矩阵理论

王磊

自动化科学与电气工程学院

第三章 矩阵分解

Jordan 分 解

第三章 矩阵分解

更一般的问题:线性变换T在什么样的基下的矩阵最简单？换而言之,在相似变换下矩阵A能有怎么样的最简形式？

定义3.8.1（λ矩阵） 以λ多项式为元素的矩阵称为**λ矩阵**，记为$A(\lambda)$，即$A(\lambda)=\left[a_{ij}(\lambda)\right]_{m\times n}$，$a_{ij}(\lambda)\in P_n(\lambda)$.

例3.8.1 判断$A(\lambda)$和$B(\lambda)$是否为λ矩阵，其中

$$A(\lambda)=\begin{bmatrix}1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2\end{bmatrix}$$

$$B(\lambda)=\begin{bmatrix}\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ \lambda^{-2} & \lambda^2 & -\lambda^2\end{bmatrix}$$

例3.8.1 判断$A(\lambda)$和$B(\lambda)$是否为λ矩阵, 其中

$$A(\lambda)=\begin{bmatrix} 1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2 \end{bmatrix}$$

$$B(\lambda)=\begin{bmatrix} \lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ \lambda^{-2} & \lambda^2 & -\lambda^2 \end{bmatrix}$$

注1: 数字矩阵是特殊的λ矩阵；复方阵A的特征矩阵$\lambda I-A$是λ矩阵.

注2: λ矩阵和数字矩阵一样有加、减、乘等运算且具有相同的运算规律. 我们同样可定义正方λ矩阵的行列式、子式及λ矩阵的秩等.

注2: λ矩阵和数字矩阵一样有加、减、乘等运算且具有相同的运算规律. 我们同样可定义正方λ矩阵的行列式、子式及λ矩阵的秩等.

定义3.8.2（λ矩阵的秩）λ矩阵$A(\lambda)$中非零子式的最高阶数r 定义为$A(\lambda)$的**秩**, 记为$\mathrm{rank}\big(A(\lambda)\big) = r$.

例3.8.2 求$A(\lambda)=\begin{bmatrix}\lambda & 0\\ 0 & \lambda+1\end{bmatrix}$的行列式和秩.

例3.8.3 设$A\in\mathbb{C}^{n\times n}$，则$f_A(\lambda)=|\lambda I-A|$是关于$\lambda$的一元$n$次多项式. 因此, A的特征矩阵$\lambda I-A$的秩为n, 即$\lambda I-A$总是满秩的.

定义3.8.3（λ矩阵的逆矩阵）设$A(\lambda)$是n阶λ方阵，若存在n阶λ方阵$B(\lambda)$满足$A(\lambda)B(\lambda)=B(\lambda)A(\lambda)=I$，则称$\lambda$矩阵$A(\lambda)$是**可逆**的，并称$B(\lambda)$为$A(\lambda)$的**逆矩阵**，记作$A(\lambda)^{-1}$.

【思考】：在数字矩阵中，满秩矩阵和可逆矩阵是等价命题. 这一结论适用于λ矩阵吗？

定理3.8.1 n阶λ方阵$A(\lambda)$可逆的充分必要条件是它的行列式$|A(\lambda)|$为非零常数.

定义3.8.4（初等变换）下列三种变换称为λ矩阵的**初等变换**:

（1）λ矩阵的两行（列）互换位置；

（2）λ矩阵的某一行（列）乘以非零常数k；

（3）λ矩阵的某一行（列）的$\varphi(\lambda)$倍加到另一行（列），其中$\varphi(\lambda) \in P_n(\lambda)$.

定义: λ矩阵初等变换指以下三类变换:

1) 任两行/列互换 ($P(i,j)$);

$$
P(i,j)=\begin{array}{c}
\left[\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 0 & & \cdots & & 1 & \cdots & & \cdots \\
 & & & & 1 & & & & & & \\
 & & & \vdots & & \ddots & & \vdots & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & \cdots & & 0 & \cdots & & \cdots \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right]
\end{array}
\begin{array}{l}
\\ \\ \\ \text{第i行} \\ \\ \\ \\ \text{第j行} \\ \\ \\ \\
\end{array}
$$

定义: λ矩阵初等变换指以下三类变换:

2) 用不为零的数k乘某行/列($P(i(k))$);

$$P(i(k))=\begin{array}{cl}\begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & k & & \cdots & & \cdots & \cdots \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \text{第i行} \\ \\ \\ \\ \\ \\ \end{array}\end{array}$$

定义:λ矩阵初等变换指以下三类变换:

3) 用λ的多项式$\varphi(\lambda)$乘j行/列并加到第i行/列上去（$P(i,j(\varphi))$）

$$P(i,j(\varphi))=\left[\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 1 & & \cdots & & \varphi(\lambda) & \cdots & & \cdots \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & \vdots & & & \\
 & & & & & & 1 & & & & \\
 & & & & & & & 1 & \cdots & & \cdots \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right]\begin{array}{l}
 \\
 \\
 \\
\text{第}i\text{行} \\
 \\
 \\
 \\
\text{第}j\text{行} \\
 \\
 \\

\end{array}$$

定义3.8.5（λ矩阵相抵）若λ矩阵$A(\lambda)$经过有限次初等变换化为$B(\lambda)$，则称$A(\lambda)$与$B(\lambda)$**相抵**，记为$A(\lambda) \cong B(\lambda)$.

注3: λ矩阵相抵则其秩相同, 反之则不然, 这**与数字矩阵是有区别的**.

例3.8.4 考查$A(\lambda)$和$B(\lambda)$是否相抵, 其中

$$A(\lambda)=\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}, B(\lambda)=\begin{bmatrix}1 & -\lambda\\ 1 & \lambda\end{bmatrix}$$

定义3.8.6（行列式因子）设λ矩阵$A(\lambda)$的秩为r, 对于正整数$1 \leq k \leq r$, $A(\lambda)$的全部k阶子式的首1最大公因式称为k阶行列式因子, 记为$D_k(\lambda)$.

例3.8.5 求$A(\lambda)=\begin{bmatrix} 1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2 \end{bmatrix}$各阶行列式因子.

例3.8.5 求$A(\lambda)=\begin{bmatrix}1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2\end{bmatrix}$各阶行列式因子.

若将$A(\lambda)$进行相抵变换得

$$\begin{bmatrix}1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2\end{bmatrix}\to\begin{bmatrix}1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda+\lambda^2\end{bmatrix}$$

比较变换前后的行列式因子:

定理3.8.2 相抵的λ矩阵具有相同的秩和相同的各阶行列式因子.

定理3.8.3（Smith标准形） 设λ矩阵$A(\lambda)$的秩为r，则

$$A(\lambda) \cong \begin{bmatrix} d_1(\lambda) & & & & & \\ & \ddots & & & & \\ & & d_r(\lambda) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

式中，$d_i(\lambda)$是首1多项式，且$d_i(\lambda)|d_{i+1}(\lambda)$，称此标准形为$A(\lambda)$的**Smith标准形**.

例3.8.6 设$A(\lambda)=\begin{bmatrix}1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2\end{bmatrix}$, 求其Smith标准形.

Smith标准型与行列式因子之间的关系：

$$
\begin{aligned}
D_1(\lambda) &= d_1(\lambda) \\
D_2(\lambda) &= d_1(\lambda)d_2(\lambda) \\
&\vdots \\
D_r(\lambda) &= d_1(\lambda)d_2(\lambda)\cdot\cdots\cdot d_r(\lambda)
\end{aligned}
$$

或

$$
\begin{aligned}
d_1(\lambda) &= D_1(\lambda) \\
d_2(\lambda) &= D_2(\lambda)/D_1(\lambda) \\
&\vdots \\
d_r(\lambda) &= D_r(\lambda)/D_{r-1}(\lambda)
\end{aligned}
$$

推论3.8.1 λ矩阵的Smith标准形是唯一的.

定义3.8.7（不变因子） 在λ矩阵$A(\lambda)$的Smith标准形中，$d_1(\lambda),\cdots,d_r(\lambda)$由$A(\lambda)$唯一确定的，称为$A(\lambda)$的**不变因子**.

例3.8.7 求下列λ矩阵的不变因子, 其中

$$A(\lambda)=\begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}, B(\lambda)=\begin{bmatrix}\lambda & 1\\ 0 & \lambda\end{bmatrix}$$

推论3.8.2 λ矩阵$A(\lambda)$与$B(\lambda)$相抵当且仅当它们有完全一致的不变因子.

注4: 初等变换不改变λ矩阵的不变因子.

再考察例3.8.5:

$$A(\lambda)=\begin{bmatrix}1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2\end{bmatrix}\cong\begin{bmatrix}1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda(\lambda+1)\end{bmatrix}$$

$$d_1(\lambda)=1$$
$$d_2(\lambda)=\lambda$$
$$d_3(\lambda)=\lambda(\lambda+1)$$

再考察例3.8.5:

$$A(\lambda)=\begin{bmatrix}1-\lambda & \lambda^2 & \lambda \\ \lambda & \lambda & -\lambda \\ 1+\lambda^2 & \lambda^2 & -\lambda^2\end{bmatrix}\cong\begin{bmatrix}1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda(\lambda+1)\end{bmatrix}$$

$$d_1(\lambda)=1=\lambda^0(\lambda+1)^0$$
$$d_2(\lambda)=\lambda=\lambda^1(\lambda+1)^0$$
$$d_3(\lambda)=\lambda(\lambda+1)=\lambda^1(\lambda+1)^1$$

定义3.8.8（初等因子） 设λ矩阵$A(\lambda)$的不变因子为$d_1(\lambda),\cdots,d_r(\lambda)$, 且有分解式

$$\begin{cases} d_1(\lambda)=(\lambda-\lambda_1)^{e_{11}}(\lambda-\lambda_2)^{e_{12}}\cdot\cdots\cdot(\lambda-\lambda_s)^{e_{1s}} \\ d_2(\lambda)=(\lambda-\lambda_1)^{e_{21}}(\lambda-\lambda_2)^{e_{22}}\cdot\cdots\cdot(\lambda-\lambda_s)^{e_{2s}} \\ \vdots \\ d_r(\lambda)=(\lambda-\lambda_1)^{e_{r1}}(\lambda-\lambda_2)^{e_{r2}}\cdot\cdots\cdot(\lambda-\lambda_s)^{e_{rs}} \end{cases}$$

则所有幂指数大于零的因子$(\lambda-\lambda_j)^{e_{ij}}$, $i=1,\cdots,r$, $j=1,\cdots,s$, 统称为λ矩阵$A(\lambda)$的**初等因子**（组）.

例 3.8.8 设 $A(\lambda) = \mathrm{diag}(\lambda(\lambda+1), \lambda^2, (\lambda+1)^2, 0)$. 求$A(\lambda)$的初等因子、不变因子和Smith标准形.

注5: 若$A(\lambda) \cong B(\lambda)$，则$A(\lambda)$与$B(\lambda)$有完全一致的行列式因子, 进而有相同的不变因子. 所以，$A(\lambda)$与$B(\lambda)$有完全一致的初等因子，这表明初等变换不改变$A(\lambda)$的初等因子.

【思考】： 若$A(\lambda)$与$B(\lambda)$有完全一致的初等因子, 则$A(\lambda) \cong B(\lambda)$成立吗?

例3.8.9 考查$A(\lambda)$与$B(\lambda)$是否相抵, 其中

$$A(\lambda)=\begin{bmatrix}1 & 0 & 0\\ 0 & (\lambda-2)(\lambda-3) & 0\\ 0 & 0 & 0\end{bmatrix}$$

$$B(\lambda)=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & (\lambda-2)(\lambda-3)\end{bmatrix}$$

定理3.8.4 λ矩阵$A(\lambda) \cong B(\lambda)$当且仅当它们有完全相同的初等因子, 且$\operatorname{rank}\big(A(\lambda)\big) = \operatorname{rank}\big(B(\lambda)\big)$.

定理3.8.5 设λ矩阵$A(\lambda)$为对角矩阵, 即

$$A(\lambda) = \operatorname{diag}(A_1(\lambda), \cdots, A_s(\lambda))$$

则$A_1(\lambda), \cdots, A_s(\lambda)$初等因子的全体就是$A(\lambda)$的全部初等因子, 其中$A_i(\lambda)$, $i = 1, \cdots, s$是适当阶数的λ矩阵.

例3.8.10 求特征矩阵$\lambda I - A$的Smith标准形, 其中

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

例3.8.11 求

$$A(\lambda)=\begin{bmatrix}\lambda-a & -1 & 0 & \cdots & 0\\ 0 & \lambda-a & -1 & \cdots & 0\\ 0 & 0 & \lambda-a & \ddots & \vdots\\ \vdots & \vdots & \vdots & \ddots & -1\\ 0 & 0 & 0 & \cdots & \lambda-a\end{bmatrix}_{n\times n}$$

的初等因子、不变因子和Smith标准形.

定理3.8.6 复方阵A和B相似当且仅当它们的特征矩阵相抵.

定理3.8.6 复方阵A和B相似当且仅当它们的特征矩阵相抵.

综上, 对于方阵$A, r(\lambda I - A) = n$, 有

$$A \sim B \iff (\lambda I - A) \cong (\lambda I - B)$$

$$\iff \lambda I - A\text{与}\lambda I - B\text{有相同的不变因子}$$

$$\iff \lambda I - A\text{与}\lambda I - B\text{有相同的初等因子}$$

思考:方阵A可对角化的充要条件?

思考：方阵A可对角化的充要条件？

推论3.8.3 复方阵A是单纯矩阵的充分必要条件是它的特征矩阵$\lambda I - A$的初等因子为一次的.

推论3.8.4 复方阵A是单纯矩阵的充分必要条件是它的特征矩阵$\lambda I - A$的不变因子无重根.

例3.8.12 判断矩阵A是否为单纯矩阵，其中

$$A=\begin{bmatrix} a & 1 & & & \\ & a & 1 & & \\ & & a & \ddots & \\ & & & \ddots & 1 \\ & & & & a \end{bmatrix}_{n\times n}$$

定义3.8.9（Jordan块） 设$A=[a_{ij}]\in\mathbb{C}^{n\times n}$，其特征矩阵$\lambda I-A$的初等因子为$(\lambda-\lambda_1)^{n_1}$，$(\lambda-\lambda_2)^{n_2}$，$\cdots$, $(\lambda-\lambda_t)^{n_s}$. 对$(\lambda-\lambda_i)^{n_i}$作$n_i$阶矩阵

$$J_i=\begin{bmatrix}\lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i\end{bmatrix}_{n_i\times n_i}$$

则称矩阵J_i（$i=1,\cdots,s$）为A的**Jordan块**.

例3.8.13 判断下列矩阵是否为Jordan块, 其中

$$A_1=\begin{bmatrix} i+1 & 1 & \\ & i+1 & 1 \\ & & i+1 \end{bmatrix}, A_2=\begin{bmatrix} 1 & 1 & \\ & 1 & 0 \\ & & 1 \end{bmatrix}$$

$$A_3=\begin{bmatrix} 3 & & & \\ & 3 & 1 & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}, A_4=\begin{bmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}$$

例3.8.14 求Jordan块$J_i=\begin{bmatrix}\lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{bmatrix}_{n_i\times n_i}$的最小多项式.

定义3.8.10（Jordan标准形） 设$A=[a_{ij}]\in\mathbb{C}^{n\times n}$，其特征矩阵$\lambda I-A$的初等因子为$(\lambda-\lambda_1)^{n_1}$，$(\lambda-\lambda_2)^{n_2}$，$\cdots$，$(\lambda-\lambda_t)^{n_s}$. 其对应的Jordan块分别记为$J_1,\cdots,J_s$，则由$s$个Jordan块作$n$阶对角块矩阵$J=diag(J_1,\cdots,J_s)$称为$A$的**Jordan标准形**（或**Jordan法式**）.

例3.8.15 设

$$(\lambda I - A) \cong \mathrm{diag}(\lambda(\lambda+1), \lambda^2, (\lambda+1)^2, 1,1,1)$$

求A的Jordan标准形.

定理3.8.7 （Jordan标准形定理）设矩阵J是复方阵A的Jordan标准形, 则矩阵A与矩阵J相似.

例3.8.16 求矩阵A的Jordan标准形J，并求P使得$P^{-1}AP=J$, 其中

$$A=\begin{bmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$$

例3.8.16 求矩阵A的Jordan标准形J，并求P使得$P^{-1}AP = J$, 其中

$$A = \begin{bmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{bmatrix}$$

例3.8.16 求矩阵A的Jordan标准形J，并求P使得$P^{-1}AP=J$, 其中

$$A=\begin{bmatrix}-1 & -2 & 6\\ -1 & 0 & 3\\ -1 & -1 & 4\end{bmatrix}$$

解: $\lambda I-A\cong\begin{bmatrix}1 & 0 & 0\\ 0 & \lambda-1 & 0\\ 0 & 0 & (\lambda-1)^2\end{bmatrix}$

例3.8.16

设$P=[\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3]$,则$A[\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3]=[\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3]J$. 整理得

$$\begin{cases} A\boldsymbol{p}_1=\boldsymbol{p}_1 \\ A\boldsymbol{p}_2=\boldsymbol{p}_2 \\ A\boldsymbol{p}_3=\boldsymbol{p}_2+\boldsymbol{p}_3 \end{cases}$$

由$A\boldsymbol{p}_i=\boldsymbol{p}_i$解得两个线性无关的向量为

$$\boldsymbol{p}_1=[3,0,1]^T和\boldsymbol{p}_2=[0,3,1]^T.$$

例3.8.17 求复方阵A的最小多项式（利用初等因子）.

定理3.8.7（Frobenious定理）设$A \in \mathbb{C}^{n\times n}$，其特征矩阵$\lambda I - A$的Smith标准形为$\mathrm{diag}(d_1(\lambda),\ldots,d_n(\lambda))$，则$A$的最小多项式$m_A(\lambda) = d_n(\lambda)$.

例3.8.18 求以下矩阵的最小多项式, 其中

$$A_1=\begin{bmatrix}3&&&\\&3&1&\\&&3&1\\&&&3\end{bmatrix},A_2=\begin{bmatrix}2&1&&\\&2&0&\\&&3&1\\&&&3\end{bmatrix}$$

第三章 矩阵分解

奇 异 值 分 解

引理3.9.1 设$A \in \mathbb{C}^{m\times n}$, 则

$$\operatorname{rank}(A^H A) = \operatorname{rank}(AA^H) = \operatorname{rank}(A).$$

注1: $N(A^H A) = N(A)$; $N(AA^H) = N(A^H)$.

引理3.9.2 设$A \in \mathbb{C}^{m\times n}$，则$A^H A$与$AA^H$都是半正定Hermite矩阵.

注2: 矩阵A是半正定Hermite矩阵还有以下命题等价:

（1）对于任意n阶可逆矩阵P，$P^H AP$是半正定Hermite矩阵；

（2）Hermite矩阵A的n个特征值均为非负数；

（3）存在n阶可逆矩阵P使得$P^H AP=\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$，其中$r=\operatorname{rank}(A)$；

（4）存在秩为$\operatorname{rank}(A)$的矩阵Q使得$A=Q^H Q$；

（5）存在n阶Hermite矩阵G使得$A=G^2$；

（6）Hermite矩阵A的所有顺序主子式均非负.

引理3.9.3 设$A \in \mathbb{C}_r^{m\times n}$, 则$A^H A$与$AA^H$的所有非零特征值完全相同且非零特征值的个数均为$r$.

例3.9.1 设$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, 计算$A^H A$和$AA^H$的特征值.

例3.9.1 设$A=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}$, 计算$A^HA$和$AA^H$的特征值.

$$A^HA=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix}, AA^H=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}.$$

注2: 在计算A^HA和AA^H的特征值时，可优先考虑计算阶数较小的矩阵.

定义3.9.1（奇异值）设$A \in \mathbb{C}_r^{m\times n}$，$A^H A$的特征值满足$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$，$\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$，称$\sigma_i = \sqrt{\lambda_i}$为$A$的**奇异值**．特别地，称$\sigma_i, i = 1, \cdots, r,$为$A$的**正奇异值**.

例3.9.2 计算$A=\begin{bmatrix}1 & 2\\ 0 & 1\\ \mathrm{i} & 0\end{bmatrix}$的正奇异值.

例3.9.2 计算$A=\begin{bmatrix}1 & 2\\ 0 & 1\\ \mathrm{i} & 0\end{bmatrix}$的正奇异值.

$$A^HA=\begin{bmatrix}1 & 0 & -\mathrm{i}\\ 2 & 1 & 0\end{bmatrix}\begin{bmatrix}1 & 2\\ 0 & 1\\ \mathrm{i} & 0\end{bmatrix}=\begin{bmatrix}2 & 2\\ 2 & 5\end{bmatrix}.$$

$$|\lambda I-A^HA|=\begin{vmatrix}\lambda-2 & -2\\ -2 & \lambda-5\end{vmatrix}=\lambda^2-7\lambda+6.$$

求得A^HA的特征值:$\lambda_1=1,\lambda_2=6.$

由此得到A的奇异值是$\sigma_1=\sqrt{\lambda_1}=1,\sigma_2=\sqrt{\lambda_2}=\sqrt{6}.$

【思考】：复方阵的特征值与奇异值有何异同？

【思考】：复方阵的特征值与奇异值有何异同？

例如：$A=\begin{bmatrix}0 & 0\\ 0 & 2\end{bmatrix}$, $B=\begin{bmatrix}0 & k\\ 0 & 1\end{bmatrix}$, $C=\begin{bmatrix}1 & k\\ 0 & 1\end{bmatrix}$, 其奇异值分别是什么？

【思考】：复方阵的特征值与奇异值有何异同？

例如：$A=\begin{bmatrix}0 & 0\\ 0 & 2\end{bmatrix}$，$B=\begin{bmatrix}0 & k\\ 0 & 1\end{bmatrix}$，$C=\begin{bmatrix}1 & k\\ 0 & 1\end{bmatrix}$，其奇异值分别是什么？

A的奇异值为0, 2;

B的奇异值为$\sigma_1=\sqrt{k^2+1}$, $\sigma_2=0$;

C的奇异值为

$$\sigma_1=\sqrt{\frac{1}{2}\left(k^2+2+k\sqrt{k^2+4}\right)},\ \sigma_2=\frac{1}{\sigma_1}.$$

【思考】: 复方阵的特征值与奇异值有何异同?

命题3.9.1 设$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$和$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$分别是$n$阶正规矩阵$A$的奇异值和特征值，则$\sigma_i = |\lambda_i|, i = 1, \cdots, n$.

命题3.9.2 设σ_i和$\lambda_i (i = 1, \cdots, n)$分别是$n$阶复方阵$A$的奇异值和特征值, 则

$$\sum_{i=1}^{n} |\lambda_i|^2 \leq \sum_{i=1}^{n} |\sigma_i|^2$$

定理3.9.1（奇异值分解）设$A \in \mathbb{C}_r^{m\times n}$，则存在酉矩阵$U \in \mathbb{C}^{m\times m}$与酉矩阵$V \in \mathbb{C}^{n\times n}$使得

$$A = U \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} V^H$$

式中，$\Sigma_r = \mathrm{diag}(\sigma_1, \cdots, \sigma_r)$，$\sigma_i (i = 1, \cdots, r)$是矩阵$A$的正奇异值.

注3: 若$A = UDV^H$是奇异值分解, 则酉矩阵U的列向量为AA^H的特征向量，酉矩阵V的列向量为A^HA的特征向量.

例3.9.3 求矩阵$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}$的奇异值分解.

注3: 若$A = UDV^H$是奇异值分解, 则酉矩阵U的列向量为AA^H的特征向量，酉矩阵V的列向量为A^HA的特征向量.

例3.9.3 求矩阵$A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}$的奇异值分解.

解: 首先有

$$A^HA = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}.$$

解：$|\lambda I - A^H A| = \begin{vmatrix} \lambda - 5 & -2 \\ -2 & \lambda - 5 \end{vmatrix} = (\lambda - 7)(\lambda - 3).$

由此得到$A^H A$的特征值：$\lambda_1 = 7, \lambda_2 = 3.$

求得$A^H A$关于特征值$\lambda_1 = 7$的单位特征向量：

$$v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

求得$A^H A$关于特征值$\lambda_2 = 3$的单位特征向量：

$$v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

令$V=V_1=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}, \Sigma=\begin{bmatrix}\sqrt{7} & 0\\ 0 & \sqrt{3}\end{bmatrix}$,

$$令U_1=AV_1\Sigma^{-1}=\begin{bmatrix}\frac{3}{\sqrt{14}} & \frac{1}{\sqrt{6}}\\ \frac{2}{\sqrt{14}} & -\frac{2}{\sqrt{6}}\\ \frac{1}{\sqrt{14}} & \frac{1}{\sqrt{6}}\end{bmatrix}=[u_1 \quad u_2].$$

将U_1中列向量u_1, u_2扩张成$\mathbb{R}^3$中的标准正交基, 设 $u_3 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$. 由$u_1, u_2, u_3$相互正交, 且$u_3$是单位向量, 得到

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 0 \\ x_1 - 2x_2 + x_3 = 0 \\ |x_1|^2 + |x_2|^2 + |x_3|^2 = 1 \end{cases} \Rightarrow u_3 = \frac{1}{\sqrt{21}} \begin{bmatrix} 2 \\ -1 \\ -4 \end{bmatrix}.$$

由此得到酉矩阵U：

$$U=\begin{bmatrix}\frac{3}{\sqrt{14}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{21}}\\ \frac{2}{\sqrt{14}} & -\frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{21}}\\ \frac{1}{\sqrt{14}} & \frac{1}{\sqrt{6}} & -\frac{4}{\sqrt{21}}\end{bmatrix}.$$

$$A=\begin{bmatrix}\frac{3}{\sqrt{14}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{21}}\\ \frac{2}{\sqrt{14}} & -\frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{21}}\\ \frac{1}{\sqrt{14}} & \frac{1}{\sqrt{6}} & -\frac{4}{\sqrt{21}}\end{bmatrix}\begin{bmatrix}\sqrt{7} & 0\\ 0 & \sqrt{3}\\ 0 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{bmatrix}.$$

注4: 读者也可先尝试构造酉矩阵V，然后再依据V构造酉矩阵U，由此给出奇异值分解证明和相应的分解步骤.

例3.9.4 求$A=\begin{bmatrix}1 & 0 & 0 & -1\\ 0 & 1 & 0 & 1\end{bmatrix}$的奇异值分解.

注4: 读者也可先尝试构造酉矩阵V，然后再依据V构造酉矩阵U，由此给出奇异值分解证明和相应的分解步骤.

例3.9.4 求$A=\begin{bmatrix}1 & 0 & 0 & -1\\ 0 & 1 & 0 & 1\end{bmatrix}$的奇异值分解.

解:首先有

$$AA^H=\begin{bmatrix}1 & 0 & 0 & -1\\ 0 & 1 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 1\\ 0 & 0\\ -1 & 1\end{bmatrix}=\begin{bmatrix}2 & -1\\ -1 & 2\end{bmatrix}$$

由此得到AA^H的特征值:$\lambda_1=3,\lambda_2=1.$

求得AA^H关于特征值$\lambda_1=3$的单位特征向量:

$$u_1=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}.$$

求得AA^H关于特征值$\lambda_2=1$的单位特征向量:

$$u_2=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}.$$

令$U=[u_1\quad u_2]=\frac{1}{\sqrt{2}}\begin{bmatrix}1&1\\-1&1\end{bmatrix}.$

因为A^HA与AA^H有相同的正特征值，所以A^HA的所有特征值是$\lambda_1=3,\lambda_2=1,\lambda_3=\lambda_4=0.$

进而求得

$$V=[v_1 \quad v_2 \quad v_3 \quad v_4]=\begin{bmatrix} \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{3}} \\ 0 & 0 & 1 & 0 \\ \frac{2}{\sqrt{6}} & 0 & 0 & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

进而求得

$$V=\begin{bmatrix}v_1 & v_2 & v_3 & v_4\end{bmatrix}=\begin{bmatrix}\frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{3}}\\ 0 & 0 & 1 & 0\\ \frac{2}{\sqrt{6}} & 0 & 0 & \frac{1}{\sqrt{3}}\end{bmatrix}.$$

$$U\begin{bmatrix}\Sigma & O\\ O & O\end{bmatrix}V^H=\begin{bmatrix}0 & 1 & 0 & 1\\ 1 & 0 & 0 & -1\end{bmatrix}\neq A$$

定理3.9.2（极分解）任意复方阵A必有如下分解:

$$A = GW$$

式中，G为半正定Hermite矩阵，W为酉矩阵．当矩阵A可逆时，G是正定Hermite矩阵，此时该极分解式唯一.

注5: 上述分解式称为**左极分解**；若$A = VG$，其中V是酉矩阵，G是半正定Hermite矩阵，则称分解式$A = VG$为**右极分解**.

【应用】: 奇异值分解在矩阵特征值、广义逆矩阵等矩阵分析和计算方面有着重要应用，而且在图像处理、机器学习等领域有着广泛应用.

（1）**最小二乘问题**

设$A \in \mathbb{C}^{m\times n}$, $\boldsymbol{b} \in \mathbb{C}^m$, 求$\boldsymbol{x}^* \in \mathbb{C}^n$使得

$$\|A\boldsymbol{x}^* - \boldsymbol{b}\| = \min_{\boldsymbol{x}\in\mathbb{C}^n}\|A\boldsymbol{x} - \boldsymbol{b}\|$$

一般而言$\boldsymbol{b} \notin R(A)$, 否则必存在$\boldsymbol{x}^* \in \mathbb{C}^n$使得$A\boldsymbol{x}^* = \boldsymbol{b}$.

设A的奇异值分解为:

$$A = U\sum V^H.$$

于是有

$$Ax - b = U\sum V^H x - b = U(\sum V^H x - U^H b).$$

令$y = V^H x, c = U^H b$, 则有

$$Ax - b = U(\sum y - c).$$

因为U是酉矩阵, 不改变向量的长度, 所以有

$$\|Ax - b\| = \|\sum y - c\|.$$

即有$\min\limits_{x\in\mathbb{C}^n}\|Ax - b\| = \min\limits_{x\in\mathbb{C}^n}\|\sum y - c\|.$

由

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \Rightarrow \Sigma y = \begin{bmatrix} \sigma_1 y_1 \\ \vdots \\ \sigma_r y_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

所以有

$$\Sigma y - c = [\sigma_1 y_1 - c_1 \quad \cdots \quad \sigma_r y_r - c_r \quad -c_{r+1} \quad \cdots \quad -c_m]^T$$

所以有

$$\Sigma y - c = \begin{bmatrix} \sigma_1 y_1 - c_1 & \cdots & \sigma_r y_r - c_r & -c_{r+1} & \cdots & -c_m \end{bmatrix}^T$$

当$\sigma_1 y_1 - c_1 = \cdots = \sigma_r y_r - c_r = 0$时, $\|\Sigma y - c\|$达到最小. 由此可解出: $y_1, y_2, \cdots, y_r$.

再由$y = V^H x$, 得$x = Vy$.

（2）**图像压缩**

问题 存储矩阵$A \in \mathbb{C}^{m\times n}$, 须存储$mn$个数据

希望用尽可能少的数据来逼近A

考察秩1的矩阵如何存储？

（2）图像压缩

问题 存储矩阵$A \in \mathbb{C}^{m\times n}$, 须存储$mn$个数据

希望用尽可能少的数据来逼近A

考察秩1的矩阵如何存储?

例如, 若$\mathrm{rank}(A) = 1$, 则有

$$A = uv^T, u \in \mathbb{C}^m, v \in \mathbb{C}^n$$

即只要用$m + n$个数据就能表示A.

一幅图像就对应一个矩阵:

$$A=\begin{bmatrix} f(1,1) & f(1,2) & \cdots & f(1,n) \\ f(2,1) & f(2,2) & \cdots & f(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ f(m,1) & f(m,2) & \cdots & f(m,n) \end{bmatrix}$$

$$=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

其中$f(x,y)$表示点(x,y)处图像的灰度或强度.

设A的奇异值分解为

$$A = U\begin{bmatrix}\sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0\end{bmatrix}V^H$$

$$= \begin{bmatrix}u_1 & \cdots & u_m\end{bmatrix}\begin{bmatrix}\sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0\end{bmatrix}\begin{bmatrix}v_1^H \\ \vdots \\ v_n^H\end{bmatrix}$$

$$= \sigma_1 u_1 v_1^H + \cdots + \sigma_r u_r v_r^H$$

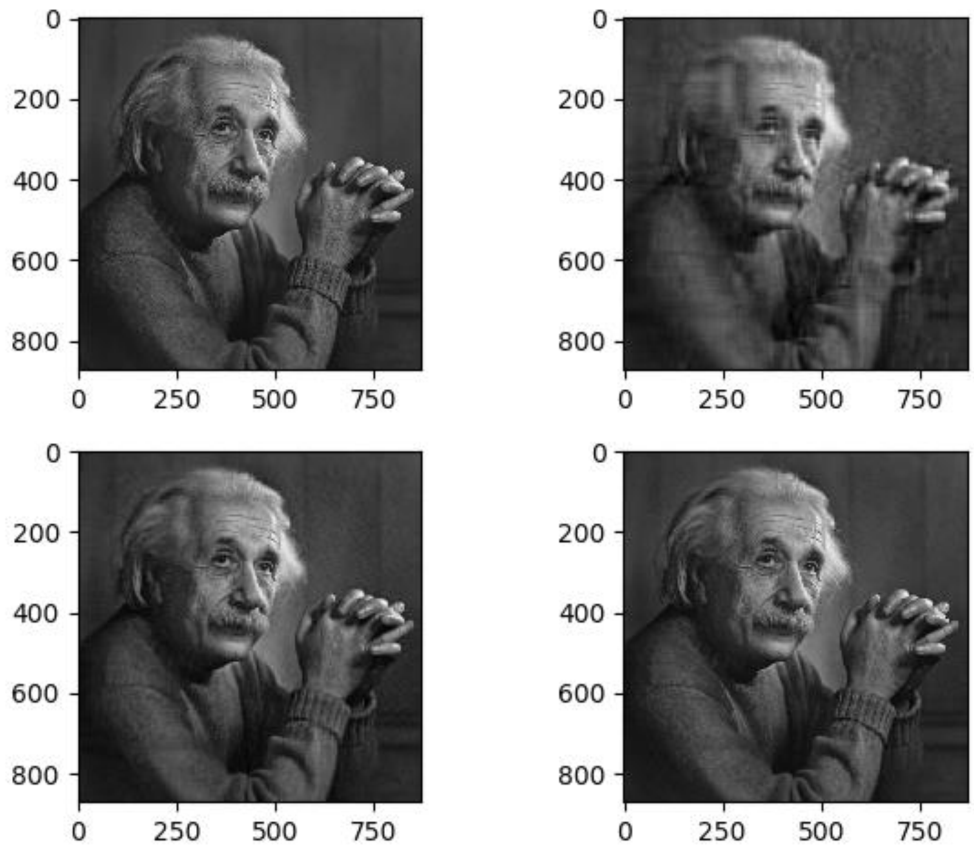
设A的奇异值分解为

$$A = U\begin{bmatrix}\sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0\end{bmatrix}V^H$$

$$= \begin{bmatrix}u_1 & \cdots & u_m\end{bmatrix}\begin{bmatrix}\sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0\end{bmatrix}\begin{bmatrix}v_1^H \\ \vdots \\ v_n^H\end{bmatrix}$$

$$= \sigma_1 u_1 v_1^H + \cdots + \sigma_r u_r v_r^H$$

（2）图像压缩



$k = 20$

$k = 50$

$k = 100$

课 堂 测 试

已知$A=\begin{pmatrix}1&0&0\\1&1&0\\1&1&1\end{pmatrix}$.问：无需计算，判断矩阵$A$能否分解为两个正规矩阵的乘积，请说明理由.