Lecture 32: General Jordan canonical forms - Cayley-Hamilton

Recall: Last time we talked about minimal polynomials of matrices

① Given $A \in Mat_{n\times n}(\mathbb{K}) \rightsquigarrow$

subring gen by $\mathbb{K}$ & A. (commutative!)

$$\Psi: \mathbb{K}[x] \longrightarrow \mathbb{K}[A] \subset End_{\mathbb{K}}(\mathbb{K}^n)$$

$$P(x) = \sum_{i=0}^{l} a_i x^i \longmapsto P(A) = \sum a_i A^i \qquad (A^0 = Id.)$$

- Ψ ring homomorphism, & $\mathbb{K}$-linear
- $\operatorname{Ker}\Psi \neq (0) \rightsquigarrow \operatorname{Ker}\Psi = (q_A)$ q_A monic in $\mathbb{K}[x]$ = minimal poly of A.

Obs: $q_A = q_{G^{-1}AG}$ for all $G \in Gl_n(\mathbb{K})$

② We defined Rational Normal forms & Jordan canonical forms (for $\mathbb{K} = \overline{\mathbb{K}}$) viewing $\mathbb{K}^n$ as a $\mathbb{K}[A]$-module ($\rightsquigarrow \mathbb{K}[x]$-module: $(\mathbb{K}^n)_{tor} = \mathbb{K}^n$)

- CASE 1: $\mathbb{K}^n$ cyclic ($\exists\, v \in \mathbb{K}^n$ st $\{v, Av, \ldots, A^k v, \ldots\}$ spans $\mathbb{K}^n$)

- CASE 2: Use classification of modules over PID to reduce to the cyclic case.

CASE 1: $\mathbb{K}^n$ is cyclic (gen by $\{v, Av, \ldots\}$)

- $q_A = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ ($\deg q_A = \underline{\min}\ d$ st $\{v, Av, \ldots, A^d\}$ is l d)

① $\exists\ B = \{v, Av, A^2v, \ldots, A^{n-1}v\}$ basis with

$$[A]_{BB} = \begin{bmatrix} 0 & & & -a_0 \\ 1 & \ddots & 0 & -a_1 \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & -a_{d-1} \end{bmatrix} = C_{q_A}$$

$[A]_{BB} = C_{(E,B)}\, A\, C_{(B,E)}$

(Companion matrix for the polynomial q_A.)

$q_A = x^n + a_{n-1}x^{n-1} + \cdots + a_0$

② If $\mathbb{K} = \overline{\mathbb{K}}$, char $\mathbb{K} = 0$ and $q_A = (x-\alpha)^n$ (cyclic case!), then $\exists\ w \in \mathbb{K}^n$ st

$$B = \{w, (A-\alpha I)(w), \ldots, (A-\alpha I)^{n-1}(w)\}$$

is a basis for some $v \in \mathbb{K}^n$

$$[A]_{BB} = \begin{bmatrix} \alpha & & & 0 \\ 1 & \alpha & & \\ & 1 & \ddots & \\ 0 & & 1 & \alpha \end{bmatrix} = J(\alpha, n) \quad \text{(Jordan block)}$$

CASE 2: $\mathbb{K}^n$ is not cyclic

We break $\mathbb{K}^n$ into cyclic $\mathbb{K}[A]$-modules

① $\mathbb{K}^n = V_1 \oplus \cdots \oplus V_r$ $\quad \mathbb{K}[A] \circlearrowright V_i = (\mathbb{K}^n)_{q_i}$ cyclic

$\exists\ \mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ ($\mathcal{B}_i$ basis for V_i) with

$$[A]_{\mathcal{B}\mathcal{B}} = \begin{bmatrix} C_{q_1} & & 0 \\ & \ddots & \\ 0 & & C_{q_r} \end{bmatrix}$$

with $q_1 \mid q_2 \mid \cdots \mid q_r$ so $q_r = q_A$.

(Rational Normal Form)

② $\mathbb{K} = \overline{\mathbb{K}}$ char $\mathbb{K} = 0$ $\quad q_A = \prod_{i=1}^{s} (x - \alpha_i)^{m_i}$ $\quad \alpha_i$'s distinct.

$\rightsquigarrow V = V_1' \oplus \cdots \oplus V_s'$ with $V_i' = (\mathbb{K}^n)_{(x-\alpha_i)^{m_i}}$ so $q_{A|_{V_i'}} = (x - \alpha_i)^{m_i}$

$\exists\ \mathcal{B}' = \mathcal{B}_1' \cup \cdots \cup \mathcal{B}_s'$ ($\mathcal{B}_i'$ basis for V_i') with

$$[A]_{\mathcal{B}'\mathcal{B}'} = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_s \end{pmatrix} \quad \& \quad A_i = [A|_{V_i'}]_{\mathcal{B}_i'\mathcal{B}_i'} = \begin{bmatrix} J(\alpha_i, m_1^{(i)}) & & 0 \\ & \ddots & \\ 0 & & J(\alpha_i, m_{s_i}^{(i)}) \end{bmatrix}$$

with $m_i = m_1^{(i)} \geq \cdots \geq m_{s_i}^{(i)}$ (Jordan canonical form)

Obs: Written additively, we get : $[A]_{BB} = D + N$

- D = diagonal part. $\rightsquigarrow q_D = \prod_{i=1}^{r} (x - \alpha_i)$
- N = nilpotent ($N^{\dim V} = 0$). $\rightsquigarrow q_N = x^\ell$ for some ℓ.
- $[A]_{BB}$, D & N commute

$\rightsquigarrow$ more general version of Jordan canonical forms, over perfect fields (Eg char 0).

Theorem: Let $\mathbb{K}$ be a perfect field, $n \in \mathbb{Z}_{\geq 1}$, $A \in \mathrm{Mat}_{n\times n}(\mathbb{K})$

Then, $\exists !$ $A_S, A_N \in \mathrm{Mat}_{n\times n}(\mathbb{K})$ st

(1) $A = A_S + A_N$. (Jordan-Chevalley decomposition of A)

(2) A_S & A_N are polynomial in A. , (2^*) A, A_S, A_N commute

(3) A_S is semisimple & A_N is nilpotent.

- A_N nilpotent means $q_{A_N}(x) = x^\ell$ for some ℓ
- A_S semisimple means $(q_{A_S}, q_{A_S}') = 1$. (q_{A_S}' = formal derivative of the polynomial q_A)

(Alternatively $q_{A_S} = \prod_{i=1}^{r} f_i(x)$ f_i = distinct monic irreducibles.)

Q: How to build A_S & A_N? $A = A_S + A_N$ A_N nilp, A_S semis $A_N = P(A)$, $A_S = Q(A)$.

Easy case: $q_A = (x-\lambda)^d \rightsquigarrow A_N = (A - \lambda I_n)$ & $A_S = A - A_N = \lambda I_n$

- A & A_N commute, A & A_S commute, A_N & A_S commute ✓
- A_N nilpotent $(A_N)^d = (A-\lambda I_n)^d = 0$ in $\mathrm{Mat}_{n\times n}(\mathbb{K})$
- Claim: $q_{A_S} = (x-\lambda)$ (so $(q_{A_S}, q'_{A_S}) = 1$)
  - $A - A_N - \lambda I_n = A - (A - \lambda I_n) - \lambda \mathrm{Id} = 0$ ✓
  - $x - \lambda$ is irreducible so $\ker \Psi = (x-\lambda)$ ($\ker \Psi \neq K[x]$)
- For $\mathbb{K} = \overline{\mathbb{K}}$, char 0, see HW 11 (semisimple = diagonalizable)
- For perfect fields, you need Galois Theory.

Q: Uniqueness? $A_S + A_N = A'_S + A'_N \rightsquigarrow A_S - A'_S = A'_N - A_N$

- $A'_N = P(A)$, $A_N = Q(A)$ so A'_N & A_N commute $\Longrightarrow A'_N - A_N$ is nilpotent
- $A_S = R(A)$, $A'_S = T(A)$ so they commute.

Fact Two diagonalizable matrices that commute can be diagonalized simultaneously
(HW 11) So $A_S - A'_S$ is semisimple & nilpotent $\Longrightarrow A_S - A'_S = 0. = A_N - A'_N$

Characteristic Polynomial

Fix V an n-dim'l $\mathbb{K}$-vector space & $A: V \longrightarrow V$ a k-linear map.

We have
$$\begin{aligned} \mathbb{K}[X] &\longrightarrow \mathbb{K}[A] \\ x &\longmapsto A \end{aligned}$$

<u>Def</u>: We define the <u>characteristic polynomial</u> of A as
$$\boxed{\chi_A} = \det(xI_n - A)$$

<u>Obs</u>: If $A \sim C$ (similar, ie $C = G^{-1}AG$ for $G \in GL_n(\mathbb{K})$) then, $\chi_C = \chi_A$.

Indeed:
$$\begin{aligned} \chi_C &= \det(xI_n - G^{-1}AG) \\ &= \det(G^{-1}(xI - A)G) = \det G^{-1} \det(xI_n - A) \det G \\ &= \chi_A \end{aligned}$$

Theorem (Cayley-Hamilton) $\chi_A(A) = 0$ (ie $q_A \mid \chi_A$)

Proof: We'll use the Rational Normal form of A.

$$[A]_{BB} = \begin{bmatrix} C_{q_1} & & 0 \\ & \ddots & \\ 0 & & C_{q_r} \end{bmatrix} \qquad q_1 \mid q_2 \mid \cdots \mid q_r = q_A.$$

$$\det(xI_n - A) = \det \begin{bmatrix} xI_n - C_{q_1} & & 0 \\ & \ddots & \\ 0 & & xI_n - C_{q_r} \end{bmatrix}$$

$$\underset{\text{block decomp}}{=} \chi_{C_{q_1}} \cdots\cdots \underbrace{\chi_{C_{q_r}}}_{= q_r}$$

$= q_r$ by Lemma below.

So $q_A \mid \chi_A$. $\square$

Lemma: $\chi_{C_f} = f$ for any monic polynomial $f \in \mathbb{K}[x]$.

Pf/ By induction on degree of f

- $\deg f = 1$: $\chi_{C_{x+a_0}} = \det(x + a_0) = x + a_0$ $\quad C_{x+a_0} = [-a_0]$

- $\deg f = m \rightsquigarrow f = x^m + a_{m-1}x^{m-1} + \cdots + a_0$

$$C_f = \begin{bmatrix} 0 & & & -a_0 \\ 1 & 0 & & -a_1 \\ & \ddots & \ddots & \vdots \\ & & 0 & \\ & & 1 & -a_{m-1} \end{bmatrix} \rightsquigarrow xI_m - C_f = \begin{bmatrix} x & 0 & \cdots & & a_0 \\ -1 & x & & & a_1 \\ & -1 & x & & a_2 \\ & & \ddots & \ddots & \vdots \\ & & & -1 & x + a_{m-1} \end{bmatrix}$$

$\det(xI_m - C_f)$ is computed by column expansion:

$$\chi_{C_f} = x \underbrace{\det \begin{bmatrix} x & & \cdots & a_1 \\ -1 & x & & a_2 \\ & \ddots & \ddots & \vdots \\ & & x & \\ & & -1 & x + a_{m-1} \end{bmatrix}}_{\substack{\| \\ \chi_{C_{\frac{f-a_0}{x}}} \underset{\mathrm{IH}}{=} \frac{f-a_0}{x}}} + 1 \underbrace{\det \begin{bmatrix} 0 & 0 & \cdots & & a_0 \\ -1 & x & & & a_2 \\ 0 & -1 & x & & \\ & & \ddots & x & \\ & & & -1 & x + a_{m-1} \end{bmatrix}}_{(-1)^{m-1+1} a_0 \det \begin{bmatrix} -1 & x & \cdots & 0 \\ 0 & -1 & x & \\ & & \ddots & x \\ & & & -1 \end{bmatrix} = (-1)^{m+m-2} a_0} = x\frac{(f - a_0)}{x} + a_0 = f$$

$\rightarrow$ size $= (m-2)^2$

. Alternative Proof of CH:

To show: $\underbrace{\chi_A(A)}_{\in \mathrm{End}_{\mathbb{K}}(\mathbb{K}^n)}(v) = 0 \quad \forall v$ in V. Assume $\dim V = n$.

Pick any $v \in V$ & consider $V' = K[A]\cdot v = \mathrm{Sp}(\{v, Av, A^2v, \dots\})$

- We know we can find d with $B' = \{v, Av, \dots, A^{d-1}v\}$ a basis for V'

- We extend B' to a basis B of V. Then we get

$$[A]_{BB} = \begin{array}{c|cc} & d & n-d \\ \hline d & A_1 & A_2 \\ n-d & 0 & A_3 \end{array} \qquad \text{where } A_1 = \begin{bmatrix} 0 & & 0 & -a_0 \\ 1 & 0 & & \vdots \\ & 1 & \ddots & \vdots \\ 0 & & 1 & -a_{d-1} \end{bmatrix} \quad \text{So } A_1 = C_{q_{A|V'}}.$$

$\Rightarrow$ Block decomp gives: $\chi_A = \chi_{A_1} \cdot \chi_{A_3} \underset{\text{Lemma}}{\uparrow =} q_{A|V'} \cdot \chi_{A_3} = \chi_{A_3}\, q_{A|V'}$

So $\chi_A(A)(v) = \left(\chi_{A_3}(A) \cdot q_{A|V'}(A)\right)(v)$

$$= \chi_{A_3}(A)\Big(\underbrace{q_{A|V'}(A)(v)}_{=0 \quad (v \in V')}\Big) = 0$$

$\square$