Announcements

• The assignment is due by 5 PM, Monday, November 30, 2015.
• You may work with another student in any section on this assignment and submit just one writeup. Please be sure to record both your names in your submission. Partnered assignments mean that you work together on the assignment, with the goal being to cut down on time spent on the writeup. The intent is not for you to split up the problems and work on them separately.
• Please remember to write your TF’s name on the front of your assignment.
• Readings: Markov Chain and MDP Handouts.

Goals

This assignment will give you a feel for modeling, solving, and analyzing Markov Chains, Stochastic Optimization, and Markov Decision Processes. It will also immerse you in the Mushroom Kingdom universe.

Contents

1 Let’s gamble 2
2 Let’s go clubbing 3
3 Let’s play Monopoly 6
4 Let’s help find that song 10
5 Let’s sell lemonade 11
6 Let’s turn in the assignment 13
1 Let’s gamble

Mario and Luigi are hanging out one afternoon and decide to put their coins where their game is. “The rules are simple,” Mario says. “At each round, we play a Nintendo game. If I win, you give me a coin. If you win, I give you a coin. We play until I lose all my coins or you lose all your coins.”

Mario has $M$ coins at the start and Luigi has $L$ coins. They are equally matched in all Nintendo games, such that each wins in a game with probability $1/2$.

This game can be described as follows: States $i$ denote the amount Mario has won or lost at any given time. Since Mario can win no more than $L$ coins (all of Luigi’s coins) and lose no more than $M$ coins (all his coins), we have states $i$ for $-M \leq i \leq L$. To model the absorbing states, the transition probability is:

$$p_{-M-M} = 1 \quad \text{and} \quad p_{L-L} = 1$$

For all states in between $-M$ and $L$, there is equal probability of winning or losing a coin, and the transition probability $p_{ij}$ is:

$$p_{ij} = \begin{cases} 
1/2, & \text{if } -M < i < L, j = i+1 \\
1/2, & \text{if } -M < i < L, j = i-1 \\
0, & \text{otherwise}
\end{cases}$$

Task 1

Mario wants you to help him figure out his probability of winning. Answer the following questions.

1. Let $P(S_t = i)$ denote the probability that the chain is in state $i$ at time $t$. What is $\lim_{t \to \infty} P(S_t = i)$ for any transient state $i$?

2. Which (if any) of the states are transient? Briefly justify your answer.

3. Which (if any) of the states are recurrent? Briefly justify your answer.

4. Let $W_t$ be a random variable representing Mario’s gains after $t$ steps. Since the game is fair, the expectation $E[W_t]$ is 0 for all $t$. Clearly then, $\lim_{t \to \infty} E[W_t] = 0$ as well.

   Derive an expression for:

   $$\lim_{t \to \infty} E[W_t]$$

   in terms of $q_i = \lim_{t \to \infty} P(S_t = i)$. (Note: remember some of the states may be recurrent states.)

5. Solve this expression for the probability that Mario wins.

End Task 1

2 Let’s go clubbing

Mario and Luigi are running a new club called ‘Peaches’. They expect the club to draw a whole lot of customers, and want to know how long the lines of people waiting to get in are likely to be.

By Mushroom Kingdom’s restrictions, the length of the line can never exceed $n$ people long. At any time step, exactly one of the following events occur:

- If the line is not empty, then with probability $\alpha$ the person in the front of the line gets into the club.
- If the line has fewer than $n$ customers, with probability $\sigma$ a new customer joins the line.
The line remains unchanged with probability $1 - \alpha - \sigma$ if the line is neither empty nor full, probability $1 - \sigma$ if it is empty, and probability $1 - \alpha$ if it is full.

Assume $\alpha, \sigma > 0$. This can be described as follows: Let $S_t$ denote the number of customers in line at time $t$. We have $n + 1$ states (from no customers to $n$ customers) with the following transition probabilities:

$$
\begin{align*}
    p_{i,i+1} &= \sigma & \text{if } i < n \\
    p_{i,i-1} &= \alpha & \text{if } i > 0 \\
    p_{i,i} &= \begin{cases} 
        1 - \sigma & \text{if } i = 0 \\
        1 - \sigma - \alpha & \text{if } 1 \leq i \leq n - 1 \\
        1 - \alpha & \text{if } i = n
    \end{cases}
\end{align*}
$$

All other entries in the transition matrix are 0.

---

### Task 2

1. Does there exist a unique stationary distribution (i.e., steady state probabilities that are independent of the start state)? Justify your answer.

2. Derive an expression for the stationary distribution $\pi$ as a function of $\alpha$ and $\sigma$. You can think of solving for the stationary distribution as solving the recurrence out one equation at a time; through substituting you should be able to isolate $\pi_0$ (or any other variable).

3. How does the stationary distribution look when $\alpha > \sigma$? When $\alpha < \sigma$? When $\alpha = \sigma$? You may wish to sketch a few graphs to help illustrate your description.

4. Assume that $n = 5$ and the line starts with 3 people when the club first opens. Under Mario’s estimates for $\alpha = \frac{1}{4}$ and $\sigma = \frac{1}{10}$, what is the probability that the line has 2 people after 3 time steps? After 10 time steps? After 150 time steps? How do these probabilities compare to the stationary distribution? (Hint: you will want to use MATLAB or another mathematical software here.)

5. Present a 1-2 sentence argument for why looking at steady-state probabilities is useful in this setting. Then, present a 1-2 sentence argument for why steady-state probabilities are not useful for this setting (hint: what if the club is only open for 6 hours each night and each time step is 1 hour?).

---

### 3 Let’s play Monopoly

After a long night of partying at the club the previous day, Mario and Luigi decide to take it easy today by inviting Peach over to play a game of Monopoly. Bowser had tapped in on the phone call, and upon hearing this news plans to join in on the fun. Being competitive, Bowser asks you for some support on how to win at this game. He assures you that he is not intending to cause any harm, but just wants to trade in his reputation of being brawny for being brainy.

You agree to help him. After thinking about this a little bit, you realize that while luck is a large part of Monopoly, there is some optimization that can be done. In particular, you think that knowing how likely a player is to land in any given square on the board may be useful information based on which to make informed decisions about which properties to go for. You analyze the game board as shown in Figure 1.¹

¹http://www.worldofmonopoly.co.uk/history/images/bd-usa.jpg
Figure 1: Monopoly game board
You think you can model the problem as a Markov Chain. As a good applied mathematician you decide to start with the following simplifying assumptions:

- A player in Jail stays there until he or she rolls doubles (e.g., both dice show same number) or has spent three turns in Jail. The player rolling doubles will advance by the number represented by the doubles roll.
- There is no need to model a “Get out of Jail” card.
- Community Chest and Chance cards are never discarded. The stack of cards is reshuffled following every draw from the stack (e.g., with replacement).
- Community Chest and Chance cards will not portal the player to another square.
- There is no doubles go twice rule. If you roll a double, you do not get to go again (however it does impact jail behavior as aforementioned).
- You do not need to model houses. Any questions about monopolies are to get you thinking about what sorts of interpretations we might derive from the steady state distribution.

You ask yourself the following questions.

1. Give a complete description of a Markov Chain that models a player moving about the board, where at every turn a player roll a pair of fair dice. Your description should be complete but concise, e.g., by making use of appropriate mathematical notation to avoid enumerating the transition probabilities from every state. Be sure to explain how your Markov Chain deals with the jail situation, e.g., “Just Visiting”, “In Jail”, and “Go to Jail”.

2. Does the Markov Chain have a unique stationary distribution? Justify your answer.

After some work, you have found the stationary distribution. You wish to now make use of this information to come up with a good strategy for Bowser.

3. After you share your findings with Bowser, he buys St. James Place. Then to prove he is a good friend, he gives two “Get out of jail free” cards to his opponents. Explain why Bowser may not be the good friend he says he is.

4. (Optional) In thinking of possible strategies, you would like to measure the cost-effectiveness of various monopolies (i.e. getting all of one block of properties). Making use of the stationary distribution, describe a simple method for measuring this. You need not worry about second order effects such as others’ loss in capital in critical stages of the game, etc. (Hint: you should take into account the gains derived from having the property but also the cost of acquiring the property.)

5. (Optional) How may you use this cost-effectiveness measure of monopolies to formulate a (simple) strategy for making trades with other players? (Note: remember that monopolies are hard to obtain by landing on the entire block before everyone else but arbitrary trades among players are allowed.)

---

4 Let’s help find that song

Enterprising plumbers that they are, Mario and Luigi have parlayed their heroic success into launching a new product line – selling portable music players online. To their (and our) surprise, the portable music player is selling extremely well, especially among Toads. The reviews for the product have been mostly positive, but
there is one issue that keeps coming up. The playlist is sorted by song name, but there is no simple way to search and seek to a particular song (there is also no LCD screen on the device). “I was listening to the Overworld theme,” one blue Toad writes. “But I had this strong urge to hear the Underworld theme. I have 50 songs in the list between these two, and had to press next like 50 times. Maybe I should’ve gotten an iPod with a scroll wheel.”

A yellow Toad replies: “There is a shuffle mode that seeks to a random song in the playlist. Maybe there is a randomized algorithm for finding the songs we want?”

Mario lets you know about the situation and you think you can help. In particular, you think the following idea may make sense. First you listen to the current song long enough to identify it (remember, there is no lcd), and if it is close in the playlist to the desired song, then you press ‘next’ or ‘previous’ a number of times until you find the desired song. Since the playlist is circular, both ‘next’ and ‘previous’ will eventually lead to the desired song. If the current song is far away, you can go into shuffle mode and hit next until you identify a song that is close to the desired song on the playlist, at which point you switch to regular mode and then seek sequentially.

You would like to figure out when one should go into shuffle mode and when one should just seek for the desired song sequentially in regular mode. Furthermore, you like to know if this idea works well in practice, that is, how long on average does it take to find a song following this method (and given that you know the order of the songs in your playlist)? To begin answering these questions, you decide to model the problem as a Markov decision process (MDP).
Task 4

1. Model the state space of the MDP given \( N \) songs. How many states are there?

2. Model the set of actions. Be precise about what the action does and how it relates to actual operations on the music player.

3. Model the reward function \( R(s, a) \). Justify how you have assigned rewards to states.

4. Model the transition function \( P(s, a, s') \). Recall that the playlist is circular.

5. What objective criterion do you think will work well for solving this model? Do you think solutions will be different for different criterions (in your particular model)? Justify your answers.

End Task 4

5 Let’s sell lemonade

Bowser envies Mario and Luigi’s entrepreneurial success and decides to start his own business, forcing his Koopa Troopas to run a lemonade stand. There are two types of customers in the world, those who will pay $3 for lemonade and those who will pay $1.50 for water (though lemonade customers will buy water if that is the only option). The morning of each day, Bowser will choose how many bottles of water to buy at a cost of $1 each, how many bottles of lemonade to buy at a cost of $2 each (did you really expect the lemonade to be homemade?), and how much to advertise (one $0.50 advertisement draws one lemonade customer, but there are only 10 advertisements available). Each day Bowser is on a budget of $55.

Bowser’s only problem is the weather. Sunny days in the mushroom kingdom mean there are 20 water customers and 10 lemonade customers. Rainy days mean there are only 10 water customers and 5 lemonade customers. Hot days mean there are 20 water customers and 15 lemonade customers. Bowser does not have a meteorologist, but he knows that it is only sunny 50% of the time, with 25% of the time being rainy and 25% of the time being hot.

You want to investigate Bowser’s purchasing problem. Your first step is to review the deterministic AMPL model we already put together in lemonade.mod and lemonade.dat.

Task 5

1. By editing lemonade.dat and running the deterministic model using lemonade.run, determine the ideal number of water bottles, lemonade bottles, and advertisements that Bowser should buy on a sunny day. Also list the profit this would result in. Do the same for rainy days and hot days.

2. lemonade_test.mod, lemonade_test.dat, and lemonade_test.run provide a quick way to see what would happen if demand were fixed and Bowser bought a certain number of water bottles, lemonade bottles, and advertisements (i.e. it immediately calculates how many of each he would sell and his profit). Use this and your answers from above to create a chart showing all 9 combinations of possibilities (e.g. Bowser prepares for a sunny day, but it is actually a rainy day. Or Bowser prepares for a hot day but it is actually sunny). Be sure to list how much he purchases of each good, how much he sells of each good, and his resulting profit.

3. For each of the 3 preparation options, calculate the expected profit given the variability of the weather. (e.g. If Bowser decides to prepare for a sunny day, what is his expected profit?)

4. If Bowser prepares for the expected demand, how much of each good does he buy? What is his expected profit?

\[\text{note: you may need to edit the option solver './cplex'; line depending on your system}\]
5. You decide to use stochastic optimization to solve Bowser’s problem. Edit `lemonade.mod` and `lemonade.dat` in the appropriate way. Paste the new files as part of your answer. What is the optimal stochastic solution? List how much of each good Bowser would buy and his expected profit.

6. In one sentence explain how to calculate the Expected Value of the Stochastic Solution (VSS). Then calculate the VSS here. In one more sentence give a plain English explanation of what the VSS means in Bowser’s case.

7. In one sentence explain how to calculate the Expected Value of Perfect Information (EVPI). Then calculate the EVPI here. In one more sentence give a plain English explanation of what the EVPI means in Bowser’s case.

-------------------------------------------
End Task 5

-------------------------------------------

6. Let’s turn in the assignment

Final Task 6

You must turn in your assignment by 5 PM, Monday, November 30, 2015. You may use 1 late day. For AMPL exercises, this includes any model and data files you have created (include the text of the files in the writeup please!). When writing down the solution from AMPL, always include both the objective value and the values assigned to variables (when the program is feasible and bounded, of course!)

Gather the AMPL model and data files you have created for this assignment, as well as a script file containing the AMPL commands you used to solve the problems. Upload your files to the course isite dropbox before 5 PM, Monday, November 30, 2015. Please scan and submit your homework if you completed it on paper. If you feel it may be hard to read, then you better re-write or re-scan everything so that we can give you credit for your work.

End Task 6

Congratulations on completing your final AM121 assignment!