标量方程对向量的导数.

$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}_{m\times 1}$

定义:

1° 分母布局:

$$\frac{\partial f(\vec{y})}{\partial \vec{y}} = \begin{bmatrix} \frac{\partial f(\vec{y})}{\partial y_1} \\ \frac{\partial f(\vec{y})}{\partial y_2} \\ \vdots \\ \frac{\partial f(\vec{y})}{\partial y_m} \end{bmatrix}_{\underset{\circ}{m}\times 1}$$

2° 分子布局:

$$\frac{\partial f(\vec{y})}{\partial \vec{y}} = \left[\frac{\partial f(\vec{y})}{\partial y_1}, \frac{\partial f(\vec{y})}{\partial y_2}, \cdots, \frac{\partial f(\vec{y})}{\partial y_m} \right]_{1\times \underset{\circ}{m}}$$

向量方程对向量的导数.

$\vec{f}(\vec{y}) = \begin{bmatrix} f_1(\vec{y}) \\ f_2(\vec{y}) \\ \vdots \\ f_n(\vec{y}) \end{bmatrix}_{n\times 1}$, $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}_{m\times 1}$

分母布局

$$\frac{\partial \vec{f}(\vec{y})_{n\times 1}}{\partial \vec{y}_{m\times 1}} = \begin{bmatrix} \frac{\partial \vec{f}(\vec{y})}{\partial y_1} \\ \frac{\partial \vec{f}(\vec{y})}{\partial y_2} \\ \vdots \\ \frac{\partial \vec{f}(\vec{y})}{\partial y_m} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1(\vec{y})}{\partial y_1} & \frac{\partial f_2(\vec{y})}{\partial y_1} & \cdots & \frac{\partial f_n(\vec{y})}{\partial y_1} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_1(\vec{y})}{\partial y_m} & \cdots & \cdots & \frac{\partial f_n(\vec{y})}{\partial y_m} \end{bmatrix}_{m\times n}$$

分母布局, 第一步先往分母靠.

分母布局 → $\frac{\partial A\vec{y}}{\partial \vec{y}} = A^T$

验证: $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $A\vec{y} = \begin{bmatrix} a_{11}y_1 + a_{12}y_2 \\ a_{21}y_1 + a_{22}y_2 \end{bmatrix}$ ← 向量方程

$$\frac{\partial A\vec{y}}{\partial \vec{y}} = \begin{bmatrix} \frac{\partial A\vec{y}}{\partial y_1} \\ \frac{\partial A\vec{y}}{\partial y_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial(a_{11}y_1 + a_{12}y_2)}{\partial y_1} & \frac{\partial(a_{21}y_1 + a_{22}y_2)}{\partial y_1} \\ \frac{\partial(a_{11}y_1 + a_{12}y_2)}{\partial y_2} & \frac{\partial(a_{21}y_1 + a_{22}y_2)}{\partial y_2} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} = A^T$$

分母布局 $\dfrac{\partial \vec{y}^T A \vec{y}}{\partial \vec{y}} = A\vec{y} + A^T\vec{y}$ （A对称时 $= 2A\vec{y}$）

验证：$\vec{y_1} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$，$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$，$\vec{y}^T A \vec{y} = [y_1 \ y_2]\begin{bmatrix} a_{11}y_1 + a_{12}y_2 \\ a_{21}y_1 + a_{22}y_2 \end{bmatrix}$

$$= y_1(a_{11}y_1 + a_{12}y_2) + y_2(a_{21}y_1 + a_{22}y_2)$$

$$\frac{\partial \vec{y}^T A \vec{y}}{\partial \vec{y}} = \begin{bmatrix} \dfrac{\partial \vec{y}^T A \vec{y}}{\partial y_1} \\ \dfrac{\partial \vec{y}^T A \vec{y}}{\partial y_2} \end{bmatrix} = \begin{bmatrix} (a_{11}y_1 + a_{12}y_2) + a_{11}y_1 + a_{21}y_2 \\ a_{12}y_1 + (a_{21}y_1 + a_{22}y_2) + a_{22}y_2 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}y_1 + a_{12}y_2 \\ a_{21}y_1 + a_{22}y_2 \end{bmatrix} + \begin{bmatrix} a_{11}y_1 + a_{21}y_2 \\ a_{12}y_1 + a_{22}y_2 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= A\vec{y} + A^T\vec{y}$$

线性回归中，要找到 y_1, y_2 s.t. $J = \sum_{i=1}^{n}[z_i - (y_1 + y_2 x_i)]^2$ 最小

令 $\vec{x} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$，$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ 则 $\hat{\vec{z}} = \vec{x} \cdot \vec{y}$ 又 $\vec{z} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$

$$J = [\vec{z} - \hat{\vec{z}}]^T[\vec{z} - \hat{\vec{z}}] = [\vec{z} - \vec{x}\vec{y}]^T[\vec{z} - \vec{x}\vec{y}]$$

$$= [\vec{z}^T - \vec{y}^T\vec{x}^T] \cdot [\vec{z} - \vec{x}\vec{y}]$$

$$= \vec{z}^T\vec{z} - \vec{y}^T\vec{x}^T\vec{z} - \vec{z}^T\vec{x}\vec{y} + \vec{y}^T\vec{x}^T\vec{x}\vec{y}$$

又∵ $\vec{y}^T\vec{x}^T\vec{z}$ 与 $\vec{z}^T\vec{x}\vec{y}$ 均为1×1矩阵 且 $(\vec{y}^T\vec{x}^T\vec{z})^T = \vec{z}^T\vec{x}\vec{y}$ ∴ $\vec{y}^T\vec{x}^T\vec{z} = \vec{z}^T\vec{x}\vec{y}$

故 $J = \vec{z}^T\vec{z} - 2\vec{z}^T\vec{x}\vec{y} + \vec{y}^T\underbrace{\vec{x}^T\vec{x}}_{\text{对称}}\vec{y}$

∴ $\dfrac{\partial J}{\partial \vec{y}} = 0 - 2(\vec{z}^T\vec{x})^T + 2\vec{x}^T\vec{x}\vec{y} = 0$ 时，有 $\vec{y} = (\vec{x}^T\vec{x})^{-1}\vec{x}^T\vec{z}$