8.5 奇异值分解——酉相抵标准型

高等代数 https://gdfzu.club

定理

设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 和 $\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \ldots, \tilde{\varepsilon}_n$ 是线性空间 V 的两组基，且

$$(\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \ldots, \tilde{\varepsilon}_n) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)P.$$

又设 $\eta_1, \eta_2, \ldots, \eta_m$ 和 $\tilde{\eta}_1, \tilde{\eta}_2, \ldots, \tilde{\eta}_m$ 是线性空间 U 的两组基，且

$$(\tilde{\eta}_1, \tilde{\eta}_2, \ldots, \tilde{\eta}_m) = (\eta_1, \eta_2, \ldots, \eta_m)Q.$$

设 $\varphi \in \mathcal{L}(V, U)$，

$$\begin{aligned} \varphi(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) &= (\eta_1, \eta_2, \ldots, \eta_m)A_{m\times n}, \\ \varphi(\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \ldots, \tilde{\varepsilon}_n) &= (\tilde{\eta}_1, \tilde{\eta}_2, \ldots, \tilde{\eta}_m)B_{m\times n}. \end{aligned}$$

则 $B = Q^{-1}AP$，即 A、B 相抵。
反之，若 A、B 相抵，则 A、B 是同一个线性映射在不同基下的矩阵。

内积空间上的线性映射与矩阵酉相抵

定理

设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 和 $\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \ldots, \tilde{\varepsilon}_n$ 是内积空间 V 的两组标准正交基，且

$$(\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \ldots, \tilde{\varepsilon}_n) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)P.$$

又设 $\eta_1, \eta_2, \ldots, \eta_m$ 和 $\tilde{\eta}_1, \tilde{\eta}_2, \ldots, \tilde{\eta}_m$ 是内积空间 U 的两组标准正交基，且

$$(\tilde{\eta}_1, \tilde{\eta}_2, \ldots, \tilde{\eta}_m) = (\eta_1, \eta_2, \ldots, \eta_m)Q.$$

则 P, Q 是酉矩阵。设 $\varphi \in \mathcal{L}(V, U)$，

$$\begin{aligned} \varphi(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) &= (\eta_1, \eta_2, \ldots, \eta_m)A_{m\times n}, \\ \varphi(\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \ldots, \tilde{\varepsilon}_n) &= (\tilde{\eta}_1, \tilde{\eta}_2, \ldots, \tilde{\eta}_m)B_{m\times n}. \end{aligned}$$

则 $B = Q^{-1}AP = Q^HAP$。

酉相抵

定义

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

$$B = UAV,$$

则称 A 与 B酉相抵。若进一步要求上述矩阵都是实矩阵，则称 A 与 B正交相抵。

- 酉相抵关系是一个等价关系。
- 若 A 与 B 酉相抵，则 $A^H A$ 与 $B^H B$ 酉相似。

奇异值

定理

设 $A \in \mathbb{C}^{m\times n}$，$r(A)=r$。则 $\overline{A}^T A$ 恰有 r 个正特征值，且其余特征值都为 0。

定义

设 $A \in \mathbb{C}^{m\times n}$，$r(A)=r$。记 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ 为 $\overline{A}^T A$ 的全部非零特征值的平方根。称 $\sigma_1, \ldots, \sigma_r$ 为矩阵 A 的奇异值。

奇异值分解

- 设 $A \in \mathbb{C}^{m\times n}$，则

$$\lambda^{m-n} \det(\lambda E_n - \overline{A}^T A) = \det(\lambda E_m - A\overline{A}^T),$$

即 $\overline{A}^T A$ 与 $A\overline{A}^T$ 有相同的非零特征值，所以 A 与 $\overline{A}^T$ 有相同的奇异值。

定理

设 $A \in \mathbb{C}^{m\times n}$，$\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$，其中 $\sigma_1, \ldots, \sigma_r$ 是 A 的全部奇异值，则存在酉矩阵U, V，使得

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \overline{V}^T.$$

实数域下的奇异值分解

推论

设 $A \in \mathbb{R}^{m\times n}$，$\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$，其中 $\sigma_1, \ldots, \sigma_r$ 是 A 的全部奇异值，则存在正交矩阵 P 和 Q，使得

$$A = P \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} Q^T.$$

例子

例

设 $A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$，求 A 的一个奇异值分解。

奇异值分解的几何意义

例

设 $A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$，求线性映射 $x \mapsto Ax$ 将单位球

$$\{x \in \mathbb{R}^3 \mid \|x\| = 1\}$$

映成的像集合。

- 像集合是椭圆，u_1 是长轴所在方向，u_2 是短轴所在方向；σ_1 是长轴长，σ_2 是短轴长。

奇异值分解的外积形式

定理

设 $A \in \mathbb{C}^{m\times n}$ 是一个 $m\times n$ 阶复方阵，$r(A)=r$，$\sigma_1 \geq \cdots \geq \sigma_r > 0$ 是 A 的所有非零奇异值，$A = USV^H$ 是 A 的奇异值分解，且 $s_{jj}=\sigma_j$，$j=1,\ldots,r$。记 U 矩阵的第 j 列为 u_j，V 矩阵的第 j 列为 v_j。则

$$A = \sigma_1 u_1 v_1^H + \cdots + \sigma_r u_r v_r^H.$$

奇异值与特征值

- 当 A 酉相似于对角矩阵时，A 的奇异值是 A 的非零特征值的模。
- 当 A不酉相似于对角矩阵时，特征值和奇异值没有特别明显的关系。
- 例子：$\begin{pmatrix} 0 & n_1 & 0 \\ 0 & 0 & n_2 \\ 0 & 0 & 0 \end{pmatrix}$
- 奇异值的“数值稳定性”。

第一大奇异值 σ_1

- $\sigma_1 = \max_{X \neq 0} \frac{\|AX\|}{\|X\|}$;
- $\sigma_1 \geq |\lambda_i|, \ \forall i = 1, \ldots, n$;
- $\sigma_1(A)$ 也称为 A 的 算子范数。