

שאלה 1

(a) ההסתברות שדוגמה תהיה נכונה הינה :

$$p(Y = y|X = x) = \text{softmax}(wx + b)_y = \frac{e^{w_y x + b_y}}{\sum_j e^{w_j x + b_j}}$$

(b) נמצא את המינימום של הלוג :

$$\begin{aligned} \text{loss: } L(w, b) &= \arg \min \left[\sum_t -\log(p(Y = y|X = x)) \right] = \arg \min [\text{softmax}(wx + b)_y] \\ &= \arg \min \left[-\sum_t \log \left(\frac{e^{w_y x + b_y}}{\sum_{j=1}^k e^{w_j x + b_j}} \right) \right] = \arg \min \left[-\sum_t \left[\log(e^{w_y x + b_y}) - \log \left(\sum_{j=1}^k e^{w_j x + b_j} \right) \right] \right] \\ &= \arg \min \left[-\sum_t \left(w_y x + b_y - \log \left(\sum_{j=1}^k e^{w_j x + b_j} \right) \right) \right] = \arg \min \left[\sum_t \left(\log \left(\sum_{j=1}^k e^{w_j x + b_j} \right) - w_y x - b_y \right) \right] \end{aligned}$$

(c) כלל העדכון יהיה על פי הנגזרת של פונקציית ההפסד לפי w ולפי b . כלומר נחשב את הגראדיאנטים L ו- $\frac{\partial}{\partial b} L$:

$$L = \sum_t \left[\log \left(\sum_{j=1}^k e^{w_j x_t + b_j} \right) - w_{y_t} x_t - b_{y_t} \right]$$

$$\frac{\partial}{\partial w} L:$$

$$i \neq y_t \quad \frac{\partial}{\partial w} L = \sum_t \left[\frac{1}{\sum_{j=1}^k e^{w_j x_t + b_j}} \cdot e^{w_{y_t} x_t + b_{y_t}} \cdot x_t - x_t \right] = x_t \cdot \text{softmax}(wx_t + b)_i - x_t$$

$$i = y_t \quad \frac{\partial}{\partial w} L = \sum_t \left[\frac{1}{\sum_{j=1}^k e^{w_j x_t + b_j}} \cdot e^{w_i x_t + b_i} \cdot x_t \right] = x_t \cdot \text{softmax}(wx_t + b)_{y_t}$$

$$\frac{\partial}{\partial b} L:$$

$$i \neq y_t \quad \frac{\partial}{\partial b} L = \sum_t \left[\frac{1}{\sum_{j=1}^k e^{w_j x_t + b_j}} \cdot e^{w_i x_t + b_i} \right] = \text{softmax}(wx_t + b)_i$$

$$i = y_t \quad \frac{\partial}{\partial b} L = \sum_t \left[\frac{1}{\sum_{j=1}^k e^{w_j x_t + b_j}} \cdot e^{w_{y_t} x_t + b_{y_t}} - 1 \right] = \text{softmax}(wx_t + b)_{y_t} - 1$$

וזאת נציב בנוסחה : $w^t = w^{t-1} - \mu \cdot \frac{\partial}{\partial w} L$ ו- $b^t = b^{t-1} - \mu \cdot \frac{\partial}{\partial b} L$ עבור עדכון סטוכסטי.

פלט:

