The Galperin–Waksman proof of Jordan canonical form

It suffices to prove:

LEMMA 1. *If* $0 \in \sigma(\underline{A})$*, then there exists a basis with respect to which*

$$A = \left(\begin{array}{c|c} B & 0 \\ \hline 0 & \begin{matrix} 0 & 1 & & & 0 \\ & 0 & 0 & & \\ & & 0 & 1 & \\ & & & 0 & \ddots \\ 0 & & & & \ddots \end{matrix} \end{array}\right).$$

Here the bottom right block is in Jordan canonical form with diagonal elements all equal to 0, and the matrix B is nonsingular, except when it is nonexistent (i.e., when $\sigma(\underline{A}) = \{0\}$).

PROOF OF THEOREM FROM LEMMA 1: Argue by induction on $\dim \mathcal{V}$. Choose $\lambda \in \sigma(\underline{A})$ and apply the lemma to $\underline{A} - \lambda$ in the role of $\underline{A}$. Choose the basis for $\operatorname{dom} \underline{B}$ so as to put B into Jordan canonical form. (This is possible by the inductive hypothesis. If $\dim \mathcal{V} = 1$ (the start of the induction), then the Jordan theorem is trivial.) The result is a Jordan form for $\underline{A} - \lambda$. Add λ (times $\underline{1}$) to get a Jordan form for $\underline{A}$. (Note that λ doesn't appear as an eigenvalue of $\underline{B} + \lambda$, since $\underline{B}$ is nonsingular.)

PROOF OF LEMMA: Let $V_1 = \operatorname{ran} \underline{A}$. (Note that this and other subspaces will not be denoted by script letters in this proof. V_1 should not be confused with $\mathcal{V}(\lambda_1)$.) Since $\underline{A}$ is singular, $V_1 \neq V_0 \equiv \mathcal{V}$.

Let $V_2 \equiv \operatorname{ran} \left(\underline{A}\big|_{V_1}\right) \equiv$ image of V_1 under $\underline{A}$;

$$\vdots$$

$$V_j \equiv \operatorname{ran} \left(\underline{A}\big|_{V_{j-1}}\right) \equiv \underline{A}\left[V_{j-1}\right] = \operatorname{ran} \underline{A}^j;$$

$$\vdots$$

Note by induction that $V_j \subseteq V_{j-1}$ (i.e., $\underline{A}[V_{j-1}] \subseteq \underline{A}[V_{j-2}]$).

Since $\dim \mathcal{V} < \infty$, eventually $\exists h : V_h = V_{h+1} = V_{h+2} = \ldots$. Thus $\underline{B} \equiv \underline{A}\big|_{V_h}$ is nonsingular, with $\operatorname{ran} \underline{B} = \operatorname{dom} \underline{B} = V_h$. For $j \leq h$, we have $V_j \subset V_{j-1}$ properly.

Let $N_j \equiv \ker \left(\underline{A}\big|_{V_{j-1}}\right)$. Then $V_j + N_j \subset V_{j-1}$. [A peek ahead: Vectors in N_j are, of course, 0-eigenvectors of $\underline{A}$. Eventually it will be seen that they are those eigenvectors that stand at the head of a Jordan chain of j vectors; e.g., for

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

the first row and column belong to a basis vector in N_3, the last row and column to one in N_1.] The sum $V_j + N_j$ need not be direct (i.e., $V_j \cap N_j \neq \{\vec{0}\}$, perhaps). Let $\tilde{N}_j \subset N_j$ be a direct complement of V_j within $V_j + N_j$ — so that $V_j \oplus \tilde{N}_j = V_j + N_j \subset V_{j-1}$.

Now $V_j \oplus \tilde{N}_j$ need not be all of V_{j-1}. However, if $j = h$, then it is: If $\vec{w} \in V_{h-1}$, then $\underline{A}\vec{w} \in V_h = V_{h+1}$, so $\underline{A}\vec{w} = \underline{A}\vec{x}$ for some $\vec{x} \in V_h \subset V_{h-1}$. Therefore $\underline{A}(\vec{w} - \vec{x}) = \vec{0}$, and $\vec{w} - \vec{x} \in V_{h-1}$; that is, $\vec{n} \equiv \vec{w} - \vec{x} \in N_h$. Thus $\vec{w} = \vec{x} + \vec{n}$, where $\vec{x} \in V_h$ and $\vec{n} \in N_h$, as claimed. (Incidentally, in this case $\tilde{N}_h = N_h$, since V_h can't contain kernel vectors.)

Thus we have $V_{h-1} = V_h \oplus \tilde{N}_h$ and also $V_{h-1} \oplus \tilde{N}_{h-1} \subset V_{h-2}$. I claim that

$$V_{h-2} = V_{h-1} \oplus \tilde{N}_{h-1} \oplus L_{h-1},$$

where L_{h-1} is a subspace which $\underline{A}$ maps isomorphically onto $\tilde{N}_h$. [Thus an element of L_{h-1} is the *second* vector in a Jordan chain of h vectors.] (Proof of claim postponed to end of proof.) We next look at $V_{h-2} \oplus \tilde{N}_{h-2} \subset V_{h-3}$. I claim (again postponing proof) that

$$V_{h-3} = V_{h-2} \oplus \tilde{N}_{h-2} \oplus L_{h-2},$$

where $\underline{A}$ maps L_{h-2} isomorphically onto $\tilde{N}_{h-1} \oplus L_{h-1}$. [Thus an element of L_{h-2} is either the second vector in a Jordan chain of length $h-1$, or the third vector in a chain of length h.] We continue in this way until we get to $V_0 \equiv \mathcal{V}$. Thus we have a decomposition

$$\begin{aligned} \mathcal{V} &= V_1 \oplus \tilde{N}_1 \oplus L_1 \\ &= V_2 \oplus \tilde{N}_2 \oplus L_2 \oplus \tilde{N}_1 \oplus L_1 \\ &= \ \dots \\ &= V_h \oplus \tilde{N}_h \oplus \tilde{N}_{h-1} \oplus L_{h-1} \oplus \cdots \oplus \tilde{N}_1 \oplus L_1 \end{aligned}$$

And the action of $\underline{A}$ on these subspaces is

$$
\begin{array}{ccc}
0 & & \\
0 & \tilde{N}_h & V_h \\
0 & \tilde{N}_{h-1} & L_{h-1} \\
 & \tilde{N}_{h-2} & L_{h-2} \\
0 & \vdots & \vdots \\
0 & \tilde{N}_3 & L_3 \\
0 & \tilde{N}_2 & L_2 \\
 & \tilde{N}_1 & L_1
\end{array}
$$

To get a Jordan basis, start by choosing a basis for each $\tilde{N}_j$; let the basis for L_{h-1} be the inverse images of the basis vectors for $\tilde{N}_h$; let the basis vectors for L_{h-2} be the inverse images of the basis vectors for $\tilde{N}_{h-1} \oplus L_{h-1}$; etc. The matrix of $\underline{A}$ with respect to this basis has the form claimed in the lemma (with $V_h = \operatorname{dom} \underline{B}$). Indeed, the basis vectors in each chain of inverse images belong to a single Jordan block

$$
\begin{array}{c}
\tilde{N}_j \\ L_{j-1} \\ L_{j-2} \\ \vdots \\ L_2 \\ L_1
\end{array}
\begin{pmatrix}
0 & 1 & & & & \text{\Large 0} \\
 & 0 & 1 & & & \\
 & & 0 & 1 & & \\
 & & & \ddots & \ddots & \\
 & & & & \ddots & 1 \\
\text{\Large 0} & & & & & 0
\end{pmatrix}.
$$

It remains to prove the “claims”:

Lemma 2. *If $V_j = V_{j+1} \oplus M$, then $\exists L \subset V_{j-1}$ such that $V_{j-1} = (V_j + N_j) \oplus L$ and $\underline{A}$ maps L isomorphically onto M.* (This has been applied in cases where $M \equiv \tilde{N}_{j+1} \oplus L_{j+1}$; we then used the fact that $V_j + N_j = V_j \oplus \tilde{N}_j$.)

Proof: Choose a basis $\{\vec{y}_1, \ldots, \vec{y}_m\}$ for M. Since $M \subset V_j \equiv \operatorname{ran} \left(\underline{A}\big|_{V_{j-1}} \right)$, $\exists \vec{z}_1, \ldots, \vec{z}_m \in V_{j-1}$ such that $\vec{y}_j = \underline{A}\vec{z}_j$. The $\vec{z}$'s are independent, and $L \equiv \operatorname{span} \{\vec{z}_1, \ldots, \vec{z}_m\}$ is a subspace of V_{j-1} mapped isomorphically onto M. We need to show:

(A) $\underline{(V_j + N_j) \cap L = \{\vec{0}\}}$: $\vec{x} \in (V_j + N_j) \cap L \;\Rightarrow\; \vec{x} = \vec{v} + \vec{n}$ and $\underline{A}\vec{x} \in M$; the first of these implies $\underline{A}\vec{x} = \underline{A}\vec{v} + \underline{A}\vec{n} = \underline{A}\vec{v} \in V_{j+1}$. Thus $\underline{A}\vec{x} \in V_{j+1} \cap M = \{\vec{0}\}$ (since the sum of these two subspaces is assumed direct). Therefore, $\vec{x} = \vec{0}$, since $\underline{A}\big|_L$ is an isomorphism.

(B) $\underline{(V_j + N_j) + L = V_{j-1}}$: Let $\vec{w} \in V_{j-1}$. Then $\underline{A}\vec{w} \in V_j$, so $\underline{A}\vec{w} = \vec{v} + \vec{m}$ where $\vec{v} \in V_{j+1}$, $\vec{m} \in M$. Then $\vec{v} = \underline{A}\vec{u}$ (for some $\vec{u} \in V_j$), and $\vec{m} = \underline{A}\vec{l}$ (for some $\vec{l} \in L$). That is, $\underline{A}\vec{w} = \underline{A}\vec{u} + \underline{A}\vec{l}$, where $\vec{w}, \vec{u}, \vec{l} \in V_{j-1}$. Thus $\vec{w} - \vec{u} - \vec{l} \in N_j \equiv \ker \left(\underline{A}\big|_{V_{j-1}} \right)$. Therefore, $\vec{w} \in N_j + V_j + L$, QED.

Uniqueness: Obviously Jordan blocks can be placed on the diagonal in any order — each arrangement corresponding to a certain permutation of basis vectors. Beyond this, however, the JCF of $\underline{A}$ is unique — it's characterized by listing the lengths of all Jordan chains associated with each $\lambda_\nu \in \sigma(\underline{A})$. The reason is that these lengths are determined by the dimensions of the spaces V_j for the operator $\underline{A} - \lambda_\nu$, which are defined independently of any choice of basis. Indeed, let s_j $(j = 1, \ldots, h)$ be the number of chains of length j in $\mathcal{U}_\nu$. Then

$$\begin{aligned} \dim\, \mathcal{V} - \dim\, V_1 &= \dim\, (\tilde{N}_1 \oplus L_1) \\ &= s_1 + \cdots + s_L \end{aligned}$$

(these are the vectors at the tail ends of chains);

$$\begin{aligned} \dim\, V_1 - \dim\, V_2 &= \dim(\tilde{N}_2 \oplus L_2) \\ &= s_2 + \cdots + s_L \end{aligned}$$

(these are the vectors second from the end of a chain);

$$\begin{aligned} &\vdots \\ \dim\, V_{h-1} - \dim\, V_h &= \dim\, \tilde{N}_h \\ &= s_h \end{aligned}$$

(these are the eigenvectors at the heads of chains of the maximum length, h). These equations can be solved for the s_j.