不变子空间与广义不变子空间*

(I) 存在与唯一性定理

孙　继　广

(中国科学院计算中心)

摘　　要

本文讨论与特征值和广义特征值问题相联系的某些子空间. 在本文中，我们定义了矩阵对的"广义特征值方阵对"和"广义特征矩阵"，并由此建立了广义不变子空间的概念；建立了对应的子空间存在与唯一的充分必要条件；给出了广义不变子空间与 G. W. Stewart 定义的"收缩对"的关系.

§1. 引　　言

在数值代数领域里，关于与特征值和广义特征值相连系的某些子空间的讨论，近几年在国际上正引起注意(参见 [1]—[6]). 在 1974 年的温哥华会议上，J. H. Wilkinson 专门就此作过一个特别报告[1]. 可是总的来说，这方面的研究还不成熟，有关的基本概念、扰动分析和数值计算仍有不少问题.

为了研究矩阵的不变子空间，本文定义了矩阵的"特征值方阵"和"特征矩阵"(分别是矩阵的特征值和特征向量的自然拓广)；此外，定义了矩阵对的"广义特征值方阵对"和"广义特征矩阵"(分别是广义特征值和广义特征向量的自然拓广)，并由此建立了广义不变子空间的概念. 本文的目的，在于研究对于不变子空间和广义不变子空间应该怎样提出问题，才能保证相应的解(是子空间)的存在性、唯一性和稳定性，以便给相应的子空间的数值计算，提供必要的代数基础.

关于存在性和唯一性，在 § 2 中建立了必要与充分条件. § 2 还论证了本文所定义的广义不变子空间与 G. W. Stewart[5] 所定义的"收缩子空间对"的关系. 关于稳定性，在 § 3 中赋于了特定的涵义，并建议一种具有二次收敛速度的迭代技术，用来逼近相应的子空间的扰动界限，从而得到了相应的扰动定理，并在末尾给出了一个实例.

以下用 $\mathbb{C}^{m\times n}$ 表示全体 m 行 n 列复数矩阵所成的空间. 用大写拉丁字母表示矩阵. $I^{(k)}$ 表示 k 行列单位方阵，e_i 表示 $I^{(k)}$ 的第 i 列，$N^{(k)}=(0, e_1, \cdots, e_{k-1})$. $\dotplus$ 表示矩阵的直和. A^T 表示 A 的转置，A^H 表示 A^T 的共轭. $\lambda(A)$ 表示方阵 A 的全体特征值所成的集合. 对于复变数 λ 和二同阶方阵 A 与 B，如果 $\det(A+\lambda B)\neq 0$，则称 $A+\lambda B$ 为正

* 1978 年 4 月 19 日收到.

则束([8], 第 XII 章). 本文仅限于讨论正则束. $\lambda(A, B)$ 表示 $A+\lambda B$ 的全体广义特征值所成的集合. $\|\cdot\|$表示矩阵模的相容族([7], 174—175). 对于 $Z_1\in \mathbf{C}^{n\times l}$, $1\leqslant l\leqslant n$, $\mathscr{R}(Z_1)$ 表示 $\mathbf{C}^{n\times n}$ 内由 Z_1 的 l 列所张成的线性子空间. $\mathscr{R}(Z_1)^{\perp}$ 表示 $\mathscr{R}(Z_1)$ 的酉正交补子空间. $\cap$表示二集合的交, ϕ 表示空集.

定义 1. 设 $A\in \mathbf{C}^{n\times n}$, $A_1\in \mathbf{C}^{l\times l}$, $1\leqslant l\leqslant n$. 如果存在 $Z_1\in \mathbf{C}^{n\times l}$ 适合 (i) $\operatorname{rank}(Z_1)=l$; (ii) $AZ_1=Z_1A_1$, 则 Z_1 叫做A的一个对应于 l 阶特征值方阵 A_1 的(右)特征矩阵.

定义 1 表明:

1) $\mathscr{R}(Z_1)$ 是A的一个 l 维不变子空间. 反之,如果在A的一个 l 维不变子空间 $\mathscr{K}$ 中任取一组基,并用这组基的 l 个列向量构成矩阵 $Z_1\in \mathbf{C}^{n\times l}$, 则 Z_1 是A的一个特征矩阵, 而且 $\mathscr{K}=\mathscr{R}(Z_1)$. 因此, A的一个 l 维不变子空间就是由A的某一个 l 阶特征矩阵的列向量所张成的子空间. 关于A的不变子空间的研究, 归根结底是关于A的特征矩阵的研究.

2) 当 $l=1$ 时,特征值方阵与特征矩阵分别是特征值与特征向量.

定义 2. 设 $A, B\in \mathbf{C}^{n\times n}$, $A_1, B_1\in \mathbf{C}^{l\times l}$, $1\leqslant l\leqslant n$, $A+\lambda B$ 与 $A_1+\lambda B_1$ 都是正则束. 如果存在 $Z_1\in \mathbf{C}^{n\times l}$ 适合

$$\text{(i)}\quad \operatorname{rank}(Z_1)=l \tag{1.1}$$

和

$$\text{(ii)}\quad AZ_1B_1=BZ_1A_1, \tag{1.2}$$

则称 Z_1 为方阵对 (A, B) 的对应于 l 阶广义特征值方阵对(以下简称广义特征对) (A_1, B_1) 的一个 l 阶广义特征矩阵. $\mathscr{R}(Z_1)$ 叫做 (A, B) 的一个 l 维广义不变子空间.

定义 2 表明:

1) 广义不变子空间,一般不是不变子空间; 但当 $B=I$ 时, 从 § 2 中的定理 2.1 知, B_1 必非奇异, 于是相应的广义不变子空间 $\mathscr{R}(Z_1)$ 便成了通常的不变子空间. 对于 (A, B) 的广义不变子空间的研究,归根结底是关于 (A, B) 的广义特征矩阵的研究.

2) 当 $l=1$ 时, 广义特征对和广义特征矩阵分别是广义特征值数对(即广义特征值的齐次参数)和广义特征向量.

§ 2. 存在性与唯一性

设 $A, B\in \mathbf{C}^{n\times n}$, $A_1, B_1\in \mathbf{C}^{l\times l}$, $1\leqslant l\leqslant n$, $A+\lambda B$ 与 $A_1+\lambda B_1$ 皆为正则束,则必存在 n 阶满秩方阵 P 与 Q, 和 l 阶满秩方阵 P_1 与 Q_1, 使得([8], 第 XII 章)

$$(A, B)=P(J_A, J_B)Q, \tag{2.1}$$

$$(A_1, B_1)=P_1(J_{A_1}, J_{B_1})Q_1. \tag{2.2}$$

(2.1) 中的

$$J_A=\begin{pmatrix} I^{(u)} & 0\\ 0 & J\end{pmatrix},\quad J_B=\begin{pmatrix} N_u & 0\\ 0 & I^{(n')}\end{pmatrix}, \tag{2.3}$$

$$J=J^{(1,1)}\dotplus\cdots\dotplus J^{(1,p_1)}\dotplus\cdots\dotplus J^{(r,1)}\dotplus\cdots\dotplus J^{(r,p_r)}, \tag{2.4}$$

$$N_u = N^{(u_1)} \dotplus \cdots \dotplus N^{(u_\alpha)},\ u_1 \geqslant \cdots \geqslant u_\alpha,\ u = \sum_{i=1}^{\alpha} u_i. \tag{2.5}$$

(2.4) 中的每一个 $J^{(i,j)} = \lambda_i I^{(n(i,j))} + N^{(n(i,j))}$ 为 Jordan 块，$k > j$ 时 $n(i, j) \geqslant n(i, k)$，$i \neq j$ 时 $\lambda_i \neq \lambda_j$. $\sum_{i=1}^{r} \sum_{j=1}^{p_i} n(i, j) = n',\ u + n' = n$. (2.2) 中的

$$J_{A_1} = \begin{pmatrix} I^{(v)} & 0 \\ 0 & J_1 \end{pmatrix},\quad J_{B_1} = \begin{pmatrix} N_v & 0 \\ 0 & I^{(l')} \end{pmatrix}, \tag{2.6}$$

$$J_1 = J_1^{(1,1)} \dotplus \cdots \dotplus J_1^{(1,q_1)} \dotplus \cdots \dotplus J_1^{(s,1)} \dotplus \cdots \dotplus J_1^{(s,q_s)}, \tag{2.7}$$

$$N_v = N^{(v_1)} \dotplus \cdots \dotplus N^{(v_\beta)},\ v_1 \geqslant \cdots \geqslant v_\beta,\ v = \sum_{i=1}^{\beta} v_i. \tag{2.8}$$

(2.7) 中的每一个 $J_1^{(i,j)} = \mu_i I^{(l(i,j))} + N^{(l(i,j))}$ 为 Jordan 块，$k > j$ 时 $l(i, j) \geqslant l(i, k)$，$i \neq j$ 时 $\mu_i \neq \mu_j$. $\sum_{i=1}^{s} \sum_{j=1}^{q_i} l(i, j) = l',\ v + l' = l$.

下面的定理 2.1 揭示了广义特征对与广义特征矩阵的结构特点. 定理 2.2 是定理 2.1 的特例，揭示了特征值方阵和特征矩阵的结构特点. 当 $l = 1$ 时，定理 2.1 和定理 2.2 的结论是熟知的.

定理 2.1. 设 $A, B \in \mathbf{C}^{n\times n}$，$A_1, B_1 \in \mathbf{C}^{l\times l}$，$1 \leqslant l \leqslant n$，$A + \lambda B$ 与 $A_1 + \lambda B_1$ 皆为正则束，(A, B) 与 (A_1, B_1) 在相抵下的标准形分别为 (2.1) 与 (2.2). 则存在 $Z_1 \in \mathbf{C}^{n\times l}$，使得 (1.1) 和 (1.2) 成立的必要充分条件是

① $\beta \leqslant \alpha$，$v_i \leqslant u_i (1 \leqslant i \leqslant \beta)$;

② $s \leqslant r$，经重排可使 $\mu_i = \lambda_i (1 \leqslant i \leqslant s)$;

③ $q_i \leqslant p_i (1 \leqslant i \leqslant s)$，$l(i, j) \leqslant n(i, j) (1 \leqslant i \leqslant s,\ 1 \leqslant j \leqslant q_i)$.

证明.

将 (2.1) 与 (2.2) 代入 (1.2) 中，令 $X_1 = QZ_1P_1 \in \mathbf{C}^{n\times l}$，则 (1.1) 与 (1.2) 分别为

(i)′ $\operatorname{rank}(X_1) = l$,

(ii)′ $J_A X_1 J_{B_1} = J_B X_1 J_{A_1}$.

于是只需证明：(i)′, (ii)′ $\Longleftrightarrow$ ①, ②, ③.

“$\Rightarrow$”:

1° 记

$$X_1 = \begin{array}{c} \begin{pmatrix} X_2 & Y \\ Z & X_3 \end{pmatrix}\begin{array}{l} u \\ n' \end{array} \\ \begin{array}{cc} v & l' \end{array} \end{array}. \tag{2.9}$$

可证 $X_1 = X_2 \dotplus X_3$，并且 $u \geqslant v$，$n' \geqslant l'$. 因而必须 $\operatorname{rank}(X_2) = v$，$\operatorname{rank}(X_3) = l'$.

为此，将 (2.3)，(2.6) 和 (2.9) 代入 (ii)′ 中，得

$$Y = N_u Y J_1,\quad JZN_v = Z \tag{2.10}$$

和

$$X_2 N_v = N_u X_2,\quad JX_3 = X_3 J_1. \tag{2.11}$$

把 (2.5) 和(2.7)代入(2.10)的第一式，并将 Y 适当分块为 (Y_{ij})，$1 \leqslant i \leqslant \alpha, 1 \leqslant j \leqslant s$，得

$$(Y_{ij}) = [N^{(u_1)}, \cdots, N^{(u_\alpha)}](Y_{ij})[J_{1,\mu_1}, \cdots, J_{1,\mu_s}],$$

其中 $J_{1,\mu_j}=J_1^{(j,1)}\dotplus\cdots\dotplus J_1^{(j,q_j)}$, $1\leqslant j\leqslant s$, 见(2.7). 于是

$$Y_{ij}=N^{(u_i)}Y_{ij}J_{1,\mu_j},\quad 1\leqslant i\leqslant\alpha,\ 1\leqslant j\leqslant s. \tag{2.12}$$

再将 Y_{ij} 分块为 $Y_{ij}=(Y_{ij}^{(1)},\cdots,Y_{ij}^{(q_j)})$, 代入(2.12), 则有

$$Y_{ij}^{(t)}=N^{(u_i)}Y_{ij}^{(t)}J_1^{(j,t)},\ 1\leqslant t\leqslant q_j, \tag{2.13}$$

其中 $Y_{ij}^{(t)}\in C^{u_i\times l(j,t)}$. 比较(2.13)两端的对应元素, 易知 $Y_{ij}^{(t)}=0$, 因而 $Y=0$. 同理知 $Z=0$.

于是 $X_1=X_2\dotplus X_3$. 由此可得 $u\geqslant v$, $n'\geqslant l'$;因若不然,则由 rank $(X_1)=\text{rank}(X_2)+\text{rank}(X_3)$ 立即可推知 rank $(X_1)<l$, 与 (i)′ 矛盾. 再由 (i)′ 可知,必有 $\text{rank}(X_2)=v$, rank $(X_3)=l'$. 这样就把要证的"$\Longrightarrow$"拆成了两个问题:

I. $X_2\in C^{u\times v}$, $v\leqslant u$, 适合 rank $(X_2)=v$ 及 $N_uX_2=X_2N_v$, $\Rightarrow$①;

II. $X_3\in C^{n'\times l'}$, $l'\leqslant n'$, 适合 rank $(X_3)=l'$ 及 $JX_3=X_3J_1$, $\Rightarrow$② 与 ③.

问题 I 只是 II 的特例,所以以下只需论证 II.

2° 证明 ②, 即任一 $\mu_j(1\leqslant j\leqslant s)$ 必等于某一 $\lambda_i(1\leqslant i\leqslant r)$.

把(2.4)与(2.7)代入(2.11)的第二式,并将 X_3 适当分块为 (X_{ij}), $1\leqslant i\leqslant r$, $1\leqslant j\leqslant s$, 得

$$[J_{\lambda_1},\cdots,J_{\lambda_r}](X_{ij})=(X_{ij})[J_{1,\mu_1},\cdots,J_{1,\mu_s}], \tag{2.14}$$

其中 $J_{\lambda_i}=J^{(i,1)}\dotplus\cdots\dotplus J^{(i,p_i)}$, $1\leqslant i\leqslant r$,见(2.4). 于是,若存在某一 $\mu_j\neq\lambda_i$(对所有的 $1\leqslant i\leqslant r$), 则由(2.14)中的

$$J_{\lambda_i}X_{ij}=X_{ij}J_{1,\mu_j},\quad i=1,2,\cdots,r$$

知[8] $X_{ij}=0$ $(1\leqslant i\leqslant r)$. 即 rank $(X_3)<l'$, 这与 II 中的假设条件矛盾.因此证明了②.

于是(2.14)便成了

$$[J_{\lambda_1},\cdots,J_{\lambda_r}](X_{ij})=(X_{ij})[J_{1,\lambda_1},\cdots,J_{1,\lambda_s}],\quad s\leqslant r.$$

因当 $i\neq j$ 时 $\lambda_i\neq\lambda_j$, 所以[8]: 当 $i\neq j$ 时, $X_{ij}=0$, 即

$$X_3=\begin{array}{c}\begin{pmatrix}X_{11} & & \\ & \ddots & \\ & & X_{ss}\\ 0 & \cdots & 0\end{pmatrix}\begin{array}{l}n_1\\ \\ n_s\\ \Big\}\sum_{j=s+1}^{r}n_j\end{array}\\ \begin{array}{ccc}l_1 & \quad & l_s\end{array}\end{array},\quad \sum_{i=1}^{r}n_i=n',\ \sum_{j=1}^{s}l_j=l', \tag{2.15}$$

其中 $n_i=\sum_{k=1}^{p_i}n(i,k)$, $l_j=\sum_{k=1}^{q_j}l(j,k)$, X_{ii} 满足

$$J_{\lambda_i}X_{ii}=X_{ii}J_{1,\lambda_i},\quad 1\leqslant i\leqslant s;$$

并由此可知,每一个 $X_{ii}=(\nabla_{\sigma\tau}^{(i)})$, $1\leqslant\sigma\leqslant p_i$, $1\leqslant\tau\leqslant q_i$. 其中 $\nabla_{\sigma\tau}^{(i)}\in C^{n(i,\sigma)\times l(i,\tau)}$ 满足

$$J^{(i,\sigma)}\nabla_{\sigma\tau}^{(i)}=\nabla_{\sigma\tau}^{(i)}J_1^{(i,\tau)},\quad 1\leqslant\sigma\leqslant p_i,\ 1\leqslant\tau\leqslant q_i,$$

因而每一块 $\nabla_{\sigma\tau}^{(i)}$ 都是形如

$$\begin{pmatrix}a & b & \cdots & c\\ & \ddots & \ddots & \vdots\\ & & \ddots & b\\ 0 & & & a\end{pmatrix}\text{或}\begin{pmatrix}a & b & \cdots & c\\ & \ddots & \ddots & \vdots\\ & & \ddots & b\\ 0 & & & a\end{pmatrix}\text{或}\begin{pmatrix}a & b & \cdots & c\\ & \ddots & \ddots & \vdots\\ 0 & & \ddots & b\\ & & & a\end{pmatrix} \tag{2.16}$$

者[8]．(2.16) 所示的这种特殊形式的上三角矩阵，特称之为 ∇-块．

3° 从 (2.15) 知，$\mathrm{rank}\,(X_3) = \sum_{i=1}^{s} \mathrm{rank}\,(X_{ii})$，然而 $\mathrm{rank}\,(X_{ii}) \leqslant \min\,(l_i, n_i)$，并且应有 $\mathrm{rank}\,(X_3) = l'$，由此立得：对于 $1 \leqslant i \leqslant s$，有 $l_i \leqslant n_i$，从而 $\mathrm{rank}\,(X_{ii}) = l_i$．

于是，对于每一个 i $(1 \leqslant i \leqslant s)$，已经得知

$$X_{ii} = (\nabla_{\sigma\tau}^{(i)}) \in \mathbb{C}^{n_i \times l_i},\ l_i \leqslant n_i,\ \mathrm{rank}\,(X_{ii}) = l_i,$$

其中每一 ∇-块 $\nabla_{\sigma\tau}^{(i)} \in \mathbb{C}^{n(i,\sigma)\times l(i,\tau)}(1 \leqslant \sigma \leqslant p_i, 1 \leqslant \tau \leqslant q_i)$，并且 $n(i,1) \geqslant \cdots \geqslant n(i,p_i)$，$l(i,1) \geqslant \cdots \geqslant l(i,q_i)$．由此可证 ③，即 $q_i \leqslant p_i$，$l(i,j) \leqslant n(i,j)$ $(1 \leqslant j \leqslant q_i)$．

事实上，假若 $q_i > p_i$，则 X_{ii} 中的第 1，$l(i,1)+1$，$l(i,1)+l(i,2)+1$，$\cdots$，$l(i,1)+\cdots+l(i,q_{i-1})+1$ 诸列线性相关(因为这 q_i 个列向量，都只在相同的 p_i 个位置上可能有非零元素，而 $p_i < q_i$)，这与 $\mathrm{rank}\,(X_{ii}) = l$ 矛盾，所以 $q_i \leqslant p_i$；进而，假若对于某一 j，$l(i,j) > n(i,j)$，则由 $n(i,j) \geqslant n(i,j+1) \geqslant \cdots \geqslant n(i,p_i)$ 知，X_{ii} 中的第 1，$l(i,1)+1$，$l(i,1)+l(i,2)+1$，$\cdots$，$l(i,1)+\cdots+l(i,j-1)+1$ 诸列线性相关(因为这 j 个列向量，都只在相同的 $j-1$ 个位置上可能有非零元素，而 $j-1<j$)，这亦与 $\mathrm{rank}\,(X_{ii}) = l$ 矛盾，所以 $l(i,j) \leqslant n(i,j)$ $(1 \leqslant j \leqslant q_i)$．

“$\Longleftarrow$”：

设已有 ①，②，③，今在

$$X_1 = \begin{array}{cc} \begin{pmatrix} X_2 & 0 \\ 0 & X_3 \end{pmatrix} & \begin{matrix} u \\ n' \end{matrix} \\ \begin{matrix} v \quad & l' \end{matrix} & \end{array} \quad (X_3 \text{ 如 } (2.15) \text{ 所示}) \tag{2.17}$$

中取

$$X_2 = \begin{pmatrix} \begin{pmatrix} I^{(v_1)} \\ 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} I^{(v_\beta)} \\ 0 \end{pmatrix} \\ 0 & \cdots & 0 \end{pmatrix} \begin{matrix} u_1 \\ \\ u_\beta \\ \} \sum_{i=\beta+1}^{\alpha} u_i \end{matrix}, \quad X_{ii} = \begin{pmatrix} \begin{pmatrix} I^{(l(i,1))} \\ 0 \end{pmatrix} & & \\ & \ddots & \\ & & \begin{pmatrix} I^{(l(i,q_i))} \\ 0 \end{pmatrix} \\ 0 & \cdots & 0 \end{pmatrix} \begin{matrix} n(i,1) \\ \\ n(i,q_i) \\ \} \sum_{j=q_i+1}^{p_i} n(i,j) \end{matrix},$$

则 (2.17) 所示的 X_1 必满足 (i)′ 与 (ii)′．定理证毕．

在 (2.1) 与 (2.2) 中，特别地取 $B = I^{(n)}$，$B_1 = I^{(l)}$，$Q = P^{-1}$，$Q_1 = P_1^{-1}$，则 $J_B = I^{(n)}$，$J_{B_1} = I^{(l)}$，并且 J_A 与 J_{A_1} 分别是 A 与 A_1 的 Jordan 标准形：$J_A = J$，$J_{A_1} = J_1$，J 与 J_1 分别如 (2.4) 与 (2.7) 所示．于是从定理 2.1 可立即推出

定理 2.2. 设 $A \in \mathbb{C}^{n\times n}$，$A_1 \in \mathbb{C}^{l\times l}$，$1 \leqslant l \leqslant n$，$A$ 与 A_1 的 Jordan 形分别为 (2.4) 与 (2.7)．则存在 $Z_1 \in \mathbb{C}^{n\times l}$，使得 (i) $\mathrm{rank}\,(Z_1) = l$，(ii) $AZ_1 = Z_1A_1$ 成立的必要充分条件是定理 2.1 中的 ② 与 ③．

上述定理表明：

① 所谓“A 的对应于它的特征值 $\lambda_1, \cdots, \lambda_l$ 的不变子空间”(见 [1]，[2]，[6])，其确切的涵义是 A 的对应于某一 l 阶特征值方阵 A_1 的特征矩阵 Z_1 所张成的子空间，A_1 以 $\lambda_1, \cdots, \lambda_l$ 为其全部特征值，且其结构必按定理 2.2 依赖于 A 的结构，否则，相应的不变子

空间是不存在的.

② 当定理中的必要与充分条件被满足时,相应的不变(或广义不变)子空间一般不是唯一的,在相应的特征(或广义特征)矩阵中,容许不超过 $\sum_{i=1}^{s}\sum_{j=1}^{p_i}\sum_{k=1}^{q_i}\min(n(i,j),l(i,k))$ 个(在广义的情形,再加上 $\sum_{i=1}^{\alpha}\sum_{j=1}^{\beta}\min(u_i,v_j)$ 个)的复参数.

下面两条定理,是在满足存在性条件的前提下,讨论相应的子空间的唯一性条件.

定理 2.3. 设 (A_1, B_1) 是 (A, B) 的某一广义特征对, 即标准形 (2.1)—(2.3) 和 (2.6) 满足定理 2.1 的条件 ①, ②, ③. 则 (A, B) 存在唯一的与 (A_1, B_1) 相对应的广义不变子空间的必要充分条件是

(a) $\beta=\alpha$;

(b) 若对某一 $j(2\leqslant j\leqslant\alpha)$ 有 $v_j<u_j$, 则对所有的 $k(1\leqslant k\leqslant j-1)$, 必有 $v_k=v_j$;

(c) $q_i=p_i$, $i=1,2,\cdots,s$;

(d) 若对某一对 $i(1\leqslant i\leqslant s)$ 与 $j\ (2\leqslant j\leqslant p_i)$ 有 $l(i,j)<n(i,j)$, 则对所有的 $k(1\leqslant k\leqslant j-1)$, 必有 $l(i,k)=l(i,j)$.

证明.

为了方便,现改记 $\alpha=p_0$, $\beta=q_0$, $u=n_0$, $v=l_0$, 对于 $1\leqslant j\leqslant\alpha$, 记 $u_j=n(0,j)$, 对于 $1\leqslant j\leqslant\beta$, 记 $v_j=l(0,j)$. 于是定理中的条件 (a)—(d) 可改述为

(c)′ $q_i=p_i$, $i=0,1,2,\cdots,s$;

(d)′ 若对某一对 $i(0\leqslant i\leqslant s)$ 与 $j(2\leqslant j\leqslant p_i)$ 有 $l(i,j)<n(i,j)$, 则对所有的 $k(1\leqslant k\leqslant j-1)$, 必有 $l(i,k)=l(i,j)$.

充分性.

将 (2.1) 与 (2.2) 代入 (1.2) 中,令 $QZ_1P_1=X_1$, 并将 (2.9) 中的 X_2 改记为 X_{00}. 定理 2.1 已证明了

$$X_1=\begin{array}{c}\left(\begin{array}{cccc}X_{00}&&&\\&X_{11}&&\\&&\ddots&\\&&&X_{ss}\\0&0&\cdots&0\end{array}\right)\begin{array}{l}n_0\\n_1\\ \\n_s\\ \}\ \sum\limits_{j=s+1}^{r}n_j\end{array}\\ \begin{array}{cccc}l_0\quad&l_1&\quad&l_s\end{array}\qquad\qquad\end{array}\tag{2.18}$$

由 (c)′ 知,对于 $0\leqslant i\leqslant s$, 每一个

$$X_{ii}=\begin{array}{c}\left(\begin{array}{ccc}\nabla_{1,1}^{(i)}&\cdots&\nabla_{1,p_i}^{(i)}\\ \cdots&&\\ \nabla_{p_i,1}^{(i)}&\cdots&\nabla_{p_i,p_i}^{(i)}\end{array}\right)\begin{array}{l}n(i,1)\\ \\ n(i,p_i)\end{array}\\ \begin{array}{ccc}l(i,1)&\quad&l(i,p_i)\end{array}\qquad\end{array}\in\mathbf{C}^{n_i\times l_i},\ \operatorname{rank}(X_{ii})=l_i.$$

其中每一个 $\nabla_{\sigma\tau}^{(i)}$ 都是 ∇-块(见 (2.16)). 由 (d)′ 知, 若用 $\nabla_{\sigma\tau}^{(i)\prime}$ 表示 $\nabla_{\sigma\tau}^{(i)}$ 的前 $l(i,\sigma)$ 行所成的子矩阵,则

$$\nabla_{\sigma\tau}^{(i)} = \begin{cases} \nabla_{\sigma\tau}^{(i)'} & \text{当 } n(i,\sigma) = l(i,\sigma) \text{ 时}, \\ \begin{array}{c} \begin{pmatrix} \nabla_{\sigma\tau}^{(i)'} \\ 0 \end{pmatrix} \begin{array}{l} l(i,\sigma) \\ n(i,\sigma) - l(i,\sigma) \end{array} \\ l(i,\tau) \end{array} & \text{当 } n(i,\sigma) > l(i,\sigma) \text{ 时}. \end{cases}$$

因此，若令

$$X'_{ii} = \begin{array}{c} \begin{pmatrix} \nabla_{1,1}^{(i)'} & \cdots & \nabla_{1,p_i}^{(i)'} \\ \cdots & & \\ \nabla_{p_i,1}^{(i)'} & \cdots & \nabla_{p_i,p_i}^{(i)'} \end{pmatrix} \begin{array}{l} l(i,1) \\ \\ l(i,p_i) \end{array} \\ \begin{array}{cc} l(i,1) & l(i,p_i) \end{array} \end{array} \in \mathbf{C}^{l_i \times l_i}, \text{ 有 } \operatorname{rank}(X'_{ii}) = l_i,$$

以及

$$I_i = \begin{array}{c} \begin{pmatrix} I_{i,1} & & \\ & \ddots & \\ & & I_{i,p_i} \end{pmatrix} \begin{array}{l} n(i,1) \\ \\ n(i,p_i) \end{array} \\ \begin{array}{cc} l(i,1) & l(i,p_i) \end{array} \end{array} \in \mathbf{C}^{n_i \times l_i}, \tag{2.19}$$

其中

$$I_{i,j} = \begin{cases} I^{(n(i,j))} & \text{当 } n(i,j) = l(i,j) \text{ 时}, \\ \begin{pmatrix} I^{(l(i,j))} \\ 0 \end{pmatrix} n(i,j) - l(i,j) & \text{当 } n(i,j) > l(i,j) \text{ 时}, \end{cases} \tag{2.20}$$

则

$$X_{ii} = I_i X'_{ii}.$$

从而，若再记

$$I' = \begin{array}{c} \begin{pmatrix} I_0 & & \\ & \ddots & \\ & & I_s \\ 0 & \cdots & 0 \end{pmatrix} \begin{array}{l} n_0 \\ \\ n_s \\ \} \sum_{j=s+1}^{r} n_j \end{array} \\ \begin{array}{cc} l_0 & l_s \end{array} \end{array} \in \mathbf{C}^{n \times l} \tag{2.21}$$

和

$$X'_1 = \begin{pmatrix} X'_{00} & & \\ & \ddots & \\ & & X'_{ss} \end{pmatrix} \in \mathbf{C}^{l \times l}, \text{ 有 } \operatorname{rank}(X'_1) = l, \tag{2.22}$$

便有

$$X_1 = I' X'_1.$$

因此得到了 (A, B) 的对应于广义特征对 (A_1, B_1) 的一个广义特征矩阵

$$Z_1 = Q^{-1} X_1 P_1^{-1} = Q^{-1} I' \cdot X'_1 P_1^{-1}, \tag{2.23}$$

其中 $Q^{-1}I' \in \mathbf{C}^{n \times l}$, $\operatorname{rank}(Q^{-1}I') = l$, $X'_1 P_1^{-1} \in \mathbf{C}^{l \times l}$, $\operatorname{rank}(X'_1 P_1^{-1}) = l$.

若 (A, B) 与 (A_1, B_1) 另有一分解

$$(A, B) = \tilde{P}(J_A, J_B)\tilde{Q}, \tag{2.1$'$}$$

$$(A_1, B_1) = \tilde{P}_1(J_{A_1}, J_{B_1})\tilde{Q}_1, \tag{2.2$'$}$$

其中 $\tilde{P}$ 与 $\tilde{Q}$ 是 n 阶满秩方阵，$\tilde{P}_1$ 与 $\tilde{Q}_1$ 是 l 阶满秩方阵. 则按上面同样的论证，可得 (A, B) 的对应于广义特征对 (A_1, B_1) 的另一个广义特征矩阵

$$\tilde{Z}_1 = \tilde{Q}^{-1} \tilde{X}_1 \tilde{P}_1^{-1} = \tilde{Q}^{-1} I' \cdot \tilde{X}'_1 \tilde{P}_1^{-1}, \tag{2.24}$$

其中 I' 见 (2.21)，$\widetilde{X}_1$ 和 $\widetilde{X}'_1$ 有如 (2.18) 和 (2.22) 之形式.

比较 (2.1) 与 (2.1)′，得到

$$P^{-1}\widetilde{P}J_A = J_A Q\widetilde{Q}^{-1}, \quad P^{-1}\widetilde{P}J_B = J_B Q\widetilde{Q}^{-1}.$$

从 $J_AJ_B = J_BJ_A$ 和

$$P^{-1}\widetilde{P}J_AJ_B = J_AQ\widetilde{Q}^{-1}J_B, \; P^{-1}\widetilde{P}J_BJ_A = J_BQ\widetilde{Q}^{-1}J_A$$

知，$Q\widetilde{Q}^{-1}$ 必满足

$$J_AQ\widetilde{Q}^{-1}J_B = J_BQ\widetilde{Q}^{-1}J_A.$$

由此可导出(参看定理 2.1 的证明)

$$Q\widetilde{Q}^{-1} = \begin{array}{cc} \begin{pmatrix} Y_{00} & & & \\ & Y_{11} & & \\ & & \ddots & \\ & & & Y_{rr} \end{pmatrix} & \begin{matrix} n_0 \\ n_1 \\ \\ n_r \end{matrix} \\ \begin{matrix} n_0 & n_1 & \quad & n_r \end{matrix} & \end{array} \in \mathbb{C}^{n\times n}, \tag{2.25}$$

此处每一个

$$Y_{ii} = \begin{array}{cc} \begin{pmatrix} \overset{*}{\nabla}{}^{(i)}_{1,1} & \cdots & \overset{*}{\nabla}{}^{(i)}_{1,p_i} \\ \cdots & & \\ \overset{*}{\nabla}{}^{(i)}_{p_i,1} & \cdots & \overset{*}{\nabla}{}^{(i)}_{p_i,p_i} \end{pmatrix} & \begin{matrix} n(i,1) \\ \\ n(i,p_i) \end{matrix} \\ \begin{matrix} n(i,1) & \quad & n(i,p_i) \end{matrix} & \end{array} \in \mathbb{C}^{n_i\times n_i}, \; \operatorname{rank}(Y_{ii}) = n_i, \; 0 \leqslant i \leqslant r$$

其中每一个 $\overset{*}{\nabla}{}^{(i)}_{\sigma\tau}$ 都是 ∇-块.

将 (2.25) 代入 (2.24) 中，知

$$\widetilde{Z}_1 = Q^{-1}\begin{pmatrix} Y_{00} & & \\ & \ddots & \\ & & Y_{rr} \end{pmatrix} I' \cdot \widetilde{X}'_1\widetilde{P}_1^{-1}. \tag{2.26}$$

注意到

$$Y_{ii}I_i = \begin{array}{cc} \begin{pmatrix} \overset{*}{\nabla}{}^{(i)}_{1,1}I_{i,1} & \cdots & \overset{*}{\nabla}{}^{(i)}_{1,p_i}I_{i,p_i} \\ \cdots & & \\ \overset{*}{\nabla}{}^{(i)}_{p_i,1}I_{i,1} & \cdots & \overset{*}{\nabla}{}^{(i)}_{p_i,p_i}I_{i,p_i} \end{pmatrix} & \begin{matrix} n(i,1) \\ \\ n(i,p_i) \end{matrix} \\ \begin{matrix} l(i,1) & \quad & l(i,p_i) \end{matrix} & \end{array},$$

而 $\overset{*}{\nabla}{}^{(i)}_{\sigma,\tau}I_{i,\tau}$ 表示 $\overset{*}{\nabla}{}^{(i)}_{\sigma,\tau}$ 的前 $l(i,\tau)$ 列所成的子矩阵；所以，如果用 $\overset{*}{\nabla}{}^{(i)'}_{\sigma,\tau}$ 表示 $\overset{*}{\nabla}{}^{(i)}_{\sigma,\tau}$ 中的前 $l(i,\sigma)$ 行、前 $l(i,\tau)$ 列所成的子矩阵，则

$$\overset{*}{\nabla}{}^{(i)}_{\sigma,\tau}I_{i,\tau} = \begin{cases} \overset{*}{\nabla}{}^{(i)'}_{\sigma,\tau} & \text{当 } n(i,\sigma) = l(i,\sigma) \text{ 时}, \\ \begin{array}{cc} \begin{pmatrix} \overset{*}{\nabla}{}^{(i)'}_{\sigma,\tau} \\ 0 \end{pmatrix} & \begin{matrix} l(i,\sigma) \\ n(i,\sigma) - l(i,\sigma) \end{matrix} \\ l(i,\tau) & \end{array} & \text{当 } n(i,\sigma) > l(i,\sigma) \text{ 时}. \end{cases} \tag{2.27}$$

(2.27) 右端的下式，是因为据 (d)′，对所有的 $\tau(1 \leqslant \tau \leqslant p_i)$，恒有 $l(i,\tau) \leqslant l(i,\sigma)$. 记

$$Y'_{ii} = \begin{array}{cc} \begin{pmatrix} \overset{*}{\nabla}{}^{(i)'}_{1,1} & \cdots & \overset{*}{\nabla}{}^{(i)'}_{1,p_i} \\ \cdots & & \\ \overset{*}{\nabla}{}^{(i)'}_{p_i,1} & \cdots & \overset{*}{\nabla}{}^{(i)'}_{p_i,p_i} \end{pmatrix} & \begin{matrix} l(i,1) \\ \\ l(i,p_i) \end{matrix} \\ \begin{matrix} l(i,1) & \quad & l(i,p_i) \end{matrix} & \end{array} \in \mathbb{C}^{l_i\times l_i}, \text{ 有 } \operatorname{rank}(Y'_{ii}) = l_i.$$

显然有

$$Y_{ii}I_i = I_i Y'_{ii}, \quad 0 \leqslant i \leqslant s,$$

从而

$$\begin{pmatrix} Y_{00} & & \\ & \ddots & \\ & & Y_{rr} \end{pmatrix} I' = I' \begin{pmatrix} Y'_{00} & & \\ & \ddots & \\ & & Y'_{ss} \end{pmatrix}.$$

代入 (2.26) 中，得到

$$\tilde{Z}_1 = Q^{-1} I' \cdot \begin{pmatrix} Y'_{00} & & \\ & \ddots & \\ & & Y'_{ss} \end{pmatrix} \tilde{X}'_1 \tilde{P}_1^{-1}, \tag{2.28}$$

其中 $\begin{pmatrix} Y'_{00} & & \\ & \ddots & \\ & & Y'_{ss} \end{pmatrix} \tilde{X}'_1 \tilde{P}^{-1}$ 为 l 阶满秩方阵.

比较 (2.23) 和 (2.28) 知，$\mathscr{R}(\tilde{Z}_1) = \mathscr{R}(Z_1) = \mathscr{R}(Q^{-1}I')$，即 $\mathscr{R}(Z_1)$ 唯一.

必要性.

先证 (c)′. 反证之. 假定对某一 $k(0 \leqslant k \leqslant s)$，$q_k < p_k$. 据充分性证明中的记号，在 (2.18) 中，当 $i \neq k$ 时，取 $X_{ii} = I_i$ (见 (2.19))；同时取

$$X_{kk}^{(1)} = \begin{array}{c} \begin{pmatrix} I_{k,1} & & \\ & \ddots & \\ & & I_{k,q_k} \\ 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{pmatrix} \begin{array}{l} n(k,1) \\ \\ n(k,q_k) \\ n(k,q_k+1) \\ \} \sum_{j=q_k+2}^{p_k} n(k,j) \end{array} \\ \begin{array}{ll} l(k,1) & l(k,q_k) \end{array} \end{array} \tag{2.29}$$

和

$$X_{kk}^{(2)} = \begin{array}{c} \begin{pmatrix} I_{k,1} & & \\ & \ddots & \\ & & I_{k,q_k} \\ \nabla_l \; 0 & \cdots & 0 \\ 0 \;\; 0 & \cdots & 0 \end{pmatrix} \begin{array}{l} n(k,1) \\ \\ n(k,q_k) \\ n(k,q_k+1) \\ \} \sum_{j=q_k+2}^{p_k} n(k,j) \end{array} \\ \begin{array}{ll} l(k,1) & l(k,q_k) \end{array} \end{array} \tag{2.30}$$

其中 $I_{k,j}$ 见 (2.20)，而

$$\nabla_l = \begin{cases} \begin{pmatrix} I \\ 0 \end{pmatrix} \begin{array}{l} l(k,1) \\ n(k,q_k+1) - l(k,1) \end{array} & \text{当 } n(k,q_k+1) > l(k,1) \text{ 时}, \\ I^{(l(k,1))} & \text{当 } n(k,q_k+1) = l(k,1) \text{ 时}, \\ \underbrace{(0,\ I)}_{l(k,1)}\, n(k,q_k+1) & \text{当 } n(k,q_k+1) < l(k,1) \text{ 时}, \end{cases} \tag{2.31}$$

于是得到方程 $J_A X_1 J_{B_1} = J_B X_1 J_{A_1}$ 的秩为 l 的两个解

$$X_1^{(j)} = \begin{pmatrix} I_1 & & & \\ & \ddots & & \\ & & X_{kk}^{(j)} & \\ & & & \ddots \\ & & & & I_s \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}, \quad j = 1, 2. \tag{2.32}$$

从而得到方程 $AZ_1B_1 = BZ_1A_1$ 的秩为 l 的两个解

$$Z_1^{(j)} = Q^{-1}X_1^{(j)}P_1^{-1}, \quad j = 1, 2.$$

为了比较 $\mathscr{R}(Z_1^{(1)})$ 与 $\mathscr{R}(Z_1^{(2)})$,只需比较 $Q^{-1}X_1^{(1)}$ 与 $Q^{-1}X_1^{(2)}$. 假定 $\mathscr{R}(Z_1^{(1)})=\mathscr{R}(Z_1^{(2)})$,则必存在满秩 l 阶方阵 R,使 $Q^{-1}X_1^{(1)}R=Q^{-1}X_1^{(2)}$,于是 $X_1^{(1)}R=X_1^{(2)}$,即 $\mathscr{R}(X_1^{(1)}) = \mathscr{R}(X_1^{(2)})$. 然而从 (2.29)—(2.32) 可看出 $\mathscr{R}(X_1^{(1)}) \neq \mathscr{R}(X_1^{(2)})$,所以 $\mathscr{R}(Z_1^{(1)}) \neq \mathscr{R}(Z_1^{(2)})$. 这就证明了 (c)'.

再证 (d)'. 亦反证之. 假定在 (c)' 已满足的条件下,(d)' 不满足,即存在一组三个自然数 i_0, j_0, k_0: $0 \leqslant i_0 \leqslant s$, $1 \leqslant k_0 < j_0 \leqslant p_{i_0}$, 使得

$$n(i_0, j_0) > l(i_0, j_0), \quad l(i_0, k_0) > l(i_0, j_0),$$

则可在 (2.18) 中,当 $i \neq i_0$ 时,取 $X_{ii} = I_i$ (见 (2.19)); 同时取

$$X_{i_0,i_0}^{(1)} = \begin{array}{c} \begin{pmatrix} I_{i_0,1} & & & & \\ & \ddots & & & \\ & & I_{i_0,j_0} & & \\ & & & \ddots & \\ & & & & I_{i_0,p_{i_0}} \end{pmatrix} \begin{array}{l} n(i_0, 1) \\ \\ n(i_0, j_0) \\ \\ n(i_0, p_{i_0}) \end{array} \\ l(i_0, 1)\ l(i_0, j_0)\ l(i_0, p_{i_0}) \end{array}$$

和

$$X_{i_0,i_0}^{(2)} = \begin{array}{c} \begin{pmatrix} I_{i_0,1} & & & & \\ & \ddots & & & \\ \nabla_l' & & I_{i_0,j_0} & & \\ & & & \ddots & \\ & & & & I_{i_0,p_{i_0}} \end{pmatrix} \begin{array}{l} n(i_0, 1) \\ \\ n(i_0, j_0) \\ \\ n(i_0, p_{i_0}) \end{array} \\ l(i_0, 1)\ l(i_0, j_0)\ l(i_0, p_{i_0}) \end{array},$$

其中 $I_{i_0,j}$ 见 (2.20),而

$$\nabla_l = \begin{cases} \begin{pmatrix} I \\ 0 \end{pmatrix} \begin{array}{l} l(i_0, 1) \\ n(i_0, j_0) - l(i_0, 1) \end{array} & \text{当 } n(i_0, j_0) > l(i_0, 1) \text{ 时}, \\ I^{(l(i_0,1))} & \text{当 } n(i_0, j_0) = l(i_0, 1) \text{ 时}, \\ \underbrace{(0, I)}_{l(i_0, 1)} n(i_0, j_0) & \text{当 } n(i_0, j_0) < l(i_0, 1) \text{ 时}, \end{cases}$$

于是得到方程 $J_AX_1J_{B_1} = J_BX_1J_{A_1}$ 的秩为 l 的两个解

$$X_1^{(j)} = \begin{pmatrix} I_1 & & & \\ & \ddots & & \\ & & X_{i_0,i_0}^{(1)} & & \\ & & & \ddots & \\ & & & & I_s \\ 0 & \cdots\cdots & & & 0 \end{pmatrix}, \quad j = 1, 2.$$

从而得到方程 $AZ_1B_1 = BZ_1A_1$ 的秩为 l 的两个解

$$Z_1^{(j)} = Q^{-1}X_1^{(j)}P_1^{-1}, \quad j = 1, 2.$$

与证明 (c)' 同理,可证 $\mathscr{R}(Z_1^{(1)}) \neq \mathscr{R}(Z_1^{(2)})$. 这就证明了 (d)'. 定理证毕.

下述定理是定理 2.3 的推论.

定理 2.4. 设 A_1 是 A 的某一特征值方阵,即 A 和 A_1 的 Jordan 形(2.4) 和 (2.7) 满足定理 2.1 中的条件 ②,③. 则 A 存在唯一的与 A_1 相对应的不变子空间的必要充分条件是

定理 2.3 中的 (c) 与 (d).

从上述定理可以看出，为了保证"A 的对应于它的特征值 $\lambda_1, \cdots, \lambda_l$ 的不变子空间"(见 [1], [2], [6]) 是唯一的，$\lambda_1, \cdots, \lambda_l$ 必须恰好是 A 的某一具有定理 2.4 所述的特定结构的特征值方阵 A_1 的全部特征值.

为了研究与广义特征值问题相连系的子空间，G. W. Stewart[5] 曾定义了"收缩对"(deflating pair)，即

定义 3. 设 $A, B \in \mathbf{C}^{n\times n}$，$\mathscr{X}$ 与 $\mathscr{Y}$ 是二子空间. 如果

$$\dim(\mathscr{X}) = \dim(\mathscr{Y})$$

并且

$$A\mathscr{X}, B\mathscr{X} \subset \mathscr{Y},$$

则 $\mathscr{X}$ 与 $\mathscr{Y}$ 叫做 (A, B) 的一个收缩对.

下述定理揭示了收缩对与广义不变子空间的关系.

定理 2.5. 设 $A, B \in \mathbf{C}^{n\times n}$，$A + \lambda B$ 是正则束，则对于子空间 $\mathscr{X}$，存在某一子空间 $\mathscr{Y}$，使得 $\mathscr{X}$ 与 $\mathscr{Y}$ 是 (A, B) 的收缩对 $\Longleftrightarrow \mathscr{X}$ 是 (A, B) 的一个广义不变子空间 $\mathscr{R}(Z_1)$.

为了证明定理 2.5，先证下面两条定理，它们对于讨论扰动问题也是有用的，所以单独列出.

定理 2.6. 设 $A, B \in \mathbf{C}^{n\times n}$，$A + \lambda B$ 是正则束. 如果存在 n 阶满秩方阵 W 与 $Z = (Z_1, Z_2)$，$Z_1 \in \mathbf{C}^{n\times l}$，$1 \leqslant l \leqslant n$，使得

$$W(A, B)Z = \left(\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}\right), \quad A_{11}, B_{11} \in \mathbf{C}^{l\times l}, \tag{2.33}$$

则 Z_1 必是 (A, B) 的对应于某一 l 阶广义特征对 (A_1, B_1) 的广义特征矩阵，此处 (A_1, B_1) 是与 (A_{11}, B_{11}) 相抵的某一矩阵对.

证明.

由 $A + \lambda B$ 为正则束知，$A_{11} + \lambda B_{11}$ 亦为正则束. 因而存在 l 阶满秩方阵 P_1 与 Q_1，使得

$$(A_{11}, B_{11}) = P_1(J_{A11}, J_{B11})Q_1,$$

其中 J_{A11} 与 J_{B11} 有如 (2.6) 所示的形式. 令

$$(A_1, B_1) = Q_1^{-1}(J_{A11}, J_{B11}), \tag{2.34}$$

则由 $J_{A11}J_{B11} = J_{B11}J_{A11}$ 知 $A_{11}B_1 = B_{11}A_1$，即

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}\begin{pmatrix} I \\ 0 \end{pmatrix} B_1 = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}\begin{pmatrix} I \\ 0 \end{pmatrix} A_1. \tag{2.35}$$

将 (2.33) 代入 (2.35) 中，即得 $AZ_1B_1 = BZ_1A_1$. $\operatorname{rank}(Z_1) = l$ 是显然的. 再由 (2.34) 知 $A_1 + \lambda B_1 = (P_1Q_1)^{-1}(A_{11} + \lambda B_{11})Q_1^{-1}$ 为正则束，且 (A_1, B_1) 与 (A_{11}, B_{11}) 相抵. 证毕.

定理 2.7. 设 $A, B \in \mathbf{C}^{n\times n}$，$A + \lambda B$ 是正则束. 如果 $Z_1 \in \mathbf{C}^{n\times l}(1 \leqslant l \leqslant n)$ 是 (A, B) 的对应于 l 阶广义特征对 (A_1, B_1) 的广义特征矩阵，则必存在 n 阶满秩方阵 W 与 $Z = (Z_1, Z_2)$，使得 (2.33) 成立，并且 (A_{11}, B_{11}) 是与 (A_1, B_1) 相抵的某一方阵对.

证明.

1° 先证,存在 $W_1 \in \mathbf{C}^{n\times l}$ 和 $R_1, S_1 \in \mathbf{C}^{l\times l}$, $\operatorname{rank}(W_1) = l$, 使得

$$AZ_1 = W_1R_1, \quad BZ_1 = W_1S_1. \tag{2.36}$$

利用 (2.1) 和 (2.2), 令 $X_1 = QZ_1P_1$, 则由定理 2.1 的证明,知 X_1 必形如 (2.17). 于是(参看 (2.3) 与 (2.6))

$$J_AX_1 = \begin{pmatrix} X_2 & 0 \\ 0 & JX_3 \end{pmatrix}, \quad J_BX_1 = \begin{pmatrix} N_uX_2 & 0 \\ 0 & X_3 \end{pmatrix}.$$

注意到 N 与 ∇-块可交换,因而

$$N_uX_2 = X_2N_v, \quad JX_3 = X_3J.$$

令

$$R_1' = \begin{pmatrix} I^{(v)} & 0 \\ 0 & J_1 \end{pmatrix}, \quad S_1' = \begin{pmatrix} N^{(v)} & 0 \\ 0 & I^{(l')} \end{pmatrix},$$

则有

$$J_AX_1 = X_1R_1', \quad J_BX_1 = X_1S_1'.$$

将 $X_1 = QZ_1P_1$ 代入上式,并令

$$W_1 = PQZ_1P_1, \quad R_1 = P_1R_1'P_1^{-1}, \quad S_1 = P_1S_1'P_1^{-1},$$

则 (2.36) 成立,且显然 $\operatorname{rank}(W_1) = l$.

2° 取一秩为 $n-l$ 的 $Z_2 \in \mathbf{C}^{n\times(n-l)} \cap \mathscr{R}(Z_1)^\perp$ 和秩为 $n-l$ 的 $W_2 \in \mathbf{C}^{n\times(n-l)} \cap \mathscr{R}(W_1)^\perp$, 记 $Z = (Z_1, Z_2)$ 和 $W = (W_1, W_2)^H$, 则由 (2.36) 立知 (2.33) 成立, 其中 $A_{ij} = W_i^HAZ_j$, $B_{ij} = W_i^HBZ_j$, $1 \leqslant i \leqslant j \leqslant 2$.

3° 用 W_1^H 从左乘以 $AZ_1B_1 = BZ_1A_1$ 的两端,得

$$A_{11}B_1 = B_{11}A_1. \tag{2.37}$$

由定理假设知, $A_1 + \lambda B_1$ 与 $A_{11} + \lambda B_{11}$ 皆为正则束,因而存在 $\tau \neq 0$, 使得

$$\det(A_{11} + \tau B_{11}) \neq 0, \quad \det(A_1 + \tau B_1) \neq 0.$$

于是从 (2.37) 可得

$$(A_{11}, B_{11})(A_1 + \tau B_1) = (A_{11} + \tau B_{11})(A_1, B_1),$$

即

$$(A_{11}, B_{11}) = (A_{11} + \tau B_{11})(A_1, B_1)(A_1 + \tau B_1)^{-1}.$$

这表示 (A_{11}, B_{11}) 与 (A_1, B_1) 相抵. 证毕.

定理 2.5 的证明

“$\Longrightarrow$”

设 $\dim(\mathscr{X}) = \dim(\mathscr{Y}) = l$. 今在 $\mathbf{C}^{n\times l}$ 内分别取秩为 l 的 $Z_1 \in \mathscr{X}$ 与 $W_1 \in \mathscr{Y}$, 于是 $\mathscr{R}(Z_1) = \mathscr{X}$, $\mathscr{R}(W_1) = \mathscr{Y}$. 再分别在 $\mathscr{X}^\perp$ 与 $\mathscr{Y}^\perp$ 内取秩为 $n-l$ 的 Z_2 与 $W_2 \in \mathbf{C}^{n\times(n-l)}$. 令 $Z = (Z_1, Z_2)$ 和 $W = (W_1, W_2)^H$, 则有 (2.33) 成立. 据定理 2.6, $\mathscr{X} = \mathscr{R}(Z_1)$ 是 (A, B) 的一个广义不变子空间.

“$\Longleftarrow$”

据定理 2.7 证明之 1° (见 (2.36)), 记 $\mathscr{R}(Z_1) = \mathscr{X}$ 和 $\mathscr{R}(W_1) = \mathscr{Y}$, 则 $\mathscr{X}$ 与 $\mathscr{Y}$ 构成 (A, B) 的一个收缩对. 证毕.

附带指出，利用 Stewart 定义的“收缩对”去研究相应的子空间的存在、唯一性定理，似乎是不方便的.

参 考 文 献

[1] J. H. Wilkinson, Invariant subspaces, Proceedings of the International of Mathematicians, Vancouver, 1974, 443—448.

[2] G. H. Golub. J. H. Wilkinson, Ill-conditioned eigensystems and the computation of the Jordan canonical form, *SIAM Rev.*, **18**(1976), 578—619.

[3] G. W. Stewart, Error bounds for invariant subspaces of closed linear operators, *SIAM J. Numer. Anal.*, **8**(1971), 796—808.

[4] ————, On the sensitivity of the eigenvalue problem $Ax=\lambda Bx$, *SIAM J. Numer. Anal.*, **9**(1972), 669—686.

[5] ————, Error and perturbation bounds for subspaces associated with certain eigenvalue problems, *SIAM Rev.*, **15**(1973), 727—764.

[6] ————, Simultaneous iteration for computing invariant subspaces of non-Hermitian matrices, *Numer. Math.*, **25**(1976), 123—136.

[7] ————, Introduction to Matrix Computations, 1973.

[8] Ф. Р. Гантмахер, Теория Матриц (有中译本，柯召译，高等教育出版社出版，1955 年).

INVARIANT SUBSPACES AND GENERALIZED INVARIANT SUBSPACES

(I) EXISTENCE AND UNIQUENESS THEOREMS

Sun Ji-guang

(Computing Center, Chinese Academy of Sciences)

Abstract

This paper discusses some subspaces associated with eigenvalue and generalized eigenvalue problems.

In this paper, we define "pair of generalized eigenvalue matrices" and "generalized eigenmatrix" of a pair of matrices, and from these definitions we establish the concept of generalized invariant subspaces; establish the necessary and sufficient conditions for the existence and uniqueness of corresponding subspaces; give the relationship of generalized invariant subspaces with the "deflating pair" which was defined by G. W. Stewart.