MULTIVARIATE OLS, SQUARED ERROR + RIDGE

Model

$\hat{y}_i = \theta x_i \qquad \theta = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix} \{k$

Cost

$$J(\theta) = \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2 + \lambda \sum_{j=1}^{k} \theta_j^2$$

$$= \frac{1}{N} \sum_{i=1}^{N} (\theta x_i - y_i)^2 + \lambda \sum_{j=1}^{k} \theta_j^2$$

∇

$$\frac{dJ}{d\theta} = \frac{1}{N} \sum_{i=1}^{N} 2(\theta x_i - y_i) x_i + \lambda \sum_{j=1}^{k} 2\theta_j$$

$$= \frac{2}{N} (\theta X - Y) X^T + 2\lambda\theta$$

Pseudocode

$X = [\cdots] \quad Y = [\cdots] \quad R = 0.001$

$\lambda = 0.1 \quad \theta = \text{random}()$

for i in iterations:

$$\text{cost} = \frac{1}{N} \text{sum}\left[(\theta x_i - y_i)^2\right] + \lambda \cdot \text{sum}(\theta^2)$$

$$\theta -= R \cdot \frac{2}{N} (\theta X - Y) X^T + 2\lambda\theta$$