Leture 5: Jordan Canonical Form

Matrix Analysis

The defective matrices are not diagonalizable, For example

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

But these defective matrices can be similar to some upper triangular matrices whose form is very close to a diagonal matrix. This form is Jordan canonical form (or Jordan normal form).

Jordan canonical form is very useful to understand the matrix structures and matrix functions.

Before introduce Jordan canonical form, we need to understand the concepts of *minimal polynomials* and *invariant subspaces*.

5.1 Minimal Polynomials

Theorem 5.1.1. *Let $A \in M_n$. Then there exists a unique monic annihilate polynomial $q_A(x)$ of minimum degree. If $p(x)$ is any annihilate polynomial, then $q_A(x)$ divides $p(x)$. [remarks: if $p(A)=0$, then $p(x)$ is called an annihilate polynomial of A. "monic" means the highest order coefficient of a polynomial is "1"]*

Proof. For matrix A, the characteristic polynomial $p_A(x)$ is an annihilate polynomial, that is $p_A(A) = 0$, assume $q_A(x)$ is a minimal degree annihilate polynomial which is monic, then $q_A(A)=0$. by the Euclidean algorithm

$$p_A(x) = q_A(x)h(x) + r(x)$$

where $\deg r(x) < \deg q_A(x)$. We know

$$p_A(A) = q_A(A)h(A) + r(A).$$

Hence $r(A) = 0$, and by the minimality assumption $r(x) \equiv 0$. Thus $q_A(x)$ divides $p_A(x)$ and also any polynomial for which $p(A) = 0$.

To establish that $q_A(x)$ is unique, suppose $q(x)$ is another monic polynomial of the same degree for which $q(A) = 0$. Then

$$r(x) = q(x) - q_A(x)$$

is a polynomial of degree <u>less than</u> $q_A(x)$ for which $r(A) = q(A) - q_A(A) = 0$. This cannot be true. □

Definition 5.1.1. The polynomial $q_A(x)$ in the theorem above is called the *minimal polynomial*.

Corollary 5.1.1. *If $A, B \in M_n$ are similar, then they have the same minimal polynomial.*

Proof. let $B = S^{-1}AS$, then

$$q_A(B) = q_A(S^{-1}AS) = S^{-1}q_A(A)S = 0$$

If there is a minimal polynomial for B of smaller degree, say $q_B(x)$, then $q_B(A) = 0$ by the same argument. This contradicts the minimality of $q_A(x)$.

□

Corollary 5.1.2. *For the minimal polynomial $q_A(x)$,* $q_A(\lambda)=0$ *iff* λ *is the eigenvalue of* A.

Proof. $q_A(\lambda)=0 \Rightarrow p_A(\lambda)=q_A(\lambda)h(\lambda)=0$

$\Rightarrow \lambda$ is the eigenvalue.

λ is the eigenvalue. $\Rightarrow Ax=\lambda x (x\neq 0)$

$\Rightarrow 0=q_A(A)x=q_A(\lambda)x$

$\Rightarrow q_A(\lambda)=0$

□

From **Corollary 5.1.2**, if $p_A(t)=\prod_{i=1}^{k}(t-\lambda_i)^{m_i}$ then the minimal polynomial $q_A(t)$ has the form

$$q_A(t)=\prod_{i=1}^{k}(t-\lambda_i)^{r_i},\ 1\le r_i\le m_i.$$

5.2 Invariant subspaces

We have considered the subspaces V of C^n that are invariant under the matrix $A \in M_n(C)$. This means that $AV \subset V$.

We now consider a specific type of invariant subspace that will lead to the so-called Jordan canonical form.

Definition 5.2.1. Let $A \in M_n(C)$ with spectrum $\sigma(A) = \{\lambda_1, \ldots, \lambda_k\}$. the *generalized eigenspace* pertaining to λ_i is

$$V_{\lambda_i} = \{x \in \mathbb{C}_n \mid (A - \lambda_i I)^n x = 0\}$$

If the span of the eigenvectors (eigenspace) pertaining to λ_i is not equal to V_{λ_i} then, there must be a positive power p and a vector x such that $(A - \lambda_i I)^p x = 0$ but that $y = (A - \lambda_i I)^{p-1} x \neq 0$. Thus y is an eigenvector pertaining to λ_i. We say x is a *generalized eigenvector of order p*.

For this reason we will call $V_{\lambda i}$ the space of *generalized eigenvectors* pertaining to λ_i.

Theorem 5.2.1. *Generalized eigenspace* $V_{\lambda i}$ *is an invariant subspace of A.*

Proof. For $\forall x \in V_{\lambda_i}$

$$(A - \lambda_i I)^n x = 0$$

$$\Rightarrow (A - \lambda_i I)^n Ax = A(A - \lambda_i I)^n x = 0$$

$$\Rightarrow Ax \in V_{\lambda_i}$$

□

5.3 The Jordan canonical Form

Definition 5.3.1. Let λ ∈ C. A *Jordan block* $J_k(\lambda)$ is a $k \times k$ upper triangular matrix of the form

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & \\ & & & 0 \\ & \lambda & 1 & \\ 0 & & \ddots & 1 \\ & & & \lambda \end{bmatrix}.$$

A *Jordan matrix* is any matrix of the form

$$J = \begin{bmatrix} J_{n_1}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & J_{n_k}(\lambda_k) \end{bmatrix}.$$

where the matrices J_{ni} are Jordan blocks.

Note that

$$J_k(0) = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} = (J_k(\lambda) - \lambda I)$$

And it is *nilpotent.*

$$J_k^2(0) = \begin{bmatrix} 0 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & 0 \\ & & & & 0 \end{bmatrix}, \ldots \ldots, \; J_k^k(0) = 0$$

Theorem 5.31

If $A \in M_n(C)$ has only one eigenvalue λ, and $geom(\lambda) = 1$. Then there exists $x \in C_n$, such that $\{(A-\lambda)^{n-1}x, \ldots, (A-\lambda)x, x\}$ is a basis of V_λ and $A \sim J$, where

$$J = \begin{pmatrix} \lambda & 1 & \ldots & \ldots \\ 0 & \lambda & 1 & \ldots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \lambda \end{pmatrix}$$

Proof:

$(A - \lambda I)y = 0$

$$\Rightarrow U^* \begin{pmatrix} 0 & * & \ldots & * \\ 0 & 0 & * & * \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix} Uy = U^*TUy = 0,$$ since $geom(\lambda) = 1$, so

$rank(T) = n - 1$. Let $z = Uy$, set $z = \begin{pmatrix} 1 \\ 0 \\ \ldots \\ 0 \end{pmatrix}$, then solve $y = U^*z$

For equation $(A-\lambda I)^{n-1}x = y = U^* z$,

$$\Rightarrow U^* \begin{pmatrix} 0 & * & \dots & * \\ 0 & 0 & * & \dots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}^{n-1} Ux = U^* z$$

$$\Rightarrow \begin{pmatrix} 0 & \dots & 0 & * \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} Ux = \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}$$

Let $v = Ux$, then $v = \begin{pmatrix} 0 \\ \dots \\ 0 \\ 1/* \end{pmatrix}$, and $x = U^* v$. It's straightforward that $(A-\lambda I)^k x \neq 0$, for $k < n$.

If $\{(A-\lambda I)^{n-1}x, \ldots, (A-\lambda I)x, x\}$ is not linear independent. There exists $\alpha_1, \ldots, \alpha_n \in C$ not all zeros, such that:

$\alpha_1(A-\lambda I)^{n-1}x + \cdots + \alpha_n x = 0 \qquad (3.1)$

Multiply by $(A-\lambda I)^{n-1}$ on both sides, we got

$\alpha_1(A-\lambda I)^{2n-2} + \cdots + \alpha_n(A-\lambda I)^{n-1}x = 0$

Since $(A-\lambda I)^n = P_A(\lambda) = 0$, the above equation becomes:

$\alpha_n(A-\lambda I)^{n-1}x = 0$, however, we just proved that

$(A-\lambda I)^{n-1}x = y \neq 0$. So $\alpha_n = 0$.

Then multiply $(A-\lambda I)^{n-2}$ to equation (3.1), we get $\alpha_{n-1} = 0$, and so forth.

Eventually $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, contradiction to linear dependence.

Therefore the set $\{(A-\lambda I)^{n-1}x, \ldots, (A-\lambda I)x, x\}$ is linearly independent, and there are n vectors in the set, thus it is a basis.

Let $P = ((A-\lambda I)^{n-1}x, \dots, (A-\lambda I)x, x)$,
$AP = (A(A-\lambda I)^{n-1}x, \dots, A(A-\lambda I)x, Ax)$
$\Rightarrow = ((A-\lambda I+\lambda)(A-\lambda I)^{n-1}x, \dots, (A-\lambda I+\lambda)(A-\lambda I)x, (A-\lambda I)x+\lambda x)$
$\Rightarrow$
$((A-\lambda I)^n x+\lambda(A-\lambda I)^{n-1}x, \dots, (A-\lambda I)^2+\lambda(A-\lambda I)x, (A-\lambda I)x+\lambda x)$

$$= ((A-\lambda I)^{n-1}x, \dots, (A-\lambda I)x, x) \begin{pmatrix} \lambda & 1 & \dots & \dots \\ 0 & \lambda & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda \end{pmatrix} = PJ$$

That is equivalent to $A = PJP^{-1}$