Ch7 矩阵分析初步

§1. 赋范线性空间

Def. 设 $F=\mathbb{R}$ or $\mathbb{C}$, V 为 F 上之一个线性空间,

若 V 上之实函数 $\|\cdot\|$, 满足

(1) 对 $\forall x\in V$, $\|x\|\geq 0$ 且 $\|x\|=0 \Leftrightarrow x=0$ 正定性

(2) $\forall k\in F$, $x\in V$, $\|kx\|=|k|\cdot\|x\|$ 齐次性

(3) $\forall x,y\in V$, $\|x+y\|\leq\|x\|+\|y\|$. 三角不等式

称 V 是 $\boxed{\text{赋范线性空间}}$, $\|x\|$ 是 x 之 $\boxed{\text{范数}}$

常用范数:

① $\|x\|_\infty = \max\limits_{1\le j\le n} |x_j|$, l_∞ or 最大范数

② $\|x\|_1 = \sum\limits_{j=1}^{n} |x_j|$ l_1 or 和范数

③ $\|x\|_2 = \left(\sum\limits_{j=1}^{n} |x_j|^2\right)^{1/2}$ l_2 or 欧几里德范数

④ $\|x\|_p = \left(\sum\limits_{j=1}^{n} |x_j|^p\right)^{1/p}$ l_p or Hölder 范数

例: $x = \begin{pmatrix} -6 \\ 1 \\ 3 \\ -2 \end{pmatrix}$.

$\|x\|_\infty = 6.$

$\|x\|_1 = 12$

$\|x\|_2 = \sqrt{6^2+1+3^2+2^2} = \sqrt{50}$

$\|x\|_p = \cdots$

例. 设 $\|\cdot\|_\alpha$ 为 F^m 上之一种向量范数.

给定 $A\in F^{m\times n}$, $rank(A)=n$ (列满秩).

对 $\forall x\in F^n$, 定义 $\|x\|_\beta=\|Ax\|_\alpha$

则 $\|\cdot\|_\beta$ 为 F^n 的范数.

证: (1) $\forall x\in F^n$, $Ax=\sum_{i=1}^{n}x_iA_i$

故 $\|x\|_\beta=\|Ax\|_\alpha\geq 0$ 且 $\|x\|_\beta=0 \Leftrightarrow Ax=0 \Leftrightarrow x=0$

(2) $\forall k$, $\|kx\|_\beta=\|kAx\|_\alpha=|k|\,\|Ax\|_\alpha=|k|\cdot\|x\|_\beta$.

(3) $\|x+y\|_\beta=\|A(x+y)\|_\alpha\leq\|Ax\|_\alpha+\|Ay\|_\alpha=\|x\|_\beta+\|y\|_\beta$.

Def. 称 $\|x\|_\alpha$ 与 $\|x\|_\beta$ 等价，

如果 ∃ 正常数 $c_1 < c_2$，

使：$c_1\|x\|_\alpha \le \|x\|_\beta \le c_2\|x\|_\alpha$.

Th. 有限维线性空间中的任何两种

$\boxed{\text{向量范数都等价.}}$

§2. 矩阵范数.

Def. $\|\cdot\|$ 为 $F^{n\times n}$ 上非负实函数. 若对 $\forall A, B\in F^{n\times n}$,
有 (1) $\|A\|\geq 0$ 且 $\|A\|=0 \Longleftrightarrow A=0$

(2) $\forall k\in F$, $\|kA\|=|k|\cdot\|A\|$.

(3) $\|A+B\|\leq\|A\|+\|B\|$

(4) $\|AB\|\leq\|A\|\cdot\|B\|$

称 $\|\cdot\|$ 为 $F^{n\times n}$ 上之一个范数.

常用矩阵范数

① $\|A\|_F = \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{\frac{1}{2}} = \sqrt{tr(A^*A)}$ Frobinues 范数

② $\|A\|_\infty = \max\limits_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$ 行范数（行模和最大值）

③ $\|A\|_1 = \max\limits_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$ 列范数（列模和最大值）

④ $\|A\|_2 = \left(\rho(A^*A)\right)^{\frac{1}{2}}$ 谱范数

⑤ $\|A\| = \sup\limits_{0\ne x\in F^n} \dfrac{\|Ax\|_\alpha}{\|x\|_\alpha}$ 向量范数 $\|\cdot\|_\alpha$ 导出矩阵或算子范数

Def. 矩阵范数 $\|\cdot\|$ 与向量范数 $\|\cdot\|_\alpha$ 相容：

如果对 $\forall A \in F^{n\times n}$, $x \in F^n$, $\|Ax\|_\alpha \le \|A\| \cdot \|x\|_\alpha$.

Th. ① 对于 $F^{n\times n}$ 上的每种矩阵范数，

都存在 F^n 上的与它相容的向量范数

② $F^{n\times n}$ 上的任意两种矩阵范数等价.

即，$\exists\, c_1 > c_2 > 0$，$\ni$ $c_2\|A\|_\beta \le \|A\|_\alpha \le c_1\|A\|_\beta$.

§3. 向量和矩阵序列

(一) 基本概念

Def. 设 $x^{(k)} = (x_1^{(k)}, \cdots, x_n^{(k)}) \in F^n$, $k=1,2,\cdots$

若 $\lim\limits_{k\to\infty} x_i^{(k)} = x_i$, $i = 1\sim n$.

则称 $\{x^{(k)}\}$ $\boxed{\text{收敛}}$ 于 $x = (x_1, \cdots, x_n)$,

记为 $\lim\limits_{k\to\infty} x^{(k)} = x$ or $x^{(k)} \to x$.

若 $\{x^{(k)}\}$ 不收敛, 称 $\boxed{\text{发散}}$.

Th. $\lim\limits_{k\to\infty} x^{(k)} = x \iff \lim\limits_{k\to\infty} \|x^{(k)} - x\| = 0$

其中 $\|\cdot\|$ 为 F^n 中任一范数

Def. $A_k = (a_{ij}^{(k)})_{n\times n} \in F^{n\times n}$. 如果 $\lim\limits_{k\to\infty} a_{ij}^{(k)} = a_{ij}$,

称 $\lim\limits_{k\to\infty} A_k = A$.

Th. $\lim\limits_{k\to\infty} A_k = A \iff \lim\limits_{k\to\infty} \|A_k - A\| = 0$

坐标收敛 $\iff$ 范数收敛

(二) 运算性质.

① $\lim\limits_{k\to\infty} A_k = A$, $\lim\limits_{k\to\infty} B_k = B$, $\lim\limits_{k\to\infty} a_k = a$, $\lim\limits_{k\to\infty} b_k = b$,

则 $\lim\limits_{k\to\infty} a_k A_k + b_k B_k = aA + bB$.

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

② $\lim\limits_{k\to\infty} A_k = A$, 则对 $\forall \|\cdot\|$, $\|A_k\|$ 有界.

③ $\lim\limits_{k\to\infty} A_k = A$, P 可逆, 则 $\lim\limits_{k\to\infty} P^{-1} A_k P = P^{-1} A P$

④ $\lim\limits_{k\to\infty} A_k = A$. A_k^{-1} 及 A^{-1} 存在. 则 $\lim\limits_{k\to\infty} A_k^{-1} = A^{-1}$.

(三) 幂收敛

Def. 若 $E, A, A^2, \cdots$ 收敛，称 A $\boxed{\text{幂收敛}}$

设 $A \sim J_A = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix}$，其中 $J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$，

i.e. ∃可逆阵 P，s.t. $P^{-1}AP = J_A$，$A = PJ_AP^{-1}$

故 $A^m = PJ_A^m P^{-1}$，$J_A^m = \begin{pmatrix} J_1^m & & \\ & \ddots & \\ & & J_s^m \end{pmatrix}$.

Th. $\{A^m\}$ 收敛 $\Longleftrightarrow$ $\{J_A^m\}$ 收敛 $\Longleftrightarrow$ $\{J_i^m\}$ 收敛

$\Longleftrightarrow$ $\boxed{|\lambda| \le 1 \text{ 且若 } |\lambda| = 1 \text{ 则 } \lambda = 1 \text{ 且其对应之 Jordan 块均为1阶}}$

$$J_i^m = (\lambda_i I - N_i)^m = \sum_{k=1}^{m \to n_i - 1} C_m^k \lambda_i^{m-k} N_i^k$$

$$= \begin{pmatrix} \lambda_i^m & C_m^1\lambda_i^{m-1} & C_m^2\lambda_i^{m-2} & \cdots & C_m^{n_i-1}\lambda_i^{m-n_i+1} \\ & \lambda_i^m & C_m^1\lambda_i^{m-1} & \cdots & C_m^{n_i-2}\lambda_i^{m-n_i+2} \\ & & \lambda_i^m & \ddots & \vdots \\ & & & \ddots & \\ & & & & \lambda_i^m \end{pmatrix}_{n_i\times n_i}$$

其中 $C_m^k = \dfrac{m!}{k!(m-k)!}$ $(k\le m)$

$C_m^k = 0$ $(k > m)$

① 如果 J_i^m 收敛 $\Rightarrow |\lambda_i|\le 1$，且若 $|\lambda_i| = 1$，则由 $|C_m^1\lambda_i^{m-1}| = m \Rightarrow \begin{cases}\lambda_i = 1 \\ n_i = 1\end{cases}$

否则 $\{C_m^1\lambda_i^m\}$ 发散

② 若 $|\lambda_i| < 1$，则由 $\lim\limits_{m\to\infty}\left|\dfrac{C_{m+1}^k\lambda_i^{m+1-k}}{C_m^k\lambda_i^{m-k}}\right| = \lim\limits_{m\to\infty}\left|\dfrac{m+1}{m+1-k}\right|\cdot|\lambda_i| = |\lambda_i|$

$\Rightarrow \sum\limits_{m=1}^{\infty} C_m^k\lambda_i^{m-k}$ 收敛 $\Rightarrow \lim\limits_{m=\infty} C_m^k\lambda_i^{m-k} = 0$

推论. 设 $S_m = \sum_{k=0}^{m} A^k$

则 $\boxed{\{S_m\} \text{ 收敛} \iff \lim_{k\to\infty} A^k = 0}$

证: "$\Longrightarrow$" $\lim_{k\to\infty} A^k = \lim_{k\to\infty} (S_k - S_{k-1}) = 0$

"$\Longleftarrow$" $\lim_{k\to\infty} A^k = 0 \Longrightarrow |\lambda_i| < 1$

$\Longrightarrow \sum J_i^k$ 收敛

$\Longrightarrow \sum J^k$ 收敛

$\Longrightarrow \sum A^k$ 收敛.

Th. $A\in\mathbb{C}^{n\times n}$，$\lambda$为$A$之特征值，$\|\cdot\|$为$A$之任一范数，

则 $\boxed{|\lambda|\le\|A\|.}$

证：令 $B=\frac{1}{\|A\|+\varepsilon}A$，其中$\varepsilon$为任意之实数.

$$\|B\|=\frac{\|A\|}{\|A\|+\varepsilon}<1 \Longrightarrow \|B^k\|\le\|B\|^k\to 0$$

$$\Longrightarrow \lim_{k\to\infty}B^k=0$$

$$\Longrightarrow \frac{|\lambda|}{\|A\|+\varepsilon}<1$$

$$\Longrightarrow |\lambda|<\|A\|+\varepsilon \Longrightarrow |\lambda|\le\|A\|.$$

推论 (Neumann 引理)

设 $\|A\|<1$，则 A 幂收敛，$I-A$ 可逆，

且 $(I-A)^{-1}=I+A+A^2+\cdots=\sum_{k=0}^{\infty}A^k$.

证. $\|A\|<1 \Rightarrow |\lambda|<1 \Rightarrow A$ 幂收敛

$\Rightarrow I-A$ 无零特征值

$\Rightarrow I-A$ 可逆.

$$(I-A)(I+A+A^2+\cdots)=I+A+A^2+\cdots+A^k+\cdots-A-A^2-\cdots-A^k\cdots$$

$$=I$$

§4. 矩阵幂级数

一、矩阵级数

def. $\{A_k\}$, $S_n=\sum_{k=1}^{n}A_k$, 若 $\lim_{n\to\infty}S_n=S$,

称矩阵级数 $\boxed{\sum_{k=1}^{\infty}A_k \text{ 收敛于 } S,}$ 记为 $\sum_{k=1}^{\infty}A_k=S$

否则称 $\sum_{k=1}^{\infty}A_k$ 发散.

性质. ① 若 $\sum_{k=1}^{\infty}A_k$ 收敛 $\Longrightarrow \lim_{k\to\infty}A_k=0$

② 若 $\sum_{k=1}^{\infty}A_k=A$, $\sum_{k=1}^{\infty}B_k=B$

$\Rightarrow \sum_{k=1}^{\infty}(A_k+B_k)=A+B$, $\sum_{k=1}^{\infty}\alpha A_k=\alpha A$

二. 幂级数 $\sum_{k=0}^{\infty}\alpha_k A^k$ 的收敛问题

引理. 设 $J=\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{pmatrix}$,

$f(t)=\sum_{k=0}^{\infty}a_k t^k$ 为收敛半径为 r 的幂级数

则当 $\boxed{|\lambda|<r \text{ 时}, \sum_{k=0}^{\infty}a_k J^k \text{ 收敛}}$, 且和矩阵为:

$$\begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & f(\lambda) & f'(\lambda) & \cdots & \frac{f^{(n-2)}(\lambda)}{(n-2)!} \\ & & f(\lambda) & & \vdots \\ & & & \ddots & f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix}.$$

收敛半径:

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

or $R=\lim_{n\to\infty}\frac{1}{\sqrt[n]{|a_n|}}$

证：记 $S_m(\lambda)=\sum_{k=0}^{m} a_k\lambda^k$.

$$S_m=\sum_{k=0}^{m} a_k J^k,\qquad J^k=\begin{pmatrix} \lambda^k & C_k^1\lambda^{k-1} & \cdots & C_k^{n-1}\lambda^{k-n+1} \\ & \lambda^k & \ddots & \vdots \\ & & \ddots & C_k^1\lambda^{k-1} \\ & & & \lambda^k \end{pmatrix}_{n\times n}$$

则

$$S_m=\begin{pmatrix} \sum_{k=0}^{m} a_k\lambda^k & \sum_{k=0}^{m} a_k C_k^1\lambda^{k-1} & \cdots & \sum_{k=0}^{m} a_k C_k^{n-1}\lambda^{k-n+1} \\ & \sum_{k=0}^{m} a_k\lambda^k & \sum_{k=0}^{m} a_k C_k^1\lambda^{k-1}\ \cdots & \vdots \\ & & \ddots & \sum_{k=0}^{m} a_k C_k^1\lambda^{k-1} \\ & & & \sum_{k=0}^{m} a_k\lambda^k \end{pmatrix}_{n\times n}$$

$$=\begin{pmatrix} \overset{\to f(\lambda)}{S_m(\lambda)}, & \overset{\to f'(\lambda)}{S_m'(\lambda)}, & \overset{\to \frac{f''(\lambda)}{2!}}{\frac{S_m''(\lambda)}{2!}}, & \cdots, & \overset{\to \frac{f^{(n-1)}(\lambda)}{(n-1)!}}{\frac{S_m^{(n-1)}(\lambda)}{(n-1)!}} \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & S_m'(\lambda) \\ & & & & S_m(\lambda) \end{pmatrix}$$

$S_m(t)=\sum_{k=0}^{m} a_k t^k$ 的收敛半径为 r.

$|\lambda|<r \Rightarrow \underbrace{S_m(\lambda)}_{\to f(\lambda)},\cdots,\underbrace{S_m^{(n-1)}(\lambda)}_{\to f^{(n-1)}(\lambda)}$ 都收敛

例. 设 $f(t) = \sum_{k=0}^{\infty} \left(\frac{t}{4}\right)^k$. 求 $f(J)$, 其中 $J = \begin{pmatrix} 3 & 1 & & \\ & 3 & 1 & \\ & & 3 & 1 \\ & & & 3 \end{pmatrix}$

解. $f(t) = \left(1 - \frac{t}{4}\right)^{-1}$, 收敛半径为 4.

故 $f(J) = \sum_{k=0}^{\infty} \left(\frac{J}{4}\right)^k$ 收敛,

且和为: $f(J) = \begin{pmatrix} f(3), & f'(3), & \frac{f''(3)}{2!} & \frac{f'''(3)}{3!} \\ & f(3) & f'(3) & \frac{f''(3)}{2!} \\ & & f(3) & f'(3) \\ & & & f(3) \end{pmatrix}$

$f'(t) = \frac{1}{4}\left(1 - \frac{t}{4}\right)^{-2}$, $f''(t) = \frac{1}{8}\left(1 - \frac{t}{4}\right)^{-3}$. $f'''(t) = \frac{3}{32}\left(1 - \frac{t}{4}\right)^{-4}$;

$\Rightarrow f(3) = 4$, $f'(3) = 4$, $f''(3) = 8$, $f'''(3) = 24$

$\Rightarrow f(J) = \begin{pmatrix} 4 & 4 & 4 & 4 \\ & 4 & 4 & 4 \\ & & & \\ & & & \end{pmatrix}$

Th (Lagrange-Sylvester 定理)

设 $f(t)=\sum\limits_{k=0}^{\infty} a_k t^k$，收敛半径为 r.

A 之 Jordan 标准形 $J_A=\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix}$，$J_i=\begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$ $i=1\sim s$,

变换阵为 P，$A=PJ_AP^{-1}$.

若对所有 $i=1\sim s$，$|\lambda_i|<r$，则 $\sum\limits_{k=0}^{\infty} a_k A^k$ 收敛

且和为：$\boxed{f(A)=\sum\limits_{k=0}^{\infty} a_k A^k = P\begin{pmatrix} f(J_1) & & \\ & \ddots & \\ & & f(J_s) \end{pmatrix}P^{-1}}$

其中 $f(J_i)=\begin{pmatrix} f(\lambda_i), & f'(\lambda_i), & \cdots, & \frac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!} \\ & f(\lambda_i) & \cdots & \vdots \\ & & \ddots & \\ & & & f(\lambda_i) \end{pmatrix}$，$i=1\sim s$.

例. 已知 $f(t)=2-t+2t^3$. 求 $f(A)$. 其中 $A=\begin{pmatrix} 1 & 1 & & & & \\ & 1 & 1 & & & \\ & & 1 & & & \\ & & & 2 & 1 & \\ & & & & 2 & 1 \\ & & & & & 2 \end{pmatrix}$

解. 可将 $f(t)$ 看作幂级数 $f(t)=\sum_{k=0}^{\infty} a_k t^k$,

其中 $a_0=2$, $a_1=-1$, $a_2=2$, $a_k=0$ $(k\geq 3)$

$f(1)=3$, $f'(1)=5$, $f''(1)=12$,

$f(2)=16$, $f'(2)=23$ $f''(2)=24$.

故 $f(A)=\begin{pmatrix} 3 & 5 & 6 & & & \\ & 3 & 5 & & & \\ & & 3 & & & \\ & & & 16 & 23 & 12 \\ & & & & 16 & 23 \\ & & & & & 16 \end{pmatrix}$

§5. 矩阵函数.

一. 基本函数及性质.

$e^z=\sum\limits_{k=0}^{\infty}\frac{z^k}{k!}$, $\sin z=\sum\limits_{k=0}^{\infty}(-1)^k\frac{z^{2k+1}}{(2k+1)!}$, $\cos z=\sum\limits_{k=0}^{\infty}(-1)^k\cdot\frac{z^{2k}}{(2k)!}$

在整个复平面上收敛.

故 $\sum\limits_{k=0}^{\infty}\frac{A^k}{k!}$, $\sum\limits_{k=0}^{\infty}(-1)^k\frac{A^{2k+1}}{(2k+1)!}$, $\sum\limits_{k=0}^{\infty}(-1)^k\frac{A^{2k}}{(2k)!}$ 收敛

记为 e^A, $\sin A$, $\cos A$.

类似: $\ln(1+z)=\sum\limits_{k=1}^{\infty}(-1)^{k-1}\frac{z^k}{k}$, $|z|<1$, 定义: $\mathrm{Ln}(E+A)=\sum\limits_{k=1}^{\infty}(-1)^{k-1}\frac{A^k}{k}$.

$\rho(A)<1$

$(1+z)^\alpha=\sum\limits_{k=0}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}z^k$. $|z|<1$, $(E+A)^\alpha=\sum\limits_{k=0}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}A^k$

性质：$e^{\lambda E}=e^{\lambda}E$，$\sin(\lambda E)=(\sin\lambda)E$，$\cos(\lambda E)=(\cos\lambda)E$

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

解. $|\lambda E-A|=(\lambda-1)(\lambda-2)$，$\lambda_1=1$，$\alpha_1=\begin{pmatrix}1\\0\end{pmatrix}$

$\lambda_2=2$，$\alpha_2=\begin{pmatrix}1\\1\end{pmatrix}$

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

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

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

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

性质：① 设 $\boxed{A\text{ 特征值为 }\lambda_1,\cdots,\lambda_n}$

则 $\boxed{e^A \quad \cdots \quad \cdots \quad e^{\lambda_1},\cdots,e^{\lambda_n}}$

$\sin A \quad \cdots \quad \cdots \quad \sin\lambda_1,\cdots,\sin\lambda_n$

$\cos A \quad \cdots \quad \cdots \quad \cos\lambda_1,\cdots,\cos\lambda_n$

② e^A 的性质：

Ⅰ 若 $AB=AB \Rightarrow e^A\cdot e^B=e^{A+B}$

Ⅱ $(e^A)^{-1}=e^{-A}$

Ⅲ $\det A = e^{\operatorname{tr} A}$

证：$AB=BA \Rightarrow (A+B)^m=\sum\limits_{k=0}^{m} C_m^k A^{m-k}B^k$

$$\Rightarrow e^{A+B}=\sum_{m=0}^{\infty}\frac{(A+B)^m}{m!}=\sum_{m=0}^{\infty}\frac{1}{m!}\sum_{k=0}^{m}C_m^k A^{m-k}B^k=\sum_{m=0}^{\infty}\sum_{k=0}^{m}\frac{A^{m-k}B^k}{k!(m-k)!}=\sum_{k=0}^{\infty}\sum_{m=k}^{\infty}\frac{B^k}{k!}\frac{A^{m-k}}{(m-k)!}=\sum_{k=0}^{\infty}\frac{A^k}{k!}\sum_{j=0}^{\infty}\frac{B^j}{j!}=e^A\cdot e^B$$

③ $e^{iA} = \cos A + i\sin A$

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

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

④ 若 $AB = BA$，则

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

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

Th. 设可对角化阵 A 之谱分解 $A = \sum_{i=1}^{s} \lambda_i P_i$，幂级数 $f(t)$ 之收敛半径为 $r > \rho(A)$，则 $f(A) = \sum_{i=1}^{s} f(\lambda_i) P_i$.

二. 函数矩阵的微分和积分

Def. 函数矩阵 $A(t)=(a_{ij}(t))_{m\times n}$.

…… 极限: $\lim\limits_{t\to t_0} a_{ij}(t)=a_{ij}$, $i=1\sim m$, $j=1\sim n$.

称 $A(t)$ 在 t_0 处之极限为 $A=(a_{ij})_{m\times n}$

记为: $\lim\limits_{t\to t_0} A(t)=A$.

类似可定义 $A(t)$ 在某点之连续性, 可微, 可积

性质: 设 $\lim\limits_{t\to t_0} A(t)=A$, $\lim\limits_{t\to t_0} B(t)=B$

① $\lim\limits_{t\to t_0}(A(t)+B(t))=A+B$, ($A(t), B(t)$ 可加)

② $\lim\limits_{t\to t_0} A(t)B(t)=AB$, ($A(t), B(t)$ 可乘)

性质: $A'(t) = (a_{ij}'(t))_{m\times n}$. $\int_a^b A(t)\,dt = \left(\int_a^b a_{ij}(t)\,dt\right)_{m\times n}$.

① $(aA(t) + bB(t))' = aA'(t) + bB'(t)$

② $(A(t)B(t))' = A'(t)B(t) + A(t)B'(t)$

③ $\left(\int_a^t A(s)\,ds\right)' = A(t)$

④ $\int_a^t A'(s)\,ds = A(t) - A(a)$

⑤ $\int_a^t BA(s)\,ds = B\int_a^t A(s)\,ds$. 其中B为常数阵

例: $x(t) = (x_1(t), \cdots, x_n(t))^T$. $A(t) = (a_{ij}(t))_{n\times n}$ ~~对称~~ 求 $x(t)^T A(t) x(t)$ 导数

解: $(x^TAx)' = (x^T)'Ax + x^T(Ax)' = (x^T)'Ax + x^TA'x + \underbrace{x^TAx'}_{(x^TAx')^T = (x^T)'Ax}$

$= 2x^TAx' + x^TA'x$

§4. 矩阵函数的计算

Th. 设 $J=\begin{pmatrix}\lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{pmatrix}_{s\times s}$ 则 $\boxed{e^{Jt}=e^{\lambda t}\begin{pmatrix}1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{s-1}}{(s-1)!} \\ & \ddots & & & \vdots \\ & & \ddots & & \\ & & & \ddots & t \\ & & & & 1\end{pmatrix}}$

证: $J=\lambda E+N$, 其中 $N=\begin{pmatrix}0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0\end{pmatrix}_{s\times s}$

$$e^{Jt}=e^{\lambda tE+Nt}=e^{\lambda t}\cdot e^{tN}$$

$$=e^{\lambda t}\sum_{k=0}^{\infty}\frac{(tN)^k}{k!}=e^{\lambda t}\cdot\sum_{k=0}^{s-1}\frac{t^kN^k}{k!}$$

$$=e^{\lambda t}\begin{pmatrix}1 & t & \cdots & \frac{t^{s-1}}{(s-1)!} \\ & 1 & t & \ddots \\ & & \ddots & t \\ & & & 1\end{pmatrix}$$

Th. 设 $A=PJ_AP^{-1}$, $J_A=\begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix}$,

则 $e^{At}=P\begin{pmatrix} e^{J_1t} & & \\ & \ddots & \\ & & e^{J_st} \end{pmatrix}P^{-1}$.

证: $e^{At}=\sum_{k=0}^{\infty}\frac{(At)^k}{k!}=\sum_{k=0}^{\infty}\frac{t^k}{k!}(PJ_AP^{-1})^k$

$=P\sum_{k=0}^{\infty}\frac{t^k}{k!}J_A^k\,P^{-1}=Pe^{J_At}P^{-1}$

$=P\begin{pmatrix} e^{J_1t} & & \\ & \ddots & \\ & & e^{J_st} \end{pmatrix}P^{-1}$

例. 求 e^{Jt}，其中 $J=\begin{pmatrix} J_1 & \\ & J_2 \end{pmatrix}$，$J_1=\begin{pmatrix} -2 & 1 & \\ & -2 & 1 \\ & & -2 \end{pmatrix}$，$J_2=\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}$

解. $e^{J_1 t}=e^{-2t}\begin{pmatrix} 1 & t & \frac{t^2}{2} \\ & 1 & t \\ & & 1 \end{pmatrix}$，$e^{J_2 t}=e^{t}\begin{pmatrix} 1 & t \\ & 1 \end{pmatrix}$

$$e^{Jt}=\begin{pmatrix} e^{J_1 t} & \\ & e^{J_2 t} \end{pmatrix}=\begin{pmatrix} e^{-2t} & te^{-2t} & \frac{t^2}{2}e^{-2t} & & \\ & e^{-2t} & te^{-2t} & & \\ & & e^{-2t} & & \\ & & & e^{t} & te^{t} \\ & & & & e^{t} \end{pmatrix}$$

二. 一般矩阵函数计算

设 $m_A(\lambda)=\lambda^m+a_1\lambda^{m-1}+\cdots+a_{m-1}\lambda+a_m$,

则 A 的任何次幂可由 $E, A, \cdots, A^{m-1}$ 线性表出.

故由矩阵幂级数定义的矩阵函数 $f(A)$ 可由不超 m 次的多项式表示.

Def. 设 $m_A(\lambda)=(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_s)^{k_s}$, 其中 $\lambda_1 \cdots \lambda_s$ 两两不同.

对 $f(t)=\sum_{k=0}^{\infty}a_k t^k$, 称

$$\left\{\boxed{\begin{matrix} f(\lambda_1), & f'(\lambda_1), & \cdots, & f^{(k_1-1)}(\lambda_1) \\ \vdots & & & \\ f(\lambda_s), & f'(\lambda_s), & \cdots, & f^{(k_s-1)}(\lambda_s) \end{matrix}}\right.$$

为 $\boxed{f(t)\text{ 在 }A\text{ 之谱上之函数值}}$.

Th. 设 $f(t)=\sum_{k=0}^{\infty} a_k t^k$, $g(t)=\sum_{k=0}^{\infty} b_k t^k$. 它们收敛半径大于 $\rho(A)$,

则 $f(A)=g(A) \iff f(t)$ 与 $g(t)$ 在 A 的谱上的值相同.

即. $f(\lambda_i)=g(\lambda_i), \cdots, f^{(k_i-1)}(\lambda_i)=g^{(k_i-1)}(\lambda_i)$
$i=1\sim s$

证明. $A=P\begin{pmatrix}J_1&&\\&\ddots&\\&&J_t\end{pmatrix}P^{-1}$.

$$f(A)=P\begin{pmatrix}f(J_1)&&\\&\ddots&\\&&f(J_t)\end{pmatrix}P^{-1}, \quad g(A)=P\begin{pmatrix}g(J_1)&&\\&\ddots&\\&&g(J_t)\end{pmatrix}P^{-1}$$

$$f(A)=g(A) \iff f(J_i)=g(J_i),\ i=1\sim t.$$

$$f(J_i)=\begin{pmatrix}f(\lambda_i)&f'(\lambda_i)&\cdots&\frac{f^{(k_i-1)}(\lambda_i)}{(k_i-1)!}\\&f(\lambda_i)&\ddots&\vdots\\&&\ddots&\\&&&f(\lambda_i)\end{pmatrix},$$

对 $\forall \lambda_i$, 对应的 Jordan 块 $J_{i_1},\cdots,J_{i_p}$. 且 J_{i_1} 的阶数最大. 则

$f(J_{i_1})=g(J_{i_1}) \Rightarrow f(J_{i_2})=g(J_{i_2})$

......

例. 已知 $m_A(\lambda)=\lambda^2(\lambda-\pi)(\lambda+\pi)$,

证明: $\sin A = A-\frac{1}{\pi^2}A^3$.

证. $\boxed{设\ f(t)=\sin t,\quad g(t)=t-\frac{1}{\pi^2}t^3.}$

直接计算:
$$f(0)=\sin(0)=0=g(0).$$
$$f'(0)=\cos 0=1=g'(0)$$
$$f(\pi)=\sin\pi=0=g(\pi)$$
$$f(-\pi)=\sin(-\pi)=0=g(-\pi)$$

从而 $\sin A = A-\frac{1}{\pi^2}A^3$

(1) 待定系数法.

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

法一. Lagrarge-Sylvester 定理

法二. $f_A(\lambda)=(\lambda-1)(\lambda-2)^2$, 直接验证 $m_A(\lambda)=(\lambda-1)(\lambda-2)^2$.

令 $f(\lambda)=e^{\lambda t}$, $g(\lambda)=a_0+a_1\lambda+a_2\lambda^2$, 其中 a_0, a_1, a_2 待定

$$\left.\begin{aligned} f(1)=g(1) &\Rightarrow e^t=a_0+a_1+a_2 \\ f(2)=g(2) &\Rightarrow e^{2t}=a_0+2a_1+4a_2 \\ f'(2)=g'(2) &\Rightarrow te^{2t}=a_1+4a_2 \end{aligned}\right\} \Rightarrow \begin{cases} a_0=4e^t-2e^{2t}+2te^{2t} \\ a_1=-4e^t+4e^{2t}-3te^{2t} \\ a_2=e^t-2e^{2t}+te^{2t} \end{cases}$$

(2) Lagrange 插值法.

命题: 设 $m_A(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_m)$ 无重根

$f(t)$ 为任一幂级数, 收敛半径 $r > \rho(A)$

$$\boxed{\begin{aligned}&\text{则 } f(A)=\sum_{i=1}^{m} f(\lambda_i)L_i(A),\\ &\text{其中 } L_i(\lambda)=\frac{(\lambda-\lambda_1)\cdots(\lambda-\lambda_{i-1})(\lambda-\lambda_{i+1})\cdots(\lambda-\lambda_m)}{(\lambda_i-\lambda_1)\cdots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\cdots(\lambda_i-\lambda_m)},\end{aligned}}\quad i=1\sim m$$

命题2: 设 $m_A(\lambda)=(\lambda-\lambda_1)^{k_1}\cdots(\lambda-\lambda_s)^{k_s}$, $\sum_{i=1}^{s}k_i=m\le n$

则 $f(A)=\sum_{i=1}^{s}\varphi_i(A)\left(a_{i1}E+a_{i2}(A-\lambda_iE)+\cdots+a_{ik_i}(A-\lambda_iE)^{k_i-1}\right)$

其中 $\varphi_i(A)=(A-\lambda_1E)^{k_1}\cdots(A-\lambda_iE)^{k_{i-1}}(A-\lambda_{i+1}E)^{k_{i+1}}\cdots(A-\lambda_sE)^{k_s}$

$$a_{ij}=\frac{1}{(j-1)!}\frac{d^{j-1}}{d\lambda^{j-1}}\left((\lambda-\lambda_i)^{k_i}\frac{f(\lambda)}{m(\lambda)}\right)\Bigg|_{\lambda=\lambda_i},\quad i=1\sim s,\ j=1\sim k.$$

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

解. 法一. Lagrange - Sylvester 定理

法二. 待定系数法

法三. $f_A(\lambda)=|\lambda E-A|=(\lambda+1)(\lambda-3)$,

$m_A(\lambda)=(\lambda+1)(\lambda-3)$

令 $f(\lambda)=e^{\lambda t}$, 则 $f(A)=f(\lambda_1)L_1(A)+f(\lambda_2)L_2(A)$

其中 $L_1(A)=\dfrac{A-3E}{-1-3}=-\dfrac{1}{4}\begin{pmatrix}-2&2\\2&-2\end{pmatrix}$

$L_2(A)=\dfrac{A+E}{3+1}=\dfrac{1}{4}\begin{pmatrix}2&2\\2&2\end{pmatrix}$

$$f(A)=e^{-t}\cdot\frac{1}{4}\begin{pmatrix}-2&2\\2&-2\end{pmatrix}+e^{3t}\cdot\frac{1}{4}\begin{pmatrix}2&2\\2&2\end{pmatrix},$$

$$=\frac{1}{2}\begin{pmatrix}e^{3t}+e^{-t}&e^{3t}-e^{-t}\\e^{3t}-e^{-t}&e^{3t}+e^{-t}\end{pmatrix}$$

§5. 矩阵函数在微分方程组的应用

一. 常系数线性齐次微分方程组的解

Th. 考虑定理问题：
$$\begin{cases} \dfrac{dx}{dt} = Ax \\ x(t)\big|_{t=t_0} = x(t_0) \end{cases} \quad (*)$$

类比于：
$$\begin{cases} \dfrac{dx}{dt} = \lambda x \\ x(t_0) = x_0 \end{cases} \Rightarrow x = e^{\lambda(t-t_0)} x_0$$

其中 $x(t) = (x_1(t), \cdots, x_n(t))^T$ 为未知向量函数.

$x(t_0) = (x_1(t_0), \cdots, x_n(t_0))^T$, A 为给定的常数矩阵.

则 $(*)$ 有唯一解：

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

证. 验证 $\frac{dx(t)}{dt} = \frac{d\left(e^{A(t-t_0)}x(t_0)\right)}{dt} = Ae^{A(t-t_0)}x(t_0) = Ax(t)$,

且 $x(t_0) = e^{A\cdot 0}x(t_0) = x(t_0)$, $\Longrightarrow$ $x(t)$ 为 (*) 的解.

设 $y(t)$ 亦为 (*) 的解, 令 $z(t) = \cancel{e^{-A(t-t_0)}y(t)} = e^{-At}y(t)$

则 $\frac{dz(t)}{dt} = -Ae^{-At}y(t) + e^{-At}y'(t)$

$= -Ae^{-At}y(t) + e^{-At}Ay(t) = 0$

$\Longrightarrow$ $z(t)$ 为常向量.

即. $z(t) = z(t_0)$, $\forall t$.

$\Longrightarrow e^{-At}y(t) = e^{-At_0}y(t_0) = e^{-At_0}x(t_0)$

$\Longrightarrow y(t) = e^{A(t-t_0)}x(t_0) = x(t)$. 即解唯一.

二. 线性常系数非齐次微分方程组的解

设 $A=(a_{ij})_{n\times n}$，$B=(b_{ij})_{n\times m}$，是常数矩阵.

$u(t)=(u_1(t),\cdots,u_m(t))^T$ 是已知向量函数.

$x(t)=(x_1(t),\cdots,x_n(t))^T$ 未知.

考虑
$$\begin{cases}\dfrac{dx(t)}{dt}=Ax(t)+Bu(t)\\[2mm] x(t)\big|_{t=t_0}=x(t_0)\end{cases} \qquad (*2)$$

Th. (*2) 有唯一解

$$\boxed{x(t)=e^{A(t-t_0)}x(t_0)+\int_{t_0}^{t}e^{A(t-s)}Bu(s)\,ds.}$$

证：设 $x(t)=e^{At}c(t)$ 为 *2 之解

则 $\dfrac{dx(t)}{dt}=Ae^{At}c(t)+e^{At}c'(t)$

从而 $Ae^{At}c(t)+e^{At}c'(t)=Ae^{At}c(t)+Bu(t)$

$\Rightarrow e^{At}c'(t)=Bu(t)$

$\Rightarrow c'(t)=e^{-At}Bu(t)$

$\Rightarrow c(t)=\int_{t_0}^{t}e^{-As}Bu(s)ds+C$，其中 C 为常数向量.

又 $x(t)\big|_{t=t_0}=x(t_0)$，

故 $e^{At_0}c(t_0)=e^{At_0}C=x(t_0)\Rightarrow C=e^{-At_0}x(t_0)$

从而 $x(t)=e^{A(t-t_0)}x(t_0)+\int_{t_0}^{t}e^{A(t-s)}Bu(s)ds.$

三. n阶常系数微分方程的解

考虑n阶常系数齐次微分方程之定解问题:

$$\boxed{\begin{cases} y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = 0 \\ y^{(i)}(t)\big|_{t=0} = y_0^{(i)}, \quad i = 0, 1, \cdots, n-1 \end{cases}} \qquad (*3)$$

令 $\begin{cases} x_1 = y \\ x_2 = y' = x_1' \\ \vdots \\ x_n = y^{(n-1)} = x_{n-1}' \end{cases}$ 则 $\begin{cases} x_1' = x_2 \\ x_2' = x_3 \\ \vdots \\ x_n' = -a_n x_1 - a_{n-1} x_2 - \cdots - a_1 x_n \end{cases}$

令 $x(t) = (x_1(t), \cdots, x_n(t))^T$, $x(0) = (x_1(0), \cdots, x_n(0))^T = (y_0, y_0', \cdots, y_0^{(n-1)})^T$

则 (*3) 等价于: $\begin{cases} \dfrac{dx}{dt} = Ax(t) \\ x(t)\big|_{t=0} = x(0) \end{cases}$ (*4) 其中 $A = \begin{pmatrix} 0, & 1, & 0 & \cdots & 0 \\ 0 & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & & 1 \\ -a_n & -a_{n-1} & & \cdots & -a_1 \end{pmatrix}$

且 (*3) 的解为 (*4) 解之第一个分量.

$$\boxed{\begin{aligned} y(t) &= (1,0,\cdots,0)\,x(t) = (1,0,\cdots,0)\,e^{At}x(0) \\ &= (1,0,\cdots,0)\,e^{At}\begin{pmatrix} y_0 \\ y_0' \\ \vdots \\ y_0^{(n-1)} \end{pmatrix}. \end{aligned}}$$

对于 n 阶常系数非齐次方程之定理问题

$$\begin{cases} y^{(n)} + a_1 y^{(n-1)} + \cdots + a_n y = u(t) \\ y^{(i)}(t)\big|_{t=0} = y_0^{(i)}, \quad i = 0,1,\cdots,n-1 \end{cases}$$

(*5)

$$B = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

类似可得: (*5) 的解是 $\begin{cases} \dfrac{dx(t)}{dt} = Ax(t) + Bu(t) \\ x(t)\big|_{t=0} = x(0) \end{cases}$ 解之第一个分量.

故 $$\boxed{y(t) = (1,0,\cdots,0)\left(e^{At}x(0) + \int_0^t e^{A(t-s)}Bu(s)\,ds\right)}$$