第一章　矩阵的分解

矩阵分解是将一个矩阵分解为比较简单的或者具有某种特性的若干矩阵的和或者乘积，往往分解出的矩阵我们可以方便的研究其矩阵的秩、特征值、奇异值等信息，这为对于原始矩阵的研究或者处理带来极大的便利性。

本节着重会介绍下面的几种矩阵分解方法:

- 矩阵的三角分解
- 矩阵的谱分解
- 矩阵的满秩分解（最大秩分解）
- 矩阵的奇异值分解

注意：本节内容与上一章知识完全没有关系，但需要了解线性代数中对于矩阵的特征值、特征向量、实对称矩阵的性质等内容，如果遗忘了，请翻阅第零章的相关知识。

1.1　矩阵的三角分解

1.1.1　常见的三角矩阵及其性质

在开始本节内容之前，先来认识一下常见的三角矩阵

- 正线上三角阵

$$\boldsymbol{R} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

- 单位上三角阵

$$\boldsymbol{R}=\begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & 1 & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

- 正线下三角阵

$$\boldsymbol{R}=\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

- 单位下三角阵

$$\boldsymbol{R}=\begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & a1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix}$$

常见的三角矩阵具有下面的性质:

定理 1.1.1: 三角阵的性质

1. 上 (下) 三角矩阵 $\boldsymbol{R}$ 的逆也为上 (下) 三角矩阵，对角元是原来元素的倒数。
2. 两个上 (下) 三角矩阵 $\boldsymbol{R_1}, \boldsymbol{R_2}$ 的乘积 $\boldsymbol{R_1}\boldsymbol{R_2}$ 也是上 (下) 三角矩阵。
3. 酉矩阵 $\boldsymbol{U}$ 的逆 $\boldsymbol{U}^{-1}$ 也是酉矩阵。
4. 两个酉矩阵之积 $\boldsymbol{U}_1\boldsymbol{U}_2$ 也是酉矩阵。

牢记上面的性质，这些性质会在后面反复用到。

1.1.2 n 阶矩阵的三角分解

定理 1.1.2

设 $\boldsymbol{A} \in \mathbb{C}_n^{n\times n}$，则 $\boldsymbol{A}$ 可以唯一的分解为

$$\boldsymbol{A} = \boldsymbol{U}_1\boldsymbol{R}$$

其中 $\boldsymbol{U}_1$ 是酉矩阵，$\boldsymbol{R}$ 是正线上三角矩阵。

或 $\boldsymbol{A}$ 可以唯一地分解为

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}_2$$

其中 $\boldsymbol{L}$ 是正线下三角矩阵，$\boldsymbol{U}_2$ 是酉矩阵

这就是 n 阶矩阵的三角分解。证明过程//TODO，(后面如果没时间了就看看课本吧，这个证明可以看看)

除此之外，还有下面的推论：

推论 1.1.3

设 $\boldsymbol{A} \in \mathbb{R}_n^{n\times n}$，则 $\boldsymbol{A}$ 可以唯一地分解为

$$\boldsymbol{A} = \boldsymbol{Q}_1\boldsymbol{R}$$

其中，$\boldsymbol{Q}_1$ 是正交矩阵，$\boldsymbol{R}$ 是正线上三角矩阵。

或，$\boldsymbol{A}$ 可以唯一分解为

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{Q}_2$$

其中，$\boldsymbol{L}$ 是正线下三角矩阵，$\boldsymbol{Q}_2$ 是正交矩阵.

推论 1.1.4

设 $\boldsymbol{A}$ 是实对称正定矩阵，则存在唯一的正线上三角矩阵，使得

$$\boldsymbol{A} = \boldsymbol{R}^T\boldsymbol{R}$$

推论 1.1.5

设 $\boldsymbol{A}$ 是正定 Hermite 矩阵，则存在唯一的正线上三角矩阵，使得

$$\boldsymbol{A} = \boldsymbol{R}^H\boldsymbol{R}$$

证明省略，后面的不考，本小节结束。

1.1.3 任意矩阵的三角分解

定理 1.1.6

设 $\boldsymbol{A}$ 为行满秩矩阵或者列满秩矩阵，则

- 设 $\boldsymbol{A} \in \mathbb{C}_n^{m\times n}$，则存在 m 阶酉矩阵 $\boldsymbol{U}$ 即 n 阶正线上三角矩阵 $\boldsymbol{R}$，使得

$$\boldsymbol{A} = \boldsymbol{U}\begin{bmatrix}\boldsymbol{R}\\ 0\end{bmatrix}$$

- 设 $\boldsymbol{A} \in \mathbb{C}_m^{m\times n}$，则存在 n 阶酉矩阵 $\boldsymbol{U}$ 即 m 阶正线下三角矩阵 $\boldsymbol{L}$，使得

$$\boldsymbol{A} = (\boldsymbol{L} \quad 0)\boldsymbol{U}$$

最后还有一个较为重要的在后面奇异值分解中会用到的定理：

定理 1.1.7

设 $\boldsymbol{A} \in \mathbb{C}_r^{m\times n}$，则存在酉矩阵 $\boldsymbol{U} \in \mathbb{U}m \times m$ 和 $\boldsymbol{V} \in \mathbb{U}^{n\times n}$ 及 r 阶正线下三角矩阵 $\boldsymbol{L}$，使得

$$\boldsymbol{A} = \boldsymbol{U}\begin{bmatrix}\boldsymbol{L} & 0\\ 0 & 0\end{bmatrix}\boldsymbol{V}$$

1.2 矩阵的谱分解

接下来了解一下矩阵的谱分解，矩阵的谱分解与矩阵的特征值与特征向量关系密切。

1.2.1 单纯矩阵的谱分解

在介绍单纯矩阵的谱分解前，需要明白何为单纯矩阵，而想要明白何为单纯矩阵，又需要了解下面的两个概念：代数重复度和几何重复度。

定义 1.2.1: 代数重复度

设 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 是 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 的相异特征值，其重数分别为 $r_1, r_2, \cdots, r_k$，则称 r_i 为矩阵 $\boldsymbol{A}$ 的特征值 λ_i 的**代数重复度**。

代数重复度与第一章提到的代数重数概念十分相似，其实他们就是一回事，如果能明白这个关系，下面的几何重复度理解起来也就没什么难度了。

定义 1.2.2: 几何重复度

齐次方程组 $\boldsymbol{Ax} = \lambda_i \boldsymbol{x}$ $(i = 1, 2, \cdots, k)$ 的解空间 $\boldsymbol{V}_{\lambda_i}$ 称为 $\boldsymbol{A}$ 的对应于特征值 λ_i 的特征空间，则 $\boldsymbol{V}_{\lambda_i}$ 的维数称为 $\boldsymbol{A}$ 的特征值 λ_i 的**几何重复度**。

了解了这两个概念，接下来就要介绍何为单纯矩阵了：

定义 1.2.3

若矩阵 $\boldsymbol{A}$ 的每个特征值的代数重复度与几何重复度相等，则称矩阵 $\boldsymbol{A}$ 为**单纯矩阵**。

在了解了何为单纯矩阵之后，接下来介绍何为单纯矩阵的谱分解：

定理 1.2.1: 单纯矩阵谱分解

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 是单纯矩阵，则 $\boldsymbol{A}$ 可分解为一系列幂等矩阵 $\boldsymbol{A}_i$ $(i = 1, 2, \cdots, n)$ 的加权和，即

$$\boldsymbol{A} = \sum_{i=1}^{n} \lambda_i \boldsymbol{A}_i$$

其中，λ_i $(i = 1, 2, \cdots, n)$ 是 $\boldsymbol{A}$ 的特征值，并且 $\boldsymbol{A}_i$ 还有如下的性质：

1. 幂等性：$\boldsymbol{A}_i^2 = \boldsymbol{A}_i$
2. 分离性：$\boldsymbol{A}_i \boldsymbol{A}_j = 0$ $(i \neq j)$
3. 可加性：$\sum\limits_{i=1}^{n} \boldsymbol{A}_i = \boldsymbol{E}_n$

这些性质的证明就略过了，感兴趣可以翻翻书。

上面介绍的是如何把单纯矩阵进行谱分解，下面的定理说的就是谱分解中的矩阵 $\boldsymbol{A}_i$ 与单纯矩阵之间的关系：

定理 1.2.2

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，它有 k 个相异特征值 $\lambda_i\ (i=1,2,\cdots,k)$，则 $\boldsymbol{A}$ 是单纯矩阵的充要条件是存在 k 个矩阵 $\boldsymbol{A}_i\ (i=1,2,\cdots,k)$ 满足

1. $\boldsymbol{A}_i\boldsymbol{A}_j = \begin{cases} \boldsymbol{A}_i, & i=j \\ 0, & i\neq j \end{cases}$
2. $\sum_{i=1}^k \boldsymbol{A}_i = \boldsymbol{E}_n$
3. $\boldsymbol{A} = \sum_{i=1}^k \lambda_i \boldsymbol{A}_i$

1.2.2 正规矩阵及其分解

了解了单纯矩阵的分解之后，下面是正规矩阵的分解

定义 1.2.4: 正规矩阵定义

若 n 阶复矩阵 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}\boldsymbol{A}^H = \boldsymbol{A}^H\boldsymbol{A}$$

则称 $\boldsymbol{A}$ 为**正规矩阵**。

正规矩阵有下面的性质：

1. 若 $\boldsymbol{A}$ 为正规矩阵且 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 酉相似，则 $\boldsymbol{B}$ 也为正规矩阵
2. (Schur 定理) 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，则存在酉矩阵 $\boldsymbol{U}$，使得

   $$\boldsymbol{A} = \boldsymbol{U}\mathbb{R}\boldsymbol{U}^H$$

   其中，$\mathbb{R}$ 是一个上三角矩阵且主对角线上的元素为 $\boldsymbol{A}$ 的特征值。
3. 设 $\boldsymbol{A}$ 是正规矩阵且是三角矩阵，则 $\boldsymbol{A}$ 是对角矩阵。

什么样的矩阵能够成为正规矩阵呢？

定理 1.2.3

n 阶复矩阵 $\boldsymbol{A}$ 是正规矩阵的充要条件是 $\boldsymbol{A}$ 与对角阵酉相似，即存在 n 阶酉矩阵 $\boldsymbol{U}$，使得

$$\boldsymbol{A} = \boldsymbol{U} diag(\lambda_1, \lambda_2, \cdots, \lambda_n) \boldsymbol{U}^H$$

其中，$\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值。

或者可以用下面这个定理来判断：

定理 1.2.4

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，它有 k 个相异特征值 λ_i $(i = 1, 2, \cdots, k)$，则 $\boldsymbol{A}$ 是单纯矩阵的充要条件是存在 k 个矩阵 $\boldsymbol{A}_i$ $(i = 1, 2, \cdots, k)$ 满足

1. $\boldsymbol{A}_i\boldsymbol{A}_j = \begin{cases} \boldsymbol{A}_i, & i = j \\ 0, & i \neq j \end{cases}$
2. $\sum_{i=1}^{k} \boldsymbol{A}_i = \boldsymbol{E}_n$
3. $\boldsymbol{A} = \sum_{i=1}^{k} \lambda_i \boldsymbol{A}_i$
4. $\boldsymbol{A}_i^H = \boldsymbol{A}_i$ $(i = 1, 2, \cdots, k)$

1.3 矩阵的满秩分解

1.3.1 满秩分解的定义

在了解了矩阵的谱分解之后，再来看一种较为简单的分解——矩阵的满秩分解，也叫最大秩分解。

定理 1.3.1

设 $\boldsymbol{A} \in \mathbb{C}_r^{m\times n}$，则存在矩阵 $\boldsymbol{B} \in \mathbb{C}_r^{m\times r}, \boldsymbol{D} \in \mathbb{C}_r^{r\times n}$，使得

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{D}$$

满秩分解很好理解，即，把一个矩阵分解成一个行满秩矩阵和一个列满秩矩阵相乘的形式即可。

1.3.2 满秩分解的步骤

对一个矩阵进行满秩分解的步骤如下:

1. 进行初等行变换，化为**行最简阶梯形矩阵** $\widetilde{\boldsymbol{A}}$
2. 找到 $\boldsymbol{A}$ 的一个极大线性无关组，这就是列满秩矩阵 $\boldsymbol{B}$
3. 在 $\widetilde{\boldsymbol{A}}$ 中的所有非零行构成矩阵 $\boldsymbol{D}$

1.4 矩阵的奇异值分解

在本节的最后了解一下矩阵的奇异值分解，矩阵的奇异值分解常用来进行图像处理，对图像进行压缩，因此具有较为重要的应用场景。

1.4.1 为什么需要进行奇异值分解?

前面对于矩阵的分解，大多数都有要求，要求矩阵需要是方阵，这就又对矩阵的种类进行了限制，如果想要对于任意的矩阵进行分解，由第一节的最后一个定理可知，存在酉矩阵 $\boldsymbol{U},\boldsymbol{V}$ 和 r 阶上三角矩阵或者下三角矩阵 $\boldsymbol{R}(L)$，使得

$$\boldsymbol{A}=\boldsymbol{U}\begin{bmatrix}\boldsymbol{R} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}\boldsymbol{V}$$

但这样子依旧要求是方阵，奇异值分解就是在这个基础上而来。

如何把一个矩阵变成方阵呢? 这里采用的方法是研究 $\boldsymbol{A}^H\boldsymbol{A}$，这的确变成了一个方阵，可为什么要这样处理呢? 这就要给出下面的定理

定理 1.4.1

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

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

上面的定理是为下面矩阵的奇异值做铺垫，上面的这些定理仿佛在说:“我们可以用 $\boldsymbol{A}^H\boldsymbol{A}$ 或者 $\boldsymbol{A}\boldsymbol{A}^H$ 来研究 $\boldsymbol{A}$”。

1.4.2 奇异值

定义 1.4.1

设 $\boldsymbol{A} \in \mathbb{C}_r^{m\times n}, \boldsymbol{A}^H\boldsymbol{A}$ 的特征值为

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

则称 $\sigma_i = \sqrt{\lambda_i}\ (i = 1, 2, \cdots, r)$ 为 $\boldsymbol{A}$ 的**正奇异值**。

有了奇异值的概念之后，就有了下面奇异值分解的定理：

定理 1.4.2: 奇异值分解

设 $\boldsymbol{A} \in \mathbb{C}_r^{m\times n}, \sigma_1, \sigma_2, \cdots, \sigma_r$ 是 $\boldsymbol{A}$ 的 r 个正奇异值，则存在酉矩阵 $\boldsymbol{U} \in \mathbb{C}^{m\times m}$ 和 $\boldsymbol{V} \in \mathbb{C}^{n\times n}$，使得

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{D} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix} \boldsymbol{V}$$

其中,$\boldsymbol{D} = diag(\sigma_1, \sigma_2, \cdots, \sigma_r), |\sigma_i| = \sigma_i$

这个定理与第一节最后的定理很类似，不同之处就是将左上角的元素换成了对角矩阵。最后简单介绍一下酉等价的概念，以及两个酉等价矩阵的奇异值关系。

定义 1.4.2

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

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{B}\boldsymbol{V}$$

则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 酉等价，若两个矩阵酉等价，则他们有相同的正奇异值。

1.4.3 求矩阵奇异值分解的方法

求矩阵 $\boldsymbol{A}$ 的奇异值分解的方法如下：

1. 求 $\boldsymbol{A}^H\boldsymbol{A}$ 的特征值以及特征向量
2. 构造酉矩阵 $\boldsymbol{V}$
3. 构造酉矩阵 $\boldsymbol{U}$

4. 给出矩阵 $\boldsymbol{A}$ 的奇异值分解