CmpE 343: Introduction to Probability and Statistics for Computer Engineers (Fall 2020)

Hazar ÇAKIR
2017400093

Bonus Homework

Due January 10, 2021 by 11:59pm on Moodle

In order to determine the maximum point, I use the second approach which is an **evolutionary approach**. My code finds the maximum point **in generally 5 or sometimes 6 iterations**. For the first sampling I took 25 samples which is a high number. For the next iterations, **my sample number is 15**. I will explain why I choose beginning more than following iterations. Let me explain my code.

Firstly I decide initial values to start the algorithm. My initial variance value is 0.5 and my initial sample number is 25. Then for the first sampling, I use an Uniform Distribution because at the beginning I do not have any assumption about my data. I take samples between -0.75 and 0.75 because when I analyze my input function, those points are the beginning of the curves. In order to not take unnecessary samples out of the box, I choose a smaller set of interval. I keep that value in **sampleSet** and I keep their rewards in **values**. Then I decrease the sample number to 15.

I iterate in a while loop till the range of my sample space becomes smaller than 0.01. Firstly I plot the graph and the samples that I got. Then I control the exit situation. I iterate through **sampleSet** and determine the **range** (the distance between max and min). When the range is smaller than 0.01, I break and exit from the loop. The current **sampleSet** is my final sample set which indicates max value.

Otherwise means I should continue iteration. What I do first is to determine the most successful 5 samples. I decided that from 15 samples, 5 samples is a good ratio. Those 5 samples will be the determinator of the next generation. I iterate through my **sampleSet** and keep 5 samples' indices in a list namely **Is**. This list is in increasing order regarding their rewards.

Determining the next generation's **mean** and **variance** is the most crucial point. In order to make a correct guess, I inspect the **Is** list which are the most successful samples. I inspect the **variance of the Is** specifically, because from that data, I can understand where I am exactly. The boundary between less and more is chosen as **0.012.**

High variance in Is means I took samples from two different peaks. From there I can understand that both can be my maximum point. So in that case, I won't decrease the variance of the distribution too much **(divided by 1.75)** in order to be able to detect false peak. While choosing mean value, I apply several approaches, but what I decide to implement is to take **the maximum value as the mean**. I realized that in most cases, the maximum value is taken from the global maximum. Moreover, in order to decrease the probability to miss the global maximum, I increased the initial number of samples and constrained the interval of sampling. In that way, I can increase the density and determine the maximum global.

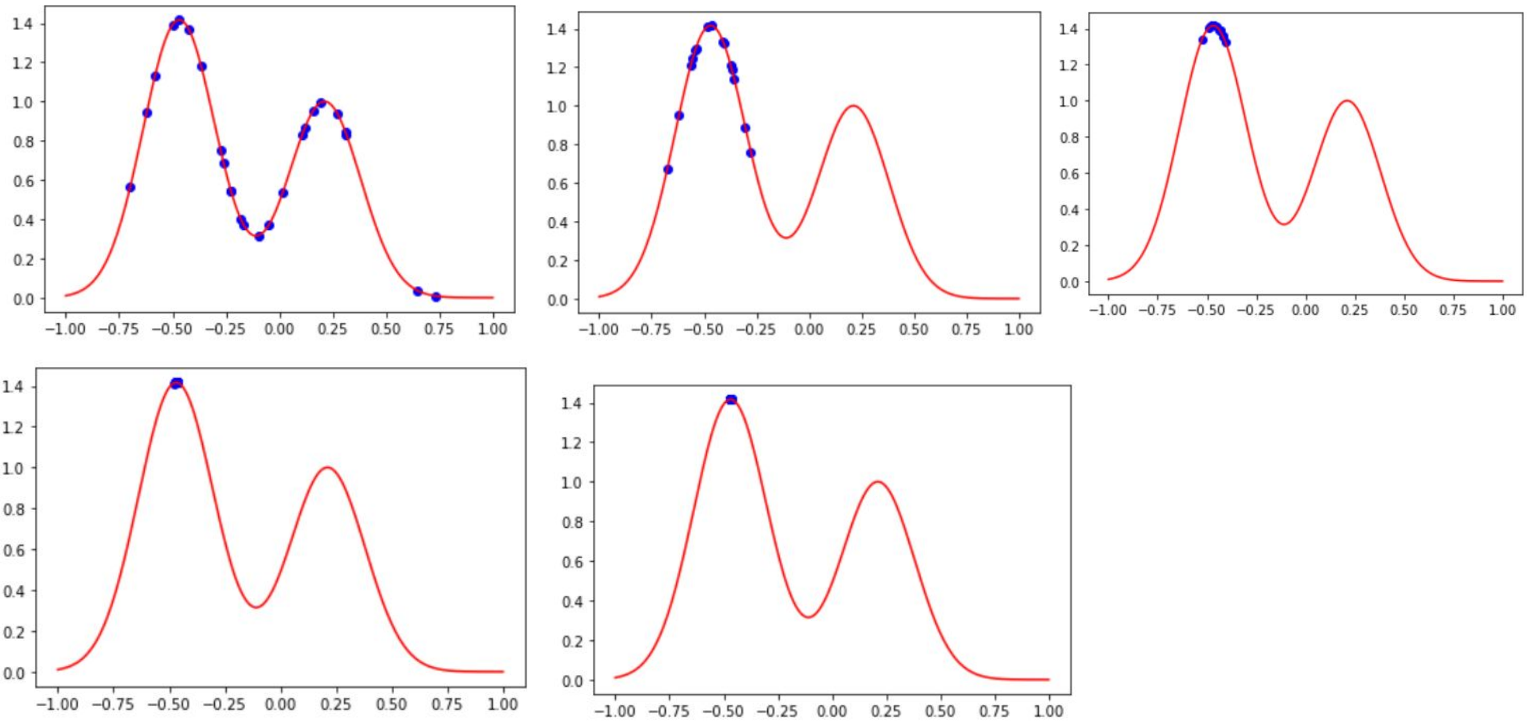
In the case of the **low variance in Is**, I decreased variance of the distribution drastically as **dividing it to 4.2.** This is to decrease the number of the sampling. In this case, all of my samples are close to each other, so I can take **the mean of those samples** in order to get to the max value.

When I implement my algorithm, I determine the constants like 4.2, 0.012 or 1.75 empirically. I played with my data several times to determine optimum values. It works about

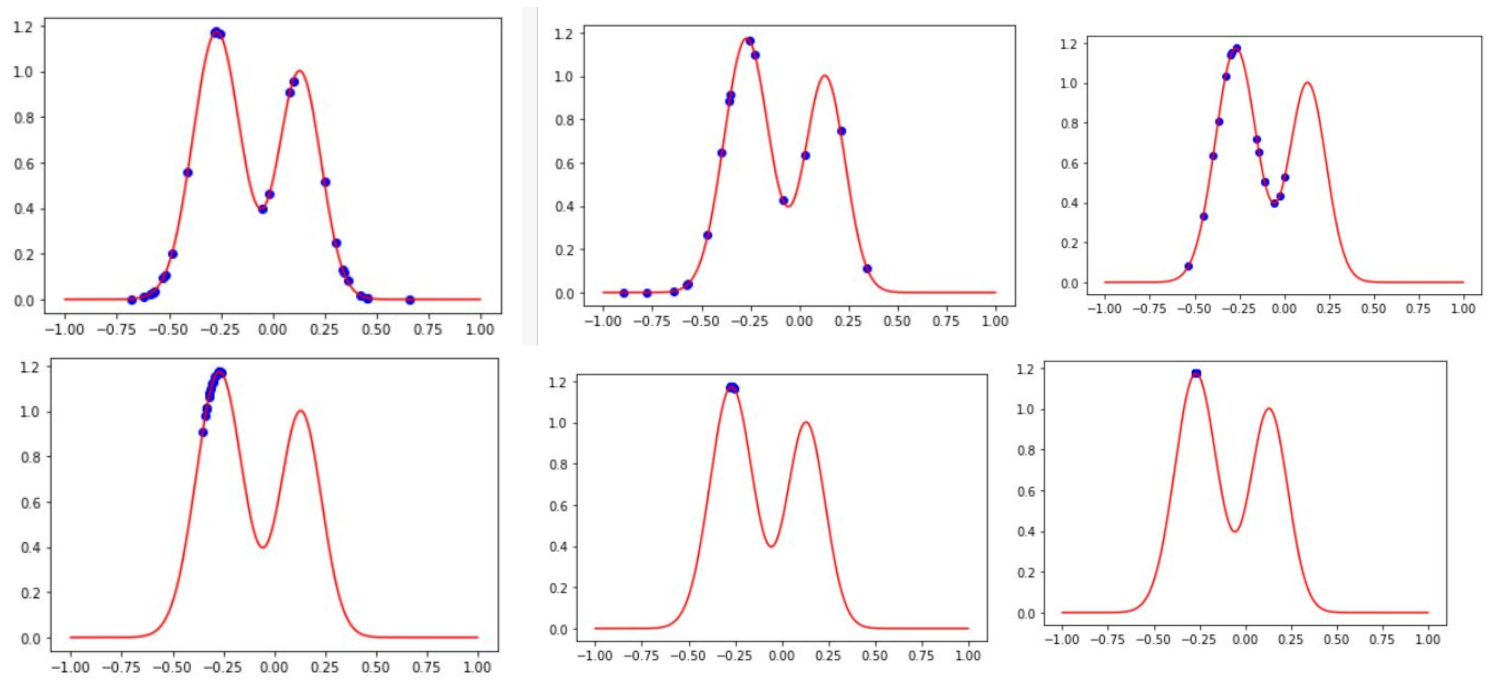
97%, because I take random sampling, when the global maximums peak is too narrow, It can be missed from my algorithm. It resolves in a solution in 5 iterations most of the time and sometimes goes to 6 iterations. The value of the max point is close to about **0.0004** in the worst case. I think it is a successful algorithm.

Examples:

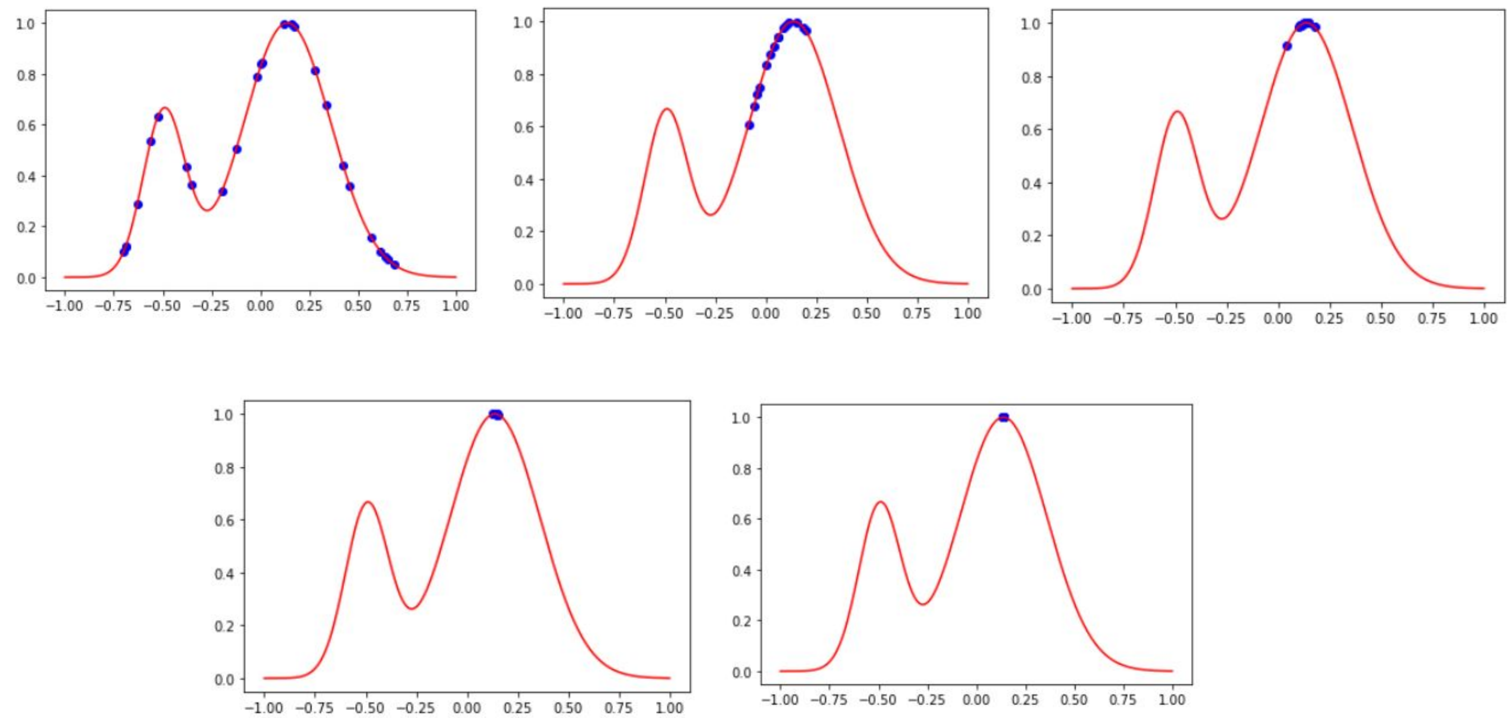
1)



2)

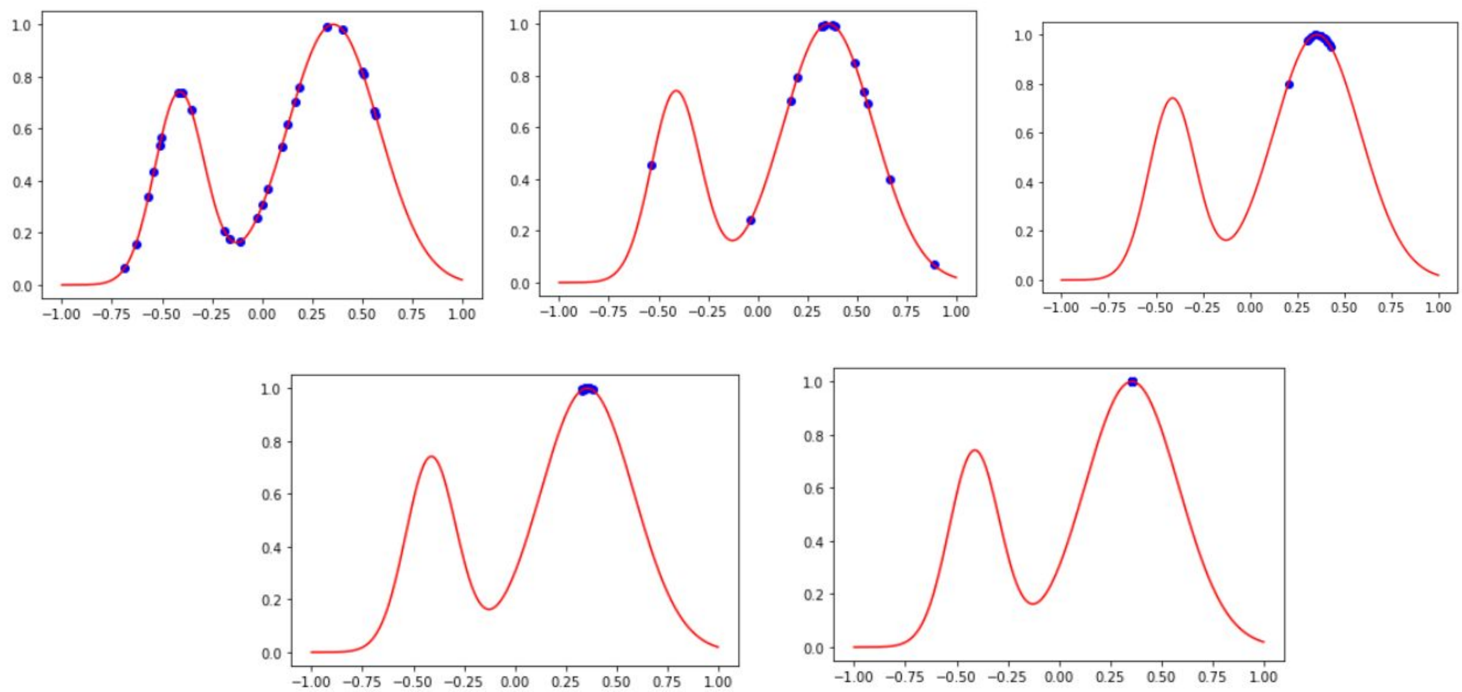


3)



1.0 0.8 0.6 0.4 0.2 0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
1.0 0.8 0.6 0.4 0.2 0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
1.0 0.8 0.6 0.4 0.2 0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
1.0 0.8 0.6 0.4 0.2 0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
1.0 0.8 0.6 0.4 0.2 0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00

4)



1.0 0.8 0.6 0.4 0.2 0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
1.0 0.8 0.6 0.4 0.2 0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
1.0 0.8 0.6 0.4 0.2 0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
1.0 0.8 0.6 0.4 0.2 0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
1.0 0.8 0.6 0.4 0.2 0.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00

5)