

# Machine Learning Foundations/hw4

連少甫/r07922107

January 8, 2019

1.

The screenshot shows the Coursera interface for the course 'Machine Learning Foundations (Machine Learning Foundations)---Algorithmic Foundations'. The user is logged in as SHAOFU, LIEN. The course progress bar shows 'Three Learning Principles' with a list of quizzes: 'Occam's Razor' (10 min), 'Sampling Bias' (11 min), 'Data Snooping' (12 min), 'Power of Three' (8 min), and '作業四' (Assignment 4) (20 questions), which is currently selected and marked as completed. The main content area displays the quiz result for '作業四' with a score of 100.00% and a message: '我們會保留您的最高分數。查看最新提交內容' (We will keep your highest score. View latest submission content). A '再次參加' (Retake) button is visible.

2.

已知加入正規化後的公式為： $E_{aug}(w) = E_{in}(W) + \frac{\lambda}{N} w^T w$

接著對 $E_{aug}(w)$ 微分得到下式：

$$\nabla E_{aug}(w) = \nabla E_{in}(W) + \frac{2\lambda w}{N}$$

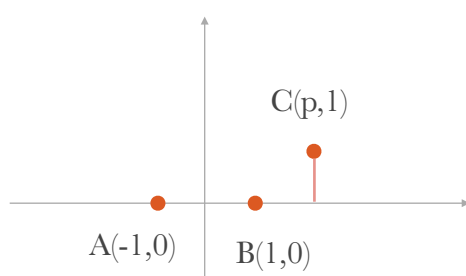
已之更新公式為： $w_{t+1} = w_t - \eta \nabla E_{aug}(w)$

將 $\nabla E_{aug}(w)$ 替換掉後得到更新公式如下：

$$w_{t+1} = (1 - \frac{2\eta\lambda}{N})w_t - \eta \nabla E_{in}(w(t))$$

3.

固定AB時： $e_C = 1$ ，如粉紅色的線條所示。

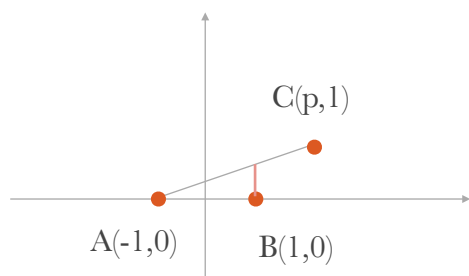


固定AC時：

$$\text{AC直線方程式為 } y = \frac{1}{p+1}(x+1)$$

$$x=1 \text{ 時 } y = \frac{2}{p+1}$$

$$e_B = \left(\frac{2}{p+1}\right)^2, \text{ 如粉紅色的線條所示。}$$



固定BC時：

$$\text{BC直線方程式為 } y = \frac{1}{1-p}(1-x)$$

$$e_A = \left(\frac{2}{1-p}\right)^2$$

$$\text{總error} = \frac{1}{3}(e_A + e_B + e_C) = \frac{1}{3}\left(1 + \left(\frac{2}{p+1}\right)^2 + \left(\frac{2}{1-p}\right)^2\right)$$

4.

step1: 從  $\{1, 2, 3, \dots, N+K\}$  中選取一個整數，記為  $s$

step2: 計算  $\nabla E_{in} = 2 * (wx_s - y)x_s$

step3: 計算  $\nabla E_{aug} = 2\lambda w$

$$w_{t+1} = w_t - \eta(\nabla E_{in} + \nabla E_{aug})$$

和第三題的更新公式主要差在  $E_{in}$  的算法不同。

Bonus:

令  $x_i$  為  $\sin(ax)$  上的採樣點  $i = 1, 2, 3, \dots, N$

$$\text{error} = (\sin(ax_1) - wx_1)^2 + (\sin(ax_2) - wx_2)^2 + \dots + (\sin(ax_N) - wx_N)^2$$

對error中的  $w$  微分得下式：

$$\nabla \text{error} = 2(wx_1^2 - x_1 \sin(ax_1)) + 2(wx_2^2 - x_2 \sin(ax_2)) + \dots + 2(wx_N^2 - x_N \sin(ax_N))$$

當  $\nabla error = 0$ 時， $w = \frac{x_1 \sin(ax_1) + x_2 \sin(ax_2) + \dots + x_N \sin(ax_N)}{x_1^2 + x_2^2 + \dots + x_N^2}$ ，對分子積分，

$$\int_0^{2\pi} x \sin(dx) = (\sin(2\pi) - 2\pi \cos(2\pi) + c) - (\sin(0) - 0 * \cos(0) + c) = -2\pi，所以當n夠大$$

時， $w = \frac{-2\pi}{x_1^2 + x_2^2 + \dots + x_N^2} \approx 0$ 。所以任取一點 $x$ 其deterministic noise= $|\sin(ax)|$