Week 14: Jordan Canonical Forms (Textbook §7.1, 7.2)

Recall: Any $A \in M_{n\times n}(\mathbb{C})$ can be changed into a Jordan canonical form J by a change of basis, i.e. $\exists$ invertible $Q \in M_{n\times n}(\mathbb{C})$ s.t.

$$Q^{-1}AQ = J = \begin{pmatrix} \boxed{\begin{matrix}\lambda_1 & 1 & \\ & \ddots & 1 \\ & & \lambda_1\end{matrix}} & & \\ & \ddots & \\ & & \boxed{\begin{matrix}\lambda_N & 1 & \\ & \ddots & 1 \\ & & \lambda_N\end{matrix}} \end{pmatrix}$$

Jordan blocks

We now address the

Question: Why can we always do that?

To answer it we need to study more carefully about the "generalized eigenspaces", i.e. for any eigenvalue $\lambda \in \mathbb{C}$ of A

$$\boxed{K_\lambda := \{ x \in \mathbb{C}^n \mid (A - \lambda I)^p x = 0 \text{ for some positive integer } p \}}$$

Lemma: (a) K_λ is a T-invariant subspace for $T = L_A : \mathbb{C}^n \to \mathbb{C}^n$.

(b) $E_\lambda \subset K_\lambda$

(c) For any $\mu \neq \lambda$, $L_{A-\mu I} : K_\lambda \to K_\lambda$ is one-to-one.

Proof: (a) Claim: K_λ is a subspace. (Exercise!)

Claim: K_λ is T-invariant.

Suppose $x \in K_\lambda$. Then $\exists\, p \geq 1$ s.t. $(A-\lambda I)^p x = 0$

$$\Rightarrow (A-\lambda I)^p A x = A \underbrace{(A-\lambda I)^p x}_{=0} = 0$$

commutes

Thus, $Ax = L_A x = Tx \in K_\lambda$.

(b) is obvious (take $p=1$).

(c) Suppose NOT. Then $\exists\ x \in K_\lambda$ s.t. $x \neq 0$ and $\boxed{(A-\mu I)x = 0}$

Since $x \in K_\lambda$, we can choose $p \geq 1$ to be the smallest positive integer s.t. $(A-\lambda I)^p x = 0$.

Define $y = (A-\lambda I)^{p-1} x \ \boxed{\neq 0}$ since p is "smallest".

Clearly, $y \in E_\lambda$.

Moreover, $y \in E_\mu$ since

$$(A-\mu I)y = (A-\mu I)(A-\lambda I)^{p-1}x = (A-\lambda I)^{p-1}\underbrace{(A-\mu I)x}_{=0}$$

commutes

Therefore $y \in E_\lambda \cap E_\mu$ but distinct eigenspaces have intersection $= \{0\}$. Therefore, $y = 0$ Contradiction!

$\square$

Lemma: Suppose $\lambda \in \mathbb{C}$ is an eigenvalue of A with multiplicity m.

Then, $\boxed{K_\lambda = N(A-\lambda I)^m}$ and $\boxed{\dim K_\lambda \leq m}$

This provides a way to find K_λ.

Proof: Let $W = K_\lambda$, which is a T-invariant subspace by previous lemma, where $T = L_A$. Suppose $\dim W = d$.

Claim: $d \leq m$

Consider the restriction $T_W : W \to W$, by (c) of previous lemma, T_W has no eigenvalue other that λ. Therefore

$$(-1)^d (t-\lambda)^d = \text{char. poly.}(T_W) \mid \text{char. poly.}(T) = (-1)^n (t-\lambda)^m \cdots\cdots$$

$$\Longrightarrow \quad d \leq m.$$

For the rest, since $N(A-\lambda I)^m \subset K_\lambda$ by definition and by Cayley-Hamilton theorem, $(A-\lambda I)^d x = 0 \ \ \forall x \in K_\lambda$. Done!

$\square$

Now we are ready to prove one of the main results about Jordan canonical forms:

<u>Theorem</u>: Let $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ be the distinct eigenvalues of $A \in M_{n\times n}(\mathbb{C})$.

Then, $\boxed{\mathbb{C}^n = K_{\lambda_1} \oplus K_{\lambda_2} \oplus \cdots\cdots \oplus K_{\lambda_k}}$

<u>Proof</u>: We split the proof into 2 steps:

<u>Step 1</u>: Show that $\boxed{\mathbb{C}^n = K_{\lambda_1} + K_{\lambda_2} + \cdots\cdots + K_{\lambda_k}}$

i.e. $\forall\, x \in \mathbb{C}^n$, there exists $v_i \in K_{\lambda_i}$ s.t.

$$x = v_1 + v_2 + \cdots\cdots + v_k .$$

The proof is by induction on k, the number of eigenvalues.

<u>When $k=1$</u>: char poly of $A = (-1)^n (t-\lambda_1)^n$.

Cayley-Hamilton $\Rightarrow$ $(A-\lambda_1 I)^n = O$

Previous lemma $\Rightarrow$ $K_{\lambda_1} = N(A-\lambda_1 I)^n = \mathbb{C}^n$. <u>done!</u>

<u>Assume the result is true for $k-1$ distinct eigenvalues.</u>

Now, suppose there are k distinct eigenvalues $\lambda_1, \dots, \lambda_k$.

char. poly of $A = (-1)^n (t-\lambda_1)^{m_1} (t-\lambda_2)^{m_2} \cdots (t-\lambda_k)^{m_k}$

Define $\boxed{W = R(A-\lambda_k I)^{m_k}}$

<u>Claim</u>: W is T-invariant where $T = L_A$

<u>Pf</u>: Let $y = (A-\lambda_k I)^{m_k} x \in W$

$$\Rightarrow \quad Ay = A(A-\lambda_k I)^{m_k} x = (A-\lambda_k I)^{m_k}(Ax) \in W .$$

For each $i \neq k$, by previous lemma (c),

(*) $\boxed{L_{(A-\lambda_k I)^{m_k}} : K_{\lambda_i} \xrightarrow{\cong} K_{\lambda_i} \quad \text{is an isomorphism.}}$

$\Rightarrow \quad \boxed{K_{\lambda_i} \subseteq W}$ for each $\lambda_i \neq \lambda_k$.

Consider $T_W : W \to W$, which has eigenvalues

$\lambda_1, \dots, \lambda_{k-1}$ since $K_{\lambda_i} \subset W$ for $i < k$.

To see why λ_k is not an eigenvalue of T_W:

suppose $\exists\, v \in W$ s.t. $T_W v = Av = \lambda_k v$.

Since $v \in W$, $v = (A - \lambda_k I)^{m_k} y$ for some $y \in \mathbb{C}^n$

$\Rightarrow (A - \lambda_k I) v = (A - \lambda_k I)^{m_k+1} y = 0$

ie $y \in K_{\lambda_k} \overset{\text{lemma}}{\Longrightarrow} (A - \lambda_k I)^{m_k} y = v = 0$. Contradiction!

Induction hypothesis satisfied by T_W.

Let $x \in \mathbb{C}^n$, then $(A - \lambda_k I)^{m_k} x \in W$. By induction hypothesis $\exists\, w_i \in K_{\lambda_i}$ s.t.

$$(A - \lambda_k I)^{m_k} x = w_1 + w_2 + \dots + w_{k-1}$$

By (*), $w_i = (A - \lambda_k I)^{m_k} v_i$ for some unique $v_i \in K_{\lambda_i}$

$$\Rightarrow (A - \lambda_k I)^{m_k} \underbrace{(x - v_1 - v_2 - \dots - v_{k-1})}_{= v_k \in K_{\lambda_k}} = 0$$

Therefore,
$$x = \underset{K_{\lambda_1}}{\underset{\cap}{v_1}} + \underset{K_{\lambda_2}}{\underset{\cap}{v_2}} + \dots + \underset{K_{\lambda_{k-1}}}{\underset{\cap}{v_{k-1}}} + \underset{K_{\lambda_k}}{\underset{\cap}{v_k}}.$$

Step 1 done. □

Step 2: Show that if β_i is an ordered basis for K_{λ_i}, then $\beta = \beta_1 \cup \dots \cup \beta_k$ is an ordered basis for $\mathbb{C}^n$.

Note: $\beta_i \cap \beta_j = \phi$ for $i \neq j$ since $L_{A - \lambda_i I} : K_{\lambda_j} \xrightarrow{\cong} K_{\lambda_j}$ for $\lambda_i \neq \lambda_j$.

β clearly spans $\mathbb{C}^n$ by Step 1.

Let $\#\beta = q \geq n$. But by Lemma,

$$q = \sum \dim K_{\lambda_i} \leq \sum m_i = n$$

Hence, $q = n$ and β is a basis for $\mathbb{C}^n$. Moreover, $\boxed{\dim K_{\lambda_i} = m_i}$

Finally, we just need to pick some "good basis" β_i for each K_{λ_i}.

This is given by "cycles", here λ is an eigenvalue

$$\gamma = \{ \underbrace{(A - \lambda I)^{p-1} x}_{v_1}, \underbrace{(A - \lambda I)^{p-2} x}_{v_2}, \dots, \underbrace{x}_{v_p} \}$$

and that $p \geq 1$ is the "smallest" s.t. $(A - \lambda I)^p x = 0$

Observe that:

$$(A - \lambda I) v_1 = 0 \implies A v_1 = \lambda v_1$$

$$(A - \lambda I) v_2 = (A - \lambda I)^{p-1} x = v_1 \implies A v_2 = v_1 + \lambda v_2$$

$$(A - \lambda I) v_3 = (A - \lambda I)^{p-2} x = v_2 \implies A v_3 = v_2 + \lambda v_3$$

$$\vdots \qquad\qquad\qquad \vdots$$

Therefore, let $W = \operatorname{span} \gamma$

$$[L_A|_W]_\gamma = \left(\boxed{\begin{matrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{matrix}} \right)$$

a Jordan block!

Question: Why should γ be linearly independent?

Lemma: γ is linearly independent (as long as $x \neq 0$).

Proof: Let $W = \operatorname{span} \gamma$, which is $\mathcal{U}$-invariant, where

$$\mathcal{U} = L_{(A - \lambda I)} .$$

We prove the lemma by induction on p.

The case for $p = 1$ is trivial since $x \neq 0$.

Assume lemma holds for $\# \gamma \leq p-1$.

Now, if $\# \gamma = p$, consider the cycle

$$\gamma' = \{ (A - \lambda I)^{p-1} x, \dots, (A - \lambda I) x \} \quad \text{with} \quad \boxed{\# \gamma' = p-1}$$

which by induction hypothesis is linearly independent!

Note that $\gamma' = \mathcal{U}(\gamma)$. $\Rightarrow$ γ' is a basis for $R(\mathcal{U}|_W)$

$$\boxed{\text{Rank-nullity Theorem}} \Rightarrow \underbrace{\operatorname{nullity}(\mathcal{U}|_W)}_{\geq 1} + \underbrace{\operatorname{rank}(\mathcal{U}|_W)}_{= p-1} = \dim W \leq p$$

$$\Rightarrow \quad p \leq \dim W = \# \gamma = p \text{, ie } \gamma \text{ is linearly indep.}$$

$\square$

By similar ideas, one can prove that

Lemma: If $\gamma_1, \gamma_2, \dots, \gamma_q$ are cycles of generalized eigenvectors of A corresponding to λ generated by $v_i \in \gamma_i$,

assume the initial vectors

$$\{ (A - \lambda I)^{p_1 - 1} v_1, (A - \lambda I)^{p_2 - 1} v_2, \dots, (A - \lambda I)^{p_q - 1} v_q \}$$

form a linearly indep. subset,

then $\gamma = \gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_q$ is linearly independent.

Proof: Exercise (see textbook Thm. 7.6)

The following theorem completes the picture.

Theorem: Every K_λ has an ordered basis consisting of disjoint union of cycles:

$$\gamma = \gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_q .$$

disjoint cycles

Proof: Basically by induction on $\dim K_\lambda$. (see textbook Thm. 7.7 for more details). $\square$