第 5 章　矩阵分析

Matrix Theory

黄 正 华

Email: huangzh@whu.edu.cn

武汉大学　数学与统计学院

December 29, 2014

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

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.

Outline

1. 向量范数及矩阵范数
   - 向量范数
   - 矩阵范数
2. 矩阵序列与矩阵级数
3. 矩阵的微分与积分
4. 矩阵函数

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.

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.

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 1.1

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: $\|\boldsymbol{x}\| \geqslant 0$. 当且仅当 $\boldsymbol{x} = \mathbf{0}$ 时, $\|\boldsymbol{x}\| = 0$;

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 1.1

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: $\|\boldsymbol{x}\| \geqslant 0$. 当且仅当 $\boldsymbol{x} = \mathbf{0}$ 时, $\|\boldsymbol{x}\| = 0$;
2. 齐次性: 对任意 $\lambda \in F$, $\|\lambda \boldsymbol{x}\| = |\lambda| \|\boldsymbol{x}\|$;

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 1.1

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: $\|\boldsymbol{x}\| \geqslant 0$. 当且仅当 $\boldsymbol{x} = \boldsymbol{0}$ 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

$$\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|,$$

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 1.1

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: $\|\boldsymbol{x}\| \geqslant 0$. 当且仅当 $\boldsymbol{x} = \boldsymbol{0}$ 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

$$\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|,$$

则称 $\|\boldsymbol{x}\|$ 为线性空间 V 中向量 $\boldsymbol{x}$ 的范数(norm).

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 1.1

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: $\|\boldsymbol{x}\| \geqslant 0$. 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

$$\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|,$$

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.

基本性质

1. 对任意的 $\boldsymbol{x} \in V$, 有

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

基本性质

1. 对任意的 $\boldsymbol{x} \in V$, 有

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

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
$$\Big| \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \Big| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|. \tag{1}$$

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

$$\big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|. \tag{1}$$

事实上, (1) 式等价于

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

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
$$\big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|. \tag{1}$$

事实上, (1) 式等价于
$$-\|\boldsymbol{x} - \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|.$$

而
$$-\|\boldsymbol{x} - \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \Leftrightarrow \|\boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{y}\|, \tag{2}$$
$$\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\| \Leftrightarrow \|\boldsymbol{x}\| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y}\|. \tag{3}$$

由

$$\|\boldsymbol{y}\| = \|\boldsymbol{x} + \boldsymbol{y} - \boldsymbol{x}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y} - \boldsymbol{x}\| = \|\boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{y}\|, \tag{4}$$

$$\|\boldsymbol{x}\| = \|\boldsymbol{x} - \boldsymbol{y} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y}\|, \tag{5}$$

得证 (1) 式成立. □

由

$$\|\boldsymbol{y}\| = \|\boldsymbol{x}+\boldsymbol{y}-\boldsymbol{x}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}-\boldsymbol{x}\| = \|\boldsymbol{x}\| + \|\boldsymbol{x}-\boldsymbol{y}\|, \tag{4}$$

$$\|\boldsymbol{x}\| = \|\boldsymbol{x}-\boldsymbol{y}+\boldsymbol{y}\| \leqslant \|\boldsymbol{x}-\boldsymbol{y}\| + \|\boldsymbol{y}\|, \tag{5}$$

得证 (1) 式成立. □

☞ 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.

由

$$\|\boldsymbol{y}\| = \|\boldsymbol{x}+\boldsymbol{y}-\boldsymbol{x}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}-\boldsymbol{x}\| = \|\boldsymbol{x}\| + \|\boldsymbol{x}-\boldsymbol{y}\|, \tag{4}$$

$$\|\boldsymbol{x}\| = \|\boldsymbol{x}-\boldsymbol{y}+\boldsymbol{y}\| \leqslant \|\boldsymbol{x}-\boldsymbol{y}\| + \|\boldsymbol{y}\|, \tag{5}$$

得证 (1) 式成立. □

☞ 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.

线性空间 $\mathbb{C}^n$ 的范数

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

1. 2-范数

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

线性空间 $\mathbb{C}^n$ 的范数

任取 $\boldsymbol{x} \in \mathbb{C}^n$, 设 $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}}$, 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

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

线性空间 $\mathbb{C}^n$ 的范数

任取 $\boldsymbol{x} \in \mathbb{C}^n$, 设 $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}}$, 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
$$\|\boldsymbol{x}\|_\infty = \max_{1 \leqslant i \leqslant n} |\xi_i|.$$

线性空间 $\mathbb{C}^n$ 的范数

任取 $\boldsymbol{x} \in \mathbb{C}^n$, 设 $\boldsymbol{x} = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}}$, 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
$$\|\boldsymbol{x}\|_\infty = \max_{1 \leqslant i \leqslant n} |\xi_i|.$$

以上三种范数都是以下 p-范数的特例:
$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |\xi_i|^p\right)^{\frac{1}{p}}, \qquad 1 \leqslant p < +\infty.$$

- p-范数定义式中, 若取 $0 < p < 1$, 则不满足三角不等式.

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

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

- 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 $\boldsymbol{\alpha} = (1, 0, 0)^{\mathrm{T}}$, $\boldsymbol{\beta} = (0, 1, 0,)^{\mathrm{T}}$, 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.

- 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 $\boldsymbol{\alpha} = (1,0,0)^{\mathrm{T}}$, $\boldsymbol{\beta} = (0,1,0,)^{\mathrm{T}}$, 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.

Theorem 1.2

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

$$\lim_{p\to\infty} \|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty. \tag{6}$$

Theorem 1.2

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

$$\lim_{p\to\infty} \|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty. \tag{6}$$

证: 因

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

Theorem 1.2

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

$$\lim_{p\to\infty} \|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty. \tag{6}$$

证: 因

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

Theorem 1.2

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

$$\lim_{p\to\infty} \|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty. \tag{6}$$

证: 因

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

故有

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

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

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.

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 1.3

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

$$K_1\|\boldsymbol{x}\|_\beta \leqslant \|\boldsymbol{x}\|_\alpha \leqslant K_2\|\boldsymbol{x}\|_\beta.$$

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 1.3

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

$$K_1\|\boldsymbol{x}\|_\beta \leqslant \|\boldsymbol{x}\|_\alpha \leqslant K_2\|\boldsymbol{x}\|_\beta.$$

由

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

表明, 任何范数 $\|\boldsymbol{x}\|_p$ 均与 $\|\boldsymbol{x}\|_\infty$ 等价,

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 1.3

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

$$K_1\|\boldsymbol{x}\|_\beta \leqslant \|\boldsymbol{x}\|_\alpha \leqslant K_2\|\boldsymbol{x}\|_\beta.$$

由

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

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

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 1.3

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

$$K_1\|\boldsymbol{x}\|_\beta \leqslant \|\boldsymbol{x}\|_\alpha \leqslant K_2\|\boldsymbol{x}\|_\beta.$$

由

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

表明, 任何范数 $\|\boldsymbol{x}\|_p$ 均与 $\|\boldsymbol{x}\|_\infty$ 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 $\|\boldsymbol{x}\|_1$, $\|\boldsymbol{x}\|_2$ 和 $\|\boldsymbol{x}\|_\infty$ 彼此等价.

Outline

1. 向量范数及矩阵范数
   - 向量范数
   - 矩阵范数
2. 矩阵序列与矩阵级数
3. 矩阵的微分与积分
4. 矩阵函数

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.

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

$$\|\boldsymbol{A}\|_{V_1} = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|,$$

$$\|\boldsymbol{A}\|_{V_2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}} = \sqrt{\operatorname{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})} \triangleq \|\boldsymbol{A}\|_F, \qquad \text{(Frobenius 范数)}$$

$$\|\boldsymbol{A}\|_{V_\infty} = \max_{i,j}|a_{ij}|, \qquad \text{(契比雪夫范数)}$$

$$\|\boldsymbol{A}\|_{V_p} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^p\right)^{\frac{1}{p}}, \quad 1 \leqslant p < +\infty.$$

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

$$\|\boldsymbol{A}\|_{V_1} = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|,$$

$$\|\boldsymbol{A}\|_{V_2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}} = \sqrt{\operatorname{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})} \triangleq \|\boldsymbol{A}\|_F, \qquad \text{(Frobenius 范数)}$$

$$\|\boldsymbol{A}\|_{V_\infty} = \max_{i,j}|a_{ij}|, \qquad \text{(契比雪夫范数)}$$

$$\|\boldsymbol{A}\|_{V_p} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^p\right)^{\frac{1}{p}}, \quad 1 \leqslant p < +\infty.$$

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

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

$$\|\boldsymbol{A}\|_{V_1} = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|,$$

$$\|\boldsymbol{A}\|_{V_2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}} = \sqrt{\operatorname{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})} \triangleq \|\boldsymbol{A}\|_F, \qquad \text{(Frobenius 范数)}$$

$$\|\boldsymbol{A}\|_{V_\infty} = \max_{i,j}|a_{ij}|, \qquad \text{(契比雪夫范数)}$$

$$\|\boldsymbol{A}\|_{V_p} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^p\right)^{\frac{1}{p}}, \quad 1 \leqslant p < +\infty.$$

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.

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

$$\|\boldsymbol{A}\|_{V_1}=\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|,$$

$$\|\boldsymbol{A}\|_{V_2}=\Big(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\Big)^{\frac{1}{2}}=\sqrt{\operatorname{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})}\triangleq\|\boldsymbol{A}\|_F, \qquad \text{(Frobenius 范数)}$$

$$\|\boldsymbol{A}\|_{V_\infty}=\max_{i,j}|a_{ij}|, \qquad \text{(契比雪夫范数)}$$

$$\|\boldsymbol{A}\|_{V_p}=\Big(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^p\Big)^{\frac{1}{p}},\quad 1\leqslant p<+\infty.$$

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, $\|\boldsymbol{A}\|_{V_p}$ 在 $p=1$, $p=2$, $p\to\infty$ 时分别成为 $\|\boldsymbol{A}\|_{V_1}$, $\|\boldsymbol{A}\|_{V_2}$, $\|\boldsymbol{A}\|_{V_\infty}$.

F-范数

Frobenius 范数

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

简称为 F-范数, 是最常见的范数之一.

F-范数

Frobenius 范数

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

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

$$(\boldsymbol{A} \mid \boldsymbol{B}) = \operatorname{tr}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}) = \sum_{i=1}^{m}\sum_{j=1}^{n}\overline{b_{ij}}a_{ij}$$

所诱导的范数.

F-范数

Frobenius 范数

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

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

$$(\boldsymbol{A} \mid \boldsymbol{B}) = \operatorname{tr}(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}) = \sum_{i=1}^{m}\sum_{j=1}^{n}\overline{b_{ij}}a_{ij}$$

所诱导的范数. 即

$$\|\boldsymbol{A}\|_F^2 = (\boldsymbol{A} \mid \boldsymbol{A}) = \operatorname{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2.$$

Theorem 1.4

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

$$\|\boldsymbol{A}\|_F = \|\boldsymbol{U}\boldsymbol{A}\|_F = \|\boldsymbol{A}\,\boldsymbol{V}\|_F = \|\boldsymbol{U}\boldsymbol{A}\,\boldsymbol{V}\|_F.$$

Theorem 1.4

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

$$\|\boldsymbol{A}\|_F = \|\boldsymbol{U}\boldsymbol{A}\|_F = \|\boldsymbol{A}\,\boldsymbol{V}\|_F = \|\boldsymbol{U}\boldsymbol{A}\,\boldsymbol{V}\|_F.$$

称之为 F-范数的酉不变性.

Theorem 1.4

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

$$\|\boldsymbol{A}\|_F = \|\boldsymbol{U}\boldsymbol{A}\|_F = \|\boldsymbol{A}\,\boldsymbol{V}\|_F = \|\boldsymbol{U}\boldsymbol{A}\,\boldsymbol{V}\|_F.$$

称之为 F-范数的酉不变性.

证: 由 $\|\boldsymbol{A}\|_F = \sqrt{\operatorname{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})}$,

Theorem 1.4

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

$$\|\boldsymbol{A}\|_F = \|\boldsymbol{U}\boldsymbol{A}\|_F = \|\boldsymbol{A}\,\boldsymbol{V}\|_F = \|\boldsymbol{U}\boldsymbol{A}\,\boldsymbol{V}\|_F.$$

称之为 F-范数的酉不变性.

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

$$\|\boldsymbol{U}\boldsymbol{A}\|_F = \sqrt{\operatorname{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{A})}$$

Theorem 1.4

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

$$\|\boldsymbol{A}\|_F = \|\boldsymbol{U}\boldsymbol{A}\|_F = \|\boldsymbol{A}\,\boldsymbol{V}\|_F = \|\boldsymbol{U}\boldsymbol{A}\,\boldsymbol{V}\|_F.$$

称之为 F-范数的酉不变性.

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

$$\|\boldsymbol{U}\boldsymbol{A}\|_F = \sqrt{\operatorname{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{A})} = \sqrt{\operatorname{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})} = \|\boldsymbol{A}\|_F.$$

Theorem 1.4

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

$$\|\boldsymbol{A}\|_F = \|\boldsymbol{U}\boldsymbol{A}\|_F = \|\boldsymbol{A}\,\boldsymbol{V}\|_F = \|\boldsymbol{U}\boldsymbol{A}\,\boldsymbol{V}\|_F.$$

称之为 F-范数的酉不变性.

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

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

其他同理. □

Definition 1.5

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 $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$, $\lambda \in \mathbb{C}^n$ 都满足

1. $\|\boldsymbol{A}\| \geqslant 0$. 当且仅当 $\boldsymbol{A} = \boldsymbol{O}$ 时, $\|\boldsymbol{A}\| = 0$;

Definition 1.5

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 $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$, $\lambda \in \mathbb{C}^n$ 都满足

1. $\|\boldsymbol{A}\| \geqslant 0$. 当且仅当 $\boldsymbol{A} = \boldsymbol{O}$ 时, $\|\boldsymbol{A}\| = 0$;
2. $\|\lambda \boldsymbol{A}\| = |\lambda|\|\boldsymbol{A}\|$;

Definition 1.5

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 $\boldsymbol{A}$, $\boldsymbol{B}\in\mathbb{C}^{n\times n}$, $\lambda\in\mathbb{C}^n$ 都满足

1. $\|\boldsymbol{A}\|\geqslant 0$. 当且仅当 $\boldsymbol{A}=\boldsymbol{O}$ 时, $\|\boldsymbol{A}\|=0$;
2. $\|\lambda\boldsymbol{A}\|=|\lambda|\|\boldsymbol{A}\|$;
3. $\|\boldsymbol{A}+\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|+\|\boldsymbol{B}\|$;

Definition 1.5

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 $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$, $\lambda \in \mathbb{C}^n$ 都满足

1. $\|\boldsymbol{A}\| \geqslant 0$. 当且仅当 $\boldsymbol{A} = \boldsymbol{O}$ 时, $\|\boldsymbol{A}\| = 0$;
2. $\|\lambda\boldsymbol{A}\| = |\lambda|\|\boldsymbol{A}\|$;
3. $\|\boldsymbol{A}+\boldsymbol{B}\| \leqslant \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$;
4. $\|\boldsymbol{A}\boldsymbol{B}\| \leqslant \|\boldsymbol{A}\|\|\boldsymbol{B}\|$.

Definition 1.5

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 $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$, $\lambda \in \mathbb{C}^n$ 都满足

1. $\|\boldsymbol{A}\| \geqslant 0$. 当且仅当 $\boldsymbol{A} = \boldsymbol{O}$ 时, $\|\boldsymbol{A}\| = 0$;
2. $\|\lambda \boldsymbol{A}\| = |\lambda|\|\boldsymbol{A}\|$;
3. $\|\boldsymbol{A} + \boldsymbol{B}\| \leqslant \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$;
4. $\|\boldsymbol{A}\boldsymbol{B}\| \leqslant \|\boldsymbol{A}\|\|\boldsymbol{B}\|$.

则称 $\|\cdot\|$ 为相容的矩阵范数, 或简称为矩阵范数.

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

$$\|\boldsymbol{Ax}\| \leqslant \|\boldsymbol{A}\|\|\boldsymbol{x}\|.$$

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

$$\|\boldsymbol{Ax}\| \leqslant \|\boldsymbol{A}\|\|\boldsymbol{x}\|.$$

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.

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

$$\|\boldsymbol{Ax}\| \leqslant \|\boldsymbol{A}\|\|\boldsymbol{x}\|.$$

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 1.6

如果矩阵范数 $\|\boldsymbol{A}\|_m$ 和向量范数 $\|\boldsymbol{x}\|_v$ 对一切矩阵 $\boldsymbol{A}$ 和向量 $\boldsymbol{x}$ 都满足

$$\|\boldsymbol{Ax}\|_v \leqslant \|\boldsymbol{A}\|_m\|\boldsymbol{x}\|_v,$$

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

Example 1.7

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

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 1.8

设 $\|\boldsymbol{A}\|_m$ 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.

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 1.8

设 $\|\boldsymbol{A}\|_m$ 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 $\boldsymbol{a}\in\mathbb{C}^n$ 是非零向量, 对任意的 $\boldsymbol{x}\in\mathbb{C}^n$, 定义

$$\|\boldsymbol{x}\|_v = \|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m.$$

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 1.8

设 $\|\boldsymbol{A}\|_m$ 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 $\boldsymbol{a}\in\mathbb{C}^n$ 是非零向量, 对任意的 $\boldsymbol{x}\in\mathbb{C}^n$, 定义

$$\|\boldsymbol{x}\|_v = \|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m.$$

先证明 $\|\boldsymbol{x}\|_v$ 是 $\mathbb{C}^n$ 中的向量范数.

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 1.8

设 $\|\boldsymbol{A}\|_m$ 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 $\boldsymbol{a}\in\mathbb{C}^n$ 是非零向量, 对任意的 $\boldsymbol{x}\in\mathbb{C}^n$, 定义

$$\|\boldsymbol{x}\|_v = \|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m.$$

先证明 $\|\boldsymbol{x}\|_v$ 是 $\mathbb{C}^n$ 中的向量范数.

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

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 1.8

设 $\|\boldsymbol{A}\|_m$ 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 $\boldsymbol{a}\in\mathbb{C}^n$ 是非零向量, 对任意的 $\boldsymbol{x}\in\mathbb{C}^n$, 定义

$$\|\boldsymbol{x}\|_v = \|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m.$$

先证明 $\|\boldsymbol{x}\|_v$ 是 $\mathbb{C}^n$ 中的向量范数.

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

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

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 1.8

设 $\|\boldsymbol{A}\|_m$ 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 $\boldsymbol{a}\in\mathbb{C}^n$ 是非零向量, 对任意的 $\boldsymbol{x}\in\mathbb{C}^n$, 定义

$$\|\boldsymbol{x}\|_v = \|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m.$$

先证明 $\|\boldsymbol{x}\|_v$ 是 $\mathbb{C}^n$ 中的向量范数.

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

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

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

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 1.8

设 $\|\boldsymbol{A}\|_m$ 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 $\boldsymbol{a}\in\mathbb{C}^n$ 是非零向量, 对任意的 $\boldsymbol{x}\in\mathbb{C}^n$, 定义

$$\|\boldsymbol{x}\|_v=\|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m.$$

先证明 $\|\boldsymbol{x}\|_v$ 是 $\mathbb{C}^n$ 中的向量范数.

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

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

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

2. 对任意 $\lambda\in\mathbb{C}$, 有

$$\|\lambda\boldsymbol{x}\|_v=\|\lambda\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m$$

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 1.8

设 $\|\boldsymbol{A}\|_m$ 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 $\boldsymbol{a}\in\mathbb{C}^n$ 是非零向量, 对任意的 $\boldsymbol{x}\in\mathbb{C}^n$, 定义

$$\|\boldsymbol{x}\|_v = \|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m.$$

先证明 $\|\boldsymbol{x}\|_v$ 是 $\mathbb{C}^n$ 中的向量范数.

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

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

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

2. 对任意 $\lambda\in\mathbb{C}$, 有

$$\|\lambda\boldsymbol{x}\|_v = \|\lambda\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m = |\lambda|\|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m$$

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 1.8

设 $\|\boldsymbol{A}\|_m$ 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 $\boldsymbol{a}\in\mathbb{C}^n$ 是非零向量, 对任意的 $\boldsymbol{x}\in\mathbb{C}^n$, 定义

$$\|\boldsymbol{x}\|_v=\|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m.$$

先证明 $\|\boldsymbol{x}\|_v$ 是 $\mathbb{C}^n$ 中的向量范数.

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

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

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

2. 对任意 $\lambda\in\mathbb{C}$, 有

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

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

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

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

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

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

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

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

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

又

$$\|\boldsymbol{A}\boldsymbol{x}\|_v = \|\boldsymbol{A}\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m$$

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

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

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

又

$$\|\boldsymbol{A}\boldsymbol{x}\|_v = \|\boldsymbol{A}\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m \leqslant \|\boldsymbol{A}\|_m \|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m$$

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

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

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

又

$$\|\boldsymbol{A}\boldsymbol{x}\|_v = \|\boldsymbol{A}\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m \leqslant \|\boldsymbol{A}\|_m \|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m = \|\boldsymbol{A}\|_m \|\boldsymbol{x}\|_v,$$

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

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

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

又

$$\|\boldsymbol{A}\boldsymbol{x}\|_v = \|\boldsymbol{A}\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m \leqslant \|\boldsymbol{A}\|_m \|\boldsymbol{x}\boldsymbol{a}^{\mathrm{H}}\|_m = \|\boldsymbol{A}\|_m \|\boldsymbol{x}\|_v,$$

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

Theorem 1.9

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

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

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

Example 1.10

设 $\boldsymbol{A}=[a_{ij}]\in\mathbb{C}^{n\times n}$, 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 $\|\boldsymbol{A}\|_1$, $\|\boldsymbol{A}\|_\infty$, $\|\boldsymbol{A}\|_2$, 则有

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

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

$$\|\boldsymbol{A}\|_2=\max_{\|\boldsymbol{x}\|_2=1}\|\boldsymbol{A}\boldsymbol{x}\|_2=\left[\rho(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})\right]^{\frac{1}{2}}=\sigma_1. \qquad \text{(谱范数)}$$

其中 σ_1 为 $\boldsymbol{A}$ 的最大奇异值.

Theorem 1.11

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

Theorem 1.11

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

Theorem 1.12

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

$$K_1\|\boldsymbol{A}\|_a \leqslant \|\boldsymbol{A}\|_b \leqslant K_2\|\boldsymbol{A}\|_a.$$

Example 1.13

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

Example 1.13

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

解: (1) $\|\boldsymbol{A}\|_1=6$.

Example 1.13

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

解: (1) $\|\boldsymbol{A}\|_1 = 6.$

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

Example 1.13

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

解: (1) $\|\boldsymbol{A}\|_1 = 6.$

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

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

Example 1.13

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

解: (1) $\|\boldsymbol{A}\|_1=6$.

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

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

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

Example 1.13

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

解: (1) $\|\boldsymbol{A}\|_1 = 6$.

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

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

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

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

Example 1.13

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

解: (1) $\|\boldsymbol{A}\|_1 = 6$.

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

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

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

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

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

Example 1.13

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

解: (1) $\|\boldsymbol{A}\|_1=6$.

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

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

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

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

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

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

Exercise 1.14 (P.176 习题一 11)

证明:

$$\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$$

Exercise 1.14 (P.176 习题一 11)

证明:

$$\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$$

证: 矩阵 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的特征值非负, 记为 $\lambda_1, \lambda_2, \cdots, \lambda_n$.

Exercise 1.14 (P.176 习题一 11)

证明:

$$\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$$

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

$$\|\boldsymbol{A}\|_2^2 = \rho(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \max\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$$

Exercise 1.14 (P.176 习题一 11)

证明:

$$\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$$

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

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

Exercise 1.14 (P.176 习题一 11)

证明:

$$\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$$

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

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

从而

$$\|\boldsymbol{A}\|_2^2 = \lambda_{\max} \leqslant \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

Exercise 1.14 (P.176 习题一 11)

证明:

$$\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$$

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

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

从而

$$\|\boldsymbol{A}\|_2^2 = \lambda_{\max} \leqslant \lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$$

Exercise 1.14 (P.176 习题一 11)

证明:

$$\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$$

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

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

从而

$$\|\boldsymbol{A}\|_2^2 = \lambda_{\max} \leqslant \lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \|\boldsymbol{A}\|_F^2.$$

Exercise 1.14 (P.176 习题一 11)

证明:

$$\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$$

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

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

从而

$$\|\boldsymbol{A}\|_2^2 = \lambda_{\max} \leqslant \lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \|\boldsymbol{A}\|_F^2.$$

另一方面,

$$\|\boldsymbol{A}\|_2^2 = \lambda_{\max} \geqslant \frac{1}{n}(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$$

Exercise 1.14 (P.176 习题一 11)

证明:

$$\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$$

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

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

从而

$$\|\boldsymbol{A}\|_2^2 = \lambda_{\max} \leqslant \lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \|\boldsymbol{A}\|_F^2.$$

另一方面,

$$\|\boldsymbol{A}\|_2^2 = \lambda_{\max} \geqslant \frac{1}{n}(\lambda_1 + \lambda_2 + \cdots + \lambda_n) = \frac{1}{n}\|\boldsymbol{A}\|_F^2.$$

Exercise 1.14 (P.176 习题一 11)

证明:

$$\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$$

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

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

从而

$$\|\boldsymbol{A}\|_2^2 = \lambda_{\max} \leqslant \lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}) = \|\boldsymbol{A}\|_F^2.$$

另一方面,

$$\|\boldsymbol{A}\|_2^2 = \lambda_{\max} \geqslant \frac{1}{n}(\lambda_1 + \lambda_2 + \cdots + \lambda_n) = \frac{1}{n}\|\boldsymbol{A}\|_F^2.$$

得证 $\frac{1}{\sqrt{n}}\|\boldsymbol{A}\|_F \leqslant \|\boldsymbol{A}\|_2 \leqslant \|\boldsymbol{A}\|_F.$ □

Exercise 1.15 (P.176 习题二 3.)

设 $\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上矩阵的算子范数, $\rho(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的谱半径, 则对任意的矩阵 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, 均有

$$\rho(\boldsymbol{A}) \leqslant \|\boldsymbol{A}\|.$$

Exercise 1.15 (P.176 习题二 3.)

设 $\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上矩阵的算子范数, $\rho(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的谱半径, 则对任意的矩阵 $\boldsymbol{A}\in\mathbb{C}^{n\times n}$, 均有

$$\rho(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|.$$

证: 设 $\|\cdot\|_v$ 是 $\mathbb{C}^n$ 上与矩阵范数 $\|\cdot\|$ 相容的向量范数.

Exercise 1.15 (P.176 习题二 3.)

设 $\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上矩阵的算子范数, $\rho(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的谱半径, 则对任意的矩阵 $\boldsymbol{A}\in\mathbb{C}^{n\times n}$, 均有

$$\rho(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|.$$

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 $\boldsymbol{A}$ 的特征值, $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量, 则 $\boldsymbol{x}\neq\boldsymbol{0}$, 从而 $\|\boldsymbol{x}\|_v>0$.

Exercise 1.15 (P.176 习题二 3.)

设 $\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上矩阵的算子范数, $\rho(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的谱半径, 则对任意的矩阵 $\boldsymbol{A}\in\mathbb{C}^{n\times n}$, 均有

$$\rho(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|.$$

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 $\boldsymbol{A}$ 的特征值, $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量, 则 $\boldsymbol{x}\neq\boldsymbol{0}$, 从而 $\|\boldsymbol{x}\|_v>0$.

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

$$\|\boldsymbol{A}\boldsymbol{x}\|_v=|\lambda|\,\|\boldsymbol{x}\|_v.$$

Exercise 1.15 (P.176 习题二 3.)

设 $\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上矩阵的算子范数, $\rho(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的谱半径, 则对任意的矩阵 $\boldsymbol{A}\in\mathbb{C}^{n\times n}$, 均有

$$\rho(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|.$$

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 $\boldsymbol{A}$ 的特征值, $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量, 则 $\boldsymbol{x}\neq\boldsymbol{0}$, 从而 $\|\boldsymbol{x}\|_v>0$.

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

$$\|\boldsymbol{A}\boldsymbol{x}\|_v=|\lambda|\,\|\boldsymbol{x}\|_v.$$

而 $\|\boldsymbol{A}\boldsymbol{x}\|_v\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{x}\|_v$, 于是

$$|\lambda|\,\|\boldsymbol{x}\|_v\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{x}\|_v.$$

Exercise 1.15 (P.176 习题二 3.)

设 $\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上矩阵的算子范数, $\rho(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的谱半径, 则对任意的矩阵 $\boldsymbol{A}\in\mathbb{C}^{n\times n}$, 均有

$$\rho(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|.$$

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 $\boldsymbol{A}$ 的特征值, $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量, 则 $\boldsymbol{x}\neq\boldsymbol{0}$, 从而 $\|\boldsymbol{x}\|_v>0$.

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

$$\|\boldsymbol{A}\boldsymbol{x}\|_v=|\lambda|\,\|\boldsymbol{x}\|_v.$$

而 $\|\boldsymbol{A}\boldsymbol{x}\|_v\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{x}\|_v$, 于是

$$|\lambda|\,\|\boldsymbol{x}\|_v\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{x}\|_v.$$

又 $\|\boldsymbol{x}\|_v>0$, 故

$$|\lambda|\leqslant\|\boldsymbol{A}\|.$$

Exercise 1.15 (P.176 习题二 3.)

设 $\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上矩阵的算子范数, $\rho(\boldsymbol{A})$ 为 $\boldsymbol{A}$ 的谱半径, 则对任意的矩阵 $\boldsymbol{A}\in\mathbb{C}^{n\times n}$, 均有

$$\rho(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|.$$

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 $\boldsymbol{A}$ 的特征值, $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量, 则 $\boldsymbol{x}\neq\boldsymbol{0}$, 从而 $\|\boldsymbol{x}\|_v>0$.

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

$$\|\boldsymbol{A}\boldsymbol{x}\|_v=|\lambda|\,\|\boldsymbol{x}\|_v.$$

而 $\|\boldsymbol{A}\boldsymbol{x}\|_v\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{x}\|_v$, 于是

$$|\lambda|\,\|\boldsymbol{x}\|_v\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{x}\|_v.$$

又 $\|\boldsymbol{x}\|_v>0$, 故

$$|\lambda|\leqslant\|\boldsymbol{A}\|.$$

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

Outline

1 向量范数及矩阵范数

2 矩阵序列与矩阵级数
- 向量序列的极限
- 矩阵序列的极限
- 矩阵级数

3 矩阵的微分与积分

4 矩阵函数

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.

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 2.1

设 $\boldsymbol{x}^{(k)}, \boldsymbol{x}_0 \in \mathbb{C}^n, k = 1, 2, \cdots,$ 若

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\| \to 0 \qquad (k \to \infty),$$

则称向量序列 $\{\boldsymbol{x}^{(k)}\}$ 收敛于向量 $\boldsymbol{x}_0$,

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 2.1

设 $\boldsymbol{x}^{(k)}, \boldsymbol{x}_0 \in \mathbb{C}^n, k = 1, 2, \cdots,$ 若

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\| \to 0 \qquad (k \to \infty),$$

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

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 2.1

设 $\boldsymbol{x}^{(k)}, \boldsymbol{x}_0 \in \mathbb{C}^n, k = 1, 2, \cdots$, 若

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\| \to 0 \qquad (k \to \infty),$$

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

$$\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}_0,$$

或

$$\boldsymbol{x}^{(k)} \to \boldsymbol{x}_0 \qquad (k \to \infty).$$

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 2.1

设 $\boldsymbol{x}^{(k)}, \boldsymbol{x}_0 \in \mathbb{C}^n$, $k = 1, 2, \cdots$, 若

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\| \to 0 \qquad (k \to \infty),$$

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

$$\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}_0,$$

或

$$\boldsymbol{x}^{(k)} \to \boldsymbol{x}_0 \qquad (k \to \infty).$$

☞ 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.

Example 2.2

设 $\boldsymbol{x}_0 = (1, 1, \cdots, 1)^{\mathrm{T}}$, 且

$$\boldsymbol{x}^{(k)} = \left(1 + \frac{1}{2^k}, 1 + \frac{1}{3^k}, \cdots, 1 + \frac{1}{(n+1)^k}\right)^{\mathrm{T}}, \qquad k = 1, 2, \cdots$$

试证: $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}_0$.

Example 2.2

设 $\boldsymbol{x}_0 = (1, 1, \cdots, 1)^{\mathrm{T}}$, 且

$$\boldsymbol{x}^{(k)} = \left(1 + \frac{1}{2^k}, 1 + \frac{1}{3^k}, \cdots, 1 + \frac{1}{(n+1)^k}\right)^{\mathrm{T}}, \qquad k = 1, 2, \cdots$$

试证: $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}_0$.

证: 由于

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\|_\infty = \max\left\{\frac{1}{2^k}, \frac{1}{3^k}, \cdots, \frac{1}{(n+1)^k}\right\}$$

Example 2.2

设 $\boldsymbol{x}_0 = (1, 1, \cdots, 1)^{\mathrm{T}}$, 且

$$\boldsymbol{x}^{(k)} = \left(1 + \frac{1}{2^k}, 1 + \frac{1}{3^k}, \cdots, 1 + \frac{1}{(n+1)^k}\right)^{\mathrm{T}}, \qquad k = 1, 2, \cdots$$

试证: $\lim\limits_{k \to \infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}_0$.

证: 由于

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\|_\infty = \max\left\{\frac{1}{2^k}, \frac{1}{3^k}, \cdots, \frac{1}{(n+1)^k}\right\} = \frac{1}{2^k}$$

Example 2.2

设 $\boldsymbol{x}_0 = (1, 1, \cdots, 1)^{\mathrm{T}}$, 且

$$\boldsymbol{x}^{(k)} = \left(1 + \frac{1}{2^k}, 1 + \frac{1}{3^k}, \cdots, 1 + \frac{1}{(n+1)^k}\right)^{\mathrm{T}}, \qquad k = 1, 2, \cdots$$

试证: $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}_0$.

证: 由于

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\|_\infty = \max\left\{\frac{1}{2^k}, \frac{1}{3^k}, \cdots, \frac{1}{(n+1)^k}\right\} = \frac{1}{2^k} \to 0 \qquad (k \to \infty),$$

Example 2.2

设 $\boldsymbol{x}_0 = (1, 1, \cdots, 1)^{\mathrm{T}}$, 且

$$\boldsymbol{x}^{(k)} = \left(1 + \frac{1}{2^k}, 1 + \frac{1}{3^k}, \cdots, 1 + \frac{1}{(n+1)^k}\right)^{\mathrm{T}}, \qquad k = 1, 2, \cdots$$

试证: $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}_0$.

证: 由于

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\|_\infty = \max\left\{\frac{1}{2^k}, \frac{1}{3^k}, \cdots, \frac{1}{(n+1)^k}\right\} = \frac{1}{2^k} \to 0 \qquad (k \to \infty),$$

所以 $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}_0$. □

Theorem 2.3

设 $\boldsymbol{x}^{(k)} = \left(\xi_1^{(k)}, \xi_2^{(k)}, \cdots, \xi_n^{(k)}\right)^{\mathrm{T}}$, $\boldsymbol{x}_0 = \left(\xi_1, \xi_2, \cdots, \xi_n\right)^{\mathrm{T}} \in \mathbb{C}^n$, 则向量序列 $\left\{\boldsymbol{x}^{(k)}\right\}$ 收敛于向量 $\boldsymbol{x}_0$ 的充分必要条件是: 每一个坐标分量序列 $\{\xi_i^{(k)}\}$ 收敛于 ξ_i, 即

$$\lim_{k\to\infty} \xi_i^{(k)} = \xi_i, \qquad i = 1, 2, \cdots, n. \tag{7}$$

Theorem 2.3

设 $\boldsymbol{x}^{(k)} = \left(\xi_1^{(k)}, \xi_2^{(k)}, \cdots, \xi_n^{(k)}\right)^{\mathrm{T}}$, $\boldsymbol{x}_0 = \left(\xi_1, \xi_2, \cdots, \xi_n\right)^{\mathrm{T}} \in \mathbb{C}^n$, 则向量序列 $\left\{\boldsymbol{x}^{(k)}\right\}$ 收敛于向量 $\boldsymbol{x}_0$ 的充分必要条件是: 每一个坐标分量序列 $\{\xi_i^{(k)}\}$ 收敛于 ξ_i, 即

$$\lim_{k\to\infty} \xi_i^{(k)} = \xi_i, \qquad i = 1, 2, \cdots, n. \tag{7}$$

证: 必要性. 设 $\|\boldsymbol{x}\| = \|\boldsymbol{x}\|_\infty = \max\limits_{1\leqslant i\leqslant n} |\xi_i|$,

Theorem 2.3

设 $\boldsymbol{x}^{(k)}=\left(\xi_1^{(k)},\xi_2^{(k)},\cdots,\xi_n^{(k)}\right)^{\mathrm{T}}$, $\boldsymbol{x}_0=\left(\xi_1,\xi_2,\cdots,\xi_n\right)^{\mathrm{T}}\in\mathbb{C}^n$, 则向量序列 $\left\{\boldsymbol{x}^{(k)}\right\}$ 收敛于向量 $\boldsymbol{x}_0$ 的充分必要条件是: 每一个坐标分量序列 $\{\xi_i^{(k)}\}$ 收敛于 ξ_i, 即

$$\lim_{k\to\infty}\xi_i^{(k)}=\xi_i,\qquad i=1,2,\cdots,n. \tag{7}$$

证: 必要性. 设 $\|\boldsymbol{x}\|=\|\boldsymbol{x}\|_\infty=\max\limits_{1\leqslant i\leqslant n}|\xi_i|$, 则

$$\|\boldsymbol{x}^{(k)}-\boldsymbol{x}_0\|_\infty\to 0\qquad(k\to\infty).$$

Theorem 2.3

设 $\boldsymbol{x}^{(k)} = \left(\xi_1^{(k)}, \xi_2^{(k)}, \cdots, \xi_n^{(k)}\right)^{\mathrm{T}}$, $\boldsymbol{x}_0 = \left(\xi_1, \xi_2, \cdots, \xi_n\right)^{\mathrm{T}} \in \mathbb{C}^n$, 则向量序列 $\left\{\boldsymbol{x}^{(k)}\right\}$ 收敛于向量 $\boldsymbol{x}_0$ 的充分必要条件是: 每一个坐标分量序列 $\{\xi_i^{(k)}\}$ 收敛于 ξ_i, 即

$$\lim_{k\to\infty} \xi_i^{(k)} = \xi_i, \qquad i = 1, 2, \cdots, n. \tag{7}$$

证: 必要性. 设 $\|\boldsymbol{x}\| = \|\boldsymbol{x}\|_\infty = \max\limits_{1\leqslant i\leqslant n} |\xi_i|$, 则

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\|_\infty \to 0 \qquad (k \to \infty).$$

由于

$$|\xi_i^{(k)} - \xi_i| \leqslant \max_{1\leqslant i\leqslant n} |\xi_i^{(k)} - \xi_i|$$

Theorem 2.3

设 $\boldsymbol{x}^{(k)} = \left(\xi_1^{(k)}, \xi_2^{(k)}, \cdots, \xi_n^{(k)}\right)^{\mathrm{T}}$, $\boldsymbol{x}_0 = \left(\xi_1, \xi_2, \cdots, \xi_n\right)^{\mathrm{T}} \in \mathbb{C}^n$, 则向量序列 $\left\{\boldsymbol{x}^{(k)}\right\}$ 收敛于向量 $\boldsymbol{x}_0$ 的充分必要条件是: 每一个坐标分量序列 $\{\xi_i^{(k)}\}$ 收敛于 ξ_i, 即

$$\lim_{k\to\infty} \xi_i^{(k)} = \xi_i, \qquad i = 1, 2, \cdots, n. \tag{7}$$

证: 必要性. 设 $\|\boldsymbol{x}\| = \|\boldsymbol{x}\|_\infty = \max\limits_{1\leqslant i\leqslant n} |\xi_i|$, 则

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\|_\infty \to 0 \qquad (k \to \infty).$$

由于

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

Theorem 2.3

设 $\boldsymbol{x}^{(k)} = \left(\xi_1^{(k)}, \xi_2^{(k)}, \cdots, \xi_n^{(k)}\right)^{\mathrm{T}}$, $\boldsymbol{x}_0 = \left(\xi_1, \xi_2, \cdots, \xi_n\right)^{\mathrm{T}} \in \mathbb{C}^n$, 则向量序列 $\left\{\boldsymbol{x}^{(k)}\right\}$ 收敛于向量 $\boldsymbol{x}_0$ 的充分必要条件是: 每一个坐标分量序列 $\{\xi_i^{(k)}\}$ 收敛于 ξ_i, 即

$$\lim_{k\to\infty} \xi_i^{(k)} = \xi_i, \qquad i = 1, 2, \cdots, n. \tag{7}$$

证: 必要性. 设 $\|\boldsymbol{x}\| = \|\boldsymbol{x}\|_\infty = \max\limits_{1\leqslant i\leqslant n} |\xi_i|$, 则

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\|_\infty \to 0 \qquad (k \to \infty).$$

由于

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

所以

$$|\xi_i^{(k)} - \xi_i| \to 0 \qquad (k \to \infty).$$

Theorem 2.3

设 $\boldsymbol{x}^{(k)} = \left(\xi_1^{(k)}, \xi_2^{(k)}, \cdots, \xi_n^{(k)}\right)^{\mathrm{T}}$, $\boldsymbol{x}_0 = \left(\xi_1, \xi_2, \cdots, \xi_n\right)^{\mathrm{T}} \in \mathbb{C}^n$, 则向量序列 $\left\{\boldsymbol{x}^{(k)}\right\}$ 收敛于向量 $\boldsymbol{x}_0$ 的充分必要条件是: 每一个坐标分量序列 $\{\xi_i^{(k)}\}$ 收敛于 ξ_i, 即

$$\lim_{k\to\infty} \xi_i^{(k)} = \xi_i, \qquad i = 1, 2, \cdots, n. \tag{7}$$

证: 必要性. 设 $\|\boldsymbol{x}\| = \|\boldsymbol{x}\|_\infty = \max\limits_{1\leqslant i\leqslant n} |\xi_i|$, 则

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}_0\|_\infty \to 0 \qquad (k \to \infty).$$

由于

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

所以

$$|\xi_i^{(k)} - \xi_i| \to 0 \qquad (k \to \infty).$$

即

$$\lim_{k\to\infty} \xi_i^{(k)} = \xi_i, \qquad i = 1, 2, \cdots, n.$$

充分性. 如果 $\lim\limits_{k\to\infty}\xi_i^{(k)}=\xi_i$, $(i=1,2,\cdots,n)$, 则

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

充分性. 如果 $\lim\limits_{k\to\infty}\xi_i^{(k)}=\xi_i$, $(i=1,2,\cdots,n)$, 则

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

即

$$\|\boldsymbol{x}^{(k)}-\boldsymbol{x}_0\|_\infty\to 0\qquad(k\to\infty).$$

充分性. 如果 $\lim\limits_{k\to\infty}\xi_i^{(k)}=\xi_i$, $(i=1,2,\cdots,n)$, 则

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

即

$$\|\boldsymbol{x}^{(k)}-\boldsymbol{x}_0\|_\infty\to 0\qquad(k\to\infty).$$

再有范数的等价性知, 上述结论对任意向量范数均成立. $\square$

Example 2.4

试考察下列向量序列的收敛性:

1. $\boldsymbol{x}^{(k)} = \left(\frac{1}{2^k}, \frac{\sin k}{k}\right)^{\mathrm{T}}$, $k = 1, 2, \cdots$.
2. $\boldsymbol{y}^{(k)} = \left(\sum\limits_{i=1}^{k} \frac{1}{2^i}, \sum\limits_{i=1}^{k} \frac{1}{i}\right)^{\mathrm{T}}$, $k = 1, 2, \cdots$.

Example 2.4

试考察下列向量序列的收敛性:

1. $\boldsymbol{x}^{(k)} = \left(\frac{1}{2^k}, \frac{\sin k}{k}\right)^{\mathrm{T}}, \ k = 1, 2, \cdots.$
2. $\boldsymbol{y}^{(k)} = \left(\sum\limits_{i=1}^{k} \frac{1}{2^i}, \sum\limits_{i=1}^{k} \frac{1}{i}\right)^{\mathrm{T}}, \ k = 1, 2, \cdots.$

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

Example 2.4

试考察下列向量序列的收敛性:

1. $\boldsymbol{x}^{(k)} = \left(\frac{1}{2^k}, \frac{\sin k}{k}\right)^{\mathrm{T}}, \ k = 1,\ 2,\ \cdots.$
2. $\boldsymbol{y}^{(k)} = \left(\sum\limits_{i=1}^{k} \frac{1}{2^i}, \sum\limits_{i=1}^{k} \frac{1}{i}\right)^{\mathrm{T}}, \ k = 1,\ 2,\ \cdots.$

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

$$\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \left(\lim_{k\to\infty} \frac{1}{2^k}, \lim_{k\to\infty} \frac{\sin k}{k}\right)^{\mathrm{T}}$$

Example 2.4

试考察下列向量序列的收敛性:

1. $\boldsymbol{x}^{(k)} = \left(\frac{1}{2^k}, \frac{\sin k}{k}\right)^{\mathrm{T}}, \ k = 1, 2, \cdots.$
2. $\boldsymbol{y}^{(k)} = \left(\sum\limits_{i=1}^{k} \frac{1}{2^i}, \sum\limits_{i=1}^{k} \frac{1}{i}\right)^{\mathrm{T}}, \ k = 1, 2, \cdots.$

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

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

Example 2.4

试考察下列向量序列的收敛性:

1. $\boldsymbol{x}^{(k)} = \left(\frac{1}{2^k}, \frac{\sin k}{k}\right)^{\mathrm{T}}$, $k = 1, 2, \cdots$.
2. $\boldsymbol{y}^{(k)} = \left(\sum\limits_{i=1}^{k} \frac{1}{2^i}, \sum\limits_{i=1}^{k} \frac{1}{i}\right)^{\mathrm{T}}$, $k = 1, 2, \cdots$.

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

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

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

Example 2.4

试考察下列向量序列的收敛性:

1. $\boldsymbol{x}^{(k)} = \left(\frac{1}{2^k}, \frac{\sin k}{k}\right)^{\mathrm{T}}$, $k = 1,\, 2,\, \cdots$.
2. $\boldsymbol{y}^{(k)} = \left(\sum\limits_{i=1}^{k} \frac{1}{2^i}, \sum\limits_{i=1}^{k} \frac{1}{i}\right)^{\mathrm{T}}$, $k = 1,\, 2,\, \cdots$.

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

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

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

(2) 当 $k \to \infty$ 时,

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

为调和级数, 是发散的.

Example 2.4

试考察下列向量序列的收敛性:

1. $\boldsymbol{x}^{(k)} = \left(\frac{1}{2^k}, \frac{\sin k}{k}\right)^{\mathrm{T}}, k = 1, 2, \cdots.$
2. $\boldsymbol{y}^{(k)} = \left(\sum\limits_{i=1}^{k} \frac{1}{2^i}, \sum\limits_{i=1}^{k} \frac{1}{i}\right)^{\mathrm{T}}, k = 1, 2, \cdots.$

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

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

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

(2) 当 $k \to \infty$ 时,

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

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

Outline

1 向量范数及矩阵范数

2 矩阵序列与矩阵级数
- 向量序列的极限
- 矩阵序列的极限
- 矩阵级数

3 矩阵的微分与积分

4 矩阵函数

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

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

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

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

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

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

Definition 2.5

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

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

Definition 2.5

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

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

则称矩阵序列 $\{\boldsymbol{A}_k\}$ 收敛于矩阵

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

Definition 2.5

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

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

则称矩阵序列 $\{\boldsymbol{A}_k\}$ 收敛于矩阵

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

并称 $\boldsymbol{A}$ 是序列 $\{\boldsymbol{A}_k\}$ 在 $k\to\infty$ 时的极限,

Definition 2.5

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

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

则称矩阵序列 $\{\boldsymbol{A}_k\}$ 收敛于矩阵

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

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

$$\lim_{k\to\infty} \boldsymbol{A}_k = \boldsymbol{A}.$$

Example 2.6

已知

$$\boldsymbol{A}_k = \begin{bmatrix} \dfrac{1}{k} & \dfrac{2k^2-1}{3k^2+4} \\ \left(1+\dfrac{1}{k}\right)^k & \cos\dfrac{1}{k^3} \end{bmatrix},$$

求 $\{\boldsymbol{A}_k\}$ 在 $k \to \infty$ 时的极限.

Example 2.6

已知

$$A_k = \begin{bmatrix} \dfrac{1}{k} & \dfrac{2k^2-1}{3k^2+4} \\ (1+\dfrac{1}{k})^k & \cos\dfrac{1}{k^3} \end{bmatrix},$$

求 $\{A_k\}$ 在 $k \to \infty$ 时的极限.

解:

$$\lim_{k\to\infty} A_k = \begin{bmatrix} \lim\limits_{k\to\infty} \dfrac{1}{k} & \lim\limits_{k\to\infty} \dfrac{2k^2-1}{3k^2+4} \\ \lim\limits_{k\to\infty} (1+\dfrac{1}{k})^k & \lim\limits_{k\to\infty} \cos\dfrac{1}{k^3} \end{bmatrix}$$

Example 2.6

已知

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

求 $\{\boldsymbol{A}_k\}$ 在 $k \to \infty$ 时的极限.

解:

$$\lim_{k\to\infty} \boldsymbol{A}_k = \begin{bmatrix} \lim\limits_{k\to\infty} \dfrac{1}{k} & \lim\limits_{k\to\infty} \dfrac{2k^2-1}{3k^2+4} \\ \lim\limits_{k\to\infty} (1+\dfrac{1}{k})^k & \lim\limits_{k\to\infty} \cos\dfrac{1}{k^3} \end{bmatrix} = \begin{bmatrix} 0 & \dfrac{2}{3} \\ \mathrm{e} & 1 \end{bmatrix}. \quad \square$$

Theorem 2.7

设有矩阵序列 $\boldsymbol{A}_k \in \mathbb{C}^{m\times n}$ 和矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$, 则 $\lim\limits_{k\to\infty} \boldsymbol{A}_k = \boldsymbol{A}$ 的充要条件是

$$\lim_{k\to\infty} \|\boldsymbol{A}_k - \boldsymbol{A}\| = 0,$$

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

Theorem 2.8

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

$$\|\boldsymbol{A}_k\| \leqslant M.$$

Theorem 2.9

若 $\boldsymbol{A}_k, \boldsymbol{B}_k, \boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{m\times n}$, $\alpha_k, \beta_k, \alpha, \beta \in \mathbb{C}$, 且有 $\lim\limits_{k\to\infty} \boldsymbol{A}_k = \boldsymbol{A}$, $\lim\limits_{k\to\infty} \boldsymbol{B}_k = \boldsymbol{B}$, $\lim\limits_{k\to\infty} \alpha_k = \alpha$, $\lim\limits_{k\to\infty} \beta_k = \beta$, 则

1. $\lim\limits_{k\to\infty} \left(\alpha_k \boldsymbol{A}_k + \beta_k \boldsymbol{B}_k\right) = \alpha \boldsymbol{A} + \beta \boldsymbol{B}$;

Theorem 2.9

若 $\boldsymbol{A}_k$, $\boldsymbol{B}_k$, $\boldsymbol{A}$, $\boldsymbol{B} \in \mathbb{C}^{m\times n}$, α_k, β_k, α, $\beta \in \mathbb{C}$, 且有 $\lim\limits_{k\to\infty} \boldsymbol{A}_k = \boldsymbol{A}$, $\lim\limits_{k\to\infty} \boldsymbol{B}_k = \boldsymbol{B}$, $\lim\limits_{k\to\infty} \alpha_k = \alpha$, $\lim\limits_{k\to\infty} \beta_k = \beta$, 则

1. $\lim\limits_{k\to\infty} \left(\alpha_k \boldsymbol{A}_k + \beta_k \boldsymbol{B}_k\right) = \alpha \boldsymbol{A} + \beta \boldsymbol{B}$;
2. $\lim\limits_{k\to\infty} \left(\boldsymbol{A}_k \boldsymbol{B}_k\right) = \boldsymbol{A}\boldsymbol{B}$;

Theorem 2.9

若 $\boldsymbol{A}_k$, $\boldsymbol{B}_k$, $\boldsymbol{A}$, $\boldsymbol{B} \in \mathbb{C}^{m\times n}$, α_k, β_k, α, $\beta \in \mathbb{C}$, 且有 $\lim\limits_{k\to\infty} \boldsymbol{A}_k = \boldsymbol{A}$, $\lim\limits_{k\to\infty} \boldsymbol{B}_k = \boldsymbol{B}$, $\lim\limits_{k\to\infty} \alpha_k = \alpha$, $\lim\limits_{k\to\infty} \beta_k = \beta$, 则

1. $\lim\limits_{k\to\infty} \left(\alpha_k \boldsymbol{A}_k + \beta_k \boldsymbol{B}_k\right) = \alpha\boldsymbol{A} + \beta\boldsymbol{B}$;
2. $\lim\limits_{k\to\infty} \left(\boldsymbol{A}_k \boldsymbol{B}_k\right) = \boldsymbol{A}\boldsymbol{B}$;
3. $\lim\limits_{k\to\infty} \boldsymbol{A}_k^{-1} = \boldsymbol{A}^{-1}$, 若 $\boldsymbol{A}_k^{-1}$, $\boldsymbol{A}^{-1}$ 存在.

Outline

1. 向量范数及矩阵范数
2. 矩阵序列与矩阵级数
   - 向量序列的极限
   - 矩阵序列的极限
   - 矩阵级数
3. 矩阵的微分与积分
4. 矩阵函数

Definition 2.10

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

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

为矩阵级数, 其中 $\boldsymbol{A}_k$ 为矩阵级数的一般项.

Definition 2.11

称 $\boldsymbol{S}_m = \sum\limits_{k=1}^{m} \boldsymbol{A}_k$ 为矩阵级数 $\sum\limits_{k=1}^{\infty} \boldsymbol{A}_k$ 的部分和.

Definition 2.11

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

$$\lim_{m\to\infty} \boldsymbol{S}_m = \boldsymbol{S},$$

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

Definition 2.11

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

$$\lim_{m\to\infty} \boldsymbol{S}_m = \boldsymbol{S},$$

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

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

Definition 2.11

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

$$\lim_{m\to\infty} \boldsymbol{S}_m = \boldsymbol{S},$$

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

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

不收敛的级数称为是发散的.

Definition 2.11

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

$$\lim_{m\to\infty} \boldsymbol{S}_m = \boldsymbol{S},$$

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

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

不收敛的级数称为是发散的.

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

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

都收敛.

Example 2.12

已知 $\boldsymbol{A}_k = \begin{bmatrix} \frac{1}{2^k} & \frac{\pi}{4^k} \\ 0 & \frac{1}{(k+1)(k+2)} \end{bmatrix}$, $k = 0, 1, 2, \cdots$. 问矩阵级数 $\sum\limits_{k=0}^{\infty} \boldsymbol{A}_k$ 是否收敛? 若收敛, 求其和.

Example 2.12

已知 $\boldsymbol{A}_k = \begin{bmatrix} \frac{1}{2^k} & \frac{\pi}{4^k} \\ 0 & \frac{1}{(k+1)(k+2)} \end{bmatrix}$, $k = 0, 1, 2, \cdots$. 问矩阵级数 $\sum\limits_{k=0}^{\infty} \boldsymbol{A}_k$ 是否收敛? 若收敛, 求其和.

解: 因为

$$\sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{1-\frac{1}{2}} = 2,$$

Example 2.12

已知 $\boldsymbol{A}_k = \begin{bmatrix} \frac{1}{2^k} & \frac{\pi}{4^k} \\ 0 & \frac{1}{(k+1)(k+2)} \end{bmatrix}$, $k = 0, 1, 2, \cdots$. 问矩阵级数 $\sum_{k=0}^{\infty} \boldsymbol{A}_k$ 是否收敛? 若收敛, 求其和.

解: 因为

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

Example 2.12

已知 $\boldsymbol{A}_k = \begin{bmatrix} \frac{1}{2^k} & \frac{\pi}{4^k} \\ 0 & \frac{1}{(k+1)(k+2)} \end{bmatrix}$, $k = 0, 1, 2, \cdots$. 问矩阵级数 $\sum\limits_{k=0}^{\infty} \boldsymbol{A}_k$ 是否收敛? 若收敛, 求其和.

解: 因为

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

又

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

即

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

即

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

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

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

1. $\sum\limits_{k=1}^{\infty} (\boldsymbol{A}_k + \boldsymbol{B}_k) = \boldsymbol{S} + \boldsymbol{T}$;

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

1. $\sum\limits_{k=1}^{\infty} (\boldsymbol{A}_k + \boldsymbol{B}_k) = \boldsymbol{S} + \boldsymbol{T}$;
2. $\sum\limits_{k=1}^{\infty} \alpha \boldsymbol{A}_k = \alpha \boldsymbol{S}$, $\forall \alpha \in \mathbb{C}$;

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

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

Definition 2.13

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

Definition 2.13

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

Theorem 2.14

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

Example 2.15

设 $\boldsymbol{A}=\begin{bmatrix}0.2 & 0.5 & 0.1\\ 0.1 & 0.5 & 0.3\\ 0.2 & 0.4 & 0.2\end{bmatrix}$, 试证矩阵级数 $\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 绝对收敛.

Example 2.15

设 $\boldsymbol{A}=\begin{bmatrix}0.2 & 0.5 & 0.1\\ 0.1 & 0.5 & 0.3\\ 0.2 & 0.4 & 0.2\end{bmatrix}$, 试证矩阵级数 $\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 绝对收敛.

证: 取矩阵的范数为 $\|\cdot\|_\infty$.

Example 2.15

设 $\boldsymbol{A} = \begin{bmatrix} 0.2 & 0.5 & 0.1 \\ 0.1 & 0.5 & 0.3 \\ 0.2 & 0.4 & 0.2 \end{bmatrix}$, 试证矩阵级数 $\boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}^2 + \cdots + \boldsymbol{A}^k + \cdots$ 绝对收敛.

证: 取矩阵的范数为 $\|\cdot\|_\infty$. 因

$$\|\boldsymbol{A}\| = \|\boldsymbol{A}\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$$

Example 2.15

设 $\boldsymbol{A} = \begin{bmatrix} 0.2 & 0.5 & 0.1 \\ 0.1 & 0.5 & 0.3 \\ 0.2 & 0.4 & 0.2 \end{bmatrix}$, 试证矩阵级数 $\boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A}^2 + \cdots + \boldsymbol{A}^k + \cdots$ 绝对收敛.

证: 取矩阵的范数为 $\|\cdot\|_\infty$. 因

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

Example 2.15

设 $\boldsymbol{A}=\begin{bmatrix}0.2 & 0.5 & 0.1\\ 0.1 & 0.5 & 0.3\\ 0.2 & 0.4 & 0.2\end{bmatrix}$, 试证矩阵级数 $\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 绝对收敛.

证: 取矩阵的范数为 $\|\cdot\|_\infty$. 因

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

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

Example 2.15

设 $\boldsymbol{A}=\begin{bmatrix}0.2 & 0.5 & 0.1\\ 0.1 & 0.5 & 0.3\\ 0.2 & 0.4 & 0.2\end{bmatrix}$, 试证矩阵级数 $\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 绝对收敛.

证: 取矩阵的范数为 $\|\cdot\|_\infty$. 因

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

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

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

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

Example 2.15

设 $\boldsymbol{A}=\begin{bmatrix}0.2 & 0.5 & 0.1\\0.1 & 0.5 & 0.3\\0.2 & 0.4 & 0.2\end{bmatrix}$, 试证矩阵级数 $\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 绝对收敛.

证: 取矩阵的范数为 $\|\cdot\|_\infty$. 因

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

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

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

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

又 $\|\boldsymbol{A}^k\|\leqslant\|\boldsymbol{A}\|^k$,

Example 2.15

设 $\boldsymbol{A}=\begin{bmatrix}0.2 & 0.5 & 0.1\\0.1 & 0.5 & 0.3\\0.2 & 0.4 & 0.2\end{bmatrix}$, 试证矩阵级数 $\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 绝对收敛.

证: 取矩阵的范数为 $\|\cdot\|_\infty$. 因

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

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

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

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

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

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

收敛,

Example 2.15

设 $\boldsymbol{A}=\begin{bmatrix}0.2 & 0.5 & 0.1\\ 0.1 & 0.5 & 0.3\\ 0.2 & 0.4 & 0.2\end{bmatrix}$, 试证矩阵级数 $\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 绝对收敛.

证: 取矩阵的范数为 $\|\cdot\|_\infty$. 因

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

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

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

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

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

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

收敛, 故原级数绝对收敛. □

Example 2.16

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, 则矩阵级数

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

绝对收敛.

Example 2.16

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, 则矩阵级数

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

绝对收敛.

证: 因为 $\left\|\dfrac{\boldsymbol{A}^k}{k!}\right\| = \dfrac{\|\boldsymbol{A}^k\|}{k!} \leqslant \dfrac{\|\boldsymbol{A}\|^k}{k!}$,

Example 2.16

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, 则矩阵级数

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

绝对收敛.

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

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

Example 2.16

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, 则矩阵级数

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

绝对收敛.

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

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

由正项级数比较法知, 级数 $\sum\limits_{k=1}^{\infty} \left\| \frac{\boldsymbol{A}^k}{k!} \right\|$ 收敛,

Example 2.16

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, 则矩阵级数

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

绝对收敛.

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

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

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

Example 2.16

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, 则矩阵级数

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

绝对收敛.

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

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

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

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

Example 2.16

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, 则矩阵级数

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

绝对收敛.

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

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

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

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

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

同理有

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

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

同理有

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

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

从而欧拉公式

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

对任意的复数或复矩阵都成立.

同理有

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

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

从而欧拉公式

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

对任意的复数或复矩阵都成立. 特别地,

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

Outline

1. 向量范数及矩阵范数
2. 矩阵序列与矩阵级数
3. 矩阵的微分与积分
   - 函数矩阵及其极限
   - 函数矩阵的微分和积分
   - 纯量函数关于矩阵的导数
   - 矩阵对矩阵的导数
4. 矩阵函数

Definition 3.1

以实变量 t 的函数为元素的矩阵

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

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 $\boldsymbol{A}(t)$ 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.

函数矩阵 $\boldsymbol{A}(t)$ 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

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)=\left[\frac{\mathrm{d}}{\mathrm{d}t}a_{ij}(t)\right]_{m\times n};$$

$$\int \boldsymbol{A}(t)\,\mathrm{d}t=\left[\int a_{ij}(t)\,\mathrm{d}t\right]_{m\times n};$$

$$\int_a^b \boldsymbol{A}(t)\,\mathrm{d}t=\left[\int_a^b a_{ij}(t)\,\mathrm{d}t\right]_{m\times n}.$$

Definition 3.2

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

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

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

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

Definition 3.3

如果所有的元素 $a_{ij}(t)$ 在 $t = t_0$ 连续, 即

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

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

Definition 3.3

如果所有的元素 $a_{ij}(t)$ 在 $t = t_0$ 连续, 即

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

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

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

容易验证下列等式是成立的.

设 $\lim\limits_{t\to t_0} \boldsymbol{A}(t) = \boldsymbol{A}$, $\lim\limits_{t\to t_0} \boldsymbol{B}(t) = \boldsymbol{B}$.

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

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

容易验证下列等式是成立的.

设 $\lim\limits_{t\to t_0}\boldsymbol{A}(t)=\boldsymbol{A}$, $\lim\limits_{t\to t_0}\boldsymbol{B}(t)=\boldsymbol{B}$.

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

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

2. 设 k 为常数, 则

$$\lim_{t\to t_0}[k\boldsymbol{A}(t)]=k\boldsymbol{A}.$$

容易验证下列等式是成立的.

设 $\lim\limits_{t \to t_0} \boldsymbol{A}(t) = \boldsymbol{A}$, $\lim\limits_{t \to t_0} \boldsymbol{B}(t) = \boldsymbol{B}$.

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

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

2. 设 k 为常数, 则

$$\lim_{t \to t_0} [k\boldsymbol{A}(t)] = k\boldsymbol{A}.$$

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

$$\lim_{t \to t_0} [\boldsymbol{A}(t)\boldsymbol{B}(t)] = \boldsymbol{A}\boldsymbol{B} = \lim_{t \to t_0} \boldsymbol{A}(t) \lim_{t \to t_0} \boldsymbol{B}(t).$$

Outline

1. 向量范数及矩阵范数
2. 矩阵序列与矩阵级数
3. 矩阵的微分与积分
   - 函数矩阵及其极限
   - 函数矩阵的微分和积分
   - 纯量函数关于矩阵的导数
   - 矩阵对矩阵的导数
4. 矩阵函数

Definition 3.4

若函数矩阵 $\boldsymbol{A}(t)$ 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 $\boldsymbol{A}(t)$ 在 t_0 处 (或在区间 (a, b) 上)可微.

Definition 3.4

若函数矩阵 $\boldsymbol{A}(t)$ 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 $\boldsymbol{A}(t)$ 在 t_0 处 (或在区间 (a,b) 上)可微. 并记为

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

Example 3.5

已知 $\boldsymbol{A}(t) = \begin{bmatrix} \sin t & 2t^3 \\ 2\sqrt{t} & \mathrm{e}^{2t} \end{bmatrix}$, 求 $\dfrac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}$.

Example 3.5

已知 $\boldsymbol{A}(t) = \begin{bmatrix} \sin t & 2t^3 \\ 2\sqrt{t} & \mathrm{e}^{2t} \end{bmatrix}$, 求 $\dfrac{\mathrm{d}\boldsymbol{A}(t)}{\mathrm{d}t}$.

解:

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

关于函数矩阵, 有下面的求导法则:

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

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

关于函数矩阵, 有下面的求导法则:

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

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

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

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

关于函数矩阵, 有下面的求导法则:

1. 若 $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$ 都是 $m \times n$ 阶矩阵可微矩阵, 则
$$\frac{\mathrm{d}}{\mathrm{d}t}[\boldsymbol{A}(t)+\boldsymbol{B}(t)] = \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t) + \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{B}(t).$$
2. 若 $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$ 分别是 $m \times n$ 阶及 $n \times r$ 阶矩阵, 则
$$\frac{\mathrm{d}}{\mathrm{d}t}[\boldsymbol{A}(t)\boldsymbol{B}(t)] = \left[\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)\right]\boldsymbol{B}(t) + \boldsymbol{A}(t)\left[\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{B}(t)\right].$$
3. 若 $\boldsymbol{A}(u)$ 可微, 且 $u = f(t)$ 关于 t 可微, 则
$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}\big(f(t)\big) = f'(t)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(u).$$

关于函数矩阵, 有下面的求导法则:

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

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

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

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

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}\big(f(t)\big) = f'(t)\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(u).$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\boldsymbol{A}^{-1}(t)\big) = -\boldsymbol{A}^{-1}(t)\left[\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)\right]\boldsymbol{A}^{-1}(t).$$

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

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

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

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

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

即

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

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

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

即

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

两边左乘以 $\boldsymbol{A}^{-1}(t)$, 得

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

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

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

即

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

两边左乘以 $\boldsymbol{A}^{-1}(t)$, 得

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

☞ 注意矩阵乘法不满足交换律. 例如

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

Definition 3.6

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

Definition 3.6

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

$$\int_a^b \boldsymbol{A}(t)\,\mathrm{d}t = \Big[\int_a^b a_{ij}(t)\,\mathrm{d}t\Big]_{m\times n}.$$

Theorem 3.7

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[\int_a^t \boldsymbol{A}(\tau)\,\mathrm{d}\tau\Big] = \boldsymbol{A}(t).$$

Theorem 3.7

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[\int_a^t \boldsymbol{A}(\tau)\,\mathrm{d}\tau\Big] = \boldsymbol{A}(t).$$

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

$$\int_a^t \Big[\frac{\mathrm{d}}{\mathrm{d}s}\boldsymbol{A}(s)\Big]\,\mathrm{d}s = \boldsymbol{A}(t) - \boldsymbol{A}(a), \qquad t \in [a, b].$$

Outline

1. 向量范数及矩阵范数

2. 矩阵序列与矩阵级数

3. 矩阵的微分与积分
   - 函数矩阵及其极限
   - 函数矩阵的微分和积分
   - 纯量函数关于矩阵的导数
   - 矩阵对矩阵的导数

4. 矩阵函数

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.

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

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

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

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

下面我们将这一概念推广到一般情形.

Definition 3.8

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

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

Definition 3.8

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

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

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

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

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

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{X}} = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \vdots \\ \dfrac{\partial f}{\partial x_m} \end{bmatrix}.$$

此即梯度.

Outline

1. 向量范数及矩阵范数

2. 矩阵序列与矩阵级数

3. 矩阵的微分与积分
   - 函数矩阵及其极限
   - 函数矩阵的微分和积分
   - 纯量函数关于矩阵的导数
   - 矩阵对矩阵的导数

4. 矩阵函数

Definition 3.9

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

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

其中

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

Example 3.10

设 $\boldsymbol{X} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$, 求向量 $\boldsymbol{X}^{\mathrm{T}} = (x_1, x_2, \cdots, x_n)$ 对向量 $\boldsymbol{X}$ 的导数.

Example 3.10

设 $\boldsymbol{X}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, 求向量 $\boldsymbol{X}^{\mathrm{T}}=(x_1,x_2,\cdots,x_n)$ 对向量 $\boldsymbol{X}$ 的导数.

解:

$$\frac{\mathrm{d}\boldsymbol{X}^{\mathrm{T}}}{\mathrm{d}\boldsymbol{X}}=\begin{bmatrix}\dfrac{\partial \boldsymbol{X}^{\mathrm{T}}}{\partial x_1}\\ \dfrac{\partial \boldsymbol{X}^{\mathrm{T}}}{\partial x_2}\\ \vdots\\ \dfrac{\partial \boldsymbol{X}^{\mathrm{T}}}{\partial x_n}\end{bmatrix}$$

Example 3.10

设 $\boldsymbol{X}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, 求向量 $\boldsymbol{X}^{\mathrm{T}}=(x_1,x_2,\cdots,x_n)$ 对向量 $\boldsymbol{X}$ 的导数.

解:

$$\frac{\mathrm{d}\boldsymbol{X}^{\mathrm{T}}}{\mathrm{d}\boldsymbol{X}}=\begin{bmatrix}\dfrac{\partial \boldsymbol{X}^{\mathrm{T}}}{\partial x_1}\\ \dfrac{\partial \boldsymbol{X}^{\mathrm{T}}}{\partial x_2}\\ \vdots\\ \dfrac{\partial \boldsymbol{X}^{\mathrm{T}}}{\partial x_n}\end{bmatrix}=\begin{bmatrix}1 & 0 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & 1\end{bmatrix}$$

Example 3.10

设 $\boldsymbol{X}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, 求向量 $\boldsymbol{X}^{\mathrm{T}}=(x_1,x_2,\cdots,x_n)$ 对向量 $\boldsymbol{X}$ 的导数.

解:

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

Outline

1. 向量范数及矩阵范数
2. 矩阵序列与矩阵级数
3. 矩阵的微分与积分
4. 矩阵函数
   - 矩阵多项式
   - 矩阵函数

Theorem 4.1

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

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_t \end{bmatrix} = \mathrm{diag}(\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_t),$$

则

$$f(\boldsymbol{A}) = \mathrm{diag}[f(\boldsymbol{A}_1), f(\boldsymbol{A}_2), \cdots, f(\boldsymbol{A}_t)].$$

证:

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

Theorem 4.2

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

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

Theorem 4.2

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

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

证: 由

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

Theorem 4.2

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

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

证: 由

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

得

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

矩阵多项式的计算方法

要计算 $f(\boldsymbol{A}) = a_0\boldsymbol{I} + a_1\boldsymbol{A} + \cdots + a_m\boldsymbol{A}^m$,

矩阵多项式的计算方法

要计算 $f(\boldsymbol{A}) = a_0\boldsymbol{I} + a_1\boldsymbol{A} + \cdots + a_m\boldsymbol{A}^m$, 可以找一个 $\boldsymbol{A}$ 的零化多项式 $\varphi(\lambda)$, 且 $\partial\big(\varphi(\lambda)\big) < m$,

矩阵多项式的计算方法

要计算 $f(\boldsymbol{A}) = a_0\boldsymbol{I} + a_1\boldsymbol{A} + \cdots + a_m\boldsymbol{A}^m$, 可以找一个 $\boldsymbol{A}$ 的零化多项式 $\varphi(\lambda)$, 且 $\partial\big(\varphi(\lambda)\big) < m$, 再由多项式带余除法有

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

则

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

要计算 $f(\boldsymbol{A}) = a_0\boldsymbol{I} + a_1\boldsymbol{A} + \cdots + a_m\boldsymbol{A}^m$, 可以找一个 $\boldsymbol{A}$ 的零化多项式 $\varphi(\lambda)$, 且 $\partial\big(\varphi(\lambda)\big) < m$, 再由多项式带余除法有

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

则

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

又 $\varphi(\boldsymbol{A}) = \boldsymbol{O}$, 故

$$f(\boldsymbol{A}) = r(\boldsymbol{A}).$$

矩阵多项式的计算方法

要计算 $f(\boldsymbol{A}) = a_0\boldsymbol{I} + a_1\boldsymbol{A} + \cdots + a_m\boldsymbol{A}^m$, 可以找一个 $\boldsymbol{A}$ 的零化多项式 $\varphi(\lambda)$, 且 $\partial\big(\varphi(\lambda)\big) < m$, 再由多项式带余除法有

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

则

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

又 $\varphi(\boldsymbol{A}) = \boldsymbol{O}$, 故

$$f(\boldsymbol{A}) = r(\boldsymbol{A}).$$

从而把计算 $f(\boldsymbol{A})$ 转化为计算次数较低的 $r(\boldsymbol{A})$.

矩阵多项式的计算方法

要计算 $f(\boldsymbol{A}) = a_0\boldsymbol{I} + a_1\boldsymbol{A} + \cdots + a_m\boldsymbol{A}^m$, 可以找一个 $\boldsymbol{A}$ 的零化多项式 $\varphi(\lambda)$, 且 $\partial\big(\varphi(\lambda)\big) < m$, 再由多项式带余除法有

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

则

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

又 $\varphi(\boldsymbol{A}) = \boldsymbol{O}$, 故

$$f(\boldsymbol{A}) = r(\boldsymbol{A}).$$

从而把计算 $f(\boldsymbol{A})$ 转化为计算次数较低的 $r(\boldsymbol{A})$.

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.

Outline

1. 向量范数及矩阵范数
2. 矩阵序列与矩阵级数
3. 矩阵的微分与积分
4. 矩阵函数
   - 矩阵多项式
   - 矩阵函数

Lemma 4.3

设 r 阶方阵为 $\boldsymbol{H} = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$,

Lemma 4.3

设 r 阶方阵为 $\boldsymbol{H} = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$, 则 *(1)* 当 $m \geqslant r$ 时, $\boldsymbol{H}^m = \boldsymbol{O}$;

Lemma 4.3

设 r 阶方阵为 $\boldsymbol{H}=\begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$, 则 (1) 当 $m \geqslant r$ 时, $\boldsymbol{H}^m=\boldsymbol{O}$; (2) 当 $m<r$ 时,

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

即 $\boldsymbol{H}$ 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

$$
\boldsymbol{H}^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ & 0 & 0 & 1 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix},
$$

例如

$$\boldsymbol{H}^2=\begin{bmatrix}0&1&0&0\\&0&1&0\\&&0&1\\&&&0\end{bmatrix}\begin{bmatrix}0&1&0&0\\&0&1&0\\&&0&1\\&&&0\end{bmatrix}=\begin{bmatrix}0&0&1&0\\&0&0&1\\&&0&0\\&&&0\end{bmatrix},$$

$$\boldsymbol{H}^3=\begin{bmatrix}0&0&1&0\\&0&0&1\\&&0&0\\&&&0\end{bmatrix}\begin{bmatrix}0&1&0&0\\&0&1&0\\&&0&1\\&&&0\end{bmatrix}=\begin{bmatrix}0&0&0&1\\&0&0&0\\&&0&0\\&&&0\end{bmatrix},$$

例如

$$\boldsymbol{H}^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ & 0 & 0 & 1 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix},$$

$$\boldsymbol{H}^3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ & 0 & 0 & 1 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix},$$

$$\boldsymbol{H}^4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix},$$

例如

$$
\boldsymbol{H}^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ & 0 & 0 & 1 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix},
$$

$$
\boldsymbol{H}^3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ & 0 & 0 & 1 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix},
$$

$$
\boldsymbol{H}^4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{bmatrix},
$$

$\boldsymbol{H}^m = \boldsymbol{O},\ m \geqslant 4.$

由引理 4.3 可知,

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

Lemma 4.4

若 $f(z)=\sum\limits_{m=0}^{\infty} a_m z^m$, 则

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

Lemma 4.4

若 $f(z)=\sum\limits_{m=0}^{\infty} a_m z^m$，则

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

证: 因为

$$\mathrm{C}_m^s a_m \lambda^{m-s}=\frac{m!}{s!(m-s)!}a_m \lambda^{m-s}$$

Lemma 4.4

若 $f(z)=\sum\limits_{m=0}^{\infty}a_m z^m$, 则

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

证: 因为

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

Lemma 4.4

若 $f(z)=\sum\limits_{m=0}^{\infty}a_m z^m$，则

$$\frac{1}{s!}f^{(s)}(z)\bigg|_{z=\lambda}=\sum_{m=s}^{\infty}\mathrm{C}_m^s a_m\lambda^{m-s}.$$

证: 因为

$$\mathrm{C}_m^s a_m\lambda^{m-s}=\frac{m!}{s!(m-s)!}a_m\lambda^{m-s}=\frac{1}{s!}a_m\frac{\mathrm{d}^s}{\mathrm{d}z^s}z^m\bigg|_{z=\lambda},$$

所以

$$\sum_{m=s}^{\infty}\mathrm{C}_m^s a_m\lambda^{m-s}=\frac{1}{s!}\sum_{m=s}^{\infty}a_m\frac{\mathrm{d}^s}{\mathrm{d}z^s}z^m\bigg|_{z=\lambda}$$

Lemma 4.4

若 $f(z)=\sum\limits_{m=0}^{\infty} a_m z^m$, 则

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

证: 因为

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

所以

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

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

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

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

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

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

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

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

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

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

$$\vdots$$

$$f^{(s)}(z) = \sum_{m=s}^{\infty} m(m-1)\cdots(m-s+1) a_m z^{m-s}$$

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

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

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

$$\vdots$$

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

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

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

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

$$\vdots$$

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

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

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

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

$$\vdots$$

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

所以

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

所以

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

故

$$\left.\frac{1}{s!}f^{(s)}(z)\right|_{z=\lambda}=\sum_{m=s}^{\infty}\mathrm{C}_m^s a_m \lambda^{m-s}. \qquad \square$$

Theorem 4.5

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 $\boldsymbol{J}=\begin{bmatrix}\lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda\end{bmatrix}$.

则当 $|\lambda|<R$ 时, 级数 $\sum_{m=0}^{\infty}a_m\boldsymbol{J}^m$ 收敛, 且

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

证: 矩阵 $\boldsymbol{J}$ 可以写成 $\lambda\boldsymbol{I}+\boldsymbol{H}$,

证: 矩阵 $\boldsymbol{J}$ 可以写成 $\lambda\boldsymbol{I}+\boldsymbol{H}$, 因为 $\boldsymbol{I}$ 是单位矩阵, 所以 $\boldsymbol{I}$ 可以与 $\boldsymbol{H}$ 交换.

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

$$\boldsymbol{J}^m = (\lambda\boldsymbol{I}+\boldsymbol{H})^m = \lambda^m\boldsymbol{I}+\mathrm{C}_m^1\lambda^{m-1}\boldsymbol{H}+\cdots+\mathrm{C}_m^{m-1}\lambda\boldsymbol{H}^{m-1}+\boldsymbol{H}^m$$

$$= \begin{bmatrix} \lambda^m & \mathrm{C}_m^1\lambda^{m-1} & \cdots & \mathrm{C}_m^{r-1}\lambda^{m-r+1} \\ & \ddots & \ddots & \vdots \\ & & \lambda^m & \mathrm{C}_m^1\lambda^{m-1} \\ & & & \lambda^m \end{bmatrix}. \qquad \text{(对照公式 (10))}$$

所以

$$\sum_{m=0}^{\infty} a_m \boldsymbol{J}^m = \begin{bmatrix} \sum\limits_{m=0}^{\infty} a_m \lambda^m & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^1 \lambda^{m-1} & \cdots & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^{r-1} \lambda^{m-r+1} \\ & \sum\limits_{m=0}^{\infty} a_m \lambda^m & \cdots & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^{r-2} \lambda^{m-r+2} \\ & & \ddots & \vdots \\ & & & \sum\limits_{m=0}^{\infty} a_m \lambda^m \end{bmatrix}$$

所以

$$\sum_{m=0}^{\infty} a_m \boldsymbol{J}^m = \begin{bmatrix} \sum\limits_{m=0}^{\infty} a_m\lambda^m & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^1 \lambda^{m-1} & \cdots & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^{r-1} \lambda^{m-r+1} \\ & \sum\limits_{m=0}^{\infty} a_m\lambda^m & \cdots & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^{r-2} \lambda^{m-r+2} \\ & & \ddots & \vdots \\ & & & \sum\limits_{m=0}^{\infty} a_m\lambda^m \end{bmatrix}$$

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

所以

$$\sum_{m=0}^{\infty} a_m \boldsymbol{J}^m = \begin{bmatrix} \sum\limits_{m=0}^{\infty} a_m\lambda^m & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^1 \lambda^{m-1} & \cdots & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^{r-1} \lambda^{m-r+1} \\ & \sum\limits_{m=0}^{\infty} a_m\lambda^m & \cdots & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^{r-2} \lambda^{m-r+2} \\ & & \ddots & \vdots \\ & & & \sum\limits_{m=0}^{\infty} a_m\lambda^m \end{bmatrix}$$

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

最后一个等号的成立需要 $|\lambda| < R$.

所以

$$\sum_{m=0}^{\infty} a_m \boldsymbol{J}^m = \begin{bmatrix} \sum\limits_{m=0}^{\infty} a_m\lambda^m & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^1 \lambda^{m-1} & \cdots & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^{r-1} \lambda^{m-r+1} \\ & \sum\limits_{m=0}^{\infty} a_m\lambda^m & \cdots & \sum\limits_{m=0}^{\infty} a_m \mathrm{C}_m^{r-2} \lambda^{m-r+2} \\ & & \ddots & \vdots \\ & & & \sum\limits_{m=0}^{\infty} a_m\lambda^m \end{bmatrix}$$

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

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

Theorem 4.6

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

Theorem 4.6

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

证: 设 $\boldsymbol{A}$ 的若当标准形为

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

Theorem 4.6

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

证: 设 $\boldsymbol{A}$ 的若当标准形为

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

其中 $\boldsymbol{J}_i$ 是特征值为 λ_i 的 r_i 阶若当块.

Theorem 4.6

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

证: 设 $\boldsymbol{A}$ 的若当标准形为

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

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

于是

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

因为 $\rho(\boldsymbol{A}) < R$, 故 $|\lambda_i| < R$, 由定理 4.5 知 $\sum\limits_{m=0}^{\infty} a_m \boldsymbol{J}_i^m$ 收敛,

于是

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

因为 $\rho(\boldsymbol{A}) < R$, 故 $|\lambda_i| < R$, 由定理 4.5 知 $\sum\limits_{m=0}^{\infty} a_m \boldsymbol{J}_i^m$ 收敛, 因而推得 $\sum\limits_{m=0}^{\infty} a_m \boldsymbol{A}^m$ 收敛. □

Definition 4.7

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

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

Definition 4.7

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

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

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

Definition 4.7

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

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

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

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

Definition 4.7

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

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

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

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

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

例如, 对于如下函数的幂级数展开式

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

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

相应地有矩阵函数

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

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

例如, 对于如下函数的幂级数展开式

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

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

相应地有矩阵函数

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

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

以及

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

定理 4.6 的证明过程给出了求 $f(\boldsymbol{A})$ 的一个方法.

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

其中

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

定理 4.6 的证明过程给出了求 $f(\boldsymbol{A})$ 的一个方法.

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

其中

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

这里 r_i 是若当块 $\boldsymbol{J}_i$ 的阶数.

实际计算中, 若当块的阶数多为 1, 2, 3, 即分别为

$$\boldsymbol{J}_i = \lambda_i, \qquad \boldsymbol{J}_i = \begin{bmatrix} \lambda_i & 1 \\ & \lambda_i \end{bmatrix}, \qquad \boldsymbol{J}_i = \begin{bmatrix} \lambda_i & 1 & 0 \\ & \lambda_i & 1 \\ & & \lambda_i \end{bmatrix},$$

实际计算中, 若当块的阶数多为 1, 2, 3, 即分别为

$$J_i = \lambda_i, \qquad J_i = \begin{bmatrix} \lambda_i & 1 \\ & \lambda_i \end{bmatrix}, \qquad J_i = \begin{bmatrix} \lambda_i & 1 & 0 \\ & \lambda_i & 1 \\ & & \lambda_i \end{bmatrix},$$

此时对应地有

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

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

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

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

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

则

$$f(\boldsymbol{A}) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2!}f''(\lambda) \\ & f(\lambda) & f'(\lambda) \\ & & f(\lambda) \end{bmatrix}$$

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

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

则

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

Example 4.8

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

Example 4.8

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

解: 先求 $\boldsymbol{A}$ 的若当标准形.

Example 4.8

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

解: 先求 $\boldsymbol{A}$ 的若当标准形. 因

$$\lambda \boldsymbol{I}-\boldsymbol{A}=\begin{bmatrix} \lambda-2 & 0 & 0 \\ -1 & \lambda-1 & -1 \\ -1 & 1 & \lambda-3 \end{bmatrix}$$

Example 4.8

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

解: 先求 $\boldsymbol{A}$ 的若当标准形. 因

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

Example 4.8

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

解: 先求 $\boldsymbol{A}$ 的若当标准形. 因

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

所以若当标准形为

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

Example 4.8

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

解: 先求 $\boldsymbol{A}$ 的若当标准形. 因

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

所以若当标准形为

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

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

由 $\boldsymbol{AP}=\boldsymbol{PJ}$ 可求得

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

由 $\boldsymbol{AP}=\boldsymbol{PJ}$ 可求得

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

(1) 记 $f(z)=z^{20}$.

由 $\boldsymbol{AP} = \boldsymbol{PJ}$ 可求得

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

(1) 记 $f(z) = z^{20}$. 则

$$f(\boldsymbol{J}_1) = 2^{20},$$

由 $\boldsymbol{AP}=\boldsymbol{PJ}$ 可求得

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

(1) 记 $f(z)=z^{20}$. 则

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

由 $\boldsymbol{AP} = \boldsymbol{PJ}$ 可求得

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

(1) 记 $f(z) = z^{20}$. 则

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

故

$$\boldsymbol{A}^{20} = \boldsymbol{P}\big(f(\boldsymbol{J}_1) \oplus f(\boldsymbol{J}_2)\big)\boldsymbol{P}^{-1}$$

由 $\boldsymbol{AP}=\boldsymbol{PJ}$ 可求得

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

(1) 记 $f(z)=z^{20}$. 则

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

故

$$\boldsymbol{A}^{20}=\boldsymbol{P}\big(f(\boldsymbol{J}_1)\oplus f(\boldsymbol{J}_2)\big)\boldsymbol{P}^{-1}=\begin{bmatrix}1&0&1\\1&1&0\\0&1&0\end{bmatrix}\begin{bmatrix}2^{20}&0&0\\&2^{20}&20\cdot 2^{19}\\&&2^{20}\end{bmatrix}\begin{bmatrix}0&1&-1\\0&0&1\\1&-1&1\end{bmatrix}$$

由 $\boldsymbol{AP}=\boldsymbol{PJ}$ 可求得

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

(1) 记 $f(z)=z^{20}$. 则

$$f(\boldsymbol{J}_1)=2^{20},\qquad f(\boldsymbol{J}_2)=\begin{bmatrix}f(\lambda_i)&f'(\lambda_i)\\&f(\lambda_i)\end{bmatrix}=\begin{bmatrix}2^{20}&20\cdot2^{19}\\&2^{20}\end{bmatrix}.$$

故

$$\boldsymbol{A}^{20}=\boldsymbol{P}\big(f(\boldsymbol{J}_1)\oplus f(\boldsymbol{J}_2)\big)\boldsymbol{P}^{-1}=\begin{bmatrix}1&0&1\\1&1&0\\0&1&0\end{bmatrix}\begin{bmatrix}2^{20}&0&0\\&2^{20}&20\cdot2^{19}\\&&2^{20}\end{bmatrix}\begin{bmatrix}0&1&-1\\0&0&1\\1&-1&1\end{bmatrix}$$

$$=2^{20}\begin{bmatrix}1&0&0\\10&-9&10\\10&-10&11\end{bmatrix}.$$

(2) 记 $f(z) = \mathrm{e}^z$.

(2) 记 $f(z)=\mathrm{e}^z$. 则

$$f(\boldsymbol{J}_1)=\mathrm{e}^2,$$

(2) 记 $f(z)=\mathrm{e}^z$. 则

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

(2) 记 $f(z)=\mathrm{e}^z$. 则

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

故

$$\mathrm{e}^{\boldsymbol{A}}=\boldsymbol{P}f(\boldsymbol{J})\boldsymbol{P}^{-1}$$

(2) 记 $f(z) = \mathrm{e}^z$. 则

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

故

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

(2) 记 $f(z)=\mathrm{e}^z$. 则

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

故

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

(3) 记 $f(z) = \sin z$.

(3) 记 $f(z) = \sin z$. 则

$$f(\boldsymbol{J}_1) = \sin 2,$$

(3) 记 $f(z) = \sin z$. 则

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

(3) 记 $f(z)=\sin z$. 则

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

故

$$\sin \boldsymbol{A}=\boldsymbol{P}f(\boldsymbol{J})\boldsymbol{P}^{-1}$$

(3) 记 $f(z)=\sin z$. 则

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

故

$$\sin \boldsymbol{A}=\boldsymbol{P}f(\boldsymbol{J})\boldsymbol{P}^{-1}=\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} \sin 2 & 0 & 0 \\ & \sin 2 & \cos 2 \\ & & \sin 2 \end{bmatrix}\begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

(3) 记 $f(z) = \sin z$. 则

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

故

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

☞ 矩阵 $\boldsymbol{P}$ 是怎么求得的?

☞ 矩阵 $\boldsymbol{P}$ 是怎么求得的?

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

☞ 矩阵 $\boldsymbol{P}$ 是怎么求得的?

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

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

得 $[\boldsymbol{A}\boldsymbol{p}_1, \boldsymbol{A}\boldsymbol{p}_2, \boldsymbol{A}\boldsymbol{p}_3] = [2\boldsymbol{p}_1, 2\boldsymbol{p}_2, \boldsymbol{p}_2 + 2\boldsymbol{p}_3]$.

☞ 矩阵 $\boldsymbol{P}$ 是怎么求得的?

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

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

得 $[\boldsymbol{A}\boldsymbol{p}_1,\boldsymbol{A}\boldsymbol{p}_2,\boldsymbol{A}\boldsymbol{p}_3]=[2\boldsymbol{p}_1,2\boldsymbol{p}_2,\boldsymbol{p}_2+2\boldsymbol{p}_3]$. 于是有

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{p}_1&=2\boldsymbol{p}_1, && \text{即} \quad (2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_1=\boldsymbol{0},\\
\boldsymbol{A}\boldsymbol{p}_2&=2\boldsymbol{p}_2, && \text{即} \quad (2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_2=\boldsymbol{0},\\
\boldsymbol{A}\boldsymbol{p}_3&=\boldsymbol{p}_2+2\boldsymbol{p}_3, && \text{即} \quad (2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2.
\end{aligned}$$

☞ 矩阵 $\boldsymbol{P}$ 是怎么求得的?

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

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

得 $[\boldsymbol{A}\boldsymbol{p}_1,\boldsymbol{A}\boldsymbol{p}_2,\boldsymbol{A}\boldsymbol{p}_3]=[2\boldsymbol{p}_1,2\boldsymbol{p}_2,\boldsymbol{p}_2+2\boldsymbol{p}_3]$. 于是有

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{p}_1&=2\boldsymbol{p}_1, && \text{即} \quad (2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_1=\boldsymbol{0},\\
\boldsymbol{A}\boldsymbol{p}_2&=2\boldsymbol{p}_2, && \text{即} \quad (2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_2=\boldsymbol{0},\\
\boldsymbol{A}\boldsymbol{p}_3&=\boldsymbol{p}_2+2\boldsymbol{p}_3, && \text{即} \quad (2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2.
\end{aligned}$$

由

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

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

$$\boldsymbol{\alpha}_1=(1,1,0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(0,1,1)^{\mathrm{T}}.$$

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

$$\boldsymbol{\alpha}_1=(1,1,0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(0,1,1)^{\mathrm{T}}.$$

又

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

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

$$\boldsymbol{\alpha}_1=(1,1,0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(0,1,1)^{\mathrm{T}}.$$

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 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2$ 有解, $-\boldsymbol{p}_2$ 需满足 $b_1=0$, $b_2=b_3$. 则 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_2=\boldsymbol{0}$ 的解可取为

$$\boldsymbol{p}_2=\boldsymbol{\alpha}_2=(0,1,1)^{\mathrm{T}},$$

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

$$\boldsymbol{\alpha}_1=(1,1,0)^{\mathrm{T}},\quad \boldsymbol{\alpha}_2=(0,1,1)^{\mathrm{T}}.$$

又

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

故要使方程组 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2$ 有解, $-\boldsymbol{p}_2$ 需满足 $b_1=0,\ b_2=b_3$. 则 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_2=\boldsymbol{0}$ 的解可取为

$$\boldsymbol{p}_2=\boldsymbol{\alpha}_2=(0,1,1)^{\mathrm{T}},$$

从而方程组 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_1=\boldsymbol{0}$ 的解可取为

$$\boldsymbol{p}_1=\boldsymbol{\alpha}_1=(1,1,0)^{\mathrm{T}}.$$

再求解非齐次线性方程组 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2$,

再求解非齐次线性方程组 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2$, 由

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

再求解非齐次线性方程组 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2$, 由

$$\left[2\boldsymbol{I}-\boldsymbol{A},\ -\boldsymbol{p}_2\right]=\left[\begin{array}{rrr:r} 0 & 0 & 0 & 0\\ -1 & 1 & -1 & -1\\ -1 & 1 & -1 & -1\end{array}\right]\sim\left[\begin{array}{rrr:r} 1 & -1 & 1 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right],$$

可取

$$\boldsymbol{p}_3=(1,0,0)^{\mathrm{T}}.$$

再求解非齐次线性方程组 $(2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2$, 由

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

可取

$$\boldsymbol{p}_3=(1,0,0)^{\mathrm{T}}.$$

另由上述过程可知 $\boldsymbol{P}$ 不唯一.

Example 4.9

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

解: 先求 $\boldsymbol{A}$ 的若当标准形. 因

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

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

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

解: 先求 $\boldsymbol{A}$ 的若当标准形. 因

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

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

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

所以若当标准形为

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

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

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

得 $[\boldsymbol{A}\boldsymbol{p}_1, \boldsymbol{A}\boldsymbol{p}_2, \boldsymbol{A}\boldsymbol{p}_3] = [\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_2 + \boldsymbol{p}_3]$.

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

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

得 $[\boldsymbol{A}\boldsymbol{p}_1,\boldsymbol{A}\boldsymbol{p}_2,\boldsymbol{A}\boldsymbol{p}_3]=[\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_2+\boldsymbol{p}_3]$. 于是有

$$\begin{aligned}\boldsymbol{A}\boldsymbol{p}_1&=\boldsymbol{p}_1, &&\text{即}\quad (\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_1=\boldsymbol{0},\\ \boldsymbol{A}\boldsymbol{p}_2&=\boldsymbol{p}_2, &&\text{即}\quad (\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_2=\boldsymbol{0},\\ \boldsymbol{A}\boldsymbol{p}_3&=\boldsymbol{p}_2+\boldsymbol{p}_3, &&\text{即}\quad (\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2.\end{aligned}$$

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

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

得 $[\boldsymbol{A}\boldsymbol{p}_1, \boldsymbol{A}\boldsymbol{p}_2, \boldsymbol{A}\boldsymbol{p}_3] = [\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_2 + \boldsymbol{p}_3]$. 于是有

$$\begin{aligned} \boldsymbol{A}\boldsymbol{p}_1 &= \boldsymbol{p}_1, & &\text{即} \quad (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{p}_1 = \boldsymbol{0}, \\ \boldsymbol{A}\boldsymbol{p}_2 &= \boldsymbol{p}_2, & &\text{即} \quad (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{p}_2 = \boldsymbol{0}, \\ \boldsymbol{A}\boldsymbol{p}_3 &= \boldsymbol{p}_2 + \boldsymbol{p}_3, & &\text{即} \quad (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{p}_3 = -\boldsymbol{p}_2. \end{aligned}$$

即 $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ 是方程组 $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 的解.

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

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

得 $[\boldsymbol{A}\boldsymbol{p}_1,\boldsymbol{A}\boldsymbol{p}_2,\boldsymbol{A}\boldsymbol{p}_3]=[\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_2+\boldsymbol{p}_3]$. 于是有

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{p}_1&=\boldsymbol{p}_1, &&\text{即}\quad (\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_1=\boldsymbol{0},\\
\boldsymbol{A}\boldsymbol{p}_2&=\boldsymbol{p}_2, &&\text{即}\quad (\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_2=\boldsymbol{0},\\
\boldsymbol{A}\boldsymbol{p}_3&=\boldsymbol{p}_2+\boldsymbol{p}_3, &&\text{即}\quad (\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2.
\end{aligned}$$

即 $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ 是方程组 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的解. 由

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

取齐次方程组 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的基础解系为 $\boldsymbol{\alpha}_1=(1,1,0)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2=(1,0,1)^{\mathrm{T}}$.

可以选取 $\boldsymbol{p}_1 = \boldsymbol{\alpha}_1$.

可以选取 $\boldsymbol{p}_1 = \boldsymbol{\alpha}_1$. 但是不能简单地令 $\boldsymbol{p}_2 = \boldsymbol{\alpha}_2$, 因为 $\boldsymbol{p}_2$ 还要保证 $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{p}_3 = -\boldsymbol{p}_2$ 有解.

可以选取 $\boldsymbol{p}_1 = \boldsymbol{\alpha}_1$. 但是不能简单地令 $\boldsymbol{p}_2 = \boldsymbol{\alpha}_2$, 因为 $\boldsymbol{p}_2$ 还要保证 $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{p}_3 = -\boldsymbol{p}_2$ 有解.

令

$$\boldsymbol{p}_2 = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 = k_1\begin{bmatrix}1\\1\\0\end{bmatrix} + k_2\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}k_1 + k_2\\k_1\\k_2\end{bmatrix},$$

其中 k_1, k_2 是不全为零的待定常数.

可以选取 $\boldsymbol{p}_1=\boldsymbol{\alpha}_1$. 但是不能简单地令 $\boldsymbol{p}_2=\boldsymbol{\alpha}_2$, 因为 $\boldsymbol{p}_2$ 还要保证 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{p}_3=-\boldsymbol{p}_2$ 有解.

令

$$\boldsymbol{p}_2=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2=k_1\begin{bmatrix}1\\1\\0\end{bmatrix}+k_2\begin{bmatrix}1\\0\\1\end{bmatrix}=\begin{bmatrix}k_1+k_2\\k_1\\k_2\end{bmatrix},$$

其中 k_1, k_2 是不全为零的待定常数. 由

$$(\boldsymbol{I}-\boldsymbol{A},-\boldsymbol{p}_2)=\left[\begin{array}{ccc:c}-1&1&1&-k_1-k_2\\-2&2&2&-k_1\\1&-1&-1&-k_2\end{array}\right]\sim\left[\begin{array}{ccc:c}1&-1&-1&-k_2\\0&0&0&k_1+2k_2\\0&0&0&0\end{array}\right],$$

可以选取 $\boldsymbol{p}_1 = \boldsymbol{\alpha}_1$. 但是不能简单地令 $\boldsymbol{p}_2 = \boldsymbol{\alpha}_2$, 因为 $\boldsymbol{p}_2$ 还要保证 $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{p}_3 = -\boldsymbol{p}_2$ 有解.

令

$$\boldsymbol{p}_2 = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 = k_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} k_1 + k_2 \\ k_1 \\ k_2 \end{bmatrix},$$

其中 k_1, k_2 是不全为零的待定常数. 由

$$(\boldsymbol{I} - \boldsymbol{A}, -\boldsymbol{p}_2) = \left[\begin{array}{ccc:c} -1 & 1 & 1 & -k_1 - k_2 \\ -2 & 2 & 2 & -k_1 \\ 1 & -1 & -1 & -k_2 \end{array}\right] \sim \left[\begin{array}{ccc:c} 1 & -1 & -1 & -k_2 \\ 0 & 0 & 0 & k_1 + 2k_2 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

故要使方程组 $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{p}_3 = -\boldsymbol{p}_2$ 有解, 需满足 $k_1 + 2k_2 = 0$, 即

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

可以选取 $\boldsymbol{p}_1 = \boldsymbol{\alpha}_1$. 但是不能简单地令 $\boldsymbol{p}_2 = \boldsymbol{\alpha}_2$, 因为 $\boldsymbol{p}_2$ 还要保证 $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{p}_3 = -\boldsymbol{p}_2$ 有解.

令

$$\boldsymbol{p}_2 = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 = k_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} k_1 + k_2 \\ k_1 \\ k_2 \end{bmatrix},$$

其中 k_1, k_2 是不全为零的待定常数. 由

$$(\boldsymbol{I} - \boldsymbol{A}, -\boldsymbol{p}_2) = \left[\begin{array}{ccc:c} -1 & 1 & 1 & -k_1 - k_2 \\ -2 & 2 & 2 & -k_1 \\ 1 & -1 & -1 & -k_2 \end{array}\right] \sim \left[\begin{array}{ccc:c} 1 & -1 & -1 & -k_2 \\ 0 & 0 & 0 & k_1 + 2k_2 \\ 0 & 0 & 0 & 0 \end{array}\right],$$

故要使方程组 $(\boldsymbol{I} - \boldsymbol{A})\boldsymbol{p}_3 = -\boldsymbol{p}_2$ 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

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

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

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

可取 $\boldsymbol{p}_3=(1,0,0)^{\mathrm{T}}$.

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

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

可取 $\boldsymbol{p}_3=(1,0,0)^{\mathrm{T}}$. 故可取

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

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

Example 4.10

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

Example 4.10

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

解: $f(z)=\left(1-\frac{z}{3}\right)^{-1}$, 其收敛半径为 3.

Example 4.10

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

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

Example 4.10

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

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

Example 4.10

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

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

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

因

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

因

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

所以

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

因

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

所以

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

故

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

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

Definition 4.11

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 $\boldsymbol{A}$ 的谱. $\boldsymbol{A}$ 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 $m_1 + m_2 + \cdots + m_s = m$.

Definition 4.11

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 $\boldsymbol{A}$ 的谱. $\boldsymbol{A}$ 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

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

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

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

称为 $f(\lambda)$ 在 $\boldsymbol{A}$ 上的谱值.

Definition 4.11

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 $\boldsymbol{A}$ 的谱. $\boldsymbol{A}$ 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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

知 $f(\boldsymbol{J})=p(\boldsymbol{J})$.

于是

$$f(\boldsymbol{A}) = \boldsymbol{Q} f(\boldsymbol{J}) \boldsymbol{Q}^{-1}$$

于是

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

于是

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

于是

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

于是

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

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

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

于是

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

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

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

令

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

于是

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

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

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

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 $a_0, a_1, \cdots, a_{m-1}$.

Exercise 4.12 (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.

Exercise 4.12 (P. 211 习题一 9 (1))

计算级数 $\sum\limits_{k=0}^{\infty}\begin{bmatrix}0.1 & 0.7\\0.3 & 0.6\end{bmatrix}^k$ 的和.

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

Exercise 4.12 (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 $\boldsymbol{A} = \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}$, 由

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

Exercise 4.12 (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 $\boldsymbol{A} = \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}$, 由

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

得特征值为 $\lambda_{1,2} = \dfrac{0.7 \pm \sqrt{1.09}}{2}$,

Exercise 4.12 (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 $\boldsymbol{A}=\begin{bmatrix}0.1 & 0.7\\0.3 & 0.6\end{bmatrix}$, 由

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

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

Exercise 4.12 (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 $\boldsymbol{A} = \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}$, 由

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

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

$$\sum_{k=0}^{\infty} \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}^k = (\boldsymbol{I} - \boldsymbol{A})^{-1}$$

Exercise 4.12 (P. 211 习题一 9 (1))

计算级数 $\sum\limits_{k=0}^{\infty}\begin{bmatrix}0.1 & 0.7\\ 0.3 & 0.6\end{bmatrix}^k$ 的和.

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

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

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

$$\sum_{k=0}^{\infty}\begin{bmatrix}0.1 & 0.7\\ 0.3 & 0.6\end{bmatrix}^k=(\boldsymbol{I}-\boldsymbol{A})^{-1}=\begin{bmatrix}0.9 & -0.7\\ -0.3 & 0.4\end{bmatrix}^{-1}$$

Exercise 4.12 (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 $\boldsymbol{A} = \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}$, 由

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

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

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