



矩阵论及其应用

主讲老师：徐跃东

2022-2023学年第二学期

矩阵分解

□ 矩阵分解：

➤ 常见矩阵标准型及分解

三角分解

满秩分解

谱分解

➤ Schur分解与正规矩阵

➤ 奇异值分解

➤ 应用范例

Schur分解与正规矩阵

□ 概念回顾

➤ 矩阵对角化

可对角化的一般矩阵

$$A = P\Lambda P^{-1}$$

实对称矩阵

$$A = P\Lambda P^{-1} = P\Lambda P^T$$

什么样的矩阵可以具有和实对称矩阵类似的分解？

Schur分解

□ 定理: $A \in \mathbb{C}^{n \times n}$ 为可逆矩阵, 则存在酉矩阵 U 和主对角线上元素皆正的上三角矩阵 R , 使得 $A=UR$ 。 (称 $A=UR$ 为矩阵 A 的酉分解)

Schur分解

□证明:

把矩阵 A 按照列分块 $A=(\alpha_1, \alpha_2, \cdots, \alpha_n)$, 则向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关, 由Schmidt正交化方法可得到

$$\beta_1 = \alpha_1,$$

$$\beta_2 = \alpha_2 - k_{21}\beta_1$$

.....

$$\beta_n = \alpha_n - k_{n, n-1}\beta_{n-1} - \cdots - k_{n1}\beta_1.$$

其中 $k_{ij} = (\alpha_i, \beta_j) / (\beta_j, \beta_j)$, $(j < i)$, 即有

$$\alpha_1 = \beta_1,$$

$$\alpha_2 = k_{21}\beta_1 + \beta_2$$

.....

$$\alpha_n = k_{n1}\beta_1 + k_{n2}\beta_2 + \cdots + k_{n, n-1}\beta_{n-1} + \beta_n$$

Schur分解

□证明:

$$(\alpha_1, \alpha_2, \dots, \alpha_n) = (\beta_1, \beta_2, \dots, \beta_n) \begin{pmatrix} 1 & k_{21} & \cdots & k_{n1} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & k_{nn-1} \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

再将 $\beta_1, \beta_2, \dots, \beta_n$ 单位化得 $\gamma_i = \frac{1}{\|\beta_i\|} (i = 1, 2, \dots, n)$

$$A = (\gamma_1, \gamma_2, \dots, \gamma_n) \begin{pmatrix} \|\beta_1\| & & & \\ & \|\beta_2\| & & \\ & & \ddots & \\ & & & \|\beta_n\| \end{pmatrix} \begin{pmatrix} 1 & k_{21} & \cdots & k_{n1} \\ & 1 & \ddots & \vdots \\ & & \ddots & k_{n,n-1} \\ & & & 1 \end{pmatrix}$$

Schur分解

□证明:

令 $Q = (\gamma_1, \gamma_2, \dots, \gamma_n)$, 则 Q 是酉 (正交) 矩阵, 而令

$$R = \begin{pmatrix} \|\beta_1\| & & & \\ & \|\beta_2\| & & \\ & & \ddots & \\ & & & \|\beta_n\| \end{pmatrix} \begin{pmatrix} 1 & k_{21} & \dots & k_{n1} \\ & 1 & \ddots & \vdots \\ & & \ddots & k_{n,n-1} \\ & & & 1 \end{pmatrix} = \begin{pmatrix} \|\beta_1\| & \frac{(\alpha_2, \beta_1)}{\|\beta_1\|} & \dots & \frac{(\alpha_n, \beta_1)}{\|\beta_1\|} \\ & \|\beta_2\| & \ddots & \vdots \\ & & \ddots & \frac{(\alpha_n, \beta_{n-1})}{\|\beta_{n-1}\|} \\ & & & \|\beta_n\| \end{pmatrix}$$

因为 $\|\beta_i\| > 0 (i=1, 2, \dots, n)$ 是正实数, 所以 R 是正线上三角矩阵,

因此 A 有 QR 分解。

Schur分解

□例：矩阵A的UR分解 $A = \begin{pmatrix} 3 & 14 & 9 \\ 6 & 43 & 3 \\ 6 & 22 & 15 \end{pmatrix}$

解 A的列向量

$$\alpha_1 = \begin{pmatrix} 3 \\ 6 \\ 6 \end{pmatrix}, \alpha_2 = \begin{pmatrix} 14 \\ 43 \\ 22 \end{pmatrix}, \alpha_3 = \begin{pmatrix} 9 \\ 3 \\ 15 \end{pmatrix}$$

将其正交化得

$$\beta_1 = \begin{pmatrix} 3 \\ 6 \\ 6 \end{pmatrix} \quad \beta_2 = \alpha_2 - \frac{16}{3}\beta_1 = \begin{pmatrix} -2 \\ 11 \\ 10 \end{pmatrix}$$

$$\beta_3 = \alpha_3 + \frac{3}{5}\beta_2 - \frac{5}{3}\beta_1 = \frac{1}{5} \begin{pmatrix} 14 \\ -2 \\ -5 \end{pmatrix}$$

Schur分解

再单位化得

$$\gamma_1 = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \gamma_2 = \frac{1}{15} \begin{pmatrix} -2 \\ 11 \\ -10 \end{pmatrix}, \gamma_3 = \begin{pmatrix} 14 \\ -2 \\ -5 \end{pmatrix}$$

故得正交矩阵 $Q=(\gamma_1, \gamma_2, \gamma_3)$ $Q = \frac{1}{15} \begin{pmatrix} 5 & -2 & 14 \\ 10 & 11 & -2 \\ 10 & -10 & -5 \end{pmatrix}$

以及上三角矩阵

$$R = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 16/3 & 5/3 \\ 0 & 1 & -3/5 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 9 & 48 & 15 \\ 0 & 15 & -9 \\ 0 & 0 & 3 \end{pmatrix}$$

所以

$$A = \frac{1}{15} \begin{pmatrix} 5 & -2 & 14 \\ 10 & 11 & -2 \\ 10 & -10 & -5 \end{pmatrix} \begin{pmatrix} 9 & 48 & 15 \\ 0 & 15 & -9 \\ 0 & 0 & 3 \end{pmatrix}$$

Schur分解

□ 定理: 设矩阵 $A \in \mathbb{C}^{m \times n}$ 是列满秩的矩阵, 则矩阵 A 可以分解为 $A=QR$, 其中 $Q \in \mathbb{C}^{m \times n}$ 的列向量是标准正交的向量组, $R \in \mathbb{C}^{n \times n}$ 是主对角线上元素为正数的上三角形矩阵。

Schur分解

证明 当 A 为列满秩时,有 $m \geq k$.将 A 的列向量扩充为空间 $\mathbb{C}^m$ 的基,得到可逆矩阵 $(A \mid A_1) \in \mathbb{C}^{m \times m}$,由定理3.7,有 UR 分解

$$(A \mid A_1) = UR.$$

将酉矩阵 $U \in \mathbb{C}^{m \times m}$ 分块为 $U = (Q \mid Q_1)$, $Q \in \mathbb{C}^{m \times k}$.

将上三角矩阵 R 分块为 $R = \begin{pmatrix} R_1 & R_2 \\ 0 & R_3 \end{pmatrix}$,

其中 $R_1 \in \mathbb{C}^{k \times k}$ 是上三角矩阵,由于

$$(A \mid A_1) = (Q \mid Q_1) \begin{pmatrix} R_1 & R_2 \\ 0 & R_3 \end{pmatrix} = (QR_1 \mid QR_2 + Q_1R_3),$$
$$A = QR_1.$$

有

$A = QR_1$ 即为满足题意的 QR 分解.

Schur分解

□ **定理:** 对矩阵 $A \in \mathbb{C}^{n \times n}$, 存在酉矩阵 U 和上三角矩阵 T , 使得

$$U^H A U = T = \begin{bmatrix} \lambda_1 & & * & \\ & \lambda_2 & * & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \quad \text{上三角}$$

其中 λ 为矩阵 A 的特征值

$$\begin{aligned} P^{-1} A P &= J & A &= P J P^{-1} \\ & & &= U \underbrace{(R J R^{-1})}_T U \end{aligned}$$

□ 证明

证明 因为 $A \in \mathbb{C}^{n \times n}$, 故 A 可相似于 Jordan 标准形,

$$A = PJP^{-1}.$$

又 $P \in \mathbb{C}^{n \times n}$ 为可逆矩阵, 则由定理 3.7, P 有 UR 分解 $P = UR$, 则

$$A = PJP^{-1} = URJR^{-1}U^H.$$

令 $T = RJR^{-1}$, 则 T 是一个上三角形矩阵, 即有

$$U^H AU = T.$$

由相似矩阵特征值相等, 上三角形矩阵 T 的主对角线元素是 A 的全部特征值: $\lambda_1, \lambda_2, \dots, \lambda_n$.

正规矩阵

□ 定义 n 阶复矩阵 A 如果满足等式 $AA^H = A^H A$, 则称 A 是正规矩阵或规范矩阵.

例：下列哪些矩阵是正规矩阵？

- 实对称矩阵 $A^T = A$;
- 反实对称矩阵 $A^T = -A$;
- 正交矩阵 $A^T = A^{-1}$;
- 酉矩阵 $A^H = A^{-1}$;
- Hermite 矩阵 $A^H = A$;
- 反Hermite矩阵 $A^H = -A$;
- 对角矩阵

正规矩阵

□ 定义 n 阶复矩阵 A 如果满足等式 $AA^H = A^HA$, 则称 A 是正规矩阵或规范矩阵.

□ 证明: 设 A 为正规矩阵, B 酉相似于 A , 求证 B 也是正规矩阵.

证明 设 U 为酉矩阵, 使 $B = U^H A U$, 则

$$B^H B = U^H A^H U U^H A U = U^H A^H A U,$$

又

$$B B^H = U^H A U U^H A^H U = U^H A A^H U,$$

由 A 为正规矩阵, $A^H A = A A^H$, 得

$$B^H B = B B^H.$$

由定义, B 为正规矩阵.

正规矩阵

□ 定理: $A \in \mathbb{C}^{n \times n}$ 正规 $\Leftrightarrow A$ 酉相似于对角形

证明 必要性: 若 A 满足 $A^H A = A A^H$, 则有 Schur 分解, 存在酉矩阵 U , 使 $A = U T U^H$, 从例 9 推出上三角形矩阵 T 满足: $T^H T = T T^H$, 由

$$T = \begin{pmatrix} \lambda_1 & t_{12} & \cdots & t_{1n} \\ & \lambda_2 & \cdots & t_{2n} \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix},$$

比较 $T T^H$ 与 $T^H T$ 的第 i 行, 第 i 列元素:

$$\begin{cases} (T T^H)_{ii} = \sum_{j=1}^n |t_{ij}|^2 + |\lambda_i|^2, \\ (T^H T)_{ii} = \sum_{j=1}^{i-1} |t_{ji}|^2 + |\lambda_i|^2, \end{cases} \quad i = 1, 2, \dots, n.$$

由 $(T^H T)_{ii} = (T T^H)_{ii}$ 得出: $t_{ij} = 0 (i \neq j)$, 因此 T 为对角矩阵:

正规矩阵

□ 定理: $A \in \mathbb{C}^{n \times n}$ 正规 $\Leftrightarrow A$ 酉相似于对角形

$$T = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

充分性: 因为对角矩阵是正规矩阵, 又 A 酉相似于正规矩阵, 由例 9, A 是正规矩阵.

□

□ 推论: $A \in \mathbb{C}^{n \times n}$ 是正规矩阵的充分必要条件是 A 有 n 个线性无关的特征向量构成空间 $\mathbb{C}^n$ 的标准正交基.

$$A = U \Lambda U^H \quad Au = u \lambda \\ \Rightarrow Au_i = \lambda_i u_i$$

正规矩阵

$$A^H = A$$

□ **定理：** Hermite 矩阵的特征值是实数，而且属于不同特征值的特征向量是正交的。

证明 设 A 为 Hermite 矩阵, 则由于 A 是正规矩阵, 所以存在酉矩阵 U , 使

$$U^H A U = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix},$$

作为 Hermite 矩阵, $A^H = A$, 即有

$$\begin{bmatrix} \bar{\lambda}_1 & & \\ & \bar{\lambda}_2 & \\ & & \ddots & \\ & & & \bar{\lambda}_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

即 $\bar{\lambda}_i = \lambda_i, i = 1, 2, \dots, n$, 因此 λ_i 为实数.

又设 A 的谱为 $\{\lambda_1, \lambda_2, \dots, \lambda_s\}$, 由 A 正规,

$$C^n = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots \oplus V_{\lambda_s},$$

设 C^n 的标准正交基 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是 A 的 n 个特征向量, 则 $V_{\lambda_1} = L\{\alpha_1, \alpha_2, \dots, \alpha_{r_1}\}$, $V_{\lambda_2} = L\{\alpha_{r_1+1}, \dots, \alpha_{r_1+r_2}\}, \dots, V_{\lambda_s} = L\{\alpha_{r_1+r_2+\dots+r_{s-1}+1}, \dots, \alpha_n\}$, 因此当 $\lambda \neq \lambda_i$ 时, V_{λ_i} 与 V_{λ_j} 为彼此正交的子空间, 这说明 A 关于不同特征值的特征向量是正交的.

正规矩阵

□ **定理：**酉矩阵的特征值的模长为 1，即分布在复平面的单位圆上.

证明 设 A 为酉矩阵, 则 A 正规, 即存在酉矩阵 U , 使

$$A = U \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} U^H,$$

则

$$A^H A = I_n \Leftrightarrow U \begin{bmatrix} |\lambda_1|^2 & & & \\ & |\lambda_2|^2 & & \\ & & \ddots & \\ & & & |\lambda_n|^2 \end{bmatrix} U^H = I_n,$$

即

$$|\lambda_i|^2 = 1, i = 1, 2, \dots, n.$$

下面我们讨论正规矩阵的谱分解, 给出一个矩阵是正规矩阵的另一个充分必要条件.

正规矩阵

□ 定理：（正规矩阵谱分解） 设 $A \in \mathbb{C}^{n \times n}$ ， A 的谱为 $\{\lambda_1, \lambda_2, \dots, \lambda_s\}$, ($s \leq n$)，则 A 是正规矩阵的充分必要条件是 A 有如下的谱分解

$$A = \sum_{i=1}^s \lambda_i P_i \quad ;$$

其中

$$(1) \quad P_i P_j = \begin{cases} P_i, & i = j \\ 0, & i \neq j \end{cases}$$

$$(2) \quad \sum_{i=1}^s P_i = I$$

$$(3) \quad P_i^H = P_i$$

正规矩阵

□证明:

充分性: 由 P_i 满足的性质

$$AA^H = \left(\sum_{i=1}^s \lambda_i P_i \right) \left(\sum_{j=1}^s \bar{\lambda}_j P_j^H \right) = \sum_{i=1}^s |\lambda_i|^2 P_i = A^H A,$$

所以 A 为正规矩阵.

正规矩阵

□ 必要性: 设 λ_i 为 A 的 r_i 重根, $\sum_{i=1}^s r_i = n$. 由 A 为正规矩阵, 存在酉矩阵 U , 使得

$$A = U \begin{bmatrix} \underbrace{\left. \begin{matrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_1 \end{matrix} \right\} r_1}_{\text{①}} & & \\ & \underbrace{\left. \begin{matrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_2 \end{matrix} \right\} r_2}_{\text{②}} & & \\ & & \underbrace{\left. \begin{matrix} \lambda_s & & \\ & \ddots & \\ & & \lambda_s \end{matrix} \right\} r_s}_{\text{③}} & & \\ & & & & \end{bmatrix} U^H$$

$$= U \begin{bmatrix} \lambda_1 I_{r_1} & & \\ & \lambda_2 I_{r_2} & \\ & & \ddots \\ & & & \lambda_s I_{r_s} \end{bmatrix} U^H = U \underbrace{\left(\sum_{i=1}^s \lambda_i \right)}_{\text{④}} \begin{bmatrix} 0 & & \\ & \ddots & \\ & & I_{r_i} & \\ & & & 0 & \ddots & \\ & & & & & 0 \end{bmatrix} U^H$$

正规矩阵

$$= \sum_{i=1}^s \lambda_i U \begin{bmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & I_{r_i} & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} U^H, \quad P_i = U \begin{bmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & I_{r_i} & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} U^H,$$

则 P_i 满足 (3.16) 式 ~ (3.18) 式, 而且

$$A = \sum_{i=1}^s \lambda_i P_i.$$

在正规矩阵的谱分解中, 矩阵 P_i 不仅是幂等矩阵, 而且是 Hermite 矩阵, 在后面章节, 我们将会看到 P_i 不仅是投影, 而且是正交投影矩阵.

正规矩阵

□例：求正规矩阵 $A = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}$ 的谱分解

由矩阵A的特征多项式 $|\lambda E - A| = (\lambda - 1)^3(\lambda + 3)$

得A的特征值 $\lambda_{1,2,3} = 1, \lambda_4 = -3$

对于 $\lambda_1 = 1$, 相应的线性无关的特征向量为

$$x_1 = (1, 1, 0, 0)^T, x_2 = (1, 0, 1, 0)^T, x_3 = (-1, 0, 0, 1)^T$$

对于 $\lambda_2 = -3$, 相应的特征向量为 $x_4 = (-1, -1, -1, 1)^T$

将 x_1, x_2, x_3 标准正交化得：

$$\alpha_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right)^T, \alpha_2 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0 \right)^T, \alpha_3 = -\left(\frac{1}{\sqrt{12}}, \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{12}}, \frac{3}{\sqrt{12}} \right)^T$$

正规矩阵

□例：求正规矩阵 $A = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}$ 的谱分解

将 x_1, x_2, x_3 标准正交化得：

$$\alpha_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right)^T, \alpha_2 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0 \right)^T, \alpha_3 = -\left(\frac{1}{\sqrt{12}}, \frac{1}{\sqrt{12}}, \frac{1}{\sqrt{12}}, \frac{3}{\sqrt{12}} \right)^T$$

将 x_4 标准化 $\alpha_4 = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right)^T$

记 $U_1 = (\alpha_1, \alpha_2, \alpha_3), U_2 = (\alpha_4)$

$$P_1 = U_1 U_1^H, P_2 = U_2 U_2^H$$

则 $A = \lambda_1 P_1 + \lambda_2 P_2 = P_1 - 3P_2$

正规矩阵

□ 正规矩阵谱分解的几何意义

► 如果2阶实正规矩阵 A 有两个相同的特征值 λ , 则 $A = \lambda I$ 就是它的谱分解。如果 A 有两个不同的特征值 λ_1 与 λ_2 , 则其谱分解为 $A = \lambda_1 P_1 + \lambda_2 P_2$. 因此, 对任意 $\alpha \in R^2$, 有

$$A\alpha = \lambda_1 P_1 \alpha + \lambda_2 P_2 \alpha.$$

计算内积可得 $(P_1 \alpha, P_2 \alpha) = (P_1 \alpha)^T P_2 \alpha = \alpha^T P_1^T P_2 \alpha = 0$, 所以 $\lambda_1 P_1 \alpha$ 与 $\lambda_2 P_2 \alpha$ 是正交的向量. 所以上述公式将 A 分解成了两个正交向量的和。因此, 二维正规矩阵的谱分解实际上是平面的正交投影变换的推广.

矩阵分解

□ 矩阵分解：

➤ 常见矩阵标准型及分解

三角分解

满秩分解

谱分解

➤ Schur分解与正规矩阵

➤ 奇异值分解

➤ 应用范例

奇异值分解

□ 概念回顾

➤ 矩阵简化

对角化

Jordan阵化

LU (LDU) 分解

满秩分解

谱分解

✓

Schur分解

✓

奇异值分解

任意非正规矩阵怎么简化为对角矩阵?

奇异值分解

定理： 设 $A \in C_r^{m \times n} (r > 0)$, 则存在 $S \in C_m^{m \times n}, T \in C_n^{n \times n}$, 使得

$$SAT = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \quad \text{右式称为矩阵A的等价标准型}$$

酉等价： 设 $A, B \in C^{m \times n}$, 若存在 m 阶酉矩阵 U 和 n 阶酉矩阵 V , 使得 $U^H A V = B$, 则称 A 与 B 酉等价。

矩阵的奇异值分解就是矩阵在酉等价下的一种标准型。

奇异值分解

□ 引理1 设 $A \in C^{m \times n}$, $A^H A$ 与 AA^H 的特征值均为非负实数。

证明 设 λ 是 $A^H A$ 的特征值, x 是相应的特征向量, 则 $A^H Ax = \lambda x$

由于 $A^H A$ 为 Hermite 矩阵, 故 λ 是实数。

$$0 \leq (Ax, Ax) = (Ax)^H (Ax) = \lambda x^H x$$

$$\because x^H x > 0, \therefore \lambda \geq 0$$

同理可证 AA^H 的特征值也是非负实数。

奇异值分解

□ 引理2 设 $A \in C_r^{m \times n}$, 则 $\text{rank}(A^H A) = \text{rank}(A A^H) = \text{rank}(A)$

证明 设 x 是方程组 $A^H A x = 0$ 的非0解, $Ax \in C^m$

则由 $(Ax, Ax) = x^H (A^H Ax) = 0$ 得 $Ax = 0$;

反之, $Ax = 0$ 的解也是 $A^H Ax = 0$ 的解;

因此, 线性方程组 $Ax = 0$ 与 $A^H Ax = 0$ 同解。

$$\text{rank}(A) = \text{rank}(A^H A)$$

用 A^H 替换 A , 得 $\text{rank}(A) = \text{rank}(A^H A) = \text{rank}(A A^H)$

奇异值分解

□性质

设 $A \in \mathbb{C}_r^{m \times n}$,

对于Hermite 矩阵 $A^H A$, AA^H , 设 $A^H A$, AA^H 有 r 个非0特征值, 分别记为

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

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_r > \mu_{r+1} = \cdots = \mu_n = 0$$

$$\text{则 } \lambda_i = \mu_i, i = 1, 2, \cdots, r$$

即: $A^H A$ 与 AA^H 非0特征值相同, 并且非零特征值的个数为 $\text{rank}(A)$

奇异值分解

□ 定义: $A \in \mathbb{C}^{m \times n}$, 秩 $(A) = r$, 设 $A^H A$ 的特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$, $\lambda_{r+1} = \lambda_{r+2} = \dots = \lambda_n = 0$, 则矩阵的奇异值

$$\sigma_i = \sqrt{\lambda_i}, \quad i = 1, 2, \dots, r.$$

注: A 的正奇异值个数恰等于 $\text{rank}(A)$, 并且 A 与 A^H 有相同的奇异值。

奇异值分解

□ 定理 酉等价的矩阵有相同的奇异值

证明 设 $A, B \in C_r^{m \times n}$, A 与 B 酉等价, 则

存在酉矩阵 $U \in C^{m \times m}$, $V \in C^{n \times n}$, 使

$$A = UB$$

由 $A^H A = (UB)^H (UB) = B^H B$,

所以 $A^H A$ 与 $B^H B$ 是酉相似的, 有相同的特征值,
故 A 与 B 有相同的奇异值。

奇异值分解

□ **定理** 任何矩阵 $A \in \mathbb{C}^{m \times n}$, 秩 $(A) = r$, 则存在酉矩阵 $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$, 使得

$$A = U \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & 0 \end{bmatrix} V^H \quad \Delta = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ & & & & 0 \end{bmatrix}$$

注: $\begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix}$ 称为矩阵 A 的酉等价标准形.

奇异值分解

□证明：设矩阵 $A^H A$ 的特征值为

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

则存在 n 阶酉矩阵 V ，使得

$$V^H (A^H A) V = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \begin{bmatrix} \Sigma^2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \quad (2)$$

将 V 分块为

$$V = (V_1 \quad V_2)$$

其中 V_1, V_2 分别是 V 的前 r 列与后 $n-r$ 列.

奇异值分解

并改写②式为

$$A^H A V = V \begin{bmatrix} \Sigma^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

则有

$$A^H A V_1 = V_1 \Sigma^2, \quad A^T A V_2 = \mathbf{0} \quad \text{③}$$

由③的第一式可得

$$V_1^H A^H A V_1 = \Sigma^2, \quad \text{或者} (A V_1 \Sigma^{-1})^H (A V_1 \Sigma^{-1}) = I_r$$

由③的第二式可得

$$(A V_2)^H (A V_2) = \mathbf{0} \quad \text{或者} \quad A V_2 = \mathbf{0}$$

令 $U_1 = A V_1 \Sigma^{-1}$, 则 $U_1^H U_1 = I_r$, 即 U_1 的 r 个列是两两正交的单位向量. 记

奇异值分解

$$U_1 = (\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_r)$$

因此可将 $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_r$ 扩充成标准正交基, 记增添的向量为 $\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m$, 并构造矩阵

$$U_2 = (\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m)$$

则 $U = (U_1, U_2) = (\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_r, \mathbf{u}_{r+1}, \cdots, \mathbf{u}_m)$

是 m 阶正交矩阵, 且有

于是可得 $U_1^H U_1 = I_r$, $U_2^H U_1 = \mathbf{O}$

$$U^H A V = U^H (A V_1, A V_2) = \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} (U_1 \Sigma, \mathbf{O}) = \begin{bmatrix} \Sigma & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$$

矩阵U, V的空间性质

□ $V = [v_1, v_2, \dots, v_r, \dots, v_n] = [V_1 \ V_2] \in \mathbb{C}^{n \times n}$ 的列向量是空间 $\mathbb{C}^n$ 的标准正交基。

➤ V_2 的列向量是空间 $N(A)$ 的标准正交基。

➤ V_1 的列向量是空间 $N^\perp(A)$ 的标准正交基。

□ $U = [u_1, u_2, \dots, u_r, \dots, u_m] = [U_1 \ U_2] \in \mathbb{C}^{m \times m}$ 的列向量是空间 $\mathbb{C}^m$ 的标准正交基。

➤ U_1 的列向量是 $R(A)$ 的标准正交基。

➤ U_2 的列向量是 $R^\perp(A)$ 的标准正交基。

矩阵U, V的空间性质

□ 推论：在矩阵A的奇异值分解 $A=UDV^H$ 中，U的列向量为 AA^H 的特征向量，V的列向量为 A^HA 的特征向量。

$$\begin{aligned}\because AA^H &= (UDV^H)(UDV^H)^H \\ &= UDV^HVDU^H = UD^2U^H\end{aligned}$$

$$\therefore (AA^H)U = UD^2 = U\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_r, 0, \dots, 0)$$

$$\text{记 } U = (u_1, u_2, \dots, u_n)$$

$$\text{则 } (AA^H)u_i = \lambda_i u_i, \quad i = 1, 2, \dots, n$$

说明：此定理仅是奇异值分解的必要条件，但不是充分条件。

奇异值分解方法

□ 利用矩阵 $A^H A$ 求解

➤ 求矩阵 AA^H 的正相似对角矩阵及酉相似矩阵 U ;

$$U^H (AA^H) U = \begin{pmatrix} \Delta^2 & 0 \\ 0 & 0 \end{pmatrix}$$

➤ 拆分矩阵: $U = (U_1, U_2), U_1 \in C^{m \times r}, U_2 \in C^{m \times (m-r)},$
 $V_1 = A^H U_1 \Delta^{-1} \in C^{n \times r}$

➤ 扩充 V_1 为酉矩阵 $V = (V_1, V_2)$

➤ 构造奇异值分解 $A = U \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} V^H$

奇异值分解方法

□例 求矩阵A的奇异值分解

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

利用矩阵 AA^H 求解

$$AA^H = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \sigma_1 = \sqrt{5}, \Delta = (\sqrt{5})$$

AA^H 的特征值 $\lambda_1 = 5, \lambda_2 = \lambda_3 = 0$

奇异值分解方法

□例 求矩阵A的奇异值分解

对应的特征向量分别为

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix};$$

取 $\mathbf{U} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, $\mathbf{U}_1 = \mathbf{x}_1$, $\mathbf{U}_2 = (\mathbf{x}_2, \mathbf{x}_3)$

$$\text{令 } \mathbf{V}_1 = \mathbf{A}^H \mathbf{U}_1 \Delta^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \frac{1}{\sqrt{5}} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix},$$

奇异值分解方法

□例 求矩阵A的奇异值分解

$$\text{取 } V_2 = \begin{pmatrix} \frac{-2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}, \text{ 则 } V = (V_1, V_2) = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}$$

$$\text{因此 } A = U \Sigma V^H$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{5} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}$$



Thanks