ML HW 1

No:

Date: /

1.

1-(a)

$$L(w,b) = \frac{1}{2\times 5}\sum_{i=1}^{5}\left(y_i - (w^T x_i + b)\right)^2$$

$$w' = \arg\min_{w} L(w,b)$$

when $\dfrac{\partial L(w,b)}{\partial w} = 0$

$$\bar{x} = \frac{\sum_{i=1}^{5} x_i}{5} = 3$$

$$w' = \frac{\sum_{i=1}^{5}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{5}(x_i - \bar{x})^2}$$

$$\bar{y} = \frac{\sum_{i=1}^{5} y_i}{5} = 3.36$$

$$= \frac{(1.2-3.36)(1-3)+(2.4-3.36)(2-3)+(3-3.36)(3-3)+(4.1-3.36)(4-3)+(5.6-3.36)(5-3)}{2^2+1^2+0^2+1^2+2^2}$$

$$= \frac{(-2.16)\cdot(-2)+(-0.96)(-1)+0+(0.74)(1)+(2.24)(2)}{10}$$

$$= \frac{1}{10}\left[4.32+0.96+0.74+4.48\right] = \frac{10.5}{10} = 1.05$$

$$b' = \bar{y} - w\bar{x} = 3.36 - 1.05\times 3 = 3.36 - 3.15 = 0.21$$

Scanned with CamScanner

No:
Date: /

1-(b) To find w, b that minimize L(w,b)

$$\Rightarrow \frac{\partial L(w,b)}{\partial w} = 0 \Rightarrow \frac{\partial \frac{1}{2N}\sum_{i=1}^{N}(y_i - b - wx_i)^2}{\partial w} = 0$$

$$\Rightarrow 2\cdot\frac{1}{2N}\sum_{i=1}^{N}(y_i - b - wx_i)x_i = 0 \Rightarrow \sum_{i=1}^{N}(y_i - b - wx_i)x_i = 0$$

$$\Rightarrow \sum_{i=1}^{N} y_i x_i - \sum_{i=1}^{N} wx_i^2 - \sum_{i=1}^{N} bx_i = 0$$

$$\Rightarrow \sum_{i=1}^{N} y_i x_i - \sum_{i=1}^{n}(\bar{y} - b\bar{x})x_i - w\sum_{i=1}^{N} x_i^2 = 0$$

$$\Rightarrow \sum_{i=1}^{N} y_i x_i - \sum_{i=1}^{N} \bar{y} x_i + \sum_{i=1}^{N} w\bar{x}x_i - w\sum_{i=1}^{N} x_i^2 = 0$$

$$\Rightarrow w\sum_{i=1}^{N} x_i(\bar{x} - x_i) + \sum_{i=1}^{N} x_i(y_i - \bar{y}) = 0$$

$$\Rightarrow w = \frac{\sum_{i=1}^{N} x_i(y_i - \bar{y})}{\sum_{i=1}^{N} x_i(x_i - \bar{x})} = \frac{\sum_{i=1}^{N} x_i y_i - \bar{y}\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} x_i^2 - \bar{x}\sum x_i}$$

$$= \frac{\sum_{i=1}^{N} xy - N\bar{x}\bar{y} - N\bar{x}\bar{y} + N\bar{x}\bar{y}}{\sum x_i^2 - N\bar{x}^2}$$

$$= \frac{\sum xy - \sum x_i\bar{y} - \sum \bar{x}y_i + \sum \bar{x}\bar{y}}{\sum x_i^2 - 2N\bar{x}^2 + N\bar{x}^2}$$

$$\therefore w = \frac{\sum(x_i y_i - x_i\bar{y} - \bar{x}y_i + \bar{x}\bar{y})}{\sum(x_i - \bar{x})^2} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

$$b = \bar{y} - \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}\bar{x}$$

Scanned with CamScanner

ML HW1-2

No:

Date: /

1-(C)

$$L_2(W,b) = \frac{1}{2N}\sum_{i=1}^{N}\left(y_i-(W^T x_i+b)\right)^2+\frac{\lambda}{2}\|W\|^2$$

$$\frac{\partial\left(L_1(W,b)+\frac{\lambda}{2}|W|^2\right)}{\partial W}=-\frac{1}{N}\sum_{i=1}^{N}(y_i-b-Wx_i)x_i+\frac{\partial\frac{\lambda}{2}\|W\|^2}{\partial W}=0$$

$$\Rightarrow -\frac{1}{N}\sum\left(y_i x_i-bx_i-Wx_i^2\right)+\lambda W=0$$

$$\Rightarrow -\left(\sum y_i x_i-\sum W x_i^2-\sum b x_i\right)+\lambda N W=0$$

$$\Rightarrow -\sum y_i x_i+\sum(\bar{y}-W^T\bar{x})x_i+\sum W x_i^2+\lambda N W=0$$

$$\Rightarrow -\sum y_i x_i+\sum\bar{y}x_i-W^T\bar{x}\sum x_i+W^T\sum x_i^2+\lambda N W=0$$

$$\Rightarrow -\sum y_i x_i+\sum\bar{y}x_i+W^T\left(\sum x_i(x_i-\bar{x})+\lambda N\right)=0$$

$$\Rightarrow W=\frac{\sum y_i x_i-\sum\bar{y}x_i}{\sum x_i(x_i-\bar{x})+\lambda N}=\frac{\sum(y_i-\bar{y})(x_i-\bar{x})}{\sum(x_i-\bar{x})^2+\lambda N}$$

$$b=\bar{y}-W^T\bar{x}$$

No:

Date: /

2、由1.(2)把 x_i 代換成 x_i+n_i

→ optimal weight: $W_0 = E\left(\frac{\sum(x_i+n_i-\bar{x})(y_i-\bar{y})}{(\sum(x_i+n_i-\bar{x})^2}\right)$

由1.(3)知,加上 regularization 的 optimal weight

$$W_1 = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2+\sigma^2 N}$$

若 $W_0 = W_1$,則可證得 minimizing 兩者 Loss 的 w, b 相同

分子: $E\left(\sum_{i=1}^{N}(x_i+n_i-\bar{x})(y_i-\bar{y})\right) = \sum_{i=1}^{N}\left[E(x_i+n_i-\bar{x})(y_i-\bar{y})\right]$

$= \sum_{i=1}^{N}\left(E\left(x_iy_i - x_i\bar{y} + \cancel{n_iy_i} - \cancel{n_i\bar{y}} - \bar{x}y_i + \bar{x}\bar{y}\right)\right)$

$\because E(n_i)=0$

$= \sum_{i=1}^{N}\left(E(x_iy_i - x_i\bar{y} - \bar{x}y_i + \bar{x}\bar{y})\right) = \sum_{i=1}^{N} E\left((x_i-\bar{x})(y_i-\bar{y})\right)$

$= \sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})$ 同 W_1 的分子

分母: $E\left(\sum_{i=1}^{N}(x_i+n_i-\bar{x})^2\right) = \sum_{i=1}^{N}\left[E(x_i-\bar{x}+n_i)^2\right]$

$= \sum_{i=1}^{N}\left[E(x_i-\bar{x})^2 + 2E(x_i-\bar{x})(n_i) + E(n_i^2)\right]$

$= \sum_{i=1}^{N}\left(E(x_i-\bar{x})^2 + \sigma^2\right) = \sum_{i=1}^{N}(x_i-\bar{x})^2 + N\sigma^2$

同 W_1 的分母

故得證 ※

Scanned with CamScanner

No:
Date: /

3.

(a) Let $A = \sum_{i=1}^{N} g_k(x_i) y_i$

$$e_k = \frac{1}{N}\left[\sum_{i=1}^{N}\left(g_k(x_i) - y_i\right)^2\right] = \frac{1}{N}\left(\sum_{i=1}^{N}\left(g_k(x_i)\right)^2 - 2\cdot\sum_{i=1}^{N} g_k(x_i)\cdot y_i + \sum_{i=1}^{N} y_i^2\right)$$

$$= \frac{1}{N}\left(NS_k - 2\cdot A + Ne_0\right) \Rightarrow Ne_k = NS_k - 2A + Ne_0$$

$$\Rightarrow A = \frac{N(S_k - e_k + e_0)}{2}$$ //

Na_k

$\sum_{i=1}^{N} a_k g(x)$

(b) Let $F = \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{k=1}^{K} a_k g_k(x_i) - y_i\right)^2$

$N \sum g(x) a_k$

$$\frac{\partial F}{\partial a} = \frac{1}{N}\sum_{i=1}^{N}\left[2\cdot\left(\sum_{k=1}^{K} a_k g_k(x_i) - y_i\right)\cdot\sum_{k=1}^{K} g_k(x_i)\right] = 0$$

$$\to \sum_{i=1}^{N}\left[\left(\sum_{k=1}^{K} a_k g_k(x_i)\right)\cdot\sum_{i=1}^{N} g_k(x_i) - \sum_{k=1}^{K} g_k(x_i)\cdot y_i\right] = 0$$

$$\to \sum_{k=1}^{K}\left[\left(\sum_{i=1}^{N} a_k g_k(x_i)\right)\cdot\sum_{i=1}^{N} g_k(x_i) - \sum_{i=1}^{N} g_k(x_i)\cdot y_i\right] = 0$$

$$\to \sum_{k=1}^{K}\left(a\cdot NS_k - \frac{N}{2}(S_k - e_k + e_0)\right) = 0$$

$$\to a = \frac{\frac{N}{2}(S_k - e_k + e_0)}{NS_k} = \frac{S_k - e_k + e_0}{2S_k}$$

Scanned with CamScanner