

模型：

$$h(x) = wx + b$$

代价函数：

$$J = \frac{1}{n} \sum_{i=1}^n (y_i - h(x_i))^2 = \frac{1}{n} \sum_{i=1}^n (y_i - wx_i - b)^2$$

J 分别对 w 和 b 求偏导：

$$\frac{\partial J}{\partial w} = -\frac{2}{n} \sum_{i=1}^n ((y_i - wx_i - b)x_i) = -\frac{2}{n} \sum_{i=1}^n (x_i y_i - wx_i^2 - bx_i) = 0$$

$$\frac{\partial J}{\partial b} = -\frac{2}{n} \sum_{i=1}^n (y_i - wx_i - b) = -\frac{2}{n} \left(\sum_{i=1}^n y_i - \sum_{i=1}^n b - \sum_{i=1}^n wx_i \right) = -2(\bar{y} - b - w\bar{x}) = 0$$

解得

$$b = \bar{y} - w\bar{x}$$

带回原式，有

$$\begin{aligned} \sum_{i=1}^n (x_i y_i - wx_i^2 - \bar{y}x_i + w\bar{x}x_i) &= 0 \\ \sum_{i=1}^n x_i y_i - w \sum_{i=1}^n x_i^2 - \sum_{i=1}^n \bar{y}x_i + w \sum_{i=1}^n \bar{x}x_i &= 0 \end{aligned}$$

因为

$$\sum_{i=1}^n x_i = n\bar{x}$$

所以

$$\sum_{i=1}^n x_i y_i - w \sum_{i=1}^n x_i^2 - n\bar{x}\bar{y} + nw\bar{x}^2 = 0$$

所以

$$w = \frac{\sum_1^n x_i y_i - n\bar{x}\bar{y}}{\sum_1^n x_i^2 - n\bar{x}^2}$$

因为

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n (x_i y_i - \bar{x}y_i - x_i \bar{y} + \bar{x}\bar{y}) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}$$

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2$$

所以

$$w = \frac{\sum_1^n ((x_i - \bar{x})(y_i - \bar{y}))}{\sum_1^n (x_i - \bar{x})^2}$$

$$b = \bar{y} - \left(\sum_{i=1}^n x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \right) \bar{x}$$

代价函数

$$J=\frac{1}{n}\sum_{i=1}^n\left(y_i-h(x_i)\right)^2=\frac{1}{n}\Big(\left(y_1-h(x_1)\right)^2+\left(y_2-h(x_2)\right)^2+\cdots+\left(y_n-h(x_n)\right)^2\Big)$$

$$h(X)=\left(h(x_1),h(x_2),\ldots,h(x_n)\right)$$

$$Y-h(X)=\left(y_1-h(x_1),y_2-h(x_2),\ldots,y_n-h(x_n)\right)$$

$$\left(Y-h(X)\right)^T\left(Y-h(X)\right)=\left(y_1-h(x_1)\right)^2+\left(y_2-h(x_2)\right)^2+\cdots+\left(y_n-h(x_n)\right)^2$$

所以

$$J=\frac{1}{n}\sum_{i=1}^n\left(y_i-h(x_i)\right)^2=\frac{1}{n}\left(Y-h(X)\right)^T\left(Y-h(X)\right)$$

设参数向量为 W,则

$$h(X)=WX$$

$$J=\frac{1}{n}(Y-WX)^T(Y-WX)$$

对 W 求导

$$\frac{\partial J}{\partial W}=\frac{2}{n}X^T(Y-WX)=0$$

$$X^TY=X^T XW$$

$$W=\left(X^TX\right)^{-1}X^TY$$

$$J=\frac{1}{n}\sum_{i=1}^n\left(y_i-h(x_i)\right)^2$$

$$h(x)=w_0+w_1x_1+w_2x_2$$

$$h(x)=w_0+w_1x_1+w_2x_2+w_3x_1x_2+w_4x_1^2+w_5x_2^2$$

$$z=[z_1,z_2,z_3,z_4,z_5]=[x_1,x_2,x_1x_2,x_1^2,x_2^2]$$

$$h(x)=w_0+w_1z_1+w_2z_2+w_3z_3+w_4z_4+w_5z_5$$

$$h(x)=w_0+w_1x_1+w_2x_1^2$$

$$h(x)=w_0+w_1z_1+w_2z_2$$

$$h(x) = 119.6 - 54x + 28.5x^2$$

$$h(x) = -22.8 + 116.99x$$