Answer: The log-likelihood $l(\theta) = \log[\Pi p(y^{(i)}|x^{(i)};\theta)] = \sum_i(-\lambda + y\log\lambda) + c(y)$. Using the GLM assumption that $\lambda = e^{\eta} = e^{\theta^T x}$ we get that $l(\theta) = \sum_i \left(-e^{\theta^T x^{(i)}} + y^{(i)}\theta^T x^{(i)}\right) + c(y)$

Then $\nabla_\theta l(\theta) = \sum_i \left(-x^{(i)} e^{\theta^T x^{(i)}} + y^{(i)} x^{(i)}\right)$ and so our gradient ascent update rule to find $\hat{\theta}_{MLE}$ is $\theta \mapsto \theta + \alpha \nabla_\theta l(\theta) = \theta + \alpha \sum_i \left(-x^{(i)} e^{\theta^T x^{(i)}} + y^{(i)} x^{(i)}\right)$