同济大学

高等代数与解析几何

教师： 蒋志洪

理学院数学系

§9.5 埃尔米特(实对称)矩阵与酉(正交)相似标准形

定义. 设A和B是数域$\mathbb{F}$上的方阵, 如果存在酉阵(或实正交阵)U使得$U^{-1}AU = B$, 则称A和B为酉相似(或正交相似).

定义. 设复阵$H = (h_{ij})_{n\times n}$, 如果$\overline{H}^t = H$, 则称H为埃尔米特(Hermite)阵.

说明. 一个埃尔米特阵的对角线元素都是实数, 一个实的埃尔米特阵是实对称阵, 所以实对称阵是埃尔米特阵的一个特例.

定义. 设V是内积空间, $\mathbf{H}$是V上线性变换, 如果

$$(\mathbf{H}(\alpha), \beta) = (\alpha, \mathbf{H}(\beta)),\ \alpha, \beta \in V$$

则称$\mathbf{H}$为V上的埃尔米特变换(或对称变换).

命题. 设$\mathbf{H}$是内积空间V上线性变换, $e_1, e_2, \cdots, e_n$是V的一个标准正交基, $\mathbf{H}$在基$e_1, e_2, \cdots, e_n$下的矩阵为$H = (h_{ij})_{n\times n}$, 则$\mathbf{H}$是埃尔米特变换的充分必要条件H是埃尔米特阵.

证明. 由于$\mathbf{H}$在基$e_1, e_2, \cdots, e_n$下的矩阵为$H = (h_{ij})_{n\times n}$, 所以

$$\mathbf{H}(e_i) = \sum_{k=1}^{n} h_{ki}e_k \ (i = 1, 2, \cdots, n)$$

再由于$e_1, e_2, \cdots, e_n$是V的一个标准正交基, 所以

$$(\mathbf{H}(e_j), e_i) = (\sum_{k=1}^{n} h_{kj}e_k, e_i) = h_{ij} \quad (e_j, \mathbf{H}(e_i)) = (e_j, \sum_{k=1}^{n} h_{ki}e_k) = \overline{h}_{ji}$$

因此当$\mathbf{H}$是埃尔米特变换时, 则对所有$i, j = 1, 2, \cdots, n$, 都有$(\mathbf{H}(e_j), e_i) = (e_j, \mathbf{H}(e_i))$; 由此推出矩阵$H = (h_{ij})$满足$\overline{h}_{ji} = h_{ij}$, 即$\overline{H}^T = H$, 所以$H = (h_{ij}$是埃尔米特阵.

反之, 当$H = (h_{ij}$是埃尔米特阵时, 则$\overline{h}_{ji} = h_{ij}$, 由此推出对于所有$i, j = 1,2,\cdots,n$, 都有$(\mathbf{H}(e_j), e_i) = (e_j, \mathbf{H}(e_i))$, 对于一般向量$\alpha, \beta \in V$, 由于$e_1, e_2, \cdots, e_n$是$V$的一个标准正交基, 所以

$$\alpha = \sum_{i=1}^{n} x_i e_i, \ \ \beta = \sum_{i=1}^{n} y_i e_i$$

于是

$$\begin{aligned}
(\mathbf{H}(\alpha), \beta) &= (\sum_{i=1}^{n} x_i \mathbf{H}(e_i), \sum_{j=1}^{n} y_j e_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \overline{y}_j (\mathbf{H}(e_i), e_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \overline{y}_j (e_i, \mathbf{H}(e_j)) \\
&= (\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j \mathbf{H}(e_j)) \\
&= (\alpha, \mathbf{H}(\beta))
\end{aligned}$$

所以$(\mathbf{H}(\alpha), \beta) = (\alpha, \mathbf{H}(\beta))$. 为此, 即$\mathbf{H}$是埃尔米特变换.

命题. 设$\mathbf{H}$是V上的埃尔米特变换, 则对任何$v \in V$, $(\mathbf{H}(v), v)$必为实数.

证明. 由埃尔米特变换的定义推出$(\mathbf{H}(v), v) = (v, \mathbf{H}(v)$, 由内积定义推出$\overline{(\mathbf{H}(v), v)} = (v, \mathbf{H}(v)$, 所以$(\mathbf{H}(v), v) = \overline{(\mathbf{H}(v), v)}$ 因此$(\mathbf{H}(v), v)$是实数.

命题. 设$\mathbf{H}$是V上的埃尔米特变换, 则$\mathbf{H}$的特征值都是实数.

证明. 设λ是$\mathbf{H}$的一个特征值, α是相应的特征向量, 则

$$\overline{\lambda}(\alpha, \alpha) = (\alpha, \lambda\alpha) = (\alpha, \mathbf{H}(\alpha)) = (\mathbf{H}(\alpha), \alpha) = (\lambda\alpha, \alpha) = \lambda(\alpha, \alpha)$$

因为特征向量$\alpha \neq \mathbf{0}$, 所以$(\alpha, \alpha) \neq 0$, 由此推出$\lambda = \overline{\lambda}$, 即$\lambda$是实数.

命题. 设$\mathbf{H}$是内积空间V上的埃尔米特变换, 则属于不同特征值的特征向量正交.

证明. 设λ, μ是$\mathbf{H}$的不同特征值, α, β是相应的特征向量. 则

$$\lambda(\alpha, \beta) = (\lambda\alpha, \beta) = (\mathbf{H}(\alpha), \beta) = (\alpha, \mathbf{H}(\beta)) = (\alpha, \mu\beta) = \overline{\mu}(\alpha, \beta)$$

由于$\mathbf{H}$的特征值是实数, 所以$\overline{\mu} = \mu$. 又因为$\lambda \neq \mu$, 所以$(\alpha, \beta) = 0$.

命题. 设V_1是$\mathbf{H}$的不变子空间, 则$V_1^{\perp}$也是$\mathbf{H}$的不变子空间.

证明. 对于$u \in V_1^{\perp}, v \in V_1$, 由于$V_1$是$\mathbf{H}$ 的不变子空间, 所以$\mathbf{H}(v) \in V_1$, 因此

$$(\mathbf{H}(u), v) = (u, \mathbf{H}(v)) = 0$$

所以$\mathbf{H}(u) \in V_1^{\perp}$, 即$V_1^{\perp}$是$\mathbf{H}$ 的不变子空间.

练习. 习题9.5: 1,6.

定理. 设A复矩阵, 必存在酉阵U, 使得;

$$U^{-1}AU=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}=T.$$

其中T是上三角阵, 对角线元素$\lambda_1,\lambda_2,\cdots,\lambda_n$恰是$A$的全部特征值.

证明. 对A的阶数n用归纳法. 当$n=1$时, A是复数, 定理当然成立.

归纳假设定理对$(n-1)$阶复阵成立. 现在设A是n阶复阵. 在复数域$\mathbb{C}$中, A总有特征值, 设为λ_1, 相应的特征向量设为α_1. 于是在$\mathbb{C}^n$中, 由α_1出发, 可以获得一组标准正交基$u_1,u_2,\cdots,u_n$, 其中$u_1=\frac{\alpha_1}{\|\alpha_1\|}$, 且$Au_1=\lambda_1u_1$. 设$Au_i=\sum\limits_{k=1}^{n}t_{ki}u_k(i=2,3,\cdots,n)$, 写成矩阵即得

$$A(u_1,u_2,\cdots,u_n)=(u_1,u_2,\cdots,u_n)\begin{pmatrix}\lambda_1 & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & t_{n2} & \cdots & t_{nn}\end{pmatrix}.$$

记$U_1=(u_1,u_2,\cdots,u_n)$, 它是酉阵, 而且

$$U_1^{-1}AU_1=\begin{pmatrix}\lambda_1 & t\\ 0 & B\end{pmatrix}$$

其中B是$(n-1)$阶复阵, $t=(t_{12},\cdots,t_{1n})$.
用归纳法假设, 存在$(n-1)$阶酉阵U_2, 使得

$$U_2^{-1}BU_2=\begin{pmatrix}\lambda_2 & * & \cdots & *\\ 0 & \lambda_3 & \cdots & *\\ \vdots & \vdots & \cdots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}$$

令

$$U_3=\begin{pmatrix}1 & 0\\ 0 & U_2\end{pmatrix},\ U=U_1U_3$$

由于U_1,U_3是n阶酉阵, 所以U是n阶酉阵.

$$U^{-1}AU = U_3^{-1}U_1^{-1}AU_1U_3 = U_3^{-1}\begin{pmatrix} \lambda_1 & t \\ 0 & B \end{pmatrix}U_3$$
$$= \begin{pmatrix} \lambda_1 & * \\ 0 & U_2^{-1}BU_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

所以A与上三角阵T酉相似, 由于A和T有相同的特征值, 所以T的主对角线元素皆为A的特征值.

练习. 9.5: 3.

命题. 设H是埃尔米特阵, U是酉阵, 则$U^{-1}HU$是埃尔米特阵.

证明. 由于

$$\overline{U^{-1}HU}^T = \overline{\overline{U}^T HU}^T = \overline{U}^T\overline{H}^T(U^T)^T = U^{-1}HU$$

所以$U^{-1}HU$是埃尔米特阵.

定理. 一个复阵H是埃尔米特阵的充分必要条件H的特征值都是实数并且H有n个特征向量$u_1, u_2, \cdots, u_n$ 使得$u_1, u_2, \cdots, u_n$构成$\mathbb{C}^n$的标准正交基.

证明. (必要性) 由于H酉相似于一个上三角阵, 所以存在酉阵U, 使得$U^{-1}HU = T$是上三角阵, 由于H为埃尔米特阵, 所以T也是埃尔米特阵, 即$\overline{T}^t = T$, 由此推出T是对角阵, 而且对角线上的元素都是实数.

如果记$U = (u_1, u_2, \cdots, u_n)$和$T = \mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}$, 由于$U$是酉阵, 所以它的列向量组$u_1, u_2, \cdots, u_n$构成$\mathbb{C}^n$ 的标准正交基, 由$HU = UT$ 推出

$$Hu_i = \lambda_i u_i \quad (i = 1, 2, \cdots, n).$$

(充分性)设$u_1, u_2, \cdots, u_n$是H的特征向量而且构成$\mathbb{C}^n$的标准正交基, 则$U = (u_1, \cdots, u_n)$是酉阵, 由于H有n个线性无关的特征向量, 所以H相似对角阵, 即$U^{-1}HU = T$是一个对角阵并且对角线上元素都是H的特征值(都是实数), 所以T是一个实对称矩阵, 因此T是一个埃尔米特阵, 由于$H = UTU^{-1}$, 所以H是一个埃尔米特阵.

推论. 如果A是实对称阵, 则存在正交阵O使得

$$O^{-1}AO = \mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}$$

其中$\lambda_1, \cdots, \lambda_n$是$A$的特征值.

- 对于实对称矩阵A, 求正交矩阵O和对角阵Λ使得$O^{-1}AO = \Lambda$的方法:

(1) 求出A的所有不同特征值$\lambda_1, \cdots, \lambda_s$. 即求出$A$的特征多项式

$$\Delta_A(\lambda) = (\lambda - \lambda_1)^{d_1}(\lambda - \lambda_2)^{d_2} \cdots (\lambda - \lambda_s)^{d_s}$$

的根.

(2) 对于每个$\lambda_i (i = 1, 2, \cdots, s)$, 求出它的特征子空间

$$V_{\lambda_i} = \{X \in \mathbb{R}^n | (\lambda_i E - A)X = 0\}$$

的基$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_{d_i}}$. 即求线性方程组$(\lambda_i E - A)X = 0$的基础解系.

(3) 对于$i = 1, 2, \cdots, s$, 将$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_{d_i}}$正交化, 单位化得$V_{\lambda_i}$中标准正交基$u_{i_1}, \cdots, u_{i_{d_i}}$.

(4) 对于$i = 1, 2, \cdots, s$, 将它们合并起来写出$O = (u_1, u_2, \cdots, u_n)$和$\Lambda = \mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}$

例. 设

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

求正交矩阵O, 使得$O^{-1}AO$为对角阵, 并求出此对角阵.

解. (1) A的特征多项式

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

所以A的特征值为$\lambda=1,-3$

(2) 求特征值$\lambda_1 = 1$的特征子空间V_1的基, 解齐次方程组$(1E - A)X = 0$, 即

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

它的基础解系就是V_1的基

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

求特征值$\lambda_2 = -3$ 的特征子空间V_{-3}的基, 解齐次方程组$(-3E - A)X = 0$, 即

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

它的基础解系就是V_{-3}的基

$$\alpha_4 = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

(3) 将V_1中基向量$\alpha_1, \alpha_2, \alpha_3$正交化, 得

$$\beta_1 = \alpha_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \beta_2 = \alpha_2 - \frac{\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1 = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}$$

$$\beta_3 = \alpha_3 - \frac{(\alpha_3, \beta_1)}{\beta_1, \beta_1)}\beta_1 - \frac{(\alpha_3, \beta_2)}{(\beta_2, \beta_2)}\beta_2 = \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ 1 \end{pmatrix}$$

将$\beta_1, \beta_2, \beta_3$单位化

$$u_1 = \frac{\beta_1}{\|\beta_1\|} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}, \quad u_2 = \frac{\beta_2}{\|\beta_2\|} = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ 0 \end{pmatrix}, \quad u_3 = \frac{\beta_3}{\|\beta_3\|} = \begin{pmatrix} -\frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{12}} \\ \frac{3}{\sqrt{12}} \end{pmatrix}$$

将V_{-3}中基向量α_4正交化:

$$\beta_4 = \alpha_4$$

将β_4单位化

$$u_4 = \frac{\beta_4}{\|\beta_4\|} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

(4) 写出正交阵

$$O=(u_1,u_2,u_3,u_4)=\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{12}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} & -\frac{1}{2} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{12}} & -\frac{1}{2} \\ 0 & 0 & \frac{3}{\sqrt{12}} & \frac{1}{2} \end{pmatrix}$$

和对角阵

$$O^{-1}AO=\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -3 \end{pmatrix}$$

说明. 对于实对称阵A, 存在正交阵O使得$O^{-1}AO$为对角阵, 事实上我们可要求$|O|=1$. 因为如果$|O|=-1$, 则可取对角阵$S=\mathrm{diag}\{-1,1,1,\cdots,1\}$. 它显然是正交阵. 于是取$Q=OS$. 作为正交阵的积, 它仍是正交阵, 但是$|Q|=|O||S|=1$, 并且$Q^{-1}AQ=O^{-1}AO$.

练习. 9.5: 2(2),4,5.

- 问题: 一个复矩阵酉相似于对角阵的条件

定义. n阶复阵N, 如果满足$\overline{N}^t N = N\overline{N}^t$, 则称$N$为正规阵.

例. 埃尔米特阵(或实对称阵)H, 酉阵(或实正交阵)U, 斜埃尔米特阵K (满足条件:$\overline{K}^t = -K$)都是正规阵

解. 请同学们自己验证.

例. $N = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ 是正规阵.

解.

$$\overline{N}^t N = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = N\overline{N}^t$$

注意. N不是上述三种类型的矩阵.

定理. 复阵A酉相似于对角阵的充分必要条件是A为正规阵.

证明. 下面是这个定理证明的步骤, 请同学们给出详细证明.

(1) 如果U是酉阵, A是正规阵, 则$U^{-1}AU$是正规阵

(2) 如果A是复数矩阵, 则A酉相似上三角阵

(3) 如果一个上三角阵R是正规阵, 则R 是对角阵.