0.1 由乘法交换性诱导的同时性质

命题 0.1 (矩阵乘法可交换的基本性质)

若两个矩阵或线性变换 A, B 乘法可交换, 即 $AB = BA$, 则有 $(AB)^m = A^m B^m$, $f(A)g(B) = g(B)f(A)$ 以及二项式定理

$$(A+B)^m = A^m + \mathrm{C}_m^1 A^{m-1}B + \cdots + \mathrm{C}_m^{m-1}AB^{m-1} + B^m$$

等成立, 其中 $m \geqslant 1$, $f(x), g(x)$ 为多项式.

特别地, 一个矩阵或线性变换 A 一定与其自身可交换, 从而也满足 $f(A)g(A) = g(A)f(A)$, 其中 $f(x), g(x)$ 为多项式.

证明 证明是显然的.

□

0.1.1 特征子空间互为不变子空间

命题 0.2 (特征子空间互为不变子空间)

1. 设 φ, ψ 是复线性空间 V 上乘法可交换的线性变换, 即 $\varphi\psi = \psi\varphi$, 求证: φ 的特征子空间是 ψ 的不变子空间, ψ 的特征子空间是 φ 的不变子空间.
2. 若 n 阶复矩阵 A, B 乘法可交换, 即 $AB = BA$, 则 A, B 的特征子空间互为不变子空间.

注 这个命题的结论对一般的数域是不成立的. 例如, $A = I_2$, $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, 显然 A, B 乘法可交换, 但它们在有理数域或实数域上没有公共的特征向量. 事实上, B 在有理数域或实数域上都没有特征值 (它的特征值是 $\pm\mathrm{i}$), 从而也没有特征向量, 所以更谈不上公共的特征向量了. 为了这个命题的结论推广到数域 $\mathbb{F}$ 上, 我们必须假设 A, B 的特征值都在 $\mathbb{F}$ 中.

证明

1. 由代数基本定理以及线性方程组的求解理论可知, $n(n \geqslant 1)$ 维复线性空间上的线性变换或 n 阶复矩阵至少有一个特征值和特征向量. 任取线性变换 φ 的一个特征值 λ_0, 设 V_0 是特征值 λ_0 的特征子空间, 则对任意的 $\alpha \in V_0$, 有

$$\varphi\psi(\alpha) = \psi\varphi(\alpha) = \psi(\lambda_0\alpha) = \lambda_0\psi(\alpha),$$

即 $\psi(\alpha) \in V_0$, 因此 V_0 是 ψ 的不变子空间. 同理可证 ψ 的特征子空间是 φ 的不变子空间.

2.

□

命题 0.3

设 V 为 n 维复线性空间, S 是 $\mathcal{L}(V)$ 的非空子集, 满足: S 中的全体线性变换没有非平凡的公共不变子空间. 设线性变换 φ 与 S 中任一线性变换乘法均可交换, 证明: φ 是纯量变换.

证明 任取 φ 的特征值 λ_0 及其特征子空间 V_0. 任取 $\psi \in S$, 则 $\varphi\psi = \psi\varphi$, 由命题 0.2可知 V_0 是 $\psi-$ 不变子空间, 从而是 S 中全体线性变换的公共不变子空间. 又 $V_0 \neq 0$(特征向量均非零), 故 $V_0 = V$, 从而 $\varphi = \lambda_0 I_V$ 为纯量变换.

□

0.1.2 有公共的特征向量

命题 0.4

1. 设 φ, ψ 是复线性空间 V 上乘法可交换的线性变换, 求证: φ, ψ 至少有一个公共的 (复) 特征向量.
2. 若 n 阶复矩阵 A, B 乘法可交换, 即 $AB = BA$, 则 A, B 至少有一个公共的 (复) 特征向量.

注 这个命题的结论对一般的数域是不成立的. 例如, $A = I_2, B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, 显然 A, B 乘法可交换, 但它们在有理数域或实数域上没有公共的特征向量. 事实上, B 在有理数域或实数域上都没有特征值 (它的特征值是 $\pm\mathrm{i}$), 从而也没有特征向量, 所以更谈不上公共的特征向量了. 为了这个命题的结论推广到数域 $\mathbb{F}$ 上, 我们必须假设 A, B 的特征值都在 $\mathbb{F}$ 中.

证明

1. 任取 φ 的特征值 λ_0 及其特征子空间 V_0, 由命题 0.2可知, V_0 是 $\psi-$ 不变子空间. 将线性变换 ψ 限制在 V_0 上, 由于 V_0 是维数大于零的复线性空间, 故由命题??可知 $\psi|_{V_0}$ 至少有一个特征值 μ_0 及其特征向量 $\alpha \in V_0$, 从而 $\varphi(\alpha) = \lambda_0\alpha, \psi(\alpha) = \mu_0\alpha$, 于是 α 就是 φ, ψ 的公共特征向量.
2.

□

命题 0.5

1. 设 φ, ψ 是数域 $\mathbb{F}$ 上线性空间 V 上的乘法可交换的线性变换, 且 φ, ψ 的特征值都在 $\mathbb{F}$ 中, 求证: φ, ψ 的特征子空间互为不变子空间, 并且 φ, ψ 至少有一个公共的特征向量.
2. 若数域 $\mathbb{F}$ 上的 n 阶矩阵 A, B 乘法可交换, 且它们的特征值都在 $\mathbb{F}$ 中, 则 A, B 的特征子空间互为不变子空间, 并且 A, B 在 $\mathbb{F}^n$ 中至少有一个公共的特征向量.

证明

1. 由线性方程组的求解理论可知, 若数域 $\mathbb{F}$ 上的线性变换或 $\mathbb{F}$ 上的矩阵在 $\mathbb{F}$ 中有一个特征值, 则在 $\mathbb{F}$ 上的线性空间或 $\mathbb{F}$ 上的列向量空间中必存在对应的特征向量. 任取线性变换 φ 的一个特征值 $\lambda_0 \in \mathbb{F}$, 设 V_0 是特征值 λ_0 的特征子空间, 则对任意的 $\alpha \in V_0$, 有
$$\varphi\psi(\alpha) = \psi\varphi(\alpha) = \psi(\lambda_0\alpha) = \lambda_0\psi(\alpha),$$
即 $\psi(\alpha) \in V_0$, 因此 V_0 是 $\psi-$ 不变子空间. 取 V_0 的一组基并扩张为 V 的一组基, 则 ψ 在这组基下的表示矩阵为分块对角矩阵 $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$, 其中 A 是 $\psi|_{V_0}$ 在给定基下的表示矩阵, 于是 $|\lambda I_V - \psi| = |\lambda I - A||\lambda I - B|$. 因为 ψ 的特征值都在 $\mathbb{F}$ 中, 故 A 的特征值都在 $\mathbb{F}$ 中, 于是 $\psi|_{V_0}$ 的特征值都在 $\mathbb{F}$ 中. 任取 $\psi|_{V_0}$ 的一个特征值 $\mu_0 \in \mathbb{F}$ 及其特征向量 $\alpha \in V_0$, 则 $\varphi(\alpha) = \lambda_0\alpha, \psi(\alpha) = \mu_0\alpha$, 于是 α 就是 φ, ψ 的公共特征向量.
2.

□

0.1.3 可同时相似上三角化

命题 0.6 (矩阵的上三角化)

1. 设数域 $\mathbb{F}$ 上的 n 阶矩阵 A 的特征值都在 $\mathbb{F}$ 中, 求证: A 在 $\mathbb{F}$ 上可上三角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}AP$ 是上三角矩阵.
2. 设数域 $\mathbb{F}$ 上线性空间 V 上的线性变换 φ 的特征值都在 $\mathbb{F}$ 中, 则存在 V 的一组基, 使得 φ 在这组基下的表示矩阵是上三角矩阵.

证明

1. 对阶数进行归纳. 当 $n=1$ 时结论显然成立, 设对 $n-1$ 阶矩阵结论成立, 现对 n 阶矩阵 A 进行证明. 设 $\lambda_1\in\mathbb{F}$ 是 A 的一个特征值, 则由线性方程组的求解理论可知, 存在特征向量 $e_1\in\mathbb{F}^n$, 使得 $Ae_1=\lambda_1 e_1$. 由基扩张定理, 可将 e_1 扩张为 $\mathbb{F}^n$ 的一组基 $\{e_1,e_2,\cdots,e_n\}$, 于是

$$(Ae_1,Ae_2,\cdots,Ae_n)=(e_1,e_2,\cdots,e_n)\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix},$$

其中 A_1 是 $\mathbb{F}$ 上的 $n-1$ 阶矩阵. 令 $P=(e_1,e_2,\cdots,e_n)$, 则 P 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 且由上式可得 $AP=P\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}$, 即 $P^{-1}AP=\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}$. 由此可得 $|\lambda I_n-A|=(\lambda-\lambda_1)|\lambda I_{n-1}-A_1|$, 又 A 的特征值全在 $\mathbb{F}$ 中, 从而 A_1 的特征值也全在 $\mathbb{F}$ 中, 故由归纳假设, 存在 $\mathbb{F}$ 上的 $n-1$ 阶可逆矩阵 Q, 使得 $Q^{-1}A_1Q$ 是上三角矩阵. 令

$$R=P\begin{pmatrix}1 & O\\ O & Q\end{pmatrix},$$

则 R 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 且

$$R^{-1}AR=\begin{pmatrix}1 & O\\ O & Q\end{pmatrix}^{-1}\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\begin{pmatrix}1 & O\\ O & Q\end{pmatrix}=\begin{pmatrix}\lambda_1 & *\\ O & Q^{-1}A_1Q\end{pmatrix}$$

是上三角矩阵.

2.

□

命题 0.7

1. 设 A,B 是数域 $\mathbb{F}$ 上的 n 阶矩阵, 满足: $AB=BA$ 且 A,B 的特征值都在 $\mathbb{F}$ 中, 求证: A,B 在 $\mathbb{F}$ 上可同时上三角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}AP$ 和 $P^{-1}BP$ 都是上三角矩阵.
2. 设数域 $\mathbb{F}$ 上线性空间 V 上的线性变换 φ,ψ 乘法可交换, 且它们的特征值都在 $\mathbb{F}$ 中, 则存在 V 的一组基, 使得 φ,ψ 在这组基下的表示矩阵都是上三角矩阵.

♠

证明

1. 对阶数进行归纳. 当 $n=1$ 时结论显然成立, 设对 $n-1$ 阶矩阵结论成立, 现对 n 阶矩阵进行证明. 因为 $AB=BA$ 且 A,B 的特征值都在 $\mathbb{F}$ 中, 故由命题 0.5可知, A,B 有公共的特征向量 $e_1\in\mathbb{F}^n$, 不妨设

$$Ae_1=\lambda_1 e_1,\ Be_1=\mu_1 e_1,$$

其中 $\lambda_1,\mu_1\in\mathbb{F}$ 分别是 A,B 的特征值. 由基扩张定理, 可将 e_1 扩张为 $\mathbb{F}^n$ 的一组基 $\{e_1,e_2,\cdots,e_n\}$. 令 $P=(e_1,e_2,\cdots,e_n)$, 则 P 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 从而有

$$\begin{aligned}A(e_1,e_2,\cdots,e_n)=(e_1,e_2,\cdots,e_n)\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\Leftrightarrow AP=P\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\Leftrightarrow P^{-1}AP=\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix},\\ B(e_1,e_2,\cdots,e_n)=(e_1,e_2,\cdots,e_n)\begin{pmatrix}\lambda_1 & *\\ O & B_1\end{pmatrix}\Leftrightarrow BP=P\begin{pmatrix}\lambda_1 & *\\ O & B_1\end{pmatrix}\Leftrightarrow P^{-1}BP=\begin{pmatrix}\lambda_1 & *\\ O & B_1\end{pmatrix}.\end{aligned}\tag{1}$$

其中 A_1,B_1 是 $\mathbb{F}$ 上的 $n-1$ 阶矩阵. 由 $AB=BA$ 及(1)式可得到

$$\begin{aligned}&\left(P^{-1}AP\right)\left(P^{-1}BP\right)=P^{-1}ABP=P^{-1}BAP=\left(P^{-1}BP\right)\left(P^{-1}AP\right)\\ \Leftrightarrow&\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\begin{pmatrix}\lambda_1 & *\\ O & B_1\end{pmatrix}=\begin{pmatrix}\lambda_1 & *\\ O & B_1\end{pmatrix}\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\\ \Leftrightarrow&\begin{pmatrix}\lambda_1 & *\\ O & A_1B_1\end{pmatrix}=\begin{pmatrix}\lambda_1 & *\\ O & B_1A_1\end{pmatrix}\end{aligned}$$

从而 $A_1B_1=B_1A_1$. 又由(1)式可得

$$|\lambda I_n-A|=|\lambda-\lambda_1|\,|\lambda I_{n-1}-A_1|,\quad |\lambda I_n-B|=|\lambda-\lambda_1|\,|\lambda I_{n-1}-B_1|.$$

因此 A_1,B_1 的特征值也是 A,B 的特征值. 又由于 A,B 的特征值都在 $\mathbb{F}$ 中, 故 A_1,B_1 的特征值都在 $\mathbb{F}$ 中. 故由归纳假设, 存在 $\mathbb{F}$ 上的 $n-1$ 阶可逆矩阵 Q, 使得 $Q^{-1}A_1Q$ 和 $Q^{-1}B_1Q$ 都是上三角矩阵. 令

$$R=P\begin{pmatrix}1&O\\O&Q\end{pmatrix},$$

则 R 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 且

$$R^{-1}AR=\begin{pmatrix}1&O\\O&Q\end{pmatrix}^{-1}\begin{pmatrix}\lambda_1&*\\O&A_1\end{pmatrix}\begin{pmatrix}1&O\\O&Q\end{pmatrix}=\begin{pmatrix}\lambda_1&*\\O&Q^{-1}A_1Q\end{pmatrix},$$

$$R^{-1}BR=\begin{pmatrix}1&O\\O&Q\end{pmatrix}^{-1}\begin{pmatrix}\mu_1&*\\O&B_1\end{pmatrix}\begin{pmatrix}1&O\\O&Q\end{pmatrix}=\begin{pmatrix}\mu_1&*\\O&Q^{-1}B_1Q\end{pmatrix}$$

都是上三角矩阵.

2.

□

命题 0.8 (一族两两可交换的一般域上的矩阵可同时上三角化)

给定域 $\mathbb{F}$ 和指标集 Λ, 设 $A_\lambda\in\mathbb{F}^{n\times n},\lambda\in\Lambda$ 且两两可交换且特征值都属于 $\mathbb{F}$, 则存在可逆矩阵 $P\in\mathbb{F}^{n\times n}$, 使得

$$P^{-1}A_\lambda P\text{ 是上三角矩阵},\forall\lambda\in\Lambda.$$

笔记 证明的想法是对有限的量归纳, 即矩阵降阶. 本结果将综合运用几何方法和矩阵方法.

注 因为数量矩阵的特征子空间就是全空间, 将其限制在特征子空间上, 维数并未下降, 所以需要分类讨论.

证明 **Step 1** 若 $\forall\lambda\in\Lambda$, 都有 A_λ 是数量矩阵, 此时结论显然成立.

Step 2 任取一个非数量矩阵 $A_1\in\mathbb{F}^{n\times n}$, 再任取 A_1 的一个特征子空间 V_1, 则 $1\leqslant\dim V_1<n$, 否则 A_1 就是纯量阵. 由命题 0.2可知,V_1 是 $A_\lambda-$ 不变子空间,$\forall\lambda\in\Lambda$. 因此可考虑线性变换 $A_\lambda|_{V_1},\lambda\in\Lambda$. 下对矩阵阶数进行归纳证明.

当 $n=1$ 时, 结论显然成立. 假设命题对小于等于 $n-1$ 的情况都成立, 考虑 n 的情形.

注意到 $A_\lambda|_{V_1},\lambda\in\Lambda$ 两两乘法可交换, 故由归纳假设可知, 存在 V_1 的一组基, 使 $A_\lambda|_{V_1}$ 在这组基下有上三角表示矩阵 $\widetilde{A}_\lambda,\lambda\in\Lambda$. 将这组基扩充为 V 的一组基, 于是在新的基下,$A_\lambda\,(\lambda\in\Lambda)$ 有表示矩阵 $\begin{pmatrix}\widetilde{A}_\lambda&\widetilde{B}_\lambda\\&\widetilde{C}_\lambda\end{pmatrix},\lambda\in\Lambda$.

又由于 $A_\lambda,\lambda\in\Lambda$ 两两乘法可交换, 故对 $\forall\lambda,\mu\in\Lambda$, 有

$$\begin{pmatrix}\widetilde{A}_\lambda&\widetilde{B}_\lambda\\&\widetilde{C}_\lambda\end{pmatrix}\begin{pmatrix}\widetilde{A}_\mu&\widetilde{B}_\mu\\&\widetilde{C}_\mu\end{pmatrix}=\begin{pmatrix}\widetilde{A}_\mu&\widetilde{B}_\mu\\&\widetilde{C}_\mu\end{pmatrix}\begin{pmatrix}\widetilde{A}_\lambda&\widetilde{B}_\lambda\\&\widetilde{C}_\lambda\end{pmatrix}$$

$$\Longleftrightarrow\begin{pmatrix}\widetilde{A}_\lambda\widetilde{A}_\mu&*\\&\widetilde{C}_\lambda\widetilde{C}_\mu\end{pmatrix}=\begin{pmatrix}\widetilde{A}_\mu\widetilde{A}_\lambda&*\\&\widetilde{C}_\mu\widetilde{C}_\lambda\end{pmatrix}$$

即 $\widetilde{A}_\lambda,\widetilde{C}_\lambda,\lambda\in\Lambda$ 两两乘法可交换. 从而由归纳假设可知, 对 $\forall\lambda\in\Lambda$, 存在可逆阵 $\widetilde{P}_\lambda$, 使得 $(\widetilde{P}_\lambda)^{-1}\widetilde{C}_\lambda\widetilde{P}_\lambda$ 是上三角阵. 取 $P=\begin{pmatrix}I&O\\O&\widetilde{P}_\lambda\end{pmatrix}\in\mathbb{F}^{n\times n}$, 则此时对 $\forall\lambda\in\Lambda$, 就有

$$P^{-1}A_\lambda P=\begin{pmatrix}\widetilde{A}_\lambda&\widetilde{B}_\lambda\\&(\widetilde{P}_\lambda)^{-1}\widetilde{C}_\lambda\widetilde{P}_\lambda\end{pmatrix}.$$

而 $\widetilde{A}_\lambda,(\widetilde{P}_\lambda)^{-1}\widetilde{C}_\lambda\widetilde{P}_\lambda$ 都是上三角阵, 故 $P^{-1}A_\lambda P$ 也是上三角阵. 因此由数学归纳法可知, 结论成立.

□

命题 0.9 (一族两两可交换的复数 (实数) 域上的矩阵可同时酉 (正交) 上三角化)

1. 给定指标集 Λ, 设 $A_\lambda \in \mathbb{C}^{n\times n}, \lambda \in \Lambda$ 且两两可交换, 则存在酉矩阵 $P \in \mathbb{C}^{n\times n}$, 使得 $P^{-1}A_\lambda P$ 是上三角矩阵, $\forall \lambda \in \Lambda$.
2. 给定指标集 Λ, 设 $A_\lambda \in \mathbb{R}^{n\times n}, \lambda \in \Lambda$ 且两两可交换且特征值都是实数. 则存在正交矩阵 $P \in \mathbb{R}^{n\times n}$, 使得 $P^{-1}A_\lambda P$ 是上三角矩阵, $\forall \lambda \in \Lambda$.

笔记 证明的想法是对有限的量归纳, 即矩阵降阶. 本结果将综合运用几何方法和矩阵方法.

注 因为数量矩阵的特征子空间就是全空间, 将其限制在特征子空间上, 维数并未下降, 所以需要分类讨论.

证明

1. 设 $V=\mathbb{C}^n$ 且 A_λ 是 V 上线性变换.

   Step 1 若 A_λ 都是数量矩阵, 则结果已经成立.

   Step 2 取某个非数量矩阵 A_1 和一个特征子空间 V_1 且 $1 \leqslant \dim V_1 < n$. 由交换性知 V_1 是所有 A_λ 不变子空间, 因此 $A_\lambda|_{V_1}, \lambda \in \Lambda$ 也是一族更低维度的两两可交换的矩阵. 于是我们就将维度降了下去, 从而可以使用归纳法来完成证明. 即:

   当 $n=1$, 命题显然成立, 假设命题对小于等于 $n-1$ 都成立, 当 n 时, 由归纳假设, $A_\lambda|_{V_1}, \lambda \in \Lambda$ 也是一族两两可交换的矩阵, 从而存在 V_1 的一族标准正交基, 使得在这组基下 $A_\lambda|_{V_1}$ 有上三角的表示矩阵 $\widetilde{A}_\lambda, \lambda \in \Lambda$. 将这组基扩充到 $\mathbb{C}^n$ 使得构成一组标准正交基, 则在新的标准正交基下, A_λ 有表示矩阵 $\begin{pmatrix} \widetilde{A}_\lambda & \widetilde{B}_\lambda \\ 0 & \widetilde{C}_\lambda \end{pmatrix}, \lambda \in \Lambda$. 由

$$\begin{pmatrix} \widetilde{A}_\lambda & \widetilde{B}_\lambda \\ 0 & \widetilde{C}_\lambda \end{pmatrix}\begin{pmatrix} \widetilde{A}_\mu & \widetilde{B}_\mu \\ 0 & \widetilde{C}_\mu \end{pmatrix} = \begin{pmatrix} \widetilde{A}_\mu & \widetilde{B}_\mu \\ 0 & \widetilde{C}_\mu \end{pmatrix}\begin{pmatrix} \widetilde{A}_\lambda & \widetilde{B}_\lambda \\ 0 & \widetilde{C}_\lambda \end{pmatrix}, \mu, \lambda \in \Lambda.$$

   知 $\widetilde{C}_\lambda, \lambda \in \Lambda$ 也是两两可交换的矩阵. 因此存在酉矩阵 $\widetilde{P}$, 使得每一个 $\left(\widetilde{P}\right)^{-1}\widetilde{C}_\lambda\widetilde{P}$ 都是上三角的. 然后我们取酉矩阵 $P=\begin{pmatrix} E & 0 \\ 0 & \widetilde{P} \end{pmatrix} \in \mathbb{C}^{n\times n}$, 就有 $P^{-1}A_\lambda P = \begin{pmatrix} \widetilde{A}_\lambda & \widetilde{B}_\lambda \\ 0 & \left(\widetilde{P}\right)^{-1}\widetilde{C}_\lambda\widetilde{P} \end{pmatrix}, \forall \lambda \in \Lambda$ 都是上三角矩阵, 我们完成了证明.

2. 设 $V=\mathbb{R}^n$ 且 A_λ 是 V 上线性变换.

   Step 1 若 A_λ 都是数量矩阵, 则结果已经成立.

   Step 2 取某个非数量矩阵 A_1 和一个特征子空间 V_1 且 $1 \leqslant \dim V_1 < n$. 由交换性知 V_1 是所有 A_λ 不变子空间, 因此 $A_\lambda|_{V_1}, \lambda \in \Lambda$ 也是一族更低维度的两两可交换的矩阵. 于是我们就将维度降了下去, 从而可以使用归纳法来完成证明. 即:

   当 $n=1$, 命题显然成立, 假设命题对小于等于 $n-1$ 都成立, 当 n 时, 由归纳假设, $A_\lambda|_{V_1}, \lambda \in \Lambda$ 也是一族两两可交换的矩阵, 从而存在 V_1 的一族标准正交基, 使得在这组基下 $A_\lambda|_{V_1}$ 有上三角的表示矩阵 $\widetilde{A}_\lambda, \lambda \in \Lambda$. 将这组基扩充到 $\mathbb{R}^n$ 使得构成一组标准正交基, 则在新的标准正交基下, A_λ 有表示矩阵 $\begin{pmatrix} \widetilde{A}_\lambda & \widetilde{B}_\lambda \\ 0 & \widetilde{C}_\lambda \end{pmatrix}, \lambda \in \Lambda$. 由

$$\begin{pmatrix} \widetilde{A}_\lambda & \widetilde{B}_\lambda \\ 0 & \widetilde{C}_\lambda \end{pmatrix}\begin{pmatrix} \widetilde{A}_\mu & \widetilde{B}_\mu \\ 0 & \widetilde{C}_\mu \end{pmatrix} = \begin{pmatrix} \widetilde{A}_\mu & \widetilde{B}_\mu \\ 0 & \widetilde{C}_\mu \end{pmatrix}\begin{pmatrix} \widetilde{A}_\lambda & \widetilde{B}_\lambda \\ 0 & \widetilde{C}_\lambda \end{pmatrix}$$

   知 $\widetilde{C}_\lambda, \lambda \in \Lambda$ 也是两两可交换的矩阵. 因此存在正交矩阵 $\widetilde{P}$, 使得每一个 $\left(\widetilde{P}\right)^{-1}\widetilde{C}_\lambda\widetilde{P}$ 都是上三角的. 然后我们取正交矩阵 $P=\begin{pmatrix} E & 0 \\ 0 & \widetilde{P} \end{pmatrix} \in \mathbb{R}^{n\times n}$, 就有 $P^{-1}A_\lambda P = \begin{pmatrix} \widetilde{A}_\lambda & \widetilde{B}_\lambda \\ 0 & \left(\widetilde{P}\right)^{-1}\widetilde{C}_\lambda\widetilde{P} \end{pmatrix}, \forall \lambda \in \Lambda$ 都是上三角矩阵, 我们完成了证明.

□

0.1.4 可同时相似对角化

命题 0.10

1. 设 φ,ψ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 上的线性变换, 满足: $\varphi\psi=\psi\varphi$ 且 φ,ψ 都可对角化, 求证: φ,ψ 可同时对角化, 即存在 V 的一组基, 使得 φ,ψ 在这组基下的表示矩阵都是对角矩阵.
2. 设 A,B 是数域 $\mathbb{F}$ 上的 n 阶矩阵, 满足: $AB=BA$ 且 A,B 都在 $\mathbb{F}$ 上可对角化, 则 A,B 在 $\mathbb{F}$ 上可同时对角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}AP$ 和 $P^{-1}BP$ 都是对角矩阵.

证明

1. 对空间维数进行归纳. 当 $n=1$ 时结论显然成立, 设对维数小于 n 的线性空间结论成立, 现对 n 维线性空间进行证明. 设 φ 的全体不同特征值为 $\lambda_1,\cdots,\lambda_s\in\mathbb{F}$, 对应的特征子空间分别为 $V_1,\cdots,V_s$, 则由 φ 可对角化可知
$$V=V_1\oplus\cdots\oplus V_s.$$
若 $s=1$, 则 $\varphi=\lambda_1 I_V$ 为纯量变换, 此时只要取 V 的一组基, 使得 ψ 在这组基下的表示矩阵为对角矩阵, 则 φ 在这组基下的表示矩阵为 $\lambda_1 I_n$, 结论成立. 若 $s>1$, 则 $\dim V_i<n$. 注意到 $\varphi\psi=\psi\varphi$ 且 φ,ψ 的特征值都在 $\mathbb{F}$ 中, 由命题 0.2可知 V_i 都是 $\psi-$ 不变子空间. 考虑线性变换的限制 $\varphi|_{V_i},\psi|_{V_i}$: 它们乘法可交换, 且由可对角化线性变换的性质可知它们都可对角化, 故由归纳假设可知, $\varphi|_{V_i},\psi|_{V_i}$ 可同时对角化, 即存在 V_i 的一组基, 使得 $\varphi|_{V_i},\psi|_{V_i}$ 在这组基下的表示矩阵都是对角矩阵. 将 V_i 的基拼成 V 的一组基, 则 φ,ψ 在这组基下的表示矩阵都是对角矩阵, 即 φ,ψ 可同时对角化.
2.

□

引理 0.1

任何域上的可对角化矩阵限制到不变子空间上仍然是可对角化矩阵.

证明 注意到矩阵可对角化等价于极小多项式可以分解为一次式的积. 设 A 的极小多项式是 p, W 是矩阵 A 一个不变子空间且 p_W 是 $A|_W$ 极小多项式. 注意到 $p(A|_W)=0$, 故 $p_W|p$. 从而 p_W 也是一次式的积, 故 $A|_W$ 可对角化.

□

命题 0.11 (一族两两可交换的可对角化矩阵可同时相似对角化)

给定域 $\mathbb{F}$, 设 $A_\lambda\in\mathbb{F}^{n\times n},\lambda\in\Lambda$ 且两两可交换. 若每一个 $A_\lambda,\lambda\in\Lambda$ 都可以在 $\mathbb{F}$ 上相似对角化, 则存在可逆矩阵 $P\in\mathbb{F}^{n\times n}$, 使得
$$P^{-1}A_\lambda P \text{ 是对角矩阵}, \forall\lambda\in\Lambda.$$

证明 设 $V=\mathbb{F}^n$ 且 A_λ 是 V 上线性变换.

Step 1 若 A_λ 都是数量矩阵, 则结果已经成立.

Step 2 取某个非数量矩阵 A_1, 于是有 $V=\bigoplus\limits_{i=1}^{s}V_i$, 这里 $s\geqslant 2$ 且 V_i 是属于 A_1 不同特征值的特征子空间. 显然由交换性, 对每一个 $i=1,2,\cdots,s$, V_i 是所有 A_λ 不变子空间, 且 $A_\lambda|_{V_i},\lambda\in\Lambda$ 是一族两两可交换的矩阵, 由引理 0.1知它们也是可对角化的. 注意到 $1\leqslant\dim V_i<n, i=1,2,\cdots,s$, 所以我们的维度降下去了, 因此可使用归纳法.

当 $n=1$, 命题显然成立, 假设命题对 $\leqslant n-1$ 都成立, 当 n 时, 由归纳假设, 对每一个 $i=1,2,\cdots,s$, 存在 V_i 的一个基使得 $A_\lambda|_{V_i},\lambda\in\Lambda$ 在这个基下表示矩阵是对角矩阵. 于是把这些基合起来构成一个新的基, 我们就得到在这个新的基下 $A_\lambda,\lambda\in\Lambda$ 都是对角矩阵.

□

命题 0.12 (一族两两可交换的复正规 (实对称) 矩阵可同时酉 (正交) 相似对角化)

1. 给定指标集 Λ, 设 $A_\lambda\in\mathbb{C}^{n\times n},\lambda\in\Lambda$ 且两两可交换且复正规, 则存在酉矩阵 $P\in\mathbb{C}^{n\times n}$, 使得
$$P^{-1}A_\lambda P\text{ 是对角矩阵}, \forall\lambda\in\Lambda.$$
2. 给定指标集 Λ, 设 $A_\lambda\in\mathbb{R}^{n\times n},\lambda\in\Lambda$ 且两两可交换且实对称, 则存在正交矩阵 $P\in\mathbb{R}^{n\times n}$, 使得
$$P^{-1}A_\lambda P\text{ 是对角矩阵}, \forall\lambda\in\Lambda.$$

笔记 注意到 $(Ax,y)=(x,A^Ty),(Ax,y)=(x,A^*y)$ 在全空间成立则在子空间也成立, 所以 $A^T|_V=(A|_V)^T, A^*|_V=(A|_V)^*$, 所以一个实对称变换限制在不变子空间也是实对称的, 一个复正规变换限制在不变子空间也是复正规的.

证明

1. 设 $V=\mathbb{C}^n$ 且 A_λ 是 V 上线性变换.

   Step 1 若 A_λ 都是数量矩阵, 则结果已经成立.

   Step 2 取某个非数量矩阵 A_1, 于是有 $V=\bigoplus\limits_{i=1}^{s}V_i$, 这里 $s\geqslant 2$ 且 V_i 是属于 A_1 不同特征值的特征子空间. 显然由交换性, 对每一个 $i=1,2,\cdots,s$, V_i 是所有 A_λ 不变子空间, 从而 V_i 两两正交, 且 $A_\lambda|_{V_i},\lambda\in\Lambda$ 是一族两两可交换的正规矩阵, 于是可酉对角化的. 注意到 $1\leqslant\dim V_i<n, i=1,2,\cdots,s$. 所以我们的维度降下去了, 因此可使用归纳法.

   当 $n=1$, 命题显然成立, 假设命题对 $\leqslant n-1$ 都成立, 当 n 时, 由归纳假设, 对每一个 $i=1,2,\cdots,s$, 存在 V_i 的一个标准正交基使得 $A_\lambda|_{V_i},\lambda\in\Lambda$ 在这个基下表示矩阵是对角矩阵. 由于 V_i 两两正交, 于是把这些基合起来构成一个新的正交基, 我们就得到在这个新的基下 $A_\lambda,\lambda\in\Lambda$ 都是对角矩阵.
2. 设 $V=\mathbb{R}^n$ 且 A_λ 是 V 上线性变换.

   Step 1 若 A_λ 都是数量矩阵, 则结果已经成立.

   Step 2 取某个非数量矩阵 A_1, 于是有 $V=\bigoplus\limits_{i=1}^{s}V_i$, 这里 $s\geqslant 2$ 且 V_i 是属于 A_1 不同特征值的特征子空间. 显然由交换性, 对每一个 $i=1,2,\cdots,s$, V_i 是所有 A_λ 不变子空间, 从而 V_i 两两正交, 且 $A_\lambda|_{V_i},\lambda\in\Lambda$ 是一族两两可交换的实对称矩阵, 由引理 0.1知它们也是可正交相似对角化的. 注意到 $1\leqslant\dim V_i<n, i=1,2,\cdots,s$, 所以我们的维度降下去了, 因此可使用归纳法.

   当 $n=1$, 命题显然成立, 假设命题对 $\leqslant n-1$ 都成立, 当 n 时, 由归纳假设, 对每一个 $i=1,2,\cdots,s$, 存在 V_i 的一个标准正交基使得 $A_\lambda|_{V_i},\lambda\in\Lambda$ 在这个基下表示矩阵是对角矩阵. 由于 V_i 两两正交, 于是把这些基合起来构成一个新的正交基, 我们就得到在这个新的基下 $A_\lambda,\lambda\in\Lambda$ 都是对角矩阵.

□

引理 0.2 (根子空间维数)

设 $A\in\mathbb{F}^{n\times n}$ 且 $\mathbb{F}$ 是域. 若 $\lambda\in\mathbb{F}$ 是 A 的 r 重特征值, 则我们有
$$\dim\operatorname{Ker}(\lambda E-A)^m=r,\forall m\geqslant r. \tag{2}$$

证明 注意到
$$\dim\operatorname{Ker}(\lambda E-A)^r=n-\operatorname{rank}\left((\lambda E-A)^r\right).$$
而秩不随域扩张而改变, 故不妨设 $\mathbb{F}$ 是代数闭域. 则不妨设 A 为 Jordan 标准型
$$\begin{pmatrix} J_{n_1}(\lambda) & & & \\ & \ddots & & \\ & & J_{n_s}(\lambda) & \\ & & & J \end{pmatrix},\quad \sum_{i=1}^{s}n_i=r,$$

这里 J 没有特征值 λ. 于是直接计算有

$$(A-\lambda E)^r = \begin{pmatrix} J_{n_1}^r(0) & & & \\ & \ddots & & \\ & & J_{n_s}^r(0) & \\ & & & (J-\lambda E_{n-r})^r \end{pmatrix} = \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & (J-\lambda E_{n-r})^r \end{pmatrix},$$

$$\det(J-\lambda E_{n-r})^r \neq 0.$$

现在知 $\operatorname{rank}(A-\lambda E)^r = n-r$, 这就得到了 (2) 的 $m=r$ 的情况. 注意到当 $m>r$ 时, 上述矩阵 0 块在次方时不会继续丢失阶数, 而 $(J-\lambda E)^m$ 仍可逆, 因此秩不会再次减少, 所以 (2) 对任何 $m \geqslant r$ 都成立, 这就完成了证明.

□

0.1.5 其他

定理 0.1 (公共特征值)

给定域 $\mathbb{F}$, 指标集 Λ 和 $A_\lambda \in \mathbb{F}^{n\times n}, \lambda \in \Lambda$.

1. 若 $A_\lambda, \lambda \in \Lambda$ 两两可交换;
2. 若 $\mathbb{F}=\mathbb{R}, n$ 为奇数.

则 A_λ 在 $\mathbb{F}$ 上有公共特征向量.

♡

笔记 某种角度上说, 存在公共特征向量和可同时相似上三角化是等价的.

1. 由一族两两可交换的一般域上的矩阵可同时上三角化, 不妨设 $A_\lambda, \lambda \in \Lambda$ 都是上三角矩阵, 注意观察第一列知他们有公共的特征向量 e_1.

2. 证明的关键在于奇数次实多项式必有实根, 从而奇数矩阵必有实特征值.

证明 **Step 1** 取 $\{A_\lambda\}_{\lambda\in\Lambda}$ 的极大无关组 $A_1, A_2, \cdots, A_k$. 如果能找到 $A_1, A_2, \cdots, A_k$ 的公共特征向量 $\alpha \in \mathbb{R}^n$, 则对任何 $A_\lambda = \sum_{i=1}^k c_i A_i$, 都有

$$A_\lambda \alpha = \sum_{i=1}^k c_i A_i \alpha = \left(\sum_{i=1}^k c_i \lambda_i\right) \alpha,$$

这里 λ_i 是 $A_i, 1 \leqslant i \leqslant k$ 关于特征向量 α 的特征值. 于是问题变成只需要对有限个矩阵讨论.

Step 2 对于两两可交换的实矩阵 $A_1, A_2, \cdots, A_k$. 取 $\{A_\lambda\}_{\lambda\in\Lambda}$ 的极大无关组 $A_1, A_2, \cdots, A_k$. 设 $\lambda \in \mathbb{R}$ 是 A_1 的奇数重特征值, 记重数为 $r \in \mathbb{N}$ 并设 $W = \operatorname{Ker}((\lambda E - A_1)^r)$. 由引理 0.2知 $\dim W = r$ 且显然有 $A_j|_W, j = 2, 3, \cdots, k$ 两两可交换.

注意到 $r=n$ 是可能的, 真正降下去的是矩阵个数, 所以我们对矩阵个数归纳. 命题对 $k=1$ 显然成立, 假设命题对 $\leqslant k-1$ 成立, 当 k 时我们运用归纳假设知 $A_j|_W, j = 2, 3, \cdots, k$ 有公共特征向量, 即存在实数 $\lambda_i, 2 \leqslant i \leqslant k$, 使得

$$W' = \bigcap_{i=2}^k \operatorname{Ker}(\lambda_i I - A_i|_W) \neq \{0\}.$$

Step 3 由可交换性知 W' 是 A_1 不变子空间, 所以断言 $A_1|_{W'}$ 在 $\mathbb{C}$ 上特征值只有 λ. 事实上由 **Step 1**, 显然 $W' \subset W$. 又设在复数域上 $A_1\beta = \mu\beta, \beta \neq 0, \mu \in \mathbb{C}$, 则

$$0 = (\lambda E - A_1)^r \beta = (\lambda E - A_1)^{r-1}(\lambda\beta - A_1\beta) = (\lambda-\mu)(\lambda E - A_1)^{r-1}\beta = \cdots = (\lambda-\mu)^r \beta,$$

故 $\mu = \lambda$.

Step 4 取 W' 中 A_1 一个特征向量 γ, 则 γ 是 $A_1, A_2, \cdots, A_k$ 公共特征向量, 这就完成了证明.

□

定理 0.2 (Netwon 公式)

设

$$s_j=\sum_{i=1}^{n}x_i^j,\quad \sigma_j=\sum_{1\leqslant i_1<i_2<\cdots<i_j\leqslant n}x_{i_1}x_{i_2}\cdots x_{i_j}, j=1,2,\cdots,n,$$

其中 $\sigma_j(j=1,2,\cdots,n)$ 称为 $x_1,x_2,\cdots,x_n$ 的所有**初等对称多项式**. 我们有

$$s_k-\sigma_1s_{k-1}+\sigma_2s_{k-2}+\cdots+(-1)^{k-1}\sigma_{k-1}s_1+(-1)^kk\sigma_k=0, 1\leqslant k\leqslant n;$$

$$s_k-\sigma_1s_{k-1}+\sigma_2s_{k-2}+\cdots+(-1)^{n-1}\sigma_{n-1}s_{k-n+1}+(-1)^n\sigma_ns_{k-n}=0, k>n.$$

证明

□

定理 0.3

对域 $\mathbb{F}$, 设 $C\in\mathbb{F}^{n\times n}$, 则 C 是幂零矩阵的充要条件是 $\mathrm{tr}(C^k)=0,\forall k\in\mathbb{N}$.

证明 **必要性**: 若 C 是幂零矩阵, 则 C 的特征值全为零, 从而 $\mathrm{tr}(C^k)=\sum_{i=1}^{n}\lambda_i^k=0\ (k\geqslant 1)$.

充分性: 证法一:不妨设 $\mathbb{F}$ 是代数闭的, 设 C 的全部特征值是 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 这里可能相同. 则 C^k 的特征值为 λ_i^k, 从而

$$\mathrm{tr}(C^k)=\sum_{i=1}^{n}\lambda_i^k=0,\forall k\in\mathbb{N},$$

即 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的一切幂和为 0. 由Netwon 公式, 我们有 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 构成的所有初等对称多项式为 0. 由 Vieta 定理, 我们有 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 都是 $\lambda^n=0$ 的根, 这就证明了 C 的全部特征值都是 0, 故 C 是幂零矩阵.

证法二:设 A 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$. 若 $\mathrm{tr}(A^k)=0\ (1\leqslant k\leqslant n)$, 则

(i) 若 $\lambda_i(1\leqslant i\leqslant n)$ 全为零, 则结论显然成立.

(ii) 若 $\lambda_i(1\leqslant i\leqslant n)$ 不全为零, 则不妨设 $\lambda_1,\cdots,\lambda_k\neq 0,\lambda_{k+1}=\cdots=\lambda_n=0$. 再不妨设

$$\begin{gathered}x_1=\lambda_1=\cdots=\lambda_{i_1},\\ x_2=\lambda_{i_1+1}=\cdots=\lambda_{i_2},\\ \cdots\cdots\cdots\cdots\\ x_{s+1}=\lambda_{i_s+1}=\cdots=\lambda_k,\end{gathered}$$

其中 $1\leqslant i_1,\cdots,i_s\leqslant k$, 并且 $x_1,x_2,\cdots,x_{s+1}$ 两两互异. 再由 $\mathrm{tr}(A^k)=0\ (1\leqslant k\leqslant n)$ 可得

$$\begin{cases}\lambda_1+\lambda_2+\cdots+\lambda_k=0\\ \lambda_1^2+\lambda_2^2+\cdots+\lambda_k^2=0\\ \cdots\cdots\cdots\cdots\\ \lambda_1^n+\lambda_2^n+\cdots+\lambda_k^n=0\end{cases}\Rightarrow\begin{cases}c_1x_1+c_2x_2+\cdots+c_{s+1}x_{s+1}=0\\ c_1x_1^2+c_2x_2^2+\cdots+c_{s+1}x_{s+1}^2=0\\ \cdots\cdots\cdots\cdots\\ c_1x_1^n+c_2x_2^n+\cdots+c_{s+1}x_{s+1}^n=0\end{cases},\tag{3}$$

其中 $c_1=i_1,c_k=i_k-i_{k-1}\ (k=2,3,\cdots,s),c_{s+1}=k-i_s$. 显然 $(c_1,c_2,\cdots,c_{s+1})'$ 非零. 注意到方程组(3)的系数矩阵行列式为 Vandermonde 行列式非零, 故方程组(3)只有零解, 这与 $(c_1,c_2,\cdots,c_{s+1})'$ 非零矛盾!

□

命题 0.13

给定数域 $\mathbb{F}$, $A,B\in\mathbb{F}^{n\times n}$, A,B 特征值都在 $\mathbb{F}$ 中.

1. 若 $AB=0$, 证明存在可逆 $P\in\mathbb{F}^{n\times n}$ 使得 $P^{-1}AP,P^{-1}BP$ 都是上三角矩阵.
2. 若 $\mathrm{rank}(AB-BA)\leqslant 1$, 证明存在可逆 $P\in\mathbb{F}^{n\times n}$ 使得 $P^{-1}AP,P^{-1}BP$ 都是上三角矩阵.

3. 若存在 $a,b\in\mathbb{F}$ 使得 $AB-BA=aA+bB$, 证明存在可逆 $P\in\mathbb{F}^{n\times n}$ 使得 $P^{-1}AP,P^{-1}BP$ 都是上三角矩阵.

笔记 此类问题, 我们可以先找一个公共的非平凡的不变子空间 (或公共特征向量), 从而可以从这个不变子空间中取一组基, 再扩充为全空间的基. 于是原矩阵在这组基下的表示矩阵就可同时上三角化, 并且表示矩阵的 $(1,1)$ 元可以确定, 接下来只需考虑表示矩阵的其余的 $n-1$ 阶矩阵即可, 这样就对原矩阵进行了降阶. 最后再利用数学归纳法就能立刻得到证明.

证明

1. 若 A 可逆, 则 $B=0$. 此时因为 $AB=BA=0$. 运用可交换矩阵可同时上三角化即证, 当然, 由定理 0.1, 此时 A,B 有公共特征向量. 若 A 不可逆, 于是我们不难证明 $\mathrm{Ker}A$ 是 B 的不变子空间 (当然也是 A 的不变子空间), 所以存在 $B|_{\mathrm{Ker}A}$ 的特征向量 $\alpha\in\mathrm{Ker}A$. 此时 α 也是 A 的特征向量 (特征值为 0), 于是 A,B 总有公共特征向量.
   那么我们运用归纳法, 当 $n=1$ 显然, 假定命题对 $n-1$ 时成立, 当 n 时, 设 α 是 A,B 公共特征向量, 将其扩充为 $\mathbb{F}^n$ 一组基, 在这组基下, A,B 分别有表示矩阵
   $$\widetilde{A}=\begin{pmatrix}c_1&\beta^T\\0&A_{n-1}\end{pmatrix},\widetilde{B}=\begin{pmatrix}c_2&\gamma^T\\0&B_{n-1}\end{pmatrix},c_1,c_2\in\mathbb{F}.$$
   由 $AB=0$ 知 $A_{n-1}B_{n-1}=0$. 故可对 A_{n-1},B_{n-1} 用归纳假设, 即存在可逆 $P_{n-1}\in\mathbb{F}^{(n-1)\times(n-1)}$ 使得 $P_{n-1}^{-1}A_{n-1}P_{n-1}$, $P_{n-1}^{-1}B_{n-1}P_{n-1}$ 是上三角矩阵. 于是取 $P=\begin{pmatrix}1&0\\0&P_{n-1}\end{pmatrix}\in\mathbb{F}^{n\times n}$, 就有
   $$P^{-1}\widetilde{A}P=\begin{pmatrix}c_1&\beta^TP_{n-1}\\0&P_{n-1}^{-1}A_{n-1}P_{n-1}\end{pmatrix},P^{-1}\widetilde{B}P=\begin{pmatrix}c_2&\gamma^TP_{n-1}\\0&P_{n-1}^{-1}B_{n-1}P_{n-1}\end{pmatrix}$$
   都是上三角矩阵. 这就证明了存在可逆 $P\in\mathbb{F}^{n\times n}$ 使得 $P^{-1}AP,P^{-1}BP$ 都是上三角矩阵.
2. 若 $\mathrm{rank}(AB-BA)=0$, 这是可交换矩阵可同时上三角化的直接结果. 若 $\mathrm{rank}(AB-BA)=1$, 我们假定 A 是不可逆的, 否则用 $A-\lambda I,\lambda\in\mathbb{F}$(这里 λ 是此时 A 的特征值) 代替 A. 下证 $\mathrm{Ker}A,\mathrm{Im}A$ 必有一个是 B 的不变子空间.
   Step 1 若 $\mathrm{Ker}A\subset\mathrm{Ker}(AB-BA)$, 则设 $x\in\mathrm{Ker}A$, 我们有 $ABx=BAx=0$, 这就证明了 $\mathrm{Ker}A$ 是 B 不变子空间.
   Step 2 若 $\mathrm{Ker}A\not\subset\mathrm{Ker}(AB-BA)$ 不成立, 取 $x\in\mathrm{Ker}A$ 使得 $y=(AB-BA)x\neq 0$. 注意到 $\mathrm{rank}(AB-BA)=1$, 故 $\mathrm{Im}(AB-BA)=\mathrm{span}\{y\}$. 但是
   $$y=(AB-BA)x=ABx-BAx=ABx\in\mathrm{Im}A,$$
   这就说明 $\mathrm{Im}(AB-BA)\subset\mathrm{Im}A$. 于是对 $u=Av\in\mathrm{Im}A$, 我们有
   $$Bu=BAv=ABv-(AB-BA)v\in\mathrm{Im}A,$$
   从而确实了 $\mathrm{Im}A$ 是 B 不变子空间.
   注意到 A,B 有一个公共非平凡不变子空间. 所以取这个不变子空间一组基扩充为全空间的基, 则 A,B 有表示矩阵
   $$\widetilde{A}=\begin{pmatrix}A_1&A_3\\0&A_2\end{pmatrix},\widetilde{B}=\begin{pmatrix}B_1&B_3\\0&B_2\end{pmatrix}.$$
   注意到
   $$r(AB-BA)=r\begin{pmatrix}A_1B_1-B_1A_1&*\\0&A_2B_2-B_2A_2\end{pmatrix}$$
   $$\overset{\text{矩阵秩的基本公式 (3)}}{\geqslant} r\begin{pmatrix}A_1B_1-B_1A_1&0\\0&A_2B_2-B_2A_2\end{pmatrix}$$

$$= r(A_1B_1 - B_1A_1) + r(A_2B_2 - B_2A_2),$$

我们有

$$r(A_1B_1 - B_1A_1) \leqslant 1, r(A_2B_2 - B_2A_2) \leqslant 1,$$

现在阶降下去了, 类似我们前面的归纳证明我们知道 A, B 可同时相似上三角化.

3. 如果 $a = b = 0$, 那么这是可交换矩阵可同时上三角化的直接结果. 不失一般性, 不妨设 $a \neq 0$. 我们可以用 $\frac{1}{a}B$ 代替 B 来不妨设 $a = 1$. 那么记 $C = AB - BA = A + bB$, 则成立等式 $CB - BC = C$. 显然只需证明 C, B 可同时相似上三角化即可, 因为此时 A 也就被相似上角化. 注意到

$$\operatorname{tr}(CB - BC) = \operatorname{tr}(C) = 0, \operatorname{tr}(CCB - CBC) = \operatorname{tr}(C^2) = 0,$$

$$\operatorname{tr}(CCCB - CCBC) = \operatorname{tr}(C^3) = 0, \operatorname{tr}(CCCCB - CCCBC) = \operatorname{tr}(C^4) = 0,$$

$$\cdots\cdots\cdots\cdots$$

于是我们证明了 $\operatorname{tr}(C^k) = 0, \forall k \in \mathbb{N}$. 由定理 0.3, 我们知道 C 是幂零矩阵.

假设 B, C 没有公共特征向量且 $Bx = \lambda x, x \neq 0, \lambda \in \mathbb{F}$, 则 $Cx \neq 0$ 且

$$CBx - BCx = Cx \Rightarrow (B - (\lambda - 1)I)Cx = 0,$$

因此 B 有特征值 $\lambda - 1$. 这样 B 也应该有特征值 $\lambda - 2$, 如此下去 B 会有无穷多个特征值, 这不可能! 所以 B, C 应该有公共特征向量 α. 将 α 扩充为全空间的一组基, B, C 在这组基下有表示矩阵

$$\widetilde{B} = \begin{pmatrix} \lambda & * \\ 0 & B_1 \end{pmatrix}, \widetilde{C} = \begin{pmatrix} 0 & * \\ 0 & C_1 \end{pmatrix}, \lambda \in \mathbb{F}, C_1 \in \mathbb{F}^{(n-1)\times(n-1)},$$

则 $C_1B_1 - B_1C_1 = C_1$. 这样阶数就降下去了. 所以可以类似前面归纳完成证明.

□

引理 0.3 (Hilbert 矩阵逆矩阵元素和)

设

$$H = \begin{pmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{pmatrix}$$

, 这里 H 是 n 阶矩阵. 证明 H^{-1} 所有元素和为 n^2.

♡

笔记 这里涉及经典技巧, 即分解 $J = AH + HB$, 需要积累.

证明 取 n 阶矩阵

$$J = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}, A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n \end{pmatrix}, B = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n-1 \end{pmatrix}$$

, 则 $AH + HB = J = \mathbf{1}_n\mathbf{1}_n^T$. 于是

$$\mathbf{1}_n^T H^{-1}\mathbf{1}_n = \operatorname{tr}\left(\mathbf{1}_n^T H^{-1}\mathbf{1}_n\right) = \operatorname{tr}\left(H^{-1}\mathbf{1}_n\mathbf{1}_n^T\right) = \operatorname{tr}\left(H^{-1}J\right) = \operatorname{tr}\left(H^{-1}(AH + HB)\right)$$

$$= \operatorname{tr}\left(H^{-1}AH + B\right) = \operatorname{tr}(A) + \operatorname{tr}(B) = \frac{n(n+1)}{2} + \frac{n(n-1)}{2} = n^2.$$

□

例题 0.1 设 A, B 都是 n 阶矩阵且 $AB = BA$. 若 A 是幂零矩阵, 求证: $|A + B| = |B|$.

证明 证法一: 由命题 0.7可知, A, B 可同时上三角化, 即存在可逆矩阵 P, 使得 $P^{-1}AP$ 和 $P^{-1}BP$ 都是上三角矩阵. 因为上三角矩阵的主对角元是矩阵的特征值, 而幂零矩阵的特征值全为零, 所以 $|P^{-1}AP + P^{-1}BP| = |P^{-1}BP|$, 即有 $|A + B| = |B|$.

证法二: 先假设 B 是可逆矩阵, 则 $|A + B| = |I_n + AB^{-1}||B|$, 只要证明 $|I_n + AB^{-1}| = 1$ 即可. 由 $AB = BA$ 可知 $AB^{-1} = B^{-1}A$, 再由 A 是幂零矩阵容易验证 AB^{-1} 也是幂零矩阵, 从而其特征值全为零. 因此 $I_n + AB^{-1}$ 的特征值全为 1, 故 $|I_n + AB^{-1}| = 1$.

对于一般的矩阵 B, 可取到一列有理数 $t_k \to 0$, 使得 t_kI_n+B 是可逆矩阵. 由可逆情形的证明可得 $|A+t_kI_n+B| = |t_kI_n + B|$. 注意到上式两边都是 t_k 的多项式, 从而关于 t_k 连续. 将上式两边同时取极限, 令 $t_k \to 0$, 即得结论.

□