PRODUCTION PLANNING AND INVENTORY MANAGEMENT UNDER UNCERTAINTIES: STOCHASTIC MODELS AND NUMERICAL SOLUTIONS WITH APPLICATIONS IN THE PAPER INDUSTRY

K. Karen Yin*, and Hu Liu
Department of Wood and Paper Science
University of Minnesota, St. Paul, MN 55108

Abstract

This paper is concerned with problem formulation and solution procedure for production planning and inventory management of systems under uncertainties. Having continuous dynamics intertwined with discrete-event interventions, the production system is modeled by finite-state continuous-time Markov chains. Discretizing the Hamilton-Jacobi-Bellman (HJB) equations satisfied by the value functions and using an approximation procedure yield the optimal solution. The inventory problem under consideration is formulated by a Markov Decision Process (MDP) model. The optimal policy is obtained by using the policy-improvement algorithm.

Keywords

Production planning, Inventory management, Stochastic models, Markov chains.

Introduction

Production planning and inventory management have attracted growing attention in recent years. Numerous papers have been published (see, e.g., Applequist et al., 1997, Balasubramanian and Grossmann, 2002, Bassett et al., 1997; Gupta et al., 2000, Pekny and Miller, 1990, Petkov and Mararas, 1997 among many others). For optimization problems under uncertainty, two most important characteristics (Bertsekas, 1976) not present in their deterministic counterpart are the need of considering risks in the model formulation and the possibility of information gathering during the decision process. This work concerns problem formulation, numerical algorithms, and solution procedure for production planning and inventory management under uncertainties. By incorporating different kinds of uncertainties into the mathematical model and by constructing a multi-period decision process, we seek the optimal policy that minimizes the expected costs over a given time span.

Production Planning

Many systems contain continuous dynamics intertwined with discrete events or subject to discrete-event interventions, which lead to jump discontinuity in their evolution. To better understand and more effectively deal with uncertainties from various sources require stochastic models that can characterize the unique feature of each major event in such hybrid systems.

There is a rich literature on both theory and computational algorithm development for manufacturing systems under uncertainties; see, e.g., Akella and Kumar (1986), Bertsekas (1976, 1987), Gershwin (1994), Sethi and Zhang (1994), Yin and Zhang (1997,1998) and the references therein. In this work we adopt a class of continuous-time, finite-state stochastic models to describe the dynamics of such hybrid systems. We provide the dynamic programming equation, and present numerical scheme that leads to an approximation of the optimal policy. The objective function used includes both production and holding costs and can be easily extended to include other costs. We consider two types of uncertainty, demand and production capacity, and formulate them using finite-state Markov chains. Such an approach allows us to quantitatively describe the random and jump behavior that is common in many stochastic systems.

Problem Formulation

Consider a manufacturing system that produces r different products. Let $\mathbf{u}_t \in \mathrm{R}^r$ denote the production rates that may vary with time and the state or capacity of the machine. Therefore $\mathbf{u}_t \geq 0$ and is subject to the random production capacity that is a random process. With the total surplus (the inventory/shortage level) $\mathbf{x}_t \in \mathrm{R}^r$ and the random demand rate $\mathbf{z}_t \in \mathrm{R}^r$, the system is given by

---

* To whom all correspondence should be addressed

$$\dot{\mathbf{x}}_t = \mathbf{u}_t - \mathbf{z}_t \qquad \mathbf{x}_0 = \mathbf{x}, \tag{1}$$

where $\mathbf{x} \in R^r$ is the initial surplus level.

We consider the cost functional J defined by

$$J(\mathbf{x}, \alpha, \mathbf{u}(\cdot), \mathbf{z}) = E\int_0^\infty e^{-\rho t}[h(\mathbf{x}_t) + c(\mathbf{u}_t)]dt \tag{2}$$

where $\rho > 0$ is the discount rate, $h(\cdot)$ represents the holding cost, $c(\cdot)$ denotes the production cost, and $\mathbf{x}$, α, and $\mathbf{z}$ are the initial surplus, the initial capacity, and the initial demand, respectively. The expectation E is taken over both random machine capacity and random demand. We seek the optimal production rate, $\mathbf{u}_t$, to minimize the objective function (2), subject to dynamics described by (1), the capacity α (t), and certain production constraints and the given initial conditions.

Specification of Demand and Capacity Processes

Denote the set of the values of demand by $Z=\{\mathbf{z}^1,\ldots,\mathbf{z}^d\}$. Assume that the demand process $\mathbf{z}(t) \in Z$ and that the transitions among the states are random. The production capacity $\alpha(t)$ is also a random process having state space $M=\{\alpha^1,\ldots,\alpha^m\}$. Suppose the transition between the states of production capacity may depend on the production rate $\mathbf{u}_t$. Now we face a joint stochastic process $(\alpha(t), \mathbf{z}(t))$.

It is conceivable that at any given time the production capacity determines the set of all possible rates of production, $\mathbf{u}_t$. For each state of the capacity, $\alpha^1 \le \alpha^l \le \alpha^m$ ($1 < l < m$), without loss of generality, denote the set of production rate constraints by U_i, Then the production rate at time t is subject to the constraint $\mathbf{u}_t \in U_{\alpha(t)}$.

We further assume that the random processes of demand rate $\mathbf{z}(\cdot)$ and capacity $\alpha(\cdot)$ are both finite-state continuous-time Markov chains. Consider the demand $\mathbf{z}(t) \in Z$ to be governed by a generator $Q_d = (q^d_{ij})$; a $d \times d$ matrix. To model the dependence of α (t) on the production rate, let $\alpha(t) \in M$ be governed by an infinitesimal generator $Q_m(\mathbf{u})$ that depends on production rates; $Q_m(\mathbf{u}) = (q^m_{ij}(\mathbf{u}))$ is an $m \times m$ matrix. Let $\Gamma^m = \{U = (\mathbf{u}^1,\ldots,\mathbf{u}^m)$: $\mathbf{u}^1 \in U_i.\}$. The two Markov chains, $\alpha(\cdot)$ and $\mathbf{z}(\cdot)$, are generated by $Q_m(\mathbf{u})$ and Q_d, respectively, i.e., for any functions ϕ on M and ψ on Z,

$$\begin{aligned} Q^m\phi(\cdot)(j) &= \Sigma_{j_1 \ne j}\, q^m_{jj_1}(\mathbf{u}^j)[\phi(j_1) - \phi(j)], \\ Q^d\psi(\cdot)(j) &= \Sigma_{j_1 \ne j}\, q^d_{jj_1}[\psi(j_1) - \psi(j)]. \end{aligned} \tag{3}$$

Let $\mathcal{A}$ denote the set of all admissible controls. Our objective is to find an admissible feedback control policy $\mathbf{u}(\cdot) \in \mathcal{A}$ that minimizes the cost function $J(\mathbf{x};\alpha, \mathbf{u}(\cdot), \mathbf{z})$.

System of HJB Equations

Define the value function $v(\cdot)$ as the minimum of the cost over $\mathbf{u}(\cdot) \in \mathcal{A}$, i.e.,

$$v(\mathbf{x}, \alpha, \mathbf{z}) = \inf_{\mathbf{u}(\cdot) \in \mathcal{A}(\alpha, \mathbf{z})} J(\mathbf{x}, \alpha, \mathbf{u}(\cdot), \mathbf{z}). \tag{4}$$

The use of Markov chains results in a total of $|M|$ value functions. With a dynamic programming approach, it can be shown that the value functions are convex and satisfy the Hamilton-Jacobi-Bellman (HJB) equations (see Fleming and Rishel, 1975; Sethi and Zhang, 1994):

$$\begin{aligned} \rho\, v(\mathbf{x}, \alpha, \mathbf{z}) &= \min_{\mathbf{u} \in \Gamma}\{(\mathbf{u} - \mathbf{z}) \cdot \nabla v(\mathbf{x}, \alpha, \mathbf{z}) + [h(\mathbf{x}) \\ &+ c(\mathbf{u})]\} + Q^m v(\mathbf{x}, \cdot, \mathbf{z})(\alpha) + Q^d v(\mathbf{x}, \alpha, \cdot)(\mathbf{z}) \end{aligned} \tag{5}$$

where $\mathbf{x} \in R^r$, $\alpha \in M$, $\mathbf{z} \in Z$, $\mathbf{a} \cdot \mathbf{b}$ denotes the inner product of the vectors $\mathbf{a}$ and $\mathbf{b}$, $\nabla f(\mathbf{x})$ is the gradient of f.

Solving (5) leads to the optimal production policy $\mathbf{u}^*$. Similar to many other controlled Markovian systems, the closed-form solution of the HJB equations is difficult or even impossible to obtain. Therefore, we resort to numerical procedure.

Numerical Procedure for the Optimal Policy

Using the numerical methods developed in Kushner (1990) and Kushner and Dupuis (1992) (see also Yin and Zhang, 1998), we discretize (5) by discretizing the space R^r with grid $\Delta > 0$, which yields a discrete space R_Δ^r. Let $\{\mathbf{e}_i\}_{i=1}^r = \{e_1, \ldots, e_r\}$ be a standard basis for the Euclidean space R^r. Using Δ as the step size, we approximate the value function $v(\mathbf{x};\ \alpha;\ \mathbf{z})$ by a sequence of functions $v^\Delta(\mathbf{x};\alpha;\mathbf{z})$; and its partial derivatives v_{xj} $(\mathbf{x};\ \alpha;\ \mathbf{z})$ by the corresponding finite differences, then write the HJB equation (5) in terms of $v(\mathbf{x};\ \alpha;\ \mathbf{z})$: This newly obtained equation can be expressed in the form

$$v^\Delta_{n+1}(\mathbf{x}, \mathbf{z}, \alpha) = \Gamma(v^\Delta_{n+1}(\mathbf{x}, \mathbf{z}, \alpha)).$$

The value function can be approximated by the value iteration procedure. Starting from an initial value of $\mathbf{x}$ and an arbitrary initial guess v^Δ_0, the procedure calls for repeated iterations until certain convergence criterion is satisfied.

Applications

The models and numerical algorithms were applied to a papermaking process. Using real demand data collected from a large paper manufacturer, we obtained the midterm production plans for different situations. The optimal strategies obtained allow us to make production decisions sequentially throughout the process lifespan.

Inventory Management

We are interested in inventory policies capable of handling situation of random demand and periodical replenishments. Observing the randomness and regularity in the inventory process, we choose to describe it with discrete-time finite-state Markov chains, and use Markov decision process (MDP) model to determine the optimal policy.

State Space and Transition Probabilities of the Markov Chain

To establish the mathematical model requires specifying the key elements of the discrete time Markov chain, which entails designating its state space and prescribing the dependence relations among the random variables based on the real process data.

Let d_1; d_2;...represent the successive demands for a particular product. Assume that d_n are i.i.d. random variables whose future values are unknown. Let $\tilde{X}_n$ denote the stock of the product on hand at the end of the nth period. The states of the stochastic process, $\{\tilde{X}_n\}$, consist of the possible values of its stock size. The amounts in stock, $\tilde{X}_0, \tilde{X}_1, \ldots$, constitute a Markov chain whose transition probability matrix is determined by the demand and the replenishment policy. If $\tilde{X}_n$ is a continuous random variable, we may discretize it via certain transformation to simplify the solution procedure. The resulting discrete random variable, X_n, also indicates the level of the stocks but takes values in $M = \{0,\ldots,m\}$. Such discretization allows us to model this inventory system by an $(m + 1)$-state Markov chain.

To completely define a Markov process $\{X_n\}$ requires specifying its initial state and its transition probability matrix $\mathbf{P} = \| P_{ij} \|$. For the inventory management problem, the former is usually available, whereas the latter is affected by the random demand as well as the replenishment activities.

Decisions, Actions and Policies

The inventory system evolves over time according to the joint effect of the probability laws and the sequence of decisions and actions. It fits to the general finite-state discrete-time Markov decision processes. The stock on hand is recorded at the end of each period. Subsequently, a decision is made and an action is taken. The question needs to be answered is which decision should be chosen at any given time and state. An inventory policy is a rule that prescribes decisions to be made for each state of the system during the entire time period of interest. Characterized by the values $\{\delta_0(R), \delta_1(R),\ldots,\delta_m(R)\}$, any policy R specifies decisions $\delta_i(R) = k$, $(k = 0, 1, \ldots, K)$; for all states i, $(i = 0, 1, \ldots, m)$, at every time instant. We want to choose a policy that minimizes the long-run expected (average) cost.

A policy R requires that the decision $\delta_0(R)$ be made whenever the system is in state i. Effected by this policy as well as the random demand, the system will move to a new state j according to the corresponding probabilities P_{ij}. For countable items, the determination of the transition probability matrix is relatively straightforward. For products measured by weights, we can discretize them as described above.

Markov Decision Process

The evolution of the system is affected by the random demands as well as the replenishment activities governed by the inventory policy. Let X_n be the state of the system at time n; let Δ_n be the decision/action chosen. Under any fixed policy R; the pair $Y_n = (X_n, \Delta_n)$ forms a two-dimensional Markov chain. For a given feedback policy, the decision $\delta_i(R) = k$ is prescribed for every state $i = 0, 1, \ldots, m$. Consequently, when the system is in state i and the policy R is used and an action based on the decision $\delta_i(R) = k$ is excised, the probability of its moving to state j at the next time period P_{ij} can be obtained. Starting from X_0, the realization of the underlying stochastic process is X_0, $X_1,\ldots$, and the decisions made are Δ_0, Δ_1, ... Note that $\Delta_n = \delta_{Xn}(R) \in \{0, 1, \ldots, K\}$ if the feedback policy is used. The sequences of observed states and decisions made are the so-called Markov Decision Process.

The Long-run Expected Cost

We seek the optimal policy in the sense that it will minimize the (long-run) expected average cost per unit time. It should be noted that another consideration in practice is that the policy should be relatively simple and easily implementable. Suppose a cost $C_{Xn\Delta n}$ is incurred when the process is in state X_n and a decision Δ_n is made. A function of both $X_n = 0, 1, \ldots, m$, and $\Delta_n = 0, 1, \ldots, K$; $C_{Xn\Delta n}$ is also a random variable. Its long-run expected average cost per unit time over a period of N is

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} E[C_{X_n \Delta_n}] = \sum_{i=0}^{m} \sum_{k=0}^{K} \pi_{ik} C_{ik} \tag{6}$$

where π_{ik} is the stationary (limiting) probability distribution associated with the transition probabilities.

The Policy Improvement Algorithm

We resort to the Policy-Improvement Algorithm in this work to seek for the optimal policy. Let $g(R)$ represent the long-run expected average cost following any given policy R, i.e. $g(R) = \sum_{i=0}^{m} \pi_i C_{ik}$. Denote $v_i^n(R)$ the total expected cost of a system starting in state i and evolving in a period of length n. By definition, it satisfies the following recursive formula

$$v_i^n(R) = C_{ik} + \sum_{j=0}^{m} P_{ij}(k) v_j^{n-1}(R). \tag{7}$$

Note that C_{ik}, the cost incurred in the first time period, is also an expected cost. It can be shown that (Hillier and Lieberman, 1999) for $i = 0, 1, \ldots, m$

$$g(R) = C_{ik} - v_i(R) + \sum_{j=0}^{m} P_{ij}(k) v_j(R). \tag{8}$$

For a system of $m + 1$ states, Eq. (8) consists of $m + 1$ simultaneous equations but $m + 2$ unknowns, $g(R)$ and $v_i(R)$ ($i = 0, 1, \ldots, m$). To obtain a unique solution, it is customary to specify $v_m(R) = 0$: Solving the system of equations (8) yields the long-run expected average cost per unit time $g(R)$ if the policy R is used. An optimal policy is one that results in the lowest cost $g(R^*)$. A policy improvement algorithm allows us to obtain the optimal policy. The procedure begins by choosing an arbitrary policy R_1. For the given policy R_1, the transition probabilities $P_{ij}(k)$ are available hence the expected costs $C_{ik}(R_1)$ can be computed. Subsequently, the values of $g(R_1)$, $v_0(R_1)$; $v_1(R_1)$; : : :; $v_{m-1}(R_1)$ can be obtained from Eq. (8). In the second step, the current values of $v_i(R_1)$ are used to find an improved policy R_2. Specifically, for each state i, choose such decision $\delta_i(R_2)$ that makes the right-hand-side of Eq.(8) a minimum,

$$\delta_i(R_2) = \arg\min_k \{C_{ik}(R_2) - v_i(R_1) + \sum_{j=0}^{m} P_{ij}(k_{R_2}) v_j(R_1)\} \quad \text{for all } i, \tag{9}$$

where $\arg\min_k f(k)$ is the value of $k \in \{0, 1, \ldots, K\}$ that minimizes $f(k)$. The set of the best decisions for all states ($i = 0, 1, \ldots, m$) constitute the second, or the improved policy R_2. Repeating this iteration procedure until the two successive R's are the same.

Applications

Those outlined above have been applied in the determination of inventory policy for a large paper manufacturer. Using real customer demand data, we have shown (Yin et al., 2002) that the MDP model consistently yields better policies than the traditional replenishment ones. In general, it results in lower average inventory level and/or lower stockout and requires fewer reorders to be placed.

Summary

In this paper, we formulate production planning as a stochastic control problem driven by Markovian noise. Stochastic differential equations are used to describe the system dynamics. To quantify the stochastic and jump behavior, we model the random demand and capacity processes with two finite-state continuous-time Markov chains. The long-run expected cost is considered in our formulation. The numerical procedure designed requires discretizing the original functional space and approximating the value function by its discretized version. The iteration results provide us with the optimal production rates under different situations.

We resort to Markov decision process models for decision making in the inventory management stage. It requires designating the possible decisions and the corresponding actions taken, identifying the Markov decision model related to the underlying system by defining its state space and the transition probabilities, specifying the cost function and evaluating its individual component, and then using the policy improvement algorithm to obtain the optimal policy.

References

Akella, R. and Kumar, P. R. (1986). Optimal control of production rate in a failure-prone manufacturing system, *IEEE Trans. Automat. Control*, **AC-31**, 116-126.

Applequist, G., Samikoglu, O., Pekny, J., and Reklaitis, G. (1997). Issues in the use, design and evolution of process scheduling and planning systems, *ISA Transactions*, **36**, 2, 81-121.

Balasubramanian, J. and Grossmann, I.E. (2002). Scheduling to minimize expected completion time in flowshop plants with uncertain processing times, *Comput. & Chem. Eng.*, **26**, 41-57.

Bassett, M. H., Pekny, J. F., and Reklaitis, G. V.(1997). Using detailed scheduling to obtain realistic operating policies for a batch processing facility, *Ind. Eng. Chem. Res.*, **36**, 5, 1717-1726.

Bertsekas, D. (1976). *Dynamic Programming and Stochastic Control*. Academic Press, New York, NY.

Bertsekas, D. (1987). *Dynamic Programming: Deterministic and Stochastic Models*, Prentice-Hall.

Fleming, W. H. and Rishel, R. W. (1975). *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York.

Gershwin, S. B. (1994). *Manufacturing Systems Engineering*, Prentice-Hall, Englewood Cliffs.

Gupta, A., Maranas, C. D., and McDonald, C. M. (2000). Mid-term supply chain planning under demand uncertainty: customer demand satisfaction and inventory management, *Comput. & Chem. Eng.*, **24**, 12, 2613-2621.

Hillier, F. S. and Lieberman, G. J. (1999). *Introduction to Operations Research*, 6ed, Holden-Day, Inc. Oakland, CA.

Kushner H. J. (1990). Numerical methods for stochastic control problems in continuous time, *SIAM J. Control Optim.* **28**, 999-1048.

Kushner, H. J. and Dupuis, P. G. (1992). *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, New York.

Pekny, J. and Miller, D. (1990). Exact solution of the no-wait flowshop scheduling problem with a comparison to heuristic methods, *Comput. & Chem. Eng.*, **14**, 9, 1009-1023.

Petkov, S. B. and Mararas, C. D. (1997). Multiperiod planning and scheduling of multiproduct batch plants under demand uncertainty, *Ind. Eng. Chem. Res.*, **36**, 11, 4864-4881.

Sethi, S. P. and Zhang, Q. (1994). *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhauser, Boston.

Yin, G. G. and Zhang, Q. (1998). *Continuous Markov Chains and Applications, A Singular Perturbation Approach*, Springer, New York.

Yin, G. G. and Zhang, Q. eds. (1997). *Mathematics of Stochastic Manufacturing Systems*, American Mathematical Society, Providence.

Yin, K., Liu, H. and Johnson, N. (2002). Markovian inventory policy with applications to the paper industry, to appear in *Comput. & Chem. Eng.*