HHU Düsseldorf
Math.-Nat. Fakultät
Prof. Dr. Nils Detering
Dr. Nicole Hufnagel

Summer semester 2024

Markov Chains

Problem sheet 6

Markov chains: limiting distributions. Markov Decision Problems basics

Problems to be discussed (in parts) during the exercise sections:

Problem 1

Let $(X_n)_{n\geq 0}$ be a Markov chain with state space $\mathcal{S} = \{1, 2, 3\}$ and transition matrix

$$P = \begin{pmatrix} \frac{1}{4} & 0 & \frac{3}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

(a) Is this Markov chain irreducible and aperiodic?

(b) Compute for all $i, j \in \mathcal{S}$ the limit

$$\lim_{n\to\infty} \mathbb{P}[X_n = j \mid X_0 = i].$$

Problem 2

A store manages its inventory of a single product using a controlled Markov Chain. The store can decide whether to order new stock or not at the end of each day. The inventory level can be between 0 and 3 units. Each day, the demand for the product follows a random process. If the demand exceeds the available inventory, the excess demand is lost (i.e., it results in lost sales). The goal is to study the limiting distribution of the inventory levels under a specific ordering policy.

States and Actions

- **States**: The state of the system is the inventory level, which can be 0, 1, 2, or 3.
- **Actions**: At the end of each day, the store can either:
  - **Order** one unit of the product.
  - **Do not order** any new stock.

Demand Distribution: The daily demand D for the product follows the distribution:

$$\begin{aligned} \mathbb{P}(D=0) &= 0.5 \\ \mathbb{P}(D=1) &= 0.3 \\ \mathbb{P}(D=2) &= 0.2 \end{aligned}$$

Transition Probabilities: Let s be the current inventory level and a be the action taken (0 for not ordering, 1 for ordering one unit). The next state s' is determined by the current state, the action taken, and the demand D. For example: $\mathbb{P}(s'=3|s=2, a=1) = P(D=0) = 0.5$

Ordering Policy: Consider a simple policy where the store orders one unit whenever the inventory level is 0 or 1, and does not order otherwise.

(a) **Define the Transition Matrix**: Construct the transition matrix for the Markov Chain under the given policy.

(b) **Find the Limiting Distribution**: Calculate the limiting distribution of the inventory levels under this policy.

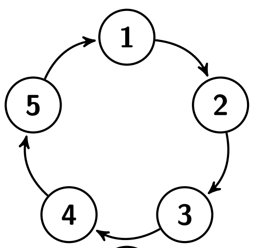
(c) **Interpret the Results**: Discuss the implications of the limiting distribution for the store's inventory management.

Problems to be handed in by:
Thursday, July 11th, 11:59 p.m., online via Ilias.

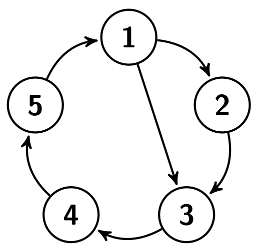
Problem 1 *(6 points)*

Which of the following Markov chains with corresponding transition graph are irreducible and aperiodic? Give a short explanation of your answer.

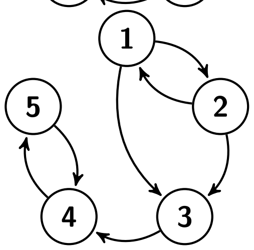
a)



b)



c)



Which of these Markov chains converges to a unique stationary distribution?

Problem 2 *(8 points)*

Solve the python programming exercise in the Jupyter file CreditRatings.ipynb

Problem 3 *(8 points)*

Solve the python programming exercise in the Jupyter file WindSpeed.ipynb

For the programming problems, please execute your code and then save your Jupyter notebook with all output in pdf. Merge this pdf with all the other problems you hand in. Then upload as usual to Ilias.

Problem 4 *(8 points)*

A warehouse robot needs to navigate through a grid-like warehouse to collect multiple items and deliver them to a designated delivery spot. The warehouse is represented as a 5×5 grid, and the robot can move in four directions: up, down, left, and right. Some cells in the grid may contain obstacles, which the robot cannot pass through.

The robot starts at a given position in the warehouse and must collect items placed at specific (fixed, see grid below) locations and then deliver them to the delivery spot. Each movement costs the robot a certain amount of energy (reward -1), collecting an item provides a reward (+10), and delivering the item to the delivery location also provides a reward (+20), hitting an obstacle leads to reward of -5. The goal is to maximize the total reward collected.

Grid Layout

- **S**: Starting position of the robot
- **D**: Delivery spot
- **O**: Obstacle
- **I**: Item location

Grid:

```
S . . . O
. O I . .
. . . O .
O I . . .
. . O . D
```

(a) **Define the States and Actions to obtain a MDP**: List all possible states and actions for the robot.

(b) **Reward Function**: Construct the reward function for the given grid.

(c) **Transition Function**: Define the transition function for the robot's movements.

(d) **Optimal Policy**: Describe the optimal policy for the robot to follow.