Lecture 25

The Minimal Polynomial of a Matrix in Jordan Normal Form

Definition

If $f_1(x), \ldots, f_k(x) \in F[x]$ then the least common multiple $\mathrm{lcm}(f_1, \ldots, f_k)$ is the lowest degree monic polynomial divisible by each of $f_1, f_2, \ldots, f_k$.

Remark In high school we learned that the least common multiple $\mathrm{lcm}(n_1, n_2, \ldots, n_k)$ of k whole numbers $n_1, n_2, \ldots, n_k$ was the smallest whole number divisible by each $n_1, n_2, \ldots, n_k$. The lcm is also the product over all primes p of p^e where p^e is the biggest power of p dividing one of the n_i's.

Problem 1 Find $\mathrm{lcm}(12, 100)$

Solution $12 = 2^2 \times 3$, $100 = 2^2 \times 5^2$

$\mathrm{lcm} = 2^2 \times 3 \times 5^2 = 300$

Problem 2 Find $\mathrm{lcm}(x^2, x^2+2x)$

$$x^2 = x^2, \quad x^2+2x = x(x+2)$$

$$\mathrm{lcm}(x^2, x^2+2x) = x^2(x+2)$$

Proposition 1 Suppose A is block diagonal

$$A = \left(\begin{array}{c|c|c|c|c} A_1 & 0 & 0 & 0 & 0 \\ \hline 0 & A_2 & 0 & 0 & 0 \\ \hline 0 & 0 & \cdot & 0 & 0 \\ \hline 0 & 0 & 0 & \cdot & \\ \hline 0 & 0 & 0 & 0 & A_m \end{array}\right)$$

Then

1. The characteristic polynomial of A is the product of the characteristic polynomials of the blocks.

$$h_A(x) = h_{A_1}(x)\, h_{A_2}(x) \cdots h_{A_m}(x)$$

2. The minimal polynomial of A is the lcm of the minimal polynomials of the blocks

$$m_A(x) = \operatorname{lcm}(m_{A_1}(x), m_{A_2}(x), \ldots, m_{A_m}(x))$$

Proof 1. is obvious

$$\det(xI-A) = \det\begin{pmatrix} xI-A_1 & 0 & 0 & 0 & 0 \\ 0 & xI-A_2 & 0 & 0 & 0 \\ 0 & 0 & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & 0 \\ 0 & 0 & 0 & 0 & xI-A_m \end{pmatrix}$$

and the determinant of a block diagonal matrix is the product of the determinants of the blocks

To prove 2 we need the following formula.
Let A be as on page 2 and $f(x) \in F[x]$
then

$$f(A) = \begin{pmatrix} f(A_1) & 0 & 0 & 0 & 0 \\ 0 & f(A_2) & 0 & 0 & 0 \\ 0 & 0 & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & 0 \\ 0 & 0 & 0 & 0 & f(A_m) \end{pmatrix}$$

To see this first check it for powers of A, A^l and then check it for linear combinations $c_1 A_1 + c_2 B_1$.

Hence, $f(A)=0 \iff f(A_1)=0, f(A_2)=0, \ldots, f(A_m)=0$

Since $f(A_1)=0$, $f(x)$ is divisible by $m_{A_1}(x)$.

Since $f(A_2)=0$, $f(x)$ is divisible by $m_{A_2}(x)$.

...

Finally since $f(A_m)=0$, $f(x)$ is divisible by $m_{A_m}(x)$.

Hence $f(x)$ is divisible by $g(x) = \mathrm{lcm}(m_1, m_2, \ldots, m_k)$

Since $f(A_i)=0 \iff f(x)$ is divisible by $m_{A_i}(x)$, $1 \le i \le m$ it follows from the definition of the lcm that $g(x)$ is the lowest degree polynomial satisfying $g(A_i)=0$, $1 \le i \le m$.

Now suppose A is in Jordan normal form so the blocks A_i above are Jordan blocks $J_{k(\lambda)}$

$$J(\lambda) = \overbrace{\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}}^{k(\lambda)} \Bigg\} k(\lambda)$$

Suppose $J(\lambda)$ is $k(\lambda)$ by $k(\lambda)$

Lemma

The minimal polynomial of $J(\lambda)$ is $(x-\lambda)^{k(\lambda)}$

Proof Clearly the characteristic polynomial of $J(\lambda)$ is $(x-\lambda)^{k(\lambda)}$. Since the minimal polynomial of any matrix divides the characteristic polynomial (by Cayley - Hamilton) the minimal polynomial $m_{J(\lambda)}(x)$ must be a power of $x-\lambda$. But in the Proposition on pg 9 of Lecture 24 we saw that

$$(A-\lambda I)^{k(\lambda)-1} v_{k(\lambda)} = v_1 = \text{the eigenvector associated the block}$$

Hence $(T-\lambda I)^{k(\lambda)-1} \neq 0$ so $m_J(\lambda) = (x-\lambda)^{k(\lambda)}$

$\square$

Now we can prove our main theorem

Theorem Let A be an n by n matrix. Suppose the characteristic polynomial $h_A(x)$ is given by

$$h_A(x) = (x-\lambda_1)^{e_1}(x-\lambda_2)^{e_2}\cdots(x-\lambda_\ell)^{e_\ell}.$$

Let $k(\lambda_j)$, $1 \le j \le \ell$, be the size of the biggest Jordan λ_j-block. Then

$$m_A(x) = (x-\lambda_1)^{k(\lambda_1)}(x-\lambda_2)^{k(\lambda_2)}\cdots(x-\lambda_\ell)^{k(\lambda_\ell)}$$

Proof

Since $(x-\lambda_i)$ and $(x-\lambda_j)$ are relatively prime $\operatorname{lcm}((x-\lambda_i)^a, (x-\lambda_j)^b) = (x-\lambda_i)^a(x-\lambda_j)^b$

So if $m_{\lambda_j}(x)$ is the minimal polynomial associated to all the Jordan λ_j-blocks we have

$$m_A(x) = m_{\lambda_1}(x)\, m_{\lambda_2}(x) \cdots m_{\lambda_\ell}(x)$$

But the lcm of a finite collection of powers of $x-\lambda$ is $(x-\lambda)^{\text{biggest power}}$

so

$$m_{\lambda_j}(x) = (x-\lambda_j)^{k(\lambda_j)}$$

$\square$

Example

$$A = \left(\begin{array}{cc|cc|ccc} 3 & 1 & & & & & \\ 0 & 3 & & & & & \\ \hline & & 3 & 1 & & & \\ & & 0 & 3 & & & \\ \hline & & & & 5 & 1 & 0 \\ & & & & 0 & 5 & 1 \\ & & & & 0 & 0 & 5 \end{array}\right)$$

$\lambda(3) = 2$, $k(5) = 3$ so

$$m_A(x) = (x-3)^2 (x-5)^3$$

Remark We will often compute $m_A(x)$. Then we will know the sizes of the biggest Jordan λ-blocks