

Chapter 5

矩阵分析

Matrix Theory

December 25, 2014

黄正华, 数学与统计学院, 武汉大学

5.1

Contents

1	向量范数及矩阵范数	1
1.1	向量范数	2
1.2	矩阵范数	4
2	矩阵序列与矩阵级数	8
2.1	向量序列的极限	8
2.2	矩阵序列的极限	9
2.3	矩阵级数	10
3	矩阵的微分与积分	13
3.1	函数矩阵及其极限	13
3.2	函数矩阵的微分和积分	14
3.3	纯量函数关于矩阵的导数	15
3.4	矩阵对矩阵的导数	16
4	矩阵函数	17
4.1	矩阵多项式	17
4.2	矩阵函数	18

5.2

分析 Analysis, 一般称为“数学分析”, 是微积分的 Pro 版本.

矩阵分析 讨论矩阵序列的极限、矩阵级数、矩阵的微分与积分等分析概念.

5.3

1 向量范数及矩阵范数

1.1 向量范数

在第二章中我们用内积定义了向量的长度:

$$\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}|\boldsymbol{x})},$$

它是几何向量长度概念的一种推广. 本节采用公理化的方法把向量长度的概念进一步推广, 主要讨论向量范数、矩阵范数及其应用.

Definition 1. 设 V 是数域 F 上的线性空间. 若对任一向量 $\boldsymbol{x} \in V$, 都有一实数 $\|\boldsymbol{x}\|$ 与之对应, 且满足以下条件:

1. 正定性: $\|\boldsymbol{x}\| \geq 0$. 当且仅当 $\boldsymbol{x} = \mathbf{0}$ 时, $\|\boldsymbol{x}\| = 0$;
2. 齐次性: 对任意 $\lambda \in F$, $\|\lambda\boldsymbol{x}\| = |\lambda|\|\boldsymbol{x}\|$;
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 $\|\boldsymbol{x}\|$ 为线性空间 V 中向量 \boldsymbol{x} 的范数(norm). 定义了范数的线性空间 V 又叫做线性赋范空间.

5.4

基本性质

1. 对任意的 $\boldsymbol{x} \in V$, 有

$$\|-\boldsymbol{x}\| = |-1|\|\boldsymbol{x}\| = \|\boldsymbol{x}\|.$$

2. 对任意 $\boldsymbol{x}, \boldsymbol{y} \in V$, 有

$$\left| \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \right| \leq \|\boldsymbol{x} - \boldsymbol{y}\|. \quad (1)$$

事实上, (1) 式等价于

$$-\|\boldsymbol{x} - \boldsymbol{y}\| \leq \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\|.$$

而

$$-\|\boldsymbol{x} - \boldsymbol{y}\| \leq \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \Leftrightarrow \|\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{y}\|, \quad (2)$$

$$\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\| \Leftrightarrow \|\boldsymbol{x}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y}\|. \quad (3)$$


由

5.5

$$\|\boldsymbol{y}\| = \|\boldsymbol{x} + \boldsymbol{y} - \boldsymbol{x}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y} - \boldsymbol{x}\| = \|\boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{y}\|, \quad (4)$$

$$\|\boldsymbol{x}\| = \|\boldsymbol{x} - \boldsymbol{y} + \boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y}\|, \quad (5)$$

得证 (1) 式成立. \square

 注意向量范数是在一般线性空间 V 中给出的, 定义中的向量是广义向量, 就是说对矩阵、多项式、连续函数等都可以定义向量范数.

下文讨论的向量范数, 仅限于线性空间 \mathbb{C}^n 上讨论.

5.6

线性空间 \mathbb{C}^n 的范数

范数的具体形式是多种多样的.

任取 $\boldsymbol{x} \in \mathbb{C}^n$, 设 $\boldsymbol{x} = (\xi_1, \xi_2, \dots, \xi_n)^T$, 常用的范数有

1. 2-范数

$$\|\boldsymbol{x}\|_2 = \left(\sum_{i=1}^n |\xi_i|^2 \right)^{\frac{1}{2}}.$$

2. 1-范数

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^n |\xi_i|.$$

3. ∞ -范数

$$\|\boldsymbol{x}\|_\infty = \max_{1 \leq i \leq n} |\xi_i|.$$

以上三种范数都是以下 p -范数的特例:

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^n |\xi_i|^p \right)^{\frac{1}{p}}, \quad 1 \leq p < +\infty.$$

5.7

- p -范数定义式中, 若取 $0 < p < 1$, 则不满足三角不等式. 例如 $p = \frac{1}{2}$, 则

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^n |\xi_i|^p \right)^{\frac{1}{p}} = \left(\sum_{i=1}^n \sqrt{|\xi_i|} \right)^2,$$

对 $\boldsymbol{\alpha} = (1, 0, 0)^T$, $\boldsymbol{\beta} = (0, 1, 0)^T$, 则

$$\|\boldsymbol{\alpha} + \boldsymbol{\beta}\|_{\frac{1}{2}} = 4, \quad \|\boldsymbol{\alpha}\|_{\frac{1}{2}} = 1, \quad \|\boldsymbol{\beta}\|_{\frac{1}{2}} = 1.$$

- 1-范数和 2-范数显然是 p -范数在 $p = 1$ 和 $p = 2$ 的特殊情形.
- 2-范数 $\|\boldsymbol{x}\|_2$ 也叫 Euclid 范数, 就是通常意义下的向量的长度.

5.8

Theorem 2. 对任意向量 $\boldsymbol{x} \in \mathbb{C}^n$,

$$\lim_{p \rightarrow \infty} \|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty. \quad (6)$$

证: 因

$$\left(\max_{1 \leq i \leq n} |\xi_i|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |\xi_i|^p \right)^{\frac{1}{p}} \leq \left(n \cdot \max_{1 \leq i \leq n} |\xi_i|^p \right)^{\frac{1}{p}},$$

故有

$$\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_p \leq n^{\frac{1}{p}} \|\boldsymbol{x}\|_\infty.$$

令 $p \rightarrow \infty$, 注意到 $n^{\frac{1}{p}} \rightarrow 1$, 即得 (6) 式成立. \square

5.9

范数的等价性

按照不同方式规定的范数, 其值一般不相同, 但在各种范数下考虑向量序列的收敛性时, 却表现出明显的一致性, 这就是向量范数的等价性.

Definition 3. 称范数 $\|\mathbf{x}\|_\alpha$ 与 $\|\mathbf{x}\|_\beta$ 等价, 如果存在 $K_2 \geq K_1 > 0$, 使得对一切 $\mathbf{x} \in \mathbb{C}^n$ 都有

$$K_1\|\mathbf{x}\|_\beta \leq \|\mathbf{x}\|_\alpha \leq K_2\|\mathbf{x}\|_\beta.$$

由

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_p \leq n^{\frac{1}{p}}\|\mathbf{x}\|_\infty$$

表明, 任何范数 $\|\mathbf{x}\|_p$ 均与 $\|\mathbf{x}\|_\infty$ 等价, 因而任何两种 p -范数都是彼此等价的.

特别地, 前述三种常用范数 $\|\mathbf{x}\|_1$, $\|\mathbf{x}\|_2$ 和 $\|\mathbf{x}\|_\infty$ 彼此等价.

5.10

1.2 矩阵范数

注意到前述的向量范数是定义在线性空间 V 上的, $m \times n$ 阶矩阵作为线性空间 $\mathbb{C}^{m \times n}$ 中的向量, 自然也可以依照该定义产生范数.

与向量 $\mathbf{x} \in \mathbb{C}^n$ 的几种范数相对应, 对矩阵 $\mathbf{A} = [a_{ij}] \in \mathbb{C}^{m \times n}$ 有:

$$\begin{aligned} \|\mathbf{A}\|_{V_1} &= \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|, \\ \|\mathbf{A}\|_{V_2} &= \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}} = \sqrt{\text{tr}(\mathbf{A}^H \mathbf{A})} \triangleq \|\mathbf{A}\|_F, \quad (\text{Frobenius 范数}) \\ \|\mathbf{A}\|_{V_\infty} &= \max_{i,j} |a_{ij}|, \quad (\text{契比雪夫范数}) \\ \|\mathbf{A}\|_{V_p} &= \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^p \right)^{\frac{1}{p}}, \quad 1 \leq p < +\infty. \end{aligned}$$

可以证明它们都是线性空间 $\mathbb{C}^{m \times n}$ 上的向量范数.

下标 V 即 Vector. 同样地, $\|\mathbf{A}\|_{V_p}$ 在 $p = 1, p = 2, p \rightarrow \infty$ 时分别成为 $\|\mathbf{A}\|_{V_1}, \|\mathbf{A}\|_{V_2}, \|\mathbf{A}\|_{V_\infty}$.

5.11

F-范数

Frobenius 范数

$$\|\mathbf{A}\|_{V_2} = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}} = \sqrt{\text{tr}(\mathbf{A}^H \mathbf{A})} \triangleq \|\mathbf{A}\|_F$$

简称为 F-范数, 是最常见的范数之一. 它是酉空间 $\mathbb{C}^{m \times n}$ 中的内积

$$(\mathbf{A} | \mathbf{B}) = \text{tr}(\mathbf{B}^H \mathbf{A}) = \sum_{i=1}^m \sum_{j=1}^n \overline{b_{ij}} a_{ij}$$

所诱导的范数. 即

$$\|\mathbf{A}\|_F^2 = (\mathbf{A} | \mathbf{A}) = \text{tr}(\mathbf{A}^H \mathbf{A}) = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2.$$

5.12

Theorem 4. 设 $A \in \mathbb{C}^{m \times n}$, 对酉矩阵 $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$, 恒有

$$\|A\|_F = \|UA\|_F = \|AV\|_F = \|UAV\|_F.$$

称之为 F -范数的酉不变性.

证: 由 $\|A\|_F = \sqrt{\text{tr}(A^H A)}$, 得

$$\|UA\|_F = \sqrt{\text{tr}(A^H U^H U A)} = \sqrt{\text{tr}(A^H A)} = \|A\|_F.$$

其他同理. □

5.13

矩阵范数

Definition 5. 在矩阵空间 $\mathbb{C}^{n \times n}$ 上的任一实函数, 记为 $\|\cdot\|$, 如果对所有的 $A, B \in \mathbb{C}^{n \times n}$, $\lambda \in \mathbb{C}$ 都满足

1. $\|A\| \geq 0$. 当且仅当 $A = O$ 时, $\|A\| = 0$;
2. $\|\lambda A\| = |\lambda| \|A\|$;
3. $\|A + B\| \leq \|A\| + \|B\|$;
4. $\|AB\| \leq \|A\| \|B\|$.

则称 $\|\cdot\|$ 为相容的矩阵范数, 或简称为矩阵范数.

若将向量 x 视为矩阵, 根据相容性矩阵范数定义的第 4 条 $\|AB\| \leq \|A\| \|B\|$, 应有

$$\|Ax\| \leq \|A\| \|x\|.$$

它们各自取不同的范数仍能使这个不等式成立吗? 两者怎样取才能协调? 这就是矩阵范数与向量范数相容的概念.

Definition 6. 如果矩阵范数 $\|A\|_m$ 和向量范数 $\|x\|_v$ 对一切矩阵 A 和向量 x 都满足

$$\|Ax\|_v \leq \|A\|_m \|x\|_v,$$

则称矩阵范数 $\|A\|_m$ 和向量范数 $\|x\|_v$ 是相容的.

Example 7. 设 $x = (\xi_1, \xi_2, \dots, \xi_n)^T \in \mathbb{C}^n$, $A = [a_{ij}] \in \mathbb{C}^{n \times n}$, 试证明 F -范数 $\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2\right)^{\frac{1}{2}}$ 与向量的欧氏范数 $\|x\|_2 = \left(\sum_{i=1}^n |\xi_i|^2\right)^{\frac{1}{2}}$ 是相容的.

对于 $\mathbb{C}^{n \times n}$ 上任意给定的一种矩阵范数, 是否一定存在与之相容的向量范数?

Theorem 8. 设 $\|A\|_m$ 是 $\mathbb{C}^{n \times n}$ 上的一种相容的矩阵范数, 则在 \mathbb{C}^n 上必存在一个与它相容的向量范数.

证: 设 $a \in \mathbb{C}^n$ 是非零向量, 对任意的 $x \in \mathbb{C}^n$, 定义

$$\|x\|_v = \|xa^H\|_m.$$

先证明 $\|x\|_v$ 是 \mathbb{C}^n 中的向量范数.

5.15

5.16

1. 当 $\mathbf{x} \neq \mathbf{0}$ 时, $\mathbf{x}\mathbf{a}^H \neq \mathbf{O}$, 所以

$$\|\mathbf{x}\|_v = \|\mathbf{x}\mathbf{a}^H\|_m > 0.$$

当且仅当 $\mathbf{x} = \mathbf{0}$ 时, 才有 $\mathbf{x}\mathbf{a}^H = \mathbf{O}$, 即 $\|\mathbf{x}\|_v = \|\mathbf{x}\mathbf{a}^H\|_m = 0$.

2. 对任意 $\lambda \in \mathbb{C}$, 有

$$\|\lambda\mathbf{x}\|_v = \|\lambda\mathbf{x}\mathbf{a}^H\|_m = |\lambda|\|\mathbf{x}\mathbf{a}^H\|_m = |\lambda|\|\mathbf{x}\|_v.$$

5.17

3 对任何 $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, 都有

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_v &= \|(\mathbf{x} + \mathbf{y})\mathbf{a}^H\|_m = \|\mathbf{x}\mathbf{a}^H + \mathbf{y}\mathbf{a}^H\|_m \\ &\leq \|\mathbf{x}\mathbf{a}^H\|_m + \|\mathbf{y}\mathbf{a}^H\|_m \\ &= \|\mathbf{x}\|_v + \|\mathbf{y}\|_v. \end{aligned}$$

所以 $\|\mathbf{x}\|_v$ 是 \mathbb{C}^n 中的向量范数.

又

$$\|\mathbf{A}\mathbf{x}\|_v = \|\mathbf{A}\mathbf{x}\mathbf{a}^H\|_m \leq \|\mathbf{A}\|_m \|\mathbf{x}\mathbf{a}^H\|_m = \|\mathbf{A}\|_m \|\mathbf{x}\|_v,$$

得证矩阵范数 $\|\mathbf{A}\|_m$ 和向量范数 $\|\mathbf{x}\|_v = \|\mathbf{x}\mathbf{a}^H\|_m$ 是相容的. \square

5.18

Theorem 9. 设 $\mathbf{A} \in \mathbb{C}^{m \times n}$, $\mathbf{x} \in \mathbb{C}^n$, 且在 \mathbb{C}^n 中已规定了向量的某种范数 $\|\mathbf{x}\|$, 则与向量范数 $\|\mathbf{x}\|$ 相容的矩阵范数可以取为向量 $\mathbf{A}\mathbf{x}$ 的范数的最大值, 即

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|.$$

其中向量范数 $\|\mathbf{x}\|$ 与 $\|\mathbf{A}\mathbf{x}\|$ 可以相同, 也可以不同.

5.19

Example 10. 设 $\mathbf{A} = [a_{ij}] \in \mathbb{C}^{n \times n}$, 当向量范数分别取 $\|\cdot\|_1, \|\cdot\|_\infty, \|\cdot\|_2$ 时, 其对应的算子范数也分别记为 $\|\mathbf{A}\|_1, \|\mathbf{A}\|_\infty, \|\mathbf{A}\|_2$, 则有

$$\|\mathbf{A}\|_1 = \max_{\|\mathbf{x}\|_1=1} \|\mathbf{A}\mathbf{x}\|_1 = \max_j \sum_{i=1}^n |a_{ij}|, \quad (\text{列范数})$$

$$\|\mathbf{A}\|_\infty = \max_{\|\mathbf{x}\|_\infty=1} \|\mathbf{A}\mathbf{x}\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|, \quad (\text{行范数})$$

$$\|\mathbf{A}\|_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2 = [\rho(\mathbf{A}^H\mathbf{A})]^{1/2} = \sigma_1. \quad (\text{谱范数})$$

其中 σ_1 为 \mathbf{A} 的最大奇异值.

5.20

Theorem 11. 矩阵 $\mathbf{A} \in \mathbb{C}^{m \times n}$ 的任意一种范数, 都是 \mathbf{A} 的元素的连续函数.

Theorem 12. 对于 $\mathbb{C}^{m \times n}$ 上的任意两种范数 $\|\mathbf{A}\|_a$ 及 $\|\mathbf{A}\|_b$, 必存在 $K_1 > 0$, $K_2 > 0$, 使得对任意的 $\mathbf{A} \in \mathbb{C}^{m \times n}$, 都有

$$K_1\|\mathbf{A}\|_a \leq \|\mathbf{A}\|_b \leq K_2\|\mathbf{A}\|_a.$$

5.21

Example 13. 设 $\mathbf{A} = \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix}$, 试计算 $\|\mathbf{A}\|_1, \|\mathbf{A}\|_\infty, \|\mathbf{A}\|_2, \|\mathbf{A}\|_F$.

解: (1) $\|\mathbf{A}\|_1 = 6$.

(2) $\|\mathbf{A}\|_\infty = 7$.

(3) 由 $\mathbf{A}^H\mathbf{A} = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} 10 & -14 \\ -14 & 20 \end{bmatrix}$, 得

$$\begin{vmatrix} \lambda - 10 & 14 \\ 14 & \lambda - 20 \end{vmatrix} = \lambda^2 - 30\lambda + 4,$$

则 $\mathbf{A}^H\mathbf{A}$ 的特征值为 $\lambda_1 = 15 + \sqrt{221}, \lambda_2 = 15 - \sqrt{221}$. 故

$$\|\mathbf{A}\|_2 = [\rho(\mathbf{A}^H\mathbf{A})]^{\frac{1}{2}} = \sqrt{15 + \sqrt{221}} \approx 5.46.$$

(4) $\|\mathbf{A}\|_F = \sqrt{1^2 + (-2)^2 + (-3)^2 + 4^2} = \sqrt{30} \approx 5.477$.

或者 $\|\mathbf{A}\|_F = \sqrt{\text{tr}(\mathbf{A}^H\mathbf{A})} = \sqrt{\lambda_1 + \lambda_2} = \sqrt{30}$. □

5.22

Exercise 14 (P.176 习题一 11). 证明:

$$\frac{1}{\sqrt{n}}\|\mathbf{A}\|_F \leq \|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F.$$

证: 矩阵 $\mathbf{A}^H\mathbf{A}$ 的特征值非负, 记为 $\lambda_1, \lambda_2, \dots, \lambda_n$. 因 $\|\mathbf{A}\|_2 = [\rho(\mathbf{A}^H\mathbf{A})]^{\frac{1}{2}}$, 故

$$\|\mathbf{A}\|_2^2 = \rho(\mathbf{A}^H\mathbf{A}) = \max\{\lambda_1, \lambda_2, \dots, \lambda_n\} \triangleq \lambda_{\max}.$$

从而

$$\|\mathbf{A}\|_2^2 = \lambda_{\max} \leq \lambda_1 + \lambda_2 + \dots + \lambda_n = \text{tr}(\mathbf{A}^H\mathbf{A}) = \|\mathbf{A}\|_F^2.$$

另一方面,

$$\|\mathbf{A}\|_2^2 = \lambda_{\max} \geq \frac{1}{n}(\lambda_1 + \lambda_2 + \dots + \lambda_n) = \frac{1}{n}\|\mathbf{A}\|_F^2.$$

得证 $\frac{1}{\sqrt{n}}\|\mathbf{A}\|_F \leq \|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F$. □

5.23

Exercise 15 (P.176 习题二 3.). 设 $\|\cdot\|$ 为 $\mathbb{C}^{n \times n}$ 上矩阵的算子范数, $\rho(\mathbf{A})$ 为 \mathbf{A} 的谱半径, 则对任意的矩阵 $\mathbf{A} \in \mathbb{C}^{n \times n}$, 均有

$$\rho(\mathbf{A}) \leq \|\mathbf{A}\|.$$

证: 设 $\|\cdot\|_v$ 是 \mathbb{C}^n 上与矩阵范数 $\|\cdot\|$ 相容的向量范数. 又设 λ 为 \mathbf{A} 的特征值, \mathbf{x} 为 \mathbf{A} 的属于特征值 λ 的特征向量, 则 $\mathbf{x} \neq \mathbf{0}$, 从而 $\|\mathbf{x}\|_v > 0$.

由 $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ 有

$$\|\mathbf{A}\mathbf{x}\|_v = |\lambda| \|\mathbf{x}\|_v.$$

而 $\|\mathbf{A}\mathbf{x}\|_v \leq \|\mathbf{A}\| \|\mathbf{x}\|_v$, 于是

$$|\lambda| \|\mathbf{x}\|_v \leq \|\mathbf{A}\| \|\mathbf{x}\|_v.$$

又 $\|\mathbf{x}\|_v > 0$, 故

$$|\lambda| \leq \|\mathbf{A}\|.$$

由 λ 的任意性, 得证 $\rho(\mathbf{A}) \leq \|\mathbf{A}\|$. □

5.24

2 矩阵序列与矩阵级数

2.1 向量序列的极限

矩阵分析理论的建立和数学分析一样,也是以极限理论为基础而形成的,下面讨论向量序列的极限运算.

Definition 16. 设 $\mathbf{x}^{(k)}, \mathbf{x}_0 \in \mathbb{C}^n, k = 1, 2, \dots$, 若


$$\|\mathbf{x}^{(k)} - \mathbf{x}_0\| \rightarrow 0 \quad (k \rightarrow \infty),$$

则称向量序列 $\{\mathbf{x}^{(k)}\}$ 收敛于向量 \mathbf{x}_0 , 或称向量 \mathbf{x}_0 是向量序列 $\{\mathbf{x}^{(k)}\}$ 在 $k \rightarrow \infty$ 时的极限. 记为

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}_0,$$

或

$$\mathbf{x}^{(k)} \rightarrow \mathbf{x}_0 \quad (k \rightarrow \infty).$$

 由于有限维空间的向量范数是等价的,因此上述定义中的范数是任意的向量范数.

5.25

Example 17. 设 $\mathbf{x}_0 = (1, 1, \dots, 1)^T$, 且

$$\mathbf{x}^{(k)} = \left(1 + \frac{1}{2^k}, 1 + \frac{1}{3^k}, \dots, 1 + \frac{1}{(n+1)^k}\right)^T, \quad k = 1, 2, \dots$$

试证: $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}_0$.

证: 由于

$$\|\mathbf{x}^{(k)} - \mathbf{x}_0\|_\infty = \max \left\{ \frac{1}{2^k}, \frac{1}{3^k}, \dots, \frac{1}{(n+1)^k} \right\} = \frac{1}{2^k} \rightarrow 0 \quad (k \rightarrow \infty),$$

所以 $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}_0$. □

5.26

Theorem 18. 设 $\mathbf{x}^{(k)} = (\xi_1^{(k)}, \xi_2^{(k)}, \dots, \xi_n^{(k)})^T, \mathbf{x}_0 = (\xi_1, \xi_2, \dots, \xi_n)^T \in \mathbb{C}^n$, 则

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}_0 \Leftrightarrow \lim_{k \rightarrow \infty} \xi_i^{(k)} = \xi_i \quad (i = 1, 2, \dots, n). \quad (7)$$

证: 必要性. 设 $\|\mathbf{x}\| = \|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |\xi_i|$, 则

$$\|\mathbf{x}^{(k)} - \mathbf{x}_0\|_\infty \rightarrow 0 \quad (k \rightarrow \infty).$$

由于

$$|\xi_i^{(k)} - \xi_i| \leq \max_{1 \leq i \leq n} |\xi_i^{(k)} - \xi_i| = \|\mathbf{x}^{(k)} - \mathbf{x}_0\|_\infty,$$

所以

$$|\xi_i^{(k)} - \xi_i| \rightarrow 0 \quad (k \rightarrow \infty).$$

即

$$\lim_{k \rightarrow \infty} \xi_i^{(k)} = \xi_i, \quad i = 1, 2, \dots, n.$$

5.27

充分性. 如果 $\lim_{k \rightarrow \infty} \xi_i^{(k)} = \xi_i, (i = 1, 2, \dots, n)$, 则

$$\lim_{k \rightarrow \infty} \max_{1 \leq i \leq n} |\xi_i^{(k)} - \xi_i| = 0,$$

即

$$\|\mathbf{x}^{(k)} - \mathbf{x}_0\|_\infty \rightarrow 0 \quad (k \rightarrow \infty).$$

再有范数的等价性知, 上述结论对任意向量范数均成立. □

5.28

Example 19. 试考察下列向量序列的收敛性:

1. $\mathbf{x}^{(k)} = \left(\frac{1}{2^k}, \frac{\sin k}{k}\right)^T, k = 1, 2, \dots$

2. $\mathbf{y}^{(k)} = \left(\sum_{i=1}^k \frac{1}{2^i}, \sum_{i=1}^k \frac{1}{i}\right)^T, k = 1, 2, \dots$

解: (1) 当 $k \rightarrow \infty$ 时, $\frac{1}{2^k} \rightarrow 0, \frac{\sin k}{k} \rightarrow 0$, 所以

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \left(\lim_{k \rightarrow \infty} \frac{1}{2^k}, \lim_{k \rightarrow \infty} \frac{\sin k}{k}\right)^T = (0, 0)^T = \mathbf{0}.$$

故向量序列 $\mathbf{x}^{(k)}$ 收敛, 且收敛于 $\mathbf{0}$.

(2) 当 $k \rightarrow \infty$ 时,

$$\sum_{i=1}^k \frac{1}{i} = 1 + \frac{1}{2} + \dots + \frac{1}{k} + \dots$$

为调和级数, 是发散的. 故向量序列 $\mathbf{y}^{(k)}$ 发散. □

5.29

2.2 矩阵序列的极限

任给 $\mathbb{C}^{m \times n}$ 中的矩阵序列 $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k, \dots$, 记为 $\{\mathbf{A}_k\}$, 其中

$$\mathbf{A}_k = \begin{bmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \cdots & a_{1n}^{(k)} \\ a_{21}^{(k)} & a_{22}^{(k)} & \cdots & a_{2n}^{(k)} \\ \vdots & \vdots & & \vdots \\ a_{m1}^{(k)} & a_{m2}^{(k)} & \cdots & a_{mn}^{(k)} \end{bmatrix}.$$

显然, $\{\mathbf{A}_k\}$ 中各矩阵的对应元素构成了 $m \times n$ 个标量序列 $\{a_{ij}^{(k)}\}$.

5.30

Definition 20. 给定 $\mathbb{C}^{m \times n}$ 中的矩阵序列 $\{\mathbf{A}_k\}$, 如果 $k \rightarrow \infty$ 时, $m \times n$ 个序列 $\{a_{ij}^{(k)}\}$ 都收敛, 即

$$\lim_{k \rightarrow \infty} a_{ij}^{(k)} = a_{ij}, \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n,$$

则称矩阵序列 $\{\mathbf{A}_k\}$ 收敛于矩阵

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

并称 \mathbf{A} 是序列 $\{\mathbf{A}_k\}$ 在 $k \rightarrow \infty$ 时的极限, 记作

$$\lim_{k \rightarrow \infty} \mathbf{A}_k = \mathbf{A}.$$

Example 21. 已知

$$\mathbf{A}_k = \begin{bmatrix} \frac{1}{k} & \frac{2k^2 - 1}{3k^2 + 4} \\ (1 + \frac{1}{k})^k & \cos \frac{1}{k^3} \end{bmatrix},$$

求 $\{\mathbf{A}_k\}$ 在 $k \rightarrow \infty$ 时的极限.

解:

$$\lim_{k \rightarrow \infty} \mathbf{A}_k = \begin{bmatrix} \lim_{k \rightarrow \infty} \frac{1}{k} & \lim_{k \rightarrow \infty} \frac{2k^2 - 1}{3k^2 + 4} \\ \lim_{k \rightarrow \infty} (1 + \frac{1}{k})^k & \lim_{k \rightarrow \infty} \cos \frac{1}{k^3} \end{bmatrix} = \begin{bmatrix} 0 & \frac{2}{3} \\ e & 1 \end{bmatrix}. \quad \square$$

5.32

Theorem 22. 设有矩阵序列 $\mathbf{A}_k \in \mathbb{C}^{m \times n}$ 和矩阵 $\mathbf{A} \in \mathbb{C}^{m \times n}$, 则 $\lim_{k \rightarrow \infty} \mathbf{A}_k = \mathbf{A}$ 的充要条件是

$$\lim_{k \rightarrow \infty} \|\mathbf{A}_k - \mathbf{A}\| = 0,$$

其中 $\|\cdot\|$ 是 $\mathbb{C}^{m \times n}$ 上的任一矩阵范数.

5.33

Theorem 23. 收敛的矩阵序列 $\{\mathbf{A}_k\}$ 是有界的, 即存在正数 $M > 0$, 使得对一切 k 都有

$$\|\mathbf{A}_k\| \leq M.$$

5.34

Theorem 24. 若 $\mathbf{A}_k, \mathbf{B}_k, \mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times n}$, $\alpha_k, \beta_k, \alpha, \beta \in \mathbb{C}$, 且有 $\lim_{k \rightarrow \infty} \mathbf{A}_k = \mathbf{A}$,

$\lim_{k \rightarrow \infty} \mathbf{B}_k = \mathbf{B}$, $\lim_{k \rightarrow \infty} \alpha_k = \alpha$, $\lim_{k \rightarrow \infty} \beta_k = \beta$, 则

1. $\lim_{k \rightarrow \infty} (\alpha_k \mathbf{A}_k + \beta_k \mathbf{B}_k) = \alpha \mathbf{A} + \beta \mathbf{B}$;
2. $\lim_{k \rightarrow \infty} (\mathbf{A}_k \mathbf{B}_k) = \mathbf{A} \mathbf{B}$;
3. $\lim_{k \rightarrow \infty} \mathbf{A}_k^{-1} = \mathbf{A}^{-1}$, 若 $\mathbf{A}_k^{-1}, \mathbf{A}^{-1}$ 存在.

5.35

2.3 矩阵级数

Definition 25. 设有矩阵序列 $\mathbf{A}_k \in \mathbb{C}^{m \times n}$, 称

$$\sum_{k=1}^{\infty} \mathbf{A}_k = \mathbf{A}_1 + \mathbf{A}_2 + \cdots + \mathbf{A}_k + \cdots$$

为矩阵级数, 其中 \mathbf{A}_k 为矩阵级数的一般项.

5.36

Definition 26. 称 $\mathbf{S}_m = \sum_{k=1}^m \mathbf{A}_k$ 为矩阵级数 $\sum_{k=1}^{\infty} \mathbf{A}_k$ 的部分和. 若矩阵序列 $\{\mathbf{S}_m\}$ 收敛, 且

$$\lim_{m \rightarrow \infty} \mathbf{S}_m = \mathbf{S},$$

则称矩阵级数 $\sum_{k=1}^{\infty} \mathbf{A}_k$ 是收敛的, 级数的和为 \mathbf{S} , 记作

$$\mathbf{S} = \sum_{k=1}^{\infty} \mathbf{A}_k.$$

不收敛的级数称为是发散的.

显然, $\sum_{k=1}^{\infty} \mathbf{A}_k$ 收敛的充分必要条件是对应的 $m \times n$ 个数值级数

$$\sum_{k=1}^{\infty} a_{ij}^{(k)} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n)$$

都收敛.

5.37

Example 27. 已知 $\mathbf{A}_k = \begin{bmatrix} \frac{1}{2^k} & \frac{\pi}{4^k} \\ 0 & \frac{1}{(k+1)(k+2)} \end{bmatrix}$, $k = 0, 1, 2, \dots$. 问矩阵级数

$\sum_{k=0}^{\infty} \mathbf{A}_k$ 是否收敛? 若收敛, 求其和.

解: 因为

$$\sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{1 - \frac{1}{2}} = 2, \quad \sum_{k=0}^{\infty} \frac{\pi}{4^k} = \frac{\pi}{1 - \frac{1}{4}} = \frac{4\pi}{3},$$

又

$$\begin{aligned} \sum_{k=0}^m \frac{1}{(k+1)(k+2)} &= 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \dots + \frac{1}{m+1} - \frac{1}{m+2} \\ &= 1 - \frac{1}{m+2} \rightarrow 1 \quad (m \rightarrow \infty), \end{aligned}$$

即

$$\sum_{k=0}^{\infty} \frac{1}{(k+1)(k+2)} = 1.$$

5.38

所以矩阵级数 $\sum_{k=0}^{\infty} \mathbf{A}_k$ 收敛, 且其和为 $\begin{bmatrix} 2 & \frac{4\pi}{3} \\ 0 & 1 \end{bmatrix}$. □

5.39

如果在矩阵空间 $\mathbb{C}^{m \times n}$ 中, $\sum_{k=1}^{\infty} \mathbf{A}_k = \mathbf{S}$, $\sum_{k=1}^{\infty} \mathbf{B}_k = \mathbf{T}$, 则有

1. $\sum_{k=1}^{\infty} (\mathbf{A}_k + \mathbf{B}_k) = \mathbf{S} + \mathbf{T}$;
2. $\sum_{k=1}^{\infty} \alpha \mathbf{A}_k = \alpha \mathbf{S}$, $\forall \alpha \in \mathbb{C}$;
3. $\sum_{k=1}^{\infty} \mathbf{A}_k \mathbf{x} = \mathbf{S} \mathbf{x}$, $\forall \mathbf{x} \in \mathbb{C}^n$.

5.40

Definition 28. 如果对 $\mathbb{C}^{m \times n}$ 上的某种范数 $\|\cdot\|$, 标量级数 $\sum_{k=1}^{\infty} \|\mathbf{A}_k\|$ 收敛, 则称级数 $\sum_{k=1}^{\infty} \mathbf{A}_k$ 为绝对收敛.

Theorem 29. 若级数 $\sum_{k=1}^{\infty} \mathbf{A}_k$ 绝对收敛, 则它也一定是收敛的.

5.41

Example 30. 设 $\mathbf{A} = \begin{bmatrix} 0.2 & 0.5 & 0.1 \\ 0.1 & 0.5 & 0.3 \\ 0.2 & 0.4 & 0.2 \end{bmatrix}$, 试证矩阵级数 $\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots + \mathbf{A}^k + \dots$

绝对收敛.

证: 取矩阵的范数为 $\|\cdot\|_\infty$. 因

$$\|\mathbf{A}\| = \|\mathbf{A}\|_\infty = \max_i \sum_{j=1}^n |a_{ij}| = 0.9 < 1,$$

又 $\|\mathbf{I}\| = \|\mathbf{I}\|_\infty = 1$. 则

$$\|\mathbf{I}\| + \|\mathbf{A}\| + \|\mathbf{A}\|^2 + \cdots + \|\mathbf{A}\|^k + \cdots$$

是公比为 $\|\mathbf{A}\| = 0.9$ 的等比级数, 故收敛.

又 $\|\mathbf{A}^k\| \leq \|\mathbf{A}\|^k$, 由正项级数比较法审敛法知

$$\|\mathbf{I}\| + \|\mathbf{A}\| + \|\mathbf{A}^2\| + \cdots + \|\mathbf{A}^k\| + \cdots$$

收敛, 故原级数绝对收敛. □

注意: 绝对收敛的定义中, 只要求存在某种范数使标量级数 $\sum_{k=1}^{\infty} \|\mathbf{A}^k\|$ 收敛. 比如此例若取范数为 $\|\cdot\|_1$, 由 $\|\mathbf{A}\|_1 = 1.4 > 1$, 就无法判断矩阵级数是否绝对收敛.

Example 31. 设 $\mathbf{A} \in \mathbb{C}^{n \times n}$, 则矩阵级数

$$\mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \cdots + \frac{\mathbf{A}^k}{k!} + \cdots$$

绝对收敛.

证: 因为 $\left\| \frac{\mathbf{A}^k}{k!} \right\| = \frac{\|\mathbf{A}^k\|}{k!} \leq \frac{\|\mathbf{A}\|^k}{k!}$, 且

$$\|\mathbf{I}\| + \|\mathbf{A}\| + \frac{\|\mathbf{A}\|^2}{2!} + \cdots + \frac{\|\mathbf{A}\|^k}{k!} + \cdots = e^{\|\mathbf{A}\|} - 1 + \|\mathbf{I}\|,$$

由正项级数比较法知, 级数 $\sum_{k=1}^{\infty} \left\| \frac{\mathbf{A}^k}{k!} \right\|$ 收敛, 所以级数 $\sum_{k=1}^{\infty} \frac{\mathbf{A}^k}{k!}$ 绝对收敛. □

级数 $\sum_{k=1}^{\infty} \frac{\mathbf{A}^k}{k!}$ 的和记为 $e^{\mathbf{A}}$, 称为 \mathbf{A} 的矩阵指数函数, 即

$$e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \cdots + \frac{\mathbf{A}^k}{k!} + \cdots.$$

同理有

$$\sin \mathbf{A} = \mathbf{A} - \frac{\mathbf{A}^3}{3!} + \cdots + (-1)^{k-1} \frac{\mathbf{A}^{2k-1}}{(2k-1)!} + \cdots, \quad (8)$$

$$\cos \mathbf{A} = \mathbf{I} - \frac{\mathbf{A}^2}{2!} + \cdots + (-1)^k \frac{\mathbf{A}^{2k}}{(2k)!} + \cdots. \quad (9)$$

从而欧拉公式

$$e^{i\mathbf{A}} = \cos \mathbf{A} + i \sin \mathbf{A}$$

对任意的复数或复矩阵都成立. 特别地,

$$e^{i\pi} + 1 = 0.$$

3 矩阵的微分与积分

3.1 函数矩阵及其极限

Definition 32. 以实变量 t 的函数为元素的矩阵

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

称为函数矩阵, 其中 $a_{ij}(t)$ 都是定义在 $[a, b]$ 上的实函数.

函数矩阵 $\mathbf{A}(t)$ 在 $[a, b]$ 上有界、有极限、连续、可微、可积等概念, 可用其 $m \times n$ 个元素 $a_{ij}(t)$ 同时在 $[a, b]$ 上有界、有极限、连续、可微、可积来定义. 例如

$$\begin{aligned} \frac{d}{dt} \mathbf{A}(t) &= \left[\frac{d}{dt} a_{ij}(t) \right]_{m \times n}; \\ \int \mathbf{A}(t) dt &= \left[\int a_{ij}(t) dt \right]_{m \times n}; \\ \int_a^b \mathbf{A}(t) dt &= \left[\int_a^b a_{ij}(t) dt \right]_{m \times n}. \end{aligned}$$

Definition 33. 如果所有的元素 $a_{ij}(t)$ 在 $t \rightarrow t_0$ 时, 极限存在, 记为常数 a_{ij} , 即

$$\lim_{t \rightarrow t_0} a_{ij}(t) = a_{ij},$$

则称矩阵 $\mathbf{A}(t)$ 在 $t \rightarrow t_0$ 时, 极限存在, 且极限值为 \mathbf{A} (常量矩阵), 即

$$\lim_{t \rightarrow t_0} \mathbf{A}(t) = \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Definition 34. 如果所有的元素 $a_{ij}(t)$ 在 $t = t_0$ 连续, 即

$$\lim_{t \rightarrow t_0} a_{ij}(t) = a_{ij}(t_0),$$

则称矩阵 $\mathbf{A}(t)$ 在 $t = t_0$ 连续, 且记为

$$\lim_{t \rightarrow t_0} \mathbf{A}(t) = \mathbf{A}(t_0) = \begin{bmatrix} a_{11}(t_0) & a_{12}(t_0) & \cdots & a_{1n}(t_0) \\ a_{21}(t_0) & a_{22}(t_0) & \cdots & a_{2n}(t_0) \\ \vdots & \vdots & & \vdots \\ a_{m1}(t_0) & a_{m2}(t_0) & \cdots & a_{mn}(t_0) \end{bmatrix}.$$

容易验证下列等式是成立的.

$$\text{设 } \lim_{t \rightarrow t_0} \mathbf{A}(t) = \mathbf{A}, \lim_{t \rightarrow t_0} \mathbf{B}(t) = \mathbf{B}.$$

1. 若 $\mathbf{A}(t), \mathbf{B}(t)$ 都是 $m \times n$ 阶矩阵, 则

$$\lim_{t \rightarrow t_0} [\mathbf{A}(t) + \mathbf{B}(t)] = \mathbf{A} + \mathbf{B} = \lim_{t \rightarrow t_0} \mathbf{A}(t) + \lim_{t \rightarrow t_0} \mathbf{B}(t).$$

2. 设 k 为常数, 则

$$\lim_{t \rightarrow t_0} [k\mathbf{A}(t)] = k\mathbf{A}.$$

3. 若 $\mathbf{A}(t), \mathbf{B}(t)$ 分别是 $m \times n$ 阶及 $n \times r$ 阶矩阵, 则

$$\lim_{t \rightarrow t_0} [\mathbf{A}(t)\mathbf{B}(t)] = \mathbf{A}\mathbf{B} = \lim_{t \rightarrow t_0} \mathbf{A}(t) \lim_{t \rightarrow t_0} \mathbf{B}(t).$$

3.2 函数矩阵的微分和积分

Definition 35. 若函数矩阵 $\mathbf{A}(t)$ 中所有的元素 $a_{ij}(t)$ 在 t_0 处 (或在区间 (a, b) 上) 可微, 则称函数矩阵 $\mathbf{A}(t)$ 在 t_0 处 (或在区间 (a, b) 上) 可微. 并记为

$$\begin{aligned} \mathbf{A}'(t_0) &= \left. \frac{d\mathbf{A}(t)}{dt} \right|_{t=t_0} \\ &= \lim_{\Delta t \rightarrow 0} \frac{\mathbf{A}(t_0 + \Delta t) - \mathbf{A}(t_0)}{\Delta t} \\ &= \begin{bmatrix} a'_{11}(t_0) & a'_{12}(t_0) & \cdots & a'_{1n}(t_0) \\ a'_{21}(t_0) & a'_{22}(t_0) & \cdots & a'_{2n}(t_0) \\ \vdots & \vdots & & \vdots \\ a'_{m1}(t_0) & a'_{m2}(t_0) & \cdots & a'_{mn}(t_0) \end{bmatrix}. \end{aligned}$$

Example 36. 已知 $\mathbf{A}(t) = \begin{bmatrix} \sin t & 2t^3 \\ 2\sqrt{t} & e^{2t} \end{bmatrix}$, 求 $\frac{d\mathbf{A}(t)}{dt}$.

解:

$$\frac{d\mathbf{A}(t)}{dt} = \begin{bmatrix} \frac{d}{dt}(\sin t) & \frac{d}{dt}(2t^3) \\ \frac{d}{dt}(2\sqrt{t}) & \frac{d}{dt}(e^{2t}) \end{bmatrix} = \begin{bmatrix} \cos t & 6t^2 \\ \frac{1}{\sqrt{t}} & 2e^{2t} \end{bmatrix}. \quad \square$$

关于函数矩阵, 有下面的求导法则:

1. 若 $\mathbf{A}(t), \mathbf{B}(t)$ 都是 $m \times n$ 阶矩阵可微矩阵, 则

$$\frac{d}{dt}[\mathbf{A}(t) + \mathbf{B}(t)] = \frac{d}{dt}\mathbf{A}(t) + \frac{d}{dt}\mathbf{B}(t).$$

2. 若 $\mathbf{A}(t), \mathbf{B}(t)$ 分别是 $m \times n$ 阶及 $n \times r$ 阶矩阵, 则

$$\frac{d}{dt}[\mathbf{A}(t)\mathbf{B}(t)] = \left[\frac{d}{dt}\mathbf{A}(t) \right] \mathbf{B}(t) + \mathbf{A}(t) \left[\frac{d}{dt}\mathbf{B}(t) \right].$$

3. 若 $\mathbf{A}(u)$ 可微, 且 $u = f(t)$ 关于 t 可微, 则

$$\frac{d}{dt}\mathbf{A}(f(t)) = f'(t) \frac{d}{dt}\mathbf{A}(u).$$

4. 若 $\mathbf{A}(t)$ 与 $\mathbf{A}^{-1}(t)$ 都可微, 则

$$\frac{d}{dt}(\mathbf{A}^{-1}(t)) = -\mathbf{A}^{-1}(t) \left[\frac{d}{dt}\mathbf{A}(t) \right] \mathbf{A}^{-1}(t).$$

证: (4) 注意到 $\mathbf{A}(t)\mathbf{A}^{-1}(t) = \mathbf{I}$, 两端对 t 求导, 得

$$\left[\frac{d}{dt}\mathbf{A}(t)\right]\mathbf{A}^{-1}(t) + \mathbf{A}(t)\left[\frac{d}{dt}(\mathbf{A}^{-1}(t))\right] = \mathbf{O},$$

即

$$\mathbf{A}(t)\left[\frac{d}{dt}(\mathbf{A}^{-1}(t))\right] = -\left[\frac{d}{dt}\mathbf{A}(t)\right]\mathbf{A}^{-1}(t)$$

两边左乘以 $\mathbf{A}^{-1}(t)$, 得

$$\frac{d}{dt}(\mathbf{A}^{-1}(t)) = -\mathbf{A}^{-1}(t)\left[\frac{d}{dt}\mathbf{A}(t)\right]\mathbf{A}^{-1}(t). \quad \square$$

注意矩阵乘法不满足交换律. 例如

$$\begin{aligned} \frac{d}{dt}\mathbf{A}^2(t) &= \frac{d}{dt}[\mathbf{A}(t)\mathbf{A}(t)] = \mathbf{A}'(t) \cdot \mathbf{A}(t) + \mathbf{A}(t) \cdot \mathbf{A}'(t) \\ &\neq 2\mathbf{A}(t) \cdot \mathbf{A}'(t). \end{aligned}$$

Definition 37. 若 $\mathbf{A}(t)$ 中所有的元素 $a_{ij}(t)$ 在区间 $[a, b]$ 上可积, 则称函数矩阵 $\mathbf{A}(t)$ 在区间 $[a, b]$ 上可积, 并规定

$$\int_a^b \mathbf{A}(t) dt = \left[\int_a^b a_{ij}(t) dt \right]_{m \times n}.$$

Theorem 38. 1. 若 $\mathbf{A}(t)$ 在区间 $[a, b]$ 上连续, 则对任一 $t \in (a, b)$, $\int_a^t \mathbf{A}(\tau) d\tau$ 可微, 且

$$\frac{d}{dt} \left[\int_a^t \mathbf{A}(\tau) d\tau \right] = \mathbf{A}(t).$$

2. 若 $\mathbf{A}(t)$ 在区间 $[a, b]$ 上可微, 则

$$\int_a^t \left[\frac{d}{ds} \mathbf{A}(s) \right] ds = \mathbf{A}(t) - \mathbf{A}(a), \quad t \in [a, b].$$

3.3 纯量函数关于矩阵的导数

在场论中, 我们对数量函数 $f(x, y, z)$ 定义梯度为

$$\mathbf{grad} f = \nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right),$$

这可以理解为数量函数 $f(x, y, z)$ 对向量 (x, y, z) 的导数.

注意梯度是向量, 有时也记为

$$\mathbf{grad} f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}.$$

下面我们将这一概念推广到一般情形.

Definition 39. 设 $\mathbf{X} = [x_{ij}] \in \mathbb{R}^{m \times n}$ 为变量矩阵, $f(\mathbf{X})$ 为矩阵 \mathbf{X} 的数量函数, 即看成是 $m \times n$ 元函数, 即

$$f(\mathbf{X}) = f(x_{11}, x_{12}, \dots, x_{1n}, x_{21}, x_{22}, \dots, x_{2n}, \dots, x_{m1}, x_{m2}, \dots, x_{mn}).$$

则规定数量函数 $f(\mathbf{X})$ 对于矩阵 \mathbf{X} 的导数为

$$\frac{df}{d\mathbf{X}} = \left[\frac{\partial f}{\partial x_{ij}} \right]_{m \times n} = \begin{bmatrix} \frac{\partial f}{\partial x_{11}} & \frac{\partial f}{\partial x_{12}} & \cdots & \frac{\partial f}{\partial x_{1n}} \\ \frac{\partial f}{\partial x_{21}} & \frac{\partial f}{\partial x_{22}} & \cdots & \frac{\partial f}{\partial x_{2n}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f}{\partial x_{m1}} & \frac{\partial f}{\partial x_{m2}} & \cdots & \frac{\partial f}{\partial x_{mn}} \end{bmatrix}.$$

特别地, 若变量矩阵为 m 维向量 $\mathbf{X} = [x_i]_{m \times 1}$, 这时数量函数为 $f(x_1, x_2, \dots, x_n)$,

5.57

则

$$\frac{df}{d\mathbf{X}} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{bmatrix}.$$

此即梯度.

5.58

3.4 矩阵对矩阵的导数

Definition 40. 设 $\mathbf{A} = [a_{ij}]_{m \times n}$, $\mathbf{X} = [x_{kl}]_{p \times q}$, 且 \mathbf{A} 中的各元素 a_{ij} 是矩阵 \mathbf{X} 中各元素 x_{kl} 的可微函数, 则矩阵 \mathbf{A} 对矩阵 \mathbf{X} 的导数定义为

$$\frac{d\mathbf{A}}{d\mathbf{X}} = \begin{bmatrix} \frac{\partial \mathbf{A}}{\partial x_{11}} & \frac{\partial \mathbf{A}}{\partial x_{12}} & \cdots & \frac{\partial \mathbf{A}}{\partial x_{1q}} \\ \frac{\partial \mathbf{A}}{\partial x_{21}} & \frac{\partial \mathbf{A}}{\partial x_{22}} & \cdots & \frac{\partial \mathbf{A}}{\partial x_{2q}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial \mathbf{A}}{\partial x_{p1}} & \frac{\partial \mathbf{A}}{\partial x_{p2}} & \cdots & \frac{\partial \mathbf{A}}{\partial x_{pq}} \end{bmatrix},$$

其中

$$\frac{\partial \mathbf{A}}{\partial x_{kl}} = \begin{bmatrix} \frac{\partial a_{11}}{\partial x_{kl}} & \frac{\partial a_{12}}{\partial x_{kl}} & \cdots & \frac{\partial a_{1n}}{\partial x_{kl}} \\ \frac{\partial a_{21}}{\partial x_{kl}} & \frac{\partial a_{22}}{\partial x_{kl}} & \cdots & \frac{\partial a_{2n}}{\partial x_{kl}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial a_{m1}}{\partial x_{kl}} & \frac{\partial a_{m2}}{\partial x_{kl}} & \cdots & \frac{\partial a_{mn}}{\partial x_{kl}} \end{bmatrix}, \quad k = 1, 2, \dots, p; \quad l = 1, 2, \dots, q.$$

5.59

5.60

Example 41. 设 $\mathbf{X} = (x_1, x_2, \dots, x_n)^T$, 求向量 $\mathbf{X}^T = (x_1, x_2, \dots, x_n)$ 对向量 \mathbf{X} 的导数.

解:

$$\frac{d\mathbf{X}^T}{d\mathbf{X}} = \begin{bmatrix} \frac{\partial \mathbf{X}^T}{\partial x_1} \\ \frac{\partial \mathbf{X}^T}{\partial x_2} \\ \vdots \\ \frac{\partial \mathbf{X}^T}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbf{I}_n. \quad \square$$

5.61

4 矩阵函数

4.1 矩阵多项式

Theorem 42. 设 $f(\lambda) = a_0 + a_1\lambda + \cdots + a_m\lambda^m = \sum_{k=0}^m a_k\lambda^k$, \mathbf{A} 为分块对角阵

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & & & \\ & \mathbf{A}_2 & & \\ & & \ddots & \\ & & & \mathbf{A}_t \end{bmatrix} = \text{diag}(\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_t),$$

则

$$f(\mathbf{A}) = \text{diag}[f(\mathbf{A}_1), f(\mathbf{A}_2), \dots, f(\mathbf{A}_t)].$$

5.62

证:

$$\begin{aligned} & f(\mathbf{A}) \\ &= a_0\mathbf{I} + a_1\mathbf{A} + \cdots + a_m\mathbf{A}^m \\ &= a_0 \begin{bmatrix} \mathbf{I}_1 & & & \\ & \mathbf{I}_2 & & \\ & & \ddots & \\ & & & \mathbf{I}_t \end{bmatrix} + a_1 \begin{bmatrix} \mathbf{A}_1 & & & \\ & \mathbf{A}_2 & & \\ & & \ddots & \\ & & & \mathbf{A}_t \end{bmatrix} + \cdots + a_m \begin{bmatrix} \mathbf{A}_1^m & & & \\ & \mathbf{A}_2^m & & \\ & & \ddots & \\ & & & \mathbf{A}_t^m \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=0}^m a_k \mathbf{A}_1^k & & & \\ & \sum_{k=0}^m a_k \mathbf{A}_2^k & & \\ & & \ddots & \\ & & & \sum_{k=0}^m a_k \mathbf{A}_t^k \end{bmatrix} = \begin{bmatrix} f(\mathbf{A}_1) & & & \\ & f(\mathbf{A}_2) & & \\ & & \ddots & \\ & & & f(\mathbf{A}_t) \end{bmatrix}. \quad \square \end{aligned}$$

5.63

Theorem 43. 设 $\mathbf{A} \in \mathbb{C}^{n \times n}$, \mathbf{P} 为 n 阶非奇异矩阵, $f(\lambda)$ 为多项式. 则

$$f(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \mathbf{P}^{-1}f(\mathbf{A})\mathbf{P}.$$

证: 由

$$(P^{-1}AP)^k = (P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP) = P^{-1}A^kP,$$

得

$$\begin{aligned} f(P^{-1}AP) &= \sum_{k=0}^m a_k (P^{-1}AP)^k \\ &= \sum_{k=0}^m a_k (P^{-1}A^kP) \\ &= P^{-1} \left(\sum_{k=0}^m a_k A^k \right) P \\ &= P^{-1} f(A) P. \quad \square \end{aligned}$$

5.64

矩阵多项式的计算方法

要计算 $f(A) = a_0I + a_1A + \cdots + a_mA^m$, 可以找一个 A 的零化多项式 $\varphi(\lambda)$, 且 $\partial(\varphi(\lambda)) < m$, 再由多项式带余除法有

$$f(\lambda) = \varphi(\lambda)g(\lambda) + r(\lambda),$$

则

$$f(A) = \varphi(A)g(A) + r(A).$$

又 $\varphi(A) = O$, 故

$$f(A) = r(A).$$

从而把计算 $f(A)$ 转化为计算次数较低的 $r(A)$.

这里的零化多项式 $\varphi(\lambda)$ 可以取为特征多项式, 或最小多项式等.

5.65

4.2 矩阵函数

Lemma 44. 设 r 阶方阵为 $H = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$, 则 (1) 当 $m \geq r$ 时,

$H^m = O$; (2) 当 $m < r$ 时,

$$H^m = \begin{bmatrix} 0 & \cdots & 0 & 1 & & \\ & \ddots & & \ddots & \ddots & \\ & & \ddots & & \ddots & 1 \\ & & & \ddots & & 0 \\ & & & & \ddots & \vdots \\ & & & & & 0 \end{bmatrix} \left. \vphantom{\begin{bmatrix} 0 & \cdots & 0 & 1 & & \\ & \ddots & & \ddots & \ddots & \\ & & \ddots & & \ddots & 1 \\ & & & \ddots & & 0 \\ & & & & \ddots & \vdots \\ & & & & & 0 \end{bmatrix}} \right\} m \text{ 个 } 0$$


即 H 的幂次每增加 1, 主对角线上方一排 1 就向右上方平移一个位置.

例如

$$\begin{aligned} \mathbf{H}^2 &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ & 0 & 0 & 1 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix}, \\ \mathbf{H}^3 &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ & 0 & 0 & 1 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix}, \\ \mathbf{H}^4 &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix}, \\ \mathbf{H}^m &= \mathbf{O}, \quad m \geq 4. \end{aligned}$$

由引理 44 可知,

$$\sum_{m=0}^{\infty} a_m \mathbf{H}^m = \begin{bmatrix} a_0 & a_1 & \cdots & a_{r-1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & a_1 \\ & & & a_0 \end{bmatrix}. \quad (10)$$

 这个结论解决了一类简单的矩阵幂级数的求和问题. 其特殊之处在于矩阵 \mathbf{H} 是个若当块, 且对角线元素为 0. 如果若当块对角线不是 0, 结果又如何呢?

Lemma 45. 若 $f(z) = \sum_{m=0}^{\infty} a_m z^m$, 则

$$\frac{1}{s!} f^{(s)}(z) \Big|_{z=\lambda} = \sum_{m=s}^{\infty} C_m^s a_m \lambda^{m-s}.$$

证: 因为

$$C_m^s a_m \lambda^{m-s} = \frac{m!}{s!(m-s)!} a_m \lambda^{m-s} = \frac{1}{s!} a_m \frac{d^s}{dz^s} z^m \Big|_{z=\lambda},$$

所以

$$\sum_{m=s}^{\infty} C_m^s a_m \lambda^{m-s} = \frac{1}{s!} \sum_{m=s}^{\infty} a_m \frac{d^s}{dz^s} z^m \Big|_{z=\lambda} = \frac{1}{s!} f^{(s)}(z) \Big|_{z=\lambda}. \quad \square$$

另证: 设 $f(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m + \cdots = \sum_{m=0}^{\infty} a_m z^m$, 则

$$f'(z) = a_1 + 2a_2 z + \cdots + ma_m z^{m-1} + \cdots = \sum_{m=1}^{\infty} ma_m z^{m-1},$$

$$f''(z) = 2a_2 + \cdots + m(m-1)a_m z^{m-2} + \cdots = \sum_{m=2}^{\infty} m(m-1)a_m z^{m-2},$$

⋮

$$\begin{aligned} f^{(s)}(z) &= \sum_{m=s}^{\infty} m(m-1)\cdots(m-s+1)a_m z^{m-s} \\ &= \sum_{m=s}^{\infty} \frac{m!}{(m-s)!} a_m z^{m-s} \\ &= s! \sum_{m=s}^{\infty} \frac{m!}{s!(m-s)!} a_m z^{m-s} \\ &= s! \sum_{m=s}^{\infty} C_m^s a_m z^{m-s}. \end{aligned}$$

所以

$$\frac{1}{s!} f^{(s)}(z) = \sum_{m=s}^{\infty} C_m^s a_m z^{m-s}.$$

故

$$\frac{1}{s!} f^{(s)}(z) \Big|_{z=\lambda} = \sum_{m=s}^{\infty} C_m^s a_m \lambda^{m-s}. \quad \square$$

5.70

5.71

Theorem 46. 设 $f(z) = \sum_{m=0}^{\infty} a_m z^m$ 的收敛半径为 R , 又 r 阶若当块为 $\mathbf{J} =$

$$\begin{bmatrix} \lambda & 1 & & & \\ & \ddots & \ddots & & \\ & & \lambda & 1 & \\ & & & & \lambda \end{bmatrix}. \text{ 则当 } |\lambda| < R \text{ 时, 级数 } \sum_{m=0}^{\infty} a_m \mathbf{J}^m \text{ 收敛, 且}$$

$$\sum_{m=0}^{\infty} a_m \mathbf{J}^m = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2!} f''(\lambda) & \cdots & \frac{1}{(r-1)!} f^{(r-1)}(\lambda) \\ & f(\lambda) & f'(\lambda) & \cdots & \frac{1}{(r-2)!} f^{(r-2)}(\lambda) \\ & & \ddots & \ddots & \vdots \\ & & & f(\lambda) & f'(\lambda) \\ & & & & f(\lambda) \end{bmatrix}.$$

5.72

证: 矩阵 \mathbf{J} 可以写成 $\lambda \mathbf{I} + \mathbf{H}$, 因为 \mathbf{I} 是单位矩阵, 所以 \mathbf{I} 可以与 \mathbf{H} 交换. 于是 $m > r$ 时,

$$\begin{aligned} \mathbf{J}^m &= (\lambda \mathbf{I} + \mathbf{H})^m = \lambda^m \mathbf{I} + C_m^1 \lambda^{m-1} \mathbf{H} + \cdots + C_m^{m-1} \lambda \mathbf{H}^{m-1} + \mathbf{H}^m \\ &= \begin{bmatrix} \lambda^m & C_m^1 \lambda^{m-1} & \cdots & C_m^{r-1} \lambda^{m-r+1} \\ & \ddots & \ddots & \vdots \\ & & \lambda^m & C_m^1 \lambda^{m-1} \\ & & & \lambda^m \end{bmatrix}. \end{aligned} \quad (\text{对照公式 (10)})$$

所以

$$\sum_{m=0}^{\infty} a_m \mathbf{J}^m = \begin{bmatrix} \sum_{m=0}^{\infty} a_m \lambda^m & \sum_{m=0}^{\infty} a_m C_m^1 \lambda^{m-1} & \cdots & \sum_{m=0}^{\infty} a_m C_m^{r-1} \lambda^{m-r+1} \\ & \sum_{m=0}^{\infty} a_m \lambda^m & \cdots & \sum_{m=0}^{\infty} a_m C_m^{r-2} \lambda^{m-r+2} \\ & & \ddots & \vdots \\ & & & \sum_{m=0}^{\infty} a_m \lambda^m \end{bmatrix}$$

$$= \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2!} f''(\lambda) & \cdots & \frac{1}{(r-1)!} f^{(r-1)}(\lambda) \\ & f(\lambda) & f'(\lambda) & \cdots & \frac{1}{(r-2)!} f^{(r-2)}(\lambda) \\ & & \ddots & \ddots & \vdots \\ & & & f(\lambda) & f'(\lambda) \\ & & & & f(\lambda) \end{bmatrix}.$$

最后一个等号的成立需要 $|\lambda| < R$. 得证级数 $\sum_{m=0}^{\infty} a_m \mathbf{J}^m$ 收敛. \square

Theorem 47. 设幂级数 $f(z) = \sum_{m=0}^{\infty} a_m z^m$ 的收敛半径为 R , 且 $\mathbf{A} \in \mathbb{C}^{n \times n}$ 的谱半径 (即 \mathbf{A} 的特征值模的最大值) 为 $\rho(\mathbf{A})$, 则当 $\rho(\mathbf{A}) < R$ 时, 矩阵幂级数 $\sum_{m=0}^{\infty} a_m \mathbf{A}^m$ 收敛.

证: 设 \mathbf{A} 的若当标准形为

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & & & \\ & \mathbf{J}_2 & & \\ & & \ddots & \\ & & & \mathbf{J}_s \end{bmatrix} = \mathbf{J}_1 \oplus \mathbf{J}_2 \oplus \cdots \oplus \mathbf{J}_s,$$

其中 \mathbf{J}_i 是特征值为 λ_i 的 r_i 阶若当块. 则存在满秩方阵 \mathbf{P} , 使 $\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}$. 于是

$$\begin{aligned} \sum_{m=0}^{\infty} a_m \mathbf{A}^m &= \sum_{m=0}^{\infty} a_m (\mathbf{P}\mathbf{J}\mathbf{P}^{-1})^m \\ &= \sum_{m=0}^{\infty} a_m \mathbf{P}\mathbf{J}^m \mathbf{P}^{-1} \\ &= \mathbf{P} \left(\sum_{m=0}^{\infty} a_m \mathbf{J}^m \right) \mathbf{P}^{-1} \\ &= \mathbf{P} \left(\sum_{m=0}^{\infty} a_m \mathbf{J}_1^m \oplus \sum_{m=0}^{\infty} a_m \mathbf{J}_2^m \oplus \cdots \oplus \sum_{m=0}^{\infty} a_m \mathbf{J}_s^m \right) \mathbf{P}^{-1}. \end{aligned}$$

因为 $\rho(\mathbf{A}) < R$, 故 $|\lambda_i| < R$, 由定理 46 知 $\sum_{m=0}^{\infty} a_m \mathbf{J}_i^m$ 收敛, 因而推得 $\sum_{m=0}^{\infty} a_m \mathbf{A}^m$ 收敛. \square

☞ 请回顾上述推导中, 由特殊到一般的渐进过程:

$$\sum_{m=0}^{\infty} a_m \mathbf{H}^m \implies \sum_{m=0}^{\infty} a_m \mathbf{J}^m \implies \sum_{m=0}^{\infty} a_m \mathbf{A}^m.$$

5.76

Definition 48. 设幂级数 $\sum_{m=0}^{\infty} a_m z^m$ 的收敛半径为 R , 且对任意的 $|z| < R$, 幂级数收敛于 $f(z)$, 即

$$f(z) = \sum_{m=0}^{\infty} a_m z^m \quad (|z| < R).$$

如果 $\mathbf{A} \in \mathbb{C}^{n \times n}$ 满足 $\rho(\mathbf{A}) < R$, 则记矩阵幂级数 $\sum_{m=0}^{\infty} a_m \mathbf{A}^m$ 的和为 $f(\mathbf{A})$, 即

$$f(\mathbf{A}) = \sum_{m=0}^{\infty} a_m \mathbf{A}^m.$$

称 $f(\mathbf{A})$ 为自变量 \mathbf{A} 的矩阵函数.

5.77

例如, 对于如下函数的幂级数展开式

$$(1 - z)^{-1} = \sum_{m=0}^{\infty} z^m \quad (R = 1),$$

$$\ln(1 + z) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m+1} z^{m+1} \quad (R = 1).$$

相应地有矩阵函数

$$(\mathbf{I} - \mathbf{A})^{-1} = \sum_{m=0}^{\infty} \mathbf{A}^m \quad (\rho(\mathbf{A}) < 1),$$

$$\ln(\mathbf{I} + \mathbf{A}) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m+1} \mathbf{A}^{m+1} \quad (\rho(\mathbf{A}) < 1).$$

以及

- 矩阵指数函数: $e^{\mathbf{A}} = \mathbf{I} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \cdots + \frac{\mathbf{A}^k}{k!} + \cdots$.
- 矩阵正弦函数: $\sin \mathbf{A} = \mathbf{A} - \frac{\mathbf{A}^3}{3!} + \cdots + (-1)^{k-1} \frac{\mathbf{A}^{2k-1}}{(2k-1)!} + \cdots$,
- 矩阵余弦函数: $\cos \mathbf{A} = \mathbf{I} - \frac{\mathbf{A}^2}{2!} + \cdots + (-1)^k \frac{\mathbf{A}^{2k}}{(2k)!} + \cdots$.

定理 47 的证明过程给出了求 $f(\mathbf{A})$ 的一个方法.

5.78

$$f(\mathbf{A}) = \mathbf{P} \left(f(\mathbf{J}_1) \oplus f(\mathbf{J}_2) \oplus \cdots \oplus f(\mathbf{J}_s) \right) \mathbf{P}^{-1},$$

其中

$$f(\mathbf{J}_i) = \sum_{m=0}^{\infty} a_m \mathbf{J}_i^m = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!} f''(\lambda_i) & \cdots & \frac{1}{(r_i-1)!} f^{(r_i-1)}(\lambda_i) \\ & f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{1}{(r_i-2)!} f^{(r_i-2)}(\lambda_i) \\ & & \ddots & \ddots & \vdots \\ & & & f(\lambda_i) & f'(\lambda_i) \\ & & & & f(\lambda_i) \end{bmatrix}.$$

这里 r_i 是若当块 \mathbf{J}_i 的阶数.

实际计算中, 若当块的阶数多为 1, 2, 3, 即分别为

$$\mathbf{J}_i = \lambda_i, \quad \mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 \\ & \lambda_i \end{bmatrix}, \quad \mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & 0 \\ & \lambda_i & 1 \\ & & \lambda_i \end{bmatrix},$$

此时对应地有

$$f(\mathbf{J}_i) = f(\lambda_i), \quad f(\mathbf{J}_i) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) \\ & f(\lambda_i) \end{bmatrix}, \quad f(\mathbf{J}_i) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!}f''(\lambda_i) \\ & f(\lambda_i) & f'(\lambda_i) \\ & & f(\lambda_i) \end{bmatrix}.$$

比如设 $f(\mathbf{A}) = e^{\mathbf{A}}$, 假定 \mathbf{A} 的为若当矩阵:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ & 2 & 1 \\ & & 2 \end{bmatrix},$$

则

$$f(\mathbf{A}) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2!}f''(\lambda) \\ & f(\lambda) & f'(\lambda) \\ & & f(\lambda) \end{bmatrix} = \begin{bmatrix} e^2 & e^2 & \frac{1}{2}e^2 \\ & e^2 & e^2 \\ & & e^2 \end{bmatrix}.$$

Example 49. 设 $\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 3 \end{bmatrix}$, 求下列矩阵函数: (1) \mathbf{A}^{20} ; (2) $e^{\mathbf{A}}$; (3) $\sin \mathbf{A}$.

解: 先求 \mathbf{A} 的若当标准形. 因

$$\lambda \mathbf{I} - \mathbf{A} = \begin{bmatrix} \lambda - 2 & 0 & 0 \\ -1 & \lambda - 1 & -1 \\ -1 & 1 & \lambda - 3 \end{bmatrix} \sim \begin{bmatrix} 1 & & \\ & \lambda - 2 & \\ & & (\lambda - 2)^2 \end{bmatrix}.$$

所以若当标准形为

$$\mathbf{J} = \begin{bmatrix} 2 & 0 & 0 \\ & 2 & 1 \\ & & 2 \end{bmatrix}.$$

则 $\mathbf{J}_1 = 2, \mathbf{J}_2 = \begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}$.

由 $\mathbf{AP} = \mathbf{PJ}$ 可求得

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{P}^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}.$$

(1) 记 $f(z) = z^{20}$. 则

$$f(\mathbf{J}_1) = 2^{20}, \quad f(\mathbf{J}_2) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) \\ & f(\lambda_i) \end{bmatrix} = \begin{bmatrix} 2^{20} & 20 \cdot 2^{19} \\ & 2^{20} \end{bmatrix}.$$

故

$$\begin{aligned} \mathbf{A}^{20} &= \mathbf{P}(f(\mathbf{J}_1) \oplus f(\mathbf{J}_2))\mathbf{P}^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2^{20} & 0 & 0 \\ & 2^{20} & 20 \cdot 2^{19} \\ & & 2^{20} \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \\ &= 2^{20} \begin{bmatrix} 1 & 0 & 0 \\ 10 & -9 & 10 \\ 10 & -10 & 11 \end{bmatrix}. \end{aligned}$$

(2) 记 $f(z) = e^z$. 则

$$f(\mathbf{J}_1) = e^2, \quad f(\mathbf{J}_2) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) \\ & f(\lambda_i) \end{bmatrix} = \begin{bmatrix} e^2 & e^2 \\ & e^2 \end{bmatrix}.$$

故

$$\begin{aligned} e^{\mathbf{A}} &= \mathbf{P}f(\mathbf{J})\mathbf{P}^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} e^2 & 0 & 0 \\ & e^2 & e^2 \\ & & e^2 \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \\ &= e^2 \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}. \end{aligned}$$

(3) 记 $f(z) = \sin z$. 则

$$f(\mathbf{J}_1) = \sin 2, \quad f(\mathbf{J}_2) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) \\ & f(\lambda_i) \end{bmatrix} = \begin{bmatrix} \sin 2 & \cos 2 \\ & \sin 2 \end{bmatrix}.$$

故

$$\begin{aligned} \sin \mathbf{A} &= \mathbf{P}f(\mathbf{J})\mathbf{P}^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \sin 2 & 0 & 0 \\ & \sin 2 & \cos 2 \\ & & \sin 2 \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \sin 2 & 0 & 0 \\ \cos 2 & \sin 2 - \cos 2 & \cos 2 \\ \cos 2 & -\cos 2 & \sin 2 + \cos 2 \end{bmatrix}. \quad \square \end{aligned}$$

 矩阵 \mathbf{P} 是怎么求得的?

设 $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3]$, 则有

$$\mathbf{A}[\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3] = [\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3] \begin{bmatrix} 2 & 0 & 0 \\ & 2 & 1 \\ & & 2 \end{bmatrix},$$

得 $[Ap_1, Ap_2, Ap_3] = [2p_1, 2p_2, p_2 + 2p_3]$. 于是有

$$\begin{aligned} Ap_1 &= 2p_1, & \text{即 } (2I - A)p_1 &= \mathbf{0}, \\ Ap_2 &= 2p_2, & \text{即 } (2I - A)p_2 &= \mathbf{0}, \\ Ap_3 &= p_2 + 2p_3, & \text{即 } (2I - A)p_3 &= -p_2. \end{aligned}$$

由

$$2I - A = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & -1 \\ -1 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

知齐次方程组 $(2I - A)x = \mathbf{0}$ 的基础解系为

$$\alpha_1 = (1, 1, 0)^T, \quad \alpha_2 = (0, 1, 1)^T.$$

又

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & b_1 \\ -1 & 1 & -1 & b_2 \\ -1 & 1 & -1 & b_3 \end{array} \right] \sim \left[\begin{array}{ccc|c} 1 & -1 & 1 & -b_3 \\ 0 & 0 & 0 & b_3 - b_2 \\ 0 & 0 & 0 & b_1 \end{array} \right],$$

故要使方程组 $(2I - A)p_3 = -p_2$ 有解, $-p_2$ 需满足 $b_1 = 0, b_2 = b_3$. 则 $(2I - A)p_2 = \mathbf{0}$ 的解可取为

$$p_2 = \alpha_2 = (0, 1, 1)^T,$$

从而方程组 $(2I - A)p_1 = \mathbf{0}$ 的解可取为

$$p_1 = \alpha_1 = (1, 1, 0)^T.$$

再求解非齐次线性方程组 $(2I - A)p_3 = -p_2$, 由

$$[2I - A, -p_2] = \left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ -1 & 1 & -1 & -1 \\ -1 & 1 & -1 & -1 \end{array} \right] \sim \left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right],$$

可取

$$p_3 = (1, 0, 0)^T.$$

另由上述过程可知 P 不唯一, 比如可以取 $p_1 = \alpha_2$ 等.

Example 50. 设 $A = \begin{bmatrix} 2 & -1 & -1 \\ 2 & -1 & -2 \\ -1 & 1 & 2 \end{bmatrix}$, 求矩阵 A 的 Jordan 标准形 J , 并求变换矩阵 P , 使 $P^{-1}AP = J$.

解: 先求 A 的若当标准形. 因

$$\lambda I - A = \begin{bmatrix} \lambda - 2 & 1 & 1 \\ -2 & \lambda + 1 & 2 \\ 1 & -1 & \lambda - 2 \end{bmatrix} \xrightarrow[r_2 + 2r_3]{r_1 - (\lambda - 2)r_2} \begin{bmatrix} 0 & \lambda - 1 & -\lambda^2 + 4\lambda - 3 \\ 0 & \lambda - 1 & 2\lambda - 2 \\ 1 & -1 & \lambda - 2 \end{bmatrix}$$

$$\begin{aligned} & \xrightarrow[c_3 - (\lambda - 2)c_1]{c_2 + c_1} \begin{bmatrix} 0 & \lambda - 1 & -\lambda^2 + 4\lambda - 3 \\ 0 & \lambda - 1 & 2\lambda - 2 \\ 1 & 0 & 0 \end{bmatrix} \xrightarrow{r_1 \leftrightarrow r_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda - 1 & 2\lambda - 2 \\ 0 & \lambda - 1 & -\lambda^2 + 4\lambda - 3 \end{bmatrix} \\ & \xrightarrow{r_3 - r_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda - 1 & 2\lambda - 2 \\ 0 & 0 & -\lambda^2 + 2\lambda - 1 \end{bmatrix} \xrightarrow[c_3 - 2c_2]{r_3 \times (-1)} \begin{bmatrix} 1 & & \\ & \lambda - 1 & \\ & & (\lambda - 1)^2 \end{bmatrix}. \end{aligned}$$

所以若当标准形为

$$J = \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 1 \\ & & 1 \end{bmatrix}.$$

设 $P = [p_1, p_2, p_3]$, 则有

$$A[p_1, p_2, p_3] = [p_1, p_2, p_3] \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 1 \\ & & 1 \end{bmatrix},$$

得 $[Ap_1, Ap_2, Ap_3] = [p_1, p_2, p_2 + p_3]$. 于是有

$$\begin{aligned} Ap_1 &= p_1, & \text{即 } (I - A)p_1 &= 0, \\ Ap_2 &= p_2, & \text{即 } (I - A)p_2 &= 0, \\ Ap_3 &= p_2 + p_3, & \text{即 } (I - A)p_3 &= -p_2. \end{aligned}$$

即 p_1, p_2 是方程组 $(I - A)x = 0$ 的解. 由

$$I - A = \begin{bmatrix} -1 & 1 & 1 \\ -2 & 2 & 2 \\ 1 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

取齐次方程组 $(I - A)x = 0$ 的基础解系为 $\alpha_1 = (1, 1, 0)^T$, $\alpha_2 = (1, 0, 1)^T$. 可以选取 $p_1 = \alpha_1$. 但是不能简单地令 $p_2 = \alpha_2$, 因为 p_2 还要保证 $(I - A)p_3 = -p_2$ 有解.

令

$$p_2 = k_1 \alpha_1 + k_2 \alpha_2 = k_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} k_1 + k_2 \\ k_1 \\ k_2 \end{bmatrix},$$

其中 k_1, k_2 是不全为零的待定常数. 由

$$(I - A, -p_2) = \left[\begin{array}{ccc|c} -1 & 1 & 1 & -k_1 - k_2 \\ -2 & 2 & 2 & -k_1 \\ 1 & -1 & -1 & -k_2 \end{array} \right] \sim \left[\begin{array}{ccc|c} 1 & -1 & -1 & -k_2 \\ 0 & 0 & 0 & k_1 + 2k_2 \\ 0 & 0 & 0 & 0 \end{array} \right],$$

故要使方程组 $(I - A)p_3 = -p_2$ 有解, 需满足 $k_1 + 2k_2 = 0$, 即

$$k_1 = -2k_2 (\neq 0).$$

取 $k_1 = 2, k_2 = -1$, 则

$$\mathbf{p}_2 = 2\boldsymbol{\alpha}_1 - 1\boldsymbol{\alpha}_2 = (1, 2, -1)^T.$$

再求解非齐次线性方程组 $(\mathbf{I} - \mathbf{A})\mathbf{p}_3 = -\mathbf{p}_2$, 由

5.92

$$[\mathbf{I} - \mathbf{A}, -\mathbf{p}_2] = \left[\begin{array}{ccc|c} -1 & 1 & 1 & -1 \\ -2 & 2 & 2 & -2 \\ 1 & -1 & -1 & 1 \end{array} \right] \sim \left[\begin{array}{ccc|c} 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right],$$

可取 $\mathbf{p}_3 = (1, 0, 0)^T$. 故可取

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 0 & -1 & 0 \end{bmatrix},$$

使 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$. □

5.93

Example 51. 设 $f(z) = \sum_{m=0}^{\infty} \left(\frac{z}{3}\right)^m$, 求 $f(\mathbf{A})$, 其中 $\mathbf{A} = \begin{bmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}$.

解: $f(z) = \left(1 - \frac{z}{3}\right)^{-1}$, 其收敛半径为 3. 注意到 \mathbf{A} 为若当矩阵, 其特征值为 2, 故谱半径 $\rho(\mathbf{A}) = 2 < 3$, 所以 $f(\mathbf{A}) = \sum_{m=0}^{\infty} \left(\frac{\mathbf{A}}{3}\right)^m$ 收敛, 且

$$f(\mathbf{A}) = \begin{bmatrix} f(2) & f'(2) & \frac{1}{2!}f''(2) & \frac{1}{3!}f'''(2) \\ & f(2) & f'(2) & \frac{1}{2!}f''(2) \\ & & f(2) & f'(2) \\ & & & f(2) \end{bmatrix}.$$

5.94

因

$$f'(z) = \frac{1}{3}\left(1 - \frac{z}{3}\right)^{-2}, \quad f''(z) = \frac{2}{9}\left(1 - \frac{z}{3}\right)^{-3}, \quad f'''(z) = \frac{2}{9}\left(1 - \frac{z}{3}\right)^{-4},$$

所以

$$f(2) = 3, \quad f'(2) = 3, \quad f''(2) = 6, \quad f'''(2) = 18.$$

故

$$f(\mathbf{A}) = \begin{bmatrix} 3 & 3 & 3 & 3 \\ & 3 & 3 & 3 \\ & & 3 & 3 \\ & & & 3 \end{bmatrix}.$$

当然, 也可以由 $f(\mathbf{A}) = \left(\mathbf{I} - \frac{\mathbf{A}}{3}\right)^{-1}$ 得到答案. □

5.95

Definition 52. 设 $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\lambda_1, \lambda_2, \dots, \lambda_s$ 是 \mathbf{A} 的谱. \mathbf{A} 的最小多项式为 m 次多项式 $m(\lambda)$,

$$m(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_s)^{m_s},$$

其中 $m_1 + m_2 + \cdots + m_s = m$. 记

$$l = \max\{m_1, m_2, \dots, m_s\},$$

设 $f(\lambda)$ 是一个给定的具有 $l-1$ 阶导数的函数, 则我们把下列 m 个值

$$\left. \begin{array}{l} f(\lambda_1), f'(\lambda_1), \dots, f^{(m_1-1)}(\lambda_1), \\ f(\lambda_2), f'(\lambda_2), \dots, f^{(m_2-1)}(\lambda_2), \\ \vdots \\ f(\lambda_s), f'(\lambda_s), \dots, f^{(m_s-1)}(\lambda_s), \end{array} \right\} \quad (11)$$

称为 $f(\lambda)$ 在 \mathbf{A} 上的谱值. 若这些值都为有限值, 则称函数 $f(\lambda)$ 在 \mathbf{A} 的谱上给定.

我们试图把计算一般矩阵函数 $f(\mathbf{A})$, 转化为计算矩阵多项式. 构造多项式 $p(\lambda)$, 使

$$\left. \begin{array}{l} p(\lambda_1) = f(\lambda_1), p'(\lambda_1) = f'(\lambda_1), \dots, p^{(m_1-1)}(\lambda_1) = f^{(m_1-1)}(\lambda_1), \\ p(\lambda_2) = f(\lambda_2), p'(\lambda_2) = f'(\lambda_2), \dots, p^{(m_2-1)}(\lambda_2) = f^{(m_2-1)}(\lambda_2), \\ \vdots \\ p(\lambda_s) = f(\lambda_s), p'(\lambda_s) = f'(\lambda_s), \dots, p^{(m_s-1)}(\lambda_s) = f^{(m_s-1)}(\lambda_s), \end{array} \right\} \quad (12)$$

即函数 $f(\lambda)$ 与多项式 $p(\lambda)$ 在 \mathbf{A} 上的谱值相同.

设 $\mathbf{A} = \mathbf{Q}\mathbf{J}\mathbf{Q}^{-1}$, 其中 \mathbf{J} 是 \mathbf{A} 的若当标准形, \mathbf{Q} 是相似变换矩阵. 由

$$f(\mathbf{J}_i) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{1}{(r-1)!} f^{(r-1)}(\lambda_i) \\ & \ddots & \ddots & \vdots \\ & & f(\lambda_i) & f'(\lambda_i) \\ & & & f(\lambda_i) \end{bmatrix},$$

知 $f(\mathbf{J}) = p(\mathbf{J})$.

于是

$$f(\mathbf{A}) = \mathbf{Q}f(\mathbf{J})\mathbf{Q}^{-1} = \mathbf{Q}p(\mathbf{J})\mathbf{Q}^{-1} = p(\mathbf{Q}\mathbf{J}\mathbf{Q}^{-1}) = p(\mathbf{A}).$$

对于任一多项式 $g(\lambda)$, 存在次数小于 m 的多项式 $p(\lambda)$, 使

$$g(\mathbf{A}) = p(\mathbf{A}).$$

令

$$p(\lambda) = a_0 + a_1\lambda + \cdots + a_{m-1}\lambda^{m-1},$$

由方程组 (12) 可以确定待定系数 a_0, a_1, \dots, a_{m-1} .

Exercise 53 (P. 211 习题一 9 (1)). 计算级数 $\sum_{k=0}^{\infty} \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}^k$ 的和.

解: 记 $\mathbf{A} = \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}$, 由

$$|\lambda \mathbf{I} - \mathbf{A}| = \begin{vmatrix} \lambda - 0.1 & -0.7 \\ -0.3 & \lambda - 0.6 \end{vmatrix} = \lambda^2 - 0.7\lambda - 0.15,$$

得特征值为 $\lambda_{1,2} = \frac{0.7 \pm \sqrt{1.09}}{2}$, 故谱半径 $\rho(\mathbf{A}) = \frac{0.7 + \sqrt{1.09}}{2} < 1$, 从而

$$\sum_{k=0}^{\infty} \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}^k = (\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} 0.9 & -0.7 \\ -0.3 & 0.4 \end{bmatrix}^{-1} = \frac{1}{0.15} \begin{bmatrix} 0.4 & 0.7 \\ 0.3 & 0.9 \end{bmatrix}. \quad \square$$