The Jordan Form

Philip Pennance[1] Version: February 1985.

1. Reminder.

   Let V be a vector space and $A : V \to V$ linear and I the identity mapping. Suppose that V is a direct sum

   $$V = U \oplus W.$$

   If bases for U and W are conjoined to form a basis for V then the matrix of A can be split into four blocks

   $$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

   and clearly,

   (a) $A_{12} = O \iff U$ is A invariant

   (b) $A_{21} = O \iff V$ is A invariant

2. Example.

   Suppose
   $$V = U \oplus W.$$
   where U, V have bases $\{u_1, u_2\}$ and $\{w_1, w_2\}$ respectively. If

   $$\begin{aligned} Au_1 &= \alpha_1 u_1 + \beta_1 u_2 + \gamma_1 w_1 + \delta_1 w_2 \\ Au_2 &= \alpha_2 u_1 + \beta_2 u_2 + \gamma_2 w_1 + \delta_2 w_2 \\ Av_1 &= \alpha_3 u_1 + \beta_3 u_2 + \gamma_3 w_1 + \delta_3 w_2 \\ Av_2 &= \alpha_4 u_1 + \beta_4 u_2 + \gamma_4 w_1 + \delta_4 w_2 \end{aligned}$$

   then the matrix of A is

   $$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\ \delta_1 & \delta_2 & \delta_3 & \delta_4 \end{bmatrix}.$$

   If U is A invariant, then the matrix reduces to

   $$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ 0 & 0 & \gamma_3 & \gamma_4 \\ 0 & 0 & \delta_3 & \delta_4 \end{bmatrix},$$

   and to

   $$\begin{bmatrix} \alpha_1 & \alpha_2 & 0 & 0 \\ \beta_1 & \beta_2 & 0 & 0 \\ 0 & 0 & \gamma_3 & \gamma_4 \\ 0 & 0 & \delta_3 & \delta_4 \end{bmatrix}$$

   if both U and W are A invariant.

3. Facts from Linear Algebra..

   Let V be a finite dimensional vector space and $A : V \to V$ linear.

   (a) The kernel and image of A are A invariant.

   (b) $A_{|\ker A^i}$ is nilpotent.

   (c) $\ker A^n \subseteq \ker A^{n+1}$

   (d) A $\ker A^n \subseteq \ker A^{n-1}$

4. Definition.

   Let λ be a scalar. Define subspaces (X_i) by

   $$X_i = \ker(A - \lambda I)^i, \quad i \geq 1.$$

5. Remark.

   By (3c)

   $$i < j \implies X_i \subseteq X_j.$$

   Moreover, since V is finite dimensional, there exists a smallest k such that $X_{k+1} = X_k$.

6. Lemma.

   If $X_{k+1} = X_k$, then

   $$X_{k+l} = X_k, \text{ for all } l > 0.$$

   Proof.

   By induction. If $l = 1$ the statement is true by hypothesis. Let $l \in \mathbb{N}$ and assume that $X_{k+l-1} = X_k$. We must show

[1] https://pennance.us

that $X_{k+l} = X_k$. Suppose to the contrary that this is not the case. Then there is a vector $v \in X_{k+l} \setminus X_{k+l-1}$. Hence,

$$(A-\lambda I)^{k+l} v = O, \quad (A-\lambda I)^{k+l-1} v \neq O$$

Let $w = A - \lambda I)v$. Then

$$(A-\lambda I)^{k+l} w = O, \quad (A-\lambda I)^{k+l-1} w \neq O$$

contradicting the induction hypothesis.

7. Lemma.

   Let k be the smallest integer such that $X_{k+1} = X_k$. Then

   $$V = \ker(A-\lambda I)^k \oplus \mathrm{Im}(A-\lambda I)^k, k \geq 1.$$

   Proof.

   From basic linear algebra, the dimensions of both sides are equal. It therefore suffices to show that

   $$\ker(A-\lambda I)^k \cap \mathrm{Im}(A-\lambda I)^k = \{O\}.$$

   Let $v \in \ker(A-\lambda I)^k \cap \mathrm{Im}(A-\lambda I)^k$. Then

   $$(A-\lambda I)^k v = O$$

   and there exists w such that

   $$v = (A-\lambda I)^k w.$$

   It follows that

   $$\mathrm{Im}(A-\lambda I)^{2k} w = O.$$

   By the preceeding claim,

   $$(A-\lambda I)^{2k} w = (A-\lambda I)^k w$$

   and so $v = O$.

8. Lemma.

   Let $\mu_1, \dots \mu_r$ be the distinct eigenvalues of A. Let

   $$V_i = \ker(A - \mu_i I)^k.$$

   Then

   (a) $A_{|V_i} \subseteq V_i$.

   (b) $A_{|V_i} = \mu_i I_{|V_i} + X_i$ where X_i is nilpotent.

   (c)

   $$\begin{aligned} V &= V_1 \oplus V_2 \oplus \cdots \oplus V_r \\ &= \bigoplus_{i=1}^{r} \ker(A - \mu_i I)^k. \end{aligned}$$

   (d) $\dim V_i$ is the multiplicity of μ_i

Proof of (a).

$$\begin{aligned} (A-\mu_i I)V_i &\overset{(3d)}{\subseteq} \ker(A-\mu_i I)^{n-1} \\ &\overset{(3c)}{\subseteq} \ker(A-\mu_i I)^n \\ &= V_i. \end{aligned}$$

Since $\mu_i I)V_i \subseteq V_i$ we have

$$A_| V_i = (A - \mu_i I + \mu_i I)V_i \subseteq V_i.$$

Proof of (b).

$$A_{|V_i} = \mu_i I + (A - \mu_i I_{|V_i})$$

and $(A - \mu_i I)_{|V_i}$ is nilpotent by (3b) .

Proof of (c).

We use induction on the dimension of V. We first show that μ_i is the only eigenvalue of $A_{|V_i}$. Suppose that there exists a non zero vector $v \in V_i$ and $(A - \mu_j I)v = O$. Then

$$(A - \mu_i I)v = (\mu_j - \mu_i)v.$$

It follows that

$$O = (A - \mu_i I)^k v = (\mu_j - \mu_i)^k v$$

and so $\mu_j = \mu_i$. Now, by (7)

$$\begin{aligned} V &= \ker(A - \mu_i I)^k \oplus \mathrm{Im}(A - \mu_i I)^k \\ &= V_i \oplus V' \end{aligned}$$

where $V' = (A - \mu_i I)^n$. Since μ_i is the only eigenvalue of $A_{|V_i}$. it must be that $\dim V' < \dim V$ and (8c) follows by induction on the dimension of V

Proof of (d).

Since μ_i is the only eigenvalue of $A_{|V_i}$. The characteristic polinomial

$$P_{A|V_i}(x) = (x - \mu_i)^{n_i}$$

where $n_i = \dim V_i$. Since

$$P_A(x) = P_{A|V_i}(x) P_{A|V'}(x),$$

the multiplicity of μ_i is given by adding to n_i to the multiplicity of μ_i in $P_{A|V'}$. The latter is zero since μ_i is not an eigenvalue of $A_{|V'}$. (If $(A - \mu_1 I)v = 0$ for some $v \in V'$, then $(A - \mu_i I)^n v = O$. But $(A - \mu_i I)^n_| A - \mu_i I)^n$ is surjective and hence injective. It follows that $v = O$ which is a contradiction.

9. Corollary.

   A matrix of $A \in \mathrm{End}(V)$ has block diagonal form

$$\begin{bmatrix} \mu_1 I_{n_1} + N_1 & & O \\ & \ddots & \\ O & & \mu_r I_{n_r} + N_r \end{bmatrix}$$

   where the dimension of the i'th block is the multiplicity of eigenvalue μ_u and N_i is nilpotent.

   It will be shown that by a suitable choice of basis the matrix of A can be further reduced

10. Definition.

    Let $n \geq 1$ be a natural number and λ a scalar. The *Jordan Block* $J_n(\lambda)$ is the $n \times n$ matrix

$$\lambda I_n + S_n$$

    where I_n is the $n \times n$ identity matrix and S_n is the matrix defined in terms of its action on a basis $\{\delta_1, \delta_2, \ldots \delta_n\}$ by

$$S_n(\delta_1) = O$$
$$S_n(\delta_j) = \delta_{j-1}.$$

11. Example.

$$J_\lambda(4) = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}.$$

12. Claim.

    The matrix S_n is nilpotent with index n.

    Proof.

    Notice that n iterates of S_n kill the basis and $n-1$ do not.

13. Corollary.

    Powers of Jordan blocks can be computed as follows:

$$\begin{aligned} J_n^m(\lambda) &= (\lambda I + S)^m \\ &= \lambda^m I + \binom{m}{1} \lambda^{m-1} S + \cdots \\ &+ \binom{m}{n-1} \lambda^{m-n+1} S^{n-1}. \end{aligned}$$

14. Definition.

    A *Jordan matrix* is a direct sum of Jordan blocks.

15. Example.

    The Jordan matrix $J_3(5) \oplus J_2(2) \oplus J_2(3)$ is

$$\left[\begin{array}{ccc|cc|cc} 5 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{array}\right]$$

16. Remark.

Since

$$(A \oplus B)^n = A^n \oplus B^n,$$

powers of Jordan matrices are easy to compute.

17. Lemma.

Let V be a finite dimensional vector space. If $N \in \operatorname{End}(V)$ is nilpotent with index p. Let $v \in V$ such that $N^p v = O$ and $N^{p-1} v \neq O$. Then

$$vN = (v, Nv, N^2 v, \dots N^{p-1} v)$$

is an independent list.

Proof.

If the vectors are not independent, one of them, say $N^i v$, can be written as a linear combination of the later ones. I.e.,

$$N^i v = \sum_{j=i+1}^{p-1} c_j N^j v.$$

Applying N^{p-i-1} to both sides yields

$$\begin{aligned} N^{p-1} &= \sum_{j=i+1}^{p-1} c_j N^{j+p-i-1} v \\ &= O \end{aligned}$$

contradicting the definition of p.

18. Lemma.

Let X_v be the subspace generated by the vectors

$$(v, Nv, N^2 v, \dots N^{k-1} v)$$

where $k < p$ is the smallest integer such that $N^k v = 0$. The matrix of $N_{|X_v}$ relative to the basis

$$\delta_1 = N^{k-1} v, \delta_2 = N^{k-2} v, \dots, \delta_k = v$$

is the matrix S_k defined in (10). Specifically, S_k is the Jordan block

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

with zeros a diagonal entries.

Proof.

Observe that

$$\begin{aligned} &N(\delta_1) = O \\ &N(\delta_j) = \delta_{j-1}, 1 < j \leq k \end{aligned}$$

19. Lemma (Basis extension).

Let N be nilpotent of index p and $(v_i)_{i=1}^{l}$ is a maximal length sequence of vectors such that

$$N^{p-1} v_1, N^{p-1} v_2, \dots, N^{p-1} v_l$$

are independent. For each i let v_{iN} be the list

$$(N v_i, N^2 v_i, \dots N^{p-1} v_i)$$

then there exist $w_1, w_2, \dots w_n \in V$ such that

$$v_{1N} \cup v_{2N} \cup \dots, v_{lN} \cup w_{1N} \cup \dots \cup w_{nN}$$

is a basis for V.

Proof.

By induction on p for arbitrary V. If $p = 1$ then $N^{p-1} = I$ and the result is trivial since any linearly independent set can be extended to a basis. Assume the proposition is true for nilpotent operators of index $p-1$ and let N be nilpotent with index p. Let $v_1, v_2, \dots v_l \in V$. a maximal length list of vectors such that

$$N^{p-1} v_1, N^{p-1} v_2, \dots N^{p-1} v_l$$

are independent. Then letting $\bar{v}_i = N v_i$, we see that $(\bar{v}_i)_{i=1}^{l}$ is a maximal length independent list such that

$$N^{p-2}(\bar{v}_1), N^{p-2}(\bar{v}_2), \dots, N^{p-2}(\bar{v}_l)$$

is independent list. Notice that $\bar{v}_i \in \ker N^{p-1}$ and $N_{|\ker N^{p-1}}$ is nilpotent of index $p-1$. By induction, there exist $w_1, w_2, \dots w_n$ such that such that

$$\bar{\mathcal{B}} = \bar{v}_{1N} \cup \cdots \cup \bar{v}_{l,N} \cup w_{1N} \cdots \cup w_{lN}$$

is a basis for $\ker N^{p-1}$. We claim

$$\mathcal{B} = \{v_1, v_2, \dots v_l\} \cup \bar{\mathcal{B}}$$

is a basis for V. Before proving this, Note that for each i

$$\begin{aligned} v_i \cup \bar{v}_{iN} &= (v_i, \bar{v}_i, N\bar{v}_1, N^2\bar{v}_i, \dots N^{p-2}\bar{v}_i) \\ &= (v_i, Nv_i, N^2v_i, \dots N^{p-1}v_1 \\ &= v_{iN} \end{aligned}$$

and it can be concluded that

$$\mathcal{B} = v_{1N} \cup v_{2N} \cup \dots, v_{lN} \cup w_{1N} \cup \dots \cup w_{nN}$$

is a basis for V.

Linear Independence:

$$\begin{aligned} &\sum c_i v_i + \sum\{c_z z : z \in \bar{\mathcal{B}}\} = O \\ &\implies \sum c_i v_i = -\sum\{c_z z : z \in \bar{\mathcal{B}}\} \\ &\implies \sum c_i v_i = -\sum\{c_z z : z \in \bar{\mathcal{B}}\} \\ &\implies \sum c_i N^{p-1} v_i = -\sum\{c_z N^{p-1} z : z \in \bar{\mathcal{B}}\} \\ &\implies \sum c_i N^{p-1} v_i = O \end{aligned}$$

where the final implication follows from the fact that $\bar{\mathcal{B}}$ is a basis for $\ker N^{p-1}$. Therefore $c_i = 0$ since $N^{p-1}v_i$ are independent. Hence $\sum\{c_z z : z \in \bar{\mathcal{B}}\} = O$ and so $c_z = 0$ since $\bar{\mathcal{B}}$ is an independent set.

Spanning:

Let $v \in V$. . By the maximality of $(v_1, \dots v_l)$, the vectors

$$N^{p-1}v, N^{p-1}v_1, \dots N^{p-1}v_l$$

are dependent, and so there exist constants c_i such that

$$N^{p-1}v = \sum\{c_i N^{p-1} v_i : 1 \le i \le l\}.$$

and it follows that

$$v - \sum_{i=1}^{l} c_i v_i \in \ker N^{p-1}.$$

Hence there exist constants c_z such that

$$v - \sum_{i=1}^{l} c_i v_i = \sum\{c_z z : z \in \bar{\mathcal{B}}\}$$

and so $v \in \operatorname{span} \mathcal{B}$.

20. Theorem (Jordan).

    Let $A \in \mathbb{C}^{n \times n}$, then

    (a) $A \sim J$, a Jordan matrix.

    (b) The diagonal elements of J are the eigenvalues of A.

    Proof.

    By (8c)

    $$A \sim \bigoplus_{1}^{r} (\mu_i I_i + N)$$

    where the N_i are nilpotent. By lemmas (19) and (18) each N_i is similar to a direct sum of Jordan blocks with zeros as the diagonal entries. Let P_i invertible such hath $P_i^{-1} N_i P_i = S_i$. Then

    $$P_i^{-1}(\mu_i I_i + N_i) P_i = S_i$$

    and so A is similar to

    $$\begin{bmatrix} P_1^{-1} & & \\ & \ddots & \\ & & P_r^{-1} \end{bmatrix} \begin{bmatrix} \mu_1 I + N_1 & & \\ & \ddots & \\ & & \mu_r I + N_r \end{bmatrix} \begin{bmatrix} P_1 & & \\ & \ddots & \\ & & P_r \end{bmatrix}$$

    I.e.,

    $$A \sim \begin{bmatrix} \mu_1 I + S_1 & & \\ & \ddots & \\ & & \mu_r I + S_r \end{bmatrix}$$

    a direct sum of Jordan blocks.

21. Remark.

    Since permutation of blocks is a similarity operation, two Jordan matrices are similar if they have the same blocks.