3b

Yiqiao Yin

1/31/2021

Normal Approximation

Normal approximation: suppose that $y_1, ..., y_5$ are independent samples from a Cauchy distribution with unknown center θ and known scale 1: $p(y_i|\theta) \approx 1/(1 = (y_i - \theta)^2)$.

Assume that the prior distribution for θ is uniform on $[0, 1]$. Given the observations $(y_1, ..., y_5) = (-2, -1, 0, 1.5, 2.5)$

(a) Determine the derivative (1st order and 2nd order)

According to the problem here we have running index i to range in $i = 1, 2, ..., 5$. In other words, we can find the log-posterior densityh

$$\log p(\theta|y_i) = -5 \log \pi - \sum_{i=1}^{5} \log(1 + (\theta - y_i)^2)$$

for $\theta \in [0, 1]$ and negative infinity if $\theta \notin [0, 1]$.

Next, we can find the 1st and the 2nd order derivative to be the following:

$$\frac{\partial \log p(\theta|y_i)}{\partial \theta} = -2 \sum_{i=1}^{5} \frac{\theta - y_i}{1 + (\theta - y_i)^2}$$

$$\frac{\partial^2 \log p(\theta|y_i)}{\partial \theta^2} = 2 \sum_{i=1}^{5} \frac{(\theta - y_i)^2 - 1}{[1 + (\theta - y_i)^2]^2}$$

and both the above results are defined on $\theta \in [0, 1]$

(b) Find the posterior mode of θ

To solve the optimal solution, we set the 1st order derivative to 0 and solve for the $\theta_{\max}$. In other words, we write the following

$$\frac{\partial \log L(y_i|\theta)}{\partial}\bigg|_{\theta=\theta_{\max}} = 0$$

and this is achieved at

$$\sum_{i=1}^{5} \frac{\theta_{\max} - y_i}{1 + (\theta_{\max} - y_i)^2} = 0$$

This can be achieved by Newton's method. Let us recall Newton's method: Consider real variable x, we have function $f(x)$ and its first order derivative of $f'(x)$. In this case, an inital guess x_0 can be what we use to start the algorithm. Then each step k, we updates the variable x by the following:

$$x_k = x_{k-1} - \frac{f(x)}{f'(x)}$$

To adapt this algorithm in this method, let us rewrite using θ and log-posterior. In this case, let us suppose x to be θ and $f(x)$ to be the first order derivative of log-posterior. Hence, let us write

$$\theta_k = \theta_{k=1} - \frac{\frac{\partial}{\partial\theta}\left(\log L(y_i|\theta)\right)}{\frac{\partial^2}{\partial\theta^2}\left(\log L(y_i|\theta)\right)}$$

```r
theta = 1
totalN = 100
f = function(x) {
  y = c(-2, -1, 0, 1.5, 2)
  value1 = 0
  value2 = 0
  for (i in 1:length(y)) {
    value1 = value1 + (x-y[i])/(1+(x-y[i])^2)
    value2 = value2 - ((x-y[i])^2-1)/(1+(x-y[i])^2)^2
  }

  return(c(value1,value2))
} # End of function

eta = 0.01
pb = progress::progress_bar$new(total = totalN)
thetaList = matrix(0L, 1, totalN)
secondOrderLogPosterior = c()
for (i in 1:totalN) {
  v = f(theta)
  secondOrderLogPosterior = c(secondOrderLogPosterior, v[2])
  theta = theta-eta*v[1]/v[2]
  thetaList[i] = theta
  pb$tick()
} # End of loop

print(theta)
```

```
## [1] 0.1167713
```

(c) Construct the normal approximation

```r
# sampling distribution
sampleDistribution = function(y, theta) {
  d0 = NULL
  for (i in 1:length(theta)) { d0 = c(d0, prod(dcauchy(y, theta[i]))) }
  return(d0)
} # finished function definition
```

```r
# data
y = c(-2, -1, 0, 1.5, 2)
step = 0.01
theta = seq(0, 1, step)[-1]

# unnormalized f
unnormalizedDistribution = sampleDistribution(y, theta)

# normalized f
normalizedDistribution = unnormalizedDistribution / (step*sum(unnormalizedDistribution))

# approximated
approximatedAnnormalizedDistribution = sampleDistribution(
  y, secondOrderLogPosterior)
normalizedApproxDist = approximatedAnnormalizedDistribution /
  (step*sum(approximatedAnnormalizedDistribution))

# plot
plot(1:totalN, normalizedDistribution,
     type = "l", xlab = "theta", ylab = "Distribution", col = "blue",
     main = "Normal Approximation based on 2nd Deriv.
     vs. Exact Normal \n(Blue: Exact, Red: Approx.)")
lines(1:totalN, normalizedApproxDist, col = "red")
```


Normal Approximation based on 2nd Deriv.
vs. Exact Normal
(Blue: Exact, Red: Approx.)