Barcelona GSE - Stochastic Models and Optimization

Problem Set 3 - Solutions

Problem 1 (Inventory Control with Forecasts - [B05] Exercise 4.3)

Consider an inventory control problem with no fixed ordering cost. At the beginning of each period k the inventory manager, in addition to knowing the current inventory level x_k, receives an accurate forecast that the demand w_k will be selected in accordance with one out of two probability distributions P_l, P_s (large demand, small demand). The a priori probability of a large demand forecast, q, is also known to the manager.

(a) Obtain the optimal ordering policy for the standard Newsvendor problem (i.e., single period and accounting for inventory holding and backorder costs, but *not* for variable ordering cost). Is it an open-loop policy like in the no-forecast case, or a closed-loop one?

(b) Obtain the optimal inventory replenishment policy for the standard multi-period problem (accounting for variable ordering, inventory holding, and backorder costs).

Solution: *Part (a).* In the standard Newsvendor problem the manager seeks to solve the optimization problem:

$$\min_q \mathbb{E}\big\{h(q-D)^+ + b(D-q)^+\big\}.$$

In this setting, though, the distribution of the demand depends on the forecast $y \in \{l, s\}$. In any case, as derived in the previous problem set, the optimal order q^* has to satisfy:

$$F_y(q^*) = \frac{b}{b+h}.$$

Since the manager knows from which distribution D will be drawn, he can solve for the optimal policy:

$$q^*(y) = \begin{cases} F_l^{-1}\left(\frac{b}{b+h}\right) & \text{if } y = l \\ F_s^{-1}\left(\frac{b}{b+h}\right) & \text{if } y = s. \end{cases}$$

As the policy is a function of the forecast/state, it is a closed-loop policy.

Part (b). The augmented system dynamics are given by the equations:

$$x_{k+1} = x_k + u_k - w_k,$$

and

$$y_{k+1} = \xi_k,$$

where ξ_k takes on values 1 (large demand) and 2 (small demand) with probability q and $1-q$, respectively. The DP algorithm is given by:

$$J_N(x_N, y_N) = 0,$$

and

$$
\begin{aligned}
J_k(x_k, y_k) = \min_{u_k \geq 0} \mathbb{E}_{w_k} \big\{ & cu_k + b \max(0, w_k - x_k - u_k) + h \max(0, x_k + u_k - w_k) \\
& + qJ_{k+1}(x_k + u_k - w_k, 1) + (1 - q) J_{k+1}(x_k + u_k - w_k, 2) \mid y_k \big\}.
\end{aligned}
$$

Since the functions $J_k(x_k, 1)$ and $J_k(x_k, 2)$ can be easily seen to be convex, the optimal policy is:

$$
\mu_k^*(x_k, y_k) = \begin{cases} S_k^i - x_k, & \text{if } y_k = i,\ x_k < S_k^i \\ 0, & \text{otherwise}, \end{cases}
$$

where S_k^i minimizes the convex function:

$$
cz + \mathbb{E}_w \big\{ b \max(0, w - z) + h \max(0, z - w) + qJ_{k+1}(z - w, 1) + (1 - q) J_{k+1}(z - w, 2) \mid y_k = i \big\}.
$$

Problem 2 (Inventory Pooling)

Consider the single period multi-location Newsvendor model: n different locations face independent and Normally distributed demands with mean μ and variance σ^2. The goal is to cover these demands with the minimum expected cost. The inventory holding cost and backorder cost parameters, h and b respectively, are the same in every location.

One approach is to cover each demand from an individual inventory repository at each location; we call this the *decentralized system*. Clearly, this is equivalent to having n standard (single period, single location) Newsvendor problems. Thus, the optimal order quantity at each location is Q^* and the optimal expected cost of the whole system equal to $nG(Q^*)$.

An alternative strategy is to satisfy all demands from a central inventory repository; we call this the *pooled system*. Let Q_p^* be the optimal order quantity in the pooled system and $G_p(Q_p^*)$ the optimal expected cost.

Show that $nG(Q^*)/G_p(Q_p^*) = \sqrt{n}$. How do you interpret this result?

Solution: Similarly to the solution of the standard Newsvendor model, the optimal order quantity in the pooled system is

$$Q_p^* = \min\left\{Q : \mathbb{P}\Big(\sum_{i=1}^{n} D_i < Q\Big) \geq \frac{b}{b+h}\right\}.$$

Using the property that $\sum_{i=1}^{n} D_i = \sqrt{n}D_1 + \mu(n-\sqrt{n})$ and defining the variable $x = \sqrt{n}[\,Q - \mu(n+\sqrt{n})]$, the optimal order quantity can be shown to be equal to

$$\begin{aligned} Q_p^* &= \min\left\{\sqrt{n}D_1 + \mu(n+\sqrt{n}) : \mathbb{P}(D_1 < x) \geq \frac{b}{b+h}\right\} \\ &= \sqrt{n}\min\left\{x : \mathbb{P}(D_1 \leq x) \geq \frac{b}{b+h}\right\} + \mu(n-\sqrt{n}) \\ &= \sqrt{n}Q^* + \mu(n-\sqrt{n}). \end{aligned}$$

Now we can use this result to calculate the cost associated to the pooled model:

$$\begin{aligned} G_p(Q_p^*) &= b\mathbb{E}_{\{D_i\}}\Big[\Big(\sum_{i=1}^{n} D_i - Q_p^*\Big)^+\Big] + h\mathbb{E}_{\{D_i\}}\Big[\Big(Q_p^* - \sum_{i=1}^{n} D_i\Big)^+\Big] \\ &= b\mathbb{E}\big[\,(\sqrt{n}D_1 + \mu(n-\sqrt{n}) - \sqrt{n}Q^* + \mu(n-\sqrt{n})^+\big] \\ &+ h\mathbb{E}\big[\,(\sqrt{n}Q^* + \mu(n-\sqrt{n}) - \sqrt{n}D_1 - \mu(n-\sqrt{n}))^+\big] \\ &= b\sqrt{n}\mathbb{E}[\,(D_1 - Q^*)^+] + h\sqrt{n}\mathbb{E}[\,(Q^* - D_1)^+] \\ &= \sqrt{n}G(Q^*). \end{aligned}$$

We can now compare the expected costs of the two approaches:

$$\frac{nG(Q^*)}{\sqrt{n}G(Q^*)} = \sqrt{n}.$$

Hence, by pooling the demands from different locations their stochastic fluctuations around the mean cancel out to some extent, and it is possible to achieve (statistical) economies of scale. The above result determines exactly how much we gain by doing so.

Problem 3 (Asset Selling with Maintenance Cost - [B05] Exercise 4.16)

Suppose that a person wants to sell a house and an offer comes at the beginning of each day. We assume that successive offers are independent and an offer is w_j with probability $p_j,\ j = 1, \dots, n$, where w_j are given nonnegative scalars. Any offer not immediately accepted is not lost but may be accepted at any later date. Also, a maintenance cost c is incurred for each day that the house remains unsold. The objective is to maximize the seller's profit when there is a deadline to sell the house within N days. Characterize the optimal policy.

Solution: This is variation of the example of asset selling with retained offers, without interest rate but with maintenance cost. The dynamics would have the same form, i.e.,

$$x_{k+1} = \max(x_k, w_k),$$

where x_k is the best offer ever received and w_k is the offer received at period k. The DP formulation of this optimal stopping problem is as follows:

$$J_N(x_N) = x_N,$$

and

$$J_k(x_k) = \max\left[x_k, -c + \sum_{j=1}^{n} p_j J_{k+1}\big(\max(x_k, w_j)\big)\right],$$

where c is the daily maintenance cost that is incurred if the property is not sold.

Notice that

$$J_{N-1}(x) = \max\left[x, -c + \sum_{j=1}^{n} p_j \max(x, w_j)\right],$$

which implies that the one-step stopping set is

$$T_{N-1} = \left\{x \;\middle|\; x \geq -c + \sum_{j=1}^{n} p_j \max(x, w_j)\right\},$$

or equivalently,

$$T_{N-1} = \big\{x \mid c \geq h(x)\big\},$$

where

$$h(x) = \sum_{j=1}^{n} p_j \max(x, w_j) - x.$$

Since past offers are retained and $h(\cdot)$ is a nonincreasing function, the one-step stopping set is absorbing. Consequently, the stopping sets of different stages are the same.

Finally, we show that the optimal policy is a threshold policy. Since $h(\cdot)$ is nonincreasing, $h(x) \leq c$ implies that $h(y) \leq c$, for all $y \geq x$. Function $h(\cdot)$ is also continuous, so let x^* such that $h(x^*) = c$. The one-step stopping set can be equivalently written as follows:

$$T_{N-1} = \big\{x \mid x \geq x^*\big\}.$$

Summarizing, the optimal policy is to sell the asset to the first offer that exceeds x^*.

Problem 4 (An Investment Problem - [B05] Exercise 4.13)

An investor has the opportunity to make N sequential investments: at time k he may invest any amount $u_k \geq 0$ that does not exceed his current wealth x_k (defined to be his initial wealth, x_0, plus his gain or minus his loss thus far). He wins his investment back and as much more with probability p, where $1/2 < p < 1$, and he loses his investment with probability $(1-p)$. Find the optimal investment strategy.

Solution: Let x_k be the investor's wealth at the beginning of period k, and u_k the investment he makes during period k, with $0 \leq u_k \leq x_k$. Consider also the random variable w_k, which takes the value -1 if the investment at period k is lost, and +1 otherwise. We have that $\mathbb{P}(w_k = 1) = 1 - \mathbb{P}(w_k = -1) = p$.

The system evolves according to the dynamic equation: $x_{k+1} = x_k + u_k w_k$. The investor wishes to maximize his (expected) wealth at the end of the investment period, i.e., $g_N(x_N) = \ln(x_N)$ and $g_k(x_k, u_k, w_k) = 0$, for $k < N$.

The optimal investment strategy follows from showing that $J_k(x_k) = A_k + \ln(x_k)$. We prove this using mathematical induction.

Basis: for $k = N$ the result holds with $A_k = 0$, since $J_N(x_N) = g_N(x_N)$.

Inductive step: assume that $J_{k+1}(x_{k+1}) = A_{k+1} + \ln(x_{k+1})$. We have that

$$J_k(x_k) = \max_{u_k \in U_k(x_k)} \mathbb{E}_{w_k}\left[J_{k+1}(x_k + u_k w_k)\right] = \max_{0 \leq u_k \leq x_k} \left\{A_{k+1} + p\ln(x_k + u_k) + (1-p)\ln(x_k - u_k)\right\}.$$

The first-order optimality condition implies that

$$u_k^* = \mu_k^*(x_k) = (2p-1)x_k.$$

The fact that $1/2 < p < 1$ implies that u_k^* is in the interior of the constraint set. It can be easily verified that an optimal solution cannot be on the boundaries of the constraint set, so the Extreme Value Theorem implies that the above stationary point constitutes the unique optimal solution. Substituting u_k^* back to $J_k(x_k)$, we get:

$$J_k(x_k) = A_{k+1} + \ln\left(2p^p(1-p)^{1-p}\right) + \ln(x_k),$$

and the result follows.

Problem 5 (A Scheduling Problem - [B05] Exercise 4.28)

Consider a quiz contest where a person is given a list of N questions and can answer these questions in any order she chooses. Question i will be answered correctly with probability p_i, and the contestant will then receive a reward $R_i > 0$; if the question is not answered correctly then the quiz terminates and the contestant is allowed to keep her previous earnings minus a penalty $F_i \geq 0$. The problem is to choose the ordering of questions so as to maximize expected rewards.

(a) Use an interchange argument to show that it is optimal to answer the questions in order of decreasing $\big(p_i R_i - (1-p_i)F_i\big)/(1-p_i)$.

(b) Solve a variant of the problem where there is a no-cost option to stop answering questions.

Solution: *Part (a).* Using an interchange argument similar to the one in Example 4.5.1 of [B05], we assume that i and j are the k-th and $(k+1)$-st questions in an optimally ordered list L:

$$L = (i_0, \ldots, i_{k-1}, i, j, i_{k+2}, \ldots, i_{N-1}).$$

We consider the list L' with questions i and j interchanged:

$$L' = (i_0, \ldots, i_{k-1}, j, i, i_{k+2}, \ldots, i_{N-1}).$$

Then, we can compare the expected rewards associated to the two lists:

$$\begin{aligned}\mathbb{E}\{\text{reward of } L\} &= \mathbb{E}\big\{\text{reward of } \{i_0, \ldots, i_{k-1}\}\big\} \\ &\quad + p_{i_0}, \ldots, p_{i_{k-1}}(p_i R_i - (1-p_i)F_i + p_i p_j R_j - p_i(1-p_j)F_j) \\ &\quad + p_{i_0}, \ldots, p_{i_{k-1}} p_i p_j \mathbb{E}\big\{\text{reward of } \{i_{k+2}, \ldots, i_{N-1}\}\big\},\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\{\text{reward of } L'\} &= \mathbb{E}\big\{\text{reward of } \{i_0, \ldots, i_{k-1}\}\big\} \\ &\quad + p_{i_0}, \ldots, p_{i_{k-1}}(p_j R_j - (1-p_j)F_j + p_j p_i R_i - p_j(1-p_i)F_i) \\ &\quad + p_{i_0}, \ldots, p_{i_{k-1}} p_i p_j \mathbb{E}\big\{\text{reward of } \{i_{k+2}, \ldots, i_{N-1}\}\big\}.\end{aligned}$$

We know that L is optimally ordered, so it must be

$$\mathbb{E}\{\text{reward of } L\} \geq \mathbb{E}\{\text{reward of } L'\},$$

which implies that

$$p_i R_i - (1-p_i)F_i + p_i p_j R_j - p_i(1-p_j)F_j \geq p_j R_j - (1-p_j)F_j + p_j p_i R_i - p_j(1-p_i)F_i.$$

With some manipulation, from the above we get:

$$\frac{p_i R_i - (1-p_i)F_i}{1-p_i} \geq \frac{p_j R_j - (1-p_j)F_j}{1-p_j}.$$

Thus, it is optimal to answer the questions in order of decreasing $\big(p_i R_i - (1-p_i)F_i\big)/(1-p_i)$.

Part (b). Clearly, it is optimal for the participant to answer the questions in the order identified in Part (a), and exercise the no-cost option to stop as soon as the expected reward becomes less than or equal to the expected loss, i.e.,

$$p_i R_i \leq (1-p_i)F_i.$$