Lecture 32: General Jordan canonical forms - Cayley-Hamilton

<u>Recall</u>: Last time we talked about minimal polynomials of matrices $A \in Mat_{n\times n}(\mathbb{K})$

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

$$P(X) \longmapsto P(A)$$

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

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

<u>Obs</u>: $q_{G^{-1}AG} = q_A$ for any $G \in GL_n(\mathbb{K})$

- If $\mathbb{K}^n$ is cyclic ($\exists\, v \in \mathbb{K}^n$ with $\{v, Av, A^2v, \ldots\}$ generating $\mathbb{K}^n$), then $q_A = x^n + a_{n-1}x^{n-1} + \cdots + a_0$

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

$$[A]_{\mathcal{B}\mathcal{B}} = \begin{bmatrix} 0 & & & -a_0 \\ 1 & \ddots & 0 & -a_1 \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & -a_{d-1} \end{bmatrix} = C_{q_A} \text{ Companion matrix for polynomial } q_A.$$

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

(2) If $\mathbb{K} = \overline{\mathbb{K}}$, char $\mathbb{K} = 0$ and $q_A = (x-\alpha)^n$, then $\mathbb{K}^n$ is cyclic & $\mathcal{B} = \{v, (A-\alpha I)(v), \ldots, (A-\alpha I)^{n-1}(v)\}$ is a basis for some $v \in \mathbb{K}^n$

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

In the noncyclic case, 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

$B = B_1 \cup \cdots \cup B_r$ $\quad$ (B_i basis for V_i) with

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

with $q_1 \mid q_2 \mid \cdots \mid q_r$

$q_r = q_A$.

(Rational Normal Form)

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

$\leadsto$ $V = V_1' \oplus \cdots \oplus V_s'$ $\quad$ with $q_{A|_{V_i'}} = (x-\alpha_i)^{m_i}$

$\exists$ bases $B_1', \ldots, B_s'$ of $V_1', \ldots, V_s'$ with $B = \bigcup_{i=1}^{s} B_i'$

$$[A]_B = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_s \end{pmatrix}$$

Furthermore, $A_i = [A|_{V_i'}]_{B_i' B_i'}$

can be further decomposed as:

$$A_i = \begin{bmatrix} J(\alpha_i, m_1^{(i)}) & & 0 \\ & \ddots & \\ 0 & & J(\alpha_i, m_{s_i}^{(i)}) \end{bmatrix}$$

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

(Jordan canonical form)

Q: What else can we say?

§1 General Jordan form

Obs: Written additively, we get:

$$[A]_{\mathcal{B}\mathcal{B}} = D + N$$

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

There is a more general version of Jordan canonical forms, that work over perfect fields (char 0 is enough).

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*) $[A_S, A] = [A_N, A] = [A_S, A_N] = \mathbb{O}$

(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 $(q_{A_S}, q_{A_S}') = 1$.

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

Easy case: $q_A = (x-\lambda)^d$ for

$\rightsquigarrow A_N = (A - \lambda I_n)$ & $A_S = A - A_N$

- 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 $\mathbb{K} = \overline{\mathbb{K}}$, see HW11 (semisimple = diagonalizable)
- For perfect fields, you need Galois Thm.

- __Uniqueness:__ $A_s + A_N = A'_s + A'_N$

$$A_S - A'_S = A'_N - A_N$$

- $A'_N = P(A)$, $A_N = Q(A)$ 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__ (HW11): Two diagonalizable matrices that commute can be diagonalized simultaneously. So $A_S - A'_S$ is semisimple. Only one semisimple and nilpotent matrix $= 0$.

§2 __Characteristic polynomial:__

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

$$K[X] \longrightarrow K[A]$$
$$x \longmapsto A$$

__Def:__ We define the __characteristic polynomial__ of A as

$$\chi_A = \det(xI_n - A)$$

Obs: If $A \sim C$, meaning $C = G^{-1} A G$ for $G \in GL_n(\mathbb{K})$, then, $\chi_C = \chi_A$.

Indeed:
$$\chi_C = \det(x I_n - G^{-1}AG) = \det\big(G^{-1}(xI - A)G\big) = \det G^{-1} \det(xI_n - A) \det G = \chi_A$$

Lemma: If $\varphi : \mathbb{K} \longrightarrow \mathbb{K}'$ is a homomorphism of rings between 2 fields, then: $\chi_{\varphi(M)} = \varphi(\chi_M)$

(Here, φ extends to $\varphi : \mathbb{K}[x] \longrightarrow \mathbb{K}'[x]$)

Proof: Exercise.

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

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

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

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

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

- $\deg f = 1$: $\chi_{C_{x+a_0}} = \det(x+a_0) = x+a_0 \qquad C_{x+a_0} = [-a_0]$
- $\deg f = m \rightsquigarrow f = x^m + a_{m-1}x^{m-1} + \cdots + a_0$

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

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

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

$$= x\, \chi_{C_{\frac{f-a_0}{x}}} + (-1)^{2m-2} a_0$$

$$\underset{\text{IH}}{=} x \frac{f-a_0}{x} + a_0 = f(x) - a_0 + a_0 = f(x) \qquad \square$$

Alternative Proof of CH:

To show: $\chi_A(A)(v) = 0 \quad \forall v \text{ in } V.$

Pick any $v \in V$ & consider $V' = K[A]\cdot v$ that is, the vector space spanned by $\{v, Av, A^2v, \dots\}$.

- We know we can find d with $B' = \{v, Av, \dots, A^{d-1}v\}$ a basis for V'. Assume $\dim V = n$.
- We extend B' to a basis B of V. Then we get

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

So $A_1 = C_{q_{A|_{V'}}}$.

Then using the same block decomposition, we get.

$$\chi_A = \chi_{A_1} \cdot \chi_{A_3} \underset{\substack{\uparrow \\ \text{Lemma}}}{=} q_{A|_{V'}} \cdot \chi_{A_3} = \chi_{A_3}\, q_{A|_{V'}}$$

So
$$\chi_A(A)(v) = \left(\chi_{A_3}(A)\cdot q_{A|_{V'}}(A)\right)(v)$$
$$= \chi_{A_3}(A)\Big(\underbrace{q_{A|_{V'}}(A)(v)}_{=0\ (v\in V')}\Big) = 0$$

$\square$

Obs: The result is true for matrices over any commutative ring R. We can show this by proving the polynomial

$$\chi_A(x) = x^n + b_{n-1}x^{n-1} + \dots + b_0 \quad \text{where } b_i \in \mathbb{Z}[a_{ij}]$$

vanishes on $A = (a_{ij})$ inside a dense open set of $\mathrm{Mat}_{n\times n}(R)$.

We can pick diagonalizable matrices as such set.

- Alternative proof: HW 11 Problem 18.

$\square$