STAT580 HOMEWORK6

Xingche Guo

4/16/2017

Problem1

(a)

$$\because \quad L(\theta) = \prod_{i=1}^{n} \frac{1}{\pi\{1 + (x_i - \theta)^2\}}$$

$$\therefore \quad l(\theta) = \sum_{i=1}^{n} \log \frac{1}{\pi\{1 + (x_i - \theta)^2\}}$$

$$\therefore \quad l(\theta) = -\sum_{i=1}^{n} \log \pi\{1 + (x_i - \theta)^2\} = -\sum_{i=1}^{n} \log\{1 + (x_i - \theta)^2\} - n \log \pi$$

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

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

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

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

(b)

Proof:

$$\because \quad I(\theta) = -E(l^{''}(\theta)) = 2\sum_{i=1}^{n} E\frac{1 - (\theta - x_i)^2}{\{1 + (\theta - x_i)^2\}^2}$$

$$\therefore \quad I(\theta) = 2nE\frac{1 - (X - \theta)^2}{\{1 + (X - \theta)^2\}^2} = 2n\int_R \frac{1 - (x - \theta)^2}{\pi\{1 + (x - \theta)^2\}^3} dx$$

$$\therefore \quad I(\theta) = \frac{2n}{\pi}\int_R \frac{1 - x^2}{(1 + x^2)^3} dx$$

$$Set: \quad I_k = \int_R \frac{1}{(1 + x^2)^k} dx$$

$$\because \quad \int_R \frac{1}{(1 + x^2)^k} dx + \int_R x d(\frac{1}{(1 + x^2)^k}) = \frac{x}{(1 + x^2)^k}|_{-\infty}^{+\infty}$$

$$\therefore \quad \int_R \frac{1}{(1+x^2)^k}dx - \int_R \frac{2kx^2}{(1+x)^{k+1}}dx = 0$$

$$\therefore \quad I_k = 2kI_k - 2kI_{k+1}$$

$$\therefore \quad I_{k+1} = \frac{2k-1}{2k}I_k$$

$$\because \frac{1}{\pi}I_1 = \int_R \frac{1}{\pi(1+x^2)}dx = 1$$

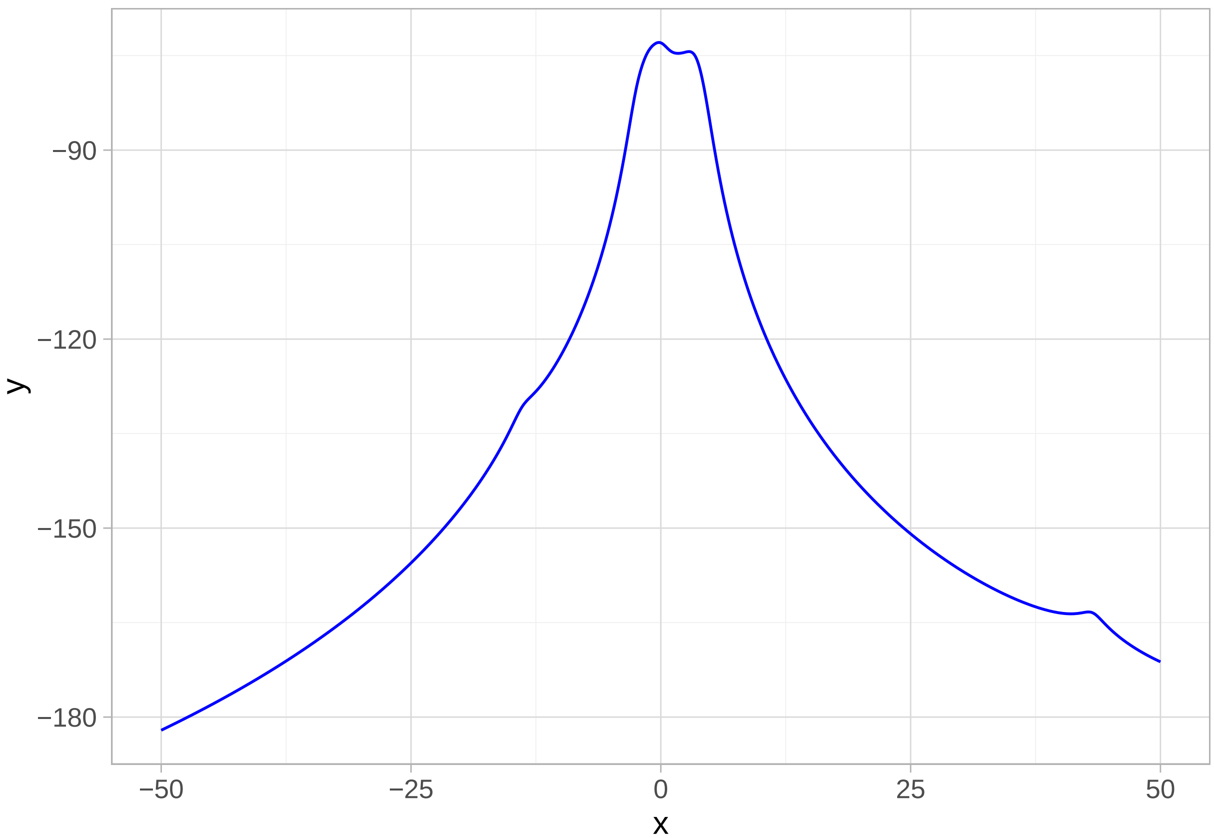
$$\therefore \quad I_1 = \pi; \quad I_2 = \frac{\pi}{2}; \quad I_3 = \frac{3\pi}{8}$$

$$I(\theta) = \frac{2n}{\pi}\int_R \frac{1-x^2}{(1+x^2)^3}dx = \frac{2n}{\pi}\int_R \frac{2-(1+x^2)}{(1+x^2)^3}dx = \frac{2n}{\pi}(2I_3 - I_2)$$

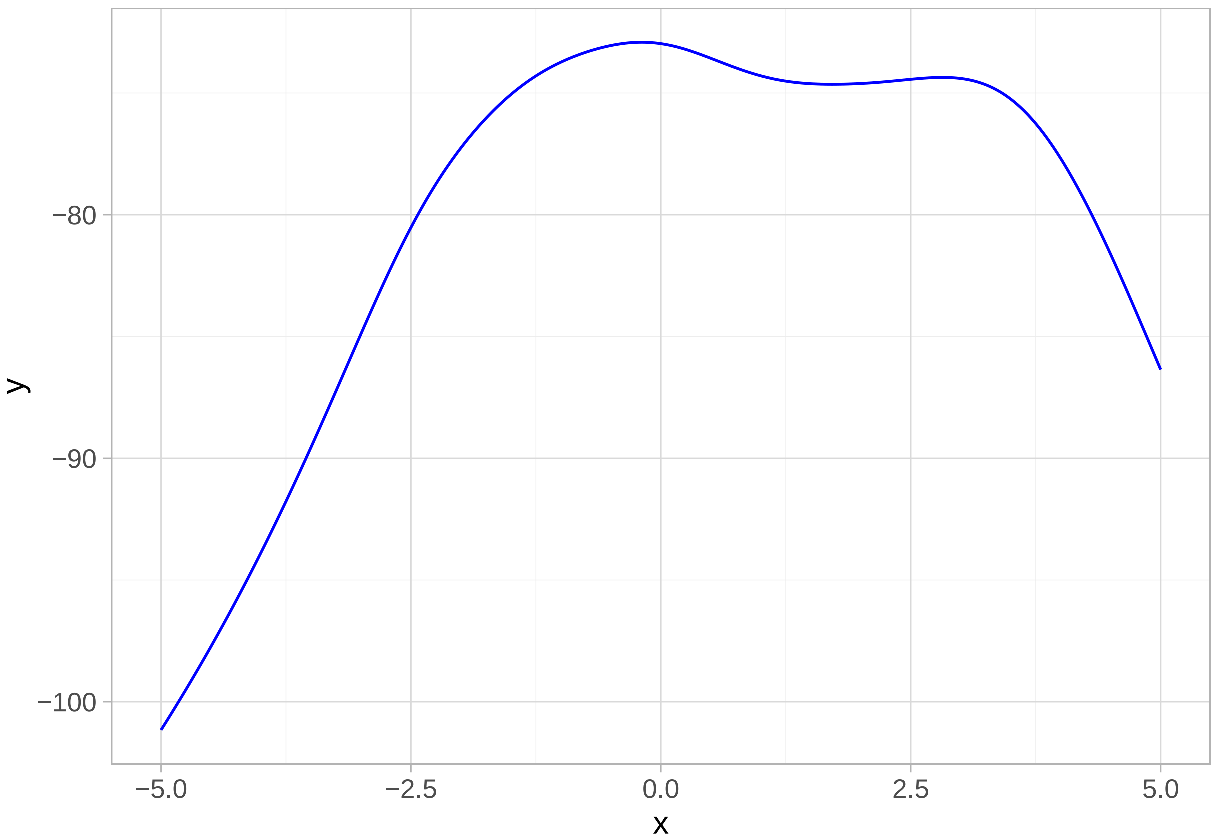
$$\therefore \quad I(\theta) = \frac{2n}{\pi}(\frac{3\pi}{4} - \frac{\pi}{2}) = \frac{n}{2}$$

(c)

```
log_cauchy <- function(a){
  X <- c(-13.87, -2.53, -2.44, -2.40, -1.75, -1.34, -1.05, -0.23, -0.07,
         0.27, 1.77, 2.76, 3.29, 3.47, 3.71, 3.80, 4.24, 4.53, 43.21, 56.75)
  n <- length(X)
  out <- -n*log(pi) - sum( log( 1 + (a-X)^2 )  )
  return(out)
}



library(ggplot2)

theta <- seq(-50,50,length.out=1000)
f_theta <- do.call("c",  lapply(theta, log_cauchy) )
d <- data.frame(x = theta, y = f_theta)
ggplot(data = d, aes(x = x, y = y)) + geom_line(color = "blue") +theme_light()
```




```
theta <- seq(-5,5,length.out=1000)
f_theta <- do.call("c",  lapply(theta, log_cauchy) )
d <- data.frame(x = theta, y = f_theta)
ggplot(data = d, aes(x = x, y = y)) + geom_line(color = "blue") +theme_light()
```



(d)

By Newton-Raphson method:

$$\theta_{k+1} = \theta_k - \frac{l^{'}(\theta_k)}{l^{''}(\theta_k)} = \theta_k - \frac{\sum_{i=1}^{n} \frac{\theta_k - x_i}{1+(\theta_k - x_i)^2}}{\sum_{i=1}^{n} \frac{1-(\theta_k - x_i)^2}{\{1+(\theta_k - x_i)^2\}^2}}$$

```
NR_cauchy <- function(p0, e = 1e-6){
  X <- c(-13.87, -2.53, -2.44, -2.40, -1.75, -1.34, -1.05, -0.23, -0.07,
         0.27, 1.77, 2.76, 3.29, 3.47, 3.71, 3.80, 4.24, 4.53, 43.21, 56.75)
f1 <- function(a){
  out <- sum(  (a-X)/( 1+(a-X)^2 )  )
  return(out)
}
f2 <- function(a){
  out <- sum(  (1-(a-X)^2)/( 1+(a-X)^2 )^2  )
  return(out)
}

while(1){
  p <- p0 - f1(p0)/f2(p0)
  if( abs(p-p0) < e || f1(p)<1e-12 ) break
  p0 <- p
}
```

```
return(p)

}

theta1_0 <- c(-11,-1,0,1.4,4.1,4.8,7,8,38)
theta1 <- do.call("c",  lapply(theta1_0, NR_cauchy) )
f_theta1 <- do.call("c",  lapply(theta1, log_cauchy) )
theta1
```

```
## [1] -16.4933088  -0.2414496  -0.1963366   1.7135868   2.8174722  -4.7577042
## [7]  41.9978045 -67.5685235  42.7953775
```

```
f_theta1
```

```
## [1] -138.61111  -72.91951  -72.91584  -74.64202  -74.36046  -99.54545
## [7] -163.48426 -193.62097 -163.31289
```

(e)

We need to use Fisher-Scoring method first:

$$\theta_{k+1} = \theta_k + \frac{l^{'}(\theta_k)}{I(\theta_k)} = \theta_k + \frac{-2\sum_{i=1}^{n}\frac{\theta_k - x_i}{1+(\theta_k - x_i)^2}}{\frac{n}{2}}$$

After several iterations, we change back to the Newton-Raphson method.

```
FS_NR_cauchy <- function(p0, e = 1e-8, fisher = 200){
  X <- c(-13.87, -2.53, -2.44, -2.40, -1.75, -1.34, -1.05, -0.23, -0.07,
         0.27, 1.77, 2.76, 3.29, 3.47, 3.71, 3.80, 4.24, 4.53, 43.21, 56.75)
  f1 <- function(a){
    out <- sum(  (a-X)/( 1+(a-X)^2 )  )
    return(out)
  }
  f2 <- function(a){
    out <- sum(  (1-(a-X)^2)/( 1+(a-X)^2 )^2  )
    return(out)
  }
  I <- length(X)/2
  for (i in 1:fisher){
    p <- p0 - 2*f1(p0)/I
    p0 <- p
  }

  while(1){
    p <- p0 - f1(p0)/f2(p0)
    if( abs(p-p0) < e || f1(p)<1e-12 ) break
    p0 <- p
  }
  return(p)
}

theta2_0 <- c(-11,-1,0,1.4,4.1,4.8,7,8,38)
theta2 <- do.call("c",  lapply(theta2_0, FS_NR_cauchy) )
```

```r
f_theta2 <- do.call("c",  lapply(theta2, log_cauchy) )
theta2
```

```
## [1] -0.1922866 -0.1922866 -0.1922866 -0.1922866  2.8174722  2.8174722
## [7]  2.8174722  2.8174722 55.7159997
```

```r
f_theta2
```

```
## [1]  -72.91582  -72.91582  -72.91582  -72.91582  -74.36046  -74.36046
## [7]  -74.36046  -74.36046 -173.24378
```

```r
optimise(log_cauchy, c(-1,1), maximum = TRUE)
```

```
## $maximum
## [1] -0.19227
##
## $objective
## [1] -72.91582
```

Problem2

(a)

```r
x <- rep(c(0.02,0.06,0.11,0.22,0.56,1.10), each = 2)
y <- c(47,76,97,107,123,139,152,159,191,201,200,207)

z <- 1/y
u <- 1/x
beta0 <- coef(lm(z~u))
theta1_0 <- as.numeric( 1/beta0[1] )
theta2_0 <- as.numeric( beta0[2]*theta1_0 )
theta0 <- c(theta1_0,theta2_0)
theta0
```

```
## [1] 195.80270885   0.04840653
```

(b)

$$Set\ \ l(\boldsymbol{\theta}) = \sum_{i=1}^{n}(y_i - \frac{\theta_1 x_i}{x_i+\theta_2})^2$$

$$l^{'}(\boldsymbol{\theta}) = \begin{pmatrix} -2\sum_{i=1}^{n}(y_i - \frac{\theta_1 x_i}{x_i+\theta_2})(\frac{x_i}{x_i+\theta_2}) \\ 2\sum_{i=1}^{n}(y_i - \frac{\theta_1 x_i}{x_i+\theta_2})\frac{\theta_1 x_i}{(x_i+\theta_2)^2} \end{pmatrix}$$

$$l^{''}(\boldsymbol{\theta}) = \begin{pmatrix} 2\sum_{i=1}^{n}(\frac{x_i}{x_i+\theta_2})^2 & 2\sum_{i=1}^{n}(\frac{x_i y_i}{(x_i+\theta_2)^2} - \frac{2\theta_1 x_i^2}{(x_i+\theta_2)^3}) \\ 2\sum_{i=1}^{n}(\frac{x_i y_i}{(x_i+\theta_2)^2} - \frac{2\theta_1 x_i^2}{(x_i+\theta_2)^3}) & 2\sum_{i=1}^{n}(\frac{3(\theta_1 x_i)^2}{(x_i+\theta_2)^4} - \frac{2\theta_1 x_i y_i}{(x_i+\theta_2)^3}) \end{pmatrix}$$

$$\therefore\ \ \boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - (l^{''}(\boldsymbol{\theta}_k))^{-1} l^{'}(\boldsymbol{\theta}_k)$$

```r
f1 <- function(a){
  p1 <- (-2)*sum( ( y - (a[1]*x)/(x+a[2]) ) * (x/(x+a[2]))  )
  p2 <- 2*sum( ( y - (a[1]*x)/(x+a[2]) ) * ( (a[1]*x)/(x+a[2])^2 )  )
  out <- matrix(c(p1,p2),nrow=2)
  return(out)
}

f2 <- function(a){
  p11 <- 2*sum( (x/(x+a[2]))^2   )
  p12 <- 2*sum(   (x*y)/(x+a[2])^2  -  (2*a[1]*x^2)/(x+a[2])^3  )
  p22 <- 2*sum(  3*(x*a[1])^2/(x+a[2])^4  -  (2*a[1]*x*y)/(x+a[2])^3   )
  out <- matrix(c(p11,p12,p12,p22),nrow=2)
  return(out)
}

NR_MM <- function(p0, e = 1e-6){
  while(1){
    p <- p0 - solve(f2(p0))%*%f1(p0)
    if( norm(p-p0,"F") < e || norm(f1(p),"F")<1e-12 ) break
    p0 <- p
  }
  return(p)
}

theta1 <- NR_MM(theta0)
theta1
```

```
##              [,1]
## [1,] 212.68374314
## [2,]   0.06412128
```

```r
f <- function(a){
  out <- sum( ( y - a[1]*x/(x+a[2]) )^2 )
  return(out)
}

f(theta0)
```

```
## [1] 1920.643
```

```r
f(theta1)
```

```
## [1] 1195.449
```

(c)

By Steepest Descent Algorithm:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \alpha_k l^{'}(\boldsymbol{\theta}_k)$$

```r
SD_MM <- function(p0, a, e = 1e-6, max_i = 100000000){
  b <- a
  for(i in 1:max_i){
    p <- p0 - b*f1(p0)
```

```
    while(f(p)>f(p0)){
      b <- b/2
      p <- p0 - b*f1(p0)
    }
    if( norm(p-p0,"F") < e || norm(f1(p),"F")<1e-8 ) break
    p0 <- p
    b <- a
  }
  if(i == max_i) cat("exceed the maximal iteration times, didn't find the minimum value.")
  return(p)
}


theta2 <- SD_MM(theta0, a = 0.15, e = 1e-7, max_i = 4000000)
theta2
```

```
##              [,1]
## [1,] 212.6290204
## [2,]   0.0640693
```

```
f(theta2)
```

```
## [1] 1195.456
```

(d)

By Gauss-Newton Algorithm:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + (A_k^T A_k)^{-1} A_k^T Z_k$$

$$where \quad A_k = \{f_1^{'}(\boldsymbol{\theta}_k), \dots, f_n^{'}(\boldsymbol{\theta}_k)\}^T; \quad Z_k = \{y_1 - f_1(\boldsymbol{\theta}_k), \dots, y_n - f_n(\boldsymbol{\theta}_k)\}^T$$

$$where \quad f_i(\boldsymbol{\theta}) = \frac{\theta_1 x_i}{x_i + \theta_2}; \quad f_i^{'}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{x_i}{x_i+\theta_2} \\ -\frac{\theta_1 x_i}{(x_i+\theta_2)^2} \end{pmatrix}$$

```
Z <- function(a){
  out <- y - a[1]*x/(x+a[2])
  out <- as.matrix(out,ncol=1)
  return(out)
}

A <- function(a){
  out1 <- x/(x+a[2])
  out2 <- -a[1]*x/(x+a[2])^2
  out <- cbind(out1,out2)
  return(out)
}

GN_MM <- function(p0, e = 1e-6, max_i = 100000000){
  for(i in 1:max_i){
    p <- p0 + solve( t(A(p0)) %*% A(p0) ) %*% t(A(p0)) %*% Z(p0)
    if( norm(p-p0,"F") < e) break
    p0 <- p
    if(i == max_i) cat("exceed the maximal iteration times, didn't find the minimum value.")
  }
```

```r
  return(p)
}

theta3 <- GN_MM(theta0)
theta3
```

```
##               [,1]
## out1 212.68374313
## out2   0.06412128
```

```r
f(theta3)
```

```
## [1] 1195.449
```