Chapter 8 Jordan Forms

Jordan Canonical Forms

No.
Date: / /

Example 1 $\exists$ basis $\{v_1, v_2, v_3, v_4, v_5\}$

$$A = \left[\begin{array}{ccc|cc} 8 & 1 & & & \\ & 8 & 1 & & \\ & & 8 & & \\ \hline & & & 8 & 1 \\ & & & & 8 \end{array}\right] \iff A - 8I = \left[\begin{array}{ccc|cc} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & & \\ \hline & & & 0 & 1 \\ & & & & 0 \end{array}\right]$$

$$\begin{aligned} A\mathbf{v}_1 &= 8\mathbf{v}_1 & (A-8I)\mathbf{v}_1 &= \mathbf{0} \\ A\mathbf{v}_2 &= 8\mathbf{v}_2 + \mathbf{v}_1 & (A-8I)\mathbf{v}_2 &= \mathbf{v}_1 \\ A\mathbf{v}_3 &= 8\mathbf{v}_3 + \mathbf{v}_2 & (A-8I)\mathbf{v}_3 &= \mathbf{v}_2 \\ A\mathbf{v}_4 &= 8\mathbf{v}_4 & (A-8I)\mathbf{v}_4 &= \mathbf{0} \\ A\mathbf{v}_5 &= 8\mathbf{v}_5 + \mathbf{v}_4 & (A-8I)\mathbf{v}_5 &= \mathbf{v}_4 \end{aligned}$$

- $P_A(t) = -(t-8)^5$, $\lambda = 8$, $m = 5$, $d = 2$
- Jordan matrix
  Jordan block $= \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$ (近似对角矩陣)
- $\begin{cases} d_i \text{ 個 blocks} \\ m_i = \text{size 和} \end{cases}$ (for λ_i)
- $\forall i$, $d_i = m_i$, 则可对角化

$(A-8I)$: $\mathbf{v}_3 \to \mathbf{v}_2 \to \mathbf{v}_1 \to 0$ (string)
$\mathbf{v}_5 \to \mathbf{v}_4 \to 0$

$$\begin{tikzcd} x \arrow[r, "A"] \arrow[d, "A-\lambda I"'] & Ax \arrow[d, "A-\lambda I"] \\ v \arrow[r, "A"] & Av \end{tikzcd}$$

Lemma 1 $A\big(C(A-\lambda I)\big) \subset C(A-\lambda I)$

証: $v \in C(A-\lambda I)$, $\exists\, x \in \mathbb{R}^n$, $(A-\lambda I)\, x = v$ $\quad A(A-\lambda I) = (A-\lambda I)A = A^2 - \lambda A$

$(A-\lambda I)Ax = Av \in C(A-\lambda I)$

定理 3 (Cayley-Hamilton) $P_A(A) = 0$

$\exists$ basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_5\}$

$$A = \left[\begin{array}{ccc|cc} \lambda_1 & 1 & & & \\ & \lambda_2 & 1 & & \\ & & \lambda_3 & & \\ \hline & & & \lambda_4 & 1 \\ & & & & \lambda_5 \end{array}\right]$$

$$\begin{aligned} (A-\lambda_1 I)\mathbf{v}_1 &= \mathbf{0} \\ (A-\lambda_2 I)\mathbf{v}_2 &= \mathbf{v}_1 \\ (A-\lambda_3 I)\mathbf{v}_3 &= \mathbf{v}_2 \\ (A-\lambda_4 I)\mathbf{v}_4 &= \mathbf{0} \\ (A-\lambda_5 I)\mathbf{v}_5 &= \mathbf{v}_4 \end{aligned}$$

証: $P_A(t) = -(t-\lambda_1)(t-\lambda_2)(t-\lambda_3)(t-\lambda_4)(t-\lambda_5)$

$P_A(A) = -(A-\lambda_1 I)(A-\lambda_2 I)(A-\lambda_3 I)(A-\lambda_4 I)(A-\lambda_5 I)$

$(A-\lambda_i I)(A-\lambda_j I) = (A-\lambda_j I)(A-\lambda_i I)$

$$\begin{aligned} P_A(A)\mathbf{v}_1 &= \mathbf{0}, & \because (A-\lambda_1 I)\mathbf{v}_1 &= \mathbf{0} \\ P_A(A)\mathbf{v}_2 &= \mathbf{0}, & (A-\lambda_1 I)(A-\lambda_2 I)\mathbf{v}_2 &= \mathbf{0} \\ P_A(A)\mathbf{v}_3 &= \mathbf{0}, & (A-\lambda_1 I)(A-\lambda_2 I)(A-\lambda_3 I)\mathbf{v}_3 &= \mathbf{0} \\ P_A(A)\mathbf{v}_4 &= \mathbf{0}, & (A-\lambda_4 I)\mathbf{v}_4 &= \mathbf{0} \\ P_A(A)\mathbf{v}_5 &= \mathbf{0}, & (A-\lambda_4 I)(A-\lambda_5 I)\mathbf{v}_5 &= \mathbf{0} \end{aligned}$$

$\Rightarrow \forall \mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_5\mathbf{v}_5 \in \mathbb{R}^5$, $P_A(A)\mathbf{x} = \mathbf{0}$

$\Rightarrow P_A(A) = O$

定理 2 $P_A(t) = \det(A - tI)$ 有 n 個實根，則 $A \sim J_n$ （實矩陣）

證 設 $P_A(t) = -(t-5)^9(t-8)^3(t-9)$

$$\begin{cases} \lambda_1 = 5, & m_1 = 9, & d_1 = 4 \quad (\text{block 個數}) \\ \lambda_2 = 8, & m_2 = 3, & d_2 = 1 \\ \lambda_3 = 9, & m_3 = 1, & d_3 = 1 \end{cases}$$

$$A \sim,\ A' \sim \begin{array}{c} \left[\begin{array}{cccc|ccc|c|c|ccc|c} 5 & 1 & & & & & & & & & & & \\ & 5 & 1 & & & & & & & & & & \\ & & 5 & 1 & & & & & & & & & \\ & & & 5 & & & & & & & & & \\ \hline & & & & 5 & 1 & & & & & & & \\ & & & & & 5 & 1 & & & & & & \\ & & & & & & 5 & & & & & & \\ \hline & & & & & & & 5 & & & & & \\ \hline & & & & & & & & 5 & & & & \\ \hline & & & & & & & & & 8 & 1 & & \\ & & & & & & & & & & 8 & 1 & \\ & & & & & & & & & & & 8 & \\ \hline & & & & & & & & & & & & 9 \end{array}\right] \\ \begin{array}{ccccccccccccc} w_1 & w_2 & w_3 & u_1 & w_4 & w_5 & u_2 & v_1 & v_2 & w_6 & w_7 & w_8 & w_9 \end{array} \end{array}$$

$N(A-5I)$

$v_1 \to 0$

$v_2 \to 0$

$(A-5I):\ u_1 \to w_3 \to w_2 \to w_1 \to 0$

$u_2 \longrightarrow w_5 \to w_4 \to 0$

$w_6 \to 3w_6$

$w_7 \to 3w_7 + w_6$

$w_8 \to 3w_8 + w_7$

$C(A-5I)$ $\quad w_9 \to 4w_9$

(1) (Induction on n), $n=1$, $A \sim [a]$

(2) $n = 13$, $\lambda = 5$, $\begin{cases} \dim N(A-5I) = r = 4 > 0, \ \dim C(A-5I) = n - r = 9 < n \\ \dim N(A-5I) \cap C(A-5I) = 2, \end{cases}$ （設 $\supset \{w_1, w_4\}$ ）

(3) $A(C(A-5I)) \subset C(A-5I) \Rightarrow A|_{C(A-5I)} = A' : C(A-5I) \to C(A-5I)$ （ A' eigen $\Rightarrow$ A eigen

$\exists$ $C(A-5I)$ 之基底 $\{w_1, w_2, \ldots, w_9\}$, $A' \sim J_9$ （Induction : $9 < n$ ）

(4) $N(A-5I)$ 〃 $\{w_1, w_4, v_1, v_2\}$ （extend w_1, w_4 ）

(5) $\begin{cases} w_3 \\ w_4 \end{cases} \in C(A-5I)$, $\exists$ $\begin{matrix} u_1 \xrightarrow{\quad} w_3 \\ u_2 \xrightarrow[A-5I]{} w_5 \end{matrix}$

(6) $\{w_1, w_2, \ldots, w_9, u_1, u_2, v_1, v_2\}$ 獨立 $\overset{(甲)}{\Rightarrow}$ $\mathbb{R}^{13}$ 之基底，且 $A \sim J_{13}$

$$c_1w_1 + c_2w_2 + c_3w_3 + c_4w_4 + c_5w_5 + c_6w_6 + c_7w_7 + c_8w_8 + c_9w_9 + b_1u_1 + b_2u_2 + a_1v_1 + a_2v_2 = 0$$

$(A-5I)\downarrow$

$$0 + c_2w_1 + c_3w_2 + c_5w_4 + 3c_6w_6 + 3c_7w_7 + 3c_8w_8 + 4c_9w_9 + b_1w_3 + b_2w_5 + c_7w_6 + c_8w_7 = 0$$

$$\Rightarrow \begin{cases} c_2 = c_3 = c_5 = c_8 = c_9 = b_1 = b_2 = 0 \\ 3c_6 + c_7 = 0 \\ 3c_7 + c_8 = 0 \end{cases} \Rightarrow c_7 = c_6 = 0$$ （ $C(A-5I)$ 基底）

$\Rightarrow$ $c_1w_1 + c_4w_4 + a_1v_1 + a_2v_2 = 0 \Rightarrow c_1 = c_4 = a_1 = a_2 = 0$ （ $N(A-5I)$ 基底） **得証**

(2) 如何求基底？ (1) $\begin{cases} (a)\ \{w_1, w_4\} \to \{w_2, w_3, u_1, w_5, u_2\} \\ (b)\ \{u_1, u_2\} \to \{w_3, w_2, w_1, w_5, w_4\} \end{cases}$ (2) $\{w_1, w_4\} \to \{v_1, v_2\}$

Example 2

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

$$P_A(t) = t^6 + 3t^5 - 10t^3 - 15t^2 - 9t - 2 = (t+1)^5(t-2)$$

$$\begin{cases} \lambda_1 = -1, & m_1 = 5 \\ \lambda_2 = 2, & m_2 = 1 \end{cases}$$

$N((A+I)^i) = \{ v \in \mathbb{R}^6 \mid (A+I)^i v = 0 \}$, $i = 1, 2, 3$ $\qquad d_1 = \gamma_1 = 2$

$$A + I = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & -7 & 4 & -3 & 1 & -3 \\ -3 & 13 & -7 & 6 & 2 & 9 \\ -2 & 14 & -7 & 5 & 2 & 10 \\ 1 & -18 & 11 & -11 & 3 & -6 \\ -1 & 19 & -11 & 10 & -2 & 8 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 1 & 3/2 \\ 0 & 0 & 1 & 0 & 2 & 3/2 \\ 0 & 0 & 0 & 1 & 0 & -1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = x_5 \begin{bmatrix} 1 \\ -1 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} 1 \\ -3/2 \\ -3/2 \\ 1/2 \\ 0 \\ 1 \end{bmatrix}$$

$\gamma_2 = 4$

$$(A+I)^2 = \begin{pmatrix} 1 & -1 & 0 & 1 & -2 & -3 \\ -2 & -16 & 9 & -11 & 4 & -3 \\ -1 & 37 & -18 & 17 & 2 & 21 \\ 1 & 35 & -18 & 19 & -2 & 15 \\ -1 & -53 & 27 & -28 & 2 & -24 \\ 2 & 52 & -27 & 29 & -4 & 21 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1/2 & 3/2 & -2 & -5/2 \\ 0 & 1 & -1/2 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, x_3 \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} \frac{-3}{2} \\ \frac{-1}{2} \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$+ x_6 \left[\frac{5}{2} \; -\frac{1}{2} \; 0 \; 0 \; 0 \; 1 \right]^T$

$$(A+I)^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -54 & 27 & -27 & 0 & -27 \\ 0 & 108 & -54 & 54 & 0 & 54 \\ 0 & 108 & -54 & 54 & 0 & 54 \\ 0 & -162 & 81 & -81 & 0 & -81 \\ 0 & 162 & -81 & 81 & 0 & 81 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 1 & -1/2 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ \frac{1}{2} \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ \frac{-1}{2} \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} 0 \\ \frac{-1}{2} \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$\gamma_3 = 5$

$\lambda_1 = -1$ $\qquad d_1 = 2$, $m_1 = 5$ $\qquad \begin{cases} v_{1,1} \\ v_{2,1} \end{cases}$ 满足

$\gamma_3 = 5$, $\gamma_2 = 4$, $\gamma_1 = 2$

$(A+I)$: $v_{1,1} \to v_{1,2} \to v_{1,3} \to 0$, $\quad * \begin{cases} (A+I)^3 v_{1,1} = 0 \\ (A+I)^2 v_{1,1} \neq 0 \end{cases}$

$$\mathbf{v}_{1,1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_{1,2} = \begin{bmatrix} 1 \\ 0 \\ -3 \\ -2 \\ 1 \\ -1 \end{bmatrix}, \mathbf{v}_{1,3} = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 1 \\ -1 \\ 2 \end{bmatrix}.$$

$\underset{A}{v_{2,1}} \to v_{2,2} \to 0$, $\quad \begin{cases} (A+I)^2 v_{2,1} = 0 \\ (A+I) v_{2,1} \neq 0 \end{cases}$

$$\mathbf{v}_{2,1} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_{2,2} = \begin{bmatrix} 1 \\ 1 \\ -4 \\ -2 \\ 5 \\ -4 \end{bmatrix}$$

(至少有1满足*)

$\mathrm{span}(v_{1,1}, v_{1,2}, v_{1,3})$

$\lambda_2 = 2$ $\qquad d_2 = 1$, $m_2 = 1$ $\qquad v = [0 \; 1 \; -2 \; -2 \; 3 \; -3]^T$

$$P^{-1}AP = J = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & -2 & 0 & 0 & 1 \\ -4 & 2 & -1 & -3 & 0 & -2 \\ -2 & 0 & 1 & -2 & 0 & -2 \\ 5 & 0 & -1 & 1 & 0 & 3 \\ -4 & 0 & 2 & -1 & 0 & -3 \end{pmatrix}$$

$\bullet \leftarrow v_{2,1} \quad \bullet \leftarrow \bullet \leftarrow v_{1,1}$

$A+I$

$v_{2,2} \; v_{2,1} \quad v_{1,3} \; v_{1,2} \; v_{1,1} \quad v$

$A+I$

Jordan Normal Form

§1. Jordan's Theorem

Definition The n by n matrix $J_{\lambda,n}$ with λ's on the diagonal, 1's on the superdiagonal and 0's elsewhere is called a *Jordan block* matrix. A *Jordan* matrix or matrix in *Jordan normal form* is a block matrix that is has Jordan blocks down its block diagonal and is zero elsewhere.

Theorem Every matrix over $\mathbf{C}$ is similar to a matrix in Jordan normal form, that is, for every A there is a P with $J = P^{-1}AP$ in Jordan normal form.

§2. Motivation for proof of Jordan's Theorem

Consider Jordan block $A = J_{\lambda,n}$, for example,

$$A = J_{5,3} = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix}.$$

We see that

$$\begin{aligned} A\mathbf{e}_1 &= 5\mathbf{e}_1 \\ A\mathbf{e}_2 &= \mathbf{e}_1 + 5\mathbf{e}_2 . \\ A\mathbf{e}_3 &= \mathbf{e}_2 + 5\mathbf{e}_3 \end{aligned}$$

Writing $A_5 = A - 5I$ this becomes:

$$\begin{aligned} A_5\mathbf{e}_1 &= \mathbf{0} \\ A_5\mathbf{e}_2 &= \mathbf{e}_1 . \\ A_5\mathbf{e}_3 &= \mathbf{e}_2 \end{aligned}$$

which can be conveniently rewritten as a *string* of length 3 with value 5:

$$\mathbf{e}_3 \xrightarrow{A_5} \mathbf{e}_2 \xrightarrow{A_5} \mathbf{e}_1 \xrightarrow{A_5} \mathbf{0}$$

Since $A_5\mathbf{e}_1 = \mathbf{0}$, $\mathbf{e}_1$ is an eigenvector with value 5. $(A_5)^2\mathbf{e}_2 = \mathbf{0}$ and $(A_5)^3\mathbf{e}_3 = \mathbf{0}$ and so $\mathbf{e}_2$ and $\mathbf{e}_3$ are called generalized eigenvectors. Although there is no basis of eigenvectors, there is a basis of generalized eigenvectors.

Definition Define $A_\lambda = A - \lambda I$. Call $\mathbf{v} \neq \mathbf{0}$ a *generalized eigenvector* with value λ for A if $(A_\lambda)^p\mathbf{v} = \mathbf{0}$ for some natural p. If $p = 1$, $\mathbf{v}$ is called an *eigenvector*.

§3. Proof of Jordan's Theorem

Introduction to the proof Although there is no basis of eigenvectors, we show there is a basis of generalized eigenvectors. More specifically we find a collection of strings:

$$\begin{array}{ccccccc} \mathbf{w}_{1,n_1} & \xrightarrow{A_{\lambda_1}} & \cdots & \xrightarrow{A_{\lambda_1}} & \mathbf{w}_{1,1} & \xrightarrow{A_{\lambda_1}} & \mathbf{0} \\ & & \vdots & & \vdots & & \vdots \\ \mathbf{w}_{k,n_k} & \xrightarrow{A_{\lambda_k}} & \cdots & \xrightarrow{A_{\lambda_k}} & \mathbf{w}_{k,1} & \xrightarrow{A_{\lambda_k}} & \mathbf{0} \end{array}$$

such that the $\mathbf{w}_{i,j}$'s form a basis. With respect to this basis the matrix of A is in Jordan normal form because the i-th string generates a Jordan block J_{λ_i,n_i}, and conversely a Jordan matrix generates a collection of strings of basis vectors. Accordingly we concern ourselves with generating strings of basis vectors.
The proof we give is due to Filippov (see *Linear Algebra and Its Applications* by G. Strang).

Proof Let A be n by n. The case $n = 1$ is trivial. By "strong" induction, assume every smaller size matrix can be put in Jordan normal form, which by the comments above, amounts to the existence of strings.

A has an eigenvector $\mathbf{v}$ with value λ. Since $A_\lambda\mathbf{v} = \mathbf{0}$, we have $r \stackrel{\text{def}}{=} \dim \operatorname{Ker} A_\lambda > 0$. By the Rank+Nullity Theorem (or directly, since the row reduced form of A_λ has r free variables there must be $n - r$ pivots) we have $\dim \operatorname{Range} A_\lambda = n - r < n$. Call $W = \operatorname{Range} A_\lambda$.

Step 1 $A_\lambda(W) \subseteq W$ so A_λ induces a transformation $T: W \to W$. Since $\dim(W) < n$, the matrix of T is of smaller size than n so by induction there are strings:

$$\begin{array}{ccccccc}
\mathbf{w}_{1,n_1} & \xrightarrow{A_{\lambda_1}} & \cdots & \xrightarrow{A_{\lambda_1}} & \mathbf{w}_{1,1} & \xrightarrow{A_{\lambda_1}} & \mathbf{0} \\
 & & \vdots & & \vdots & & \vdots \\
\mathbf{w}_{k,n_k} & \xrightarrow{A_{\lambda_k}} & \cdots & \xrightarrow{A_{\lambda_k}} & \mathbf{w}_{k,1} & \xrightarrow{A_{\lambda_k}} & \mathbf{0}
\end{array}$$

where the $\mathbf{w}_{i,j}$'s form a basis for W — here we used the fact that $(A_\lambda)_{\mu_i} = A_{\lambda+\mu_i} \stackrel{\text{def}}{=} A_{\lambda_i}$.

Step 2 Let $q = \dim(W \cap \operatorname{Ker} A_\lambda)$. Since $\mathbf{w}_{j,1} \in \operatorname{Ker} A_{\lambda_j}$, q of the above strings are A_λ strings, say the first q: $\lambda_j = \lambda$ for $1 \le j \le q$. At the other end of these strings, $\mathbf{w}_{j,n_j} \in W = \operatorname{Range} A_\lambda$ so there are $\mathbf{y}_j$ with $\mathbf{y}_j \xrightarrow{A_\lambda} \mathbf{w}_{j,n_j}$ for $1 \le j \le q$.

Step 3 Since $\operatorname{Ker} A_\lambda$ is r dimensional and meets W on a q dimensional subspace, some $r-q$ dimensional subspace Z of $\operatorname{Ker} A_\lambda$ meets W only at $\mathbf{0}$. Let $\mathbf{z}_1, \ldots, \mathbf{z}_{r-q}$ be a basis for Z.
This gives $q + (n-r) + (r-q) = n$ vectors in strings:

$$\begin{array}{ccccccccc}
\mathbf{y}_1 & \xrightarrow{A_\lambda} & \mathbf{w}_{1,n_1} & \xrightarrow{A_\lambda} & \cdots & \xrightarrow{A_\lambda} & \mathbf{w}_{1,1} & \xrightarrow{A_\lambda} & \mathbf{0} \\
 & & & & \vdots & & \vdots & & \vdots \\
\mathbf{y}_q & \xrightarrow{A_\lambda} & \mathbf{w}_{q,n_q} & \xrightarrow{A_\lambda} & \cdots & \xrightarrow{A_\lambda} & \mathbf{w}_{q,1} & \xrightarrow{A_\lambda} & \mathbf{0} \\
 & & \mathbf{w}_{q+1,n_{q+1}} & \xrightarrow{A_{\lambda_{q+1}}} & \cdots & \xrightarrow{A_{\lambda_{q+1}}} & \mathbf{w}_{q+1,1} & \xrightarrow{A_{\lambda_{q+1}}} & \mathbf{0} \\
 & & & & \vdots & & \vdots & & \vdots \\
 & & \mathbf{w}_{k,n_k} & \xrightarrow{A_{\lambda_k}} & \cdots & \xrightarrow{A_{\lambda_k}} & \mathbf{w}_{k,1} & \xrightarrow{A_{\lambda_k}} & \mathbf{0} \\
 & & & & & & \mathbf{z}_1 & \xrightarrow{A_\lambda} & \mathbf{0} \\
 & & & & & & \vdots & & \vdots \\
 & & & & & & \mathbf{z}_{r-q} & \xrightarrow{A_\lambda} & \mathbf{0}
\end{array}$$

It suffices to show they are linearly independent, so assume

$$\sum_i a_i \mathbf{y}_i + \sum_{i,j} b_{ij} \mathbf{w}_{i,j} + \sum_i c_i \mathbf{z}_i = \mathbf{0}.$$

Applying A_λ gives a linear combination, L, in $\mathbf{w}_{i,j}$'s as one can see by referring to the strings above. Using $A_{\lambda_r} \mathbf{w}_{s,r} = \mathbf{w}_{s,r-1}$ together with $A_\lambda = A_{\lambda_r} + (\lambda_r - \lambda) I$ shows $A_\lambda \mathbf{w}_{s,r} = \mathbf{w}_{s,r-1} + (\lambda_r - \lambda)\mathbf{w}_{s,r}$, hence the coefficient of the $\mathbf{w}_{j,n_j}$ for $1 \le j \le q$ in linear combination L is a_j. By linear independence of the $\mathbf{w}_{i,j}$'s we obtain $a_j = 0$. So

$$\sum_{i,j} b_{ij} \mathbf{w}_{i,j} + \sum_i c_i \mathbf{z}_i = \mathbf{0}.$$

But $\sum_{i,j} b_{ij} \mathbf{w}_{i,j} = \mathbf{0}$ and $\sum_i c_i \mathbf{z}_i = \mathbf{0}$ since W and Z meet only at $\mathbf{0}$. By linear independence in W and Z, $b_{ij} = 0$ and $c_i = 0$. ∎

Computing the Jordan Canonical Form

Let A be an n by n square matrix. If its characteristic equation $\chi_A(t) = 0$ has a repeated root then A may not be diagonalizable, so we need the Jordan Canonical Form. Suppose λ is an eigenvalue of A, with multiplicity r as a root of $\chi_A(t) = 0$. The the vector $\boldsymbol{v}$ is an eigenvector with eigenvalue λ if $A\boldsymbol{v} = \lambda\boldsymbol{v}$ or equivalently

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

The trouble is that this equation may have fewer then r linearly independent solutions for $\boldsymbol{v}$. So we generalize and say that $\boldsymbol{v}$ is a *generalized eigenvector* with eigenvalue λ if

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

for some positive integer k. Now one can prove that there are exactly r linearly independent generalized eigenvectors. Finding the Jordan form is now a matter of sorting these generalized eigenvectors into an appropriate order.

To find the Jordan form carry out the following procedure for each eigenvalue λ of A. First solve $(A - \lambda I)\boldsymbol{v} = 0$, counting the number r_1 of linearly independent solutions. If $r_1 = r$ good, otherwise $r_1 < r$ and we must now solve $(A - \lambda I)^2\boldsymbol{v} = 0$. There will be r_2 linearly independent solutions where $r_2 > r_1$. If $r_2 = r$ good, otherwise solving $(A - \lambda I)^3\boldsymbol{v} = 0$ gives $r_3 > r_2$ linearly independent solutions, and so on. Eventually one gets $r_1 < r_2 < \cdots < r_{N-1} < r_N = r$. The number N is the size of the largest Jordan block associated to λ, and r_1 is the total number of Jordan blocks associated to λ. If we define $s_1 = r_1$, $s_2 = r_2 - r_1$, $s_3 = r_3 - r_2, \ldots, s_N = r_N - r_{N-1}$ then s_k is the number of Jordan blocks of size at least k by k associated to λ. Finally put $m_1 = s_1 - s_2$, $m_2 = s_2 - s_3, \ldots, m_{N-1} = s_{N-1} - s_N$ and $m_N = s_N$. Then m_k is the number of k by k Jordan blocks associated to λ. Once we've done this for all eigenvalues then we've got the Jordan form!

To find P such that $J = P^{-1}AP$ is the Jordan form then we need to work a bit harder. We do the following for each eigenvalue λ. First find the Jordan block sizes associated to λ by the above process. Put them in decreasing order $N_1 \geq N_2 \geq N_3 \geq \cdots \geq N_k$. Now find a vector $\boldsymbol{v}_{1,1}$ such that $(A-\lambda I)^{N_1}\boldsymbol{v}_{1,1} = 0$ but $(A-\lambda I)^{N_1-1}\boldsymbol{v}_{1,1} \neq 0$. Define $\boldsymbol{v}_{1,2} = (A-\lambda I)\boldsymbol{v}_{1,1}$, $\boldsymbol{v}_{1,3} = (A - \lambda I)\boldsymbol{v}_{1,2}$, and so on until we get $\boldsymbol{v}_{1,N_1}$. We can't go further as $(A-\lambda I)\boldsymbol{v}_{1,N_1} = 0$. If we only have one block we're OK, otherwise we can find a vector $\boldsymbol{v}_{2,1}$ such that $(A - \lambda I)^{N_2}\boldsymbol{v}_{2,1} = 0$, $(A - \lambda I)^{N_2-1}\boldsymbol{v}_{2,1} \neq 0$ and (this

is important!) $\boldsymbol{v}_{2,1}$ is not linearly dependent on $\boldsymbol{v}_{1,1}, \ldots, \boldsymbol{v}_{1,N_1}$. Define $\boldsymbol{v}_{2,2} = (A - \lambda I)\boldsymbol{v}_{2,1}$ etc., until we get to $\boldsymbol{v}_{2,N_2}$. If $k = 2$ this is the end, if not then choose $\boldsymbol{v}_{3,1}$ with $(A - \lambda I)^{N_3}\boldsymbol{v}_{3,1} = 0$, $(A - \lambda I)^{N_3 - 1}\boldsymbol{v}_{3,1} \neq 0$ and $\boldsymbol{v}_{3,1}$ not linearly dependent on $\boldsymbol{v}_{1,1}, \ldots, \boldsymbol{v}_{1,N_1}, \boldsymbol{v}_{2,1}, \ldots \boldsymbol{v}_{2,N_2}$. Keep going! Eventually we get r linearly independent vectors $\boldsymbol{v}_{1,1}, \boldsymbol{v}_{1,2}, \ldots, \boldsymbol{v}_{k,N_k}$. Let

$$P_\lambda = (\boldsymbol{v}_{k,N_k} \ \cdots \ \boldsymbol{v}_{1,1})$$

be the n by r matrix whose columns are these vectors in **reverse** order. Once we've done this for all eigenvalues λ stick the matrices P_λ together horizontally to get an n by n matrix P. Then P will be non-singular, and $P^{-1}AP = J$, the Jordan form.

A worked example

To illustrate this method, I give a reasonably sized example (6 by 6) which I hope will make things clear, and I hope is safely too big come up on any exam! I have used MAPLE in the computations; only a truly hardy soul would try this one by hand!

Let

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

The characteristic polynomial of this matrix is

$$\chi_A(t) = t^6 + 3t^5 - 10t^3 - 15t^2 - 9t - 2 = (t+1)^5(t-2)$$

and so its eigenvalues are -1 with multiplicity 5, and 2 with multiplicity 1. I'll deal with $\lambda = -1$ first. We first solve $(A + I)\boldsymbol{v} = 0$. The matrix

$$A + I = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & -7 & 4 & -3 & 1 & -3 \\ -3 & 13 & -7 & 6 & 2 & 9 \\ -2 & 14 & -7 & 5 & 2 & 10 \\ 1 & -18 & 11 & -11 & 3 & -6 \\ -1 & 19 & -11 & 10 & -2 & 8 \end{pmatrix}$$

has REF

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 1 & 3/2 \\ 0 & 0 & 1 & 0 & 2 & 3/2 \\ 0 & 0 & 0 & 1 & 0 & -1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence $(A+I)\boldsymbol{v}$ has 2 linearly independent solutions, i.e., $r_1 = 2$. As $r_1 < r = 5$ then we must solve $(A+I)^2\boldsymbol{v} = 0$. Now

$$(A+I)^2 = \begin{pmatrix} 1 & -1 & 0 & 1 & -2 & -3 \\ -2 & -16 & 9 & -11 & 4 & -3 \\ -1 & 37 & -18 & 17 & 2 & 21 \\ 1 & 35 & -18 & 19 & -2 & 15 \\ -1 & -53 & 27 & -28 & 2 & -24 \\ 2 & 52 & -27 & 29 & -4 & 21 \end{pmatrix}$$

whose REF is

$$\begin{pmatrix} 1 & 0 & -1/2 & 3/2 & -2 & -5/2 \\ 0 & 1 & -1/2 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The system $(A+I)^2\boldsymbol{v}$ has $r_2 = 4$ linearly independent solutions. As $r_2 < r$, then we now consider $(A+I)^3\boldsymbol{v}$. Now

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -54 & 27 & -27 & 0 & -27 \\ 0 & 108 & -54 & 54 & 0 & 54 \\ 0 & 108 & -54 & 54 & 0 & 54 \\ 0 & -162 & 81 & -81 & 0 & -81 \\ 0 & 162 & -81 & 81 & 0 & 81 \end{pmatrix}$$

and it's easy to see (!) that the REF of this matrix is

$$\begin{pmatrix} 0 & 1 & -1/2 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence $(A + I)^3\boldsymbol{v} = 0$ has $r_3 = 5$ linearly independent solutions, and as $r_3 = r$ we conclude this part of the proceedings. We calculate $s_1 = r_1 = 2$, $s_2 = r_2 - r_1 = 2$ and $s_3 = r_3 - r_2 = 1$; also $m_3 = s_3 = 1$, $m_2 = s_2 - s_3 = 1$ and $m_1 = s_1 - s_2 = 0$. Hence, associated to $\lambda = -1$, there is a 2 by 2 and a 3 by 3 Jordan block. As the other eigenvalue $\lambda = 2$ has multiplicity 1, then there's just a 1 by 1 Jordan block associated to $\lambda = 2$. Hence the Jordan canonical form of A is J=

$$\begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Let's compute the matrix P. We've already done most of the work for $\lambda = -1$. The Jordan blocks have sizes $N_1 = 3$ and $N_2 = 2$. We start by finding a vector $\boldsymbol{v}_{1,1}$ with $(A+I)^3\boldsymbol{v}_{1,1} = 0$ but $(A+I)^2\boldsymbol{v}_{1,1} \neq 0$. Looking at the REFs of these matrices we see that we can choose

$$\boldsymbol{v}_{1,1} = (1 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0)^t.$$

Now

$$\boldsymbol{v}_{1,2} = (A+I)\boldsymbol{v}_{1,1} = (1 \;\; 0 \;\; -3 \;\; -2 \;\; 1 \;\; -1)^t$$

and

$$\boldsymbol{v}_{1,3} = (A+I)\boldsymbol{v}_{1,2} = (1 \;\; -2 \;\; -1 \;\; 1 \;\; -1 \;\; 2)^t.$$

(As a check one verifies $(A+I)\boldsymbol{v}_{1,3} = 0$.) The next block is 2 by 2, so one must find $\boldsymbol{v}_{2,1}$ with $(A+I)^2\boldsymbol{v}_{2,1} = 0$, $(A+I)\boldsymbol{v}_{2,1} \neq 0$, and such that $\boldsymbol{v}_{2,1}$ is not linearly dependent on $\boldsymbol{v}_{1,1}$, $\boldsymbol{v}_{1,2}$ and $\boldsymbol{v}_{1,3}$. The vector

$$\boldsymbol{v}_{2,1} = (1 \;\; 1 \;\; 2 \;\; 0 \;\; 0 \;\; 0)^t$$

fits the bill, and

$$\boldsymbol{v}_{2,2} = (A+I)\boldsymbol{v}_{2,1} = (1 \;\; 1 \;\; -4 \;\; -2 \;\; 5 \;\; -4)^t.$$

Again one checks that $(A+I)\boldsymbol{v}_{2,2} = 0$ The matrix P_{-1} is the 6 by 5 matrix with columns $\boldsymbol{v}_{2,2}$, $\boldsymbol{v}_{2,1}$, $\boldsymbol{v}_{1,3}$, $\boldsymbol{v}_{1,2}$ and $\boldsymbol{v}_{1,1}$ in that order and so

$$P_{-1} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -2 & 0 & 0 \\ -4 & 2 & -1 & -3 & 0 \\ -2 & 0 & 1 & -2 & 0 \\ 5 & 0 & -1 & 1 & 0 \\ -4 & 0 & 2 & -1 & 0 \end{pmatrix}.$$

One must now consider $\lambda = 2$. As this is a simple root, P_2 is just an eigenvector with eigenvalue 2. One such is

$$P_2 = (0 \;\; 1 \;\; -2 \;\; -2 \;\; 3 \;\; -3)^t$$

and sticking together P_{-1} and P_2 gives

$$P = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & -2 & 0 & 0 & 1 \\ -4 & 2 & -1 & -3 & 0 & -2 \\ -2 & 0 & 1 & -2 & 0 & -2 \\ 5 & 0 & -1 & 1 & 0 & 3 \\ -4 & 0 & 2 & -1 & 0 & -3 \end{pmatrix}.$$

One now checks that $P^{-1}AP = J$ as required!

RJC 25/1/95