矩阵对角化和 Jordan 标准形

曾焰

版本 1.0.1，最后修改于 2014-04-14

摘要

蓝以中 [1, 2] 关于矩阵对角化和 Jordan 标准形的相关内容的摘要笔记。

目录

1 线性变换的特征值与特征向量

本节内容参见蓝以中 [1, page 296-312]。

1.1 特征值与特征向量的计算法

设 V 是数域 K 上的 n 维线性空间，$\mathcal{A}$ 是 V 内一个线性变换。我们需要解决下面两个问题：

1) 决定 K 内所有 $\mathcal{A}$ 的特征值 λ;

2) 对于属于特征值 λ 的特征子空间 V_λ，找出它的一组基。

我们把**计算线性变换 $\mathcal{A}$ 的特征值和特征向量的步骤**归纳如下：

1) 在 V 中给定一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，求 $\mathcal{A}$ 在这组基下的矩阵 A。

2) 计算特征多项式 $f(\lambda) = |\lambda E - A|$。

3) 求 $f(\lambda) = 0$ 的属于数域 K 的那些根

$$\lambda_1, \lambda_2, \cdots, \lambda_s.$$

4) 对每个 λ_i $(i = 1, 2, \cdots, s)$ 求齐次线性方程组

$$(\lambda_i E - A)X = 0$$

的一个基础解系（例如可用矩阵消元法）。这个齐次线性方程组具体写出来就是

$$\begin{bmatrix} \lambda_i - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda_i - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda_i - a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0.$$

其中的 λ_i 是在步骤 3) 中求出的，是已知数，不是未知量。

5) 以步骤 4) 中求出的基础解系为坐标写出 V 中一个向量组，它就是 $\mathcal{A}$ 的属于特征值 λ_i 的特征子空间 V_{λ_i} 的一组基。

值得注意的是，两个 n 阶方阵 A，B 乘积一般不可交换：$AB \neq BA$，但我们可以证明它们的特征多项式却是一样。实际上，我们可以得到更一般的结果。

命题 1. 设 A 是数域 K 上 $n \times m$ 矩阵，B 是 K 上 $m \times n$ 矩阵，则

$$\lambda^m |\lambda E_n - AB| = \lambda^n |\lambda E_m - BA|,$$

特别地，当 $m = n$ 时 $|\lambda E - AB| = |\lambda E - BA|$。

证明. 借助于分块矩阵运算的技巧。 □

1.2 具有对角形矩阵的线性变换

定理 1. 数域 K 上 n 维线性空间 V 内一个线性变换 A 的矩阵可对角化的充分必要条件是，A 有 n 个线性无关的特征向量。

证明. 根据定义即可得证，且易见用这 n 个线性无关的特征向量的坐标做列向量的矩阵就是对角化 A 的矩阵。 □

命题 2. 线性变换 A 的属于不同特征值的特征向量线性无关。

证明. 可用数学归纳法得证。 □

定理 2. 设 A 是数域 K 上 n 维线性空间 V 的线性变换，$\lambda_1, \lambda_2, \cdots, \lambda_k$ 是 A 的全部互不相同的特征值。则 A 的矩阵可对角化的充分必要条件是

$$V = V_{\lambda_1} \bigoplus V_{\lambda_2} \bigoplus \cdots \bigoplus V_{\lambda_k}$$

在 A 的矩阵可对角化的情况下，在每个 V_{λ_i} 中任取一组基，合并后即为 V 的一组基，在该组基下 A 的矩阵为对角矩阵。

注 1. 这是用空间分解的语言对定理*1*的另一种表述。

注 2. 因为 $V = V_{\lambda_1} \bigoplus V_{\lambda_2} \bigoplus \cdots \bigoplus V_{\lambda_k}$ 的充分必要条件是 $\sum_{i=1}^{k} \dim V_{\lambda_i} = \dim V$，所以定理*2*的条件是否得到满足可通过计算特征值和特征子空间的一组基立刻得到解决。

1.3 不变子空间

命题 3. 设 A 是数域 K 上 n 维线性空间 V 内的一个线性变换。在 V 内存在一组基 $\varepsilon_1, \cdots, \varepsilon_n$，使 A 在这组基下的矩阵成准对角形的充分必要条件是，V 可以分解成 A 的不变子空间 $M_1, M_2, \cdots, M_s$ 的直和

$$V = M_1 \bigoplus M_2 \bigoplus \cdots \bigoplus M_s$$

命题 4. 设 A 是数域 K 上 n 维线性空间 V 内的一个线性变换。如果 A 的矩阵可对角化，则对 A 的任意不变子空间 M, $A|_M$ 的矩阵也可对角化。

证明. 考虑证明 $V = (M \cap V_{\lambda_1}) \bigoplus (M \cap V_{\lambda_2}) \bigoplus \cdots \bigoplus (M \cap V_{\lambda_k})$。 □

2 实对称矩阵的对角化

本节内容参见蓝以中 [2, page 26-38]。

命题 5. 实对称矩阵 A 的特征多项式在复数域内的根都是实数。

证明. 设 λ 是 A 的特征多项式在复数域内的一个根，并设 $x \in \mathbb{C}^n \setminus \{0\}$ 满足 $Ax = \lambda x$。通过取共轭转置，我们有

$$\bar{x}' A x = \bar{x}' \lambda x = \lambda \|x\|^2$$

及

$$\bar{x}'\bar{A}'x = \bar{\lambda}\bar{x}'x = \bar{\lambda}\|x\|^2$$

比较两式并由 $\bar{A}' = A$ 及 $\|x\| \neq 0$, 我们可推知 $\lambda = \bar{\lambda}$。 □

推论 1. 欧氏空间 V 内任一对称变换 A 至少有一个特征值。

注 3. 从定理*3*的证明来看，这一推论是实对称矩阵能够对角化的直接原因。但更高的观点可以揭示定理*3* 成立的本质原因。根据谱分解定理（可参见 *Lax[4, page 70-73]* 或蓝以中 *[3, page 349]*），给定复数域上的一个 $n \times n$ 方阵 A，$\mathbb{C}^n$ 中的每一个向量都可分解成 A 的特征值和广义特征值的和。换言之，$\mathbb{C}^n$ 可分解为 A 的特征子空间和广义特征子空间的直和：

$$\mathbb{C}^n = N_{d_1}(\lambda_1)\bigoplus\cdots\bigoplus N_{d_k}(\lambda_k)$$

这里的 $N_d(\lambda) = \ker[(\lambda I - A)^d]$ 并且 d_i *(*$i = 1, \cdots, k$*)* 是使得 $N_d(\lambda_i) = N_{d+1}(\lambda_i)$ 的最小正整数，称作**特征值** λ **的指标**（易证对于有限维线性空间而言，特征值的指标总是存在的，且 $N_{d_i} = N_{d_i+1} = N_{d_i+2} = \cdots$）。从谱分解定理这一"高观点"来看，关键性的问题就变成："为什么对于对称变换这个特例，$d_i = 1$ *(*$i = 1, \cdots, k$*)*?" 事实上，假定存在某个 $d_i > 1$，则对于任何 $x \in N_2(\lambda_i)\setminus N_1(\lambda_i)$，我们有

$$(\lambda_i I - A)x \neq 0,\ (\lambda_i I - A)^2 x = 0$$

但后一条件意味着 $((\lambda_i I - A)x, (\lambda_i I - A)x) = ((\lambda_i I - A)^2 x, x) = 0$，也即 $(\lambda_i I - A)x = 0$，矛盾。此观察即直接说明了对称变换是如何迫使特征值的指标为 *1* 的。

命题 6. 设 A 是欧氏空间 V 内的一个对称变换，则 A 的对应于不同特征值的特征向量互相正交。

定理 3. 设 A 是 n 维欧氏空间 V 内的一个对称变换，则在 V 内存在一组标准正交基，使 A 在此组基下的矩阵成对角形。

证明. 对 V 的维数 n 作数学归纳法。设 λ_1 是 A 的一个特征值，η_1 是对应于 λ_1 的单位特征向量：

$$A\eta_1 = \lambda_1\eta_1, (\eta_1, \eta_1) = 1.$$

由 η_1 生成的线性空间 $M = L(\eta_1)$ 是 A 的一维不变子空间。易见 M 的正交补 $M^\perp$ 是 A 的 $n-1$ 维不变子空间，且 $A|_{M^\perp}$ 仍为对称变换。按归纳假设，在 $M^\perp$ 内存在一组标准正交基 $\eta_2, \cdots, \eta_n$，使

$$A\eta_i = \lambda_i\eta_i \ \ (i = 1, 2, \cdots, n).$$

易知 $\eta_1, \eta_2, \cdots, \eta_n$ 为 V 的一组标准正交基，使得 A 在此组基下的矩阵成对角形。 □

注 4. 正如注*3*所述，数学归纳法的运用掩盖了 A 的矩阵能够对角化的本质原因：对称变换的特征值的指标必然为 *1*。

推论 2. 设 A 是一个 n 阶实对称矩阵，则存在 n 阶正交矩阵 T，使

$$T^{-1}AT = T'AT = D$$

为对角矩阵。

用正交矩阵化实对称矩阵成对角形的算法。给定 n 阶实对称矩阵 A，我们已知存在 n 阶正交矩阵 T，使 $T^{-1}AT = T'AT = D$ 成对角形（推论2）。设 A 的特征多项式的全部互不相同的根是 $\lambda_1, \lambda_2, \cdots, \lambda_k$，它们是 A 的全部特征值（根据命题5，它们都是实数）。由定理2，A 可对角化说明

$$\mathbb{R}^n = V_{\lambda_1} \bigoplus V_{\lambda_2} \bigoplus \cdots \bigoplus V_{\lambda_k}$$

再由命题6，V_{λ_i} 与 V_{λ_j} $(i \neq j)$ 的向量互相正交。因此，只要在每个 V_{λ_i} 中取一组标准正交基（全由特征值为 λ_i 的特征向量组成），合并后为 $\mathbb{R}^n$ 内 n 个两两正交的单位向量组，即为 $\mathbb{R}^n$ 的一组标准正交基，在此基下 A 的矩阵成对角形 D，而 T 即为从 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ $(\varepsilon_i = (0, \cdots, 0, \overbrace{1}^{i}, 0, \cdots, 0)^T)$ 到此组基的过渡矩阵，T 的列向量即为此组基在 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标。根据这些分析，我们把T **和** D **的具体计算方法**归纳为以下几个步骤：

1) 计算特征多项式 $f(\lambda) = |\lambda I - A|$，并求出它的全部根（两两不同者）$\lambda_1$，$\lambda_2$，$\cdots$，$\lambda_k$。

2) 对每个 λ_i，求齐次线性方程组 $(\lambda_i I - A)X = 0$ 的一个基础解系 X_{i1}，X_{i2}，$\cdots$，X_{it_i}，它们即为解空间 M_{λ_i} 的一组基。

3) 在欧氏空间 $\mathbb{R}^n$ 内将 $X_{i1}, X_{i2}, \cdots, X_{it_i}$ 正交化：

$$\begin{aligned}
Y_{i1} &= X_{i1}, \\
Y_{i2} &= X_{i2} - \frac{(X_{i2}, Y_{i1})}{(Y_{i1}, Y_{i1})} Y_{i1}, \\
Y_{i3} &= X_{i3} - \frac{(X_{i3}, Y_{i1})}{(Y_{i1}, Y_{i1})} Y_{i1} - \frac{(X_{i3}, Y_{i2})}{(Y_{i2}, Y_{i2})} Y_{i2}, \\
\cdots &\cdots \cdots
\end{aligned}$$

再把所得的 $Y_{i1}, Y_{i2}, \cdots, Y_{it_i}$ 在 $\mathbb{R}^n$ 内单位化，得 V_{λ_i} 的一组标准正交基 $Z_{i1}, Z_{i2}, \cdots, Z_{it_i}$。所寻求的正交矩阵 T 应为 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\mathbb{R}^n$ 的标准正交基

$$Z_{11}, Z_{12}, \cdots, Z_{1t_1}, Z_{21}, Z_{22}, \cdots, Z_{2t_2}, \cdots, Z_{k1}, Z_{k2}, \cdots, Z_{kt_k}$$

的过渡矩阵：

$$\begin{aligned}
& (Z_{11}, Z_{12}, \cdots, Z_{1t_1}, Z_{21}, Z_{22}, \cdots, Z_{2t_2}, \cdots, Z_{k1}, Z_{k2}, \cdots, Z_{kt_k}) \\
= & (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) T \\
= & T
\end{aligned}$$

所以只要把上述向量（写成竖列形式）作为列向量依次排列，即得正交矩阵 T，而此时相应的对角矩阵 D 应为

$$D = \mathrm{diag}\{\overbrace{\lambda_1, \cdots, \lambda_1}^{t_1}, \overbrace{\lambda_2, \cdots, \lambda_2}^{t_2}, \cdots, \overbrace{\lambda_k, \cdots, \lambda_k}^{t_k}\}$$

3 矩阵的 Jordan 标准形

本节内容参见蓝以中 [2, page 72-99]。

3.1 幂零线性变换的 Jordan 标准形

根据谱分解定理，

$$\mathbb{C}^n = N_{d_1}(\lambda_1)\bigoplus N_{d_2}(\lambda_2)\bigoplus\cdots\bigoplus N_{d_k}(\lambda_k)$$

这里的 $N_d(\lambda)=\ker[(\lambda I-A)^d]$ 并且 d_i $(i=1,\cdots,k)$ 是使得 $N_d(\lambda_i)=N_{d+1}(\lambda_i)$ 的最小正整数。从此结果出发，一个自然而然的想法是研究幂零线性变换的最简单的矩阵表示。

命题 7. *设 A 是数域 K 上 n 维线性空间 V 内的一个幂零线性变换，则 A 的特征多项式为 $f(\lambda)=\lambda^n$，从而 A 有唯一的特征值 $\lambda_0=0$.*

给定一个幂零线性变换 A,取 V 中任意非零向量 α,则存在最小正整数 k,使得 $A^{k-1}\alpha\neq 0$,但 $A^k\alpha=0$。易证向量组 α, $A\alpha$, $\cdots$, $A^{k-1}\alpha$ 线性无关。由此向量组张成的线性空间 $I(\alpha)=L(\alpha,A\alpha,\cdots,A^{k-1}\alpha)$ 称作由 α 生成的 A 的**循环不变子空间**。在 $I(\alpha)$ 的基 $A^{k-1}\alpha$, $A^{k-2}\alpha$, $\cdots$, $A\alpha$, α（这组基称作**循环基**）下，$A|_{I(\alpha)}$ 的矩阵为

$$\begin{aligned}
& A(A^{k-1}\alpha,A^{k-2}\alpha,\cdots,A\alpha,\alpha)\\
= & (A^k\alpha,A^{k-1}\alpha,\cdots,A^2\alpha,A\alpha)\\
= & (A^{k-1}\alpha,A^{k-2}\alpha,\cdots,A\alpha,\alpha)J
\end{aligned}$$

这里

$$J=\begin{pmatrix} 0 & 1 & & & \\ & 0 & \ddots & \bigcirc & \\ & & \ddots & \ddots & \\ & \bigcirc & & \ddots & 1 \\ & & & & 0 \end{pmatrix}$$

反过来说，如果 M 是 A 的一个不变子空间，且 M 内存在一组基使得 $A|_M$ 在此组基下的矩阵为 J，则存在 α 使得 $M=I(\alpha)$。

由于 V 的维数有限，上述制造 A 的“互不相同”的循环不变子空间的过程会在进行若干次后穷尽全空间。于是全空间 V 可分解成若干个 A 的循环不变子空间的直和，在每一个循环不变子空间上，A 的矩阵都形如 J，而在全空间 V 上，A 的矩阵是以形如 J 的矩阵为对角元素的准对角矩阵。

于是我们可定义形如 J 的矩阵为 **Jordan 块**，并定义以 Jordan 块为对角线元素的准对角矩阵为 **Jordan 形矩阵**。

命题 8. *设 A 是数域 K 上 n 维线性空间 V 内的一个幂零线性变换，则在 V 内存在一组基，使 A 在该组基下的矩阵成 Jordan 形矩阵的充分必要条件是 V 可分解为 A 的循环不变子空间的直和：*

$$V=I(\alpha_1)\bigoplus I(\alpha_2)\bigoplus\cdots\bigoplus I(\alpha_s)$$

证明. 若 A 的矩阵在一组基下为 Jordan 形矩阵，则命题3说明 V 可分解为不变子空间的直和。再由本节开始时的讨论，可知这些不变子空间必是循环不变子空间，必要性得证。充分性是显而易见的。 □

定理 4. 设 A 是数域 K 上 n 维线性空间 V 内一幂零线性变换，则在 V 内存在一组基，使在该组基下 A 的矩阵成 *Jordan* 形矩阵。

证明. 只需证 V 可分解为 A 的循环不变子空间的直和即可，而这可以通过商空间的手段降低维数，通过数学归纳法得证。

当 $n=1$ 时，A 的任一特征向量 ε 为 V 的一组基，且 $A\varepsilon=0$（幂零线性变换的特征值必为 0），于是 A 在 ε 下的矩阵为 (0)，自然是 Jordan 形矩阵。

设命题对维数小于 n 的线性空间已成立，则当 $\dim V=n$ 时，不妨设 $A\neq 0$，则 A 在商空间 $\bar{V}=V/V_{\lambda_0}$（λ_0 为 A 的唯一特征值 0，V_{λ_0} 为其特征子空间）内的诱导线性变换仍为 $\bar{V}$ 内幂零线性变换。按归纳假设，我们有

$$\bar{V}=I(\bar{\alpha}_1)\bigoplus I(\bar{\alpha}_2)\bigoplus\cdots\bigoplus I(\bar{\alpha}_s)$$

这里 $I(\bar{\alpha}_i)$ 是 A 在商空间 $\bar{V}$ 内的诱导线性变换的一个 k_i 维循环不变子空间（$i=1,2,\cdots,s$）。易证 $I(\alpha_i)=L(\alpha_i,A\alpha_i,\cdots,A^{k_i}\alpha_i)$ 是 A 在 V 内的 k_i+1 维循环不变子空间，且 $A^{k_i}\alpha_i\in V_{\lambda_0}$。还易证 $A^{k_1}\alpha_1$, $A^{k_2}\alpha_2$，$\cdots$，$A^{k_s}\alpha_s$ 为 V_{λ_0} 内线性无关向量组，故可将其扩充为 V_{λ_0} 内的一组基：

$$A^{k_1}\alpha_1,A^{k_2}\alpha_2,\cdots,A^{k_s}\alpha_s,\beta_1,\beta_2,\cdots,\beta_t$$

最后可证明

$$V=I(\alpha_1)\bigoplus I(\alpha_2)\bigoplus\cdots\bigoplus I(\alpha_s)\bigoplus I(\beta_1)\bigoplus\cdots\bigoplus I(\beta_t)$$

□

注 5. 此证明过程实际给出了在 V 中找一组基，使 A 在该组基下的矩阵成 *Jordan* 形的具体计算方法。这是 *Jordan* 形理论的其他证明方法所未能给出的。细节参见蓝以中 *[3, page 336，340-341]*。

3.2 一般线性变换的 Jordan 标准形

命题 9. 设 A 是数域 K 上 n 维线性空间 V 内的一个线性变换。如果存在 $\lambda_0\in K$，使 $A-\lambda_0 I$ 是一个幂零线性变换，则在 V 内存在一组基，使 A 在这组基下的矩阵成为如下的 *Jordan* 标准形：

$$J=\begin{pmatrix} J_1 & & & \bigcirc \\ & J_2 & & \\ & & \ddots & \\ \bigcirc & & & J_s \end{pmatrix},\ J_i=\begin{pmatrix} \lambda_0 & 1 & & & \\ & \lambda_0 & \ddots & \bigcirc & \\ & & \ddots & \ddots & \\ & \bigcirc & & \ddots & 1 \\ & & & & \lambda_0 \end{pmatrix}$$

证明. 对 $A-\lambda_0 I$ 直接应用上节关于幂零矩阵的结论即可。 □

下面设 A 为 V 中任一线性变换，又设 A 有一特征值 $\lambda_0\in K$。令 $B=A-\lambda_0 I$，定义 V 的两串子空间序列如下：

$$M_0=\{0\},\ M_i=\ker(B^i)\ (i=1,2,\cdots);$$
$$N_0=V,\ N_i=\mathrm{Im}(B^i)\ (i=1,2,\cdots).$$

我们有如下简单的事实：
1) M_i, N_i 间有如下包含关系：

$$\{0\} = M_0 \subseteq M_1 \subseteq M_2 \subseteq \cdots ;\ V = N_0 \supseteq N_1 \supseteq N_2 \supseteq \cdots .$$

2) $\dim M_i + \dim N_i = n$ （$i = 0, 1, 2, \cdots$）。
3) 存在一个最小正整数 k，使得

$$M_0 \subset M_1 \subset \cdots \subset M_k = M_{k+1} = M_{k+2} = \cdots ,$$

$$N_0 \supset N_1 \supset \cdots \supset N_k = N_{k+1} = N_{k+2} = \cdots .$$

4) 对上述的最小正整数 k，有 $V = M_k \bigoplus N_k$，且 M_k 和 N_k 均为 A 的不变子空间。
5) B 在 M_k 上的限制 $B|_{M_k}$ 为幂零线性变换，所以在 M_k 内存在一组基，使得 $A|_{M_k}$ 在这组基下的矩阵为 Jordan 标准形。

总结起来，我们有如下命题：

命题 10. 设 A 是 n 维线性空间 V 内的一个线性变换，λ_0 是 A 的一个特征值。令 $B = A - \lambda_0 I$，又设

$$M_i = \ker(B^i),\ N_i = Im(B^i),$$

这里 $i = 0, 1, 2, \cdots$。则存在正整数 k，使 $V = M_k \bigoplus N_k$，且 M_k, N_k 为 A, B 的不变子空间，$B|_{M_k}$ 为幂零线性变换，又有 $M_k = M_{k+1} = M_{k+2} = \cdots$，$N_k = N_{k+1} = N_{k+2} = \cdots$。

定理 5. 设 A 是数域 K 上 n 维线性空间 V 内的一个线性变换，其特征多项式 $f(\lambda)$ 的根全属于 K，则在 V 内存在一组基，使在该组基下 A 的矩阵成为如下的准对角形

$$J = \begin{pmatrix} J_1 & & & \bigcirc \\ & J_2 & & \\ & & \ddots & \\ \bigcirc & & & J_s \end{pmatrix},\ J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & \ddots & \bigcirc & \\ & & \ddots & \ddots & \\ & \bigcirc & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}$$

J 称为 A 的 **Jordan** 标准形。

证明. 沿用命题10的记号，取定 A 的一个特征值 λ_0，我们可将 V 分解成直和 $M_k \bigoplus N_k$。由 $B|_{M_k}$ 是幂零矩阵（这是关键），我们可在 M_k 内找到一组基，使得 $A|_{M_k}$ 在该组基下的矩阵是 Jordan 标准形。对 N_k 使用归纳假设（因其维数小于 n），可得 N_k 的一组基，使得 $A|_{N_k}$ 在该组基下的矩阵是 Jordan 标准形。将两组基合并起来即完成证明。 □

下述定理对于证明 Jordan 标准形的唯一性和计算 Jordan 标准形的具体步骤至关重要。

定理 6. 设 A 是数域 K 上 n 维线性空间 V 内的一个线性变换，其特征多项式 $f(\lambda)$ 的根全属于 K。又设 J 是 A 的任一 *Jordan* 标准形（J 的存在性已由定理5保证）。则对 A 的任一特征值 λ_0，以及

$B = A - \lambda_0 I,\ M_i = \ker(B^i)$ （$i = 0, 1, 2, \cdots$），J 中以 λ_0 为特征值且阶为 l 的 *Jordan* 块的个数为（$r(\cdot)$ 表示一个矩阵的秩）

$$2\dim M_l - \dim M_{l+1} - \dim M_{l-1} = r(B^{l+1}) + r(B^{l-1}) - 2r(B^l).$$

故除了主对角线上 *Jordan* 块的排列次序可以变化以外，*Jordan* 标准形是由 A 唯一决定的。

证明. 由线性变换的像与核的维数关系，我们有 $\dim M_l + r(B^l) = \dim V$ （$l = 0, 1, 2, \cdots$），故我们只需证明关于 $\dim M_k$ 的结论成立。我们设 A 的 Jordan 标准形 J 有如下形式

$$J = \begin{pmatrix} J_1 & & & \bigcirc \\ & J_2 & & \\ & & \ddots & \\ \bigcirc & & & J_s \end{pmatrix},\ J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & \ddots & \bigcirc & \\ & & \ddots & \ddots & \\ & \bigcirc & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}$$

其中 J_i 为 n_i 阶 Jordan 块（这里的 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 可以有相同的）。则

$$r(B^l) = r[(J - \lambda_0 I_{n\times n})^l)] = \sum_{i=1}^{s} r[(J_i - \lambda_0 I_{n_i\times n_i})^l].$$

因为

$$(J_i - \lambda_0 I_{n_i\times n_i})^l = \begin{pmatrix} \lambda_i - \lambda_0 & 1 & & & \\ & \lambda_i - \lambda_0 & \ddots & \bigcirc & \\ & & \ddots & \ddots & \\ & \bigcirc & & \ddots & 1 \\ & & & & \lambda_i - \lambda_0 \end{pmatrix}^l_{n_i\times n_i},$$

(i) 若 $\lambda_i \neq \lambda_0$，则 $r[(J_i - \lambda_0 I_{n_i\times n_i})^l] = n_i$;
(ii) 若 $\lambda_i = \lambda_0$，则利用幂零 Jordan 块的乘法性质有

$$r[(J_i - \lambda_0 I_{n_i\times n_i})^l] = \begin{cases} n_i - l, & l < n_i, \\ 0, & l \geq n_i. \end{cases}$$

因此有

$$r[(J_i - \lambda_0 I_{n_i\times n_i})^l] - r[(J_i - \lambda_0 I_{n_i\times n_i})^{l+1}] = \begin{cases} 0, & \lambda_i \neq \lambda_0, \\ 0, & \lambda_i = \lambda_0,\ l \geq n_i, \\ 1, & \lambda_i = \lambda_0,\ l < n_i. \end{cases}$$

因此，

$$
\begin{aligned}
& r(B^l) - r(B^{l+1}) \\
= & \sum_{i=1}^{s} \{r[(J_i - \lambda_0 I_{n_i\times n_i})^l] - r[(J_i - \lambda_0 I_{n_i\times n_i})^{l+1}]\} \\
= & \sum_{\lambda_i=\lambda_0,\, n_i>l} \{r[(J_i - \lambda_0 I_{n_i\times n_i})^l] - r[(J_i - \lambda_0 I_{n_i\times n_i})^{l+1}]\} \\
= & \sum_{\lambda_i=\lambda_0,\, n_i>l} 1 \\
= & J \text{ 中以 } \lambda_0 \text{ 为特征值而阶数 } \geq l+1 \text{ 的 Jordan 块的个数.}
\end{aligned}
$$

于是 J 中以 λ_0 为特征值的 l 阶 Jordan 块的个数是 $r(B^{l-1}) + r(B^{l+1}) - 2r(B^l)$。 □

注 6. *命题10中使得 $M_k = M_{k+1}$ 和 $N_k = N_{k+1}$ 成立的最小正整数 k，根据定义也是特征值 λ_0 的指标 d_0，也即，使得 $N_d(\lambda_0) = N_{d+1}(\lambda_0)$ 成立的最小正整数（此处 $N_d(\lambda_0) := \ker[(A-\lambda_0 I)^d] = M_d$）。$d_0$ 也能在计算 Jordan 标准形的过程中确定。事实上，由 $\dim M_k = \dim V - \dim N_k = \dim V - r(B^k)$ 易知，d_0 即是使得 $r(B^k) = r(B^{k+1})$ 的最小正整数。由此推断，它同时也是特征值 λ_0 的 Jordan 块的最高阶数，因为以 λ_0 为特征值的 l 阶 Jordan 块的个数在 $l > d_0$ 时为 $r(B^{l+1}) + r(B^{l-1}) - 2r(B^l) = 0$，在 $l = d_0$ 时为 $r(B^{d_0-1}) - r(B^{d_0}) > 0$。总结起来，我们有*

$$
\begin{aligned}
& \text{特征值 } \lambda_0 \text{ 的指标 } d_0 \\
= & \text{使得 } r[(A-\lambda_0 I)^k] = r[(A-\lambda_0 I)^{k+1}] \text{ 成立的最小正整数} \\
= & \lambda_0 \text{ 的 Jordan 块的最高阶数。}
\end{aligned}
$$

设在 n 维线性空间 V 内给定线性变换 $\mathcal{A}$，**可按如下步骤计算 $\mathcal{A}$ 的 Jordan 标准形**（假设其存在）：

1) 先求 $\mathcal{A}$ 在 V 的一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的矩阵 A;

2) 求出 A 的全部不同特征值 $\lambda_1, \lambda_2, \cdots, \lambda_s$（假设都属于数域 K）;

3) 对每个 λ_i，令 $B = A - \lambda_i I$，由公式

$$r(B^{l+1}) + r(B^{l-1}) - 2r(B^l)$$

计算出以 λ_i 为特征值，阶为 l 的 Jordan 块个数。为此，令 $l = 1, 2, \cdots$，逐次计算。从 A 的 Jordan 形 J 的特征多项式容易看出：以 λ_i 为特征值的 Jordan 块阶数之和等于特征值 λ_i 的重数，由此即可知道是否已经找出全部以 λ_i 为特征值的 Jordan 块；或者从 $r(B^l) - r(B^{l+1})$ 等于 J 中以 λ_i 为特征值而阶 $\geq l+1$ 的 Jordan 块的个数这一点作出判断。

4) 将所获得的 Jordan 块按任意次序排列成准对角形 J，即为所求。

至于如何找到一组基，使得 $\mathcal{A}$ 在该组基下的矩阵为 Jordan 标准形，参阅蓝以中 [3, page 336, 340-341, 349-351]。

3.3 最小多项式

命题 11. *给定数域 K 上的 Jordan 块*

$$J=\begin{pmatrix} \lambda_0 & 1 & & & \\ & \lambda_0 & \ddots & \bigcirc & \\ & & \ddots & \ddots & \\ & \bigcirc & & \ddots & 1 \\ & & & & \lambda_0 \end{pmatrix}_{n\times n}$$

又设 $g(x)$ 是 K 上一个 m 次多项式。则 $g(x)$ 是 J 的化零多项式的充分必要条件是 λ_0 是 $g(x)$ 的一个零点，且其重数 $\geq J$ 的阶 n。

证明. 只需注意若 $g(x)$ 在 $\mathbb{C}$ 内可分解为

$$g(x)=a_0(x-\mu_1)^{e_1}(x-\mu_2)^{e_2}\cdots(x-\mu_k)^{e_k},$$

其中 μ_1，μ_2，$\cdots$，μ_k 两两不同，则

$$g(J)=a_0(J-\mu_1)^{e_1}(J-\mu_2)^{e_2}\cdots(J-\mu_k)^{e_k}.$$

且

$$(J-\mu_iI)^{e_i}=\begin{pmatrix} \lambda_0-\mu_i & 1 & & & \\ & \lambda_0-\mu_i & \ddots & \bigcirc & \\ & & \ddots & \ddots & \\ & \bigcirc & & \ddots & 1 \\ & & & & \lambda_0-\mu_i \end{pmatrix}^{e_i}$$

□

定理 7 (Hamilton-Cayley 定理)**.** *设 A 是数域 K 上的 n 阶方阵，$f(\lambda)=|\lambda I-A|$ 为 A 的特征多项式，则 $f(A)=0$。*

证明. 选用 A 的 Jordan 标准形：

$$J=\begin{pmatrix} J_1 & & & \bigcirc \\ & J_2 & & \\ & & \ddots & \\ \bigcirc & & & J_s \end{pmatrix},\ J_i=\begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & \ddots & \bigcirc & \\ & & \ddots & \ddots & \\ & \bigcirc & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}$$

其中 J_i 为 n_i 阶 Jordan 块（这里的 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 可以有相同的）。则 $f(\lambda)=|\lambda I-J|=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$。因为 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 可有相同的，$f(\lambda)$ 的每个根 λ_i 的重数 $\geq$Jordan 块 J_i 的阶数 n_i。现在

$$f(J)=\mathrm{diag}\{f(J_1),f(J_2),\cdots,f(J_s)\},$$

则由本节开始时的命题可知 $f(J)=0$，从而 $f(A)=0$。 □

命题 12. 设 A 是数域 K 上的 n 阶方阵，A 的最小多项式 $\varphi(x)$ 与 A 看作 $\mathbb{C}$ 上的 n 阶方阵时的最小多项式 $\psi(x)$ 有相同次数。换言之，A 在 K 内的任一最小多项式也是 A 在 $\mathbb{C}$ 内的最小多项式。

定理 8 (最小多项式)**.** 设 A 是数域 K 上的 n 阶方阵。设 A 的特征多项式在 $\mathbb{C}$ 内全部互不相同的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_k$。$A$ 在 $\mathbb{C}$ 内的 *Jordan* 标准形 J 中以 λ_i 为特征值的 *Jordan* 块的最高阶数为 l_i，则 A（在 K 内）的最小多项式是唯一的，它就是

$$\varphi(x) = (x-\lambda_1)^{l_1}(x-\lambda_2)^{l_2}\cdots(x-\lambda_k)^{l_k}.$$

推论 3. 设 A 是数域 K 上的 n 阶方阵且 A 的特征多项式的根全属于 K，则 A 在 K 内相似于对角矩阵的充要条件是它的最小多项式没有重根。

证明. 从 A 的 Jordan 标准形直接看出。 □

参考文献

[1] 蓝以中：《高等代数简明教程（上册）》。北京：北京大学出版社，2002.8。

[2] 蓝以中：《高等代数简明教程（下册）》。北京：北京大学出版社，2002.8。

[3] 蓝以中：《高等代数学习指南》。北京：北京大学出版社，2008.7。

[4] P. Lax. *Linear algebra and its applications*, 2nd Edition, Wiley-Interscience, 2007.