控制系统的状态空间分析

（第 三 册）

（日）绪方胜 廖 新著　华东工程学院译

武汉钢铁学院电气自动化教研室翻印

1979年1月

第五章　矩　阵　分　析

5—1　引　言

第二章中我们介绍了关系矩阵分析的一些初步知识。在本章中我们再介绍一些本书其余部分所需用到的关于矩阵分析的补充材料。特别是我们将讨论方阵的函数、二次型以及埃尔米特形式等问题。

本章的以下各节将作如下安排：5－2节介绍矩阵多项式，Frabanius定理、Cayley—Hamilton定理、最小多项式以及矩阵的指数函数。5—3节给出Sylvester内插公式。5—4节中讨论实数二次型和埃尔米特形式，并将介绍关于实数对称矩阵和埃尔米特矩阵的定理。我们还将讨论实数二次型和埃尔米特形式的"定"、"半定"和"不定"等概念。最后5—5节介绍关于实数二次型和埃尔米特形式的"定"和"半定"的Sylvester准则。

5—2　方阵的函数

本节将讨论方阵的多项式，Frobenius定理，Cayley－Hamilton定理、最小多项式以及矩阵的指数函数。

<u>方阵的多项式</u>　设A_0，A_1，A_2，……，A_p和B_0，B_1，B_2，……，B_g为n×n矩阵。设x为标量，定义f(x)，g(x)和h(x)如下：

$$f(x)=\sum_{i=0}^{p}A_i x^i$$

$$g(x)=\sum_{j=0}^{g}B_j x^j$$

$$h(x)=f(x)g(x)=\sum_{i=0}^{p}\sum_{j=0}^{g}A_i B_j x^{i+j}$$

在上列 $h(x)$ 之方程中，如果我们用一个 $n\times n$ 矩阵 C 代表 x，则在计算时必须十分小心，因为一般的说，矩阵的乘法不服从交换律。我们有下列关系：

$$\sum_{i=0}^{p}\sum_{j=0}^{q} A_i B_j C^{i+j}$$

$$=A_1\left(\sum_{j=0}^{g} B_j C^j\right)C+\cdots\cdots+A_p\left(\sum_{j=0}^{q} B_j C^j\right)C^p$$

$$=\left(\sum_{i=0}^{p} B_i C^i\right)\left(\sum_{j=0}^{g} B_j C^j\right) \text{假定 } B_j \text{ 和 } C \text{ 可交换}$$

$$=\sum_{i=0}^{p}\sum_{j=0}^{q} A_i B_j C^{i+j} \neq \left(\sum_{i=0}^{p} A_i C^i\right)\left(\sum_{j=0}^{q} B_j C^j\right)$$

假定 B_j 和 C 不可交换

然而值得注意的是，假定 $\sum_{j=0}^{q} B_j C^j=0$ 则有

$$\sum_{i=0}^{p}\sum_{j=0}^{q} A_i B_j C^{i+j}=0$$

同样我们有

$$\sum_{i=0}^{p}\sum_{j=0}^{q} C^{i+j} A_i B_j$$

$$=\left(\sum_{i=0}^{p} C^i A_i\right)B_0+C\left(\sum_{i=0}^{p} C^i A_i\right)B_1+\cdots\cdots+$$

$$+C^q\left(\sum_{i=0}^{p} C^i A_i\right)B_g=\left(\sum_{i=0}^{p} C^i A_i\right)\left(\sum_{j=0}^{q} C^j B_j\right)$$

假定 A_j 和 C 可交换

$$\sum_{i=0}^{p}\sum_{j=0}^{q} C^{i+j} A_i B_j \neq \left(\sum_{i=0}^{p} C^i A_i\right)\left(\sum_{j=0}^{q} C^j B_j\right)$$

假设 A_j 和 C 不可交换

注意，假若 $\sum_{i=0}^{p} C^i A_j=0$ 则 $\sum_{i=0}^{p}\sum_{j=0}^{q} C^{i+j} A_i B_j=0$

以上这些结果在后面有时常要用到。

5—3

Frobenius 定理　下面将介绍关于 $n\times n$ 矩阵 A 之多项式之特征值的 Frobenius 定理。

定理 5—1（Frobenius 定理）　假设一个 $n\times n$ 矩阵 A 之特征值为 λ_1，λ_2，……，λ_n。设 $P(X)$ 为 X 之任意多项式。

即
$$P(x)=a_0X^p+a_1X^{p-1}+\cdots\cdots+a_p$$
则矩阵多项式 $P(A)$ 之特征值为 $P(\lambda_1)$，$P(\lambda_2)$，……，$P(\lambda_n)$，其中
$$P(A)=a_0A^p+a_1A^{p-1}+\cdots\cdots+a_pI$$

证：　设　$p(x)-\lambda=0$ 之根为 μ_1，μ_2，……，μ_p。

则　$p(x)-\lambda=a_0(x-\mu_1)(x-\mu_2)\cdots\cdots(x-\mu_p)=a_0\prod_{i=1}^{p}(x-\mu_i)$

考虑到
$$a_0(\mu_1+\mu_2+\cdots\cdots+\mu_p)=-a_1$$
$$a_0(\mu_1\mu_2+\mu_2\mu_3+\cdots\cdots+\mu_{p-1}\mu_p)=a_2$$
$$\cdots\cdots\qquad\cdots\cdots\qquad\cdots\cdots$$
$$a_0\mu_1\mu_2\cdots\cdots\mu_p=(-)^p(a_p-\lambda)$$

故有
$$\begin{aligned}P(A)-\lambda I&=a_0A^p+a_1A^{p-1}+\cdots\cdots+(a_p-\lambda)I\\&=a_0(A-\mu_1I)(A-\mu_2I)\cdots\cdots(A-\mu_pI)\end{aligned}$$

取 $P(A)-\lambda I$ 之行列式
$$|P(A)-\lambda I|=a_0^n|A-\mu_1I||A-\mu_2I|\cdots\cdots|A-\mu_pI|$$

由于
$$|A-\lambda I|=(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots\cdots(\lambda_n-\lambda)$$

故得
$$|P(A)-\lambda I|=a_0^n\left[\prod_{j=1}^{n}(\lambda_j-\mu_1)\right]\left[\prod_{j=1}^{n}(\lambda_j-\mu_2)\right]\cdots\cdots$$
$$\cdots\cdots\left[\prod_{j=1}^{n}(\lambda_j-\mu_p)\right]=\prod_{i=1}^{p}a_0^n\prod_{j=1}^{n}(\lambda_j-\mu_i)$$

$$=\prod_{j=1}^{n}\left[a_0\prod_{i=1}^{p}(\lambda_j-\mu_i)\right]=\prod_{j=1}^{n}\left[P(\lambda_j)-\lambda\right]$$

$$=\left[P(\lambda_1)-\lambda\right]\left[P(\lambda_2)-\lambda\right]\cdots\cdots\left[P(\lambda_n)-\lambda\right]$$

这就证明了$P(A)$的特征值为$P(\lambda_1)$，$P(\lambda_2)$，……，$P(\lambda_n)$。

Cayley-Hamilton 定理　在一个$n\times n$矩阵A之多项式的简化中Cayley-Hamilton定理有很重要的作用，此定理表明，A满足它的特征方程

我们先举一个例，然后再来讨论定理本身·考虑下列矩阵A。

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

特征方程为

$$|A-\lambda I|=-\lambda^3-3\lambda^2+7\lambda+17=0$$

注意，$(A-\lambda I)$和$\mathrm{adj}(A-\lambda I)$分别为$\lambda$之一次和二次多项式，即

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

而

$$\mathrm{adj}(A-\lambda I)=\begin{bmatrix}\lambda^2+4\lambda+3 & 2\lambda+6 & -4\\ 3\lambda+7 & \lambda^2+2\lambda-3 & -2\lambda+2\\ \lambda+1 & 2 & \lambda^2-7\end{bmatrix}$$

$$=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\lambda^2+\begin{bmatrix}4 & 2 & 0\\ 3 & 2 & -2\\ 1 & 0 & 0\end{bmatrix}\lambda+\begin{bmatrix}3 & 6 & -4\\ 7 & -3 & 2\\ 1 & 2 & -7\end{bmatrix}$$

注意，一般言$\mathrm{adj}(A-\lambda I)$为$\lambda$之$(n-1)$次多项式，其中在本例中$n=3$·如果我们以$\mathrm{adj}(A-\lambda I)$乘$(A-\lambda I)$，则得

5—5

$$(A-\lambda I)\,adj(A-\lambda I)$$

$$=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\lambda^{3}+\begin{bmatrix}-3&0&0\\0&-3&0\\0&0&-3\end{bmatrix}\lambda^{2}+\begin{bmatrix}7&0&0\\0&7&0\\0&0&7\end{bmatrix}\lambda+\begin{bmatrix}17&0&0\\0&17&0\\0&0&17\end{bmatrix}$$

$$=-\lambda^{3}I-3\lambda^{2}I+7\lambda I+17I$$

$$=(-\lambda^{3}-3\lambda^{2}+7\lambda+17)I \qquad (5-1)$$

在方程（5—1）中若以A代替λ，则

$$(-A^{3}-3A^{2}+7A+17)I$$

$$=-\begin{bmatrix}3&14&12\\27&-11&-38\\13&6&-31\end{bmatrix}-3\begin{bmatrix}7&0&-4\\-2&7&8\\-2&2&9\end{bmatrix}+7\begin{bmatrix}1&2&0\\3&-1&-2\\1&0&-3\end{bmatrix}$$

$$+17\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}$$

因此，若在特征多项式中以A代入，并乘以单位矩阵，则其结果为零·此例可说明下述定理：

定理5—2（Cayley-Hamilton） 设A为n×n矩阵，而其特征方程为

$$|A-\lambda I|=(-1)^{n}(\lambda^{n}+C_{1}\lambda^{n-1}+\cdots\cdots+C_{n-2}\lambda^{2}+C_{n-1}\lambda+C_{n})=0$$

矩阵A满足其特征方程，或即

$$A^{n}+C_{1}A^{n-1}+\cdots\cdots+C_{n-2}A^{2}+C_{n-1}A+C_{n}I=0$$

证： 为了证明此定理，我们首先注意到adj(A−λI)为λ之n－1次多项式，即

$$-adj(A-\lambda I)=B_{1}\lambda^{n-1}+B_{2}\lambda^{n-2}+\cdots\cdots+B_{n-2}\lambda^{2}+B_{n-1}\lambda+B_{n}$$

其中

$$B_1 = (-)^{n-1} I$$

还有

$$(A-\lambda I)\mathrm{adj}(A-\lambda I) = [\mathrm{adj}(A-\lambda I)](A-\lambda I)$$
$$= |A-\lambda I| I$$

于是得

$$|A-\lambda I| I = (-)^n (I\lambda^n + C_1 I\lambda^{n-1} + \cdots\cdots + C_{n-2} I\lambda^2 + C_{n-1} I\lambda + C_n I)$$
$$= (-I\lambda + A)(B_1\lambda^{n-1} + B_2\lambda^{n-2} + \cdots\cdots + B_{n-2}\lambda^2 + B_{n-1}\lambda + B_n)$$
$$= (B_1\lambda^{n-1} + B_2\lambda^{n-2} + \cdots\cdots + B_{n-2}\lambda^2 + B_{n-1}\lambda + B_n)(-I\lambda + A)$$

从上列方程可见A和B_i($i=0,1,2,\cdots\cdots,n+1$)可交换，并可见在$A-\lambda I$和$\mathrm{adj}(A-\lambda I)$两者中只要其一为零，则二者之乘积亦为零。若在上列最末方程中以A代入，则显然$A-\lambda I$为零，故

$$A^n + C_1 A^{n-1} + \cdots\cdots + C_{n-2} A^2 + C_{n-1} A + C_n I = 0$$

于是定理得证。

最小多项式　如定理5—2所述，每一$n\times n$矩阵A均满足其本身的特征方程。然而特征方程并不一定就是A所满足的次数最低的标准方程，具有A作为其根之次数最低的多项式，就称为最小多项式；也就是说，$n\times n$矩阵A的最小多项式可定义为一个次数最低的多项式$\phi(\lambda)$。

$$\phi(\lambda) = \lambda^m + a_1\lambda^{m-1} + \cdots\cdots + a_m \qquad m \leq n$$

其中有$\phi(A) = 0$，即

$$\phi(A) = A^m + a_1 A^{m-1} + \cdots\cdots + a_m I = 0$$

在$n\times n$矩阵的多项式的计算中，最小多项式将起很重要的作用。

定理5—3　设一个λ之多项式$d(\lambda)$为$n\times n$矩阵$(A-\lambda I)$之所有$(n-1)$行之式之最大公因式·将定$d(\lambda)$中最高次项的系数选择为$(-1)^n$，则最小多项式$\phi(\lambda)$由下式给定·

$$\phi(\lambda)=\frac{|A-\lambda I|}{d(\lambda)}$$

证：根据假定矩阵$adj(A-\lambda I)$之最大公因式为$d(\lambda)$，故

$$adj(A-\lambda I)=d(\lambda)B(\lambda)$$

其中$B(\lambda)$之n^2个元素（均为λ之函数）之最大公因式为1·由于

$$(A-\lambda I)adj(A-\lambda I)=|A-\lambda I|I$$

故得　$$d(\lambda)(A-\lambda I)B(\lambda)=|A-\lambda I|I$$

由此可见$|A-\lambda I|$可为$d(\lambda)$所除，设我们令

$$|A-\lambda I|=d(\lambda)\psi(\lambda)$$

则$\psi(\lambda)$之最高次项系数为1·由于我们有

$$(A-\lambda I)B(\lambda)=\psi(\lambda)I$$

故　$$\psi(A)=0$$

我们注意到$\psi(\lambda)$可写成

$$\psi(\lambda)=g(\lambda)\phi(\lambda)+\alpha(\lambda)$$

其中$\alpha(\lambda)$之次数比$\phi(\lambda)$之次数低·既有$\psi(A)=0$，故$\alpha(A)=0$·但因$\phi(\lambda)$为最小多项式，故$\alpha(\lambda)$必恒等于零，即

$$\psi(\lambda)=g(\lambda)\phi(\lambda)$$

注意，因为$\phi(A)=0$，故可写

$$\phi(\lambda)I=(A-\lambda I)C(\lambda)$$

于是

$$\psi(\lambda)I=g(\lambda)\phi(\lambda)I=g(\lambda)(A-\lambda I)C(\lambda)$$

$$B(\lambda)=g(\lambda)C(\lambda)$$

我们考虑到因为前面已假定 $B(\lambda)$ 之 n^2 个元素之最大公因式为 1。

故 $\quad g(\lambda)=1, \quad \psi(\lambda)=\phi(\lambda)$

因此而得

$$\phi(\lambda)=\frac{|A-\lambda I|}{d(\lambda)}$$

于是定理得证。

还可证明，相似矩阵具有相等的最大多项式。也即对于一个非奇异（非退化）矩阵 T，$T^{-1}AT$ 和 A 的最小多项式是相同的。（其证明可参看问题 A—5—3）。

一个 $n\times n$ 矩阵 A 之最小多项式 $\phi(\lambda)$ 可按下列步骤求取：

1.写出 $adj(A-\lambda I)$ 并将其元素写成分解为因子的 λ 的多项式。

2.确定 $adj(A-\lambda I)$ 所有元素之最大公因式 $d(\lambda)$，并将其最高次项之系数定为 $(-1)^n$。如果没有公因式，则 $d(\lambda)=(-1)^n$。

3.将 $|A-\lambda I|$ 被 $d(\lambda)$ 除，所得结果就是最小多项式 $\phi(\lambda)$。

在 A 具有不相重的特征值的情况下，最小多项式和特征多项式只有一个常系数 $(-)^n$ 之差别。

注意，若 A 之多重特征值是联成若当链的，则最小多项式和特征多项式也是相同的。举例如下。设

$$A=\begin{bmatrix}2 & 1 & 4\\ 0 & 2 & 0\\ 0 & 3 & 1\end{bmatrix}$$

特征多项式为

$$|A-\lambda I|=\begin{bmatrix}2-\lambda & 1 & 4\\ 0 & 2-\lambda & 0\\ 0 & 3 & 1-\lambda\end{bmatrix}=(2-\lambda)^2(1-\lambda)$$

故 A 之特征值为 2，2，1。可以得 A 之约当标准形为

5—9

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

可见多重特征值是联成约当链的。下面我们来说明对于此矩阵 A，其最小多项式和特征多项式是相同的。为了计算最小多项式，先求 $\mathrm{adj}(A-\lambda I)$ 如下：

$$\mathrm{adj}(A-\lambda I)=\begin{bmatrix} (2-\lambda)(1-\lambda) & \lambda+11 & -4(2-\lambda) \\ 0 & (2-\lambda)(1-\lambda) & 0 \\ 0 & -3(2-\lambda) & (2-\lambda)^2 \end{bmatrix}$$

上列矩阵各元素无公共因子，故 $d(\lambda)=(-1)^2=-1$，除了一个负号之外，$\phi(\lambda)$ 和特征多项式是相同的，即

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

$$=\lambda^3-5\lambda^2+8\lambda-4$$

简单的计算即可证明下述事实

$$A^3-5A^2+8A-4I=0$$

而 $$A^2-3A+2I\neq 0$$

注意，若 A 之多重特征值不是联成约当链的，则最小多项式之次数将比特征多项式之次数低。考虑下例

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

特征多项式为

$$|A-\lambda I|=\begin{bmatrix} 2-\lambda & 0 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 3 & 1-\lambda \end{bmatrix}=(2-\lambda)^2(1-\lambda)$$

简单的计算表明 A 有三个特征矢量，而其若当标准形为

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

可见这里多重特征值是不相联的。为了求最小多项式，我们先计算 $adj(A-\lambda I)$，即

$$adj(A-\lambda I)=\begin{bmatrix} (2-\lambda)(1-\lambda) & 0 & 0 \\ 0 & (2-\lambda)(1-\lambda) & 0 \\ 0 & 3(2-\lambda) & (2-\lambda)^2 \end{bmatrix}$$

很容易看出

$$d(\lambda)=2-\lambda$$

故

$$\phi(\lambda)=\frac{|A-\lambda I|}{d(\lambda)}=\frac{(2-\lambda)^2(1-\lambda)}{2-\lambda}=\lambda^2-3\lambda+2$$

为了检验，我们计算一下 $\phi(A)$

$$\phi(A)=A^2-3A+2I$$

$$=\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 9 & 1 \end{bmatrix}-3\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 3 & 1 \end{bmatrix}+2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

对于这个A，其最小多项式比特征多项式低一次

<u>A之逆</u>　注意，利用最小多项式可将一个非奇异（非退化）矩阵A之逆表示为具有标量系数之A之多项式。设有非奇异（非退化）矩阵A，其最小多项式 $\phi(A)$ 为

$$\phi(A)=A^m+a_1A^{m-1}+\cdots\cdots+a_mI=0$$

对于非退化矩阵 $a_m\neq 0$，故

$$I=\frac{1}{a_m}(A^m+a_1A^{m-1}+\cdots\cdots+a_{m-1}A)$$

乘以 A^{-1} 得

$$A^{-1}=-\frac{1}{a_m}(A^{m-1}+a_1A^{m-2}+\cdots\cdots+a_{m-1}I) \qquad (5-2)$$

故A' 可按方程（5—2）进行数值计算

下面介绍一个关于n×n矩阵A之无穷级数的定理。

定理5—4　设有复变量z的幂级数

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

假定此级数不论z取何值均收敛，则n×n矩阵A之矩阵级数$\sum\limits_{k=0}^{\infty}C_k A^k$绝对收敛。

证：我们先证明，当并且只有当$\sum\limits_{k=0}^{\infty}||C_k A^k||$为收敛时，无穷级数$\sum\limits_{k=0}^{\infty}C_k A^K$才是绝对收敛的。定义$A^k=(a_{ij}^k)$，假设$\sum\limits_{k=0}^{\infty}C_k A^k$为绝对收敛，则存在如下之常数M。

$$\sum_{k=0}^{\infty}C_k|a_{ij}^k|\sum_{k=0}^{N}|C_k||a_{ij}^k|<M$$

$$(N\geqslant 0;\quad ij=1,2,\cdots\cdots,n)$$

其中M与i，j，N均无关，由于

$$\sum_{k=0}^{N}||C_k A^k||\leqslant\sum_{k=0}^{N}|C_k|\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}^k|<n^2M$$

故可得结论$\sum\limits_{k=0}^{\infty}||C_k A^k||$为收敛。

反之，假设$\sum\limits_{k=0}^{\infty}||C_k A^k||$为收敛，则由$\sum\limits_{k=0}^{\infty}C_k A^k$之第（i，j）元素组成之级数为绝对收敛，因为

$$|C_k A_{ij}^k|\leqslant||C_k A^k||$$

于是我们证明了，当而且仅当$\sum\limits_{k=0}^{\infty}||C_k A^k||$为收敛时，级数$\sum\limits_{k=0}^{\infty}C_k A^k$才是绝对收敛的。

现在根据定理中的假定，我们有

$$\sum_{k=0}^{\infty}||C_kA^k||\leqslant\sum_{k=0}^{\infty}|C_k|(||A||)^k<M'$$

其中M为一常数。于是 $\sum_{k=0}^{\infty}||C_kA^k||$ 为收敛。于是定理得证。

方阵的指数函数　在标量的情况下e^z定义如下：

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

由于下列幂级数

$$\sum_{k=0}^{\infty}\frac{1}{k!}z^k$$

不论z取何值均收敛，故下列矩阵幂级数

$$\sum_{k=0}^{\infty}\frac{1}{k!}A^k$$

对任何方阵A均为绝对收敛。因此对于每个方阵A我们均可定义其指数函数为

$$e^A=I+A+\frac{1}{2!}A^2+\frac{1}{2!}A^3+\cdots\cdots$$

其中我们用I代替A^0。同样，由于cos z和sinz也都是对于任何z均收敛，故可定义

$$\cos A=I-\frac{1}{2!}A^2+\frac{1}{4!}A^4-\cdots\cdots$$

$$\sin A=A-\frac{1}{3!}A^3+\frac{1}{5!}A_5-\cdots\cdots$$

根据A之指数函数之定义，At之指数函数可写成

$$e^{At}=I+At+\frac{1}{2!}A^2t^2+\cdots\cdots+\frac{1}{n!}A^nt^n+\cdots\cdots$$

此级数对于任意有限的t值均为绝对收敛（参看问题A—5—5）。

我们再来看一下e^{At}对于t的导数。对t微分e^{At}可得出

$$\frac{d}{dt}(e^{At}) = A + A^2t + \cdots\cdots + A^n\frac{nt^{n-1}}{n!} + \cdots\cdots$$

$$= A(I + At + \cdots\cdots + A^{n-1}\frac{t^{n-1}}{(n-1)!} + \cdots\cdots)$$

$$= Ae^{At} = (I + At + \cdots\cdots + A^{n-1}\frac{t^{n-1}}{(n-1)!} + \cdots)A$$

$$= e^{At}A$$

此关系式和标量的情况是相同的。我们再看 $e^{At}e^{AS}$。

$$e^{At}e^{AS} = (\sum_{n=0}^{\infty}\frac{A^nt^n}{n!})(\sum_{n=0}^{\infty}\frac{A^nS^n}{n!})$$

$$= \sum_{n=0}^{\infty}A^n(\sum_{k=0}^{\infty}\frac{t^kS^{n-k}}{k!(n-k)!})$$

$$= \sum_{n=0}^{\infty}A^n\frac{(S+t)^n}{n!} = e^{A(s+t)}$$

故得 $\quad e^{At}e^{As} = e^{A(s+t)}$

如果 $\quad s = -t$，则

$$e^{At}e^{-At} = e^{-At}e^{At} = e^{A(t-t)} = I$$

故 e^{At} 之逆为 e^{-At}。由于存在逆矩阵，故 e^{At} 是非退化的。一般言 $e^{(A+B)t}$ 和 $e^{At}e^{Bt}$ 是不相同的，除非 A 和 B 可交换，即

$$AB = BA$$

我们来考察一下，$e^{(A+B)t}$ 与 $e^{At}e^{Bt}$ 之间的差别。我们有

$$e^{(A+B)t} = I + (A+B)t + \frac{(A+B)^2}{2}t^2 + \frac{(A+B)^3}{3!}t^3 + \cdots\cdots$$

而

$$e^{At}e^{Bt} = (I + At + \frac{A^2t^2}{2} + \frac{A^3t^3}{3!} + \cdots\cdots)$$

$$(I + Bt + \frac{B^2t^2}{2} + \frac{B^3t^3}{3!} + \cdots\cdots)$$

$$= I + (A+B)t + \frac{A^2t^2}{2} + ABt^2 + \frac{B^2t^2}{2} +$$

$$+ \frac{A^3t^3}{3!} + \frac{A^2Bt^3}{2} + \frac{AB^2t^3}{2} + \frac{B^3t^3}{3!} + \cdots\cdots$$

于是

$$e^{(A+B)t} - e^{At}e^{Bt}$$

$$= \frac{(A+B)(A+B)}{2}t^2 - \frac{A^2+2AB+B^2}{2}t^2 +$$

$$+ \frac{(A+B)(A+B)(A+B)}{3!}t^3 - \frac{A^3+3A^2B+3AB^2+B^3}{3!}t^3 + \cdots$$

$$= \frac{BA-AB}{2}t^2 \frac{BA^2+ABA+B^2A+BAB-2A^2B-2AB^2}{3!}t^3 + \cdots\cdots$$

如果 A 和 B 可交换，即 A B = B A，则可见对于所有 t 均有

$$e^{(A+B)t} - e^{At}e^{Bt} = 0$$

故 $e^{(A+B)t} = e^{At}e^{Bt}$ 如果 A 和 B 可交换。

如果 A 和 B 不可交换，则 $e^{(A+B)t}$ 不等于 $e^{At}e^{Bt}$。

作为一个例子，我们对于下列 A 计算 e^{At}

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

我们看到

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = B + C$$

其中

$$B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

我们看到 B 和 C 是可交换的即 B C = C B 故有

$$e^{At} = e^{(B+C)t} = e^{Bt}e^{Ct}$$

5—15

e^{Ct} 可给出如下：

$$e^{ct}=e^{It}=I+It+\frac{I^2t^2}{2}+\cdots\cdots+\frac{I^nt^n}{n!}+\cdots=e^tI$$

e^{Bt} 则由下式给出

$$e^{Bt}=I+Bt+\frac{B^2t^2}{2}+\cdots\cdots+\frac{B^nt^n}{n!}+\cdots\cdots$$

其中

$$B^2=3B,\ B^3=3B^2=3^2B,\quad B^4=3^2B^2=3^3B,\cdots$$

故
$$\begin{aligned}e^{Bt}&=I+tB+\frac{t^2}{2}(3B)+\frac{t^3}{3!}(3^2B)+\cdots\cdots\\&=I+\frac{1}{3}\left[3t+\frac{(3t)^2}{2!}+\frac{(3t)^3}{3!}+\cdots\cdots\right]B\\&=I+\frac{1}{3}(e^{3t}-1)B\end{aligned}$$

故我们有

$$e^{At}=\left[I+\frac{1}{3}(e^{3t}-1)B\right]e^tI=e^tI+\frac{1}{3}(e^{4t}-e^t)B$$

$$=\begin{bmatrix}\frac{1}{3}(e^{4t}+2e^t) & \frac{1}{3}(e^{4t}-e^t) & \frac{1}{3}(e^{4t}-e^t)\\ \frac{1}{3}(e^{4t}-e^t) & \frac{1}{3}(e^{4t}-2e^t) & \frac{1}{3}(e^{4t}-e^t)\\ \frac{1}{3}(e^{4t}-e^t) & \frac{1}{3}(e^{4t}-e^t) & \frac{1}{3}(e^{4t}-2e^t)\end{bmatrix}$$

上面所说明的直接计算 e^{At} 的方法，只在一些简单的情况能实际应用·在6—4节中我们将讨论计算 e^{At} 之别的方法。

下面介绍一个关于A之指数函数的定理·

定理5—5　对于一个具有n个不同的特征矢量的 $n\times n$ 矩阵A，e^A 可写成

$$e^A=PMP^{-1}\qquad(P=n\times n\text{ 非退化矩阵})$$

其中

$$M=\begin{bmatrix} e^{\lambda_1} & & & 0 \\ & e^{\lambda_2} & \ddots & \\ 0 & & & e^{\lambda n} \end{bmatrix}$$

而 λ_1，λ_2，……，λn 则是A之特征值，我们并不一定要各不相同。

证：　　e^A 可给出如下：

$$e^A=\sum_{k=0}^{\infty}\frac{1}{k!}A^k$$

由于A可写成为

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

其中

$$D=\begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & \ddots & \\ 0 & & & \lambda n \end{bmatrix}$$

我们得

$$A^2=PDP^{-1}\,PDP^{-1}=PD^2P^{-1}$$

因此 A^k 可写成为

$$A^k=PD^kP^{-1}$$

其中

$$D^k=\begin{bmatrix} \lambda_1^k & & & 0 \\ & \lambda_2^k & \ddots & \\ 0 & & & \lambda_n^k \end{bmatrix}$$

故有

$$\sum_{k=0}^{\infty}\frac{1}{k!}D^k=\begin{bmatrix} \sum_{k=0}^{\infty}\frac{1}{k!}\lambda_1^k & & & 0 \\ & \sum_{k=0}^{\infty}\frac{1}{k!}\lambda_2^k & \ddots & \\ 0 & & & \sum_{k=0}^{\infty}\frac{1}{k!}\lambda_n^k \end{bmatrix}$$