矩阵论复习

Haijun Xiong

2022 年 12 月 3 日

1 线性空间与线性变换

1.1 线性空间

定义 1.1 设 V 是一个以 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\cdots$ 为元素的非空集合，F 是一个数域。在其中定义两种运算，一种叫加法：$\forall\boldsymbol{\alpha},\boldsymbol{\beta}\in V$，$\boldsymbol{\alpha}+\boldsymbol{\beta}\in V$；另一种叫数量乘法：$\forall k\in F$，$\boldsymbol{\alpha}\in V$，$k\boldsymbol{\alpha}\in V$，并且满足八条运算法则：

(1) 加法交换律：$\boldsymbol{\alpha}+\boldsymbol{\beta}=\boldsymbol{\beta}+\boldsymbol{\alpha}$；

(2) 加法结合律：$(\boldsymbol{\alpha}+\boldsymbol{\beta})+\boldsymbol{\gamma}=\boldsymbol{\alpha}+(\boldsymbol{\beta}+\boldsymbol{\gamma})$；

(3) V 中存在零元素：$\exists\boldsymbol{\alpha}_0\in V$，$\forall\boldsymbol{\alpha}\in V$, 使得 $\boldsymbol{\alpha}+\boldsymbol{\alpha}_0=\boldsymbol{\alpha}$，记 $\boldsymbol{\alpha}_0=\mathbf{0}$；

(4) V 中存在负元素：$\forall\boldsymbol{\alpha}\in V$，$\exists\boldsymbol{\beta}\in V$, 使得 $\boldsymbol{\alpha}+\boldsymbol{\beta}=\mathbf{0}$，记 $\boldsymbol{\beta}=-\boldsymbol{\alpha}$；

(5) 数乘结合律：$(kl)\,\boldsymbol{\alpha}=k\,(l\boldsymbol{\alpha})$；

(6) 存在 $1\in F$：$1\cdot\boldsymbol{\alpha}=\boldsymbol{\alpha}$；

(7) 分配律：$(k+l)\,\boldsymbol{\alpha}=k\boldsymbol{\alpha}+l\boldsymbol{\alpha}$；

(8) 分配律：$k\,(\boldsymbol{\alpha}+\boldsymbol{\beta})=k\boldsymbol{\alpha}+k\boldsymbol{\beta}$。

则称 V 为数域 F 上的**线性空间**，V 中元素称为**向量**。F 为**实**（**复**）**数域**时，称 V 是**实**（**复**）**线性空间**。

定理 1.1 线性空间 V 有以下性质：

(1) V 中的零元素惟一；

(2) V 中任一元素的负元素惟一；

(3) 设 0 为数零，$\mathbf{0}$ 为 V 中零向量，则

(i) $0\cdot\boldsymbol{\alpha}=\mathbf{0}$;

(ii) $k\cdot\mathbf{0}=\mathbf{0}$，$k\in F$；

(iii) 若 $k\cdot\boldsymbol{\alpha}=\mathbf{0}$，则一定有 $k=0$ 或者 $\boldsymbol{\alpha}=\mathbf{0}$；

(iv) $(-1)\,\boldsymbol{\alpha}=-\boldsymbol{\alpha}$。

定义 1.2 设 V 是线性空间，若存在一组线性无关的向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 使空间中任一向量可由它们线性表示，则称向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 为 V 的一组**基**。基所含向量个数为 V 的**维数**，记为 $\dim V=n$，$n<+\infty$ 或者 $n=+\infty$。

定理 1.2 n 维线性空间中任意 n 个线性无关的向量构成的向量组都是空间的基。

定义 1.3 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是线性空间 $V_n(F)$ 的一组基，$\forall\boldsymbol{\beta}\in V$

$$\boldsymbol{\beta}=\sum_{i=1}^{n}x_i\boldsymbol{\alpha}_i=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} \tag{1.1}$$

则称数 $x_1,x_2,\cdots,x_n$ 是 $\boldsymbol{\beta}$ 在基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 下的坐标，(1.1) 式中向量 $\begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}^{\mathrm{T}}$ 为 $\boldsymbol{\beta}$ 的坐标向量，也简称为坐标。

定理 1.3 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 n 维线性空间 $V_n(F)$ 的一组基，$V_n(F)$ 中向量 $\boldsymbol{\beta}_i$ 在该基下坐标为 X_i，$i=1,2,3,\cdots,m$，则 $V_n(F)$ 中向量组 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m\}$ 线性相关的充要条件是其坐标向量组 $\{X_1,X_2,\cdots,X_m\}$ 是 F^n 中的线性相关组。

定义 1.4 设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$，$\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$ 是 n 维线性空间 $V_n(F)$ 的两组基，若有矩阵 $\boldsymbol{C}\in F^{n\times n}$，使

$$\begin{bmatrix}\boldsymbol{\beta}_1 & \boldsymbol{\beta}_2 & \cdots & \boldsymbol{\beta}_n\end{bmatrix}=\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\boldsymbol{C} \tag{1.2}$$

则称矩阵 $\boldsymbol{C}$ 是从基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 到基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$ 的**过渡矩阵**（**基变换矩阵**）。

定理 1.4 设线性空间 $V_n(F)$ 的一组基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 到另一组基 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$ 的过渡矩阵为 $\boldsymbol{C}$，$V_n(F)$ 中向量 $\boldsymbol{\alpha}$ 在两组基下坐标分别为 $\boldsymbol{X}$，$\boldsymbol{Y}$，则有

$$\boldsymbol{X}=\boldsymbol{C}\boldsymbol{Y} \tag{1.3}$$

定义 1.5 设 $V_n(F)$ 为线性空间，W 是 V 的非空子集合。若 W 的元素关于 V 中加法与数乘向量法运算也构成线性空间，则称 W 是 V 的一个子空间。

定理 1.5 设 W 是线性空间 $V_n(F)$ 的非空子集合，则 W 是 $V_n(F)$ 的子空间的充分必要条件是

(1) 若 $\boldsymbol{\alpha}$，$\boldsymbol{\beta}\in W$，则 $\boldsymbol{\alpha}+\boldsymbol{\beta}\in W$；

(2) 若 $\boldsymbol{\alpha}\in W$，$k\in F$，则 $k\boldsymbol{\alpha}\in W$。

定理 1.6 设 W_1，W_2 是线性空间 V 的子空间，则有

(1) W_1 与 W_2 的交集 $W_1\cap W_2=\{\boldsymbol{\alpha}|\boldsymbol{\alpha}\in W_1\text{且}\boldsymbol{\alpha}\in W_2\}$ 是 V 的子空间，称为 W_1 与 W_2 的**交空间**；

(2) W_1 与 W_2 的和 $W_1+W_2=\{\boldsymbol{\alpha}|\boldsymbol{\alpha}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2，\boldsymbol{\alpha}_1\in W_1，\boldsymbol{\alpha}_2\in W_2\}$ 是 V 的子空间，称为 W_1 与 W_2 的**和空间**。

定理 1.7 设 W_1 和 W_2 是线性空间 V 的子空间，则有如下维数公式：

$$\dim(W_1) + \dim(W_2) = \dim(W_1 + W_2) + \dim(W_1 \cap W_2) \tag{1.4}$$

定义 1.6 设 W_1，W_2 是线性空间 V 的子空间，$W = W_1 + W_2$，如果 $W_1 \cap W_2 = \{\mathbf{0}\}$，则称 W 是 W_1 与 W_2 的直和子空间。记为 $W = W_1 \oplus W_2$

定理 1.8 设 W_1 和 W_2 是线性空间 V 的子空间，$W = W_1 + W_2$，则成立以下等价条件：

(1) $W = W_1 \oplus W_2$；

(2) $\forall \boldsymbol{X} \in W$，$\boldsymbol{X}$ 有惟一的表示式：$\boldsymbol{X} = \boldsymbol{X}_1 + \boldsymbol{X}_2$，其中 $\boldsymbol{X}_1 \in W_1$，$\boldsymbol{X}_2 \in W_2$；

(3) W 中零向量的表达式惟一，即只要 $\mathbf{0} = \boldsymbol{X}_1 + \boldsymbol{X}_2$，$\boldsymbol{X}_1 \in W_1$，$\boldsymbol{X}_2 \in W_2$，就有 $\boldsymbol{X}_1 = \mathbf{0}$，$\boldsymbol{X}_2 = \mathbf{0}$；

(4) 维数公式 $\dim W = \dim W_1 + \dim W_2$。

1.2 内积空间

定义 1.7 对数域 F 上的 n 维线性空间 $V_n(F)$，定义的一个从 $V_n(F)$ 中向量到数域 F 的二元运算，记为 $(\boldsymbol{\alpha}, \boldsymbol{\beta})$，即 $(\boldsymbol{\alpha}, \boldsymbol{\beta})$：$V_n(F) \to F$，如果满足

(1) **对称性**：$(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \overline{(\boldsymbol{\beta}, \boldsymbol{\alpha})}$，其中 $\overline{(\boldsymbol{\beta}, \boldsymbol{\alpha})}$ 表示复数 $(\boldsymbol{\beta}, \boldsymbol{\alpha})$ 的共轭；

(2) **线性性**：$(k_1\boldsymbol{\alpha}, k_2\boldsymbol{\beta}) = k_1\overline{k_2}(\boldsymbol{\alpha}, \boldsymbol{\beta})$，$(\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2, \boldsymbol{\beta}) = (\boldsymbol{\alpha}_1, \boldsymbol{\beta}) + (\boldsymbol{\alpha}_2, \boldsymbol{\beta})$；

(3) **正定性**：$(\boldsymbol{\alpha}, \boldsymbol{\alpha}) \geq 0$，$(\boldsymbol{\alpha}, \boldsymbol{\alpha}) = 0$ 的充分必要条件是 $\boldsymbol{\alpha} = \mathbf{0}$。

则称 $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ 是 $V_n(F)$ 的一个**内积**，并称其中定义了内积的线性空间 $[V_n(F); (\boldsymbol{\alpha}, \boldsymbol{\beta})]$ 为**内积空间**。

如果 $V_n(F)$ 是**实数域**$\mathbb{R}$ 上的线性空间 $V_n(\mathbb{R})$，则 $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}$ 为实内积，$[V_n(\mathbb{R}); (\boldsymbol{\alpha}, \boldsymbol{\beta})]$ 为**欧氏空间**；同理 $V_n(F)$ 是**复数域**$\mathbb{C}$ 上的线性空间 $V_n(\mathbb{C})$ 时，$(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{C}$ 为复内积，称 $[V_n(\mathbb{C}); (\boldsymbol{\alpha}, \boldsymbol{\beta})]$ 为**酉空间**。

定义 1.8 设 $[V_n(F); (\boldsymbol{\alpha}, \boldsymbol{\beta})]$ 为内积空间，称

$$\|\boldsymbol{\alpha}\| = \sqrt{(\boldsymbol{\alpha}, \boldsymbol{\alpha})} \tag{1.5}$$

为向量 $\boldsymbol{\alpha}$ 的长度，若 $\|\boldsymbol{\alpha}\| = 1$，则称 $\boldsymbol{\alpha}$ 为**单位向量**。

定理 1.9 (Cauchy 不等式) 设 $[V_n(F); (\boldsymbol{\alpha}, \boldsymbol{\beta})]$ 为内积空间，则对空间中任意向量 $\boldsymbol{\alpha}$，$\boldsymbol{\beta} \in V_n(F)$，都有

$$|(\boldsymbol{\alpha}, \boldsymbol{\beta})|^2 \leq (\boldsymbol{\alpha}, \boldsymbol{\alpha})(\boldsymbol{\beta}, \boldsymbol{\beta})$$

其中等式成立的充分必要条件是 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 线性相关。*Cauchy-Schwarz* 不等式可以写为：

$$|(\boldsymbol{\alpha}, \boldsymbol{\beta})| \leq \|\boldsymbol{\alpha}\| \cdot \|\boldsymbol{\beta}\|$$

由此可以得出三角不等式:

$$\|\boldsymbol{\alpha}+\boldsymbol{\beta}\| \leq \|\boldsymbol{\alpha}\|+\|\boldsymbol{\beta}\|$$

$$\|\boldsymbol{\alpha}-\boldsymbol{\beta}\| \leq \|\boldsymbol{\alpha}\|+\|\boldsymbol{\beta}\|$$

定义 1.9 在内积空间中，若向量 $\boldsymbol{\alpha}$，$\boldsymbol{\beta}$ 满足

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=0$$

则称向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 是**正交**的。

定理 1.10 不含零向量的正交向量组是线性无关的。

定义 1.10 在内积空间 $[V_n(F);(\boldsymbol{\alpha},\boldsymbol{\beta})]$ 中，若一组基 $\{\epsilon_1,\varepsilon_2,\cdots,\varepsilon_n\}$ 满足条件

$$\left(\varepsilon_i,\varepsilon_j\right)=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}$$

则称向量 $\{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\}$ 为 $[V_n(F);(\boldsymbol{\alpha},\boldsymbol{\beta})]$ 的**标准正交基**。

定理 1.11 (Gram-Schmidt 标准正交化方法) 设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 是内积空间 $[V_n(F);(\boldsymbol{\alpha},\boldsymbol{\beta})]$ 中线性无关的向量组，则由如下方法:

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1$$

$$\boldsymbol{\beta}_k=\boldsymbol{\alpha}_k-\sum_{i=1}^{k-1}\frac{(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_i)}{(\boldsymbol{\beta}_i,\boldsymbol{\beta}_i)}\boldsymbol{\beta}_i,\ k=2,3,\cdots,n \tag{1.6}$$

所得向量组 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n\}$ 是正交向量组。之后再标准化为:

$$\varepsilon_i=\frac{\boldsymbol{\beta}_i}{\|\boldsymbol{\beta}_i\|},\ i=1,3,\cdots,n$$

把正交化方法和标准化方法结合在一起，可得从一组基 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$ 得到标准正交基的方法

$$\begin{cases}\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1,\ \varepsilon_1=\dfrac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|}\\ \boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-(\boldsymbol{\alpha}_2,\varepsilon_1)\,\varepsilon_1,\ \varepsilon_2=\dfrac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|}\\ \cdots\\ \boldsymbol{\beta}_n=\boldsymbol{\alpha}_n-\displaystyle\sum_{i=1}^{n-1}(\boldsymbol{\alpha}_n,\varepsilon_i)\,\varepsilon_i,\ \varepsilon_n=\dfrac{\boldsymbol{\beta}_n}{\|\boldsymbol{\beta}_n\|}\end{cases}$$

用矩阵运算可表示为

$$\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}=\begin{bmatrix}\varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n\end{bmatrix}\begin{bmatrix}\|\boldsymbol{\beta}_1\| & (\boldsymbol{\alpha}_2,\varepsilon_1) & \cdots & (\boldsymbol{\alpha}_n,\varepsilon_1)\\ & \|\boldsymbol{\beta}_2\| & \cdots & (\boldsymbol{\alpha}_n,\varepsilon_2)\\ & & \ddots & \vdots\\ & & & \|\boldsymbol{\beta}_n\|\end{bmatrix} \tag{1.7}$$

标准正交基是内积空间中十分方便的基。$\forall\boldsymbol{\alpha}$，$\boldsymbol{\beta}\in V_n(F)$，$\boldsymbol{\alpha}=\begin{bmatrix}\varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n\end{bmatrix}\boldsymbol{X}$，$\boldsymbol{\beta}=\begin{bmatrix}\varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n\end{bmatrix}\boldsymbol{Y}$，则由前面讨论可知：

(1) 当 $V_n(F)$ 为欧氏空间 $V_n(\mathbb{R})$ 时，$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}=\boldsymbol{Y}^{\mathrm{T}}\boldsymbol{X}$；

(2) 当 $V_n(F)$ 为酉空间 $V_n(\mathbb{C})$ 时，$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{Y}^{\mathcal{H}}\boldsymbol{X}$。

1.3 线性变换

定义 1.11 设 $V_n(F)$ 是一个线性空间，若有 $V_n(F)$ 上的对应关系 T，使 $\forall\boldsymbol{\alpha}\in V_n(F)$，都有确定的向量 $\boldsymbol{\alpha}'=T(\boldsymbol{\alpha})\in V_n(F)$ 与之对应，则称 T 为 $V_n(F)$ 上一个变换. 又若 T 对线性空间中的线性运算，对于 $\forall k_1$，$k_2\in F$，$\forall\boldsymbol{\alpha}_1$，$\boldsymbol{\alpha}_2\in V_n(F)$ 满足

$$T(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2)=k_1T(\boldsymbol{\alpha}_1)+k_2T(\boldsymbol{\alpha}_2) \tag{1.8}$$

则称 T 是线性空间 $V_n(F)$ 上的一个**线性变换**。

线性空间 $V_n(F)$ 上的相似变换 T：$\forall\boldsymbol{\alpha}\in V_n(F)$，$T(\boldsymbol{\alpha})=\lambda\boldsymbol{\alpha}$ 是线性变换，其中 $\lambda\in F$。

当 $\lambda=0$ 时，$\forall\boldsymbol{\alpha}\in V_n(F)$，$T(\boldsymbol{\alpha})=\mathbf{0}$，称 T 为**零变换**；

当 $\lambda=1$ 时，$\forall\boldsymbol{\alpha}\in V_n(F)$，$T(\boldsymbol{\alpha})=\boldsymbol{\alpha}$，称 T 为**恒等变换**。

定理 1.12 设 T 是线性空间 $V_n(F)$ 上的线性变换，则有：

(1) $R(T)=\{\boldsymbol{\beta}|\forall\boldsymbol{\alpha}\in V_n(F)$，使 $\boldsymbol{\beta}=T(\boldsymbol{\alpha})\}$ 是 $V_n(F)$ 的子空间，称为 T 的**像空间**；

(2) $N(T)=\{\boldsymbol{\alpha}|T(\boldsymbol{\alpha})=0\}$ 是 $V_n(F)$ 的子空间，称为 T 的**零空间**。

子空间 $R(T)$ 的维数 $\dim R(T)$ 为线性变换 T 的**秩**，$\dim N(T)$ 为 T 的**零度**。

定义 1.12 设 T_1 与 T_2 都是线性空间 $V_n(F)$ 上的线性变换，定义如下运算：

(1) 变换的乘积 T_1T_2：

$$\forall\boldsymbol{\alpha}\in V_n(F)\,,\quad (T_1T_2)(\boldsymbol{\alpha})=T_1(T_2(\boldsymbol{\alpha}))$$

(2) 变换的加法 T_1+T_2：

$$\forall\boldsymbol{\alpha}\in V_n(F)\,,\quad (T_1+T_2)(\boldsymbol{\alpha})=T_1(\boldsymbol{\alpha})+T_2(\boldsymbol{\alpha})$$

(3) 数乘变换 kT：

$$\forall\boldsymbol{\alpha}\in V_n(F)\,,\ k\in F,\quad (kT)(\boldsymbol{\alpha})=kT(\boldsymbol{\alpha})$$

(4) 可逆变换：对变换 T_1，如果存在变换 T_2，使：

$$T_1\cdot T_2=T_2\cdot T_1=I(\text{恒等变换})$$

则称 T_1 为可逆变换，T_2 是 T_1 的逆变换，记为 $T_2=T_1^{-1}$。

定义 1.13 设 T 是线性空间 $V_n(F)$ 上的线性变换，$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 为 $V_n(F)$ 的基，若存在 n 阶方阵 $\boldsymbol{A} \in F^{n\times n}$，使

$$T\begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix} = \begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\boldsymbol{A} \tag{1.9}$$

则称 $\boldsymbol{A}$ 为 T 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下的**矩阵**。

又 $\forall\boldsymbol{\alpha} \in V_n(F)$，设 $\boldsymbol{\alpha}$ 与 $T(\boldsymbol{\alpha})$ 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下坐标分别是 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$，即

$$\boldsymbol{\alpha} = \begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\boldsymbol{X},\ T(\boldsymbol{\alpha}) = \begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\boldsymbol{Y}$$

可以得到

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} \tag{1.10}$$

定理 1.13 设 T_1 与 T_2 是线性空间 $V_n(F)$ 上的两个线性变换，对基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$，它们的矩阵分别是 $\boldsymbol{A}_1$ 与 $\boldsymbol{A}_2$，则在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下，

(1) $T_1 + T_2$ 的矩阵为 $(\boldsymbol{A}_1 + \boldsymbol{A}_2)$；

(2) T_1T_2 的矩阵为 $\boldsymbol{A}_1\boldsymbol{A}_2$；

(3) kT_1 的矩阵为 $k\boldsymbol{A}_1$；

(4) T_1 为可逆变换的充分必要条件是 $\boldsymbol{A}_1$ 为可逆矩阵，且 T_1^{-1} 的矩阵为 $\boldsymbol{A}_1^{-1}$。

定理 1.14 设线性空间 $V_n(F)$ 的基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 到基 $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$ 的过渡矩阵为 $\boldsymbol{C}$。又 $V_n(F)$ 上线性变换 T 在上述两组基下矩阵分别为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$，则有

$$\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} \tag{1.11}$$

定义 1.14 设 T 是线性空间 $V_n(F)$ 上的线性变换，W 是 $V_n(F)$ 的子空间，如果 $\forall\boldsymbol{\alpha} \in W$，有 $T(\boldsymbol{\alpha}) \in W$，即值域 $T(W) \in W$，则称 W 是 T 的不变子空间。

定义 1.15 设 T 是内积空间 $[V_n(F) : (\boldsymbol{\alpha}, \boldsymbol{\beta})]$ 上的线性变换，如果 T 不改变向量的内积，即 $\forall\boldsymbol{\alpha}$，$\boldsymbol{\beta} \in [V_n(F) : (\boldsymbol{\alpha}, \boldsymbol{\beta})]$ 都有

$$(T(\boldsymbol{\alpha}), T(\boldsymbol{\beta})) = (\boldsymbol{\alpha}, \boldsymbol{\beta})$$

则称 T 为**内积空间上的正交变换**，当空间为欧式空间时称 T 为**正交变换**；若空间为酉空间，则称 T 为**酉变换**。正交（酉）变换在标准正交基下的矩阵称为正交（酉）矩阵。

定理 1.15 设 T 是内积空间上的线性变换，则下列命题等价：

(1) T 是正交（酉）变换；

(2) T 保持向量长度不变；

(3) T 把空间 $V_n(F)$ 的标准正交基变换为标准正交基；

(4) 正交变换关于任一标准正交基的矩阵 $\boldsymbol{C}$ 满足 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C} = \boldsymbol{C}\boldsymbol{C}^{\mathrm{T}} = \boldsymbol{I}$；酉变换关于任一标准正交基的矩阵 $\boldsymbol{U}$ 满足 $\boldsymbol{U}^{\mathcal{H}}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^{\mathcal{H}} = \boldsymbol{I}$。

定理 1.16 正交矩阵 $\boldsymbol{C}$ 和酉矩阵 $\boldsymbol{U}$ 有如下性质：

(1) 正交矩阵的行列式为 ± 1；酉矩阵的行列式的模长为 1；

(2) $\boldsymbol{C}^{-1} = \boldsymbol{C}^{\mathrm{T}}$；$\boldsymbol{U}^{-1} = \boldsymbol{U}^{\mathcal{H}}$；

(3) 正交（酉）矩阵的逆矩阵与乘积仍然是正交（酉）矩阵；

(4) n 阶正交（酉）矩阵的列和行向量组是欧氏（酉）空间 $\mathbb{R}^n$ ($\mathbb{C}^n$) 中的标准正交基。

2 Jordan 标准形介绍

2.1 线性变换的对角矩阵表示

定义 2.1 设 T 为线性空间 V_n 上的线性变换，如果存在 $\boldsymbol{\xi} \in V_n(F)$ 和数 $\lambda \in F$，$\boldsymbol{\xi} \neq 0$，使得 $T(\boldsymbol{\xi}) = \lambda\boldsymbol{\xi}$，则称数 λ 为 T 的**特征值**，向量 $\boldsymbol{\xi}$ 为线性变换 T 的对应于特征值 λ 的**特征向量**。

定理 2.1 设 $V_n(F)$ 上线性变换 T 在基 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 下矩阵为 $\boldsymbol{A}$，则 $\boldsymbol{A}$ 的特征值 λ 就是变换 T 的特征值；若 $\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的特征向量，则 $\boldsymbol{\xi} = \begin{bmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\boldsymbol{X}$ 就是 T 的特征向量。

定义 2.2 设 λ 为线性变换 T 的特征值，$\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_t$ 是 T 对应于 λ 的特征向量的极大线性无关组，则称子空间 $V_\lambda = L\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_t\}$ 为 T 关于 λ 的**特征子空间**。

定理 2.2 设 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 $V_n(F)$ 上线性变换 T 的 s 个互异的特征值。V_{λ_i} 是 λ_i 的特征子空间，$i = 1, 2, \cdots, s$，则有

(1) V_{λ_i} 是 T 的不变子空间；

(2) $\lambda_i \neq \lambda_j$ 时，$V_{\lambda_i} \cap V_{\lambda_j} = \{\boldsymbol{0}\}$；

(3) 若 λ_i 是 T 的 k_i（代数重数）重特征值，则 $\dim V_{\lambda_i}$(几何重数）$\leq k_i$，$i = 1, 2, \cdots, s$。

定理 2.3 线性变换 T 有对角矩阵表示的充分必要条件是 T 有 n 个线性无关的特征向量。

定理 2.4 线性变换 T 有对角矩阵表示的充分必要条件是：

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s} = V_n(F)$$

(1) 幂等矩阵：$\boldsymbol{A}^2 = \boldsymbol{A}$，$\boldsymbol{A}$ 相似于对角矩阵 $\begin{bmatrix}\boldsymbol{I}_r & \\ & \boldsymbol{0}\end{bmatrix}$，其中 r 为矩阵 $\boldsymbol{A}$ 的秩；

(1) 乘方矩阵：$\boldsymbol{A}^2 = \boldsymbol{I}$，$\boldsymbol{A}$ 相似于对角矩阵 $\begin{bmatrix}\boldsymbol{I}_s & \\ & -\boldsymbol{I}_t\end{bmatrix}$，其中 $s + t = n$。

2.2 Jordan 矩阵介绍

定义 2.3 形如

$$\boldsymbol{J}(\lambda) = \begin{bmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{bmatrix} \tag{2.1}$$

的 r 阶方阵称为一个 r 阶 ***Jordan* 块**．由若干个 *Jordan* 块 $\boldsymbol{J}_i(\lambda_i)$ 构成的准对角矩阵

$$\boldsymbol{J} = \begin{bmatrix}\boldsymbol{J}_1(\lambda_1) & & & \\ & \boldsymbol{J}_2(\lambda_2) & & \\ & & \ddots & \\ & & & \boldsymbol{J}_m(\lambda_m)\end{bmatrix} \tag{2.2}$$

称为 ***Jordan* 矩阵**。

定理 2.5 在复数域上，每个 n 阶方阵 $\boldsymbol{A}$ 都相似于一个 *Jordan* 矩阵，即存在可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J}_{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{J}_1(\lambda_1) & & & \\ & \boldsymbol{J}_2(\lambda_2) & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s(\lambda_s) \end{bmatrix}$$

其中

$$\boldsymbol{J}_i(\lambda_i) = \begin{bmatrix} \boldsymbol{J}_{i1}(\lambda_i) & & & \\ & \boldsymbol{J}_{i2}(\lambda_i) & & \\ & & \ddots & \\ & & & \boldsymbol{J}_{it_i}(\lambda_i) \end{bmatrix} \in \mathbb{C}^{k_i \times k_i}$$

$\boldsymbol{J}_{ij}(\lambda_i)$ 为 n_i 阶 *Jordan* 块，$\sum\limits_{j=1}^{t_i} n_j = k_i$。$\boldsymbol{J}_i(\lambda_i)$ 是 $k_i \times k_i$ 阶 *Jordan* 矩阵，$\sum\limits_{i=1}^{s} k_i = n$。若不计较 *Jordan* 块的排列次序，则每个方阵的 ***Jordan* 标准形**$\boldsymbol{J}_{\boldsymbol{A}}$ 是惟一的。

定理 2.6 (Jordan 化方法) 计算步骤归纳如下：

(1) 求 $\boldsymbol{A}$ 的特征多项式

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = (\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_s)^{k_s}$$

$\lambda_1, \lambda_2, \cdots, \lambda_s$ 互异，从而 λ_i 为 $\boldsymbol{A}$ 的 k_i 重特征值，其代数重数 k_i 决定 *Jordan* 矩阵 $\boldsymbol{J}_i(\lambda_i)$ 的阶数为 k_i；

(2) 对 λ_i，由 $(\boldsymbol{A} - \lambda_i I)X = \boldsymbol{0}$，求 $\boldsymbol{A}$ 的线性无关的特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{t_i}$。$\lambda_i$ 的几何重数 $\dim V_{\lambda_i} = t_i$ 决定 $\boldsymbol{J}_i(\lambda_i)$ 中有 t_i 个 *Jordan* 块；

(3) 若 λ_i 的代数重数等于几何重数：$k_i \dim V_{\lambda_i}$，λ_i 对应的 *Jordan* 矩阵为 t_i 阶对角矩阵。若 $\dim V_{\lambda_i} = t_i < k_i$，则在 V_{λ_i} 中选择适当特征向量 $\boldsymbol{\alpha}_i$。由

$$\begin{cases} (\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{\alpha}_i = \boldsymbol{0} \\ (\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_i \\ (\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{\beta}_3 = \boldsymbol{\beta}_2 \\ \qquad \vdots \\ (\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{\beta}_{n_j} = \boldsymbol{\beta}_{n_j - 1} \end{cases} \tag{2.3}$$

求 *Jordan* 链 $\boldsymbol{\alpha}_i, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_{n_i}$，确定 $\boldsymbol{J}_i(\lambda_i)$ 中 *Jordan* 块 $\boldsymbol{J}_{ij}(\lambda_i)$ 的阶数 n_i，从而得到了 $\boldsymbol{J}_{\boldsymbol{A}}$ 的结构。

(4) 所有 *Jordan* 链构成矩阵 $\boldsymbol{P}$，必有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J}_{\boldsymbol{A}}$$

2.3 最小多项式

定义 2.4 设 $\boldsymbol{A} \in F^{n\times n}$，$a_i \in F$，$g(\lambda) = a_m\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0$ 是一个多项式，则称矩阵 $g(\boldsymbol{A}) = a_m\boldsymbol{A}^m + a_{m-1}\boldsymbol{A}^{m-1} + \cdots + a_1\boldsymbol{A} + a_0\boldsymbol{I}$ 为 $\boldsymbol{A}$ 的矩阵多项式。

定理 2.7 设 $\boldsymbol{A} \in F^{n\times n}$，$g(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的矩阵多项式，则有如下结果：

(1) 若 λ_0 是 $\boldsymbol{A}$ 的特征值，则 $g(\lambda_0)$ 是 $g(\boldsymbol{A})$ 的特征值；

(2) 如果 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$：$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$，则 $g(\boldsymbol{A})$ 相似于 $g(\boldsymbol{B})$：$\boldsymbol{P}^{-1}g(\boldsymbol{A})\boldsymbol{P} = g(\boldsymbol{B})$；

(3) 如果 $\boldsymbol{A}$ 为准对角矩阵，则 $g(\boldsymbol{A})$ 也是准对角矩阵。而且若

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k \end{bmatrix}, \boldsymbol{A}_i\text{为方子块}$$

则

$$g(\boldsymbol{A}) = \begin{bmatrix} g(\boldsymbol{A}_1) & & & \\ & g(\boldsymbol{A}_2) & & \\ & & \ddots & \\ & & & g(\boldsymbol{A}_k) \end{bmatrix}$$

定理 2.8 ($g(\boldsymbol{A})$ 计算)

$$\boldsymbol{A} = \boldsymbol{P}\begin{bmatrix} \boldsymbol{J}_1(\lambda_1) & & & \\ & \boldsymbol{J}_2(\lambda_2) & & \\ & & \ddots & \\ & & & \boldsymbol{J}_k(\lambda_k) \end{bmatrix}\boldsymbol{P}^{-1} \to g(\boldsymbol{A}) = \boldsymbol{P}\begin{bmatrix} g(\boldsymbol{J}_1) & & & \\ & g(\boldsymbol{J}_2) & & \\ & & \ddots & \\ & & & g(\boldsymbol{J}_k) \end{bmatrix}\boldsymbol{P}^{-1}$$

$$\boldsymbol{J}(\lambda) = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \to g(\boldsymbol{J}) = \begin{bmatrix} g(\lambda) & g'(\lambda) & \dfrac{g''(\lambda)}{2!} & \cdots & \dfrac{g^{(r-1)}(\lambda)}{(r-1)!} \\ & g(\lambda) & g'(\lambda) & \ddots & \vdots \\ & & g(\lambda) & \ddots & \dfrac{g''(\lambda)}{2!} \\ & & & \ddots & g'(\lambda) \\ & & & & g(\lambda) \end{bmatrix}$$

定理 2.9 (Cayley-Hamilton) 对 n 阶方阵 $\boldsymbol{A}$，若存在多项式 $g(\lambda)$，使矩阵 $g(\boldsymbol{A}) = \boldsymbol{0}$，则称 $g(\lambda)$ 为矩阵 $\boldsymbol{A}$ 的化零多项式。设 $\boldsymbol{A} \in F^{n\times n}$，则方阵 $\boldsymbol{A}$ 的特征多项式就是 $\boldsymbol{A}$ 的化零多项式，即若

$$f(\lambda) = |\lambda\boldsymbol{I} - \boldsymbol{A}| = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

则有 $f(\boldsymbol{A}) = \boldsymbol{0}$。

定义 2.5 设 T 是线性空间 $V_n(T)$ 上的线性变换，$m_T(\lambda)$ 是一个关于 λ 的多项式，如果 $m_T(\lambda)$ 是 T 的**最小多项式**，则满足

(1) $m_T(\lambda)$ 最高次项系数为 *1*；

(2) $m_T(\lambda)$ 是 T 的一个化零多项式，即 $m_T(T) = 0$；

(3) $m_T(\lambda)$ 是 T 的化零多项式中次数最低的多项式。

定理 2.10 T 的特征多项式 $f(\lambda)$ 与最小多项式 $m_T(\lambda)$，即若

$$f(\lambda) = |\lambda \boldsymbol{I} - \boldsymbol{A}| = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_s)^{r_s}$$

则有

$$m_T(\lambda) = (\lambda - \lambda_1)^{t_1}(\lambda - \lambda_2)^{t_2} \cdots (\lambda - \lambda_s)^{t_s} \tag{2.4}$$

其中

$$1 \le t_i \le r_i,\ i = 1, 2, \cdots, s$$

定理 2.11 设变换 T 的特征多项式为

$$f(\lambda) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_s)^{r_s}$$

又 T 的 *Jordan* 标准形中关于特征值 λ_i 的 *Jordan* 块的最高阶数为 $\overline{n_i}$，则 T 的最小多项式

$$m_T(\lambda) = (\lambda - \lambda_1)^{\overline{n_1}}(\lambda - \lambda_2)^{\overline{n_2}} \cdots (\lambda - \lambda_s)^{\overline{n_s}}$$

定理 2.12 线性变换 T 可对角化的充分必要条件是 T 的最小多项式 $m_T(\lambda)$ 是一次因子的乘积，即

$$m_T(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_s)$$

3 矩阵的分解

3.1 常见的矩阵标准型与分解

定义 3.1 (矩阵标准型) 常见的几种标准型：

(1) 等价标准型：对于 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，存在可逆矩阵 $\boldsymbol{P} \in \mathbb{C}^{m\times m}$，$\boldsymbol{Q} \in \mathbb{C}^{n\times n}$，使得

$$\boldsymbol{A} = \boldsymbol{P}\begin{bmatrix} \boldsymbol{I}_r & 0 \\ 0 & 0 \end{bmatrix}\boldsymbol{Q}$$

(2) 相似标准型：对于 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，存在可逆矩阵 $\boldsymbol{P} \in \mathbb{C}^{n\times n}$，使得

$$\boldsymbol{A} = \boldsymbol{P}\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}\boldsymbol{P}^{-1} \text{或} \boldsymbol{A} = \boldsymbol{P}\boldsymbol{J}_A\boldsymbol{P}^{-1}$$

(3) 正交相似：一个实对称矩阵 $\boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A})$ 一定可正交相似于对角形：即存在正交矩阵 $\boldsymbol{C}$ 使

$$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

(4) 酉相似：一个 *Hermite* 矩阵 $\boldsymbol{A}(\boldsymbol{A}^{\mathcal{H}} = \boldsymbol{A})$ 一定可酉相似于对角形：即存在酉矩阵 $\boldsymbol{U}$，使

$$\boldsymbol{U}^{\mathcal{H}}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

定义 3.2 (三角分解) 设 $\boldsymbol{A} \in F^{n\times n}$

(1) 若 $\boldsymbol{L}$，$\boldsymbol{U} \in F^{n\times n}$ 分别是下三角矩阵和上三角矩阵，$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$，则称 $\boldsymbol{A}$ 可作 ***LU* 分解**；对于增广矩阵 $\left[\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{I}_n \end{array}\right]$，只用第 i 行乘数 k 加到第 j 行 $(i < j)$ 型初等变换将 $\boldsymbol{A}$ 化为上三角形 $\boldsymbol{U}$，可以得到 $\left[\begin{array}{c|c} \boldsymbol{U} & \boldsymbol{P} \end{array}\right]$，可知 $\boldsymbol{P}\boldsymbol{A} = \boldsymbol{U}$，于是有 $\boldsymbol{L} = \boldsymbol{P}^{-1}$，则 $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$。

(2) 若 $\boldsymbol{L}$，$\boldsymbol{V} \in F^{n\times n}$ 分别是对角元素为 1 的下三角矩阵和上三角矩阵，$\boldsymbol{D}$ 为对角矩阵。$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{V}$，则称 $\boldsymbol{A}$ 可作 ***LDV* 分解**；通过 *LU* 分解得到 $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$，通过每行除以对应的对角线上元素的值，将 $\boldsymbol{U}$ 的对角线元素化为 1，得到 $\boldsymbol{U} = \boldsymbol{D}\boldsymbol{V}$，则 $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{V}$。

定义 3.3 (满秩分解) 设 $\boldsymbol{A} \in F^{m\times n}$，$\operatorname{rank}(\boldsymbol{A}) = r$，若存在秩为 r 的矩阵 $\boldsymbol{B} \in F^{m\times r}$，$\boldsymbol{C} \in F^{r\times n}$，使

$$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C} \tag{3.1}$$

则称 (3.1) 式为矩阵 $\boldsymbol{A}$ 的**满秩分解**，对任何非零矩阵 $\boldsymbol{A} \in F^{m\times n}$，都存在满秩分解。

定理 3.1 (满秩分解求法) 设 $\boldsymbol{A} \in F^{m\times n}$，$\operatorname{rank}(\boldsymbol{A}) = r$，而且 $\boldsymbol{A}$ 的前 r 列线性无关，则它们是 $\boldsymbol{A}$ 的列向量的极大无关组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r\}$，设 $\boldsymbol{A}_1 = \begin{bmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r \end{bmatrix}$，则 $\operatorname{rank}(\boldsymbol{A}_1) = r$，$\boldsymbol{A}_1 \in F^{m\times r}$。又 $\boldsymbol{A}$ 的后 $n-r$ 列 $\{\boldsymbol{\alpha}_{r+1}, \boldsymbol{\alpha}_{r+2}, \cdots, \boldsymbol{\alpha}_n\}$ 可表示为列向量极大无关组的线性组合，设 $\boldsymbol{A}_2 = \begin{bmatrix} \boldsymbol{\alpha}_{r+1} & \boldsymbol{\alpha}_{r+2} & \cdots & \boldsymbol{\alpha}_n \end{bmatrix}$，则 $\boldsymbol{A}_2 = \boldsymbol{A}_1\boldsymbol{S}$，其中 $\boldsymbol{S}_{r\times(n-r)} = \begin{bmatrix} \boldsymbol{X}_{r+1} & \boldsymbol{X}_{r+2} & \cdots & \boldsymbol{X}_n \end{bmatrix}$，$\boldsymbol{X}_j$ 满足 $\boldsymbol{\alpha}_j = \begin{bmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_r \end{bmatrix} \boldsymbol{X}_j$，$j = r+1, r+2, \cdots, n$。因此 $\boldsymbol{A} = \left[\begin{array}{c|c} \boldsymbol{A}_1 & \boldsymbol{A}_2 \end{array}\right] = \left[\begin{array}{c|c} \boldsymbol{A}_1 & \boldsymbol{A}_1\boldsymbol{S} \end{array}\right]$，即 $\boldsymbol{A} = \boldsymbol{A}_1 \left[\begin{array}{c|c} \boldsymbol{I}_r & \boldsymbol{S} \end{array}\right]$，所以 $\boldsymbol{B} = \boldsymbol{A}_1$，$\boldsymbol{C} = \left[\begin{array}{c|c} \boldsymbol{I}_r & \boldsymbol{S} \end{array}\right]$ 即为满秩分解。

(1) 用行初等变换把 $\boldsymbol{A}$ 化为 *Hermite* 标准形；

(2) 依 *Hermite* 标准形中向量 $\boldsymbol{e}_i$ 所在的列的位置第 j_i 列，相应地取出 $\boldsymbol{A}$ 的第 j_i 列 $\boldsymbol{\alpha}_{j_i}$，得到 $\boldsymbol{A}$ 的列向量极大无关组 $\{\boldsymbol{\alpha}_{j_1}, \boldsymbol{\alpha}_{j_2}, \cdots, \boldsymbol{\alpha}_{j_r}\}$，$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{\alpha}_{j_1} & \boldsymbol{\alpha}_{j_2} & \cdots & \boldsymbol{\alpha}_{j_r} \end{bmatrix}$；

(3) $\boldsymbol{A}$ 的 *Hermite* 标准形中非零行构成矩阵 $\boldsymbol{C}$。得到 $\boldsymbol{A}$ 的满秩分解：$\boldsymbol{A} = \boldsymbol{B}\boldsymbol{C}$。

定义 3.4 (可对角化矩阵的谱分解) 对方阵 $\boldsymbol{A} \in F^{n\times n}$，设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为矩阵 $\boldsymbol{A}$ 的 n 个特征值。$\boldsymbol{A}$ 互异的特征值集合 $\{\lambda_1, \lambda_2, \cdots, \lambda_s\}$ 称为矩阵 $\boldsymbol{A}$ 的**谱**，其中 λ_i 为 $\boldsymbol{A}$ 的 r_i 重特征值（$i = 1, 2, \cdots, s$），$\sum\limits_{i=1}^{s} r_i = n$。当 $\boldsymbol{A}$ 可相似于对角形时，则有可逆矩阵 $\boldsymbol{P}$，使

$$\boldsymbol{A} = \boldsymbol{P} \begin{bmatrix} \lambda_1 &&&&&&&& \\ & \ddots &&&&&&& \\ && \lambda_1 &&&&&& \\ &&& \lambda_2 &&&&& \\ &&&& \ddots &&&& \\ &&&&& \lambda_2 &&& \\ &&&&&& \ddots && \\ &&&&&&& \lambda_s & \\ &&&&&&&& \ddots & \\ &&&&&&&&& \lambda_s \end{bmatrix} \boldsymbol{P}^{-1} \tag{3.2}$$

首先分解对角矩阵

$$
\begin{bmatrix}
\lambda_1 &&&&&&&& \\
& \ddots &&&&&&& \\
&& \lambda_1 &&&&&& \\
&&& \lambda_2 &&&&& \\
&&&& \ddots &&&& \\
&&&&& \lambda_2 &&& \\
&&&&&& \ddots && \\
&&&&&&& \lambda_s & \\
&&&&&&&& \ddots \\
&&&&&&&&& \lambda_s
\end{bmatrix}
$$

$$
= \lambda_1 \begin{bmatrix} \boldsymbol{I}_{r_1} &&& \\ & \mathbf{0} && \\ && \ddots & \\ &&& \mathbf{0} \end{bmatrix} + \lambda_2 \begin{bmatrix} \mathbf{0} &&& \\ & \boldsymbol{I}_{r_2} && \\ && \ddots & \\ &&& \mathbf{0} \end{bmatrix} + \cdots + \lambda_s \begin{bmatrix} \mathbf{0} &&& \\ & \mathbf{0} && \\ && \ddots & \\ &&& \boldsymbol{I}_{r_s} \end{bmatrix}
$$

$$
= \sum_{i=1}^{s} \lambda_i \begin{bmatrix} \mathbf{0} &&&& \\ & \ddots &&& \\ && \boldsymbol{I}_{r_i} && \\ &&& \ddots & \\ &&&& \mathbf{0} \end{bmatrix}
$$

令 $\boldsymbol{Q}_1 = \begin{bmatrix} \boldsymbol{I}_{r_1} &&& \\ & \mathbf{0} && \\ && \ddots & \\ &&& \mathbf{0} \end{bmatrix}$，$\boldsymbol{Q}_2 = \begin{bmatrix} \mathbf{0} &&& \\ & \boldsymbol{I}_{r_2} && \\ && \ddots & \\ &&& \mathbf{0} \end{bmatrix}$，$\boldsymbol{Q}_s = \begin{bmatrix} \mathbf{0} &&& \\ & \mathbf{0} && \\ && \ddots & \\ &&& \boldsymbol{I}_{r_s} \end{bmatrix}$，则 $\boldsymbol{Q}_i$ 满足以下性质：

(1) $\sum_{i=1}^{s} \boldsymbol{Q}_i = \boldsymbol{I}_n$；

(2) $\boldsymbol{Q}_i^2 = \boldsymbol{Q}_i$，$i = 1, 2, \cdots, s$；

(3) $\boldsymbol{Q}_i \cdot \boldsymbol{Q}_j = 0$，$i \neq j$。

带入 (3.4) 式，则有：

$$
\boldsymbol{A} = \boldsymbol{P}\left(\sum_{i=1}^{s} \lambda_i \boldsymbol{Q}_i\right)\boldsymbol{P}^{-1} = \sum_{i=1}^{s} \lambda_i (\boldsymbol{P}\boldsymbol{Q}_i\boldsymbol{P}^{-1})
$$

令 $\boldsymbol{P}_i = \boldsymbol{P}\boldsymbol{Q}_i\boldsymbol{P}^{-1}$，有 $\boldsymbol{A} = \sum_{i=1}^{s} \lambda_i \boldsymbol{P}_i$，并且 $\boldsymbol{P}_i$ 具有以下性质：

(1) $\sum_{i=1}^{s} \boldsymbol{P}_i = \boldsymbol{I}_n$;

(2) $\boldsymbol{P}_i^2 = \boldsymbol{P}_i$，$i = 1, 2, \cdots, s$;

(3) $\boldsymbol{P}_i \cdot \boldsymbol{P}_j = 0$，$i \neq j$。

定理 3.2 方阵 $\boldsymbol{P} \in F^{n\times n}$，若满足 $\boldsymbol{P}^2 = \boldsymbol{P}$，则称 $\boldsymbol{P}$ 为幂等矩阵。幂等矩阵 $\boldsymbol{P}$ 有如下性质：

(1) $\boldsymbol{P}^{\mathcal{H}}$ 和 $(\boldsymbol{I} - \boldsymbol{P})$ 仍为幂等矩阵；

(2) $\boldsymbol{P}$ 的特征值为 1 或者是 0，而且 $\boldsymbol{P}$ 可相似于对角矩阵；

(3) $F^n = N(\boldsymbol{P}) \oplus R(\boldsymbol{P})$。

定理 3.3 (可对角化矩阵的谱分解) 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，$\boldsymbol{A}$ 的谱为 $\{\lambda_1, \lambda_2, \cdots, \lambda_s\}$，则 $\boldsymbol{A}$ 可对角化的充分必要条件是 $\boldsymbol{A}$ 有如下分解式

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

其中方阵 $\boldsymbol{P}_i \in \mathbb{C}^{n\times n}$，满足如下条件：

(1) $\boldsymbol{P}_i^2 = \boldsymbol{P}_i$，$i = 1, 2, \cdots, s$;

(2) $\boldsymbol{P}_i \cdot \boldsymbol{P}_j = 0$，$i \neq j$;

(3) $\sum_{i=1}^{s} \boldsymbol{P}_i = \boldsymbol{I}_n$。

定理 3.4 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 是半正定的 *Hermite* 矩阵，$\mathrm{rank}(\boldsymbol{A}) = k$，则 $\boldsymbol{A}$ 可被分解为下列矩阵的和

$$\boldsymbol{A} = \boldsymbol{v}_1\boldsymbol{v}_1^{\mathcal{H}} + \boldsymbol{v}_2\boldsymbol{v}_2^{\mathcal{H}} + \cdots + \boldsymbol{v}_k\boldsymbol{v}_k^{\mathcal{H}}$$

其中 $\boldsymbol{v}_i \in F^n$，$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k\}$ 是空间 F^n 中非零的正交向量组。

3.2 Schur 分解与正规矩阵

定理 3.5 (UR 分解) 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 为可逆矩阵，则存在酉矩阵 $\boldsymbol{U} \in \mathbb{C}^{n\times n}$ 和主对角线上元素皆为正的上三角矩阵

$$\boldsymbol{R} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix}，r_{ii} > 0；i = 1, 2, \cdots, n$$

使

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

其中 $\boldsymbol{U}$ 为 $\boldsymbol{A}$ 的列向量组标准正交化 $\begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \end{bmatrix}$，$\boldsymbol{R} = \begin{bmatrix} \|\boldsymbol{\beta}_1\| & (\boldsymbol{\alpha}_2, \varepsilon_1) & \cdots & (\boldsymbol{\alpha}_n, \varepsilon_1) \\ & \|\boldsymbol{\beta}_2\| & \cdots & (\boldsymbol{\alpha}_n, \varepsilon_2) \\ & & \ddots & \vdots \\ & & & \|\boldsymbol{\beta}_n\| \end{bmatrix}$

定理 3.6 (QR 分解) 设 $\boldsymbol{A} \in \mathbb{C}^{m\times k}$ 是一个列满秩 $(m \geq k)$ 的矩阵，即 $\text{rank}(\boldsymbol{A}) = k$，则 $\boldsymbol{A}$ 可被分解为

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

其中 $\boldsymbol{Q} \in \mathbb{C}^{m\times k}$，$\boldsymbol{Q}$ 的列向量是 A 的列空间的标准正交基，$\boldsymbol{R} \in \mathbb{C}^{k\times k}$ 是一个可逆的上三角矩阵。

定理 3.7 (Schur 分解) 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，则存在酉矩阵 $\boldsymbol{U}$ 和上三角矩阵 $\boldsymbol{T}$，使得

$$\boldsymbol{U}^{\mathcal{H}}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{T} = \begin{bmatrix} \lambda_1 & t_{12} & \cdots & t_{1n} \\ & \lambda_2 & \cdots & t_{2n} \\ & & \ddots & \vdots \\ & & & \lambda_n \end{bmatrix}$$

其中 λ_i 为矩阵 $\boldsymbol{A}$ 特征值，$i = 1, 2, \cdots, n$。

定义 3.5 若矩阵 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 满足 $\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{\mathcal{H}}$，则称 $\boldsymbol{A}$ 是一个**正规矩阵**。常见的正规矩阵有：

(1) 对角矩阵；

(2) 对称与反对称矩阵：$\boldsymbol{A}^T = \boldsymbol{A}$，$\boldsymbol{A}^T = -\boldsymbol{A}$；

(3) *Hermite* 与反 *Hermite* 矩阵：$\boldsymbol{A}^{\mathcal{H}} = \boldsymbol{A}$，$\boldsymbol{A}^{\mathcal{H}} = -\boldsymbol{A}$；

(4) 正交矩阵与酉矩阵：$\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}_n$，$\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{\mathcal{H}} = \boldsymbol{I}_n$。

定理 3.8 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 是正规矩阵的充分必要条件是 $\boldsymbol{A}$ 酉相似于对角矩阵，即存在酉矩阵 $\boldsymbol{U} \in \mathbb{C}^{n\times n}$，使

$$\boldsymbol{U}^{\mathcal{H}}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{T} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \tag{3.3}$$

推论：$\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 是正规矩阵的充分必要条件是 $\boldsymbol{A}$ 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。

定理 3.9 (正规矩阵的谱分解) 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，$\boldsymbol{A}$ 的谱为 $\{\lambda_1, \lambda_2, \cdots, \lambda_s\}$，$s \leq n$，则 $\boldsymbol{A}$ 是正规矩阵的充分必要条件是 $\boldsymbol{A}$ 有如下谱分解

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

其中方阵 $\boldsymbol{P}_i \in \mathbb{C}^{n\times n}$，满足如下条件：

(1) $\boldsymbol{P}_i^2 = \boldsymbol{P}_i$，$\boldsymbol{P}_i^{\mathcal{H}} = \boldsymbol{P}_i$，$i = 1, 2, \cdots, s$；

(2) $\boldsymbol{P}_i \cdot \boldsymbol{P}_j = 0$，$i \neq j$；

(3) $\displaystyle\sum_{i=1}^{s} \boldsymbol{P}_i = \boldsymbol{I}_n$。

3.3 矩阵的奇异值分解

定理 3.10 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，则矩阵 $\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 和矩阵 $\boldsymbol{A}\boldsymbol{A}^{\mathcal{H}} \in \mathbb{C}^{m\times m}$ 具有如下性质：

(1) $\mathrm{rank}(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}\boldsymbol{A}^{\mathcal{H}}) = r$

(2) $\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}$ 和 $\boldsymbol{A}\boldsymbol{A}^{\mathcal{H}}$ 的非零特征值相等。

(3) $\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathcal{H}}$ 都是半正定矩阵，当 $\mathrm{rank}(\boldsymbol{A}) = n$ 时，$\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}$ 为正定矩阵，当 $\mathrm{rank}(\boldsymbol{A}) = m$ 时，$\boldsymbol{A}\boldsymbol{A}^{\mathcal{H}}$ 为正定矩阵。

定义 3.6 对于 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，$\mathrm{rank}(\boldsymbol{A}) = r$，矩阵 $\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}$ 的特征值为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$，$\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$，称正数 $\sigma_i = \sqrt{\lambda_i}$ 为矩阵 $\boldsymbol{A}$ 的**奇异值**，简称 $\boldsymbol{A}$ 的**奇值**。

定理 3.11 矩阵 $\boldsymbol{A}$ 的奇异值具有如下性质：

(1) $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 为正规矩阵时，$\boldsymbol{A}$ 的奇异值为 $\boldsymbol{A}$ 的特征值的模 $|\lambda_i|$，$i = 1, 2, \cdots, r$；

(2) $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 为正定的 *Hermite* 矩阵时，$\boldsymbol{A}$ 的奇异值等于 $\boldsymbol{A}$ 的特征值；

(3) 若存在酉矩阵 $\boldsymbol{U} \in \mathbb{C}^{m\times m}$，$\boldsymbol{V} \in \mathbb{C}^{n\times n}$，矩阵 $\boldsymbol{B} \in \mathbb{C}^{m\times n}$，使 $\boldsymbol{U}\boldsymbol{A}\boldsymbol{V} = \boldsymbol{B}$，则称 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 酉等价，酉等价的矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 有相同的奇异值。

定理 3.12 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，$\mathrm{rank}(\boldsymbol{A}) = r$，$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ 是矩阵 $\boldsymbol{A}$ 的奇异值，则存在酉矩阵 $\boldsymbol{U} \in \mathbb{C}^{m\times m}$，$\boldsymbol{V} \in \mathbb{C}^{n\times n}$，分块矩阵 $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Delta} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \in \mathbb{C}^{m\times n}$，使

$$\boldsymbol{A} = \boldsymbol{U} \begin{bmatrix} \boldsymbol{\Delta} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \boldsymbol{V}^{\mathcal{H}}$$

其中 $\boldsymbol{\Delta} = \begin{bmatrix} \boldsymbol{\sigma_1} & & & \\ & \boldsymbol{\sigma_2} & & \\ & & \ddots & \\ & & & \boldsymbol{\sigma_r} \end{bmatrix}$

奇异值分解的求法

(1) 由特征多项式 $|\lambda\boldsymbol{I} - \boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}| = \boldsymbol{0}$ 求得特征值 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$，$\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$，以及每个特征值对应的特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$；

(2) 对特征向量进行施密特正交化和单位化，得到正交向量组 $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$，则 $\boldsymbol{V} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix}$；

(3) 对于奇异值 $\sigma_1, \sigma_2, \cdots, \sigma_r$，有 $\boldsymbol{\mu}_i = \frac{1}{\sigma_i}\boldsymbol{A}\boldsymbol{v}_i$，则得到 $\mathbb{C}^m$ 中标准正交的向量组 $\{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \cdots, \boldsymbol{\mu}_r\}$，把它扩充为 $\mathbb{C}^m$ 中的标准正交基 $\{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \cdots, \boldsymbol{\mu}_r, \boldsymbol{\mu}_{r+1}, \cdots \boldsymbol{\mu}_m\}$，则 $\boldsymbol{U} = \begin{bmatrix} \boldsymbol{\mu}_1 & \boldsymbol{\mu}_2 & \cdots & \boldsymbol{\mu}_r & \boldsymbol{\mu}_{r+1} & \cdots & \boldsymbol{\mu}_m \end{bmatrix}$。

定理 3.13 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，$\boldsymbol{A}$ 的奇异值为 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$，令 $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \cdots, \boldsymbol{\mu}_m$ 是相应于奇异值的左奇异向量，$\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ 是相应于奇异值的右奇异向量，则

$$\boldsymbol{A} = \sigma_1\boldsymbol{\mu}_1\boldsymbol{v}_1^{\mathcal{H}} + \sigma_2\boldsymbol{\mu}_2\boldsymbol{v}_2^{\mathcal{H}} + \cdots + \sigma_r\boldsymbol{\mu}_r\boldsymbol{v}_r^{\mathcal{H}} \tag{3.4}$$

定理 3.14 (极分解) 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，$\operatorname{rank}(\boldsymbol{A}) = r$，则 $\boldsymbol{A}$ 可以被分解为

$$\boldsymbol{A} = \boldsymbol{U\Sigma V}^{\mathcal{H}} = \boldsymbol{U\Sigma U}^{\mathcal{H}}\boldsymbol{UV}^{\mathcal{H}} = (\boldsymbol{U\Sigma U}^{\mathcal{H}})(\boldsymbol{UV}^{\mathcal{H}}) = \boldsymbol{PQ}$$

其中 $\boldsymbol{P}$ 是秩为 r 的 $n\times n$ 阶半正定矩阵，$\boldsymbol{Q}$ 是 $n\times n$ 阶的酉矩阵。特别地，若 $r = n$，则 $\boldsymbol{P}$ 为正定矩阵。

4 矩阵的广义逆

4.1 矩阵的左逆与右逆

定义 4.1 (矩阵的左逆) 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，若存在矩阵 $\boldsymbol{B} \in \mathbb{C}^{n\times m}$，使得

$$\boldsymbol{B}\boldsymbol{A} = \boldsymbol{I}_n$$

则称 $\boldsymbol{A}$ 是**左可逆**的，称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的一个**左逆矩阵**，记为 $\boldsymbol{A}_L^{-1}$，且下列条件是等价的。

(1) $\boldsymbol{A}$ 是左可逆的；

(2) $\boldsymbol{A}$ 的零空间 $N(\boldsymbol{A}) = \{\boldsymbol{0}\}$；

(3) $m \geq n$，$\operatorname{rank}(\boldsymbol{A}) = n$，即 $\boldsymbol{A}$ 是列满秩的；

(4) $\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}$ 是可逆的，$(\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathcal{H}}$ 是 $\boldsymbol{A}$ 的一个左逆矩阵。

定义 4.2 (矩阵的右逆) 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，若存在矩阵 $\boldsymbol{C} \in \mathbb{C}^{n\times m}$，使得

$$\boldsymbol{A}\boldsymbol{C} = \boldsymbol{I}_m$$

则称 $\boldsymbol{A}$ 是**右可逆**的，称 $\boldsymbol{C}$ 为 $\boldsymbol{A}$ 的一个**右逆矩阵**，记为 $\boldsymbol{A}_R^{-1}$，且下列条件是等价的。

(1) $\boldsymbol{A}$ 是右可逆的；

(2) $\boldsymbol{A}$ 的列空间 $R(\boldsymbol{A}) = \boldsymbol{C}^m$；

(3) $m \leq n$，$\operatorname{rank}(\boldsymbol{A}) = m$，即 $\boldsymbol{A}$ 是行满秩的；

(4) $\boldsymbol{A}\boldsymbol{A}^{\mathcal{H}}$ 是可逆的，$\boldsymbol{A}^{\mathcal{H}}(\boldsymbol{A}\boldsymbol{A}^{\mathcal{H}})^{-1}$ 是 $\boldsymbol{A}$ 的一个右逆矩阵。

定理 4.1 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ 是左可逆的，$\boldsymbol{B} \in \mathbb{C}^{n\times m}$ 是 $\boldsymbol{A}$ 的一个左逆矩阵，则线性方程组 $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{b}$ 有形如 $\boldsymbol{X} = \boldsymbol{B}\boldsymbol{b}$ 解的充分必要条件是

$$(\boldsymbol{I}_m - \boldsymbol{A}\boldsymbol{B})\boldsymbol{b} = 0$$

若上式成立，则方程组有惟一解

$$\boldsymbol{X} = \left(\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathcal{H}}\boldsymbol{b}$$

定理 4.2 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ 是右可逆的，则线性方程组 $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{b}$ 对任何 $\boldsymbol{b} \in \boldsymbol{C}^m$ 都有解，且对 $\boldsymbol{A}$ 的任意一个右逆矩阵 $\boldsymbol{A}_R^{-1}$，$\boldsymbol{X} = \boldsymbol{A}_R^{-1}\boldsymbol{b}$ 是其解。特别地，$\boldsymbol{X} = \boldsymbol{A}^{\mathcal{H}}\left(\boldsymbol{A}\boldsymbol{A}^{\mathcal{H}}\right)^{-1}\boldsymbol{b}$ 是方程组 $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{b}$ 的一个解。

4.2 广义逆矩阵

定义 4.3 (加号广义逆) 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，若存在矩阵 $\boldsymbol{G} \in \mathbb{C}^{n\times m}$，使得

(1) $\boldsymbol{A}\boldsymbol{G}\boldsymbol{A} = \boldsymbol{A}$；

(2) $\boldsymbol{G}\boldsymbol{A}\boldsymbol{G} = \boldsymbol{G}$；

(3) $(\boldsymbol{AG})^{\mathcal{H}} = \boldsymbol{AG}$;

(4) $(\boldsymbol{GA})^{\mathcal{H}} = \boldsymbol{GA}$。

则称 $\boldsymbol{G}$ 为 $\boldsymbol{A}$ 的 ***Moore-Penrose* 广义逆**或**加号广义逆**，简称为 $\boldsymbol{A}$ 的 ***M-P* 逆**。$\boldsymbol{A}$ 的任意 *M-P* 逆记为 $\boldsymbol{A}^{+}$。

定理 4.3 若矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ 存在 *M-P* 广义逆，则 $\boldsymbol{A}$ 的 *M-P* 逆是惟一的。

定理 4.4 任意矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ 都存在 *M-P* 广义逆 $\boldsymbol{A}^{+}$。设 $\operatorname{rank}(\boldsymbol{A}) = r$，$\boldsymbol{A}$ 的一个满秩分解为

$$\boldsymbol{A} = \boldsymbol{BC},\ \boldsymbol{B} \in \mathbb{C}^{m\times r},\ \boldsymbol{C} \in \mathbb{C}^{r\times n},\ \operatorname{rank}(\boldsymbol{B}) = \operatorname{rank}(\boldsymbol{C}) = r$$

则 $\boldsymbol{A}^{+} = \boldsymbol{C}_R^{-1}\boldsymbol{B}_L^{-1} = \boldsymbol{C}^{\mathcal{H}}\left(\boldsymbol{CC}^{\mathcal{H}}\right)^{-1}\left(\boldsymbol{B}^{\mathcal{H}}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathcal{H}}$

定理 4.5 任意矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}, \operatorname{rank}(\boldsymbol{A}) = r$，$\boldsymbol{A}$ 的一个奇异值分解为

$$\boldsymbol{A} = \boldsymbol{U}\begin{bmatrix} \boldsymbol{\Delta} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{V}^{\mathbb{H}}$$

则 $\boldsymbol{A}^{+} = \boldsymbol{V}\begin{bmatrix} \boldsymbol{\Delta}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}\boldsymbol{U}^{\mathbb{H}}$，其中 $\boldsymbol{\Delta}$ 为对角线由 $\boldsymbol{A}$ 的正奇异值所构成的对角矩阵，$\boldsymbol{\Delta} \in \mathbb{C}^{r\times r}$。

定理 4.6 设 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，$\lambda \in \mathbb{C}^{n}$，则 $\boldsymbol{A}^{+}$ 满足以下性质：

(1) $(\boldsymbol{A}^{+})^{+} = \boldsymbol{A}$;

(2) $(\boldsymbol{A}^{+})^{\mathcal{H}} = \left(\boldsymbol{A}^{\mathcal{H}}\right)^{+}$;

(3) $(\lambda\boldsymbol{A})^{+} = \lambda^{+}\boldsymbol{A}^{+}$，其中 $\lambda^{+} = \begin{cases} \dfrac{1}{\lambda}, & \lambda \neq 0 \\ 0, & \lambda = 0 \end{cases}$;

(4) 若 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ 是列满秩的，则 $\boldsymbol{A}^{+} = \boldsymbol{A}_L^{-1} = \left(\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathcal{H}}$;

(5) 若 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ 是行满秩的，则 $\boldsymbol{A}^{+} = \boldsymbol{A}_R^{-1} = \boldsymbol{A}^{\mathcal{H}}\left(\boldsymbol{AA}^{\mathcal{H}}\right)^{-1}$;

(6) 若 $\boldsymbol{A}$ 有满秩分解 $\boldsymbol{A} = \boldsymbol{BC}$，则 $\boldsymbol{A}^{+} = \boldsymbol{C}^{+}\boldsymbol{B}^{+} = \boldsymbol{C}_R^{-1}\boldsymbol{B}_L^{-1} = \boldsymbol{C}^{\mathcal{H}}\left(\boldsymbol{CC}^{\mathcal{H}}\right)^{-1}\left(\boldsymbol{B}^{\mathcal{H}}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathcal{H}}$;

(7) $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A}^{+})$;

(8) $\operatorname{rank}(\boldsymbol{A}^{+}\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{AA}^{+}) = \operatorname{rank}(\boldsymbol{A})$。

4.3 投影变换

定义 4.4 设 $\mathbb{C}^{n} = \boldsymbol{L} \oplus \boldsymbol{M}$，$\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$，$\boldsymbol{y} \in \boldsymbol{L}$，$\boldsymbol{z} \in \boldsymbol{M}$。如果线性变换 $\sigma : \mathbb{C}^{n} \to \mathbb{C}^{n}$ 满足 $\sigma(\boldsymbol{x}) = \boldsymbol{y}$，则称 σ 是从 $\mathbb{C}^{n}$ 沿子空间 $\boldsymbol{M}$ 到子空间 $\boldsymbol{L}$ 上的**投影变换**，投影变换在 $\mathbb{C}^{n}$ 空间的一组基下的矩阵称为**投影矩阵**。

投影变换 σ 把 $\mathbb{C}^{n}$ 映射成子空间 $\boldsymbol{L}$，故子空间 $\boldsymbol{L}$ 称为投影子空间，显然它就是 σ 的像空间 $\boldsymbol{R}(\sigma)$，子空间 $\boldsymbol{M}$ 是投影变换的核空间 $\boldsymbol{N}(\sigma)$，这时 $\mathbb{C}^{n}$ 空间的直和分解为

$$\mathbb{C}^{n} = \boldsymbol{R}(\sigma) \oplus \boldsymbol{N}(\sigma)$$

定理 4.7 $\mathbb{C}^n$ 空间上的线性变换 σ 是投影变换的充分必要条件是 σ 是幂等变换，即 $\sigma^2=\sigma$。

推论：$\mathbb{C}^n$ 空间上的线性变换 σ 是投影变换的充分必要条件是 σ 关于某组基下的矩阵 $\boldsymbol{A}$ 为幂等矩阵，即 $\boldsymbol{A}^2=\boldsymbol{A}$。

求法：(1) 在子空间 $\boldsymbol{L}$ 和 $\boldsymbol{M}$ 中分别取定基底 $\{\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_r\}$ 和 $\{\boldsymbol{z}_{r+1},\boldsymbol{z}_{r+2},\cdots,\boldsymbol{z}_n\}$，得到矩阵 $\boldsymbol{B}=\begin{bmatrix}\boldsymbol{y}_1 & \boldsymbol{y}_2 & \cdots & \boldsymbol{y}_r\end{bmatrix}$ 和矩阵 $\boldsymbol{C}=\begin{bmatrix}\boldsymbol{z}_{r+1} & \boldsymbol{z}_{r+2} & \cdots & \boldsymbol{z}_n\end{bmatrix}$

(2) 于是投影矩阵 $\boldsymbol{A}=\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{0}\end{array}\right]\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{C}\end{array}\right]^{-1}$

定义 4.5 设 σ 是 $\mathbb{C}^n$ 空间的投影变换，$\mathbb{C}^n=\boldsymbol{R}(\sigma)\oplus\boldsymbol{N}(\sigma)$。如果 $\boldsymbol{R}(\sigma)$ 的正交补子空间 $\boldsymbol{R}(\sigma)^\perp=\boldsymbol{N}(\sigma)$，则称 σ 是 $\mathbb{C}^n$ 空间的**正交投影变换**。正交投影变换在 $\mathbb{C}^n$ 空间的一组**标准正交基**下的矩阵称为**正交投影矩阵**。

定理 4.8 $\mathbb{C}^n$ 空间上的线性变换 σ 是正交投影变换的充分必要条件是 σ 关于某组基下的矩阵 $\boldsymbol{A}$ 为幂等的 *Hermite* 矩阵，即 $\boldsymbol{A}^2=\boldsymbol{A}$，$\boldsymbol{A}^{\mathcal{H}}=\boldsymbol{A}$

求法：$\boldsymbol{A}=\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{0}\end{array}\right]\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{C}\end{array}\right]^{-1}=\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{0}\end{array}\right]\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{C}\end{array}\right]^{-1}\begin{bmatrix}\boldsymbol{B}^{\mathcal{H}}\\ \boldsymbol{C}^{\mathcal{H}}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{B}^{\mathcal{H}}\\ \boldsymbol{C}^{\mathcal{H}}\end{bmatrix}=\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{0}\end{array}\right]\left(\begin{bmatrix}\boldsymbol{B}^{\mathcal{H}}\\ \boldsymbol{C}^{\mathcal{H}}\end{bmatrix}\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{C}\end{array}\right]\right)^{-1}\begin{bmatrix}\boldsymbol{B}^{\mathcal{H}}\\ \boldsymbol{C}^{\mathcal{H}}\end{bmatrix}=\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{0}\end{array}\right]\begin{bmatrix}\boldsymbol{B}^{\mathcal{H}}\boldsymbol{B} & \boldsymbol{B}^{\mathcal{H}}\boldsymbol{C}\\ \boldsymbol{C}^{\mathcal{H}}\boldsymbol{B} & \boldsymbol{C}^{\mathcal{H}}\boldsymbol{C}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{B}^{\mathcal{H}}\\ \boldsymbol{C}^{\mathcal{H}}\end{bmatrix}=\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{0}\end{array}\right]\begin{bmatrix}\boldsymbol{B}^{\mathcal{H}}\boldsymbol{B} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{C}^{\mathcal{H}}\boldsymbol{C}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{B}^{\mathcal{H}}\\ \boldsymbol{C}^{\mathcal{H}}\end{bmatrix}=\left[\begin{array}{c|c}\boldsymbol{B} & \boldsymbol{0}\end{array}\right]\begin{bmatrix}\left(\boldsymbol{B}^{\mathcal{H}}\boldsymbol{B}\right)^{-1} & \boldsymbol{0}\\ \boldsymbol{0} & \left(\boldsymbol{C}^{\mathcal{H}}\boldsymbol{C}\right)^{-1}\end{bmatrix}\begin{bmatrix}\boldsymbol{B}^{\mathcal{H}}\\ \boldsymbol{C}^{\mathcal{H}}\end{bmatrix}=\boldsymbol{B}\left(\boldsymbol{B}^{\mathcal{H}}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathcal{H}}$

定理 4.9 设 $\boldsymbol{A}\in\mathbb{C}^{m\times n}$，则

(1) $(\boldsymbol{A}^+\boldsymbol{A})^2=\boldsymbol{A}^+\boldsymbol{A}$，$(\boldsymbol{A}^+\boldsymbol{A})^{\mathcal{H}}=\boldsymbol{A}^+\boldsymbol{A}$

(2) $\begin{cases}\mathbb{C}^n=\boldsymbol{R}(\boldsymbol{A}^+)\oplus\boldsymbol{N}(\boldsymbol{A})\\ \boldsymbol{R}(\boldsymbol{A}^+)^\perp=\boldsymbol{N}(\boldsymbol{A})\end{cases}$

(3) $(\boldsymbol{A}\boldsymbol{A}^+)^2=\boldsymbol{A}\boldsymbol{A}^+$，$(\boldsymbol{A}\boldsymbol{A}^+)^{\mathcal{H}}=\boldsymbol{A}\boldsymbol{A}^+$

(4) $\begin{cases}\mathbb{C}^m=\boldsymbol{R}(\boldsymbol{A})\oplus\boldsymbol{N}(\boldsymbol{A}^+)\\ \boldsymbol{R}(\boldsymbol{A})^\perp=\boldsymbol{N}(\boldsymbol{A}^+)\end{cases}$

定理 4.10 设 W 是 $\mathbb{C}^n$ 的子空间，$\boldsymbol{x}_0\in\mathbb{C}^n$，$\boldsymbol{x}_0\notin W$。如果 σ 是 $\mathbb{C}^n$ 空间向 W 的正交投影变换，则 $\sigma(\boldsymbol{x}_0)$ 是 W 中离 $\boldsymbol{x}_0$ 最近的向量，即对欧几里得范数 $\|\cdot\|$，都有

$$\|\sigma(\boldsymbol{x}_0)-\boldsymbol{x}_0\|\le\|\boldsymbol{y}-\boldsymbol{x}_0\|,\ \forall\boldsymbol{y}\in W$$

4.4 最佳的最小二乘解

定理 4.11 $\mathbb{A}\in\mathbb{C}^{m\times n}$，$\boldsymbol{b}\in\mathbb{C}^m$，则 $\boldsymbol{x}_0=\boldsymbol{A}^+\boldsymbol{b}$ 是线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的最佳的最小二乘解。

定理 4.12 $\mathbb{A}\in\mathbb{C}^{m\times n}$，$\boldsymbol{B}\in\mathbb{C}^{m\times k}$，则 $\boldsymbol{X}_0=\boldsymbol{A}^+\boldsymbol{B}$ 是线性方程组 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$ 的最佳的最小二乘解。

5　矩阵分析

5.1　向量范数

定义 5.1 设 V 是数域 F 上的线性空间，且对于 V 的任一个向量 $\boldsymbol{x}$，对应一个非负实数 $\|\boldsymbol{x}\|$，满足以下条件：

(1) 正定性：$\|\boldsymbol{x}\| \geq 0$，$\|\boldsymbol{x}\| = 0$ 当且仅当 $\boldsymbol{x} = \boldsymbol{0}$；

(2) 齐次性：$\|a\boldsymbol{x}\| = |a| \cdot \|\boldsymbol{x}\|$，$a \in F$；

(3) 三角不等式：对任意 $\boldsymbol{x}$，$\boldsymbol{y} \in V$，都有 $\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$。

则称 $\|\boldsymbol{x}\|$ 为向量 $\boldsymbol{x}$ 的**范数**，$[V; \|\cdot\|]$ 为**赋范空间**。

定理 5.1 (重要的向量范数) 在 n 维酉空间 $\mathbb{C}^n$ 中，复向量 $\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$，有

(1) 2-范数：$\|\boldsymbol{x}\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$；

(2) 1-范数：$\|\boldsymbol{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n|$；

(3) ∞-范数：$\|\boldsymbol{x}\|_\infty = \max\limits_i |x_i|$；

(4) p-范数：$\|\boldsymbol{x}\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p} = \left(\sum\limits_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$。

5.2　矩阵范数

定义 5.2 在 $F^{n\times n}$ 上定义一个非负实值函数，使得对任意矩阵 $\boldsymbol{A} \in F^{n\times n}$ 是数域 F，对应一个非负实数 $\|\boldsymbol{A}\|$，满足以下四个条件：

(1) 正定性：$\|\boldsymbol{A}\| \geq 0$，$\|\boldsymbol{A}\| = 0$ 当且仅当 $\boldsymbol{A} = \boldsymbol{0}$；

(2) 齐次性：$\|a\boldsymbol{A}\| = |a| \cdot \|\boldsymbol{A}\|$，$a \in F$；

(3) 三角不等式：对任意 $\boldsymbol{A}$，$\boldsymbol{B} \in F^{n\times n}$，都有 $\|\boldsymbol{A}+\boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$；

(4) 相容性：对任意 $\boldsymbol{A}$，$\boldsymbol{B} \in F^{n\times n}$，都有 $\|\boldsymbol{A}\boldsymbol{B}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\|$。

则称 $\|\boldsymbol{A}\|$ 为**矩阵$\boldsymbol{A}$** 的**范数**。

定义 5.3 设 $\|\boldsymbol{x}\|$ 是向量范数，$\|\boldsymbol{A}\|$ 是矩阵范数，若

$$\|\boldsymbol{A}\boldsymbol{x}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|$$

则称矩阵范数 $\|\boldsymbol{A}\|$ 与向量范 $\|\boldsymbol{x}\|$ 数是**相容**的。

定理 5.2 设 $\|\boldsymbol{x}\|$ 是向量范数，则

$$\|\boldsymbol{A}\| = \max_{x\neq 0} \left\{ \frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \right\}$$

是与向量范数 $\|\boldsymbol{x}\|$ 相容的矩阵范数。称其为由向量范数 $\|\boldsymbol{x}\|$ 所诱导的诱导范数。

定理 5.3 (重要的矩阵范数和诱导范数) 设 $\boldsymbol{A}=\left(a_{ij}\right)\in\mathbb{C}^{n\times n}$，$\boldsymbol{A}$ 的奇异值为 $\sigma_1,\sigma_2,\cdots,\sigma_r$，由 $\|\boldsymbol{x}\|_p$ 所诱导的矩阵范数称为矩阵 p-范数。常用的 p-范数为 $\|\boldsymbol{A}\|_1$，$\|\boldsymbol{A}\|_2$ 与 $\|\boldsymbol{A}\|_\infty$，则有

(1) F-范数：$\|\boldsymbol{A}\|_F=\sqrt{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}|a_{ij}|^2}=\sqrt{\operatorname{tr}\left(\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}\right)}=\sqrt{\sigma_1^2+\sigma_2^2+\cdots+\sigma_r^2}$；

(2) 列和范数：$\|\boldsymbol{A}\|_1=\max\limits_{j}\left(\sum\limits_{i=1}^{n}|a_{ij}|\right)$；

(3) 谱范数：$\|\boldsymbol{A}\|_2=\sqrt{\lambda_1}=\sigma_1$，$\lambda_1$ 为 $\boldsymbol{A}^{\mathcal{H}}\boldsymbol{A}$ 的最大特征值；

(4) 行和范数：$\|\boldsymbol{A}\|_\infty=\max\limits_{i}\left(\sum\limits_{j=1}^{n}|a_{ij}|\right)$。

5.3 向量序列和矩阵序列的极限

定义 5.4 设 $\boldsymbol{x}^{(k)}=\begin{bmatrix}x_1^{(k)} & x_2^{(k)} & \cdots & x_n^{(k)}\end{bmatrix}^T$，$k=1,2,\cdots$ 是 $\mathbb{C}^n$ 空间的一个向量序列，如果当 $k\to+\infty$ 时，它的 n 个分量数列都收敛，即

$$\lim_{k\to+\infty}x_i^{(k)}=a_i,\ i=1,2,\cdots,n$$

则称向量序列 $\left\{\boldsymbol{x}^{(k)}\right\}$ 是按分量收敛的。向量 $\boldsymbol{\alpha}=\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}^T$ 是它的极限，记为 $\lim\limits_{k\to+\infty}\boldsymbol{x}^{(k)}=\boldsymbol{\alpha}$ 或 $\boldsymbol{x}^{(k)}\to\boldsymbol{\alpha}$。如果至少有一个分量数列是发散的，则称向量序列是发散的。

定义 5.5 设 $\left\{\boldsymbol{x}^{(k)}\right\}$ 是 $\mathbb{C}^n$ 空间的一个向量序列，$\|\boldsymbol{x}\|$ 是 $\mathbb{C}^n$ 空间的一个向量范数。如果存在向量 $\boldsymbol{\alpha}\in\mathbb{C}^n$，当 $k\to+\infty$ 时，$\left\|\boldsymbol{x}^{(k)}-\boldsymbol{\alpha}\right\|\to 0$，则称向量序列**按向量范数收敛于$\boldsymbol{\alpha}$**。

定理 5.4 设 $\left\{\boldsymbol{x}^{(k)}\right\}$ 是 $\mathbb{C}^n$ 空间的一个向量序列，它按分量收敛的充分必要条件是它按 $\mathbb{C}^n$ 空间的任意一个向量范数收敛。

定义 5.6 设 $\boldsymbol{A}^{(k)}=\left(a_{ij}^{(k)}\right)\in\mathbb{C}^{n\times n}$，$\left\{\boldsymbol{A}^{(k)}\right\}$ 一个矩阵序列，如果当 $k\to+\infty$ 时，它的 n^2 个数列 $\left\{a_{ij}^{(k)}\right\}$ 都收敛，即

$$\lim_{k\to+\infty}a_{ij}^{(k)}=a_{ij},\ i,j=1,2,\cdots,n$$

则称矩阵序列 $\left\{\boldsymbol{A}^{(k)}\right\}$ **按元素数列收敛**。向量 $\boldsymbol{A}=\left(a_{ij}\right)\in\mathbb{C}^{n\times n}$ 是它的极限，记为 $\lim\limits_{k\to+\infty}\boldsymbol{A}^{(k)}=\boldsymbol{A}$ 或 $\boldsymbol{A}^{(k)}\to\boldsymbol{A}$。如果至少有一个元素数列是发散的，则称该矩阵序列发散。

矩阵序列收敛满足以下性质：

(1) 设 $\boldsymbol{A}^{(k)}\to\boldsymbol{A}$，$\boldsymbol{B}^{(k)}\to\boldsymbol{B}$，则

$$\alpha\boldsymbol{A}^{(k)}+\beta\boldsymbol{B}^{(k)}\to\alpha\boldsymbol{A}+\beta\boldsymbol{B}$$

(2) 设 $\boldsymbol{A}^{(k)}\to\boldsymbol{A}$，$\boldsymbol{B}^{(k)}\to\boldsymbol{B}$，则

$$\boldsymbol{A}^{(k)}\boldsymbol{B}^{(k)}\to\boldsymbol{A}\boldsymbol{B}$$

(3) 设 $\boldsymbol{A}^{(k)}$ 与 $\boldsymbol{A}$ 都是可逆矩阵，且 $\boldsymbol{A}^{(k)} \to \boldsymbol{A}$，则

$$\left(\boldsymbol{A}^{(k)}\right)^{-1} \to \boldsymbol{A}^{-1}$$

定义 5.7 设 $\left\{\boldsymbol{A}^{(k)}\right\}$ 是 $\mathbb{C}^{n\times n}$ 空间的一个矩阵序列，$\|\boldsymbol{A}\|$ 是 $\mathbb{C}^{n\times n}$ 空间的一个矩阵范数。如果存在矩阵 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，当 $k \to +\infty$ 时，$\left\|\boldsymbol{A}^{(k)} - \boldsymbol{A}\right\| \to 0$，则称矩阵序列**按矩阵范数收敛于$\boldsymbol{A}$**。

定理 5.5 设 $\left\{\boldsymbol{A}^{(k)}\right\}$ 是 $\mathbb{C}^{n\times n}$ 空间的一个矩阵序列，它按元素数列收敛的充分必要条件是它按 $\mathbb{C}^{n\times n}$ 空间的任意一个矩阵范数收敛。

定义 5.8 矩阵序列 $\left\{\boldsymbol{A}^{(k)}\right\}$ 称为有界的，如果存在常数 $M > 0$，使得对所有 $k > 0$ 都有

$$|a_{ij}^{(k)}| < M,\ (i = 1, 2, \cdots, m;\ j = 1, 2, \cdots, n)$$

定义 5.9 设 $\boldsymbol{A}$ 为方阵，且当 $k \to +\infty$ 时有 $\boldsymbol{A}^k \to \boldsymbol{0}$，则称 $\boldsymbol{A}$ 为收敛矩阵。

5.4 矩阵幂级数

定义 5.10 设矩阵 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 的全部特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，则称 $\rho(\boldsymbol{A}) = \max\limits_i |\lambda_i|$ 为 $\boldsymbol{A}$ 的谱半径。由定义知，$\boldsymbol{A}$ 的全部特征值分布在复平面上以原点为中心，$\rho(\boldsymbol{A})$ 为半径的圆盘上。并且有 $\rho\left(\boldsymbol{A}^k\right) = \rho(\boldsymbol{A})^k$，$\rho(k\boldsymbol{A}) = |k|\rho(\boldsymbol{A})$，$\rho(\boldsymbol{A}) = \rho\left(\boldsymbol{A}^T\right)$

定理 5.6 $\boldsymbol{A}^k \to \boldsymbol{0}\,(k \to \infty)$ 的充分必要条件是 $\rho(\boldsymbol{A}) < 1$。

定理 5.7 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，则对 $\mathbb{C}^{n\times n}$ 上任意一种矩阵范数 $\|\boldsymbol{A}\|$，都有

$$\rho(\boldsymbol{A}) \le \|\boldsymbol{A}\|$$

即 $\boldsymbol{A}$ 的谱半径是 $\boldsymbol{A}$ 的任意一种矩阵范数的下界。

定义 5.11 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，$a_k \in \mathbb{C}$，$k = 0, 1, 2, \cdots$ 称

$$a_0\boldsymbol{I} + a_1\boldsymbol{A} + a_2\boldsymbol{A}^2 + \cdots + a_k\boldsymbol{A}^k + \cdots$$

为矩阵 $\boldsymbol{A}$ 的幂级数，记为 $\sum\limits_{k=0}^{\infty} a_k\boldsymbol{A}^k$。

定义 5.12 矩阵幂级数 $\sum\limits_{k=0}^{\infty} a_k\boldsymbol{A}^k$ 的前 $N+1$ 项的和 $S_N = \sum\limits_{k=0}^{\infty} a_k\boldsymbol{A}^k$ 称为**矩阵幂级数的部分和**。若矩阵幂级数 $\sum\limits_{k=0}^{\infty} a_k\boldsymbol{A}^k$ 的部分和序列 $\{S_N(\boldsymbol{A})\}$ 收敛，则称 $\sum\limits_{k=0}^{\infty} a_k\boldsymbol{A}^k$ 收敛；否则，称其为发散。若 $\lim\limits_{N\to\infty} S_N(\boldsymbol{A}) = S$，则称 S 为 $\sum\limits_{k=0}^{\infty} a_k\boldsymbol{A}^k$ 的**和矩阵**。

定理 5.8 若复变量 z 的幂级数 $\sum_{k=0}^{\infty} a_k z^k$ 的收敛半径为 R，而方阵 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 的谱半径为 $\rho(\boldsymbol{A})$，则

(1) 当 $\rho(\boldsymbol{A}) < R$ 时，矩阵幂级数 $\sum_{k=0}^{\infty} a_k \boldsymbol{A}^k$ 收敛；

(2) 当 $\rho(\boldsymbol{A}) > R$ 时，矩阵幂级数 $\sum_{k=0}^{\infty} a_k \boldsymbol{A}^k$ 发散；

(3) 当计算 $\boldsymbol{A}$ 的特征值比较困难时，由于 $\boldsymbol{A}$ 的每个范数都是谱半径 $\rho(\boldsymbol{A})$ 的上界，只要能找到一种特殊的矩阵范数 $\|\boldsymbol{A}\|$，使 $\|\boldsymbol{A}\| < R$，便可断定该矩阵幂级数是收敛的。(优先考虑行和、列和范数)

5.5 矩阵函数

定义 5.13 设 $f(z)$ 是复变量的解析函数，$f(z) = \sum_{k=0}^{\infty} a_k z^k$ 的收敛半径为 R。如果矩阵 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 的谱半径 $\rho(\boldsymbol{A}) < R$，则称

$$f(\boldsymbol{A}) = \sum_{k=0}^{\infty} a_k \boldsymbol{A}^k$$

为 $\boldsymbol{A}$ 的**矩阵函数**。常用的矩阵函数：

(1) $\exp(\boldsymbol{A}) = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{A}^k$，收敛域是整个复平面上；

(2) $\cos \boldsymbol{A} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\boldsymbol{A}^{2k}$，收敛域是整个复平面上；

(3) $\sin \boldsymbol{A} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\boldsymbol{A}^{2k+1}$，收敛域是整个复平面上；

(4) $(\boldsymbol{I} - \boldsymbol{A})^{-1} = \sum_{k=0}^{\infty} \boldsymbol{A}^k$，收敛域是复平面 $|z| < 1$；

(5) $\ln(\boldsymbol{I} + \boldsymbol{A}) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}\boldsymbol{A}^k$，收敛域是复平面 $|z| < 1$；

定理 5.9 (Jordan 标准形法) 设 $f(z)$ 是复变量 z 的解析函数，$\boldsymbol{A} \in \mathbb{C}^{n\times n}$，且存在可逆阵 $\boldsymbol{P}$，使得

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1} = \boldsymbol{P}\,\mathrm{diag}(\boldsymbol{J}_1, \boldsymbol{J}_2, \cdots, \boldsymbol{J}_m)\,\boldsymbol{P}^{-1}$$

则 $f(\boldsymbol{A}) = \boldsymbol{P}f(\boldsymbol{J})\,\boldsymbol{P}^{-1} = \boldsymbol{P}\,\mathrm{diag}(f(\boldsymbol{J}_1), f(\boldsymbol{J}_2), \cdots, f(\boldsymbol{J}_m))\,\boldsymbol{P}^{-1}$

其中

$$f\left(\boldsymbol{J}_i\right)=\begin{bmatrix} f\left(\lambda_i\right) & f'\left(\lambda_i\right) & \dfrac{f''\left(\lambda_i\right)}{2!} & \cdots & \dfrac{f^{(n_i-1)}\left(\lambda_i\right)}{\left(n_i-1\right)!} \\ & f\left(\lambda_i\right) & f'\left(\lambda_i\right) & \ddots & \vdots \\ & & f\left(\lambda_i\right) & \ddots & \dfrac{f''\left(\lambda_i\right)}{2!} \\ & & & \ddots & f'\left(\lambda_i\right) \\ & & & & f\left(\lambda_i\right) \end{bmatrix}$$

若 $\boldsymbol{A}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则 $f\left(\boldsymbol{A}\right)$ 的特征值为 $f\left(\lambda_1\right),f\left(\lambda_2\right),\cdots,f\left(\lambda_n\right)$

定理 5.10 (最小多项式法) 设 n 阶矩阵 $\boldsymbol{A}$ 的最小多项式为

$$m_{\boldsymbol{A}}\left(\lambda\right)=\left(\lambda-\lambda_1\right)^{n_1}\left(\lambda-\lambda_2\right)^{n_2}\cdots\left(\lambda-\lambda_s\right)^{n_s}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 为 $\boldsymbol{A}$ 的所有不同特征值，$\sum\limits_{i=1}^{s}n_i=m$，$f\left(\lambda\right)$ 是复变量 λ 的解析函数，令

$$g\left(\lambda\right)=c_0+c_1\lambda+\cdots+c_{m-1}\lambda^{m-1}$$

则 $f\left(\boldsymbol{A}\right)=g\left(\boldsymbol{A}\right)$ 的充分必要条件是

$$g^{(j)}\left(\lambda_i\right)=f^{(j)}\left(\lambda_i\right),\ i=1,2,\cdots,s,\ j=0,1,2,\cdots,n_i-1$$

5.6 函数矩阵的微分与积分

定义 5.14 现在考虑矩阵元素是实变量 t 的实函数的矩阵

$$\boldsymbol{A}\left(t\right)=\begin{bmatrix} a_{11}\left(t\right) & a_{12}\left(t\right) & \cdots & a_{1n}\left(t\right) \\ a_{21}\left(t\right) & a_{22}\left(t\right) & \cdots & a_{2n}\left(t\right) \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\left(t\right) & a_{m2}\left(t\right) & \cdots & a_{mn}\left(t\right) \end{bmatrix}$$

$\boldsymbol{A}\left(t\right)$ 中所有的元素 $a_{ij}\left(t\right)$ 定义在同一区间 $[a,b]$ 上。$\boldsymbol{A}\left(t\right)$ 在区间 $[a,b]$ 上有界，连续，可微，可积是指所有的 $a_{ij}\left(t\right)$ 在 $[a,b]$ 上有界，连续，可微，可积。例如函数矩阵 $\boldsymbol{A}\left(t\right)$ 的极限、微分与积分定义为

$$\begin{aligned} \lim_{t\to t_0}\boldsymbol{A}\left(t\right) &= \left[\lim_{t\to t_0}a_{ij}\left(t\right)\right]_{m\times n} \\ \frac{\mathrm{d}\boldsymbol{A}\left(t\right)}{\mathrm{d}t} &= \left[\frac{\mathrm{d}a_{ij}\left(t\right)}{\mathrm{d}t}\right]_{m\times n} \\ \int_a^b\boldsymbol{A}\left(t\right)\,\mathrm{d}t &= \left[\int_a^b a_{ij}\left(t\right)\,\mathrm{d}t\right]_{m\times n} \end{aligned}$$

定理 5.11 (微分性质) 函数矩阵的微分满足以下性质：

(1) $\frac{\mathrm{d}}{\mathrm{d}t}\left(a\boldsymbol{A}(t)+b\boldsymbol{B}(t)\right)=a\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)+b\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{B}(t)$

(2) $\frac{\mathrm{d}}{\mathrm{d}t}\left(k(t)\boldsymbol{A}(t)\right))=\frac{\mathrm{d}k(t)}{\mathrm{d}t}\boldsymbol{A}(t)+k(t)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)$

(3) $\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{A}(t)\boldsymbol{B}(t)\right))=\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)\right)\cdot\boldsymbol{B}(t)+\boldsymbol{A}(t)\cdot\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{B}(t)\right)$

(4) $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}^{-1}(t)=-\boldsymbol{A}^{-1}(t)\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)\right)\boldsymbol{A}^{-1}(t)$

(5) $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}^{2}(t)=\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)\right)\boldsymbol{A}(t)+\boldsymbol{A}(t)\left(\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)\right)$

(6) $\frac{\mathrm{d}}{\mathrm{d}t}\exp(\boldsymbol{A}t)=\boldsymbol{A}\exp(\boldsymbol{A}t)=\exp(\boldsymbol{A}t)\boldsymbol{A}$

(7) $\frac{\mathrm{d}}{\mathrm{d}t}\cos(\boldsymbol{A}t)=-\boldsymbol{A}\sin(\boldsymbol{A}t)=-\sin(\boldsymbol{A}t)\boldsymbol{A}$

(8) $\frac{\mathrm{d}}{\mathrm{d}t}\sin(\boldsymbol{A}t)=\boldsymbol{A}\cos(\boldsymbol{A}t)=\cos(\boldsymbol{A}t)\boldsymbol{A}$

(9) $\frac{\mathrm{d}}{\mathrm{d}t}\ln(\boldsymbol{I}+\boldsymbol{A}t)=\boldsymbol{A}(\boldsymbol{I}+\boldsymbol{A}t)^{-1}=(\boldsymbol{I}+\boldsymbol{A}t)^{-1}\boldsymbol{A}$

定理 5.12 (积分性质) 函数矩阵的积分满足以下性质：

(1) $\int(a\boldsymbol{A}(t)+b\boldsymbol{B}(t))\,\mathrm{d}t=a\int\boldsymbol{A}(t)\,\mathrm{d}t+b\int\boldsymbol{B}(t)\,\mathrm{d}t$（$a,b$ 为任意实数）

(2) $\int\boldsymbol{C}\cdot\boldsymbol{A}(t)\,\mathrm{d}t=\boldsymbol{C}\cdot\int\boldsymbol{A}(t)\,\mathrm{d}t$（$\boldsymbol{C}$ 为常数矩阵）

(3) $\int\boldsymbol{A}(t)\boldsymbol{B}'(t)\,\mathrm{d}t=\boldsymbol{A}(t)\boldsymbol{B}(t)-\int\boldsymbol{A}'(t)\boldsymbol{B}(t)\,\mathrm{d}t$（$\boldsymbol{B}'(t)$ 表示 $\boldsymbol{B}(t)$ 的导数）

5.7 矩阵函数的应用

定理 5.13 一阶线性常系数齐次微分方程组

$$\begin{cases}\boldsymbol{x}'(t) & = & \boldsymbol{A}\boldsymbol{x}(t)\\ \boldsymbol{x}(t_0) & = & \boldsymbol{C}\end{cases}$$

的解为

$$\boldsymbol{x}(t)=\exp(\boldsymbol{A}(t-t_0))\boldsymbol{C}$$

定理 5.14 一阶线性常系数非齐次线性方程组

$$\begin{cases}\boldsymbol{x}'(t) & = & \boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{f}(t)\\ \boldsymbol{x}(t_0) & = & \boldsymbol{C}\end{cases}$$

的解为

$$\boldsymbol{x}(t)=\exp(\boldsymbol{A}(t-t_0))\boldsymbol{C}+\int_{t_0}^{t}\exp(\boldsymbol{A}(t-\tau))\boldsymbol{f}(\tau)\,\mathrm{d}\tau$$

6 矩阵的 Kronecker 积与 Hadamard 积

6.1 Kronecker 积与 Hadamard 积的定义

定义 6.1 设 $\boldsymbol{A}=(a_{ij})\in\mathbb{C}^{m\times n}$，$\boldsymbol{B}=(b_{ij})\in\mathbb{C}^{s\times t}$，则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的 *Kronecker* 积 $\boldsymbol{A}\otimes\boldsymbol{B}$ 定义为

$$\boldsymbol{A}\otimes\boldsymbol{B}=\begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots & a_{2n}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \cdots & a_{mn}\boldsymbol{B} \end{bmatrix}\in\mathbb{C}^{ms\times nt}$$

如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 是同阶矩阵，即 $\boldsymbol{A}=(a_{ij})\in\mathbb{C}^{m\times n}$，$\boldsymbol{B}=(b_{ij})\in\mathbb{C}^{m\times n}$，则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的 *Hadamard* 积 $\boldsymbol{A}\circ\boldsymbol{B}$ 定义为

$$\boldsymbol{A}\circ\boldsymbol{B}=(a_{ij}b_{ij})\in\mathbb{C}^{m\times n}$$

定理 6.1 (K-积的基本性质) 设 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 是矩阵，k 是数，则下列性质成立

(1) $(k\boldsymbol{A})\otimes\boldsymbol{B}=\boldsymbol{A}\otimes(k\boldsymbol{B})=k(\boldsymbol{A}\otimes\boldsymbol{B})$

(2) $\boldsymbol{A}\otimes(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{A}\otimes\boldsymbol{B}+\boldsymbol{A}\otimes\boldsymbol{C}$

(3) $(\boldsymbol{B}+\boldsymbol{C})\otimes\boldsymbol{A}=\boldsymbol{B}\otimes\boldsymbol{A}+\boldsymbol{C}\otimes\boldsymbol{A}$

(4) $(\boldsymbol{A}\otimes\boldsymbol{B})\otimes\boldsymbol{C}=\boldsymbol{A}\otimes(\boldsymbol{B}\otimes\boldsymbol{C})$

(5) $(\boldsymbol{A}\otimes\boldsymbol{B})^{\mathcal{H}}=\boldsymbol{A}^{\mathcal{H}}\otimes\boldsymbol{B}^{\mathcal{H}}$

(6) $\operatorname{rank}(\boldsymbol{A}\otimes\boldsymbol{B})=\operatorname{rank}(\boldsymbol{A})\operatorname{rank}(\boldsymbol{B})$

定理 6.2 (H-积的基本性质) 设 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$ 是同阶矩阵，k 是数，则下列性质成立

(1) $(k\boldsymbol{A})\circ\boldsymbol{B}=\boldsymbol{A}\circ(k\boldsymbol{B})=k(\boldsymbol{A}\circ\boldsymbol{B})$

(2) $\boldsymbol{A}\circ(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{A}\circ\boldsymbol{B}+\boldsymbol{A}\circ\boldsymbol{C}$

(3) $(\boldsymbol{B}+\boldsymbol{C})\circ\boldsymbol{A}=\boldsymbol{B}\circ\boldsymbol{A}+\boldsymbol{C}\circ\boldsymbol{A}$

(4) $(\boldsymbol{A}\circ\boldsymbol{B})\circ\boldsymbol{C}=\boldsymbol{A}\circ(\boldsymbol{B}\circ\boldsymbol{C})$

(5) $(\boldsymbol{A}\circ\boldsymbol{B})^{\mathcal{H}}=\boldsymbol{A}^{\mathcal{H}}\circ\boldsymbol{B}^{\mathcal{H}}$

(6) $\boldsymbol{A}\circ\boldsymbol{B}=\boldsymbol{B}\circ\boldsymbol{A}$

定理 6.3 设矩阵 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{C}$，$\boldsymbol{D}$ 是使下列运算有意义的矩阵，则有

$$(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D})=(\boldsymbol{A}\boldsymbol{C})\otimes(\boldsymbol{B}\boldsymbol{D})$$

若 $\boldsymbol{A}\in F^{m\times m}$，$\boldsymbol{B}\in F^{n\times n}$，则

$$\boldsymbol{A}\otimes\boldsymbol{B}=(\boldsymbol{I}_m\otimes\boldsymbol{B})(\boldsymbol{A}\otimes\boldsymbol{I}_n)$$

6.2 Kronecker 积与 Hadamard 积的性质

定理 6.4 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是使下列运算有意义的矩阵，对 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的 *Kronecker* 积矩阵 $\boldsymbol{A}\otimes\boldsymbol{B}$，下列性质成立。

(1) 当 $\boldsymbol{A}$，$\boldsymbol{B}$ 分别可逆时，$\boldsymbol{A}\otimes\boldsymbol{B}$ 和 $\boldsymbol{B}\otimes\boldsymbol{A}$ 都为可逆矩阵，而且有

$$(\boldsymbol{A}\otimes\boldsymbol{B})^{-1}=\boldsymbol{A}^{-1}\otimes\boldsymbol{B}^{-1}$$

$$(\boldsymbol{B}\otimes\boldsymbol{A})^{-1}=\boldsymbol{B}^{-1}\otimes\boldsymbol{A}^{-1}$$

(2) 当方阵 $\boldsymbol{A}\in F^{m\times m}$，$\boldsymbol{B}\in F^{n\times n}$ 时，方阵 $\boldsymbol{A}\otimes\boldsymbol{B}$ 的行列式

$$|\boldsymbol{A}\otimes\boldsymbol{B}|=|\boldsymbol{B}\otimes\boldsymbol{A}|=|\boldsymbol{A}|^n\,|\boldsymbol{B}|^m$$

(3) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 *Hermite* 矩阵，则 $\boldsymbol{A}\otimes\boldsymbol{B}$ 和 $\boldsymbol{B}\otimes\boldsymbol{A}$ 都是 *Hermite* 矩阵；

(4) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是酉矩阵，则 $\boldsymbol{A}\otimes\boldsymbol{B}$ 和 $\boldsymbol{B}\otimes\boldsymbol{A}$ 都是酉矩阵。

定理 6.5 设 $\boldsymbol{A}$，$\boldsymbol{B}$ 使下列运算有意义，则有

(1) 设 $\boldsymbol{A}$，$\boldsymbol{B}$ 为同阶矩阵，且 $\boldsymbol{A}$ 等价于 $\boldsymbol{B}$，则对任意单位矩阵 $\boldsymbol{I}$，$(\boldsymbol{A}\otimes\boldsymbol{I})$ 等价于 $(\boldsymbol{B}\otimes\boldsymbol{I})$；

(2) 设方阵 $\boldsymbol{A}\in F^{m\times m}$，$\boldsymbol{B}\in F^{n\times n}$，如果 $\boldsymbol{A}$ 相似于 $\boldsymbol{J}_A$，$\boldsymbol{B}$ 相似于 $\boldsymbol{J}_B$，则 $\boldsymbol{A}\otimes\boldsymbol{B}$ 相似于 $\boldsymbol{J}_A\otimes\boldsymbol{J}_B$。

定理 6.6 设方阵 $\boldsymbol{A}\in F^{m\times m}$，$\boldsymbol{A}$ 的特征值是 λ_i，相应的特征向量是 $\boldsymbol{x}_i$，$i=1,2,\cdots,m$。方阵 $\boldsymbol{B}\in F^{n\times n}$，$B$ 的特征值是 μ_i，相应的特征向量是 $\boldsymbol{y}_i$，$i=1,2,\cdots,n$，则

(1) $\boldsymbol{A}\otimes\boldsymbol{B}$ 的特征值是 $\lambda_i\mu_j$，对应的特征向量是 $\boldsymbol{x}_i\otimes\boldsymbol{y}_j$，$i=1,2,\cdots,m$；$j=1,2,\cdots,n$。

(2) $\boldsymbol{A}\otimes\boldsymbol{I}_n+\boldsymbol{I}_m\otimes\boldsymbol{B}$ 的特征值是 $\lambda_i+\mu_j$，对应的特征向量是 $\boldsymbol{x}_i\otimes\boldsymbol{y}_j$，$i=1,2,\cdots,m$；$j=1,2,\cdots,n$。其中 $(\boldsymbol{A}\otimes\boldsymbol{I}_n+\boldsymbol{I}_m\otimes\boldsymbol{B})$ 称为方阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的 *Kronecker* 和，记为

$$\boldsymbol{A}\oplus\boldsymbol{B}=\boldsymbol{A}\otimes\boldsymbol{I}_n+\boldsymbol{I}_m\otimes\boldsymbol{B}$$

定理 6.7 设 $f(z)$ 是解析函数，$\boldsymbol{A}\in F^{n\times n}$，且 $f(\boldsymbol{A})$ 存在，则

$$f(\boldsymbol{I}_m\otimes\boldsymbol{A})=\boldsymbol{I}_m\otimes f(\boldsymbol{A})$$

$$f(\boldsymbol{A}\otimes\boldsymbol{I}_m)=f(\boldsymbol{A})\otimes\boldsymbol{I}_m$$

特例 $f(z)=\exp(z)$：

$$\exp(\boldsymbol{I}_m\otimes\boldsymbol{A})=\boldsymbol{I}_m\otimes\exp(\boldsymbol{A})$$

$$\exp(\boldsymbol{A}\otimes\boldsymbol{I}_m)=\exp(\boldsymbol{A})\otimes\boldsymbol{I}_m$$

6.3 矩阵的向量化算子和 Kronecker 积

定义 6.2 设 $\boldsymbol{A} \in F^{m\times n}$ 是 $m \times n$ 阶矩阵，$\boldsymbol{A} = (\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n)$，其中 $\boldsymbol{A}_i \in F^m$ 是 $\boldsymbol{A}$ 的第 i 列，则 $\boldsymbol{A}$ 的向量算子 $\mathrm{Vec}\,(\boldsymbol{A})$，定义为：

$$\mathrm{Vec}\,(\boldsymbol{A}) = \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \\ \vdots \\ \boldsymbol{A}_n \end{bmatrix} \in F^{mn}$$

定理 6.8 设矩阵 $\boldsymbol{A} \in F^{m\times k}$，$\boldsymbol{B} \in F^{k\times s}$，$\boldsymbol{C} \in F^{s\times n}$，则

$$\mathrm{Vec}\,(\boldsymbol{ABC}) = \left(\boldsymbol{C}^{\mathrm{T}} \otimes \boldsymbol{A}\right) \mathrm{Vec}\,(\boldsymbol{B})$$

推论：设 $\boldsymbol{A} \in F^{m\times k}$，$\boldsymbol{X} \in F^{k\times s}$，$\boldsymbol{C} \in F^{s\times n}$，则有

(1) $\mathrm{Vec}\,(\boldsymbol{AX}) = (\boldsymbol{I}_n \otimes \boldsymbol{A})\,\mathrm{Vec}\,(\boldsymbol{X})$

(2) $\mathrm{Vec}\,(\boldsymbol{XC}) = \left(\boldsymbol{C}^{\mathrm{T}} \otimes \boldsymbol{I}_k\right) \mathrm{Vec}\,(\boldsymbol{X})$

定理 6.9 (用向量化算子求解矩阵方程) 求解矩阵方程 $\boldsymbol{AX} + \boldsymbol{XB} = \boldsymbol{D}$, 其中 $\boldsymbol{A} \in F^{m\times m}$，$\boldsymbol{B} \in F^{n\times n}$，$\boldsymbol{D} \in F^{m\times n}$

(1) 用向量化算子 Vec 作用在等式两边，有 $\left(\boldsymbol{I}_n \otimes \boldsymbol{A} + \boldsymbol{B}^{\mathrm{T}} \otimes \boldsymbol{I}_m\right) \mathrm{Vec}\,(\boldsymbol{X}) = \mathrm{Vec}\,(\boldsymbol{D})$

(2) 令 $\boldsymbol{G} = \left(\boldsymbol{I}_n \otimes \boldsymbol{A} + \boldsymbol{B}^{\mathrm{T}} \otimes \boldsymbol{I}_m\right)$，求解线性方程 $\boldsymbol{G}\,\mathrm{Vec}\,(\boldsymbol{X}) = \mathrm{Vec}\,(\boldsymbol{D})$

定理 6.10 (用向量化算子求解矩阵方程) 求解矩阵方程 $\boldsymbol{AX} - \boldsymbol{XA} = k\boldsymbol{X}$, 其中 $\boldsymbol{A}$，$\boldsymbol{X} \in F^{n\times n}$

(1) 用向量化算子 Vec 作用在等式两边，有 $\left(\boldsymbol{I} \otimes \boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}} \otimes \boldsymbol{I}\right) \mathrm{Vec}\,(\boldsymbol{X}) = k\,\mathrm{Vec}\,(\boldsymbol{X})$

(2) 令 $\boldsymbol{H} = \left(\boldsymbol{I} \otimes \boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}} \otimes \boldsymbol{I}\right)$，求解线性方程 $(k\boldsymbol{I} - \boldsymbol{H})\,\mathrm{Vec}\,(\boldsymbol{X}) = \boldsymbol{0}$

定理 6.11 (用向量化算子求解矩阵微分方程) 微分方程组

$$\begin{cases} \boldsymbol{X}'(t) = \boldsymbol{AX}(t) + \boldsymbol{X}(t)\,\boldsymbol{B} \\ \boldsymbol{X}(0) = \boldsymbol{C} \end{cases}$$

其中 $\boldsymbol{A} \in \mathbb{C}^{m\times m}$，$\boldsymbol{B} \in \mathbb{C}^{n\times n}$，$\boldsymbol{X} \in \mathbb{C}^{m\times n}$

(1) 用向量化算子 Vec 作用在方程两边，有 $\mathrm{Vec}\left(\boldsymbol{X}'(t)\right) = \left(\boldsymbol{I}_n \otimes \boldsymbol{A} + \boldsymbol{B}^{\mathrm{T}} \otimes \boldsymbol{I}_m\right) \mathrm{Vec}\,(\boldsymbol{X}(t))$ 和 $\mathrm{Vec}\,(\boldsymbol{X}(0)) = \mathrm{Vec}\,(\boldsymbol{C})$

(2) 令 $\boldsymbol{Y}(t) = \mathrm{Vec}\,(\boldsymbol{X}(t))$，$\boldsymbol{C}_1 = \mathrm{Vec}\,(\boldsymbol{C})$，$\boldsymbol{G} = \left(\boldsymbol{I}_n \otimes \boldsymbol{A} + \boldsymbol{B}^{\mathrm{T}} \otimes \boldsymbol{I}_m\right)$，则原方程组等价于

$$\begin{cases} \boldsymbol{Y}'(t) = \boldsymbol{GY}(t) \\ \boldsymbol{Y}(0) = \boldsymbol{C}_1 \end{cases}$$

(3) 通过求解普通微分方程的方法得到 $\boldsymbol{Y}(t) = \exp(\boldsymbol{G}t)\,\boldsymbol{C}_1$，从而 $\boldsymbol{X}(t) = \exp(\boldsymbol{A}t) \cdot \boldsymbol{C} \cdot \exp(\boldsymbol{B}t)$