

ML Foundation: HW3

b04902053 鄭淵仁

December 30, 2017

1

QUIZ

作業三

20 questions

Your Score

200/200 points (100%)

We keep your highest score.

[View Latest Submission](#)

[Take it again](#)

2

Claim: $H^2 = H$

Proof:

$$\begin{aligned} H^2 &= (X(X^T X)^{-1} X^T)^2 \\ &= (X(X^T X)^{-1} X^T)(X(X^T X)^{-1} X^T) \\ &= X(X^T X)^{-1} [(X^T X)(X^T X)^{-1}] X^T \\ &= X(X^T X)^{-1} X^T \\ &= H \end{aligned}$$

With the claim above, we can prove that:

$$\begin{aligned} (I - H)^2 &= I^2 - 2IH + H^2 \\ &= I - 2H + H \\ &= I - H \end{aligned}$$

3

TODO: 看不懂題目在寫什麼？QQ

4

$$\hat{E}_2(\Delta u, \Delta v) = E(u, v) + \nabla E(u, v) \cdot (\Delta u, \Delta v) + \frac{1}{2}(\Delta u, \Delta v)^T \nabla^2 E(u, v) (\Delta u, \Delta v)$$

Set the partial differences of $\hat{E}_2(\Delta u, \Delta v)$ be 0, we have :

$$\begin{cases} 0 = \frac{\partial \hat{E}_2(\Delta u, \Delta v)}{\partial \Delta u} = \frac{\partial E}{\partial u} + \frac{1}{2} \left(2 \frac{\partial^2 E}{\partial u^2} \Delta u + 2 \frac{\partial^2 E}{\partial u \partial v} \Delta v \right) \\ \quad \quad \quad = \frac{\partial E}{\partial u} + \frac{\partial^2 E}{\partial u^2} \Delta u + \frac{\partial^2 E}{\partial u \partial v} \Delta v \\ 0 = \frac{\partial \hat{E}_2(\Delta u, \Delta v)}{\partial \Delta v} = \frac{\partial E}{\partial v} + \frac{\partial^2 E}{\partial v^2} \Delta v + \frac{\partial^2 E}{\partial v \partial u} \Delta u \end{cases}$$

Simplify the equations :

$$\begin{cases} 0 = \frac{\partial E}{\partial u} + \frac{\partial^2 E}{\partial u^2} \Delta u + \frac{\partial^2 E}{\partial u \partial v} \Delta v \\ 0 = \frac{\partial E}{\partial v} + \frac{\partial^2 E}{\partial v^2} \Delta v + \frac{\partial^2 E}{\partial v \partial u} \Delta u \end{cases}$$

Now combine the two equations to one equation by vector (u, v) :

$$\begin{aligned} 0 &= \nabla E(u, v) + \nabla^2 E(u, v) \cdot (\Delta u, \Delta v) \\ -\nabla^2 E(u, v) \cdot (\Delta u, \Delta v) &= \nabla E(u, v) \\ (\Delta u, \Delta v) &= -(\nabla^2 E(u, v))^{-1} \nabla E(u, v) \end{aligned}$$

Q.E.D.

5

$$\max_h \prod_{n=1}^N h_y(x_n) = \max_w \prod_{n=1}^N \frac{\exp(w_{y_n}^T x_n)}{\sum_{k=1}^K \exp(w_k^T x_n)}$$

Take natural log on the it :

$$\begin{aligned} & \max_w \ln \prod_{n=1}^N \frac{\exp(w_{y_n}^T x_n)}{\sum_{k=1}^K \exp(w_k^T x_n)} \\ &= \max_w \sum_{n=1}^N \ln \frac{\exp(w_{y_n}^T x_n)}{\sum_{k=1}^K \exp(w_k^T x_n)} \\ &= \max_w \sum_{n=1}^N \left(\ln(\exp(w_{y_n}^T x_n)) - \ln \sum_{k=1}^K \exp(w_k^T x_n) \right) \\ &= \min_w \sum_{n=1}^N \left(\ln \sum_{k=1}^K \exp(w_k^T x_n) - w_{y_n}^T x_n \right) \end{aligned}$$

Therefore the E_{in} is :

$$E_{in} = \frac{1}{N} \sum_{n=1}^N \left(\ln \sum_{k=1}^K \exp(w_k^T x_n) - w_{y_n}^T x_n \right)$$

6

First we compute :

$$\begin{aligned} & \frac{\partial \left(\sum_{n=1}^N \left(\ln \sum_{k=1}^K \exp(w_k^T x_n) \right) \right)}{\partial w_i} \\ &= \sum_{n=1}^N \left(\frac{\exp(w_i^T x_n)}{\sum_{k=1}^K \exp(w_k^T x_n)} x_n \right) \\ &= \sum_{n=1}^N (h_i(x_n) x_n) \end{aligned}$$

Therefore the answer is :

$$\begin{aligned} \frac{\partial E_{in}}{\partial w_i} &= \frac{\partial \left(\frac{1}{N} \sum_{n=1}^N \left(\ln \sum_{k=1}^K \exp(w_k^T x_n) - w_{y_n}^T x_n \right) \right)}{\partial w_i} \\ &= \frac{1}{N} \sum_{n=1}^N ((h_i(x_n) x_n) - [[y_n = i]] x_n) \\ &= \frac{1}{N} \sum_{n=1}^N (((h_i(x_n)) - [[y_n = i]]) x_n) \end{aligned}$$

7

Figure 1 shows the histogram of E_{out} .

8

9

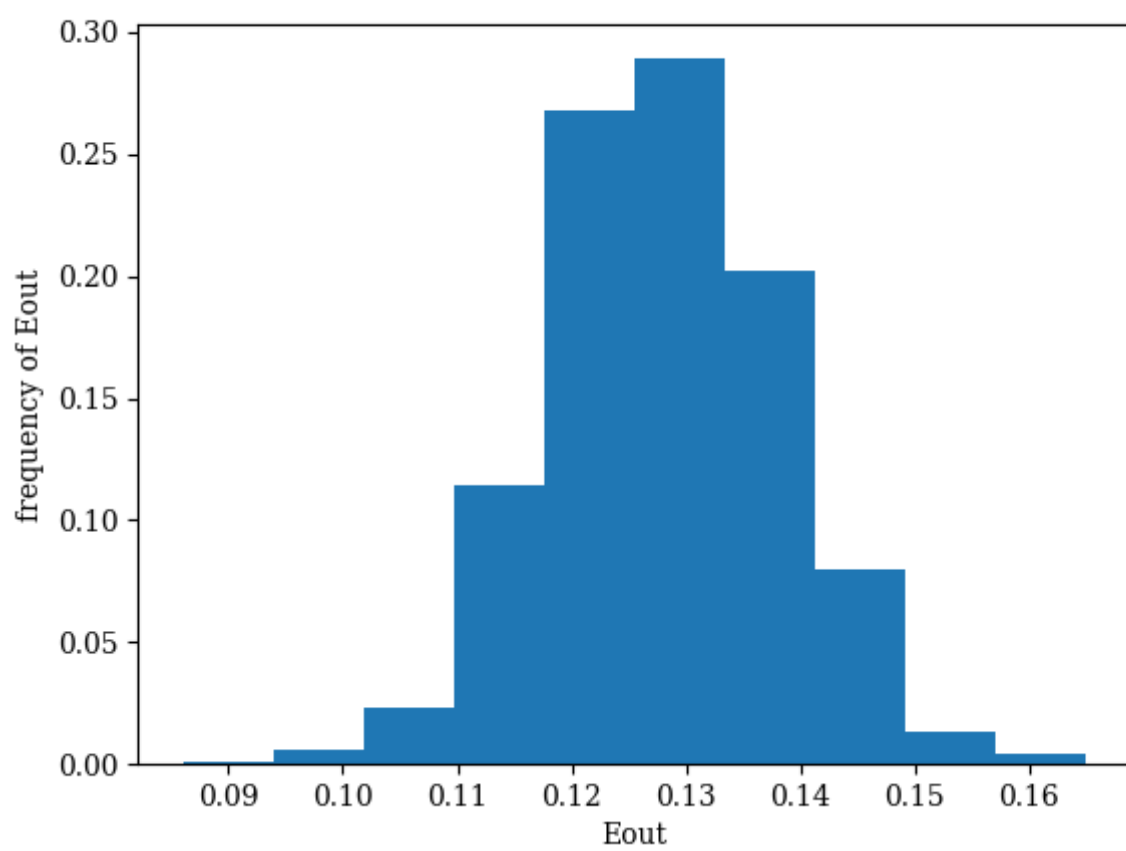


Figure 1: Histogram of E_{out}