규제가 있는 선형 모델

○ 릿지 회귀 비용함수

$$J(\theta) = MSE(\theta) + \frac{1}{2}\alpha\sum_{i=1}^{n}\boxed{{\theta_i}^2}$$

○ 라쏘 회귀 비용함수

$$J(\theta) = MSE(\theta) + \alpha\sum_{i=1}^{n}\boxed{|\theta_i|}$$

노름(norm)

: 두 벡터 사이의 거리를 측정하는 방법

$$\|\mathbf{x}\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

○ l_1 노름

$$d_1(\mathbf{p}, \mathbf{q}) = \|\mathbf{p} - \mathbf{q}\|_1 = \sum_{i=1}^{n} |p_i - q_i|, \text{ where } (\mathbf{p}, \mathbf{q}) \text{ are vectors } \mathbf{p} = (p_1, p_2, \dots, p_n) \text{ and } \mathbf{q} = (q_1, q_2, \dots, q_n)$$

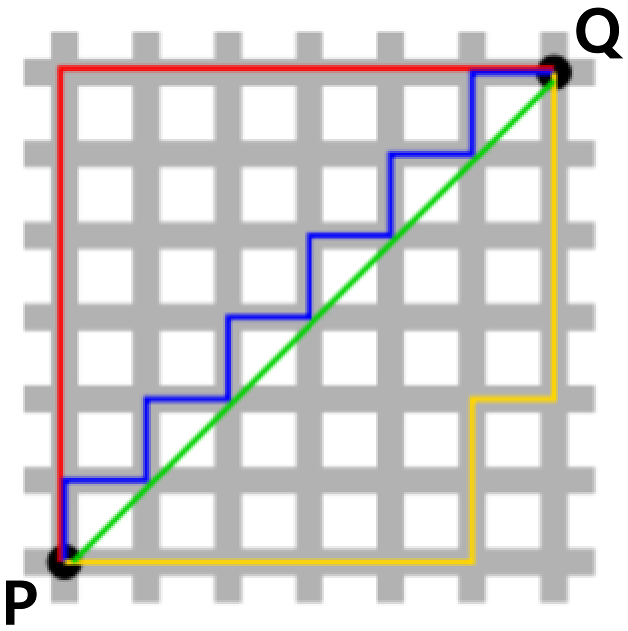
ex) p = (3, 5, 1) q = (1, 2, 4)

$$l_1\ norm = |3 - 1| + |5 - 2| + |1 - 4| = 8$$

○ l_2 노름

$$\|\boldsymbol{x}\|_2 := \sqrt{x_1^2 + \cdots + x_n^2}.$$

○ l_1 노름과 l_2 노름의 차이



l_1 : 빨간색, 파란색, 노란색과 같이 다양한 거리

l_2 : 초록색과 같이 가장 짧은 거리

규제가 있는 선형 모델

○ 릿지 회귀 비용함수

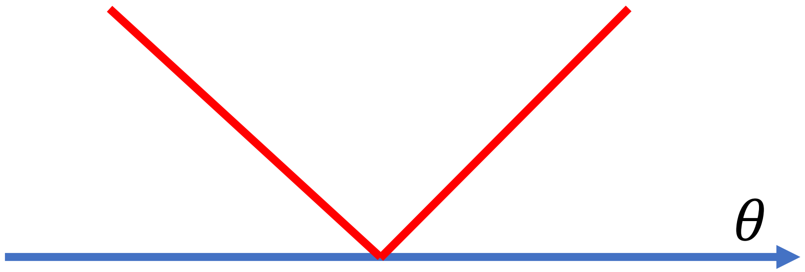
$$J(\theta) = MSE(\theta) + \frac{1}{2}\,\alpha \sum_{i=1}^{n} {\theta_i}^2$$

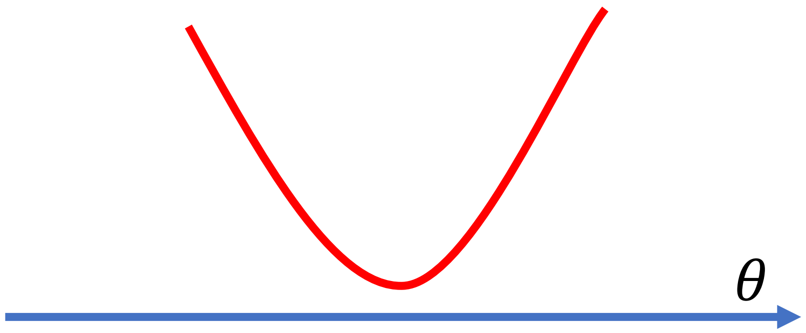
○ 라쏘 회귀 비용함수

$$J(\theta) = MSE(\theta) + \alpha \sum_{i=1}^{n} |\theta_i|$$

○ 라쏘 회귀 비용함수

$$J(\theta) = MSE(\theta) + \alpha \sum_{i=1}^{n} |\theta_i|$$





$$g(\theta, J) = \nabla_{\theta} MSE(\theta) + \alpha \begin{pmatrix} sign(\theta_1) \\ sign(\theta_2) \\ \cdots \\ sign(\theta_n) \end{pmatrix}, \quad sign(\theta_i) = \begin{cases} -1 & when \ \theta_i < 0 \\ 0 & when \ \theta_i = 0 \\ +1 & when \ \theta_i > 0 \end{cases}$$

○ 엘라스틱넷 회귀 비용함수

$$J(\theta) = MSE(\theta) + r\alpha \sum_{i=1}^{n} |\theta_i| + \frac{1-r}{2} \alpha \sum_{i=1}^{n} {\theta_i}^2$$