9.4. Wednesday for MAT3040

9.4.1. Jordan Normal Form

Theorem 9.3 — Jordan Normal Form. Suppose that $T: V \to V$ has minimial polynomial

$$m_T(x) = \prod_{i=1}^{k}(x-\lambda_i)^{e_i},$$

then there exists a basis $\mathcal{A}$ such that

$$(T)_{\mathcal{A},\mathcal{A}} = \text{diag}(J_1,\dots,J_\ell),$$

where each block J_i is a square matrix of the form

$$J_i = \begin{bmatrix} \mu_i & 1 & & \\ & \mu_i & \ddots & \\ & & \ddots & 1 \\ & & & \mu_i \end{bmatrix}.$$

R By primary decomposition theorem,

$$V = V_1 \oplus \cdots \oplus V_k, \quad \text{where } V_i = \ker((T-\lambda_i I)^{e_i}),\ i=1,\dots,k,$$

and each V_i is T-invariant.

We pick basis $\mathcal{B}_i$ for each subspace V_i, then $\mathcal{B} := \cup_{i=1}^{k}\mathcal{B}_i$ is a basis of V, and

$$(T)_{\mathcal{B},\mathcal{B}} = \begin{pmatrix} (T\mid_{V_1})_{\mathcal{B}_1,\mathcal{B}_1} & 0 & \cdots & 0 \\ 0 & (T\mid_{V_2})_{\mathcal{B}_2,\mathcal{B}_2} & \vdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \vdots & (T\mid_{V_k})_{\mathcal{B}_k,\mathcal{B}_k} \end{pmatrix}$$

with $m_{T\mid_{V_i}}(x) = (x-\lambda_i)^{e_i}$.

Therefore, it suffices to show the Jordan normal form holds for the linear operator

T with minimal polynomial $m_T(x) = (x - \lambda)^e$.

Firstly, we consider the case where the minimal polynomial has the form x^m:

Proposition 9.5 Suppose $T : V \to V$ is such that $m_T(x) = x^m$, then the theorem (9.3) holds, i.e., there exists a basis $\mathcal{A}$ such that

$$(T)_{\mathcal{A},\mathcal{A}} = \text{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_\ell),$$

where each block J_i is a square matrix of the form

$$J_i = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

Proof. • Suppose that $m_T(x) = x^m$, then it is clear that

$$\{0\} := \ker(T^0) \le \ker(T) \le \ker(T^2) \le \cdots \le \ker(T^m) := V$$

Furthermore, we have $\ker(T^{i-1}) \subsetneq \ker(T^i)$ for $i = 1, \dots, m$: Note that $\ker(T^{m-1}) \subsetneq \ker(T^m) := V$ due to the minimality of $m_T(x)$; and $\ker(T^{m-2}) \subsetneq \ker(T^{m-1})$ since otherwise for any $\boldsymbol{x} \in \ker(T^m)$,

$$T^{m-1}(T\boldsymbol{x}) = \boldsymbol{0} \implies T\boldsymbol{x} \in \ker(T^{m-1}) = \ker(T^{m-2}) \implies T^{m-2}(T\boldsymbol{x}) = T^{m-1}(\boldsymbol{x}) = \boldsymbol{0},$$

i.e., $\boldsymbol{x} \in \ker(T^{m-1})$, which contradicts to the fact that $\ker(T^{m-1}) \subsetneq \ker(T^m)$. Proceeding this trick sequentially for $i = m, m-1, \dots, 1$, we proved the disired result.

• Then construct the quotient space $W_i = \ker(T^i)/\ker(T^{i-1})$ and define $\mathcal{B}_i'$ to be a basis of W_i:

$$\mathcal{B}_i' = \{a_1^i + \ker(T^{i-1}), \dots, a_{\ell_i}^i + \ker(T^{i-1})\}$$

Construct $\mathcal{B}_i = \{a_1^i, \dots, a_{\ell_i}^i\}$, then we claim that $B := \cup_{i=1}^m \mathcal{B}_i$ forms a basis of V:

- First proof the case $m = 2$ first: let $U \le V$ ($\dim(V) < \infty$), and $\mathcal{B}_1 = \{a_1^1, \dots, a_{k_1}^1\}$ be a basis of U, and

$$\mathcal{B}_2' = \{a_1^2 + U, \dots, a_{k_2}^2 + U\}$$

be a basis of V/U. Then to show the statement suffices to show that

$$\bigcup_{i=1}^{2} \{a_1^i, \dots, a_{k_i}^i\} \text{ forms a basis of } V.$$

It's clear that $\cup_{i=1}^{2}\{a_1^i, \dots, a_{k_i}^i\}$ spans V. Furthermore, $\dim(V) = \dim(U) + \dim(V/U) = k_1 + k_2$, i.e., $\cup_{i=1}^{2}\{a_1^i, \dots, a_{k_i}^i\}$ contains correct amount of vectors. The proof is complete.

- This result can be extended from 2 to general m, thus the claim is shown.

- For $i < m$, consider the set $S_i = \{T(\boldsymbol{w}_j) + \ker(T^{i-1}) \mid \boldsymbol{w}_j \in B_{i+1}\}$. Note that
  - Since $T^{i+1}(\boldsymbol{w}_j) = \mathbf{0}$, $T^i(T(\boldsymbol{w}_j)) = \mathbf{0}$, we imply $T(\boldsymbol{w}_j) \in \ker(T^i)$, i.e., $S_i \subseteq W_i$.
  - The set S_i is linearly independent: consider the equation

$$\sum_j k_j(T(\boldsymbol{w}_j) + \ker(T^{i-1})) = \mathbf{0}_{W_i} \iff T\left(\sum_j k_j \boldsymbol{w}_j\right) + \ker(T^{i-1}) = \mathbf{0}_{W_i}$$

i.e.,

$$T\left(\sum_j k_j \boldsymbol{w}_j\right) \in \ker(T^{i-1}) \iff T^{i-1}(T(\sum_j k_j \boldsymbol{w}_j)) = \mathbf{0}_V,$$

i.e., $\sum_j k_j \boldsymbol{w}_j \in \ker(T^i)$, i.e.,

$$\sum_j k_j \boldsymbol{w}_j + \ker(T^i) = \mathbf{0}_{W_{i+1}} \iff \sum_j k_j(\boldsymbol{w}_j + \ker(T^i)) = \mathbf{0}_{W_{i+1}}.$$

Since $\{\boldsymbol{w}_j + \ker(T^i), \ \forall j\}$ fomrs a basis of W_{i+1}, we imply $k_j = 0, \forall j$.

From $\mathcal{B}_{i+1}$ we construct S_i, which is linearly independent in W_i. Therefore, we imply $|T(\mathcal{B}_{i+1})| \le |\mathcal{B}_i|$ for $\forall i < m$ (why?).

- Now we start to construct a basis $\mathcal{A}$ of V:
  - Start with $\mathcal{B}_m' := \{u_1^m + \ker(T^{m-1}), \dots, u_{\ell_m}^m + \ker(T^{m-1})\}$, and $\mathcal{B}_m = \{u_1^m, \dots, u_{\ell_m}^m\}$.

- By the previous result,

$$\{T(u_1^m) + \ker(T^{m-2}), \ldots, T(u_{\ell_m}^m) + \ker(T^{m-2})\}$$

  is linear independent in W_{m-1}. By basis extension, we get a basis $\mathcal{B}'_{m-1}$ of W_{m-1}, and let

$$\mathcal{B}_{m-1} = \{T(u_1^m), \ldots, T(u_{\ell_m}^m)\} \cup \xi_{m-1}$$

  where $\xi_{m-1} := \{u_1^{m-1}, \ldots, u_{\ell_{m-1}}^{m-1}\}$

- Continue the process above to obtain $\mathcal{B}_{m-2}, \ldots, \mathcal{B}_1$, and $\cup_{i=1}^m \mathcal{B}_i$ forms a basis of V:

$\mathcal{B}_1$	$\mathcal{B}_2$	$\cdots$	$\mathcal{B}_{m-1}$	$\mathcal{B}_m$
$\{T^{m-1}(u_1^m), \ldots, T^{m-1}(u_{\ell_m}^m)\}$ $\{T^{m-2}(u_1^{m-1}), \ldots, T^{m-2}(u_{\ell_{m-1}}^{m-1})\}$ $\vdots$ $\{T(u_1^2), \ldots, T(u_{\ell_2}^2)\}$ $\{u_1^1, \ldots, u_{\ell_1}^1)\}$	$\{T^{m-2}(u_1^m), \ldots, T^{m-2}(u_{\ell_m}^m)\}$ $\{T^{m-3}(u_1^{m-1}), \ldots, T^{m-3}(u_{\ell_{m-1}}^{m-1})\}$ $\vdots$ $\{u_1^2, \ldots, u_{\ell_2}^2)\}$	$\cdots$ $\cdots$ $\cdots$	$\{T(u_1^m), \ldots, T(u_{\ell_m}^m)\}$ $\{u_1^{m-1}, \ldots, u_{\ell_{m-1}}^{m-1}\}$	$\{u_1^m, \ldots, u_{\ell_m}^m\}$

- Now construct the ordered basis $\mathcal{A}$:

$$\mathcal{A} = \left\{ \begin{array}{ccccc} T^{m-1}(u_1^m) & \cdots & T^2(u_1^m) & T(u_1^m) & u_1^m \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ T^{m-1}(u_{\ell_m}^m) & \cdots & T^2(u_{\ell_m}^m) & T(u_{\ell_m}^m) & u_{\ell_m}^m \\ & T^{m-2}(u_1^{m-1}) & \cdots & T(u_1^{m-1}) & u_1^{m-1} \\ & \vdots & \ddots & \vdots & \vdots \\ & T^{m-2}(u_{\ell_{m-1}}^{m-1}) & \cdots & T(u_{\ell_{m-1}}^{m-1}) & u_{\ell_{m-1}}^{m-1} \\ & & \vdots & \ddots & \vdots \\ & & & & u_1^1 \\ & & & & \vdots \\ & & & & u_{\ell_1}^1 \end{array} \right\}$$

– Then the diagonal entries of $(T)_{\mathcal{A},\mathcal{A}}$ should be all zero, since

$$T(T^{i-1}(u_j^i)) = T^i(u_j^i) = 0, \forall i = 1,\dots,m, j = 1,\dots,\ell_i,$$

and every entry on the superdiagonal is 1:

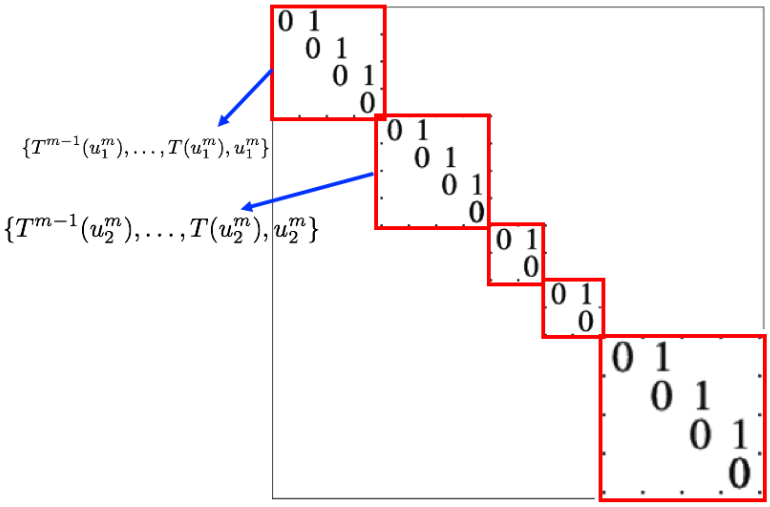

Figure 9.1: Illustration for $(T)_{\mathcal{A},\mathcal{A}}$

■

Then we consider the case where $m_T(x) = (x-\lambda)^e$:

Corollary 9.3 Suppose $T : V \to V$ is such that $m_T(x) = (x-\lambda)^e$, then the theorem (9.3) holds, i.e., there exists a basis $\mathcal{A}$ such that

$$(T)_{\mathcal{A},\mathcal{A}} = \text{diag}(J_1,\dots,J_\ell),$$

where each block J_i is a square matrix of the form

$$J_i = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}.$$

Proof. Suppose that $m_T(x) = (x-\lambda)^e$. Consider the operator $U := T - \lambda I$, then $m_U(x) = x^e$.

By applying proposition (9.5),

$$(U)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_\ell),$$

where

$$J_i = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

Or equivalently,

$$(T)_{\mathcal{A},\mathcal{A}} - \lambda(I)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_\ell)$$

i.e.,

$$(T)_{\mathcal{A},\mathcal{A}} = \operatorname{diag}(\boldsymbol{K}_1, \dots, \boldsymbol{K}_\ell),$$

where

$$\boldsymbol{K}_i = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

■

R The Jordan Normal Form Theorem (9.3) follows from our arguments using the primary decomposition.

Corollary 9.4 Any matrix $A \in M_{n \times n}(\mathbb{C})$ is similar to a matrix of the Jordan normal form

$$\operatorname{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_\ell).$$

9.4.2. Inner Product Spaces

Definition 9.8 [Bilinear] Let V be a vector space over $\mathbb{R}$. A bilinear form on V is a mapping

$$F : V \times V \to \mathbb{R}$$

satisfying

1. $F(\boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w}) = F(\boldsymbol{u}, \boldsymbol{w}) + F(\boldsymbol{v}, \boldsymbol{w})$
2. $F(\boldsymbol{u}, \boldsymbol{v} + \boldsymbol{w}) = F(\boldsymbol{u}, \boldsymbol{v}) + F(\boldsymbol{u}, \boldsymbol{w})$
3. $F(\lambda \boldsymbol{u}, \boldsymbol{v}) = \lambda F(\boldsymbol{u}, \boldsymbol{v}) = F(\boldsymbol{u}, \lambda \boldsymbol{v})$

We say

- F is symmetric if $F(\boldsymbol{u}, \boldsymbol{v}) = F(\boldsymbol{v}, \boldsymbol{u})$
- F is non-degenerate if $F(\boldsymbol{u}, \boldsymbol{w}) = \mathbf{0}$ for $\forall \boldsymbol{u} \in V$ implies $\boldsymbol{w} = 0$
- F is positive definite if $F(\boldsymbol{v}, \boldsymbol{v}) > 0$ for $\forall \boldsymbol{v} \neq \mathbf{0}$