§5 矩阵的奇异值分解

程光辉

2019 年 11 月 15 日

定理 1 设 $\boldsymbol{A} \in \boldsymbol{C}_r^{m\times n}$，则有

(1) $\mathbf{rank}(\boldsymbol{A}) = \mathbf{rank}(\boldsymbol{A}^H\boldsymbol{A}) = \mathbf{rank}(\boldsymbol{A}\boldsymbol{A}^H)$

(2) $\boldsymbol{A}^H\boldsymbol{A}, \boldsymbol{A}\boldsymbol{A}^H$ 的特征值均为非负实数

(3) $\boldsymbol{A}^H\boldsymbol{A}, \boldsymbol{A}\boldsymbol{A}^H$ 的非零特征值相同.

证明： (1) 因为 $\mathbf{rank}(\boldsymbol{A}^H\boldsymbol{A}) \leq \mathbf{rank}(\boldsymbol{A})$，只需证明 $\mathbf{rank}(\boldsymbol{A}^H\boldsymbol{A}) \geq \mathbf{rank}(\boldsymbol{A})$ 即可.

令 $\mathbf{rank}(\boldsymbol{A}^H\boldsymbol{A}) = \boldsymbol{r}$，则 $\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ 的解空间为 $\boldsymbol{n} - \boldsymbol{r}$ 维，记为 $\boldsymbol{X}$. 对 $\forall \boldsymbol{x}_1 \in \boldsymbol{X}$，都有

$$\boldsymbol{x}_1^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x}_1 = (\boldsymbol{A}\boldsymbol{x}_1)^H(\boldsymbol{A}\boldsymbol{x}_1) = \mathbf{0},$$

即

$$\boldsymbol{A}\boldsymbol{x}_1 = \mathbf{0}.$$

也就是说 $\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ 的解空间是 $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ 的解空间的子空间，进而有

$$\mathbf{rank}(\boldsymbol{A}^H\boldsymbol{A}) \geq \mathbf{rank}(\boldsymbol{A}),$$

得证.

(2) 设 $\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{\alpha} = \boldsymbol{\lambda}\boldsymbol{\alpha}$，其中 $\boldsymbol{\lambda}$ 为任意特征值，$\boldsymbol{\alpha}$ 为对应的特征向量. 于是有

$$(\boldsymbol{A}\boldsymbol{\alpha}, \boldsymbol{A}\boldsymbol{\alpha}) = (\boldsymbol{\alpha}, \boldsymbol{A}^H\boldsymbol{A}\boldsymbol{\alpha}) = (\boldsymbol{\alpha}, \boldsymbol{\lambda}\boldsymbol{\alpha}) = \boldsymbol{\lambda}(\boldsymbol{\alpha}, \boldsymbol{\alpha}) \geq \mathbf{0},$$

即 $\boldsymbol{\lambda} \geq \mathbf{0}$.

(3) 设 $\boldsymbol{A}^H\boldsymbol{A}$ 的特征值为

$$\boldsymbol{\lambda}_1 \geq \boldsymbol{\lambda}_2 \geq \cdots \geq \boldsymbol{\lambda}_r > \boldsymbol{\lambda}_{r+1} = \cdots = \boldsymbol{\lambda}_n = \mathbf{0},$$

对应的特征向量分别为 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$.

因为 $A^HA\alpha_i=\lambda_i\alpha_i\ (\neq 0)$，$i=1,2,\cdots,r$，则

$$(AA^H)A\alpha_i=A(A^HA\alpha_i)=A(\lambda_i\alpha_i)=\lambda_i(A\alpha_i),$$

则 A^HA 的非零特征值 λ_i，$i=1,2,\cdots,r$ 也是 AA^H 的非零特征值，对应的特征向量为 $A\alpha_i$.

同理可证，AA^H 的非零特征值也是 A^HA 的非零特征值.

设 $y_1,\cdots,y_p$ 是 A^HA 的非零特征值 λ 对应的特征子空间 V_λ 的一组基，则有 $Ay_1,\cdots,Ay_p$ 是 AA^H 的非零特征值 λ 对应的特征向量.

令

$$k_1Ay_1+k_2Ay_2+\cdots+k_pAy_p=0,$$

左乘 A^H，有

$$k_1A^HAy_1+k_2A^HAy_2+\cdots+k_pA^HAy_p=\lambda(k_1y_1+k_2y_2+\cdots+k_py_p)=0,$$

即只有 $k_1=k_2=\cdots=k_p=0$，则 $Ay_1,\cdots,Ay_p$ 线性无关. 也就是说，A^HA 的特征子空间 V_λ 的维数小于等于 AA^H 的特征子空间 V_λ 的维数.

同理可证，AA^H 的特征子空间 V_λ 的维数小于等于 A^HA 的特征子空间 V_λ 的维数.

综上，得证.

定义 1 设 $A\in C_r^{m\times n}$，A^HA 的特征值为

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

则称 $\sigma_i=\sqrt{\lambda_i}\quad(i=1,2,\cdots,r)$ 为 A 的正奇异值.

定义 2 设 $A,B\in C^{m\times n}$，如果存在酉矩阵 $U\in U^{m\times m}$ 和 $V\in U^{n\times n}$，使得

$$A=UBV,$$

则称 A 与 B 酉等价.

定理 2 若 A 与 B 酉等价，则 A 与 B 有相同的正奇异值.

证明: 因为 A 与 B 酉等价，则存在酉矩阵 $U\in U^{m\times m}$ 和 $V\in U^{n\times n}$，使得 $A=UBV$，进而有

$$AA^H=UBV(UBV)^H=UBVV^HB^HU^H=UBB^HU^H,$$

所以 AA^H 和 BB^H 相似，有相同的特征值，即 A 与 B 有相同的正奇异值.

定理 3 设 $A \in C_r^{m\times n}$，$\sigma_1, \sigma_2, \cdots, \sigma_r$ 是矩阵 A 的 r 个正奇异值，则存在酉矩阵 $U \in U^{m\times m}$ 和 $V \in U^{n\times n}$，使得

$$A = U \begin{bmatrix} D & O \\ O & O \end{bmatrix} V^H,$$

其中 $D = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r)$.

证明：因为 $A^H A$ 为 Hermite 半正定矩阵，故特征值为 $\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2, 0, \cdots, 0$. 令 $v_1, v_2, \cdots, v_n$ 是对应的单位正交特征向量组，而且

$$V_1 = [v_1, v_2, \cdots, v_r]$$
$$V_2 = [v_{r+1}, v_{r+2}, \cdots, v_n].$$

因此，有 $A^H A V_1 = V_1 D^2$，进而

$$D^{-1} V_1^H A^H A V_1 D^{-1} = U_1^H U_1 = E_r,$$

其中 $U_1 = A V_1 D^{-1} \in U_r^{m\times r}$，即 $D = U_1^H A V_1$，$U_1 D = A V_1$.

另有 $A^H A V_2 = V_2 O$，使得 $V_2^H A^H A V_2 = O$，因此 $A V_2 = O$.

选择合适 $U_2 \in U_{(m-r)}^{m\times(m-r)}$，使得 $U = [U_1, U_2]$ 为酉矩阵. 于是

$$U^H A V = \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} A[V_1, V_2] = \begin{bmatrix} U_1^H A V_1 & U_1^H A V_2 \\ U_2^H A V_1 & U_2^H A V_2 \end{bmatrix} = \begin{bmatrix} D & O \\ U_2^H U_1 D & O \end{bmatrix} = \begin{bmatrix} D & O \\ O & O \end{bmatrix},$$

从而有

$$A = U \begin{bmatrix} D & O \\ O & O \end{bmatrix} V^H.$$

推论 1 设 $A \in C_n^{n\times n}$, $\sigma_1, \sigma_2, \cdots, \sigma_n$ 是矩阵 A 的 n 个正奇异值, 则存在酉矩阵 $U, V \in U^{n\times n}$，使得

$$A = U D V^H,$$

其中 $D = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n)$.

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

解：求 $A^H A$ 的特征值和特征向量，

$$A^H A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

则特征值为 $\lambda_1 = 5, \lambda_2 = \lambda_3 = 0$. 对应的特征向量为

$$\alpha_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \alpha_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

因此，有

$$V_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, V_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

构造酉矩阵 U，有

$$U_1 = AV_1D^{-1} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}, U_2 = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}, U = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}.$$

综上，有

$$A = \begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

引理 1 设 $A \in C_r^{n\times n}$ 是半正定矩阵，且 $k \geq 1$ 是给定的整数，则存在唯一的半正定 Hermite 矩阵 B 使得 $B^k = A$. 同时还有

(1) $BA = AB$,

(2) $\mathrm{rank}(A) = \mathrm{rank}(B)$, 进而若 A 正定，则 B 也正定,

(3) 若 A 是实矩阵，则 B 也实矩阵.

定理 4 设 $A \in C_n^{n\times n}$，则必存在酉矩阵 $U \in U^{n\times n}$ 与两个正定 Hermite 矩阵 H_1，H_2，使得

$$A = H_1U = UH_2,$$

且这种分解式是唯一的.

证明： (存在性) 因为 $A \in C_n^{n\times n}$，所以 A^HA 为正定 Hermite 矩阵，则矩阵 A 的奇异值均为正数，其奇异值分解为

$$A = U_1DV_1^H = U_1DU_1^HU_1V_1^H = H_1U,$$

其中 $H_1 = U_1DU_1^H$ 为 Hermite 正定矩阵，$U = U_1V_1^H$ 为酉矩阵.

(唯一性) 令 $\boldsymbol{A}=\boldsymbol{H}_{11}\boldsymbol{U}_1=\boldsymbol{H}_{12}\boldsymbol{U}_2$，其中 $\boldsymbol{H}_{11}$，$\boldsymbol{H}_{12}$ 为正定 **Hermite** 矩阵，$\boldsymbol{U}_1$，$\boldsymbol{U}_2$ 为酉矩阵. 因此，有

$$\begin{gathered}\boldsymbol{H}_{12}^{-1}\boldsymbol{H}_{11}=\boldsymbol{U}_2\boldsymbol{U}_1^H\\ \boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}=\boldsymbol{U}_1\boldsymbol{U}_2^H,\end{gathered}$$

上面两个等式的右侧矩阵互为共轭转置，进而

$$\boldsymbol{H}_{12}^{-1}\boldsymbol{H}_{11}=\left(\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\right)^H=\boldsymbol{H}_{12}^H\left(\boldsymbol{H}_{11}^{-1}\right)^H=\boldsymbol{H}_{12}\boldsymbol{H}_{11}^{-1},$$

即

$$\boldsymbol{H}_{11}^2=\boldsymbol{H}_{12}^2,$$

再由引理 1，知

$$\boldsymbol{H}_{11}=\boldsymbol{H}_{12},\quad \boldsymbol{U}_1=\boldsymbol{U}_2,$$

得证.

推论 2 设 $\boldsymbol{A}\in\boldsymbol{R}_n^{n\times n}$，则必存在唯一的正交矩阵 $\boldsymbol{Q}$ 与两个实对称正定矩阵 $\boldsymbol{H}_1$，$\boldsymbol{H}_2$，使得

$$\boldsymbol{A}=\boldsymbol{H}_1\boldsymbol{Q}=\boldsymbol{Q}\boldsymbol{H}_2.$$

定理 5 设 $\boldsymbol{A}\in\boldsymbol{C}_r^{n\times n}$, 则必存在酉矩阵 $\boldsymbol{U}\in\boldsymbol{U}^{n\times n}$ 与两个半正定 **Hermite** 矩阵 $\boldsymbol{H}_1$，$\boldsymbol{H}_2$，使得

$$\boldsymbol{A}=\boldsymbol{H}_1\boldsymbol{U}=\boldsymbol{U}\boldsymbol{H}_2.$$

证明： 因为 $\boldsymbol{A}\in\boldsymbol{C}_r^{n\times n}$，则由奇异值分解可得

$$\boldsymbol{A}=\boldsymbol{U}_1\begin{bmatrix}\boldsymbol{D}&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{O}\end{bmatrix}\boldsymbol{V}_1^H=\boldsymbol{U}_1\begin{bmatrix}\boldsymbol{D}&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{O}\end{bmatrix}\boldsymbol{U}_1^H\boldsymbol{U}_1\boldsymbol{V}_1^H=\boldsymbol{H}_1\boldsymbol{U},$$

其中 $\boldsymbol{H}_1=\boldsymbol{U}_1\begin{bmatrix}\boldsymbol{D}&\boldsymbol{O}\\ \boldsymbol{O}&\boldsymbol{O}\end{bmatrix}\boldsymbol{U}_1^H$ 为半正定 **Hermite** 矩阵，$\boldsymbol{U}=\boldsymbol{U}_1\boldsymbol{V}_1^H$ 为酉矩阵.