HW4 — ISYE6416

Ashish Dhiman — ashish.dhiman9@gatech.edu

May 7, 2023

Contents

1 Question 1: Box Mulller methods

1.1 part a: Joint density

We are given that X,Y are jointly distributed as $X, Y \sim \mathcal{N}(0, I_2)$ and $R = \sqrt{X^2+Y^2}$, and $\Theta = \text{angle}[(X,Y)]$

Hence joint pdf is:

$$\implies f_{XY}(x,y) = \frac{1}{2\pi}\exp(-\frac{x^2+y^2}{2})$$

If we

$$\begin{aligned}
\text{We have, } X = R\cos\Theta, X = R\sin\Theta \\
\text{By change of variable} \\
f_{R\Theta}(r,\theta) = f_{XY}(r\cos\theta, r\sin\theta)|J| \quad \text{(J is Jacobian)} \\
\implies J = \begin{bmatrix} \frac{\partial X}{R} & \frac{\partial X}{\Theta} \\ \frac{\partial Y}{R} & \frac{\partial Y}{\Theta} \end{bmatrix} = \begin{bmatrix} \cos\Theta & -R\sin\Theta \\ \sin\Theta & R\cos\Theta \end{bmatrix} \\
\implies det(J) = R(\cos^2\Theta + \sin^2\Theta) = R^2 \\
\text{Hence,} \\
f_{R\Theta}(r,\theta) = \frac{1}{2\pi}\exp(-\frac{r^2\cos^2\theta + r^2\sin^2\theta}{2}) * r \\
= \frac{1}{2\pi}\exp(-\frac{r^2}{2}) * r
\end{aligned} \tag{1}$$

By Marginalisation of the joint pdf:

$$\begin{aligned}
f_\Theta(\theta) = \int_{r=0}^{\inf} f_{R\Theta}(r,\theta) * dr \\
= \frac{1}{2\pi}\int_{r=0}^{\inf} \exp(-\frac{r^2}{2}) * r * dr \\
= \frac{1}{2\pi}\int_{t=0}^{\inf} \exp(-\frac{t}{2}) * \frac{dt}{2} \\
= \frac{1}{2\pi} * 1
\end{aligned} \tag{2}$$

$$\begin{aligned}
f(\mathring{}) = \int_{\theta=0}^{2\pi} f_{R\Theta}(r,\theta) * d\theta \\
= \frac{1}{2\pi}\exp(-\frac{r^2}{2}) * r \int_{\theta=0}^{2\pi} 1 * d\theta \\
= \frac{1}{2\pi} * \exp(-\frac{r^2}{2}) * r * 2\pi
\end{aligned} \tag{3}$$

$$\implies f_{R\Theta}(r,\theta) = \frac{1}{2\pi}\exp(-\frac{r^2}{2}) * r = f_R(r) * f_\Theta(\theta)$$

Hence R and Θ are independent.

1.2 part b: $\Theta \sim$ unif$[0, 2\pi]$ and random variable $R^2 \sim \exp(2)$

From above, we know pdf for Θ is

$$f_\Theta(\theta) = \frac{1}{2\pi} = \frac{1}{2\pi - 0}$$

Also pdf of unif$[a,b] = \frac{1}{b-a}$. Hence we can say that $\Theta \sim$ unif$[0, 2\pi]$.
Similarly, for pdf of $Z = R^2$

$$\begin{aligned} f_Z(z) = f_R(z)|\frac{\partial \sqrt{z}}{\partial z}| \\ = \exp(-\frac{z}{2}) * \sqrt{z} * |\frac{-1}{2\sqrt{z}}| \\ = \exp(-\frac{z}{2}) * \frac{1}{2} \\ \implies Z = R^2 \sim \exp(1/2) \\ \text{as pdf of exponetial variable is } \lambda \exp(-\lambda x) \end{aligned} \tag{4}$$

2 Question 2: Monte Carlo & Bayes

We are given the random vector $Y = (U, V)$ has joint pdf $f_{UV}(u, v)$.

To approximate the marginal pdf $f_U(u) = \int f_{UV}(u, v)dv$, we take a random sample $y_1, \ldots, y_n$ from Y and form the average

$$A_n = \frac{1}{n}\sum_{i=1}^{n} \frac{f_{UV}(u, v_i)g(u_i)}{f_{UV}(u_i, v_i)}$$

Now,

$$\begin{aligned} f_U(u) = \int f_{UV}(u, v)dv \\ E_{g(u)}[f_U(u)] = \int_u f_U(u)g_U(u)du \\ = \int_u f_U(u)\frac{f_{UV}(u, v)}{f_{UV}(u, v)}g_U(u)du \\ = \int_u (\int_v f_{UV}(u, v)dv)\frac{f_{UV}(u, v)}{f_{UV}(u, v)}g_U(u)du \\ = E[\frac{(\int_v f_{UV}(u, v)dv)}{f_{UV}(u, v)}g_U(u)] \\ \text{Converting to Monte Carlo} \\ E[f_U(u)] = E(\frac{f_{UV}(u, v_i)g(u_i)}{f_{UV}(u_i, v_i)}) \\ \text{Now taking expectation of } E_n \\ E[A_n] = \frac{1}{n}\sum_{i=1}^{n} E(\frac{f_{UV}(u, v_i)g(u_i)}{f_{UV}(u_i, v_i)}) \\ \implies E[A_n] = \frac{1}{n}\sum_{i=1}^{n} E[f_U(u)] \end{aligned} \tag{5}$$

Hence A_n is an unbiased estimator

3 Question 3: Metropolis Hastings

We need to run MCMC chain for estimating Parameter for binomial distribution is probability of success $\theta \in [0, 1]$, $n = 20$. with the observed data vector having $S_n = 5$.

3.1 part a: $\pi(\theta|Y)$ with prior $\pi(\theta) = 2\cos^2(4\pi\theta)$

Code for the implementation is given below

```
def prior(x):
    return 2*np.square(np.cos(4*np.pi*x))

def likelihood(x,S,n):
    return np.power(x,S)*np.power(1-x,n-S)

def MHalgo(T,func_prior,func_likelihood,fx_proposal,start,theta_grid):
    chain = np.zeros(T)
    acceptance_arr = np.zeros(T)
    histogram = np.zeros(11)
    histogram[6] = 1
    chain[0]=start
    for t in range(1,T):
        theta_old = chain[t-1]
        #propose new point
        theta_proposal = fx_proposal(theta_old)
        #get acceptance ratio
        posterior_proposal = func_likelihood(theta_proposal)*func_prior(theta_proposal)
        posterior_old = func_likelihood(theta_old)*func_prior(theta_old)
        A = (posterior_proposal)/(posterior_old) #for symmteric proposal
        #accept or reject
        u = np.random.uniform()
        if u<=min(1,A):
            chain[t]=theta_proposal
            acceptance_arr[t] = 1
        else:
            chain[t] = theta_old
        histogram[np.argmin(np.abs(theta_grid-chain[t]))]+=1
    histogram = histogram/histogram.sum()
    print (f"AR:{np.mean(acceptance_arr):.4f}")
    return chain,np.mean(acceptance_arr),histogram

def proposal(old,sigma=0.2):
    return np.random.normal(loc=old,scale=sigma)
```

Discretizing θ to $0, 0.1 \ldots 1$, gives the following True Posterior

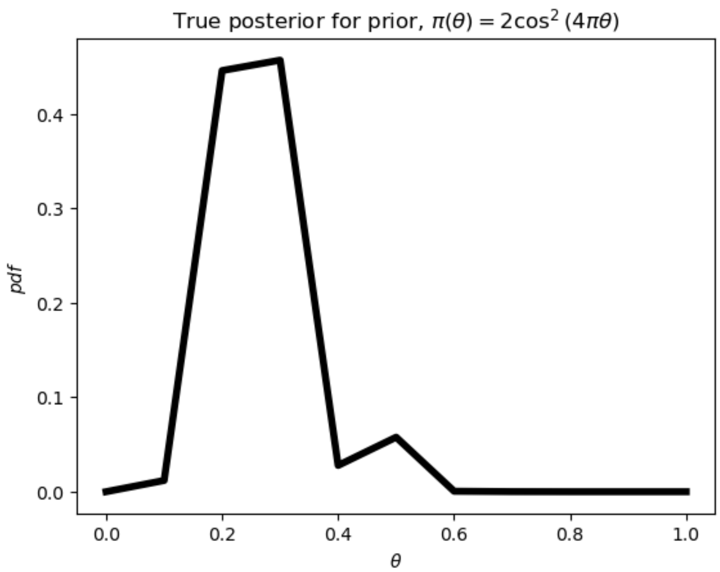

Figure 1: True Posterior

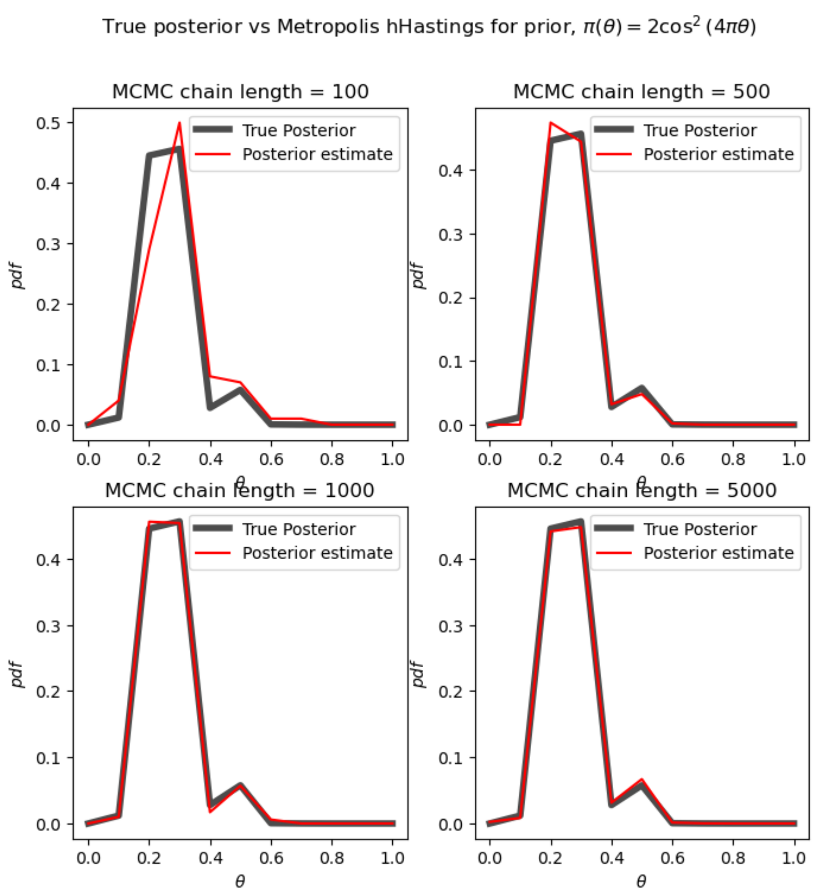

Figure 2: Estimated Posterior from MH

3.2 part b: numerical intergration

By using the posterior estimated from MH, we get:

$$\mathbb{E}^{\pi(\theta|Y)}\{\theta\} = 0.26863$$

$$\mathbb{E}^{\pi(\theta|Y)}\{[\theta - 1/2]^2\} = 0.059311$$

3.3 part c: Posterior with uniform prior

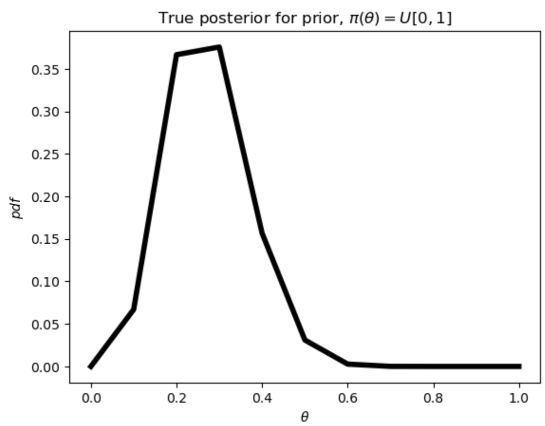

Figure 3: True Posterior

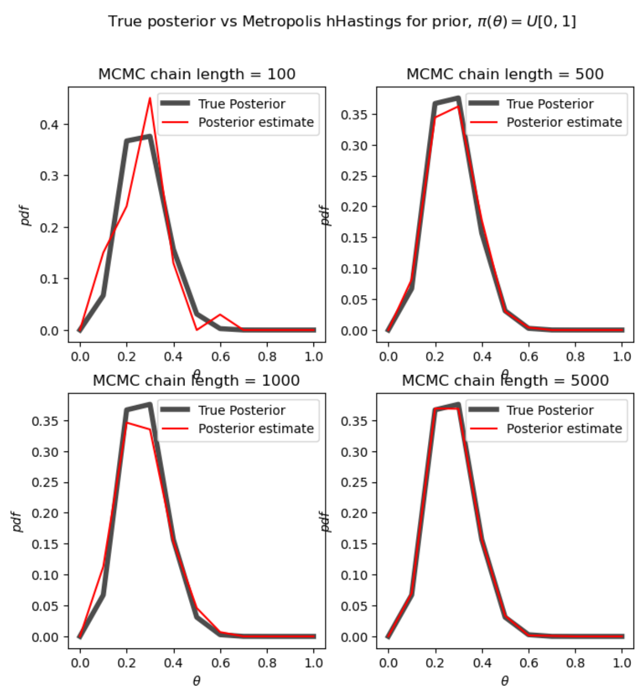

Figure 4: Estimated Posterior from MH

Similarly numerical integration values are:

$$\mathbb{E}^{\pi(\theta|Y)}\{\theta\} = 0.27470$$

$$\mathbb{E}^{\pi(\theta|Y)}\{[\theta - 1/2]^2\} = 0.05986$$

Note Posterior in this case is proportional to likelihood of Beta function, $beta(5+1, 25+1)$ which gives true mean value of theta as 0.27.

4 Question 4: Change Point detection

We have the generative model as follows:

$$
\begin{aligned}
n &\sim \text{Uniform}(1, 2, \ldots, N) \\
\lambda_i &\sim \text{Gamma}(a, b), i = 1, 2 \\
x_i | n, \lambda_1, \lambda_2 &\sim \begin{cases} \text{Poisson}(\lambda_1) & 1 \leq i \leq n \\ \text{Poisson}(\lambda_2) & n < i \leq N \end{cases}
\end{aligned}
$$

$$
\mathbb{P}(\lambda_1, \lambda_2, n | x_{1:N}) \propto \mathbb{P}(x_{1:n} | \lambda_1, n)\mathbb{P}(x_{n+1:N} | \lambda_2, n)\mathbb{P}(\lambda_1)\mathbb{P}(\lambda_2)\mathbb{P}(n).
$$

4.1 part a: posterior distribution $P(\lambda_1 | n, \lambda_2, x_{1:N})$

$$
\begin{aligned}
P(\lambda_1 | n, \lambda_2, x_{1:N}) &= \frac{P(\lambda_1, n, \lambda_2, x_{1:N})}{P(n, \lambda_2, x_{1:N})} \\
&= \frac{P(\lambda_1, n, \lambda_2 | x_{1:N})}{P(n, \lambda_2 | x_{1:N})} \\
\propto \frac{\mathbb{P}(x_{1:n} | \lambda_1, n)\mathbb{P}(x_{n+1:N} | \lambda_2, n)\mathbb{P}(\lambda_1)\mathbb{P}(\lambda_2)\mathbb{P}(n)}{P(n, \lambda_2 | x_{1:n}, x_{n+1:N})} \\
\propto \frac{\mathbb{P}(x_{1:n} | \lambda_1, n)\mathbb{P}(x_{n+1:N} | \lambda_2, n)\mathbb{P}(\lambda_1)\mathbb{P}(\lambda_2)\mathbb{P}(n)}{\frac{P(x_{1:n}, x_{n+1:N} | n, \lambda_2)P(n, \lambda_2)}{P(x_{1:N})}} \\
&\propto \mathbb{P}(x_{1:n} | \lambda_1, n) * P(\lambda_1) \\
&\text{By Generative model:} \\
P(x_{1:n} | \lambda_1, n) &= \prod_{i=1}^{n} p(x_i | \lambda_1, n) \qquad (6) \\
&= \prod_{i=1}^{n} \frac{\lambda_1^{x_i} \exp(-\lambda_1)}{x_i!} \\
&= \lambda_1^{\sum_{i=1}^{n} x_i} * \exp(-n\lambda_1) \\
&\text{and} \\
P(\lambda_1) = \Gamma(a, b) &= \frac{1}{\Gamma(a)} b^a \lambda_1^{a-1} \exp(-b\lambda_1) \\
\implies P(\lambda_1 | n, \lambda_2, x_{1:N}) &\propto \lambda_1^{a + \sum_{i=1}^{n} x_i - 1} * \exp(-(n+b)\lambda_1) \\
&\sim \Gamma(a + \sum_{i=1}^{n} x_i, n + b)
\end{aligned}
$$

4.2 part b: posterior distribution $P(\lambda_2|n, \lambda_1, x_{1:N})$

Similar to part a above, by using symmetry we get

$$\begin{aligned} P(\lambda_2|n, \lambda_2, x_{1:N}) &\propto \mathbb{P}(x_{n+1:N}|\lambda_2, n) * P(\lambda_2) \\ \implies P(x_{n+1:N}|\lambda_2, n) &= \lambda_2^{\sum_{i=n+1}^{N} x_i} * \exp(-(N-n)\lambda_2) \\ \implies P(\lambda_2) = \Gamma(a, b) &= \frac{1}{\Gamma(a)} b^a \lambda_2^{a-1} \exp(-b\lambda_2) \\ \text{Hence} \\ P(\lambda_2|n, \lambda_2, x_{1:N}) &\sim \Gamma(a + \sum_{i=n+1}^{N} x_i, (N-n) + b) \end{aligned} \tag{7}$$

4.3 part c: posterior distribution $P(n|\lambda_1, \lambda_2, x_{1:N})$

$$\begin{aligned} P(n|\lambda_1, \lambda_2, x_{1:N}) &= \frac{P(\lambda_1, n, \lambda_2, x_{1:N})}{P(\lambda_1, \lambda_2, x_{1:N})} \\ &= \frac{P(\lambda_1, n, \lambda_2|x_{1:N})}{P(\lambda_1, \lambda_2|x_{1:N})} \\ &\propto \frac{\mathbb{P}(x_{1:n}|\lambda_1, n)\mathbb{P}(x_{n+1:N}|\lambda_2, n)\mathbb{P}(\lambda_1)\mathbb{P}(\lambda_2)\mathbb{P}(n)}{\frac{P(x_{1:N}|\lambda_1,\lambda_2)*P(\lambda_1,\lambda_2)}{P(x_{1:N})}} \\ &= \frac{\mathbb{P}(x_{1:n}|\lambda_1, n)\mathbb{P}(x_{n+1:N}|\lambda_2, n)P(n)}{P(x_{1:N}|\lambda_1, \lambda_2)} \\ &\propto \mathbb{P}(x_{1:n}|\lambda_1, n)\mathbb{P}(x_{n+1:N}|\lambda_2, n)P(n) \\ \implies P(n|\lambda_1, \lambda_2, x_{1:N}) \sim \Gamma(a + \sum_{i=1}^{n} x_i, n + b) &* \Gamma(a + \sum_{i=n+1}^{N} x_i, +b) * U[1, N] \\ &\propto \exp(-n\lambda_1 + n\lambda_2)\lambda_1^{\sum_{i=1}^{n} x_i} \lambda_2^{\sum_{i=n+1}^{N} x_i} \end{aligned} \tag{8}$$

4.4 part d: posterior distribution $P(n|\lambda_1, \lambda_2, x_{1:N})$

For

$$N = 20, a = 1.2, b = 1.2$$

The posteriors are as follows:

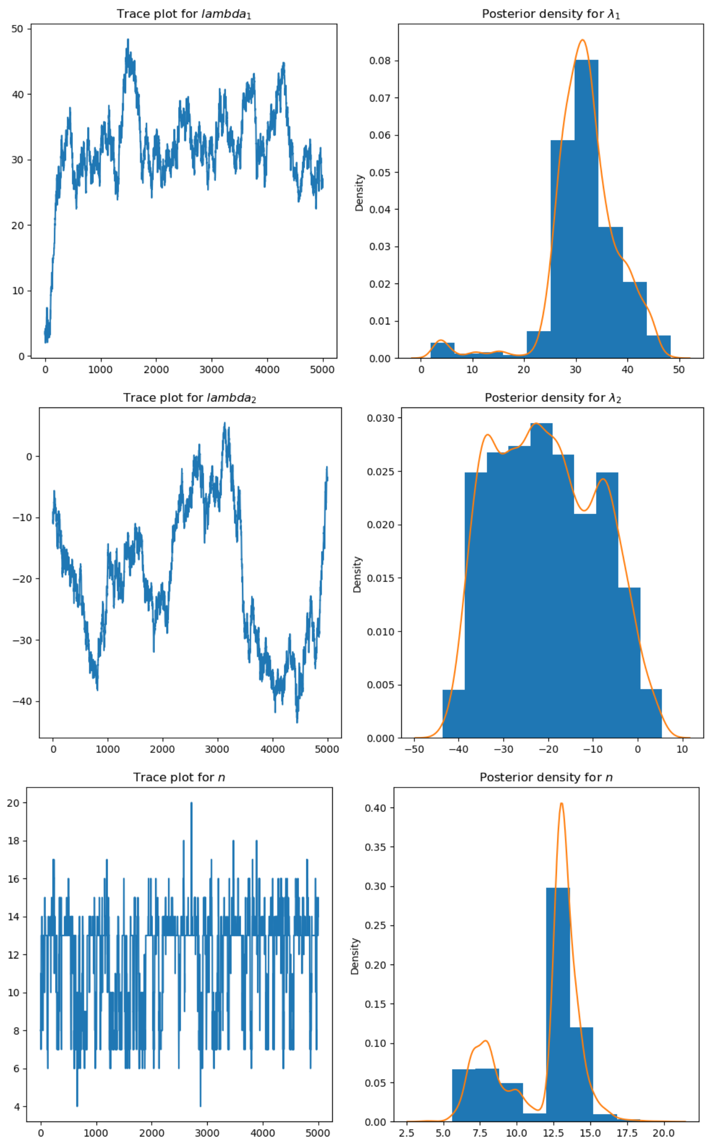

Figure 5: Estimated Posteriors from MH

5 References:

1. **Collaborators**: Yibei, Dipendra