从 Jordan 矩阵再探不变子空间

吴方班高代习题课

2022 年 5 月 23 日

1 Jordan 块的情况

设 $\sigma : V \longrightarrow V$ 为复 n 维线性空间上的线性变换。满足 σ 在一组基 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 下的矩阵为一个 Jordan 块矩阵，

$$J = \begin{pmatrix} \lambda_0 & 0 & \cdots & 0 & 0 \\ 1 & \lambda_0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \lambda_0 \end{pmatrix}$$

1.1 V 不能分解为非平凡不变子空间的直和

设线性变换 $\sigma : V \longrightarrow V$ 在一组基 $\alpha_1, \cdots, \alpha_r$ 下的矩阵为一个 Jordan 块 $J_{r\times r}$。则

- 包含 α_1 的 σ-子空间只有 V 本身；

- V 的任一 σ-子空间都包含 α_r；

- V 不能分解为两个非平凡 σ-子空间的直和。

1.2 可能的不变子空间

设 σ 的唯一特征值为 λ，特征向量为 $k\alpha_r, k \neq 0$。

$$\lambda E - J = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0 \end{pmatrix}$$

为幂零阵，$(\lambda E - J)^r = 0$。（每乘一个 $\lambda E - J$ 都会多出一个 0 列。）

所以

$$V_{\lambda^0} = \{X \in V \mid (\lambda E - J)^0 X = 0\} = \{0\}$$

$$V_{\lambda^1} = \{X \in V \mid (\lambda E - J) X = 0\} = L(\alpha_r)$$

$$V_{\lambda^2} = \{X \in V \mid (\lambda E - J)^2 X = 0\} = L(\alpha_{r-1}, \alpha_r)$$

$$\cdots$$

$$V_{\lambda^r} = \{X \in V \mid (\lambda E - J)^r X = 0\} = L(\alpha_1, \cdots, \alpha_{r-1}, \alpha_r)$$

$$= V_{\lambda^{r+1}} = V_{\lambda^{r+2}} = \cdots = V$$

都为 σ-子空间。（同时也是 $(\lambda\cdot id-\sigma)$-子空间。）

设 W 为 V 任意一个 σ-子空间，任取 W 中的一个基向量 $X=k_1\alpha_1+\cdots+k_r\alpha_r\in W$，设 X 的系数 $k_1,\cdots,k_r$ 中左起第一个非零系数为 $k_{i(X)}$，则 $L(\alpha_{i(X)},\cdots,\alpha_r)=V_{\lambda^{r+1-i(X)}}\subset W$。则 $W=\bigcup\limits_X V_{\lambda^{r+1-i(X)}}=V_{\lambda^{r+1-i_{min}}}$，其中 $i_{min}=\min\{i(X)\mid X\text{为基向量}\}$。

所以，上面 V_{λ^i} 给出了 V 的所有 σ-子空间，并满足

$$V_{\lambda^0}\subset V_{\lambda^1}\subset V_{\lambda^2}\subset\cdots V_{\lambda^{r-1}}\subset V_{\lambda^r}=V_{\lambda^{r+1}}=\cdots=V$$

1.3　特征子空间和根子空间

V_{λ_1} 为特征值 λ 的特征子空间，V_{λ_r} 为特征值 λ 的根子空间。

1.4　核和值域

当特征值 $\lambda\neq 0$ 时，线性变换 σ 为可逆变换，它的核和值域分别为：

$$\ker\sigma=\{0\}$$

$$\mathrm{im}\ \sigma=V$$

当特征值 $\lambda=0$ 时，则线性变换 σ 的核和值域分别为：

$$\ker\sigma=L(\alpha_r)=V_{\lambda^1}$$

$$\mathrm{im}\ \sigma=L(\alpha_2,\cdots,\alpha_r)=V_{\lambda^{r-1}}$$

1.5 $V = \ker\sigma \oplus \mathbf{im}\ \sigma$

由上知，$V = \ker\sigma \oplus \mathrm{im}\ \sigma$ 当且仅当 $\lambda \neq 0$ 或 $\lambda = 0$，$r = 1$，即 σ 为可逆变换或零变换。

2 唯一特征值对应多个 Jordan 块的情况

设 $\sigma : V \longrightarrow V$ 为复 n 维线性空间上的线性变换。取 V 的一组基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$，使得 σ 在这组基下的矩阵为一个 Jordan 矩阵，

$$J = \mathrm{diag}(J_1, \cdots, J_m)$$

其中 $J_1, \cdots, J_m$ 都是特征值 λ 对应的 Jordan 块构成的 Jordan 阵，m 为 λ 的几何重数，即 $\lambda = \lambda_0$ 的线性无关特征向量的个数。设 Jordan 块 J_j 的阶数为 r_j，这里不妨设 $1 \leq r_1 \leq r_2 \leq \cdots \leq r_m$，则 $\sum\limits_j r_j = n$ 为 λ 的代数重数。

2.1 V 的不变子空间的直和分解

V 可以分为 Jordan 块对应的 σ-子空间的直和，

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$$

其中 V_j 为 Jordan 块 J_j 在 J 所处的行（或列）对应的基向量，设为 $\{\alpha_1,\alpha_2,\cdots,\alpha_{r_j}\}\subset\{\epsilon_1,\cdots,\epsilon_n\}$, r_j 为 J_j 的阶数。则每个 V_j 都是 σ-子空间，这是由于

$$\begin{cases}\sigma(\alpha_1)=\lambda\alpha_1+\alpha_2\in V_j\\ \sigma(\alpha_2)=\lambda\alpha_2+\alpha_3\in V_j\\ \qquad\cdots\\ \sigma(\alpha_{r_j-1})=\lambda\alpha_{r_j-1}+\alpha_{r_j}\in V_j\\ \sigma(\alpha_{r_j})=\lambda\alpha_{r_j}\in V_j\end{cases}$$

2.2　可能的不变子空间

所以不同 V_j 的 $\sigma\mid_{V_j}$-子空间的直和为 σ-子空间。这样，我们可以找到了 V 的一些 σ-子空间，共有 $\prod_j(r_j+1)$ 个。但显然，不是 V 的全部不变子空间。例如，σ 是一个数乘变换，则任意子空间都是 σ-子空间。事实上，由 σ 的两个线性无关的特征向量的和生成的一维子空间是 σ-不变的，但不属于任何 V_j 的 $\sigma\mid_{V_j}$-子空间的直和。

2.3 特征子空间和根子空间

设 V_j 为 Jordan 块 J_j 在 J 所处的行（或列）对应的基向量，为 $\{\alpha_1^j, \alpha_2^j, \cdots, \alpha_{r_j}^j\} \subset \{\epsilon_1, \cdots, \epsilon_n\}$,

$$V_{\lambda^0} = \{X \in V \mid (\lambda E - J)^0 X = 0\} = \{0\}$$

$$V_{\lambda^1} = \{X \in V \mid (\lambda E - J) X = 0\} = L(\alpha_{r_1}^1, \alpha_{r_2}^2, \cdots, \alpha_{r_m}^m)$$

$$V_{\lambda^2} = \{X \in V \mid (\lambda E - J)^2 X = 0\} = L(\alpha_{r_1-1}^1, \alpha_{r_1}^1, \alpha_{r_2-1}^2, \alpha_{r_2}^2, \cdots, \alpha_{r_m-1}^m, \alpha_{r_m}^m)$$

$$\cdots$$

$$V_{\lambda^{r_m}} = \{X \in V \mid (\lambda E - J)^{r_m} X = 0\} = L(\alpha_1^1, \cdots, \alpha_{r_1}^1, \cdots, \alpha_1^m, \cdots, \alpha_{r_m}^m) = L(\epsilon_1, \cdots, \epsilon_n)$$

$$= V_{\lambda^{r_m+1}} = V_{\lambda^{r_m+2}} = \cdots = V$$

所以，

$$V_{\lambda^0} \subset V_{\lambda^1} \subset V_{\lambda^2} \subset \cdots V_{\lambda^{r_m-1}} \subset V_{\lambda^{r_m}} = V_{\lambda^{r_m+1}} = \cdots = V$$

其中 r_s 为 λ 对应阶数最多 Jordan 块的阶数。

λ 的特征子空间为 V_{λ^1}。

λ 的根子空间 $\{X \in V \mid (\lambda E - J)^n X = 0\} = V$。这是因为 $n = \sum_j r_j \geq r_s$。

2.4 核和值域

当特征值 $\lambda \neq 0$ 时，线性变换 σ 为可逆变换，它的核和值域分别为：

$$\ker \sigma = \{0\}$$

$$\mathrm{im}\ \sigma = V$$

当特征值 $\lambda = 0$ 时，则线性变换 σ 的核和值域分别为：

$$\ker \sigma = L(\alpha_{r_1}^1, \alpha_{r_2}^2, \cdots, \alpha_{r_m}^m) = V_{\lambda^1}$$

$$\mathrm{im}\ \sigma = L(\alpha_2^1, \cdots, \alpha_{r_1}^1, \cdots, \alpha_2^m, \cdots, \alpha_{r_m}^m) = V_{\lambda^{r_m - 1}}$$

2.5 $V = \ker \sigma \oplus \mathbf{im}\ \sigma$

由上知，$V = \ker \sigma \oplus \mathrm{im}\ \sigma$ 当且仅当 $\lambda \neq 0$ 或 $\lambda = 0$，$r_1 = \cdots = r_m = 1$，即 σ 为可逆变换或零变换。

3 一般 Jordan 矩阵的情况：多个特征值，每个特征值对应多个 Jordan 块

设 $\sigma : V \longrightarrow V$ 为复 n 维线性空间上的线性变换。取 V 的一组基 $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$，使得 σ 在这组基下的矩阵为一个 Jordan 矩阵，

$$\begin{aligned} J &= \operatorname{diag}(J_{11}, \cdots, J_{1m_1}, J_{21}, \cdots, J_{2m_2}, \cdots, J_{s1}, \cdots, J_{sm_s}) \\ &= \operatorname{diag}(A_1, A_2, \cdots, A_s) \end{aligned}$$

其中 $A_i = \operatorname{diag}(J_{i1}, \cdots, J_{im_i})$ 是特征值 $\lambda = \lambda_i$ 对应的 Jordan 块构成的 Jordan 阵，m_i 为 $\lambda = \lambda_i$ 的几何重数, A_i 的阶数是 $\lambda = \lambda_i$ 的代数重数 n_i。$\lambda_1, \lambda_2, \cdots, \lambda_s$ 互不相同。

3.1 V 的不变子空间的直和分解

V 可以分为 Jordan 块对应的 σ-子空间的直和，

$$V = V_{11} \oplus \cdots \oplus V_{1m_1} \oplus V_{21} \oplus \cdots \oplus V_{2m_2} \oplus \cdots \oplus V_{s1} \oplus \cdots \oplus V_{sm_s}$$

其中 V_{ij} 为 Jordan 块 J_{ij} 在 J 所处的行（或列）对应的基向量，设为 $\{\alpha_1, \alpha_2, \cdots, \alpha_{r_{ij}}\} \subset \{\epsilon_1, \cdots, \epsilon_n\}$, r_{ij} 为 J_{ij} 的阶数。则每个 V_{ij}

是 σ-子空间，这是由于

$$\begin{cases} \sigma(\alpha_1) = \lambda_i \alpha_1 + \alpha_2 \in V_{ij} \\ \sigma(\alpha_2) = \lambda_i \alpha_2 + \alpha_3 \in V_{ij} \\ \cdots \\ \sigma(\alpha_{r_{ij}-1}) = \lambda_i \alpha_{r_{ij}-1} + \alpha_{r_{ij}} \in V_{ij} \\ \sigma(\alpha_{r_{ij}}) = \lambda_i \alpha_{r_{ij}} \in V_{ij} \end{cases}$$

V 也可以分为特征值对应的 σ-子空间的直和，

$$V = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_s$$

其中 $\mathcal{V}_i = V_{i1} \oplus \cdots \oplus V_{im_1}$ 也是 σ-子空间，这是由于不变子空间的直和还是不变子空间。

3.2 特征子空间和根子空间

设 V_{ij} 为 Jordan 块 J_{ij} 在 J 所处的行（或列）对应的基向量，为 $\{\alpha_1^{i,j},\alpha_2^{i,j},\cdots,\alpha_{r_{ij}}^{i,j}\}\subset\{\epsilon_1,\cdots,\epsilon_n\}$，

$$V_{\lambda_i^0}=\{X\in V\mid(\lambda E-J)^0X=0\}=\{0\}$$

$$V_{\lambda_i^1}=\{X\in V\mid(\lambda E-J)X=0\}=L(\alpha_{r_{i1}}^{i,1},\cdots,\alpha_{r_{im_1}}^{i,m_i})$$

$$V_{\lambda_i^2}=\{X\in V\mid(\lambda E-J)^2X=0\}=L(\alpha_{r_{i1}-1}^{i,1},\alpha_{r_{i1}}^{i,1},\cdots,\alpha_{r_{im_1}-1}^{i,m_i},\alpha_{r_{im_1}}^{i,m_i})$$

$$\cdots$$

$$V_{\lambda_i^{r_{m_i}}}=\{X\in V\mid(\lambda E-J)^{r_{m_i}}X=0\}=L(\alpha_1^{i,1},\cdots,\alpha_{r_1}^{i,1},\cdots,\alpha_1^{i,m_i},\cdots,\alpha_{r_m}^{i,m_i})=L(\epsilon_{i1},\cdots,\epsilon_{in_i}$$

$$=V_{\lambda_i^{r_{m_i}+1}}=V_{\lambda_i^{r_{m_i}+2}}=\cdots=\mathcal{V}_i$$

所以，

$$\{0\}=V_{\lambda_i^0}\subset V_{\lambda_i^1}\subset V_{\lambda_i^2}\subset\cdots V_{\lambda_i^{r_{m_i}-1}}\subset V_{\lambda_i^{r_{m_i}}}=V_{\lambda^{r_{m_i}+1}}=\cdots=\mathcal{V}_i$$

其中 r_{m_i} 为 $\lambda=\lambda_i$ 对应阶数最多 Jordan 块的阶数。

$\lambda=\lambda_i$ 的特征子空间为 $V_{\lambda_i^1}$。

$\lambda=\lambda_i$ 的根子空间 $\{X\in V\mid(\lambda E-J)^{n_i}X=0\}=V$。这是因为 $n_i=\sum_j r_{ij}\geq r_{m_i}$。

3.3 核和值域

当特征值 $\lambda_i \neq 0$ 时，线性变换 $\sigma_i = \sigma|_{\mathcal{V}_i}$ 为可逆变换，它的核和值域分别为：

$$\ker \sigma_i = \{0\}$$

$$\text{im } \sigma_i = \mathcal{V}_i$$

当特征值 $\lambda_i = 0$ 时，则线性变换 $\sigma_i = \sigma|_{\mathcal{V}_i}$ 的核和值域分别为：

$$\ker \sigma_i = L(\alpha_{r_{i1}}^{i,1}, \alpha_{r_{i2}}^{i,2}, \cdots, \alpha_{r_{im_i}}^{i,m_i}) = V_{\lambda_i^1}$$

$$\text{im } \sigma_i = L(\alpha_2^{i,1}, \cdots, \alpha_{r_{i1}}^{i,1}, \cdots, \alpha_2^{i,m_i}, \cdots, \alpha_{r_{m_i}}^{i,m_i}) = V_{\lambda^{r_m - 1}}$$

3.4 $V = \ker \sigma \oplus \textbf{im } \sigma$

由上知，$V = \ker \sigma \oplus \text{im } \sigma$ 当且仅当特征值 0 对应的线性变换为零变换，即 $\lambda = 0$ 的代数重数等于几何重数。

4 一般复矩阵的情况

设 $\sigma : V \longrightarrow V$ 为复 n 维线性空间上的线性变换。设 σ 在一组基 $\beta_1, \cdots, \beta_n$ 下的矩阵为 A。我们总可以取 V 的一组基 $\epsilon_1, \cdots, \epsilon_n$，使得 σ 在这组基下的矩阵为一个 Jordan 矩阵 J。设由 $\beta_1, \cdots, \beta_n$ 到 $\epsilon_1, \cdots, \epsilon_n$ 的过渡矩阵为 P，则

$$P^{-1}AP = J$$

所以只需要考虑一般 Jordan 矩阵的情况即可。

综上，

定理. 设 $\sigma : V \longrightarrow V$ 为复 n 维线性空间上的线性变换。则 $V = \ker\sigma \oplus \operatorname{im}\sigma$ 当且仅当 σ 的特征值 $\lambda = 0$ 的代数重数等于几何重数。

推论. 设 $\sigma : V \longrightarrow V$ 为复 n 维线性空间上的可对角化线性变换，则 $V = \ker\sigma \oplus \operatorname{im}\sigma$。

推论. 设 $\sigma : V \longrightarrow V$ 为复 n 维线性空间上的幂等线性变换，则 $V = \ker\sigma \oplus \operatorname{im}\sigma$。

5 相似 Jordan 标准形的过渡矩阵

设 A 是一个 n 阶复矩阵。我们知道如果 A 可对角化，则 A 有 n 个线性无关的特征向量，不妨设为 $\beta_1, \beta_2, \cdots, \beta_n$。则 $P = (\beta_1, \beta_2, \cdots, \beta_n)$ 即为过渡矩阵，此时 $P^{-1}AP$ 是一个对角阵。

但如果 A 不可对角化，则 A 只有 $\sum m_i < n$ 个线性无关的特征向量，构成不了一个可逆过渡方阵。但从以上对不变子空间的分

析中，我们有以下结论。

定理. 设 A 是一个 n 阶复矩阵,则上述中 $\alpha_1^{i,1},\cdots,\alpha_{r_1}^{i,1},\cdots,\alpha_1^{i,m_i},\cdots,\alpha_{r_m}^{i,m_i}$ 构成了一个过渡矩阵 P，使得 $P^{-1}AP$ 是一个 Jordan 矩阵。

求 P 的过程.

- 求特征值。$|\lambda E-A|=0$ 的根给出所有特征值 $\lambda_1,\cdots,\lambda_s$；

- 求特征向量。对每个 λ_i，$\mathrm{Rank}(\lambda_i E-A)=n-m_i$。所以，$(\lambda_i E-A)X=0$ 给出了 m_i 个线性无关特征向量，每个 Jordan 块对应其中一个特征向量，不妨设为 $\alpha_1,\cdots\alpha_{m_i}$；

- 求向量 β_k 使得 $(\lambda_i E-A)\beta_k=\alpha_k$。对每个 λ_i 和 λ_i 的每个特征向量 α_k，考虑 $(\lambda_i E-A)X=\alpha_k$。如果对应 α_k 的 Jordan 块是一阶的，则 $(\lambda_i E-A)X=\alpha_k$ 无解；如果对应 α_k 的 Jordan 块大于一阶，则 $(\lambda_i E-A)X=\alpha_k$ 可解出一个向量 β_k，满足 $(\lambda_i E-A)^2\beta_k=0$；

- 通过不断迭代。对 r_i 阶的 Jordan 块对应的特征向量 α_k 可以求出 r_i 个线性无关的向量，设为 $\xi_1,\xi_2=(\lambda E-A)\xi_1,\cdots,\xi_{r_i}=(\lambda E-A)^{r_i-1}\xi_1=\alpha_k$，$(\lambda-A)^{r_i}\xi_1=0$（对比 P220 习题 10 和 11）；

- 这样我们就将 $\sum m_i$ 个线性无关的向量扩充到 n 个线性无关向量，就得到所需要的过渡矩阵 P。

例子 P243-6-7：

$$A = \begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix}$$

$f(\lambda) = |\lambda E - A| = \lambda(\lambda+1)^2 = 0$ 知特征值 $\lambda = 0, \lambda = -1$（二重）。

$(0E - A)X = 0$ 求得特征向量 $\alpha_1 = (2, 1, -1)^T$；

$(-1E - A)X = 0$ 求得特征向量 $\alpha_2 = (3, 3, -4)^T$；

$(0E - A)X = \alpha_1$ 无解；

$(-1E - A)X = \alpha_2$ 求得向量 $(1, 2, 0)^T + k\alpha_2$，可取 $\alpha_3 = (1, 2, 0)^T$。

这样就得到得过渡矩阵 $P = \alpha_1, \alpha_2, \alpha_3$，使得

$$P^{-1}AP = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

思考. 设 A 是一个数域 $\mathbb{F}$ 上的 n 阶方阵，我们知道特征多项式 $f(\lambda)=|\lambda E-A|$ 在数域 $\mathbb{F}$ 上的每个不可约因式都决定了一个友矩阵，则 A 相似于对角块为友矩阵的准对角块矩阵。问这里的过渡矩阵 P 怎么求？

（欢迎指正其中可能出现的错误。）