CS 215 Assignment 4 Report

Diwan Anuj Jitendra - 170070005 Soumya Chatterjee - 170070010

November 10, 2018

Problem 1

Part (a)

We have to find the probability that $X_1^2 + X_2^2 \leq 1$ given that X_1 and X_2 are iid with Uniform$(-1, -1)$

$$\begin{aligned}
P\left\{X_1^2 + X_2^2 \leq 1\right\} &= \int_x P\left\{X_1 = x, -\sqrt{1-x^2} \leq X_2 \leq \sqrt{1-x^2}\right\} dx \\
&= \int_{-1}^{1} p(X_1 = x) P\left\{-\sqrt{1-x^2} \leq X_2 \leq \sqrt{1-x^2}\right\} dx \\
&= \int_{-1}^{1} \frac{1}{2} \frac{2\sqrt{1-x^2}}{2} dx \\
&= \int_{-1}^{1} \frac{\sqrt{1-x^2}}{2} dx \\
&= \frac{1}{4} x\sqrt{1-x^2} + \frac{1}{4}\sin^{-1} x \bigg|_{-1}^{1} \\
&= \frac{\pi}{4}
\end{aligned}$$

Part (b)

To estimate π, we can generate N samples from a 2-dimensional random vector $X = (X_1, X_2)$ where X_1 and X_2 are drawn from Uniform$(-1, 1)$ distribution. Then find the fraction of samples that lie within the circle $X_1^2 + X_2^2 \leq 1$.

By part (a), this fraction is $\frac{\pi}{4}$ for large number of samples. Hence,

$$\text{Estimate for } \pi = \hat{\theta} = 4 \times \frac{\text{No. of points inside the circle}}{\text{N}}$$

Part (c)

N	Estimate of π
10	2.000000
100	3.240000
1000	3.208000
10000	3.131600
100000	3.142480
1000000	3.141868
10000000	3.141377
100000000	3.141560

MATLAB runs out of memory when it has to allocate space for 10^9 samples. Since all the samples are not required to be present in memory at all times(we simply need to count how many of them are inside the circle), we can generate smaller sized samples of size 10^8(in batches) and use them to estimate π. The initial code without batches couldn't handle large sample sizes. By taking batches of 10^8, we can handle cases for sample sizes of 10^9 or larger.

Part (d)

We need to find M such that,

$$\begin{aligned}
P\left(\pi - 0.01 \leq \hat{\theta} \leq \pi + 0.01\right) &\geq 0.95 \\
P\left(\pi - 0.01 \leq \frac{4 \times \#\text{inside}}{M} \leq \pi + 0.01\right) &\geq 0.95 \\
P\left(\frac{M(\pi - 0.01)}{4} \leq \#\text{inside} \leq \frac{M(\pi + 0.01)}{4}\right) &\geq 0.95 \\
P\left(\left\lceil\frac{M(\pi - 0.01)}{4}\right\rceil \leq \#\text{inside} \leq \left\lfloor\frac{M(\pi + 0.01)}{4}\right\rfloor\right) &\geq 0.95
\end{aligned}$$

Since X_1 and X_2 are independent and Uniform$(-1,1)$ random variables, (X_1, X_2) is uniformly distributed on the $[-1,1] \times [-1,1]$ square. Hence since #inside is a Binomial$(M, \frac{\pi}{4})$ random variable, we have

$$P\left(\left\lceil\frac{M(\pi - 0.01)}{4}\right\rceil \leq \#\text{inside} \leq \left\lfloor\frac{M(\pi + 0.01)}{4}\right\rfloor\right) = \sum_{i=\left\lceil\frac{M(\pi-0.01)}{4}\right\rceil}^{\left\lfloor\frac{M(\pi+0.01)}{4}\right\rfloor} \binom{M}{i}\left(\frac{\pi}{4}\right)^i\left(1 - \frac{\pi}{4}\right)^{M-i}$$

Thus, we need M which satisfies

$$\sum_{i=\left\lceil\frac{M(\pi-0.01)}{4}\right\rceil}^{\left\lfloor\frac{M(\pi+0.01)}{4}\right\rfloor} \binom{M}{i}\left(\frac{\pi}{4}\right)^i\left(1 - \frac{\pi}{4}\right)^{M-i} \geq 0.95$$

To find the required value of M, we apply binary search on the values of $M \in [1, 10^8]$ and keep the lowest M_{ans} for which $P \geq 0.95$. For each value of M, the LHS of the previous equation is calculated and compared with 0.95. If it exceeds 0.95, we set the upper bound to $M - 1$ and update M_{ans} if M is smaller than the recorded one, else set lower bound to $M + 1$. This is done because a sequential search takes a huge amount of time(85 s to search till the obtained value!).

Estimate $\pi = 3.142150$. For $M = 103482$. Probability=0.950137

The value of M_{ans} is not exact but very close to the actual value(computed with sequential gives a similar answer). Since the function that we are computing is not strictly monotonic, neither sequential search nor binary search will give exact answers. However the obtained estimate with binary search is a good one and is computed much faster than the sequential method.