Licenciatura Thesis

A Sequential Allocation Model with random opportunities

Isaías Alfredo Nhavane

Student at Eduardo Mondlane University
Mathematics and Computer Science Department
Maputo-Mozambique

Written under the supervision of **Prof. Dr. Dr. h. c. K. Hinderer** during a stay at the Institut für Mathematische Stochastik der Universität Karlsruhe, Germany.

To my parents.
To my wife.
To my daughters.

Every valley shall be exalted, and every mountain and hill
shall be made low: and the crooked shall be made straight,
and the rough places plain.

Isaiah, XL, 4

Acknowledgements

We are grateful to **Prof. Dr. Dr. h. c. Karl Hinderer**, for his valuable assistance during our work. His teaching and dedication were decisive factors for our success, and were a constant inspiration to us. Most of the new results and corresponding proofs contained in this work were obtained following his hints.

Our gratitude is extended also to **Priv. Doz. Dr. Dieter Kadelka** and **Dipl.- Math. oec. Alfred Müller** and **Mr. Markus Wagner** for their support, suggestions and aid throughout our activity.

We are grateful for the support obtained concerning the presentation and the English language. However, we are responsible for all errors which possibly remain.

The support given by the **Institut für Mathematische Stochastik der Universität Karlsruhe, Germany** as a whole should be mentioned and we are grateful for it.

This work was done under the financial support of the **Eduardo Mondlane University**, and for this many thanks from our side.

We are especially indebted to **Dipl.- Math. Mário Getimane**, he was originator of this project.

I. A. Nhavane

Abstract

The present work deals with an investment model where the investment opportunities are random.

We have initially K units of a resource (capital or fertilizer or energy, etc) available for investment. At certain times $\nu = 0, 1, \ldots, N-1$ an opportunity to invest will occur with probability p. As soon as an opportunity arises, we must decide how much of our available resource to invest.

If we invest a, then we obtain the profit $u(a)$, the amount a then becomes unavailable for further investment.

If no investment opportunity arises, we obtain c (if $c \geq 0$) or must pay $-c$ (if $c < 0$), e.g. as the management cost per time unit.

The problem is to decide how much to invest at each opportunity so as to maximize the total expected profit over the N time periods. N is called the horizon. This model has been investigated (with $c = 0$, $\beta = 1$ and $V_0 \equiv 0$) by Derman/Lieberman/Ross (1975), cited as D/L/R.

The resource may be measured in discrete units or in real numbers. In another model (called the "continuous-time case" by D/L/R) the investment opportunities arise at random times, namely according to a renewal process.

We now give a survey on the contents of the present work.

To introduce the reader in this field of mathematics we present in detail in the first chapter, what is dynamic programming and its intuitive background. Also we talk about the advantages of the use of (personal) computer in dynamic programming. The use of PC's played also an important role in the present work. We wrote numerous programs in TURBO-PASCAL VERSION 6.0 (cf. Appendix C) for illustration and control of theoretical results and for finding conjectures about the structure of the solution.

In the second chapter we collect some auxiliary results from mathematics and dynamic programming, without proofs. The purpose of this exposition is to be helpful in the main proofs.

Mainly the proofs use the induction principle, combined with the value

iteration.

We begin Chapter 3 with the discrete-time discrete-state case, where we present two models, one which is the model of D/L/R, (for the continuous state case), and another one which seems to be simpler. In Theorem 3.1 we present the basic solution procedure (value iteration and optimality criterion) for our model.

In Theorem 3.3 we show how both models are related. We also prove that our model and the model of D/L/R lead to the same result, cf. Lemma 3.1.

Afterwards we discuss structure properties of the value functions and of the optimal policy. The Theorems 3.4 and 3.5 show that two natural expectations in an investment process are fulfilled: (a) it is natural to expect that "if we invest more, we get more," (b) it is natural to expect that with increasing horizon we shall have increasing returns. The structure properties about concavity, convexity and monotonicity of the value functions and of the optimal policy are presented later in the continuous-state version.

The treatment of the last three cases uses only the first model.

For the discrete-time continuous-state version the respective model is described, and again the structure properties are discussed as in the case before, cf. Theorems 3.7, 3.8, 3.9, and 3.10.

If u and V_0 (terminal reward) are convex it is shown that if an opportunity presents itself we must invest all which we have at hand or nothing, and that V_n is convex. In the next Theorem, 3.10, it is shown that if u and V_0 are concave, in general there is not an explicit solution of the optimal policy but V_n is concave. In Theorem 3.11 the continuity of V_n in (s,p) is proved.

Case $p = 1$ is the classical case, where always an opportunity occurs. If $p \to 1$, then $V_n(s,p) \to V_n(s,1)$, this is a consequence of Theorem 3.11.

D/L/R give bounds of the value functions (Proposition 4). In Theorem 3.12 we give another bound for V_n.

While Theorem 3.10 does not yield the explicit solution of the optimal policy and the value functions, a special case where the optimal policy and the value functions can be completely specified is when $u(a) = a^\alpha$, for $0 < \alpha < 1$ and $V_0(s) = d_0 \cdot s^\alpha + e_0$, for some $d_0 \in \mathbb{R}_+$ and for some $e_0 \in \mathbb{R}$ (Example 1). Of particular interest is the case $\alpha = 0.5$ (Example 3).

Another special case is when $u(a) = \ln a$ and $V_0(s) = d_0 \cdot \ln s + e_0$, for some $d_0 \in \mathbb{R}^+$, and for some $e_0 \in \mathbb{R}$ (Example 2).

In all these examples we discuss the monotonicity and the convergence of the sequences (d_n) and (e_n).

In the renewal process case with discrete or continuous states we present the respective models and we discuss some results in D/L/R. We also prove some structural properties and state some conjectures we obtained from our program, for the case the times between opportunities have a geometric distribution (the discrete counterpart of Poisson process). The model for the renewal process case needs a two-dimensional state space.

Some of the previous results, e.g. closed solutions and structure properties, remain true even when the probability of an opportunity, the management cost, the utility function and the discounting change from period to period. In this non-stationary case we present the model for the discrete-time discrete and continuous-state. As before we present closed solutions for two special cases (Examples 1 and 2) and the convergence and the monotonicity of the sequences (d_n) and (e_n).

In a discussion of the reduction of problems from mathematical formulation to computer code, we explain with more details a flow chart for a general allocation process and a self-explanatory flow chart (cf. Appendix B) is used.

In the first appendix the notation used throughout the work is presented.

The new results in the present work concern the following topics:

(i) More realistic modelling by inclusion of a fixed cost c, and using an arbitrary function V_0 and an arbitrary β.

(ii) Use of simpler model than in D/L/R.

(iii) Continuity of V_n in (s,p), in particular if $p \to 1 \Rightarrow V_n(s,p) \to V_n(s,1)$.

(iv) Convergence and monotonicity of the sequences (d_n) and (e_n).

(v) Modelling the problem for the non-stationary case.

(vi) Modelling the renewal case with geometrically distributed times between opportunities.

Contents

Chapter 1

Introduction

1.1 Dynamic Programming

Dynamic Programming (DP) is a mathematical technique which is useful in many types of sequential decision problems.

What is dynamic programming ?

We will attempt to answer this question throughout this introduction.

Everybody can fill in with more details the following questions and can make some kind of answers to them. What is the best way to travel from where you live to where you work ? Or what is the best route to follow when you travel from home to any city ? Or what is the best itinerary for a vehicle that delivers food from a supplier to some shop ?

It is obvious that to answer such questions everyone will have difficulties, because: (1) the questions have not defined what is the meaning of "best", and (2) the questions do not tell what should be the form of the answer and how much details are needed. "Best" may have several meanings: least distance, least time, least cost, and usually this must be achieved under constraints.

We see that independently of the observations mentioned above, for answering those questions we must take decisions sequentially. Often, in order to achieve an objective, the decision process involves several decisions to be taken at different times. The mathematical technique of solving such a sequence of inter-related decision problems over a period of time is called *dynamic programming*.

Dynamic programming is a way of solving decision problems by finding an optimal strategy.

Dynamic programming uses recursion to solve complex problems, which can be subdivided into a series of sub-problems. The word *dynamic* is used because time is explicitly taken into consideration.

Dynamic programming differs from many of the techniques of Operations Research in that there is no universal algorithm which can be applied to all problems.

In dynamic programming there are deterministic and stochastic models. We now explain the deterministic model.

1.1.1 Deterministic Dynamic Programming Features

What are the characteristics of a deterministic **DP** problem ?

The principal objective of dynamic programming is to maximize (or minimize) a function of N variables and of a sequential structure subject to one or more constraints of a sequential structure. The number N is called the *horizon*.

We are now presenting in details the *intuitive background* for the N-stage deterministic dynamic programming problem $DP_N(s_0)$ with initial state s_0; cf. Figure 1.1.

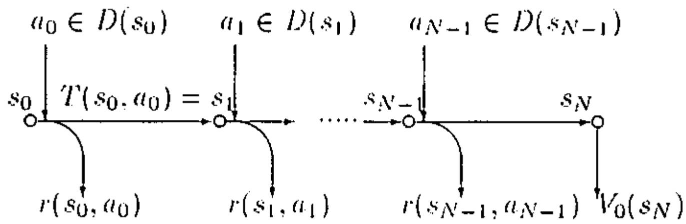


Figure 1.1. Development of states s_ν and actions a_ν.

The function to be maximized is defined as follows. Consider a system which starts at time $\nu = 0$ in some point s_0 of a set S. We call s_0 the *initial state* and S the *state space* . In the general theory, S is an arbitrary set. In our work S is either one-dimensional or two-dimensional (in Sections 3.3.4 and 3.3.5). The system moves at times $\nu = 1, 2, \ldots, N$ to states $s_1, \ldots, s_N$ under the influence of actions $a_0, a_1, \ldots, a_{N-1}$, taken from a set A, called the *action space*.

The interval between times ν and $\nu + 1$ is called ν-th period and "at stage n" means "at time $\nu = N - n$". When the system reaches state s, the next action is chosen from a set $D(s) \subset A$, the *set of admissible actions for state* s; and $D := \{(s, a) \in S \times A : a \in D(s)\}$ is called the *constraint set*. For the transition from one state to the next one we use the function $T : D \to S$, the so-called *transition function*. This means the following: if at time ν the system is in state s_ν and action $a_\nu \in D(s_\nu)$ is taken, then the system moves to the next state $s_{\nu+1} := T(s_\nu, a_\nu)$. At time ν a *one-stage reward* $r(s_\nu, a_\nu) \in \mathbb{R}$ is obtained (negative rewards are costs.)

Moreover, if the movement of the system ends at time N in state s_N, then a *terminal reward* $V_0(s_N)$ occurs. The same monetary units, obtained at different times, will have different cash value due to interest and/or devaluation. Thus we must take into consideration a so-called *discount factor* $\beta \in \mathbb{R}^+$. This means: the reward $r(s_\nu, a_\nu)$ obtained at time ν and the terminal reward $V_0(s_N)$ at time N enter the balance relative to time $\nu = 0$ as $\beta^\nu r(s_\nu, a_\nu)$ and $\beta^N V_0(s_N)$, respectively. In most applications early rewards are profitable which means $\beta < 1$.

For each state $s_0 \in S$ and each (admissible) action sequence $y = (a_\nu)_0^{N-1}$, we have to maximize the following function:

$$y \longrightarrow V_{Ny}(s_0) := \sum_{\nu=0}^{N-1} \beta^\nu r(s_\nu, a_\nu) + \beta^N V_0(s_N).$$

These sub-problems are then solved sequentially until the original problem is finally solved.

The N-stage value function $V_N : S \to (-\infty, +\infty]$ is defined by

$$V_N(s) := \sup\{V_{Ny}(s_0) : y \in A^N(s)\}.$$

The basic Theorem of deterministic **DP** (cf. Theorem 2.6) says, that the value functions can be computed recursively by the so-called *value iteration*

$$V_n(s) = \sup_{a \in D(s)} \{r(s,a) + \beta \cdot V_{n-1}(T(s,a))\}, \qquad s \in S, n \in \mathbb{N}.$$

A sequence $y^* = (a_\nu^*)_0^{N-1}$ of N decisions which achieve the optimum of the system, starting in s_0, is called an s_0-optimal action sequence.

A *decision rule* is a mapping $s \to f(s)$ such that $f(s) \in D(s)$ for all s. An N-*stage policy* is a sequence $\pi := (f_N, f_{N-1}, \ldots, f_1)$ of decision rules where f_k is used at stage k, which means k periods before the end of the decision process. $D_n^*(s)$ is the set of "optimal actions" at stage n in state s, i.e. the set of maximum points of $a \to r(s,a) + \beta \cdot V_{n-1}(T(s,a))$.

A *maximizer* at stage n is a decision rule f_n such that $f_n(s) \in D_n^*(s)$ for all s. The set of all decision rules is denoted by F, hence F^N is the set of N-stage policies.

1.1.2 Stochastic Dynamic Programming Features

From the many models of stochastic dynamic programming we only need the so called *control models with independent* (not necessary identically distributed) *disturbances*. For the moment assume that the disturbances are also identically distributed.

The case of an N-stage stochastic model DP_N is the one where the transition from s_ν to $s_{\nu+1}$ (where $s \in S$ is the *state* and S is the *state space*) is specified by a transition function T which is "disturbed" by some random variable $X_{\nu+1}$, taking in our applications values in a *finite set*, the disturbance space M, and where the variables $X_1, X_2, \cdots, X_N$ are i.i.d. (independent and identically distributed.) We put $X := X_1$.

As in the deterministic model, the *intuitive background* for such a model can be explained in details with a figure; cf. Figure 1.2.

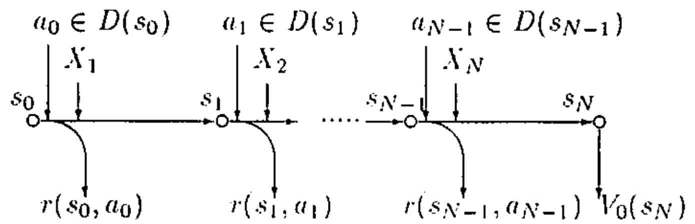


Figure 1.2. Development of states s_ν and actions a_ν, "disturbed" by $X_{\nu+1}$.

The transition law for the stochastic model has the following intuitive meaning: if at time ν we are in state s_ν and take action a_ν from the *set* $D(s_\nu) \subset A$, *the set of admissible actions at state* s_ν, then the system moves to the new random state $s_{\nu+1} := T(s_\nu, a_\nu, X_{\nu+1})$; $0 \le \nu \le N-1$. Thus T is defined on $D \times M$.

If at time ν, we are in state s_ν, take the action a_ν and if the disturbance $x_{\nu+1}$ occurs, then we obtain the reward $\hat{r}(s_\nu, a_\nu, x_{\nu+1})$. The *one-stage reward* is the expected value of $\hat{r}$, i.e. $r(s, a) := E\hat{r}(s, a, X)$.

The notions of *decision rule*, *N-stage policy*, *maximizer at stage* n, F and F^N are defined exactly as in the deterministic case.

For each initial state $s_0 \in S$ and each action policy $\pi = (\varphi_\nu)_0^{N-1}$, we have the N-stage *random* reward

$$R_{N\pi}(s_0) := \sum_{\nu=0}^{N-1} \beta^\nu \hat{r}(\xi_\nu, \varphi_\nu(\xi_\nu), X_{\nu+1}) + \beta^N V_0(\xi_N).$$

Here the state random variables ξ_ν are generated by $\xi_0 \equiv s_0, \pi$ and T as

$$\xi_{\nu+1} = T(\xi_\nu, \varphi_\nu(\xi_\nu), X_{\nu+1}), \quad 0 \le \nu \le N-1.$$

Moreover, the expected N-stage reward for initial state s and policy π is defined by the real number

$$V_{N\pi}(s) := ER_{N\pi}(s).$$

Consequently the N-stage value function $V_N : S \to (-\infty, +\infty]$ is defined by

$$V_N(s) := \sup_{a \in D(s)} V_{N\pi}(s),$$

which is the maximal expected N-stage reward for initial state s over the set F^N of all N-stage policies. We will show in Chapter 2, Theorem 2.6, that V_N can be computed recursively by the so-called **value iteration**

$$V_n(s) := \sup_{a \in D(s)} \{r(s,a) + \beta \cdot EV_{n-1}(T(s,a,X))\}, \qquad s \in S, n \in \mathbb{N}.$$

$D_n^*(s)$ is now the set of maximum points of

$$a \to W_n(s,a) := r(s,a) + \beta \cdot EV_{n-1}(T(s,a,X)).$$

1.1.3 The Inverse Principle of Optimality

In the literature very often the basic solution procedure of **DP** is said to be valid because of the following *Principle of Optimality* (Bellman and Dreyfus (1962), p. 15) which can be stated as follows:

> (a) (Deterministic problem) If the sequence $(a_\nu)_0^{N-1}$ is s_0-optimal for the N-stage problem and if $s_1 := T(s_0, a_0)$, then $(a_\nu)_1^{N-1}$ is s_1-optimal for the $(N-1)$-stage problem.
>
> (b) (Stochastic problem) If the policy $(f, \sigma) \in F \times F^{N-1}$ is s_0-optimal for the N-stage problem and if $s_1 := T(s_0, f(s_0))$, then σ is s_1-optimal for the $(N-1)$-stage problem.

The principle tells us that having chosen some initial s_0 and a_0, we do not examine all policies involving that particular choice of s_0 and a_0, but rather only those policies which are optimal for the $N-1$ stage problem, resulting from (s_0, a_0).

However, the principle of optimality is only a necessary condition for an N-stage optimal policy and hence of little use. On the other hand, the following *Inverse Principle of Optimality* (Hinderer (1993), supplement II.2.2) is a sufficient condition. It is in fact equivalent to the sufficiency part of the **Optimality Criterion**, and hence much more useful than the Principle of Optimality.

Inverse Principle of Optimality:

> If a^* is an optimal initial action for DP_N at state s_0 (i.e. if $a^* \in D_N^*(s_0)$) and if $(a_\nu^*)_1^{N-1}$ is an s_1-optimal action sequence for DP_{N-1}, where $s_1 := T(s_0, a_0)$, then $(a_\nu^*)_0^{N-1}$ is s_0-optimal for DP_N.

Based on this inverse principle of optimality, the solution begins by a one-stage problem and adds sequentially a series of one-stage problems that are solved until the overall optimum of the initial problem is obtained. The solution procedure is based on a *backward process* and a *forward process*.

In the first process, the problem of computing $V_n(s)$ and maximizers $s \to f_n(s)$ is solved by solving the problem for the last stage and working backwards towards the first stage, making optimal decisions at each stage of the problem. And in the second process, the problem of computing an s_0-optimal action sequence $(a_\nu^*)_0^{N-1}$ is solved by computing, using the maximizers from the backward process, recursively the actions a_ν^*.

The essential advantage of dynamic programming is that it transforms one (parametric) problem in N variables into N (parametric) problems, each in one variable.

1.2 General Reflections on the Use of (Personal) Computers in Dynamic Programming

Meanwhile, despite its theoretical and practical appeal, dynamic programming has not evolved into a standard methodology, primarily due to the lack of software specifically designed to support this technique and to the great diversity of problems.

While dynamic programming is a frequently used method in theoretical studies, it is so far among the main methods of Operations Research probably the least used methodology in computational and application efforts.

On the other hand, in many cases there do not exist solution methods other than **DP**, and then it is important to make as much as possible use of the structure of the solution in order to make dynamic programming a computationally efficient methodology.

1.2.1 Usage of (Personal) Computers in Dynamic Programming

The maximal expected reward $V_n(s)$ for the n-stage model with initial state s can be computed recursively by the value iteration VI (or DP algorithm):(cf. Chapter 2)

$$V_n(s) = \max_{a \in D(s)} \{r(s,a) + \beta \cdot EV_{n-1}(T(s,a,X))\}.$$

We use this VI for the stochastic model and also for the deterministic case. The latter is obtained when $|M| = 1$.

Obviously the solution of an N-stage DP_N problems consists of two parts:

1. the sequence of value functions $s \to V_n(s), n \in \mathbb{N}$, or at least V_N;

2. a sequence of maximizers f_n at stage $n, 1 \leq n \leq N$ (or more generally the sequence of sets $D_n^*(s)$ of optimal actions at stage n in the state s.)

We see that the recursive nature of VI lends itself easily to the implemention on computers. However, it is quickly realized that many problems of realistic size require huge memory and/or execution time. On the other hand, the VI can be accelerated if $a \rightarrow W_n(s, a)$ is convex (bang-bang maximizers) or concave, also if there exist monotone maximizers or maximizers of other simple structure.

The correctness of computer code for the **DP** can be controlled by theoretically obtained structural results. Also it is useful to retain structural properties when aggregation methods are used for approximate computations.

1.2.2 Advantages Due to the Use of Computer in Dynamic Programming

In general a **DP** problem has not an explicit solution. As far as hand computation is concerned, time and accuracy considerations usually rule out this method.

Once a digital computer with its enormous speed is available, numerical methods are important and assume a certain feasibility.

Computers are very important in particular because their graphic display of results helps to investigate the structural features of the solution, e.g. monotonocity, convexity, concavity, and so on. The structural properties comprise the following:

1. monotonicity of $V_n(s)$ in n and/or s;

2. concavity or convexity of $s \rightarrow V_n(s)$ and/or $a \rightarrow W_n(s, a)$;

3. monotonicity and/or Lipschitz-continuity of $s \rightarrow f_n(s)$;

4. sensitivity analysis, i.e. the influence of cost parameters or transition law parameters;

5. form of $D_n^*(s)$;

6. existence of limits of $V_n(s)$ and $f_n(s)$ $(s \rightarrow \infty, n \rightarrow \infty)$;

7. guess for speed of convergence, asymptotics;

8. support for finding a closed solution of f_n and/or V_n;

9. generation of counterexamples;

10. properties of "sufficiently" fine approximate models often carry over to continuous models, but not vice versa.

The support consists often in discovering, enhancing or disproving conjectures.

Numbers stored in the computer's memory are usually to ten or more significant decimal figures, depending upon what is desired. Consequently, there is feasibility and reliability to compare two values.

Chapter 2

Mathematics and Dynamic Programming Foundations

Throughout this chapter, we shall present some auxiliary results from mathematics and dynamic programming, without proofs. The justification of these Theorems, Lemmas, etc, can be found in the books of Hinderer, K., (1993), Roberts, A. W. and D. E. Varberg, (1973), and other books referenced in the bibliography.

2.1 Mathematical Background

2.1.1 Convex and Concave Functions

We assume that our functions $f : I \rightarrow \mathbb{R}$ are defined on some interval I of the real line $\mathbb{R}$.

Definition 2.1 *A function $f : I \rightarrow \mathbb{R}$ is called* **convex** *if*

$$f[\lambda \cdot x + (1 - \lambda) \cdot y] \leq \lambda \cdot f(x) + (1 - \lambda) \cdot f(y) \tag{2.1}$$

for all $x, y \in I$ and λ in the open interval $(0, 1)$. (One could equivalently take λ to be in the closed interval $[0, 1]$.)

The function f is called **strictly convex** provided that the inequality (2.1) is strict for $x \neq y$.

Definition 2.2 *A function $f : I \rightarrow \mathbb{R}$ is called* **concave** *if*

$$f[\lambda \cdot x + (1 - \lambda) \cdot y] \geq \lambda \cdot f(x) + (1 - \lambda) \cdot f(y) \tag{2.2}$$

for all $x, y \in I$ and λ in the open interval $(0, 1)$.

Also, f is called **strictly concave** provided that the inequality (2.2) is strict for $x \neq y$.

Observe that $-f$ is **convex**, iff f is **concave**. Therefore any result on convex functions yields immediately a "dual" result for concave functions, and vice versa.

Obviously a **linear function** $x \rightarrow c \cdot x + d$ is convex and concave.

Theorem 2.1 *If $f : I \rightarrow \mathbb{R}$ and $g : I \rightarrow \mathbb{R}$ are convex and $\alpha \geq 0$, then $f + g$ and $\alpha \cdot f$ are convex.*

Of course, it is also true that the sum of concave functions is a concave function and the product of a concave function with a non-negative number is a concave function. Moreover, any finite sum of convex [concave] functions is convex [concave].

Theorem 2.2 *If $f : I \rightarrow \mathbb{R}$ and $g : I \rightarrow \mathbb{R}$ are both non-negative, decreasing [increasing], and convex, then $x \rightarrow h(x) = f(x) \cdot g(x)$ also exhibits these three properties.*

Theorem 2.3 *Let $f_\alpha : I \rightarrow \mathbb{R}$, $\alpha \in B$, be an arbitrary family of convex functions and let $f(x) = \sup_\alpha f_\alpha(x)$. If $J = \{x \in I : f(x) < \infty\}$ is non-empty, then J is an interval and f is convex on J.*

In particular, the supremum of a family of linear functions is convex. Similarly if f_α is concave for all $\alpha \in B$, then $f(x) = \inf_\alpha f_\alpha(x)$ is concave on $\{x \in I : f(x) > -\infty\}$.

Theorem 2.4 *If D is convex and if $D(s)$ is bounded for all s, and if $W : D \rightarrow \mathbb{R}$ is concave, then $s \rightarrow W^*(s) := \sup_{a \in D(s)} W(s, a)$ is concave.*

The results about maxima and minima of convex and concave functions illustrate why convex and concave functions are specially interesting in the theory of dynamic programming.

Theorem 2.5 *If $f : I \rightarrow \mathbb{R}$ is convex, then any relative minimum of f in I is also a global minimum of f over I.*

More informations on the maximum or minimum of convex or concave functions are as follows:

Lemma 2.1 *If $f : [a, b] \rightarrow \mathbb{R}$ is convex, then it attains its maximum at the point a or at the point b.*

Theorem 2.6 *If $f : I \to \mathbb{R}$ is concave, then any relative maximum of f in I is also a global maximum of f over I.*

Lemma 2.2 *If $f : [a,b] \to \mathbb{R}$ is concave, then it attains its minimum at the point a or at the point b.*

Lemma 2.3 *If $f : I \to \mathbb{R}$ is strictly convex [strictly concave], it has at most one minimum [maximum] point.*

Lemma 2.4 *If v is concave on an interval $I \subset \mathbb{R}$, then $x \to v(x+h) - v(x)$ is decreasing for all $h > 0$.*

2.1.2 Other results from Mathematics

Lemma 2.5 (Linear first order difference equation) *If the sequence $(b_n)_0^\infty$ of real numbers satisfies*

$$b_n \leq [\geq] \quad c + \alpha \cdot b_{n-1}, \qquad n \in \mathbb{N},$$

for some real c and α, then

$$b_n \leq [\geq] \quad c \cdot \sigma_n(\alpha) + \alpha^n \cdot b_0, \qquad n \in \mathbb{N},$$

where

$$\sigma_n(x) := \sum_{\nu=0}^{n-1} x^\nu = \begin{cases} \frac{1-x^n}{1-x}, & x \in \mathbb{R} - \{1\}, \\ n, & x = 1. \end{cases}$$

2.2 Dynamic Programming Background

2.2.1 Stationary Control Model

Definition 2.3 *A* **(stationary)** *N-stage control model (CM) with finite disturbance space and i.i.d. disturbance variables $X_\nu, 1 \leq \nu \leq N$, is a tupel (S, A, D, T, M, p, r, V_0, β) of the following kind:*

(i) S is a non-empty arbitrary set, the **state space**.

(ii) A is a non-empty arbitrary set, the **action space**.

(iii) D is a subset of $S \times A$ such that all s-sections $D(s) := \{a \in A : (s,a) \in D\}$ of D are non-empty. D is called the **constraint set** *and $D(s)$ is called the* **set of admissible actions** *for state s.*

(iv) T is a mapping from D into S, the **transition function**.

(v) M is a non-empty finite set, the **disturbance space**.

(vi) p is the **discrete density** *of the disturbance random variables.*

(vii) r is a finite function on D, the (one-stage) **reward function**.

(viii) V_0 is a finite function on S, the **terminal reward function**.

(ix) β is a real positive number, the **discount factor**.

Theorem 2.7 (Basic Theorem) *(a) The value functions V_n satisfy the following recursion, called* **value iteration (VI)***:*

$$V_n(s) = \sup_{a \in D(s)} \{r(s,a) + \beta \cdot EV_{n-1}(T(s,a,X))\}.$$

(b) The following **Optimality Criterion (OC)** *holds: if $f_n(s)$ is a maximum point of $a \to W_n(s,a) := r(s,a) + \beta \cdot EV_{n-1}(T(s,a,X))$ for all $s \in S$ and $1 \leq n \leq N$, then the policy $(f_n)_N^1$ is optimal for DP_N.*

In 2.6 and in the following we put $X := X_1$. As the random variables X_ν are discrete, the VI reads, if $p(x) := P(X = x)$,

$$V_n(s) = \sup_{a \in D(s)} \{r(s,a) + \beta \cdot \sum_{x \in M} V_{n-1}(T(s,a,x)) \cdot p(x)\}.$$

In case $A \subset \mathbb{R}$ we say that D has *interval form* if

$$D(s) = [d_1(s), d_2(s)], \quad s \in S,$$

for two functions d_1 and $d_2 \geq d_1$ from S into A.

Theorem 2.8 *Assume*

(a) S is an interval in $\mathbb{R}$, $A \subset \mathbb{R}$.
(b) D has interval form for continuous functions d_1 and d_2.
(c) $\begin{cases} (s,a) \longrightarrow T(s,a,x) \text{ is continuous for all } x \in M, \\ D \longrightarrow S. \end{cases}$
(d) r and V_0 are bounded and continuous.
Then
(i) V_n is continuous and bounded for all $n \in \mathbb{N}$.
(ii) For each stage n there exists a smallest [largest] maximizer g_n [h_n].

Lemma 2.6 *If $s \to D(s)$ is increasing and $s \to W(s,a)$ is increasing for all a, then $s \to \sup_{a \in D(s)} \{W(s,a)\}$ is increasing.*

Theorem 2.9 *If $V_{k+1} \geq V_k$ for some $k \in \mathbb{N}_0$, then $n \to V_n(s)$ is increasing for all $s \in S$ and for $n \geq k$.*

For $a \in A$ the set $D_a := \{s \in S : (s,a) \in D\}$ is called the *a-section* of the constraint set D. As an example, if $S = A = \mathbb{R}_+$ and $D(s) = [0, s]$ for all s, then $D_a = [a, +\infty)$. In general, D_a may be empty for some a. If W is a function on D, the function $s \to W(s,a)$ is defined on D_a, provided that D_a is non-empty. A property of $s \to W(s,a)$ "for all a" is to be understood as "for all $a \in A$" such that $D_a \neq \emptyset$.

Theorem 2.10 *Assume $S \subset \mathbb{R}$. Then $s \to V_n(s)$ is increasing for all $n \in \mathbb{N}$, if*

(a) $D(\cdot)$ is increasing, i.e. $s \leq s'$ implies $D(s) \subset D(s')$.
(b) $s \to T(s, a, x)$ is increasing on D_a for all a and all $x \in M$.
(c) $s \to r(s, a)$ is increasing, and
(d) V_0 is increasing.

Remark
2.1. If $A \subset \mathbb{R}$ and if D has interval form, then $D(\cdot)$ is increasing, if d_2 is increasing and d_1 is decreasing.

Definition 2.4 *W has increasing differences iff $s \to W(s, a') - W(s, a)$ is increasing on $\{s \in D : (s, a) \in D, (s, a') \in D\}$, for all $a \leq a'$.*

Theorem 2.11 (Serfozo's criterion) *Assume $S \subset \mathbb{R}, A \subset \mathbb{R}, D$ has interval form with increasing d_1 and d_2 and that $W : D \to \mathbb{R}$ has increasing differences. If $W(s, \cdot)$ has a smallest [largest] maximum point $g(s)$ $[h(s)]$, then g is increasing [h is increasing].*

Lemma 2.7 *Assume $S \subset \mathbb{R}^n$, $A = \mathbb{R}$ and $D(s) := [0, d(s)]$ for a continuous function d from S into $\mathbb{R}$. If $W : D \to \mathbb{R}$ is continuous and bounded, then*

$$s \longrightarrow W^*(s) := \max_{0 \leq a \leq d(s)} W(s, a)$$

is continuous.

2.2.2 Non-Stationary Control Model

Definition 2.5 *A (non-stationary) N-stage control model (CM) with finite disturbance space and independent disturbance variables X_ν is a sequence of tupels $(S_n, A_n, D_n, T_n, M_n, p_n, r_n, \beta_n)$, $1 \leq n \leq N$, and S_0 and V_0 of the following kind:*

(i) $S_n, A_n, D_n, T_n, M_n, r_n, V_0, \beta_n$ have the same meaning and interpretation as in the stationary case, but they depend on n, which denotes stage n.

(ii) p_n is the **discrete density** *of the disturbance random variable at stage n.*

Note that the random variable $\xi_n := X_{N-n+1}$, which is the disturbance at stage n, has the discrete density p_n.

Theorem 2.12 (Basic Theorem) *(a) The value functions V_n satisfy the following recursion, called* **value iteration (VI)***:*

$$V_n(s) = \sup_{a \in D_n(s)} \{r_n(s,a) + \beta_n \cdot EV_{n-1}(T_n(s,a,\xi_n))\}, \quad s \in S_n, 1 \le n \le N.$$

(b) The following **Optimality Criterion (OC)** *holds: if $f_n(s)$ is a maximum point of $a \to W_n(s,a) := r_n(s,a) + \beta_n \cdot EV_{n-1}(T_n(s,a,\xi_n))$ for all $s \in S_n$ and $1 \le n \le N$, then the policy $(f_n)_N^1$ is optimal for DP_N.*

Remark
2.2. The expectation also depends on n, as in each stage there may be different discrete densities.

As the random variables X_n are discrete, the VI reads, if $p_n := P_n(X = x)$,

$$V_n(s) = \sup_{a \in D_n(s)} \{r_n(s,a) + \beta_n \cdot \sum_{x \in M_n} V_{n-1}(T_n(s,a,x)) \cdot p_n(x)\}.$$

Chapter 3

A Sequential Allocation Model with random opportunities

3.1 Allocation Problems

Problems of allocation arise whenever we can use a resource in a way to obtain a maximum possible profit. Suppose that there is available a certain quantity of an *economic resource*. This may represent money, machines, water for agricultural and industrial purposes or for hydroelectric power, fuel for a ship or plane, and so on. This resource can be used in different ways, using all or part of it in any way a certain return is derived. The fundamental problem is that of dividing at certain times the available resource so as to maximize the sum of returns. Another (non-temporal) interpretation is the allocation (at one time) to different activities.

3.2 A Stochastic Sequential Allocation Model

Sequential allocation problems are standard examples in Dynamic Programming.

The present work deals with a model where the investment opportunities are random.

This model has been introduced by Derman/Lieberman/Ross (1975), cited as D/L/R. The present work contains the complete proofs of main results of D/L/R (for a slightly more general model) and of some new results, new examples and numerous computations.

We have initially K units of capital available for investment. At times $\nu = 0, 1, \ldots, N-1$ an opportunity to invest will occur with probability p. The opportunities are assumed to be stochastically independent. As soon as an opportunity arises, we must decide how much of our available resource to invest.

If we invest a, then we obtain an (expected) profit $u(a)$, the amount a then becomes unavailable for investment.

If no investment opportunity arises, we obtain c (if $c \geq 0$) or must pay $-c$ (if $c < 0$), e.g. as management costs. We assume that these costs also arise after the resource s_ν becomes zero. A more realistic model would assume that no costs arise when s_ν becomes zero. But then the analysis of the model would become much more difficult. Moreover, in some cases, e.g. in Examples of $\ln a$, a^α (if $d_0 > 0$), and $h_1 \cdot \sqrt{a} + h_2 \cdot \sqrt{s-a} + h_3 \cdot \sqrt{s}$ (if $d_0 > 0$), one never reaches $s_\nu = 0$, and then our model is completely realistic.

If after N periods, s_N of the initial capital is left, we obtain a terminal reward $V_0(s_N) \geq 0$. In reality one will often have $V_0 \leq u$.

The problem is to decide how much to invest at each opportunity so as to maximize the total expected profit.

In order to facilitate the comparison with D/L/R, we give a translation table for notations.

D/L/R	Our notation
D	K
N	N
p	p
y	a
$P(y)$	$u(a)$
A	s (Model 1), t (Model 2)
	y (continuous-time)
$y_n(A)$	$f_n(s)$
$V(n, A)$	$V_n(s)$
$\hat{V}$	EV
t	t (continuous-time)

Table 3.1. Translation of notation.

The resource may be measured:

1. in discrete units *(discrete state case)*, or

2. in real numbers *(continuous state case)*.

The investment opportunities may arise:

1. at times $\nu = 0, 1, \ldots, N$ *(discrete time case)*, or

2. at random times $0 \leq T_1 \leq T_2 \leq \cdots$, e.g. according to a Poisson process *(renewal process case)*.

3.3 Stationary problems

Thus we shall treat four different cases.

3.3.1 The Discrete-Time Discrete-State Version

Here we present two different formulations of the problem as CM's: one, that corresponds in case $c = 0$ to the approach in D/L/R (1975), p. 1121-1122, (for the continuous state case), and a new one, which seems to be simpler. We begin with the latter, called **Model 1**:

$$S := \{0, 1, \ldots, K\}$$

is the *state space*, where K is the maximal (initial) capital for investment, and $s_\nu \in S$ is the remaining capital at time ν.

$$A := \{0, 1, \ldots, K\}$$

is the *action space*, where $a_\nu \in A$ is the capital to invest at time ν, ***in case an opportunity arises*** at that time.

$$D(s) := \{0, 1, \ldots, s\}$$

is the *set of admissible actions at state s.*

$$D := \{(s, a) \in S \times A : a \in D(s)\} = \{(s, a) \in S \times A : 0 \leq a \leq s\}$$

is the *constraint set.*

The i.i.d. disturbance random variables are

$$X_{\nu+1} := \begin{cases} 0, & \text{if no opportunity occurs at time } \nu, \\ 1, & \text{if an opportunity occurs at time } \nu, \quad \nu = 0, 1, \ldots, N-1. \end{cases}$$

Therefore, the *disturbance space* is $M = \{0, 1\}$. Put also $p := P(X_\nu = 1)$, $q := 1 - p$. We assume $0 < p \leq 1$.

Note that $p = 1$ corresponds to the classical case that always an opportunity occurs.

$$T : D \times M \longrightarrow S$$

is the *transition function*, given by

$$s_{\nu+1} = T(s_\nu, a_\nu, x_{\nu+1}) := \begin{cases} s_\nu - a_\nu, & \text{if } x_{\nu+1} = 1, \\ s_\nu, & \text{if } x_{\nu+1} = 0, \end{cases}$$

$$= s_\nu - a_\nu \cdot x_{\nu+1}.$$

As

$$\hat{r}(s,a,x) := \begin{cases} u(a), & \text{if } x = 1, \\ c, & \text{if } x = 0, \quad c \in \mathbb{R}, \end{cases}$$

the one-stage reward is

$$r(s,a) := E\hat{r}(s,a,X) = p \cdot \hat{r}(s,a,1) + q \cdot \hat{r}(s,a,0) = p \cdot u(a) + q \cdot c.$$

The terminal value function $V_0 \geq 0$ and the discount factor $\beta > 0$ are arbitrary.

A case where $V_0 \equiv 0$ is realistic, is a company who promotes some project and who requires that a resource remaining at time N must be returned. (Then V_n is the maximal expected utility for the project, not for the company.) In some other processes unallocated resources will have a value, and this value will be taken as $V_0(s)$.

We make the following *assumption*:

- V_0 is increasing.

This is a natural assumption as an increase in the terminal capital s_N will imply an increase of $V_0(s_N)$. On the other hand, u will not be increasing in all applications. For example in the use of fertilizers for agricultural purpose.

Remark

3.1. D/R/L (1975) use $c = 0$, $V_0 \equiv 0$ and $\beta = 1$, but it is mentioned in Remark (4), p. 1124, that (in the continuous state case) the results remain true for arbitrary β. In reality one will have $V_0 \not\equiv 0$ as the remaining resource s_N will not be worthless. (Note also that by chance it may happen that not a single investment opportunity arises during the time from $\nu = 0$ to $\nu = N - 1$.) A natural assumption may be $V_0 \leq u$.

Now let $V_n(s)$ denote the maximal expected n-stage reward, if the initial capital is s and *before it* is known whether or not at time $\nu = 0$ an opportunity arises. As the disturbance space M is finite, we obtain from Theorem 2.7 the following result:

Theorem 3.1 *(a) V_N may be computed recursively by the value iteration*

$$\begin{aligned} V_n(s) &= q \cdot c + \beta \cdot q \cdot V_{n-1}(s) + p \cdot \max_{0 \leq a \leq s} \{u(a) + \beta \cdot V_{n-1}(s-a)\} \\ &= q \cdot (c + \beta \cdot V_{n-1}(s)) + p \cdot \max_{0 \leq a \leq s} \{w_n(s,a)\}, \end{aligned} \tag{3.1}$$

where $w_n(s,a) := u(a) + \beta \cdot V_{n-1}(s-a)$.

(b) (**Optimality Criterion**) *If $f_n(s)$ is a maximum point of $a \to w_n(s,a)$ for all $s \in S$ and $1 \leq n \leq N$, then the policy $(f_n)_N^1$ is optimal for DP_N.* □

Let us now present **Model 2**:

In contrast to Model 1 we now indicate in the state s_ν whether or not an opportunity occurs at time ν. Thus our states are $s_\nu = (\varepsilon_\nu, t_\nu)$, where

$$\varepsilon_\nu = \begin{cases} 0, & \text{if no opportunity occurs at time } \nu, \\ 1, & \text{if an opportunity occurs at time } \nu, \end{cases}$$

and t_ν is the available capital at time ν.

Therefore, the *state space* is

$$S := \{0, 1\} \times \mathbb{N}_{0,K},$$

where K is the maximal initial capital.

$$A := \{0, 1, \ldots, K\}$$

is the *action space*, $a_\nu \in A$ is the capital to invest at time ν.

$$D(s) = D(\varepsilon, t) := \{a \in A : 0 \leq a \leq t\}$$

is the *set of admissible actions at state s*.

$$D := \{(s, a) \in S \times A : 0 \leq a \leq t\}$$

is the *constraint set*. Note that in case $\varepsilon = 0$ no action is required.

The i.i.d. disturbance random variables are

$$X_{\nu+1} := \begin{cases} 0, & \text{if no opportunity occurs at time } \nu + 1, \\ 1, & \text{if an opportunity occurs at time } \nu + 1, \end{cases} \quad \nu = 0, 1, \ldots, N-1,$$

and M is as in Model 1.

$$T(\varepsilon, t, a, x) := \begin{cases} (x, t - a), & \text{if } \varepsilon = 1, \\ (x, t), & \text{if } \varepsilon = 0, \end{cases}$$

$$= (x, t - a \cdot \varepsilon), \quad (\varepsilon, t, a) \in D,$$

is the *transition function*, and

$$r(\varepsilon, t, a) = \begin{cases} u(a), & \text{if } \varepsilon = 1, \\ c, & \text{if } \varepsilon = 0, \end{cases}$$

$$= \varepsilon \cdot u(a) + (1 - \varepsilon) \cdot c$$

is the one-stage reward.

The terminal reward function is denoted by $\check{V}_0$. As it does not matter for the terminal reward whether or not an opportunity occurs at time N, we define $\check{V}_0(\varepsilon, t) := V_0(t)$, where V_0 is the terminal reward function of Model 1.

As in Model 1 we assume that $\check{V}_0$ is increasing. Also $\beta > 0$.

Now let $\tilde{V}_n(\varepsilon, t)$ denote the maximal expected n-stage reward when the initial capital is t and when at time $\nu = 0$ no opportunity occurs ($\varepsilon = 0$) or an opportunity occurs ($\varepsilon = 1$).

Again we have a CM with finite disturbance space. Hence the following counterpart to Theorem 3.1 holds:

Theorem 3.2 *(a) $\tilde{V}_N$ may be computed recursively by the value iteration*

$$\begin{aligned} \tilde{V}_n(\varepsilon, t) &= (1-\varepsilon) \cdot c + \max_{0 \le a \le t} \{\varepsilon \cdot u(a) + \beta \cdot E\tilde{V}_{n-1}(X, t - a \cdot \varepsilon)\} \\ &= \max_{0 \le a \le t} \{\tilde{W}_n(\varepsilon, t, a)\}. \end{aligned} \tag{3.2}$$

*(b) (***Optimality Criterion***) If $f_n^*(s)$ is a maximum point of $a \to \tilde{W}_n(\varepsilon, t, a)$ for all $s = (\varepsilon, t) \in S$ and $1 \le n \le N$, then the policy $(f_n)_N^1$ is optimal for DP_N.* □

Using the abbreviations

$$v_n(t) := \tilde{V}_n(1, t),$$

$$w_n(t) := \tilde{V}_n(0, t),$$

the *value iteration* has the form:

$$v_n(t) = \max_{0 \le a \le t} \{u(a) + \beta \cdot p \cdot v_{n-1}(t-a) + \beta \cdot q \cdot w_{n-1}(t-a)\}, \tag{3.3}$$

$$w_n(t) = c + \beta \cdot p \cdot v_{n-1}(t) + \beta \cdot q \cdot w_{n-1}(t). \tag{3.4}$$

Note that $w_1(t) = c + \beta \cdot p \cdot v_0(t) + \beta \cdot q \cdot w_0(t)$.

It is intuitively clear, that $V_n(s)$ and $v_n(s)$, $w_n(s)$ are related as follows:

Theorem 3.3

$$V_n(s) \quad = \quad p \cdot v_n(s) + q \cdot w_n(s), \quad n \in \mathbb{N}_0. \tag{3.5}$$

Proof. For $n = 0$ the equality is obvious because $V_0(s) = v_0(s) = w_0(s)$.

Now we assume that the theorem holds for $n \le k$, for some $k \in \mathbb{N}_0$. Then for $n = k+1$, using (3.1) we have

$$V_{k+1}(s) \quad = \quad q \cdot c + \beta \cdot q \cdot V_k(s) + p \cdot \max_{0 \le a \le s} \{u(a) + \beta \cdot V_k(s-a)\}. \tag{3.6}$$

On the other hand, we have

$$\begin{aligned} p \cdot v_{k+1}(s) + q \cdot w_{k+1}(s) \quad &= \quad p \cdot \max_{0 \le a \le s} \{u(a) + \beta \cdot p \cdot v_k(s-a) + \beta \cdot q \cdot \\ &\cdot w_k(s-a)\} + \beta \cdot q \cdot (p \cdot v_k(s) + q \cdot w_k(s)) + q \cdot c. \end{aligned}$$

This may be written, using the induction assumption in (3.5) for $n = k$ as $q \cdot c + \beta \cdot q \cdot V_k(s) + p \cdot \max_{0 \le a \le s} \{u(a) + \beta \cdot V_k(s-a)\}$. We see that this is equal to (3.6), and in this way the theorem is proved. □

As consequence of this theorem, we have the following result:

Proposition 3.1

$$w_n(s) = \beta \cdot V_{n-1}(s) + c, \quad n \in \mathbb{N}. \tag{3.7}$$

Proof. Using Theorem 3.3, we can replace in (3.4) the expression $p \cdot v_{n-1}(s) + q \cdot w_{n-1}(s)$, by $V_{n-1}(s)$, and consequently $w_n(s) = \beta \cdot V_{n-1}(s) + c$. □

Remark

3.2. Numerically, (3.1) is easier than (3.3) and (3.4), as in (3.1) we have only one sequence of functions $(V_n)_1^N$.

In D/L/R, p. 1122, (formula (2)) the value iteration is written in a different form as in (3.3) and (3.4). That both formulae lead to the same result follows from:

Lemma 3.1 *For* $n \in \mathbb{N}$, *we have*

$$w_n(s) = c \cdot \sigma_n(\beta \cdot q) + \beta \cdot p \cdot \sum_{\nu=0}^{n-1} (\beta \cdot q)^\nu \cdot v_{n-\nu-1}(s) + (\beta \cdot q)^n \cdot V_0(s), \tag{3.8}$$

Proof. For $n = 1$, using (3.4) we have

$$\begin{aligned} w_1(s) &= c + \beta \cdot p \cdot v_0(s) + \beta \cdot q \cdot w_0(s) \\ &= c \cdot \sigma_1(\beta \cdot q) + \beta \cdot p \cdot v_0(s) + (\beta \cdot q)^1 \cdot V_0(s). \end{aligned}$$

Now we assume that (3.8) holds for $n \le k$, for $k \in \mathbb{N}$. Then for $n = k+1$, using (3.4) and the assumption above, we get

$$\begin{aligned} w_{k+1}(s) &= c + \beta \cdot p \cdot v_k(s) + \beta \cdot q \cdot w_k(s) \\ &= c + \beta \cdot p \cdot v_k(s) + \beta \cdot q \cdot (c \cdot \sigma_k(\beta \cdot q) + \beta \cdot p \cdot \sum_{\nu=0}^{k-1} (\beta \cdot q)^\nu \cdot v_{k-\nu-1}(s) \\ &\quad + (\beta \cdot q)^k \cdot V_0(s)) \\ &= c \cdot \left(1 + \sum_{\nu=0}^{k-1} (\beta \cdot q)^{\nu+1}\right) + \beta \cdot p \cdot \left(v_k(s) + \sum_{\nu=0}^{k-1} (\beta \cdot q)^{\nu+1} \cdot v_{k-\nu-1}(s)\right) \\ &\quad + (\beta \cdot q)^{k+1} \cdot V_0(s) \\ &= c \cdot \sum_{\nu=0}^{k} (\beta \cdot q)^\nu + \beta \cdot p \cdot \sum_{\nu=0}^{k} (\beta \cdot q)^\nu \cdot v_{k-\nu}(s) + (\beta \cdot q)^{k+1} \cdot V_0(s) \\ &= c \cdot \sigma_{k+1}(\beta \cdot q) + \beta \cdot p \cdot \sum_{\nu=0}^{k} (\beta \cdot q)^\nu \cdot v_{k-\nu}(s) + (\beta \cdot q)^{k+1} \cdot V_0(s). \end{aligned}$$

Therefore, (3.8) holds for $n = k + 1$, and thus the lemma holds for all $n \in \mathbb{N}$. □

However, the numerical solution by using (3.3) and (3.4) is easier than using (3.8) (which is essentially formula (2) in D/L/R): since by using (3.3) and (3.4), to obtain the solution (values of $V_n(s)$), we must store in the memory of the computer only the values of the last functions, while by using (3.8), we must store the values of all the functions $V_\nu(s)$ for $0 \leq \nu \leq n-1$.

The natural expectation in an investment process is that the n-stage maximal return will increase in s and n. It is logical that "if we invest more we get more." Also it is natural to expect that with time increasing we shall have increasing returns. The following theorems show that these natural expectations are fullfilled.

	s										
n	0	1	2	3	4	5	6	7	8	9	10
0	0.00	0.45	0.71	0.83	0.89	0.93	0.95	0.96	0.97	0.98	0.98
1	1.00	1.73	2.16	2.36	2.50	2.58	2.65	2.69	2.72	2.75	2.76
2	1.85	2.76	3.30	3.58	3.79	3.94	4.04	4.11	4.17	4.21	4.24
3	2.57	3.60	4.20	4.58	4.84	5.04	5.17	5.28	5.36	5.42	5.47
4	3.19	4.29	4.93	5.39	5.68	5.94	6.10	6.23	6.34	6.42	6.49
5	3.71	4.86	5.53	6.06	6.38	6.67	6.87	7.01	7.14	7.25	7.33
6	4.15	5.33	6.02	6.60	6.95	7.26	7.50	7.66	7.80	7.93	8.02
7	4.53	5.73	6.43	7.05	7.42	7.74	8.01	8.19	8.35	8.49	8.60
8	4.85	6.06	6.77	7.42	7.80	8.14	8.44	8.63	8.80	8.95	9.08
9	5.12	6.34	7.05	7.72	8.12	8.47	8.79	8.99	9.17	9.33	9.48
10	5.35	6.58	7.29	7.98	8.38	8.75	9.08	9.29	9.48	9.65	9. 80

Table 3.2. $V_n(s)$ for $u(a) = a/\sqrt{a^2+4}$, $p = 0.25$, $c = 1$, $K = 10$, $N = 10$, $\beta = 0.85$, and $V_0(s) = s/\sqrt{s^2+4}$.

Theorem 3.4 *$s \longrightarrow V_n(s)$ is increasing and finite for all $n \in \mathbb{N}$.*

Proof. We show that the assumptions in Theorem 2.10 are fulfilled:

a) If $s \leq s'$ then $D(s) \subset D(s')$, as $D(s) := \{0, \ldots, s\}$, if we choose $s' = s + 1 \Rightarrow D(s') := \{0, \ldots, s, s+1\}$, consequently $D(s) \subset D(s')$ and therefore $D(\cdot)$ is increasing.

b) $s \rightarrow T(s, a, x)$ is increasing for all a and x, as $T(s, a, x) := s - a \cdot x$.

c) $s \rightarrow r(s, a)$ is increasing for all a, as $r(s, a) = p \cdot u(a) + q \cdot c$ does not depend on s.

d) V_0 is increasing by assumption.

In this way the theorem is proved. □

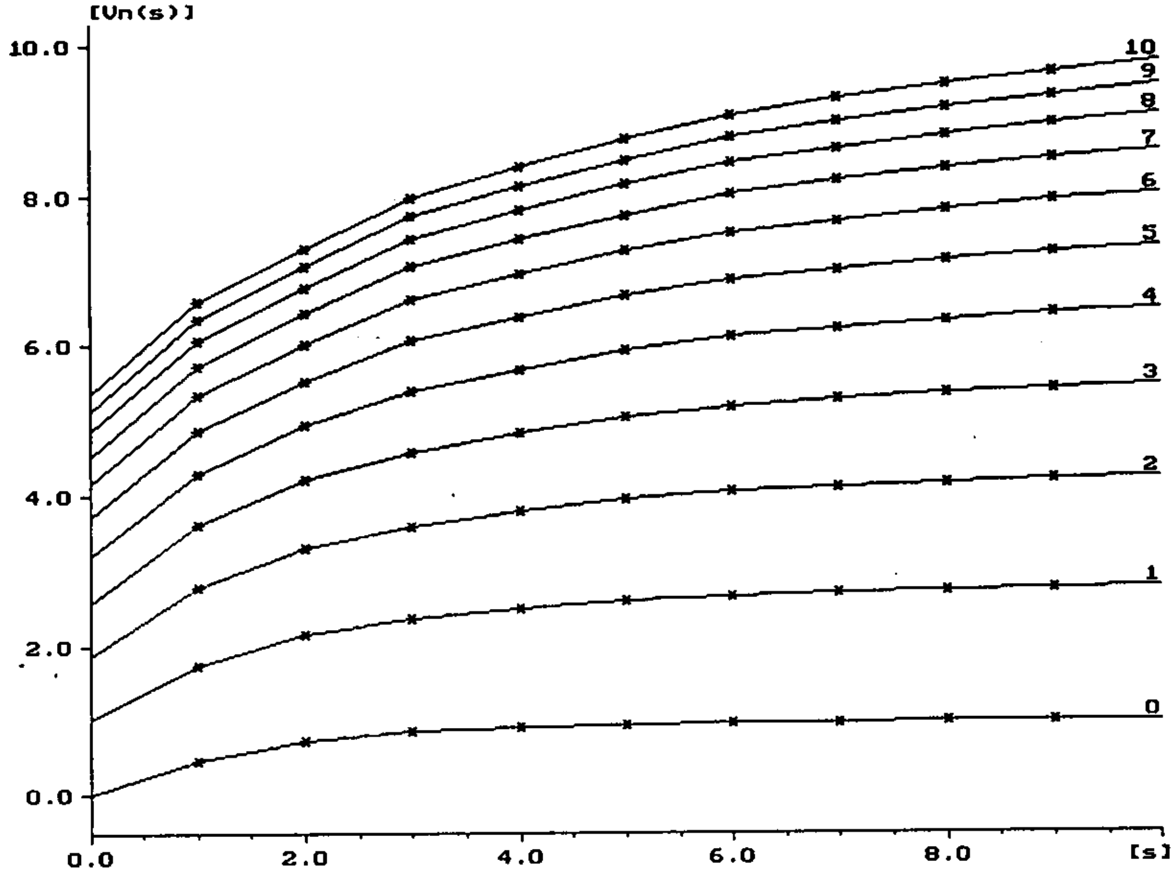


Figure 3.1. $s \to V_n(s)$ for $u(a) = a/\sqrt{a^2+4}$, $p = 0.25$, $c = 1$, $K = 10$, $N = 10$, $\beta = 0.85$, and $V_0(s) = s/\sqrt{s^2+4}$.

Theorem 3.5 *$V_n(s)$ is increasing in n for all s, if either*

(i)

$$\beta \le 1$$

and

$$V_0(s) \le \frac{p \cdot u(s) + q \cdot c}{1 - \beta \cdot q}, \qquad \text{for all } s,$$

or

(ii)

$$\beta = 1$$

and

$$c \geq -\frac{p}{q} \cdot u(0).$$

Proof. Using Theorem 2.9, it is sufficient to prove that $V_1 \geq V_0$. For $n = 1$, we have

$$V_1(s) = q \cdot c + \beta \cdot q \cdot V_0(s) + p \cdot \max_{0 \leq a \leq s} \{u(a) + \beta \cdot V_0(s - a)\}.$$

We get a lower bound for $V_1(s)$, choosing $a = 0$ or $a = s$. Therefore, using $a = s$ to prove the first assertion, we obtain

$$V_1(s) \geq q \cdot c + \beta \cdot q \cdot V_0(s) + p \cdot u(s) + p \cdot \beta \cdot V_0(0).$$

Thus $V_1 \geq V_0$ if

$$q \cdot c + q \cdot \beta \cdot V_0(s) + p \cdot u(s) + p \cdot \beta \cdot V_0(0) \geq V_0(s).$$

From the last inequality, we get

$$V_0(s) \leq \frac{p \cdot u(s) + q \cdot c}{1 - \beta \cdot q}.$$

as $p \cdot \beta \cdot V_0(0) \geq 0$.

To prove the second assertion, we will use $a = 0$. Then

$$V_1(s) \geq q \cdot c + q \cdot V_0(s) + p \cdot u(0) + p \cdot V_0(s).$$

Thus $V_1 \geq V_0$, if

$$q \cdot c + q \cdot V_0(s) + p \cdot u(0) + p \cdot V_0(s) \geq V_0(s)$$

$$q \cdot c + p \cdot u(0) \geq 0$$

$$c \geq -\frac{p}{q} \cdot u(0).$$

□

Remarks

3.3. (i) holds if $\beta \leq 1$ and $V_0 = d_0 \cdot u + e_0$ and $0 \leq d_0 \leq p/(1 - \beta \cdot q)$, and $e_0 \leq q \cdot c/(1 - \beta \cdot q)$, (ii) holds if $\beta = 1$ and $u(0) \geq 0$.

3.4. Conditions (i) and (ii) are only sufficient, not necessary. For example, when $u(a) = \sqrt{a}$, and $V_0(s) = d_0 \cdot \sqrt{s}$, $n \to V_n(s)$ is increasing in n iff $\sqrt{\lambda} \cdot d_0 \leq p$, where $\lambda := (1 - \beta \cdot q)^2 - p^2 \cdot \beta^2$ (cf. Example 1).

3.5. The Tables 3.2 and 3.3 show that $s \to V_n(s)$ and $n \to V_n(s)$ are increasing, if $u(a) = a/\sqrt{a^2+4}, p = 0.25, c = 1, K = 10, N = 10, \beta = 0.85$ and $V_0(s) = s/\sqrt{s^2+4}$, and $u(a) = a^2, p = 0.5, c = 4, K = 10, N = 10, \beta = 0.80$, and $V_0(s) = 0.5 \cdot s^2 + 2$, respectively.

3.6. The Figures 3.1 and 3.2 show that $s \to V_n(s)$ and $n \to V_n(s)$ are increasing, if $u(a) = a/\sqrt{a^2+4}, p = 0.25, c = 1, K = 10, N = 10, \beta = 0.85$, and $V_0(s) = s/\sqrt{s^2+4}$.

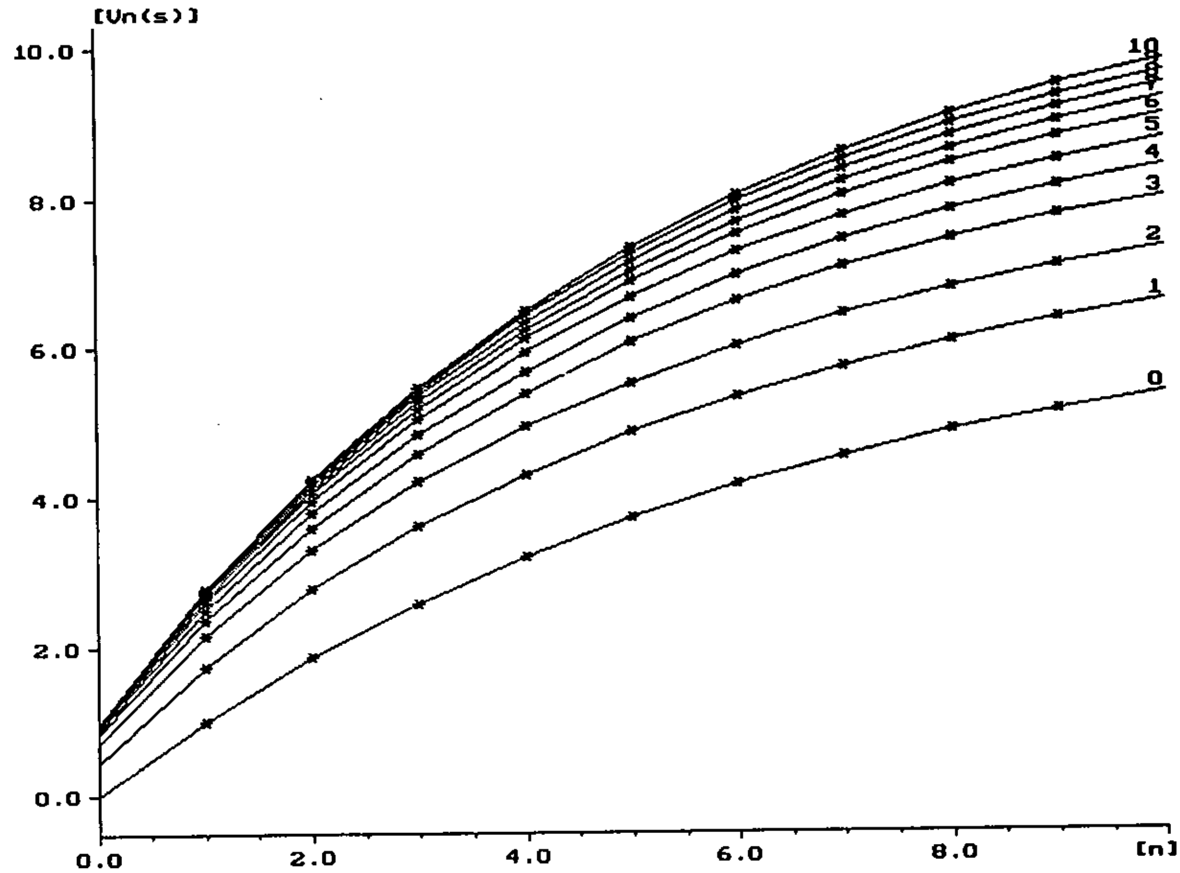


Figure 3.2. $n - V_n(s)$ for $u(a) = a/\sqrt{a^2+4}$, $p = 0.25$, $c = 1$, $K = 10$, $N = 10$, $\beta = 0.85$, and $V_0(s) = s/\sqrt{s^2+4}$.

n	s										
	0	1	2	3	4	5	6	7	8	9	10
0	2.00	2.50	4.00	6.50	10.00	14.50	20.00	26.50	34.00	42.50	52.00
1	3.60	4.30	6.40	9.90	14.80	21.10	28.80	37.90	48.40	60.30	73.60
2	4.88	5.66	8.00	11.90	17.16	24.38	32.96	43.10	54.80	68.06	82.88
3	5.90	6.72	9.15	13.21	18.90	26.20	35.14	45.69	57.87	71.68	87.10
4	6.72	7.55	10.02	14.15	19.92	27.34	36.42	47.14	59.51	73.53	89.20
5	7.38	8.21	10.70	14.85	20.66	28.13	37.26	48.04	60.49	74.60	90.37
6	7.90	8.73	11.23	15.39	21.21	28.70	37.85	48.67	61.15	75.29	91.10
7	8.32	9.16	11.65	15.82	21.65	29.14	38.30	49.13	61.62	75.78	91.60
8	8.66	9.49	11.99	16.16	21.99	29.49	38.65	49.48	61.98	76.14	91.97
9	8.93	9.76	12.26	16.43	22.26	29.76	38.92	49.76	62.25	76.42	92.25
10	9.14	9.97	12.47	16.64	29.47	29.96	39.14	49.97	62.47	76.64	92.47

Table 3.3. $V_n(s)$ for $u(a) = a^2$, $p = 0.5$,
$c = 4$, $N = 10$, $K = 10$, $\beta = 0.80$, and
$V_0(s) = 0.5 \cdot s^2 + 2$.

Now we shall present the other three cases. We will use only Model 1.

3.3.2 The Discrete-Time Continuous-State Version

This case differs from case 1 only in the state space, action space and the constraint set, as follows:

$$S = A := [0, K],$$

for $K \in \mathbb{R}$.

$$D(s) = [0, s]$$

is the *set of admissible actions at state s*. Therefore

$$D = \{(s, a) \in [0, K]^2 : 0 \leq a \leq s\}.$$

Note that

$$(s, a) \longrightarrow T(s, a) := s - a$$

is measurable. Moreover,

$$(s, a) \longrightarrow r(s, a) = p \cdot u(a) + q \cdot c$$

is measurable. As before we assume that V_0 is increasing.

As in the first case, let $V_n(s)$ denote the maximal expected n-stage reward, if the initial capital is s and *before it* is known whether or not at time $\nu = 0$ an opportunity arises. Using Theorem 2.7, we obtain the following result:

Theorem 3.6 *(a) V_N may be computed recursively by the value iteration*

$$V_n(s) = q \cdot c + \beta \cdot q \cdot V_{n-1}(s) + p \cdot \sup_{0 \le a \le s} \{u(a) + \beta \cdot V_{n-1}(s-a)\}. \tag{3.9}$$

*(b) (**Optimality Criterion**) If f_n is a maximizer at stage n for $1 \le n \le N$, then the policy $(f_n)_N^1$ is optimal for DP_N.* □

Remark

3.7. The word "maximizer" in Theorem 3.6 b) includes the property that $s \to f_n(s)$ is measurable. This was not necessary in the corresponding Theorems 3.1 b) and 3.2 b) for Case 1.

The results about the structure properties found in the first case and the proofs remain valid for the continuous-state version, as follows.

Theorem 3.7 *$s \longrightarrow V_n(s)$ is increasing for all $n \in \mathbb{N}$.* □

Theorem 3.8 *$V_n(s)$ is increasing in n for all s, if either*

(i)

$$\beta \le 1,$$

and

$$V_0(s) \le \frac{p \cdot u(s) + q \cdot c}{1 - \beta \cdot q}, \qquad \textit{for all } s,$$

or

(ii)

$$\beta = 1.$$

and

$$c \ge -\frac{p}{q} \cdot u(0).$$

□

When u and V_0 are convex functions, we show in 3.9 that the optimal policy is to invest nothing or all which we have at hand (cf. Table 3.4), in case an opportunity presents itself.

When u and V_0 are concave, it is not possible to describe the structure of the optimal policy, (except for special cases of $a^\alpha, \alpha \in (0,1)$ and $\ln a$, where it is possible to find explicit solutions, cf. Examples 1 and 2 below), but V_n is concave for all n, as shown in 3.10 below.

Theorem 3.9 *Assume that u and V_0 are convex. Then V_n is finite and convex and*

$$s \to f_n(s) := \begin{cases} 0, & \textit{if} \quad u(s) - u(0) \le \beta \cdot (V_{n-1}(s) - V_{n-1}(0)), \\ s, & \textit{else}, \end{cases}$$

is a bang-bang maximizer.

Proof. We prove by induction that V_n is finite and convex for all $n \in \mathbb{N}$.

Firstly, this is true for $n = 0$ by assumption.

Now we assume, that V_{n-1} is finite and convex. Then $a \to \beta \cdot V_{n-1}(s-a)$, which is defined on [0,s], is convex by Theorem 2.1. As u is convex, also $a \to W_n(s,a) := u(a) + \beta \cdot V_{n-1}(s-a)$ is convex on [0,s] by Theorem 2.1.

It follows by Lemma 2.1, that $a \to W_n(s,a)$ assumes on [0,s] its maximum either for $a = 0$ or $a = s$. More precisely, the smallest maximum point $f_n(s)$ has the form given above. Therefore, $V_n = \sup_{0 \le a \le s} W_n(s,a) = max(W_n(s,0), W_n(s,s)) < \infty$. The two functions $s \to W_n(s,0) = u(0) + \beta \cdot V_{n-1}(s)$ and $s \to W_n(s,s) = u(s) + \beta \cdot V_{n-1}(0)$ are convex, as u and V_{n-1} are convex. Now V_n is convex, by Theorem 2.3.

Thus all functions V_n are convex, and step $n-1 \to n$ of the proof verifies the assertion about f_n. □

	s										
n	0	1	2	3	4	5	6	7	8	9	10
1	0	1	2	3	4	5	6	7	8	9	10
2	0	1	2	3	4	5	6	7	8	9	10
3	0	1	2	3	4	5	6	7	8	9	10
4	0	1	2	3	4	5	6	7	8	9	10
5	0	1	2	3	4	5	6	7	8	9	10
6	0	1	2	3	4	5	6	7	8	9	10
7	0	1	2	3	4	5	6	7	8	9	10
8	0	1	2	3	4	5	6	7	8	9	10
9	0	1	2	3	4	5	6	7	8	9	10
10	0	1	2	3	4	5	6	7	8	9	10

Table 3.4. $f_n(s)$ for $c = -8$, $u(a) = c^a$, $p = 0.75$, $N = 10$, $K = 10$, $\beta = 0.95$, and $V_0(s) = c^s$.

Theorem 3.10 *Assume that u and V_0 are concave. Then*

(i) $s \to V_n(s)$ and $a \to W_n(s,a), s \in S$, are concave.

(ii) If there exists a smallest maximizer f_n at stage n, then $s \to f_n(s)$ is increasing.

(iii) If there exists a smallest maximizer f_n at stage n, then $n \to f_n(s)$ is decreasing.

Remark

3.8. If u and V_0 are in addition continuous at $s = 0$ and $s = K$ and hence continuous, then the smallest maximizer f_n at stage n. can be found by maximizing the *concave* functions $a \to W_n(s,a)$ on $[0,s]$, $s \in [0,K]$.

Proof. (i) The proof is similar as the one for 3.9.

Thus we only show that V_n and $a \to W_n(s, a)$ are concave, provided that V_{n-1} is concave. Firstly, $(s, a) \to V_{n-1}(s - a)$ is concave, as a concave function of an affine function is concave. As u is concave, $(s, a) \to W_n(s, a)$ is concave on the convex set D, by the "dual" of Theorem 2.1.

Now it follows that $a \to W_n(s, a)$ is concave. Moreover, $s \to \sup_{0 \le a \le s} W_n(s, a)$ is finite and concave by Theorem 2.4, as $D(s) = [0, s]$ is bounded. Now it follows by the "dual" of Theorem 2.1 that V_n is concave.

(ii) For fixed n, we have

$$\begin{aligned} V_n(s) &= q \cdot c + \beta \cdot q \cdot V_{n-1}(s) + p \cdot \sup_{0 \le a \le s} \{u(a) + \beta \cdot V_{n-1}(s - a)\} \\ &=: \sup_{0 \le a \le s} \{W(s, a)\}. \end{aligned}$$

Using the Serfozo's criterion (Theorem 2.11), we have to prove that

$$W(s, a') - W(s, a) \le W(s', a') - W(s', a),$$

for $s' > s$ and $a' > a$, $s, s' \in S$ and $a, a' \in D(s)$. This is equivalent to

$$V_{n-1}(s - a') - V_{n-1}(s - a) \le V_{n-1}(s' - a') - V_{n-1}(s' - a),$$

and this follows, as V_{n-1} is concave, from Lemma 2.4, by putting $x := s - a'$, $h := a' - a$, and $x' := s' - a' > x$.

(iii) This assertion is proved in D/L/R, we can use this proof, replacing $V(n, s)$ by $(V_n(s) - \beta \cdot q \cdot V_{n-1}(s))/p$ and $\hat{V}(n, s)$ by $\beta \cdot V_n(s)$. □

Remark

3.9. In some models, with continuous state one could use $a_\nu :=$ portion of s_ν invested, hence $\hat{D}(s) = [0, 1]$, $\hat{T}(s, a, x) = s \cdot (1 - a \cdot x)$, $\hat{f}_n(s) \in [0, 1]$ and hence $\hat{f}_n(s) = f_n(s)/s$. If $\hat{f}_n$ is a bang-bang maximizer, then $f_n(s) \in \{0, 1\}$ and vice versa.

	s										
n	0	1	2	3	4	5	6	7	8	9	10
1	0	1	1	2	2	3	3	4	4	5	5
2	0	1	1	2	2	3	3	4	4	5	5
3	0	1	1	2	2	3	3	3	4	4	5
4	0	1	1	2	2	3	3	3	4	4	5
5	0	1	1	2	2	3	3	3	4	4	5
6	0	1	1	2	2	3	3	3	4	4	5
7	0	1	1	2	2	3	3	3	4	4	4
8	0	1	1	2	2	3	3	3	4	4	4
9	0	1	1	2	2	3	3	3	4	4	4
10	0	1	1	2	2	3	3	3	4	4	4

Table 3.5. $f_n(s)$ for $c = 1$, $u(a) = \ln(a+1)$,
$p = 0.25$, $N = 10$, $K = 10$, $\beta = 0.90$, and
$V_0(s) = \ln(s+1) + 3$.

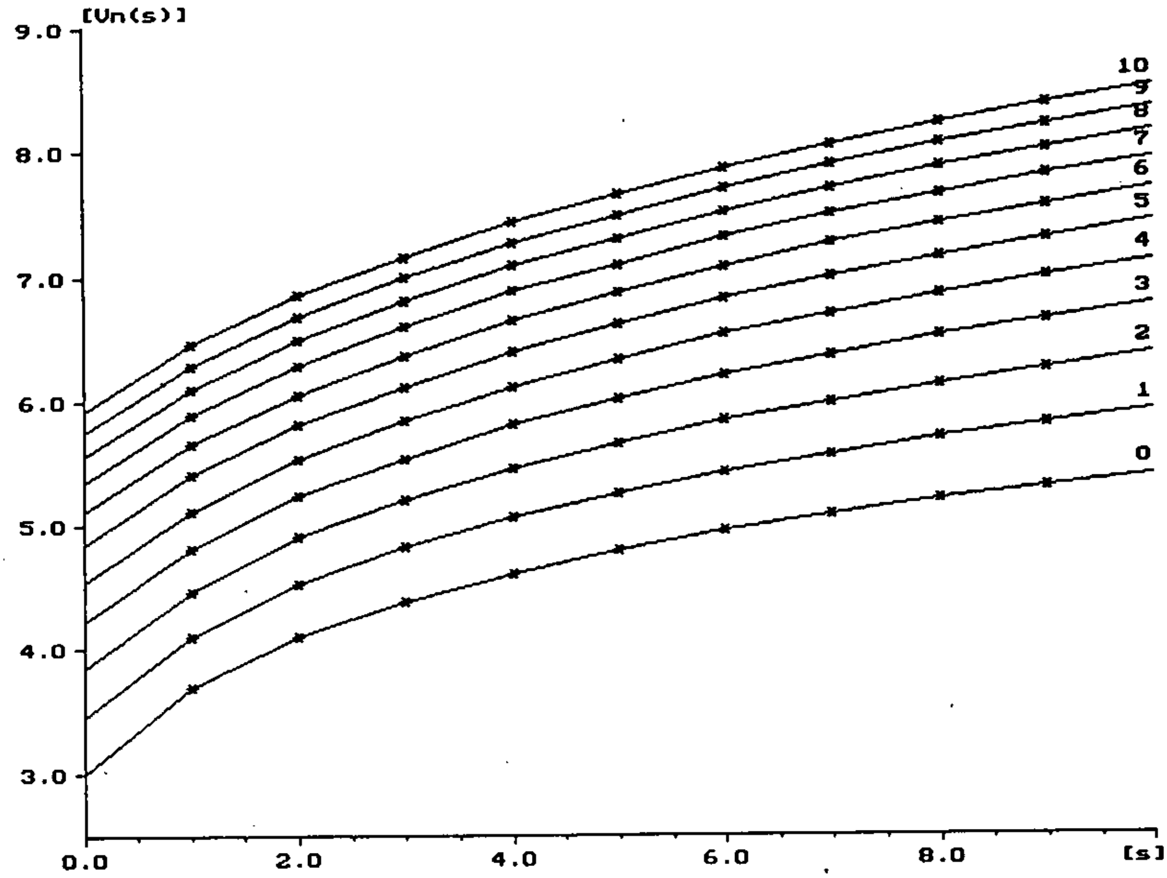


Figure 3.3. $s \to V_n(s)$ for $c = 1$, $u(a) = \ln(a+1)$,
$p = 0.25$, $N = 10$, $K = 10$, $\beta = 0.90$, and
$V_0(s) = \ln(s+1) + 3$.

Remark
3.10. At first sight it seems strange that in the example of Figure 3.4 $V_n(s,p)$ is decreasing and not increasing in p. That this can really happen, can be seen in the case of Example 2 when

$$V_1(s) \quad = \quad (p + \beta \cdot d_0) \ln s + q \cdot c + p \cdot \ln \lambda_0 + \beta \cdot e_0$$

$$= \quad p \cdot (\ln s - c + \ln \lambda_0) + \beta \cdot d_0 \cdot \ln s + \beta \cdot e_0,$$

and λ_0 and e_0 are independent of p. Therefore $p \to V_1(s,p)$ is decreasing if $s \leq s^* := e^{c/\lambda_0}$, and increasing, if $s \geq s^*$.

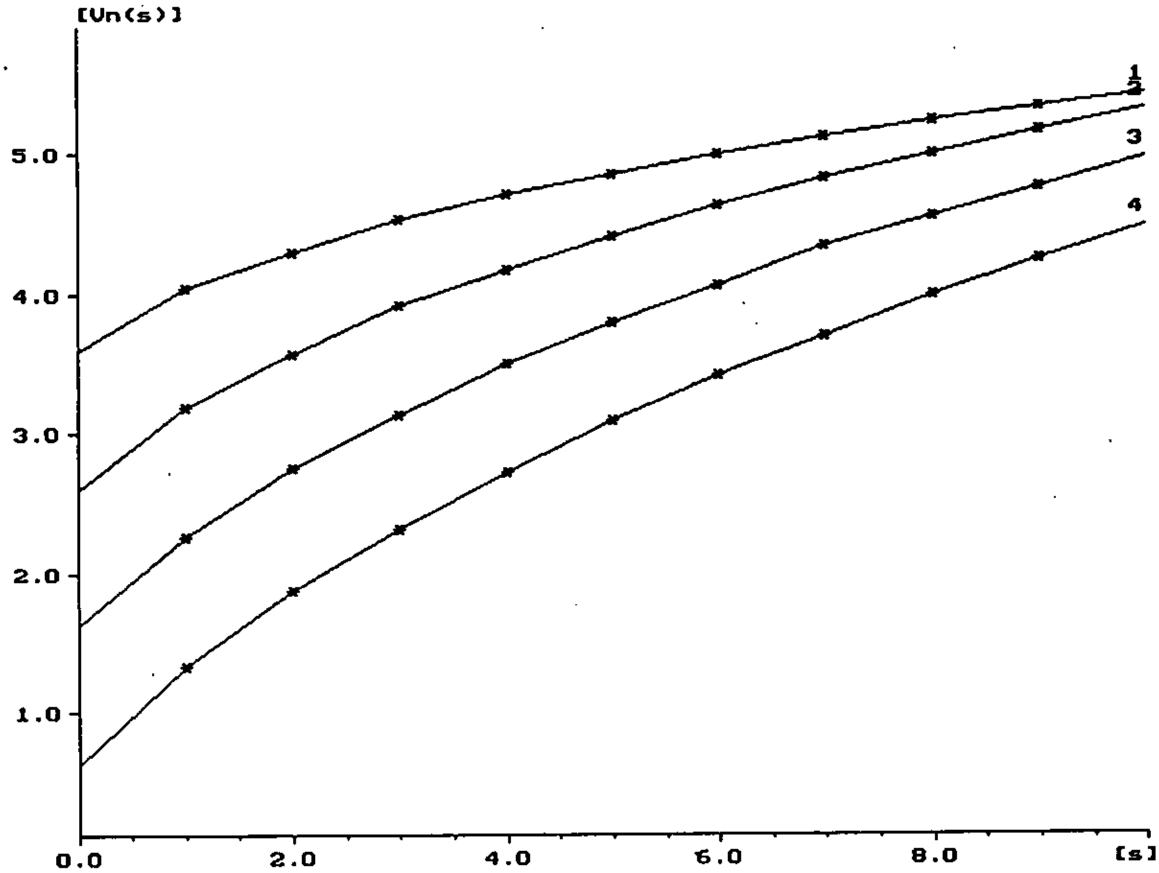


Figure 3.4. $s \to V_n(s)$ for $c = 1$, $u(a) = \ln(a+1)$,
$p_1 = 0.25$, $p_2 = 0.50$, $p_3 = 0.75$, $p_4 = 1.0$, $n = 7$,
$K = 10$, $\beta = 0.90$, and $V_0(s) = \ln(s+1) + 3$.

Theorem 3.11 *If u and V_0 are continuous, then*

$$(s,p) \longrightarrow V_n(s,p)$$

is continuous, and hence $V_n(s,p)$ is also continuous in each variable.

Remark
3.11. We use repeatedly, without further mentioning, that the sum and the product of continuous functions are continuous, and that a continuous function of a continuous function is continuous.

Proof. Case $n = 1$.

As V_0 and $(s,a) \to s - a$ are continuous, $(s,a) \to V_0(s-a)$ is continuous. As u is continuous, also $(s,a) \to u(a) + \beta \cdot V_0(s-a)$ is continuous.

Now it follows from Lemma 2.7 with $n = 1$ and $d(s) := s$, that

$$s \to \sup_{0 \le a \le s} \{u(a) + \beta \cdot V_0(s-a)\}$$

is continuous. Finally, it follows from

$$V_1(s,p) = q \cdot c + \beta \cdot q \cdot V_0(s) + p \cdot \sup_{0 \le a \le s} \{u(a) + \beta \cdot V_0(s-a)\},$$

that $(s,p) \to V_1(s,p)$ is continuous.

Now assume that $(s,p) \to V_k(s,p)$ is continuous. Then

$$(s,p,a) \to W(s,p,a) := u(a) + \beta \cdot V_k(s-a,p)$$

is continuous. Using Lemma 2.7 with $n = 2$, $d(s,p) := s$, we see that

$$(s,p) \to \sup_{0 \le a \le s} W(s,p,a)$$

is continuous. Now

$$V_{k+1}(s,p) = q \cdot c + \beta \cdot q \cdot V_k(s,p) + p \cdot \sup_{0 \le a \le s} \{u(a) + \beta \cdot V_k(s-a,p)\},$$

shows that $(s,p) \to V_{k+1}(s,p)$ is continuous. □

In principle, $V_n(s) = \infty$ is possible. A sufficient condition for $V_n(s) < \infty$ is continuity of u and V_0, as V_n is continuous on the compact interval $[0, K]$ by Theorem 3.11, hence V_n is bounded by Theorem 2.8. Another result is

Theorem 3.12 *If $\beta < 1$ and if u is increasing, then there exists $V(s) := \lim_{n\to\infty} V_n(s)$ and $V(s) \le \frac{1}{1-\beta} \cdot (q \cdot |c| + p \cdot |u|)$.*

Proof. (a) V_n is increasing by Theorem 3.7, as V_0 is increasing.

(b) From the VI, we obtain

$$\begin{aligned} V_n(s) &= q \cdot c + \beta \cdot q \cdot V_{n-1}(s) + p \cdot \sup_{0 \le a \le s} \{u(a) + \beta \cdot V_{n-1}(s-a)\} \\ &\le q \cdot c + \beta \cdot q \cdot V_{n-1}(s) + p \cdot (\sup_{0 \le a \le s} u(a) + \beta \cdot \sup_{0 \le a \le s} V_{n-1}(s-a)). \end{aligned}$$

As u and V_{n-1} are increasing, then $\sup_{0 \le a \le s} u(a) = u(s)$, *and* $\sup_{0 \le a \le s} V_{n-1}(s-a) = V_{n-1}(s)$. Therefore

$$V_n(s) \le q \cdot c + \beta \cdot V_{n-1}(s) + p \cdot u(s).$$

Now the assertion follows from Lemma 2.5 with $b_n := V_{n-1}(s)$, s fixed, c replaced by $q \cdot c + p \cdot u(s)$ and $\alpha := \beta$. Thus

$$\begin{aligned} V_n(s) &\le (q \cdot c + p \cdot u(s)) \cdot \sigma_n(\beta) + \beta^n \cdot V_0(s) \\ &\le (q \cdot |c| + p \cdot |u(s)|) \cdot \sigma_n(\beta) + \beta^n \cdot |V_0(s)|. \end{aligned} \tag{3.10}$$

(c) By (a) and (b), V_n converges, and letting n go to infinity in (3.10) we obtain the upper bound of $V(s)$. □

3.3.3 Closed Solution

While the theorems above yield the structure of the optimal policy and of the value functions V_n and can be used, in the obvious manner, to simplify the necessary computations, they do not present a closed-form expression for f_n and V_n.

Example 1: A special case for which the optimal policy and the value functions can be completely specified is when $u(a) := a^\alpha$ and $V_0(s) := d_0 \cdot s^\alpha + e_0$, $0 < \alpha < 1$, $d_0 \in \mathbb{R}_+, e_0 \in \mathbb{R}$. Then the VI reads for $n = 1$

$$\begin{aligned} V_1(s) &= q \cdot c + q \cdot \beta \cdot (d_0 \cdot s^\alpha + e_0) + p \cdot \sup_{0 \le a \le s} \{a^\alpha + \beta \cdot (e_0 + d_0 \cdot (s-a)^\alpha)\}, \\ &= q \cdot c + \beta \cdot e_0 + q \cdot \beta \cdot d_0 \cdot s^\alpha + p \cdot \sup_{0 \le a \le s} \{a^\alpha + \beta \cdot d_0 \cdot (s-a)^\alpha\}. \end{aligned}$$

For $d_0 > 0$, we have for $0 \le a \le s$ and $g(a) := a^\alpha + \beta \cdot d_0 \cdot (s-a)^\alpha$,

$$\begin{aligned} g'(a) = 0 \quad \Leftrightarrow \quad \alpha \cdot a^{\alpha-1} - \alpha \cdot \beta \cdot d_0 \cdot (s-a)^{\alpha-1} &= 0 \\ a^{\alpha-1} - \beta \cdot d_0 \cdot (s-a)^{\alpha-1} &= 0 \\ \left(\frac{a}{(\beta \cdot d_0)^{\frac{1}{\alpha-1}}} \right)^{\alpha-1} &= (s-a)^{\alpha-1} \\ \frac{a}{(\beta \cdot d_0)^{\frac{1}{\alpha-1}}} + a &= s. \end{aligned}$$

Put $\rho := \frac{1}{1-\alpha}$. Then

$$g'(a) = 0 \quad \Leftrightarrow \quad \frac{a}{(\beta \cdot d_0)^{-\rho}} + a = s$$

$$
\begin{aligned}
(\beta \cdot d_0)^\rho \cdot a + a &= s \\
a^* &= \frac{s}{1+(\beta \cdot d_0)^\rho} \in (0, s).
\end{aligned}
$$

As g is concave, $a^* := s/(1+(\beta \cdot d_0)^\rho)$ is the unique maximum point of g by Theorem 2.6. For $d_0 = 0$ the function g has obviously the unique maximum point $a^* := s$. It follows that for $d_0 > 0$ and also for $d_0 = 0$

$$
s \longrightarrow f_1(s) := \frac{s}{1+(\beta \cdot d_0)^\rho}
$$

is the unique maximizer at stage $n = 1$. Moreover, we get

$$
\begin{aligned}
V_1(s) &= q \cdot c + \beta \cdot e_0 + \beta \cdot q \cdot d_0 \cdot s^\alpha + p \cdot \left(\frac{s}{1+(\beta \cdot d_0)^\rho}\right)^\alpha + \\
&\quad + \beta \cdot p \cdot d_0 \cdot \left(s - \frac{s}{1+(\beta \cdot d_0)^\rho}\right)^\alpha .
\end{aligned}
$$

Put $\delta := 1 + (\beta \cdot d_0)^\rho$, then

$$
\begin{aligned}
V_1(s) &= q \cdot c + \beta \cdot e_0 + \beta \cdot q \cdot d_0 \cdot s^\alpha + p \cdot s^\alpha \left(\frac{1}{\delta^\alpha} + \frac{(\beta \cdot d_0)^\rho}{\delta^\alpha}\right) \\
&= q \cdot c + \beta \cdot e_0 + \beta \cdot q \cdot d_0 \cdot s^\alpha + \frac{p \cdot s^\alpha}{\delta^\alpha} \cdot (1 + (\beta \cdot d_0)^\rho) \\
&= q \cdot c + \beta \cdot e_0 + \beta \cdot q \cdot d_0 \cdot s^\alpha + p \cdot s^\alpha \cdot \delta^{1-\alpha} \\
&= q \cdot c + \beta \cdot e_0 + (\beta \cdot q \cdot d_0 + p \cdot \delta^{1-\alpha}) \cdot s^\alpha \\
&= q \cdot c + \beta \cdot e_0 + \left(\beta \cdot q \cdot d_0 + p \cdot (1 + (\beta \cdot d_0)^\rho)^{1-\alpha}\right) \cdot s^\alpha .
\end{aligned}
$$

Let be

$$
\begin{aligned}
d_1 &:= \beta \cdot q \cdot d_0 + p \cdot (1 + (\beta \cdot d_0)^\rho)^{1-\alpha}, \\
e_1 &:= q \cdot c + \beta \cdot e_0,
\end{aligned}
$$

then

$$
V_1(s) = e_1 + d_1 \cdot s^\alpha,
$$

and

$$
s \longrightarrow f_1(s) = \frac{s}{1+(\beta \cdot d_0)^\rho}.
$$

Therefore, substituting d_1 for d_0 above, we obtain $V_2(s) = d_2 \cdot s^\alpha + e_2$, where $d_2 \in \mathbb{R}^+$, $e_2 \in \mathbb{R}$ and maximizer f_2, also substituting d_2 for d_0 above, we obtain $V_3(s) = d_3 \cdot s^\alpha + e_3$, where $d_3 \in \mathbb{R}^+$, $e_3 \in \mathbb{R}$ and maximizer f_3, etc. More formally we obtain from the VI by induction on n, using the maximum point a^* of the function g above, the following result:

Proposition 3.2 *Assume that $\alpha \in (0,1)$, $u(a) = a^\alpha$, $a \in A$. $V_0(s) = d_0 \cdot s^\alpha + e_0, s \in S$, for some $d_0 \in \mathbb{R}_+$ and for some $e_0 \in \mathbb{R}$. Then the following holds:*

(i) V_n is of the form

$$V_n(s) = d_n \cdot s^\alpha + e_n, \qquad s \in S, n \in \mathbb{N},$$

for some $d_n \in \mathbb{R}^+$ which satisfy the recursion

$$d_n = \beta \cdot q \cdot d_{n-1} + p \cdot \left(1 + (\beta \cdot d_{n-1})^\rho\right)^{1-\alpha}, \qquad n \in \mathbb{N}, \tag{3.11}$$

and

$$e_n = q \cdot c \cdot \sigma_n(\beta) + \beta^n \cdot e_0 \qquad \text{for } e_0 \in \mathbb{R}.$$

(ii)

$$s \longrightarrow f_n(s) := \frac{s}{1 + (\beta \cdot d_{n-1})^\rho}$$

is the unique maximizer at stage $n \in \mathbb{N}$. □

Note that e_n has a closed solution, as using Lemma 2.5 we obtain

$$e_n = q \cdot c \cdot \sum_{i=0}^{n-1} \beta^i + \beta^n \cdot e_0.$$

Obviously e_n converges towards $q \cdot c/(1-\beta)$ if $n \to \infty$, iff $\beta < 1$.

Lemma 3.2 *Let G be an increasing function from $I \in \mathbb{R}$ into $\mathbb{R}$, select $x_0 \in I$ and define $(x_n)_1^\infty$ recursively by $x_{n+1} = G(x_n), n \in \mathbb{N}$.*

If $x_1 \geq x_0$ then (x_n) is increasing, and if $x_1 \leq x_0$ then (x_n) is decreasing.

Proof. We show that (x_n) is increasing if $x_1 \geq x_0$. (The other result is proved in the same way.)

We have to show that

$$x_n \leq x_{n+1}, \quad \text{for } n \in \mathbb{N}. \tag{3.12}$$

For $n = 0$, (3.12) holds by assumption. Now assume that (3.12) holds for some $n \in \mathbb{N}_0$. $G(x_n) \leq G(x_{n+1})$ by isotonocity of G, and thus $x_{n+2} = G(x_{n+1}) \geq G(x_n) = x_{n+1}$. Therefore (3.12) holds also for $n+1$ and therefore by induction for all n. □

Proposition 3.3 *Let $\beta \leq 1$. Put $K(\beta) := (1 - \beta \cdot q)^\rho - (\beta \cdot p)^\rho, 0 < \beta \leq 1$. Then (d_n) is increasing, if either $p \geq d_0$ or if $p < d_0$ and $\beta \geq \beta^*$, where β^* is the unique solution of $K(\beta) = (\frac{p}{d_0})^\rho$.*

Proof. Applying Lemma 3.2, (d_n) is increasing iff $d_0 \leq d_1$. Now

$$\begin{aligned} d_0 &\leq d_1 \\ \Leftrightarrow \quad d_0 &\leq \beta \cdot q \cdot d_0 + p \cdot (1 + (\beta \cdot d_0)^\rho)^{1-\alpha} \\ \Leftrightarrow \quad d_0 - \beta \cdot q \cdot d_0 &\leq p \cdot (1 + (\beta \cdot d_0)^\rho)^{1-\alpha} \\ \Leftrightarrow \quad d_0 \cdot (1 - \beta \cdot q) &\leq p \cdot (1 + (\beta \cdot d_0)^\rho)^{1-\alpha} \end{aligned} \tag{3.13}$$

Case 1: $d_0 = 0$. Then (3.13) holds.

Now assume $d_0 > 0$. Then, as $\beta \leq 1$, we have $1 - \beta \cdot q > 0$, and hence (3.13) is equivalent to

$$\begin{aligned} d_0^{\frac{1}{1-\alpha}} \cdot (1 - \beta \cdot q)^{\frac{1}{1-\alpha}} &\leq p^{\frac{1}{1-\alpha}} \cdot (1 + (\beta \cdot d_0)^\rho) \\ \Leftrightarrow \quad d_0^\rho \cdot (1 - \beta \cdot q)^\rho &\leq p^\rho \cdot (1 + (\beta \cdot d_0)^\rho) \\ \Leftrightarrow \quad d_0^\rho \cdot (1 - \beta \cdot q)^\rho - (p \cdot \beta \cdot d_0)^\rho &\leq p^\rho \\ \Leftrightarrow \quad K(\beta) := (1 - \beta \cdot q)^\rho - (\beta \cdot p)^\rho &\leq (\frac{p}{d_0})^\rho. \end{aligned} \tag{3.14}$$

Obviously K decreases on $(0, 1]$ from 1 to 0.

Case 2: $p \geq d_0 > 0$. Then (3.14) holds for all $\beta \in (0, 1]$.

Case 3: $p < d_0$. Then (3.14) holds iff $\beta \geq \beta^*$, where β^* is the unique solution of $K(\beta) = (\frac{p}{d_0})^\rho$.

Now the proposition follows, as we can combine cases 1 and 2 to the condition $p \geq d_0$. □

Theorem 3.13 *If* $\beta < 1$, *then*

$$d_n \rightarrow \frac{p}{((1 - \beta \cdot q)^\rho - (\beta \cdot p)^\rho)^{1-\alpha}}.$$

Proof. (a) (d_n) is increasing or decreasing, and $d_n \geq 0, \forall n$. Therefore d_n is converging, if d_n is bounded above.

(b) (d_n) is bounded above.

Let $h(x) := 1 + x^\gamma - (1 + x)^\gamma$, with $0 < \gamma < 1$, and $x \geq 0$. We have $h(0) = 0$ and h is increasing, as $h'(x) = \gamma \cdot (x^{\gamma-1} - (1 + x)^{\gamma-1}) > 0$, for all $x > 0$, therefore $h(x) \geq 0$. Consequently

$$1 + x^\gamma \leq (1 + x)^\gamma. \tag{3.15}$$

Combining (3.11) and (3.15), with $x := (\beta \cdot d_{n-1})^\rho$, $\gamma := 1 - \alpha$, we obtain

$$\begin{aligned} (1 + (\beta \cdot d_{n-1})^\rho)^{1-\alpha} &\leq 1 + ((\beta \cdot d_{n-1})^\rho)^{1-\alpha} \\ &\leq 1 + \beta \cdot d_{n-1}. \end{aligned}$$

Thus

$$\begin{aligned} d_n &\le \beta \cdot q \cdot d_{n-1} + p \cdot (1 + \beta \cdot d_{n-1}) \\ &\le \beta \cdot d_{n-1} + p. \end{aligned}$$

Applying Lemma 2.5, we get

$$\begin{aligned} d_n &\le p \cdot \sigma_n(\beta) + \beta^n \cdot d_0 \\ &\le d_0 + \frac{p}{1-\beta}, \end{aligned}$$

as $\beta^n \cdot d_0 \le d_0$.

(c) Taking the limit for n to infinity in

$$d_n = \beta \cdot q \cdot d_{n-1} + p \cdot \left(1 + (\beta \cdot d_{n-1})^\rho\right)^{1-\alpha},$$

we obtain

$$d = \beta \cdot q \cdot d + p \cdot \left(1 + (\beta \cdot d)^\rho\right)^{1-\alpha}.$$

Here we have used that $x \to (1+x)^{1-\alpha}$ is continuous. Now we have

$$\begin{aligned} d &= \beta \cdot q \cdot d + p \cdot \left(1 + (\beta \cdot d)^\rho\right)^{1-\alpha} \\ d^\rho \cdot (1 - \beta \cdot q)^\rho &= p^\rho + (\beta \cdot d \cdot p)^\rho \\ d^\rho \cdot ((1 - \beta \cdot q)^\rho - (\beta \cdot p)^\rho) &= p^\rho \\ d &= \frac{p}{\left((1-\beta \cdot q)^\rho - (\beta \cdot p)^\rho\right)^{1-\alpha}}, \end{aligned}$$

as $(1 - \beta \cdot q)^\rho - (\beta \cdot p)^\rho > 0$. □

Therefore

Proposition 3.4 *If $\beta < 1$ and putting $\lambda := \left((1-\beta \cdot q)^\rho - (\beta \cdot p)^\rho\right)^{1-\alpha}$, then for $n \to \infty$*

(i)

$$V_n(s) \to \frac{p}{\lambda} \cdot s^\alpha + \frac{q \cdot c}{1-\beta},$$

and

(ii)

$$f_n(s) \to \frac{s \cdot \lambda}{\lambda + \beta \cdot p}.$$

Proof. By Theorem 3.13 d_n converges towards p/λ, and e_n converges towards $q \cdot c/(1-\beta)$, if $\beta < 1$. Then replacing these expressions in Proposition 3.2 (i) and (ii), we obtain

$$V_n(s) \to \frac{p}{\lambda} \cdot s^\alpha + \frac{q \cdot c}{1-\beta},$$

and

$$f_n(s) \to \frac{s \cdot \lambda}{\lambda + \beta \cdot p}.$$

□

Example 2: Another special case for which the optimal policy and the value functions V_n can be made explicit is when $S = A = (0, s), D(s) = (0, s)$, $u(a) := \ln a$, $a \in (0, s)$, and $V_0(s) := d_0 \cdot \ln s + e_0$, $s \in (0, K]$, $d_0 \in \mathbb{R}^+$, $e_0 \in \mathbb{R}$. Then the VI reads for $n = 1$

$$\begin{aligned} V_1(s) &= q \cdot c + \beta \cdot q \cdot (e_0 + d_0 \cdot \ln s) + p \cdot \sup_{0<a<s} \{\ln a + \beta \cdot (e_0 + d_0 \cdot \ln(s - a))\} \\ &= q \cdot c + \beta \cdot e_0 + \beta \cdot q \cdot d_0 \cdot \ln s + p \cdot \sup_{0<a<s} \{\ln a + \beta \cdot d_0 \cdot \ln(s - a)\}. \end{aligned}$$

As $d_0 > 0$,

$$g(a) := \ln a + \beta \cdot d_0 \cdot \ln(s - a),$$

has the derivative

$$g'(a) = \frac{1}{a} - \frac{\beta \cdot d_0}{s - a}, \qquad 0 < a < s.$$

As g is strictly concave, it attains its unique maximum a^*, iff

$$\begin{aligned} g'(a) = 0 \qquad \Leftrightarrow \qquad \frac{1}{a} - \frac{\beta \cdot d_0}{s - a} &= 0 \\ s - a &= a \cdot \beta \cdot d_0 \\ a^* &= \frac{s}{1 + \beta \cdot d_0}. \end{aligned}$$

Thus

$$s \longrightarrow f_1(s) := \frac{s}{1 + \beta \cdot d_0}.$$

Note that $f_1(s) \in D(s) = (0, s)$, as $s > 0$ and as $1 + \beta \cdot d_0 > 1$. Moreover

$$\begin{aligned} V_1(s) &= q \cdot c + \beta \cdot e_0 + \beta \cdot q \cdot d_0 \cdot \ln s + p \cdot \ln \frac{s}{1 + \beta \cdot d_0} + \\ &\quad + p \cdot \beta \cdot d_0 \cdot \ln \left(s - \frac{s}{1 + \beta \cdot d_0} \right) \\ &= q \cdot c + \beta \cdot e_0 + \beta \cdot q \cdot d_0 \cdot \ln s - p \cdot \ln(1 + \beta \cdot d_0) + p \cdot \ln s \\ &\quad + \beta \cdot p \cdot d_0 \cdot \ln(\beta \cdot d_0) + \beta \cdot p \cdot d_0 \cdot \ln s - p \cdot \beta \cdot d_0 \cdot \ln(1 + \beta \cdot d_0) \\ &= q \cdot c + \beta \cdot e_0 + (\beta \cdot d_0 + p) \cdot \ln s - p \cdot (1 + \beta \cdot d_0) \cdot \\ &\quad \cdot \ln(1 + \beta \cdot d_0) + p \cdot \beta \cdot d_0 \cdot \ln(\beta \cdot d_0) \\ &= q \cdot c + p \cdot \ln \frac{(\beta \cdot d_0)^{\beta d_0}}{(1 + \beta \cdot d_0)^{1 + \beta d_0}} + \beta \cdot e_0 + (\beta \cdot d_0 + p) \cdot \ln s. \end{aligned}$$

Put $\lambda_0 := (\beta \cdot d_0)^{\beta d_0}/(1+\beta \cdot d_0)^{1+\beta d_0}$, then

$$V_1(s) = q \cdot c + p \cdot \ln \lambda_0 + \beta \cdot e_0 + (\beta \cdot d_0 + p) \cdot \ln s,$$

therefore

$$V_1(s) = e_1 + d_1 \cdot \ln s, \qquad d_1 > 0,$$

where

$$e_1 = q \cdot c + p \cdot \ln \lambda_0 + \beta \cdot e_0,$$
$$\lambda_0 = \frac{(\beta \cdot d_0)^{\beta d_0}}{(1+\beta \cdot d_0)^{1+\beta d_0}},$$
$$d_1 = p + \beta \cdot d_0,$$

and

$$s \longrightarrow f_1(s) = \frac{s}{1+\beta \cdot d_0}.$$

Therefore, replacing d_1 for d_0 above, we obtain $V_2(s) = d_2 \cdot \ln s + e_2$, and maximizer f_2, for $d_2 > 0$, and $e_2 \in \mathbb{R}$, etc. More formally we obtain from the VI by induction on n, using the maximum point a^* of the function g above, the following result:

Proposition 3.5 *Assume that $S = A = (0, K]$, $D(s) = (0, s)$, $u(a) = \ln a, a \in (0, s)$, $V_0(s) = d_0 \cdot \ln s + e_0, s \in (0, K]$, for some $d_0 \in \mathbb{R}^+$ and for some $e_0 \in \mathbb{R}$. Then the following holds:*

(i) V_n is of the form

$$V_n(s) = d_n \cdot \ln s + e_n, \qquad n \in \mathbb{N}, s \in (0, K],$$

for numbers $d_n > 0$ which satisfy the recursion

$$d_n = p \cdot \sigma_n(\beta) + \beta^n \cdot d_0, \qquad n \in \mathbb{N},$$

and

$$e_n = q \cdot c + p \cdot \ln \lambda_{n-1} + \beta \cdot e_{n-1}, \qquad e_{n-1} \in \mathbb{R},$$
$$\lambda_{n-1} = \frac{(\beta \cdot d_{n-1})^{\beta d_{n-1}}}{(1+\beta \cdot d_{n-1})^{1+\beta d_{n-1}}}.$$

(ii)

$$s \longrightarrow f_n(s) := \frac{s}{1+\beta \cdot d_{n-1}} = \frac{s}{q + d_n}$$

is the unique maximizer at stage $n \in \mathbb{N}$. □

Using Lemma 2.5, we find the closed solution of d_n:

$$d_n = p \cdot \sum_{i=0}^{n-1} \beta^i + \beta^n \cdot d_0.$$

As d_n has a closed solution, if $n \to \infty$, it converges towards $p/(1-\beta)$ iff $\beta < 1$.

We have

$$e_n = q \cdot c + \beta \cdot e_{n-1} + p \cdot \ln \lambda_{n-1}, \tag{3.16}$$

with

$$\lambda_{n-1} := \frac{(\beta \cdot d_{n-1})^{\beta d_{n-1}}}{(1 + \beta \cdot d_{n-1})^{1+\beta d_{n-1}}}.$$

As $d_n \to p/(1-\beta)$ if $\beta < 1$, it implies that $\lambda := \lim_{n\to\infty} \lambda_n$ exists.

Put $\mu := q \cdot c + p \cdot \ln \lambda$, then

$$\forall \epsilon > 0, \exists \quad n_0(\epsilon) : q \cdot c + p \cdot \ln \lambda_{n-1} \leq \mu + \epsilon, \quad \forall n \geq n_0(\epsilon). \tag{3.17}$$

Replacing (3.16) in (3.17), we obtain

$$e_n \leq \beta \cdot e_{n-1} + \mu + \epsilon, \quad \forall n \geq n_0(\epsilon).$$

It follows by induction that

$$\begin{aligned} e_{n_0+k} &\leq \beta \cdot e_{n_0+k-1} + \mu + \epsilon, \quad k \geq n_0 \\ &\leq \mu + \epsilon + \beta \cdot (\mu + \epsilon + \beta \cdot e_{n_0+k-2}) \\ &\leq \cdots \\ &\leq (\mu + \epsilon) \cdot \sigma_k(\beta) + \beta^k \cdot e_{n_0}, \quad \forall k \in \mathbb{N}. \end{aligned} \tag{3.18}$$

If $k \to \infty$, the right hand side of (3.18) converges towards $(\mu + \epsilon)/(1-\beta)$, thus $e_n \leq (\mu + \epsilon)/(1-\beta)$ for all n large enough. Thus

$$\limsup_{n\to\infty} e_n \leq \frac{\mu + \epsilon}{1-\beta}.$$

As this holds for all $\epsilon > 0$, we have

$$\limsup_{n\to\infty} e_n \leq \frac{\mu}{1-\beta}.$$

Analogously $\forall \epsilon > 0, \exists \quad n_0(\epsilon) : q \cdot c + p \cdot \ln \lambda_{n-1} \geq \mu - \epsilon$, proceeding in same manner, we get

$$\liminf_{n\to\infty} \geq \frac{\mu}{1-\beta}.$$

As $\mu/(1-\beta) \leq \liminf_{n\to\infty} e_n \leq \limsup_{n\to\infty} \leq \mu/(1-\beta)$, we have $\liminf_{n\to\infty} e_n = \limsup_{n\to\infty} e_n = \mu/(1-\beta)$ and hence

$$e_n \to \frac{\mu}{1-\beta}.$$

Therefore,

Proposition 3.6 *If $\beta < 1$, then for $n \to \infty$*

(i)

$$V_n(s) \to \frac{1}{1-\beta} \cdot (p \cdot \ln s + \mu),$$

and

(ii)

$$f_n(s) \to \frac{s \cdot (1-\beta)}{1-\beta \cdot q}.$$

Proof. The proof uses the convergence of d_n and e_n. As d_n converges towards $p/(1-\beta)$ and e_n converges towards $\mu/(1-\beta)$ if $\beta < 1$. Then replacing these expressions in Proposition 3.5 (i) and (ii), we obtain

$$V_n(s) \to \frac{1}{1-\beta} \cdot (p \cdot \ln s + \mu),$$

and

$$f_n(s) \to \frac{s \cdot (1-\beta)}{1-\beta \cdot q}.$$

□

Remark

3.12. Even if a recursion for a sequence (d_n) has a closed solution it may be better to use the recursion for computations. The closed solution may be useful for studying properties of sequence (d_n).

Example 3: Here we assume that $r(s,a) := q \cdot c + p \cdot (h_1 \cdot \sqrt{a} + h_2 \cdot \sqrt{s-a} + h_3 \cdot \sqrt{s})$, $V_0(s) = d_0 \cdot \sqrt{s} + e_0$, for some $d_0 \in \mathbb{R}_+$, $e_0 \in \mathbb{R}$ and $h_1, h_2, h_3 \geq 0$. Then the VI reads for $n = 1$

$$\begin{aligned} V_1(s) &= q \cdot c + \beta \cdot q \cdot (d_0 \cdot \sqrt{s} + e_0) + p \cdot \sup_{0 \leq a \leq s} \{h_1 \cdot \sqrt{a} + h_2 \cdot \sqrt{s-a} + \\ &\quad + h_3 \cdot \sqrt{s} + \beta \cdot (d_0 \cdot \sqrt{s-a} + e_0)\} \\ &= q \cdot c + \beta \cdot e_0 + \beta \cdot q \cdot d_0 \cdot \sqrt{s} + p \cdot h_3 \cdot \sqrt{s} + p \cdot \sup_{0 \leq a \leq s} \{h_1 \cdot \sqrt{a} + (h_2 + \\ &\quad + \beta \cdot d_0) \cdot \sqrt{s-a}\}. \end{aligned} \tag{3.19}$$

Put $g(a) := h_1 \cdot \sqrt{a} + (h_2 + \beta \cdot d_0) \cdot \sqrt{s-a}$. Then g is concave, and analogously as in the cases above, we obtain

$$a^* = \frac{h_1^2 \cdot s}{h_1^2 + (h_2 + \beta \cdot d_0)^2},$$

as the unique maximum point of g.

Substituting a^* in (3.19), we get

$$V_1(s) = q \cdot c + \beta \cdot e_0 + \left(\beta \cdot q \cdot d_0 + p \cdot \left(h_3 + \sqrt{h_1^2 + (h_2 + \beta \cdot d_0)^2}\right)\right) \cdot \sqrt{s}.$$

Therefore

$$V_1(s) = d_1 \cdot \sqrt{s} + e_1,$$

where

$$d_1 = \beta \cdot q \cdot d_0 + p \cdot \left(h_3 + \sqrt{h_1^2 + (h_2 + \beta \cdot d_0)^2}\right),$$

and

$$e_1 = q \cdot c + \beta \cdot e_0.$$

More formally, we obtain

Proposition 3.7 *If* $r(s,a) = q \cdot c + p \cdot (h_1 \cdot \sqrt{a} + h_2 \cdot \sqrt{s-a} + h_3 \cdot \sqrt{s})$, $V_0(s) = d_0 \cdot \sqrt{s} + e_0$, *for some* $d_0 \in \mathbb{R}_+$, *for some* $e_0 \in \mathbb{R}$ *and* $h_1, h_2, h_3 \geq 0$, *then*

(i) V_n *is of the form*

$$V_n(s) = d_n \cdot \sqrt{s} + e_n, \quad n \in \mathbb{N},$$

for some $d_n \in \mathbb{R}^+$ *which satisfy the recursion*

$$d_n = \beta \cdot q \cdot d_{n-1} + p \cdot \left(h_3 + \sqrt{h_1^2 + (h_2 + \beta \cdot d_{n-1})^2}\right), \quad n \in \mathbb{N}$$

and

$$e_n = q \cdot c \cdot \sigma_n(\beta) + \beta^n \cdot e_0, \quad e_0 \in \mathbb{R}.$$

(ii)

$$s \to f_n(s) := \frac{h_1^2 \cdot s}{h_1^2 + (h_2 + \beta \cdot d_{n-1})^2}$$

is the unique maximizer at stage n. □

3.3.4 Allocation times as renewal process: Discrete-State Version

In this model it is assumed that opportunities arrive at random times $0 = T_0 < T_1 < T_2 < \cdots$, but only opportunities before some given time $t_{max} \in \mathbb{N}$ are allowed. As a consequence, the state s_ν at the time of the ν-th allocation, $\nu = 0, 1, 2, \ldots$, must consist of the remaining resource y_ν and the remaining time t_ν.

Now our *state space* is $S := \mathbb{N}_{0,t_{max}} \times \mathbb{N}_{0,K}$, where K has the same meaning as before and $t_{max} \in \mathbb{N}$. The *state* is $s_\nu = (t_\nu, y_\nu)$, where t_ν is the time available for further investment and y_ν is the available resource, both before the ν-th investment has been made, $\nu \in \mathbb{N}_0$.

A and a are defined as before. We assume that the disturbance random variables X_ν are i.i.d. and $X_\nu > 0$, hence that the sequence (T_n), where $T_n := \sum_{i=1}^{n} X_i$ is the time of the n-th allocation, is a *renewal process.*

$X_{\nu+1}$ is the time until the next opportunity, and its distribution on $(0, \infty)$ can be arbitrary. Important special case:

> $X \sim geo(p)$. This is a discrete time counterpart of the model where $S = [0, t_{max}] \times [0, K]$ and $X \sim exp(\lambda)$, for $\lambda \in \mathbb{R}^+$; this corresponds to the case of D/L/R, as (T_n) is a Poisson process.

The *disturbance space* is $M = \{0, 1, \ldots\}$.

$$D(t, y) = \{0, 1, \ldots, y\}$$

is the *set of admissible actions at state* $s = (t, y)$.

The *transition function* is given by

$$s_{\nu+1} = T(t_\nu, y_\nu, a_\nu, x_{\nu+1}) = \begin{cases} ((t_\nu - x_\nu)^+, y_\nu - a_\nu), & \text{if } t_\nu > 0, \\ (0, y_\nu), & \text{if } t_\nu = 0. \end{cases}$$

As

$$\tilde{r}(t, y, a, x) = \begin{cases} u(a), & \text{if } t > 0, \\ 0, & \text{if } t = 0, \end{cases}$$

the one-stage reward is

$$r(t, y, a) = \begin{cases} u(a), & \text{if } t > 0, \\ 0, & \text{if } t = 0. \end{cases}$$

The terminal reward $V_0(t_N, y_N)$ is defined as $v_0(y_N) \geq 0$, e.g. $V_0(t_N, y_N) = d_0 \cdot u(y_N)$, for some $d_0 \in \mathbb{R}_+$. We assume $v_0(0) = 0$, and no discounting, i.e. $\beta = 1$.

Now we assume that the process of opportunities stops when $t = 0$ is reached. Let $V_n(t,y)$ denote the maximal expected reward when the initial state is (t,y) and when the process stops after the n-th opportunity or after reaching $t = 0$, whichever occurs first. Then the VI is of the form:

$$V_n(t,y) = \begin{cases} \max\limits_{0\le a\le y} \{u(a) + EV_{n-1}((t-X)^+, y-a)\}, & \text{if } t > 0, \\ v_0(y), & \text{if } t = 0. \end{cases} \tag{3.20}$$

This VI is different from the VI of D/L/R (p. 1128).

Now we assume that $X \sim geo(p), p \in (0,1)$ so that the process, starting in (t,y), $t \in \mathbb{N}$, stops after at most $n := t$ offers. Therefore

$$\forall n \ge t: \quad V_n(t,y) = V_t(t,y) = \lim_{n\to\infty} V_n(t,y) =: V(t,y).$$

The VI simplifies to

$$V(t,y) = \begin{cases} \max\limits_{0\le a\le y} \{u(a) + EV((t-X)^+, y-a)\}, & \text{if } t > 0, \\ v_0(y), & \text{if } t = 0. \end{cases} \tag{3.21}$$

If $p(x) := P(X = x), x \in \mathbb{N}$, we obtain if $t > 0$

$$\begin{aligned} V(t,y) &= \max_{0\le a\le y} \{u(a) + \sum_{x=1}^{t} p\cdot(1-p)^{x-1}\cdot V(t-x, y-a)\} \qquad (3.22)\\ &=: \max_{0\le a\le y} W(t,y,a), \quad t \in \mathbb{N}_{0,t_{max}}, y \in \mathbb{N}_{0,K}, \end{aligned}$$

Note that (3.22) can be solved numerically by "*recursion in state space*", starting with $V(0,y) = v_0(y)$ for $0 \le y \le K$: if $V(t',y')$ is known for $0 \le t' \le t-1$ and all $y' \in \mathbb{N}_{0,K}$, $V(t,y)$ can be computed for all y by (3.22).

Moreover, if $f(t,y)$ is a maximum point of $a \to W(t,y,a)$ for all $(t,y) \in S$, then $(f,f,\ldots)$ is a stationary infinite-stage optimal policy.

Remark

3.13. If $X \sim exp(\lambda)$, then

$$\begin{aligned} V(t,y) &= \lim_{n\to\infty} V_n(t,y) \\ &= \text{the maximal expected reward when the process goes on until it reaches } t = 0. \end{aligned}$$

3.14. D/L/R present one result about the concavity and the monotonicity of value functions and the optimal policy.

Now we prove some structural results for the general model where X has an arbitrary distribution, and where the VI is given by (3.20).

Theorem 3.14 *Assume $v_0(0) = 0$. Then*
(a) $V(t, y)$ is increasing in t.
(b) $V(t, y)$ is increasing in y, if v_0 is increasing.

Proof. (a_1) We prove by induction on n that $t \to V_n(t, y)$ is increasing for all y.

Firstly,

$$t \to V_0(t, y) = \begin{cases} v_0(y), & \text{if } t > 0, \\ 0, & \text{if } t = 0, \end{cases}$$

is increasing, as $v_0(y) \geq 0$.

Now assume, $t \to V_{n-1}(t, y)$ is increasing, $\forall y$. Then, as $t \to (t - x)^+$ is increasing for all x, and as an increasing function of an increasing function is also increasing, $t \to V_{n-1}((t - X)^+, y)$ is increasing. This implies that $t \to EV_{n-1}((t - X)^+, y - a)$ is increasing for all y, a. As the supremum of increasing functions is increasing, $t \to V_n(t, y)$ is increasing on $(0, t_{max}]$ by (3.20).

Moreover, for $t > 0$ we have, as $u(0) \geq 0$ and $(t - X)^+ \geq 0$,

$$\begin{aligned} V_n(t, y) &\geq u(0) + EV_{n-1}((t - X_1)^+, y) \\ &\geq V_{n-1}(0, y) = 0 \\ &= v_0(y) = V_n(0, y), \end{aligned}$$

hence $t \to V_n(t, y)$ is increasing on $[0, t_{max}]$.

(a_2) As the limit of a sequence of increasing functions on S is increasing, then $t \to V(t, y)$ is increasing.

(b_1) Firstly,

$$y \to V_0(t, y) = \begin{cases} v_0(y), & \text{if } t > 0, \\ 0, & \text{if } t = 0, \end{cases}$$

is increasing, as $v_0(y) \geq 0$. Moreover, $y \to V_n(0, y) = v_0(y)$ is increasing.

Now assume, $y \to V_{n-1}(t', y)$ is increasing, $\forall t'$. Assume $t > 0$. Then, by (3.20) and Lemma 2.6, and by the justification used in (a_1), $y \to V_n(t, y)$ is increasing on $[0, K]$.

(b_2) As the limit of a sequence of increasing functions on S is increasing, then $y \to V(t, y)$ is increasing. □

3.3.5 Allocation time as renewal process: Continuous-State Version

This is the last case: instead of discrete-state, we have continuous-state, therefore

$$S := [0, t_{max}] \times [0, K].$$

Also, instead $X \sim geo(p)$, $p \in (0, 1)$, we have $X \sim exp(\lambda)$, $\lambda \in \mathbb{R}^+$ or other distribution, e.g. *gamma distribution*.

3.4 Non-stationary problems

Some of the previous results remain valid even when the probability of an opportunity, the management cost, the utility function and the discounting change from period to period - *Non-stationary case.*

We treat the discrete and the continuous state case jointly.

3.4.1 The Discrete-Time Discrete and Continuous-State Version

S, $s_\nu \in S$, A, $a_\nu \in A$, $D(s)$, and D are defined as in the stationary CM.

The independent random variables X_ν and the set M are defined as in the stationary case, but now the distribution of X_ν may depend on ν. Then $p_n := P(X_{N-n+1} = 1)$, $1 \leq n \leq N$ is the probability of an opportunity at *stage* n, e.g. $p_n = \alpha^{N-n}$, $\alpha \in (0,1)$. Note that if $n \to p_n$ is increasing this means that at later times ν smaller chances for opportunities occur. Also we put $q_n := 1 - p_n$.

$$T : D \times M \longrightarrow S$$

is the *transition function* independent of n, given by

$$s_{\nu+1} = s_\nu - a_\nu \cdot X_{\nu+1}.$$

We also allow that the discount factor, the management cost and the utility function depend on n, i.e.

$$\tilde{r}_n(s,a,x) := \begin{cases} u_n(a), & \text{if } x = 1, \\ c_n & \text{if } x = 0, \end{cases}$$

for functions u_n and $c_n \in \mathbb{R}$. Therefore

$$r_n(s,a) = p_n \cdot u_n(a) + q_n \cdot c_n,$$

is the one-stage reward.

V_0, and $\beta_n > 0$ are arbitrary. V_0 is assumed to be increasing.

Note also that if p_n, c_n, u_n and β_n do not depend on n we return to the stationary case.

Now let $V_n(s)$ denote the maximal expected n-stage reward, if the initial capital is s and *before it* is known whether or not at time $\nu = 0$ an opportunity arises. Using Theorems 2.12 and 2.8, we obtain the following result:

Theorem 3.15 (Continuous-State Case) *(a) V_N may be computed recursively by the value iteration*

$$\begin{aligned} V_n(s) &= q_n \cdot c_n + q_n \cdot \beta_n \cdot V_{n-1}(s) + p_n \cdot \sup_{0 \le a \le s} \{u_n(a) + \beta_n \cdot V_{n-1}(s-a)\} \\ &=: q_n \cdot c_n + q_n \cdot \beta_n \cdot V_{n-1}(s) + p_n \cdot \sup_{0 \le a \le s} \{W_n(s,a)\}, \end{aligned} \tag{3.23}$$

where $W_n(s,a) := u_n(a) + \beta_n \cdot V_{n-1}(s-a)$.

(b) **(Optimality Criterion)** *If $f_n(s)$ is a maximum point of $a \to W_n(s,a)$, for $1 \le n \le N$ and $s \in S$, then the policy $(f_n)_N^1$ is optimal for DP_N.*

(c) If u_n and V_0 are continuous, then there exists a smallest maximizer f_n at stage n, and V_n is continuous for all n. □

Remarks

3.15. The results (a) and (b) hold also in the discrete-state case, *sup* can be replaced by *max*.

3.16. It is easy to see that under the same or similar conditions, the Theorems 3.7, 3.8, 3.9 and 3.10, about monotonicity, convexity, and concavity of V_n and f_n hold also in the non-stationary case.

3.4.2 Closed Solution

Example 1: Let be $u_1(a) := b_0 \cdot a^\alpha$ and $V_0(s) := d_0 \cdot s^\alpha + e_0$ for some $d_0 \in \mathbb{R}_+$, $e_0 \in \mathbb{R}$, $b_0 \in \mathbb{R}^+$ and $\alpha \in (0,1)$. Then the VI reads for $n = 1$

$$\begin{aligned} V_1(s) &= q_1 \cdot c_1 + q_1 \cdot \beta_1 \cdot (d_0 \cdot s^\alpha + e_0) + p_1 \cdot \sup_{0 \le a \le s} \{b_0 \cdot a^\alpha + \beta_1 \cdot (e_0 + d_0 \cdot \\ &\quad \cdot (s-a)^\alpha)\}, \\ &= q_1 \cdot c_1 + \beta_1 \cdot e_0 + q_1 \cdot \beta_1 \cdot d_0 \cdot s^\alpha + p_1 \cdot \sup_{0 \le a \le s} \{b_0 \cdot a^\alpha + \beta_1 \cdot d_0 \cdot \\ &\quad \cdot (s-a)^\alpha\}, \end{aligned}$$

for $d_0 > 0$, we have

$$\begin{aligned} & g(a) := b_0 \cdot a^\alpha + \beta_1 \cdot d_0 \cdot (s-a)^\alpha, \\ g'(a) = 0 \quad \Leftrightarrow \quad & \alpha \cdot b_0 \cdot a^{\alpha-1} - \alpha \cdot \beta_1 \cdot d_0 \cdot (s-a)^{\alpha-1} = 0. \end{aligned}$$

As g is concave by Theorem 2.6, as in the stationary case we obtain

$$a^* := s \Big/ \left(1 + \left(\frac{\beta_1 \cdot d_0}{b_0}\right)^\rho\right)$$

as the unique maximum point of g, for $d_0 \ge 0$, where

$$\rho := \frac{1}{1-\alpha}.$$

Moreover, we get

$$V_1(s) = q_1 \cdot c_1 + \beta_1 \cdot e_0 + \beta_1 \cdot q_1 \cdot d_0 + p_1 \cdot \left(1 + (\frac{\beta_1 \cdot d_0}{b_0})^\rho\right)^{1-\alpha} \cdot s^\alpha.$$

Let be

$$\begin{aligned} d_1 &:= \beta_1 \cdot q_1 \cdot d_0 + p_1 \cdot \left(1 + \left(\frac{\beta_1 \cdot d_0}{b_0}\right)^\rho\right)^{1-\alpha}, \\ e_1 &:= q_1 \cdot c_1 + \beta_1 \cdot e_0, \end{aligned}$$

then

$$V_1(s) = e_1 + d_1 \cdot s^\alpha,$$

and

$$s \longrightarrow f_1(s) = \frac{s}{1 + \left(\frac{\beta \cdot d_0}{b_0}\right)^\rho}.$$

More formally, if $u_n(a) = b_{n-1} \cdot a^\alpha, \alpha \in (0,1)$, and $b_{n-1} \in \mathbb{R}^+$, for all $n \in \mathbb{N}$ we obtain from the VI by induction on n, using the maximum point a^*, the following result:

Proposition 3.8 *Assume that for some $\alpha \in (0,1)$, $u_n(a) = b_{n-1} \cdot a^\alpha$, for all n, $a \in A$, $V_0(s) = d_0 \cdot s^\alpha, s \in S$, for some $b_{n-1} \in \mathbb{R}^+$, for some $d_0 \in \mathbb{R}_+$ and for some $e_0 \in \mathbb{R}$. Then the following holds:*

(i) V_n is of the form

$$V_n(s) = d_n \cdot s^\alpha + e_n, \qquad s \in S, n \in \mathbb{N},$$

for some $d_n \in \mathbb{R}^+$ which satisfy the recursion

$$d_n = \beta_n \cdot q_n \cdot d_{n-1} + p_n \cdot \left(1 + \left(\frac{\beta_n \cdot d_{n-1}}{b_{n-1}}\right)^\rho\right)^{1-\alpha}; \qquad n \in \mathbb{N}, b_{n-1} \in \mathbb{R}^+,$$

and

$$e_n = q_n \cdot c_n + \beta_n \cdot e_{n-1} \qquad \text{for } e_{n-1} \in \mathbb{R}.$$

(ii)

$$s \longrightarrow f_n(s) := \frac{s}{1 + \left(\frac{\beta_n \cdot d_{n-1}}{b_{n-1}}\right)^\rho}$$

is the unique maximizer at stage $n \in \mathbb{N}$. □

Note that e_n has the following closed solution

$$e_n = \sum_{i=0}^{n-1} (q \cdot c)_{i+1} \cdot \beta^i + \beta^n e_0, \qquad \text{for } e_0 \in \mathbb{R}.$$

Lemma 3.3 *Let $I \subset \mathbb{R}$. Assume that $H_n : I \to \mathbb{R}$, $n \in \mathbb{N}_0$, is increasing and $H_n \leq H_{n+1}$ $[H_n \geq H_{n+1}]$ for all n. If $x_0 \in I$ and $x_{n+1} := H_n(x_n)$, $n \in \mathbb{N}_0$ and $x_1 \geq x_0$ $[x_1 \leq x_0]$. Then (x_n) is increasing [decreasing].*

Proof. Case 1: $x_1 \geq x_0$.

We show by induction on n that

$$x_{n+1} \geq x_n. \tag{3.24}$$

At first, (3.24) holds for $n = 0$ by assumption. Now assume, that (3.24) holds for some $n \in \mathbb{N}_0$. Then $x_{n+2} = H_{n+1}(x_{n+1}) \geq H_{n+1}(x_n)$, as H_{n+1} is increasing. As $H_{n+1} \geq H_n$, we obtain $H_{n+1}(x_n) \geq H_n(x_n) = x_{n+1}$. Altogether we have $x_{n+2} \geq x_{n+1}$. Thus (3.24) holds for $n+1$.

Case 2: $x_1 \leq x_0$

The proof of this case goes through exactly as in the case 1.

Now the proof is complete. □

As

$$\begin{aligned} d_n &= \beta_n \cdot q_n \cdot d_{n-1} + p_n \cdot \left(1 + \left(\frac{\beta_n \cdot d_{n-1}}{b_{n-1}}\right)^\rho\right)^{1-\alpha} \\ &=: H_n(d_{n-1}), \end{aligned}$$

we have

$$H_n(x) = \beta_n \cdot q_n \cdot x + p_n \cdot \left(1 + \left(\frac{\beta_n \cdot x}{b_{n-1}}\right)^\rho\right)^{1-\alpha}.$$

Lemma 3.4 *Assume that $b_n = b \leq 1$, for all $n \in \mathbb{N}_0$, and $\beta_n = \beta$, for all $n \in \mathbb{N}$, and that (p_n) is increasing [decreasing]. Then $H_n(x) \leq H_{n+1}(x)$ $[H_n(x) \geq H_{n+1}(x)]$ for all $x \in \mathbb{R}_+$.*

Proof. We will show that the first assertion holds. We have

$$H_n(x) \leq H_{n+1}(x)$$

for all $x \in \mathbb{R}_+$.

Let be

$$\delta := \left(1 + \left(\frac{\beta \cdot x}{b}\right)^\rho\right)^{1-\alpha},$$

therefore

$$\beta \cdot q_n \cdot x + p_n \cdot \delta \leq \beta \cdot q_{n+1} \cdot x + p_{n+1} \cdot \delta$$

$$\beta \cdot x + p_n \cdot (\delta - \beta \cdot x) \leq \beta \cdot x + p_{n+1} \cdot (\delta - \beta \cdot x)$$

$$p_n \cdot (\delta - \beta \cdot x) \leq p_{n+1} \cdot (\delta - \beta \cdot x).$$

As $p_n \leq p_{n+1}$, it is enough to show that

$$\delta - \beta \cdot x \geq 0, \tag{3.25}$$

for all $x \in \mathbb{R}_+$. Then as

$$\left(1 + \left(\frac{\beta \cdot x}{b}\right)^\rho\right)^{1-\alpha} - \beta \cdot x \geq 0$$

$$\left(1 + \left(\frac{\beta \cdot x}{b}\right)^\rho\right)^{1-\alpha} \geq \beta \cdot x$$

$$1 + \left(\frac{\beta \cdot x}{b}\right)^\rho \geq (\beta \cdot x)^{\frac{1}{1-\alpha}}$$

$$1 + \left(\frac{\beta \cdot x}{b}\right)^\rho \geq (\beta \cdot x)^\rho,$$

we see that (3.25) is always positive, as $(\beta \cdot x/b)^\rho \geq (\beta \cdot x)^\rho$. Thus the first assertion is proved.

The second assertion is proved exactly as the first assertion. □

Applying Lemmas 3.3 and 3.4, and observing that H_n is increasing we obtain the following result:

Theorem 3.16 *If (p_n). is increasing [decreasing] and $b_n = b \leq 1$, for all $n \in \mathbb{N}_0$, and $\beta_n = \beta$ for all $n \in \mathbb{N}$, then (d_n) is increasing [decreasing].* □

Example 2: Now we assume that $S = A = (0, K]$, $D(s) = (0, s)$, $u_1(a) = b_0 \cdot \ln a$, $V_0(s) = d_0 \cdot \ln s + e_0$, for some $d_0 \in \mathbb{R}^+$, $e_0 \in \mathbb{R}$, $a \in (0, s)$, $b_0 \in \mathbb{R}^+$, and $s \in (0, K]$. Then the VI reads for $n = 1$

$$\begin{aligned} V_1(s) &= q_1 \cdot c_1 + \beta_1 \cdot q_1 \cdot (e_0 + d_0 \cdot \ln s) + p_1 \cdot \sup_{0<a<s} \{b_0 \cdot \ln a + \beta_1 \cdot (e_0 + \\ &\quad + d_0 \cdot \ln(s - a))\} \\ &= q_1 \cdot c_1 + \beta_1 \cdot e_0 + \beta_1 \cdot q_1 \cdot d_0 \cdot \ln s + p_1 \cdot \sup_{0<a<s} \{b_0 \cdot \ln a + \beta_1 \cdot d_0 \cdot \\ &\quad \cdot \ln(s - a)\}. \end{aligned}$$

As $d_0 > 0$, then

$$g(a) = b_0 \cdot \ln a + \beta_1 \cdot d_0 \cdot \ln(s - a).$$

It has a derivative

$$g'(a) = \frac{b_0}{a} - \frac{\beta_1 \cdot d_0}{s-a}, \qquad 0 < a < s.$$

As g is strictly concave, it attains its unique maximum a^*, iff

$$a^* = \frac{s}{1 + \frac{\beta_1 \cdot d_0}{b_0}}.$$

Then

$$s \longrightarrow f_1(s) := \frac{s}{1 + \frac{\beta_1 \cdot d_0}{b_0}}.$$

Thus

$$\begin{aligned} V_1(s) &= q_1 \cdot c_1 + \beta_1 \cdot e_0 + \beta_1 \cdot q_1 \cdot d_0 \cdot \ln s + p_1 \cdot \ln \frac{s}{1 + \frac{\beta_1 \cdot d_0}{b_0}} + \\ &\quad + p_1 \cdot \beta_1 \cdot d_0 \cdot \ln \left(s - \frac{s}{1 + \frac{\beta_1 \cdot d_0}{b_0}} \right) \\ &= q_1 \cdot c_1 + p_1 \cdot \ln \frac{\left(\frac{\beta_1 \cdot d_0}{b_0}\right)^{\frac{\beta_1 \cdot d_0}{b_0}}}{\left(1 + \frac{\beta_1 \cdot d_0}{b_0}\right)^{1 + \frac{\beta_1 \cdot d_0}{b_0}}} + \beta \cdot e_0 + (\beta_1 \cdot d_0 + p) \cdot \ln s. \end{aligned}$$

Put $\lambda_0 := \left(\frac{\beta_1 \cdot d_0}{b_0}\right)^{\frac{\beta_1 \cdot d_0}{b_0}} / \left(1 + \frac{\beta_1 \cdot d_0}{b_0}\right)^{1 + \frac{\beta_1 \cdot d_0}{b_0}}$, then

$$V_1(s) = q_1 \cdot c_1 + p_1 \cdot \ln \lambda_0 + \beta_1 \cdot e_0 + \left(p_1 + \frac{\beta_1 \cdot d_0}{b_0} \right) \cdot \ln s.$$

Therefore

$$V_1(s) = e_1 + d_1 \cdot \ln s, \qquad d_1 > 0,$$

where

$$e_1 = q_1 \cdot c_1 + p_1 \cdot \ln \lambda_0 + \beta_1 \cdot e_0,$$

$$d_1 = p_1 + \frac{\beta_1 \cdot d_0}{b_0},$$

and

$$s \longrightarrow f_1(s) = \frac{s}{1 + \frac{\beta_1 \cdot d_0}{b_0}}.$$

More formally, if $u_n(a) = b_{n-1} \cdot \ln a$, and $b_{n-1} \in \mathbb{R}^+$, for all $n \in \mathbb{N}$, we obtain the following result:

Proposition 3.9 *Assume that $S = A = (0, K]$, $D(0, s)$, $u_n(a) = b_{n-1} \cdot \ln a$, for all $n \in \mathbb{N}$, $a \in (0, s)$, $V_0(S) = d_0 \cdot \ln s + e_0$, $s \in (0, K]$, for some $d_0 \in \mathbb{R}^+$, for some $e_0 \in \mathbb{R}$ and for some $b_{n-1} \in \mathbb{R}^+$. Then the following holds:*

(i) V_n is of the form

$$V_n(s) = d_n \cdot \ln s + e_n, \qquad n \in \mathbb{N}, s \in (0, K],$$

for some $d_n > 0$ which satisfy the recursion

$$d_n = p_n + \frac{\beta_n \cdot d_{n-1}}{b_{n-1}}, \qquad n \in \mathbb{N},$$

and

$$e_n = q_n \cdot c_n + p_n \cdot \ln \lambda_{n-1} + \beta_n \cdot e_{n-1}, \qquad e_{n-1} \in \mathbb{R},$$

$$\lambda_{n-1} = \frac{\left(\frac{\beta_n \cdot d_{n-1}}{b_{n-1}}\right)^{\frac{\beta_n d_{n-1}}{b_{n-1}}}}{\left(1 + \frac{\beta_n \cdot d_{n-1}}{b_{n-1}}\right)^{1+\frac{\beta_n d_{n-1}}{b_{n-1}}}}, \qquad b_{n-1} \in \mathbb{R}^+.$$

(ii)

$$s \longrightarrow f_n(s) := \frac{s}{1 + \frac{\beta_n \cdot d_{n-1}}{b_{n-1}}}$$

is the maximizer at stage $n \in \mathbb{N}$. □

3.5 Flow Chart For General Allocation Process

In the discussion of the reduction of problems from mathematical formulation to computer code, we shall explain the process with more details so that it can be easily programmed by someone who is not familiar with the original mathematical problem or technique. Later the flow chart will follow a similar way and will be assumed to be self-explanatory (cf. Appendix B).

Step1. The basic code will use the value iteration (3.1) to compute a tabular function $s \to V_n(s)$ using the whole function V_{n-1}. In order to begin the initial step of calculation, we must store V_0 as a tabular function. We can now have the computer determine $V_1(s)$ using V_0 in the same manner as it determines $V_n(s)$ from V_{n-1}.

Step 2. The index n will denote the number of stages that we are considering. The index n will be increased as the calculation progresses (see step 16). We start with $n = 1$.

Step 3. We shall compute a table of values representing the function $V_n(s)$ at discrete points s. The initial argument for which we compute $V_n(s)$ is $s = 0$.

After computing and storing $V_n(0)$, we compute $V_n(1)$, and then $V_n(2)$, and so on, until the table is complete.

Step 4. The cell *max* will contain the "best return so far" as we test various actions seeking that which maximizes $a \rightarrow W_n(s,a)$. Setting this cell initially to a large negative number (denoted by $-\infty$, we can use, in stage n, $W_n(s,0)$,) we guarantee that the smallest action tested will be accepted as the "best so far."

Step 5. We use $f_n(s)$ to denote our *smallest* allocation decision given a quantity of initial resource s at stage n. Since, to begin with $n = 1$ and $s = 0$, we test 0 as the initial candidate for $f_n(s)$.

Step 6. We have now specified the stage n, the resource s and the allocation $f_n(s)$. Using the reward function and the optimal $(n-1)$-stage return, $V_{n-1}(s-a)$, we compute the total return associated with the given decision in states and we store this number in location *aux*.

Step 7. Compare this number with the number in cell *max*, the best of all previously tested actions for this particular state and stage. If the current decision yields a smaller return than for some previous one, go to step 9. If this is the best allocation decision tested thus far, perform step 8.

Step 8. Replace the contents of cell *max* by the greater return that has just been stored in cell *aux*. Cell *maxa* is to contain the "best action so far," hence we place a in cell *maxa*.

Step 9. Having examined the effect of the allocation of quantity a to the $n-$th stage, we now prepare to test the larger allocation $a + 1$.

Step 10. Is this allocation greater than our resource s ? If so, this decision is not admissible and we go to step 11. If $a + 1$ is an admissible decision, return to step 6 to evaluate this decision, and to compare it with the previous decisions.

Step 11. We have now compared all decisions for a specific initial resource s. Store the maximum attainable return, $V_n(s)$, and the smallest decision yielding this return, $f_n(s)$.

Step 12. Increase the initial resource by 1. We now have a new problem involving the same number of states but a greater initial resource.

Step 13. If the new problem involves a resource greater than K, we have completed the computation of the table of values of $V_n(s)$ and go on to step 14. If this new s is admissible, we begin the entire maximization process over again by returning to step 4.

Step 14. Now we have the results, for stage n which should be stored in cells $V_n(s)$ and $f_n(s)$.

Step 15. From this point on, we shall use the newer table, $V_n(s)$, to compute $V_{n+1}(s)$.

Step 16. We now proceed to the next stage and prepare to solve a family of problems involving the same number of states.

Step 17. If we have just computed $V_n(s)$, n is increased by 1. If $n + 1$ is greater than the horizon N, we stop and declare the calculation completed and we go on to step 18. If the new n is less than or equal N, we return to step 3.

Step 18. Now we show the result in form of an Output.

This completes our analysis of the actual operations within the computer.

Appendix A

Used Notations and its Meanings

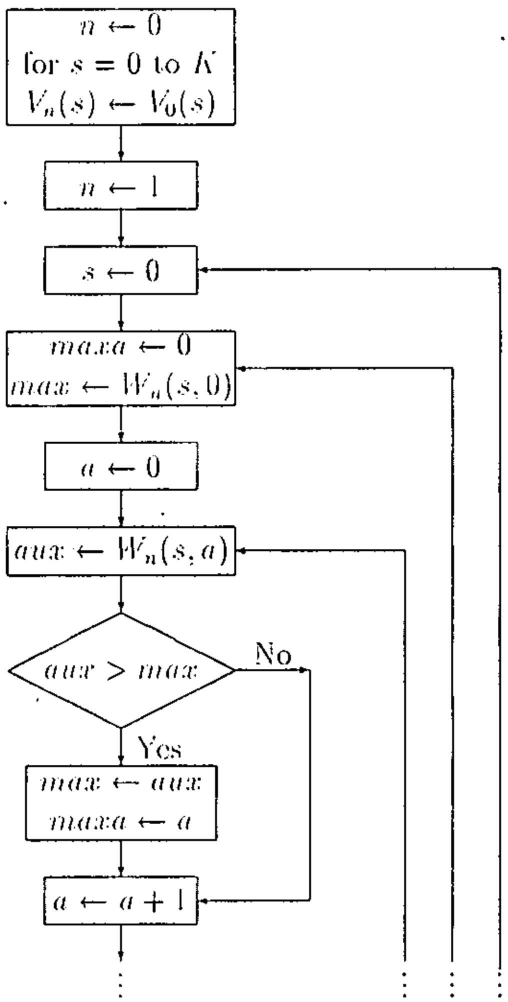
$(a_i)_1^N = (a_1, a_2 \ldots, a_N)$
$\mathbb{R}_+ = [0, +\infty[$
$\mathbb{R}^+ =]0, +\infty[$
$\mathbb{N}_m = \{1, 2, \ldots, m\}$
$\mathbb{N}_{0,m} = \{0, 1, \ldots, m\}$
$|A|$ = number of elements of the set A.
$A := B$, A is defined to be B.
$A =: B$, B is defined to be A.
$(0, 1] =]0, 1]$
If A is some set, then A^2 means the cartesian product $A \times A := \{(a_1, a_2) : a_1, a_1 \in A\}$.
EX = expectation of random variable X.
$x^+ = max(0, x)$ for $x \in \mathbb{R}$
For functions from B to G we used in general the notations $\left\{ \begin{array}{l} x \rightarrow f(x) \\ B \rightarrow G \end{array} \right.$ or $f(\cdot)$:

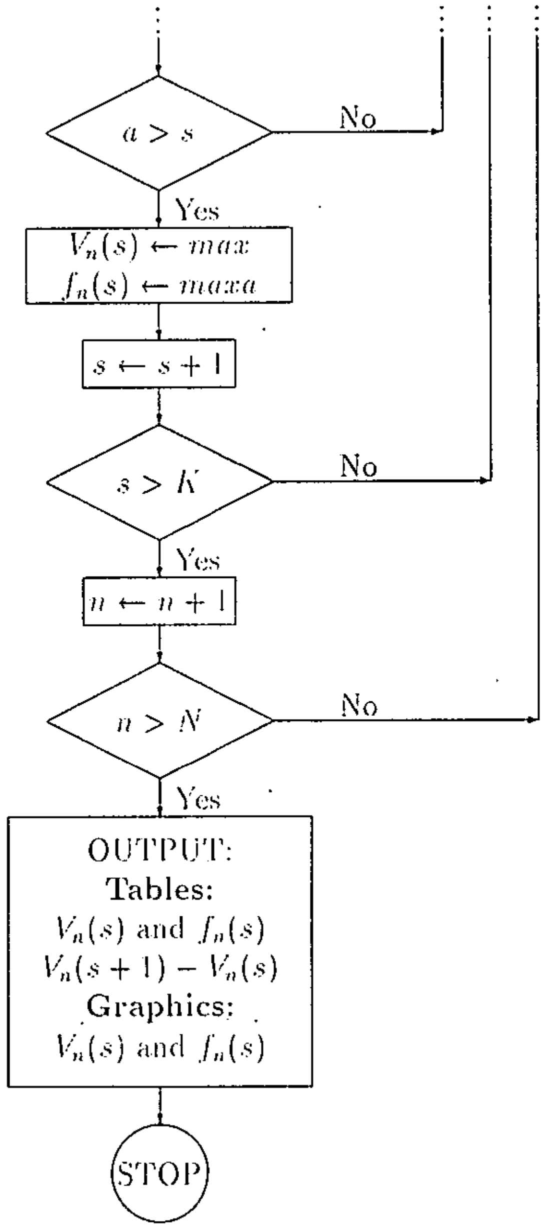
Note that $f(x)$ denotes only the value of f at x, not the function f.

Appendix B

Flow Chart

Numerical Inputs:	p	-	probability of occurance of an opportunity
	K	-	units of capital available
	N	-	horizon, β - discount factor
	c	-	management cost




$a > s$
No
Yes
$V_n(s) \leftarrow max$
$f_n(s) \leftarrow maxa$
$s \leftarrow s + 1$
$s > K$
No
Yes
$n \leftarrow n + 1$
$n > N$
No
Yes
OUTPUT:
Tables:
$V_n(s)$ and $f_n(s)$
$V_n(s+1) - V_n(s)$
Graphics:
$V_n(s)$ and $f_n(s)$
STOP

Appendix C

List of Programs

During the work we wrote numerous programs in TURBO-PASCAL VERSION 6.0:

1. Program to calculate $V_n(s)$ and $f_n(s)$, for the discrete-time discrete-state version, with inputs: N, K, p, β, d and c.

   Files: THESIS1.PAS and THESIS8.PAS.

2. Program to calculate $V_n(s,p)$ and $f_n(s,p)$, for the discrete-time discrete-state version, with inputs: N, K, β, d *stage* and c.

   File: THESIS2.PAS.

3. Program to calculate the values of the function $K(\beta) := (1-\beta\cdot q)^\rho - (\beta\cdot p)^\rho$ where $\rho := 1/(1-\alpha)$, for $p = 0.1 - 1.0$ and $\beta = 0.1 - 1.0$, with input: α.

   File: THESIS3.PAS.

4. Program to calculate the values of the function $h := (1 + (\beta \cdot d_0)^\rho)/(1 - \beta \cdot q)^\rho$, for $p = 0.1 - 1.0$ and $\beta = 0.1 - 1.0$, with inputs: α and d_0.

   File: THESIS4.PAS.

5. Program to calculate $V_n(s,p)$ and $f_n(s,p)$, for $p = 0.25, 0.50, 0.75, 1.00$, with inputs: N, K, β, d, c and *stage*.

   File: THESIS5.PAS.

6. Program to calculate $V_n(s)$ and $f_n(s)$, for the non-stationary case, discrete-time discrete-state version, with inputs: N, K, β, d, p and c.

   File: THESIS6.PAS.

7. Program to calculate $V(t,y)$ and $f(t,y)$, for the allocation time as renewal process discrete-state version, with inputs: t_{max}, K, p, d and α.

   File: THESIS7.PAS.

Bibliography

[1] Bellman, R. E., and S. E. Dreyfus. *"Applied Dynamic Programming"*, 1962, p. 3-92

[2] Denardo, E. V., *"Dynamic Programming: Models and Applications"*, 1983, p. 34-63

[3] Derman, G., G. J. Lieberman. and S. M. Ross. *"A Stochastic Sequential Allocation Model"*, Operation Research, Vol. 23, 1975, p. 1120-1130

[4] Deuermeyer, B. L., and G. L. Curry. *"On a Language for Discrete Dynamic Programming and a Microcomputer Implementation"*, Computers Operations Research, Vol. 16, Nr. 1, 1989, p. 1-11

[5] Hadley, G., *"Nonlinear and Dynamic Programming"*, 1964, p. 83-87, 423-432

[6] Hinderer, K., *"Looking at the Structure of Dynamic Programs with PC Support - a Review"*, Methods of Operations Research, 1988, p. 75-85

[7] Hinderer, K., *"Increasing Lipschitz Continuous Maximizers of some Dynamic Programs"*, Annals of Operations Research, Nr. 29, 1991, p. 565-586

[8] Hinderer, K., *"Dynamic Optimization"*, 1993, p. 19-49. To be published.

[9] Hinderer, K., *"Stochastik für Informatiker und Ingenieure"*, 1991, p. 10.1-10.5. Script, unpublished.

[10] Press, H. W., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *"Numerical Recipes - The Set of Scientific Computing"*, 1986, p. 245-259

[11] Roberts, A. W., and D. E. Varberg, *"Convex Functions"*, 1973, p. 1-36

[12] Sasieni, M., A. Yaspan, and L. Friedman, *"Operations Research - Methods and Problems"*, 1960, 2th Edition, p. 183-185, 270-274

[13] Sharma, J. K., "*Mathematical Models in Operations Research*", 1989, p. 491-494

[14] Smith, D. K., "*Dynamic Programming - A Practical Introduction*", 1991, p. 11-18, 37-50

[15] Sniedovich, M., "*Dynamic Programming*", 1991, p. 283-305, 307-321