HW#3

1. $L(W) = \frac{1}{2}\left[(QW - Y)^T (QW - Y) + \lambda W^T I_{0+M} W\right]$

W – weight matrix $((M+1) \times 1)$

Y – label vector $(M \times 1)$

I_{0+M} – Identity Matrix $((M+1) \times (M+1))$ $I_{0+M} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & & 1 & \\ & & & \ddots \\ 0 & & & 1 \end{bmatrix}$

$\frac{\partial L(W)}{\partial W} = 2Q^T(QW - Y) + 2\lambda I_{0+M} W = 0$

$\therefore W^* = (Q^T Q + \lambda I_{0+M})^{-1} Q^T Y$

2. $\sigma \to \infty$ $k(y, z) = 1 \Rightarrow \phi_i(x) = \frac{1}{M}$ & $\phi_0(x) = 1$

$Q = \begin{bmatrix} 1 & \frac{1}{M} & \frac{1}{M} & \cdots \\ 1 & \frac{1}{M} & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & \frac{1}{M} & \frac{1}{M} & \frac{1}{M} \end{bmatrix}$ $Q^T = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ \frac{1}{M} & \frac{1}{M} & \frac{1}{M} & & \frac{1}{M} \\ \frac{1}{M} & \vdots & \vdots & & \vdots \\ \vdots & & & & \\ \frac{1}{M} & & & & \frac{1}{M} \end{bmatrix}$

$(M+1) \times M$

$Q^T Y = [\Sigma y_i, \bar{y}, \bar{y}, \cdots \bar{y}]$ $((M+1) \times 1)$

$W^* = [\frac{\Sigma y_i}{M}, 0, 0, 0 \quad]$
$= [\bar{y}, 0, 0, \cdots 0]$

For $\phi_i(x) = \frac{1}{M}$, we know that all the features are in the same space and expect linear regression. Under this condition, it might be good to only update bias to adjust model and eliminate penalty term by expecting $W_1 - W_M = 0$

$$\bar{y} = \frac{\sum_{i=1}^{M} y_i}{M}$$

3. $L(w) = \frac{1}{2}[(QW - Y)^T (QW - Y)]$

It will be possibe to achieve $QW - Y = 0$

$\sigma \to 0 \quad k(y,z) = \begin{cases} 0 & y \neq z \\ 1 & y = z \end{cases} \quad \therefore \phi_i(x_j) = \begin{cases} 1 & \text{when } i = j \\ 0 & \text{when } i \neq j \end{cases}$

$\& \ \phi_0(x_j) = 1$

$$\therefore Q = \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & & & \vdots \\ \vdots & 0 & 0 & 1 & & \vdots \\ \vdots & & & & \ddots & \\ 1 & & & & & 1 \end{bmatrix}}_{M \times (M+1)} \qquad Q^T = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 0 & & \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \\ \vdots & \vdots & \vdots & & \\ 0 & & & & 1 \end{bmatrix}$$

If there is no penalty, we expect $W = Q^{-1}Y$ & ignore bias term.

$\therefore W^* = [y_1, y_2, y_3, \cdots y_M]$

4. $L(w) = \frac{1}{2}[(QW - Y)^T (QW - Y) + \lambda W I_{0+m} W]$

$w^* = (Q^T Q + \lambda I_{0+m})^{-1} Q^T Y \qquad Q^T Q = \begin{bmatrix} M & 1 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ 1 & & & 1 & \\ & & & & 1 \end{bmatrix}$

$Q^T Y = [\Sigma y_i, y_1, y_2, y_3 \cdots y_M]$

$\therefore w^* = [\ \bar{y}\ , \frac{y_1}{1+\lambda}, \frac{y_2}{1+\lambda} \cdots \frac{y_M}{1+\lambda}\ ]$