第五章

矩阵分析

5.1 矩阵级数

微积分的基础是数列极限的收敛理论及其衍生出来的级数理论。矩阵可看成一个“超数”，因此类比可得矩阵序列与矩阵级数，只要找到度量两个“超数”距离的适当工具。在矩阵里，这就是范数。

5.1.1矩阵序列的极限

一、矩阵序列的收敛性

定义5. 1. 1 设有$P^{m\times n}$中的矩阵序列 $\{A^{(k)}\}$,

$$A^{(k)}=\left[a_{ij}^{(k)}\right]\in C^{m\times n}.$$

若 $\lim\limits_{k\to\infty} a_{ij}^{(k)}=a_{ij}\,,\quad i=1,\cdots,m;\ j=1,\cdots,n$

则称矩阵序列$\{A^{(k)}\}$收敛于$A=(a_{ij})\in C^{m\times n}$

或称A为矩阵序列$\{A^{(k)}\}$的**极限**，记作

$$\lim_{k\to\infty} A^{(k)} = A \quad \text{或} \quad A^{(k)} \to A\ (k\to\infty)$$

若矩阵序列$\{A^{(k)}\}$不收敛，则称它是**发散的**。

例1

$$A^{(k)} = \begin{bmatrix} \dfrac{2k^2+k+1}{k^2} & \dfrac{\sin k}{k} \\ e^{-k}\sin k & \dfrac{3k^2+1}{k^2} \end{bmatrix} \qquad (k=1,2,\cdots)$$

$$A^{(k)} = \begin{bmatrix} \dfrac{2k^2+k+1}{k^2} & \dfrac{\sin k}{k} \\ e^{-k}\sin k & \dfrac{3k^2+1}{k^2} \end{bmatrix} \quad (k=1,2,\cdots)$$

则

$$\lim_{k\to\infty} A^{(k)} = \begin{bmatrix} \lim\limits_{k\to\infty} \dfrac{2k^2+k+1}{k^2} & \lim\limits_{k\to\infty} \dfrac{\sin k}{k} \\ \lim\limits_{k\to\infty} e^{-k}\sin k & \lim\limits_{k\to\infty} \dfrac{3k^2+1}{k^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

根据矩阵序列收敛性的定义，可证明下列性质。

定理 设 $\lim\limits_{k\to+\infty} A^{(k)} = A$, $\lim\limits_{k\to+\infty} B^{(k)} = B$,

$\forall\ a,b\in C$, 则有

（1） $\lim\limits_{k\to+\infty}(aA^{(k)}+bB^{(k)}) = aA+bB$

（2） $\lim\limits_{k\to+\infty} A^{(k)}B^{(k)} = AB$

（3） $\lim\limits_{k\to+\infty}(A^{(k)})^{-1} = A^{-1}$

$A^{-1},(A^{(k)})^{-1}$ 均存在， $i=1,2......$

例2 设

$$A^{(k)}=\begin{bmatrix} k^{-1} & \\ & e^{-k} \end{bmatrix},\quad k=1,2,\cdots$$

则

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

$$\left(A^{(k)}\right)^{-1}=\begin{bmatrix} k & \\ & e^{k} \end{bmatrix},\quad k=1,2,\cdots$$

但 A 不可逆。

用矩阵的范数理论来研究矩阵序列的收敛性是最常用、最简单的方法。

定理 5. 1. 1： 矩阵序列 $\{A^{(k)}\}\in F^{n\times n}$ 收敛于A的

充分必要条件是

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

其中，$\left\|A^{(k)}-A\right\|$为任意一种矩阵范数。

证明：取矩阵范数

$$\|A\|_1=\max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}^{(k)}|$$

必要性：设

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

那么由定义可知对每一对i, j都有

$$\lim_{k\to\infty}\left|a_{ij}^{(k)} - a_{ij}\right| = 0$$

$$(i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, n)$$

从而有

$$\lim_{k\to\infty}\max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}^{(k)}-a_{ij}|=0$$

上式记为

$$\lim_{k\to\infty}\left\|A^{(k)}-A\right\|_1=0$$

充分性：设

$$\lim_{k\to\infty}\left\|A^{(k)}-A\right\|_1=\lim_{k\to\infty}\max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}^{(k)}-a_{ij}|=0$$

那么对每一对i,j都有

$$\lim_{k\to\infty}\left|a_{ij}^{(k)}-a_{ij}\right|=0\ (i=1,\cdots,n;\ j=1,\cdots,n)$$

即 $\lim\limits_{k\to\infty} a_{ij}^{(k)} = a_{ij}\,(i=1,\cdots,n;\ j=1,\cdots,n)$

故有 $\lim\limits_{k\to\infty} A^{(k)} = A = [a_{ij}]$

现在已经证明了定理对于所设的范数成立，如果$\|A\|_{\alpha}$是另外任意一种范数，那么由范数的等价性可知

$$d_1\left\|A^{(k)}-A\right\|_1 \le \left\|A^{(k)}-A\right\|_\alpha \le d_2\left\|A^{(k)}-A\right\|_1$$

这样，当 $\lim\limits_{k\to\infty}\left\|A^{(k)}-A\right\|_1 = 0$时，

同样可得 $$\lim_{k\to\infty}\left\|A^{(k)}-A\right\|_\alpha = 0$$

因此定理对于任意一种范数都成立。

二、谱半径与矩阵范数

根据矩阵的诱导范数的含义，结合特征值，

设 (λ, x) 为 A 的任意特征对，则

$$|\lambda| \, \|x\|_{\alpha} = \|\lambda x\|_{\alpha}$$

$$= \|Ax\|_{\alpha} \leq \|A\|_{\beta} \|x\|_{\alpha}$$

从而 $|\lambda| \leq \|A\|_{\beta}$

这说明矩阵特征值的模都不超过它的范数。

定义5. 1. 2 设 $A\in P^{n\times n}$ 的特征值 $\lambda_1,\cdots,\lambda_n$ ，称

$$\rho(A)\equiv\max_i|\lambda_i|$$

为矩阵 A 的谱半径。

注：$\rho(A^k)=\left[\rho(A)\right]^k$

定理5.1.2 对 $A\in P^{n\times n}$ 的任意矩阵范数 $\|A\|$，恒有

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

定理5.1.2给出了矩阵谱半径的的一个上界，那么矩阵谱半径的下界呢？

定理5. 1. 3 对 $A \in P^{n\times n}$，存在 $P^{n\times n}$ 上矩阵范数 $\|\bullet\|_M$，对任意 $\varepsilon > 0$，恒有

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

三、(幂)收敛矩阵

最常见的矩阵序列是方阵的幂构成的矩阵序列。

定义 设 $A \in C^{n\times n}$， 若有 $\lim\limits_{k\to\infty} A^k = 0,$

则称 A为（幂）收敛矩阵。

联想到等比数列 $\{\ q^n\ \}$ 收敛当且仅当 $|q| < 1$

类似地，我们有

定理5.1.4 已知矩阵序列：$A, A^2, \cdots, A^k, \cdots,$

则 $$\lim_{k\to\infty} A^k = 0 \Leftrightarrow \rho(A) < 1$$

证明 必要性：已知A为收敛矩阵，则

$$[\rho(A)]^k = \rho(A^k) \le \|A^k\|$$

其中$\|\cdot\|$为$C^{n\times n}$上任一矩阵范数，即

$$\lim_{k\to+\infty}[\rho(A)]^k = 0$$

故 $$\rho(A) < 1.$$

充分性：由于$\rho(A)<1$，则存在正数ε 使得

$\rho(A)+\varepsilon<1$，由定理5.1.3,存在矩阵范数$\|\cdot\|$,使得

$$\|A\|\le\rho(A)+\varepsilon<1$$

从而 $\lim\limits_{k\to+\infty}\|A\|^k=0$

又$\|A^k\|\le\|A\|^k$，所以 $\lim\limits_{k\to+\infty}\|A^k\|=0.$

故

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

由于谱半径不易计算，联系到谱半径不超过任何一种矩阵范数，实际常用范数来判断矩阵是否是收敛矩阵。如果很难找到这样的范数，再计算出矩阵的所有特征值，进而得到谱半径。

推论 若对矩阵A的某一种范数$\|A\|<1$,

则 $$\lim_{k\to\infty} A^k = 0$$

例3 判断矩阵A是否为收敛矩阵，其中

$$A=\begin{pmatrix} 0.2 & 0.1 & 0.2 \\ 0.5 & 0.4 & 0.4 \\ 0.1 & 0.3 & 0.2 \end{pmatrix}$$

解： 由 $\|A\|_1 = \max\limits_{1\le j\le n} \sum\limits_{i=1}^{n} |a_{ij}| = 0.8 < 1$

可知A是收敛矩阵。

例4 设

$$A=\begin{pmatrix}0 & a & a\\ a & 0 & a\\ a & a & 0\end{pmatrix}, a\in R$$

试问a为何值时A为收敛矩阵。

解: 由题意, $|\lambda I-A|=(\lambda+a)^2(\lambda-2a)$

$$\rho(A)=|2a|<1\Rightarrow\ |a|<\frac{1}{2}$$

故, $|a|<\dfrac{1}{2}$时, A是收敛矩阵。

思考 证明

$$A=\begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/5 & 2/5 & 1/5 & 1/5 \\ 1/6 & 1/6 & 3/6 & 1/6 \\ 1/7 & 1/7 & 1/7 & 4/7 \end{pmatrix}$$

的谱半径 $\rho(A)=1.$

5.1.2、矩阵级数

1、矩阵级数及其收敛性

定义5. 1. 3 设矩阵序列 $\{A^{(k)}\}$, $A^{(k)} \in C^{m\times n}, k=0,1,\cdots$

定义**矩阵级数**为

$$\sum_{k=0}^{\infty} A^{(k)} = A^{(0)} + A^{(1)} + \cdots + A^{(k)} + \cdots$$

定义5. 1. 4 定义**矩阵级数的部分和**为

$$S_N = \sum_{k=0}^{N} A^{(k)} = A^{(0)} + A^{(1)} + \cdots + A^{(N)}$$

因此，可构成部分和序列 $S_0, S_1, \cdots, S_N, \cdots$

如果矩阵级数的部分和序列收敛于**A**，即

$$\lim_{N\to\infty} S_N = A$$

则称矩阵级数收敛于A，记作

$$\sum_{k=0}^{\infty} A^{(k)} = A.$$

不收敛的矩阵级数称为发散的。

2. 矩阵级数收敛的等价定义

定理5.1.5 矩阵级数收敛当且仅当相应的mn个数项级数是收敛的。

即设

$$A^{(k)}=(a_{ij}^{(k)}), A=(a_{ij})$$

则

$$\sum_{k=0}^{\infty} A^{(k)}=A \Leftrightarrow \sum_{k=0}^{\infty} a_{ij}^{(k)}=a_{ij}$$

例5 已知

$$A^{(k)}=\begin{pmatrix} \dfrac{1}{2^k} & \dfrac{\pi}{4^k} \\ 0 & \dfrac{1}{(k+1)(k+2)} \end{pmatrix} \quad (k=0,1,\cdots)$$

研究矩阵级数 $\sum_{k=0}^{\infty} A^{(k)}$ 的收敛性。

解

$$S_N = \sum_{k=0}^{N} A^{(k)}$$

$$= \begin{pmatrix} 2-\left(\frac{1}{2}\right)^N & \frac{\pi}{3}\left[4-\left(\frac{1}{4}\right)^N\right] \\ 0 & 1-\frac{1}{N+2} \end{pmatrix}$$

故有

$$S = \begin{pmatrix} 2 & \frac{4\pi}{3} \\ 0 & 1 \end{pmatrix}$$

3、收敛矩阵级数的性质：

（1）若矩阵级数$\sum_{k=0}^{\infty} A^{(k)}$收敛，则$\lim_{k\to\infty} A^{(k)} = 0$

（2）若矩阵级数$\sum_{k=0}^{\infty} A^{(k)}=S_1, \sum_{k=0}^{\infty} B^{(k)} = S_2$　则

$$\sum_{k=0}^{\infty}(aA^{(k)} + bB^{(k)}) = aS_1 + bS_2 \ ,\forall a,b \in C.$$

（3）设 $P \in C^{m\times m}, Q \in C^{n\times n}$，若矩阵级数

$\sum_{k=0}^{\infty} A^{(k)}$收敛，则 $\sum_{k=0}^{\infty} PA^{(k)}Q$ 收敛，且

$$\sum_{k=0}^{\infty} PA^{(k)}Q = P\left(\sum_{k=0}^{\infty} A^{(k)}\right)Q.$$

4. 绝对收敛

定义5.1.5 $P^{m\times n}$ 中的矩阵级数 $\sum_{k=0}^{\infty} A^{(k)}$ 称为绝对收敛的，如果数项级数

$$\sum_{k=0}^{\infty} a_{ij}^{(k)} \quad (i=1,2,\cdots,m;\quad j=1,2,\cdots,n)$$

都绝对收敛。这里 $A^{(k)}=(a_{ij}^{(k)})$.

利用高等数学中的相应结果，可得到

若矩阵级数绝对收敛，则该矩阵级数一定收敛。

同数项级数相吻合的是，判定矩阵级数是否绝对收敛可借助范数理论转化为判定正项级数的敛散性。

定理5.1.6　$P^{n\times n}$ 中的矩阵级数 $\sum_{k=0}^{\infty} A^{(k)}$ 绝对收敛的充要条件是正项级数 $\sum_{k=0}^{\infty} \| A^{(k)} \|$ 收敛，这里矩阵范数是任意的。

于是，判断一个矩阵级数是绝对收敛的问题归结为判定一个正项级数是否收敛的问题。从而使问题大大简化。

5.1.3矩阵幂级数

定义5. 1. 6 $P^{n\times n}$ 中的矩阵级数

$$\sum_{k=0}^{\infty} c_k A^k = c_0 I + c_1 A + c_2 A^2 + \cdots + c_k A^k + \cdots$$

称为矩阵 A 的幂级数。这里 $c_k \in P$

由前可知矩阵的幂级数是实变量的幂级数 $\sum_{k=0}^{\infty} a_k x^k$ 以及复变量的幂级数 $\sum_{k=0}^{\infty} c_k z^k$ 的推广，因此讨论矩阵幂级数的收敛性问题自然就与复变量的幂级数的收敛半径联系起来。

定理5. 1. 7 设幂级数 $\sum_{k=0}^{\infty} c_k z^k$ 的收敛半径为 R，则

(1) 当 $\rho(A) < R$ 时幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 绝对收敛；

(2) 当 $\rho(A) > R$ 时幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 发散。

注：$\rho(A) = R$，此定理失效，需用其他方法。

推论5.1.1 若幂级数$\sum_{k=0}^{\infty} c_k z^k$在整个平面上都收敛，则对任意的$A \in C^{n\times n}$，矩阵幂级数$\sum_{k=0}^{\infty} C_k A^k$收敛。

推论 设幂级数$\sum_{k=1}^{\infty} c_k z^k$的收敛半径为$R$，$A \in C^{n\times n}$，若存在$C^{n\times n}$上的某一矩阵范数$\|\cdot\|$使得$\|A\| < R$，则矩阵幂级数$\sum_{k=0}^{\infty} c_k A^k$绝对收敛。

例 6（1）求下面级数的收敛半径

$$\sum_{k=1}^{\infty}\frac{x^k}{2^k\cdot k}=\frac{x}{2\cdot 1}+\frac{x^2}{2^2\cdot 2}+\frac{x^3}{2^3\cdot 3}+\cdots+\frac{x^k}{2^k\cdot k}+\cdots$$

（2）设

$$A=\begin{bmatrix}1 & 4\\ -1 & -3\end{bmatrix}$$

判断幂级数$\sum_{k=1}^{\infty}\frac{A^k}{2^k\cdot k}$的敛散性。

解　设此级数的半径为R, 利用公式

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = \frac{1}{R}$$

容易求得此级数的收敛半径为 2。而 $\rho(A)=1.$

所以由上面的定理可知矩阵幂级数

$$\sum_{k=1}^{\infty}\frac{A^k}{2^k\cdot k} = \frac{A}{2\cdot 1}+\frac{A^2}{2^2\cdot 2}+\frac{A^3}{2^3\cdot 3}+\cdots+\frac{A^k}{2^k\cdot k}+\cdots$$

绝对收敛。

例7 试判断矩阵幂级数

$$\sum_{k=1}^{\infty}\frac{k}{6^k}\begin{pmatrix}1 & -8\\ -2 & 1\end{pmatrix}^k$$

的敛散性。

解 令 $A=\frac{1}{6}\begin{pmatrix}1 & -8\\ -2 & 1\end{pmatrix}$，可求得$A$的特征值为

$\lambda_1=\frac{5}{6},\ \ \lambda_2=-\frac{1}{2}$，故$\rho(A)<1$ 由于题设中幂级数

收敛半径为1。所以题设中幂级数绝对收敛。

最后讨论特殊的诺伊曼(Neumann)级数, 即

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

幂级数 $\sum_{k=0}^{\infty} z^k$ 的收敛半径是 **1**，并且收敛于

$$\sum_{k=0}^{\infty} z^k = \lim_{k\to\infty} \frac{1-z^k}{1-z} = \frac{1}{1-z} = (1-z)^{-1}$$

所以我们通过类比可以得到

推论6.1.6 矩阵幂级数（**Neumann**级数）

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

绝对收敛的充分必要条件是 $\rho(A) < 1$

且其和为 $(I - A)^{-1}$

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

证：

设 $\sum_{k=0}^{\infty} A^k$ 收敛，则有 $\lim_{k=0} A^k = 0$，即A为收敛矩阵，

从而有 $\rho(A) < 1$ 。

设 $\rho(A) < 1$,又幂级数 $\sum_{k=0}^{\infty} z^k$ 的收敛半径是**1**，

因此，**Neumann**级数绝对收敛。

当**Neumann**级数收敛时，$\rho(A) < 1$，故 $(I - A)$ 可逆。

又

$$S^{(N)}(I-A)$$
$$=\left(I+A+\cdots+A^{N}\right)(I-A)=I-A^{N+1}$$

故

$$S^{(N)}=\left(I-A^{N+1}\right)(I-A)^{-1}$$
$$=(I-A)^{-1}-A^{N+1}(I-A)^{-1}$$

从而

$$S=\lim_{N\to\infty}S^{(N)}=(I-A)^{-1}$$

例8 设 $A=\begin{pmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{pmatrix}$

试判断矩阵幂级数$\sum_{k=0}^{\infty} A^k$的敛散性，若收敛求其和。

解 由 $\|A\|_\infty = 0.9 < 1$ 知 $\rho(A) < 1$

$\therefore$ 幂级数$\sum_{k=0}^{\infty} A^k$的收敛，且

$$\sum_{k=1}^{\infty} A^k = (I-A)^{-1} = \frac{1}{0.15}\begin{pmatrix} 0.4 & 0.7 \\ 0.3 & 0.9 \end{pmatrix}$$

5.2 矩阵函数

矩阵函数在力学、控制理论及信号处理等学科中具有重要应用。类比普通函数，矩阵函数的特殊之处在于其自变量与因变量都是方阵。对应于矩阵函数的多种表示方式(幂级数、Jordan表示、多项式表示、积分表示等)，定义矩阵函数的方式也很多。

5.2.1 矩阵函数的定义

定义5. 2. 1 设一元函数 $f(z)$可展开为收敛半径为 R 的幂级数，即

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

矩阵 A 的谱半径 $\rho(A)<R$，则矩阵函数

$f(A)$ 即为相应的矩阵幂级数(收敛时)的和，即

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

在高等数学和复变函数中，有幂级数展开式：

$$e^z=\sum_{k=0}^{\infty}\frac{z^k}{k!},\qquad (R=+\infty)$$

$$\sin z=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}z^{2k+1},\qquad (R=+\infty)$$

$$\cos z=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}z^{2k},\qquad (R=+\infty)$$

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

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

相应地，我们有矩阵函数：

$$e^{A}=\sum_{k=0}^{\infty}\frac{A^{k}}{k!},\quad (\forall A\in P^{n\times n})$$

$$\sin A=\sum_{k=0}^{\infty}\frac{(-1)^{k}}{(2k+1)!}A^{2k+1},\quad (\forall A\in P^{n\times n})$$

$$\cos A=\sum_{k=0}^{\infty}\frac{(-1)^{k}}{(2k)!}A^{2k},\quad (\forall A\in P^{n\times n})$$

$$(E-A)^{-1}=\sum_{k=0}^{\infty}A^{k},\quad (\rho(A)<1)$$

$$\ln(E+A)=\sum_{k=0}^{\infty}\frac{(-1)^{k}}{k+1}A^{k+1},\quad (\rho(A)<1)$$

以及含参数的矩阵函数:

$$e^{At} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k, \quad (\forall A \in P^{n\times n})$$

$$\sin(At) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} t^{2k+1} A^{2k+1}, \quad (\forall A \in P^{n\times n})$$

$$\cos(At) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} t^{2k} A^{2k}, \quad (\forall A \in P^{n\times n})$$

5.2.2 矩阵函数的计算

由矩阵函数的定义，矩阵函数的计算转化为矩阵幂级数和的计算，主要就是矩阵幂的计算。

首先联想到线性代数中矩阵的对角化问题，因此我们想到利用**特征值分解**来计算矩阵函数。对角矩阵的对角元就是矩阵的特征值，而相似矩阵就是相应的特征向量构成的矩阵。按这个思路，对任意矩阵则使用**Jordan分解**。由于计算比较复杂，最后我们给出**待定系数法**。

1. 利用相似对角化计算

设A可对角化，即存在可逆矩阵P 使得

$$P^{-1}AP = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

$$\Rightarrow A = Pdiag(\lambda_1, \lambda_2, \cdots, \lambda_n)P^{-1}$$

$$f(A) = \sum_{k=0}^{\infty} c_k A^k = Pdiag(\sum_{k=0}^{\infty} c_k \lambda_1{}^k, \cdots, \sum_{k=0}^{\infty} c_k \lambda_n{}^k)P^{-1}$$

$$= P\begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} P^{-1}$$

同理可得

$$f(At) = P\begin{pmatrix} f(\lambda_1 t) & & \\ & \ddots & \\ & & f(\lambda_n t) \end{pmatrix} P^{-1}$$

例1 设$A=\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}$，求$e^{A}$,$\sin A$.

解 $|\lambda I-A|=(\lambda-1)(\lambda-2)$

$$\Rightarrow \lambda=1,\lambda=2$$

因此，A可对角化，且与

$$\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$$

相似。

解$(I-A)x=0$可得对应于λ_1的特征向量

$$\alpha_1=\begin{pmatrix}1 & 0\end{pmatrix}^T$$

解$(2I-A)x=0$可得对应于λ_2的特征向量

$$\alpha_2=\begin{pmatrix}1 & 1\end{pmatrix}^T$$

令$P=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$,

则$P^{-1}AP=\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$,

即$A=P\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}P^{-1}$.

当$f(z)=e^z$时，

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

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

$$=\begin{pmatrix} e & e^2-e \\ 0 & e^2 \end{pmatrix}$$

当$f(z)=\sin z$时，

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

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

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

例2 设

$$A=\begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{pmatrix}$$

求 $e^{A}, e^{tA}(t\in R), \cos A$

解 该矩阵的特征多项式为

$$\varphi(\lambda)=\begin{vmatrix} \lambda-4 & -6 & 0 \\ 3 & \lambda+5 & 0 \\ 3 & 6 & \lambda-1 \end{vmatrix}=(\lambda-1)^{2}(\lambda+2)$$

求得特征值 $\lambda_1=-2$，$\lambda_2=\lambda_3=1$，对应的特征向量

矩阵论 MATRIX THEORY

$$X_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \quad X_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \quad X_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

构造矩阵

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

求得

$$P^{-1} = \begin{pmatrix} 1 & 2 & 0 \\ -1 & -1 & 0 \\ -1 & -2 & 1 \end{pmatrix}$$

则有

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

因此有

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

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

$$\cos A=P\begin{pmatrix} \cos(-2) & 0 & 0 \\ 0 & \cos 1 & 0 \\ 0 & 0 & \cos 1 \end{pmatrix}P^{-1}$$

2. 待定系数法

设$A\in C^{n\times n}$的特征多项式或零化多项式为

$$\varphi(\lambda)=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_s)^{r_s},\ r_1+r_2+\cdots+r_s=m\le n.$$

将$f(\lambda)$表示为 $f(\lambda)=q(\lambda)\varphi(\lambda)+r(\lambda)$

其中，$r(\lambda)$为次数低于n 的多项式，记作

$$r(\lambda)=b_0+b_1\lambda+\cdots+b_{n-1}\lambda^{n-1},$$

由Cayley-Hamilton定理知$f(A)=r(A)$,

利用$r^{(p)}(\lambda_i)=f^{(p)}(\lambda_i)\ (p=0,1,\cdots,r_i-1;i=1,2,\cdots s)$

求出待定系数$b_0,\ b_1,\cdots,b_{n-1}$.

矩阵论 MATRIX THEORY

若$f(z)=\sum_{k=0}^{\infty} a_k z^k$，求$f(A)$及$f(At)$

(1) 求矩阵A的特征多项式（或最小多项式）

$$\varphi(\lambda)=|\lambda I-A|=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_s)^{r_s}$$

其中，λ_i互异，$r_1+r_2+\cdots+r_s=n.$

(2) 设$r(\lambda)$为一$n-1$次待定系数多项式

$$r(\lambda)=b_0+b_1\lambda+\cdots+b_{n-1}\lambda^{n-1}.$$

(3)　利用 $r^{(p)}(\lambda_i) = f^{(p)}(\lambda_i)$

$$(p = 0, 1, \cdots, r_i - 1;\ i = 1, 2, \cdots s)$$

或 $r^{(p)}(\lambda_i) = f^{(p)}(\lambda_i t)$

$$(p = 0, 1, \cdots, r_i - 1;\ i = 1, 2, \cdots s)$$

对方程组求解得 $b_0, b_1, \cdots b_{n-1}$.

(4) $f(A) = r(A) = b_0 I + b_1 A + \cdots b_{n-1} A^{n-1}$.

例3 $A=\begin{pmatrix}-1 & 0 & 1\\ 1 & 2 & 0\\ -4 & 0 & 3\end{pmatrix}$. 求$e^{At}$, $\cos A$.

解 （1）A的特征多项式 $\varphi(\lambda)=(\lambda-1)^2(\lambda-2)$.

（2）设 $r(\lambda)=b_0+b_1\lambda+b_2\lambda^2$.

对于$f(\lambda)=e^{\lambda t}$, $f'(\lambda)=te^{\lambda t}$

由$r(1)=f(1), r'(1)=f'(1), r(2)=f(2)$得

$$\begin{cases} b_0 + b_1 + b_2 = e^t, \\ b_1 + 2b_2 = te^t, \\ b_0 + 2b_1 + 4b_2 = e^{2t}. \end{cases} \Rightarrow \begin{cases} b_0 = e^{2t} - 2te^t, \\ b_1 = -2e^{2t} + 2e^t + 3te^t, \\ b_2 = e^{2t} - e^t - te^t. \end{cases}$$

从而 $e^{At} = b_0 I + b_1 A + b_2 A^2$

$$= \begin{pmatrix} e^t - 2te^t & 0 & te^t \\ -e^{2t} + e^t + 2te^t & e^{2t} & e^{2t} - e^t - te^t \\ -4te^t & 0 & e^t + 2te^t \end{pmatrix}.$$

对于$f(\lambda)=\cos\lambda,\quad f'(\lambda)=-\sin\lambda$

由$r(1)=f(1), r'(1)=f'(1), r(2)=f(2)$得

$$\begin{cases} b_0+b_1+b_2=\cos 1, \\ b_1+2b_2=-\sin 1, \\ b_0+2b_1+4b_2=\cos 2 \end{cases} \Rightarrow \begin{cases} b_0=2\sin 1+\cos 2, \\ b_1=-3\sin 1+2\cos 1-2\cos 2, \\ b_2=\sin 1-\cos 1+\cos 2. \end{cases}$$

从而 $\cos A = b_0 I + b_1 A + b_2 A^2$

$$= \begin{pmatrix} 2\sin 1 + \cos 2 & 0 & -\sin 1 \\ -2\sin 1 + \cos 1 - \cos 2 & \cos 2 & \sin 1 - \cos 1 + \cos 2 \\ 4\sin 1 & 0 & -2\sin 1 + \cos 1 \end{pmatrix}.$$

5.2.3 常用矩阵函数的性质

定理5.2.1 设$A \in C^{n\times n}$，则

(1) $\sin(-A) = -\sin A$, $\cos(-A) = \cos A$;

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

$$\sin A = \frac{1}{2i}\left(e^{iA} - e^{-iA}\right).$$

定理5. 2. 2 设$A, B \in C^{n\times n}$，且$AB = BA$，则

(1) $e^{A+B} = e^{A}e^{B} = e^{B}e^{A}$;

(2) $\sin(A+B) = \sin A\cos B + \cos A\sin B$;

$\cos(A+B) = \cos A\cos B - \sin A\sin B$.

(3) $\cos 2A = \cos^2 A - \sin^2 A$,

$\sin 2A = 2\sin A\cos B$.

定理 $\det e^{A} = e^{trA}$, $\left(e^{A}\right)^{-1} = e^{-A}$.

注 对任一方阵A, e^{A} 总可逆，但$\sin A, \cos A$未必.

5.3 矩阵的微分与积分

实际使用时，矩阵函数与函数矩阵的微分、积分常常同时出现。研究矩阵函数和函数矩阵的微分、积分，这对研究微分方程组以及优化问题等非常重要。

5.3.1 函数矩阵的微积分

定义5. 3. 1 设有函数矩阵 $A(t)=(a_{ij}(t))\in C^{m\times n}$.
称矩阵 $A(t)$ 可微，如果其每个元素 $a_{ij}(t)$ 都是可微函数，且导数为

$$A'(t)=\frac{d}{dt}A(t)\equiv(a'_{ij}(t))$$

称矩阵 $A(t)$ 在 $[a,b]$ 上可积，如果其每个元素 $a_{ij}(t)$ 都在 $[a,b]$ 上可积，且积分为

$$\int_a^b A(t)\,dt \equiv \left(\int_a^b a_{ij}(t)\,dt\right)$$

容易验证矩阵微分具有下列性质：

定理 5.3.1 设 $A(t)$ 和 $B(t)$ 都是可微矩阵，则

(1) $\frac{d}{dt}(A(t)\pm B(t))=\frac{d}{dt}A(t)\pm\frac{d}{dt}B(t)$

(2) $\frac{d}{dt}(k(t)\cdot A(t))=\frac{d}{dt}k(t)\cdot A(t)+k(t)\cdot\frac{d}{dt}A(t)$

这里 $k(t)$ 为可微函数。

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

当 $u=f(t)$ 关于 t 可微时，有

(4) $$\frac{d}{dt}A(u)=f'(t)\frac{d}{du}A(u)$$

(5) $$\frac{d}{dt}(A^{-1}(t))=-A^{-1}(t)\cdot\frac{d}{dt}A(t)\cdot A^{-1}(t)$$

这里 $A^{-1}(t)$ 为可微矩阵。

矩阵论 MATRIX THEORY

类似地，设有函数矩阵 $A(t)=(a_{ij}(t))\in\mathbb{C}^{m\times n}$。称矩阵 $A(t)$ 二阶可微，如果其每个元素 $a_{ij}(t)$ 都是二阶可微函数，且二阶导数为

$$A''(t)=\frac{d}{dt}A'(t)\equiv(a''_{ij}(t))$$

一般地，不难给出矩阵的高阶导数。

容易验证矩阵积分具有下列性质：

定理 5.3.2 设 $A(t)$ 和 $B(t)$ 都在 $[a,b]$ 上可积，则

(1) $\int_a^b (A(t) \pm B(t))\,dt = \int_a^b A(t)\,dt \pm \int_a^b B(t)\,dt$

(2) $\int_a^b \lambda \cdot A(t)\,dt = \lambda \cdot \int_a^b A(t)\,dt, \quad \lambda \in C$

(3) $\int_a^b P \cdot A(t) \cdot Q\,dt = P \cdot \int_a^b A(t)\,dt \cdot Q$

这里 P、Q 为常量矩阵。

(4)　设 $A(t)$ 在 $[a,b]$ 上连续，则成立微积分基本定理：

$$\frac{d}{dt}\int_a^t A(s)\,ds = A(t)$$

(5)　设 $A'(t)$ 在 $[a,b]$ 上连续，则成立牛顿-莱布尼兹公式：

$$\int_a^b A'(t)\,dt = A(b) - A(a)$$

例 1 已知函数矩阵

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

试计算

(1) $\dfrac{d}{dx}A(x), \dfrac{d^2}{dx^2}A(x), \dfrac{d^3}{dx^3}A(x)$

(2) $\dfrac{d}{dx}\left|A(x)\right|$

(3) $\dfrac{d}{dx}A^{-1}(x)$

解

$$\frac{d}{dx}A(x)=\begin{bmatrix}0 & 2x\\ 1 & 0\end{bmatrix}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}A(x)=\begin{bmatrix}0 & 2\\ 0 & 0\end{bmatrix}$$

$$\frac{d^3}{dx^3}A(x)=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

由于 $|A(x)| = -x^3$，所以

$$\frac{d}{dx}|A(x)| = -3x^2$$

下面求 $A^{-1}(x)$。由伴随矩阵公式可得

$$A^{-1}(x) = \frac{1}{|A(x)|}A^{*}(x)$$

$$= -\frac{1}{x^3}\begin{bmatrix} 0 & -x^2 \\ -x & 1 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{x} \\ \frac{1}{x^2} & -\frac{1}{x^3} \end{bmatrix}$$

再求 $\frac{d}{dx}A^{-1}(x)$

$$\frac{d}{d}A^{-1}(x) = \begin{bmatrix} 0 & -\frac{1}{x^2} \\ -\frac{2}{x^3} & \frac{3}{x^4} \end{bmatrix}$$

例 2 已知函数矩阵

$$A(x)=\begin{bmatrix} \sin x & \cos x & x \\ \dfrac{\sin x}{x} & e^{x} & x^{2} \\ 1 & 0 & x^{3} \end{bmatrix}$$

试求

(1) $\lim\limits_{x\to 0} A(x)$ (2) $\dfrac{\mathrm{d}}{\mathrm{d}x}A(x)$ (3) $\dfrac{\mathrm{d}^2}{\mathrm{d}x^2}A(x)$

(4) $\dfrac{\mathrm{d}}{\mathrm{d}x}|A(x)|$ (5) $\left|\dfrac{\mathrm{d}}{\mathrm{d}x}A(x)\right|$

例 3 已知函数矩阵

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

试求 $\int_0^x A(x)dx$, $\left(\int_0^{x^2} A(x)dx\right)'$

解

$$\int_0^x A(x)dx = \begin{bmatrix} \int_0^x \sin x dx & -\int_0^x \cos x dx \\ \int_0^x \cos x dx & \int_0^x \sin x dx \end{bmatrix}$$

$$= \begin{bmatrix} 1-\cos x & -\sin x \\ \sin x & 1-\cos x \end{bmatrix}$$

同样可以求得

$$\left(\int_0^{x^2} A(x)dx\right)' = 2x \begin{bmatrix} \sin x^2 & -\cos x^2 \\ \cos x^2 & \sin x^2 \end{bmatrix}$$

关于矩阵的指数函数与三角函数，具有以下几个常用的性质：

$$(1)\ \frac{d}{dt}e^{At}=Ae^{At}=e^{At}A;$$

$$(2)\ \frac{d}{dt}(\sin At)=A(\cos At)=(\cos At)A;$$

$$(3)\ \frac{d}{dt}(\cos At)=-A(\sin At)=-(\sin At)A.$$

5.3.2 矩阵数量值函数对矩阵的导数

定义5. 3. 2 设有矩阵数量值函数 $f(X), X \in C^{m\times n}$.

函数 $f(X)$ 对 X 的导数为**梯度矩阵**

$$\frac{df}{dX} \equiv \left(\frac{\partial f}{\partial x_{ij}}\right) = \begin{pmatrix} \frac{\partial f}{\partial x_{11}} & \cdots & \frac{\partial f}{\partial x_{1n}} \\ \cdots & & \cdots \\ \frac{\partial f}{\partial x_{m1}} & \cdots & \frac{\partial f}{\partial x_{mn}} \end{pmatrix}$$

以向量为自变量的函数的导数——梯度向量

$$f : C^n \to F$$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \frac{df}{dx} = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix} = \mathbf{grad}\, f$$

例4

$$\boldsymbol{a}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad \boldsymbol{x}=\begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} \quad f(\boldsymbol{x})=\boldsymbol{a}^T\boldsymbol{x}=\boldsymbol{x}^T\boldsymbol{a} \quad \frac{df}{dx}=?$$

$$f(\boldsymbol{x})=a_1\xi_1+a_2\xi_2+\cdots+a_n\xi_n \quad \frac{\partial f}{\partial \xi_i}=a_i \quad (i=1,\cdots,n)$$

$$\frac{df}{dx}=\begin{pmatrix} \dfrac{\partial f}{\partial \xi_1} \\ \vdots \\ \dfrac{\partial f}{\partial \xi_n} \end{pmatrix}=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}=\boldsymbol{a}$$

例5

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \qquad X=\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix}$$

$$f(X)=\mathrm{tr}(AX) \qquad \frac{df}{dX}=?$$

$$f(X)=\mathrm{tr}(AX)=a_{11}x_{11}+a_{12}x_{21}+a_{13}x_{31}+a_{21}x_{12}+a_{22}x_{22}+a_{23}x_{32}$$

$$\frac{\partial f}{\partial x_{ij}}=a_{ji} \qquad (i=1,2,3; \quad j=1,2)$$

$$\frac{df}{dX}=\left(\frac{\partial f}{\partial x_{ij}}\right)_{3\times 2}=(a_{ji})_{3\times 2}=A^T$$

例6 已知

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, x=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, f(x)=x^{\mathrm{T}}Ax, 求 \frac{\mathrm{d}f(x)}{\mathrm{d}x}.$$

解:

$$f(x)=a_{11}x_1^2+a_{12}x_1x_2+a_{13}x_1x_3+a_{21}x_1x_2+a_{22}x_2^2$$

$$+a_{23}x_2x_3+a_{31}x_1x_3+a_{32}x_2x_3+a_{33}x_3^2$$

所以

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x}=\begin{bmatrix}\dfrac{\partial f}{\partial x_1}\\ \dfrac{\partial f}{\partial x_2}\\ \dfrac{\partial f}{\partial x_3}\end{bmatrix}=A^{\mathrm{T}}x+Ax=(A^{\mathrm{T}}+A)x.$$

特别地，当A为对称矩阵时，

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x}=2Ax.$$

5.3.3 矩阵值函数对矩阵的导数

定义5. 3. 3设矩阵值函数 $F(X)=(f_{ij}(X))_{s\times t}, X\in C^{m\times n}$ 的元素 $f_{ij}(X)$ 都是矩阵数量值函数。矩阵值函数 $F(X)$ 对 X 的导数指的是 $ms\times nt$ 矩阵

$$\frac{dF}{dX}=\begin{pmatrix} \frac{\partial F}{\partial x_{11}} & \cdots & \frac{\partial F}{\partial x_{1n}} \\ \cdots & & \cdots \\ \frac{\partial F}{\partial x_{m1}} & \cdots & \frac{\partial F}{\partial x_{mn}} \end{pmatrix}$$

其中

$$\frac{\partial F}{\partial x_{ij}}=\begin{pmatrix}\frac{\partial f_{11}}{\partial x_{ij}} & \cdots & \frac{\partial f_{1t}}{\partial x_{ij}}\\ \cdots & & \cdots\\ \frac{\partial f_{s1}}{\partial x_{ij}} & \cdots & \frac{\partial f_{st}}{\partial x_{ij}}\end{pmatrix}$$

例7

$$x=\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}\qquad \frac{dx^T}{dx}=?\qquad \frac{dx}{dx^T}=?$$

解 $F(x)=x^T=\begin{pmatrix}\xi_1 & \xi_2 & \cdots & \xi_n\end{pmatrix}$

$$=\begin{pmatrix}e_1^T x & e_2^T x & \cdots & e_n^T x\end{pmatrix}$$

$$F(x)=(f_1(x) \quad f_2(x) \quad \cdots \quad f_n(x))$$

$$\frac{dx^T}{dx}=\begin{pmatrix}\dfrac{\partial x^T}{\partial \xi_1}\\ \dfrac{\partial x^T}{\partial \xi_2}\\ \vdots \\ \dfrac{\partial x^T}{\partial \xi_n}\end{pmatrix}$$

$$\frac{\partial x^T}{\partial \xi_1}=\begin{pmatrix}\dfrac{\partial f_1(x)}{\partial \xi_1} & \dfrac{\partial f_2(x)}{\partial \xi_1} & \cdots & \dfrac{\partial f_n(x)}{\partial \xi_1}\end{pmatrix}$$

$$=\begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix}$$

$$\frac{\partial x^T}{\partial \xi_n}=\begin{pmatrix}\dfrac{\partial f_1(x)}{\partial \xi_n} & \dfrac{\partial f_2(x)}{\partial \xi_n} & \cdots & \dfrac{\partial f_n(x)}{\partial \xi_n}\end{pmatrix}$$

$$=\begin{pmatrix}0 & 0 & \cdots & 1\end{pmatrix}$$

故

$$\frac{dx^T}{dx} = I_n \qquad \frac{dx}{dx^T} = I_n$$

例8

$$\boldsymbol{a}=\begin{pmatrix}a_1\\a_2\\a_3\\a_4\end{pmatrix}\quad X=\begin{pmatrix}x_{11}&x_{12}&x_{13}&x_{14}\\x_{21}&x_{22}&x_{23}&x_{24}\end{pmatrix}\quad \frac{d(Xa)^T}{dX}=?$$

解

$$F(X)=(Xa)^T$$

$$\frac{d(Xa)^T}{dX}=\begin{pmatrix}\dfrac{\partial(Xa)^T}{\partial x_{11}}&\dfrac{\partial(Xa)^T}{\partial x_{12}}&\dfrac{\partial(Xa)^T}{\partial x_{13}}&\dfrac{\partial(Xa)^T}{\partial x_{14}}\\[2ex]\dfrac{\partial(Xa)^T}{\partial x_{21}}&\dfrac{\partial(Xa)^T}{\partial x_{22}}&\dfrac{\partial(Xa)^T}{\partial x_{23}}&\dfrac{\partial(Xa)^T}{\partial x_{24}}\end{pmatrix}$$

矩阵论 MATRIX THEORY

$$\boldsymbol{a}=\begin{pmatrix}\boldsymbol{a}_1\\ \boldsymbol{a}_2\\ \vdots\\ \boldsymbol{a}_n\end{pmatrix}\qquad \boldsymbol{X}=\begin{pmatrix}\boldsymbol{x}_{11} & \boldsymbol{x}_{12} & \boldsymbol{x}_{13} & \boldsymbol{x}_{14}\\ \boldsymbol{x}_{21} & \boldsymbol{x}_{22} & \boldsymbol{x}_{23} & \boldsymbol{x}_{24}\end{pmatrix}$$

$$\begin{matrix} e_1^T X\boldsymbol{a} & e_2^T X\boldsymbol{a} \\ \updownarrow & \updownarrow \\ \big(f_{11}(X) & f_{12}(X)\big) \\ \updownarrow & \updownarrow \end{matrix}$$

$$F(X)=(Xa)^T=\begin{pmatrix}\sum_{k=1}^{4}x_{1k}a_k & \sum_{k=1}^{4}x_{2k}a_k\end{pmatrix}$$

$$\frac{\partial (Xa)^T}{\partial x_{11}}=\begin{pmatrix}\frac{\partial f_{11}}{\partial x_{11}} & \frac{\partial f_{12}}{\partial x_{11}}\end{pmatrix}$$

$$=\begin{pmatrix}a_1 & 0\end{pmatrix}$$

$$\begin{matrix} & (a_2 \quad 0) & (a_3 \quad 0) & (a_4 \quad 0) \end{matrix}$$

$$\frac{d(Xa)^T}{dX}=\begin{pmatrix}\frac{\partial (Xa)^T}{\partial x_{11}} & \frac{\partial (Xa)^T}{\partial x_{12}} & \frac{\partial (Xa)^T}{\partial x_{13}} & \frac{\partial (Xa)^T}{\partial x_{14}}\\ \frac{\partial (Xa)^T}{\partial x_{21}} & \frac{\partial (Xa)^T}{\partial x_{22}} & \frac{\partial (Xa)^T}{\partial x_{23}} & \frac{\partial (Xa)^T}{\partial x_{24}}\end{pmatrix}$$

$$\begin{matrix}(0 \quad a_1) & (0 \quad a_2) & (0 \quad a_3) & (0 \quad a_4)\end{matrix}$$

$$\frac{d(Xa)^T}{dX} = \left(\begin{array}{cc|cc|cc|cc} a_1 & 0 & a_2 & 0 & a_3 & 0 & a_4 & 0 \\ \hline 0 & a_1 & 0 & a_2 & 0 & a_3 & 0 & a_4 \end{array}\right)$$

例9

$$\boldsymbol{a}=\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_n \end{pmatrix} \qquad X=\begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \end{pmatrix} \qquad \frac{d(Xa)}{dX}=?$$

解

$$F(X)=Xa$$

$$\frac{d(Xa)}{dX}=\begin{pmatrix} \frac{\partial(Xa)}{\partial x_{11}} & \frac{\partial(Xa)}{\partial x_{12}} & \frac{\partial(Xa)}{\partial x_{13}} & \frac{\partial(Xa)}{\partial x_{14}} \\ \frac{\partial(Xa)}{\partial x_{21}} & \frac{\partial(Xa)}{\partial x_{22}} & \frac{\partial(Xa)}{\partial x_{23}} & \frac{\partial(Xa)}{\partial x_{24}} \end{pmatrix}$$

矩阵论 MATRIX THEORY

$$\boldsymbol{a}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\quad X=\begin{pmatrix}x_{11}&x_{12}&x_{13}&x_{14}\\x_{21}&x_{22}&x_{23}&x_{24}\end{pmatrix}$$

$$F(X)=Xa=\begin{pmatrix}\sum_{k=1}^{4}x_{1k}a_k\\\sum_{k=1}^{4}x_{2k}a_k\end{pmatrix}\leftrightarrow\begin{pmatrix}f_{11}(X)\\f_{21}(X)\end{pmatrix}\begin{matrix}\leftrightarrow e_1^T X\boldsymbol{a}\\\leftrightarrow e_2^T X\boldsymbol{a}\end{matrix}$$

$$\frac{\partial(Xa)^T}{\partial x_{11}}=\begin{pmatrix}\dfrac{\partial f_{11}}{\partial x_{11}}\\\dfrac{\partial f_{21}}{\partial x_{11}}\end{pmatrix}=\begin{pmatrix}a_1\\0\end{pmatrix}$$

$$\begin{matrix}&\begin{pmatrix}a_2\\0\end{pmatrix}&\begin{pmatrix}a_3\\0\end{pmatrix}&\begin{pmatrix}a_4\\0\end{pmatrix}\end{matrix}$$

$$\frac{d(Xa)}{dX}=\begin{pmatrix}\dfrac{\partial(Xa)}{\partial x_{11}}&\dfrac{\partial(Xa)}{\partial x_{12}}&\dfrac{\partial(Xa)}{\partial x_{13}}&\dfrac{\partial(Xa)}{\partial x_{14}}\\\dfrac{\partial(Xa)}{\partial x_{21}}&\dfrac{\partial(Xa)}{\partial x_{22}}&\dfrac{\partial(Xa)}{\partial x_{23}}&\dfrac{\partial(Xa)}{\partial x_{24}}\end{pmatrix}$$

$$\begin{matrix}\begin{pmatrix}0\\a_1\end{pmatrix}&\begin{pmatrix}0\\a_2\end{pmatrix}&\begin{pmatrix}0\\a_3\end{pmatrix}&\begin{pmatrix}0\\a_4\end{pmatrix}\end{matrix}$$

$$\frac{d(Xa)}{dX}=\left(\begin{array}{c:c:c:c} a_1 & a_2 & a_3 & a_4 \\ 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 \\ a_1 & a_2 & a_3 & a_4 \end{array}\right)$$

5.4 一阶线性常系数微分方程组

在线性控制系统中，常涉及线性微分方程组的问题，矩阵函数在其中有着重要应用，它不仅使线性微分方程组定解问题表示形式比较简单，而且可使线性微分方程组求解得到简化。

5.4.1 一阶线性常系数齐次微分方程组

齐次微分方程

$$\frac{dx}{dt}=ax(t),\ x(t_0)=x_0$$

的通解为

$$x(t)=e^{a(t-t_0)}x_0$$

将 $x(t)$、x_0 推广到向量，将系数 a 推广到任意矩阵，结论仍然成立

定理5. 4. 1 线性常系数齐次微分方程组

$$\frac{d\,x(t)}{d\,t} = Ax(t),\ x(t_0) = x_0$$

的通解为

$$\boxed{x(t) = e^{A(t-t_0)}x_0}$$

这里 $x(t) = (x_1(t), \cdots, x_n(t))^T$ ， $A = (a_{ij})_{n\times n}$ 是常数矩阵， $x(t_0) = (x_1(t_0), \cdots, x_n(t_0))^T$

证明 由于

$$\frac{d}{dt}(e^{-At}x(t)) = e^{-At}(-A)x(t) + e^{-At}\frac{d}{dt}x(t) = 0$$

两边积分得 $$\int_{t_0}^{t}\frac{d}{dt}(e^{-As}x(s))ds = 0$$

$$e^{-At}x(t) - e^{-At_0}x(t_0) = 0$$

因此 $$x(t) = e^{A(t-t_0)}x(t_0)$$

例 1 求常系数线性齐次微分方程组

$$\begin{cases} \dfrac{dx_1}{dt} = 2x_1 \\ \dfrac{dx_2}{dt} = x_1 + x_2 + x_3 \\ \dfrac{dx_3}{dt} = x_1 - x_2 + 3x_3 \end{cases}$$

在下列初始条件下的解：

$$x(0) = (x_1(0), x_2(0), x_3(0))^T = (1,1,1)^T$$

解: 方程组的矩阵形式为

$$\frac{d\,x(t)}{d\,t} = A\,x(t)$$

这里 $x = (x_1, x_2, x_3)^T, \frac{dx}{dt} = (\frac{dx_1}{dt}, \frac{dx_2}{dt}, \frac{dx_3}{dt})^T,$

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 3 \end{bmatrix}$$

根据定理5.4.1，其解为

$$x(t)=e^{At}x(0)\qquad (t_0=0)$$

这时

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

因此所求微分方程组的解为

$$x(t)=e^{At}x(0)=e^{2t}(1,1+t,1+t)^T$$

5.4.2 一阶线性常系数非齐次微分方程组

非齐次微分方程 $\frac{dx}{dt} = ax(t) + f(t),\ x(t_0) = x_0$

的通解为

$$x(t) = e^{a(t-t_0)}x_0 + \int_{t_0}^{t} e^{a(t-s)} f(s) ds$$

将 $x(t)$、x_0、$f(t)$ 推广到向量，将系数 a 推广到任意矩阵，结论仍然成立。

定理 5.4.2 线性常系数非齐次微分方程组

$$\frac{dx(t)}{dt}=Ax(t)+f(t),\ x(t_0)=x_0$$

的通解为

$$x(t)=e^{A(t-t_0)}x_0+\int_{t_0}^{t}e^{A(t-s)}f(s)ds$$

这里 $f(t)=(f_1(t),\cdots,f_n(t))^T$，其他与定理6.4.1相同。

证明 用 e^{-At} 乘方程两边，并整理得

$$e^{-At}\frac{dx(t)}{dt}+e^{-At}(-A)x(t)=e^{-At}f(t)$$

即 $$\frac{d}{dt}\left(e^{-At}x(t)\right)=e^{-At}f(t)$$

两边积分得

$$e^{-At}x(t)-e^{-At_0}x(t_0)=\int_{t_0}^{t}e^{-As}f(s))ds$$

再用 e^{At} 乘方程两边，并整理即得结果。

例 2 求常系数线性非齐次微分方程组

$$\frac{dx(t)}{dt} = Ax(t) + f(t)$$

满足初始条件 $x(0) = (1,1,1)^T$ 的解，这里

$$A = \begin{bmatrix} 3 & -1 & 1 \\ 2 & 0 & -1 \\ 1 & -1 & 2 \end{bmatrix}, f(t) = (0,0,e^{2t})^T$$

解：矩阵 A 的特征值分解为 $A = P\Lambda P^{-1}$ ，

这里

$$P=\begin{pmatrix}1 & 1 & 2\\ 5 & 1 & 1\\ 2 & 0 & 1\end{pmatrix},\quad \Lambda=\begin{pmatrix}0 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{pmatrix}$$

因此

$$e^{At}=Pe^{\Lambda t}P^{-1}x(0)=-\frac{1}{6}\begin{pmatrix}-1+3e^{2t}-8e^{3t}\\ -5+3e^{2t}-4e^{3t}\\ -2-4e^{3t}\end{pmatrix}$$

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

$$\int_0^t e^{A(t-s)}f(s)ds=-\frac{1}{6}\begin{pmatrix}\frac{1}{2}+(9t+\frac{15}{2})e^{2t}-8e^{3t}\\ \frac{5}{2}+(9t+\frac{3}{2})e^{2t}-4e^{3t}\\ 1+3e^{2t}-4e^{3t}\end{pmatrix}$$

因此所求微分方程组的解为

$$x(t)=e^{At}x_0+\int_0^t e^{A(t-s)}f(s)ds$$

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

5.4.3 Lyapunov方程

定义5. 4. 1 形如

$$AX + XB = F$$

的方程称为Lyapunov方程。

定理5. 4. 3 Lyapunov方程

$$AX + XB = F$$

其中, $A \in \mathbf{C}^{m\times m}, B \in \mathbf{C}^{n\times n}, F \in \mathbf{C}^{m\times n}$,如果$A$和$B$的所有特征值都具有负实部，则该方程有唯一解

$$X = -\int_0^{+\infty} \mathrm{e}^{At} F \mathrm{e}^{Bt} \mathrm{d}t.$$

定理5. 4. 4 设$A \in \mathbf{C}^{m\times m}, B \in \mathbf{C}^{n\times n}, F \in \mathbf{C}^{m\times n}$,则矩阵微分方程

$$\begin{cases} \dfrac{\mathrm{d}X}{\mathrm{d}t} = AX + XB \\ X(0) = F \end{cases}$$

的解为 $X(t) = \mathrm{e}^{At} F \mathrm{e}^{Bt}$.