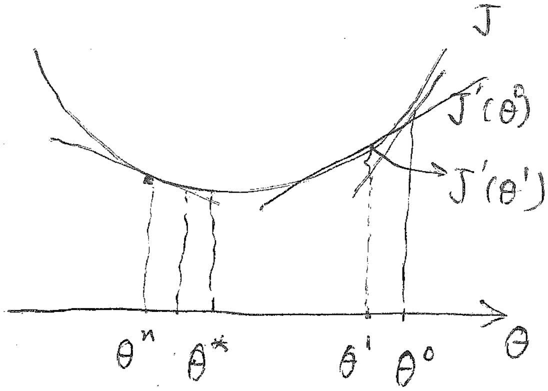

∂ : Learing rate.

$$\theta^{1} = \theta^{0} - \partial J'(\theta^{0})$$

$$\theta^{2} = \theta^{1} - \partial J'(\theta^{1})$$

$$\theta^{n+1} = \theta^{n} - \partial \underbrace{J'(\theta^{n})}_{<0}$$

$$h(x) = \frac{e^{x}}{1+e^{x}} = \frac{1}{1+e^{-x}} \qquad h(x,\theta) = \frac{e^{\theta^T x}}{1+e^{\theta^T x}} = h(\theta^T x^{(i)})$$

$$h'(x) = \frac{d}{dx}\left(\frac{1}{1+e^{-x}}\right) = \frac{e^{-x}}{(1+e^{-x})^2} = \frac{1}{1+e^{-x}} \, \frac{e^{-x}}{1+e^{-x}} = h(x)(1-h(x))$$

$$\frac{\partial}{\partial \theta_j} J(\theta) = \frac{\partial}{\partial \theta_j}\left\{ -\frac{1}{m}\sum_{i=1}^{m} y^{(i)} \log h(x^{(i)},\theta) + (1-y^{(i)}) \log(1-h(x^{(i)},\theta)) \right\}$$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left\{ \frac{y^{(i)} \frac{\partial}{\partial \theta_j} h(x^{(i)},\theta)}{h(x^{(i)},\theta)} + \frac{(1-y^{(i)})(-1)\frac{\partial}{\partial \theta_j} h(x^{(i)},\theta)}{1-h(x^{(i)},\theta)} \right\}$$

$$= -\frac{1}{m}\sum_{i=1}^{m} \frac{y^{(i)} \cancel{h(x^{(i)},\theta)} (1-h(x^{(i)},\theta)) \boxed{\frac{\partial}{\partial \theta_j}(\theta^T x^{(i)})}}{\cancel{h(x^{(i)},\theta)}} = x_j^{(i)}$$

$$- \frac{(1-y^{(i)}) \, h(x^{(i)},\theta) \cancel{(1-h(x^{(i)},\theta))} \frac{\partial}{\partial \theta_j}(\theta^T x^{(i)})}{\cancel{1-h(x^{(i)},\theta)}}$$

$$= -\frac{1}{m}\sum_{i=1}^{m}\left\{ y^{(i)}(1-h(x^{(i)},\theta)) x_j^{(i)} - (1-y^{(i)}) h(x^{(i)},\theta) x_j^{(i)} \right\}$$

$$= -\frac{1}{m}\sum_{i=1}^{m} (y^{(i)} - h(x^{(i)},\theta)) x_j^{(i)}$$

$$= \frac{1}{m}\sum_{i=1}^{m} (h(x^{(i)},\theta) - y^{(i)}) x_j^{(i)}$$