CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION CFSTI
DOCUMENT MANAGEMENT BRANCH 410.11

LIMITATIONS IN REPRODUCTION QUALITY

ACCESSION # AD 604283

- [X] 1. WE REGRET THAT LEGIBILITY OF THIS DOCUMENT IS IN PART UNSATISFACTORY. REPRODUCTION HAS BEEN MADE FROM BEST AVAILABLE COPY.
- [ ] 2. A PORTION OF THE ORIGINAL DOCUMENT CONTAINS FINE DETAIL WHICH MAY MAKE READING OF PHOTOCOPY DIFFICULT.
- [ ] 3. THE ORIGINAL DOCUMENT CONTAINS COLOR, BUT DISTRIBUTION COPIES ARE AVAILABLE IN BLACK-AND-WHITE REPRODUCTION ONLY.
- [ ] 4. THE INITIAL DISTRIBUTION COPIES CONTAIN COLOR WHICH WILL BE SHOWN IN BLACK-AND-WHITE WHEN IT IS NECESSARY TO REPRINT.
- [ ] 5. LIMITED SUPPLY ON HAND: WHEN EXHAUSTED, DOCUMENT WILL BE AVAILABLE IN MICROFICHE ONLY.
- [ ] 6. LIMITED SUPPLY ON HAND: WHEN EXHAUSTED DOCUMENT WILL NOT BE AVAILABLE.
- [ ] 7. DOCUMENT IS AVAILABLE IN MICROFICHE ONLY.
- [ ] 8. DOCUMENT AVAILABLE ON LOAN FROM CFSTI (TT DOCUMENTS ONLY).
- [ ] 9.

NBS 9/64

PROCESSOR: BL

SOME PROBLEMS IN THE THEORY OF DYNAMIC PROGRAMMING

Richard Bellman

P-455

2 November 1953

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA

SOME PROBLEMS IN THE THEORY OF DYNAMIC PROGRAMMING

Richard Bellman

Summary: The theory of dynamic programming treats problems involving multi-stage processes by means of a transformation of the problem from the space of decisions to the space of functions. This is accomplished by deriving a functional equation whose solution is equivalent to the solution of the original problem. To illustrate this approach most clearly, free of extraneous analytic details, ~~we consider~~ a simple but nontrivial multi-stage investment problem. We show how exact solutions may be obtained in some cases, approximate solutions in others, and how these approximate solutions may be used to obtain more accurate solutions in the general case. Of particular importance is the decrease in the number of independent variables made possible by this approach. This is not only important from the theoretical standpoint, but is also of great value in reducing the cost in time and effort of numerical computation.

1. Introduction.

The purpose of this paper is to provide an introduction to a class of mathematical techniques that are useful in treating a variety of problems arising in the planning of multi-stage processes. These are programming problems, to use the currently popular terminology, and the adjective "dynamic" emphasizes that in these problems time plays an important role. This characterization of our subject matter is not merely one of nomenclature--the problems before us are of a conceptually distinct type. Moreover, the mathematical techniques employed prove to be especially powerful for the resolution of problems of this type.

The multi-stage problems in which we are interested are composed of sequences of operations in which the outcome of preceding operations may be used to guide the course of future operations. Two types of operations may be distinguished immediately, those in which the outcome is predictable on the basis of a probability distribution, and those in which the outcome is completely determined. Depending upon one's point of view, either type may be considered to be an approximation to the reality represented by the other.

Since most of the problems which arise are of an entirely novel character, frequently offering formidable mathematical difficulties, we have restricted ourselves to the consideration of a simple yet important problem in order not to obscure our techniques by extraneous analytic and algebraic complications.

The basic idea underlying our analysis is that of replacing the decision problem by a functional equation. We show then that the solution of the functional equation yields a solution of the original programming problem. It is precisely at this point that the question of uniqueness of solution acquires an economic as well as a mathematical importance.

Using the functional equation approach we consider a very simple version of an optimal allocation problem. We begin with a statement of the problem and then contrast the classical approach with the dynamic programming approach. We next supply an existence and uniqueness proof, and then derive some important properties of the solution.

Following this we formulate the stochastic version of the same problem, and then in the concluding section discuss the essential features of the mathematical techniques employed.

2. The Formulation of a Problem of Optimal Resource Allocation.

A very simple problem relating to the allocation of resources is the following: We have x dollars which may be split into two parts, y and x-y. From y we obtain a return of g(y), from x-y a return of h(x-y). How does one choose y so as to maximize the total return?

The analytic problem is that of finding the maximum of $g(y) + h(x-y)$ subject to the constraint $0 \leq y \leq x$.

Let us now complicate the problem by changing this one-stage problem into an N-stage problem in the following way: At the end of the first stage let us assume that we have left, as a result of our division into y and x-y, the quantity of money $ay + b(x-y)$, where $0 \leq a,b \leq 1$, and we are to continue this process for N-1 additional stages. How does one allocate at each stage in order to maximize the total return?

There is no difficulty in setting this problem up in classical form. Let $y_1, y_2, \cdots, y_N$ be the sequence of choices at each successive stage. The total return will be

$$(2.1) \qquad R(y_1, y_2, \cdots, y_N) = \sum_{i=1}^{N} g(y_i) + \sum_{i=1}^{N} h(x_i - y_i)$$

where

$$(2.2) \qquad \begin{aligned} x_1 &= x, & 0 &\leq y_1 \leq x_1, \\ x_2 &= ay_1 + b(x_1 - y_1), & 0 &\leq y_2 \leq x_2 \\ &\vdots \\ x_N &= ay_{N-1} + b(x_{N-1} - y_{N-1}), & 0 &\leq y_N \leq x_N \end{aligned}$$

The problem is to maximize R subject to the above constraints. Since several of the y may be 0 or x_1, endpoints of their allowable intervals, any naive application of calculus is a bit hazardous.

3. <u>The Functional Equation Approach to Optimal Allocation.</u>

Since the above formulation presents all the unpleasant features usually involved in a maximization problem over an N-dimensional region, let us cast about for an alternative approach. The clue to another formulation is the observation that at any stage it is necessary only to choose the corresponding y in order to continue the process.

Let us set

(3.1) $f_N(x)$ = total return obtained from an N-stage process starting with x dollars and using an optimal procedure.

Whatever division of x into y and x-y is optimal on the first step, the remaining amount of money $ay + b(x-y)$ must be treated by an optimal procedure for the next N-1 steps if a maximum return is to be obtained for N steps. Thus the return for the N-step process due to this initial division is $g(y) + h(x-y) + f_{N-1}(ay + b(x-y))$. Since an optimal selection of y subject to $0 \leq y \leq x$ would maximize this, we must have the equations

(3.2)
$$f_1(x) = \max_{0 \leq y \leq x} \left[g(y) + h(x-y) \right],$$
$$f_N(x) = \max_{0 \leq y \leq x} \left[g(y) + h(x-y) + f_{N-1}(ay + b(x-y)) \right] \quad (N \geq 2).$$

We have thus reduced the original programming problem to the solution of the system of functional equations given in (3.2).

If we are interested in the case where N is large, we obtain a great simplification by making the approximation, $f_N(x) \cong f_{\infty}(x)$. The case of an unbounded number of operations, which is, of course, meaningless in any practical situation, possesses a very important mathematical property of invariance, because after any finite number of initial operations, there still remains a process with an infinite number of steps.

Introducing

(3.3) $f(x)$ = total return obtained from the infinite process starting with a quantity x and using an optimal policy,

we obtain, formally, the equation

$$(3.4) \qquad f(x) = \operatorname*{Max}_{0 \leq y \leq x} \left[g(y) + h(x-y) + f(ay + b(x-y)) \right]$$

for $x \geq 0$.

The advantage of the above equation over the system in (3.2) lies in the fact that we have now a single unknown function. The disadvantage lies in the difficulty usually associated with infinite processes, namely that of proving existence, uniqueness, and attainability. We shall show in the next section that in this problem this is a minor difficulty.

4. Existence and Uniqueness.

Here we establish the existence and uniqueness of the solution of (3.4), under certain natural assumptions of continuity. The method of successive approximations which we employ is equivalent to the obvious idea of showing that as $N \longrightarrow \infty$ we have $f_N(x) \to f(x)$. There are alternative approximation techniques, also based upon economic principles, which we discuss elsewhere, in [4].

Theorem 1. *If*

(4.1) (a) $g(y)$ *and* $h(y)$ *are continuous and monotone increasing in* $[0,x_0]$ *with* $g(0) = h(0) = 0$, *and*

(b) $$\sum_{n=0}^{\infty} \left[g(a^n x_0) + g(b^n x_0) + h(a^n x_0) + h(b^n x_0) \right] < \infty,$$

there is a unique solution to (3.4) *which is bounded in* $[0,x_0]$ *and is continuous at* $x = 0$ *with the value* 0 *at this point. This solution is actually continuous throughout the whole interval* $[0,x_0]$.

As we shall see from the proof the condition that g and h are monotone increasing is not needed. Use of this condition, however, which is satisfied in any application, simplifies the proof notationally.

Proof: We use the successive approximations defined by (3.2). Since $f_1 \geq 0$ it is clear from (3.2) that $f_2 \geq f_1$, which yields, via an induction, the obvious result that $0 \leq f_1 \leq f_2 < \cdots \leq f_n$. Let $c = \text{Max}(a,b)$ and let us show that

$$(4.2) \qquad f_n(x) \leq \sum_{n=0}^{\infty} \left[g(c^n x) + h(c^n x) \right] .$$

This will be proved inductively, using the fact that the inequality clearly holds for $n = 1$. Since $ay + b(x-y) \leq cx$, we have

$$(4.3) \qquad f_{n-1}(x) \leq \max_{0 \leq y \leq x} \left[g(y) + h(x-y) + \sum_{n=1}^{\infty} \left[g(c^n x) + h(c^n x) \right] \right]$$

$$\leq g(x) + h(x) + \sum_{n=1}^{\infty} \left[g(c^n x) + h(c^n x) \right]$$

$$\leq \sum_{n=0}^{\infty} \left[g(c^n x) + h(c^n x) \right]$$

Hence the sequence $\{f_n(x)\}$ is monotone increasing and bounded, which means that $f_n(x)$ converges to a function $f(x)$. The function satisfies the equation

$$(4.4) \qquad f(x) = \sup_{0 \le y \le x} \left[g(y) + h(x-y) + f(ay + b(x-y)) \right] .$$

In order to prove that the maximum is attained, we must prove that f is continuous. To accomplish this, we use the fact that each f_n is a continuous function of x, and establish the uniform convergence of the series $\sum_{n=0}^{\infty} |f_{n+1} - f_n|$ by means of the following technique:

Let $y_n \equiv y_n(x)$ be a point where the maximum is attained in the expression for f_{n+1}, i.e.,

$$(4.5) \qquad f_{n+1}(x) = \max_{0 \le y \le x} \left[g(y) + h(x-y) + f_n(ay + b(x-y)) \right]$$
$$= g(y_n) + h(x-y_n) + f_n(ay_n + b(x-y_n)).$$

Then we have the following obvious inequalities,

$$(4.6) \qquad f_{n+1}(x) = g(y_n) + h(x-y_n) + f_n(ay_n + b(x-y_n))$$
$$\ge g(y_{n-1}) + h(x-y_{n-1}) + f_n(ay_{n-1} + b(x-y_{n-1})),$$
$$f_n(x) = g(y_{n-1}) + h(x-y_{n-1}) + f_{n-1}(ay_{n-1} + b(x-y_{n-1}))$$
$$\ge g(y_n) + h(x-y_n) + f_{n-1}(ay_n + b(x-y_n)).$$

It follows that

$$(4.7) \qquad f_{n+1} - f_n \geq f_n(ay_{n-1} + b(x-y_{n-1})) - f_{n-1}(ay_{n-1} + b(x-y_{n-1}))$$

$$\leq f_n(ay_n + b(x-y_n)) - f_{n-1}(ay_n + b(x-y_n)).$$

Hence

$$(4.8) \qquad f_{n+1}(x) - f_n(x) \leq \text{Max}\Big[\left|f_n(ay_{n-1} + b(x-y_{n-1})) - f_{n-1}(ay_{n-1} + b(x-y_{n-1}))\right| ,$$

$$\left|f_n(ay_n + b(x-y_n)) - f_{n-1}(ay_n + b(x-y_n))\right|\Big] .$$

Let for $n \geq 1$

$$(4.9) \qquad u_n(x) = \underset{0 \leq y \leq x}{\text{Max}} \left|f_{n+1}(y) - f_n(y)\right| .$$

From (4.8) we obtain

$$(4.10) \qquad \left|f_{n+1}(x) - f_n(x)\right| \leq \text{Max}\left[u_n(cx), u_n(cx)\right] = u_n(cx).$$

Hence $u_{n+1}(x) \leq u_n(cx)$ for $n \geq 1$, while for $n = 0$, we have

$$(4.11) \qquad u_0(x) = \underset{0 \leq y \leq x}{\text{Max}} f_1(y) \leq g(x) + h(x).$$

From this we conclude that the series $\sum u_n$ is majorized by the series $\sum\left[g(c^n x) + h(c^n x)\right]$. Hence the series $\sum u_n$ converges uniformly in the interval $[0,x_0]$, which means that f_n converges uniformly to $f(x)$, which must then be continuous.

This completes the proof of existence. To establish uniqueness, we proceed similarly. Let $F(x)$ be another solution of the equation which is continuous at $x = 0$. Then as in (4.5) - (4.6) we obtain the inequality

$$(4.12) \qquad |f(x) - F(x)| \leq \text{Max} \left[f(ay + b(x-y)) - F(ay + b(x-y)), \; f(az + b(x-z)) - F(az + b(x-z)) \right].$$

where $y = y(x)$ is a point at which the maximum of $g(y) + h(x-y) + f(ay + b(x-y))$ is attained and z is a corresponding point associated with F. Defining $v(x) = \sup_{0 \leq y \leq x} |f(y) - F(y)|$ we obtain from (4.12) the series of inequalities

$$(4.13) \qquad v(x) \leq v(cx) \leq v(c^2x) \leq v(c^nx) \quad (n=1,2,\cdots).$$

Since $v(x)$ is continuous at 0 and has the value 0 there, we see that as $n \longrightarrow \infty$ (4.13) leads to the conclusion that $v(x)$ is identically zero throughout $[0,x_0]$.

This completes the proof of existence and uniqueness.

5. Analytic Results g and h both Convex.

Since the basic functional equation is nonlinear, it is too much to expect that the solution may in general be obtained in explicit form. Instead we must focus our attention upon deriving simple general qualitative properties of the solution from simple assumptions concerning g and h. The term solution is used to include both the function satisfying the functional equation and the policy yielding the optimal return. The fact that this duality exists is of tremendous importance in connection with both theoretical and computational investigations.

In this section we obtain a simple consequence of convexity.

Theorem 2. If

$$(5.1)\qquad \text{(a)}\ g(0) = h(0) = 0,$$

$$\text{(b)}\ g'(x),\ h'(x) \geq 0,\quad g''(x),\ h''(x) \geq 0 \ \text{for}\ 0 \leq x \leq x_0, \ \text{and}$$

$$\text{(c)}\ \sum_{n=0}^{\infty} \left[g(c^n x) + h(c^n x) \right] < \infty$$

where $c = \mathrm{Max}\,(a,b)$, the optimal policy consists in choosing $y = 0$ or x for x in $[0,x_0]$.

Proof: The proof is readily obtained by showing that each of the functions f_n is convex. It is clear that $g(y) + h(x-y)$ is convex as a function of y for $0 \leq y \leq x$. Hence

$$(5.2)\qquad f_1(x) = \underset{0 \leq y \leq x}{\mathrm{Max}} \left[g(y) + h(x-y) \right] = \mathrm{Max} \left[g(x),\ h(x) \right].$$

From this it follows that $f_1(x)$ is convex, and hence that $g(y) + h(x-y) + f_1(ay + b(x-y))$ as a function of y is convex in $[0,x]$. Thus

$$(5.3)\qquad f_2(x) = \mathrm{Max} \left[g(x) + f_1(ax),\ \ h(x) + f_1(bx) \right],$$

which is again a convex function. It is clear now that the general result follows inductively.

The resulting functional equation

$$(5.4)\qquad f(x) = \mathrm{Max} \left[g(x) + f(ax),\ h(x) + f(bx) \right]$$

may be solved in certain special cases, which we shall not enter into here.

6. Analytic Results.

Let us now return to the equation

$$(6.1)\qquad f(x) = \max_{0 \le y \le x} \left[g(y) + h(x-y) + f(ay + b(x-y)) \right], \qquad f(0) = 0$$

and assume that g and h are both concave increasing functions of x. The problem is now much more complex, and in general, the optimal y will not be at an end point.

We shall prove

Theorem 3. *Let*

$$(6.2)\qquad \begin{aligned} &(a)\quad g(0) = h(0) = 0, \\ &(b)\quad g'(x),\ h'(x) \ge 0 \quad \text{for } x \ge 0, \\ &(c)\quad g''(x),\ h''(x) \ge 0 \quad \text{for } x \ge 0, \end{aligned}$$

and consider the sequence of approximations to f defined by

$$(6.3)\qquad \begin{aligned} f_0(x) &= \max_{0 \le y \le x} \left[g(y) + h(x-y) \right] \\ f_{n+1}(x) &= \max_{0 \le y \le x} \left[g(y) + h(x-y) + f_n(ay + b(x-y)) \right], \quad n=1,2,\cdots. \end{aligned}$$

For each n, there is a unique $y_n = y_n(x)$ that yields the maximum. If $b \le a$, we have $y_1 \le y_2 \le y_3 \cdots$, and the reverse inequalities for $b \ge a$. In particular, if $y_n(x) = x$ for some n in the case $b \le a$, then $y_m(x) = x$ for $m \ge n$, and the solution of the original equation in (6.1) will be furnished by $y = x$

This result is important in connection with determining approximate solutions, since it is quite simple to determine y_1, y_2, and even y_3 numerically.

We shall begin by assuming that all the maxima occur within the interval $[0,x]$ and shall then consider the case in which one $y_n(x) = x$. Considering the function $f_1(x)$, we see that its maximum, y, is determined by the equation

$$(6.4) \qquad g'(y) = h'(x-y).$$

Since the left-hand side is monotone increasing and the right-hand side is monotone decreasing, there is at most one solution. If we assume $h'(x) > g'(0)$, $g'(x) > h'(0)$, there will be exactly one solution of (6.4), which we call $y_1 = y_1(x)$. Differentiating (6.4), we obtain

$$(6.5) \qquad y_1' g''(y_1) = (1 - y_1')h''(x - y_1),$$

which yields

$$(6.6) \qquad y_1' = \frac{h''(x - y_1)}{g''(y_1) + h''(x - y_1)} \quad 0,$$

and

$$(6.7) \qquad 1 - y_1' > 0.$$

Turning to the expression for f, we have

(6.8) $f_1(x) = g(y_1) + h(x - y_1)$,

whence

(6.9) $f_1'(x) = g'(y_1) + (1 - y_1')h'(x - y_1) = h'(x - y_1)$,

using (6.4). Thus $f_1'(x) > 0$ and $f_1''(x) = (1 - y_1')h''(x - y_1) < 0$, which means that $f_1(x)$ is concave.

Let us now turn to the function $f_2(x)$,

(6.10) $$f_2(x) = \max_{0 \le y \le x} \left[g(y) + h(x-y) + f_1(ay + b(x-y)) \right] .$$

Assuming that there is a maximum inside the interval, we obtain

(6.11) $g'(y) - h'(x-y) + (a-b)f_1'(ay + b(x-y)) = 0$,

which we write

(6.12) $g'(y) + (a-b)f_1'(ay + b(x-y)) = h'(x-y)$.

The left-hand side is again strictly decreasing and the right-hand side strictly increasing, so that there is at most one solution which we call $y_2 = y_2(x)$, if it exists. Note that if there is no solution of (6.12), then

(6.13) $f_2(x) = g(x) + f_1(ax)$.

Let us, however, assume that there is a solution. Then

$$(6.14) \qquad f_2(x) = g(y_2) + h(x - y_2) + f_1(ay_2 + b(x - y_2)),$$

whence, as above, using (6.12),

$$(6.15) \qquad f_2'(x) = h'(x - y_2) + bf_1'(ay_2 + b(x - y_2)).$$

Using (6.12) again, this may be written

$$(6.16) \qquad f_2' = \frac{ah'(x - y_2) - bg'(y_2)}{a - b} .$$

This procedure is perfectly general, and we obtain, under our assumption concerning the existence of an internal maximum,

$$(6.17) \qquad f_n' = \frac{ah'(x - y_n) - bg'(y_n)}{a - b} , \quad n=1,2,3,\cdots .$$

We now wish to show that if $b < a$, then $y_1 \leq y_2 \leq \cdots$, and, conversely, if $a < b$, that $y_1 \geq y_2 \geq \cdots$. The two cases are really one, since we may interchange the roles of y and $x - y$ if we so wish. Since $f_1' > 0$, we see, on comparing (6.12) and (6.4), that $y_1 < y_2$.

The equation for y_3 is

$$(6.18) \qquad g'(y) + (a-b)f_2'(ay + b(x-y)) = h'(x-y).$$

If we can show that $f_2'(x) > f_1'(x)$, the same argument as that for y_1, y_2 shows that $y_3 > y_2$. Comparing (6.8) and (6.15), we see that $f_2' > f_1'$, since $h'(x-y_2) > h'(x-y_1)$.

To obtain the result for general n, always assuming that the maxima occur at inner points, we use (6.17). We know that $f_n'(x) > f_{n-1}'(x)$ implies that $y_{n+1} > y_n$. Since the function

$$(6.19) \qquad r(y) = \frac{ah'(x-y) - bg'(y)}{a - b}$$

is monotone increasing in y and $y_n > y_{n-1}$, via an inductive hypothesis, it follows that $f'_n > f'_{n-1}$, and thus that $y_{n+1} > y_n$.

Let us now consider the situation in which some $y_n(x) = x$. If $n = 1$, it is easy to see that $y_n(x) = x$, $n \geq 1$, since $y_1(x) = x$ means that $g'(y) \geq h'(x-y)$ for $0 \leq y \leq x$. Since

$$(6.20) \qquad \frac{\partial}{\partial y}\left[g(y) + h(x-y) + f_1(ay + b(x-y))\right]$$

$$= g'(y) - h'(x-y) + (a-b)f'_1(ay + b(x-y))$$

and $a \geq b$, we see that this expression is positive if $g'(y) \geq h'(x-y)$ for $0 \leq y \leq x$. Hence, $y_2(x) = x$, and, similarly, $y_n(x) = x$.

Let us now take the case in which $y_2(x) = x$, $y_1(x) \neq x$. Since $y_2(x) = x$ implies that $g'(y) - h'(x-y) + (a-b)f'_1(ay + b(x-y)) \geq 0$ for all $0 \leq y \leq x$, we have, in particular, $g'(x) - h'(0) + (a-b)f'_1(ax) \geq 0$. Since

$$(6.21) \qquad \begin{aligned} f'_2(x) &= g'(x) + af'_1(ax) \\ &= g'(x) - h'(0) + (a-b)f'_1(ax) + h'(0) + bf'_1(ax) \\ &\geq h'(0) \end{aligned}$$

and $f'_1(x) = h'(x-y) \leq h'(0)$, we see that $f'_2(x) \geq f'_1(x)$. This, as above, implies that $y_3 \geq y_2 = x$, and the process continues.

Let us note, finally, that if $g'(y) \geq h'(x-y)$ for all y in $[0,x]$, then $g'(y) \geq h'(z-y)$ for y in $[0,z]$ for all $z \leq x$.

In closing, this discussion of the functional equation, let us observe that if an interior maximum exists, we must have

$$(6.22) \qquad g'(y) - h'(x-y) + (a-b)f'(ay+b(x-y)) = 0,$$

and

$$(6.23) \qquad f'(x) = h'(x-y) + bf'(ay + b(x-y)).$$

This system of functional equations for y and f(x) may be solved explicitly if g and h are quadratic, which is a fact of some use in obtaining approximate solutions. In general, however, the system does not seem to be of much use for this particular equation. However, for other closely related equations, equations similar to (6.22) and (6.23) play a very important role in determining the solution, as we shall show elsewhere.

7. Stochastic Case.

Thus far we have considered a situation in which the outcome of a particular division of resources is completely determinate. Let us now briefly sketch the modifications required to treat the case where there is a probability distribution of outcomes. Let us assume if a division into y and x-y is made, there is a probability p_1, which in some cases might very well be a function of y, that the return will be $g_1(y) + h_1(x-y)$ with $a_1y + b_1(x-y)$ dollars available for distribution in the next stage, and a probability p_2 that the return will be $g_2(y) + h_2(x-y)$ with $a_2y + b_2(x-y)$ dollars remaining.

Since we are dealing with stochastic variables, it is necessary to introduce the metric of probability theory and speak of expected values. Let us define

(7.1) f(x) = expected total return starting with x dollars and using an optimal policy for an infinite number of stages.

Using the same arguments as before, we see that (formally at least) f(x) satisfies the functional equation,

$$f(x) = \max_{0 \le y \le x} \Big\{ p_1 \left[g_1(y) + h_1(x-y) + f(a_1 y + b_1(x-y)) \right] + p_2 \left[g_2(y) + h_2(x-y) + f(a_2 y + b_2(x-y)) \right] \Big\}$$

Precisely the same techniques as applied above may now be used to show existence and uniqueness, and to derive results corresponding to the previous analytic results.

8. Conclusion.

Let us now attempt to abstract some of the essential features of the preceding problem, features which are common to a large class of problems amenable to the techniques of the theory of dynamic programming.

(8.1) (a) We have an economic system described at any time by a set of parameters, p, which we call the state variables.

(b) At certain times we are to choose one of a set of decisions, D, which will have the effect of transforming these state variables into a similar set. The outcome of a decision may or may not be completely known.

(c) The choice of decisions is governed by the desire to maximize some function of the final state variables, a criterion function.

What is desired is a rule which will yield the optimal decision at each stage, knowing the values of the state variables and the permissible decisions at that stage.

In the problem treated above, there were for the N-stage process three state variables, namely (a) the quantity of money available for division at each stage, (b) the number of remaining stages, and (c) the return obtained from the preceding stages. Our aim was to maximize the state variable of (c) at the end of the process.

We have purposely left the description a bit loose, since we feel that it is the spirit of the technique that is important. Once grasped, the proper formulation of any particular problem becomes a matter of ingenuity, a quality which cannot be altogether banished from mathematics.

In order to obtain a mathematical formulation of the general problem, let us define

(8.2) $f(P)$ = the function of the final state variables obtained using an optimal policy starting with the initial variables represented by P.

Let us also represent the transformation effected by a choice by $P' = T_k(P)$ where k represents the parameters describing the particular choice. To obtain a functional equation governing the process, we use the following obvious

Principle of Optimality: An optimal policy has the property that whatever the initial state and the initial decision may be, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

The mathematical translation of this statement is the equation

$$(8.3) \qquad f(P) = \max_k f(T_k(P)) \ .$$

This equation may now be used, as in the preceding sections, to determine various properties of optimal policies under suitable assumptions concerning

$T_k(P)$. Of particular importance from the mathematical and practical viewpoints are those criterion functions that possess invariant properties and that reduce the number of state variables required. Expected return is an important function of this type, perhaps the most important, since a knowledge of the return over the preceding stages is not needed in order to determine the optimal continuation. Infinite processes are important for a like reason, since the number of stages remaining stays constant.

A more extensive and intensive study of the various types of functional equations arising from (8.3) will appear subsequently, [6], [7]. The interested reader may also wish to consult the references listed below.

REFERENCES

1. Arrow, K. J., T. E. Harris, and J. Marschak. "Optimal Inventory Policy," Cowles Commission Discussion Paper No. 44, 1951.

2. Bellman, R. "On the Theory of Dynamic Programming," Proc. Nat. Acad. Sci., Vol. 38 (1952), pp. 716-719.

3. ———. "An Introduction to the Theory of Dynamic Programming," RAND Corporation Report No. 245, 1953.

4. ———. "On Some Computational Problems in the Theory of Dynamic Programming," RAND Corporation Paper No. 423, presented at Symposium on Numerical Techniques, Santa Monica, 1953.

5. ———. "Bottleneck Problems and Dynamic Programming," Proc. Nat. Acad. Sci., Vol. 39 (1953), p. 947.

6. ———. "Some Functional Equations in the Theory of Dynamic Programming," Proc. Nat. Acad. Sci. (to appear).

7. Bellman, R. and S. Lehman. "On the Continuous Gold-mining Equation," Proc. Nat. Acad. Sci. (to appear).

8. Bellman, R., I. Glicksberg, and O. Gross. "On Some Variational Problems Occurring in the Theory of Dynamic Programming," Proc. Nat. Acad. Sci., Vol. 39 (1953).

9. Bellman, R., I. Glicksberg, and O. Gross. "On Some Variational Problems Occurring in the Theory of Dynamic Programming," Rendiconti del Circolo Matematico di Palermo (forthcoming).

10. Bellman, R., T. E. Harris, and H. N. Shapiro. "Studies on Functional Equations Occurring in Decision Processes," RAND Paper No. 382, 1952.

11. Dvoretzky, A. J., J. Kiefer, and J. Wolfowitz. "The Inventory Problems, Parts I and II," Econometrica, Vol. 20, Nos. 2 and 3 (1952), pp. 187-222; 450-466.

12. Karlin, S., and H. N. Shapiro. "Decision Processes and Functional Equations," RAND Research Memorandum No. 933, 1952.