矩阵理论

王磊

自动化科学与电气工程学院

矩 阵 函 数

定义4.7.1(矩阵函数) 设复幂级数$\sum_{m=0}^{\infty} c_m z^m$的收敛半径为$r$. 当$|z|<r$时, 幂级数$\sum_{m=0}^{\infty} c_m z^m$收敛于函数$f(z)$, 即

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

若复方阵A满足$\rho(A)<r$, 称收敛的矩阵幂级数$\sum_{m=0}^{\infty} c_m A^m$为矩阵函数, 记为$f(A)$.

常见矩阵函数有

$e^{A} = \sum_{m=0}^{\infty} \frac{1}{m!} A^{m}, \forall A \in \mathbb{C}^{n\times n}$ (矩阵指数函数)

$\sin A = \sum_{m=0}^{\infty} \frac{(-1)^{m}}{(2m+1)!} A^{2m+1}, \forall A \in \mathbb{C}^{n\times n}$ (矩阵正弦函数)

$\cos A = \sum_{m=0}^{\infty} \frac{(-1)^{m}}{(2m)!} A^{2m}, \forall A \in \mathbb{C}^{n\times n}$ (矩阵余弦函数)

$(I-A)^{-1} = \sum_{m=0}^{\infty} A^{m}, \forall \rho(A) < 1$

$\ln(I+A) = \sum_{m=0}^{\infty} \frac{(-1)^{m}}{m+1} A^{m+1}, \forall \rho(A) < 1$

命题4.7.1 设$A \in \mathbb{C}^{n\times n}$，则以下结论成立：

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

（2）$e^{\mathrm{i}A} = \cos A + \mathrm{i}\sin A$；

（3）$\cos A = \frac{1}{2}(e^{\mathrm{i}A} + e^{-\mathrm{i}A})$；

（4）$\sin A = \frac{1}{2\mathrm{i}}(e^{\mathrm{i}A} - e^{-\mathrm{i}A})$.

注1: 尽管指数函数满足$e^a e^b = e^{a+b} = e^b e^a$, 但这一性质对矩阵指数函数一般不成立.

例4.7.1 已知$A=\begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix}$和$B=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$，分别计算$e^{A}, e^{B}, e^{A+B}, e^{A}e^{B}, e^{B}e^{A}$，并比较$e^{A+B}, e^{A}e^{B}$和$e^{B}e^{A}$.

例4.7.1 已知$A=\begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix}$和$B=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$，分别计算$e^A, e^B, e^{A+B}, e^Ae^B, e^Be^A$，并比较$e^{A+B}, e^Ae^B$和$e^Be^A$.

$$\boldsymbol{e}^A=\begin{bmatrix}e & 1-e\\ 0 & 1\end{bmatrix}, \boldsymbol{e}^B=\begin{bmatrix}e & e-1\\ 0 & 1\end{bmatrix},$$

$$\boldsymbol{e}^A\boldsymbol{e}^B=\begin{bmatrix}e^2 & (e-1)^2\\ 0 & 1\end{bmatrix},$$

$$\boldsymbol{e}^B\boldsymbol{e}^A=\begin{bmatrix}e^2 & -(e-1)^2\\ 0 & 1\end{bmatrix},$$

$$\boldsymbol{e}^{A+B}=\begin{bmatrix}e^2 & 0\\ 0 & 0\end{bmatrix}.$$

注2:矩阵函数的定义式提供了一种计算矩阵函数的方法.在采用定义法计算矩阵函数时应巧妙地应用矩阵的最小多项式进行简化计算.

定理4.7.1 设$A, B \in \mathbb{C}^{n\times n}$, 若$AB = BA$, 则

$$e^{A}e^{B} = e^{B}e^{A} = e^{A+B}.$$

推论4.7.1 设$A \in \mathbb{C}^{n\times n}$, 则$e^{A}e^{-A} = e^{-A}e^{A} = I$, 即$(e^{A})^{-1} = e^{-A}$.

推论4.7.2 设$A \in \mathbb{C}^{n\times n}$, 则$\sin^2 A + \cos^2 A = I$.

注3: 推论4.7.1表明无论矩阵A是否可逆, 矩阵指数函数e^{A}必可逆, 且其逆矩阵为e^{-A}.

若矩阵函数$f(A)$的自变量由矩阵A换成At，其中t为标量参数，则有矩阵函数表达式为

$$f(At)=\sum_{m=0}^{\infty}c_m(At)^m=\sum_{m=0}^{\infty}c_m t^m A^m\ ,t\rho(A)<r$$

我们称之为含参矩阵函数.

这类含参矩阵函数常在线性常微分方程求解等应用中遇到.

例4.7.2 设$A=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$，求$e^{At}$.

例4.7.3 设$A \in \mathbb{C}^{4\times 4}$, 其特征值分别为$\pi$, $-\pi$, 0, 0, 求$\sin A$和$\cos A$.

定理4.7.2 设复方阵A与B相似,即存在可逆矩阵P使得$P^{-1}AP = B$. 若$f(A)$是矩阵函数,则

$$f(A) = Pf(B)P^{-1}.$$

基于定理4.7.2，我们对矩阵A分两种情况讨论：

（1）A是单纯矩阵；（2）A非单纯矩阵.

（1）若A是单纯矩阵，则存在可逆矩阵P使得$P^{-1}AP=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$，进而

$$f(A)=P\mathrm{diag}(f(\lambda_1),\cdots,f(\lambda_n))P^{-1}$$

$$f(At)=P\mathrm{diag}(f(\lambda_1 t),\cdots,f(\lambda_n t))P^{-1}$$

此时，矩阵函数和含参矩阵函数仍是单纯矩阵.

$$e^A=P\mathrm{diag}(e^{\lambda_1},\cdots,e^{\lambda_n})P^{-1}$$

$$\sin A=P\mathrm{diag}(\sin\lambda_1,\cdots,\sin\lambda_n)P^{-1}$$

例4.7.4 设矩阵$A=\begin{bmatrix}1 & 2\\ 0 & 2\end{bmatrix}$, 求矩阵函数$\sin(A)$.

例4.7.4 设矩阵$A=\begin{bmatrix}1 & 2\\ 0 & 2\end{bmatrix}$，求矩阵函数$\sin(A)$.

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

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

$$\sin A=\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}\sin\left(\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}\right)\begin{bmatrix}1 & -2\\ 0 & 1\end{bmatrix}$$

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

思考:除了定义法和对角化分解方法求解单纯矩阵的矩阵函数式,还可以采用什么方法进行计算?

思考：除了定义法和对角化分解方法求解单纯矩阵的矩阵函数式，还可以采用什么方法进行计算？

利用单纯矩阵的谱分解也可方便地计算矩阵函数.

（1）若A是单纯矩阵，且有谱分解$A=\sum_{j=1}^{m}\lambda_j E_j$，则有

$$f(A)=\sum_{j=1}^{m}f(\lambda_j)E_j,$$

$$f(At)=\sum_{j=1}^{m}f(\lambda_j t)E_j.$$

例4.7.4 设矩阵$A=\begin{bmatrix}1 & 2\\ 0 & 2\end{bmatrix}$, 求矩阵函数$\sin(A)$.

（2）若A为非单纯矩阵，则存在可逆矩阵P使得$P^{-1}AP=J$，其中J是矩阵A的Jordan标准形，可表示为$J=\mathrm{diag}(J_1,\cdots,J_s)$，$J_i(i=1,\cdots,s)$为Jordan块，可表示为

$$J_i(\lambda_i)=\begin{bmatrix} \lambda_i & 1 & & & 0 \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ 0 & & & & \lambda_i \end{bmatrix}_{n_i\times n_i}$$

此时, $A^k = PJ^kP^{-1}$, 式中,

$$J^k = \mathrm{diag}(J_1^k, \cdots, J_s^k)$$

$J_i^k(i = 1, \cdots, s)$满足

$$J_i^k = \begin{bmatrix} \lambda_i^k & C_k^1\lambda_i^{k-1} & C_k^2\lambda_i^{k-2} & \cdots & C_k^{n_i-1}\lambda_i^{k-n_i+1} \\ & \lambda_i^k & C_k^1\lambda_i^{k-1} & \cdots & C_k^{n_i-2}\lambda_i^{k-n_i+2} \\ & & \ddots & \ddots & \vdots \\ & & & \lambda_i^k & C_k^1\lambda_i^{k-1} \\ & & & & \lambda_i^k \end{bmatrix} \in \mathbb{C}^{n_i \times n_i}$$

将上式代入矩阵函数$f(A)=\sum_{m=0}^{\infty}c_mA^m$，最终得

$$f(A)=P\begin{bmatrix} f(J_1(\lambda_1)) & & 0 \\ & \ddots & \\ 0 & & f(J_s(\lambda_s)) \end{bmatrix}P^{-1}$$

$$f(J_i(\lambda_i))=\begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{1}{(n_i-1)!}f^{n_i-1}(\lambda_i) \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & f'(\lambda_i) \\ 0 & & & f(\lambda_i) \end{bmatrix}_{n_i\times n_i}$$

上式称为Sylvester公式.

推论4.7.3 设$A \in \mathbb{C}^{n\times n}$的特征值$\lambda_1, \cdots, \lambda_n$, $f(z) = \sum_{m=0}^{\infty} c_m z^m$的收敛半径为$r$. 当$\rho(A) < r$时, 矩阵函数$f(A)$的特征值为$f(\lambda_1), \cdots, f(\lambda_n)$.

例4.7.5 设矩阵$A=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 2 & -5 & 4\end{bmatrix}$，求矩阵函数$\sin A$.

例4.7.5 设矩阵$A=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 2 & -5 & 4\end{bmatrix}$, 求矩阵函数$\sin A$.

所以A的Jordan标准形J及相似变换矩阵P为:

$$J=\begin{bmatrix}2 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}\quad P=\begin{bmatrix}1 & 1 & 0\\ 2 & 1 & 1\\ 4 & 1 & 2\end{bmatrix}$$

$$\sin J=\begin{bmatrix}\sin 2 & 0 & 0\\ 0 & \sin 1 & \cos 1\\ 0 & 0 & \sin 1\end{bmatrix}$$

推论4.7.4 设矩阵$A\in\mathbb{C}^{n\times n}$有分解式$A=PJP^{-1}$，其中$J=\mathrm{diag}(J_1,\cdots,J_s)$是$A$的Jordon标准形，幂级数$f(z)=\sum_{m=0}^{\infty}c_m z^m$的收敛半径为$r$. 当$t\rho(A)<r$时，含参矩阵函数

$$f(At)=P\mathrm{diag}(f(J_1(\lambda_1 t),\cdots,f(J_s(\lambda_s t))P^{-1}$$

$$f(J_i(\lambda_i t))=\begin{bmatrix} f(\lambda_i t) & tf'(\lambda_i t) & \cdots & \dfrac{t^{n_i-1}}{(n_i-1)!}f^{n_i-1}(\lambda_i t) \\ & f(\lambda_i t) & \ddots & \vdots \\ & & \ddots & tf'(\lambda_i t) \\ 0 & & & f(\lambda_i t) \end{bmatrix}_{n_i\times n_i}$$

例4.7.6 设矩阵$A=\begin{bmatrix}2 & 1\\ 0 & 2\end{bmatrix}$, 求矩阵函数$\cos(At)$.

例4.7.6 设矩阵$A=\begin{bmatrix}2 & 1\\ 0 & 2\end{bmatrix}$, 求矩阵函数$\cos(At)$.

$$\cos(At)=\begin{bmatrix}\cos 2t & t\cos' 2t\\ 0 & \cos 2t\end{bmatrix}$$
$$=\begin{bmatrix}\cos 2t & -t\sin 2t\\ 0 & \cos 2t\end{bmatrix}.$$

定理4.7.3 设矩阵$A \in \mathbb{C}^{n\times n}$的最小多项式的阶数为$l$，幂级数$f(z)=\sum_{m=0}^{\infty} c_m z^m$的收敛半径为$r$. 若$\rho(A)<r$，定义矩阵函数$f(A)$，则必存在唯一的$(l-1)$次矩阵多项式

$$p(A)=\beta_0 I+\beta_1 A+\cdots+\beta_{l-1}A^{l-1}$$

使得$f(A)=p(A)$.

例4.7.8 考查$f(z)$、$p(z)$和矩阵A

$$f(z)=e^z,$$

$$p(z)=\beta_0+\beta_1 z,$$

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

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

定义4.7.2(谱上给定) 设$\lambda_1,\cdots,\lambda_s$是$n$阶复方阵$A$的$s$个互异特征值，$m_A(\lambda)=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}$ 是A的最小多项式，$\deg(m_A(\lambda))=l$.

若复函数$f(z)$及其各阶导数$f^{(i)}(z)$在$z=\lambda_i$处的n_i个值$f^{(i)}(\lambda_i)$均有界，$j=0,1,\cdots,n_i-1$，

- 称$f(z)$在矩阵A的谱上给定(或谱上有定义)，
- 称$\lambda_1,\cdots,\lambda_s$为谱点，
- $f^{(i)}(\lambda_i)$为$f(z)$在矩阵A上的谱值.

例4.7.7 考查函数$f(z)$在矩阵A的谱上是否有定义，其中

$$f(z)=\frac{1}{(z-3)(z-4)},A=\begin{bmatrix}2&0&0\\0&1&1\\0&0&1\end{bmatrix}$$

若$f(z)=\frac{1}{(z-1)(z-4)}$,

问$f(z)$在矩阵A的谱上是否给定?

定义4.7.3(谱上一致) 设复方阵A的最小多项式为

$$m_A(\lambda)=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}, \deg(m_A(\lambda))=l.$$

若函数$f(\lambda)$和$p(\lambda)$在谱上给定且满足

$$\begin{cases} f(\lambda_i)=p(\lambda_i) \\ f'(\lambda_i)=p'(\lambda_i) \\ \vdots \\ f^{(n_i-1)}(\lambda_i)=p^{(n_i-1)}(\lambda_i) \end{cases}, i=1,\cdots,s$$

则称函数$f(\lambda)$和$p(\lambda)$在矩阵A的谱上一致.

定理4.7.4 设$A \in \mathbb{C}^{n\times n}$, 幂级数$f(z) = \sum_{m=0}^{\infty} c_m z^m$与多项式$p(z) = \sum_{i=0}^{k} \beta_i z^i$在矩阵$A$的谱上给定, 则$f(A) = p(A)$的充分必要条件是$f(z)$和$p(z)$在矩阵$A$的谱上一致.

根据定理4.7.4的谱上一致性条件，我们可列出l个独立方程. 由此可求解出定理4.7.3中的l个系数$\beta_0,\beta_1,\cdots,\beta_{l-1}$.

这就是计算矩阵函数的第二种方法，我们称之为**谱上一致性方法**.

例4.7.9 设矩阵$A=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 2 & -5 & 4\end{bmatrix}$, 求矩阵函数$e^{A}$.

例4.7.9 设矩阵 $A=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 2 & -5 & 4\end{bmatrix}$, 求矩阵函数 $\sin A$.

解: 方阵A的最小多项式为

$$m_A(\lambda)=(\lambda-2)(\lambda-1)^2$$

则设多项式$p(\lambda)=a_2\lambda^2+a_1\lambda+a_0$

由$\begin{cases} f(2)=p(2)\\ f(1)=p(1)\\ f'(1)=p'(1)\end{cases} \Rightarrow \begin{cases} 4a_2+2a_1+a_0=\sin 2\\ a_2+a_1+a_0=\sin 1\\ 2a_2+a_1=\cos 1\end{cases}$

例4.7.9 设矩阵 $A=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 2 & -5 & 4\end{bmatrix}$，求矩阵函数 $\sin A$.

解得

$$\begin{cases} a_2 = \sin 2 - \sin 1 - \cos 1 \\ a_1 = -2\sin 2 + 2\sin 1 + 3\cos 1 \\ a_0 = \sin 2 - 2\cos 1 \end{cases}$$

$$e^A = a_2A^2 + a_1A + a_0I$$

注4: 利用谱上一致性方法计算矩阵函数不仅适用于单纯矩阵，也适用于非单纯矩阵.

设$f(At)=\sum_{m=0}^{\infty} c_m' A^m$，其中，$c_m'=c_m t^m$.

故含参矩阵函数$f(At)$仍可以唯一地由$l-1$次矩阵多项式$p_t(A)$表示，即

$$\begin{aligned} f(At) &= \alpha_0(t)I+\alpha_1(t)A+\cdots+\alpha_{l-1}(t)A^{l-1} \\ &= p_t(A) \end{aligned}$$

式中，$\alpha_i(t)(i=0,1,\cdots,l-1)$为待定含参系数.

利用Sylvester公式，可得如下方程组

$$p_t^{(j)}(\lambda_i) = t^j f^{(j)}(\lambda_i t),$$

$$j = 0, \cdots, n_i - 1, i = 1, \cdots, s$$

通过求解上式方程组可确定待定的l个系数$\alpha_i(t)$，$i = 0,1, \cdots, l-1.$

例4.7.10 设矩阵$A=\begin{bmatrix}2&1&0\\0&2&1\\0&0&2\end{bmatrix}$，求矩阵函数$e^{At}$.

$$\begin{cases}\alpha_0(t)+2\alpha_1(t)+4\alpha_2(t)=e^{2t}\\ \alpha_1(t)+4\alpha_2(t)=te^{2t}\\ 2\alpha_2(t)=t^2e^{2t}\end{cases}$$

注5:利用谱上一致性方法计算矩阵函数时,可不选用矩阵的最小多项式,而选用矩阵的任一零化多项式.此时须把选用的零化多项式当作矩阵的最小多项式,然后仿照上面的方法求解即可.

再次考查例4.7.3 设$A \in \mathbb{C}^{4\times 4}$, 其特征值分别为

π, $-\pi$, 0, 0, 求$\sin A$和$\cos A$.

例4.7.11 设函数$f(z)=\frac{1}{z}$,矩阵$A=\begin{bmatrix}2&1&0\\0&2&1\\0&0&2\end{bmatrix}$,求$f(A)$.

方法一：Sylvester公式. 由Sylvester公式知，

$$f(2)=\frac{1}{2},\ f'(2)=-\frac{1}{z^2}\bigg|_{z=2}=-\frac{1}{4},\ \frac{1}{(3-1)!}f^2(2)=\frac{1}{z^3}\bigg|_{z=2}=\frac{1}{8}$$

$$f(A)=\begin{bmatrix}\frac{1}{2}&-\frac{1}{4}&\frac{1}{8}\\0&\frac{1}{2}&-\frac{1}{4}\\0&0&\frac{1}{2}\end{bmatrix}$$

例4.7.11 设函数 $f(z)=\frac{1}{z}$, 矩阵 $A=\begin{bmatrix}2&1&0\\0&2&1\\0&0&2\end{bmatrix}$, 求 $f(A)$.

方法二：谱上一致性方法.

$$\begin{cases}\alpha_0+2\alpha_1+4\alpha_2=\frac{1}{2}\\ \alpha_1+4\alpha_2=-\frac{1}{4}\\ 2\alpha_2=\frac{1}{4}\end{cases}\Rightarrow \alpha_0=\frac{3}{2},\ \ \alpha_1=-\frac{3}{4}\text{和}\alpha_2=\frac{1}{8}.$$

$$f(A)=\frac{3}{2}I-\frac{3}{4}A+\frac{1}{8}A^2=\begin{bmatrix}\frac{1}{2}&-\frac{1}{4}&\frac{1}{8}\\0&\frac{1}{2}&-\frac{1}{4}\\0&0&\frac{1}{2}\end{bmatrix}$$

Sylvester公式和谱上一致性方法所需条件为$f(z)$在矩阵A的谱上给定.

尽管在定义或结论中我们仍要求矩阵的谱半径小于收敛半径，实际上我们可移除这一条件.

谱上给定条件比定义4.7.1的幂级数要求弱.

因此，我们可拓宽定义4.7.1矩阵函数的定义.

定义4.7.4(矩阵函数) 设复方阵A的最小多项式为$m_A(\lambda)=(\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_s)^{n_s}, \deg(m_A(\lambda))=l.$ 若函数$f(\lambda)$在矩阵A的谱上给定,则矩阵函数$f(A)$定义为$f(A)=p(A)$,式中,$p(A)=\beta_0 I+\beta_1 A+\cdots+\beta_{l-1}A^{l-1}$,$l$个系数$\beta_0,\beta_1,\cdots,\beta_{l-1}$由以下方程组决定

$$p^{(j)}(\lambda_i)=f^{(j)}(\lambda_i),$$

$$j=0,\cdots,n_i-1,$$

$$i=1,\cdots,s$$

注6：矩阵函数也可以利用Sylvester公式定义，这里就不再赘述了. 特别地，当函数$f(z)$能够展开为“z”的幂级数时，矩阵函数的定义4.7.1和定义4.7.4一致.