矩阵分析简介

数列

$$a_k \to a, \quad b_k \to b$$

1、 $a_k \to a \; \underset{\not\Leftarrow}{\Rightarrow} \; |a_k| \to |a|$

$a_k \to a \; \rightleftarrows \; |a_k - a| \to 0$

2、 $\alpha a_k \pm \beta b_k \to \alpha a \pm \beta b$

$a_k b_k \to ab$

$a_k^{-1} b_k \to a^{-1} b, (a_k, a \neq 0)$

3、非零数列极限可能为零

矩阵序列

?

矩阵序列收敛 $\Leftrightarrow$ mn个数列收敛

定义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。

矩阵序列收敛 $\Leftrightarrow$ 元素收敛

例1　**讨论矩阵序列** $\{A_k\}_{k=1}^{\infty}$**的收敛性。**

解：**只需求出它的每一个元素的极限即可，极限为：**

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

矩阵序列收敛的性质和数列收敛性质相似

矩阵序列的收敛相当于mn**个数列同时收敛**

可以用初等分析的方法来研究它

同时研究mn个数列的极限很繁琐，我们可以利用矩阵范数来研究矩阵序列的极限。

矩阵收敛 $\Leftrightarrow$ 元素收敛 $\Leftrightarrow$ 范数收敛

定理1 设 $\{A_k\}_{k=1}^{\infty}$ 为 $\mathbf{C}^{m\times n}$ 中的矩阵序列，$\|\cdot\|$ 为 $\mathbf{C}^{m\times n}$ 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。

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

此结论只是充分条件，反过来不一定成立。

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

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

但是矩阵序列 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1

设 $\{A_k\}_{k=1}^{\infty}$ 和 $\{B_k\}_{k=1}^{\infty}$ 为 $\mathrm{C}^{m\times n}$ 中的矩阵序列，并且

$$\lim_{k\to\infty} A_k = A \qquad \lim_{k\to\infty} B_k = B$$

则

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

证：由

$$\|(\alpha A_k + \beta B_k) - (\alpha A + \beta B)\| = \|\alpha(A_k - A) - \beta(B_k - B)\|$$

$$\le |\alpha| \cdot \|A_k - A\| + |\beta| \cdot \|B_k - B\|$$

由定理1，即结论成立。

性质2

设 $\{A_k\}_{k=1}^{\infty}$和 $\{B_k\}_{k=1}^{\infty}$分别为 $C^{m\times n}$ 和 $C^{n\times l}$ 中的矩阵序列，

并且

$$\lim_{k\to\infty} A_k = A \;, \quad \lim_{k\to\infty} B_k = B$$

则

$$\lim_{k\to\infty} A_k B_k = AB$$

证 由

$$\|A_k B_k - AB\| = \|A_k B_k - A_k B + A_k B - AB\|$$

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

由定理1和推论可知，结论成立。

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

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

又有 $\lim\limits_{k\to\infty}\tilde{\boldsymbol{A}}_k=\lim\limits_{k\to\infty}\tilde{\boldsymbol{A}}\neq 0$,

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

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

条件 $A_k(k=1,2,\cdots)$ 和A均为可逆的是不可少的。

矩阵序列可逆不能保证极限一定可逆

例:

$$A_k=\begin{pmatrix}1+\dfrac{1}{k} & 1\\ 1 & 1\end{pmatrix},\quad \text{对于}\mathbf{A}_k(k=1,2,\cdots)\text{都有}(A_k)^{-1}=\begin{pmatrix}k & k\\ -k & k+1\end{pmatrix}$$

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

数列

$$a_k \to a, \quad b_k \to b$$

1、 $a_k \to a \Leftrightarrow |a_k| \to |a|$

$a_k \to a \Leftrightarrow |a_k - a| \to 0$

2、

$\alpha a_k \pm \beta b_k \to \alpha a \pm \beta b$

$a_k b_k \to ab$

$a_k^{-1} b_k \to a^{-1} b, (a_k, a \neq 0)$

3、非零数列极限可能为零

矩阵序列

$$A_k \to A, \quad B_k \to B$$

1、 $A_k \to A \Leftrightarrow |A_k| \to |A|$

$A_k \to A \Leftrightarrow |A_k - A| \to 0$

2、

$\alpha A_k \pm \beta B_k \to \alpha A \pm \beta B$

$A_k B_k \to AB$

$A_k^{-1} B_k \to A^{-1} B,\ A_k, A$可逆

3、可逆矩阵列极限可能奇异

非零矩阵列极限可能为零

矩阵序列收敛 $\Leftrightarrow$ mn个数列收敛

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

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

证 必要性 由定理1知 $\lim\limits_{k\to\infty}\mathbf{A}^k=0$ 的充分必要条件是对任意一种矩阵范数 $\|\cdot\|$ 均有$\lim\limits_{k\to\infty}\left\|\mathbf{A}^k\right\|=0$。因此对充分大的$k$，必有 $\left\|A^k\right\|<1$

因此得

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

利用矩阵谱半径的定义以及相容矩阵范数的性质有： $\rho(\mathbf{A})<1$

充分性 根据定理2.8，对于 $\varepsilon=\dfrac{1}{2}\left(1-\rho(\mathbf{A})\right)>0$一定存在一种相容的矩阵范数$\|\cdot\|$，使得 $\rho(\mathbf{A})+\varepsilon=\dfrac{1}{2}\left(1+\rho(\mathbf{A})\right)<1$ 则

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

于是，$\lim\limits_{k\to\infty}\left\|\mathbf{A}^k\right\|=0$ 根据定理1 即知 $\lim\limits_{k\to\infty}\mathbf{A}^k=0$。

推论 设 $A\in \mathrm{C}^{n\times n}$， $\lim\limits_{k\to\infty}\mathbf{A}^k=0$ 的充分必要条件是存在$\mathrm{C}^{n\times n}$上的某种范数 $\|\cdot\|$ ，使得 $\|A\|<1$

$$\boxed{\rho(\mathbf{A})<1} \Longleftrightarrow \boxed{\lim_{k\to\infty}\mathbf{A}^k=\mathbf{0}} \quad \underset{\not\Longrightarrow}{\Longleftarrow} \quad \boxed{\|A\|<1}$$

判断一个矩阵是否为收敛矩阵:

1、若 A^k 容易计算，则利用其判断收敛性

2、判断矩阵的某种范数是否小于1

3、计算矩阵的谱半径

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

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

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 $\lim\limits_{k\to\infty}\mathbf{A}^k=\boldsymbol{0}$。

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

2、矩阵级数

定义2　设 $\{\mathbf{A}_k\}_{k=1}^{\infty}$ 为$\mathbf{C}^{m\times n}$中的矩阵序列，称

$$\mathbf{A}_1+\mathbf{A}_2+\cdots+\mathbf{A}_k+\cdots$$

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

定义3　记 $\mathbf{S}_k=\sum_{i=1}^{k}\mathbf{A}_i$，称之为矩阵级数 $\sum_{k=1}^{\infty}\mathbf{A}_k$ 的前k项部分和。若矩阵序列 $\{\mathbf{S}_k\}_{k=1}^{\infty}$ 收敛且 $\lim_{k\to\infty}\mathbf{S}_k=\mathbf{S}$，则称矩阵级数 $\sum_{k=1}^{\infty}\mathbf{A}_k$ 收敛，而矩阵S称为矩阵级数的和矩阵，记为 $\mathbf{S}=\sum_{k=1}^{\infty}\mathbf{A}_k$。不收敛的矩阵级数称为发散的。

显然，和 $\sum_{k=1}^{\infty}A_k=\mathbf{S}=(s_{ij})$ 的意义指的是：$\sum_{k=1}^{\infty}a_{ij}^{(k)}=s_{ij}$ $(i=1,2,\cdots,m,\quad j=1,2,\cdots,n)$ 即$m\times n$个数项级数 $\sum_{k=1}^{\infty}a_{ij}^{(k)}$ 均为收敛的。

例　研究矩阵级数 $\sum_{k=0}^{\infty} A_k$ 的收敛性，其中

解：因为

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

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 $\sum_{k=1}^{\infty} A_k$ 收敛，且和为S。

例3 设A为n阶方阵，证明：矩阵级数 $\mathbf{I}_n+\mathbf{A}+\mathbf{A}^2+\cdots+\mathbf{A}^k+\cdots$ 收敛($A_0=I_n$)的充要条件是 $\rho(\mathbf{A})<1$, 而且矩阵级数收敛时有

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

证： 必要性

$$S_k=I_n+A+A^2+\cdots+A^{k-1}$$
$$S_{k+1}=I_n+A+\cdots+A^{k-1}+A^k$$

若矩阵级数 $\sum_{k=1}^{\infty}\mathbf{A}_k$ 收敛，则

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

根据例2，$\rho(\mathbf{A})<1$。

充分性

$$S_k=I_n+A+A^2+\cdots+A^{k-1}$$
$$AS_k=A+A^2+\cdots++A^{k-1}+A^k$$

$\rho(\mathbf{A})<1$, 则I-A可逆且 $\lim_{k\to\infty}\mathbf{A}^k=0$

$$\lim_{k\to\infty}\left(S_k-AS_k\right)=\lim_{k\to\infty}\left(I_n-A^k\right)=I_n$$

则矩阵级数收敛且 $\lim_{k\to\infty}S_k=(\boldsymbol{I}_n-\boldsymbol{A})^{-1}$

★ 级数收敛的必要条件是通项的极限为0.

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

思路：

判断是否是收敛矩阵，若是则收敛到 $(I-A)^{-1}$ 否则不收敛

解：

$$\|\boldsymbol{A}\|_{\infty}=0.9<1 \Rightarrow \rho(\mathbf{A})<1 \Rightarrow \sum_{k=0}^{\infty} \boldsymbol{A}^k \text{收敛，且有}$$

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

数项级数

1、数列所有项的和 $\sum_{k=1}^{\infty} a_k$

2、前n项和 $S_n = \sum_{k=1}^{n} a_k$

3、级数收敛 $\Leftrightarrow \lim_{n\to\infty} S_n = S$

4、数项级数

$$S = 1 + a + \cdots + a^n + \cdots$$

收敛 $\Leftrightarrow |a| < 1$

矩阵级数

1、矩阵列所有项的和 $\sum_{k=1}^{\infty} A_k$

2、前n项和 $S_n = \sum_{k=1}^{n} A_k$

3、级数收敛 $\Leftrightarrow \lim_{n\to\infty} S_n = S$

4、矩阵级数

$$S = I + A + \cdots + A^n + \cdots$$

收敛 $\Leftrightarrow \rho(A) < 1$

矩阵级数收敛 $\Leftrightarrow$ mn个数项级数收敛

矩阵级数　类比于　数项级数

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

性质5 矩阵级数 $\sum_{k=1}^{\infty}\mathbf{A}_k$ 为绝对收敛 ⟺ 正项级数 $\sum_{k=1}^{\infty}\|\mathbf{A}_k\|$ 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。

$$\underbrace{\begin{pmatrix} a_{11}^{(1)} & \cdots & a_{1n}^{(1)} \\ \vdots & \ddots & \vdots \\ a_{m1}^{(1)} & \cdots & a_{mn}^{(1)} \end{pmatrix}}_{A_1} + \underbrace{\begin{pmatrix} a_{11}^{(2)} & \cdots & a_{1n}^{(2)} \\ \vdots & \ddots & \vdots \\ a_{m1}^{(2)} & \cdots & a_{mn}^{(2)} \end{pmatrix}}_{A_2} + \cdots + \underbrace{\begin{pmatrix} a_{11}^{(k)} & \cdots & a_{1n}^{(k)} \\ \vdots & \ddots & \vdots \\ a_{m1}^{(k)} & \cdots & a_{mn}^{(k)} \end{pmatrix}}_{A_k} + \cdots$$

$$\left|a_{ij}^{(1)}\right| \le \|A_1\|_\infty \qquad \|A_1\|_\infty \le \left|a_{11}^{(1)}\right| + \cdots + \left|a_{mn}^{(1)}\right|$$

$$\left|a_{ij}^{(2)}\right| \le \|A_2\|_\infty \qquad \|A_2\|_\infty \le \left|a_{11}^{(2)}\right| + \cdots + \left|a_{mn}^{(2)}\right|$$

$$\vdots \qquad \vdots$$

$$\left|a_{ij}^{(k)}\right| \le \|A_k\|_\infty \qquad \|A_k\|_\infty \le \left|a_{11}^{(k)}\right| + \cdots + \left|a_{mn}^{(k)}\right|$$

$$\sum_{k=1}^{l}\left|a_{ij}^{(k)}\right| \le \sum_{k=1}^{l}\|A_k\|_\infty \qquad \sum_{k=1}^{l}\|A_k\|_\infty \le M_{11} + \cdots + M_{mn}$$

性质 4　若矩阵级数 $\sum_{k=1}^{\infty} \mathbf{A}_k$ 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。

绝对收敛 $\underset{\not\longleftarrow}{\longrightarrow}$ 收敛

性质6

设 $\sum_{k=1}^{\infty}\mathbf{A}_k$ 为$\mathrm{C}^{m\times n}$中的绝对收敛的矩阵级数，$\sum_{k=1}^{\infty}\mathbf{B}_k$ 为$\mathrm{C}^{n\times 1}$中的绝对收敛的级数，并且 $\mathbf{A}=\sum_{k=1}^{\infty}\mathbf{A}_k$，$\mathbf{B}=\sum_{k=1}^{\infty}\mathbf{B}_k$，则

$$\sum_{k=1}^{\infty}\mathbf{A}_k \bullet \sum_{k=1}^{\infty}\mathbf{B}_k$$

按任何方式排列得到的级数也是绝对收敛的，且和为AB.

证：
$$\sum_{k=1}^{\infty} A_k \cdot \sum_{k=1}^{\infty} B_k = (A_1 + A_2 + \cdots A_p + \cdots)(B_1 + B_2 + \cdots + B_p + \cdots)$$
$$= A_1B_1 + A_1B_2 + A_1B_3 + \cdots + A_1B_p + \cdots$$
$$+ A_2B_1 + A_2B_2 + A_2B_3 + \cdots + A_2B_p + \cdots$$
$$+ A_3B_1 + A_3B_2 + A_3B_3 + \cdots + A_3B_p + \cdots$$
$$+ \cdots$$
$$+ A_pB_1 + A_pB_2 + A_pB_3 + \cdots + A_pB_p + \cdots$$
$$+ \cdots$$
$$= \sum_{k=1}^{\infty} C_k, \quad C_k = \sum_{i,j} A_iB_j$$

绝对收敛

有限项和

从而

绝对收敛

$$\sum_{k=1}^{p} \|\mathbf{C}_k\|_\infty \le \sum_{k=1}^{p} \sum_{i,j} \|\mathbf{A}_i\mathbf{B}_j\|_\infty \le \sum_{k=1}^{N_i} \|\mathbf{A}_k\|_\infty \cdot \sum_{k=1}^{N_i} \|\mathbf{B}_k\|_\infty \le \mathrm{M}_{\mathbf{A}} \cdot \mathrm{M}_{\mathbf{B}}$$

$$\sum_{k=1}^{\infty} \boldsymbol{C}_k = \sum_{k=1}^{\infty} \boldsymbol{A}_k \cdot \sum_{k=1}^{\infty} \boldsymbol{B}_k =$$

$$\begin{aligned}
&= \boldsymbol{A}_1\boldsymbol{B}_1 + \boldsymbol{A}_1\boldsymbol{B}_2 + \boldsymbol{A}_1\boldsymbol{B}_3 + \cdots + \boldsymbol{A}_1\boldsymbol{B}_p + \cdots \\
&+ \boldsymbol{A}_2\boldsymbol{B}_1 + \boldsymbol{A}_2\boldsymbol{B}_2 + \boldsymbol{A}_2\boldsymbol{B}_3 + \cdots + \boldsymbol{A}_2\boldsymbol{B}_p + \cdots \\
&+ \boldsymbol{A}_3\boldsymbol{B}_1 + \boldsymbol{A}_3\boldsymbol{B}_2 + \boldsymbol{A}_3\boldsymbol{B}_3 + \cdots + \boldsymbol{A}_3\boldsymbol{B}_p + \cdots \\
&+ \cdots \\
&+ \boldsymbol{A}_p\boldsymbol{B}_1 + \boldsymbol{A}_p\boldsymbol{B}_2 + \boldsymbol{A}_p\boldsymbol{B}_3 + \cdots + \boldsymbol{A}_p\boldsymbol{B}_p + \cdots \\
&+ \cdots
\end{aligned}$$

$$\boldsymbol{C}_1 = \boldsymbol{A}_1\boldsymbol{B}_1, \quad \boldsymbol{C}_2 = \boldsymbol{A}_1\boldsymbol{B}_2 + \boldsymbol{A}_2\boldsymbol{B}_1 + \boldsymbol{A}_2\boldsymbol{B}_2, \quad \boldsymbol{C}_p = \boldsymbol{A}_p \sum_{i=1}^{p} \boldsymbol{B}_i + \sum_{i=1}^{p-1} \boldsymbol{A}_i \boldsymbol{B}_p,$$

$$\sum_{k=1}^{p} \boldsymbol{C}_k = \sum_{k=1}^{p} \boldsymbol{A}_k \cdot \sum_{k=1}^{p} \boldsymbol{B}_k \Rightarrow \sum_{k=1}^{\infty} C_k = AB$$

性质7 设$P \in C^{p\times m}$ 和 $Q \in C^{n\times q}$ 为给定矩阵，如果

$$\sum_{k=0}^{\infty} A^k \text{ 收敛} \Rightarrow \sum_{k=0}^{\infty} PA^kQ \text{ 收敛}$$

$$\sum_{k=0}^{\infty} A^k \text{ 绝对收敛} \Rightarrow \sum_{k=0}^{\infty} PA^kQ \text{ 绝对收敛}$$

且有等式 $$\sum_{k=0}^{\infty} PA^kQ = P\left(\sum_{k=0}^{\infty} A^k\right)Q$$

乘积运算和无穷和运算可交换

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

即 $\sum_{k=0}^{\infty} PA^kQ$ 收敛

$$\sum_{k=0}^{\infty} A^k\text{绝对收敛} \xrightarrow{\text{性质5}} \sum_{k=0}^{\infty} \left\|A^k\right\| \text{ 收敛} \Rightarrow \sum_{k=0}^{\infty} \|P\|\left\|A^k\right\|\|Q\|\text{收敛}$$

$$\xrightarrow{\text{正项级数比较法}} \sum_{k=1}^{\infty} \left\|PA^kQ\right\|\text{收敛} \xrightarrow{\text{性质5}} \sum_{k=0}^{\infty} PA^kQ \text{ 绝对收敛}$$

数项级数	矩阵项级数
1、绝对收敛 $\Rightarrow$ 收敛且级数和不变	1、绝对收敛 $\Rightarrow$ 收敛且级数和不变
2、$\sum_{k=1}^{\infty} a_k$ 绝对收敛 $\Leftrightarrow$ 正项级数 $\sum_{k=1}^{\infty} \lvert a_k \rvert$ 收敛	2、$\sum_{k=1}^{\infty} A_k$ 绝对收敛 $\Leftrightarrow$ 正项级数 $\sum_{k=1}^{\infty} \lVert A_k \rVert$ 收敛
3、$\sum_{k=1}^{\infty} a_k, \sum_{k=1}^{\infty} b_k$ 绝对收敛到 a, b 则 $\sum_{k=1}^{\infty} a_k \sum_{k=1}^{\infty} b_k = ab$	3、$\sum_{k=1}^{\infty} A_k, \sum_{k=1}^{\infty} B_k$ 绝对收敛到 A, B 则 $\sum_{k=1}^{\infty} A_k \sum_{k=1}^{\infty} B_k = AB$
4、$\sum_{k=1}^{\infty} a_k$ 绝对收敛，$\sum_{k=1}^{\infty} b a_k c$ 绝对收敛	4、$\sum_{k=1}^{\infty} A_k$ 绝对收敛，$\sum_{k=1}^{\infty} P A_k Q$ 绝对收敛

矩阵项级数绝对收敛 $\Leftrightarrow$ mn个数项级数绝对收敛

二、矩阵幂级数

矩阵幂级数的一般形式：

$$\sum_{k=0}^{\infty} a_k \boldsymbol{A}^k$$

问题：

1、判断矩阵幂级数是否收敛，若收敛，求收敛矩阵（解析函数）

2、解析函数的矩阵幂级数展开形式

矩阵幂级数收敛条件

矩阵幂级数$\sum_{k=0}^{\infty} a_k \boldsymbol{A}^k$绝对收敛 $\xLeftrightarrow{\text{性质5}}$ 正项级数$\sum_{k=0}^{\infty} \|a_k \boldsymbol{A}^k\|$收敛

$$\|a_k \boldsymbol{A}^k\| = |a_k| \|\boldsymbol{A}^k\| \le |a_k| \|\boldsymbol{A}\|^k$$

范数的齐次性 $\Updownarrow$

正项级数$\sum_{k=0}^{\infty} |a_k| \|\boldsymbol{A}^k\|$收敛

正项级数 $\Uparrow$ 比较判别法

幂级数$\sum_{k=0}^{\infty} a_k x^k$的收敛半径为r

$\rho(A) < r \Rightarrow$ 存在范数$\|A\| < r$

$\Rightarrow$ 正项级数$\sum_{k=0}^{\infty} |a_k| \|A\|^k$收敛

矩阵幂级数发散条件

矩阵幂级数 $\sum_{k=0}^{\infty} a_k A^k$ 发散 $\overset{\text{性质5}}{\Longleftrightarrow}$ 正项级数 $\sum_{k=0}^{\infty} \left\| a_k A^k \right\|$ 发散

$\Updownarrow$ 范数的齐次性

正项级数 $\sum_{k=0}^{\infty} |a_k| \left\| A^k \right\|$ 发散

$\Uparrow$ 正项级数比较判别法

$$|a_k|\left(\rho(A)\right)^k = |a_k|\,\rho\left(A^k\right) \le |a_k|\left\|A^k\right\| = \left\|a_k A^k\right\|$$

$$\left.\begin{array}{l} \text{幂级数} \sum_{k=0}^{\infty} a_k x^k \text{的收敛半径为r} \\ \rho(A) > r \end{array}\right\} \Longrightarrow \text{正项级数} \sum_{k=0}^{\infty} |a_k|\, \rho(A)^k \text{发散}$$

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

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

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

证　（1）如果 $\rho(\mathbf{A})<r$，根据矩阵范数的性质，对于 $\varepsilon=\frac{1}{2}[r-\rho(\mathbf{A})]>0,$ 一定存在一种相容的矩阵范数 $\|\cdot\|$，使得

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

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

$$\|a_k\boldsymbol{A}^k\|=|a_k|\|\boldsymbol{A}^k\|\le|a_k|\|\boldsymbol{A}\|^k$$

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

再利用矩阵级数性质2知，$\sum_{k=0}^{\infty}a_k\mathbf{A}^k$ 绝对收敛。

（2）如果 $\rho(A)>r$，设 $\boldsymbol{A}\boldsymbol{x}=\lambda_i\boldsymbol{x}$，其中 $|\lambda_i|=\rho(A)$，且 $\boldsymbol{x}$ 为单位长度特征向量。下面用反证法证明矩阵幂级数 $\sum_{k=0}^{\infty}a_k\mathbf{A}^k$ 发散。如果它是收敛的，则利用矩阵收敛的性质4知，数项级数

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

也收敛。但 $|\lambda_i|=\rho(\mathbf{A})>r$，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:

设 $\sum_{k=0}^{\infty} a_k (z-z_0)^k$ 为收敛半径为r的幂级数，A为n阶方阵，

1、如果A的特征值均落在收敛圆内，即$|\lambda - z_0| < r$，其中 λ 为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;

2、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。

幂级数的和函数与收敛圆内的解析函数的关系

1、幂级数的和函数是收敛圆内的解析函数

2、一个圆内解析的函数可以展开成收敛的幂级数

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 $A\in C^{n\times n}$ 的特征值落在收敛圆 $|z-z_0|<r$ 内时，定义

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

并称之为 A 关于解析函数 $f(z)$ 的矩阵函数。

解析函数:

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots \quad (|z| < +\infty)$$

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots \quad (|z| < +\infty)$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \cdots \quad (|z| < +\infty)$$

$$(1-z)^{-1} = 1 + z + z^2 + z^3 + \cdots (|z| < 1)$$

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \cdots \quad (|z| < 1)$$

矩阵函数:

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots \quad (\rho(\boldsymbol{A}) < +\infty)$$

$$\sin A = A - \frac{A^3}{3!} + \frac{A^5}{5!} - \cdots \quad (\rho(\boldsymbol{A}) < +\infty)$$

$$\cos A = I - \frac{A^2}{2!} + \frac{A^4}{4!} - \cdots \quad (\rho(\boldsymbol{A}) < +\infty)$$

$$(I-A)^{-1} = I + A + A^2 + A^3 + \cdots (\rho(\boldsymbol{A}) < 1)$$

$$\ln(I+A) = A - \frac{A^2}{2} + \frac{A^3}{3} - \cdots \quad (\rho(\boldsymbol{A}) < 1)$$

利用矩阵A的Jordan分解求矩阵函数 $f(A)$ 的具体表达式

$$A = TJT^{-1}$$

对于单项式函数:

$$f(A) = A^n = \left(TJT^{-1}\right)^n = TJ^nT^{-1} = Tf(J)T^{-1}$$

对于多项式函数:

$$\begin{aligned} f(A) &= \alpha_n A^n + \cdots + \alpha_1 A + \alpha_0 I \\ &= \alpha_n TJ^nT^{-1} + \cdots + \alpha_1 TJT^{-1} + \alpha_0 I \ = Tf(J)T^{-1} \end{aligned}$$

对于初等函数:

$$f(A) = Tf(J)T^{-1}$$

求矩阵函数$f(J)$的具体表达式

Jordan标准型J为块对角矩阵 $J=\operatorname{diag}\left(J_1,\ldots,J_k\right)$

对于单项式函数:

$$f(J)=J^n=\operatorname{diag}\left(f(J_1),\ldots,f(J_k)\right)=\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}$$

对于多项式函数:

$$f(J)=\alpha_n J^n+\cdots+\alpha_1 J+\alpha_0 I$$

$$=\operatorname{diag}\left(f(J_1),\ldots,f(J_k)\right) \qquad J^n=\operatorname{diag}\left(J_1^n,\ldots,J_k^n\right)$$

对于初等函数:

$$f(J)=\operatorname{diag}\left(f(J_1),\ldots,f(J_k)\right)=\begin{pmatrix} J_1^n & & \\ & \ddots & \\ & & J_k^n \end{pmatrix}$$

Jordan标准型J为块对角矩阵，有:

$$J=\operatorname{diag}(J_1,\ldots,J_k) \qquad f(J)=\operatorname{diag}\left(f(J_1),\ldots,f(J_k)\right)$$

利用矩阵的Jordan分解求矩阵函数 $f(A)$ 的具体表达式

1、$f(z)=\sum_{k=0}^{\infty}a_k(z-z_0)^k$ 为收敛半径为r的幂级数

2、A 为n阶方阵，$A=TJT^{1}$ 为其Jordan分解，$\boldsymbol{J}=\mathbf{diag}\left(\boldsymbol{J}_1,\ \boldsymbol{J}_2,\ \cdots,\ \boldsymbol{J}_s\right)$

3、 当A的特征值均落在收敛圆内

则矩阵幂级数 $\sum_{k=0}^{\infty}a_k(\boldsymbol{A}-z_0\boldsymbol{I})^k$ 绝对收敛， 并且和矩阵为

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

关键是如何计算

$$f(J_i)=\sum_{k=0}^{\infty}a_k\left(J_i-z_0I\right)^k \longleftarrow \sum_{k=0}^{\infty}a_kJ_i^k \longleftarrow J_i^k$$

因为 $\quad J=\lambda I+N$ 其中 $N=\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix},$

$$N^0=I$$

$$N^k=\begin{array}{cl} \begin{array}{cccccc} & & & k+1\text{列} & & \end{array} & \\ \begin{pmatrix} 0 & \cdots & 0 & 1 & & \boldsymbol{0} \\ & \ddots & & \ddots & \ddots & \\ & & \ddots & & \ddots & 1 \\ & & & \ddots & & 0 \\ & & & & \ddots & \vdots \\ & & & & & 0 \end{pmatrix} & \begin{array}{l} \\ \\ n-k\text{行} \\ \\ \\ \\ \end{array} \end{array} \qquad k\le n-1,$$

$$N^k=0, \qquad k\ge n$$

$$\boldsymbol{J} = \lambda \boldsymbol{I} + \boldsymbol{N} \quad \text{其中}\ \boldsymbol{N} = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix},\ (\lambda \boldsymbol{I})\boldsymbol{N} = \boldsymbol{N}(\lambda \boldsymbol{I})$$

$$J^k = (\lambda I + N)^k = \sum_{i=0}^{k} C_k^i \lambda^{k-i} N^i = \lambda^k I_n + C_k^1 \lambda^{k-1} N + \cdots + C_k^i \lambda^{k-i} N^i + \cdots$$

$$= \begin{pmatrix} \lambda^k & C_k^1 \lambda^{k-1} & \cdots & C_k^i \lambda^{k-i} & \cdots & C_k^{n-1} \lambda^{k-n+1} \\ & \lambda^k & C_k^1 \lambda^{k-1} & \ddots & & \vdots \\ & & & & & C_k^i \lambda^{k-i} \\ & & \ddots & \ddots & & \vdots \\ & & & \lambda^k & & C_k^1 \lambda^{k-1} \\ & & & & & \lambda^k \end{pmatrix}$$

进一步，有

$$
\mathbf{S}_{m+1}=\sum_{k=0}^{m}a_k\boldsymbol{J}^k=\begin{pmatrix}
\sum\limits_{k=0}^{m}a_k\lambda^k & \sum\limits_{k=0}^{m}a_kC_k^1\lambda^{k-1} & \cdots & \sum\limits_{k=0}^{m}a_kC_k^i\lambda^{k-i} & \sum\limits_{k=0}^{m}a_kC_k^{n-1}\lambda^{k-n+1} \\
 & \sum\limits_{k=0}^{m}a_k\lambda^k & \sum\limits_{k=0}^{m}a_kC_k^1\lambda^{k-1} & \cdots & \sum\limits_{k=0}^{m}a_k\,C_k^i\lambda^{k-i} \\
 & & \ddots & \ddots & \vdots \\
 & & & \sum\limits_{k=0}^{m}a_k\lambda^k & \sum\limits_{k=0}^{m}a_k\,C_k^1\lambda^{k-1} \\
 & & & & \sum\limits_{k=0}^{m}a_k\lambda^k
\end{pmatrix}
$$

注意到 $C_k^i=\dfrac{k(k-1)\cdots(k-i+1)}{i!}$， 记 $S_{m+1}(\lambda)=\sum\limits_{k=0}^{m}a_k\lambda^k$，则

$$\sum_{k=0}^{m}a_kC_k^i\lambda^{k-i}=\sum_{k=i}^{m}a_kC_k^i\lambda^{k-i}=\frac{1}{i!}\sum_{k=i}^{m}a_kk(k-1)\cdots(k-i+1)\lambda^{k-i}$$

$$=\frac{1}{i!}\left[\sum_{k=i}^{m}a_k\left(\lambda^k\right)^{(i)}\right]=\frac{1}{i!}\left(\sum_{k=i}^{m}a_k\lambda^k\right)^{(i)}=\frac{S_{m+1}^{(i)}(\lambda)}{i!}$$

其中 $S_{m+1}^{(i)}(\lambda)$ 表示 $S_{m+1}(\lambda)$ 对λ的 i阶导数。

$$
f(\boldsymbol{J})=\lim_{m\to\infty}\mathbf{S}_{m+1}=\begin{pmatrix}
S_{m+1}(\lambda) & S'_{m+1}(\lambda) & \cdots & \dfrac{S^{(i)}_{m+1}(\lambda)}{i!} & \cdots & \dfrac{S^{(n-1)}_{m+1}(\lambda)}{(n-1)!} \\
 & S_{m+1}(\lambda) & S'_{m+1}(\lambda) & & \ddots & \\
 & & \ddots & & & \dfrac{S^{(i)}_{m+1}(\lambda)}{i!} \\
 & & & \ddots & \ddots & \vdots \\
 & & & & \ddots & S'_{m+1}(\lambda) \\
 & & & & & S_{m+1}(\lambda)
\end{pmatrix}
$$

根据幂级数性质知，有 $\lim\limits_{m\to\infty}S^{(i)}_{m+1}(\lambda)=f^{(i)}(\lambda)$ ，因此有

引理1 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$\lambda$的$n$阶Jordan块阵，且 $|\lambda|<r$，则

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

推论2 设 $f(z)=\sum_{k=0}^{\infty}a_k(z-z_0)^k$ **是收敛半径为** r **的幂级数，** J **是特征值为** λ **的Jordan块，且** $|\lambda-z_0|<r$ **，则**

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

推论3 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 的n阶Jordan块阵，且 $|t\lambda|<r$，则

$$f(t\mathbf{J})=\begin{pmatrix} f(t\lambda) & t f'(t\lambda) & \cdots & \dfrac{t^{n-1}f^{(n-1)}(t\lambda)}{(n-1)!} \\ & f(t\lambda) & \ddots & \vdots \\ & & \ddots & t f'(t\lambda) \\ & & & f(t\lambda) \end{pmatrix}$$

利用矩阵的Jordan分解求矩阵函数 $f(A)$ 的具体表达式

定理3 设 $f(z)=\sum_{k=0}^{\infty}a_k(z-z_0)^k$ 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，$J=\mathbf{diag}(J_1, J_2, \cdots, J_s)$。当$A$的特征值均落在收敛圆内时，即 $|\lambda-z_0|<r$，其中 λ 为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，并且和矩阵为

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

其中$f(J_i)$的定义如表达式（1）。

计算矩阵函数值的基本计算步骤:

1、计算Jordan分解 $A = TJT^{-1} = T\mathrm{diag}\left(J_1, \cdots, J_s\right)T^{-1}$

2、求 $f(A) = T\mathrm{diag}\left(f(J_1), \cdots, f(J_s)\right)T^{-1}$

例4 设 $\mathbf{A}=\begin{pmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{pmatrix}$，求 $\sin A$

解 根据矩阵A的Jordan分解

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

因此，由

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

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

例5 设 $A=\begin{pmatrix}3 & 1 & -1\\ 1 & 2 & -1\\ 2 & 1 & 0\end{pmatrix}$，求 e^{At}

解 根据矩阵的Jordan 分解

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

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

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

我们还可以证明

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

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

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

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

（0） $e^{\boldsymbol{A}}e^{\boldsymbol{B}} = e^{\boldsymbol{B}}e^{\boldsymbol{A}} = e^{\boldsymbol{A}+\boldsymbol{B}}$

（1） $\sin(\mathbf{A}+\mathbf{B}) = \sin\mathbf{A}\cos\mathbf{B} + \cos\mathbf{A}\sin\mathbf{B}$

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

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

需要指出的是，对任何n阶方阵A，e^A总是可逆矩阵。

$$e^{A}e^{-A}=e^{-A}e^{A}=e^{O}=I$$

$\sin A$和$\cos A$不一定可逆. 例如 $\boldsymbol{A}=\begin{pmatrix}\pi & 0\\ 0 & \dfrac{\pi}{2}\end{pmatrix}$

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

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

三、函数矩阵的微积分

1、元素为函数的矩阵的微分和积分

2、数量函数对向量变量或矩阵变量的导数

3、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 $\mathbf{A}(t)=\left(a_{ij}(t)\right)_{m\times n}$ 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$\mathbf{A}(t)$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

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

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

（1） $$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A}(t)+\mathbf{B}(t))=\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{A}(t)+\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{B}(t)$$

（2） $$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A}(t)\mathbf{B}(t))=(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{A}(t))\cdot\mathbf{B}(t)+\mathbf{A}(t)\cdot(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{B}(t))$$

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

（4） 当$A^{-1}(t)$为可微矩阵时，有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{A}^{-1}(t))=-\mathbf{A}^{-1}(t)(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{A}(t))\mathbf{A}^{-1}(t)$$

（5） 当$u=f(t)$关于t可微时，有

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

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

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

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

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

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

（4）由于$A(t)^{-1}A(t)=I$，两端对t求导得

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

从而

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

由于$\dfrac{\mathrm{d}}{\mathrm{d}t}\big(\boldsymbol{A}(t)\big)$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：

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

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：

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

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

$$\neq$$

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

只有当 $\boldsymbol{A}(t)\frac{d}{dt}\left(\boldsymbol{A}(t)\right) = \frac{d}{dt}\left(\boldsymbol{A}(t)\right)\boldsymbol{A}(t)$ 时成立。

定理5 设n阶方阵A与t无关，则有

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

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

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

证， 只证(1)，(2, 3)的证明与(1)类似。

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

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

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

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

$$\int_{t_0}^{t_1}\mathbf{A}(t)dt=\left(\int_{t_0}^{t_1}a_{ij}(t)dt\right)_{m\times n}$$

积分性质：

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

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

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

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

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

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

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

2 相对于矩阵变量的微分

定义7 设 $\boldsymbol{X}=\left(x_{ij}\right)_{m\times n}$，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 $\dfrac{\partial f}{\partial x_{ij}}\left(i=1,2,\cdots,m;\quad j=1,2,\cdots,n\right)$都存在，定义 $f(\boldsymbol{X})$对矩阵$\boldsymbol{X}$的导数为

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

例6 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 $\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}^T}$，$\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}$，和 $\dfrac{\mathrm{d}^2 f}{\mathrm{d}\boldsymbol{x}^2}$。

解 根据定义有

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

特别地，以x为自变量的函数的导数，一般记为：

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

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

数量函数对向量变量的二阶导数称为函数 f 的Hessian矩阵，它显然是对称的。 记为

$$\boldsymbol{H}(\boldsymbol{x})=\nabla^2 f(\boldsymbol{x}) = \frac{\mathrm{d}^2 f}{\mathrm{d}\boldsymbol{x}^2} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial \xi_1^2} & \dfrac{\partial^2 f}{\partial \xi_1 \partial \xi_2} & \cdots & \dfrac{\partial^2 f}{\partial \xi_1 \partial \xi_n} \\ \dfrac{\partial^2 f}{\partial \xi_2 \partial \xi_1} & \dfrac{\partial^2 f}{\partial \xi_2^2} & \cdots & \dfrac{\partial^2 f}{\partial \xi_2 \partial \xi_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial \xi_n \partial \xi_1} & \dfrac{\partial^2 f}{\partial \xi_n \partial \xi_2} & \cdots & \dfrac{\partial^2 f}{\partial \xi_n^2} \end{pmatrix}$$

例6-1，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

$$f(\boldsymbol{x})=\sum_{i=1}^{n} a_i\xi_i \text{，} \frac{\partial f}{\partial \xi_j}=a_j \text{，} (j=1,2,\cdots,n)$$

所以

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

例6-2，设 $A=\left(a_{ij}\right)_{m\times n}$ 为常矩阵，$\boldsymbol{X}=\left(x_{ij}\right)_{n\times m}$ 为矩阵变量，且

$f(X)=\mathrm{tr}(AX)$ 求 $\dfrac{\partial f}{\partial \boldsymbol{X}}$。

分析：

$$\boldsymbol{AX}=\begin{pmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mm} \end{pmatrix}$$

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \qquad \boldsymbol{X}=\begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix}$$

解：由于 $\boldsymbol{AX}=\left(\sum_{k=1}^{n}a_{ik}x_{kj}\right)_{m\times m}$，

所以 $f(X)=\mathrm{tr}(AX)=\sum_{s=1}^{m}\sum_{k=1}^{n}a_{sk}x_{ks}$

而 $\left(\frac{\partial f}{\partial x_{ij}}\right)_{n\times m}=\left(a_{ji}\right)_{n\times m}\quad(i=1,2,\cdots,n\quad j=1,2,\cdots,m)$，

故 $\frac{\partial f}{\partial \boldsymbol{X}}=\left(\frac{\partial f}{\partial x_{ij}}\right)_{n\times m}=\left(a_{ji}\right)_{n\times m}=\boldsymbol{A}^{T}$

例7 设$\boldsymbol{x}=(\xi_1,\xi_2,\cdots\xi_n)^T$,$\mathbf{A}=\left(a_{ij}\right)_{n\times n}$,$n$元函数$f(\boldsymbol{x})=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$，求$\dfrac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}$。

解 因

$$f(\boldsymbol{x})=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\xi_i\,\xi_j=\xi_1\sum_{j=1}^{n}a_{1j}\xi_j+\cdots+\xi_k\sum_{j=1}^{n}a_{kj}\xi_j$$

$$+\cdots+\ \xi_n\sum_{j=1}^{n}a_{nj}\xi_j$$

所以

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

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

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

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

特别地，当A为对称矩阵时,

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}=2Ax$$

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

解：因为

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

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

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

$$=x^T\left(A^TA\right)x-2\left(A^Tb\right)^Tx+b^Tb$$

从而,由例6-1、例7，可得

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

矩阵函数在微分方程中的应用

一阶线性常系数齐次微分方程组的定解问题：

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

给定初始条件：$x_i(0)$，$(i = 1, 2, \cdots, n)$

记 $\mathbf{A} = (a_{ij}) \in \mathbf{C}^{n\times n}$,

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

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

则上述微分方程组可写成：

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

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

利用矩阵微分的性质有

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

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

方程（2）意味着

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

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

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

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

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

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

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

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = \boldsymbol{A}\boldsymbol{X}(t) + \boldsymbol{F}(t) & (4) \\ \boldsymbol{X}(t_0) = (x_1(t_0), x_2(t_0), \ldots, x_n(t_0))^T & (5) \end{cases}$$

这里 $\boldsymbol{F}(t) = (f_1(t), f_2(t), \cdots, f_n(t))^T$ 是已知向量函数，A 和 X 意义同前。改写方程为　并以 e^{-At} 左乘方程两边，即

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

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

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

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

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

例8 求定解问题 $\begin{cases} \dfrac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = \mathbf{A}\boldsymbol{X}(t) \\ \boldsymbol{X}(0) = (1,1,1)^T \end{cases}$ 其中 $\mathbf{A} = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & -1 & 2 \end{pmatrix}$

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

故A有三个不同的特征根，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：

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

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

$$\boldsymbol{X} = e^{\mathbf{A}t}\boldsymbol{X}(0) = \boldsymbol{T}\begin{pmatrix} 1 & & \\ & e^{2t} & \\ & & e^{3t} \end{pmatrix}\boldsymbol{T}^{-1}\boldsymbol{X}(0)$$

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

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

例9 求定解问题 $\begin{cases} \dfrac{\mathrm{d}\boldsymbol{X}(t)}{\mathrm{d}t} = \mathbf{A}\boldsymbol{X}(t) + \boldsymbol{F}(t) \\ \boldsymbol{X}(0) = (1,1,1)^T \end{cases}$ 的解, 其中矩阵A见例8。

$$\boldsymbol{F}(t) = (0, 0, e^{2t})^T$$

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

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