STAT230 HW 6
University of California, Berkeley

Thibault Doutre, Student ID 26980469

February 29, 2016

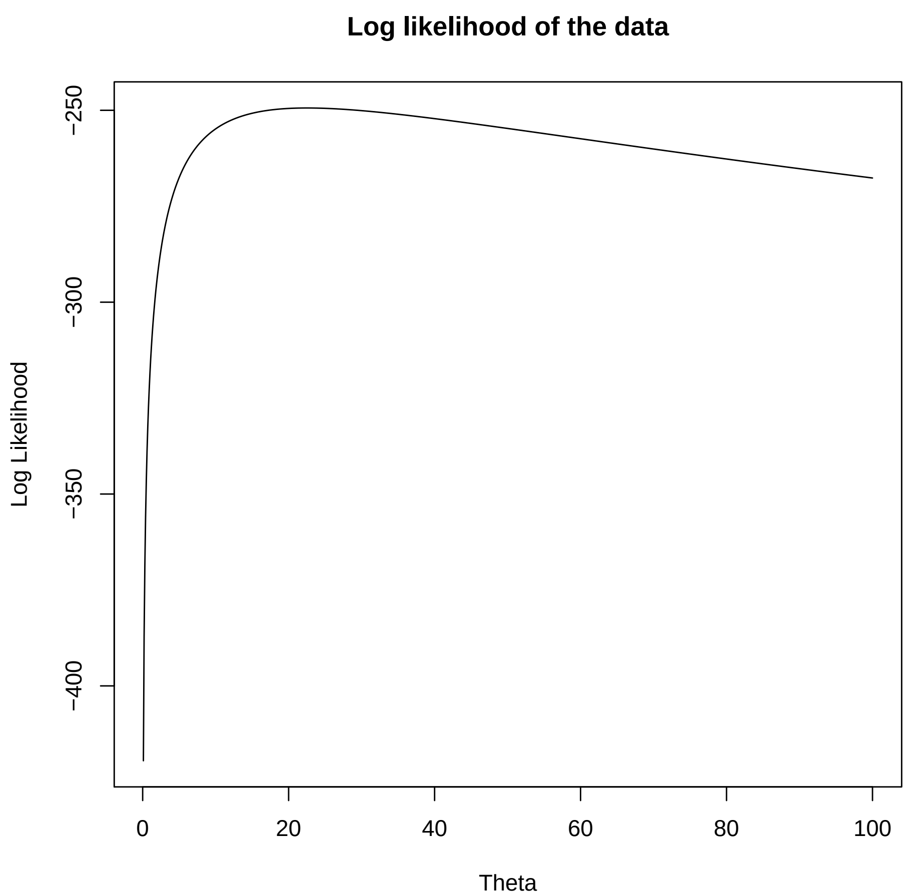
1

The log likelihood is equal to:

$$
\begin{aligned}
l(\theta|X) &= log\prod_{i=1}^{n}\frac{\theta}{(\theta+X_i)^2} &(1)\\
&= \sum_{i=1}^{n} log\frac{\theta}{(\theta+X_i)^2} &(2)\\
&= \sum_{i=1}^{n} log(\theta) - 2log(\theta+X_i) &(3)\\
&= n\times log(\theta) - 2\sum_{i=1}^{n} log(\theta+X_i) &(4)
\end{aligned}
$$

```
## # Load data
data = read.table('mle.txt')
data = unlist(data, use.names = FALSE)
n = length(data)

## # Compute log likelihood
log_likelihood_aux = function(theta){
  n*log(theta) - 2*sum(log(theta+data))
}
log_likelihood = Vectorize(log_likelihood_aux)

# Plot
```

```
thetas = seq(from = 0.1, to = 100, by = 0.1)
lls = log_likelihood(thetas)
plot(thetas,lls,
     main = "Log likelihood of the data",
     ylab = "Log Likelihood",
     xlab = "Theta",
     type="l")
```

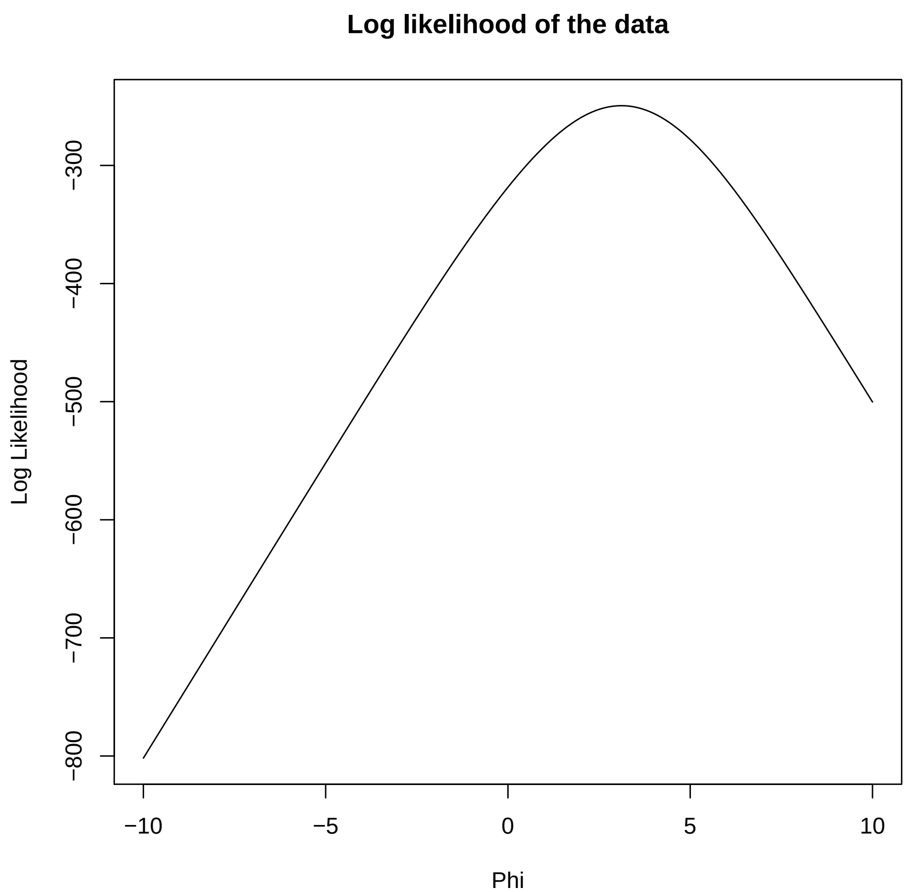


2

First, I parametrize again l with $\exp(\phi) = \theta$, so the likelihood is defined on R. I obtain:

$$l(\phi|X) = n\phi - 2\sum_{i=1}^{n} log(\exp(\phi) + X_i) \tag{5}$$

```
## # Change variable theta<-exp(phi)
log_likelihood_phi_aux = function(phi){
  n*phi - 2*sum(log(exp(phi)+data))
}
log_likelihood_phi = Vectorize(log_likelihood_phi_aux)

# Plot
phis = seq(from = -10, to = 10, by = 0.1)
lls = log_likelihood_phi(phis)
plot(phis,lls,
     main = "Log likelihood of the data",
     ylab = "Log Likelihood",
     xlab = "Phi",
     type="l")
```



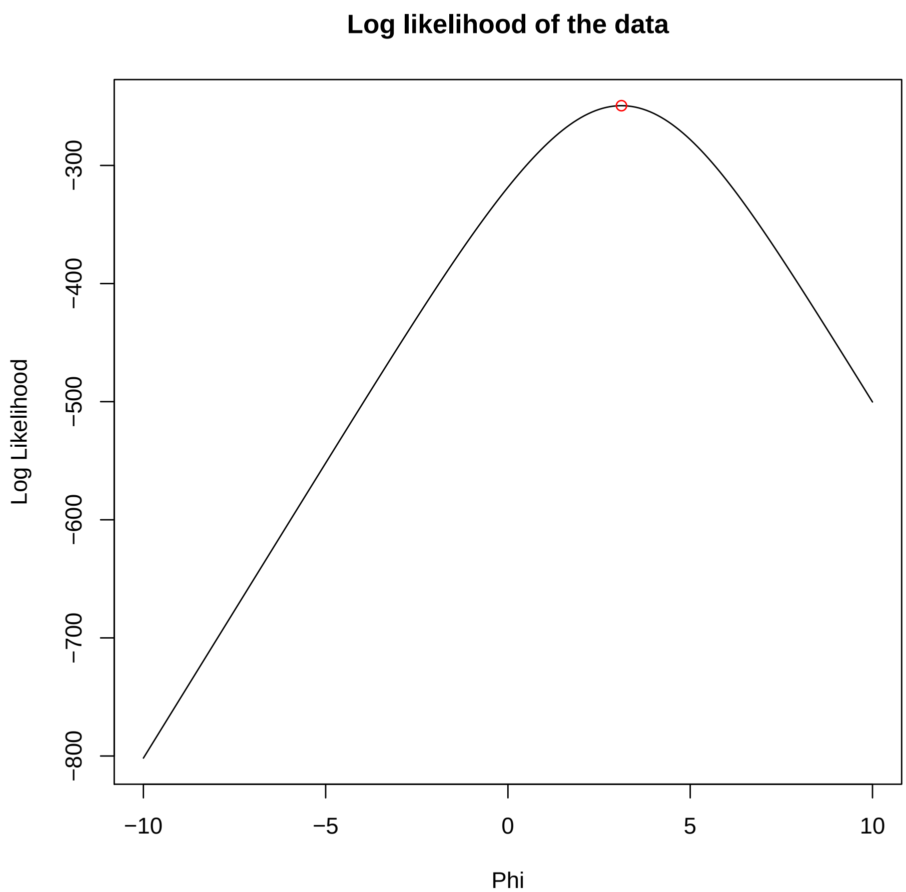
Then, I define my function to minimize, to be the opposite of the likelihood.

```
# Function to optimize
f = function(phi){
  -log_likelihood_phi_aux(phi)
}
```

Now, I can find the optimum with the optim function:

```
opt = optim(0,f,
            method="Brent",
            lower = 0,
```

```
             upper = 10)
opt

## $par
## [1] 3.113949
##
## $value
## [1] 249.3968
##
## $counts
## function gradient
##       NA       NA
##
## $convergence
## [1] 0
##
## $message
## NULL

# Plot
plot(phis,lls,
     main = "Log likelihood of the data",
     ylab = "Log Likelihood",
     xlab = "Phi",
     type="l")
points(opt$par,
       -opt$value,
       col = "red")
```



3

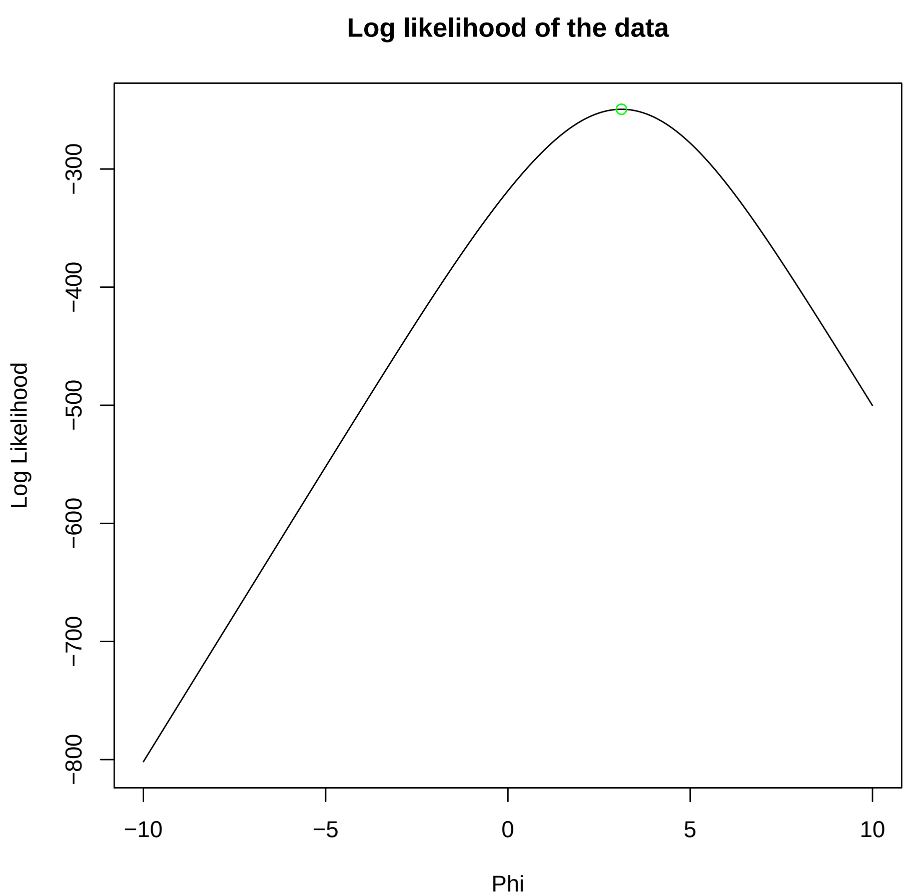
In order to optimize the log likelihood we can also compute the first two derivatives of the likelihood, and compute the Newton Raphson algorithm. I have coded it by hand: The first two derivatives are equal to:

$$\frac{\partial l(\phi|X)}{\partial \phi} = n - 2\sum_{i=1}^{n} \frac{1}{1 + X_i \exp(-\phi)} \tag{6}$$

$$\frac{\partial^2 l(\phi|X)}{\partial \phi^2} = -2\exp(-\phi)\sum_{i=1}^{n} \frac{X_i}{(1 + X_i \exp(-\phi))^2} \tag{7}$$

```
## # Optimization using derivatives

# # Compute first derivative
D_f = function(phi){
  -n + 2*sum(1/(exp(-phi)*data+1))
}
# # Compute second derivative
D2_f = function(phi){
  + 2*exp(-phi)*sum(data/(exp(-phi)*data+1)^2)
}

# # Newton raphson: one variable
newton_raphson = function(start,f,Df,D2f,eps){
  x_old=start-2*eps
  x_new=start
  while (abs(x_new-x_old)>eps){
    x_old = x_new
    x_new = x_old - Df(x_old)/D2f(x_old)
  }
  return(list(par = x_new, value = f(x_new)))
}

eps=10e-5
nr = newton_raphson(0,f,D_f,D2_f,eps)
nr

## $par
## [1] 3.113949
##
## $value
## [1] 249.3968

# Plot
plot(phis,lls,
     main = "Log likelihood of the data",
     ylab = "Log Likelihood",
     xlab = "Phi",
     type="l")

points(nr$par,
```

```
       -nr$value,
       col = "green")
```



The estimated value of theta is then:

```
theta_hat = exp(nr$par)
theta_hat

## [1] 22.50976
```

4

In order to compute the standard error of $\hat{\theta}$, I use the fact that the asymptotic variance can be computed as $var(\hat{\theta}) = -(\frac{\partial^2 l(\theta|X)}{\partial \theta^2})^{-1}$, where:

$$\frac{\partial^2 l(\theta|X)}{\partial \theta^2} = -\frac{n}{\theta^2} + 2\sum_{i=1}^{n} \frac{1}{(\theta + X_i)^2} \tag{8}$$

Then, the SE can be computed as: $SE(\hat{\theta}) = (var(\hat{\theta}))^{1/2}$

```
## # Standar error on theta

D2_log_likelihood_theta = function(theta){
  -n/theta^2+2*sum(1/(theta+data)^2)
}

# # Variance
var_hat = -1/D2_log_likelihood_theta(theta_hat)
# # Standard Error
se_hat = sqrt(var_hat)
se_hat

## [1] 5.488464
```