The Philippine Statistician

Vol. xxi, Nos. 1 and 2, 1972

DYNAMIC PROGRAMMING AND OPTIMAL CONTROL IN MANAGEMENT

PAUL DE MUNTER [1]

1. *Inventory Problem*

In order to meet a known monthly demand, a firm must stock up every month at known prices. The quantity to be stocked is limited by the warehouse capacity. Knowing in advance prices and demand, what quality should be bought at the beginning of each month in order to meet the monthly demand and to make a maximum profit?

Methods of solution depend on the duration of the process —finite or infinite—and on the nature of the final goal—fixed or random.

Consider the simple case where the process extends over 6 months and the final goal is a nil stock.

Suppose that prices and demand are known in advance for the 6 month period; they are given in the following table.

Month Number	1	2	3	4	5	6
Demand	7	6	2	4	8	6
Prices	10	20	10	16	20	10

Also suppose that the maximum capacity of the warehouse is 12 and that the initial stock is 1.

The problem is to find a sequence of 6 numbers, each representing what is to be bought at the beginning of each month, at prices given in the table, in order to meet the demand, also given in the table, and so that the total stock never exceeds 12, the maximum capacity of the warehouse.

For example, one can decide to buy exactly what is needed to meet the coming month demand. Since there is an initial

[1] UNESCO Visiting Professor, U.P. Statistical Center

stock of 1 and the demand is 7, the first purchase is 6. Afterwards, purchases are equal to demand and the policy is

6, 6, 2, 4, 8, 6

The total cost of this policy is

$$(6).(10) + (6).(20) + (2).(10) + (4).(16) + (8).(20) + (6).(10) = 484$$

Let us consider another policy, for example,

11, 1, 12, 2, 0, 6

for which the total cost is 342 only. We leave to the reader to check that this policy actually meets demand. It is obviously a better policy than the first one. Moreover, it is the best one! In other words, it is impossible to find another sequence of 6 numbers which would satisfy the conditions and for which the total cost would be strictly smaller.

The problem is how to find this best solution. Instead of giving here details about the method, we shall emphasize the basic ideas of dynamic programming on which numerical methods of solution are grounded.

As for the first policy considered above, the best policy is also obtained step by step. Whereas, in the first case, we consider at a time only one month, in the dynamic programming approach we first consider 2 months and find the optimal purchase for the first month over a 2-month period. Then starting from this optimal result we add one more month and find the next optimal purchase over the 3-month period. This procedure is continued until the end of the 6-month period under consideration.

The method is based on the "principle of optimality", i.e., a theorem in the Calculus of Variations which, roughly, says that an optimal path is made up of disjoint optimal subpaths.

Let $v(n)$ be the value of step number n and let $f(n)$ be the optimal total value of the process up to step n. Then,

$$\begin{aligned} f(n) &= \text{opt } (v(1) + \ldots + v(n)) \\ &= \text{opt } (v(n) + \text{opt } (v(1) + \ldots + v(n\text{-}1))) \\ &= \text{opt } (v(n) + f(n\text{-}1)) \end{aligned}$$

Optimalization must be peformed according to the constraints. In our example, the stock at the beginning of each month must be at most 12 and at least equal to the demand.

From the statement of the pinciple of optimality, one can guess that the solution can be obtained starting either from the first two months or from the last two months.

In the case considered here, computations are simple. If the maximum warehouse capacity 12 is replaced by 25 and if sixth month demand 6 is replaced by 14, then computations become rather involved.

2. *Allocation Problem*

Five salesmen are to be distributed in five areas. It is well known what profit make 1, 2, 3, 4 or 5 salesmen when assigned to each area. They are given in the following table.

Areas / Nb of Sm	1	2	3	4	5
0	0	0	0	0	0
1	4	4	2	3	3
2	7	6	6	4	5
3	8	9	5	9	4
4	10	8	7	5	7
5	14	9	10	9	9

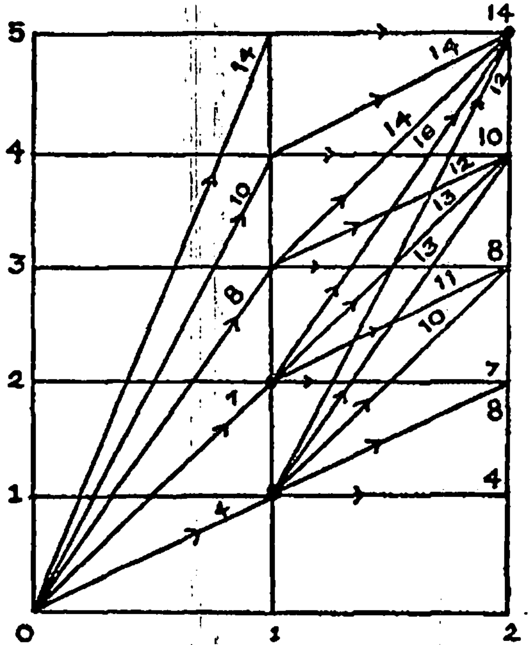
For example, if 5 salesmen were assigned to area 1, the total profit would be 14. If 4 salesmen are assigned to area 1 and 1 salesman to area 2, the total profit is also 14. If 4 salesmen are in area 3 and 1 in area, 1, the profit is 11.

The best distribution of salesmen can be found by listing all 126 possible distributions and by picking up the one with maximum profit.

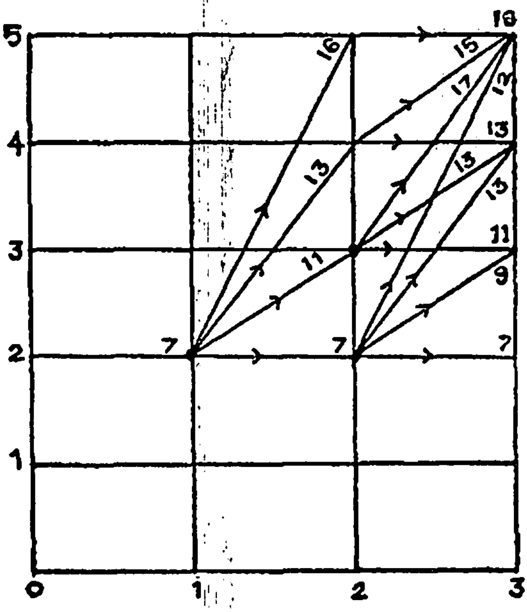
A dynamic programming approach needs only to consider 19 non-trivial distributions. A graphical solution which illustrates the application of the principle of optimality is given in this paper.

As the first step, two areas are taken into account and the *cumulative* numbers of salesmen are plotted on the graph. The value attached to each path is the profit from the cumulative number of salesmen up to the area considered.

Since maximum profit 16 is attained for the path passing through the point "two salesmen in area 1" this profit is selected for the optimal policy.

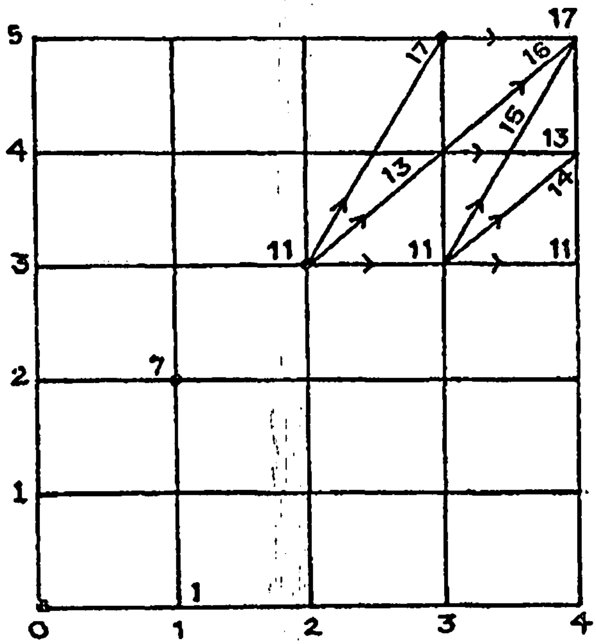


The second step gives all the paths starting from the first optimal point.



Since maximum profit 17 is attained for the path passing through the point "3 salesmen up to area 2" this point is selected. So far, area 1 is given 2 salesmen and area 2 is given 1.

The third step describes all the possible paths starting from the last optimal point.



Since maximum profit 17 is attained for the path passing through the point "5 salesmen up to area 3" this point is selected. Area 3 is given 2 salesmen. No more salesmen being available, areas 4 and 5 are given no salesmen. The final and best solution is to allocate

2 salesmen to area 1

1 salesman to area 2

2 salesmen to area 3

0 salesman to area 4

0 salesman to area 5

the total profit of this distribution being $7 + 4 + 6 = 17$.

3. *An Equipment Replacement Problem Including Probabilistic Consideration.*

Suppose that the state of a system is the age of a piece of equipment (counted in years). Call Bib that piece of equipment. Every year, a decision is to be made whether to keep Bib or to replace it. The cost of each decision is known in advance.

To make our model more realistic, the probability for Bib i years old to become (i + 1) years old can be introduced. These probabilities can be obtained from previous experience with a piece of equipment similar to Bib.

What is sought in the problem is a sequence of decisions (k = keep, r = replace)

$$k, k, \ldots, k, r, k, \ldots$$

which extends over a fixed or undetermined period of time and makes the total profit maximum. This will be expressed in terms of "return over n years, when Bib is i years old". Return over one year is equal to productivity minus maintenance or to trade-in value minus purchase price, depending on what decision is made.

To apply the principle of optimality to this problem, consider the optimal return after n years of a i years old Bib. Denote this number by f(n,i). After one year, Bib is either (i = 1) years old or 1 year old (if we replaced old Bib by a new one). The values of this one year step are v(k), v(r), respectively. Let the respective probalities be denoted by p(k), p(r).

According to the principle of optimality, the optimal return over n years is equal to the actual value of a one year step plus the optimal return over the (n — 1) remaining years. Therefore, the optimal return over n years is

$$p(k)\ (v(k) + f(n-1, i+1)$$

or

$$p(r)\ (v(r) + f(n-1, 1)\ )$$

and decision k or r is made according to which number is greater.

Computations are rather involved and have to be carried out step by step starting with n = 1, i = 0, 1, 2, ..., then n = 2, i = 0, 1, 2, ... etc.

4. *Optimum Control*

A chemical A is added to a tank at constant rate and during a fixed period of time T. The reaction in the tank occurs at a certain pH which determines the quality of the final product.

Denote the pH at time t by $x(t)$. It can be controlled by the strength $u(t)$ of a certain component of A. If the variation of x during a certain period of time is proportional to x itself and to u, then the reaction occurs at the desired pH.

Maintaining u over T at the convenient level costs $C(u)$, a well known number.

The problem is to find u so that $C(u)$ is minimum. According to the principle of optimality, u optimal over T must be optimal over each subinterval of T.

Let $C(u, t, h)$ be the cost of u over an arbitrary period of time h, starting a time t. Let $C^*(u, T\text{-}t\text{-}h)$ be the optimal cost of u over the remaining period of time. Therefore

$$C^*(u, T-t) = \min_u \left(C(u,t,h) + C^*(u,T-t-h)\right)$$

which is very similar to relations obtained in Sections 1 and 3.

Given C, it is now a purely mathematical problem to solve for u.

For example, if $C(u)$ is the integral, from 0 to T, of a quadratic function in x and u, then it can be found that u is merely proportional to x.

5. *Bibliography*

Lee (E.B.), Markus (L.). Foundations of optimal control theory. Wiley.
Bellman (R.E.), Dreyfus (S.E.). Applied Dynamic Programming. Princeton U.P.
Dreyfus (S.E.). Dynamic Programming and the Calculus of Variations. Academic Press.
Roberts (S.N.). Dynamic Programming in Chemical Engineering and Process Control. Academic Press.
Kaufmann (A.), Cruon (R.). Dynamic Programming. Academic Press.
Bellman (R.). Mathematical Theory of Control Processes. Academic Press.