(c) $h_\theta(x) = g(\theta^T x) = e^{\eta = \theta^T x} \Rightarrow p(y^{(i)} | x^{(i)}; \theta) = \frac{e^{-e^{\theta^T x}} (e^{\theta^T x})^y}{y!}$

$$\mathcal{L}(\theta) = p(y | x; \theta)$$

$$= \prod_{i=1}^{m} p(y^{(i)} | x^{(i)}; \theta)$$

$$= \prod_{i=1}^{m} \frac{e^{-g(\eta)} \times (g(\eta))^{y^{(i)}}}{y!}$$

~~log~~ $l(\theta) = \log \mathcal{L}(\theta) = \sum_{i=1}^{m} \left\{ -g(\eta) + y^{(i)} \log g(\eta) + \log y! \right\}$

$$\Rightarrow \frac{\partial l(\theta)}{\partial \theta_j} = \sum_{i=1}^{m} \left\{ -\frac{\partial g(\eta)}{\partial \eta} \times \frac{\partial \eta}{\partial \theta_j} + \frac{y^{(i)}}{g(\eta)} \times \frac{\partial g(\eta)}{\partial \eta} \times \frac{\partial \eta}{\partial \theta_j} \right\}$$

$$\cancel{= \sum_{i=1}^{m} \left\{ -\frac{\partial (\theta^T x^{(i)})}{\partial \theta_j} + \frac{y^{(i)}}{g(\eta)} \times \partial(\theta^T x^{(i)}) \right.}$$

$$= \sum_{i=1}^{m} \left\{ -1 + \frac{y^{(i)}}{g(\eta)} \right\} \times \frac{\partial g(\eta)}{\partial \eta} \times \frac{\partial \eta}{\partial \theta_j}$$

$$= \sum_{i=1}^{m} \frac{(y^{(i)} - h_\theta(x^{(i)}))}{\cancel{g(\eta)}} \times \cancel{g(\eta)} \times x_j^{(i)}$$

$$\left(g(\eta) = h_\theta(x^{(i)}),\ \frac{\partial g(\eta)}{\partial \eta} = g(\eta) \because g(\eta) = e^{\eta} \right)$$

$$= \underline{\sum_{i=1}^{m} (y^{(i)} - h_\theta(x^{(i)})) x_j^{(i)}}$$

$\Rightarrow$ update rule, (gradient ascent)

$$\boxed{\theta_j := \theta_j + \alpha \sum_{i=1}^{m} (y^{(i)} - h_\theta(x^{(i)})) x_j^{(i)}}$$

in practice divide by m