1

MATRIX ANALYSIS
MA353

Feb 7, 2019

(1) Complex Numbers

"Traditional way to think"

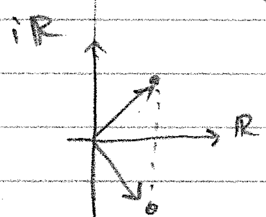
- $\mathbb{C}$ writing "$a+ib$" is equiv to writing a pair $(a,b) \in \mathbb{R}^2$, $a, b \in \mathbb{R}$
- Addition $(a,b) + (c,d) = (a+c, b+d)$
- Multiplication $(a,b)\cdot(c,d) = (ac-bd, ad+bc)$

same as $(a+ib)(c+id) = (ac-bd) + i(ad+bc)$

Think of i as $\sqrt{-1}$, a multiplication helper.

"better" multiplication than element wise multiplication, because there are properties we're familiar with... in real multiplication.

"i" ~ $(0,1)$, "1" ~ $(1,0)$, "0" ~ $(0,0)$



- Conjugate of $a+ib$ is $a-ib$ → reflecting about the real axis.

"New way to think abt complex no."

$$a+ib \sim \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = a\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{I} + b\underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{\text{"}i\text{"}}$$

real-entried matrices

$$= aI + b\mathring{I}$$

So

$$(aI + b\mathring{I})(cI + d\mathring{I}) \stackrel{?}{=} (cI + d\mathring{I})(aI + b\mathring{I}) \quad ✓$$

$$\rightarrow \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \text{ \& } \begin{pmatrix} c & -d \\ d & c \end{pmatrix}$$ commute! same as complex no. multiplication

Addition works exactly the same way as before. → adding matrices.

Taking the adjoint (conjugating) ≡ Transpose them.

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = MM^T = \begin{pmatrix} a^2+b^2 & 0 \\ 0 & a^2+b^2 \end{pmatrix} = \det(M)\cdot I$$

$$= (a^2+b^2) I$$

Similarly $\rightarrow (a+ib)(a-ib) = a^2+b^2$

Note $\boxed{A\cdot[\ \ ] = (\det A) I}$

Note, $\det A = 0 \Leftrightarrow a = b = 0$

$A\cdot[\ \ ]/\det A = I \Rightarrow [\ \ ] = A^{-1} \equiv$ "reciprocal" of complex no.

Note $A^T = A^{-1}$ here if $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$

$$\underbrace{\frac{1}{3+5i}}_{\text{number}} = \frac{3-5i}{3-5i}\,\frac{1}{3+5i} = \underbrace{\frac{1}{3^2+5^2}}_{1/\text{length}}\underbrace{(3-5i)}_{\text{new number...}}$$

We will show that there always exists a <u>classical adjoint</u> B s.t. $AB = (\det A). I$

<u>Geometric interpretation</u>

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac-bd \\ bc+ad \end{pmatrix}$$ $\rightarrow$ multiplication is some linear transformation

"$a+ib$" "$c+id$" = "$(a+ib)(c+id)$"

<u>Think</u> normalize

$$A = \frac{1}{\sqrt{a^2+b^2}}\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \rightarrow \text{is a } \boxed{\text{Unitary matrix}}$$

$\rightarrow$ preserves length, orthogonality, angles...

Can write $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \to$ rotation by θ

So complex number multiplication $\equiv$ rotate + scale by length.

↳ Euler notation $re^{i\theta}$

↳ $\boxed{re^{i\varphi}\rho e^{i\theta} = r\rho e^{i(\theta+\varphi)}} \to$ captures multiplication.

↳ Formula $\boxed{e^{i\theta} = \cos\theta + i\sin\theta}$

Why complex numbers? $\to$ Closed – completes fundamental theorem of algebra

↳ poly. w/ real coeffs. $p(x) = \sum^n a_i x_i^{i\text{th}}$

Let $z \in \mathbb{C}$ be a root $\sum a_i z_0^i = 0$

So $\left(\sum^n a_i z_0^i\right)^* = 0 = \sum^n (a_i z_0^i)^*$ $\qquad (AB)^T = B^T A^T$

$$= \sum^n (a_i)^* (z_0{}^i)^* = 0$$

$$= \sum^n (a_i)^* (z_0^*)^i = 0$$

$$= \sum^n (a_i)(z_0^*)^i = 0$$

So if z_0 is a root, then z_0^* is also a root.

So $P(x) = (x-r_1)(x-r_2)\cdots \underbrace{(x-z_1)(x-z_1^*)}_{\text{quadratic}} \cdots (x-z_n)(x-z_n^*)$

So "real" fundamental theorem of algebra $\Rightarrow$ any poly ~~e~~ factors as ... linear terms + quadratic terms

Feb 12, 2019

Review of Linear spaces & Linear Functions

- Linear Space $\equiv$ Vector Space $\to$ Ex $\mathbb{C}^n$, $\mathbb{P}_n$ (Poly. deg $\leq n$ complex coeffs)
  $\mathbb{P} \to$ polynomials no restrictions on degrees
  closed under addition & multiplication scalar
- Coordinate system $\equiv$ Basis

More examples

- $C(X, \mathbb{C})$ $\to$ continuous functions $[0,1] \mapsto \mathbb{C}$

$\mathbb{P}_n$ subspace of $\mathbb{P}$ subspace of $\mathbb{C}$

- Space of $n \times k$ matrixes $\mathbb{M}_{n\times k}(\mathbb{C})$

- $\mathcal{G}^2 \to$ geometric vectors in a plane

$\boxed{\text{Linear Function}}$ $L : V \mapsto W$

$$\Leftrightarrow \begin{cases} L(\alpha \cdot v) = \alpha \cdot L(v) \\ L(v_1 + v_2) = L(v_1) \hat{+} L(v_2) \end{cases}$$

operations in W

Ex

- $\underbrace{\mathbb{M}_{2\times3}} : \mathbb{C}^3 \mapsto \mathbb{C}^2$

$$A = [V_1 \ \ V_2 \ \ V_3]$$

$$A\begin{pmatrix}\alpha\\ \beta\\ \gamma\end{pmatrix} = \alpha V_1 + \beta V_2 + \gamma V_3$$

- $\varphi : \mathbb{C}^3 \longmapsto \mathbb{P}_2$ linear

"Atrix"

We write $\boxed{\Phi = [\, x^2 + ix \quad 3x - i \quad ix^2 + 2x - 7 \,]}$

$$\hookrightarrow \varphi\begin{pmatrix}\alpha\\ \beta\\ \gamma\end{pmatrix} = \alpha(x^2 + ix) + \beta(3x - i) + \gamma(ix^2 + 2x - 7)$$

"Atrix" $\varphi : \mathbb{C}^n \longmapsto V$ $\rightarrow$ always linear.

Define $\varphi = [V_1 \; V_2 \dots V_n]$

$$\varphi(a^i_?) = \sum_{i=1}^{n} a_i V_i = a^i V_i$$

$\boxed{\text{Theorem}}$

↳ Every linear function $\Psi : \mathbb{C}^n \xrightarrow{\text{lin}} V$ is an atrix

key $\Psi(\alpha^i) = \alpha^i \Psi(\vec{e}_i)$

So $\Psi = [\Psi(\vec{e}_i)]$

So atrices are linear functions.

Why are atrices useful? $\rightarrow$ $\begin{cases} \text{linear independence} \\ \text{span} \\ \text{bases.} \end{cases}$

- $Im(\Psi) = span(\{V_i\})$ If $\Psi = [V_1 \dots V_m]$

Linear independence. $\boxed{a^i V_i = 0_V \iff a^i = 0 \;\; \forall i}$

- $\Psi = [V_1 \dots V_m]$ injective when V_i's linearly independent
  $\Updownarrow$
  $ker(\Psi) = 0_V$

- "Basis" = "linear independence" + "span"

- Ψ bijective $\iff$ $\{V_i\}$ is a basis of V

$\boxed{\text{Elementary Column Operations}}$ $\begin{cases} \text{swap} \\ \text{scale} \\ \text{add one} \cdot \alpha \text{ to another.} \end{cases}$

- $[V_1 \dots V_n] : \mathbb{C}^n \mapsto V$

__Note__ Let $\varphi = [V_1 \cdots V_n] \quad \mathbb{C}^n \longmapsto V$

Consider $\underbrace{\mathbb{C}^m \xrightarrow[\text{lin}]{\underset{n\times m}{M}} \mathbb{C}^n \xrightarrow[\text{lin}]{\varphi} V}_{} \xrightarrow[\text{lin}]{} W$

"Every atrix can be combined with a __matrix__"

__Ex__ $\left.\begin{array}{l} A = [V_1 \ V_2 \ V_3]_{V\times 3} \\ M = [\vec{v}_1 \ \vec{v}_2]_{3\times 2} \end{array}\right\} \quad A \circ M = \left[A\vec{v}_1 \overset{?}{\cdot} A\vec{v}_2\right]_{V\times 2}$

↳ Use the standard tuples

$(A\circ M)(\vec{e}_1) = A(M(\vec{e}_1)) = A\vec{v}_1$

So $\boxed{A \circ \underset{[v_k]}{M} = [A(\vec{v}_1) \dots A(\vec{v}_k)]}$

Elementary Column operation

$\boxed{Ex}$

$A = [V_1 \ V_2 \ V_3] \circ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = [V_3 \ V_2 \ V_1] = B$

"swap" is equiv to composing with bijection

→ "jectivity" is preserved.

→ $\operatorname{Im} A = \operatorname{Im} B$

→

$[V_1 \ V_2 \ V_3] \circ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3i & 0 \\ 0 & 0 & 1 \end{pmatrix} = [V_1 \ \ 3iV_2 \ \ V_3]$

$[V_1 \ V_2 \ V_3] \circ \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = [V_1 - 4V_2 \ \ V_2 \ \ V_3]$

So, performing elementary column operations don't alter

$$\left\{ \begin{matrix} \text{`jectivity} \\ \text{Image} \end{matrix} \right\} \quad \text{(composing w/ bijection)}$$

$$\operatorname{Im}[V_1 \dots V_n] = \operatorname{Im}[V_1 \dots V_n \; 0_V]$$ same thing

$$\operatorname{Im}[V_1 \dots V_n] = \operatorname{Im}[V_1 \dots V_{n-1}] \Leftrightarrow V_n \in \operatorname{Ran}[V_1 \dots V_{n-1}]$$

- $(\Leftarrow)$ If $V_n \in \operatorname{Ran}(V_1 \dots V_n) = \operatorname{Im}(V_1 \dots V_n)$

then $\operatorname{Im}[V_1 \dots V_n]$

$$= \operatorname{Im}\left[V_1 \dots V_{n-1} \; \sum^{n-1} V_i a_i\right]$$

$$= \operatorname{Im}[V_1 \dots V_{n-1} \; 0_V] \quad \text{E.C.O}$$

$$= \operatorname{Im}[V_1 \dots V_{n-1}]$$

- $(\Rightarrow)$ $V_n \in \operatorname{Ran}[V_1 \dots V_n] = \operatorname{Ran}[V_1 \dots V_{n-1}]$

"Ran" $\equiv$ "Im"

If not all $\{V_i\}$'s are null, by removing some of them, we can arrive at a linearly independent column with the same <u>span</u>.

Consequently, If $\{V_i\}$ <u>spans</u> V then $\{V_i\}$ contains a <u>basis</u>.

<u>Note</u> NOT THIS $\{0_V\}$ $\rightarrow$ not a basis! Ever!

$$\boxed{\text{Every basis of } V \text{ has the same dimension}}$$

A B

basis

Suppose $[V_1 \dots V_{13}]$ and $[W_1 \dots W_{17}]$ ~~span~~ V

$$[V_1 \dots V_{13}] : \mathbb{C}^{13} \underset{\text{linear}}{\overset{\text{bijection}}{\longmapsto}} V \underset{\text{linear}}{\overset{\text{bijection}}{\longleftarrow}} [W_1 \dots W_{17}]^{-1}\ \mathbb{C}^{17}$$

So $$[V_1 \dots V_{13}] : \mathbb{C}^{13} \underset{\text{lin}}{\overset{\text{bij}}{\longmapsto}} V \underset{\text{lin}}{\overset{\text{bij}}{\longrightarrow}} [W_1 \dots W_{17}]^{-1}\ \mathbb{C}^{17}$$

So $$A\mathbb{C}^{13} \overset{A}{\longmapsto} V \overset{B^{-1}}{\longmapsto} \mathbb{C}^{17}$$

Note $AB^{-1}\ \mathbb{C}^{13} \longmapsto \mathbb{C}^{17}$ has to be a linear bijection

~~But~~ **So** AB^{-1} is a 17×13 matrix

But this can't happen. $\Rightarrow$ Cardinality true.

So

$\begin{cases} \text{Span sets have at least } \dim(V) \text{ elements} \\ \text{linearly independent set have at } \underline{\text{most}} \dim(V) \text{ elements.} \end{cases}$

↳ Claim basis $V_1 \dots V_{13}$
lin. ind $W_1 \dots W_{17}$

$$\begin{tikzcd} \mathbb{C}^{17} \arrow[d, "\text{inj}"', "{[W_1 \dots W_{17}]}"] \arrow[dd, bend right=60] \\ V \arrow[d, shift right, "\text{bij}"'] \\ \mathbb{C}^{13} \arrow[u, shift right, "{[V_1 \dots V_{13}]^{-1}}"'] \end{tikzcd}$$

$[\quad]_{13\times 17}$ → wide matrix cannot be inj!

So

- Every lin. ind in finite dim can be enlarged to a basis

↳ Start w. $[V_1 \dots V_n]$ lin. ind.

If → Span → stop

↳ not Span → $\exists V_{n+1}$ s.t. $V_{n+1} \notin \operatorname{Im}[V_1 \dots V_n]$

→ Bigger matrix $A = [V_1 \dots V_{n+1}]$

Look at .tex

(9)

<u>Claim</u> $[V_1 \ldots V_n\ V_{n+1}]$ lin. ind.

$\hookrightarrow$ since $V_{n+1} \notin \mathrm{Im}[V_1, \ldots V_n]$

$\rightarrow [V_1, \ldots V_n\ V_{n+1}]$ lin. ind $(\Leftarrow)$

keep repeating until

$\#[\ ] = \dim(V)$

$V_1 \ldots V_n$ lin. ind.

$\Leftrightarrow \sum a_i V_i = 0 \Leftarrow a_i = 0 \quad \forall i$

$\Leftrightarrow \begin{cases} V_1 \neq 0_V \\ \text{no } V_i \text{ is a lin ind of the previous } V_i\text{'s} \end{cases}$

<u>Next time</u> Isomorphisms, linear functions & matrices ...

$\rightarrow$ <u>change of basis</u>

then Rank-Nullity theorem

<u>Then</u> $\rightarrow$ <u>Decomposition</u> of vector space w/ a basis

---

<u>Feb 14, 2019</u>

$\boxed{\text{Products of Vector Spaces}}$

$\hookrightarrow$ <u>Recall</u> $\mathbb{C}^2 \begin{pmatrix} a \\ b \end{pmatrix}$, $\mathbb{C}^5 \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \varepsilon \end{pmatrix}$ $\begin{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \varepsilon \end{pmatrix} \end{pmatrix} \rightarrow \mathbb{C}^7$

<u>So</u> $\mathbb{C}^2 \times \mathbb{C}^5 \rightsquigarrow \mathbb{C}^7$

<u>Note</u> $\mathbb{C}^5 \times \mathbb{C}^2 \neq \mathbb{C}^7$

<u>So</u> ... $\mathbb{C}^2 \times \mathbb{C}^5 \longrightarrow \mathbb{C}^7$

<u>So</u> ...

$$\alpha \begin{pmatrix} \bigcirc \\ \square \end{pmatrix} \underset{\sim}{+} \beta \begin{pmatrix} \hat{\bigcirc} \\ \hat{\square} \end{pmatrix} = \begin{pmatrix} \alpha \bigcirc + \beta \hat{\bigcirc} \\ \alpha \square + \beta \hat{\square} \end{pmatrix}$$

So, what if we have 2 vector spaces V & W

$$V \times W = \left\{ \begin{pmatrix} \vec{v} \\ \vec{w} \end{pmatrix} \middle| \ \vec{v} \in V, \ \vec{w} \in W \right\}$$

Redefine addition

$$\left\{ \begin{aligned} \begin{pmatrix} V_1 \\ W_1 \end{pmatrix} \tilde{+} \begin{pmatrix} V_2 \\ W_2 \end{pmatrix} &\triangleq \begin{pmatrix} V_1 \overset{V}{+} V_2 \\ W_1 \overset{W}{+} W_2 \end{pmatrix} \\ \alpha \begin{pmatrix} V \\ W \end{pmatrix} &\triangleq \begin{pmatrix} \alpha V \\ \alpha W \end{pmatrix} \end{aligned} \right\}$$ This makes $V \times W$ a vector space.

Neutral element $\rightarrow \begin{pmatrix} \phi_V \\ \phi_W \end{pmatrix}$

Suppose V is 3 dimensional, W is 5 dimensional } finite-dim

claim $\dim(V \times W) = \dim(V) + \dim(W)$

Proof let ~~suppose~~ $\{v_1, v_2, v_3\}$ be a basis of V

$\{w_1 \ldots w_5\}$ be a basis of W

Consider $\begin{pmatrix} v_1 \\ \phi_W \end{pmatrix}, \begin{pmatrix} v_2 \\ \phi_W \end{pmatrix}, \begin{pmatrix} v_3 \\ \phi_W \end{pmatrix}, \begin{pmatrix} \phi_V \\ w_1 \end{pmatrix} \ldots \begin{pmatrix} \phi_V \\ w_5 \end{pmatrix}$

Claim these span $V \times W$

Given $\begin{pmatrix} V \\ W \end{pmatrix}$ $\quad V = \sum_1^3 a_i v_i \ , \ W = \sum_1^5 b_i w_i$

$$\begin{pmatrix} V \\ W \end{pmatrix} = a_1 \begin{pmatrix} v_1 \\ \phi_W \end{pmatrix} + a_2 \begin{pmatrix} v_2 \\ \phi_W \end{pmatrix} + \ldots\ldots + b_5 \begin{pmatrix} \phi_V \\ W_5 \end{pmatrix} \Rightarrow \text{span}$$

Lin-indep $a_1 \begin{pmatrix} v_1 \\ \phi_W \end{pmatrix} + \cdots a_8 \begin{pmatrix} \phi_V \\ w_5 \end{pmatrix}$

$$= \begin{pmatrix} \sum^3 a_i v^i \\ \sum^5 a_i w^i \end{pmatrix} = \begin{pmatrix} \phi_V \\ \phi_W \end{pmatrix} \iff \begin{cases} a = b = c = 0 \\ d = e = f = g = i = 0 \end{cases}$$

So all coeffs are zero

$\Rightarrow$ So $\dim(V + W) = \dim(V) + \dim(W) = 8$

Caution Vector products are $\boxed{\text{not}}$ commutative

- $\boxed{W \times V \neq V \times W}$ $\quad W \times V \sim V \times W$ But they're not the same
- Not associative

$$\boxed{V \times (W \times Z) \neq (V \times W) \times Z}$$ (natural $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$)

Def
$$\boxed{V_1 \times \ldots \times V_n = \left\{ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \middle| \; v_i \in V_i \right\}}$$

Note

$$\begin{pmatrix} v \\ w \end{pmatrix} \rightsquigarrow \begin{pmatrix} v \\ \phi_W \end{pmatrix} + \begin{pmatrix} \phi_V \\ w \end{pmatrix}$$ "subspace"

Note $\left\{ \begin{pmatrix} v \\ \phi_W \end{pmatrix} \middle| \; v \in V \right\} < V \times W$

$\left\{ \begin{pmatrix} \phi_V \\ w \end{pmatrix} \middle| \; w \in W \right\} < V \times W$

Every object in $V \times W$ can be written in a unique way with $\begin{pmatrix} v \\ \phi_W \end{pmatrix} + \begin{pmatrix} \phi_V \\ w \end{pmatrix}$

→ this is something along the line of basis...

If Z & U are subspaces of a vector space W
then
$$Z \boxplus U \triangleq \{ z \underset{W}{+} u \mid z \in Z, u \in U \}$$
is a subspace of W

↳ idea → decomposing a vector space...

Question: is $Z \boxplus U$ a subspace? → not empty $0_Z + 0_U$

Yes. $\alpha(\underbrace{z}_{\in Z} + \underbrace{u}_{\in U}) = \underbrace{\alpha z}_{\in Z} + \underbrace{\alpha u}_{\in U} \in W$

$$(z_1 + u_1) + (z_2 + u_2) = \underbrace{(z_1 + z_2)}_{\in Z} + \underbrace{(u_1 + u_2)}_{\in U}$$

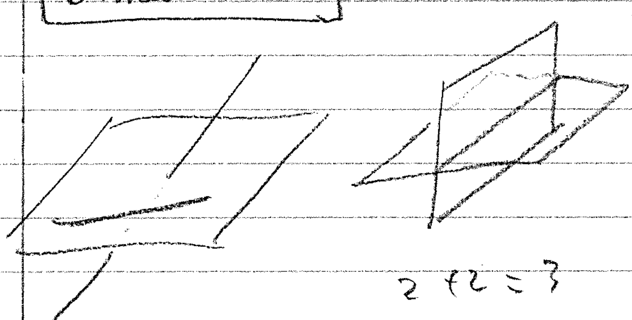
So $Z \boxplus U$ subspace

$\boxed{\text{Question}}$ is $Z \boxplus U$ commutative? Yes
$= U \boxplus Z$

↳ $\boxed{Z \boxplus U = U \boxplus Z}$

Question: Associativity? Yes. $\boxed{(Z \boxplus U) \boxplus W = Z \boxplus (U \boxplus W)}$

$\boxed{\text{Dimensions?}}$ → they don't add! { span, linear ind



$2 + 2 = 3$

$2 + 1 = 2$

4. Suppose Z_1 & Z_2 are subspaces of W

Consider $Z_1 \overset{\otimes}{\times} Z_2 \xrightarrow{\varphi} Z_1 \overset{\oplus}{+} Z_2 \leq W$

Defined by $\varphi\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \triangleq z_1 + z_2$

1 → φ is a linearity function. ⟶ φ is linear

$$\varphi\left(\alpha\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \overset{\otimes}{+} \beta\begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix}\right) = \varphi\begin{pmatrix} \alpha z_1 \overset{\oplus}{+} \beta \hat{z}_1 \\ \alpha z_2 \overset{\oplus}{+} \beta \hat{z}_2 \end{pmatrix} = \alpha z_1 \overset{\oplus}{+} \beta \hat{z}_1 \overset{\oplus}{+} \alpha z_2 \overset{\oplus}{+} \beta \hat{z}_2$$

$$= \alpha \varphi\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \overset{\oplus}{+} \beta \varphi\begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} = \ldots$$

2 If W finite dim, then Z_1, Z_2 finite dim

So $Z_1 \times Z_2$ and $Z_1 + Z_2 \leq W$ finite dim

By Rank-Nullity Theorem (fin → fin, linear

$$\underbrace{\dim(Z_1 \times Z_2)}_{} = \dim(Z_1 + Z_2) + \dim(\ker(\varphi))$$

$$\dim(Z_1) + \dim(Z_2) = \dim(Z_1 + Z_2) + \dim(\ker \varphi)$$

What is $\ker(\varphi)$?

$$\ker(\varphi) = \left\{ \begin{pmatrix} z \\ -z \end{pmatrix} \middle| \begin{array}{l} z \in Z_1 \\ +z \in Z_2 \end{array} \right\} = \left\{ \begin{pmatrix} z \\ -z \end{pmatrix} \middle| z \in \underbrace{\overset{?}{Z_1 \cap Z_2}}_{} \right\}$$

also a subspace of W

From Assignment 1, $\boxed{\dim\left\{\begin{pmatrix} u \\ -u \end{pmatrix} \middle| u \in U\right\} = \dim(U)}$

start with

So $\dim(\ker(\varphi)) = \dim(Z_1 \cap Z_2)$

↳ A sieve formula

$$\boxed{\dim(Z_1 + Z_2) = \dim(Z_1) + \dim(Z_2) - \dim(Z_1 \cap Z_2)}$$

when $Z_1 \cap Z_2$ is trivial, then $Z_1 + Z_2$ is direct.

$\rightarrow \dim(\ker \varphi) = 0 \rightarrow \varphi$ is injective

But φ is also surjective by def

$\Rightarrow \varphi$ is a bijection. $\rightarrow Z_1 \otimes Z_2 \sim Z_1 \oplus Z_2$

isomorphic

Feb 19 2019

Warm up $W_1 + W_2 = V$ direct when $W_1 \cap W_2 = \{0_V\}$

φ bijective $\Leftrightarrow$ $a = 0_V$
$b = 0_V$

Contraposition
If $\{0_V\} \subsetneq W_1 \cap W_2$, then $\exists\, u \neq 0_V$ s.t. $u \in W_1 \cap W_2$

then $u + (-u) = 0_V$ (contradiction)

$\boxed{\text{Why direct Sum?}}$ $\longrightarrow$ generalize the idea of Basis.

↳ decomposes a big vector space into understandable ones.

Now let V be a f.dim vector space

$L \in \mathcal{L}(V)$ $\quad (L : V \xrightarrow{\text{lin}} V)$

$$\begin{array}{ccccc} \mathrm{Im}(L^2) & \subseteq & \mathrm{Im}(L) & \subseteq & V \\ \| & & \| & & \| \\ L^2[V] & \subseteq & L[V] & \subseteq & I[V] \\ & & & & \| \\ & & & & L^0[V] \end{array}$$

What about $L^2[V]$? $\to$ $L[V] \supset L[L[V]]$

<u>Claim</u> $\boxed{\text{if } \mathrm{Im}(L^p) = \mathrm{Im}(L^{p+1}), \text{ then } \mathrm{Im}(L^{p+1}) = \mathrm{Im}(L^{p+2})}$

<u>automatic</u>

$(\mathrm{Im}\, L^{p+2} \subset \mathrm{Im}\, L^{p+1})$

↳ <u>i.e.</u> if $\mathrm{Im}(L^p) = \mathrm{Im}(L^{p+1})$ then

then $\mathrm{Im}(L^{p+1}) \subset \mathrm{Im}(L^{p+2})$

<u>To Show</u> $L^{p+1}(x) \in \mathrm{Im}\, L^{p+2}$ $\forall x \in V$, <u>supp</u> $\mathrm{Im}(L^p) = \mathrm{Im}(L^{p+1})$

$$L^{p+1}(x) = L(L^p(x)) = L(L^{p+1}(v)) \text{ for some } v \in V$$

$$= L^{p+2}(v) \in \mathrm{Im}(L^{p+2})$$

<u>~~Stop~~</u> <u>So</u>

Claim true

"If you get equality, then you'll get equality forever..."

<u>Theorem</u> If $\dim(V) = n$, and $L \in \mathcal{L}(V)$, then

$\mathrm{Im}\, L^n = \mathrm{Im}\, L^{n+1} = \mathrm{Im}\, L^{n+2} = \mathrm{Im}^{n+3} = \cdots$

and $\ker L^n = \ker L^{n+1} = \ker L^{n+2} = \cdots$

"losing dimension"

↳ <u>Kernell</u> grows... $\{0_V\} \subseteq \ker(L) \subseteq \ker(L^2) \subseteq \ldots = \ldots =$

Image shrinks... $= \ldots = \ldots \ldots \subseteq \mathrm{Im}(L^2) \subseteq \mathrm{Im}(L) \subseteq V$

What about $\boxed{\begin{array}{c} Im(L^n) \text{ and } ker(L^n) \text{ and } V? \\ \text{for any } n \geq = dim(V) \end{array}}$

Claim

$\boxed{\text{Theorem}}$ $\boxed{\begin{array}{l} \text{If } V \text{ is finite dimensional and } L \in \mathcal{L}(V) \\ \quad dim(V) = n \\ \text{Then } \quad V = Im(L^n) \oplus Ker(L^n) \end{array}}$

can't we sieve formula ..

- By rank-nullity

$$dim(V) = dim(Im(L^n)) + dim(ker(L^n))$$

- If we can show that $Ran\, L^n \cap ker\, L^n = \{0_V\}$

then the sum is direct, and

$$dim(\underbrace{Ran\, L^n \oplus ker\, L^n}_{\text{subspace of } V}) = dim(Ran\, L^n) + dim(ker\, L^n)$$

$$= n = dim(V)$$

So $\cancel{Ran}(L^n) \oplus ker(L^n) = V$

- Now if $x \in Im\, L^n \cap ker\, L^n$, then

$$x \in L^n(v), \; v \in V, \; L^n(x) = 0_V$$

So $L^n(L^n(v)) = 0_V = L^{2n}(v)$

So $v \in ker(L^{2n}) \Rightarrow v \in ker(L^n) \Rightarrow x = 0_V$

$\boxed{\text{Note}}$ only <u>true</u> if $n = \dim(V)$

If $\dim V = 3$, then $\ker(L^2) \overset{?}{\oplus} \operatorname{Ran}(L^2) \to$ may not give V.

"The loss in dimension" $V \to \operatorname{Im}(L) \to \operatorname{Im}(L^2) \to \cdots$

↳ $\boxed{\text{Weyr number}}$ → can change each step.

What about gain in dimension kernel?

↳ By nullity, Weyr for loss of dim of range (image) is the same of # gain in dimension of kernel ...

$$\begin{tikzcd} \overset{a}{W_1} \times \overset{b}{W_2} \arrow[d, shift right, "\varphi"'] \\ W_1 \oplus W_2 = V \arrow[u, shift right, "\varphi^{-1}"'] \arrow[r, "?"] & {} \end{tikzcd}$$

<u>Projection</u> (a,b) $\xrightarrow{\pi_2}$ $b \in W_2$, (a,b) $\xrightarrow{\pi_1}$ $a \in W_1$

What are projection on V? $W_1 \oplus W_2 \longrightarrow V$

"change of basis" $\overset{?}{\longleftarrow}$ $a+b$

$$\boxed{\begin{tikzcd} W_1 \oplus W_2 \arrow[r, "\varphi^{-1}"] \arrow[dr, "\varphi^{-1}"] & W_1 \times W_2 \arrow[r, "\pi_1"] & W_1 & a \\ & W_1 \times W_2 \arrow[r, "\pi_2"] & W_2 & b \end{tikzcd}}$$

$(a+b)$

Ex Let $E : W_1 \oplus W_2 \longmapsto W_1$ "$E(a+b) = a$"

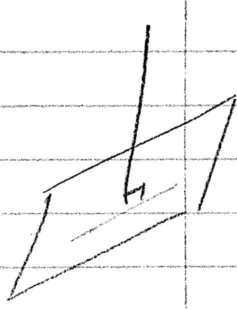
defined as $E(v) = a \in W_1$ s.t $\exists\, b \in W_2$ s.t $a+b = v$

So $E(\underbrace{E(v)}_{a}) = E(v) = a$ $\underset{W_1}{a} = \underset{W_1}{a} + \underset{W_2}{0}$ unique since $\oplus$

true for any $v \in V$

So $\boxed{E^2 = E}$ → "idempotent" (squares to themselves)

↳ Note $\left\{ \begin{array}{l} \text{Ran}(E_1) = \text{Im}(E_1) = W_1 \\ \ker(E_1) = W_2 \end{array} \right\}$



So

Decomposition of $W_1 \oplus W_2$ gives rise to idempotent whose $\text{Im} = W_1$, $\ker = W_2$

→ Are there other idempotents like this? No

↳ For idempotent (F) → $\ker(F) \oplus \text{Im}(F) = V$

In fact Idempotent ⟷ direct sum

↳ coding space by functions $E_1 + \dots + E_k = I$

↳ "resolution of Identity"

Ex $\begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix} + \begin{pmatrix} 0 & & \\ & 1 & \\ & & 0 \end{pmatrix} + \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix}$

↓ span$\{\vec{e}_1\}$ ↓ span$(\vec{e}_2)$ ↘ Span$(\vec{e}_3)$

Why functions code for space? → So we can do algebra on space!

Feb 21 2019

$\boxed{\text{Idempotents}} \to$ linear operator.

Operator: linear fn from vector space onto itself.

$$E : V \mapsto V$$

Def

$\left\{ \text{Idempotents are operators with the property } E^2 = E, \text{ i.e. } E \circ E = E \right\}$

Last time

$\boxed{\begin{array}{l} \text{If } V = W \oplus Z, \text{ then there exists an idempotent } E \in \mathcal{L}(V) \text{ such that } \\ \quad W = \mathrm{Im}(E) \to \text{"projection"} \\ \quad Z = \ker(E) \end{array}}$

In fact, $V = W \oplus Z$ then there are at least 2 idempotents E, $W = \mathrm{Im}(E)$, $Z = \ker(E)$
and F, $W = \ker(F)$, $Z = \mathrm{Im}(F)$

Now ? $\to$ Is every idempotent $E \in \mathcal{L}(V)$ generated this way?

The answer is Yes $\to$ Two parts to answer $\left\{ \begin{array}{l} \text{existence (I)} \\ \text{uniqueness} \end{array} \right\}$

I// Theorem: (II)

$$\boxed{\begin{array}{l} \text{If } E \circ E = E^2 = E, \text{ then } V = \mathrm{Ran}(E) \oplus \ker(E) \\ \qquad\qquad\qquad\qquad\quad V = \mathrm{Im}(E) \oplus \ker(E) \end{array}}$$

Proof 1) Show $\mathrm{Im}(E) + \ker(E) = V$
2) Show $\mathrm{Im}(E) \oplus \ker(E) = V$

(1) First, $v = \underset{\mathrm{Im}(E)}{\underset{\Uparrow}{E(v)}} + \underset{\ker(E)}{\underset{\Uparrow}{(v - E(v))}}$. Now $E(v - E(v)) = Ev - E^2v$
$= Ev - Ev = 0$

Now (2) directness. Show $Im(E) \oplus ker(E) = V$

i.e. $Im(E) \cap ker(E) = \{0_V\}$

Proof Let $x \in Im(E) \rightarrow E(\underbrace{E(x)}_{\text{generic element in image}}) = 0_V$

so $E(x) = 0_V$

so $E(x) \in ker(E)$

By Product

$$\boxed{E(E(x)) = E(x) \text{ for all } x \in V \iff E = E^2}$$

i.e. E is an idempotent multiply when it acts as an identity for on its own image.

Theorem

$\in Im(E)$

$E(w) = w$

$E(E(v)) = Ev$

$$\boxed{\text{if } E^2 = E \text{ and } G^2 = G \text{ and in } \mathcal{L}(V) \; \left\{ \begin{array}{l} Im(G) = Im(E) \\ ker(E) = ker(G) \end{array} \right. \text{ then } E = G}$$

Proof Let $W = Im(E) = Im(G)$ Note $W \oplus Z = V$
$Z = ker(E) = ker(G)$

$v \in V$

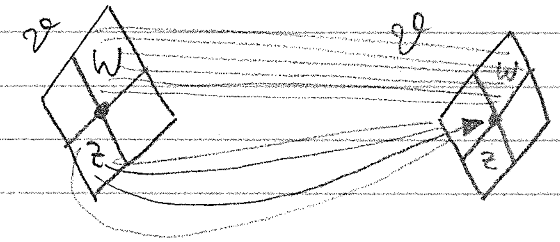
Then $E(v) = E(w+z) = E(w) + \underbrace{E(z)}_{0_V} = E(w) = w$

where $v = w + z$ is unique

Can do the same thing with $G \rightarrow G(v) = w = E(v) \; \forall v$

$\rightarrow G = E$

$\boxed{\text{Picture of idempotent}}$ → think "<u>Projections</u>"



$W = \mathrm{Im}(E)$
$Z = \ker(E)$

<u>So</u>

$\boxed{\text{For each decomposition } V = W \oplus Z, \ \exists \text{ a unique idempotent whose Im is } W \text{ and kernel is } Z}$

<u>Now</u>, notice:

$$w + z \xleftarrow{I} v = w + z \quad \begin{matrix} \xrightarrow{E} w \\ \xrightarrow{F} z \end{matrix}$$

<u>So</u> $E + F = I$

<u>Now</u>, E, F idempotents s.t. $\mathrm{Ran}(E) = \ker(F)$
$\ker(E) = \mathrm{Im}(F)$, then

$E + F = I$ ⇝ identity

<u>So</u>

(1) If $E^2 = E \in \mathcal{L}(V)$, then $(I - E)^2 = (I - E)$

(2) $\cancel{\mathrm{Ran}}(E) = \ker(I - E)$
$\ker(E) = \mathrm{Im}(I - E)$

<u>So</u> idempotents <u>come in pairs</u>

<u>Summary</u> If $V = W \oplus Z$, then there is a pair of idempotents ...
more concretely ...

$\Rightarrow$ $\boxed{\text{If } V = W \oplus Z \text{, then there exist unique idempotents } E \,\&\, F \text{ such that } E + F = I \text{ and } \begin{cases} W = \operatorname{Im}(E) = \ker(F) \\ Z = \ker(E) = \operatorname{Im}(F) \end{cases}}$

$\Leftarrow$ $\boxed{\text{If } \exists \text{ exist idempotents} \ldots E \,\&\, F \ldots \text{such th.. then } V = W \oplus Z}$

All great, but what about $W \oplus Z \oplus U = V$?

Well $W \oplus (Z \oplus U) = V \Rightarrow$

By $\exists$ unique $E_W^2 = E_W$ such that $\operatorname{Im}(E_W) = W$

$\ker(E_W) = Z \oplus U$

(★) Similarly $E_V^2 = E_V$ s.t. $\operatorname{Im}(E_V) = U$

$\ker(E_V) = Z \oplus W$

$E_Z^2 = E_Z$ s.t. $\operatorname{Im}(E_Z) = Z$

$\ker(E_Z) = U \oplus W$

So there exists

Observation ① $E_W + E_U + E_u = ?$

$$(E_Z + E_U + E_W)(v) = (E_W + E_U + E_Z)(u + w + z)$$

$$= E_W(w + u + z) + E_u(w + u + z) + E_Z(u + w + z)$$

$$= w + u + z$$

$$= v$$

So $\boxed{E_W + E_U + E_Z = I}$

Obs ② $E_\omega(E_z(v)) = E_\omega(z) = 0_V$

So $E_\omega \circ E_z$ is null.

In fact compositions of different E_i's get 0_V

compositions of the same E_i's get E_i

Now other direction

If we know $E_w^2 = E_w$, $E_u^2 = E_u$, $E_z^2 = E_z$ such that ... (page 22), then

$$V = W \oplus U \oplus Z$$

Result ~~(Suppose E_W, E_U, E_Z satisfy (*) for some $W, U, Z \leq V$~~

Suppose E_1 E_2 E_3 are idempotents such that

$E_1 + E_2 + E_3 = I$ and $E_i E_j = \delta_j^i E_i$, then

(**) $Im(E_1) \oplus Im(E_2) \oplus Im(E_3) = V$ and

$ker(E_1) = Im(E_2) \oplus Im(E_3)$, etc

Once we established (*) and (**), we have the following assertions are equivalent for ss $W, Z, U \leq V$

① $W \oplus Z \oplus U = V$

② There are unique idempotents $E_1, E_2, E_3 \in \mathcal{L}(V)$ such that

$Im(E_1) = W$, $Im(E_2) = Z$, $Im(E_3) = U$

and $E_1 + E_2 + E_3 = I$, $E_j E_i = \delta_j^i E_i$

1 $\to$ 2 is already shown.

2 $\to$ 1 We will show now.

(1) We have $\mathrm{Im}(E_1) + \mathrm{Im}(E_2) + \mathrm{Im}(E_3) = V$

$$v = \underset{\mathrm{Im}(I) = V}{\underset{\cap}{id(v)}} = (E_1 + E_2 + E_3)(v) = \underset{\mathrm{Im}(E_1)}{\underset{\cap}{E_1(v)}} + \underset{\mathrm{Im}(E_2)}{\underset{\cap}{E_2(v)}} + \underset{\mathrm{Im}(E_3)}{\underset{\cap}{E_3(v)}}$$

So $V = \mathrm{Ran}(E_1) + \mathrm{Ran}(E_2) + \mathrm{Ran}(E_3)$

(2) Show directness. Suppose $x_1 + x_2 + x_3 = 0_V$, $x_i \in \mathrm{Im}(E_i)$

To show: $x_i = 0_V \ \forall i$

Then $E_1(x_1) + E_2(x_2) + E_3(x_3) = 0_V$

So $E_1(E_1(x_1) + E_2(x_2) + E_3(x_3)) = E_1(0_V) = 0_V$

So $E_1(E_1(x_1)) = 0_V$

So $E(x_1) = 0_V$

$x_1 = 0_V$

Same thing, get $x_2 = x_3 = x_1 = 0_V$

So get directness. $\mathrm{Im}(E_1) \oplus \mathrm{Im}(E_2) \oplus \mathrm{Im}(E_3) = V$

(there's more ...)

Note
- for finite dim vector space $\sum E_i = I \not\Rightarrow E_i E_j = \delta_j^i E_i$
- Resolution of identity, code direct sum as algebra ...

Next, kernel

Note $E_1(E_2(v)) = 0_V$

So $\mathrm{Im}(E_2) \subset \ker(E_1)$

Similarly, $\mathrm{Im}(E_3) \subset \ker(E_1)$

So $\mathrm{Im}(E_3) + \mathrm{Im}(E_2) \subset \ker E_1$

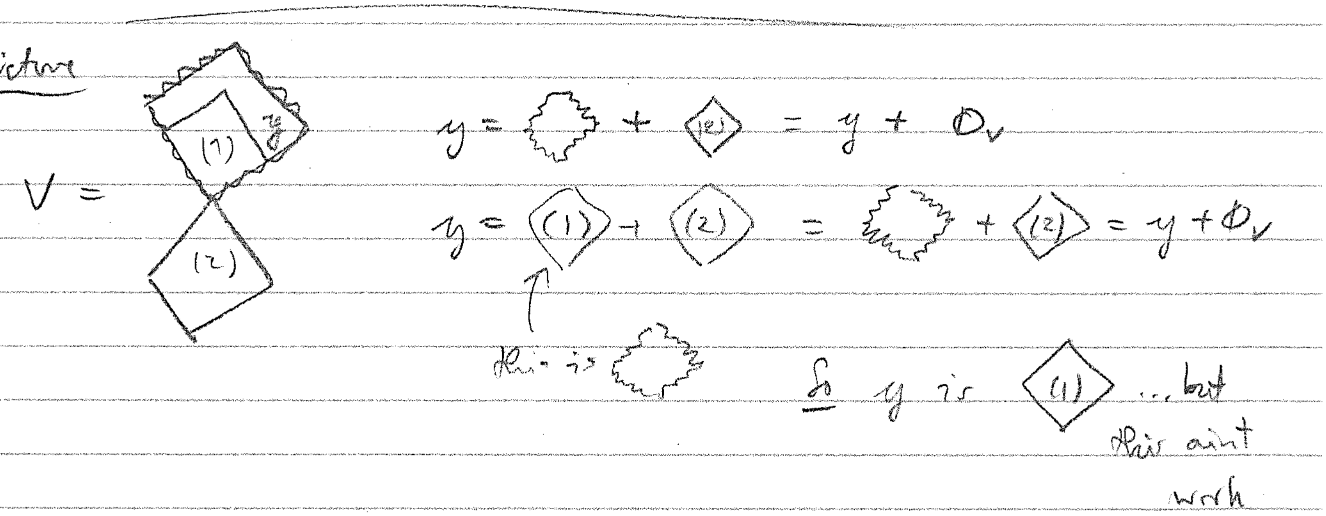
But this is also a direct sum, so

$\mathrm{Im}(E_3) \oplus \mathrm{Im}(E_2) \subset \ker E_1$ → cannot be proper → must be equality

$$\begin{array}{ccc} & \oplus & & \oplus \\ \text{Now} & \mathrm{Im}(E_k) & = & \mathrm{Im}(E_1) \\ \hline & V & = & V \end{array}$$

So $\mathrm{Im}(E_2) \oplus \mathrm{Im}(E_3) = \ker(E_1)$ $\square$

Picture



Feb 26, 2019

Review of idempotents.

TFAE for $A \in \mathcal{L}(V)$

1. $A^2 = A$
2. $Ax = x$ for $x \in Im(A)$
3. $(I-A)^2 = (I-A)$
4. $Im(A) = \ker(I-A)$
5. $Im(I-A) = \ker(A)$

Today ~ $\boxed{\text{Representing linear functions as matrices}}$

Consider $\underset{f.d.}{V} \longmapsto V$ and coordinate system, i.e. order basis.

$\beta = (v_1, \dots, v_m)$ ordered basis of V (coord. sys. of V)

Then if $z \in V$, $z = \sum_{i=1}^{m} a_i v_i$ unique. But long notation → new notation

$$[z]_\beta = (a_1 \dots a_m)^T$$

So consider atrix $[V_1 \dots V_m]$, then $[V_1 \dots V_m][z]_\beta = z$

i.e. $\underbrace{[V_1 \dots V_m]}\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = a_1 V_1 + a \dots + a_m V_m = z$

$$\boxed{A_\beta [z]_\beta = z}$$

~ ~~give the address of z_i~~ gives z from $[z]_\beta$

Image

$$\begin{array}{ccc} z & & \underset{f.d.}{V} \longmapsto V \\ {\scriptstyle [\ ]_\beta}\downarrow & & A_\beta \uparrow \downarrow A_\beta^{-1} = [\ ]_\beta \\ [z]_\beta & & \mathbb{C}^m \end{array}$$

Note $[V_1 \dots V_m] = A_\beta$ is a bijection, obviously

More fully. Consider $\mathcal{L} : V \xmapsto{\text{lin}} V$

Get
$$\begin{tikzcd}
V \arrow[rr, mapsto, "\mathcal{L}", "\text{linear}"'] \arrow[d, shift left, "{[\ ]_\beta}"] & & V \arrow[d, shift left, "{[\ ]_\beta}"] \\
\mathbb{C}^m \arrow[u, shift left, "A_\beta"] \arrow[rr, "M"'] \arrow[rr, bend left=60] & & \mathbb{C}^m \arrow[u, shift left, "A_\beta"]
\end{tikzcd}
\Bigg\} \text{ isomorphisms}$$

M is a composition of linear fns mapping $\mathbb{C}^m \mapsto \mathbb{C}^m$, so M is a matrix. Note $A_\beta^{-1} = [\ ]_\beta$ are isomorphisms

We note $\quad M \equiv [\mathcal{L}]_{\beta\leftarrow\beta}$

Def $\quad \boxed{[\mathcal{L}]_{\beta\leftarrow\beta} := [\ ]_\beta \circ \mathcal{L} \circ A_\beta = A_\beta^{-1} \circ \mathcal{L} \circ A_\beta}$

Now
$$\begin{tikzcd}
z \arrow[r] \arrow[d] & \mathcal{L}(z) \arrow[d] \\
{[z]_\beta} \arrow[r] & {[\mathcal{L}(z)]_\beta}
\end{tikzcd}$$

Claim $\quad [\mathcal{L}(z)]_\beta = [\mathcal{L}]_{\beta\leftarrow\beta}\,[z]_\beta$

$$\begin{array}{ccc} \| & & \| \\ [\ ]_\beta \circ \mathcal{L}(z) & \text{vs} & ([\ ]_\beta \circ \mathcal{L} \circ \underbrace{A_\beta) \circ [\ ]_\beta}_{I}\, z \end{array}$$

So $\quad ([\ ]_\beta \circ \mathcal{L})(z) = ([\ ]_\beta \circ \mathcal{L})(z) \longrightarrow$ true.

So, observe
$$\begin{tikzcd}
(z) \arrow[r, "\mathcal{L}"] \arrow[d, leftrightarrow] & \mathcal{L}(z) \\
{[z]_\beta} \arrow[r, "M"] & {[\mathcal{L}(z)]_\beta}
\end{tikzcd}$$

$\boxed{\text{Consider 2 bases}}$

$$\begin{tikzcd}
\mathbb{C}^m \arrow[rr, "{[\mathcal{L}]_{\gamma\leftarrow\gamma}}"] \arrow[dd, bend right=60, "\text{isomorphism}"'] \arrow[dr, shift left, "A_\gamma"] & & \mathbb{C}^m \arrow[dl, shift left, "A_\gamma"] \\
& \mathcal{V} \arrow[ul, shift left, "{[\,]_\gamma}"] \arrow[dl, shift left, "{[\,]_\beta}"] \arrow[r, "\mathcal{L}"] & \mathcal{V} \arrow[u, shift left, "{[\,]_\gamma}"] \arrow[d, shift left, "{[\,]_\beta}"] \\
\mathbb{C}^m \arrow[ur, shift left, "A_\beta"] \arrow[rr, "{[\mathcal{L}]_{\beta\leftarrow\beta}}"'] & & \mathbb{C}^n \arrow[u, shift left, "A_\beta"] \arrow[uu, bend right=60, "\text{isomorphism}"']
\end{tikzcd}$$

Are $[\mathcal{L}]_{\gamma\leftarrow\gamma}$ the same as $[\mathcal{L}]_{\beta\leftarrow\beta}$

$$\begin{aligned}
[\mathcal{L}]_{\gamma\leftarrow\gamma} &= [\;]_\gamma \circ \mathcal{L} \circ A_\gamma \\
&= [\;]_\gamma \circ \left(A_\beta \circ [\mathcal{L}]_{\beta\leftarrow\beta} \circ [\;]_\beta \right) \circ A_\gamma \\
&= \left([\;]_\gamma \circ A_\beta \right) \circ [\mathcal{L}]_{\beta\leftarrow\beta} \circ \left([\;]_\beta \circ A_\gamma \right)
\end{aligned}$$

<u>Note</u>

$$\left.\begin{aligned}
[\;]_\gamma \circ A_\beta &= \left([\;]_\beta \circ A_\gamma \right)^{-1} \\
&= A_\gamma^{-1} \circ [\;]_\beta^{-1} \\
&= [\;]_\gamma \circ A_\beta
\end{aligned}\right\} \quad \underline{\text{nice}}$$

<u>But note</u> $[\;]_\gamma \circ A_\beta$ and $[\;]_\beta \circ A_\gamma$ are invertible matrices $\mathbb{C}^m \longrightarrow \mathbb{C}^m$.

<u>So</u> if we choose different bases.

$$\boxed{[\mathcal{L}]_{\gamma\leftarrow\gamma} = M^{-1} \circ [\mathcal{L}]_{\beta\leftarrow\beta} \circ M}$$

$M \to$ invertible matrix

So $[\mathcal{L}]_{\gamma\leftarrow\gamma}$ is similar to $[\mathcal{L}]_{\beta\leftarrow\beta}$

?

<u>Question</u> If $B \sim [\mathcal{L}]_{\beta\leftarrow\beta}$, does there exist basis γ such that $B = [\mathcal{L}]_{\gamma\leftarrow\gamma}$

$\boxed{\text{The answer is YES}}$

If $B = K^{-1}[\mathcal{L}]_{\beta\leftarrow\beta} K$, we will show that there is a basis γ of V such that

$K = [\;]_\beta \circ A_\gamma$, so that $B = [\mathcal{L}]_{\gamma\leftarrow\gamma}$.

i.e. we will show that there is a basis γ of V s.t.

$$A_\beta \circ K = A_\gamma$$

Next $\underbrace{\underbrace{A_\beta}_{V \xleftarrow{} \mathbb{C}^m \text{ isom}} \circ \underbrace{K}_{\mathbb{C}^m \xleftarrow{} \mathbb{C}^m \text{ isom}}}_{V \xleftarrow{} \mathbb{C}^m \text{ isom}}$. So $A_\beta \circ K$ must be a bijective matrix & so is

$$A_\gamma = [\underbrace{\;\ldots\;}_{\text{some basis } \gamma \text{ of } V}]$$

So similar matrices represent the same linear function. But which one?
What is this linear function?

Sure the answer is unsatisfying... Let $A \in M_{153}$

$\hookrightarrow$ itself $A = [A]_{\mathcal{E}\leftarrow\mathcal{E}}$

$$\begin{tikzcd} \mathbb{C}^{153} \arrow[r, "A"] \arrow[d, shift left, "{[\,]_{\mathcal{E}} = I}"] & \mathbb{C}^{153} \arrow[d, shift left, "{[\,]_{\mathcal{E}}}"] \\ \mathbb{C}^{153} \arrow[r, "A"] \arrow[u, shift left, "I = A_{\mathcal{E}}"] & \mathbb{C}^{153} \arrow[u, shift left, "A_{\mathcal{E}} = I"'] \end{tikzcd}$$

$\mathcal{E} \to$ std basis

So if $A \sim B$ then B represents A with respect to some different basis (not the standard basis)

$\boxed{\text{Big Idea}}$ Similarity is all about change of basis.
Similar matrices ~~give~~ represent the same linear fn.
The goal is to find better representation...

$\#$

$\boxed{Ex}$ $\mathbb{P}_3 \xrightarrow{L} \mathbb{P}_3$

defined as $L(p) = p' + 2p$

Find $[\mathcal{L}]_{\beta \leftarrow \beta}$ where $\beta = \{1, x+1, x^2, x^3\}$

$A_\beta = [\, 1 \quad x+1 \quad x^2 \quad x^3 \,]$

First column of $[\mathcal{L}]_{\beta \leftarrow \beta} = [\;]_\beta \circ \mathcal{L} \circ A_\beta \, e_1$

$= [\mathcal{L}(e_1)]_{\beta_1}$

So i^{th} column of $[\mathcal{L}]_{\beta \leftarrow \beta} = [\mathcal{L}(e_i)]_\beta$

So

1^{st} col of $[\mathcal{L}]_{\beta \leftarrow \beta} = (2\ 0\ 0\ 0)^T$
2^{nd} col of $[\mathcal{L}]_{\beta \leftarrow \beta} = (1\ 2\ 0\ 0)^T$
3^{rd} col of $[\mathcal{L}]_{\beta \leftarrow \beta} = (-2\ 2\ 2\ 0)^T$
4^{th} col of $[\mathcal{L}]_{\beta \leftarrow \beta} = (0\ 0\ 3\ 2)^T$

So $[\mathcal{L}]_{\beta \leftarrow \beta} = \begin{pmatrix} 2 & 1 & -2 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

Side note

$\alpha [L]_{\beta \leftarrow \beta} + [M]_{\beta \leftarrow \beta} = [\alpha L + M]_{\beta \leftarrow \beta}$

$[L]_{\beta \leftarrow \beta} \circ [M]_{\beta \leftarrow \beta} = [L \circ M]_{\beta \leftarrow \beta}$

Notice

~~$\mathcal{L}(V) \longmapsto [\mathcal{L}]_{\beta \leftarrow \beta}$ is injective~~

Note $\mathcal{L} \in \mathcal{L}(V) \longmapsto M_n(\mathbb{C}) \ni [L]_{\beta \leftarrow \beta}$

This map is isomorphic. But there's also multiplicative. C
$\Rightarrow$ called an algebra ... $\mathcal{L}(V)$ & M_n are algebras ...

BLOCK MATRICES

Feb 28, 2019

Start with Cartesian product of 2 vector spaces

$$V \times W \xrightarrow[\text{lin}]{L} V \times W$$

$$\begin{pmatrix} v \\ w \end{pmatrix} \downarrow \begin{pmatrix} v \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ w \end{pmatrix}$$

So formal...

$$\boxed{\begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{pmatrix} v \\ w \end{pmatrix} \triangleq \begin{pmatrix} F_{11} v + F_{12} w \\ F_{21} v + F_{22} w \end{pmatrix}}$$

So

$$\left.\begin{aligned} F_{11} &: V \longmapsto V \\ F_{12} &: W \longmapsto V \end{aligned}\right\} \text{ since } F_{11}(v) + F_{12}(w) \in V$$

and

$$\left.\begin{aligned} F_{21} &: V \longmapsto W \\ F_{22} &: W \longmapsto W \end{aligned}\right\} \text{ since } F_{21}(v) + F_{22}(w) \in W$$

In diagram

	V	W
V	$\swarrow$	$\swarrow$
W	$\swarrow$	$\swarrow$

If F_{ij} are all linear, then $\begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$ is a linear fn as well

$$\begin{bmatrix} F_{12} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} : V \times W \longrightarrow V \times W$$

$\boxed{Q}$ → can every linear fn be represented this way?

Note $\begin{array}{|c|c|} \hline F_{11} & F_{12} \\ \hline F_{21} & F_{22} \\ \hline \end{array} \begin{pmatrix} v \\ 0 \end{pmatrix} = \begin{pmatrix} F_{11}(v) \\ F_{21}(v) \end{pmatrix}$

In particular $\pi_1\left(\underbrace{\boxplus}_{L}\begin{pmatrix} v \\ 0 \end{pmatrix}\right) = F_{11}(v)$ (π : projection)

So

Given $L : V \times W \xrightarrow{\text{lin.}} V \times W$

Let $F_{11}(v) := \pi_1 \circ L\begin{pmatrix} v \\ 0 \end{pmatrix}$

$F_{12}(w) := \pi_1 \circ L\begin{pmatrix} 0 \\ w \end{pmatrix}$

$F_{21}(v) := \pi_2 \circ L\begin{pmatrix} v \\ 0 \end{pmatrix}$

$F_{22}(w) := \pi_2 \circ L\begin{pmatrix} 0 \\ w \end{pmatrix}$

Do we have $L = \begin{array}{|c|c|} \hline F_{11} & F_{12} \\ \hline F_{21} & F_{22} \\ \hline \end{array}$?

Well

$$\begin{cases} F_{11}(v) + F_{12}(w) = \pi_1 \circ L\begin{pmatrix} v \\ 0 \end{pmatrix} + \pi_1 \circ L\begin{pmatrix} 0 \\ w \end{pmatrix} = \pi_1 \circ L\begin{pmatrix} v \\ w \end{pmatrix} \\ F_{21}(v) + F_{22}(w) = \pi_2 \circ L\begin{pmatrix} v \\ 0 \end{pmatrix} + \pi_2 \circ L\begin{pmatrix} 0 \\ w \end{pmatrix} = \pi_2 \circ L\begin{pmatrix} v \\ w \end{pmatrix} \end{cases}$$

by linearity

So $\begin{pmatrix} F_{11}(v) + F_{12}(w) \\ F_{21}(v) + F_{22}(w) \end{pmatrix} = \begin{pmatrix} \pi_1 \circ L\begin{pmatrix} v \\ w \end{pmatrix} \\ \pi_2 \circ L\begin{pmatrix} v \\ w \end{pmatrix} \end{pmatrix} = L\begin{pmatrix} v \\ w \end{pmatrix}$ obviously...

So If we start with any L, we can represent L as a matrix of linear functions...

■ But we don't often work with Cartesian products. So we want to use this idea to break a vector space into direct sums

$\boxed{\text{What about direct sums?}}$

■ Bad notation Given $W \oplus Z = V$. Let us write $\begin{pmatrix} w \\ z \end{pmatrix}_{\oplus}$ instead of $w+z$

so $\begin{pmatrix} w \\ z \end{pmatrix}_{+} \sim w+z$ so that we can mimick our previous idea

$\begin{pmatrix} w \\ z \end{pmatrix}_{+}$ is actually $\varphi \begin{pmatrix} w \\ z \end{pmatrix} = w+z$, $W \times Z \longrightarrow W \oplus Z$

■ Recall

$$\begin{array}{c|c|c|} & W & Z \\ \hline W & \swarrow & \swarrow \\ \hline Z & \swarrow & \swarrow \\ \hline \end{array}$$

Given
$F_{11} : W \longmapsto W$
$F_{12} : Z \longmapsto W$
$F_{21} : W \longmapsto Z$
$F_{22} : Z \longmapsto Z$

we can define linear fun

$$\begin{array}{|c|c|} \hline F_{11} & F_{12} \\ \hline F_{21} & F_{22} \\ \hline \end{array} : V \longmapsto V \quad \text{by} \quad \begin{array}{|c|c|} \hline F_{11} & F_{12} \\ \hline F_{21} & F_{22} \\ \hline \end{array} \begin{pmatrix} w \\ z \end{pmatrix}_{+} = \begin{pmatrix} \cdots \\ \cdots \end{pmatrix}_{+}$$

same as before...

■ Now, we want to find something that is similar to π_i (projection)

→ the answer is idempotent (analogous to projections)

★ Think

$$\begin{array}{ccccc} W & \oplus & Z & = & V \\ \| & & \| & & \\ \operatorname{Ran}(E_1) & & \operatorname{Ran}(E_2) & & \end{array}$$

where E_i are idempotents.

where $E_2 = I - E$

■ If $L = \begin{array}{|c|c|} \hline F_{11} & F_{12} \\ \hline F_{21} & F_{22} \\ \hline \end{array}$, then

$$L \begin{pmatrix} w \\ 0 \end{pmatrix}_{+} = \begin{pmatrix} F_{11}(w) \\ F_{21}(w) \end{pmatrix}_{+} \begin{matrix} \in W \\ \in Z \end{matrix}$$

so

$E \circ L(w) = F_{11}(w) \in W$

$(I-E) \circ L(w) = F_{21}(w) \in Z$

Same thing $\quad L\begin{pmatrix} 0 \\ z \end{pmatrix}_+ = \begin{pmatrix} F_{12}(z) \\ F_{22}(z) \end{pmatrix}_+$

So $\quad E(L(z)) = F_{12}(z) \in W$

$\quad (I-E)\,L(z) = F_{22}(z) \in Z$

So can again check that

$$L = \begin{array}{c|c|c} & W & Z \\ \hline W & E\circ L\big|_W & E\circ L\big|_Z \\ \hline Z & (I-E)\circ L\big|_W & (I-E)\circ L\big|_Z \end{array}$$

restricted maps...

So if "$L : V \longmapsto V$ and wrt to the composition $V = W \oplus Z$ L has the block form $\begin{array}{|c|c|} \hline A & B \\ \hline C & D \\ \hline \end{array}$"

then it means

$$A : E\circ L\big|_W : W \longmapsto W \quad \text{etc}$$

where E is an idempotent with $Im(E) = W$
$\ker(E) = Z$

Caveat direct sum is commutative
BUT here order matters!

[Order of the direct sum matters in our exploration here.]

Example ex $\quad G$: idem $V \longmapsto V$, $\quad V = \operatorname{Ran} G \oplus \ker G$

So

$$G = \begin{array}{c|c|c} & Im\,G & \ker G \\ \hline Im\,G & I\big|_{Im\,G} & 0 \\ \hline \ker G & 0 & 0 \end{array}$$

Properties of Block-representation...

Algebra of Block-matrices

↳ All our block-matrices go from $V \longmapsto V$ and $V = W \oplus Z$

①
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} + \begin{bmatrix} P & Q \\ R & T \end{bmatrix} = \begin{array}{c|cc} & W & Z \\ \hline W & ?_{11} & ?_{12} \\ Z & ?_{21} & ?_{22} \end{array}$$
$$(L) \qquad\qquad (S) \qquad\qquad (L+S)$$

$W = \mathrm{Im}(E)$, $E^2 = E$

$Z = \ker(E)$

$$?_{11}(w) = E\Big[(L+S)(w)\Big] = EL(w) + ES(w) = A(w) + P(w)$$

So... the algebra is very simple :-)

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} + \begin{bmatrix} P & Q \\ R & T \end{bmatrix} = \begin{bmatrix} A+P & B+Q \\ C+R & D+T \end{bmatrix}$$

② Scaling also works...
$$\alpha \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \alpha A & \alpha B \\ \alpha C & \alpha D \end{bmatrix}$$

③ Composition
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \circ \begin{bmatrix} P & Q \\ R & T \end{bmatrix} = \begin{array}{c|cc} & W & Z \\ \hline W & ?_{11} & \\ Z & & \end{array}$$
$$L \qquad\qquad S \qquad\qquad L\circ S$$

$$?_{11}(w) = E\Big[L\circ S(w)\Big] = ELS(w)$$
$$= EL\Big[E + (id - E)\Big]S(w)$$
$$= ELES(w) + EL(id-E)S(w)$$
$$\cancel{= ELE(ES(w)) + EL(id-E)(id-E)S(w)}$$
$$= \underbrace{EL}\, \underset{\substack{\uparrow \\ W}}{P(w)} + EL\, \underset{\substack{\uparrow \\ Z}}{\underbrace{R(w)}}$$
$$= A(P(w)) + B(R(w))$$

So $\boxed{?_{11}(w) = [A \circ P + B \circ R](w)}$ multiply like matrices!

So
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \circ \begin{bmatrix} P & Q \\ R & T \end{bmatrix} = \begin{bmatrix} A \circ P + B \circ R & A \circ Q + B \circ T \\ C \circ P + D \circ R & C \circ Q + D \circ T \end{bmatrix}$$

(4) Suppose I have $L = \begin{array}{c|cc} & w & z \\ \hline w & A & B \\ z & C & D \end{array}$. If I pick a basis β for W and γ for Z

then

$$[L]_{\beta \| \gamma \leftarrow \beta \| \gamma} = \begin{bmatrix} c_1 & \cdots & \end{bmatrix}$$

concatenate

$C_1 = [L(b_1)]_{\beta\|\gamma}$ $\quad b_1 \to$ first vector of β_γ

$$= \left[\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{pmatrix} b_1 \\ 0 \end{pmatrix}_+ \right]_{\beta\|\gamma} = \left[\begin{pmatrix} A(b_1) \\ C(b_1) \end{pmatrix}_+ \right]_{\beta\|\gamma}$$

$\in W$

$\in Z$

$A(b_1) =$ lin comb of b_i
$C(b_i) =$ lin co of g_i

same thing

$$= \begin{bmatrix} [A(b_1)]_\beta \\ [C(b_1)]_\gamma \end{bmatrix}$$

If we keep doing this ... $\to$ Note $W \sim \beta$, $Z \sim \gamma$

$$[L]_{\beta\|\gamma \leftarrow \beta\|\gamma} = \begin{array}{c|cc} & \beta\{W & \gamma\{Z \\ \hline W\{\beta & [A]_{\beta\leftarrow\beta} & [B]_{\beta\leftarrow\gamma} \\ Z\{\gamma & [C]_{\gamma\leftarrow\beta} & [D]_{\gamma\leftarrow\gamma} \end{array}$$

Note $[C]_{\gamma\leftarrow\beta}$ is generated

(1) Take sth in beta
(2) $C(b)$
(3) coord in γ

$$\to [C]_{\gamma\leftarrow\beta} = \begin{bmatrix} [C(\beta_1)]_\gamma & [C(\beta_2)]_\gamma \ldots & [C(\beta_n)]_\gamma \end{bmatrix}$$

Nov 5, 2019

Representing linear functions as matrix-like stuff...

$$\mathcal{L}: \overset{F_1}{V_1} \oplus \overset{F_2}{V_2} \oplus \overset{F_3}{V_3} \longmapsto \overset{E_1}{W_1} \oplus \overset{E_2}{W_2}$$

	V_1	V_2	V_3
W_1	$\mathcal{L}_{11}$	$\mathcal{L}_{12}$	$\mathcal{L}_{13}$
W_2	$\mathcal{L}_{21}$	$\mathcal{L}_{22}$	$\mathcal{L}_{23}$

$$\mathcal{L}_{13} = E_1 \circ \mathcal{L}\big|_{V_2} = E_1 \circ \mathcal{L} \circ F_2\big|_{V_2}$$

Back to idempotents

Consider idempotent matrices. $E = E^2 \in M_n$

Ex $\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$ is such a matrix.

↳ we know that $\mathrm{Im}(E) \oplus \ker(E) = \mathbb{C}^n$

So $E: \mathrm{Im}(E) \oplus \ker(E) \longmapsto \mathrm{Im}(E) \oplus \ker(E)$

So we can represent E as

		E: $\mathrm{Im}\,E$	$I-E$: $\ker E$
(E)	$\mathrm{Im}\,E$	E_{11}	E_{12}
$(I-E)$	$\ker E$	E_{21}	E_{22}

$$E_{12} = E \circ E\big|_{\ker E} = 0$$

$$E_{22} = \underbrace{(I-E)\circ E}_{0}\big|_{\ker E} = 0$$

$$E_{11} = \underbrace{E \circ E}_{E}\big|_{\mathrm{Im}\,E} = Id$$

$$E_{21} = \underbrace{(I-E)\circ E}_{0}\big|_{\mathrm{Im}\,E} = 0$$

So "E" $= \boxed{\begin{matrix} Id & 0 \\ 0 & 0 \end{matrix}}$

Idea

Can pick a basis for $\mathrm{Im}(E)$
& $\ker(E)$
→ then concatenate to get basis for whole space

$$[E]_{\Gamma_1 \| \Gamma_2 \leftarrow \Gamma_1 \| \Gamma_2} = \begin{bmatrix} [Id]_{\Gamma_1 \leftarrow \Gamma_1} & [0]_{\Gamma_1 \leftarrow \Gamma_2} \\ [0]_{\Gamma_2 \leftarrow \Gamma_1} & [0]_{\Gamma_2 \leftarrow \Gamma_2} \end{bmatrix}$$

Now $[Id]_{\Gamma_1 \leftarrow \Gamma_1} = I$

$[0]_{\Gamma_i \leftarrow \Gamma_j} = 0$ So $E = \begin{array}{|c|c|}\hline I & 0 \\ \hline 0 & 0 \\ \hline\end{array}$

(1) So I can write $[E]_{\Gamma_0 \leftarrow \Gamma_0} = E = E^2 \in M_n$

$\uparrow$ standard basis

(2) But we also know:

$$[E]_{\Gamma_1 \| \Gamma_2 \leftarrow \Gamma_1 \| \Gamma_2} = \begin{array}{|c|c|}\hline I & 0 \\ \hline 0 & 0 \\ \hline\end{array}$$

From (1) & (2), we have a theorem: Every idempotent $E \in M_n$ is similar to a matrix of the form

$$\begin{pmatrix} \underbrace{1 \; 1 \; 1}_{k} & & & 0 \\ & \ddots & & \\ 0 & & 0 & \\ & & & 0 \; 0 \end{pmatrix} \quad \text{where } k = \text{Rank}(E) = \text{Tr}(E)$$

So $\text{Rank}(E) = \text{Trace}(E)$

consequence? { $\hookrightarrow$ So since $\text{Rank}(E) = \text{Trace}(E)$, if $\text{Trace}(M)$ is not an integer or negative ... then M is NOT idempotent

(PRE-) $\boxed{\text{Corollary}}$ Suppose $\underbrace{E_1 + E_2 + \dots + E_m}_{m \text{ idempotents}} = I_{\mathbb{C}^n}$

Then $\text{Trace}(E_1 + E_2 + \dots + E_m) = \text{Trace}(I_{\mathbb{C}^n}) = n$

$\hookrightarrow \text{Tr}(E_1) + \text{Tr}(E_2) + \dots + \text{Tr}(E_m) = n$

So

$$\boxed{\text{Rank}(E_1) + \text{Rank}(E_2) + \dots + \text{Rank}(E_m) = n}$$

Now, consider $\operatorname{Im}(E_1 + \dots + E_m)$

and

$\operatorname{Im} E_1 + \operatorname{Im} E_2 + \dots + \operatorname{Im} E_m$

Consider $(E_1 + \dots E_m)(x) = E_1 x + E_2 x + \dots + E_m x \in \sum \operatorname{Im} E_i$

So $\mathbb{C}^n = \operatorname{Im}(I_{\mathbb{C}^n}) = \operatorname{Im}\left(\sum_1^m C_i\right) \subseteq \sum_1^m \operatorname{Im} C_i \subseteq \mathbb{C}^n$

So $\sum_1^m \operatorname{Im} E_i = \mathbb{C}^n$ (i)

But we also know $\sum_1^m \dim(\operatorname{Im}(E_i))$

$= \sum_1^m \operatorname{rank}(E_i) = n$ (ii)

From (i) & (ii), then

$\boxed{\bigoplus_{i=1}^m \operatorname{Im}(E_i) = \mathbb{C}^n}$ ← direct sum

In particular,

$$\dim\left(\bigoplus_{i=2}^m \operatorname{Im}(E_i)\right) = n - \dim(\operatorname{Im} E_1)$$

$$= \text{nullity}(E_1) = \dim(\ker E_1)$$

Now,

$$\operatorname{Im}\left(\sum_2^m E_i\right) = \operatorname{Im}(I - E_1)$$

$$\cap \qquad \qquad \parallel$$

Now $\bigoplus_{i=2}^m \operatorname{Im}(E_i) = \ker(E_1)$

So Since $\ker(E_1) \subseteq \bigoplus_{i=2}^m \operatorname{Im}(E_i)$

But $\dim \underset{\ker}{\wedge} (E_1) = \dim\left(\bigoplus_{i=2}^m \operatorname{Im}(E_i)\right)$

$$\left.\right\} \boxed{\ker(E_1) = \bigoplus_2^m \operatorname{Im} E_i}$$

This implies $\operatorname{Im} E_2 \subseteq \ker E_1$

So $E_1\left[E_2(\square)\right] = \mathcal{O}_n \quad \forall \square$

Hence $E_1 \circ E_2 = \mathcal{O}$

So, in general $\boxed{E_i \circ E_j = \mathcal{O}_{n\times n} \quad \text{if} \quad i \neq j}$

So Corollary: If $E_1 \ldots E_m$ are idempotents and $\sum_1^m E_i = I$, then $\left\{ \begin{array}{l} E_i E_j = 0 \text{ for } i \neq j \\ \quad \& \\ \bigoplus_{i=1}^m \operatorname{Im}(E_i) = \mathbb{C}^n \end{array} \right\}$

Note

$\bigoplus_{i=2}^m \operatorname{Im} E_i = \ker(E_1)$

Note $\operatorname{Trace}(AB) = \operatorname{Trace}(BA)$

So

$\left.\begin{array}{l} \operatorname{Trace}(S^{-1}AS) = \operatorname{Trace}(ASS^{-1}) \\ \operatorname{Trace}(B) = \operatorname{Trace}(A) \end{array}\right\}$ $\operatorname{Trace}(AB^T) = \operatorname{Trace} b_A \cdot b_B = \operatorname{Trace}(B^T A)$

$\hookrightarrow$ use indices.

So similar matrices have the same trace & rank.

Now, let's revisit Ξ $\to$ what we've done is change of basis to get a better matrix (simpler...)

Now, suppose $\mathcal{L} : V \xrightarrow{\text{lin}} V \to$ finite dim

Suppose $\{0\} \neq \underset{\underset{V}{\neq}}{W} \in \operatorname{Lat}(\mathcal{L})$

$V = W \oplus \boxed{?}$

Let $w_1 \ldots w_m \in W$ basis of W, then $\operatorname{span}\{w_{m+1} \ldots w_n\} \to$ basis of $\square$

Now think V $\quad v_1 \ldots v_{37} \mid v_{38} \ldots v_{155}$

Then $\operatorname{span}(v_1 \ldots v_{37}) + \operatorname{span}(v_{38} \ldots v_{155}) = V$

because $V = \underbrace{\sum^{37} a_i v_i}_{W_1} + \underbrace{\sum_{38}^{155} a_i v_i}_{W_2}$ $\quad (\checkmark)$

$$\begin{array}{cc} W_1 & W_2 \\ \cap & \cap \\ W_1 & W_2 \end{array}$$

So $\operatorname{span}(v_1 \ldots v_{37}) \oplus \operatorname{span}(v_{38} \ldots v_{155}) = V$

since $\dim W_1 + \dim W_2 = V$

$$\begin{array}{c} E \quad I-E \\ \text{Now} \quad V = W \oplus Z \end{array} \qquad (W \in \text{Lat}(\mathcal{L}))$$

$$\text{Then} \quad \begin{array}{c|c|c|} & W & Z \\ W & & \\ \hline (I-E)\,Z & \theta & \\ \hline \end{array} \qquad ? = (I-E) \circ \mathcal{L}\big|_W = 0$$

Note $\mathcal{L}(w) \in W = \operatorname{Im} E = \ker(I-E)$

$\rightarrow (I-E) \circ \mathcal{L}\big|_W = \theta$

In fact, converse also true

$(I-E) \circ \mathcal{L}\big|_W = \theta \Rightarrow \operatorname{Im}(\mathcal{L}|_W) \subseteq \ker(I-E) = \operatorname{Ran}(E) = W$

So $W \in \text{Lat}(\mathcal{L})$

So ...

Now, consider

So $\begin{array}{c|c|c} & W & Z \\ \hline W & & \\ \hline Z & 0 & \end{array}$ $\Leftrightarrow$ $W \in Lat(\mathcal{L})$

Observe $\begin{array}{c|c|c} & W & Z \\ \hline W & & 0 \\ \hline Z & & \end{array}$ $\Leftrightarrow$ $Z \in Lat(\mathcal{L})$

So if $W, Z \in Lat(\mathcal{L})$, $W \oplus Z = V$

Then $\begin{array}{c|c|c} & W & Z \\ \hline W & A & 0 \\ \hline Z & 0 & B \end{array}$ $\rightarrow$ in essence, there's no mixing W & Z

$\boxed{\text{Q: how do we know there's an invariant subspace to start?}}$

$\mathcal{L}: V \longmapsto V \rightarrow$ fin-dim (m)

How do I find a good invariant subspace.

We can construct one.. For $v_0 \neq 0$, we want $v_0 \in W \in Lat(\mathcal{L})$ small

Want $v_0 \in W$

$\mathcal{L}(v_0) \in W$ — If $m = \dim(V)$

$\mathcal{L}^2(v_0) \in W$ — then there are at most m

$\mathcal{L}^3(v_0) \in W$ — things...

$\vdots$

$\mathcal{L}^n(v_0) \in W$

Say $\mathcal{L}^k(v_0)$ is the 1st one that is a lin. combo. of previous ones, i.e.

$$\mathcal{L}^n(v_0) = \sum_0^{n-1} a_i \mathcal{L}^i(v_0)$$

Then $\mathcal{L}^{n+1}(v_0) = \mathcal{L}\left(\sum a_i \mathcal{L}^n(v_0)\right)$

$= \sum \mathcal{L}^{i+1}(v_0)\, a_i \in Span\left(\mathcal{L}^i(v_0)\right)_{i=1 \cdots n-1}$

So other way &

Consider $(\mathcal{L}^3 - a_0\mathcal{L}^0 - a_1\mathcal{L}^1 - a_2\mathcal{L}^2)\, v_0 = 0$ → complex r_i

So $(\mathcal{L} - r_1 I)(\mathcal{L} - r_2 I)(\mathcal{L} - r_3 I)(v_0) = 0$

but since $v_0 \neq 0$

At least one of $\mathcal{L} - r_1 I$, $\mathcal{L} - r_2 I$, $\mathcal{L} - r_3 I$ is NOT injective.

If $\mathcal{L} - r_i I$ not injective

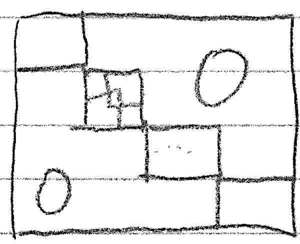
$\Rightarrow \mathcal{L}(\hat{v}) = r_i \hat{v}$ for some $\hat{v} \neq 0$

eigenvalue　eigenvector

$\boxed{\text{We have just proved, every linear fn on fin dim vector space (on } \mathbb{C}\text{) has an eigenvector!}}$

→ $\exists$ 1 dim invariant subspace.

So we can always have $\{0\} \neq W \subsetneq V$



Mar 7, 2019

Idea

$\mathbb{P} \longmapsto$ "$\mathbb{P}(L)$"

$c_0 z^0 + c_1 z^1 + \dots + c_n z^n$ $\qquad$ $c_0 I + c_1 \mathcal{L} + \dots + c_n \mathcal{L}^n$

$(\mathcal{L} : V \xrightarrow{\text{lin}} V)$

$$\mathbb{P} \xrightarrow{P_{\mathcal{L}}} \mathcal{L}(V)$$

what sort of properties does this map have?

$\boxed{\text{Properties of } \mathbb{P} \xrightarrow{\rho_{\mathcal{L}}} \mathcal{L}(V)}$

1. Linear

2. Multiplicative (product of ~~linear~~ polynomial → products of lin. fn)

Consider something from previous proof

$$\rightarrow \left(-a_0 I - a_1 \mathcal{L} - \dots + \mathcal{L}^n\right)(v_0) = \theta_V$$

So $\quad \left(P(L)\right)(v_0) = \theta_V$

By Fundamental theorem of algebra,

unique ... → $P(z) = \underbrace{(z-\lambda_1)^{\mu_1} \dots (z-\lambda_k)^{\mu_k}}_{\text{distinct}}$, λ_k are the roots, μ_k multiplicity of root λ_k

So since $\rho_{\mathcal{L}}$ multiplicative

Hence $\quad (\mathcal{L}-\lambda_1 I)^{\mu_1} \dots (\mathcal{L}-\lambda_k I)^{\mu_k}$

So $\quad (\mathcal{L}-\lambda_1 I)^{\mu_1} \dots (\mathcal{L}-\lambda_k I)^{\mu_k}(v_0) = 0$

i.e. $\quad \underbrace{(\mathcal{L}-\lambda_1 I_1) \dots (\mathcal{L}-\lambda_1 I)}_{\mu_1} \dots \underbrace{(\mathcal{L}-\lambda_k I) \dots (\mathcal{L}-\lambda_k I)}_{\mu_k}(v_0) = \theta_V$

So $\quad (\mathcal{L}-\lambda_1 I_1) \dots (\mathcal{L}-\lambda_1 I) \dots (\mathcal{L}-\lambda_k I) \dots (\mathcal{L}-\lambda_k I)$
not injective for at least one i,

i.e. there exists some $v_i \neq 0$ such that $(\mathcal{L}-\lambda_i I)(v_i) = \theta_V$

or $\quad \mathcal{L}(v_i) = \lambda_i v_i$

So one of λ_i is an eigenvalue of $\mathcal{L}$.

(★) $\boxed{\text{Every } \mathcal{L} : \mathcal{V} \xrightarrow{\text{lin}} \mathcal{V} \text{ has an eigenvalue}}$ finite dim

$\boxed{\text{Every } \mathcal{L} : \mathcal{V} \xrightarrow{\text{lin}} \mathcal{V} \text{ has a 1-dimensional invariant subspace}}$ fin dim

Consider $V = W_1 \oplus \cancel{Z_2}$ $\qquad \mathcal{V}$ 2-dimensional

(1-dim) $\quad \cap \; Lat(\mathcal{L})$

1-dim

Then

$$\mathcal{L} = \begin{array}{c|cc} & W_1 & \cancel{Z_2} \\ \hline W_1 & A & B \\ \cancel{Z_2} & 0 & D \end{array} = \begin{pmatrix} "\lambda" & B \\ 0 & D \end{pmatrix}$$

What about D? $\quad D : Z \to Z$ or ~~W~~ W_2

If W_2 not one-dimensional then $D : \cancel{Z_2} \longmapsto \cancel{Z_2}$, then there is an $\cancel{Z_2} = W_2 \oplus \hat{W}_4$

1-dim invariant, $\in Lat(D)$

Then

$$\mathcal{L} = \begin{array}{c|cc} & W_1 & Z \\ \hline W_1 & \lambda & B \\ Z & 0 & D \end{array} = \begin{array}{c|ccc} & W_1 & W_2 & \\ \hline W_1 & "\lambda_1" & B_1 & B_2 \\ W_2 & 0 & "\lambda_2" & \\ \vdots & 0 & 0 & \hat{D} \end{array} = \ldots \quad \boxed{\text{UPPER TRIANGULAR MATRIX}}$$

$\boxed{\text{Schur's Theorem}}$ $\quad \boxed{\text{Every } n \times n \text{ matrix is } \underline{\text{similar}} \text{ to an upper triangular matrix}}$

Proof by induction $\quad$ Base case $n = 1 \to$ trivial

* Suppose n_0 is smallest size for which there is a counter example A $(n_0 \geq 2)$

- Now, A lin $\to$ A has an eigenvector by (★), called w_0 and Span(w_0) is invariant under A. Then

$\mathbb{C}^{n_0} = \mathrm{Span}(w) \oplus Z$

w.r.t to this decomposition, A has the form
$$\begin{array}{c|c|c} & W & Z \\ \hline W & \lambda & B \\ \hline Z & 0 & D \end{array}$$

Pick a basis of Z: $\overbrace{w_1 \ldots w_{n_0-1}}^{\Gamma_2}, \overbrace{w_0}^{\Gamma_1} = \Gamma$

$w_0 \ldots w_{n_0-1}$

A $\quad w_0 \ \vdots \ w_{n_0-1}$

Recall $[A]_\Gamma = \begin{bmatrix} [A(w_0)]_\Gamma & \cdots & [A(w_{n_0-1})]_\Gamma \end{bmatrix} = \left[\begin{array}{c|c} \lambda & \\ \hline 0 & \\ 0 & B \\ 0 & \\ 0 & \end{array}\right]$

So
$$[A]_\Gamma = \begin{array}{cc} & \begin{array}{cc} (1) & (n_0-1) \\ \Gamma_1 & \Gamma_2 \end{array} \\ \begin{array}{c} (1)\ \Gamma_1 \\ (n_0-1)\ \Gamma_2 \end{array} & \left[\begin{array}{c|c} \lambda & \ast \\ \hline 0 & [T]_{\Gamma_2 \to \Gamma_2} \end{array}\right] \end{array}$$

A is the smallest counter exp

So $A \sim [A]_\Gamma$ $\qquad$ Since B smaller than A, $B = S \circ T \circ S^{-1}$ for some upper Δ^r, T

But A is a counter example →

So
$$A \sim \left[\begin{array}{c|c} \lambda & M \\ \hline 0 & S^{-1}TS \end{array}\right] = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & S^{-1} \end{array}\right] \left[\begin{array}{c|c} \lambda & MS^{-1} \\ \hline & T \\ & 0 \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & S \end{array}\right]$$

inverses

So
$$A \sim \left[\begin{array}{c|c} \lambda & MS^{-1} \\ \hline & T \end{array}\right]$$ ~ upper triangular → contradiction

Now, back to polynomials ...

(1) For each $v_0 \in V$, $v_0 \neq 0_v$, there is some $P \neq 0_p \in \mathbb{P}$ such that

$$(P(L))(v_0) = 0_v$$

Consider $P \xmapsto{P_L} P(L)$

$$\{ P \in \mathbb{P} \mid (P(L))(v_0) = 0_v \} \subseteq \mathbb{P}$$

Closed under scaling — closed under scaling

if $(P(L))(v_0) = 0$ then $(\beta \cdot P(L))(v_0) = 0$

~~closed under multiplication~~ absorption property w.r.t multiplication

in particular

$$(q(L) \circ P(L))(v_0) = 0_v \text{ for any } q \in \mathbb{P}$$

i.e

$$(q \cdot P)(L)(v_0) = 0_v \text{ for any } q \in \mathbb{P}$$

closed under addition — closed under addition

$\boxed{\text{Ideals}}$ ← subspace that is closed under multiplication

There's the smallest unique monic polynomial

$\exists$ unique $P(L)$ smallest that kills (v_0)

Now, each v_0, can associate a polynomial to it.

Is there a polynomial that kills them all?

Consider
$$\begin{array}{ccc} V & \xrightarrow{L} & V \\ & \Big\downarrow \text{linear + multiplicative} & \\ \mathbb{C}^n & \xrightarrow{[L]_{\Gamma \in \Gamma}} & \mathbb{C}^n \end{array}$$

$$[P(L)]_{\Gamma \in \Gamma} = P\left([L]_{\Gamma \in \Gamma}\right)$$

Recall $[L^2]_{\Gamma \in \Gamma} = [L \circ L]_{\Gamma \in \Gamma}$

$$= [L]_{\Gamma \in \Gamma} \circ [L]_{\Gamma \in \Gamma} = [L]^2_{\Gamma \in \Gamma}$$

Can ~~also~~ ~~add~~ $[\mathcal{L}^2+\mathcal{L}]_{\Gamma\in\Gamma} = [\mathcal{L}]^2_{\Gamma\in\Gamma} + [\mathcal{L}]_{\Gamma\in\Gamma}$

Leads to this

$$[P(\mathcal{L})]_{\Gamma\in\Gamma} = P\left([\mathcal{L}]_{\Gamma\in\Gamma}\right)$$

To kill linear fn $\to$ kill matrix $\in \mathbb{M}_{n\times n}$

But a matrix is just a $v_0 \in V$

So something like $A^k - a_0 A^0 - \cdots = \mathcal{O}_{n\times n}$

Look at $\{P \in \mathbb{P} \mid (P(A)) = \mathcal{O}_{n\times n}\}$

↳ an ideal again ... (

↳ $\exists$ single minimal monic element that kills A ...

But if $P(A) = \mathcal{O}_{n\times n}$ then $(P(A))(v_1) = \mathcal{O}_V$

So $P(A) \in \{P \in \mathbb{P} \mid (P(A))(v_0) = \mathcal{O}_V\}$

So all $P(A)(v_0)$ are divisors ~~of~~ $P(A)$ when $P(A) = \mathcal{O}_{n\times n}$

Mar 12, 2019

Review Let $M \in M_n = n^2$ - dimensional vector space (k)

Look at $I, M, M^2, M^3, \ldots$ eventually come to the first power of M such that ~~that~~ $\to$ smallest $k_0 \in \mathbb{N}$ such that M^{k_0} is a lin. combination of $I, M, M^2, \ldots M^{k_0-1}$

Note $k_0 \leq n^2$ | So $M^{k_0} = \sum_{i}^{k_0-1} a_i M^i$ So

So $M^{k_0} + (-a_{k_0-1})M^{k_0-1} + (-a_{k_0-2})M^{k_0-2} + \dots + (-a_0)I = \mathcal{O} = P(M)$

↳ This polynomial annihilates matrix M. In fact, it is the smallest degree (monic) polynomial to do so.

Now

for any matrix we can do this → now look at the set of all polys that annihilate M.

$$W = \{ p \in \mathbb{P} \mid P(M) = \mathcal{O} \}$$

This is a vector space. Recall: $\mathbb{P} \xrightarrow[\text{linear, multiplicative}]{f_M} \mathbb{M}_n$

(non-trivial) subspace absorption property $p \longmapsto p(M)$

So W is also an ideal. Why? → singly generated.

↳ We say that it is generated by a single core monic polynomial.

Called: $\boxed{\mu_M \to \text{the minimal polynomial of } M.}$

global

Recall that

$$Z = \{ p \in \mathbb{P} \mid (P(M))(v_0) = \mathcal{O} \} \text{ is also an ideal of } \mathbb{P}$$

[Here $M \in \mathbb{M}_n$, $\mathcal{O} \neq v_0 \in \mathbb{C}^n$]

Local → and we have the minimal poly μ_{M,v_0} as the generator of this ideal.

μ_M is a multiple of μ_{M,v_0} ...

Observe that if $(\mu_M(M))(v_0) = \mathcal{O}(v_0) = \mathcal{O}$

So the global μ_M is in Z, generated by μ_{M,v_0}

So, μ_{M,v_0} divides μ_M

So $Z \subseteq W$

Also note that a polynomial has at most finitely many monic divisors

Suppose $\mu_M(z) = (z - r_1)^{\gamma_1} \cdots (z - r_k)^{\gamma_k}$

Consider $\mu_{M,v_1}(z), \dots \mu_{M,v_{135}}(z)$

Consider $\left\{ v \,\middle|\, (\mu_{M,v_7}(M))(v) = 0 \right\} = \ker\left(\mu_{M,v_7}(M)\right)$

$= \text{subspace}$

The union of these 135 subspaces is

→ the whole space. So, by PSET 2, one of these subspaces equals the whole union, i.e. equals the whole space.

So $\mu_{M,v_i}(M) = 0$ for at least one i

i.e. μ_{M,v_i} annihilates M

But μ_{M,v_i} divides $\mu_M = q \cdot \mu_{M,v_i}$

So $\boxed{\mu_{M,v_i} = \mu_M \text{ for at least one } i}$

Why do we want to look at these polynomials?

(1) Suppose $\mu_A(z) = (z - \alpha_1)^{\gamma_1} \cdots (z - \alpha_k)^{\gamma_k}$

Then $(A - \alpha_1 I)^{\gamma_1} \circ \dots \circ (A - \alpha_k I)^{\gamma_k} = 0$

But none of these factors can be invertible (proof by contradiction)

So ~~$\forall i\ (A - \alpha_i I)$~~ So $(A - \lambda I)v = 0 \Rightarrow A(v) = \lambda v \to v$ eigenvector
$v \neq 0$

So $\boxed{\text{All roots of } \mu_A \text{ are eigenvalues of } A}$ Theorem/2

? <u>Are there other eigenvalues that are NOT roots of μ_A?</u>

<u>Consider</u> $(z-\lambda) \to$ eigen- of A and $\lambda \neq \alpha_i$ NO

<u>So</u> $\mu_A(z)$ and $(z-\lambda)$ relatively prime. $\to \exists\, p, q$ such that

$$p(z)(z-\lambda) + q(z)\mu_A(z) = \mathbb{1}$$

<u>So</u> $p(A)(A-\lambda I) + \underbrace{q(A)(\mu_A(A))}_{\Theta} = I$

<u>So</u> $p(A)(A-\lambda I) = I$ $\quad$ ker $(A-\lambda I) = \{\Theta\}$

<u>So</u> $(A-\lambda I)$ invertible $\to$ λ not eigenvalue of A
$(p^{-1}(A) = (A-\lambda I))$

<u>So</u> $\boxed{\text{All roots of } \mu_A \text{ are } \underline{\text{exactly}} \text{ eigenvalues of } A}$

much easier than finding eigenvalues using characteristic polynomials since the degrees are often much smaller.

//

<u>Recap</u>

$\mu_A(z)$ $\qquad$ $\cancel{\mu_{A,v}(z)}$ $\quad$ characteristic polynomial

$I, A, A^2, \ldots$ $\qquad$ $v, Av, A^2v, \ldots$

at most n^2 $\qquad$ at most n

<u>But</u> $\mu_{A,v}(z)$ divides $\mu_A(z)$

<u>So</u> <u>Thm.</u> $\boxed{\text{For } A \in M_n,\ \deg(\mu_A) \leq n}$

Block-diagonal matrices

Observe $T = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$ $T^2 = \begin{pmatrix} A^2 & 0 \\ 0 & D^2 \end{pmatrix}$ $\cdots$ $T^0 = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$

So $P(T) = \begin{pmatrix} P(A) & 0 \\ 0 & P(D) \end{pmatrix}$

What is μ_T?

P annihilates $T \iff P$ annihilates A and D
$\iff \mu_A | P$ and $\mu_D | P$
$\iff P$ is a common multiple of μ_A, μ_D

So $\boxed{\mu_T = LCM(\mu_A, \mu_D)}$

For a diagonal matrix, $T = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_k \end{pmatrix}$

So $\mu_T = LCM(z-\alpha_1, z-\alpha_2, \ldots, z-\alpha_k)$

some α_i can be repeated

So $\boxed{\mu_T = LCM(\;) = (z-\lambda_1)\ldots(z-\lambda_k)}$ → power 1

where $\lambda_i \neq \lambda_j \;\forall\, i,j$

VERY IMPORTANT

$\boxed{?}$ Diagonalizability

Recall (we shown)

$$\begin{tikzcd} {}_\Gamma V \arrow[r, "\mathcal{L}"] \arrow[d, shift left] & V_\Gamma \arrow[d, shift left] \\ \mathbb{C}^n \arrow[u, shift left] \arrow[r, "{[\mathcal{L}]}"'] & \mathbb{C}^n \arrow[u, shift left] \end{tikzcd}$$

lin / multi.

$\Rightarrow \boxed{[P(\mathcal{L})]_{\Gamma\leftarrow\Gamma} = P([\mathcal{L}]_{\Gamma\leftarrow\Gamma})}$

i.e. $\boxed{\text{Similar matrices have the same minimal poly.}}$

"Polynomials belong to linear fn, not the basis/matrix"

So, for $P([\mathcal{L}]_{\Gamma\leftarrow\Gamma}) = 0$, $P(\mathcal{L}) = 0$

If $A \in M_n$ and $a_0 I + a_1 A + \dots + a_k A^k = 0$

Apply transposition $\to$ $\sum (a_i A^i)^T$

$\hookrightarrow \sum_k^i a_k (A^k)^T = \sum_k^i a_k (A^T)^k = 0$

So if $P(A) = 0$, then $P(A^T) = 0$

But if $P(A^T) = 0$, then $P(A) = 0$

So $\boxed{\mu_{A^T} = \mu_A}$

---

what abt $T = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \rightsquigarrow T^2 = \begin{pmatrix} A^2 & \\ 0 & D^2 \end{pmatrix} \dots$

$$T^0 = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

So $P(T) = \begin{pmatrix} P(A) & \\ 0 & P(D) \end{pmatrix}$

not least

only 1 direction $\rightarrow$ So $\left\{ P(T) \text{ annihilates } T \Rightarrow P(A) = P(D) = 0 \right.$

$\Rightarrow P = $ common multiple of μ_A, μ_D

So $(\mu_A \cdot \mu_D)(T) = \mu_A(T) \cdot \mu_D(T)$

$$= \begin{pmatrix} 0 & \\ 0 & \mu_A(D) \end{pmatrix} \cdot \begin{pmatrix} \mu_D(A) & \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

So $\mu_A \cdot \mu_D$ annihilates T So $\left.\begin{matrix}\mu_A \\ \mu_D\end{matrix}\right| T \,|\, \mu_A \mu_D$

Therefore, $\boxed{LCM(\mu_A, \mu_D) \,|\, T \,|\, \mu_A \mu_D}$

if $\mu(A), \mu_D$ relatively prime ... then "=" $\mu_A \mu_D$

Mar 14, 2019 **Review**

1. Minimal Polys
2. Eigenvalues are exactly the roots of the minimal polynomial
3. Local v Global min poly
   (a) Local divides Global
   (b) One of locals is global
4. Similar matrices have the same minimal poly.
   $(\mu_L = \mu_{[L]_{B \leftarrow \Gamma}})$
5. $\boxed{\mu_A = \mu_{A^T} \Rightarrow \sigma_{\mathbb{C}}(T) = \sigma_{\mathbb{C}}(T^T)}$ — same T & same A & B
6. if $T = \begin{bmatrix} A & O \\ O & B \end{bmatrix}$ then $\mu_T = LCM(\mu_A, \mu_B)$ — gap due to multiplicity
7. if $T = \begin{bmatrix} A & D \\ O & B \end{bmatrix}$ then $\boxed{LCM(\mu_A, \mu_B) \mid \mu_T \mid \mu_A \mu_B}$

   (b) $\boxed{\sigma_{\mathbb{C}}(T) = \sigma_{\mathbb{C}}(A) \cup \sigma_{\mathbb{C}}(B)}$

   lin factors of T are exactly those that appear in μ_A or μ_B (or both)

   (a) if μ_A, μ_B relatively prime $\Rightarrow$ $\boxed{LCM(\mu_A, \mu_B) = \mu_A \mu_B = \mu_T}$

(6a) μ_T of $T = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$ is a product of $(x-\alpha_1)\cdots(x-\alpha_n)$ — distinct

Ex: $T = \begin{pmatrix} 3 & & & \\ & 4 & & \\ & & 1 & \\ & & & 3 \end{pmatrix} \rightarrow \mu_T(z) = (z-3)(z-4)(z-1)$

$\hookrightarrow \sigma_{\mathbb{C}}(T) = \{1, 3, 4\}$

What if

$\begin{pmatrix} A & D \\ O & B \end{pmatrix} = T = \left(\begin{array}{cc|cc} 3 & 1 & 2 & 8 \\ & 4 & 3 & 1 \\ \hline & & 2 & 7 \\ & & & 3 \end{array}\right) \Rightarrow \sigma_{\mathbb{C}}(T) = \sigma_{\mathbb{C}}\left[\begin{pmatrix} 3 & 1 \\ 0 & 4 \end{pmatrix}\right] \cup \sigma_{\mathbb{C}}\left[\begin{pmatrix} 2 & 7 \\ 0 & 3 \end{pmatrix}\right]$

$= \sigma_{\mathbb{C}}([3]) \cup \sigma_{\mathbb{C}}([4]) \cup \sigma_{\mathbb{C}}([2]) \cup \sigma_{\mathbb{C}}([3])$

$= \{3, 4, 2\} \longrightarrow$ just the diagonal!

So what is $\mu_T(z)$?

$\mu_T(z) = (z-2)^{1}(z-3)^{1\text{ or }2}(z-4)^{1}$

we $LCM(\mu_A, \mu_B) \mid \mu_T \mid \mu_A \mu_B$ {

$\boxed{\text{Observe}}$ If $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $p(z) = (z-a)(z-d) - bc$ then p annihilates T

Proof $p(T) = (T - aI)(T - dI) - bcI$ ⟶ We can show this is the minimal poly if....

$$= \begin{pmatrix} 0 & b \\ c & d-a \end{pmatrix}\begin{pmatrix} a-d & b \\ c & 0 \end{pmatrix} \underbrace{- bcI} = \begin{pmatrix} bc & 0 \\ 0 & bc \end{pmatrix} - \begin{pmatrix} bc & 0 \\ 0 & bc \end{pmatrix} = [0]$$

Thm $\boxed{\text{if } T = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ is not a multiple of } I_2 \text{, then } \mu_T = (z-a)(z-d) - bc}$

Back to example

$\mu_A = \mu_{\left[\begin{smallmatrix} 3 & 1 \\ 0 & 4 \end{smallmatrix}\right]}(z) = (z-3)(z-4)$

$LCM = (z-2)(z-3)(z-4)$

$\mu_B = \mu_{\left[\begin{smallmatrix} 2 & 2 \\ 0 & 3 \end{smallmatrix}\right]}(z) = (z-2)(z-3)$

So $(z-2)(z-3)(z-4) \mid \mu_T \mid (z-2)(z-3)^2(z-4)$

So $\mu_T(z) = (z-2)(z-3)^{1 \text{ or } 2}(z-4)$

$\boxed{\text{Question}}$ $\begin{pmatrix} 3 & 1 \\ 0 & 4 \end{pmatrix} \overset{?}{\rightsquigarrow} \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$ or $\begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}$

Well $\begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix} \sim \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}$ So $\begin{pmatrix} 3 & 1 \\ 0 & 4 \end{pmatrix} \overset{?}{\sim} \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$ Yes.

Why? Same minimal polynomial. $(z-3)(z-4)$

■ Suppose p_1, p_2, relatively prime.

Then $\ker\left(p_1(L)\right) \cap \ker\left(p_2(L)\right) = \{\theta\}$

"Proof" $\rightsquigarrow$ $\underline{\exists\ q_1, q_2 \text{ s.t.}}$ $q_1(z)\,p_1(z) + q_2(z)\,p_2(z) = \mathbb{1}$

So $q_1(L)\,p_1(L) + q_2(L)\,p_2(L) = I$

So $\left(q_1(L)\,p_1(L) + q_2(L)\,p_2(L)\right)(v) = v$

So $q_1(L)\left[\left(p_1(L)\right)(v)\right] + q_2(L)\left[\left(p_2(L)\right)(v)\right] = v$

if $v \in \ker(p_1(L)) \cap \ker(p_2(L))$ then

$$q_1(L)\,\vec{\theta} + q_2(L)\,\vec{\theta} = v \Rightarrow v = \vec{\theta}$$

So $\boxed{\{\theta\} = \ker(p_1(L)) \cap \ker(p_2(L))}$ if p_1, p_2 relatively prime

$\boxed{\text{Observe}}$

If $\mu_A(z) = \underbrace{(z-\lambda_1)^{k_1}}_{p_1} \cdots \underbrace{(z-\lambda_m)^{k_m}}_{p_2}$

Since p_1, p_2 relatively prime $\Rightarrow$ $\ker(p_1(A)) \cap \ker(p_2(A)) = \{\theta\}$

Also $p_1(A)\,p_2(A) = \mu_A(A) = \theta = p_2(A)\,p_1(A)$

so $\begin{cases} \operatorname{Im}(p_1(A)) \subseteq \ker(p_2(A)) \\ \operatorname{Im}(p_2(A)) \subseteq \ker(p_1(A)) \end{cases}$

Yet also $q_1(z)\,p_1(z) + q_2(z)\,p_2(z) = 1$ for some q_1, q_2

So $p_1(z)\cdot q_1(z) + p_2(z)\cdot q_2(z) = 1$

So $p_1(A)\,q_1(A) + p_2(A)\,q_2(A) = I$

hence, $\underbrace{p_1(A)\big((q_1(A))(y)\big)}_{\in\, \operatorname{Im}(p_1(A)) \subset \ker(p_2(A))} + \underbrace{p_2(A)\big((q_2(A))(y)\big)}_{\in\, \operatorname{Im}(p_2(A)) \subset \ker(p_1(A))} = y \quad \forall y \in V$

So

$$\ker(p_1(A)) + \ker(p_2(A)) = V$$

But since $\ker(p_2(A)) \cap \ker(p_1(A)) = \{0\}$

So $\boxed{V = \ker(p_1(A)) \oplus \ker(p_2(A))}$ (*)

Observe that $p_1(A) \leftrightarrow A$, ($p_1(A)$ commutes with A)

So $\ker(p_1(A)) \in Lat(A)$

Similarly, $\ker(p_2(A)) \in Lat(A)$

So, with respect to the decomposition, A can be expressed as

$$A = \boxed{\begin{matrix} C & 0 \\ 0 & D \end{matrix}} = \begin{array}{c|c|c} & \ker(p_1(A)) & \ker(p_2(A)) \\ \hline \ker(p_1(A)) & C & 0 \\ \hline \ker(p_2(A)) & 0 & D \end{array}$$

Observe what are μ_C & μ_D?? $\left.\begin{matrix}\mu_C = p_1 \\ \mu_D = p_2\end{matrix}\right\} \to$ so we keep going

Primary decomposition

Nov 19, 2019

Lemma

Summarize. If q_1, q_2 relatively prime and $q_1 \cdot q_2$ annihilates $A : V \mapsto V$ then $V = \ker(q_1(A)) \oplus \ker(q_2(A))$ and w.r.t this decomp A has the form

$$A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix} \quad \begin{cases} q_1 \text{ annihilates } C \\ q_2 \text{ annihilates } D \end{cases}$$

"Proof Sketch" $\to$ (1) $q_1(A)\, q_2(A) = 0 = q_2(A)\, q_1(A) \Rightarrow \begin{cases} \operatorname{Im}(q_2(A)) \subset \ker q_1 \\ \operatorname{Im}(q_1(A)) \subset \ker q_2 \end{cases}$

(2) $\exists\ f, g$ $fq_1 + gq_2 = \mathbb{1} \Rightarrow q_1(A)f(A) + q_2(A)g(A) = I$

So $\underbrace{q_1(A)f(A)(x)} + q_2(A)g(A)(x) = x$

$$\underbrace{q_1(A)f(A)(x)}_{\substack{\cap \\ \operatorname{Im} q_1(A) \\ \cap \\ \ker(q_2(A))}} + \underbrace{q_2(A)g(A)(x)}_{\substack{\cap \\ \operatorname{Im} q_2(A) \\ \cap \\ \ker(q_1(A))}} = x$$

So $\ker(q_1(A)) + \ker(q_2(A)) = V$

(3) for any $y \neq 0$, $\underbrace{f(A)q_1(A)(y)} + \underbrace{g(A)q_2(A)(y)} = y$

cannot both be zero unless $y = 0$

$\Rightarrow \ker(q_1(A)) \cap \ker(q_2(A)) = \{0\}$

So $\ker(q_1(A)) \oplus \ker(q_2(A)) = V$

(4) $\left.\begin{array}{l} \ker(q_1(A)) \in \operatorname{Lat}(A) \\ \ker(q_2(A)) \in \operatorname{Lat}(A) \end{array}\right\}$

Hence wrt $V = \ker(q_1(A)) \oplus \ker(q_2(A))$,

$$A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix} \Rightarrow q_1(A) = \begin{pmatrix} q_1(B) & 0 \\ 0 & q_1(C) \end{pmatrix}$$

So,

$$A\begin{pmatrix} x \\ y \end{pmatrix}_+ = \begin{pmatrix} Bx \\ Cy \end{pmatrix}_+ = \underset{\substack{\ker(q_1(A)) \\ \Downarrow}}{Bx} + \underset{\substack{\ker(q_2(A)) \\ \Downarrow}}{Cy}$$

• $q_1(B) : \ker(p_1(A)) \mapsto \ker(p_1(A))$

Now

$$q_1(A)\begin{pmatrix} x \\ y \end{pmatrix}_+ = q_1(B)(x) + q_1(C)(y)$$

But $x \in \ker(q_1(A)) \Rightarrow q_1(B)(x) + q_1(C)(y) = q_1(A)(y)$

when $y = 0 \Rightarrow \boxed{q_1(B)(x) = 0}\ \forall x \in \ker(q_1(A))$

$\underline{\text{So}}$ $q_1(B) = 0$: $\ker(q_1(A)) \longmapsto \ker(q_1(A))$

$\boxtimes$ $\underline{\text{So}}$ q_1 annihilates B

Same argument $\Rightarrow$ q_2 annihilates C

Corollary

$\boxed{\begin{array}{l}\text{Under the hypotheses of the previous lemma} \\ \text{If } q_1, q_2 \text{ are monic and } \mu_A = q_1 \cdot q_2 \text{ then} \\ q_1 = \mu_B,\ q_2 = \mu_C\end{array}}$

$\underline{\text{"Proof"}}$ $\mu_A = \mu_{\boxed{\begin{smallmatrix} B & 0 \\ 0 & C \end{smallmatrix}}} = LCM(\mu_B, \mu_C)$

We know that μ_B divides q_1 since q_1 annihilates B

μ_C divides q_2 since q_2 annihilates C

But since q_1, q_2 relatively prime $\Rightarrow$ μ_B, μ_C relatively prime

$\Rightarrow$ $LCM(\mu_B, \mu_C) = \mu_B \cdot \mu_C$ $\underline{\text{So}}$ $\mu_A = \mu_B \cdot \mu_C$

But we also know $\mu_A = q_1 \cdot q_2 = \mu_B \cdot \mu_C$ and $\mu_B | q_1$, $\mu_C | q_2$ and that q_1, q_2 are monic

$\underline{\text{So}}$ $q_1 = \mu_B$ and $q_2 = \mu_C$

//

$\boxed{\text{Primary Decomposition Theorem}}$

Suppose $A : V \to V$ and $\mu_A(z) = (z - \lambda_1)^{p_1} \dots (z - \lambda_k)^{p_k}$ (distinct), then

$$V = \ker(A - \lambda_1 I)^{p_1} \oplus \dots \oplus \ker(A - \lambda_k I)^{p_k}$$

and wrt this decomposition, A has the form — $\boxed{\begin{matrix} A_1 & & \\ & A_2 & \\ & & \ddots \\ & & & A_k \end{matrix}}$

where

$\mu_{A_i} = (z - \lambda_i)^{p_i}$

"Proof" Induct on k.

Base case "$k=1$" holds true trivially

Inductive hypothesis → Suppose "$k = 1, 2, \ldots, m$" hold

Show Case "$k = m+1$" holds.

Then $\mu_A(z) = \underbrace{(z-\lambda_1)^{p_1} \cdots (z-\lambda_m)^{p_m}}_{q_1} \underbrace{(z-\lambda_{m+1})^{p_{m+1}}}_{q_2}$

By the lemma + cor we have $A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$ w.r.t $V = \overbrace{\ker(q_1(A))}^{W_1} \oplus \underbrace{\ker(q_2(A))}_{W_2}$

and $\mu_B = q_1$, $\mu_C = q_2$

By inductive hyp, $\underset{\text{"new V"}}{\ker(q_1(A))} = \ker(B-\lambda_1 I)^{p_1} \oplus \cdots \oplus \ker(B-\lambda_m I)^{p_m}$

new $V = \ker(q_1(A)) = W_1$

and

$$B = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{pmatrix}$$ w.r.t this decomposition, with $\mu_{B_i}(z) = (z-\lambda_i)^{p_i}$

By assignment (?), A can be represented by

$$A = \left(\begin{array}{ccc|c} B_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & B_m & \\ \hline & 0 & & C \end{array}\right)$$ w.r.t the decomposition

$$V = \overbrace{\ker(B-\lambda_1 I)^{p_1} \oplus \cdots \oplus \ker(B-\lambda_m I)^{p_m}}^{\ker(q_1(A))} \oplus \ker(A-\lambda_{m+1} I)$$

So, the next step is to show $\ker(\underline{B}-\lambda_i I)^{p_i} = \ker(\underline{A}-\lambda_i I)^{p_i}$

Now, $\ker(A-\lambda_i I)^{p_i} \subset \ker\big(\underbrace{(A-\lambda_1 I)^{p_1} \cdots (A-\lambda_m I)^{p_m}}_{\text{commute} \ldots}\big)$ since

Such that I can write $(A-\lambda_1 I)^{p_1} \cdots \overset{\text{move}}{(A-\lambda_m I)^{p_m}} (A-\lambda_i I)^{p_i}$

(1) So $\ker(A-\lambda_i I)^{p_i} \subset \ker(q_1(A)) = W_1$

(2) Similarly, $\ker(B-\lambda_j I)^{p_j} \subset \ker(q_1(A)) = W_1$

Recall

$$A = \begin{array}{c|cc} & W_1 & W_2 \\ \hline W_1 & B & 0 \\ W_2 & 0 & C \end{array}$$

So

$$(A-\lambda_i I)^{p_i} = \begin{array}{c|cc} & W_1 & W_2 \\ \hline W_1 & (B-\lambda_i I)^{p_i} & 0 \\ W_2 & 0 & \text{-Stuff-} \end{array}$$

Now $x \in W_1$ is in $\ker(A-\lambda_i I)^{p_i} \Leftrightarrow (A-\lambda_i I)^{p_i}(x) = 0 \Leftrightarrow \boxed{}\begin{pmatrix} x \\ 0 \end{pmatrix} = 0$

$\Leftrightarrow (B-\lambda_i I)^{p_i}(x) = 0 \Leftrightarrow x \in \ker\left((B-\lambda_i I)^{p_i}\right)$

So $x \in \ker(A-\lambda_i I)^{p_i} \Leftrightarrow x \in \ker\left((B-\lambda_i I)^{p_i}\right)$

So $\boxed{\ker\left((A-\lambda_i I)^{p_i}\right) = \ker\left((B-\lambda_i I)^{p_i}\right)}$ for $i = 1 \ldots m$ //

So, what does this theorem say?

Think if $(A-\lambda I)^p = 0$ and $N = A - \lambda I$, N is nilpotent

$\Leftrightarrow A = \lambda I + N$, N nilpotent

Corollary

→ Every $L: V \xrightarrow{\text{lin}} V$, findim can be expressed as

$L = L_1 \oplus L_2 \oplus L_3 \oplus \cdots \oplus L_m$, where $L_i = \lambda_i I + N_i$

distinct nilpotent

NILPOTENTS

Structures of Nilpotents: such that $N^8 = \theta \to \mu_N = z^8$ Global

Given a nilpotent $N : V \longmapsto V$ and $v_0 \in V$, we write

$$\langle v_0 \rangle \equiv \text{span}\{v_0, N(v_0), N^2(v_0), \dots, N^7(v_0)\}$$

$$= \{P(N)(v_0) \mid \deg(P) \le 7\} \in Lat(N)$$

Theorem For any N as above, there exist some $v_1, \dots v_k \in V$ such that

$$V = \langle v_0 \rangle \oplus \langle v_1 \rangle \oplus \dots \oplus \langle v_k \rangle$$

we will look at proof later, but wrt this decomposition, N has the form

$$N = \begin{array}{c} \\ \langle v_0\rangle \\ \langle v_1\rangle \\ \vdots \\ \langle v_k\rangle \end{array} \begin{array}{c} \begin{array}{cccc} \langle v_0\rangle & \langle v_1\rangle & \dots & \langle v_k\rangle \end{array} \\ \begin{pmatrix} * & \theta & \dots & \theta \\ \theta & * & & \vdots \\ \vdots & & \ddots & \theta \\ \theta & \dots & \theta & * \end{pmatrix} \end{array}$$

$\to$ N is block-diagonal

Suppose $\underbrace{v_0, N(v_0), \dots, N^5(v_0)}_{\text{lin. ind}}, N^6(v_0)$

$(v_0, N(v_0) \dots$

$\mathcal{B}_1$

So $N^6(v_0) = \sum_0^5 a_i N^i(v_0)$

So $\underbrace{(N^6 - a^5 N^5 \dots - a_0 I)}(v_0) = \theta$

So $q(N)(v_0) = \theta$ So $q_{(v_0)}(N) = N^k$ but deg at most 6

But local is some z^k, $k < 8$ but deg at least 6

So $N^6 = \theta$

So $(v_0, N(v_0) \ldots N^5(v_0))$ = basis of $\langle v_0 \rangle$

So

$$\begin{array}{c|cccc} & v_0 & N(v_0) & \ldots & N^5(v_0) \\ \hline v_0 & & & & \\ N(v_0) & & & & \\ \vdots & & & & \\ & & & & \\ N^5(v_0) & & & & \end{array}$$

$$N(\alpha) = \begin{array}{c} v_0 \\ N(v_0) \\ \vdots \\ \\ N^5(v_0) \end{array}\begin{pmatrix} 0 & 0 & & 0 \\ 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ 0 & 0 & & 0 \\ 0 & 0 & & 0 \end{pmatrix}$$

Mar 21, 2019

$\boxed{\text{Recall}}$ Primary Decomposition Theorem

If $\mathcal{L}: \mathcal{V} \mapsto \mathcal{V}$ has minimal polynomial $\mu_{\mathcal{L}} = (z-\lambda_1)^{p_1} \ldots (z-\lambda_k)^{p_k}$ then $\mathcal{V} = \ker(\mathcal{L}-\lambda_1 I)^{p_1} \oplus \cdots \oplus \ker(\mathcal{L}-\lambda_k I)^{p_k}$ and wrt this decomposition $\mathcal{L}$ has the form $\mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \ldots \oplus \mathcal{L}_k$ with
$\mu_{\mathcal{L}_i}(z) = (z-\lambda_i)^{p_i}$

$$\mathcal{L} = \begin{pmatrix} \mathcal{L}_1 & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & \mathcal{L}_k \end{pmatrix}$$

1. Next, observe that

$$\mu_A(z) = (z-\alpha)^p \quad \text{then} \quad A = \alpha I - \underbrace{N}_{\text{nilpotent}} \quad \text{where } \mu_N(z) = z^p$$

we say that
"N is a nilpotent of order p."

2. if N is nilpotent of order p. $N: \mathcal{W} \mapsto \mathcal{W}$, and $\mathcal{O} \neq w_0 \in \mathcal{W}$ then $\mu_{N,w_0}(z) = z^m$, for some $m \leq p$.

Consider $\underbrace{\overset{\mathcal{O}}{\overset{\#}{w_0}}, N(w_0), N^2(w_0), \ldots, N^{m-1}(w_0)}_{\text{lin. ind (since otherwise we get another } \mu_N)}, \mathcal{O}, \mathcal{O}, \mathcal{O}, \ldots$

Recall $\langle w_0 \rangle_N = \{ q(N)(w_0) \mid q \in \mathbb{P} \} \leftarrow$ cyclic invariant subspace $\in Lat(N)$

So $= \{ q(N)(w_0) \mid q \in \mathbb{P}_{m-1} \}$ & this invariant ss of N has basis $w_0 \ldots N^{m-1}(w_0)$

Let $\Gamma =$ basis set $w_0, N(w_0), \ldots N^{m-1}(w_0)$.

<u>Note</u>: $N\Big|_{\langle w_0\rangle} : \langle w_0 \rangle \longmapsto \langle w_0 \rangle$ & $\left[N\Big|_{\langle w_0\rangle} \right]_{\Gamma \leftarrow \Gamma} = \begin{bmatrix} 0 & 0 & & & 0 & 0 \\ 1 & 0 & & & & 0 \\ 0 & 1 & \ddots & & & \vdots \\ 0 & & \ddots & \ddots & & \vdots \\ \vdots & & & & 1 & 0 \end{bmatrix}$

Let $\eta : N^{m-1}(w_0) \ldots w_0 \quad N^{m-1} \ldots w_0 \rightarrow$ write basis in reverse

<u>then</u> $\left[N\Big|_{\langle w_0\rangle} \right]_{\eta \leftarrow \eta} = \begin{array}{c} N^{m-1} \\ \\ \\ \\ w_0 \end{array} \begin{bmatrix} 0 & 1 & 0 & & 0 \\ \vdots & 0 & 1 & & \\ & & 0 & \ddots & \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & & & 0 \end{bmatrix}$

<u>Theorem</u> $\boxed{\text{Cyclic Decomposition for Nilpotents}}$

If $\underset{\text{f.d.}}{N : V} \longmapsto \underset{\text{f.d.}}{V}$ nilpotent then $V = \bigoplus_{i=1}^{k} \langle v_i \rangle_N$

for some non-zero v_i.

In fact, we can choose any v_1 to start with.

(and wrt to this decomposition, N has the form $N_1 \oplus N_2 \oplus \ldots \oplus N_k$)

$\boxed{\text{Corollary}}$ If $N : \underset{\text{f.d.}}{V} \mapsto \underset{\text{f.d.}}{V}$ nilpotent then there is a basis of Ψ of V such that $[N]_{\Psi \leftarrow \Psi} = \begin{bmatrix} \boxed{\begin{smallmatrix} 0 & 1 & \\ & \ddots & 1 \\ & & 0 \end{smallmatrix}} & & & & \bigcirc \\ & \boxed{\begin{smallmatrix} 0 & 1 & \\ & \ddots & 1 \\ & & 0 \end{smallmatrix}} & & & \\ & & \boxed{\begin{smallmatrix} 0 & 1 & \\ & \ddots & 1 \\ & & 0 \end{smallmatrix}} & & \\ \bigcirc & & & \boxed{\begin{smallmatrix} 0 & \ddots & 1 \end{smallmatrix}} & \\ & & & & \boxed{0} \end{bmatrix}$

$\boxed{\begin{smallmatrix} 0 & 1 & & \bigcirc \\ & \ddots & \ddots & \\ & & & 1 \\ \bigcirc & & & 0 \end{smallmatrix}}$

{ Ψ is made by concatenation }

may or may not be there

Now $\rightarrow$ "improve", look at

$$[\alpha I + N]_{\Psi \leftarrow \Psi} = \begin{bmatrix} \boxed{\begin{smallmatrix} \alpha & 1 & \\ & \ddots & 1 \\ & & \alpha \end{smallmatrix}} & & & & \bigcirc \\ & \boxed{\begin{smallmatrix} \alpha & 1 & \\ & \ddots & 1 \\ & & \alpha \end{smallmatrix}} & & & \\ \bigcirc & & \ddots & & \\ & & & \boxed{\begin{smallmatrix} \alpha & 1 & \\ & \ddots & 1 \\ & & \alpha \end{smallmatrix}} & \\ & & & & \boxed{\alpha} \end{bmatrix}$$

$$\| $$

$$[\alpha I]_{\Psi \leftarrow \Psi} + [N]_{\Psi \leftarrow \Psi}$$

<u>where</u> each $\boxed{\begin{smallmatrix} \alpha & 1 & \\ & \ddots & 1 \\ & & \alpha \end{smallmatrix}}$ is called a "Jordan block"

" $J_{\alpha, \text{size}}$ " $\leftarrow$ denoted

$\begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \rightarrow J_{\alpha, 1}$

Jordan canonical form theorem

unique up to changing order of appearance

{ Every matrix is similar to a direct sum of Jordan blocks }

Consider $\{0\} \subseteq \ker(N) \subseteq \ker(N^2) \subseteq \ker(N^3) \subseteq \dots \subseteq \ker(N^8) \subseteq \ker(N$

$$N = \begin{pmatrix} J_{0,9} & & \\ & \ddots & \\ & & J_{0,1} \end{pmatrix}$$

8 7 7 6

↑ # of blocks

Given $\{0\} \subseteq \ker(N) \subseteq \dots \subseteq V$

if 10 10 8 6 ....

10 blocks, no $J_{0,1}$ blocks, 2 $J_{0,2}$ blocks, 2 $J_{0,3}$ blocks ...

So we can reconstruct $N = \begin{pmatrix} J & & \\ & \ddots & \\ & & J \end{pmatrix}$ by Weyr characteristic...

↳ So, no matter how decomposition is different, ~~the~~ this Jordan construction is determined by just N.

So Similar N, M nilpotents then they have the same Weyr characteristics → they have the same Jordan form

But we also shown that if N, M have the same Jordan-block decomposition they are similar

Step 1

↓

...how true ...r nilpotents

$\boxed{\text{So}}$ Weyr characteristic of N completely determines the Jordan block sizes and their multiplicities, i.e. completely determines the J-form of N (up to order of appearance)

⇒ Similar matrices have the same Weyr characteristics; so similar nilpotent have the same J-form

Conversely ...

Conversely, if 2 nilpotents have the same J-form, then they are similar to it, and so to each other

$\boxed{\text{So}}$ { The J-form is a "complete invariant" for similarity of nilpotent }

Now, what abt $A = \alpha I + N$?

{ If $A = \alpha I + N$ then the Weyr characteristic of $A - \alpha I$ determines the Jordan form of A. }

What if we don't know α? What if all we know is
$A =$ something $\cdot I$ + nilpotent

→ Use Trace! → $\boxed{\dfrac{Tr(A)}{\text{size of } A}}$ → the trace doesn't change in different representations...

("Trace" is similarity - invariant)

$\boxed{\text{Step 2}}$ Similar ~~(nilpotent)~~ matrices have the same Jordan forms.

(not just restricted to nilpotents)

April 2, 2019

$\boxed{\text{Diagonalizability}}$

① Recall Every $n \times n$ matrix is triangularizable, i.e. is similar to an upper or lower triangular matrix.

Δ-ty

Equivalently / Consequence $L \in \mathcal{L}(V)$ can be represented by Δ^r matrix for some basis

② Which matrices are diagonalizable?

<u>Reverse engineer</u>: If $A \sim D$, D diagonal, then, $\mu_A = \mu_D$.

and $\mu_D = (z - d_{i_1}) \cdots (z - d_{i_k})$, μ_D has no repeated roots

distinct d_i

i.e. all roots have multiplicity 1

$\boxed{\text{Observe}}$

<u>Ex</u> $J_{\lambda,3} = \begin{bmatrix} \lambda & 1 & 0 \\ & \lambda & 1 \\ & & \lambda \end{bmatrix}$ $\quad (z-\lambda) \mid \mu_J(z) \mid (z-\lambda)^3$

<u>Know</u>

$$(J - \lambda I)^3 = 0$$

$$J_{\lambda,3} \Rightarrow \mu_J(z) = (z-\lambda)^3$$

So $\boxed{J_{\lambda,p} \Rightarrow \mu_J(z) = (z-\lambda)^p}$

<u>Now</u>, if $A \sim J_1 \oplus J_2 \oplus \cdots \oplus J_k$ then $\mu_A = LCM(\mu_{J_1}, \ldots, \mu_{J_k})$

So in particular, $\mu_{J_i} \mid \mu_A$ for each i

So if μ_A has no repeated roots, then no μ_{J_i} has repeated roots

i.e. all J_i is 1×1

So $A \sim J_1 \oplus \cdots \oplus J_k$ is similar to a diagonal matrix

$\boxed{\text{Theorem}}$

TFAE

1. A is diagonalizable
2. μ_A has no repeated roots
3. There is a basis of $\mathbb{C}^n$ made up entirely of eigenvectors of A (an A-eigenbasis of $\mathbb{C}^n$)

How to check if μ_A has no repeated roots? $\to$ look at μ_A, μ_A'

if μ_A has repeated roots, $\mu_A(\text{root}) = \mu_A'(\text{root}) = 0$

How to check if μ_A and μ_A' have no common roots? $\Rightarrow$ division algorithm
→ find gcd (...) (or Euclidean algorithm)

If gcd $= \mathbb{1}$ then no common roots...

matrix $A \longrightarrow$ $\boxed{\sigma_{\mathbb{C}}(p(A)) = p[\sigma_{\mathbb{C}}(A)]}$

Spectral Mapping Theorem

Prop
$$\begin{cases} \sigma_{\mathbb{C}}(\alpha(A)) = \alpha[\sigma_{\mathbb{C}}(A)] \\ \sigma_{\mathbb{C}}(A+\beta I) = \sigma_{\mathbb{C}}(A) + \beta \end{cases}$$

Suppose A invertible. what is $\sigma_{\mathbb{C}}(A^{-1})$?

$$A^{-1}(v) = \gamma v \quad , v \neq 0$$

$$\Updownarrow$$

$$1/\gamma \, v = Av \quad (\gamma \neq 0 \text{ since} \ldots$$

Theorem $\boxed{\gamma \in \sigma_{\mathbb{C}}(A^{-1}) \Leftrightarrow 1/\gamma \in \sigma_{\mathbb{C}}(A)}$

So
$$\sigma_{\mathbb{C}}(A^{-1}) = \frac{1}{\sigma_{\mathbb{C}}(A)} = \frac{1}{z}\left[\sigma_{\mathbb{C}}(A)\right]$$

$$\downarrow$$

$$\frac{1}{z}[A]$$

SMT holds for rational fn...

What about simultaneous...? Commute...?

April 4, 2019

$\boxed{\text{Simultaneous } \Delta\text{-ability, Diag-ability}}$

① Sim. Diag-bility

$V \xrightarrow{A, M} V$ Suppose $A \xrightarrow{S}$ diagonal

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad B \xrightarrow{S}$ diagonal

$\mathbb{C}^n \longrightarrow \mathbb{C}^n$ i.e.

$S^{-1}AS$ diag

$S^{-1}BS$ diag

Then

$$S^{-1}ASS^{-1}BS = S^{-1}ABS \iff \boxed{AB = BA}$$

$$S^{-1}BS\, S^{-1}AS = S^{-1}BAS$$

Necessary conditions for simultaneous diagonalizability

(1) Individual diagonalizability

(2) Commutativity

In fact, there are also sufficient.

(†) Theorem TFAE for a collection of $\mathcal{F} \in M_n$

① $\mathcal{F}$ is a commutative collection and all elements $\in \mathcal{F}$ are individually diagonalizable

② There is invertible S such that $S^{-1}AS$ diagonal for every $A \in \mathcal{F}$

$AB = BA$
$BC = CB$

Pf We prove Lemma

A diagonalizable $\iff A = \alpha_1 E_1 + \dots + \alpha_k E_k$ for some idempotents s.t $E_1 + \dots + E_k = I$ (non zero)

$(1 \Rightarrow 2)$ Pf A diag-able $\iff S^{-1}AS$ diag $= \beta_1 F_1 + \dots + \beta_m F_m$, $F_i \to$ idemp $\neq 0$, $F_1 + \dots + F_m = I$

So $A = S\left[\beta_1 F_1 + \dots + \beta_m F_m\right] S^{-1}$

$$= \beta_1 \underbrace{S F_1 S^{-1}}_{\text{idemp}} + \dots + \beta_m S F_m S^{-1}$$

$$= \beta_1 \mathcal{E}_1 + \dots + \beta_m \mathcal{E}_m$$

Now, $\mathcal{E}_1 + \dots + \mathcal{E}_m = S^{-1}(F_1 + \dots + F_m) S = I$

(2) $\Rightarrow$ (1) Suppose $A = \alpha_1 \mathcal{E}_1 + \dots + \alpha_k \mathcal{E}_k$ where $\mathcal{E}_1 + \dots + \mathcal{E}_k = I$, $\mathcal{E}_i$ idemp.
then

$$\bigoplus \operatorname{Im} \mathcal{E}_i = V = \mathbb{C}^p \quad (*)$$

Let Γ_i = a basis of $\operatorname{Im} \mathcal{E}_i$, $\Gamma = \Gamma_1 \| \Gamma_2 \| \dots \| \Gamma_k$ a basis of $\mathbb{C}^p$

then $[A]_{\Gamma \leftarrow \Gamma}$? Express A as a block matrix w.r.t $(*)$

$$A_{11} = \mathcal{E}_1 \circ A \circ \mathcal{E}_1 \Big|_{\operatorname{Im}(\mathcal{E}_1)} = \alpha_1 \mathcal{E}_1 \Big|_{\operatorname{Im}(\mathcal{E}_1)} = \alpha I_{\operatorname{Ran} \mathcal{E}_1}$$

So...

$$A = \begin{bmatrix} \alpha_1 I & & \\ & \ddots & \\ & & \alpha_k I \end{bmatrix} \text{ w.r.t } \Gamma \longrightarrow [A]_{\Gamma \leftarrow \Gamma} = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_k \end{pmatrix}$$

So A diagonalizable. $\square$

$\boxed{\text{Corollary (to the proof) The } \alpha\text{'s are exactly the eigenvalues of } A}$

<u>and</u> $\operatorname{Im} \mathcal{E}_i = \mathcal{E}_A(\alpha_i)$ $\rightsquigarrow$ eigenspace of A corresponding to α_i

<u>and</u> the rep^n of A as

$A = \alpha_1 \mathcal{E}_1 + \dots + \alpha_k \mathcal{E}_k$ is <u>unique</u>

distinct

$\square$

If $A = \begin{pmatrix} \alpha_1 I & & \\ & \ddots & \\ & & \alpha_k I \end{pmatrix}$ wrt $\oplus \operatorname{Im} \mathcal{E}_i = \mathbb{C}^p$

then

$$p(A) = \boxed{\begin{matrix} p(\alpha_1) I & & \\ & \ddots & \\ & & p(\alpha_k) I \end{matrix}}$$

α_i are distinct

Consider $\mathcal{E}_1 = \boxed{\begin{matrix} I & 0 & \\ 0 & \ddots & \\ & & 0 \end{matrix}}$

Now, $p(z) = \dfrac{(z-\alpha_2)\cdots(z-\alpha_k)}{(\alpha_1-\alpha_2)\cdots(\alpha_1-\alpha_k)}$

sends A to $\mathcal{E}_1$

Thm $\boxed{\text{Each spectral idempotent of a diagonalizable } A \text{ is a polynomial in } A \text{, i.e. is of the form } p(A) \text{ for some poly. } p.}$

$\boxed{\text{Pf of Thm } \dagger}$ → by induction ... Induct on n

(1)→(2) . Base case : $n = 1$ trivially true

. Suppose holds for all $1 \le n < n_0$. Show holds for n_0.

If $\mathcal{F}$ contains only scalar multiples of identity ⟹ done ...

So, let us assume $\mathcal{F}$ contains $A \neq \alpha I$

then $A = \sum^k \alpha_i \mathcal{E}_i$ for some ... ... distinct α_i

Since distinct α_i ..., $k \geq 2$

Recall $\mathcal{E}_i = p_i(A)$, for some poly p_i

and every $B \in \mathcal{F}$ commutes with A

So B commutes with $p_i(A)$, i.e with every $\mathcal{E}_i$.

So $B = \begin{bmatrix} \end{bmatrix}$ $\qquad B_{11} = \mathcal{E}_1 B \mathcal{E}_1 \Big|_{\operatorname{Im} \mathcal{E}_1} = \ldots$

$$B_{ij} = \mathcal{E}_i \circ B \circ \mathcal{E}_j \Big|_{\mathrm{Im}(\mathcal{E}_j)} = \underbrace{\mathcal{E}_i \circ \mathcal{E}_j}_{\delta^i_j} \circ B \Big|$$

So B block diagonal: $B = \begin{bmatrix} B_{11} & & \\ & \ddots & \\ & & B_{kk} \end{bmatrix}$

A B commute, so A_{ij} commutes w B_{ij}

$\Rightarrow$ ~~BE~~ B_{ij} individually diagonalizable ... (by hypothesis)

concatenate basis $\Rightarrow$ B diagonalizable $\Rightarrow$ //

$\square$

April 9, 2019

$\boxed{\text{INNER PRODUCTS}}$

Lots and lots of way to measure distance ... $\left\{\begin{array}{l}\text{Pythagorean theorem} \\ \text{sum of + x ...}\end{array}\right.$

~~But~~

Pythagorean distance is associated with the dot product ...

$$\mathrm{dist}\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix}\right)^2 = \begin{pmatrix} x-a \\ y-b \end{pmatrix} \cdot \begin{pmatrix} x-a \\ y-b \end{pmatrix}$$

Also, $\left\|\begin{pmatrix} x \\ y \end{pmatrix}\right\| \cdot \left\|\begin{pmatrix} a \\ b \end{pmatrix}\right\| \cdot \cos\theta = \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix}$ $\rightsquigarrow$ there's angles captured and distance,

$\Rightarrow$ makes Pythagorean distance stand out ...

Defn An inner product $\varphi: V \times W \mapsto \mathbb{C}$ on a vector space V is a function satisfying the following conditions ...

① φ is ~~bilinear~~ partially linear in the first slot & partially conjugate linear in the second

i.e.,

$$\begin{cases} \varphi(\alpha v_1 + \beta v_2, w) = \alpha \varphi(v_1, w) + \beta \varphi(v_2, w) \quad \text{and} \\ \varphi(v, \alpha w_1 + \beta w_2) = \bar{\alpha}(\varphi)(v, w_1) + \bar{\beta} \varphi(v, w_2) \end{cases}$$

Standard inner product on $\mathbb{C}^n$:

$$\left\langle \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \right\rangle = a\bar{\alpha} + b\bar{\beta} + c\bar{\gamma} = \ldots$$

so that $\left\langle \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right\rangle = |a|^2 + |b|^2 + |c|^2 = \left\| \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right\|^2$

---

(2) $\overline{\varphi(v,w)} = \varphi(w,v)$

(3) $\varphi(v,v) \geq 0$, equality holds iff $v = 0$ (positive definite)

\#

$\boxed{\text{An inner product space is a vector space with an inner product}}$

Observe

If (V, φ) is an inner product space, then

$\|v\| := \sqrt{\varphi(v,v)}$ defines a "norm" (i.e. a sizing function)

Such that

(i) $\| \ \| : V \longrightarrow [0, \infty)$

(ii) $\|v\| = 0 \iff v = 0$

(iii) $\|\alpha v\| = |\alpha| \cdot \|v\|$

(iv) $\|u + v\| \leq \|u\| + \|v\|$ (triangle inequality)

(v) $\|\langle v, u \rangle\| \leq \|v\| \cdot \|u\|$

Pf of (iv) $\|u + v\| = \cdots = \sqrt{\|u\|^2 + 2\operatorname{Re}\varphi(u,v) + \|v\|^2}$

$$\cancel{\left(\leq \sqrt{\|u\|^2 + 2|\varphi(u,v)| + \|v\|^2} \right)}$$

$$\leq \sqrt{(\|u\| + \|v\|)^2} = \|u\| + \|v\|$$

☒ Can it be that we always have $|\varphi(u,w)| \leq \|u\| \cdot \|w\|$?

Yes! $\rightarrow$ $\boxed{\text{Cauchy-Schwartz inequality}}$

In any product space (V, φ) $\quad |\varphi(v,w)| \leq \sqrt{\varphi(v,v)} \cdot \sqrt{\varphi(w,w)}$

$$0 \leq \varphi(tu+w, tu+w) = \varphi(tu,tu) + 2\operatorname{Re}\varphi(tu,w) + \varphi(w,w)$$

for $t \in \mathbb{R}$ $\quad = t^2\varphi(u,u) + 2t\operatorname{Re}\varphi(u,w) + \varphi(w,w)$

Hence no real roots except 0 $\Rightarrow \Delta < 0$

So $\quad (2\operatorname{Re}\varphi(u,w))^2 - 4\varphi(u,u)\varphi(w,w) \leq 0$

So $\quad \operatorname{Re}\varphi(u,w) \leq \sqrt{\varphi(u,u)}\sqrt{\varphi(w,w)}$ $\boxed{?}$ But we're not done...

Know, for $|\alpha| = 1$,

$$\operatorname{Re}\left(\varphi(\alpha u, w)\right) \leq |\alpha|\sqrt{\varphi(u,u)}\sqrt{\varphi(w,w)}$$

$$\operatorname{Re}\left(\alpha \underbrace{\varphi(u,w)}_{re^{i\theta}}\right) \leq \sqrt{\varphi(u,u)}\sqrt{\varphi(w,w)}$$

for $\alpha = e^{-i\theta}$, I get

$\operatorname{Re}(r) \leq \sqrt{\varphi(u,u)}\sqrt{\varphi(w,w)}$, but $r = |\varphi(u,w)|$ real

So $\quad \operatorname{Re}|\varphi(u,w)| \leq \sqrt{\varphi(u,u)}\sqrt{\varphi(w,w)}$

So $\quad |\varphi(u,w)| \leq \sqrt{\varphi(u,u)}\sqrt{\varphi(w,w)}$ ▨

▶ for $V = C[0,1]$
we can def inner product of 2 fns $\quad \varphi(f,g) = \int_0^1 f \cdot \bar{g}$

April 11, 2019

ORTHOGONALITY

Recall: Inner product spaces $(V, \langle , \rangle)$

"Def" ORTHOGONAL $\Leftrightarrow$ $\langle , \rangle = 0$

☒ For any $S \subset V$, let $S^\perp = \{ v \in V \mid \langle v, s \rangle = 0 \text{ for every } s \in S \}$

① Observe that • $\theta_V \in S^\perp$

② • $V^\perp = \{ v \in V \mid \langle v, w \rangle \overset{0}{=} \text{ for all } w \in V \} \ni \theta_V$

Suppose $v_0 \in V^\perp$, then $\langle v_0, v_0 \rangle = 0$, thus $v_0 = \theta$

③ • $(V \setminus \{\theta_V\})^\perp = \dots = \{\theta_V\}$

④ • $(V \setminus \{v_0\})^\perp = \dots = \{\theta_V\}$

⑤ • $(V \setminus \{v_0, v_1 \dots v_n\})^\perp = \dots = \{\theta_V\}$

uncountably many

⑥ For any $S \subset V$: $S^\perp \triangleleft V$ Pf $\begin{cases} \theta_V \in S^\perp \\ \text{partial lin. in } 1^{st} \text{ slot of } \langle v, w \rangle \end{cases}$

⑦ Thm $\boxed{\text{If } W \triangleleft V \text{ then } W \oplus W^\perp = V}$ $\rightarrow$ f.d only...

If we knew $W + W^\perp = V$ then the sum is direct, since $W \cap W^\perp = \{\theta\}$

☒ finite-dim Let $w_1, w_2, \dots, w_k$ be a basis for W and let $v \in V$

To show: $\exists \alpha_1 \dots \alpha_k \in \mathbb{C}$ s.t. $\alpha_1 w_1 + \dots + \alpha_k w_k + y = v$ for some $y \in W^\perp$

i.e. $\exists \alpha_1 \dots \alpha_k \in \mathbb{C}$ such that $v - \alpha_1 w_1 - \dots - \alpha_k w_k \in W^\perp$

i.e. $\exists \alpha_1 \dots \alpha_k \in \mathbb{C}$ s.t.

$$\langle v - \alpha_1 w_1 - \dots - \alpha_k w_k, w_1 \rangle = 0$$
$$\langle v - \alpha_1 w_1 - \dots - \alpha_k w_k, w_2 \rangle = 0$$
$$\vdots$$

So $\langle v - \alpha_1 w_1 - \dots - \alpha_k w_k , w_j \rangle = 0$

So $\langle v, w_j \rangle = \alpha_1 \langle w_1, w_j \rangle + \alpha_2 \langle w_2, w_j \rangle + \dots + \alpha_k \langle w_k, w_j \rangle$

GRAMIAN MATRIX $\rightsquigarrow$

i.e. $$\begin{bmatrix} \langle w_1, w_1 \rangle & \langle w_2, w_1 \rangle & \dots & \langle w_k, w_1 \rangle \\ \vdots & & & \\ \langle w_1, w_k \rangle & & & \langle w_k, w_k \rangle \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} = \begin{bmatrix} \langle v, w_1 \rangle \\ \vdots \\ \langle v, w_k \rangle \end{bmatrix}$$

We shall show it is invertible. But since it's square, we can just show it is injective.

Suppose $G(\gamma) = 0$ then

$$\sum_i \gamma_i \langle w_i, w_1 \rangle = \sum_i \gamma_i \langle w_i, w_2 \rangle = \dots = \sum_i \gamma_i \langle w_i, w_k \rangle = 0$$

So $\sum_i \gamma_i w_i$ is $\perp$ $w_1, \dots, w_k$

In particular, $\sum_i \gamma_i w_i \perp \sum_i \gamma_i w_i$, i.e. $\sum \gamma_i w_i = 0$

So $\gamma_i = 0 \ \forall i$ (since w_i form a basis)

So, the Gramian is injective, hence it is invertible. ✓

⇒ Note that in infinite dim → this theorem does not hold

Ex ↳ Let $V = \mathbb{C}^{\mathbb{N}} : \{ f : \mathbb{N} \to \mathbb{C} \mid f = \text{fnct} \}$ ← sequences of complex numbers

$$= \boxed{\begin{matrix} 1 & 2 & 3 & \dots \\ \alpha_1 & \alpha_2 & \alpha_3 & \dots \end{matrix}} \quad \alpha_i \in \mathbb{C}$$

We claim that V is a vector space (this is easy to check)

Let $V_2 < V$ be defined by ℓ^2

$$(\alpha_1 \ldots\ldots \alpha \ldots) \in V_2 \iff \sum_1^\infty |\alpha_i|^2 < \infty$$

Define $\langle (\alpha_1 \ldots\ldots), (\beta_1 \ldots) \rangle = \alpha_1\overline{\beta_1} + \alpha_2\overline{\beta_2} + \ldots$

Now, we show that this sum doesn't diverge

Consider $|\alpha_1\overline{\beta_1}| + |\alpha_2\overline{\beta_2}| + \ldots$ i.e. $|\alpha_1|\cdot|\overline{\beta_1}| + |\alpha_2||\overline{\beta_2}| + \ldots$

Consider partial sums... $|\alpha_1|\cdot|\beta_1| + \ldots + |\alpha_{34}||\beta_{34}|$

$$= \left\langle \begin{pmatrix} |\alpha_1| \\ \vdots \\ |\alpha_{34}| \end{pmatrix}, \begin{pmatrix} |\beta_1| \\ \vdots \\ |\beta_{34}| \end{pmatrix} \right\rangle \le \sqrt{\underbrace{\left\| \begin{pmatrix} |\alpha_1|^2 \\ \vdots \\ |\alpha_{34}|^2 \end{pmatrix} \right\|}_{A} \cdot \underbrace{\left\| \begin{pmatrix} |\beta_1|^2 \\ \vdots \\ |\beta_{34}|^2 \end{pmatrix} \right\|}_{B}}$$

$$A \le \sum^\infty |\alpha_i|^2 \qquad B \le \sum^\infty |\beta_i|^2$$

But $\sum^\infty |\alpha_i|^2 < \infty$, so both A, B converges.

→ so any inner product converges...

Therefore the inner product is well defined.

Back to proof. Let $W \subseteq \ell^2$ be defined by

W the set of all terminating sequences in ℓ^2

So (we can check) $W < \ell^2$, and $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix} \in W$, and

$\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix} \in W$, $\begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ \vdots \end{pmatrix} \in W$ ...

$W + W^\perp = W$

So $\begin{pmatrix} \alpha_1 \\ \vdots \end{pmatrix} \in W^\perp \Rightarrow \alpha_j = 0 \ \forall j \Rightarrow W^\perp = \{0\}$

Yet $\begin{pmatrix} 1 \\ 1/2 \\ 1/4 \\ \vdots \end{pmatrix} \in \ell^2, \notin W$ so $\boxed{W + W^\perp \neq \ell^2}$

(read Axler)

April 16, 2019 **Recall**, $W < V$, V finite dimensional inner product space...
$W \oplus W^{\perp} = V$

Observation $\boxed{\text{Orthonormal list in inner product space is linearly independent.}}$

Suppose $\alpha_1 V_1 + \dots + \alpha_k V_k = 0$
where $V_1 \dots V_k$ mutually orthogonal unit vectors

Then $0 = \langle 0, V_i \rangle$

$= \alpha_i \langle V_i, V_i \rangle + 0 = \alpha_i$

So all V_i's are linearly independent.

Thm $\boxed{\text{Every finite dimensional inner product space (nontrivial) has an orthonormal basis, and in fact any orthonormal list can be enlarged to one.}}$

$\boxed{\text{Riesz Representation Theorem}}$

Suppose V is fdips and $\overbrace{\rho : V \xrightarrow{\text{lin}} \mathbb{C}}^{\text{a linear functional on } V}$.
Then there exists ~~only~~ exactly one $w_0 \in V$ s.t.

$\rho(v) = \langle v, w_0 \rangle$, all v, i.e. $\rho(\,) = \langle (\,), w_0 \rangle$

Proof $\underbrace{\text{Rank}(\rho)}_{0 \text{ or } 1} + \text{Nullity}(\rho) = \dim(V)$

If rank $(\rho) = 0$, then $\rho \equiv 0$, so $w_0 = 0$ (unique) (easy)
If rank $(\rho) = 1$
then $\ker(\rho) \oplus \ker(\rho)^{\perp} = V$

then $\ker(\rho)^\perp$ has dimension one. Then $\ker(\rho)^\perp = \text{span}(z_0)$, z_0 unit vector...

then $\rho(z_0) \neq 0 \Rightarrow \frac{1}{\rho(z_0)} z_0 \in \text{span}(z_1)$

and $\rho\Big(\underbrace{\frac{1}{\rho(z_0)} z_0}_{z_1}\Big) = \frac{1}{\rho(z_0)} \rho(z_0) = 1$

And so, $\text{span}(z_0) = \text{span}(z_1)$

Thus, any $v \in V$, $v = \underset{\ker(\rho)}{\underset{\cap}{w}} + \underset{\ker(\rho)^\perp}{\underset{\cap}{\alpha z_1}}$ (uniquely)

So $\rho(v) = \rho(w + \alpha z_1) = 0 + \alpha \rho(z_1) = \alpha \overset{?}{=} \langle v, ??\rangle$

So now, $\langle w + \alpha z_1, \boxed{??} \rangle$ how to get α?

$= \langle w + \alpha z_1, ?? \rangle$

$= \langle w, ?z_1 \rangle + \alpha \langle z_1, ?z_1 \rangle$

0 So, $? = \frac{1}{\|z_1\|^2}$

And thus, $\boxed{??} = \frac{1}{\|z_1\|^2} z_1$

So, let $w_0 = \frac{z_1}{\|z_1\|^2}$

then $\rho(v) = \rho(w + \alpha z_1) = \langle v, \frac{z_1}{\|z_1\|^2} \rangle = (\alpha) = \langle v, w_0 \rangle$

for all v. //

Recap, if $v = w + \alpha z_1$, then $w_0 = \frac{z_1}{\|z_1\|^2}$ $\forall v$

<u>Uniqueness</u> Assume $\rho() = \langle (), w_0 \rangle = \langle (), w_1 \rangle$ then $0 = \langle (), w_0 - w_1 \rangle \Leftrightarrow w_0 = w_1$. //

$\boxed{\text{Consequences?}}$

① Suppose $\mathcal{L} : V \xrightarrow[\text{lin}]{\text{f dim'l}} V$. Consider

$f() := \langle \mathcal{L}(), y_0 \rangle$, $f : V \xmapsto{\text{lin}} \mathbb{C}$, So, f must be like

$$f() = \langle (), w_0 \rangle \quad \text{for some } \overset{\boxed{\text{UNIQUE}}}{w_0} \in V, \ w_0 \text{ depends on } y_0 \text{ and } \mathcal{L}.$$

So let us write $w_0 = w_{\mathcal{L}}(y_0)$

Consider $w_{\mathcal{L}} : V \longmapsto V$

Now $f() = \langle (), w_0 \rangle = \langle \mathcal{L}(), y_0 \rangle$

$$\langle x, \alpha w_{\mathcal{L}}(y_0) \rangle = \bar{\alpha} \langle x, w_{\mathcal{L}}(y_0) \rangle$$

So $\langle x, w_{\mathcal{L}}(\alpha y_0) \rangle = \langle \mathcal{L}x, \alpha y_0 \rangle = \bar{\alpha} \langle \mathcal{L}x, y_0 \rangle = \bar{\alpha}\langle x, w_{\mathcal{L}}(y_0)\rangle$

So $\boxed{w_{\mathcal{L}}(\alpha y_0) = \alpha w_{\mathcal{L}}(y_0)}$ (~~conjugate~~ linear)

- what about additivity?

$\langle x, w_{\mathcal{L}}(y_1+y_2) \rangle = \langle \mathcal{L}(x), y_1+y_2 \rangle = \langle x, w_{\mathcal{L}}(y_1) \rangle + \langle x, w_{\mathcal{L}}(y_2) \rangle$

$\langle x, y \rangle - \langle x, z \rangle = 0$
$\langle x, y-z \rangle = 0 \ \forall x$
$\Rightarrow y - z = 0$

Thus, we can show $w_{\mathcal{L}}(y_1+y_2) = w_{\mathcal{L}}(y_1) + w_{\mathcal{L}}(y_2)$ by subtracting ...

And so, additivity checks ...

So $W_{\mathcal{L}} : V \xmapsto{\text{lin}} V$ has the property

$$\langle \mathcal{L}(), y \rangle = \langle (), W_{\mathcal{L}}(y) \rangle$$

$W_{\mathcal{L}}$ is the adjoint of $\mathcal{L}$, denoted $\mathcal{L}^*$ □

Note $\mathcal{L}^*$ depends on the inner product, but not on the basis.

Now, $\overset{\text{fdips}}{V} \xmapsto{\;L\;} \overset{\text{fdips}}{U}$

Let Γ be orthonorm basis of V $\quad v_1 \dots v_n$

Π be orthnorm basis of U $\quad u_1 \dots u_m$

then

$$\begin{tikzcd}
V \arrow[r, shift left, "L"] \arrow[d, shift left, "{[\;]_\Gamma}"] & U \arrow[l, shift left, "L^*"] \arrow[d, shift right, "{[\;]_\Pi}"'] \\
\mathbb{C}^n \arrow[u, shift left, "A_\Gamma"] \arrow[r, shift left, "{[L]_{\Pi\leftarrow\Gamma}}"] & \mathbb{C}^m \arrow[u, shift right, "A_\Pi"'] \arrow[l, shift left, "{[L^*]_{\Gamma\leftarrow\Pi}}"]
\end{tikzcd}$$

What is the relationship between $[L]$ & $[L^*]$? Claim:

$$\boxed{\underbrace{[L^*]_{\Gamma\leftarrow\Pi}}_{n\times m} = \left(\overline{\underbrace{[L]_{\Pi\leftarrow\Gamma}}_{m\times n}}\right)^T}$$

Let write this $\quad A = (\bar{B})^T \qquad$ do this entry-wise ...

$$A_{ij} = A[\![i,j]\!] = A(e_j)\cdot e_i = A(e_j) \,\bar{\cdot}\, e_i$$

(★) Observation $\quad \alpha_1 v_1 + \dots + \alpha_n v_n \quad , v_i$ orthonormal

then $\langle \alpha_1 v_1 + \dots + \alpha_n v_n , \beta_1 v_1 + \dots + \beta_n v_n \rangle$

$$= \alpha_1 \bar{\beta}_1 + \dots + \alpha_n \bar{\beta}_n = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \cdot \begin{pmatrix} \bar{\beta}_1 \\ \vdots \\ \bar{\beta}_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \,\bar{\cdot}\, \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

standard inner product

↳ like dot product!

April 18, 2019

Def^n $L : V \xrightarrow{\text{lin}} V$ f.dips is <u>normal</u> if $LL^* = L^*L$; $L^* \leftrightarrow L$

is <u>self-adjoint</u> if $L = L^*$; $L^* = L$

is <u>unitary</u> if $LL^* = L^*L = I$ i.e. $L^* = L^{-1}$

NORMAL OPERATOR | SELF-ADJOINT | UNITARY

(Peak into Future) (1) $\mathcal{L}$ is normal $\Leftrightarrow$ there is an ortho-basis Γ of V such that $[\mathcal{L}]_{\Gamma\leftarrow\Gamma}$ is <u>diagonal</u>

1 SPECTRAL THEOREM $\Leftrightarrow$ there is an ortho-eigenbasis of L for V (eigenvect

2 $L = L^* \Leftrightarrow \mathcal{L}$ is normal & $\sigma_{\mathbb{C}}(\mathcal{L}) \subset \mathbb{R}$

3 $L^* = L^{-1} \Leftrightarrow \mathcal{L}$ is normal & $\sigma_{\mathbb{C}}(\mathcal{L}) \subset \mathbb{T} = \{e^{i\theta} \mid \theta \in \mathbb{R}\}$

(all eigenvalues have modulus 1)

☆ **Schur's Thm** For each $\mathcal{L} : V \xrightarrow{\text{lin}} V$ f.dips there is an ortho-basis Γ of V s.t. $[\mathcal{L}]_{\Gamma\leftarrow\Gamma}$ = upper-triangular

PF 1-dim ✓

<u>2-dim</u> : There is a unit ~~eigenvector~~ eigenpair (λ, v_0) for $\mathcal{L}$. → unit vector

dim = 1 ← Let $W = \text{Span}(v_0)$ ⇝ 1 dimensional

then $V = W \oplus W^\perp$. Let v_1 be a unit vector in $W^\perp$

Then $\Gamma = (v_0, v_1)$ is an orth-basis of V

then

$$[\mathcal{L}]_{\Gamma\leftarrow\Gamma} = \begin{array}{cc} & v_0 \\ v_0 & \begin{pmatrix} \lambda & m \\ 0 & m \end{pmatrix} \end{array} \rightsquigarrow \text{upper triangular...}$$

<u>3-dim</u>

$$[\mathcal{L}]_{\Gamma\leftarrow\Gamma} = \left[\begin{array}{c|c} \lambda & \\ \hline 0 & \\ 0 & M \end{array}\right] \longrightarrow M = \mathcal{E}_{W^\perp} \circ \mathcal{L}\Big|_{W^\perp}$$

By 2-dim, there is an basis v_1, v_2 of $W^\perp$ such that (φ) ortho

$$[M]_{\varphi\leftarrow\varphi} = \left[\begin{array}{c|c} ? & ? \\ \hline 0 & ? \end{array}\right]$$

Thus, let $\Gamma = (v_0, v_1, v_2)$. then Γ = ortho. basis of $\mathcal{L}$

<u>Thus</u> $[\mathcal{L}]_{\Gamma\leftarrow\Gamma} = \begin{bmatrix} [\lambda I] & [4] \\ 0 & [M] \end{bmatrix} = \begin{bmatrix} \lambda & \cdots \\ 0 & 0 \ \ \ddots \end{bmatrix} \sim$ <u>upper triangular</u>.

Towards the Spectral Theorem

① Suppose $M \in M_{m\times m}$, and
$$M = \begin{array}{c|c|c|c} & k & & \\ k & A & B & m-k \\ m-k & C & D & k \end{array}$$
then
$$M^* = (\overline{M})^T = \begin{pmatrix} \overline{A} & \overline{B} \\ \overline{C} & \overline{D} \end{pmatrix}^T = \begin{pmatrix} \overline{A}^T & \overline{C}^T \\ \overline{B}^T & \overline{D}^T \end{pmatrix} = \begin{pmatrix} A^* & C^* \\ B^* & D^* \end{pmatrix}$$

② Suppose M is normal, then $M \leftrightarrow M^*$ then

$$MM^* = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} A^* & C^* \\ B^* & D^* \end{pmatrix} = \begin{pmatrix} AA^* + BB^* & \cdots \\ \cdots & \cdots \end{pmatrix}$$
$$\parallel$$
$$M^*M = \begin{pmatrix} A^* & C^* \\ B^* & D^* \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A^*A + C^*C & \cdots \\ \cdots & \cdots \end{pmatrix}$$

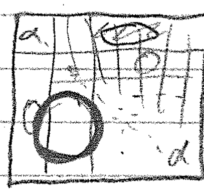
So $AA^* + BB^* = A^*A + C^*C$

So $\mathrm{Trace}(AA^* + BB^*) = \mathrm{Trace}(A^*A + C^*C)$

So $\cancel{\mathrm{Trace}(AA^*)} + \mathrm{Trace}(BB^*) = \cancel{\mathrm{Trace}(A^*A)} + \mathrm{Trace}(C^*C)$

Note $\boxed{\mathrm{Trace}(AA^*) = 0 \Leftrightarrow A = 0}$ N.B. $\mathrm{Tr}(BB^*) = \sum |B_{ij}|^2$

So

③ Suppose  is normal, then  is diagonal ...

 → every normal has a representation that is diagonal ...

④ Converse?

SPECTRAL THEOREM

✪ Equivalent statements:

① A is normal then there is an ortho-basis of the underlying space made up entirely of the eigenvectors of A

i.e. ② $A: {}_{\Gamma}V \xrightarrow[\text{norm}]{\text{lin}} {}_{\Gamma}V$, then $\exists$ ortho basis s.t $[A]_{\Gamma\Gamma}$ is diagonal

i.e. ③ $\mathcal{E}_{A,\lambda_1} \oplus \mathcal{E}_{A,\lambda_2} \oplus \cdots \oplus \mathcal{E}_{A,\lambda_k} = V$ → direct sum + orthonormal

$\boxed{\text{direct sum}} + \boxed{\mathcal{E}_{A,\lambda_1}^{\perp} = \bigoplus_{2}^{k} \mathcal{E}_{A,\lambda_j}}$

if A normal

i.e. (4) $\boxed{A = \lambda_1 E_1 + \dots + \lambda_k E_k}$ → idempotent → also self adjoint

ortho-resolution of identity

$$A \sim \begin{pmatrix} \lambda_{\lambda_1} & & \\ & \ddots & \\ & & \lambda_{\ell_{\lambda_k}} \end{pmatrix}$$

where $E_i^2 = E_i = E_i^*$ and $\sigma(E_i) \subset \{0,1\}$ ← superfluous...

⇒ $\left\{ \begin{array}{l} E_i \text{ is a non-negative self-adjoint} \\ \text{or positive-semidefinite} \end{array} \right\}$

$\boxed{\text{Terminology}}$ Non-negative self-adjoint idempotents are called ORTHO-PROJECTION

So
{ diagonalizable ⇒ resolution of identity = lin. comb. of idempotents
{ normal ⇒ resolution of identity = lin. comb. of idempotents that are ORTHO-PROJECTIONS.

Recall for idempotents. $V = \mathrm{Im}(E) \oplus \ker(E)$
For ORTHO PROJECTIONS ⇒ $\mathrm{Im}(E) \perp \ker(E)$

SPECTRAL THEOREM FOR MATRICES

$\Gamma = (f_1 \dots f_n) =$ ortho-basis f[or] $\mathbb{C}$

$$\begin{array}{ccc} \mathbb{C}^n_\Gamma & \xrightarrow[\text{normal}]{M} & \mathbb{C}^n_\Gamma \\ \uparrow A_\Gamma & & \downarrow [\,]_\Gamma \\ \mathbb{C}^n_e & \xrightarrow[{[M]_{\Gamma\leftarrow\Gamma}}]{} & \mathbb{C}^n_e \end{array}$$

$A_\Gamma = \begin{bmatrix} f_1 & f_2 & \dots & f_n \end{bmatrix} : \mathbb{C}^n \xrightarrow[\text{bij}]{} \mathbb{C}^n$

square matrix
bijective
orthonormal columns
} ⇒ Unitary

Unitary matrices

$$U^* U = I \iff \begin{bmatrix} -\bar{c}_1- \\ -\bar{c}_2- \end{bmatrix} \begin{bmatrix} | & \\ c_1 & \\ | & \end{bmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$$

$\iff \langle c_1, c_1 \rangle = 1 \,;\, \langle c_2, c_1 \rangle = 0 \Rightarrow U$ normal & $\sigma_U \in \{e^{i\theta} \mid \theta \in \mathbb{R}\}$

Unitary Matrices, cont

$U^*U = I \Leftrightarrow \langle U^*Ux, y\rangle = \langle Ix, y\rangle = \langle x, y\rangle \quad \forall x, y$

(proof: $\langle Ax, y\rangle = 0 \ \forall x, y \Leftrightarrow A = 0$)

$\Leftrightarrow \langle U^*Ux, x\rangle = \langle x, x\rangle \quad \forall x$

(proof: ⇐ Note $\|U\vec{x}\|^2 = \|\vec{x}\|^2$

then note that $\langle x, y\rangle = \frac{1}{4}\left[\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\right]$

then $\langle Ux, Uy\rangle = \dots$ then linearity, then preserve lgth

then $\langle x, y\rangle = \langle Ux, Uy\rangle \quad \forall x, y$)

Equivalent Statements

Note $\langle Ux, Uy\rangle = \langle x, y\rangle \Leftrightarrow \|U(x)\| = \|x\| \Leftrightarrow \langle Ux, Ux\rangle = \langle x, x\rangle$

$\Updownarrow$

$\langle U^*Ux, y\rangle = \langle x, y\rangle \Leftrightarrow U^*U = I$

$\Updownarrow$

U normal & $\sigma(U) = \Pi$ $\Leftrightarrow$ columns form ortho-basis $\Leftrightarrow$ Unitary

$\Updownarrow$ $\Updownarrow$

U^* unitary $\Leftrightarrow$ Rows of U form ortho-basis

Back to $A_\beta = \begin{bmatrix} f_1 & \dots & f_n \end{bmatrix} : \mathbb{C}^n \xrightarrow{bij} \mathbb{C}^n \quad A_\beta^* = $ ...

orthobasis $\Rightarrow A_\beta$ unitary, so $[\ ]_\beta$ unitary

So
$$\begin{tikzcd} M \arrow[d, bend left, "{[\ ]_\Gamma}"] \\ {[M]_{\Gamma\leftarrow\Gamma}} \arrow[u, bend left, "A_\Gamma"] \end{tikzcd}$$

So $\boxed{\underbrace{U^*}_{U^{-1}} \circ M \circ \underbrace{U}_{\text{unitary}} = \underbrace{[M]_{\Gamma\leftarrow\Gamma}}_{\text{diagonal}}}$

columns of U form Γ

So $\boxed{\text{spectral theorem for matrices}}$ Diagonalizable, $S^{-1}AS = B$ such that S is unitary

M is normal, then M is unitarily equivalent to diagonal

What if $W^* M W = \overset{D}{\text{diagonal}}$, W unitary?

Does this mean M is normal? $\rightarrow$ $\boxed{\text{Yes}}$

$M = W \circ D \circ W^*$, then $M^* = W \circ \underbrace{D^*}_{\text{diagonal}} \circ W^*$

So $M^* M = W D^* W^* W D W^* = W D^* D W^*$

$ M M^* = W D W^* W D^* W^* = W D D^* W^*$

$D^* D = D D^*$

So $M^* M = M M^*$

$\boxed{\text{Spectral Theorem for Matrices}}$

$[M] \in M_{n\times n}$ normal $\Leftrightarrow$ it is unitarily equivalent to diagonal

Schur's Theorem - extension

Every $n\times n$ matrix is <u>unitarily ~~equivalent~~</u> to a triangular

---

$\boxed{\text{Back to normal matrices}}$

A normal, then $A = \lambda_1 E_1 + \dots + \lambda_k E_k$

ortho - resolution of identity

☒ What else do we get from Spectral Theorem?

A normal, then $A = U^* D U$ → unitary

$D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_k \end{bmatrix}$

then $A^2 = U^* D^2 U$ ...

So $A^n = U^* D^n U$

→ polynomial.

In particular, $P(A) = U^* P(D) U$

"$\boxed{\text{Defn}}$" $f(A) := U^* \begin{bmatrix} f(d_1) & & \\ & \ddots & \\ & & f(d_k) \end{bmatrix} U$ → whenever this makes sense..

~~Thm~~ But what if we have different U's?

$\boxed{\text{Better definition}}$

$$\boxed{f(A) := f(\lambda_1) E_1 + \dots + f(\lambda_k) E_k}$$

$E \sim \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 0 \end{pmatrix}$

↑ this definition removes ambiguity

<u>Thm</u>

$$f(A) = U^* \begin{bmatrix} f(d_1) & & \\ & \ddots & \\ & & f(d_k) \end{bmatrix} U$$

this becomes a theorem.

Suppose A is normal & $\sigma_A \subset [0,\infty) \subset \mathbb{R}$

then $\sqrt{A}\cdot\sqrt{A} = \left(\sum \sqrt{\lambda_i}\, E_i\right)\left(\sum \sqrt{\lambda_j}\, E_j\right)$

$= \lambda_1 E_1 + \lambda_2 E_2 \cdots + \lambda_k E_k = A$

So if A is normal & $\sigma_A \subset [0,\infty) \subset \mathbb{R}$
then
$\sqrt{A}\circ\sqrt{A} = A$, where $\sqrt{A} = \sqrt{\lambda_1}\,E + \ldots + \sqrt{\lambda_k}\,E_k$

and so $\sqrt{A}$ is also normal, with non negative spectrum
and so ...

So

$\boxed{\text{Every } \underline{\text{non-negative}} \text{ normal matrix } A \text{ has a } \underline{\text{nonnegative normal}} \text{ square root } \sqrt{A}}$

In fact

$\boxed{A \text{ normal} + \sigma_A \text{ real} \Rightarrow A \text{ self adjoint}}$

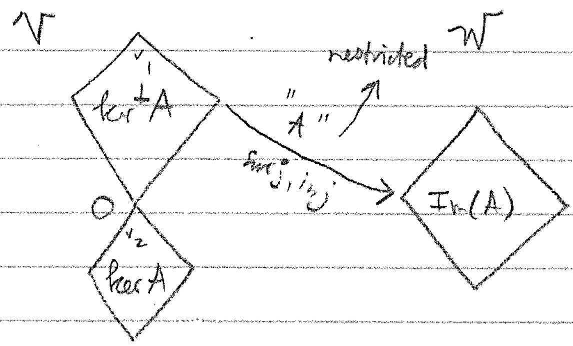
Pf
$$\left\{\begin{aligned} A &= \lambda_1 E_1 + \ldots + \lambda_k E_k \\ A^* &= \bar\lambda_1 E_1^* + \ldots + \bar\lambda_k E_k^* \\ A^* &= \lambda_1 E_1 + \ldots + \lambda_k E_k = A \text{ so } A \text{ self adj} \end{aligned}\right.$$

"normal matrices are like numbers ... we can apply functions to them ... they're like a number system ... not really, but they're quite nice"

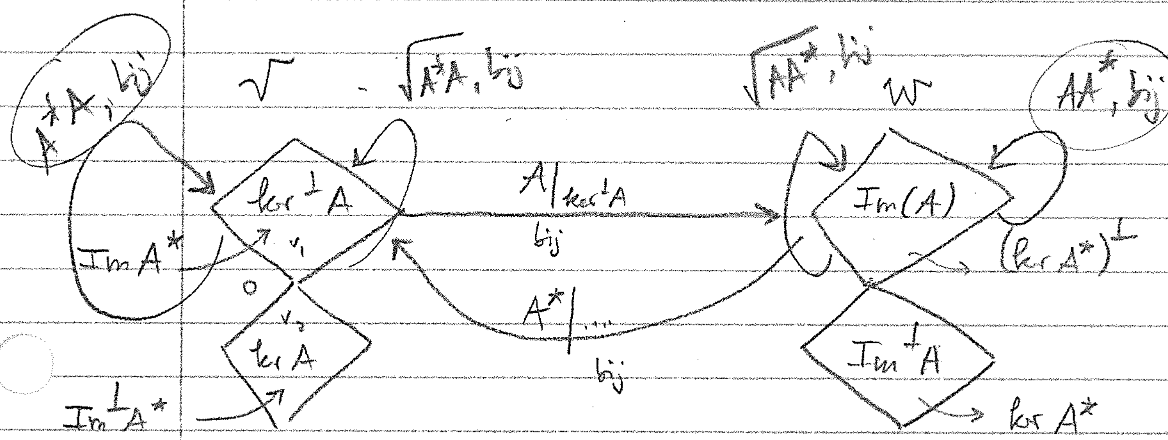
May 7, 2019

Spatial of Adjoint

$A : V \longmapsto W, \quad V = \ker A \oplus \ker^{\perp} A$



But if $Av_1 = 0$, then $v_1 \in \ker A$
But $v_1 \in \ker^{\perp} A$
So $v_1 = 0$
$\rightarrow$ $\boxed{\text{"A" is a bijection}}$



$A|_{\ker^{\perp} A}$ is bijective

But $\ker(A^*) = (Im A)^{\perp}$
$(\ker(A))^{\perp} = (\ker((A^*)^*))^{\perp} = Im(A^*)$

<u>Now</u> $Im(A^*A) \overset{?}{=} Im(A^*) = \ker^{\perp} A$

It's true that $Im(A^*A) \subset Im(A^*) = \ker^{\perp} A$

and $\ker(A^*A) \supset \ker A$

<u>But notice</u> $A^*A(v_1 + v_2) = A^*Av_1 = 0 \iff v_1 = 0$

<u>So</u> $\boxed{\ker(A^*A) = \ker A}$

Alternatively, $\langle A^*Ax, x \rangle = \langle Ax, Ax \rangle = \|Ax\|^2$
if $A^*Ax = 0$, then $\|Ax\|^2 = 0$, then $Ax = 0$
and if $Ax = 0 \rightarrow A^*Ax = 0 \rightarrow \langle , \rangle = 0$

So $\boxed{Ax = 0 \Leftrightarrow A^*Ax = 0 \text{ , i.e. } \ker(A) = \ker(A^*A)}$

So now $\ker(A^*A) = \ker(A)$

$$\to \left(\ker(A^*A)\right)^\perp = \left(\ker(A)\right)^\perp$$

$$\to \mathrm{Im}\left(\underbrace{(A^*A)^*}\right) = \mathrm{Im}\, A^*$$

But A^*A non negative

$$\boxed{\mathrm{Im}(A^*A) = \mathrm{Im}(A^*) = \ker^\perp(A)}$$

Now...

$$x \in \ker(A^*) \Leftrightarrow A^*x = 0 \Leftrightarrow \|A^*x\|^2 = 0$$

$$\to \langle A^*x, A^*x\rangle = 0 \Leftrightarrow \langle AA^*x, x\rangle = 0$$

$$x \in \ker(A) \Leftrightarrow \dots \Rightarrow \langle A^*Ax, x\rangle = 0$$

So, if A normal $\Rightarrow \ker(A) = \ker(A^*)$

$$\qquad\qquad\qquad\qquad\qquad \| $$

$$\qquad\qquad\qquad\qquad (\mathrm{Im}(A))^\perp$$

So $\boxed{A \text{ normal} \Rightarrow \ker(A) = (\mathrm{Im} A)^\perp}$

Or $\boxed{A \leftrightarrow A^* \to \ker(A) = (\mathrm{Im} A)^\perp}$

$x \in \ker(A)$

Now since A^*A nonnegative, there is exactly one $\sqrt{A^*A}$ nonnegative

--- AA^* --- --- --- $\sqrt{AA^*}$ --- ---

Observe: $x \in \ker \sqrt{A^*A} \Leftrightarrow \left\|\sqrt{A^*A}\, x\right\|^2 = 0$

$$\Rightarrow \left\langle \sqrt{A^*A}\,x, \sqrt{A^*A}\,x \right\rangle = 0$$

But $\sqrt{A^*A}$ self adj

$$\to \left\langle \left(\sqrt{A^*A}\right)^*\left(\sqrt{A^*A}\right)x, x\right\rangle = 0 \Leftrightarrow \langle A^*Ax, x\rangle = 0 \Leftrightarrow \boxed{\|Ax\|^2 = 0}$$

So $\boxed{\ker\left(\sqrt{A^*A}\right) = \ker(A)}$

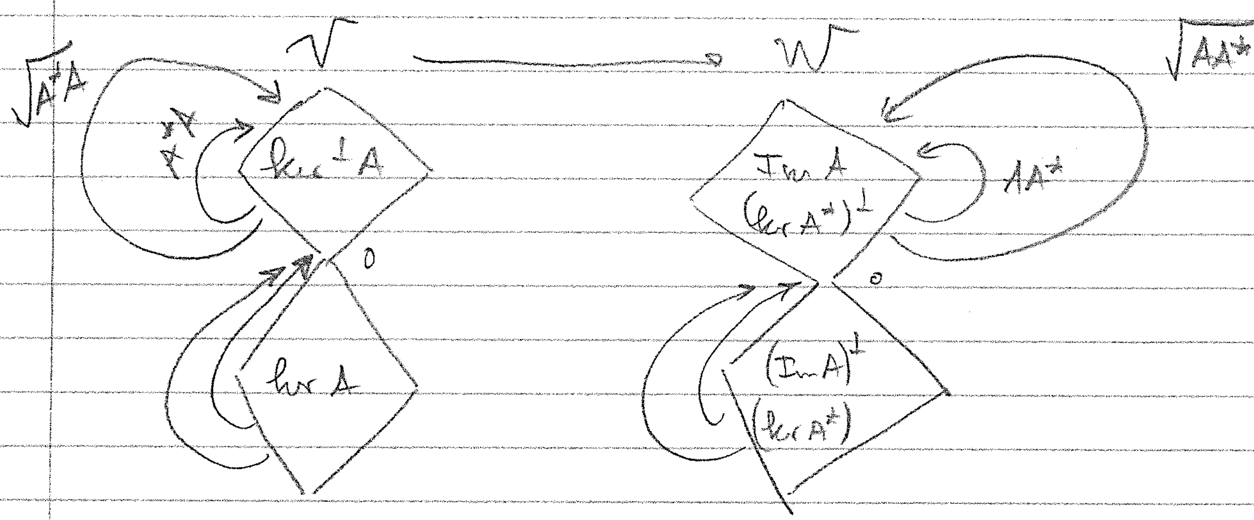
Now $\operatorname{Im}\left(\sqrt{A^*A}\right) = \left(\ker\left(\sqrt{A^*A}\right)^*\right)^{\perp}$

$$= \left(\ker\sqrt{A^*A}\right)^{\perp}$$

$$= \left(\ker(A)\right)^{\perp}$$

$$\boxed{\operatorname{Im}\sqrt{A^*A} = \operatorname{Im}\left(A^*\right)}$$

Picture, again

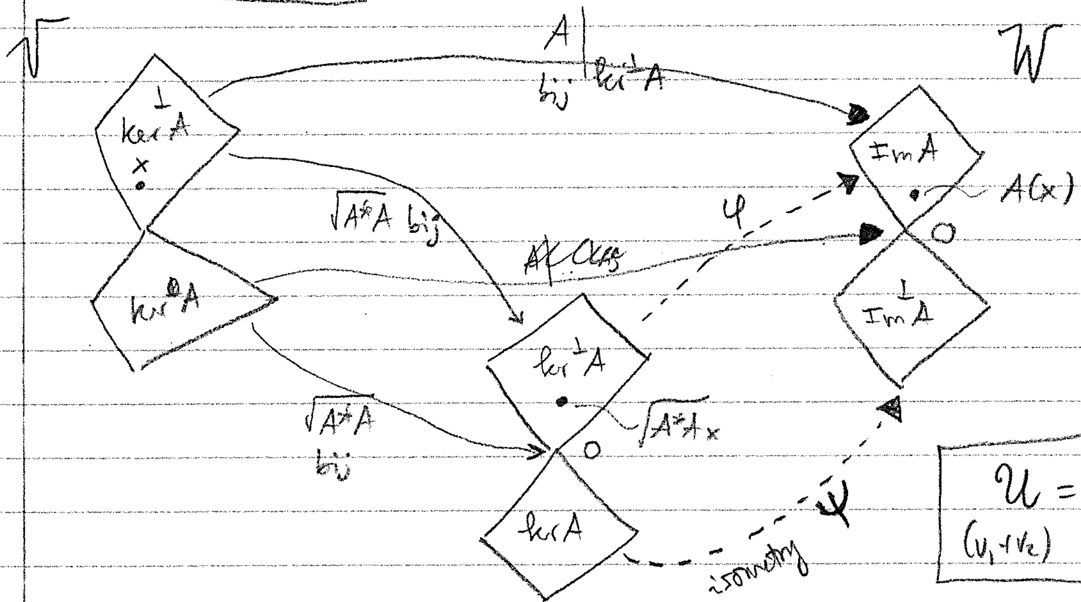


$\boxed{\text{Repeat argument}}$ $\boxed{\text{Theorem}}$

$$\| \sqrt{A^*A}\, x \| = \| Ax \|$$

$\in V$ $\in W$

Another picture



$$\underset{(v_1+v_2)}{\mathcal{U}} = \varphi(v_1) + \psi(v_2)$$

Now, $\underline{\|\sqrt{A^*A}x\| = \|Ax\|}$

Define new function

$\varphi: \ker^\perp A \longmapsto \mathrm{Im}\, A$

as $\varphi\left[\sqrt{A^*A}(x)\right] := A(x)$

lin, bij   lin, bij

Claim 1 φ is linear ... because $\varphi = A \circ \left[\sqrt{A^*A}\Big|_{\ker^\perp(A)}\right]^{-1}$

Claim 2 φ is an isometry (length preserving)

Since length preserving ...

If W is V, then $\dim(\ker A) = \dim(\mathrm{Im}^\perp A)$

then, we can define an isometry $\ker A \longmapsto \mathrm{Im}^\perp A$

To do this $\underset{\text{ortho-B}}{\Gamma = (u_1, \ldots u_k)}$ ... $\underset{\text{ortho-B}}{\Gamma' = (w_1 \ldots w_k)}$, send the basis elements

So $\quad z = \sum \alpha_i u_i \rightarrow \|z\|^2 = \sum |\alpha_i|^2$

$$\mathcal{L}(z) = \sum \alpha_i \mathcal{L}(u_i) = \sum \alpha_i w_i \Rightarrow \|\mathcal{L}(z)\|^2 = \sum |\alpha_i|^2 = \|z\|^2$$

↳ Isometry ...

This means we can extend ⊛ to $\ker A \longmapsto \operatorname{Im}^\perp(A)$

⊛ By rank-nullity, $\dim(\ker(A)) = \dim(\operatorname{Im}^\perp A) = \dim V - \dim(\operatorname{Im} A)$

Pick any isometry Ψ from $\ker A$ to $\operatorname{Im}^\perp A$ (linear)

and define

$$\boxed{\mathcal{U} : V \longmapsto V \text{ by } \mathcal{U}(\underbrace{v_1}_{\in \ker^\perp A} + \overbrace{v_2}^{\in \ker A}) = \varphi(v_1) + \Psi(v_2)}$$

Now

$$x \xrightarrow{A} Ax$$

☑ Let $x \in \ker^\perp A$

$$x \xrightarrow{\sqrt{A^*A}} \sqrt{A^*A}\, x \xrightarrow{\varphi} Ax$$

☑ Let $x \in \ker A$

$$x \xrightarrow{A} 0$$

$$x \xrightarrow{\sqrt{A^*A}} 0 \xrightarrow{\Psi} 0$$

So we have just shown that $\boxed{A = \overbrace{\mathcal{U}}^{\text{unitary}} \circ \sqrt{A^*A}}$

This $\mathcal{U}$ is an isometry, surj, ... $\Rightarrow$ $\mathcal{U}$ is unitary

What if start at adjoint?

$$A^* = \hat{\mathcal{U}} \circ \sqrt{AA^*}$$

$$\hookrightarrow A = \underline{\sqrt{AA^*} \circ (\hat{\mathcal{U}})^*} =$$

So $\boxed{A = \sqrt{AA^*} \circ \mathcal{W}}$ $\leadsto$ unitary...

So

> For any $A : V \xrightarrow{\text{lin}} V$, there exists a unitary $\mathcal{U}_1$
> $\mathcal{U}_1 : V \longrightarrow V$ such that
> $$A = \sqrt{\cancel{A^*A^*}} \circ \mathcal{U}_1$$
> and there exists a unitary $\mathcal{U}_2 : V \longmapsto V$ such that
> $$A = \mathcal{U}_2 \circ \sqrt{A^*A}$$

Now, think complex numbers... $z = |z| e^{i\theta} = \sqrt{z\bar{z}}\, e^{i\theta}$

$\Rightarrow$ This called $\boxed{\text{POLAR DECOMPOSITION}}$

Now

$$A = \mathcal{U} \circ \sqrt{A^*A} \longrightarrow \boxed{\cancel{\mathcal{U}^*A = \sqrt{A^*A}}}$$

$\cancel{A\mathcal{U}}$

$\cancel{A\mathcal{U}^* \sqrt{A}}$

$$= \big(\underbrace{\mathcal{U} \circ \sqrt{A^*A}\ \mathcal{U}^*}_{\text{nonnegative, called } N}\big)\mathcal{U}$$

$$N^2 = (\mathcal{U}A^*)(A\mathcal{U}^*)$$

$\cancel{\&}$ Now $AA^* = \mathcal{U}(A^*A)\mathcal{U}^* = (\mathcal{U}A^*)(A\mathcal{U}^*) = N^2 \Rightarrow N = \sqrt{AA^*}$

$\therefore$ $A = N\mathcal{U} = \sqrt{AA^*}\,\mathcal{U}$

$\therefore$ $\boxed{A = \mathcal{U}\sqrt{A^*A} = \sqrt{AA^*}\,\mathcal{U}}$

$\therefore$ We can use the same unitary in the polar decomp.
So can just use $\mathcal{U}_1 = \mathcal{U}_2$.

Note $\mathcal{U}$ is not necessarily unique

If A is invertible, then there's a unique choice for $\mathcal{U}$.

Now, if $A \in M_n$, then $A = \mathcal{U} \cdot \underbrace{\sqrt{A^*A}}_{\text{nonnegative, self adj, normal}}$

$= \mathcal{U}\,(\hat{\mathcal{U}} D \hat{\mathcal{U}}^{-1})$

diag, nonnegative.

This says

$$A = (\mathcal{U}\hat{\mathcal{U}})\,\hat{D}\,.\,\hat{\mathcal{U}}^{-1}$$

diag

SVD for square matrix

$$A = \mathcal{U}_1 D\, \mathcal{U}_2$$

unitaries

, and $D = \begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{bmatrix}$

where s_i are eigen vals of $\sqrt{A^*A}$

s' are called "singular values" of A

This is $\boxed{\text{SINGULAR VALUE DECOMPOSITION}}$

for square matrices

(?) What about $\sqrt{AA^*}$?

May 9, 2019

$\boxed{\text{SVD}}$

For any $A \in \mathbb{M}_n$ we can always find an orthonormal basis $v_1 \dots v_n$ such that $A(v_1) \dots A(v_n)$ orthogonal

i.e. there is a unitary $W \in \mathbb{M}_n$ (whose columns are $v_1 \dots v_n$) such that

$$A \circ W = A \circ [v_1 \dots v_n] = [Av_1 \ Av_2 \dots Av_n]$$

$$= \underbrace{\begin{bmatrix} | & & | \\ z_1 & \cdots & z_n \\ | & & | \end{bmatrix}}_{\text{unitary}} \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix} \quad (\alpha_i \geq 0)$$

"scaled each columns by a nonnegative factor"

i.e. s.t. $\boxed{A \circ W = \mathcal{U} \circ \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix}}$ nonnegative

i.e. s.t. $\boxed{A = \mathcal{U} \circ \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix} \circ W^*}$

☒ If $A = \mathcal{U} \underset{\geq 0}{D} W$, $\mathcal{U}, W$ unitary, D diagonal

then $A^* = W^* D \mathcal{U}^*$

So $\underset{\geq 0}{A^*A} = W^* D^2 W = \underbrace{(W^* D W)}_{\geq 0}^2$ → Similarly $\boxed{\mathcal{U} D \mathcal{U}^* = \sqrt{AA^*}}$

So $\boxed{W^* D W = \sqrt{A^*A} \longleftrightarrow \alpha_i\text{'s are exactly the eigenvalues of } \sqrt{A^*A}}$

☒ Now, $\mathcal{U} W W^* D W = \mathcal{U} W \sqrt{A^*A}$

So $A = \mathcal{U} W \sqrt{A^*A}$ → polar decomposition

<u>Now</u> $\left\{ \begin{array}{l} W^* D W = \sqrt{A^*A} \\ U D U^* = \sqrt{AA^*} \end{array} \right\}$

$$\Rightarrow \boxed{\sqrt{A^*A} \cong \sqrt{AA^*} \quad \text{(unitarily equivalent)}}$$

In particular, $U D^2 U^* = AA^*$
$W^* D^2 W = A^*A$

<u>So</u>

$$\boxed{A^*A \cong AA^* \quad \text{(unitary equivalent)}}$$

$\Big\{$ So... $\sqrt{A^*A}$ & $\sqrt{AA^*}$ have the same eigenvalues (eigen properties...) $\Big\}$
Same with A^*A & AA^*

What if A is not square, i.e. $M_{m \times n}$?

Say, $\hat{A} = \begin{bmatrix} A_{4\times 7} \\ 0 \end{bmatrix}_{7\times 7} \underbrace{\begin{bmatrix} v_1 & \cdots & v_7 \end{bmatrix}}_{\text{orthonormal basis s.t.}} = \underbrace{\begin{pmatrix} A(v_1) \\ 0 \end{pmatrix} \cdots \begin{pmatrix} A(v_7) \\ 0 \end{pmatrix}}_{\text{orthogonal}}$

with $\hat{A}(v_1)$ and $\hat{A}(v_7)$ written above the first and last columns.

$\left\{ \begin{array}{l} A : \mathbb{C}^7 \longmapsto \mathbb{C}^4 \\ \hat{A} : \mathbb{C}^7 \longmapsto \mathbb{C}^7 \end{array} \right. \Rightarrow$ orthonormal basis s.t.

And $\hat{A}(v_1) \ldots \hat{A}(v_7)$ orthogonal exactly when $A(v_1) \ldots A(v_7)$ orthogonal...

$$\rightarrow \quad \underset{4\times 7}{\boxed{\;A\;}} \underbrace{\begin{bmatrix} v_1 & \cdots & v_7 \end{bmatrix}}_{W} = \begin{bmatrix} A(v_1) & \cdots & A(v_7) \end{bmatrix}_{4\times 7}$$

$A : \mathbb{C}^7 \longmapsto \mathbb{C}^4$

<u>So</u> to put 0's at the back

$$\boxed{\ A\ }\ \underbrace{W \cdot \overbrace{\text{Permutation}}^{P}}_{\text{unitary},\ \hat{W}^*} = \underbrace{\begin{bmatrix} | & | & | & | & | & 0 & 0 & 0 \\ & & & & & 0 & 0 & 0 \\ & & & & & 0 & 0 & 0 \\ & & & & & 0 & 0 & 0 \end{bmatrix}_{4\times 7}}_{\text{unitary}} \begin{bmatrix} \alpha_1 & & & & & & \\ & \alpha_2 & & & & & \\ & & \alpha_3 & & & & \\ & & & \ddots & & & \\ & & & & 0 & & \\ & & & & & 0 & \\ & & & & & & \alpha_7 \end{bmatrix}$$

<u>So</u>

$$\underset{4\times7}{\boxed{\ A\ }} = \underset{4\times7}{\begin{bmatrix} \mathcal{U} & 0 \end{bmatrix}} \underset{7\times7}{\begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ \hline & & 0\ 0\ \alpha_7 \end{bmatrix}} \underset{7\times7}{\begin{bmatrix} \hat{W} \end{bmatrix}}$$

$$= \underset{4\times4}{\boxed{\ \mathcal{U}\ }} \begin{bmatrix} \alpha_1 & & & \\ & \ddots & & \\ & & & \alpha_4 \end{bmatrix}\!\!\begin{array}{|c|} 0 \end{array} \underset{7\times7}{\begin{bmatrix} \hat{W} \end{bmatrix}}$$

same shape as A

For $\begin{bmatrix} A \end{bmatrix}$ (tall), do same thing, with A^*

What are the α_i's?

$$D = \begin{bmatrix} S & 0 \end{bmatrix}$$

$$A = \mathcal{U} D W$$
$$A^* = W^* D^* \mathcal{U}^* \quad \Rightarrow \quad D^* = \begin{bmatrix} S \\ \hline 0 \end{bmatrix}$$

$$\Rightarrow \sqrt{A^*A} = W^* \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} W \qquad \rightarrow \quad D^*D = \begin{bmatrix} S^2 & 0 \\ 0 & 0 \end{bmatrix} \rightarrow \sqrt{D^*D} = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$$

$\begin{bmatrix} A^* \end{bmatrix} \begin{bmatrix} A \end{bmatrix}$

$\sqrt{AA^*}$

$\begin{bmatrix} A \end{bmatrix} \begin{bmatrix} A^* \end{bmatrix}$ $\rightsquigarrow$ $DD^* = \boxed{S}$

~~→ In fact, one~~

<u>The point is</u>

↳ Eigenvalues of A^*A, AA^* are still the same, 4 of them ...

→ SVD for non square

In fact, can derive the non-square version of polar decomposition ... using the same recipe.

Consequences

① Recall Hilbert-Schmidt norm
← Frobenius...

$$\|A\|^2_{F/HS} = tr(A^*A)$$
$$= tr\left((W^*D^*U^*)(UDW)\right)$$
$$= tr\left(\underbrace{W^*}\underbrace{D^*DW}\right)$$

Now, since $tr(BC) = tr(CB)$, so...

So $\boxed{\|A\|^2_{F/HS} = tr\left[D^*DWW^*\right] = tr(D^*D) = \sum s_i^2(A)}$

no need square matrices for this to work ...

sum of squares of singular values of A

② Easy to check... $\boxed{\|U_1 A U_2\|^2_{F/HS} = \dots = \|A\|^2_{F/HS}}$

"sandwiching A with unitaries don't change size."

③ Let $\mathcal{F}_k$ be $n \times n$ matrices of rank k. Given $A \in M_n$, what is the $\inf \|A - F\|_{F/HS}$ for $F \in \mathcal{F}_k$.

i.e. How small can $\|A - F\|_{HS}$ be?

Observe...

<u>Observe</u> $A = \mathcal{U} D \mathcal{W}$

then $$\|\mathcal{U}D\mathcal{W} - F\|_{HS} = \|\mathcal{U}D\mathcal{W} - \mathcal{U}\mathcal{U}^*F\mathcal{W}^*\mathcal{W}\|_{HS}$$
$$= \|\mathcal{U}(D - \mathcal{U}^*F\mathcal{W}^*)\mathcal{W}\|_{HS}$$

HS norm is unitarily invariant

$$= \|D - \underbrace{\mathcal{U}^*F\mathcal{W}^*}_{\Downarrow \atop \Uparrow \atop F_k}\|_{HS}$$

▨ Now the question becomes how small can we make $\|D - \overbrace{\mathcal{U}^*F\mathcal{W}^*}^{T}\|_{HS}$?

<u>So</u> $$\Big\| \underbrace{\begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_5 \end{pmatrix}}_{D} - T \Big\|_{HS}^2 \leq \Big\| \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_5 \end{pmatrix} - \underbrace{\begin{pmatrix} \alpha_1 & & & & \\ & \ddots & & & \\ & & \alpha_3 & & \\ & & & 0 & \\ & & & & 0 \end{pmatrix}}_{\Uparrow \atop F_3} \Big\|_{HS}^2$$

Let T rank 3, D rank 5 $\qquad = \boxed{\alpha_4^2 + \alpha_5^2}$

▨ Turns out there's a minimum approximation

<u>Detour</u>... if $V = W \oplus W^\perp$, $\underset{\uparrow \atop \text{(ortho-projection)}}{P_W(z)}$ is an element of W

<u>and</u>
$$z = P_W(z) + (I - P_W)(z)$$

orthogonal

then $\boxed{\|z\|^2 = \|P_W(z)\|^2 + \|(I-P_W)(z)\|^2}$ (Pythagorean)

$\cancel{\|z\|^2 = \|P_W(z) \ldots (I-P_W)}$ <u>So</u> $P_W(z)$ is an element of W that is closest to z.

$\in W$

$\boxed{\text{Strange Question}}$

Suppose $v_1 \dots v_7 \in \mathbb{C}^7$. How do we find a 3D subspace $W < \mathbb{C}^7$ s.t.

$$(i) \quad \| P_W(v_1)\|^2 + \dots + \| P_W(v_7)\|^2 \text{ is } \begin{cases} \max ? \\ \min ? \end{cases}$$

$$+ \qquad\qquad +$$

$$\underbrace{(ii) \quad \| P_{W^\perp}(v_1)\|^2 + \dots + \| P_{W^\perp}(v_7)\|^2}_{\|v_1\|^2 + \dots + \|v_7\|^2} \text{ is } \begin{cases} \min ? \\ \max ? \end{cases}$$

What if (v_i) is $(s_i e_i) \longrightarrow$ multiples of standard basis?

$$s_1 \geq s_2 \geq \dots \geq s_7 \geq 0$$

Given $W < V$, $\dim W = 3$. Take an ONB of W, say

$$w_1 = (w_{11} \dots w_{17})^T \quad w_2 = (w_{21} \dots w_{27})^T \quad w_3 = (w_{31} \dots w_{37})^T$$

$$P_W(s_1 e_1) = \langle s_1 e_1, w_1 \rangle w_1 + \langle s_1 e_1, w_2 \rangle w_2 + \langle s_1 e_1, w_3 \rangle w_3$$

$$= s_1 \left[\overline{w_{11}} \cdot w_1 + \overline{w_{21}}\, w_2 + \overline{w_{31}}\, w_3 \right]$$

So $$\| P_W(s_1 e_1)\|^2 = \underbrace{|s_1|^2}_{s_1^2} \left(|w_{11}|^2 + |w_{21}|^2 + |w_{31}|^2 \right)$$

$\vdots$

$$\| P_W(s_j e_j)\|^2 = s_j^2 \overbrace{\left(|w_{1j}|^2 + |w_{2j}|^2 + |w_{3j}|^2 \right)}^{t_j}$$

Check $\boxed{t_1 + \dots + t_7 = 3}$

So $$\sum_j \| P_W(s_j e_j)\|^2 = \sum s_j^2 t_j = s_1^2 t_1 + s_2^2 t_2 + \dots s_7^2 t_7$$

To make this as small as possible ... $\rightarrow$ need all the weight on $s_4, s_5, s_6 \ldots$

$\rightarrow$ need $W = \text{span}(e_4, e_5, e_6)$

Back to exp

$$\left\| \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_7 \end{pmatrix} - \begin{pmatrix} | & & | \\ c_1 & - & c_7 \\ | & & | \end{pmatrix} \right\|^2$$

$$= \left\| \begin{bmatrix} | & | & & | \\ s_1 e_1 - c_1 & s_2 e_2 - c_2 & \ldots & s_5 e_5 - c_5 \\ | & | & & | \end{bmatrix} \right\|^2$$

$$= \| s_1 e_1 - c_1 \|^2 + \ldots + \| s_5 e_5 - c_5 \|^2$$

now, $\underbrace{\text{span}\begin{pmatrix} c_1, c_2, c_3 \\ c_4, c_5 \end{pmatrix}}_{W}$ is 3D

the closest to $s_1 e_1$ for c_1 is to set $P_W(s_1 e_1) = c_1$

So $c_j = P_W(s_j e_j)$

So

$$\| (I - P_W)(s_1 e_1) \|^2 + \ldots + \| (I - P_W)(s_5 e_5) \|^2$$

$$= \| P_{W^\perp}(s_1 e_1) \|^2 + \ldots + \| P_{W^\perp}(s_5 e_5) \|^2$$

"minimizing one is maximizing the other"