REVUE FRANÇAISE D'AUTOMATIQUE, INFORMATIQUE, RECHERCHE OPÉRATIONNELLE. RECHERCHE OPÉRATIONNELLE

G. V. RAO

An optimal decision procedure using dynamic programming technique

Revue française d'automatique, informatique, recherche opérationnelle. Recherche opérationnelle, tome 6, nº V2 (1972), p. 79-86

<http://www.numdam.org/item?id=RO_1972__6_2_79_0>

R.A.I.R.O.
(6e année, octobre 1972, V-2, p. 79-86)

AN OPTIMAL DECISION PROCEDURE USING DYNAMIC PROGRAMMING TECHNIQUE *

par G. V. Rao (1)

Summary. — *In this paper an attempt is made to study an optimization problem and it is solved using the dynamic programming approach. A stochastic model is developed for the minimization of the over-all loss and explicit solutions are derived for gamma, exponential, normal and beta distributions. A numerical example is solved to illustrate the procedure.*

1. INTRODUCTION

In his novel, Arnold Bennett (1957) describes the following interesting situation. Many tradesmen in the Five Towns had formed clubs so that their customers pay so much during every month to the tradesmen, who charged them (customers) nothing for keeping it, and at the end of the agreed number of months they purchased goods worth of the total amount deposited by them. Denry, a philanthropist of the Five Towns, wanted to start a new club and as a special inducement and to prove superior advantages to the ordinary clubs, wanted to allow his customers to spend their (customers') full nominal subscription to the club as soon as they had actually paid only half of it. Thus, for example, after paying fifty dollars (ten dollars a month), the customers could spend one hundred dollars in Denry's chosen shops and Denry would settle with the shops immediately, while collecting the balance every month from his customers. These benefits to the customers were without any charge whatsoever to them. Now the factor that influences the decision to be made is the loss incurred to him due to the possible discontinuation of payment by the customers during the later part of the planning horizon (i.e., after the purchase of articles from Denry's chosen shops). Also there is a certain amount of remuneration received by Denry from the shop-keepers to whom Denry would give new customers. They (shop-keepers) were to allow him certain

* Supported by a U.G.C. Junior Research Fellowship.
(1) University of Bombay, Bombay.

percentage of discount on all the transactions affected through Denry. The problem faced to Denry was how long he should make the collection in order to minimize his total loss, taking into consideration the loss incurred due to the possible discontinuation of payment by the customers during the second half of the planning horizon, some incidental expenses (may be fixed) and the commission allowed by the shop-keepers.

Of course, Denry was lucky enough to loose nothing without applying an optimization technique, but in a situation of this type, the dynamic programming technique would give an useful solution.

2. STATEMENT OF THE PROBLEM

We shall re-state the above problem with some modifications as follows :

Consider a situation where customers deposit a certain amount of money every month with a trading concern continuously over a duration of time, say N months. A the end of $N/2$ (N is assumed to be even) months, the trading concern will issue credit cards to the customers worth of double the amount of what they have actually deposited till then, under the assumption that they (customers) will continue to pay their subscription upto the end of N months, so that the customers can purchase goods from an authorized dealer equivalent to that amount. There is a certain amount of loss due to the possible discontinuation of payment by the customers soon after they are being issued with credit cards for the purchase of articles (i.e., after the purchase of articles from the dealer). However, the dealer would pay a certain percentage of commission to the trading concern as remuneration towards the new business offered to the dealer by the trading concern. There may be some additional incidental expenses also.

Because of these physical characteristics of the process, it may not be very economical (i.e., the loss incurred by the trading concern may be very high as compared to the commission allowed by the dealer) to continue the process for a length tof N months and a decision has to be made as to when the process should be revised. If the revision costs were zero, then one would like to revise the process more frequently (i.e., there may be many cycles of shorter duration). However, this is not so, and will result in cycles of longer duration.

3. FORMULATION OF THE PROBLEM

Suppose, at the beginning of the process, we know the state of the system, i.e., the values of the parameters α and β (say) which will determine the over-all loss if we continue the process over the duration of n, say, months where α

may be the factor contributed by the loss incurred due to the possible discontinuation of payment during the later part of the planning horizon and β may be the factor contributed by the gain due to the percentage of commission allowed by the dealer. Then the total loss incurred during this cycle consisting of n months be denoted by $1\ (\alpha, \beta, n)$.

Our problem is to minimize the expected total loss when the process is continued over a duration of N months. Once we know the parameters α and β and the revision policy we intend to adopt (which tells us when to revise the process when α and β are known), the expected over-all loss over the next N months will be a function of α, β and N only. Let us denote this policy, when the optimal decision procedure is used, by the function $f_N(\alpha, \beta)$, the immediate loss (when the process is carried over n months) being $1(\alpha, \beta, n)$.

The loss over the remaining duration of (N-n) months (i.e., at the end of the first cycle) will depend upon the new values of the parameters α, β, when the process is revised, and the duration, N-n, of the process. Thus, using the principle of optimality (Bellman, 1957), the optimal loss over the remaining duration of (N-n) months would be

$$f_{N-n}(\gamma, \omega) \tag{3.1}$$

where γ and ω are the resultant parameters at the end of n months (i.e., the parameters at the start of the next cycle) corresponding to the original parameters α and β respectively.

Since we do not know the exact values of the new parameters γ and ω, we shall have to estimate the values of these parameters on the basis of the past information. That is, we assume the joint probability distribution of parameters γ and ω is of the form $dG(\gamma, \omega)$.

Therefore, the expected loss over the remaining duration of (N-n) months will be (using (3.1)).

$$\int_\gamma \int_\omega f_{N-n}(\gamma, \omega)\, dG(\gamma, \omega) \tag{3.2}$$

Hence, we can derive the recurrence relation for the over-all N month loss as

$$f_N(\alpha, \beta) = \underset{0 \leqslant n \leqslant N}{\text{Min}} \left[1(\alpha, \beta, n) + \int \int_\omega f_{N-n}(\gamma, \omega)\, dG(\gamma, \omega) \right] \tag{3.3}$$

4. SOLUTION OF THE PROBLEM

Generally, we are interested in finding the optimal policy which minimizes the over-all loss when the process continues over a long duration of time, i.e., when N tends to infinity. Under certain conditions (Howard 1961, White 1969), $f_N(\alpha, \beta)$ takes the form (1)

$$f_N(\alpha, \beta) = Ng + f(\alpha, \beta) \tag{4.1}$$

as $N \to \infty$ and where g is the average loss per unit time.

Using (4.1) and for large N, we can write the equation (3.3) as

$$f(\alpha, \beta) = \underset{n \geqslant 0}{\text{Min}} \left[1(\alpha, \beta, n) - ng + \int\int f(\gamma, \omega)\, dG(\gamma, \omega) \right] \tag{4.2}$$

Substituting for

$$\int_\gamma \int_\omega f(\gamma, \omega)\, dG(\gamma, \omega) = \mu \tag{4.3}$$

we can write (4.2) as

$$f(\alpha, \beta) = \underset{n \geqslant 0}{\text{Min}} \left[1(\alpha, \beta, n) - ng + \mu \right] \tag{4.4}$$

If the loss over the initial cycle of duration of n months is of the form

$$1(\alpha, \beta, n) = \alpha n^2 - \beta n + m \tag{4.5}$$

where m is a constant (this factor may be the fixed incidental cost), then, substituting (4.5) in (4.4), we have

$$f(\alpha, \beta) = \underset{n \geqslant 0}{\text{Min}} \left[\alpha n^2 - \beta n + m - ng + \mu \right] \tag{4.6}$$

$$\frac{\partial f(\alpha, \beta)}{\partial n} = 0 \text{ gives}$$

$$n = \frac{\beta + g}{2\alpha} \tag{4.7}$$

and the optimal value of the over-all loss is

$$f(\alpha, \beta) = \mu + m - \frac{(\beta + g)^2}{4\alpha} \tag{4.8}$$

(1) $f_N(\alpha, \beta) = Ng + f(\alpha, \beta) + \delta_N(\alpha, \beta)$ where $\delta_N(\alpha, \beta) = 0$ as $N \to \infty$ and $/\delta_N(\alpha, \beta)/$ need only be uniformly bounded for all α, β and N.

Using this result in the definition of μ (i.e., in (4.3)), we have

$$4m = \int_{\gamma=0}^{\infty} \int_{\omega=0}^{\infty} \frac{(\omega + g)^2}{\gamma} dG(\gamma, \omega) \tag{4.9}$$

Once we have solved (4.9) for the value of g, we can obtain the value of n from (4.7).

5. SOLUTION OF g WHEN α AND β ARE INDEPENDENTLY DISTRIBUTED

Let the two parameters α and β be independently distributed so that the joint probability distribution G (γ, ω) (the resultant parameters γ and ω are also independently distributed) can be expressed as

$$dG(\gamma, \omega) = d\o(\gamma)\, d\theta(\omega) \tag{5.1}$$

Hence the equation (4.9) reduces to

$$4m = \left[\int_{\gamma=0}^{\infty} \frac{1}{\gamma} d\O(\gamma)\right] \left[\int_{\omega=0}^{\infty} (\omega + g)^2 \, d\theta(\omega)\right] \tag{5.2}$$

Substituting $k = \int_{\gamma=0}^{\infty} \frac{1}{\gamma} d\O(\gamma)$ in (5.2), we get

$$\int_{\omega=0}^{\infty} (\omega + g)^2 \, d\theta(\omega) = 4mk^{-1} \tag{5.3}$$

6. PARTICULAR CASES

6.1. α-gamma and β-gamma.

Let

$$d\O(\gamma) = \frac{b(b\gamma)^{a-1} \exp\{-b\gamma\}\, d\gamma}{\Gamma(a)} \text{ for } \gamma > 0 \tag{6.1.1}$$

and

$$d\O(\omega) = \frac{c(c\omega)^{d-1} \exp\{-c\omega\}\, d\omega}{\Gamma(\mathrm{d})} \text{ for } \omega > 0 \tag{6.1.2}$$

Then, we can obtain

$$k = \frac{b}{a-1} \tag{6.1.3}$$

and

$$\int_0^{\infty} (\omega + g)^2 \, d\theta(\omega) = \frac{d(d+1)}{c} + 2g\frac{d}{c} + g^2 \tag{6.1.4}$$

Using (6.1.3) and (6.1.4) in (5.3), we get

$$g^2 + 2g\frac{d}{c} + \frac{d(d+1)}{c^2} = \frac{4m(a-1)}{b} \tag{6.1.5}$$

(6.1.5) can be solved to obtain the value of g.

6.2. α-gamma and β-exponential.

Let

$$d\theta(\omega) = \lambda \exp\{-\lambda\omega\}\, d\omega, \text{ for } \omega > 0, \tag{6.2.1}$$

Then

$$\int_0^\infty (\omega + g)^2\, d\theta(\omega) = g^2 + 2g\lambda^{-1} + 2\lambda^{-2} \tag{6.2.2}$$

From (6.2.2), (6.1.3) and (5.3), we have

$$g^2 + 2g\lambda^{-1} + 2\lambda^{-2} = \frac{4m(a-1)}{b} \tag{6.2.3}$$

Thus, we can obtain the value of g from (6.2.3).

6.3. α-gamma and β-normal.

Let

$$d\theta(\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\omega^2/2\sigma^2\}\, d\omega, -\infty < \omega < \infty \tag{6.3.1}$$

With θ(0) so small that negative values of β are not significant.

Then

$$\int_0^\infty (\omega + g)^2\, d\theta(\omega) = g^2 + 2g\sigma\sqrt{\frac{2}{\pi}} + \sigma^2 \tag{6.3.2}$$

so that using (6.1.3) and (6.3.2) in (5.3), we get

$$g^2 + 2g\sigma\sqrt{\frac{2}{\pi}} + \sigma^2 = \frac{4m(a-1)}{b} \qquad (6.3.3))$$

Hence we can get the solution of g from (6.3.3).

6.4. α-beta and β-beta.

Let

$$d\varnothing(\gamma) = \frac{1}{B(a, b)}\gamma^{a-1}(1-\gamma)^{b-1}\, d\gamma, \text{ for } 0 < \gamma < 1 \tag{6.4.1}$$

and

$$d\theta(\omega) = \frac{1}{B(c, d)}\omega^{c-1}(1-\omega)^{d-1}\, d\omega, \text{ for } 0 < \omega < 1 \tag{6.4.2}$$

Then

$$k = \frac{a + b - 1}{a - 1} \tag{6.4.3}$$

and

$$\int_0^1 (\omega + g)^2 \, d\theta(\omega) = g^2 + 2g \frac{c}{c + d} + \frac{c(c + 1)}{(c + d)(c + d + 1)} \tag{6.4.4}$$

so that substituting (6.4.3) and (6.4.4) in (5.3), we get

$$g^2 + 2g \frac{c}{c + d} + \frac{c(c + 1)}{(c + d)(c + d + 1)} = \frac{4m(a - 1)}{a + b - 1}$$

which can be solved to obtain the value of g.

7. NUMERICAL SOLUTION

In the previous section, under 6.2, we have assumed that α follows a gamma type distribution with parameters a and b and β follows an exponential distribution with parameter λ. In particular, let $a = 3$, $b = 1$ and $\lambda = 1$. Let $m = 10$, so that from (6.2.3) we obtain

$$g^2 + 2g + 2 = 80.$$

Solving this for g and taking the positive value (since we are interested in the average loss per unit time), we have $g = 7.89$. Then, for $\alpha = 0.5$ and $\beta = 2.5$, the optimal value of n (from (4.7)) is approximately equal to 10, so that the optimal decision policy is to collect the amount over a duration of 10 months initially, and then to carry the process depending upon the new values of the parameters.

Here, one may suppose that α depends upon the probability of non-payment of amount by the customers during the second-half of the cycle and β depends upon the amount of monthly deposit by the customers and the rate of commission allowed by the dealer to the trading concern. The constant m may be some fixed incidental cost and revision cost, if any.

With some modifications, the problem can also be viewed as a maximization problem (White, 1962) and a solution to the problem can be obtained. Such of these problems may arise when a manufacturer wants to introduce a new product to the market. To capture the market, initially, the manufacturer would like to give some incentives to the customers whereby he may incur some loss but once the market is captured, he will be able to make-up the loss and then run profitably. In such a situation, the use of optimization technique will help the decision maker to run the process most efficiently.

ACKNOWLEDGEMENTS

The author is extremely grateful to Professor M. C. Chakrabarti for suggesting this problem and for his guidance during the preparation of the paper. Thanks are also due to the referee for his useful suggestions.

REFERENCES

ARNOLD Bennett, *The Card*, World Distributions Ltd., London (1957).

BELLMAN R., *Dynamic Programming*, Princeton University Press, Princeton (1957).

HOWARD R. A., *Dynamic Programming and Markov Processes*, John Wiley & Sons, New York (1961).

WHITE D. J. « Optimal revision periods », *J. Math. Ans. and Appl.*, (1962), vol. 4, pp 353-365.

WHITE D. J., *Dynamic Programming*, Oliver and Boyd, London (1969).