Jordan 标准型简介 (根子空间分解)

袁懿

2025 年 12 月 12 日

1 补充内容

1.1 特征子空间

定义 1 (特征子空间)**.** 设 $\mathscr{A}$ 是 V 上的线性变换，λ_0 是 $\mathscr{A}$ 的一个特征值.
则 $V_{\lambda_0} := \{\alpha \in V | \mathscr{A}\alpha = \lambda_0\alpha\}$ 是 $\mathscr{A}$ 的属于特征值 λ_0 的**特征子空间**.

验证 V_{λ_0} 为 V 的子空间：

证明. **1. 非空**

由于 λ_0 是特征值，存在非零向量 $\alpha_0 \in V$ 使得 $\mathscr{A}\alpha_0 = \lambda_0\alpha_0$，因此 $\alpha_0 \in V_{\lambda_0}$. 另外，零向量 $\mathbf{0} \in V$ 满足 $\mathscr{A}\mathbf{0} = \mathbf{0} = \lambda_0\mathbf{0}$，所以 $\mathbf{0} \in V_{\lambda_0}$. 故 $V_{\lambda_0} \neq \emptyset$.

2. 加法封闭性

任取 $\alpha, \beta \in V_{\lambda_0}$，则

$$\mathscr{A}\alpha = \lambda_0\alpha, \quad \mathscr{A}\beta = \lambda_0\beta.$$

考虑 $\alpha + \beta$：

$$\mathscr{A}(\alpha + \beta) = \mathscr{A}\alpha + \mathscr{A}\beta = \lambda_0\alpha + \lambda_0\beta = \lambda_0(\alpha + \beta).$$

因此 $\alpha + \beta \in V_{\lambda_0}$.

3. 数乘封闭性

任取 $\alpha \in V_{\lambda_0}$ 和 $k \in \mathbb{K}$，则 $\mathscr{A}\alpha = \lambda_0\alpha$.

考虑 $k\alpha$：

$$\mathscr{A}(k\alpha) = k\mathscr{A}\alpha = k(\lambda_0\alpha) = \lambda_0(k\alpha).$$

因此 $k\alpha \in V_{\lambda_0}$. □

命题 1 (特征子空间的性质)**.** $\lambda_1 \neq \lambda_2, V_{\lambda_1} \cap V_{\lambda_2} = \{\mathbf{0}\}$.

证明. 任取向量 $\alpha \in V_{\lambda_1} \cap V_{\lambda_2}$，则根据特征子空间的定义有：

$$\mathscr{A}\alpha = \lambda_1\alpha \quad 且 \quad \mathscr{A}\alpha = \lambda_2\alpha.$$

由此可得：

$$\lambda_1\alpha = \lambda_2\alpha.$$

移项整理：

$$(\lambda_1 - \lambda_2)\alpha = \mathbf{0}.$$

由于 $\lambda_1 \neq \lambda_2$，所以 $\lambda_1 - \lambda_2 \neq 0$，从而可以推出：

$$\alpha = \mathbf{0}.$$

因此，$V_{\lambda_1} \cap V_{\lambda_2}$ 中只包含零向量，即：

$$V_{\lambda_1} \cap V_{\lambda_2} = \{\mathbf{0}\}.$$

□

推论 1. 设 $\lambda_1, \lambda_2, \ldots, \lambda_k$ 是 $\mathscr{A}$ 的 k 个互不相同的特征值，则它们的特征子空间的和是直和：

$$V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_k} = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}.$$

定义 2 (几何重数)**.** 设 $\mathscr{A}$ 是数域 $\mathbb{K}$ 上线性空间 V 的线性变换，λ_0 是 $\mathscr{A}$ 的一个特征值. 特征子空间 V_{λ_0} 的维数称为特征值 λ_0 的**几何重数**，记作 $\mathrm{gm}(\lambda_0)$，即：

$$\mathrm{gm}(\lambda_0) = \dim V_{\lambda_0}.$$

几何重数表示属于特征值 λ_0 的线性无关特征向量的最大个数.

定义 3 (代数重数)**.** 设 $\mathscr{A}$ 的特征多项式为 $f(\lambda) = \det(\mathscr{A} - \lambda I)$，特征值 λ_0 作为 $f(\lambda)$ 的根的重数称为 λ_0 的**代数重数**，记作 $\mathrm{am}(\lambda_0)$.

定理 1 (几何重数与代数重数的关系)**.** 设 $\mathscr{A}$ 是 n 维线性空间 V 上的线性变换，λ_0 是 $\mathscr{A}$ 的特征值，则

$$1 \leq \mathrm{gm}(\lambda_0) \leq \mathrm{am}(\lambda_0) \leq n.$$

证明. 设 $\mathrm{gm}(\lambda_0)=d$，取 V_{λ_0} 的一组基 $\alpha_1,\alpha_2,\dots,\alpha_d$。将其扩充为 V 的一组基：

$$\alpha_1,\dots,\alpha_d,\alpha_{d+1},\dots,\alpha_n.$$

线性变换 $\mathscr{A}$ 在这组基下的矩阵具有如下分块形式：

$$A=\begin{pmatrix}\lambda_0 I_d & B\\ 0 & C\end{pmatrix},$$

其中 I_d 是 d 阶单位矩阵. 计算特征多项式：

$$f(\lambda)=\det(A-\lambda I_n)=\det\begin{pmatrix}(\lambda_0-\lambda)I_d & B\\ 0 & C-\lambda I_{n-d}\end{pmatrix}=(\lambda_0-\lambda)^d\cdot\det(C-\lambda I_{n-d}).$$

可见 $(\lambda_0-\lambda)^d$ 整除 $f(\lambda)$，因此 $\mathrm{am}(\lambda_0)\ge d=\mathrm{gm}(\lambda_0)$. 另外，由于 λ_0 是特征值，至少有一个特征向量，故 $\mathrm{gm}(\lambda_0)\ge 1$, 显然代数重数不超过维数 n. □

毫无疑问，我们希望 $\mathscr{A}$ 在某组基下有对角形式的标准型，有吗？很遗憾，并没有. 但是对一些特殊的 $\mathscr{A}$ 可以做到，具体来说是下面这个定理：

定理 2 (可对角化的充要条件)**.** *设 $\mathscr{A}$ 是 n 维线性空间 V 上的线性变换，$\lambda_1,\lambda_2,\dots,\lambda_k$ 是 $\mathscr{A}$ 的全部互异特征值，则下列条件等价：*
(1) $\mathscr{A}$ 可对角化（即存在 V 的一组基，使得 $\mathscr{A}$ 在这组基下的矩阵为对角矩阵）.
(2) V 可以分解为特征子空间的直和：

$$V=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_k}.$$

(3) 所有特征值的几何重数之和等于 n：

$$\sum_{i=1}^{k}\mathrm{gm}(\lambda_i)=n.$$

(4) 每个特征值的几何重数等于其代数重数：

$$\mathrm{gm}(\lambda_i)=\mathrm{am}(\lambda_i),\quad i=1,2,\dots,k.$$

证明. (1) $\Rightarrow$ (2): 若 $\mathscr{A}$ 可对角化，则存在由特征向量组成的基, 每个基向量属于某个 V_{λ_i}. 将这些基向量按特征值分类，属于 λ_i 的那些向量组成 V_{λ_i} 的一组基, 因此 V 是这些特征子空间的直和.
(2) $\Rightarrow$ (3): 由直和分解，$n=\dim V=\sum_{i=1}^{k}\dim V_{\lambda_i}=\sum_{i=1}^{k}\mathrm{gm}(\lambda_i)$.

(3) $\Rightarrow$ (4): 由定理 1，$\mathrm{gm}(\lambda_i) \le \mathrm{am}(\lambda_i)$，且 $\sum_{i=1}^{k}\mathrm{am}(\lambda_i) = n$（因为特征多项式是 n 次多项式），若 $\sum_{i=1}^{k}\mathrm{gm}(\lambda_i) = n$，则必须有 $\mathrm{gm}(\lambda_i) = \mathrm{am}(\lambda_i)$ 对每个 i 成立.

(4) $\Rightarrow$ (1): 在每个特征子空间 V_{λ_i} 中取一组基,这些基向量合并起来构成 V 的一组基(因为 $\sum_{i=1}^{k}\mathrm{gm}(\lambda_i) = \sum_{i=1}^{k}\mathrm{am}(\lambda_i) = n$，且不同特征值的特征子空间交为零)，在这组基下，$\mathscr{A}$ 的矩阵为对角矩阵，对角线元素为相应的特征值. □

1.2 根子空间

那对于几何重数小于代数重数的矩阵，我们能不能寻找到一个 V 的直和分解，让它在某组基下有简单的形式？接下来这一节会给出答案.

定义 4 (根子空间)**.** 设 $\mathscr{A}$ 是数域 $\mathbb{K}$ 上线性空间 V 的线性变换，λ_0 是 $\mathscr{A}$ 的一个特征值，定义 λ_0 的**根子空间**为：

$$W_{\lambda_0} := \{\alpha \in V \mid \exists m \in \mathbb{N}^*, (\mathscr{A} - \lambda_0 I)^m \alpha = 0\}.$$

同样的，验证 W_{λ_0} 为 V 的子空间:

证明. 它显然包含特征子空间所以非空, 下证它对加法数乘封闭:

设 $\alpha, \beta \in W_{\lambda_0}$，则存在正整数 m, n 使得:

$$(\mathscr{A} - \lambda_0 I)^m \alpha = 0, \quad (\mathscr{A} - \lambda_0 I)^n \beta = 0.$$

取 $p = \max\{m, n\}$，则:

$$(\mathscr{A} - \lambda_0 I)^p(\alpha + \beta) = (\mathscr{A} - \lambda_0 I)^p \alpha + (\mathscr{A} - \lambda_0 I)^p \beta = 0 + 0 = 0.$$

故 $\alpha + \beta \in W_{\lambda_0}$.

对任意 $k \in \mathbb{K}$，$\alpha \in W_{\lambda_0}$：

$$(\mathscr{A} - \lambda_0 I)^m (k\alpha) = k(\mathscr{A} - \lambda_0 I)^m \alpha = 0.$$

故 $k\alpha \in W_{\lambda_0}$, 因此 W_{λ_0} 是子空间. □

定理 3 (根子空间的维数等于代数重数)**.** 设 $\mathscr{A}$ 是复数域 $\mathbb{C}$ 上 n 维线性空间 V 的线性变换，λ_0 是 $\mathscr{A}$ 的特征值，其代数重数为 d，则根子空间 W_{λ_0} 的维数等于 d，即：

$$\dim W_{\lambda_0} = \mathrm{am}(\lambda_0) = d.$$

证明. 老师在课上（或许）证明过 $\mathscr{A}$ 在适当的基下的矩阵 A 是上三角形矩阵，并且可以要求 $\mathscr{A}$ 的特征值在 A 的主对角线上按如下顺序排列：先将 n_i 个 λ_i 全部排完，再将其余特征值按任意顺序排列,$\mathcal{A}$ 可以写成分块形式:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ O & A_{22} \end{pmatrix}$$

其中

$$A_{11} = \begin{pmatrix} \lambda_i & * & \cdots & * \\ & \lambda_i & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_i \end{pmatrix}_{n_i \times n_i}$$

是主对角线元全为 λ_i 的 n_i 阶方阵，A_{22} 是主对角元等于其余特征值 $\lambda_j (j \neq i)$ 的 $n - n_i$ 阶上三角形矩阵, 对正整数 $k \geq n_i$，

$$(A - \lambda_i I)^k = \begin{pmatrix} (A_{11} - \lambda_i I_{(n_i)})^k & * \\ 0 & (A_{22} - \lambda_i I_{(n-n_i)})^k \end{pmatrix}$$

其中 $(A_{11} - \lambda_i I_{(n_i)})^k = 0$；而 $(A_{22} - \lambda_i I_{(n-n_i)})^k$ 是上三角形矩阵且对角元 $(\lambda_j - \lambda_i)^k$ 全不为 0，因而是 $n - n_i$ 阶可逆方阵.

因此，对正整数 $k \geq n_i$，有 $\operatorname{rank}(A - \lambda_i I)^k = n - n_i$，$n$ 元线性方程组 $(A - \lambda_i I)^k X = 0$ 的解空间 $\operatorname{Ker}(A - \lambda_i I)^k$ 的维数等于 n_i，从而 $\dim \operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^k = n_i$.

对任意正整数 $k_1 \leq k_2$ 和 $\beta \in V$，

$$(\mathscr{A} - \lambda_i \mathscr{I})^{k_1}(\beta) = 0 \Rightarrow (\mathscr{A} - \lambda_i \mathscr{I})^{k_2}(\beta) = (\mathscr{A} - \lambda_i \mathscr{I})^{k_2 - k_1}(\mathscr{A} - \lambda_i \mathscr{I})^{k_1}(\beta) = 0$$

这说明

$$\operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^{k_1} \subseteq \operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^{k_2}.$$

特别, 当 $k \geq n_i$ 时, 由 $\operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^{n_i} \subseteq \operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^k$ 及 $\dim \operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^k = n_i = \dim \operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^{n_i}$ 知

$$\operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^k = \operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^{n_i}.$$

对 $\mathscr{A}$ 的属于特征值 λ_i 的每个根向量 β, 存在正整数 k 使 $(\mathscr{A} - \lambda_i \mathscr{I})^k(\beta) = 0$, 即 $\beta \in \operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^k$. 当 $k \leq n_i$ 时, $\beta \in \operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^k \subseteq \operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^{n_i}$; 当 $k \geq n_i$ 时, $\beta \in \operatorname{Ker}(\mathcal{A} - \lambda_i \mathscr{I})^k = \operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^{n_i}$. 这说明了 $\mathscr{A}$ 的属于特征值 λ_i 的所有根向量都含于 n_i 维子空间 $\operatorname{Ker}(\mathscr{A} - \lambda_i \mathscr{I})^{n_i}$.

反过来，$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{n_i}$ 中的非零向量 β 都满足条件 $(\mathscr{A}-\lambda_i\mathscr{I})^{n_i}(\beta)=0$，都是 $\mathscr{A}$ 的属于特征值 λ_i 的根向量.

因此，$\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{n_i}$ 是由 $\mathscr{A}$ 的属于特征值 λ_i 的全体根向量与零向量共同组成的子空间，维数为 n_i. □

定理 4 (全空间可分解为根子空间)**.** 设 $\lambda_1,\cdots,\lambda_t$ 是 n 维复线性空间 V 的线性变换 $\mathscr{A}$ 的全部不同的特征值，则

$$V=W_{\lambda_1}\oplus\cdots\oplus W_{\lambda_t}.$$

证明. 先证明各根子空间 $W_{\lambda_i}\,(1\le i\le t)$ 的和是直和，为此需要证明：对任意一组 $\beta_i\in W_{\lambda_i}\,(1\le i\le t)$，若

$$\beta_1+\beta_2+\cdots+\beta_t=0,$$

则必有 $\beta_1=\beta_2=\cdots=\beta_t=0$.

对每个 i，令 $m_i=n_i$（即特征值 λ_i 的代数重数），考虑多项式

$$f_i(\lambda)=\prod_{j\neq i}(\lambda-\lambda_j)^{m_j}.$$

注意到当 $j\neq i$ 时，$(\lambda_i-\lambda_j)^{m_j}\neq 0$，故 $f_i(\lambda_i)\neq 0$, 另一方面，对任意 $j\neq i$，由于 $\beta_j\in W_{\lambda_j}=\mathrm{Ker}(\mathscr{A}-\lambda_j\mathscr{I})^{m_j}$，故

$$(\mathscr{A}-\lambda_j\mathscr{I})^{m_j}\beta_j=0.$$

于是

$$f_i(\mathscr{A})\beta_j=\prod_{\substack{k\neq i\\k\neq j}}(\mathscr{A}-\lambda_k\mathscr{I})^{m_k}\cdot(\mathscr{A}-\lambda_j\mathscr{I})^{m_j}\beta_j=0,$$

即当 $j\neq i$ 时，$f_i(\mathscr{A})\beta_j=0$, 而对 β_i，由于 $f_i(\lambda_i)\neq 0$，且 $\beta_i\in W_{\lambda_i}$，可证 $f_i(\mathscr{A})\beta_i\neq 0$ 除非 $\beta_i=0$.

现在假设 $\beta_1+\cdots+\beta_t=0$, 对每个固定的 i，用 $f_i(\mathscr{A})$ 作用于此等式两边：

$$f_i(\mathscr{A})(\beta_1+\cdots+\beta_t)=0.$$

由上述讨论，$f_i(\mathscr{A})\beta_j=0$（当 $j\neq i$），故得

$$f_i(\mathscr{A})\beta_i=0.$$

但 $f_i(\lambda)$ 与 $(\lambda-\lambda_i)^{m_i}$ 互素（因为 $f_i(\lambda_i)\neq 0$），故存在多项式 $u(\lambda),v(\lambda)$ 使得

$$u(\lambda)f_i(\lambda)+v(\lambda)(\lambda-\lambda_i)^{m_i}=1.$$

代入 $\lambda=\mathscr{A}$，并作用到 β_i 上：

$$u(\mathscr{A})f_i(\mathscr{A})\beta_i+v(\mathscr{A})(\mathscr{A}-\lambda_i\mathscr{I})^{m_i}\beta_i=\beta_i.$$

由于 $\beta_i\in W_{\lambda_i}=\mathrm{Ker}(\mathscr{A}-\lambda_i\mathscr{I})^{m_i}$，第二项为零；又 $f_i(\mathscr{A})\beta_i=0$，故第一项也为零，于是 $\beta_i=0$。由 i 的任意性，即证得直和.

而我们知道 $dimV=n=\sum_{i=1}^{t}n_i=\sum_{i=1}^{t}dim(W_{\lambda_i})$.

故 $V=W_{\lambda_1}\oplus\cdots\oplus W_{\lambda_t}$. □

我们自然地会去思考, 通过这样的直和分解把 V 的基分类，$\mathscr{A}$ 在这组基下的矩阵表示如何?

定理 5 (根子空间分解的矩阵表示)**.** 设 $\lambda_1,\cdots,\lambda_t$ 是 n 维复线性空间 V 上线性变换 $\mathscr{A}$ 的全部不同特征值，W_{λ_i} 是 $\mathscr{A}$ 属于 λ_i 的根子空间，且 $V=W_{\lambda_1}\oplus\cdots\oplus W_{\lambda_t}$。取 W_{λ_i} 的一组基 $\alpha_{i1},\cdots,\alpha_{in_i}$，将它们合并为 V 的一组基：

$$\alpha_{11},\cdots,\alpha_{1n_1},\alpha_{21},\cdots,\alpha_{2n_2},\cdots,\alpha_{t1},\cdots,\alpha_{tn_t},$$

则 $\mathscr{A}$ 在这组基下的矩阵为准对角阵

$$A=\begin{pmatrix}A_1&&&\\&A_2&&\\&&\ddots&\\&&&A_t\end{pmatrix},$$

其中 A_i 是 $\mathscr{A}$ 在 W_{λ_i} 上的限制 $\mathscr{A}|_{W_{\lambda_i}}$ 在基 $\alpha_{i1},\cdots,\alpha_{in_i}$ 下的矩阵，且 A_i 的特征值均为 λ_i.

证明. 由根子空间分解 $V=W_{\lambda_1}\oplus\cdots\oplus W_{\lambda_t}$ 可知，每个 W_{λ_i} 都是 $\mathscr{A}$ 的不变子空间，即对任意 $\alpha\in W_{\lambda_i}$，有 $\mathscr{A}(\alpha)\in W_{\lambda_i}$.

考虑 $\mathscr{A}$ 在 W_{λ_i} 上的限制 $\mathscr{A}|_{W_{\lambda_i}}$，取 W_{λ_i} 的一组基 $\alpha_{i1},\cdots,\alpha_{in_i}$. 设

$$\mathscr{A}(\alpha_{ij})=a_{1j}^{(i)}\alpha_{i1}+a_{2j}^{(i)}\alpha_{i2}+\cdots+a_{n_ij}^{(i)}\alpha_{in_i},\quad j=1,\cdots,n_i.$$

记 $A_i=(a_{kj}^{(i)})_{n_i\times n_i}$，则 $\mathscr{A}|_{W_{\lambda_i}}$ 在基 $\alpha_{i1},\cdots,\alpha_{in_i}$ 下的矩阵就是 A_i.

由于 W_{λ_i} 中每个向量都是 $\mathscr{A}$ 的属于特征值 λ_i 的根向量，即存在正整数 m 使得 $(\mathscr{A}-\lambda_i\mathscr{I})^m\alpha=0$，特

别地，$\mathscr{A}|_{W_{\lambda_i}}$ 的特征值均为 λ_i，故 A_i 的特征值均为 λ_i.
现在，将各 W_{λ_i} 的基合并为 V 的一组基，按定理中顺序排列, 由于 W_{λ_i} 是 $\mathscr{A}$ 的不变子空间，且各 W_{λ_i} 的和是直和，所以 $\mathscr{A}$ 在这组基下的矩阵具有准对角形式，对角块依次为 $A_1, A_2, \cdots, A_t$，而非对角块均为零. 即

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_t \end{pmatrix}.$$

□

这已经得到了比较简单的形式, 接下来我们要做的就是把根子空间进一步分解，使得 A_i 有比较简单的形式，而这个问题抽象为:
V 为全空间，$\mathscr{B}$ 为 V 上的映射，$\mathscr{B}^n = O$, 我们令 $ker\mathscr{B}^i = V_i$，则有自然的包含关系：

$$\{\mathbf{0}\} = V_0 \subseteq V_1 \subseteq ... \subseteq V_n = V$$

能否找到一组基使得 $\mathscr{B}$ 在其上有简单的形式？我们将在下一节解决这个问题.

1.3 根子空间分解

1.3.1 基本思路

由于 $\mathscr{B}$ 是幂零的，其特征值全为 0, 我们要构造的基应当由若干个 $\mathscr{B}$-循环链组成，每个循环链形如：

$$v, \mathscr{B}v, \mathscr{B}^2 v, \dots, \mathscr{B}^{k-1} v, \quad \text{满足} \mathscr{B}^k v = 0,$$

且这些循环链彼此不相交（即它们生成的循环子空间的直和等于 V）. 在每个循环子空间上，$\mathscr{B}$ 的矩阵是一个幂零 Jordan 块：

$$J_k(0) = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}_{k \times k}.$$

1.3.2 构造步骤

记 $d_i = \dim V_i - \dim V_{i-1}$（$i = 1, \ldots, m$），则 $d_1 \geq d_2 \geq \cdots \geq d_m > 0$, 定义 $d_{m+1} = 0$. 下面递归构造向量组：

1. 对 $k = m, m-1, \ldots, 1$，归纳选取向量 $u_1^{(k)}, u_2^{(k)}, \ldots, u_{s_k}^{(k)}$，其中 $s_k = d_k - d_{k+1}$，满足：

   - $u_j^{(k)} \in V_k$；
   - 在商空间 V_k/V_{k-1} 中，向量组

     $$\{u_1^{(k)} + V_{k-1}, \ldots, u_{s_k}^{(k)} + V_{k-1}\}$$

     与来自 V_{k+1} 的向量

     $$\{\mathscr{B}u_1^{(k+1)} + V_{k-1}, \ldots, \mathscr{B}u_{s_{k+1}}^{(k+1)} + V_{k-1}\}$$

     的并构成 V_k/V_{k-1} 的一组基.

2. 对每个 $u_j^{(k)}$，生成一个长度为 k 的 **Jordan 链**：

$$\left\{u_j^{(k)},\ \mathscr{B}u_j^{(k)},\ \ldots,\ \mathscr{B}^{k-1}u_j^{(k)}\right\}.$$

所有这样的 Jordan 链中的向量构成 V 的一个向量组，记作 $\mathcal{J}$.

引理 1 (线性无关性)**.** *向量组 $\mathcal{J}$ 线性无关.*

证明. 假设存在系数 $c_{j,k,\ell} \in \mathbb{F}$（$k = 1, \ldots, m$, $j = 1, \ldots, s_k$, $\ell = 0, \ldots, k-1$）使得

$$\sum_{k=1}^{m} \sum_{j=1}^{s_k} \sum_{\ell=0}^{k-1} c_{j,k,\ell} \mathscr{B}^\ell u_j^{(k)} = 0. \tag{*}$$

我们对指数 $r = m, m-1, \ldots, 0$ 归纳证明所有系数为零. 设 $r = m$, 在 $(*)$ 式两边作用 $\mathscr{B}^{m-1}$。注意到若 $k < m$, 则 $\ell \leq k-1 \leq m-2$, 故 $\mathscr{B}^{m-1}\mathscr{B}^\ell u_j^{(k)} = \mathscr{B}^{m-1+\ell}u_j^{(k)} = 0$(因为 $u_j^{(k)} \in V_k$, 而 $m-1+\ell \geq m-1 \geq k$), 于是仅 $k = m$ 且 $\ell = 0$ 的项留下：

$$\sum_{j=1}^{s_m} c_{j,m,0} \mathscr{B}^{m-1} u_j^{(m)} = 0.$$

由构造，$\{\mathscr{B}^{m-1}u_j^{(m)}\}_{j=1}^{s_m}$ 线性无关（它们属于 V_1 且像在 V_m/V_{m-1} 中无关），故 $c_{j,m,0} = 0$ 对所有 j.

现设对某个 r（$1 \leq r \leq m$），我们已经证明：对一切 $k > r$ 及一切 j, ℓ 有 $c_{j,k,\ell} = 0$，且对 $k = r$，当 $\ell < m - r$ 时 $c_{j,r,\ell} = 0$, 在 $(*)$ 式两边作用 $\mathscr{B}^{r-1}$。此时：

- 对 $k > r$：系数已为零，贡献为零.

- 对 $k=r$: 仅 $\ell=m-r$ 的项可能非零(因为 $\ell<m-r$ 的系数为零), 且 $\mathscr{B}^{r-1}\mathscr{B}^{m-r}u_j^{(r)}=\mathscr{B}^{m-1}u_j^{(r)}$。但注意 $u_j^{(r)}\in V_r$，而 $\mathscr{B}^{m-1}u_j^{(r)}=0$（因 $m-1\geq r$），故这些项实际为零.

- 对 $k<r$：$\ell\leq k-1\leq r-2$，故 $\mathscr{B}^{r-1}\mathscr{B}^{\ell}u_j^{(k)}=\mathscr{B}^{r-1+\ell}u_j^{(k)}=0$（因为 $r-1+\ell\geq r-1\geq k$）.

于是得到

$$\sum_{j=1}^{s_r} c_{j,r,m-r}\mathscr{B}^{m-1}u_j^{(r)}=0.$$

但由构造，$\{\mathscr{B}^{m-1}u_j^{(r)}\}_{j=1}^{s_r}$ 线性无关（它们属于 V_1 且是 $\{u_j^{(r)}\}$ 在 $\mathscr{B}^{m-1}$ 下的像，而 $\{u_j^{(r)}\}$ 在商空间中的像与来自 V_{r+1} 的像合并后线性无关），因此 $c_{j,r,m-r}=0$.
逐步下降 r，最终得到所有系数为零, 故 $\mathcal{J}$ 线性无关. □

由于 $\mathcal{J}$ 中向量个数为

$$\sum_{k=1}^{m} ks_k=\sum_{k=1}^{m} k(d_k-d_{k+1})=\sum_{k=1}^{m} d_k=\dim V,$$

因此 $\mathcal{J}$ 是 V 的一组基，称为 $\mathscr{B}$ 的 **Jordan 基**. 将基 $\mathcal{J}$ 按如下顺序排列：对每个 k 和 j，将链

$$u_j^{(k)},\ \mathscr{B}u_j^{(k)},\ \ldots,\ \mathscr{B}^{k-1}u_j^{(k)}$$

中的向量顺序排列。在这组基下，变换 $\mathscr{B}$ 的矩阵为分块对角形，每个块对应一个 Jordan 链。具体地，对长度为 k 的链，对应的矩阵块为 $k\times k$ 的 **幂零 Jordan 块**:

$$J_k(0)=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&1\\0&0&0&\cdots&0\end{pmatrix}_{k\times k}.$$

于是 $\mathscr{B}$ 的矩阵为

$$J=\operatorname{diag}\big(J_{k_1}(0),J_{k_2}(0),\ldots,J_{k_t}(0)\big),$$

其中 $k_1,k_2,\ldots,k_t$ 是所有 Jordan 链的长度（按重数排列），且满足 $k_1+k_2+\cdots+k_t=\dim V$。此矩阵称为 $\mathscr{B}$ 的 **Jordan 标准型**（幂零情形）.

定理 6 (幂零变换的 Jordan 分解)**.** *设 $\mathscr{B}:V\to V$ 为幂零线性变换，则存在 V 的一组基，使得 $\mathscr{B}$ 在这组基下的矩阵为 Jordan 标准型 J，且 J 在相似意义下唯一（不计 Jordan 块的排列顺序）.*

于是我们可以得到最终的结果:

1.4 一般矩阵的 Jordan 标准型

设 A 是复数域 $\mathbb{C}$ 上的 n 阶矩阵，其特征多项式为

$$p_A(\lambda) = \det(\lambda I - A) = \prod_{i=1}^{k} (\lambda - \lambda_i)^{m_i},$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_k$ 是 A 的互异特征值，m_i 是 λ_i 的代数重数.

定理 7 (Jordan 标准型)**.** *存在可逆矩阵 $P \in \mathrm{GL}_n(\mathbb{C})$，使得*

$$P^{-1}AP = J = \mathrm{diag}(J_1, J_2, \dots, J_t),$$

其中每个 J_i 是形如

$$J(\lambda, r) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}_{r \times r}$$

的 Jordan 块，λ 是 A 的特征值，r 是 Jordan 块的阶数。矩阵 J 称为 A 的 Jordan 标准型，在相似意义下唯一（不计 Jordan 块的排列顺序）。

Rmk: 1. 特征值 λ_i 的几何重数（即特征子空间 $ker(A - \lambda_i I)$ 的维数）等于 λ_i 对应的 Jordan 块个数. (这是因为 $ker(A - \lambda_i I)$ 中每个元是 Jordan 链的“终点”)

2. 每个 Jordan 块 $J(\lambda_i, r)$ 对应一条长度为 r 的 Jordan 链:

$$\mathbf{v}_1, (A - \lambda_i I)\mathbf{v}_1, \dots, (A - \lambda_i I)^{r-1}\mathbf{v}_1,$$

其中 $(A - \lambda_i I)^r \mathbf{v}_1 = \mathbf{0}$ 但 $(A - \lambda_i I)^{r-1}\mathbf{v}_1 \neq \mathbf{0}$.

3. 矩阵 A 可对角化当且仅当所有 Jordan 块的阶数均为 1.

1.5 例子

设矩阵

$$A=\begin{pmatrix}1&1&1&1&1&1\\0&2&0&1&2&3\\0&0&2&0&1&-1\\0&0&0&2&1&2\\0&0&0&0&2&0\\0&0&0&0&0&2\end{pmatrix}$$

求其 Jordan 标准形 J 及可逆矩阵 P 使得 $P^{-1}AP=J$.

步骤 1：求特征多项式与特征值

A 的特征多项式为:

$$p_A(\lambda)=(\lambda-1)(\lambda-2)^5$$

故特征值为 1（1 重）、2（5 重）.

步骤 2：处理特征值 1 的根子空间设 $V=F^{6\times 1}$，$\mathscr{A}:X\mapsto AX$ 是 V 上的线性变换.

特征值 1 的根子空间 $W_1=\ker(\mathscr{A}-\mathscr{I})$（1 维），解方程组 $(A-I)X=0$，得基础解:

$$X_1=(1,0,0,0,0,0)^{\mathrm{T}}$$

此为 W_1 的基，满足 $AX_1=X_1$.

步骤 3：处理特征值 2 的根子空间

特征值 2 的根子空间 $W_2=\ker(\mathscr{A}-2\mathscr{I})^3$（由 $\operatorname{rank}(A-2I)^3=1$ 得 $\dim W_2=5$）.

记 $\mathscr{B}=(\mathscr{A}-2\mathscr{I})|_{W_2}$（幂零变换），需确定 $\ker\mathscr{B},\ker\mathscr{B}^2,\ker\mathscr{B}^3$ 的基:

1. 解 $(A-2I)Y=0$（求 $\ker\mathscr{B}$ 的基）：得基 $S_1=\{Y_1,Y_2\}=\{(1,1,0,0,0,0)^{\mathrm{T}},(1,0,1,0,0,0)^{\mathrm{T}}\}$.

2. 解 $(A-2I)^2Y=0$（扩充为 $\ker\mathscr{B}^2$ 的基）：将 S_1 扩充为 $M_2=S_1\cup S_2$，其中 $S_2=\{Y_3,Y_4\}=\{(0,0,0,1,0,0)^{\mathrm{T}},(0,0,3,0,2,-1)^{\mathrm{T}}\}$. 3. 解 $(A-2I)^3Y=0$（扩充为 $\ker\mathscr{B}^3$ 的基）：将 M_2 扩充为 $M_3=M_2\cup S_3$，其中 $S_3=\{Y_5\}=\{(0,0,0,0,1,-2)^{\mathrm{T}}\}$.

步骤 4：构造 Jordan 链

计算幂零变换作用下的向量:

$$\mathscr{B}^2(Y_5)=(A-2I)^2(Y_5)=(-3,-3,0,0,0,0)^{\mathrm{T}}$$

$$\mathscr{B}(Y_5)=(A-2I)(Y_5)=(1,1,0,0,0,0)^{\mathrm{T}}$$

$$\mathscr{B}(Y_4) = (A-2I)(Y_4) = (4,1,3,0,0,0)^{\mathrm{T}}$$

W_2 的 Jordan 链为：

长度 3 的链：$\mathscr{B}^2(Y_5) \to \mathscr{B}(Y_5) \to Y_5$

长度 2 的链：$\mathscr{B}(Y_4) \to Y_4$

长度 1 的链：Y_2

步骤 5：确定 Jordan 标准形与可逆矩阵将 W_1 的基 $\{X_1\}$ 与 W_2 的 Jordan 基合并，得 V 的基 $N = \{X_1, \mathscr{B}^2(Y_5), \mathscr{B}(Y_5), Y_5, \mathscr{B}(Y_4), Y_4\}$.

对应的 Jordan 标准形为：

$$J = \operatorname{diag}(1,\ J_3(2),\ J_2(2)) = \begin{pmatrix} 1 & & & & & \\ & 2 & 1 & & & \\ & & 2 & 1 & & \\ & & & 2 & & \\ & & & & 2 & 1 \\ & & & & & 2 \end{pmatrix}$$

可逆矩阵 P 以 N 中向量为列：

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

满足 $P^{-1}AP = J$.