CS 760 Homework 3 by Cheng-Wei Lu

Problem 3.1

(a)

I set the iteration number of $\hat{\theta}$ update to 1000. Step size = 1/20. I change the step size by multiply it with 3/4 every 50 iterations. The stop criteria is that when iteration number reaches 1000 or when the improvement for likelihood is less than $10^(- 3)$, the parameter update will stop.

(b)

It took about 18 seconds for my computer to converge.

(c)

$$\hat{\theta} = \begin{pmatrix} 1.802(\text{class 1}) \\ 0.567(\text{class 2}) \\ -0.645(\text{class 3}) \\ 2.769(\text{sex}) \\ -1.372(\text{age}) \\ -0.217(\text{siblings/spouses aboard}) \\ -0.041(\text{parents/children aboard}) \\ 0.079(\text{fare}) \\ -0.274(\text{intercept}) \end{pmatrix}$$

(d)

likelihood = −390.5596

(e)

The distribution of $\hat{\theta}$ is $N(\theta^*, \frac{1}{N} I_{\theta^*}^{-1})$, where the θ^* is the real θ and the variance is represented by a covariance matrix $\frac{1}{N} I_{\theta^*}^{-1}$, which is of the following value :

```
[[ 1.22051416e+10  1.22051416e+10  1.22051416e+10 -4.44930791e-06
  -1.44864551e-04 -1.06086203e-05  8.97691961e-07 -1.43621015e-05
  -1.22051416e+10]
 [ 1.22051416e+10  1.22051416e+10  1.22051416e+10 -1.55204483e-05
  -1.17915114e-04 -9.54766623e-06 -5.10799684e-07 -2.53167065e-06
  -1.22051416e+10]
 [ 1.22051416e+10  1.22051416e+10  1.22051416e+10 -2.17357854e-05
  -1.06585162e-04 -9.67655103e-06 -7.51375495e-07 -2.58029852e-07
  -1.22051416e+10]
 [-5.68619406e-06 -1.67573344e-05 -2.29726716e-05  4.59177223e-05
  -7.08245260e-06 -2.35996616e-06 -2.01564456e-06 -4.16479576e-07
   1.02967288e-05]
 [-1.47441131e-04 -1.20491693e-04 -1.09161742e-04 -7.08245260e-06
   6.09370981e-05  3.96723477e-06  3.89623267e-07  1.36233277e-06
   5.92298739e-05]
 [-9.68498535e-06 -8.62403132e-06 -8.75291613e-06 -2.35996616e-06
   3.96723477e-06  3.88650985e-06 -7.85593406e-07 -5.88115590e-07
   3.80054581e-06]
 [-5.20618251e-07 -1.92910990e-06 -2.16968571e-06 -2.01564456e-06
   3.89623267e-07 -7.85593406e-07  2.37075559e-06 -8.72230171e-07
   1.14175355e-06]
 [-1.49046181e-05 -3.07418721e-06 -8.00546416e-07 -4.16479576e-07
   1.36233277e-06 -5.88115590e-07 -8.72230171e-07  6.99207938e-06
  -1.24205151e-06]
 [-1.22051416e+10 -1.22051416e+10 -1.22051416e+10  9.05984262e-06
   5.66532947e-05  4.72418071e-06 -2.76556665e-07 -1.78456807e-06
   1.22051416e+10]]
```

Problem 3.2

(a)

By the invariance property of MLE, we know that $\hat{w} = x^T\hat{\theta}$.

(b)

Since $\hat{\theta} \xrightarrow{d} N(\theta^*, \frac{1}{N}I_{\theta^*}^{-1})$, and $\hat{w} = x^T\hat{\theta}$. It implies $\hat{w} \xrightarrow{d} N(\theta^*, \frac{1}{N}x^T I_{\theta^*}^{-1}x)$.