homework-3

```
library(bis557)
```

QUESTION 1

```
set.seed(1234)
X <- matrix(rnorm(9), nrow = 3, ncol = 3)
XX <- t(X) %*% X
svd(XX)$d
#> [1] 8.2407073 1.2242344 0.1353224

p <- c(0.9999, 0.1, 1e-15)
D <- p * (1-p)
H <- t(X) %*% diag(D) %*% X
svd(H)$d
#> [1] 5.058342e-02 5.383925e-04 4.540159e-16
```

From the above results we can know the linear Hessian $X^T X$ is well conditioned, as the ratio of the largest singular value to the smallest value is approximately 61. However the longistic variation has the ratio around 10^{14}, which means it is ill-conditioned.

QUESTION 2

glm_new() implements a first-order solution for the GLM maximum likelihood problem using only gradient information. In the following examples, a constant step size and a Momentum adaptive step size are used to fit the GLM model.

```
library(bis557)

#devtools::build_vignettes()
n <- 5000
p <- 3
beta <- c(-1, 0.2, 0.1)
X <- cbind(1, matrix(rnorm(n * (p- 1)), ncol = p - 1))
eta <- X %*% beta
lambda <- exp(eta)
y <- rpois(n, lambda = lambda)
#Example 1
fit_linear_model <- glm_new(X, y, mu_fun = function(eta) exp(eta), var_fun = identity,
                            step = 0.0001, method = "constant")
fit_linear_model$coefficients
#>             [,1]
#> [1,] -0.95730122
#> [2,]  0.19535022
#> [3,]  0.03065324
#Example 2
fit_linear_model2 <- glm_new(X, y, mu_fun = function(eta) exp(eta), var_fun = identity, step = 0.0001,
                             friction = 0.9, method = "momentum")
```

```
fit_linear_model2$coefficients
#>              [,1]
#> [1,] -0.94106260
#> [2,]  0.22389656
#> [3,]  0.05124123
```

QUESTION 3

The function multi_class() is built to classify multi-categorical variables based on logistic regression. It returns the class with the highest probability and the corresponding coefficients.

```
library(bis557)
data(iris)
X <- cbind(1, as.matrix(iris[, -5]))
y <- ifelse(iris$Species==levels(iris$Species)[1], 1, ifelse(iris$Species==levels(iris$Species)[2], 2, 3

fit_linear_model <- multi_class(X, y, itr = 50, tolerance = 1e-10)
fit_linear_model$class
#>   [1] "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1"
#>  [19] "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1"
#>  [37] "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "1" "2" "2" "2" "2" "2" "2"
#>  [55] "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "3" "2" "2"
#>  [73] "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "3" "2" "2" "2" "2" "2" "2" "2" "2"
#>  [91] "2" "2" "2" "2" "2" "2" "2" "2" "2" "2" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3"
#> [109] "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3"
#> [127] "3" "3" "3" "3" "3" "3" "3" "2" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3" "3"
#> [145] "3" "3" "3" "3" "3" "3"
fit_linear_model$coefficients
#>             [,1]       [,2]      [,3]       [,4]       [,5]
#> [1,] -22.349934 18.5270872 14.826355 -34.661353 -46.625047
#> [2,]   7.378487 -0.2453567 -2.796568   1.313643  -2.778344
#> [3,] -42.637804 -2.4652202 -6.680887   9.429385  18.286137
```

REFERENCE

Textbook CASL P124 - 138.