MATH 2040 Lecture 20 (21/11/2016)

§ Jordan canonical forms (§ 7.1 in textbook)

Question: What else can we do if NOT diagonalizable?

Schur's Lemma: Assume char. poly. splits over $\mathbb{F}$.

$$\Rightarrow \qquad [T]_\beta = \begin{pmatrix} \ddots & * \\ 0 & \ddots \end{pmatrix} \qquad \text{upper triangular}$$

Prototype (Non-diagonalizable / $\mathbb{C}$)

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

Ans: "canonical form(s)": Jordan modeled on these

Theorem: $T : V \to V$ s.t. char. poly. splits / $\mathbb{F}$ (automatic when $\mathbb{F} = \mathbb{C}$)

$\Rightarrow \exists$ basis β for V s.t. $[T]_\beta = J$

$$J = \begin{pmatrix} \boxed{A_1} & & & 0 \\ & \boxed{A_2} & & \\ & & \ddots & \\ 0 & & & \boxed{A_k} \end{pmatrix} \qquad \boxed{A_i} = \overbrace{\begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}}^{m} \Big\} m$$

"Jordan canonical form"

"Jordan block" corr. to λ with size m

ASSUME: $\boxed{\mathbb{F} = \mathbb{C}}$ ($\Rightarrow$ char. poly. splits)

Examples (of smaller matrices)

$n=1$: $A = (a)$

$n=2$: $\begin{pmatrix} \boxed{1} & \\ & \boxed{2} \end{pmatrix}$, $\begin{pmatrix} \boxed{1} & \\ & \boxed{1} \end{pmatrix}$, $\left(\boxed{\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}}\right)$

$$2 = 1+1 \\ = 2+0$$

$n=3$: $\begin{pmatrix} \boxed{1} & & \\ & \boxed{2} & \\ & & \boxed{3} \end{pmatrix}$, $\left(\boxed{\begin{matrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{matrix}}\right)$, $\begin{pmatrix} \boxed{1} & \\ & \boxed{\begin{matrix} 1 & 1 \\ & 1 \end{matrix}} \end{pmatrix}$

$\begin{pmatrix} \boxed{1} & \\ & \boxed{\begin{matrix} 2 & 1 \\ & 2 \end{matrix}} \end{pmatrix}$ .....

$$3 = 1+1+1 \\ = 3+0 \\ = 2+1$$

$n=4$: Exercise. Write down "all" examples.

$$4 = 1+1+1+1 \\ = 3+1 \\ = 2+2 \\ = 2+1+1$$

Goal: If $\exists\, \beta$ s.t. $[T]_\beta$ = Jordan can. form

then what does the vectors in β have to satisfy?

$\Rightarrow$ "generalized eigenvectors"

1st: We look at one Jordan block:

$$A = \underbrace{\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}}_{m \times m}$$

Jordan block

Observation:

(1) char. poly of $A = f(t) = (-1)^m (t-\lambda)^m$

$\Rightarrow$ only 1 eigenvalue $= \lambda$

(2) $E_\lambda = \text{span}\{\vec{e}_1\}$ dim $= 1$

$\because E_\lambda = N(A - \lambda I)$

$$= N\begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} = \text{span}\{\vec{e}_1\}.$$

Q: $\vec{e}_1$ is e.vector, but what $\vec{e}_2, \dots, \vec{e}_m$?

Let $\{\vec{e}_1, \dots, \vec{e}_m\}$ std basis for $\mathbb{C}^m$.

$$\begin{cases} A\vec{e}_1 = \lambda \vec{e}_1 \\ A\vec{e}_2 = 1\cdot\vec{e}_1 + \lambda\cdot\vec{e}_2 \\ A\vec{e}_3 = 1\cdot\vec{e}_2 + \lambda\cdot\vec{e}_3 \\ \vdots \qquad \vdots \\ A\vec{e}_m = 1\cdot\vec{e}_{m-1} + \lambda\cdot\vec{e}_m \end{cases} \Rightarrow \begin{cases} (A - \lambda I)\vec{e}_1 = \vec{0} \\ (A - \lambda I)\vec{e}_2 = \vec{e}_1 \\ (A - \lambda I)\vec{e}_3 = \vec{e}_2 \\ \vdots \qquad \vdots \\ (A - \lambda I)\vec{e}_m = \vec{e}_{m-1} \end{cases}$$

i.e. $\boxed{(A - \lambda I)^i \vec{e}_i = \vec{0} \quad \forall i}$

Def^n: $\vec{v} \in V$ is a generalized eigenvector corr. to λ, & A

if (1) $\vec{v} \neq \vec{0}$

(2) $\exists\, p \in \mathbb{N}$ s.t. $(A - \lambda I)^p \vec{v} = \vec{0}$

Note: e.vectors $\Rightarrow$ gen. e.vectors ; p depends on $\vec{v}$

$p = 1$

Def$^{\underline{n}}$: $K_\lambda = \{ \vec{v} \in V : \text{gen. e.vectors corr. to } \lambda \}$

"generalized eigenspace of λ"

Remark: λ has to be an eigenvalue!

$\because$ take any $\underset{\neq \vec{0}}{\vec{v}} \in K_\lambda \Rightarrow \exists\, p$ "1" s.t. $(T-\lambda I)^p \vec{v} = \vec{0}$

the smallest integer

then $(T-\lambda I)\left[\underbrace{(T-\lambda I)^{p-1}\vec{v}}_{\neq \vec{0}}\right] = \vec{0}$

$\Rightarrow$ is e.vector with e.value $= \lambda$

Remember: $T: V \to V$ diagonalizable / $\mathbb{F}$

$\Leftrightarrow \quad V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$

(not diag. $\Rightarrow \quad V \supsetneq E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$)

Theorem: $T: V \to V$ s.t. char. poly. splits (e.g. $\mathbb{F} = \mathbb{C}$)

$$\Rightarrow \quad \boxed{V = K_{\lambda_1} \oplus \cdots \oplus K_{\lambda_k}}$$

Lemma: (Properties of K_λ)

1) K_λ is T-invariant (1) subspace (2) & $E_\lambda \subseteq K_\lambda$ (3)

2) If $\mu \neq \lambda$, then

$$T-\mu I\big|_{K_\lambda} : K_\lambda \xrightarrow{\cong} K_\lambda \quad \text{1-1 \& onto.}$$

(i.e. isomorphism)

3) Let m = multiplicity of λ.

$$\dim K_\lambda \underset{(=)}{\leq} m \qquad \text{and} \qquad K_\lambda = N\left((T-\lambda I)^m\right)$$

Proof: 1) ✓ 3 $E_\lambda \subseteq K_\lambda$ $\because$ trivial $(p=1)$

✓ 2 K_λ subspace :

i) $\vec{0} \in K_\lambda$ $\because$ trivial

ii) closed under $\cdot$: $\vec{v} \in K_\lambda \Rightarrow (T-\lambda I)^p \vec{v} = \vec{0}$ for some p

$$\Rightarrow (T-\lambda I)^p(\mu \vec{v}) = \vec{0}$$

$$\Rightarrow \mu \cdot \vec{v} \in K_\lambda$$

iii) closed under $+$:

$$\vec{v}, \vec{w} \in K_\lambda \Rightarrow (T-\lambda I)^p \vec{v} = \vec{0}, \; (T-\lambda I)^q \vec{w} = \vec{0}$$

$$\Rightarrow (T-\lambda I)^{p+q}(\underbrace{\vec{v}+\vec{w}}_{\in K_\lambda}) = \vec{0}$$

1 : K_λ T-invariant!

Take $\vec{v} \in K_\lambda \Rightarrow (T-\lambda I)^p \vec{v} = \vec{0}$ for some $p \geq 1$

$$(T-\lambda I)^p \, T\vec{v} = T\,\underbrace{(T-\lambda I)^p \vec{v}}_{=\vec{0}} = \vec{0}$$

commutes

(2) $^{(*)}$ <u>Remember</u>: $E_\lambda \cap E_\mu = \{\vec{0}\}$ if $\lambda \neq \mu$.

<u>Claim</u>: $T - \mu I : K_\lambda \longrightarrow K_\lambda$ is 1-1 ($\Rightarrow$ onto)

Assume $\underset{\neq \vec{0}}{\vec{v}} \in K_\lambda$ s.t.

$$(T - \mu I)\vec{v} \overset{(*)}{=} \vec{0} \quad (\Rightarrow \vec{v} \in E_\mu)$$

Choose smallest p s.t. $(T - \lambda I)^p \vec{v} = \vec{0}$.

$$(T - \mu I)\big(\underline{(T - \lambda I)^{p-1}\vec{v}}\big) \overset{(*)}{=} \vec{0}$$

commutes

$\hookrightarrow$ ① $\in E_\mu$ $\overset{(*)}{\Rightarrow} = \vec{0}$

② $\in E_\lambda$ contradicts the choice of p.

(3) $d = \dim K_\lambda \leq m$.

$\because$ $\underbrace{\text{char poly. of } T|_{K_\lambda}}_{(2)\ \Downarrow\ g(t) = (-1)^d (t-\lambda)^d}$ divides $\underbrace{\text{char. poly. of } T}_{f(t) = (-1)^n (t-\lambda)^m \cdots\cdots}$

$$\Rightarrow \quad d \leq m.$$

$K_\lambda = N\big((T - \lambda I)^m\big)$: "$\supseteq$" trivial

"$\subseteq$" By Cayley-Hamilton,

$$(T - \lambda I)^{\underset{\wedge\!\!\!\mid\mid}{d}} = 0 \quad \text{on } K_\lambda$$

$$\Rightarrow \quad (T - \lambda I)^m = 0 \quad \text{on } K_\lambda$$