14D006 Stochastic Models and Optimization

Problemset 3

Daniel Bestard, Michael Cameron, Hans-Peter Höllwirth, Akhil Lohia

March 9, 2017

1 Inventory Control with Forecasts

2 Inventory Pooling

We want to compare the optimal expected cost between a decentralized and a pooled inventory system with n locations. Let the demand at each location i be independently, normally distributed with mean μ and variance σ^2, i.e. $D_1, ...D_N \overset{iid}{\sim} N(\mu, \sigma^2)$. Then $\sum_{i=1}^{N} D_i \overset{d}{=} \sqrt{n}D_1 + \mu(n - \sqrt{n})$.

Let h and b denote the inventory holding cost and backorder cost, respectively. We already know that if $G(Q^*) = \mathbb{E}_D[h(Q^* - D)_+ + b(D - Q^*)_+]$ denotes the optimal expected cost for each location (with optimal order quantity $Q^* = \inf \left\{ Q \geq 0 : \mathbb{P}\left(D_1 \leq Q\right) \geq \frac{b}{b+h} \right\}$), then the optimal expected cost in the decentralized inventory system is

$$G_D(Q^*) = \sum_{i=1}^{N} G(Q^*) = nG(Q^*)$$

To find the optimal expected cost in the pooled inventory system $G_P(Q_P^*)$, we first need to determine the optimal order quantity Q_P^* under this system.

$$
\begin{aligned}
Q_P^* &= \inf\left\{Q \geq 0 : \mathbb{P}\left(\sum_{i=1}^{N} D_i \leq Q\right) \geq \frac{b}{b+h}\right\} \\
&= \inf\left\{Q \geq 0 : \mathbb{P}\left(\sqrt{n}D_1 + \mu(n-\sqrt{n}) \leq Q\right) \geq \frac{b}{b+h}\right\} \\
&= \inf\left\{\sqrt{n}X + \mu(n-\sqrt{n}) \geq 0 : \mathbb{P}\left(D_1 \leq X\right) \geq \frac{b}{b+h}\right\} \\
&= \sqrt{n}\inf\left\{X \geq 0 : \mathbb{P}\left(D_1 \leq X\right) \geq \frac{b}{b+h}\right\} + \mu(n-\sqrt{n}) \\
&= \sqrt{n}Q^* + \mu(n-\sqrt{n})
\end{aligned}
$$

With this we can now determine optimal expected cost in the pooled inventory system $G_P(Q_P^*)$:

$$
\begin{aligned}
G_P(Q_P^*) &= \mathbb{E}_{\{D\}}\left[h(Q_P^* - \sum_{i=1}^{N} D_i)_+ + b(\sum_{i=1}^{N} D_i - Q_P^*)_+\right] \\
&= \mathbb{E}_{\{D\}}\big[h((\sqrt{n}Q^* + \mu(n-\sqrt{n})) - (\sqrt{n}D_1 + \mu(n-\sqrt{n})))_+ \\
&+ b((\sqrt{n}D_1 + \mu(n-\sqrt{n})) - (\sqrt{n}Q^* + \mu(n-\sqrt{n})))_+\big] \\
&= \mathbb{E}_{\{D\}}\big[h(\sqrt{n}Q^* - \sqrt{n}D_1)_+ + b(\sqrt{n}D_1 - \sqrt{n}Q^*)_+\big] \\
&= \sqrt{n} \cdot \mathbb{E}_{\{D\}}\big[h(Q^* - D_1)_+ + b(D_1 - Q^*)_+\big] \\
&= \sqrt{n}G(Q^*)
\end{aligned}
$$

It follows that

$$
\frac{G_D(Q^*)}{G_P(Q_P^*)} = \frac{nG(Q^*)}{\sqrt{n}G(Q^*)} = \frac{n}{\sqrt{n}} = \sqrt{n}
$$

This result means that the expected optimal cost of the decentralized system is $\sqrt{n}$-times as large as the expected optimal cost of the pooled inventory system under the given assumptions.

3 An Investment Problem

We want to find the optimal investment strategy in a situation where the investor can make N sequential investment decisions, each resulting in one of two possible outcomes: Either (1) the invested money doubles (with probability $1/2 < p < 1$, or (2) the invested money is lost (with probability $1-p$).

Let x_k denote the wealth at time k (initial wealth x_0), let $u_k \in U_k(x_k) = [0, x_k]$ be the investment (decision) at time k and let

$$w_k(u_k) = \begin{cases} 2u_k & \text{with probability } p \\ 0 & \text{with probability } (1-p) \end{cases}$$

be the outcome of the investment at time k.

First, note that the state transition has the following form:

$$x_{k+1} = f(x_k, u_k, w_k) = x_k - u_k + w_k = \begin{cases} x_k + u_k & \text{with probability } p \\ x_k - u_k & \text{with probability } (1-p) \end{cases}$$

The DP algorithm has the form (note that we can consider the logarithm of the investor's wealth after the N^{th} investment as the objective function without changing the optimal strategy):

$$J_N(x_N) = \ln(x_N)$$

$$\begin{aligned} J_k(x_k) &= \max_{u_k \in U_k(x_k)} \mathbb{E}_{w_k}\left[J_{k+1}(x_{k+1})\right] \\ &= \max_{u_k \in U_k(x_k)} \left[pJ_{k+1}(x_k + u_k) + (1-p)J_{k+1}(x_k - u_k)\right] \end{aligned}$$

We want to find the strategy that yields the maximum expected wealth at time N, given initial wealth x_0: $J_0(x_0)$.

First, consider the optimal strategy for a one-period investment horizon, i.e. $J_0(x_0) = J_{N-1}(x_{N-1}) = \max_{u_k \in U_k(x_k)} \left[p \ln(x_k + u_k) + (1-p)\ln(x_k - u_k)\right]$. We can find the optimal investment strategy u^*_{N-1} using first-order conditions:

$$\begin{aligned} \frac{\partial}{\partial u_{N-1}} J_{N-1}(x_{N-1}) &= 0 \\ \frac{\partial}{\partial u_{N-1}} \left[p \ln(x_{N-1} + u_{N-1}) + (1-p)\ln(x_{N-1} - u_{N-1})\right] &= 0 \\ \frac{p}{x_{N-1} + u^*_{N-1}} - \frac{1-p}{x_{N-1} - u^*_{N-1}} &= 0 \\ p(x_{N-1} - u^*_{N-1}) &= (1-p)(x_{N-1} + u^*_{N-1}) \\ u^*_{N-1} &= (2p-1)x_{N-1} \end{aligned}$$

It follows that $J_{N-1}(x_{N-1})$ has the form:

$$\begin{aligned}
J_{N-1}(x_{N-1}) &= \max_{u_{N-1}} \mathbb{E}_{w_{N-1}} [\ln J_N(x_N)] \\
&= \max_{u_{N-1}} [p \ln(x_{N-1} + u_{N-1}) + (1-p)\ln(x_{N-1} - u_{N-1})] \\
&= p \ln(x_{N-1} + (2p-1)x_{N-1}) + (1-p)\ln(x_{N-1} - (2p-1)x_{N-1}) \\
&= p \ln(2p \cdot x_{N-1}) + (1-p)\ln((2-2p) \cdot x_{N-1}) \\
&= p \ln(2p) + \ln(x_{N-1}) + (1-p)\ln(2-2p) + \ln(x_{N-1}) \\
&= p \ln(2p) - p\ln(2-2p) + \ln(2-2p) + \ln(x_{N-1}) \\
&= A_{N-1} + \ln(x_{N-1})
\end{aligned}$$

where $A_{N-1} = p\ln(2p) - p\ln(2-2p) + \ln(2-2p)$ is independent of x_{N-1}. We see that $J_{N-1}(x_{N-1})$ contains the term $\ln(x_{N-1})$ and so the optimal strategy (known as "Kelly strategy") generalizes to all $k = 0, 1, .., (N-1)$ in the multi-period problem:

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

and therefore $J_k(x_k) = A_k + \ln(x_k)$ for some A_K independent of x_k.

4 Asset Selling with Maintenance Cost

Let's start by setting some notation. Let $w_0, w_1, ..., w_{N_1}$ be the offers the household gets. Given that the problem does not specify anything about investing the money from selling the house we will assume that such money does not get any interest rate from any possible investment. Let T denote the terminate state. Let $x_k = T$ at some time $k \leq N-1$ denote that the house has already been sold. Likewise, let $x_k \neq T$ at some time $k \leq N-1$ denote that the house has not been sold yet.

Let's now identify the variables that define the DP algorithm that we will used in this exercise:

- x_k: state of the system at period k, which is the amount of money that the household gets at such period.
- u_k: control variable at period k. In this case the household can take two action: sell or not sell. That is $u_k \in \{u_k^1(\text{sell}), u_k^2(\text{not sell})\}$
- w_k: uncertainty at time k, which corresponds to the offers.

Taking into account the fact that the rejects offers can be accepted in the future, the dynamics of the DP algorithm are:

$$x_{k+1} = \begin{cases} T & if\ x_k = T,\ or\ if\ x_k \neq T\ and\ u_k = u^1 \\ max(x_k, w_k) & otherwise \end{cases}$$

and the reward function is

$$\mathop{\mathbb{E}}_{w_k} \left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) \right\}$$

where

$$g_N(x_N) = \begin{cases} x_N & if\ x_N \neq T \\ 0 & otherwise \end{cases}$$

$$g_k(x_k, u_k, w_k) = \begin{cases} x_k & if\ x_N \neq T\ and\ u_k = u^1 \\ -c & if\ x_N \neq T\ and\ u_k = u^2 \\ 0 & if\ x_N = T \end{cases}$$

Therefore, the DP algorithm is,

$$J_N(x_N) = \begin{cases} x_N & if\ x_N \neq T \\ 0 & otherwise \end{cases}$$

$$\begin{aligned} J_k(x_k) &= \begin{cases} \underset{u_k \in U_k(x_k)}{max} \mathop{\mathbb{E}}_{w_k} \{ g_k(x_k, u_k, w_k) + J_{k+1}(x_{k+1}) \} & if\ x_N \neq T \\ 0 & if\ x_N = T \end{cases} \\ &= \begin{cases} max \mathop{\mathbb{E}}_{w_k} \{ g_k(x_k, u_k^1, w_k) + J_{k+1}(x_{k+1}),\ g_k(x_k, u_k^2, w_k) + J_{k+1}(x_{k+1}) \} & if\ x_N \neq T \\ 0 & if\ x_N = T \end{cases} \\ &= \begin{cases} max \{ \text{Revenue from selling, Expected revenue from not selling} \} & if\ x_N \neq T \\ 0 & if\ x_N = T \end{cases} \\ &= \begin{cases} max \{ x_k + 0, \mathop{\mathbb{E}}_{w_k} \{ J_{k+1}(w_k) - c \} \} & if\ x_N \neq T \\ 0 & if\ x_N = T \end{cases} \end{aligned}$$

This leads to the one-stopping set of the following form:

$$\begin{aligned} T_{N-1} &= \{x | x > \mathop{\mathbb{E}}_{w_k} (J_N(x_N) - c)\} \\ &= \{x | x > \mathop{\mathbb{E}}_{w_k} (max(x_{N-1}, w_{N-1}) - c)\} \\ &= \{x | x > \sum_{j=1}^{n} p_j max(x_{N-1}, w_{j,N-1}) - c\} \\ &= \{x | x > \sum_{j=1}^{n} p_j max(x, w_j) - c\} \end{aligned}$$

Note that the last line of the previous mathematical development comes from the fact that the function $\sum_{j=1}^{n} p_j max(x_{N-1}, w_{j,N-1}) - c$ is non-increasing which implies that T_{N-1} is absorbing in the sense that if a state belongs to T_{N-1} and termination is not selected, then the next state will also be in T_{N-1}. If we define $\bar{\alpha}$ to be

$$\bar{\alpha} = \sum_{j=1}^{n} p_j max(x, w_j) - c$$

then the optimal policy is defined by,

$$\mu_k^*(x_k) = \begin{cases} u^1 & if \ x_k > \bar{\alpha} \\ u^2 & if \ x_k < \bar{\alpha} \end{cases}$$

5 Scheduling Problem

(a)
We solve this problem using list indexing. Let L be the optimal order of answering the questions. Let i and j be the k^{th} and $(k+1)^{st}$ questions in the optimally ordered list L.

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

We can then calculate the expected return for answering the questions in this order.

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

Now we consider the is were quiestions i and j are interchanged.

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

Since L is optimal, $\mathbb{E}[reward\ of\ L] \geq \mathbb{E}[reward\ of\ L']$ This is equivalent to

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

So questions should be answered in order of decreasing $\frac{p_i R_i - (1-p_i)F_i}{1-p_i}$ (b)