BST 232: Homework 3

Phuc Vu

Question 1

From slide, we have:

$$U(\boldsymbol{\beta}) = \frac{\partial l(\boldsymbol{\beta} \mid \mathbf{y})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \mathbf{x_i} - \frac{2\exp(-(y_i - \mathbf{x_i}'\boldsymbol{\beta}))}{1+\exp(-(y_i - \mathbf{x_i}'\boldsymbol{\beta}))}\mathbf{x_i}$$

Note that:

$$\begin{aligned}
\mathcal{I}(\boldsymbol{\beta}) &= -\frac{\partial^2 l(\boldsymbol{\beta} \mid \mathbf{y})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta'}} \\
&= -\sum_{i=1}^{n} \frac{-2\exp(-(y_i - \mathbf{x_i}'\boldsymbol{\beta})}{(1+\exp(-(y_i - \mathbf{x_i}'\boldsymbol{\beta})))^2}\mathbf{x_i}\mathbf{x_i}' \\
&= \sum_{i=1}^{n} \frac{2\exp(-(y_i - \mathbf{x_i}'\boldsymbol{\beta})}{(1+\exp(-(y_i - \mathbf{x_i}'\boldsymbol{\beta})))^2}\mathbf{x_i}\mathbf{x_i}'
\end{aligned}$$

```
## [1] 19
##            [,1]
##     4.873236111
## BMI 0.002394433
```

$\hat{\boldsymbol{\beta}} = (4.873, 0.0024)$ which is relatively close to the $\hat{\boldsymbol{\beta}}' = (4.865, 0.0028)$ from Homework 2

Question 2

a)

From slides:

$$\mathcal{L}(\boldsymbol{\beta}, \sigma^2 \mid \mathbf{y}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(\frac{-1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}))$$

b)

$$\mathcal{L}(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2 \mid \mathbf{y}) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp(\frac{-1}{2\hat{\sigma}^2}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}))$$

Note that $(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = SSE$ and $\hat{\sigma}^2 = \frac{1}{n}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = \frac{SSE}{n}$, therefore:

$$\begin{aligned}
\mathcal{L}(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2 \mid \mathbf{y}) &= \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp(\frac{-1}{2\hat{\sigma}^2}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})) \\
&= \frac{1}{\sqrt{2\pi\frac{SSE}{n}}} \exp(\frac{-1}{2\frac{SSE}{n}} SSE) \\
&= \frac{\sqrt{n}}{\sqrt{2\pi SSE}} \exp(\frac{-n}{2})
\end{aligned}$$

c)

$$
\begin{aligned}
\frac{\mathcal{L}(\hat{\boldsymbol{\beta}_{\boldsymbol{R}}}, \hat{\sigma}_R^2)}{\mathcal{L}(\hat{\boldsymbol{\beta}_{\boldsymbol{F}}}, \hat{\sigma}_F^2)} &= \frac{\frac{\sqrt{n}}{\sqrt{2\pi SSE_R}} \exp(\frac{-n}{2})}{\frac{\sqrt{n}}{\sqrt{2\pi SSE_F}} \exp(\frac{-n}{2})} \\
&= \frac{\sqrt{SSE_F}}{\sqrt{SSE_R}}
\end{aligned}
$$

d) With $LR = \frac{\sqrt{SSE_F}}{\sqrt{SSE_R}}$, we have:

$$
\begin{aligned}
F &= \frac{(SSE_R - SSE_F)/(df_R - df_F)}{SSE_F/df_F} \\
&= \frac{SSE_R/(df_R - df_F)}{SSE_F/df_F} - \frac{SSE_F/(df_R - df_F)}{SSE_F/df_F} \\
&= \frac{SSE_R/SSE_F}{(df_R - df_F)/df_F} - \frac{df_F}{df_R - df_F} \\
&= \frac{LR^2}{(df_R - df_F)/df_F} - \frac{df_F}{df_R - df_F} \\
&= g(LR)
\end{aligned}
$$

Note that $\frac{d}{dLR} g(LR) = 2LR/((df_R - df_F)/df_F) > 0$, therefore, g is indeed monotone

3)

a) $E(Y_i) = \beta_0 + \beta_1 \times TIME + \beta_2 \times TIME^2 + \beta_3 \times Treatment$

```
##
## Call:
## lm(formula = uptake ~ time + I(time^2) + treatment, data = epa)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.20821 -0.06062  0.01984  0.04540  0.49503
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.370774   0.052381  45.261  < 2e-16 ***
## time        -0.086652   0.024942  -3.474  0.00104 **
## I(time^2)   -0.000997   0.003994  -0.250  0.80386
## treatment   -0.165000   0.027671  -5.963  2.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1035 on 52 degrees of freedom
## Multiple R-squared:  0.8052, Adjusted R-squared:  0.794
## F-statistic: 71.64 on 3 and 52 DF,  p-value: < 2.2e-16
```

b) Null hypothesis $H_0 : \beta_1 = \beta_2 = 0$\ Alternative hypothesis: $H_A : \beta_1 \neq 0$ or $\beta_2 \neq 0$

```
## Analysis of Variance Table
##
## Model 1: uptake ~ treatment
## Model 2: uptake ~ time + I(time^2) + treatment
```

```
##   Res.Df     RSS Df Sum of Sq      F    Pr(>F)
## 1     54 2.48024
## 2     52 0.55742  2    1.9228 89.687 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Based on the $F-$test with $p-$value < 0.05, we are able to reject the null hypothesis. There is statistically significant result that $\beta_1 \neq 0$ or $\beta_2 \neq 0$

c) Let variable $Treat = 0$ when the treatment is 1 and $Treat = 1$ when the treatment is 2

$$
\begin{aligned}
E(Y_i) = & (1 - Treat) \times (\beta_0 + \beta_1 \times TIME_i + \beta_2 \times TIME_i^2) + \\
& Treat \times (\gamma_0 + \gamma_1 \times TIME_i + \gamma_2 \times TIME_i^2) \\
& = \beta_0 + \beta_1 \times TIME_i + \beta_2 \times TIME_i^2 - \\
& Treat \times (\beta_0 + \beta_1 \times TIME_i + \beta_2 \times TIME_i^2 - \gamma_0 - \gamma_1 \times TIME_i - \gamma_2 \times TIME_i^2) \\
& = \beta_0 + \beta_1 \times TIME_i + \beta_2 \times TIME_i^2 + (\gamma_0 - \beta_0) \times Treat + \\
& (\gamma_1 - \beta_1) \times Treat \times TIME_i + (\gamma_2 - \beta_2) \times Treat \times TIME_i^2 \\
& = \alpha_0 + \alpha_1 \times TIME_i + \alpha_2 \times TIME_i^2 + \alpha_3 \times Treat + \alpha_4 \times Treat \times TIME_i + \alpha_5 \times Treat \times TIME_i^2
\end{aligned}
$$

With $\alpha_0 = \beta_0$, $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$, $\alpha_3 = \gamma_0 - \beta_0$, $\alpha_4 = \gamma_1 - \beta_1$, $\alpha_5 = \gamma_2 - \beta_2$.

Let $\boldsymbol{\beta}$ contains all of the regression coefficients from the above model. Then $\boldsymbol{\beta} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)'$ If the null hypothesis is $H_0 : \beta_1 = \gamma_1$ and $\beta_2 = \gamma_2$, we have this is tantamount to $\alpha_4 = 0$ and α_5=0. Therefore the $\boldsymbol{C}$ correspond to this null hypothesis would be

$$
\begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}
$$

d)

```
## Analysis of Variance Table
##
## Model 1: uptake ~ time + I(time^2) + treatment
## Model 2: uptake ~ time * treatment + I(time^2) * treatment + treatment
##   Res.Df     RSS Df Sum of Sq      F    Pr(>F)
## 1     52 0.55742
## 2     50 0.41877  2   0.13866 8.2776 0.0007847 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Based on the $F-$test with $p-$value < 0.05, we are able to reject the null hypothesis. There is statistically significant result that $\beta_1 \neq \gamma_1$ or $\beta_2 \neq \gamma_2$

Code

Important: The blank R chunk below automatically compiles all of the code chunks you wrote above into one long appendix. Do not remove it or write anything in it (unless you want to delete the code appendix for some reason). Again, ensure that you've labeled each chunk so that the teaching team knows which code pertains to which question.

```
knitr::opts_chunk$set(
  tidy = T, results = 'hold', echo = FALSE, warning = FALSE, message = FALSE,
  out.width = "50%"
)
```

```r
# This code cell sets the template format.
# You can change `out.width` to change the plot size.
# Note you may need to run install.packages("formatR") for this cell to work.

# You can also load your libraries here, like so:
library(ggplot2)
library(dplyr)


# Code for question 1
# helper functions
  score_i_calculate<-function(y,X,beta,i){
    score<-as.matrix(t(X[i,])-(2*exp(-(y[i]-X[i,]%*%beta))/(1+exp(-(y[i]-X[i,]%*%beta))))%*%t(X[i,]))
    return(score)
  }

  information_i_calculate<-function(y,X,beta,i){
    scaler<-as.numeric(exp(-(y[i]-X[i,]%*%beta)*2)/(1+exp(-(y[i]-X[i,]%*%beta)))^2)
    matrix<-X[i,]%*%t(X[i,])

    info<-matrix*scaler
    return(info)
  }


  distance_calculation<-function(beta_1,beta_2){
    distance<-sum((beta_1-beta_2)^2)
    return(distance)
  }

  score_calculate<-function(y,X,beta){
    scores<-lapply(1:length(y),function(x) score_i_calculate(y=y,X=X,beta=beta,i=x))
    return(Reduce(`+`,scores))
  }

  information_calculate<-function(y,X,beta){
    infos<-lapply(1:length(y),function(x) information_i_calculate(y=y,X=X,beta=beta,i=x))
    return(Reduce(`+`,infos))
  }

# main function

MLE_logistic<-function(X=NULL,y=NULL,initial_beta=NULL,epsilon=0.001){

  # data preparation
  X <- cbind(rep(1,dim(X)[1]),X)
  if (is.null(initial_beta)){initial_beta=as.matrix(rep(0,dim(X)[2]))}
  distance<-1
  i=1
  new_beta<-initial_beta
  while((distance>epsilon)&(i<1000)){
    old_beta<-new_beta
    new_beta<-old_beta+solve(information_calculate(y,X,old_beta))%*%t(score_calculate(y,X,old_beta))
```

```
    distance<-distance_calculation(new_beta,old_beta)
    i<-i+1
  }
  print(i)
  return(new_beta)
}
data<-read.csv("/Users/quangphucvu/Downloads/hers.csv")
data<-na.omit(data)
Y<-log(data$LDL)
X<-as.matrix(data |> select(c(BMI)))
MLE_logistic(X=X,y=Y,initial_beta = c(4.94,0),epsilon = 0.001)
epa<-read.table("epa.dat",header=TRUE)
model3a<- lm(uptake~time + I(time^2)+treatment,data=epa)
summary(model3a)
model3b<-lm(uptake~treatment,data=epa)
anova(model3b,model3a)
model3c<-lm(uptake~time*treatment + I(time^2)*treatment+treatment,data=epa)
anova(model3a, model3c)
```