第四章　矩阵的因子分解

4.1 初等矩阵

4.1.1 初等矩阵

- **定义 4.1.1** 设 $u, v \in \mathbb{C}^n$，σ 为一复数，如下形式的矩阵

$$\boldsymbol{E}(u, v, \sigma) = \boldsymbol{I} - \sigma u v^{\mathrm{H}} \tag{4.1.1}$$

称为初等矩阵。

- 例

$$\begin{aligned}
u &= \begin{bmatrix} 1 \\ 2 \end{bmatrix}, v = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \sigma = 2, \\
\boldsymbol{E}(u, v, \sigma) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & 4 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 6 & 8 \\ 12 & 16 \end{bmatrix} \\
&= \begin{bmatrix} -5 & -8 \\ -12 & -15 \end{bmatrix}
\end{aligned}$$

- **定理 4.1.1** 初等矩阵 $\boldsymbol{E}(u,v,\sigma)$ 具有如下性质：
  (1) $\det(\boldsymbol{E}(u,v,\sigma)) = 1-\sigma v^{\mathrm{H}}u$；
  (2) 如果 $\sigma v^{\mathrm{H}}u \neq 1$，则 $\boldsymbol{E}(u,v,\sigma)$ 可逆，并且其逆矩阵也是初等矩阵
  $$\boldsymbol{E}(u,v,\sigma)^{-1} = \boldsymbol{E}(u,v,\tau)$$
  其中 $\tau = \dfrac{\sigma}{\sigma v^{\mathrm{H}}u-1}$。
  (3) 对任意非零向量 $a,b \in \mathbb{C}^n$，可适当选取 u,v 和 σ 使得
  $$\boldsymbol{E}(u,v,\sigma)a = b$$
  证明：（3）取u,v,σ满足$v^{\mathrm{H}}a \neq 0$,$\sigma u = \frac{a-b}{v^{\mathrm{H}}a}$，则
  $$\boldsymbol{E}(u,v,\sigma)a = (\boldsymbol{I}-\sigma uv^{\mathrm{H}})a = a-\frac{a-b}{v^{\mathrm{H}}a}v^{\mathrm{H}}a = b$$

- **例 4.1.1**

$$
\begin{aligned}
\boldsymbol{P}(i,j) &= \left[\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 0 & & \cdots & & 1 & & & \\
 & & & & 1 & & & & & & \\
 & & & \vdots & & \ddots & & \vdots & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & \cdots & & 0 & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \\ \text{第}i\text{行} \\ \\ \\ \\ \text{第}j\text{行} \\ \\ \\ \\
\end{array} \\
&= \boldsymbol{I} - \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & -1 & 0 & \cdots & 0 \end{bmatrix} \\
&= \boldsymbol{I} - (e_i - e_j)(e_i - e_j)^{\mathrm{T}} \\
&= \boldsymbol{E}(e_i - e_j, e_i - e_j, 1)
\end{aligned}
$$

- **例 4.1.2**

$$
\begin{aligned}
\boldsymbol{P}(i(k)) &= \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & k & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \text{第}i\text{行} \\
&= \boldsymbol{I} - (1-k)\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \\
&= \boldsymbol{I} - (1-k)(e_i)(e_i)^{\mathrm{T}} \\
&= \boldsymbol{E}(e_i, e_i, 1-k)
\end{aligned}
$$

4.1.2 初等下三角矩阵

- 令 $u = l_i = (0, \cdots, 0, l_{i+1,i}, \cdots, l_{ni})^{\mathrm{T}} \in \mathbb{C}^n$， $v = e_i, \sigma = 1$，则

$$\boldsymbol{L}_i = \boldsymbol{L}_i(l_i) = \boldsymbol{E}(l_i, e_i, 1)$$

称为**初等下三角矩阵**。即

$$\boldsymbol{L}_i = \boldsymbol{L}_i(l_i) = \boldsymbol{I} - l_i e_i^{\mathrm{T}} = \begin{bmatrix} 1 & & & & & 0 & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & -l_{i+1,i} & 1 & & & \\ & & \vdots & & \ddots & & \\ 0 & & \vdots & 0 & & \ddots & \\ & & -l_{ni} & & & & 1 \end{bmatrix} \tag{4.1.4}$$

由定理4.1.1知 $\det(L_i) = 1$，并且

$$\boldsymbol{L}_i^{-1} = \boldsymbol{E}(l_i, e_i, -1) = \boldsymbol{L}_i(-l_i) = \begin{bmatrix} 1 & & & & & 0 & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & l_{i+1,i} & 1 & & & \\ & & \vdots & & \ddots & & \\ 0 & & \vdots & 0 & & \ddots & \\ & & l_{ni} & & & & 1 \end{bmatrix}$$

对初等下三角矩阵(4.1.4)，当 $i < j$ 时，有

$$\boldsymbol{L}_i(l_i)\boldsymbol{L}_j(l_j) = \boldsymbol{I} - l_i e_i^T - l_j e_j^T = \begin{bmatrix} 1 & & & & & & & \\ & \ddots & & & & & 0 & \\ & & 1 & & & & & \\ & & -l_{i+1,i} & \ddots & & & & \\ & & \vdots & & 1 & & & \\ & 0 & \vdots & 0 & -l_{j+1,j} & \ddots & & \\ & & & & \vdots & & & \\ & & -l_{ni} & & -l_{nj} & & 0 & 1 \end{bmatrix}$$

例

$$\begin{bmatrix} 1 & 0 & 0 \\ -\frac{4}{3} & 1 & 0 \\ -\frac{5}{3} & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 6 & 9 \\ 4 & 9 & 10 \\ 5 & 123 & 11 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ 0 & 1 & -2 \\ 0 & 113 & -4 \end{bmatrix}$$

用初等下三角矩阵 $\boldsymbol{L}_i$ 左乘一个矩阵 $\boldsymbol{A}$，等于从 $\boldsymbol{A}$ 的第 k 行减去第 i 行乘以 $l_{ki}(k=i+1,\cdots,n)$。对于 $\boldsymbol{A}=(a_{ij})$，如果 $a_{ij}\neq 0$，取

$$l_{ki}=\frac{a_{kj}}{a_{ij}}, \qquad k=i+1,\cdots,n$$

则 $\boldsymbol{L}_i\boldsymbol{A}$ 的第 $(i+1,j),\cdots,(n,j)$ 元素全为零。这就是消去法的一步。

4.1.3 Householder矩阵

- 在(4.1.1)中取 $u = v = w, \sigma = 2$，并且 w 是单位向量，即 $\|w\| = 1$，初等矩阵

$$\boldsymbol{H}(w) = \boldsymbol{E}(w, w, 2) = \boldsymbol{I} - 2ww^{\mathrm{H}}$$

称为Householder矩阵或初等Hermite矩阵。

- **定理 4.1.2** Householder矩阵 $\boldsymbol{H}(w)$ 具有如下性质：

(1) $\det(\boldsymbol{H}(w)) = -1$；

(2) $\boldsymbol{H}(w)^{\mathrm{H}} = \boldsymbol{H}(w) = \boldsymbol{H}(w)^{-1}$；

(3) 设 $a, b \in \mathbb{C}^n$ 且 $a \neq b$，则存在单位向量 w 使得 $\boldsymbol{H}(w)a = b$ 的充分必要条件是

$$a^{\mathrm{H}}a = b^{\mathrm{H}}b,\ a^{\mathrm{H}}b = b^{\mathrm{H}}a \tag{4.1.8}$$

并且若上述条件成立，则使 $\boldsymbol{H}(w)a = b$ 成立的单位向量 w 可取为

$$w = e^{\mathrm{i}\theta}(a-b)/\|a-b\|$$

其中 θ 为任一实数。

证明 (3)必要性。

$$\begin{aligned} b^{\mathrm{H}}b &= a^{\mathrm{H}}\boldsymbol{H}(w)^{\mathrm{H}}\boldsymbol{H}(w)a = a^{\mathrm{H}}\boldsymbol{H}(w)^{-1}\boldsymbol{H}(w)a = a^{\mathrm{H}}a \\ b^{\mathrm{H}}a &= a^{\mathrm{H}}\boldsymbol{H}(w)^{\mathrm{H}}a = a^{\mathrm{H}}\boldsymbol{H}(w)a = a^{\mathrm{H}}b \end{aligned} \tag{1}$$

充分性。

因为 $a \neq b$，则取 $w = e^{\mathrm{i}\theta}(a-b)/\|a-b\|$，有

$$\boldsymbol{H}(w)a = \left(\boldsymbol{I} - 2\frac{(a-b)(a-b)^{\mathrm{H}}}{\|a-b\|^2}\right)a = a - \frac{2(a^{\mathrm{H}}a - b^{\mathrm{H}}a)(a-b)}{a^{\mathrm{H}}a + b^{\mathrm{H}}b - a^{\mathrm{H}}b - b^{\mathrm{H}}a}$$

由条件(4.1.8)即得 $\boldsymbol{H}(w)a = b$

- 对于非零向量 $a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \neq e_1$，若令

$$\sigma = \begin{cases} \|a\|, & a_1 = 0 \\ -e^{\mathrm{i}(\arg a_1)}\|a\|, & a_1 \neq 0 \end{cases}$$

并取

$$w = (a - \sigma e_1)/\|a - \sigma e_1\|$$

则有

$$\boldsymbol{H}(w)a = \sigma e_1$$

4.2 满秩分解

- **定理 4.2.1 (满秩分解定理)** 设 $m\times n$ 矩阵 $\boldsymbol{A}$ 的秩为 $r>0$，则存在 $m\times r$ 矩阵 $\boldsymbol{B}$ 和 $r\times n$ 矩阵 $\boldsymbol{C}$ 使得
$$\boldsymbol{A}=\boldsymbol{B}\boldsymbol{C} \tag{4.2.2}$$
并且 rank$(\boldsymbol{B})$ =rank$(\boldsymbol{C})=r$。
证明:
$$\begin{aligned}\boldsymbol{A} &= \boldsymbol{P}\begin{bmatrix}\boldsymbol{I}_r & 0\\ 0 & 0\end{bmatrix}\boldsymbol{Q}\\ &= \boldsymbol{P}\begin{bmatrix}\boldsymbol{I}_r\\ 0\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_r & 0\end{bmatrix}\boldsymbol{Q}\end{aligned}$$
取
$$\boldsymbol{B}=\boldsymbol{P}\begin{bmatrix}\boldsymbol{I}_r\\ 0\end{bmatrix},\boldsymbol{C}=\begin{bmatrix}\boldsymbol{I}_r & 0\end{bmatrix}\boldsymbol{Q}$$

- 注意：满秩分解不惟一。对任一r阶非奇异矩阵$\boldsymbol{D}$，若令$\boldsymbol{B}_1=\boldsymbol{B}\boldsymbol{D},\boldsymbol{C}_1=\boldsymbol{D}^{-1}\boldsymbol{C}$，则有$\boldsymbol{A}=\boldsymbol{B}_1\boldsymbol{C}_1$

- **例 4.2.2** 求矩阵

$$\boldsymbol{A}=\begin{bmatrix} 0 & 2 & 1 & -1 \\ 1 & 2 & 3 & 0 \\ -2 & -2 & -5 & -1 \end{bmatrix}$$

的一个满秩分解。

解 取

$$\boldsymbol{P}(1,2)=\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

则

$$\boldsymbol{P}(1,2)\boldsymbol{A}=\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 2 & 1 & -1 \\ -2 & -2 & -5 & -1 \end{bmatrix}$$

令

$$\boldsymbol{L}_1=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

则

$$\boldsymbol{L}_1\boldsymbol{P}(1,2)\boldsymbol{A}=\begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 2 & 1 & -1 \end{bmatrix}=\boldsymbol{A}^{(1)}$$

取

$$\boldsymbol{L}_2=\begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}$$

则

$$\boldsymbol{L}_2\boldsymbol{L}_1\boldsymbol{P}(1,2)\boldsymbol{A}=\begin{bmatrix}1&2&3&0\\0&2&1&-1\\0&0&0&0\end{bmatrix}=\boldsymbol{A}^{(2)}$$

因此

$$\boldsymbol{A}=\boldsymbol{P}(1,2)\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\boldsymbol{A}^{(2)}=\boldsymbol{P}(1,2)\begin{bmatrix}1&0&0\\0&1&0\\-2&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&1&1\end{bmatrix}\begin{bmatrix}1&2&3&0\\0&2&1&-1\\0&0&0&0\end{bmatrix}$$

$$=\boldsymbol{P}(1,2)\begin{bmatrix}1&0&0\\0&1&0\\-2&1&1\end{bmatrix}\begin{bmatrix}1&2&3&0\\0&2&1&-1\\0&0&0&0\end{bmatrix}=\left[\begin{array}{cc|c}0&1&0\\1&0&0\\-2&1&1\end{array}\right]\left[\begin{array}{cccc}1&2&3&0\\0&2&1&-1\\\hline 0&0&0&0\end{array}\right]$$

令 $\boldsymbol{B}=\begin{bmatrix}0&1\\1&0\\-2&1\end{bmatrix}$，$\boldsymbol{C}=\begin{bmatrix}1&2&3&0\\0&2&1&-1\end{bmatrix}$，则 $\boldsymbol{A}=\boldsymbol{B}\boldsymbol{C}$。

4.3 三角分解

- $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，如果$\boldsymbol{A}$的对角线下方的元素全为零，即$a_{ij}=0, \forall i>j$，则称$\boldsymbol{A}$为上三角矩阵；如果$\boldsymbol{A}$的对角线上方的元素全为零，即$a_{ij}=0, \forall i<j$，则称$\boldsymbol{A}$为下三角矩阵；对角元全为1的上三角矩阵称为单位上三角矩阵；对角元全为1的下三角矩阵称为单位下三角矩阵。
- $\boldsymbol{A}, \boldsymbol{B}$是上（下）三角矩阵，$\boldsymbol{A}+\boldsymbol{B}$，$\boldsymbol{AB}$ 仍是上（下）三角矩阵，并且$\boldsymbol{A}$可逆的充分必要条件是$\boldsymbol{A}$的对角元均非零。当$\boldsymbol{A}$可逆时，其逆矩阵也是上（下）三角矩阵。
- 两个单位上（下）三角矩阵的乘积仍是单位上（下）三角矩阵，并且单位上（下）三角矩阵的逆矩阵也是单位上（下）三角矩阵。
- 设$\boldsymbol{A}$是n阶矩阵，如果有下三角矩阵$\boldsymbol{L}$和上三角矩阵$\boldsymbol{U}$使得$\boldsymbol{A}=\boldsymbol{LU}$，则称$\boldsymbol{A}$能作三角分解，并且称$\boldsymbol{A}=\boldsymbol{LU}$为$\boldsymbol{A}$的三角分解或$LU$分解。

- $\boldsymbol{A}\in\mathbb{C}^{n\times n}$，在$\boldsymbol{A}$中选取$k$行$k$列$(1\le k\le n)$，由这些行和列相交处的元素按原来相关位置构成的$k$阶行列式称为$\boldsymbol{A}$的$k$阶子式。$\boldsymbol{A}$中由第$i_1$行，第$i_2$行，$\cdots$，第$i_k$行和第$i_1$列，第$i_2$列，$\cdots$，第$i_k$列组成的$k$阶子式称为$k$阶主子式。特别地，当

  $$i_1=1,i_2=2,\cdots,i_k=k$$

  时，称为k阶顺序主子式，记为Δ_k
- 例

  $$\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}$$

  2阶子式$\begin{bmatrix}1&2\\4&5\end{bmatrix},\begin{bmatrix}1&3\\4&6\end{bmatrix},\begin{bmatrix}1&2\\7&8\end{bmatrix},\begin{bmatrix}1&3\\7&9\end{bmatrix},\begin{bmatrix}4&5\\7&8\end{bmatrix},\begin{bmatrix}5&6\\8&9\end{bmatrix},\begin{bmatrix}2&3\\5&6\end{bmatrix},\begin{bmatrix}2&3\\8&9\end{bmatrix},\begin{bmatrix}4&6\\7&9\end{bmatrix}$

  2阶主子式$\begin{bmatrix}1&2\\4&5\end{bmatrix},\begin{bmatrix}1&3\\7&9\end{bmatrix},\begin{bmatrix}5&6\\8&9\end{bmatrix}$

  2阶顺序主子式$\begin{bmatrix}1&2\\4&5\end{bmatrix}$

- **定理 4.3.1 ($\boldsymbol{LU}$分解定理)** 设 $\boldsymbol{A}$ 是 n 阶非奇异矩阵，则存在唯一的单位下三角矩阵 $\boldsymbol{L}$ 和上三角矩阵 $\boldsymbol{U}$ 使得
$$\boldsymbol{A} = \boldsymbol{LU} \tag{4.3.1}$$
的充分必要条件是 $\boldsymbol{A}$ 的所有顺序主子式均非零，即
$$\Delta_k \neq 0, k = 1, \cdots, n-1$$
证明 必要性。如果存在单位下三角矩阵 $\boldsymbol{L}$ 和上三角矩阵 $\boldsymbol{U}$ 使得 $\boldsymbol{A} = \boldsymbol{LU}$，记
$$\boldsymbol{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & u_{nn} \end{bmatrix}$$
则 $|\boldsymbol{A}| = |\boldsymbol{LU}| = |\boldsymbol{U}| = u_{11}u_{22}\cdots u_{nn}$。因为 $\boldsymbol{A}$ 非奇异，所以 $u_{ii} \neq 0$。将 $\boldsymbol{A} = \boldsymbol{LU}$ 分块写成
$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{L}_{11} & 0 \\ \boldsymbol{L}_{21} & \boldsymbol{L}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}_{11} & \boldsymbol{U}_{12} \\ 0 & \boldsymbol{U}_{22} \end{bmatrix}$$

其中 $\boldsymbol{A}_{11}, \boldsymbol{L}_{11}, \boldsymbol{U}_{11}$ 分别为 $\boldsymbol{A}, \boldsymbol{L}, \boldsymbol{U}$ 的 k 阶顺序主子矩阵，于是

$$\boldsymbol{A}_{11} = \boldsymbol{L}_{11}\boldsymbol{U}_{11}$$

从而 $|\boldsymbol{A}_{11}| = \Delta_k = |\boldsymbol{U}_{11}| = u_{11}\cdots u_{kk} \neq 0\ (k = 1, 2, \cdots, n)$，并且

$$u_{11} = a_{11},\ u_{kk} = \frac{\Delta_k}{\Delta_{k-1}}, k = 2, \cdots, n$$

充分性。对矩阵的阶数作归纳法证明分解式(4.3.1)存在。当矩阵的阶为1时结论显然成立。设对 $n-1$ 阶矩阵有分解式(4.3.1)。对 n 阶矩阵 $\boldsymbol{A}$，记

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{n-1} & \beta \\ \alpha & a_{nn} \end{bmatrix}$$

其中 $\boldsymbol{A}_{n-1}$ 为 $\boldsymbol{A}$ 的 $n-1$ 阶顺序主子矩阵。根据定理的条件，$\boldsymbol{A}_{n-1}$ 是非奇异矩阵，则有

$$\begin{bmatrix} \boldsymbol{I}_{n-1} & 0 \\ -\alpha\boldsymbol{A}_{n-1}^{-1} & 1 \end{bmatrix}\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I}_{n-1} & 0 \\ -\alpha\boldsymbol{A}_{n-1}^{-1} & 1 \end{bmatrix}\begin{bmatrix} \boldsymbol{A}_{n-1} & \beta \\ \alpha & a_{nn} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{n-1} & \beta \\ 0 & a_{nn} - \alpha\boldsymbol{A}_{n-1}^{-1}\beta \end{bmatrix}$$

从而

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I}_{n-1} & 0 \\ \alpha\boldsymbol{A}_{n-1}^{-1} & 1 \end{bmatrix}\begin{bmatrix} \boldsymbol{A}_{n-1} & \beta \\ 0 & a_{nn} - \alpha\boldsymbol{A}_{n-1}^{-1}\beta \end{bmatrix}$$

由归纳假设，存在 $n-1$ 阶单位下三角矩阵 $\boldsymbol{L}_{n-1}$ 和上三角矩阵 $\boldsymbol{U}_{n-1}$ 使得 $\boldsymbol{A}_{n-1}=\boldsymbol{L}_{n-1}\boldsymbol{U}_{n-1}$。于是可得

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{I}_{n-1} & 0\\ \alpha\boldsymbol{A}_{n-1}^{-1} & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{L}_{n-1}\boldsymbol{U}_{n-1} & \beta\\ 0 & a_{nn}-\alpha\boldsymbol{A}_{n-1}^{-1}\beta\end{bmatrix}$$

$$=\begin{bmatrix}\boldsymbol{I}_{n-1} & 0\\ \alpha\boldsymbol{A}_{n-1}^{-1} & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{L}_{n-1} & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{U}_{n-1} & \boldsymbol{L}_{n-1}^{-1}\beta\\ 0 & a_{nn}-\alpha\boldsymbol{A}_{n-1}^{-1}\beta\end{bmatrix}$$

令

$$\boldsymbol{L}=\begin{bmatrix}\boldsymbol{I}_{n-1} & 0\\ \alpha\boldsymbol{A}_{n-1}^{-1} & 1\end{bmatrix}\begin{bmatrix}\boldsymbol{L}_{n-1} & 0\\ 0 & 1\end{bmatrix},\ \boldsymbol{U}=\begin{bmatrix}\boldsymbol{U}_{n-1} & \boldsymbol{L}_{n-1}^{-1}\beta\\ 0 & a_{nn}-\alpha\boldsymbol{A}_{n-1}^{-1}\beta\end{bmatrix}$$

即得 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}$，其中 $\boldsymbol{L}$ 是单位下三角矩阵，$\boldsymbol{U}$是上三角矩阵。因此矩阵的阶为 n 时分解式(4.3.1)也存在。

下面证明唯一性。如果

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}=\widetilde{\boldsymbol{L}}\widetilde{\boldsymbol{U}}$$

其中 $\boldsymbol{L}$，$\widetilde{\boldsymbol{L}}$ 为 n 阶单位下三角矩阵， $\boldsymbol{U}$，$\widetilde{\boldsymbol{U}}$ 为 n 阶可逆上三角矩阵，则

$$\widetilde{\boldsymbol{L}}^{-1}\boldsymbol{L}=\widetilde{\boldsymbol{U}}\boldsymbol{U}^{-1}$$

上式左边的矩阵是单位下三角矩阵，而右边的矩阵是上三角矩阵。因此 $\widetilde{\boldsymbol{L}}^{-1}\boldsymbol{L}=\widetilde{\boldsymbol{U}}\boldsymbol{U}^{-1}=\boldsymbol{I}$。于是 $\boldsymbol{L}=\widetilde{\boldsymbol{L}},\boldsymbol{U}=\widetilde{\boldsymbol{U}}$。这就证明了唯一性。 $\square$

- 例

$$\begin{bmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -\frac{2}{5} & 1 & 0 \\ \frac{4}{5} & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{bmatrix} = \begin{bmatrix} 5 & 2 & -4 \\ 0 & \frac{1}{5} & -\frac{2}{5} \\ 0 & -\frac{2}{5} & \frac{9}{5} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 & -4 \\ 0 & \frac{1}{5} & -\frac{2}{5} \\ 0 & -\frac{2}{5} & \frac{9}{5} \end{bmatrix} = \begin{bmatrix} 5 & 2 & -4 \\ 0 & \frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 \end{bmatrix}$$

于是

$$\begin{aligned} \begin{bmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 \\ -\frac{2}{5} & 1 & 0 \\ \frac{4}{5} & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 5 & 2 & -4 \\ 0 & \frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{5} & 1 & 0 \\ -\frac{4}{5} & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 & -4 \\ 0 & \frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{5} & 1 & 0 \\ -\frac{4}{5} & -2 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 & -4 \\ 0 & \frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

非奇异上三角矩阵

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & u_{n-1,n-1} & u_{n-1,n} \\ 0 & \cdots & & 0 & u_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} u_{11} & & & & \\ & u_{22} & & 0 & \\ & & \ddots & & \\ & 0 & & u_{n-1,n-1} & \\ & & & & u_{nn} \end{bmatrix} \begin{bmatrix} 1 & \frac{u_{12}}{u_{11}} & \frac{u_{13}}{u_{11}} & \cdots & \frac{u_{1n}}{u_{11}} \\ & 1 & \frac{u_{23}}{u_{22}} & \cdots & \frac{u_{2n}}{u_{22}} \\ & & \ddots & \ddots & \vdots \\ & 0 & & 1 & \frac{u_{n-1,n}}{u_{n-1,n-1}} \\ & & & & 1 \end{bmatrix} \tag{4.3.5}$$

- **定理 4.3.2 (*LDU*分解定理)** 设 $\boldsymbol{A}$ 是 n 阶非奇异矩阵,则存在唯一的单位下三角矩阵 $\boldsymbol{L}$，对角矩阵 $\boldsymbol{D}$ =diag$(d_1, d_2, \cdots, d_n)$ 和单位上三角矩阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{A} = \boldsymbol{LDU} \tag{4.3.6}$$

的充分必要条件是 $\boldsymbol{A}$ 的所有顺序主子式均非零，即 $\Delta_k \neq 0$ $(i = 1, \cdots, n-1)$，并且

$$d_1 = a_{11},\ d_k = \frac{\Delta_k}{\Delta_{k-1}}, k = 2, \cdots, n$$

分解式(4.3.6)称为矩阵 $\boldsymbol{A}$ 的 ***LDU* 分解**。

- **定义 4.3.1** 设 e_i 是 n 阶单位矩阵的第 i 列 $(i = 1, 2, \cdots, n)$，以 $e_1, e_2, \cdots, e_n$ 为列作成的矩阵 $[e_{i_1}, e_{i_2}, \cdots, e_{i_n}]$ 称为n阶**排列矩阵**，其中 $i_1, i_2, \cdots, i_n$ 是 $1, 2, \cdots, n$ 的一个排列。

- **定理 4.3.3** 设 $\boldsymbol{A}$ 是 n 阶非奇异矩阵，则存在排列矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{PA} = \boldsymbol{L}\widetilde{\boldsymbol{U}} = \boldsymbol{LDU} \tag{4.3.8}$$

其中 $\boldsymbol{L}$ 是单位下三角矩阵，$\widetilde{\boldsymbol{U}}$ 是上三角矩阵，$\boldsymbol{U}$ 是单位上三角矩阵，$\boldsymbol{D}$ 是对角矩阵。

- 例

$$\boldsymbol{A} = \begin{bmatrix} 0 & 4 & 6 \\ 0 & -3 & -5 \\ 1 & -3 & -6 \end{bmatrix}$$

不满足三角分解的条件。而

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 4 & 6 \\ 0 & -3 & -5 \\ 1 & -3 & -6 \end{bmatrix} = \begin{bmatrix} 1 & -3 & -6 \\ 0 & -3 & -5 \\ 0 & 4 & 6 \end{bmatrix}$$

满足三角分解的条件。

4.4 QR 分解

- 定理4.4.2　设 $\boldsymbol{A}$ 是 $m\times n$ 实(复)矩阵，且其 n 个列向量线性无关，则存在 m 阶正交(酉)矩阵 $\boldsymbol{Q}$ 和 n 阶非奇异实(复)上三角矩阵 $\boldsymbol{R}$ 使得$\boldsymbol{QA}=\begin{bmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{bmatrix}$

证明：令 $\boldsymbol{A}=\boldsymbol{A}^{(0)}=[\alpha_1,\alpha_2,\cdots,\alpha_n]$，存在$n$阶Householder矩阵 $\boldsymbol{H}_1$ 使得

$$\boldsymbol{H}_1\alpha_1=k_1e_1,\quad |k_1|=\|\alpha_1\|>0$$

$$\boldsymbol{A}^{(1)}=\boldsymbol{H}_1\boldsymbol{A}^{(0)}=[\boldsymbol{H}_1\alpha_1,\boldsymbol{H}_1\alpha_2,\cdots,\boldsymbol{H}_1\alpha_n]=\left[\begin{array}{c|ccc} k_1 & \beta_2 & \cdots & \beta_n\\ \hline 0 & & & \\ \vdots & \alpha_2^{(1)} & \cdots & \alpha_n^{(1)}\\ 0 & & & \end{array}\right]$$

有$n-1$阶Householder矩阵 $\tilde{\boldsymbol{H}}_2$ 使得$\tilde{\boldsymbol{H}}_2\alpha_2^{(1)}=k_2\tilde{e}_1$，$|k_2|=\|\alpha_2^{(1)}\|>0$，令$\boldsymbol{H}_2=\begin{bmatrix}1 & \\ & \tilde{\boldsymbol{H}}_2\end{bmatrix}$，易证$\boldsymbol{H}_2$是$n$阶Householder矩阵，且有

$$\boldsymbol{A}^{(2)}=\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{A}^{(0)}=\left[\begin{array}{cc|ccc} k_1 & \beta_2 & \beta_3 & \cdots & \beta_n\\ 0 & k_2 & \gamma_3 & \cdots & \gamma_n\\ \hline 0 & 0 & & & \\ \vdots & \vdots & \alpha_3^{(2)} & \cdots & \alpha_n^{(2)}\\ 0 & 0 & & & \end{array}\right]$$

如此继续，最终可得 $\boldsymbol{A}^{(n)}$，并且 $\boldsymbol{A}^{(n)}$ 的第 n 行以下元素全为零，即

$$\boldsymbol{H}_n\cdots\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{A}^{(0)}=\boldsymbol{A}^{(n)}=\begin{bmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{bmatrix}$$

$\boldsymbol{R}$ 为 n 阶非奇异实(复)上三角矩阵，$\boldsymbol{Q}=\boldsymbol{H}_n\cdots\boldsymbol{H}_2\boldsymbol{H}_1$为 m 阶正交(酉)矩阵。

- 定理4.4.1　设$\boldsymbol{A}$是n阶非奇异实（复）矩阵，则存在n阶正交(酉) 矩阵 $\boldsymbol{Q}$ 和n阶非奇异实(复)上三角矩阵 $\boldsymbol{R}$ 使得

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R} \tag{4.4.1}$$

除去相差一个对角元绝对值(模)全等于1的对角矩阵因子外，分解式(4.4.1)惟一。

惟一性证明：设矩阵 $\boldsymbol{A}$ 有两个 QR 分解

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R} = \boldsymbol{Q}_1\boldsymbol{R}_1$$

其中 $\boldsymbol{Q}, \boldsymbol{Q}_1$ 为正交(酉)矩阵， $\boldsymbol{R}, \boldsymbol{R}_1$ 为非奇异上三角矩阵，令$\boldsymbol{D} = \boldsymbol{R}_1\boldsymbol{R}^{-1}$，则

$$\begin{aligned} \boldsymbol{Q} &= \boldsymbol{Q}_1\boldsymbol{D} \\ \boldsymbol{R} &= \boldsymbol{D}^{-1}\boldsymbol{R}_1 \end{aligned}$$

由$\boldsymbol{I} = \boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q} = (\boldsymbol{Q}_1\boldsymbol{D})^{\mathrm{H}}(\boldsymbol{Q}_1\boldsymbol{D}) = \boldsymbol{D}^{\mathrm{H}}\boldsymbol{D}$，知$\boldsymbol{D}^{-1} = \boldsymbol{D}^{\mathrm{H}}$， 又$\boldsymbol{D}^{-1}$是上三角矩阵而$\boldsymbol{D}^{\mathrm{H}}$是下三角矩阵，所以 $\boldsymbol{D} = \mathrm{diag}(d_1, \cdots, d_n)$ 为对角矩阵，再由

$$\boldsymbol{D}^{\mathrm{H}}\boldsymbol{D} = \mathrm{diag}(\bar{d}_1, \cdots, \bar{d}_n)\mathrm{diag}(d_1, \cdots, d_n) = \mathrm{diag}(\bar{d}_1 d_1, \cdots, \bar{d}_n d_n) = \boldsymbol{I}$$

可知$|d_i| = 1, \forall i \in \{1, \cdots, n\}$，即$\boldsymbol{D}$是对角元绝对值(模)全等于1的对角矩阵。

- 注意：如果在非奇异矩阵 $\boldsymbol{A}$ 的 QR 分解中规定上三角矩阵的各个对角元的符号(例如全为正数)，则由

$$\boldsymbol{R} = \boldsymbol{D}^{-1}\boldsymbol{R}_1 = \mathrm{diag}\left(\frac{1}{d_1}, \cdots, \frac{1}{d_n}\right)\boldsymbol{R}_1$$

可知$\boldsymbol{A}$的 QR 分解是惟一的.

- 例：求$\boldsymbol{A}=\begin{bmatrix}0&3&1\\0&4&-2\\2&1&2\end{bmatrix}$的$QR$分解

解：令$\alpha_1^{(0)}=\begin{bmatrix}0&0&2\end{bmatrix}^{\mathrm{T}}$，取$k_1=\|\alpha_1^{(0)}\|=2$，则

$$w_1=\frac{\alpha_1^{(0)}-k_1e_1}{\|\alpha_1^{(0)}-k_1e_1\|}=\frac{\begin{bmatrix}-2&0&2\end{bmatrix}^{\mathrm{T}}}{\sqrt{8}},\ \boldsymbol{H}_1=\boldsymbol{I}-2w_1w_1^{\mathrm{H}}=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}$$

$$\boldsymbol{H}_1\boldsymbol{A}=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}\begin{bmatrix}0&3&1\\0&4&-2\\2&1&2\end{bmatrix}=\begin{bmatrix}2&1&2\\0&4&-2\\0&3&1\end{bmatrix}$$

再令$\alpha_2^{(1)}=\begin{bmatrix}4&3\end{bmatrix}^{\mathrm{T}}$，取$k_2=\|\alpha_2^{(1)}\|=5$，则

$$\tilde{w}_2=\frac{\alpha_2^{(1)}-k_2\tilde{e}_1}{\|\alpha_2^{(1)}-k_2\tilde{e}_1\|}=\frac{\begin{bmatrix}-1&3\end{bmatrix}^{\mathrm{T}}}{\sqrt{10}},\tilde{\boldsymbol{H}}_2=\boldsymbol{I}-2\tilde{w}_2\tilde{w}_2^{\mathrm{H}}=\frac{1}{5}\begin{bmatrix}4&3\\3&-4\end{bmatrix}$$

$$\boldsymbol{H}_2=\frac{1}{5}\begin{bmatrix}5&0&0\\0&4&3\\0&3&-4\end{bmatrix},\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{A}=\frac{1}{5}\begin{bmatrix}5&0&0\\0&4&3\\0&3&-4\end{bmatrix}\begin{bmatrix}2&1&2\\0&4&-2\\0&3&1\end{bmatrix}=\begin{bmatrix}2&1&2\\0&5&-1\\0&0&-2\end{bmatrix}$$

于是

$$\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$$

$$\boldsymbol{Q}=\boldsymbol{H}_1\boldsymbol{H}_2=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}\frac{1}{5}\begin{bmatrix}5&0&0\\0&4&3\\0&3&-4\end{bmatrix}=\frac{1}{5}\begin{bmatrix}0&3&-4\\0&4&3\\5&0&0\end{bmatrix},\boldsymbol{R}=\begin{bmatrix}2&1&2\\0&5&-1\\0&0&-2\end{bmatrix}$$

- 定义　$m\times r$矩阵$\boldsymbol{Q}_1$，$m\geq r$，如果$\boldsymbol{Q}_1^{\mathrm{H}}\boldsymbol{Q}_1=\boldsymbol{I}_r$，则称$\boldsymbol{Q}_1$为列正交规范矩阵。
- 定理4.4.3　设 $\boldsymbol{A}$ 是 $m\times n$ 的矩阵，且 $\mathrm{rank}(\boldsymbol{A})=r>0$,则存在 m 阶正交（酉）矩阵 $\boldsymbol{Q}$ 和 $r\times n$ 行满秩矩阵 $\boldsymbol{R}$ 使得$\boldsymbol{QA}=\begin{bmatrix}\boldsymbol{R}\\ \mathbf{0}\end{bmatrix}$或$\boldsymbol{A}$有分解$\boldsymbol{A}=\boldsymbol{Q}_1\boldsymbol{R}$，其中，$\boldsymbol{Q}_1$ 是 $m\times r$ 列正交规范矩阵，$\boldsymbol{R}$ 是 $r\times n$ 行满秩矩阵.

4.5 Schur定理与正规矩阵

- 定义4.5.1 设 $\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{n\times n}(\mathbb{C}^{n\times n})$,如果存在 n 阶正交(酉)矩阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{U}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{B}\ \ (\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{B})$$

则称 $\boldsymbol{A}$ 正交（酉）相似于$\boldsymbol{B}$

- 定理4.5.1(Schur定理) 任何一个 n 阶复矩阵 $\boldsymbol{A}$ 都酉相似于一个上三角矩阵，即存在一个 n 阶酉矩阵 $\boldsymbol{U}$ 和一个 n 阶上三角矩阵 $\boldsymbol{R}$ 使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{R} \tag{4.5.1}$$

其中 $\boldsymbol{R}$ 的对角元是 $\boldsymbol{A}$ 的特征值，它们可以按照要求的次序排列.

证明：用数学归纳法。

- 定义4.5.2　设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$,如果
$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}} = \boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} \tag{4.5.5}$$
则称 $\boldsymbol{A}$ 为正规矩阵
- 定义1.6.7　设$\boldsymbol{A} \in \mathbb{C}^{n\times n}$，如果$\boldsymbol{A}^{\mathrm{H}} = \boldsymbol{A}$，则称 $\boldsymbol{A}$ 为Hermite矩阵；如果$\boldsymbol{A}^{\mathrm{H}} = -\boldsymbol{A}$，则称 $\boldsymbol{A}$ 为反Hermite矩阵
- 对角矩阵、Hermite矩阵、反Hermite矩阵、正交（酉）矩阵都是正规矩阵

- 定理4.5.2　n阶矩阵 $\boldsymbol{A}$ 酉相似于一个对角矩阵的充分必要条件为 $\boldsymbol{A}$ 是正规矩阵

证明：必要性。若存在酉矩阵 $\boldsymbol{U}$ 和对角矩阵 $\boldsymbol{\Lambda}$ 使得 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}}$,则

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Lambda}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{U}\boldsymbol{\Lambda}^{\mathrm{H}}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{\Lambda}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$$

这说明 $\boldsymbol{A}$ 是正规矩阵.
充分性。由Schur定理知存在酉矩阵 $\boldsymbol{U}$ 使得 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^{\mathrm{H}}$,其中 $\boldsymbol{R}$ 是上三角矩阵.记

$$\boldsymbol{R}=\begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & r_{33} & \cdots & r_{3n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & r_{nn} \end{bmatrix}$$

因为 $\boldsymbol{U}\boldsymbol{R}^{\mathrm{H}}\boldsymbol{R}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{U}\boldsymbol{R}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{R}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{U}\boldsymbol{R}\boldsymbol{R}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}$,所以 $\boldsymbol{R}^{\mathrm{H}}\boldsymbol{R}=\boldsymbol{R}\boldsymbol{R}^{\mathrm{H}}$.比较

$$\begin{bmatrix} \bar{r}_{11} & 0 & \cdots & 0 \\ \bar{r}_{12} & \bar{r}_{22} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \bar{r}_{1n} & \bar{r}_{2n} & \cdots & \bar{r}_{nn} \end{bmatrix}\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{bmatrix}=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{bmatrix}\begin{bmatrix} \bar{r}_{11} & 0 & \cdots & 0 \\ \bar{r}_{12} & \bar{r}_{22} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \bar{r}_{1n} & \bar{r}_{2n} & \cdots & \bar{r}_{nn} \end{bmatrix}$$

两边的对角元素

$$\begin{aligned} &\bar{r}_{11}r_{11}=r_{11}\bar{r}_{11}+r_{12}\bar{r}_{12}+\cdots+r_{1n}\bar{r}_{1n} \\ &\bar{r}_{12}r_{12}+\bar{r}_{22}r_{22}=r_{22}\bar{r}_{22}+r_{23}\bar{r}_{23}+\cdots+r_{2n}\bar{r}_{2n} \\ &\cdots \\ &\bar{r}_{1n}r_{1n}+\bar{r}_{2n}r_{2n}+\cdots+\bar{r}_{nn}r_{nn}=r_{nn}\bar{r}_{nn} \end{aligned}$$

即得 $\boldsymbol{R}=\boldsymbol{\Lambda}=\mathrm{diag}(r_{11},r_{22},\cdots,r_{nn})$,则 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}}$

- 推论4.5.1　若 $\boldsymbol{A}$ 是 n 阶Hermite矩阵,则 $\boldsymbol{A}$ 必酉相似于实对角矩阵，即存在 n 阶酉矩阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{\Lambda} \tag{4.5.6}$$

其中 $\boldsymbol{\Lambda}=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$，$\lambda_i(i=1,\cdots,n)$ 是 $\boldsymbol{A}$ 的实特征值。式(4.5.6)称为Hermite矩阵$\boldsymbol{A}$的谱分解式。

证明：由定理4.5.2知，存在 n 阶酉矩阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{\Lambda}=\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

因为

$$\boldsymbol{U}\boldsymbol{\Lambda}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}}$$

则$\boldsymbol{\Lambda}^{\mathrm{H}}=\boldsymbol{\Lambda}$，即$\forall i\in\{1,\cdots,n\},\bar{\lambda}_i=\lambda_i$是实数，因此，$\boldsymbol{\Lambda}=\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$ 是实对角矩阵。再令ξ_i $(i=1,\cdots n)$是$\boldsymbol{U}$的第i列，即

$$\boldsymbol{U}=\begin{bmatrix}\xi_1 & \xi_2 & \cdots & \xi_n\end{bmatrix}$$

则由$\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}\boldsymbol{\Lambda}$，有

$$\begin{bmatrix}\boldsymbol{A}\xi_1 & \boldsymbol{A}\xi_2 & \cdots & \boldsymbol{A}\xi_n\end{bmatrix}=\begin{bmatrix}\xi_1 & \xi_2 & \cdots & \xi_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}=\begin{bmatrix}\lambda_1\xi_1 & \lambda_2\xi_2 & \cdots & \lambda_n\xi_n\end{bmatrix}$$

于是$\boldsymbol{A}\xi_i=\lambda_i\xi_i$，这意味着$\lambda_i$是$\boldsymbol{A}$的特征值。

- 推论4.5.2　若 $\boldsymbol{A}$ 是 n 阶实对称矩阵，则 $\boldsymbol{A}$ 正交相似于实对角矩阵，即存在 n 阶正交矩阵 $\boldsymbol{Q}$ 使得

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{\Lambda} \tag{4.5.13}$$

其中 $\boldsymbol{\Lambda}=\operatorname{diag}(\lambda_1,\cdots,\lambda_n),\lambda_i$ $(i=1,\cdots,n)$ 是 $\boldsymbol{A}$ 的实特征值。

4.6 奇异值分解

- 引理4.6.1　设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$,则

$$\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}) = \operatorname{rank}(\boldsymbol{A}) \tag{4.6.1}$$

证明：如果 $x \in \mathcal{N}(\boldsymbol{A})$，即$\boldsymbol{A}x = 0$，则$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}x = 0$，$x \in \mathcal{N}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$；反过来，如果 $x \in \mathcal{N}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$，即$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}x = 0$，那么 $x^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}x = 0$，$(\boldsymbol{A}x)^{\mathrm{H}}(\boldsymbol{A}x) = 0$,于是 $\boldsymbol{A}x = 0$，这意味着$x \in \mathcal{N}(\boldsymbol{A})$。因此

$$\dim(\mathcal{N}(\boldsymbol{A})) = \dim(\mathcal{N}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}))$$

而

$$\begin{aligned}\dim(\mathcal{N}(\boldsymbol{A})) &= n - \operatorname{rank}(\boldsymbol{A})\\ \dim(\mathcal{N}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})) &= n - \operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})\end{aligned}$$

故

$$\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A})$$

同理可证

$$\operatorname{rank}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}) = \operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}) = \operatorname{rank}(\boldsymbol{A})$$

- 引理4.6.2　设 $\boldsymbol{A}\in\mathbb{C}^{m\times n}$,则
  (1) $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的特征值均为非负实数
  (2) $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的非零特征值相同，并且非零特征值的个数(重特征值按重数计算)等于rank($\boldsymbol{A}$)

  证明：(1)设 λ 是 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的任一特征值， $x\neq 0$ 为相应的特征向量，则

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}x=\lambda x$$

$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$是Hermite矩阵，由推论4.5.1的证明过程可知 λ 是实数，并且

$$\lambda x^{\mathrm{H}}x=x^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}x=(\boldsymbol{A}x)^{\mathrm{H}}(\boldsymbol{A}x)\geq 0$$

由于 $x^{\mathrm{H}}x>0$,因此 $\lambda\geq 0$.
类似可以证明 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的特征值均为非负实数.
(2) 由节2.4的结论(⟨⟨矩阵论⟩⟩67页第1行)知， $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的非零特征值相同，并且由

$$\boldsymbol{P}^{-1}\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}\boldsymbol{P}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A},\quad \operatorname{rank}\left(\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}\right)=\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$$

和引理4.6.1知，非零特征值的个数等于rank$(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\operatorname{rank}(\boldsymbol{A})$

- 定义　设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, $\mathrm{rank}(\boldsymbol{A}) = r$,称$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$或$\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$的$r$个非零特征值$\lambda_i$ $(i = 1, \cdots, r)$的算术平方根$\sigma_i = \sqrt{\lambda_i}$ $(i = 1, \cdots, r)$为$\boldsymbol{A}$的奇异值
- 定理4.6.1　若 $\boldsymbol{A}$ 是正规矩阵，则 $\boldsymbol{A}$ 的奇异值是 $\boldsymbol{A}$ 的非零特征值的模.
证明：设 $\boldsymbol{A}$ 是 n 阶正规矩阵，由定理4.5.2知存在 n 阶酉矩阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{A} = \boldsymbol{U}\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)\boldsymbol{U}^{\mathrm{H}}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的特征值，不妨设$\boldsymbol{A}$的非零特征值为$\lambda_1, \cdots, \lambda_r$，则

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} = \boldsymbol{U}\mathrm{diag}(|\lambda_1|^2, |\lambda_2|^2, \cdots, |\lambda_n|^2)\boldsymbol{U}^{\mathrm{H}}$$

于是 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的特征值为$|\lambda_1|^2, |\lambda_2|^2, \cdots, |\lambda_n|^2$，$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$的非零特征值为$|\lambda_1|^2, \cdots, |\lambda_r|^2$，则 $\boldsymbol{A}$ 的奇异值是 $|\lambda_1|, \cdots, |\lambda_r|$

- 定理4.6.2 设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，且 $\mathrm{rank}(\boldsymbol{A})=r$, 则存在 m 阶酉矩阵 $\boldsymbol{V}$ 和 n 阶酉矩阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{V}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\begin{bmatrix}\boldsymbol{\Sigma} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} \tag{4.6.4}$$

其中 $\boldsymbol{\Sigma}=\mathrm{diag}(\sigma_1,\cdots,\sigma_r)$,且 $\sigma_1\geq\cdots\geq\sigma_r>0$.

证明：因为 $\mathrm{rank}(\boldsymbol{A})=r$,由引理4.6.2，可设 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的特征值是$\sigma_1^2\geq\cdots\geq\sigma_r^2>0,\sigma_{r+1}^2=\cdots=\sigma_n^2=0$，因为$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$是Hermite矩阵，由推论4.5.1知存在$n$阶酉矩阵$\boldsymbol{U}$使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\begin{bmatrix}\boldsymbol{\Sigma}^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix},\ \boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{U}\begin{bmatrix}\boldsymbol{\Sigma}^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}$$

记 $\boldsymbol{U}=[\boldsymbol{U}_1,\boldsymbol{U}_2]$,其中 $\boldsymbol{U}_1$ 是 $n\times r$ 矩阵. 上式可改写为$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}[\boldsymbol{U}_1\ \ \boldsymbol{U}_2]=[\boldsymbol{U}_1\ \ \boldsymbol{U}_2]\begin{bmatrix}\boldsymbol{\Sigma}^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}$，则有

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1=\boldsymbol{U}_1\boldsymbol{\Sigma}^2,\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_2=0,\boldsymbol{U}_2^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_2=0,(\boldsymbol{A}\boldsymbol{U}_2)^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_2=0,\boldsymbol{A}\boldsymbol{U}_2=0 \tag{4.6.8}$$

取$m\times r$矩阵$\boldsymbol{V}_1=\boldsymbol{A}\boldsymbol{U}_1\boldsymbol{\Sigma}^{-1}$，由

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1=\boldsymbol{U}_1\boldsymbol{\Sigma}^2,\boldsymbol{U}_1^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1=\boldsymbol{\Sigma}^2,\boldsymbol{\Sigma}^{-1}\boldsymbol{U}_1^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1\boldsymbol{\Sigma}^{-1}=\boldsymbol{I}_r,\boldsymbol{V}_1^{\mathrm{H}}\boldsymbol{V}_1=\boldsymbol{I}_r$$

知， $\boldsymbol{V}_1$ 是列正交规范矩阵，那么$\boldsymbol{\Sigma}=\boldsymbol{V}_1^{-1}\boldsymbol{A}\boldsymbol{U}_1=\boldsymbol{V}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1$。再取$m\times(m-r)$ 矩阵 $\boldsymbol{V}_2$ 使 $\boldsymbol{V}=[\boldsymbol{V}_1\ \ \boldsymbol{V}_2]$ 是酉矩阵，则

$$\boldsymbol{V}_2^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1=\boldsymbol{V}_2^{\mathrm{H}}\boldsymbol{V}_1\boldsymbol{\Sigma}=0 \tag{4.6.10}$$

由(4.6.8)和(4.6.10)有

$$\boldsymbol{V}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\begin{bmatrix}\boldsymbol{V}_1^{\mathrm{H}}\\ \boldsymbol{V}_2^{\mathrm{H}}\end{bmatrix}\boldsymbol{A}\begin{bmatrix}\boldsymbol{U}_1 & \boldsymbol{U}_2\end{bmatrix}=\begin{bmatrix}\boldsymbol{V}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1 & \boldsymbol{V}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_2\\ \boldsymbol{V}_2^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_1 & \boldsymbol{V}_2^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}_2\end{bmatrix}=\begin{bmatrix}\boldsymbol{\Sigma} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}$$

例4.6.1　作出矩阵 $\boldsymbol{A}$ 的奇异值分解

$$\boldsymbol{A}=\begin{bmatrix}0 & 1\\ -1 & 0\\ 0 & 2\\ 1 & 0\end{bmatrix}$$

解：因为$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\begin{bmatrix}2 & 0\\ 0 & 5\end{bmatrix}$，则 $\boldsymbol{A}$ 的非零奇异值为$\sqrt{5},\sqrt{2}$，$\boldsymbol{\Sigma}=\begin{bmatrix}\sqrt{5} & 0\\ 0 & \sqrt{2}\end{bmatrix}$，对应于特征值5和2的标准正交特征向量为 $u_1=\begin{bmatrix}0\\ 1\end{bmatrix}, u_2=\begin{bmatrix}1\\ 0\end{bmatrix}$,则$\boldsymbol{U}=\boldsymbol{U}_1=[u_1\ \ u_2]=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$，构造$\boldsymbol{V}_1=\boldsymbol{A}\boldsymbol{U}_1\boldsymbol{\Sigma}^{-1}=\begin{bmatrix}\frac{1}{\sqrt{5}} & 0\\ 0 & -\frac{1}{\sqrt{2}}\\ \frac{2}{\sqrt{5}} & 0\\ 0 & \frac{1}{\sqrt{2}}\end{bmatrix}$，再取$\boldsymbol{V}_2=\begin{bmatrix}-\frac{2}{\sqrt{5}} & 0\\ 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{5}} & 0\\ 0 & \frac{1}{\sqrt{2}}\end{bmatrix}$使得$\boldsymbol{V}=[\boldsymbol{V}_1\ \ \boldsymbol{V}_2]$为酉矩阵。于是 $\boldsymbol{A}$ 的奇异值分解为

$$\boldsymbol{A}=\boldsymbol{V}\begin{bmatrix}\boldsymbol{\Sigma}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{U}^{\mathrm{H}}$$

即

$$\boldsymbol{A}=\begin{bmatrix}0 & 1\\ -1 & 0\\ 0 & 2\\ 1 & 0\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} & 0\\ 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ \frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} & 0\\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\sqrt{5} & 0\\ 0 & \sqrt{2}\\ 0 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}^{\mathrm{H}}$$

Thank you!