Assignment 1

CS 229 2018 Fall

Eric Lagnese
September 4, 2022

Question 1: *Find the Hessian:*

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

Notes: To me the most helpful thing in solving this problem was using mathematicalmonk for reference. Specifically lessons 15.3-15.6. If you get stuck this is an excellent resource however I will go through step-by-step on how I solved this problem and possible explain some things that weren't trivial to you before.

no title Lets move from this summation to some value for i. It will just look easier notation wise because this simplification will apply to every iteration of i. The equation will look like:

$$J(\theta) = -\frac{1}{m}y \cdot log(\frac{1}{1 + e^{-\theta^T x}}) + (1 - y)log(1 - \frac{1}{1 + e^{-\theta^T x}})$$

First I'm going to use some shorthand. $\alpha = \frac{1}{1+e^{-\theta^T x}}$ So we have:

$$J(\theta) = -\frac{1}{m}y \cdot log(\alpha) + (1 - y)log(1 - \alpha)$$

Now we ultimately need to take two partial derivatives. One with respect to θ_j and the other with respect to θ_k.

$$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{\partial}{\partial \theta_j} - \frac{1}{m}y \cdot log(\alpha) + (1 - y)log(1 - \alpha)$$

First lets break it up and deal with $y^{(i)}log(\alpha)$ first. This is simple if you use the home work hint: $\alpha' = \alpha(1 - \alpha)(\theta^T x)'$ (todo explain this little jump you made)

$$\frac{\partial}{\partial \theta_j}y \cdot log(\alpha) = y \cdot \frac{1}{\alpha}\alpha'$$

$$y^{(i)}\frac{1}{\alpha}\alpha' = y \cdot \frac{1}{\alpha}\alpha(1 - \alpha)(\theta^T x)'$$

If we look at $(\theta^T x)'$, the best way I thought about its simplification is if you expand it and then try to take the partial derivative. So $(\theta^T x)'$ becomes $\theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + ... + \theta_n x_n$. Now if you were to take the derivative with respect to θ_1 for example, most of the items would go to zero because they are constants in that dimension. So your result would be x_1. Now do the same but for θ_j and that's how you get $(\theta^T x)' = x_j$.

$$y\frac{1}{\alpha}\alpha' = y\frac{1}{\alpha}\alpha(1-\alpha)(x_j)$$

Reorganizing:

$$\begin{aligned} &= y\frac{x_j\alpha(1-\alpha)}{\alpha} \\ &= y \cdot x_j(1-\alpha) \end{aligned}$$

That should be simplified enough for us to move on. Now lets do the same for $(1-y)log(1-\alpha)$.

$$\begin{aligned} &= \frac{\partial}{\partial\theta_j}(1-y)log(1-\alpha) \\ &= (1-y)\frac{1}{1-\alpha}(-\alpha') \end{aligned}$$

Again using the substitution of the hint $(\alpha' = \alpha(1-alpha)(\theta^T x)')$:

$$\begin{aligned} &= (1-y)\frac{1}{1-\alpha}(-\alpha(1-\alpha)(\theta^T x))' \\ &= (-1+y)\frac{\alpha(1-\alpha)}{1-\alpha}(\theta^T x)' \\ &= (-1+y)\frac{\alpha(1-\alpha)}{1-\alpha}(x_j)\frac{\cdot\frac{1}{(1-\alpha)}}{\cdot\frac{1}{(1-\alpha)}} \\ &= (-1+y)\alpha(x_j) \end{aligned}$$

Putting it all together we have:

$$\begin{aligned} &= \frac{1}{m}[y \cdot x_j(1-\alpha) + (y-1)\alpha(x_j)] \\ &= \frac{1}{m}[(y - y\alpha) + (-\alpha + y \cdot \alpha)]x_j \\ &= \frac{1}{m}[(y-\alpha)x_j] \end{aligned}$$

substituting we get

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

Now we to apply $\frac{\partial}{\partial \theta_k}$ to this equation. I will start with the simplified equation using α's.

$$= \frac{\partial}{\partial \theta_k} \frac{1}{m}[(y - \alpha)x_j]$$

Take the derivative one at a time.

$$\begin{aligned} &= \frac{\partial}{\partial \theta_k} y x_j \\ &= 0 \end{aligned}$$

$$\begin{aligned} &= \frac{\partial}{\partial \theta_k} - \alpha x_j \\ &= -\alpha(1 - \alpha)(\theta^T x)' x_j \\ &= -\alpha(1 - \alpha) x_k x_j \end{aligned}$$

Put that all back together and see if you can simplify.

$$\begin{aligned} &= 0 - \alpha(1 - \alpha) x_k x_j \\ &= -\alpha(1 - \alpha) x_k x_j \end{aligned}$$

Turns out we really can't simplify it much... But that's OK. This now is in the form of a quadratic. (For more information check out: TODO link kahn academy or explain)

TODO explain most of the time I am not moving it to matrice space and am just dropping the $^{(i)}$ because it is convenient.

$$\frac{\partial^2 J(\theta)}{\partial \theta_j \, \partial \theta_k} = -\frac{1}{m} \sum_{i=1}^{m} -\alpha(1 - \alpha) x_k x_j$$

To change this to matrices we would get TODO get an explaination for how this works

$$\frac{\partial^2 J(\theta)}{\partial \theta_j \, \partial \theta_k} = x_j^T \frac{1}{m}[\alpha(1 - \alpha)] x_k$$

Now lets prove $z^T H z \geq 0$.

$$H = diag(\frac{1}{m}[\alpha(1 - \alpha)])$$

Given $1 > \alpha > 0$ we can say that $\alpha > 0$ and $(1 - \alpha) > 0$. So now $H > 0$. Now we have to determine what the rest will result in. The best way to determine this is actually in