

1.

 $h_\theta: \mathbb{R}^d \rightarrow \mathbb{R}$, where

$$h_\theta(x) = \theta^T x$$

 $\theta, x \in \mathbb{R}^d$

$$\theta: d \times 1$$

$$x: d \times 1$$

$$J(\theta) = \frac{1}{m} \sum_{i=1}^m (h_\theta(x_i) - y_i)^2$$

$$\text{令 } e = h_\theta(x_i) - y_i$$

$$J(\theta) = \frac{1}{m} e^T e \quad \text{标量形式}$$

$$J(\theta) = \frac{1}{m} (h_\theta(x_i) - y_i)^T (h_\theta(x_i) - y_i)$$

$$J(\theta) = \frac{1}{m} (\theta^T x - y_i)^T (\theta^T x - y_i)$$

$$J(\theta) = \frac{1}{m} (x^T \theta - y^T) (\theta^T x - y)$$

$$J(\theta) = \frac{1}{m} (x^T \theta \theta^T x - x^T \theta y - y^T \theta^T x + y^T y)$$

$$\begin{aligned}
 x_i &\Rightarrow d \times 1 \Rightarrow \begin{bmatrix} x_1 \\ \vdots \\ x_d \end{bmatrix} \\
 x_i &= m \times d \\
 \theta &= d \times 1 \Rightarrow \text{symmetric} \\
 \theta &= m \times m
 \end{aligned}$$

$$\theta^T x = \text{标量}$$

$$\begin{aligned}
 &\Rightarrow m \times m \times m \times 1 = m \times 1 \\
 &\Rightarrow m \times m \times m \times 1
 \end{aligned}$$

$$J(\theta) = \frac{1}{m} (x^T \theta \theta^T x - 2y^T \theta^T x + y^T y)$$

$$2. \quad \nabla J(\theta) = \frac{1}{m} \frac{\partial (x^T \theta \theta^T x - 2y^T \theta^T x + y^T y)}{\partial \theta}$$

$$= \frac{2}{m} \times \frac{\partial (x^T \theta \theta^T x)}{\partial \theta} - \frac{\partial (2y^T \theta^T x)}{\partial \theta} + \frac{\partial (y^T y)}{\partial \theta}$$

$$= \frac{2}{m} x \left[x x^T \theta^T - x^T y \right]$$