利用矩阵的秩讨论若当标准型

首先， 对于如下 $r\times r$ 的若当块矩阵

$$\mathbf{J}=\begin{pmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda \end{pmatrix}$$

任给$\eta\in$C，考虑矩阵 $Q(\eta)=\eta\cdot\mathbf{E}_{r\times r}-\mathbf{J}$, 那么我们如下简单性质：

性质 1. 如果$\eta\neq\lambda$, 那么 $Q(\eta)$为可逆矩阵.

性质 2. 当 $1\leq m\leq r$ 时，$\text{rank}(Q(\lambda)^m)=r-m$.

性质 3. 当 $m\geq r$ 时 $Q(\lambda)^m=\mathbf{0}$.

设矩阵 $\mathbf{A}$ 为 $n\times n$ 的矩阵，它的若当标准型 $\mathbf{J}=\text{diag}(\mathbf{J}_1,\mathbf{J}_2,\ldots,\mathbf{J}_K)$,即存在可逆矩阵 P 使得 $\mathbf{A}=\mathbf{PJP}^{-1}$ 成立，其中 $\mathbf{J}_i=\begin{pmatrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_i \end{pmatrix}$, 并且 $\mathbf{J}_i$ 的阶数为 r_i , $i=1,2,\ldots,K$. 很明显，对于不同的 i，相应的若当块的对角元素可能是相同的。

很自然，我们有如下的简单关系：

$$r_1+r_2+\ldots+r_K=n$$

下面我们讨论一下矩阵$(\lambda\cdot\mathbf{E}_{n\times n}-\mathbf{A})^m$ 的秩。由于存在可逆矩阵 P 使得 $\mathbf{A}=\mathbf{PJP}^{-1}$ 成立，我们只需要分析矩阵$(\lambda\cdot\mathbf{E}_{n\times n}-\mathbf{J})^m$ 的秩就可以了。

当λ 不为 $\mathbf{A}$ 的特征值时，$(\lambda\cdot\mathbf{E}_{n\times n}-\mathbf{J})^m$ 为可逆矩阵，这对于我们进一步的讨论没有任何意义。因此，我们只考虑λ 是 $\mathbf{A}$ 特征值的情形, 并且不妨设在 A 的若当标准型中$\lambda=\lambda_i=\lambda_{i+1}=\ldots=\lambda_{i+s-1}$ 所对应的若当块为 $\mathbf{J}_i, \mathbf{J}_{i+1},\ldots, \mathbf{J}_{i+s-1}$ 共 s 个，那么

$$\text{rank}((\lambda\cdot\mathbf{E}_{n\times n}-\mathbf{A})^m)=\text{rank}((\lambda\cdot\mathbf{E}_{n\times n}-\mathbf{J})^m)=\sum_{i=1}^{K}\text{rank}((\lambda\cdot\mathbf{E}_{r_i\times r_i}-\mathbf{J}_i)^m)$$

很明显，当 $j<i$ 或者 $j\geq i+s$ 时 $\text{rank}((\lambda\cdot\mathbf{E}_{r_j\times r_j}-\mathbf{J}_j)^m)=r_j$;

对于 $i\leq j\leq i+s-1$ 的情形，我们需要区分 $1\leq m\leq r_j$ 和 $m>r_j$ 的情况。

根据性质 2，当 $i\leq j\leq i+s-1$ 且 $1\leq m\leq r_j$ 时，$\text{rank}((\lambda\cdot\mathbf{E}_{r_j\times r_j}-\mathbf{J}_j)^m)=r_j-m$;

根据性质 3，当 $i\leq j\leq i+s-1$ 且 $m\geq r_j$ 时, $\text{rank}((\lambda\cdot\mathbf{E}_{r_j\times r_j}-\mathbf{J}_j)^m)=0$;

如果对于 $x\in$R 引入记号 $(x)_+=\max\{x,0\}$, 那么我们有:

当 $i\leq j\leq i+s-1$ 时，$\text{rank}((\lambda\cdot\mathbf{E}_{r_j\times r_j}-\mathbf{J}_j)^m)=(r_j-m)_+$.

因此

$$\text{rank}((\lambda\cdot\mathbf{E}_{n\times n}-\mathbf{A})^m)=\sum_{i=1}^{K}\text{rank}((\lambda\cdot\mathbf{E}_{r_i\times r_i}-\mathbf{J}_i)^m)=\sum_{j=1}^{i-1}r_j+\sum_{j=i+s}^{K}r_j+\sum_{j=i}^{i+s-1}(r_j-m)_+$$

$$=n-\sum_{j=i}^{i+s-1}(r_j-(r_j-m)_+)\ =n-\sum_{j=i}^{i+s-1}\min(r_j,m)$$

所以我们不妨记 $d_m=n-\operatorname{rank}((\lambda\cdot\mathbf{E}_{n\times n}-\mathbf{A})^m)$，那么 $d_m=\sum_{j=i}^{i+s-1}\min(r_j,m)$

设 $r=\max(r_i, r_{i+1},\ldots, r_{i+s-1})$,那么明显有 $d_1<d_2<\cdots<d_r=d_{r+1}=d_{r+2}=\cdots$.
进一步，设对应于 A 的特征值λ的所有若当块中阶数为 k 的个数有 t_k 个，那么我们有

$$d_m=\sum_{j=i}^{i+s-1}\min(r_j,m)=\sum_{k=1}^{r}t_k\cdot\min(k,m)$$

对于 $m=1,2,\ldots,r$,我们得到如下的线性方程组：

$$\begin{cases} t_1+t_2+t_3+\cdots+t_r=d_1 \\ \quad t_2+t_3+\cdots+t_r=d_2-d_1 \\ \qquad t_3+\cdots+t_r=d_3-d_2 \\ \qquad\quad \cdots \\ \qquad\qquad t_r=d_r-d_{r-1} \end{cases}$$

因此，我们有 $t_1=2d_1-d_2$, $t_r=d_r-d_{r-1}$,当 $1<k<r$ 时 $t_k=2d_k-d_{k-1}-d_{k+1}$.当 $r=1$ 时，公式 $t_r=d_r-d_{r-1}$ 无法定义。事实上，可以补充定义 $d_0=0$,那么公式 $t_k=2d_k-d_{k-1}-d_{k+1}$,对于 $k=1,\ldots,r$ 都成立。因为当 $k=1,\ldots,r-1$ 显然是成立的,而当 $k=r$ 时，由于 $d_{r+1}=d_r$,因此公式 $t_r=2d_r-d_{r-1}-d_{r+1}=d_r-d_{r-1}$,仍然是成立的。综合上面的讨论，我们得到如下基于特征矩阵幂的秩决定矩阵若当标准型的算法：
第一步：对于给定矩阵特征值λ, 对于 $m=1,2,\ldots$, 计算 $d_m=n-\operatorname{rank}((\lambda\cdot\mathbf{E}_{n\times n}-\mathbf{A})^m)$. 显然有
$d_1<d_2<\cdots<d_r=d_{r+1}=d_{r+2}=\cdots$， 因此当 d_m 不再严格增加的 m 值置为 r;
第二步：补充定义 $d_0=0$;那么根据公式
$t_k=2d_k-d_{k-1}-d_{k+1}$, $k=1,2,\ldots,r$
计算出矩阵特征值λ的阶数为 k 的若当块的个数 t_k.

下面我们在讨论以下根据特征矩阵幂的秩决定最小多项式 $m_{\mathbf{A}}(\lambda)$。

设矩阵 $\mathbf{A}$ 为 $n\times n$ 的矩阵，它的特征多项式可以写为

$$\varphi_{\mathbf{A}}(\lambda)=(\lambda-\lambda_1)^{n_1}\ldots(\lambda-\lambda_L)^{n_L},$$

其中特征根各不相同。由于最小多项式和特征多项式的根相同，只是每个根的重数不一样，因此我们的目标就是决定在最小多项式 $m_{\mathbf{A}}(\lambda)$每个特征根的指数。设

$$m_{\mathbf{A}}(\lambda)=(\lambda-\lambda_1)^{m_1}\ldots(\lambda-\lambda_L)^{m_L},$$

很明显 $1\le r_l\le n_l$, $l=1,2,\ldots,L$.
对于特征值λ_l, $l=1,2,\ldots,L$, 我们考虑矩阵$(\lambda_l\cdot\mathbf{E}_{n\times n}-\mathbf{A})^m$ 的秩。对于 A 的若当标准型作前面的假设 。设矩阵 $\mathbf{A}$ 的若当标准型为 $\mathbf{J}=\operatorname{diag}(\mathbf{J}_1,\mathbf{J}_2,\ldots,\mathbf{J}_K)$,即存在可逆矩阵 P 使得 $\mathbf{A}=\mathbf{PJP}^{-1}$ 成立，

其中 $\mathbf{J}_i=\begin{pmatrix}\lambda_i & 1 & \cdots & 0\\ 0 & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & 1\\ 0 & \cdots & 0 & \lambda_i\end{pmatrix}$, 并且 $\mathbf{J}_i$ 的阶数为 r_i, $i=1,2,\ldots,K$. 很明显，对于不同的 i，相应的

若当块的对角元素可能是相同的。因此 A 的最小多项式 $m_{\mathbf{A}}(\lambda)$和 **J** 的最小多项式 $m_{\mathbf{J}}(\lambda)$是完全相同的。这样我们来讨论 **J** 的最小多项式 $m_{\mathbf{J}}(\lambda)$。　有最小多项式的定义

$$m_{\mathbf{J}}(\mathbf{J})=(\mathbf{J}-\lambda_1\cdot\mathbf{E}_{n\times n})^{m_1}\ldots(\mathbf{J}-\lambda_L\cdot\mathbf{E}_{n\times n})^{m_L}=0$$

由于 **J** 为块对角矩阵，因此$(\mathbf{J}-\lambda_l\cdot\mathbf{E}_{n\times n})^{m_l}$, l=1,2,…,L,也为具有和 **J** 相同分块的对角矩阵。

而具有相同分块对角矩阵的乘积也为具有相同分块的对角矩阵，而且乘积的每个对角块为相乘矩阵对角块的乘积。

因此，$m_{\mathbf{J}}(\mathbf{J})=(\mathbf{J}-\lambda_1\cdot\mathbf{E}_{n\times n})^{m_1}\ldots(\mathbf{J}-\lambda_L\cdot\mathbf{E}_{n\times n})^{m_L}=\mathrm{diag}(m_{\mathbf{J}}(\mathbf{J}_1), m_{\mathbf{J}}(\mathbf{J}_2),\ldots,m_{\mathbf{J}}(\mathbf{J}_K))=0$

对于给定的矩阵 A 的特征值λ_l, l=1,2,…,L, 假设对应得若当标准型中的若当块为所对应的若当块为 $\mathbf{J}_i, \mathbf{J}_{i+1},\ldots, \mathbf{J}_{i+s-1}$ 共 s 个。那么对于 $i\le k\le i+s-1$，由

$$m_{\mathbf{J}}(\mathbf{J}_k)=(\mathbf{J}_k-\lambda_1\cdot\mathbf{E}_{r_k\times r_k})^{m_1}\ldots(\mathbf{J}_k-\lambda_L\cdot\mathbf{E}_{r_k\times r_k})^{m_L}$$

$$=\prod_{\substack{j=1\\ j\neq l}}^{L}(\mathbf{J}_k-\lambda_j\cdot\mathbf{E}_{r_k\times r_k})^{m_j}\cdot(\mathbf{J}_k-\lambda_k\cdot\mathbf{E}_{r_k\times r_k})^{m_l}=0.$$

根据性质 1，欲使等式成立，因为等式右边的第一部分为可逆矩阵，因此只有等式右边的第二部分为 0，也就是$(\mathbf{J}_k-\lambda_k\cdot\mathbf{E}_{r_k\times r_k})^{m_l}=0$，根据性质 3，我们得到 $m_l\ge r_k$,　$i\le k\le i+s-1$.

因此，我们可以得到 $m_l\ge\max\{r_i, r_{i+1},\ldots, r_{i+s-1}\}$, $m_{\mathbf{J}}(\mathbf{J}_k)=0$。再根据最小多项式的定义，零化多相式的次数必须是最小的，因此我们得到 $m_l=\max\{r_i, r_{i+1},\ldots, r_{i+s-1}\}$。

这也就说，假如我们知道若当标准型，我们实际上是可以知道最小多项式的。但是，在实际很多情形我们不知道若当标准型，但我们仍然可以知道最小多项式。事实上，由前面的讨论我们知道 m_l 也就是计算数列 $d_m=n-\mathrm{rank}((\lambda_l\cdot\mathbf{E}_{n\times n}-\mathbf{A})^m)$, m=1,2,…, 的值，当 d_m 不再增加时的 m 的值即为 m_l 的值，即

$$d_1<d_2<\cdots<d_m=d_{m+1}=\cdots$$

因此我们得到利用特征矩阵的秩计算最小多项式的方法：

对于 A 的特征值λ_l，计算 $d_m=n-\mathrm{rank}((\lambda\cdot\mathbf{E}_{n\times n}-\mathbf{A})^m)$, m=1,2,…, 的值，

$$d_1<d_2<\cdots<d_m=d_{m+1}=\cdots$$

当 d_m 不再增加时的 m 的值即为最小多项式 $m_{\mathbf{A}}(\lambda)=(\lambda-\lambda_1)^{m_1}\ldots(\lambda-\lambda_L)^{m_L}$ 中包含的因式

$(\lambda-\lambda_l)^{m_l}$ 的次数 m_l.

因此我们给出了通过计算特征矩阵$(\lambda_l\cdot\mathbf{E}_{n\times n}-\mathbf{A})^m$, m=1,2,…,的秩可以确定矩阵若当标准型和最小多项式的。当然，前提条件我们必需知道矩阵的所有不同特征根。

逐次满秩分解计算矩阵幂级数的秩

定义：设 $\mathbf{A}\in C^{n\times n}$ 时，使得 $\mathrm{rank}\mathbf{A}^k=\mathrm{rank}\mathbf{A}^{k+1}$ 成立的最小正整数 k 称为 **A** 的指标。

当矩阵 **A** 阶数较高时，求 **A** 的指标是不容易的。另外一方面，当 A 严重病态时，求 A 的高次幂会使得病态更为严重。因此可以考虑采取 Cline 给出的逐次满秩分解方法求取 A 的指标。该方法每一步都作较小阶矩阵分解，有限次后可以确定出矩阵的指标。

定理　设 $\mathbf{A}\in C^{n\times n}$,　令 **A** 的满秩分解为 $\mathbf{A}=\mathbf{G}_1\mathbf{H}_1$, 而 $\mathbf{H}_i\mathbf{G}_i$ 的满秩分解为

$$\mathbf{H}_i\mathbf{G}_i=\mathbf{G}_{i+1}\mathbf{H}_{i+1}\ (i=1,2,\ldots)$$

则 **A** 的指标为 k 的充要条件是 $\mathbf{H}_k\mathbf{G}_k$ 非奇异。

证明 可以推得 $\mathbf{A}^i=\mathbf{G}_1\ldots\mathbf{G}_i\mathbf{H}_i\ldots\mathbf{H}_1$

$$A^{i+1}=\mathbf{G}_1\ldots\mathbf{G}_i(\mathbf{H}_i\mathbf{G}_i)\mathbf{H}_i\ldots\mathbf{H}_1$$

设 $\mathrm{rank}\mathbf{G}_i=\mathrm{rank}\mathbf{H}_i=r_i$.利用 $\mathbf{G}_i^{(1)}\mathbf{G}_i=\mathbf{I}$, $\mathbf{H}_i\mathbf{H}_i^{(1)}=\mathbf{I}$ $(i=1,2,\ldots,k)$ 得

$$\mathrm{rank}\,\mathbf{A}^i=r_i,\ \mathrm{rank}\mathbf{A}^{i+1}=\mathrm{rank}(\mathbf{H}_i\mathbf{G}_i)$$

从而 $\mathbf{A}$ 的指标为 k 的充要条件为 $\mathbf{H}_k\mathbf{G}_k$ 非奇异。