Northeastern University, Department of Mathematics

MATH G5110: Applied Linear Algebra and Matrix Analysis.

• Instructor: **He Wang** Email: **he.wang@northeastern.edu**

§8 Jordan Canonical Form

Contents

$\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$

$n\times n$ matrix $A = PBP^{-1}$

$\Leftrightarrow AP = PB$

$\Leftrightarrow [AP_1\ AP_2 \cdots AP_m] = [\vec{v}_1 \cdots \vec{v}_m\ \vec{v}_{m+1} \cdots \vec{v}_{2?} \cdots] \begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_m \end{bmatrix}$

$= [P_1B_1 \cdots P_mB_m]$

1. Block diagonal

An $n \times n$ matrix B is a **block diagonal matrix** if

$\Leftrightarrow AP_i = P_iB_i$

$$B = \begin{bmatrix} B_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & B_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & B_m \end{bmatrix}$$

Denote $V_i := \text{Span}\{P_i\}$

$\Leftrightarrow T_A(V_i) \subset V_i$

$A = PBP^{-1}$

Theorem 1. *An $n \times n$ matrix A is similar to a block diagonal matrix B if and only if there exists a decomposition of*

$$\mathbb{F}^n = V_1 \oplus V_2 \oplus \cdots \oplus V_m$$

such that V_i is invariant under T_A.

The following non-diagonalizable matrices are called **Jordan blocks** of size 1, 2, 3, 4, ...

$$J_{\lambda,1} = \begin{bmatrix} \lambda \end{bmatrix}, \quad J_{\lambda,2} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}, \quad J_{\lambda,3} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}, \quad J_{\lambda,4} = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}, \ldots$$

Definition 2. An $n \times n$ **Jordan normal matrix** (**Jordan normal form**) is a block diagonal matrix

$$J = \begin{bmatrix} \boxed{J_{n_1}(\lambda_1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boxed{J_{n_2}(\lambda_2)} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boxed{J_{n_m}(\lambda_m)} \end{bmatrix}$$

such that all diagonal matrices $J_{n_i}(\lambda_i)$ are of the form

J_{λ_i, s_j}

$s_j \geq s_{j+1}$

$$J_{n_i}(\lambda_i) = \begin{bmatrix} \lambda_i & * & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \lambda_i & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & * \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \lambda_i \end{bmatrix}$$

Ex: $\begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ & 2 & 1 & 0 & 0 & 0 \\ & & 2 & 0 & 0 & 0 \\ & & & 2 & 1 & 0 \\ & & & & 2 & 0 \\ & & & & & 2 \end{bmatrix}$

where $* = 1$ or 0.

$\mathbb{F} = \mathbb{R}, \mathbb{C}$

Theorem 3. *Every $n \times n$ matrix A with n eigenvalues in a field $\mathbb{F}$ is similar to a matrix J in Jordan normal matrix, that is $A = PJP^{-1}$.*

To be proved.

$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

2. Nilpotent matrix

Definition 4. An $n \times n$ matrix A is called **nilpotent of degree** m if

$A^m = 0$ and $A^{m-1} \neq 0$ for some $m > 0$.

Proposition 5. *If A is nilpotent, then zero is the only eigenvalue of A.*

$\Leftrightarrow$

$A\vec{v} = \lambda \vec{v}$ $\vec{v} \neq \vec{0}$

$\vec{0} = A^m \vec{v} = \lambda^m \vec{v} \Rightarrow \lambda^m = 0 \Rightarrow \lambda = 0$

$J_n(0)$ is nilpotent!

$(J_{0,k})^{k-1} = \begin{bmatrix} & & 1 \\ & 0 & \\ & & \end{bmatrix}$

$$J_{0,k}\vec{x} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{k-1} \\ x_k \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_k \\ 0 \end{bmatrix}$$

$J_{0,k}^{k-1} \vec{x} = \begin{bmatrix} x_k \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

$(J_{0,k})^k = 0$

Lemma 6.
- *$J_{0,k}$ is nilpotent of degree k.*
- *Suppose a Jordan matrix $J = J_n(\lambda)$ with the same entry λ on diagonal, then there exist a number m such that $(J - \lambda I_n)^m = \mathbf{0}$.*

$\parallel$

the size of the largest Jordan block in $J_n(\lambda)$

$$J_{\lambda,k} - \lambda I = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} = J_{0,k}$$

Suppose $\boxed{A \text{ is similar to a Jordan block } J_{\lambda,n}.}$ $A = PJ_{\lambda,n}P^{-1}$

$\Leftrightarrow AP = PJ_{\lambda n}$ $\qquad P = [\vec{b}_1 \cdots \vec{b}_n]$

$$\Leftrightarrow [A\vec{b}_1 \cdots A\vec{b}_n] = [\vec{b}_1 \cdots \vec{b}_n]\begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} = [\lambda\vec{b}_1, \vec{b}_1 + \lambda\vec{b}_2, \cdots, \vec{b}_{n-1} + \lambda\vec{b}_n]$$

$$\begin{cases} A\vec{b}_1 = \lambda\vec{b}_1 \\ A\vec{b}_2 = \vec{b}_1 + \lambda\vec{b}_2 \\ \vdots \\ A\vec{b}_n = \vec{b}_{n-1} + \lambda\vec{b}_n \end{cases} \Leftrightarrow \begin{cases} (A-\lambda I)\vec{b}_1 = \vec{0} \\ (A-\lambda I)\vec{b}_2 = \vec{b}_1 \\ \vdots \\ (A-\lambda I)\vec{b}_n = \vec{b}_{n-1} \end{cases} \Rightarrow \begin{cases} \vec{b}_1 = (A-\lambda I)^{n-1}\vec{b}_n \\ \vec{b}_2 = (A-\lambda I)^{n-2}\vec{b}_n \\ \vdots \\ \vec{b}_{n-2} = (A-\lambda I)^2\vec{b}_n \\ \vec{b}_{n-1} = (A-\lambda I)\vec{b}_n \end{cases}$$

- Denote $N = A - \lambda I$
- $\{\underbrace{N^{n-1}\vec{b}_n}_{b_1}, \underbrace{N^{n-2}\vec{b}_n}_{b_2}, \cdots, \underbrace{N\vec{b}_n}_{b_{n-1}}, \vec{b}_n\}$ is called the Jordan chain.
- To find $P = [\vec{b}_1 \cdots \vec{b}_n]$ $\Leftrightarrow$ to find $\vec{b}_n$ such that $\vec{b}_n \in \ker N^n$ and $\vec{b}_n \notin \ker N^{n-1}$

$(A-\lambda I)\vec{v}=\vec{0}$

Definition 7. A non-zero vector $\vec{v}$ is called a **generalized eigenvector** of A if

$$(A-\lambda I)^k\vec{v}=\vec{0}$$

for some $k \geq 1$.

Remark:

(1) Any eigenvector is a generalized vector. $k=1$

(2) A generalized vector can exist only for the regular eigenvalue λ.

$\det(A-\lambda I)=0 \iff \boxed{\det (A-\lambda I)^k}=0$

$\qquad \| $

$\qquad (\det(A-\lambda I))^k$

(3) Let V_λ be the set of all generalized eigenvectors together with $\vec{0}$. Then V_λ is a subspace of $\mathbb{F}^n$.

(4) A Jordan chain is independent if and only if $\vec{b}_n \neq \vec{0}$

Ex.

(5) A is similar to a Jordan block $J_{\lambda,n}$ if and only if

$A=PJ_{\lambda,n}P^{-1}$ s.t. $P=[\vec{b}_1 \cdots \vec{b}_n] \iff$ to find $\vec{b}_n$ such that $\vec{b}_n \in \ker N^n$ and $\vec{b}_n \notin \ker N^{n-1}$

$\iff \exists$ Jordan chain.

We need to find the structure of a nilpotent matrix. We want to show that any nilpotent matrix is similar to $J_n(0)$. For example,

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \ldots$$

$\{3, 2, 1\}$ $\quad (2, 2, 1, 1) \quad (4, 2)$

That is $J_3 \oplus J_2 \oplus J_1$, or $J_2 \oplus J_2 \oplus J_1 \oplus J_1$, or $J_4 \oplus J_2, \ldots$

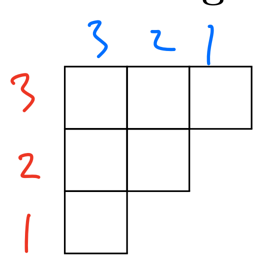
Proposition 8. *There is a one-to-one corresponding between $J_n(0)$ and* ***partition*** *of n, $(n_1, n_2, \ldots, n_k)$ such that*

$$n = n_1 + n_2 + \cdots + n_k \text{ and } n_1 \geq n_2 \geq \cdots \geq n_k \geq 1$$

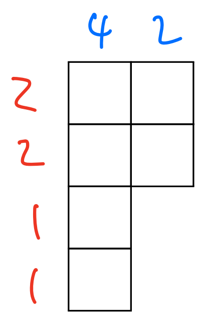
dual.

In the above examples, the partition of 6 are $(3, 2, 1)$, or $(2, 2, 1, 1)$, or $(4, 2)$...

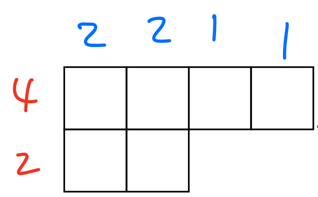
We can use the **Young diagram** to describe the partitions.



or



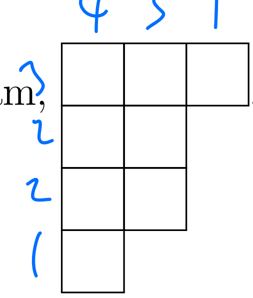
or



...

For another example, the Jordan matrix corresponds to partition $(n_1, n_2, n_3) = (3, 2, 2, 1)$ is

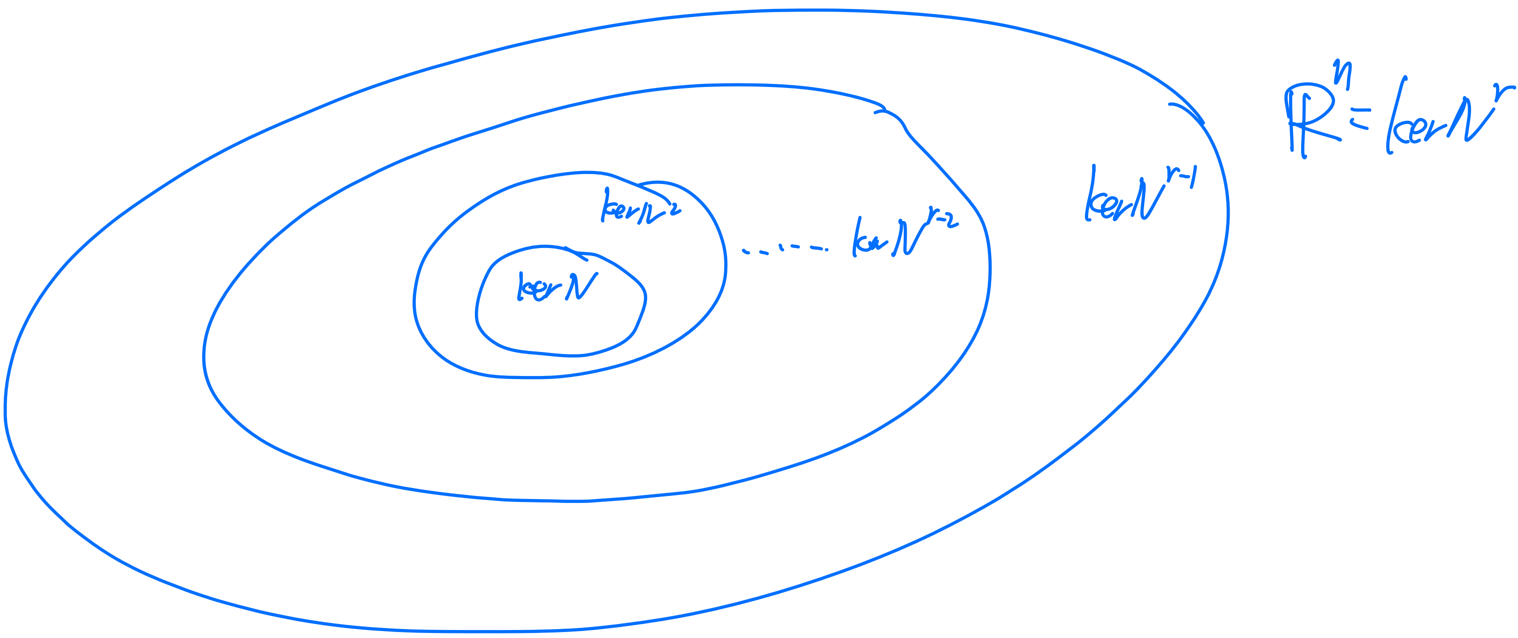
$$B = \begin{bmatrix} 0 & 1 & 0 & & & & & \\ 0 & 0 & 1 & & & & & \\ 0 & 0 & 0 & & & & & \\ & & & 0 & 1 & & & \\ & & & 0 & 0 & & & \\ & & & & & 0 & 1 & \\ & & & & & 0 & 0 & \\ & & & & & & & 0 \end{bmatrix}$$

By Young diagram , the dual partition is $(s_1, s_2, s_3) = (4, 3, 1)$.

$N^r = 0$ Goal $N = P J_n(0) P^{-1}$

Lemma 9. *Let N be an $n \times n$ nilpotent of degree r. Then we have strict inclusions*

$$\ker N \subset \ker N^2 \subset \cdots \subset \ker N^r = \mathbb{F}^n$$



Goal $N = P J_n(0) P^{-1}$

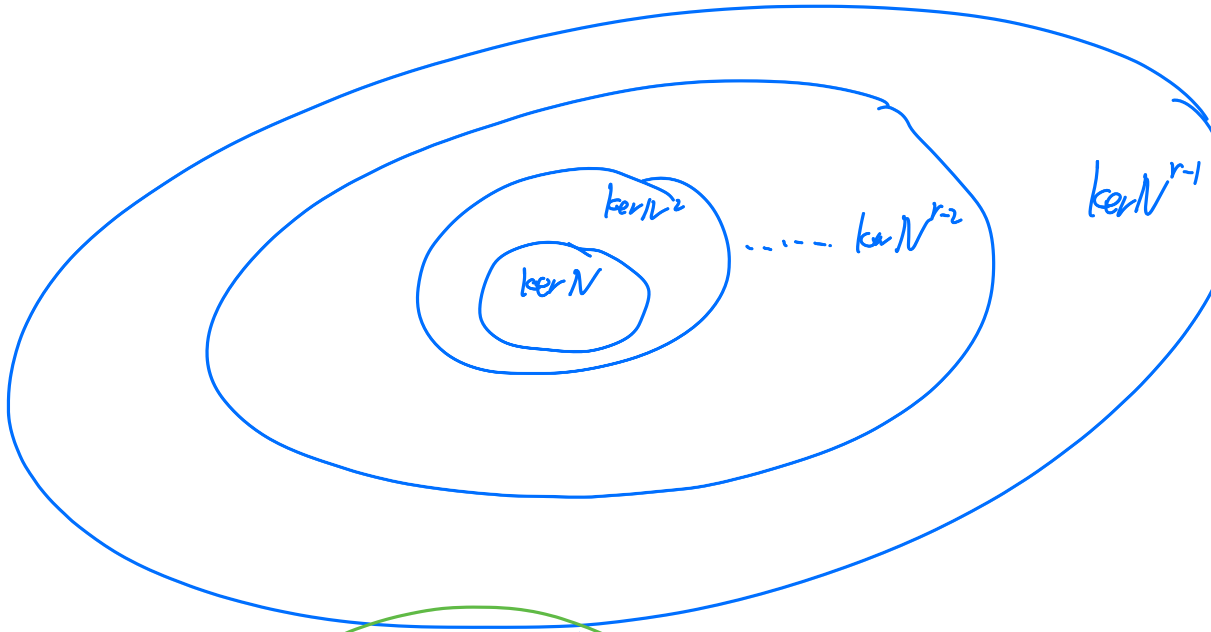
P

Theorem 10. *Let N be an $n \times n$ nilpotent matrix of degree r. Then there exist vectors $\vec{v}_1, \ldots, \vec{v}_s$ and integers $n_1, \ldots, n_s$ with $1 \leq n_s \leq \cdots \leq n_1 = r$ such that $N^{n_i - 1}\vec{v}_i \neq \vec{0}$ and $N^{n_i}\vec{v}_i = \vec{0}$ for all $i = 1, 2, \ldots, s$ and vectors*

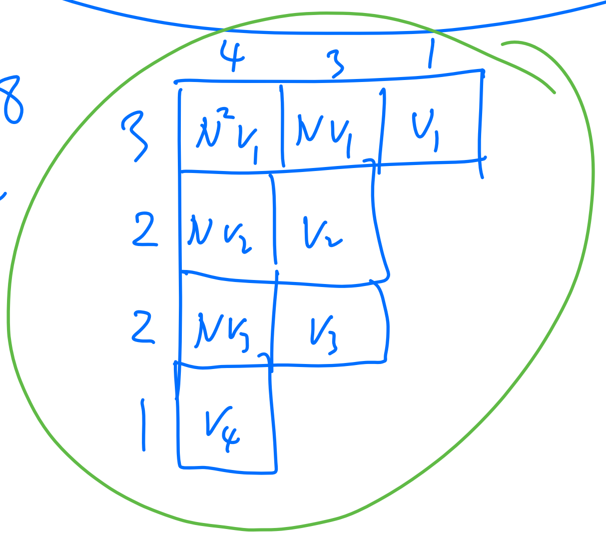
$$\begin{gathered} N^{n_1-1}\vec{v}_1, \ldots, \ldots, \ldots N\vec{v}_1,\ \vec{v}_1, \\ N^{n_2-1}\vec{v}_2 \ldots, \ldots, N\vec{v}_2,\ \vec{v}_2, \\ \vdots \\ N^{n_s-1}\vec{v}_s \ldots, N\vec{v}_s,\ \vec{v}_s \end{gathered}$$

form a basis for $\mathbb{F}^n$.

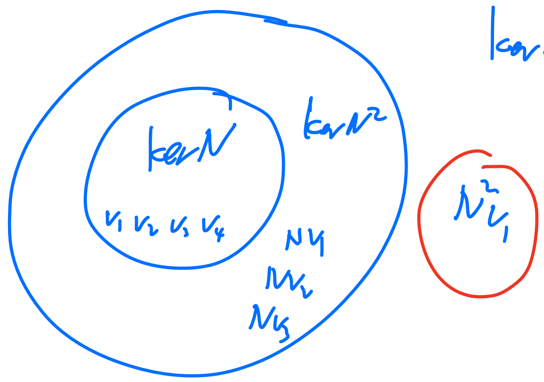
$\mathbb{R}^n = \ker N^r$



$\mathbb{R}^8$



$\ker N^3 = \mathbb{R}^8$



$$\mathbb{R}^8 = W_1 \oplus W_2 \oplus W_3$$

Step 1

$$\ker N \subset \ker N^2 \subset \ker N^3 = \mathbb{R}^8$$

dim: $m_1 = 4 < m_2 = 7 < m_3 \overset{8}{=} n$

$m_1 = s_1$, $m_2 = m_1 + s_2$, $m_3 = m_2 + s_3$

Step 2 $\quad W_1 \oplus W_2 \oplus W_3$

dim: $s_1 = 4$, $s_2 = 3$, $s_3 = 1$

dual of $(s_1\ s_2\ s_3)$

<u>Want:</u> $(n_1\ n_2\ n_3\ n_4) = (3, 2, 2, 1)$

Proof. By Lemma **??**, there are strict inclusions

$$\ker N \subset \ker N^2 \subset \cdots \subset \ker N^r = \mathbb{F}^n$$

Hence, there exist direct decompositions

$$\ker N^i = \ker N^{i-1} \oplus W_i$$

Hence

$$\mathbb{F}^n = W_r \oplus W_{r-1} \oplus \cdots \oplus W_2 \oplus W_1$$

where $W_1 = \ker N$.
Denote the dimension of each null space as $m_i = \dim \ker N^i$ for $i = 1, 2, ..., r$.
Then denote $\dim W_i = s_i$ where $s_1 = m_1$, $s_2 = m_2 - m_1$, $s_3 = m_3 - m_2$,..., $s_r = m_r - m_{r-1}$.
Choose a basis $\{\vec{w}_{r,1}, .., \vec{w}_{r,s_r}\}$ for W_r.
Extend $\{N\vec{w}_{r,1}, .., N\vec{w}_{r,s_r}\}$ to be a basis for W_{r-1} by adding $\{\vec{w}_{r-1,1}, .., \vec{w}_{r-1,s_{r-1}-s_r}\}$.
Keep extending until to W_1, we extended

$$\{N^{r-1}\vec{w}_{r,1}, .., N^{r-1}\vec{w}_{r,s_r}, N^{r-2}\vec{w}_{r,1}, .., N^{r-2}\vec{w}_{r-2,s_r}, ...\}$$

to be a basis for W_1 by adding $\{\vec{w}_{1,1}, .., \vec{w}_{1,s_1-s_2}\}$
Claim, the set

$$\begin{array}{lll}
N^{r-1}\vec{w}_{r1} \dots & N\vec{w}_{r,1}, & \vec{w}_{r,1}, \\
\vdots & & \\
N^{r-1}\vec{w}_{r,s_r} \dots & N\vec{w}_{r,s_r} & \vec{w}_{r,s_r} \\
N^{r-2}\vec{w}_{r-1,1}, \dots & \vec{w}_{r-1,1}, & \\
\vdots & & \\
N^{r-2}\vec{w}_{r-1,,s_{r-1}-s_r} \dots & \vec{w}_{r-1,s_{r-1}-s_r} & \\
\quad\vdots & & \\
\vec{w}_{1,1}, & & \\
\vdots & & \\
\vec{w}_{1,s_1-s_2} & &
\end{array}$$

is a basis for $\mathbb{F}^n$. $\square$

Corollary 11. *N is nilpotent if and only if N is similar to a Jordan canonical matrix $J_n(0)$.*

Corollary 12. *Let N be an $n \times n$ nilpotent matrix. Then $\lambda I + N$ is similar to a Jordan canonical matrix $J_n(\lambda)$.*

3. **Jordan Canonical Form**

Let A be an $n \times n$ matrix.

Theorem 13. *If* $\ker A \cap \operatorname{im} A = \{0\}$, *then* $\mathbb{F}^n = \ker A \oplus \operatorname{im} A$.

Remark: (1) The assumption is needed.

(2) Both $\ker A$ are $\operatorname{im} A$ invariant under T_A.

Theorem 14. *Let A be an $n \times n$ matrix with an eigenvalue λ. Denote the set of all generalized eigenvectors of A corresponding to λ, together with $\{\vec{0}\}$ by V_λ. Then, there exists m such that*

$$V_\lambda = \ker(A - \lambda I)^m$$

and

$$\mathbb{F}^n = \ker(A - \lambda I)^m \oplus \operatorname{im}(A - \lambda I)^m.$$

Both $\ker(A - \lambda I)^m$ *and* $\operatorname{im}(A - \lambda I)^m$ *are invariant under* T_A.

Theorem 15. *Let A be an $n \times n$ matrix with n eigenvalues. The distinct eigenvalues are $\lambda_1, \ldots, \lambda_k$. Then, there exist numbers $m_1, m_2, \ldots, m_k$ such that*

$$\mathbb{F}^n = \ker(A - \lambda_1 I)^{m_1} \oplus \cdots \oplus \ker(A - \lambda_k I)^{m_k}$$

and each $\ker(A - \lambda_i I)^{m_i}$ *is invariant under* T_A.

Theorem 16 (Block Diagonalization)**.** *Every $n \times n$ matrix A with n eigenvalues in a field $\mathbb{F}$ is similar to a block diagonal matrix, where each block has a single eigenvalue.*
More precisely, suppose $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the distinct eigenvalues of A. Then there is an invertible matrix $P \in \mathbb{F}^{n\times n}$ such that

$$P^{-1}AP = \begin{bmatrix} B_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & B_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & B_k \end{bmatrix}$$

where the matrix $B_i - \lambda_i I$ is nilpotent for $i = 1, 2, \ldots, k$.

Theorem 17. *Every $n \times n$ matrix A with n eigenvalues in a field $\mathbb{F}$ is similar to a matrix J in Jordan normal matrix, that is $A = PJP^{-1}$.*

Algorithm and example

Let A be an $n \times n$ matrix with distinct real eigenvalues $\lambda_1, \ldots, \lambda_p$ such that

$$f_A(\lambda) = \det(A - \lambda I) = (\lambda_1 - \lambda)^{k_1}(\lambda_2 - \lambda)^{k_2} \cdots (\lambda_p - \lambda)^{k_p}.$$

Suppose $k_1 + k_2 + \cdots + k_p = n$. (This is always true if $\mathbb{F}$ is algebraic closed, e.g., when $\mathbb{F} = \mathbb{C}$).

Algorithm of computing Jordan Normal form of a matrix:

Step 1. Find all eigenvalues λ_i and their algebraic multiplicity $am(\lambda_i) = k_i$.

Step 2. For each eigenvalue λ_i, calculate $m_j = \dim\ker(A - \lambda_i I)^j$ for $j = 1, 2, ...$ until $\dim\ker(A - \lambda_i I)^s = k_i$.

Step 3. From $m_1, ..., m_s$ we can calculate $s_j = m_j - m_{j-1}$, then use Young diagram calculate $n_1, ..., n_t$. Now we have determined the Jordan normal form J.

Step 4. To calculate the matrix P such that $A = PJP^{-1}$, we calculate $\mathbf{rref}(A - \lambda I)^j$ for each $\lambda = \lambda_i$.

Step 5. Find vectors $\{\vec{w}_{r,1}, ...\vec{w}_{r,s_r}\}$,..., $\{\vec{w}_{1,1}, ...\vec{w}_{1,s_1-s_2}\}$ such that

Example 18.

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & -8 & 4 & -3 & 1 & -3 \\ -3 & 13 & -8 & 6 & 2 & 9 \\ -2 & 14 & -7 & 4 & 2 & 10 \\ 1 & -18 & 11 & -11 & 2 & -6 \\ -1 & 19 & -11 & 10 & -2 & 7 \end{bmatrix}$$

(1) Find the Jordan normal form J of A.

Step 1, calculate all eigenvalues of A, which are $\lambda = 2$ with algebraic multiplicity 1 and $\lambda = -1$ with algebraic multiplicity 5. We know that the Jordan form looks like:

$$J = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & * & 0 & 0 & 0 \\ 0 & 0 & -1 & * & 0 & 0 \\ 0 & 0 & 0 & -1 & * & 0 \\ 0 & 0 & 0 & 0 & -1 & * \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Step 2, Calculate $m_i = \dim\ker((A+I)^i)$ we have $m_1 = 2, m_2 = 4, m_3 = 5$ which is the algebraic multiplicity $am(-1)$.

So, $s_1 = 2, s_2 = 2, s_3 = 1$ and by Young diagram

$B^2\vec{v}_1$	$B\vec{v}_1$	$\vec{v}_1$
$B\vec{v}_2$	$\vec{v}_2$	

.

$n_1 = 3, n_2 = 2.$

$$\text{So, } J = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

(2) Find matrix matrix P such that $A = PJP^{-1}$.

Step 1.

$$\mathbf{rref}(A+I) = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 1 & \frac{3}{2} \\ 0 & 0 & 1 & 0 & 2 & \frac{3}{2} \\ 0 & 0 & 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{rref}(A+I)^2 = \begin{bmatrix} 1 & 0 & -\frac{1}{2} & \frac{3}{2} & -2 & -\frac{5}{2} \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{rref}(A+I)^3 = \begin{bmatrix} 0 & 1 & -\frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\vec{v}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$ is the vector in $\ker(A+I)^3$ but not in $\ker(A+I)^2$

Calculate $(A+I)\vec{v}_1 = \begin{bmatrix} 1 & 0 & -3 & -2 & 1 & -1 \end{bmatrix}^T$ and $(A+I)^2\vec{v}_1 = \begin{bmatrix} 1 & -2 & -1 & 1 & -1 & 2 \end{bmatrix}^T$

$\vec{v}_2 = \begin{bmatrix} 0 & 1 & -2 & -2 & 3 & -3 \end{bmatrix}^T$ is the vector in $\ker(A+I)^2$ but not in $\ker(A+I)$ and not dependent on $\vec{v}_1$, $(A+I)\vec{v}_1$ and $(A+I)^2\vec{v}_1$

Step 2.

$$\mathbf{rref}(A+2I) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & 0 & 0 & -\frac{2}{3} \\ 0 & 0 & 0 & 1 & 0 & -\frac{2}{3} \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$ A basis for $\ker(A+2I)$ is $\begin{bmatrix} 0 & 1 & -2 & -2 & 3 & -3 \end{bmatrix}^T$

Step 3.

Hence matrix P is $P = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & -2 & 0 & 0 & 1 & 1 \\ -2 & -1 & -3 & 0 & -4 & 2 \\ -2 & 1 & -2 & 0 & -2 & 0 \\ 3 & -1 & 1 & 0 & 5 & 0 \\ -3 & 2 & -1 & 0 & -4 & 0 \end{bmatrix}$

Using Matlab directly A=sym(A) and [P, J] = jordan(A) will give us the result

$$J = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \quad P = \begin{bmatrix} 0 & -\frac{9}{2} & -7 & -7 & \frac{3}{2} & \frac{5}{2} \\ -1 & 9 & 3 & 1 & 0 & 0 \\ 2 & \frac{9}{2} & 18 & \frac{5}{2} & -\frac{9}{2} & -\frac{3}{2} \\ 2 & -\frac{9}{2} & \frac{17}{2} & 2 & -\frac{3}{2} & -1 \\ -3 & \frac{9}{2} & -6 & \frac{3}{2} & \frac{9}{2} & \frac{3}{2} \\ 3 & -9 & \frac{7}{2} & -\frac{3}{2} & -3 & -\frac{1}{2} \end{bmatrix}$$

4. **Cayley-Hamilton Theorem**

Definition 19. An **annihilating polynomial** for a square matrix A is a non-zero polynomial $p(t)$ such that $p(A) = 0$.

Theorem 20. *Then there exists an annihilating polynomial for any $n \times n$ matrix A.*

The degree of the annihilating polynomial is n^2. In fact, the degree can be smaller.

Theorem 21 (**Cayley-Hamilton Theorem**). *If $f(t)$ is the characteristic polynomial of A, then $f(A) = \mathbf{0}$.*

Proof. Suppose $f_A(t) = \det(A - tI) = (\lambda_1 - t)^{k_1}(\lambda_2 - t)^{k_2} \cdots (\lambda_p - t)^{k_p}$.
If A is diagonalizable, (i.e., $A = PDP^{-1}$), the proof is easy. Since f is a polynomial,

$$f(A) = Pf(D)P^{-1} = P \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_p) \end{bmatrix} P^{-1} = P\mathbf{0}P^{-1} = \mathbf{0}.$$

In general, we use Jordan normal forms decomposition $A = PJP^{-1}$. We only need to show that $f(J) = \mathbf{0}$.

$$f(J) = \begin{bmatrix} f(J_{\lambda_1}(k_1)) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & f(J_{\lambda_2}(k_2)) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & f(J_{\lambda_m}(k_p)) \end{bmatrix}$$

Each matrix $f(J_{\lambda_i}(k_i)) = (\lambda_1 I - J_{\lambda_i}(k_i))^{k_1} \cdots (\lambda_i I - J_{\lambda_i}(k_i))^{k_i} \cdots (\lambda_p I - J_{\lambda_i}(k_i))^{k_p} = \mathbf{0}$, since $(\lambda_i I - J_{\lambda_i}(k_i))^{k_i} = \mathbf{0}$ by Lemma **??** $\square$

Wrong proof: $f(t) = \det(A - tI)$. So, $f(A) = \det(A - AI) = \det(0) = 0$. (Why?)

5. **Minimal polynomial**

By Cayley-Hamilton Theorem, we know that we can find annihilating polynomial of A with degree $\leq n$.

Definition 22. The smallest degree annihilating polynomial of A is called the **minimal polynomial** of A.

Theorem 23 (Minimal Polynomial Theorem)**.** *Consider $\mathbb{F} = \mathbb{C}$. The eigenvalues of A are the roots of the minimal polynomial $f(t)$ of A.*

Corollary 24. *The minimal polynomial $f(t)$ of A has the form*

$$f(t) = (t - \lambda_1)^{p_1}(t - \lambda_2)^{p_2} \cdots (t - \lambda_m)^{p_m}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_m$ be the distinct eigenvalues of A and the exponents p_k is the largest block size for each eigenvalue..