Problem 2

1. Objective function with new label

Because

Then we have the loss function

Because

We have

For the logistic regression with , its similar that

With the same minimization process

When it is clear that . When , for we have and for we have

Therefore, the logistic regression loss function for is equal to the logistic regression loss function for .

1. Gradient
2. Hessian Matrix

Because we have the diagonal matrix such that

Because sigmoid function ranges from 0 to 1

1. Convex

Because and

That is, the Hessian is the linear combination of the product of a squared term

Therefore, the negative log-likelihood is convex because the Hessian is positive semi definite

1. Newton update

Because we have a substitution