矩阵、向量求导法则

目录

1. 向量、矩阵对元素求导

1.1 行向量对元素求导

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

1.2 列向量对元素求导

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

1.3 矩阵对元素求导

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

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

2. 元素对向量、矩阵求导

2.1 元素对行向量求导

设 y 是元素，$\mathbf{x}^T = [x_1 \;\; \cdots \;\; x_q]$ 是 q 维行向量，则 $\dfrac{\partial y}{\partial \mathbf{x}^T} = \left[\dfrac{\partial y}{\partial x_1} \;\; \cdots \;\; \dfrac{\partial y}{\partial x_q} \right]$。

2.2 元素对列向量求导

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

2.3 元素对矩阵求导

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

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

3. 向量对向量求导

3.1 行向量对列向量求导

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

$$\frac{\partial \mathbf{y}^T}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_n}{\partial x_1} \\ \vdots & & \\ \dfrac{\partial y_1}{\partial x_p} & \cdots & \dfrac{\partial y_n}{\partial x_p} \end{bmatrix} \text{。}$$

3.2 列向量对行向量求导

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

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}^T} = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_q} \\ \vdots & & \\ \dfrac{\partial y_m}{\partial x_1} & \cdots & \dfrac{\partial y_m}{\partial x_q} \end{bmatrix} \text{。}$$

3.3 行向量对行向量求导

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

$$\frac{\partial \mathbf{y}^T}{\partial \mathbf{x}^T} = \begin{bmatrix} \dfrac{\partial \mathbf{y}^T}{\partial x_1} & \cdots & \dfrac{\partial \mathbf{y}^T}{\partial x_q} \end{bmatrix} \text{。}$$

3.4 列向量对列向量求导

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

4. 矩阵对向量求导

4.1 矩阵对行向量求导

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

$$\frac{\partial Y}{\partial \mathbf{x}^T} = \begin{bmatrix} \dfrac{\partial Y}{\partial x_1} & \cdots & \dfrac{\partial Y}{\partial x_q} \end{bmatrix} \text{。}$$

4.2 矩阵对列向量求导

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

$$\frac{\partial Y}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial y_{11}}{\partial \mathbf{x}} & \cdots & \dfrac{\partial y_{1n}}{\partial \mathbf{x}} \\ \vdots & & \vdots \\ \dfrac{\partial y_{m1}}{\partial \mathbf{x}} & \cdots & \dfrac{\partial y_{mn}}{\partial \mathbf{x}} \end{bmatrix} \text{。}$$

5. 矩阵对向量求导

5.1 行向量对矩阵求导

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

$$\frac{\partial \mathbf{y}^T}{\partial X} = \begin{bmatrix} \dfrac{\partial \mathbf{y}^T}{\partial x_{11}} & \cdots & \dfrac{\partial \mathbf{y}^T}{\partial x_{1q}} \\ \vdots & & \\ \dfrac{\partial \mathbf{y}^T}{\partial x_{p1}} & \cdots & \dfrac{\partial \mathbf{y}^T}{\partial x_{pq}} \end{bmatrix} \circ$$

5.2 列向量对矩阵求导

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

$$\frac{\partial \mathbf{y}}{\partial X} = \begin{bmatrix} \dfrac{\partial y_1}{\partial X} \\ \vdots \\ \dfrac{\partial y_m}{\partial X} \end{bmatrix} \circ$$

5.3 矩阵对矩阵求导

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

$= \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_q \end{bmatrix}$ 是 $p \times q$ 矩阵，则

$$\frac{\partial Y}{\partial X} = \begin{bmatrix} \dfrac{\partial Y}{\partial \mathbf{x}_1} & \cdots & \dfrac{\partial Y}{\partial \mathbf{x}_q} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \mathbf{y}_1^T}{\partial X} \\ \vdots \\ \dfrac{\partial \mathbf{y}_m^T}{\partial X} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \mathbf{y}_1^T}{\partial \mathbf{x}_1} & \cdots & \dfrac{\partial \mathbf{y}_1^T}{\partial \mathbf{x}_q} \\ \vdots & & \vdots \\ \dfrac{\partial \mathbf{y}_m^T}{\partial \mathbf{x}_1} & \cdots & \dfrac{\partial \mathbf{y}_m^T}{\partial \mathbf{x}_q} \end{bmatrix} \circ$$

6. 例

例 设 $\dfrac{\partial A}{\partial X} = \begin{bmatrix} 2xy & y^2 & y \\ x^2 & 2xy & x \end{bmatrix}$，$X = \begin{bmatrix} x \\ y \end{bmatrix}$，根据（12）矩阵对列向量求导

法则，有

$$\frac{\partial^2 A}{\partial X^2} = \begin{bmatrix} \dfrac{\partial(2xy)}{\partial X} & \dfrac{\partial(y^2)}{\partial X} & \dfrac{\partial y}{\partial X} \\[2mm] \dfrac{\partial(x^2)}{\partial X} & \dfrac{\partial(2xy)}{\partial X} & \dfrac{\partial x}{\partial X} \end{bmatrix} = \begin{bmatrix} 2y & 0 & 0 \\ 2x & 2y & 1 \\ 2x & 2y & 1 \\ 0 & 2x & 0 \end{bmatrix} 。$$

例 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（15）矩阵对矩阵求导法则，有

$$\frac{\partial Y}{\partial X} = \begin{bmatrix} \dfrac{\partial[a\ b\ c]}{\partial\begin{bmatrix} u \\ v \\ w \end{bmatrix}} & \dfrac{\partial[a\ b\ c]}{\partial\begin{bmatrix} x \\ y \\ z \end{bmatrix}} \\[6mm] \dfrac{\partial[d\ e\ f]}{\partial\begin{bmatrix} u \\ v \\ w \end{bmatrix}} & \dfrac{\partial[d\ e\ f]}{\partial\begin{bmatrix} x \\ y \\ z \end{bmatrix}} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial a}{\partial u} & \dfrac{\partial b}{\partial u} & \dfrac{\partial c}{\partial u} & \dfrac{\partial a}{\partial x} & \dfrac{\partial b}{\partial x} & \dfrac{\partial c}{\partial x} \\[2mm] \dfrac{\partial a}{\partial v} & \dfrac{\partial b}{\partial v} & \dfrac{\partial c}{\partial v} & \dfrac{\partial a}{\partial y} & \dfrac{\partial b}{\partial y} & \dfrac{\partial c}{\partial y} \\[2mm] \dfrac{\partial a}{\partial w} & \dfrac{\partial b}{\partial w} & \dfrac{\partial c}{\partial w} & \dfrac{\partial a}{\partial z} & \dfrac{\partial b}{\partial z} & \dfrac{\partial c}{\partial z} \\[2mm] \dfrac{\partial d}{\partial u} & \dfrac{\partial e}{\partial u} & \dfrac{\partial f}{\partial u} & \dfrac{\partial d}{\partial x} & \dfrac{\partial e}{\partial x} & \dfrac{\partial f}{\partial x} \\[2mm] \dfrac{\partial d}{\partial v} & \dfrac{\partial e}{\partial v} & \dfrac{\partial f}{\partial v} & \dfrac{\partial d}{\partial y} & \dfrac{\partial e}{\partial y} & \dfrac{\partial f}{\partial y} \\[2mm] \dfrac{\partial d}{\partial w} & \dfrac{\partial e}{\partial w} & \dfrac{\partial f}{\partial w} & \dfrac{\partial d}{\partial z} & \dfrac{\partial e}{\partial z} & \dfrac{\partial f}{\partial z} \end{bmatrix} 。$$