Assignment 2

Deep Reinforcement Learning

Spring 2022

A Newspaper Salesman Problem

Let's say you want to sell newspapers. At the beginning of each day i, you choose the number n_i of newspapers to buy. After you buy them, you have d_i customers. You buy the newspapers at price p and sell them at a price $p' > p$. Any unsold newspapers at the end of the day can be recycled, in which case you receive $.1 * p$ per newspaper.

Part 1

Write an expression for your profit $r(i)$ on day i, keeping in mind that the number of newspapers that are recycled is given by $\max\{0, n - d\}$.

Part 2

Assume there is a limit to the number of newspapers you can buy a day, which is fixed and given by K. What is the action space?

Part 3

In this problem, let's assume you start off with 0 dollars, and can buy up to K newspapers regardless if your balance is positive or negative. Write an equation for Bellman's Equation. Let's say your policy π is either to buy 10 newspaper or 20 newspapers, depending on your balance (i.e. your state s). Give an equation for Bellman's Equation. Note that your action policy should be deterministic.

Exploration Exploitation Tradeoff

Part 1

In your own words, what is the exploration exploitation tradeoff?

Part 2

In a dynamic programming problem, if the variance is high, is it better to explore or exploit? Does it depend on the amount of risk you are willing to take? What does risk mean to you in the context of dynamic programming problems? Please explain.

Portfolio Investing

Let's say you start off with a balance of 100 dollars. There are two stocks you'd like to invest in, stock s_1 and stock s_2. At the beginning of each day, you can decide how much money m_1 to invest in stock s_1 and m_2 to invest in stock s_2. You must invest all of your money each day, but you can re-allocate at the end of each day. That is, on each day i, you have $m_1(i) + m_2(i) = b_i$, where b_i is your initial balance on day i (i.e. the state).

Part 1

Suppose on day i, the return of stocks one and two, respectively, are given by $r_1(i)$ and $r_2(i)$. In this problem, by return, we actually mean a multiplier. So if you invest 10 dollars in stock s_1 on day i, you will end up with $10 * r_1(i)$ at the end of the day. We assume that r_1 and r_2 are always between .9 and 1.1. What is the action space?

Part 2

Suppose on day i, you invest $m_1(i)$ and $m_2(i)$ in each stock. Write an equation for your *profit* (not balance) at the end of the day $r(i)$.

Part 3

Suppose now that you have three options at the beginning of each day. You can either invest all your money in stock 1, all your money in stock 2, or fifty percent in each stock. Write an equation for Bellman's Equation. Note that we are considering non-stochastic policies for this problem.