COMP9417 – Machine Learning

Tutorial: Regression I

Wanqing Yang - z5325987

Question 1.

(b) Because $L(w_0, w_1) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - (w_0 + w_1 x_i)\right)^2$

Then $\frac{\partial L(w_0,w_1)}{\partial w_0} = \frac{1}{n}\sum_{i=1}^{n}(2w_0 + 2w_1 x_i - 2y_i)$ $\frac{\partial L(w_0,w_1)}{\partial w_1} = \frac{1}{n}\sum_{i=1}^{n}\left(2w_1 x_i^2 + 2w_0 x_i - 2x_i y_i\right)$

Let $\frac{\partial L(w_0,w_1)}{\partial w_0} = 0$ and $\frac{\partial L(w_0,w_1)}{\partial w_1} = 0$

Then $w_0 = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$ …… ① $w_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$ …… ②

Because $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ …… ③ and $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ …… ④

Substituting equation ①②③④ into $y = w_0 + w_1 x$

$\frac{1}{n}\sum_{i=1}^{n} y_i = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} + \frac{1}{n}\sum_{i=1}^{n} x_i \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$

$\frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\left(\frac{n\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - n\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i + n\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)^2 \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i - \left(\sum_{i=1}^{n} x_i\right)^2}\right)$

$\frac{1}{n}\sum_{i=1}^{n} y_i = \frac{\sum_{i=1}^{n} y_i (n\sum_{i=1}^{n} x_i - (\sum_{i=1}^{n} x_i)^2)}{n\sum_{i=1}^{n} x_i - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{1}{n}\sum_{i=1}^{n} y_i$

Therefore, the point $(\bar{x}, \bar{y})$ is always on the least squares regression line.

(c) Because $L(w_0, w_1) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - (w_0 + w_1 x_i)\right)^2 + \lambda w_1^2$

Then $\frac{\partial L(w_0,w_1)}{\partial w_0} = \frac{1}{n}\sum_{i=1}^{n}(2w_0 + 2w_1 x_i - 2y_i)$ and

$\frac{\partial L(w_0,w_1)}{\partial w_1} = \frac{1}{n}\sum_{i=1}^{n}\left(2w_1 x_i^2 + 2w_0 x_i - 2x_i y_i\right) + 2\lambda w_1$

Let $\frac{\partial L(w_0,w_1)}{\partial w_0} = 0$ and $\frac{\partial L(w_0,w_1)}{\partial w_1} = 0$

Then $w_0 = \frac{1}{n}\sum_{i=1}^{n} y_i - \frac{\left(1-\frac{1}{n}\right)(\sum_{i=1}^{n} x_i)^2 \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x^2 - \left(\sum_{i=1}^{n} x_i\right)^2 + n^2\lambda}$ and

$w1 = \frac{\left(1-\frac{1}{n}\right)\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 + n\lambda}$

Question 2.

(g) (i) This is true.

$MSE(w) = \frac{1}{n}\left\|y - Xw\right\|_2^2 = \frac{1}{n}(y\text{-}Xw)^T(y - Xw)$

$SSE(w) = \left\|y - Xw\right\|_2^2 = (y\text{-}Xw)^T(y - Xw)$

Because $\frac{\partial MSE(w)}{\partial \mathrm{w}} = \frac{\partial \frac{1}{\mathrm{n}}(y\text{-}Xw)^T(y-Xw)}{\partial \mathrm{w}} = \frac{\partial \frac{1}{\mathrm{n}}(y^Ty-y^Txw-w^Tx^Ty+w^Tx^Txw)}{\partial \mathrm{w}}$

$= \frac{1}{\mathrm{n}}\left(\frac{\partial y^Ty}{\partial \mathrm{w}} - 2\frac{\partial(x^Ty)w}{\partial \mathrm{w}} + \frac{\partial w^Tx^Txw}{\partial \mathrm{w}}\right)$

$= \frac{1}{\mathrm{n}}(-2x^Ty + 2x^Txw)$ …………… (a) Hint1 & Hint2

$\frac{\partial \mathrm{SSE}(w)}{\partial \mathrm{w}} = \frac{\partial(y\text{-}Xw)^T(y-Xw)}{\partial \mathrm{w}} = \frac{\partial(y^Ty-y^Txw-w^Tx^Ty+w^Tx^Txw)}{\partial \mathrm{w}}$

$= \frac{\partial y^Ty}{\partial \mathrm{w}} - 2\frac{\partial(x^Ty)w}{\partial \mathrm{w}} + \frac{\partial w^Tx^Txw}{\partial \mathrm{w}}$

$= -2x^Ty + 2x^Txw$ …………… (a) Hint1 & Hint2

Then $\frac{\partial \mathrm{SSE}(w)}{\partial \mathrm{w}} = n\frac{\partial MSE(w)}{\partial \mathrm{w}}$

Therefore, both MSE and SSE are minimized at point $w = (X^TX)^{-1}X^Ty$

(ii) False.

Because both MSE and SSE are minimized at point $w = (X^TX)^{-1}X^Ty$.

But $MSE(w) = \frac{1}{\mathrm{n}}SSE(w)$

Therefore, $min_{w\in R^p}MSE(w) \neq min_{w\in R^p}SSE(w)$