SHAPES

$y:$ $\hat{y}:$

$$\begin{bmatrix} \\ \\ \\ \\ \end{bmatrix}_{N \times 1} \quad \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix}_{N \times 1}$$

$$x_i \begin{bmatrix} \\ \\ \end{bmatrix}_{D \times 1} \Rightarrow X \in R^{N \times D}$$

$$\theta \begin{bmatrix} \\ \end{bmatrix}_{D \times 1}$$

$\therefore$ Rewriting ③ in vector form

$$\frac{\partial J(\theta)}{\partial \theta} = \underset{D \times N}{X^T} * \underset{N \times 1}{(\hat{y} - y)}$$