3.5 正规矩阵

定理（Schur 定理）

任一复方阵均可酉相似于上三角矩阵.

证明：（只需把第一章中的证明所用的基底取成标准正交基即可.）

证法二：已知对任意 $A \in \mathbb{C}^{n\times n}$， 存在非奇异矩阵 $P \in \mathbb{C}^{n\times n}$ 使得

$$P^{-1}AP = J$$

为上三角矩阵. 现对 P 的列向量组实施 Gram-schmidt 标准正交化，即对 P 作正交-三角分解如下

$$P = UT ,$$

其中 U 为酉矩阵， T 是(正实数对角线的)上三角矩阵. 于是

定义 （正规矩阵）一个复方阵若与它的共轭转置矩阵可交换，则称为正规矩阵．即 $A \in \mathbb{C}^{n\times n}$ 称为正规矩阵，若

$$AA^{\mathrm{H}} = A^{\mathrm{H}}A\ .$$

例 （正规矩阵的典型例子）

Hermit 矩阵（对称矩阵）；反 Hermit 矩阵（反对称矩阵）；酉矩阵（正交矩阵）

矩阵正规性的酉相似不变性：设 A 是正规矩阵，且 A 酉相似于 B，则 B 也是正规矩阵．

上三角的正规矩阵是对角矩阵：上三角矩阵若还是正规矩阵，则必为对角矩阵．

定理 $A \in \mathbb{C}^{n\times n}$ 可酉相似对角化，当且仅当 A 是正规矩阵．

回忆，此定理等价于说，n 阶复方阵有 n 个相互正交的特征向量的充要条件，是该矩阵为正规矩阵.

定理（三种特殊正规矩阵的特征值）设 A 是正规矩阵，则

（1） A 是 Hermit 矩阵的充要条件是 A 的特征值是实数.

（2） A 是反 Hermit 矩阵的充要条件是 A 的特征值的实部为零.

（2） A 是酉矩阵的充要条件是 A 的特征值的模等于 1.

正规矩阵酉相似对角化的具体求法

设 A 是 n 阶正规矩阵. 由上定理知

(1) A 的代数重数等于 A 的几何重数. 也就是说, 特征根 c 是特征多项式 $|\lambda I_n - A|$ 的 r 重根, 等价于齐次线性方程组

$$(cI_n - A)x = 0$$

的解空间 $\ker(cI_n - A)$ (称为 c 的特征子空间) 的维数是 r ; 或者说相应于 r 重特征根 c , 恰有 r 个线性无关的特征向量. 更简洁地

$$r = \dim(\ker(cI_n - A)) = n - \operatorname{rank}(cI_n - A).$$

(2) 相应于不同特征值的特征子空间相互正交. 设 λ_1 , λ_2 是两个不同的特征值, 则

$$\ker(\lambda_1 I_n - A) \perp \ker(\lambda_2 I_n - A).$$

问题：已知n阶正规矩阵A，求对角矩阵Λ及酉矩阵U，使得$U^H AU = \Lambda$.

解法：

第一步. 求出$|\lambda I_n - A|$的所有不同的根，记为$\lambda_1, \ldots, \lambda_p$.

第二步. 对每个λ_i，求齐次线性方程组$(\lambda_i I_n - A)x = 0$的一个基础解系，记为

$$v_1^i, \ldots, v_{r_i}^i.$$

第三步. 对向量组$v_1^i, \ldots, v_{r_i}^i$实施Gram-Schmidt标准正交化，得

$$u_1^i, \ldots, u_{r_i}^i.$$

第四步. 拼出Λ及U如下：

$$\Lambda = \operatorname{diag}\{\lambda_i I_{r_i}, i = 1, \ldots, p\},$$

$$U = \begin{bmatrix} v_1^1 & \cdots & v_{r_1}^1 & \cdots & v_1^i & \cdots & v_{r_i}^i & \cdots & v_1^p & \cdots & v_{r_p}^p \end{bmatrix}.$$

3.6 Hermit 矩阵与 Hermit 二次型

定理（Hermit 矩阵的内积刻画） 设 $A \in \mathbb{C}^{n\times n}$. 则下列条件等价

（1） $A^{\mathrm{H}} = A$， 即 A 为 Hermit 矩阵.

（2） 对任意的 $x, y \in \mathbb{C}^n$， $\langle Ax, y\rangle = \langle x, Ay\rangle$.

定义（Hermit 二次型）

给定 n 阶 Hermit 矩阵 A. 映射 $x \mapsto \langle Ax, x\rangle = x^{\mathrm{H}}Ax$ 确定从 $\mathbb{C}^n$ 到 $\mathbb{R}$ 的函数，称为 Hermit 二次型. 写成分量形式，有

$$x^{\mathrm{H}}Ax = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{x}_i x_j .$$

定义（Hermit 矩阵的合同）
设 A 和 B 是两个 n 阶 Hermit 矩阵，P 是 n 阶非奇异矩阵．若有

$$P^{\mathrm{H}} AP = B,$$

则称 A 和 B 关于 P 合同．此时也简称 A 与 B 合同，并称 P 为相应的合同矩阵或合同变换矩阵．

注（Hermit 矩阵的酉相似也是合同）

注（Hermit 二次型在坐标变换下的动态） Hermit 二次型与坐标变换两个概念的结合，导致 Hermit 矩阵的合同概念．向量 $x \in \mathbb{C}^n$ 在 $\mathbb{C}^n$ 的标准基 I_n 下的坐标就是 x 自己：$x = I_n x$．它在 $\mathbb{C}^n$ 的新的基底 P 下的坐标 $\tilde{x}$ 满足：$x = P\tilde{x}$．于是

$$x^{\mathrm{H}} Ax = (P\tilde{x})^{\mathrm{H}} A(P\tilde{x}) = \tilde{x}^{\mathrm{H}} \underbrace{\left(P^{\mathrm{H}} AP\right)}_{=:B} \tilde{x}.$$

由 Hermit 是正规矩阵，可找到酉矩阵 U，将 Hermit 二次型 $x^{\mathrm{H}}Ax$ 通过正交坐标变换 $x=U\tilde{x}$ 化为“平方和”形式

$$\begin{aligned} x^{\mathrm{H}}Ax &= \tilde{x}^{\mathrm{H}}(U^{\mathrm{H}}AU)\tilde{x} \\ &= \lambda_1\bar{x}_1x_1+\lambda_2\bar{x}_2x_2+\cdots+\lambda_n\bar{x}_nx_n, \\ &= \lambda_1|x_1|^2+\lambda_2|x_2|^2+\cdots+\lambda_n|x_n|^2 \end{aligned}$$

这里，$\lambda_1,\lambda_2,\ldots,\lambda_n$ 是 A 的 n 个特征值.

Hermit 正定矩阵与 Hermit 正定二次型

定义（正定矩阵与正定二次型）

给定 Hermit 矩阵 A，如果对任意 $x \in \mathbb{C}^n$，$x \neq 0$，有 $x^{\mathrm{H}} Ax > 0$（≥ 0），则称 $x^{\mathrm{H}} Ax$ 为 Hermit 正定(半正定)二次型，A 为 Hermit 正定(半正定)矩阵.

正定(半正定)矩阵的特征值为正(非负)实数；正定(半正定)矩阵开平方

应用于向量组的 Gram 矩阵：任意向量组均可改造成正交组.

Hermit 矩阵特征值的极值刻画

事实：Hermit 矩阵 A 的 n 个特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 都是实数，不妨设

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n .$$

又 $\langle Ax, x\rangle = x^{\mathrm{H}} Ax$ 总是实数.

定理（Hermit 矩阵特征值的极值刻画）

$$\lambda_{\min}(A) = \lambda_1 = \min_{x\in\mathbb{C}^n, x\ne 0} \frac{x^{\mathrm{H}} Ax}{x^{\mathrm{H}} x} = \min_{x\in\mathbb{C}^n, \|x\|=1} x^{\mathrm{H}} Ax ,$$

$$\lambda_{\max}(A) = \lambda_n = \max_{x\in\mathbb{C}^n, x\ne 0} \frac{x^{\mathrm{H}} Ax}{x^{\mathrm{H}} x} = \max_{x\in\mathbb{C}^n, \|x\|=1} x^{\mathrm{H}} Ax .$$

注（Hermit 矩阵特征值与 Rayleigh 商）

从 $\mathbb{C}^n \setminus \{0\}$ 到 $\mathbb{R}$ 的函数

$$x \mapsto R(x) = \frac{x^{\mathrm{H}} A x}{x^{\mathrm{H}} x}$$

称为Hermit矩阵 A 的Rayleigh商. A 的其他特征值 $\lambda_2, \ldots, \lambda_{n-1}$ 也可通过 A 的 Rayleigh 商用极值来刻画. 对此我们不再详述.

注（极值刻画将特征多项式求根与二次型求极值联系起来）

矩阵的奇异值分解

引理（标准酉空间中的Gram矩阵）

设 $A \in \mathbb{C}^{m\times n}$， $\operatorname{rank}(A) = r$. 则

（1） $A^{\mathrm{H}}A$ 和 AA^{H} 分别是 n 阶和 m 阶 Hermit 半正定矩阵.

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

定理（矩阵的奇异值分解）

设 $A \in \mathbb{C}^{m\times n}$. 则存在 m 阶酉矩阵 U 和 n 阶酉矩阵 V，使得

$$U^{H}AV=\begin{bmatrix}\Sigma_r & 0_{r\times(n-r)}\\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)}\end{bmatrix},$$

其中 $r=\operatorname{rank}(A)$，r 阶对角矩阵

$$\Sigma_r=\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix},\quad \sigma_i>0,\quad i=1,\dots,r.$$

推论 $A^{H}A$ 与 AA^{H} 有相同的非零（正实数）特征值

$$\lambda_i=\lambda_i(A^{H}A)=\lambda_i(AA^{H}),\quad i=1,\dots,r.$$

且

$$\sigma_i=\sqrt{\lambda_i}>0,\quad i=1,\dots,r.$$

注（奇异值分解的几何意义）

由 $A\in\mathbb{C}^{m\times n}$ 决定的从 $\mathbb{C}^n$ 到 $\mathbb{C}^m$ 的线性映射 $x\mapsto y=Ax$，在 $\mathbb{C}^n$ 中的标准正交基 V（入口基）和 $\mathbb{C}^m$ 中的标准正交基 U（出口基）下的矩阵表示为

$$\begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}.$$

以下讨论奇异值的几何意义和极值刻画.

定义（高维空间中的椭球面）（复的也完全可以!）

设 S 是 n 阶 Hermit 正定矩阵，c 是大于 0 的实数. 则 $\mathbb{C}^n$ 中的点集

$$\left\{x \mid x \in \mathbb{C}^n, x^{\mathrm{H}} S x = c\right\}$$

称为 $\mathbb{C}^n$ 中的一个椭球面.

$$\left\{x \mid x \in \mathbb{C}^n, x^{\mathrm{H}} S x \leq c\right\}$$

称为椭球体.

定理（线性映射把椭球映射成椭球）

设 A 是 n 阶非奇异矩阵. 则任一椭球面在映射 $x \mapsto y = Ax$ 下的像也是椭球.

设 A 是 n 阶非奇异矩阵，记 A 的奇异值分解为

$$U^{\mathrm{H}}AV=\Sigma=\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_n\end{bmatrix}.$$

现计算单位球

$$\{x \mid x\in\mathbb{C}^n, \|x\|=1\}$$

在 A 下的像．由

$$\begin{aligned}\|x\|^2 &= \|A^{-1}y\|^2=\|(U\Sigma V^{\mathrm{H}})^{-1}y\|^2\\ &= \|V\Sigma^{-1}(U^{\mathrm{H}}y)\|^2=\|\Sigma^{-1}\tilde{y}\|^2\\ &= \frac{|\tilde{y}_1|^2}{\sigma_1^2}+\frac{|\tilde{y}_2|^2}{\sigma_2^2}+\cdots+\frac{|\tilde{y}_n|^2}{\sigma_n^2},\end{aligned}$$

这里 $\tilde{y}=U^{\mathrm{H}}y$，即 $y=U\tilde{y}$.

也就是说，在新的标准正交基U下，单位球的像的方程是标准的椭球面方程

$$\frac{|\tilde{y}_1|^2}{\sigma_1^2}+\frac{|\tilde{y}_2|^2}{\sigma_2^2}+\cdots+\frac{|\tilde{y}_n|^2}{\sigma_n^2}=1.$$

由于

$$U\Sigma=AV\ ,\quad U=AV\Sigma^{-1}=\begin{bmatrix}Av_1\frac{1}{\sigma_1} & Av_2\frac{1}{\sigma_2} & \cdots & Av_n\frac{1}{\sigma_n}\end{bmatrix},$$

可知，

$$\sigma_i=\|Av_i\|,\quad i=1,2,\ldots,n.$$

换言之，标准正交出口基$U=\begin{bmatrix}u_1 & u_2 & \cdots & u_n\end{bmatrix}$就是标准正交入口基$V=\begin{bmatrix}v_1 & v_2 & \cdots & v_n\end{bmatrix}$在$A$下的像的单位化.

因此，若σ_1是最大奇异值，则v_1就是单位球面上被A映的最远的向量：映射成到椭球面最长的半轴$Av_1=u_1\sigma_1$.