INDENG 290: Machine Learning with Applications in Electronic Markets

Two-Part Assignment

Instructor: Matteo Santamaria
`matteosantamaria@berkeley.edu`

Fall 2024

This assignment consists of two parts with separate due dates. Part I: Read the questions below and familiarize yourself with the code in `assignment.py`. Your task is to write two test cases for each question. You must also submit work proving your test cases are correct. I have indicated in `assignment.py` where exactly to implement your test cases. Once you have done so, you should submit the file to Gradescope. On submission, your test cases will be run against an autograder. You can view your results and resubmit as often as you'd like before the deadline. Part I is due **October 7th**.

For the second part of the assignment you must implement code to solve the three questions. This time you need only submit `assignment.py` (no second submission). Again, your code will be run against and autograder and you can resubmit as often as you'd like. Part II is due **November 4th**.

> You may use any functions in the Python standard library but you cannot rely on any third-party dependencies. Your code should run on Python 3.10.

Question 1. Markov Chain on Infinite State Space

Consider a state space $\mathcal{S} = \mathbb{N}_0 = \{0, 1, 2, \dots\}$ with transition kernel

$$\Pr\left(X_{n+1} = x_{n+1} \mid X_n = 0\right) = \begin{cases} \frac{1}{2} & \text{if } x_{n+1} = 1 \\ \frac{1}{2} & \text{if } x_{n+1} = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\Pr\left(X_{n+1} = x_{n+1} \mid X_n = x_n \neq 0\right) = \begin{cases} \frac{1}{2} & \text{if } x_{n+1} = x_n + 1 \\ \frac{1}{2x_n} & \text{if } x_{n+1} < x_n \\ 0 & \text{otherwise} \end{cases}$$

Determine $\Pr(X_n = x)$ for any $x \in \mathcal{S}$ given a initial state distribution π_0 and positive integer n.

Question 2. Finite-Horizon Markov Decision Process

You are a stockbroker and a client has just sent you an order to sell $2n \in \mathbb{N}$ shares on their behalf. The client gives you a horizon $T \in \mathbb{N}$ to fully execute her order. The stock price at time $t \in \{0, 1, \ldots, T\}$ is denoted $S_t \in \mathbb{N}_0$. To simplify the analysis we assume every timestep the stock rises by \$1.

The exchange has a few peculiar rules. At each timestep you observe S_t and submit an order $A_t \in \mathbb{N}_0$. After you have submitted your order, K_t buyers arrive where $K_t \sim \text{Binomial}(n, \frac{1}{2})$.

- If $A_t \leq K_t$ then your order is fully executed.
- If $A_t > K_t$ then only K_t shares are sold.
- If $A_t > K_t$ then you must pay $c\,(A_t - K_t)$ to the exchange.

$c \in \mathbb{R}_+$ is a known parameter. Lastly, at time $t = T$ you are forced to liquidate all your remaining shares at the price $\frac{1}{2}S_T$. What is total expected revenue under an optimal policy? Short selling is not allowed.

Question 3. Grid World Markov Decision Process

Consider a $N \times N$ grid where N is odd. We label the grid with $(0, 0)$ in the top-left corner and $(N-1, N-1)$ in the bottom-right corner. An agent navigating the environment starts in state $s_0 = (0, 0)$ and has four actions available to her: $\{\texttt{L}, \texttt{R}, \texttt{D}, \texttt{U}\}$.

$(0,0)$				
		`U`		
	`L`	S_t	`R`	
		`D`		
$(4,0)$				$(4,4)$

If the agent is against a wall and tries to move that direction she remains in place.

You are given a reward specification

$$r(s,a) = \begin{cases} +1 & s = (N-1, N-1) \\ -\alpha & s = \left(\frac{N}{2}, \frac{N}{2}\right) \\ -\beta & s = (N-1, 0) \\ 0 & \text{otherwise} \end{cases}$$

with parameters $\alpha, \beta \geq 0$. After choosing an action and moving in that direction (if possible), the agent is blown either `L` or `D` each with probability $1/4$ (she remains in place with probability $1/2$). What is total expected reward

$$\mathbb{E}\left[\sum_{t=0}^{T} \left(\frac{1}{2}\right)^t r(S_t, A_t) \mid S_0 = s_0\right]$$

under an optimal policy?