Briana Zhao
U93987617
EC414
HW2

1

$$Loss = (y - Xw)^T(y - Xw) + \lambda \|w\|_2^2$$

$$\min_{\vec{w}} \left[(y - Xw)^T(y - Xw) + \lambda \|w\|_2^2 \right]$$

ⓐ $$\frac{\partial}{\partial w}\left[(y - Xw)^T(y - Xw) \right] = -2X^T(y - Xw)$$

ⓑ $$\frac{\partial}{\partial w}\left[\lambda w^T w \right] = 2\lambda w$$

Combining ⓐ + ⓑ and setting to 0:

$$-2X^T(y - Xw) + 2\lambda w = 0$$
$$-\cancel{2}X^T y + \cancel{2}X^T X w + \cancel{2}\lambda w = 0$$
$$X^T X w + \lambda w = X^T y$$

$$\boxed{\vec{w} = (X^T X + \lambda I)^{-1} X^T y}$$

2

$$p(y|x,w) = \frac{1}{\sigma\sqrt{2\pi}} e^{\sum_{i=1}^{n} \frac{-(y_i - x_i^T w)^2}{2\sigma^2}}$$

$$p(w|\mu,b) = \frac{1}{2b} e^{-\frac{\|w-\mu\|_1}{b}} \longrightarrow \frac{1}{2b} e^{-\frac{\|w\|_1}{b}}$$ since mean is 0

$$\underset{\vec{w}}{\operatorname{argmax}} \left[\frac{1}{\sigma\sqrt{2\pi}} e^{\sum_{i=1}^{n} \frac{-(y_i - x_i^T w)^2}{2\sigma^2}} \cdot \left(\frac{1}{2b} e^{-\frac{\|w-\mu\|_1}{b}} \right) \right]$$

$$= \underset{\vec{w}}{\operatorname{argmax}} \left[\ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n - \sum_{i=1}^{n} \overbrace{\frac{1}{2\sigma^2}}^{\text{constant}} (y_i - x_i^T w)^2 + \ln\left(\frac{1}{2b}\right) - \frac{\|w-\mu\|_1}{b} \right]$$

$$= \underset{\vec{w}}{\operatorname{argmax}} \left[\underbrace{\ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n + \ln\left(\frac{1}{2b}\right)}_{\text{not depedent on } \vec{w}} - \sum_{i=1}^{n} (y_i - x_i^T w)^2 - \frac{\|w-\mu\|_1}{b} \right]$$

$$= \underset{\vec{w}}{\operatorname{argmax}} \left[-\sum_{i=1}^{n} (y_i - x_i^T w)^2 - \frac{\|w\|_1}{b} \right]$$

$$= \underset{\vec{w}}{\operatorname{argmin}} \left[\sum_{i=1}^{n} (y_i - x_i^T w)^2 - \frac{1}{b} \sum_{j=1}^{d} |w_j| \right] \qquad \lambda = \frac{1}{b}$$

In Lasso, we want to minimize $\sum_{i=1}^{n}(y_i - x_i^T\omega)^2 + \lambda\sum_{j=1}^{d}|\omega_j|$

therefore, $\underset{\vec{\omega}}{\operatorname{argmin}}\left[\sum_{i=1}^{n}(y_i - x_i^T\omega)^2 + \overbrace{\frac{1}{b}}^{\lambda}\sum_{j=1}^{d}|\omega_j|\right]$ derived from MAP with

a Laplace prior is the same formula as Lasso Linear Regression