

EE365 Homework 5

1. *A refined inventory model.* In this problem we consider an inventory model that is more refined than the one you've seen in the lectures. The amount of inventory at time t is denoted by $q_t \in \{0, 1, \dots, C\}$, where $C > 0$ is the maximum capacity. New stock ordered at time t is $u_t \in \{0, 1, \dots, C - q_t\}$. This new stock is added to the inventory instantaneously. The demand that arrives after the new inventory, before the next time period, is $d_t \in \{0, 1, \dots, D\}$. The dynamics are

$$q_{t+1} = (q_t + u_t - d_t)_+.$$

The unmet demand is $(q_t + u_t - d_t)_- = \max(-q_t - u_t + d_t, 0)$. The demands d_0, d_1, \dots are IID.

We now describe the stage cost, which does not depend on t , as a sum of terms.

The *ordering cost* is

$$g^{\text{order}}(u) = \begin{cases} 0 & u = 0 \\ p^{\text{fixed}} + p^{\text{whole}}u & 1 \leq u \leq u^{\text{disc}} \\ p^{\text{fixed}} + p^{\text{whole}}u^{\text{disc}} + p^{\text{disc}}(u - u^{\text{disc}}) & u > u^{\text{disc}} \end{cases}$$

where p^{fixed} is the fixed ordering price, p^{whole} is the wholesale price for ordering between one and u^{disc} units, and $p^{\text{disc}} < p^{\text{whole}}$ is the discount price for any amount ordered above u^{disc} .

The *storage cost* is $g^{\text{store}}(q) = s^{\text{lin}}q + s^{\text{quad}}q^2$, where s^{lin} and s^{quad} are positive.

The *negative revenue* is $g^{\text{rev}}(q, u, d) = -p^{\text{rev}} \min(q + u, d)$, where $p^{\text{rev}} > 0$ is the retail price.

The *cost for unmet demand* is $g^{\text{unmet}}(q, u, d) = p^{\text{unmet}}(q + u - d)_-$, where $p^{\text{unmet}} > 0$.

The terminal cost is a *salvage cost*, which is $g^{\text{sal}}(q) = -p^{\text{sal}}q$ where $p^{\text{sal}} > 0$ is the salvage price.

We now consider a specific problem with

$$T = 50, \quad C = 20, \quad D = 4, \quad q_0 = 10,$$

and demand distribution

$$\mathbf{Prob}(d_t = 0, 1, \dots, 4) = (0.2, 0.25, 0.25, 0.2, 0.1).$$

The cost function parameters are

$$\begin{aligned} p^{\text{fixed}} &= 4, & p^{\text{whole}} &= 2, & p^{\text{disc}} &= 1.6, & u^{\text{disc}} &= 6, & s^{\text{lin}} &= 0.1, \\ s^{\text{quad}} &= 0.05, & p^{\text{rev}} &= 3, & p^{\text{unmet}} &= 3, & p^{\text{sal}} &= 1.5 \end{aligned}$$

- (a) Solve the MDP and report J^* .

- (b) Plot the optimal policy for several interesting values of t , and describe what you see. Does the optimal policy converge as t goes to zero? If so, give the steady state optimal policy.
 - (c) Plot $\mathbf{E} g^{\text{order}}(u_t)$, $\mathbf{E} g^{\text{store}}(q_t)$, $\mathbf{E} g^{\text{rev}}(q_t, u_t, d_t)$, $\mathbf{E} g^{\text{unmet}}(q_t, u_t, d_t)$, and $\mathbf{E} g^{\text{sal}}(q_t)$, versus t , all on the same plot.
2. *Exiting the market with incomplete fulfillment.* You have S identical stocks that you can sell in one of T periods. The price fluctuates randomly; we model this as the prices being IID from a known distribution. At each period, you are told the price, and then decide whether to sell one stock or wait. However, the market is not necessarily liquid, so if you decide to sell a stock, there is some chance that no buyer is willing to buy your stock. This random event is modeled as a Bernoulli random variable with probability 0.6 that no buyer is willing to buy your stock, and a probability of 0.4 that you sell it. Note that you can decide to sell a maximum of one stock per period. Your goal is to maximize the expected revenue from the sales.
- (a) Model this problem as a Markov decision problem that you can use to find the optimal selling policy: that is, give the set of states \mathcal{X} , the set of actions \mathcal{U} , and the set of disturbances \mathcal{W} . Give the dynamics function f , and reward function g . There is no terminal reward. What is the information pattern of the policy?
 - (b) The stock prices are independent random variables following a discretized log-normal distribution with the log-prices having mean 0 and standard deviation 0.2. In particular, the prices take 15 values, ranging from 0.6 to 2.0 in increments of 0.1, and the probability mass function of p is proportional to the probability density function of the stated log-normal distribution. Compute the optimal policy and the optimal expected revenue for $T = 50$ periods and $S = 10$. Plot value functions at times $t = 0, 45, 48, 49, 50$, and policies at times $t = 0, 20, 40, 45$. What do you observe?
 - (c) Let us define a threshold policy in which you decide to sell only when the stock's price is greater than the expected value of the prices. With the same parameters as in question (b), compute the expected revenue of this threshold policy. What can you notice in comparison to question (b)?
 - (d) For the optimal policy and the threshold policy described in (c), compute the probability that we have unsold stocks after the final time T .
 - (e) Find a policy for which the probability of having unsold stock after the final time T is less than 0.005. Compute the expected revenue of this policy, and compare it to the expected revenues of the two previous policies you calculated in this question.
3. *Appliance scheduling with fluctuating real-time prices.* An appliance has C cycles, $c = 1, \dots, C$, that must be run, in order, in $T \geq C$ time periods, $t = 0, \dots, T - 1$. A schedule consists of a sequence $0 \leq t_1 < \dots < t_C \leq T - 1$, where t_c is the time period in which cycle c is run. Each cycle c uses a (known) amount of energy $e_c > 0$, $c = 1, \dots, C$, and, in each period t , there is an energy price p_t . The total energy

cost is then $J = \sum_{c=1}^C e_c p_{t_c}$. In the lecture on deterministic finite-state control, we considered an example of this type of problem, where the prices are known ahead of time. Here, however, we assume that the prices are independent log-normal random variables, with known means, \bar{p}_t , and variances, σ_t^2 , $t = 0, \dots, T-1$. You can think of \bar{p}_t as the predicted energy price (say, from historical data), and p_t as the actual realized real-time energy price.

The following questions pertain to the specific problem instance defined in `appliance_sched_data.m`.

- (a) *Minimum mean cost schedule.* Find the schedule that minimizes $\mathbf{E} J$. Give the optimal value of $\mathbf{E} J$, and show a histogram of J (using Monte Carlo simulation). Here you do not know the real-time prices; you only know their distributions.
- (b) *Optimal policy with real-time prices.* Now suppose that right before each time period t , you are told the real-time price p_t , and then you can choose whether or not to run the next cycle in time period t . (If you have already run all cycles, there is nothing you can do.) Find the optimal policy, μ^* . Find the optimal value of $\mathbf{E} J$, and compare it to the value found in part (a). Give a histogram of J .

You may use Monte Carlo (or simple numerical integration) to evaluate any integrals that appear in your calculations. For simulations, the following facts will be helpful: If $z \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$, then $w = \exp z$ is log-normal with mean μ and variance σ^2 given by

$$\mu = e^{\tilde{\mu} + \tilde{\sigma}^2/2}, \quad \sigma^2 = (e^{\tilde{\sigma}^2} - 1) e^{2\tilde{\mu} + \tilde{\sigma}^2}.$$

We can solve these equations for

$$\tilde{\mu} = \log \left(\frac{\mu^2}{\sqrt{\mu^2 + \sigma^2}} \right), \quad \tilde{\sigma}^2 = \log(1 + \sigma^2/\mu^2).$$