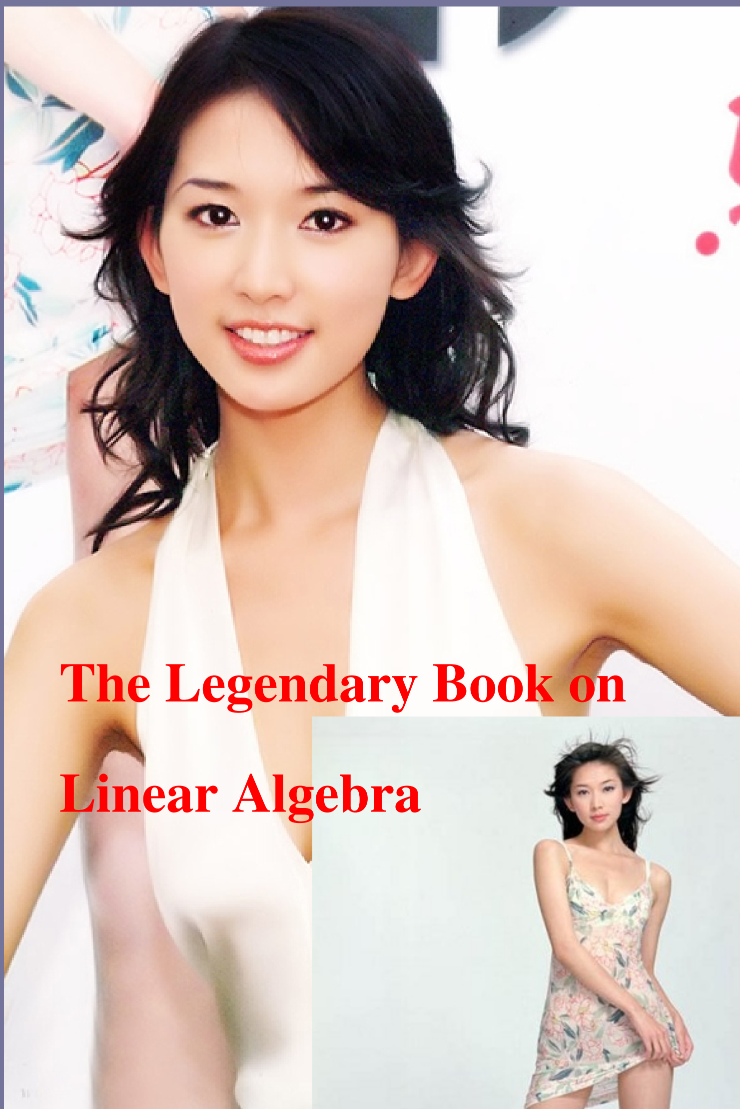
# 高等代数葵花宝典

# The Legendary Book on Linear Algebra



# 目录

# 1

# 将打洞进行到底

打洞是高等代数里面最基本的功夫, 江湖上混的没有不知道的（不知道的都被灭了 ）. 可以说几乎所有的矩阵技巧都离不开打洞, 所以怎样强调它的重要性都不过分. 下面这个定理就告诉了我们什么是打洞.

**定理1.1.** 设

$$M=\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

是一个方阵, 其中 $A$ 是可逆的子方阵, 那么有结论

$$|M|=|A|\cdot|D-CA^{-1}B|.$$

结论不难记, 顺时针走一圈就可以了.

证明思路就是利用 $A$ 的可逆性来“打洞”, 干掉 $B,C$ 之一:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \xrightarrow{\text{第一行左乘以 } -CA^{-1} \text{ 加到第二行}} \begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix},$$

也就是

$$\begin{pmatrix} I_n & 0 \\ -CA^{-1} & I_m \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix},$$

两边求行列式即可.

当然, 如果是 $B,C,D$ 之一可逆也可以类似地打洞.

注意到把上式写成

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}=\begin{pmatrix} I_n & 0 \\ -CA^{-1} & I_m \end{pmatrix}^{-1}\begin{pmatrix} A & B \\ 0 & D-CA^{-1}B \end{pmatrix}$$

的话, 这就很像一个分块的“ LU 分解”, 真正的 LU 分解其实和这个是一回事, 这里就不具体写了.

有时候根据问题的特点, 把打洞倒过来用也很有用处:

**定理1.2.** 设 $A$ 是 $n\times m$ 矩阵, $B$ 是 $m\times n$ 矩阵, 则 $AB$ 和 $BA$ 的特征多项式只差一个因子 $\lambda^{n-m}$, 即

$$\lambda^m|\lambda I_n-AB|=\lambda^n|\lambda I_m-BA|.$$

只需要证明 $\lambda\neq 0$ 的情况. 观察到 $\lambda I_n-AB=\lambda I_n-AI_m^{-1}B$, 这不正好是 $D-CA^{-1}B$ 的形式么? 所以利用前面的结论有

$$|\lambda I_n-AB|=|I_m|\cdot|\lambda I_n-AI_m^{-1}B|=\begin{vmatrix}I_m & B\\ A & \lambda I_n\end{vmatrix}.$$

但是当 $\lambda\neq 0$ 时 $\lambda I_n$ 也是可逆的, 所以（现在变成从 $I_m$ 出发绕一圈）

$$\begin{vmatrix}I_m & B\\ A & \lambda I_n\end{vmatrix}=|\lambda I_n|\cdot|I_m-B\frac{1}{\lambda}I_nA|=\lambda^{n-m}\cdot|\lambda I_m-BA|,$$

即

$$|\lambda I_n-AB|=\begin{vmatrix}I_m & B\\ A & \lambda I_n\end{vmatrix}=\lambda^{n-m}|\lambda I_m-BA|.$$

这就完成了证明.

这个思路也可以采用另一种叙述, 先证明 $\lambda=1$ 的时候结论成立, 然后用 $A/\lambda$ 替换 $A$ 就得到结论对于 $\lambda\neq 0$ 的情形都成立.

$\lambda=1$ 的话问题就变成证明

$$|I_m-AB|=|I_n-BA|.$$

提升以后就是

$$\begin{vmatrix}I_m & B\\ A & I_n\end{vmatrix}.$$

事情很简单, 结论很显然.

打洞有很多很多很多....重要的应用, 特别是当 $M$ 是对称矩阵的时候, 如果你用 $A$ 打两次洞干掉 $B$ 和 $C$ 就会发现这恰好是一个合同变换:

$$\begin{pmatrix}I_n & 0\\ -CA^{-1} & I_m\end{pmatrix}\begin{pmatrix}A & B\\ C & D\end{pmatrix}\begin{pmatrix}I_n & -A^{-1}B\\ 0 & I_m\end{pmatrix}=\begin{pmatrix}A & 0\\ 0 & D-CA^{-1}B\end{pmatrix}.$$

其中左边行变换打掉 $C$, 右边列变换打掉 $B$. 不难验证当 $M$ 对称时有

$$\begin{pmatrix}I_n & 0\\ -CA^{-1} & I_m\end{pmatrix}=\begin{pmatrix}I_n & -A^{-1}B\\ 0 & I_m\end{pmatrix}'$$

成立. 特别强调的是, 对称矩阵 $M$ 的打洞有特别重要的意义: 由于 $M$ 可以看作一个“内积”的度量矩阵, 所以两边打洞实际上就是在这个“内积”下做 Schmidt 正交化. 化二次型为标准形的配方法和矩阵法都源自于此. 这里简要描述一下矩阵方法的算法, 详细的叙述请查阅教科书.

**定理1.3.** [化二次型为标准形的算法] 设 $A=(a_{ij})$ 是一个 $n$ 阶对称矩阵, 现在要把它合同为对角形.
(1) 如果 $a_{11}\neq 0$, 那就用 $a_{11}$ 两次打洞“合同”掉第一行和第一列的其它元素, 然后考虑右下角的 $n-1$ 阶的矩阵.
(2) 如果 $a_{11}=0$ 但是某个 $a_{ii}\neq 0$, 那就交换第 $i$ 行和第 1 行, 交换第 $i$ 列和第 1 列, 把 $a_{ii}$ 变到 $a_{11}$ 的位置上来, 然后执行步骤 (1).
(3) 如果 $A$ 的对角线上都是 0, 但是某个 $a_{ij}\,(i<j)$ 不是 0, 那就把第 $j$ 行加到第 $i$ 行, 第 $j$ 列加到第 $i$ 列, 这样 $a_{ii}$ 的位置上就出现了 $2a_{ij}$, 从而可以执行步骤 (2), 进而执行步骤 (1).

由于每进行一次步骤 (1) 要处理的矩阵都降了一阶, 所以经过有限步以后就可以把 $A$ 变成对角形.

这个算法说白了就是一句话: 制造非零的对角元来打洞“合同”掉非对角元. 其实就是不断地做 Schmidt 正交化.

正定矩阵是最容易化为标准形的对称矩阵. 因为正定矩阵的对角元总不是 0（想一想, 为什么? ）, 所以只需要步骤 (1) 就可以化为标准形. 半正定矩阵的打洞也很简单, 虽然对角元可能出现 0, 但是我们有下面的引理:

**引理1.4.** 如果半正定矩阵 $A$ 的某个对角元是 0, 那么该对角元所在的行和列所有元素都是 0.

证明: 由于 $A$ 半正定, 所以有平方根分解 $A=P'P$. 记 $P=\{v_1,\dots,v_n\}$, 则 $A=\big((v_i,v_j)\big)$. $a_{ii}=0$ 说明 $v_i=0$, 从而第 $i$ 行第 $i$ 列都是 0.

所以半正定矩阵化为标准形本质上也只需要步骤 (1), 只不过对角线上遇到 0 的时候不用打洞, 自动跳过去继续考虑右下角的矩阵.

**定理1.5.** 设

$$M=\begin{pmatrix}A & B\\ C & D\end{pmatrix}$$

是一个半正定矩阵, 则可以用合同变换干掉 $B,C$, 把 $M$ 化为准对角形

$$M=\begin{pmatrix}A & 0\\ 0 & *\end{pmatrix}.$$

注意不需要 $A$ 可逆的假设!

道理已经讲过了, 把算法描述一下:
(1) 如果 $a_{11}\neq 0$, 则打洞消去 $B$ 的第一行和 $C$ 的第一列并考虑 $a_{22}$. 如果 $a_{11}=0$, 根据引理 1.4, $B$ 的第一行和 $C$ 的第一列本来就是 0, 什么也不用干, 自动考虑 $a_{22}$.
(2) 重复上面的方法, 如果 $a_{22}\neq 0$, 则打洞消去 $B$ 的第二行和 $C$ 的第二列并考虑 $a_{33}$. 如果 $a_{22}=0$, 什么也不用干, 自动考虑 $a_{33}$.

...
这样从 $a_{11}$ 一直到 $a_{nn}$, 最多 $n$ 次就能合同掉 $B$ 和 $C$. 注意 $A$ 不变, 而 $D$ 在变.

有了上面这些准备, 下面这个定理就变得很容易了.

➤ **定理1.6.** 两个 $n$ 阶半正定矩阵 $A, B$ 可以同时合同于对角形, 即存在可逆矩阵 $T$ 使得 $T'AT, T'BT$ 都是对角矩阵.

证明: 首先做合同变换把 $A$ 化成标准形

$$A \sim \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix},$$

这时 $B$ 仍然是半正定的. 所以不妨从一开始就假设 $A$ 就是如上的标准形, 并设

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \quad B_{12} = B_{21}'.$$

我们要在保持 $A$ 的形状的前提下把 $B$ 化成标准形, 根据上面的定理, 可以用 $B_{22}$ 的对角元打洞合同掉 $B_{21}$ 和 $B_{12}$, 把 $B$ 化为

$$\begin{pmatrix} * & 0 \\ 0 & B_{22} \end{pmatrix}.$$

由于每一步都是把后面的行或列加到前面的行或列上去, 所以均保持 $A$ 的形状不变. 现在只需要在保持

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

形状的前提下把

$$B \sim \begin{pmatrix} * & 0 \\ 0 & B_{22} \end{pmatrix}$$

对角化. 找两个合适的正交矩阵 $T, S$,

$$\begin{pmatrix} T & 0 \\ 0 & S \end{pmatrix}$$

就可以办到.

定理证完了. 怎么样, 会打洞是不是很爽呢? 总的讲, 打洞要天天打, 月月打, 见题就打. 特别是对称矩阵的时候, 没有条件创造条件也要打.

➤ **练习1.1.** 设 $A$ 是正定矩阵, 求证 $|A| \le a_{11}a_{22}\cdots a_{nn}$. 这个结论有什么几何意义?

➤ **练习1.2.** 书上有这样一个定理: 对称矩阵 $A$ 是正定的当且仅当 $A$ 的顺序主子式都大于 0. 看看这个是怎样打洞的? 和 LU 分解定理比较一下.

2

# Jordan 块的运算

Jordan 块形式比较简单, 而且有比较好的运算规则, 用它来讨论往往可以起到简化的效果, 所以熟悉 Jordan 块的运算很重要. 下面是一个简单的介绍.

## Jordan 块的多项式

想来你一定知道这个小规律（不知道的用力将头撞墙三下  ）：设

$$N = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{n\times n},$$

那么 $N^2$ 就是把 $N$ 中的 1 向右上方平移一层, $N^k$ 就平移 $k-1$ 层, $N^n$ 就变成零矩阵了. 用这个规律我们可以很快算出一般的 Jordan 块

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

的多项式来: 对于给定的 $m$ 次多项式 $f(x)$, 在 $\lambda_0$ 点作 $f$ 的 Taylor 展开: $f(x) = a_0 + a_1(x-\lambda_0) + \cdots + a_m(x-\lambda_0)^m$, 那么

$$f(J) = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ 0 & a_0 & \ddots & \vdots \\ \vdots & & \ddots & a_1 \\ 0 & 0 & \cdots & a_0 \end{pmatrix}.$$

可见 Jordan 块的多项式是很好算的, 其形状是一个上三角的分层矩阵.

### Jordan 块多项式的标准形

继续, 怎样求 Jordan 块的多项式的 Jordan 标准形呢? 沿用上一小节的符号, 记

$$A = f(J) - f(\lambda_0)I = \begin{pmatrix} 0 & a_1 & \cdots & a_{n-1} \\ & 0 & \ddots & \vdots \\ & & \ddots & a_1 \\ & & & 0 \end{pmatrix},$$

根据计算公式我们知道 $f(J)$ 的 Jordan 标准形中阶为 $i$ 的 Jordan 块的个数是 $\mathrm{r}(A^{i+1})+\mathrm{r}(A^{i-1})-2\mathrm{r}(A^i)$, 所以下面的任务就是计算 $r_i = \mathrm{r}(A^i)$. 关键是注意到上三角分层矩阵的秩由它最靠近主对角线的非零的那一“层”决定, 与更高的“层”无关. 设 $a_k$ 所在的层是最靠近主对角线的非零的层, 由于 $A = a_kN^k + \cdots + a_{n-1}N^{n-1}$, 所以

$$A^i = (a_kN^k + \cdots + a_{n-1}N^{n-1})^i = a_k^iN^{ki} + \cdots \text{（更高次项）}.$$

作带余除法 $n = qk + r$, 这里 $0 \le r < k$. 显然

$$r_i = \mathrm{r}(A^i) = \begin{cases} n - ki & 1 \le i \le q, \\ 0 & i > q. \end{cases}$$

那么根据 Jordan 块个数的计算公式, $f(J)$ 的阶是 $i$ 的 Jordan 块的个数是 $r_{i+1} + r_{i-1} - 2r_i$, 得出下面的表格:

| 阶数 | $i \le q-1$ | $i = q$ | $i = q+1$ | $i \ge q+1$ |
|---|---|---|---|---|
| 个数 | 0 | $k-r$ | $r$ | 0 |

这结果很有意思, $f(J)$ 分解为阶数相邻或者相同的一些 Jordan 块, 而且 $\lambda_0$ 是 $f(x) - f(\lambda_0)$ 的几阶零点 $f(J)$ 就会分解为几块. 下面是两个值得记住的结论:

**推论2.1.** 可逆矩阵 $A$ 的任意次幂 $A^m$ 与 $A$ 有相同的 Jordan 形, 仅仅把对角元换成 $\lambda_i^m$.

这是因为这时候 $k = 1, q = n, r = 0$.

**推论2.2.** 对于不可逆矩阵, 如果特征值 0 对应的 Jordan 块都是一阶的, 那么它的平方与 $A$ 有相同的 Jordan 形, 仅仅把对角元换成 $\lambda_i^2$; 如果特征值 0 有阶数大于等于 2 的 Jordan 块, 那么如果这个块是 $2q$ 阶的, 就分解为 2 个 $q$ 阶的子块, 如果它是 $2q+1$ 阶的, 就分解为一个 $q$ 阶和一个 $q+1$ 阶的子块.

这是因为 Jordan 块的阶 $n \ge 2$ 的时候 $a_2 \ne 0, k = 2$, 所以分 $n$ 是奇偶两种可能.

### 与 Jordan 块交换的矩阵

设

$$J_n=\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1\\ & & & \lambda\end{pmatrix}_{n\times n},\quad J_m=\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1\\ & & & \lambda\end{pmatrix}_{m\times m}.$$

这一小节来解矩阵方程 $J_nX=XJ_m$. 设 $X=(x_{ij})$, 把矩阵方程写成

$$\lambda I_nX+J_n(0)X=\lambda XI_m+XJ_m(0),$$

即

$$J_n(0)X=XJ_m(0).$$

由于

$$[J_n(0)X]_{ij}=\begin{cases}x_{i+1,j} & i=1,2,\dots,n-1.\\ 0 & i=n.\end{cases}$$

$$[XJ_m(0)]_{ij}=\begin{cases}x_{i,j-1}, & j=2,3,\dots,m.\\ 0 & j=1.\end{cases}$$

所以如果给 $X$ 补上第 0 列和第 $n+1$ 行, 并规定这一列和这一行元素都是 0, 那么就有

$$x_{i+1,j+1}=x_{ij},\quad i=1,2,\dots,n,\ j=0,1,\dots,m-1.$$

从而 $X$ 就形如

$$[0\quad Y]\quad 或者\quad \begin{bmatrix}Y\\0\end{bmatrix}.$$

这里 $Y$ 是一个 $r=\min\{n,m\}$ 的上三角分层方阵, 形状为

$$Y=\begin{pmatrix}a_0 & a_1 & \cdots & a_{r-1}\\ 0 & a_0 & \ddots & \vdots\\ \vdots & & \ddots & a_1\\ 0 & 0 & \cdots & a_0\end{pmatrix}_{r\times r}.$$

零矩阵是 $(n-m)\times m$ 或者 $n\times(m-n)$ 阶的, 对应于 $n\geq m$ 和 $n\leq m$ 两种情形.

很容易看出来 $X$ 中含有 $r$ 个未知的参数, 所以矩阵方程 $J_nX=XJ_m$ 的解空间是 $r$ 维的. 不仅如此, 当 $m=n$ 时,

$$X=\begin{pmatrix}a_0 & a_1 & \cdots & a_{n-1}\\ 0 & a_0 & \ddots & \vdots\\ \vdots & & \ddots & a_1\\ 0 & 0 & \cdots & a_0\end{pmatrix}_{n\times n}=a_0I+a_1(J_n-\lambda I)+\cdots+a_{n-1}(J_n-\lambda I)^{n-1}$$

是 $J_n$ 的多项式, 这就说明与一个 Jordan 块 $J$ 交换的矩阵可以表示为 $J$ 的多项式.

### 实 Jordan 块开平方

这一节讨论的是 $\lambda_0$ 是实数的时候 $J$ 能否在实数范围内开平方的问题. 即是否存在实矩阵 $A$ 使得 $A^2=J$. 约定 $J$ 的阶 $n>1$, 我们有下面的结论

**定理2.3.** 当 $\lambda_0$ 大于 0 时 $J$ 可以在实数范围开平方, 小于等于 0 时则不能.

证明: 当 $\lambda_0>0$ 的时候有一个巧妙的办法给出了求 $A$ 的算法: 记$J=\lambda_0 I+N$, 则 $N^n=0$ 考虑 $\sqrt{x+\lambda_0}$ 的 $n$ 次 Taylor 多项式 $P(x)$, 则 $P(x)^2-(x+\lambda_0)=x^nQ(x)$, 而且 $Q(x)$ 也是一个多项式（不要以为是个幂级数）. 我们还有 $P(x),Q(x)$ 都是实系数的. 那么

$$P(N)^2=N+\lambda_0 I+N^nQ(N)=N+\lambda_0 I=J,$$

所以 $A=P(N)$ 满足要求. 当 $\lambda_0=0$ 的时候可以这样做, 设 $A^2=N$, 那么 $A^{2n-2}=N^{n-1}\neq 0$, $A^{2n}=N^n=0$. 但是 $A^{2n}=0$ 意味着 $A^n=0$（极小多项式次数不会大于 $n$）, 从而 $n\geq 2$ 时有 $A^{2n-2}=A^nA^{n-2}=0$, 矛盾!

最后我想多说一点废话. Jordan 标准形的理论告诉我们复数域上的线性变换本质上只有两种, 那就是数乘变换和移位变换:

$$A\varepsilon_1=\varepsilon_2,\ A\varepsilon_2=\varepsilon_3,\ \cdots,\ A\varepsilon_{n-1}=\varepsilon_n,\ A\varepsilon_n=0.$$

复数域上的任何线性变换都可以分解为数乘变换和移位变换的组合, 而数乘变换是很平凡的变换, 所以可以说移位运算是最重要的线性变换. Halmos 说过在泛函分析里面最重要的泛函仍然是移位运算（至少你要想想紧算子, Toeplitz 算子, Hankel 算子）.

不说泛函了, 我懂的不多, 再说就露馅了. 不过用移位算子来看 Jordan 块的碎裂就很自然了. 比如一个 $2q$ 阶的 0 特征值的 Jordan 块 $A$, 它的轨道本来是传递的, 但是 $A^2$ 的轨道则是在跳:

$$\begin{aligned}&A^2\varepsilon_1=\varepsilon_3, \quad A^2\varepsilon_3=\varepsilon_5, \quad \cdots, \quad A^2\varepsilon_{2n-3}=\varepsilon_{2n-1}, \quad A^2\varepsilon_{2n-1}=0.\\&A^2\varepsilon_2=\varepsilon_4, \quad A^2\varepsilon_4=\varepsilon_6, \quad \cdots, \quad A^2\varepsilon_{2n-2}=\varepsilon_{2n}, \quad A^2\varepsilon_{2n}=0.\end{aligned}$$

所以 $A^2$ 的轨道有两条, 相应地 Jordan 块就有两个.

看来线性变换和群在集合上的作用是很相似的, 就是一些不动点（相当于特征向量）和非平凡的轨道（Jordan 块）的组合. 把作用限制在子群上（相当于子模 $A^k$）轨道还可以继续分解. 有些矩阵不能对角化的原因就在于它们有非平凡的轨道. 或者说, 组成成分里含有移位算子.

3

# 秩不等式

秩不等式问题有比较模式化的解决方法, 就是利用分块矩阵的初等变换, 其原理就是下面这个引理:

**引理3.1.** 设

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

其中 $A,B$ 都是方阵, 那么 $\mathrm{r}(M)\geq \mathrm{r}(A)+\mathrm{r}(B)$.

证明思路就是先把 $A,B$ 都化成标准形: 设 $P_1AQ_1=E_r$, $P_2BQ_2=E_s$, 那么

$$\begin{pmatrix} P_1 & 0\\ 0 & P_2\end{pmatrix}\begin{pmatrix} A & 0\\ C & B\end{pmatrix}\begin{pmatrix} Q_1 & 0\\ 0 & Q_2\end{pmatrix}=\begin{pmatrix} E_r & 0\\ D & E_s\end{pmatrix}.$$

然后用 $E_r$ 干掉 $D$ 的前 $r$ 列, 用 $E_s$ 干掉 $D$ 的后 $s$ 行, $M$ 就变成

$$\begin{pmatrix} E_r & 0 & 0\\ 0 & * & 0\\ 0 & 0 & E_s\end{pmatrix},$$

这样就证明了引理.

在证明有关秩的不等式的时候, 基本思路就是从准对角矩阵出发, 用初等变换化成三角形的分块矩阵, 来应用上面这个引理.

下面是秩不等式里面最基本的结论:

➤ **定理3.2** (Frobenius **不等式**)**.** 设 $A$ 是数域 $F$ 上的 $m\times n$ 矩阵, $B$ 是 $n\times k$ 矩阵, $C$ 是 $k\times s$ 矩阵, 则

$$\mathrm{r}(AB)+\mathrm{r}(BC)\leq \mathrm{r}(ABC)+\mathrm{r}(B).$$

证明:

$$\begin{pmatrix} ABC & 0\\ 0 & B\end{pmatrix}\to\begin{pmatrix} ABC & AB\\ 0 & B\end{pmatrix}\to\begin{pmatrix} 0 & AB\\ -BC & B\end{pmatrix}.$$

比较两段的矩阵的秩即可.

下面是论坛中讨论过的比较有代表性的两个问题.

**例题(北京大学2005).** 设 $A$ 是数域 $F$ 上 $n$ 维向量空间上的线性变换, 求证

$$A^3=I \Leftrightarrow \mathrm{r}(I-A)+\mathrm{r}(I+A+A^2)=n.$$

注意到存在 $u(x)$, $v(x)$ 使得 $u(x)(1-x)+v(x)(1+x^2+x)=1$, 考虑大的 $2n$ 阶的方阵

$$\begin{aligned}\begin{pmatrix} I-A & 0 \\ 0 & I+A+A^2 \end{pmatrix} &\xrightarrow{\text{第一列乘以 } u(A) \text{ 加到第二列}} \begin{pmatrix} I-A & u(A)(I-A) \\ 0 & I+A+A^2 \end{pmatrix}\\ &\xrightarrow{\text{第二行乘以 } v(A) \text{ 加到第一列}} \begin{pmatrix} I-A & I \\ 0 & I+A+A^2 \end{pmatrix}\\ &\xrightarrow{\text{用 } I \text{ 打洞干掉两边的矩阵}} \begin{pmatrix} 0 & I \\ A^3-I & 0 \end{pmatrix}.\end{aligned}$$

所以这个矩阵的秩是 $n$ 当且仅当 $A^3-I=0$, 这就得到了证明.

**例题(科大教材习题).** 设 $A,B$ 是数域 $F$ 上的 $n$ 阶方阵且满足 $AB=BA=0, \mathrm{r}(A)=\mathrm{r}(A^2)$, 求证

$$\mathrm{r}(A+B)=\mathrm{r}(A)+\mathrm{r}(B).$$

思路: 只要证明 $\mathrm{r}(A+B)\geq \mathrm{r}(A)+\mathrm{r}(B)$ 即可. 用老办法, 从

$$\begin{pmatrix} A+B & 0 \\ 0 & 0 \end{pmatrix}$$

出发, 通过初等变换化为形如

$$\begin{pmatrix} B & * \\ 0 & A \end{pmatrix}$$

的矩阵, 就得到了结论. 在变换前先注意两点:
(1) 由于 $A^2$ 的列向量都是 $A$ 的列向量的线性组合, 所以 $\mathrm{r}(A)=\mathrm{r}(A^2)$ 说明 $A$ 的列向量组和 $A^2$ 的列向量组是等价的, 也就是 $A$ 的列向量可以被 $A^2$ 的列向量线性表示出来. 设 $\alpha_i$ 是 $A$ 的第 $i$ 个列向量, 那么线性方程组 $A^2X=\alpha_i$ 有解 $X_i$, 令 $n\times n$ 矩阵 $P=\{X_1,\ldots,X_n\}$, 那么 $A^2P=A$.
(2) 当 $\mathrm{r}(A)=\mathrm{r}(A^2)$ 时有 $\mathrm{r}(A)=\mathrm{r}(A^2)=\mathrm{r}(A^3)=\cdots$
这个在 Frobenius 不等式

$$\mathrm{r}(B)+\mathrm{r}(ABC)\geq \mathrm{r}(AB)+\mathrm{r}(BC)$$

中令 $B=C=A$ 即可.

下面进行变换:

$$\begin{aligned}\begin{pmatrix} A+B & 0 \\ 0 & 0 \end{pmatrix} &\xrightarrow{\text{第一行左乘以 } A \text{ 加到第二行}} \begin{pmatrix} A+B & 0 \\ A^2 & 0 \end{pmatrix}\\ &\xrightarrow{\text{第一列右乘以 } A \text{ 加到第二列}} \begin{pmatrix} A+B & A^2 \\ A^2 & A^3 \end{pmatrix}\\ &\xrightarrow{\text{第二列右乘以 } -P \text{ 加到第一列}} \begin{pmatrix} B & A^2 \\ 0 & A^3 \end{pmatrix}.\end{aligned}$$

从而 $\mathrm{r}(A+B) \geq \mathrm{r}(B)+\mathrm{r}(A^3)=\mathrm{r}(B)+\mathrm{r}(A)$, 得证.

下面再给一种解法:

可以假设 $A$ 是 Jordan 标准形, 那么 $\mathrm{r}(A)=\mathrm{r}(A^2)$ 说明 $A$ 的 0 特征值对应的 Jordan 块都是一阶的, 从而 $A$ 形如

$$\begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix},$$

这里 $A_1$ 是可逆的. $AB=BA=0$ 说明 $B$ 形如

$$\begin{pmatrix} 0 & 0 \\ 0 & B_1 \end{pmatrix},$$

这样就得到了证明.

大家也可以去看看当初论坛上的讨论:

http://old.math.org.cn/forums/index.php?showtopic=63488&hl=

**定理3.3.** 设 $A_1,\dots,A_m$ 为 $n$ 阶方阵, 适合条件 $A_1+A_2+\cdots+A_m=I$. 试证明下面三个条件等价:

(1) $A_1,\dots,A_m$ 都是幂等矩阵;

(2) $\mathrm{r}(A_1)+\mathrm{r}(A_2)+\cdots+\mathrm{r}(A_m)=n$;

(3) 当 $1\leq i<j\leq n$ 时有 $A_iA_j=A_jA_i=0$.

证明: (1) ⇒ (2): 用熟知的结论, 幂等矩阵的秩等于它的迹, 可得

$$\sum_{i=1}^m \mathrm{r}(A_i)=\sum_{i=1}^m \mathrm{tr}(A_i)=\mathrm{tr}(I_n)=n.$$

(2) ⇒ (3): 考虑大的准对角方阵

$$\begin{pmatrix} 0 & & & \\ & A_1 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix} \xrightarrow{\text{都加到第一行}} \begin{pmatrix} 0 & A_1 & \cdots & A_m \\ & A_1 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix} \xrightarrow{\text{都加到第一列}} \begin{pmatrix} I_n & A_1 & \cdots & A_m \\ A_1 & A_1 & & \\ \vdots & & \ddots & \\ A_m & & & A_m \end{pmatrix}$$

$$\xrightarrow{\text{打洞干掉第一行和第一列}} \begin{pmatrix} I_n & 0 & \cdots & 0 \\ 0 & A_1-A_1^2 & \cdots & -A_1A_m \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -A_mA_1 & \cdots & A_m-A_m^2 \end{pmatrix}$$

所以 $\mathrm{r}(A_1)+\mathrm{r}(A_2)+\cdots+\mathrm{r}(A_m)=n$ 当且仅当右下角的大的矩阵为0, 即 $A_i$ 都是幂等矩阵而且互相正交.

(3) ⇒ (1): too easy, 省略.

定理 3.3 是很常考的一个定理, 因为它是数理统计中的很基本的 Cochran 定理的矩阵版本:

**定理3.4** (Cochran **定理**). 设 $X_1,\dots,X_n$ 是相互独立且服从正态分布 $N(0,1)$ 的随机变量, 若
$$Q_1+Q_2+\cdots+Q_k=\sum_{i=1}^{n}X_i^2,$$
其中 $Q_i$ 是秩为 $n_i$ 的 $X_1,\dots,X_n$ 的二次型, 则下面三个条件等价:
(1) $Q_1,\dots,Q_k$ 相互独立;
(2) $Q_i\sim\chi^2(n_i)$.
(3) $\sum\limits_{i=1}^{k}n_i=n$.

**例题**(北京大学2007). $A,B$ 是两个 $n$ 阶矩阵满足 $AB=BA$. 求证
$$\mathrm{r}(A)+\mathrm{r}(B)\ge\mathrm{r}(A+B)+\mathrm{r}(AB).$$

证明: 设 $X$ 是齐次线性方程组 $AX=0$ 的解空间, $Y$ 是齐次线性方程组 $BX=0$ 的解空间, $Z$ 是齐次线性方程组 $ABX=BAX=0$ 的解空间, $W$ 是齐次线性方程组 $(A+B)X=0$ 的解空间, 那么我们有 $X\subset Z, Y\subset Z$. 从而 $X+Y\subset Z$. 而且 $X\cap Y\subset W$. 由维数公式,
$$\dim X+\dim Y=\dim X\cap Y+\dim(X+Y)\le\dim W+\dim Z.$$
从而
$$n-\mathrm{r}(A)+n-\mathrm{r}(B)\le n-\mathrm{r}(A+B)+n-\mathrm{r}(AB),$$
即 $\mathrm{r}(A)+\mathrm{r}(B)\ge\mathrm{r}(A+B)+\mathrm{r}(AB)$.

记得以前在论坛上有个网友发了一个帖子, 说他写了一篇论文, 证明了这样一个结论: 设 $A,B$ 都是 $n$ 阶方阵, 而且 $AB=BA$, 那么有如下的不等式成立:
$$\mathrm{r}(A^2)+\mathrm{r}(B^2)\ge 2\mathrm{r}(AB).$$
而且他说这是均值不等式在矩阵里的推广. 请大家思考一下这个问题, 看看结论成立不?

# 4

# 交结数: 刻画相似程度的不变量

以下都约定 $F\subset K$ 是两个数域, $A,B$ 分别是 $F$ 上的 $n$ 阶, $m$ 阶方阵. 这一节要讨论的是矩阵方程 $AX=XB$. 具体一点, 是这个矩阵方程解空间的维数. 这个维数又叫做 $A$ 与 $B$ 的交结数（interwine number）, 用 $i(A,B)$ 来表示. 下面就会看到, 交结数是一个刻画 $A$ 和 $B$ 相似程度的量, 它包含了很丰富的信息.

首先要指出的是交结数是个不依赖域的概念. 换句话说, 设

$$U_F=\left\{\ X\in \mathrm{Mat}_{n\times m}(F)\ \middle|\ AX=XB\ \right\},$$

$$U_K=\left\{\ Y\in \mathrm{Mat}_{n\times m}(K)\ \middle|\ AY=YB\ \right\},$$

那么我们有

$$\dim U_F=\dim U_K=i(A,B).$$

这个道理很简单, 解矩阵方程就是解齐次线性方程组, 大家求基础解系的过程是一样的, 在 $F$ 内就能完成. 只不过 $U_F$ 是基础解系的 $F-$ 线性组合, $U_K$ 是基础解系的 $K-$ 线性组合.

由上面的讨论可以得到

➤ **定理4.1.** [相似性和域无关] 如果 $A,B$ 在 $K$ 上相似, 它们就在 $F$ 上也相似.

证明: $A,B$ 在 $K$ 上相似说明矩阵方程 $AX=XB$ 在 $K$ 上有一个可逆解, 现在要证明它在 $F$ 上也有一个可逆解. 设 $X_1,\dots,X_s$ 是 $U_F$ 的一组基, 那么它们也构成 $U_K$ 的一组基. $AX=XB$ 在 $K$ 上有一个可逆解说明存在 $X_1,\dots,X_s$ 的 $K-$ 线性组合

$$a_1X_1+a_2X_2+\cdots+a_sX_s,\quad a_i\in K$$

使得 $a_1X_1+a_2X_2+\cdots+a_sX_s$ 可逆. 我们要证明一定有 $X_1,\dots,X_s$ 的 $F-$ 线性组合

$$b_1X_1+b_2X_2+\cdots+b_sX_s,\quad b_i\in F$$

使得 $b_1X_1+b_2X_2+\cdots+b_sX_s$ 可逆. 自然我们要考察 $s$ 个变元 $t_1,\dots,t_s$ 的多元多项式

$$f(t_1,t_2,\cdots,t_s)=\det|t_1X_1+t_2X_2+\cdots+t_sX_s|.$$

$f$ 是一个数域 $F$ 上的多元多项式（当然也是 $K$ 上的）, $f(a_1,a_2,\cdots,a_s)\neq 0$ 说明 $f$ 不是零多项式. 由于数域有无穷多个元素, 所以必定存在 $F$ 中的 $b_1,\dots,b_s$ 使得 $f(b_1,b_2,\cdots,b_s)\neq 0$（想一想为什么? ）. 这就说明了 $A,B$ 在 $F$ 上也是相似的.

你也许要说, 真是一堆 , 用不变因子直接不就得出结论吗? 还用不着数域的假设. 可是你再看看下面这个问题:

➤ **定理4.2.** 设 $A_1,\dots,A_m,B_1,\dots,B_m$ 是 $F$ 上的 $n$ 阶方阵, 如果存在 $K$ 上的可逆矩阵 $S$ 使得 $S^{-1}A_iS=B_i\,(i=1,2,\dots,m)$, 那么必存在 $F$ 上的可逆矩阵 $T$ 使得 $T^{-1}A_iT=B_i\,(i=1,2,\dots,m)$. 这里同样不需要 $F$ 是数域的假设.

这个定理来自群表示论, 意义就是两个表示在 $K$ 上等价就在 $F$ 上等价. 显然这个时候不变因子的概念就不好使了, 因为没法说明 $T$ 是"公共"的. 但是只要 $F$ 有无限多个元素, 定理 4.1 的办法就能用, 所以它还是很有应用前景的.

证明: 沿用前面的记号, 考虑向量空间

$$U_F=\left\{\,T\in \mathrm{Mat}_n(F)\,\middle|\,A_iT=TB_i,\ i=1,2,\dots,m\,\right\},$$

$$U_K=\left\{\,S\in \mathrm{Mat}_n(K)\,\middle|\,A_iS=SB_i,\ i=1,2,\dots,m\,\right\}.$$

其实就是把 $m$ 个线性方程组联立起来. 所以和上面同样的道理, $U_F$ 和 $U_K$ 的维数是一样的. 设 $T_1,\dots,T_s$ 是 $U_F$ 和 $U_K$ 共同的一组基, 考察多元多项式

$$f(t_1,t_2,\cdots,t_s)=\det|\,t_1T_1+t_2T_2+\cdots+t_sT_s\,|,$$

$S$ 的存在说明 $f$ 不是零多项式, 基域有无限多个元素就保证了存在 $(a_1,\dots,a_s)\neq(0,\dots,0)$ 使得 $f(a_1,a_2,\cdots,a_s)\neq 0$, 所以 $T=a_1T_1+a_2T_2+\cdots+a_sT_s$ 满足要求.

至于 $F$ 是有限域的情形, 要用到比较深的知识, 文档里就不写了.

接下来是个很有用的结论:

➤ **定理4.3.** 如果 $A,B$ 没有共同的特征值, 那么矩阵方程 $AX=XB$ 只有零解.

证明: 设 $X$ 满足 $AX=XB$, 那么 $A^2X=AXB=XB^2$, 用归纳法可得对任何正整数 $k$ 都有 $A^kX=XB^k$. 从而对任何多项式 $f(x)$ 都有 $f(A)X=Xf(B)$. 特别令 $f(x)$ 是 $A$ 的特征多项式, 由 Hamilton-Cayley 定理, $f(A)=0$. 但是这时 $f(B)$ 是可逆矩阵, 必有 $X=0$!（想一想, 为什么? ）

这个结论有个很好的应用, 就是当一个矩阵 $A$ 形如

$$\begin{pmatrix}A_1&0\\0&A_2\end{pmatrix}$$

（$A_1$ 和 $A_2$ 没有共同的特征值）的时候, 与 $A$ 交换的矩阵 $B$ 也必然形如

$$\begin{pmatrix}B_1&0\\0&B_2\end{pmatrix}$$

（验证一下!）. 反复地应用这个结论可以推广到对角线上有 $n$ 个矩阵的情形. 换言之, 当把 $A$ 按特征值进行分块的时候与 $A$ 交换的矩阵 $B$ 也就同时被分块了, 这就起到了简化问题的作用. 例题 3 的第二个证明就是用的这一招, 而且这一招还会反复出现, 所以一定要注意.

最后来给出交结数的确切值:

➤ **定理4.4.**

$$i(A,B)=\sum_{i=1}^{p}\sum_{j=1}^{q}\deg\Big(f_i(x),g_j(x)\Big).$$

这里 $f_i(x)$ 跑遍 $A$ 的不变因子, $g_j(x)$ 跑遍 $B$ 的不变因子.

证明大意: 可以在 $A,B$ 都是 Jordan 标准型的前提下进行计算,设

$$A=\begin{pmatrix}J_{k_1}(\lambda_1)&&&\\&J_{k_2}(\lambda_2)&&\\&&\ddots&\\&&&J_{k_r}(\lambda_r)\end{pmatrix},\quad B=\begin{pmatrix}L_{m_1}(\mu_1)&&&\\&L_{m_2}(\mu_2)&&\\&&\ddots&\\&&&L_{m_s}(\mu_s)\end{pmatrix}.$$

相应地把 $X$ 分为 $rs$ 个子块 $X_{ij}$, $X_{ij}$ 的阶是 $k_i\times L_j$. 那么 $AX=XB$ 就变成了

$$J_iX_{ij}=X_{ij}L_j.\quad i=1,2,\dots,r.\ j=1,2,\dots,s.$$

整个矩阵方程 $AX=XB$ 的解空间的维数就是各个方程 $J_iX_{ij}=X_{ij}L_j$ 的解空间维数的和. 根据第 2 章的结论, 矩阵方程 $J_iX_{ij}=X_{ij}L_j$ 的解空间的维数是

$$\gcd\Big((x-\lambda_i)^{k_i},(x-\mu_j)^{m_j}\Big)$$

的次数. 所以矩阵方程 $AX=XB$ 的解空间的维数就是两两求出 $A$ 和 $B$ 的初等因子的最大公因式, 然后把次数加起来. 进一步这个结果还等于两两求出不变因子的最大公因式, 然后把次数相加, 这就证明了定理.

可以看到, $A$ 和 $B$ “相似”程度越高, 交结数就越大, 可以说交结数刻画了 $A$ 和 $B$ 之间的相似程度. 用模的语言可以更好的解释清楚: 交结数就是从 $A-$ 模 $V$ 到 $B-$ 模 $W$ 的模同态组成的向量空间的维数. 至于定理 4.3, 可以解释为当矩阵方程有非零解的时候, 由模同态基本定理, 有模同构

$$W/\mathrm{Ker}X\cong\mathrm{Im}X.$$

即 $B$ 在某个商空间上的诱导变换和 $A$ 在某个子空间上的限制是相似的, 从而 $A,B$ 有共同的特征值.

➤ **练习4.1.** 已知方阵 $A,B$ 的特征值都是正数且满足 $A^2=B^2$, 求证 $A=B$.

也许最重要的不是答案, 而是想法, 建议大家看一下论坛上的讨论:

http://old.math.org.cn/forums/index.php?showtopic=63012

看看你找到了多少种证明.

5

# 覆盖定理

➤ **定理5.1.** 设基域 $F$ 有无穷多个元素（比如数域）, 那么 $F$ 上的有限维向量空间 $V$ 不能被它的有限个真子空间覆盖. 也就是不存在真子空间 $V_1,\ldots,V_m$ 使得 $V\subseteq V_1\cup V_2\cup\cdots\cup V_m$.

证明: 设 $\varepsilon_1,\ldots,\varepsilon_n$ 是 $V$ 的一组基, 考察

$$\alpha_i=\varepsilon_1+i\varepsilon_2+\cdots+i^{n-1}\varepsilon_n,$$

由 Vandermonde 行列式的知识知道无穷向量序列 $\alpha_1,\alpha_2,\cdots,\alpha_n\cdots$ 中任何 $n$ 个都构成 $V$ 的一组基. 而有限多个真子空间只能包含其中有限多个向量, 从而有无穷多个 $\alpha_i$ 不能被覆盖.

这个定理用到的地方很少, 不过每次应用都可以带来惊喜. 以下是两个例子.

## 09 年南开考研高代最后一题

➤ **例题**(2009南开大学). 设 $F$ 是一个数域, $T$ 是 $\mathrm{Mat}_n(F)$ 上的线性变换满足对任何 $A,B\in\mathrm{Mat}_n(F)$, $T(AB)=T(A)T(B)$ 和 $T(AB)=T(B)T(A)$ 至少有一个成立. 求证: 要么对所有的 $A,B\in\mathrm{Mat}_n(F)$ 都有 $T(AB)=T(A)T(B)$ 成立, 要么对所有的 $A,B\in\mathrm{Mat}_n(F)$ 都有 $T(AB)=T(B)T(A)$ 成立.

证明: 对于一个固定的矩阵 $A$, 如果对任何矩阵 $B$ 都有 $T(AB)=T(A)T(B)$ 成立, 就称 $A$ 是正手型的. 相应地, 如果对任何矩阵 $B$ 都有 $T(AB)=T(B)T(A)$ 成立, 就称 $A$ 是反手型的. 证明分两步, 首先证明 $\mathrm{Mat}_n(F)$ 中的方阵必然是正手型或者反手型之一, 其次证明正手型和反手型不能同时出现, 就完成了证明.

首先对任一矩阵 $A$, 设 $U_A$ 是所有满足 $T(AB)=T(A)T(B)$ 的方阵 $B$ 构成的集合, $W_A$ 是所有满足 $T(AC)=T(C)T(A)$ 的方阵 $C$ 构成的集合, 不难验证 $U_A,W_A$ 都是 $\mathrm{Mat}_n(F)$ 的子空间. 而且根据已知, $\mathrm{Mat}_n(F)=U_A\cup W_A$. 从而 $U_A,W_A$ 不可能都是 $\mathrm{Mat}_n(F)$ 的真子空间, 所以必有 $U_A=\mathrm{Mat}_n(F)$ 或者 $W_A=\mathrm{Mat}_n(F)$ 之一成立. 从而 $A$ 不是正手型就是反手型.

其次, 设 $U_1$ 是全体正手型的 $A$ 构成的集合, $U_2$ 是全体反手型的 $A$ 构成的集合, 显然 $U_1,U_2$ 也都是子空间. 注意由刚刚证得的结论, $\mathrm{Mat}_n(F)=U_1\cup U_2$, 从而 $U_1,U_2$ 不能都是真子空间, 所以要么 $U_1=\mathrm{Mat}_n(F)$, 要么 $U_2=\mathrm{Mat}_n(F)$. 如果前者成立（都是正手型的）, 那么对任何 $A,B$ 都有

$$T(AB)=T(A)T(B).\quad \text{（因为 } A \text{ 是正手型的）}$$

同理如果后者成立（都是反手型的）, 则

$$T(AB) = T(B)T(A). \quad （因为\ A\ 是反手型的）$$

这样就得到了证明.

这个证明很巧吧? 我猜南开的老师要的就是这个证明.

### 极小零化多项式=极小多项式的向量

**定理5.2.** 设 $A$ 是 $n$ 维向量空间 $V$ 上的线性变换, 则存在 $V$ 中的向量 $v$, 其极小零化多项式恰好等于 $A$ 的极小多项式 $m(x)$.

证明: 熟知 $V$ 中任一元素 $v$ 的极小零化多项式 $m_v(x)$ 整除 $A$ 的极小多项式 $m(x)$. 但是 $m(x)$ 只有有限多个因子, 所以当 $v$ 跑遍 $V$ 时, $m_v(x)$ 只有有限多个不同的可能 $m_1(x), m_2(x), \ldots, m_k(x)$. 考虑

$$V_i = \left\{ v \in V \,\middle|\, m_i(A)v = 0 \right\},$$

那么 $V$ 属于 $V_1, \ldots, V_k$ 的并, 从而 $V_1, \ldots, V_k$ 不能都是真子空间, 所以必有某个 $V_i$ 使得 $V = V_i$, 那么就有 $m(x) = m_i(x)$, 得证.

### 特征 = 极小

一个变换的特征多项式和极小多项式相等是一个很重要的情形, 值得大书特书一番. 下面这个问题在论坛上经常被问起:

**定理5.3.** $A$ 的特征多项式 $f(x)$ 与极小多项式 $m(x)$ 相等的充要条件是存在 $v \in V$ 使得 $v, Av, \ldots, A^{n-1}v$ 构成 $V$ 的一组基.

必要性: 由已知, $m(x)$ 的次数是 $n$, 根据定理 5.2, 有一个向量 $v$ 的极小零化多项式 $m_v(x)$ 等于 $m(x)$, 从而 $v, Av, \ldots, A^{n-1}v$ 必然线性无关. 否则的话设

$$a_0 v + a_1 Av + \cdots + a_{n-1}A^{n-1}v = 0,$$

这就说明 $v$ 的阶 $m_v(x)$ 的次数小于等于 $n-1$, 矛盾!

充分性: 显然 $m_v(x)$ 是 $n$ 次的, 而 $m_v$ 又是极小多项式 $m(x)$ 的因子, 所以 $m(x)$ 也是 $n$ 次的, 从而 $m(x)$ 与 $f(x)$ 相等.

**定理5.4.** 矩阵 $A$ 的特征多项式 $f(x)$ 与极小多项式 $m(x)$ 相等的充要条件是 $A$ 的任何特征值对应的特征子空间都是一维的.

其实这个定理很直观, $A$ 的特征多项式与极小多项式相等等价于 $A$ 的 Jordan 标准型里面任何特征值对应的 Jordan 块都只有一块, 也就等价于任何特征值空间都是一维的.

6

# 把交换矩阵表示为多项式

何时任何与矩阵 $A$ 交换的矩阵总能表示为 $A$ 的多项式? 下面将从变换和矩阵两个角度入手, 对这个问题进行一下深入的探讨, 力求把想法写清楚.

## 变换的角度

如果任何与 $A$ 交换的变换 $B$ 总能表示为 $A$ 的多项式, 那么 $A$ 应该有怎样的性质? 首先由于 $A$ 的不变子空间也是 $f(A)$ 的不变子空间, 所以如果 $V$ 有 $A-$ 不变子空间的直和分解

$$V = V_1 \bigoplus V_2 \bigoplus \cdots \bigoplus V_s$$

的话, 那么它必然也是 $V$ 的 $B-$ 不变子空间直和分解. 换句话说, 任何与 $A$ 交换的变换 $B$ 不能把 $V_i$ 中的向量映到别的 $V_j$ 中去. 从而我们得到第一个有用的信息: 设

$$V = V_1 \bigoplus V_2 \bigoplus \cdots \bigoplus V_s$$

是 $V$ 的 $A-$ 不变子空间直和分解, 那么当 $i \neq j$ 时不存在非零变换 $C : V_i \to V_j$ 使得 $AC = CA$.

若不然, 有这样的 $C$ 存在, 那么当 $k$ 不等于 $i$ 时定义 $B$ 在所有的 $V_k$ 上都是0, 在 $V_i$ 上就定义为 $C$, 则不难验证 $AB = BA$, 但是 $V_i$ 不是 $B$ 的不变子空间, $B$ 是不能表示为 $A$ 的多项式的. 这是一个重要的发现, 这让我们想起 **??** 讨论过的:

设 $A, B$ 分别是复数域上有限维向量空间 $V$ 和 $W$ 上的线性变换, 那么存在非零的线性映射$C : V \to W$ 使得 $CA = BC$ 的充要条件是 $A$ 和 $B$ 有共同的特征值.

所以我们得到如下重要结论: 当 $A$ 有不变子空间直和分解时, $A$ 在任何两个不变子空间上的限制不能有共同的特征值. 进一步, 由于 $A$ 总是可以分解为循环不变子空间的直和, 每个子空间下呈 Jordan 形, 我们立刻得到 $A$ 的 Jordan 标准形中任何特征值 $\lambda$ 对应的 Jordan 块只能有一块, 或者说 $A$ 的极小多项式与特征多项式相同, 即存在 $V$ 中的向量 $v$ 上使得 $v, Av, \cdots, A^{n-1}v$ 构成 $V$ 的一组基.

现在我们已经找到了 $A$ 应该满足的必要条件: $A$ 有一个循环向量. 那么这个条件是不是充分的呢? 答案是肯定的:

证明: 设 $V$ 在 $A$ 的作用下成为一个循环空间, $v$ 是一个生成元, 我们断言循环空间上一个与 $A$ 交换的变换 $B$ 由它在生成元 $v$ 处的值 $Bv$ 完全决定: 由于 $V$ 中的任何向量 $m$ 可以表示为 $v, Av, \cdots, A^{n-1}v$ 的线性组合, 也就是存在多项式 $f(x)$ 使得 $m = f(A)v$. 特别有 $Bv = f(A)v$. 不难验证就有 $B = f(A)$ 成立. 从而我们已经从变换的角度完全解决了这一问题.

**定理6.1.** 任何与 $A$ 交换的矩阵都能写成 $A$ 的多项式的充要条件是 $V$ 在 $A$ 的作用下有一个循环向量.

### 矩阵的角度

在定理 4.4 中令 $A=B$, 那么 $i(A,A)$ 就是与 $A$ 交换的矩阵构成的空间的维数.

$$i(A,A)=\sum_{i=1}^{p}\sum_{j=1}^{p}\deg\Big(f_i(x),f_j(x)\Big)\geq\sum_{i=1}^{p}\deg\Big(f_i(x),f_i(x)\Big)=\sum_{i=1}^{p}\deg f_i(x)=n.$$

等号当且仅当 $p=1$ 时成立. 这个时候 $A$ 只有一个不变因子, 即 $A$ 的特征多项式等于其极小多项式, 从而 $I,A,A^2,\dots,A^{n-1}$ 线性无关, $A$ 的多项式构成一个 $n$ 维的空间, 恰好等于与 $A$ 交换的矩阵构成的空间的维数, 这就从矩阵的角度证明了定理 6.1.

**例题.** 如果矩阵 $A$ 有 $n$ 个互不相同的特征值, 那么与 $A$ 交换的矩阵一定可以表示为 $A$ 的多项式.

证明: 设 $Av_i=\lambda_i v_i$, 那么 $v_1+v_2+\cdots+v_n$ 就是一个循环向量.

也可以这样做, $A$ 必然是可以对角化的, 那么与 $A$ 交换的矩阵必然也是对角矩阵（按特征值分块的小技巧）, 所以问题就是一个在 $n$ 个点处找插值多项式的问题, 剩下的就不写了.

## 双中心化子定理

**定理6.2 (双中心化子).** 设 $A$ 是任一给定的方阵, 如果方阵 $C$ 和每一个与 $A$ 可交换的方阵都可交换, 则 $C$ 可以表示为 $A$ 的多项式.

证明: 首先把 $A,C$ 都看成某个向量空间上的线性变换. 如果结论成立的话, 显然下面的条件是必须满足的:

**引理6.3.** $A$ 的不变子空间直和分解必定也是 $C$ 的不变子空间直和分解.

引理的证明: 设 $V$ 在 $A$ 的作用下分解为不变子空间的直和

$$V=V_1\oplus V_2\oplus\cdots\oplus V_r,$$

$P$ 是从 $V$ 到 $V_1$ 上的投影, 则有 $AP=PA$, 那么根据已知有 $CP=PC$, 所以 $V_1$ 也是 $C$ 的不变子空间, 类似地所有 $V_i$ 都是 $C$ 的不变子空间, 引理得证.

现在立刻想到用上定理 6.1 的结论, 根据 Frobenius 有理标准形的理论, $V$ 可以分解为循环不变子空间的直和

$$V=\{v_1\}\oplus\{v_2\}\oplus\cdots\oplus\{v_r\},$$

而 $C$ 在每个 $\{v_i\}$ 上的限制与 $A$ 在 $\{v_i\}$ 上的限制交换, 所以可以表示为 $A$ 的多项式 $C = g_i(A)$. 问题是这些 $g_i(x)$ 相等吗? 如果是同一个多项式 $g(x)$ 那问题就OK了, 在每一个不变子空间上都有 $C = g(A)$, 当然全空间也是. 但是这个不那么显然, 所以得分析的再细一点. 约定每个 $\{v_i\}$ 对应的极小多项式 $f_i(x)$ 是 $A$ 的第 $i$ 个不变因子, 次数从大到小排列. 我们有

**引理6.4.** 存在与 $A$ 交换的变换 $B$ 使得 $Bv_1 = v_2$.

我不打算写出详细的证明过程, 而是把想法写清楚. 每个 $\{v_i\}$ 都是一个"循环群", 生成元是 $v_i$, 阶是 $f_i(x)$, 特别 $v_2$ 的阶是 $v_1$ 的阶的因子. 请仿照下面这个问题进行类比来完成引理的证明:

**引理6.5.** 设 $G = <g>$ 和 $H = <h>$ 是两个循环群, 且 $|H|$ 是 $|G|$ 的因子, 则存在从 $G$ 到 $H$ 的群同态 $\varphi$ 满足 $\varphi(g) = h$.

最后的证明: 我们有

$$v_1 \xrightarrow{B} v_2 \xrightarrow{C} g_2(A)v_2,$$
$$v_1 \xrightarrow{C} g_1(A)v_1 \xrightarrow{B} g_1(A)v_2.$$

由于 $CB = BC$ 所以两条路径的结果应该是一样的. 也就是 $g_1(A)v_2 = g_2(A)v_2$. 再引用一遍那句话: 循环空间上一个与 $A$ 交换的变换由它在生成元 $v$ 处的值完全决定. 所以 $g_1(A)$ 和 $g_2(A)$ 相等（在 $\{v_2\}$ 上）, 那么 $g_1(x)$ 和 $g_2(x)$ 可以取成相同的多项式. 类似地, 所有的 $g_i(x)$ 全部可以取成同样的多项式.

这个定理的意义就是在环 $\mathrm{End}_k(V)$ 中, 对变换 $A$ 生成的子环 $R$ 求两次中心化子以后又得到了 $R$, 所以叫做双中心化子定理. 在半单代数结构定理中还会见到双中心化子定理的身影.

# 7 分解与提升

空间的分解与提升是高等代数中最重要的思想, 处理与线性变换有关的问题的时候, 基本的想法就是先找合适的子空间实现降阶, 然后转移到商空间（使用归纳假设）, 最后拉回到全空间得出结论. 其实教材上的很多定理已经体现了这一点.

### Hamilton-Cayley 定理

**定理7.1.** 设 $A$ 是数域 $F$ 上的 $n$ 阶方阵, $f(x)$ 是 $A$ 的特征多项式, 则 $f(A)=0$.

这个降阶并不难, 把 $A$ 看成 $\mathbb{C}$ 上的矩阵, 取特征向量 $A\alpha=\lambda\alpha$ 并扩充为一组基, 则在这组基下 $A$ 的矩阵形如

$$\begin{pmatrix} \lambda & C \\ 0 & B \end{pmatrix}.$$

其中 $B$ 是 $A$ 在商空间 $\bar{V}=V/\{\alpha\}$ 上的诱导变换对应的矩阵. 设 $B$ 的特征多项式为 $g(x)$, 那么 $f(x)=(x-\lambda)g(x)$. 对降阶后的矩阵 $B$ 使用归纳假设: $g(B)=0$, 那么

$$f(A)=(A-\lambda)g(A)=\begin{pmatrix} 0 & C \\ 0 & B-\lambda \end{pmatrix}\begin{pmatrix} g(\lambda) & * \\ 0 & 0 \end{pmatrix}=0.$$

所以 Hamilton-Cayley 定理是最简单的降阶.

### 同时上三角化

**定理7.2.** 如果 $\mathbb{C}$ 上的方阵 $A,B$ 满足 $AB=BA$, 则 $A,B$ 可以同时相似于上三角形. 即存在可逆矩阵 $T$ 使得 $T^{-1}AT, T^{-1}BT$ 都是上三角矩阵.

这个降阶和上面完全一样, 就是找共同的特征向量, 开拓为一组基以后对商空间上的诱导变换使用归纳假设得出结论.

类似的还有

**定理7.3.** 如果 $A,B$ 都是可对角化的矩阵, 而且 $AB=BA$, 则 $A,B$ 可以同时对角化. 即存在可逆矩阵 $T$ 使得 $T^{-1}AT, T^{-1}BT$ 都是对角矩阵.

道理完全一样, 就是把空间分解为 $A$ 的特征子空间的直和, 在每一个子空间上使用归纳假设即可. 不过这里有一个小小的陷阱: 必须论证一个可对角化的变换在其任一不变子空间上的限制仍然是一个可对角化的变换（虽然简单但并不显然）.

➤ **定理7.4.** 当 $\mathrm{r}(AB-BA)\le 1$时 $A,B$ 可以同时上三角化.

首先我们可以假设 $A$ 不是可逆的, 否则的话就以 $A-\lambda I$ 代替 $A$. 这个假定的目的就是保证 $\mathrm{Ker}A$ 和 $\mathrm{Im}A$ 都是非平凡的 $A-$ 不变子空间. 我们要论证 $\mathrm{Ker}A$ 和 $\mathrm{Im}A$ 中至少有一个是 $B$ 的不变子空间, 这样就可以降低空间的维数, 便于应用归纳法. 记齐次线性方程组 $(AB-BA)X=0$ 的解空间为 $X$, 分两种情况讨论:

(1) $\mathrm{Ker}A\subset X$. 设$x\in\mathrm{Ker}A$, 立刻有 $ABx=BAx=0$. 从而 $\mathrm{Ker}A$ 也是 $B$ 的不变子空间.

(2) $\mathrm{Ker}A\not\subseteq X$. 从而存在向量 $x\in\mathrm{Ker}A$ 但是 $y=(AB-BA)x\ne 0$. 注意到 $\mathrm{r}(AB-BA)=1$, 所以 $\mathrm{Im}(AB-BA)=\{y\}$. 但是

$$y=(AB-BA)x=ABx-BAx\in\mathrm{Im}A$$

这就说明 $\mathrm{Im}(AB-BA)\subset\mathrm{Im}A$. 那么对任何 $u=Av\in\mathrm{Im}A$, 我们有

$$BAv=ABv-(AB-BA)v\in\mathrm{Im}A.$$

从而 $\mathrm{Im}A$ 确实是 $B$ 的不变子空间.

剩下的就可以对空间的维数进行归纳来证明结论了, 不妨设 $\mathrm{Im}A$ 是 $A,B$ 共同的不变子空间, 那么取 $\mathrm{Im}A$ 的一组基然后开拓, $A$ 和 $B$ 的矩阵就形如

$$\begin{pmatrix}A_1 & A_3\\ 0 & A_2\end{pmatrix}\quad 和\quad \begin{pmatrix}B_1 & B_3\\ 0 & B_2\end{pmatrix}.$$

显然

$$\mathrm{r}(AB-BA)\ge \mathrm{r}(A_1B_1-B_1A_1)+\mathrm{r}(A_2B_2-B_2A_2),$$

所以由归纳假设 $A_1$ 和 $B_1$ 可以同时上三角化, $A_2$ 和 $B_2$ 可以同时上三角化, 这就完成了证明.

### 正规矩阵的结构

正规矩阵有很好的标准形的关键就在于当 $M$ 是它的不变子空间的时候, $M^{\perp}$ 也是其不变子空间. 这就可以把 $V$ 分解为一些“不可约”的不变子空间的直和, 即每一个直和项不再含有更小的非零不变子空间（否则这个直和项就可以再进行直和分解）. 所以对于酉空间上的正规矩阵 $A$, 由于“不可约”的不变子空间必须是一维的特征向量, 所以 $A$ 可以对角化. 欧式空间上的正规矩阵 $A$ 的“不可约”不变子空间是一维或者二维的, 一维的是数乘, 二维的是一个带伸缩因子的旋转（想象平面上复数乘法的几何意义）, 所以 $A$ 的标准形是一些

$$\lambda\quad 和\quad \begin{pmatrix}a & b\\ -b & a\end{pmatrix}\qquad（伸缩因子\ \sqrt{a^2+b^2},\ 转角\ \cos\theta=\frac{a}{\sqrt{a^2+b^2}}\ ）$$

的直和.

## Frobenius 标准形

**定理7.5.** 设 $A$ 是 $F$ 上 $n$ 维向量空间 $V$ 上的线性变换, 则存在首 1 正次数的多项式 $p_1,p_2,\cdots,p_m\in F[x]$ 以及 $A-$ 循环不变子空间 $V_1,V_2,\cdots,V_m$ 使得

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

并且 $p_i$ 恰好是 $A$ 在 $V_i$ 上的限制 $A|V_i$ 的极小多项式（当然也是特征多项式）. 此外还有
(1) $p_1(x)p_2(x)\cdots p_m(x)=f(x)$, 这里 $f(x)$ 是 $A$ 的特征多项式.
(2) $p_1(x)=m(x)$ 是 $A$ 的极小多项式, $p_m|p_{m-1}|\cdots|p_2|p_1$.

Frobenius 标准形是最能体现分解与提升思想的定理, 它难就难在最后一步的"拉回"上. 教材上的证明确实不容易读懂（比如科大教材上的叙述）. 这有两方面的原因, 一个是证明不用模的语言来表述的话确实是麻烦, 另一个就是作者没有把思路讲清楚. 下面主要讲思路, 细节大家自己去看.

还是对空间的维数归纳, 首先根据定理 5.2, 有一个向量 $v_1$ 的极小化零多项式恰好等于 $A$ 的极小多项式 $m(x)$, 设 $v_1$ 生成的循环子空间为 $V_1$, 我们的思路就是在商空间 $V/V_1$ 中使用归纳假设, 再拉回到 $V$ 中来得出结论. 具体讲, 就是根据归纳假设, $V/V_1$ 可以分解为循环不变子空间的直和

$$V/V_1=\bar{V_2}\oplus\bar{V_3}\oplus\cdots\oplus\bar{V_m}.$$

这里 $\bar{V_i}$ 由 $\bar{v_i}$ 生成, 每个 $\bar{v_i}$ 在 $V/V_1$ 中的极小化零多项式是 $p_i(x)\,(i\geq 2)$, 且有 $p_m|p_{m-1}|\cdots|p_2$. 当然 $p_i|m(x)=p_1(x)$ 也是显然的. 现在的问题是怎么拉回全空间 $V$ 呢? 显然应该取 $\bar{v_i}$ 的代表元 $v_i$, 设 $v_i$ 生成的 $A-$ 不变子空间为 $V_i$, 去证明这样的 $V_i$ 符合要求. 但是这个代表元不能随便取, 我们有下面的引理:

**引理7.6.** 对于 $i\geq 2$, 存在 $\bar{v_i}$ 在 $V$ 中的代表元 $v_i$ 使得 $v_i$ 在 $V$ 中的极小化零多项式等于 $\bar{v_i}$ 在 $V/V_1$ 中的极小化零多项式.

让我们先承认引理 7.6 的正确性, 把证明继续下去:

设 $v_i$ 生成的 $A-$ 不变子空间为 $V_i$, 下面只需要证明

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

首先得证明

$$V=V_1+V_2+\cdots+V_m.$$

这个简单, 关键是证明

$$V_1+V_2+\cdots+V_m$$

是直和, 即零向量的表示法唯一. 由于 $V_i$ 中的向量都可以表示为 $v_i$ 的"多项式"加上一个 $v_1$ 的"多项式", 所以可以假设存在 $r_1(x),r_2(x),\ldots,r_m(x)\in F[x]$ 使得

$$r_1v_1+r_2v_2+\cdots+r_mv_m=0,$$

转移到商空间 $V/V_1$ 中去:

$$r_2\bar{v}_2+\cdots+r_m\bar{v_m}=\bar{0}.$$

然而由归纳假设,

$$\bar{V_2}\bigoplus\cdots\bigoplus\bar{V_m}$$

是直和, 所以必须

$$r_i\bar{v}_i=0,\quad i\geq 2.$$

又由于 $p_i$ 是 $\bar{v}_i$ 的极小零化多项式, 所以

$$p_i(x)\,|\,r_i(x)$$

然而 $v_i$ 的极小化零多项式也是 $p_i$, 从而

$$r_iv_i=0,$$

从而每一个和项都是零向量, 这就证明了直和.

引理 7.6 的证明: $p_i\bar{v}_i=0$ 说明 $p_iv_i\in V_1$, 即存在多项式 $f_i(x)$ 使得

$$p_iv_i=f_iv_1.$$

由于 $p_i\,|\,m(x)$, 所以可以设 $m(x)=p_i(x)q_i(x)$, 在两边同时乘以 $q_i(x)$:

$$0=f_iq_iv_1.$$

由于 $m(x)$ 是 $v_1$ 的极小化零多项式, 所以

$$m(x)\,|\,f_i(x)q_i(x).$$

即

$$p_i(x)\,|\,f_i(x).$$

令 $f_i(x)=p_i(x)l_i(x)$, 那么

$$p_i(x)(v_i-l_iv_1)=0,$$

容易验证 $v_i-l_iv_1$ 的阶就是 $p_i(x)$, 这就完成了引理 7.6 的证明.

8

# 正规矩阵

正规矩阵的谱完全决定了它的结构,这是正规矩阵有很好的性质的根本原因.这一章收录几个和正规矩阵有关的问题.

➤ **定理8.1.** 证明 $A$ 是正规变换当且仅当 $A^*$ 可以表示为 $A$ 的多项式.

证明: 设 $\lambda_1,\dots,\lambda_s$ 是 $A$ 的全部不同的特征值, $V_1,\dots,V_s$ 是对应的特征子空间, $P_1,\dots,P_s$ 是 $V$ 到 $V_1,\dots,V_s$ 的射影, 则 $P_i^2=P_i, P_iP_j=0\,(i\neq j)$.
我们有

$$A=\lambda_1P_1+\lambda_2P_2+\cdots+\lambda_sP_s,$$

那么

$$A^2=\lambda_1^2P_1+\cdots+\lambda_s^2P_s.$$

可见对任何正整数 $k$,

$$A^k=\lambda_1^kP_1+\cdots+\lambda_s^kP_s,$$

从而对任何多项式 $f(x)$,

$$f(A)=f(\lambda_1)P_1+\cdots+f(\lambda_s)P_s.$$

设 $p(x)$ 是 Lagrange 插值多项式

$$p(\lambda)=\prod_{i\neq 1}\frac{\lambda-\lambda_i}{\lambda_1-\lambda_i},\quad p(\lambda_1)=1,\quad p(\lambda_i=0),\quad i=2,\cdots,s.$$

则 $p(A)=\sum\limits_{i=1}^{s}p(\lambda_i)P_i=P_1$, 即 $P_1$ 可以表示为 $A$ 的多项式, 同理每个 $P_i$ 都可以表示为 $A$ 的多项式, 从而

$$A^*=\bar\lambda_1P_1+\cdots+\bar\lambda_sP_s$$

可以表示为 $A$ 的多项式.

**引理8.2.** 设 $A$ 是正规矩阵, $B$ 是方阵. 如果 $AB=BA$, 那么有 $A^*B=BA^*$.

证明: 把 $A,B$ 都看成酉空间 $V$ 上的线性变换, 由于 $V$ 可以分解为 $A$ 的特征子空间的直和, 所以 $A,B$ 交换等价于 $A$ 的特征子空间都是 $B$ 的不变子空间. 但是 $A$ 的特征子空间也恰好是 $A^*$ 的特征子空间, 所以 $A^*$ 的特征子空间也都是 $B$ 的不变子空间, 所以 $A^*$ 和 $B$ 交换.

其实这个引理也可以由前面的定理得到:

$$AB=BA\Rightarrow f(A)B=Bf(A)\Rightarrow A^*B=BA^*.$$

**引理8.3.** 设 $A,B$ 是正规矩阵而且 $AC=CB$, 那么 $A^*C=CB^*$.

证明: 考察

$$X=\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix},\quad S=\begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix}.$$

那么 $X$ 仍是正规矩阵而且 $XS=SX$. 所以由引理 8.2, $X^*S=SX^*$, 也就是 $A^*C=CB^*$.

➤ **定理8.4 (正规矩阵复相似等价于酉相似).** 设 $A,B$ 是正规矩阵且存在可逆矩阵 $T$ 使得 $T^{-1}AT=B$, 则存在酉矩阵 $U$ 使得 $U^{-1}AU=B$.

思路: 设可逆矩阵 $T$ 使得 $T^{-1}AT=B$, 要找一个酉矩阵 $U$ 使得 $U^{-1}AU=B$. 怎么从相似过渡到酉相似呢? 用极分解, 取正定 Hermite 矩阵 $S$ 和酉矩阵 $U$ 使得 $T=SU$, 那么$U^{-1}(S^{-1}AS)U=B$. 如果能证明 $S$ 与 $A$ 交换, 那结论就成立了. 而 $S$ 又是 $TT^*$ 的平方根, 所以只要证明 $A$ 与 $TT^*$ 交换. (这里你需要知道, 一个矩阵 $A$ 和一个正定矩阵 $P$ 交换当且仅当 $A$ 和 $P$ 的平方根交换.)

证明: 由已知, $AT=TB$, 那么根据引理 8.3 就有

$$A^*T=TB^*.$$

两边求共轭矩阵:

$$T^*A=BT^*.$$

那么

$$T(T^*A)=T(BT^*)=(TB)T^*=ATT^*.$$

即 $A$ 与 $TT^*$ 交换, 这就完成了证明.

➤ **定理8.5 (实矩阵酉相似等价于正交相似).** 设 $A,B$ 是实矩阵, $U$ 是酉矩阵, $U^{-1}AU=B$, 则存在正交矩阵 $O$ 使得 $O^{-1}AO=B$.

思路: 设 $U=P+iQ$, $P,Q$ 是实矩阵, 那么由 $AU=UB$ 可以推出 $AP=PB, AQ=QB$. 从而对任何实数 $\lambda$ 都有 $A(P+\lambda Q)=(P+\lambda Q)B$, 特别取一个实数 $\lambda$ 使得 $T=P+\lambda Q$ 可逆, 那么就得到了一个可逆的实矩阵 $T$ 使得 $T^{-1}AT=B$. 所以只要证明 $A$ 与 $TT^*$ 交换, 就可以复制上面的极分解的手段了.

证明: 首先由 $U^{-1}AU=B$ 可得 $U^*A=BU^*$, 从而

$$(P^*+\lambda Q^*)A=B(P^*+\lambda Q^*),$$

也就是$T^*A=BT^*$. 那么 $ATT^*=TBT^*=TT^*A$, 得证.

9

# 几种特殊的矩阵

## 幂零矩阵

矩阵幂零等价于所有的特征根都是0, 要验证这一点, 只要证明对所有的 $1\leqslant k\leqslant n$ 都有 $\mathrm{tr}A^k=0$. 为什么呢? 原因是迹函数 $\mathrm{tr}A,\mathrm{tr}A^2,\cdots,\mathrm{tr}A^n$ 完全决定了 $A$ 的特征多项式. 设 $\sigma_k$ 是 $\lambda_1,\ldots,\lambda_n$ 的 $k$ 次初等对称多项式, 那么

$$\mathrm{tr}A^k=\lambda_1^k+\lambda_2^k+\cdots+\lambda_n^k=s_k$$

根据 Newton 恒等式有

$$\begin{aligned}\sigma_1&=s_1\\ s_1\sigma_1-2\sigma_2&=s_2\\ s_2\sigma_1-s_1\sigma_2+3\sigma_3&=s_3\\ &\cdots\cdots\\ s_k\sigma_1-s_{k-1}\sigma_2+\cdots+(-1)^{k+1}k\sigma_k&=s_k\end{aligned}$$

用 Cramer 法则依次解出 $\sigma_1,\ldots,\sigma_n$ 来:

$$\sigma_k=\frac{1}{k!}\begin{vmatrix}s_1&1&0&0&\cdots&0\\ s_2&s_1&1&0&\cdots&0\\ s_3&s_2&s_1&1&\cdots&0\\ \vdots&\vdots&\vdots&\ddots&\ddots&\vdots\\ s_{k-1}&s_{k-2}&\cdots&\cdots&\ddots&1\\ s_k&s_{k-1}&\cdots&\cdots&\cdots&s_1\end{vmatrix}$$

从而 $s_1,\ldots,s_n$ 完全决定了 $\sigma_1,\ldots,\sigma_n$, 也就完全决定了 $A$ 的特征多项式.

## Gram 矩阵

设 $\alpha_1,\ldots,\alpha_s$ 是 $n$ 维欧式空间中的 $s$ 个线性无关的向量, 称 $s\times s$ 矩阵

$$G_s=\begin{pmatrix}(\alpha_1,\alpha_1)&(\alpha_1,\alpha_2)&\cdots&(\alpha_1,\alpha_s)\\ (\alpha_2,\alpha_1)&(\alpha_2,\alpha_2)&\cdots&(\alpha_2,\alpha_s)\\ \vdots&\vdots&\ddots&\vdots\\ (\alpha_s,\alpha_1)&(\alpha_s,\alpha_2)&\cdots&(\alpha_s,\alpha_s)\end{pmatrix}=\begin{pmatrix}G_{s-1}&x\\ x'&(\alpha_s,\alpha_s)\end{pmatrix}$$

为向量组 $\alpha_1,\ldots,\alpha_s$ 的 Gram 矩阵.

下面来算 $G_s$ 的行列式. 以 $\alpha_1,\ldots,\alpha_s$ 为列向量排成 $n\times s$ 矩阵 $A$, 那么 $G_s = A'A$. 设 $\alpha_1,\ldots,\alpha_s$ 施密特正交化后得到的向量为 $\beta_1,\ldots,\beta_s$. 以 $\beta_1,\ldots,\beta_s$ 为列向量排成 $n\times s$ 矩阵 $B$, 那么 $B=AQ$, 这里 $Q$ 是一个对角线上都是1的上三角矩阵. 从而 $B'B = Q'G_sQ$. 注意这里

$$B'B = \begin{pmatrix} (\beta_1,\beta_1) & (\beta_1,\beta_2) & \cdots & (\beta_1,\beta_s) \\ (\beta_2,\beta_1) & (\beta_2,\beta_2) & \cdots & (\beta_2,\beta_s) \\ \vdots & \vdots & \ddots & \vdots \\ (\beta_s,\beta_1) & (\beta_s,\beta_2) & \cdots & (\beta_s,\beta_s) \end{pmatrix} = \begin{pmatrix} (\beta_1,\beta_1) & & & \\ & (\beta_2,\beta_2) & & \\ & & \ddots & \\ & & & (\beta_s,\beta_s) \end{pmatrix}$$

而且 $|Q'| = |Q| = 1$. 所,以在 $B'B = Q'G_sQ$ 两边求行列式即得 $|G_s| = \|\beta_1\|^2\|\beta_2\|^2\cdots\|\beta_s\|^2$.

从上面的式子可以看出来 $G_s$ 的行列式就是 $\alpha_1,\ldots,\alpha_s$ 张成的平行多面体体积的平方.

### 循环矩阵

对于给定的 $\mathbb{C}^n$ 中的向量 $\alpha = (a_0,a_1,\cdots,a_{n-1})$, 定义由 $\alpha$ 生成的循环矩阵

$$circ[\alpha] = circ[a_0,a_1,\cdots,a_{n-1}] = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix}.$$

设所有循环矩阵组成的集合为 $S$.

➤ **定理9.1.** 所有循环矩阵可以同时对角化.

证明: 设 $\pi = circ[0,1,\cdots,0]$, 换言之 $\pi$ 就是轮换 $(123\cdots n)$ 对应的置换矩阵, 那么 $\pi^n = I$, 不难验证 $circ[\alpha] = a_0I + a_1\pi + a_2\pi^2 + \cdots + a_{n-1}\pi^{n-1}$, 从而所有循环矩阵都可以表示为 $\pi$ 的多项式, 所以只要把 $\pi$ 对角化, 所有的循环矩阵就同时被对角化了. 显然 $\pi$ 是可以被对角化的, 因为 $\pi$ 的极小多项式没有重根, 这就证明了结论.

特别注意 $S$ 作为一个代数仅由一个元素 $\pi$ 生成, 但是看作一个向量空间却是 $n$ 维的. 这是因为 $\pi$ 的极小多项式为 $x^n-1$, 所以 $I,\pi,\cdots,\pi^{n-1}$ 线性无关.

### Hilbert 矩阵

考虑 $\mathbb{R}[x]$ 的由$1,x,\cdots,x^{n-1}$ 生成的 $n$ 维子空间 $V$. 在内积

$$<f,g>=\int_0^1 f(x)g(x)\,\mathrm{d}x$$

下 $V$ 成为一个欧式空间. 这个内积在基$1,x,\cdots,x^{n-1}$ 下的度量矩阵为

$$G = (\frac{1}{i+j+1}),\quad 0\le i,j\le n-1.$$

$G$ 又叫做 Hilbert 矩阵, 它是正定矩阵, 因为内积的度量矩阵总是正定的.
下面来算 $G$ 的行列式: 更一般地, 考虑 $n$ 阶矩阵

$$G=(\frac{1}{\alpha_i+\beta_j}),\qquad 1\le i,j\le n$$

下面证明

$$\det(G)=\frac{\prod\limits_{1\le i<j\le n}(\alpha_i-\alpha_j)(\beta_i-\beta_j)}{\prod\limits_{1\le i,j\le n}(\alpha_i+\beta_j)}$$

这是因为设

$$\det(G)\prod_{1\le i,j\le n}(\alpha_i+\beta_j)=F(\alpha_1,\cdots,\alpha_n,\beta_1,\cdots,\beta_n)$$

那么 $F$ 是一个 $n^2-n$ 次的多项式. 如果某个 $\alpha_i=\alpha_j$, 那么 $G$ 的第 $i,j$ 两行相同, $\det(G)=0$, 所以 $F$ 有因子

$$\prod_{1\le i<j\le n}(\alpha_i-\alpha_j)$$

类似地, $F$ 还有因子

$$\prod_{1\le i<j\le n}(\beta_i-\beta_j)$$

而

$$\prod_{1\le i<j\le n}(\alpha_i-\alpha_j)(\beta_i-\beta_j)$$

已经是一个 $n^2-n$ 次的多项式, 再比较一下首项系数就可得证.

特别令 $\alpha_i=i+\frac{1}{2},\beta_j=j+\frac{1}{2}$, 就可得到 $G$ 的行列式为

$$\det(G)=\frac{\prod\limits_{0\le i<j\le n-1}(i-j)^2}{\prod\limits_{0\le i,j\le n-1}(i+j+1)}=\frac{[1!2!\cdots(n-1)!]^3}{n!(n+1)!\cdots(2n-1)!}.$$

10

# 解题的艺术

按特征值分块的技巧在这个文档里已经出现好几次了, 下面继续介绍两个例子.

➤ **例题.** 伴随矩阵 $A^*$ 总可以表示为 $A$ 的多项式.

这个问题在论坛上出现过多次了, 可以参看周不通老师的帖子

http://old.math.org.cn/forums/index.php?showtopic=40761

这里给出用纯矩阵的技巧的一种解法. 首先 $\mathrm{r}(A) < n-1$ 和 $\mathrm{r}(A) = n$ 的情形是很容易的, $\mathrm{r}(A) = n-1$ 的情形才是关键. 下面就考虑这种情形.

显然只要对 $A$ 是 Jordan 标准形的情况加以证明即可. 把 $A$ 按特征值分块为

$$A = \begin{pmatrix} B & 0 \\ 0 & J \end{pmatrix},$$

这里 $B$ 是 $m$ 阶可逆矩阵, $J$ 是一个特征值为 0 的 $l$ 阶 Jordan 块, $m+l=n$. 注意由于有 $\mathrm{r}(A) = n-1$ 的前提, 所以 $A$ 只能有一个特征值为 0 的 Jordan 块.

现在来看 $A^*$. 由于 $AA^* = A^*A = 0$, 所以 $A^*$ 也被同时分块为

$$A^* = \begin{pmatrix} 0 & 0 \\ 0 & C \end{pmatrix}.$$

（因为 $B$ 可逆, 所以左上角也应该是 0）只需要算 $C$, 也就是 $J$ 的伴随矩阵乘以 $|B|$. 而 $J$ 只有左下角的 0 的代数余子式是 $(-1)^{l+1}|B|$, 其它的都是 0. 所以

$$C = |B|J^* = (-1)^{l+1}|B| \begin{pmatrix} 0 & 0 & \cdots & 1 \\ & 0 & \ddots & \vdots \\ & & \ddots & 0 \\ & & & 0 \end{pmatrix}_{l\times l} = (-1)^{l+1}|B|J^{l-1}.$$

好了, 现在的任务是找一个多项式 $f(x)$ 使得

$$f(B) = 0, \quad f(J) = (-1)^{l+1}|B|J^{l-1}.$$

设 $B$ 的特征多项式是 $g(x)$, 那么

$$f(x) = (-1)^{n+1}x^{l-1}g(x)$$

就合乎要求（注意 $g(x)$ 的常数项是 $(-1)^m|B|$）.

这里的 $f(x)$ 一眼看不大出来, 我是这么找的: 设

$$f(x)=a_0+a_1x+\cdots+a_{l-1}x^{l-1}+a_lx^l+a_{l+1}x^{l+1}+\cdots$$

并代入 $f(J)=(-1)^m|B|J^{l-1}$ 就会发现 $a_0=a_1=\cdots=a_{l-2}=0$, $a_{l-1}=(-1)^{l+1}|B|$, 所以

$$f(x)=(-1)^{m+l+1}x^{l-1}\Big((-1)^m|B|+a_lx+a_{l+1}x^2+\cdots\Big),$$

这样就凑出来了.

➤ **例题.** 设 $A,B$ 都是 $n$ 阶矩阵, $C=AB-BA$, 如果 $AC=CA$, 求证 $C$ 是幂零的.

这个问题的传统解法是归纳证明 $\mathrm{tr}C^k=0$. 不过让我们考虑另外一种解法: 在 $C$ 是 Jordan 标准型的前提下来考虑问题, 并且把相同特征值的 Jordan 块合并在一起. 即

$$C=\begin{pmatrix}C_1&&&\\&C_2&&\\&&\ddots&\\&&&C_r\end{pmatrix}.$$

这样不同的 $C_i$ 没有共同的特征值. 这个时候与 $A$ 也被分块为

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

如果这个时候再把 $B$ 划分为 $r^2$ 个子矩阵 $B=(B_{ij})$ 的话, 有意思的事情发生了:

$$C_i=A_iB_{ii}-B_{ii}A_i.$$

这说明 $C_i$ 的迹是 0. 但是 $C_i$ 是上三角矩阵而且对角线上是它仅有的特征值 $\lambda_i$, 所以 $\lambda_i=0$, 从而 $C$ 所有特征值都是 0, 从而幂零.

下面这个例子的解法来自 Halmos.

➤ **例题.** $A,B$ 是数域 $F$ 上的 $n$ 阶方阵, 若 $I-BA$ 可逆, 则 $I-AB$ 也可逆, 并求其逆.

首先用纯形式的推导找出这个逆来: 把矩阵看成数, 作幂级数展开:

$$\begin{aligned}(I-AB)^{-1}&=I+AB+(AB)^2+\cdots\\&=I+A(I+BA+(BA)^2+\cdots)B\\&=I+A(I-BA)^{-1}B.\end{aligned}$$

所以 $(I-AB)^{-1}=I+A(I-BA)^{-1}B$, 这就从形式上找出来了结果, 剩下的只是验证而已.

这个做法确实很巧妙, 而且不用记忆答案, 用到的时候立刻就能推出来.

➤ **例题.** 求证在 $n$ 维欧式空间中两两夹钝角的向量的个数的最大值是 $n+1$.

老题目了, 以前做过一个解答:

http://old.math.org.cn/forums/index.php?showtopic=64867&hl=s

➤ **例题.** 设 $A$ 是一个正定矩阵, 求证 $A$ 的元素中的最大者必然出现在对角线上.

正定矩阵的另一个名字是内积的度量矩阵, 一定不要忘记这一点. 而内积是有 Schwarz 不等式的:

$$|(\alpha,\beta)|^2\le(\alpha,\alpha)(\beta,\beta).$$

也就是 $a_{ij}^2\le a_{ii}a_{jj}$. 所以 $a_{ij}\le\max\{a_{ii},a_{jj}\}$.

最后这个题目来自 Horn 的书:

➤ **例题.** 设 $V$ 是一个向量空间, 是否存在 $V$ 上的线性变换 $A,B$ 使得 $AB-BA=I$?

答案不难, $V$ 是有限维的话就没有, 两边求迹就得出矛盾. $V$ 是无限维的话可能有, 比如在 $C^1(\mathbb{R})$ 上定义

$$Af=f'(x),\quad Bf=xf(x),$$

不难验证就有 $AB-BA=I$ 成立.

H.Wielandt 给出了如下证明:

如果有满足条件的 $A,B$ 存在, 不难用归纳假设证明

$$A^{k+1}B-BA^{k+1}=(k+1)A.$$

设 $\|\cdot\|$ 是 $\mathrm{Mat}_n(F)$ 上任一范数, 那么

$$(k+1)\|A^k\|\le2\|A^{k+1}\|\cdot\|B\|\le2\|A^k\|\cdot\|A\|\cdot\|B\|.$$

这样必有某个 $k$ 使得 $\|A^k\|=0$. 但是上面第一个不等号告诉我们 $A^{k-1}$ 的范数被 $A^k$ 的范数所“控制”, 所以 $A^k=0\Rightarrow A^{k-1}=0\Rightarrow\cdots\Rightarrow A=0$, 这就导出了矛盾.

这个证明的好处就在于它可以应用到 Banach 空间上去, 在量子力学中有如下方程

$$XP-PX=\frac{ih}{2\pi}I.$$

$X$ 是位置算子, $P$ 是动量算子, $h$ 是普朗克常数. Wielandt 的证明告诉我们这样的算子一定不是有界算子, 这解决了一个物理学中长久以来悬而未决的问题.

# 后记

经过反复的思考和修改, 文档终于定稿了. 回顾整个写作过程, 真是很辛苦. 最初积累的素材有很多, 经过半年左右的组织、改写、创作和再创作, 最终浓缩成了这个小小的宝典. 虽然只有短短的三十几页, 内容也只选取了几个有代表意义的课题, 但是我想已经把我想要表达和值得表达的都包含在内了. 与其炖上一锅大杂烩, 倒不如几样精致的小菜来得有滋味.

写文档的目的就是帮助大家更好的理解高等代数的思想和方法, 同时兼顾应对研究生入学考试. 希望用最短的篇幅, 最简洁的语言在最短的时间内帮助读者提高高等代数的水平. 当然这个文档只是讲述了我认为比较重要的部分, 要学好高代或者考好研仅看这些是不够的.

显然我的水平使我只能写到这个程度了, 很多精彩的内容无法讲述, 错误更是在所难免. 很欢迎大家来信指教: zlxida@sina.com.

最后特别感谢周不通老师, 是他的众多帖子帮助我走进了高等代数的大门.