Hw1

Name - Saurabh Maheshwari

ID - 915481507

Contribution - 100%

This document was created using R Markdown

Answer 1) *R Script file named - hw1_1.R*

part a, b) refer the appended page

part c) refer to the attached figure

part d) After running newton raphson, it is found that different starting points lead to different MLEs. The stopping criteria used here is as follows:

```
  i) absolute difference between successive theta < 1e-6

 ii) gradient is less than 1e-12
```

Refer to the attached figure for trend in theta w.r.t each starting point

Below are the MLEs for corresponding starting points

```
##           -11            -1             0           1.4           4.1
## -5.510873e+13 -1.922866e-01 -1.922866e-01  1.713587e+00  2.817472e+00
##           4.8             7             8            38
## -5.191959e+13  4.104085e+01 -7.685208e+13  4.279538e+01
```

part e) Fisher scoring followed by tuning with newton raphson was done for each starting point. Refer to the figure for trend in theta after each iterations.

Below are the MLEs for corresponding starting points. The MLEs in this case converge to the same value for most of the starting points than as observed by only applying newton raphson. Hence, fisher scoring seems to be more robust with better convergence power.

```
##           -11            -1             0           1.4           4.1
## -1.183545e+11 -1.763491e+00  2.380311e-01  1.521033e+00  7.330081e+00
##           4.8             7             8            38
##  5.625336e+01  1.084586e+01  4.279538e+01  3.819039e+01
```

Answer 2) *R Script file named - hw1_2.R*

part a) refer to the figure attached

part b) Theta_MLE = 0.05844. Refer to the appended page for derivation part c, d) Using Newton-Raphson following thetas were obtained. The figure showing the trend of each theta is attached

```
## 0.0584406061404241                -2.7                  2.7
##          -0.011972          -2.666700            2.873095
```

part e) The figure showing different groups having different thetas is attached. Following are the values demarkating end value for each interval and starting value for the next interval. The starting value for first interval was -3.142 (-pi)

```
##  [1] -2.826 -2.763 -2.605 -2.415 -2.384 -2.258 -2.226 -1.468 -1.437 -0.837
## [11]  0.489  1.942  2.194  2.258  2.447  2.479  2.510  2.984  3.142
```

Following are the number of points corresponding to a particular MLE

```
## -3.093 -2.786 -2.667 -2.508 -2.388 -2.297 -2.232 -1.658 -1.447 -0.953
##     11      2      5      6      1      4      1     24      1     19
## -0.012  0.791  2.004  2.236  2.361  2.475  2.514  2.873   3.19
##     42     46      8      2      6      1      1     15      5
```

Answer 3) *R Script file named - hw1_3.R*

part a) Estimated value of theta1 and theta2 using least squares is as follows

```
## [1] "theta1 =  195.80270884775 ;  theta2 =  0.0484065338672541"
```

part b, c, d) implemented codes are provided in the R script files. Newton Raphson method and gauss newton are faster to reach optima as compared to steepest descent method. The trends for each convergence is depicted in the attached figure. Also, the step size for steepest descent and gauss-netwon was chosen after hit and trial.

thetas after newton raphson (took 8 iterations to converge) were found out to be:

```
## [1] "theta1_nr =  212.683743142536 ;  theta2_nr =  0.0641212816815671"

## [1] "objective value =  1195.44881443936"
```

thetas after steepest descent (20000 iterations) were found out to be:

```
## [1] "theta1_sd =  196.111707915376 ;  theta2_sd =  0.0496111837305381"

## [1] "objective value =  1888.23743633296"
```

thetas after Gauss-Newton method (took 148 iterations to converge) were found out to be:

```
## [1] "theta1_gn =  212.683730744393 ;  theta2_gn =  0.0641212645659096"

## [1] "objective value =  1195.44881443983"
```

Ans. 1)

a) $P(x-\theta) = \dfrac{1}{\pi\{1+(x-\theta)^2\}}$

$$l(\theta) = \log \prod_{i=1}^{n} \frac{1}{\pi\{1+(x-\theta)^2\}}$$

$$= \sum_{i=1}^{n} \log \frac{1}{\pi\{1+(x-\theta)^2\}}$$

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

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

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

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

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

b) $I(\theta) = E\left[(l'\theta)^2\right]$

$$= E\left(\frac{2(x-\theta)}{1+(x-\theta)^2}\right) = \int_{-\infty}^{\infty} \left(\frac{2(x-\theta)}{1+(x-\theta)^2}\right)^2 \cdot \frac{1}{\pi\{1+(x-\theta)^2\}}\, dx$$

$$= \frac{4}{\pi} \int_{-\infty}^{\infty} \frac{(x-\theta)^2}{(1+(x-\theta)^2)^3} dx$$

let $t = x - \theta \Rightarrow dt = dx$

$$\Rightarrow I_X(\theta) = \frac{4}{\pi} \int_{-\infty}^{\infty} \frac{t^2}{(1+t^2)^3} dt$$

let $u = \frac{\cancel{t^2}1}{(1+t^2)} \Rightarrow t = (1/u - 1)^{1/2} \Rightarrow dt = \frac{1}{2}(1/u-1)^{-1/2} \cdot (-\frac{1}{u^2}) du$

$$\Rightarrow I_X(\theta) = \frac{8}{\pi} \int_{0}^{\infty} \frac{t^2}{(1+t^2)^3} dt$$

$$= \frac{8}{\pi} \int_0^1 \cancel{u}^3 \cdot \frac{(1-u)}{\cancel{u}} \cdot \frac{1}{2} \left(\frac{1-u}{u}\right)^{-1/2} \cdot \frac{1}{\cancel{u^2}} du$$

$$= \frac{8}{2\pi} \int_0^1 (1-u)^{1/2} \cdot u^{1/2} du$$

$$= \frac{4}{\pi} \int_0^1 u^{3/2-1} \cdot (1-u)^{3/2-1} du \quad (\because \text{Beta Integral})$$

$$= \frac{4}{\pi} \frac{\Gamma(3/2)^2}{\Gamma(3/2+3/2)} = \frac{4}{\pi} \frac{(0.5\sqrt{\pi})^2}{2!} = 1/2$$

$$\Rightarrow I_{\underline{X}}(\theta) = \sum_{i=1}^{n} 1/2 = n/2$$

Ans. 2)

$$p(x) = \frac{1-\cos(x-\theta)}{2\pi}$$

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

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

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

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

$$= \mp \sum_{i=1}^{n} \frac{\cancel{}\cos(x_i-\theta)}{(1-\cos(x_i-\theta))^2} = -\sum_{i=1}^{n} \frac{1}{1-\cos(x_i-\theta)}$$

(b) $$E(x) = \int_0^{2\pi} x \cdot p(x)\, dx = \int_0^{2\pi} \frac{x\,(1-\cos(x-\theta))}{2\pi}\, dx$$

$$= \pi - \frac{1}{2\pi}\int_0^{2\pi} x\cos(x-\theta)\, dx$$

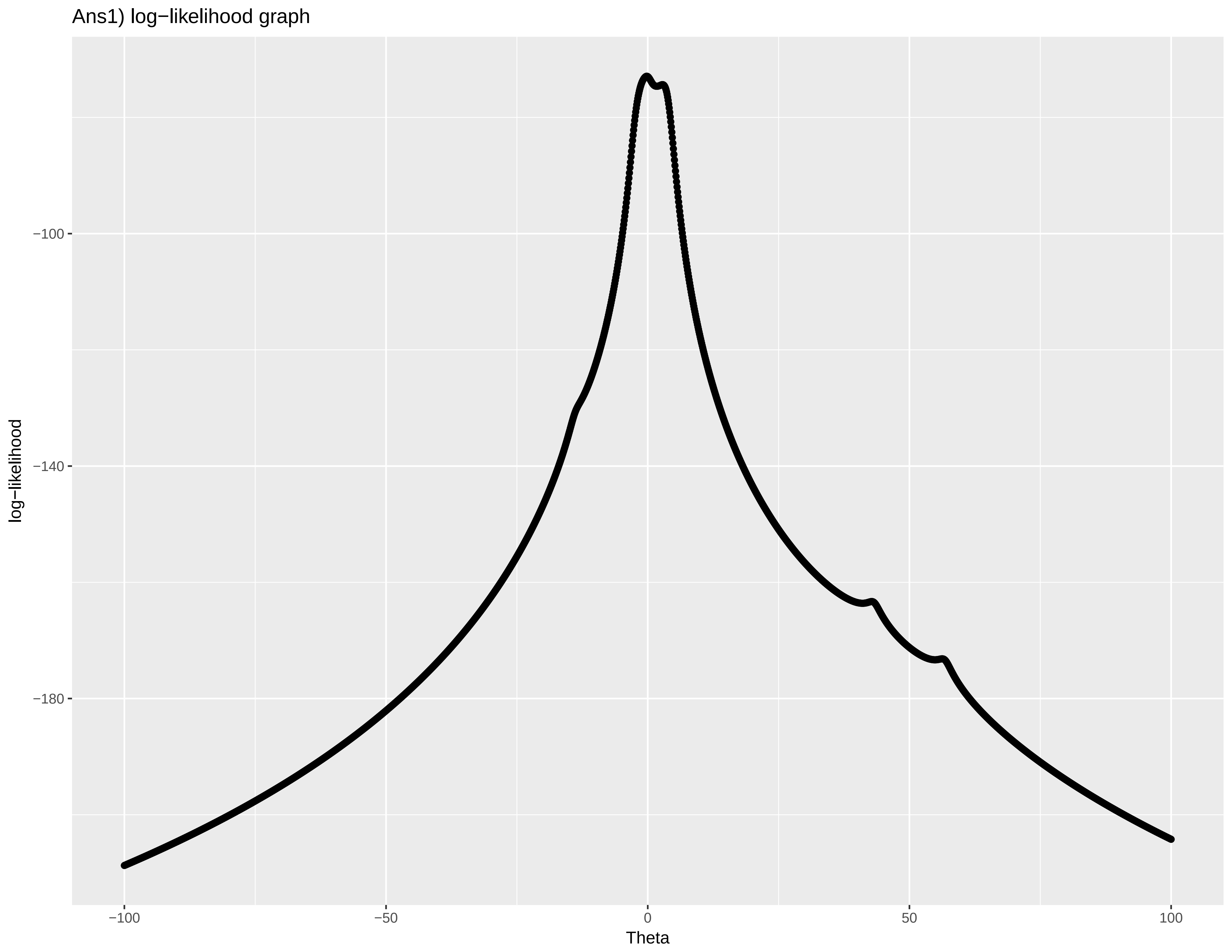
$$= \pi - \frac{1}{2\pi}\Big[x\sin(x-\theta) - \cos(x-\theta)\Big]\Big|_0^{2\pi}$$

$$= \pi - \frac{1}{2\pi}\left[-2\pi\sin\theta - \cancel{\cos\theta} + \cancel{\cos\theta}\right]$$

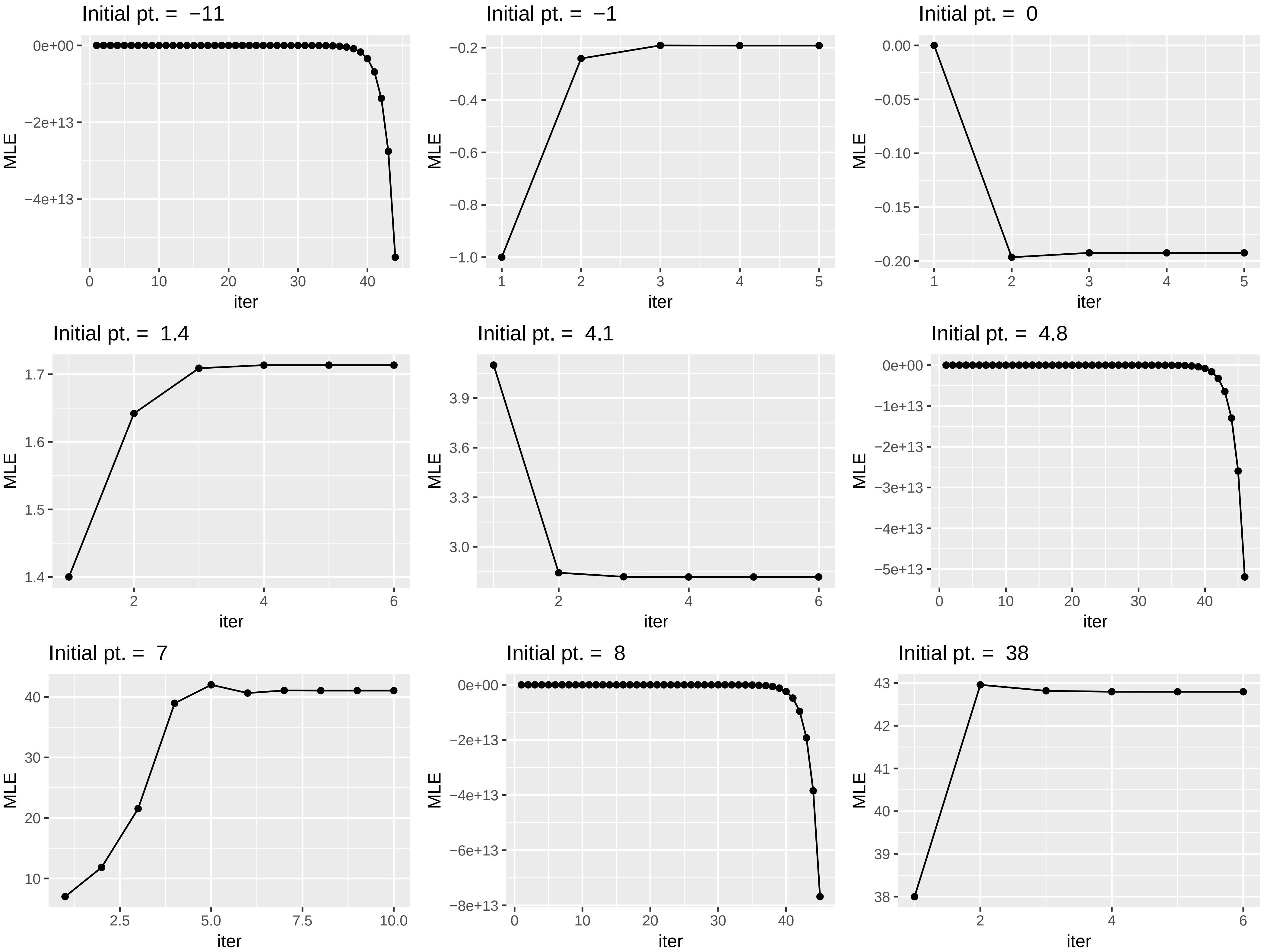
$$= \pi + \sin\theta$$

$\Rightarrow$ $E(x) = \frac{1}{n}\sum x_i \Rightarrow \pi + \sin\theta = 3.2$

$\Rightarrow$ ~~θ_{MLE}~~ $\theta_{moment} = 0.05844$

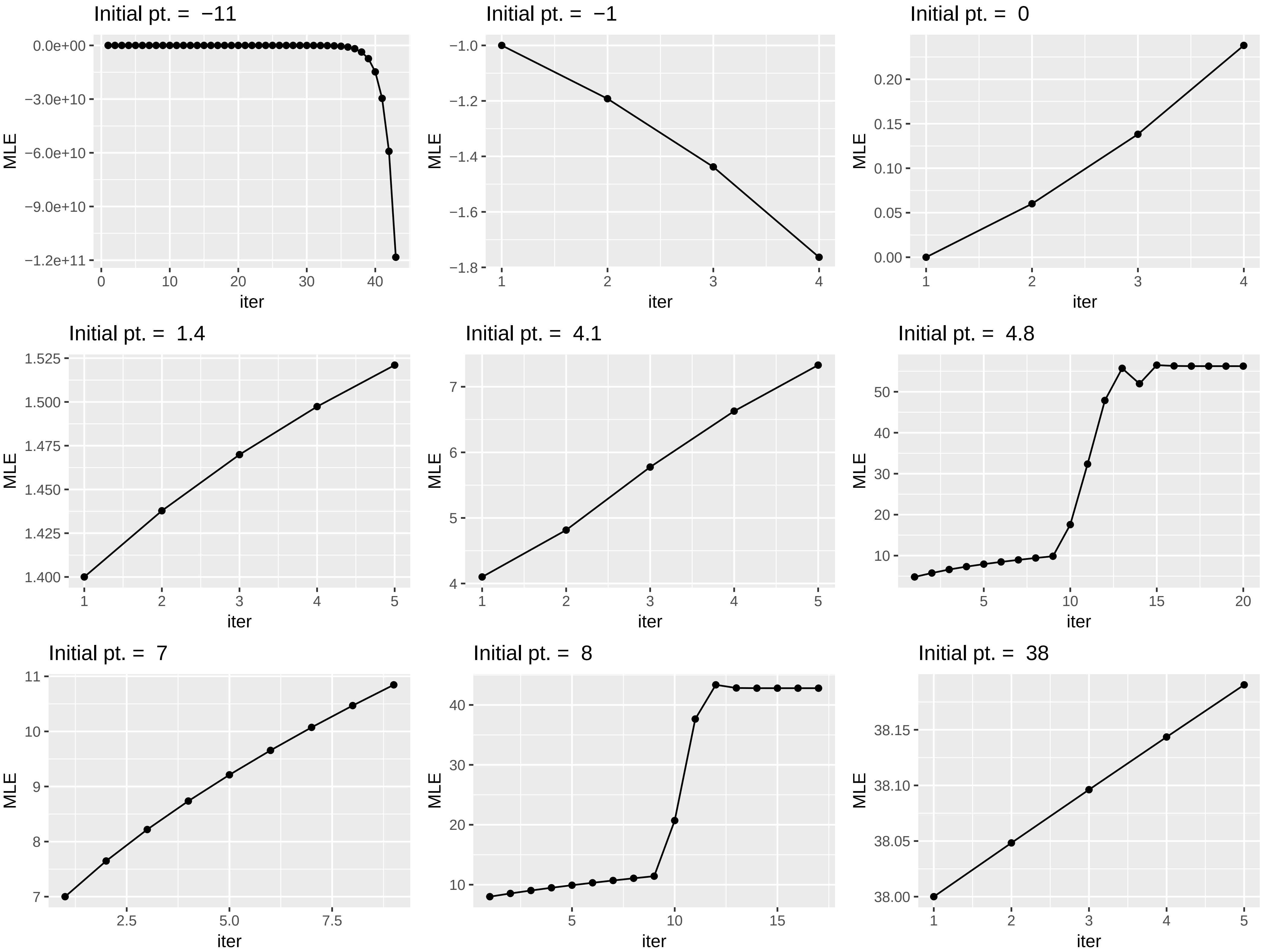

Ans1) log−likelihood graph
log−likelihood
−100
−140
−180
−100
−50
0
50
100
Theta

Ans 1) Part c

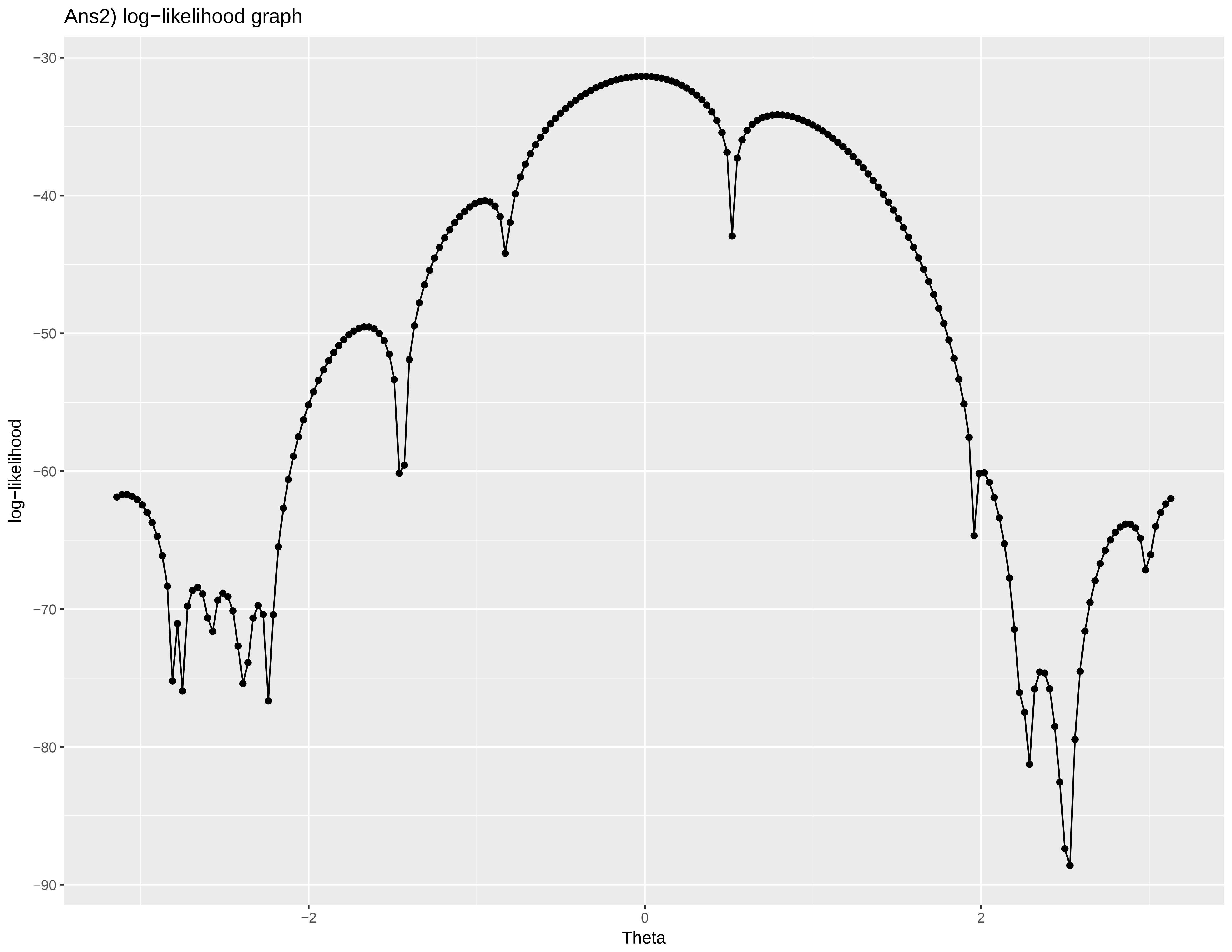


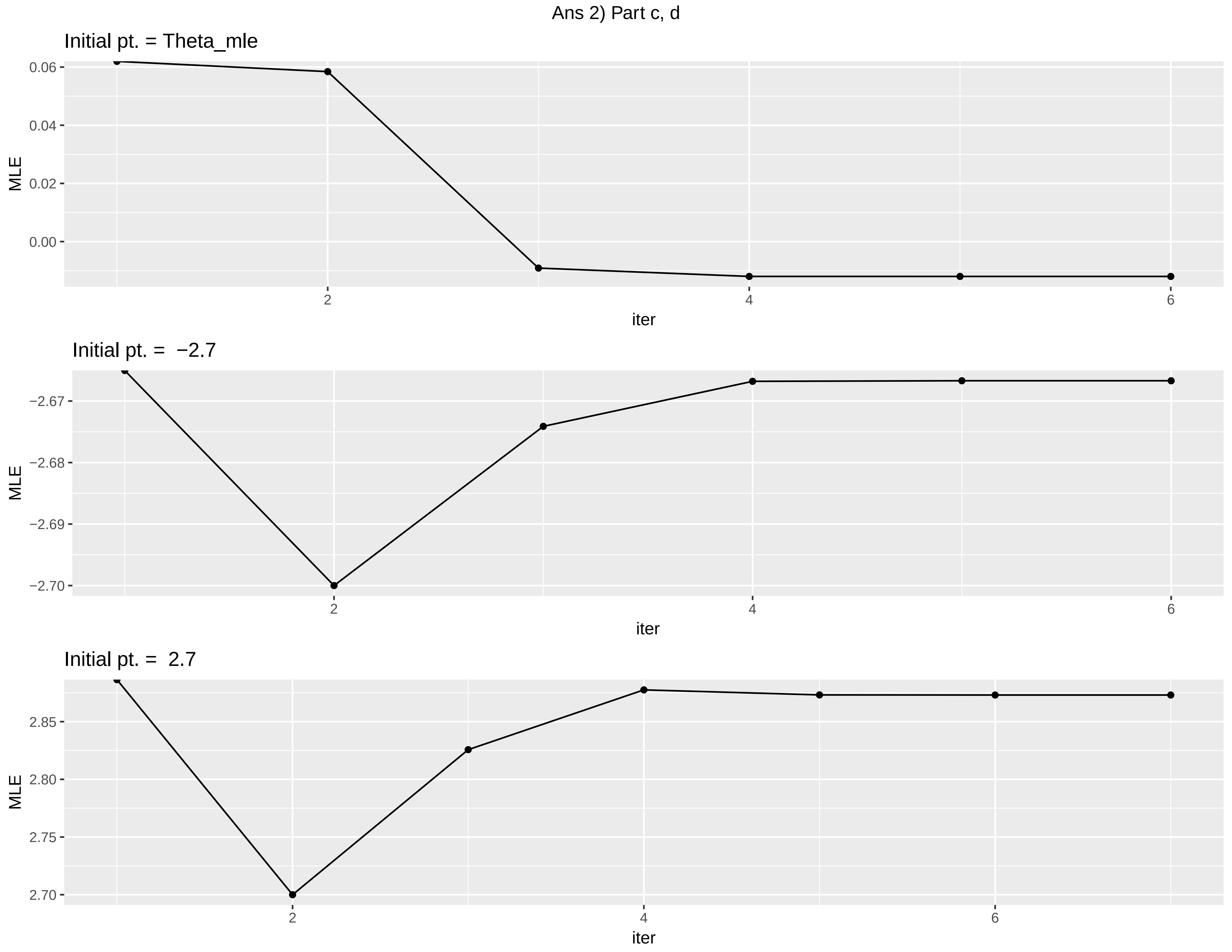
Initial pt. = −11
Initial pt. = −1
Initial pt. = 0
Initial pt. = 1.4
Initial pt. = 4.1
Initial pt. = 4.8
Initial pt. = 7
Initial pt. = 8
Initial pt. = 38
MLE
iter

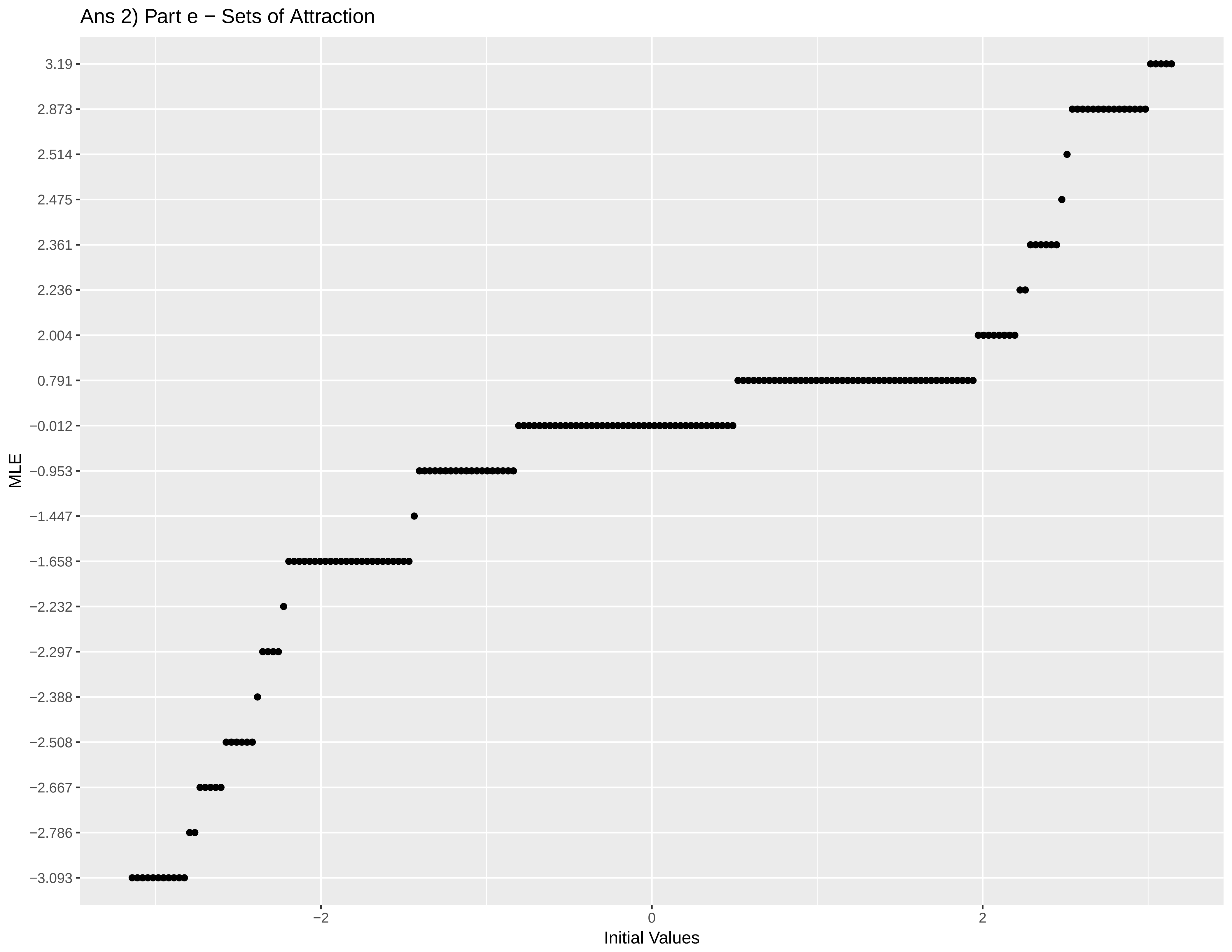
Ans 1) Part d



Initial pt. = −11
Initial pt. = −1
Initial pt. = 0
Initial pt. = 1.4
Initial pt. = 4.1
Initial pt. = 4.8
Initial pt. = 7
Initial pt. = 8
Initial pt. = 38
MLE
iter


Ans2) log−likelihood graph
−30
−40
−50
−60
−70
−80
−90
log−likelihood
−2
0
2
Theta


Ans 2) Part c, d
Initial pt. = Theta_mle
0.06
0.04
0.02
0.00
MLE
2
4
6
iter
Initial pt. = −2.7
−2.67
−2.68
−2.69
−2.70
MLE
2
4
6
iter
Initial pt. = 2.7
2.85
2.80
2.75
2.70
MLE
2
4
6
iter


Ans 2) Part e − Sets of Attraction
MLE
3.19
2.873
2.514
2.475
2.361
2.236
2.004
0.791
−0.012
−0.953
−1.447
−1.658
−2.232
−2.297
−2.388
−2.508
−2.667
−2.786
−3.093
−2
0
2
Initial Values

Ans 3) Part b, c, d