矩阵分析与应用

第十讲　矩阵分析及其应用之二

信息工程学院
吕旌阳

2006-12-1

本讲主要内容

- 矩阵函数的值的计算（续）
- 矩阵函数的一般定义
- 矩阵函数的性质
- 矩阵的微分和积分

2. 数项级数求和法。

利用首一多项式 $\psi(\lambda)$ ，且满足 $\psi(A)=0$，即

$$A^m + b_1 A^{m-1} + \cdots + b_{m-1} A + b_m I = 0$$

或者 $A^m = k_0^{(0)} I + k_1^{(0)} A + \cdots + k_{m-1}^{(0)} A^{m-1} \quad \left(k_i^{(0)} = -b_{m-i}\right)$

可以求出 $A^{m+1} = A^m A = k_0^{(1)} I + k_1^{(1)} A + \cdots + k_{m-1}^{(1)} A^{m-1}$

$$\vdots$$

$$A^{m+l} = k_0^{(l)} I + k_1^{(l)} A + \cdots + k_{m-1}^{(l)} A^{m-1}$$

$$\vdots$$

于是 $f(A) = \sum_{k=0}^{\infty} c_k A^k = \left(c_0 I + c_1 A + \cdots + c_{m-1} A^{m-1}\right) + c_m \left(k_0^{(0)} I + k_1^{(0)} A + \cdots + k_{m-1}^{(0)} A^{m-1}\right) + \cdots$

$$= \left(c_0 + \sum_{l=0}^{\infty} c_{m+l} k_0^{(l)}\right) I + \left(c_1 + \sum_{l=0}^{\infty} c_{m+l} k_1^{(l)}\right) A + \cdots + \left(c_{m-1} + \sum_{l=0}^{\infty} c_{m+l} k_{m-1}^{(l)}\right) A^{m-1}$$

例7： $A=\begin{bmatrix}\pi & 0 & 0 & 0\\ & -\pi & 0 & 0\\ & & 0 & 1\\ & & & 0\end{bmatrix}$ ，求 $\sin A$

解：$\varphi(\lambda)=|\lambda I-A|=\lambda^4-\pi^2\lambda^2$，取 $\psi(\lambda)=\varphi(\lambda)$

$$\psi(A)=0\Rightarrow A^4=\pi^2A^2,\ A^5=\pi^2A^3, A^7=\pi^4A^3,\cdots$$

$$\sin A=A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\frac{1}{7!}A^7+\cdots$$

$$=A+\left[-\frac{1}{3!}+\frac{\pi^2}{5!}-\frac{\pi^4}{7!}+\cdots\right]A^3$$

$$=A+\frac{1}{\pi^3}[\sin\pi-\pi]A^3=A-\frac{1}{\pi^2}A^3=\begin{bmatrix}0 & 0 & 0 & 0\\ & 0 & 0 & 0\\ & & 0 & 1\\ & & & 0\end{bmatrix}$$

$$\because A^3=\mathrm{diag}\left(\pi^3,-\pi^3,0,0\right)$$

3. 对角阵法

设 $P^{-1}AP = \mathrm{diag}(\lambda_1, \cdots, \lambda_n) = \Lambda$ ， 则 $A^k = P\Lambda^k P^{-1}$,

且有 $\sum_{k=0}^{N} c_k A^k = P\sum_{k=0}^{N} c_k \Lambda^k P^{-1}$

$$= P\,\mathrm{diag}\left(\sum_{k=0}^{N} c_k \lambda_1^k, \cdots, \sum_{k=0}^{N} c_k \lambda_n^k\right) P^{-1}$$

于是

$$f(A) = \sum_{k=0}^{\infty} c_k A^k = P \bullet \mathrm{diag}\left(f(\lambda_1), \cdots, f(\lambda_n)\right) \bullet P^{-1}$$

例8： $P^{-1}AP=\Lambda$ ：

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

$$e^{tA}=P\bullet \mathrm{diag}\left(e^{\lambda_1 t},\cdots,e^{\lambda_n t}\right)\bullet P^{-1}$$

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

4.Jordan标准型法

设 $P^{-1}AP = J = \mathrm{diag}\left(J_1, \cdots, J_s\right),\ J_i = \lambda_1 I + I^{(1)}$

易证 $I^{(k)}I^{(1)} = I^{(1)}I^{(k)} = I^{(k+1)}, I^{(m_i)} = O$

$$k \le m_i - 1: J_i^k = \lambda_i^k I + C_k^1 \lambda_i^{k-1} I^{(1)} + \cdots + C_k^{k-1} \lambda_i I^{(k-1)} + I^{(k)}$$

$$k \ge m_i : J_i^k = \lambda_i^k I + C_k^1 \lambda_i^{k-1} I^{(1)} + \cdots + C_k^{m_i-1} \lambda_i^{k-m_i+1} I^{(m_i-1)}$$

$$f\left(J_i\right) = \sum_{k=0}^{\infty} c_k J_i^k = f\left(\lambda_i\right) I + \frac{f'\left(\lambda_i\right)}{1!} I^{(1)} + \cdots + \frac{f^{(m_i-1)}\left(\lambda_i\right)}{(m_i-1)!} I^{(m_i-1)}$$

$$f\left(A\right) = \sum_{k=0}^{\infty} c_k A^k = P \bullet \sum_{k=0}^{\infty} c_k J^k \bullet P^{-1} = P \bullet diag\left(f(J_1), \cdots, f(J_s)\right) \bullet P^{-}$$

三、矩阵函数的一般定义

展开式 $f(z)=\sum c_k z^k,\quad (|z|<r, r>0)$，要求

（1） $f^{(k)}(0)$ 存在 $\quad (k=0,1,2,\cdots)$

（2） $\lim\limits_{k\to\infty}\dfrac{f^{(k+1)}(\xi)}{(k+1)!}z^{k+1}=0 \quad (|z|<r)$

对于一元函数 $f(z)=\dfrac{1}{z}$ 等，还不能定义矩阵函数。

基于矩阵函数值的Jordan标准形算法，拓宽定义

矩阵函数的一般定义

设 $P^{-1}AP = J = \mathrm{diag}\left(J_1, \cdots, J_s\right), \ J_i = \lambda_1 I + I^{(1)}$

如果 $f(z)$ 在 λ_i 处有 $m_i - 1$ 阶导数，令

$$f\left(J_i\right) = \sum_{k=0}^{\infty} c_k J_i^k = f\left(\lambda_i\right) I + \frac{f'\left(\lambda_i\right)}{1!} I^{(1)} + \cdots + \frac{f^{(m_i-1)}\left(\lambda_i\right)}{(m_i-1)!} I^{(m_i-1)}$$

$$f\left(A\right) = \sum_{k=0}^{\infty} c_k A^k = P \bullet \sum_{k=0}^{\infty} c_k J^k \bullet P^{-1} = P \bullet diag\left(f(J_1), \cdots, f(J_s)\right) \bullet P^{-}$$

称 $f\left(A\right)$ 为对应于 $f\left(z\right)$ 的矩阵函数

[注] ① 拓宽定义不要求$f(z)$能展为“z”的幂级数，但要求在A 的特征值 λ_i （重数为m_i）处有 m_i-1 阶导数,后者较前者弱！

② 当能够展为“z”的幂级数时, 矩阵函数的拓宽定义与级数原始定义是一致的.

例9：$A=\begin{bmatrix}2&1&&\\&2&1&\\&&2&1\\&&&2\end{bmatrix}, f(z)=\dfrac{1}{z}$，求 $f(A)$

解：$f(z)=\dfrac{1}{z}, f'(z)=-z^{-2}, f''(z)=2z^{-3}, f'''(z)=-6z^{-4}$

$$f(A)=f(J)$$

$$=f(2)\cdot I+f'(2)\cdot I^{(1)}+\frac{f''(2)}{2!}\cdot I^{(2)}+\frac{f'''(2)}{3!}\cdot I^{(3)}$$

$$=\begin{bmatrix}0.5&-0.25&0.125&-0.0625\\&0.5&-0.25&0.125\\&&0.5&-0.25\\&&&0.5\end{bmatrix}$$

例10：$A=\begin{bmatrix}1&1&0\\0&1&0\\0&0&2\end{bmatrix}$，$f(z)=\sqrt{z}$ ，求 $f(A)$

解：$f(z)=\sqrt{z},\ f'(z)=\dfrac{1}{2\sqrt{z}}$

$$J_1=\begin{bmatrix}1&1\\&1\end{bmatrix}:\quad f(J_1)=f(1)\cdot I+f'(J_1)\cdot I^{(1)}=\begin{bmatrix}1&1/2\\&1\end{bmatrix}$$

$$J_2=[2]:\quad f(J_2)=f(2)\cdot I=\left[\sqrt{2}\right]$$

$$f(A)=f(J)=\begin{bmatrix}f(J_1)&\\&f(J_2)\end{bmatrix}=\begin{bmatrix}1&1/2&0\\&1&0\\&&\sqrt{2}\end{bmatrix}$$

四、矩阵函数的性质

级数定义或拓宽定义给出的矩阵函数具有下列性质：

(1) $f(z)=f_1(z)+f_2(z) \Rightarrow f(A)=f_1(A)+f_2(A)$

$$f^{(l)}(\lambda_i)=f_1^{(l)}(\lambda_i)+f_2^{(l)}(\lambda_i)$$

$$\Rightarrow f^{(l)}(J_i)=f_1^{(l)}(J_i)+f_2^{(l)}(J_i)$$

$$f(A)=P\bullet\left\{\begin{bmatrix} f_1(J_1) & & \\ & \ddots & \\ & & f_1(J_s)\end{bmatrix}+\begin{bmatrix} f_2(J_1) & & \\ & \ddots & \\ & & f_2(J_s)\end{bmatrix}\right\}\bullet P^{-1}$$

$$=f_1(A)+f_2(A)$$

(2) $f(z) = f_1(z) \cdot f_2(z)$

$$\Rightarrow f(A) = f_1(A) \cdot f_2(A) = f_2(A) \cdot f_1(A)$$

$$f_1(J_i) \cdot f_2(J_i) = \left[f_1 \cdot I + f_1' \cdot I^{(1)} + \frac{f_1''}{2!} \cdot I^{(2)} + \cdots + \frac{f_1^{(m_i-1)}}{(m_i-1)!} \cdot I^{(m_i-1)} \right] \cdot$$

$$\left[f_2 \cdot I + \frac{f'}{1!} \cdot I^{(1)} + \frac{f_2''}{2!} \cdot I^{(2)} + \cdots + \frac{f_2^{(m_i-1)}}{(m_i-1)!} \cdot I^{(m_i-1)} \right]$$

$$= (f_1 f_2) \cdot I + \frac{f_1' f_2 + f_1 f_2'}{1!} \cdot I^{(1)} + \frac{f_1'' f_2 + 2 f_1' f' + f_1 f_2''}{2!} \cdot I^{(2)} + \cdots$$

$$= (f_1 f_2) \cdot I + \frac{(f_1 f_2)'}{1!} \cdot I^{(1)} + \frac{(f_1 f_2)''}{2!} \cdot I^{(2)} + \cdots + \frac{(f_1 f_2)^{(m_i-1)}}{(m_i-1)!} \cdot I^{(m_i-1)}$$

$$= f(J_i)$$

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

$$=P\begin{bmatrix} f_1(J_1) & & \\ & \ddots & \\ & & f_1(J_s) \end{bmatrix}P^{-1}\bullet P\begin{bmatrix} f_2(J_1) & & \\ & \ddots & \\ & & f_2(J_s) \end{bmatrix}P^{-1}$$

$$= f_1(A)\bullet f_2(A)$$

4、矩阵的微分和积分

定义: 如果矩阵 $A(t)=\left(a_{ij}(t)\right)_{m\times n}$,的每一个元素 $a_{ij}(t)$

是变量t 的可微函数，则$A(t)$关于t的导数(微商)定义为

$$\frac{dA(t)}{dt}=\left(a_{ij}'(t)\right)_{m\times n}，\quad 或者 \quad A'(t)=\left(a_{ij}'(t)\right)_{m\times n}$$

定理8：设 $A(t),B(t)$ 可导，则有

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

(2) $A_{m\times n}, f(t)$ 可导 $\frac{d}{dt}\left[f(t)A(t)\right]=f'(t)A(t)+f(t)A'(t)$

(3) $A_{m\times n}, A_{n\times l}$: $\frac{d}{dt}\left[A(t)B(t)\right]=A'(t)B(t)+A(t)B'(t)$

证明：（3）左 $=\frac{d}{dt}\left(\sum_k a_{ik}(t)b_{kj}(t)\right)_{m\times l}$

$$=\left(\sum_k a'_{ik}(t)b_{kj}(t)+\sum_k a_{ik}(t)b'_{kj}(t)\right)_{m\times l}$$

$$=\left(\sum_k a'_{ik}(t)b_{kj}(t)\right)_{m\times l}+\left(\sum_k a_{ik}(t)b'_{kj}(t)\right)_{m\times l}$$ =右

定理9：设 $A_{n\times n}$ 为数量矩阵，则有

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

（2） $$\frac{d}{dt}\cos(tA)=-A\bullet\sin(tA)=-\sin(tA)\bullet A$$

（3） $$\frac{d}{dt}\sin(tA)=A\bullet\cos(tA)=\cos(tA)\bullet A$$

证明：（1） $e^{tA}=\sum_{n=0}^{\infty}\frac{1}{n!}(tA)^n$ 绝对收敛

$$\left(e^{tA}\right)_{ij}=\delta_{ij}+\frac{t}{1!}(A)_{ij}+\frac{t^2}{2!}\left(A^2\right)_{ij}+\cdots+\frac{t^k}{k!}\left(A^k\right)_{ij}+\cdots$$ 绝对收敛

$$\frac{d}{dt}\left(e^{tA}\right)_{ij}=0+\left(A\right)_{ij}+\frac{t}{1!}\left(A^2\right)_{ij}+\cdots+\frac{t^{k-1}}{(k-1)!}\left(A^k\right)_{ij}+\cdots$$

绝对收敛

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

绝对收敛

$$=\begin{cases} A\left[I+\dfrac{t}{1!}A+\cdots+\dfrac{t^{k-1}}{(k-1)!}A^{k-1}+\cdots\right] & =Ae^{tA} \\ \left[I+\dfrac{t}{1!}A+\cdots+\dfrac{t^{k-1}}{(k-1)!}A^{k-1}+\cdots\right]A & =e^{tA}A \end{cases}$$

定义: 如果矩阵 $A(t)=\left(a_{ij}(t)\right)_{m\times n}$ 的每一个元素 $a_{ij}(t)$ 在 $[t_0,t]$ 上可积，称$A(t)$可积，记为

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

(1) $\int_{t_0}^{t}\left[A(\tau)+B(\tau)\right]d\tau=\int_{t_0}^{t}A(\tau)d\tau+\int_{t_0}^{t}B(\tau)d\tau$

(2) A为常数矩阵：$\int_{t_0}^{t}\left[A\bullet B(\tau)\right]d\tau=A\bullet\left[\int_{t_0}^{t}B(\tau)d\tau\right]$

B为常数矩阵：$\int_{t_0}^{t}\left[A(\tau)\bullet B\right]d\tau=\left[\int_{t_0}^{t}A(\tau)d\tau\right]\bullet B$

(3) 设 $a_{ij}(t)\in C[t_0,t_1],a\in[t_0,t_1]$ 则 $:\frac{d}{dt}\int_{a}^{t}A(\tau)d\tau=A(t)$

(4) 设 $a_{ij}'(t)\in C[t_0,t_1]$, 则 $:\int_{t_0}^{t_1}A'(\tau)d\tau=A(t_1)-A(t_0)$

其它微分概念

函数对矩阵的导数(包括向量)

定义:设$X=(\xi_{ij})_{m\times n}$,mn元函数

$$f(X)=f(\xi_{11},\xi_{12},\cdots,\xi_{1n},\cdots,\xi_{m\times n})$$

定义$f(X)$对矩阵X的导数为

$$\frac{df}{dX}=\left(\frac{\partial f}{\partial \xi_{ij}}\right)_{m\times n}=\begin{bmatrix} \frac{\partial f}{\partial \xi_{11}} & \cdots & \frac{\partial f}{\partial \xi_{1n}} \\ \vdots & & \vdots \\ \frac{\partial f}{\partial \xi_{m1}} & \cdots & \frac{\partial f}{\partial \xi_{mn}} \end{bmatrix}$$

例11: $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $f(x)=f(\xi_1,\xi_2,\cdots,\xi_n)$ $\quad \dfrac{df}{dx}=\begin{bmatrix}\dfrac{\partial f}{\partial \xi_1}\\ \vdots\\ \dfrac{\partial f}{\partial \xi_n}\end{bmatrix}$

例12: $A=(a_{ij})_{m\times n}$, $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $f(x)=x^TAx$, 求 $\dfrac{df}{dx}$

$$f(x)=\xi_1\sum_{j=1}^{n}a_{1j}\xi_j+\cdots+\xi_k\sum_{j=1}^{n}a_{kj}\xi_j+\cdots+\xi_n\sum_{j=1}^{n}a_{nj}\xi_j$$

$$\frac{\partial f}{\partial \xi_k}=\xi_1a_{1k}+\cdots+\xi_{k-1}a_{k-1,k}+\left(\sum_{j=1}^{n}a_{kj}\xi_j+\xi_k a_{kk}\right)$$

$$+\xi_{k+1}a_{k+1,k}+\cdots+\xi_n a_{nk}\quad=\sum_{j=1}^{n}a_{kj}\xi_j+\sum_{i=1}^{n}a_{ik}\xi_i$$

$$\therefore \frac{df}{dx} = \left(A + A^T\right) x$$

如果 $A = A^T$ ，有 $\frac{df}{dx} = 2Ax$

例13: $\boldsymbol{X}=\left(\xi_{ij}\right)_{m\times n}$: $f(\boldsymbol{X})=\left[\operatorname{tr}(\boldsymbol{X})\right]^2$ 求 $\left.\dfrac{df}{d\boldsymbol{X}}\right|_{\boldsymbol{X}=\boldsymbol{I}_n}$

解：

$$f(\boldsymbol{X})=\left(\xi_{11}+\xi_{22}+\cdots+\xi_{nn}\right)^2$$

$$\frac{df}{d\boldsymbol{X}}=2\left(\xi_{11}+\xi_{22}+\cdots+\xi_{nn}\right)\boldsymbol{I}_n$$

$$\left.\frac{df}{d\boldsymbol{X}}\right|_{\boldsymbol{X}=\boldsymbol{I}_n}=2n\boldsymbol{I}_n$$

例14: $A \in R^{m\times n}, b \in R^m$，若 $x \in R^n$ 使得 $\|Ax-b\|_2 = \min$,

则 $A^T A x = A^T b$

解： $f(x) = \|Ax-b\|_2^2 = (Ax-b)^T (Ax-b)$

$$= x^T A^T A x - 2b^T A x + b^T b$$

$$g(x) = b^T A x = b_1 \sum_{j=1}^{n} a_{1j}\xi_j + \cdots + b_m \sum_{j=1}^{n} a_{mj}\xi_j$$

$$\frac{dg}{dx} = \begin{bmatrix} \dfrac{\partial g}{\partial \xi_1} \\ \vdots \\ \dfrac{\partial g}{\partial \xi_n} \end{bmatrix} = \begin{bmatrix} b_1 a_{11} + \cdots + b_m a_{m1} \\ \vdots \\ b_1 a_{1n} + \cdots + b_m a_{mn} \end{bmatrix} = A^T b$$

$$\frac{df}{dx} = 2A^T A x - 2A^T b = 0 \quad \Rightarrow A^T A x = A^T b$$

【注】 $r(A^T A) = r(A) \Rightarrow r(A^T A | A^T b) = r(A^T A) \Rightarrow A^T A x = A^T b$ 有解

5、函数矩阵对矩阵的导数

定义:设 $X=\left(\xi_{ij}\right)_{m\times n}$, $f_{kl}(X)=f_{kl}\left(\xi_{11},\xi_{12},\cdots,\xi_{1n},\cdots,\xi_{m\times n}\right)$

$$F=\begin{bmatrix} f_{11} & \cdots & f_{1s} \\ \vdots & & \vdots \\ f_{r1} & \cdots & f_{rs} \end{bmatrix}, \qquad \frac{\partial F}{\partial \xi_{ij}}=\begin{bmatrix} \frac{\partial f_{11}}{\partial \xi_{ij}} & \cdots & \frac{\partial f_{1s}}{\partial \xi_{ij}} \\ \vdots & & \vdots \\ \frac{\partial f_{r1}}{\partial \xi_{ij}} & \cdots & \frac{\partial f_{rs}}{\partial \xi_{ij}} \end{bmatrix},$$

定义 $$\frac{dF}{dX}=\begin{bmatrix} \frac{\partial F}{\partial \xi_{11}} & \cdots & \frac{\partial F}{\partial \xi_{1n}} \\ \vdots & & \vdots \\ \frac{\partial F}{\partial \xi_{m1}} & \cdots & \frac{\partial F}{\partial \xi_{mn}} \end{bmatrix}$$

- 可表示为

$$\frac{dF}{dX}=\left(\frac{1}{dX}\right)\otimes dF$$

例15: $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$, $F(x)=\left[f_1(x), f_2(x), \cdots, f_l(x)\right]$

$$\frac{dF}{dx}=\begin{bmatrix}\frac{\partial f_1}{\partial \xi_1} & \cdots & \frac{\partial f_l}{\partial \xi_1}\\ \vdots & & \vdots\\ \frac{\partial f_1}{\partial \xi_n} & \cdots & \frac{\partial f_l}{\partial \xi_n}\end{bmatrix}$$

例16: $A=(a_{ij})_{n\times n}$, $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$ $Ax=\begin{bmatrix}\sum_{j=1}^{n} a_{1j}\xi_j\\ \vdots\\ \sum_{j=1}^{n} a_{nj}\xi_j\end{bmatrix}$

$$\frac{d(Ax)}{dx^T}=\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{bmatrix}=A$$

作业

- **P163：1、2、5、6**
- **P170：4、5、6**