Assignment 7

Bodhi Nguyen

February 20th, 2021

1 Problem 3: Worker

Let the state space be a tuple representing whether you have a job and also your skill level. The first part of the tuple is a boolean representing whether you have a job. The second part of the tuple is continuous representing the current skill level $sk \in (0, \infty)$. The action set is also continuous in that you choose a fraction $\alpha \in (0, 1)$ that represents the fraction of the day spent working. Let the total number of minutes in a work day be T. Assume that on each night, you lose your job with probability p, and if you don't have a job in the morning, you are offered a job with probability $h(s)$. Then, the reward on each day given that you have a job is

$$R(s = (true, sk), a = \alpha) = f(sk)\alpha T$$

.

and the reward given that you don't have a job is just 0. The transition probabilities are dependent on whether you have a job or not, which can take a couple cases.

1. Case 1: You have a job today and have a job the next day with probability $(1 - p) + p * h(sk)$
2. Case 2: You have a job and lose it with no job the next day: $p(1 - h(sk))$
3. Case 3: You don't have a job that day and get one the next day with probability $h(sk)$
4. Case 4: You don't have a job and don't get one with probability $1 - h(sk)$

Assuming g() is deterministic then the skill level transitions are also deterministic. So, let sk_t be the skill level today (t in minutes). Then if you have a job, for the next day, $sk_{t+T} = g(sk)(1 - \alpha)Tsk_t$ and if you don't have a job that day, $sk_{t+T} = sk_t e^{-\lambda T}$.

$$P(s_{t+T} = (true, g(sk_t)(1 - \alpha)Tsk_t)|s_t = (true, sk_t), a = \alpha) = (1 - p) + p * h(sk_t) \tag{1}$$

$$P(s_{t+T} = (false, g(sk_t)(1 - \alpha)Tsk_t)|s_t = (true, sk_t), a = \alpha) = p(1 - h(sk)) \tag{2}$$

$$P(s_{t+T} = (true, sk_t e^{-\lambda T})|s_t = (false, sk_t), a = \alpha) = h(sk) \tag{3}$$

$$P(s_{t+T} = (false, sk_t e^{-\lambda T})|s_t = (false, sk_t), a = \alpha) = (1 - h(sk) \tag{4}$$

The optimal policy should look something like a threshold where you need to spend lots of time learning to raise skill level first, to dampen the effect of losing your job and to get higher wages, but then after a certain skill level you wouldn't spend any time at all on learning. With multiple skills, the learning might have to be spread out more, and with multiple job options it probably pays less to learn since you're less at risk of losing your job.