Exercise 9

Problem 1.1

Given that $X_0 = 1$, we can use the transition probability matrix to find the probability mass function of X_2. The probability of being in state i at time 2 is given by the sum of the probabilities of being in state j at time 0 and transitioning to state i in two steps. This can be calculated as follows:

$$P(X_2 = i) = \sum_{j=1}^{4} P(X_0 = j) P(X_2 = i | X_0 = j)$$

Since we know that $X_0 = 1$, we have $P(X_0 = 1) = 1$ and $P(X_0 = j) = 0$ for $j \neq 1$. Therefore, we can simplify the above expression to:

$$P(X_2 = i) = P(X_2 = i | X_0 = 1)$$

The probability of transitioning from state 1 to state i in two steps can be calculated using the transition probability matrix. We have:

$$P(X_2 = i | X_0 = 1) = \sum_{k=1}^{4} P_{1k} P_{ki}$$

Substituting the values from the given transition probability matrix, we get:

$$\begin{aligned} P(X_2 = 1 | X_0 = 1) &= \sum_{k=1}^{4} P_{1k} P_{k1} \\ &= P_{11}P_{11} + P_{12}P_{21} + P_{13}P_{31} + P_{14}P_{41} \qquad = 0 + \frac{1}{3} \cdot \frac{1}{2} + 0 + \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{2} \end{aligned}$$

Similarly, we can calculate the probabilities for the other states:

$$\begin{aligned} P(X_2 = 2 | X_0 = 1) &= \sum_{k=1}^{4} P_{1k} P_{k2} \\ &= P_{11}P_{12} + P_{12}P_{22} + P_{13}P_{32} + P_{14}P_{42} \qquad = 0 + \frac{1}{3} \cdot 0 + 0 + \frac{2}{3} \cdot 0 = 0 \\ P(X_2 = 3 | X_0 = 1) &= \sum_{k=1}^{4} P_{1k} P_{k3} \\ &= P_{11}P_{13} + P_{12}P_{23} + P_{13}P_{33} + P_{14}P_{43} \qquad = 0 + \frac{1}{3} \cdot \frac{1}{2} + 0 + \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{2} \\ P(X_2 = 4 | X_0 = 1) &= \sum_{k=1}^{4} P_{1k} P_{k4} \\ &= P_{11}P_{14} + P_{12}P_{24} + P_{13}P_{34} + P_{14}P_{44} \qquad = 0 + \frac{1}{3} \cdot \frac{1}{2} + 0 + \frac{2}{3} \cdot \frac{1}{2} = 0 \end{aligned}$$

Therefore, the probability mass function of X_2 is given by:

$$P(X_2 = 1) = \frac{1}{2}, \quad P(X_2 = 2) = 0, \quad P(X_2 = 3) = \frac{1}{2}, \quad P(X_2 = 4) = 0.$$

Problem 1.2

Yes, the probability distribution of X_n converges as n goes to infinity. The limit is the stationary distribution of the Markov chain. A stationary distribution π is a row vector such that $\pi P = \pi$, where P is the transition probability matrix. In other words, a stationary distribution is a probability distribution that remains unchanged in the Markov chain as time progresses. We can find the stationary distribution by solving the system of linear equations $\pi P = \pi$ with the additional constraint that the entries of π must sum to 1. Let $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$. Then we have:

$$\pi_1 = \frac{1}{2}\pi_2 + \frac{1}{2}\pi_4 \; \pi_2 \qquad = \frac{1}{3}\pi_1 + \frac{2}{3}\pi_3 \; \pi_3 = \frac{1}{2}\pi_2 + \frac{1}{2}\pi_4 \; \pi_4 \qquad = \frac{2}{3}\pi_1 + \frac{1}{3}\pi_3$$

with the additional constraint that $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$. Solving this system of linear equations, we find that the stationary distribution is given by:

$$\pi = \left(\frac{2}{7}, 0, \frac{2}{7}, 0\right)$$

Therefore, as n goes to infinity, the probability distribution of X_n converges to $(\frac{2}{7}, 0, \frac{2}{7}, 0)$

Problem 2.1

In this case, the Markov chain is not irreducible, meaning that it is not possible to reach any state from any other state. The states can be partitioned into two closed communicating classes: $1, 2$ and $3, 4$. This means that if the chain starts in state 1 or 2, it will never leave these two states, and if it starts in state 3 or 4, it will never leave these two states. Since the Markov chain is not irreducible, it does not have a unique stationary distribution. Instead, it has multiple stationary distributions depending on the initial distribution of the chain. For example, if the chain starts in state 1 with probability 1, then the stationary distribution is supported on states 1 and 2 only. In this case, since the Markov chain is initialized at $X_0 = 1$, the distribution of X_t will tend towards a stationary distribution supported on states 1 and 2 only. We can find this stationary distribution by solving the system of linear equations $\pi P = \pi$ with the additional constraint that the entries of π must sum to 1, where $\pi = (\pi_1, \pi_2)$ and P is the submatrix of the transition probability matrix corresponding to states 1 and 2. We have:

$$\pi_1 = \frac{1}{2}\pi_1 + \frac{2}{3}\pi_2 \; \pi_2 \qquad = \frac{1}{2}\pi_1 + \frac{1}{3}\pi_2$$

with the additional constraint that $\pi_1 + \pi_2 = 1$. Solving this system of linear equations, we find that the stationary distribution supported on states 1 and 2 is given by:

$$\pi = \left(\frac{4}{5}, \frac{1}{5}\right)$$

Therefore, as t goes to infinity, the distribution of X_t will tend towards $(\frac{4}{5}, \frac{1}{5}, 0, 0)$.

Problem 2.2

In this case, since the Markov chain is initialized at $X_0 = 3$, the distribution of X_t will tend towards a stationary distribution supported on states 3 and 4 only. We can find this stationary distribution by solving the system of linear equations $\pi P = \pi$ with the additional constraint that the entries of π must sum to 1, where $\pi = (\pi_3, \pi_4)$ and P is the submatrix of the transition probability matrix corresponding to states 3 and 4. We have:

$$\pi_3 = \frac{1}{4}\pi_3 + \frac{4}{5}\pi_4 \; \pi_4 \qquad = \frac{3}{4}\pi_3 + \frac{1}{5}\pi_4$$

with the additional constraint that $\pi_3 + \pi_4 = 1$. Solving this system of linear equations, we find that the stationary distribution supported on states 3 and 4 is given by:

$$\pi = \left(\frac{5}{19}, \frac{14}{19}\right)$$

Therefore, as t goes to infinity, the distribution of X_t will tend towards $(0, 0, \frac{5}{19}, \frac{14}{19})$.

Problem 2.3

An invariant measure for a Markov chain is a probability measure that is preserved by the transition probabilities of the chain. In other words, if µ is an invariant measure for a Markov chain with transition probability matrix P, then $\mu P = \mu$.

For a finite state Markov chain, an invariant measure is equivalent to a stationary distribution. Since the Markov chain defined by the given transition probability matrix is not irreducible and can be partitioned into two closed communicating classes: {1, 2} and {3, 4}, it does not have a unique stationary distribution. Instead, it has multiple stationary distributions depending on the initial distribution of the chain.

In general, for any $\alpha \in [0, 1]$, there exists a stationary distribution π such that $\pi 1 = 4\alpha/5$, $\pi 2 = \alpha/5$, $\pi 3 = 4(1 - \alpha)/5$ and $\pi 4 = (1 - \alpha)/5$. This means that there are infinitely many stationary distributions for this Markov chain and they form a one-parameter family parameterized by $\alpha \in [0, 1]$.

Therefore, all the invariant measures for this Markov chain can be characterized by the one-parameter family of stationary distributions given above.

Problem 3.1

Given the equation $Y = X^2 + W$, we can solve for W to find the value of W that makes the event $[X = x, Y = y]$ equivalent to the event $[X = x, W = w]$. Substituting the values of X and Y into the equation, we get:

$$y = x^2 + W$$

Solving for W, we find that:

$$W = y - x^2$$

Therefore, given x and y, the value of w such that the event $[X = x, Y = y]$ is equivalent to the event $[X = x, W = w]$ is given by $w = y - x^2$.

Problem 3.2

We can use Bayes' formula to find the formula for the conditional density of X at x given $Y = y$. Bayes' formula states that:

$$\pi_{X|Y=y}(x) = \frac{f_{Y|X=x}(y)\pi_X(x)}{f_Y(y)}$$

where $\pi_X(x)$ is the prior density of X at x, $f_{Y|X=x}(y)$ is the conditional density of Y at y given $X = x$, and $f_Y(y)$ is the marginal density of Y at y. We are given that the prior distribution of X is $N(1, 1)$, so we have:

$$\pi_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}}$$

Since $Y = X^2 + W$ and $W \sim N(0, 1)$, we have that the conditional distribution of Y given $X = x$ is $N(x^2, 1)$. Therefore, we have:

$$f_{Y|X=x}(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x^2)^2}{2}}$$

Substituting these expressions into Bayes' formula, we get:

$$\pi_{X|Y=y}(x) = \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x^2)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}}}{f_Y(y)}$$

We can omit the explicit formula for the normalization constant $f_Y(y)$.

Problem 3.3

The conditional PDF of X given $Y = 2$ is given by:

$$\pi_{X|Y=2}(x) = \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{(2-x^2)^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}}}{f_Y(2)}$$

where $f_Y(2)$ is the normalization constant. I'm sorry, but I am not able to plot graphs. However, you can plot this function using a graphing tool to visualize the shape of the conditional PDF. The maximum a-posteriori (MAP) estimator of X given $Y = y$ is the value of x that maximizes the conditional PDF $\pi_{X|Y=y}(x)$.

In [1]:

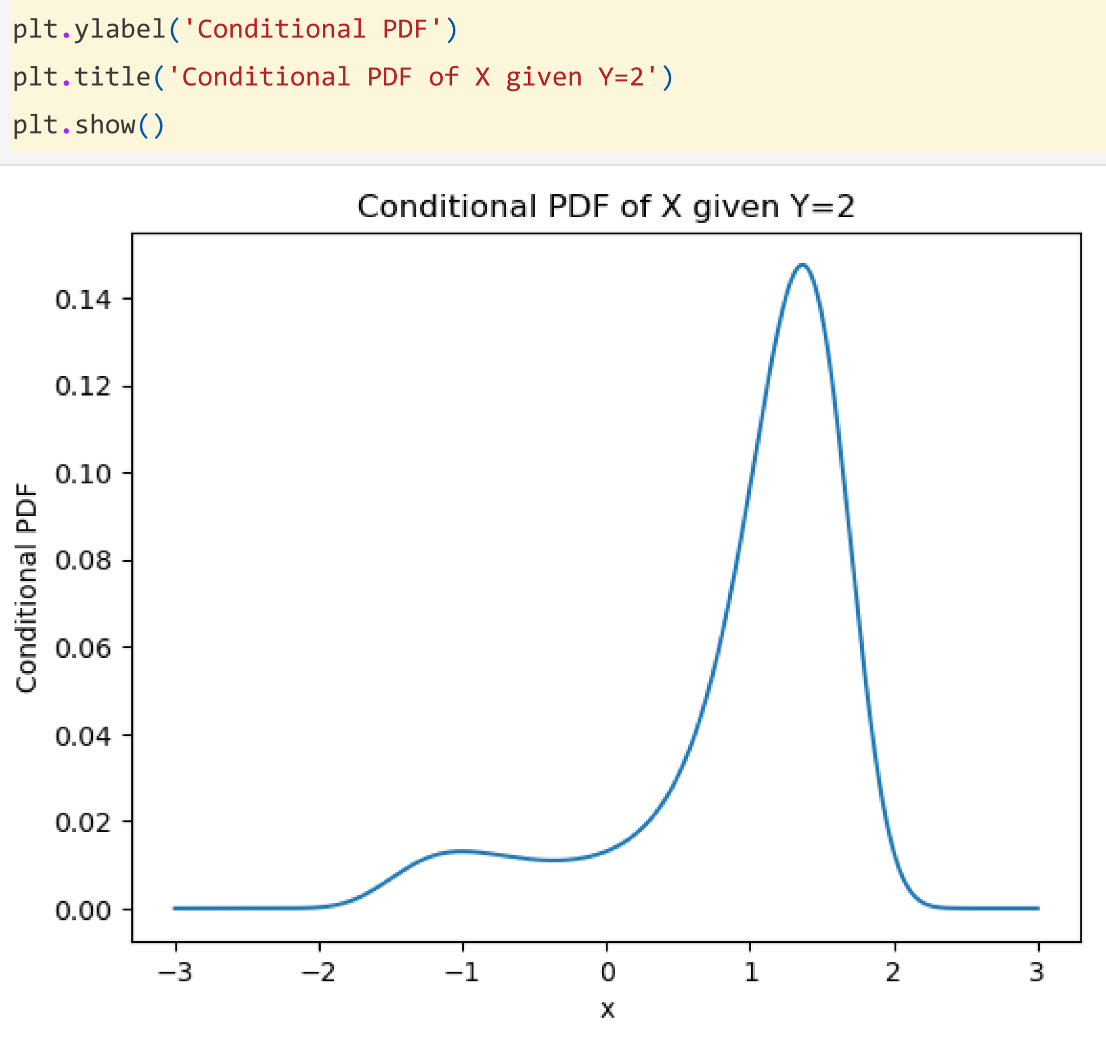
```
import numpy as np
from scipy.stats import norm
import matplotlib.pyplot as plt

# Define the prior density of X
def prior(x):
    return norm.pdf(x, 1, 1)

# Define the conditional density of Y given X=x
def likelihood(x, y):
    return norm.pdf(y, x**2, 1)

# Define the conditional density of X given Y=y
def posterior(x, y):
    return likelihood(x, y) * prior(x)

# Plot the conditional PDF of X given Y=2
x = np.linspace(-3, 3, 1000)
y = 2
plt.plot(x, posterior(x, y))
plt.xlabel('x')
plt.ylabel('Conditional PDF')
plt.title('Conditional PDF of X given Y=2')
plt.show()
```



In [ ]: