# LA12 Ch5
# The Jordan Canonical Form

KYB

Thrn, it's a Fact

*mathrnfact@gmail.com*

February 9, 2021

# Overview

## Definition (Nilpotent )

$V$ : v.sp. $T : V \to V$ linear is nilpotent if $T^k = 0$ for some $k > 0$. If such $k$ exists, there is the smallest $k > 0$ such that $T^k = 0$ and we say $T$ is nilpotent of index $k$.

## Example

For $A \in \mathbb{C}^{n\times n}$ with e.val $\lambda$, we know that $\mathcal{N}(A-\lambda I)^k$ is stable. Let $m$ be the smallest number such that $\mathcal{N}(A-\lambda I)^k = \mathcal{N}(A-\lambda I)^{k+1}$ and $V = \mathcal{N}(A-\lambda I)^m$ Then $T : V \to V$ by $T(x) = (A-\lambda I)x$ is nilpotent.

## Theorem (249)

*Let $T : V \to V$ be linear. If $x \in V$ is satisfying that $T^{k-1}(x) \neq 0$ and $T^k(x) = 0$, then $\{x, T(x), \cdots, T^{k-1}(x)\}$ is linearly independent.*

# Theorem (250)

*Let $T : V \to V$ be nilpotent operator of index $k$. Suppoe $x_0 \in V$ is any vector with $V^{k-1}(x) \neq 0$ and define $S =$ span$\{x_0, T(x_0), \cdots, T^{k-1}(x_0)\}$. Then $S$ is invariant under $T$. If $k < \dim(V)$, there exists a subspace $W$ of $V$ such that $W$ is invariant under $T$ and $V = S \oplus W$.*

To prove the existence of $W$, we argue by induction on $k$.

## Step1

$k = 1$.
Then $T(x) = 0$, Thus $W = V$.

## Step2

Suppose the result holds for all nilpotent operators of index $k-1$.
$R = \mathcal{R}(T)$ is invariant under $T$ and $T\big|_R : R \to R$ is nilpotent of index $k-1$. Define $S_0 = S \cap R$ and $y_0 = T(x_0)$. Then

$$S_0 = \text{span} T(x_0), \cdots, T^{k-1}(x_0) = \text{span} y_0, \ldots, T^{k-2}(y_0).$$

Then $\dim(S_0) = k-1$, by I.H, there is $W_0$ such that $R = S_0 \oplus W_0$.

## Step3

We have $S = \{x \in V | T(x) \in S_0\}$. Define $W_1 = \{x \in V | T(x) \in W_0\}$($W_1$ is not desired one). Since $W_0$ is invariant under $T$, $W_0 \subset W_1$.

## Step4

Claim)

- $V = S + W_1$
- $S \cap W_1 = \text{span}\{T^{k-1}(x_0)\}$
- Extend $\{T^{k-1}(x_0)\}$ to a basis $\mathcal{B}$ for $W_1$ and let $W = \text{span}\big(\mathcal{B} - \{T^{k-1}(x_0)\}\big)$.

Then $V = S \oplus W$.

## Theorem (251)

*Let* $T : V \to V$ *be a nilpotent operator of index* $k$. *Then there exists* $\{x_1, \cdots, x_s\} \subset V$ *and integers* $r_1, \cdots, r_s$ *with* $1 \leq r_s \leq \cdots \leq r_1 = k$ *such that*

$$T^{r_i - 1}(x_i) \neq 0, T^{r_i}(x_i) = 0 \text{ for all } i = 1, \cdots, s,$$

*and*

$$\begin{array}{llll} x_1, & T(x_1), & \cdots, & T^{r_1 - 1}(x_1), \\ \vdots & \vdots & & \vdots \\ x_s, & T(x_s), & \cdots, & T^{r_s - 1}(x_s) \end{array}$$

*form a basis for* $V$.

## Step1

We can choose a nonzero vector $x_1 \in V$ such that $T^{k-1}(x_1) \neq 0$ and $T^k(x_1) = 0$. Let $r_1 = k$ and $S_1 =$ span$\{x_1, T((x_1), \cdots, T^{r_1-1}(x_1)\}$.

## Step2

By Thm250, $S_1$ is invariant under $T$ and there is $W_1$ such that $V = S_1 \oplus W_1$. Note that $\dim(W_1) < \dim(V)$ and $T\big|_{W_1}$ is nilpotent of index $W_2 \leq r_1$. Applying Step2 on $W_1$ and then we get $S_1 =$ span$x_2, T(x_2), \cdots, T^{r_2-1}(x_2)$.
Continue in this fashion to find $x_1, \cdots, x_s$.

## Observation

For convenient, assume $k = \dim(V)$. Then $\{x_1, \cdots, T^{k-1}(x_1)\}$ is a basis for $V$. Label $u_1 = T^{k-1}(x_1), \cdots, u_k = x_1$. Then $T(u_1) = 0, T(u_j) = u_{j-1}$ for $j = 2, \cdots, k$. In terms of the isomorphism bwtween $V$ and $F^k$, $u_j \mapsto e_j$, then

$$[T] = J = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$$

In general, for $s > 1$, we can find a basis $\{u_1, \cdots, u_n\}$ such that

$$\begin{aligned} u_1 &= T^{r_1-1}(x_1), \cdots, u_{r_1} = x_1, \\ u_{r_1+1} &= T^{r_2-1}(x_2), \cdots, u_{r_2} = x_2, \\ &\vdots \\ u_{n-r_s+1} &= T^{r_s-1}(x_s), \cdots, u_{r_s} = x_s \end{aligned}$$

with $J_i$, and

$$[A] = J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix}$$

## Lemma (232)

*$A \in F^{n \times n}$. If $A$ has the form*

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

*where $B \in F^{k \times k}$, $C \in F^{k \times (n-k)}$, and $D \in F^{(n-k)\times(n-k)}$, then*

$$\det(A) = \det(B)\det(D) \qquad\qquad p_A(r) = p_B(r)p_D(r).$$

# Proof

It suffices to show that $\det(A) = \det(B)\det(D)$.

$$\det(A) = \sum_{\tau \in S_n} \sigma(\tau) A_{\tau(1)1} \cdots A_{\tau(n)n}$$

Consider $\tau_1 \in S_k$ and $\tau_2 \in S_{n-k}$. Then we can make $\tau \in S_n$ by

$$\tau(i) = \begin{cases} \tau_1(i) & \text{for } i = 1, \ldots, k \\ k + \tau_2(i-k) & \text{for } i = k+1, \cdots, n \end{cases}$$

# Continue

Moreover, we can regard $\tau_1, \tau_2$ as elements, and then $\tau = \tau_1\tau_2$, and $\sigma(\tau) = \sigma(\tau_1)\sigma(\tau_2)$. If $\tau \in S_n$ is not a form $\tau_1\tau_2$, there $i \le k < j$ such that $\tau(i) = j, \tau(j) = i$. Then $A_{\tau(i)i} = 0$. Thus

$$\begin{aligned}
\det(A) &= \sum_{\tau \in S_n} \sigma(\tau) A_{\tau(1)1} \cdots A_{\tau(n)n} \\
&= \sum_{\tau_1, \tau_2} \sigma(\tau_1)\sigma(\tau_2) \left(B_{\tau_1(1)1} \cdots B_{\tau_1(k)k}\right) \left(D_{\tau_2(k+1)k+1} \cdots D_{\tau_2(n)n}\right) \\
&= \det(B)\det(D)
\end{aligned}$$

# Recall

- $A \in F^{n\times n}$, define $N = \mathcal{N}(A)$. If there exists a subspace $R$ of $F^n$ such that $R$ is invariant under $A$ and $F^n = N \oplus R$, then $R = \text{col}(A)$.
- TFAE.
  1. there exists a subspace $R$ of $F^n$ such that $R$ is invariant under $A$ and $F^n = N \oplus R$
  2. $\mathcal{N}(A) \cap \text{col}(A) = \{0\}$, in which case $R = \text{col}(A)$.
- $\mathcal{N}(A) \cap \text{col}(A) = \{0\}$ if and only if $\mathcal{N}(A^2) = \mathcal{N}(A)$.

  $A \in F^{n\times n}$, $\lambda \in F$.
- If $S$ is a subspace of $F^n$, then $S$ is invariant under $A$ iff $S$ is invariant under $A - \lambda I$.

## Theorem (233)

$\lambda$ : *e.val of* $A \in F^{n\times n}$ *and suppose* $\mathcal{N}((A-\lambda I)^2) = \mathcal{N}(A-\lambda)$. *Then* *m.geo*$(\lambda) =$ *m.alg*$(\lambda)$.

## Proof

Let $k = \text{m.geo}(\lambda), m = \text{m.alg}(\lambda)$. We can find bases $\{x_1, \cdots, x_k\}$ and $\{x_{k+1}, \cdots, x_n\}$ for $N = \mathcal{N}(A-\lambda I)$ and $R = \text{col}(A-\lambda)$, respectively. Then $F^n = N \oplus R$. Then $\{x_1, \cdots, x_n\}$ is a basis for $F^n$ and if put $X = [x_1|\cdots|x_n]$,

$$X^{-1}AX = \left[\begin{array}{c|c} \lambda I & 0 \\ \hline 0 & D \end{array}\right]$$

Then $p_A(r) = (r-\lambda)^k p_D(r)$.

# Continue

On the other hand, m.alg$(\lambda) = m$ implies $p_A(r) = (r-\lambda)^m q(r)$ where $q(\lambda) \neq 0$. To prove $k = m$, it suffices to show that $p_D(\lambda) \neq 0$. Let $X_1 = [x_1|\cdots|x_k], X_2 = [x_{k+1}|\cdots|x_n]$. Suppose not. Then there is a nonzero $u \in F^{n-k}$ such that $Du = \lambda u$.

$$A[X_1|X_2] = [X_1|X_2]\left[\begin{array}{c|c} \lambda I & 0 \\ \hline 0 & D \end{array}\right] \Rightarrow AX_1 = \lambda X_2, AX_2 = X_2 D.$$

Define $x = X_2 u$.

$$Ax = AX_2u = X_2Du = X_2(\lambda u) = \lambda(X_2u) = \lambda x.$$

## Corollary (234)

*If $A \in \mathbb{C}^{n\times n}$ and for each e.val $\lambda$ of $A$,*

$$\mathcal{N}((A-\lambda I)^2) = \mathcal{N}(A-\lambda I),$$

*then $A$ is diagonalizable.*

How about $A$ is not diagonalizable?
A: Jordan Canonical Form

## Theorem (235)

*$A \in \mathbb{C}^{n\times n}$, and let $\lambda$ be e.val of $A$ with m.alg$(\lambda) = m$. Then:*

1. *there exists a smallest $k > 0$ such that*
$$\mathcal{N}((A-\lambda I)^{k+1}) = \mathcal{N}((A-\lambda I)^k);$$
2. *for all $l > k$, $\mathcal{N}((A-\lambda I)^l) = \mathcal{N}((A-\lambda I)^k)$;*
3. *$N = \mathcal{N}((A-\lambda I)^k)$ and $R =$ col$((A-\lambda I)^k)$ are invariant under $A$;*
4. *$\mathbb{C}^n = N \oplus R$;*
5. *$\dim(N) = m$.*

## Definition (Generalized eigenspaces)

For such $k$, $G_\lambda(A) = \mathcal{N}((A-\lambda I)^k)$.

1.
there exists a smallest $k > 0$ such that

$$\mathcal{N}((A - \lambda I)^{k+1}) = \mathcal{N}((A - \lambda I)^{k})$$

2.

for all $l > k$, $\mathcal{N}((A - \lambda I)^l) = \mathcal{N}((A - \lambda I)^k)$;

3.
$N = \mathcal{N}((A - \lambda I)^k)$ and $R = \text{col}((A - \lambda I)^k)$ are invariant under $A$;

# 4.

$\mathbb{C}^n = N \oplus R;$

## 5.

$$\dim(N) = m.$$

## Lemma (237)

*Let $A \in \mathbb{C}^{n\times n}$. If $\lambda_i$'s are distinct e.vals of $A$ $(i = 1, \cdots, s)$ and if*

$$x_1 + \cdots + x_s = 0, x_i \in G_{\lambda_i}(A)$$

*then $x_1 = \cdots = x_s = 0$.*

## Theorem (238)

*If*

$$p_A(r) = (r-\lambda_1)^{m_1} \cdots (r-\lambda_t)^{m_t}$$

*where $\lambda_i$'s are the distinct e.vals of $A$. Then $\mathbb{C}^n = \bigoplus G_{\lambda_i}(A)$.*

# Observation

For convenient, assume m.alg$(\lambda) = n$ for $A \in \mathbb{C}^{n \times n}$. Let $N = \mathcal{N}((A - \lambda I)^n)$ Then there is a smallest $k \leq n$ such that

$$\mathcal{N}\left((A - \lambda I)^k\right) = \mathcal{N}((A - \lambda I)^n)$$

Then $x \in N$, $(A - \lambda I)^k x = 0$. But there exists at least one nonzero $x \in N$ such that $(A - \lambda I)^{k-1} x \neq 0$.

## Observation

Thus we can find $\{x_1, \cdots, x_s\}$ and $1 \leq r_1 \leq \cdots \leq r_s = k$ such that

$$(A - \lambda I)^{r_i - 1)} x_i = \neq 0, (A - \lambda I)^{r_i)} x_i = 0$$

and

$$\begin{array}{lll} x_1, (A - \lambda I)x_1, \cdots & , (A - \lambda I)^{r_1 - 1} x_1, \\ \vdots \qquad\quad \vdots & \qquad\qquad \vdots \\ x_s, (A - \lambda I)x_s, \cdots & , (A - \lambda I)^{r_s - 1} x_s \end{array}$$

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

# Observation

Assume $s = 1$. Then $\{x, (A - \lambda I)x, \cdots, (A - \lambda I)^{n-1}x\}$ is a basis for $\mathbb{C}^n$. Define $x_i = (A - \lambda I)^{n-i}x$ for $i = 1, \cdots, n$. Let $X = [x_1|\cdots|x_n]$. Then

$$(A - \lambda I)X = [0|x_1|\cdots|x_{n-1}]$$

$$\begin{aligned} AX &= [\lambda x_1|\cdots|\lambda x_n] + [0|x_1|\cdots|x_{n-1}] \\ &= [x_1|\cdots|x_n]\begin{bmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{bmatrix} + [x_1|\cdots|x_n]\begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} \end{aligned}$$

# Observation

$$AX = X \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix} = XB$$

$$X^{-1}AX = B$$

## Observation

In general, let e.vals $\lambda_i$ with m.alg$\lambda_i = m_i$ $(i = 1, \cdots t)$. For each $i$ let $k_i$ be the nilpotent index of $A - \lambda_i$. Define $N_i = \mathcal{N}((A - \lambda_i)_i^k)$. Then we can find $k_i \geq r_1^{(i)} \geq \cdots \geq r_1^{(s_i)}$ and $x_1^{(i)}, \cdots, x_{s_i}^{(i)}$ such that

$$\begin{array}{lll} x_1^{(i)}, (A - \lambda_i I)x_1^{(i)}, \cdots, & (A - \lambda_i I)^{r_1^{(i)}-1} x_1^{(i)}, \\ \vdots \quad\quad \vdots & \vdots \\ x_{s_i}^{(i)}, (A - \lambda_i I)x_{s_i}^{(i)}, \cdots, & (A - \lambda_i I)^{r_{s_i}^{(i)}-1} x_{s_i}^{(i)} \end{array}$$

form a basis for $N_i$. Define $X_j^{(i)} = [(A - \lambda_i I)^{r_j^{(i)}-1} x_1^{(i)} | \cdots | x_j^{(i)}]$

Then $AX_j^{(i)} = X_j^{(i)} J_j^{(i)}$ where

$$J_j^{(i)} = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{bmatrix}$$

Now define $X_i = [X_1^{(i)} | \cdots | X_{r_i}^{(i)}]$ and $AX_i = X_i B_i$ where

$$B_i = \begin{bmatrix} J_1^{(i)} & & & \\ & J_2^{(i)} & & \\ & & \ddots & \\ & & & J_{s_i}^{(i)} \end{bmatrix}$$

Finally, define $X = [x_1|\cdots|X_t]$. Then $AX = XJ$ where

$$J = \begin{bmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_t \end{bmatrix}$$

We call each $B_i$ Jordan blocks. $A$ is diagonalizable if and only if each Jordan blocks are $1 \times 1$ matrix.

# Example

$$A = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & -1 & -2 & 0 & 1 & -2 \\ -4 & 2 & 0 & -2 & 2 & 1 & 2 & 2 \\ 1 & 11 & -2 & -6 & -13 & -2 & 3 & -13 \\ 1 & 4 & -1 & -2 & -4 & -1 & 1 & -5 \\ -4 & -10 & 3 & 6 & 15 & 2 & -2 & 15 \\ 0 & 10 & -2 & -6 & -11 & -2 & 3 & -11 \\ -1 & -4 & 1 & 2 & 3 & 1 & -1 & 4 \end{bmatrix}$$

# Example

$$J = \left[\begin{array}{ccc|c|cc|cc} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

# Ex5.4.1

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

# Ex5.4.5

$$A = \begin{bmatrix} -3 & 1 & -4 & -4 \\ -17 & 1 & -17 & -38 \\ -4 & -1 & -3 & -14 \\ 4 & 0 & 4 & 10 \end{bmatrix} \in \mathbb{R}^{4\times 4}$$

# Ex5.4.7

$$A = \begin{bmatrix} -7 & 1 & 24 & 4 & 7 \\ -9 & 4 & 21 & 3 & 6 \\ -2 & -1 & 11 & 2 & 3 \\ -7 & 13 & -18 & -6 & -8 \\ 3 & -5 & 6 & 3 & 5 \end{bmatrix} \in \mathbb{R}^{5\times 5}$$

# Ex5.4.4

Let $A$ be a $4 \times 4$ matrix.

(a) $p_A(r) = (r-1)(r-2)(r-3)(r-4)$

(b) $p_A(r) = (r-1)^2(r-2)(r-3)$

(c) $p_A(r) = (r-1)^2(r-2)^2$

(d) $p_A(r) = (r-1)^3(r-2)$

(e) $p_A(r) = (r-1)^4$

# The End