

* سوال اول

$$h_{\theta}(x) = b + w_1 x + w_2 x^2$$

x	3	2	1	0
y	1	-1	-1	2

$$X = \begin{bmatrix} 1 & 3 & 9 \\ 1 & 2 & 4 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 2 \end{bmatrix}, \quad \theta = \begin{bmatrix} b \\ w_1 \\ w_2 \end{bmatrix}$$

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^i) - y^i)^2 + \frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2$$

$$\text{عرف: } \min_{\theta} \left[\frac{1}{2m} \sum_{i=1}^m (y_i - x_i^T \theta)^2 + \frac{\lambda}{2m} \|\theta\|_2^2 \right]$$

برای تحقق شدن این هدف کافیه از عبارت فوق نسبت به θ گزینش گرفته و گزینش را برابر با صفر قرار دهیم و به روش Normal Equation پاسخ (θ_{ridge}^*) را پیدا کنیم.

$$\frac{\partial J(\theta)}{\partial \theta} = -\frac{1}{m} X^T (y - X\theta) + 2\lambda \theta \stackrel{\text{set}}{=} 0$$

$$\Rightarrow \theta_{ridge}^* = (X^T X + 2m\lambda \cdot I)^{-1} X^T y$$

در مضمون فوق $\lambda = 0.1$, $m = 4$, X , y ماتریس های فوق هستند. (یادآوری) $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$\theta_{ridge}^* = \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 \\ 9 & 4 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & 9 \\ 1 & 2 & 4 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} + 0.8 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 \\ 9 & 4 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 2 \end{bmatrix}$$

$$= \left(\begin{bmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{bmatrix} + \begin{bmatrix} 0.8 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0.8 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 \\ 9 & 4 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{5195}{12136} & -\frac{75}{328} & \frac{275}{12136} \\ -\frac{75}{328} & \frac{235}{328} & -\frac{75}{328} \\ \frac{275}{12136} & -\frac{75}{328} & \frac{1095}{12136} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 2 & 1 & 0 \\ 9 & 4 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{6295}{12136} \\ -\frac{375}{328} \\ \frac{4655}{12136} \end{bmatrix}$$

در نهایت نتایج به دست آمده به صورت زیر است :

$$\theta_{ridge}^* = \begin{bmatrix} b \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0.5187046803 \\ -1.1432926829 \\ 0.3835695452 \end{bmatrix}$$