STUDIES
ON
STOCHASTIC DYNAMIC OPTIMIZATION MODELS WITH APPLICATIONS

by

KATSUSHIGE SAWAKI

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF ENGINEERING

AT

KYOTO UNIVERSITY

Kyoto, Japan

JANUARY, 1997

ABSTRACT

This dissertation develops wide classes of stochastic dynamic optimization models that are characterized by many stochastic processes and several fields of applications. Chapter 2 studies a certain class of dynamic programs and its applications. Chapter 3 deals with inventory control models, including the study of optimal policies with fixed inventory holding costs and for price differential products with no carrying over of any remaining inventory to the next day. In Chapter 4 airline seat allocation models are analyzed to derive an optimal booking policy. In Chapter 5 we consider portfolio selection problems related to allocating firms' or individuals' wealth (money) among available assets. Money is also able to be treated as inventory. In Chapter 6 we propose new software reliability growth models based on counting processes for instruction execution in software. Chapter 7 summarizes conclusions drawn from the previous chapters.

ACKNOWLEDGEMENT

I would like to express sincere gratitude to Professor Masao Fukushima, Department of Applied Mathematics and Physics, Graduate School of Engineering, Kyoto University, for his invaluable comments and continuous encouragement in accomplishing this work.

I wish to express my thanks to Professor Toshihide Ibaraki and Professor Toshiharu Hasegawa for their helpful comments and suggestions.

I would like to thank Father Hans-Jügen Marx, President of Nanzan University, for his constant encouragement. Thanks also go to Nanzan University for providing financial support and for allowing me the time to complete my work. I also would like to thank Mr. Indra Malela for his word processing of my draft and overcoming difficulties of time.

Last but not least, I must thank my wife, Emiko, and my three children for their support and understanding during the period of completing this work

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The combined theories of dynamic programming and Markov decision processes have been applied to many managerial decision problems, including inventory resource allocation, portfolio management, and machine maintenance problems.

When an information generating process is described by a Markov process and a multi-stage decision process is able to be applied through the technique of dynamic programming, the possibility of handling the dynamic system in a multi-period model as well as the risk precautionary motive has been widely recognized. However, two problems are often pointed out as the reasons why the combined theory of dynamic programming and Markov decision theory has been abandoned as a multi-stage decision making process: first, it is practically correct that a dynamical system can be approximated by a Markov process; secondly, the larger the size of the problem becomes, the larger the computational burden of an algorithm based on dynamic programming.

This thesis aims at overcoming the theoretical defects of Markov decision processes by considering many stochastic processes; in regard to dynamic programming we develop a theory for a general class of dynamic programs that yield policies that are simple and ε-optimal. The formulation of our general class of dynamic programs is motivated by consideration of the special structure that the partially observable model possesses.

In this thesis an important special class of stochastic dynamic models that requires a series of sequential decisions is widely considered from various aspects. These decisions must be made sequentially over both discrete periods and continous time. However, not all decisions need be made at the beginning of the period; instead, we have to choose a "policy" that determines what we should do in each period as a function of the information that is then available.

Organization of the Dissertation

It has been recognized that stochastic dynamic optimization models are primarily concerned with the aspects of problem solving related to formulating multi-stage decision making processes, evaluating the predicted effects of certain risk environments and deriving optimal policies. Each different mathematical model may focus on one or more of these aspects. Therefore, it is helpful to organize these models according to their primal functions and applicable fields.

Chapter 2 provides a formulation of a general class of dynamic programs in which there are distinguished subsets of policies and value functions. An algorithm, called generalized policy improvement, is used to find ε-optimal policies. Piecewise linear dynamic programs and partially observable Markov decision processes treated in Sections 2.3 and 2.4 of Chapter 2 are special cases of such simple dynamic programs. In Section 2.6 we show that partially observable Markov decision processes can be transformed into piecewise linear dynamic programs.

In Chapter 3 we show how evaluative and predictive inventory models can be combined in certain special cases of dynamic stochastic models to derive optimal policies. In particular, Section 3.3 analyses an inventory control problem of allocating products between two types of prices.

In Chapter 4 we consider airline seat allocation models with stochastic demands over a discrete time horizon in Section 4.3 and a continous time horizon in Section 4.2, respectively. There is a strong similarity between ordinary inventory control and airline seat management. Airline seats are also inventory products which are perishable or can not be carried over for future use, and the total amount of the products is fixed.

In Chapter 5 we consider the multi-asset version of the consumption and portfolio selection problem that is solved by using stochastic dynamic programming. The analysis begins with a static formulation of the intertemporal model with various risk measures in Section 5.2 and then derives the continous-time formulation for semi-martingale processes in Section 5.3. Since the Black-Scholes model in 1973 was a break through paper in the

field of modern finance theory in the 1970s, we show another derivation of the option pricing formula in Section 5.4.

Chapter 6 is a study of developing a new software reliability growth model focused on software module structures. It is shown in Section 6.2 that software developed by an object-oriented approach has a better quality in terms of the proposed measure of reliability compared with one developed by functional decomposition. In Section 6.3 this model is extended to one predicated on counting processes of instruction executions.

In Chapter 7 we summarize the results drawn from the previous chapters as conslusion.

Chapter 2

Simple Dynamic Programs with Applications

2.1 Introduction

Blackwell [17], Denardo [31], Strauch [118] et al. consider a general class of monotone contractive dynamic programs. In this chapter we consider a special class of Denardo's dynamic programs which satisfies the monotone and contraction assumption. Brumelle [19] and Brumelle and Putterman [20] develops a theory, as well as an algorithm for a state increment dynamic programming which is applied to the continues time model where the state dynamics is described by differential equations. The concepts of "state increment" is similiar to the one of simple partition in this chapter in the sense that a convex polyhedral cell of a simple partition corresponds to rectangular block of a state increment dynamic programming.

In Section 2.2, we consider a class of dynamic programs, based upon Sawaki [101], in which there are distinguished subsets of policies and value functions, respectively called simple policies and simple value functions. An algorithm called generalized policy improvement is used to find ε-optimal policies. This algorithm has the property that only simple functions and policies are generated. When formulated as a dynamic program, it has an uncountable state space. However, the sets of simple policies and simple value functions can be chosen so that they are easily represented in a computer.

Section 2.3 considers a modification of dynamic programs which satisfies the monotone and contraction assumptions (see Sawaki [101]). This class of dynamic programs is

characterized by the piecewise linearity that the cost function is piecewise linear whenever the terminal cost functon is piecewise linear. Sawaki and Ichikawa [107] points out that partially observable Markov decision processes have this property.

An algorithm based on policy improvement is developed to construct ε-optimal policies and ε-optimal cost functions. This algorithm has the advantage of involving only linear functions. A numerical example is also presented.

In Section 2.4 we consider an optimal control problem for partially observable Markov decision processes with finite states, signals and actions over an infinite horizon. It is shown that there are ε-optimal piecewise-linear value functions and piecewise-contant policies which are simple. Simple means that there are only finitely many pieces, each of which is defined on a convex polyhedral set. An algorithm based on the method of successive approximation is developed to compute ε-optimal policy and ε-optimal cost. Furthermore, a special class of atationary policies, called finitely transient, will be considered. It will be shown that such policies have attractive properties which enable us to convert a partially observable Markov decision chain into s usual finite state Markov one.

Section 2.5 is related to theoritcal development for more general classes of partially observable Markov and semi-Markov decision processes with imperfect information structures. The approach taken is to considere such processes with imperfect information states in terms of the probability distributions of those states, which themselves form Markov processes and are generated from a Bayes' rule. Those studies have possible applications in inventory control, queuing, machine maintenance problems, etc.

Section 2.6 considers how partially observable Markov decision processes may be trasformed into piecewise linear ones, which have many advantages in that they are easily represented in a computer. Also we refer Sawaki [100] to specify how to find the products of simple partitions on which cost functions are piecewise linear.

2.2 Generalized Policy Improvement for Simple Dynamic Programs

An algorithm for dynamic programs was developed in [11] and [20]. This algorithm, called generalized policy improvement, includes policy improvement [11], [17] and successive approximation [11] as special cases. In this paper we consider a class of dynamic programs, called simple, with the property that the generalized policy improvement algorithm stays within a certain subset of value functions and policies. Simple dynamic programs are defined in Subsection 2.2.1. Conditions that ensure the existence of an ε-optimal policy within the distinguished subset of policies and an algorithm for finding such a policy are given in Subsection 2.2.3.

Piecewise linear dynamic programs, discussed in Subsection 2.2.2 are a special case of simple dynamic programs. In this type of dynamic program the distinguished subsets of value functions and policies used by the algorithm are easily stored in a computer even for uncountable state space problems. Partially observable Markov decision processes [6], [46], [110], [114], are piecewise linear dynamic programs. The piecewise linear structure was first noted by Sondik [115], Sawaki [107] and [98].

2.2.1 Simple Dynamic Programs

A simple dynamic program is a special case of a dynamic program which satisfies the monotonicity and contraction assumptions of Denardo [31]. These assumptions and Denardo's notation are now reviewed. The *state space* Ω is an arbitrary set. Let V be the set of alll bounded real value functions on Ω. An element of V is a *value function.* The norm defined by $\|v\| = \sup\{|v(x)| : x \in \Omega\}$ makes V a Banach space. For u and v in V we write $u \leq v$ if $u(x) \leq v(x)$ for each $x \in \Omega$. The norm of V is monotone in the sense that $0 \leq u \leq v$ implies $\|u\| \leq \|v\|$.

For each $x \in \Omega$ there is a set D_x of *decisions.* Let Δ be the Cartesian product $\times_{x\in\Omega} D_x$. An element $\delta \in \Delta$ is a *policy.* The *return function* h assigns a real number to each triplet $(x, d, v) \in \bigcup_{x\in\Omega}\{x\} \times D_x \times V$. In a Markov decision process the return function $h(x, d, v)$ can be interpreted as the value of choosing decision d when in state x if a terminal reward $v(z)$ is received whenever the pair (x, d) causes a transition to the state z. The return function is assumed to satisfy the contraction and monotonicity assumptions. The *contraction assumption* is that for some $\beta \in [0, 1), |h(x, d, u) - h(x, d, v)| \leq \beta|u - v|$ for each $u \in V, v \in V, x \in \Omega$, and $d \in D_x$. The *monotonicity assumption* is that for each $x \in \Omega$ and $d \in D_x, h(x, d, v) \leq h(x, d, u)$ whenever $u \leq v$ in V. For $\delta \in \Delta$ define $H_\delta : V \to V$ by $(H_\delta v)(x) = h(x, \delta(x), v)$ for $v \in V$ and $x \in \Omega$. Assume for each $v \in V$ there is some $\delta \in \Delta$ such that $H_\delta v = \sup\{H_\delta v; \delta \in \Delta\}$. (Denardo in Corollary 2 of

Theorem 1 [31] gives a useful sufficient condition for this hold.) Define $H_* : V \to V$ by $H_* v = \sup\{H_\delta v : \delta \in \Delta\}$. Here we deviate slightly from Denardo by using H_* instead of A.

An operator $H : V \to V$ is *monotone* if $u \leq v$ implies $Hu \leq HV$, and is a *contraction* if for some $\beta \in [0, 1), \|Hu - Hv\| \leq \beta\|u - v\|$ for each u and v in V. Denardo verifies that H_* and H_δ are monotone contraction operators.

By Banach's fixed point theorem for contractions, for each $\delta \in \Delta$ there is a unique $v_\delta \in V$ such that $H_\delta v_\delta = v_\delta$. The function v_δ is called the *value of the policy* δ. Similarly v^*, called the *optimal value*, is uniquely defined by $H_* v^* = v^*$. Denardo shows that $v^* = \sup\{v_\delta : \delta \in \Delta\}$. If $\|v_\delta - v^*\| \leq \varepsilon$ then δ is an *ε-optimal policy*, and if $\|v - v^*\| \leq \varepsilon$ then δ is an *ε-optimal value function.*

The objects defined so far and the assumptions that have been imposed are collectively called a *contractive monotone dynamic program. A simple dynamic program* is a contractive monotone dynamic program which has a subset of value functions $V' \subseteq V$ and a subset of policies $\Delta' \subseteq \Delta$ which satisfy the following two conditions :

1. $H_\delta v \in V'$ whenever $\delta \in \Delta'$ and $v \in V'$;

2. if $v \in V'$, then there exists some $\delta \in \Delta'$ such that $H_\delta v = H_* v$.

Elements of V' and Δ' are *simple value function* and *simple policies*, respectively.

2.2.2 Piecewise Linear Dynamic Programs

Let R^N be N-dimensional Euclidean space. Any set of the form $\{x \in R^N : Kx < b, Lx <\leq d\}$ where K and L are N by N matrices and b and d are in R^N is called a *convex polyhedron.* Suppose $A \subseteq R^N$. A collection $P = \{B_1, B_2, \ldots, B_m\}$ of subsets of A is a *partition* of A if $B_i \cap B_j = \emptyset$ for $i \neq j$ and if $\bigcup_{i=1}^m B_i = A$. Each member B_i of a partition P is a *cell.* If each cell of a partition is a convex polyhedron, then the partition is *simple.* The product of two partition P_1 and P_2 is $P_1 \cdot P_2 = \{B \cap D : B \in P_1, D \in P_2\}$. The product of $P_1 \cdot P_2 \ldots P_m$ is defined inductively by $\prod_{i=1}^n P_i = P_n \cdot \prod_{i=1}^{n-1} P_i$. Clearly, the finite product of simple partitions is simple. A partition $\hat{P}$ is *finer* than a partition P if each cell of P has a partition which is a subset of $\hat{P}$.

Suppose that $\Omega \subseteq R^N$ and let V be as defined in the previous subsection. A function $v \in V$ is a *piecewise linear* if there exists a simple partition $\{B_1, B_2, \ldots, B_m\}$ of Ω and a set of vectors $\{v_1, v_2, \ldots, v_m\}$ such that $v(x) = v_i \cdot x$ for $x \in B_i, i = 1, 2, \ldots, m$. *Piecewise affine* functions are defined analogously as functions which are affine on each cell of a simple partition. A policy $\delta \in \triangle$ is *piecewise constant* if there is a simple partition $\{B_1, B_2, \ldots, B_m\}$ of Ω and a set of decisions $\{d_1, d_2, \ldots, d_m\}$ such that $\delta(x) = d_i$ for $x \in B_i$ and $d_i \in \bigcap_{x \in B_i} D_x, i = 1, 2, \ldots, m$. A *piecewise linear dynamic program* is a simple dynamic program with V' as the set of all piecewise linear functions in V and $\triangle'$ as the set of all piecewise constant policies in $\triangle$.

Although the cannonical example for the rest of our paper is a piecewise linear dynamic program and the particular case of a partially observable Markov decision process, other simple dynamic programs are also of interest. The paper by Denardo and Rothblum [32] discusses simple dynamic programs with V' as the set of affine functions in V and $\triangle'$ as the set of constant policies in $\triangle$. An example of theoretical (rather than computational) interest arises when $D_x = D$ for $x \in \Omega$ where D is a measurable space, V' is the set of Borel measurable function in V, and $\triangle'$ is the set of measurable policies in $\triangle$. An economic model motivated by Walras [130] provides still another example. Let Ω be the set of possibleprice vectors of N securities (or N commodities). Assume that there is a finite set D of decisions, each of which can be implemented at any $x \in \Omega$. For each $d \in D$ there is a corresponding stochastic matrix P_d. If the price of vector is x in one period, then it is $P_d x$ in the next period if decision d is chosen. The return function is $h(x, d, v) = r_d \cdot x - \xi_d + \beta v(P_d x)$ for $x \in \Omega, d \in D, v \in V, 0 \leq \beta \leq 1$, and vector $r_d \in R^n$ and $\xi \in R$. The term $r_d \cdot x - \xi$ is the immediate reward if decision d is chosen while in the state x. Since $h(x, d, v)$ is piecewise affine in x whenever v is piecewise affine, it can be shown by an argument analogous to that in Theorem 2.2.1 which follows that this is a simple dynamic program with V' as the set of piecewise affine functions and $\triangle'$ as the set of piecewise constant policies.

Our motivation for studying piecewise linear and piecewise constant functions is that they can be conveniently represented in a computer. This property is shared by some other possible choices of V' such as piecewise affine or even piecewise plynominal. For example, a simple partition $\{B_1.B_2, \ldots, B_m\}$ can easily be stored in computer. Each cell

of B_i of the partition is characterized by a list of inequalities. An inequality consists of a vector, a number which is the righthand side, and an indication of the type of inequality. A piecewise linear function v requires, in addition to partition, a vector v_i for each cell of partition. Affine, piecewise affine, and piecewise constant functions (e.g. policies) can similarly be stored in a computer. The intersecton of two cells can be performed by combining the corresponding lists of inequalities. Thus it is easy to form product partitions. To avoid replicating a list of inequalities which is in several cells, it is convenient to address the lists indirectly. Emptyness of a cell can be checked using a Phase I linear program. Since the number of variables (N if $\Omega \subseteq R^N$), we actually check the dual problem for unboundedness.

The next theorem provides a sufficient condition for a monotone contractive dynamic program to be a piecewise linear dynamic program. The proof is contructive and provides an algorithm for computing $H_\delta v$ for $\delta \in \triangle'$ and $v \in V'$ and an algorithm for computing $H_* v$ and finding $\delta \in \triangle'$ such that $H_\delta v = H_* v$ for $v \in V'$. In order to code these algorithms, a subroutine computing $h(\cdot, d, v)$ for $d \in D$ and piecewise linear v is needed. An example of an algorithm evaluating a return function is described in Subsection 2.2.3.

Theorem 2.2.1 *Suppose that a monotone contractive dynamic program has the property that $\Omega \subseteq R^N$ and that for each $x \in \Omega, D_x$ is the same finite set $D = \{1, 2, \ldots, p\}$. Let V' be the set of piecewise linear value functions and let $\triangle'$ be the set of piecewise contant policies. If $h(\cdot, d, v) \in V'$ for each $d \in D$ and $v \in V$ then the dynamic program is piecewise linear.*

Proof Choose $v \in V'$ and $\delta \in \triangle'$. Suppose $\delta(x) = d_i$ for $x \in B_i$ where $\{B_1, B_2, \ldots, B_m\}$ is a simple partition of Ω. For $i = 1, 2, \ldots, m$ the return function $h(\cdot, d_i, v)$ is piecewise linear, say $h(x, d_i, v) + w_{ij} \cdot x$ for $x \in C_{ij}$ where $\{C_{i1}, C_{i2}, \ldots, C_{in}\}$ is a simple partition of Ω. Let $P_i = \{C_i^j \ capB_i; j = 1, 2, \ldots, n\}$. Note that P_i is a simple partition of B_i and that $P = \bigcup_{i=1}^m P_i$ is a simple partition of Ω In addition, $(H_\delta v)(x) = w_{ij} \cdot x$ for $x \in B_i \cap C_{ij}$ which is a cell in P. Thus $H_\delta v \in V'$.

Let $v \in V'$. We next show how to find $\delta \in \triangle'$ such that $H_\delta v = H_* v$. For $d \in D, h(\cdot, d, v)$ is piecewise linear, say $h(x, d, v) = r_{jd} \cdot x$ for x in the j-th cell of a simple partition P_d. Form the product partition $X_{d=1}^p P_d = P$. Let the cellls of $P = \{B_1, B_2 m \ldots, B_m\}$

and let $\alpha_{id} = \gamma_{jd}$ if B_i is a subset of the j-th cell of P_d. For each $d \in D$, P is finer than P_d so that, $h(x, d, v) = \alpha_{id} \cdot x$ for $x \in B_i, i = 1, 2, \ldots, m$. For $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, p$ define the convex polyhedrons

$$G_{ij} = \{x \in B_i; \alpha_{ij}x > \alpha_{id}x \quad \text{for} \quad d = 1, 2, \ldots, j-1 \quad \text{and} \quad \alpha_{ij}x > \alpha_{id}x \quad \text{for} \quad d = j+1, \ldots, p\}.$$

Then $Q_i = \{G_{ij} : j = 1, 2, \ldots, p\}$ is a simple partition of B_i and $Q = \bigcup_{i=1}^{m} Q_i$ is a simple partition of Ω with the property that

$$(H_*v)(x) = \alpha_{ij} \cdot x \quad \text{if} \quad x \in G_{ij} \quad \text{which is a cell of } Q.$$

The policy $\delta \in \Delta'$ defined by $\delta(x) = j$ for $x \in G_{ij}$ satisfies $H_\delta v = H_* v$.

2.2.3 The Generalized Policy Improvement Algorithm for Simple Dynamic Programs

First we review some properties of iterates of operators from Denardo [31]. If H is a contraction operator in V with contraction coefficient β, then for each $v \in V$, $\|H^n v - \hat{v}\| \to 0$ as $n \to \infty$ where $\hat{v}$ is the unique fixed point of H. This algorithm for approximating the fixed point $\hat{v}$ of H is called successive aproximation. A termination criterion is given by

$$\|v - Hv\| \le (1-\beta)\varepsilon \quad \text{implies} \quad \|v - \hat{v}\| \le \varepsilon \tag{2.1}$$

An upper bound on the number of iterations starting from v required to obatin an ε-approximation to$\hat{v}$ can be derived from $\|v - Hv\| \le (1-\beta)\varepsilon/\beta^n$ implies $\|H^n v - \hat{v}\| \le \varepsilon$. Restating this implication explicitly in terms of n, we have

$$\|H^n v - \hat{v}\| \le \varepsilon \quad \text{for} \quad n > \log \frac{(1-\beta)^2}{2\beta\|v - Hv\|} (\log \beta)^{-1}. \tag{2.2}$$

if H is a monotone operator and $v \le Hv$, then $H^n v \le H^{n+1} v$ for $n = 0, 1, 2, \ldots$ So in this case successive approximation generates a monotone sequence of functions which coverge to the fixed point of H.

In a simple dynamic program successive approximation provides a means of approximating either v_s or $v^*$by iterating H_δ or H_*, respectively, until (2.1) is satisfied. If $v \in V'$, then $H^n v \in V'$ for each n. If in addition $\delta \in \Delta'$, then $H_\delta^n v \in V'$ for each N. Thus the

following algorithm, called generalized policy improvement by Brumalle [19], only involves functions in V' and policies in Δ'.

Algorithm

Step 0 Start with $v_0 \in V'$ satisfying $v_0 \le H_* v_0$. Set $n = 0$.

Step 1 Find $\delta_n \in \Delta'$ such that $H_{\delta_n} v_n = H_* v_n$.

Step 2 If $\|v_n - H_{\delta_n} v_n\| \le (1-\beta)\varepsilon$ then go to Step 4.

Step 3 Otherwise choose some positive integer k_n and evaluate $v_{n+1} := H_{\delta_n}^{k_n} v_n$. Increment n by 1 and go to Step 1.

Step 4 δ_n is an ε-optimal policy and the value functions v_n and $H_* v_n$ are ε-optimal ($v_n \le H_* v_n \le v^*$).

As noted above, v_{δ_n} can be approximated by iterating H_{δ_n}.

Provided that a v_0 with the properties specified in the Step 0 can be found, the other steps can be performed by the definition of a simple dynamic program. We next argue thet the termination criterion in Step 2 will eventually be satisfied and that δ and v_{n+1} have the properties stated in the Step 4. Since v^* is the unique fixed point of H_*, it follows by Theorem 2.2.2 of Brumelle [19] that v_n is increasing and $\sup_n v_n \le v^*$; by Theorem 3.2 [19] $H_n^* v_0 \le v_n \le v^*$; since $H_*$is a contraction operator, $\lim_n H_*^n v_0 = v^*$; and since the norm of V is monotone, $\lim_n v_n = v^*$. Consequently, the termination criterion in Step 2 will eventually be satisfied, and by (2.1)it ensures that v_n is an ε-optimal value function. By Theorem 2.2.2 [19],
$v_n \le H_{\delta_n} v_n \le H_{\delta_n}^2 v_n \le \ldots \le v^*$. But $H_{\delta_n}^k v_n \to v_{\delta_n}$. Thus $v_n \le v_{\delta_n} \le v^*$ and v_{δ_n} is an ε-optimal policy. In addition to showing that the algorithm converges, this argument verifies the following theorem.

Theorem 2.2.2 *A simple dynamic program has simple ε-optimal value functions and simple ε-optimal policies, provided that there exists some $v_0 \in V'$ such that $v_0 \le H_* v_0$.*

If each $k_n = 1$, then the algorithm reduces to successive approximation and Step 1 becomes: evaluate $H_* v_n$. If in Step 2 $\lim_{k \to \infty} H_{\delta_n}^k v_n = v_{n+1}$ can be evaluated, then

$v_{n+1} = v_{\delta_n}$ and the method is policy improvement. It is for this reason that the method is called generalized policy improvement. However, V' is not necessarilly closed in V, and v_δ is not necessarilly in V' even for $\delta \in \Delta'$. Thus k_n must be finite in order to involve only functions in V'.

The question of how best to establish the appropriate values of the parameters k_n in the algorithm is succesive approximation which converges linearly by (2.1). However, the effort per iteration is small. If each $k_n = \infty$, then the policy improvement is known to converge quadratically in some situations [20], [85]. However, $k_n = \infty$ might take us outside of V'. It seems reasonable to take k_n small, perhaps even 1, in the early iterations, and then to later increase k_n so that $U_{\delta_N}^{k_n} v_n$ approximates v_{δ_n} in order to take advantage of the super-linear canvesgence.

In the remainder of the paper we discuss the implimentation of the algorithm for piecewise linear programs and for partially observable Markov decision processes. Provided the return function can be computed, Theorem 2.2.1 provides algorithms for performing Steps 1 and 3.

We next show how to compute $\|v_n - v_{n+1}\|$. Let P_n and P_{n+1} be simple partitions corresponding to the piecewise linear functions v_n and v_{n+1}. Then each function is piecewise linear with respect to $\{B_1, B_2, \ldots, B_m\} := P_n \times P_{n+1}$. Let $v_n(x) = w_i \cdot x$ and $v_{n+1}(x) = w_i' \cdot x$ for $x \in B_i$. The quantities $M_i = \max\{|w_i \cdot x - w_i' \cdot x| : x \in B_i\}$ can be computed by linear programming. Thus $\|v_n - v_{n+1}\| = \sup M_i$ can be computed. Step 2 need not be performed each iteration. Since (2.2) provides an upper bound on the number of remaining iterations in terms of $\|v_n - v_{n+1}\|$, a reasonable procedure would be to compute this upper bound and then do some fraction, say 10%, of third number of iterations before next checking the termination criterion in Step 2.

In Step 0 a suitable v_0 must be found. For a partially observable Markov decision process one can choose $v_0(x) = -M/(1-\beta)$ for each $x \in \Omega$, where $M = \max\{|r(d,i)| : d \in D, i \in S\}$.

2.3 A Modification of Piecewise Linear Dynamic Programs and Their Applications

First, we shall formulate a general dynamic programming problem under the setting of Denardo [31]. Secondly, a piecewise linear dynamic program will be defined. It is a special class of general dynamic programs which satisfies the monotonicity and contraction assuptions.

The state space Ω is an arbitrary set of a real linear space. For each $x \in \Omega$ there is a set A_x of actions. Let Δ be the Cartesian product $X_{x \in \Omega} A_x$. An element $\delta \in \Delta$ is a policy. There is always an optimal stationary policy among a general class of policies in a contractive monotone dynamic program by Denardo [31] or Blackwell [17]. It suffices to consider only the class of stationary policies. Let V be the set of all bounded real valued functions on Ω. An element of V is a cost function. V is a Banach space with the norm $\|v\| = \sup_{x \in \Omega} |v(x)|$. For $u, v \in V$ if $u(x) \leq v(x)$ for all $x \in \Omega$. The loss function h is defined to be a mapping from $U_{x \in \Omega} x \times A_x \times V$ to a real number. Our objective function to be minimized is somehow ambigous, unless that the loss function h is specified. In a Markov decision process, however, $h(x,a,v)$ can be written as $h(x,a,v) = c(x,a) + \beta \int_\Omega v(y) q(dy|x,a)$ where $c(x,a)$ is the immediate cost, β the discount factor and $q(\cdot|x,a)$ the transition probability on Ω given x and a. Therefore, note that the system dynamics as well as the objective function is concealed behind our formulation. Assume that the loss function satisfies the monotonicity and contrction assumptions, that is for each $x \in \Omega$ and $a \in A_x h(x,a,u) \leq h(x,a,v)$ whenever $u \leq v$ in V, and for some $\beta \in [0,1), |h(x,a,u) - h(x,a,v)| \leq \beta \|u - v\|$ for each $u, v \in V, x \in \Omega$ and $a \in A_x$. For $\delta \in \Delta$ define $U_\delta * V \to V$ by $(U_\delta v)(x) = h(x, \delta(x), v)$ for $v \in V$ and $x \in \Omega$. Assume that there is some $\bar{\delta} \in \Delta$ such that $U_{\bar{\delta}} v = \inf_{\delta \in \Delta} U_\delta v$. Also, define $U_* : V \to V$ by $U_* v = \inf_{\delta \in \Delta} U_\delta v$. If $\delta(x) = a$ for each $a \in \Omega$, then we write $U_a = U_\delta$. Denardo [31] verifies that U_* and U_δ are monotone contraction operators. By Banch's fixed point theorem, for each $\delta \in \Delta$ there is a unique $v \in V$ such that $U_\delta v^\delta = v^\delta$. Similarly there is $v^* \in V$ such that $U_* v^* = v^*$. Such v^δ and v^* are called the cost of the policy δ and the optimal cost, respectively. Denardo [31] shows that $v^* = \inf_{\delta \in \Delta} v^*$. If $\|v^\delta - v^*\| \leq \varepsilon$, then δ is an ε-optimal policy, and if $\|v - v^*\| \leq \varepsilon$, then v is an ε-optimal cost function. Our purpose is to find such ε-optimal policy and ε-optimal cost function.

Any set of the form $\{x \in \Omega : \ell_{ij}(x) < (or \leq) d_j, j = 1,2,\ldots,n_i\}, i = 1,2,\ldots,m$, where ℓ_{ij} is a linear functional and d_j a real number is called a convex polyhedron a collection $P = \{E_1, E_2, \ldots, E_m\}$ of subsets of Ω is apartition if $E_i \cap E_j = \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^m E_i = \Omega$. Each member of a partition is a cell. m is the number of cells in partition. If each cell of a partition is a convex polyhedron, then the partition is called simple. The product of two partitionP_1 and P_2 is $P_1 \cdot P_2 = \{E \cap D : E \in P_1, D \in P_2\}$. The product of $P_1 \cdot P_2 \cdots P_m$ is defined by $\prod_{i=1}^m P_i = P_m \cdot \prod_{i=1}^{m-1} P_i$. Plainly, the finite product of simple partitions is again simple. A vector valued function v on Ω is piecewise linearif there exists a simple partition $\{E_1, E_2, \ldots, E_m\}$ of Ω and m linear functions $v_1, v_2, \ldots, v_m$ such that $v(x) = v_i(x)$ for all $x \in E_i, i = 1,2,\ldots,m$. A piecewise linear contant policy is simple and easily reprensented in computer. For example a a bang bang control is such piecewise contant policy. The paper Denardo and Rothblum [32] discusses affine (but not piecewise) dynamic programs.

Although v^* is not necessarily piecewise linear and δ^* is not necessarily piecewise constant, we will show for a class of dynamic program having the structure described in the following assumption that there are ε-optimal piecewise linear cost function and piecewise contant policies.

Assumptions I For each a, $(U_a v)(x)$ is piecewise linear on Ω, provided that v is piecewise linear on Ω.

The following theorem shows that the structure in Assumption I implies how U_* and U_δ preserve the pieecewise linearity of loss functions and the piecewise contant of policies. Assume from now on that $A_x = A = \{a_1, a_2, \ldots, a_p\}$ for all $x \in \Omega$ is finite.

Theorem 2.3.1 *Suppose that Assumption I holds that v is piecewise linear. Then*

(i) $U_\delta v$ is pievewise linear whenever δ is piecewise contant;

(ii) $U_* v$ is piecewise linear; and

(iii) there exists a piecewise constant policy δ such that $U_\delta = U_* v$.

Proof

(i) Suppose that δ is piecewise constant with respect to a simple partition $\{E_i\}$. Let E_i be an arbitrary but fixed cell from the partition and suppose that $\delta(x) = a$ for $x \in E_i$. then

$$(U_\delta v)(x) = (U_a v)(x) \quad for x \in E_i.$$

From Assumption I, $U_a v$ is piecewise linear for each a. Hence $U_\delta v$ is piecewise linear on each cell E_i, and is consequently piecewise linear on Ω.

(ii)-(iii) The functions $U_a v$ are piecewise linear by Assumption I. Suppose that $U_a v$ is piecewise linear with respect to the simple partition P_a. Let $P = \prod_{a \in A} P_a$. Then P is finer than each P_a, and also each $U_a v$ is piecewise linear with respect to P. For each $F \in P$ and $a \in A$, there is some linear functional α_F^a such that

$$(U_a v)(x) = \alpha_F^a(x) \quad \text{for} \quad x \in F.$$

For each $F \in P$, define the sets $G_F^b, b \in A = \{1,2,\ldots,p\}$, by $G_F^b = \{x : \alpha_F^b x < \alpha_F^a x, a = 1,2,\ldots,b-1$ and $\alpha_F^b x \leq \alpha_F^a x, a = b+1,\ldots,p\}$. Then $\{G_F^a : a \in A\} = P_F$ is a partition of F and $\hat{P} = \prod_{F \in P} P_F$ is a partition of Ω with the property that

$$(U*_v)(x) = \alpha_F^a(x) \quad \text{if} \quad x \in G_F^a \in \hat{P}.$$

The policy δ defined by $\delta(x) = a$ for $x \in G_F^a \in \hat{P}$ satisfies $U_\delta v = U_* v$.

Corollary Suppose that Assumption I holds and that $v^0 \in V$ is piecewise linear.

(i) Define $v^n(x) = (U_\delta v^{n-1})(x), n = 1,2,\ldots$, for piecewise constant δ.

(ii) Define $v^n(x) = (U_* v^{n-1})(x), n = 1,2,\ldots$.
Then v^n is piecewise linear and there exists a piecewise constant stationay policy δ_n satisfying $U_{\delta_n} v^{n-1}$.

We next consider the effects of iterating monotone contraction mappings such as U_* and U_δ, citing some results of Denardo [31].

Lemma 2.3.1 *Suppose that U is a contraction mapping on V with contraction coefficient $\beta < 1$. Let $v^0 \in V$ be given and define the functions $v^n, n = 1, 2, \ldots$ by*

$$v^n(x) = (Uv^{n-1})(x).$$

Then

(i) $\{v^n\}$ converges in norm to the fixed point $\hat{v}$ of U; i.e., $U\hat{v} = \hat{v}$.

Now assume that U is also monotone.

(ii) If $v^1 \leq v^0$, then $\{v^n\}$ is monotonically decreasing to $\hat{v}$.

(iii) If $v^1 \geq v^0$, then $\{v^n\}$ is monotonically increasing to $\hat{v}$.

Remarks 1 *The fixed point $\hat{v}$ need not to be piecewise linear since the cells in the limiting partition are not necessarily finite in number nor polyhedral.*

Examples

Model 1. A markov decision process (Blackwell [17])

Let Ω be a bounded convex polyhedron in R^N and the loss function $h(x, a, v) = c(x, a) + \beta \int_\Omega v(x')q(dx'|x, a)$ as mentioned in the preceding section. Assume that $c(x, a) = c^a \cdot x$, which may be interpreted to be the expectation of c^a if x is a probability vector. Also assume that for each convex polyhedron $B \subset \Omega$

$$q^a(B, x) = \int_B x'q(dx'|x, a)$$

is piecewise linear in x with respect to a simple partition $P^a(B) = \{E_j(a, B), j = 1, 2, \ldots, m_{a,B}\}$ for each a where the integral of the vector x' is defined componentwise. These two assumptions imply Assumption I.

We explicitly check that Assumption I is satisfied. Let $a \in A$ be arbitrary but fixed and suppose that v is piecewise linear with respect to a simple partition $\{E_i, i = 1, 2, \ldots, m\}$. Let $P^a = \prod_{i=1}^m P^a(E_i) = \{\bar{E}_j^a; j = 1, 2, \ldots, r\}$, the product partition, which is again simple.

$$\begin{aligned}
(U_a v)(x) &= c^a \cdot x + \beta \int_\Omega v(x')q(dx'|x, a)
\end{aligned}$$

$$\begin{aligned}
&= c^a \cdot x + \beta \sum_{i=1}^m \int_{E_i} (v_i x')q(dx'|x, a) \\
&= c^a x + \beta \sum_{i=1}^m v_i \cdot (\int_{E_i} x'q(dx'|x, a)) \\
&= c^a \cdot x + \beta \sum_{i=1}^m v_i q^a(E_i, x) \\
&= [c^a + \beta \sum_{i=1}^m v_i \lambda_{ij}^a]x \quad \text{for} \quad x \in E_j(a, E_i)
\end{aligned}$$

where $\lambda_{ij}^a \cdot x = q^a(E_i, x)$ for $x \in E_j(a, E_i)$ and the index j depends on i for each $a \in A$. The third equality is obtained from the fact that the integral of the inner product is equal to the inner product of the integral if v_i does not depend on x, a and each componentwise integral is well defined. $U_a v$ is linear on each $\bar{E}_j^a$. Hence $U_a v$ is piecewise linear with respect to the simple partition $P^a = \{\bar{E}_j^a, j = 1, 2, \ldots, r\}$, which satisfies Assumption I.

Model 2. A partially observable Markov Decision Process (Sawaki and Ichikawa [107], Dynkin [39])

We will show that a partially observable Markov decision process is a special case of model 1. Consider a Markov decision process with state space $\{1, 2, \ldots, N\}$, with finite action set A, with the probability transition matrices p^a and with immediate cost vectors c^a. Let Z_n be the state at the n-th transition. Assume that the process $\{Z_n, n = 0, 1, 2, \ldots\}$ cannot be observed, but at each transition a signal θ is transmitted to the decision maker. The set of possible signals θ is assumed to be finite. For each n, given that $Z_n = j$ and that action a is to be implemented, the signal θ_n is independent of the history of the signals and actions $\{\theta_0, a_0, \theta_1, a_1, \ldots, \theta_{n-1}, a_{n-1}\}$ prior to the n-th transition and has conditional probability denoted by $\gamma_{j\theta}^a = P[\theta_n = \theta | Z_n = j, a_{n-1} = a]$.

Let $\Omega = \{x = (x_1, x_2, \ldots, x_N) : \sum_{i=1}^N x_i = 1, x_i \geq 0, \forall_i\} \subset R^N$. Define the i-th component of X_n, the random variable of x, to be

$$P[Z_n = i | \theta_0, a_0, \theta_1, a_1, \ldots, \theta_{n-1}, a_{n-1}, \theta_n], \quad i = 1, 2, \ldots, N.$$

It can be shown (see Dynkin [39]) that

$$P[Z_{n+1} = j | \theta_0, a_0, \theta_1, \ldots, \theta_n, a_n, \theta_{n+1}] = P[Z_{n+1} = j | \theta_{n+1}, a_n, X_n].$$

Thus X_n represents a sufficient statistics for the complete past history $\{\theta_0, a_0, \ldots, a_{n-1}, \theta_n\}$. It follows that $\{X_n : n = 0, 1, 2, \ldots\}$ is a Markov process (see Dynkin [39]), called the observed process. Its immediate cost is $c(x,a) = c^a \cdot x$. Its probability transition function is determined by the following calculation. For each measurable subset $B \subseteq \Omega, x \in \Omega$, and $a \in A$,

$$\begin{aligned}
q(B|x,a) &= P[X_{n+1} \in B | X_n = x, a_n = a] \\
&= \sum_\theta P[X_{n+1} \in B | \theta_{n+1} = \theta, X_n = x, a_n = a] \\
&\quad \cdot \sum_j P[\theta_{n+1} = \theta | Z_{n+1} = j, X_n = x, a] \cdot P[Z_{n+1} = j | X_n = x, a_n = a] \\
&= \sum_\theta P[X_{n+1} \in B | \theta_{n+1} = \theta, X_n = x, a_n = a] \\
&\quad \cdot \sum_j \gamma^a_{j\theta} \sum_i P[Z_{n+1} = j | Z_n = i, X_n = x, a_n = a] P[Z_n = i | X_n = x, a_n = a] \\
&= \sum_\theta P[X_{n+1} \in B | \theta_{n+1} = \theta, X_n = x, a_n = a] \sum_j \gamma^a_{j\theta} \sum_i P^a_{ij} x_i \\
&= \sum_\theta P[X_{n+1} \in B | \theta_{n+1} = \theta, X_n = x, a_n = a] \mathbf{1} P^a(\theta) x
\end{aligned}$$

where $\mathbf{1} = (1, 1, \ldots, 1)$ and $P^a(\theta) = [P^a_{ij(\theta)}] = [P^a_{ji}\gamma^a_{i\theta}]$.

Define the vector $T(x|\theta, a)$ by

$$T(x|\theta,a) = \frac{P^a(\theta)x}{\mathbf{1}P^a(\theta)x}$$

Note that $T(X_n|\theta, a) = X_{n+1}$, and that

$$P[X_{n+1} \in B | \theta_{n+1} = \theta, X_n = x, a_n = a] = \begin{cases} 1, & \text{if} \quad T(x|\theta,a) \in B \\ 0, & \text{if} \quad \text{otherwise} \end{cases}$$

So,

$$q(B|x,a) = \sum_{\theta \in \phi^a(B,x)} \mathbf{1}P^a(\theta)x$$

where $\phi^a(B,x) = \{\theta : T(x|\theta,a) \in B\}$.

Finally, we show that the observed process $\{X_n\}$ is a special case of Model 1; i.e., $q^a(B,x) = \int_B x' q(dx'|x,a)$ is piecewise linear in x for each convex polyhedral set $B \subset \Omega$ and action $a \in A$. Using the previously computed $q(B|x,a)$ we have

$$\begin{aligned}
q^a(B,x) &= \int_B x' q(dx'|x,a) \\
&= \sum_{\theta \in \phi^a(B,x)} T[x|\theta,a] \mathbf{1}P^a(\theta)x \\
&= \sum_{\theta \in \phi^a(B,x)} \frac{P^a(\theta)x}{\mathbf{1}P^a(\theta)x} \mathbf{1}P^a(\theta)x \\
&= \sum_{\theta \in \phi^a(B,x)} P^a(\theta)x
\end{aligned}$$

which can be shown to be piecewise linear (see Brumelle and Sawaki [22]).

Theorem 2.3.2 *For each iteration, $n = 0, 1, 2, \ldots$, in the algorithm,*

$$y^n \geq U_{\delta^n} y^n \geq U^2_{\delta^n} y^n \geq \ldots \geq U^{k_n}_{\delta^n} y^n = y^{n+1}.$$

In other words, $\{y^n\}$ is a decreasing sequence.

Proof First, it is true for $n = 0$. Since $y^0 \geq U_{\delta^0} u^0$ and since U_{δ^0} is monotone, it follows that $y^0 \geq U_{\delta^0} y^0 \geq U^2_{\delta^0} y^0 \geq \ldots \geq U^{k_0}_{\delta^0} y^0 = y^1 \geq U_{\delta^0} y^1$. By definition δ^1 satisfies $U_{\delta^1} y^1 = U_* y^1$. However, $U_* y^1 \leq U_{\delta^0} y^1 \leq y^1$, and so not only is the theorem established for $n = 0$, but we have also shown that $U_{\delta^1} y^1 \leq y^1$.

Now suppose $U_{\delta^n} y^n \leq y^n$. The same argument as in the first paragraph establishes the theorem for n and also that $U_{\delta_{n+1}} y^{n+1} \leq y^{n+1}$. Hence the proof is completed by induction.

Corollary $y^n \geq v^*$ for $n = 1, 2, \ldots$.

Proof For an arbitrary $n, y^n \geq U_{\delta^n} y^n \geq U_* y^n$. Since U_* is monotone, $y^n \geq U^j_* y^n$ for each j. By Lemma 2.3.1, $U^j_* y^n$ decreases monotonically and converges to v^* as $j \to \infty$. Consequently, $y^n \geq v^*$ and the proof is complete.

We next show that if the algorithm terminates then it will provide an ε-optimal cost function and an ε-optimal policy.

Theorem 2.3.3 *If $\|y^n - y^{n+1}\| \leq (1-\beta)\varepsilon$, then $\|y^n - v^*\| \leq \varepsilon$, i.e., y^n is ε-optimal. Moreover, δ^n is also ε-optimal and $v^* \leq v^{\delta^n} \leq y^n$.*

Proof Note that $U_{\delta^n} y^n = U_* y^n$ and that by the previous corollary $y^n \geq v^*$.

$$\begin{aligned}
\|y^n - v^*\| &\leq \|y^n - U_* y^n\| + \|U_* y^n - U_* v^*\| \\
&\leq \|y^n - U_{\delta^n} y^n\| + \beta \|y^n - v^*\| \\
&\leq \|y^n - U^m_{\delta^n} y^n\| + \beta \|y^n - v^*\| \quad \text{for} \quad m = 1, 2, \ldots,
\end{aligned}$$

because

$$y^n \geq U_{\delta^N} y^n \geq U^m_{\delta^n} y^n \quad \text{for} \quad m = 1, 2, \ldots. \quad \text{(Theorem 2.3.2.)}$$

Thus

$$(1-\beta)\|y^n - v^*\| \leq \|y^n - U^m_{\delta^n} y^n\| = \|y^n - y^{n+1}\| \leq (1-\beta)\varepsilon,$$

and so $\|y^n - v^*\| \leq \varepsilon$.

The last statement in the theorem follows by Theorem 2.3.2 and Corollary.

Theorem 2.3.4 *Suppose that $\{y^n\}$ is a sequence of costs generated by the algorithm.*

(i) y^n is converges pointwise to $y \in V$.

(ii) $y = U_* y$, i.e., y is optimal.

In other words, the algorithm converges.

Proof

(i) First of all we shall show that $\{y^n\}$ is bounded below. By Theorem 2.3.2 we have $y^n \geq U_{\delta^n}^m y^n$ for each $m = 1, 2, \ldots$. It is well known (see [17] and [31]) that $U_{\delta^n}^m y^n \rightarrow v^{\delta^n}$ as $m \rightarrow \infty$. Therefore $y^n \geq v^{\delta^n}$. From $v^{\delta^n} \geq v^* \in V$, there exists an M such that $\|v^{\delta^n}\| \leq M$. Hence $y^n(x) \geq -M$ for all x. From Theorem 2.3.2 y^n is a decreasing sequence. Hence y^n converges pointwise.

(ii) By a choice of y^0 and Theorem 2.3.2 we know that

1. $y^n \geq U_{\delta^n} y^n \geq U_* y^n$

   To show The other way we have

2.

$$\begin{aligned} y^n &= U_{\delta^{n-1}}^m y^{n-1} && \text{(By definition of } y^n\text{)} \\ &\leq U_{\delta^{n-1}} y^{n-1} && (U_\delta^m y \leq Uy, \quad \forall y \in V) \\ &= U_* y^{n-1} && \text{(By definition of } \delta^{n-1}\text{).} \end{aligned}$$

Then from (1) and (2), we obtain

$$U_* y^n \leq y^n \leq U_* y^{n-1}.$$

From the statement (i) $y^n \rightarrow y$. Since a contraction mapping U_* is continuous, $U_* y^n \rightarrow U_* y$. Therefore, we must have

$$U_* y = y$$

which completes the proof. □

A Numerical Example

This subsection presents a numerical example for Model 2, partially observable Markov decision processes, especially in the case of two dimensions, $\Omega = \{(x_1, x_2) | x_1 + x_2 = 1, x_1, x_2 \geq 0\}, A = \{1, 2\}$ and $\theta = \{1, 2\}$. The necessary data are shown in Table 2.1. To specify the stopping rule we choose $\beta = 0.8$ and $\varepsilon = 0.01$. Therefore, if $\|y^n - y^{n-1}\| \geq 0.002$, then the algorithm stops and y^n is ε-optimal.

Set $x_1 = x$. To start the algorithm an initial piecewise constant policy δ^0 and an initial piecewise linear function y^0 satisfying $y^0 \geq U_{\delta^0} y^0$ must be found. Choose a policy δ^0 minimizing $c^a \cdot x$; thus $\delta^0(x) = 1$ if $x \leq \frac{2}{3}$, $\delta^0(x) = 2$ if $x > \frac{2}{3}$. Set an initial cost function $y^0(x) = (0, 0)(x, 1 - x)^T$ with the partition $\{[0, 1]\}$, which is piecewise linear and satisfies $y^0 \geq U_{\delta^0} y^0$. Also set $k_n = 1$ for all n. The computational results programmed in FORTRAN are shown in Table 2.2. We may observe from Table 2.2 that the algorithm converges at period $n = 35$, and an ε-optimal cost is $-15.166 - 3.826x$ if $x \leq 0.571$ and $-16.732 - 1.086x$ if $x > 0.571$. An ε-optimal policy $\delta^{30}(x) = 1$ if $x \leq 0.571$ and $\delta^{30}(x) = 2$ if $x > 0.571$. Table 2.2 also shows that an ε-optimal policy converges (at $n = 10$) much faster than an ε-optimal cost does.

The goal of this section is to generate and construct ε-optimal cost and ε-optimal policies in a sequential fashion for a general class of dynamic programs. Toward this end we have taken advantage of the properties of piecewise linear cost functions and piecewise constant policies. These properties guarantee that the algorithm involves only piecewise linear and constant functions which belong to the class of linear programs. Finally we should also emphasis the importance of the algorithm capable for solving continuous state dynamic programs. Many sequential decision problems under uncertainty often turn out to have a probability vector as their state space, which is no longer finite nor countably infinite, but continuous. Therefore, the algorithm developed in this section will become more important in the field of sequential decision problem under uncertainty.

Table 2.1: Data for A Numerical Example

actions	c^a	p^a		γ^a	
a = 1	(-5,-1)	0.7	0.3	0.75	0.25
		0.9	0.1	0.60	0.40
a = 2	(-4,-3)	0.5	0.5	0.30	0.70
		0.4	0.6	0.40	0.60

Table 2.2: A List of Optimal Values and Partitions

Periods n	cost function y^n	Policies and	Partitions	$\|y^n - y^{n-1}\|$
1	-1-4x	1	[0.00,0.666]	5.00
	-3-x	2	(0.666,1.00]	
2	-4.12-3.84x	1	[0.00,0.579]	2.96
	-5.179-1.08x	2	(0.579,1.00]	
3	-6.35-3.827x	1	[0.00,0.572]	2.22
	-7.92-1.086x	2	(0.572,1.00]	
5	-9.53-3.826x	1	[0.00,0.571]	1.411
	-11.095-1.086x	2	(0.571,1.00]	
10	-13.324-3.826x	1	[0.00,0.571]	0.462
	-14.889-1.086x	2	(0.571,1.00]	
20	-14.975-3.826x	1	[0.00,0.571]	0.05
	-16.540-1.086x	2	(0.571,1.00]	
30	-15.152-3.826x	1	[0.00,0.571]	0.005
	-16.717-1.086x	2	(0.571,1.00]	
35	-15.166-3.826x	1	[0.00,0.571]	0.001
	-16.732-1.086x	2	(0.571,1.00]	

2.4 Optimal Control for Partially Observable Markov Decision Processes over an Infinite Horizon

The partially observable Markov process, intoduced by Dynkin [39], consists of two stochastic processes, the core process $\{X_n, n = 1, 2, \ldots\}$, which cannot directly oberved, and the signal process $\{S_n, n = 1, 2, \ldots\}$ which becomes known at each decision epoch $n = 1, 2, \ldots$. The core process is a Markov chain and the signal process is probabilistically related to the core process by the conditional probability $\gamma_{i\theta}$ of observing a signal θ given that the core process is in state i. Dynkin shows that the state occupancy probability represents a sufficient statistic for the complete past history. Åström [5] also considered a similar model with finite states and finite actions over a finite horizon, using the method of successive approximation to find ε-optimal cost vectors, however, it is only applicable to problems in two dimensions. Smallwood and Sondik [114] have independently obtained similar results. Later Sondik [116] extended this model to the infinite horizon and introduced the class of finitely transient policies. White [131] has considered a partially observable semi-Markov process with a finite horizon where the controler konows the times of the core process transition. Sawaragi and Yoshikawa [110] also studied the partially observable control problem with countable states, uncountable action sets and infinite horizon, where they have explicitly showed that such partially observable models can be transformed into an ordinary complete observable one.

In this section, under the setting of [114], we shall consider an optimal control problem with discounted cost over an infinite horizon. We introduce three concepts of simple partitions, simple policies, and piecewise linear functions. Using only these concepts we present an algorithm to find an approximation to the optimal cost function. We also show that we can construct an ε-optimal simple stationary policy. We are guaranteed to obtain an ε-approximation of the optimal cost function in finite steps, and each step we only need to find a finite number of vectors by linear programming. Also, an application to a machine maintenance model will be discussed.

Furthermore, in this section a special class called finite transient, of stationary policies will be considered. We shall show that such policies have very attractive properties and are useful for approximating an optimal policy. If policies are finitely transient, partially observable Markov decision processes can be reduced without loss of generality into finite states Markov decision processes with complete observation.

Sondik [116] has originally introduced the concept of finite transientness of policies for the model with finite sets of states, signals and actions over infinite horizon. However, many patrs of his paper are unclear. These will be revised and clarified by giving a different definitions of finitely transient policies. The same notations and symbols as in Sondik's paper are adopted here except where confusion accurs.

2.4.1 Statement of the Problem

Consider a Markov decision process (called the core process) with state set $\Omega = \{1, 2, \ldots, N\}$, with finite action set A with probability transition matrices $\{P^a, a \in A\}$, and with immediate cost vectors $\{q^a, a \in A\}$. Let X_n be the state at the n-th transition. Assume that the process $\{X_n, n = 0, 1, 2, \ldots\}$ cannot be observed, but in each transition a signal is transmitted to the decision maker. The set of possible signals $S = \{1, 2, \ldots, \theta\}$ is assumed to be finite. For each n, given that $X_n = j$ and that action a is to be implemented, the signal θ_n is independent of the history of the signals and actions $\{\theta_0, a_0, \theta_1, a_1, \ldots, \theta_{n-1}, a_{n-1}\}$ prior to the n-th transition and has conditional probability denoted by $\gamma^a_{j\theta} = P[\theta_n = \theta | X_n = j, a]$. At time $n = 0, 1, 2, \ldots$,let $\pi = (\pi_i)$ be the state probability (N-vector). For a transition probability $p^a = (P^a_{ij})$ and an information structure $\Gamma^a = \text{diag}(\gamma^a_{j\theta})$ put $Q^a_\theta = p^a \Gamma^a_\theta$.

If the current state information vector is π, a signal θ is observed and action a has been chosen, then the next state information is given by

$$T(\pi|\theta, a) = \frac{\pi Q^a_\theta}{\{\theta|\pi, a\}} \tag{2.3}$$

where

$$\{\theta|\pi, a\} = \pi Q^a_\theta \mathbf{1}, \quad \mathbf{1} = (1, \ldots, 1)^T.$$

Let

$$\Pi = \{\pi \in R\ N : \sum_{i=1}^{N} \pi_i = 1, \quad \pi_i > 0 \quad \text{for all } i\}$$

We define Δ as the family of mappings $\delta : T \times \Pi \to A$ where $T = [0, \infty)$. Each element of Δ is called a policy. Given an initial distribution $\pi(0)$ and a policy δ, the subsequent

information vectors $\pi(n)$ form a Markov process. Our discounted control problem for an initial distribution $\pi(0)$ is described by

$$\min_{\delta \in \Delta} E_\delta[\sum_{n=0}^{\infty} \beta^n \pi(n) q^{\delta(n, \pi(n))}],$$

where E is the expectation with respect to the signal, $\beta, 0 \le \beta < 1$, is the discount factor and the cost at time n is given by the inner product πq^a with action a. Let $C(\pi|\delta)$ be a *cost of a stationary policy* δ at an initial value π. Then it is well known (see [17], [31]) that $C(\pi|\delta)$ satisfies

$$C(\pi|\delta) = \pi q^\delta + \beta \sum_\theta \{\theta|\pi, \delta\} C(T(\pi|\delta, \delta)|\delta). \tag{2.4}$$

Let $C^*(\pi)$ be the *optimal cost*, then the following is true (see [17], [29]).

Theorem 2.4.1 *There exists an optimal stationay policy δ^* with $C(\pi|\delta = C^*(\pi)$. Also, $C^*(\pi)$ satisfies*

$$C^*(\pi) = \min_{a \in A}\{\pi q^a + \beta \sum_{\delta \in S} \{\delta|\pi, a\} C^*(T(\pi|\delta, a))\} \tag{2.5}$$

for any $\pi \in \Pi$.

An *ε-optimal cost function* C is one satisfying

$$\|c^* - C\| = \sup_{\pi \in \Pi} \|C^*(\pi) - C(\pi|\cdot)\| \le \varepsilon. \tag{2.6}$$

A policy δ such that $C = C(\cdot|\delta)$ satisfying (2.6) is an *ε-optimal policy and its cost function* we define simple partitions, simple pilicies and piecewise linear functions.

Definition 1 A partition $\{V_i\}_{i=1}^m$ of all π is called *simple* if each V_i is a convex polyhedral set, where a convex polyhedral set is the solution set of a finite system of linear inequalities, i.e.,

$$V_i = \{\pi \in \Pi : v_{ij}\pi < 0, j = 1, 2, \ldots, n_i\}$$

where $v_{ij} \in R^N$ and $v_{ij}\pi$ is the inner product of v_{ij} and π.

Remarks 2 *Inequalities of the form $v\pi < 0$ contains those of the form $v\pi < \alpha, \alpha$ scalar. In fact $v\pi < \alpha$ is equivalent to $(v - \alpha\mathbf{1})\pi < 0$.*

Lemma 2.4.1 *Let $P = \{V_i\}$ and $P_2 = \{W_j\}$ be two simple partitions of π. Then, the product partition $P_1 \cdot P_2 = \{V_i \cap W_j\}$ is again simple.*

Proof Here we omit $V_i \cap W_j$ if $V - i \cap W_j = \emptyset$. The sets $V_i \cap W_i$ are disjoint and are convex polihedral sets. Hence $P_1 \cdot P_2$ is simple.

Definition 2 A stationary policy δ is called *simple* with respect to a simple partition $\{V_i\}$ if $\delta(\pi) = a_i$ on $V_i, i = 1, 2, \ldots, m$.

Definition 3 A real valued function f on π is called *piecewise linear* if $f(\pi) = f_i\pi$ on $V_i, i = 1, 2, \ldots, m$, where $\{V_i\}$ is a simple partition $f_i \in R^N$.

Example: Define an information structure as a mapping from the set of states (unobservable) of the core process to the set of distinctive signals δ. The decision maker chooses an information structure from the set of available structures decides upon an action for the system.

Let $a = (a_1, a_2)$ be the pair of actions, a_1 for the system control and a_2 for information acquisition. More precisely, we have

$$P_{ij}^a(\delta) = P_{ij}^{a_1}\gamma_{j\theta}^{a_2}$$

$$\pi q^a = \sum_{i=1} \pi_i \sum_{j=1} P_{ij}^{a1} \sum_{\theta=1}^{\theta} \gamma_{j\theta}^{a_2} q(i, j, a_1, a_2)$$

where $q(i, j, \theta, a_1, a_2)$ is the immediate cost of the core process when a state of the core process moves from i to j and a signal θ observed under actions a_1 for the system and a_2 for the information structure, and $\pi = (\pi_1, \ldots, \pi_N)$ is the probability vector with an interpretation π_i is the probability that the core process is in state i.

Consider a machine maintenance and repair model (e.g. Smallwood and Sondik [114]) as an application of partially observable models. But this model is a modification of Smallwood and Sondik's. The machine consists of two internal components. The states of the core process $X_n = i, i = 1, 2, 3$, have the following interpretation. If $i = 1$, then both components are broken down, if $i = 2$ either one is broken down and if $i = 3$ both of them are working. Assume that the machine produces M finished products at each period and the machine cannot be inspected. The actions a_1 for the machine control are to repair and to repair the machine. The actions a_2 for information acquisition are the numbers of a sample to choose out of the M finished products. The signals δ are the number of defective products in the sample, which forms the signal process $\{\theta_n, n = 1, 2, \ldots\}$. The core process $\{X_n, n = 1, 2, \ldots\}$ is the unknown states of the components of the machine. Let $\pi_i = P\{X_n = i\}, i = 1, 2, 3$ and put $\pi = (\pi_1, \pi_2, \pi_3)$. Then, the process

$\{(X_n, \theta_n), n = 1, 2, \ldots\}$ becomes a partially observable machine maintenace and repair model with actions $a = (a_1, a_2)$ and immediate cost πq^a.

2.4.2 Finitely Transient Policies

In this subsection a special class of simple stationary policies, called finitely transient, will be studied. The class of such policies has very attractive properties even though all stationary policies do not belong to such a class.

Define, for a simple policy δ,

$$D^k = \bigcup_{\delta}\{\pi : T\{\pi|\delta, \delta\} \in D_{k-1}\}, \quad k = 1, 2, \ldots \tag{2.7}$$

where

$$D^0 = \bigcup_i D_i^0 = \bigcup_{i,j}\{\pi \in \Pi : v_{ij}\pi = 0\}$$

which forms the boundary set of the partition $\{V_i\}$ corresponding to a simple policy δ. Let $V^k = \{V_j^k\}_{j=1}^m$ be the collections of sets whose boundaries are $\bigcup_{L=0}^k D^L$ and then V^k is a refinement of V^{k-1}, $k \geq 1$, where $V^0 = \{V_j\}$.

Definition 4 A simple policy δ is called *finitely transient* if there is an integer $k < \infty$ such that

$$T(V_j^k|\theta, \delta) \subset V_{\nu(j,\theta)}^k \quad \text{for all } \theta$$

where $T(V|\theta, \delta) = \{T(\pi|\theta, \delta) : \pi \in V\}$ and $\nu(j, \theta)$ is the index of the set containing $T(\pi|\theta, \delta)$ for $\pi \in V_j^k$.

Lemma 2.4.2 *Let k_δ be the smallest such integer. $D^k = \emptyset$ for all $k \geq k_\delta$ if only if δ is finitely transient with the index k.*

Proof Suppose that δ is finitely transient with the index k_δ, that is,

$$T(V_j^{k_\delta}|\theta, \delta) \subset V_{\nu(j,\theta)}^k \quad \text{for all } \theta.$$

$D^{k_\delta} = \emptyset$ because $T(V_j^{k_\delta}|\theta, \delta)$ is the set of all possible state information at the k_δ-th period and D_i^k is open in Π for all i, k. Let $\mathcal{L}_\delta$ be the set function defined as $\mathcal{L}_\delta(B) = U_\delta\{\pi : T(\pi|\theta, \delta) \in B\}$

$$\begin{aligned} D^k &= \bigcup_{\theta}\{\pi : T(\pi|\theta, \delta) \in D^{k-1}\} \\ &= \mathcal{L}_\delta(D^{k-1}) \\ &= \mathcal{L}_\delta^k(D^0) \end{aligned}$$

If $D^k = \mathcal{L}_\delta^k(D^0) = \emptyset$, then

$$D^{k+1} = \mathcal{L}_\delta(\emptyset) = \emptyset$$

Hence, by induction $D^k = \emptyset$ for all $k \geq k_\delta$.

Conversely, $D^k = \emptyset$ for all $k \geq k_\delta$ and that

$$T(V_j^k|\theta,\delta) \not\subset V_{\nu(j,\theta)}^k \quad \text{for some } \theta.$$

So, there exists $\pi^1, \pi^2 \in V_j^k$ such that for some θ, $T(\pi^1|\theta,\delta)$ and $T(\pi^2|\theta,\delta)$ do not belong to the same set $V_{\nu(j,\theta)}^k$.

Then, there is a constant λ, $0 < \lambda < 1$, such that $\lambda' T(\pi^1|\theta,\delta) + (1-\lambda')T(\pi^2|\theta,\delta) \in D^k$ and λ' is given by

$$\begin{aligned} \lambda' &= \frac{\lambda \pi^i Q_\theta^a}{\pi Q_\theta^a} \\ \lambda' T(\pi^1|\theta,\delta) + (1-\lambda')T(\pi^2|\theta,\delta) &= T(\lambda\pi^1 + (1-\lambda)\pi^2|\theta,\delta). \end{aligned}$$

By letting $\pi = \lambda\pi^1 + (1-\lambda)\pi^2$, we obtain

$$T(\pi|\theta,\delta) \in D^k$$

which is contradiction.

Lemma 2.4.3 *Let $Q_{\theta_1}^{a_1} Q_{\theta_2}^{a_2} \cdots Q_{\theta_k}^{a_k} = (Q_\theta^a)^k$ and $\mathbf{0}$ be a zero row vector. A simple policy δ is finitely transient if there exists an integer $k < \infty$ such that*

$$V_{ij}(Q_\theta^a)^k > \mathbf{0} \quad V_{ij}(Q_\theta^a)^k < \mathbf{0} \quad \text{for all } \theta, a, i, j.$$

Proof

$$\begin{aligned} D^k &= \bigcup_\theta \{\pi : T(\pi|\theta,\delta) \in D^{k-1}\} \\ &= \bigcup_\theta \bigcup_{i,j} \{\pi : v_{ij}(Q_\theta^a)^k \pi = 0\}. \end{aligned}$$

Since $\pi_i \geq 0$ and $\sum_i \pi_i = 1$, $D^k = \emptyset$ if $v_{ij}(Q_\theta^a)^k > \mathbf{0}$ or $\mathbf{0}$ for all θ, a, i, j. By Lemma 2.4.2, this completes the proof. □

Remarks 3 *In Lemmas 2.4.2 and 2.4.3, the assumption concerning δ being simple is crucial. A counter example is presented as follows:*

suppose that there are only two states $N = 2$ and $\pi_2 = 1 - \pi_2 \geq 0$.

Define,

$$\delta(\pi_1) = \begin{cases} a_1, & \text{if } \pi_1 \text{ is rational;} \\ a_2, & \text{otherwise} \end{cases}$$

which is stationary but not simple. Then D^δ is the uncountable discountinuous set which never becomes empty. Therefore, a finitely many partition $\{V_i\}$ does not exists.

Theorem 2.4.2 *Let δ be a simple policy. Then, the following are equivalent.*

(i) δ is finitely transient with the index k_δ.

(ii) $C(\pi|\delta)$ is piecewise linear.

Proof

(i)→(ii) Suppose that we have a finitely many partition $V^k = \{V_j^k\}$ for $k \geq k_\delta$. Let $\overline{C}(\pi|\delta) = \pi\alpha_j$, $\pi \in V_j^k$ and $\alpha_j = q^{a_j} + \beta \sum_\delta Q_\delta^{a_j} \alpha_{\nu(j,\theta)}$. Then

$$\begin{aligned} \overline{C}(\pi|\delta) &= \pi\alpha_j, \quad \pi \in V_j^k \\ &= \pi(q^{a_j} + \beta \sum_\theta Q_\theta^{a_j} \alpha_{\nu(j,\theta)}) \\ &= \pi q^\delta + \beta \sum_\theta \{\pi|\theta,\delta\} \frac{\pi Q_\theta^\delta}{\{\pi|\theta,\delta\}} \alpha_{\nu(j,\theta)} \quad \text{for all } \pi \in V_j^k \\ &= \pi q^\delta + \beta \sum_\theta \{\pi|\theta,\delta\} T(\pi|\theta,\delta)\alpha_{\nu(j,\theta)} \quad \text{for } \delta \text{ finitely transient} \\ &= \pi q^\delta + \beta \sum_\theta \{\pi|\theta,\delta\} \overline{C}(T(\pi|\theta,\delta)|\delta) \\ &\equiv (U_\delta \overline{C}(\pi) \end{aligned}$$

Since $C(\cdot|\delta)$ is unique solution of $U_\delta, C(\cdot|\delta) = \overline{C}(\cdot|\delta)$.

(ii)→(i) From piecewise linearity of $C(\cdot|\delta)$, we have $C(\pi|\delta) = \alpha\pi_j$ for $\pi \in V_j^k$ with the partition $\{V_i^k\}$ for $k \geq k_delta$ and $\delta(\pi) = a_j$, $\pi \in V_j^k$. So

$$C(T(\pi|\theta,\delta)|\delta) = T(\pi|\theta,\delta)\alpha_{\nu(j,\theta)} \quad \text{for } \pi \in V_j^k.$$

Then, we must have $T(\pi|\theta,\delta) \in V_{\nu(j,\theta)}^k$ for all $\pi \in V_j^k$ and all θ.

So

$$T\{V_j^k|\theta,\delta\} \subset V_{\nu(j,\theta)}^k \quad \text{for all } \theta.$$

Corollary If a policy δ is finitely transient with the simple partition $\{V_j\}$, then its cost $C(\pi|\delta)$ can be computed by solving the following equations.

$$C(\pi|\delta) = \pi\alpha_j \quad \text{for all } \pi \in V_j, j = 1, 2, \ldots, m \tag{2.8}$$

$$\alpha_j = q^{a_j} + \beta \sum_\theta Q_\theta^{a_j} \alpha_{\nu(j,\theta)}, \quad j = 1, 2, \ldots, m. \tag{2.9}$$

The proof immediately follows from Theorem 2.4.1. Note that the set of equations (2.9) has a unique bounded solution and that m need not be equal to the number of actions.

2.4.3 Properties of U_a and U_*

This subsection is a study of the properties of U_a and U_*. Most of these properties will be used later in the development of the algorithm to find ε-optimal approximations to C^* and δ^*.

Let F the space of real valued fucntions on Π with sup norm. Then F is a Banach space (B-space). Let Π be equipped with Euclidean norm, and let C be the subset of continuous functions in F. Then C is a closed linear subspace (hence is itself a B-space) of F. Define operators U_a, U_* on F by

$$(U_a f)(\pi) = \pi q^a + \beta \sum_{\theta \in S} \{\theta|\pi, a\} f(T(\pi|\theta, a)), \quad f \in F,$$
$$(U_* f)(\pi) = \min_{a \in A} \{\pi q^a + \beta \sum_{\theta \in S} \{\theta|\pi, a\} f(T(\pi|\theta, a))\}.$$

Lemma 2.4.4

(i) U_a, U_* are contraction mappings with contraction coefficient β.

(ii) U_a, U_* are monotone, i.e., if $f, g \in F$ with $f \leq g$, then $U_* f \leq U_* g$ and $U_a f \leq U_a g$.

(iii) They map C into itself, thus fixed points of these operators are continuous functions.

Proof The properties (i), (ii) are standard. (See [17], [90]). (iii) U_a clearly maps C into self. $U_* f)(\pi)$ is the minimum of finite number of continuous, hence it also continuous, provided f is continuous.

From Lemma 2.4.4 we get some information on C^* and δ^*.

Lemma 2.4.5 *The fixed point of U_* exists and is the optimal cost function C^*, which is continuous.*

Before stating our main results, we need two lemmas.

Lemma 2.4.6 *Let f be a piecewise linear function w.r.t. $\{V_i\}$ on Π. Define a stationary policy δ_f by $U_* f$, namely, $\delta_f(\pi) = a_i$ if a_i minimizes $(U_a f)(\pi)$. Then δ_f is simple.*

Proof Let $\{V_i\}$ be the simple partition for f. Define

$$V_i(a, \theta) = \{\pi \in \Pi : T(\pi|\theta, a) \in V_i\}.$$

Then for each $a, \theta, \{V_i(a, \theta)\}$ is a simple partition. In fact $V_i(a, \theta)$ is given by

$$\frac{\pi Q_\theta^a v_{ij}}{\{\theta|\pi, a\}} < 0, \quad j = 1, 2, \ldots, n_i,$$

or equivalently,

$$\pi Q_\theta^a v_{ij} < 0, \quad j = 1, 2, \ldots, n_i,$$

where v_{ij} characterizes V_i. Let $\{V_{i,\theta}^a\}$ be a simple partition defined by $\cap_{i,\theta} V_i(a, \theta)$ (see Lemma 2.4.1), then $U_a f$ is linear on each $V_{i,\theta}^a$.
More precisely,

$$(U_a f)(\pi) = \pi q^a + \beta \sum_{\theta \in S} \pi Q_\theta^a \sum_i X_{V_{i,\theta}^a}(\pi) f_i,$$

where

$$X_{V_{i,\theta}^a}(\pi) = \begin{cases} 1, & \text{if } \pi \in V_{i,\theta}^a \\ 0, & \text{otherwise,} \end{cases}$$

and f_i is a vector defining f.
Since δ is defined by minimizing finite number of piecewise linear functions, it is simple.

Lemma 2.4.7

(i) If f is piecewise linear, then $U_* F$ is piecewise linear.

(ii) If f is concave, then $U_* f$ is also concave.

Proof $U_a f$ has the same property as f's. By the definition of $U_* f$, the desired results are obtained.

Theorem 2.4.3 *Let $f_0 \in F$, and define*

$$f_n(\pi) = (U_* f_{n-1})(\pi).$$

Let δ_n be the decision rule at stage n defined by $U_* f_{n-1}$.

(i) f_n converge to C_*.

(ii) If f_0 is piecewise linear, then so is f_n for any n. Furthermore, δ_n is simple.

(iii) If f_0 is concave, then f_n is concave.

(iv) if $f_1 \le f_0$, then $f_n \downarrow C^*$. If $f_1 \ge f_0$, then $f_n \uparrow C^*$.

Proof The assertions follow from Lemmas 2.4.4 - 2.4.7

Remarks 4 *If we take $f_0(\pi) = C(\pi|\delta)$ for some stationary policy δ, then $f_n \downarrow C^*$. In particular, if we take $\delta(\pi) = a$ for all π, thus $C(\pi|\delta = f_0(\pi) = \pi(I - \beta P^a)^{-1}q^a$, then f_n is continuous concave and piecewise linear and $f_n \downarrow C^*$. Hence C^* is continuous and concave.*

Remarks 5 *Let $f_0(\pi) = \min_{a\in A} \pi q^a$, then f_0 is piecewise linear, concave and continuous. Hence (ii) and (iii) hold. Since f_n corresponds to the optimal cost for the n-period problem with discounting, this case is essentially equivalent to the results in [116]. If we further assume $q^a \ge 0$ for any $a \in A$, then $f_n \uparrow C^*$.*

Next we shall discuss the rate on convergence.

Lemma 2.4.8 *Let $f \in F$. If $\|f - U_*f\| \le (1-\beta)\varepsilon$, then $\|C^* - f\| \le \varepsilon$.*

Proof

$$\begin{aligned} \|C^* - f\| &\le \|U_*c^* - U_*f\| + \|U_*f - f\| \\ &\le \beta\|C^* - f\| + \|U_*f - f\|. \end{aligned}$$

After arranging, the result is obtained.

Theorem 2.4.4 *If $\beta^n\|f_0 - U_*f_0\| \le (1-\beta)\varepsilon$, then $\|C^* - f_n\| \le \varepsilon$.*

Proof Since we have

$$\begin{aligned} \|f_n - U_*f_n\| &\le \|U_*f_{n-1} - U_*^2 f_{n-1}\| \\ &\le \beta\|f_{n-1} - U_*f_{n-1}\| \\ &\vdots \\ &\le \beta^n\|f_0 - U_*f_0\|, \end{aligned}$$

the theorem follows directly from Lemma 2.4.8.

Remarks 6 *If we calculate $\|f_0 - U_*f_0\|$, then Theorem 2.4.4 tells us when to stop. Furthermore, at each step n we know from $\|f_n - U_*f_n\|$ how many steps (at most) we have to go after step n.*

2.4.4 Algorithm

Since Π is uncountable, it is far from trivial to calculate $C(\pi|\delta)$ which may not be a piecewise linear function of π, except the case that δ is finitely transient. In this subsection we shall approximate $C(\pi|\delta)$ by using the method of successive approximation.

The method of successive approximation is a wellknown and popular method for solving equations. The method is to start with a cost function f_0, and to iterate U_*, constructing a sequence of cost functions $f_n = U_*f_{n-1}, n = 1, 2, \ldots$. By Lemma 2.4.4, U_* is a contraction mapping with fixed point C^* and by Theorem 2.4.3, $\{f_n\}$ converge to c^*. By Theorem 2.4.4, n can be chosen sufficiently large, so that f_n is an ε-optimal cost function. In fact by taking logarithms of the expression in Theorem 2.4.4,

$$n > \log[\frac{(1-\beta)\varepsilon}{\|f_0 - f_1\|} / \log\beta$$

is adequate.

The next theorem provides a means of constructing an ε-optimal policy from an ε'-optimal cost function and specifies the relationship between ε and ε'. The algorithm will first construct an ε'-optimal cost function. From this cost function, an ε-optimal policy is constructed.

Let f_0 be piecewise linear, and let δ_n be defined by U_*f_{n-1}, i.e., $\delta_n(\pi) = a_1$ if a_1 minimizes $(U_a f_{n-1})(\pi)$. Then δ_n is simple, and satisfies $U_*f_{n-1} = U_{\delta_n} f_{n-1}$, where U_δ for a stationary policy δ is defined by

$$(U_\delta f)(\pi) = \pi q^{\delta(\pi)} + \beta \sum_{\theta\in S} \{\theta|\pi, \delta(\pi)\} f(T(\pi|\theta, \delta(\pi))).$$

Theorem 2.4.5 *If $\|C^* - f_{n-1}\| \le \frac{1-\beta}{2\beta}$, then $\|C^* - C(\cdot|\delta_n)\| \le \varepsilon$.*

Proof It is easy to show that U_δ for any stationary policy δ is contraction mapping and that the fixed point is $C(\cdot|\delta)$, i.e., $C(\pi|\delta) = U_\delta C(\cdot|\delta)(\pi)$.
We obtain

$$\|C^* - C(\cdot|\delta_n)\| = \|U_{\delta_n} C(\cdot|\delta_n) - U_*C^*\|$$

$$\leq \|U_{\delta_n}C(\cdot|\delta_n) - U_{\delta_n}C^*\| + \|U_{\delta_n}C^* - U_{\delta_n}f_{n-1}\| + \|U_* f_{n-1} - U_* C^*\|$$
$$\leq \beta\|C(\cdot|\delta_n) - C^*\| + \beta\|C^* - f_{n-1}\| + \beta\|f_{n-1} - C^*\|.$$

Here we used the equality $U_* f_{n-1} = U_{\delta_n} f_{n-1}$. Rearranging the above inequality we obtain

$$(1-\beta)\|C(\cdot|\delta - n) - C^*\| \leq 2\beta\|c^* - f_{n-1}\| \leq (1-\beta)\varepsilon.$$

Hence $\|C(\cdot|\delta_n) - C^*\| \leq \varepsilon$.

If the state space is uncountable, or even countably infinite, then this procedure is difficult to implement on a computer. However, since the partially observable Markov decision process has the structure of piecewise linearity and f_0 is piecewise linear, then each f_n is piecewise linear and each δ_n constructed as in the previous theorem is simple (by Lemma 2.4.6). In this case, the cost functions and policies can be specified by a finite number of items - the inequalities describing each cell of a simple partition and the corresponding action or linear function.

Algorithm to Find an ε-optimal Simple Policy:

(i) Start with any piecewise linear function f_0.

(ii) Compute $f_1 = U_* f_0$.

(iii) Choose an integer n such that

$$\beta^n\|f_0 - f_1\| \leq (1-\beta)\varepsilon',$$

where $\varepsilon' = (1-\beta)\varepsilon/2\beta$. I.e., choose $\hat{n}$ larger than

$$\log[\frac{(1-\beta)^2\varepsilon}{2\beta\|F_0 - f_1\|}]/\log\beta.$$

(iv) Compute $f_n = U^* f_{n-1}$ successively until $n = \hat{n}$.

(v) Consequently, we obtain $f_{\hat{n}}$ such that

$$U_\delta f_{\hat{n}}\| \leq \varepsilon'.$$

(vi) Construct a policy δ satisfying

$$U_\delta f_{\hat{n}} = U^* f_{\hat{n}}.$$

Then δ is ε-optimal.

Remarks 7 *The algorithm can be started with $f_0 \equiv 0$.*

Remarks 8 *The termination criterion, $n \equiv \hat{n}$, in the algorithm has the advantage that $\|f_0 - f_1\|$ is computed only once. However, it has the disadvantage that $\hat{n}$ will probably be larger than necessary, causing unnecessary iterations.*

An alternative would be to compute $\|f_n - f_{n-1}\|$ at each iteration and stop whenever $\|f_n - f_{n-1}\| \leq (1-\beta)\varepsilon'/\beta$. Theorem 2.4.2 guarantees that f_n is an ε'-optimal cost function. However, the computation of $\|f_n - f_{n-1}\|$ will, in general, be expensive.

The best procedure is undoubtedly to check $\|f_n - f_{n-1}\|$ at some, but not all, iterations. For example, $\hat{n}$ might be computed based on $\|f_n - f_{n-1}\|$. Then at some iteration n near $\frac{\hat{n}}{2}$, recompute n based on $\|f_n - f_{n-1}\|$.

2.5 Partially Observable Markov Decision Processes with Abstract Spaces

A number of chapters have resulted from the marriage between the areas of dynamic programming and Markov decision processes. In developing a theory for optimal control it was natural to rely on Markov decision theory by making the assumption that the system can be observed at each stage. In other words, the observer of the system is assumed to have complete information about the state of the system at the time when transitions occur. Such a system is said to be observed under certainty or a complete (perfect) state information.

Since it may be difficult and expensive to obtain such a complete state information, it is more practical to consider a system with an incomplete information state. This, for instance, is how problems of statistics, reliability, relevancy, etc. are described in accounting reports. One interpretation of a Markov decision process under uncertainly is as follows:

Suppose there is an information structure which is a mapping from the set of states of the unobservable system to the set of distinctive signals available, where the states of the system form a Markov process. If the information structure is perfect, there is a one to one and onto mapping which provides an ordinary Markov decision process under certainly, i.e., additional information about the state of the system is not needed. However, if the set of signals is singleton, a completely uncertain situation results.

In this chapter more general classes of incomplete (imperfect) information structures, called partially observable Markov decision processes, will be considered and these will be extended into semi-Markov decision processes. These studies have possible applications in inventory control, queuing, machine maintenance problems, etc. The approach taken is to consider Markov and semi-Markov processes with incomplete information states in terms of the probability distributions of those states, which themselves form Markov processes and are generated from a Bayesian formula. The problem on dynamic programming side is to select an action to be performed, observe the signal generated from the information structure and revise the state information as a result of transitions, in a sequential fashion. They are generally based on a Bayesian formula.

This model formulation is the "best" that can be expected in the sense that no futher information is available about the state of the system.

In this section we shall extend these into semi-Markov decision processes. These are at least two approaches, that is, the discounted approach due to dynamic programming and the nondiscounted one due to average cost criteria. The semi-Markov decision processes are mainly studied using the second approach (Miller [75], Ross [88], Lippman [63]). Ross

2.5.1 Partially Observable Markov and Semi-Markov Processes

Consider a control process, referred to as a partially observable Markov process, which is described by a pair of random sequences $\{X_t, S_t\}$ where the process $\{X_t\}$ is not able to be observed but the signal process $\{S_t\}$ becomes known to the observer at each decision epoch t. The decision maker chooses an information structure from a set of available structures and decises upon an action for the system.

Policies for information acquisition and system control are sought to minimize expected costs over an infinite horizon.

Model Formulation

Let $(\Omega, \mathbf{F}, P)$ be a probability space on which a semi-Markov jump process $\{X_t, 0 \leq t < \infty\}$ mapping from Ω to X a separable metric space is defined, where Ω is a non-empty

Borel subset of a complete separable space and $\mathbf{F}$ is the σ-field with respect to Ω. Let S_t be a random variable mapping from Ω to a signal space S which is assumed to be a Borel subset of a complete separable metric space. Let $(S, \mathbf{S}, \mu)$ be the probability signal space where $\mathbf{S}$ is a σ-field of S and for each $M \in \mathbf{S}$, $\mu(M) = P\{S^{-1}(M)\}$.

It is of interest to interpret the signal space S in such a way that if $s = (\theta, \tau)$ for θ output and τ time between transitions, then $S = \theta \times R^+$ where θ is the set of outputs and R^+ is a non-negative real line, and furthermore $\mu = \mu_1 \times \mu_2$ where μ_1 is the same measure as in Ω, and μ_2 is the counting measure for a Markov chain and a Lebesgue measure for (continuous time) semi-Markov processes. The process is described by the pair of random variables $\{X_t, S_t, 0 \leq t < \infty\}$ where the process $\{X_t\}$ cannot be observed but the process $\{S_t\}$ becomes known to the observer at each time t, $0 \leq t < \infty$.

Let Π be the set of probability distributions of X_n and Π_n be a random variable of the distribution of X_n at the n-th epoch. Let A be the set of actions a which is assumed to be finite. Let $\mathbf{H}_n = \{X_0, a_0, S_0, \ldots, X_n, a_n, S_n\}$ and $H_n = \{S_n = s, a_n = a, H_{n-1}\}$, $n = 1, 2, \ldots$ where H_0 is given. Note that $H_n \subset \mathbf{H}_n$. Assume that there is some decision rule δ_n such that $\delta_n(H_{n-1}) = a_n$ for each n=1,2,$\cdots$. A sequence $R = \{\delta_n\}_{n=1}^{\infty}$ is called a policy and $f = \{\delta\}_{n=1}^{\infty}$ is called a stationary policy. It is assumed that for every $x \in \Omega$, $a \in A$ there is a known probability measure $Q^a(\cdot, \cdot|x)$ on $F \times S$ such that $\Pr\{X_{n+1} \in \Gamma, S_{n+1} \in M | X_n = x, a_n = a, S_n = s, \mathbf{H}_{n-1} = Q^a(\Gamma, M|x)$ for every $\Gamma \times M \in F \times S$ and all histories H_{n-1}. Also, assume that $Q^a(\Gamma, \cdot|x)$ is absolutely continuous with respect to the measure μ. Then, by the Radon-Nikodym's Theorem, there exists $q^a(\Gamma, \cdot|x)$ such that

$$Q_a(\Gamma, M|x) = \int_M q^a(\Gamma, s|x)\mu(ds), \quad x, \Gamma \in \mathbf{F}, M \in S.$$

Lemma 2.5.1 *For each $\Gamma \in \mathbf{F}$ and $s \in S$*

$$P\{X_n \in \Gamma | \Pi_{n-1} = \pi, S_n = s, a_n = a\} = \frac{\int_\Omega q^a(\Gamma, s|x)\Pi(dx)}{\int_\Omega q^a(\Omega, s|x)\Pi(dx)} \equiv T_\Gamma\{\Pi|s, a\}$$

Proof $\quad P\{X_n \in \Gamma | \Pi_{n-1} = \pi, S_n = s, a_n = a\}$

$$= \frac{P\{X_n \in \Gamma, S_n = s | \Pi_{n-1} = \pi, a_n = a\}}{P\{S_n = s | \Pi_{n-1} = \pi, a_n = a\}}$$

$$= \frac{\int_\Omega P\{X_n \in \Gamma S_n = s | X_{n-1} = x, a_n = a\} P\{dx | \Pi_{n-1} = \pi, a_n = a\}}{\int_\Omega P\{\Omega, S_n = s | X_{n-1} = x', a_n = a\} P\{dx' | \Pi_{n-1} = \pi, a_n = a\}}$$

$$= \frac{\int_\Omega q^a(\Gamma, s(x)\pi(dx)}{\int_\Omega q^a(\Omega, s|x')\pi(dx')} \equiv T_\Gamma\{\pi|s,a\}$$

because X_{n-1} is independent of a_n due to the fact that $\delta_n(\Pi_{n-1}) = a_n$

Remarks 9 *Suppose that* Ω *is countable,* $S_n = (\theta_n, \tau_n)$ *for* τ_n *the time between transitions and* $P\{\theta_n = \theta|x_n = i\tau_n = t, a_n = a, \mathbf{H}_{n-1}\} = \gamma^a_{i\theta}$.

Then it follows that

$$P\{X_n = j|\Pi_{n-1} = \pi, \theta_n = \theta, \tau_n = t, a_n = a\} = \frac{\sum_i \pi_i P^a_{ij} f^a_{ij}(t)\gamma^a_{j\theta}}{\sum_{i,j} \pi_i P^a_{ij} f^a_{ij}(t)\gamma^a_{j\theta}}$$

where P^a_{ij} *are transition probabilities of* $\{X_n\}$,

and $\int_0^t f^a_{ij}(t)dt = P\{\tau_n \le t|X_{n-1} = i, X_n = j, a_n = a\}$

Lemma 2.5.2 *Let* $H_n = [S_n = s, a_n = a, H_{n-1}]$, $n = 1, 2, \cdots$

Then

$$P\{X_n \in \Gamma|H_n\} = P\{X_n \in \Gamma|\Pi_{n-1} = \pi_{n-1}, S_n = s, a_n = a\}$$

Proof Let $\pi_n(\Gamma) = P\{X_n \in \Gamma|H_n\}$.

$$\begin{aligned}
\pi_n(\Gamma) &= P\{X_n \in \Gamma|H_n\} = P\{X_n \in \Gamma|S_n = s, a_n = a, H_{n-1}\} \\
&= \frac{P\{X_n \in \Gamma, S_n = s|a_n = a, H_{n-1}\}}{P\{S_n = s|a_n = a, H_{n-1}\}} \\
&= \frac{\int_\Omega P\{X_n \in \Gamma, S_n = s|X_{n-1} = x, a_n = a, H_{n-1}\}P\{dx|a_n = a, H_{n-1}\}}{\int_\Omega P\{\Omega, S_n = s|X_{n-1} = x', a_n = a, H_{n-1}\}P\{dx'|a_n = a, H_{n-1}\}} \\
&= \frac{\int_\Omega q^a(\Gamma, s|x)\pi_{n-1}(dx)}{\int_\Omega q^a(\Omega, s|x')\pi_{n-1}(dx')} \equiv T_\Gamma\{\pi|s,a\} \\
&= P\{X_n \in \Gamma|\Pi_{n-1} = \pi_{n-1}, s_n = s, a_n = a\}
\end{aligned}$$

due to Lemma 2.5.1. Therefore, the distribution π_n is a sufficient statistic with respect to H_n in the sense that Π_n represents all the informations on the past history of observations of the Markov process. It is important to note that $\{\Pi_n\}$ itself forms a Markov chain because it depends only on one step transition probabilities and π_{n-1}.

2.5.2 Control Model and Optimal Policies

If the process is in $x \in \Omega$, action a is chosen and a signal s is observed, then two things occur:

(i) we incur a cost $c(x, a)$ which is a bounded Borel measurable function on $\Omega \times A$, where if one allows the cost to depend also on the next state visited and the signal observed, then $c(x, a)$ should be interpreted as an expected cost as

$$c(x,a) = \int_s \int_\Omega c(x, a, x', s)Q^a(dx', s|x)\mu(ds).$$

(ii) the next state of the process is chosen according to the transition probability

$$Q^a(\Gamma, S|x) = Pr\{X_{n+1} \in \Gamma, S|X_n = x, a_n = a\} \quad \text{for every } \Gamma \in F.$$

Since the states of the system cannot be observed and distributions π_n which themselves form a Markov chain represent all the available information about the history of the system, π_n are used as states of dynamic programming.

For any policy R, define

$$C_R(\pi) = E_R[\sum_{n=1}^{\infty} \Pi_{i=1}^{n-1}\beta(S_i)c(X_n, a_n)|\Pi_1 = \pi]$$

and

$$C*(\pi) = \inf_R C_R(\pi), \pi \in \Pi$$

where $\Pi_{i=1}^n \beta(S_i) = 1$ for $n > 1$ and $0 \le \beta(S_n) < 1$ for all S_n, $n = 1, 2, \cdots$. Accordingly, $\beta(S_n)$ is called a discount factor which depends on a signal S_n received at the nth epoch. It is of interest to note that we have $\beta(S_n) = \beta$ for Markov decision processes and $\beta(S_n) = e^{-\beta\tau n}$ for semi-Markov decision processes with $S_n = (\theta_n, \tau_n)$ for τ_n the time between transitions.

Condition 1 For a given $\epsilon > 0$ there exists $\delta \in S$ and $\beta_0(\delta)$ such that

$$0 \le \beta(s) \le \beta_0(\delta) < 1 \quad \text{for all } s \in S \backslash \delta$$

and

$$\int_\Omega Q^a(\Omega, \delta|x)\pi(dx) \le 1 - \epsilon \quad \text{for all } a \epsilon A,\ \pi \in \Pi.$$

Lemma 2.5.3 *For any policy* R $C_R(\pi)$ *is bounded for all* $\pi \in \Pi$

Proof There exists a constant K such that $|c(x, a)| \le K$ for all $x \in \Omega$

$$C_R(\pi) \le KE_R[\sum_{n=1}^{\infty} \Pi_{i=1}^{n-1}\beta(S_i)|\Pi_1 = \pi]$$

$$= K\sum_{n=1}^{\infty} E_R[\Pi_{i=1}^{n-1}\beta(S_i)|\Pi_1 = \pi]$$

$$= K\sum_{n=1}^{\infty} \Pi_{i=1}^{n-1}E_R[\beta(S_i)|\Pi_{i-1} = \pi]$$

because S_i is conditionally independent for a given X_i.

$$E_R[\beta(S_i)|\Pi_{i-1} = \pi] = \int_\Omega E_R[\beta(S_i)|X_{i-1} = x]Pr\{dx|\Pi_{i-1} = \pi\}$$
$$= \int_\Omega\int_S \beta(S)Q^a(\Omega, s|x)\mu(ds)\pi(dx)$$

where the policy R is assumed to assign an action a at the $(n-1)$-th epoch,

$$= 1 - \int_\Omega\int_S(1-\beta(s))Q^a(\Omega, s|x)\mu(ds)\pi(dx)$$
$$= 1 - \int_\Omega\int_{S/\delta}(1-\beta(s))Q^a(\Omega, s|x)\mu(ds)\pi(dx)$$
$$- \int_\Omega\int_\delta(1-\beta(s))Q^a(\Omega, s|x)\mu(ds)\pi(dx)$$
$$\leq 1 - (1-\beta_0(\delta))\int_\Omega\int_{S/\delta} Q^a(\Omega, s|x)\mu(ds)\pi(dx)$$
$$= 1 - (1-\beta_0(\delta))\int_\Omega[1 - Q^a(\Omega, \delta|x)\pi(dx)$$
$$< 1 - \epsilon(1-\beta_0(\delta))$$

Therefore

$$C_R(\pi) < K\sum_{n=1}^{\infty}[1-\epsilon(1-\beta_0(\delta))]^{n-1} = K/(1-\beta_0(\delta))\epsilon$$

for $\epsilon > 0$ and $\beta_0(\delta) < 1$.

We have the following generalized form of Bellman's dynamic equation.

Theorem 2.5.1 $C*(\pi) = \min_a\{\int_\Omega[c(x,a)+\int_S \beta(s)C*(T\{\pi|s,a\})Q^a(\Omega, s|x)\mu(ds)]\pi(dx)\}$.

The proof is omitted.

For a stationary policy f mapping from Π to A, define

$$(T_f u)(\pi) = \int_\Omega[c(x, f(\pi)) + \int_S \beta(s)\mu(T\{\pi|s, f(\pi)\})Q^{f(\pi)}(Q, s|x)u(ds)]\pi(dx)$$

and

$$(T_\beta u)(\pi) = \min_a\{\int_\Omega[c(x,a) + \int_S \beta(s)\mu(T\{\pi|s,a\})Q^a(\Omega, s|x)\mu(ds)]\pi(dx)\}$$

Let $B(\Pi)$ denote the set of all bounded functions with respect to the *sup norm*. Then $C_R(\pi) \in B(\Pi)$ for all R.

Note that $B(\Pi)$ is complete.

Lemma 2.5.4 *For every $u, v \in B(\Pi)$ and stationary policy f*

(i) $u \leq v$ implies $T_f u \leq T_f v$

(ii) $T_f C_f = C_f$

(iii) $T_f^n u$ uniformly converges to C_f for all $u \in B(\Pi)$.

The proof is very easy and then is omitted.

Lemma 2.5.5 *Let $B(a|\pi)$ be the probability of choosing action a when the state information is π.*

Then

$$\sum_{a\in A} p(a|\pi)(T_a C_R)(\pi) \geq C_R(\pi) \quad \textit{for any } p(\cdot|\cdot).$$

Theorem 2.5.2 *For every $u, v \in B(\Pi)$ and the sup norm $\|\cdot\|$*

$$\|T_\beta u - T_\beta v\| \leq \beta'\|u - v\|$$

where $\beta' = 1 - \epsilon(1-\beta_0(\delta))$ for $\epsilon > 0, \beta_0(\delta) < 1$ of Condition I.

Theorem 2.5.3 *A stationary policy f, selecting the action minimizing the right hand side of Theorem 2.5.1, is optimal.*

Theorem 2.5.4 *For any stationary policy f, let f' be a policy selecting the action minimizing $T_a C_f$. Then*

$$C_{f'}(\pi) \leq C_f(\pi) \quad \textit{for all } \pi$$

Particularly, if $C_{f'}(\pi) = C_f(\pi)$, then f is optimal.

Lemma 2.5.6 *For $\pi^i \in \Pi$, $i = 1, 2, \ldots, m$ and $\sum_{i=1}^m \lambda_i = 1$, $\lambda_i \geq 0$ for all i.*

Then

$$T\{\sum_{i=1}^m \lambda_i\pi^i|s,a\} = \sum_{i=1}^m \lambda_i'\{\pi^i|s,a\} \textit{for every } s,\ a$$

where

$$\lambda_i' = \frac{\lambda_i\int_\Omega q^a(\Omega, s|x)\pi^i(dx)}{\int_\Omega q^a(\Omega, s|x)\pi(dx)} \quad \textit{and} \quad \pi = \sum_{i=1}^m \lambda_i\pi^i \in \Pi\sum_{i=1}^m \lambda_i' = 1.$$

The proof is easy and is omitted.

Theorem 2.5.5 *Let $(T_\beta^n 0)(\pi) = C^n(\pi)$ for all $\in \Pi$. Then $C^n(\pi)$ is piecewise-linear and concave in π and then $C*(\pi)$ is concave in π as $n \to \infty$.*

Proof The proof is trivial for $n = 1$. Assume for $n-1$.

Let $(T_\beta^n 0) = \min_a (T_a C^{n-1})(\pi)$. Let $\pi^1, \pi^2 \in \Pi$ and λ, $0 \le \lambda \le 1$.

Let $\pi = \lambda\pi^1 + (1-\lambda)\pi^2$

$$\begin{aligned}(T_aC^{n-1})(\pi) &= \lambda\int_\Omega c(x,a)\pi^1(dx) + (1-\lambda)\int_\Omega c(x,a)\pi^2(dx)\\ &+\int_\Omega\int_S \beta(s)C^{n-1}T\{\lambda\pi^1+(1-\lambda)\pi^2|s,a\})Q^a(\Omega,s|x)\mu(ds)(dx)\\ &= \lambda\int_\Omega c(s,a)\pi^1(dx)+(1-\lambda)\int_\Omega c(x,a)\pi^2(dx)+\int_\Omega\int_S\beta(s)C^{n-1}\lambda'\{T\pi^1|s,a\}\\ &+(1-\lambda')T\{\pi^1(s,a)\}Q^a(\Omega,s|x)\mu(ds)\pi(dx)\\ &= \lambda[\int_\Omega c(x,a)\pi^1(dx)+\int_\Omega\int_S\beta(s)C^{n-1}(T\{\pi^1|s,a\})Q^a(\Omega,s|x)\mu(ds)\pi^1(dx)]\\ &+(1-\lambda)[\int_\Omega c(x,a)\pi^2(dx)+\int_\Omega\int_S\beta(s)C^{n-1}(T\{\pi^2|s,a\})Q^a(\Omega,s|x)\mu(ds)\pi^2(dx)]\\ &= \lambda T_aC^{n-1}(\pi^1)+(1-\lambda)T_aC^{n-1}(\pi^2).\end{aligned}$$

Since $C^n = \min_{a\in A} T_aC^{n-1}(\pi)$ and A is finite, $C^n(\pi)$ is piecewise-linear and concave in π and so $C^*(\pi)$ is concave as $n \to \infty$.

Bounds for optimal expected costs

Let us consider two extreme cases, (i) the states of the system are observable (complete state information), (ii) the state of the system are not observable at all (no observation).

(i) Complete state information

Define

$$C'(x) = \inf_R E_R[\sum_{n=1}^\infty \beta(S_i)c(x_n,a_n)|X_1 = x]$$

and let

$$C'(\pi) = \int_\Omega C'(x)\pi(dx)$$

where π is the distribution of X_1 for the partially observable model.

(ii) No observation

Define

$$C''(\pi) = \inf_R E_R[\sum_{n=1}^\infty \epsilon\beta(\tau_1+\cdots+\tau_{n-1})c(X_n,a_n)|\Pi_1 = \pi]$$

where control policy is based only on a prior state information and so no a posterior state information is obtainable.

Theorem 2.5.6 *$C'(\pi) \le C*(\pi) \le C''(\pi)$ for all $\pi \in \Pi$.*

Proof Let $H'_n = \{\pi_0, X_0, a_0, \cdots, X_n, S_n, a_n\}$ for the complete state information case and $H''_n = \{\pi_0, a_0, \cdots, a_n\}$ for the no observation case.

Then, for the action space A define as follows;

$$\begin{aligned}\delta_n &: H'_n \to A \quad \text{for the complete information}\\ \delta_n &: H_n \to A \quad \text{for the partial observation, and}\\ \delta_n &: H''_n \to A \quad \text{for no observation.}\end{aligned}$$

Obviously

$$H''_n \subset H_n \subset H'_n.$$

Then

$$C'(\pi) \le C*(\pi) \le C''(\pi) \quad \text{for all} \quad \pi \in \Pi$$

Remarks 10 *Under Condition I $C''(\pi)$ is again bounded by $K/\epsilon(1-\beta_0(\delta))$. Hence, $C(\pi) \le C*(\pi) \le C''(\pi) \le K/\epsilon(1-\beta_0(\delta))$. It is of interest to note that $C'(\pi)$ is linear in π. $C''(\pi) - C'(\pi)$ represents a value of complete information and $C''(\pi) - C*(\pi)$ is a value of partial information.*

2.6 Transformation of Partially Observable Markov Decision Processes into Piecewise Linear Ones

It is well known (see Sawaragi and Yoshikawa [110], Dynkin [39]) that partially observable Markov decision processes (abbreviated by POMP) with finite (or at most countable) states can be transformed into completely observable Markov decision processes (abbreviated by MP) with continuous states. But the state space of transformed MP becomes the set of probability vectors which is no longer finite nor countable but continuous (continuum). Then it is almost impossible to compute an optimal cost and its corresponding policy of continuous state MP in the form of dynamic programming. Sawaki [96] recently discusses piecewise linear MP. In this section it is shown that such POMP are actually piecewise linear MP with complete state observations. Since piecewise linear MP have many advantages for applications and implementation in a computer, it is important and of interest to provide a justification of the transformation of POMP to completely observable MP, which enables us to handle uncountable (continuous) state space MP and lightens a computational burdens. Also it will be shown how to find the coefficients of piecewise linear functions and to handle the product of simple partitions for the purpose of computer implementation.

2.6.1 Piecewise Linear Markov Decision Processes

Piecewise linear MP are special cases of the general MP with finite actions which satisfy the monotone contradiction mapping assumption of Denardo [31]. Under the setting of Blackwell [17] the general MP with finite actions are defined by the four subjects (Ω, A, q, c), where Ω is a linear vector *state space*, A is the finite *set of actions* $a \in A$, $q(\cdot|x,a)$ is the one step *transition probability* on Ω for each pair $(x,a) \in \Omega \times A$, and $c(\cdot,\cdot)$ is the bounded *immediate cost* on (Ω, A). Define a policy $\delta : \Omega \to A$. Our expected discounted total cost $V^\delta(x)$ at an initial state x under a statinary policy δ is written as

$$V^\delta(x) = E\left\{\sum_{n=1}^{\infty} \beta^{n-1} c(X_n, \delta(X_n))|X_1 = x\right\},$$

where $\{X_n : n = 1, 2, \ldots\}$ is a Markov chain with transition probability $q(\cdot|x, \delta(x))$ and $\beta, 0 \leq \beta < 1$, is the discount factor. Define the optimal cost V^* by

$$V^*(x) = \inf_{\delta \in \Delta} V^\delta(x) \quad \text{for all} \quad x \in \Omega,$$

where Δ is a family of stationary policies. It is well known that there always exists an optimal policy δ^* which is stationary, and $V^{\delta^*} = V^*$ satisfies $V^* = U_* V^*$, where

$$U_* v(x) = \min_a \left\{c(x,a) + \beta \int_\Omega v(x') q(dx'|x,a)\right\}$$

for $v \in B(\Omega)$ the set of bounded functions on Ω. Also, define $U_\delta : B(\Omega) \to B(\Omega)$ by

$$(U_\delta v)(x) = c(x, \delta(x)) + \beta \int_\Omega v(x') q(dx'|x, \delta(x)).$$

We write $U_\delta = U_a$ if $\delta(x) = a$ for $x \in \Omega$.

A collection $P = \{E_1, E_2, \ldots, E_m\}$ of subsets of Ω is a partition of Ω if $E_i \cap E_j = \emptyset$ for $i \neq j$ and if $\cup_{i=1}^m E_i = \Omega$. Each member of partition P is a *cell*. If each cell of a partition

is a convex polyhedron then the partition is called *simple*. A function v is *piecewise linear* if there exists a simple partition $P = \{E_1, E_2, \ldots, E_n\}$ such that $v(x) = v_i(x)$ for all $x \in E_i, i = 1, 2, \ldots, m$ and each $v - i$ is the restriction to E_i of a linear functions on Ω. A policy $\delta \in \Delta$ is *piecewise constant* if there is a simple partition $\{E_1, E_2, \ldots, E_m\}$ of Ω such that $\delta(x) = a_i$ for all $x \in E_i$.

Definition MP are called *piecewise linear* if there is exists a simple partition $P = \{E_1, E_2, \ldots, E_m\}$ of Ω such that $(U_\delta v)(x)$ is piecewise linear for v piecewise linear and δ piecewise constant.

Lemma 2.6.1 *If MP is piecewise linear, then $U_* v$ is piecewise linear and there exists a piecewise constant policy δ such that $U_\delta v = U_* v$.*

Proof Suppose that $U_\delta v$ is piecewise linear with a simple partition $\{E_1, E_2, \ldots, E_m\}$ and that $U_\delta v = U_a v$ for $x \in E_i$ an arbitrary but fixed cell. Therefore, there exists a simple partition P^a for each $a \in A$. Then we may suppose that $U_a v$ is piecewise linear with respect to the simple partition P^a. Let $P = \prod_{a \in A} P^a$ which is the product of the simple partitions. Since the product of simple partitions is again simple, P is simple and finer than each P^a, and so each $U_a v$ is piecewise linear with respect to P. For this refined partition P, there is some linear functional α_F^a such that for each $F \in P$ and $a \in A$

$$(U_a v)(x) = \alpha_F^a(x) \quad \text{for} \quad x \in F.$$

For each $F \in P$, define the sets $G_F^b, b \in A$, by $G_F^b = \{x \in F : \alpha_F^b x = \min_a \alpha_F^a x\}$. Then $\{G_F^a : a \in A\} = P^F$ is a partition of F. Put $\hat{P} \equiv \cup_{F \in P} P^F$ and then $\hat{P}$ is a partition of Ω with the property that

$$(U_* V)(x) = \alpha_F^a(x) \quad \text{if} \quad x \in G_F^a \in \hat{P}.$$

The policy δ defined by $\delta(x) = a$ for $x \in G_F^a \in \hat{P}$ satisfies $U_\delta V = U_* V$.

Corollary Suppose that MP is piecewise linear with contraction mapping U which is either $U - \delta$ or U_*. Let $v^n = U v^{n-1}$ for $n = 1, 2, \ldots$, and v^n be piecewise linear. Then v^n is piecewise linear and the stationary policy δ_n, defined by $U_{\delta_n} = U_* v^{n-1}$, is piecewise constant. Furthermore v^n converges in norm to the fixed point V^* or V^δ corresponding to U_* or U_δ, respectively.

Remarks 11 *The fixed points V^* or V^δ need not to be piecewise linear and δ^* need not to be piecewise constant since the number of cells in the limiting partition is not necessarily finite.*

2.6.2 Partially Observable Markov Decision Processes

First, we shall introduce POMP and provide a lemma to be used for transformation of POMP into piecewise linear MP.

Consider Markov decision process (called the core process) with state set $\{1, 2, \ldots, N\}$, with action set A, with probability transition matrices $\{P_{ij}^a\}$ and with immediate cost vectors h^a. Let Z_n be the state at the n-th transition. Assume that the process $\{Z_n, n = 0, 1, 2, \ldots\}$ cannot be observed, but at each transition a signal is transmitted to the decision maker. The set of possible signal Θ is assumed to be finite. For each n, given that $Z_n = j$ and that action a is to be implemented, the signal θ_n is independent of the history of the signals and actions $\{\theta_0, a_0, \theta_1, a_1, \ldots, \theta_{n-1}, a_{n-1}\}$ prior to the n-th transition and has conditional probability denoted by $\gamma_{j\theta}^a = P[\theta_n = \theta | Z_n = j, a]$.

Let $\Omega = \{x = (x_1, x_2, \ldots, x_N) : \sum_{i=1}^N x_i = 1, x_i \geq 0, \forall i\}$. Define the i-th component of X_n, the random variable of x, to be

$$P[Z_n = i | \theta_0, a_0, \theta_1, \ldots, \theta_{n-1}, a_{n-1}, \theta_n], \quad i = 1, 2, \ldots, N.$$

It can be shown (see Dynkin [39]) that

$$P[Z_{n+1} = j | \theta_0, a_0, \theta_1, \ldots, \theta_n, a_n, \theta_{n+1}] = P[Z_{n+1} = j | \theta_{n+1}, a_n, X_n]$$

Thus X_n represent a sufficient statistic for the complete past history $\{\theta_0, a_0, \ldots, a_{n-1}, \theta_n\}$. It follows that $\{X_n : n = 0, 1, 2, \ldots\}$ is Markov decision process (see Dynkin [39]), called the observed process. Its immediate cost is $c(x, a) = h^a x$. Its action set is A. Its probability transition function is determined by the following calculation: For each measure subset $B \subseteq \Omega, x \in \Omega$, and $a \in A$,

$$\begin{aligned} q(B|x,a) &= P[X_{n+1} \in B | X_n = x, a_n = a] \\ &= \sum_\theta P[X_{n+1} \in B | \theta_{n+1} = \theta, X_n = x, a_n = a] \sum_j \gamma_{j\theta}^a \sum_i P_{ij}^a x_i \\ &= \sum_\theta P[X_{n+1} \in B | \theta_{n+1} = \theta, X_n = x, a_n = a] \mathbf{1} P^a(\theta) x \end{aligned}$$

where $\mathbf{1} = (1, 1, \ldots, 1)$ and $P^a(\theta) = [P_{ij}^a \gamma_{j\theta}^a]^T$. Define the vector $T(x|\theta, a)$ by

$$T(x|\theta, a) = \frac{P^a(\theta)x}{\mathbf{1}P^a(\theta)x}.$$

Note that $T(X_n|\theta_{n+1}, a_n) = X_{n+1}$, and that

$$P[X_{n+1} \in B | \theta_{n+1} = \theta, X_n = x, a_n = a] = \begin{cases} 1, & \text{if } T(x|\theta, a) \in B, \\ 0, & \text{otherwise.} \end{cases}$$

So,

$$q(B|x,a) = \sum_{\theta \in \Phi^a(B,x)} \mathbf{1} P^a(\theta) x,$$

where $\Phi^a(B, x) = \{\theta : T(x|\theta, a) \in B\}$.

Next, we show that $q^a(B, x) \equiv \int_B x' q(dx'|x, a)$ is piecewise linear in x for each convex polyhedral set $B \subseteq \Omega$ and action $a \in A$. Using the previously computed $q(B|x, a)$ we have

$$\begin{aligned} q^a(B,x) &\equiv \int_B x' q(dx'|x,a) = \sum_{\theta \in \Phi^a(B,x)} T(x|\theta,a) \mathbf{1} P^a(\theta) x && (2.10) \\ &= \sum_{\theta \in \Phi^a(B,x)} \frac{P^a(\theta)x}{\mathbf{1}P^a(\theta)x} \mathbf{1} p^a(\theta) x = \sum_{\theta \in \Phi^a(B,x)} p^a(\theta) x. && (2.11) \end{aligned}$$

Thus it is sufficient to verify that the set valued function $\Phi^a(B, \cdot) : \Omega \to 2^\Theta$ is piecewise constant on Ω, where 2^Θ is the power set of Θ. To this we need

Lemma 2.6.2 *For each signal θ, action a, and set $B \in \Omega$, define*

$$E_\theta^{B,a} = \{x \in \Omega : T(x|\theta, a) \in B\}.$$

Then for any subset of signals $\psi \subseteq \Theta$, we have

$$\Phi^a(B, x) = \psi \quad \text{if and only if} \quad x \in \bigcap_{\theta \in \psi} E_\theta^{B,a} \cap \bigcap_{\theta \in \psi^c} (E^{B,a})^c.$$

Proof Note that $E_\theta^{B,a} = \{x : \theta \in \Phi^a(B, x)\}$. Thus if $x \in E_\theta^{B,a}$ for $\theta \in \psi$, then $\theta \in \Phi^a(B, x)$. On the other hand, is $x \in (E_\theta^{B,a})^c$ for $\theta \in \psi^c$, then $\theta \in \Phi^a(B, x)$. Consequently, $\psi^c \subseteq (\Phi^a(B, x))^c$. It follows that $\psi = \Phi^a(B, x)$.

Conversely, suppose that $\Phi^a(B, \hat{x}) = \psi$. Then $\hat{x} \in B_\theta^{B,a}$ for each $\theta \in \psi$ and $\hat{x} \in (E_\theta^{B,a})^c$ for each $\theta \in \psi^c$, which completes the proof. □

The next theorem shows that POMP is actually piecewise linear MP and provides a formula for computing the cost function which is convinient for computer implementation.

Theorem 2.6.1 *Suppose $v(x) = v_i(x)$ for $x \in E_i$ with a simple partition $P_v = \{E_1, E_2, \ldots, E_m\}$ of Ω. Then*

$$(U_a v)(x) = \left\{h^a + \beta \sum_i v_i \sum_{\theta \in \psi_i} p^a(\theta)\right\} x \quad for \quad x \in \cap_{i=1}^m E^a_{B_i}(\psi_i),$$

where $E^a_B(\psi) = \{x \in \Omega : \Phi^a(B, x) = \psi\}$ and $\cap_{i=1}^m E^a_{B_i}(\psi_i)$ is a cell in the partition P of Ω defined by

$$P = \prod_{i=1}^m \{E^a_{B_i}(\psi) : \psi \in 2^\theta\}$$

That is, $U_a v$ is piecewise linear with the partition P.

Proof First observed from equation (2.11) and Lemma 2.6.2,

$$\begin{aligned}\int_B x' q(dx'|x, a) &= \sum_{\theta \in \psi} T(x|\theta, a) P^a(\theta) a \\ &= \sum_{\theta \in \psi} P^a(\theta) x \quad \text{for} \quad x \in E^a_B(\psi).\end{aligned}$$

Therefore we have

$$\begin{aligned}(U_a v)(x) &= h^a x + \beta \int_\Omega v(x') q(dx'|x, a) \\ &= h^a x + \beta \sum_{i=1}^m v_i \int_{B_i} x' q(dx'|x, a) \\ &= \left\{h^a + \beta \sum_{i=1}^m v_i \sum_{\theta \in \psi} P^a(\theta)\right\} x \quad \text{for} \quad x \in E^a_{B_i}(\psi).\end{aligned}$$

Lemma 2.6.2 gives an explicit representation of $E^a_B(\psi)$ and $q^a(B, x)$ is piecewise linear with respect to the partition $\{E^a_B(\psi) : \psi \to 2^\Theta\}$, where it is assumed that $q^a(B, x) = 0$ if $E^a_B(\psi) = \emptyset$ for all ψ. Although this partition is not simple, it can easily be refined to a simple partition as in the next paragraph.

Suppose that $B \subseteq \Omega$ is a convex polyhedral set. Since for $X \in \Omega = \{x : \sum x_i = 1, x_i \geq 0 \;\; \forall i\}$ an inequality $lx \leq b$ can be rewritten as $lx - b = (l - b\mathbf{1})x \leq 0$, we can without loss of generality assume that B has the representation

$$B = \{x \in \Omega : Kx < \mathbf{0}, Lx \leq \mathbf{0}\}$$

for some matrices K and L, where $\mathbf{0} = (0, 0, \ldots, 0)^T$. With this representation of B,

$$\begin{aligned}E^{B,a}_\theta &= \{x \in \Omega : T(x|\theta, a) \in B\} \\ &= \left\{x \in \Omega : K \frac{P^a(\theta)x}{\mathbf{1}P^a(\theta)x} < \mathbf{0}, L \frac{P^a(\theta)x}{\mathbf{1}P^a(\theta)x} \leq \mathbf{0}\right\} \\ &= \{x \in \Omega : KP^a(\theta)x < \mathbf{0}, LP^a(\theta)x \leq \mathbf{0}\} \\ &= \{x \in \Omega : K^a(\theta)x < \mathbf{0}, L^a(\theta) \leq \mathbf{0}\},\end{aligned}$$

where $K^a(\theta) = KP^a(\theta)$ and $L6a(\theta) = LP^a(\theta)$. SO each $E^{B,a}_\theta$ is a convex polyhedral set. Each $(E^{B,a}_\theta)^c$ can be represented as a union of disjoint convex polyhedral sets. It follows that $E^a_B(\psi)$ is a union disjoint polyhedral sets, say $E^a_B(\psi) = \cup_{j=1}^{n_\psi} \{E^a_B(\psi)\}$. Thus $q^a(B, x)$ is piecewise linear with respect to the simple partition $\{E_j(\psi) : j = 1, 2, \ldots, n_\psi, \psi \in 2^\Theta\}$.

Our motivation for studying piecewise linear MP which include POMP as special cases is that they are easily represented in a computer in terms of piecewise linear costs and piecewise constant policies as well as simple partitions. For example, a simple partition $P = \{E_1, E_2, \ldots, E_n\}$ can easily be stored in a computer as:

$$E_i = \{x : K^i x < b^i, L^i x \leq d^i\}, \quad i = 1, 2, \ldots, n,$$

where each b^i and d^i is an N-dimensional vector and each K^i and L^i is a matrix with N-dimensional rows. A piecewise cost function $f(x) = f_i x$ and a piecewise constant policy $\delta(x) = a_i$ on E_i. This situation will be denoted by

$$\begin{aligned}&f \sim \{(f_i; K^i, b^i; L^i, d^i), \quad i = 1, 2, \ldots, n\}, \\ &\delta \sim \{(a_i; K^i, b^i; L^i, d^i), \quad i = 1, 2, \ldots, n\}, \\ &E_i \sim (K^i, b^i; L^i, d^i).\end{aligned}$$

Our MP requires a performance of product of simple partitions. This can be performed by combining the corresponding lists of inequalities as the intersection of two cells. Thus it is easy to form product partitions.

Chapter 3

Optimal Policies in Inventory Control Problems

3.1 Introduction

In this chapter we consider two types of dynamic inventory control problems. The first type is of a classical inventory control problem with fixed ordering costs. The second one is of products which can not carried over to the future demand.

In Section 3.2 we consider a dynamic stochastic inventory model with fixed inventory holding and shortage costs in addition to a fixed ordering cost. We discuss a sufficient condition for the (s, S) policy to be optimal in the class of such stochastic inventory models. Furthermore, we explore how such a sufficient condition can be rewritten when the demand distribution is specified. Several examples like uniform, exponential, normal and gamma distribution functions are treated.

The main result of this paper is to show on the basis of Ishigaki and Sawaki [49] that the (s, S) policy is still optimal under a simple condition even if the fixed inventory costs are involved. Even though Aneja nd Noori [3] consider a similar model only with fixed inventory shortage cost, our proof for the optimality of an (s, S) policy in the multi-period model is different from and much simpler than theirs. It is well known (see Scarf [111], and Veinott [124], [125], [126]) that the (s, S) policy is optimal for the stochastic inventory control problem with fixed and proportional production costs. As to dynamic stochastic inventory control the concept of K-convexity is crucial to the existence of an optimal policy which is (s, S) type. However, if the inventory cost includes a fixed cost,

the (s, S) policy is no longer optimal. For example, Aneja and Noori [3] discuss a sufficient condition for the (s, S) policy to be optimal if the inventory shortage cost has a fixed part but inventory holding cost does not have such a fixed one.

In Section 3.3 we deal with the problem of selling a fixed number of units of certain products that cannot be carried over and are not storable for consumers. Hotel rooms, airplane seats, and concert seats are examples of such products that are sold at multiple prices under certain restrictions (see Sawaki [106]).

This chapter analyses the problem of allocation products between two types of prices when the demands for the types of product are stochastically dependent. We derive a simple formula for determining how many products to sell at each price. In addition, we provide three interesting examples of cases in which demand distribution is specified.

Vacant hotel rooms, seats in passenger planes, and seats in concert halls are examples of what are referred to as "inventory" in their respective business circles. The special feature of such inventory items is that the saleable total capacity is fixed in advance and it is impossible to carry over any remaining inventory to the next day. In order to counteract this physical property of inventory that is incapable of being carried over to the next day, businesses with such inventory aim at guaranteeing demand by setting multiple prices on inventory items of identical quality and issuing a variety of discount tickets at a relatively early stage. Users, on the other hand, also have different preferences in regard to inventory items of identical quality. Thus, for example, people who use passenger planes for tourism purposes generally make reservations at a comparatively early time and at the cheapest discount fare possible, as opposed to those who use passenger planes for business trips.

In Section 3.3 we hypothesize a case in which, when the amount of saleable inventory is fixed in advance, two types of demand for such inventory items occur as a result of differences in the time that demand arises and in profitability. We consider the problem of deciding how to distribute the inventory items between these two types of demands.

3.2 On the (s, S) Policy with Fixed Inventory Costs

In this section we consider a finite dynamic stochastic inventory problem with a single item. We need the following assumptions and notations:

- The unsatisfied demand is lost.
- If the demand is less than stock level, then holding cost incurrs at the end of each period. This holding cost consists of two parts, the fixed holding cost $[B_1]$ and the proportional holding cost $[h]$.
- If the demand is bigger than the stock level, then shortage cost incurrs. This shortage cost again consists of two parts, the fixed shortage cost $[B_2]$ and the proportional shortage cost $[p]$.
- If an order is taken, then the ordering cost incurrs. This ordering cost consists of the fixed ordering cost $[K]$ and the proportional cost $[c]$.
- Demand of each period is given by the random variable which has the probability density function (p.d.f.) $\phi(\xi)$. We assume that p.d.f. $\phi(\xi)$ is differentiable.
- Both cost functions and p.d.f. of demand are identical over the periods.

Let us assume that the planning horizon is discrete, finite and consists of N periods. At first we consider the expected cost over n periods $(n \leq N)$. If the stock level immediately after an ordering is y, then the sum of the expected holding and shortage costs to be charged during a period is given by

$$\begin{aligned} L(y) &= h\int_0^y (y-\xi)\phi(\xi)d\xi + B_1\int_0^y \phi(\xi)d\xi \\ &\quad + p\int_y^\infty (\xi-y)\phi(\xi)d\xi + B_2\int_y^\infty \phi(\xi)d\xi \end{aligned} \tag{3.1}$$

where we assume that B_2 is not equal to B_1. If $B_1 = B_2$, then it is easy to see from equation (3.1) that the sum of the fixed holding and shortage cost is independent of y. Therefore, this model reduces to the classical stochastic inventory model only with a fixed cost. Let $C_n(x)$ be the minimum of the expected total discount cost over n-periods when x is the starting inventory level before an ordering at the beginning of period n. Then we

have from the principle of the optimality,

$$C_n(x) = \min_{y\geq x}\{H(y-x) + L(y) + \rho\int_0^\infty C_{n-1}([y-\xi]^+)\phi(\xi)d\xi\} \tag{3.2}$$

$$[y-\xi]^+ = \begin{cases} 0, & \text{if } y \leq \xi \\ y-\xi, & \text{otherwise.} \end{cases}$$

where $n = 1, 2, \ldots, N$, ρ is the discount factor, $0 < \rho \leq 1$, $C_0(x) = 0$ for all x and $H(\cdot)$ is defined as follows:

$$H(y-x) = \begin{cases} 0, & \text{if } y - x \leq 0; \\ K + c\cdot(y-x), & \text{otherwise.} \end{cases}$$

The objective of this section is to find an optimal inventory policy which minimizes the expected total discounted cost. To prove the optimality of an (s, S) policy for the multi-period model, we first consider the single-period model of this problem.

3.2.1 Single-Period Model

We discuss the optimality of the (s, S) policy in a single period model. $N = 1$, equation (3.2) reduces to

$$C_1(x) = \min_{y\geq x}\{H(y-x) + L(y)\}. \tag{3.3}$$

Theorem 3.2.1 *For all noninceasing demand density functions, a necessary and sufficient condition for the optimality of an (s, S) policy for the single-period problem is that*

Condition (A)

$$\frac{\phi'(y)}{\phi(y)}\left\{\begin{array}{c} \leq \\ \geq \end{array}\right\}\frac{h+p}{B_2-B_1} \quad \text{for all} \quad y \in \Re^+ \quad \text{if} \quad B_2 - B_1\left\{\begin{array}{c} > \\ < \end{array}\right\}0,$$

where $\Re^+ = \{y | y \geq 0\}$.

Proof (Necessity) It suffices to prove for the case $B_2 - B_1 < 0$ because the proof of the case $B_2 - B_1 > 0$ can be applied to Aneja and Noori's result as $B = B_2 - B_1$. Let $F_1(y|x)$ be the quantity inside the braces of the righthand side of equation (3.3) and put $G_1(y)$ as follows. If $y > x$, then we have $G_1(y) = F_1(y|x) - K + cx$

$$G_1(y) = cy + h\int_0^y (y-\xi)\phi(\xi)d\xi + B_1\int_0^y \phi(\xi)d\xi + p\int_y^\infty (\xi-y)\phi(\xi)d\xi + B_2\int_y^\infty \phi(\xi)d\xi$$

In this case the first and second derivatives of function $G_1(y)$ are as follows:

$$G_1'(y) = c + h\int_0^y \phi(\xi)d\xi + B_1\phi(y) - p\int_y^\infty \phi(\xi)d\xi - B_2\phi(y)$$
$$G_1''(y) = h\phi(y) + B_1\phi'(y) + p\phi(y) - B_2\phi'(y) = (h+p)\phi(y) - (B_2 - B_1)\phi'(y). \quad (3.4)$$

From the Condition (A),

$$(h+p)\phi(y) - (B_2 - B_1)\phi'(y) \geq 0.$$

That is

$$G_1''(y) \geq 0$$

so $G_1(y)$ is convex in y.

If $y = x$, then we have $G_1(y) - cy = F_1(y)$

$$\begin{aligned} G_1(x) &= cx + h\int_0^x (x-\xi)\phi(\xi)d\xi + B_1\int_0^x \phi(\xi)d\xi \\ &\quad + p\int_x^\infty (\xi - x)\phi(\xi)d\xi + B_2\int_x^\infty \phi(\xi)d\xi \end{aligned} \quad (3.5)$$

The first and second derivatives of equation (3.5) are as follows:

$$\begin{aligned} G_1'(x) &= c + h\int_0^x \phi(\xi)d\xi + B_1\phi(x) - p\int_x^\infty \phi(\xi)d\xi - B_2\phi(x) \\ G_1''(x) &= h\phi(x) + B_1\phi'(x) + p\phi(x) - B_2\phi'(x) \\ &= (h+p)\phi(x) - (B_2 - B_1)\phi'(x) \end{aligned} \quad (3.6)$$

Since this equation is identical with the equation (3.4), equation (3.4) and equation (3.6) implies that $G_1(y)$ is a convex function of $y \geq x$.

Therefore,

$$F_1(y) = \begin{cases} G_1(y) + K - cx, & \text{if} \quad y > x; \\ G_1(x) - cx, & \text{if} \quad y = x, \end{cases}$$

and

$$C_1(x) = \begin{cases} G_1(S) + K = G_1(s), & \text{if} \quad x < s; \\ G_1(x) - cx, & \text{if} \quad x \geq s, \end{cases} \quad (3.7)$$

where

$$S = \arg\{\inf\{G_1(y)\}\},$$
$$s = \min\{z | G_1(S) + K = G_1(z)\}.$$

Consequently, under the condition (A) an optimal inventory policy is as follows:

(I) if $x < s$, order $S - x$

(II) if $x \geq s$, do not order.

Such a policy is called the (s, S) policy.

(Sufficiency) We shall show that the following statement. Suppose that the condition (A) does not hold, that is, there exists $y \in \Re^+$ such that

$$\frac{\phi'(y)}{\phi(y)} \begin{Bmatrix} > \\ < \end{Bmatrix} \frac{h+p}{B_2 - B_1} \quad \text{if} \quad B_2 - B_1 \begin{Bmatrix} > \\ < \end{Bmatrix} 0,$$

then there are values of parameters K and c for which any (s, S) policy cannot be optimal. We prove it for the case of $B_2 - B_1 > 0$. The proof for the case $B_2 - B_1 < 0$ is similar and is hence omitted. From assumption, we can find y^* such that

$$\phi'(y) < 0 \quad \text{for all} \quad y > y^*,$$

where $\phi(y)$ is assumed to be continuously differentiable. Since $\phi'(y)$ is continuous and negative for sufficiently large y, there exists y_0 such that

$$\frac{\phi'(y)}{\phi(y)} = \frac{h+p}{B_2 - B_1} \quad (3.8)$$

Let y_0 satisfying (3.8). For all y with $y \geq y_0$ an inequality

$$\frac{\phi'(y)}{\phi(y)} \leq \frac{h+p}{B_2 - B_1}$$

holds, which implies that

$$G_1''(y) \geq 0. \quad (3.9)$$

On the other hand, in the left neighbourhood of y_0, we have

$$G_1''(y) \leq 0 \quad \text{for} \quad y \in (y_0 - \delta, y_0), \quad \delta > 0. \quad (3.10)$$

Now consider the function $f(y) = G_1'(y) - (c - p)$. Thus,

$$f(y) = (h+p)\Phi(y) - (B_2 - B_1)\phi(y),$$

where Φ is the cumulative distribution function (c.d.f.) of ϕ. Since $f(y) = G_1''(y)$, it follows from (3.9) and (3.10) that $f(y)$ attains a local minimum at y_0, where we assume that this minimum is also global.

By an appropriate choice of c, we can ensure that $G'(y) = f(y) + (c-p) = 0$ at y_1 and y_2 such that $G_1''(y_1) < 0$ and $G_1''(y_2) > 0$. Thus, the function $G_1'(y)$ has at least two consecutive zeros, one at y_1 where it is concave and the other at y_2 where $G_1(y)$ is convex and there are no zeros at y_2. So we can choose an appropriate $K < G_1(y_1) - G_1(y_2)$. Summing up the above argument we obtain the optimal inventory policy as follows.

(i) order $(y_2 - x)$, if $a \le x \le b$

(ii) not oder, otherwise

which is no longer an (s, S) policy. Thus condition (A) is necessary for the (s, S) policy to be optimal.

Remarks 1 *Theorem 3.2.1 concludes that condition (A) is necessary an sufficient condition for the (s, S) policy to be optimal. If the rigthand side of condition (A) is positive, the condition (A) holds for all nondecreasing p.d.f.. Furthermore, note that pur model includes Aneja and Noori [3] type $(B_2 = B, B_1 = 0)$ and Scarf [111] type $(B_2 = B_2 = 0)$.*

3.2.2 Multi-Period Model

In this subsection we shall showw that condition (A) is also sufficient for the (s, S) policy to be optimal in the multi-period model. This is not true in Aneja and Noori [3] because our proof is difficult from theirs.

Define $G_n(y)$ by

$$G_n(y) = cy + L(y) + \rho \int_0^y C_{n-1}([y-\xi]^+)\phi(\xi)d\xi$$

where $n = 1, 2, \ldots, N$. This definition is corresponding to G_1 in We prove Theorem 3.2.2 by using properties of a K-convex function which is defined as follows.

Definition [K-convexity [111]] Let $K \ge 0$, and let $G_n(x)$ be a differentiable function. We say that $G_n(x)$ is K-convex if

$$K + G_n(a+x) - G_n(x) - aG_n'(x) \ge 0, \quad x > 0, \forall x \forall n$$

Before presenting Theorem 3.2.2 we prepare Proposition 1 and 2 which proofs can be found in their references.

Proposition 1 (Scarf [111])

1. 0-convex function is ordinary convex.
2. If $f(x)$ is K-convex, then $f(x+h)$ is K-convex for all h.
3. If f and g are K_1-convex and K_2-convex, respectively, then $(\alpha f + \beta g)$ is $(\alpha K_1 + \beta K_2)$-convex for all α and β positive.
4. If $g_n(x)$ is K-convex, so is $\int_0^\infty g_n(x\xi)\phi(\xi)d\xi$.

Proposition 2 (Denardo [31]) Let $h(y)$ be convex and nondecreasing on Y. Let $C(x)$ be K-convex on a set $X \supseteq \{h(y) | y \in Y\}$. If all elements $a < c$ of X have $C(a) \le C(c) + K$, then $C[h(y)]$ is K-convex on Y.

Theorem 3.2.2 *If condition (A) holds, then C_n is K-convex.*

Proof The proof is by induction on n. For $n = 1$, C_1 is K-convex because equation (3.7) satisfies the definition of K-convexity (from convexity of G_1). For $n = k$, we assume that C_k is K-convex. From equation (3.2), $C_{k+1}(x)$ is

$$C_{k+1}(x) = \min_{y \ge y}\{H(y-x) + L(y) + \rho \int_0^\infty C_k([y-\xi]^+)\phi(\xi)\}. \tag{3.11}$$

Let $F_{k+1}(y|x)$ be the quantity inside the braces of the righthand side of equation (3.11) and put $G_{k+1}(y)$ as follows.

$$G_{k+1}(y) = \begin{cases} F_{k+1}(y|x) + K - cx, & \text{if} \quad y > x; \\ F_{k+1}(x|x) - cx, & \text{if} \quad y = x. \end{cases}$$

For $y \ge x$,

$$G_{k+1}(y) = cy + L(y) + \rho \int_0^\infty C_k([y-\xi]^+)\phi(\xi)d\xi.$$

Since the first term plus the second term is $G_1(y)$, so it is 0-convex. The K-convexity of the third term derived from Proposition 1 and 2. Because we can take $h(y) = [y-\xi]^+$ and $C(x) = C_k(x)$, respectively, in Proposition 2, and take $g_n(x) = C_k(x)$ in Proposition 1. Thus $G_{k+1}(y)$ is a combination of a convex and a K-convex function and is, therefore, K-convex. So is $C_{k+1}(x)$.

Theorem 3.2.3 *If the p.d.f. of demand, $\phi(\xi)$, satisfies condition (A), then a (s_n, S_n) policy is optimal for our multi-period inventory problem.*

Proof From Theorem 3.2.2, we established $C_n(x)$ is K-convex for all n and hence, the optimal policy for the n-period problem is (s_n, S_n) where:

$$G_n(S_n) = \min_y G_n(y), \quad G_n(s_n) = K + G_n(S_n).$$

This policy states that when inventory on hand is equal or below the reorder point s_n, sufficient stock is ordered to raise the inventory level to the order-up-to-level S_n. The minimum expected total cost of following such a policy would be:

$$C_n(x) = \begin{cases} K + c(S_n - x) + C_n(S_n) = K - cx + G_n(S_n), & \text{if } x < s_n; \\ -cx + G_n(x), & \text{if } x > s_n, \end{cases}$$

which is a (s_n, S_n) policy.

Examples

In this subsection we explore the condition (A) when p.d.f. is specified. If p.d.f. is uniform, exponential, normal or gamma, the condition (A) can be rewritten as in We discuss two cases.

Case 1

$$\frac{\phi'(y)}{\phi(y)} \le \frac{h+p}{B_2 - B_1} \text{for all } y \in \Re^+, B_2 - B_1 > 0,$$

If $\sup_{0<y<\infty} \frac{\phi'(y)}{\phi(y)} \le \frac{h+p}{B_2-B_1}$ holds, then the condition (A) is immediately satisfied. Therefore any uniform and exponential distributions satisfy the condition (A).

On the other hand if ϕ is a normal distribution with the mean μ and variance σ^2, then the condition (A) reduce to

$$y \ge \mu - \frac{(h+p)\sigma^2}{B_2 - B_1} > 0.$$

Since $\Pr\{-4\sigma + \mu < y < 4\sigma + \mu\} \approx 1$ and $\Pr\{0 \le y < \infty\} \approx 1$ in our model, the condition (A) satisfies

$$4\sigma \le \mu \quad \text{and} \quad \frac{(h+p)\sigma}{B_2 - B_1} > 4. \tag{3.12}$$

If $(h+p)$ is large enough compared with $|B_2 - B_1|$, then the inequality (3.12) may possibly hold.

For a gamma distribution with parameter (α, v) condition (A) reduces to

$$y \ge \frac{(v-1)(B_2 - B_1)}{\alpha(B_2 - B_1) + h + p}. \tag{3.13}$$

Therefore if

$$0 < v < 1 \tag{3.14}$$

holds, then the condition (A) holds. Furthermore, if (3.14) does not hold but $(h+p)$ is large enough compared with $|B_2 - B_1|$, then the inequality (3.13) may possibly hold.

Case 2

$$\frac{\phi'(y)}{\phi(y)} \ge \frac{h+p}{B_2 - B_1} \text{for all } y \in \Re^+, B_2 - B_1 < 0,$$

If $\in_{0<y<\infty} \frac{\phi'(y)}{\phi(y)} \ge \frac{h+p}{B_2-B_1}$ holds, then a uniform distribution immediately satisfies the condition (A).

For the exponential distribution with mean $1/\alpha$, the condition (A) reduces to

$$\alpha < |\frac{h+p}{B_2 - B_1}| < 1. \tag{3.15}$$

In this case, if $(h+p)$ is large enough compared with $|B_2 - B_1|$, then the inequality (3.15) may possibly hold.

For the normal distribution the condition (A) reduce to

$$y \le \mu - \frac{(h+p)\sigma^2}{B_2 - B_1}(< \mu).$$

Since $\Pr\{-4\sigma + \mu < y < 4\sigma + \mu\} = 1$ and $\Pr\{0 \le y < \infty\} = 1$ in our model, the condition (A) satisfies

$$4\sigma \le \mu \quad \text{and} \quad -\frac{(h+p)\sigma}{B_2 - B_1} > 4. \tag{3.16}$$

If $(h+p)$ is large enough compared with $|B_2 - B_1|$, then the inequality (3.16) may possibly hold.

For the gamma distribution the condition (A) reduces to

$$y \le \frac{(v-1)(B_2 - B_1)}{\alpha(B_2 - B_1) + h + p}. \tag{3.17}$$

Therefore if

$$\alpha(B_2 - B_1) + h + p < 0, \quad |\alpha(B_2 - B_1) + h + p| \to 0, \quad \text{and } v > 1,$$

then the inequality (3.17) may possibly hold. Summing up the above discussion, we have the following proposition which is also summarized in

Proposition 3

Case 1 For any uniform ditribution or exponential distribution, the condition (A) holds.

If ϕ is a normal distribution with (3.12), or a gamma distribution with (3.13), tthe condition (A) holds.

Case 2 For any uniform distribution, the condition (A) holds.

If ϕ is an exponential distribution with (3.15), a normal distribution with (3.16), or a gamma distribution with (3.17), then the condition (A) holds.

Conclusion

In this section we have shown under the condition (A) that the (s, S) policy is optimal for finite period stochastic inventory models with fixed inventory holding and shortage costs in addition to a fixed ordering cost. It is found that our proof of this result is diiferent from and much simpler than Aneja and Noori [3]. This section also provides an answer to the question of how robust the class of (s, S) policies is for the stochastic inventory models with fixed costs.

Furthermore, we have demonstrated that the condition (A) is necessary and sufficient for the (s, S) policy to be optimal. However, this condition may restrict on the class of probability density functions of demands. When the probability density function of demand is specified like uniform, exponential, normal or gamma, we have discussed in how the condition (A) can be rewritten to and whether it holds or not.

3.3 Inventory Control for Price Differentiable Products with No Carrying Over

Vacant hotel rooms, seats in passenger planes, and seats in concert halls are examples of what are referred to as "inventory" in their respective business circles. The special feature of such inventory items is that the saleable total capacity is fixed in advance and it is impossible to carry over any remaining inventory to the next day. In order to counteract this physical property of inventory that is incapable of being carried over to the next day, businesses with such inventory aim at guaranteeing demand by setting multiple prices on inventory items of identical quality and issuing a variety of discount tickets at a relatively early stage. Users, on the other hand, also have different preferences in regard to

inventory items of identical quality. Thus, for example, people who use passenger planes for tourism purposes generally make reservations at a comparatively early time and at the cheapest discount fare possible, as opposed to those who use passenger planes for business trips.

In this section we hypothesize a case in which, when the amount of saleable inventory is fixed in advance, two types of demands for such inventory items occur as a result of differences in the time that demand arises and in profitability. We consider the problem of deciding how to distribute the inventory items between these two types of demands.

We define the two types of demands for inventory items that we are dealing with in this study as follows. The demand that arises at an early period we refer to as "early demand," and its profitability is low. The demand that arises at a late stage we refer to as "late demand," and its profitability is high. Since what we are dealing with here is the sale of space, two special features of these inventory items are that they cannot be carried over to the next day and that any demand that is not filled is lost forever. Businesses that deal in this kind of inventory items seek a control policy that will maximize expected total revenue, something that will stir up demand by offering the goods at an early period at low rates but without thereby missing out on the demand that produces high profit. In this section we formulate this type of decision problem in general form, as a problem of maximizing expected revenue. The information obtained in this way is desirable because it enables one to establish the upper limit to the amount of inventory items to be allocated to the lower-profitability early demand in cases where it is anticipated that the more highly profitable late demand will arise at a later future time. This information is also applicable to bargain sales held after the peak selling period of seasonal goods has been passed. A sales strategy often used in bargain sales is that of setting an upper limit on the saleable quantity (amount of inventory) and indicating that the goods can be purchased by the first whatever number of customers. The idea behind this sales strategy can be seen as one of arousing demand that has passed its peak but at the same time preventing those customers who are willing to purchase similar goods at the normal prices from shifting to the purchase of bargain goods.

Beckmann [9], Brumelle et. al. [21], Rothstein [92], and Sawaki [96] have developed similar discussions in regard to airline seat management. Belobaba [12] and Sawaki [96] have provided overviews of airline seat management and yield management. Liberman and Yechiali [62] have considered hotel room inventory control. In all these previous models the authors assume the independence of two types of demands, but in this section that kind of independence will not be assumed. Accordingly, it recognizes the possibility that early demand that has been unable to purchase at a discount price will be willing to purchase at the normal price and will shift to late demand.

In the following subsection a model formulation is followed by an analysis of what the optimal inventory policy might be. We pay particular attention to what the suficient conditions are for a simple optimal policy to exist. In the second subsection we discuss in detail the optimal inventory policy of this model when demand distribution has been specified. Then in the third subsection, we discuss an overbooking model, and then the conclusion, we bring together the information obtained by the first model and then touch upon the range of application of the model and directions for future expansion.

3.3.1 Model Formulation and Optimal Policy

Let us express the total capacity of presently saleable inventory items as C, let it be a fixed and given value, and let it be expressed in the late demand units to be described next. The two types of demands for this inventory items will be X and Y, which we shall call, respectively, the early demand (demand that is realized early) and late demand (demand that is realized at a late stage). At the beginning of the planning period both X and Y are random variables, and we posit the distribution function of X to be $F(x)$ and posit the conditional distribution of Y when $X = x$ is given as $G(y|x)$. We posit the decision variable to be I, which we assume to be the upper limit of the amount of inventory to be allocated to early demand. In other words, the revenue from early demand will be $\min\{I, X\}$. Accordingly, when the allocation of inventory to late demand is $Q(I)$, then

$$Q(I) = C - \alpha \min\{I, X\} \tag{3.18}$$

where α is the exchange rate of one early demand unit into late demand, $0 \le \alpha \le 1$. If $\alpha = 1$, then one early demand unit is equal to one late demand unit, and if $0 \le \alpha < 1$, then the case of returned goods or overbooking is hypothesized. In the case of passenger

planes, if a change in the size of the seating is possible (taking early demand as group tourist economy seats and late demand as individual passenger business seats), then for example, α becomes $2/3$. In addition, the following symbols are used:

p_1 = lower price of early demand

p_2 = normal price of late demand

h = unit holding cost of the product unsold

s_1 = unit shortage cost for early demand

s_2 = unit shortage cost for late demand

Assumption 1 $0 < p_1 + s_1 < \alpha(p_2 + s_2)$

It is possible to assume $p_1 < p_2, s_1 < s_2$ in order to guarantee that the profitability of early demand will be smaller than the profitability of later demand, but in order to obtain the optional inventory control given below, Assumption 1 is sufficient.

If we assume $T(I)$ to be the expected total profit when an inventory level of I units has been allocated to early demand, then $T(I)$ is given by the following equation:

$$\begin{aligned} T(I) &= p_1 \cdot E\left[\min\{X, I\}\right] + p_2 \cdot E\left[E_{Y|X}\left[\min\{Y, Q(I)\}\right]\right] \\ &\quad - h \cdot E\left[E_{Y|X}\left[\max\{Q(I) - Y, 0\}\right]\right] \\ &\quad - s_1 \cdot E\left[\max\{X - I, 0\}\right] - s_2 \cdot E\left[E_{Y|X}\left[\max\{Y - Q(I), 0\}\right]\right] \end{aligned} \tag{3.19}$$

The problem is to allocate C amount of inventory between early and late demands so as to maximize expected total profit under the condition $0 \le \alpha I \le C$. Figure 3.1 shows the fluctuations in amount of inventory when $\alpha = 1,\ X > I,\ Y < C - I$. In the case described in Figure 3.1 the result was that the allocation of inventory to early demand was too little, and in consequence unsold inventory remained even after late demand was satisfied. Figure 3.2, on the contrary, illustrates what happens when $\alpha = 1,\ X < I,\ Y > C - I$ so that the amount of inventory allocated to early demand is too great and the more profitable late demand is lost. In order to find the optimal allocation of inventory between the two demands, let us prepare the following assumption.

Assumption 2 $P\{Y \ge C - \alpha \cdot I \mid X \ge I\}$ is increasing in I.

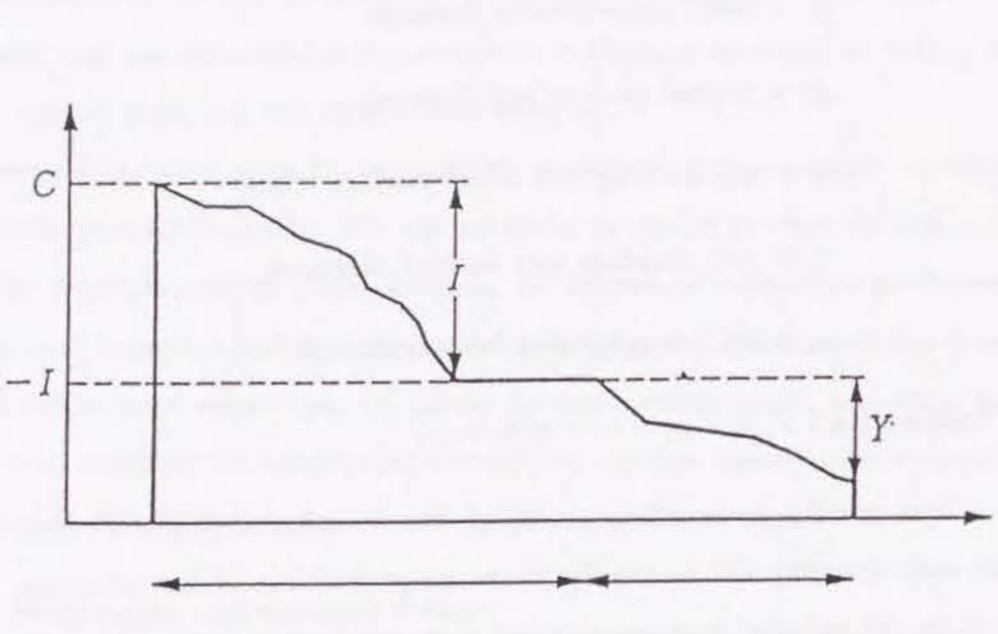


Figure 3.1 Inventory Fluctuation ($\alpha = 1, X > I, Y < C - I$)

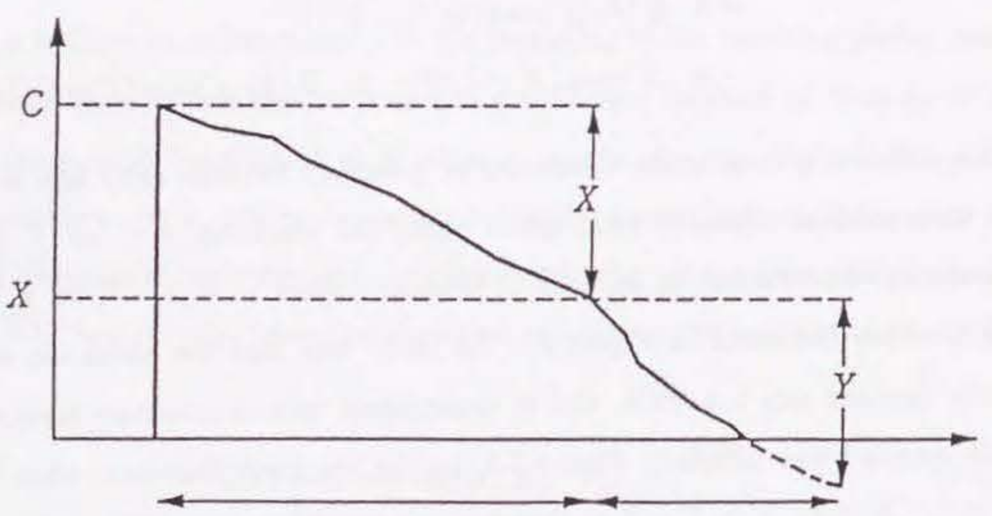


Figure 3.2 Inventory Fluctuation ($\alpha = 1, X < I, Y < C - I$)

Lemma 3.3.1 *Under Assumptions 1 and 2, $T(I)$ is unimodal in I.*

Proof Taking the derivative of (3.19) we obtain

$$\frac{dT(I)}{dI} = \bar{F}(I)\{(p_1 + s_1 + \alpha h) - \alpha(p_2 + s_2 + h)P\{Y > C - \alpha I \mid X \geq I\}\} \quad (3.20)$$

Since $\bar{F}(\cdot) \geq 0$ and the content of the bracket is decreasing in I, the sign of the bracket is changed only once, that is, there exists I^* such that $dT(I)/dI = 0$, and then $T(I)$ is increasing in $I < I^*$ and is decreasing in $I > I^*$, which implies that $T(I)$ is unimodal in I.

Theorem 3.3.1 *Assume that $F(I) < 1$ for all I, $P[Y > C] < (p_1+s_1+\alpha h)/\alpha(p_2+s_2+h)$ and $P[Y > (1-\alpha)C \mid X \geq C] > (p_1 + s_1 + \alpha h)/\alpha(p_2 + s_2 + h)$. Then, an optimal upper limit I^* for early demand is given by*

$$I^* = \max\left\{0 \leq \alpha \cdot I \leq C \mid P[Y > C - \alpha \cdot I \mid X \geq I] \leq \frac{p_1 + s_1 + \alpha \cdot h}{\alpha(p_2 + s_2 + h)}\right\} \quad (3.21)$$

Remarks 2

(i) From Theorem 3.3.1 we conclude that when $P[Y > C] \geq (p_1 + s_1 + \alpha \cdot h)/\alpha(p_2 + s_2 + h)$, $I^* = 0$. This implies that an optimal allocated value must be 0 when the late demand is sufficiently larger, compared with the price ratio. Conversely, when $P[Y > (1-\alpha)C \mid X \geq C] \leq (p_1 + s_1 + \alpha \cdot h)/\alpha(p_2 + s_2 + h)$, then $I^* = C$.

(ii) Note that an optimal upper limit for early demand depends only on the relative price ratio, but does not specific values of prices.

Corollary If $\alpha = 1$ and X and Y are stochastically independent, then an optimal upper limit I^* of the product allocation for the early demand is given by

$$I^* = \begin{cases} C & \text{if } \bar{G}(0) \leq r \\ C - \bar{G}^{-1}(r) & \text{if } \bar{G}(C) < r < \bar{G}(0) \\ 0 & \text{if } r \leq \bar{G}(C) \end{cases} \quad (3.22)$$

where $\bar{G} = 1 - G$ and $r \equiv \frac{p_1+s_1+h}{p_2+s_2+h}$.

Remarks 3

(i) An optimal protection limit for late demand $C - I^*$ is equal to $\bar{G}^{-1}(r)$, which does not depend on C and F.

(ii) I^*/C is increasing in C.

(iii) I^* depends only on G and r, and not on F.

3.3.2 Examples

This subsection is a discussion of special cases of the model presented in the preceding subsection and cases in which the distribution function of demand is specified.

Example 1 Early demand is sufficient.

Let $Pr(X > I) = 1$ for all I in which $\alpha = 1$ and $I \leq C$.
Then $Pr[Y > C - I | X > I] = Pr[Y > C - I] = \bar{G}(C - I)$.
When $\alpha = 1$, $T(I)$ in Equation (3.19) becomes

$$\begin{aligned} T(I) &= p_1 \cdot I + p_2 \cdot E[\min\{Y, C - I\}] \\ &\quad - h \cdot E[\max\{C - I - Y, 0\}] \\ &\quad - s_1 \cdot E[X - I] - s_2 \cdot E[\max\{Y - C - I, 0\}] \end{aligned} \tag{3.23}$$

It is easy to see that Equation (3.23) is concave in I. Taking the derivative of (3.23) we have

$$\begin{aligned} \frac{dT(I)}{dI} &= p_1 - p_2 P[Y > C - I] + hP[Y < C - I] + s_1 - s_2 P[Y > C - I] \\ &= (p_1 + s_1 + h) - (p_2 + s_2 + h)P[Y > C - I] \end{aligned} \tag{3.24}$$

Hence, an optimal upper limit for early demand is given by

$$I^* = C - \bar{G}^{-1}\left(\frac{p_1 + s_1 + h}{p_2 + s_2 + h}\right) \tag{3.25}$$

which is essentially the same equation as in Theorem 3.3.1, but does not require Assumption 2.

Example 2 Demand follows bivariate normal distribution.

Let $\alpha = 1$ and X and Y are bivariately normally distributed. When X is positively correlated with Y, Assumption 2 holds. In this example we provide numerically optimally

Table 3.1: Optimal allocation values, $(I^*, C - I^*)$

ρ	C					
	40	60	80	100	120	140
Independent ($\rho = 0$)	(13, 27)	(33, 27)	(53, 27)	(73, 27)	(93, 27)	(133, 27)
Dependent ($\rho = 0.9$)	(13, 27)	(32, 28)	(49, 31)	(65, 35)	(81, 39)	(97, 43)
$I^*/C(\rho = 0)$	0.33	0.55	0.66	0.73	0.77	0.80
($\rho = 0.9$)	0.33	0.53	0.61	0.65	0.67	0.69

distributed values for early and late demands. The means and standard deviations are as follows:

$$\mu_X = 70, \quad \mu_Y = 30, \quad \sigma_X = 26.5, \quad \sigma_Y = 11.5,$$
$$(p_1 + s_1 + h)/(p_2 + s_2 + h) = 0.6$$

Optimal allocation values for early and late demands are summarized in Table 3.1.

Note that when $\rho = 0$, $C - I^*$ is always equal to 27. In a case of $\rho = 0$ optimal upper limits for early demand are bigger than ones in a case of $\rho = 0.90$.

Example 3 Demand is exponentially distributed.

In this example we assume that X is exponentially distributed with mean $1/\lambda$ and $Y = X/\beta$, $\beta > 0$, $\alpha = 1$. $1/\beta$ of early demand turns into late demand.
$P\{Y > C - I \mid X \geq I\} = e^{-\beta\lambda C} e^{\lambda(\beta+1)I}$ is increasing in I

$$I^* = \frac{\beta}{\beta - 1} C + \log\left(\frac{p_1 + s_1 + h}{p_2 + s_2 + h}\right) \leq C. \tag{3.26}$$

If X and Y are independent, then we have

$$\begin{aligned} P\{Y > C - I\} &= e^{-\lambda\beta(C - I)} \\ I^{**} &= C + \log\left(\frac{p_1 + s_1 + h}{p_2 + s_2 + h}\right)^{1/(\lambda\beta)} \end{aligned} \tag{3.27}$$

From the argument above, we obtain the finding that $I^* < I^{**}$ for $\beta > 0$.

3.3.3 An Overbooking Model

Suppose that there is only one type of product, say a normal fare class. With each fare booked we associate a random variable

$$Z_i = \begin{cases} 1 & \text{if the } i^{th} \text{ fare booked confirms} \\ 0 & \text{if the } i^{th} \text{ fare cancels} \end{cases}$$

Assume that $Z_1, Z_2, \cdots$ are independent and identically distributed with $EZ_i = \beta$. If I seats are booked, then $N(I) = \sum_{i=1}^{I} Z_i$ seats are confirmed. Then the expected total revenue is

$$R(I) = p_2E[N(I)] - (p_2 + p_0)E[N(I) - C \mid N(I) > C]P\{N(I) > C\} \tag{3.28}$$

where p_0 is the penalty cost due to overbooking.

Theorem 3.3.2 *If Z_i is independent and identically distributed, then an optimal booking limit is given by*

$$I^*(C) = \max\left\{I \mid P[N(I) > C] < \frac{p_0}{p_2 + p_0}\right\} \tag{3.29}$$

Proof It follows from the fact that the incremental expected revenue is given by

$$\begin{aligned} \Delta R(I) &= R(I+1) - R(I) \\ &= p_2\beta - (p_2 + p_0)\beta P\{N(I) > C\} \end{aligned} \tag{3.30}$$

and $N(I)$ is increasing in I, which implies that $R(I)$ is concave in (I).

Examples:

(i) Z_i is Bernoulli random variable. Set $p_0 = p_2$ and $E(Z_i) = \beta = 1/2$. So $I^*(C) = 2C$.

(ii) $N(I)$ is normally distributed with mean $I\beta$ and variance $I\beta(1-\beta)$.

$$\begin{aligned} I^*(C) &= \max\left[I \mid P\left\{N(I) \geq \frac{C - \beta I}{\sqrt{I\beta(1-\beta)}}\right\} < \frac{p_0}{p_2 + p_0}\right] \\ &= \begin{cases} \frac{C}{\beta} + \xi - \sqrt{(\xi + \frac{C}{\beta})^2 - (\frac{C}{\beta})^2} & \text{if } \frac{p_0}{p_0+p_2} \leq \frac{1}{2} \\ \frac{C}{\beta} + \xi + \sqrt{(\xi + \frac{C}{\beta})^2 - (\frac{C}{\beta})^2} & \text{if } \frac{p_0}{p_0+p_2} > \frac{1}{2} \end{cases} \end{aligned} \tag{3.31}$$

where z is defined by $Pr\{Z > z\} = \frac{p_0}{p_2+p_0}$ for Z normally distributed and

$$\xi = \frac{z^2}{2} \cdot \frac{1-\beta}{\beta}$$

3.3.4 Conclusion

In this section we have analyzed the best way to allocate a fixed amount of inventory to meet two types of demands differing in the times at which they arise. We have shown that, when a higher-profitability late demand is expected to arise at a future stage, it is possible to maximize expected revenue obtained from the allocation of total inventory by establishing an upper limit on the amount of inventory allocated to lower-profitability early demand. In those cases in which there is a stochastic dependency between early demand and late demand, if there is a certain type of monotonicity in the conditional distribution (Assumption 2), then the optimal allocation of inventory to early demand I^* is given by the equation

$$I^* = \frac{\beta}{\beta - 1}C + \log\left(\frac{p_1 + s_1 + h}{p_2 + s_2 + h}\right) \leq C \tag{3.32}$$

If there is the possibility that customers who have been unable to purchase the inventory items during the early demand period will purchase them as late demand items at higher price, then setting an upper limit to the allocation of inventory items to lower-profitability early demand agrees with our economic intuition.

The model dealt with here is a single-period static model; it is not a dynamic inventory control model involving close observation of the process in which early demand is realized, and revision of the upper limit of inventory allocation. A model for sequentially revising the upper limit of inventory allocation while constantly monitoring the reduction in inventory can probably be considered next. In addition, by way of extension of the model, instead of discussing only two types of demands we could discuss an inventory allocation policy for the more general case of N types of demands. In either of these later cases the model would be more complex and more difficult than the model dealt with here, but it definitely is an important and real problem for all businesses that sell goods of such a nature that inventory cannot be carried over and total inventory capacity is fixed.

Chapter 4

Airline Seat Allocation Models

4.1 Introduction

In an era of increasing pricing freedom, airline companies no longer offer seats for sale at a single fare. Recognizing that different groups of consumers have different willingness to pay for the same seat, airline companies offer seats at a wide range of air fares. However, to prevent consumers willing to pay high fares from buying seats at low fares, the airlines attach various restrictions to their tickets such as early time bookings, Saturday night stayovers or reduced service without food to discount fares (see Belobaba [13], Sawaki [104]).

The process of determining fares, associated restrictions and the number of seats to offer at a given fare is referred to as "Airline Revenue Management" (see [12]). Within this area of airline revenue management, the decision process determining the number of seats to be protected for various classes of passengers is called the airline seat allocation problem. The key idea of the seat allocation problemm is to limit the number of discounted seats when there is a strong demand expected from high fare consumers, so as to maximize the expected revenue per flight. Airline revenue management includes

(i) determination of high and low fare levels,

(ii) determination of restrictions associated with low fares, and

(iii) dynamic monitoring of seat for sale on a given flight and readjusting the allocation of seats between high and low fares so as to maximize expected revenues.

The purpose of this chapter is to analyze some of the first element (i), determination of the initial allocation of seats between high and low fares. This static model with one period is of great importance to airlines as the two classes of passengers have very different time patterns of booking. High fare passengers tend to book in the last days or hours before a departure time.

In Section 4.2 we consider a dynamic airline seat allocation problem for a single flight with two fare classes, based upon Sawaki [104]. The problem is formulated as an N-step dynamic problem and aims at deriving optimal policies. We also explore some analytical properties of such an optimal seat allocation policy and the associated expected revenue. The model also extends the existing literature in two ways. First, it is a dynamic version with the cost of lost sales. Second, it is formulated under the setting of Markov decision processes which explicitly take into account the periods remaining until departure and permite reopening of fare classes.

In Section 4.3 we consider the airline seat allocation between high and low fares with and without stochastic cancellations. We also analyze the problem of simultaneously determining allocation and overbooking levels for two different classes of passengers, which also extends the existing literature in three ways. First, the cost of lost sales, which has been ignored in the existing literature, is explicitly incorporated into the model. Second, the over booking phenomenon is also explicitly treated. Third, the concept of spill rate is clarified into the passeger spill rate and the flight rate. It is found that the results obtained here are in closer agreement with actual airline practice.

Section 4.4 examines the problem, treated in [21], of allocating airline seats between two nested fare classes when the demands for the classes are stochastically dependent. The well known simple seat allotment formula of Littlewood which requires the assumption of statistical independence between demands is generalized to a formula which requires only a much weaker monotonic association assumption. The model employed here is also used to examine the problems of full fare passenger spillage and passenger upgrades from the discount class.

4.2 A Dynamic Airline Seat Allocation Model

The deregulation of the North America airlines allowed the airline industry to undergo major changes in price competition, skillful dynamic pricing policies and seat allocation management. These changes helped to stimulate the demand for air travelling. On the other hand, the deregulation challenges airline companies with an important managerial problem of determining an optimal booking policy which allocation optimally the seats of an airline among the various fare classes.

It is in an airline company's interest to control the booking process by selling the right seats to the right passengers at the right prices and timing to maximize the total revenue acquired from a single airplane.

This section addresses a dynamic airline seat allocation problem. This dynamic model is very important because it allows us to monitor dynamicly available seats and readjust the seat allocation among fare classes on a given flight. Another advantage of dynamic models is of the dynamic reallocation of seats as time progresses, reflecting the actual booking progress. Optimal seat allocation for a dynamic booking limit revision process is in fact different from the allocations derived previously for a static case (see [21], [65], [92], [96]).

In this respect, the section intends to make a significant contribution to the existing literature by dealing with dynamic aspects of airline seat inventory control.

Liberman and Yechiali [62] presents a model for determining an overbooking policy for a hotel with a single fare class. Rothstein [92], and [93] formulate an airline overbooking model with a single fare class as a Markov decision process which allowed for dynamic adjustment of overbooking levels as the day offlight departure approached, but lacks a formal derivation and investigation into the properties of the associated optimal expected revenue. The first significant result on the seat allocation problem was presented by Littlewood [65] who proposed a simple seat allocation rule by using the marginal revenue analysis. Richter [86] also proposed a seat allocation model in a simplified manner. More recently, Belobaba [12] generalized the results above to more than two fare classes. None of these works but [21] allowed both for the two fare classes. Belobaba [12] is also a recent good survey article on airline seat management.

These works of Belobaba [12], Sawaki [96], Brumelle et al [21], Rothstein [92], Littlewood [65], Richter [86] are based on common assumptions as follows;

1. single flight leg : Bookings are made on the basis of a single departure and landing.
2. independent demand : The demands for the different fare classes are mutually independent.
3. low fare booking first : The lowest fare reservations requests arrive first, followed by the next lowest, etc.
4. no cancellations : Cancellations,'no-show'and overbooking are not considered.
5. limited information : The decision to close a class is based only on the number of current bookings.
6. nested classes : Any fare class can be booked into seats not taken by bookings in lower fare classes.

While assumption 6 is a common practice in airline reservation systems today, assumptions 1 through 5 are restrictive. These sometimes overly restrictive assumptions serve the purpose of making the problem tractable.

The purpose of this section is to deal with the problems above, that is, a fomal derivation for an optimal dynamic seat allocation rule and investigation into the properties of the associated expected revenue. The impact of uncertain demand on the optimal policy and associated revenue is also explored. In Subsection 4.2.1, we formulate a dynamic seat allocation problem with no cancellation allowed. In Subsection 4.2.2, we show the existence of an optimal policy and discuss properties of optimal policy and associated revenue.

4.2.1 Dynamic Seat Allocation Problems

In this subsection, we consider dynamic models for airline seat management shere there are N periods before the flight departure. Model is in discrete time $t = 1, 2, \ldots, N$. We can time 1 the initial time and N the departure time. The time intervals need not be evenly spaced. It might be best to space them widely at first when the demand for seats is relatively low, and then make the intervals smaller as the departure time nears. This might be important for the following assumption (C) to hold.

Suppose that there are two fare classes, low and high fares.

Assumption (A) No passengers cancel their reservations.

Assumption (B) In each decision period, low fare booking occur earlier than the high fares.

Assumption (C) High and low fare demands are both independent of demands in other periods.

Assumption (D) The refused passengers are not picked up by other flights of the same carrier.

Assumption (A) will be relaxed later. Assumption (B) is not so restrictive since we do assume "early bird" in each period, but do not over the entire periods. Assumption (D) is for making the flight revenue maximization criterion reasonable. In each period, after observing the number of seats available, we determine the number of seats to allocate for low fare demands, then accept for booking of low fare demands up to the number of seats allocated and accept the high fare bookings as many as available. Our objective is to maximize the expected total revenue obtained from the flight over the entire periods. Indices i and j denote low and high fare demands, respectively. We use the following notations;

p_1, p_2 = high and low fares, respectively,

s_t = the number of unbooked seats remaining at the beginning of period t,

$L_t(s)$ = the maximum number of seats to allocate for low fare demand in period t out of s seats available,

$q_i^2(t)$ = probability that i low fare bookings occur in period t,

$q_{ij}^1(t)$ = probability that j high fare bookings occur in period t, given i low fare bookings requested.

Note that high fare demand depends on low fare which allows the possibility of passengers switching from low fare seats to high fare, so called " grade up ". We also use the notation

$a \vee b = \max\{a, b\}, a \wedge b = \min\{a, b\}$ and $a^+ = \max\{a, 0\}$. Any seats remaining at the end of each period are passed on to the next period. Then we have the following seat inventory identity:

$$s_{t+1} = [s_t - i \wedge L_t(s_t) - j]^+ \tag{4.1}$$

where t=1, 2, ·, N-1 and i, j denote low and high fare demands, respectively. We wish to determine $L_t(s)$ for each t and s in a sequential order so as to maximize the expected total revenue. Let $V_t(s)$ be the maximum expected revenue obtained if there are s seats remaining at the beginning of period t and the flight departs at period N. Then, we obtain the optimality equation from the principle of optimality

$$\begin{aligned} V_t(s) &= \max_{0 \le L \le s} \{p_2 E_i(i \wedge L) + p_1 E_i E_{j|i}[j \wedge (s - i \wedge L)] \\ &\quad + E_i E_{j|i} V_{t+1}([s - i \wedge L - j]^+)\} \end{aligned} \tag{4.2}$$

where $L \equiv L_t(s)$, $t = 1, 2, \cdots, N$, $V_t(0) = 0$ for all t and $V_{N+1}(\cdot) \equiv 0$. Equation (4.3) can explicitly be rewritten as follows;

$$\begin{aligned} V_t(s) &= \max\{p_2 \sum_{i=0}^{L_t(s)-1} i q_i^2(t) + p_2 L_t(s) \sum_{i=L_t(s)}^{\infty} q_i^2(t) & (4.3) \\ &\quad + p_1 (\sum_{i=0}^{L_t(s)-1} q_i^2(t) [\sum_{j=0}^{s-i-1} j q_{ij}^1(t) + (s - i) \sum_{j=s-i}^{\infty} q_{ij}^1(t)]) \\ &\quad + p_1 (\sum_{i=L_t(s)}^{\infty} q_i^2(t) [\sum_{j=0}^{s-L_t(s)-1} j q_{ij}^1(t) + (s - L_t(s)) \sum_{j=s-L_t(s)}^{\infty} q_{ij}^1(t)]) \\ &\quad + \sum_{i=0}^{L_t(s)-1} q_i^2(t) \sum_{j=0}^{s-i} q_{ij}^1(t) V_{t+1}(s - i - j) \\ &\quad + \sum_{i=L_t(s)}^{\infty} q_i^2(t) \sum_{j=0}^{s-L_t(s)} q_{ij}^1(t) V_{t+1}(s - L_t(s) - j)\} \\ &\equiv \max_{0 \le L_t(s) \le s} \{R_t(s, L_t(s))\} & (4.4) \end{aligned}$$

where $R_t(\cdot, \cdot)$ denotes the expression inside the braces of equation (4.3) or (4.4).

Assumption (E) $L_t(s)$ is nondecreasing in s for each t.

Assumption (F) For each t $\sum_{i=L+1}^{\infty} q_i^1(t) q_{i,s-L-1}^2(t) \ge q_{L+1}^2(t) \sum_{j=s-L-1}^{\infty} q_{L+1,j}^1(t)$ for all $L \le s$.

To establish the existence of a simple optimal seat allocation policy under assumptions (E), (D) we need the following lemma.

Lemma 4.2.1

(i) $R_t(s, L_t(s))$ is concave in L for each t and s, and nonincreasing in s for each t.

(ii) $V_t(s)$ is nondecreasing and concave in s for each t, and nonincreasing in t for each s.

Proof Let $\Delta R_t(L) = R_t(s, L_t(s)+1) - R_t(s, L_t(s))$. We wish to establish the concavity of $R_t(\cdot,\cdot)$ in $L_t(s)$ for fixed s and t. It will be sufficient to demonstrate that $\Delta R_t(L)$ is nonincreasing in L to establish concavity. Let $S_t(L) = p_2 E_i(i \wedge L) + p_1 E_i E_{j|i}[j \wedge (s - i \wedge L)]$ and $T_t(L) = E_i E_{j|i} V_{t+1}([s - i \wedge L - j]^+)$. Then, $R_t(s, L) = S_t(L) + T_t(L)$. Letting $\Delta S_t(L) = S_t(L+1) - S_t(L)$ and $\Delta T_t(L) = T_t(L+1) - T_t(L)$, we have $\Delta R_t(L) = \Delta S_t(L) + \Delta T_t(L)$. We will examine $\Delta S_t(L)$ and $\Delta T_t(L)$ separately. After calculating and rearranging $\Delta S_t(L)$ term by term, we have

$$\Delta S_t(L) = p_2 \sum_{i=L+1}^{\infty} q_i^2(t) - p_1 \sum_{i=L+1}^{\infty} q_i^2(t) \sum_{j=s-L}^{\infty} q_{ij}^1(t). \tag{4.5}$$

$$\begin{aligned} \Delta^2 S_t(L) &= S_t(L+1) - \Delta S_t(L) \\ &= -p_2 q_{L+1}^2(t) - p_1 [\sum_{i=L+1}^{\infty} q_i^2(t) q_{i,s-L-1}^1(t) - q_{L+1}^2(t) \sum_{j=s-L-1}^{\infty} q_{L+1,j}^1(t)] \\ &< 0 \quad (\text{ by assumption (F)}) \end{aligned}$$

which implies that $S_t(L)$ is concave in L. Collecting terms and noting that $V_t(0) = 0$, we have

$$\begin{aligned} \Delta T_t(L) &= T_t(L+1) - T_t(L) \\ &= \sum_{i=L+1}^{\infty} q_i^2(t) \sum_{j=0}^{s-L-1} q_{ij}^1(t) [V_{t+1}(s-L-1-j) - V_{t+1}(s-L-j)]. \end{aligned} \tag{4.6}$$

Combining equation (4.5) and (4.6) we have

$$\begin{aligned} \Delta R_t(L) &= \sum_{i=L+1}^{\infty} q_i^2(t) \{p_2 - p_1 \sum_{j=s-L}^{\infty} q_{ij}^1(t) \\ &\quad + \sum_{j=0}^{s-L-1} q_{ij}^1(t) [V_{t+1}(s-L-1-j) - V_{t+1}(s-L-j)]\}. \end{aligned}$$

For an inductive argument letting $t = N$, we have

$$R_N(s, L_N(s)) = S_N(L_N(s))$$

which has been shown to be concave in L. Then

$$V_N(s) = \max_{0 \le L(s) \le s} \{S_N(L(s))\}.$$

Since $S_t(L(s))$ is strictly concave in L, there exists a unique solution L^* such that

$$V_N(s) = \begin{cases} S_N(s) & for \quad s \le L^*, \\ S_N(L^*) & for \quad s > L^*, \end{cases} \tag{4.7}$$

which is certainly concave and nondecreasing in s. Suppose that $V_{t+1}(s)$ is concave and nondecreasing in s, and consider $\Delta T_t(L)$ given by (4.6). By the induction step, $\Delta^2 T_t(L) = \Delta T_t(L+1) - \Delta T_t(L)$ is negative, and hence $T_t(L)$ is concave in L. So is $R_t(s, L)$. By assumption (E) $R_t(s, L(s))$ is nondecreasing in s. Hence, we have

$$V_t(s) = \max_{0 \le L \le s} \{R_t(s, L(s))\}$$

which is concave and nondecreasing in s for each t. It is obvious from the definition that $V_t(s)$ is nonincreasing in t for each s. □

4.2.2 Optimal Seat Allocation Rules

In this subsection we show under assumptions (C), (D) that there exists a simple optimal policy-a control limit type of seat allocation for low fares. Under rather restrictive conditions some analytical properties of the optimal policy are explored. Also, a special case with sufficient low fare demands is analyzed.

Theorem 4.2.1 *There exists a sequence of optimal seat allocations for low fares at each period $(L_N^*(s), L_{N-1}^*(s), \cdots, L_2^*(s), L_1^*(s))$ such that*

$$\begin{aligned} L_t^*(s) &= \max\{L \le s : \sum_{i=L+1}^{\infty} q_i^2(t) [p_1 \sum_{j=s-L-1}^{\infty} q_{ij}^1(t) + \sum_{j=0}^{s-L-1} q_{ij}^1(t) V_{t+1}(s-L-j)] \\ &\le \sum_{i=L+1}^{\infty} q_i^2(t) [p_2 + \sum_{j=0}^{s-L-1} q_{ij}^1(t) V_{t+1}(s-L-1-j)]\}, \quad t = 1, 2, \ldots, N. \end{aligned} \tag{4.8}$$

Proof For $t = N$ we have shown by equation (4.7) in Lemma 4.2.1 that $V_N(s)$ is maximized at L_N^* given by

$$L_N^*(s) = \max\{L \le s : p_1 \sum_{i=L+1}^{\infty} q_i^2(N) \sum_{j=s-L-1}^{\infty} q_{ij}^1(N) \le p_2 \sum_{i=L+1}^{\infty} q_i^2(N)\}$$

For $t < N$ we have

$$V_t(s) = \max_L\{R_t(s, L(s))\}.$$

From Lemma 4.2.1, $R_t(\cdot, L)$ is concave in L and

$$\begin{aligned}\Delta R_t(L) &= \sum_{i=L+1}^{\infty} q_i^2(t)\{p_2 - p_1 \sum_{j=s-L}^{\infty} q_{ij}^1(t) \\ &\quad + \sum_{j=0}^{s-L-1} q_{ij}^1(t)[V_{t+1}(s-L-1-j) - V_{t+1}(s-L-j)]\}\end{aligned}$$

Then, by the same argument used with $R_N(s, L)$, we see that $R_t(s, L(s))$ is maximized at

$$L_t^*(s) = \max\{L : \Delta R_t(s, L) \geq 0\}.$$

Hence, we obtained the sequence of optimal decision rules $\{L_t^*(s)\}_{t=1}^N$. □

Corollary Suppose that $q_{ij}^1(t) = q_j^1$ for all i and t and $N = 1$. The one-period optimal seat allocation rule can be reduced to as follows :

$$L^*(s) = \max\{L : p_2/p_1 \geq \sum_{j=s-L+1}^{\infty} q_j^1\} \quad (4.9)$$

where $L^*(s) = 0$ if the above set is empty. This is a discrete version of a simple seat allocation model studied by many authors [12], [34], [45], [58], [62]. Equations (4.8) and (4.9) can be interpreted as follows ; if once the expected marginal revenue from high fares is larger than or equal to the one from low fares, we should stop allocating seats to low fare demands.

Theorem 4.2.2 *Assume that $q_i^2(t) = q_i^2, q_{ij}^1(t) = q_j^1$ for all i, j, t and $\Delta V_t(L) = V_t(L) - V_t(L-1)$ is non-increasing in t for each L. Then $L_t^*(s)$ is non-decreasing in t for each s.*

Proof Under the assumptions canceling the $\sum q_i^2$ terms $L_t^*(s)$ simplifies to

$$\begin{aligned}L_t^*(s) &= \max\{L \leq s : p_1 \sum_{j=s-L+1}^{\infty} q_j^1 + \sum_{j=0}^{s-L-1} q_j^1 V_{t+1}(s-L-j) \\ &\quad \leq p_2 + \sum_{j=0}^{s-L-1} q_j^1 V_{t+1}(s-L-1-j)\} \\ &= \max\{L \leq s : p_1 \sum_{j=s-L+1}^{\infty} q_j^1 + \sum_{j=0}^{s-L-1} q_j^1 \Delta V_{t+1}(s-L-j) \leq p_2\} \\ &\leq \max\{L \leq s : p_1 \sum_{j=s-L+1}^{\infty} q_j^1 \sum_{j=0}^{s-L-1} q_j^1 \Delta V_{t+2}(s-L-j) \leq p_2\}\end{aligned}$$

$$\begin{aligned}&= max\{L \leq s : p_1 \sum_{j=s-L+1}^{\infty} q_j^1 + \sum_{j=0}^{s-L-1} q_j^1 V_{t+2}(s-L-j) \\ &\quad \leq p_2 + \sum_{j=0}^{s-L-1} q_j^1 V_{t+2}(s-L-1-j)\} \\ &= L_{t+1}^*(s)\end{aligned}$$

quad□

Remarks 1 *Theorem 4.2.2 implies that as the time becomes closer to the flight departure, we should allocate more seats to low fare demands whenever there are still s seats available. The assumption that $\Delta V_t(s)$ is non-increasing in t is intuitively reasonable since an additional seat available in some period is more likely to be utilized than one made available in the next period.*

A special case : low fare demands are large enough

We shall consider a special case in which low fare demands are large enough to sell as many as we desire up to the maximum level allocated for low fares, and in which demand distributions are independent of time t, that is

$$q_i^2(t) = q_i^2, \quad q_{ij}^1(t) = q_j^1 \quad \text{and} \quad \sum_{i=L}^{\infty} q_i^2 = 1 \quad \text{for} \quad all \quad L \leq s.$$

Let $\overline{V}_t(s)$ be the maximum revenue corresponding to the special case. Then, equation (4.3) can simplify to

$$\begin{aligned}\overline{V}_t(s) &= \max_{0 \leq L \leq s}\{p_2 L + E_{j|L}[p_1(j \wedge (s-L)) + V_{t+1}((s-L-j)^+)]\} \\ &= \max\{p_2 L + p_1 \sum_{j=0}^{s-L-1} j q_j^1 + p_1(s-L) \sum_{j=s-L}^{\infty} q_j^1 \\ &\quad + \sum_{j=0}^{s-L} q_j^1 \overline{V}_{t+1}(s-L-j)\} \\ &\equiv \max\{\overline{R}_t(s, L)\}\end{aligned}$$

where

$$\overline{R}_t(s, L) = p_2 L + p_2 \sum_{j=0}^{s-L-1} j q_j^1 + p_1(s-L) \sum_{j=s-L}^{\infty} q_j^1 + \sum_{j=0}^{s-L} q_j^1 \overline{V}_{t+1}(s-L-j).$$

Lemma 4.2.2 *$\overline{V}_t(s)$ is quasi-concave in s.*

Proof The proof is again by induction on t. For $t = N$, we have $\overline{R}_N(s, L) = p_2L + p_1E[j \wedge (s - L)]$ which is clearly concave in L. Assume for $t+1$ that $\overline{V}_{t+1}(\cdot)$ is quasi-concave. $(s - L - j)^+$ is monotone decreasing in L which is quasi-concave. $\overline{V}_{t+1}[(s - L - 1)^+]$ is quasi-concave. Hence, $\overline{R}_t(s, L) = p_2 + p_1E(j \wedge (s - L)) + E\overline{V}_{t+1}[(s - L - j)^+]$ is quasi-concave. So is $\overline{V}_t(s)$. □

It is important to note that Lemma 4.2.2 holds without assumption (F). In this case, the optimal allocation for low fares is as follows :

$$\overline{L}_t(s) = \max\{L \le s : p_1 \sum_{j=s-L}^{\infty} q_j^1 + \sum_{j=0}^{s-L} q_j^1[\overline{V}_{t+1}(s - L - j) - \overline{V}_{t+1}(s - L - 1 - j)] \le p_2\}.$$

Finaly, we shall investigate the impact of uncertainty on high fare demands. Let E and $\overline{E}$ be the expectations with respect to the probabilities q_j^1 and $\overline{q}_j^1$,respectively. To emphasize the dependence on q_j, we write

$$V_t(s, q) = \max\{p_2L + p_1E[j \wedge (s - L)] + EV_{t+1}((s - L - j)^+, q)\}$$

and

$$V_t(s, \overline{q}) = \max\{p_2L + p_1\overline{E}[j \wedge (s - L)] + \overline{E}V_{t+1}((s - L - j)^+, \overline{q})\}.$$

Theorem 4.2.3 *If $E[j] = \overline{E}[j]$ and $\sum_{k=0}^{l}\sum_{j=0}^{k} q_j^1 \le \sum_{k=0}^{l}\sum_{j=0}^{k} \overline{q}_j^1$ for all l, then we have $V_t(s, q) \ge V_t(s, \overline{q})$ for all t and s.*

Proof Since $(j \wedge (s - L))$ is concave in j and $V_{t+1}(s, q)$ is also concave in s, we have from the second degree of stochastic dominace

$$E(j \wedge (s - L)) \ge \overline{E}(j \wedge (s - L))$$

and

$$\sum_{j=0}^{s-L} q_j^1 V_{t+1}(s - L - j, q) \ge \sum_{j=0}^{s-L} \overline{q}_j^1 V_{t+1}(s - L - j, \overline{q}).$$

Hence, we obtain

$$\begin{aligned} V_t(s, q) &= \max\{p_2L + p_1E[j \wedge (s - L)] + EV_{t+1}(s - L - j, q)\} \\ &\ge \max\{p_2L + p_1\overline{E}[(j \wedge (s - L)] + \overline{E}V_{t+1}(s - L - j, \overline{q})\} \\ &= V_t(s, \overline{q}). \end{aligned}$$

Remarks 2 *Since $E[j] = \overline{E}[j]$ and $\sum_{k=0}^{l}\sum_{j=0}^{k} q_j^1 \le \sum_{k=0}^{l}\sum_{j=0}^{k} \overline{q}_j^1$ implies that variance of the probability q_j^1 is smaller than of $\overline{q}_j^1$, Theorem 4.2.3 asserts that increasing the riskiness of demand distribution is sacrificial to the airline's expected revenue. As a result of this, the airline company turns out to set the higher fare price for increased uncertainty of the fare demand to protect the revenue.*

This section shows that there is a simple optimal seat allocation policy which is computationally feasible. In comparison to static models, our dynamic model takes the time periods remaining until departure and allows fare classes to reopen after closing.

Prior work on this problem falls into one of two categories and there are two different approaches to the problem. First, mathematical programming has been applied into those works together incorporated with network optimization. (See Ladany [58]).

Second, those works are based upon some restrictive assumptions (Rothstein [93], Littlewood [65], Belobaba [12], Curry [28], Brumelle et al. [21]). We follow this latter approach. We formulate the problem as a Markovian sequential decision problem with discrete time. The discretization is practical but makes it difficult to keep track of the booking process of passengers. Possible future research is to incorporate explicity the overbooking phenomenon into the model.

4.3 An Analysis of Airline Seat Allocation

In this section we simultaneously consider the seat allocation between high and low fares passengers and the overbooking problem, which also extends the existing literature in three ways. First, the cost of lost sales which has been ignored in the existing literature (see [2],[5],[8]) is explicitly incorporated into the model. Second, the overbooking phenomenon is also explicitly treated. Third, the concept of spill rate is clarified into the passenger spill rate and the flight spill rate. It is found that the results obtained here are in more close agreement with actual airline practice.

4.3.1 A Simple Seat Allocation Model

In this subsection, we consider a rather simple seat allocation model in which there are two class of passengers, low and high fares passengers. Assume that they do not both cancel their booking reservations. So, in this case the airline company does not have to overbook to hold out against the cancellations of their passengers' bookings. We make three assumptions as follows :

Assumption (A) The low fare and high fare demand are independent to each other.

Assumption (B) Demand for low fares occurs earlier than for high fares, e.g. low fare demand has a minimum advance booking requirement.

Assumption (C) The refused passengers do not pick up other flights of the same carrier (called the total loss of the spilled sales).

Assumption (B) is known as early birds. Assumption (C) excludes the possibility that the denied low fare passengers may then purchase a high fare ticket, which is called "grade up". Define X and Y the random variable of the number of high fare demand with distribution function $F(x)$ and Y the random variable of the number of high fare demand with distribution function $G(y)$, respectively. We use the following notations :

$p_1 =$ the high fare,

$p_2 =$ the low fare,

$C =$ the airplane capacity,

$\pi_1 =$ the cost of goodwill loss per high fare passenger due to the shortage of the capacity,

$L =$ the number of seats allocated to low fare passengers, so at least $(C - L)$ seats are available to high fare passengers,

$a \wedge b = \min(a, b), a^+ = \max(a, 0)$, and $E=$ expectation operator. All variables are treated as continuous.

Defining $ER(L)$ the expected total revenue per flight when L seats are allocated to the low fare demand, $ER(L)$ can be written as follows ;

$$ER(L) = E_Y[p_2(Y \wedge L)] + E_Y E_{X|Y}[p_1(X \wedge (C - Y \wedge L)) - \pi_1(X - C + Y \wedge L)^+] \quad (4.10)$$

which can be rewritten as

$$\begin{aligned} ER(L) &= p_2[\int_0^L y dG(y) + L\overline{G}(L)] \\ &+\overline{G}(L)\{p_1 \int_0^{G-L} x dF(x) + \int_{G-L}^{\infty} [p_1(C - L) - \pi_1(x - C + L)] dF(x)\} \quad (4.11) \end{aligned}$$

where $\overline{F}(\cdot) = 1 - F(\cdot)$ and $\overline{G}(\cdot) = 1 - G(\cdot)$. The problem is of chosing L to maximize $ER(L)$ subject to $0 \le L \le C$. Let f and g be density functions of X and Y, respectively. $f(\cdot)$ and $g(\cdot)$ are assumed to be positive on $(0, C)$.

Proposition 1 If $g(L)/\overline{G}(L) \le f(C - L)/\overline{F}(C - L)$ for all L, then an optimal number of seats to allocate to low fare demand L^* is given by

$$L^* = \begin{cases} 0, & \text{if } p_2 < (p_1 + \pi_1)\overline{F}(C), \\ C - \overline{F}^{-1}(\frac{p_2}{p_1+\pi_1}), & \text{if } (p_1 + \pi_1)\overline{F}(C) \le p_2 \le (p_1 + \pi_1)\overline{F}(0), \\ C, & \text{otherwise.} \end{cases} \quad (4.12)$$

Proof $ER(L)$ is differentiable. We shall show that $ER(L)$ is strictly concave with respect to L under the condition that $g(L)/\overline{G}(L) \le f(C - L)/\overline{F}(C - L)$. Hence, if it possensses a maximum, it is a unique one. Differentiating $ER(L)$ results in

$$\frac{dER(L)}{dL} = \overline{G}(L)[p_2 - (p_1 + \pi_1)\overline{F}(C - L)]. \quad (4.13)$$

Differentiating $ER(L)$ twice, we have

$$\frac{d^2ER(L)}{dL^2} = -g(L)p_2 + (p_1 + \pi_1)[g(L)\overline{F}(C - L) - f(C - L)\overline{G}(L)] < 0.$$

Note that $\overline{F}(C - L)$ is monoton increasing in L and then the inverse of $\overline{F}(\cdot)$ exists. Since $g(\cdot) > 0$ implies $\overline{G}(\cdot) > 0$, letting $dER(L)/dL = 0$ with condition $0 \le L \le C$ yields equation (4.12).

Remarks 3 *Note that the optimal allocation to low fare demand* L^* *is independent of the distribution of low fare demand, provided* $\overline{G}(\cdot) > 0$. L^* *is decreasing as* π_1 *increases, and depends only on the relative low fare* $p_2/(p_1 + \pi_1)$.

It is difficult to provide an economic interpretation on the condition for Proposition 1, $g(L)/\overline{G}(L) \le f(C - L)/\overline{F}(C - L)$, which looks like a failure rate often appearing in reliability theory. Insted of doing that, we consider a special case where there is strong demand from low fare passengers, that is, $\overline{G}(C) = 1$. For instance, the peak season may have the demand distribution satisfying $\overline{G}(C) = 1$.

Assume that we can sell as many low fare seats as we desire up to the capacity C of the plane. Let $R(L, X)$ be the revenue obtained if L seats are to low fare demand and

$C-L$ to high fare demand and a demand for high fare X is realized. Then, for $0 \le L \le C$ we have

$$R(L, X) \;=\; p_2 L + p_1(C-L) \wedge X - \pi_1(X - C + L)^+. \qquad (4.14)$$

Note that for each $X, R(L, X)$ is concave in L without any condition. So, the expected revenue $ER(L) \equiv E[R(L, X)]$ is also concave in L. Because of the concavity, the optimal number of low fares L^* to maximize the expected revenue can be determined by looking at the incremental expected revenue from selling an additional low fare and stop selling when this becomes negative. The following corollary immediately follows from the fact that the assumption of Proposition 1 is satisfied if $\overline{G}(C) = 1$ for all $C > 0$.

Corollary If there is an unlimited low fare demand, one should sell only L^* low fares and protect $C - L^*$ seats for high fares, where L^* is given by

$$L^* = \min\{L \ge 0 \;:\; \frac{p_2}{p_1 + \pi_1} \le \overline{F}(C - L)\}.$$

4.3.2 Optimal Seat Allocation with Overbooking

In this subsection we treat with an optimal seat allocation model allowing overbooking and cancellations. Assume that high fare passengers may cancel their reservations but low fare ones can not. Hence, only high fare demands are overbooked. We must determine both the maximum level to accept the reservations from low fare passengers and the overbooking ratio for high fare passengers. A sequence of operations is as follows ; (i) choose the number of low fare seats to reserve, (ii) observe the number of realized booking progress of low fare passengers, and them (iii) determine the number of high fare seats to protect.

In addition to the notations listed in Subsection 4.3.1, we use the following:

Z = the number of cancellations of high fare reservations, $Z \le X$,

π_2 = the cost per denied boarding due to overbooking,

$K(y)$ = the number of seats allowed to book for high fare passengers when $Y = y$,

which is assumed to satisfy $K(y) = (C - y \wedge L)(1 + \alpha)$, $0 \le \alpha \le 1$.

Remembering L as the seat allocation for low fares, a pair of (L, α) is called a booking strategy and set $B = (L, \alpha)$. We may possibly have $L + K(y) \ge C$ but $L \le C$.

Remarks 4

(i) If high fare passengers do not cancel their reservations, so that the airline does not necessarily overbook, then the booking strategy can be written as $B = (L, 0)$, which is reduced to $K(y) = C - y \wedge L$ as same as in Subsection 4.3.1.

(ii) α can be interpreted as an overbooking ratio for high fare passengers because

$$\frac{y \wedge L + (C - y \wedge L)(1+\alpha) - C}{C - y \wedge L} = \frac{\alpha(C - y \wedge L)}{C - y \wedge L} = \alpha$$

So, $\alpha(C - y \wedge L)$ seats are overbooked for high fares. $\alpha = 0$ corresponds to no overbooking with which cases are treated in Subsection 4.3.1.

Assumption(D) High fare passengers may cancel their reservations independently with the equal probability $(1 - \theta)$.

It is well known under assumption (D) that the probability distribution of the number of cancellations Z, given the number of reservations x is binomial with the mean $x(1-\theta)$, that is, $H(z \mid x) = P\{Z \le z \mid X = x\} = \sum_{k=0}^{z}\binom{x}{k}(1-\theta)^k\theta^{x-k}$, $z = 0, 1, \cdots, x$. So, a passenger may board on with probability θ. Put $\beta \equiv (1 + \alpha)$ and $B = (L, \beta)$ in stead of (L, α). Define $ER(L, \beta)$ the expected profit obtained from a flight when a booking strategy $B = (L, \beta)$ is used. Then, we obtain

$$\begin{aligned}
ER(L,\beta) \;=\;& p_1 E_Y E_{X|Y} E_{Z|X}[X \wedge ((C-Y) \wedge L)\beta - Z] + p_2 E[Y \wedge L] \qquad (4.15)\\
& - \pi_1 E_Y E_{X|Y}[X - ((C-Y) \wedge L)\beta]^+\\
& - (p_1 + \pi_2) E_Y E_{X|Y} E_{Z|X}[Y \wedge L + X \wedge ((C-Y) \wedge L)\beta) - Z - C]^+\\
=\;& p_1\{E[X - Z \mid X \le (C-Y)\beta]P[X \le (C-Y)\beta]\\
& + E[(C-Y)\beta - Z \mid X > (C-Y)\beta]P[X > (C-Y)\beta]\}P[Y \le L]\\
& + p_1\{E[X - Z \mid X \le (C-L)\beta]P[X \le (C-L)\beta]\\
& + E[(C-L)\beta - Z \mid X > (C-L)\beta]P[X > (C-L)\beta]\}P[Y > L]\\
& + p_2\{E[Y \mid Y \le L]P[Y \le L] + E[L \mid Y > L]P[Y > L]\}
\end{aligned}$$

$$
\begin{aligned}
& -\pi_1\{E[X-(C-Y)\beta \mid Y \leq L]^+ P[Y \leq L] \\
& +E[X-(C-L)\beta \mid Y > L]^+ P[Y > L]\} \\
& -(p_1+\pi_2)\{E[Y+X-Z-C \mid X \leq (C-Y)\beta]P[X \leq (C-Y)\beta] \\
& +E[Y+(C-Y)\beta - Z - C \mid X > (C-Y)\beta]P[X > (C-Y)\beta]\}P[Y \leq L] \\
& -(p_1+\pi_2)\{E[L+X-Z-C \mid X \leq (C-L)\beta]P[X \leq (C-L)\beta] \\
& +E[L+(C-L)\beta - Z - C \mid X > (C-L)\beta]P[X > (C-L)\beta]\}P[Y > L] \\
= \; & p_1\theta\{\overline{G}(L)[\int_0^{(G-L)\beta} x dF(x) + \overline{F}((C-L)\beta)(C-L)\beta] \\
& +\int_0^L [\int_0^{(C-y)\beta} x dF(x) + \overline{F}((C-y)\beta)(C-y)\beta] dG(y)\} \\
& +p_2 \int_0^L y dG(y) + p_2 L\overline{G}(L) \\
& -\pi_1\{\overline{G}(L)\int_{(C-L)\beta}^{\infty}(x-(C-L)\beta)dF(x) \\
& +\int_0^L \int_{(C-y)\beta}^{\infty}(x-(C-y)\beta)dF(x)dG(y)\} \\
& -(p_1+\pi_2)\{\overline{G}(L)[\int_{C-L}^{(C-L)\beta}(L+\theta_x - C)dF(x) \\
& \qquad\qquad +\overline{F}((C-L)\beta)(L-C)(1-\beta\theta)] \\
& +\int_0^L [\int_{C-y}^{(C-y)\beta}(y+\theta_x-C)dF(x) + \overline{F}((C-y)\beta)(y-C)(1-\beta\theta)]dG(x)\}.
\end{aligned}
$$

The first two terms are the revenues for the high and low fares, respectively. The third one is the cost of lost sales due to the booking limit. The last one is the cost of denied booking which occurs whenever the sum of the numbers of low fares and of confirmed high fares is larger than the capacity of the plane. Since the number of cancelled booking is binomially distributed, $E[X-Z] = E_Z E_X[X - E(Z \mid X) \mid Z] = E[X - X(1-\theta) = \theta E[X]]$. The problem is to find an optimal booking strategy (L^*, β^*) so as to maximize $ER(L, \beta)$. After taking and rearranging the partial derivatives with respect to L and β, we have*

$$
\begin{aligned}
p_2 - \pi_2 H * F((C-L^*)\beta^*) &= \overline{F}((C-L^*)\beta^*)[\beta^*(p_1\theta + \pi_1) \\
& \quad +\alpha\pi_2 H(\alpha(C-L) \mid (C-L)\beta)],
\end{aligned} \tag{4.16}
$$

$$
\overline{G}(L^*)\overline{F}((C-L^*)\beta^*)Q(L^*,\beta^*) = \int_0^{L^*} \overline{F}((C-y)\beta^*)Q(y,\beta^*)dG(y), \tag{4.17}
$$

where

$$
Q(L,\beta) = (C-L)[p_1\theta + \pi_1 - \pi_2 H(\alpha(C-L) \mid (C-L)\beta)], \tag{4.18}
$$

$$
Q(y,\beta) = (C-y)[p_1\theta + \pi_1 - \pi_2 H(\alpha(C-y) \mid (C-y)\beta)], \tag{4.19}
$$

and

$$
H * F(u) = \int_0^u H(L + x - C \mid x)dF(x). \tag{4.20}
$$

An optimal booking strategy, seat allocation for low fares and overbooking ratio for high fares, must jointly satisfy equations (4.16) and (4.17). Note that such an optimal booking strategy is no longer independent of $G(\cdot)$, the probability distribution of low fare demands. Equations (4.18) and (4.19) are the net profits obtained from high fare passengers when $Z \leq (C - y \wedge L)\beta$. Equation (4.17) can, therefore, be interpreted as follows ; under the optimal booking strategy the expected profit from high fare demands when $Y \geq L$ should be equal to the one when $Y < L$. It seems to us that finding a closed form of an optimal booking strategy jointly satisfying equations (4.16) and (4.17) is almost impossible. For this reason we consider a special case of overbooking problems where there is only one class of fares, say high fares.

A special case of overbooking problems

Suppose that there is only one class of fares, say a high fare class. With each fare booked we associate a random variable

$$
D_i = \begin{cases} 1 & \text{if the } i\text{-th fare booked confirms,} \\ 0 & \text{if the } i\text{-th fare cancels.} \end{cases}
$$

Assume that $\{D_1, D_2, \cdots\}$ are independent and identically distributed with mean $ED_i = \theta$. If B seats are booked, then $N(B) = \sum_{i=1}^{B} D_i$ seats are confirmed and the distribution of $N(B)$ is binomial with mean $B\theta$. The revenue as a function of the number of seats available, L, and the number of fares booked, B, is assumed to be

$$
R(L,B) = p_1 N(B) - (p_1+\pi_2)[N(B)-L]^+ - \pi_1(X-B)^+.
$$

Since $E[N(B)] = \theta B$, the expected revenue function is

$$
ER(L,B) = p_1\theta B - (p_1+\pi_2)E[N(B)-L]^+ - \pi_1 E[X-B]^+.
$$

We first compute the incremental revenue from an additional booking

$$
\Delta R(L,B) = R(L,B+1) - R(L,B)
$$

$$= p_1[N(B+1) - N(B)] - (p_1 + \pi_1)[(N(B+1) - L)^+ - (N(B) - L^+)]$$
$$-\pi_1[(X - (B+1))^+ - (X - B)^+]$$
$$= p_1 D_{B+1} - (p_1 + \pi_1)Z_B + \pi_1 Z'_B$$

where

$$Z_B = \begin{cases} D_{B+1} & \text{if } N(B) > L, \\ 0 & \text{if } N(B) \leq L, \end{cases}$$

and

$$Z'_B = \begin{cases} 1 & \text{if } X > B, \\ 0 & \text{if } X \leq B. \end{cases}$$

Now, the incremental expected revenue can be obtained.

$$\begin{aligned} \Delta ER(L,B) &\equiv ER(L,B+1) - ER(L,B) = E\Delta R(L,B) \\ &= p_1 E(D_{B+1}) - (p_1 + \pi_1)EZ_B - \pi_1 EZ'_B \\ &= p_1\theta - (p_1 + \pi_1)\theta P\{N(B) > L\} + P\{X > B\}. \end{aligned}$$

Since $N(B)$ is increasing in B, so is $P\{N(B) > L\}$. $P\{X > B\}$ is also decreasing in B. Therefore, $\Delta ER(L,B)$ is decreasing in B, which implies that $ER(L,B)$ is concave in B. So, we should book fares so long as $ER(L,B+1) - ER(L,B)$ is positive. Hence, we arrive at the following theorem.

Theorem 4.3.1 *If we have L seats protected for high fares, it is optimal to book up to $B^*(L)$ where*

$$B^*(L) = \min\{B \geq L_i \; ; \; p_1\theta \leq (p_1 + \pi_2)\theta P\{N(B) > L\} + \pi_1\overline{F}(B)\}. \quad (4.21)$$

Since $P\{N(B) > L\}$ can be evaluated from a table of binomial distributions and $\overline{F}(B)$ is given, $B^*(L)$ can easily be calculated. Note that $P\{N(B) > L\}$ and $\overline{F}(B)$ in equation (4.21) is strictly decreasing in B. So, there exists a unique solution satisfying (4.21).

4.3.3 Spill Rates and Overbooking

If passengers may cancel their reservations without any penalties, airline companies tend to overbook. However, such overbookings cause them compensation costs. If a small

number of seats are allocated to each fare class to prevent them from overbooking, they also lose refused boarding passengers resulting in a cost of goodwill lost, which is called spilling passengers of the airline. Hence, a booking strategy must be balanced between overbooking and spilling passengers. In this subsection we discuss the concepts of spill rates and the expected number of overbookings when you must obtain a certain number of " confirmed " seats, say the number of seats equal to the airplane capacity.

There are two possible interpretations of the term " spill rate " in the airline context. The first is that the spill rate is the expected proportion of flights on which some high fare reservations must be refused because of prior low fare bookings, which is often used in " Airline Yield Management " articles. (For Example, see [6], [62].) The second is that the spill rate is the expected proportion of high fare reservations that must be refused out of the total number of such reservations, which seems to be more meaningful since it relates more closely to the amount of high fare revenues lost. We provide formulae for calculating the spill rate under either interpretation and consider such spill rates associated with use of the revenue maximizing seat allocation rule.

The proportion of *flights* refusing high-fare reservations, called the *flight spill rate*, can be expressed as

$$\begin{aligned} R_1 &= P\{X + L \wedge Y > C\} \quad (4.22) \\ &= \overline{G}(L)\overline{F}(C-L) + \int_0^L \overline{F}(C-y)dG(y). \end{aligned}$$

It is easy to show that when S seats are available for high-fare passengers the expected proportion of reservations refused will be $\overline{F}(S)E[X - S \mid X > S]/E[X]$. Thus we have for the expected proportion of high fare reservations refused R_2 called the *passenger spill rate* ;

$$R_2 = \frac{1}{E[X]}\{\overline{G}(L)\overline{F}(S)u_{C-L} + \int_0^{C-S} \overline{F}(C-y)u_{C-y}dG(y)\} \quad (4.23)$$

where

$$u_s = [(1/\overline{F}(S))\int_S^\infty x dF(x)] - S.$$

Consider the simple seat allocation model discussed in Subsection 4.3.1 where the revenue maximizing rule is used to determine L^*, the number of low fare seats to protect. Let S^* be the number of high fare seats to allocate. In this case we have $S^* = C - L^* =$

$\overline{F}^{-1}(p_2/(p_1+\pi_1))$ from equation (4.12). $\overline{F}(S^*) = p_2/(p_1+\pi_1)$. In the extreme case that low fare demand always exceeds the allocation of low fare seats, we have $\overline{G}(L^*) = 1$, and the above formulae become :

$$R_1 = p_2/(p_1+\pi_1), \quad and$$

$$R_2 = [p_2/(p_1+\pi_1)](u_{S^*}/E[X]).$$

Now, further assume that $p_2/(p_1+\pi_1) = 0.4$ (a typical ratio) and that high fare demand is approximately normally distributed with mean 100 seats and standard deviation 20 seats. This will yield a spill rate (expressed as a percentage) of 40 %, if the first interpretation (R_1) is used. It seems to us that this figure is abnormally too high. However, if the second interpretation is used, we get $u_{S^*} \approx 32$ and then $R_2 \approx (0.4)(32)/100{=}0.128$. A spill rate of 12.8 % seems in closer agreement with actual airline practice.

Given the probability distribution for high fare demands, the problem of determining the seat allocation for each fare class is also of determining either spill rate. The more seats are protected, the smaller spill rate we have. However, the more bookings we accept, the higher probability of overbookings we have, while at that time the spill rate is close to zero. So, the seat allocation problem is of trade off between a loss spill rate and a high overbooking ratio. Let us consider the following probability. What is the probability that we must accept a certain number of bookings, say B, in order to obtain the number of confirmed seats which is naturally equal to the capacity of airline seats $C, B \geq C$. Let $q(B;C)$ be such probabilities. It is easy to see that such a random variable follows a negative binomial distribution, that is,

$$q(B;C) \;=\; \begin{pmatrix} B-1 \\ C-1 \end{pmatrix} \theta^C(1-\theta)^{B-C}. \tag{4.24}$$

with mean C/θ and variance $C(1-\theta)/\theta^2$. Note that variance rapidly increases as $\theta \to 0$. This suggests that airline companies should make effort of reducing the cancellation probability. For example, the restriction on tickets or on booking procedures must be imposed. Equation (4.24) can be expressed in terms of binomial distribution $b(n, c)$, that is,

$$q(B;C) = \theta b(B-1, C-1).$$

which can be evaluated from a table of binomial distributions.

4.4 Allocation of Airline Seats between Stochastically Dependent Demands

This section deals with the problem of setting a limit for bookings of airline seats in a " discount " fare class when there is stochastic dependency between demands for the discount seats and demands for " full fare " seats. Specifically, the problem is examined as follows;

> *Given probability distributions of forecasted demand for discount and full fare passengers for a given leg of a future flight, determine the stopping rule for discount booking that maximizes the expected total flight revenue.*

The main accomplishment of the present section is the introduction of a seat allocation model that allows for demand dependency between fare classes. Also presented here are an extension of the model to allow for control of the full fare passenger *spillage* (the rejection of full fare reservation requests when a flight is fully booked), consideration of the impact on passenger *goodwill* of seat allocation decisions, and a rigorous proof of a formula for optimal seat allocations in the special case that dependency arises because of upgrades.

A useful approach to the seat allocation problem was suggested in 1972 by Littlewood [65]. He proposed that an airline should continue to sell discount seats as long as the following condition held :

$$\rho_B \;\geq\; \rho_Y P[Y > \kappa - \eta], \tag{4.25}$$

where ρ_B is average revenue from discount passengers, $P[\cdot]$ denotes probability, Y is full fare demand, κ is the number of seats available for the two fare classes, and η is the number of discount seats sold. The intuition here is cleat – sell an additional discount seat as long as the discount revenue equals or exceeds the *expected* full fare revenue from the seat.

A second interpretation of (4.25) will prove useful in the sequel. If discount fare demand reaches the limit η on every flight, the probability $P[Y > \kappa - \eta]$ is the expected proportion of flights on which full fare demands is turned away, or *spilled*. The *actual* proportion of flights on which such spillage occurs is termed the *flight spill rate*, thus the probability above represents the highest possible flight spill rate given the distribution of

Y demand, or *maximum flight spill rate* (maximum because discount bookings might not reach the limit η on every flight). When discount and full fare demands are independent, the maximum spill rate will increase as η increases. Littlewood's rule specifies that the optimal booking limit is the largest value of η for which this maximal rate does not exceed the ratio of discount to full fare. With discount fares in the range of 30 % to 60 % of full fares, the rule implies turning away one or more full fare booking requests on 30 % to 60 % of all flights when discount demands are high – proportions that seem higher than most airline managers would accept.

A continuous version of Littlewood's rule was derived by Bhata and Parekh [14] in 1973. Ritcher [86] in 1982 gave a marginal analysis which proved that (4.25) gives an optimal allocation. Note of this early work allowed for the possibility of dependencies between classes of demand.

More recently, Belobaba [13] proposed a generalization of the marginal analysis approach to more than two fare classes. In the same work [pp. 143-150], Belobaba discussed the possible impact of demand dependencies on booking limits and showed with numerical examples that, in a three fare class problem, the booking limit for the lowest fare class will be reduced as the correlation between demands for the two upper fare classes increases. He did not attempt to directly analyze that problem nor to examine the problem of determining the booking limit *between* two dependent classes (the problem examined here). Belobaba also proposed a seat allocation formula for the case that demand dependency arises because of upgrades.

Before proceeding with a detailed analysis of the dependent demand case, we offer the following brief intuitive argument. The case considered fare is much the same as that considered in deriving Littlewood's rule except that, here, the full fare demand distribution must be modified as each discount demand occurs because of the dependency between the demands. That is, after observing $B \geq \eta$ the full fare demand distribution becomes $P[Y > \kappa - \eta \mid B \geq \eta]$. It seems reasonable to conjecture that the optimal booking limit will be obtained simply by replacing the probability in Littlewood's formula with this conjecture is valid as long as the discount and full fare demands are *monotonically associated* – a condition that will be precisely defined later.

The section is organized as follows. Subection 4.4.1 describes a general seat allocation model that will form the basis for later analyses. In Subsection 4.4.2, the allocation model is used to derive the above generalization of Littlewood's rule for the dependent demand case.

Notation

Our notation will follow the convention that capital letters are random variables or functions, and Greek letters are paramenters in the model. We will use the notation $a \vee b = \max(a, b)$, $a \wedge b = \min(a, b)$, and $a^+ = a \vee 0$. The indicator operator I is defined for logical propositions as 1 if the proposition is true and 0 otherwise. For example,

$$I_{[\alpha<\beta]} = \begin{cases} 1, & \text{if } \alpha < \beta, \\ 0, & \text{if } \alpha \geq \beta. \end{cases} \tag{4.26}$$

The operator E denotes stochastic expectation.

4.4.1 A Seat Allocation Model

This subsection presents a general model for the seat allocation problem that will serve as a basis for the specific analyses of later subsections. It is similar in structure to *optimal stopping* models described in Chow, Robbins and Siegmund [25], and Derman and Sacks [34] and this correspondence will be used to characterize instances of the problem for which a simple rule yield an optimal solution – the so-called monotone class of problems.

As discussed in the introduction, the following restrictions are placed on the demand and decision processes. There is an initial demand for discount seats which is followed by demand for full fare seats. Once the decision is made to stop satisfying requests for discount seats, no further requests will be accepted.

It is assumed that the decision to close discount sales is made knowing only the number of requests currently accepted. The decision cannot ,for example, involve any observations of the times at which the demands occur or the arrival rate of demands. Also, the decision cannot use any observations of full fare sales.

Note that one way of relaxing the static nature of this restricted model would be to simply re-apply the model as demand forecasts are updated or other changes occur. Empirical work by Mayer [68] showed that such a heuristic application of Littlewood's rule results in only slight losses in revenues relative to more complex dynamic optimization methods. Also, simulation work by Titze and Greisshaber [121] shows that, in practice, the strict booking sequence assumption can be relaxed somewhat.

Let ρ_B and ρ_Y denote the average revenues per discount and full fare booking, respectively. The revenue resulting from booking $(B \wedge \eta)$ discount passengers will be $\rho_B(B \wedge \eta)$. Define $F(\eta)$ to be the seats remaining after discount bookings are closed, so that $F(\eta) = \kappa - (\eta \wedge B)$. It is assumed that there is now an additional demand for a total of $Y(\eta)$ full fare demand might depend on the decision variable η, as is the case when a proportion of customers denied discount bookings elect to *upgrade* to full fare bookings. Moreover, it is not assumed that B and $Y(\eta)$ are independent.

By accepting as much of the demand $Y(\eta)$ as possible, an additional revenue of $\rho_Y(Y \wedge F(\eta))$ will be generated. In the case that a goodwill cost or penalty is incurred for turning away full fare demands, the unsatisfied portion of this demand will incur a total cost of $\rho_G(Y - F(\eta))^+$; where ρ_G is the goodwill cost per rejected full fare passenger.

Combining the above revenues and costs gives the net revenue function

$$\begin{aligned} R(\eta) &= \rho_B(B \wedge \eta) + \rho_Y(Y(\eta) \wedge F(\eta)) \qquad (4.27) \\ &\quad -\rho_G(Y(\eta) - F(\eta))^+, \end{aligned}$$

whose expectation is to be maximized as a function of η. Since it is not possible to allocate more than the available capacity, $R(\eta)$ is only defined for η such that

$$0 \le \eta \le \kappa. \qquad (4.28)$$

Now suppose $\eta - 1$ requests have been accepted from the discount demand, and an additional discount request is received ; (i.e., $B \ge \eta$). If bookings stop at $\eta - 1$, then the expected revenue is $E[R(\eta - 1) \mid B \ge \eta]$. If the additional request is accepted, expected revenue will be $E[R(\eta) \mid B \ge \eta]$. It is useful to write the expected incremental gain, $G(\eta)$, of accepting an additional request. It follows from (4.27) that

$$\begin{aligned} G(\eta) &= E[R(\eta) \mid B \ge \eta] - E[R(\eta - 1) \mid B \ge \eta] \qquad (4.29) \\ &= \rho_B + \rho_Y E[Y(\eta) \wedge (\kappa - \eta) - Y(\eta - 1) \wedge (\kappa - \eta + 1) \mid B \ge \eta] \\ &\quad +\rho_G E[(Y(\eta) - (\kappa - \eta))^+ - (Y(\eta - 1) - (\kappa - \eta + 1))^+ \mid B \ge \eta], \end{aligned}$$

provided $P[B \ge \eta] > 0$ so that the conditional expectations are defined. If B is less than η then the decision to accept or reject the ηth request can never arise and $G(\eta)$ is not defined. The domain of G is also limited to that of $R(\eta)$, as specified by condition (4.28), above.

The gain function $G(\eta)$ is just the first difference of the expected revenue function. Clearly, a booking limit of η will be preferred to $\eta - 1$ whenever $G(\eta)$ is positive. Furthermore, if $G(\eta)$ is nonnegative for all η up to some η^*, and nonpositive thereafter, then η^* will be optimal.

The seat allocation problem belongs to a class of stochastic optimization problems known as optimal stopping problems. Within that class of problems, those defined as *monotone* have particularly simple solutions. The seat allocation problem is monotone if the following conditions are satisfied :

1. There is some η^* such that the gain $G(\eta)$ is nonnegative (and defied) for $\eta \le \eta^*$ and nonpositive (or not defined) for $\eta > \eta^*$; and
2. $\mid Y(\eta) - Y(\eta - 1) \mid$ is bounded in η.

If the model is monotone, then the expected revenue will be maximized by accepting up to η^* requests from the B demand; that is, by *protecting* $\kappa - \eta^*$ seats for full fare customers. The expected revenue is maximized in the sense that no policy which only uses the information obtained by observing $I_{[B \ge \eta]}$ can do better.

The following subsections will consider applications of the above model to specific allocation problems which are monotone.

4.4.2 Specific Seat Allocation Problems

This subsection specializes the above model to three variants of the seat allocation problem with dependent demands. In the first variant, it is assumed that there are no penalties for refused bookings and that there are no penalties for refused bookings and that full fare demand is not influenced by the discount booking level η. The second considers the loss of goodwill associated with full fare passenger spillage by introducing a penalty for refused bookings. The third deals with the upgrades case in which ultimate full fare demand is influenced by the discount booking level.

A Simple Seat Allocation Model with Dependent Demand

The model analyzed here is the usual seat allocation model except that the demands of the two fare classes, Y and B are allowed to be stochastically dependent.

With reference to the general revenue model (4.27), assume no penalty costs, so $\rho_G = 0$, and assume that full fare demand is not influenced by the booking limit assigned to discount fares, so that $Y(\eta) = Y$, for $\eta = 1, ..., \kappa$. Note that since demand is an integer, $Y > \kappa - \eta$ is the same as $Y \geq \kappa - \eta + 1$.

Using these properties, the gain associated with increasing the discount booking limit from $\eta - 1$ to η, given by (4.29), can be simplified to

$$\begin{aligned} G(\eta) &= \rho_B + \rho_Y E[(Y \wedge (\kappa - \eta)) \\ &\quad -(Y \wedge (\kappa - \eta + 1)) \mid Y > \kappa - \eta, B \geq \eta] \cdot P[Y > \kappa - \eta \mid B \geq \eta] \\ &= \rho_B - \rho_Y P[Y > \kappa - \eta \mid B \geq \eta]. \end{aligned} \tag{4.30}$$

This expression has a familiar interpretation : when an additional seat is sold to a discount customer, there will be a certain gain of one discount fare, and if full fare demand exceeds the new lower protection level, there will be a loss of one full fare.

The expected gain is positive whenever $G(\eta) > 0$, or equivalently whenever

$$P[Y > \kappa - \eta \mid B \geq \eta] < \frac{\rho_B}{\rho_Y}. \tag{4.31}$$

If it is the case that

$$P[Y > \kappa - \eta \mid B \geq \eta] \quad \text{is nondecreasing in } \eta, \tag{4.32}$$

then $G(\eta)$ is nonincreasing in η, and the problem is monotone.

Henceforth, property (4.32) will be referred to as the *monotonic association* property. Loosely speaking, this property specifies that as the discount booking limit increases, the full fare spill rate tends to increase.

A suitable η to satisfy the definition of an optimal solution in a monotone problem is

$$\begin{aligned} \eta^* &= \max\{\eta \ : \ G(\eta) > 0\} \\ &= \max\left\{0 \leq \eta \leq \kappa \ : \ P[Y > \kappa - \eta \mid B \geq \eta] < \frac{\rho_B}{\rho_Y}\right\}, \end{aligned} \tag{4.33}$$

where we will adopt the convention that $\eta^* = 0$ if $P[Y > \kappa] \geq \rho_B/\rho_Y$ so that the maximum is over the empty set. (Recall that the domain of G consists of those η between 0 and κ such that $P[B \geq \eta] > 0$.) It is thus optimal to sell at most η^* seats to customers requesting discount fares.

Note that the probability $P[Y > \kappa - \eta \mid B \geq \eta]$ can be interpreted as the *maximal flight spill rate* as was the analogous term in Littlewood's rule (4.25). But then (4.33) is just a generalization of the fact that this rate should be just less than the discount/full fare ratio.

If the demands are independent, then (4.32) clearly holds, and the optimality condition becomes

$$\eta^* = \max\left\{0 \leq \eta \leq \kappa \ : \ P[Y > \kappa - \eta] < \frac{\rho_B}{\rho_Y}\right\}. \tag{4.34}$$

In this case there is not a unique η^* which is optimal. The η^* defined by (4.33) is the smallest. The largest optimal discount booking limit is obtained by permitting equality in (4.33) :

$$\eta^{**} = \max\left\{0 \leq \eta \leq \kappa \ : \ P[Y > \kappa - \eta \mid B \geq \eta] < \frac{\rho_B}{\rho_Y}\right\}. \tag{4.35}$$

This is just Littlewood's rule (1) except that now dependency between discount and full fare demands is allowed, subject to the monotonic association property (4.32). The following subsection illustrates the effect of such dependency.

4.4.2.1 Example : Seat Allocation with Dependent Demands

Table 4.1 presents an example of optimal discount seat booking limits for a range of cabin capacities and for both independent and dependent demands. For this example, the discount fare was fixed at 60 % of the full fare, and discrete approximations to bivariate normal distributions were used to model the discount/full joint probability functions. (Recall that monotonic association condition is satisfied by positively correlated bivariate normal random variables.) The mean combined demand was 100 seats in all calculations. In the dependent demand case, a high correlation ($\rho = 0.9$ between discount and full fare demands was used in order to obtain approximate upper bounds on the revenue increases that result from taking dependency into account.

Table 4.1: Effect of Demand Dependency on Discount Seat Booking Limits

Booking	Cabin Capacity					
	46	60	80	100	120	140
Discount booking limit: $\eta(\rho = 0)^a$	19	33	53	73	93	113
Full fare protection: $\kappa - \eta$	27	27	27	27	27	27
Discount booking limit: $(\rho = 0.9)^b$	19	32	49	65	81	97
Full fare protection:	27	28	31	35	39	43
Percent revenue increase[c]	0	0.08	0.54	1.25	1.27	0.71

[a] Independent: correlation = 0. For all calculations, mean demands were 70 discount and 30 full, and standard deviations were nominally 26.5 discount and 11.5 full. The standard deviations varied slightly between cases because of the discretization procedure.

[b] Dependent: correlation = 0.9.

[c] Revenue increase achieved by allowing for dependency.

With reference to Table 4.1, note that in the independent case, the optimal discount booking limits correspond to a fixed *protection* level of 27 seats for full fare passengers at all cabin capacities. The discount booking limits are increased as capacity increases in order to keep the maximum flight spill rate for full fares in balance with the discount/full fare ratio, as discussed earlier. Since the mean demands are being held constant for all cabin capacities, it appears that increased capacity is being allocated exclusively to discount demands. Recall, however, that unsold discount seats can be sold to full fare

passengers. In the capacity = 140 case, for example, the majority of flights will have discount demands of less than 113 seats, and full fare seating capacity will be accordingly larger than 27 seats most of the time. It is only when discount demands reach 113 seats that the marginal revenue considerations expressed by equation (4.34) dictate closing down discount sales.

In this example, the optimal full fare protection level increases with capacity when discount and full fare demands are dependent (the $\rho = 0.9$ case). The same spill rate balancing considerations are acting here ; however, because of the positive correlation between demands, the discount booking limits are not increased as much as in the independent case. (The information that discount demand has exceeded some value should imply an increased probability of higher full fare demand and should lead to higher protection levels for full fare seats.) In fact, it has been shown by Mcgill [69] that, with bivariate normal demand distributions, the optimal discount booking limit always decreases as correlation increases.

With small cabin capacities relative to demand (capacities of 46 seats or fewer), the booking limits in the dependent case are the same as those in the independent case, as there is no revenue benefit from taking dependency into account. This i because with small capacities the discount demand is almost certain to exceed the discount seat booking limit, and so $P[Y > \kappa - \eta \mid B \geq \eta] \cong P[Y > \kappa - \eta]$. For large capacities relative to demand (somewhat greater than 140 seats), the optimal discount booking limit in the independent case will be substantially lower than that in the dependent case ; however, the corresponding revenue benefits will be negligible as there is ample space for both fare classes under most realizations of the demand process.

Implementation

The optimal booking rule for the dependent demand case (4.33) is simple to implement as a planning tool if some joint distribution such as the bivariate normal is assumed to hold for the demands. In this case it is straightforward to calculate the conditional distribution $P[Y > \kappa - \eta \mid B \geq \eta]$ for enough values of η to solve the optimality condition. It is then possible to study the impact of hypothesized shifts in the demand distribution or in other parameters in much the same way as in the example above.

Implementation of the dependent demand booking rule as a control tool in a reservations system is also possible, but less straightforward. Estimation of the conditional demand distributions can be done, as above, by using a joint demand distribution. In this case, however the parameters of the distribution must be obtained by fitting to historical data and, possibly, by adjusting for anticipated market conditions. This fitting process is not straightforward since 1) demand data from a history of previous flights will be censored whenever demand reaches a booking limit or the capacity of the aircraft, and 2) the parameters of the demand distribution depend on external factors like fares, competition and time fo year. The same problems with the estimation problem in the dependent demand case.

The spill rate interpretation of the optimal allocation rules suggests a second, simpler, application. Over a series of flights for which the underlying demand distributions are considered stable (e.g., within one season, mid-week flights) the optimal allocation rule in either the independent or dependent demand cases specifies that the *maximal spill rate* should be as close as possible to, without exceeding, the discount/full fare ratio. An *observed* proportion that is too high would indicate that the discount booking limits have generally been too high. Similarly, a proportion that is too low would indicate that booking limits have been too low. This approach has two significant advantages. First, there is no requirement for modeling the demand distribution, and second, there is little computational difference between the independent and dependent demnad cases. To see the second point note that the observed maximal spill rate in the independent case will simply be the proportion of flights on which full fare demand exceeded the protection level $\kappa - \eta$. In the dependent case, it will be the proportion of those flights on which the discount booking limit was also exceeded. This technique does not provide a practical way of controlling bookings on individual flights since airlines perform such control based on individual forecasts of demand and other factors ; however, it does provide a simple way of monitoring past performance relative to theoretically optimal booking limits.

4.4.3 Full Fare Passenger Goodwill and the Spill Rate

Airlines are justifiably concerned about the impact of discount seat allocation policies on the number of full fare reservation requests that must be returned away. This number

expressed as a proportion of total full fare demand is the *passenger spill rate*, or simply *spill rate*. The related, but different. proportion of flights on which one or more reservation requests are turned away, or the *flight spill rate* has been discussed earlier.

With monotonically associated demands, full fare spill rates are most severe when discount demands are sufficiently high that the discount booking limit is always reached. Under these circumstances, the *maximal* flight spill rate and the *actual* flight spill rate will be the same. If an optimal seat allocation rule is used (in either the independent or dependent demand case), the flight spill rate will be close to the discount/full fare ratio.

For example, consider the independent demand case with a plane capacity of 100 seats in Table 4.1. If mean low fare demand is significantly higher than 70 seats so that the discount booking limit of 73 seats is reached most the time, and full fare mean demand remains at 30 seats, then the flight spill rate is approximately 60%, since $\rho_B/\rho_Y = 0.60$.

In this example, the full fare passengers are essentially being booked into a fixed allocation of 27 seats. In a report prepared by Boeing Computer Services, Harmer [42] gives a simple formula relating passenger and flight spill rates under these circumstances, when the full fare class has a normal demand distribution.

Let α denote the fixed allocation of seats, and let r_P andr_F denote the passenger and flight spill rates respectively. Then

$$r_P \;=\; (\sigma_Y/\mu_y)(\phi(z_\alpha) - z_\alpha r_F); \tag{4.36}$$

where μ_Y and σ_y are the mean and standard deviation respectively of the demand distribution, $\Phi(\cdot)$ is the standard normal probability density, and z_α is the standardized allocation $(\alpha - \mu_Y)/\sigma_Y$. Upon applying this formula to the example above, it is found that the passenger spill rate corresponding to the 60% flight spill rate is 21%.

It is difficult to obtain reliable data on actual airline passenger spill rates, but it is hard to imagine that airline managers would tolerate turning away 21% of their best customers, even given the high demand for discount fares assumed in the example.

There thus appears to be a substantial discrepancy between spill rates corresponding to optimal booking limits and the spill rates that would be tolerated by airlines. Possible explanations for this discrepancy include the following:

1. Optimal allocation rules may simply not be used by many airlines.

2. The airlines may be compensating for demand dependencied, either deliberately or on a trial-and-error basis, by lowering discount booking limits below those specified by the simple allotment rule.

3. The discount and full fare demands may overlap in time to a sufficient degree, that the observed full fare demand can be used to adjust the discount booking limit.

4. Voluntary "bumping" of discount passengers may be used to permit high overbooking levels for full fare passengers, thus reducing the effective full fare spill rate.

5. The discount booking limits may be adjusted downward in an ad hoc fashion to compensate for the perceived extra value of full fare passengers above and beyond thier higher fares. (Full fare passengers are predominantly composed of business travelers who can be expected to travel more frequently than the discount, predominantly leisure, travelers. Low spill rates can be seen then as a way of promoting future earnings from these customers by maintaining passenger *goodwill*.)

The latter case, which recognizes the goodwill benefits associated with serving the full fare passenger, is now examined.

The effect of not being able to accommodate a full fare passenger can be viewed in two ways. First, a *goodwill premium* of ρ_G can be included in the full fare.Alternatively, the revenue derived from a full fare can be kept at ρ_Y, and a *loss* of goodwill can be incurred for each full fare customer not accommodated. The argument used in Subection 4.4.2 can be applied to this version of the revenue model to derive the optimality condition.

$$\begin{aligned} \eta* &= \max\{\eta \geq 0 : G(\eta) > 0\} \\ &= \max\Big\{0 \leq \eta \leq k : \\ &\qquad P[Y > k - \eta | B \geq \eta] < \frac{\rho_B}{\rho_Y + \rho_G}\Big\}. \end{aligned} \tag{4.37}$$

It is clear from (4.33) with ρ_Y replaced by $\rho_Y + \rho_G$, and from (4.37), that the optimal allocation will be identical with either interpretation of goodwill. That is, the same number of seats should be protected whether a loss of goodwill is incurred worth ρ_G per full fare customer denied a booking or whether a gain of goodwill is accrued worth ρ_G per full fare customer booked. In either case, the *incorporation of goodwill considerations will increase the full fare protection level and reduce the full fare spill rate.*

To illustrate one implication of (4.37), consider an airline that wishes to limit its passenger spill rate to 3%. From formula (4.36), using the same assumptions as the example given above, a passenger spill rate of 3% corresponds to a flight spill rate of 15%. And this in turn corresponds to a goodwill premium of $\rho_G \approx 3\rho_Y$ (the solution to: $0.15 = 0.6\rho_Y/(\rho_Y + \rho_G)$). Thus a goodwill premium of three times the full fare would be required to justify restricting the spill rate to 3% (assuming none of the other spill rate control methods, mentioned previously, are being used). It is not clear whether such a high premium is justified. Such a justification would depend upon an airline's assessment of the proportion of thier full fare customers who might be lost permanently to competitors after failing to obtain a booking.

Perhaps one of the chief uses of Equation (4.37) would be, as in this example, to impute the goodwill premium implied by a particular spill rate policy.

4.4.4 Upgrades

We now examine the case in which the dependency between discount and full fare demands arises because of a tendency for some discount fare customers to upgrade to full fares if denied a discount reservation. If this context, it will be assumed that the upgrading tendency is the *only* source of dependency and that the initial B and Y demands (i.e., before upgrading) are independent. Under these circumstances, the *ultimate* Y demand will depend both on the B demand and on the booking limit set for the B demand. It is this dependency on the booking limit that necessitates an analysis separate from and more involved than that for the dependent demand case discussed in Subsection 4.4.2. Note that the optimality condition derived here was previously proposed without formal proof by Belobaba [12, p. 130, equation 5.53] and that a similar result has been obtained independently by Pfeifer [84] using different methods.

The purpose here is to provide a formal proof of the result within the context of a general model for the seat allocation problem.

To model the upgrading, define

$$D_i = \begin{cases} 1 & \text{if the } i\text{th customer would upgrade if denied a discount fare,} \\ 0 & \text{otherwise.} \end{cases} \tag{4.38}$$

Assume that $\{D_1, D_2, \cdots\}$ are independent and identically distributed with $ED_i = \gamma$ being the probability that a customer denied a discount fare will upgrade. Also assume independence of the process $\{D_1, D_2, \cdots\}$ of upgrades, the demand B for discount fares, and the demand Y for fares *exclusive of the upgrades*. Let $U(\eta)$ denote the total number of upgrades when the discount booking limit is η; that is,

$$U(\eta) = \sum_{i=\eta+1}^{i=B} D_i. \tag{4.39}$$

This quantity is, of course, zero if $B \leq \eta$. Identification of this model with the general revenue model (4.27) is the same as in Subsection 4.4.2 except that now

$$Y(\eta) = Y + U(\eta) \tag{4.40}$$

is the sum of the full fare demand and upgrades.

To motivate the optimality condition, marginal analysis can be used as in Belobaba. If a discount fare customer is booked, then the revenue is ρ_B. If a discount fare customer cannot be booked, then with probability γ there is an upgrade generating revenue ρ_Y, and with probability $1-\gamma$ there is no upgrade. In the later case, the booking decision will have no impact on revenue if $B \leq \eta$. However, if $B > \eta$, then additional revenue ρ_Y will be obtained if the seat being considered is used either by some other upgrade or by a full fare customer. This analysis leads one to conjecture that it is optimal to book a discount fare customer if

$$\rho_B \;>\; \gamma\rho_Y + (1-\gamma)P[(Y+U(\eta)) > k-\eta | B \geq \eta]. \tag{4.41}$$

To verify this optimality condition, compute $G(\eta)$ from (4.29). Let$H(\eta) = [Y(\eta) \wedge (k-\eta)] - [Y(\eta-1) \wedge (k-\eta+1)]$. To evaluate H consider two cases. First suppose that $Y(\eta) > k-\eta$. Then $Y(\eta-1) \geq k-\eta+1$ and $H(\eta) = -1$. Second, suppose that $Y(\eta) \leq k-\eta$. Then $Y(\eta-1) \leq k-\eta+1$ and $H(\eta) = Y(\eta) - Y(\eta-1) = -D_\eta$. Thus (5) reduces to

$$\begin{aligned} G(\eta) &= \rho_B - \rho_Y P[Y(\eta) > k-\eta|B \geq \eta] - \rho_Y P[Y(\eta) \leq k-\eta|B \geq \eta]E[D_{eta}] \\ &= \rho_B - (1-\gamma)\rho_Y P[Y(\eta) > k-\eta|B \geq \eta] - \gamma\rho_Y, \end{aligned} \tag{4.42}$$

where the assumption that D_η is independent of B and of Y is used to obtain the first equation.

It remains to be shown that the problem is monotone by establishing that $G(\eta)$ is nonincreasing in η. Using the fact that $D_i \leq 1$ gives

$$\begin{aligned} P[Y(\eta-1) > k-\eta+1|B \geq \eta-1] &= P[(Y + \sum_{i=\eta}^{i=B} D_i) > k-\eta+1|B \geq \eta-1] \\ &\leq P[(Y + \sum_{i=\eta+1}^{i=B} D_i) > k-\eta|B \geq \eta-1] \\ &= P[Y(\eta) > k-\eta|B \geq \eta-1]. \end{aligned} \tag{4.43}$$

By conditioning on whether $B = \eta-1$ or $B \geq \eta$, and manipulating the conditional probabilities, $P[Y(\eta) > k-\eta|B \geq \eta-1]$ can be written as

$$\begin{aligned} P[Y(\eta) > k-\eta|B \geq \eta] \;&+\; P[B = \eta-1|B \geq \eta-1] \\ &\cdot\; (P[Y(\eta) > k-\eta|B = \eta-1] - P[Y(\eta) > k-\eta|B \geq \eta]). \end{aligned} \tag{4.44}$$

The difference in the last term cannot be positive since

$$\begin{aligned} P[Y(\eta) > k-\eta|B \geq \eta] &\geq P[Y > k-\eta|B \geq \eta] \\ &= P[Y(\eta) > k-\eta|B = \eta-1], \end{aligned} \tag{4.45}$$

where the assumption that Y and$B$ are independent and the observation that $U(\eta) = 0$ if $B = \eta-1$ are used to obtain the last equation. Replacing the difference in (4.44) by 0, and using the inequality (4.43), shows that

$$P[Y(\eta-1) > k-\eta+1|B > \eta-1] \leq P[Y(\eta) > k-\eta|B > \eta], \tag{4.46}$$

and so $G(\eta)$ is nonincreasing. Then, from (4.39) and (4.42), $G(\eta)$ will be positive as long as

$$P[(Y+U(\eta)) > k-\eta|B \geq \eta] < \frac{\rho_B - \gamma\rho_Y}{(1-\gamma)\rho_Y}, \tag{4.47}$$

which is equivalent to (4.41). Define $\eta*$ to be the largest $\eta (0 \leq \eta \leq k)$ satisfying (4.47). As with optimality condition (4.33), set $\eta* = 0$ if no η can satisfy (4.47). This will be the case, for example, when γ is sufficiently large that the right hand side of (4.33) is nonpositive. This $\eta*$ satisfies the condition in the definition of a monotone problem and $|Y(\eta) - Y(\eta-1)| \leq 1$. Hence the problem is monotone and it is optimal to book discount fares up to $\eta*$.

Implementation

The comments made earlier regarding implementation of the dependent demand solution apply again here. In the present case estimation of the joint distribution of $Y + U(\eta)$and B will be somewhat easier aince Y and B can be estimated independently and then Y adjusted by the binomial distribution $U(\eta)$ for each η. Altenatively, spill rate control approach could be applied with no change except for adjustment of the discount/full fare ratio as indicated in (4.47).

A numerical example of the use of the upgrades formula is provided in Belobaba [13, pp. 138-139].

4.4.5 Summary

This subsection has presented a simple resource allocation model and applied it to airline seat allocation problems. For ease of reference, the main results are summarized below:

1. When discount and full fare demands are bivariate normal with positive correlation, optimal discount seat booking limit will be less than or equal to that specified by Littlewood's rule (independent demand). The optimal limit will decrease as the correlation increases.

2. With monotonically associated discount and full fare demands B and Y, respectively, cabin capacity k, discount fare ρ_B, full fare ρ_Y, and full fare goodwill premium ρ_G; it is optimal to lomit discount fare bookings to $\eta*$ seats, where:

$$\eta* = \max\left\{0 \leq \eta \leq k : P[Y > k - \eta | B \geq \eta] < \frac{\rho_B}{\rho_Y + \rho_G}\right\}.$$

   Again, this will result in a lower discount seat booking limit.

3. When the discount and *initial* full fare demands are independent but the presence of upgrades creates a dependency between discount and*ultimate* full fare demand, results (4.37) and (4.47) can be combined to obtain the following optimal discount seat allocation:

$$\eta* = \max\left\{0 \leq \eta \leq k : P[Y + U(\eta) > k - \eta | B \geq \eta] < \frac{\rho_B - \gamma(\rho_Y + \rho_G)}{(1-\gamma)(\rho_Y + \rho_G)}\right\},$$

   where γ is the upgrade probability, and $U(\eta)$ is the total number of upgrades given discount allocation η.

Once again, this implies lower discount seat booking limits.

It has been shown that these conditions are optimal among all policies that use only the information $B > \eta$. Given stable fares, the only possible justification for changing an optimal booking limit is a perceived shift in the joint demand distribution for discount and full fares. Thus, for example, the occurrence of a sudden "flurry" of discount demand at some point in the booking process cannot in itself justify a change in the booking limit unless it can be validly associated with a change in the joint demand distribution. If it is decided that such a change has occured, a reasonable response is to simply recalculate the optimal booking limit on the basis of the new joint demand distribution and seat capacity remaining for the flight. More sophisticated dynamic modeling is required to optimally account for the possibility of periodic revision of the joint demand distribution on the basis of more information than $B > \eta$.

The three variants of optimal booking conditions given above all suggest lower discount booking limits than those implied by Littlewood's rule for independent demands. This is important since these results are more easily reconciled with the low full fare passenger spill rates actually observed in practice. Numerical examples suggest that the revenue gains from application of these conditions may be modest (e.g., 1.3% in the extreme 0.9 correlation case in Table 4.1). However, given the largely fixed cost, low margin nature of airline operations in competitive markets, such revenue gains represent almost pure profit and thus are greatly magnified in terms of profit impact.

With the exception of independent demands, this paper has retained many of the strong assumptions required in earlier work on the two-fare allocation problem. One direction for further work is in developing the type of dynamic policy mentioned above, while another is in estimating joint demand distributions on the basis of data that has been censored by the presence of booking limits. The authors are currently pursuing both of these topics.

Chapter 5

Optimal Portfolio Selection Models

5.1 Introduction

In this chapter we demonstrate that when there are more than two assets, we show how to derive an optimal portfolio so as to maximize the expected utility function defined on the wealth of an investor. The key idea is a trade off between return and risk. This observation is one of the motivation to characterize optimal portfolios which have the minimum risk for the various levels of expected rate of return.

In Section 5.2 we develop an asset allocation model with various risk measures which is quite different from the mean-variance porfolio models. From institutional investors' perspective the purpose of investing is to achieve a target level of rate of return to meet the cash flows of the business. A situation unfavorable to this purpose is penalized as a risk. The model developed here is in closer agreement with actual practice in Japanese financial institutions. (Refer to Sawaki [105])

In Section 5.3 we treat with a systematic approach of optimal consumption and portfolio selection models under the setting of stochastic optimal control. Stochastic processes for the asset prices are semi-martingale. Main results obtained for the model are an optimal policy for consumption and portfolio selection, an equation that the expected excess return of each asset should satisfy, and the closed solution for a special class of risk-averse utility functions. Those results are driven by Sawaki and Lin [108].

In Section 5.4 we discuss optimal exercise policies for a discrete time option model in which state of the economy follows a Markov chain and stock prices fluctuate according to the distribution of the product of independent positive random variables. We show under

some specific assumptions that there exist a simple optimal exercise policy which depends only on the stock price and the state of economy. Furthermore, a simple alternative derivation of the Black and Scholes' option pricing formula is presented by the means of an analysis developed by Sawaki [103].

5.2 An Asset Allocation Model with Various Risk Measures

For the majority of institutional investors the main objective of asset allocation is achieving sufficient return to satisfy the demand for cash that is generated by the business activities of the enterprise. What is the risk to intitutional investors in the light of this objective ? Previous analysis have measured risk in the terms of mean and variance types (Markowitz 1959), safety-first criterion types (Kataoka 1963 ; Pyle and Turnovsky 1970), absolute-deviation types (Konno and Yamazaki 1991), and, more generally, by the shape of utility function. The problems pointed out by these measures of risk were, one, the need for an enormous amount of computation and an enormous amount of data input labor for large-scale problems, and two, the difficulty of identifying investors' risk preferences and the discrepancy between investment behavior among institutional investors.

In this section risk is defined as an allocation performance that falls short of the target return of asset allocation made on the assumption of institutional investors, and a new asset allocation model is developed that imposes a penalty when such unfavourable situations have been generated. The advantages of this approuch are, one, it avoids the problems mentioned above, and two, it becomes possible to avoid handling expressly the utility function that is the risk preference of investors. Recognizing as a risk a situation that falls short of the target return seems to be an approach that agrees more closely with actual practice in the investment arena by institutional investors in Japan. This sort of formulation is an optimization problem that belongs to the class of stochastic programmings with recourse and the actual problem is one of large-scale mathematical programming. The most essential characteristic of an asset allocation model is that, to counter the fact that there is uncertainty about asset returns, the optimization technique based on mathematical programming basically assumes the handling of a deterministic amount. Hence, the computational algorithm for the asset allocation model also involves a development of the procedures on how to transform a stochastic quantity into a deterministic quantity.

In Subsection 5.2.2 an asset allocation model that maximizes terminal wealth under the constraint that short of the target return is formulated as an asset allocation model with penalty costs, and we lay down the conditions that must be satisfied by optimal allocation. This is followed by a more detailed analysis of cases in which there are two types os asset class and cases in which the distribution of the rate of return is normal distribution. In Subsection 5.2.3 we propose more general risk measures that embrace the various risk measures that have been suggested in the past, and we formulate an optimization problem that deals with a trade off between risk and return. We touch upon the possibility of constructing, by making the optimization problem the parameter and creating a return-and-risk set, a new efficient frontier and a capital asset pricing model. In Subsection 5.2.4, by way of conclusion we shall list the advantages of the new asset allocation model that is proposed here the differences between it and other models. In addition, we shall look at this asset allocation model from the perspective of performance assesment of asset allocation.

5.2.1 An Asset Allocation Model with Penalty Costs

In this subsection we wish to formulate an asset allocation model that assumes intitutional investors. In the Markowitz-type mean-variance model, risk was measured by the variance of portfolio returns. Institutional investors do not necessarily have such votality as their main concern. Rather, they see as a risk an insufficiency of returns that would act as an obstacle to business activities, or circumstances that would necessitate a transfer of funds among several accounts that should be managed independently, and they consider the principle aim of asset allocation to be ensuring of a sufficient cash flow to avoid such an undesirable situation. In conformity with the actual practice of asset allocation by institutional investors, let us define as risk situation in which the asset allocation performance of institutional investors falls short of the target return in advance, and let us assume that, when this undesirable situation has arisen, the institutional investors impose a penalty cost on their own objective function.

Consider the case of n types of risk assets and one type of riskless asset. there is,

therefor, an asset class of $(n+1)$ types, and the institutional investor makes a decision regarding asset allocation under this class. We make use of this following notation :

r_i = the rate of return of risky asset i, $i = 1, 2, \dots, n$.

r_0 = the rate of return of riskless asset.

x_i = the fraction of the asset allocation invested in asset i.

R_0 = the target rate of return of an asset allocation.

$C(\cdot)$ = the penalty function defined on the set of rates of return.

μ_i = the expected rate of return of risky asset $i, i = 1, 2, \dots, n$.

Let $x = (x_0, x_1, \dots, x_n)$ be an asset allocation where $\sum_{i=0}^{n} = 1$ and short sales are allowed. We assume that $R^0 \geq r^0$. Define $R_i = r_i - r_0, \bar{R}_i = \mu_i - r_0, i = 1, 2, \dots, n$, and $\bar{R}^0 = R^0 - r_0$.

The objective function is to choose an asset allocation of maximizing the total rate of return $\sum_{i=0}^{n} r_i x_i$ subject to $\sum_{i=0}^{n} r_i x_i \geq R^0$ and $\sum_{i=0}^{n} x_i = 1$. However, the objective function contains random variables and so does the constraint. Therefore, by using the relation $x_0 = 1 - \sum_{i=0}^{n} x_i$, this conditional optimization problem may be transformed to the unconditional problem as follows :

$$\max E[\sum_{i=1}^{n} \bar{R}_i x_i - C((\sum_{i=1}^{n} R_i x_i - \bar{R}^0)^-)], \tag{5.1}$$

where $(x)^- = \min\{x, 0\}$ and E denotes the expectation operator.

Assumption. $C(x)$ is decreasing in x for $x < 0$ and $C(x) = 0$ for $x \geq 0$, and is continuously differentiable.

From the first condition of optimality we have

$$\bar{R}_i - E[C'((\sum_{i=1}^{n} R_i x_i^* - \bar{R}^0)^-) R_i] = 0, \quad i = 1, 2, \dots, n.$$

By multiplying x_i and summing up with respect to i, an optimal asset allocation x^* shoul satisfy

$$\sum_{i=1}^{n} \bar{R}_i x_i^* - E[C'((R_i x_i^* - \bar{R}^0)^-) \sum_{i=1}^{n} R_i x_i^*] = 0. \tag{5.2}$$

Note that $C'(\cdot) \leq 0$.

When there are two asset classes

We assume that there are only two asset classes, risky asset and riskless asset. Let μ be the expected rate of return of the risky asset and x the associated fraction of the asset allocation. The penalty cost is assumed to be piecewise linear as follows : for some $p > 0$

$$C(y) = \begin{cases} 0, & y \geq 0 \\ -p \cdot y, & y < 0 \end{cases} \tag{5.3}$$

Under this assumption equation (5.1) can be rewritten as

$$\max\{(\mu - r_0)x + r_0 + pE[(rx + r_0(1-x) - R^0)^-]\}. \tag{5.4}$$

Let $K(t)$ be the realization of the portfolio rate return $rx + r_0(1-x)$ which value is just equal to R^0, that is,

$$K(x) = \frac{R^0}{x} - \frac{r_0(1-x)}{x}. \tag{5.5}$$

The equality of the bracket of equation (5.4) can be reduced to

$$(\mu - r_0)x + r_0 + p\int_{-\infty}^{K(x)} ((r - r_0)x + r_0) - R^0)dF(f) \tag{5.6}$$

where $F(r)$ is the distribution function of the rate of return of risky asset. By taking the derivative of equation (5.6) an optimal allocation to the risky asset x^* must satisfy

$$(\mu - r_0) + p\int_{-\infty}^{K(x^*)} rdF(r) = pr_0F(K(x^*)). \tag{5.7}$$

To explore the existence of a solution in equation (5.7), we put $y \equiv K(x)$. It is easily seen that

$$K'(x) = \frac{r_o - R^0}{x^2}0 \quad \text{for} R^0 > r_0$$

and

$$K'' = \frac{2(R^0 - r_0)}{x^3}.$$

The inverse function $x = K^{-1}(y)$ exists for all y except $y = r_0$.

$K(x)$ is decreasing in x, convex for $x > 0$ and concave for $x < 0$. If no short sales are allowed, $R^0 \leq K(x) < \infty$ for $0 < x < 1$, and if short sales are allowed, $-\infty < K(x) < \infty$. Equation (5.7) can be rewritten as

$$\mu - r_0 + p\int_{-\infty}^{y} rdF(r) = pr_0F(y). \tag{5.8}$$

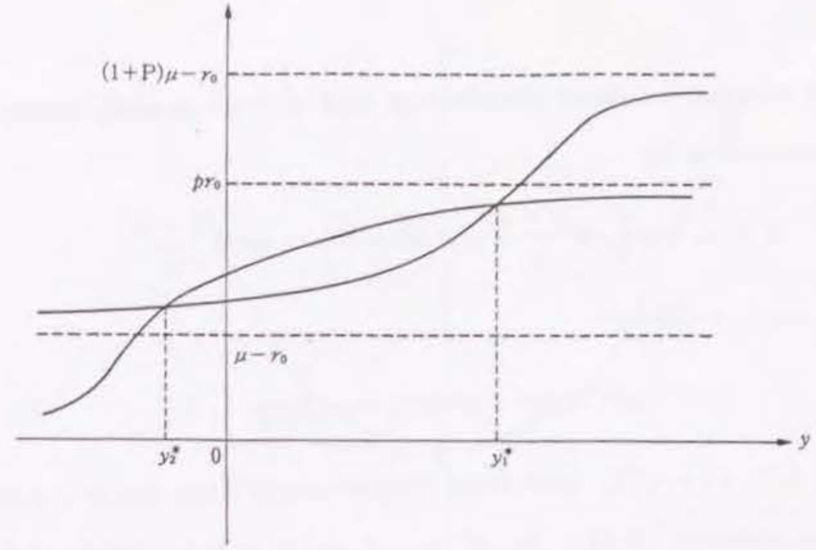


Figure 5.1: A Pair of Optimal Values

Next, we shall investigate whether or not a solution of equation (5.8) exists. It can be shown that the limit of the left hand side of equation (5.8) goes to $(1+p)\mu - r_0$ as $y \to \infty$ and $\mu - r_0$ as $y \to -\infty$. On the other hand the limit of the right hand side goes to pr_0 as $y \to \infty$ and) as $y \to -\infty$. Note that we have $(1+p)\mu - r_0 > pr_0$ and $pr_0 > \mu - r_0$ for p large enough. If solution of equation (5.8) exist, then there exist at least two solutions, say y_1^*, y_2^*. (See Figure 5.1). If no short sales are allowed, the solution is unique. So we reach to the following proposition.

Proposition 1 If no sales are allowed and $\mu - r_0 + p\int_{-\infty}^{R_0} rdF(r) < pr_0$ holds, then there exists an optimal solution. If short sales are allowed, then there exist at least two optimal allocations.

The Case of the Rate of Return Normally Distributed

Suppose that the rate of return is normally distributed with mean μ and variance σ^2. Putting $u = (r - \mu)/\sigma, u$ is normally distributd with mean 0 and varaince 1. Then we obtain

$$\begin{aligned}\int_{-\infty}^{y} rdF(r) &= \int_{-\infty}^{(y-\mu)\sigma} (\mu + \sigma u)\phi(u)du \\ &= \mu\Phi\frac{y-\mu}{\sigma} - \sigma\phi\frac{y-\mu}{\sigma}\end{aligned}$$

where $\Phi(\cdot)$ is the normal standard distribution and $\phi(\cdot)$ its density function. Hence, equation (5.8) turns out to be

$$\mu - r_0 + p\left\{\mu\Phi\frac{y-\mu}{\sigma} - \sigma\phi\frac{y-\mu}{\sigma}\right\} = pr_0\Phi\frac{y-\mu}{\sigma}. \tag{5.9}$$

Putting $z = (y-\mu)/\sigma$, we have

$$(\mu - r_0)(1 + p\Phi(z)) = p\sigma\phi(z). \tag{5.10}$$

If $p\sigma\sqrt{2\pi} \geq (\mu - r_0)(1+\mu/2)$, then from Proposition 1 there exists a pair (z_1^*, z_2^*) of solutions satisfying equation (5.10). So, $y_i^* = \mu + \sigma z_i^*, i = 1, 2$. Hence, the associated optimal asset allocation for the risky asset is given by

$$x_i^* = \frac{R^0 - r}{\mu - r + \sigma z_i^*}, \quad i = 1, 2. \tag{5.11}$$

Note: When you look at (5.11) carefully, you can see that z_i^* depends on penalty cost p when the target return R^0 has not been met. When $z_1^* > 0, z_2^* < 0$, if σ increases the allocation ratio to risky assets, x_1^* decreases and x_2^* increases. The higher you set the target return, the more x_i^* increases. If you set target return at μ, you get $0 < x_1^* < 1$, but x_2^* depends on the size of z_2^*.

5.2.2 Various Risk Measures

In this subsection we propose more general risk measures that can embrace the various risk measures put forward in the past, and we formulate the problem of a trade off between risk and return. Let us define the function m of such risk measures as the following :

$$m(r; \alpha^-, \alpha^+, k^-, k^+) = \begin{cases} \alpha^- r + \frac{(k^-)^2}{2} - \alpha^- k^-, & r \leq k^- \\ \frac{r^2}{2}, & k^- < r \leq k^+ \\ \alpha^+ r + \frac{(k^+)^2}{2} - \alpha^+ k^+, & r > k^+ \end{cases} \tag{5.12}$$

where the parameter α^-, α^+, k^- and k^+ must satisfy $\alpha^- \leq k^- \leq k^+ \leq \alpha^+$, the risk measure $m(\cdot; \alpha^-, \alpha^+, k^-, k^+)$ is a convec and piecewise quadratic function. The existing risk measures are the special cases of our risk measure. (See Figure 5.2)

(i) If $k^- \to \infty$ and $k^+ \to \infty, m(\cdot;\cdot)$ is the variance which is Markowitz type.

(ii) If $\alpha^- = k^-$ and $\alpha^+ = k^+, m(\cdot;\cdot)$ reduces to the risk measure introduced by King [56].

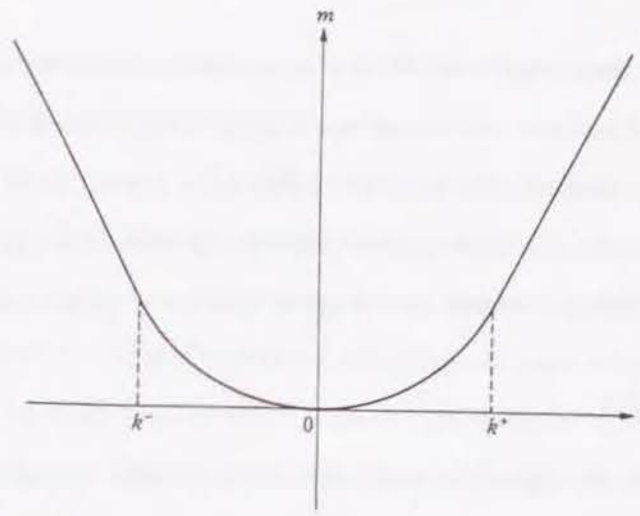


Figure 5.2: A Generalized Risk Measure

(iii) If k^- and k^+ both converge to 0, $\alpha^- = -1$, $\alpha^+ = 1$, then $m(\cdot;\cdot)$ turns out to be equal to the risk measure proposed by Konno and Yamazaki [59].

(iv) if $k^- \to \infty$ and $k^+ \to 0$, then $m(\cdot;\cdot)$ becomes the lower semi-variances risk measure.

(v) If $\alpha^- = p, k^- = R^0$, and $k^+ = \alpha^+ = R^0$, then $m(\cdot;\cdot)$ come to be equal to the penalty function $p(\cdot)$.

Next, let us define the risk measures by defining the probability of ıalling short of the target return. Thus, if we take the risk measure m when $R \equiv \sum_{i=0}^n r_i x_i$ as

$$m(R; R^0) = \frac{1}{\Pr\{R \geq R^0\}} - 1 = \frac{\Pr\{R < R^0\}}{\Pr\{R \geq R^0\}} \tag{5.13}$$

then, when the target return has been definitely achieved, $m(\cdot;\cdot) = 1$, and when the probability of achieving the target return approaches 0, $m(\cdot;\cdot)$ becomes infinitely large. Also, m has the desirable characteristic that a monotone increasing function of R^0. In regard to such a variety of risk measures and the target return R^0, let us define the following conditional optimization problem.

$$\min_x E[m(\sum_{i=0}^n r_i x_i - \sum_{i=0}^n \mu_i x_i; \alpha^-, \alpha^+, k^-, k^+)] \tag{5.14}$$

subject to

$$\sum_{i=0}^n \mu_i x_i \geq R^0, \quad \sum_{i=0}^n x_i = 1.$$

To solve this problem, if we take R^0 as the parameter, then we can draw the efficient frontier on the (R^0, m) plane. As there is a CAPM model for when m is the standard deviation, so it is also theoritically possible to discuss a capital asset pricing model on the (R^0, m) plane. It is especially to be noted that the optimization problem of m defined in equation (5.12) still remains within the range of quadratic programming.

By Way of Conclusion

We have proposed, in regard to asset allocation models (portfolio selection models), an approach different from previous ones. We analyzed in particular a model by means of measures different from the standard deviation as a risk measure related to an investor's return. As a result we have not defined the investor's risk preference as information on utility function. We have also derived an equation that an optimal asset allocation ought to satisfy under a framework different from a mean-variance type of model. There are at least two points in which there is an important difference between our model and mean-variance models. The first point is that, whereas the previous types of models are trade off between the average value or return and standard deviation, in our model we have a trade off between target return and the penalty costs for a shortage in it. The second point is that it uses all information concerning the probability distribution of the rate of return of assets. This is an especially important point when the probability distribution of return does not follow normal distribution.

A method that acknowledges as a risk a situation that falls short of target return and imposes penalty costs on such a risk is known as stochastic programming with recourse. Problems based to a realistically meaningful extent on large-scale stochastic programming require an enormous amount of computation to obtain an optimalsolution. For this reason we have prepared numerical examples based on the computational algorithm of Rockafellar and Wets, which splits stochastic programming by means of the scenario method into subproblems based on normal mathematical programs.

An optimization problem that takes into account a trade off between target return and penalty costs for a shortage in it would seem to be closer to the invesment environment of institutional investors in Japan, and more faithful to the risk measures of people who are not allocating their own funds. Also, target return does not differ widely from

institutional investor to institutional investor ; rather, it is strongly regulated by the term structure of the bond market and such security-market benchmarks as the Tokyo Stock Index and the Nikkei average. Mindful that these are business indicators common to institutional investors, and that our model is independent of the utility function that is the risk preference of investors, we present in our model, we believe, a more objective norm form the perspective of performance evaluation of asset allocation. Assuming that utility function varies with each investor makes it possible to carry out a comparison of performance evaluation in reliance upon the information that is more commonly shared among investors, that of the probability distribution of return.

5.3 Optimal Portfolio Selection and Asset Pricing Models for Semi-Martingale Processes

The study of an individual investor's optimal consumption and portfolio selection was one of the principal themes in finance theory. The chief problem was the question of whether an optimal policy exists in regard to consumption and portfolio selection when a utility function and an asset price process in continuous time have been given. One had to prepare an answer to the question: If it does exist, what properties does that optimal policy possess? (See Merton [70] and Cox and Huang [26]) Just as the static model CAPM depends on the mean-variance type of portfolio selection, so the dynamic asset pricing model (the Intertemporal Capital Asset Pricing Model) presupposes consumption and portfolio selection in continuous time. Stochastic processes of asset prices almost all assume geometric Brownian motion. But two problems are often pointed out as the reasons why geometric Brownian motion is abandoned as a stochastic process of asset prices: first, while it has been empirically tested that there is a time series correlation in asset return, Brownian motion does not have such a time series correlation; and secondly, volatility does not depend on asset price and time.

This study aims at overcoming the theoretical defects of the geometric Brownian motion by considering the semi-martingale as a stochastic process of asset pricing and, also in regard to utility function, it examines a consumption and portfolio selection model under a general class. An intertemporal model that takes Merton [70] as its starting point has been developed by means of specification and generalization with regard to the asset price process and the utility function of the investor. (See Aase [1], Back [7], Cox, Ingersoll and Ross [27], and Ingersoll [48]) In our study we develop an intertemporal asset pricing model in a semi-martingale. In particular we derive an equation that the value for the expected excess return should satisfy, and we indicate that this equation represents a modification of the accepted CAPM formula. There are two benefits in examining a consumption and portfolio selection model for occasions in which asset prices follow a semi-martingale. The first benefit is that an optimal policy for consumption and portfolio selection possesses robust properties from the asset price process. The second benefit is that, when investors in the market are all using this optimal policy, an asset pricing formula that is derived from the demand-supply conditions of the market is similar to the CAPM formula.

This study is made up of four subsections. In the first subsection we explain asset price processes and give examples of semi-martingales. In the second subsection we formulate a consumption and portfolio selection model as the optimal control problem, and we discuss the optimal policy for consumption and portfolio selection. We follow this by deriving an equation that the expected excess return for each asset should satisfy from demand-supply conditions, and we go into its economic implications. In the third subsection we express in concrete fashion the results obtained in Subsection 5.3.2, restricting the utility functions to a special class of risk-averse utility functions and assuming asset prices follow a geometric Brownian motion. Finally, in Subsection 5.3.4 we draw the conclusions of this study and at the same time touch upon directions future research might take.

5.3.1 Asset Price Processes

It is well known that the geometric Brownian motion is not always supported as an asset price process. If asset prices are not described by means of geometric Brownian motion, the derivation of an optimal policy for consumption and portfolio selection becomes extremely difficult. Still, it is possible to investigate the analytical properties of an optimal policy in even more general stochastic processes, say semi-martingale. From the point of view of explaining the Black Monday of October 1987 and of constructing a model that includes actual cases in which asset prices jump as a result of public announcement of information, the semi-martingale has definite theoretical advantages.

Let us take time t continuously and assume it to be an element of a closed interval $[0,T]$. $t \in [0,T]$. There are n types of risky assets, indicated by subscript i. We use the following symbols:

$$\begin{aligned} P_i(t) &= \text{the price of asset } i \text{ at time } t \\ x_i(t) &= \text{the investment ratio of asset } i \text{ at time } t \\ W(t) &= \text{the wealth of the investor at time } t \\ C(t) &= \text{the instantaneous consumption rate at time } t \end{aligned}$$

The price processes are described by the stochastic differential equation as follows:

$$\frac{dP_i(t)}{P_i(t_-)} = dM_i(t), \quad i = 1, 2, \cdots, n, \tag{5.15}$$

where $M_i(t)$ is a semi-martingale process and $P(t_-)$ denotes the left-hand limit at time t.

For the riskless asset, indicated by subscript 0, with instantaneous rate of return r we have

$$\frac{dP_0(t)}{P_0(t)} = r dt. \tag{5.16}$$

The closed solution of equation (5.15) is known (see [74]) and given by

$$P_i(t) = P_i(0) e^{M_i(t) - <M_i^c, M_i^c>_t/2} \Pi_{0 \le s \le t}(1 + \triangle M_i(s)) e^{-\triangle M_i(s)}, \tag{5.17}$$

where $< M_i^c, M_i^c >$ is the bounded variation process for the continuous part of M_i and $\triangle M_i(s)$ denotes the size of the jump of the process $M_i(\cdot)$ at time t. Thus, in this case, $P(t) \ge 0$ with probability 1 if and only if $\triangle M_i \ge -1$.

The change in wealth at time t when the investor follows a consumption and portfolio selection policy $x(t) = (x_0(t), x_1(t), \cdots, x_n(t))^T$ satisfies

$$\begin{aligned} dW(t) &= \sum_{i=0}^{n} x_i(t) W(t_-) dM_i(t) - C(t) dt \\ &= \sum_{i=1}^{n} x_i(t) W(t_-) dM_i(t) + \left[W(t_-) r (1 - \sum_{i=1}^{n} x_i(t)) - C(t) \right] dt. \end{aligned} \tag{5.18}$$

Whenever the stochastic process $M(t)$ can be decomposed as the sum of a local martingale and a bounded variation process denoted by $< M, M >_t$, then $M(t)$ is called a semi-martingale. Let us consider four examples of semi-martingale processes.

Example 1 (geometric Brownian motion)

We replace equation (1) by

$$dM_i(t) = \mu_i dt + \sigma_i dZ_i(t), \tag{5.19}$$

where $Z_i(t)$ is a Wiener process (Brownian motion) with mean 0 and variance t. Then, the price of asset i follows geometric Brownian motion with mean $\mu_i t$ and variance $\sigma_i^2 t$, that is,

$$dP_i(t) = \mu P_i(t)dt + \sigma_i P_i(t)dZ_i(t).$$

Putting $dP_i^M(t) = \sigma_i P_i(t)dZ_i(t)$ and $dP_i^B(t) = \mu_i P_i(t)dt$, we obtain the decomposition as follows:

$$P_i(t) = P_i(0) + P_i^M(t) + P_i^B(t),$$

where P_i^M is the martingale and P_i^B the bounded variation which is denoted by $< M_i, M_i >_t$.

Example 2 (sub-martingale)

Denote the wealth of the investor at time t by $W(t)$, which satisfies

$$W(t) = \sigma Z(t), \tag{5.20}$$

where $Z(t)$ is the Weiner process with the mean 0 and variance t. Since $W(t)$ is martingale, $W^2(t)$ is a sub-martingale. Put $\hat{M}(t) = W^2(t) - \sigma^2 t$ and $< M, M >_t = \sigma^2 t$, and then we have the decomposition as follows:

$$\begin{aligned} W^2(t) &= (W^2(t) - \sigma^2 t) + \sigma^2 t \\ &= \hat{M}(t) + < M, M >_t, \end{aligned}$$

where $\hat{M}(t)$ is the martingale and $< M, M >_t$ the bounded variation process.

Example 3 (Poisson process)

Let $N(t)$be a the Poisson process and put $\hat{M}(t) \equiv N(t) - \lambda t$ and $< M, M >_t \equiv \lambda t$. The Poisson process has the following decomposition.

$$\begin{aligned} N(t) &= (N(t) - \lambda t) + \lambda t \\ &= \hat{M}(t) + < M, M >_t, \end{aligned} \tag{5.21}$$

which implies that the Poisson process is also a semi-martingale.

Example 4 (geometric Brownian motion with jumps)

Consider an example of a semi-martingale that is a composite of examples 1 and 2. Such a stochastic process for the asset price is convenient to describe practical situations that can occur when information on a new technology becomes public.

Let $N_{ik}(t)$ be the number of price changes of jump size β_{ik} at time t which follows a Poisson process with the intensity $\lambda_{ik}(t)$. We assume that β_{ik} are given for $k = \pm 1, \pm 2, \cdots, \pm m$, and $\beta_{ik} > 0$ stands for the jump up and $\beta_{ik} < 0$ the jump down. In connection with equation (1) set

$$\frac{dP_i(t)}{P_i(t_-)} = dM_i(t) = \mu_i dt + \sigma_i dZ_i(t) + \sum_{k=-m}^{m} \beta_{ik} dN_{ik}(t). \tag{5.22}$$

The continuous part of the asset price follows geometric Brownian motion and the discrete part follows Poisson jump process. From examples 1 and 3, the continuous and discrete parts are both semi-martingales and so the sum of them is also a semi-martingale whose decomposition becomes $M_i(t) = \hat{M}_i(t) + < M_i, M_i >_t$, where

$$d\hat{M}_i(t) \equiv P_i(t_-)\sigma_i dZ_i(t) + P_i(t_-) \sum_{k=-m}^{m} \beta_{ik} \left[dN_{ik}(t) - \lambda_{ik}(t)dt\right], \tag{5.23}$$

$$d < M_i, M_i >_t \equiv P_i(t_-) \left(\mu_i dt + \sum_{k=-m}^{m} \beta_{ik}\lambda_{ik}(t)dt \right). \tag{5.24}$$

The martingale part consists of the continuous part $P_i(t)\sigma_i dZ_i(t)$ and discrete parts $P_i(t) \sum_{k=-m}^{m} \beta_{ik}$
$[dN_{ik}(t) - \lambda_{ik}(t)dt]$, respectively. Note that $\Pr\{N_{ik}(t) = 1\} = \int_0^t \lambda_{ik}(s)ds + o(t)$, the probability that the asset i makes a jump is $\lambda_{ik}(t)dt$, and $E[dN_{ik}(t)] = \lambda_{ik}(t)dt$.

5.3.2 The Optimal Control Problem and Asset Pricing Model

As mentioned in Subsection 5.3.1, a semi-martingale can be represented by the sum of a local martingale and a bounded variation. Consider a portfolio $x(t) = (x_0(t), x_1(t), \cdots, x_n(t))^T$ consisting of n risky assets and one riskless asset. The wealth generated by this portfolio is given by

$$W(t) = W(0) + \int_0^t \sum_{i=1}^{n} x_i W(s_-) dM_i(s) + \int_0^t \left[W(s_-) r (1 - \sum_{i=1}^{n} x_i) - C(s) \right] ds$$

$$
\begin{aligned}
&\equiv W(0)+W^{M}(t)+W^{B}(t) \\
&\equiv W(0)+W^{Mc}(t)+W^{Md}(t)+W^{B}(t) \qquad (5.25)
\end{aligned}
$$

with

$$
\begin{aligned}
W^{M}(t) &\equiv W^{Mc}(t)+W^{Md}(t) \\
W^{Mc}(t) &\equiv \int_0^t \sum_{i=1}^n x_i W(s_-) dM_i^c(t) \\
W^{Md}(t) &\equiv \int_0^t \sum_{i=1}^n x_i W(s_-) dM_i^d(t) \qquad (5.26) \\
W^{B}(t) &\equiv \int_0^t \sum_{i=1}^n x_i W(s_-) d < M_i, M_i >_s + \int_0^t \left[W(s_-) r(1-\sum_{i=1}^n x_i) - C(s) \right] ds,
\end{aligned}
$$

where $W^{Mc}(t)$ and $W^{Md}(t)$ are the continuous and discrete parts of the local martingale, respectively.

Our problem is to find an optimal policy with respect to consumption and portfolio selection so as to maximize the expected utility. We need the generalized Ito lemma. We get the bounded variation of the continuous part of the wealth

$$
< W^{Mc}, W^{Mc} >= \int_0^t \sum_{i=1}^n \sum_{j=1}^n x_i x_j W^2(s_-) d < M_i^c, M_j^c >_s .
$$

Let $u(C(t),t)$ be a utility function defined on the consumption and $B(W(T))$ the bequest function defined on the wealth. Given the dynamics of the wealth changes, we define an optimal control problem as follows:

$$
\sup_{C(t),x(t)} E[\int_0^T u(C(t),t)dt + B(W(T))]. \qquad (5.27)
$$

To solve this problem, define the derived utility function $J(W,t)$

$$
J(W,t) = \sup E \left[\int_t^T u(C(s),s)ds + B(W(T)) \mid W(t)=W \right]. \qquad (5.28)
$$

Applying the generalized Ito lemma, we obtain

$$
\begin{aligned}
J(W,t) &= J(W_0,0) + \int_0^t J_W(W(s_-),s)dW(s) + \int_0^t J_s(W(s_-),s)ds \\
&\quad + \frac{1}{2}\int_0^t J_{WW}(W(s_-),s)d < W^{Mc}, W^{Mc} >_s \\
&\quad + \sum_{0\le s\le t} \{J(W(s),s) - J(W(s_-),s) - J_W(W(s_-),s)\triangle W(s)\} \\
&= J(W_0,0) + \int_0^t J_s(W(s_-),s)ds
\end{aligned}
$$

$$
\begin{aligned}
&+ \int_0^t J_W(W(s_-),s)\left[\sum_{i=1}^n x_i W(s_-)(dM_i^c + dM_i^d) + W(s_-)r(1-\sum_{i=1}^n x_i) - C \right] dt \\
&+ \frac{1}{2}\int_0^t J_{WW}(W(s_-),s)\sum_{i=1}^n\sum_{j=1}^n x_i x_j W^2(s_-)d < M_i^c, M_j^c >_s \\
&+ \sum_{0\le s\le t} \{J(W(s),s) - J(W(s_-),s)\} - \sum_{0\le s\le t} J_W(W(s_-),s)\triangle W(s). \qquad (5.29)
\end{aligned}
$$

Note that $W(t)-W(t_-) = W(t_-)\sum_{i=1}^n x_i dM_i^d$, and let τ_{jl} be the length of the sub-interval with the partition of the interval $[0,t]$,

$$
\begin{aligned}
\sum_{0\le s\le t} J_W(W(s_-),s)\triangle W(s) &= \lim_{l\to\infty} \sum_{i=1}^n \sum_{j=1}^l \sum_{\tau_{jl}<t} J_W(W(\tau_{jl}),s)(W(\tau_{jl}) - W(\tau_{jl-})) \\
&= \int_0^t J_W(W(s_-),s)\sum_{i=1}^n W(s_-)x_i dM_i^d,
\end{aligned}
$$

which equals the third term to cancel to each other. The fifth term can be rewritten

$$
\begin{aligned}
&\sum_{0\le s\le t} \{J(W(s),s) - J(W(s_-),s)\} \\
&= \lim_{l\to\infty} \sum_{i=1}^n \sum_{j=1}^l \sum_{\tau_{jl}<t} \{J(W(\tau_{jl}) + W(\tau_{jl})x_i\triangle M_i) - J(W(\tau_{jl}))\} dM_i^d \\
&= \int_0^t \sum_{i=1}^n \{J(W(s_-)+W(s_-)x_i\triangle M_i) - J(W(s_-),s)\} dM_i^d(s).
\end{aligned}
$$

Hence, upon taking expectations, we are left with

$$
\begin{aligned}
E(dJ(W,t)) &= J_t(W,t) + J_W(W,t)W\sum_{i=1}^n x_i dM_i^c + J_W(Wr(1-\sum_{i=1}^n x_i) - C)dt \\
&\quad + \frac{1}{2}J_{WW}W^2\sum_{i=1}^n\sum_{j=1}^n x_i x_j d < M_i^c, M_j^c >_t \\
&\quad + \sum_{i=1}^n E\left\{ [J_W(W(1+x_i\triangle M_i),t) - J_W(W(t),t)]\, dM_i^d(t) \right\}, \qquad (5.30)
\end{aligned}
$$

where $J_t(\cdot)$, $J_W(\cdot)$, and $J_{WW}(\cdot)$ present the first and the second partial derivatives, respectively. The Bellman's optimal equation associated with problem (5.27) now turns out to be

$$
\begin{aligned}
0 &= \sup_{C(t),x(t)} E\{u(C(t),t)dt + dJ\} \\
&= \sup_{C(t),x(t)} E\Big\{ u(C(t),t)dt + J_t dt + J_W W\sum_{i=1}^n x_i dM_i^c \\
&\quad + J_W\left[Wr(1-\sum_{i=1}^n x_i) - C \right] dt + \frac{1}{2}J_{WW}W^2\sum_{i=1}^n\sum_{j=1}^n x_i x_j d < M_i^c, M_j^c >_t \\
&\quad + \sum_{j=1}^n [J(W(1+x_i\triangle M_i),t) - J(W,t)]\, dM_i^d(t) \Big\}.
\end{aligned}
$$

The first order condition that an optimal policy must satisfy is

$$
\begin{aligned}
0 &= u_c - J_W; & (5.31)\\
0 &= J_W \cdot W(E[dM_i^c] - rdt) \\
&\quad + J_{WW} \cdot W^2 \sum_{j=1}^{n} x_j E[d < M_i^c, M_j^c >_t] \\
&\quad + W E[J_{\hat{W}} \Delta M_i dM_i^d(t)] \quad i = 1, 2, \cdots, n, & (5.32)
\end{aligned}
$$

where $u_c = \partial u / \partial c$, and $\hat{W} = W(1 + x_i \Delta M_i)$. Let $\Gamma = [\Gamma_{ij}]$ be the $n \times n$ matrix with an element $E[d < M_i^c, M_j^c >_t]$. Then, solving equation (5.32) we obtain an optimal portfolio

$$x_i^* = -\frac{J_W}{J_{WW} W} \sum_{j=1}^{n} \Gamma_{ij}^{-1} \left\{ E[dM_j^c] - rdt \right\} - \frac{1}{J_{WW} W} \sum_{j=1}^{n} \Gamma_{ij}^{-1} E[J_{\hat{W}} \Delta M_j dM_j^d(t)]. \quad (5.33)$$

When the discrete part of the asset price process vanishes, the equation (5.33) above reduces to

$$x_i^* = -\frac{J_W}{J_{WW} W} \sum_{j=1}^{n} \Gamma_{ij}^{-1} \left\{ E[dM_j^c] - rdt \right\}. \quad (5.34)$$

Now, suppose that all the investors in the market have chosen the portfolio satisfying equation (5.33) as well as (5.32) with respect to the consumption. What relation does the expected excess return of the risky asset satisfy under the market equilibrium? Let D_i^l be the demand of asset i investor l wishes to possess. Then, the amount of demand of asset i is

$$
\begin{aligned}
D_i &= \sum_{l=1}^{L} D_i^l \\
&= \sum_{l=1}^{L} x_i^l W^l \\
&= \sum_{l=1}^{L} \frac{-J_W^l}{J_{WW}^l} \sum_{j=1}^{n} \Gamma_{ij}^{-1} \left\{ E[dM_j^c] - rdt \right\} - \sum_{l=1}^{L} \frac{1}{J_{WW}^l} \sum_{j=1}^{n} \Gamma_{ij}^{-1} E[J_{\hat{W}}^l \Delta M_j dM_j^d(t)] \\
&\equiv A \sum_{j=1}^{n} \Gamma_{ij}^{-1} \left\{ E[dM_j^c] - rdt \right\} + \sum_{j=1}^{n} \Gamma_{ij}^{-1} B_j, & (5.35)
\end{aligned}
$$

where $A = -\sum_l J_W^l / J_{WW}^l$ and $B_j = \sum_l E(J_{\hat{W}}^l \Delta M_j dM_j^d(t)) / J_{WW}^l$. Let x_i be the percentage of asset i among the total asset volume, called market portfolio. The amount of supply for asset i, S_i must equal $S_i = x_i S$.

$$
\begin{aligned}
S_i = x_i S &= D_i \\
&= A \sum_{j=1}^{n} \Gamma_{ij}^{-1} \left\{ E[dM_j^c] - rdt \right\} + \sum_{j=1}^{n} \Gamma_{ij}^{-1} B_j, \quad i = 1, 2, \cdots, n. & (5.36)
\end{aligned}
$$

Rearranging equation (5.36) with respect to the expected excess return of asset i, we obtain

$$E[dM_i] - rdt = \frac{1}{A} \sum_{j=1}^{n} S_j E d < M_i^c, M_j^c >_t - \frac{B_i}{A}. \quad (5.37)$$

Note that even if $E[\Delta M_i dM_i^d(t)]$ is positive, the jump up of asset price does not mean an increase in the expected excess return of the asset.

Using equation (5.32) we obtain the following for A:

$$A = \sum_l \frac{-J_W^l}{J_{WW}^l} = -\sum_l \frac{u_C^l}{u_{CC}^l} \frac{1}{\partial u / \partial W} > 0 \qquad \text{for } \partial u / \partial W > 0.$$

It is possible to give the following economic interpretation with regard to equation (5.37).

1) Since A is the inverse of risk averse function (risk tolerance) resulting from derived utility, risk tolerance decreases as A increases. When this happens, expected excess return also decreases.

2) If we assume the risk measure of asset i to be $d < M_i^c, M_i^c >_t$, then the excess rate of return increases as risk increases.

3) Equation (5.37) tells us that an increase in money supply for asset i pushes up the excess rate of return of that asset in the equilibrium.

The conclusion that has been arrived at here is similar in many respects to the conclusion arrived at in geometric Brownain motion. For example, from equations (5.32) and (5.33), optimal decisions on consumption and portfolio selection are dependent only through derived utility functions and, formally at least, are independent. Also, as long as the parameters in equation (5.34) do not depend on time t, the optimal portfolio is also time independent. In this sense, optimal consumption and portfolio selection can be described as structurally robust after the asset price process.

5.3.3 A Consumption/Portfolio Selection Model under the HARA Type Utility Function

In this subsection we shall specify the class of utility function and explain in fuller detail a consumption/portfolio selection model when asset prices follow a geometric Brownian motion. The optimal policy regarding consumption and portfolio selection derived in the preceding subsection depends on the unknown derived utility function $J(W,t)$, and it is not a closed solution. In order to obtain a closed solution we must substitute equations (5.32) and (5.34) for equation (5.31), and solve the second degree partial differential equation regarding $J(W,t)$ in order to find $J(W,t)$. But unless the utility function is specified and the asset price process is also something simple, it is almost impossible to find a closed solution. In this subsection, therefore, we shall give a concrete form to the class of utility function and derive a closed solution to the problem of an optimal consumption and portfolio selection when asset prices follow geometric Brownian motion.

Consider a continuous time stochastic model in which an individual decides on the optimal consumption rule and portfolio selection that would maximize the expected utility within consumption and portfolio selection consisting of n types of risky assets and one type of safe asset. We assume that changes in the rate of return of the ith risky asset will follow geometric Brownian motion:

$$\frac{dP_i(t)}{P_i(t)} = \mu_i dt + \sigma_i dZ_i(t), \quad i = 1, 2, \cdots, n. \tag{5.38}$$

We define $E[dZ_i(t)dZ_j(t)] = \rho_{ij}dt$.

Using the Ito lemma and Bellman equation, we then get

$$\begin{aligned} 0 &= \max_{C(t),\omega(t)} \left[U(C(t),t) + J_W(W,t)\left(\sum_{i=1}^{n} x_i(t)W(\mu_i - r) + rW - C(t)\right) \right. \\ &\left. + \frac{1}{2}J_{WW}(W,t)\sum_{i=1}^{n}\sum_{j=1}^{n} x_i(t)Wx_j(t)W\sigma_{ij} + J_t(W,t) \right]. \end{aligned} \tag{5.39}$$

Here, subscription denotes the partial derivative for the variables, and σ_{ij} is the $(i-j)$th element of the variance-covariance matrix.

The boundary condition is

$$J(W(T),T) = B(W(T),T).$$

The first order condition that an optimal consumption and portfolio must satisfy is

$$U_c(C^*(t),t) = J_W(W,t); \tag{5.40}$$

$$0 = (\mu_i - r)J_W(W,t) + J_{WW}(W,t)\sum_{j=1}^{N} x_j^*(t)W\sigma_{ij}. \tag{5.41}$$

The decision making on consumption is independent of the portfolio selection. They indirectly depend on each other only through the derived utility function $J(W,t)$.

Define the HARA type of utility function as follows:

$$U[C(t),t] = \frac{1-\gamma}{\gamma}\left[(\frac{\beta}{1-\gamma}C(t) + \eta)^\gamma + \zeta \right] e^{-\rho t}. \tag{5.42}$$

The parameters are restricted to take their values in the following intervals:

$$-\infty < \gamma < \infty, \quad \gamma \neq 1, \quad \beta > 0, \quad \frac{\beta}{1-\gamma}C + \eta > 0.$$

Also, define the bequest function for the finite planning horizon model as

$$B[W(T),T] = \hat{\kappa}\frac{1-\gamma}{\gamma}\left[(\frac{\hat{\beta}}{1-\gamma}W(T) + \hat{\eta})^\gamma + \hat{\zeta} \right] e^{-\rho T} \tag{5.43}$$

with

$$\hat{\kappa} > 0, \quad -\infty < \gamma < \infty, \quad \gamma \neq 1, \quad \hat{\beta} > 0, \quad \frac{\hat{\beta}}{1-\gamma}W + \hat{\eta} > 0.$$

An optimal consumption rule is given by

$$C^*(t) = \frac{1-\gamma}{\beta}\left(\frac{J_W(W,t)e^{\rho t}}{\beta}\right)^{1/(\gamma-1)} - \frac{1-\gamma}{\beta}\eta; \tag{5.44}$$

$$U(C^*(t),t) = \frac{1-\gamma}{\gamma}\left[\left(\frac{J_W(W,t)e^{\rho t}}{\beta}\right)^{\gamma/(\gamma-1)} + \zeta\right] e^{-\rho t}. \tag{5.45}$$

Substituting (5.44) and (5.45) into (5.39), we are left with the partial differential equations

$$\begin{aligned} 0 &= \frac{(1-\gamma)^2}{\gamma}\left[\frac{J_W(W,t)e^{\rho t}}{\beta}\right]^{\gamma/(\gamma-1)} e^{-\rho t} + \left(\frac{1-\gamma}{\gamma}\right)\zeta e^{-\rho t} + (rW + \frac{(1-\gamma)}{\beta}\eta)J_W(W,t) \\ &- \frac{1}{2}(\mu - r\mathbf{1})'\Sigma^{-1}(\mu - r\mathbf{1})\frac{J_W^2(W,t)}{J_{WW}(W,t)} + J_t(W,t), \end{aligned} \tag{5.46}$$

where $\mu = [\mu_1, \mu_2, \cdots, \mu_n]^T$.

In order to solve (5.46), let us try separation of variables and so assume that

$$J(W,t) = e^{-\rho t}\left[A(W+B)^\gamma + D\right],$$

where all parameters A, B, and D are unknown. For the infinite planning horizon the partial differential equation (5.46) can be rewritten as follows:

$$\begin{aligned} 0 &= \frac{(1-\gamma)^2}{\gamma}\left(\frac{\beta}{\gamma}\right)^{\gamma/(1-\gamma)} A^{\gamma/(\gamma-1)}(W+B)^\gamma + \frac{1-\gamma}{\gamma}\zeta + (rW + \frac{(1-\gamma)}{\beta}\eta)\gamma A(W+B)^{\gamma-1} \\ &+ \frac{1}{2}\frac{\gamma}{1-\gamma}A(\mu - r\mathbf{1})'\Sigma^{-1}(\mu - r\mathbf{1})(W+B)^\gamma - \rho\left(A(W+B)^\gamma + D\right). \end{aligned} \tag{5.47}$$

Table 5.1: A List for Optimal Values for HARA Type Utility Functions

Horizon	Infinite($T = \infty$)	Finite($T < \infty$)
HARA Type$U(\cdot)$	$\frac{1-\gamma}{\gamma}\left[(\frac{\beta}{1-\gamma}C(t)+\eta)^\gamma+\zeta\right]e^{-\rho t}$	same as left side
Bequest$B(\cdot)$	none	$\hat{\kappa}\frac{1-\gamma}{\gamma}\left[(\frac{\hat{\beta}}{1-\gamma}W(T)+\hat{\eta})^\gamma+\hat{\zeta}\right]e^{-\rho T}$
Derived Uti. Func.$J(\cdot)$	$[A(W+B)^\gamma+D]e^{-\rho t}$	$[A(t)(W+B(t))^\gamma+D(t)]e^{-\rho t}$
A or $A(t)$	$\frac{1-\gamma}{\gamma}\left(\frac{\beta}{1-\gamma}\right)^\gamma\left(\frac{1}{a}\right)^{1-\gamma}$	$\frac{1-\gamma}{\gamma}\left(\frac{\beta}{1-\gamma}\right)^\gamma\left(\frac{1+c_1}{a}\right)^{1-\gamma}$
B or $B(t)$	$\frac{1-\gamma}{r}\frac{\eta}{\beta}$	$\frac{1-\gamma}{r}\frac{\eta}{\beta}+\frac{1-\gamma}{\beta}c_2$
D or $D(t)$	$\frac{1-\gamma}{\gamma}\frac{\zeta}{\rho}$	$\frac{1-\gamma}{\gamma}\frac{\zeta}{\rho}+c_3$
Opt. Consumpt. $C^*(t)$	$a\left[W+\frac{1-\gamma}{r}\frac{\eta}{\beta}\right]-\frac{1-\gamma}{\beta}\eta$	$\frac{a\left[W+\frac{1-\gamma}{\beta}(\frac{\eta}{r}+c_2)\right]}{1+c_1}-\frac{1-\gamma}{\beta}\eta$
Portfolio $x_i^*(t)W$	$\frac{\left[W+\frac{1-\gamma}{r}\frac{\eta}{\beta}\right]}{1-\gamma}\sum_{j=1}^n\sigma^{ij}(\mu_j-r)$	$\frac{\left[W+\frac{1-\gamma}{\beta}(\frac{\eta}{r}+c_2)\right]}{1-\gamma}\sum_{j=1}^n\sigma^{ij}(\mu_j-r)$

As in (5.46), we assume that the derived utility function has the form

$$J(W,t) = e^{-\rho t}\left[A(t)(W+B(t))^\gamma + D(t)\right], \tag{5.48}$$

where $A(t)$, $B(t)$ and $D(t)$ are time dependent unknown parameters. Then, we obtain from (5.47)

$$\begin{aligned}
0 \;=\;& \frac{(1-\gamma)^2}{\gamma}\left(\frac{\beta}{\gamma}\right)^{\gamma/(1-\gamma)} A(t)^{\gamma/(\gamma-1)}(W+B(t))^\gamma \\
&+\frac{1-\gamma}{\gamma}\zeta + (rW + \frac{(1-\gamma)}{\beta}\eta)\gamma A(t)(W+B(t))^{\gamma-1} \\
&+\frac{1}{2}\frac{\gamma}{1-\gamma}A(t)(\mu - r\mathbf{1})'\Sigma^{-1}(\mu - r\mathbf{1})(W+B(t))^\gamma - \rho\left(A(t)(W+B(t))^\gamma + D(t)\right) \\
&+A'(t)(W+B(t))^\gamma + \gamma A(t)B'(t)(W+B(t))^{\gamma-1} + D'(t).
\end{aligned} \tag{5.49}$$

Table 5.1 is a list of optimal consumption, portfolio, and derived utility function for HARA type utility functions.

Here, we define the parameters as follows:

$$\begin{aligned}
a \;&\equiv\; \frac{\gamma}{1-\gamma}\left[\frac{\rho}{\gamma} - r - \frac{1}{2(1-\gamma)}(\mu - \mathbf{r1})'\mathbf{\Sigma}^{-1}(\mu - \mathbf{r1})\right]; \\
c_1 \;&\equiv\; \left[a\hat{\kappa}^{\frac{1}{1-\gamma}}\left(\frac{\hat{\beta}}{\beta}\right)^{\gamma/(1-\gamma)} - 1\right]e^{-a(T-t)}; \\
c_2 \;&\equiv\; \left[\frac{\beta}{\hat{\beta}}\hat{\eta} - \frac{\eta}{r}\right]e^{-r(T-t)}; \\
c_3 \;&\equiv\; \frac{1-\gamma}{\gamma}\left(\hat{\kappa}\hat{\zeta} - \frac{\zeta}{\rho}\right)e^{-\rho(T-t)}.
\end{aligned}$$

Coefficient a is a constant that is a time independent; we can see that coefficients c_1, c_2, and c_3 are dependent on the planned period. Hence , since $C^*(t)$ and W must be positive in regard to a finite planned period, we end up with

$$W \geq \frac{B}{r}\left(\frac{1}{a} - r\right).$$

Note, however, that when $B/r > 0$, then $1/a < r$, and contrariwise, when $B/r < 0$, then $1/a > r$.

For cases in which the planned period is finite, we get

$$W \geq \left(\frac{1-\gamma}{\beta}\right)\left[\frac{\zeta}{a}(1-e^{-a(T-t)}) - \frac{\zeta}{r}(1-e^{-r(T-t)}) + \zeta\hat{\kappa}\left(\frac{\hat{\beta}}{\beta}\right)^{\gamma/(1-\gamma)} e^{-a(T-t)} - \frac{\beta}{\hat{\beta}}\hat{\zeta}e^{-r(T-t)}\right].$$

From Table 5.1 the following is clear. In regard to an infinite planning horizon ($T \to \infty$), (1) the individual's derived utility $J(W,t)$ is the result of multiplying the linear of wealth $A(W+B)^\gamma + D$ by the discount rate $e^{-\rho t}$; (2) the parameters A, B, and D are constants that are not time dependent; (3) the optimal consumption $C^*(t)$ is one in which wealth W is proportional to a, after wealth W has been added to the coefficient B; (4) the optimal consumption is not time dependent; and (5) optimal consumption and the decision making on asset allocation are separated.

Setting

$$A = \frac{\left[W + \frac{1-\gamma}{\beta}(\frac{\eta}{r} + c_2)\right]}{1-\gamma}$$

for the case of a finite period, the optimal sum invested in a risky asset is given by

$$x_i W = A\sum_{j=1}^{n}\sigma^{ij}(\mu_j - r). \tag{5.50}$$

In this case, putting

$$A = \sum_{l=1}^{L}\frac{\left[W^l + \frac{1-\gamma^l}{\beta^l}(\frac{\eta^l}{r} + c_2^l)\right]}{1-\gamma^l} \quad \text{for investor } l,$$

equation (5.37) can be reduced to

$$\mu_i - r = \frac{1}{A}\sum_{j=1}^{n}\sigma_{ij}S_j,$$

which is an equation that the expected excess return should satisfy for each asset.

When a Bequest Utility Function Does Not Depend on the Amount of the Bequest

Here we shall see what happens when we seek the derived utility function that corresponds to the HARA type utility function, in those cases in which a bequest utility function regarding a finite planned period does not depend on the amount of the bequest. At such times, the conditions that guarantee that the bequest amount W is not negative and that from equation (5.43) the utility function of the bequest $B(W,T)$ does not depend on the amount of the bequest, are $\hat{\eta}=0$ and $\hat{\kappa}=0$. Under these conditions the derived utility function of the HARA type utility function when the amount of the bequest does not depend on the bequest utility function becomes

$$\lim_{\hat{\kappa}\to 0} J(W,t)\mid_{\hat{\eta}=0}=\frac{1-\gamma}{\gamma}$$

$$\left[\left(\frac{1-e^{-a(T-t)}}{a}\right)^{1-\gamma}\left(\frac{\beta}{1-\gamma}\left(W+\eta\frac{1-e^{-r(T-t)}}{r}\right)\right)^{\gamma}+\zeta\left(\frac{1-e^{-\rho(T-t)}}{\rho}\right)\right]e^{-\rho t}. \quad (5.51)$$

When this happens, the optimal consumption rule $C^*(t)$ in a finite planned becomes

$$\lim_{\hat{\kappa}\to 0} C^*(t)=\frac{a}{1-e^{-a(T-t)}}\left[W+\frac{(1-\gamma)\eta}{\beta}\left(\frac{1-e^{-r(T-t)}}{r}\right)\right]-\frac{(1-\gamma)\eta}{\beta}. \quad (5.52)$$

The optimal sum to be invested in risky asset i is given in

$$\lim_{\hat{\kappa}\to 0} x_i^*(t)W=\frac{\sum_{j=1}^{N}\sigma^{ij}(\mu_j-r)}{1-\gamma}\left[W+\frac{(1-\gamma)\eta}{\beta}\left(\frac{1-e^{-r(T-t)}}{r}\right)\right]. \quad (5.53)$$

From this it is easily seen that the utility of the bequest, regardless of the amount of the bequest, is

$$\lim_{\gamma\to 0} B(W(T),T)\mid_{\hat{\eta}=0}=0.$$

5.3.4 Conclusion

In this section we have discussed an optimal policy regarding consumption and portfolio selection when asset prices follow a semi-martingale. Then we derived an equation that the expected rate of return should satisfy when investors in the market have identical utility functions and agree on the parameters of stochastic processes that describe asset prices. Finally, when we gave utility function classes in concrete, we derived a closed solution for an optimal consumption and portfolio as well as the derived utility. What was learned from this study is that the analytical properties of the optimal policy regarding

consumption and portfolio selection are 1) it is robust after an asset price stochastic process; and 2) it takes the form of an addition of a discrete part to past results. We gave an equation similar to the intertemporal capital asset pricing model of Merton [73]. Instead of introducing, as Merton does, a state variable, we have expressed it as the supply of each asset.

Future research should derive a closed solution for an optimal policy regarding consumption and portfolio selection as an example of semi-martingale, but using examples other than those of geometric Brownian motion and similar stochastic processes, and develop an intertemporal capital asset pricing model in those cases. Also, even in the case of well-known stochastic processes, it would probably be interesting to carry out the same kind of analysis in regard to other classes of utility functions. Other future research tasks would include studying what happens when variables besides wealth are introduced as state variables in derived utility functions, or analyzing utility functions that are not additive in regard to time. Just so long as we rely on methods of dynamic programming in continuous times, we shall probably encounter the problem of a trade-off between the generalization of utility functions and asset prices, and the richness of conclusions that have been obtained in that area.

5.4 Optimal Exercise Policies for Call Options and Their Valuation

An American call option, simply called an option, is a right to buy a share of stock at any time during a stated interval for a stated price, the exercise price. Suppose that you own an option to buy one share of stock at a fixed exercise price, say c and you have n days to the maturity date. If you exercise the option on a day when the stock price is s and sell it in the open market, then your profit is s-c. The problem here is to find which strategy maximizes your expected profit. In other words, how should we choose a stopping rule in order to maximize our expected profit? Ross [89] and Taylor [119] study the stock option model under the setting of an optimal stopping problem in which price changes are independent identically distributed, that is, a random walk, and consequently the model has a single state. The model proposed by them leads to the unrealistic conclusion that for a fixed time the stock price is negative with a positive probability. To avoid this defect we modify their model so that stock prices are assumed to change according to the distribution of the product of independent positive random variables, which excludes the possibility of stock prices becoming negative. We also assume that the distribution of stock prices depends on the state of the economy which follows a Markov chain. In subection 5.4.1 we formulate the stock option model with multiple states as an optimal stopping problem. In subection 5.4.2 we show under some assumptions that there exists a simple optimal exercise policy which depends only on the current stock price and the state of the economy.Furthermore, properties of the optimal policy and its bounds are investigated in Subsection 5.4.3. Most results are distribution-free under the assumption that both the distribution of price changes and the state transition probability have a monotone property with respect to a state of the economy.

Finally, in Subsection 5.4.4 we propose a new, but simple, derivation of the Black and Scholes' option pricing formula with some concluding remarks.

5.4.1 Formulation of a stock option model

Let $\{1, 2, \cdots, N\}$ be the set of states of the economy and i or j denote one of these states. The economy changes according to a discrete time finite state Markov chain with a one step transition matrix $\{P_{ij}\}$. Let S_t be the stock price on the day t and suppose

$$S_{t+1} = S_t \cdot X_{t+1}^i = S_0 X_1^{i_1} X_2^{i_2} \cdots X_{t+1}^i$$

provided that the states of the economy from day 1 through $t+1, (i_1, i_2, \cdots, i)$ are observed, where $X_1, X_2, \cdots,$ are independent positive random variables with finite means.

Remarks 1 *If we have $S_{t+1} = S_t + X_{t+1}$ and X_t are independent and identically distributed, then the process S_t is a random walk which reduces to be the case of [89]. This random walk hypothesis leacs to the unrealistic conclusion that for a fixed day the stock price can be negative with a positive probability.*

Note that our model avoids this defect and that the price process in our model is a martingale if $E(X_{t+1}) = 1$ for all t, where E stands for the expectation operator. Consider now an option that entitles the holder to buy the stock at any time before the maturity date at a fixed exercise price, say c, regardless of what the market price might be. Let

T be the maturity date of the option. Suppose that we have already bought an option on day 1. If $S_t > c$ on day t, the option holder may exercise his option, buy the stock at the stated exercise price, and resell it in the market at the market price S_t, which gives him the profit $S_t - c$. If $S_t \le c$, no one exercises his option and no such profit is possible. Thus the expected profit ti the option holder is $E\{\max(S_t - c, 0)\}$. The problem is to find a stopping time $t*$ to maximize $E\{\max(S_t - c, 0)\}$ with respect t, $1 \le t \le T$. Let$F_i(\cdot)$ be the probability distribution of X_t^i. If we let $V_t(s, i)$ denote the maximum profit when the stock price is s, the state of the economy is i on day t, and the option has $(T - t)$ additional days to go , then from the principle of optimality $V_t(\cdot, \cdot)$ must satisfy

$$V_t(s, i) = \max\left\{s - c, \sum_{j=1}^N P_{ij} \int_0^\infty V_{t+1}(sx, j) dF_i(x)\right\} \tag{5.54}$$

with the boundary condition

$$V_T(s, i) = \max\{s - c, 0\} \tag{5.55}$$

on the maturity date. To establish an optimal exercise policy we need the following assumption. Assumption

(i) $F_1(x) \ge F_2(x) \ge \cdots \ge F_N(x)$ for all x.

(ii) For each k, $\sum_{j=k}^N P_{ij}$ is increasing in i.

Lemma 5.4.1 *(i) $V_t(s, i)$ is increasing, convex and continuous in s for each i, t.*

(ii) $V_t(s, i)$ is increasing in i and decreasing in t for each s.

Proof The proof is by induction on t. For $t = T$the statements (i) and (ii) certainly hold. Assume that $V_{t+1}(s, i)$ is increasing, convex and aontinuous in s for each i and is increasing in t for each s. Since $V_{t+1}(sx, j)$ is increasing, convex and continuous in s for each $x > 0$, so is $\sum_j P_{ij} \int_0^\infty V_{t+1}(sx, j) dF_i(x)$. Hence, $V_t(x, i)$ is increasing, convex and continuous in s for each i. On the other hand,

$$\begin{aligned}\sum_{j=1}^N P_{ij} \int_0^\infty V_{t+1}(sx, j) dF_i(x) &= \int_0^\infty \sum_{k=1}^N [V_{t+1}(sx, k) - V_{t+1}(sx, k-1)] dF_i(x) \sum_{j=k}^N P_{ij} \\ &\le \int_0^\infty \sum_{k=1}^N [V_{t+1}(sx, k) - V_{t+1}(sx, k-1)] dF_{i+1}(x) \sum_{j=K}^N P_{ij}\end{aligned}$$

$$\begin{aligned}
&\leq \int_0^\infty \sum_{k=1}^N [V_{t+1}(sx,k) - V_{t+1}(sx,k-1)]dF_{i+1}(x) \sum_{j=k}^N P_{i+1j} \\
&= \sum_{j=1}^N P_{i+1J} \int_0^\infty V_{t+1}(sx,j)dF_{i+1}(x),
\end{aligned}$$

where the first and second inequalities follow from Assumption (i), (ii), respectively, and $V_t(st,0) \equiv 0$. Therefore, we obtain

$$\begin{aligned}
V_t(s,i) &= \max\left\{s-c, \sum_j P_{ij} \int_0^\infty V_{t+}(sx,j)dF_i(x)\right\} \\
&\leq \max\left\{s-c, \sum_j \int_0^\infty V_{t+1}(sx,j)dF_{i+1}(x)\right\} \\
&= V_t(s,i+1).
\end{aligned}$$

which asserts that $V_t(s,i)$ is increasing in i for each s,t. That $V_t(s,i)$ is decreasing in t is immediately apparent from the fact that a higher value of t has the less chance of exercising options.

5.4.2 An optimal exercise policy

In this subsection we shall show under certain conditions that there is a simple optimal exercise policy which can be specified by the single value $s_t(i)$ at each day t with the state i, which in words says, do exercise the option if $s < s_t(i)$, do not exercise, otherwise. To establish this result we need the following lemma in addition to Lemma 5.4.1.

Lemma 5.4.2 *If $\mu_N \equiv \int_0^\infty x dF_N(x) \leq 1$, then $V_t(s,i) - s$ is decreasing in s for each i,t.*

Proof The proof is again by induction on t. For $t = T$ we have $V_T(s,i) - s = \max\{-c, -s\}$ which is plainly decreasing in s. Assume the assertion for $t+1$. Then, for t we have

$$V_t(s,i) - s = \max\left\{-c, \sum_j P_{ij} \int_0^\infty [V_{t+1}(sx,j) - sx]dF_i(x) + s(\mu_i - 1)\right\},$$

where μ_i is the mean of X_t^i. By the induction assumption for $t+1, V_{t+1}(sx,j) - sx$ is decreasing in s for each $x > 0$. Assumption (i) implies that $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_N$. If $\mu_N \leq 1$, then $\mu_i \leq 1$ for all i. Therefore, $s(\mu_i - 1)$ is decreasing in s. This completes the induction arguments.

For each t and i define

$$s_t(i) = \inf\{s : V_t(s,i) - s \leq -c\}, \tag{5.56}$$

where we take $s_t(i)$ to equal ∞ when this set is empty.

Theorem 5.4.1 *If $\mu_N \leq 1$, then there exits an optimal exercise policy as follows: If the stock price is s on day t and $s > s_t(i)$, then exercise the option, otherwise do not exercise.*

Proof If the stock price is s and the state is in i on day t, it is optimal to exercise the option if $V_t(s,i) \leq s - c$ because $V_t(s,i) \geq s - c$ from equation (1). Since for $\mu_N \leq 1, V_t(s,i) - s$ is decreasing in s for each i by Lemma 5.4.2, it follows that for all $s > s_t(i) V_t(s,i) - s \leq V_t(s_t(i),i) - s_t(i) \leq c$, which asserts that it is optimal to exercise the option when at price s and in state i on day t $s > s_t(i)$.

Theorem 5.4.2 *If $\mu_N \leq 1$, then $s_t(i)$ is increasing in i for each t and decreasing in t for each i.*

Proof From Lemma 5.4.1 (ii) $V_t(s,i)$ is increasing in i and from Lemma 5.4.2 $V_t(s,i) - s$ is decreasing in s. Hence, for each t

$$\begin{aligned}
s_t(i) &= \inf\{s : V_t(s,i) - s \leq c\} \\
&\leq \inf\{s : Vt(s,i+1) - s \leq -c\} \\
&= s_t(i+1).
\end{aligned}$$

Furthermore, from Lemma 5.4.1 (ii) $V_t(s,i)$ is decreasing in t, which implies that

$$\begin{aligned}
s_t(i) &= \inf\{s : V_t(s,i) - s \leq -c\} \\
&\geq \inf\{s : V_t(s,i) - s \leq -c\} \\
&= s_{t+1}(i)
\end{aligned}$$

Theorem 5.4.3 *If $(c/s)F_i(c/s) > 1 - \int_{c/s}^\infty x dF_i(x)$ for each i and s, then it is never optimal to exercise the option before the maturity.*

Proof At the maturity we have

$$\begin{aligned}
s_T(i) &= \inf\{s : V_T(s,i) - s \leq -c\} \\
&= inf\{s : \max\{s-c, 0\} - s \leq -c\} \\
&= c < \infty \quad \text{for each } i.
\end{aligned}$$

Therefore, the set $\{s : V_T(s,i) - s \le -c\}$ never becomes empty. Since $s_t(i)$ is decreasing in t from Theorem 5.4.2, we only have to show that $s_{T-1}(i) = \infty$ for each i. For $t = T-1$ we have

$$\begin{aligned} V_{T-1}(s,i) &= \max\left\{s-c, \sum_j P_{ij}\int_0^\infty V_T(sx,j)dF_i(x)\right\} \\ &= \max\left\{s-c, s\int_{c/s}^\infty xdF_i(x) - c\left[1-F_i\left(\frac{c}{s}\right)\right]\right\} \\ &= s-c+\max\left\{0, s\left(\int_{c/s}^\infty xdF_i(x)-1\right)+cF_i\left(\frac{c}{s}\right)\right\}. \end{aligned}$$

Since the quantity of the backet is positive, we obtain $V_{T-1}(s,i) > s-c$, which implies that $s_{T-1}(i) = \infty$ for each i.

Remarks 2

(i) Note that

$$\begin{aligned} s\left(\int_{c/s}^\infty xdF_i(x)-1\right)+cF_i\left(\frac{c}{s}\right) &= s\left(\int_0^\infty xdF_i(x)-1\right) - s\int_0^{c/s} xdF_i(x) \\ &\quad +cF_i\left(\frac{c}{s}\right) \\ &\ge s(\mu_i-1) - s\int_0^{c/s}\left(\frac{c}{s}\right)dF_i(x) \\ &\quad +cF_i\left(\frac{c}{s}\right) \\ &= s(\mu_i-1) \\ &\ge s(\mu_i-1) > 0, \quad \textit{for } \mu_1 > 1. \end{aligned}$$

This implies that $\mu_1 > 1$ is a sufficient condition for Theorem 5.4.3.

(ii) It is of interest from an investor's point of view to mention an implication of the optimal exercise policy established in Theorem 5.4.2 and 5.4.3. If an investor infers that the price of stock is expected to increase in the mean $\mu_1 > 1$, he should do nothing until the day of maturity. On that day he should make a purchase at the lower of the call price or the market price. If he expects $\mu_N \le 1$, follows the optimal exercise policy in Theorem 5.4.2 may give him the expected value $V_t(s,i)$. In words, should the price rise on the day with the stock prices $s < s_t(i)$ and the state i, he should exercise the option to buy at the stated price and immediately resell in the open stock market so as to receive a capital gain. Interpreting $V_1(s,i)$ as the value of the option at the begining of the first day when the stock price is s and the state is in i and also denoting the purchasing price of the option by V, we should buy the option when $V_1(s,i) > V$, and should not buy it, otherwise.

5.4.3 Properties of the optimal value and optimal policy

In this subsection we explore some analytical properties of an optimal exercise policy and its value. We write $P^1 > P^2$ whenever $\sum_{j=k}^N P_{ij}^1 \ge \sum_{j=k}^N P_{ij}^2$ for all k. To present the dependency of $P = [P_{ij}]$ on $V(\cdot,\cdot)$, we define

$$V_t(s,i,P) = \max\left\{s-c, \sum_j P_{ij}\int_0^\infty V_{t+1}(sx,j,P)dF_i(x)\right\}$$

and

$$s_t(i,P) = \inf\{s : V_t(s,i,P) - s \le -c\}.$$

Property 1 If $P^1 > P^2$, then we have

(i) $V_t(s,i,P^1) \ge V_t(s,i,P^2)$ for each s,i,t.

(ii) $s_t(i,P^1 \ge s_t(i,P^2)$ for each i,t.

Proof The proof is by induction on t. For $t = TV_t(s,i)$ is constant with respect to $P = [P_{ij}]$. So assertion (i) holds with equation. Assume for $t+1$ that $V_{t+1}(s,i,P^1) \ge V_{t+1}(s,i,P^2)$ for $P^1 > P^2$. Since $V_t(s,i)$ is increasing in i, by the first degree of stochastic dominance we can easily show that

$$\sum_j P_{ij}^1\int_0^\infty V_{t+1}(sx,j,P^1)dF_i(x) = \sum_j P_{ij}^2\int_0^\infty V_{t+1}(sx,j,P^2)dF_i(x)$$

which implies that $V_t(s,i,P^1) \ge V_t(s,i,P^2)$ for each s,i,t. Assertion (ii) immediately follows from assertion (i) because

$$\begin{aligned} s_t(i,P^1) &= \inf\{s : V_t(s,i,P^1) - s \le -c\} \\ &\ge \inf\{s : V_{t+1}(s,i,P^2) - s \le -c\} \\ &= s_t(i,P^2) \end{aligned}$$

Similarly, we use the notation $V_t(\cdot,\cdot,F)$ for emphasising the dependency on $F_i(\cdot)$.

Property 2 If $\mu_i^1 = \mu_i^2$ for all i and each i $\int_0^x F_i^1(y)dy \le \int_0^x F_i^2(y)dy$ for all x, then we have

(i) $V_t(s,i,F^1) \le V_t(s,i,F^2)$ for each s, i, t.

(ii) $s_t(i,F^1) \le s_t(i,F^2)$ for each i, t.

Proof The proof is again by induction on n. For $t = TV_T(\cdot,\cdot)$ and $s_T(\cdot)$ are both constant with respect to F. Since from Lemma 5.4.1 (i) $V_t(s,i,F)$ is increasing and convex in s, it can be shown that the second degree of stochastic dominance and the induction assumption for $t+1$ imply

$$\int_0^\infty V_{t+1}(sx,j,F^1)dF_i^1(x) \le \int_0^\infty V_{t+1}(sx,j,F^2)dF_i^2(x)$$

Therefore, we have $V_t(s,i,F^1) \le V_t(s,i,F^2)$. Then,

$$\begin{aligned} s_t(i,F^1) &= \inf\{s : V_t(s,i,F^1) - s \le -c\} \\ &\le \inf\{s : V_{t+1}(s,i,F^2) - s \le -c\} \\ &= s_t(i,F^2). \end{aligned}$$

5.4.4 An alternative derivation of the option pricing formula

In Subsection 5.4.3 we demonstrate in Theorem 5.4.3 that it is never optimal to exercise the option before maturity, provided that $(c/s)F_i(c/s) > 1 - \int_{c/s}^\infty x dF_i(x)$ for each i and s. Using this fact, we propose an alternative derivation of the Black and Scholes' option pricing formula.

Suppose that there is only one state, which allows us to eliminate the state variable i from our notation. Also, suppose that $\ln X_t$ is independently normally distributed with the mean μ and variance σ^2, that is, the stock price follows a geometric random walk. $E[X] > 1$ implies $\exp(\mu + \sigma^2/2) > 1$, which can be rewritten as $\mu > -\sigma^2/2$. If $E[X] > 1$, it has been shown from remarks for Theorem 5.4.3 that it is optimal never to exercise the option before maturity. Therefore, we have the maximum expected gain $V(s|c,T,\mu,\sigma)$ as follows:

$$\begin{aligned} V(s|c,T,\mu,\sigma) &= E[\max\{S_T - c, 0\}] \\ &= \int_c^\infty x dF_{S_T}(x) - c\int_c^\infty dF_{S_T}(x), \end{aligned} \quad (5.57)$$

where $F_{S_T}(\cdot)$ is the probability distribution of $S_T = s \cdot X_1 \cdot X_2 \cdots X_T$. Note that

$$\begin{aligned} F_{S_T}(x) &= Pr\{s \cdot X_1 \cdot X_2 \cdots X_T \le x\} \end{aligned}$$

$$\begin{aligned} &= \Pr\{\ln X_1 + \cdots + \ln X_T \le \ln \frac{x}{s}\} \\ &= \Phi\left\{\frac{\ln x/s - \mu T}{\sigma\sqrt{T}}\right\}, \end{aligned}$$

where $\Phi(\cdot)$ is the standard normal distribution. So, Equation (4) can be rewritten as follows:

$$\begin{aligned} V(s|c,T,\mu,\sigma) &= \int_{(\ln c/s - \mu T)/\sigma\sqrt{T}}^\infty s e^{\sigma\sqrt{T}z + \mu T} \frac{1}{\sqrt{2\pi}} e^{-z^2/2dz} + \int_{(\ln c/s - \mu T)/\sigma\sqrt{T}}^\infty d\Phi(z) \\ &= s e^{(\mu+\sigma^2/2)T}\Phi\left(\frac{\ln c/s + (\mu+\sigma^2)T}{\sigma\sqrt{T}}\right) - c\Phi\left(\frac{\ln c/s + \mu T}{\sigma\sqrt{T}}\right). \end{aligned} \quad (5.58)$$

It is easy to see that $\frac{\partial V}{\partial s} > 0$, $\frac{\partial V}{\partial c} < 0$, $\frac{\partial V}{\partial T} > 0$, $\frac{\partial V}{\partial \mu} > 0$, and $\frac{\partial V}{\partial \sigma} > 0$.

Now, we are ready to derive the Black and Scholes' option pricing formula [3]. Let r be the riskless rate of interest. Since each $\ln X_t$ is normally distributed with mean μ and variance σ^2, the expected value of the maturity price of the stock with the initial price s is $s \cdot \exp\{\mu + \sigma^2/2\}T$. On the other hand, if we invest x dollars today with the rate of interest r, we can expect to receive $s \cdot \exp\{rT\}$ at the maturity date T. For any risk neutral investor or from the no arbitrage condition, these two investment opportunities must be equivalent, that is,

$$s \cdot \exp\left\{\mu + \frac{\sigma^2}{2}\right\}T = s \cdot \exp\{rT\}. \quad (5.59)$$

Hence, we obtain a relation

$$\mu = r - \frac{\sigma^2}{2}. \quad (5.60)$$

Substituting (7) into (5) we have

$$V(s|c,T,\mu,\sigma) = s\Phi\left(\frac{\ln c/s + (r+\sigma^2/2)T}{\sigma\sqrt{T}}\right) - ce^{-rT}\Phi\left(\frac{\ln c/s + (r-\sigma^2/2)T}{\sigma\sqrt{T}}\right),$$

which is exactly the same equation as the Black and Scholes' option pricing formula. This implies that the model presented here includes the Black and Scholes' model as a special case which can be derived from a framework of optimal stopping problems. It is worthwhile to note that the expected value derived from an optimal exercise policy is coincident with the option pricing formula under the condition that the underlying stock has no divided with the mean rate of return $E[X] > 1$. This approach gives us a deeper understanding of the Black and Scholes' model.

Chapter 6

Two Software Reliability Growth Models Based Upon Module Structures

6.1 Introduction

There have proposed quite a few Software Reliability Growth Models (hereafter abbreviated by SRGM) to estimate, interpret and monitor failure behavior of large-scale software systems [41], [51], [132]. Those traditional models are formulated in a way of birds eye watch of target software. That is, the models handle the software as a black-box entity and interpret its defect behavior in the manner of macro. In other words, the traditional SRGM studied so far assumes failure occurrence and detection process as a stochastic or deterministic process.

Those traditional SRGM's have been applied to large-scale software developed through life-cycle processes. The SRGM's, however, have never applied to software developed in a modern process which is based on, for example, the object-oriented paradigm or data abstraction concept. In other words, it is an open question that the application of SRGM's to such type of software would success or fail (see [79], [82]).

The purpose of this chapter is to provide an approach of formulating a new SRGM by emphasizing module structure of software and to answer this open question. To consider the module structure, in this context, means to divide types of instruction in software into several types according to inter or intra module decomposition. In concrete, in the

process of model formulation, we first categorized instructions into multiple types such as data access, subprogram call, data flow via global data, etc, we, then, assume that the number of instruction executions is a stochastic counting process as well as the number of failures occurred. For the last decade, theoretical achievements in SRGM have been applied to the practice of software development. Data abstraction technique [64] and object-oriented design methods [18] are such example to provide some theoretical bases of the quality and measurement of software.

In this chapter we formulate a new SRGM which is much more general than the one appeared in the existing literatures and explains detect behavior of software. We also show that software developed by data abstraction techniques is more reliable than one developed by functional decomposition in terms of failure rates as well as the variance of the number of failure occurred. Those results are in closer agreements with actual software development practice. In Section 6.2 we formulate a new SRGM by considering module structures of software and detection rate of failures. In Section 6.3 we discuss a comparison of data abstraction software with functional decomposition one. Section 6.4 follows with some concluding remarks.

We have developed a new software reliability growth model based on a counting processes for instruction execution in a software. Through the analysis using the proposed model, we conclude that higher software reliability can be achieved with data abstraction techniques than with functional decomposition, under reasonable assumptions. We note also that the exponential NHPP model, which was developed through experience, is a special case of our theoretical model, and that the results of our model therefore agree with that of NHPP.

6.2 A Software Reliability Model Based Upon Module Structures and Error Detection Rates

In this section we formulate a new SRGM which explains detect behavior of software. Before doing so we need some notations and assumptions. The set of operations consists of functions and instructions in software. A set of operations or data is called a module. A software is defined as a set of such decomposed modules. Object-oriented software's modules consists of data and operations. On the other hand, functionally decomposed software consists only of operations. Any software carrys out jobs to communicate each other among modules .

Failure of software can occur either in communications between modules or in executing internal instructions within modules. Failure occurred within modules can be classified into two types due to access to either local or global data. We assume that multiple faults never turn out to be one failure.

We use the following notations :

$$X_n = \begin{cases} 1 & \text{if inter module communication occurs at } n^{\text{th}} \text{ instruction} \\ 0, & \text{otherwise,} \end{cases}$$

$$Y_n = \begin{cases} 1 & \text{if accessed to local data at } n^{\text{th}} \text{ instruction} \\ 0, & \text{otherwise,} \end{cases}$$

$$Z_n = \begin{cases} 1 & \text{if accessed to global data at } n^{\text{th}} \text{ instruction} \\ 0, & \text{otherwise.} \end{cases}$$

$N(t)$ = the number of instructions executed by time t.

Let t be a continuous time in a real line. So, $N(t)$ is a counting process whose sample path is non-decreasing in t. Let $X(t)$ be the number of intermodule communications by t. Let $Y(t)$ be the number of accesses to local data and $Z(t)$ the number of accesses to global data by t, respectively. Then, we have

$$X(t) = \sum_{i=1}^{N(t)} X_i, \quad Y(y) = \sum_{i=1}^{N(t)} Y_i, \text{ and } Z(t) = \sum_{i=1}^{N(t)} Z_i.$$

Assumption I $\{X_n\}$, and $\{Y_n\}$ and $\{Z_n\}$ are independently and identically distributed and also independent of $N(t)$, that is, $N(t)$ is a stopping time with respect to $\{X_n\}$, and $\{Y_n\}$ and $\{Z_n\}$.

Let p be the probability of intermodule communication, q the probability of accessing to local data and r the probability of accessing to global data at each instruction, respectively. So, we have

$$\begin{aligned} E[X_n] &= \Pr\{X_n = 1\} = p \\ E[Y_n] &= \Pr\{Y_n = 1\} = q \\ E[Z_n] &= \Pr\{Z_n = 1\} = r. \end{aligned}$$

Define

$$S(t) = X(t) + Y(t) + Z(t) \tag{6.1}$$

which presents the total sum of number of inter-communications and intra-instructions on the module decomposed software. Under *Assumption I*, applying Wald's formula into equation (6.1), we obtain

$$\begin{aligned} E[S(t)] &= E[X(t)] + E[Y(t)] + E[Z(t)] \\ &= (p+q+r)E[N(t)]. \end{aligned} \tag{6.2}$$

To count the number of failures on the software we define the following random variables :

$$O_j^1 = \begin{cases} 1 & \text{if a failure is observed at the } j\text{th inter module communication} \\ 0, & \text{otherwise,} \end{cases}$$

$$O_j^2 = \begin{cases} 1 & \text{if a failure is observed at the } j\text{th local data access} \\ 0, & \text{otherwise,} \end{cases}$$

and

$$O_j^3 = \begin{cases} 1 & \text{if a failure is observed at the } j\text{th global data access} \\ 0, & \text{otherwise.} \end{cases}$$

Define

$$\begin{aligned} \Pr\left\{O_j^1 = 1\right\} &= \alpha_1 \\ \Pr\left\{O_j^2 = 1\right\} &= \alpha_2 \\ \Pr\left\{O_j^3 = 1\right\} &= \alpha_3. \end{aligned}$$

Assumption II $\{O_j^1\}$, $\{O_j^2\}$ and $\{O_j^3\}$ are independently and identically distributed and also independent of $S(t)$, that is, $S(t)$ is a stopping time with respect to O_j^1, O_j^2 and O_j^3, respectively.

Denote $O(t)$ the total number of failures observed by time t, which is given by

$$O(t) = \sum_{j=1}^{S(t)} O_j^1 + \sum_{j=1}^{S(t)} O_j^2 + \sum_{j=1}^{S(t)} O_j^3. \tag{6.3}$$

Let $H(t)$ be the mean value function of $O(t)$. Applying Wald's formula into equation (6.3), again, we obtain

$$\begin{aligned} H(t) &= E[O(t)] \\ &= E[S(t)]\left\{E(O_j^1)+E(O_j^2)+E(O_j^3)\right\} \\ &= (p+q+r)E[N(t)](\alpha_1+\alpha_2+\alpha_3) \\ &= \alpha(p+q+r)E[N(t)], \end{aligned} \tag{6.4}$$

where $\alpha = \alpha_1 + \alpha_2 + \alpha_3$.

Defining $\lambda(t)$ a failure rate at time t, we have

$$\lambda(t) = \frac{dH(t)/dt}{a - H(t)} = \frac{\alpha(p+q+r)dE[N(t)]/dt}{a - H(t)} \tag{6.5}$$

where $a \equiv \lim_{t\to\infty} H(t) > 0$ the total number of faults embedded in the software.

To distinguish where the observed failures come from we put

$$O^1(t) = \sum_{i=1}^{S(t)} O_i^1, \quad O^2(t) = \sum_{i=1}^{S(t)} O_i^2, \text{ and } O^3(t) = \sum_{i=1}^{S(t)} O_i^3.$$

Let $H^1(t)$ be the mean value function of failures caused by inter-module communications, $H^2(t)$ the mean value function of failures by local data access and $H^3(t)$ by global data access.

$$\begin{aligned} H^1(t) &= \alpha_1(p+q+r)E[N(t)] \\ H^2(t) &= \alpha_2(p+q+r)E[N(t)] \\ H^3(t) &= \alpha_3(p+q+r)E[N(t)]. \end{aligned}$$

Let $\lambda^1(t)$, $\lambda^2(t)$ and $\lambda^3(t)$ be the failure rate corresponding to $H^1(t)$, $H^2(t)$ and $H^3(t)$, respectively.

$$\lambda_i(t) = \frac{dH^i(t)/dt}{a - H(t)}, \quad i = 1,2,3.$$

where $H(t)$ the reliability function of the entire software is given by $H(t) = H^1(t) + H^2(t) + H^3(t)$. The failure rate of the entire software is given by

$$\begin{aligned} \lambda(t) &= \frac{dH(t)/dt}{a - H(t)} \\ &= \frac{dH^1(t)/dt}{a - H(t)} + \frac{dH^2(t)/dt}{a - H(t)} + \frac{dH^3(t)/dt}{a - H(t)} \\ &\equiv \lambda_1(t) + \lambda_2(t) + \lambda_3(t). \end{aligned}$$

Let T_1, T_2 and T_3 be the inter failure times due to inter module communications, local data and global data accesses, respectively. Let $R(t)$ be the reliability function of the entire software, $R^1(t)$, $R^2(t)$ and $R^3(t)$ the reliability functions for T_1, T_2 and T_3, respectively. We have the relation of the reliability function and failure rates $\lambda(t)$, $\lambda_1(t)$, $\lambda_2(t)$, $\lambda_3(t)$ as follows :

$$\begin{aligned} R(t) &= \Pr\{\min(T_1,T_2,T_3) > t\} \\ &= \Pr\{T_1 > t, T_2 > t, T_3 > t\} \\ &= \Pr\{T_1 > t\} \cdot \Pr\{T_2 > 1\} \cdot \Pr\{T_3 > t\} \\ &= R^1(t) \cdot R^2(t) \cdot R^3(t) \\ &= \exp\left\{-\int_0^t \lambda_1(\tau)d\tau\right\} \cdot \exp\left\{-\int_0^t \lambda_2(\tau)d\tau\right\} \cdot \exp\left\{-\int_0^t \lambda_3(\tau)d\tau\right\} \\ &= \exp\left\{-\int_0^t [\lambda_1(\tau)+\lambda_2(\tau)+\lambda_3(\tau)]\,d\tau\right\} \\ &= \exp\left\{-\int_0^t \lambda(\tau)d\tau\right\}. \end{aligned} \tag{6.6}$$

Note that the software reliability function based on module structure is of series system. The reliability function is uniquely determined by and non-increasing in the failure rate.

6.3 A Comparison of Object Oriented Software with Functional Decomposition Software

In this section we describe functional decomposition and object oriented softwares. Defining their reliability functions or failure rates, we discuss a comparison between two softwares. It can be shown under some well-accepted assumptions that object-oriented software is better than functional decomposition one in a sence of software reliability growth models.

A functional decomposition software consists of the set of modules and global data, where each module shares global data with each other, and jobs can be carried out through intermodule communications and accesses to global data. Each module has no local data. On the other hand, a object oriented software consists only of the set of modules where each module posesses a sequence of operations and local data but does not have global data. Each module carries out instructions or statements within the module and communicates with other modules. Note that each module in the object-oriented software has a similar structure with a functional decomposition software. However, global data of functional decomposition software are not decomposed but shared with all modules.

One advantage of module partition in software design process can reduce the possibility of putting faults by partitioning modules in smaller size. Access to data in codes of intra-module in object-oriented software has a smaller probability of putting faults than in functional decomposition software. Under those considerations we assume that

(1) the object-oriented and functional decomposition software have both the same probability of putting a fault in the inter module communications,

(2) the object-oriented software has the smaller probability of putting a fault at the data access instruction,

(3) the functional decomposition software has a higher probability of accessing the global data and

(4) the objected-oriented software has a higher probability of inter module communications.

We use hereafter the subscripts o and f to present the object-oriented and functional decomposition softwares, respectively. Hence, $H_o(t)$ and $H_f(t)$ are the mean value functions for object-oriented and functional decomposition softwares. From equation (6.4) we have

$$\begin{aligned} H_o(t) &= \alpha_o(p_o + q_o + r_o)E[N(t)] \\ H_f(t) &= \alpha_f(p_f + q_f + r_f)E[N(t)]. \end{aligned}$$

Similarly, we have the corresponding failure rates

$$\begin{aligned} \lambda_o(t) &= \frac{dH_o(t)/dt}{a - H_o(t)} \\ \lambda_f(t) &= \frac{dH_f(t)/dt}{a - H_f(t)}. \end{aligned}$$

Proposition 1 If $\alpha_o/\alpha_f < (p_f + q_f + r_f)/(p_o + q_o + r_o)$, then $\lambda_o(t) < \lambda_f(t)$ for all t. Object-oriented software's reliability is higher than functional decomposition's.

Proof Taking the difference of the failure rates between two softwares, we obtain

$$\begin{aligned} \lambda_o(t) - \lambda_f(t) &= \frac{dH_o(t)/dt}{a - H_o(t)} - \frac{dH_f(t)/dt}{a - H_f(t)} \\ &= \left[\frac{\alpha_o(p_o + q_o + r_o)}{a - H_o(t)} - \frac{\alpha_f(p_f + q_f + r_f)}{a - H_f(t)}\right]\frac{dE[N(t)]}{dt} \\ &= \left[\frac{\alpha_o(p_o + q_o + r_o)}{a - \alpha_o(p_o + q_o + r_o)} - \frac{\alpha_f(p_f + q_f + r_f)}{a - \alpha_f(p_f + q_f + r_f)}\right]\frac{dE[N(t)]}{dt} \\ &= \frac{a}{K}\left[\alpha_o(p_o + q_o + r_o) - \alpha_f(p_f + q_f + r_f)\right]\frac{dE[N(t)]}{dt} \\ &< 0 \end{aligned}$$

where K is the common denominator appeared in the calculation and $K > 0$. Since $N(t)$ is a counting process, $dE[N(t)]/dt > 0$. By definition of reliability function $\lambda_o(t) < \lambda_f(t)$ implies $R_o(t) > R_f(t)$.

(*Q.E.D.*)

5ßNext, we shall compute the variance of the number of faults observed by time t, which is the risk measure of the reliability of software system. From equation (6.3) and *Assumption II* we obtain

$$\begin{aligned} Var[O(t)] &= Var\left[\sum_{i=1}^{S(t)}(O_i^1 + O_i^2 + O_i^3)\right] \\ &= Var\left[\sum_{i=1}^{S(t)} O_i^1\right] + Var\left[\sum_{i=1}^{S(t)} O_i^2\right] + Var\left[\sum_{i=1}^{S(t)} O_i^3\right] \\ &= E[S(t)]Var(O_i^1) + (E[O_i^1])^2 Var[S(t)] \\ &\quad + E[S(t)]Var(O_i^2) + (E[O_i^2])^2 Var[S(t)] \\ &\quad + E[S(t)]Var(O_i^3) + (E[O_i^3])^2 Var[S(t)] \\ &= E[S(t)]\left\{Var(O_i^1 + O_i^2 + O_i^3)\right\} \\ &\quad + Var[S(t)]\left\{(E[O_i^1])^2 + (E[O_i^2])^2 + (E[O_i^3])^2\right\} \end{aligned}$$

From *Assumption I, II* we have for $m = 1, 2, 3$

$$\begin{aligned} Var(O_i^m) &= E[(O_i^m)^2] - E[O_i^m]^2 \\ &= \alpha_m - (\alpha_m)^2 = \alpha_m(1 - \alpha_m). \end{aligned}$$

Hence,

$$E[S(t)] = (p + q + r)E[N(t)]$$

$$
\begin{aligned}
Var[S(t)] &= Var[X(t)+Y(t)+Z(t)] \\
&= Var\left[\sum_{i=1}^{N(t)} X_i\right] + Var\left[\sum_{i=1}^{N(t)} Y_i\right] + Var\left[\sum_{i=1}^{N(t)} Z_i\right] \\
&= E[N(t)]Var(X_i) + (E(X_i))^2 Var[N(t)] \\
&\quad + E[N(t)]Var(Y_i) + (E(Y_i))^2 Var[N(t)] \\
&\quad + E[N(t)]Var(Z_i) + (E(Z_i))^2 Var[N(t)] \\
&= E[N(t)]\{Var(X_i) + Var(Y_i) + Var(Z_i)\} \\
&\quad + Var[N(t)]\left\{(E[X_i])^2 + (E[Y_i])^2 + (E[Z_i])^2\right\} \\
&= E[N(t)]\left\{p - p^2 + q - q^2 + r - r^2\right\} \\
&\quad + Var[N(t)]\left\{p^2 + q^2 + r^2\right\}
\end{aligned}
$$

So, the variance of $O(t)$ can be written as

$$
\begin{aligned}
Var[O(t)] &= (p+q+r)E[N(t)]\{\alpha_1(1-\alpha_1) + \alpha_2(1-\alpha_2) + \alpha_3(1-\alpha_3)\} \\
&\quad + \{E[N(t)]\{p(1-p) + q(1-q) + r(1-r)\} \\
&\quad + Var[N(t)]\left\{p^2 + q^2 + r^2\right\}\} \cdot \left\{\alpha_1^2 + \alpha_2^2 + \alpha_3^2\right\}
\end{aligned}
$$

Putting $A \equiv \alpha_1^2 + \alpha_2^2 + \alpha_3^2$, $B \equiv p + q + r$, $C \equiv p^2 + q^2 + r^2$ and $\alpha = \alpha_1 + \alpha_2 + \alpha_3$, the equation above can be rewritten as

$$
Var[O(t)] = (\alpha B - Ac) \cdot E[N(t)] + AC \cdot Var[N(t)].
$$

For each m and i $0 < \alpha_i^m < 1$, $0 < p, q, r < 1$ and $\alpha B > AC$, which guarantees $Var[O(t)] > 0$.

Assumption III $Var[N(t)] \geq E[N(t)]$

Under *Assumption III* we obtain the following proposition which insists that object oriented software is better than functional decomposition software in a sense of the variance of the number of faults observed by time t.

Proposition 2 If $\alpha_o/\alpha_f < (p_f+q_f+r_f)/(p_o+q_o+r_o)$ and $(\alpha_{1f}^2+\alpha_{2f}^2+\alpha_{3f}^2)(p_f^2+q_f^2+r_f^2) < (\alpha_{1o}^2+\alpha_{2o}^2+\alpha_{3o}^2)(p_o^2+q_o^2+r_o^2)$ then

$$
Var_o[O(t)] \leq Var_f[O(t)]
$$

Proof Taking the difference of the variances between two softwares, we have

$$
\begin{aligned}
Var_o[O(t)] - Var_f[O(t)] &= E[N(t)](\alpha_o B_o - \alpha_f B_f) \\
&\quad + \{Var[O(t)] - E[N(t)]\}(A_f C_f - A_o C_o) \\
&< 0
\end{aligned}
$$

where $A \equiv \alpha_1^2+\alpha_2^2+\alpha_3^2$, $B \equiv p+q+r$, $C \equiv p^2+q^2+r^2$ and $\alpha_o B_o < \alpha_f B_f$, $A_f C_f < A_o C_o$ from the *assumption*.

(*Q.E.D.*)

Remarks 1 *If the testing procedure has been done by the equally qualified engineers, C_f almost equals to C_o. Then Proposition 2 holds for $A_o > A_f$.*

Remarks 2 *In this section software is treated as the set of decomposed modules and is classified to object-oriented and functional decomposition softwares. Executed instructions on the softwares can be considered to be carried out by inter module communications or accesses to data. Under those considerations we formulate a new software reliability growth model emphasizing upon its module structures.*

It has been shown under some assumptions that object-oriented software is better than functional decomposition software in the sense of their reliability functions as well as the variance of the number of faults observed. This result obtained here is in closer agreement with actual practice in software design and development processes. So, our model can provide some theoretical base to software engineers who have intuitively recognized this practical fact.

6.4 A New Software Reliability Growth Model Predicated on Counting Processes for Instruction Execution

If we develop an SRGM which represents internal structure of the software in detail, the values of various parameters estimated from the defect behavior naturally give measures of the goodness of structure of target software and even causes and types of faults can be estimated. With such a model, we can have not only the estimation of the total number of faults but also quantitative measure for the goodness of software structure with such a model. Therefore the model can be actively used to control the software development process [8].

Following the discussion above, we propose a new SRGM which represents structure of target software. Our model has, in concrete, the following features:

1. Execution of instruction is assumed to be a counting process.
2. Instructions in software are classified into several classes.
3. A set of parameters for instruction classes represents the structue of a target software.

In case instruction execution obeys Poisson process, let us preview the formula obtained in the next section for reliability of software at time t, $R(t)$:

$$R(t) = \prod_{i \in S} \exp[-\frac{p_i}{a + p_i} \lambda q_i t]$$

where

$S = \{G, L, C, O\}$, G, L, C, O denotes global, local, communication, and others, respectively,

a is the total number of faults,

p_i is the probability that failure occurs at the first execution of a class i instruction, proportional to the number of faults in class i instructions (we call p_i *the initial probability of failure*),

λ is the rate of instruction execution,

q_i is the probability that an instruction execution is of class i,

β is a constant which denotes failure detection probability.

In this section, we develop an SRGM for sequential-processing software in two steps:

- We derive a reliability function of the given time t for a single class of instructions, as in traditional models.
- We categorize the instructions into several classes to represent the structure of the code, and then extend the model to the general case.

Throughout this section, we assume that *a single failure is always caused by a single fault.*

That is, a failure never be caused by a combination of multiple faults.

The Basic Model

In general, an SRGM yields a continuous-time formula. Since failure can take place at any moment of time, it is natural to build our SRGM as a continuous-time model. First we focus on the series of time instants of instruction execution and derive a discrete-time formula. Then the discrete-time formula is translated into the corresponding continuous-time formula in the case that instruction execution is randomly distributed according to Poisson process.

Instruction Execution and Failure

Figure 6.1 illustrates the basic assumptions of the relationships between instruction execution, occurrence of failures, and detection of failure. Let us define random variables, M_i and O_i as the following:

$$O_i = \begin{cases} 1 & \text{if a failure occurs at the epoch } i; \\ 0 & \text{otherwise} \end{cases}$$

$$M_i = \begin{cases} 1 & \text{if a failure is detected at the epoch } i; \\ 0 & \text{otherwise} \end{cases}$$

where the epoch i is the time instance of the i-th instruction execution. α_i represents the probability of failure occurrence at the epoch i, and its decreases when the reliability grows. β is assumed to be a constant independent of time and represents failure-detection probability. The probability of failure occurrence is assumed only depends on time through the number of remaining failures.

The figure assumes the following scenario:

1. At epoch 2 and 3, failures occurred, ($M_2 = M_3 = 1$) but were not detected ($O_2 = O_3 = 0$).
2. At epoch 5, a failure also occurred ($M_5 = 1$) and was detected ($O_5 = 1$). The fault caused the failure was fixed and the total number of faults was reduced by one. Then α_6 reduced.
3. At epoch 6, a failure occurred and was detected (same as above).

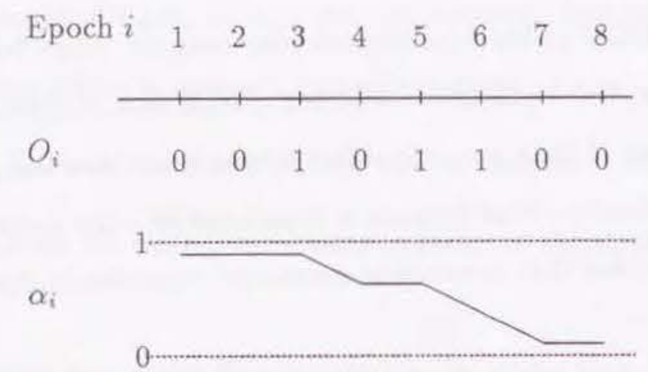


Figure 6.1: Relationship between Instruction Execution, Failure Occurrence, and Its Probability in Software

The Mean Value Function for Failures in Discrete Time

Let p be the initial probability of failure occurrence, that is,

$$\begin{aligned} p &= \Pr\{\text{One of faults in the code is executed}\} \\ &= \frac{(\text{the total number of faults})}{(\text{the total number of instructions})} \times (\text{constant}). \end{aligned}$$

Then, α_i is defined as:

$$\begin{aligned} \alpha_i &= \frac{p \cdot \sum_{m=0}^{\infty} \cdot \Pr\{\text{the total number of faults} = m\}}{(\text{the total number of faults})} \\ &= \frac{p \cdot E[\text{the number of remaining faults}]}{(\text{the total number of faults})} \\ &= \frac{(a - H_i)}{a} \cdot p \end{aligned} \tag{6.7}$$

where a is the total number of faults in software. The mean value function, H_i, is

$$H_i = E\,[\text{the number of faults detected at } i\text{-th event (fixed til } (i-1)\text{-th event)}]$$

To derive H_i, let $Q_i(k)$ be a probability mass function:

$$Q_i(k) = \Pr\{\text{the number of faults detected til } (i-1)\text{-th event} = k\}.$$

Then, the expectation of k becomes H_i:

$$H_i = \sum_{k=0}^{\infty} k \cdot Q_i(k) \quad (\text{the mean value with respect to } k).$$

Note that we have the following recurrent equation, $Q_i(k)$:

$$\begin{aligned} Q_i(k) &= \Pr\{(k \text{ at } (i-1)\text{-th event}) \wedge (\text{no new failure are detected at } (i-1)\text{-th event})\} \\ &\quad + \Pr\{(k-1 \text{ at } (i-1)\text{-th event}) \wedge (\text{a new failure is detected at } (i-1)\text{-th event})\} \\ &= \Pr\{\text{no new failure are detected at } (i-1)\text{-th event} \mid k \text{ at } (i-1)\text{-th event}\} \\ &\quad \times Q_{i-1}(k) \\ &\quad + \Pr\{[\text{a new failure is detected at } (i-1)\text{-th event}] \wedge [k-1 \text{ at } (i-1)\text{-th event}]\} \\ &\quad \times Q_{i-1}(k-1) \\ &= \{(1-\alpha_i) + \alpha_i(1-\beta)\} \cdot Q_{i-1}(k) + \alpha_i \beta Q_{i-1}(k-1), \end{aligned} \tag{6.8}$$

where $i = 1, 2, 3, \ldots,$ $k = 0, 1, 2, \ldots, a \wedge (i-1)$, and $Q_i(k) = 0$ otherwise. Let $Q_i^*(z)$ be a probability-generating function of $Q_i(k)$. That is,

$$Q_i^*(z) = \sum_{k=0}^{\infty} z^k Q_i(k).$$

Equation (6.8) becomes

$$\begin{aligned} Q_i^*(z) &= (1-\alpha_i)Q_{i-1}^*(z) + \alpha_i z Q_{i-1}^*(z) \\ &= \{(1-\alpha_i) + \alpha_i z\} Q_{i-1}^*(z) \\ &= \prod_{n=2}^{i} \{(1-\alpha_n) + \alpha_n z\} Q_1^*(z). \end{aligned}$$

Note that $Q_1(0) = 1$ and $Q_1^*(z) = 1$, then,

$$\begin{aligned} Q_1^*(z) &= 1 \\ Q_i^*(z) &= \prod_{n=2}^{i} (1 - \alpha_n + \alpha_n z). \end{aligned}$$

We have H_i in terms of α_n:

$$\begin{aligned} H_1 &= 0 \\ H_i &= \sum_{k=0}^{\infty} k \cdot Q_i(k) = \lim_{z \to 1} \frac{dQ_i^*(z)}{dz} \\ &= \lim_{z \to 1} \sum_{k=2}^{i} \{\alpha_k \prod_{m=2, m \neq k}^{i} (1 - \alpha_m + \alpha_m z)\} \\ &= \sum_{n=2}^{i} \alpha_n \quad (i \geq 2). \end{aligned}$$

From equation (6.7),

$$\begin{aligned}\alpha_1 &= p \\ \alpha_i &= p \cdot \frac{a - H_i}{a} \\ &= p - \frac{p \sum_{n=2}^{i} \alpha_n}{a} \quad (i \geq 2).\end{aligned}$$

Now we take the difference as:

$$\alpha_{i+1} - \alpha_i = -\frac{p}{a}\alpha_{i+1}$$

$$\alpha_{i+1} = \frac{1}{1 + \frac{p}{a}}\alpha_i.$$

Hence,

$$\alpha_i = \frac{p}{(1 + \frac{p}{a})^{i-1}}.$$

From (6.7), again:

$$H_i = a \cdot \{1 - \frac{1}{(1 + \frac{p}{a})^{i-1}}\} \quad (i \geq 2).$$

Since $\frac{p\beta}{a} > 0$,

$$\lim_{i \to \infty} \alpha_i = 0$$

$$\lim_{i \to \infty} H_i = a.$$

The derivation so far does not depend on any specific probability distribution of consecutive instruction executions.

The Continuous-Time Formula

If we assume that the executions of software instructions are carried out according to a Poisson process with intensity λ, we have a quite simple form for the reliability function. Since a Poisson process is known to represent random arrival of customers at a service facility in terms of queuing theory, this assumption is reasonable in most cases of instruction executions in software.

$$\Pr\{\text{the number of instructions executed in time interval } t = i\} = \frac{(\lambda t)^i}{i!}e^{-\lambda t}.$$

Thus, mean value function of failures at a given time t is

$$H(t) = \sum_{i=0}^{\infty} H_{i+1} \cdot \frac{(\lambda t)^i}{i!}e^{-\lambda t}.$$

Hence,

$$H(t) = a(1 - \exp[-\frac{p}{a+p}\lambda t]).$$

The failure rate defined by

$$d(t) = \frac{\frac{dH(t)}{dt}}{a - H(t)}.$$

can be obtained:

$$d(t) = \frac{p}{a+p}\lambda.$$

Finally, we have the reliability, $R(t) = \exp[-\int_0^t d(\tau)d\tau]$:

$$R(t) = \exp[-\frac{p}{a+p}\lambda t].$$

6.4.1 Extension of the Basic Model for Multiple Classes of Software Instructions

We now classify instructions into the following four categories:

- Global data access.
- Local data access.
- Inter-module communication.
- Others.

Here, a module is defined as a set of subprograms (procedures or functions) and/or data definitions. A software system consists of multiple number of modules. Inter-module communication is defined as follows:

> When an instruction in module M_1 is a call to a subprogram p in module M_2, inter-module communication occurs between M_1 and M_2, and is defined to be the set of actions that are the parameter passings of p's invocation, along with the return value in the case of a function call.

Let p_i be the initial probability of failure occurrence, and q_i be the probability of instruction execution, and suffixes such as G, L, C, O mean global data access, local data access, inter-module communication, and others, respectively. Say p_G denotes the initial probability of failure occurrence in global data access. If we assume that executions of the different types of instructions is independent of one another, the failure rate can be obtained:

$$d(t) = \sum_{i \in S} \frac{p_i}{a + p_i} \lambda q_i$$

where

$$S = \{G, L, C, O\}.$$

Reliability is defined as:

$$R(t) = \prod_{i \in S} \exp[-\frac{p_i}{a + p_i} \lambda q_i t].$$

6.4.2 Data Abstraction versus Functional Decomposition from the Perspective of Our Model

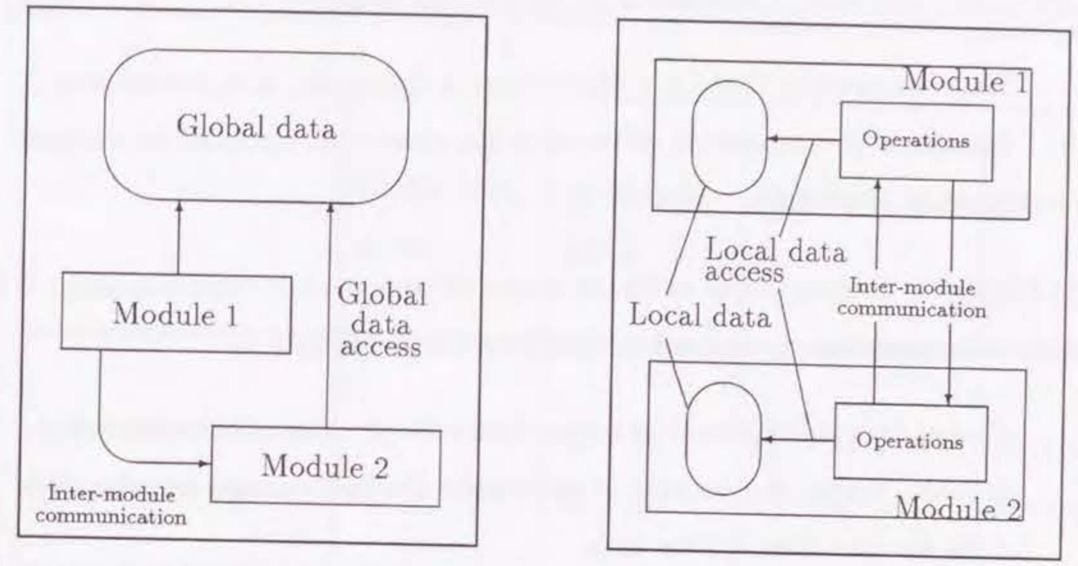
This section discusses the reliability of data abstraction software and functional decomposition software. Data abstraction software is the one which is designed by abstract data type technique [64] or an object-oriented design method such as Booch's [18]. Software developed by the design method of functional decomposition is called functional decomposition software. Sets of parameters for both types of software are defined according to the SRGM proposed in the previous section. Then, their reliabilities are compared with each other under several well-accepted assumptions.

6.4.3 Data Abstraction Software and Functional Decomposition Software

Structure of functional decomposition software and data abstraction software is illustrated in Figures 6.2 (a) and (b), respectively. As shown in the figure, functional decomposition software consists of a set of global data and modules that are sets of subprograms. Data abstraction software, on the other hand, is composed of modules which have internal local data and operations (procedures and functions) to access the internal data. No global data exist in data abstraction software.

Which is More Reliable?

Let p_i^j be the initial probability of failure occurrence and q_i^j be the probability of instruction execution, where $i \in S = \{G, L, C, O\}$ and $j \in T = \{d, f\}$. $G, L, C,$ and



(a) Functional Decomposition Software (b) Data Abstraction Software

Figure 6.2: Structure of Software

O denote global, local, communication and others, respectively. d denotes *data abstraction software* and f denote *functional decomposition software*. For example, p_G^f is the initial probability of failure occurrence at the time of global data access in functional decomposition software.

Let $d_d(t)$ and $d_f(t)$ mean failure rates of data abstraction software and functional decomposition software at time t. That is,

$$d_d(t) = \sum_{i \in S} \frac{p_i^d}{a + p_i^d} \lambda q_i^d$$

$$d_f(t) = \sum_{i \in S} \frac{p_i^f}{a + p_i^f} \lambda q_i^f.$$

Assumptions Made for Comparison

The following assumptions are made for comparison of these two types of software:

1. Assumptions on the probability of instruction execution

1a) Data abstraction software does not have global data in its components, which are accessed by multiple number of modules.

It is assumed that the data structure is completely hidden in data abstraction software.

1b) There is global data in functional decomposition software.

Since we assume that data abstraction is done only in a limited way in functional decomposition software, it has global data accessed by multiple number of modules.

1c) In functional decomposition software, a part of inter-module communication in data abstraction software is realized as data flow via global data. n

Global data in functional decomposition software, thus, can be categorized into two types: one because of incomplete abstraction; and and the other data for data flow via the data.

2. Assumptions on initial probability of failure occurrence.

2a) Initial probability is 0 if it is for instructions never executed ($p_G^d = 0$ as $q_G^d = 0$.)

2b) Initial probabilities of failure occurrences in local data access, inter module communication, and others are the same.

We assume that both types of software are designed and coded by programmers with the same level of skill. Hence, there is an equal chance to introduce faults in codes for those three types of instructions. As a result, all of their initial probabilities are equal.

Table 6.1 summarizes, based on the assumptions above, initial probability of failure occurrence and probability of each instruction. q_c in the table represents probability of flowing data via global data, q_L represents probability of accessing global data which are declared as a result of incomplete data abstraction in functional decomposition software.

Let Δd be the difference between the failure rates of data abstraction software and functional decomposition software, that is,

$$\begin{aligned}\Delta d &= d_d(t) - d_f(t) \\ &= \sum_{i \in S} \frac{p_i^d}{a + p_i^d} \lambda q_i^d - \sum_{i \in S} \frac{p_i^f}{a + p_i^f} \lambda q_i^f \\ &= \frac{p_L^d - p_G^f}{(a + p_L^d)(a + p_G^f)} a \lambda q_L + \frac{p_C^d - p_G^f}{(a + p_C^d)(a + p_G^f)} a \lambda q_C. \qquad (6.9)\end{aligned}$$

Table 6.1: Initial Probability of Failure Occurrence and Probability of Instruction Execution

	Global data access	Local data access	Communication	Others
Initial probability of failure occurrence	$p_G^d = 0$ p_G^f	$p_L^d = p_L^f$ $p_L^f = p_L^d$	$p_C^d = p_C^f$ $p_C^f = p_C^d$	$p_O^d = p_O^f$ $p_O^f = p_O^d$
Probability of instruction execution	$q_G^d = 0$ $q_G^f = q_C + q_L$	$q_L^d = q_L + q_L^f$ q_L^f	$q_C^d = q_C + q_C^f$ q_C^f	$q_O^d = q_O^f$ $q_O^f = q_O^d$

The denominator of each term of equation (6.9) is positive and the parameters a, λ, q_L and q_C also take positive value. Numerators, therefore, determine the sign of Δd. Due to the following reasons, the value of Δd is always negative.

1. $p_G^f > p_L^d$ because the possibility to introduce faults into instructions which access local data is smaller compared with codes for global data access. That is, abstracted data is decomposed into fine chunks compared with global data, and they are protected from illegal access from outside by operations. Hence, faults can be easily introduced into instructions which concern global data.

2. $p_G^f > p_C^d$ because faults introduction is, compared with inter-module communication, easier in codes for data flow via global data, which makes use of side-effects.

Based on the discussion above, we can objectively conclude

Data abstraction technique contributes to achieve highly reliable software.

6.4.4 Validation of Our Model

Our Model and the NHPP Model

Our SRGM can be coneived as a refinement of the exponential NHPP model [41] whose mean value function of failures is represented as:

$$H(t) = a(1 - e^{-bt})$$

when we define b as:

$$b = \frac{p}{a+p}\lambda.$$

That is, our SRGM explains failure behavior of software more precisely than the exponential NHPP model does. The exponential NHPP model is, from experience, from its application to real project, said to explain failure behavior of large-scale software. We can conclude, from this fact, and the discussion in Section 6.4.3, that:

> Data abstraction is a useful technique to enhance reliability of large-scale software.

How Can We Demonstrate Usefulness of Our Model?

The next questiob is that how we can demonstrate the usefulness os our SRGM other than the formal proof above. The most common and popular way to demonstrate usefulness of SRGM's is to

1. gather failure data in a real project,
2. estimate parameters of the SRGM, and
3. validate the difference between the estimated value of parameters (particularly the total number of faults) and an actual set of value observed.

So far, data from real projects (preferably from third party) detailed enough to be utilize in our SRGM is not available. Instead, we are now validating model by simulation. The steps in our simulation experiment is as follows:

1. Generate pseudon failure data based on Gonpertz curve.
2. Estimate parameters of our SRGM, using the data generated.

3. Validate the estimated total number of faults with the value which is set at the first step.

Chapter 7

Conclusion

This dissertation develops many stochastic dynamic optimization models involving sequential decisions. Some of these models require a special structure in order to be analysed by means of dynamic programming. In cases where this structure exists, dynamic programming is a very useful and practical technique for solving problems. In cases where uncertainty is involved in a control system described by a Markov process, we can apply dynamic programming to the control system. In earlier chapters, especially in chapters 3, 4, and 5, we present examples of stochastic dynamic models that analyse the dynamic aspect of system behaviour. However, the dynamic programming approach is both more difficult and at the same time simpler than other optimization approaches such as linear and non-linear optimization techniques. It is difficult to formulate certain optimizing models involving many decision variables and state descriptions because the computational requirements increase even though computer-aided search methods may be available.

In Chapter 2 we consider a class of dyamic programs in which there are distinguished subsets of policies and value functions, respectively called simple policies and simple value functions. An algorithm called generalized policy improvement is used to find ε-optimal policies. This algorithm has the property that only simple functions and policies are generated. When formulated as a dynamic program, it has an uncountable state space. However, the sets of simple policies and simple value functions can be chosen so that they are easily represented in a computer. As a special class of these simple dynamic programs we analyse piecewise linear dynamic programs and partially observable Markov decision processes. We also demonstrate how partially observable Markov processes may

be transformed into piecewise linear ones. Moreover, we specify how to find the simple policies and simple value functions.

In Chapter 3 we deal with problems of how many units of a particular product to produce each period and how to allocate price differentiable products between two types of demands. The main result of this chapter is to show the existence of a simple optimal policy even if fixed inventory costs are involved and even when the demands for the types of products are stochastically dependent. In addition, we provide several interesting examples in which demand distribution is specified. This inventory control model is one example of a sequentially interrelated decision that must be made over time.

In Chapter 4 we consider airline seat allocation between high and low fares with and without stochastic cancellations. Here we consider a dynamic airline seat allocation problem for a single flight with two fare classes. The problem is formulated as an N-step dynamic problem and aims at deriving optimal policies. We also explore some analytical properties of such an optimal seat allocation policy and the associated expected revenue. The model also extends the existing literature in two ways. First, it is a dynamic version with the cost of lost sales. Second, it is formulated under a setting of Markov decision processes which explicitly take into account the periods remaining until departure and permit reopening of fare classes. We also examine the problem of allocating airline seats between two nested fare classes when the demands for the classes are stochastically dependent. The well-known simple seat formula of Littlewood, which requires the assumption of statistical independence between demands, is generalized to a formula that requires only a much weaker monotonic association assumption. The model employed here is also used to examine the problems of full-fare passenger spillage and passenger upgrades from the discount class.

Chapter 5 develops an asset allocation model with various risk measures that is quite different from the mean-variance portfolio models. From the perspective of institutional investors the purpose of investing is to achieve a target level of rate of return that meets the cash flows of the business. A situation unfavorable to this aim is penalized as a risk. The model developed here is in closer agreement with actual practice in Japanese financial institutions. We discuss an optimal policy regarding consumption and portfolio selection when asset prices follow semi-martingales. Then we derive an equation that the expected rate of returns should satisfy when investors in the market have identical utility functions and agree on the parameters of stochastic processes that describe asset prices. Future research should derive a closed solution for an optimal policy regarding consumption and portfolio selection, but using examples other than those of geometric Brownian motion and similar stochastic processes, and developing an intertemporal capital asset pricing model in those cases. Other future research tasks would include studying what happens when variables besides wealth are introduced as state variables in derived utility functions. So long as we rely on methods of dynamic programming in continuous time, we shall probably encounter the problem of a trade-off between the wide perspectives gained by generalizing on utility functions and asset prices, and the richness of conclusions that can be obtained by a more specific model.

In Chapter 6 we develop a new SRGM by representing execution of software instructions as a counting process. Using the SRGM, we discuss whether or not the data abstraction technique contributes to enhance the reliability of software. The conclusion that *it really does* is formally drawn under well-accepted assumptions on software development. That is, we have *objectively* supported an instinctive apprehension good software developers have.

Future research includes the following topics: (1) Generalization of the model (in the process of our model formulation, described in Section 6.4, we assume that instruction execution is a Poisson process). A more generalized process such as a non-homogeneous Poisson process can be used to formulate a more generalized model. (2) Deriving a family of models (in the current form of our SRGM, α_i is to result a family of models based on our model formulation framework). This fact is a kind that various NHPP models can be derived by defining the mean value function of failures from one to another. Thus, our SRGM is a meta model whose parameter is α_i. (3) Demonstration of the usefulness of our SRGM by real project data.

This dissertation has been developed on the basis of an algorithmic method of dynamic programming in Chapter 2 and four areas of applications in chapters 3,4,5 and 6 using stochastic dynamic optimizing models. These applications are formulated in a form amenable to dynamic programming techniques. The formulation is used to derive certain simple properties of the expected value function, which then motivate the construction of

the computational algorithm.

Most of the applications in which these dynamic programming techniques are used are examples of a sequential decision problem in which the optimal policy has a special structural property that simplifies the procedure for constructing simple stationary policies. Of course, we should not emphasize only the analytical results discussed in this dissertation, but also the fact that we have developed the computational requirement of an algorithm that is solvable in practical applications. Finally, a most important area for future research should include the estimation of parameters for the stochastic dynamic models and should extend to include uncertainty about the stochastic process itself. The problem becomes more difficult, but it nevertheless remains important, when the stochastic process and the system dynamics come to include mutual interdependence and uncertainty in the process parameters.

Bibliography

References

[1] Aase, K., "Optimum Portfolio Diversification in a General Continuous-Time Model", *Stochastic Processes and their Applications*, 18, 81-98, (1984).

[2] Aase, K., "Contingent Claims Valuation When the Security Price is a Combination of an Ito Process and a Random Point Process", *Stochastic Processes and their Applications*, 28, 185-220, (1988).

[3] Aneja Yash and Noori, H.A., "The Optimality of (s, S) Policies for a Stochastic Inventory Problem with Proportional and Lump Sum Penalty Cost", *Management Science*, 33, 750-755, (1987).

[4] Aoki, M., *Optimization of Stochastic System*, Academic Press, New York, (1967).

[5] Åstrom, K.J., *Introduction to Stochastic Control Theory*, Academic Press, New York, (1970).

[6] Åstrom, K.J., "Optimal Control of Markov processes with Incomplete State Information", *J. Math. Anal. A*, 10, 174-205, (1965).

[7] Back, K., "Asset Pricing for General Process", *J. of Mathematical Economics*, 20, 371-395, (1991).

[8] Basili, V.R.: *Software Development: A Paradigm for the Future, Proc. COMPSAC '89*, 471-485, (1989).

[9] Beckmann, M.J.,"Decision and Team Problems in Airline Reservations," *Econometrica* 26, 134-145, (1958).

[10] Bellman R., *Introduction to the Mathematical Theory of Control Precesses*, Academic Press, New York, (1971).

[11] Bellman R., *Dynamic Programming*, Princeton University Press, Princeton, New Jersy, (1957).

[12] Belobaba, P.P., "Airline Yield Management : An Overview of Seat Inventory Control," *Transportation Science* 21 ; 2, 63-73, (1987).

[13] Belobaba, P.P., "Air Travel Demand and Airline Seat Inventory Management", *Ph.D. Dissertation*, Report R87-7, Flight Transportation Laboratory, Massachusetts Institute of Technology, Cambrige, Mass., (1987).

[14] Bhata, A.V. and Parekh, S.C., "Optimal Allocation of Seats by Fare", Presentation by Trans World Airlines to AGIFORS Reservations Study Group, (1973).

[15] Black, F. and Scholes, "The Pricing of Option and Corporate Liabilities", *J. of Polotical Economy* 81, 637-659, (1973).

[16] Bjork, T., Myhrman, J., and Persson, M., "Optimal Consumption with Stochastic Prices in Continuous Times", *J. of Appli. Prob.*, 24, 35-47, (1987).

[17] Blackwell, D., "Discrete Dynamic Programming", *Annals of Matimatical Statistics*, 33, 719-726, (1962).

[18] Booch, G., *Software Engineering with Ada*, 2nd ed. Benjamin/Cummings Pub. Co., Menlo Park, CA, (1987).

[19] Brumelle, S.L., "Generalized Policy Improvement for Monotone Dynamic Programs", *Working Paper* 503, Faculty of Commerce, University of British COlumbia, Vancouver, (1977).

[20] Brumelle, S.L. and Puterman, M.L., "On the Convergence of Newton's Method for Operators with Supports", *Operation Research Center Report ORC* 76-12, University of Californis, Berkeley, (1976).

[21] Brumelle, S.L., McGill, J.I., Oum, T.H., Tretheway, M.W. and Sawaki, K., "Allocation of Airline Seats Between Stochastically Dependent Demands," *Transportation Science* 24, 183-192 (1990).

[22] Brumelle, S.L. and Sawaki, K., "Generalized Policiy Improvement for Simple Dynamic Programs", *Working Paper* 546, Faculty of Commerce, University of British Columbia, Vancouver, (1978).

[23] Breeden, D.T., "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities", *J. of Financial Economics*, 7, 265-296, (1979).

[24] Chitgopekar, S., "Continuous Time Markov Sequential Control Processes", *SIAM J.Control*, 7, 367-389, (1989).

[25] Chow, Y.S., Robin, H. and Siegmund, D., *Great Expectations: The Theory of Optimal Stopping*, Houghton Mifflin Company, Boston, (1971).

[26] Cox, J. and Huang, C.F., "Optimal Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process", *J. of Economic Theory*, 49, 33-83, (1989).

[27] Cox, J., Ingersoll, J., and Ross, S., "An Intertemporal General Equilibrium Model of Asset Prices", *Econometrica*, 53, 363-384, (1985).

[28] Curry, R.E., "Optimum Airline Seat Allocation with Fare Classes Nested by Origins and Destinations", *Transportation Science* 24, 193-203, (1990).

[29] Davis, M.H.A. and Varaiya, P., "Dynamic Programming Conditions for Partially Observable Stochastic Systems", *SIAM J.Control*, 11, 226-261, (1973).

[30] DeGroot, M.H., *Optimal Statistical Decisions,* McGraw-Hill, New York, (1970).

[31] Denardo, E.V., "Contraction Mapping in the Theory Underlying Dynamic Programming", *SIAM Review*, 9, 165-177, (1967).

[32] Denardo, E.V. and Rothblum, U.G., "Affine Dynamic Programming", presented at the 1977 International Conference on Dynamic Programming, University of British Columbia, (1976).

[33] Derman, C., "On Sequential Decisions and Markov Chain", *Management Science*, 9, 16-24, (1962).

[34] Derman, C. and Sacks, J.,"Replacement of Periodically Inspected Equipment," *Naval Research Logistics Quarterly* 7, 597-607, (1960).

[35] Denardo, E.V., *Dynamic Programming*, Prentice-Hall, Englewood Cliffs,N.J., (1982).

[36] Detemple, J., and Murthy, S., "Intertemporal Asset Pricing with Heterogeneous Beliefs", *J. of Economic Theory*, 62, 294-320, (1994).

[37] Dothan, M.U., *Prices in Financial Markets*, Oxford University Press, New York, (1990).

[38] Duffie, D., and Skiadas, C., "Continuous-time Security Pricing: A Utility Gradient Approach", *J. of Mathematical Economics*, 23, 101-131, (1994).

[39] Dynkin, E. B., Controlled Random Sequences, *Theory of Probability and its applications*, X, 1-14, (1965).

[40] Fox, B. L., "Markov Renewal Programming by Linear Fractional Programming", *SIAM J. Appl. Math.*, 14, 1418-1432, (1966).

[41] Goel, A. L. and Okumoto, K.: *Time-Dependent Error Detection Rate Model for Software Reliability and Other Performance Measures, IEEE Trans. Reliability*, R-28, 206-211, (1979).

[42] Harmer, D.L., "A Description of the Surplus Seats Analysis System", The Consulting Division–Boeing Computer Services Inc., Seattle, Wash., (1976).

[43] Harrison, W. H. and Sweeney, P. F., *Good News, Bad News: Experience Building a Software Development Environment Using the Object-Oriented Paradigm, Proc. OOPSLA '89*, 85-94, (1989).

[44] Harrison, J., and Pliska, S., "Martingales and Stochastic Integral in the Theory of Continuous Trading", *Stochastic Processes and their Applications*, 11, 215-260, (1981).

[45] Hersh, M. and Ladany, S.P., "Optimal Seat Allocation for Flights with Intermediate Stops," *Computers and Operations Research* 5 ; 11,31-37, (1978).

[46] Hockstra, D., "Partially Observable Markov Decision Processes with Applications", *Technical Report* 156, Department of Operations Research, Stanford University, Stanford, September, (1973).

[47] Howard, R. A., *Dynamic programming and Markov Processes,* Wieley, New York, (1960).

[48] Ingersoll, J.E., Jr., *Theory of Financial Decision Making,* Totowa, NJ: Rowman & Littlefield, (1987).

[49] Ishigaki, T. and Sawaki, K., "On the (s, S) Policy with Fixed Inventory Holding and Shortage Costs", *J. Operations Research Society of Japan*, 34, 47-57, (1991).

[50] Jewell, W., "Markov Renewal Programming I and II", *Operation Research*, 2, 938-971, (1963).

[51] Jelinski, Z. and Moranda, P.: *Statistical Computer Performance Evaluation in Software Reliability Research*, Freiberger, W. Ed., Academic Press Inc., New York, (1972).

[52] Kakalik, J.S., "Optimum Policies for Partially Observable Markov Systems," *Technical Report TR-18,* Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts, (October 1965).

[53] Karatzas, I., "Optimization Problems in the Theory of Continuous Trading", *SIAM J. Control and Optimization*, 27, 1221-1259, (1989).

[54] Karatzas, I., and Shreve, S.E., *Brownian Motion and Stochastic Calculus*, 2nd, Springer-Verlag, New York, (1991).

[55] "Revenue Control: Mining Gold at the Margin", *Airline Executive*, 32-34, (January 1987).

[56] King, A.J., "Asymetric Risk Measures and Tracking Models for Portfolio Optimization", *Research Report*, IBM T.J. Watson Research Center, (1992).

[57] Kraft, D.J.H., Oum, T.H. and Tretheway, M.W., "Airline Seat Management", *Logistics and Transportation Review* 22:2, 115-130, (1986).

[58] Ladany, S., "Dynamic Operating Rules for Motel Reservation," *Decision Science* 7, 829-840, (1976).

[59] Konno, H. andYamazaki, H., "Mean-Absolute Deviation Portfolio Optimization Model and Its Applications to Tokyo Stock Market", *Management Science*, 37, 529-531, (1991).

[60] Lembersky, M. R., "On Maximal Rewards and ϵ-Optimal Policies in Continuous Time Markov Decision Chains", *Ann. of Statistics.*, 2, 159-169, (1974).

[61] Li, Yu-Ku, "Information Structure and Optimal Policy", Computer and Information Science Research Center, The Ohio State University, Ohio, (September 1970).

[62] Liberman, V. and Yechiali, U., "On the Hotel Overbooking Problems-An Inventory System with Stochastic Cancellations," *Management Science* 24, 1117-1126, (1978).

[63] Lippman, S. A., "Semi-Markov Decision Processes with Unbounded Rewards', *Management Science.*, 19, 717-731, (1973).

[64] Liskov, B. and Zilles, S., "Programming with Abstract Data Types", *SIGPLAN Notices*, 9, No. 4, 50-59, (1974).

[65] Littlewood, K., "Forecasting and Control of Passenger Bookings," *AGIFORS Symposium Proceeding* 12, 95-117, (1972).

[66] Lucas, R.E., Jr., "Asset Prices in an Exchange Economy", *Econometrica*, 46, 1429-1445, (1978).

[67] Maitra, A., "Discounted Dynamic Programming on Compact Metric Space", *Sankhya.*, 30, 211-216, (1968).

[68] Mayer, M., "Seat Allocation, or a Simple Model of Seat Allocation Via Sophisticated Ones", AGIFORS *Symp. Proc.* 16, 103-135, (1976).

[69] McGill, J.I., "Optimization and Estimation Problems in Airline Yield Management", *Ph.D. Dissertation*, University of British Columbia, Vancouver, B.C., Canada, (1989).

[70] Merton, R.C., "Lifetime Portfolio Selection under Uncertainty : The Continuous-Time Case", *Review of Economics and Statistics*, 51, 247-257, (1969).

[71] Merton, R.C., "Optimum Consumption and Portfolio Rules in a Continuous-Time Case", *J. of Economic Theory*, 3, 373-413, (1971).

[72] Merton, R.C., "An Intertemporal Capital Asset Pricing Model", *Econometrica,* 41, 867-887, (1973).

[73] Merton, R.C., *Continuous-Time Finance*, Rev. ed., Blackwall Pub., Massachuseets, (1992).

[74] Meyer, P.A., "Un cours sur les intégrales stochastiques", Séminaire de Probabilité X, *Lecture Notes in Mathematics*, 511, Springer-Verlag, (1976).

[75] Miller, B.L., "Finite State Continuous Time Markov Decision Processes with a finite planning Horizon", *SIAM J. Control.*, 6, 266-280, (1968).

[76] Mitten, L.G., "Composition Principles for Synthsis of Optimal Multistage Processes," *J. of Operations Research.,* 12, 610-619, (1964).

[77] Morton, R., "Optimal Control of Stationary Markov Processes", *Adv. Appl. Prob.,* 5, 18-19, (1973).

[78] Musa, D.J., "The Measurement and Management of Software Reliability: How Good Are They and How Can They Be Improved?", *IEEE Trans. Softw. Eng.*, SE-6, 489-500, (1980).

[79] Noro, M., Goto, K. and Sawaki, K., *A Software Reliability Growth Model Predicated on Instruction Execution* (in Japanese), *IPSJ SIG Notes*, 93-se-91, 41-48, (1993).

[80] Noro, M., Goto, K. and Sawaki, K., "A Software Reliability Growth Model Based on Module Structures", *Stochastic Models in Engineering, Technology and Management*, (ed. by S. Osaki and D.N. Murthy), World Scientific, (1993).

[81] Noro, M., Goto, K. and Sawaki, K., "A Software Reliability Growth Model Predicated on Counting Processes for Instruction Execution", 「情報処理学会論文誌」第３５巻２号, 2623-2630, (1994).

[82] Ohba, M., "Software Reliability Analysis Models", *IBM J. Res. & Dev.*, 28, 428-443, (1984).

[83] Oum, T.H.,and Tretheway, M.W., "Airline Seat Management", *Logist. Trans. Rev.* 22, 115-130, (1986).

[84] Pfeifer, P.E.,"The Airline Discount Fare Allocation Problem", *Decision Sci.* 20, 149-157, (1989).

[85] Puterman, M.L. and Brumelle, S.L., "On the Convergence of Policy Iteration in Stationary Dynamic Programming", *Working Paper* 392, Faculty of COmmerce, University of British Columbia, (1976).

[86] Richter, H., "The Differential Revenue Method to Determine Optimal Seat Allocations by Fare Type," *AGIFORS Symposium Proceeding* 22, 339-362, (1982).

[87] Rishel, R., "Necessary and Sufficient Dynamic Programming Conditions for Continuous Time Stochastic optimal Control", *SIAM J. on Control.,* 8, 559-571, (1970).

[88] Ross, S.M., Arbitrary State Markovian Decision Processes, *Annals of Mathematical Statistics.*, 39, 2118-2122, (1968).

[89] Ross, S.M., *Introduction to Stochastic Dynamic Programming*, Academic Press, New York, (1983).

[90] Ross, S.M., *Applied Probability Models with Optimization Applications*, Holden Day, 1970.

[91] Ross, S.M., "Average Cost Semi-Markov Decision Processes', *J. Applied Probability.,* 7, 649-656, (1970).

[92] Rothstein, M., "An Airline Overbooking Model," *Transportation Science* 5, 180-192, (1971).

[93] Rothstein,M., "OR and the Airline Overbooking Problem," *Operations Research* 32(2), 237-248, (1985).

[94] Samuelson, P.A., "Lifetime Portfolio Selection by Dynamic Stochastic Programming", *Review of Economics and Statistics*, 51, 239-246, (1969).

[95] Sawaki, K., "Partially Observable Markov Decision Processes with Abstract Spaces", アカデミア第III集（経済・経営学編５０）, 151-165, (1976).

[96] Sawaki, K., "An Analysis of Airline Seat Allocation," *J. of Operations Research Society of Japan* 32, No.4, 411-419, (1989).

[97] Sawaki, K., "Some Extentions of Dynamic Economic Lot Size Model with Backlogging", *J. Operations Research Society of Japan*, 14, 1-18, (1971).

[98] Sawaki, K., "Piecewise Linear Markov Decision Processes with Application", *Recent Development in Markov Decision Processes*, Academic Process, (1980).

[99] Sawaki, K., "Piecewise Linear Dynamic Programs with Applications", *J. Operations Research Society of Japan*, 23, 91-110, (1980).

[100] Sawaki, K., "Transformation of Partially Observable Markov Decision Processes into Piecewise Linear Ones", *J. Mathematical Analysis and Applications*, 91, 112-118, (1983).

[101] Sawaki, K., "On the Class of Closed Dynamic Programs", *J. Operations Research Society of Japan*, 27, 97-110, (1984).

[102] Sawaki, K., "On the (s, S) Policy with Fixed Inventory Holding and Shortage Costs", *J. Operations Research Society of Japan*, 34, 47-57, (1991).

[103] Sawaki, K., "Optimal Exercise Policies for Call Options and Their Valuation", *J. Computer and Mathematics with Applications*, 24, No. 1/2, 141-146, (1992).

[104] Sawaki, K., "A Dynamic Airline Seat Allocation Model", 「南山経営研究」第7巻3号, 415-424, (1993).

[105] Sawaki, K., "An Asset Allocation Model with Various Risk Measures", 「南山経営研究」第9巻2号, 585-594, (1994).

[106] Sawaki, K., "Inventory Control for Price Differentiable Products with No Carrying Over", *Stochastic Modelling in Innovative Manufacturing, Lecture Notes in Economics and Mathematical Systems*, Springer, (1997) (forthcoming).

[107] Sawaki K. and Ichikawa I., "Optimal Control for Partially Observable Markov Decision Processes Over an Infinite Horizon", *J. of Operations Research Society of Japan*, 21, 1-16, (1978).

[108] Sawaki, K. and Lin, T.J., "Optimal Portfolio Selection and Asset Pricing Models for Semi-Martiangle Processes", *Modern Portfolio Theory and Its Applications*, 45-64, Center for Academic Societies Japan, (1996).

[109] Sawaki, K. and Nishida, T., "On the Impact of Uncertainly in Dynamic Job Search", *J. of Operations Research Society of Japan*, 27, 32-42, (1984).

[110] Sawaragi, Y. and Yoshikawa, T., "Discrete-Time Markovian Decision Process with Incomplete State Observation", *Annals of Mathematical Statistic*, 41, 78-79, (1970).

[111] Scarf, H., "The Optimality of (s, S) Policies in the Dynamic Inventory Problem", Chapter 13, *Mathematical Method in the Social Science*, Stanford University Press, Stanford, (1960).

[112] Schweitzer, P. L., "Iterative Solution of the Functional Equations of Undiscounted Markov Renewal Programming", *J. Math Anal. Appl.*, 39, 495-501, (1971).

[113] Shlifer, R.and Vardi, Y., "An Airline Overbooking Policy," *Transportation Science* 9, 101-104, (1975).

[114] Smallwood, R. D. and Sondik, E. J., "The Optimal Control of Partially Observable Markov Processes over a Finite Horizon", *Operations Research.*, 21, 1071-1088, (1973).

[115] Sondik, E.J., *The Optimal Control of Partially Observable Markov Processes, Ph.D. Thesis*, Departement of Electrical Engineering, Stanford University, Stanfoord, (1971)

[116] Sondik, E.J., "The Optimal Control of Partially Observable Markov Processes over the Infinite Horizon: Discounted Costs", Department of Engineering Economic Systems, Stanford University, Stanford, (June 1973).

[117] Stone, L., "Necessary and Sufficient Conditions for Optimal Control of Semi-Markov Jump Processes", *SIAM J. Control.*, 11, 187-201, (1973).

[118] Strauch, R.E., "Negative Dynamic Programming", *Annals of Mathematical Statistics.*, 34, 871-890, (1966).

[119] Taylor, H.M., "Evaluating a Call Option and Optimal Timing Strategy in the Stock Market", *Manegement Science* 14, 111-120, (1967).

[120] Taylor, H.M., "Bounds for Stopped Partial Sums", *Annals of Mathematical Statistics* 43, 733-747, (1972).

[121] Titze, B. and Greisshaber, R.,"Realistic Passenger Booking Behaviours and the Simple Low Fare/High Fare Seat Allotment Model", *AGIFORS Symp. Proc.* 23, 197-223, (1983).

[122] Tretheway, M.W, "Frequent Flyer Programs: Marketing Bonanza or Anti-Competitive Tool?", *Proc. Can. Trans. Res. Forum.* 24,443-446, (1989).

[123] Veinott, A.F., "Discrete Dynamic Programming with Sensitive Discount Optimality Criteria", *Annals of Mathmatical Statistics.*, 40, 1635-1660, (1969).

[124] Veinott, A.F., Jr., "On the Optimality of (s, S) Inventory Policies: New Conditions and a New Proof", *SIAM J. on Applied Mathematics*, 14, 1067-1083, (1966).

[125] Veinott, A.F., Jr., "Optimal Policy for a Multi-product Dynamic, Non-stationary Inventory Problem", *Management Science*, 12 206-222, (1965).

[126] Veinott, A.F., Jr. and Wagner, H.M., "The Status Quo of Mathemathical Inventory Theory", *Management Science*, 12, 745-777, (1966).

[127] Wald, A., *Statistics Decision Functions*, John Wiley New York, (1950).

[128] "Yield Management Now Control Tactical Marketing", *Lloyd's Aviation Economist* 12-13, (May 1985).

[129] Tohma, Y. et al., "Structural Approach to the Estimation of the Number of Residental Software Faults Based on the Hyper-Geometric Distribution", *IEEE Trans. Softw. Eng.*, SE-15, 345-355, (1989).

[130] Walras, L., *Elements D'economie Politique Pure*, L. Corbaz, Lausane, (1874).

[131] White, C.C., "Procedure for the Solution of a Finite Problem, Partially Observed, Semi-Markov Optimazation Problem", *Operation Research*, 24, 348-358, (1976).

[132] Yamada, S. and Osaki, S., "Software Reliability Growth Modeling: Models and Applications", *IEEE Trans. Soft. Eng.*, SE-11, 1431-1437, (1985).

[133] Yamada, S., "Software Quality/Reliability Measurement and Assessment: Software Reliability Growth Model and Data Analysis", *J. Inf. Process.*, 14, 254-266, (1991).

[134] Yamada, S. and Takahashi, M.: *Introduction to Software Management Model* (in Japanese), 254, Kyoritsu-Shuppan, Tokyo, (1993).