CSE-512 ML (HW 6)

Name: Venkata Subba Narasa Bharath, Meadam

SBU ID: 112672986

Q1) (a) $f(\theta) = \frac{1}{m} \sum_{i=1}^{m} \left(\sum_{k=1}^{K} y_{ik} x_i^T \theta_k - \log \sum_{k=1}^{K} \exp(x_i^T \theta_k) \right)$

For gradient, $f'(\theta) = 0$

$$f'(\theta) = \frac{1}{m} \frac{\partial}{\partial \theta} \left(\sum_{i=1}^{m} \left(\sum_{k=1}^{K} y_{ik} x_i^T \theta_k - \log \sum_{k=1}^{K} \exp(x_i^T \theta_k) \right) \right)$$

$$= \frac{1}{m} \sum_{i=1}^{m} \frac{\partial}{\partial \theta} \left(\left(\sum_{k=1}^{K} y_{ik} x_i^T \theta_k - \log \sum_{k=1}^{K} \exp(x_i^T \theta_k) \right) \right)$$

$$= \frac{1}{m} \left(\sum_{k=1}^{K} y_{ik} x_i^T \right) - \frac{1}{m} \left(\frac{1}{\sum_{k=1}^{K} \exp(x_i^T \theta_k)} \frac{\partial}{\partial \theta} \left(\exp(x_i^T \theta_k) \right) \right)$$

$$= \frac{1}{m} \left(\sum_{k=1}^{K} y_{ik} x_i^T \right) - \frac{1}{m} \left(\frac{\exp(x_i^T \theta_k) * x_i^T}{\sum_{k=1}^{K} \exp(x_i^T \theta_k)} \right)$$

$$f'(\theta) = \frac{1}{m} \left(\sum_{k=1}^{m} y_{ik} x_i^T - \frac{\exp(x_i^T \theta_k) * x_i^T}{\sum_{k=1}^{K} \exp(x_i^T \theta_k)} \right)$$

(c) $f(\theta) = \log\left(\sum_{i=1}^{m} \exp(\theta_i - D)\right)$

For $f(\theta) \leq 1$

$$\Rightarrow \log\left(\sum_{i=1}^{m} \exp(\theta_i - D)\right) \leq 1$$

$$\Rightarrow \sum_{i=1}^{m} e^{(\theta_i - D)} \leq e$$

$$\Rightarrow \sum_{i=1}^{m} \frac{e^{\theta_i}}{e^{D}} \leq e$$

$$\Rightarrow \frac{1}{e^{D}} \sum_{i=1}^{m} e^{\theta_i} \leq e$$

$$\sum_{i=1}^{m} e^{\theta_i} \leq e \cdot e^{D}$$

$$\sum_{i=1}^{m} e^{\theta_i} \leq e^{D+1}$$

$$\log\left(\sum_{i=1}^{m} e^{\theta_i}\right) \leq D + 1$$

$$D \geq \log\left(\sum_{i=1}^{m} e^{\theta_i}\right) - 1$$ → Prevent underflow

For $f(\theta) \geq 1$

$$\Rightarrow \log\left(\sum_{i=1}^{m} \exp(\theta_i - D)\right) \geq$$

$$\Rightarrow \sum_{i=1}^{m} e^{(\theta_i - D)} \geq e$$

$$\Rightarrow \sum_{i=1}^{m} \frac{e^{\theta_i}}{e^{D}} \geq e$$

$$\Rightarrow \frac{1}{e^{D}} \sum_{i=1}^{m} e^{\theta_i} \geq e$$

$$\Rightarrow \sum_{i=1}^{m} e^{\theta_i} \geq e \cdot e^{D}$$

$$\Rightarrow \sum_{i=1}^{m} e^{\theta_i} \geq e^{D+1}$$

$$\Rightarrow \log\left(\sum_{i=1}^{m} e^{\theta_i}\right) \geq D+1$$

$$D \leq \log\left(\sum_{i=1}^{m} e^{\theta_i}\right) - 1$$ → Prevent overflow