A decision support system for forecasting and optimal procurement

Neven Miculnic

April 26, 2016

ABSTRACT

Optimal procurement in most industries involve forecasting of two quantities: prices of raw materials and customer's demand. The aim of this work is to integrate forecasts into production planning models, with the aim of minimizing overall procurement, holding and production costs under demand satisfaction constraints. The decision support system should allow the decision maker to integrate qualitative, unstructured information through a simple interface, for scenario selection and solution refinement

1

PROBLEM DEFINITION

1.1 INTRODUCTION

This problem is variation of lot sizing problem in stochastic setting. Variation is as follows. We have stochastic prediction for raw material supply cost $\mathbf{s}$ at future time moment t. The *factory* converts all bought raw materials at time t and stores them into a warehouse. Number of discrete time moments is denoted by n.

One time moment storage costs fixed amount h, that is the holding cost.

Likewise demand for the product is equally stochastic, denoted as $\mathbf{d}$. In case there's no products in storage to satisfy demand we pay backlogging cost denoted as b.

Our aim is to optimize our procurement policy by variating $\mathbf{x}$, that is product amount we buy at time moment t. However we are constrained by the maximum raw materials we can buy each day by $\mathbf{x}_{\text{max}}$.

1.2 FORMAL PROBLEM DEFINITION

1.2.1 *Definitions*

Every input value is known beforehand, and it's assumed for random variables their distributions are known.

$\mathbf{s} = \left[S_1, S_2, \ldots, S_n\right]^\intercal$	Supply cost random vector
$\mathbf{x} = \left[x_1, x_2, \ldots, x_n\right]^\intercal$	Procurement quantity vector
$X_i = \sum_{t \leq i} x_t$	Cumulative procurement quantity sum
$\mathbf{d} = \left[Y_1, Y_2, \ldots, Y_n\right]$	Demand random vector, Y_t is a random variable
$D_i = \sum_{t \leq i} d_t$	Cumulative demand random variable
b	backlogging cost
h	holding cost

1.2.2 *Variables*

$\mathbf{x}$ is our decision variable, as described previously.

1.2.3 *Constraints*

$$\begin{aligned} \mathbf{x} &\in \mathbb{N}_0^n \\ \mathbf{x} &\leq \mathbf{x}_{\max} \\ D_i &\geq 0 \ \forall t \end{aligned}$$

1.3 OBJECTIVE FUNCTION

First let's define auxiliary function $c(t)$ which determines holding or backlogging cost at t:

$$c(t) = \begin{cases} h\left(X_i - D_i\right) & X_i \geq D_i \\ b\left(D_i - X_i\right) & X_i < D_i \end{cases} \tag{1}$$

Therefore our minimizing objective function is:

$$\min f(\mathbf{x}) = \mathbf{s}^\intercal \mathbf{x} + \sum c(t) \tag{2}$$

Because the function is stochastic our aim is to minimize $\mathrm{E}\left[f(\mathbf{x})\right]$.

2

DEEPER ANALYSIS

First we are going to analise problem deeply without presuming any independence or probability distribution on random variables. Later in subsequent chapters we are going to focus more on where demand at time t has independent Gaussian distribution and mean.

2.1 COST FUNCTION

The minimizing function is:

$$\mathrm{E}\left[\mathbf{s}^\mathsf{T}\mathbf{x} + \sum c(t)\right]$$

Due to linerality of expectation and x being variable it's equal to:

$$\mathbf{x}\,\mathrm{E}\left[\mathbf{s}^\mathsf{T}\right] + \mathrm{E}\left[\sum c(t)\right]$$

Therefore only needed modeling information for supply cost is its expectation $\mathrm{E}\left[\mathbf{s}\right]$. The other part is more trickier since D_i and D_j aren't usually independent.

2.2 HANDLING DEMAND COST NON-LINEARITY

As we can see in equation 1 we have non-linearity depending whether we're satifying all demand or are we backlogging demand at time t. Therefore here are two possible solutions for minimizing objective function 2.

Simulation

We generate multiple scenarios for demand vector, d according to probability distribution. For small n and relatively small number of outcomes in each random variable we can exhaustedly model each scenario, scale it appropriately and feed to MIP solver[1]

Safety net approach

Alternatively, we can artificially add new constraints and avoiding backlogging with arbitrary probably. This model assumes backlogging cost are significantly greater than storage cost, that is backlogging penalty is severe.

Thus we chose values arbitrary realizations of random variables D_t and add additional constraints of the form:

$$X_t \geq D_i \ \forall t \tag{3}$$

which reduces cost function 1 to:

$$c(t) = h\left(X_i - D_i\right) \ \ X_i \geq D_i$$

and enables are faster solving approaches. Given parameters D_i chosen and their underlying distribution we derive non-increasing function $p(t)$ which represent probability of not breaking newly introduced constraints 3

Min-cost max flow with uncertaintiy

Since deterministic case can be modeled with min-cost max flow, we can investigate further in this directi

Stochastic Dynamic Programmnig

TODO.. maybe, I've got recommendation by a friend, need to read some papers.

1 There's a trick on using binary variable for discontinuity in cost function 1