矩阵分析简介

矩阵微分——相对于数量变量的微分

对于n维函数向量 $\mathbf{a}(t)=\left(a_1(t),a_2(t),\cdots a_n(t)\right)^T$

对数量变量 t 的导数为：

$$\frac{d\mathbf{a}}{dt}\triangleq\left(\frac{da_1}{dt},\frac{da_2}{dt},\cdots,\frac{da_n}{dt}\right)^T$$

对于m×n阶函数矩阵 $\mathbf{A}(t)=\left(a_{i,j}(t)\right)_{m\times n}$

对数量变量 t 的导数为：

$$\frac{d\mathbf{A}}{dt}\triangleq\left(\frac{da_{i,j}}{dt}\right)_{m\times n}$$

矩阵微分——相对于数量变量的微分

设 $\mathbf{A}(t),\mathbf{B}(t),\mathbf{A}^{-1}(t)$ 及数量函数 $\lambda(t)$ 对数量变量t均可导，则：

1) $$\frac{d(\mathbf{A}\pm\mathbf{B})}{dt}=\frac{d\mathbf{A}}{dt}\pm\frac{d\mathbf{B}}{dt} \qquad \forall \mathbf{A}、\mathbf{B}\in \mathrm{F}^{m\times n}$$

2) $$\frac{d\lambda\mathbf{A}}{dt}=\frac{d\lambda}{dt}\mathbf{A}+\lambda\frac{d\mathbf{A}}{dt} \qquad \forall \mathbf{A}\in \mathrm{F}^{m\times n}$$

3) $$\frac{d\mathbf{AB}}{dt}=\frac{d\mathbf{A}}{dt}\mathbf{B}+\mathbf{A}\frac{d\mathbf{B}}{dt} \qquad \forall \mathbf{A}\in \mathrm{F}^{m\times n},\mathbf{B}\in \mathrm{F}^{n\times p}$$

4) $$\frac{d\mathbf{A}^{-1}}{dt}=-\mathbf{A}^{-1}\frac{d\mathbf{A}}{dt}\mathbf{A}^{-1} \qquad \forall \mathbf{A}\in \mathrm{F}_n^{n\times n}$$

3) $$\frac{d\mathbf{AB}}{dt}=\frac{d\mathbf{A}}{dt}\mathbf{B}+\mathbf{A}\frac{d\mathbf{B}}{dt} \qquad \forall \mathbf{A}\in F^{m\times n}, \mathbf{B}\in F^{n\times p}$$

第一步：将AB以转换为向量乘积矩阵

第二步：证明向量乘积求导公式：

$$\mathbf{AB}=\left(\mathbf{a}_i^T\mathbf{b}_j\right)_{m\times p}$$

$$\frac{d(\mathbf{a}^T\mathbf{b})}{dt}=\frac{d\mathbf{a}^T}{dt}\mathbf{b}+\mathbf{a}^T\frac{d\mathbf{b}}{dt}$$

$$\frac{d\mathbf{AB}}{dt}=\left(\frac{d\mathbf{a}_i^T\mathbf{b}_j}{dt}\right)_{m\times p}=\frac{d\mathbf{A}}{dt}\mathbf{B}+\mathbf{A}\frac{d\mathbf{B}}{dt}$$

第三步：代入乘积矩阵，得出求导结果

第一步：AB以转换为向量乘积矩阵：

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \\ \mathbf{a}_4^T \\ \mathbf{a}_5^t \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{pmatrix} = \begin{pmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 & \mathbf{b}_4 \end{pmatrix}$$

$$\mathbf{AB} = \begin{pmatrix} \mathbf{a}_1^T\mathbf{b}_1 & \mathbf{a}_1^T\mathbf{b}_2 & \mathbf{a}_1^T\mathbf{b}_3 & \mathbf{a}_1^T\mathbf{b}_4 \\ \mathbf{a}_2^T\mathbf{b}_1 & \mathbf{a}_2^T\mathbf{b}_2 & \mathbf{a}_2^T\mathbf{b}_3 & \mathbf{a}_2^T\mathbf{b}_4 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_5^T\mathbf{b}_1 & \mathbf{a}_5^T\mathbf{b}_2 & \mathbf{a}_5^T\mathbf{b}_3 & \mathbf{a}_5^T\mathbf{b}_4 \end{pmatrix} = \left(\mathbf{a}_i^T\mathbf{b}_j\right)_{m\times p}$$

第二步：向量乘积求导

$$\frac{d(\mathbf{a}^T\mathbf{b})}{dt} = \frac{d\sum_n a_n b_n}{dt} = \sum_n \frac{d(a_n b_n)}{dt}$$

$$= \sum_n \left(\frac{da_n}{dt} b_n + a_n \frac{db_n}{dt} \right) = \sum_n \left(\frac{da_n}{dt} b_n \right) + \sum_n \left(a_n \frac{db_n}{dt} \right)$$

$$= \frac{d\mathbf{a}^T}{dt}\mathbf{b} + \mathbf{a}^T \frac{d\mathbf{b}}{dt}$$

第三步：代入乘积矩阵，得出求导结果

$$\mathbf{AB} = \left(\mathbf{a}_i^t \mathbf{b}_j\right)_{m\times p} \qquad \frac{d(\mathbf{a}^t\mathbf{b})}{dt} = \frac{d\mathbf{a}^t}{dt}\mathbf{b} + \mathbf{a}^t\frac{d\mathbf{b}}{dt}$$

$$\frac{d\mathbf{AB}}{dt} = \left(\frac{d\mathbf{a}_i^t\mathbf{b}_j}{dt}\right)_{m\times p} = \left(\frac{d\mathbf{a}^t}{dt}\mathbf{b} + \mathbf{a}^t\frac{d\mathbf{b}}{dt}\right)_{m\times p}$$

$$= \left(\frac{d\mathbf{a}^t}{dt}\mathbf{b}\right)_{m\times p} + \left(\mathbf{a}^t\frac{d\mathbf{b}}{dt}\right)_{m\times p} = \frac{d\mathbf{A}}{dt}\mathbf{B} + \mathbf{A}\frac{d\mathbf{B}}{dt}$$

矩阵微分——相对于数量变量的微分

例:求二次型 $\mathbf{x}^T A\mathbf{x}$ 对t 的导数

其中 $\mathbf{x}=\mathbf{x}(t)$ 是n维函数向量$A=A^T\in R^{n\times n}$是数字矩阵

矩阵微分——相对于向量变量的微分

1，数量函数的导数

2，函数向量的导数

矩阵微分——相对于向量变量的微分

1，数量函数的导数

设 $f(\mathbf{x}) = g(x_1, x_2, \cdots x_n)$ 是以向量$\mathbf{x} = (x_1, x_2, \cdots x_n)^t$ 为自变量的数量函数，定义

f成绩=g(出勤，作业，考试，报告)

$$\frac{df(\mathbf{x})}{d\mathbf{x}} \triangleq \begin{pmatrix} \dfrac{\partial g}{\partial x_1} \\ \dfrac{\partial g}{\partial x_2} \\ \vdots \\ \dfrac{\partial g}{\partial x_n} \end{pmatrix} = \left(\frac{\partial}{\partial x_i} \right)_{n\times 1} \otimes g \; = \underline{\nabla g(x_1, x_2, \cdots x_n)}$$

数量场的梯度

矩阵微分——相对于向量变量的微分

1，数量函数的导数

设 $f(\mathbf{x}) = g(x_1, x_2, \cdots x_n)$ 是以向量 $\mathbf{x} = (x_1, x_2, \cdots x_n)^t$ 为自变量的数量函数，定义

$$\frac{df(\mathbf{x})}{d\mathbf{x}^t} \triangleq \begin{pmatrix} \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} & \cdots & \frac{\partial g}{\partial x_n} \end{pmatrix}$$

数量函数，对列向量的导数 ——列向量

对行向量的导数 ——行向量

矩阵微分——相对于向量变量的微分

1，数量函数的导数

设 $f(\mathbf{x}) = g(x_1, x_2, \cdots x_n)$ ，$h(\mathbf{x}) = k(x_1, x_2, \cdots x_n)$ ，有：

$$\frac{d(f(\mathbf{x}) \pm h(\mathbf{x}))}{d\mathbf{x}} = \frac{df(\mathbf{x})}{d\mathbf{x}} \pm \frac{dh(\mathbf{x})}{d\mathbf{x}}$$

$$\frac{d\left(f(\mathbf{x})h(\mathbf{x})\right)}{d\mathbf{x}} = \frac{df(\mathbf{x})}{d\mathbf{x}}h(\mathbf{x}) + f(\mathbf{x})\frac{dh(\mathbf{x})}{d\mathbf{x}}$$

矩阵微分——相对于向量变量的微分

1，数量函数的导数

例：求函数 $f(\mathbf{x})=\mathbf{x}^t\mathbf{x}$ 对 $\mathbf{x}$ 的导数

矩阵微分——相对于向量变量的微分

2，函数向量的导数

设 $\mathbf{x}=(x_1,x_2,\cdots x_n)^t$ 且 $\mathbf{a}(\mathbf{x})=\left(\alpha_1(\mathbf{x}),\alpha_2(\mathbf{x}),\cdots\alpha_m(\mathbf{x})\right)^t$

是向量 $\mathbf{x}$ 的函数向量，定义：

$$\frac{d\mathbf{a}(\mathbf{x})}{d\mathbf{x}^t}\triangleq\begin{pmatrix}\frac{\partial\alpha_1}{\partial x_1} & \frac{\partial\alpha_1}{\partial x_2} & \cdots & \frac{\partial\alpha_1}{\partial x_n}\\ \frac{\partial\alpha_2}{\partial x_1} & \frac{\partial\alpha_2}{\partial x_2} & \cdots & \frac{\partial\alpha_2}{\partial x_n}\\ & \vdots & & \\ \frac{\partial\alpha_m}{\partial x_1} & \frac{\partial\alpha_m}{\partial x_2} & \cdots & \frac{\partial\alpha_m}{\partial x_n}\end{pmatrix}_{m\times n}$$

$$\frac{d\mathbf{a}(\mathbf{x})^t}{d\mathbf{x}}\triangleq\left(\frac{d\mathbf{a}(\mathbf{x})}{d\mathbf{x}^t}\right)^t$$

列对行、行对列

矩阵微分——相对于向量变量的微分

2，函数向量的导数

设 $\mathbf{a}(\mathbf{x}),\mathbf{b}(\mathbf{x})\in F^m,\mathbf{x}\in F^n,\lambda(\mathbf{x})\in F$ 则

1) $$\frac{d(\mathbf{a}\pm\mathbf{b})}{d\mathbf{x}^t}=\frac{d\mathbf{a}}{d\mathbf{x}^t}\pm\frac{d\mathbf{b}}{d\mathbf{x}^t}$$

单独一个向量，通常特指列向量

2) $$\frac{d\lambda\mathbf{a}}{d\mathbf{x}^t}=\frac{d\lambda}{d\mathbf{x}^t}\mathbf{a}+\lambda\frac{d\mathbf{a}}{d\mathbf{x}^t}$$

3) $$\frac{d\mathbf{a}^t\mathbf{b}}{d\mathbf{x}^t}=\mathbf{b}^t\frac{d\mathbf{a}}{d\mathbf{x}^t}+\mathbf{a}^t\frac{d\mathbf{b}}{d\mathbf{x}^t}$$

a^tb 求导后的形式还是p^tq

4) $$\frac{d\mathbf{a}^t\mathbf{b}}{d\mathbf{x}}=\frac{d\mathbf{a}^t}{d\mathbf{x}}\mathbf{b}+\frac{d\mathbf{b}^t}{d\mathbf{x}}\mathbf{a}$$

行列相对

矩阵微分——相对于向量变量的微分

$$\frac{d\mathbf{a}^t\mathbf{b}}{d\mathbf{x}^t} = \left(\frac{\partial(\mathbf{a}^t\mathbf{b})}{\partial x_1}, \frac{\partial(\mathbf{a}^t\mathbf{b})}{\partial x_2}, \cdots, \frac{\partial(\mathbf{a}^t\mathbf{b})}{\partial x_n}\right)$$

$$\frac{d(\mathbf{a}^t\mathbf{b})}{dt} = \frac{d\mathbf{a}^t}{dt}\mathbf{b} + \mathbf{a}^t\frac{d\mathbf{b}}{dt}$$

$$= \left(\frac{\partial\mathbf{a}^t}{\partial x_1}\mathbf{b} + \mathbf{a}^t\frac{\partial\mathbf{b}}{\partial x_1}, \cdots, \frac{\partial\mathbf{a}^t}{\partial x_n}\mathbf{b} + \mathbf{a}^t\frac{\partial\mathbf{b}}{\partial x_n}\right)$$

$$= \left(\frac{\partial\mathbf{a}^t}{\partial x_1}\mathbf{b}, \cdots, \frac{\partial\mathbf{a}^t}{\partial x_n}\mathbf{b}\right) + \left(\mathbf{a}^t\frac{\partial\mathbf{b}}{\partial x_1}, \cdots, \mathbf{a}^t\frac{\partial\mathbf{b}}{\partial x_n}\right)$$

$$= \left(\frac{d\mathbf{a}^t}{d\mathbf{x}}\mathbf{b}\right)^t + \mathbf{a}^t\frac{d\mathbf{b}}{d\mathbf{x}^t} = \mathbf{b}^t\frac{d\mathbf{a}}{d\mathbf{x}^t} + \mathbf{a}^t\frac{d\mathbf{b}}{d\mathbf{x}^t}$$

矩阵微分——相对于向量变量的微分

2，函数向量的导数

例：**a**）求行向量 $\mathbf{x}^t A$ 对 $\mathbf{x}$ 的导数

b）求列向量 $B\mathbf{x}$ 对 $\mathbf{x}^t$ 的导数

c）求二次型 $\mathbf{x}^t A\mathbf{x}$ 对 $\mathbf{x}$ 的导数

d）求数量函数 $\mathbf{p}^t A\mathbf{x}$ 对 $\mathbf{x}$ 的导数

矩阵微分——复合函数微分

1，数量函数的求导公式

2，向量函数的求导公式

矩阵微分——复合函数微分

1，数量函数的求导公式

设$f = f(\mathbf{y}), \mathbf{y} = \mathbf{y}(t)$, 若$f, t$为数量变量，$\mathbf{y}$为向量变量，则：

$$\frac{df}{dt} = \frac{df}{d\mathbf{y}^t}\frac{d\mathbf{y}}{dt} = \frac{d\mathbf{y}^t}{dt}\frac{df}{d\mathbf{y}}$$

设$f = f(\mathbf{y}), \mathbf{y} = \mathbf{y}(\mathbf{x})$, 若$f$为数量变量，$\mathbf{y}, \mathbf{x}$为向量变量，则：

否则，倒着写

$$\frac{df}{d\mathbf{x}} = \frac{d\mathbf{y}^t}{d\mathbf{x}}\frac{df}{d\mathbf{y}} \qquad \frac{df}{d\mathbf{x}^t} = \frac{df}{d\mathbf{y}^t}\frac{d\mathbf{y}}{d\mathbf{x}^t}$$

微分行，正常，顺着写

设$f = f(\mathbf{x}, \mathbf{y}), \mathbf{y} = \mathbf{y}(\mathbf{x})$, 若$f$为数量变量，$\mathbf{y}, \mathbf{x}$为向量变量，则：

$$\frac{df}{d\mathbf{x}} = \frac{\partial f}{\partial \mathbf{x}} + \frac{d\mathbf{y}^t}{d\mathbf{x}}\frac{\partial f}{\partial \mathbf{y}} \qquad \frac{df}{d\mathbf{x}^t} = \frac{\partial f}{\partial \mathbf{x}^t} + \frac{\partial f}{\partial \mathbf{y}^t}\frac{d\mathbf{y}}{d\mathbf{x}^t}$$

矩阵微分——复合函数微分

1，数量函数的求导公式

例：求$f=(\mathbf{x}-\mathbf{a}-B\mathbf{z})^t(\mathbf{x}-\mathbf{a}-B\mathbf{z})$对$\mathbf{a}$的导数：

矩阵微分——复合函数微分

2， 向量函数的求导公式

设$\mathbf{z}=\mathbf{z}(\mathbf{y}),\mathbf{y}=\mathbf{y}(t)$,若$t$为数量变量， $\mathbf{y},\mathbf{z}$为向量变量， 则：

$$\frac{d\mathbf{z}}{dt}=\frac{d\mathbf{z}}{d\mathbf{y}^t}\frac{d\mathbf{y}}{dt}$$

设$\mathbf{z}=\mathbf{z}(\mathbf{y}),\mathbf{y}=\mathbf{y}(\mathbf{x})$,若$\mathbf{x},\mathbf{y},\mathbf{z}$为向量变量， 则：

否则， 倒着写

$$\frac{d\mathbf{z}^t}{d\mathbf{x}}=\frac{d\mathbf{y}^t}{d\mathbf{x}}\frac{d\mathbf{z}^t}{d\mathbf{y}} \qquad \frac{d\mathbf{z}}{d\mathbf{x}^t}=\frac{d\mathbf{z}}{d\mathbf{y}^t}\frac{d\mathbf{y}}{d\mathbf{x}^t}$$

微分行， 正常，顺着写

设$\mathbf{z}=\mathbf{z}(\mathbf{x},\mathbf{y}),\mathbf{y}=\mathbf{y}(\mathbf{x})$,若$\mathbf{x},\mathbf{y},\mathbf{z}$为向量变量， 则：

$$\frac{d\mathbf{z}^t}{d\mathbf{x}}=\frac{\partial\mathbf{z}^t}{\partial\mathbf{x}}+\frac{d\mathbf{y}^t}{d\mathbf{x}}\frac{\partial\mathbf{z}^t}{\partial\mathbf{y}} \qquad \frac{d\mathbf{z}}{d\mathbf{x}^t}=\frac{\partial\mathbf{z}}{\partial\mathbf{x}^t}+\frac{\partial f}{\partial\mathbf{y}^t}\frac{d\mathbf{z}}{d\mathbf{x}^t}$$

矩阵微分——复合函数微分

2，向量函数的求导公式

例：求$f=(\mathbf{Ax}-\mathbf{b})^t\mathbf{R}(\mathbf{Ax}-\mathbf{b})$对$\mathbf{x}$的导数

其中，$\mathbf{A},\mathbf{R}$为常数矩阵，$\mathbf{b}$是常数向量

书面作业，全体都有，下周一交

1: 求二次型 $\mathbf{x}^T A\mathbf{x}$ 对标量$\boldsymbol{t}$ 的导数

其中 $\mathbf{x}=\mathbf{x}(t)$ 是n维函数向量 $A=A^t\in R^{n\times n}$是数字矩阵

2 $f(\mathbf{x})=g(x_1,x_2,\cdots x_n)$,$h(\mathbf{x})=k(x_1,x_2,\cdots x_n)$为标量函数

2.1证明：
$$\frac{d\left(f(\mathbf{x})h(\mathbf{x})\right)}{d\mathbf{x}}=\frac{df(\mathbf{x})}{d\mathbf{x}}h(\mathbf{x})+f(\mathbf{x})\frac{dh(\mathbf{x})}{d\mathbf{x}}$$

2.2 求函数 $f(\mathbf{x})=\mathbf{x}^t\mathbf{x}$ 对 $\mathbf{x}$ 的导数

2.3 求$f=(\mathbf{x}-\mathbf{a}-B\mathbf{z})^t(\mathbf{x}-\mathbf{a}-B\mathbf{z})$对$\mathbf{a}$的导数：

4，函数向量的导数

求：a）求行向量 $\mathbf{x}^t A$ **对** $\mathbf{x}$ **的导数**

b）求列向量 $B\mathbf{x}$ **对** $\mathbf{x}^t$ **的导数**

c）求二次型 $\mathbf{x}^t A\mathbf{x}$ **对** $\mathbf{x}$ **的导数**

d）求数量函数 $\mathbf{p}^t A\mathbf{x}$ **对** $\mathbf{x}$ **的导数**

5 向量函数的求导 求 $f = (\mathbf{Ax}-\mathbf{b})^t \mathbf{R}(\mathbf{Ax}-\mathbf{b})$ 对 $\mathbf{x}$ 的导数

其中，$\mathbf{A},\mathbf{R}$ 为常数矩阵，$\mathbf{b}$ 是常数向量

6，求以下三个代价函数J对于变量a的导数：

$$J(\mathbf{a}) = \sum_{\mathbf{y}\in\mathcal{Y}} \left(-\mathbf{a}^t\mathbf{y}\right), \quad J(\mathbf{a}) = \sum_{\mathbf{y}\in Y} \frac{(\mathbf{a}^t\mathbf{y}-b)^2}{\|\mathbf{y}\|^2} \quad J(\mathbf{a}) = \|\mathbf{Ya}-\mathbf{b}\|^2$$