Monte Carlo Optimization Report

Ravish Kamath: 213893664

22 April, 2023

Introduction

In mathematics, optimization is seen as finding the extremas of a function $h(\theta)$ over a domain Θ. This is seen often in calculus courses, where we learn to solve $\frac{\partial h(\theta)}{\partial \theta} = 0$ (assuming the function h is differentiable). In statistics, we are usually focused on finding the maximum (such as maximum likelihood), hence we will focus on maximization problems: $\arg\max_{\theta \in \Theta} h(\theta)$. Similarly to integration using Monte Carlo methods, we can use the same approach to calculate very complicated h functions or irregular domain Θ using a rather stochastic approach. In this report we will focus on one numerical method approach (Newton-Raphson) and 3 stochastic methods:

- Basic Uniform Monte Carlo optimization
- Monte Carlo Optimization using g distribution
- Simulated Annealing

We will look at the effectiveness of each method on **5 separate functions**.

Example 1

Goal

Find the arg max on the domain $[0, 2\pi]$:

$$\sin(\sqrt{x}) \tag{1}$$

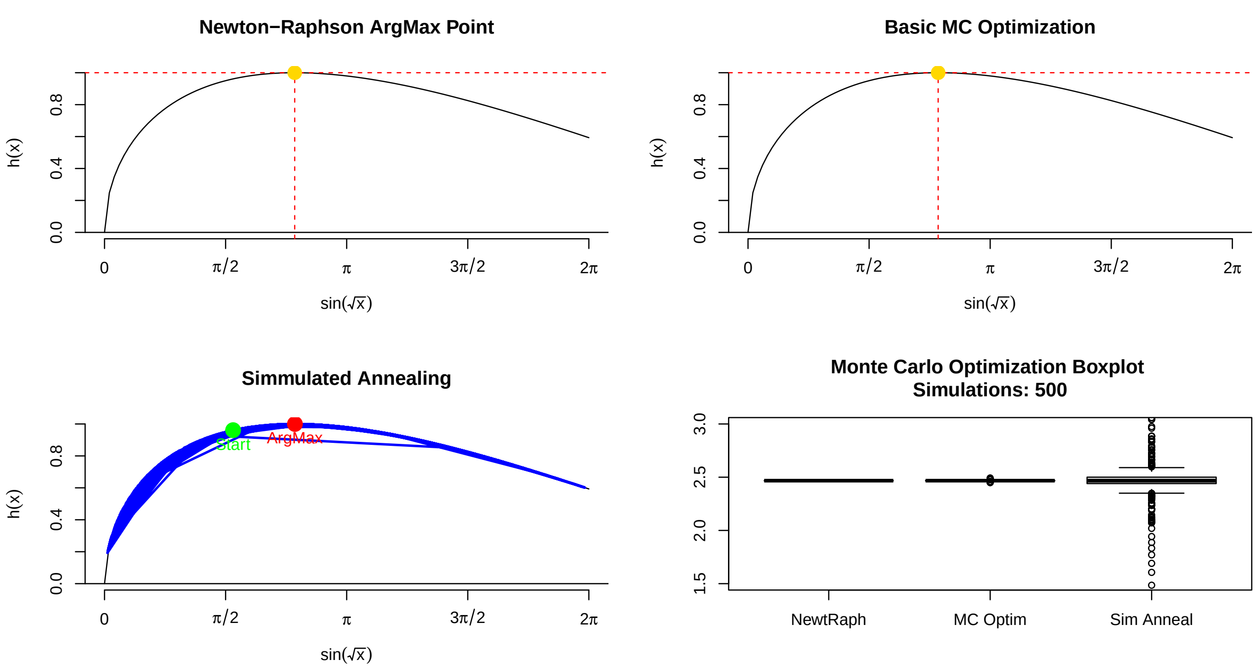


Figure 1: Finding the Arg max for (1)

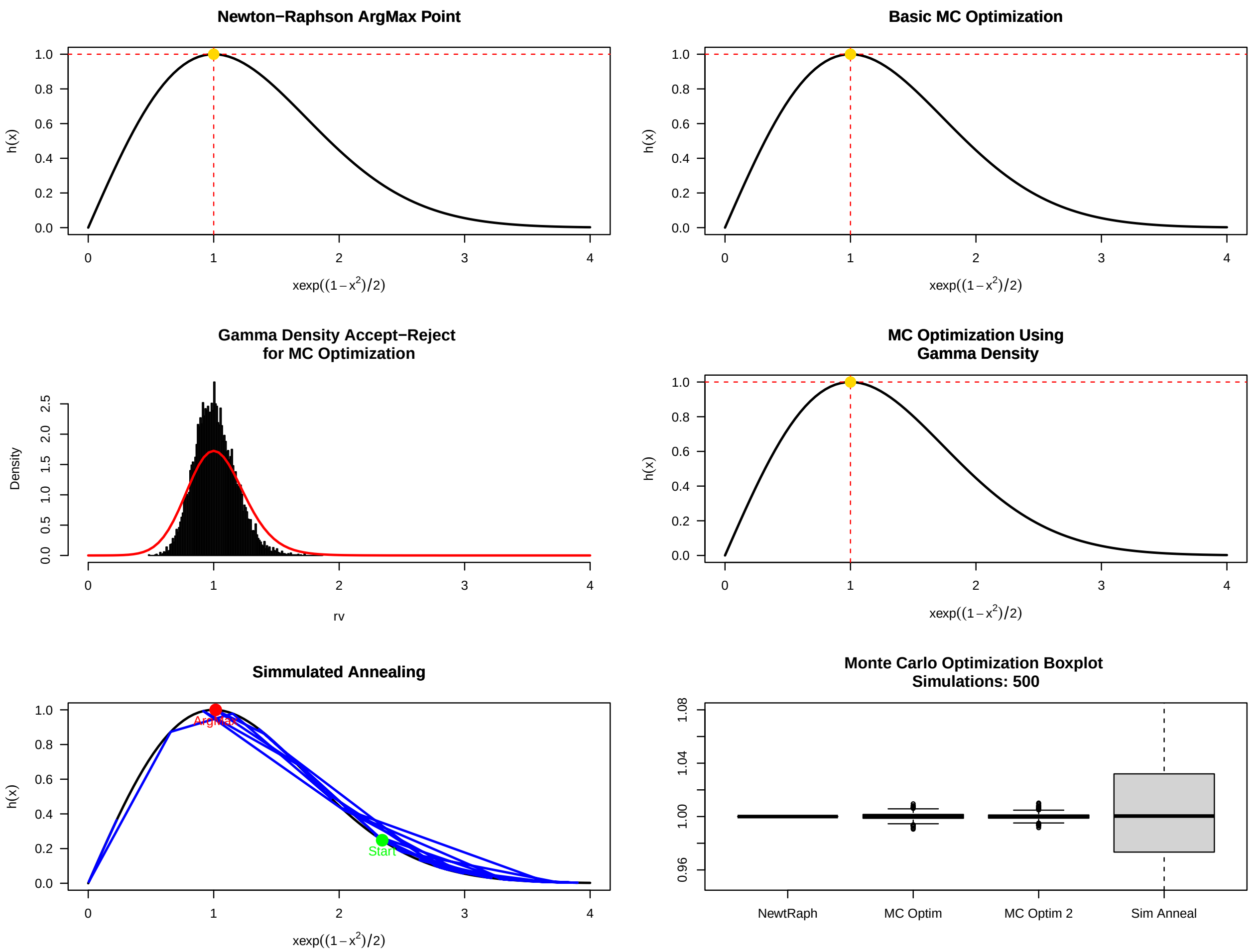
We start with a very basic function to make it easier to understand how each method works. Based on the plots in the previous page, we can see that each method was able to capture the maximum of our given function quite accurately in one simulation. However, looking at the box plot in the bottom right of Figure 1, we can see that the simulated annealing seems to have quite a bit of variance, when locating the arg max. Based on this result, it looks like either the Newton-Raphson method or the basic Monte Carlo Method is the most effective method to solve the arg max. We will now look at another example, where we try the Monte Carlo method, but using a different distribution for our Random Variables.

Example 2

Goal

Find the arg max on the domain $[0, 4]$:

$$x \exp\left\{\frac{(1-x^2)}{2}\right\} \tag{2}$$



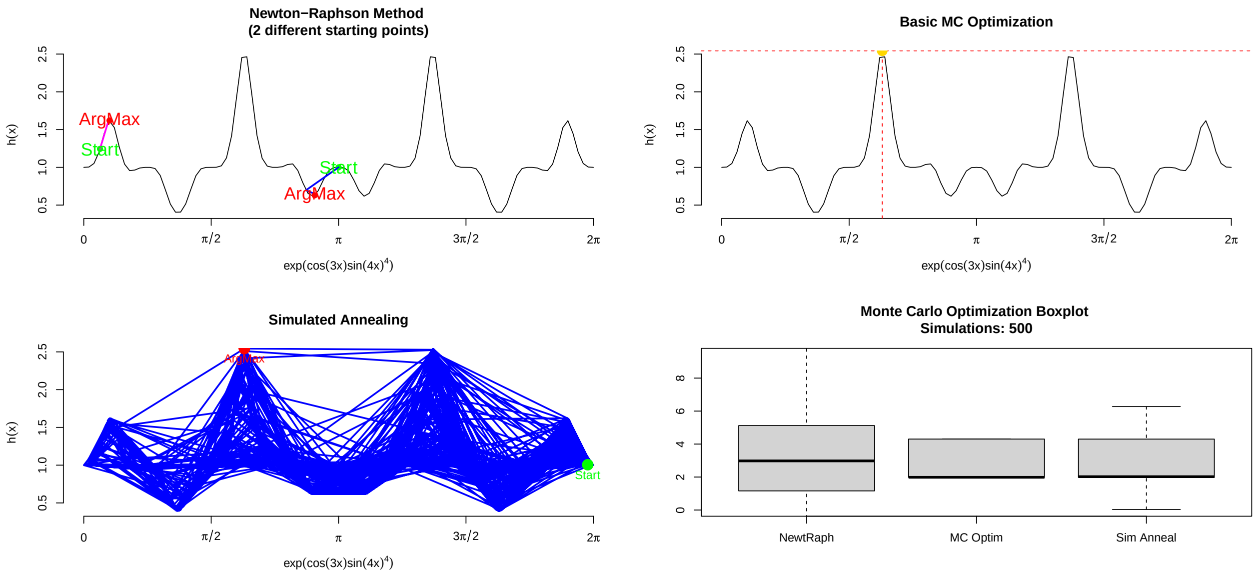
For this function we used all four methods to see how each one performs on finding the arg max. Clearly, we can see that each method is pretty effective at finding the arg max, especially the Newton-Raphson and MC Optimization method. Although, the size of the box-plot does not do justice, by using gamma distribution, it was actually more effective when compared to the basic method. Although these plots indicate that Newton-Raphson and basic MC optimization are the best method, let us now provide more complicated examples, where they are maybe the best choice for finding the arg max.

Example 3

Goal

Find the $\arg\max$ on the domain $[0, 2\pi]$:

$$\exp\{\cos(3x)\sin(4x)^4\} \tag{3}$$



This time we get very interesting results. With the Newton-Raphson method, I tried different points to see how effective it was. Unfortunately, at both instances it was not able to capture either of the arg max. With the basic MC optimization method, it was actually able to capture global maximum, hence it is a better approach compared to the Newton-Raphson. Finally, we can now see the power of the Simulation Annealing. This method was able to move around the function and find the global maximum, rather than any local maximum. Based off the box plot, we can see that the simulation annealing overpowers the other 2 methods.

Example 4

Goal

Find the $\arg\max$ on the domain $[0, 8]$ of the likelihood associated with the mixture model:

$$X \sim \pi\Gamma(\alpha_1, \beta_1) + (1 - \pi)\Gamma(\alpha_2, \beta_2)$$

where $\alpha_1 = 38$, $\alpha_2 = 50$, $\beta_1 = 20$, $\beta_2 = 11$ and $\pi = 1/3$.

Looking at the plots in Figure 2, we can see that the Newton Raphson and the Simulated Annealing do a better job compared to the Basic Monte Carlo Optimization. It is quite surprising that the Newton method was able to find the global maximum in this function, despite have another close local maximum. With the box plot, I decided to try a randomization of each starting point for both the Newton and the Simulated Annealing method. Although both are effective, it seems that Newton is the best method for this function. Furthermore, we can definitely agree that the Basic MC optimization method, is not that effective when it comes to mixture models in general. In the next example, we will focus on comparing the effectiveness of the Newton-Raphson method vs. Simulation Annealing with a mixture model that contains a location parameter. This may affect the results we just saw in the gamma mixture.

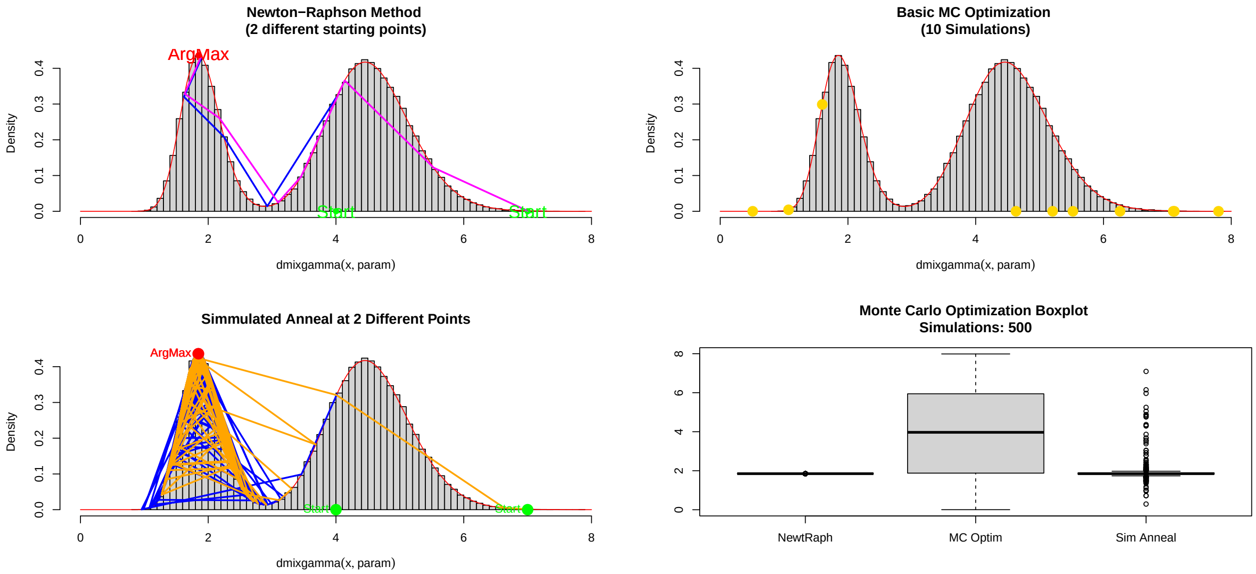


Figure 2: Finding Argmax of Gamma Mixture Distribution

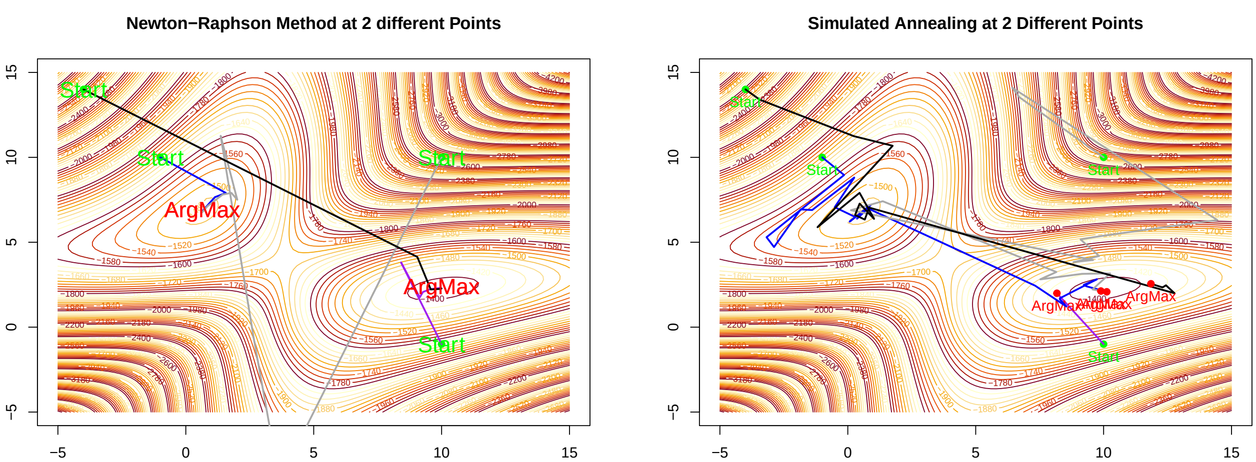
Example 5

Goal

Find the arg max on the domain $[-5, 15]$ of the likelihood associated with the mixture model:

$$X \sim \pi \text{Logis}(\mu_1, s_1) + (1 - \pi)\text{Logis}(\mu_2, s_2)$$

where $\mu_1 = 2, \mu_2 = 5, s_1 = 1, s_2 = 3$.



In this example, we chose 4 random points on the map to see how effective each method are. As we can see in the Newton Raphson-Method, some of the points we chose, seem to believe that at the maximum at [0,6] seems to be the arg max. However, the true maximum is actually situated at around [9,2]. Clearly, the Newton Raphson method is not able to identify the arg max. However, we can see that the Simulated Annealing computes a better job at finding the arg max. Hence we can see that handling bi modal or mixture data, the Simulated Annealing seems to be the better choice out of the options.