Homework 8

Ziyang (Bob) Ding

31 March, 2020

Problem 1

Solve the following by defining an appropriate Markov chain and using first-step analysis:

- (a) **In repeated (fair) coin tosses:**
  - i. **Calculate the expected number of tosses until 2 successive heads.**
  - ii. **Of the two patterns HHT and HTH, which do you expect to observe first and why.**
- (b) **Rat maze: for a rat located as shown performing a random walk calculate the probability of finding food before exiting.Note that the exit is on the wall, ie the rat can move down 1 step without exiting the maze.**

Proof

- (a)
  - Denote $f(n)$ as expected step needed from already having n consecutive heads to achieving 2 consecutive heads. So we are interested in $f(0)$. By Komogorov extension theorem, we have

    $$\begin{aligned} f(2) &= 0 \\ f(1) &= 1 + \frac{1}{2}f(2) + \frac{1}{2}f(0) = 1 + \frac{1}{2}f(0) \\ f(0) &= 1 + \frac{1}{2}f(1) + \frac{1}{2}f(0) \end{aligned}$$

    We have that $f(0) = 6$
  - HTH is going to be faster as if you get T after H, you lose everything and have to start from the bottom. However, for HTH, if you get H after another H, you don't lose anything. The rest of the game is symmetrical so we don't need to worry. Therefore, HTH is faster.

- (b)

I don't like this question, especially in latex: Label the rooms using 1,2,3(food),4(Rat),5,6,7,8(exit). Denote $f(n)$ as probability that n will go to exit and been absorbed before it reaches the food. Therefore, we have:

$$
\begin{aligned}
f(8) &= 1 \\
f(7) &= \frac{1}{2} + \frac{1}{2} f(4) \\
f(6) &= \frac{1}{2} f(3) + \frac{1}{2} f(5) \\
f(5) &= \frac{1}{3} f(2) + \frac{1}{3} f(4) + \frac{1}{3} f(6) \\
f(4) &= \frac{1}{3} f(1) + \frac{1}{3} f(5) + \frac{1}{3} f(7) \\
f(3) &= 0 \\
f(2) &= \frac{1}{2} f(1) + \frac{1}{2} f(3) + \frac{1}{3} f(5) \\
f(1) &= \frac{1}{2} f(2) + \frac{1}{2} f(4)
\end{aligned}
$$

Solve for solution. We get the answer

Problem 2

Define a Markov chain on $0, \dots, N$ by:

$$
P_{i,j} = \begin{cases} 1 & i = j = 0 \\ 1/i & 0 \le j < i \le N \\ 0 & \text{otherwise} \end{cases}
$$

i.e. from state j the chain is equally likely to go to $0, \dots, j-1$

- (a) **Determine the fundamental matrix for the transient states.**

$$
\sum_{i=1}^{\infty} G^i = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots 0 & 0 \\
1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots 0 & 0 \\
1/2 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & \dots 0 & 0 \\
1/2 & 1/3 & 1/4 & 0 & 0 & 0 & 0 & 0 & \dots 0 & 0 \\
1/2 & 1/3 & 1/4 & 1/5 & 0 & 0 & 0 & 0 & \dots 0 & 0 \\
1/2 & 1/3 & 1/4 & 1/5 & 1/6 & 0 & 0 & 0 & \dots 0 & 0 \\
1/2 & 1/3 & 1/4 & 1/5 & 1/6 & 1/7 & 0 & 0 & \dots 0 & 0 \\
1/2 & 1/3 & 1/4 & 1/5 & 1/6 & 1/7 & 1/8 & 0 & \dots 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots 0 & 0 \\
1/2 & 1/3 & 1/4 & 1/5 & 1/6 & 1/7 & 1/8 & 1/9 & \dots 1/n & 0
\end{bmatrix}
$$

- (b) **Determine the distribution of the last positive integer visited.**

Proof

- (a)

- (b)

Problem 3

Choose one of the above problems (either 1(a) **ii,** 1b, **or** 2b **) and validate your results by:**

- (a) **Computing large powers of the transition matrix.**
- (b) **Simulating realizations of the Markov chain.**

Proof

- (a)

- (b)

Problem 4

The Gibbs sampler defined in class cycles through coordinates in a fixed order. Alternatively, we may define a random-scan Gibbs sampler, which iteratively chooses $i \in \{1, \ldots, d\}$ **at random (according to probabilities** p_i **say), and sets**

$$\theta^{(n+1)} = \left(\theta_1^{(n)}, \ldots, \theta_{i-1}^{(n)}, \theta_i^*, \theta_{i+1}^{(n)}, \ldots, \theta_d^{(n)}\right)$$

with

$$\theta_i^* \sim \pi\left(\theta_i | \theta_1^{(n)}, \ldots, \theta_{i-1}^{(n)}, \theta_{i+1}^{(n)}, \ldots, \theta_d^{(n)}\right)$$

- (a) **Show that this also produces a Markov chain with stationary distribution** π
- (b) **Give sufficient conditions for this chain to have limiting distribution** π

Proof

- (a)

$$\pi(\theta_x)P(\theta_x, \theta_y) = \pi(\theta_y)P(\theta_y, \theta_x)$$

The chain is reversible. Besides,

$$\begin{aligned} P\left(\theta_{n+1} \in A | \theta_0, \ldots, \theta_n\right) &= P\left(\theta_{n+1} \in A | \theta_n\right) \\ &= \int_A K\left(\theta_n, d\theta\right) \end{aligned}$$

In this way, we know that the chain is memoriless. Therefore, it is a markov chain with distribution π.

- (b)

By theorem, we need the chain to be π-invariant, π-irreducible, aperiodic, and Harris recurrent. To be π-invariant.

Problem 5

The famous "braking data" of Tukey (1977) is available in R under the dataset name cars. It gives the speeds traveled (mph) and braking distances (feet) for 50 cars. It is thought that a good model for this dataset is a quadratic model:

$$y_{ij} = a + bx_i + cx_i^2 + \epsilon_{ij} \quad \text{for } i = 1, \ldots, k; \quad j = 1 \ldots, n_i$$

- (a) **Write down the likelihood function assuming** $\epsilon_{ij} \sim N\left(0, \sigma^2\right)$

*(b) **Obtain estimates of** a, b, c **and** σ^2 **from a standard linear regression.**

*(c) **View the likelihood in part (a) as a posterior distribution under flat priors, and construct a Metropolis-Hastings algorithm to sample from it using a Metropolized independence sampler with proposal distributions selected based on your estimates in (b). Use normals for** a, b, c **and inverse-gamma for** σ^2

*(d) **Make histograms of the posterior distributions of the parameters. Show any plots or diagnostics used to monitor convergence.**

*(e) **Consider robustness by modifying the error distribution to** $\epsilon_{ij} \sim t_4\left(0, \sigma^2\right)$ **and re-running your analysis. Do you need to modify your proposal distributions?**

Proof

- (a)

Denote X_i as aligning $[1, x_i, x_i^2]$ vertically for n_i times, so we can form k $n_i \times 3$ matricies. Then, we concatenate X_i vertically to form a X as $\sum_{i=1}^k n_i \times 3$ matrix. Similarly, we align all the corresponding y_{ij} vertically, and we get a $\sum_{i=1}^k n_i$ dimensional vector. The likelihood is that

$$\begin{aligned}
Y &= X\beta + I\epsilon \\
\beta &= (a, b, c)^\top \\
\epsilon &\sim N(0, I\sigma^2) \\
\mathcal{L}(Y; X, \beta) &= (2\pi)^{-\sum_{i=1}^k n_i/2} \sigma^{-\sum_{i=1}^k n_i} \exp\left\{-\frac{1}{2}\sigma^{-2\sum_{i=1}^k n_i} \beta^T X^T X \beta\right\}
\end{aligned}$$

- (b)

$$(a, b, c) \sim \mathcal{N}\left((X^T X)^{-1} XY, (X^T X)^{-1}\sigma^2\right)$$

- (c)

- (d)

- (e)