GZH007Y 40/2

高等工程数学

Advanced Engineering Mathematics

王 泳

中国科学院大学

2017.10.14

只要一门科学分支能提出大量的问题，它就充满着生命力，而问题缺乏则预示着独立发展的终止或衰亡

— 大卫·希尔伯特 (David Hilbert 1862-1943)

德国数学家、科学家

真与可证是两个概念

不完备性定理

1. 任意一个包含算术系统在内的形式系统中，都存在一个命题，它在这个系统中既不能被证明也不能被否定。

2. 任意一个包含算术系统的形式系统自身不能证明它本身的无矛盾性。

— 库尔特·哥德尔（Kurt Godel 1906-1978)

奥地利数学家、逻辑学家和哲学家

“上帝是存在的，因为数学无疑是相容的；魔鬼也是存在的，因为我们不能证明这种相容性。”

—外尔（Hermann Weyl)

矩阵理论

矩阵理论在自然科学、工程技术、控制理论和社会经济学等领域的应用日趋深广，应用矩阵的理论和方法来解决工程技术和社会经济领域中的实际问题也越来越普遍。

线性空间与线性变换

- 线性空间
- 基变换与坐标变换
- 子空间与维数定理
- 线性空间的同构
- 线性变换的概念
- 线性变换的矩阵表示

内积空间

- 欧氏空间
- 正交基及子空间的正交关系
- 内积空间的同构
- 正交变换
- 复内积空间（酉空间）
- 正规矩阵

矩阵的标准形

- 矩阵的相似对角形
- 矩阵的约当标准形
- 最小多项式

矩阵函数及其应用

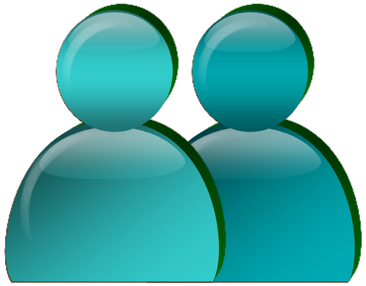
- 向量范数
- 矩阵范数
- 向量和矩阵的极限
- 矩阵幂级数
- 矩阵函数
- 矩阵的微分与积分
- 常用矩阵函数的性质

第四章：矩阵函数及其应用

1. 向量范数
2. 矩阵范数
3. 向量和矩阵的极限
4. 矩阵幂级数
5. 矩阵函数
6. 矩阵的微分与积分

向量范数

1. 向量范数
2. 矩阵范数
3. 向量和矩阵的极限
4. 矩阵幂级数
5. 矩阵函数
6. 矩阵的微分与积分



■ 定义4.1 向量范数

设 V 是数域 P 上的线性（向量）空间，若对于 V 中任一向量 $\boldsymbol{x}$，都有一非负实数 $\|\boldsymbol{x}\|$ 与之对应，并且满足下列三个条件：

（1） 正定性：当 $\boldsymbol{x} \neq \boldsymbol{\theta}$ 时，有 $\|\boldsymbol{x}\| > 0$；

（2） 齐次性：对于任何 $a \in P$，都有 $\|a\boldsymbol{x}\| = |a| \cdot \|\boldsymbol{x}\|$；

（3） 三角不等式：对于任何 $\boldsymbol{x}, \boldsymbol{y} \in V$，都有 $\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$，

则称非负实数 $\|\boldsymbol{x}\|$ 为向量 $\boldsymbol{x}$ 的范数。简言之，向量的范数是定义在线性空间上的非负实值函数。定义了范数的线性空间被称为赋范线性空间。

例如，实内积空间及酉空间中定义的向量长度 $|\boldsymbol{x}| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}$ 就是一种向量范数，因为 $|\boldsymbol{x}|$ 满足定义 4.1 中 $\|\boldsymbol{x}\|$ 的所有条件。

向量范数

酉空间 $\mathbf{C}^n$ 的向量范数 $\|\boldsymbol{x}\|$ 用向量长度 $|\boldsymbol{x}|$ 来定义时，被记作 $\|\boldsymbol{x}\|_2$，即

$$\|\boldsymbol{x}\|_2=|\boldsymbol{x}|=\sqrt{\boldsymbol{x}\boldsymbol{x}^H}=\sqrt{\sum_{i=1}^{n}\xi_i\overline{\xi_i}}=\sqrt{\sum_{i=1}^{n}|\xi_i|^2}$$

其中 $\boldsymbol{x}=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbf{C}^n$。$\|\boldsymbol{x}\|_2$ 被称为欧氏范数或 2-范数。

其实，在酉空间 $\mathbf{C}^n$ 上的赋范，实质上就是在 $\mathbf{C}^n$ 上定义满足上述三个条件的一个实函数，因此 $\mathbf{C}^n$ 上可以找到无穷多种范数。

■ 例4.1

若对酉空间 $\mathbf{C}^n$ 的每个向量 $\boldsymbol{x}=(\xi_1,\xi_2,\cdots,\xi_n)$，定义

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

则易证它是 $\mathbf{C}^n$ 的向量的范数。$\|\boldsymbol{x}\|_\infty$ 被称为∞-范数。

■ **例4.2**

对于任意的 $\boldsymbol{x}=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathrm{C}^n$，规定

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

则 $\|\boldsymbol{x}\|_1$ 也是 C^n 的向量的范数。$\|\boldsymbol{x}\|_1$ 被称为 1-范数。

一般地，可以证明：若定义

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

则 $\|\boldsymbol{x}\|_p$ 也是酉空间 C^n 的向量范数，并被称为向量 $\boldsymbol{x}$ 的 p-范数。

向量范数

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

显然，当 $p=1$ 时，即为例 4.2 的范数 $\|\boldsymbol{x}\|_1$；当 $p=2$ 时，即为前面讲过的范数 $\|\boldsymbol{x}\|_2$；当 $p \to \infty$ 时，即为例 4.1 的范数 $\|\boldsymbol{x}\|_\infty$，即

$$\|\boldsymbol{x}\|_\infty = \lim_{p\to\infty} \|\boldsymbol{x}\|_p = \max_{1\le i\le n} |\xi_i|$$

证明：

当 $\boldsymbol{x} = \boldsymbol{\theta}$ 时，上式显然成立。故只需对非零向量 $\boldsymbol{x}$ 加以证明，令 $w = \max_{1\le i\le n} |\xi_i|$，则有

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

其中 $\beta_i = \left|\frac{\xi_i}{w}\right| \le 1$，又因为至少有一个 $\beta_i = 1$，所以有

$$1 \le \sum_{i=1}^{n} \beta_i^p \le n$$

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

因此

$$1\le\left(\sum_{i=1}^{n}\beta_i^{\,p}\right)^{\frac{1}{p}}\le n^{\frac{1}{p}}$$

又因为 $\lim\limits_{p\to\infty}n^{\frac{1}{p}}=1$，所以有

$$\lim_{p\to\infty}\left(\sum_{i=1}^{n}\beta_i^{\,p}\right)^{\frac{1}{p}}=1$$

从而

$$\lim_{p\to\infty}\|\boldsymbol{x}\|_p=\lim_{p\to\infty}\left(\sum_{i=1}^{n}|\xi_i|^p\right)^{\frac{1}{p}}=w\lim_{p\to\infty}\left(\sum_{i=1}^{n}\beta_i^{\,p}\right)^{\frac{1}{p}}=w=\max_{1\le i\le n}|\xi_i|$$

即

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

命题得证

在无限维向量空间中，两个向量范数是可以不等价的

- **定理4.1**

对于任何有限维向量空间 V 上定义的任意两种向量范数 $\|x\|_a$ 及 $\|x\|_b$，都存在两个与 x 无关的正的常数 C_1，C_2，使得对 V 中任一向量 x，都有

$$\|x\|_a \le C_1\|x\|_b,\quad \|x\|_b \le C_2\|x\|_a$$

满足上面两个不等式的两种向量范数被称为等价的。

其实，上面的两个不等式也可以合二为一，即存在两个与 x 无关的正的常数 K_1，K_2 对 V 中任一向量 x，都有

$$\boxed{K_1\|x\|_a \le \|x\|_b \le K_2\|x\|_a}$$

定理 4. 1 可以叙述为：有限维向量空间上的不同向量范数是等价的。

同一个向量 $x \in \mathbf{C}^n$，按不同公式所定义的向量范数，其大小一般是不相等的。例如，当取 $x=(1,1,\cdots,1)\in \mathbf{C}^n$，则

$$\|x\|_2=\sqrt{n},\quad \|x\|_1=n,\quad \|x\|_\infty=1$$

但在考虑向量序列的收敛性时，则效果是一致的，即此序列在一种范数定义下收敛，则在另一种范数定义下亦收敛，这也就是所谓向量范数的等价性。

思考题　（15分）

设$\|x\|_a$与$\|x\|_b$是C^n上的两种向量范数，k_1、k_2是正常数，求证下列函数

$$（1）\max\left\{\|x\|_a,\|x\|_b\right\}$$

$$（2）k_1\|x\|_a+k_2\|x\|_b$$

都是向量范数。

矩阵范数

1. 向量范数
2. 矩阵范数
3. 向量和矩阵的极限
4. 矩阵幂级数
5. 矩阵函数
6. 矩阵的微分与积分

■ 定义4.2 矩阵范数

在 $P^{n\times n}$ 上定义一个非负实值函数 $\|A\|$（对每个 $A\in P^{n\times n}$），如果对任意的 $A,B\in P^{n\times n}$ 都满足下面的四个条件：

（1） 正定性：若 $A\neq 0$（零矩阵），则 $\|A\|>0$；

（2） 齐次性：对于任何 $a\in P$，都有 $\|aA\|=|a|\cdot\|A\|$；

（3） 三角不等式：对于任何 $A,B\in P^{n\times n}$，都有 $\|A+B\|\leq\|A\|+\|B\|$；

（4） $\|AB\|\leq\|A\|\cdot\|B\|$

则非负实函数 $\|A\|$ 被称为 $n\times n$ 方阵的范数。

与向量的情形一样，矩阵也可以有各种各样的范数，而且在大多数情况下，矩阵范数常和向量范数混合在一起使用，因此这就需要考察矩阵范数与向量范数的关系。

■ 定义4.3 矩阵范数与向量范数的相容

若对任意 $\boldsymbol{A} \in P^{n\times n}$ 及 n 维列向量 $x \in P^n$，方阵范数 $\|\boldsymbol{A}\|$ 能与某种向量范数 $\|\boldsymbol{x}\|_a$ 满足关系式

$$\|\boldsymbol{Ax}\|_a \le \|\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|_a$$

则称方阵范数 $\|\boldsymbol{A}\|$ 与向量范数 $\|\boldsymbol{x}\|_a$ 是相容的。

可以证明：

（1） $P^{n\times n}$ 上的每一种方阵范数，在 P^n 上都存在与之相容的向量范数；

（2） $P^{n\times n}$ 上任意两种方阵范数 $\|\boldsymbol{A}\|_\alpha$，$\|\boldsymbol{A}\|_\beta$ 都是等价的，即存在两个与 $\boldsymbol{A}$ 无关的正数 C_1，C_2，使得对 $P^{n\times n}$ 中任一方阵 $\boldsymbol{A}$，都有

$$\|\boldsymbol{A}\|_\alpha \le C_1 \|\boldsymbol{A}\|_\beta，\quad \|\boldsymbol{A}\|_\beta \le C_2 \|\boldsymbol{A}\|_\alpha$$

或写成

$$C_1 \|\boldsymbol{A}\|_\alpha \le \|\boldsymbol{A}\|_\beta \le C_2 \|\boldsymbol{A}\|_\alpha$$

矩阵范数

设 $A=\left(a_{ij}\right)_{n\times n}\in \mathbf{C}^{n\times n}$，则比较常用的矩阵范数有：

（1）$\|A\|_1=\max\limits_{1\le j\le n}\sum\limits_{i=1}^{n}\left|a_{ij}\right|$（列模和最大者）；

（2）$\|A\|_\infty=\max\limits_{1\le i\le n}\sum\limits_{j=1}^{n}\left|a_{ij}\right|$（行模和最大者）；

（3）$\|A\|_2=\sqrt{\lambda_{A^HA}}$（$\lambda_{A^HA}$ 是 A^HA 的最大特征值）；

$$\|A\|_\infty=\max_{1\le i\le n}\left\{\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}}_{\|A\|_1=\max\limits_{1\le j\le n}}\right.$$

因为 $\left(A^HA\right)^H=A^HA$，即 A^HA 是厄米特矩阵，它对应的二次型

$$f(x)=x^HA^HAx=(Ax)^H(Ax)=y^Hy\ge 0$$

是正定或半正定的，因此它的特征值都大于或等于零。

（4）$\|A\|_F=\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2}=\underline{\sqrt{\mathrm{tr}\left(A^H A\right)}}$ 是一种与向量范数 $\|x\|_2=\sqrt{\sum_{i=1}^{n}|\xi_i|^2}$ 相容的方阵范数，被称为 Frobenius 范数（$x\in\mathrm{C}^n$），也被称为 F-范数。

F-范数的优点之一是矩阵乘以酉矩阵 U 后 F-范数不变（在实矩阵的情况下是乘以正交矩阵后不变），即

$$\|UA\|_F=\|A\|_F=\|AU\|_F$$

证明：

根据 F-范数的定义得 $$\|UA\|_F^{\ 2}=\mathrm{tr}\left[(UA)^H(UA)\right]=\mathrm{tr}\left[A^H A\right]=\|A\|_F^{\ 2}$$

即 $$\|UA\|_F=\|A\|_F$$

又因为 $\|A\|_F=\|A^H\|_F$，且 U^H 也是酉矩阵，故根据上面得到的结果可以推得

$$\|AU\|_F=\left\|(AU)^H\right\|_F=\|U^H A^H\|_F=\|A^H\|_F=\|A\|_F$$

命题得证

所以，综合可得 $$\|UA\|_F=\|A\|_F=\|AU\|_F$$

由此可知，A 的酉相似矩阵的 F-范数都是相同的，即

若 $B = U^H A U$，则 $\|B\|_F = \|A\|_F$。

思考题　（15分）

若 T 为正交矩阵，又 $A \in \mathbf{R}^{n \times n}$，证明：

（1）$\|T\|_2 = 1$；

（2）$\|A\|_2 = \|AT\|_2$

向量和矩阵的极限

1. 向量范数
2. 矩阵范数
3. 向量和矩阵的极限
4. 矩阵幂级数
5. 矩阵函数
6. 矩阵的微分与积分

■ 定义4.4 向量序列的极限

若 $\boldsymbol{x}^{(m)}=\left(\xi_1^{(m)},\xi_2^{(m)},\cdots,\xi_n^{(m)}\right)\in\mathbf{C}^n$（$m=1,2,\cdots$），如果存在极限

$$\lim_{m\to\infty}\xi_i^{(m)}=\xi_i\quad(m=1,2,\cdots),$$

则称酉空间 $\mathbf{C}^n$ 的向量序列 $\left\{\boldsymbol{x}^{(m)}\right\}$ 收敛于向量 $\boldsymbol{x}=\left(\xi_1,\xi_2\cdots,\xi_n\right)$，并且记为

$$\lim_{m\to\infty}\boldsymbol{x}^{(m)}=\boldsymbol{x}\ \text{或}\ \boldsymbol{x}^{(m)}\to\boldsymbol{x}\quad(m\to\infty)$$

换言之，向量序列的极限是按照坐标序列的极限来定义的，当向量序列不收敛时，也叫做发散的。

■ **定理4.2** $\lim\limits_{m\to\infty} \boldsymbol{x}^{(m)} = \boldsymbol{x} \Leftrightarrow \lim\limits_{m\to\infty} \left\| \boldsymbol{x}^{(m)} - \boldsymbol{x} \right\| = 0$（对任一向量范数 $\|\bullet\|$ 成立）

证明：

利用向量范数的等价性，易知只要对一种向量范数进行证明，则对任何一种向量范数也能成立，为此，取向量范数 $\|\boldsymbol{x}\|_\infty$。

（充分性）如对向量范数 $\|\boldsymbol{x}\|_\infty$ 有 $\lim\limits_{m\to\infty} \left\| \boldsymbol{x}^{(m)} - \boldsymbol{x} \right\|_\infty = 0$

则由 $\left\| \boldsymbol{x}^{(m)} - \boldsymbol{x} \right\|_\infty = \max\limits_{1\le i\le n} \left| \xi_i^{(m)} - \xi_i \right| \to 0$（$m \to \infty$）

可知，对每个（$i = 1, 2, \cdots, n$）有 $\xi_i^{(m)} \to \xi_i$，

因此 $\lim\limits_{m\to\infty} \boldsymbol{x}^{(m)} = \boldsymbol{x}$

$$\lim_{m\to\infty} \boldsymbol{x}^{(m)} = \boldsymbol{x} \Leftrightarrow \lim_{m\to\infty} \left\|\boldsymbol{x}^{(m)} - \boldsymbol{x}\right\| = 0$$ （对任一向量范数$\|\bullet\|$成立）

（必要性）若有$\lim_{m\to\infty} \boldsymbol{x}^{(m)} = \boldsymbol{x}$，则由定义，这相当于

$$\xi_i^{(m)} - \xi_i \to 0 \quad (i = 1, 2, \cdots, n)$$

故对任意给定的正数ε，都有正数M_i存在，使得$m > M_i$时，都有

$$\left|\xi_i^{(m)} - \xi_i\right| < \varepsilon \quad (i = 1, 2, \cdots, n)$$

若取$M = \max_{1\le i\le n}\{M_i\}$，则当$m > M$时，对每个$i$值，上述不等式都成立，从而，$m > M$时，

$$\left\|\boldsymbol{x}^{(m)} - \boldsymbol{x}\right\|_\infty = \max_{1\le i\le n}\left|\xi_i^{(m)} - \xi_i\right| < \varepsilon$$

这就证明了

$$\lim_{m\to\infty}\left\|\boldsymbol{x}^{(m)} - \boldsymbol{x}\right\|_\infty = 0$$

命题得证

$\lim\limits_{m\to\infty} \boldsymbol{x}^{(m)} = \boldsymbol{x} \Leftrightarrow \lim\limits_{m\to\infty} \left\|\boldsymbol{x}^{(m)} - \boldsymbol{x}\right\| = 0$ （对任一向量范数$\|\bullet\|$成立）

定理 4.2 表明：向量序列$\left\{\boldsymbol{x}^{(m)}\right\}$收敛于向量$\boldsymbol{x}$，当且仅当对任何一种向量范数$\|\bullet\|$，序列$\left\{\left\|\boldsymbol{x}^{(n)} - \boldsymbol{x}\right\|\right\}$收敛于零。因此$n$维向量序列的收敛问题，借助范数为工具，就归结为实数序列的收敛问题。

■ 定义4.5 矩阵序列的极限

若$A_m = \left(a_{ij}^{(m)}\right) \in \mathrm{C}^{n\times n}$（$m = 1, 2, \cdots$），如果存在极限

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

则称矩阵序列$\{A_m\}$收敛于矩阵$A = \left(a_{ij}\right) \in \mathrm{C}^{n\times n}$，并且记为

$$\lim_{m\to\infty} A_m = A \text{ 或 } A_m \to A \quad (m \to \infty)$$

当矩阵序列不收敛时，也称为发散的。

由此可见，矩阵序列$\{A_m\}$的收敛性，相当于n^2个数值序列$\{a_{ij}^{(m)}\}$（$i,j=1,2,\cdots,n$）的收敛性。

例如，若

$$A_m=\begin{bmatrix}\frac{1}{m} & \frac{\sqrt{2m-1}}{3m+2}\\ 1-\frac{1}{m^2} & \sqrt{-1}\cos\frac{\pi}{m}\end{bmatrix}\in\mathbf{C}^{2\times 2}\quad(m=1,2,\cdots)$$

则有

$$\lim_{m\to\infty}A_m=\begin{bmatrix}0 & \frac{\sqrt{2}}{3}\\ 1 & \sqrt{-1}\end{bmatrix}$$

■ **定理4.3** $\lim\limits_{m\to\infty} A_m = A \Leftrightarrow \lim\limits_{m\to\infty} \|A_m - A\| = 0$（对任一矩阵范数 $\|\bullet\|$ 成立）

证明：

由矩阵范数的等价性，我们只需要对 F-范数进行证明。因此可得

$$\lim_{m\to\infty} A_m = A \Leftrightarrow \lim_{m\to\infty}\left(a_{ij}^{(m)} - a_{ij}\right) = 0(\forall i, j) \Leftrightarrow \lim_{m\to\infty}\sqrt{\sum_{i,j=1}^{n}\left|a_{ij}^{(m)} - a_{ij}\right|^2} = \lim_{m\to\infty}\|A_m - A\|_F = 0$$

命题得证

定理 4.3 表明：矩阵序列 $\{A_m\}$ 收敛于矩阵 A，当且仅当对任一矩阵范数 $\|\bullet\|$，序列 $\{\|A_m - A\|\}$ 收敛于零，特别地 $A_m \to \mathbf{0}$（零矩阵），当且仅当 $\|A_m\| \to 0$（$m \to \infty$）。

$\mathbf{C}^{n\times n}$ 中收敛的矩阵序列有下列基本性质:

(1) 若 $\lim\limits_{m\to\infty} \boldsymbol{A}_m = \boldsymbol{A}$，则对 $\mathbf{C}^{n\times n}$ 中任何矩阵范数 $\|\bullet\|$，$\|\boldsymbol{A}_m\|$ 有界。

证明:

根据矩阵范数的等价性，我们只需要在 F-范数下做出证明。由于 $\boldsymbol{A}_m \to \boldsymbol{A}$，故由定理 4.3，有 $\|\boldsymbol{A}_m - \boldsymbol{A}\|_F \to \mathbf{0}$（零矩阵），所以数列 $\|\boldsymbol{A}_m - \boldsymbol{A}\|_F$ 有界，故有非负实数 C_1，使得

$$\|\boldsymbol{A}_m - \boldsymbol{A}\|_F < C_1$$

从而有

$$\|\boldsymbol{A}_m\|_F = \|\boldsymbol{A}_m - \boldsymbol{A} + \boldsymbol{A}\|_F \le \|\boldsymbol{A}_m - \boldsymbol{A}\|_F + \|\boldsymbol{A}\|_F \le C_1 + \|\boldsymbol{A}\|_F = C_1 + C_2 = C$$

即 $\|\boldsymbol{A}_m\|_F$ 是有界的（这里 $\boldsymbol{A}$ 是给定矩阵，因此 $\|\boldsymbol{A}\|_F = C_2$ 也是常数）

命题得证

这条性质表明：矩阵序列有极限，则矩阵的任一范数都有界。

向量和矩阵的极限

（2）若 $A_m \to A$，$B_m \to B$，又 $a_m \to a$，$b_m \to b$（这里 $\{a_m\}$，$\{b_m\}$ 为数列），则有

$$\boxed{\lim_{m\to\infty}(a_m A_m + b_m B_m) = aA + bB}, \quad \lim_{m\to\infty} A_m B_m = AB$$

思考题

证明： 这两个公式的证明是类似的。我们只出第二式的证明。

因为在任何一种取定的矩阵范数 $\|\cdot\|$ 下

（15分）

$$\begin{aligned}\|A_m B_m - AB\| &= \|A_m B_m - A_m B + A_m B - AB\| \\ &\le \|A_m B_m - A_m B\| + \|A_m B - AB\| \\ &\le \|A_m\| \bullet \|B_m - B\| + \|A_m - A\| \bullet \|B\|\end{aligned}$$

但由性质（1）知 $\|A_m\|$ 有界，又由定理 4.3 知 $\|A_m - A\| \to 0$，$\|B_m - B\| \to 0$，而 $\|B\|$ 是确定的常数，当然是有界的，故由上面的不等式和说明，即有：

对任意给定的正数 ε，都有正数 M 存在，当 $m > M$ 时，$\|A_m B_m - AB\| < \varepsilon$，即

$\lim\limits_{m\to\infty} \|A_m B_m - AB\| = 0$，故 $\lim\limits_{m\to\infty} A_m B_m = AB$

命题得证

这条性质表明：矩阵序列有极限，则矩阵序列的线性组合和乘积都有极限。

（3） 若 $A_m \to A$，且 A_m^{-1} 及 A^{-1} 都存在，则 $\lim\limits_{m\to\infty} A_m^{-1} = A^{-1}$

证明：

对二阶矩阵，可知

$$\lim_{m\to\infty} |A_m| = |A|$$

成立；应用归纳法且按照一行（列）展开，则可证明此式对 n 阶矩阵也成立。

再注意到伴随矩阵 A_m^* 的元素都是 $n-1$ 阶行列式，因此有

$$\lim_{m\to\infty} A_m^{-1} = \lim_{m\to\infty} \frac{1}{|A_m|} A_m^* = \frac{1}{|A|} A^* = A^{-1}$$

命题得证

这条性质表明：矩阵序列有极限，则矩阵的逆组成的矩阵序列也有极限。

下面考察由矩阵 $A\in\mathbf{C}^{n\times n}$ 的幂所形成的序列 $\{A^m\}: E, A, A^2, \cdots, A^m, A^{m+1}, \cdots$，研究其收敛于零矩阵的条件。

■ **定理4.4** $\lim\limits_{m\to\infty} A^m = 0$（零矩阵）的充分条件是有某一矩阵范数 $\|\bullet\|$，使得 $\|A\|<1$

证明：

根据矩阵范数的定义（4），即有

$$\|A^m\| \le \|A^{m-1}\| \bullet \|A\| \le \|A^{m-2}\| \bullet \|A\|^2 \le \cdots \le \|A\|^m$$

因此，若 $\|A\|<1$，则 $A^m \to 0$，从而 $\|A^m\| \to 0$，根据定理 4.3 便得 $\lim\limits_{m\to\infty} A^m = 0$（零矩阵）

命题得证

矩阵序列 $\{A^m\}$ 收敛于零矩阵与矩阵的特征值有着重要联系。

■ **定理4.5** $\lim\limits_{m\to\infty} A^m = 0$（零矩阵）的充分必要条件是 A 的所有特征值的模都小于 1

矩阵特征值与矩形范数间存在一个基本关系

■ 定理4.6

矩阵 A 的每个特征值 λ 的模 $|\lambda|$，都不大于矩阵 A 的任何一种范数 $\|A\|$，即 $|\lambda| \le \|A\|$

证明：

证法二：设 $\|\bullet\|$ 是矩阵的任意一种范数，$\|\bullet\|_a$ 是与之相容的向量范数，若 λ 是矩阵 A 的任一特征值，则有

$$AX = \lambda X\ (X \neq 0) \qquad (4.2)$$

一定存在吗?

（4.2）式两边取向量范数 $\|\bullet\|_a$，则有 $\|AX\|_a = \|\lambda X\|_a$，从而，根据定义 4.1 和定义 4.3 可得

$$|\lambda| \bullet \|X\|_a = \|\lambda X\|_a = \|AX\|_a \le \|A\| \bullet \|X\|_a \qquad (4.3)$$

由于 $X \neq 0$，$\|X\|_a > 0$，（4.3）式可变为

命题得证

$$|\lambda| \le \|A\|$$

矩阵幂级数

1. 向量范数
2. 矩阵范数
3. 向量和矩阵的极限
4. 矩阵幂级数
5. 矩阵函数
6. 矩阵的微分与积分

■ 定义4.6 矩阵级数

若给定 $\mathbf{C}^{n\times n}$ 中一矩阵序列 $\boldsymbol{A}_0$， $\boldsymbol{A}_1$， $\boldsymbol{A}_2$， $\cdots$， $\boldsymbol{A}_m$ $\cdots$，

则和式 $\boldsymbol{A}_0+\boldsymbol{A}_1+\boldsymbol{A}_2+\cdots+\boldsymbol{A}_m+\cdots$

被称为矩阵级数，也常简写为 $\sum\limits_{m=0}^{\infty}\boldsymbol{A}_m$。

令

$$S_N=\sum_{m=0}^{N}\boldsymbol{A}_m$$

若矩阵序列 $\{S_N\}$ 收敛于矩阵 $\boldsymbol{S}$，则称矩阵级数 $\sum\limits_{m=0}^{\infty}\boldsymbol{A}_m$ 收敛，且其和为 $\boldsymbol{S}$，记为 $\boldsymbol{S}=\sum\limits_{m=0}^{\infty}\boldsymbol{A}_m$

显然 $\sum\limits_{m=0}^{\infty}\boldsymbol{A}_m$ 收敛的充要条件是 n^2 个数值级数 $\sum\limits_{m=0}^{\infty}(\boldsymbol{A}_m)_{ij}$ 收敛（$i,j=1,2,\cdots,n$）。这里 $(\boldsymbol{A}_m)_{ij}$ 表示 n 阶矩阵 $\boldsymbol{A}_m$ 位于第 i 行、第 j 列上的元素。

当这 n^2 个数值级数绝对收敛时，则称矩阵级数 $\sum\limits_{m=0}^{\infty}\boldsymbol{A}_m$ 是绝对收敛的。

关于矩阵级数的收敛问题，有下列基本性质：

（1） 若矩阵级数 $\sum_{m=0}^{\infty} A_m$ 绝对收敛，则它一定收敛，且任意交换各项的次序所得的新级数仍收敛，和也不改变。

（2） 矩阵级数 $\sum_{m=0}^{\infty} A_m$ 绝对收敛的充要条件是对任意一种矩阵范数 $\|\bullet\|$，正项级数 $\sum_{m=0}^{\infty} \|A_m\|$ 收敛。

证明：

由于矩阵范数的等价性，如果能在矩阵范数 $\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$ 下证明这一性质，则对其它任何矩阵范数的情形，也容易得知这性质成立。

矩阵幂级数

（充分性）现假设正项级数 $\sum_{m=0}^{\infty}\|A_m\|_1$ 收敛，则由等式 $\|A_m\|_1=\max_{1\le j\le n}\sum_{i=1}^{n}\left|(A_m)_{ij}\right|$

推知 $\left|(A_m)_{ij}\right|\le\|A_m\|_1$ （$i,j=1,2,\cdots,n$）

故由正项级数的比较判别法，得知级数 $\sum_{m=0}^{\infty}\left|(A_m)_{ij}\right|$ （$i,j=1,2,\cdots,n$）

也收敛。因此，矩阵级数 $\sum_{m=0}^{\infty}A_m$ 绝对收敛。

（必要性）若矩阵级数 $\sum_{m=0}^{\infty}A_m$ 绝对收敛，则正项级数 $\sum_{m=0}^{\infty}\left|(A_m)_{ij}\right|$ （$i,j=1,2,\cdots,n$）都收敛，

从而由这 n^2 个级数相加所构成的级数也收敛，即 $\sum_{m=0}^{\infty}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left|(A_m)_{ij}\right|\right)$ 收敛。

但由于 $$\|A_m\|_1=\max_{1\le j\le n}\sum_{i=1}^{n}\left|(A_m)_{ij}\right|\le\sum_{i=1}^{n}\sum_{j=1}^{n}\left|(A_m)_{ij}\right|$$

命题得证

所以，根据正项级数的比较判别法得知级数 $\sum_{m=0}^{\infty}\|A_m\|_1$ 绝对收敛，因而正项级数 $\sum_{m=0}^{\infty}\|A_m\|$ 绝对收敛。

矩阵幂级数

（3） 若 $P,Q\notin\mathrm{C}^{n\times n}$ 为给定矩阵，如果矩阵级数 $\sum\limits_{m=0}^{\infty}A_m$ 收敛（或绝对收敛），则级数 $\sum\limits_{m=0}^{\infty}PA_mQ$ 也收敛（或绝对收敛），且有等式

$$\sum_{m=0}^{\infty}PA_mQ=P\left(\sum_{m=0}^{\infty}A_m\right)Q \tag{4.5}$$

证明：

设 $\sum\limits_{m=0}^{\infty}A_m$ 收敛于方阵 S，即 $\lim\limits_{N\to\infty}\sum\limits_{m=0}^{N}A_m=S=\sum\limits_{m=0}^{\infty}A_m$

而由等式 $\sum\limits_{m=0}^{N}PA_mQ=P\left(\sum\limits_{m=0}^{N}A_m\right)Q$ 取极限，即得 $\lim\limits_{N\to\infty}\sum\limits_{m=0}^{N}PA_mQ=P\left(\lim\limits_{N\to\infty}\sum\limits_{m=0}^{N}A_m\right)Q=PSQ$

即 $\sum\limits_{m=0}^{\infty}PA_mQ$ 收敛，且有（4.5）式成立。

现设 $\sum_{m=0}^{\infty} A_m$ 绝对收敛，则由性质(2)可知，$\sum_{m=0}^{\infty} \|A_m\|$ 也收敛。又因为 $\|PA_mQ\| \le \|P\| \cdot \|A_m\| \cdot \|Q\|$

$\|P\|$、$\|Q\|$ 都是有界实数，所以由比较判别法即知正项级数 $\sum_{m=0}^{\infty} \|PA_mQ\|$ 收敛。故根据性质(2)

可知矩阵级数 $\sum_{m=0}^{\infty} PA_mQ$ 绝对收敛。

命题得证

注意：

基本性质（3）对于 P，Q 分别为 n 维行向量与列向量的情形，可以证明也能成立。

■ 定义 方阵幂级数

若已知 n 阶复数方阵序列 $\{A^m\}$ 及复数序列 $\{c_m\}$，则方阵级数

$$\sum_{m=0}^{\infty} c_m A^m$$

被称为方阵 A 的幂级数。

方阵幂级数是矩阵幂级数的特例，因此矩阵幂级数的收敛性质也适用于方阵幂级数。

■ 定义 方阵谱半径

如果 λ_1， λ_2， $\cdots$， λ_n 为方阵 $A \in \mathrm{C}^{n\times n}$ 的全部特征值，则

$$\rho(A) = \max_{1\le i\le n}\{|\lambda_i|\}$$

被称为方阵 A 的谱半径。

■ **定理4.7**

若 $A\in\mathbf{C}^{n\times n}$，则对于任意给定的正数 ε，都有某一方阵范数 $\|\bullet\|$，使得

$$\|A\|\le\rho(A)+\varepsilon$$

■ **定理4.8**

若复变数幂级数 $\sum_{m=0}^{\infty}c_m z^m$ 的收敛半径为 R，而矩阵 $A\in\mathbf{C}^{n\times n}$ 的谱半径为 $\rho(A)$，则：

（1） 当 $\rho(A)<R$ 时，方阵幂级数 $\sum_{m=0}^{\infty}c_m A^m$ 绝对收敛；

（2） 当 $\rho(A)>R$ 时，方阵幂级数 $\sum_{m=0}^{\infty}c_m A^m$ 发散。

■ 推论 1

若复变数幂级数 $\sum_{m=0}^{\infty} c_m (z-\lambda_0)^m$ 的收敛半径为 R，则对于方阵 $A \in \mathbf{C}^{n\times n}$，当其特征值 λ_1，λ_2，…，λ_n 满足

$$|\lambda_i - \lambda_0| < R \quad (i=1,2,\cdots,n)$$

时，方阵幂级数

$$\sum_{m=0}^{\infty} c_m (A-\lambda_0 E)^m \tag{4.6}$$

绝对收敛；若有某一个 λ_i 使得 $|\lambda_i - \lambda_0| > R$，则方阵幂级数（4.6）发散。

■ 推论 2

若复变数幂级数 $\sum_{m=0}^{\infty} c_m z^m$ 在整个复平面上都收敛，则对任意的方阵 $A \in \mathbf{C}^{n\times n}$，方阵幂级数 $\sum_{m=0}^{\infty} c_m A^m$ 也收敛。

矩阵函数

1 向量范数

2 矩阵范数

3 向量和矩阵的极限

4 矩阵幂级数

5 矩阵函数

6 矩阵的微分与积分

在复变函数中，复变数幂级数

$$\sum_{m=0}^{\infty}\frac{z^m}{m!},\quad \sum_{m=1}^{\infty}(-1)^{m-1}\frac{z^{2m-1}}{(2m-1)!},\quad 1+\sum_{m=1}^{\infty}(-1)^m\frac{z^{2m}}{(2m)!}$$

在整个复平面上收敛，因而它们都有确定的和，并一次用 e^z，$\sin z$，$\cos z$ 来表示。

复变数初等函数 e^z，$\sin z$，$\cos z$ 可以用幂级数来定义：

$$e^z=\sum_{m=0}^{\infty}\frac{z^m}{m!}$$

$$\sin z=\sum_{m=1}^{\infty}(-1)^{m-1}\frac{z^{2m-1}}{(2m-1)!}$$

$$\cos z=1+\sum_{m=1}^{\infty}(-1)^m\frac{z^{2m}}{(2m)!}$$

利用定理 4.8 的推论 2，可知对任何方阵 $A \in \mathbf{C}^{n\times n}$，下列各方阵幂级数都收敛

$$\sum_{m=0}^{\infty} \frac{A^m}{m!}, \quad \sum_{m=1}^{\infty} (-1)^{m-1} \frac{A^{2m-1}}{(2m-1)!}, \quad E + \sum_{m=1}^{\infty} (-1)^m \frac{A^{2m}}{(2m)!}$$

因此，可用它们来定义下列三个方阵函数：

$$e^A = \sum_{m=0}^{\infty} \frac{A^m}{m!}$$

$$\sin A = \sum_{m=1}^{\infty} (-1)^{m-1} \frac{A^{2m-1}}{(2m-1)!}$$

$$\cos A = E + \sum_{m=1}^{\infty} (-1)^m \frac{A^{2m}}{(2m)!}$$

分别被称为方阵 A 的指数函数、正弦函数和余弦函数

矩阵函数

同样地，由

$$\ln(1+z)=\sum_{m=1}^{\infty}\frac{(-1)^{m-1}}{m}z^m \quad（|z|<1）$$

$$(1+z)^a=1+\sum_{m=1}^{\infty}\frac{a(a-1)\cdots(a-m+1)}{m!}z^m \quad（|z|<1）$$

（a为任意实数）

可定义方阵函数$\ln(1+A)$和$(1+A)^a$为：

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

$$(E+A)^a=E+\sum_{m=1}^{\infty}\frac{a(a-1)\cdots(a-m+1)}{m!}A^m \quad（\rho(A)<1）$$

特别地，$(E-A)^{-1}=\sum_{m=0}^{\infty}A^m$。

思考题 （15分）

一般地，若复变数幂级数 $\sum_{m=0}^{\infty} c_m z^m$ 的收敛半径为 R，其和为 $f(z)$，即

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

则由定理 4.8 就可以定义矩阵函数

$$f(A)=\sum_{m=0}^{\infty} c_m A^m \quad (\rho(A)<R, \quad A\in \mathbf{C}^{n\times n})$$

现在的问题是如何求解矩阵函数 $f(A)$！本节主要讨论用矩阵 A 的标准形来计算矩阵函数 $f(A)$ 的方法，同时也简单介绍用多项式来表示矩阵函数 $f(A)$ 的方法。

■ **定理4.9**

若对任一方阵 X，幂级数 $\sum_{m=0}^{\infty} c_m X^m$ 都收敛，其和为 $f(X)=\sum_{m=0}^{\infty} c_m X^m$，则当 X 为分块对角矩阵

$$X=\begin{bmatrix} X_1 & & & \\ & X_2 & & \\ & & \ddots & \\ & & & X_k \end{bmatrix}$$

时，即有

$$f(X)=\begin{bmatrix} f(X_1) & & & \\ & f(X_2) & & \\ & & \ddots & \\ & & & f(X_k) \end{bmatrix}$$

证明：

$$f(X)=\sum_{m=0}^{\infty}c_m X^m=\lim_{N\to\infty}\sum_{m=0}^{N}c_m\begin{bmatrix}X_1{}^m & & & \\ & X_2{}^m & & \\ & & \ddots & \\ & & & X_k{}^m\end{bmatrix}$$

$$=\begin{bmatrix}\lim\limits_{N\to\infty}\sum\limits_{m=0}^{N}c_m X_1{}^m & & & \\ & \lim\limits_{N\to\infty}\sum\limits_{m=0}^{N}c_m X_2{}^m & & \\ & & \ddots & \\ & & & \lim\limits_{N\to\infty}\sum\limits_{m=0}^{N}c_m X_k{}^m\end{bmatrix}$$

$$=\begin{bmatrix}f(X_1) & & & \\ & f(X_2) & & \\ & & \ddots & \\ & & & f(X_k)\end{bmatrix}$$

命题得证

■ 定理4.10

若

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

是收敛半径为R的复变数幂级数，又

$$J_0=\begin{bmatrix}\lambda_0 & & & \\ 1 & \lambda_0 & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_0\end{bmatrix}$$

是n阶约当块，则当$|\lambda_0|<R$时，方阵幂级数$\sum_{m=0}^{\infty}c_m J_0^m$绝对收敛，且其和为

$$f(J_0)=\begin{bmatrix} f(\lambda_0) & & & & \\ f'(\lambda_0) & \ddots & & & \\ \frac{1}{2!}f''(\lambda_0) & \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \ddots & \\ \frac{1}{(n-1)!}f^{(n-1)}(\lambda_0) & \cdots & \frac{1}{2!}f''(\lambda_0) & f'(\lambda_0) & f(\lambda_0) \end{bmatrix}$$

矩阵函数

现在，可以用矩阵 A 的标准形来计算矩阵函数 $f(A)$ 了。按照 A 是否相似于对角形矩阵，分两种情形来讨论：

（1） 若 A 相似于对角形矩阵

$$A = P\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}P^{-1}$$

简记为 $A = PJP^{-1}$，这里 λ_i 是 A 的特征值（$i = 1, 2, \cdots, n$）。

（2） 当 A 不能与对角形矩阵相似时，A 必可与其约当标准形相似：

$$A = P\begin{bmatrix} J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_k(\lambda_k) \end{bmatrix}P^{-1}$$

其中约当块 $J_i = \begin{bmatrix} \lambda_i & & & \\ 1 & \lambda_i & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_i \end{bmatrix}$

（1） 若 A 相似于对角形矩阵

$$A = P\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}P^{-1}$$

简记为 $A = PJP^{-1}$，这里 λ_i 是 A 的特征值（$i = 1, 2, \cdots, n$）。

根据定理 4.8，如果复变数幂级数 $\sum_{m=0}^{\infty} c_m z^m$ 的收敛半径为 R，则当 $|z| < R$ 时，此级数绝对收敛，其和设为 $f(z)$。又当方阵 A 的谱半径 $\rho(A) < R$ 时，方阵幂级数 $\sum_{m=0}^{\infty} c_m A^m$ 也绝对收敛，且其和为

$$f(A) = \sum_{m=0}^{\infty} c_m A^m = f\left(PJP^{-1}\right) = \sum_{m=0}^{\infty} c_m \left(PJP^{-1}\right)^m$$

$$= \sum_{m=0}^{\infty} c_m PJ^m P^{-1} = P\left(\sum_{m=0}^{\infty} c_m J^m\right)P^{-1} = Pf(J)P^{-1}$$

根据定理 4.9 得

$$f(A)=P\begin{bmatrix} f(\lambda_1) & & & \\ & f(\lambda_2) & & \\ & & \ddots & \\ & & & f(\lambda_n) \end{bmatrix}P^{-1} \qquad (4.11)$$

并且可见 $f(A)$ 的特征值为 $f(\lambda_1)$，$f(\lambda_2)$，…，$f(\lambda_n)$，因而，在 A 相似于对角形矩阵时，与复变数幂级数

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

相应的矩阵函数

$$f(A)=\sum_{m=0}^{\infty}c_m A^m \quad (\rho(A)<R)$$

也相似于一个对角形矩阵，而矩阵函数 $f(A)$ 就是一个由（4.11）式确定的矩阵。

特别地，有

$$e^{A}=P\begin{bmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{bmatrix}P^{-1} \qquad \sin A=P\begin{bmatrix} \sin\lambda_1 & & & \\ & \sin\lambda_2 & & \\ & & \ddots & \\ & & & \sin\lambda_n \end{bmatrix}P^{-1}$$

$$\cos A=P\begin{bmatrix} \cos\lambda_1 & & & \\ & \cos\lambda_2 & & \\ & & \ddots & \\ & & & \cos\lambda_n \end{bmatrix}P^{-1}$$

这里只需要分别取 $f(z)=e^{z}$，$\sin z$，$\cos z$，并假定 A 相似于对角形矩阵（此外，对 A 再无其它要求），由此也可以看出 e^{A}，$\sin A$，$\cos A$ 这三个矩阵的特征值分别是 e^{λ_i}，$\sin\lambda_i$，$\cos\lambda_i$（$i=1,2,\cdots,n$）。

例4.3

设

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

求 e^{A}， $\sin A$， $\cos A$。

解：

因

$$|\lambda E-A|=\begin{vmatrix}\lambda & -1\\ 0 & \lambda-2\end{vmatrix}=\lambda(\lambda-2)$$

故 A 有不同的特征值 $\lambda_1=0$， $\lambda_2=2$，所以 A 能相似于对角形矩阵，不难求得相应于两个特征值的特征向量分别是

$$a_1=\begin{bmatrix}1\\ 0\end{bmatrix},\quad a_2=\begin{bmatrix}1\\ 2\end{bmatrix}$$

令

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

则有

$$P^{-1}=\frac{1}{2}\begin{bmatrix}2 & -1\\ 0 & 1\end{bmatrix},\quad A=P\begin{bmatrix}0 & \\ & 2\end{bmatrix}P^{-1}$$

因此有

$$e^{A}=P\begin{bmatrix}e^{0} & \\ & e^{2}\end{bmatrix}P^{-1}=\begin{bmatrix}1 & -\frac{1}{2}+\frac{1}{2}e^{2}\\ 0 & e^{2}\end{bmatrix}$$

$$\sin A=P\begin{bmatrix}\sin 0 & \\ & \sin 2\end{bmatrix}P^{-1}=\begin{bmatrix}0 & \frac{1}{2}\sin 2\\ 0 & \sin 2\end{bmatrix}$$

$$\cos A=P\begin{bmatrix}\cos 0 & \\ & \cos 2\end{bmatrix}P^{-1}=\begin{bmatrix}1 & -\frac{1}{2}+\frac{1}{2}\cos 2\\ 0 & \cos 2\end{bmatrix}$$

（2） 当 $\boldsymbol{A}$ 不能与对角形矩阵相似时，$\boldsymbol{A}$ 必可与其约当标准形相似：

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

其中约当块

$$\boldsymbol{J}_i=\begin{bmatrix} \lambda_i & & & \\ 1 & \lambda_i & & \\ & \ddots & \ddots & \\ & & 1 & \lambda_i \end{bmatrix}_{n_i\times n_i}$$

由初级因子 $(\lambda-\lambda_i)^{n_i}$ 所决定。

又 $n_1+n_2+\cdots+n_k=n$，因此，应用定理 4.9 可得

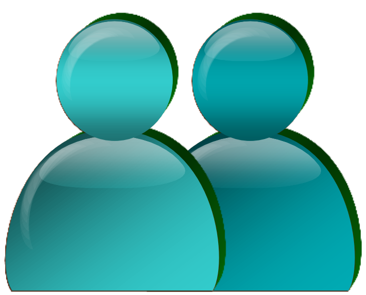
$$f(A)=\sum_{m=0}^{\infty}c_m A^m=\sum_{m=0}^{\infty}c_m P\begin{bmatrix} J_1^m & & & \\ & J_2^m & & \\ & & \ddots & \\ & & & J_k^m \end{bmatrix}P^{-1}$$

$$=P\left(\sum_{m=0}^{\infty}c_m\begin{bmatrix} J_1^m & & & \\ & J_2^m & & \\ & & \ddots & \\ & & & J_k^m \end{bmatrix}\right)P^{-1}=P\begin{bmatrix} f(J_1) & & & \\ & f(J_2) & & \\ & & \ddots & \\ & & & f(J_k) \end{bmatrix}P^{-1}$$

再参照定理 4. 10 的公式计算出每个 $f(J_i)$（$i=1,2,\cdots,k$），从而可得到 $f(A)$。

矩阵的微分与积分

1. 向量范数
2. 矩阵范数
3. 向量和矩阵的极限
4. 矩阵幂级数
5. 矩阵函数
6. 矩阵的微分与积分



■ 定义 函数矩阵的导数

若函数矩阵 $\boldsymbol{A}(z)=\left(a_{ij}(z)\right)_{m\times n}$ 的每个元素 $a_{ij}(z)$ 都是复变量 z 的函数，且都在 $z=z_0$ 或变量 z 的某个区域内可导，则定义 $\boldsymbol{A}(z)$ 的导数为

$$\frac{d}{dz}\boldsymbol{A}(z)=\left(\frac{d}{dz}a_{ij}(z)\right)_{m\times n}$$

简单记为 $\boldsymbol{A}'(z)$。

例如

$$\boldsymbol{A}(z)=\begin{bmatrix} z^2+1 & \sin z \\ 3 & e^z \end{bmatrix}$$

则

$$\frac{d}{dz}\boldsymbol{A}(z)=\begin{bmatrix} 2z & \cos z \\ 0 & e^z \end{bmatrix}$$

矩阵导数的性质：

（1）$\left[A(z)+B(z)\right]' = A'(z)+B'(z)$

（2）$\left[A(z)\bullet B(z)\right]' = A'(z)\bullet B(z)+A(z)\bullet B'(z)$

特别地，当 C 为常数矩阵时，有 $C'=\mathbf{0}$（零矩阵），且

$$\left[C\bullet A(z)\right]' = C\bullet A'(z)$$

（3）如果 $A(u)=\left(a_{ij}(u)\right)_{m\times n}$ 和变量 z 的函数 $u=f(z)$ 都可导，则

$$\frac{d}{dz}A(u)=\frac{dA(u)}{du}\bullet\frac{du}{dz}$$

（4）若 n 阶矩阵函数 $A(z)$ 可逆，且 $A(z)$ 及其逆矩阵 $A^{-1}(z)$ 都可导，则

$$\frac{dA^{-1}(z)}{dz}=-A^{-1}(z)\bullet\frac{dA(z)}{dz}\bullet A^{-1}(z)$$

证明：

因为 $$A^{-1}(z)\bullet A(z)=E$$

两边对 z 求导，即有

$$\frac{d}{dz}\left[A^{-1}(z)\bullet A(z)\right]=\frac{dA^{-1}(z)}{dz}\bullet A(z)+A^{-1}(z)\bullet\frac{dA(z)}{dz}=0$$ （零矩阵）

所以 $$\frac{dA^{-1}(z)}{dz}=-A^{-1}(z)\bullet\frac{dA(z)}{dz}\bullet A^{-1}(z)$$ **命题得证**

■ 定义 函数矩阵的积分

若函数矩阵 $\boldsymbol{A}(x)=\left(a_{ij}(x)\right)_{m\times n}$ 的每个元素 $a_{ij}(x)$ 都是实变量 x 的函数，且都在 $[a,b]$ 上可积，则 $\boldsymbol{A}(x)$ 的定积分与不定积分定义如下

$$\int_a^b \boldsymbol{A}(x)dx=\begin{bmatrix}\int_a^b a_{11}(x)dx & \cdots & \int_a^b a_{1n}(x)dx\\ \vdots & & \vdots\\ \int_a^b a_{m1}(x)dx & \cdots & \int_a^b a_{mn}(x)dx\end{bmatrix}$$

$$\int \boldsymbol{A}(x)dx=\begin{bmatrix}\int a_{11}(x)dx & \cdots & \int a_{1n}(x)dx\\ \vdots & & \vdots\\ \int a_{m1}(x)dx & \cdots & \int a_{mn}(x)dx\end{bmatrix}$$

简言之，矩阵的积分定义为每个元素的积分。

矩阵的微分与积分

易证以下各个性质成立：

（1）$\int A^T(x)dx = \left(\int A(x)dx\right)^T$；

（2）$\int [aA(x)+bB(x)]dx = a\int A(x)dx + b\int B(x)dx$（$a$，$b$ 为非零实数）

（3）$\int C\bullet A(x)dx = C\bullet\int A(x)dx$（$C$ 为非零常数矩阵）

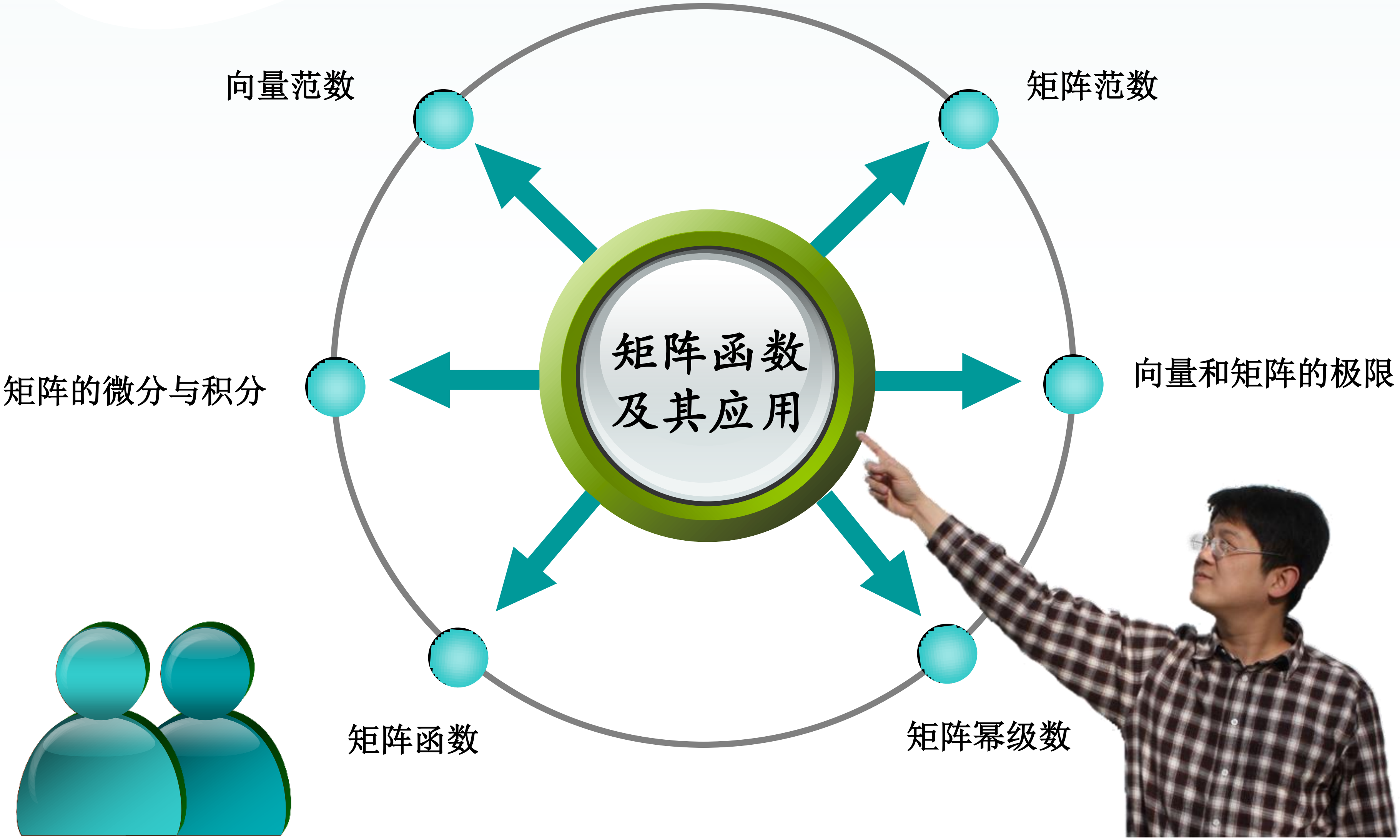
（4）$\int A(x)\bullet B'(x)dx = A(x)\bullet B(x) - \int A'(x)\bullet B(x)dx$

思考题　（15分）

设 $A(t)=\begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$

求 $\frac{d}{dt}A(t)$，$\frac{d}{dt}A^{-1}(t)$，$\frac{d}{dt}\left|A(t)\right|$，$\left|\frac{d}{dt}A(t)\right|$

总结



人民语录

使我们不快乐的大都是一些芝麻小事，因为我们可以躲闪一头大象，却很难躲开一只苍蝇

To be continued…

wangyong@ucas.ac.cn

http://people.ucas.ac.cn/~wangyong

Thank you !