理科专业基础课(71100201)

《代 数 与 几 何》

主讲人: 吕新民

第七章 线性变换

- 线性变换的定义与性质
- 线性变换与矩阵
- 特征值与特征向量
- 矩阵的对角化
- 线性变换的值域与核
- 矩阵的Jordan标准形

第1节 线性变换的定义与性质

- **定义**

定义 设V是数域P上的一个线性空间，映射$\mathcal{A}: V \to V$称为V上的一个线性变换，如果对任意$\alpha, \beta \in V$及任意的$k \in P$,

$$\mathcal{A}(\alpha+\beta)=\mathcal{A}(\alpha)+\mathcal{A}(\beta),\quad \mathcal{A}(k\alpha)=k\mathcal{A}(\alpha).$$

【例 1】(1) 设V是数域P上的一个线性空间，对于任意的$\alpha \in V$，恒等变换$\varepsilon(\alpha)=\alpha$与零变换$0(\alpha)=0$均是$V$上的线性变换.

(2) 在线性空间$P[x]_n$中，微商变换

$$\mathcal{D}(f(x)) = f'(x)$$

是$P[x]_n$上的线性变换.

(3) $C_{[a,b]}$表示闭区间$[a,b]$上的连续函数全体关于函数的加法和数乘运算构成的线性空间，则积分变换

$$\mathcal{A}(f(x)) = \int_a^x f(t)dt$$

是$C_{[a,b]}$上的线性变换.

直接验证易知线性变换具有如下性质.

性质 设V是数域P上的一个线性空间，$\mathcal{A}$为V上的一个线性变换，则

(1) $\mathcal{A}(0)=0.$

(2) $\mathcal{A}(-\alpha)=-\mathcal{A}(\alpha).$

(3) 若$\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s$，则$\mathcal{A}(\beta)=k_1\mathcal{A}(\alpha_1)+k_2\mathcal{A}(\alpha_2)+\cdots+k_s\mathcal{A}(\alpha_s)$;

(4) 线性变换把相关组变成相关组.

注: 线性变换虽然把相关组变成相关组. 但不能把无关组变成无关组. 如零变换将任何线性无关的向量组均变成线性相关的向量组.

- **线性变换的运算**

设V是数域P上的线性空间，令$L(V)$表示V的所有线性变换所构成的集合，即

$$L(V) = \{\mathcal{A} |\ \mathcal{A} : V \longrightarrow V \text{ 的线性变换}\}.$$

对任意$\mathcal{A}, \mathcal{B} \in L(V)$，规定

(1) 加法：$(\mathcal{A} + \mathcal{B})(\alpha) = \mathcal{A}(\alpha) + \mathcal{B}(\alpha)$，对于任意的$\alpha \in V$.

由规定，$\mathcal{A} + \mathcal{B}$是$V$上的线性变换，且满足以下运算规律:

①$\mathcal{A}+\mathcal{B}=\mathcal{B}+\mathcal{A}$.

②$(\mathcal{A}+\mathcal{B})+\mathcal{C}=\mathcal{A}+(\mathcal{B}+\mathcal{C})$.

③$\mathcal{A}+0=\mathcal{A}$.

④$\mathcal{A}+(-\mathcal{A})=0$ (这里的$-\mathcal{A}$定义为$(-\mathcal{A})(\alpha)=-\mathcal{A}(\alpha)$，称为$\mathcal{A}$的负变换).

(2) 数乘：$(k\mathcal{A})(\alpha)=k\mathcal{A}(\alpha)$，这里$k\in P$.

由规定，$k\mathcal{A}$仍是V上的线性变换，且满足以下运算规律:

⑤$1\cdot\mathcal{A}=\mathcal{A}$.

⑥$(k\ell)\mathcal{A}=k(\ell\mathcal{A})$，这里$k,\ell\in P$.

⑦$k(\mathcal{A}+\mathcal{B})=k\mathcal{A}+k\mathcal{B}$，这里$k\in P$.

⑧$(k+\ell)\mathcal{A}=k\mathcal{A}+\ell\mathcal{A}$，这里$k,\ell\in P$.

因此，$L(V)$关于上述定义的加法和数乘运算作成一个线性空间. $L(V)$除了上述两种运算外，在$L(V)$中还可以定义乘法运算.

(3) 乘法：$(\mathcal{A}\mathcal{B})(\alpha)=\mathcal{A}(\mathcal{B}(\alpha))$，这里的乘法运算即是通常映射的合成.

V上的变换$\mathcal{A}$称为可逆的，如果存在V上的变换

$\mathcal{B}$，使得$\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}=\varepsilon$. 直接验证知: 若$\mathcal{A}$可逆，则满足上述条件的变换$\mathcal{B}$是唯一的. 通常称$\mathcal{B}$是$\mathcal{A}$的逆变换，记作$\mathcal{A}^{-1}=\mathcal{B}$.

【例 2】设$\mathcal{A},\mathcal{B}$是线性空间V的两个线性变换，且$\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A}=\varepsilon$. 证明：对于任何正整数$n>1$，$\mathcal{A}^n\mathcal{B}-\mathcal{B}\mathcal{A}^n=n\mathcal{A}^{n-1}$.

证明：对n作数学归纳法. 当$n=2$时，

$$\begin{aligned}&\mathcal{A}^2\mathcal{B}-\mathcal{B}\mathcal{A}^2\\&=(\mathcal{A}^2\mathcal{B}-\mathcal{A}\mathcal{B}\mathcal{A})+(\mathcal{A}\mathcal{B}\mathcal{A}-\mathcal{B}\mathcal{A}^2)\\&=\mathcal{A}(\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A})+(\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A})\mathcal{A}\end{aligned}$$

$$= \mathcal{A}\varepsilon + \varepsilon\mathcal{A} = 2\mathcal{A}.$$

结论成立. 假定结论对$n = k$时成立，即

$$\mathcal{A}^k\mathcal{B} - \mathcal{B}\mathcal{A}^k = k\mathcal{A}^{k-1}.$$

现考察$n = k + 1$时的情形. 因为

$$\begin{aligned}
&\mathcal{A}^{k+1}\mathcal{B} - \mathcal{B}\mathcal{A}^{k+1} \\
&= (\mathcal{A}^{k+1}\mathcal{B} - \mathcal{A}^k\mathcal{B}\mathcal{A}) + (\mathcal{A}^k\mathcal{B}\mathcal{A} - \mathcal{B}\mathcal{A}^{k+1}) \\
&= \mathcal{A}^k(\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A}) + (\mathcal{A}^k\mathcal{B} - \mathcal{B}\mathcal{A}^k)\mathcal{A} \\
&= \mathcal{A}^k\varepsilon + k\mathcal{A}^k\varepsilon = (k+1)\mathcal{A}^k.
\end{aligned}$$

即$n = k + 1$时成立. 故对一切$n > 1$结论均成立.

第2节 线性变换与矩阵

• 线性变换的矩阵

为建立线性变换与矩阵之间的关系，先证如下两个结果.

引理1 设$\mathcal{A},\mathcal{B}$是n维线性空间V的两个线性变换，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是$V$的一组基. 证明: $\mathcal{A}=\mathcal{B}\Leftrightarrow\mathcal{A}(\varepsilon_i)=\mathcal{B}(\varepsilon_i)$，这里$i=1,2,\cdots,n$.

证明： 必要性是显然的. 下证充分性.

任取$\alpha\in V$，令$\alpha=\sum\limits_{i=1}^{n}k_i\varepsilon_i$. 于是

$$\mathcal{A}(\alpha)=\mathcal{A}(\sum_{i=1}^{n}k_i\varepsilon_i)=\sum_{i=1}^{n}k_i\mathcal{A}(\varepsilon_i)=\sum_{i=1}^{n}k_i\mathcal{B}(\varepsilon_i)=\mathcal{B}(\sum_{i=1}^{n}k_i\varepsilon_i)=\mathcal{B}(\alpha).$$

故$\mathcal{A}=\mathcal{B}$.

引理2 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是线性空间$V$的一组基，给定$\alpha_1,\alpha_2,\cdots,\alpha_n\in V$，存在唯一的线性变换$\mathcal{A}$，使得$\mathcal{A}\varepsilon_i=\alpha_i$，$i=1,2,\cdots,n$.

证明： 仅证存在性，唯一性直接由引理可得. 对任意$\alpha\in V$，令$\alpha=\sum\limits_{i=1}^{n}x_i\varepsilon_i$，作$\mathcal{A}:V\to V$为$\mathcal{A}\alpha=\sum\limits_{i=1}^{n}x_i\alpha_i$. 直接验证易知$\mathcal{A}$是$V$上的线性变换，且满

足$\mathcal{A}\varepsilon_i=\alpha_i$，$i=1,2,\cdots,n$.

上述引理表明：一个线性变换$\mathcal{A}$完全由$\mathcal{A}$在基下的象唯一确定. 因此，要想确定$\mathcal{A}$，只需确定每个$\mathcal{A}(\varepsilon_i)$即可. 反之，给定$\mathcal{A}(\varepsilon_1),\mathcal{A}(\varepsilon_2),\cdots,\mathcal{A}(\varepsilon_n)$，则$\mathcal{A}$也被唯一确定.

令

$$\begin{cases}\mathcal{A}(\varepsilon_1)=a_{11}\varepsilon_1+a_{21}\varepsilon_2+\cdots+a_{n1}\varepsilon_n\\ \mathcal{A}(\varepsilon_2)=a_{12}\varepsilon_1+a_{22}\varepsilon_2+\cdots+a_{n2}\varepsilon_n\\ \quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\ \mathcal{A}(\varepsilon_n)=a_{1n}\varepsilon_1+a_{2n}\varepsilon_2+\cdots+a_{nn}\varepsilon_n\end{cases}$$

若记$A=(a_{ij})$，则上述形式可记为

$$(\mathcal{A}(\varepsilon_1),\mathcal{A}(\varepsilon_2),\cdots,\mathcal{A}(\varepsilon_n))=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A.$$

若记

$$(\mathcal{A}(\varepsilon_1),\mathcal{A}(\varepsilon_2),\cdots,\mathcal{A}(\varepsilon_n))=\mathcal{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n),$$

则

$$\mathcal{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A.$$

给定一线性变换$\mathcal{A}$，上述矩阵A是被$\mathcal{A}$唯一确定. 反过来，给定一个矩阵A，依上述形式可唯一确定$\mathcal{A}(\varepsilon_i)$，进而可唯一确定一个线性变换$\mathcal{A}$.

定义 设V是数域P上的一个n维线性空间，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$是它的一组基，$\mathcal{A}$为$V$的一个线性变换，且

$$\mathcal{A}(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A,$$

称矩阵$A = (a_{ij})$是$\mathcal{A}$在基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的矩阵.

线性变换与矩阵之间的对应更重要地体现在下列运算上.

定理 设V是数域P上的一个n维线性空间，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$是它的一组基，则$V$的每个线性

变换$\mathcal{A}$必唯一的对应数域P上一个$n\times n$矩阵A，且

(1) 线性变换的和对应于矩阵的和.

(2) 线性变换的数乘对应于矩阵的数乘.

(3) 线性变换的积对应于矩阵的积.

(4) 可逆的线性变换对应于可逆的矩阵，且逆变换对应于逆矩阵.

证明： 仅证(3)，其余类似可证.

设$\mathcal{A},\mathcal{B}$是V上的两个线性变换，它们在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵分别为$A,B$，即

$$\mathcal{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A,$$

$$\mathcal{B}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B.$$

于是,

$$\begin{aligned}&(\mathcal{A}\mathcal{B})(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=\mathcal{A}(\mathcal{B}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n))\\&=\mathcal{A}((\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B)=(\mathcal{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n))B,\\&=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AB.\end{aligned}$$

故线性变换的乘积$\mathcal{A}\mathcal{B}$ 对应矩阵的乘积AB.

下面进一步考查V中任何一个向量ξ与$\mathcal{A}(\xi)$在同一组基下坐标之间的关系.

定 理 设V是数域P上的一个n维线性空间，$\mathcal{A}$在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵为$A$. 对于任意的$\xi\in V$，令$\xi$在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的坐标为$(x_1,x_2,\cdots,x_n)'$，$\mathcal{A}(\xi)$在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的坐标为$(y_1,y_2,\cdots,y_n)'$，则

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

证明： 因为

$$\xi = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

于是

$$\mathcal{A}(\xi) = (\mathcal{A}(\varepsilon_1), \mathcal{A}(\varepsilon_2), \cdots, \mathcal{A}(\varepsilon_n)) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

另一方面，由题设，有

$$\mathcal{A}(\xi) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

因为$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$线性无关，则

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

- **矩阵的相似**

在线性空间中，基的选取不是唯一的. 下面考察同一线性变换在两组基下坐标之间的关系.

定理 设V是数域P上的一个n维线性空间，$\mathcal{A}$在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵为$A$，而在基$\eta_1,\eta_2,\cdots,\eta_n$下的矩阵为$B$. 如果从基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$到基$\eta_1,\eta_2,\cdots,\eta_n$的过渡矩阵为$X$，则$B=X^{-1}AX$.

证明：由题设有

$$\mathcal{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A,$$

$$\mathcal{A}(\eta_1,\eta_2,\cdots,\eta_n)=(\eta_1,\eta_2,\cdots,\eta_n)B.$$

又

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X.$$

于是

$$\begin{aligned}\mathcal{A}(\eta_1,\eta_2,\cdots,\eta_n)&=\mathcal{A}[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X]\\&=[\mathcal{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)]X=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AX\\&=(\eta_1,\eta_2,\cdots,\eta_n)X^{-1}AX,\end{aligned}$$

比较可得$B=X^{-1}AX.$

对于矩阵A与B的这种关系，给出如下定义.

定义 数域P上的两个n阶方阵A,B称为相似的，

记作$A \sim B$，如果存在数域P上的n阶可逆矩阵X，使得$B = X^{-1}AX$.

相似是矩阵所有关系中最重要的关系，直接验证知这种关系具有：

(1) 反身性：$A \sim A$.

(2) 对称性：若$A \sim B$，则$B \sim A$.

(3) 传递性：若$A \sim B, B \sim C$，则$A \sim C$.

依据上述讨论知，同一个线性变换在不同基下的矩阵是相似的. 反过来，两个相似的矩阵也可以看成同一个线性变换在两组基下的矩阵.

定理 线性变换在不同基下的矩阵是相似的；反过来，如果两个矩阵相似，那么它们可以看作同一线性变换在两组基下所对应的矩阵.

证明：前一部分结论已证. 仅证后一部分即可.

设矩阵A是线性变换$\mathcal{A}$在基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的矩阵. 又设矩阵$A$与$B$相似，则存在可逆矩阵$X$，使的$B = X^{-1}AX$. 令

$$(\eta_1, \eta_2, \cdots, \eta_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)X,$$

则$\eta_1, \eta_2, \cdots, \eta_n$必是$V$ 的一组基，且

$$\mathcal{A}(\eta_1, \eta_2, \cdots, \eta_n) = (\mathcal{A}(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n))X$$

$$= (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)(AX)$$
$$= (\eta_1, \eta_2, \cdots, \eta_n)(X^{-1}AX),$$

即$\mathcal{A}$在基$\eta_1, \eta_2, \cdots, \eta_n$下的矩阵即为$B$. 故结论成立.

【例 3】在$P^{2\times 2}$中定义线性变换

$$\mathcal{A}(X) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} X \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

求$\mathcal{A}$在基$E_{11}, E_{12}, E_{21}, E_{22}$下的矩阵.

解：直接计算易得

$$\mathcal{A}(E_{11}) = a^2E_{11} + abE_{12} + acE_{21} + bcE_{22}.$$

$$\mathcal{A}(E_{12}) = acE_{11} + adE_{12} + c^2E_{21} + cdE_{22}.$$

$$\mathcal{A}(E_{21}) = abE_{11} + b^2E_{12} + adE_{21} + bdE_{22}.$$

$$\mathcal{A}(E_{22}) = bcE_{11} + bdE_{12} + cdE_{21} + d^2E_{22}.$$

故$\mathcal{A}$在基$E_{11}, E_{12}, E_{21}, E_{22}$下的矩阵为

$$A = \begin{pmatrix} a^2 & ac & ab & bc \\ ab & ad & b^2 & bd \\ ac & c^2 & ad & cd \\ bc & cd & bd & d^2 \end{pmatrix}.$$

【例 4】设$\mathcal{A}$是n维线性空间V的线性变换.

(1) 如果$\mathcal{A}^{n-1}(\xi) \neq 0$，而$\mathcal{A}^{n}(\xi) = 0$，证明: $\xi, \mathcal{A}(\xi), \cdots, \mathcal{A}^{n-1}(\xi)$构成$V$的一组基.

(2) 求$\mathcal{A}$在基$\xi, \mathcal{A}(\xi), \cdots, \mathcal{A}^{n-1}(\xi)$下的矩阵.

证明：(1)令

$$k_1\xi + k_2\mathcal{A}(\xi) + \cdots + k_n\mathcal{A}^{n-1}(\xi) = 0.$$

用$\mathcal{A}^{n-1}$作用于上式，得

$$k_1\mathcal{A}^{n-1}(\xi) = 0.$$

由题设$\mathcal{A}^{n-1}(\xi) \neq 0$，故$k_1 = 0$. 于是有

$$k_2\mathcal{A}(\xi) + \cdots + k_n\mathcal{A}^{n-1}(\xi) = 0.$$

再用$\mathcal{A}^{n-2}$作用于上式，得$k_2 = 0$. 继续这一过程，便可得$k_1 = k_2 = \cdots = k_n = 0$. 故$\xi, \mathcal{A}(\xi), \cdots, \mathcal{A}^{n-1}(\xi)$构成$V$的一组基.

(2) $\mathcal{A}$在基$\xi, \mathcal{A}(\xi), \cdots, \mathcal{A}^{n-1}(\xi)$下的矩阵为

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

第3节 特征值与特征向量

- **定义**

定义 设$\mathcal{A}$是数域P上线性空间V的一个线性变换，称$\lambda \in P$是$\mathcal{A}$的一个特征值，如果存在$0 \neq \xi \in V$，使得

$$\mathcal{A}\xi = \lambda\xi.$$

此时，称ξ是$\mathcal{A}$的从属于λ的特征向量.

由定义易知

(1) 若ξ_1与ξ_2均是$\mathcal{A}$的从属于λ的特征向量，则$\xi_1+\xi_2$也是$\mathcal{A}$的从属于λ的特征向量;

(2) 若ξ是$\mathcal{A}$的从属于λ的特征向量，则对于任意的$0 \neq k \in P$，$k\xi$均是$\mathcal{A}$的从属于λ的特征向量.

以上事实表明：$\mathcal{A}$的从属于λ 的特征向量的全体结合零向量作成V的一个子空间，记作V_λ. 通常称V_λ是V的特征子空间.

下面探讨如何求特征值与特征向量.

设V是数域P上的n维线性空间，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是它的一组基，令A是线性变换$\mathcal{A}$在这组基下的矩阵，又令X是ξ在这一组基下的坐标. 于是

$$\mathcal{A}\xi = \lambda\xi \Rightarrow AX = \lambda X,$$

即$(\lambda E - A)X = 0$. 因$\xi \neq 0$，所以$X \neq 0$，从而方程组$(\lambda E - A)X = 0$有非零解X，于是$|\lambda E - A| = 0$.

若记$A = (a_{ij})$，即有

$$|\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0.$$

通常称$f(\lambda) = |\lambda E - A|$为矩阵$A$的特征多项式，而称$f(\lambda) = |\lambda E - A| = 0$为$A$的特征方程.

在复数域$\mathbb{C}$范围内解方程$|\lambda E - A| = 0$，可得到n个特征值λ_i $(i = 1, 2, \cdots, n)$. 重根按重数计.

对每个λ_i，通过解齐次线性方程组$(\lambda_i E - A)X = 0$除零解外的通解，可得从属于λ_i的所有特征向量.

- **性质**

先观察特征多项式的某些特殊项

$$|\lambda E - A| = \lambda^n - (\sum_{i=1}^{n} a_{ii})\lambda^{n-1} + \cdots + (-1)^n |A|$$

又设$\lambda_1, \lambda_2, \cdots, \lambda_n$是$A$的$n$个特征值，从而

$$\begin{aligned}|\lambda E - A| &= (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)\\ &= \lambda^n - (\sum_{i=1}^{n} \lambda_i)\lambda^{n-1} + \cdots + (-1)^n \prod_{i=1}^{n} \lambda_i.\end{aligned}$$

比较以上关系式易得

$$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}, \quad \prod_{i=1}^{n} \lambda_i = |A|.$$

性质1 设A是数域P上的一个n阶方阵，$\lambda_1, \lambda_2, \cdots, \lambda_n$是$A$的$n$个特征值，则

(1) $\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}$，$\prod_{i=1}^{n} \lambda_i = |A|$.

(2) 若λ是A的特征值，则对于任意的正整数m，λ^m是A^m的特征值. 特别地，若$f(x) \in P[x]$，则$f(\lambda)$是$f(A)$的特征值.

(3) 若A可逆，λ是A的特征值，则λ^{-1}是A^{-1}的特征值.

证明：仅证(2)，其余类似可证.

若λ是A的特征值，则存在非零向量X，使得$AX = \lambda X$. 于是,

$$A^m X = A^{m-1}(AX) = A^{m-1}(\lambda X) = \lambda(A^{m-1}X) = \cdots = \lambda^m X.$$

故λ^m是A^m的特征值.

若$f(x) \in P[x]$，记$f(x) = a_0 + a_1 x + \cdots + a_k x^k$，则$f(A) = a_0 E + a_1 A + \cdots + a_k A^k$. 因为

$$\begin{aligned} f(A)X &= (a_0 E + a_1 A + \cdots + a_k A^k)X \\ &= a_0(EX) + a_1(AX) + \cdots + a_k(A^k X) \\ &= a_0 X + a_1(\lambda X) + \cdots + a_k(\lambda^k X) \\ &= (a_0 + a_1\lambda + \cdots + a_k\lambda^k)X = f(\lambda)X. \end{aligned}$$

故$f(\lambda)$是$f(A)$的特征值.

通常，称$\sum\limits_{i=1}^{n} a_{ii}$为矩阵$A=(a_{ij})$的迹，记作$\mathrm{Tr}(A)$.

性质2　设A是数域P上的一个n阶方阵，α,β分别是λ_1,λ_2的特征向量，则

(1) 若$\lambda_1\neq\lambda_2$，则α,β必线性无关.

(2) 若$\lambda_1\neq\lambda_2$，则$\alpha+\beta$一定不是A的特征向量.

证明：(1)由题设知$A(\alpha)=\lambda_1\alpha$，$A(\beta)=\lambda_2\beta$. 令

$$k_1\alpha+k_2\beta=0.$$

两边用A作用，得

$$k_1A(\alpha)+k_2A(\beta)=0，\text{即}k_1\lambda_1\alpha+k_2\lambda_2\beta=0.$$

现将$k_1\alpha+k_2\beta=0$两端同乘λ_1再减上式得

$$k_2(\lambda_1-\lambda_2)\beta=0.$$

因为$\lambda_1-\lambda_2\neq 0$，且$\beta\neq 0$，故$k_2=0$. 进而得$k_1=0$. 因此，$\alpha,\beta$必线性无关.

(2) 反证法. 假定$\alpha+\beta$是A的从属于λ的特征向量，则$A(\alpha+\beta)=\lambda(\alpha+\beta)$，即$A(\alpha)+A(\beta)=\lambda\alpha+\lambda\beta$. 于是，$\lambda_1\alpha+\lambda_2\beta=\lambda\alpha+\lambda\beta$，即$(\lambda_1-\lambda)\alpha+(\lambda_2-\lambda)\beta=0$. 因为$\alpha,\beta$线性无关，故$\lambda_1=\lambda_2=\lambda$，矛盾. 故$\alpha+\beta$一定不是$A$的特征向量.

上述性质表明任何一个n阶方阵A的从属于不同特征值的特征向量必线性无关.

直接由定义可得

性质3 设A, B是数域P上的两个n阶方阵，且$A \sim B$，则

(1) 对于任意的正整数m，$A^m \sim B^m$. 特别地，若$f(x) \in P[x]$，则$f(A) \sim f(B)$.

(2) $|\lambda E - A| = |\lambda E - B|$.

- **Hamilton-Cayley定理**

最后，我们介绍一个非常重要的定理.

哈密尔顿–凯莱(Hamilton-Cayley)定理 设A是数域P上的n阶方阵，$f(\lambda)=|\lambda E-A|$是A的特征多项式，则$f(A)=0$.

【例5】 设$A=\begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{pmatrix}$. 求$A$的特征值与特征向量.

解： A的特征方程为

$$|\lambda E-A|=\begin{vmatrix}\lambda-1&-1&-1&-1\\-1&\lambda-1&1&1\\-1&1&\lambda-1&1\\-1&1&1&\lambda-1\end{vmatrix}=(\lambda-2)^3(\lambda+2).$$

故A的特征值为$\lambda_1=\lambda_2=\lambda_3=2$，$\lambda_4=-2$.

当$\lambda=2$时，$(2E-A)X=0$的系数矩阵为

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

得基础解系为

$$\eta_1=\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \eta_2=\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \eta_3=\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

故从属于2的特征向量为$k_1\eta_1 + k_2\eta_2 + k_3\eta_3$ (k_1, k_2, k_3不全为零).

当$\lambda = -2$时，$(-2E - A)X = 0$的系数矩阵为

$$A = \begin{pmatrix} -3 & -1 & -1 & -1 \\ -1 & -3 & 1 & 1 \\ -1 & 1 & -3 & 1 \\ -1 & 1 & 1 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & -1 & 3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

的基础解系为

$$\eta_4 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

故从属于-2的特征向量为$k_4\eta_4$ (k_4为非零常数).

【例6】已知

$$\xi=\begin{pmatrix}1\\1\\-1\end{pmatrix}\text{是矩阵}A=\begin{pmatrix}2&-1&2\\5&a&3\\-1&b&-2\end{pmatrix}$$

的从属于λ的一个特征向量，求λ及参数a,b.

解：由定义有$A\xi=\lambda\xi$，即

$$\begin{pmatrix}2&-1&2\\5&a&3\\-1&b&-2\end{pmatrix}\begin{pmatrix}1\\1\\-1\end{pmatrix}=\lambda\begin{pmatrix}1\\1\\-1\end{pmatrix}.$$

进而有下列方程组

$$\begin{cases} 2-1-2=\lambda \\ 5+a-3=\lambda \\ -1+b+2=\lambda \end{cases}$$

解之得$\lambda=-1, a=-3, b=0.$

【例 7】 设$A=\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. 证明: 对于任何正整数$n\geq 3$，$A^n=A^{n-2}+A^2-E$. 并求$A^{100}$.

解： A的特征多项式为

$$|\lambda E-A|=\begin{vmatrix} \lambda-1 & 0 & 0 \\ -1 & \lambda & -1 \\ 0 & -1 & \lambda \end{vmatrix}=\lambda^3-\lambda^2-\lambda+1.$$

由Hamilton-Cayley知$A^3 - A^2 - A + E = 0$.

现对n作归纳. 当$n = 3$时，结论显然. 假定结论对$n = k$时成立，即$A^k = A^{k-2} + A^2 - E$. 现考察情形$n = k+1$. 因为

$$\begin{aligned} A^{k+1} &= AA^k = A(A^{k-2} + A^2 - E) \\ &= A^{k-1} - A^3 - A = A^{k-1} + A^2 - E, \end{aligned}$$

由归纳假定知对于任何正整数$n \geq 3$，结论均成立.

利用上述递推公式可得

$$\begin{aligned} A^{100} &= A^{98} + A^2 - E = A^{96} + 2(A^2 - E) \\ &= \cdots = A^2 + 49(A^2 - E) = 50A^2 - 49E. \end{aligned}$$

直接计算知

$$A^2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

故

$$A^{100} = 50A^2 - 49E$$

$$= 50 \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} - 49 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 50 & 1 & 0 \\ 50 & 0 & 1 \end{pmatrix}.$$

注: Hamilton-Cayley定理建立了一个矩阵A的降次公式，这为计算一个矩阵A的n次方提供了一种方法.

第4节 矩阵的对角化

- **矩阵可对角化的判定I**

设A是数域P上的n阶方阵，若存在可逆矩阵X，使得$X^{-1}AX$为对角矩阵Λ，称A可对角化.

定理1 设A是数域P上的n阶方阵，则A可对角化$\Leftrightarrow A$有n个线性无关的特征向量.

证明： $\Rightarrow$ 若A可对角化，则存在可逆矩阵X，使得

$$X^{-1}AX = \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

即$AX = X\Lambda$. 记$X = (X_1, X_2, \cdots, X_n)$，上式即为

$$A(X_1, X_2, \cdots, X_n) = (X_1, X_2, \cdots, X_n)\Lambda,$$

即

$$(AX_1, AX_2, \cdots, AX_n) = (\lambda_1 X_1, \lambda_2 X_2, \cdots, \lambda_n X_n).$$

于是，$AX_i = \lambda_i X_i$，其中$i = 1, 2, \cdots, n$. 结合X可逆，知$X_1, X_2, \cdots, X_n$是A的n个线性无关的特征向量，且$\lambda_1, \lambda_2, \cdots, \lambda_n$是$A$的$n$个特征值.

$\Leftarrow$ 若A有n个线性无关的特征向量$X_1, X_2, \cdots, X_n$，则必存在数$\lambda_1, \lambda_2, \cdots, \lambda_n$，使得$AX_i = \lambda_i X_i$，其中$i = 1, 2, \cdots, n$. 于是，

$$(AX_1, AX_2, \cdots, AX_n) = (\lambda_1 X_1, \lambda_2 X_2, \cdots, \lambda_n X_n).$$

记$X=(X_1,X_2,\cdots,X_n)$，上式即为

$$AX=X\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n).$$

因为$X_1,X_2,\cdots,X_n$线性无关，所以X可逆，且

$$X^{-1}AX=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n).$$

故A可对角化.

作为上述结果的直接推论有

推论 设A是数域P上的n阶方阵，若A有n个互异的特征值，则A必可对角化.

- **矩阵可对角化的判定II**

设A是一个n阶方阵，λ是A的一个特征值. λ作为特征方程$|\lambda E - A| = 0$根的重数，称为λ的代数重数，而λ的特征子空间V_λ的维数(即从属于λ的线性无关的特征向量的个数)，称为λ的几何重数.

定理2 设A是数域P上的n阶方阵，则对于A的任一特征值λ_0，λ_0的几何重数$\leq \lambda_0$的代数重数.

证明： 设λ_0的代数重数为m. 假定λ_0的几何重数大于m，则存在A的属于λ_0的$m+1$个线性无关的特征向量

$$\alpha_1, \alpha_2, \cdots, \alpha_{m+1}.$$

现将其扩充成n个线性无关的向量

$$\alpha_1, \alpha_2, \cdots, \alpha_{m+1}, \alpha_{m+2}, \cdots, \alpha_n.$$

记$T = (\alpha_1, \alpha_2, \cdots, \alpha_{m+1}, \alpha_{m+2}, \cdots, \alpha_n)$，则

$$\begin{aligned} AT &= (A\alpha_1, A\alpha_2, \cdots, A\alpha_{m+1}, A\alpha_{m+2}, \cdots, A\alpha_n) \\ &= (\lambda_0\alpha_1, \lambda_0\alpha_2, \cdots, \lambda_0\alpha_{m+1}, A\alpha_{m+2}, \cdots, A\alpha_n) \end{aligned}$$

$$= (\alpha_1, \alpha_2, \cdots, \alpha_{m+1}, \alpha_{m+2}, \cdots, \alpha_n) \begin{pmatrix} \lambda_0 & & & & \\ & \lambda_0 & & & * \\ & & \ddots & & \\ & & & \lambda_0 & \\ & & 0 & & * \end{pmatrix}$$

即

$$T^{-1}AT=\begin{pmatrix} \lambda_0 & & & & \\ & \lambda_0 & & & * \\ & & \ddots & & \\ & & & \lambda_0 & \\ & & 0 & & * \end{pmatrix}.$$

于是

$$|\lambda E-A|=\begin{vmatrix} \lambda-\lambda_0 & & & & \\ & \lambda-\lambda_0 & & & * \\ & & \ddots & & \\ & & & \lambda-\lambda_0 & \\ & & 0 & & * \end{vmatrix}=(\lambda-\lambda_0)^{m+1}g(\lambda).$$

其中$g(\lambda)$是关于λ的多项式，所以λ_0的代数重数至少是$m+1$，矛盾.

作为上述结果的直接推论有

推论 设A是数域P上的n阶方阵，则A可对角化$\Leftrightarrow$ 对于A的任一特征值λ，λ的几何重数$=\lambda$的代数重数.

显然，若λ的代数重数$=1$，则λ的几何重数$=1$. 当λ的代数重数>1时，λ的几何重数未必>1.

【例8】 设$A=\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix}$.

(1) 求A的特征值及特征向量；

(2) 求A^k $(k\in\mathbb{N})$.

解：(1) A的特征多项式

$$|\lambda E-A|=\begin{vmatrix}\lambda-2 & 0 & 0\\ -1 & \lambda-2 & 1\\ -1 & 0 & \lambda-1\end{vmatrix}=(\lambda-2)^2(\lambda-1).$$

故A的特征值为$\lambda=2$或$\lambda=1.$

当$\lambda=2$时，$(2E-A)X=0$的系数矩阵为

$$2E-A=\begin{pmatrix}0 & 0 & 0\\ -1 & 0 & 1\\ -1 & 0 & 1\end{pmatrix}\rightarrow\begin{pmatrix}1 & 0 & -1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

的基础解系为

$$\eta_1=\begin{pmatrix}0\\1\\0\end{pmatrix},\eta_2=\begin{pmatrix}1\\0\\1\end{pmatrix}.$$

故从属于$\lambda=2$的特征向量为$k_1\eta_1+k_2\eta_2$ (k_1,k_2是不同时为零的任意常数);

当$\lambda=1$时，$(E-A)X=0$的系数矩阵为

$$E-A=\begin{pmatrix}-1&0&0\\-1&-1&1\\-1&0&0\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&1&-1\\0&0&0\end{pmatrix}.$$

的基础解系为

$$\eta_3=\begin{pmatrix}0\\1\\1\end{pmatrix}.$$

故从属于$\lambda=1$的特征向量为$k_3\eta_3$ (k_3是非零的任意常数).

(2) 令$P=\begin{pmatrix}0&1&0\\1&0&1\\0&1&1\end{pmatrix}$，则$P^{-1}AP=\begin{pmatrix}2&0&0\\0&2&0\\0&0&1\end{pmatrix}$. 于是

$$A^k=P\begin{pmatrix}2&0&0\\0&2&0\\0&0&1\end{pmatrix}P^{-1}$$

$$=\begin{pmatrix}0&1&0\\1&0&1\\0&1&1\end{pmatrix}\begin{pmatrix}2&0&0\\0&2&0\\0&0&1\end{pmatrix}^k\begin{pmatrix}1&1&-1\\1&0&0\\-1&0&1\end{pmatrix}$$

$$=\begin{pmatrix}2^k&0&0\\2^k-1&2^k&-2^k+1\\2^k-1&0&1\end{pmatrix}.$$

【例9】设$A=\begin{pmatrix}3 & 2 & -2\\ -k & -1 & k\\ 4 & 2 & -3\end{pmatrix}$. 问$k$当为何值时，存在可逆矩阵$P$使$P^{-1}AP$化成对角阵? 并求出$P$和相应的对角阵.

解：由题设，A的特征多项式为

$$f(\lambda)=|\lambda E-A|=(\lambda-1)(\lambda+1)^2,$$

故$\lambda_1=\lambda_2=-1, \lambda_3=1$是$A$的特征值.

$\lambda_1=\lambda_2=-1$是重根(即-1的代数重数为2)，因此要想A可对角化，从属于-1的特征向量应有两个，于是$\mathrm{rank}(-E-A)=1$. 而

$$-E-A=\begin{pmatrix}-4 & -2 & 2\\ k & 0 & -k\\ -4 & -2 & 2\end{pmatrix}\to\begin{pmatrix}-4 & -2 & 2\\ k & 0 & -k\\ 0 & 0 & 0\end{pmatrix}.$$

欲使$\operatorname{rank}(-E-A)=1$，只有$k=0$.

解方程组$(E+A)X=0$和$(E-A)X=0$可分别得相应于$\lambda_1=\lambda_2=-1$和$\lambda_3=1$的特征向量

$$\alpha_1=\begin{pmatrix}-1\\ 2\\ 0\end{pmatrix},\alpha_2=\begin{pmatrix}1\\ 0\\ 2\end{pmatrix},\alpha_3=\begin{pmatrix}1\\ 0\\ 1\end{pmatrix},$$

故

$$P=\begin{pmatrix}-1 & 1 & 1\\ 2 & 0 & 0\\ 0 & 2 & 1\end{pmatrix}，\text{使得}P^{-1}AP=\begin{pmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

第5节 线性变换的值域与核

- **定义**

定义 设$\mathcal{A}$是线性空间V的一个线性变换.

(1) 称$\mathcal{A}^{-1}(0)=\{\xi\in V|\ \mathcal{A}(\xi)=0\}$为$\mathcal{A}$的核，有时也记作$Ker\mathcal{A}$.

(2) 称$\mathcal{A}(V)=\{\mathcal{A}(\xi)|\ \xi\in V\}$为$\mathcal{A}$的值域，有时也记作$Im\mathcal{A}$.

显然，$\mathcal{A}^{-1}(0)$与$\mathcal{A}(V)$均是V的子空间. $\mathcal{A}^{-1}(0)$的维数称为$\mathcal{A}$的零度，$\mathcal{A}(V)$的维数称为$\mathcal{A}$的秩.

定理 设$\mathcal{A}$是n维线性空间V的一个线性变换，$\mathcal{A}$

在基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的矩阵为$A$，则

(1) $\mathcal{A}(V) = L(\mathcal{A}(\varepsilon_1), \mathcal{A}(\varepsilon_2), \cdots, \mathcal{A}(\varepsilon_n))$.

(2) $\mathcal{A}$的秩=rank(A).

证明：(1)是显然的. 下证(2). 利用(1)的结果有

$$\mathcal{A}\text{的秩}=\mathrm{rank}\{\mathcal{A}(\varepsilon_1), \mathcal{A}(\varepsilon_2), \cdots, \mathcal{A}(\varepsilon_n)\}.$$

而

$$(\mathcal{A}(\varepsilon_1), \mathcal{A}(\varepsilon_2), \cdots, \mathcal{A}(\varepsilon_n)) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A.$$

因为A即是$\mathcal{A}(\varepsilon_1), \cdots, \mathcal{A}(\varepsilon_n)$在基$\varepsilon_1, \cdots, \varepsilon_n$下的坐标作为$A$的列构成的，因为$V \cong P^n$，所以$\mathcal{A}(\varepsilon_1), \cdots, \mathcal{A}(\varepsilon_n)$与$A$的列向量组具有相同的线性

相关性. 故$\mathcal{A}$的秩=rank(A).

- **值域与核的关系**

定理 设$\mathcal{A}$是n维线性空间V的一个线性变换，$\mathcal{A}(V)$的一组基的原像及$\mathcal{A}^{-1}(0)$ 的一组基合起来就是V的一组基. 特别地，$\mathcal{A}$的秩+$\mathcal{A}$的零度$=n$.

证明： 设$\mathcal{A}(V)$的一组基为$\eta_1,\cdots,\eta_r$，它们的原像为$\varepsilon_1,\cdots,\varepsilon_r$，即$\mathcal{A}(\varepsilon_i)=\eta_i$，$i=1,2,\cdots,r$. 又令$\mathcal{A}^{-1}(0)$的一组基为$\varepsilon_{r+1},\cdots,\varepsilon_s$. 下证：$\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_s$ 是V的一组基. 令

$$\ell_1\varepsilon_1+\cdots+\ell_r\varepsilon_r+\ell_{r+1}\varepsilon_{r+1}+\cdots+\ell_s\varepsilon_s=0$$

两端用$\mathcal{A}$作用，注意$\mathcal{A}(\varepsilon_i)=0$，$i=r+1,\cdots,s$，得

$$\ell_1\mathcal{A}(\varepsilon_1)+\cdots+\ell_r\mathcal{A}(\varepsilon_r)=0$$

即$\ell_1\eta_1+\cdots+\ell_r\eta_r=0$. 因$\eta_1,\eta_2,\cdots,\eta_r$是线性无关的，所以$\ell_1=\cdots=\ell_r=0$. 于是，

$$\ell_{r+1}\varepsilon_{r+1}+\cdots+\ell_s\varepsilon_s=0.$$

又因为$\varepsilon_{r+1},\cdots,\varepsilon_s$是$\mathcal{A}^{-1}(0)$的一组基，从而$\ell_{r+1}=\cdots=\ell_s=0$. 故$\varepsilon_1,\cdots,\varepsilon_r,\varepsilon_{r+1},\cdots,\varepsilon_s$是线性无关的.

对任意$\alpha \in V$，因为$\mathcal{A}(\alpha) \in \mathcal{A}(V)$，必存在数$k_1, \cdots, k_r$，使得

$$\mathcal{A}(\alpha) = k_1\mathcal{A}(\varepsilon_1) + \cdots + k_r\mathcal{A}(\varepsilon_r),$$

即$\mathcal{A}(\alpha - k_1\varepsilon_1 - \cdots - k_r\varepsilon_r) = 0$. 于是$\alpha - k_1\varepsilon_1 - \cdots - k_r\varepsilon_r \in \mathcal{A}^{-1}(0)$，因此存在数$k_{r+1}, \cdots, k_s$，使得

$$\alpha - k_1\varepsilon_1 - \cdots - k_r\varepsilon_r = k_{r+1}\varepsilon_{r+1} + \cdots + k_s\varepsilon_s,$$

即

$$\alpha = k_1\varepsilon_1 + \cdots + k_r\varepsilon_r + k_{r+1}\varepsilon_{r+1} + \cdots + k_s\varepsilon_s$$

故$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_r, \varepsilon_{r+1}, \varepsilon_{r+2}, \cdots, \varepsilon_s$构成$V$的一组基.

由题设，$\dim(V) = n$，所以$s = n$. 定理证毕.

【例 10】设$\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$是四维线性空间$V$的一组基，已知线性变换$\mathcal{A}$在这组基下的矩阵为

$$A=\begin{pmatrix}1&0&2&1\\-1&2&2&3\\1&3&5&5\\2&-2&1&-2\end{pmatrix},$$

求$\mathcal{A}V$与$\mathcal{A}^{-1}(0)$.

解： 先求$\mathcal{A}^{-1}(0)$. 设$\xi\in\mathcal{A}^{-1}(0)$，则$\mathcal{A}(\xi)=0$. 令$\xi$的坐标为$(x_1,x_2,x_3,x_4)$，则

$$\begin{pmatrix}1&0&2&1\\-1&2&1&3\\1&2&5&5\\2&-2&1&-2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

因为$\text{rank}(A)=2$，可得方程组的基础解系为

$$X_1=\begin{pmatrix}-2\\-\frac{3}{2}\\1\\0\end{pmatrix},X_2=\begin{pmatrix}-1\\-2\\0\\1\end{pmatrix}$$

若令

$$\alpha_1=(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)X_1,\alpha_2=(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)X_2$$

则α_1,α_2是$\mathcal{A}^{-1}(0)$的一组基，所以$\mathcal{A}^{-1}(0)=L(\alpha_1,\alpha_2)$.

再求$\mathcal{A}(V)$. 依题意，知

$$\mathcal{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_4)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_4)A,$$

于是

$$\begin{cases} \mathcal{A}(\varepsilon_1) = \varepsilon_1 - \varepsilon_2 + \varepsilon_3 + 2\varepsilon_4 \\ \mathcal{A}(\varepsilon_2) = 2\varepsilon_2 + 2\varepsilon_3 - 2\varepsilon_4 \\ \mathcal{A}(\varepsilon_3) = 2\varepsilon_1 + \varepsilon_2 + 5\varepsilon_3 + \varepsilon_4 \\ \mathcal{A}(\varepsilon_4) = \varepsilon_1 + 3\varepsilon_2 + 5\varepsilon_3 - 2\varepsilon_4 \end{cases}$$

因$\mathrm{rank}(A) = 2$，故$\mathcal{A}(\varepsilon_1), \mathcal{A}(\varepsilon_2), \mathcal{A}(\varepsilon_3), \mathcal{A}(\varepsilon_4)$的秩为$2$. 直接观察$\mathcal{A}(\varepsilon_1), \mathcal{A}(\varepsilon_2)$线性无关，可组成一组基，故$\mathcal{A}(V) = L(\mathcal{A}(\varepsilon_1), \mathcal{A}(\varepsilon_2))$.

第6节 矩阵的Jordan标准形

- 不变子空间

定义 设$\mathcal{A}$是数域P上线性空间V的线性变换，V的子空间W称为不变子空间，如果对任意$\xi \in W$，$\mathcal{A}(\xi) \in W$，此时，也称W为$\mathcal{A}$-子空间.

对于一个线性空间而言，前面讨论的很多子空间均是不变子空间.

【例 11】 设$\mathcal{A}$是数域P上线性空间V的线性变换.

(1) V本身与零子空间$\{0\}$均是$\mathcal{A}$-子空间.

(2) $\mathcal{A}$的值域$\mathcal{A}(V)$与核$\mathcal{A}^{-1}(0)$均是$\mathcal{A}$-子空间.

(3) $\mathcal{A}$的任何一个特征子空间都是$\mathcal{A}$-子空间.

(4) 若$\mathcal{A}$是数乘变换，则V的每个子空间都是$\mathcal{A}$-子空间.

(5) V的任意两个$\mathcal{A}$-子空间的交与和均是V的$\mathcal{A}$-子空间.

- **Jordan标准形**

定义 形式为

$$J(\lambda,t)=\begin{pmatrix}\lambda & 0 & \cdots & 0 & 0 & 0\\ 1 & \lambda & \cdots & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & \lambda & 0\\ 0 & 0 & \cdots & 0 & 1 & \lambda\end{pmatrix}_{t\times t} \text{或} \begin{pmatrix}\lambda & 1 & \cdots & 0 & 0 & 0\\ 0 & \lambda & \cdots & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & \lambda & 1\\ 0 & 0 & \cdots & 0 & 0 & \lambda\end{pmatrix}_{t\times t}.$$

的矩阵称为Jordan块，其中λ是复数.

定义 由若干个Jordan块组成的准对角矩阵称为Jordan矩阵，其一般形式为

$$J=\begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & J_s \end{pmatrix},$$

其中每个J_i是一个Jordan块.

一个n阶方阵A可对角化$\Leftrightarrow A$有n个线性无关的特征向量. 因此，若A没有n个线性无关的特征向量，则A是不能对角化. 此时，有下列更一般的结论.

定理 复数域上的每一个方阵均与一个Jordan矩

阵相似.

【例 12】设A是一个n阶下三角矩阵. 证明:

(1) 如果$a_{ii} \neq a_{jj}$ $(i \neq j, i, j = 1, 2, \cdots, n)$，那么$A$可相似对角化;

(2) 如果$a_{11} = a_{22} = \cdots = a_{nn}$，且至少有一个$a_{ij} \neq 0$ $(i > j)$，那么A不能相似对角化.

证明: (1) A的特征多项式为

$$|\lambda E - A| = (\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}),$$

即A的特征值为$a_{11}, a_{22}, \cdots, a_{nn}$. 因为$a_{ii} \neq a_{jj}$，因此$A$有$n$个互异的特征值，从而$A$可相似对角化.

(2) 反证. 假定A相似于对角阵Λ，不妨设

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n).$$

则A与Λ具有相同的特征值，但$|\lambda E - A| = (\lambda - a_{11})^n$，因此$\lambda_1 = \lambda_2 = \cdots = \lambda_n = a_{11}$，故$\Lambda = a_{11}E$，于是存在可逆矩阵$P$，使得

$$A = P^{-1}\Lambda P = P^{-1}a_{11}EP = a_{11}E$$

显然与所给的矩阵A矛盾，故A不能相似对角化.

习题讨论课七

一、填空题

1. 在$P[x]_3$中定义如下线性变换：$\mathcal{A}(f(x)) = (a_0 + a_2) + a_1 x$，其中$f(x) = a_0 + a_1 x + a_2 x^2 \in P[x]_3$，则$\mathcal{A}$在一组基$1, x + x^2, x - x^2$下的矩阵为_______.

2. 设A是n阶方阵，α, β是两个互异的n维列向量，且$A\alpha = \beta, A\beta = \alpha$，则$A$必有特征值_______.

3. 若3阶方阵满足$|A + E| = |A + 2E| = |A + 3E| = 0$，则$|A + 4E| =$_______.

4. 已知$\alpha = \begin{pmatrix} 1 \\ k \\ 1 \end{pmatrix}$是$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$的逆阵$A^{-1}$的特征向量，则$k =$_______.

5. 设$\varepsilon_1,\varepsilon_2,\varepsilon_3$是线性空间$V$的一组基，$\mathcal{A}$是$V$上的线性变换，且$\mathcal{A}\varepsilon_1=\varepsilon_1+\varepsilon_3,\mathcal{A}\varepsilon_2=\varepsilon_2+\varepsilon_3,\mathcal{A}\varepsilon_3=\varepsilon_1+\varepsilon_2+2\varepsilon_3$，则$\mathcal{A}$的零度为_______.

二、解答与证明题.

6.已知两个矩阵

$$A=\begin{pmatrix}-2&0&0\\2&x&2\\3&1&1\end{pmatrix},\quad B=\begin{pmatrix}-1&0&0\\0&2&0\\0&0&y\end{pmatrix},$$

相似.

(1) 求x,y的值;

(2) 求可逆矩阵P，使得$P^{-1}AP=B$.

7. 设A是3阶矩阵，$\alpha_1,\alpha_2,\alpha_3$是三维线性无关的列向量，它们满足$A\alpha_1=\alpha_2+\alpha_3,A\alpha_2=\alpha_3+\alpha_1,A\alpha_3=\alpha_1+\alpha_2$. 问$A$是否可对角化？如果可以对角化，求可逆矩阵$P$及对角矩阵$\Lambda$，使$P^{-1}AP=\Lambda$.

8. 设A是3阶方阵，$\beta=\begin{pmatrix}9\\18\\-18\end{pmatrix}$，且方程组$AX=\beta$的通解为

$$\eta=\begin{pmatrix}1\\2\\-2\end{pmatrix}+k_1\begin{pmatrix}-2\\1\\0\end{pmatrix}+k_2\begin{pmatrix}2\\0\\1\end{pmatrix},$$

(1) 求A的特征值和特征向量；

(2) 求矩阵A.

9. 设矩阵$A=\begin{pmatrix}1&2&-3\\-1&4&-3\\1&a&5\end{pmatrix}$的特征方程有一个二重特征根，求$a$的值，并讨论$A$是否可相似对角化.

10. 设A是一个n阶方阵，如果对于任意互异的实数a,b，A满足关系式$A^2-(a+b)A+abE=0$. 证明：A必可相似对角化.

习题讨论课七参考解答

1. **解：** 因为

$$\begin{cases} \mathcal{A}(1) = 1 \\ \mathcal{A}(x+x^2) = 1 + x = 1 + \frac{1}{2}(x+x^2) + \frac{1}{2}(x-x^2) \\ \mathcal{A}(x-x^2) = -1 + x = -1 + \frac{1}{2}(x+x^2) + \frac{1}{2}(x-x^2) \end{cases},$$

所以$\mathcal{A}$在一组基$1, x+x^2, x-x^2$下的矩阵为$\begin{pmatrix} 1 & 1 & -1 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

2. **解：** 因为$A\alpha = \beta, A\beta = \alpha$，则$A(\alpha-\beta) = -1\cdot(\alpha-\beta)$. 由题设$\alpha-\beta \neq 0$，从而$-1$必是$A$的一个特征值.

3. **解：** 由题设知，$\lambda = -1, -2, -3$是A的三个特征值. 由特征值的性质知$A+4E$的三个特征值为$3, 2, 1$，故$|A+4E| = 1\cdot 2\cdot 3 = 6$.

4. **解：** 令α是矩阵A^{-1}的从属于λ的特征向量，则$A^{-1}\alpha = \lambda\alpha$，从而

$A\alpha=\frac{1}{\lambda}\alpha$，即

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ k \\ 1 \end{pmatrix}=\frac{1}{\lambda}\begin{pmatrix} 1 \\ k \\ 1 \end{pmatrix},$$

解之得$k=-2$或1.

5. **解：** 显然

$$\mathcal{A}(\varepsilon_1,\varepsilon_2,\varepsilon_3)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

因为$\mathcal{A}$的秩=rank(A)，而

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}\to\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

所以$\mathcal{A}$的秩为2，故$\mathcal{A}$的零度为$3-\mathcal{A}$的秩$=1$.

6. **解**：(1)A, B的特征多项式分别为

$$|\lambda E - A| = (\lambda+2)[\lambda^2 - (x+1)\lambda + (x-2)]$$

$$|\lambda E - B| = (\lambda+1)(\lambda-2)(\lambda-y)$$

因为A与B相似，所以$|\lambda E - A| = |\lambda E - B|$，由此可解得$x = 0, y = -2.$

(2) A的特征值为$\lambda_1 = -1, \lambda_2 = 2, \lambda_3 = -2$，相应的特征向量分别为

$$\eta_1 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}, \quad \eta_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \eta_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

故所求$P = \begin{pmatrix} 0 & 0 & -1 \\ -2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$，使得$P^{-1}AP = B.$

7. **解**：由题设可得

$$A(\alpha_1+\alpha_2+\alpha_3)=2(\alpha_1+\alpha_2+\alpha_3)$$
$$A(\alpha_2-\alpha_1)=-(\alpha_2-\alpha_1)$$
$$A(\alpha_3-\alpha_1)=-(\alpha_3-\alpha_1)$$

所以A有特征值$\lambda=2,-1,-1$，且属于$\lambda=2$的特征向量为$\alpha_1+\alpha_2+\alpha_3$，属于$\lambda=-1$的特征向量为$\alpha_2-\alpha_1,\alpha_3-\alpha_1$. 直接验证可知向量组

$$\alpha_1+\alpha_2+\alpha_3,\alpha_2-\alpha_1,\alpha_3-\alpha_1$$

线性无关，从而A有三个线性无关的特征向量，因此A可相似对角化，且

$$P=(\alpha_1+\alpha_2+\alpha_3,\alpha_2-\alpha_1,\alpha_3-\alpha_1),$$

使得$P^{-1}AP=\begin{pmatrix}2&&\\&-1&\\&&-1\end{pmatrix}$.

8. **解：** (1) 依题意，方程组$AX=\beta$对应的齐次线性方程组$AX=0$的基础解系为

$$\xi_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \xi_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix},$$

于是$A\xi_1 = 0, A\xi_2 = 0$，即ξ_1, ξ_2是A的从属于$\lambda = 0$的两个线性无关的特征向量. 又因为

$$A\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 9 \\ 18 \\ -18 \end{pmatrix} = 9\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix},$$

即$\xi_3 = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$是$A$的从属于$\lambda = 9$的特征向量.

(2) 令

$$P = (\xi_1, \xi_2, \xi_3)，使得 P^{-1}AP = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

故$A=P\begin{pmatrix}0&0&0\\0&0&0\\0&0&9\end{pmatrix}P^{-1}=\begin{pmatrix}1&2&-2\\2&4&-4\\-2&-4&4\end{pmatrix}$.

9. **解：** A的特征多项式为

$$|\lambda E-A|=\begin{vmatrix}\lambda-1&-2&3\\1&\lambda-4&3\\-1&-a&\lambda-5\end{vmatrix}=(\lambda-2)(\lambda^2-8\lambda+18+3a).$$

(1) 若$\lambda=2$是特征方程的二重根，则$2^2-16+18+3a=0$，解得$a=-2$.

当$a=-2$时，A的特征值为$2,2,6$，因矩阵

$$2E-A=\begin{pmatrix}1&-2&3\\1&-2&3\\-1&2&-3\end{pmatrix}$$

的秩为1，从而$\lambda=2$对应的线性无关的特征向量有2个，从而A可相似对角化.

(2) 当$\lambda=2$不是二重特征根时，则$\lambda^2-8\lambda+18+3a$ 为完全平方，从而$18+3a=16$，解得$a=-\frac{2}{3}$.

当$a=-\frac{2}{3}$时，A的特征值为$2,4,4$，因矩阵

$$4E-A=\begin{pmatrix} 3 & -2 & 3 \\ 1 & 0 & 3 \\ -1 & \dfrac{2}{3} & -3 \end{pmatrix}$$

的秩为2，从而$\lambda=4$对应的线性无关的特征向量只有1个，从而A不能相似对角化.

10. **证明：** 由题设，$(A-aE)(A-bE)=0$，两边取行列式知$|A-aE|=0$或$|A-bE|=0$. 以下我们分三种情况进行讨论：

(1) 若$|A-aE|=0$，但$|A-bE|\neq 0$. 此时，$A-bE$可逆，从而$A=aE$. 故A必可相似对角化.

(2) 若$|A-aE|\neq 0$，但$|A-bE|=0$. 此时，$A-aE$可逆，从而$A=bE$. 故A必可相似对角化.

(3) 若$|A-aE|=0$，且$|A-bE|=0$. 欲证A可相似对角化，只需证A有n个线性无关的特征向量即可.

首先，因为$(A-aE)(A-bE)=0$，利用矩阵秩的性质，$\mathrm{rank}(A-aE)+\mathrm{rank}(A-bE)\leq n$. 又

$$\mathrm{rank}(A-aE)+\mathrm{rank}(A-bE)=\mathrm{rank}(A-aE)+\mathrm{rank}(bE-A)$$

$$\geq\mathrm{rank}(A-aE+bE-A)=\mathrm{rank}((b-a)E)=n.$$

故$\mathrm{rank}(A-aE)+\mathrm{rank}(A-bE)=n$.

其次，因为$|A-aE|=0$，且$|A-bE|=0$，所以a与b是A的两个互异的特征值. 从属于a的线性无关的特征向量的个数由齐次线性方程组$(A-aE)X=0$的基础解系所含解向量的个数$n-\mathrm{rank}(A-aE)$决定，从属于b线性无关的特征向量的个数由齐次线性方程组$(A-bE)X=0$的基础解系所含解向量的个数$n-\mathrm{rank}(A-bE)$决定. 因为从属于不同特征值的特征向量也是线性无关的，因此A共含有

$$n-\mathrm{rank}(A-aE)+n-\mathrm{rank}(A-bE)=n.$$

个线性无关的特征向量. 故A可对角化.