复习：2 个许尔公式（Schur）

许尔公式(1)：方阵 $\mathbf{A}_{n\times n}$ 存在可逆阵 $\mathbf{P}$ 使 $\mathbf{P}^{-1}\mathbf{AP}=\mathbf{D}=\begin{pmatrix}\lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n\end{pmatrix}$ 为上三角阵；

许尔公式(2)：方阵 $\mathbf{A}_{n\times n}$ 存在酉阵 $\mathbf{Q}$，使 $\mathbf{Q}^{-1}\mathbf{AQ}=\mathbf{D}=\begin{pmatrix}\lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n\end{pmatrix}$ 为上三角阵.

证明略.

例 1：$\mathbf{A}=\begin{pmatrix}1 & 3\\ 2 & 2\end{pmatrix}$，令 $\mathbf{P}=\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$，有 $\mathbf{P}^{-1}=\frac{1}{2}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}$，

$$\mathbf{P}^{-1}\mathbf{AP}=\frac{1}{2}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\begin{pmatrix}1 & 3\\ 2 & 2\end{pmatrix}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}=\frac{1}{2}\begin{pmatrix}3 & 5\\ 1 & -1\end{pmatrix}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}=\frac{1}{2}\begin{pmatrix}8 & 2\\ 0 & -2\end{pmatrix}=\begin{pmatrix}4 & 1\\ 0 & -1\end{pmatrix};$$

令 $\mathbf{P}=\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}$，有 $\mathbf{P}^{-1}=\begin{pmatrix}2 & -1\\ -1 & 1\end{pmatrix}$，

$$\mathbf{P}^{-1}\mathbf{AP}=\begin{pmatrix}2 & -1\\ -1 & 1\end{pmatrix}\begin{pmatrix}1 & 3\\ 2 & 2\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}=\begin{pmatrix}0 & 4\\ 1 & -1\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}=\begin{pmatrix}4 & 8\\ 0 & -1\end{pmatrix}。$$

例 2：$\mathbf{A}=\begin{pmatrix}2+\mathrm{i} & 1\\ 1 & 2-\mathrm{i}\end{pmatrix}$，令 $\mathbf{Q}=\frac{1}{\sqrt{2}}\begin{pmatrix}\mathrm{i} & 1\\ 1 & \mathrm{i}\end{pmatrix}$，有 $\mathbf{Q}^{-1}=\mathbf{Q}^{H}=\frac{1}{\sqrt{2}}\begin{pmatrix}-\mathrm{i} & 1\\ 1 & -\mathrm{i}\end{pmatrix}$，

$$\mathbf{Q}^{-1}\mathbf{AQ}=\frac{1}{\sqrt{2}}\begin{pmatrix}-\mathrm{i} & 1\\ 1 & -\mathrm{i}\end{pmatrix}\begin{pmatrix}2+\mathrm{i} & 1\\ 1 & 2-\mathrm{i}\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}\mathrm{i} & 1\\ 1 & \mathrm{i}\end{pmatrix}=\frac{1}{2}\begin{pmatrix}2-2\mathrm{i} & 2-2\mathrm{i}\\ 2 & -2\mathrm{i}\end{pmatrix}\begin{pmatrix}\mathrm{i} & 1\\ 1 & \mathrm{i}\end{pmatrix}=\frac{1}{2}\begin{pmatrix}4 & 4\\ 0 & 4\end{pmatrix}=\begin{pmatrix}2 & 2\\ 0 & 2\end{pmatrix}$$

许尔推论：　每个方阵都酉相似于上三角阵.

Hermit 分解：若 $\mathbf{A}$ 是 Hermite 阵（$\mathbf{A}^{H}=\mathbf{A}$），则存在优阵 $\mathbf{Q}$，

使 $\mathbf{Q}^{-1}\mathbf{AQ}=D=\begin{pmatrix}\lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n\end{pmatrix}$ 为对角阵

且，特征根 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 都为实数

注：$\mathbf{Q}$ 中列都是 $\mathbf{A}^{H}\mathbf{A}$ 的特征向量

证：据许尔公式 $\mathbf{Q}^{-1}\mathbf{AQ}=\begin{pmatrix}\lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n\end{pmatrix}$ 为上三角阵，$\mathbf{Q}$ 为优阵：$Q^{-1}=Q^{H}$

共轭转置后: $(\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q})^H = \begin{pmatrix} \overline{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ \overline{*} & \cdots & \overline{\lambda_n} \end{pmatrix}$为下三角阵，且可写左边如下:

$(\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q})^H = (\mathbf{Q}^H\mathbf{A}\mathbf{Q})^H = \mathbf{Q}^H\mathbf{A}^H\mathbf{Q} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$，即 $\begin{pmatrix} \lambda_1 & \cdots & (*) \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} = \begin{pmatrix} \overline{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ \overline{*} & \cdots & \overline{\lambda_n} \end{pmatrix}$

$\therefore$所有的元素（$*$）都为零，且 $\lambda_i = \overline{\lambda_i}$，即每个 λ_i 为实数（$i = 1,2,3,\cdots,n$）.

由 Hermit 分解可得如下结论:

（$\mathbf{A}^H\mathbf{A}$）**定理**: 任 $\mathbf{A}_{m\times n}$，则 $\mathbf{A}^H\mathbf{A}$ 与 $\mathbf{A}\mathbf{A}^H$ 都为 Hermite 阵.

且对于 $\mathbf{A}^H\mathbf{A}$，存在酉阵 $\mathbf{Q}$（$\mathbf{Q}^{-1} = \mathbf{Q}^H$），使:

$$\mathbf{Q}^{-1}(\mathbf{A}^H\mathbf{A})\mathbf{Q} = \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$ 为对角形

备注: $\mathbf{Q}$ 中列都是 $\mathbf{A}^H\mathbf{A}$ 的特征向量，对应特征根 $\lambda_1, \lambda_2, \cdots, \lambda_n$

例: $\mathbf{A} = \begin{pmatrix} 0 & \mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix}$ 为 hermite（$\boldsymbol{A}^H = \boldsymbol{A}$），求 $\mathbf{Q}$ 使 $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ 为对角

解: 求得 $\lambda(\mathbf{A}) = \{1, -1\}$ 都为实根，

可知特根1对应特征向量为 $(1 \quad -\mathrm{i})^{\mathrm{T}}$，特根 -1 对应特向为 $(-\mathrm{i} \quad 1)^{\mathrm{T}}$，

令 $\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -\mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix}$ 为酉阵，则 $\mathbf{Q}^{-1} = \mathbf{Q}^H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}$

可知 $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}\begin{pmatrix} 0 & \mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -\mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ 为对角形

补例: $\mathbf{B} = \begin{pmatrix} -1 & -3 & 3 & -3 \\ -3 & -1 & -3 & 3 \\ 3 & -3 & -1 & -3 \\ -3 & 3 & -3 & -1 \end{pmatrix}$ 为 hermit（$\boldsymbol{B}^H = \boldsymbol{B}$），求 $\mathbf{Q}$ 使 $\mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}$ 为对角形.

（令 $\mathbf{A}=\begin{pmatrix}3 & -3 & 3 & -3\\ -3 & 3 & -3 & 3\\ 3 & -3 & 3 & -3\\ -3 & 3 & -3 & 3\end{pmatrix}$ 为秩 1 阵 ）

解：因为 $\mathbf{B}+4\mathbf{I}=\mathbf{A}$ 为秩 1 阵，求得：$\lambda(\mathbf{A})=\{12,0,0,0\}$，特征根 12 对应的特征向量为 $(1\quad -1\quad 1\quad -1)^{\mathrm{T}}$，

可知 $\mathbf{B}=\mathbf{A}-4\mathbf{I}$ 利用"平移法"可知，$\mathbf{B},\ \mathbf{A}$ 必有有相同特征向量!

$\lambda(\mathbf{B})=\{8,-4,-4,-4\}$，且特根 8 对应特向为 $(1\quad -1\quad 1\quad -1)^{\mathrm{T}}$。

令 $\mathbf{Q}=\frac{1}{2}\begin{pmatrix}1 & 1 & 1 & -1\\ -1 & -1 & 1 & -1\\ 1 & -1 & 1 & 1\\ -1 & 1 & 1 & 1\end{pmatrix}$，$\mathbf{Q}$ 为优阵，则 $\mathbf{Q}^{-1}=\mathbf{Q}^{H}=\frac{1}{2}\begin{pmatrix}1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\\ 1 & 1 & 1 & 1\\ -1 & -1 & 1 & 1\end{pmatrix}$，

可得：

$$\mathbf{Q}^{-1}\mathbf{B}\mathbf{Q}=D=\begin{pmatrix}8 & 0 & 0 & 0\\ 0 & -4 & 0 & 0\\ 0 & 0 & -4 & 0\\ 0 & 0 & 0 & -4\end{pmatrix}\text{为对角}$$

....

由 Hermit 分解可得如下结论：

（$\mathbf{A}^{H}\mathbf{A}$）引理：任 $\mathbf{A}_{m\times n}$，则 $\mathbf{A}^{H}\mathbf{A}$ 与 $\mathbf{A}\mathbf{A}^{H}$ 都为 Hermite 阵.

且对于 $\mathbf{A}^{H}\mathbf{A}$，存在酉阵 $\mathbf{Q}$（$\mathbf{Q}^{-1}=\mathbf{Q}^{H}$），使：

$$\mathbf{Q}^{-1}(\mathbf{A}^{H}\mathbf{A})\mathbf{Q}=\mathbf{D}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}\text{为对角形}$$

且 $\mathbf{Q}$ 中列都是 $\mathbf{A}^{H}\mathbf{A}$ 的特征向量，对应特征根 $\lambda_1,\lambda_2,\cdots,\lambda_n$

备注：利用换位公式推论 "$\mathbf{AB}$ 与 $\mathbf{BA}$ 必有相同的非 0 根"

可知 $A^{H}A,\ AA^{H}$ 必有相同的非 0 特根!

又 $A^{H}A,\ AA^{H}$ 都是 hermite：$(A^{H}A)^{H}=A^{H}A,\qquad (AA^{H})^{H}=AA^{H}$

且 $A^H A \ge 0$，$AA^H \ge 0$ 都是半正定阵，它们的非 0 根都是正根.

可知 $A^H A$，AA^H 必有相同的正根

结论：设秩 rank(A)= r >0，必有 r($A^H A$)= r(AA^H)=r(A) = r(秩公式)

且 $A^H A$，AA^H 必有相同的 r 个正根，

可写全体根 $\lambda\left(A^H A\right)=\{\lambda_1,\cdots,\lambda_r,0,0,\cdots,0_{n-r}\}$，其中

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_r > 0，\quad \lambda_{r+1}=\lambda_{r+2}=\cdots=\lambda_n=0$$

必有 hermit 分解：$Q^H(A^H A)Q=\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{pmatrix}$，Q 为酉阵

其中 $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_r > 0$，$\lambda_{r+1}=\lambda_{r+2}=\cdots=\lambda_n=0$，r=rank(A)

$r<n$ 时，可写分解：

$$Q^H(A^H A)Q=\begin{pmatrix}D_r & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & \cdots & \cdots & 0\end{pmatrix}_{n\times n},\quad D_r=\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_r\end{pmatrix}$$

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_r > 0$$

奇异值定义：给定 $\mathbf{A}_{m\times n}$，则 $\mathbf{A}^H\mathbf{A}$ 与 $\mathbf{A}\mathbf{A}^H$ 有相同正根：$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_r > 0$，$r=r(\mathbf{A})$，

$\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_r}$ 叫做 $\mathbf{A}$ 的正奇异值.

全体正奇异值记作 $s_+(A)=\{\sqrt{\lambda_1},\cdots,\sqrt{\lambda_r}\}$

又记为 $s_+(A)=\{s_1,s_2\cdots,s_r\}$，$s_1=\sqrt{\lambda_1},\cdots,s_r=\sqrt{\lambda_r}$

可按大小的顺序排列：$s_1 \geqslant s_2 \geqslant \cdots \geqslant s_r$，称 $\boldsymbol{s_1}$ 为 $\mathbf{A}$ 的最大奇异值

备注，对于 n 阶方阵 $A=A_{n\times n}$，则 $\mathbf{A}^H\mathbf{A}$ 与 $\mathbf{A}\mathbf{A}^H$ 有 n 个相同非负根

记为 $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n \ge 0$

此时，$\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n}$ 叫做 $\mathbf{A}$ 的全体奇异值(含 0 奇异值)

全体奇异值记为 $s(A)=\{\, s_1, s_2 \cdots,\ s_n\ \}=\{\ \sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n}\ \}$

即，n 阶方阵 $A=A_{n\times n}$ 恰有 n 个奇异值：$\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n}$

其中 $\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_r>0$，$\lambda_{r+1}=\lambda_{r+2}=\cdots=\lambda_n=0$，r=rank(A)

………………………………………………………………………………………………….

必有分解：$Q^H(A^HA)Q=\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{pmatrix}$,Q 为酉阵

例：求下列正奇异值，①$\mathbf{A}=\begin{pmatrix}1&2\\0&0\\0&0\end{pmatrix}$；②$\mathbf{A}=\begin{pmatrix}1&1\\0&0\\1&1\end{pmatrix}$。

解：①$\mathbf{A}^H\mathbf{A}=\begin{pmatrix}1&0&0\\2&0&0\end{pmatrix}\begin{pmatrix}1&2\\0&0\\0&0\end{pmatrix}=\begin{pmatrix}1&2\\2&4\end{pmatrix}$为秩 1 阵，可知

$\lambda(A^HA)=\{tr(A^HA),\ \ 0\}=\{5,\ 0\}$，可令 $\lambda_1=5,\ \ \lambda_2=0,\therefore$ 正奇异值为 $s_1=\sqrt{\lambda_1}=\sqrt{5}$

② $\mathbf{A}^H\mathbf{A}=\begin{pmatrix}1&0&1\\1&0&1\end{pmatrix}\begin{pmatrix}1&1\\0&0\\1&1\end{pmatrix}=\begin{pmatrix}2&2\\2&2\end{pmatrix}$为秩 1 阵，可知 $\lambda(A^HA)=\{tr(A^HA),\ \ 0\}=\{4,\ 0\}$

令 $\lambda_1=5,\ \ \lambda_2=0,$，正奇异值为 $s_1=\sqrt{\lambda_1}=\sqrt{4}=2$.

练习 Ex1：求正奇异值 $(1)\mathbf{A}=\begin{pmatrix}2&3\\0&0\\0&0\end{pmatrix}$; $(2)\mathbf{A}=\begin{pmatrix}i&i\\0&0\\1&1\end{pmatrix}$, $i^2=-1$； $(3)\mathbf{A}=\begin{pmatrix}i\\i\\1\end{pmatrix}$; $(4)\mathbf{A}=\begin{pmatrix}1&1&i\end{pmatrix}$

Ex2 求方阵的全体奇异值 $s(A)=\{\, s_1, s_2\cdots,\ s_n\ \}$与特征值 $\lambda(A)=\{t_1,t_2,\cdots t_n\}$

$(1)A=\begin{pmatrix}1&1\\1&1\end{pmatrix}$; $(2)A=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$; $(2)A=\begin{pmatrix}i&1\\1&i\end{pmatrix}$, $i^2=-1$; $(4)A=-\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}$

Ex3 设矩阵 $A=A_{m\times n}$ 全体正奇异值为 $s_+(A)=\{\, s_1, s_2\cdots,\ s_r>0\}$，r=rank(A)

证明：$s_1^2+s_2^2+\cdots+s_r^2=tr(A^HA)$; $s_1^2+s_2^2+\cdots+s_r^2=\sum|a_{i,j}|^2$