

第三次习题课

3.1 第四周作业选讲

见第四周作业答案。

3.2 第五周作业选讲

习题 3.1: P186 Ex 4.6.17

设 $P_1, P_2, \dots, P_k, Q_1, Q_2, \dots, Q_k$ 都是 n 阶方阵, 并且 $P_i Q_j = Q_j P_i, \text{rank } P_i = \text{rank } P_i Q_i, 1 \leq i, j \leq k$.

证明: $\text{rank } P_1 P_2 \cdots P_k = \text{rank } P_1 \cdots P_k Q_1 \cdots Q_k$.

证明

思路: 题目给出的秩相等条件实际上可以转化为比较相应的核空间或像空间的问题, 这比直接比较矩阵乘积的秩更易利用。而各 P_i 与 Q_j 可交换, 自然考虑用归纳法逐步增加矩阵因子。以下为解答过程:

用归纳法证明。当 $k=1$, 这就是给定的条件。假设结论对 $k-1$ 成立, 记 $S = P_1 \cdots P_{k-1}, T = Q_1 \cdots Q_{k-1}$, 则 $\text{rank } S = \text{rank}(ST)$. 由可交换性有 $ST = TS$. 将矩阵视作 $\mathbb{F}^n$ 上的线性变换, 由 $\text{Ker } S \subseteq \text{Ker}(TS)$ 及两者维数相等, 可得 $\text{Ker } S = \text{Ker}(TS)$. 类似地, 由 $\text{Im}(P_k Q_k) \subseteq \text{Im } P_k$ 及 $\text{rank}(P_k Q_k) = \text{rank } P_k$, 可得 $\text{Im}(P_k Q_k) = \text{Im } P_k$. 对任意两个可相乘的矩阵 A, B 有 $\dim \text{Im } B = \dim \text{Im}(A|_{\text{Im } B}) + \dim \text{Ker}(A|_{\text{Im } B}) = \dim \text{Im}(AB) + \dim(\text{Ker } A \cap \text{Im } B)$. 于是有 $\text{rank}(AB) = \text{rank } B - \dim(\text{Ker } A \cap \text{Im } B)$, 从而可以得到

$$\begin{aligned} \text{rank}(SP_k T Q_k) &= \text{rank}(TSP_k Q_k) = \text{rank}(P_k Q_k) - \dim(\text{Ker}(TS) \cap \text{Im}(P_k Q_k)) \\ &= \text{rank } P_k - \dim(\text{Ker } S \cap \text{Im } P_k) = \text{rank}(SP_k). \end{aligned}$$

这正好就是我们所要证明的结论, 这样我们就完成了证明。□

习题 3.2: P194 Ex 4.8.3

设 W 是数域 $\mathbb{F}$ 上 n 维列向量空间 $\mathbb{F}^n$ 的子空间。证明: 存在数域 $\mathbb{F}$ 上 $m \times n$ 矩阵 A , 使得齐次方程组 $Ax = 0$ 的解空间 V_A 为 W .

证明

思路: 题目实际上要求构造一个以 W 为核空间的线性映射。由于线性映射可以通过指定一组基的像来确定, 自然想到先将 W 的一组基扩充为整个空间的基, 再通过基变换来构造所需的 A 。以下为解答过程:

若 $W = \{0\}$, 任取可逆矩阵 $A \in \mathbb{F}^{n \times n}$ 即可。若 $W = \mathbb{F}^n$, 取 A 为 $1 \times n$ 零矩阵即可。

若 $0 < \dim W = r < n$, 设 $\{\alpha_1, \dots, \alpha_r\}$ 是 W 的基。记 $n \times r$ 矩阵 $(\alpha_1, \alpha_2, \dots, \alpha_r)$ 为 B , 则 B 是列满秩的。

存在 n 阶可逆矩阵 P 使得 $B = P \begin{pmatrix} I_r \\ 0 \end{pmatrix}$ 。(可取 $P = (\alpha_1, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_n)$, 其中 $\{\alpha_{r+1}, \dots, \alpha_n\}$ 是 W 的补空间的一组基) 定义 $A = (0, I_{n-r})P^{-1} \in \mathbb{F}^{(n-r) \times n}$, 则有

$$AB = (0, I_{n-r})P^{-1}P \begin{pmatrix} I_r \\ 0 \end{pmatrix} = 0.$$

从而 $A\alpha_j = 0, j = 1, \dots, r$. 于是 $W \subseteq V_A$. 而 $\dim V_A = n - \text{rank } A = r = \dim W$, 故 $W = V_A$. $\square$

习题 3.3: P205 Ex 5.2.12

设 $\mathcal{A}: \mathbb{F}^{n \times n} \rightarrow \mathbb{F}$ 是线性映射, 并且对任意 $A, B \in \mathbb{F}^{n \times n}$, $\mathcal{A}(AB) = \mathcal{A}(BA)$. 证明: $\mathcal{A} = \lambda \text{tr}$, 其中 $\lambda \in \mathbb{F}$.

证明

思路: 由于线性映射完全由一组基上的取值决定, 自然想到利用矩阵空间的标准基 $\{E_{ij}\}$ 来研究 $\mathcal{A}$. 题目给出的条件涉及矩阵乘法, 而矩阵单位之间的乘法关系十分简单, 因此可以通过适当选取 E_{ij} , 利用条件证明 $\mathcal{A}$ 在非对角矩阵单位上的取值为零, 在所有对角矩阵单位上的取值相同。以下为解答过程:

容易验证 $E_{ij}E_{kl} = E_{il}\delta_{jk}, 1 \leq i, j, k, l \leq n$, 其中 δ_{jk} 为 Kronecker 符号, 由条件可得

$$\mathcal{A}(E_{ij}E_{kl}) = \mathcal{A}(E_{il}\delta_{jk}) = \mathcal{A}(E_{kl}E_{ij}) = \mathcal{A}(E_{kj}\delta_{il}).$$

在上式中令 $k = j = 1$, 可得对任意 $i \neq l$, 有 $\mathcal{A}(E_{il}) = \mathcal{A}(0) = 0$. 再令 $i = l \neq j = k$, 可得 $\mathcal{A}(E_{ii}) = \mathcal{A}(E_{jj}) = \lambda$, 其中 $\lambda = \mathcal{A}(E_{11})$. 故对任意 $A = (a_{ij})_{n \times n} = \sum_{i=1}^n \sum_{j=1}^n a_{ij}E_{ij} \in \mathbb{F}^{n \times n}$, 由 $\mathcal{A}$ 线性可得

$$\mathcal{A}(A) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} \mathcal{A}(E_{ij}) = \sum_{i=1}^n a_{ii} \mathcal{A}(E_{ii}) = \lambda \sum_{i=1}^n a_{ii} = \lambda \text{tr } A.$$

因此 $\mathcal{A} = \lambda \text{tr}$, 其中 $\lambda = \mathcal{A}(E_{11}) \in \mathbb{F}$. 这样我们就完成了证明. $\square$

习题 3.4: P211 Ex 5.3.4

设 U 与 V 分别是数域 $\mathbb{F}$ 上 m 维与 n 维线性空间。取定 $\alpha \in U$. 将所有满足 $\mathcal{A}(\alpha) = 0$ 的线性映射 $\mathcal{A}: U \rightarrow V$ 的集合记为 K . 证明: K 在线性映射的加法以及纯量与线性映射的乘法下成为数域 $\mathbb{F}$ 上线性空间。求 $\dim K$.

证明

思路: 所有 $U \rightarrow V$ 线性映射构成一个 mn 维线性空间, 故只需考虑 $\mathcal{A}(\alpha) = 0$ 对其维数的限制。当 $\alpha \neq 0$ 时, 自然想到将 α 扩充为 U 的一组基。由于线性映射可以通过任意指定一组基的像来唯一确定, 而 α 的像已经固定为零, 其余 $m-1$ 个基向量的像仍可任意选取, 由此便能确定 $\dim K$. 以下为解答过程:

对 $\mathcal{A}, \mathcal{B} \in K, \lambda, \mu \in \mathbb{F}$, 有 $(\lambda\mathcal{A} + \mu\mathcal{B})(\alpha) = \lambda\mathcal{A}(\alpha) + \mu\mathcal{B}(\alpha) = 0$, 因此 $\lambda\mathcal{A} + \mu\mathcal{B} \in K$. 且显然零映射也属于 K , 故 K 是所有 $U \rightarrow V$ 线性映射构成的线性空间的子空间, 所以它是一个线性空间。

当 $\alpha = 0$ 时, K 为所有 $U \rightarrow V$ 线性映射构成的集合, 故此时 $\dim K = mn$. 当 $\alpha \neq 0$ 时, 记 α 为 α_1 , 并将其扩充为 U 的一组基 $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$. 取 V 的一组基 $\{\beta_1, \beta_2, \dots, \beta_n\}$. 对任意 $\xi = \sum_{s=1}^m b_s \alpha_s \in U$ 定义 $\mathcal{E}_{ij}(\xi) = b_i \beta_j, 1 \leq i \leq m, 1 \leq j \leq n$. 这些映射构成所有 $U \rightarrow V$ 线性映射的一组基。

令 $\mathcal{B} = \{\mathcal{E}_{ij} : 2 \leq i \leq m, 1 \leq j \leq n\}$, 下面证明 $\mathcal{B}$ 是 K 的一组基. 由 $\mathcal{E}_{ij}(\alpha_1) = 0, i \neq 1$ 可知 $\mathcal{B} \subseteq K$, 并且 $\mathcal{B}$ 是线性无关的. 对任意 $\mathcal{A} \in K$, 存在 $c_{ij} \in \mathbb{F}$ 使得 $\mathcal{A} = \sum_{i=1}^m \sum_{j=1}^n c_{ij} \mathcal{E}_{ij}$. 于是由 $\mathcal{A}(\alpha_1) = 0$ 可得

$$0 = \mathcal{A}(\alpha_1) = \sum_{i=1}^m \sum_{j=1}^n c_{ij} \mathcal{E}_{ij}(\alpha_1) = \sum_{j=1}^n c_{1j} \beta_j.$$

由于 $\beta_1, \beta_2, \dots, \beta_n$ 线性无关, 故 $c_{11} = c_{12} = \dots = c_{1n} = 0$, 从而 $\mathcal{A} = \sum_{i=2}^m \sum_{j=1}^n c_{ij} \mathcal{E}_{ij}$ 可由 $\mathcal{B}$ 线性表示. 于是 $\mathcal{B}$ 是 K 的一组基, 故此时 $\dim K = (m-1)n$. $\square$

3.3 补充内容

3.3.1 空间的分解相关补充知识

幂等线性映射与直和分解 我们知道一个幂等线性映射 $P: V \rightarrow V$ 可以给出一个直和分解 $V = \text{Im } P \oplus \text{Ker } P$. 那反过来是否任意一个直和分解都可以自然地给出一个幂等线性映射呢?

命题 3.1: 直和分解与投影映射

设 $V = U \oplus W$, 则存在唯一的幂等线性映射 $P: V \rightarrow V$, 使得 $\text{Im } P = U, \text{Ker } P = W$. 这样的线性映射称为沿 W 到 U 的投影映射。

证明

由于 $V = U \oplus W$, 对任意 $v \in V$, 存在唯一的 $u \in U, w \in W$ 使得 $v = u + w$. 定义 $P(v) = u$, 不难验证 P 是线性映射, 且 $\text{Im } P = U, \text{Ker } P = W$. 又因为 $P^2(v) = P(u) = u = P(v)$, 故 $P^2 = P$.

下面证明唯一性. 设 Q 也是满足条件的幂等线性映射, 则对任意 $u \in U = \text{Im } Q$, 有 $Qu = u$, 而对任意 $w \in W = \text{Ker } Q$, 有 $Qw = 0$. 因此 $Q(u + w) = u = P(u + w)$, 故 $Q = P$. $\square$

注记

上述结论表明, 线性空间的直和分解与投影映射一一对应. 需要注意的是 $V = \text{Im } P \oplus \text{Ker } P$ 本身并不能保证 $P^2 = P$, 还需要 P 在其像空间上是恒等映射。

此外, 若 $\dim \text{Im } P = r$, 则可以分别取 $\text{Im } P$ 与 $\text{Ker } P$ 的一组基, 合并得到 V 的一组基. 在这组基下, P 的矩阵为 $\text{diag}(I_r, 0)$. 因此任意幂等矩阵都相似于这样的分块对角矩阵。

推论 3.1: 对合映射的直和分解

设 V 是特征不为 2 的数域 $\mathbb{F}$ 上的有限维线性空间, 线性映射 $T: V \rightarrow V$ 满足 $T^2 = I$, 则 $V = \text{Ker}(T - I) \oplus \text{Ker}(T + I)$.

证明

令 $P = (I + T)/2$, 由于 $T^2 = I$, 可得 $P^2 = (I + 2T + T^2)/4 = (I + T)/2 = P$. 故 P 是幂等映射. 由前面的定理可知 $V = \text{Im } P \oplus \text{Ker } P$. 容易验证 $\text{Ker } P = \text{Ker}(T + I)$. 另一方面, 由 $(T - I)(T + I) = 0$ 可知

$\text{Im } P \subseteq \text{Ker}(T - I)$. 而对任意 $v \in \text{Ker}(T - I)$, 有 $Tv = v$, 故 $Pv = (v + Tv)/2 = v$, 从而 $v \in \text{Im } P$. 因此 $\text{Im } P = \text{Ker}(T - I)$, 代入直和分解即得结论。□

Fitting 分解 在第五周作业 Ex 5.4.3 中, 我们证明了当 $\text{rank}(T^2) = \text{rank}(T)$ 时, 有 $\text{Im } T \cap \text{Ker } T = \{0\}$. 事实上, 对于任意有限维线性空间上的线性映射都有类似的结论, 这就是 Fitting 分解。

首先回顾第四周作业中的一个结论, 证明过程此处略去。

引理 3.1: 核空间与像空间的稳定性

设 V 是 n 维线性空间, $T: V \rightarrow V$ 是线性映射, 则

$$\text{Ker } T^n = \text{Ker } T^{n+1} = \text{Ker } T^{n+2} = \cdots, \text{Im } T^n = \text{Im } T^{n+1} = \text{Im } T^{n+2} = \cdots.$$

利用这一结论, 我们可以得到下面的 Fitting 分解定理。

定理 3.1: Fitting 分解

设 V 是 n 维线性空间, 而 $T: V \rightarrow V$ 是线性映射, 则 $V = \text{Ker } T^n \oplus \text{Im } T^n$. 并且直和项是两个不变子空间, 从而可以将 T 在其上的限制视为其到自身的线性变换, 并且此时 T 在 $\text{Ker } T^n$ 上的限制是幂零线性映射, 在 $\text{Im } T^n$ 上的限制是可逆线性映射。

证明

任取 $v \in \text{Ker } T^n \cap \text{Im } T^n$, 由 $v \in \text{Im } T^n$, 存在 $u \in V$ 使得 $v = T^n u$. 又由 $v \in \text{Ker } T^n$, 故 $T^{2n} u = T^n v = 0$. 由引理知 $\text{Ker } T^{2n} = \text{Ker } T^n$, 故 $T^n u = 0$, 即 $v = 0$. 因此交空间为 $\{0\}$, 从而 $V = \text{Ker } T^n \oplus \text{Im } T^n$.

下面考虑 T 在这两个子空间上的限制。显然 $T(\text{Ker } T^n) \subseteq \text{Ker } T^n$, 且 T^n 在 $\text{Ker } T^n$ 上恒为零, 故其限制是幂零线性映射。另一方面 $T(\text{Im } T^n) = \text{Im } T^{n+1} = \text{Im } T^n$. 故 T 在 $\text{Im } T^n$ 上的限制是满射。由于该空间有限维, 故这一限制也是单射, 从而可逆。□

注记

Fitting 分解说明, 任意有限维线性空间上的线性映射, 都可以分解为一个幂零部分与一个可逆部分。所以在适当选取的基下, T 的矩阵可以写为 $\text{diag}(N, A)$, 其中 N 是幂零矩阵, A 是可逆矩阵。并且这一结论不依赖于特征值的存在性, 因此在任意数域上都成立。

3.3.2 交换子与迹为零的矩阵

在第五周作业 Ex 5.2.12 中, 我们证明了若线性映射 $\mathcal{A}: \mathbb{F}^{n \times n} \rightarrow \mathbb{F}$ 满足 $\mathcal{A}(AB) = \mathcal{A}(BA)$, 则 $\mathcal{A}$ 必然是迹映射的数量倍。下面通过矩阵的交换子, 给出这一结论背后更深刻的原因。

首先我们定义矩阵的交换子:

定义 3.1: 矩阵的交换子

设 $A, B \in \mathbb{F}^{n \times n}$, 定义 A 与 B 的交换子为 $[A, B] = AB - BA$.

显然对任意 $A, B \in \mathbb{F}^{n \times n}$, 有 $\text{tr}[A, B] = 0$. 这说明任意交换子都是迹为零的矩阵。于是一个自然的问题是, 反过来是否所有迹为零的矩阵都可以由交换子线性表示? 幸运的是, 答案是肯定的, 也就是下面这个定理, 并且事实上我们还能得到一些更强的结论。

定理 3.2: 迹为零的矩阵与交换子空间

记 $\mathfrak{sl}_n(\mathbb{F}) = \{A \in \mathbb{F}^{n \times n} : \operatorname{tr} A = 0\}$, 则 $\mathfrak{sl}_n(\mathbb{F}) = \operatorname{span}_{\mathbb{F}}\{[A, B] : A, B \in \mathbb{F}^{n \times n}\}$. 即所有交换子张成的线性空间恰好是全体迹为零的矩阵构成的线性空间。

证明

记 $W = \operatorname{span}_{\mathbb{F}}\{[A, B] : A, B \in \mathbb{F}^{n \times n}\}$. 首先显然 $W \subseteq \mathfrak{sl}_n(\mathbb{F})$. 反过来, 我们知道 $E_{ij}E_{kl} = \delta_{jk}E_{il}$. 于是可知当 $i \neq j$ 时, 有 $[E_{ii}, E_{ij}] = E_{ii}E_{ij} - E_{ij}E_{ii} = E_{ij}$. 因此所有非对角矩阵单位 E_{ij} 都属于 W .

另一方面, 对任意 $i \neq j$, 有 $[E_{ij}, E_{ji}] = E_{ij}E_{ji} - E_{ji}E_{ij} = E_{ii} - E_{jj}$. 因此所有形如 $E_{ii} - E_{jj}$ 的矩阵也属于 W . 对任意 $A = (a_{ij}) \in \mathfrak{sl}_n(\mathbb{F})$, 由于 $\sum_{i=1}^n a_{ii} = 0$, 故

$$A = \sum_{i \neq j} a_{ij} E_{ij} + \sum_{i=2}^n a_{ii} (E_{ii} - E_{11}).$$

由前面的结论可知右端各项均属于 W , 因此 $A \in W$. 故 $\mathfrak{sl}_n(\mathbb{F}) \subseteq W$, 从而得到所要证明的等式。

实际上我们有一个更强的结论: 当 $\operatorname{char} \mathbb{F} = 0$ 或 $> n$ 时, 所有迹为零的矩阵都可以表示为单个交换子。

首先证明任意迹为零的矩阵都相似于对角元全为零的矩阵。对矩阵阶数 n 进行归纳, 当 $n = 1$ 时结论显然成立。假设结论对 $(n-1)$ 阶矩阵成立, 考虑 n 阶矩阵 A , 满足 $\operatorname{tr} A = 0$. 若 $A = \lambda I$, 则 $\operatorname{tr} A = n\lambda = 0$, 由 $\operatorname{char} \mathbb{F} = 0$ 或 $> n$ 可知 $\lambda = 0$, 结论显然成立。而若 A 不是数量矩阵, 则存在 $v \in \mathbb{F}^n$, 使得 v, Av 线性无关。将其扩充为一组基, 则 A 在这组基下的矩阵具有如下形式:

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

其中 $\operatorname{tr} A_1 = 0$. 由归纳假设, A_1 相似于对角元全为零的矩阵, 因此 A 也相似于对角元全为零的矩阵。即存在可逆矩阵 P , 使得 $M = P^{-1}AP = (m_{ij})$ 的对角元全为零。令 $D = \operatorname{diag}(1, 2, \dots, n)$, 定义矩阵 $E = (e_{ij})$, 其中当 $i \neq j$ 时令 $e_{ij} = m_{ij}/(i-j)$, 当 $i = j$ 时令 $e_{ii} = 0$. 由 $\operatorname{char} \mathbb{F} = 0$ 或 $> n$ 可知 E 是良定义的。且注意到 $(DE - ED)_{ij} = (i-j)e_{ij} = m_{ij}$. 因此 $M = DE - ED$, 从而 $A = PMP^{-1} = P(DE - ED)P^{-1} = (PDP^{-1})(PEP^{-1}) - (PEP^{-1})(PDP^{-1})$. 即 A 可以表示为两个矩阵的交换子。□

注记

事实上结论还可以进一步加强到对任意域都成立, 即 Albert-Muckenhoupt 定理。由于上述证明利用了特征为 0 或 $> n$ 时 $1, 2, \dots, n$ 两两不同的性质, 因此不能直接推广到任意域。对任意域情形的证明过程更为复杂, 感兴趣的同学可以自行学习了解, 此处不再给出。

由上述定理, 我们可以给出第五周作业 Ex 5.2.12 的一个新证明。

推论 3.2: 迹映射的刻画

设 $\mathcal{A} : \mathbb{F}^{n \times n} \rightarrow \mathbb{F}$ 是线性映射, 则 $\mathcal{A}(AB) = \mathcal{A}(BA)$ 对任意 A, B 成立的充分必要条件是存在 $\lambda \in \mathbb{F}$, 使得 $\mathcal{A} = \lambda \operatorname{tr}$.

证明

充分性显然, 下面证明必要性。由 $\mathcal{A}(AB) = \mathcal{A}(BA)$ 可得 $\mathcal{A}([A, B]) = 0$. 由前面的定理知 $\mathcal{A}$ 在 $\mathfrak{sl}_n(\mathbb{F})$ 上恒为零。取 $\lambda = \mathcal{A}(E_{11})$. 对任意 $A \in \mathbb{F}^{n \times n}$, 注意到 $\operatorname{tr}(A - \operatorname{tr}(A)E_{11}) = 0$, 故 $A - \operatorname{tr}(A)E_{11} \in \mathfrak{sl}_n(\mathbb{F})$. 于是由 $\mathcal{A}$ 的线性性可得 $\mathcal{A}(A) = \mathcal{A}(A - \operatorname{tr}(A)E_{11}) + \operatorname{tr}(A)\mathcal{A}(E_{11}) = \lambda \operatorname{tr}(A)$. □

注记

迹映射 $\text{tr} : \mathbb{F}^{n \times n} \rightarrow \mathbb{F}$ 是非零线性映射，其核空间就是 $\mathfrak{sl}_n(\mathbb{F})$. 于是 $\dim \mathfrak{sl}_n(\mathbb{F}) = n^2 - 1$. 因此迹为零的矩阵构成矩阵空间中的一个余维数为1的子空间。

3.3.3 复结构与复线性空间

第五周作业 Ex 5.2.10 讨论了复数与一类特殊的二阶实矩阵之间的对应关系。事实上，这种表示并非偶然，而是揭示了复线性空间与实线性空间之间的一个重要联系。

首先考虑复数单位 i 最基本的性质 $i^2 = -1$. 如果希望在一个实线性空间上实现类似于乘以 i 的运算，那么就需要寻找一个满足相同代数关系的实线性映射。

定义 3.2: 复结构

设 V 是实线性空间，若存在实线性映射 $J : V \rightarrow V$ 满足 $J^2 = -I$ ，则称 J 为 V 上的一个复结构。

下面讨论什么样的实线性空间可以具有复结构。

定理 3.3: 复结构存在的充要条件

设 V 是有限维实线性空间，则 V 上存在复结构的充分必要条件是 $\dim_{\mathbb{R}} V$ 为偶数。

证明

必要性: 若 V 上存在复结构 J ，设 $\dim_{\mathbb{R}} V = n$. 取 V 的一组基，将 J 表示为 n 阶实矩阵，仍记为 J . 由 $J^2 = -I_n$ 可得 $\det(J)^2 = (-1)^n$. 由于 $\det(J)$ 为实数可得 n 必为偶数。

充分性: 设 $\dim_{\mathbb{R}} V = 2m$ ，取 V 的一组基 $\{e_1, e_2, \dots, e_{2m}\}$. 定义线性映射 $J : V \rightarrow V$ ，使得

$$J(e_{2k-1}) = -e_{2k}, J(e_{2k}) = e_{2k-1}, 1 \leq k \leq m.$$

容易验证 $J^2(e_i) = -e_i$ 对每个基向量均成立，因此 $J^2 = -I$. 故 J 是 V 上的一个复结构。□

那么一个很自然的问题是：为什么满足 $J^2 = -I$ 的线性映射被称为复结构呢？这是因为它可以让一个实线性空间自然地成为复线性空间。

命题 3.2: 由复结构构造复线性空间

设 V 是实线性空间，而 J 是 V 上的复结构。保持 V 原有的向量加法，并对任意 $a + bi \in \mathbb{C}$ 及 $v \in V$ 定义 $(a + bi)v = av + bJ(v)$. 则 V 成为一个复线性空间，且 i 的数量乘法恰好对应线性映射 J .

证明

只需验证新定义的数量乘法满足复线性空间的公理。由于 J 是实线性映射，数量乘法对向量加法及复数加法的分配律均容易验证，并且 $1v = v$. 下面验证复数乘法的结合律。对任意 $a, b, c, d \in \mathbb{R}$ 及 $v \in V$ 有

$$\begin{aligned} (a + bi)((c + di)v) &= (a + bi)(cv + dJv) = acv + adJv + bcJv + bdJ^2v \\ &= (ac - bd)v + (ad + bc)Jv = ((a + bi)(c + di))v. \end{aligned}$$

其中第三个等式利用了 $J^2 = -I$. 故 V 成为复线性空间，且由定义有 $iv = J(v)$. □

最后作为一个简单的例子，我们再回到第五周作业 Ex 5.2.10 中的复数矩阵表示，现在我们可以从复结构的角度的来理解它。

例 3.1: 复数的实矩阵表示

考虑 $\mathbb{R}^2$ 上的线性映射 J , 其在标准基下的矩阵为

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

显然 $J^2 = -I_2$, 因此 J 给出了 $\mathbb{R}^2$ 上的一个复结构. 对任意 $z = a + bi \in \mathbb{C}$, 定义

$$\Phi(z) = aI_2 + bJ = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

由 $J^2 = -I_2$, 故对任意 $z_1 = a + bi, z_2 = c + di$ 有 $\Phi(z_1)\Phi(z_2) = (aI_2 + bJ)(cI_2 + dJ) = (ac - bd)I_2 + (ad + bc)J = \Phi(z_1 z_2)$. 同时容易验证 $\Phi(z_1 + z_2) = \Phi(z_1) + \Phi(z_2)$, 且 $\Phi(1) = I_2$. 因此复数的加法与乘法, 都可以通过这类特殊实矩阵的加法与乘法来实现.

注记

这一构造可以推广到任意具有复结构 J 的非零实线性空间 V .

考虑映射 $\Phi: \mathbb{C} \rightarrow \text{End}_{\mathbb{R}}(V)$, 定义 $\Phi(a + bi) = aI + bJ$. 考虑 $\Phi(z)\Phi(\bar{z})$ 可得 Φ 是单射, 并且保持加法与乘法, 因此 $\mathbb{C}$ 可以表示为实线性映射空间中的一个子代数. 从这个角度来看, 复结构的本质就是在实线性空间上选取一个能够扮演复数单位 i 的线性映射.

3.3.4 Sylvester 算子

在本节的讨论中 $\mathbb{F} = \mathbb{C}$. 在之前的学习中, 我们知道相似变换不仅可以让方阵上三角化, 更可以让方阵相似准对角化, 每个对角上的方块是上三角矩阵且具有相同的特征值 (这比 Jordan 标准型要弱). 即

$$A \sim \text{diag}(J(\lambda_1), \dots, J(\lambda_n)).$$

其中每个对角块 $J(\lambda_i)$ 形如 $\begin{pmatrix} \lambda_i & & \\ & \lambda_i & * \\ & & \ddots \\ & & & \lambda_i \end{pmatrix}$. 也就是下面这个命题, 相信不少同学都有所了解.

命题 3.3

我们设 $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix} \in M_p(\mathbb{C}), B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ & b_{22} & \cdots & b_{2q} \\ & & \ddots & \vdots \\ & & & b_{qq} \end{pmatrix} \in M_q(\mathbb{C})$, 且对任意 i, j ,

$a_{ii} \neq b_{jj}$. 则对任意的 $X \in \mathbb{C}^{p \times q}$, 均存在 $C \in \mathbb{C}^{p \times q}$ 使得 $AX - XB = C$. 也就是说线性映射

$$\varphi: \mathbb{C}^{p \times q} \rightarrow \mathbb{C}^{p \times q}, \varphi(X) = AX - XB$$

是线性空间的同构.

这个命题告诉我们, 对于任意的矩阵 $X \in \mathbb{C}^{p \times q}$, 均存在 $X \in \mathbb{C}^{p \times q}$ 使得 $AX - XB = C$, 进而就有 $\begin{pmatrix} A & X \\ 0 & B \end{pmatrix} = \begin{pmatrix} I_p & -C \\ 0 & I_q \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} I_p & C \\ 0 & I_q \end{pmatrix}$, 即分块方阵 $\begin{pmatrix} A & X \\ 0 & B \end{pmatrix}$ 与 $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ 相似. 所以上三角方阵可以被进一步准对角化. 这个命题有如下强化的版本:

定理 3.4: Sylvester 定理

设 $A \in M_n(\mathbb{C})$, $B \in M_m(\mathbb{C})$, 则以下条件等价:

- (1) 映射 $\varphi: M_{n \times m}(\mathbb{C}) \rightarrow M_{n \times m}(\mathbb{C})$, $\varphi(X) = AX - XB$ 是可逆的。其中映射 φ 称为 Sylvester 算子。
- (2) A 与 B 没有公共特征值。

在证明这个定理之前, 我们先给出一些关于 Kronecker 积的介绍, 方便我们后续给出 Sylvester 算子的矩阵形式。

Kronecker 积及性质 首先给出 Kronecker 积的定义。

定义 3.3: Kronecker 积

设 $A = (a_{ij}) \in M_{m \times n}(\mathbb{F})$, $B \in M_{p \times q}(\mathbb{F})$, 定义

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix} \in M_{mp \times nq}(\mathbb{F}).$$

称其为 A 与 B 的 Kronecker 积 (Kronecker product)。

命题 3.4

Kronecker 积具有以下性质:

- (1) 若矩阵阶数相容, 则 $(A \otimes B)(C \otimes D) = AC \otimes BD$.
- (2) 若 A, X, B 的阶数相容, 则 $\text{vec}(AXB) = (B^T \otimes A) \text{vec}(X)$, 其中 $\text{vec}(X)$ 表示将矩阵 X 的各列依次堆叠得到的列向量。

证明

- (1) 设 $A = (a_{ij})$, $C = (c_{jk})$, 则 $(A \otimes B)(C \otimes D)$ 的第 (i, k) 个分块为

$$\sum_j (a_{ij}B)(c_{jk}D) = \left(\sum_j a_{ij}c_{jk} \right) BD = (AC)_{ik}BD.$$

故 $(A \otimes B)(C \otimes D) = AC \otimes BD$.

- (2) 设 $A \in M_{p \times n}(\mathbb{F})$, $X \in M_{n \times m}(\mathbb{F})$, $B \in M_{m \times q}(\mathbb{F})$. 记 $X = (x_{ij})$, 则

$$X = \sum_{i=1}^n \sum_{j=1}^m x_{ij} E_{ij}.$$

由于 $\text{vec}(E_{ij}) = e_j \otimes e_i$, 由性质 (1) 可得 $(B^T \otimes A) \text{vec}(E_{ij}) = (B^T e_j) \otimes (A e_i) = \text{vec}(A E_{ij} B)$. 再由线性性, 即得 $\text{vec}(AXB) = (B^T \otimes A) \text{vec}(X)$. □

定理 3.4 的证明 下面给出定理 3.4 的完整证明。

证明

首先设 $A' = PAP^{-1}$ 和 $B' = QBQ^{-1}$, 其中 $P \in M_p(\mathbb{F})$, $Q \in M_q(\mathbb{F})$ 均是可逆矩阵。那么

$$\begin{aligned} \varphi_{A,B}(X) &= AX - XB = P^{-1}A'PX - XQ^{-1}B'Q \\ &= P^{-1}(A'PX - XQ^{-1}B'Q) = P^{-1}(A'PXQ^{-1} - PXQ^{-1}B')Q. \end{aligned}$$

考虑线性映射 $T: M_{n \times m}(\mathbb{F}) \rightarrow M_{n \times m}(\mathbb{F})$, $T(X) = PXQ^{-1}$, 则 $\varphi_{A,B}(X) = \varphi_{A',B'}(T(X))$. 进而对矩阵 A, B 进行相似变换不会改变 Sylvester 算子的相似类。此时我们可以假设 A 是上三角矩阵, B 是下三角方

阵且 $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix} \in M_p(\mathbb{C})$, $B = \begin{pmatrix} b_{11} & & & \\ b_{21} & b_{22} & & \\ \vdots & \vdots & \ddots & \\ b_{q1} & b_{q2} & \cdots & b_{qq} \end{pmatrix} \in M_q(\mathbb{C})$. 然后, 我们下面利用

Kronecker 积得到 Sylvester 算子的矩阵表示:

$$\text{vec}(\varphi_{A,B}(X)) = \text{vec}(AX - XB) = (I_q \otimes A - B^T \otimes I_p) \text{vec}(X).$$

此时我们得到了 Sylvester 算子的矩阵表示 $I_q \otimes A - B^T \otimes I_p$. 由于 A, B^T 均为上三角方阵, 故 $I_q \otimes A, B^T \otimes I_p, I_q \otimes A - B^T \otimes I_p$ 也均为上三角方阵, 且上三角方阵 $I_q \otimes A - B^T \otimes I_p$ 的对角元 (即 $\varphi_{A,B}$ 的特征值) 为

$$\{a_{ii} - b_{jj} \mid 1 \leq i \leq p, 1 \leq j \leq q\}.$$

显然 $I_q \otimes A - B^T \otimes I_p$ 可逆 $\Leftrightarrow a_i \neq b_j, \forall i, j \Leftrightarrow A$ 与 B 没有公共特征值。 □

推论 3.3

$\varphi_{A,B}$ 的特征多项式是 $\prod_{1 \leq i \leq p, 1 \leq j \leq q} (x - a_{ii} + b_{jj})$.