$$L(\theta) = P(\vec{y} \mid X; \theta) = \prod p(y^{(i)} \mid x^{(i)}; \theta)$$

$$= \prod_i h(x^{(i)})^{y^{(i)}} (1 - h(x^{(i)}))^{1-y^{(i)}}$$

find θ that max $L(\theta)$;

max the $\ell(\theta) =$

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

Apply gradient descent argument $\rightarrow$ gradient ascent

$$\theta := \theta + \alpha \nabla_\theta \ell(\theta)$$: max the quadratic

Compute partial deriv for each w.r.t θ_j:

$$\frac{\partial}{\partial \theta_j} \ell(\theta) = \sum_{i=1}^{m} (y^{(i)} - h_\theta(x^{(i)})) x_j^{(i)}$$ – ← lots of algebra

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

← (batch gradient descent)