矩阵、向量求导法则

1. 行向量对元素求导

设 $\boldsymbol{y}^T = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}$ 是 n 维行向量，x 是元素，则

$$\frac{\partial \mathbf{y}^T}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} & \cdots & \frac{\partial y_n}{\partial x} \end{bmatrix} \tag{1}$$

2. 列向量对元素求导

设 $\boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$ 是 m 维列向量，x 是元素，则

$$\frac{\partial \mathbf{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \cdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix} \tag{2}$$

3. 矩阵对元素求导

设 $Y = \begin{bmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{m1} & \cdots & y_{mn} \end{bmatrix}$ 是 $m \times n$ 矩阵，x 是元素，则

$$\frac{\partial Y}{\partial x} = \begin{bmatrix} \frac{\partial y_{11}}{\partial x} & \cdots & \frac{\partial y_{1n}}{\partial x} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_{m1}}{\partial x} & \cdots & \frac{\partial y_{mn}}{\partial x} \end{bmatrix} \tag{3}$$

4. 元素对行向量求导

设 y 是元素，$\boldsymbol{x}^T = \begin{bmatrix} x_1 & \cdots & x_q \end{bmatrix}$ 是 q 维行向量，则

$$\frac{\partial y}{\partial \mathbf{x}^T} = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \cdots & \frac{\partial y}{\partial x_q} \end{bmatrix} \tag{4}$$

5. 元素对列向量求导

设 y 是元素，$\boldsymbol{x}=\begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$ 是 p 维列向量，则

$$\frac{\partial y}{\partial \boldsymbol{x}}=\begin{bmatrix} \dfrac{\partial y}{\partial x_1} \\ \vdots \\ \dfrac{\partial y}{\partial x_p} \end{bmatrix} \tag{5}$$

6. 元素对矩阵求导

设 y 是元素，$X=\begin{bmatrix} x_{11} & \cdots & x_{1q} \\ \vdots & \ddots & \vdots \\ x_{p1} & \cdots & x_{pq} \end{bmatrix}$ 是 $p\times q$ 矩阵，则

$$\frac{\partial y}{\partial X}=\begin{bmatrix} \dfrac{\partial y}{\partial x_{11}} & \cdots & \dfrac{\partial y}{\partial x_{1q}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial y}{\partial x_{p1}} & \cdots & \dfrac{\partial y}{\partial x_{pq}} \end{bmatrix} \tag{6}$$

7. 行向量对列向量求导

设 $\boldsymbol{y}^T=\begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}$ 是 n 维行向量，$\boldsymbol{x}=\begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$ 是 p 维列向量，则

$$\frac{\partial \boldsymbol{y}^T}{\partial \boldsymbol{x}}=\begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_n}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial y_1}{\partial x_p} & \cdots & \dfrac{\partial y_n}{\partial x_p} \end{bmatrix} \tag{7}$$

8. 列向量对行向量求导

设 $\boldsymbol{y}=\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$ 是 m 维列向量，$\boldsymbol{x}^T=\begin{bmatrix} x_1 & \cdots & x_q \end{bmatrix}$ 是 q 维行向量，则

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}^T}=\begin{bmatrix}\frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_q}\\ \vdots & \ddots & \vdots\\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_q}\end{bmatrix} \tag{8}$$

9. 行向量对行向量求导

设 $\boldsymbol{y}^T=\begin{bmatrix}y_1 & \cdots & y_n\end{bmatrix}$是 n 维行向量，$\boldsymbol{x}^T=\begin{bmatrix}x_1 & \cdots & x_q\end{bmatrix}$是 q 维行向量，则

$$\frac{\partial \boldsymbol{y}^T}{\partial \boldsymbol{x}^T}=\begin{bmatrix}\frac{\partial \boldsymbol{y}^T}{\partial x_1} & \cdots & \frac{\partial \boldsymbol{y}^T}{\partial x_q}\end{bmatrix} \tag{9}$$

10. 列向量对列向量求导

设 $\boldsymbol{y}=\begin{bmatrix}y_1\\ \vdots\\ y_m\end{bmatrix}$是 m 维列向量，$\boldsymbol{x}=\begin{bmatrix}x_1\\ \vdots\\ x_p\end{bmatrix}$是 p 维列向量，则

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}}=\begin{bmatrix}\frac{\partial y_1}{\partial \boldsymbol{x}}\\ \vdots\\ \frac{\partial y_m}{\partial \boldsymbol{x}}\end{bmatrix} \tag{10}$$

11. 矩阵对行向量求导

设 $Y=\begin{bmatrix}y_{11} & \cdots & y_{1n}\\ \vdots & \ddots & \vdots\\ y_{m1} & \cdots & y_{mn}\end{bmatrix}$是 $m\times n$ 矩阵，$\boldsymbol{x}^T=\begin{bmatrix}x_1 & \cdots & x_q\end{bmatrix}$是 q 维行向量，则

$$\frac{\partial Y}{\partial \boldsymbol{x}^T}=\begin{bmatrix}\frac{\partial Y}{\partial x_1} & \cdots & \frac{\partial Y}{\partial x_q}\end{bmatrix} \tag{11}$$

12. 矩阵对列向量求导

设 $Y=\begin{bmatrix}y_{11} & \cdots & y_{1n}\\ \vdots & \ddots & \vdots\\ y_{m1} & \cdots & y_{mn}\end{bmatrix}$是 $m\times n$ 矩阵，$\boldsymbol{x}=\begin{bmatrix}x_1\\ \vdots\\ x_p\end{bmatrix}$是 p 维列向量，则

$$\frac{\partial Y}{\partial \boldsymbol{x}}=\begin{bmatrix} \frac{\partial y_{11}}{\partial \boldsymbol{x}} & \cdots & \frac{\partial y_{1n}}{\partial \boldsymbol{x}} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_{m1}}{\partial \boldsymbol{x}} & \cdots & \frac{\partial y_{mn}}{\partial \boldsymbol{x}} \end{bmatrix} \tag{12}$$

13. 行向量对矩阵求导

设 $\boldsymbol{y}^T=\begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}$ 是 n 维行向量，$X=\begin{bmatrix} x_{11} & \cdots & x_{1q} \\ \vdots & \ddots & \vdots \\ x_{p1} & \cdots & x_{pq} \end{bmatrix}$ 是 $p\times q$ 矩阵，则

$$\frac{\partial \boldsymbol{y}^T}{\partial X}=\begin{bmatrix} \frac{\partial \boldsymbol{y}^T}{\partial x_{11}} & \cdots & \frac{\partial \boldsymbol{y}^T}{\partial x_{1q}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \boldsymbol{y}^T}{\partial x_{p1}} & \cdots & \frac{\partial \boldsymbol{y}^T}{\partial x_{pq}} \end{bmatrix} \tag{13}$$

14. 列向量对矩阵求导

设 $\boldsymbol{y}=\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$ 是 m 维列向量，$X=\begin{bmatrix} x_{11} & \cdots & x_{1q} \\ \vdots & \ddots & \vdots \\ x_{p1} & \cdots & x_{pq} \end{bmatrix}$ 是 $p\times q$ 矩阵，则

$$\frac{\partial \boldsymbol{y}}{\partial X}=\begin{bmatrix} \frac{\partial y_1}{\partial X} \\ \vdots \\ \frac{\partial y_m}{\partial X} \end{bmatrix} \tag{14}$$

15. 矩阵对矩阵求导

设 $Y=\begin{bmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{m1} & \cdots & y_{mn} \end{bmatrix}=\begin{bmatrix} \boldsymbol{y}_1^T \\ \vdots \\ \boldsymbol{y}_m^T \end{bmatrix}$ 是 $m\times n$ 矩阵，$X=\begin{bmatrix} x_{11} & \cdots & x_{1q} \\ \vdots & \ddots & \vdots \\ x_{p1} & \cdots & x_{pq} \end{bmatrix}=\begin{bmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_p \end{bmatrix}$ 是 $p\times q$ 矩阵，则

$$\frac{\partial Y}{\partial X}=\begin{bmatrix} \frac{\partial Y}{\partial \boldsymbol{x}_1} & \cdots & \frac{\partial Y}{\partial \boldsymbol{x}_q} \end{bmatrix}=\begin{bmatrix} \frac{\partial \boldsymbol{y}_1^T}{\partial X} \\ \vdots \\ \frac{\partial \boldsymbol{y}_m^T}{\partial X} \end{bmatrix}==\begin{bmatrix} \frac{\partial \boldsymbol{y}_1^T}{\partial \boldsymbol{x}_1} & \cdots & \frac{\partial \boldsymbol{y}_1^T}{\partial \boldsymbol{x}_q} \\ \vdots & \ddots & \vdots \\ \frac{\partial \boldsymbol{y}_m^T}{\partial \boldsymbol{x}_1} & \cdots & \frac{\partial \boldsymbol{y}_m^T}{\partial \boldsymbol{x}_q} \end{bmatrix} \tag{15}$$

例 1

设 $\frac{\partial A}{\partial X}=\begin{bmatrix}2xy & y^2 & y\\ x^2 & 2xy & x\end{bmatrix}$， $X=\begin{bmatrix}x\\ y\end{bmatrix}$，根据矩阵对列向量求导法则式(12)，有

$$\frac{\partial^2 A}{\partial X^2}=\begin{bmatrix}\frac{\partial(2xy)}{\partial X} & \frac{\partial\left(y^2\right)}{\partial X} & \frac{\partial y}{\partial X}\\ \frac{\partial\left(x^2\right)}{\partial X} & \frac{\partial(2xy)}{\partial X} & \frac{\partial(x)}{\partial X}\end{bmatrix}=\begin{bmatrix}2y & 0 & 0\\ 2x & 2y & 1\\ 2x & 2y & 1\\ 0 & 2x & 0\end{bmatrix}$$

例 2

设 $Y=\begin{bmatrix}a & b & c\\ d & e & f\end{bmatrix}, X=\begin{bmatrix}u & x\\ v & y\\ w & z\end{bmatrix}$，根据矩阵对矩阵求导法则式(15)，有

$$\frac{\partial Y}{\partial X}=\begin{bmatrix}\frac{\partial\begin{bmatrix}a & b & c\end{bmatrix}}{\partial\begin{bmatrix}u\\ v\\ w\end{bmatrix}} & \frac{\partial\begin{bmatrix}a & b & c\end{bmatrix}}{\partial\begin{bmatrix}x\\ y\\ z\end{bmatrix}}\\ \frac{\partial\begin{bmatrix}d & e & f\end{bmatrix}}{\partial\begin{bmatrix}u\\ v\\ w\end{bmatrix}} & \frac{\partial\begin{bmatrix}d & e & f\end{bmatrix}}{\partial\begin{bmatrix}x\\ y\\ z\end{bmatrix}}\end{bmatrix}=\begin{bmatrix}\frac{\partial a}{\partial u} & \frac{\partial b}{\partial u} & \frac{\partial c}{\partial u} & \frac{\partial a}{\partial x} & \frac{\partial b}{\partial x} & \frac{\partial c}{\partial x}\\ \frac{\partial a}{\partial v} & \frac{\partial b}{\partial v} & \frac{\partial c}{\partial v} & \frac{\partial a}{\partial y} & \frac{\partial b}{\partial y} & \frac{\partial c}{\partial y}\\ \frac{\partial a}{\partial w} & \frac{\partial b}{\partial w} & \frac{\partial c}{\partial w} & \frac{\partial a}{\partial z} & \frac{\partial b}{\partial z} & \frac{\partial c}{\partial z}\\ \frac{\partial d}{\partial u} & \frac{\partial e}{\partial u} & \frac{\partial f}{\partial u} & \frac{\partial d}{\partial x} & \frac{\partial e}{\partial x} & \frac{\partial f}{\partial x}\\ \frac{\partial d}{\partial v} & \frac{\partial e}{\partial v} & \frac{\partial f}{\partial v} & \frac{\partial d}{\partial y} & \frac{\partial e}{\partial y} & \frac{\partial f}{\partial y}\\ \frac{\partial d}{\partial w} & \frac{\partial e}{\partial w} & \frac{\partial f}{\partial w} & \frac{\partial d}{\partial z} & \frac{\partial e}{\partial z} & \frac{\partial f}{\partial z}\end{bmatrix}$$