第三章：矩阵分析

矩阵的微分与积分

- 函数矩阵的微分

定义3.7：设$A(t)=\left(a_{ij}(t)\right)_{m\times n}$，如果每个$a_{ij}(t)$都在$[a,b]$上可微，则称矩阵$A(t)$在$[a,b]$上可微，并定义

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

称为矩阵$A(t)$对t的导数或微商.

函数矩阵的求导法则

- 定理3.12

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

(2) $\frac{d}{dt}\left[\lambda(t)A(t)\right]=\left[\frac{d}{dt}\lambda(t)\right]A(t)+\lambda(t)\frac{d}{dt}A(t),$

其中$\lambda(t)$为一数量函数;

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

矩阵的求导法则

(4) $\frac{d}{dt}\left[A(u(t))\right]=\frac{d}{du}\left[A(u)\right]\Big|_{u=u(t)}\cdot u'(t).$

(5) $\frac{d}{dt}\left[A(t)\right]^{-1}=-\left[A(t)\right]^{-1}\left[\frac{d}{dt}A(t)\right]\left[A(t)\right]^{-1};$

【证明】 $\because A^{-1}(t)\cdot A(t)=I$　两边对 t 求导得

$$\frac{d}{dt}[A^{-1}(t)\cdot A(t)]=(\frac{d}{dt}A^{-1}(t))\cdot A(t)+A^{-1}(t)\cdot\frac{d}{dt}(A(t))=0$$

$$\therefore [\frac{d}{dt}A^{-1}(t).]A(t)=-A^{-1}(t)\cdot\frac{d}{dt}A(t)\cdot$$

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

矩阵的求导法则

- 注意：由于矩阵的乘法不一定满足交换律，所以上式中乘法顺序一般是不能交换的

如：若k是一个常数矩阵，则有

$$\frac{d}{dt}[k\cdot A(t)]=k[\frac{d}{dt}A(t)]\qquad \frac{d}{dt}[A(t)\cdot k]=[\frac{d}{dt}A(t)]\cdot k$$

又如 $\frac{d}{dt}[A^2(t)]=\frac{d}{dt}[A(t)\cdot A(t)]$

$$=[\frac{d}{dt}A(t)]\cdot A(t)+A(t)\cdot[\frac{d}{dt}A(t)]$$

$$\neq 2A(t)[\frac{d}{dt}A(t)]$$

矩阵的求导法则

- 定理3.13：

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

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

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

矩阵的求导法则

【证明】 由 $e^{At}=\sum_{k=0}^{\infty}\frac{t^k}{k!}A^k$

$$\therefore \frac{d}{dt}e^{At}=\frac{d}{dt}\sum_{k=0}^{\infty}\frac{t^k}{k!}A^k=\sum_{k=0}^{\infty}\frac{kt^{k-1}}{k!}A^k=A\sum_{k=1}^{\infty}\frac{t^{k-1}}{(k-1)!}A^{k-1}=Ae^{At}$$

$$\therefore \frac{d}{dt}e^{At}=\frac{d}{dt}\sum_{k=0}^{\infty}\frac{t^k}{k!}A^k=\sum_{k=0}^{\infty}\frac{kt^{k-1}}{k!}A^k=[\sum_{k=1}^{\infty}\frac{t^{k-1}}{(k-1)!}A^{k-1}]A=e^{At}A$$

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

同理可证(2)(3)

注：由定理可知，任意方阵A都与 $e^{At},\sin At,\cos At$ 可交换

矩阵的求导法则

【例】已知 $\sin At=\begin{bmatrix}\sin t-2t\cos t & 0 & t\cos t\\ \sin t+2t\cos t-\sin 2t & \sin 2t & -t\cos t-\sin t+\sin 2t\\ -4t\cos t & 0 & 2t\cos t+\sin t\end{bmatrix}$

求矩阵A

【解】 由定理3.13: $\frac{d}{dt}\sin At=A\cos At$

$$=\begin{bmatrix}\cos t-2\cos t+2t\sin t & 0 & \cos t-t\sin t\\ 3\cos t-2t\sin t-2\cos 2t & 2\cos 2t & t\sin t-2\cos t+2\cos 2t\\ -4\cos t+4t\sin t & 0 & 3\cos t-2t\sin t\end{bmatrix}$$

令$t=0$得：

$$\because \cos O=\begin{bmatrix}\cos 0 & 0 & 0\\ 0 & \cos 0 & 0\\ 0 & 0 & \cos 0\end{bmatrix}=I \qquad \therefore A=\begin{bmatrix}-1 & 0 & 1\\ 1 & 2 & 0\\ -4 & 0 & 3\end{bmatrix}$$

矩阵的积分

- 定义3. 7：

设$A(t)=\left(a_{ij}(t)\right)_{m\times n}$，如果每个$a_{ij}(t)$都在$[a,b]$上可积，则称函数矩阵$A(t)$在$[a,b]$上可积，并定义

$$\int_a^b A(t)dt=\left(\int_a^b a_{ij}(t)dt\right)_{m\times n},$$

称为矩阵$A(t)$在区间$[a,b]$上的积分.

矩阵的积分运算法则

- 定理3. 14：

(1) $\int_a^b[A(t)\pm B(t)]dt=\int_a^b A(t)dt\pm\int_a^b B(t)dt$;

(2) $\int_a^b \lambda A(t)dt=\lambda\int_a^b A(t)dt$, 其中$\lambda$为常数;

(3) $\int_a^b AB(t)dt=A\int_a^b B(t)dt$,

$\int_a^b A(t)Bdt=\left(\int_a^b A(t)dt\right)B$, 其中$A,B$为常数矩阵；

微分与积分的关系

(4) $A(t)$在$[a,b]$连续 $\Rightarrow \forall t\in[a,b]$, $\frac{d}{dt}\int_a^t A(s)ds=A(t)$;

(5) $A(t)$在$[a,b]$连续可微 $\Rightarrow \int_a^b\frac{d}{dt}A(t)dt=A(b)-A(a)$.

数量函数对矩阵变量的导数

- 定义：设$f(X)$是以矩阵变量$X=\left(x_{ij}\right)_{m\times n}$为自变量的$mn$元函数，且$\frac{\partial f}{\partial x_{ij}}$存在，定义

$$\frac{df}{dX}=\left(\frac{\partial f}{\partial x_{ij}}\right)_{m\times n}=\begin{pmatrix}\frac{\partial f}{\partial x_{11}} & \frac{\partial f}{\partial x_{12}} & \cdots & \frac{\partial f}{\partial x_{1n}}\\ \frac{\partial f}{\partial x_{21}} & \frac{\partial f}{\partial x_{22}} & \cdots & \frac{\partial f}{\partial x_{2n}}\cdots\\ \cdots & \cdots & \cdots & \cdots\\ \frac{\partial f}{\partial x_{m1}} & \frac{\partial f}{\partial x_{m2}} & \cdots & \frac{\partial f}{\partial x_{mn}}\end{pmatrix},$$

注：$f(X)\in C$ $f'(X)\in C^{m\times n}$

数量函数对向量变量的导数

Remark. 当$X=(x_1,x_2,\cdots,x_n)^T$时，

$$\frac{df}{dX}=\left(\frac{df}{dx_1},\frac{df}{dx_2},\cdots,\frac{df}{dx_n}\right)^T=\mathrm{grad}f.$$

例3.12 $a=(a_1,a_2,\cdots,a_n)^T$, $x=(x_1,x_2,\cdots,x_n)^T$,

$f(x)=a^Tx=x^Ta$. 求$\dfrac{df}{dx}$.

Solution. $\because \dfrac{\partial f}{\partial x_j}=a_j \quad \therefore \dfrac{\partial f}{\partial x}=\left(\dfrac{\partial f}{\partial x_1},\dfrac{\partial f}{\partial x_2},\cdots,\dfrac{\partial f}{\partial x_n}\right)^T=a.$

数量函数对矩阵变量的导数

例3.13 $A=(a_{ij})_{m\times n}$, $X=(x_{ij})_{n\times m}$, $f(X)=tr(AX)$. 求$\dfrac{df}{dX}$.

Solution. $\because AX=\left(\sum_{k=1}^{n}a_{ik}x_{kj}\right)_{m\times m}$, $f(X)=\sum_{s=1}^{m}\left(\sum_{k=1}^{n}a_{sk}x_{ks}\right)$,

$\dfrac{\partial f}{\partial x_{ij}}=a_{ji}$,

$$\therefore \frac{\partial f}{\partial X}=\left(\frac{\partial f}{\partial x_{ij}}\right)_{n\times m}=(a_{ji})_{n\times m}=A^T.$$

例3.14 $A=(a_{ij})_{n\times n}$, $x=(x_1,x_2,\cdots,x_n)^T$, $f(x)=x^TAx$. 求$\dfrac{df}{dx}$.

数量函数对矩阵变量的导数

Solution. $\because f(x)=\sum_{k=1}^{n}\sum_{l=1}^{n}a_{kl}x_kx_l$,

$$\frac{\partial f}{\partial x_i}=\sum_{l=1}^{n}a_{il}x_l+\sum_{k=1}^{n}a_{ki}x_k=A_{i\cdot}x+A_{\cdot i}^Tx=x^TA_{i\cdot}^T+x^TA_{\cdot i}$$

$$=x^T\left(A_{i\cdot}^T+A_{\cdot i}\right),$$

$$\therefore \frac{\partial f}{\partial x}=\left(x^T\left(A_{1\cdot}^T+A_{\cdot 1}\right),x^T\left(A_{2\cdot}^T+A_{\cdot 2}\right),\cdots,x^T\left(A_{n\cdot}^T+A_{\cdot n}\right)\right)^T$$

$$=\left(x^T\left(A_{1\cdot}^T+A_{\cdot 1},A_{2\cdot}^T+A_{\cdot 2},\cdots,A_{n\cdot}^T+A_{\cdot n}\right)\right)^T$$

$$=\left(A_{1\cdot}^T+A_{\cdot 1},A_{2\cdot}^T+A_{\cdot 2},\cdots,A_{n\cdot}^T+A_{\cdot n}\right)^Tx$$

数量函数对矩阵变量的导数

$$=\left(A_{1\cdot}^T+A_{\cdot 1},A_{2\cdot}^T+A_{\cdot 2},\cdots,A_{n\cdot}^T+A_{\cdot n}\right)^Tx$$

$$=\begin{pmatrix}A_{1\cdot}+A_{\cdot 1}^T\\A_{2\cdot}+A_{\cdot 2}^T\\\cdots\\A_{n\cdot}+A_{\cdot n}^T\end{pmatrix}x=\begin{pmatrix}A_{1\cdot}\\A_{2\cdot}\\\cdots\\A_{n\cdot}\end{pmatrix}x+\begin{pmatrix}A_{\cdot 1}^T\\A_{\cdot 2}^T\\\cdots\\A_{\cdot n}^T\end{pmatrix}x$$

$$=\begin{pmatrix}A_{1\cdot}\\A_{2\cdot}\\\cdots\\A_{n\cdot}\end{pmatrix}x+\begin{pmatrix}A_{\cdot 1}&A_{\cdot 2}&\cdots&A_{\cdot n}\end{pmatrix}^Tx=\left(A+A^T\right)x.$$

特别，当$A^T=A$时，$\dfrac{df}{dx}=2Ax.$

数量函数对矩阵变量的导数

例3.15 $X=(x_{ij})_{n\times n}$, $\det X\neq 0$, $f(X)=\det X$. 证明

$$\frac{df}{dX}=\det X\left(X^{-1}\right)^T.$$

Proof. 设x_{ij}的代数余子式为X_{ij}. 把$\det(X)$按第i行展开，得

$$\det X=\sum_{k=1}^{n}x_{ik}X_{ik},$$

易见 $\dfrac{\partial f}{\partial x_{ij}}=X_{ij}$， 故

$$\frac{df}{dX}=\left(X_{ij}\right)=\det X\cdot\frac{1}{\det X}\left(X_{ij}\right)=\det X\left(X^{-1}\right)^T.$$

数量函数对矩阵变量的导数

- 例：考虑高斯分布

$$p(\boldsymbol{x};\boldsymbol{\mu},\Sigma)=(2\pi)^{-\frac{D}{2}}|\Sigma|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$

– 对数似然函数：

$$\ln p(\boldsymbol{x};\boldsymbol{\mu},\Sigma)=-\frac{D}{2}\ln 2\pi+\frac{1}{2}\ln|\Sigma^{-1}|-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$$

– 对数似然函数的导数：

$$\frac{\partial \ln p(\boldsymbol{x};\boldsymbol{\mu},\Sigma)}{\partial\boldsymbol{\mu}}=\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$$

$$\frac{\partial \ln p(\boldsymbol{x};\boldsymbol{\mu},\Sigma)}{\partial\Sigma^{-1}}=\frac{1}{2}\Sigma-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^T$$

矩阵值函数对矩阵变量的导数

- **定义：** 设$F(X)=\left(f_{ij}(X)\right)_{s\times t}$，$X=\left(x_{ij}\right)_{m\times n}$，称$F(X)$为矩阵值函数，定义$F(X)$对矩阵变量$X$的导数为

$$\frac{dF}{dX}:=\begin{pmatrix} \frac{\partial F}{\partial x_{11}} & \cdots & \frac{\partial F}{\partial x_{1n}} \\ \cdots & \cdots & \cdots \\ \frac{\partial F}{\partial x_{m1}} & \cdots & \frac{\partial F}{\partial x_{mn}} \end{pmatrix}_{m\times n}，\text{其中 } \frac{\partial F}{\partial x_{ij}}=\begin{pmatrix} \frac{\partial f_{11}}{\partial x_{ij}} & \cdots & \frac{\partial f_{1t}}{\partial x_{ij}} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_{s1}}{\partial x_{ij}} & \cdots & \frac{\partial f_{st}}{\partial x_{ij}} \end{pmatrix}_{s\times t}.$$

注1）　$F(X)\in C^{s\times t}$　$\frac{dF}{dX}\in C^{ms\times nt}$

注2）当$s=t=1$时即为函数对矩阵变量的导数

矩阵值函数对矩阵变量的导数

例3.16　$x=(\xi_1,\xi_2,\cdots,\xi_n)^T$. $\frac{dx}{dx^T}=I_n$，$\frac{dx^T}{dx}=I_n$.

例3.17　$a=(a_1,a_2,a_3,a_4)^T$，$X=\left(x_{ij}\right)_{2\times 4}$.

$$\frac{d(Xa)^T}{dX}=\begin{pmatrix} a_1 & 0 & a_2 & 0 & a_3 & 0 & a_4 & 0 \\ 0 & a_1 & 0 & a_2 & 0 & a_3 & 0 & a_4 \end{pmatrix},$$

$$\frac{d(Xa)}{dX}=\begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_1 & a_2 & a_3 & a_4 \end{pmatrix}.$$

注：更多矩阵导数可查阅 The Matrix Cookbook

第三章：矩阵分析

矩阵分析应用举例

- 求解一阶线性常系数微分方程组

$$\begin{cases} \frac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t}=A\boldsymbol{x}(t)+\boldsymbol{f}(t) \\ \boldsymbol{x}(t_0)=\boldsymbol{x}_0 \end{cases}$$

- 此微分方程组的解为

$$\boldsymbol{x}(t)=e^{A(t-t_0)}\boldsymbol{x}_0+\int_{t_0}^{t}e^{A(t-s)}\boldsymbol{f}(s)\mathrm{d}s$$

一阶线性常系数微分方程组

$$\because \frac{d(e^{-At}\vec{x}(t))}{dt}$$

$$=e^{-At}(-A)\vec{x}(t)+e^{-At}\frac{d(\vec{x}(t))}{dt}=e^{-At}\left[\frac{d(\vec{x}(t))}{dt}-A\vec{x}(t)\right]=e^{-At}\vec{f}(t)$$

$$\therefore \int_{t_0}^{t}\frac{d}{d\tau}[e^{-A\tau}\vec{x}(\tau)]d\tau=\int_{t_0}^{t}e^{-A\tau}\vec{f}(\tau)d\tau$$

即　$e^{-A\tau}\vec{x}(\tau)\Big|_{t_0}^{t}=e^{-At}\vec{x}(t)-e^{-At_0}\vec{x}(t_0)=\int_{t_0}^{t}e^{-A\tau}\vec{f}(\tau)d\tau$

即　$e^{-At}\vec{x}(t)=e^{-At_0}\vec{x}(t_0)+\int_{t_0}^{t}e^{-A\tau}\vec{f}(\tau)d\tau$

由于$-At$与At_0可换，故有：

$$\vec{x}(t)=e^{A(t-t_0)}\vec{x}_0+e^{At}\int_{t_0}^{t}e^{-A\tau}\vec{f}(\tau)d\tau$$

一阶线性常系数微分方程组

求解的一般步聚：

1）计算矩阵函数　e^{At}；$e^{A(t-t_0)}$；e^{-At}

2）计算矩阵向量乘积　$e^{A(t-t_0)}\vec{x}_0$

3）计算积分$\int_{t_0}^{t}e^{-A\tau}\vec{f}(\tau)d\tau$与矩阵向量乘积$e^{At}\int_{t_0}^{t}e^{-A\tau}\vec{f}(\tau)d\tau$

则　$\vec{x}(t)=e^{A(t-t_0)}\vec{x}_0+e^{At}\int_{t_0}^{t}e^{-A\tau}\vec{f}(\tau)d\tau$

注：$\vec{f}(t)=\vec{0}$时，解为　$\vec{x}(t)=e^{A(t-t_0)}\vec{x}_0$

一阶线性常系数微分方程组

例 3.18 求解微分方程组初值问题

$$\begin{cases} \dfrac{\mathrm{d}x_1(t)}{\mathrm{d}t} = -x_1(t) + x_3(t) + 1, \\ \dfrac{\mathrm{d}x_2(t)}{\mathrm{d}t} = x_1(t) + 2x_2(t) - 1, \\ \dfrac{\mathrm{d}x_3(t)}{\mathrm{d}t} = -4x_1(t) + 3x_3(t) + 2, \\ x_1(0) = 1,\ x_2(0) = 0,\ x_3(0) = 1. \end{cases}$$

解 记

$$\boldsymbol{A} = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 2 & 0 \\ -4 & 0 & 3 \end{bmatrix},\quad \boldsymbol{c} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},\quad \boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix},\quad \boldsymbol{f}(t) = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix},$$

一阶线性常系数微分方程组

则微分方程组可以写成式(3.8)的矩阵形式. 例 3.9 已求得

$$\mathrm{e}^{\boldsymbol{A}t} = \begin{bmatrix} \mathrm{e}^t - 2t\mathrm{e}^t & 0 & t\mathrm{e}^t \\ -\mathrm{e}^{2t} + \mathrm{e}^t + 2t\mathrm{e}^t & \mathrm{e}^{2t} & \mathrm{e}^{2t} - \mathrm{e}^t - t\mathrm{e}^t \\ -4t\mathrm{e}^t & 0 & 2t\mathrm{e}^t + \mathrm{e}^t \end{bmatrix},$$

依次计算下列各量

$$\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{c} = \begin{bmatrix} \mathrm{e}^t - t\mathrm{e}^t \\ t\mathrm{e}^t \\ \mathrm{e}^t - 2t\mathrm{e}^t \end{bmatrix},$$

一阶线性常系数微分方程组

$$\int_0^t \mathrm{e}^{-\boldsymbol{A}\tau}\boldsymbol{f}(\tau)\mathrm{d}\tau = \int_0^t \begin{bmatrix} \mathrm{e}^{-\tau} \\ -\mathrm{e}^{-\tau} \\ 2\mathrm{e}^{-\tau} \end{bmatrix} \mathrm{d}\tau = \begin{bmatrix} 1 - \mathrm{e}^{-t} \\ -1 + \mathrm{e}^{-t} \\ 2 - 2\mathrm{e}^{-t} \end{bmatrix},$$

$$\mathrm{e}^{\boldsymbol{A}t}\int_0^t \mathrm{e}^{-\boldsymbol{A}\tau}\boldsymbol{f}(\tau)\mathrm{d}\tau = \begin{bmatrix} \mathrm{e}^t - 1 \\ -\mathrm{e}^t + 1 \\ 2\mathrm{e}^t - 2 \end{bmatrix},$$

故微分方程组的解为

$$\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} \mathrm{e}^t - t\mathrm{e}^t \\ t\mathrm{e}^t \\ \mathrm{e}^t - 2t\mathrm{e}^t \end{bmatrix} + \begin{bmatrix} \mathrm{e}^t - 1 \\ -\mathrm{e}^t + 1 \\ 2\mathrm{e}^t - 2 \end{bmatrix} = \begin{bmatrix} (2-t)\mathrm{e}^t - 1 \\ (t-1)\mathrm{e}^t + 1 \\ (3-2t)\mathrm{e}^t - 2 \end{bmatrix}.$$

矩阵分析应用举例

- 求解矩阵方程

  定理3.15：给定矩阵方程

  $$AX + XB = F \quad \text{(Sylvester equation)}$$

  其中$A \in \mathbb{C}^{m\times m}, B \in \mathbb{C}^{n\times n}, F \in \mathbb{C}^{m\times n}$. 如果$A$和$B$的所有特征值具有负实部（这种矩阵称为稳定矩阵），则该矩阵方程有唯一解

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

证 记$\boldsymbol{Y}(t) = \mathrm{e}^{\boldsymbol{A}t}\boldsymbol{F}\mathrm{e}^{\boldsymbol{B}t}$，则有$\boldsymbol{Y}(0) = \boldsymbol{F}$，且

$$\frac{\mathrm{d}\boldsymbol{Y}(t)}{\mathrm{d}t} = \boldsymbol{A}\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{F}\mathrm{e}^{\boldsymbol{B}t} + \mathrm{e}^{\boldsymbol{A}t}\boldsymbol{F}\mathrm{e}^{\boldsymbol{B}t}\boldsymbol{B} = \boldsymbol{A}\boldsymbol{Y}(t) + \boldsymbol{Y}(t)\boldsymbol{B}. \tag{3.10}$$

设$\lambda_1, \lambda_2, \cdots, \lambda_m$是$\boldsymbol{A}$的$m$个特征值，$\mu_1, \mu_2, \cdots, \mu_n$是$\boldsymbol{B}$的$n$个特征值. 根据利用 Jordan 标准形求矩阵函数的方法(见 3.3 节)知，$\mathrm{e}^{\boldsymbol{A}t}$的元素是形如$t^r\mathrm{e}^{\lambda_j t}$ $(r \geqslant 0)$的项的线性组合. 因为$\boldsymbol{A}$的所有特征值λ_j的实部是负的，所以$\lim\limits_{t\to+\infty}\mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{O}$. 同理$\lim\limits_{t\to+\infty}\mathrm{e}^{\boldsymbol{B}t} = \boldsymbol{O}$. 于是

$$\lim_{t\to+\infty}\boldsymbol{Y}(t) = \lim_{t\to+\infty}\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{F}\mathrm{e}^{\boldsymbol{B}t} = \boldsymbol{O}.$$

又由于$\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{F}\mathrm{e}^{\boldsymbol{B}t}$的元素是形如$t^r\mathrm{e}^{(\lambda_i+\mu_j)t}$ $(r \geqslant 0)$的项的线性组合，且积分$\int_0^{+\infty} t^r\mathrm{e}^{(\lambda_i+\mu_j)t}\mathrm{d}t$都存在，故积分$\int_0^{+\infty}\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{F}\mathrm{e}^{\boldsymbol{B}t}\mathrm{d}t$存在.

对式(3.10)的两边从 0 到$+\infty$积分，得

$$\boldsymbol{Y}(+\infty) - \boldsymbol{Y}(0) = \boldsymbol{A}\left(\int_0^{+\infty}\boldsymbol{Y}(t)\mathrm{d}t\right) + \left(\int_0^{+\infty}\boldsymbol{Y}(t)\mathrm{d}t\right)\boldsymbol{B},$$

即

$$\boldsymbol{A}\left(-\int_0^{+\infty}\boldsymbol{Y}(t)\mathrm{d}t\right) + \left(-\int_0^{+\infty}\boldsymbol{Y}(t)\mathrm{d}t\right)\boldsymbol{B} = \boldsymbol{F}.$$

这说明$\boldsymbol{X} = -\int_0^{+\infty}\mathrm{e}^{\boldsymbol{A}t}\boldsymbol{F}\mathrm{e}^{\boldsymbol{B}t}\,\mathrm{d}t$是矩阵方程(3.9)的解.

唯一性的证明见第 7 章. 证毕

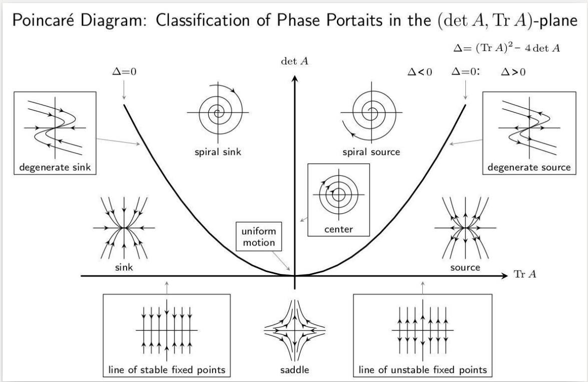
矩阵分析应用举例

- 稳定矩阵介绍
  - 考虑一阶齐次线性常系数微分方程组

    $$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{x}(t)}{\mathrm{d}t} = A\boldsymbol{x}(t) \\ \boldsymbol{x}(0) = \boldsymbol{x}_0 \end{cases}$$

  - 解为：$\boldsymbol{x}(t) = e^{At}\boldsymbol{x}_0$
  - 若A所有特征值的实部都为负，即$Re(\lambda) < 0$，则微分方程组的解$e^{At}\boldsymbol{x}_0$是渐近稳定的，即$e^{At}\boldsymbol{x}_0 \to \boldsymbol{0}$，当$t \to \infty$

稳定性示意图



https://en.wikipedia.org/wiki/Stability_theory

矩阵分析应用举例

- 引理：设$A \in \mathbb{C}^{m\times m}, B \in \mathbb{C}^{n\times n}, F \in \mathbb{C}^{m\times n}$，则矩阵微分方程

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

的解为

$$X(t) = e^{At}Fe^{Bt}$$

矩阵分析应用举例

- 推论：设$A, F \in \mathbb{C}^{n\times n}$，且$A$的所有特征值具有负实部，则矩阵方程

$$A^H X + XA = -F \text{ (Lyapunov equation)}$$

的唯一解为

$$X = \int_0^{+\infty} e^{A^H t} F e^{At}\, \mathrm{d}t$$

如果F是Hermite正定矩阵，则解矩阵X也是Hermite正定矩阵

矩阵分析应用举例

- 最小二乘问题

定理3.16：设$A \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m$. 若$\boldsymbol{x}_0 \in \mathbb{R}^n$是最小二乘解，即

$$\boldsymbol{x}_0 = \arg\min_{\boldsymbol{x}\in\mathbb{R}^n} \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2$$

则$\boldsymbol{x}_0$是方程组

$$A^T A\boldsymbol{x} = A^T \boldsymbol{b}$$

的解. 称此方程组为$A\boldsymbol{x} = \boldsymbol{b}$的法方程组

【注】只有矛盾方程才有最小二乘解的概念

矩阵分析应用举例

证　由于

$$f(\boldsymbol{x}) = \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2 = (A\boldsymbol{x} - \boldsymbol{b})^{\mathrm{T}}(A\boldsymbol{x} - \boldsymbol{b})$$
$$= \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x} - \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{b} - \boldsymbol{b}^{\mathrm{T}}A\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b},$$

若 $\boldsymbol{x}_0$ 为 $A\boldsymbol{x}=\boldsymbol{b}$ 的最小二乘解，则它应是 $f(\boldsymbol{x})$ 的极小值点，从而

$$\left.\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}}\right|_{\boldsymbol{x}_0} = \boldsymbol{0}. \tag{3.14}$$

根据例 3.12 和例 3.14，得

$$\frac{\mathrm{d}f}{\mathrm{d}\boldsymbol{x}} = 2A^{\mathrm{T}}A\boldsymbol{x} - 2A^{\mathrm{T}}\boldsymbol{b}.$$

由式(3.14)即知 $A^{\mathrm{T}}A\boldsymbol{x}_0 - A^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{0}$，故 $\boldsymbol{x}_0$ 是式(3.13)的解.　　证毕

【例】求下列方程组的最小二乘解 $\begin{bmatrix}1 & 2\\0 & 0\\2 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}$

【解】

记 $A = \begin{bmatrix}1 & 2\\0 & 0\\2 & 4\end{bmatrix}$　$\vec{x} = \begin{bmatrix}x_1\\x_2\end{bmatrix}, \vec{b} = \begin{bmatrix}1\\1\\1\end{bmatrix}$

则原方程无解，且其法方程为　$A^T A\vec{x} = A^T\vec{b}$

即 $\begin{bmatrix}1 & 0 & 2\\2 & 0 & 4\end{bmatrix}\begin{bmatrix}1 & 2\\0 & 0\\2 & 4\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}1 & 0 & 2\\2 & 0 & 4\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}$

即 $\begin{bmatrix}5 & 10\\10 & 20\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}3\\6\end{bmatrix}$

最小二乘解的通解为　$\vec{x}_0 = \begin{bmatrix}-2\\1\end{bmatrix}\lambda + \begin{bmatrix}\frac{3}{5}\\0\end{bmatrix}$

【注】最小二乘解$\boldsymbol{x}_0$一般不唯一，但$A\boldsymbol{x}_0$为同一向量

矩阵分析应用举例

- 带约束的最小二乘问题
  例3.19：设$A \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m, B \in \mathbb{R}^{k\times n}$, $\boldsymbol{d} \in \mathbb{R}^k$, 且$B\boldsymbol{x} = \boldsymbol{d}$有解，试求下列最优化问题的解

$$\min_{B\boldsymbol{x}=\boldsymbol{d}} \|A\boldsymbol{x} - \boldsymbol{b}\|_2^2$$

解：

$$\begin{pmatrix} A^T A & B^T \\ B & O \end{pmatrix}\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} A^T\boldsymbol{b} \\ \boldsymbol{d} \end{pmatrix}$$

思考题

1. 设 $X = (x_{ij})_{n\times n}$ 证明：$\frac{d(trX)}{dX} = I_n$

2. 设 $A = (a_{ij})_{n\times n}, \vec{x} = (x_1(t), x_2(t), \ldots, x_n(t))^T, f(\vec{x}) = \vec{x}^T A\vec{x}$

   证明：(1) $\frac{df}{dt} = \vec{x}^T(A^T + A)\frac{d\vec{x}}{dt}$;

   (2) $\frac{d}{dt}\vec{x}^T\vec{x} = 2\vec{x}^T\frac{d\vec{x}}{dt}$.

小结

- 矩阵的微分及性质
- 矩阵的积分及其性质
- 矩阵分析的应用举例

Assignment

- 徐仲、张凯院等.《矩阵论简明教程》,第三版.科学出版社出版.
  - 习题三：12、15、16、17、20