1 奇异值分解

为了引入奇异值的概念, 我们先证明两个引理.

Lemma 1. 设 $A \in \mathbb{C}^{m\times n}$, 则

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

证明. 利用 $Ax = 0$ 与 $A^H Ax = 0$ 同解以及 $Ax = 0$ 与 $AA^H x = 0$ 同解立刻可得. □

Lemma 2. 设 $A \in \mathbb{C}^{m\times n}$, 则

(1)$A^H A$ 与 AA^H 的特征值均为非负实数*;*

(2)$A^H A$ 与 AA^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 $\mathrm{rank}(A)$.

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

$$A^H Ax = \lambda x$$

因为 $A^H A$ 是 Hermite 矩阵, 所以 λ 是实数, 并且

$$\lambda x^H x = x^H A^H Ax = (Ax)^H Ax \geq 0$$

由于 $x^x > 0$, 所以 $\lambda \geq 0$.

同理可证 AA^H 的特征值军费非负实数.

(2) 由线性代数知识, 易知方阵 AB 与方阵 BA 的非零特征值相同, 则 $A^H A$ 与 AA^H 的非零特征值相同, 并且它们的非零特征值个数为

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

□

Theorem 1. 设 A 是 $m \times n$ 矩阵, 且 $\mathrm{rank}(A) = r$, 则存在 m 阶酉矩阵 V 和 n 阶酉矩阵 U 使得

$$V^H AU = \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}$$

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

证明. 因为 $\mathrm{rank}(A) = r$, 由 Lemma2 可设 $A^H A$ 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 $A^H A$ 是 Hermite 矩阵, 由谱分解定理可知存在 n 阶酉矩阵 U 使得

$$U^H A^H AU = \begin{pmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{pmatrix}$$

记 $U = [U_1, U_2]$, 其中 U_1 是 $n \times r$ 矩阵, 上式可改写为

$$A^H A[U_1, U_2] = [U_1, U_2] \begin{pmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{pmatrix}$$

则有

$$A^H AU_1 = U_1 \Sigma^2, A^H AU_2 = 0$$

记 $V = [V_1, V_2]$, 其中 V_1 是 $m \times r$ 矩阵,V_2 是 $m \times (m - r)$ 矩阵. 令

$$V_1 = AU_1 \Sigma^{-1}$$

则

$$V_1^H V = (\Sigma^{-1})^H U_1^H A^H AU_1 \Sigma^{-1} = I_r$$

那么我们可以找到一个 V_2 使得 $V = [V_1, V_2]$ 是酉矩阵, 则

$$V_2^H AU_1 = V_2^H V_1 \Sigma = 0$$

那么

$$V^H AU = \begin{pmatrix} V_1^H \\ V_2^H \end{pmatrix} A[U_1, U_2] = \begin{pmatrix} V_1^H AU_1 & V_1^H AU_2 \\ V_2^H AU_1 & V_2^H AU_2 \end{pmatrix} = \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}$$

□

2 矩阵级数

Definition 1. 设有矩阵序列 $\{A^{(k)}\}$, 若当 $k \to +\infty$ 时, 有 $a_{ij}^{(k)} \to a_{ij}$, 则称 $\{A^{(k)}\}$ 收敛, 并将 $A = (a_{ij})$ 称为 $\{A^{(k)}\}$ 的极限, 记为 $\lim_{k\to+\infty} A^{(k)} = A$. 不收敛的矩阵序列称为发散的.

Lemma 3. *Definition 1* 与矩阵范数意义下的收敛性是等价的