Aidan Fischer CS584 HW3 3/18/2024

4) $J(\theta) = \frac{1}{T}\sum_{t=1}^{T} J^{(t)}(\theta) = \frac{-1}{T}\sum_{t=1}^{T}\sum_{j=1}^{|V|} y_j^{(t)} \log \hat{y}_j^{(t)} = -\frac{1}{T}\sum_{t=1}^{T} \log \hat{y}_j^{(t)}$

$j: y_j^{(t)} = 1$

$$\text{perplexity} = \prod_{t=1}^{T}\left(\frac{1}{P_{LM}(x^{(t+1)} \mid x^{(t)}, \ldots, x^{(1)})}\right)^{\frac{1}{T}} = \prod_{t=1}^{T}\left(\frac{1}{\hat{y}_{x_{t+1}}^{(t+1)}}\right)^{\frac{1}{T}}$$

$$\ln(\text{perplexity}) = \ln\left(\prod_{t=1}^{T}\left(\frac{1}{\hat{y}_{x_{t+1}}^{(t+1)}}\right)^{\frac{1}{T}}\right) = \frac{1}{T}\left(\sum_{t=1}^{T} \ln\left(\frac{1}{\hat{y}_{x_{t+1}}^{(t+1)}}\right)\right) =$$

$$\frac{1}{T}\left(\sum_{t=1}^{T} -\ln\left(\hat{y}_{x_{t+1}}^{(t+1)}\right)\right)$$

$$\text{perplexity} = \exp\left(\frac{1}{T}\sum_{t=1}^{T} -\ln\left(\hat{y}_{x_{t+1}}^{(t+1)}\right)\right) = \exp(J(\theta))$$