CME241 Assignment7

Quinn Hollister

February 2022

1 Problem 1

Derive the solution to Merton's Portfolio problem for the case of the log(*) Utility function

First we can model our wealth process as

$$W_t = ((\pi_t(\mu - r) + r) \cdot W_t - c_t) * dt + \pi_t \sigma W_t dz_t \tag{1}$$

We want to maximize the expected discounted utility of wealth, so our optimal value function is given by

$$V^*(t, W_t) = \max_{\pi_t, c_t} E[\int_t^T \exp -\rho * (s-t) \log(c_s) \cdot ds + \exp(-\rho(T-t)) \cdot \epsilon^\gamma \cdot \log W_T] \tag{2}$$

$$\exp \rho(t_1 - t) * V^*(t, W_t) = \max_{\pi_t, c_t} E[\int_t^{t_1} \exp -\rho * (s-t) \log c_s \cdot ds + \exp -\rho(t_1) \cdot V^*(t_1, W_{t_1}) \tag{3}$$

$$\max_{\pi_t, c_t} E_t[d(\exp(-\rho t) V^*(t, W_t)) + \exp -\rho t \log c_t] = 0 \tag{4}$$

$$\max_{\pi_t, c_t} E_t[dV^*(t, W_t) + \log c_t \dot{d}t] = \rho V^*(t, W_t) \cdot dt \tag{5}$$

We can use the HJB optimality equations to rewrite this as a PDE:

$$\max_{\pi_t, c_t}\{\frac{\partial V^*}{\partial t} + \frac{\partial V^*}{\partial W_t}((\pi_t(\mu - r) + r) \cdot W_t - c_t) + \frac{\partial^2 V^*}{\partial W_t^2} \cdot \frac{\pi_t^2 \sigma^2 \cdot W_t^2}{2} + \log c_t\} = \\ \rho \cdot V^*(t, W_t)$$

with the terminal condition that:

$$V^*(T, W_T) = \epsilon^\gamma \cdot \log W_T \tag{6}$$

To find the optimal actions, we need to take the partials of $\Phi(t, W_t; \pi_t, c_t)$, which is the function inside the max argument, and set them to zero:

$$\frac{\partial \Phi}{\partial \pi_t} = (\mu - r) \cdot W_t \cdot \frac{\partial V^*}{\partial W_t} + 2 \cdot \frac{\partial^2 V^*}{\partial W_t^2} \cdot \frac{\pi_t \sigma^2 W_t^2}{2},$$
$$\frac{\partial \Phi}{\partial c_t} = -\frac{\partial V^*}{\partial W_t} + \frac{1}{c_t}$$

Thus, we can see that the optimal values for phi_t, c_t are given by the equations

$$c_t^* = (\frac{\partial V^*}{\partial W_t})^{-1},$$
$$\pi_t^* = \frac{-\frac{\partial V^*}{\partial W_t} \cdot (\mu - r)}{\frac{\partial^2 V^*}{\partial W_t^2} \cdot \sigma^2 \cdot W_t}$$

Now, replacing these optimal values into equation 6, we get the following PDE:

$$\frac{\partial V^*}{\partial t} = -\frac{(\mu - r)^2}{2 \cdot \sigma^2} \cdot \frac{(\frac{\partial V^*}{\partial W_t})^2}{\frac{\partial^2 V^*}{\partial W_t^2}} + \frac{\partial V^*}{\partial W_t} \cdot r \cdot W_t + \log{(\frac{\partial V^*}{\partial W_t})^{-1}} - 1 = \rho \cdot V^* \tag{7}$$

But, we also know that $\gamma = 1$, and our educated guess (based on our B.C. $V^*(T, W_T) = \epsilon \cdot \log W_T$) is:

$$V^*(t, W_t) = f(t) + log(W_t) \tag{8}$$

which has the following partial derivatives:

$$\begin{aligned}\frac{\partial V^*}{\partial t} &= f^{'}(t) \\ \frac{\partial V^*}{\partial W_t} &= \frac{1}{W_t} \\ \frac{\partial^2 V^*}{\partial W_t^2} &= -\frac{1}{W_t^2}\end{aligned}$$

So, our PDE resolves to an ODE of the following form:

$$f^{'}(t) = \frac{(\mu - r)^2}{2 \cdot \sigma^2} * \frac{W_t^{-2}}{-W_t^{-2}} + \frac{1}{W_t} \cdot r \cdot W_t + \log W_t = \rho \cdot (f(t) + \log W_t) \tag{9}$$

$$\begin{aligned}f^{'}(t) &= \nu + \rho \cdot f(t), \\ \nu &= \frac{(\mu - r)^2}{2 \cdot \sigma^2} - r + \log W_t(\rho - 1)\end{aligned}$$

This is a simple ODE, with the following solution:

$$f(t) = \frac{-\nu + (\rho \cdot \epsilon - \nu) \cdot \exp{(-\rho(T - t))}}{\rho} \tag{10}$$

Now, we can look back at our equations for our optimal actions and replace the symbolic partial derivatives with values, and we get the following forms

$$\begin{aligned}\pi_t^* =& \frac{(\mu - r)}{\sigma^2}, \\ c_t^* =& W_t\end{aligned}$$

and finally taking $\epsilon = 0$, we get the optimal value function:

$$V^* = \frac{-\nu(1 + \exp{(-\rho \cdot (T - t))})}{\rho} + \log W_t \tag{11}$$

2 Problem 2

Sketch out a rough design for the following MDP

You can model your life as taking an action α each day which states that you spend a fraction α working each day and a fraction $1 - \alpha$ of your day learning a new skill. Each minute that you spend working results in a reward rate of $f(s)$ dollars per minute where s is your current skill level. Each minute you spend learning improves your skill level by a certain factor $g(s)$ where s is your current skill level. You could lose your job any day with probability p. While unemployed you do not have access to learning, so you skill level decays exponentially with a half life λ. If you lose your job, you can be offered your job back with probability $h(s)$ where s is your current skill level.

We can easily see that our action space is defined as such: $\mathcal{A} = \alpha \in [0, 1]$. The state space seems to be a function of your employment, your wealth, and your skill, i.e. $\mathcal{S} = (U, E, W_t, S_t)$ where U, E indicates your employment state, W_t indicates your wealth, or accumulated earnings if we don't consider consumption, and S_t is your current skill level. Our transition probabilities are dictate by employment. If we're employed, then with probability p we can become unemployed. All other transitions are dictated by your actions, and thus not influenced by uncertainty. But, when you're unemployed, the probability that you will become employed is some function $h(s)$ which is obviously dependent on your current skill level, and if you remain unemployed, then your skill level will deterministically decrease by $e^{-\lambda}$ with each day. Notice that your action set is the zero set when you are unemployed.

$$\begin{aligned}
\mathcal{P}((U, W_t, S_t), (U, W_t, e^{-\lambda} \cdot S_t)) &= 1 - h(s) \\
\mathcal{P}((U, W_t, S_t), (E, W_t, S_t)) &= h(s) \\
\mathcal{P}((E, W_t, S_t), (E, W_t + \alpha * f(s), S_t + \alpha * g(s))) &= 1 - p \\
\mathcal{P}((E, W_t, S_t), (U, W_t, S_t)) &= p
\end{aligned}$$

Just for completeness, we also know that our rewards function is strictly determined by our action while we're employed (since we our rewards are zero if we're employed) and the probability that we'll fired that day,

$$R((E, W_t, S_t), \alpha, (E, W_{t+1}, S_{t+1})) = \alpha \cdot g(S_t) \tag{12}$$

It appears that the optimal strategy heavily favors increasing skill as quickly as possible early in life. The reason for this is threefold: your pay is determined by your skill level, and depending on the degree of curvature of the function, higher skill levels result in higher efficiency in accumulating wealth. Also, if you're fired then the probability that you will be rehired depends on your skill level (and the probability that you will be fired has no dependency on skill, or how much time you put into accumulating skill), thus it seems that increasing skill is paramount to staying employed and having the least amount of unemployed days (which are critical because you can't accumulate wealth, and you LOSE skill). Lastly, your skills decay exponentially while you're unemployed, so every extra day you remain unemployed REALLY hurts your current state. A rough estimate is that you want to heavily spend time developing skills when you're first starting out (low α), and only transitioning to working when the marginal increases in skills starts to matter less (will be dependent on curvature of $f(s)$, $g(s)$, $h(s)$. Thus, your work will now be much more efficient due to commanding higher labor rates. Now, if we factor consumption into the MDP, things get more interesting, especially if we have a utility function that could depend on t (i.e. discounted), i.e. we get more out of consumption when we're young because we can "enjoy" that consumption more than when we are older (and the time value of money). This might invalidate some of the optimal policy fleshed out previously, as it would put a higher weight on early wealth accumulation that could have a high utility. I think the biggest perturbation you could make to this problem set-up, would be to make the probability of losing your job depend on your action α, since obviously this would mean you would be spending less time on work, and thus not performing to such a high standard for that given day. This would obviously shift the optimal policy to a more equally weighted decision between work and learning, and would probably reflect the real world with a higher fidelity.