均方误差(MSE)：$L(\theta)=\frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)})-y^{(i)})$

特征：$x^{(i)}$

预测：$\hat{y}^{(i)}\ (h_\theta(x^{(i)}))$

真实值：$y^{(i)}$

当单一特征　　　　最优解即使 $L(\theta)$ 最小

$h_\theta(x^{(i)})=\theta_0+\theta_1 x^{(i)}$　　　　最小二乘法(正规方程法)

$$\frac{\partial L(\theta)}{\partial\theta_0}=\cancel{\frac{1}{m}}\sum_{i=1}^{m}(\theta_0+\theta_1x^{(i)}-y^{(i)})$$

$$=m\theta_0+\sum_{i=1}^{m}\theta_1x^{(i)}-\sum_{i=1}^{m}y^{(i)}=0 \quad ①$$

$$\frac{\partial L(\theta)}{\partial\theta_1}=\cancel{\frac{1}{m}}\sum_{i=1}^{m}(\theta_0+\theta_1x^{(i)}-y^{(i)})x^{(i)}$$

$$=\sum_{i=1}^{m}\theta_0(x^{(i)})+\sum_{i=1}^{m}\theta_1(x^{(i)})^2-\sum_{i=1}^{m}y^{(i)}x^{(i)}=0 \quad ②$$

① $\times\sum_{i=1}^{m}x^{(i)}$，得 $m\theta_0\sum_{i=1}^{m}x^{(i)}+\theta_1(\sum_{i=1}^{m}x^{(i)})^2-\sum_{i=1}^{m}x^{(i)}\sum_{i=1}^{m}y^{(i)}=0$ ③

② $\times\ m$，得 $m\theta_0\sum_{i=1}^{m}x^{(i)}+m\theta_1\sum_{i=1}^{m}(x^{(i)})^2-m\sum_{i=1}^{m}y^{(i)}x^{(i)}=0$ ④

④ − ③，得 $\theta_1[m\sum_{i=1}^{m}(x^{(i)})^2-(\sum_{i=1}^{m}x^{(i)})^2]=m\sum_{i=1}^{m}x^{(i)}y^{(i)}-\sum_{i=1}^{m}x^{(i)}\sum_{i=1}^{m}y^{(i)}$

$$\theta_1=\frac{m\sum_{i=1}^{m}x^{(i)}y^{(i)}-\sum_{i=1}^{m}x^{(i)}\sum_{i=1}^{m}y^{(i)}}{m\sum_{i=1}^{m}(x^{(i)})^2-(\sum_{i=1}^{m}x^{(i)})^2}$$

$$=\frac{\sum_{i=1}^{m}x^{(i)}y^{(i)}-m\bar{x}\bar{y}}{\sum_{i=1}^{m}(x^{(i)})^2-m(\bar{x})^2}$$

代回①，有　$\theta_0=\bar{y}-\theta_1\bar{x}$