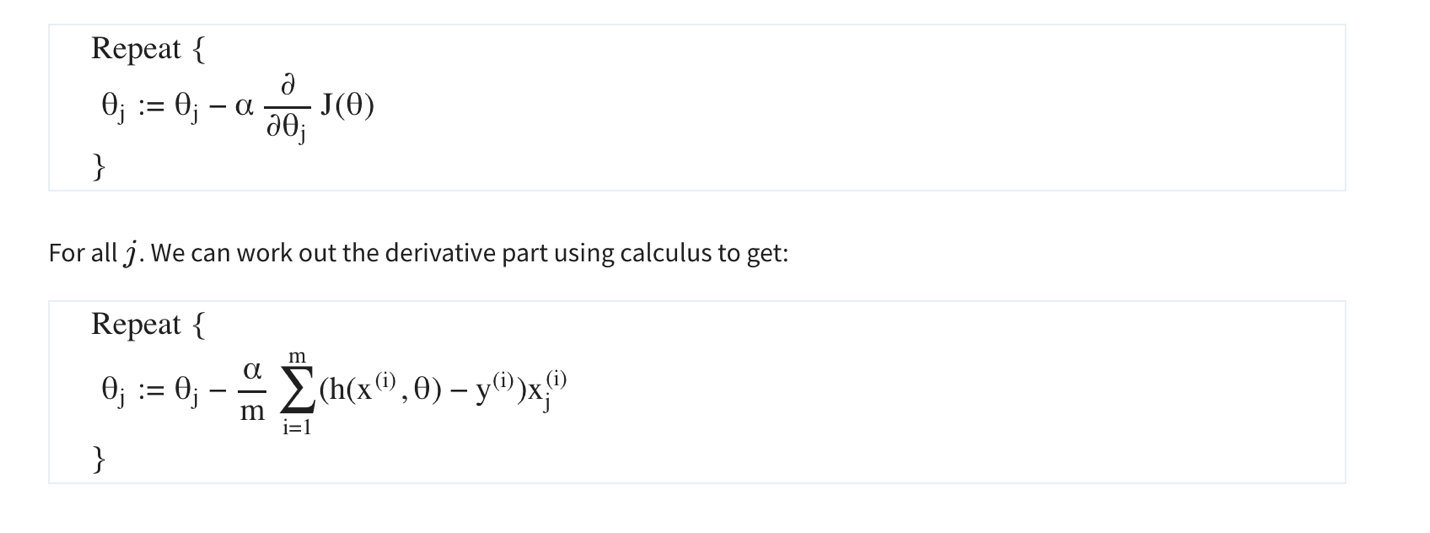
**Optional Logistic Regression: Gradient**

**This is an optional reading where I explain gradient descent in more detail. Remember, previously I gave you the gradient update step, but did not explicitly explain what is going on behind the scenes.**

The general form of gradient descent is defined as: A black text on a white background

Description automatically generated

**Partial derivative of J(θ)**

First calculate derivative of sigmoid function (it will be useful while finding partial derivative of J(θ)):A math equations on a white background

Description automatically generated

Note that we computed the partial derivative of the sigmoid function. If we were to derive h(x(i),θ)*h*(*x*(*i*),*θ*) with respect to θj*θj*​, you would get h(x(i),θ)(1−h(x(i),θ))xj(i)*h*(*x*(*i*),*θ*)(1−*h*(*x*(*i*),*θ*))*xj*(*i*)​. Note that we used the chain rule there, because we multiply by the derivative of θTx(i)*θTx*(*i*) with respect to θj*θj*​. Now we are ready to find out resulting partial derivative:A math equations on a white background

Description automatically generated

Congratulations, you now know the full derivation of logistic regression :) !