第七章　函数矩阵与矩阵微分方程

函数矩阵

定义： 以实变量 x 的函数为元素的矩阵

$$A(x)=\begin{bmatrix} a_{11}(x) & a_{12}(x) & \cdots & a_{1n}(x) \\ a_{21}(x) & a_{22}(x) & \cdots & a_{2n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}(x) & a_{m2}(x) & \cdots & a_{mn}(x) \end{bmatrix}$$

称为函数矩阵，其中所有的元素

$$a_{ij}(x),\quad i=1,2,\cdots,m;\ j=1,2,\cdots,n$$

都是定义在闭区间 $[a,b]$ 上的实函数。

函数矩阵与数字矩阵一样也有加法，数乘，乘法，转置等几种运算，并且运算法则完全相同。

例：已知

$$A=\begin{bmatrix}1-x & \sin x\\ e^x & 1+x\end{bmatrix},\quad B=\begin{bmatrix}1+x & \cos x\\ e^x & 1-x\end{bmatrix}$$

计算 $A+B,\ AB,\ A^T,\ 2^x(A-B)$

定义：设 $A(x)$ 为一个 n 阶函数矩阵，如果存在 n 阶函数矩阵 $B(x)$ 使得对于**任何** $\boldsymbol{x\in[a,b]}$都有

$$A(x)B(x)=B(x)A(x)=I$$

那么我们称 $A(x)$ 在区间 $[a,b]$ 上是可逆的。

称 $B(x)$ 是 $A(x)$ 的逆矩阵，一般记为 $A^{-1}(x)$

函数矩阵对纯量的导数和积分

定义： 如果 $A(x)=(a_{ij}(x))_{m\times n}$ 的所有各元素

$a_{ij}(x)$ 在 $x=x_0$ 处有极限，即

$$\lim_{x\to x_0} a_{ij}(x)=a_{ij} \qquad (i=1,\cdots,m;\ j=1,\cdots,n)$$

其中 a_{ij} 为固定常数。则称 $A(x)$ 在 $x=x_0$ 处有极限，且记为

$$\lim_{x\to x_0} \boldsymbol{A}(x)=\boldsymbol{A}$$

如果 $A(x)$ 的各元素 $a_{ij}(x)$ 在 $x=x_0$处连续，即

$$\lim_{x\to x_0} a_{ij}(x)=a_{ij}(x_0) \qquad (i=1,\cdots,m;\ j=1,\cdots,n)$$

则称 $A(x)$ 在 $x=x_0$处连续，且记为

$$\lim_{x\to x_0} A(x)=A(x_0)$$

其中

$$A(x_0)=\begin{bmatrix} a_{11}(x_0) & a_{12}(x_0) & \cdots & a_{1n}(x_0) \\ a_{21}(x_0) & a_{22}(x_0) & \cdots & a_{2n}(x_0) \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}(x_0) & a_{m2}(x_0) & \cdots & a_{mn}(x_0) \end{bmatrix}$$

容易验证下面的等式是成立的：

设 $$\lim_{x\to x_0} A(x)=A,\quad \lim_{x\to x_0} B(x)=B$$

则 $$(1)\quad \lim_{x\to x_0}(A(x)\pm B(x))=A\pm B$$

(2) $\lim_{x \to x_0}(kA(x)) = kA$

(3) $\lim_{x \to x_0}(A(x)B(x)) = AB$

定义： 如果 $A(x) = (a_{ij}(x))_{m \times n}$ 的所有各元素 $a_{ij}(x)(i = 1, \cdots, m; j = 1, \cdots, n)$ 在点 $x = x_0$ 处(或在区间 $[a,b]$ 上)可导，便称此函数矩阵 $A(x)$ 在点 $x = x_0$ 处(或在区间 $[a,b]$ 上)**可导**，并且记为

$$A'(x_0)=\left.\frac{\mathrm{d}A(x)}{\mathrm{d}x}\right|_{x=x_0}=\lim_{\Delta x\to 0}\frac{A(x_0+\Delta x)-A(x_0)}{\Delta x}$$

$$=\begin{bmatrix} a'_{11}(x_0) & a'_{12}(x_0) & \cdots & a'_{1n}(x_0) \\ a'_{21}(x_0) & a'_{22}(x_0) & \cdots & a'_{2n}(x_0) \\ \cdots & \cdots & \cdots & \cdots \\ a'_{m1}(x_0) & a'_{m2}(x_0) & \cdots & a'_{mn}(x_0) \end{bmatrix}$$

函数矩阵的导数运算有下列性质：

(1) $A(x)$ 是常数矩阵的充分必要条件是

$$\frac{\mathrm{d}A(x)}{\mathrm{d}x}=0$$

（**2**）设 $A(x)=(a_{ij}(x))_{m\times n},B(x)=(b_{ij}(x))_{m\times n}$

均可导，则

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{x}}[\boldsymbol{A}(\boldsymbol{x})+\boldsymbol{B}(\boldsymbol{x})]=\frac{\mathrm{d}\boldsymbol{A}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}+\frac{\mathrm{d}B(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}$$

(3) 设 $k(x)$ 是 x 的纯量函数，$A(x)$ 是函数矩阵，

$k(x)$ 与 $A(x)$ 均可导，则

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{x}}[\boldsymbol{k}(\boldsymbol{x})\boldsymbol{A}(\boldsymbol{x})]=\frac{\mathrm{d}k(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}\boldsymbol{A}(\boldsymbol{x})+\boldsymbol{k}(\boldsymbol{x})\frac{\mathrm{d}A(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}$$

特别地，当 $k(x)$ 是常数 k 时有

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{x}}[\boldsymbol{k}\boldsymbol{A}(\boldsymbol{x})]=\boldsymbol{k}\frac{\mathrm{d}\boldsymbol{A}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}$$

(4) 设 $A(x), B(x)$ 均可导，且 $A(x)$ 与 $B(x)$ 是可乘的，则

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{x}}[\boldsymbol{A}(\boldsymbol{x})\boldsymbol{B}(\boldsymbol{x})]=\frac{\mathrm{d}\boldsymbol{A}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}\boldsymbol{B}(\boldsymbol{x})+\boldsymbol{A}(\boldsymbol{x})\frac{\mathrm{d}B(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}$$

因为矩阵没有交换律，所以

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{x}}\boldsymbol{A}^2(\boldsymbol{x})\neq 2\boldsymbol{A}(\boldsymbol{x})\frac{\mathrm{d}\boldsymbol{A}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}$$

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{x}}\boldsymbol{A}^3(\boldsymbol{x})\neq 3\boldsymbol{A}^2(\boldsymbol{x})\frac{\mathrm{d}\boldsymbol{A}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}$$

(5) 设 $A(x)$ 为矩阵函数， $x=f(t)$ 是 t 的纯量函数，$A(x)$ 与 $f(t)$ 均可导，则

$$\frac{\mathrm{d}}{\mathrm{dt}}\boldsymbol{A}(\boldsymbol{x})=\frac{\mathrm{d}\boldsymbol{A}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}\boldsymbol{f}'(\boldsymbol{t})=\boldsymbol{f}'(\boldsymbol{t})\frac{\mathrm{d}\boldsymbol{A}(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}}$$

定义： 如果函数矩阵 $A(x)=(a_{ij}(x))_{m\times n}$ 的所有各

元素 $a_{ij}(x)(i=1,\cdots,m;\ j=1,\cdots,n)$ 在 $[a,b]$ 上可积，则称 $A(x)$ 在 $[a,b]$ 上可积，且

$$\int_a^b A(x)\mathrm{d}x=\begin{bmatrix}\int_a^b a_{11}(x)\mathrm{d}x & \int_a^b a_{12}(x)\mathrm{d}x & \cdots & \int_a^b a_{1n}(x)\mathrm{d}x \\ \int_a^b a_{21}(x)\mathrm{d}x & \int_a^b a_{22}(x)\mathrm{d}x & \cdots & \int_a^b a_{2n}(x)\mathrm{d}x \\ \cdots & \cdots & \cdots & \cdots \\ \int_a^b a_{m1}(x)\mathrm{d}x & \int_a^b a_{m2}(x)\mathrm{d}x & \cdots & \int_a^b a_{mn}(x)\mathrm{d}x\end{bmatrix}$$

函数矩阵的定积分具有如下性质：

$$\int_a^b kA(x)\mathrm{d}x = k\int_a^b A(x)\mathrm{d}x \qquad k \in R$$

$$\int_a^b [A(x)+B(x)]\mathrm{d}x = \int_a^b A(x)\mathrm{d}x + \int_a^b B(x)\mathrm{d}x$$

例 1： 已知函数矩阵

$$A(x) = \begin{bmatrix} 1 & x^2 \\ x & 0 \end{bmatrix}$$

试计算

(1) $\dfrac{d}{dx}A(x), \dfrac{d^2}{dx^2}A(x), \dfrac{d^3}{dx^3}A(x)$

(2) $\dfrac{d}{dx}\left|A(x)\right|$

(3) $\dfrac{d}{dx}A^{-1}(x)$

解：

$$\frac{d}{dx}A(x)=\begin{bmatrix}0 & 2x\\ 1 & 0\end{bmatrix}, \qquad \frac{d^2}{dx^2}A(x)=\begin{bmatrix}0 & 2\\ 0 & 0\end{bmatrix}$$

$$\frac{d^3}{dx^3}A(x)=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

由于 $\left|A(x)\right|=-x^3$，所以

$$\frac{d}{dx}\left|A(x)\right|=-3x^2$$

下面求 $A^{-1}(x)$。由伴随矩阵公式可得

$$A^{-1}(x)=\frac{1}{|A(x)|}A^{*}(x)$$

$$=-\frac{1}{x^{3}}\begin{bmatrix}0 & -x^{2}\\ -x & 1\end{bmatrix}=\begin{bmatrix}0 & \dfrac{1}{x}\\ \dfrac{1}{x^{2}} & -\dfrac{1}{x^{3}}\end{bmatrix}$$

再求 $\dfrac{d}{dx}A^{-1}(x)$

$$\frac{d}{d}A^{-1}(x)=\begin{bmatrix} 0 & -\frac{1}{x^2} \\ -\frac{2}{x^3} & \frac{3}{x^4} \end{bmatrix}$$

例 2：已知函数矩阵

$$A(x)=\begin{bmatrix} \sin x & \cos x & x \\ \frac{\sin x}{x} & e^x & x^2 \\ 1 & 0 & x^3 \end{bmatrix}$$

试求 (1) $\lim\limits_{x \to 0} A(x)$

(2) $\dfrac{\mathrm{d}}{\mathrm{d}x} A(x)$

(3) $\dfrac{\mathrm{d}^2}{\mathrm{d}x^2} A(x)$

(4) $\dfrac{\mathrm{d}}{\mathrm{d}x} \left| A(x) \right|$

(5) $\left| \dfrac{\mathrm{d}}{\mathrm{d}x} A(x) \right|$

$$\begin{bmatrix} \sin x & \cos x & x \\ \dfrac{\sin x}{x} & e^x & x^2 \\ 1 & 0 & x^3 \end{bmatrix}$$

例 3： 已知函数矩阵

$$A(x)=\begin{bmatrix}\sin x & -\cos x\\ \cos x & \sin x\end{bmatrix}$$

试求

$$\int_0^x A(x)dx, \quad \left(\int_0^{x^2} A(x)dx\right)'$$

解：

$$\int_0^x A(x)dx = \begin{bmatrix} \int_0^x \sin x dx & -\int_0^x \cos x dx \\ \int_0^x \cos x dx & \int_0^x \sin x dx \end{bmatrix}$$

$$= \begin{bmatrix} 1-\cos x & -\sin x \\ \sin x & 1-\cos x \end{bmatrix}$$

同样可以求得

$$(\int_0^{x^2} A(x)dx)' = 2x \begin{bmatrix} \sin x^2 & -\cos x^2 \\ \cos x^2 & \sin x^2 \end{bmatrix}$$

线性向量微分方程

定理： 设 A 是一个 n 阶常数矩阵，则微分方程组

$$\frac{dx(t)}{dt} = Ax(t)$$

满足初始条件 $x(t_0) = x_0$ 的解为

$$x = e^{A(t-t_0)}x_0$$

定理：设 A是一个 n阶常数矩阵，则微分方程组

$$\frac{dx(t)}{dt} = Ax(t) + f(t)$$

满足初始条件 $x(t_0) = x_0$ 的解为

$$x = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-\tau)} f(\tau) d\tau$$

补充：实值函数相对于实向量（矩阵）的导数

（不在考试范围）

$f(x)$ 为实值函数，$x=\left(x_1,\ x_2,\cdots,x_m\right)^T$

$$\frac{\partial f}{\partial x}\equiv\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_m}\right)^T$$

称为 $f(x)$ 的梯度向量.（反应了 f 的最大变化率）

$$\frac{\partial f}{\partial x^T}\equiv\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_m}\right)$$

$$\frac{\partial^2 f}{\partial \mathbf{x} \partial \mathbf{x}^T} \equiv \left[\frac{\partial^2 f}{\partial \mathbf{x}_i \partial \mathbf{x}_j} \right]$$

$$= \begin{bmatrix} \frac{\partial^2 f}{\partial \mathbf{x}_1 \partial \mathbf{x}_1} & \cdots & \frac{\partial^2 f}{\partial \mathbf{x}_1 \partial \mathbf{x}_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial \mathbf{x}_m \partial \mathbf{x}_1} & \cdots & \frac{\partial^2 f}{\partial \mathbf{x}_m \partial \mathbf{x}_m} \end{bmatrix}$$

$f(x)$ 的 **Hessian** 矩阵

判断某个点 x^* 是否为目标函数的局部极小点，可通过检验梯度

$$\frac{\partial f(x^*)}{\partial x} \left(\nabla_x f(x^*) \right),$$

和**Hessian** 矩阵

$$\nabla_x^2 f(x^*)$$

$f(X)$ 为实值函数， $X = [x_{ij}]_{m\times n}$

$$\frac{\partial f(X)}{\partial X} \equiv \left[\frac{\partial f(X)}{\partial x_{ij}} \right]$$

$$\boxed{\frac{\partial \mathrm{tr}(X)}{\partial X} = I}$$

$$= \begin{bmatrix} \frac{\partial f}{\partial x_{11}} & \cdots & \frac{\partial f}{\partial x_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_{m1}} & \cdots & \frac{\partial f}{\partial x_{mn}} \end{bmatrix} \quad (m\times n)$$

$$y(x) = \left[y_1(x), y_2(x), \cdots y_n(x) \right]^T$$ 实向量函数

$$x = \left(x_1, \; x_2, \cdots, x_m \right)^T$$

$$\frac{\partial y}{\partial x^T} \equiv \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_m} \end{bmatrix} \qquad (n \times m)$$

***y*(*x*) 的 Jacobian 矩阵**

$$\boldsymbol{y}(\boldsymbol{x}) = \left[y_1(\boldsymbol{x}),\ y_2(\boldsymbol{x}), \cdots y_n(\boldsymbol{x}) \right]^T$$ 实向量函数

$$\boldsymbol{x} = \left(x_1,\ x_2, \cdots, x_m \right)^T$$

$$\frac{\partial \boldsymbol{y}^T}{\partial \boldsymbol{x}} \equiv \left(\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}^T} \right)^T$$

$$= \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_n}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial y_1}{\partial x_m} & \cdots & \dfrac{\partial y_n}{\partial x_m} \end{bmatrix} \quad (m \times n)$$

$\boldsymbol{y}(\boldsymbol{x})$ 的 **Gradient (梯度) 矩阵**

梯度的性质：

(1) $x\,(m\times 1)$, $c\in R$,(常数): $\dfrac{\partial c}{\partial x}=0_{m\times 1}$

(2) $x\,(m\times 1)$, $f(x)$, $g(x)$ 为实值函数, $c_1, c_2\in R$,

$$\frac{\partial\left(c_1 f(x)+c_2 g(x)\right)}{\partial x}=c_1\frac{\partial f(x)}{\partial x}+c_2\frac{\partial g(x)}{\partial x}$$

$$\frac{\partial f(x)g(x)}{\partial x}=f(x)\frac{\partial g(x)}{\partial x}+g(x)\frac{\partial f(x)}{\partial x}$$

(3) $x\,(m\times1),\ y(x)\,(n\times1),\ f(y)$为实值函数：

$$\frac{\partial f\left(y(x)\right)}{\partial x}=\frac{\partial y^T(x)}{\partial x}\frac{\partial f(y)}{\partial y}$$

(4) $x\,(m\times1)$： $\dfrac{\partial x^T}{\partial x}=I$

(5) $x,a\,(m\times1)$： $\dfrac{\partial a^T x}{\partial x}=\dfrac{\partial x^T a}{\partial x}=a$

(6) $x\,(m\times1);\ a,\ y(x)\,(n\times1)$： $\dfrac{\partial a^T y(x)}{\partial x}=\dfrac{\partial y^T(x)}{\partial x}a$

(7) $x\ (m \times 1),\ A\ (m \times m):\quad \frac{\partial x^T A x}{\partial x} = (A + A^T)x$

(8) $x\ (m \times 1),\ y(x)\ (n \times 1),\ A\ (n \times n):$

$$\frac{\partial y^T(x) A y(x)}{\partial x} = \frac{\partial y^T(x)}{\partial x}(A + A^T)y(x)$$

(9) $x\ (m \times 1),\ y(x)\ (n \times 1),\ z(x)\ (p \times 1),\ A\ (n \times p):$

$$\frac{\partial y^T(x) A z(x)}{\partial x} = \frac{\partial y^T(x)}{\partial x} A z(x) + \frac{\partial z^T(x)}{\partial x} A^T y(x)$$

$$\frac{\partial y^T(x)Az(x)}{\partial x} = \frac{\partial y^T(x)}{\partial x}Az(x) + \frac{\partial z^T(x)}{\partial x}A^T y(x)$$

$$\left[\frac{\partial y^T(x)Az(x)}{\partial x}\right]_k = \frac{\partial\left(\sum_{i=1}^{n}\sum_{j=1}^{p} a_{ij} y_i(x) z_j(x)\right)}{\partial x_k}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ij}\frac{\partial\left(y_i(x)\right)}{\partial x_k} z_j(x) + \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ij} y_i(x)\frac{\partial\left(z_j(x)\right)}{\partial x_k}$$

$$= \left[\frac{\partial y^T(x)}{\partial x}Az(x) + \frac{\partial z^T(x)}{\partial x}A^T y(x)\right]_k$$