此前我们只研究了矩阵的代数运算，但在工程实际中，特别是涉及到多元分析时，还要用到矩阵的分析运算。

同微积分理论一样，矩阵分析的理论建立，也是以极限理论为基础的，其内容丰富，是研究数值方法和其它数学分支的重要工具。

本章讨论矩阵序列的极限运算，然后介绍矩阵序列和矩阵级数收敛的定理，矩阵幂级数的极限运算和一些矩阵函数，如 **sin***A***,cos***A***,** e^A等，最后介绍矩阵的微积分。

3.1、矩阵序列与矩阵级数

3.2、矩阵幂级数

3.3、函数矩阵的微积分

3.1.1. 矩阵序列

定义3.1 $\{A_k\}_{k=1}^{\infty}$为 $\mathbf{C}^{m\times n}$中的矩阵序列，其中 $A_k=\left(a_{ij}^{(k)}\right)$ 又 $A=\left(a_{ij}\right)\in\mathbf{C}^{m\times n}$。如果 $\lim\limits_{k\to\infty}a_{ij}^{(k)}=a_{ij}$ 对 $i=1, 2, \cdots, m$, $j=1, 2,\cdots,n$ 均成立，则称矩阵序列 $\{A_k\}_{k=1}^{\infty}$ 收敛，而A称为矩阵序列 $\{A_k\}_{k=1}^{\infty}$ 的极限，记为 $\lim\limits_{k\to\infty}A_k=A$。不收敛的矩阵序列称为发散的。

例1 讨论矩阵序列 $\{\boldsymbol{A}_k\}_{k=1}^{\infty}$ 的收敛性。其中

解：根据定义，只须求出它的每一个元素的极限即可， 因此它的极限为：

$$\lim_{k\to\infty}\boldsymbol{A}_k=\begin{pmatrix}\lim\limits_{k\to\infty}\left(1+\frac{1}{k}\right)^k & \lim\limits_{k\to\infty}\frac{\sin k}{k}\\ \lim\limits_{k\to\infty}1 & \lim\limits_{k\to\infty}e^{-k}\\ \lim\limits_{k\to\infty}\frac{2+k}{k} & \lim\limits_{k\to\infty}\sqrt[k]{k}\end{pmatrix}=\begin{pmatrix}e & 0\\ 1 & 0\\ 1 & 1\end{pmatrix}=\boldsymbol{A}$$

由矩阵序列极限的定义可以看出，矩阵序列收敛的性质和数列收敛性质相似。

由定义可见，$\mathbf{C}^{m\times n}$中的矩阵序列的收敛相当于mn个数列同时收敛。因此可以用初等分析的方法来研究它。

但同时研究mn个数列极限未免繁琐，我们可以利用矩阵范数来研究矩阵序列的极限。

定理3.1 设 $\{A_k\}_{k=1}^{\infty}$ 为 $C^{m\times n}$ 中的矩阵序列, $\|\cdot\|$ 为 $C^{m\times n}$ 中的一种矩阵范数， 则矩阵序列 $\{A_k\}_{k=1}^{\infty}$ 收敛于矩阵A的充要条件是 $\|A_k - A\|$ 收敛于零。

证：首先, 利用范数的等价性知, 对于$C^{m\times n}$ 中的任意两个矩阵范数 $\|\cdot\|_t$ 和 $\|\cdot\|_s$ ，存在常数$c_1 \geqslant c_2 > 0$，使得

$$c_2 \cdot \|A_k - A\|_t \le \|A_k - A\|_s \le c_1 \cdot \|A_k - A\|_t$$

即有

$$\lim_{k\to\infty} \|A_k - A\|_t = 0 = \lim_{k\to\infty} \|A_k - A\|_s$$

即收敛于零是一致的。

因此, 只需证明定理对一种特定的矩阵范数成立即可。

我们选取∞−范数加以证明。 根据∞−范数的定义，对于

$1\leqslant i\leqslant m$, $1\leqslant i\leqslant m$均有

$$0 \longleftarrow \left| a_{ij}^{(k)} - a_{ij} \right| \le \max_{1\le i\le m} \left\{ \sum_{j=1}^{n} \left| a_{ij}^{(k)} - a_{ij} \right| \right\} = \left\| \boldsymbol{A}_k - \boldsymbol{A} \right\|_\infty$$

$$\le \sum_{i=1}^{m} \sum_{j=1}^{n} \left| a_{ij}^{(k)} - a_{ij} \right| \le mn \cdot \max_{\substack{1\le i\le m \\ 1\le j\le n}} \left| a_{ij}^{(k)} - a_{ij} \right| \longrightarrow 0$$

因此，

$$\lim_{k\to\infty} \boldsymbol{A}_k = \boldsymbol{A} \iff \lim_{k\to\infty} \left\| \boldsymbol{A}_k - \boldsymbol{A} \right\|_\infty = 0$$

证毕

推论3.1 设 $\{\boldsymbol{A}_k\}_{k=1}^{\infty}, \boldsymbol{A} \in \mathbf{C}^{m\times n}$，并且 $\lim\limits_{k\to\infty} \boldsymbol{A}_k = \boldsymbol{A}$，则 $\lim\limits_{k\to\infty}\|\boldsymbol{A}_k\| = \|\boldsymbol{A}\|$

证：由 $\left|\|\boldsymbol{A}_k\| - \|\boldsymbol{A}\|\right| \le \|\boldsymbol{A}_k - \boldsymbol{A}\|$ 即结论成立。

需要指出的是，此结论只是充分条件，反过来不一定成立。

给定矩阵序列 $\boldsymbol{A}_k = \begin{pmatrix} (-1)^k & \frac{1}{k+1} \\ 1 & 2 \end{pmatrix}$ 和矩阵 $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$

显然 $\lim\limits_{k\to\infty}\|\boldsymbol{A}_k\|_F = \lim\limits_{k\to\infty}\sqrt{\left|-1\right|^k + 1^2 + 2^2 + \frac{1}{(k+1)^2}} = \sqrt{6} = \|\boldsymbol{A}\|_F$

但是矩阵序列$\boldsymbol{A}_k$不收敛，故更不收敛于矩阵 $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$。

性质3.1 设 $\{\boldsymbol{A}_k\}_{k=1}^{\infty}$ 和 $\{\boldsymbol{B}_k\}_{k=1}^{\infty}$ 为 $\mathbf{C}^{m\times n}$ 中的矩阵序列，并且

$$\lim_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A},\quad \lim_{k\to\infty}\boldsymbol{B}_k=\boldsymbol{B}$$

则

$$\lim_{k\to\infty}(\alpha\boldsymbol{A}_k+\beta\boldsymbol{B}_k)=\alpha\boldsymbol{A}+\beta\boldsymbol{B},\quad \forall\alpha,\beta\in\mathbf{C}$$

证 由 $\|(\alpha A_k+\beta B_k)-(\alpha A+\beta B)\|\le|\alpha|\|A_k-A\|+|\beta|\|B_k-B\|$

由定理3.1，即结论成立。

性质3.2 设 $\{\boldsymbol{A}_k\}_{k=1}^{\infty}$ 和 $\{\boldsymbol{B}_k\}_{k=1}^{\infty}$ 分别为 $\mathbf{C}^{m\times n}$ 和 $\mathbf{C}^{n\times l}$ 中的矩阵序列，

并且

$$\lim_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A},\quad \lim_{k\to\infty}\boldsymbol{B}_k=\boldsymbol{B}$$

则

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

证 由

$$\|A_kB_k-AB\|\le\|A_k\|\|B_k-B\|+\|A_k-A\|\cdot\|B\|$$

由定理3.1和推论可知，结论成立。

性质3. 3 设 $\{A_k\}_{k=1}^{\infty}\in \mathbf{C}^{n\times n}$中的矩阵序列，$\lim\limits_{k\to\infty} A_k = A$ 并且 $A_k\,(k=1,2,\cdots)$ 和 $A\in \mathbf{C}^{n\times n}$均为可逆矩阵，则 $\lim\limits_{k\to\infty} A_k^{-1} = A^{-1}$ 。

证 因为$(A_k)^{-1}$和A^{-1}存在，所以 $\lim\limits_{k\to\infty}\det\left(A_k\right)=\det\left(A\right)\neq 0$，又有 $\lim\limits_{k\to\infty}\tilde{A}_k = \lim\limits_{k\to\infty}\tilde{A}\neq 0$，

$$\tilde{A}_k = \begin{pmatrix} \det\left(\bar{A}_{11}^{(k)}\right) & \det\left(\bar{A}_{21}^{(k)}\right) & \cdots & \det\left(\bar{A}_{n1}^{(k)}\right) \\ \det\left(\bar{A}_{12}^{(k)}\right) & \det\left(\bar{A}_{22}^{(k)}\right) & \cdots & \det\left(\bar{A}_{n2}^{(k)}\right) \\ \vdots & \vdots & \cdots & \vdots \\ \det\left(\bar{A}_{1n}^{(k)}\right) & \det\left(\bar{A}_{2n}^{(k)}\right) & \cdots & \det\left(\bar{A}_{nn}^{(k)}\right) \end{pmatrix}$$

其中$\det\left(\left(\bar{A}_{ij}^{(k)}\right)_{(n-1)\times(n-1)}\right)\, i, j=1,2,\cdots,n$ 为A_k的第ij个代数余子式。

于是，$$\lim_{k\to\infty}\left(A_k\right)^{-1} = \lim_{k\to\infty}\frac{\tilde{A}_k}{\det(A_k)} = \frac{\tilde{A}}{\det(A)} = A^{-1}$$

注意，性质3.3中条件$A_k(k=1,2,\cdots)$和A均为可逆的是不可少的。

因为即使$A_k(k=1, 2, \cdots)$可逆也不能保证 A一定可逆。

例如， $A_k=\begin{pmatrix} 1+\frac{1}{k} & 1 \\ 1 & 1 \end{pmatrix}$， 对于 A_k（$k=1, 2,\cdots$） 都有

$$(A_k)^{-1}=\begin{pmatrix} k & -k \\ -k & k+1 \end{pmatrix}$$

但是 $\lim\limits_{k\to\infty} A_k=\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}=A$ 不可逆。

一般的矩阵序列 $\{\boldsymbol{A}_k\}_{k=1}^{\infty}\in\mathbf{C}^{m\times n}$， 即 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_k,\cdots$

在矩阵序列中，最常见的是由一个方阵的幂构成的序列：

$\{\boldsymbol{A}^k\}_{k=0}^{\infty}\in\mathbf{C}^{n\times n}$， 即 $\boldsymbol{I},\boldsymbol{A},\boldsymbol{A}^2,\cdots,\boldsymbol{A}^k,\cdots$

关于这样的矩阵序列有以下的概念和收敛定理。

定义 设 $\boldsymbol{A}\in\mathbf{C}^{\boldsymbol{n}\times\boldsymbol{n}}$， 若 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=\mathbf{0}$，则称$\boldsymbol{A}$为收敛矩阵。

例 2 设 $\boldsymbol{A}\in\mathbf{C}^{\boldsymbol{n}\times\boldsymbol{n}}$，证明 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=\mathbf{0}$ 的充分必要条件是 $\rho(\boldsymbol{A})<1$。

证 必要性 由定理3.1知 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=\mathbf{0}$ 的充分必要条件是对任意一种矩阵范数 $\|\cdot\|$ 均有 $\lim\limits_{k\to\infty}\left\|\boldsymbol{A}^k-\mathbf{0}\right\|=\lim\limits_{k\to\infty}\left\|\boldsymbol{A}^k\right\|=\mathbf{0}$。因此对充分大的$k$，必有 $\left\|\boldsymbol{A}^k\right\|<1$，利用矩阵谱半径的定义以及相容矩阵范数的性质有：

$$\left(\rho(\boldsymbol{A})\right)^k=\rho\left(\boldsymbol{A}^k\right)\le\left\|\boldsymbol{A}^k\right\|<1$$

因此得 $\rho(\boldsymbol{A})<1$。

充分性 根据定理2.9，对于 $\varepsilon=\frac{1}{2}\left(1-\rho\left(\boldsymbol{A}\right)\right)>0$ 一定存在一种相容的矩阵范数 $\|\cdot\|$，使得 $\|\boldsymbol{A}\|\le\rho(\boldsymbol{A})+\varepsilon$。

又根据相容矩阵范数的性质有， 再注意到上述关系式中

$$\rho\left(\boldsymbol{A}\right)+\varepsilon=\frac{1}{2}\left(1+\rho\left(\boldsymbol{A}\right)\right)<1$$

那么

$$\left\|\boldsymbol{A}^k\right\|\le\|\boldsymbol{A}\|^k\le\left(\rho(\boldsymbol{A})+\varepsilon\right)^k\le q^k<1\qquad\left(0<q<1\right)$$

于是，$\lim\limits_{k\to\infty}\left\|\boldsymbol{A}^k\right\|=0$ 根据定理3.1 即知 $\lim\limits_{k\to\infty}\boldsymbol{A}^k=0$。

推论 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$，若对$\mathbf{C}^{n\times n}$上的某种范数 $\|\cdot\|$，有 $\|\boldsymbol{A}\|<1$ 则

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

练习题 判断对下列矩阵是否有 $\lim\limits_{k\to\infty} \boldsymbol{A}^k = 0$

（1）$\boldsymbol{A} = \dfrac{1}{6}\begin{pmatrix} 1 & -8 \\ -2 & 1 \end{pmatrix}$，（2）$\boldsymbol{A} = \begin{pmatrix} 0.2 & 0.1 & 0.2 \\ 0.5 & 0.5 & 0.4 \\ 0.1 & 0.3 & 0.2 \end{pmatrix}$

解：（1）取 $\boldsymbol{B} = \begin{pmatrix} 1 & -8 \\ -2 & 1 \end{pmatrix}$，则 $\lambda(\boldsymbol{A}) = \dfrac{1}{6}\lambda(\boldsymbol{B})$，令

$$\det(\lambda \boldsymbol{I} - \boldsymbol{B}) = \begin{vmatrix} \lambda - 1 & 8 \\ 2 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^2 - 16 = (\lambda - 5)(\lambda + 3) = 0$$

得 $\lambda_1(\boldsymbol{B}) = 5$，$\lambda_2(\boldsymbol{B}) = -3$，进而得 $\lambda_1(\boldsymbol{A}) = \dfrac{5}{6}$，$\lambda_2(\boldsymbol{A}) = -\dfrac{1}{2}$。

于是，$\rho(\boldsymbol{A}) = \dfrac{5}{6} < 1$ 故 $\lim\limits_{k\to\infty} \boldsymbol{A}^k = \boldsymbol{0}$。

（2）因为 $\|\boldsymbol{A}\|_1 = 0.9 < 1$，由推论，故 $\lim\limits_{k\to\infty} \boldsymbol{A}^k = \boldsymbol{0}$

3.1.2 矩阵级数

定义3.2 设 $\{A_k\}_{k=1}^{\infty}$ 为$\mathbf{C}^{m\times n}$中的矩阵序列，称

$$A_1+A_2+\cdots+A_k+\cdots$$

为由矩阵序列 $\{A_k\}_{k=1}^{\infty}$ 构成的矩阵级数, 记为 $\sum_{k=1}^{\infty}A_k$。

定义3.3 记 $S_k=\sum_{i=1}^{k}A_i$, 称之为矩阵级数 $\sum_{k=1}^{\infty}A_k$ 的前k项部分和。若矩阵序列 $\{S_k\}_{k=1}^{\infty}$ 收敛且 $\lim_{k\to\infty}S_k=S$, 则称矩阵级数 $\sum_{k=1}^{\infty}A_k$ 收敛, 而矩阵S称为矩阵级数的和矩阵, 记为 $S=\sum_{k=1}^{\infty}A_k$。 不收敛的矩阵级数称为发散的。

显然，和 $\sum_{k=1}^{\infty}A_k=S=(s_{ij})$的意义指的是: $\sum_{k=1}^{\infty}a_{ij}^{(k)}=s_{ij}$

$(i=1,2,\cdots,m,\ j=1,2,\cdots,n)$ 即$m\times n$个数项级数 $\sum_{k=1}^{\infty}a_{ij}^{(k)}$ 均为收敛的。

练习 研究矩阵级数 $\sum_{k=1}^{\infty}\boldsymbol{A}_k$ 的收敛性。其中

解：因为

$$\boldsymbol{S}_N=\sum_{k=0}^{N}\boldsymbol{A}_k=\begin{pmatrix}\sum_{k=0}^{N}\frac{1}{(k+1)(k+2)} & \sum_{k=0}^{N}\frac{1}{2^k}\\ 0 & \frac{\pi}{3}\left(\sum_{k=0}^{N}\frac{1}{4^k}\right)\end{pmatrix},\quad k=0,1,2,\cdots,$$

于是

$$\boldsymbol{S}=\lim_{N\to\infty}\boldsymbol{S}_N=\begin{pmatrix}1 & 2\\ 0 & \frac{4}{3}\pi\end{pmatrix}$$

故矩阵级数 $\sum_{k=1}^{\infty}\mathbf{A}_k$ 收敛，且和为$\boldsymbol{S}$。

例3 设A为n阶方阵，则有

（1） $\sum_{k=0}^{\infty} A^k = I + A + A^2 + \cdots + A^k + \cdots$ 收敛($A_0=I$)的充要条件是 $\rho(A) < 1$;

（2） 当$\sum_{k=0}^{\infty} A^k$ 收敛时，有

$$\sum_{k=0}^{\infty} A^k = (I - A)^{-1},$$

而且存在$C^{n\times n}$上的算子范数 ‖·‖，使得

$$\left\| (I - A)^{-1} - \sum_{k=0}^{m} A^k \right\| \le \frac{\|A\|^{m+1}}{1 - \|A\|}$$

证 必要性 若矩阵级数 $\sum_{k=1}^{\infty} A_k$ 收敛，则有 $S=\lim_{k\to\infty} S_k$。

又级数 $$I+A+A^2+\cdots+A^k+\cdots$$

的前k项部分和与前$k+1$项部分和分别为：

$$S_k=I+A+A^2+\cdots+A^{k-1}, \qquad S_{k+1}=I+A+A^2+\cdots+A^k$$

因此 $A^k=S_{k+1}-S_k$，利用极限运算法则有

$$\lim_{k\to\infty} A^k=\lim_{k\to\infty}\left[S_{k+1}-S_k\right]=0$$

根据例2，$\rho(A)<1$。

充分性 由 $AS_k=A\left(I+A+\cdots+A^{k-1}\right)=A+A^2+\cdots+A^k$

则有 $S_k-AS_k=I-A^k$，$(I-A)S_k=I-A^k$。 由$\rho(A)<1$，可知

则存在某种范数 $\|\cdot\|$，使得 $\|A\|<1$，且（I−A）可逆。 又有

$\lim_{k\to\infty} A^k=0$，根据矩阵序列极限法则，有

$$\sum_{k=0}^{\infty} A^k=\lim_{k\to\infty} S_k=\lim_{k\to\infty}\left[(I-A)^{-1}(I-A^k)\right]=(I-A)^{-1}\lim_{k\to\infty}\left(I-A^k\right)=\left(I-A\right)^{-1}。$$

从而, 进一步有

$$\left\|(\boldsymbol{I}-\boldsymbol{A})^{-1}-\sum_{k=0}^{m}\boldsymbol{A}^{k}\right\|=\left\|\sum_{k=m+1}^{\infty}\boldsymbol{A}^{k}\right\|\leq\sum_{k=m+1}^{\infty}\|\boldsymbol{A}\|^{k}=\frac{\|\boldsymbol{A}\|^{m+1}}{1-\|\boldsymbol{A}\|}\text{。}$$

练习　计算 $\sum_{k=0}^{\infty}\boldsymbol{A}^{k}$，其中 $\boldsymbol{A}=\begin{pmatrix}0.1 & 0.7\\ 0.3 & 0.6\end{pmatrix}$。

解： 由于 $\|\boldsymbol{A}\|_{\infty}=0.9<1$, 故 $\rho(\boldsymbol{A})<1$，从而 $\sum_{k=0}^{\infty}\boldsymbol{A}^{k}$ 收敛，且有

$$\sum_{k=0}^{\infty}\boldsymbol{A}^{k}=(\boldsymbol{I}-\boldsymbol{A})^{-1}=\begin{pmatrix}0.9 & -0.7\\ -0.3 & 0.4\end{pmatrix}^{-1}=\frac{2}{3}\begin{pmatrix}4 & 7\\ 3 & 9\end{pmatrix}$$

推论3.2 设$A\in \mathbf{C}^{n\times n}$，若对$\mathbf{C}^{n\times n}$上的某种矩阵范数$\|\cdot\|$，有$\|A\|<1$，则 $\lim\limits_{k\to\infty} A^k=0$。

例4（1）已知$\sum\limits_{k=0}^{\infty} A^k=\begin{pmatrix}2 & 0\\ -4 & 2\end{pmatrix}$，求$A$。

（2）设 $A=\begin{pmatrix}0.2 & 0.5 & 0.4\\ 0.5 & 0.2 & 0.1\\ 0.2 & 0.1 & 0.3\end{pmatrix}$，求证 $\lim\limits_{k\to\infty} A^k=0$。

解 （1）由$\sum\limits_{k=0}^{\infty} A^k=(I-A)^{-1}=\begin{pmatrix}2 & 0\\ -4 & 2\end{pmatrix}$，则$(I-A)=\begin{pmatrix}2 & 0\\ -4 & 2\end{pmatrix}^{-1}$，

$$A=I-\begin{pmatrix}2 & 0\\ -4 & 2\end{pmatrix}^{-1}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}-\begin{pmatrix}\frac{1}{2} & 0\\ 1 & \frac{1}{2}\end{pmatrix}=\begin{pmatrix}\frac{1}{2} & 0\\ -1 & \frac{1}{2}\end{pmatrix}$$

（2）因为$\|A\|_\infty=0.9<1$，由推论3.2，故 $\lim\limits_{k\to\infty} A^k=0$。

性质3.4 设 $\sum_{k=1}^{\infty} A_k = A$ 和 $\sum_{k=1}^{\infty} B_k = B$，其中$A_k, A, B_k, B \in \mathbf{C}^{m\times n}$，则

$$\sum_{k=0}^{\infty} (\alpha A_k + \beta B_k) = \alpha \sum_{k=0}^{\infty} A_k + \beta \sum_{k=0}^{\infty} B_k, \qquad \forall \alpha, \beta \in \mathbf{C}$$

证 因为 $s_N = \sum_{k=0}^{N} (\alpha A_k + \beta B_k) = \alpha \sum_{k=0}^{N} A_k + \beta \sum_{k=0}^{N} B_k, \qquad \forall \alpha, \beta \in \mathbf{C}$

所以

$$\sum_{k=0}^{\infty} (\alpha A_k + \beta B_k) = \lim_{N\to\infty} s_N = \alpha \lim_{N\to\infty} \sum_{k=0}^{N} A_k + \beta \lim_{N\to\infty} \sum_{k=0}^{N} B_k = \alpha A + \beta B$$

由定理3.1，即结论成立。

矩阵级数收敛的定义与数项级数的定义没有本质的区别，我们有一些类似于数项级数的概念和结论。

定义3. 4 设 $\sum_{k=1}^{\infty} \boldsymbol{A}_k$ 为$\mathbf{C}^{m\times n}$中的矩阵级数，其中 $\boldsymbol{A}_k = \left(a_{ij}^{(k)}\right)$。如果$\sum_{k=1}^{\infty} a_{ij}^{(k)}$ 对任意的$1\leqslant \boldsymbol{i}\leqslant \boldsymbol{m}$，$1\leqslant \boldsymbol{j}\leqslant \boldsymbol{n}$均为绝对收敛的，则称矩阵级数 $\sum_{k=1}^{\infty} \boldsymbol{A}_k$ 绝对收敛。

对比矩阵级数绝对收敛的定义以及高等数学中的数项级数的绝对收敛的定义可以得出矩阵级数收敛的一些性质。

性质3. 5 若矩阵级数 $\sum_{k=1}^{\infty} \boldsymbol{A}_k$ 是绝对收敛，则它一定是收敛的，并且任意调换各项的顺序所得到的级数还是收敛的， 且级数和不变。

性质3.6 矩阵级数 $\sum_{k=1}^{\infty} A_k$ 为绝对收敛的充分必要条件是正项级数 $\sum_{k=1}^{\infty} \|A_k\|$ 收敛。

利用矩阵范数的等价性, 只需证明对于∞−范数定理成立即可。

证 必要性 如果 $\sum_{k=1}^{\infty} A_k$ 是绝对收敛的， 由定义即对任意的 $1 \leqslant i \leqslant m$，$1 \leqslant j \leqslant n$ ，$\sum_{k=1}^{\infty} a_{ij}^{(k)}$ 均绝对收敛, 即存在充分大的 N 和一个与N无关的正数M, 使得

$$\sum_{k=1}^{N} \left| a_{ij}^{(k)} \right| < M, \ (i = 1, 2, \cdots, m, \quad j = 1, 2, \cdots, n)$$

从而有

$$\sum_{k=1}^{N} \|A_k\|_{\infty} = \sum_{k=1}^{N} \left(\max_i \sum_{j=1}^{n} \left| a_{ij}^{(k)} \right| \right) = \max_i \left\{ \sum_{j=1}^{n} \sum_{k=1}^{N} \left| a_{ij}^{(k)} \right| \right\} \leq \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{N} \left| a_{ij}^{(k)} \right| < nm \cdot M$$

因此 $\sum_{k=1}^{\infty} \|A_k\|_{\infty}$ 为收敛的正项级数。

充分性 如果 $\sum_{k=1}^{\infty}\|A_k\|_\infty$ 为收敛的正项级数，那么有

$$\sum_{k=1}^{N}\left|a_{ij}^{(k)}\right| \le \sum_{k=1}^{N}\|A_k\|_\infty < M \quad i=1,2,\cdots,m, \quad j=1,2,\cdots,n$$

可知$m\times n$个级数 $\sum_{k=1}^{\infty}a_{ij}^{(k)}$ 均为绝对收敛的，利用定义4可知矩阵级数 $\sum_{k=1}^{\infty}A_k$ 是绝对收敛的。

性质3.7 设 $\sum_{k=1}^{\infty} \boldsymbol{A}_k$ 为$\mathbf{C}^{m\times n}$中的绝对收敛的矩阵级数，$\sum_{k=1}^{\infty} \boldsymbol{B}_k$ 为$\mathbf{C}^{n\times l}$中的绝对收敛的级数，并且 $\boldsymbol{A}=\sum_{k=1}^{\infty} \boldsymbol{A}_k$，$\boldsymbol{B}=\sum_{k=1}^{\infty} \boldsymbol{B}_k$，则 $\sum_{k=1}^{\infty} \boldsymbol{A}_k \sum_{k=1}^{\infty} \boldsymbol{B}_k$ 按任何方式排列得到的级数也绝对收敛，且和为$\boldsymbol{AB}$。

性质3.8 设 $P\in C^{p\times m}$和 $Q\in C^{n\times q}$为给定矩阵，如果$m\times n$型矩阵级数 $\sum_{k=0}^{\infty}A_k$收敛(或绝对收敛)，则 $p\times q$ 矩阵级数 $\sum_{k=0}^{\infty}PA_kQ$ 也收敛（或绝对收敛），且有等式

$$\sum_{k=0}^{\infty}PA_kQ = \left(\sum_{k=0}^{\infty}P A_k\right)Q$$

证 设 $\sum_{k=0}^{\infty}A_k$收敛于矩阵S，即 $S=\sum_{k=0}^{\infty}A_k=\lim_{n\to\infty}\sum_{k=0}^{n}A_k$，而由等式 $\sum_{k=0}^{n}PA_kQ=P\left(\sum_{k=0}^{n}A_k\right)Q$ 取极限即得：

$$\sum_{k=0}^{\infty}PA_kQ=\lim_{n\to\infty}\sum_{k=0}^{n}PA_kQ=P\left(\lim_{n\to\infty}\sum_{k=0}^{n}A_k\right)Q=PSQ=P\left(\sum_{k=0}^{\infty}A_k\right)Q$$

即 $\sum_{k=0}^{\infty}PA_kQ$ 收敛，且有 $\sum_{k=0}^{\infty}PA_kQ=P\left(\sum_{k=0}^{\infty}A_k\right)Q$

现设$\sum_{k=0}^{\infty} A_k$绝对收敛，由矩阵级数性质2知，$\sum_{k=0}^{\infty}\|A_k\|$收敛， 又

$$\|PA_kQ\| \le \|P\|\,\|A_k\|\,\|Q\|$$

利用比较判别法, 即知级数 $\sum_{k=0}^{\infty}\|PA_kQ\|$ 收敛， 再利用矩阵级数性质2，便知矩阵级数 $\sum_{k=0}^{\infty}PA_kQ$ 绝对收敛。

返回

3.2、矩阵幂级数

定理3. 2 设$\sum_{k=0}^{\infty} a_k t^k$ 为收敛半径为r的幂级数，$\boldsymbol{A}$为n阶方阵，则

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

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

证 （1）如果$\rho(\boldsymbol{A}) < r$，根据矩阵范数的性质，对于

$\varepsilon = \frac{1}{2}\left[r - \rho(\boldsymbol{A})\right] > 0$，一定存在一种相容的矩阵范数 $\|\cdot\|$，使得

$$\|\boldsymbol{A}\| \le \rho(\boldsymbol{A}) + \varepsilon = \rho(\boldsymbol{A}) + \frac{1}{2}r - \frac{1}{2}\rho(\boldsymbol{A}) = \frac{r + \rho(\boldsymbol{A})}{2} < r$$

因此，数项级数 $\sum_{k=0}^{\infty} |a_k| \|A\|^k$ 收敛，又

再由数项级数比较判别法可知，$\sum_{k=0}^{\infty} \|a_k \boldsymbol{A}^k\| \le \sum_{k=0}^{\infty} |a_k| \cdot \|\boldsymbol{A}\|^k$ 收敛。

再利用矩阵级数性质3. 6知，$\sum_{k=0}^{\infty} a_k \boldsymbol{A}^k$ 绝对收敛。

（2）如果 $\rho(\boldsymbol{A})>r$，设 $\boldsymbol{Ax}=\lambda_i\boldsymbol{x}$，其中 $|\lambda_i|=\rho(\boldsymbol{A})$，且设$\boldsymbol{x}$为其单位长度特征向量，亦即 $\boldsymbol{x}^H\boldsymbol{Ax}=\lambda_i$。

下面用反证法证明矩阵幂级数 $\sum_{k=0}^{\infty}a_k\boldsymbol{A}^k$ 发散。

如果它是收敛的，则利用矩阵收敛的性质3.8知，应有

$$\underline{\sum_{k=0}^{\infty}a_k\lambda_i^k}=\sum_{k=0}^{\infty}a_k\left(\boldsymbol{x}^H\boldsymbol{A}^k\boldsymbol{x}\right)=\boldsymbol{x}^H\left(\sum_{k=0}^{\infty}a_k\boldsymbol{A}^k\right)\boldsymbol{x}=\boldsymbol{x}^H\boldsymbol{S}\boldsymbol{x}\ \underline{<\infty}$$

即得出此数项级数也收敛。但 $|\lambda_i|=\rho(\boldsymbol{A})>r$，数项级数 $\sum_{k=0}^{\infty}a_kz^k$ 在收敛圆外是发散的。故 $\sum_{k=0}^{\infty}a_k\lambda_i^k$ 应该是发散的， 因此矛盾，

故结论（2）成立。

经过简单的变换便可得到如下推论:

推论3.3 设 $\sum_{k=0}^{\infty} a_k (z-z_0)^k$ 为收敛半径为$\boldsymbol{r}$的幂级数，$\boldsymbol{A}$为$\boldsymbol{n}$阶方阵，如果$\boldsymbol{A}$的特征值均落在收敛圆内，即 $|\lambda - z_0| < r$,其中 λ 为$\boldsymbol{A}$的任意特征值，则矩阵幂级数 $\sum_{k=0}^{\infty} a_k (\boldsymbol{A} - z_0 \boldsymbol{I})^k$ 绝对收敛； 若有某个 λ_{i_0} 使得 $|\lambda_{i_0} - z_0| > r$ ，则幂级数 $\sum_{k=0}^{\infty} a_k (\boldsymbol{A} - z_0 \boldsymbol{I})^k$ 发散。

根据幂级数性质，幂级数的和函数是收敛圆内的解析函数（任意次可微，在任一点处均可展成**Taylor**级数），而一个圆内解析的函数可以展开成收敛的幂级数。于是，如果 $f(z)$ 是 $|z-z_0|<r$ 内的解析函数，其展成绝对收敛的幂级数为

$$f(z)=\sum_{k=0}^{\infty}a_k(z-z_0)^k$$

则当矩阵$\boldsymbol{A}\in\mathbf{C}^{n\times n}$的特征值落在收敛圆 $|z-z_0|<r$ 内时，则由

$$f(\boldsymbol{A})\triangleq\sum_{k=0}^{\infty}a_k(\boldsymbol{A}-z_0\boldsymbol{I})^k$$

称之为$\boldsymbol{A}$关于解析函数$f(z)$的矩阵函数。

例如，对于收敛半径$r=+\infty$的幂级数

$$e^z=1+z+\frac{z^2}{2!}+\frac{z^3}{3!}+\cdots;$$

$$\sin z=z-\frac{z^3}{3!}+\frac{z^5}{5!}-\cdots;\qquad \cos z=1-\frac{z^2}{2!}+\frac{z^4}{4!}-\cdots;$$

根据上述的定义，有矩阵指数函数和矩阵三角函数$(\boldsymbol{A}\in\mathbf{C}^{n\times n})$

$$e^{\boldsymbol{A}}=\boldsymbol{I}+\boldsymbol{A}+\frac{\boldsymbol{A}^2}{2!}+\frac{\boldsymbol{A}^3}{3!}+\cdots;$$

$$\cos\boldsymbol{A}=\boldsymbol{I}-\frac{\boldsymbol{A}^2}{2!}+\frac{\boldsymbol{A}^4}{4!}-\cdots;\qquad \sin\boldsymbol{A}=\boldsymbol{A}-\frac{\boldsymbol{A}^3}{3!}+\frac{\boldsymbol{A}^5}{5!}-\cdots;$$

对于收敛半径$r=1$的幂级数

$$(1-z)^{-1}=1+z+z^2+z^3+\cdots;\qquad \ln(1+z)=z-\frac{z^2}{2}+\frac{z^3}{3}-\cdots;$$

相应的有$(\boldsymbol{A}\in\mathbf{C}^{n\times n}$，且 $\rho(\boldsymbol{A})<1)$

$$(\boldsymbol{I}-\boldsymbol{A})^{-1}=\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\boldsymbol{A}^3+\cdots;\qquad \ln(\boldsymbol{I}+\boldsymbol{A})=\boldsymbol{A}-\frac{\boldsymbol{A}^2}{2}+\frac{\boldsymbol{A}^3}{3}-\cdots;$$

练习 设A为n阶Householder矩阵, 则 $\cos\left(2\pi \mathbf{A}\right) =$ _____

注意, A 是Householder矩阵, 则满足:

$$\boldsymbol{A}^2 = \boldsymbol{A}^H \boldsymbol{A} = \boldsymbol{I} \quad \boldsymbol{A}^4 = \boldsymbol{I}, \cdots, \quad \boldsymbol{A}^{2n} = \boldsymbol{I}$$

$$\cos(2\pi \boldsymbol{A}) = \boldsymbol{I} - \frac{(2\pi \boldsymbol{A})^2}{2!} + \frac{(2\pi \boldsymbol{A})^4}{4!} - \cdots + \frac{(2\pi \boldsymbol{A})^{2n}}{(2n)!!} + \cdots$$

$$= \boldsymbol{I} - \frac{(2\pi)^2}{2!}\boldsymbol{A}^2 + \frac{(2\pi)^4}{4!}\boldsymbol{A}^4 - \cdots + \frac{(2\pi)^{2n}}{(2n)!!}\boldsymbol{A}^{2n} + \cdots$$

$$= \left(1 - \frac{(2\pi)^2}{2!} + \frac{(2\pi)^4}{4!} - \cdots + \frac{(2\pi)^{2n}}{(2n)!!} + \cdots\right)\boldsymbol{I}$$

$$= (\cos 2\pi)\boldsymbol{I} = \boldsymbol{I} 。$$

现在利用矩阵的**Jordan**分解写出矩阵函数 $f(A)$ 的具体表达式

首先介绍一个引理

引理3.1 设 $f(z)=\sum_{k=0}^{\infty}a_k z^k$ 是收敛半径为r的幂级数，J_i是特征值为λ_i的n_i阶**Jordan**块阵，且 $|\lambda_i|<r$，则

$$f(J_i)=\begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \cdots & \dfrac{f^{(n-1)}(\lambda_i)}{(n_i-1)!} \\ & f(\lambda_i) & f'(\lambda_i) & \vdots \\ & & \ddots & \\ & & \ddots & f'(\lambda_i) \\ & & & f(\lambda_i) \end{pmatrix} \qquad (3\text{-}1)$$

推论3. 4 设 $f(z)=\sum_{k=0}^{\infty}a_k(z-z_0)^k$ 是收敛半径为r的幂级数, $\boldsymbol{J_i}$是特征值为 λ_i的**Jordan**块, 且 $|\lambda_i - z_0| < r$,

$$f(\boldsymbol{J}_i)=\begin{pmatrix} f(\lambda_i) & f'(\lambda) & \cdots & \dfrac{f^{(n-1)}(\lambda_i)}{(n_i-1)!} \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & f'(\lambda_i) \\ & & & f(\lambda_i) \end{pmatrix}$$

推论3. 5 设 $f(z)=\sum_{k=0}^{\infty}a_k z^k$ 是收敛半径为r的幂级数，J是特征值为 λ_i 的n_i阶Jordan块阵, 且 $|t\lambda_i|<r$，则

$$f(tJ_i)=\begin{pmatrix} f(t\lambda_i) & t\,f'(t\lambda_i) & \cdots & \dfrac{t^{n-1}f^{(n-1)}(t\lambda_i)}{(n_i-1)!} \\ & f(t\lambda_i) & \ddots & \vdots \\ & & \ddots & t\,f'(t\lambda_i) \\ & & & f(t\lambda_i) \end{pmatrix} \qquad \text{（3-3）}$$

练习

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

$$
e^{\boldsymbol{J}_i}=\begin{pmatrix} e^{-1} & e^{-1} & \dfrac{e^{-1}}{2} & \dfrac{e^{-1}}{6} \\ & e^{-1} & e^{-1} & \dfrac{e^{-1}}{2} \\ & & e^{-1} & e^{-1} \\ & & & e^{-1} \end{pmatrix}
$$

$$
\boldsymbol{J}_i=\begin{pmatrix} -1 & 1 & 0 & 0 \\ & -1 & 1 & 0 \\ & & -1 & 1 \\ & & & -1 \end{pmatrix}
$$

$$
f\left(t\boldsymbol{J}_i\right)=\begin{pmatrix}
f\left(t\lambda_i\right) & t\,f'\left(t\lambda_i\right) & \cdots & \dfrac{t^{n-1}f^{(n-1)}\left(t\lambda_i\right)}{\left(n_i-1\right)!} \\
 & f\left(t\lambda_i\right) & \ddots & \vdots \\
 & & \ddots & t\,f'\left(t\lambda_i\right) \\
 & & & f\left(t\lambda_i\right)
\end{pmatrix}
$$

$$
e^{t\boldsymbol{J}_i}=\begin{pmatrix}
e^{-t} & te^{-t} & \dfrac{t^2}{2}e^{-t} & \dfrac{t^3}{6}e^{-t} \\
 & e^{-t} & te^{-t} & \dfrac{t^2}{2}e^{-t} \\
 & & e^{-t} & te^{-t} \\
 & & & e^{-t}
\end{pmatrix}
$$

根据上面的引理和矩阵级数的性质, 有

定理3. 3 设 $f(z)=\sum_{k=0}^{\infty}a_k z^k$ 为收敛半径为r的幂级数, A 为n阶方阵, $A=TJT^{-1}$为其Jordan分解, $J=\mathrm{diag}(J_1, J_2, \cdots, J_s)$。当$A$的特征值均落在收敛圆内时, 即 $|\lambda|<r$, 其中 λ 为A的任意特征值, 则矩阵幂级数 $\sum_{k=0}^{\infty}a_k A^k$ 绝对收敛, 并且和矩阵为

$$f(A)=T\ \mathrm{diag}\big(f(J_1), f(J_2), \cdots, f(J_s)\big)\ T^{-1} \qquad (3\text{-}4)$$

其中$f(J_i)$的定义如表达式（3-1）。

那么, 若A的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则 $f(A)$ 的特征值应为 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$。

证明

$$\left(TJT^{-1}\right)^k=\overbrace{\left(TJT^{-1}\right)\left(TJT^{-1}\right)\cdots\left(TJT^{-1}\right)}^{k}=T\overbrace{J\,J\,J\cdots J}^{k}T^{-1}=TJ^kT^{-1}$$

则 $$f(A)=\sum_{k=0}^{\infty}a_kA^k=\sum_{k=0}^{\infty}a_k\left(TJT^{-1}\right)^k=\sum_{k=0}^{\infty}a_k\left(TJ^kT^{-1}\right)=T\left(\sum_{k=0}^{\infty}a_kJ^k\right)T^{-1}$$

$$=T\ \mathbf{diag}\left(\sum_{k=0}^{\infty}a_kJ_1^k,\quad \sum_{k=0}^{\infty}a_kJ_2^k,\quad \cdots,\quad \sum_{k=0}^{\infty}a_kJ_s^k\right)T^{-1}$$

$$=T\ \mathbf{diag}\left(f(J_1),\quad f(J_2),\quad \cdots,\quad f(J_s)\right)T^{-1}$$

$$f(tA)=\sum_{k=0}^{\infty}a_kt^kA^k=\sum_{k=0}^{\infty}a_k\left(TJT^{-1}\right)^kt^k=\sum_{k=0}^{\infty}a_k\left(TJ^kT^{-1}\right)t^k$$

$$=T\left(\sum_{k=0}^{\infty}a_kt^kJ^k\right)T^{-1}$$

$$=T\ \mathbf{diag}\left(\sum_{k=0}^{\infty}a_k(tJ_1)^k,\quad \sum_{k=0}^{\infty}a_k(tJ_2)^k,\quad \cdots,\quad \sum_{k=0}^{\infty}a_k(tJ_s)^k\right)T^{-1}$$

$$=T\,\mathbf{diag}\left(f\left(tJ_1\right),\quad f\left(tJ_2\right),\quad \cdots,\quad f\left(tJ_s\right)\right)T^{-1}$$

$$\boldsymbol{f}\left(\boldsymbol{A}\right)=\boldsymbol{T}\ \mathbf{diag}\Big(f\left(\boldsymbol{J}_1\right),\ f\left(\boldsymbol{J}_2\right),\ \cdots,\ f\left(\boldsymbol{J}_s\right)\Big)\ \boldsymbol{T}^{-1}$$

$$=\boldsymbol{T}\begin{pmatrix} f\left(\boldsymbol{J}_i\right) & & & \\ & f\left(\boldsymbol{J}_i\right) & & \\ & & \ddots & \\ & & & f\left(\boldsymbol{J}_i\right) \end{pmatrix}\boldsymbol{T}^{-1}$$

其中

$$f\left(\boldsymbol{J}_i\right)=\begin{pmatrix} f\left(\lambda_i\right) & f'\left(\lambda_i\right) & \cdots & \dfrac{f^{(n-1)}\left(\lambda_i\right)}{\left(n_i-1\right)!} \\ & f\left(\lambda_i\right) & \ddots & \vdots \\ & & \ddots & f'\left(\lambda_i\right) \\ & & & f\left(\lambda_i\right) \end{pmatrix}\quad i=1,2,\cdots,s。$$

$$f(t\boldsymbol{A}) = \boldsymbol{T}\ \mathbf{diag}\Big(f(t\boldsymbol{J}_1),\quad f(t\boldsymbol{J}_2),\quad \cdots,\quad f(t\boldsymbol{J}_s)\Big)\boldsymbol{T}^{-1}$$

$$= \boldsymbol{T}\begin{pmatrix} f(t\boldsymbol{J}_1) & & & \\ & f(t\boldsymbol{J}_2) & & \\ & & \ddots & \\ & & & f(t\boldsymbol{J}_s) \end{pmatrix}\boldsymbol{T}^{-1}$$

其中

$$f(t\boldsymbol{J}_i) = \begin{pmatrix} f(t\lambda_i) & t f'(t\lambda_i) & \cdots & \dfrac{t^{n-1} f^{(n-1)}(t\lambda_i)}{(n_i - 1)!} \\ & f(t\lambda_i) & \ddots & \vdots \\ & & \ddots & t f'(t\lambda_i) \\ & & & f(t\lambda_i) \end{pmatrix} \quad i = 1, 2, \cdots, s。$$

例1 设 $\boldsymbol{A}=\begin{pmatrix}3 & 0 & 8\\ 3 & -1 & 6\\ -2 & 0 & -5\end{pmatrix}$， 求 $\sin\boldsymbol{A}$。

解 根据矩阵A的Jordan分解

$$\boldsymbol{A}=\boldsymbol{TJT}^{-1}=\begin{pmatrix}0 & 4 & 1\\ 1 & 3 & 0\\ 0 & -2 & 0\end{pmatrix}\begin{pmatrix}-1 & & \\ & -1 & 1\\ & & -1\end{pmatrix}\begin{pmatrix}0 & 1 & \frac{3}{2}\\ 0 & 0 & -\frac{1}{2}\\ 1 & 0 & 2\end{pmatrix}$$

因此，由 $\sin\boldsymbol{A}=f(\boldsymbol{A})=\boldsymbol{T}\ \mathbf{diag}\big(f(\boldsymbol{J}_1),\ f(\boldsymbol{J}_2)\big)\boldsymbol{T}^{-1}$

$$=\begin{pmatrix}0 & 4 & 1\\ 1 & 3 & 0\\ 0 & -2 & 0\end{pmatrix}\begin{pmatrix}-\sin1 & & \\ & -\sin1 & \cos1\\ & & -\sin1\end{pmatrix}\begin{pmatrix}0 & 1 & \frac{3}{2}\\ 0 & 0 & -\frac{1}{2}\\ 1 & 0 & 2\end{pmatrix}$$

$$=\begin{pmatrix}4\cos1-\sin1 & 0 & 8\cos1\\ 3\cos1 & -\sin1 & 6\cos1\\ -2\cos1 & 0 & -\sin1-4\cos1\end{pmatrix}$$

例2 设 $\boldsymbol{A}=\begin{pmatrix}3 & 1 & -1\\ 1 & 2 & -1\\ 2 & 1 & 0\end{pmatrix}$，求 $\boldsymbol{e}^{\boldsymbol{A}t}$。

解 根据矩阵的**Jordan** 分解

$$\boldsymbol{A}=\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 1\end{pmatrix}\begin{pmatrix}1 & & \\ & 2 & 1\\ & & 2\end{pmatrix}\begin{pmatrix}-1 & 0 & 1\\ 0 & -1 & 1\\ 1 & 1 & -1\end{pmatrix}$$ 则

$$\boldsymbol{e}^{\boldsymbol{A}t}=f(\boldsymbol{A}t)=\boldsymbol{T}\ \mathbf{diag}\Big(f(t\boldsymbol{J}_1),\quad f(t\boldsymbol{J}_2)\Big)\boldsymbol{T}^{-1}$$

$$e^{At}=\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 1\end{pmatrix}\begin{pmatrix}e^t & & \\ & e^{2t} & te^{2t}\\ & & e^{2t}\end{pmatrix}\begin{pmatrix}-1 & 0 & 1\\ 0 & -1 & 1\\ 1 & 1 & -1\end{pmatrix}$$

$$=\begin{pmatrix}0 & e^{2t} & (t+1)e^{2t}\\ e^t & 0 & e^{2t}\\ e^t & e^{2t} & (t+1)e^{2t}\end{pmatrix}\begin{pmatrix}-1 & 0 & 1\\ 0 & -1 & 1\\ 1 & 1 & -1\end{pmatrix}=\begin{pmatrix}(1+t)e^{2t} & te^{2t} & -te^{2t}\\ e^{2t}-e^t & e^{2t} & -e^{2t}+e^t\\ (1+t)e^{2t}-e^t & te^{2t} & e^t-te^{2t}\end{pmatrix}$$

为避免求矩阵**A**的的**Jordan** 分解，也可用有限待定系数法计算$f(A)$和$f(At)$。

有限待定系数法 设$A\in\mathbf{C}^{n\times n}$且

$$\psi(\lambda)=\det(\lambda I-A)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s} \quad \text{(3-5)}$$

其中$m_i(i=1,2,\cdots,s)$均为正整数，$\sum_{i=1}^{s} m_i=n$，$\lambda_1,\lambda_2,\cdots,\lambda_s$为$A$的不同特征值为计算矩阵函数$f(At)=\sum_{k=0}^{\infty} a_k A^k t^k$，记$f(\lambda t)=\sum_{k=0}^{\infty} a_k\lambda^k t^k$，将$f(At)$改写为

$$f(\lambda t)=p(\lambda,t)\psi(\lambda)+q(\lambda,t) \quad \text{(3-6)}$$

其中$p(\lambda,t)$是含参数t的λ的幂级数，$q(\lambda,t)$是含参数t且次数不超过n-1的λ的多项式，即

$$q(\lambda,t)=b_{n-1}(t)\lambda^{n-1}+b_{n-2}(t)\lambda^{n-2}+\cdots+b_0(t)$$

由**Hamilton-Cayley**定理知 $\psi(A)=O$，于是由（**3-5**）式得

$$f(At)=p(A,t)\psi(A)+q(A,t)=b_{n-1}(t)A^{n-1}+\cdots+b_1(t)A+b_0(t)I$$

可见，只要求出 $b_0(t),b_1(t),\cdots,b_{n-1}(t)$ 即可得到***f*(*A*t)**。注意到

$$\psi^{(j)}(\lambda_i)=0\quad(\ j=0,1,\cdots,m_i-1;i=1,2,\cdots,s\ )$$

将（**3-6**）式两端对λ求导，并利用上式，得

$$\left.\frac{\mathrm{d}^j f(\lambda t)}{\mathrm{d}\lambda^j}\right|_{\lambda=\lambda_i}=\left.\frac{\mathrm{d}^j q(\lambda,t)}{\mathrm{d}\lambda^j}\right|_{\lambda=\lambda_i}$$

即

$$t^j\left.\frac{\mathrm{d}^j f(u)}{\mathrm{d}u^j}\right|_{u=\lambda_i t}=\left.\frac{\mathrm{d}^j q(\lambda,t)}{\mathrm{d}\lambda^j}\right|_{\lambda=\lambda_i}\qquad\text{（3-7）}$$

$$(\ j=0,1,\cdots,m_i-1;i=1,2,\cdots,s\ )$$

由（**3-7**）式即得到以 $b_0(t),b_1(t),\cdots,b_{n-1}(t)$ 为未知量的线性方程组

用有限待定级数法计算矩阵函数***f*(*A*)** 和***f*(*At*)**的步骤如下：

（一）求矩阵***A***的特征多项式（**3-4**）；

（二）设 $q(\lambda,t)=b_{n-1}(t)\lambda^{n-1}+b_{n-2}(t)\lambda^{n-2}+\cdots+b_0(t)$。 根据

$$q^{(j)}(\lambda_i)=t^j f^{(j)}(\lambda)\Big|_{\lambda=\lambda_i t} \quad (j=0,1,\cdots,m_i-1;i=1,2,\cdots,s)$$

或

$$q^{(j)}(\lambda_i)=f^{(j)}(\lambda_i) \quad (j=0,1,\cdots,m_i-1;i=1,2,\cdots,s)$$

列出线性方程组**,**并求解 $b_0,b_1,\cdots,b_{n-1}$。

（三）计算***f*(*tA*) (**当取***t*=1**时**,** 为***f*(*A*) =*q*(*A*))**

$$q(\boldsymbol{A})=b_{n-1}(t)\boldsymbol{A}^{n-1}+\cdots+b_1(t)\boldsymbol{A}+b_0(t)\boldsymbol{I}$$

例3 用有限待定级数法计算（1）例1，（2）计算 $e^{\boldsymbol{A}t}$。

解 首先求 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix}\lambda-3 & 0 & -8\\ -3 & \lambda+1 & -6\\ 2 & 0 & \lambda+5\end{vmatrix}=(\lambda+1)^3$

(1) 设 $q(\lambda)=b_2\lambda^2+b_1\lambda+b_0$，$f(\lambda)=\sin\lambda$。

因此，由（3-6）可得

$$\begin{cases} q(-1)=b_2-b_1+b_0=f(-1)=-\sin 1\\ q'(-1)=-2b_2+b_1=f'(-1)=\cos 1\\ q''(-1)=2b_2=\sin 1=f''(-1)\end{cases}$$

解得

$$\begin{cases} b_0=\cos 1-\dfrac{1}{2}\sin 1\\ b_1=\sin 1+\cos 1\\ b_2=\dfrac{1}{2}\sin 1\end{cases}$$

于是， $\sin \boldsymbol{A} = f\left(\boldsymbol{A}\right) = b_2\boldsymbol{A}^2 + b_1\boldsymbol{A} + b_0\boldsymbol{I}$ ，即

$$\sin \boldsymbol{A} = \frac{1}{2}\sin 1\begin{pmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{pmatrix}^2 + \left(\sin 1 + \cos 1\right)\begin{pmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{pmatrix}$$

$$+\left(\cos 1 - \frac{1}{2}\sin 1\right)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{7}{2}\sin 1 & 0 & -8\sin 1 \\ -3\sin 1 & \frac{1}{2}\sin 1 & 6\cos 1 \\ 2\sin 1 & 0 & \frac{9}{2}\sin 1 \end{pmatrix} + \begin{pmatrix} 3\left(\sin 1 + \cos 1\right) & 0 & 8\left(\sin 1 + \cos 1\right) \\ 3\left(\sin 1 + \cos 1\right) & -\left(\sin 1 + \cos 1\right) & 6\left(\sin 1 + \cos 1\right) \\ -2\left(\sin 1 + \cos 1\right) & 0 & -5\left(\sin 1 + \cos 1\right) \end{pmatrix}$$

$$+ \begin{pmatrix} \cos 1 - \frac{1}{2}\sin 1 & 0 & 0 \\ 0 & \cos 1 - \frac{1}{2}\sin 1 & 0 \\ 0 & 0 & \cos 1 - \frac{1}{2}\sin 1 \end{pmatrix} = \begin{pmatrix} 4\cos 1 - \sin 1 & 0 & 8\cos 1 \\ 3\cos 1 & -\sin 1 & 6\cos 1 \\ -2\cos 1 & 0 & -\sin 1 - 4\cos 1 \end{pmatrix}$$

(2) 设 $q(\lambda)=b_2(t)\lambda^2+b_1(t)\lambda+b_0(t)$，$f(t\lambda)=e^{\lambda t}$ 。

因此，由（**3-6**）可得

$$\begin{cases} q(-1)=b_2-b_1+b_0=f(-t)=e^{-t} \\ q'(-1)=-2b_2+b_1 \quad =f'(-t)=te^{-t} \\ q''(-1)=2b_2 \quad =f''(-t)=t^2e^{-t} \end{cases}$$

解得

$$\begin{cases} b_0=\left(1+t+\dfrac{t^2}{2}\right)e^{-t} \\ b_1=\left(t+t^2\right)e^{-t} \\ b_2=\dfrac{t^2}{2}e^{-t} \end{cases}$$

于是，$e^{tA}=f(tA)=b_2(t)A^2+b_1(t)A+b_0(t)I$，即

$$e^{tA}=\left(1+t+\frac{t^2}{2}\right)e^{-t}\begin{pmatrix}3&0&8\\3&-1&6\\-2&0&-5\end{pmatrix}^2+\left(t+t^2\right)e^{-t}\begin{pmatrix}3&0&8\\3&-1&6\\-2&0&-5\end{pmatrix}$$

$$+\frac{t^2}{2}e^{-t}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

$$e^{\boldsymbol{A}t} = \begin{pmatrix} -\dfrac{7}{2}t^2e^{-t} & 0 & -8t^2e^{-t} \\ -3t^2e^{-t} & \dfrac{1}{2}t^2e^{-t} & 6t^2e^{-t} \\ 2t^2e^{-t} & 0 & \dfrac{9}{2}t^2e^{-t} \end{pmatrix} + \begin{pmatrix} 3\left(t+t^2\right)e^{-t} & 0 & 8\left(t+t^2\right)e^{-t} \\ 3\left(t+t^2\right)e^{-t} & -\left(t+t^2\right)e^{-t} & 6\left(t+t^2\right)e^{-t} \\ -2\left(t+t^2\right)e^{-t} & 0 & -5\left(t+t^2\right)e^{-t} \end{pmatrix}$$

$$+ \begin{pmatrix} \left(1+t+\dfrac{t^2}{2}\right)e^{-t} & 0 & 0 \\ 0 & \left(1+t+\dfrac{t^2}{2}\right)e^{-t} & 0 \\ 0 & 0 & \left(1+t+\dfrac{t^2}{2}\right)e^{-t} \end{pmatrix}$$

$$= \begin{pmatrix} \left(1+4t\right)e^{-t} & 0 & 8te^{-t} \\ 3te^{-t} & e^{-t} & 6te^{-t} \\ -2te^{-t} & 0 & \left(1-4t\right)e^{-t} \end{pmatrix}$$

我们还可以证明

（Ⅰ） $\forall \boldsymbol{A} \in \mathbf{C}^{n\times n}$，总有

（1） $\sin(-\boldsymbol{A}) = -\sin \boldsymbol{A}$， $\cos(-\boldsymbol{A}) = \cos \boldsymbol{A}$

（2） $e^{i\boldsymbol{A}} = \cos \boldsymbol{A} + \mathrm{i}\sin \boldsymbol{A}$， $\cos \boldsymbol{A} = \frac{1}{2}\left(e^{i\boldsymbol{A}} + e^{-i\boldsymbol{A}}\right)$， $\sin \boldsymbol{A} = \frac{1}{2}\left(e^{i\boldsymbol{A}} - e^{-i\boldsymbol{A}}\right)$

（Ⅱ） $\boldsymbol{A}, \boldsymbol{B} \in \mathbf{C}^{n\times n}$， 且 $\boldsymbol{AB}=\boldsymbol{BA}$，则

（1） $e^{\boldsymbol{A}+\boldsymbol{B}} = e^{\boldsymbol{A}}e^{\boldsymbol{B}} = e^{\boldsymbol{B}}e^{\boldsymbol{A}}$ $=\left(\boldsymbol{I}+\boldsymbol{B}+\frac{\boldsymbol{B}^2}{2!}+\frac{\boldsymbol{B}^3}{3!}+\cdots\right)\left(\boldsymbol{I}+\boldsymbol{A}+\frac{\boldsymbol{A}^2}{2!}+\frac{\boldsymbol{A}^3}{3!}+\cdots\right)$

$$=\left(\boldsymbol{I}+(\boldsymbol{A}+\boldsymbol{B})+\frac{1}{2!}\left(\boldsymbol{A}^2+2\boldsymbol{AB}+\boldsymbol{BA}^2\right)+\cdots\right)$$

$$=\sum_{k=0}^{\infty}\frac{1}{k!}(\boldsymbol{A}+\boldsymbol{B})^k = e^{\boldsymbol{A}+\boldsymbol{B}}$$

（2） $\sin(\boldsymbol{A}+\boldsymbol{B}) = \sin \boldsymbol{A}\cos \boldsymbol{B} + \cos \boldsymbol{A}\sin \boldsymbol{B}$

（3） $\cos(\boldsymbol{A}+\boldsymbol{B}) = \cos \boldsymbol{A}\cos \boldsymbol{B} - \sin \boldsymbol{A}\sin \boldsymbol{B}$

若$\boldsymbol{A}=\boldsymbol{B}$，则 $\cos 2\boldsymbol{A} = \cos^2 \boldsymbol{A} - \sin^2 \boldsymbol{A}$， $\sin 2\boldsymbol{A} = 2\sin \boldsymbol{A}\cos \boldsymbol{A}$

需要指出的是，对任何n阶方阵$\boldsymbol{A}$，$e^{\boldsymbol{A}}$总是可逆矩阵。

对任何n阶方阵A，$\sin A$与$\cos A$不一定可逆。

例如， 取 $A=\begin{pmatrix}\pi & 0\\ 0 & \dfrac{\pi}{2}\end{pmatrix}$，则

$$\sin A=\begin{pmatrix}\sin\pi & 0\\ 0 & \sin\dfrac{\pi}{2}\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$$ 不可逆；

$$\cos A=\begin{pmatrix}\cos\pi & 0\\ 0 & \cos\dfrac{\pi}{2}\end{pmatrix}=\begin{pmatrix}-1 & 0\\ 0 & 0\end{pmatrix}$$ 不可逆。

（1） $\det\left(e^{A}\right)=e^{\mathrm{tr}A}$，（2） $\left(e^{A}\right)^{-1}=e^{-A}$，（3） $\left\|e^{A}\right\|\leq e^{\|A\|}$

（4）若A为**Hermite**阵，则e^{iA}是酉阵。

证：（1）设A的特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$，由定理3（$A$的特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$，则矩阵$f(A)$的特征值为$f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$），知$e^A$的特征值为 $e^{\lambda_1},e^{\lambda_2},\cdots,e^{\lambda_n}$，从而

$$\det\left(e^{A}\right)=e^{\lambda_1}e^{\lambda_2}\cdots e^{\lambda_n}=e^{\lambda_1+\lambda_2+\cdots+\lambda_n}=e^{\mathrm{tr}A}。$$

（2）由于 $\det\left(e^{A}\right)=e^{\mathrm{tr}A}\neq 0$，故$e^{A}$总是可逆的，从而

$$e^{A}e^{-A}=e^{A-A}=e^{0}=I\Rightarrow\left(e^{A}\right)^{-1}=e^{-A}$$

（3）由

$$\left\|\boldsymbol{S}_N\right\| = \left\|\sum_{k=0}^{N}\frac{1}{k!}\boldsymbol{A}^k\right\| \le \sum_{k=0}^{N}\frac{1}{k!}\left\|\boldsymbol{A}\right\|^k$$，则 $N \to \infty$ 时，有

$$\left\|e^{\boldsymbol{A}}\right\| \le e^{\left\|\boldsymbol{A}\right\|}$$

（4）因为 $f\left(\boldsymbol{A}^T\right)=\left(f\left(\boldsymbol{A}\right)\right)^T$，所以

$$\left(e^{i\boldsymbol{A}}\right)^H = e^{\left(i\boldsymbol{A}\right)^H} = e^{-i\boldsymbol{A}^H} = e^{-i\boldsymbol{A}}$$，故

$$\left(e^{i\boldsymbol{A}}\right)^H e^{i\boldsymbol{A}} = e^{-i\boldsymbol{A}}e^{i\boldsymbol{A}} = e^{i(\boldsymbol{A}-\boldsymbol{A})} = e^{\boldsymbol{0}} = I$$

则$e^{i\boldsymbol{A}}$是酉阵。

值得注意：当 $\boldsymbol{AB} \neq \boldsymbol{BA}$时，$e^{\boldsymbol{A}+\boldsymbol{B}} = e^{\boldsymbol{A}}e^{\boldsymbol{B}}$ 或 $e^{\boldsymbol{A}+\boldsymbol{B}} = e^{\boldsymbol{B}}e^{\boldsymbol{A}}$ 不一定成立。

例如

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{则 } \boldsymbol{AB} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \boldsymbol{BA},$$

而且易知，$\boldsymbol{A}$和$\boldsymbol{B}$的特征值均为0；

又 $\boldsymbol{A}+\boldsymbol{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$， $\boldsymbol{A}+\boldsymbol{B}$的特征值为：$\lambda_1 = 1, \lambda_2 = -1$。

显然，有

$\boldsymbol{A}$的Jordan分解为：$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

$\boldsymbol{B}$的Jordan分解为：$\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

由定理3.3可知，

$$e^{A}=\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}e^{0}&\left(\dfrac{\mathrm{d}e^{\lambda}}{\mathrm{d}\lambda}\right)\Bigg|_{\lambda=0}\\&e^{0}\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}=\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}=\begin{pmatrix}1&0\\1&1\end{pmatrix}$$

$$e^{B}=\begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}e^{0}&\left(\dfrac{\mathrm{d}e^{\lambda}}{\mathrm{d}\lambda}\right)\Bigg|_{\lambda=0}\\&e^{0}\end{pmatrix}\begin{pmatrix}1&0\\0&1\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}1&0\\0&1\end{pmatrix}=\begin{pmatrix}1&1\\0&1\end{pmatrix}$$

则

$$e^{A}e^{B}=\begin{pmatrix}1&1\\1&2\end{pmatrix}\neq\begin{pmatrix}2&1\\1&1\end{pmatrix}=e^{B}e^{A}$$

A+***B***的Jordan分解为：

$$\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}1&1\\-1&1\end{pmatrix}\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}\dfrac{1}{2}&\dfrac{1}{2}\\\dfrac{1}{2}&-\dfrac{1}{2}\end{pmatrix}$$

从而

$$e^{\boldsymbol{A}+\boldsymbol{B}} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} e & 0 \\ 0 & e^{-1} \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} e+e^{-1} & e-e^{-1} \\ e-e^{-1} & e+e^{-1} \end{pmatrix}$$

即

$$e^{\boldsymbol{A}+\boldsymbol{B}} \neq e^{\boldsymbol{A}}e^{\boldsymbol{B}} \qquad e^{\boldsymbol{A}+\boldsymbol{B}} \neq e^{\boldsymbol{B}}e^{\boldsymbol{A}}$$

求e^{At}, **sinAt**, 其中 $A=\begin{pmatrix}1&&&\\1&1&&\\&1&1&\\&&1&1\end{pmatrix}$；解：注意到 $A^T=\begin{pmatrix}1&1&&\\&1&1&\\&&1&1\\&&&1\end{pmatrix}$,

则有

$$e^{A^Tt}=\begin{pmatrix}e^t&te^t&\frac{t^2}{2}e^t&\frac{t^3}{6}e^t\\&e^t&te^t&\frac{t^2}{2}e^t\\&&e^t&te^t\\&&&e^t\end{pmatrix};\quad e^{At}=\left(e^{A^Tt}\right)^T=\begin{pmatrix}e^t&&&\\te^t&e^t&&\\\frac{t^2}{2}e^t&te^t&e^t&\\\frac{t^3}{6}e^t&\frac{t^2}{2}e^t&te^t&e^t\end{pmatrix};$$

同理

$$\sin tA=\begin{pmatrix}\sin t&&&\\t\cos t&\sin t&&\\-\frac{t^2}{2}\sin t&t\cos t&\sin t&\\\frac{t^3}{6}\cos t&-\frac{t^2}{2}\sin t&t\cos t&\sin t\end{pmatrix};$$

三、函数矩阵的微积分

在研究微分方程组时，为了简化对问题的表达及求解过程，需要考虑以函数为元素的矩阵的微分和积分； 在研究优化等问题时，则要碰到数量函数对向量变量或矩阵变量的导数， 以及向量值或矩阵值函数对向量变量或矩阵变量的导数。

1 相对于数量变量的微分和积分

定义5 如果矩阵 $\boldsymbol{A}(t)=\left(a_{ij}(t)\right)_{m\times n}$ 的每一个元素 $a_{ij}(t)$ $i=1,2,\cdots,m$; $j=1,2,\cdots,n$ 在$[a,b]$上均为变量t的可微函数， 则称$\boldsymbol{A}(t)$可微，且导数定义为

$$\boldsymbol{A}'(t)=\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)=\left(\frac{\mathrm{d}}{\mathrm{d}t}a_{ij}(t)\right)_{m\times n}$$

例如

$$\boldsymbol{A}(t)=\begin{pmatrix} t+e^t & \sin t \\ t & 4 \end{pmatrix},\quad 则\quad \boldsymbol{A}'(t)=\begin{pmatrix} 1+e^t & \cos t \\ 1 & 0 \end{pmatrix}$$

由定义 5 可以验证矩阵导数的如下运算性质.

定理4 设$\boldsymbol{A}(t)$、$\boldsymbol{B}(t)$是可进行运算的两个可微矩阵，则以下的运算规则成立

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

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

（3）$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\alpha\boldsymbol{A}(t)\right)=\alpha\cdot\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A}(t)$$，其中 α 为任意常数

（4） 当$\boldsymbol{u}=\boldsymbol{f}(\boldsymbol{t})$关于$\boldsymbol{t}$可微时，有

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

（5）当 $\boldsymbol{A}^{-1}(t)$为可微矩阵时，有

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

由于$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}(t))$ 仍是函数矩阵，如果它仍是可导函数矩阵，则可定义其二阶导数。不难给出函数矩阵的高阶导数：

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

只就(2)、(5)证之。

（2）设 $\boldsymbol{A}(t)=\left(a_{ij}(t)\right)_{m\times n}$ ，$\boldsymbol{B}(t)=\left(b_{ij}(t)\right)_{n\times p}$ 则

$$\frac{d}{dt}(\boldsymbol{A}(t)\boldsymbol{B}(t))=\frac{d}{dt}\left(\sum_{k=1}^{n}a_{ik}(t)b_{kj}(t)\right)_{m\times p}=\left(\sum_{k=1}^{n}\left[\frac{d}{dt}\left(a_{ik}(t)b_{kj}(t)\right)\right]\right)_{m\times p}$$

$$=\left[\sum_{k=1}^{n}\left[\frac{d}{dt}\left(a_{ik}(t)\right)\cdot b_{kj}(t)+a_{ik}(t)\cdot\frac{d}{dt}\left(b_{kj}(t)\right)\right]\right]_{m\times p}$$

$$=\left[\sum_{k=1}^{n}\left(\frac{d}{dt}\left(a_{ik}(t)\right)\right)\cdot b_{kj}(t)\right]_{m\times p}+\left[\sum_{k=1}^{n}a_{ik}(t)\cdot\left(\frac{d}{dt}\left(b_{kj}(t)\right)\right)\right]_{m\times p}$$

$$=\frac{d}{dt}(\boldsymbol{A}(t))\boldsymbol{B}(t)+\boldsymbol{A}(t)\frac{d}{dt}\boldsymbol{B}(t)$$

（5） 由于$\boldsymbol{A}(t)^{-1}\boldsymbol{A}(t)=\boldsymbol{I}$， 由性质(2)，两端对$t$求导得

从而 $$\frac{d}{dt}\left(\boldsymbol{A}^{-1}(t)\right)\boldsymbol{A}(t) = -\boldsymbol{A}^{-1}(t)\frac{d}{dt}\left(\boldsymbol{A}(t)\right)\boldsymbol{A}^{-1}(t)$$ 证毕

注： $\frac{d}{dt}\left(\boldsymbol{A}^{m}(t)\right)=m\boldsymbol{A}^{m-1}(t)\frac{d}{dt}\left(\boldsymbol{A}(t)\right)$ 不一定成立。

例如, 对$\boldsymbol{m=2}$, 并取

$$\boldsymbol{A}(t)=\begin{pmatrix} t^2 & t \\ 0 & t \end{pmatrix},\quad \text{则}\ \frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{A}(t)\right)=\begin{pmatrix} 2t & 1 \\ 0 & 1 \end{pmatrix},$$

又 $$\boldsymbol{A}^2(t)=\begin{pmatrix} t^4 & t^3+t^2 \\ 0 & t^2 \end{pmatrix},\quad \frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{A}^2(t)\right)=\begin{pmatrix} 4t^3 & 3t^2+2t \\ 0 & 2t \end{pmatrix},$$

而
$$2\boldsymbol{A}(t)\frac{d}{dt}\left(\boldsymbol{A}(t)\right)=2\begin{pmatrix} t^2 & t \\ 0 & t \end{pmatrix}\begin{pmatrix} 2t & 1 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 4t^3 & 2t^2+2t \\ 0 & 2t \end{pmatrix}$$

故$\frac{d}{dt}\left(\boldsymbol{A}^2(t)\right)\neq 2\boldsymbol{A}(t)\frac{d}{dt}\left(\boldsymbol{A}(t)\right)$。只有 $\boldsymbol{A}(t)\frac{d}{dt}\left(\boldsymbol{A}(t)\right)=\frac{d}{dt}\left(\boldsymbol{A}(t)\right)\boldsymbol{A}(t)$,时成立。

定理5 设n阶方阵$\boldsymbol{A}$与t无关，则有

（1）$$\frac{\mathrm{d}}{\mathrm{d}\,t}e^{t\boldsymbol{A}} = \boldsymbol{A}e^{t\boldsymbol{A}} = e^{t\boldsymbol{A}}\boldsymbol{A}$$

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

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

证明 只证(1)，(2)和(3)的证明与(1)类似。

由 $e^{t\boldsymbol{A}} = \sum_{k=0}^{\infty}\frac{t^k}{k!}\boldsymbol{A}^k$并利用绝对收敛的级数可以逐项求导的性质得

$$\frac{\mathrm{d}\left(e^{t\boldsymbol{A}}\right)}{\mathrm{d}\,t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{k=0}^{\infty}\frac{t^k}{k!}\boldsymbol{A}^k\right) = \sum_{k=0}^{\infty}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t^k}{k!}\right)\boldsymbol{A}^k = \left(\sum_{k=1}^{\infty}\frac{t^{k-1}}{(k-1)!}\boldsymbol{A}^{k-1}\right)$$

$$= \left(\sum_{k=1}^{\infty}\frac{t^{k-1}}{(k-1)!}\boldsymbol{A}^{k-1}\right)\boldsymbol{A} \Rightarrow \boldsymbol{A}e^{t\boldsymbol{A}} = e^{t\boldsymbol{A}}\boldsymbol{A}\text{。}$$

$$\cos \boldsymbol{A} = \boldsymbol{I} - \frac{\boldsymbol{A}^2}{2!} + \frac{\boldsymbol{A}^4}{4!} - \cdots;$$

下面利用性质2， 进行矩阵计算 。

练习1 已知 $\sin t\boldsymbol{A} = \frac{1}{4}\begin{pmatrix} \sin 5t + \sin 3t & 2\sin 5t - 2\sin t \\ \sin 5t - \sin t & 2\sin 5t + 2\sin t \end{pmatrix}$, 求$A$

解： 由于

$$(\sin t\boldsymbol{A})' = \frac{1}{4}\begin{pmatrix} 5\mathrm{con}5t + 3\cos 3t & 10\cos 5t - 2\cos t \\ 5\cos 5t - \cos t & 10\cos 5t + 2\cos t \end{pmatrix} = \boldsymbol{A}\cos \boldsymbol{A}t$$

令t=0, 并注意**cos0=*I***， 则

$$\boldsymbol{A} = \frac{1}{4}\begin{pmatrix} 5\mathrm{con}0 + 3\cos 0 & 10\cos 0 - 2\cos 0 \\ 5\cos 0 - \cos 0 & 10\cos 0 + 2\cos 0 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 8 & 8 \\ 4 & 12 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$$

练习2

已知 $e^{tA}=\begin{pmatrix} 2e^{2t}-e^{t} & e^{2t}-e^{t} & e^{t}-e^{2t} \\ e^{2t}-e^{t} & 2e^{2t}-e^{t} & e^{t}-e^{2t} \\ 3e^{2t}-3e^{t} & 3e^{2t}-3e^{t} & 3e^{t}-2e^{2t} \end{pmatrix}$，求$\boldsymbol{A}$

解： 由于

$$\left(e^{tA}\right)'=\begin{pmatrix} 4e^{2t}-e^{t} & 2e^{2t}-e^{t} & e^{t}-2e^{2t} \\ 2e^{2t}-e^{t} & 4e^{2t}-e^{t} & e^{t}-2e^{2t} \\ 6e^{2t}-3e^{t} & 6e^{2t}-3e^{t} & 3e^{t}-4e^{2t} \end{pmatrix}=\boldsymbol{A}e^{\boldsymbol{A}t}$$

令t=0 ，并注意$\boldsymbol{e^0=I}$， 则

$$\boldsymbol{A}=\begin{pmatrix} 4e^{0}-e^{0} & 2e^{0}-e^{0} & e^{0}-2e^{0} \\ 2e^{0}-e^{0} & 4e^{0}-e^{0} & e^{0}-2e^{0} \\ 6e^{0}-3e^{0} & 6e^{0}-3e^{0} & 3e^{0}-4e^{0} \end{pmatrix}=\begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ 3 & 3 & -1 \end{pmatrix}$$

定义6 如果矩阵 $\boldsymbol{A}(t)=\left(a_{ij}(t)\right)_{m\times n}$ 的每一个元素$a_{ij}(t)$都是区间 $[t_0,t_1]$上的可积函数，则定义$\boldsymbol{A}(t)$在区间$[t_0, t_1]$上的积分为

$$\int_{t_0}^{t_1}\boldsymbol{A}(t)\,\mathrm{d}t=\left(\int_{t_0}^{t_1}a_{ij}(t)\,\mathrm{d}t\right)_{m\times n}$$

例如 $\boldsymbol{A}(t)=\begin{pmatrix} t+e^t & \sin t \\ t & 4 \end{pmatrix}$，则 $\int_0^1\boldsymbol{A}(t)\,\mathrm{d}t=\begin{pmatrix} \frac{1}{2}+e-1 & 1-\cos 1 \\ \frac{1}{2} & 4 \end{pmatrix}$

容易验证如下运算法则成立

（1）$\int_{t_0}^{t_1}\left(\alpha\boldsymbol{A}(t)+\beta\boldsymbol{B}(\boldsymbol{t})\right)\mathrm{d}t=\alpha\int_{t_0}^{t_1}\boldsymbol{A}(t)\,\mathrm{d}t+\beta\int_{t_0}^{t_1}\boldsymbol{B}(t)\,\mathrm{d}t \quad \forall\alpha,\beta\in\mathbf{C}$

（2）$\int_{t_0}^{t_1}\left(\boldsymbol{A}(t)\boldsymbol{B}\right)\mathrm{d}t=\int_{t_0}^{t_1}\boldsymbol{A}(t)\,\mathrm{d}t\,\boldsymbol{B}$， 其中$\boldsymbol{B}$为常数矩阵；

$\int_{t_0}^{t_1}\left(\boldsymbol{A}\,\boldsymbol{B}(t)\right)\mathrm{d}t=\boldsymbol{A}\int_{t_0}^{t_1}\boldsymbol{B}(t)\,\mathrm{d}t$， 其中$\boldsymbol{A}$为常数矩阵；

（3）当$\boldsymbol{A}(t)$在$[a，b]$上连续可微时，对任意$\boldsymbol{t}\in(\boldsymbol{a},\boldsymbol{b})$，有

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\int_a^t \boldsymbol{A}(\tau)\,\mathrm{d}\tau\right)=\boldsymbol{A}(t)$$

（4）当$\boldsymbol{A}(t)$在$[a，b]$上连续可微时，对任意$\boldsymbol{t}\in(\boldsymbol{a},\boldsymbol{b})$，有

$$\int_a^b \frac{\mathrm{d}\left(\boldsymbol{A}(t)\right)}{\mathrm{d}t}\mathrm{d}t = \boldsymbol{A}(b)-\boldsymbol{A}(a)$$

2 相对于矩阵变量的微分

定义7 设 $X=\left(x_{ij}\right)_{m\times n}$，函数 $f(\mathbf{X})=f(x_{11},x_{12},\cdots x_{1n},x_{21},\cdots,x_{mn})$ 为***mn***元的多元函数，且 $\dfrac{\partial f}{\partial x_{ij}}\,(i=1,2,\cdots,m;\quad j=1,2,\cdots,n)$ 都存在，定义 $\boldsymbol{f}(\boldsymbol{X})$ 对矩阵***X***的导数为

$$\frac{\mathrm{d}f\left(\mathbf{X}\right)}{\mathrm{d}\boldsymbol{X}}=\left(\frac{\partial f}{\partial x_{ij}}\right)_{m\times n}=\begin{pmatrix}\dfrac{\partial f}{\partial x_{11}} & \cdots & \dfrac{\partial f}{\partial x_{1n}}\\ \cdots & \ddots & \cdots\\ \dfrac{\partial f}{\partial x_{m1}} & \cdots & \dfrac{\partial f}{\partial x_{mn}}\end{pmatrix}$$

例7 设$\boldsymbol{x}=(\xi_1,\xi_2,\cdots\xi_n)^T$,$n$元函数 $f(\boldsymbol{x})=f(\xi_1,\xi_2,\cdots\xi_n)$，求 $\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}^T},\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}$。

解 根据定义有

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}^T}=\left(\frac{\partial f}{\partial\xi_1},\frac{\partial f}{\partial\xi_2},\cdots,\frac{\partial f}{\partial\xi_n}\right)$$

以x为自变量的函数的导数为：

$$\nabla f(\boldsymbol{x})=\frac{\mathbf{d}f}{\mathbf{d}\boldsymbol{x}}=\left(\frac{\partial f}{\partial\xi_1},\frac{\partial f}{\partial\xi_2},\cdots,\frac{\partial f}{\partial\xi_n}\right)^T$$

称为数量函数对向量变量的导数，即为高等数学学过的函数的梯度向量，也记为 **grag *f*** 。

例7-1，设 $\boldsymbol{a}=(a_1, a_2, \cdots, a_n)^T$ 为常向量，$\boldsymbol{x}=(\xi_1, \xi_2, \cdots, \xi_n)^T$ 为向量变量, 且 $f(\boldsymbol{x})=(\boldsymbol{x}, \boldsymbol{a})=\boldsymbol{a}^T\boldsymbol{x}=\boldsymbol{x}^T\boldsymbol{a}$, 求 $\dfrac{\partial f}{\partial \boldsymbol{x}}$。

解 由于 $f(\boldsymbol{x})=\sum\limits_{i=1}^{n} a_i\xi_i$，$\dfrac{\partial f}{\partial \xi_j}=a_j$，$(j=1, 2, \cdots, n)$

所以

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}=\begin{pmatrix}\dfrac{\partial f}{\partial \xi_1}\\ \dfrac{\partial f}{\partial \xi_2}\\ \vdots\\ \dfrac{\partial f}{\partial \xi_n}\end{pmatrix}=\begin{pmatrix}a_1\\ a_2\\ \vdots\\ a_n\end{pmatrix}=\boldsymbol{a}$$

例8 设$\boldsymbol{x}=(\xi_1, \xi_2, \cdots, \xi_n)^T$, $\boldsymbol{A}=(a_{ij})_{n\times n}$，$n$元函数$\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$，求 $\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}$。

解： 因 $f(\boldsymbol{x})=\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}a_{ij}\xi_i\,\xi_j=\xi_1\sum\limits_{j=1}^{n}a_{1j}\xi_j+\cdots+\xi_k\sum\limits_{j=1}^{n}a_{kj}\xi_j+\cdots+\xi_n\sum\limits_{j=1}^{n}a_{nj}\xi_j$

所以 $$\frac{\partial f(\boldsymbol{x})}{\partial \xi_k}=\xi_1 a_{1k}+\cdots+\xi_{k-1}a_{k-1,k}+\left(\sum_{j=1}^{n}a_{kj}\xi_j+\xi_k a_{kk}\right)+\xi_{k+1}a_{k+1,k}+\cdots+\xi_n a_{nk}$$

$$=\sum_{i=1}^{n}a_{ik}\xi_i+\sum_{j=1}^{n}a_{kj}\xi_j,\qquad k=1,2,\cdots,n$$

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}=\begin{pmatrix}\dfrac{\partial f}{\partial\xi_1}\\ \dfrac{\partial f}{\partial\xi_2}\\ \vdots\\ \dfrac{\partial f}{\partial\xi_n}\end{pmatrix}=\begin{pmatrix}\sum\limits_{j=1}^{n}a_{1j}\xi_j\\ \sum\limits_{j=1}^{n}a_{2j}\xi_j\\ \vdots\\ \sum\limits_{j=1}^{n}a_{nj}\xi_j\end{pmatrix}+\begin{pmatrix}\sum\limits_{i=1}^{n}a_{i1}\xi_i\\ \sum\limits_{i=1}^{n}a_{i2}\xi_i\\ \vdots\\ \sum\limits_{i=1}^{n}a_{in}\xi_i\end{pmatrix}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{A}^T\boldsymbol{x}=(\boldsymbol{A}+\boldsymbol{A}^T)\boldsymbol{x}$$

特别地, 当$\boldsymbol{A}$时对称矩阵时, $\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}=2\boldsymbol{A}\boldsymbol{x}$

练习2 设 $A\in\mathbf{R}^{m\times n}, \boldsymbol{b}\in\mathbf{R}^{m}$， $\boldsymbol{x}\in\mathbf{R}^{n}$, $f(\boldsymbol{x})=\|A\boldsymbol{x}-\boldsymbol{b}\|_2^2$，试求 $\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}$

解：因为

$$f(\boldsymbol{x})=\|A\boldsymbol{x}-\boldsymbol{b}\|_2^2=(A\boldsymbol{x}-\boldsymbol{b},\ A\boldsymbol{x}-\boldsymbol{b})=(A\boldsymbol{x}-\boldsymbol{b})^T(A\boldsymbol{x}-\boldsymbol{b})$$

$$=(\boldsymbol{x}^T A^T-\boldsymbol{b}^T)(A\boldsymbol{x}-\boldsymbol{b})$$

$$=\boldsymbol{x}^T A^T A\boldsymbol{x}-\boldsymbol{b}^T A\boldsymbol{x}-\boldsymbol{x}^T A^T\boldsymbol{b}+\boldsymbol{b}^T\boldsymbol{b}$$

$$=\boldsymbol{x}^T(A^T A)\boldsymbol{x}-(A^T\boldsymbol{b})^T\boldsymbol{x}-\boldsymbol{x}^T(A^T\boldsymbol{b})+\boldsymbol{b}^T\boldsymbol{b}$$

从而, 由例7-1、例8, 可得

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}=2A^T A\boldsymbol{x}-A^T\boldsymbol{b}-A^T\boldsymbol{b}=2\left(A^T A\boldsymbol{x}-A^T\boldsymbol{b}\right)$$

返回

3 矩阵函数在微分方程中的应用

在线性控制系统中，常常涉及求解线性微分方程组的问题。矩阵函数在其中有重要的应用。

我们首先讨论一阶线性常系数齐次微分方程组的定解问：

$$\begin{cases} \dfrac{\mathbf{d}x_1(t)}{\mathbf{d}t} = a_{11}x_1(t) + a_{12}x_2(t) + \cdots + a_{1n}x_n(t) \\ \dfrac{\mathbf{d}x_2(t)}{\mathbf{d}t} = a_{21}x_1(t) + a_{22}x_2(t) + \cdots + a_{2n}x_n(t) \\ \qquad\vdots \\ \dfrac{\mathbf{d}x_n(t)}{\mathbf{d}t} = a_{n1}x_1(t) + a_{n2}x_2(t) + \cdots + a_{nn}x_n(t) \end{cases}$$

给定初始条件：$x_i(0)$，$(i = 1, 2, \cdots, n)$，记 $\boldsymbol{A} = \left(a_{ij}\right) \in \mathbf{C}^{n\times n}$，

$$\boldsymbol{X}(0) = \left(x_1(0), x_2(0), \cdots, x_n(0)\right)^T$$

$$\boldsymbol{X}(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T,$$

则上述微分方程组可写成：

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = \boldsymbol{A}\boldsymbol{X}(t) & (2) \\ \boldsymbol{X}(0) = (x_1(0), x_2(0), \ldots, x_n(0))^T & (3) \end{cases}$$

利用矩阵微分的性质有

$$\frac{\mathrm{d}(e^{-\boldsymbol{A}t}\boldsymbol{X}(t))}{\mathrm{d}t} = \frac{\mathrm{d}e^{-\boldsymbol{A}t}}{\mathrm{d}t}\boldsymbol{X}(t) + e^{-\mathbf{A}t}\frac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t}$$

$$= -e^{-\boldsymbol{A}t}\boldsymbol{A}\boldsymbol{X}(t) + e^{-\mathbf{A}t}\frac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = e^{-\boldsymbol{A}t}\left(\frac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} - \boldsymbol{A}\boldsymbol{X}(t)\right)$$

方程（2）意味着 $$\frac{\mathrm{d}(e^{-\mathbf{A}t}\boldsymbol{X}(t))}{\mathrm{d}t} = \boldsymbol{0}$$

因此 **X=e^{At}C**，其中**C**为常数向量，由初始条件（**3**），**C=X(0)**

$$\boldsymbol{X}(t) = e^{\mathbf{A}t}\boldsymbol{X}(0) \qquad (4)$$

下面说明解的唯一性。

如果定解问题（2）和(3)有两个解$\boldsymbol{X}_1(t)$，$\boldsymbol{X}_2(t)$，则令$\boldsymbol{Y}(t)=\boldsymbol{X}_1(t)-\boldsymbol{X}_2(t)$，显然满足

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{Y}(t)}{\mathrm{d}t}=\boldsymbol{A}\boldsymbol{Y}(t) \\ \boldsymbol{Y}(0)=\boldsymbol{X}_1(0)-\boldsymbol{X}_2(0)=\boldsymbol{0} \end{cases}$$

由上述推导可知，$\boldsymbol{Y}(t)=e^{At}\boldsymbol{Y}(0)=\boldsymbol{0}$，即$\boldsymbol{X}_1(t)=\boldsymbol{X}_2(t)$。

综上所述，

定理6 一阶线性常系数齐次微分方程组的定解问题（2）**(3)** 有唯一解$\boldsymbol{X}(t)=\mathrm{e}^{At}\boldsymbol{X}(0)$。

最后我们考虑一阶线性常系数非齐次微分方程组的定解问题

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = \boldsymbol{AX}(t) + \boldsymbol{F}(t) & （4） \\ \boldsymbol{X}(0) = (x_1(0), x_2(0), \ldots, x_n(0))^T & （5） \end{cases}$$

这里 $\boldsymbol{F}(t) = (f_1(t), f_2(t), \cdots, f_n(t))^T$ 是已知向量函数，$\boldsymbol{A}$和$\boldsymbol{X}$意义同前。

改写方程为　并以$\boldsymbol{e}^{-\boldsymbol{A}t}$左乘方程两边，即

$$e^{-\mathbf{A}t}\left(\frac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} - \boldsymbol{AX}(t)\right) = e^{-\mathbf{A}t}\,\boldsymbol{F}(t)$$

即 $\dfrac{\mathrm{d}(e^{-\mathbf{A}t}\boldsymbol{X}(t))}{\mathrm{d}t} = e^{-\mathbf{A}t}\boldsymbol{F}(t)$ 对此方程在$[t_0, t]$上进行积分，可得

$$e^{-\mathbf{A}t}\boldsymbol{X}(t) - e^{-\mathbf{A}t_0}\boldsymbol{X}(t_0) = \int_{t_0}^{t} e^{-\mathbf{A}\tau}\boldsymbol{F}(\tau)\mathrm{d}\tau$$

$$e^{-\boldsymbol{A}t}\boldsymbol{X}(t) = e^{-\boldsymbol{A}t_0}\boldsymbol{X}(t_0) + \int_{t_0}^{t} e^{-\boldsymbol{A}\tau}\boldsymbol{F}(\tau)\mathbf{d}\tau$$

$\boldsymbol{X}(t) = e^{\mathbf{A}(t-t_0)}\boldsymbol{X}(t_0) + \int_{t_0}^{t} e^{\mathbf{A}(t-\tau)}\boldsymbol{F}(\tau)\mathrm{d}\tau$　就是上述定解问题的解。

一阶线性常系数微分方程组在$[t_0, t]$上的解：

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = \boldsymbol{AX}(t) \\ \boldsymbol{X}(t_0) = \left(x_1(t_0), x_2(t_0), \cdots, x_n(t_0)\right)^T \end{cases}$$

$$\boldsymbol{X}(t) = e^{A(t-t_0)}\boldsymbol{X}(t_0)$$

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = \boldsymbol{AX}(t) + \boldsymbol{F}(t) \\ \boldsymbol{X}(t_0) = \left(x_1(t_0), x_2(t_0), \cdots, x_n(t_0)\right)^T \end{cases}$$

$$\boldsymbol{X}(t) = e^{A(t-t_0)}\boldsymbol{X}(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}\boldsymbol{F}(\tau)\mathrm{d}\tau$$

例9 求定解问题 $\begin{cases} \dfrac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = \boldsymbol{A}\boldsymbol{X}(t) \\ \boldsymbol{X}(0) = (1,1,1)^T \end{cases}$ 其中 $\boldsymbol{A} = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & -1 & 2 \end{pmatrix}$

解 $\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda-3 & 1 & -1 \\ -2 & \lambda & 1 \\ -1 & 1 & \lambda-2 \end{vmatrix} = \lambda(\lambda-2)(\lambda-3),$

故$\boldsymbol{A}$有三个不同的特征根，$\boldsymbol{A}$可与对角形矩阵相似。 与特征根

$\lambda_1=0, \lambda_2=2, \lambda_3=3$相应的三个线性无关的特征向量分别为：

$$\boldsymbol{X}_1 = (1,5,2)^T,\ \boldsymbol{X}_2 = (1,1,0)^T,\ \boldsymbol{X}_3 = (2,1,1)^T$$

进一步得 $\boldsymbol{T} = \begin{pmatrix} 1 & 1 & 2 \\ 5 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{T}^{-1} = -\dfrac{1}{6}\begin{pmatrix} 1 & -1 & -1 \\ -3 & -3 & 9 \\ -2 & 2 & -4 \end{pmatrix}$

由定理6可得所求的解为

$$\boldsymbol{X} = e^{\mathrm{A}t}\boldsymbol{X}(0) = \boldsymbol{T}\begin{pmatrix} 1 & & \\ & e^{2t} & \\ & & e^{3t} \end{pmatrix}\boldsymbol{T}^{-1}\boldsymbol{X}(0) = \begin{pmatrix} 1 & 1 & 2 \\ 5 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & & \\ & e^{2t} & \\ & & e^{3t} \end{pmatrix}\left(-\frac{1}{6}\right)\begin{pmatrix} 1 & -1 & -1 \\ -3 & -3 & 9 \\ -2 & 2 & -4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\boldsymbol{X}(t) = -\frac{1}{6}\begin{pmatrix} -1+3e^{2t}-8e^{3t} \\ -5+3e^{2t}-4e^{3t} \\ -2-4e^{3t} \end{pmatrix}$$

例10 求定解问题 $\begin{cases}\dfrac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t}=\boldsymbol{AX}(t)+\boldsymbol{F}(t)\\ \boldsymbol{X}(0)=(1,1,1)^T\end{cases}$ 的解, 其中矩阵$\boldsymbol{A}$见例9。 $\boldsymbol{F}(t)=\left(0\,,0\,,e^{2t}\right)^T$

解 由前面讨论，该问题的解为 $\boldsymbol{X}(t)=e^{\mathbf{A}t}\boldsymbol{X}(0)+\int_0^t e^{\mathbf{A}(t-\tau)}\boldsymbol{F}(\tau)\mathrm{d}\tau$

下面计算 $\boldsymbol{P}=\int_0^t e^{\mathbf{A}(t-\tau)}\boldsymbol{F}(\tau)\mathrm{d}\tau$,由 $e^{\mathbf{A}(t-\tau)}\boldsymbol{F}(\tau)=\boldsymbol{T}e^{[\boldsymbol{J}(t-\tau)]}\boldsymbol{T}^{-1}\boldsymbol{F}(\tau)$

$$=\begin{pmatrix}1&1&2\\5&1&1\\2&0&1\end{pmatrix}\begin{pmatrix}1&&\\&e^{2(t-\tau)}&\\&&e^{3(t-\tau)}\end{pmatrix}\left(-\frac{1}{6}\right)\begin{pmatrix}1&-1&-1\\-3&-3&9\\-2&2&-4\end{pmatrix}\begin{pmatrix}0\\0\\e^{2\tau}\end{pmatrix}=\left(\frac{-1}{6}\right)\begin{pmatrix}1&1&2\\5&1&1\\2&0&1\end{pmatrix}\begin{pmatrix}-e^{2\tau}\\9e^{2\tau}\\-4e^{2\tau}\end{pmatrix}$$

$$=-\frac{1}{6}\begin{pmatrix}-e^{2\tau}+9e^{2t}-8e^{3t-\tau}\\-5e^{2\tau}+9e^{2t}-4e^{3t-\tau}\\-2e^{2\tau}-4e^{3t-\tau}\end{pmatrix}$$

。将这一结果对变量τ从0到t进行积分, 即得

$$\boldsymbol{P}=-\frac{1}{6}\begin{pmatrix}\frac{1}{2}+(9t+\frac{15}{2})e^{2t}-8e^{3t}\\\frac{5}{2}+(9t+\frac{3}{2})e^{2t}-4e^{3t}\\1+3e^{2t}-4e^{3t}\end{pmatrix}$$

因此 $\boldsymbol{X}(t)=\mathrm{e}^{At}\boldsymbol{X}(0)+\boldsymbol{P}$

$$\boldsymbol{X}(t)=-\frac{1}{6}\begin{pmatrix}-\frac{1}{2}+(9t+\frac{21}{2})e^{2t}-16e^{3t}\\-\frac{5}{2}+(9t+\frac{9}{2})e^{2t}-8e^{3t}\\-1+3e^{2t}-8e^{3t}\end{pmatrix}$$

THE END

定理2.12（Hamilton-Caylay） 设 $\boldsymbol{A}\in\mathbf{C}^{n\times n}$, $\psi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})$,

则
$$\psi(\boldsymbol{A})=\boldsymbol{0}$$

证 存在$\boldsymbol{P}\in\mathbf{C}^{n\times n}$, 使得$\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{J}$，其中$\boldsymbol{J}$是$\boldsymbol{A}$的**Jordan**标准型，

可以写成为：

$$\boldsymbol{J}=\begin{pmatrix}\lambda_1 & \delta & & \\ & \lambda_2 & \ddots & \\ & & \ddots & \delta \\ & & & \lambda_n\end{pmatrix}$$

（ δ或为1或为0）

由于λ $_1$, λ $_2$, …, λ $_n$,是的特征值, 于是

$$\psi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

$$\psi(\boldsymbol{A})=(\boldsymbol{A}-\lambda_1\boldsymbol{I})(\boldsymbol{A}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{A}-\lambda_n\boldsymbol{I})$$

$$=(\boldsymbol{PJP}^{-1}-\lambda_1\boldsymbol{I})(\boldsymbol{PJP}^{-1}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{PJP}^{-1}-\lambda_n\boldsymbol{I})$$

$$=\boldsymbol{P}(\boldsymbol{J}-\lambda_1\boldsymbol{I})(\boldsymbol{J}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{J}-\lambda_n\boldsymbol{I})\boldsymbol{P}^{-1}$$

$$\psi(\boldsymbol{A})=\boldsymbol{P}(\boldsymbol{J}-\lambda_1\boldsymbol{I})(\boldsymbol{J}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{J}-\lambda_n\boldsymbol{I})\boldsymbol{P}^{-1}$$

$$=\boldsymbol{p}\begin{pmatrix}0 & \delta & & \\ & \lambda_2-\lambda_1 & \ddots & \\ & & \ddots & \delta \\ & & & \lambda_n-\lambda_1\end{pmatrix}\begin{pmatrix}\lambda_1-\lambda_2 & \delta & & \\ & 0 & \ddots & \\ & & \ddots & \delta \\ & & & \lambda_n-\lambda_2\end{pmatrix}\cdots\begin{pmatrix}\lambda_1-\lambda_n & \delta & & \\ & \ddots & \ddots & \\ & & \lambda_{n-1}-\lambda_n & \delta \\ & & & 0\end{pmatrix}\boldsymbol{p}^{-1}$$

$$=\boldsymbol{p}\begin{pmatrix}0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & *\end{pmatrix}\begin{pmatrix}\lambda_1-\lambda_3 & \delta & & & \\ & \lambda_2-\lambda_3 & \delta & & \\ & & 0 & \delta & \\ & & & \ddots & \delta \\ & & & & \lambda_n-\lambda_3\end{pmatrix}\cdots\begin{pmatrix}\lambda_1-\lambda_n & \delta & & \\ & \ddots & \ddots & \\ & & \lambda_{n-1}-\lambda_n & \delta \\ & & & 0\end{pmatrix}\boldsymbol{p}^{-1}$$

$$=\cdots=\boldsymbol{O}_{n\times n}$$

以下用例子说明**Hamilton-Caylay**定理在简化矩阵计算中的应用

例 已知矩阵 $\boldsymbol{A}=\begin{pmatrix}-1 & 1 & 0\\ -4 & 3 & 0\\ 1 & 0 & 2\end{pmatrix}$，试计算

（**1**） $\boldsymbol{A}^7-\boldsymbol{A}^5-19\boldsymbol{A}^4+28\boldsymbol{A}^3+6\boldsymbol{A}-4\boldsymbol{I}$；

（**2**） $\boldsymbol{A}^{-1}$，（**3**） $\boldsymbol{A}^{100}$。

解 取 $\psi(\lambda)=\det\left(\lambda\boldsymbol{I}-\boldsymbol{A}\right)=\begin{vmatrix}\lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2\end{vmatrix}=\lambda^3-4\lambda^2+5\lambda-2,$

（**1**）令 $f(\lambda)=\lambda^7-\lambda^5-19\lambda^4+28\lambda^3+6\lambda-4$，则只须计算 $\boldsymbol{f(A)}$。

用 $\psi(\lambda)$ 除 $f\left(\lambda\right)$，得 $-3\lambda^2+22\lambda-8$，则

$$f\left(\lambda\right)=\left(\lambda^4+4\lambda^3+10\lambda^2+3\lambda-2\right)\psi\left(\lambda\right)-3\lambda^2+22\lambda-8,$$

由**H-C**定理知，$\psi(\boldsymbol{A})=0$，于是 $f\left(\boldsymbol{A}\right)=-3\boldsymbol{A}^2+22\boldsymbol{A}-8\boldsymbol{I}=\begin{pmatrix}-19 & 6 & 0\\ -64 & 43 & 0\\ 19 & -3 & 24\end{pmatrix}$

（2）由 $\psi(A) = \boldsymbol{A}^3 - 4\boldsymbol{A}^2 + 5\boldsymbol{A} - 2\boldsymbol{I} = \boldsymbol{O}$，可得

$$\frac{1}{2}\left(\boldsymbol{A}^3 - 4\boldsymbol{A}^2 + 5\boldsymbol{A}\right) = \boldsymbol{I},$$

进一步

$$\boldsymbol{A}^{-1}\boldsymbol{A}\left(\frac{1}{2}\left(\boldsymbol{A}^2 - 4\boldsymbol{A} + 5\boldsymbol{I}\right)\right) = \boldsymbol{A}^{-1}$$

从而

$$\boldsymbol{A}^{-1} = \frac{1}{2}\begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}^2 - 2\begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix} + \begin{pmatrix} \frac{5}{2} & 0 & 0 \\ 0 & \frac{5}{2} & 0 \\ 0 & 0 & \frac{5}{2} \end{pmatrix} = \begin{pmatrix} 3 & -1 & 0 \\ 4 & -1 & 0 \\ -\frac{3}{4} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

（3）设① $\lambda^{100}=g(\lambda)\psi(\lambda)+a\lambda^2+b\lambda+c$， 注意到,

$\psi(\lambda)=(\lambda-2)(\lambda-1)^2$ 即有, $\psi(2)=\psi(1)=\psi'(1)=0$，

又 $100\lambda^{99}=g'(\lambda)\psi(\lambda)+g(\lambda)\psi'(\lambda)+2a\lambda+b$，

分别将 $\lambda=2$，$\lambda=1$ 代入①式, 再对上式求导后将 $\lambda=1$代入,得

$$\begin{cases} 2^{100}=4a+2b+c \\ 1=a+b+c \\ 100=2a+b \end{cases} \quad \text{解得} \quad \begin{cases} a=2^{100}-101 \\ b=-2^{101}+302 \\ c=2^{100}-200 \end{cases}$$

故 $$\begin{aligned} \boldsymbol{A}^{100} &= g(\boldsymbol{A})\psi(\boldsymbol{A})+a\boldsymbol{A}^2+b\boldsymbol{A}+c\boldsymbol{I} \\ &= \left(2^{100}-\mathbf{101}\right)\boldsymbol{A}^2+\left(-2^{101}-302\right)\boldsymbol{A}+\left(2^{100}-200\right)\boldsymbol{I} \\ &= \begin{pmatrix} -199 & 100 & 0 \\ -400 & 201 & 0 \\ 201-2^{100} & 2^{100}-101 & 2^{100} \end{pmatrix} \end{aligned}$$