第一章　矩阵分析

在学习微积分的课程中，我们对于函数研究过其极限、求导、求积分等操作，能否将这些数学工具应用到矩阵中呢？这就是本章矩阵分析要干的事情。

本章内容包括：

- 矩阵序列与矩阵级数
- 矩阵函数

1.1　矩阵序列与矩阵级数

1.1.1　矩阵序列

从向量序列到矩阵序列

在第二章介绍向量范数的应用的时候，我们曾经接触过向量序列的敛散性问题：

$$\lim_{k\to\infty} x^k = \lim_{k\to\infty} \|x^k - a\| = 0$$

也在相对应的地方介绍了何为向量序列，向量序列也是通过数列的情况延伸而来，再对向量序列进行延伸，便有了矩阵序列的概念

定义 1.1.1: 矩阵序列

设 $m \times n$ 型矩阵序列为 $\boldsymbol{A}^{(k)}$，其中

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

矩阵序列的极限

与向量序列的极限类似，矩阵序列的极限定义如下：

定义 1.1.2

若

$$\lim_{k\to\infty} a_{ij}^{(k)} = a_{ij}$$

则

$$\lim_{k\to\infty} \boldsymbol{A}^{(k)} = \boldsymbol{A}$$

矩阵序列极限的性质

定理 1.1.1

设 $\lim\limits_{k\to\infty} \boldsymbol{A}^{(k)=\boldsymbol{A}}$, $\lim\limits_{k\to\infty} \boldsymbol{B}^{(k)=\boldsymbol{B},\alpha,\beta\in\mathbb{C}}$ 则

1. $\lim\limits_{k\to\infty}(\alpha\boldsymbol{A}^{(k)} + \beta\boldsymbol{B}^{(k)}) = \alpha\boldsymbol{A} + \beta\boldsymbol{B}$
2. $\lim\limits_{k\to\infty} \boldsymbol{A}^{(k)}\boldsymbol{B}^{(k)} = \boldsymbol{A}\boldsymbol{B}$
3. 当 $\boldsymbol{A}^(k)$ 与 $\boldsymbol{A}$ 都可逆时，$\lim\limits_{k\to\infty}(\boldsymbol{A}^{(k)})^{-1} = \boldsymbol{A}^{-1}$

下面的定理给出矩阵序列收敛条件：

定理 1.1.2

设 $\|\cdot\|$ 是 $\mathbb{C}^{m\times n}$ 上任一矩阵范数，$\mathbb{C}^{m\times n}$ 中矩阵序列收敛于 $\boldsymbol{A}$ 的充要条件是

$$\lim_{k\to\infty} \|\boldsymbol{A}^{(k)} - \boldsymbol{A}\| = 0$$

与向量序列的收敛判断类似，定理1.1.2指出了矩阵序列收敛条件，下面的定义将会指出何为收敛矩阵。

定义 1.1.3

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，若 $\lim\limits_{k\to\infty} \boldsymbol{A}^k = 0$($k$为正整数)，则称 $\boldsymbol{A}$ 为**收敛矩阵**.

下面举一个例子，来表示何为收敛矩阵

例 1.1.1.

$$A^{(k)} = \begin{bmatrix} \frac{1}{2^k} & 0 \\ 0 & \frac{1}{2^k} \end{bmatrix}$$

我们知道

$$\lim_{k\to\infty} \frac{1}{2^k} = 0$$

自然就可以推出

$$\lim_{k\to\infty} \boldsymbol{A}^k = 0$$

就代表矩阵 $\boldsymbol{A}$ 是一个收敛矩阵。

下面介绍一下矩阵能够成为收敛矩阵的条件

定理 1.1.3

设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，则 $\boldsymbol{A}$ 为收敛矩阵的充要条件是谱半径 $r(\boldsymbol{A}) < 1$

证明如下：

证明.

充分性证明：

由 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，再由 Jordan 标准型的知识可知，存在矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J} = diag(J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_s}(\lambda_s))$$

由矩阵乘法的知识，不难看出

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

而

$$\boldsymbol{J}_{r_i}^k(\lambda_i) = \begin{bmatrix} f_k(\lambda_i) & f_k'(\lambda_i) & \cdots & \frac{f_k^{(r_i-1)}(\lambda_i)}{(r_i-1)!} \\ & f_k(\lambda_i) & \cdots & \frac{f_k^{(r_i-2)}(\lambda_i)}{(r_i-2)!} \\ & & \ddots & \vdots \\ & & & f_k(\lambda_i) \end{bmatrix}$$

其中，$f_k(\lambda_i) = \lambda_i^k$

由于 $r(\boldsymbol{A}) < 1$，则可以知道矩阵 $\boldsymbol{A}$ 的所有特征值均小于 1，那不难看出，当 $k \to \infty$ 时，$f_k(\lambda_i) = \lambda_i^k \to 0$, 从而 $f_k^{(l)}(\lambda_i) \to 0, \boldsymbol{J}_{r_i}^k(\lambda_i) \to 0$

因此，$\boldsymbol{J} \to 0$, 故 $\boldsymbol{A} \to 0$，从而知道 $\boldsymbol{A}$ 是一个收敛矩阵，充分性证毕。

必要性证明：

由于 $\boldsymbol{A}^k \to 0$，故 $\boldsymbol{J}_{r_i}^k(\lambda_i) \to 0$, 故 $\lambda_i^k \to 0$，从而可以知道 $\boldsymbol{A}$ 的所有特征值绝对值必定小于 1，因此 $r(\boldsymbol{A}) < 1$，必要性证毕。 □

1.1.2 矩阵级数

介绍完矩阵序列，接下来介绍矩阵级数。

定义 1.1.4

设 $\boldsymbol{A}^{(k)}$ 是 $\mathbb{C}^{m\times n}$ 的矩阵序列，称

$$\sum_{k=1}^{\infty} \boldsymbol{A}^{(k)} = \boldsymbol{A}^{(1)} + \boldsymbol{A}^{(2)} + \cdots + \boldsymbol{A}^{(k)} + \cdots$$

为**矩阵级数**，称 $\boldsymbol{S}^{(N)} = \sum\limits_{k=1}^{N} \boldsymbol{A}^{(k)}$ 为矩阵级数的**部分和**，如果 $\lim\limits_{N\to\infty} \boldsymbol{S}^{(N)} = \boldsymbol{S}$，则称 $\sum\limits_{k=1}^{\infty} \boldsymbol{A}^{(k)}$ **收敛**

上面的定义是从矩阵序列的角度定义的矩阵级数，接下来，与研究数项级数类似，如何判断矩阵级数绝对收敛?

定义 1.1.5

如果 mn 个数项级数

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

都绝对收敛，则称矩阵级数 $\sum_{k=1}^{infty} \boldsymbol{A}^{(k)}$ 绝对收敛。

上面的定义是通过数项级数绝对收敛推出矩阵级数的绝对收敛，还可以通过利用范数来确定矩阵级数是否绝对收敛：

定理 1.1.4

在 $\mathbb{C}^{n\times n}$ 中，$\sum\limits_{k=1}^{\infty} \boldsymbol{A}^{(k)}$ 绝对收敛的充要条件是正项级数 $\sum\limits_{k=1}^{\infty} \|\boldsymbol{A}^{(k)}\|$ 收敛。

证明如下：

证明.

若 $\sum\limits_{k=1}^{\infty}\boldsymbol{A}^{(k)}$ 绝对收敛，则 $\sum\limits_{k=1}^{\infty}|a_{ij}^{(k)}|\le M$

因此，根据矩阵 1-范数，有

$$\sum_{k=1}^{N}\|\boldsymbol{A}^{(k)}\|_{m_1}=\sum_{k=1}^{N}\left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}^{(k)}|\right)\le mnM$$

因此，若矩阵 1-范数收敛，则任意矩阵范数一定收敛（小收敛则大收敛）若任意矩阵范数收敛，那必定矩阵 1-范数收敛，由于一定存在

$$|a_{ij}^{(k)}|\le\|\boldsymbol{A}^{(k)}\|_{m_1}$$

则可以得出 $\sum_{k=1}^{\infty}a_{ij}^{(k)}$ 绝对收敛，证毕。 □

1.1.3 矩阵幂级数

定理 1.1.5

方阵 $\boldsymbol{A}$ 的 Neumann 级数

$$\sum_{k=0}^{\infty}\boldsymbol{A}^k=\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$$

收敛的充要条件是 $r(\boldsymbol{A})<1$，且收敛时，和为 $(\boldsymbol{I}-\boldsymbol{A})^{-1}$

证明过程略，感兴趣的可以看书。

在学习微积分的无穷级数内容中，我们知道，幂级数存在收敛半径，那矩阵幂级数的收敛半径要如何求呢？下面的定理给出了答案：

定理 1.1.6

设幂级数

$$f(\boldsymbol{z})=\sum_{k=0}^{\infty}c_k\boldsymbol{z}^k$$

的收敛半径为 r，如果方阵 $\boldsymbol{A}$ 满足 $r(\boldsymbol{A})<r$，则矩阵幂级数绝对收敛，如果 $r(\boldsymbol{A})>r$，则矩阵幂级数发散。

1.2 矩阵函数

上一节是为本节内容作为铺垫的，本节将会利用矩阵级数来研究矩阵函数。

1.2.1 函数与幂级数

说到函数，可能大多数人都会想到函数都是类似于下面的形式：

$$f(x) = x, f(x) = x^2, \cdots$$

可在最一开始说，我们将会用矩阵级数的内容来研究矩阵函数，这两者存在何种关系呢？

如果还记得当时学无穷级数——函数展开成幂级数和幂级数的和函数的相关知识的话，就会知道，函数和无穷级数是离不开的，比如，你可能还记得下面这个式子：

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \cdots$$

这就建立起了一个函数与幂级数的关系，事实上，常见初等函数都可以展开成幂级数的形式，这里就不一一介绍了，具体请翻阅高等数学教材。

这种幂级数与函数之间的关系放到矩阵中依旧成立，这就是为什么说用矩阵幂级数的内容来研究矩阵函数的原因，下面给出矩阵函数的定义：

定义 1.2.1

设幂级数 $\sum\limits_{k=0}^{\infty} c_k \boldsymbol{z}^k$ 的收敛半径为 r，且当 $\boldsymbol{z} < r$ 时，幂级数收敛域 $f(\boldsymbol{z})$，即

$$f(\boldsymbol{z}) = \sum_{l=0}^{\infty} c_k \boldsymbol{z}^k, \quad |\boldsymbol{z}| < r$$

如果 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ 满足 $r(\boldsymbol{A}) < r$，则称收敛的矩阵幂级数 $\sum\limits_{k=0}^{\infty} c_k \boldsymbol{z}^k$ 的和记为矩阵函数，记为 $f(\boldsymbol{A})$，即

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

将 $f(\boldsymbol{A})$ 的方阵 $\boldsymbol{A}$ 换成 $\boldsymbol{At}, \boldsymbol{t}$ 为参数，就会得到

$$f(\boldsymbol{At}) = \sum_{k=0}^{\infty} c_k (\boldsymbol{At})^k$$

定义太长可以不看，简单来说就是定义前面的介绍，与普通的数项函数与幂级数的概念类似。

1.2.2 常见的矩阵函数

下面介绍了一些常见的矩阵函数，**这些函数的展开式应熟稔于心** (每天起床头件事，先背一遍展开式 (手动狗头))：

1.
$$e^{\boldsymbol{A}} = \boldsymbol{E} + \boldsymbol{A} + \frac{1}{2!}\boldsymbol{A}^2 + \cdots + \frac{1}{k!}\boldsymbol{A}^k + \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\boldsymbol{A}^{2k+1}$$

2.
$$\sin \boldsymbol{A} = \boldsymbol{A} - \frac{1}{3!}\boldsymbol{A}^3 + \frac{1}{5!}\boldsymbol{A}^5 - \frac{1}{7!}\boldsymbol{A}^7 + \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\boldsymbol{A}^{(2k+1)}$$

3.
$$\cos \boldsymbol{A} = \boldsymbol{E} - \frac{1}{2!}\boldsymbol{A}^2 + \frac{1}{4!}\boldsymbol{A}^4 - \frac{1}{6!}\boldsymbol{A}^6 + \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\boldsymbol{A}^{(2k)}$$

4.
$$(\boldsymbol{E} - \boldsymbol{A})^{-1} = \boldsymbol{E} + \boldsymbol{A} + \boldsymbol{A}^2 + \cdots + \boldsymbol{A}^k + \cdots = \sum_{k=0}^{\infty} \boldsymbol{A}^k$$

5.
$$\ln(\boldsymbol{E} + \boldsymbol{A}) = \boldsymbol{A} - \frac{1}{2}\boldsymbol{A}^2 + \frac{1}{3}\boldsymbol{A}^3 - \frac{1}{4}\boldsymbol{A}^4 + \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1}\boldsymbol{A}^{k+1}$$

1.2.3 矩阵函数值的计算

只有函数的形式还没有用，我们最终要像计算普通函数那样把矩阵函数的函数值给计算出来，有三种方法可以计算矩阵函数值。

相似对角化计算矩阵函数值

若矩阵 $\boldsymbol{A}$ 可以相似对角化，则存在下面的等式

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda} = diag(\lambda_1, \lambda_2, \cdots, \lambda_n) = \boldsymbol{D}$$

故

$$\begin{aligned} f(\boldsymbol{A}) &= \sum_{k=0}^{\infty} c_k \boldsymbol{A}^k = \sum_{k=0}^{\infty} c_k (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})^k = \boldsymbol{P}\left(\sum_{k=0}^{\infty} c_k \boldsymbol{D}^k\right)\boldsymbol{P}^{-1} \\ &= \boldsymbol{P}\begin{bmatrix} \sum\limits_{k=0}^{\infty} c_k\lambda_1^k & & \\ & \ddots & \\ & & \sum\limits_{k=0}^{\infty} c_k\lambda_n^k \end{bmatrix}\boldsymbol{P}^{-1} = \boldsymbol{P}\begin{bmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{bmatrix}\boldsymbol{P}^{-1} \end{aligned}$$

其中

$$f(\lambda_i)=\sum_{k=0}^{\infty}c_k\lambda_i^k$$

因此，使用相似对角化的方法求解矩阵函数值的步骤为：

1. 将矩阵进行相似对角化
2. 直接代公式算答案

注意：使用该方法前，请务必确认该矩阵是否可以相似对角化，有关相似对角化的内容，可以参考第零章相似理论的相关知识，或查阅第一章 Jordan 标准型的相关知识。

使用 Jordan 标准型的方式计算矩阵函数值

设

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}=diag(\boldsymbol{J}_1,\boldsymbol{J}_2,\cdots,\boldsymbol{J}_s),\boldsymbol{J}_i=\begin{bmatrix}\lambda_1 & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i\end{bmatrix}_{m_i\times m_i}$$

则

$$f(\boldsymbol{J}_i)=\sum_{k=0}^{\infty}a_k\boldsymbol{J}_i^k=\begin{bmatrix}f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \cdots & \frac{1}{(m_i-1)!}f^{(m_i-1)}(\lambda_i) \\ & f(\lambda_i) & \cdots & \vdots \\ & & \ddots & \frac{1}{1!}f'(\lambda_i) \\ & & & f(\lambda_i)\end{bmatrix}$$

故

$$\begin{aligned}f(\boldsymbol{A})&=\sum_{k=0}^{\infty}a_k\boldsymbol{P}\boldsymbol{J}^k\boldsymbol{P}^{-1}=(\sum_{k=0}^{\infty}a_k\boldsymbol{J}^k)\boldsymbol{P}^{-1}\\&=\boldsymbol{P}\begin{bmatrix}\sum\limits_{k=0}^{\infty}a_k\boldsymbol{J}_1^k & & \\ & \ddots & \\ & & \sum\limits_{k=0}^{\infty}a_k\boldsymbol{J}_s^k\end{bmatrix}\boldsymbol{P}^{-1}=\boldsymbol{P}\begin{bmatrix}f(\boldsymbol{J}_1) & & \\ & \ddots & \\ & & f(\boldsymbol{J}_s)\end{bmatrix}\boldsymbol{P}^{-1}\end{aligned}$$

因此，使用 Jordan 标准型的方法求解矩阵函数值的步骤为：

1. 求 Jordan 标准型
2. 代公式计算

使用数项级数求和方式计算矩阵函数值

第一种方法有局限性，第二种方法需要较大计算量，下面将要介绍的这种方法计算量不大，但是需要记住常见函数的矩阵函数展开式。

在开始说计算方法之前，首先介绍哈密尔顿-凯莱定理：

定理 1.2.1: 哈密尔顿-凯莱定理

设 $\boldsymbol{A}$ 是数域 $\mathbb{P}$ 上的一个 $n \times n$ 矩阵，$f(\lambda) = |\lambda \boldsymbol{E} - \boldsymbol{A}|$ 是 $\boldsymbol{A}$ 的特征多项式，则

$$f(\boldsymbol{A}) = \boldsymbol{A}^n - b_{n-1}\boldsymbol{A}^{n-1} - \cdots - b_1\boldsymbol{A} - b_0\boldsymbol{E} = 0$$

这个定理有什么用呢？我们继续看。

由定理，我们可以得到下面的等式关系：

$$\boldsymbol{A}^n = b_{n-1}\boldsymbol{A}^{n-1} + \cdots + b_1\boldsymbol{A} + b_0\boldsymbol{E}$$

如果我们想要计算 $\boldsymbol{A}^{n+1}$，那只需要计算

$$\boldsymbol{A}^n \cdot \boldsymbol{A}$$

的结果即可，即

$$\boldsymbol{A}^{n+1} = \boldsymbol{A}^n\boldsymbol{A} = b_0\boldsymbol{A} + b_1\boldsymbol{A}^2 + \cdots + b_{n-1}\boldsymbol{A}^n$$

而

$$\boldsymbol{A}^n = b_{n-1}\boldsymbol{A}^{n-1} + \cdots + b_1\boldsymbol{A} + b_0\boldsymbol{E}$$

因此，我们重新将式子整理一下合并同类项就可以得到一个新的式子

由此，我们会发现，再计算更高次的矩阵乘法时，我们完全可以用低阶矩阵乘积的结果来表示出来，从而简化了计算，这就是该定理的作用。

可这对计算矩阵函数值有什么用呢？由矩阵幂级数的知识，一个矩阵函数可以写成

$$f(\boldsymbol{A}) = c_0\boldsymbol{E} + c_1\boldsymbol{A} + \cdots + c_k\boldsymbol{A}^k + \cdots$$

的形式，由哈密尔顿-凯莱定理，最高次项可以用低阶来表示，因此可以重新合并同类项，最终就会变成

$$(c_0 + c_nb_0 + c_{n+1}b_0^{(1)} + \cdots)\boldsymbol{E} + (c_1 + c_nb_1 + c_{n+1}b_1^{(1)} + \cdots)\boldsymbol{A} + \cdots$$

的形式，而

$$(c_0 + c_n b_0 + c_{n+1} b_0^{(1)} + \cdots)\text{和}(c_1 + c_n b_1 + c_{n+1} b_1^{(1)} + \cdots)$$

都是数项级数，因此可以使用数项级数的方式来求解矩阵函数函数值。

因此，使用求解数项级数的方法求解矩阵函数值的步骤为:

1. 求特征方程 $|\lambda \boldsymbol{E} - \boldsymbol{A}|$ 的值，利用哈密尔顿-凯莱定理进行简化
2. 代公式计算

（如果只看上面这些抽象方法肯定看不懂，因此每一个方法应该都要加一个例子的，但是时间有限并且这些例子书上都有，所以可以看书或者 PPT 的例子就不在这里写了，等有时间再写）

1.2.4 矩阵函数的其他性质

本节的最后，再介绍一些矩阵函数的其他性质:

1. 如果 $\boldsymbol{AB} = \boldsymbol{BA}$，则 $e^{\boldsymbol{A}} e^{\boldsymbol{B}} = e^{\boldsymbol{B}} e^{\boldsymbol{A}} = e^{\boldsymbol{A}+\boldsymbol{B}}$
2. 如果 $\boldsymbol{AB} = \boldsymbol{BA}$，则
   (a) $\cos(\boldsymbol{A} + \boldsymbol{B}) = \cos \boldsymbol{A} \cos \boldsymbol{B} - \sin \boldsymbol{A} \sin \boldsymbol{B}$
   (b) $\sin(\boldsymbol{A} + \boldsymbol{B}) = \sin \boldsymbol{A} \cos \boldsymbol{B} + \cos \boldsymbol{A} \sin \boldsymbol{B}$

这些性质与代数上的性质都类似，只不过需要注意一下这些性质的成立是基于 $\boldsymbol{AB} = \boldsymbol{BA}$ 的情况下才可以，其他的就没有什么需要注意的了。