Homework 1
(excused for extra day for health reasons)

Jared Yu

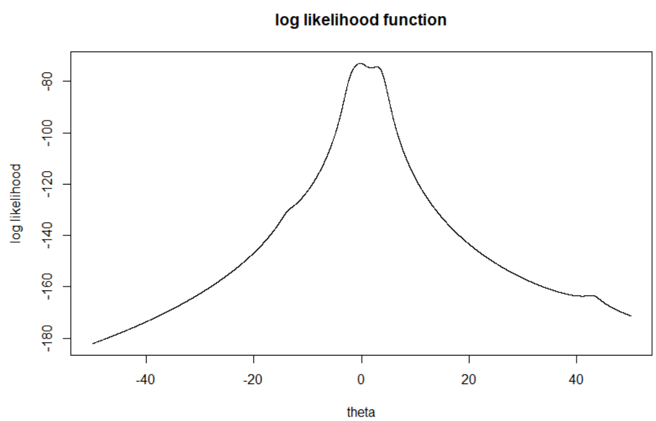
April 13, 2018

Problem 1.

a) $f_n(x_1, \cdots, x_n|\theta) = \prod_{i=1}^n 1/\pi(1+(x_i-\theta)^2)$
$= 1/\pi^n \prod_{i=1}^n (1+(x_i-\theta)^2)$
$l(\theta) = -nln\pi - \sum_{i=1}^n log(1+(x_i-\theta)^2)$
$= -nln\pi - \sum_{i=1}^n log(1+(\theta-xi)^2)$
$l'(\theta) = -\sum_{i=1}^n (2(\theta-x_i)/(1+(\theta-x_i)^2))$
$= -2 * \sum_{i=1}^n (\theta-x_i)/(1+(\theta-x_i)^2)$
$l''(\theta) = -2 * \sum_{i=1}^n [(1+(\theta-x_i)^2) - (\theta-x_i)(2(\theta-x_i))/(1+(\theta-x_i)^2)^2]$
$= -2 * \sum_{i=1}^n [1-(\theta^2-2\theta x_i+x_i^2)/(1+(\theta-x_i)^2)^2]$
$= -2 * \sum_{i=1}^n [1-(\theta-x_i)^2/(1+(\theta-x_i)^2)^2]$

b) $I(\theta) = -E(l''(\theta))$
$= 2 * \sum_{i=1}^n E[(1-(\theta-x_i)^2)/(1+(\theta-x_i)^2)^2]$
$= 2n * \int_{-\infty}^{\infty} ((1-(\theta-x)^2)/(1+(\theta-x)^2)) * (1/\pi * (1+(\theta-x)^2))dx$
$= 2n * \int_{-\infty}^{\infty} ((1-(\theta-x)^2)/(\pi(1+(\theta-x)^2)^3))$
Let $u = \theta - x$
$du = -dx$
$= 2n * \int_{-\infty}^{\infty} ((1-x^2)/(\pi * (1+x^2)^3))$
Let $x = tan\theta, dx = sec^2\theta d\theta$
$= 2n * \int_{-\pi/2}^{\pi/2} ((1-tan^2\theta)/(\pi * (1+tan^2\theta)^3)) * sec^2\theta d\theta$
$= 2n * \int_{-\pi/2}^{\pi/2} ((1-tan^2\theta)/(\pi * (sec^2\theta)^3)) * sec^2\theta d\theta$
$= 2n * \int_{-\pi/2}^{\pi/2} ((1-tan^2\theta)/(\pi * sec^4\theta))d\theta$
$= 2n/\pi * \int_{-\pi/2}^{\pi/2} cos^4\theta - sin^2\theta cos^2\theta d\theta$
$= 2n/\pi * \int_{-\pi/2}^{\pi/2} cos^2\theta(cos^2\theta - sin^2\theta)d\theta$
$= 2n/\pi * \int_{-\pi/2}^{\pi/2} cos^2\theta(cos^2\theta - (1-cos^2\theta))d\theta$

$= 2n/\pi * \int_{-\pi/2}^{\pi/2} cos^2\theta(2cos^2\theta - 1)d\theta$
$= 4n/\pi * \int_{0}^{\pi/2} 2cos^4\theta - (cos^2\theta)d\theta$
$\int cos^2\theta d\theta = \int (1 + cos2\theta/2)d\theta$
$= (1/2)(\theta + sin2\theta/2)$
$\int 2cos^4\theta d\theta = \int 2 * ((1 + cos2\theta)/2)^2 d\theta$
$= (1/2)\int (1 + cos2\theta)^2 d\theta$
$= (1/2)\int (1 + 2cos^2\theta + cos^2 2\theta d\theta$
$= (1/2)(\theta + sin2\theta) + (1/2)\int cos^2 2\theta d\theta$
$= (1/2)(\theta + sin2\theta) + (1/2)\int ((1 + cos4\theta)/2)d\theta$
$= (1/2)(\theta + sin2\theta) + (1/4)(\theta + (sin4\theta/4)) + C$
$= 3\theta/4 + (1/2)sin2\theta + (1/16)sin4\theta$
$= (4n/\pi)*(3\theta/4+(1/2)sin2\theta+(1/16)sin4\theta \mid_0^{\pi/2} -(1/2)(\theta+sin2\theta/2) \mid_0^{\pi/2}$
$)$
$= (4n/\pi)((3\pi/8 + 0 + 0) - (1/2)(\pi/2 + 0))$
$= (4n/\pi)(3\pi/8 - \pi/4)$
$= (4n/pi) * (\pi/8)$
$= n/2$

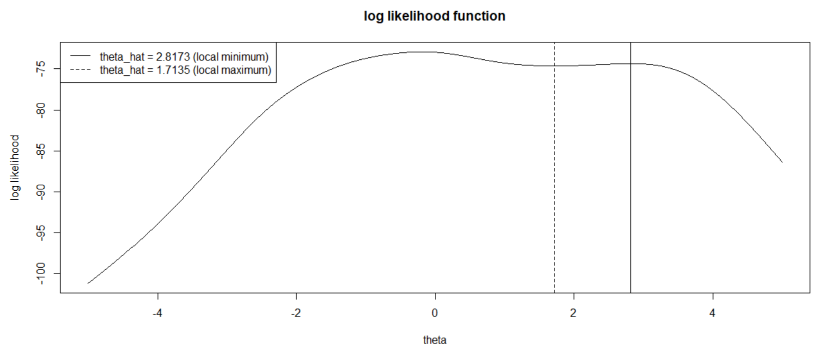


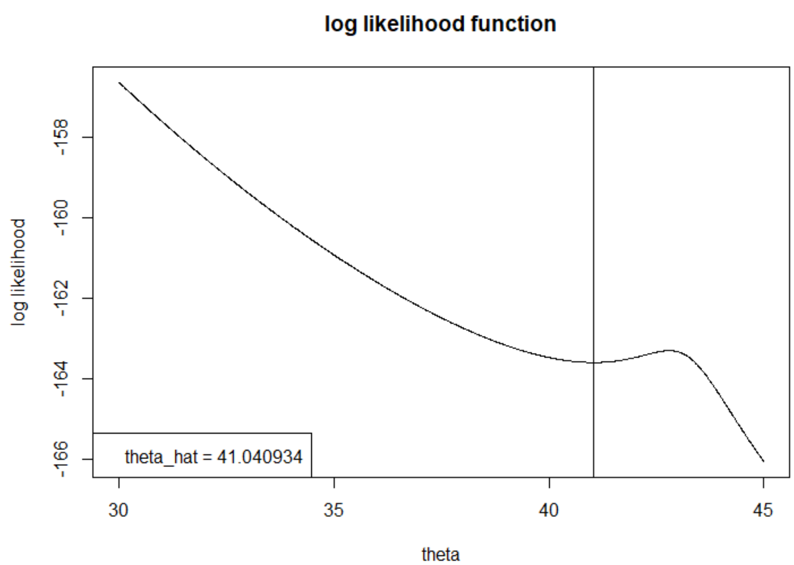
c)

Here in the plot it is possible to see that there are different local maximums and minimums which can be found along the line. Therefore, it is possible to see how the results from using the Newton-Raphson method come about.

d) After using the Newton-Raphson method with the given set of initial data points, the values are: 2.817290, 2.817499, 1.713544, 2.817641, 2.817295, 2.817536, 41.040934, 1.713432, 2.817408.

The results are mostly within an area of either 2.82 or 1.71, however there is an extreme outlier at 41.0.





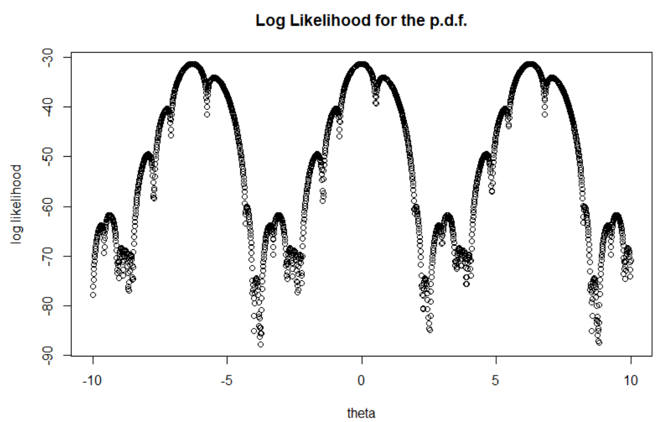
From the above two plots it's possible to understand why the previous estimators of the Newton-Raphson method were considered. The method seems to choose local minimums or maximums, which can be an issue if the goal is to obtain the global maximum instead.

e) After changing the Newton-Raphson method to include the Fisher scoring, the new results are: 2.817342, 1.713471, 1.713626, 1.713637, 2.817343, 2.817572, 2.817515, 1.713552, 2.817643.

The new results are much smoother in the sense that they no longer have any extreme outliers that are selected as false local maximums/minimums. The two values that arise are roughly 1.71 and 2.82 which are close to the global maximum. The improvement is over the previous local maximum/minimum which was considered at around 41.04.

Problem 2.

a) $f_n(x_1, \cdots, x_n|\theta) = \prod_{i=1}^{n}(1 - cos(x_i - \theta))/2\pi)$
$= \prod_{i=1}^{n}[1 - cos(x_i - \theta)]/2^n\pi^n$
$l(\theta) = \sum_{i=1}^{n} log(1 - cos(x_i - \theta)) - n * log(2\pi))$



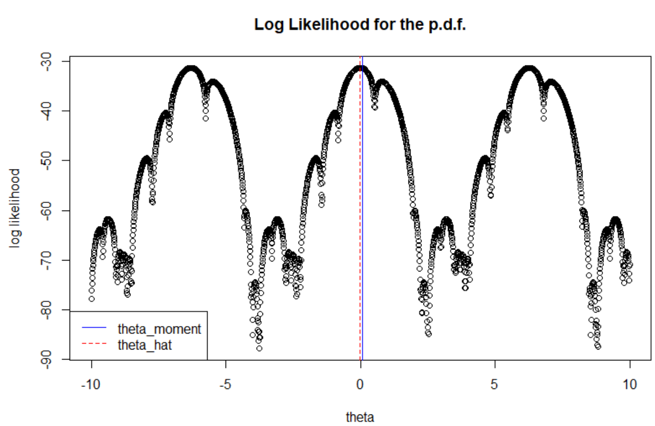
It is apparent from the plot that there are numerous local maximums within this particular range of θ which makes it difficult to be certain of a predictor.

b) $E(x) = \bar{x}$
$E(x) = \int_0^{2\pi} x * ((1 - cos(x - \theta))/2\pi)dx$
$= (1/2\pi)\int_0^{2\pi} xdx - (1/2\pi)\int_0^{2\pi} xcos(x - \theta)dx$
$= (1/2\pi)\int_0^{2\pi} xdx - (1/2\pi)\int_0^{2\pi} xd(sin(x - \theta)$
$= (1/2\pi)\int_0^{2\pi} xdx - (1/2\pi)[xsin(x - \theta)]_0^{2\pi} + (1/2\pi)\int_0^{2\pi} sin(x - \theta)dx$
$= (1/4\pi)[x^2]_0^{2\pi} - (1/2\pi)[xsin(x - \theta)]_0^{2\pi} - (1/2\pi)[cos(x - \theta)]_0^{2\pi}$
$= (1/4\pi)(4\pi^2) - (1/2\pi)[2\pi sin(2\pi - \theta] - (1/2\pi)[cos(2\pi - \theta) - cos(-\theta)]$
$= \pi - sin(2\pi - \theta) - (1/2\pi)cos(2\pi - \theta) + (1/2\pi)cos(\theta)$
Set $\bar{x} = \pi - sin(-\theta) - (1/2\pi)cos(-\theta) + (1/2\pi)cos\theta$

$= \pi + sin\theta - (1/2\pi)cos\theta + (1/2\pi)cos\theta$
$= \pi + sin\theta$
$sin\theta = \bar{x} - \pi$
$\hat{\theta}_{moment} = \arcsin(\bar{x} - \pi)$
$= 0.05844061$

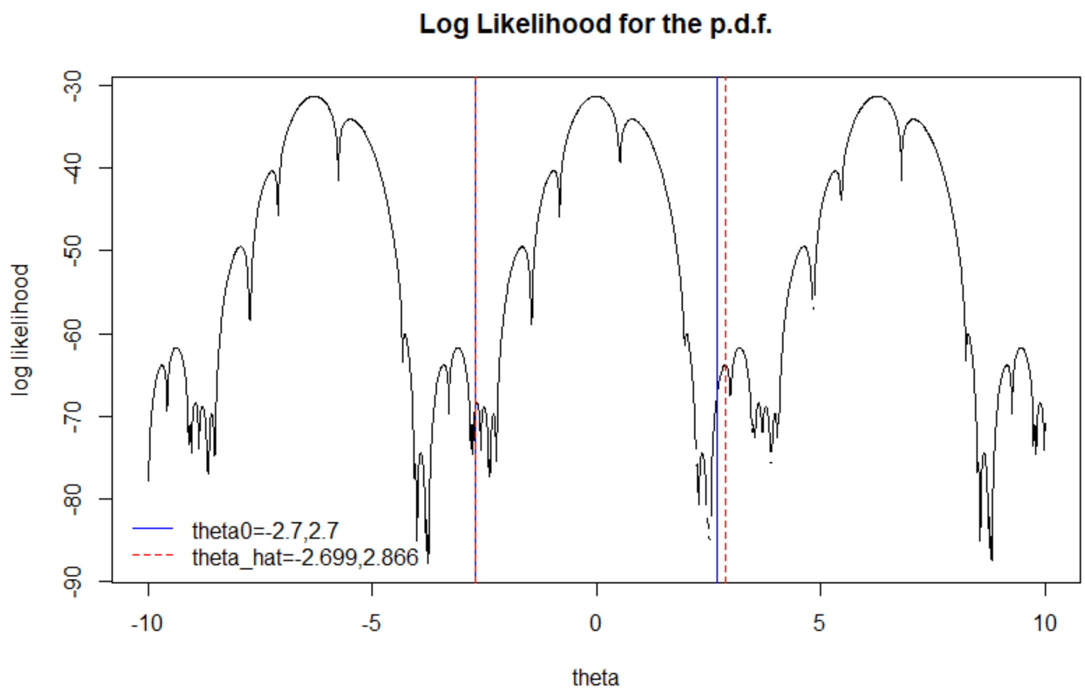
c) $l'(\theta) = \sum_{i=1}^{n}((sin(x_i - \theta)(-1))/(1 - cos(x_i - \theta)))$
$= -\sum_{i=1}^{n}(sin(x_i - \theta))/(1 - cos(x_i - \theta))$
$l''(\theta) = (-\sum_{i=1}^{n})((-cos(x_i - \theta)(1 - cos(x_i - \theta)) + sin(x_i - \theta)(sin(x_i - \theta))/(1 - cos(x_i - \theta))^2$
$= (-\sum_{i=1}^{n})((-cos(x_i - \theta)(cos^2(x_i - \theta)) + sin(^2(x_i - \theta)))/(1 - cos(x_i - \theta))^2$
$= (\sum_{i=1}^{n})((cos^2(x_i - \theta) + sin^2(x_i - \theta))cos(x_i - \theta))/(1 - cos(x_i - \theta))^2$
$= (\sum_{i=1}^{n})cos(x_i - \theta)/(1 - cos(x_i - \theta))^2$

The Newton-Raphson algorithm setting the $\theta_0 = \hat{\theta}_{moment}$ returns a value of -0.01190088.



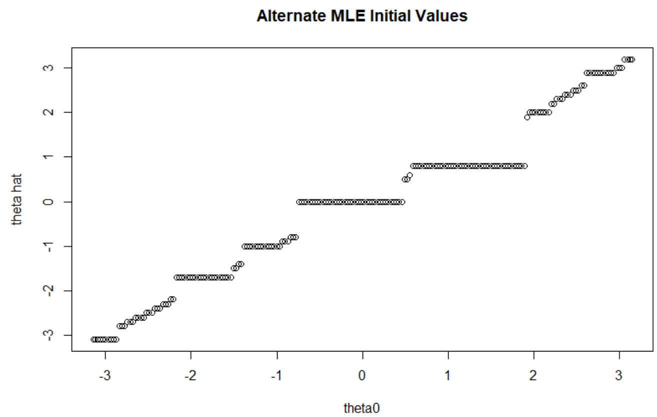
It is clear from the above chart that when $\theta_0 = \hat{\theta}_{moment}$, the estimator is capable of maximizing the function.

d) After setting $\theta_0 = -2.7$ and 2.7, the values that the Newton-Raphson algorithm returned are -2.699948 for -2.7 and 2.865892 for 2.7.



From the above graph, it is clear that the estimators are able to locate the local maximum for each point. On the left side the $\hat{\theta}$ is found right away, and on the right side the local maximum is found quickly.

e) Here the range from $-\pi$ to π is split into 200 equal parts, where the MLE is determined based on some range between the two points on the number line.



In this plot, it is possible to see the different local maximums that can be achieved within the range of $-\pi$ and π. There are two local maximums which stand out as being more dense than the others. It would also be possible to determine that there are certain cutoffs from one point along the line with another point that is further to the right.

Problem 3.

a) From the given information in the problem it is possible to determine values for $\hat{\theta}_1$ and $\hat{\theta}_2$.
$\beta_0 = 1/\theta_1$ and $\beta_1 = \theta_2/\theta_1$
$\hat{\theta}_1 = 1/\beta_0$
$\hat{\theta}_2 = \beta_1\hat{\theta}_1$

After plugging in the given data from the problem and utilizing least squares, the estimates for θ_1 is 195.802 and for θ_2 it is 0.04840226.

b) The function given in the problem can be seen as $g(\theta_1, \theta_2) = \sum_{i=1}^{n}(y_i - (\theta_1 x_i)/(x_i - \theta_2))^2$
The first partial derivative is $dg/d\theta_1 = \sum_{i=1}^{n}[-2x_i(y_i - (\theta_1 x_i)/(x_i + \theta_2))]$
$= -2\sum_{i=1}^{n} x_i y_i + 2\sum_{i=1}^{n}(\theta_1 x_i^2)/(x_i + \theta_2)$
The next partial derivative is $dg/d\theta_2 = \sum_{i=1}^{n}[2(y_i - (\theta_1 x_i)/(x_i + \theta_2)) * (\theta_1 x_i)/(x_i + \theta_2)^2]$
$= \sum_{i=1}^{n}(2y_i\theta_1 x_i)/(x_i + \theta_2)^2 - \sum_{i=1}^{n}(2\theta_1^2 x_i^2)/(x_i + \theta_2)^2$
A partial derivative taken a second time is $d^2g/d^2\theta_1 = 2\sum_{i=1}^{n}(x_i^2)/(x_i + \theta_2)$

The other partial derivative taken a second time is $d^2g/d^2\theta_2 = -2\sum_{i=1}^{n}(2y_i\theta_1x_i)/(x_i+\theta_2)^3 + 6\sum_{i=1}^{n}(\theta_1^2x_i^2)/(x_i+\theta_2)^4$
The last second derivative is $d^2g/d\theta_1d\theta_2 = -2\sum_{i=1}^{n}(\theta_1x_i^2)/(x_i+\theta_2)^2$
The last function is a 2x2 Hessian matrix, $g''(\theta_1, \theta_2) =$

$$\begin{bmatrix} d^2g/d\theta_1^2 & d^2g/d\theta_1d\theta_2 \\ d^2g/d\theta_1d\theta_2 & d^2g/d\theta_2^2 \end{bmatrix}$$

After using the Newton-Ralphson algorithm for predicting θ_1 and θ_2, the results are 215.56 for θ_1 and 0.0669 for θ_2.

c) The Steepest Ascent Method utilizes $M_t = -\alpha_t^{-1}I_p$ where I_p is the Identity matrix.
The equation for the method is: $\vec{x}_{t+1} = \vec{x}_t + \alpha_t g'(\vec{x}_t)$

Using the Steepest Ascent algorithm, the estimate for θ_1 is 195.80 and the estimate for θ_2 is 0.0596.

d) We want to minimize $\sum(y_i - (\theta_1x_i)/(x_i+\theta_2))^2$
Let $f_i(\vec{\theta}) = (\theta_1x_i)/(x_i+\theta_2)$
$f_i'(\vec{\theta}) =$

$$\begin{bmatrix} (x_i)/(x_i+\theta_2) \\ -(\theta_1x_i)/(x_i+\theta_2)^2 \end{bmatrix}$$

$\vec{A} = \vec{A}(\vec{\theta}) =$

$$\begin{bmatrix} f_1'(\vec{\theta})^T \\ \vdots \\ f_n'(\vec{\theta})^T \end{bmatrix}$$

$\vec{Z} = \vec{Z}(\vec{\theta}) =$

$$\begin{bmatrix} y_1 - f_1(\vec{\theta}) \\ \vdots \\ y_n - f_n(\vec{\theta}) \end{bmatrix}$$

The updating formula for this method is:
$\vec{\theta}_{t+1} = \vec{\theta}_t + (\vec{A}_t^T\vec{A}_t)^{-1}\vec{A}_t^T\vec{Z}_t$
where $\vec{A}_t = \vec{A}(\vec{A}_t)$, and $\vec{Z}_t = \vec{Z}(\vec{\theta}_t)$

Using the Gauss-Newton method, the estimate for θ_1 is 212.68 and the estimate for θ_2 is 0.0641.