

Normal Equation

Let $x^{(i)} \in \mathbf{R}^3$ be i -th data and $y^{(i)} \in \mathbf{R}$ be i -th label for $i = 1, 2, \dots, m$. We want to find unknown weight, $w \in \mathbf{R}^3$. y, X, w are defined as

$$y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{bmatrix}, \quad (1)$$

$$X = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} \\ x_1^{(2)} & x_2^{(2)} & x_3^{(2)} \\ \vdots & \vdots & \vdots \\ x_1^{(m)} & x_2^{(m)} & x_3^{(m)} \end{bmatrix}, \quad (2)$$

$$w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}. \quad (3)$$

Derive a normal equation, $X^T X w = X^T y$, to solve the following least squares problem.

$$\min_w \frac{1}{m} \sum_{i=1}^m |y^{(i)} - f(w; x^{(i)})|^2 \quad (4)$$

where

$$f(w; x) = w_1 x_1 + w_2 x_2 + w_3 x_3 \quad (5)$$

Proof. For $k = 1, 2, 3$,

$$\frac{\partial J}{\partial w_k} = \frac{\partial}{\partial w_k} \left[\frac{1}{m} \sum_{i=1}^m |y^{(i)} - f(w; x^{(i)})|^2 \right] \quad (6)$$

$$= -\frac{2}{m} \sum_{i=1}^m \left(y^{(i)} - f(w; x^{(i)}) \right) \frac{\partial}{\partial w_k} f(w; x^{(i)}) \quad (7)$$

$$= -\frac{2}{m} \sum_{i=1}^m \left(y^{(i)} - f(w; x^{(i)}) \right) \frac{\partial}{\partial w_k} (w_1 x_1^{(i)} + w_2 x_2^{(i)} + w_3 x_3^{(i)}) \quad (8)$$

$$= -\frac{2}{m} \sum_{i=1}^m \left(y^{(i)} - f(w; x^{(i)}) \right) x_k^{(i)} \quad (9)$$

$$= -\frac{2}{m} \sum_{i=1}^m x_k^{(i)} \left(y^{(i)} - (w_1 x_1^{(i)} + w_2 x_2^{(i)} + w_3 x_3^{(i)}) \right) \quad (10)$$

$$\nabla J(w) = \begin{bmatrix} \frac{\partial J}{\partial w_1} & \frac{\partial J}{\partial w_2} & \frac{\partial J}{\partial w_3} \end{bmatrix}^T \quad (11)$$

$$= -\frac{2}{m} X^T (y - Xw) = 0 \quad (12)$$

Therefore,

$$-\frac{2}{m} X^T (y - Xw) = 0 \quad (13)$$

$$-X^T (y - Xw) = 0 \quad (14)$$

$$X^T X w = X^T y \quad (15)$$

□

Proof. For $k = 1, 2, 3$,

$$\nabla J(w) = \nabla \left[\frac{1}{m} \|y - f(w; x)\|^2 \right] \quad (16)$$

$$= \nabla \left[\frac{1}{m} \|y - Xw\|^2 \right] \quad (17)$$

$$= \nabla \left[\frac{1}{m} (y - Xw)^T (y - Xw) \right] \quad (18)$$

$$= \frac{1}{m} \nabla (y^T y - y^T Xw - w^T X^T y + w^T X^T Xw) \quad (19)$$

$$= \frac{1}{m} \nabla (y^T y - 2w^T X^T y + w^T X^T Xw) \quad (20)$$

$$= \frac{1}{m} (\nabla(y^T y) - 2\nabla(w^T X^T y) + \nabla(w^T X^T Xw)) \quad (21)$$

$$= \frac{1}{m} (-2X^T y + 2X^T Xw) = 0 \quad (22)$$

Therefore,

$$-X^T y + X^T Xw = 0 \quad (23)$$

$$X^T Xw = X^T y \quad (24)$$

□