Robust Stochastic Lot-Sizing Using Historical Data

D. Klabjan, D. Simchi-Levi, M. Song

Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA

1. Introduction

The stochastic lot-sizing model has been studied extensively in the inventory literature. Most of the research has focused on models with complete information about the distribution of customer demand. However, in most real-world situations, the distribution of customer demand is not known; only historical data is available. Thus, a common approach is to hypothesize the general family of demand distribution and then estimate the parameters specifying the distribution using the historical data. Once the probability distribution has been identified, the inventory problem is solved following the estimated distribution. This implies that the inventory policy is determined under the assumption that the demand distribution is correct.

In this paper, we consider a different approach that recognizes that the estimated customer demand distribution may not be accurate. We analyze the single-item stochastic finite-horizon periodic review lot-sizing model, under the assumption that the demand is subject to an unknown discrete distribution and only historical demand observations (given by histograms) are available. Rather than first estimating the demand distribution and then optimizing inventory decisions, as in the classical approach, we combine these two steps to minimize the worst case expected cost over a set of all possible distributions that satisfy a certain goodness-of-fit constraint. In this way, we combine the sample fitting and the inventory optimization, and characterize a *robust inventory control policy* based on the historical data.

The notion of robust inventory control is not new in the literature. Bertsimas and Thiele [1] analyze distribution-free inventory problems, in which demand in each period is assumed to be a random variable that takes values in a given range. That is, each period demand is assumed to be a random variable controlled only by two values: the lower and the upper estimator. To capture the trade-off between robustness and optimality, a parameter is defined to control the budgets of uncertainty at every time period. They show that for a variety of problems, the structures of optimal policy remains the same as in the associated model with complete information about the distribution of customer demand. A related model is analyzed in Bienstock and Ozbay [2].

The paper by Liyanage and Shanthikumar [3] is related to our research. The authors provide concrete examples in a single period (newsvendor) setting that illustrate that separating the distribution estimation and the inventory optimization, as is done in the classical approach, may lead to suboptimal solutions. They propose the use of operational statistics where they assume the demand distribution function belongs to a specific (predetermined) family and estimate the (single) parameter of the family in the inventory optimization model.

Thus, the robust optimization approach from Bertsimas and Thiele [1] does not use any historical data except for lower and upper bound on customer demand. On the other hand, Liyanage and Shanthikumar [3] use historical data but predetermine the family of distribution, and in fact restrict to distribution characterized by a single unknown paramater. Our research combines both strategies by integrating curve fitting with robust optimization. Specifically, we consider the set of demand distributions that satisfy a certain data fitting criterion with respect to historical data and characterize the optimal policy to minimize the maximium expected cost.

The main contributions of this paper are as follows:

1. We develop a robust minimax model that only requires historical data. The set of demand distributions is directly related to the testing of data fitting. We also show that this set can be defined by a set of second order cone constraints and therefore is computationally tractable.

2. The optimal policy to the robust model has the same structure as the corresponding policy in the classical stochastic lot-sizing model. In particular, the optimal policy is base-stock for the multi-period inventory problem without fixed ordering costs, and an (s, S) policy if the fixed ordering cost is considered.

In Section 2 we describe our robust model which incorporate historical data and present the optimality equation in a compact form. The structure of the optimal policies are characterized in Section 3. Finally, some direct extensions of our results are presented in Section 4.

2. Formulation of Robust Stochastic Lot-Sizing

The classical multi-period inventory problem considers a finite planning horizon of T periods. For each period $t = 1, ..., T$, let $\tilde{D}_t$ be a random variable representing demand in that period. We assume that $\tilde{D}_t$ has a discrete distribution for any t, and $\tilde{D}_1, ..., \tilde{D}_T$ are independent but not necessarily identical. The sequence of events in the model is as follows. At the beginning of each period, t, the decision maker reviews the inventory level, x_t, and places an order for q_t (possibly zero) units.

Since lead time is assumed to be zero [1], this order arrives immediatley and hence increases the inventory level up to y_t, where $y_t = x_t + q_t$. After observing demand, $\tilde{D}_t$, the net inventory at the beginning of period $t+1$ is reduced to

$$x_{t+1} = y_t - \tilde{D}_t \qquad t = 1, ..., T-1.$$

The ordering cost in each period $t = 1, ..., T-1$ includes two components, a fixed ordering cost K if $q_t > 0$, and a unit ordering cost c_t for each unit ordered. Inventory holding cost is charged at a rate of h_t for any unit of excess inventory at the end of a period, and a unit back-order cost b_t is incurred for any unit of unsatisfied demand. Similarly to classical inventory models, we assume that all shortages are backlogged.

Thus, the total expected cost for period t given the inventory levels before and after ordering (x_t and y_t respectively) is

$$\tilde{C}_t(x_t, y_t) = K\mathbb{I}(y_t - x_t) + c_t(y_t - x_t) + E\left[h_t\left(y_t - \tilde{D}_t\right)^+ + b_t\left(y_t - \tilde{D}_t\right)^-\right] \qquad t = 1, ..., T,$$

where $x^+ = \max(x, 0)$, $x^- = \max(-x, 0)$, $\mathbb{I}(x) = 1$ if $x > 0$ and $\mathbb{I}(x) = 0$ otherwise.

In the dynamic programming formulation, we consider $\tilde{V}_t(x_t)$, $t = 1, ..., T$, which denotes the optimal expected cost over the horizon $[t, T]$, given that the inventory level at the beginning of period t is x_t and an optimal policy is adopted over the horizon $[t, T]$. We define $\tilde{V}_t(x_{T+1}) = 0$. Let $\theta \in [0, 1]$ be the discount rate. The optimality equation reads

$$\tilde{V}_t(x_t) = \min_{y_t \geq x_t}\left\{\tilde{C}_t(x_t, y_t) + \theta E\left[\tilde{V}_{t+1}\left(y_t - \tilde{D}_t\right)\right]\right\} \qquad t = 1, ..., T. \tag{1}$$

Note that the distribution of $\tilde{D}_t$, $t = 1, ..., T$ is required to solve this dynamic programming formulation.

In practice, information on the demand distribution is typically not known. Rather, the inventory manager has historical data. Let $D_{t,i}$ denote the ith possible value that the demand of period t can take, and let $N_{t,i}$ denote the number of samples that fall within $[D_{t,i}, D_{t,i+1})$. Finally, define $n_t = \sum_i N_{t,i}$. [2]

We assume that $D_{t,0} = -\infty$ and $D_{t,M_{t+1}} = +\infty$, where M_t corresponds to the number of bins a histogram representing the historical data associated with $\tilde{D}_t$. The classical approach to identify the best distribution representing the data is to use the goodness-of-fit test. In this approach, the objective is to fit a distribution that "closely" folows the observed data.

[1] It can be easily extended to the case with nonzero lead time.

[2] For example, suppose that the planning horizon is one year and the time period is one month. Because of seasonality, the demand is distributed differently in every month. The samples for the demand of period t, for example, December, could be the historical demand in December in previous years.

For this purpose, let $P_{t,i} = P\left(\tilde{D}_t \in [D_{t,i}, D_{t,i+1})\right)$ be the probability that demand in period t falls in the interval $[D_{t,i}, D_{t,i+1})$ when the fitted distribution is applied. Clearly, $n_t P_{t,i}$ is the expected number of observation that falls in this interval according to the fitted distribution.

In the classical goodness-of-fit approach, the chi-square method, the statistical test is

$$\sum_i \frac{(N_{t,i} - n_t P_{t,i})^2}{n_t P_{t,i}} \leq \chi_t^2 \qquad t = 1, ..., T,$$

where the parameter χ_t^2 controls how close the observed sample data to the estimated expected number of observations according to the fitted distribution, $P_{t,i}$. Since $P_{t,i}$ should define a probability distribution, we have $\sum_i P_{t,i} = 1$ and $P_{t,i} \geq 0$. Let $\mathbf{P}_t$ denote the vector of $(P_{t,i})_i$. The set of distributions that satisfy the chi-square test is

$$\mathcal{P}_t = \left\{ \mathbf{P}_t \middle| \mathbf{A}_t \mathbf{P}_t = \mathbf{b}_t,\ \sum_i \frac{(N_{t,i} - n_t P_{t,i})^2}{n_t P_{t,i}} \leq \chi_t^2,\ \mathbf{P}_t \geq \mathbf{0} \right\} \qquad t = 1, ..., T. \tag{2}$$

The linear constraints $\mathbf{A}_t \mathbf{P}_t = \mathbf{b}_t$ capture the fact that $\sum_i P_{t,i} = 1$. They can also be used to model more complicated properties of the distribution set, such as constraints on the expected value, any moment or desired percentiles of the distributions.

We first provide an alternative characterization of $\mathcal{P}_t$. We assume that every norm is the Euclidian norm.

Proposition 1 *The set of demand distributions $\mathcal{P}_t$ defined in (2) is equivalent to the projection of the set*

$$\left\{ (\mathbf{P}_t, \mathbf{Q}_t) \middle| \boldsymbol{A}_t \boldsymbol{P}_t = \boldsymbol{b}_t, \sum_i N_{t,i}^2 Q_{t,i} - n_t^2 \leq n_t \chi_t^2, \left\| \begin{bmatrix} P_{t,i} - Q_{t,i} \\ 2 \end{bmatrix} \right\| \leq P_{t,i} + Q_{t,i} \right\}$$

on the space of $\mathbf{P}_t$.

To minimize the maximum expected cost arising from any distribution in the set $\mathcal{P}_t$, we have the optimality equation of the robust model

$$V_t(x_t) = \min_{y_t \geq x_t} \max_{\mathbf{P}_t \in \mathcal{P}_t} \left\{ C_t(x_t, y_t) + \theta \sum_i P_{t,i} V_{t+1}(y_t - D_{t,i}) \right\} \qquad t = 1, ..., T, \tag{3}$$

where $\mathcal{P}_t$ is defined by (2), $C_t(x_t, y_t)$ denotes the cost incurred in period t

$$C_t(x_t, y_t) = K\mathbb{I}(y_t - x_t) + c_t(y_t - x_t) + \sum_i P_{t,i} \left[h_t \left(y_t - D_{t,i}\right)^+ + b_t \left(y_t - D_{t,i}\right)^- \right],$$

and $V_{T+1}(x_{T+1}) = 0$ under the assumption that the salvage cost is 0.

We next give an alternative optimality equation.

Proposition 2 *The optimality equation of the robust stochastic model (3) is equivalent to*

$$\begin{array}{rcll} V_t(x_t) & = & \displaystyle\min_{y_t,\boldsymbol{U}_t,\boldsymbol{p}_t,\boldsymbol{u}_t,\lambda_t} & K_t\mathbb{I}(y_t-x_t)+c_t(y_t-x_t)+\boldsymbol{p}_t^T\boldsymbol{b}_t-2\sum_i u_{t,i}N_{t,i}+\lambda_t\left(n_t^2+n_t\chi_t^2\right) \\ & & \text{s.t.} & \left\|\begin{bmatrix}\boldsymbol{p}_t^T-U_{t,i}\boldsymbol{A}_{t,i}-\lambda_t\\ 2u_{t,i}\end{bmatrix}\right\|\le \boldsymbol{p}_t^T-U_{t,i}\boldsymbol{A}_{t,i}+\lambda_t \qquad \forall i \\ & & & U_{t,i}\ge h_t\left(y_t-D_{t,i}\right)+\theta V_{t+1}(y_t-D_{t,i}) \qquad \forall i \\ & & & U_{t,i}\ge b_t\left(D_{t,i}-y_t\right)+\theta V_{t+1}(y_t-D_{t,i}) \qquad \forall i \\ & & & y_t\ge x_t, \end{array}$$

for any $t=1,...,T$.

Note that this is not really the standard optimality equation since $V_{t+1}(\cdot)$ is present in constraints and not the objective function.

3. Properties of Optimal Policies

We first study the linear ordering cost case ($K=0$) and then we add economies of scale ($K>0$).

3.1 Models with Variable Ordering Cost

Theorem 3 *The base-stock policy is optimal for the robust stochastic model with the linear ordering cost. In particular, let* S_t^* *be an optimal solution to the following convex programming problem*

$$\begin{array}{ll} \displaystyle\min_{y_t,\boldsymbol{U}_t,\boldsymbol{p}_t,\boldsymbol{u}_t,\lambda_t} & c_ty_t+\boldsymbol{p}_t^T\boldsymbol{b}_t-2\sum_i u_{t,i}N_{t,i}+\lambda_t\left(n_t^2+n_t\chi_t^2\right) \\ \text{s.t.} & \left\|\begin{bmatrix}\boldsymbol{p}_t^T\boldsymbol{A}_{t,i}-U_{t,i}-\lambda_t\\ 2u_{t,i}\end{bmatrix}\right\|\le \boldsymbol{p}_t^T\boldsymbol{A}_{t,i}-U_{t,i}+\lambda_t \quad \forall i \\ & U_{t,i}\ge h_t\left(y_t-D_{t,i}\right)+\theta V_{t+1}(y_t-D_{t,i}) \qquad \forall i \\ & U_{t,i}\ge b_t\left(D_{t,i}-y_t\right)+\theta V_{t+1}(y_t-D_{t,i}) \qquad \forall i. \end{array}$$

The policy orders $S_t^*-x_t$ *units in period* t *if* $x_t\le S_t^*$ *and no order is placed otherwise.*

3.2 Models with Fixed and Variable Ordering Cost

We now assume that there is also a fixed ordering cost.

Theorem 4 *An* (s,S) *policy is optimal for the robust stochastic model with fixed and variable ordering cost. In particular, let* S_t *be the optimal solution to the minimization problem*

$$\begin{array}{ll} \displaystyle\min_{y_t,\boldsymbol{U}_t,\boldsymbol{p}_t,\boldsymbol{u}_t,\lambda_t} & c_ty_t+\boldsymbol{p}_t^T\boldsymbol{b}_t-2\sum_i u_{t,i}N_{t,i}+\lambda_t\left(n_t^2+n_t\chi_t^2\right) \\ \text{s.t.} & \left\|\begin{bmatrix}\boldsymbol{p}_t^T\boldsymbol{A}_{t,i}-U_{t,i}-\lambda_t\\ 2u_{t,i}\end{bmatrix}\right\|\le \boldsymbol{p}_t^T\boldsymbol{A}_{t,i}-U_{t,i}+\lambda_t \quad \forall i \\ & U_{t,i}\ge h_t\left(y_t-D_{t,i}\right)+\theta V_{t+1}(y_t-D_{t,i}) \qquad \forall i \\ & U_{t,i}\ge b_t\left(D_{t,i}-y_t\right)+\theta V_{t+1}(y_t-D_{t,i}) \qquad \forall i, \end{array}$$

and let s_t be the smallest element of the set

$$\{s_t \mid s_t \leq S_t,\ f_t(s_t) = f_t(S_t) + K\},$$

where

$$f_t(y_t) = c_t y_t + \max_{\boldsymbol{P}_t \in \mathcal{P}_t} \sum_i P_{t,i} \left[h_t \left(y_t - D_{t,i}\right)^+ + b_t \left(y_t - D_{t,i}\right)^- + \theta V_{t+1}(y_t - D_{t,i}) \right].$$

The policy orders $S_t - x_t$ units in period t if $x_t \leq s_t$ and no order is placed otherwise.

4. Conclusions and Extensions

In this paper, we propose a robust stochastic model for the multi-period lot sizing problem, in which the demand distribution is unknown and the only available information is historical data. This robust framework based on historical data can be applied to any finite-horizon dynamic programming problem in which the stochastic inputs are subject to some discrete distribution.

We provide important theoretical insights on the structure of optimal policies. Interestingly, the proofs of Theorem 3 and Theorem 4 (not given here), only use convexity of $\mathcal{P}_t$. Thus, the optimality of the base-stock policy and the (s, S) policy holds under the more general assumption of $\mathcal{P}_t$ being convex.

References

[1] Bertsimas, D. and A. Thiele. 2004. A robust optimization approach to supply chain management. *Proceedings of 14th IPCO*, 86-100.

[2] Bienstock, D. and N. Ozbay. 2005. Computing robust basestock levels. CORC Report, Columbia University.

[3] Liyanage, L. H. and J. G. Shanthikumar. 2005. A practical inventory control policy using operational statistics. *Operations Research Letters* **33**, 341-348.