as exercise. □

Theorem 9.4 (Dimension)

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

1. $im\ T$ and $\ker T$ are finite dimensional vector space over F.
2. $\dim V = \dim \ker T + \dim imT$.

Note: $\dim \ker T$ is also called the NULLITY of T and $\dim imT$ is also called the RANK of T.

Proof. Let $n = \dim V$.

$\ker T \subset V$ is a subspace, V is a finite dimensional vector space over F so $\ker T$ is a finite dimensional vector space over F and $\dim \ker T \leq \dim V = n$. Say $m = \dim \ker T$. Let $\mathscr{B}_0 = \{v_1, \dots, v_m\}$ be a basis for $\ker T$. By the Extension Theorem $\exists \mathscr{B} = \{v_1, \dots, v_m, \dots, v_n\}$ a basis for V.

Claim 9.1. $Tv_{m+1}, \dots, Tv_n$ are linearly indep. (in particular, distinct) and

$$\mathscr{C} = \{Tv_{m+1}, \dots, Tv_n\}$$

is a basis for imT.

If we prove the claim above, then imT is a finite dimensional vector space over F of dimension $n-m$ and we are done.

Step 1: $\mathscr{C}$ spans imT:

Let $w \in imT$. By definition, $\exists v \in V \ni Tv = w$. As $\mathscr{B}$ is a basis for V $\exists \alpha_1, \dots, \alpha_n \in F \ni$

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n$$

Hence

$$\begin{aligned} w = T(v) = T(\alpha_1 v_1 + \dots + \alpha_n v_n) &= \alpha_1 T v_1 + \dots + \alpha_n T v_n \\ &= \alpha_1 0_W + \dots + \alpha_m 0_W + \alpha_{m+1} T v_{m+1} + \dots + \alpha_n T v_n \end{aligned}$$

lies w Span($\mathscr{C}$) (as $v_1, \dots, v_m \in \ker T$). need recheck

Case 2: $\mathscr{C}$ is linearly indep.

Suppose $\alpha_{m+1}, \dots, \alpha_n \in F$ and

$$\alpha_{m+1} T v_{m+1} + \dots + \alpha_n T v_n = 0_W$$

Then

$$0_W = T(\alpha_{m+1} v_{m+1} + \dots + \alpha_n v_n)$$

So $\alpha_{m+1} v_{m+1} + \dots + \alpha_n v_n \in \ker T$. By defn, $\mathscr{B}_0$ is a basis for $\ker T$. So $\exists \beta_1, \dots, \beta_m \in F \ni$

$$\alpha_{m+1} v_{m+1} + \dots + \alpha_n v_n = \beta_1 v_1 + \dots + \beta_m v_n$$

Hence

$$0 = -\beta_1 v_1 - \dots - \beta_m v_m + \alpha_{m+1} v_{m+1} + \dots + \alpha_n v_n$$

As $\mathscr{B}$ is a basis for V, it is linearly indep, so $\beta_1 = 0, \dots, \beta_m = 0, \alpha_{m+1} = 0, \dots, \alpha_n = 0$ (Coordinate Theorem) and the claim follows. □

Note: Let V be a finite dimensional vector space over F, $W \subset V$ a subspace, V/W the quotient space, then $- : V \to V/W$, $v \mapsto \overline{v} = v + W$ and $\dim V/W = \dim V - \dim W$.

§9.2 Algebra of Linear Transformation

We want to study the set of all linear transformation from a vector space over F V to a vector space over F W. Let V, W be a vector space over F. Set

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

Check: if $T, S \in L(V, W), \alpha \in F$, then $\alpha T + S \in L(V, W)$. Since we know $\mathscr{F}(V, W) = \{f : V \to W | f \text{ a map}\}$ is a vector space over F, by the Subspace Theorem, $L(V, W) \subset \mathscr{F}(V, W)$ is a subspace.

Proposition 9.5

Let V, W be a vector space over F, then $L(V, W) \subset \mathscr{F}(V, W)$ is a subspace.

Now we know if we have maps

$$f : X \to Y \quad \text{and} \quad g : y \to Z,$$

we have the COMPOSITE MAP

$$g \circ f : X \to Z \quad \text{by } (g \circ f)(x) = g\left(f(x)\right) \forall x \in X$$

where $\circ$ is called the COMPOSITION (and often omitted when clear). Then we have

Proposition 9.6

Let V, W, X, U be vector space over F, $T, T' : V \to W$, $S, S' : W \to X$, $R : X \to U$ all be linear. Then,

1. $S \circ T : V \to W$ is linear.(the composition of linear transformations is linear).
2. $R \circ (S \circ T) = (R \circ S) \circ T$ and linear.
3. $S \circ (T + T') = S \circ T + S \circ T'$ and linear.
4. $(S + S') \circ T = S \circ T + S' \circ T$ and linear.

Proof.

$$\begin{aligned}(S \circ T)(\alpha v_1 + v_2) &= S\left(T(\alpha v_1 + v_2)\right) = S(\alpha T v_1 + T v_2) \\ &= \alpha S \circ T(v_1) + S \circ T(v_2)\end{aligned}$$

$\forall v_1, v_2 \in V, \alpha \in F$.
The rest are left as exercises. □

Definition 9.7 (Linear Operator) — Let V be a vector space over F, $T : V \to V$ linear, so a linear operator is defined as

$$\begin{aligned}&T^n := \underbrace{T \circ \ldots \circ T}_{n} \quad \text{if } n \in \mathbb{Z}^+ \\ &T^0 = 1_V\end{aligned}$$

Proposition 9.8

Let V be a vector space over F. Then $L(V, V)$ under $+$ and $\circ$ of functions $V \to V$ satisfies all the axioms of a field except possibly (M3) and (M4) with

$$\begin{aligned}\text{one} = 1_V : V \to V \quad \text{by } v \mapsto v \\ \text{zero} = 0_V : v \to v \quad \text{by } v \mapsto 0\end{aligned}$$

We say $L(V, V)$ is a (non-commutative) ring of $M_n F$.

§9.3 Linear Transformation Theorems

Definition 9.9 (Properties/Consequences of Linear Transformation) — Let $T : V \to W$ be linear. We say that T is

1. a MONOMORPHISM (write mono or monic) or NONSINGULAR if T is $1 - 1$. (i.e., injective).
2. an EPIMORPHISM (write epi or epic) if T is onto (i.e., surjective).
3. an ISOMORPHISM (write iso) or INVERTIBLE if T is bijective and $T^{-1} : W \to V$ is linear. We say V, W vector spaces over F are ISOMORPHIC (write $V \cong W$ if $\exists$ an isomorphism $S : V \to W$, we also write an isomorphism $S : V \to W$ as $S : V \xrightarrow{\sim} W$

Remark 9.10. $V \cong W$ vector space over F means that we cannot take V and W apart algebraically.

Example 9.11

$F^{n+1} \cong F[t]_n$ as $F^{n+1} \to F[t]_n$ by $(\alpha_0, \ldots, \alpha_n) \mapsto \alpha_0 + \alpha_1 t_1 + \ldots + \alpha_n t^n$ is an isomorphism with inverse $F[t]_n \to F^{n+1}$ by $\alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n \mapsto (\alpha_0, \ldots, \alpha_n)$

$$\begin{aligned} T^{-1}(\alpha w_1 + w_2) &= T^{-1}(\alpha T v_1 + T v_2) = T^{-1}\left(T(\alpha v_1 + v_2)\right) \\ &= T^{-1}T(\alpha v_1 + v_2) \\ &= \alpha v_1 + v_2 \\ &= \alpha T^{-1} w_1 + T^{-1} w_2 \quad \square \end{aligned}$$

Corollary 9.12

Let $T : V \to W$ be a monomorphism. Then $V \cong imT$ via T.

Remark 9.13. If V, W, X are vector space over F, then

1. $V \cong V$
2. $V \cong W \to W \cong V$
3. $V \cong W$ and $W \cong X$ then $V \cong X$

In algebra, isomorphisms are usually easier to check than are one might assume, because the following result is often true.

Proposition 9.14

Let $T : V \to W$ be linear. Then T is an isomorphism iff T is bijective.

Proof. $(\to)$ immediate.
$(\leftarrow)$ Let $T^{-1} : W \to V$ be the set inverse of $T : V \to W$, so

$$T \circ T^{-1} = 1_W \quad \text{and } T^{-1} \circ T = 1_V$$

In particular, if $v \in V$ and $w \in W$,

$$w = Tv \quad \text{iff} \quad T^{-1}w = v$$

Let $w_1, w_2 \in W$, $\alpha \in F$. To show

$$T^{-1}(\alpha w_1 + w_2) = \alpha T^{-1} w_1 + T^{-1} w_2$$

T is onto so

$$\exists v_i \in V \ni T v_i = w_i, i = 1, \ldots$$

Hence, we have

$$\begin{aligned} T^{-1}(\alpha w_1 + w_2) &= T^{-1}(\alpha T v_1 + T v_2) = T^{-1}(T(\alpha v_1 + v_2)) \\ &= T^{-1}T(\alpha v_1 + v_2) = \alpha v_1 + v_2 \\ &= \alpha T^{-1} w_1 + T^{-1} w_2 \end{aligned}$$

$\square$

§10 Lec 10: Oct 23, 2020

§10.1 Monomorphism, Epimorphism, and Isomorphism

Corollary 10.1
Let $T : V \to W$ be a monomorphism. Then $V \cong \text{ im } T$ via T.

Definition 10.2 (Linear Map) — Let $T : V \to W$ be linear. We say T takes linearly independent sets to linearly independent sets if $v_i, i \in I$ are linearly independent in V (in particular, distinct). Then, $Tv_i, i \in I$ are linearly indep. in W. ($Tv_i \neq Tv_j$ if $i \neq j$ in I)

Theorem 10.3 (Monomorphism)
Let $T : V \to W$ be linear. Then the followings are true

1. T is $1 - 1$, so it's monomorphism.
2. T takes linearly indep. sets in V to linearly indep. sets in W.
3. $\ker T = 0 := \{0_V\}$.
4. $\dim \ker T = 0$.

Proof.

- 3) iff 4) is the defn of the 0-space.
- 1) $\to$ 2) It suffices to show that T takes finite linearly indep. sets in V to linearly indep. sets in W.
  Suppose that $v_1, \ldots, v_n \in V$ are linearly indep. and $\alpha_1, \ldots, \alpha_n \in F$ satisfy
  $$0_W = \alpha_1 Tv_1 + \ldots + \alpha_n Tv_n$$
  Then
  $$T(0_V) = 0_W = T(\alpha_1 v_1 + \ldots + \alpha_n v_n)$$
  As T is $1 - 1$
  $$0_V = \alpha_1 v_1 + \ldots + \alpha_n v_n$$
  Since $v_1, \ldots, v_n$ are linearly indep. $\alpha_i = 0, i = 1, \ldots, n$ as needed.
- 2) $\to$ 3) Let $v \in \ker T$. Then $Tv = 0_W$. If $v \neq 0$, then $\{v\}$ is linearly indep. By 2) $Tv \neq 0_W$ as then $\{Tv\}$ is linearly indep. So $v \neq 0$.
- 3) $\to$ 1) If $Tv_1 = Tv_2$, $v_1, v_2 \in V$, then
  $$0_W = Tv_1 - Tv_2 = T(v_1 - v_2)$$
  So $v_1 - v_2 = 0_V$ by 3), i.e., $v_1 = v_2$ □

Remark 10.4. The Monomorphism Theorem says $\ker T$ measures the deviation of T from being $1-1$.

Note: In the Monomorphism Theorem, we do not assume that V or W is a finite dimensional vector space over F.

Theorem 10.5 (Isomorphism)

Suppose $T : V \to W$ is linear with $\dim V = \dim W < \infty$,i.e., V, W are finite dimensional vector space over F of the same dimension. Then the followings are true

1. T is an isomorphism.
2. T is a monomorphism.
3. T is an epimorphism.
4. If $\mathscr{B} = \{v_1, \dots, v_n\}$ is a basis for V, then $\{Tv_1, \dots, Tv_n\}$ is a basis for W (so $Tv_1, \dots, Tv_n$ are distinct), i.e., T takes basis of V to basis of W.
5. There exists a basis $\mathscr{B}$ of V that maps to a basis of W.

Remark 10.6.

1. The condition that $\dim V = \dim W < \infty$ is crucial

   Come up with a counter example

2. Let $V \cong W$ with V, W be finite dimensional vector space over F. So $\dim V = \dim W$. Let $S : V \to W$ be linear. Then S may or may not be an isomorphism, e.g., if S is the zero map then it is not an isomorphism unless $V = 0$. The theorem only says that $\exists$ an isomorphism and any such satisfies the theorem.
3. Let $f : A \to B$ be a map of finite sets with $|A| = |B|$. Then f is a bijection iff f is an injection iff f is a surjection.

Proof. (of Theorem)

- 1) $\to$ 2) follows by defn.
- 2) $\to$ 3) By the Dimension Theorem

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

  Thus, T is onto iff im $T = W$ iff $\dim W = \dim$ im T (by the Corollary to the Existence Theorem) iff $\dim \ker T = 0$ iff T is $1-1$.
- 3) $\to$ 1) as 3) $\to$ 2) and 1) = 2) + 3) by the Proposition ?
- 2) $\to$ 4) Let $\{v_1, \dots, v_n\}$ be a basis for V. By the Monomorphism Theorem, $Tv_1, \dots, Tv_n$ are linearly indep. in W, so

$$n \le \dim W = \dim V = n$$

  Hence $\{Tv_1, \dots, Tv_n\}$ also spans as $\dim W = \dim V$.
- 4) $\to$ 5) $\to$ 3) are clear.

$\square$

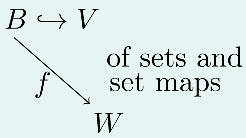
§10.2 Existence of Linear Transformation

The next result is really the defining property of finite dimensional vector space and linear transformation.

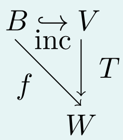
Theorem 10.7 (Existence of Linear Transformation (UPVS))
– (Universal Property of Vector Space) Let V be a finite dimensional vector space over F, $\mathscr{B} = \{v_1, \ldots, v_n\}$ a basis for V and W an arbitrary vector space over F. Let $w_1, \ldots, w_n \in W$, not necessarily distinct. Then

$$\exists! \;\; T : V \to W \text{ linear } \ni Tv_i = w_i \forall i$$

We can write this in an other way as follows:
Let $B \hookrightarrow V$ be a basis for V, V a finite dimensional vector space over F and W a vector space over F. Given a diagram,



then $\exists! T : V \to W$ linear $\ni$



commutes , i.e., $T \circ \text{inc} = f$.

Proof. Define $T : V \to W$ as follows: let $V \in V$. The $\exists! \alpha_1, \ldots, \alpha_n \in F \ni v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ by the Coordinate Theorem. Define

$$Tv = T(\alpha_1 v_1 + \ldots + \alpha_n v_n) \coloneqq \alpha_1 w_1 + \ldots + \alpha_n w_n$$

Since the α_i ARE UNIQUE, this defines a map – we say $T : V \to W$ is WELL – DEFINED. Certainly, $Tv_i = w_i, i = 1, \ldots, n$. To show T is linear, let $v = \sum_{i=1}^n \alpha_i v_i, v' = \sum_{i=1}^n \beta_i v_i$, $\alpha, \alpha_i, \beta_j \in F \forall i, j$. Then

$$\begin{aligned} T(\alpha v + v') &= T\left(\alpha \sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^n \beta_i v_i\right) \\ &= T\left(\sum_{i=1}^n (\alpha\alpha_i + \beta_i) v_i\right) = \sum_{i=1}^n (\alpha\alpha_i + \beta_i) w_i \\ &= \alpha \sum_{i=1}^n \alpha_i w_i + \sum_{i=1}^n \beta_i w_i = \alpha Tv + Tv' \end{aligned}$$

as needed. This shows existence.
Uniqueness: Let $T : V \to W$ by (*) and $S : V \to W$ linear s.t. $Sv_i = w_i \forall i$. To show $S = T$, let $v = \sum_{i=1}^n \alpha_i v_i, \alpha_i \in F$ unique, $i = 1, \ldots, n$. Then $Tv = \sum_{i=1}^n \alpha_i Tv_i = \sum_{i=1}^n \alpha_i w_i$ which is equivalent to

$$= \sum_{i=1}^n \alpha_i S v_i = S\left(\sum_{i=1}^n \alpha_i v_i\right) = Sv$$

So S is T and we have proven uniqueness. □

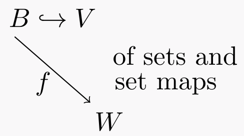
Remark 10.8. The theorem says a linear transformation from a finite dimensional vector space over F is completely determined by what it does to a fixed basis. i.e., as there are no non – trivial RELATIONS on linear combos of elements in $\mathscr{B}$, the only relation in im T will arise from the kernel of T.

§11 Lec 11: Oct 26, 2020

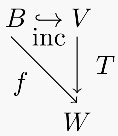
§11.1 Lec 10 (Cont'd)

Remark 11.1.
1. In the above, given $fv_i = w_i \forall i$, we say that $T : V \to W$ by $\sum \alpha_i v_i \mapsto \alpha_i w_i$ EXTENDS f linearly.
2. Let V be any vector space over F (not necessarily finite dimensional). Suppose V has a basis $\mathscr{B}$, then every $v \in V$ is a finite linear combo elements in $\mathscr{B}$. Using the same proof of UPVS, shows

   if W is a vector space over F, then given a diagram

   of set and set maps. $\exists! T : V \to W$ linear s.t.

   commutes. I.E., if $\mathscr{B} = \{v_i\}_I$ is a basis for V, $w_i \in W$, $i \in I$ (not necessarily distinct), $f : V \to W$ by $v_i \mapsto w_i \forall i \in I$. Then $\exists! T : V \to W$ linear s.t. $Tv_i = w_i \forall i \in I$. So any linear transformation from a vector space over F V having a basis is completely determined by what it does to that basis.
3. <u>Axiom</u>: Every vector space over F has a basis. This is equivalent to the Axiom of Choice.

Theorem 11.2 (Classification of Finite Dimensional Vector Space)
Let V, W be finite dimensional vector space over F. Then

$$V \cong W \iff \dim V = \dim W$$

Proof. ($\to$)Let $T : V \to W$ be an isomorphism, $\mathscr{B} = \{v_1, \ldots, v_n\}$ a basis for V (so $\dim V = n$). By the Monomorphism Theorem,

$$\mathscr{C} = \{Tv_1, \ldots, Tv_n\}$$

is linearly indep. in W. Since $|\mathscr{C}| = n$ and $\text{span}(\mathscr{C}) = w$ (as T is onto), $\mathscr{C}$ is a basis for W and $\dim W = \dim V$.
($\leftarrow$) Suppose $n = \dim V = \dim W$. Let $\mathscr{B} = \{v_1, \ldots, v_n\}$ be a basis for V, $\mathscr{C} = \{w_1, \ldots, w_n\}$ a basis for W. By the UPVS, $\exists! T : V \to W$ linear $v_i \mapsto w_i \forall i$, i.e., T takes the basis $\mathscr{B}$ of V to the basis $\mathscr{C}$ of W. By the Isomorphism Theorem, T is an isomorphism. □

Example 11.3
1. $F^{n \times m} \cong F^{m \times n} \cong F^{mn}$
2. $M_n F \cong F^{n^2}$
3. $F[t]_n \cong F^{n+1}$

Let $T : V \to W$ be linear with V, W arbitrary. Since T only tells us about im T, we replace the target W by im $T = T(V)$, i.e., view $T : V \to W$ surjective linear. Let $\mathscr{B}_0$ be a basis for $\ker T \subset V$ subspace. Then Extension. Theorem holds even when V is not finite dimensional. Extend $\mathscr{B}_0$ to a basis $\mathscr{B} = \mathscr{B}_0 \cup \mathscr{C}$ so $\mathscr{C} \cap \mathscr{B}_0 = \emptyset$ and $V = \text{span } \mathscr{B}$. By the argument proving the Dimension Theorem,

$$T(\mathscr{C}) = \{T(y) | y \in \mathscr{C}\}$$

is linearly indep. and since T is onto $T(\mathscr{C})$ is a basis for W. The new relation in $W = \text{ im } T$ comes from

$$Tx = 0, x \in \mathscr{B}_0$$

In the extra section (3), we showed

$$V/\ker T = \{\overline{v} | v \in V\}$$

where

$$\overline{v} = v + \ker T = \{v + z | z \in \ker T\}$$

is a vector space over F. In fact, $\{\overline{y} | y \in \mathscr{C}\}$ is a basis for $V/\ker T$. By the UPVS, $\exists!$ linear transformation

$$\overline{T} : V/\ker T \to W$$

given by $\overline{0} = \overline{x} \mapsto 0, x \in \mathscr{B}_0, \overline{y} \mapsto Ty, y \in \mathscr{C}$. $\overline{T}$ is clearly onto and $\overline{T}$ is $1-1$,

$$\overline{T}(\overline{v}) = T(v) \quad \forall v \in V$$

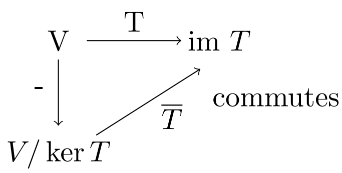
So

$$\overline{T} : V/\ker T \to W = \text{ im } T$$

is an isomorphism.
As $- : V \to V/\ker T$ by $v \mapsto \overline{v}$ is a surjective linear transformation, by definition,

$$\overline{\alpha v + v'} = \alpha \overline{v} + \overline{v'}$$

<u>*Note*</u>: $\ker - = \ker T$.
We have a commutative diagram



with - an epimorphism
$\overline{T}$ an isomorphism

Notiece if $W \neq \text{ im } T, \overline{T}$ is only a monomorphism.
We shall show that all of this is true without using bases (or the Extension Theorem in the Extra Lecture). In particular,

$$V/\ker T \cong \text{ im } T$$

§11.2 Matrices and Linear Transformations

Goal: Let V, W be finite dimensional vector spaces over F. Reduce the study of linear transformations $T : V \to W$ to matrix theory, hence often to computation (Deabstractify).

> **Remark 11.4.** In this section, all bases are ORDERED.

Set up and Notation: Let V, W be finite dimensional vector space over F. $\mathscr{B} = \{v_1, \ldots, v_n\}$ an ordered basis for V, so $\dim V = n$. $\mathscr{C} = \{w_1, \ldots, w_m\}$ an ordered basis for W, so $\dim W = m$.
Step 1: If $v \in V$, write

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

i.e., $\alpha_1, \ldots, \alpha_n$ are the unique coordinate of v relative to $\mathscr{B}$. Then let

$$[v]_{\mathscr{B}} := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in F^{n\times 1}$$

the coordinate matrix of v relative to the ordered basis $\mathscr{B}$. E.g.,

$$[v_i]_{\mathscr{B}} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{pmatrix} i^{\text{th}}$$

and set

$$v_{\mathscr{B}} := \{[v]_{\mathscr{B}} | v \in V\} = F^{n\times 1}$$

Then

$$v \to v_{\mathscr{B}} \quad \text{by } v \mapsto [v]_{\mathscr{B}} \quad \text{isomorphism}$$

as

$$v_i \mapsto e_i := \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} i^{\text{th}}, f_{n,1} = \{e_1, \ldots, e_n\}$$

the standard basis for $F^{n\times 1}$.
Step 2: Let $T : V \to W$ be linear, then

$$Tv_i \in W = \text{Span } \mathscr{C} = \text{Span}(w_1, \ldots, w_m)$$

as $\mathscr{C}$ is a basis for W. Therefore,

$$\exists! \alpha_{ij} \in F, 1 \le i \le m, 1 \le j \le n \ni$$
$$Tv_j = \sum_{i=1}^{m} \alpha_{ij} w_i, \quad j = 1, \ldots, n$$

Let $A = (\alpha_{ij} \in F^{m\times n})$, i.e., $A_{ij} = \alpha_{ij} \forall i, j$. Then the j^{th} COLUMN of A is

$$\begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} = [Tv_j]_{\mathscr{C}} \in W_{\mathscr{C}} = F^{m\times 1}$$

Step 3: Let

$$A : V_{\mathscr{B}} \to W_{\mathscr{C}} \text{ by } A\left([v]_{\mathscr{B}}\right) = A \cdot [v]_{\mathscr{B}}$$

This is a linear transformation.

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

Since

$$A\left([v_j]_{\mathscr{B}}\right) = [Tv_j]_{\mathscr{C}}, j = 1, \dots, n$$

A is the unique linear transformation s.t.

$$A[v_j]_{\mathscr{B}} = [Tv_j]_{\mathscr{C}}$$

So by UPVS,

$$A[v]_{\mathscr{B}} = [Tv]_{\mathscr{C}} \quad \forall v \in V \tag{*}$$

Definition 11.5 (Matrix Representation) — The unique matrix $A \in F^{m\times n}$ in (*) is called the matrix representation of T relative to the ordered bases, $\mathscr{B}, \mathscr{C}$. We denote A by $[T]_{\mathscr{B},\mathscr{C}}$.

Notation: if $V = W$, $\mathscr{B} = \mathscr{C}$, we usually write $[T]_{\mathscr{B}}$ for $[T]_{\mathscr{B},\mathscr{B}}$.

§12 Lec 12: Oct 28, 2020

§12.1 Lec 11 (Cont'd)

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

$$\mathscr{B} = \{v_1, \dots, v_n\} \text{ an ordered basis for } V, \dim V = n$$
$$\mathscr{C} = \{w_1, \dots, w_n\} \text{ an ordered basis for } W, \dim W = m$$

Then $\exists!\ A = [T]_{\mathscr{B},\mathscr{C}} \in F^{m\times n}$ satisfying

$$A[v]_{\mathscr{B}} = [T]_{\mathscr{B},\mathscr{C}}[v]_{\mathscr{B}} = [Tv]_{\mathscr{C}} \forall v \in V$$

Moreover, if

$$Tv_j = \sum_{i=1}^{m} \alpha_{ij} w_i, \quad j = 1, \dots, n$$

then the j^{th} column of $A = [T]_{\mathscr{B},\mathscr{C}}$ is precisely

$$[Tv_j]_{\mathscr{C}} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} \in F^{m\times 1}$$

i.e.,

$$[T]_{\mathscr{B},\mathscr{C}} = \left(\underbrace{[Tv_1]_{\mathscr{C}} \dots [Tv_n]_{\mathscr{C}}}_{\text{columns}} \right)$$

Warning: If $\mathscr{B}', \mathscr{C}'$ are two other ordered bases for V, W respectively (even the same vectors in $\mathscr{B}, \mathscr{C}$ written in a different order), then in general

$$[T]_{\mathscr{B},\mathscr{C}} \neq [T]_{\mathscr{B}',\mathscr{C}'}$$

Example 12.1 1. Let $\mathscr{B} = \{v_1, \dots, v_n\}, \mathscr{C} = \{w_1, \dots, w_n\}$ be two ordered bases for V. Let

$$T : V \to V \text{ linear by } v_i \mapsto w_i, i = 1, \dots, n$$

Then $[T]_{\mathscr{B},\mathscr{C}} = I$, the identity matrix. Moreover, if

$$Tv_j = w_j = \sum_{i=1}^{n} \alpha_{ij} v_i$$

then

$$[T]_{\mathscr{B}} = [T]_{\mathscr{B},\mathscr{B}} = (\alpha_{ij}) = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{n1} & & \alpha_{nn} \end{pmatrix}$$

2. $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $(\alpha, \beta) \mapsto (\beta, \alpha)$, $\mathscr{S} = \mathscr{S}_2 = \{e_1, e_2\}$, the standard ordered basis for $\mathbb{R}^2$. Then

$$[T]_{\mathscr{S}} = ([Te_1]_{\mathscr{S}}, [Te_2]_{\mathscr{S}}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and if $\mathscr{B}$ is the ordered bases $\mathscr{B} = \{e_2, e_1\}$ then

$$[T]_{\mathscr{S},\mathscr{B}} = ([Te_1]_{\mathscr{B}}, [Te_2]_{\mathscr{B}}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

3. Let $\mathscr{B} = \left\{1, x, x^2, x^3\right\}$ be a basis for $\mathbb{R}[x]_3$, the polynomial functions of degree ≤ 3 (and 0), and

$$D : \mathbb{R}[x]_3 \to \mathbb{R}[x]_3 \text{ differentiation}$$

Find $[D]_{\mathscr{B}}$

$$D \cdot 1 = 0 \text{ so } [D \cdot 1]_{\mathscr{B}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$Dx = 1 \text{ so } [Dx]_{\mathscr{B}} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$Dx^2 = 2x \text{ so } [Dx^2]_{\mathscr{B}} = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$

$$Dx^3 = 3x^2 \text{ so } [Dx^3]_{\mathscr{B}} = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 0 \end{pmatrix}$$

Hence,

$$[D]_{\mathscr{B}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Some more examples

Example 12.2

1. Let $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be counterclockwise rotation by an $\angle\theta$

$$T_\theta e_1 = \cos\theta e_1 + \sin\theta e_2$$
$$T_\theta e_2 = (-\sin\theta) e_1 + \cos\theta e_2$$

So

$$[T_\theta]_{\mathscr{S}} = ([T_\theta e_1]_{\mathscr{S}} [T_\theta e_2]_{\mathscr{S}}) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

2. Let $\mathscr{B} = \{v_1, v_2\}$ be an ordered basis for V and $\mathscr{C} = \{w_1, w_2, w_3\}$ an ordered basis for W. Suppose

$$T : V \to W \text{ by } \begin{cases} Tv_1 = 3w_1 + w_3 \\ Tv_2 = w_1 + 6w_2 + w_3 \end{cases}$$

then $[T]_{\mathscr{B},\mathscr{C}} = \begin{pmatrix} 3 & 1 \\ 0 & 6 \\ 1 & 1 \end{pmatrix}$

3. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the reflection about the e_1, e_2 plane. What is $[T]_{\mathscr{S}}$?

$$e_1 \mapsto e_1$$
$$e_2 \mapsto e_2$$
$$e_3 \mapsto -e_3$$

So $[T]_{\mathscr{S}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

Theorem 12.3 (Matrix Theory)

(MTT) Let V, W be finite dimensional vector space F, $\dim V = n$, $\dim W = m$, and $\mathscr{B}, \mathscr{C}$ ordered bases for V, W. Then the map

$$\phi : L(V, W) \to F^{m \times n} \text{ by } T \mapsto [T]_{\mathscr{B},\mathscr{C}}$$

is an isomorphism. In particular

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

Proof. Left as exercise (Homework). □

Using the fact that $W \to W_{\mathscr{C}}$ is an isomorphism if $w \mapsto [w]_{\mathscr{C}}$ show that

1. ϕ is linear
2. ϕ is onto
3. ϕ is $1 - 1$
4. $\dim L(V, W) = mn$

Theorem 12.4

Let V, W, U be finite dimensional vector space over F with ordered bases $\mathscr{B}, \mathscr{C}, \mathscr{D}$ respectively, $T : V \to W$, $S : W \to U$ linear. Then

$$[S \circ T]_{\mathscr{B},\mathscr{D}} = [S]_{\mathscr{C},\mathscr{D}} \cdot [T]_{\mathscr{B},\mathscr{C}}$$

Proof.

$$\begin{aligned}
[S]_{\mathscr{C},\mathscr{D}}[T]_{\mathscr{B},\mathscr{C}}[v]_{\mathscr{B}} &= [S]_{\mathscr{C},\mathscr{D}}[Tv]_{\mathscr{C}} \\
&= [S(Tv)]_{\mathscr{D}} \\
&= [(S \circ T)(v)]_{\mathscr{D}} \\
&= [S \circ T]_{\mathscr{B},\mathscr{D}} [v]_{\mathscr{B}}
\end{aligned}$$

□

Exercise: Let V, W be finite dimensional vector space over F with $\dim V = \dim W$, $\mathscr{B}, \mathscr{C}$ ordered bases of V, W respectively, $T : V \to W$ linear. Then, T is an isomorphism iff $[T]_{\mathscr{B},\mathscr{C}}$ is invertible.

Let V be a finite dimensional vector space over F, $\dim V = n$, $\mathscr{B}$ an ordered basis for V. Then

$$\phi : L(V, V) \to M_n F \text{ by } T \mapsto [T]_{\mathscr{B}}$$

satisfies all of the following: $\forall T, S \in L(V, V)$

(i) $\phi(T + S) = \phi(T) + \phi(S)$

(ii) $\phi(T \circ S) = \phi(T)\phi(S)$

(iii) $\phi(0_V) = 0_{F^{n\times 1}}$

(iv) $\phi(1_V) = 1_{F^{n\times 1}}$

By the exercise, ϕ is bijection linear transformation. Both $L(V, V)$ and $M_n F$ satisfy all the axioms of a field except (M3) and (M4). We call them (NON COMMUTATIVE) rings and since ϕ preserves all the structure i) – iv) as does its inverse(?), we say ϕ is an ISOMORPHISM of rings

Definition 12.5 (Change of Basis Matrix) — Let V be a finite dimensional vector space over F with ordered bases $\mathscr{B}, \mathscr{C}$. Then the invertible matrix $[1_V]_{\mathscr{B},\mathscr{C}}$ is called a CHANGE OF BASIS MATRIX.

Example 12.6 1. $\mathscr{S} = \{e_1, e_2\}, \mathscr{B} = \{(1,1),(2,1)\}, \mathscr{C} = \{(3,4),(6,1)\}$ ordered bases for $\mathbb{R}^2$.

$$[1_{\mathbb{R}^2}]_{\mathscr{B},\mathscr{S}} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}, \quad [1_{\mathbb{R}^2}]_{\mathscr{S}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$[1_{\mathbb{R}^2}]_{\mathscr{C},\mathscr{S}} = \begin{pmatrix} 3 & 6 \\ 4 & 1 \end{pmatrix}, \quad [1_{\mathbb{R}^2}]_{\mathscr{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

2. $\mathscr{B}$ an ordered basis for V, a finite dimensional vector space over F, $\dim V = n$, then $[1_V]_{\mathscr{B}} = I \in M_n F$

3. V a finite dimensional vector space over F, $\mathscr{B}, \mathscr{C}$ ordered bases for V, then $[1_V]_{\mathscr{B},\mathscr{C}}$ is invertible and

$$\begin{aligned} [1_V]^{-1}_{\mathscr{B},\mathscr{C}} &= [1_V]_{\mathscr{C},\mathscr{B}} \\ [1_V]_{\mathscr{B},\mathscr{C}}[1_V]_{\mathscr{C},\mathscr{B}} &= [1_V]_{\mathscr{C}} \\ &= I \\ &= [1_V]_{\mathscr{C},\mathscr{B}}[1_V]_{\mathscr{B},\mathscr{C}} \end{aligned}$$

4. Apply 3) to 1)

$$\begin{aligned} [1_V]_{\mathscr{S},\mathscr{C}} &= [1_V]^{-1}_{\mathscr{C},\mathscr{S}} = \begin{pmatrix} 3 & 6 \\ 4 & 1 \end{pmatrix}^{-1} = -\frac{1}{21}\begin{pmatrix} 1 & -6 \\ -4 & 3 \end{pmatrix} \\ [1_V]_{\mathscr{B},\mathscr{C}} &= [1_V]_{\mathscr{S},\mathscr{C}}[1]_{\mathscr{B},\mathscr{S}} \\ &= -\frac{1}{21}\begin{pmatrix} 1 & -6 \\ -4 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \\ &= -\frac{1}{21}\begin{pmatrix} -5 & -4 \\ -1 & -5 \end{pmatrix} \end{aligned}$$

Some more examples

Example 12.7 1. Any invertible matrix $A \in M_n F$ is a change of basis matrix for some ordered bases $\mathscr{B}, \mathscr{C}$ for F^n : if $A = (\alpha_{ij})$ is invertible, define

$$v_j = \sum_{i=1}^{n} \alpha_{ij} e_i, \quad \mathscr{B} = \{v_1, \dots, v_n\}$$

Then $A = [A]_{\mathscr{B},\mathscr{S}}$ since A is invertible, so $\mathscr{B}$ is linearly indep., hence a basis by counting and $A = [\mathscr{F}_v]_{\mathscr{B},\mathscr{S}}$.

2. The j^{th} column of $[1_v]_{\mathscr{B},\mathscr{C}}$, V a finite dimensional vector space over F is the j^{th} vector of $\mathscr{B}$ expressed as a linear combo of vectors in $\mathscr{C}$.

3. Generalizing (1), (3) from above example, we get the following crucial computational device: if $V = F^n, \mathscr{B}, \mathscr{C}$ ordered bases for V, then

$$[1_v]_{\mathscr{B},\mathscr{C}} = [1_v]_{\mathscr{S},\mathscr{C}} [1_v]_{\mathscr{B},\mathscr{S}} = [1_v]^{-1}_{\mathscr{C},\mathscr{S}} [1_v]_{\mathscr{B},\mathscr{S}}$$

if we only have $V \cong F^n$, then we have to use an isomorphism $V \to F^n$ – how?

Since $[1_v]_{\mathscr{B},\mathscr{S}}$ and $[1_v]_{\mathscr{C},\mathscr{S}}$ are usually (often?) easy to write down, this is quite useful. What if $V = F^{m \times n}$?

Theorem 12.8 (Change of Basis)
Let V, W be finite dimensional vector space over F with ordered bases $\mathscr{B}, \mathscr{B}'$ for V and $\mathscr{C}, \mathscr{C}'$ for W. Let $T : V \to W$ be linear. Then

$$\begin{aligned}[T]_{\mathscr{B},\mathscr{C}} &= [1_W]_{\mathscr{C}',\mathscr{C}} [T]_{\mathscr{B}',\mathscr{C}'} [1_V]_{\mathscr{B},\mathscr{B}'} \\ &= [1_W]^{-1}_{\mathscr{C},\mathscr{C}'} [T]_{\mathscr{B}',\mathscr{C}'} [1_V]_{\mathscr{B},\mathscr{B}'} \\ &= [1_W]_{\mathscr{C}',\mathscr{C}} [T]_{\mathscr{B}',\mathscr{C}'} [1_V]^{-1}_{\mathscr{B}',\mathscr{B}}\end{aligned}$$

Proof. We have

$$[1_W]^{-1}_{\mathscr{C},\mathscr{C}'} = [1_W]_{\mathscr{C}',\mathscr{C}} \text{ and } [1_V]_{\mathscr{B},\mathscr{B}'} = [1_V]^{-1}_{\mathscr{B}',\mathscr{B}}$$

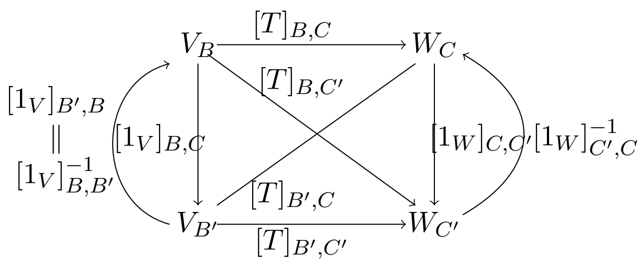
Since

$$\begin{aligned}[1_W]_{\mathscr{C}',\mathscr{C}} [T]_{\mathscr{B}',\mathscr{C}'} [1_V]_{\mathscr{B},\mathscr{B}'} &= [1_W \circ T]_{\mathscr{B}',\mathscr{C}} [1_V]_{\mathscr{B},\mathscr{B}'} \\ &= [1_W \circ T \circ 1_V]_{\mathscr{B},\mathscr{C}} \\ &= [T]_{\mathscr{B},\mathscr{C}}\end{aligned}$$

the result follows. □

To use (and remember) this, do it as follows – to let the notation help you:

$$T : V \to W$$



COMMUTES, i.e., can compose along any allowable arrows in the correct direction if we arrive at the same place in different way starting at the same place we get the same answer.
Warning: You can only reverse direction if the arrow is an isomorphism and then you can take the inverse. To remember the theorem, we write

$$T : V \to W$$

$$\begin{array}{ccc} V_B & \xrightarrow{[T]_{B,C}} & W_C \\ {\scriptstyle [1_V]_{B,B'}} \downarrow & & \downarrow {\scriptstyle [1_W]_{C,C'}} \\ V_{B'} & \xrightarrow{\phantom{[T]_{B,C}}} & W_{C'} \end{array}$$

and fill in arrows you can find in the diagram before.

§13 Lec 13: Oct 30, 2020

§13.1 Some Examples of Change of Basis

If V, W are finite dimensional vector space over F with ordered bases $\mathscr{B}, \mathscr{C}$ respectively and if $T : V \to W$ is linear

$$[Tv]_{\mathscr{C}} = [T]_{\mathscr{B},\mathscr{C}}[v]_{\mathscr{B}} \forall v \in V$$

<u>*Note*</u>: There is nothing about the bases in which v was written.

1. $V = \mathbb{R}^2$, $\mathscr{S} = \{e_1, e_2\}, \mathscr{B} = \{v_1 = (1,1), v_2 = (2,1)\}$ ordered bases. Find $[T]_{\mathscr{S}}$ in the following (equivalently, $[T]_{\mathscr{S}} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}_{\mathscr{S}} \leftrightarrow T(\alpha, \beta)$)

   (i) $T(1,1) = (2,1)$ and $T(2,1) = (1,1)$

   $$\begin{array}{ccc} V_B & \xrightarrow{[T]_B} & V_B \\ {\scriptstyle [1_V]_{B,S}} \big\downarrow & & \big\downarrow {\scriptstyle [1_V]_{B,S} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}} \\ V_S & \xrightarrow[{[T]_S}]{} & V_S \end{array}$$

   So

   $$\begin{aligned} [T]_{\mathscr{S}} &= [1_V]_{\mathscr{B},\mathscr{S}}[T]_{\mathscr{B}}[1_V]^{-1}_{\mathscr{B},\mathscr{S}} \\ &= \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} -1 & 3 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

   So $T(\alpha, \beta) = (-\alpha + 3\beta, \beta)$

   (ii) $T(1,1) = 6(1,1) + (2,1)$ and $T(2,1) = -2(1,1) + (2,1)$

   $$\begin{array}{ccc} V_B & \xrightarrow{[T]_B} & V_B \\ {\scriptstyle [1_V]_{B,S}} \big\downarrow & & \big\downarrow {\scriptstyle [1_V]_{B,S} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}} \\ V_S & \xrightarrow[{[T]_S}]{} & V_S \end{array}$$

   So

   $$[T]_{\mathscr{S}} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 6 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -8 & 16 \\ -8 & 15 \end{pmatrix}$$

   (iii) $T(1,1) = (3,1)$ and $T(2,1) = (5,1)$

   $$\begin{array}{ccc} V_B & \longrightarrow & V_B \\ \big\downarrow & \searrow^{[T]_{B,S}} & \big\downarrow {\scriptstyle [1]_{B,S}} \\ V_S & \longrightarrow & V_S \end{array}$$

   $$[T]_{\mathscr{B},\mathscr{S}} = ([T(1,1)]_{\mathscr{S}}[T(2,1)]_{\mathscr{S}}) = ([(3,1)][(5,1)]_{\mathscr{S}})$$

   So $[T]_{\mathscr{S}} = [T]_{\mathscr{B},\mathscr{S}}[1_V]^{-1}_{\mathscr{B},\mathscr{S}}$ which is equal to $\begin{pmatrix} 3 & 5 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}^{-1}$

2. Let T be a rotation about the axis $(1,1,1) \in V = \mathbb{R}^3$ of an $\angle\theta$ in the counter-clockwise direction with $(1,1,1)$ up. We will use stuff from 33A – dot product. Normalize $(1,1,1)$ to

$$v_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) = \frac{(1,1,1)}{\|(1,1,1)\|}$$

a unit vector in the DIRECTION of v_1. Find a vector $\perp$ to v_1, say

$$v_2' = (0,1,-1)$$

and normalize it to

$$v_2 = \left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

Let $v_3 = v_1 \times v_2$ the cross product of v_1, v_2. It is orthogonal to v_1 and v_2 and by the right hand rule in the correct orientation

$$v_3 = \begin{pmatrix} i & j & k \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

a unit vector (or use Gram – Schmidt and check you have $v_3 = v_1 \times v_2$ and not $-(v_1 \times v_2)$)

§13.2 Orthonormal Basis

Definition 13.1 (Orthonormal Basis) — Let $\mathscr{B} = \{v_1, v_2, v_3\}$ an ordered bases of vectors of length 1 and each $\perp$ to the others, called an ORTHONORMAL BASIS.

$$\begin{aligned} Tv_1 &= v_1 \\ Tv_2 &= \cos\theta v_2 + \sin\theta v_3 \\ Tv_3 &= -\sin\theta v_2 + \cos\theta v_3 \\ [T]_{\mathscr{B}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \\ [1_V]_{\mathscr{B},\mathscr{S}} &= \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix} \end{aligned}$$

$$\begin{array}{ccc} V_B & \xrightarrow{[T]_B} & V_B \\ {\scriptstyle [1_V]_{B,S}} \downarrow & & \downarrow {\scriptstyle [1_V]_{B,S}} \\ V_S & \xrightarrow[{[T]_S}]{} & V_S \end{array}$$

$$[T]_{\mathscr{S}} = [1_V]_{\mathscr{B},\mathscr{S}}[T]_{\mathscr{B}}[1_V]^{-1}_{\mathscr{B},\mathscr{S}} = [1_V]_{\mathscr{B},\mathscr{S}}[T]_{\mathscr{B}}[1_V]_{\mathscr{S},\mathscr{B}}$$

Since both $\mathscr{S}$ and $\mathscr{B}$ are orthonormal bases and $F = \mathbb{R}$, it turns out that

$$[1_V]^{-1}_{\mathscr{B},\mathscr{S}} = [1_V]^{\top}_{\mathscr{B},\mathscr{S}}$$

This is, however, not true in general.

3. $V = \mathbb{R}^3$, $T : V \to V$ as in 2) and $S : V \to V$ a reflection about the plane $\perp (1,2,3)$. Find $[S]_{\mathscr{S}}$ and $[S \circ T]_{\mathscr{S}}$.

   Find an orthonormal basis with $(1,2,3)$ direction of the first vector

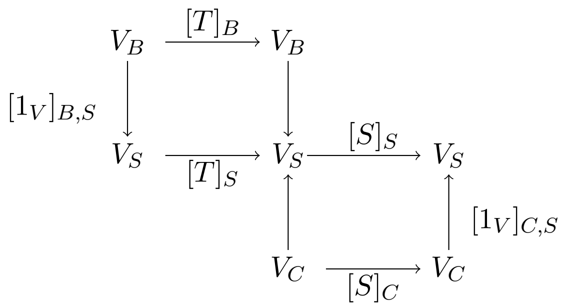
$$(1,2,3),(0,3,-2),(-13,2,3)$$

   then normalize as follows:

$$w_1 = \left(\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, \frac{3}{\sqrt{14}}\right)$$
$$w_2 = \left(0, \frac{3}{\sqrt{13}}, -\frac{2}{\sqrt{13}}\right)$$
$$w_3 = \left(\frac{-13}{\sqrt{182}}, \frac{2}{\sqrt{182}}, \frac{3}{\sqrt{182}}\right)$$

   So $\mathscr{C} = \{w_1, w_2, w_3\}$ is an orthonormal basis and

$$[S]_{\mathscr{C}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



$$[1_V]_{\mathscr{C},\mathscr{S}} = \begin{pmatrix} \frac{1}{\sqrt{14}} & 0 & \frac{13}{\sqrt{182}} \\ \frac{2}{\sqrt{14}} & \frac{3}{\sqrt{13}} & \frac{2}{\sqrt{182}} \\ \frac{3}{\sqrt{14}} & -\frac{2}{\sqrt{13}} & \frac{3}{\sqrt{182}} \end{pmatrix}$$

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

The only reason to normalize $\mathscr{C}$ to an orthonormal basis is

$$[1_V]))\mathscr{C},\mathscr{S}^{-1} = [1_V]^{\top}_{\mathscr{C},\mathscr{S}}$$

§13.3 Similarity

Definition 13.2 (Similar Matrices) — Let $A, B \in M_n F$. We say A is SIMILAR to B write $A \sim B$ if $\exists C \in M_n F$ invertible $\ni$

$$A = C^{-1}BC$$

Remark 13.3. $A, B \in M_n F$:

1. $A \sim B \to B \sim A$:
   $A = C^{-1}BC, C$ invertible $\to B = (C^{-1})^{-1}AC^{-1}$ as $CC^{-1} = I = C^{-1}C$
2. If $A \sim B$, then $\det A = \det B$. If $A = C^{-1}BC$, invertible, then

$$\begin{aligned}\det A &= \det\left(C^{-1}BC\right) = \det(C^{-1})\det B \det C \\ &= (\det C)^{-1}\det B \det C = \det B\end{aligned}$$

3. $\sim$ is an equivalence relation.

Theorem 13.4 (Similar Matrices)
Let $A, B \in M_n F$. Then $A \sim B$ iff $\exists V$ a vector space over F, $\dim V = n$, $T : V \to V$ linear and ordered bases $\mathscr{B}, \mathscr{C}$ for V s.t

$$A = [T]_{\mathscr{B}} \quad \text{and} \quad B = [T]_{\mathscr{C}}$$

i.e., $A \sim B$ iff they represent the same linear transformation relative to (possibly) different ordered bases.

§14 Lec 14: Nov 2, 2020

§14.1 Lec 13 (Cont'd)

Proof. (Of Similar Matrices Theorem) ($\leftarrow$) If $A = [T]_{\mathscr{B}}, B = [T]_{\mathscr{C}}$, then $C = [1_V]_{\mathscr{B},\mathscr{C}} \in M_nF$ is invertible with $A = C^{-1}BC$ by the Change of Basis Theorem.
($\rightarrow$) Suppose $C \in M_nF$ is invertible, $A = C^{-1}BC$. Define $V = F^n$, $T : V \to V$ by

$$T_{ij} = \sum_{i=1}^{n} A_{ij}e_i$$

with $\mathscr{S} = \{e_1, \ldots, e_n\}$ the standard basis

$$[T]_{\mathscr{S}} = A = C^{-1}BC$$

Let $w_j := \sum_{i=1}^{n}(C^{-1})_{ij}e_i$, i.e., $(C^{-1})_{ij}$ is the ij^{th} entry of C^{-1}. As C is invertible, C^{-1} exists and is invertible. Then

$$\mathscr{B} = \{w_1, \ldots, w_n\}$$

is a basis for V and $[1_V]_{\mathscr{B},\mathscr{S}} = C^{-1}$ figure here so $A = C^{-1}[T]_{\mathscr{B}}C$ and $B = [T]_{\mathscr{B}}$ works. □

§14.2 Eigenvalues and Eigenvectors

Definition 14.1 (Eigenvalues, Eigenvectors & Eigenspace) — Let $0 \neq V$ be a vector space over F, $T : V \to V$ a linear operator and $\lambda \in F$. Set

$$S_\lambda := T - \lambda 1_V : V \to V,$$

a linear operator, so

$$S_\lambda(v) = Tv - \lambda v \forall v \in V$$

We say λ is an EIGENVALUE of T if S_λ is not $1-1$, i.e., $\ker S_\lambda \neq 0$. Let

$$\begin{aligned} E_T(\lambda) := \ker S_\lambda &= \{v \in V | Tv - \lambda v = 0\} \\ &= \{v \in V | Tv = \lambda v\} \end{aligned}$$

if $E_T(\lambda) \neq 0$, we call $E_T(\lambda)$ an EIGENSPACE of V relative T, λ and any $v \in E_T(\lambda)$ an EIGENVECTOR of T relative to λ. So if $T : V \to V$ is linear, $\lambda \in F$ is an eigenvalue of T iff

$$\exists 0 \neq v \in V \ni Tv = \lambda v$$

Remark 14.2. Let $0 \neq V$ be a vector space over F and $T : V \to V$ linear

1. Eigenvalues occur as measured quantities in science and engineering, e.g., resonance, quantum number – measurable values.
2. If $\lambda \in F$ is an eigenvalue of T, then

$$0 \neq E_T(\lambda) \subset V \text{ is a subspace}$$

3. If $\lambda \in F$ an eigenvalue, any $v \in E_T(\lambda)$ is an eigenvector. In particular, any basis for $E_T(\lambda)$ consists of eigenvectors of T relative to λ. Hence

$$T\Big|_{E_T(\lambda)} = \lambda 1_{E_T(\lambda)}$$

(the notation above means we restrict the domain to $E_T(\lambda)$. In particular, if $V = E_T(\lambda)$, then $T = \lambda 1_V$.

4. If $T = 0$, then $V = E_T(\lambda)$ with eigenvalue $\lambda = 0(\lambda = 1)$.

Example 14.3 5. Let $V = \mathbb{R}^3$, $T : V \to V$ a counterclockwise rotation by an $\angle\theta, 0 < \theta < 2\pi$ around the axis determined by $0 \neq v \in V$. Then

$$T(\alpha v) = \alpha Tv = \alpha v \forall \alpha \in F$$

So $\text{Span}(v) \subset E_T(1)$. Note if $0 \neq v$ is an eigenvector with eigenvalue μ of linear $S : V \to V$, then

$$Sv \in \text{Span}(v) = Fv \text{ so } \text{Span}(v) \subset E_S(\mu)$$

Do there exist other eigenvalues of T? Ever? So the only other possibilities would be

$$\theta = \pi, \lambda = -1$$

In that case

$$E_T(-1) = \text{Span}(w_1, w_2)$$

where w_1, w_2 are linearly indep. with $w_i \perp v, i = 1, 2$. (of course, if one allows $\theta = 0, T = 1_V$.)

6. Let $0 \neq v \in V$. Suppose that

$$\mu v = Tv = \lambda v, \quad \lambda, \mu \in F$$

Then $\mu = \lambda$ so $0 \neq v \in V$ is an eigenvector of at most one eigenvalue of T – usually none. In particular,

$$E_T(\lambda) \cap E_T(\mu) = 0 \text{ if } \lambda \neq \mu$$

and we write

$$E_T(\lambda) \oplus E_T(\mu) = E_T(\lambda) + E_T(\mu)$$

and call it the DIRECT SUM of the subspace $E_T(\lambda)$ and $E_T(\mu)$.

What do you think is $W_1 \bigoplus W_2 \bigoplus W_3$?

7. Suppose $\dim V = n$, $\mathscr{B} = \{v_1, \dots, v_n\}$ is an ordered basis for V. Suppose that that

$$Tv_i = \alpha_i v_i, \qquad i = 0, \dots, n$$

$\lambda_1, \dots, \lambda_n \in F$ not necessarily distinct. Then

$$[T]_{\mathscr{B}} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_n \end{pmatrix}$$

is a DIAGONAL MATRIX, i.e., all non-diagonal entries 0. We say T is DIAGONALIZABLE if $\exists$ an ordered bases $\mathscr{C}$ for $V \ni [T]_{\mathscr{C}}$ is diagonal.

8. Suppose $\dim V = n (< \infty)$ and T is diagonalizable, i.e., $\exists$ an ordered basis $\mathscr{C} = \{w_1, \dots, w_n\}$ for V s.t.

$$[T]_{\mathscr{C}} = \begin{pmatrix} \mu_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mu_n \end{pmatrix}$$

Then $Tw_i = \mu_i w_i, i = 1, \dots, n$ and $\mathscr{C}$ is an ordered basis for V consisting of eigengenvalues for T.

Conclusion: Let V be a finite dimensional vector space over F, $T : V \to V$ linear. Then T is diagonalizable iff $\exists$ a basis for V consisting of eigenvectors of T.
Note: If T is diagonalizable, $T : V \to V$ linear, V a finite dimensional vector space over F, ordered basis $\mathscr{B}$ for V. Then $\exists C \in M_n F$, invertible, $n = \dim V \ni C^{-1}[T]_{\mathscr{B}} C$ is diagonal by the Change of Basis Theorem.

Example 14.4 9. Let V be a finite dimensional vector space over F, $n = \dim V$, $\mathscr{B}$ an ordered basis for V, $S : V \to V$ linear. Then by the Isomorphism Theorem, S is 1-1 iff S is onto. Apply this to

$$S_\lambda = T - \lambda 1_V : V \to V$$

to conclude:

λ is an eigenvalue of T iff $S_\lambda = T - \lambda 1_V$ is singular (i.e., S_λ is not 1-1)

iff

$$[S_\lambda]_{\mathscr{B}} = [T - \lambda 1_V]_{\mathscr{B}} \text{ is not invertible}$$

iff

$$\det[T - \lambda 1_V]_{\mathscr{B}} = 0 \text{ (by properties of det)}$$

iff

$$\det\left([T]_{\mathscr{B}} - \lambda [1_V]_{\mathscr{B}}\right) = 0$$

iff

$$\det\left([T]_{\mathscr{B}} - \lambda I\right) = 0$$

iff

$$\det\left(\lambda I - [T]_{\mathscr{B}}\right) = 0$$

Summary: Let V be a finite dimensional vector space over F, $\dim V = n$, $T : V \to V$ linear, $\mathscr{B}$ an ordered basis for V, $\lambda \in F$. Then, λ is an eigenvalue of T iff $\det(\lambda I - [T]_{\mathscr{B}}) = 0$.

Definition 14.5 (Characteristics Polynomial) — Let $A \in M_n F$. Define

$$f_A := \det(tI - A) \in F[t]$$

called the Characteristics Polynomial of A.

The properties of the determinant on $F[t]$ is the same as on F except that $A \in M_n F[t]$ is invertible iff $\det A \in F \setminus \{0\}$ and we assume these properties.

Proposition 14.6

If $A, B \in M_n F$ are similar, then $f_A = f_B$

Proof. If $A = C^{-1} B C$, $C \in M_n F$ in

$$\begin{aligned} f_A &= \det(C^{-1}(tI - B)C) = \det C^{-1} \det(tI - B) \det C \\ &= \det(tI - B) = f_B \end{aligned}$$

□

Warning: Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Then, A and B are not similar, but $f_A = f_B$, i.e., the converse is false.

Corollary 14.7

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

$$f_{[T]_{\mathscr{B}}} = f_{[T]_{\mathscr{C}}}$$

Proof. Change of Basis Theorem. □

Definition 14.8 (Characteristics Polynomial) — Let V be a finite dimensional vector space over F, $T : V \to V$ linear, $\mathscr{B}$ ordered basis for V. We call $f[t]_{\mathscr{B}}$ the characteristics polynomial of T. By the corollary, it is independent of $\mathscr{B}$, so we denote it by $f_T(= f_{[T]_{\mathscr{B}}})$ and write $f_T = \det(t1_V - T) := \det(tI - [T]_{\mathscr{B}})$

Theorem 14.9

Let V be a finite dimensional vector space over F, $T : V \to V$ linear. Then, the eigenvalues of T are precisely, the roots of f_T, i.e., those $\alpha \in F \ni f_T(\alpha) = 0$.

Proof. det $\lambda \in F, \mathscr{B}$ an ordered basis for V. Set $A = [T]_{\mathscr{B}}$, so $f_T = \det(tI - A)$. Then λ is a root of f_T iff evaluating f_T at λ, i.e., $f_T(\lambda)$, we have

$$f_T(\lambda) = \det(tI - A)\Big|_{t=\lambda} = 0 \iff \lambda \text{ is an eigenvalue of } T$$

i.e., expanding the polynomial $\det(tI - A)$ and plugging λ for t gives 0. □

We cannot use the following theorem if we fully prove it.

Theorem 14.10 (Cayley – Hamilton)

Let $A \in M_n F$. Then

$$f_A(A) = 0$$

plugging A into the expansion of the determinant f_A, you get 0.

Remark 14.11. By HW, we have $\left\{I, A, A^2, \ldots, A^{n^2}\right\} \subset M_n F$ is linearly dep., i.e., $\left\{I, A, \ldots, A^N\right\}$ is linearly dep. for some $N > 0$. This means $\exists 0 \neq g \in F[t]$ with deg $g \leq N$ and $g(A) = 0$ – why?

So Cayley – Hamilton's Theorem says $\{I, A, \ldots, A^n\}$ in $M_n F$ is always linearly dep. in $M_n F$ with $f_A(A)$ giving a dependence relation.
Note: If you know Cramer's Rule in determinant theory, one can prove Cayley – Hamilton follows from it. In fact, it is essentially Cramer's Rule.

Remark 14.12. Let V be a finite dimensional vector space over F, $T : V \to V$ linear. You will show in your Take home Exam. There exists a polynomial $q \in F[t]$ satisfying

1. $q \neq 0$
2. $q(A) = 0$
3. deg q is the minimal degree for a poly $g \neq 0$ in $F[t]$ to satisfy $g(A) = 0$
4. q is MONIC, i.e., leading coeff is 1.

Moreover, q is unique and called the MINIMAL POLYNOMIAL of A and denoted q_T. Using it we shows a stronger form of the Cayley – Hamilton Theorem.

§15 Lec 15: Nov 4, 2020

§15.1 Lec 14 (Cont'd)

Cayley – Hamilton (Stronger Form): Let V be a finite dimensional vector space over F, $T : V \to V$ linear, then

$$q_T | f_T \text{ in } F[t]$$

(where $q_T = q[T]_{\mathscr{B}}$, $\mathscr{B}$ an ordered basis and q_T is indep. of $\mathscr{B}$). Why does this show the other form?

Computation: Let V be a finite dimensional vector space over F, $T : V \to V$ linear. To find eigenvalues and eigenvectors of T, you must solve

$$Tv = \alpha v$$

By Matrix Theory Theorem, this is equivalent to

$$[T]_{\mathscr{B}}[v]_{\mathscr{B}} = \lambda[v]_{\mathscr{B}} \tag{*}$$

$\mathscr{B}$ an ordered basis for V. To find eigenvalues, we find the roots of f_T. To find the eigenvectors, we solve (*).

Theorem 15.1

Let $T : V \to V$ be linear and $\lambda_1, \ldots, \lambda_n$ in F distinct eigenvalues of $T, 0 \neq v_i \in E_T(\lambda_i), i = 1, \ldots, n$. Then $\{v_1, \ldots, v_n\}$ is linearly indep.

Proof. We induct on n.

- $n = 1 : v_1 \neq 0$ so $\{v\}$ is linearly indep.
- $n > 1$ – Induction Hypothesis (IH) : If $\lambda_1, \ldots, \lambda_{n-1}$ are distinct eigenvalues of $T, 0 \neq v_i \in E_T(\lambda_i), i = 1, \ldots, n-1$ then $\{v_1, \ldots, v_{n-1}\}$ is linearly indep. Suppose that

  $$0 = \alpha_1 v_1 + \ldots + \alpha_n v_n, \alpha_1, \ldots, \alpha_n \in F \tag{*}$$

  Apply the linear operator $S_{\lambda_n} = T - \lambda_n 1_V$ to (*). As

  $$S_{\lambda_n}(v_i) = Tv_i - \lambda_n v_i = \lambda_i v_i - \lambda_n v_i = (\lambda_i - \lambda_n)v_i$$

  We get

  $$S_{\lambda_n}(\alpha_1 v_1 + \ldots + \lambda_n v_n) = \alpha_1 S_{\lambda_{v_n}} v_1 + \ldots + \alpha_n S_{\lambda_{v_n}} v_n$$
  $$0 = \alpha_1(\alpha_1 - \alpha_n)v_1 + \ldots + \alpha_{n-1}(\lambda_{n-1} - \lambda_n)v_{n-1}$$

  By the IH, $\alpha_i(\lambda_i - \lambda_n) = 0, i = 1, \ldots, n-1$

  As $\lambda_i - \lambda_n \neq 0, i = 1, \ldots, n-1, \alpha_i = 0, i = 1, \ldots, n-1$. So $0 = \alpha_n v_n$. As $v_n \neq 0$, $\alpha_n = 0$ also.

□

Proof. (Alternative) Take T of (*) to get an eqn 1). Multiply (*) by λ_n to get an eqn 2). Subtract eqn 2) from eqn 1). The proof that if $\alpha_1, \ldots, \alpha_n$ are distinct then $e^{\lambda_1 x}, \ldots, e^{\lambda_n x}$ are linearly indep. □

Corollary 15.2

Let V be a finite dimensional vector space over F, $\dim V = n$ if $T : V \to V$ linear has n distinct eigenvalues, then T is diagonalizable. The converse is false, e.g., $T = 1_V$.

Corollary 15.3

If V is a finite dimensional space over F, $\dim V = n$, $T : V \to V$ linear, then T has at most n distinct eigenvalues. This also follows as any $0 \neq f \in F[t]$ has at most $\deg f$ roots.

Corollary 15.4

Let V be a vector space over F, $T : V \to V$ linear, $\lambda_1, \ldots, \lambda_n$ distinct eigenvalues of T. Set

$$w = E_T(\lambda_1) + \ldots + E_T(\lambda_n)$$

if $v_i \in E_T(\lambda_i), i = 1, \ldots, n$ satisfy

$$v_1 + \ldots + v_n = 0$$

then $v_i = 0, i = 1, \ldots n$. We write this as

$$W = E_T(\lambda_1) \oplus \ldots \oplus E_T(\lambda_n)$$

Exercise 15.1. Let V be a vector space over F, $W_1, \ldots, W_n \subset V$ subspaces. Let $W = W_1 + \ldots + W_n$. Then the followings are equivalent

1. If $w_i \in W_i, i = 1, \ldots, n$ satisfy $w_1 + \ldots + w_n = 0$ then $w_i = 0 \forall i$. We say W_i are indep.
2. If $v \in W \exists ! w_i \in W_i \ni v = w_1 + \ldots + w_n$
3. $W_i \cap \sum_{j \neq i, j=1}^{n} W_j = 0 \forall i = 1, \ldots, n$
4. If $\mathscr{B}_i$ is a basis for $W_i, i = 1, \ldots, n$ then $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_n$ is a basis for W.

If these hold for W, we say W is an (internal) direct sum of the W_i and write

$$W = W_1 \oplus \ldots \oplus W_n$$

Remark 15.5. This generalizes to $W = \oplus W_i$, general I – How. What is the proof?

Exercise 15.2. Let V be a vector space over F, $W_1, \ldots, W_n \subset V$ subspaces $\ni V = W_1 + \ldots + W_n$. Let

$$W = W_1 \times \ldots \times W_n = \{(W_1, \ldots, W_n) | w_i \subset W_i \forall i\}$$

a vector space over F via component wise operations. Show

$$v = W_1 \oplus \ldots \oplus W_n \iff T : W_1 \times \ldots \times W_n \to V$$

by $(w_1, \ldots, w_n) \mapsto w_1 + \ldots w_n$ is an isomorphism. We call W the external direct sum of the W_i.

Consequences: Let V be a finite dimensional vector space over F, $\lambda_1, \dots, \lambda_n$ distinct eigenvalues of $T : V \to V$ linear, $?_i = \dim E_T(\lambda_i), \mathscr{B}_i$ ordered basis for $E_T(\lambda_i), i = 1, \dots, n$ if

$$V = E_T(\lambda_1) + \ldots + E_T(\lambda_n)$$

then

$$V = E_T(\lambda_1) \oplus \ldots \oplus E_T(\lambda_n)$$

and $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_n$ is an ordered basis for V and

$$[T]_{\mathscr{B}} = \begin{pmatrix} \left[\lambda_1 1_{E_T(\lambda_1)}\right]_{\mathscr{B}_1} & & \\ & \ddots & \\ & & \left[\lambda_n 1_{E_T(\lambda_n)}\right]_{\mathscr{B}_n} \end{pmatrix}$$

(Block form) is a diagonal matrix. In particular,

$$f_T = \det(T1_V - T) = (t - \lambda_1)^{r_1} \ldots (t - \lambda_n)^{r_n}$$

By determinant theory,

$$\det \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \det A \det B$$

A, B square matrices and T is diagonalizable.

> **Remark 15.6.** $T : V \to V$ linear may or may not have eigenvalues
>
> 1. $V = \mathbb{R}^2, f_T = t^2 + 1$, then T has not eigenvalues.
> 2. If V is a finite dimensional vector space over $\mathbb{C}$, then T has an eigenvalue as f_T has a root by the FUNDAMENTAL THEOREM OF ALGEBRA (which we shall always assume to be true).

§15.2 Inner Product Space

We know that the dot product of vectors in $\mathbb{R}^3$ allows us to define $\perp$, $\angle$, distance, etc. We want to generalize this to "inner product spaces". When we talk about inner product spaces, we shall always assume that OUR FIELD F LIES in $\mathbb{C}$ (e.g., $\mathbb{Q}, \mathbb{R}, \mathbb{C}$) as a subfield.
Let $- : \mathbb{C} \to \mathbb{C}$ by $\alpha + \beta\sqrt{-1} \mapsto \alpha - \beta\sqrt{-1} \forall \alpha, \beta \in \mathbb{R}$ denoted complex conjugation.
Note:Let $a = \alpha + \beta\sqrt{-1}$ in $\mathbb{C}$, $\alpha, \beta \in \mathbb{R}$. Then

1. $a = \overline{a}$ iff $a \in \mathbb{R}$
2. $\overline{\overline{a}}$
3. $|a|^2 := a\overline{a} \geq 0$ in $\mathbb{R}$ as $a\overline{a} = \alpha^2 + \beta^2$ and $= 0$ iff a= 0.

   As we shall assume $F \subset \mathbb{C}$, we define:

   $$\overline{F} := \{\overline{z} \in \mathbb{C} | z \in F\}$$

   and we shall also assume that

   $$F = \overline{F}$$

   This is true if $F \subset \mathbb{R}$ or $F = \mathbb{C}$, but does not always hold UNLESS we only consider those F that do which we will.

Definition 15.7 (Inner Product Space) — Let $F \subset \mathbb{C}$ be a subfield satisfying $F = \overline{F}, V$ a vector space over F. We call V an inner product space over F, <u>write</u> V is an ips / F, under the map

$$\langle , \rangle := \langle , \rangle_V : V \times V \to F$$

<u>Write</u>: $\langle v, w \rangle$ for $\langle , \rangle(v, w)$ if $\langle , \rangle$ satisfies $\forall v_1, v_2, v_3, v \in V, \forall \alpha \in F$

1. $\langle v_1 + v_2, v_3 \rangle = \langle v_1, v_3 \rangle + \langle v_2, v_3 \rangle$
2. $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle$
3. $\langle \alpha v_1, v_2 \rangle = \alpha \langle v_1, v_2 \rangle = \langle v_1, \overline{\alpha} v_2 \rangle$
4. $\langle v, v \rangle \in \mathbb{R}$ and $\langle v, v \rangle \geq 0$ with $\langle v, v \rangle = 0$ iff $v = 0$.

If V is an inner product space over F (under $\langle , \rangle$, the LENGTH (or NORM or MAGNITUDE) of $v \in V$ is given by

$$\|v\| := \sqrt{\langle v, v \rangle} \geq 0 \in \mathbb{R}$$

<u>Note</u>: If $F < \mathbb{C}$, $\|v\|^2 \in F$, but it is possible that $\|v\| \notin F$, e.g., if $V = \mathbb{Q}^2$ a vector space over $\mathbb{Q}$ and an inner product space over $\mathbb{Q}$ under the dot product $\|(1,1)\| = \sqrt{2} \notin \mathbb{Q}$. This is a reason to work only with $F = \mathbb{R}$ or $\mathbb{C}$.

§16 Lec 16: Nov 6, 2020

§16.1 Lec 15 (Cont'd)

Properties: Let V be an inner product space over F, $\alpha \in F, v_1, v_2, v_3 \in V$.

1. $\langle 0, v\rangle = 0 = \langle w, 0\rangle, \forall v, w \in V$.
2.
   - $\langle \alpha v_1 + v_2, v_3\rangle = \alpha\langle v_1, v_3\rangle + \langle v_2, v_3\rangle$
   - $\langle v_1, \alpha v_2 + v_3\rangle = \overline{\alpha}\langle v_1, v_2\rangle + \langle v_1, v_3\rangle$
3. If $F \subset \mathbb{R}$ define the ANGLE $\theta, 0 \le \theta \le 2\pi$ between $v_1 \neq 0$ and $v_2 \neq 0$ in V by

   $$\cos\theta := \frac{\langle v_1, v_2\rangle}{\|v_1\|\|v_2\|}$$

   and if $F \not\subset \mathbb{R}$ define θ by

   $$\cos\theta := \frac{|\langle v_1, v_2\rangle|}{\|v_1\|\|v_2\|}$$

   Note: This does not make sense yet, and will not until we show

   $$\frac{|\langle v_1, v_2\rangle|}{\|v_1\|\|v_2\|} \le 1 \quad \text{for } v_1 \neq 0, v_2 \neq 0$$

4. (very useful prop) Let $v \in V$. If $\langle v, w\rangle = 0, \forall w \in V$ (or $\langle w, v\rangle = 0 \forall w \in W$), then $v = 0$.
5. Let $0 \neq x \in V$. Then

   $$\langle , x\rangle : V \to F \text{ by } v \mapsto \langle v, x\rangle$$

   is a linear transformation, i.e., linear functional, i.e., $\langle , x\rangle \in V^*$. However,

   $$\langle x, \rangle : V \to F \text{ by } v \mapsto \langle x, v\rangle$$

   is linear iff $F \subset \mathbb{R}$. In general, we say that $\langle x, \rangle$ is SESQUILINEAR as $\forall \alpha \in F, \forall v_1, v_2 \in V$

   $$\langle x, \alpha v_1 + v_2\rangle = \overline{\alpha}\langle x, v_1\rangle + \langle x, v_2\rangle$$

   Of course if $x = 0$, $\langle 0, \rangle\langle , 0\rangle \in V^*$.

Example 16.1

Let $F \subset \mathbb{C}, F = \overline{F} = \{\overline{\alpha} | \alpha \in F\}$. The following V vector space over F are inner product space over F under the given $\langle , \rangle$:

1. $V = F^n$ and $\langle , \rangle = \underbrace{\cdot}_{\text{dot product}}$, i.e., if

$$v = (\alpha_1, \dots, \alpha_n), w = (\beta_1, \dots, \beta_n), \alpha_i, \beta_i \in F, \forall i, j$$

   Then,

$$\langle v, w \rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}$$

   Note: If $F \subset \mathbb{R}$, then

$$\langle v, w \rangle = \sum_{i=1}^{n} \alpha_i \beta_i$$

2. Let $I = [\alpha, \beta], \alpha < \beta$ in $\mathbb{R}$, $V = C(I)$ with $C(I) = \{f : I \to \mathbb{R} | f \text{ cont}\}$ then

$$\langle f, g \rangle := \int_{\alpha}^{\beta} fg$$

   Think about what if $C_{\mathbb{C}} := \{f : I \to \mathbb{C} | f \text{ cont}\}$.

3. In 2), let $h \in C(I)$ satisfy $h(x) > 0 \forall x \in I$. Then

$$\langle f, g \rangle_h := \int_{\alpha}^{\beta} hfg$$

   the WEIGHTED INNER PRODUCT SPACE via h.

4. Let $A \in M_n F$. Define the adjoint of A to be A^* where

$$(A^*)_{ij} := \overline{A}_{ji}, \quad \forall i, j$$

   the conjugate transpose of A., i.e., $A^* = \overline{A}^{\top}$. So if $F \subset \mathbb{R}, A^* = A^{\top}$.

Remark 16.2. If $A = F^{m \times n}$, then A^* defined by $(A^*)_{ij} = \overline{A}_{ji}$ still makes sense and is called the ADJOINT of A. What can you say about AA^* and A^*A?

Let $V = M_n F$ under

$$\langle A, B \rangle := \text{tr}(AB^*)$$

where $\text{tr } C = \sum_{i=1}^{n} C_{ii}$. So if $F \subset \mathbb{R}, \langle A, B \rangle = \text{tr}(AB^{\top})$.

tr=trace

Example 16.3 5. Let $F = \mathbb{R}$

$$l_2 := \left\{(a_0, a_1, \ldots, a_n, \ldots) \,|a_i \in \mathbb{R} \forall i - \text{infinite seq with } \sum a_i^2 < \infty\right\}$$

a vector space over F by component wise operation (a subspace of $\mathbb{R}^{\infty}_{\text{inf}}$ – see below) and an inner product space over $\mathbb{R}$ via

$$\langle v, w \rangle := \sum_{i=0}^{\infty} a_i b_i \in \mathbb{R}$$

if $v = (a_0, a_1, \ldots), w = (b_0, b_1, \ldots)$

$$0 \le (a_i \pm b_i)^2 = a_i^2 \pm 2a_i b_i + b_i^2, \forall i \text{ so}$$
$$\mp 2 \sum_{i=0}^{\infty} a_i b_i \le \sum_{i=0}^{\infty} a_i^2 + \sum_{i=0}^{\infty} b_i^2 < \infty$$

Theorem 16.4
Let V be an inner product space over F. Then $\forall v_1, v_2 \in V, \forall \alpha \in F$, we have

1. $\|v_1\| \in \mathbb{R}$ with $\|v_1\| \ge 0$ and $\|v_1\| = 0$ iff $v_1 = 0$.
2. $\|\alpha v_1\| = |\alpha| \|v_1\|$.
3. Cauchy – Schwarz Inequality
$$|\langle v_1, v_2 \rangle| \le \|v_1\| \|v_2\|$$
4. Minkowski Inequality(special case)
$$\|v_1 + v_2\| \le \|v_1\| + \|v_2\|$$

Proof. 1) and 2) are left as exercise.
3) If $v_1 = 0$ or $v_2 = 0$, the result is immediate, so we may assume that $v_1 \ne 0, v_2 \ne 0$. We use the following important trick. Take the orthogonal projection. Let

$$v = v_2 - \underbrace{\frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} v_1}_{\text{orthogonal projection on } v_1}$$

Claim 16.1. $\langle v, \alpha v_1 \rangle = 0 \forall \alpha \in F$ (i.e., $v \perp \alpha v_1$)

$$\begin{aligned}
\langle v, \alpha v_1 \rangle &= \langle v_2 - \frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} v_1, \alpha v_1 \rangle \\
&= \langle v_2, \alpha v_1 \rangle + \langle -\frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} v_1, \alpha v_1 \rangle \\
&= \overline{\alpha} \langle v_2, v_1 \rangle - \frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} \langle v_1, \alpha v_1 \rangle \\
&= \overline{\alpha} \langle v_2, v_1 \rangle - \frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} \overline{\alpha} \|v_1\|^2 = 0
\end{aligned}$$

establishing the claim. Therefore, we have

$$\begin{aligned}
0 \leq \langle v, v\rangle &= \langle v, v_2 - \frac{\langle v_2, v_1\rangle}{\|v_1\|^2} v_1\rangle \\
&= \langle v, v_2\rangle + \langle v_1 - \frac{\langle v_2, v_1\rangle}{\|v_1\|^2} v_1\rangle = \langle v, v_2\rangle \\
&= \langle v_2 - \frac{\langle v_2, v_1\rangle}{\|v_1\|^2} v_1, v_2\rangle = \langle v_2, v_2\rangle - \frac{\langle v_2, v_1\rangle}{\|v_1\|^2}\langle v_1, v_2\rangle \\
&= \|v_2\|^2 - \frac{\overline{\langle v_1, v_2\rangle}}{\|v_1\|^2}\langle v_1, v_2\rangle = \|v_2\|^2 - \frac{|\langle v_1, v_2\rangle|^2}{\|v_1\|^2}
\end{aligned}$$

So

$$|\langle v_1, v_2\rangle|^2 \leq \|v_1\|^2 \|v_2\|^2$$

or

$$|\langle v_1, v_2\rangle| \leq \|v_1\| \|v_2\|$$

as required. □

Proof. 4.

$$\begin{aligned}
\|v_1 + v_2\|^2 &= \langle v_1 + v_2, v_1 + v_2\rangle \\
&= \|v_1\|^2 + \langle v_1, v_2\rangle + \langle v_2, v_1\rangle + \|v_2\|^2 \\
&= \|v_1\|^2 + \langle v_1, v_2\rangle + \overline{\langle v_1, v_2\rangle} + \|v_2\|^2
\end{aligned}$$

Let $\langle v_1, v_2\rangle = \alpha + \beta\sqrt{-1}, \alpha, \beta \in \mathbb{R}$. Then

$$\begin{aligned}
\|v_1 + v_2\|^2 &= \|v_1\|^2 + 2\alpha + \|v_2\|^2 \\
&\leq \|v_1\|^2 + 2\sqrt{\alpha^2 + \beta^2} + \|v_2\|^2 \\
&= \|v_1\|^2 + 2\,|\langle v_1, v_2\rangle| + \|v_2\|^2 \\
&\leq (\|v_1\| + \|v_2\|)^2
\end{aligned}$$

So, $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$. □

§17 Lec 17: Nov 9, 2020

§17.1 Lec 16 (Cont'd)

Example 17.1

Let V be an inner product space over F

1. $|\alpha_1\beta_1 + \ldots + \alpha_n\beta_n| \le \sqrt{\sum_{i=1}^n \alpha_i^2}\sqrt{\sum_{i=1}^n \beta_i^2}, \forall \alpha_i, \beta_i \in \mathbb{R}$.
2. $\int_\alpha^\beta fg \le \sqrt{\int_\alpha^\beta f^2}\sqrt{\int_\alpha^\beta g^2}, \forall f, g \in C[\alpha, \beta]$.
3. $\angle$ between nonzero vectors in V makes sense.
4. Distance between (end pts) vectors makes sense by the following:

   If V is an inner product space over F, define the distance between $v_1, v_2 \in V$ by

   $$d(v_1, v_2) := \|v_1 - v_2\| \ge 0 \in \mathbb{R}$$

   Then d satisfies $\forall v, w, x \in V$

   - $d(v, w) \ge 0 \in \mathbb{R}$ and $d(v, w) = 0$ iff $v = w$.
   - $d(v, w) = d(w, v)$
   - Triangle inequality

     $$d(v, x) \le d(v, w) + d(w, x)$$

     We call V a METRIC SPACE under d.

Example 17.2 (Metric Space)

If $v = (\alpha_1, \ldots, \alpha_n)$, $w = (\beta_1, \ldots, \beta_n) \in \mathbb{R}^n$ under the dot product, then

$$d(v, w) = \sqrt{(\alpha_1 - \beta_1)^2 + \ldots + (\alpha_n - \beta_n)^2}$$

§17.2 Orthogonal Bases

Motivation: in $\mathbb{R}^n$ (or $\mathbb{C}^n$), $\mathscr{S} = \mathscr{S}_n = \{e_1, \ldots, e_n\}$ the standard basis satisfies

$$e_i \cdot e_j = \delta_{ij} := \begin{cases} 1, \text{if } i = j, \forall i, j \\ 0, \text{if } i \ne j \end{cases}$$

Goal: Let V be a finite dimensional inner product space over F, $F = \mathbb{R}$ or $\mathbb{C}$. Find a basis $\mathscr{B} = \{v_1, \ldots, v_n\}$ for $V \ni$

$$\langle v_i, v_j \rangle = \delta_{ij}, \forall i, j \tag{*}$$

if we only want bases $\mathscr{C} = \{w_1, \ldots, w_n\}$ for $V \ni$

$$\langle w_i, w_j \rangle = 0 \forall i \ne j,$$

we can work with any subfield $F \subset \mathbb{C}$ with $F = \overline{F}$, since we do not need $\|w_i\| \in F$ for such a $\mathscr{C}$.

Example 17.3

In $\mathbb{R}^2$, let $0 \leq \theta < 2\pi$ be fixed. Then

$$\mathscr{C}_\theta = \{(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\}$$

satisfies (*)

Definition 17.4 (Orthonormal/Orthogonal) — Let V be an inner product space over F, $\emptyset \neq S \subset V$ a subset. We say

1. S is ORTHOGONAL (or OR) if

$$\langle v, w\rangle = 0 \forall v \neq w \in S$$

2. If S is an OR set, we call it ORTHONORMAL (or ON) if, in addition $\|v\| = 1 \forall v \in S$.
3. An OR set is called an OR basis if, in addition, it is a basis for V.
4. If $v, w \in V$, we say v, w are orthogonal or perpendicular if $\langle v, w\rangle = 0$ write $v \perp w$. (equivalently $\langle w, v\rangle = 0$)

Goal: If $F \subset \mathbb{C}$ is a subfield (and $F = \overline{F}$), V a finite dimensional inner product space over F, then V has an OR bases and an ON bases if $F = \mathbb{R}$ or $\mathbb{C}$.

Remark 17.5. Let V be an inner product space over F, $x, y \in V$.

1. $0 \perp x$
2. $x \perp y$ iff $y \perp x$
3. 0 is the only vector perpendicular to all $z \in V$.

Theorem 17.6

Let V be an inner product space over $F, S \subset V$ an OR set. Suppose that $0 \neq S$, then S is linearly indep. If, in addition, V is a finite dimensional inner product space over F and $|S| = \dim V$, then S is an OR basis for V.

Proof. Let $v \in \text{Span}(S)$. Then $\exists$ (distinct) $v_1, \ldots, v_n \in S, \alpha_1, \ldots, \alpha_n \in F \ni$

$$v = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

We have

$$\begin{aligned}\langle v, v_j\rangle &= \langle \alpha_1 v_1 + \ldots + \alpha_n v_n\rangle \\ &= \sum_{i=1}^n \alpha_i \langle v_i, v_j\rangle \\ &= \sum_{i=1}^n \alpha_i \delta_{ij} \|v_j\|^2 = \alpha_j \|v_j\|^2\end{aligned}$$

This is so useful, we record it as
Crucial Equation: If $\{v_1, \ldots, v_n\}, \alpha_1, \ldots, \alpha_n \in F$ then

$$\alpha_j = \frac{\langle v, v_j \rangle}{\|v_j\|^2}, j = 1, \ldots, n$$

Note: If V is not necessarily finite dimensional and S is an OR set not containing O, the same holds.
Now, suppose that $v = 0$, i.e.,

$$0 = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

so

$$\alpha_j = \frac{\langle v, v_j \rangle}{\|v_i\|^2} = \frac{\langle 0, v_j \rangle}{\|v_i\|^2} = 0, j = 1, \ldots, n$$

and the result follows. □

Note: If $\mathscr{B} = \{v_1, \ldots, v_n\}$ is an OR set, $v_i \neq 0 \forall i, V = \text{Span}\mathscr{B}$, hence a basis for V then

$$\frac{\langle v, v_j \rangle}{\|v_j\|^2}$$

is the jth coordinate of v on v_j and

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

If, in addition, $\|v_j\| \in F \forall j$, then

$$\mathscr{C} = \left\{ \frac{v_1}{\|v_1\|}, \ldots, \frac{v_n}{\|v_n\|} \right\}$$

is an ON basis and $\forall v \in V$.

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

Hence if $w_i = \frac{v_i}{\|v_i\|}$, $i = 1, \ldots, n, \mathscr{C} = \{w_1, \ldots, w_n\}$ is an ON basis and

$$v = \sum_{i=1}^{n} \langle v, w_i \rangle w_i$$

i.e., $\langle v, w_i \rangle$ is the coordinate of v and w_i for each i.

Remark 17.7. Does this look familiar?

1. Look at the proof of the Cauchy – Schwarz Inequality
2. Let $\mathscr{B} = \{v_1, \ldots, v_n\}$ be an OR basis for V a finite dimensional inner product space over F and
$$\mathscr{B}^* = \{f_1, \ldots, f_n\}$$
the dual basis for $V^* = L(V, F)$. So, $f_i(v_j) = \delta_{ij}, \forall i, j$. Then $f_i : V \to F$ is $f_i(v) = \frac{\langle v, v_i \rangle}{\|v_i\|^2}, i = 1, \ldots, n$ by Crucial Equation:
$$f_i = \langle -, \frac{v_i}{\|v_i\|^2} \rangle : V \to F$$
and if $\mathscr{C} = \{w_1, \ldots, w_n\}$ is an ON basis then
$$f_i = \langle , w_i \rangle \in \mathscr{C}^*$$
$$f_i(v) = \langle v, w_i \rangle$$
i.e., we can associate a vector in V to a linear functional.

Theorem 17.8

Let V be an inner product space over F, $\mathscr{B}$ an OR basis for V, $v \in V$. Then $\langle v, w \rangle = 0$ for all but finitely many $w \in \mathscr{B}$ and
$$v = \sum_{\mathscr{B}} \frac{\langle v, w \rangle}{\|w\|^2} w$$
is a finite sum. If, in addition, $\mathscr{B}$ is ON, then this becomes
$$v = \sum_{\mathscr{B}} \langle v, w \rangle w$$

Corollary 17.9 (Parseval's Equation)

Let V be a finite dimensional inner product space over F with ON basis $\{v_1, \ldots, v_n\}$ and $v, w \in V$. Then
$$\langle v, w \rangle = \sum_{i=1}^{n} \langle v, v_i \rangle \overline{\langle w, v_i \rangle}$$
In particular,
$$\|v\|^2 = \sum_{i=1}^{n} |\langle v, v_i \rangle|^2, \qquad \text{(Pythagorean Theorem)}$$

Proof. Hw – Take home. □

§18 Lec 18: Nov 16, 2020

§18.1 Lec 17 (Cont'd)

Example 18.1

Let $V = C[0, 2\pi]$ an inner product space over $\mathbb{R}$ via

$$\langle f, g\rangle := \int_0^{2\pi} fg$$

Let $u_0 = \frac{1}{\sqrt{2\pi}}, u_{2n} = \frac{1}{\sqrt{\pi}} \sin nx, u_{2n+1} = \frac{1}{\sqrt{\pi}} \cos nx$ for all $n \in \mathbb{Z}^+$ and set

$$S = \{u_i | i \geq 0\}$$

By calculus

$$\langle u_i, u_j\rangle = \int_0^{2\pi} u_i u_j = \delta_{ij}, \forall i, j$$

So S is ON hence linearly indep ($0 \notin S$) and a ON basis for Span S.

Note: Vectors in span S are finite linear combos of vectors in S. In particular, $C[0, 2\pi]$ is infinite dimensional (and Span $S < C[0, 2\pi]$ is a subspace). In calculus, you studied convergent series, a convergent series

$$\sum_{i=0}^{\infty} \alpha_i u_i \qquad (*)$$

is called a FOURIER SERIES, the α_i Fourier coefficients.

Warning: $S = \mathscr{B} = \cup \mathscr{B}_n, \mathscr{B}_n = \{u_i | i = 0, \ldots, 2n+1\}$ is ON but not a basis for $C[0, 2\pi]$ or even

$$V = \{f \in C[0, 2\pi] | f \text{ converges to its Fourier series}\}$$

It can be shown that $C'[0, 2\pi] \subset V$.

Note: No one knows a precise basis for $C[0, 2\pi]$ although it exists by axioms.

Remark 18.2.
1. One can modify the interval $[0, 2\pi]$ in the above with appropriate changes to the u_i.
2. Infinite ON sets are very useful.

To solve our goal about finite dimensional inner product space over F, we know show:

Theorem 18.3 (Gram-Schmidt)

Let V be an inner product space over F and $\emptyset \neq S_n = \{v_1, \ldots, v_n\} \subset V$ a linearly indep. set. Then $\exists y_1, \ldots, y_n \in V \ni$

- $y_1 = v_1$
- $T_n = \{y_1, \ldots, y_N\}$ is an OR set and linearly indep.
- Span T_n = Span S_n

Proof. We construct T_n from S_n. This construction is called the Gram – Schmidt process. $n = 1$ is clear. We proceed by induction. We may assume we have done the S_n case, i.e.,

1. $y_1, \ldots, y_n \in V, y_1 = v_1, y_i \neq 0, i = 1, \ldots, n$
2. $T_n = \{y_1, \ldots, y_n\}$ is OR. (hence linearly indep. as $0 \notin T_n$)
3. Span S_n = Span$\{y_1, \ldots, y_n\}$
4. Must extend this to the case of $n + 1$.

As in the proof of GS (where we threw away one orthogonal complement), we subtract an ORTHOGONAL PROJECTION figure here Define:

$$y_{n+1} = v_{n+1} - \sum_{k=1}^{n} \frac{\langle v_{n+1}, y_k \rangle}{\|y_k\|^2} y_k \tag{*}$$

Claim 18.1. $y_{n+1} \neq 0$: if $y_{n+1} = 0$, then $v_{n+1} \in$ Span T_n = Span$(v_1, \ldots, v_n)$ contradicting S?, is linearly indep. So $y_{n+1} \neq 0$

Claim 18.2. $\langle y_{n+1}, y_j \rangle = 0, j = 1, \ldots, n$

$$\begin{aligned}
\langle y_{n+1}, y_j \rangle &= \langle v_{n+1} - \sum_{k=1}^{n} \frac{\langle v_{n+1}, y_k \rangle}{\|y_k\|^2} y_k, y_j \rangle \\
&= \langle v_{n+1}, y_j \rangle - \sum_{k=1}^{n} \frac{\langle v_{n+1}, y_k \rangle}{\|y_k\|^2} \langle y_k, y_j \rangle \\
&= \langle v_{n+1}, y_j \rangle - \sum_{k=1}^{n} \frac{\langle v_{n+1}, y_k \rangle}{\|y_k\|^2} \delta_{kj} \|y_j\|^2 \\
&= \langle v_{n+1}, y_j \rangle - \langle v_{n+1}, y_j \rangle = 0
\end{aligned}$$

This prove the above claim.

Since $0 \notin T_{n+1} = \{y_1, \ldots, y_{n+1}\}$ and T_{n+1} is OR, it is linearly indep. As Span T_n = Span$\{v_1, \ldots, v_n\}$ and $\{v_1, \ldots, v_{n+1}\}$ is linearly indep.

$$\text{Span } T_{n+1} = \text{Span}(v_{n+1}, y_1, \ldots, y_n) = \text{Span}(v_1, \ldots, v_{n+1})$$

by the Replacement Theorem and $(*)$. The theorem follows by induction. □

Theorem 18.4 (Orthogonal)

Let V be a finite dimensional inner product space over F. Then V has an OR basis. If $F = \mathbb{R}$ or $\mathbb{C}$, then V has an ON basis.

Proof. Any basis for V can be converted to an OR basis $\mathscr{C}$ for V by the GS process if V is finite dimensional if $F = \mathbb{R}$ or $\mathbb{C}$, then $\left\{ \frac{v}{\|v\|} | v \in \mathscr{C} \right\}$ is an ON basis for V as $\|v\| \in \mathbb{R} \forall v \in \mathscr{C}$ □

Remark 18.5. Let $V = \mathbb{Q}^2$ a finite dimensional inner product space over $\mathbb{Q}$ with inner product defined by

$$\langle (\alpha_1, \alpha_2), (\beta_1, \beta_2) \rangle_{\frac{1}{3}} := \frac{1}{3}(\alpha_1 \beta_1 + \alpha_2 \beta_2)$$

i.e., WEIGHTED DOT PRODUCT by $\frac{1}{3}$. Then V has an OR basis but not any ON basis

$\| \left(\frac{a_1}{b_1}, \frac{a_2}{b_2} \right) \|_{\frac{1}{3}} \notin \mathbb{Q}$ as $3b_1^2 b_2^2 = a_1^2 b_2^2 + b_1^2 a_2^2$ has no solution in $\mathbb{Z}$.

§18.2 Examples – Computation

Example 18.6 1. $V = \mathbb{R}^3$ under $\langle , \rangle$ = dot product with $v_1 = (1,1,1), v_2 = (1,1,0), v_3 = (1,0,1)$. GS v_1, v_2, v_3 to an OR basis and then to an ON basis:

$$y_1 = (1,1,1)$$
$$y_2 = v_2 - \frac{v_2 \cdot y_1}{\|y_1\|^2} y_1$$

... some boring calculation – can refer online notes/textbook

<u>*Note*</u>:

1. It is easier to guess.
2. If instead of $F = \mathbb{R}$, we had $F = \mathbb{Q}$, we could not get an ON basis after GS-ing.

Example 18.7
$V = \mathbb{R}[x]$ (polynomial function) via

$$\langle f, g \rangle := \int_{-1}^{1} fg$$

$\mathscr{B}_n = \left\{ x^i | 0 \le i \le n \right\}$ is a basis for $\mathbb{R}[x]_n$. GS, $\mathscr{B}_n$ to an OR basis, at least start

$$\begin{aligned} g_0 &= 1 \\ g_1 &= x - \frac{\langle x, 1 \rangle}{\|1\|^2} 1 = x - \frac{\int_{-1}^{1} x}{\int_{-1}^{1} 1} = x \\ g_2 &= x^2 - \frac{\langle x^2, 1 \rangle}{\|1\|^2} 1 - \frac{\langle x^2, x \rangle}{\|x\|^2} x \\ &= x^2 - \frac{\int_{-1}^{1} x^2}{\int_{-1}^{1} 1} - \frac{\int_{-1}^{1} x^3}{\int_{-1}^{1} x^2} x = x^2 - \frac{1}{3} \\ &\vdots \end{aligned}$$

The g_i are called LEGENDRE POLYNOMIALS. You can normalize them, i.e., form $\frac{g_i}{\|g_i\|}$ to get an ON set.

These are important polynomials, g_n satisfies the ODE

$$(1 - x^2) y'' - 2xy' + n(n+1) y = 0$$

These occur in physics, e.g., converting Laplace's Equation $\nabla^2 g = 0$ into spherical coordinates in some cases in quantum mechanics in the solution of Schrodinger's Eqn for the hydrogen atom.
Flow of an (ideal fluid) past a sphere. Determination of the electric fluid due to a charged sphere. Determination of the temperature distribution in a sphere given its surface temperature. Computing $g_n's$ by GS is too difficult. There are many formulas to determine the $g_n's$. Many arise by proving the following recurrence relation:
Rodriguez Representation:

$$g_n = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n$$

Some of these are, using the appropriate ? of the binomial coefficient

$$\binom{n}{m} := \frac{n!}{m!(m-n)!}, 0 \le m \le n:$$

let $M = \frac{n}{2}$ or $\frac{n-1}{2}$ whichever one is an integer, i.e., $\left[\frac{n}{2}\right] =$ greatest integer $\le \frac{n}{2}$.

$$\begin{aligned} g_n &= 2^{\frac{1}{n}} \sum_{m=0}^{M} (-1)^m \frac{(2n-2m)!}{m!(n-m)!(n-2m)!} x^{n-2m} \\ &= 2^n \sum_{k=0}^{n} \binom{n}{k}^2 (x-1)^{n-k}(x+1)^k \\ &= \sum_{k=0}^{n} \binom{n}{k} \binom{-n-1}{k} \left(\frac{1-x}{2}\right)^k \end{aligned}$$

§19 Lec 19: Nov 18, 2020

§19.1 Lec 18(Cont'd)

Note:Gamma function:

$$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$$

where z is complex and Re $(z) > 0$ and $\Gamma(n) = (n-1)!, \forall n > 1,$.

3. GS $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$ in $M_2(\mathbb{R})$ under

$$\langle A, B \rangle = \text{tr } AB^*$$

$$y_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$y_2 = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix} - \frac{\text{tr}\left(\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^*\right)}{\text{tr}\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^*\right)}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$y_2 = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix} - \frac{\text{tr}\left(\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\right)}{\text{tr}\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\right)}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

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

4. $T : \mathbb{R}^3 \to \mathbb{R}^3$ rotation counterclockwise by $\angle\theta$ about a vector $0 \neq v_1$ as axis. Find $T(\alpha, \beta, \gamma)$ i.e., $[T]_{\mathscr{S}}$ complete v_1 to a basis $\{v_1, v_2, v_3\}$ for $\mathbb{R}^3$. GS it to an OR basis, then an ON basis $\mathscr{C}$. Compute $[T]_{\mathscr{C}}$. Then use Change of Basis to compute $[T]_l$ or guess v_2, normalize v_1, v_2 to v_1', v_2' then $v_3 \subset v_1' \times v_2'$.

   Note: If you have a basis with vectors of different lengths, it is hard to compute in this basis. If each vector in your OR basis has the same length r, you can compute.

§19.2 Orthogonal Polynomials

There are many interesting infinite sets of orthogonal polys $\{f_n\}_{n\in\mathbb{Z}^+}$. They often arise as relate α to the HYPERGEOMETRIC ODE

$$z(1-z)\frac{d^2y}{dz^2} + [\gamma - (\alpha+\beta+1)z]\frac{dy}{dz} - \alpha\beta y = 0$$

where z is a complex variable, $y = y(z), \alpha, \beta, \gamma \in \mathbb{C}$. They arise as OR sets or weighted inner product space over $\mathbb{R}$ (or $\mathbb{C}$ on an interval $[a, b]$ (or variant).

$$\int_a^b fgw = \langle f, g \rangle_w$$

where $w > 0$ in $[a, b]$.

- A very general such is the OR set of JACOBI POLYNOMIALS $\left\{P_n^{\alpha,\beta}\right\}$ under the weighted inner product space
$$\langle f, g\rangle_w = \int_{-1}^{1} fgw$$
and
$$w = \frac{(1-x)^\alpha (1+x)^\beta}{\langle \alpha, \beta\rangle - 1}$$
Often such OR sets are not orthonormalized but rather normalized "by dividing by $P_n^{\alpha,\beta}$ (1). In this case, $P_n^{\alpha,\beta}(1) = \binom{n+\alpha}{n}$. The $P_n^{\alpha,\beta}$ are solutions to the ODE.
$$0 = (1-x^2)y'' + (\beta - \alpha - (\alpha+\beta+2)x)y' + n(n+\alpha+\beta-1)y$$
used in Wigner d-matrix theory in quantum mechanics. There are many special cases of Jacobi polys.
  1. Gegenbauer polys (ultra-symmetric) polynomials, $C_n^{(\alpha)}$ where
  $$w = (1-x^2)^{\alpha - \frac{1}{2}}$$
  $$C_n^{(\alpha)} = P_n^{(\alpha-\frac{1}{2}, \alpha-\frac{1}{2})}$$
  $$(1-x^2)y'' - (2\alpha+1)xy' + n(n+2\alpha)y = 0$$
  potential theory, harmonics analysis, Newtonian's potential.
  2. Legendre polys. There are a special case of Gegenbauer polys, namely
  $$w = 1$$
  $$C_n^{\frac{1}{2}}$$
  $$\left((1-x^2)y'\right)' + n(n+1)y = 0$$
  3. Chebychev polys come in two kinds: T_n, U_n
  $$w = \frac{1}{\sqrt{1-x^2}}$$
  $$T_n = P_n^{(-\frac{1}{2}, -\frac{1}{2})}$$
  $$U_n = P_n^{(\frac{1}{2}, \frac{1}{2})}$$
  $$(1-x^2)y'' - xy' + n^2 y = 0$$
  $$(1-x^2)y'' - 3xy' + n(n+2)y = 0$$
  Least square fit, optimal control, numerical analysis.

- Laguerre polys $L_n^{(\alpha)}$ OR set with $w_\alpha(x) = x^\alpha e^{-x}, \alpha > -1$ in $\mathbb{R}$ on $[0, \infty)$
$$xy'' + (\alpha + 1 - x)y' + ny = 0, 0 \neq n \in \mathbb{Z}$$
quantum mechanics, plasma physics.

- HERMITE polys. H_n, He_n
$$w = e^{-x^2}, \text{ for } H_n \text{ on } (-\infty, \infty)$$
$$= e^{-\frac{x^2}{2}}, \text{ for } He_n \text{ on } (-\infty, \infty)$$
(H_n is called physicist Hermite polys and He_n probabilists Hermite polys).
$$0 = (e^{-\frac{1}{2}x^2} y')' + ne^{-\frac{1}{2}x^2} y = 0$$
probability, numerical analysis, physics.

Remark 19.1. Let

$$D = \text{diff} = \frac{d}{dx}, \quad p, q \text{ functions}, w > 0$$
$$L = -\frac{1}{w}\left(D(pD) + q\right), \quad \text{a linear operator}$$

Then one wants to solve

$$Lf = \lambda f$$

The solutions are called eigenfunctions in the above they are the eigenfunctions for the given ODEs.

§19.3 Orthogonal Complement

Notation: $F \subset \mathbb{C}$ a field satisfying $F = \overline{F}$.

Definition 19.2 (Distance from a Vector to a Set) — Let V be an inner product space over $F, v_1, v_2 \in V$. We know that the DISTANCE between v_1, v_2 is defined to be

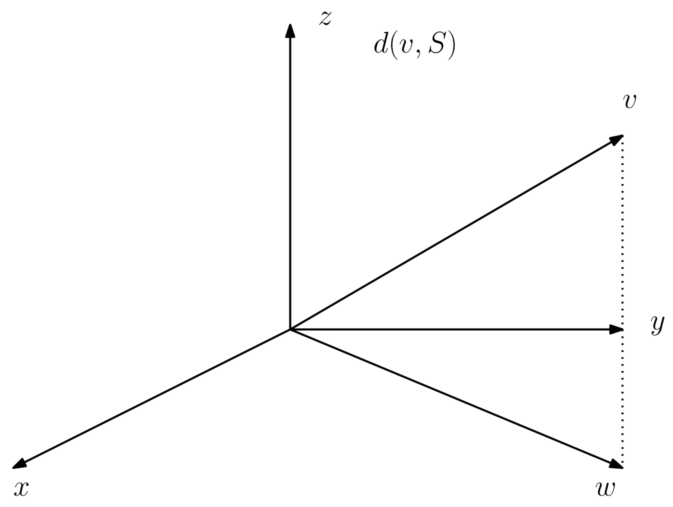
$$d(v_1, v_2) := \|v_1 - v_2\| \geq 0$$

More generally, let $\emptyset \neq S \subset V$ be a subset and $v \in V$. Define the DISTANCE of v to S by

$$d(v, S) := \inf\left\{d(v, w) | w \in S\right\}$$

if it exists and hence finite.

Problem 19.1. Let V be an inner product space over F, $S \subset V$ a finite dimensional subspaces, $v \in V$. Determine



Solution take the orthogonal projection of v to w in S

Definition 19.3 (Orthogonal Complement) — Let V be an inner product space over $F, \emptyset \neq S \subset V$ a subset of, $v \in V$. We say v is ORTHOGONAL to S, write $v \perp S$, if

$$\langle s, v \rangle = 0, \forall s \in S$$

Set:

$$S^\perp := \{v \in V | v \perp S\}$$

called the ORTHOGONAL COMPLEMENT of S in V.

Remark 19.4.

1. Compare $S^\perp$ to $S^\circ \subset V^*$, if V is an arbitrary vector space over F.
2. In $\mathbb{R}^3$(under the dot product)

$$(\text{Span} e_1)^\perp = \text{Span}(e_2, e_3)$$

3. Let V be an inner product space over $F, \emptyset \neq S \subset V$ a subset, not necessarily a subspace. Then $S^\perp \subset V$ is a subspace (if $\emptyset \neq S \subset V$ a subset with V a vector space over F, F arbitrary, then $S^\circ \subset V^*$ is a subspace).

   Proof. Hw. □

4. In 3), $S \subset S^{\perp\perp} := \left(S^\perp\right) \perp$: $S^\perp \subset S^{\perp\perp}$ so $S \subset S^{\perp\perp}$. If, in addition, $S \subset V$ is a subspace and V is a finite dimensional inner product space over F, then $S = S^{\perp\perp}$ (if V is a finite dimensional vector space over F, F arbitrary $W \subset V$ a subspace, then $W = W^{\circ\circ} = (W^\circ)^\circ$).
5. Let V be a finite dimensional inner product space over $F, S = \{v_1, \ldots, v_n\}$ an OR basis for V. Then

$$(\text{Span}(v_1, \ldots, v_r))^\perp = \text{Span}(v_{r+1}, \ldots, v_n)$$

6. Let V be an inner product space over $F, S \subset V$ a subspace. Then

$$S \cap S^\perp = 0$$

   if $v \in S \cap S^\perp$, then $\langle v, v \rangle = \|v\|^2 = 0$, so $v = 0$. In particular,

$$S + S^\perp = S \oplus S^\perp$$

   We write: $S \oplus S^\perp$ as $S \perp S^\perp$ to show it is also orthogonal. The key result (and most important result for use about general inner product space over F) is:

Theorem 19.5 (Orthogonal Decomposition)
Let V be an inner product space over $F, S \subset V$ a finite dimensional subspace, $v \in V$. Then

$$\exists! s \in S, s^\perp \in S^\perp \ni v = s + s^\perp \tag{*}$$

In particular, $V = S + S^\perp, S \cap S^\perp = 0$, so $V = S \perp S^\perp$. Moreover, if

$$v = s + s^\perp, s \in S, s^\perp \in S^\perp$$

then

$$\|v\|^2 = \|s\|^2 + \|s^\perp\|^2, \qquad \text{(Pythagorean Theorem)}$$

In addition, if V is a finite dimensional inner product space over F, then

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

§20 Lec 20: Nov 20, 2020

§20.1 Lec 19 (Cont'd)

Proof. By the OR Theorem, $\exists$ an OR basis $\mathscr{B} = \{v_1, \dots, v_n\}$ for the finite dimensional inner product space over F S.
Existence: Let $v \in V$. Define $s \in S = \operatorname{Span} \mathscr{B}$ by

$$s = \sum_{i=1}^{n} \frac{\langle v, v_i \rangle}{\|v_i\|^2} v_i$$

and set

$$s^\perp = v - s$$

Suppose we have shown $s^\perp \in S^\perp$. Then $v = s + s^\perp$ giving existence as well as $V = S + S^\perp$ and $S \cap S^\perp = 0$, i.e., $V = S \oplus S^\perp$. Repeating the previous computation, we have if $j = 1, \dots, n$ then

$$\begin{aligned}
\langle s^\perp, v_j \rangle = \langle v - s, v_j \rangle &= \langle v, v_j \rangle - \langle s, v_j \rangle \\
&= \langle v, v_j \rangle - \sum_{i=1}^{n} \frac{\langle v, v_i \rangle}{\|v_i\|^2} \langle v_i, v_j \rangle \\
&= \langle v, v_j \rangle - \sum_{i=1}^{n} \frac{\langle v, v_i \rangle}{\|v_i\|^2} \delta_{ij} \|v_j\|^2 = 0
\end{aligned}$$

Since $s^\perp \perp v_j$, $j = 1, \dots, n$ i.e., $\forall v_j \in \mathscr{B}$, if $\sum_{i=1}^{n} \alpha_i v_i \in S$, then

$$\langle s^\perp, \sum_{i=1}^{n} \alpha_i v_i \rangle = \sum_{i=1}^{n} \overline{\alpha_i} \langle s^\perp, v_i \rangle = 0$$

Thus, $s^\perp \in S^\perp$ as needed.
Uniqueness: If

$$s + s^\perp = v = r + r^\perp, r \in S, r^\perp \in S^\perp$$

$(s \in S, s^\perp \in S^\perp)$ as both $S, S^\perp$ are subspaces

$$s - r = r^\perp - s^\perp \in S \cap S^\perp = 0$$

So $s = r$ and $s^\perp = r^\perp$. $\square$

Theorem 20.1 (Pythagorean)
Let $v = s + s^\perp, s \in S, s^\perp \in S^\perp$. Then

$$\begin{aligned}
\|v\|^2 &= \langle s + s^\perp, s + s^\perp \rangle = \langle s, s \rangle + \langle s, s^\perp \rangle + \langle s^\perp, s \rangle + \langle s^\perp, s^\perp \rangle \\
&= \|s\|^2 + \|s^\perp\|^2
\end{aligned}$$

Corollary 20.2 (Bessel's Inequality)

Let V be an inner product space over $F, \mathscr{B} = \{v_1, \ldots, v_n\}$ an OR set in V with $0 \notin \mathscr{B}$. Let $v \in V$. Then

$$\sum_{i=1}^{n} \frac{|\langle v, v_j \rangle|^2}{\|v_i\|^2} \le \|v\|^2$$

with equality iff

$$v = \sum_{i=1}^{n} \frac{\langle v, v_j \rangle}{\|v_i\|^2} v_i$$

Proof. Hw. □

Remark 20.3. Let V be an inner product space over F, $S \subset V$ a finite subspace. Then by the OR Decomposition Theorem, $\forall v \in V \exists ! s \in S, s^\perp \in S^\perp \implies v = s + s^\perp$. We call s the orthogonal projection of v on S and denote it by v_S. By the proof of the OR Decomposition Theorem, if $\mathscr{B} = \{v_1, \ldots, v_n\}$ is ANY OR basis for S, then the uniqueness of v_S means

$$v_S = \sum_{i=1}^{n} \frac{\langle v, v_i \rangle}{\|v_i\|^2} v_i$$

i.e.,is INDEPENDENT of OR basis. So the ORTHOGONAL PROJECTION of v onto S.

Theorem 20.4 (Approximation)

Let V be an inner product space over $F, S \subset V$ a finite dimensional subspace, and $v \in V$. Then v_S is closer to v than any other vector in S, i.e.,

$$d(v, v_S) = \|v - v_S\| \le \|v - r\| = d(v, r)$$

in $\mathbb{R}, \forall r \in S$. Equivalently,

$$d(v, S) = d(v, v_S)$$

Moreover, if $r \in S$, then

$$\|v - v_S\| = \|v - r\| \in \mathbb{R} \iff r = v_S$$

We say v_S gives the BEST APPROXIMATION.

Proof. By the OR Decomposition Theorem (and its proof), $v = s + s^\perp$ with $s = v_S, s^\perp = v - s = v - v_S, s^\perp \in S^\perp$. Let $r \in S$. Then

$$v - r = (v - v_S) + (v_S - r) = s^\perp + (v_S - r)$$

$S \subset V$ is a subspace, so $v_S - r \in S$, hence $s^\perp \perp v_S - r$, i.e.,

$$0 = \langle s^\perp, v_S - r \rangle = \langle v - v_S, v_S - r \rangle$$

By the Pythagorean Theorem,

$$\|v - r\|^2 = \|v - v_S\|^2 + \|v_s - r\|^2 \ge \|v - v_S\|^2$$

with equality iff

$$\|v_S - r\| = 0 \iff v_s = r$$

□

Definition 20.5 (Error) — Let V be an inner product space over $F, S \subset V$ a finite dimensional subspace and $v \in S$. Then, $\|v - v_S\|$ is called the error of v not being v_S.

Problem 20.1. Let V, X be inner product space over F, $S \subset V$ a finite dimensional subspace $v \in V$, and $T : X \to V$ linear. Find $x \in X$ with $\|x\|$ minimal s.t. Tx is the best approximation to $v \in V$ in S, i.e., find $x \in X$, $\|x\|$ minimal $\ni Tx = v_S$.

§20.2 Examples of Best Approximation

Example 20.6 (Fourier Coefficient)
Let $V = C[0,\pi]$ an inner product space over $\mathbb{R}$ via $\langle f, g\rangle = \int_0^{2\pi} fg$, $u_0 = \frac{1}{\sqrt{2\pi}}$, $u_{2n-1} = \frac{\cos nx}{\sqrt{\pi}}$, $u_{2n} = \frac{\sin nx}{\sqrt{\pi}}$, $n > 0$. Set

$$S = \{u_0, \dots, u_n, \dots\}$$

an ON set (as we have seen) and let

$$\mathscr{B}_n := \{u_0, \dots, u_{2n+1}\}$$
$$V_n := \text{Span}(\mathscr{B}_n)$$

if $f \in V$, then

$$f_n := f_{v_n} = f_{\text{span } \mathscr{B}_n},$$

the function in V_n closest to f, i.e., the orthogonal projection of f onto V_n. So

$$f_n = \sum_{i=0}^{2n+1} \langle f, u_i\rangle u_i$$

where

$$\langle f, u_i\rangle = \int_0^{2\pi} f u_i, \quad \forall i \le 2n$$

called the i^{th} FOURIER COEFFICIENT. The ERROR to the actual f is

$$d(f, f_n) = \|f - f_n\| = \sqrt{\int_0^{2\pi} (f - f_n)^2}$$

One checks:

$$f_n = \frac{1}{2} 0_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

with

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) dx$$
$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx$$
$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx$$

is the BEST APPROXIMATION of f by such functions. If $\lim_{n\to\infty} \|f - f_n\| = 0$, i.e., $f = \sum_{i=0}^{\infty} \langle f, u_i\rangle u_i$ converges, we say f converges to its Fourier expansion (similar results with modest change work for ($[0, L]$).

Example 20.7

Let $V = C[-1,1]$ with $\langle f, g\rangle = \int_{-1}^{1} fg$. Let $f(x) = e^x$. Find a linear polynomial nearest f and compute $d(f,g)$ (=error) for such a g and we let $W = \text{span}(1,x) \subset V$ a finite dimensional subspace. We want f_W. To do this, we compute ON (or OR) basis for W i.e., GS $\{1,x\}$ and normalize. GS yields $1, x$ (as before) and ON it to $\frac{1}{\|1\|}, \frac{x}{\|x\|}$, i.e., $\frac{1}{\sqrt{\int_{-1}^{1} 1}}, \frac{x}{\sqrt{\int_{-1}^{1} x^2}}$ which is

$$\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x$$

Let $f = e^x$. Then

$$\begin{aligned} f_W &= \langle f, \frac{1}{\sqrt{2}}\rangle \frac{1}{\sqrt{2}} + \langle f, \frac{\sqrt{3}}{2}x\rangle \frac{\sqrt{3}}{2}x \\ &= \frac{1}{2}\int_{-1}^{1} e^z dz + \frac{3}{2}x\int_{-1}^{1} ze^z dz \\ &= \dots \\ &= \frac{1}{2}(e - \frac{1}{e}) + \frac{3}{e}x \end{aligned}$$

So, $f_W = \frac{1}{2}(e - \frac{1}{e}) + \frac{3}{e}x$. Let $\alpha = \frac{1}{2}(e-\frac{1}{e}), \beta = \frac{3}{e}x$. So $g = f_W = \alpha + \beta x$ and

$$\begin{aligned} \|f - f_W\|^2 = \|f-g\|^2 &= \int_{-1}^{1} (f-g)^2\, dz \\ &= \int_{-1}^{1} (f^2 - 2fg + g^2)\, dz \\ &= \int_{-1}^{1} \left[(e^{2x} - 2e^x(\alpha + \beta x) + \alpha^2 + 2\alpha\beta x + \beta^2 x^2\right]\, dx \\ &= \dots \text{(boring algebra)} \\ &= 1 - \frac{7}{e^2} \end{aligned}$$

So

$$d(f,g) = d(f, f_W) = \sqrt{1 - \frac{7}{e^2}} \approx .05625$$

§20.3 Hermitian Operators

Definition 20.8 (Hermitian/Self-Adjoint) — Let V be an inner product space over F, $T: V \to V$ linear. We say T is HERMITIAN or SELF-ADJOINT if

$$\langle Tv, w\rangle = \langle v, Tw\rangle, \forall v, w \in V$$

if $F \subset \mathbb{R}$ is an hermitian operator, it is also called a SYMMETRIC OPERATOR.

Example 20.9 1. Let $V = F^{n\times 1}$ be an inner product space over F via the dot product, i.e.,

$$\langle \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \rangle \coloneqq \sum_{i=1}^{n} \alpha_i \overline{\beta}_i$$

remember we always assume $F = \overline{F} \subset \mathbb{C}$. Note that some people write the dot product $v * w$ – they do not like columns.

Let $A \in M_n(F)$. As usual, we view A as a linear operator,

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

By HW, A is hermitian iff $A = A^*$ (so if $F \subset \mathbb{R} \iff A = A^t$). In fact, you will prove on the takehome the following theorem

Theorem 20.10

Let V, W be finite dimensional inner product space over F with ON bases, $T : V \to W$ linear. Then, $\exists! T^* : W \to V$ linear s.t.

$$\langle Tv, w\rangle_W = \langle v, T^* w\rangle_V, \forall v \in V, \forall w \in W$$

T^* is called the ADJOINT of T. Hence if $T : V \to V$ is a linear operator, then T is hermitian iff $T = T^*$ and T^* exists.

Example 20.11

Let $\alpha < \beta$ in $\mathbb{R}$ and $V = C[\alpha, \beta] \coloneqq \{f : [\alpha, \beta] \to \mathbb{R}/\text{cont}\}$ an inner product space over $\mathbb{R}$ by

$$\langle f, g\rangle \coloneqq \int_\alpha^\beta fg$$

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

$$\int_\alpha^\beta (fTg - gTf) = 0, \forall f, g \in V \qquad (*)$$

Note: V is not finite dimensional and (*) is a commutativity type of condition.

Example 20.12 (fancy)

$V = C^\infty[\alpha, \beta], \alpha < \beta$ in $\mathbb{R}$. (often $C^\infty[\alpha, \beta]$ vector space of convergent power series in some neighborhood of every point of (α, β) and ? open neighborhood at α, β). Again V is not finite dimensional and is an inner product space over $\mathbb{R}$ as in the above example.
Let $p \in V$ be fixed, $p(x) > 0$, and

$$W = \{f \in V | p(\alpha)f(\alpha) = 0 = p(\beta)f(\beta)\}$$

an inner product space as in the above example (e.g., $p(\alpha) = 0p(\beta)$. Fix $q \in W$ and let

$$T_{p,q} = T : W \to W \text{ the linear operator}$$

defined by

$$Tf \coloneqq (pf')' + qf$$

called a STURM LIOUVILLE operator. Then T is hermitian. Check T satisfies (*) in the above example using integration by parts.

Example 20.13

More generally, let $V = C^\infty[\alpha, \beta], \alpha < \beta \in \mathbb{R}$ an inner product space over $\mathbb{R}$ as in the above. Let $p, q, w \in V, p(x) > 0, w(x) > 0, \forall x \in [\alpha, \beta]$. Fix $a, b, c, d \in \mathbb{R} \ni$ both $a = 0 = b$ and $c = 0 = d$ are excluded. Let

$$w = \left\{f \in V | af(\alpha) + bf'(\alpha) = 0 = cf(\beta) + df'(\beta)\right\}$$

where f satisfies the boundary condition. Let W be an inner product space over $\mathbb{R}$ by the weighted inner product

$$\langle f, g \rangle_w = \int_\alpha^\beta wfg$$

Define the STURM LIOUVILLE OPERATOR:

$$T = T_{p,q,w} : W \to W \text{ by}$$

$f \mapsto -\frac{1}{w}\left((pf')' + qf\right)$. Then T is hermitian. This arises from finding eigenvalues of $T_{p,q,w}$, i.e., solutions to the ODE

$$\frac{d}{dx}\left(p\frac{dy}{dx}\right) + q(x)y = -\lambda wy$$

which have as special cases – Legendre ODE

$$(1 - x^2)y'' + 2xy' + n(n + 1) = 0$$

arising in spherical harmonic problems. Bessel's ODE:

$$x^2y'' + xy' + (x^2 - a^2)y = 0$$

$\alpha \in \mathbb{C}$ (often in $\mathbb{Z}$ or $2\alpha \in \mathbb{Z}$), i.e., one wants to find the eigenvalues of $f = y$, λ in (*) for which there is a solution and $f \in E_T(\lambda)$. Eigenvectors in function spaces are called EIGENFUNCTIONS.

§21 Lec 21: Nov 23, 2020

§21.1 Lec 20 (Cont'd)

Goal: Spectral Theorem for Hermitian Operator: 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 diagonalizable, i.e., $\exists$ a basis $\mathscr{B}$ for V consisting of eigenvectors of T, and in fact, such a $\mathscr{B}$ is ON.
Calculus Application: Let $S \subset \mathbb{R}^n$ be "nice" (open + nice boundary + ...), $x_1, \ldots, x_n$ the rectilinear coordinate functions relative to the standard basis and

$$(+) f : S \to \mathbb{R} \text{ a } C^2 - \text{ a function}$$

Calculus Theorem if f satisfies (+), then

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a), \forall^i_j, \forall a \in S$$

For each $a \in S$, associate the symmetric matrix

$$Hf(a) :=:= \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)$$

called the HESSIAN at f at a. Suppose $a \in S$ is a critical point of f, i.e.,

$$Df(a) := \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right) = (0, \ldots, 0)$$

Equivalently, $\nabla f(a) = 0$. Recall the TOTAL DERIVATIVE of f at a is the linear transformation

$$f'(a,) : \mathbb{R}^n \to \mathbb{R} \text{ given by}$$

$f'(a, v) = Df(a) \cdot v$. Now, let $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ be the eigenvalues of $Hf(a)$, so the roots of $f_{Hf(a)}$ counted with multiplicity. Since $Hf(a)$ is symmetric, by the Spectral Theorem, $m = n$ and

$$Hf(a) \sim \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \text{ in } M_n\mathbb{R}$$

$\lambda_1, \ldots, \lambda_n$ not necessarily distinct. Then, we have the 2nd Derivative Test under the above conditions at the critical point a.

1. a is a relative minimum for f at a if $\lambda_i > 0 \forall i$.
2. a is a relative maximum for f at a if $\lambda_i < 0 \forall i$.
3. a is a saddle point for f at a if $\exists i, j \ni \lambda_i > 0, \lambda_j < 0$.
4. No info if $\lambda_i = 0 \forall i$ or $\exists i \ni \lambda_i = 0$.

The total derivative $f'(a, -) : \mathbb{R}^n \to \mathbb{R}$ can be defined at $a \in S$ if it exists as the following: it is a linear transformation

$$Ta : \mathbb{R}^n \to \mathbb{R} \ni$$

$\exists$ a scalar valued function satisfying

$$f(a + v) = f(a) + \|v\| E(a, v)$$

for some $r, \ni$ if $\|v\| < r$ then

$$E(a, v) \to 0 \text{ as } \|v\| \to 0$$

Question 21.1. What is the total derivative

$$f'(a,\cdot):\mathbb{R}^n\to\mathbb{R}^m \text{ if } f:S\to\mathbb{R}^m?$$

Theorem 21.1

Let V be an inner product space over $F, T: V\to V$ linear, λ an eigenvalue of $T, 0\neq v\in E_T(\lambda)$. Then

$$\lambda=\frac{\langle Tv,v\rangle}{\|v\|^2} \text{ and } \overline{\lambda}=\frac{\langle v,Tv\rangle}{\|v\|^2}$$

In particular, $\lambda\in\mathbb{R}$ iff

$$\langle Tv,v\rangle=\langle v,Tv\rangle$$

Proof. By assumption, $Tv=\lambda v$, $\|v\|\neq 0$.So $\langle Tv,v\rangle=\langle\lambda v,v\rangle=\lambda\langle v,v\rangle=\lambda\|v\|^2$ and $\langle v,Tv\rangle=\langle v,\lambda v\rangle=\overline{\lambda}\langle v,v\rangle=\overline{\lambda}=\|v\|^2$. As $\|v\|\neq 0$, the first statement follows. Hence,

$$\lambda=\overline{\lambda}\iff\langle Tv,v\rangle=\langle v,Tv\rangle$$

□

Corollary 21.2 (Hermitian)

Let V be an inner product space over $F, T: V\to V$ linear. Suppose that T is hermitian. Then any eigenvalues of T is real, i.e., lies in $F\cap\mathbb{R}$.

Theorem 21.3 (Fundatemental Theorem of Algebra)

Let $f\in\mathbb{C}[t]\setminus\mathbb{C}$. Then f has a root in $\mathbb{C}$, i.e., $\exists\alpha\in\mathbb{C}\ni f(\alpha)=0$

Addendum: Let $f\in\mathbb{R}[t]\setminus\mathbb{R}$. As $\mathbb{R}\subset\mathbb{C},\mathbb{R}[t]\subset\mathbb{C}[t]$. So we can view $f\in\mathbb{C}[t]$. Then f has a root $\beta\in\mathbb{C}$. Of course, β may not lie in $\mathbb{R}$.
Suppose β is real, i.e., $\beta\in\mathbb{R}$. As β is a root of $f\in\mathbb{C}$

$$f=(t-\beta)g, g\in\mathbb{C}[t],\beta\in\mathbb{R}$$

Then

$$f=(t-\beta)(h), h\in\mathbb{R}[t](\text{if } \beta\in\mathbb{R})$$

Proof.

1. If $f=\sum_{i=0}^n\alpha_i t^i,\alpha_i\in\mathbb{R}\forall i$ and $\sum_{i=1}^n\alpha_i\beta^i=0$ in $\mathbb{C}$ with $\beta\in\mathbb{R}$, then every term in $\sum\alpha_i\beta^i$ lies in $\mathbb{R}$, so β is a root of f when viewed in $\mathbb{R}[t]$.
2. (Generalization) Let $F\subset K$, K a field, F a subfield of K so same $+,\cdot,0,1$ as in K (e.g., $\mathbb{R}\subset\mathbb{C}$). Let $f\in F[t],\alpha\in F$. By the DIVISION ALGORITHM,

$$f=f(t-\alpha)g+r,\quad r,g\in F[t] \text{ unique with r}=0 \text{ or deg r}<\deg(t-\alpha)\qquad (*)$$

But $\deg(t-\alpha)=1$, so $r\in F$ (a constant). Evaluate (*) at $t=\alpha$, so ($e_\alpha: F[t]\to F$ by $h\mapsto h(\alpha)$ a ring homomorphism)

$$f(\alpha)=(\alpha-\alpha)g(\alpha)+r=r$$

i.e.,

$$(+) f = (t-\alpha)g + f(\alpha)$$

So

$$\alpha \in F \text{ is a root in } F \iff$$

$(\star) f = (t-\alpha)g$ in $F[t]$ some $g \in F[t]$. So we have, viewing $F[t] \subset K[t]$. If $\beta \in K$, then

$$f = (t-\beta)h + f(\beta), h \in K[t]$$

and if $\beta \in K$ is a root of f in K, then

$$f = (t-\beta)h \in K[t]$$

So if $\beta \in K$ is a root of f with $\beta \in F$, then

$$f(\beta) = 0_K = 0_F,$$

so $(\star)$ holds.

□

Remark 21.4.

1. By the Addendum and induction, FTA says if $f \in \mathbb{C}[t] \setminus \mathbb{C}$, says $n = \deg f \geq 1$, then $\exists! \alpha_1, \ldots, \alpha_n \in \mathbb{C}$, not necessarily distinct and $\beta \in \mathbb{C} \ni$

$$f = \beta(t-\alpha_1)\ldots(t-\alpha_n)$$

i.e., f factors into a product of linear polys. We say f splits in $\mathbb{C}$ and $\alpha_1, \ldots, \alpha_n$ are the unique roots (up to multiplicity) of f in $\mathbb{C}$.

2. FTA is proven in Math 132 and math 110C. The essential analysis fact used in math 132 is if $f \in \mathbb{C}[t] \setminus \mathbb{C}$, then $|f(z)| \to \infty$ as $|z| \to \infty$ and the essential analysis fact used in math 110C is the Intermediate Value Theorem in the special case that says if $f \in \mathbb{R}[t]$ is of odd degree, then f has a real root.

3. The following fact is true: If V is a finite dimensional vector space over F, F an arbitrary field, $T: V \to V$ linear, then $\exists$ an ordered basis $\mathscr{B}$ for $V \ni [T]_{\mathscr{B}}$ is UPPER TRIANGULAR (i.e. $([T]_{\mathscr{B}})_{ij} = 0 \forall i > 1$) iff $f_T \in F[t]$ splits, i.e., factors into a product of linear terms. If this occurs, we say T is TRIANGULARIZABLE. Can you prove that if $F = \mathbb{C}$, then every such T is triangularizable? (T is diagonalizable iff q_T of the HW7/Midterm splits and has no multiple roots)

§22 Lec 22: Nov 25, 2020

§22.1 Lec 21 (Cont'd)

Definition 22.1 (T-invariant) — Let F be an arbitrary field, V a vector space over $F, W \subset V$ a subspace, $T : V \to V$ linear. We say W is T-INVARIANT (or INVARIANT under T) if

$$Tw \in W, \forall w \in W, \text{ i.e., } T(W) \subset W$$

if W is T-invariant, then we can (and do) view

$$T\Big|_W : W \to W \text{ linear}$$

Example 22.2

1. Any subspace of an eigenspace of T (if any) is T-invariant.
2. $\ker T \subset V$ is T-invariant.
3. $\operatorname{im} T \subset V$ is T-invariant.

Lemma 22.3 (Hermitian Operator (Key Lemma))
Let V be an inner product space over $F, T : V \to V$ hermitian, $S \subset V$ a T-invariant subspaces. Then

1. $S^\perp$ is T-invariant, i.e., $T(S^\perp) \subset S^\perp$.
2. $T\Big|_{S^\perp} : S^\perp \to S^\perp$ is hermitian.

Proof.

1. Let $w \in S^\perp$. To show $Tw \in S^\perp$, if $v \in S$, then $Tv \in S$ as S is T-invariant. So
$$\langle v, Tw \rangle = \langle Tv, w \rangle = 0$$
So, $Tw \in S^\perp$.
2. By 1), $T\Big|_{S^\perp} : S^\perp \to S^\perp$ is linear. As $\langle Tv, w \rangle = \langle v, Tw \rangle, \forall v, w \in V$, this is certainly true $\forall v, w \in S^\perp$.

□

Remark 22.4. Let $F = \mathbb{R}$ or $\mathbb{C}$, V a finite dimensional inner product space over $F, T : V \to V$ hermitian. By the Hermitian Corollary, if T has an eigenvalue, it is real and $\alpha \in F$ is a roof of f_T in F iff eigenvalue of T. We know f_T has a root in $\mathbb{C}[t]$ by the FTA. The key lemma should allow us to induct on $\dim V$.

Subtle Difficulty: Let V be a finite dimensional inner product space over $\mathbb{R}, T : V \to V$ hermitian. We know $f_T \in \mathbb{R}[t]$ has a root in $\mathbb{C}$, but we do not know a priori that f_T is the characteristics polynomial of an hermitian operator over an inner product space over $\mathbb{C}$, so we do not know that the roots of f_T are real.
Unfortunately, to over come this, we have use bases. There is an abstract way to do it but we cannot do it.

Theorem 22.5 (Spectral – First Version)
(for Hermitian Operator) Let $F = \mathbb{R}$ or $\mathbb{C}$, V a finite dimensional inner product space over $F, T : V \to V$ hermitian. Then $\exists$ an ON basis $\mathscr{B} = \{v_1, \dots, v_n\}$ for V with each $v_i, i = 1, \dots, n$, an eigenvector for some eigenvalues $\alpha_i \in \mathbb{R}, i = 1, \dots, n$ (not necessarily distinct). In particular, T is diagonalizable.

Proof. We prove $\mathscr{B}$ exists by induction on $\dim V = n$.
$n = 1 :$ $V = \text{Span}(v)$, any $0 \neq u \in V$. As $Tv \in \text{Span}(v), \exists \alpha \in F \ni Tv = \alpha v$,so $v \in E_T(\alpha)$. As T is hermitian, $\alpha \in \mathbb{R}$ is real by Hermitian Corollary even if $F = \mathbb{C}$. So $\mathscr{B} = \left\{ \frac{v}{\|v\|} \right\}$.
$n > 1 :$ Induction Hypothesis (IH): Let $F = \mathbb{R}$ or $\mathbb{C}$, W a finite dimensional inner product space over $F, \dim W = n-1, T_0 : W \to W$ hermitian. Then $\exists$ an ON basis for W of eigenvectors of T_0 and every eigenvalues of T_0 is real.
Let $\mathscr{C}$ be an ON basis for $n-$dimensional V, which exists as $F = \mathbb{R}$ or $\mathbb{C}$. Let $A = [T]_{\mathscr{C}} \in M_nF \subset M_n\mathbb{C}$.

$$A = A^* \text{ and } Ax \cdot y = x \cdot Ay, \forall x, y \in C^{n \times 1}$$

since T is hermitian, i.e.,

$$A : C^{n\times 1} \to C^{n \times 1} \text{ is hermitian}$$

where $C^{n\times 1}$ is an inner product space over $\mathbb{C}$ via the dot product. By the FTA, f_A has a root $\alpha \in \mathbb{C}$, hence α is an eigenvalue of hermitian $A : \mathbb{C}^{n\times 1} \to C^{n\times 1}$. Thus, $\alpha \in \mathbb{R}$ by the Hermitian Corollary. But

$$f_T = f_{[T]_{\mathscr{C}}} = f_A$$

So f_T has a root $\alpha \in \mathbb{R}$, if $F = \mathbb{R}$ or $F = \mathbb{C}$ by the Addendum. Thus, $\exists 0 \neq u \in E_T(\lambda) \subset V$ an eigenvector of T. Let $Fv = \text{Span}(v) \subset E_T(\lambda)$. Then Fv is T-invariant. By the OR Decomposition Theorem,

$$V = Fv \perp (Fv)^\perp$$

and

$$\dim V = \dim Fv + \dim (Fv)^\perp = 1 + \dim (Fv)^\perp$$

hence

$$\dim (Fv)^\perp = n - 1$$

By the Key Lemma, since Fv is T-invariant and $T : V \to V$ is hermitian. $(Fv)^\perp$ is T-invariant and

$$T\Big|_{(Fv)^\perp} : (Fv)^\perp \to (Fv)^\perp \text{ is hermitian}$$

By the IH, $(Fv)^\perp$ has an ON basis, say $\{v_2, \dots, v_n\}$ of eigenvectors for $T\Big|_{(Fv)^\perp} : (Fv)^\perp \to (Fv)^\perp$. But

$$T\Big|_{(Fv)^\perp}(v_i) = Tv_i, i = 2, \dots, n$$

So, $v_2, \dots, v_n$ are eigenvectors of $T : V \to V$ and all the eigenvalues of the $v_i, i = 2, \dots, n$ are real by IH. Since $v \perp v_i, i = 2, \dots, n, 0 \neq \|v\| \in \mathbb{R} \subset F$,

$$\mathscr{B} = \{\|v\|, v_2, \dots, v_n\}$$

is an ON basis for V of eigenvalues for T and all the eigenvalues are real and T is diagonalizable. ☐

By the HW/Takehome, we know

Theorem 22.6

Let V be a finite dimensional inner product space over $F, F = \mathbb{R}$ or $\mathbb{C}$. Let $\mathscr{B}, \mathscr{C}$ be ordered ON basis for V. Then

$$[1_V]_{\mathscr{B},\mathscr{C}} : F^{n\times 1} \to F^{n\times 1}$$

$n = \dim V$, is an ISOMETRY. In particular,

$$[1_V]_{\mathscr{B},\mathscr{C}}^{-1} = [1_V]_{\mathscr{B},\mathscr{C}}^{*}$$

$T : V \to W$ linear is called an ISOMETRY if

- T is an isomorphism.
- $\langle Tv_1, Tv_2\rangle_W = \langle v_1, v_2\rangle_V, \forall v_1, v_2 \in V$.

Theorem 22.7 (Spectral Theorem for Hermitian Operator (refined))

Let $F = \mathbb{R}$ or $\mathbb{C}$, V a finite dimensional inner product space over $F, T : V \to V$ hermitian. Then $\exists$ an ordered ON basis $\mathscr{C}$ of eigenvectors for V of T and every set of T if real. Moreover, if $\mathscr{B}$ is any ordered ON basis for V, then

$$[T]_{\mathscr{C}} = C[T]_{\mathscr{B}}C^*$$

for some invertible matrix $C \in M_nF$, i.e., $C = [1_V]_{\mathscr{B},\mathscr{C}}$.

Remark 22.8. The Spectral Theorem says, if V is a finite dimensional inner product space over $F, F = \mathbb{R}$ or $\mathbb{C}, T : V \to V$ hermitian, $\mathscr{B}$ an ordered ON basis for V, then

$$[T]_{\mathscr{B}} \sim \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, n = \dim V, \alpha_i \in \mathbb{R}, \forall i$$

if $V = \mathbb{R}^n$, this is often called the PRINCIPAL AXIS THEOREM.

e.g., It means if

$$f = \sum a_{ij}t_it_j \in \mathbb{R}[t_1, \ldots, t_n]$$

with

$$a_{ij} = a_{ji}, \forall i, j$$

This can always be arranged as $t_it_j = t_jt_i$ and we replace a_{ij}, a_{ji} with $\frac{a_{ij}+a_{ji}}{2}$ if necessary. Then we can change variables to make it look like

$$\lambda_1 I_1^2 + \ldots + \lambda_n I_n^2$$

(How? – Confer completing the square and $TAT^*, A = (a_{ij}), T^* = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}$. We want even more

Let $F = \mathbb{R}$ or $\mathbb{C}$, V a finite dimensional inner product space over $F, \dim V = n, T : V \to V$ hermitian, $\mathscr{B}$ an ordered ON basis of eigenvectors of T for V. Reordering $\mathscr{B}$ if necessary, we may assume $\lambda_1, \ldots, \lambda_k$ are all the distinct eigenvalues of T, i.e., if $j > k$ then $\exists i < k \ni \lambda_j = \lambda_i$.

Claim 22.1. Let $v \in E_T(\lambda_i)$, $w \in E_T(\lambda_j), 1 \leq i, j \leq k, i \neq j$. Then $v \perp w$: We may assume that $v \neq 0, w \neq 0$. So

$$\begin{aligned}\lambda_i \langle v, w\rangle &= \langle \lambda_i v, w\rangle = \langle Tv, w\rangle = \langle v, Tw\rangle \\ &= \langle v, \lambda_j w\rangle = \overline{\lambda_j}\langle v, w\rangle = \lambda_j \langle v, w\rangle\end{aligned}$$

as $\lambda_l \in \mathbb{R} \forall l$. Thus,

$$\left(\lambda_i - \lambda_j\right) \langle v, w\rangle = 0 \in F, \lambda_i \neq \lambda_j$$

so

$$\langle v, w\rangle = 0$$

Claim 22.2. We have

$$\begin{aligned}W &\coloneqq E_T(\lambda_1) + \ldots + E_T(\lambda_k) \\ &= E_T(\lambda_1) \oplus \ldots \oplus E_T(\lambda_k)\end{aligned} \tag{*}$$

if $w_i \in E_T(\lambda_i), i = 1, \ldots, k$ and

$$0 = w_1 + \ldots + w_k,$$

then

$$0 = \langle w_1 + \ldots + w_k, w_j\rangle = \langle w_j, w_j\rangle = \|w_j\|^2$$

by the previous claim, so $w_j = 0$ and (*) holds.

§23 Lec 23: Nov 30, 2020

§23.1 Lec 22 (Cont'd)

Note: Of course we already know this claim, but this proof is nice. Recall this is equivalent to $w = E_T(\lambda_1) + \ldots + E_T(\lambda_k)$ and

$$E_T(\lambda_i) \cap \sum_{j=1}^{k} E_T(\lambda_j) = 0, i = 1, \ldots, k$$

Also by the first claim, the DIRECT SUM DECOMPOSITION (*) of w is an ORTHOGONAL DIRECT SUM. Since $\mathscr{B}$ is a bases for V of eigenvectors for T and $\mathscr{B} \subset W$, we have

$$V = E_T(\lambda_1) \perp \ldots \perp E_T(\lambda_k) \quad (\star)$$

Genral Problem: Let V be a vector space over $F, T : V \to V$ linear operator. Can we DECOMPOSE V as

$$V = W_1 \oplus W_2 \oplus \ldots \oplus W_r \oplus \ldots$$

with each subspace W_i T-invariant, i.e., decomposition reflects the action T. This can be done if V is finite dimensional vector space over F. Then V is a finite direct sum. If $F = \mathbb{C}$, the solution is called JORDAN CANONICAL FORM.
F arbitrary is called RATIONAL CANONICAL FORM (done in 115B or 110BH).
By the OR Decomposition Theorem,

$$V = E_T(\lambda_i) \perp E_T(\lambda_i)^\perp, i = 1, \ldots, k \quad (**)$$

So

$$E_T(\lambda_i)^\perp = E_T(\lambda_i) \perp \ldots \perp E_T(\lambda_i) \perp \ldots \perp E_T(\lambda_k)$$

$i = 1, \ldots, k$ by uniqueness and, also by the OR Decomposition Theorem, as

$$V = E_T(\lambda_i) \perp E_T(\lambda_i)^\perp$$

means that $(\star)$ implies if $v \in V$, then

$$v = v_{E_T(\lambda_1)} + \ldots + v_{E_T(\lambda_k)}$$

where $v_{E_T(\lambda_i)}$ is the ORTHOGONAL PROJECTION of v onto $E_T(\lambda_i), i = 1, \ldots, k$. Define:

$$P_{\lambda_i} : V \to V \text{ by } v \mapsto v_{E_T(\lambda_i)}, i = 1, \ldots, k$$

As P_{λ_i} is the composition

$$V \to E_T(\lambda_i) \hookrightarrow V,$$
$$v \mapsto v_{E_T(\lambda_i)}$$

It is a linear operator, $i = 1, \ldots, k$. Moreover, by (**),

$$\text{im } P_{\lambda_i} = E_T(\lambda_i)$$
$$\ker P_{\lambda_i} = E_T(\lambda_i)^\perp$$

Since

$$P_{\lambda_j}(v_{E_T(\lambda_i)} = \delta_{ij} v_{E_T(\lambda_i)}, i = 1, \ldots, k$$

We see that

1. $P_{\lambda_i} P_{\lambda_j} = 0$ if $i \neq j$.
2. $P_{\lambda_i} P_{\lambda_i} = P_{\lambda_i}$.

So

$$P_{\lambda_i} P_{\lambda_j} = \delta_{ij} P_{\lambda_i} : V \to V \text{ linear}$$

The $P_{\lambda_1}, \ldots, P_{\lambda_k}$ are called ORTHOGONAL IDEMPOTENTS. We now see what we have done: Let $v \in V$. Then

$$\begin{aligned} 1_V v = v &= v_{E_T(\lambda_1)} + \ldots + v_{E_T(\lambda_k)} \\ &= P_{\lambda_1}(v) + \ldots + P_{\lambda_k}(v) = \left(P_{\lambda_1} + \ldots + P_{\lambda_k}\right)(v) \end{aligned}$$

So

$$1_V = P_{\lambda_1} + \ldots + P_{\lambda_k}$$

We also have

$$\begin{aligned} T = T \circ 1_V &= T \circ \left(P_{\lambda_1} + \ldots + P_{\lambda_k}\right) \\ &= T P_{\lambda_1} + \ldots + T P_{\lambda_k} \\ &= \lambda_1 P_{\lambda_1} + \ldots + \lambda_k P_{\lambda_k} \end{aligned}$$

as

$$\begin{aligned} \text{im } P_{\lambda_i} &= E_T(\lambda_i) \\ T\Big|_{E_T(\lambda_i)} &= \lambda_i 1_{E_T(\lambda_i)}, i = 1, \ldots, k \end{aligned}$$

We also have

$$\begin{aligned} 1_V \circ T &= \left(P_{\lambda_1} + \ldots + P_{\lambda_k}\right) T \\ &= P_{\lambda_1} T + \ldots + P_{\lambda_k} T \end{aligned}$$

and

$$P_{\lambda_i} T = T P_{\lambda_i}, i = 1, \ldots, k$$

This is called the SPECTRAL RESOLUTION of the Hermitian operator $T : V \to V$. Now, appropriately reordering $\mathscr{B}$ to $\mathscr{B}'$, we have, with

$$n_i = \dim E_T(\lambda_i), i = 1, \ldots, k$$

$$[T]_{\mathscr{B}'} = \begin{pmatrix} \lambda_1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_k & & \\ & & & & & \ddots & \\ 0 & & & & & & \lambda_k \end{pmatrix}$$

Summary(Spectral Theorem for Hermitian Operator – Full version):
Let $F = \mathbb{R}$ or $\mathbb{C}, V$ a finite dimensional inner product space over $F, T : V \to V$ hermitian, $\lambda_1, \ldots, \lambda_k$ all distinct eigenvalues of T. Then T is diagonalizable and

1. $\lambda_i \in \mathbb{R}, i = 1, \ldots, k$

2. Let $\mathscr{B}_i$ be an ordered ON basis for $E_T(\lambda_i), i = 1, \ldots, k$. Then $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_n$ is an ordered ON bases for V consisting of eigenvectors of T.

3.
$$[T]_{\mathscr{B}} = \begin{pmatrix} \lambda_1 & & & & 0 \\ & \ddots & \lambda_1 & & \\ & & & \ddots & \\ 0 & & & & \lambda_k \end{pmatrix}$$
$$n_i = \dim E_T(\lambda_i)$$
$$\dim V = n = n_1 + \ldots + n_k$$

4. $f_T = (t - \lambda_1)^{n_1} \ldots (t - \lambda_k)^{nk}$

5. $V = E_T(\lambda_1) \perp \ldots \perp E_T(\lambda_k)$

6. $1_V = P_{\lambda_1} + \ldots + P_{\lambda_k} : V \to V$ where $P_{\lambda_i} : V \to V$ linear by $v \mapsto v$

7. $P_{\lambda_i} P_{\lambda_j} = \delta_{ij} P_{\lambda_i}, i, j = 1, \ldots, k$

8. $T = \lambda_1 P_{\lambda_1} + \ldots + \lambda_k P_{\lambda_k}$

9. $TP_{\lambda_i} = P_{\lambda_i} T, i = 1, \ldots, k$

10. If $\mathscr{C}$ is an ON basis for V, then
$$\begin{aligned} [T]_{\mathscr{B}} &= [1_V]_{\mathscr{C},\mathscr{B}} [T]_{\mathscr{C}} [1_V]_{\mathscr{B},\mathscr{C}} \\ &= [1_V]_{\mathscr{C},\mathscr{B}} [T]_{\mathscr{C}} [1_V]^{-1}_{\mathscr{C},\mathscr{B}} \\ &= [1_V]_{\mathscr{C},\mathscr{B}} [T]_{\mathscr{C}} [1_V]^{*}_{\mathscr{C},\mathscr{B}} \end{aligned}$$
i.e., $[1_V]^{-1}_{\mathscr{B},\mathscr{C}} = [1_V]^{*}_{\mathscr{B},\mathscr{C}}$

Remark 23.1. One can also show that the MINIMAL POLYNOMIAL q_T of the HW/Takehome in the above is
$$q_T = (t - \lambda_1) \ldots (t - \lambda_k)$$

In fact this is a necessary and sufficient condition $\iff$ to be diagonalizable.

Remark 23.2. The Spectral Theorem for hermitian operator for $F = \mathbb{R}$, e.g., symmetric matrices, has a nice generalization:
Let F be a field with $2 \neq 0$ in F and $A \in M_n F$ a symmetric matrix, i.e., $A = A^t$. Then, $\exists$ an invertible matrix P in $M_n F \ni p^t A p$ is diagonal.

Note: in the above, we are not saying $p^t = p^{-1}$
Computation: To compute: let V be a finite dimensional vector space over F, $F = \mathbb{R}$ or $\mathbb{C}, T : V \to V$ hermitian. Find all the above:
Step 1: Find a basis for V and GS it to an OR bases, then normalize to an ON bases $\mathscr{C}$.
Step 2: Compute:
$$f_T = f_{[T]_{\mathscr{C}}} = \det(tI - [T]_{\mathscr{C}})$$
Step 3: Factor f_T, i.e., find all the roots of f_T. There are the eigenvalues of T. Since T is hermitian f_T splits and all the roots are real.

Step 4: For each eigenvalue of T, compute $E_T(\lambda)$ by solving

$$[T]_{\mathscr{C}}[v]_{\mathscr{C}} = \lambda[v]_{\mathscr{C}}$$

(equivalently row reduce $[T]_{\mathscr{C}} - \lambda I$ to row echelon form and solve).
Step 5: For each eigenvalue λ, find a basis for $E_T(\lambda_i)$ and GS to an ordered ON basis and normalize to an ordered ON basis $\mathscr{B}_\lambda$. Let $\mathscr{B} = \cup\mathscr{B}_\lambda$ an ordered ON basis of eigenvectors of T. As $\mathscr{C}$ is ON

$$[1_V]_{\mathscr{C},\mathscr{B}}[T]_{\mathscr{C}}[1_V]^*_{\mathscr{C},\mathscr{B}} \text{ is diagonal}$$

§24 Lec 24: Dec 2, 2020

§24.1 Normal Operators

We now need the following part of the Takehome

Theorem 24.1

Let V be a finite dimensional inner product space over F having an ordered ON basis $\mathscr{B}$, $T : V \to V$ linear. Then $\exists! T^* : V \to V$ linear s.t.

$$\langle Tv, w\rangle = \langle v, T^*w\rangle, \forall v, w \in V \tag{*}$$

called the ADJOINT of T. Moreover,

$$[T]^*_{\mathscr{B}} = [T^*]_{\mathscr{B}}$$

Remark 24.2. Actually, to prove (*), you do not need $\exists$ an ON basis, only an OR basis (which you know exist) if you prove it using dual bases.

Properties: Let V be a finite dimensional inner product space over F with an ON basis $\mathscr{B}, S, T : V \to V$ linear, $\lambda \in F$. Then $\forall v, w \in V$

(i) $\langle T^*v, w\rangle = \langle v, Tw\rangle$

(ii) $T^{**} := (T^*)^* = T$

(iii) $\langle v, T^*Tv\rangle = \langle Tv, Tv\rangle = \|Tv\|^2$

(iv) $\langle v, TT^*v\rangle = \langle T^*v, T^*v\rangle = \|T^*v\|^2$

(v) $(T \circ S)^* = S^* \circ T^*$

(vi) $(S + T)^* = S^* + T^*$

(vii) $(\lambda T)^* = \overline{\lambda} T^*, \forall \lambda \in F$.

Proof. Left as exercise. □

Remark 24.3. The above means: Let V be a finite dimensional inner product space over F with an ON basis. Then

$$\phi : L(V, V) \to L(V, V) \text{ by } T \to T^*$$

is a SESQUILINEAR transformation, i.e.,

$$\phi(\lambda T + S) = \overline{\lambda} T^* + S^*, \forall T, S \in L(V, V), \lambda \in F$$

and hence linear if $F \subset \mathbb{R}$ and is also bijection with inverse sesquilinear so a sesquilinear isomorphism.

Lemma 24.4 (New Key)

Let V be a finite dimensional inner product space over $F, T : V \to V$ linear. Suppose that V has an ON basis and $W \subset V$ is a T-invariant subspace. Then $W^\perp \subset V$ is T^*-invariant. In particular,

$$T^*\big|_{W^\perp} : W^\perp \to W^\perp \text{ is linear}$$

Proof. Let $w^\perp \in W^\perp$ and $x \in W$ be arbitrary. Then

$$\langle x, T^* w^\perp \rangle = \langle Tx, w^\perp \rangle = 0,$$

as $Tx \in W$ by hypothesis. So $T^* w^\perp \in W^\perp$ as needed. □

Definition 24.5 (Triangularizability) — Let V be a finite dimensional vector space over $F, T : V \to V$ linear. We say T is TRIANGULARIZABLE if $\exists$ an ordered basis $\mathscr{B}$ for $V \ni [T]_\mathscr{B}$ is upper triangular, i.e.,

$$[T]_\mathscr{B} = \begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix}$$

i.e., $([T]_\mathscr{B})_{ij} = 0$ if $i > j$.

Remark 24.6. In the above, $[T]_\mathscr{B}$ is upper triangular iff $[T]_\mathscr{B}$ is lower triangular where $\mathscr{B}'$ is an ordered basis with vectors in $\mathscr{B}$ in reverse ordered.

Theorem 24.7 (Schur)

Let V be a finite dimensional inner product space over $\mathbb{C}, T : V \to V$ linear. Then T is triangularizable. Moreover, $\exists$ an ordered ON basis $\mathscr{B}$ for $T \ni [T]_\mathscr{B}$ is upper triangular.

Proof. We induct on $n = \dim V$.

- $n = 1$: is immediate: if $\{v\}$ is a basis $\left\{\frac{v}{\|v\|}\right\}$ works.
- $n > 1$: By the FTA, the characteristics poly f_{T^*} for T^* has a root $\lambda \in \mathbb{C}$, hence λ is an eigenvalue of T^*. Let $0 \neq v \in E_{T^*}(\lambda)$. By the OR Decomposition Theorem,

  $$V = \mathbb{C}v \perp (\mathbb{C}v)^\perp$$

  and

  $$\begin{aligned} n = \dim V &= \dim \mathbb{C}v + \dim(\mathbb{C}v)^\perp \\ &= 1 + \dim(\mathbb{C}v)^\perp \end{aligned}$$

  i.e., $\dim(\mathbb{C}v)^\perp = n - 1$. $\mathbb{C}v$ is T^*-invariant as $v \in E_{T^*}(\lambda)$, so $(\mathbb{C}v)^\perp$ is $(T^*)^* = T$-invariant by New Key Lemma. So may view

  $$T\big|_{(\mathbb{C}v)^\perp} (\mathbb{C}v)^\perp \to (\mathbb{C}v)^\perp \text{ linear} \tag{*}$$

By induction, $\exists$ an ordered ON basis $\mathscr{B}_0 = \{v_1, \dots, v_{n-1}\}$ for $(\mathbb{C}v)^\perp \ni \left[T\big|_{(\mathbb{C}v)^\perp}\right]_{\mathscr{B}_0}$ is upper triangular. Let $\mathscr{B} = \left\{v_1, \dots, v_{n-1}, \frac{v}{\|v\|}\right\}$ an ordered ON basis for V. Then by (*), we have

$$\begin{pmatrix} \left[T\big|_{(\mathbb{C}v)^\perp}\right]_{\mathscr{B}_0} & * \\ & \vdots \\ & * \\ 0 \quad \dots & * \end{pmatrix} \in M_n\mathbb{C}$$

□

> **Remark 24.8.** As mentioned before, if F is arbitrary, V a finite dimensional vector space over F, then T is triangularizable $\iff$ $f_T, T : V \to V$ linear satisfies f_T splits, i.e., factors into a product of linear polys in $F[t]$.

Proof. ($\implies$) is clear as f_T is independent of a matrix representation.
($\impliedby$) is not clear and we not prove it. □

Corollary 24.9

Let V be a finite dimensional inner product space over $\mathbb{C}, T : V \to V$ linear, $\mathscr{C}$ an ordered ON basis for V. Then $\exists$ an ordered ON basis $\mathscr{B}$ for $V \ni [T]_{\mathscr{B}}$ is upper triangular and

$$[T]_{\mathscr{B}} = [1_V]_{\mathscr{C},\mathscr{B}}[T]_{\mathscr{C}}[1_V]^*_{\mathscr{C},\mathscr{B}}$$

with $[1_V]^{-1}_{\mathscr{C},\mathscr{B}} = [1_V]^*_{\mathscr{C},\mathscr{B}}$.

Proof. Theorem and HW as $\mathscr{C}, \mathscr{B}$ are ON. □

Definition 24.10 (Normal Operator) — Let V be an inner product space over $F, T : V \to V$ linear. Suppose that $T^* : V \to V$ exists, i.e.,

$$\langle Tv, w\rangle = \langle v, T^*w\rangle, \forall v, w \in V$$

with $T^* : V \to V$ linear. Then we say T is a NORMAL OPERATOR, if $TT^* = T^*T$.

§25 Lec 25: Nov 4, 2020

§25.1 Lec 24(Cont'd)

Example 25.1
1. Every hermitian operator is normal as $T = T^*$
2. Let $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be a rotation counterclockwise by $\angle\theta$ with $0 < \theta < 2\pi$ and $\theta \neq \pi$. Then T_θ has no eigenvalues in $\mathbb{R}$. Viewing $\mathbb{R}^2$ as an inner product space over $\mathbb{R}$ via the dot product.

$$T_{-\theta} = T_\theta^{-1} = T_\theta^t = T_\theta^*$$

So

$$T_\theta T_\theta^* = T_\theta^* T_\theta$$

and T_θ is normal. However, T_θ is not diagonalizable (is not even triangularizable). We shall show that this does not happen if $F = \mathbb{C}$, we start with (a replacement for the Hermitian Corollary)

Lemma 25.2 (Crucial Property of Normal Operators)
Let V be an inner product space over $F, T : V \to V$ normal, $\lambda \in F$. Let $0 \neq v \in V$. Then

$$v \in E_T(\lambda) \iff v \in E_{T^*}(\overline{\lambda})$$

i.e., λ is an eigenvalue of T with eigenvector $v \iff \overline{\lambda}$ is an eigenvalue of T^* with (the same) eigenvector v. So

$$Tv = \lambda v \iff T^* v = \overline{\lambda} v$$

if T is normal.

Proof. Suppose $S : V \to V$ is normal, $v \in V$. Then

$$\begin{aligned}\|Sv\|^2 &= \langle Sv, Sv\rangle = \langle v, S^*Sv\rangle \\ &= \langle v, SS^*v\rangle = \langle S^*v, S^*v\rangle = \|S^*v\|^2\end{aligned}$$

Hence

$$Sv = 0 \iff S^*v = 0 \text{ when S is normal} \tag{*}$$

Let $S = T - \lambda 1_V : V \to V$ linear. So λ is an eigenvalue of T iff $\ker S \neq 0$. But

$$S^* = (T - \lambda 1_V)^* = T^* - \overline{\lambda} 1_V$$

by properties of $()^*$. It follows that

$$S^*S = SS^* \text{ as } T^*T = TT^*$$

i.e., S is also normal. The result follows by (*). □

Theorem 25.3 (Spectral Theorem for Normal Operator)
Let V be a finite dimensional inner product space over $\mathbb{C}, T : V \to V$ normal. Then $\exists$ an ordered ON basis $\mathscr{C}$ for V consisting of eigenvectors of T. In particular, T is diagonalizable. Moreover, if $\mathscr{B}$ is an ordered ON basis for V, then

$$[T]_{\mathscr{C}} = [1_V]_{\mathscr{B},\mathscr{C}}[T]_{\mathscr{B}}[1_V]^*_{\mathscr{B},\mathscr{C}}$$

Proof. We induct on $n = \dim V$.

- $n = 1$ is immediate.
- $n > 1$: By the FTA, $\exists \overline{\lambda} \in \mathbb{C}$ a root of $f_{T^*} \in \mathbb{C}[t]$, hence an eigenvalue of T^*. Let $0 \neq v \in E_{T^*}(\overline{\lambda})$. By the lemma, $v \in E_T(\lambda)$. Thus, $\mathbb{C}_v$ is both T- and T^*-invariant. Hence, by New Key Lemma,

  $$(\mathbb{C}_v)^\perp \text{ is both } T^* \text{ and T-invariant}$$

  In particular,

  $$\langle x, T^*y\rangle = \langle Tx, y\rangle \quad \forall x, y \in (\mathbb{C}_v)^\perp$$

  and $\left(T\big|_{(\mathbb{C}v)^\perp}\right)^*$ is the unique linear map

  $$\left(T\big|_{(\mathbb{C}v)^\perp}\right)^* : (\mathbb{C}v)^\perp \to (\mathbb{C}v)^\perp$$

  satisfying $\forall x, y \in (\mathbb{C}v)^\perp$

  $$\begin{aligned}\langle x, \left(T\big|_{(\mathbb{C}v)^\perp}|^* y\right)\rangle_{(\mathbb{C}v)^\perp} &= \langle T\big|_{(\mathbb{C}v)^\perp} x, y\rangle_{(\mathbb{C}v)^\perp} \\ &= \langle Tx, y\rangle_V \\ &= \langle x, T^*y\rangle_V\end{aligned}$$

  It follows by the uniqueness of the adjoint that

  $$T^*\big|_{(\mathbb{C}v)^\perp} = \left(T\big|_{(\mathbb{C}v)^\perp}\right)^*$$

  Hence, we have

  $$T\big|_{(\mathbb{C}v)^\perp} : (\mathbb{C}v)^\perp \to (\mathbb{C}v)^\perp$$

  is also normal. Since

  $$\dim V = \dim \mathbb{C}v + \dim(\mathbb{C}v)^\perp = 1 + \dim(\mathbb{C}v)^\perp$$

  by the OR Decomposition Theorem, by induction $\exists$ an ON basis $\mathscr{C}_0 = \{v_2, \dots, v_n\}$ for $(\mathbb{C}v)^+$ of eigenvectors of $T\big|_{(\mathbb{C}v)^\perp}$ hence of eigenvectors of T. It follows that

  $$\mathscr{C} = \left\{\frac{v}{\|v\|}, v_2, \dots, v_n\right\}$$

  is an ON basis for V consisting of eigenvectors of T. If $\mathscr{B}$ is an ON basis for V, then $[1_V]^*_{\mathscr{B},\mathscr{C}} = [1_V]^{-1}_{\mathscr{B},\mathscr{C}}$ by Hw, so

  $$[T]_{\mathscr{C}} = [1_V]_{\mathscr{B},\mathscr{C}}[T]_{\mathscr{B}}[1_V]^*_{\mathscr{B},\mathscr{C}}$$

  by the change of basis theorem.

□

In fact, the converse is also true.

Theorem 25.4

Let V be a finite dimensional inner product space over $\mathbb{C}, T : V \to V$ linear. Then T is normal iff $\exists$ an ON basis $\mathscr{B}$ for V consisting of eigenvectors of T. In particular, T is diagonalizable if either holds.

Proof. ($\Longrightarrow$) Has been done.
($\Longleftarrow$) Let $\mathscr{B}$ has an ordered ON basis for V of eigenvectors of T. Then

$$[T]_{\mathscr{B}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, n = \dim V$$

As $\mathscr{B}$ is ON, by HW

$$[T^*]_{\mathscr{B}} = [T]^*_{\mathscr{B}} = \begin{pmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{pmatrix}$$

in $M_n\mathbb{C}$. So

$$\begin{aligned} [T^*T]_{\mathscr{B}} = [T^*]_{\mathscr{B}}[T]_{\mathscr{B}} &= \begin{pmatrix} |\lambda_1|^2 & & 0 \\ & \ddots & \\ 0 & & |\lambda_n|^2 \end{pmatrix} \\ &= [T]_{\mathscr{B}}[T^*]_{\mathscr{B}} = [TT^*]_{\mathscr{B}} \end{aligned}$$

(as $|\lambda_i|^2 = \lambda_i\overline{\lambda_i} = \overline{\lambda_i}\lambda_i \in \mathbb{C}$) By the Matrix Theory Theorem,

$$\phi : L(V, V) \to M_n\mathbb{C} \text{ by } S \mapsto [S]_{\mathscr{B}}$$

is an isomorphism, so

$$T^*T = TT^*$$

□

Remark 25.5. The result needs $F = \mathbb{C}$. Indeed if $V = \mathbb{R}^n, n > 1$, is an inner product space over $\mathbb{R}$ via the dot product and $T : V \to V$ is a rotation by an $\angle\theta, 0 < \theta < 2\pi, \theta \neq \pi$ in some plane through the origin in $\mathbb{R}^n$, then T is normal and not diagonalizable.

What is true is: Let $F = \mathbb{R}$ or $\mathbb{C}, V$ a finite dimensional inner product space over $F, T : V \to V$ linear $\exists$ an ON basis for $V \ni [T]_{\mathscr{B}}$ is triangularizable, then T is normal iff T is diagonalizable.

Remark 25.6. As in the Hermitian case, we can do more.

<u>Extension</u>: Let V be a finite dimensional inner product space over $\mathbb{C}, \dim V = n, T : V \to V$ normal, $\mathscr{C}$ an ordered basis of V of eigenvalues for normal T. After relabeling, we may assume $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T, i.e., if $j > k \exists i, 1 \leq i \leq k \ni \lambda_i = \lambda_j$.

Claim 25.1. Let $v \in E_T(\lambda_i), w \in E_T(\lambda_i), i \neq j, i \leq 1, j \leq k$. Then $v \perp w$.

Proof. We may assume that $v \neq 0$ and $w \neq 0$. As $w \in E_T(\lambda_j), w \in E_{T^*}(\overline{\lambda_j})$ by the lemma, as T is normal. Hence

$$\begin{aligned} \lambda_i\langle v, w\rangle = \langle \lambda_i v, w\rangle = \langle Tv, w\rangle &= \langle v, T^*w\rangle \\ &= \langle v, \overline{\lambda_j}w\rangle = \lambda_j\langle v, w\rangle \end{aligned}$$

Since $\lambda_i \neq \lambda_j, \langle v, w\rangle = 0$. □

§26 Lec 26: Dec 7, 2020

§26.1 Lec 25 (Cont'd)

Let V be a vector space over F, $W_i \subset V, i \in I$ subspace. Suppose that $V = \sum_I W_i$. Then V is a DIRECT SUM of the $W_i, i \in I$ write $V = \bigoplus_I W_i$ if one of the following equivalent condition hold

1. $\forall v \in V \exists! w_i \in W_i \ni w_i = 0$ almost all i and $v = \sum_I w_i$
2. If $w_i \in W_i$, almost all $w_i = 0$, and $0 = \sum_I w_i$, then $w_i = 0 \forall i \in I$
3. $\forall i \in I$
$$W_i \cap \sum_{j \in I, j \neq i} W_j = 0$$
4. If $\mathscr{B}_i$ is a basis for W_i, $i \in I$, then $\mathscr{B} = \cup \mathscr{B}_i$ is a basis for V.

If V is also an inner product space over F, and $V = \bigoplus_I W_i$ with $\langle w_i, w_j \rangle = 0 \forall i \neq j$ in I, we call V an orthogonal direct sum and write $V = \frac{1}{I} W_i$.
Since $\lambda_i \neq \lambda_j, \langle v, w \rangle = 0$. Let

$$W = E_T(\lambda_1) + \ldots + E_T(\lambda_k)$$

It is a direct OR sum for if

$$0 = w_1 + \ldots + w_k, w_i \in E_T(\lambda_i), i = 1, \ldots, k$$

then

$$\begin{aligned} 0 &= \langle 0, w_j \rangle = \langle w_1 + \ldots + w_k, w_j \rangle = \langle w_j, w_j \rangle \\ &= \|w_j\|^2 \end{aligned}$$

$j = 1, \ldots, k$. Hence $w_j = 0 \forall i$ and

$$W = E_T(\lambda_1| \oplus \ldots \oplus E_T(\lambda_k))$$

(why – uniqueness follows immediately) and $\mathscr{C}$ is a basis for V, so

$$V = E_T(\lambda_1) \perp \ldots \perp E_T(\lambda_k)$$

By the OR Decomposition Theorem,

$$E_T(\lambda_i)^\perp = E_T(\lambda_1) \perp \ldots \perp E_T(\lambda_i) \perp \ldots \perp E_T(\lambda_k)$$

and if $v \in V$

$$v = w_1 + \ldots + w_k, w_i \in W_i \text{ unique}$$

So

$$w_i = v_{E_T(\lambda_i)}$$

the OR properties of v an $E_T(\lambda_i)$ for $i = 1, \ldots, k$ by the OR Decomposition Theorem, as

$$V = E_T(\lambda_i) \perp E_T(\lambda_i)^\perp$$

Let

$$P_{\lambda_i} : V \to V \text{ by } v \mapsto v_{E_T(\lambda_i)}, i = 1, \ldots, k$$

be the composition

$$\begin{aligned} V &\to E_T(\lambda_i) \hookrightarrow V \\ v &\mapsto v_{E_T(\lambda_i)} \end{aligned}$$

a linear operator

$$\begin{aligned} \text{im } P_{\lambda_i} &= E_T(\lambda_i) \\ \ker P_{\lambda_i} &= E_T(\lambda_i)^\perp \\ P_{\lambda_i} P_{\lambda_j} &= \delta_{ij} P_{\lambda_j}, \forall i, j \end{aligned}$$

i.e., $P_{\lambda_1}, \ldots, P_{\lambda_k}$ are ORTHOGONAL IDEMPOTENTS and we see $\forall v \in V$

$$\begin{aligned} v &= P_{\lambda_1} v + \ldots + P_{\lambda_k} v \\ 1_V &= P_{\lambda_1} + \ldots + P_{\lambda_k} \end{aligned}$$

So

$$\begin{aligned} T &= T \circ 1_V = T \circ P_{\lambda_1} + \ldots + T \circ P_{\lambda_k} = \lambda_1 P_{\lambda_1} + \ldots + \lambda_k P_{\lambda_k} \\ T &= 1_V T = P_{\lambda_1} T + \ldots + P_{\lambda_k} T \\ T P_{\lambda_i} &= P_{\lambda_i} T, \forall i \end{aligned}$$

as

$$T\Big|_{E_T(\lambda_i)} = \lambda_i 1_{E_T(\lambda_i)}, i = 1, \ldots, k$$

This is the SPECTRAL RESOLUTION of T if $n_i = \dim E_T(\lambda_i), \mathscr{B}_i$ an ordered ON basis for $E_T(\lambda_i)$, $\mathscr{B}_i$ an ordered ON basis for $E_T(\lambda_i), i = 1, \ldots, k$. Then $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_k$ is an ordered ON basis for V consisting of eigenvectors of T

$$\begin{aligned} n &= \dim V = n_1 + \ldots n_k \\ f_T &= (t - \lambda_1)^{n_1} \ldots (t - \lambda_k)^{nk} \\ [T]_{\mathscr{B}} &= \begin{pmatrix} \lambda_1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_k & & \\ & & & & & \ddots & \\ 0 & & & & & & \lambda_k \end{pmatrix} \end{aligned}$$

Theorem 26.1 (Spectral Theorem for Normal Operator - Full Version)
Let $F = \mathbb{C}, V$ a finite dimensional inner product space over $\mathbb{C}, T : V \to V$ normal, $\lambda_1, \ldots, \lambda_k$ all the distinct eigenvalues of T. Then T is diagonalizable and

1. Let $\mathscr{B}_i$ be an ordered ON basis for $E_T(\lambda_i)$, $i = 1, \ldots, k$. Then $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_n$ is an ordered ON basis for V (obvious order) consisting of eigenvectors of T.

2. $$[T]_{\mathscr{B}} = \begin{pmatrix} \lambda_1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_k & & \\ & & & & & \ddots & \\ 0 & & & & & & \lambda_k \end{pmatrix}$$

   where

   $$n_i = \dim E_T(\lambda_i), i = 1, \dots, k$$
   $$\dim V = n = n_1 + \dots + n_k$$

3. $f_T = (t - \lambda_1)^{n_1} \dots (t - \lambda_k)^{nk}$
4. $V = E_T(\lambda_1) \perp \dots \perp E_T(\lambda_k)$
5. $1_V = P_{\lambda_1} + \dots + P_{\lambda_k} : V \to V$ where $P_{\lambda_i} : v \to v$ linear by $v \mapsto v_{E_T(\lambda_i), i=1,\dots,k}$ (viewed in V).
6. $P_{\lambda_i} P_{\lambda_j} = \delta_{ij} P_{\lambda_i}, i, j = 1, \dots, k$
7. $T = \lambda_1 P_{\lambda_1} + \dots + \lambda_k P_{\lambda_k}$
8. $T P_{\lambda_i} = P_{\lambda_i} T, i = 1, \dots, k$
9. If $\mathscr{C}$ is an ON basis for V then

   $$\begin{aligned} [T]_{\mathscr{B}} &= [1_V]_{\mathscr{C},\mathscr{B}} [T]_{\mathscr{C}} [1_V]_{\mathscr{B},\mathscr{C}} \\ &= [1_V]_{\mathscr{C},\mathscr{B}} [T]_{\mathscr{C}} [1_V]^{-1}_{\mathscr{C},\mathscr{B}} \\ &= [1_V]_{\mathscr{C},\mathscr{B}} [T]_{\mathscr{C}} [1_V]^{*}_{\mathscr{C},\mathscr{B}} \end{aligned}$$

   i.e., $[1_V]^{-1}_{\mathscr{B},\mathscr{C}} = [1_V]^{*}_{\mathscr{B},\mathscr{C}}$
10. $q_T = (t - \lambda_1) \dots (t - \lambda_k)$

Now T is normal so T^* is also normal with distinct eigenvalues $\overline{\lambda_1}, \dots, \overline{\lambda_k}$ and

$$E_T(\lambda_i) = E_{T^*}(\overline{\lambda_i}), i = 1, \dots, k$$

In fact, as

$$Tv = \lambda_i v \iff T^* v = \overline{\lambda_i} v$$

the orthogonal projection

$$P_{\overline{\lambda_1}}, \dots, P_{\overline{\lambda_k}}$$

for T^* satisfy

$$P_{\lambda_i} = P_{\overline{\lambda_i}}, i = 1, \dots, k$$

as

$$v_{E_T(\lambda_i)} = v_{E^*_T(\overline{\lambda_i})}$$

Hence the spectral resolution for T^* is

$$\begin{aligned} T^* &= \overline{\lambda_1} P_{\overline{\lambda_1}} + \dots + \overline{\lambda_k} P_{\overline{\lambda_k}} \\ &= \overline{\lambda_1} P_{\lambda_1} + \dots + \overline{\lambda_k} P_{\lambda_k} \end{aligned}$$

§27 Lec 27: Dec 9, 2020

§27.1 Lec 26 (Cont'd)

We make a further computation using the Spectral Resolution of normal $T : V \to V, V$ a finite dimensional inner product space over $\mathbb{C}$. This also holds for hermitian $T : V \to V, V$ a finite dimensional inner product space over $\mathbb{R}$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_k$, orthogonal idempotents $P_{\lambda_1}, \ldots, P_{\lambda_k}$ i.e, spectral resolution.

$$T = \lambda_1 P_{\lambda_1} + \ldots + \lambda_k P_{\lambda_k}$$

As $P_{\lambda_i} P_{\lambda_j} = \delta_{ij} P_{\lambda_i}$, we have

$$T^2 = (\lambda_1 P_{\lambda_1} + \ldots + \lambda_k P_{\lambda_k})(\lambda_1 P_{\lambda_1} + \ldots + \lambda_k P_{\lambda_k}) = \lambda_1^2 P_{\lambda_1} + \ldots + \lambda_k^2 P_{\lambda_k}$$

An easy induction shows

$$T^m = \lambda_1^m P_{\lambda_1} + \ldots + \lambda_k^m P_{\lambda_k}, m \in \mathbb{Z}^+$$

Since

$$1_V = P_{\lambda_1} + \ldots + P_{\lambda_k}$$

we see that if for any

$$f = a_m t^m + a_{m-1} t^{m-1} + \ldots a_0 \in F[t]$$

a poly (with $F = \mathbb{C}$ if T normal, $F = \mathbb{R}$ or $\mathbb{C}$ if T is hermitian) that

$$f(T) = a_m T^m + \ldots + a_0 1_V$$
$$f(T^*) = a_m T^{*m} + \ldots + a_0 1_V$$

and as $f(T)$ is also normal (resp hermitian)

$$f(T) = \sum_{i=1}^{k} f(\lambda_i) P_{\lambda_i}$$
$$f(T^*) = \sum_{i=1}^{k} f_i(\overline{\lambda}_i) P_{\lambda_i} \forall f \in \mathbb{C}[t]$$

Now let $m = k - 1$. Set

$$f_i = \prod_{j=1, j \neq i}^{k} \frac{(t - \lambda_j)}{\lambda_i - \lambda_j} \in \mathbb{C}[t], j = 1, \ldots, k$$

the LAGRANGE POLY associated to $\lambda_1, \ldots, \lambda_k$. By the LAGRANGE INTERPOLATION THEOREM, $\exists! g \in \mathbb{C}[t]$, deg g ¡ k, $\lambda \ni g(\lambda_i) = \overline{\lambda_i}, i = 1, \ldots, k$. Thus by the above, we have

$$g(T) = g(\lambda_1) P_{\lambda_1} + \ldots + g(\lambda_k) P_{\lambda_k} = \overline{\lambda}_1 P_{\lambda_1} + \ldots + \overline{\lambda}_k P_{\lambda_k} = T^* \qquad (\star)$$

i.e., T^* is a polynomial in T.

Proposition 27.1

Let $F = \mathbb{C}$, V a finite dimensional inner product space over $\mathbb{C}, T : V \to V$ linear. Then the following are true

1. T is normal iff $\exists g \in \mathbb{C}[t] \ni T^* = g(T)$.
2. T is isometry iff T is normal and $|\lambda| = 1$ for every eigenvalue λ of T.
3. If T is normal, then T is hermitian iff every eigenvalue of T is real.

Proof. 1. $\to$ is $(\star)$,

$$Tg(T) = g(T)T$$

T^* is normal.

2. $\to$ If T is an isometry, then $T^* = T^{-1}$. Let $\mathscr{B}$ be an ON basis for V, the cols of $[T]_{\mathscr{B}}$ corresponds to an ON basis for V and we are done via the $\phi : L(V, V) \to M_n\mathbb{C}, T \mapsto [T]_{\mathscr{B}}$, i.e. MTT. In particular, $1_V = TT^* = T^*T$, so T is normal if $v \in V$ then we know

$$v \in E_T(\lambda) \iff v \in E_{T^*}(\overline{\lambda})$$

i.e.,

$$Tv = \lambda v \iff T^*v = \overline{\lambda}v$$

So if $v \in E_T(\lambda)$, ...

We have

$$TT^* = |\lambda_1|^2 P_{\lambda_1} + \ldots + |\lambda_k|^2 P_{\lambda_k}$$

Since $|\lambda_i| = 1 \forall i$,

$$TT^* = P_{\lambda_1} + \ldots + P_{\lambda_k} = 1_V = T^*T$$

Therefore,

$$\|v\|^2 = \langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2$$

i.e., $\|v\| = \|Tv\| \forall v \in V$. By Hw, T is an isometry.

3. $\to$ is the Hermitian Corollary.

$\leftarrow$)$\lambda_i \in \mathbb{R}$ eigenvalues of normal T implies $T = T^*$ by $(\star)$.

□

§27.2 Singular Value Theorem

Theorem 27.2 (Singular Value)
Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in F^{m\times n}$. Then

$$\exists u \in U_n(F) := \{B \in M_nF | BB^* = I\}, X \in U_nF \ni$$

$$X^*AU = D := \begin{pmatrix} u_1 & & & & & 0 \\ & \ddots & & & & \\ & & u_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} \in F^{m\times n}$$

diagonal, i.e. $D_{ij} = 0 \forall i \neq j$ with $D_{ii} = 0 \forall i > r, D_{ii} = \mu_i, i \leq r$ with

$$\mu_i \gg \ldots \gg \mu_r > 0$$

and $r = \text{rank } A$

Proof. $A^*A \in M_nF$ is hermitian with non-negative real eigenvalues using problem 9 of the Take home. Let $\lambda_1, \ldots, \lambda_r$ be the positive eigenvalues ordered such that

$$\lambda_1 \gg \ldots \lambda_r > 0$$

(there can be repetitions). By the Spectral Theorem for Hermitian Operators, $\exists U \in U_nF \ni$

$$(AU)^*(AU) = U^*(A^*A)U = \begin{pmatrix} \lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} \in M_nF$$

(as $A = [A]_{\mathscr{S}_n, \mathscr{S}_m}$). Let

$$C = AU \in F^{m\times n}$$

So

$$C^*C = (AU)^*(AU) \in M_nF$$

Write

$$\lambda_i = \mu_i^2, \mu_i > 0, 1 \leq i \leq r$$

(which we can do as $\lambda_i > 0 \in \mathbb{R}$) and let

$$\lambda_i = 0 \text{ for } i > r$$

Set

$$B = \begin{pmatrix} \mu_1 & & & & & 0 \\ & \ddots & & & & \\ & & \mu_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0 \end{pmatrix} \in M_nF$$

if E is a matrix let $E^{(k)}$ denote the k^{th} column of E. Then we have

$$\begin{aligned}\lambda_i\delta_{ij} = (C^*C)_{ij} &= \sum_{l=1}^{n}(C^*)_{il}C_{l_j} = \sum_{i=1}^{n}\overline{C_{l_i}}C_{l_j} \\ &= \sum_{l=1}^{n} C_{lj}\overline{C_{l_i}} = \langle C^{(j)}, C^{(i)}\rangle\end{aligned}$$

Hence

$$C = \begin{bmatrix}C^{(1)} & \dots & C^{(r)} & 0 & 0\end{bmatrix} \in F^{m\times n}$$

satisfies $\mathscr{C}_0 = \left\{C^{(1)},\dots,C^{(r)}\right\}$ is an OR set in $F^{m\times 1}$. As $C^{(i)} \neq 0, 1 \le i \le r, \mathscr{C}_0$ is linearly independent. Therefore,

$$\text{Rank } C = r$$

with

$$\|C^{(i)}\|^2 = \langle C^{(i)}, C^{(i)}\rangle = \lambda_i = \mu_i^2$$

for $i = 1,\dots,r$. As U is invertible

$$\text{Rank } A = \text{Rank } AU = \text{Rank } C = r,$$

i.e.,

$$\text{Rank } A = r$$

as required. Now define

$$X^{(i)} := \frac{1}{\mu_i}C^{(i)} \in F^{m\times 1}, i = 1,\dots,r$$

Then $\mathscr{B}_0 = \left\{X^{(1)},\dots,X^{(r)}\right\}$ is an ON set in $F^{m\times 1}$. Extend this to an ordered ON basis

$$\mathscr{B} = \left\{X^{(1)},\dots,X^{(m)}\right\} \text{ for } F^{m\times 1}$$

Then the matrix

$$X = \begin{bmatrix}X^{(1)}\dots X^{(m)}\end{bmatrix} = [1_{F^{m\times 1}}]_{\mathscr{B},\mathscr{S}_{m,1}} \in M_mF$$

Since $\mathscr{B}, \mathscr{S}_{m,1}$ are ON bases

$$X \in U_m(F)$$

Set

$$D = \begin{pmatrix}\mu_1 & & & & & 0\\ & \ddots & & & & \\ & & \mu_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0\end{pmatrix} \in F^{m\times n}$$

as in the statement of the theorem.

$$XD = \begin{bmatrix}X^{(1)}\dots X^{(m)}\end{bmatrix}\begin{pmatrix}\mu_1 & & & & & 0\\ & \ddots & & & & \\ & & \mu_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 0\end{pmatrix}$$

$$\begin{bmatrix}\mu_1X^{(1)}\dots\mu_rX^{(r)}0\dots 0\end{bmatrix} = C = AU$$

Hence

$$X^*AU = D$$

as needed. □

§28 Lec 28: Dec 11, 2020

§28.1 Lec 27 (Cont'd)

Definition 28.1 (Singular Value Decomposition) — Let $A \in F^{m\times n}, F = \mathbb{R}$ or $\mathbb{C}$

(i) $A = XDU^*, U \in U_nF, X \in U_mF$ (so $D = X^*AU$ as $X^{-1} = X^*, U^{-1} = U^*$)

(ii) $\mu_1 \geq \ldots \geq \mu_r > 0 \in \mathbb{R}$ where

(iii)

$$D = \begin{pmatrix} \mu_1 & & & & & \\ & \ddots & & & & \\ & & \mu_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

Then $i), ii), iii)$ is called a SINGULAR VALUE DECOMPOSITION (SVD) for A, $\mu_1, \ldots, \mu_r$ the singular values of A, D 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 forms an ON basis for $F^{m\times 1}$ of eigenvectors of AA^*
3. The rows of U form an ON basis for $F^{1\times n}$ 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 an SVD of $A, X \in U_mF, U \in U_nF$. Then

$$A^*A = (XDU^*)^*(XDU^*) = UD^*X^*XDU^* = UD^*DU^*$$

as $X^*X = I$, so

$$A^*A \sim D^*D = \begin{pmatrix} d_{11}^2 & \\ & \ddots \end{pmatrix}$$

have the same eigenvalues, $d_{11}^2, \ldots,$ i.e., these are the eigenvalues of AA^*. □

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. □

Theorem 28.5 (Polar Decomposition)

Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in M_nF$. Then $\exists U^\sim \in U_nF, N \in M_NF$ hermitian (i.e., $N = N^*$) with all its (real) eigenvalues non-negative s.t.

$$A = U^\sim N$$

cf. polar form of a complex number $U^\sim \leftrightarrow e^{\sqrt{-1}\theta}, N \leftrightarrow r$.

Proof. In the Singular Value Theorem, we have $m = n$, so if

$$A = XDU^* \text{ is an SVD} \quad X, U \in U_uF,$$

We have $D = D^*$ is hermitian with non-negative eigenvalues $AU = XD$. So

$$A = XDU^* = X(U^*U)DU^* = (XU^*)(UDU^*)$$

Since

$$(XU^*)^*(XU^*) = UX^*XU^* = UU^* = I,$$

we have $XU^* \in U_nF$.
So letting $U^\sim = XU^* \in U_nF, N = UDU^*$ work. □

Exercise 28.1. In the above theorem, N is unique and U is unique if A invertible in M_nF. (as it has positive eigenvalues).

§28.2 Application of SVD

Problem 28.1. Let $F = \mathbb{R}$ or $\mathbb{C}, V$ a finite dimensional inner product space over $F, W \subset V$ a subspace

$$P_W : V \to W \text{ by } v \mapsto v_W$$

the orthogonal projection of V onto W. We know v_W is the BEST APPROXIMATION of $v \in V$ onto W. Now let X be another finite dimensional inner product space over $F, T : X \to V$ linear, $W = T(X) = \text{im } T, v \in V, x \in X$. We call

(i) X a best approximation to v via T if

$$T_x = v_W = P_W(v)$$

(ii) X an optimal approximation to v via T if it is a best approximation to v via T and $\|v\|$ is minimal among all best approximations to v via T.

In the above, find an optimal approximation of x.

Ans: Let $A = T : F^{n\times 1} \to F^{m\times 1}, A \in F^{m\times n}, v \in F^{m\times 1}$ ($F = \mathbb{R}$ or $\mathbb{C}$). Let $A = XDU^*$ be an SVD

$$D = \begin{pmatrix} \mu_1 & & & & & \\ & \ddots & & & & \\ & & \mu_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \in F^{m\times n}$$

$\mu_1 \geq \ldots \geq \mu_r > 0 \in \mathbb{R}$. Define

$$D^\dagger = \begin{pmatrix} \mu_1^{-1} & & & & & \\ & \ddots & & & & \\ & & \mu_r^{-1} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \in F^{n\times m}$$

$$A^\dagger := UD^\dagger X^* \in F^{n\times m}$$

called the Moore-Penrose generalized pseudo-inverse of A. Then

(i) rank $A =$ rank $A^\dagger$

(ii) $A^\dagger v$ is an optimal approximation in $F^{n\times 1}$ to v via A and is unique. (Hence $A^\dagger$ is well-defined, i.e., independent of SVD)

(iii) If rank $A = n$, then
$$A^\dagger = (A^*A)^{-1}A^*$$

Application (Least square): $F = \mathbb{R}$ or $\mathbb{C}$. Given date $(x_1, y_1), \ldots, (x_n, y_n) \in F^2$. Find the best line relative to this data, i.e., find

$$y = \lambda x + b, \lambda = \text{slope}$$

Let

$$A = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}, X = \begin{pmatrix} \lambda \\ b \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Solve $AX = Y$. The solution is probably inconsistent, so want optimal soln. Solve

$$\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \begin{pmatrix} \lambda \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

(Least squares approximation) Let $W = \text{im } A$. To find optimal approximation to

$$AX = Y_W$$

Then $X = A^\dagger y$ works. If rank $A = 2$, then

$$X = (A^*A)^{-1}A^*Y$$

§28.3 Smith Normal Form

Polynomials are important in analyzing linear operator $T : V \to V$, V a finite dimensional vector space over F, e.g., f_T, q_T. Algebraically, this arises from the generalization of a vector space over F.
Let $\mathbb{R}$ be a ring, i.e., axioms of a field except M3, M4 (inverse and commutativity). Let M be a set satisfying $A1 - A4$, i.e., axiom for $+$ in $\mathbb{Z}$. Then M is called a (left) $R-$Module via

$$\cdot : R \times M \to M \qquad (r, m) \mapsto rm$$

if $(M, +., \cdot)$ satisfies the axioms of a vector space over F with R replacing a field.
For linear algebra, this arises as follows: Let V be a vector space over F, a set $T : V \to V$ a linear operator. Make V into a $F[t]$-module by $\forall v \in V \forall g \in F[t]$

$$g \cdot v :\mapsto g(T)v$$

We let t in $F[t]$ act on V by

$$tv := T(v)$$

Then use module theory to break V into $v = w_1 \oplus \ldots \oplus w_r, w_i$ T-invariant $\forall i$ (and nice) if V is a finite dimensional vector space over F.
We say that $A \in F[t]^{m\times n}$ is in **Smith Normal Form** (or SNF) if A is the zero matrix or if A is a 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 element of $F[t]$ do not have multiplicative inverses.
Let $A \in M_n(F[t])$. 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 28.6. 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 28.7.

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 define over a field.
3. The elementary matrices of Type III have a restriction. The u's appearing in the definition are precisely the element in $F[t]$ having a multiplicative inverse TBA

Notation: We let

$$GL_n(F[t]) := \{A \in M_n(F[t]) | A \text{ is invertible}\}$$

Warning: A matrix in $M_n(F[t])$ having $\det(A) \neq 0$ may no longer be invertible, i.e., have an inverse. What is true is that $GL_n(F[t]) = \{A \in M_n(F[t]) | 0 \neq \det(A) \in F\}$, equivalently $GL_n(F[t])$ consist of those matrices whose determinant have a multiplicative inverse in $F[t]$.

Definition 28.8 (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 exists matrices $P \in GL_m(F[t])$ and $Q \in GL_n(F[t])$ s.t. $B = PAQ$.

Theorem 28.9

Let $A \in F[t]^{m \times n}$. Then A is equivalent to a matrix in Smith Normal Form (SNF). Moreover, there exists matrices $P \in GL_m(F[t])$ and $Q \in GL_n(F[t])$, each a product of matrices of Type I, Type II, and Type III, such that PAQ is in SNF.

Proof. The proof will, in fact, be an algorithm to find a SNF of A. Refer to `www.math.ucla.edu/~rse/115ah.1.20f/L28.pdf` – Pg. 9-10. □

Remark 28.10. The SNF derived by this algorithm is, in fact, unique. In particular, the monic polynomial $q_1 | q_2 | q_3 | \ldots | q_r$ arising in the Smith Normal Form of a matrix A are unique and are called the **invariant factors** of A. This is proven using results about determinants. It follows if $A, B \in F[t]^{m \times n}$ then $A \sim B$ if and only if they have the same SNF if and only if they have the same invariant factors.

So what good is the SNF relative to linear operators on finite dimensional vector spaces? It tells us a great deal, because the following is true: Let $A, B \in M_n(F)$. Then $A \sim B$ if and only if $tI - A \approx tI - B \in M_n(F[t])$ 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 of 1, and we shall drop these.

§28.4 Some definitions

Definition 28.11 (Companion Matrix) — Let $q = t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0$ be a monic polynomial in $F[t]$. The **companion matrix** $C(q)$ is defined to be the $n \times n$ matrix:

$$\begin{pmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{pmatrix}$$

Definition 28.12 (Invariant Factors) — 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 Smith Normal Form of $tI - [T]_{\mathscr{B}}$, it will have the form

$$\begin{pmatrix} 1 & 0 & & \ldots & \ldots & & 0 \\ 0 & 1 & & & & & 0 \\ \vdots & & \ddots & & & & \vdots \\ & & & q_1 & & & \\ & & & & q_2 & & \\ \vdots & & & & & \ddots & \vdots \\ 0 & & & \ldots & \ldots & & q_r \end{pmatrix}$$

with $q_1 | q_1 | \ldots | q_r$ are all the monic polynomials in $F[t] \setminus F$. These are called the **invariant factors** of T. They are uniquely determined by T.

Definition 28.13 (Rational Canonical Form) — The main theorem is that there exists an ordered basis $\mathscr{B}$ for V such that

$$[T]_{\mathscr{B}} = \begin{pmatrix} C(q_1) & 0 & \ldots & 0 \\ 0 & C(q_2) & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \ldots & C(q_r) \end{pmatrix}$$

and this matrix representation is unique. This is called the **rational canonical form** or RCF of T. Moreover, the minimal polynomial of T is q_r. The algorithm computes this as well as all invariant factors of T. The characteristics polynomial f_T of T is the product of $q_1 \ldots q_r$. 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 $V = W_1 \oplus \ldots \oplus W_r$ where $f_{T|W_i} = q_{T|W_i} = q_i$ and if $\dim(W_i) = n_i$, then $\mathscr{B}_i = \left\{v_i, Tv_i, \ldots, T^{n_i - 1} v_i\right\}$ is a basis for W_i (we say that the W_i are T-cyclic subspaces).

Definition 28.14 (Jordan Block/Size – Jordan Canonical Form) — 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. Say

$$q_i = (t - \lambda_1)^{r_1} \dots (t - \lambda_m)^{r_m}, i = 1, \dots, m$$

in $F[t]$, with $\lambda_1, \dots, \lambda_m$ distinct. A matrix in $M_r(F)$ of the form

$$J_r(\lambda) = \begin{pmatrix} \lambda & 0 & \dots & 0 & 0 \\ 1 & \lambda & 0 & \dots & 0 \\ 0 & 1 & \lambda & \dots & \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & \lambda \end{pmatrix}$$

is called a **Jordan block** or **size** $r \times r$ with eigenvalue λ. The one can show that $C(q_i), i = 1, \dots, m$ is similar to the following matrix in block form:

$$\begin{pmatrix} J_{r_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{r_2}(\lambda_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{r_m}(\lambda_m) \end{pmatrix}$$

Replacing each $C(q_i)$ in the rational canonical form by its Jordan blocks give what is called **Jordan Canonical Form** or **JCF** of T. It is unique up to the order of the blocks (blocks with the same eigenvalues are usually put together).

§29 Extra Lec: Nov 2/9, 2020

§29.1 Dual Bases – Dual Spaces

Let $0 \neq V$ be a vector space over F with basis $\mathscr{B}$. For each $v_0 \in \mathscr{B}$, we define a map

$$f_{v_0} : V \to F \text{ linear}$$

as follows: by the UPVS (which also holds if the basis is infinite, let fv_0 be the unique linear transformation) s.t.

$$\begin{aligned} v_0 &\mapsto 1 \\ v &\mapsto 0 \quad \forall v_0 \neq v \in \mathscr{B} \end{aligned}$$

We have

$$0 < \text{ im } fv_0 \subset F \text{ a subspace}$$

(im $fv_0 \neq 0$ as $v_0 \neq 0$). As $\dim_F F = 1$, we must have $\dim fv_0 = 1$, so $fv_0 : V \to F$ is an epimorphism and

$$\begin{aligned} \ker fv_0 &= \{w \in V | w \text{ has } v_0 \text{ coordinate} = 0\} \\ &= \text{Span}(\mathscr{B} \setminus \{v_0\}) \end{aligned}$$

So if $w \in V$, $w = \sum \alpha_v v, \alpha_v \in F$ almost all 0 with α_v unique.

$$fv_0(w) = \alpha_{v_0}$$

the coordinate of w on v_0. We can do this for each $v \in \mathscr{B}$. If $v' \in \mathscr{B}, f_V : V \to F$ is the linear transformation determined by

$$f_{v'}(v) = \delta_{vv'} = \begin{cases} i, & \text{if } v = v' \\ 0, & \text{if } v \neq v', v \in \mathscr{B} \end{cases}, \text{ the Kronecker } \delta$$

Set

$$\mathscr{B}^* := \{fv | v \in \mathscr{B}\} \; f_v \text{ is the coordinate function} f_v \text{ on } v$$

The vector space

$$V^* := L(V, F)$$

is called the DUAL SPACE of V. So by the above if $w \in V$

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

then

$$\alpha_v = f_v(w) \text{ the coordinate } w, v \in \mathscr{B}$$

so

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

Now by the UPVS, we have a unique linear transformation

$$D_{\mathscr{B}} : V \to V^{\times}$$

determined by $v \in \mathscr{B} \mapsto f_v$. So $\sum_{\mathscr{B}} \alpha_v v \mapsto \sum_{\mathscr{B}} \alpha_v f_v$ almost all $\alpha_v = 0$

Claim 29.1. *$D_{\mathscr{B}}$ is 1-1.*

Suppose $w = \sum_{\mathscr{B}} \alpha_v v \mapsto 0$ almost all $\alpha_v = 0$ i.e., $\sum_{\mathscr{B}} \alpha_v f_v = 0 \leftarrow$ in v^*
Let $v_0 \in \mathscr{B}$, then

$$0 = \left(\sum_{\mathscr{B}} \alpha_v f_v\right)(v_0) = \sum_{\mathscr{B}} \alpha_v f_v(v_0) = \sum_{\mathscr{B}} \alpha_v S_{vv_0} = \alpha v_0$$

Hence $\sum \alpha_v f_v = 0 \rightarrow \alpha_v = 0 \forall v \in \mathscr{B}$, so $w = 0$. $D_{\mathscr{B}}$ is therefore 1-1 as claimed. □
Warning: If V is not finite dimensional, then $D_{\mathscr{B}}$ is not onto, i.e., $\mathscr{B}^*$ does not span V^*. $\left(|V^*| = |F|^{|\mathscr{B}|}\right.$ and $|F| = |V|$ by UPVS if F is infinite$\left.\right)$
Note: $D_{\mathscr{B}} : V \rightarrow V^*$ depends on the choice of basis $\mathscr{B}$.

Definition 29.1 (Linear Functionals) — If V is a vector space over F, elements in $V^* = L(V, F)$ are called LINEAR FUNCTIONALS.

Fact 29.1. If S is a linearly indep. set in a vector space over F (even infinite) then S is part of a basis for V, i.e., the Extension Theorem holds (This needs the Axiom of Choice).

Example 29.2

V a vector space over F. Then followings are linear functionals

1. If $0 \neq v \in V$, then $\{v\}$ extend to a basis $\mathscr{B}$ for V and $\mathscr{B}^*$ satisfies $\mathscr{B}^*$ is linearly indep.
$$f_v(x) = S_{vx} \forall x \in \mathscr{B}$$
Let $w = \sum_{x \in \mathscr{B}} \alpha_x x, \alpha_x = 0$ almost all $x \in \mathscr{B}$. Then $f_x(w) = \alpha_x \in F \forall x \in \mathscr{B}$, $w = \sum f_x(w) x$

2. $\pi_i : F^n \to F$ by $(\alpha_1, \ldots, \alpha_n) \mapsto \alpha_i \forall i$

3. Let $\text{Int} : C[\alpha, \beta] \to \mathbb{R}, \alpha < \beta$ be given by
$$\text{Int } f \mapsto \int_\alpha^\beta f$$

4. trace: $M_n F \to F$ by
$$A \mapsto \sum_{i=1}^n A_{ii}$$
The sum of the diagonal entries of A called the TRACE of A.

   We can iterate our constructions as follows:

   Let $\mathscr{C}$ be a basis for $V^* = L(V, F)$ a vector space over F, where V is a vector space over F. Then
$$D_{\mathscr{C}} : V^* \to (V^*)^* := V^{**}$$
V^{**} is called the DOUBLE DUAL of V, is induced by
$$f_0 \in \mathscr{C} \mapsto G_{f_0} \in \mathscr{C}^*$$
the coordinate function on f_0, i.e.,
$$\sum_{\mathscr{C}} \alpha_f f \mapsto \sum_{\mathscr{C}^*} \alpha_f G_f$$
with
$$G_{f_0}(f) = \delta_{tf_0} = \begin{cases} 1 \text{ if } f = f_0 \forall f, f_0 \in \mathscr{C} \\ 0 \text{ if } f \neq f_0 \end{cases}$$
So we have
$$V \overset{\mathscr{D}_{\mathscr{B}}}{\to} V^* \overset{\mathscr{D}_{\mathscr{C}}}{\to} V^{**}$$
and the composition is a monomorphism.

<u>Wonderful Result</u>: $\exists$ a monomorphism
$$L : V \to V^{**}$$
INDEPENDENT OF CHOICE OF BASES. We know want to show this:

For each $v \in V$ define the following linear functionals on V^*
$$L_v : V^* \to F \text{ by } L_v(f) := f(v)$$
EVALUATION at v.

Check. $L_v : V^* \to F$ is linear, i.e., $L_v \in V^{**} = (V^*)^*$:

$$\begin{aligned} L_v(\alpha f + g) &= (\alpha f + g)(v) = \alpha f(v) + g(v) \\ &= \alpha L_v f + L_v g \end{aligned}$$

$\forall t, g \in V^* \forall \alpha \in F$ as needed. Now define

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

i.e., $L(v) = L_v$

Claim 29.2. L is linear.

$\forall f \in V^*$, $v, v' \in V, \alpha \in F$, we have

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

as needed.

Claim 29.3. $L : V \to V^{**}$ is monic.

Suppose $v \neq 0$. By Example TBA, $\exists f \in V^* \ni L_v(f) = f(v) \neq 0$. As L is linear, L is a monomorphism. Hence

$$L : V \to V^{**}$$

is a NATURAL or CANONICAL MONOMORPHISM, i.e., no basis is needed to define it. We now assume that V is a finite dimensional vector space over F, let

$$\begin{aligned} \mathscr{B} &= \{v_1, \ldots, v_n\} \text{ be a basis for } V \\ \mathscr{B}^* &= \{f_1, \ldots, f_n\} \subset V^* \text{ defined by } f_i(v_j) = \delta_{ij} \forall i, j \end{aligned}$$

i.e., the f_i are the coordinate functions relative to $\mathscr{B}$. Then, as before, we have a monomorphism

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

But we also have

$$\dim V^* = \dim L(V, F) = \dim V \dim F = \dim V$$

by the Matrix Theory Theorem, so $D_{\mathscr{B}}$ is an isomorphism by the Isomorphism Theorem with $\mathscr{B}^*$ a basis for V^* called the DUAL BASIS of $\mathscr{B}$. We also have

$$V \cong V^* \cong V^{**}, \text{ so } V \cong V^{**}$$

and

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

with

$$L_{v_i} := L_{f_i}, f_i \in \mathscr{B}^*$$

$$L_{f_i}(f_j) = L_{v_i}(f_j) = f_j(v_i) = \delta_{ij}$$

So $\mathscr{B}^{**}$ is the DUAL BASIS of $\mathscr{B}^*$. We also now $L : V \to V^{**}$ is now a natural isomorphism by the Isomorphism Theorem and even better that

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

EVALUATION at v. So when V is a finite dimensional vector space over F, we can and do identify L_v and v $\forall v \in V$.
Any $v \in V$ is determined by the $t \in V^*$ and every $f \in V^*$ is determined by the $L_v \in V^{\times\times}$ and

$$f(v) = L_v(f)$$

So now we have: if V is a finite dimensional vector space over F

$$\begin{aligned}\mathscr{B} &= \{v_1, \dots, v_n\} \text{ a basis for } V \\ \mathscr{B}^* &= \{f_1, \dots, f_n\} : \{f_{v_1}, \dots, f_{v_n}\} \text{ the dual basis of } \mathscr{B} \\ \mathscr{B}^{**} &= \left\{L_{f_{v_1}}, \dots, L_{f_{v_n}}\right\} = \{Lv_1, \dots, Lv_n\} \text{ the dual basis of } \mathscr{B}^*\end{aligned}$$

i.e.,

$$\begin{aligned}f_i &= f_{v_i} \\ L_{f_{v_i}} &= L_{v_i}\end{aligned}$$

and these satisfy

$$f_|(v_i) = tv_j(v_i) = \delta_{ij} = L_{f_{v_i}}(v_j) = L_{v_i}(f_|)$$

If $v \in V$, then

$$\begin{aligned}v &= \alpha_1 v_1 + \ldots + \alpha_n v_n \text{ unique } \alpha_1, \dots, \alpha_n \in F \\ f_j(v) &= f_j(\alpha_1 v_1 + \ldots + \alpha_n v_n) \\ &= \alpha_j\end{aligned}$$

So

$$\boxed{v = \sum_{i=1}^n f_i(v) v_i}$$

where $f_i(v)$ is the coordinate function relative to $\mathscr{B}$ and if $f \in V^*$, then

$$f = \beta_1 f_1 + \ldots + \beta_n f_n \text{ unique } \beta_1, \dots, \beta_n \in F$$

As

$$\begin{aligned}L_{v_1}(f) &= (\beta_1 f_1 + \ldots + \beta_n f_n)(v_j) \\ &= \beta_1 f_1(v_1) + \ldots + \beta_n f_n(v_j) = \beta_|\end{aligned}$$

And

$$\begin{aligned}f &= \beta_1 f_1 + \ldots + \beta_n f_n \\ &= L_{v_1}(f) f_1 + \ldots + L_{v_n}(f) f_n \\ &= f(v_1) f_1 + \ldots + f(v_n) f_n\end{aligned}$$

So,

$$\boxed{f = \sum f(v_i) f_i}$$

where $f(v_i)$ is the coordinate function.

§29.2 The Transpose

Let V, W be vector space over F, $T : V \to W$ linear if $g \in W^* = L(W, F)$, i.e., $g : W \to F$ linear, then the composition

$$V \xrightarrow{T} W \xrightarrow{g}$$

is a linear functional, i.e., $g \circ T \in V^*$.

Definition 29.3 (Transpose) — Let V, W be vector space over $F, T : V \to W$ linear. Define the transpose of T by

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

i.e.,

$$T^\top g := g \circ T \quad \forall g \in W^*$$

i.e.,

$$\begin{array}{ccc} \mathrm{V} & \xrightarrow{\mathrm{T}} & \mathrm{W} \\ & \searrow & \downarrow \mathrm{g} \quad \text{commutes} \\ T^t g := g \circ T & & \mathrm{p} \end{array}$$

So

$$V \xrightarrow{T} W$$

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

Claim 29.4. $T^\top : W^* \to V^*$ is linear if $g, g' \in W^*, \alpha \in F$, then

$$T^\top(\alpha g + g') = (\alpha g + g') \circ T = \alpha g T + g' T = \alpha T^\top g + T^\top g'$$

$T^\top$ is called the transpose because of the followings

Theorem 29.4

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

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

Proof. Let

$$\begin{aligned} \mathscr{B} = \{v_1, \dots, v_n\}, \quad \mathscr{B}^* &= \{f_1, \dots, f_n\} \\ \mathscr{C} = \{w_1, \dots, w_m\}, \quad \mathscr{C}^* &= \{g_1, \dots, g_m\} \end{aligned}$$

with $\mathscr{B}^*, \mathscr{C}^*$ the ordered dual bases of ordered bases $\mathscr{B}, \mathscr{C}$ of V, W respectively.
Let

$$[T]_{\mathscr{B},\mathscr{C}} = (\alpha_{ij}) \text{ and } [T^\top]_{\mathscr{C}^*,\mathscr{B}^*} = (\beta_{ij})$$

i.e.,

$$T_{v_k} = \sum_{i=1}^{m} \alpha_{ik} w_i \in W, \quad k = 1, \dots, n$$
$$T^\top g_j = \sum_{i=1}^{n} \beta_{ij} f_i \in V^*, \quad j = 1, \dots, m$$

Then computation gives

$$\begin{aligned}\left(T^\top g_j\right)(v_k) &= g_j(T_{v_k}) = g_j\left(\sum_{i=1}^{m} \alpha_{ik} w_i\right) \\ &= \sum_{i=1}^{m} \alpha_{ik} g_j(w_i) = \sum_{i=1}^{m} \alpha_{ik} \delta_{ij} = \alpha_{jk}\end{aligned}$$

and

$$\begin{aligned}\left(T^\top g\right)(v_k) &= \left(\sum_{i=1}^{n} \beta_{ij} f_i\right)(v_k) = \sum_{i=1}^{n} \beta_{ij} f_i(v_k) \\ &= \sum_{i=1}^{n} \beta_{ij} \delta_{ik} = \beta_{kj}\end{aligned}$$

Hence, $\alpha_{jk} = \beta_{kj} \forall j, k$ as needed. □

Definition 29.5 (Annihilator) — Let V be a vector space over $F, \emptyset \neq S \subset V$ a subset. The set

$$S^\circ := \left\{f \in V^* | f\big|_S = 0\right\} = \{f \in V^* | f(s) = 0 \forall s \in S\}$$

is called the annihilator of S.

Question 29.1. If V is an inner product space over F, can you find something analogous?

Claim 29.5. $S^\circ \subset V^*$ is a subspaces (even if S is not).

Proof. Let $f, g \in S^\circ, \alpha \in F$. To show $(\alpha f + g)\big|_S = 0$, let $s \in S$, then

$$(\alpha f + g)(s) = \alpha f(s) + g(s) = 0$$

so $\alpha f + g \in S^\circ$. □

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

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

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

Proposition 29.6

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

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

Question 29.2. If V is a finite dimensional inner product space over F, can you find something similar?

Proof. Let $\{v_1, \ldots, v_k\}$ be a basis for W. Extend it to $\mathscr{B} = \{v_1, \ldots, v_n\}$ a basis for V. Let $\mathscr{B}^* = \{f_1, \ldots, f_n\}$ be the dual basis of $\mathscr{B}$, i.e.,

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

Claim 29.6. $\mathscr{C} = \{f_{k+1}, \ldots, f_n\}$ is a basis for W°. Let $f \in W^\circ$. Then $\exists \beta_1, \ldots, \beta_n \in F \ni$

$$f = \sum_{i=1}^{n} \beta_i f_i = \sum_{i=1}^{n} \underbrace{f(v_i)}_{\beta_i} f_i = \sum_{i=1}^{k+1} f(v_i) f_i \in \text{Span}\, \mathscr{C}$$

As $\mathscr{C} \subset \mathscr{B}^*$ and $\mathscr{B}^*$ is linearly indep., so is $\mathscr{C}$. This proves the claim and the result follows. □

Corollary 29.7

Let V be a finite dimensional vector space over F, $W \subset V$ a subspace. Identifying V and V^{**} via $v \leftrightarrow L_v$, we have

$$W = (W^\circ)^\circ := W^{\circ\circ}$$

If V is a inner product space over F, can you find something similar?

Proof. We have $W^\circ \subset V^*$ and $W^{\circ\circ} \subset V^{**} = V$ are subspaces and by the last proposition, we have

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

If $w \in W$, then

$$L_w f = f(w) = 0, \quad \forall f \in W^\circ$$

So

$$w = L_w \in W^{\circ\circ}$$

i.e., $W \subset W^{\circ\circ}$ is a subspace. As $\dim W = \dim W^{\circ\circ}, W = W^{\circ\circ}$. □

Theorem 29.8

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. We have $\dim W = \dim W^*$

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

by the previous proposition and the Dimension Theorem. By observation,

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

Hence,

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

Application: Let $A \in F^{m\times n}$. The row (respectively column) RANK of A is the dimension of the subspace spanned by the rows (respectively column of A viewed as vectors in F^m (respectively $F^{n\times 1}$).
Using the theorems and our previous computation, we have

Claim 29.7. row rank A = col rank A.

III

115B Lectures

§30 Lec 1: Mar 29, 2021

§30.1 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 30.1 (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 30.2

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

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

+	0	1
0	0	1
1	1	0

•	0	1
0	0	0
1	0	1

Fact 30.1. Let $p > 0$ be a prime number in $\mathbb{Z}$. Then $\exists$ a field $\mathbb{F}_{p^n}$ having p^n elements write $|\mathbb{F}_{p^n}| = p^n \quad \forall n \in \mathbb{Z}^+$.

Definition 30.3 (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 30.4 (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
$$F[t] := \{\text{polys coeffs in } F\}$$
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. $\mathbb{Q} := \left\{\frac{a}{b} | a, b \in \mathbb{Z},\, b \neq 0\right\} < \mathbb{C}$ ($<$ means $\subset$ and $\neq$) with usual $+, \cdot$ of fractions. (when does $\frac{a}{b} = \frac{c}{d}$?)

5. If F is a field
$$F(t) := \left\{\frac{f}{g} | f, g \in F[t],\, g \neq 0\right\} \text{ (rational function)}$$
with usual $+, \cdot$ of fractions is a field.

Example 30.5 (Cont'd from above) 6. $\mathbb{Q}[\sqrt{-1}] := \left\{\alpha + \beta\sqrt{-1} \in \mathbb{C} | \alpha, \beta \in \mathbb{Q}\right\} < \mathbb{C}$. Then $\mathbb{Q}[\sqrt{-1}]$ is a field and

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

where $\mathbb{Z}[\sqrt{-1}] := \left\{\alpha + \beta\sqrt{-1} \in \mathbb{C},\, \alpha, \beta \in \mathbb{Z}\right\} < \mathbb{C}$. 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

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

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

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

for all $x \in I$.

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

Definition 30.6 (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 30.7. 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,...

§31 Lec 2: Mar 31, 2021

§31.1 Vector Spaces (Cont'd)

Example 31.1

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. $F^n = F^{1\times n} := \{(\alpha_1, \ldots, \alpha_n) \,|\, \alpha_i \in F\}$
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 31.2 (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 $\mathbb{Q}$.

§31.2 Subspaces

Definition 31.3 (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 31.4

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 31.5

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 $\mathbb{Z}$.

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

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

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

§31.3 Direct Sums

Problem 31.1. Can you break down an object into simpler pieces? If yes can you do it uniquely?

Example 31.6

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

Example 31.7

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

$$V = E_T(\lambda_1) \perp \ldots \perp E_T(\lambda_r) \qquad (*)$$

$E_T(\lambda_i) := \{v \in V | \, Tv = \lambda_i v\} \neq 0$ eigenspace of λ_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 31.2.

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 31.8 (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

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

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 31.9 (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 $I = \{1, \dots, n\}$, 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.

§32 Lec 3: Apr 2, 2021

§32.1 Direct Sums (Cont'd)

Definition 32.1 (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 32.2

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. □

<u>Warning</u>: 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 32.3

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. □

<u>Notation</u>: 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 32.4

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 32.5 (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 32.6 (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 32.7

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 $F, W_1, \dots, W_n \subseteq V$ 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 $x \in V$,

$$x = \sum_{v\in\mathscr{B}} \alpha_v v$$

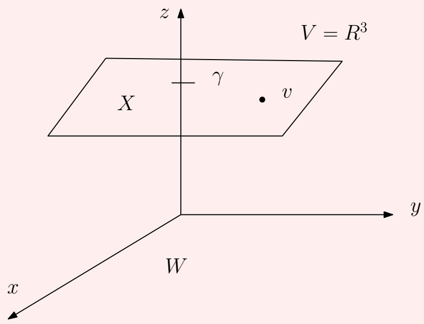
to mean α_v is the unique complement of x on v and hence $\alpha_v = 0$ for almost all $v \in \mathscr{B}$.

§32.2 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 32.8

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



So

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

Note: 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 conclusion $W + v = W + v' \iff v - v' \in W$. We can also write $W + v$ as $v + W$.

§33 Lec 4: Apr 5, 2021

§33.1 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

$$(W+v) + (W+v') := W + (v+v')$$
$$\alpha \cdot (W+v) := W + \alpha v$$

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^{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

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

Note: $\dim V = 3$, $\dim W = 2$, $\dim S = 1$ 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.
Construction: Let V be a vector space over F, $W \subseteq V$ a subspace. Define $\equiv \mod W$ called congruent mod W on V 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

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

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 33.1
$\overline{0}_V = W + 0_V = W$.

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

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

a collection of subsets of V.

§34 Lec 5: Apr 7, 2021

§34.1 Quotient Spaces (Cont'd)

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

$$x \equiv y \mod W \tag{+}$$
$$z \equiv v \mod W$$

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

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

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

$$\dim V = \dim \ker^{-}$$
$$\dim V = \dim W + \dim V/W$$
$$\dim V/W = \dim V - \dim W$$

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

Proposition 34.1

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. □

§34.2 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 34.2

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 34.3

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 $W = \text{Span}(\mathscr{C}) = \text{Span}\,(f(\mathscr{B}))$. ☐

§35 Lec 6: Apr 9, 2021

§35.1 Linear Transformation (Cont'd)

Theorem 35.1
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 35.1. 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 $\text{im } T$. 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 35.2.** Note that the proof is essentially the same as the proof of the Dimension Theorem.

Corollary 35.3 (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 35.4

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

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

$\forall x, y \in V$ 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

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

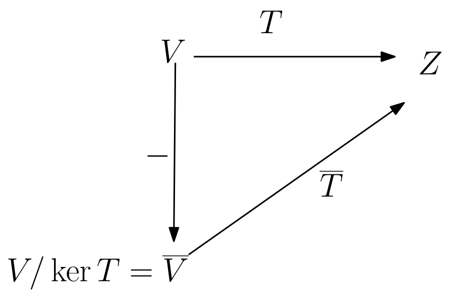
as needed. Therefore,

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

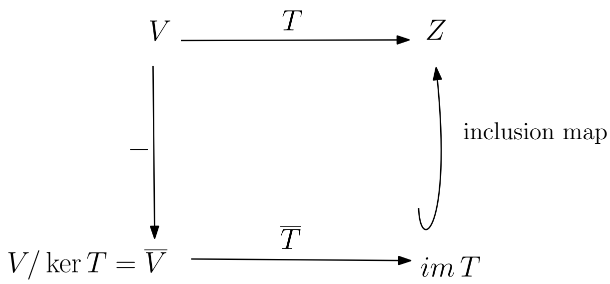
is a monomorphism, so induces an isomorphism onto im $\overline{T}$ and we recall im $\overline{T}$ = im T, so

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

and we have a commutative diagram



This can also be written as



Consequence: 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.

§35.2 Projections

Motivation: 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 35.5 (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 35.6

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

1. $\ker T$ and $\operatorname{im} T$ 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

$$E_T(\lambda) := \{v \in V \mid Tv = \lambda v\}$$

is T-invariant as $T\big|_{E_T(\lambda)} = \lambda 1_{E_T(\lambda)}$

Remark 35.7. 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 35.8

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.

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

Goal: 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.
Recall: 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 similarity, 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 35.9 (Projection) — Let V be a vector space over F, $P : V \to V$ linear. We call P a projection if $P^2 = P \circ P = P$.

Example 35.10 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

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

Lemma 35.11
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

$$v = Pv = 0$$

□

§36 Lec 7: Apr 12, 2021

§36.1 Projection (Cont'd)

Lemma 36.1

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 36.2. 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 36.3

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 36.4. 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

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

Construction: 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 $W_i = \text{im } P_{W_i}$ 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, $\text{im } P_{W_i} = W_i$.

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

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

i.e., $1_V T = T1_V$. 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 36.5.** Compare this to the case that T is diagonalizable and the W_i are the eigenspaces.

Question 36.1. 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 36.1. 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

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

§36.2 Dual Spaces

Question 36.2. 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 36.6

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

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

If $v = (\alpha_1, \dots, \alpha_n)$ 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 36.7

$\mathscr{B} = \mathscr{E}_n$ 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 $v = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n = V$. Then

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

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 36.8

Let V be a finite dimensional inner product space over $\mathbb{R}$, $\mathscr{B} = \{v_1, \ldots, v_n\}$ 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 36.9 (Dual Space) — Let V be a vector space over F. A linear transformation $f : V \to F$ is called a <u>linear functional</u>. Set

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

is called the <u>dual space</u> of V.

Proposition 36.10

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 36.11

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 $fv_0 \in \mathscr{B}^* := \{fv | v \in \mathscr{B}\}$.

§37 Lec 8: Apr 14, 2021

§37.1 Dual Spaces (Cont'd)

Example 37.1 (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 γ

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

Lemma 37.2

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

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

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

Corollary 37.3

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. $\square$

Remark 37.4.
1. If $V = \mathbb{R}^\infty_f := \{(\alpha_1, \alpha_2, \dots) \, | \alpha_i \in \mathbb{R} \text{ almost all } 0\}$, 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 37.5

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 $f, g \in V^*$, $\alpha \in F$

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

Theorem 37.6

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. $\square$

Corollary 37.7

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 $\mathscr{B} = \{v_1, \ldots, v_n\}$. Then

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

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

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

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

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

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 37.8 (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 37.9. Many people write $\langle v, f \rangle$ for $f(v)$ in the above even though $f \notin v$.

§38 Lec 9: Apr 16, 2021

§38.1 Dual Spaces (Cont'd)

Lemma 38.1

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 38.2

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

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

Claim 38.1. $\mathscr{C} := \{f_{k+1}, \ldots, f_n\}$ is a basis for S°.

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 38.3

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

$$\dim V = \dim S + \dim S^{\circ}$$
$$\dim V^* = \dim S^{\circ} + \dim S^{\circ\circ}$$

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

Remark 38.4. If V is an inner product space over $\mathbb{R}$, compare all this to $\emptyset \neq S \subseteq V$ a subset and $S^{\perp}, S^{\perp\perp}$.

§38.2 The Transpose

<u>Construction</u>: 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 38.5

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 38.6

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 38.7 (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 38.8

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 38.9

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

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

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

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

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

So

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

So we just proved...

Theorem 38.10

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

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

Definition 38.11 (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 38.12

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.

§38.3 Polynomials

Definition 38.13 (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 38.14

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

Definition 38.16 (Polynomial Degree and Leading Coefficient) — Let

$$0 \neq f = at^n + a_{n-1}t^{n-1} + \ldots + a_1t + a_0 \in F[t]$$

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

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

So

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

§39 Lec 10: Apr 19, 2021

§39.1 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 39.1 (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 39.2. If a gcd of f, g exists, then it is unique.

Remark 39.3. If $d = 1$ is a gcd of $f, g \in F[t]$, we say that f, g are relatively bear.

Remark 39.4. Compare the above with analogous in $\mathbb{Z}$.

Theorem 39.5

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 39.6 (Irreducible Polynomial) — $f \in F[t] \setminus F$ is called irreducible 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 39.7

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

Remark 39.8. 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 39.9 (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. □

Remark 39.10. In $\mathbb{Z}$ the analog of an irreducible element is called a prime element.

Remark 39.11. 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 ± 1.

Theorem 39.12 (Fundamental Theorem of Arithmetic (Polynomial Case))

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

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

unique up to order.

Proof. (Sketch) Existence: We induct on $n = \deg g \geq 1$. If g is irreducible, $a, (\text{lead } g)^{-1}g$, $a =$ 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 39.13. 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.

§40 Lec 11: Apr 21, 2021

§40.1 Minimal Polynomials

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

Remark 40.1. 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 40.2

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

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

- Left as exercise.

□

Lemma 40.3

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 $\exists a_0, \ldots, a_{n^2} \in F$ 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. □

Theorem 40.4

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 40.1. 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

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

Corollary 40.5

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. □

<u>*Note*</u>: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 $q(T) \neq 0$.

<u>Goal</u>: 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 40.6 (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 40.7.

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 $\det : \mathbb{M}_n F[t] \to F[t]$

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

for each $P \in GL_n F$

Proposition 40.8

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 λ be a root of q_T. To show λ is an eigenvalue of T, i.e., a root of f_T. As λ 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

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

So $0 \neq w \in E_T(\lambda)$ and λ is an eigenvalue of T.

Conversely, suppose λ 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 $T|_{Fv}$. □

§40.2 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

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

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

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

Check that $\bar{e}_T$ 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

$$f \cdot v := f(T)v$$
$$q_T(T) = 0$$

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.

§41 Lec 12: Apr 23, 2021

§41.1 Triangularizability

Proposition 41.1

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 41.2 (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 $\mathscr{B} = \{v_1, \dots, v_n\}$ in (*) and $\mathscr{C} = \{v_n, v_{n-1}, \dots, v_1\}$, 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 41.3. 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

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

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

Definition 41.4 (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 41.5

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 41.6

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 λ be a root of q_T in F (q_T splits in $F[t]$). Then λ 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 $A = \left[\overline{T}\right]_{\mathscr{C}}$ 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, $\exists A_{nj} \in F$, $1 \leq j \leq n-1$ satisfying

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

So

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

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

$$Tv_n = \lambda v_n$$

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

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

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

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

So

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

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 $F[t]$. □

Corollary 41.7

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 41.8 (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 41.9 (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 41.10 (Fundamental Theorem of Algebra)

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

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

§42 Lec 13: Apr 26, 2021

§42.1 Triangularizability (Cont'd)

Remark 42.1. 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 42.2. 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.

§42.2 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

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

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

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

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 42.3

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

$$f(T)Tv = Tf(T)v = 0$$

so this is immediate. □

Lemma 42.4

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. □

§43 Lec 14: Apr 28, 2021

§43.1 Primary Decomposition (Cont'd)

Proposition 43.1

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 43.2 (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:

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

Note:

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

Claim 43.1. Let $2 \leq i \leq r$. 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 43.2. $W_1 = \ker p_1^{e_1}(T)$ 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. □

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

§44 Lec 15: Apr 30, 2021

§44.1 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 44.1

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

$$q_T = \prod_{i=1}^{r} (t - \lambda_i), \quad \lambda_1, \dots, \lambda_r \text{ the distinct roots of } q_T$$

Proof. " $\Longleftarrow$ " $q_T = \prod_{i=1}^{r} (t - \lambda_i)$, $\lambda_1, \dots, \lambda_r$ the distinct roots of q_T. Let $V_i = \ker(T - \lambda_i 1_V) = E_T(\lambda_i)$, $i = 1, \dots, r$. Then by the Primary Decomposition Theorem,

$$V = V_1 \oplus \ldots \oplus V_r$$

SO T is diagonalizable.
" $\Longrightarrow$ " Let $\mathscr{B} = \{v_1, \dots, v_n\}$ be an ordered basis for V consisting of eigenvectors of T with λ_i the eigenvalue of v_i and ordered s.t.

$$\lambda_1, \dots, \lambda_r \text{ are the distinct eigenvalues of } T$$

For each j, $1 \le j \le n$, we have

$$(T - \lambda_i 1_V)\, v_j = T v_j - \lambda_i v_j = (\lambda_j - \lambda_i)\, v_j, \quad j = 1, \dots, n$$

So

$$\prod_{i=1}^{r} (T - \lambda_i 1_V)\, v_j = 0 \quad \text{for } j = 1, \dots, n$$

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

$$q_T | \prod_{i=1}^{r} (t - \lambda_i) \in F[t]$$

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

$$t - \lambda_i | q_T, \qquad i = 1, \dots, r$$

using f_T and q_T have the same roots. Therefore,

$$q_T = \prod_{i=1}^{r} (t - \lambda_i) \in F[t]$$ □

§44.2 Jordan Blocks

Definition 44.2 (Jordan Block Matrix) — $J \in \mathbb{M}_n F$ is called a Jordan block matrix of eigenvalue λ 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 44.3 (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 44.4

$J = J_n(0)$ 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 44.5 (Cont'd from above)

In any case, we have

$$\left.\begin{aligned}\dim \text{im } N^r &= n-r\\ \dim\ker N^r &= r\end{aligned}\right\}\text{if } r \le n$$

$$\left.\begin{aligned}\dim \text{im } N^r &= 0\\ \dim \text{im } \ker N^r &= n\end{aligned}\right\}\text{if } r > n$$

Lemma 44.6

Let $J = J_n(\lambda) \in \mathbb{M}_n F$. Then

1. λ 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 44.7. $J_n(\lambda)$ has only a line as an eigenspace, so among triangulariazable operator away from being diagonalizable when $n \ge 1$.

Proposition 44.8

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

$$E_\lambda(\lambda) = Fv_n$$

$v_1, \dots, v_{n-1}$ are not eigenvectors, but do satisfy

$$\begin{aligned} (T - \lambda 1_V)v_i &= v_{i+1} \qquad i = 1, \dots, n-1 \\ (T - \lambda 1_V)^{n-i} v_i &= v_n \qquad \text{, an eigenvector} \end{aligned}$$

So we can compute $v_1, \dots, v_{n-1}$ from the eigenvalue v_n.

§45 Lec 16: May 3, 2021

§45.1 Jordan Blocks (Cont'd)

Definition 45.1 (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 λ 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 λ.

<u>*Note*</u>: We should really write

$$g_n(v_n, \lambda, v_1, \ldots, v_{n-1})$$

Lemma 45.2

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). □

Definition 45.3 (Jordan Canonical Form) — $A \in \mathbb{M}_n F$ is called a matrix in Jordan canonical form (JCF) if A has the block form

$$A = \begin{pmatrix} J_{r_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{r_m}(\lambda_m) \end{pmatrix}$$

$\lambda_1, \dots, \lambda_m$ not necessarily distinct.

Definition 45.4 (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

$$g_{r_1}(v_{1,r_1}, \lambda_1) \cup \dots \cup g_{r_m}(v_{m,r_m}, \lambda_m) \qquad (\star)$$

where $g_{r_j}(v_{j,r_j}, \lambda_j)$ is an ordered sequence of generalized eigenvectors of T relative to λ_j ending at eigenvector v_{j,r_j}. The $\lambda_1, \dots, \lambda_m$ need not be distinct.

Proposition 45.5

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. □

Conclusion: 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:

$$\lambda_1, \dots, \lambda_m \text{ the distinct eigenvalues of } T$$

$$\begin{aligned} \mathscr{B} = & \, g_{r_{11}}(v_{11}, \lambda_1) \cup \dots \cup g_{r_{1,n_1}}(v_{1,n_1}, \lambda_1) \\ & \cup g_{r_{21}}(v_{21}, \lambda_2) \cup \dots \cup g_{r_{2,n_2}}(v_{2,n_2}, \lambda_2) \\ & \vdots \\ & \cup g_{r_{m,1}}(v_{m,1}, \lambda_m) \cup \dots \cup g_{r_{m,n_m}}(v_{m,r_m}, \lambda_m) \end{aligned}$$

with

$$r_{i1} \le r_{i2} \le \dots \le r_i n_i, \qquad 1 \le i \le m$$

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

$$\begin{aligned} q_T &= \prod_i \text{lcm}\left((t-\lambda_i)^{r_{ij}} | j = 1, \dots, n_i\right) \\ &= \prod_i (t - \lambda_i)^{r_{in_i}} \end{aligned}$$

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 $r_{i1} \le \dots \le r_{in_i}$,

$$\begin{aligned} (T - \lambda_i 1_V)^{r_{in_i}}\big|_{W_{ij}} &= 0, \qquad 1 \le j \le n_i \\ (T - \lambda_i 1_V)^{r_{in_i}-1}\big|_{W_{ij}} &\ne 0 \end{aligned}$$

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.

§45.2 Jordan Canonical Form

Theorem 45.6

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

$$W_i = \ker (T - \lambda_i 1_V)^{r_i}, \quad i = 1, \dots, m$$

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. □

§46 Lec 17: May 5, 2021

§46.1 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 λ 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 $(T|_{\text{im } T})^r = 0$, so

  $$T\big|_{\text{im } T} \text{ is also nilpotent}$$

  and

  $$q_{T|_{\text{im } T}} | q_T \in F[t]$$

  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

  $$q_{T|_{\text{im } T}} = t^s, \quad \text{for some } s \leq r$$

  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 $w_i \in W_i \subseteq \text{ im } T$,

$$\exists v_i \in V \ni Tv_i = w_i, \quad i = 1, \ldots, m$$

So we also have

$$T^{r_i+1}v_i = T^{r_i}Tv_i = T^{r_i}w_i = 0$$

and

$$T^{r_i}v_i = T^{r_i-1}Tv_i = T^{r_i-1}w_i \neq 0$$

Therefore, $v_i, Tv_i, \ldots, T^{r_i}v_i$ is an ordered sequence of generalized eigenvalues for T in V, and, in particular, linearly independent. For each $i = 1, \ldots, m$, let

$$V_i = \text{Span}\left\{v_i, Tv_i, \ldots, T^{r_i}v_i\right\}$$

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, $i = 1, \ldots, m$.

Note: 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 46.1. $V' = V_1 \oplus \ldots \oplus V_m$

In particular,

$$\mathscr{B}_0 = \{v_1, Tv_1, \ldots, T^{r_i}v_1, \ldots, v_m, Tv_m, \ldots, T^{r_m}v_m\}$$

is a basis for V'. □

§47 Lec 18: May 7, 2021

§47.1 Jordan Canonical Form (Cont'd)

Proof. (Cont'd) Suppose $u_i \in V_i$, $i = 1, \dots, m$ satisfies

$$u_1 + \dots + u_m = 0 \tag{1}$$

To show $u_i = 0$, $i = 1, \dots, m$. 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

$$f_i(T)v_1 + \dots + f_m(T)v_m = 0 \tag{2}$$

Since $Tf(T) = f(T)T\ \forall f \in F[t]$ and

$$w_i = Tv_i \qquad i = 1, \dots, m$$

taking T of (2) yields

$$f_1(T)w_1 + \dots + f_m(T)w_m = 0$$

As the T-invariant W_i satisfying

$$W_1 + \dots + W_m = W_1 \oplus \dots \oplus W_m \tag{*}$$

We have

$$f_i(T)w_i = 0, \qquad i = 1, \dots, m$$

Hence

$$f_i(T) = 0 \text{ on } W_i, \qquad i = 1, \dots, m$$

Thus

$$t^{r_i} = q_{T\big|_{W_i}} \big| f_i \in F[t], \quad i = 1, \dots, m$$

In particular, since $r_i \geq 1\ \forall i$, we can write

$$f_i = tg_i \in F[t], \qquad i = 1, \dots, m$$
$$\deg g_i < \deg f_i, \qquad i = 1, \dots, m \text{ if } f_i \neq 0$$

Since

$$f_i(T) = Tg_i(T) = g_i(T)T$$

and

$$w_i = Tv_i, \qquad i = 1, \dots, m$$

(2) now becomes

$$g_1(T)w_1 + \dots + g_m(T)w_m = 0 \tag{3}$$

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 $i = 1, \dots, m$ by the definition of W_i. Therefore, for each i, $i = 1, \dots, m$

$$t^{r_i} = q_{T\big|_{W_i}} \big| g_i \in F[t]$$

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

$$w_i = T v_i \in W_i, \qquad i = 1, \ldots, m$$

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 47.1. $V = \ker T + V'$

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, $1 \le i \le s$.
Using claim 1, we have

$$\begin{aligned} V &= V' \oplus \text{Span } \{u_{j_1}, \dots, u_{j_s}\} \\ &= V_1 \oplus \dots \oplus V_m \oplus F u_{j_1} \oplus \dots \oplus F u_{j_s} \end{aligned}$$

and $[T]_{\mathscr{B}}$ is in Jordan canonical form. This proves existence.
Note: $F u_{j_i}$ 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.
Uniqueness: 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!)** □

Corollary 47.1

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 47.2

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 47.3

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 47.4

Let F be an algebraically closed field. Then $A \sim A^\top$.

Proof. For any $B \in \mathbb{M}_n F$, $q_B = q_{B^\top}$. □

§47.2 Companion Matrix

Definition 47.5 (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

$$C(g) := \begin{pmatrix} 0 & 0 & \ldots & 0 & — & a_0 \\ 1 & 0 & & 0 & — & a_1 \\ 0 & 1 & & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots \\ & & & 0 & — & a_{n-2} \\ 0 & 0 & \ldots & 1 & — & a_{n-1} \end{pmatrix}$$

is called the companion matrix of g.

Example 47.6
$C(t^n) = J_n(0)$.

<u>*Note*</u>: If $f, g \in F[t]$ are monic, then

$$f = g \iff C(f) = C(g)$$

Lemma 47.7
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 47.8
Let $g \in F[t] \setminus F$ be monic. Then

$$q_{C(g)} = f_{C(g)} = g$$

In particular,

$$f_{C(g)}\left(C(g)\right) = 0$$

§48 Lec 19: May 10, 2021

§48.1 Companion Matrix (Cont'd)

Remark 48.1. 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 48.2 (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

$$V = \text{Span}\left\{v, Tv, \ldots, T^n v, \ldots\right\} = F[T]v$$

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 48.3

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

$$T^j v = \alpha_{j-1}T^{j-1}v + \alpha_{j-2}T^{j-2}v + \ldots + \alpha_1 Tv + \alpha_0 v \qquad (*)$$

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

$$v = F[T]v = \text{Span}\left\{v, Tv, \dots, T^{j-1}v\right\}$$

So

$$n = \dim V \leq j, \qquad \text{hence } n = j$$

This proves i).

ii) The computation proving i) shows

$$\mathscr{B} = \left\{v, Tv, \dots, T^{n-1}v\right\}$$

is an ordered basis for V. As

$$\begin{aligned}[T]_{\mathscr{B}} &= \begin{pmatrix}[Tv]_{\mathscr{B}} & [T^2v]_{\mathscr{B}} & \dots & [T^{n-2}v]_{\mathscr{B}} & [T^n v]_{\mathscr{B}}\end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & & 0 & * \\ 1 & 0 & & \vdots & * \\ 0 & 1 & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & * \end{pmatrix}\end{aligned}$$

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 48.4

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 $\exists v \in V \ni$

$$\mathscr{B} = \left\{v, Sv, \dots, S^{n-1}v\right\}$$

So v s an S-cyclic vector and

$$V = F[S]v$$

Fact 48.1. 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

$$\det D = \sum \text{sgn } \sigma D_{1\sigma(1)} \dots D_{n\sigma(n)}$$

§48.2 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

$$\begin{pmatrix} q_1 & 0 & \dots & & & & \\ 0 & q_2 & & & & & \\ \vdots & & \ddots & & & & \\ & & & q_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & & & & \end{pmatrix}$$

with $q_1|q_2|q_3|\dots|q_r$ in $F[t]$ and all monic, i.e., there exists a positive integer r satisfying $r \le \min(m,n)$ and $q_1|q_2|q_3|\dots|q_r$ 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 48.5. 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 48.6.

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

$$GL_n(F[t]) := \{A|A \text{ is invertible}\}$$

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 48.7 (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 48.8

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 48.9. The SNF derived by this algorithm is, in fact, unique. In particular, the monic polynomials $q_1|q_2|q_3|\ldots|q_r$ 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.

§49 Lec 20: May 12, 2021

§49.1 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 $F[t]\backslash F$. 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 $\mathscr{B}_i = \{v_i, Tv_i, \dots, T^{n_i-1}v_i\}$ 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 49.1. 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 $s_i \geq r_i$ for each $1 \leq i \leq m$ and $m+1, \ldots, m+k \geq 0$ with $\lambda_1, \ldots, \lambda_{m+k}$ 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

$$\begin{pmatrix} J_{r_1}(\lambda_1) & 0 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \ldots & J_{r_m}(\lambda_m) \end{pmatrix}$$

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

$$n_v \coloneqq \dim F[T]v \leq \dim V$$

and has ordered basis

$$\mathscr{B}_v \coloneqq \left\{v, Tv, \ldots, T^{n_v-1}v\right\}$$

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 $as f_T \neq 0$. 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 49.1. 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 $0 \ldots 0 \cdot q_1 \ldots q_r \in F \setminus \{0\}$ if there are any 0's on the diagonal, which is inseparable.

Lemma 49.2

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 49.3

Let V be a finite dimensional vector space over F, $T : V \to V$ linear $q_1 | \ldots | q_r$ 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 49.4

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 49.5

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.,

$$(\mathbb{M}_n F)[t] = \mathbb{M}_r (F[t])$$

Lemma 49.6

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) $P = NA + S$

§50 Lec 21: May 14, 2021

§50.1 Rational Canonical Form (Cont'd)

Recall from last lecture,

Lemma 50.1

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 α_{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

$$M \coloneqq P_{m-1} + P_{m-2} + \ldots + P_0$$

then

$$R \coloneqq P - AM$$

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 50.2

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 $B \approx A$.
" $\Longrightarrow$ " Suppose there exist $P_1, Q_1 \in GL_n F[t]$, hence each a product of elementary matrices by Lemma 49.5, satisfying

$$B = tB - B_0 = P_1AQ_1 = P_1\left(tI - A_0\right)Q_1$$

Applying Lemma 50.1, we can write

i) $P_1 = BP_2 + R$, $P_2 \in \mathbb{M}_n F[t]$, $R \in \mathbb{M}_n F$

ii) $Q_1 = Q_2B + S$, $Q_2 \in \mathbb{M}_n F[t]$, $S \in \mathbb{M}_n F$

Since $B = P_1AQ_1$, $P_1, Q_1 \in GL_n F[t]$, we also have

iii) $P_1A = BQ^{-1}$

iv) $AQ_1 = P_1^{-1}B$

Thus, we have

$$\begin{aligned} B = P_1AQ_1 &\overset{i)}{=} (BP_2 + R)AQ_1 = BP_2AQ_1 + RAQ_1 \\ &\overset{iv)}{=} BP_2P_1^{-1}B + RAQ_1 \overset{ii)}{=} BP_2P_1^{-1}B + RA(Q_2B + S) \\ &= BP_2P_1^{-1}B + RAQ_2B + RAS \end{aligned}$$

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) $B = BTB + RAS \in \mathbb{M}_nF[t]$

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 ≤ 1 and every entry of $RAS = R(tI - A_0)S$ has is zero or has deg ≤ 1.

Question 50.1. What about BTB?

Let $T = T_mt^m + T_{m-1}t^{m-1} + \ldots + T_0$ with $T_0, \ldots, T_m \in \mathbb{M}_nF$. Then

$$\begin{aligned} BTB &= (tI - B_0)\left(T_mt^m + T_{m-1}t^{m-1} + \ldots + T_0\right)(tI - B_0) \\ &= T_mt^{m+2} + \text{ lower terms in } t \end{aligned}$$

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 50.3

Let $A_0, B_0 \in \mathbb{M}_nF$, $A = tI - A_0$, $B = tI - B_0$ in $\mathbb{M}_nF[t]$. 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, $q_1|\dots|q_r$ the invariants of T, then

$$\begin{aligned} V &= \ker q_1(T) \oplus \dots \oplus \ker q_n(T) \\ q_r &= q_T \\ f_T &= q_1 \dots q_r \end{aligned}$$

<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 50.4

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 $\mathbb{M}_n K$.

§51 Lec 22: May 17, 2021

§51.1 Inner Product Spaces

Notation: $- : \mathbb{C} \to \mathbb{C}$ 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 51.1 (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 51.2 (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 51.3
If $F \subseteq \mathbb{R}$, then any sesquilinear map is linear and $V^\dagger = V^*$.

Remark 51.4. 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 λ_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 $w \to v = 0$, λ is injective hence monic. If V is finite dimensional then λ 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 Cauchy – Schwarz 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. Gram – Schmidt: 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 51.5 (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 51.6 (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 51.7. 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 51.8 (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 51.9

Let V, W be finite dimensional inner product space over F, $T : V \to W$ linear. Then the adjoint $T^* : W \to V$ exists.

§52 Lec 23: May 19, 2021

§52.1 Inner Product Spaces (Cont'd)

Corollary 52.1

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 52.2 (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 52.3. 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 52.4

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 52.5

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 52.6

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 $\langle x, y\rangle_V = \langle Tx, Ty\rangle_W$. □

§53 Lec 24: May 21, 2021

§53.1 Inner Product Spaces (Cont'd)

Definition 53.1 (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 53.2

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 53.3. 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 53.4 (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

$$U_n F := \{A \in GL_n F \mid AA^* = I\}$$

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

$$O_n F := \left\{A \in GL_n F \mid AA^\top = I\right\}$$

Remark 53.5.
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.

§53.2 Spectral Theory

Lemma 53.6

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 53.7

Let V be a finite dimensional inner product space over F, $T : V \to V$ linear. Then the following is true: If λ 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 53.8 (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 $\mathscr{B}_0 = \{v_1, \dots, v_{n-1}\}$ 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. □

§54 Lec 25: May 24, 2021

§54.1 Spectral Theory (Cont'd)

Definition 54.1 (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 54.2

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 54.3

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 54.4

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 54.5 (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 54.6. 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^{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

$$\begin{gathered} P_{E_T(\lambda_i)} = P_i = P_{E_T^*}(\overline{\lambda_i}) \qquad \forall i \\ TP_i = P_iT \quad \text{and} \quad T^*P_i = P_iT^* \qquad \forall i \\ 1_V = P_1 + \ldots + P_r \\ P_iP_j = \delta_{ij}P_i \qquad \forall i \\ T = \lambda_1 P_1 + \ldots + \lambda_r P_r \\ T^* = \overline{\lambda_1} P_1 + \ldots + \overline{\lambda_r} P_r \end{gathered}$$

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 54.7
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. $T^* = g(T)$.

§55 Lec 26: May 26, 2021

§55.1 Spectral Theory (Cont'd)

Remark 55.1. 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 55.2

Let V be an inner product space over F, $T : V \to V$ hermitian. If λ 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 55.3

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 55.4 (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. □

§55.2 Hermitian Addendum

Theorem 55.5

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 55.6

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., λ_i is an eigenvalue. □

§56 Lec 27: May 28, 2021

§56.1 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 56.1 (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 56.2

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 56.1. What does this say about the 2nd derivative test for C^2 function, $f : S \to \mathbb{R}$ at a critical point in the interior of S?

Theorem 56.3

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 $\mathscr{B}=\{v_1,\dots,v_n\}$ be an ordered orthonormal basis for V of eigenvectors of T

$$Tv_i=\lambda_i v_i,\qquad \lambda_i\ge 0\in\mathbb{R},\ \ i=1,\dots,n$$

Then $\exists\mu_i\in\mathbb{R}$, $\mu_i\ge 0$ s.t. $\lambda_i=\mu_i^2$, $i=1,\dots,n$. 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 $\lambda_i>0\forall i$, $\det B\ne 0$, so $B\in GL_nF$.
" $\impliedby$ " Let $\mathscr{B}$ be an ordered orthonormal basis for V of eigenvectors for S. Then

$$[S]_{\mathscr{B}}=\begin{pmatrix}\mu_1 & & 0\\ & \ddots & \\ 0 & & \mu_n\end{pmatrix},\quad \mu_i\ge 0\in\mathbb{R}\text{ and}$$

$$[T]_{\mathscr{B}}=\left[S^2\right]_{\mathscr{B}}=\begin{pmatrix}\mu_1^2 & & 0\\ & \ddots & \\ 0 & & \mu_n^2\end{pmatrix}$$

is diagonal. Therefore, $\mathscr{B}$ is also an orthonormal basis for V of eigenvectors of T. As $\mu_i^2\ge 0$ (>0 if S is invertible), T is non-negative (respectively positive if S is invertible). □

Theorem 56.4

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 56.5

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 56.6.** 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 λ 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 56.7

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. □

Notation:

- $F = \mathbb{R}$ or $\mathbb{C}$, $A \in F^{m\times n}$
- $A^{(i)}$ = the i^{th} column of A
- $A = \left[A^{(1)} \ \dots \ A^{(m)}\right]$
- $\langle , \rangle$ = the dot product on F^N for any $N \geq 1$
- $U_N(F) = \left\{U \in GL_N F \,|\, U^* = U^{-1}\right\}$

Definition 56.8 (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 56.9 (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

$$r = \text{rank}(A)$$

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

$$\lambda_i = \mu_i^2, \quad \mu_i > 0, \qquad 1 \le i \le r$$

So

$$\mu_1 \ge \ldots \ge \mu_r > 0$$

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. □

§57 Lec 28: Jun 2, 2021

§57.1 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

$$X^* A U = D$$

as needed. □

Definition 57.1 (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 57.2

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 57.3. An SVD of $A \in F^{m\times n}$, $F = \mathbb{R}$ or $\mathbb{C}$ may not be unique.

Corollary 57.4

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 57.5 (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

$$\mathscr{B} = \{v_1, \ldots, v_n\} \text{ for } V$$
$$\mathscr{C} = \{w_1, \ldots, w_m\} \text{ for } W$$
$$\mu_1 \geq \ldots \geq \mu_r > 0 \in \mathbb{R}$$

satisfying

$$Tv_i = \begin{cases} \mu_i w_i, & i = 1, \ldots, r \\ 0, & i > r \end{cases}$$

Conversely, suppose the above conditions are all satisfied. Then v_i is an eigenvector for T^*T with eigenvalue μ_i^2 for $i = 1, \ldots, r$ and eigenvalue 0 for $i = r+1, \ldots, n$. In particular, $\mu_1, \ldots, \mu_r$ are uniquely determined.

Proof. Left as exercise. □

Remark 57.6. 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 57.7 (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. □

§57.2 Least Squares

We give an application of SVD

Problem 57.1. 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 $W = T(X) = \text{im } T$. 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.

Solution:
$$\langle x, T^* y\rangle_X = \langle Tx, y\rangle_V,$$
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,
$$\begin{aligned} v_W = P_W v &= P_W\left(Tx - (Tx - v)\right) \\ &= P_W(Tx) - P_W(Tx - v) \\ &= Tx + 0 = Tx \end{aligned}$$
Conclusion: x is a best approximation to v via T if and only if $T^* T x = T^* v$.

Claim 57.1. 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.

§58 Lec 29: Jun 4, 2021

§58.1 Least Squares (Cont'd)

Claim 58.1. 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 $\operatorname{rank}(A) = n$, then

$$A^\dagger = (A^*A)^{-1}A^*$$

Proof. i) $\operatorname{rank}(A) = \operatorname{rank}(D) = \operatorname{rank}(D^\dagger) = \operatorname{rank}(A^\dagger)$ as X, U are invertible.

ii) **Case 1:** $A = D$, i.e., X, U are the appropriate identity matrices. Let $W = \operatorname{im} A$, $U = \ker D^\dagger D$, $W = \operatorname{span}\{e_i \in \mathscr{S}_{m,1} | D_{ii} \neq 0\}$

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

$$U = \ker D^\dagger D = \operatorname{Span}\{e_j \in \mathscr{S}_{n,1} | D_{jj} = 0\}$$

and we have

$$D^\dagger v \in \operatorname{Span}\{e_j \in \mathscr{S}_{n,1} | D_{jj} \neq 0\} = Y^\perp,$$

and $P_Y\left(D^\dagger v\right) = 0$.

$$D^\dagger v \text{ is optimal approximation to } v \text{ relative to } D$$

Case 2: $A = XDU^*$ 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 $A^\dagger = UD^\dagger X^*$

$$D\left(D^\dagger X^*v\right) \overset{\text{SVD}}{=} (X^*AU)\left(D^\dagger X^*v\right) = X^*A\left(A^\dagger v\right)$$

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 $\operatorname{rank}(A) = n$, then $(A^*A)^{-1}A^*v$ is the unique optimal best approximation to $Az = v$. □

Warning: 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 58.1

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 $W = \text{im } A$. To find the optimal approximation to $AX = Y_W$, $X = A^\dagger Y$ works. But $\text{rank}(A) = 2$ is most probable

$$X = (A^* A)^{-1} A^* Y$$

§58.2 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 58.2

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 $v = \sum_{i=1}^n \alpha_i v_i$, $\alpha_i \in F$, $i = 1, \dots, n$. Then

$$R(v) = \frac{\langle Av, v\rangle}{\|v\|^2} = \langle \sum_{i=1}^n \alpha_i \lambda_i v_i, \sum_{j=1}^n \alpha_j v_j\rangle / \|v\|^2$$
$$= \frac{\sum_{i,j=1}^n \lambda_i \alpha_i \overline{\alpha_j} \delta_{ij} \langle v_i, v_j\rangle}{\|v\|^2} = \frac{\sum_{i=1}^n \lambda_i - |\alpha_i|^2}{\|v\|^2}$$

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. □

Corollary 58.3

Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in \mathbb{M}_n F$. Then $\|A\| < \infty$. Moreover, if μ 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. □

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 58.4

Let $A \in GL_n F, F = \mathbb{R}$ or $\mathbb{C}$, μ 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

$$\mu_1 \ge \dots \ge \mu_n > 0$$

are the singular values of A,

$$\mu_n \ge \dots \ge \mu_1 > 0$$

are the singular values of A^{-1} as $(A^{-1})^* A^{-1} = (AA^*)^{-1}$. □

§59 Additional Materials: Jun 04, 2021

§59.1 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 59.1. 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 59.1

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 59.2. From the theorem,

1. If cond(A) is close to one, then a small relative error in b forces a small relative error in x.
2. If cond(A) is large, even a small relative error in x may cause a relatively large error in b.

Remark 59.3. 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 $A + SA \in GL_n F$, $F = \mathbb{R}$ or $\mathbb{C}$, it is true that

$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq \text{cond}(A) \frac{\|\delta A\|}{\|A\|}$$

One almost never computes 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.

§59.2 Mini-Max

Let $F = \mathbb{R}$ or $\mathbb{C}$, $A \in \mathbb{M}_n F$. We want a method to compute its eigenvalues if A is hermitian. Since A is hermitian, by the Spectral Theorem,

$$U^* A U = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad U \in U_n F$$

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

$$A v_i = \lambda_i v_i$$

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 $(F v_1)^\perp$ is A-invariant, $A\big|_{(F v_1)^\perp}$ has maximum eigenvalue λ_2 obtained from v_2, i.e.,

$$\max_{x \in (F v_1)^\perp} R_A(x) = \lambda_{n-1}$$

is obtained from $x = v_2$. The constraint is

$$\langle x, v_1 \rangle = 0$$

We can obtain λ_{n-1} without knowing v_1 or λ_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 59.4 (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 59.5. 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 $\dim j$.

§59.3 Uniqueness of Smith Normal Form

Consult Professor Elman's notes.