Homewrk2

Chenyi Yu

2018.2.8

Question 1

(a)

The density function is:

$$p(x;\theta) = \frac{1}{\pi[1+(x-\theta)^2]} \tag{1}$$

The likelihood function is:

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} p(x_i;\theta) && (2) \\ &= \prod_{i=1}^{n} \frac{1}{\pi[1+(x_i-\theta)^2]} && (3) \\ & && (4) \end{aligned}$$

The log-likelihood function is:

$$\begin{aligned} l(\theta) &= \ln(L(\theta)) && (5) \\ &= \ln(\prod_{i=1}^{n} p(x_{i=1};\theta)) && (6) \\ &= \ln(\prod_{i=1}^{n} \frac{1}{\pi[1+(x_i-\theta)^2]}) && (7) \\ &= \sum_{i=1}^{n} \ln(\frac{1}{\pi[1+(x_i-\theta)^2]}) && (8) \\ &= -n\ln\pi - \sum_{1}^{n} \ln[1+(\theta-x_i)^2] && (9) \\ & && (10) \end{aligned}$$

The first derivative of the log-likelihood function is:

$$l^{'}(\theta) = -2\sum_{i=1}^{n} \frac{\theta-x_i}{1+(\theta-x_i)^2} \tag{11}$$

The second derivative of the log-likelihood function is:

$$l^{''}(\theta) = -2\sum_{i=1}^{n} \frac{1-(\theta - x_i)^2}{[1+(\theta - x_i)^2]^2} \tag{12}$$

And,

$$\begin{align}
I(\theta) &= n\int_{-\infty}^{\infty} \frac{[p^{'}(x)]^2}{p(x)}dx \tag{13}\\
&= \frac{4n}{\pi}\int_{-\infty}^{\infty} \frac{x^2}{(1+x^2)^3}dx \tag{14}\\
& \tag{15}
\end{align}$$

Let $x = \tan(t)$, $t \in (-\frac{\pi}{2}, \frac{\pi}{2})$, then we have:

$$\begin{align}
I(\theta) &= \frac{4n}{\pi}\int_{-\pi/2}^{\pi/2} \frac{\tan^2(t)}{(1+tan^2(t))^3}d\tan(t) \tag{16}\\
&= \frac{4n}{\pi}\int_{-\pi/2}^{\pi/2} \frac{\tan^2(t)}{(\frac{1}{\sec^2(t)})^3}\frac{1}{\sec^2(t)}dt \tag{17}\\
&= \frac{4n}{\pi}\int_{-\pi/2}^{\pi/2} \frac{\sin^2(t)}{\cos^2(t)}\cos^4(t)\,dt \tag{18}\\
&= \frac{4n}{\pi}\int_{-\pi/2}^{\pi/2} \sin^2(t)\cos^2(t)\,dt \tag{19}\\
&= (\frac{4n}{\pi}) * (\frac{\pi}{8}) \tag{20}\\
&= \frac{n}{2} \tag{21}
\end{align}$$

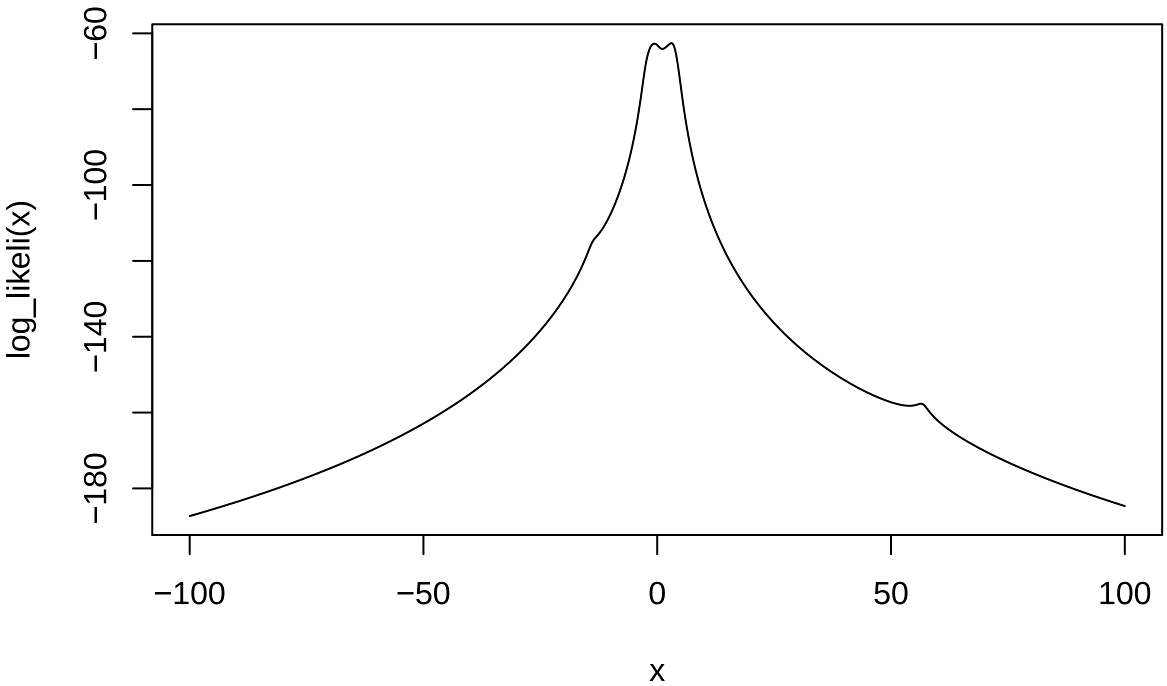
(b)

```
x <- c(1.77, -0.23, 2.76, 3.80, 3.47, 56.75, -1.34, 4.24, -2.44,
       3.29, 3.71, -2.40, 4.53, -0.07, -1.05, -13.87, -2.53, -1.75)
log_1 <- function(theta,x) {
  -log(pi)-log(1+(theta-x)^2)
}

log_likeli <- function(theta) {
  log_1(theta,1.77)+log_1(theta,-0.23)+log_1(theta,2.76)+log_1(theta,3.80)+log_1(theta,3.47)+
    log_1(theta,56.75)+log_1(theta,-1.34)+log_1(theta,4.24)+log_1(theta,-2.44)+log_1(theta,3.29)+
    log_1(theta,3.71)+log_1(theta,-2.40)+log_1(theta,4.53)+log_1(theta,-0.07)+log_1(theta,-1.05)+
    log_1(theta,-13.87)+log_1(theta,-2.53)+log_1(theta,-1.75)
}


neg_log_likeli <- function(theta) {
```

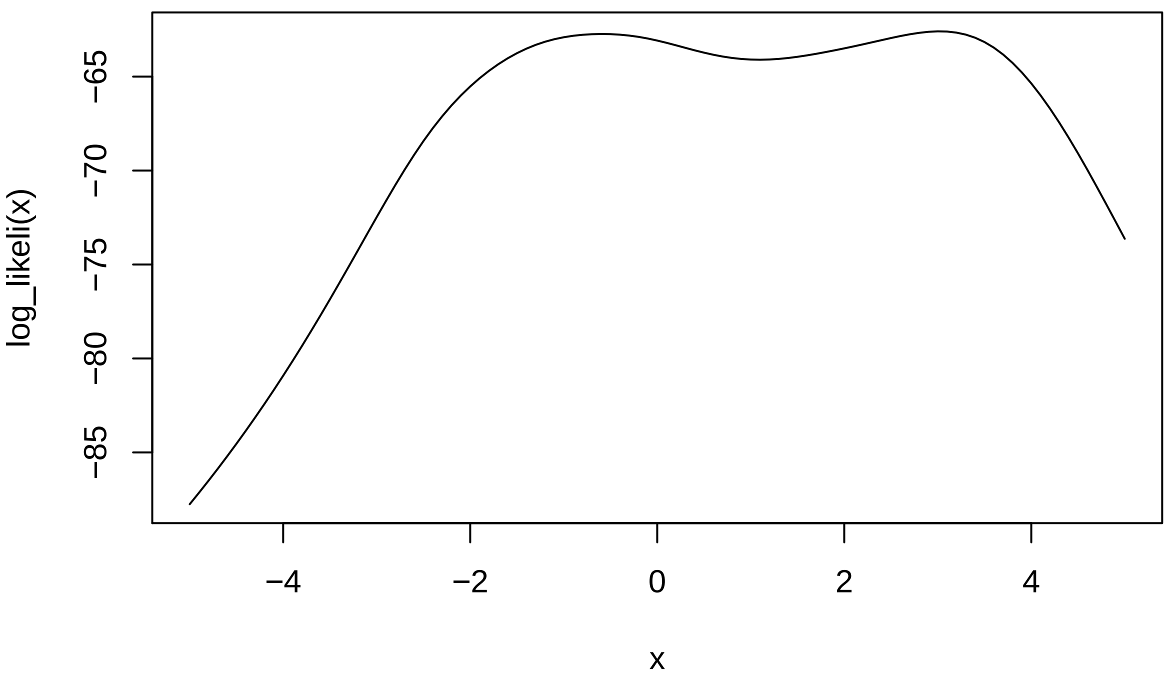
```
  -(log_1(theta,1.77)+log_1(theta,-0.23)+log_1(theta,2.76)+log_1(theta,3.80)+log_1(theta,3.47)+
      log_1(theta,56.75)+log_1(theta,-1.34)+log_1(theta,4.24)+log_1(theta,-2.44)+log_1(theta,3.29)+
      log_1(theta,3.71)+log_1(theta,-2.40)+log_1(theta,4.53)+log_1(theta,-0.07)+log_1(theta,-1.05)+
      log_1(theta,-13.87)+log_1(theta,-2.53)+log_1(theta,-1.75))
}

oc_neg_11 <- nlminb(-11,neg_log_likeli)
theta_neg_11 <- oc_neg_11$par
oc_neg_1 <- nlminb(-1,neg_log_likeli)
theta_neg_1 <- oc_neg_1$par
oc_0 <- nlminb(0,neg_log_likeli)
theta_0 <- oc_0$par
oc_1.5 <- nlminb(1.5,neg_log_likeli)
theta_1.5 <- oc_1.5$par
oc_4 <- nlminb(4,neg_log_likeli)
theta_4<- oc_4$par
oc_4.7 <- nlminb(4.7,neg_log_likeli)
theta_4.7 <- oc_4.7$par
oc_7 <- nlminb(7,neg_log_likeli)
theta_7 <- oc_7$par
oc_8 <- nlminb(8,neg_log_likeli)
theta_8 <- oc_8$par
oc_38 <- nlminb(38,neg_log_likeli)
theta_38 <- oc_38$par


v <- c(theta_neg_11,theta_neg_1,theta_0,theta_1.5,theta_4,theta_4.7,theta_7,theta_8,theta_38)
oc_MLE <- data.frame(v)
rownames(oc_MLE) <- c('-11','-1','0','1.5','4','4.7','7','8','38')
library(knitr)
kable(oc_MLE, format="markdown",col.names =c('MLE'),caption="MLE for theta using the Newton-Raphson
      method", padding=2)
```

	MLE
-11	-0.5914762
-1	-0.5914735
0	-0.5914740
1.5	3.0213454
4	3.0213454
4.7	3.0213455
7	3.0213455
8	-0.5914708
38	-0.5914735

```
curve(log_likeli,from=-100,to=100,n=10000)
```




```
curve(log_likeli,from=-5,to=5)
```



Therefore, sample mean is fairly good for a starting point in this case.

(c)

```
fixedpoint <- function(fun, x0, tol=1e-07, niter=500) {
  x_o <- x0
  x_n <- fun(x_o)
```

```
  for (i in 1:niter) {
    x_o <- x_n
    x_n <- fun(x_o)
    if ( abs((x_n-x_o)) < tol )
      return(x_n)
  }
  stop
  return('N/A')
}

start=c(-11,-1,0,1.5,4,4.7,7,8,38)
neg_log_deri <- function(theta,x) {
  2*(theta-x)/(1+(theta-x)^2)
}

g_0.64<- function(theta) {
  -0.64*(neg_log_deri(theta,1.77)+ neg_log_deri(theta,-0.23)+ neg_log_deri(theta,2.76)+
           neg_log_deri(theta,3.80)+ neg_log_deri(theta,3.47)+ neg_log_deri(theta,56.75)+
           neg_log_deri(theta,-1.34)+ neg_log_deri(theta,4.24)+ neg_log_deri(theta,-2.44)+
           neg_log_deri(theta,3.29)+ neg_log_deri(theta,3.71)+ neg_log_deri(theta,-2.40)+
           neg_log_deri(theta,4.53)+ neg_log_deri(theta,-0.07)+ neg_log_deri(theta,-1.05)+
           neg_log_deri(theta,-13.87)+ neg_log_deri(theta,-2.53)+ neg_log_deri(theta,-1.75 ))+theta
}

fix_0.64 <- NULL
for (i in c(-11,-1,0,1.5,4,4.7,7,8,38)) {
  fix_0.64 <- append(fix_0.64,fixedpoint(g_0.64,x0=i))
}

g_1<- function(theta) {
  -(neg_log_deri(theta,1.77)+ neg_log_deri(theta,-0.23)+ neg_log_deri(theta,2.76)+
      neg_log_deri(theta,3.80)+ neg_log_deri(theta,3.47)+ neg_log_deri(theta,56.75)+
      neg_log_deri(theta,-1.34)+ neg_log_deri(theta,4.24)+ neg_log_deri(theta,-2.44)+
      neg_log_deri(theta,3.29)+ neg_log_deri(theta,3.71)+ neg_log_deri(theta,-2.40)+
      neg_log_deri(theta,4.53)+ neg_log_deri(theta,-0.07)+ neg_log_deri(theta,-1.05)+
      neg_log_deri(theta,-13.87)+ neg_log_deri(theta,-2.53)+ neg_log_deri(theta,-1.75))+theta
}

fix_1 <- NULL

for (i in c(-11,-1,0,1.5,4,4.7,7,8,38)) {
  fix_1 <- append(fix_1,fixedpoint(g_1,x0=i))
}

g_0.25<- function(theta) {
  -0.25*(neg_log_deri(theta,1.77)+ neg_log_deri(theta,-0.23)+ neg_log_deri(theta,2.76)+
           neg_log_deri(theta,3.80)+ neg_log_deri(theta,3.47)+ neg_log_deri(theta,56.75)+
           neg_log_deri(theta,-1.34)+ neg_log_deri(theta,4.24)+ neg_log_deri(theta,-2.44)+
           neg_log_deri(theta,3.29)+ neg_log_deri(theta,3.71)+ neg_log_deri(theta,-2.40)+
           neg_log_deri(theta,4.53)+ neg_log_deri(theta,-0.07)+ neg_log_deri(theta,-1.05)+
           neg_log_deri(theta,-13.87)+ neg_log_deri(theta,-2.53)+ neg_log_deri(theta,-1.75))+theta
}
```

```
fix_0.25 <- NULL

for (i in c(-11,-1,0,1.5,4,4.7,7,8,38)) {
  fix_0.25 <- append(fix_0.25,fixedpoint(g_0.25,x0=i))
}

FP <- cbind(fix_0.25,fix_0.64,fix_1)

library(knitr)
kable(FP)
```

fix_0.25	fix_0.64	fix_1
-0.591473587999266	-0.591473512207137	-0.591473474719358
-0.591473601088113	-0.591473513198491	N/A
-0.59147345262032	-0.591473526770121	N/A
3.02134543948525	N/A	N/A
3.02134544141267	-0.591473525808221	N/A
3.02134544032359	-0.591473508198237	N/A
3.02134544282813	N/A	N/A
3.02134544078601	-0.591473508258934	N/A
3.02134543943626	N/A	N/A

(d)

```
I <- function(x) diag(9,nrow=length(x))

neg_log_likeli_deri <- function(theta) {
  neg_log_deri(theta,1.77)+ neg_log_deri(theta,-0.23)+ neg_log_deri(theta,2.76)+
    neg_log_deri(theta,3.80)+ neg_log_deri(theta,3.47)+ neg_log_deri(theta,56.75)+
    neg_log_deri(theta,-1.34)+ neg_log_deri(theta,4.24)+ neg_log_deri(theta,-2.44)+
    neg_log_deri(theta,3.29)+ neg_log_deri(theta,3.71)+ neg_log_deri(theta,-2.40)+
    neg_log_deri(theta,4.53)+ neg_log_deri(theta,-0.07)+ neg_log_deri(theta,-1.05)+
    neg_log_deri(theta,-13.87)+ neg_log_deri(theta,-2.53)+ neg_log_deri(theta,-1.75)
}

oc_neg_11 <- nlminb(start = -11,neg_log_likeli,neg_log_likeli_deri,I)
T_neg_11 <- oc_neg_11$par
oc_neg_1 <- nlminb(start = -1,neg_log_likeli,neg_log_likeli_deri,I)
T_neg_1 <- oc_neg_1$par
oc_0 <- nlminb(start = 0,neg_log_likeli,neg_log_likeli_deri,I)
T_0 <- oc_0$par
oc_1.5 <- nlminb(start = 1.5,neg_log_likeli,neg_log_likeli_deri,I)
T_1.5 <- oc_1.5$par
oc_4 <- nlminb(start = 4,neg_log_likeli,neg_log_likeli_deri,I)
T_4 <- oc_4$par
oc_4.7 <- nlminb(start = 4.7,neg_log_likeli,neg_log_likeli_deri,I)
T_4.7 <- oc_4.7$par
oc_7 <- nlminb(start = 7,neg_log_likeli,neg_log_likeli_deri,I)
T_7 <- oc_7$par
oc_8 <- nlminb(start = 8,neg_log_likeli,neg_log_likeli_deri,I)
```

```r
T_8 <- oc_8$par

oc_MLE <- data.frame(c(theta_neg_11,theta_neg_1,theta_0,theta_1.5,theta_4,theta_4.7,theta_7,theta_8,the
rownames(oc_MLE) <- c('-11','-1','0','1.5','4','4.7','7','8','38')
library(knitr)
kable(oc_MLE, format="markdown",col.names =c('MLE'),caption="MLE for theta using Fisher Scoring
      method", padding=2)
```

	MLE
-11	-0.5914762
-1	-0.5914735
0	-0.5914740
1.5	3.0213454
4	3.0213454
4.7	3.0213455
7	3.0213455
8	-0.5914708
38	-0.5914735

(e) Comments

We can see that for three different methods, we get different optimal points. By seting different starting points, we get different answers since the starting point determines the direction and the local maximum points.

For the univariate problem, Newton's method is relatively efficient to find the maximum. It takes least steps to find the optimal point among all methods. After applying the Fisher scoring, the process runs even faster.

For the fixed-point method, the results are various. Different starting points lead to different answers and different running speed of the process.

Question 2

(a)

The density function is:

$$f(x_i, \theta) = \frac{1 - \cos(x_i - \theta)}{2\pi} \tag{22}$$

The likelihood function is:

$$L(\theta) = \prod_{i=1}^{19} \frac{1 - \cos(x_i - \theta)}{2\pi} \tag{23}$$

The log-likelihood function is:

$$l(\theta) = -19 \log 2\pi + \sum_{n=1}^{19} \log\left(1 - \cos(x_i - \theta)\right) \tag{24}$$

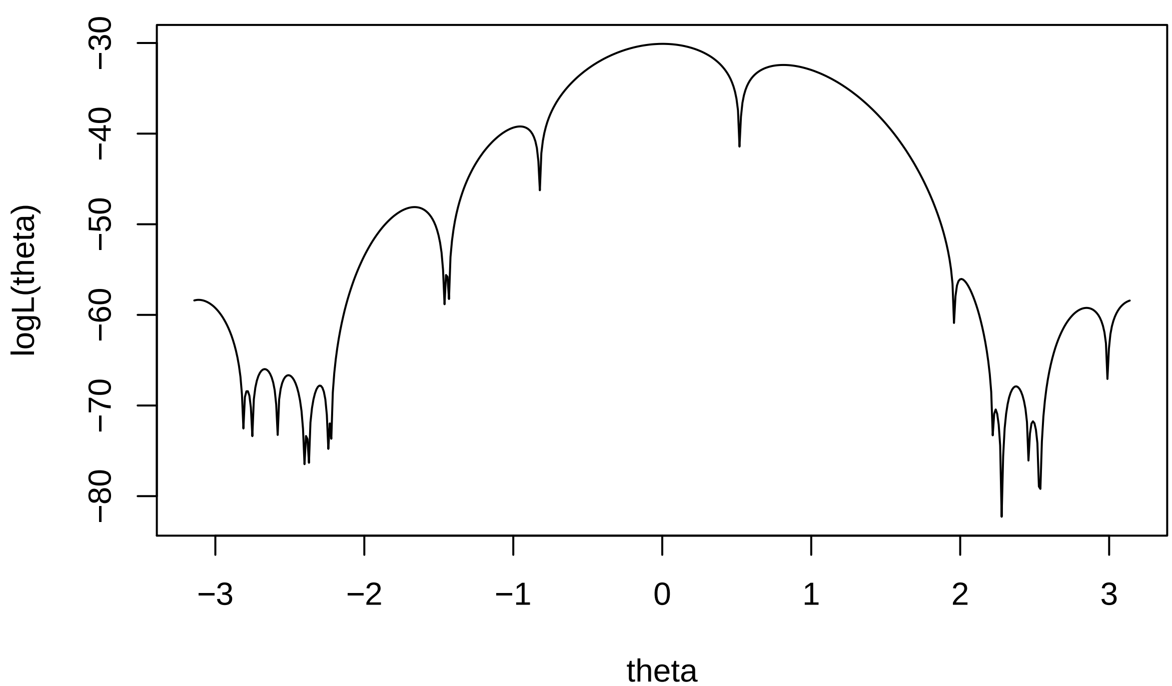
The first derivative of the log-likelihood function is:

$$l^{'}(\theta) = \sum_{i=1}^{19} \frac{\sin(\theta - x_i)}{1 - \cos(\theta - x_i)} \tag{25}$$

The second derivative of the log-likelihood function is:

$$\begin{align}
l^{''}(\theta) &= \sum_{n=1}^{19} \frac{\cos(\theta - x_i)[1 - \cos(\theta - x_i)] - \sin^2(\theta - x_i)}{[1 - \cos(\theta - x_i)]^2} \tag{26}\\
&= \sum_{n=1}^{19} \frac{\cos(\theta - x_i) - \cos^2(\theta - x_i) - \sin^2(\theta - x_i)}{[1 - \cos(\theta - x_i)]^2} \tag{27}\\
&= \sum_{n=1}^{19} \frac{\cos(\theta - x_i) - 1}{[1 - \cos(\theta - x_i)]^2} \tag{28}\\
&= \sum_{n=1}^{19} \frac{1}{\cos(\theta - x_i) - 1} \tag{29}\\
& \tag{30}
\end{align}$$

```r
x <- c(3.91, 4.85, 2.28, 4.06, 3.70, 4.04, 5.46, 3.53, 2.28, 1.96,
       2.53, 3.88, 2.22, 3.47, 4.82, 2.46, 2.99, 2.54, 0.52)

theta <- seq(-pi,pi, by=0.01)

log.lik <- function (x, theta) {
  sapply(X=theta, FUN=function(theta) sum(log((1-cos(x-theta))/(2*pi))) )
}

plot(theta, log.lik(x, theta), xlab="theta", ylab="logL(theta)", type="l")
```



(b)

$$E[X|\theta] = \frac{1}{2\pi}\int_0^{2\pi} x[1-\cos(x-\theta)]\,dx \tag{31}$$
$$= \pi - \frac{1}{2\pi}\int_0^{2\pi} x\cos(x-\theta)\,dx \tag{32}$$

Using integration by parts for the integral above, we get

$$\int_0^{2\pi} x\cos(x-\theta)\,dx = x\sin(x-\theta)\,|_0^{2\pi} - \int_0^{2\pi} sin(x-\theta)\,dx \tag{33}$$
$$= 2\pi\sin(2\pi-\theta) \tag{34}$$
$$= 2\pi\sin(-\theta)\, Since\ \sin(2\pi+x) = \sin(x) \tag{35}$$
$$= 2\pi\sin(\theta)\, Since\ \sin(-x) = \sin(x) \tag{36}$$
$$\tag{37}$$

Therefore, we have

$$E[X|\theta] = \pi - \sin(\theta) \tag{38}$$

By using the method of moment, we get

$$E[X|\theta] = \overline{X} \iff \pi - \sin(\theta) = \overline{X} \iff \theta = \arcsin(\pi - \overline{X}) \tag{39}$$

```
theta_mom <- asin(pi-mean(x))
theta_mom
```

```
## [1] -0.09539407
```

Therefore, the method-of-moments estimator is

$$\hat{\theta}_{moment} = arcsin(\pi - \overline{x}) = -0.09539 \tag{40}$$

which is close to the actual one.

(c)

```
newton <- function(f, g, x0, a, maximum) {
          x1 <- x0 - f(x0)/g(x0)
          iter <- 1
          while(abs(x1 - x0) > a & iter < maximum) {
          x0 <- x1
          x1 <- x0 - f(x0)/g(x0)
          cat("[ITER]", iter, "/", x1, fill=T)
          iter <- iter + 1
          }
          return (x1)
```

```
}

# First derivative of the log-likelihood function
deri1 <- function(theta) {
         return (-sum(sin(x-theta)/(1-cos(x-theta))))
}

# Second derivative of the log-likelihood function
deri2 <- function(theta) {
         return (-sum(1/(1-cos(x-theta))))
}

newton(deri1, deri2, theta_mom, 1e-6, 100)
```

```
## [ITER] 1 / 0.003119991
## [ITER] 2 / 0.003118157
## [ITER] 3 / 0.003118157

## [1] 0.003118157
```

(d)

```
newton(deri1, deri2, 2.7, 1e-6, 100)
```

```
## [ITER] 1 / 2.850107
## [ITER] 2 / 2.848423
## [ITER] 3 / 2.848415
## [ITER] 4 / 2.848415

## [1] 2.848415
```

If we start at $\theta_0 = 2.7$, we get the method-of-moment estimator, which converges to $\hat{\theta} = 2.848415$

```
newton(deri1, deri2, -2.7, 1e-6, 100)
```

```
## [ITER] 1 / -2.66896
## [ITER] 2 / -2.668857
## [ITER] 3 / -2.668857

## [1] -2.668857
```

If we start at $\theta_0 = -2.7$, we get the method-of-moment estimator, which converges to $\hat{\theta} = -2.668857$

(e)

Question 3

(a)

```
beetle <- data.frame(
days = c(0, 8, 28, 41, 63, 69, 97, 117, 135, 154),
```

```r
beetles = c(2, 47, 192, 256, 768, 896, 1120, 896, 1184, 1024)
)

grow_fun <- function(t,k,r){2*k / (2+(k-2) * exp(-r*t))}

model <- nls(beetles ~ grow_fun(days,k,r), data = beetle, start=list(k=1030,r=0.2), trace=TRUE)
```

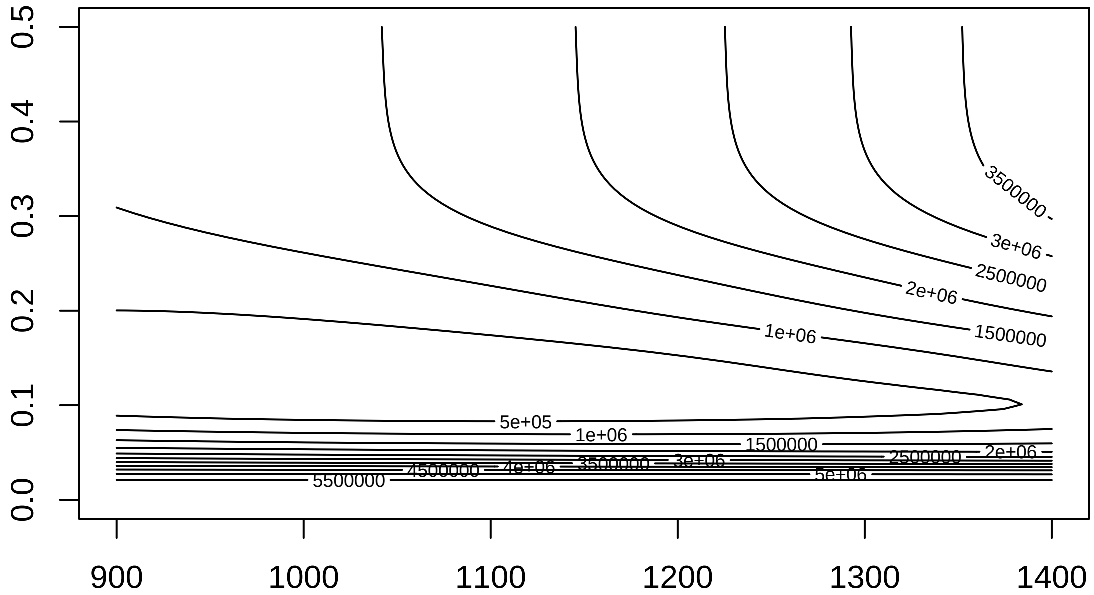
```
## 581379.6 :  1030.0    0.2
## 140584 :  947.1168989   0.1418781
## 79987.33 :  1014.9597754    0.1241012
## 73588.69 :  1043.9681300    0.1191291
## 73423.65 :  1048.8120264    0.1184152
## 73419.81 :  1049.3141568    0.1182932
## 73419.7 :  1049.3916987    0.1182726
## 73419.7 :  1049.4046222    0.1182691
## 73419.7 :  1049.4068005    0.1182685
```

```r
model
```

```
## Nonlinear regression model
##   model: beetles ~ grow_fun(days, k, r)
##    data: beetle
##         k         r
## 1049.4068    0.1183
##  residual sum-of-squares: 73420
##
## Number of iterations to convergence: 8
## Achieved convergence tolerance: 5.433e-06
```

(b)

```r
error <- function(k,r){
  return(sum((beetles-2*k/(2+(k-2)*exp(-r*days)))^2))
}
days = c(0, 8, 28, 41, 63, 69, 97, 117, 135, 154)
beetles = c(2, 47, 192, 256, 768, 896, 1120, 896, 1184, 1024)
y <- matrix(0,100,100,byrow = TRUE)
for (i in 1:100){
  for(j in 1:100){
    k <- 900+5*j
    r <- 0+0.005*i
    y[j,i] <- error(k,r)
  }
}
k <- seq(900, 1400, length.out = 100)
r <- seq(0, 0.5, length.out = 100)
contour(k, r, y)
```



(c)

```r
l <- expression(
  log(1/(sqrt(2*pi)*sigma)) - (log((2*2+2*(k-2) * exp(-r*0)) / (2*k)))^2 / (2*sigma^2) +
  log(1/(sqrt(2*pi)*sigma)) - (log((2*47+47*(k-2) * exp(-r*8)) / (2*k)))^2 / (2*sigma^2) +
  log(1/(sqrt(2*pi)*sigma)) - (log((2*192+192*(k-2) * exp(-r*28)) / (2*k)))^2 / (2*sigma^2) +
  log(1/(sqrt(2*pi)*sigma)) - (log((2*256+256*(k-2) * exp(-r*41)) / (2*k)))^2 / (2*sigma^2) +
  log(1/(sqrt(2*pi)*sigma)) - (log((2*768+768*(k-2) * exp(-r*63)) / (2*k)))^2 / (2*sigma^2) +
  log(1/(sqrt(2*pi)*sigma)) - (log((2*896+896*(k-2) * exp(-r*69)) / (2*k)))^2 / (2*sigma^2) +
  log(1/(sqrt(2*pi)*sigma)) - (log((2*1120+1120*(k-2) * exp(-r*97)) / (2*k)))^2 / (2*sigma^2) +
  log(1/(sqrt(2*pi)*sigma)) - (log((2*896+896*(k-2) * exp(-r*117)) / (2*k)))^2 / (2*sigma^2) +
  log(1/(sqrt(2*pi)*sigma)) - (log((2*1184+1184*(k-2) * exp(-r*135)) / (2*k)))^2 / (2*sigma^2) +
  log(1/(sqrt(2*pi)*sigma)) - (log((2*1024+1024*(k-2) * exp(-r*154)) / (2*k)))^2 / (2*sigma^2) )

l_k <- D(l,"k")
l_r <- D(l,"r")
l_s <- D(l,"sigma")
l_kk <- D(D(l,"k"),"k")
l_kr <- D(D(l,"k"),"r")
l_ks <- D(D(l,"k"),"sigma")
l_rr <- D(D(l,"r"),"r")
l_rs <- D(D(l,"r"),"sigma")
l_ss <- D(D(l,"sigma"),"sigma")
```

When k = 1050, r = 0.12, and sigma = 0,5,

```r
krsig<- matrix(c(1050, 0.12, 0.5))
row.names(krsig) <- c("k", "r", "sigma")
knitr::kable(krsig)
```

k	1050.00
r	0.12
sigma	0.50

```r
count <- 0
process <- TRUE
while(process){
  k <- krsig[1]
  r <- krsig[2]
  sigma <- krsig[3]
  gp <- matrix(c(eval(l_k), eval(l_r), eval(l_s)))
  gpt <- t(gp)
  ma <- matrix(c(eval(l_kk),eval(l_kr),eval(l_ks),eval(l_kr),eval(l_rr),
                eval(l_rs),eval(l_ks),eval(l_rs),eval(l_ss)),byrow=TRUE,nrow=3)
  Minv <- solve(ma)
  krsig <-  krsig - Minv %*% gp
  count <- count + 1
  if(gpt%*%gp < 1e-6 | count == 1000)
    process = FALSE
}
count
```

```
## [1] 8
```

```r
krsig2 <- matrix(c(k, r, sigma^2), ncol = 3)
colnames(krsig2) <- c("k", "r", "sigma2")
knitr::kable(krsig2)
```

k	r	sigma2
820.3801	0.19264	0.4148444

Variance of the estimates.

```r
vari <- solve(-ma)
colnames(vari) <- row.names(vari) <- c("k", "r", "sigma")
knitr::kable(vari)
```

	k	r	sigma
k	62464.5546683	-9.0597875	0.0000049
r	-9.0597875	0.0039784	0.0000000
sigma	0.0000049	0.0000000	0.0207422

Acknowledgement:

I have collaborated with Yue Gu on this assignment.