Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Technical Note—An Adjustment to the Norman-White Approach to Approximating Dynamic Programs

Evan L. Porteus,

An Adjustment to the Norman-White Approach to Approximating Dynamic Programs

EVAN L. PORTEUS
Stanford University, Stanford, California
(Received August 1977; accepted August 1978)

This note considers an adjustment to the Norman-White approach of approximating stochastic dynamic problems by using expectations The expectations are used only in approximating the transition mechanism The original expected returns are left intact. The adjustment also involves eliminating the policy improvement step proposed by Norman and White The approach appears to work surprising well on the examples considered. It may turn out to be a competitive practical approach to approximating solutions to the (s, S) inventory model

---

IN [4], Norman and White proposed that solutions to stochastic dynamic programs be approximated by using exceptions. They suggested that a deterministic problem be solved, followed by one application of Howard's [2] policy improvement step to get the approximating policy. They showed that their approach worked quite well in two examples. Morton [3] then provided an example where the error due to their approach could be made arbitrarily large. He suggested that the deterministic approach might indeed yield a good approximation of the optimal return function, and that it be used to initiate a policy iteration or a modified value iteration. This note suggests that, in their examples, Norman and White used only one of the two deterministic approaches implied by their presentation and that the approach they did not take may be preferable. Indeed, this note contends that the solution yielded by this latter deterministic approach will often be good enough so that no policy improvement steps are needed. The ensuing computational advantages are transparent.

After showing how the proposed approach follows from the Norman-White exposition, we apply it to Morton's example and find that it yields the optimal solution, rather than an approximate solution whose error can be made arbitrarily large. We then illustrate the approach using other examples. In particular, the approach appears to yield excellent results for the (s, S) inventory model. We conclude by observing that the approach can be interpreted as a state variable reduction technique in multidimensional state variable problems.

Operations Research
Vol 27, No 6, November-December 1979
0030-364X/79/2706-1203 $01 25

1. DERIVATION OF THE APPROACH

We illustrate the derivation using the discounted finite Markov decision chain case. Suppose the process is currently in state i and decision k is made. Let p_{ij}^k denote the probability that the process will move to state j and r_{ij}^k the present value of the resulting return. Let $r_i^k = \sum_j p_{ij}^k r_{ij}^k$. Assuming that the concept is meaningful for the process at hand, let $\mu\,(i, k)$ denote the expected next state. In their exposition, Norman and White propose that the following deterministic problem be solved: find the optimal return function and associated policy such that

$$f_i = \max_k(r_i^k + \alpha f_{\mu(i,k)})$$

where $\alpha(0 \leq \alpha < 1)$ is the discount factor. However, in their examples, they find the optimal return and associated policy such that $f_i = \max_k(r_{i\mu(i,k)}^k + \alpha f_{\mu(i,k)})$. In the first approach, the returns from the stochastic model are used in the deterministic approximation, whereas in the second, the returns from the deterministic approximation are used as well. This note merely proposes that the first approach be used, and that no policy improvement steps necessarily be applied. Norman and White show that this approach is a first order one: The same approach is achieved by linearizing the optimal return function in the stochastic model. Thus, the approach ought to yield good results whenever the optimal return function is sufficiently linear over the relevant ranges. It is possible to develop one sided bounds on the optimal return function when it is convex or concave. However, details are not developed here. In the next section we illustrate the use of the proposed approach.

2. EXAMPLES

Morton [3] considered an undiscounted inventory model with no backlogging, a holding cost of 1 per unit on positive ending inventory, a shortage cost of p per unit, no (relevant) ordering cost, and two possible demands: m units with probability $1/m$ and zero units otherwise. The approach here seeks the optimal relative value function f, average return g, and associated policy such that

$$f(x) + g = \min_{y \geq x}\left[\frac{(m-1)y}{m} + \frac{p(m-y)^+}{m} + f(y-1)\right]$$

where $x^+ = \max(0, x)$ and, without loss, we have assumed that $y \leq m$. It is straightforward that when $p \leq m - 1$ (Morton's condition), the use of $f(x) = f(z)$ for $0 \leq x, z \leq m$ and $g = m - 1$ solves the optimality equations and the optimal policy is as given by Morton: base stock with level equal to m.

Norman and White's Examples

The first example considered by Norman and White [4] was a discounted inventory model without backlogging due to Bellman [1]. Modifying their approach to retain the original expected one period costs yields

$$f(x) = \min_{y \geq x}[c \cdot (y - x) + L(y) + \alpha f(y - \mu)^+]$$

where c is the unit ordering cost, and L is the expected penalty cost function, μ is the expected demand. The optimal policy will be base stock. Let S denote its level. Thus we get

$$\begin{aligned} v(S) &= L(S) + \alpha v(S - \mu) \\ &= L(S) + \alpha[c\mu + v(S)] \end{aligned}$$

so $v(S) = (L(S) + \alpha c\mu)/(1 - \alpha)$ and $v(0) = (cS(1 - \alpha) + L(S) + \alpha c\mu)/(1 - \alpha)$. Minimizing this quantity over S yields the condition $\Phi(S) = 1 - c(1 - \alpha)/(\alpha p)$, where Φ is the probability distribution of demand, and p is the unit shortage cost. When compared to the correct condition, $\Phi(S) = 1 - c(1 - \alpha)/(\alpha(p - c))$, the approximate solution will be quite good when $p \gg c$. For instance, using the numerical example in [4], where $\alpha = 0.95$, $c = 1$, $p = 10$, and demand is normally distributed with mean 100 and standard deviation 10, the optimal level is 125.2 whereas the approximate level is 125.5.

Norman and White's other example, Howard's [2] automobile replacement problem, was nearly completely solved by essentially the approach suggested here (without application of a policy improvement step) as noted by Morton [3].

The (*s*, *S*) Inventory Model

Consider the discounted inventory model with backlogging where it is known that an (s, S) policy is optimal (Scarf [6], Veinott [8], Porteus [5], and Schäl [7]). The approach here yields $f(x) = \min_{y \geq x}[c(y - x) + L(y) + \alpha f(y - \mu)]$, where c incorporates a linear portion plus a setup cost and L is the expected holding and shortage cost function. It is clear that specification of s is relevant only to the extent of the number of periods until the expected stock level drops below s. Letting n denote this number, it is straightforward to derive that we wish to minimize $v(n, S) = K + cS + [L(S) + \alpha L(S - \mu) + \cdots + \alpha^{n-1}L(S - (n - 1)\mu) + \alpha^n(K + nc\mu)]/(1 - \alpha^n)$ over n and S. Having found the optimal n^* and S^*, we set $f(S) = v(n^*, S^*) - K - cS$. Using $f(s) = L(s) + \alpha f(s - \mu) = L(s) + \alpha[K + c(S - s + \mu) + f(S)]$, we estimate the optimal s by finding the smallest s such that $L(s) - (1 - \alpha)cs \leq (1 - \alpha)v(n^*, S^*) - \alpha c\mu$. Computationally, the approach is straightforward: when $n = 1$, it reduces to the usual newsboy problem. When $n > 1$, the formulas are different but essentially require little additional work.

To illustrate the development and to prepare for some numerical examples, we consider the undiscounted model and specialize the results to the case when h is the unit holding cost charged on positive ending inventory, p is the unit shortage cost, and demand is normal with mean μ and standard deviation σ. Here $L(y) = h \int_0^y (y - \xi)\, d\Phi(\xi) + p \int_y^\infty (\xi - y)\, d\Phi(\xi) = (h + p)\sigma I_N(\tilde{y}) + h(y - \mu)$, where I_N is the readily available standardized normal loss integral (see [10, 11]) and $\tilde{y} = (y - \mu)/\sigma$. Furthermore $L'(y) = -p + (h + p)\Phi(y)$. Here, we seek to minimize $v(n, S) = [L(S) + L(S - \mu) + \cdots + L(S - (n - 1)\mu) + K]/n$. Since $v(n, \cdot)$ is convex, we can find the optimal S for given n by requiring that $\Phi(S) + \Phi(S - \mu) + \cdots + \Phi(S - (n - 1)\mu) = np/(h + p)$. Furthermore, in many cases, all terms on the left except for the last one will (nearly) equal one, in which case the formula reduces to finding S so that $\Phi(S - (n - 1)\mu) = (p - (n - 1)h)/(h + p)$. An obvious by-product is that $n \le 1 + p/h$. Let S_n denote the optimal value of S for given n. To determine the optimal n, we must minimize $v(n, S_n)$ over n. This can be done directly using the expressions provided. Since $L(y) \cong h(y - \mu)$ for large y, in many cases the following simplification can be used: $v(n, S_n) \cong [K + (h + p)\sigma I_N([S_n - n\mu]/\sigma)]/n + hS_n - h\mu(n + 1)/2$. One can find the optimal n by iterating on n, or, if that is too much work, one can examine values near $\sqrt{2K/(h\mu)}$ which corresponds to using the classical EOQ formula.

Given n and S, to estimate the optimal value of s, we seek the smallest s such that $L(s) = v(n, S)$. For an estimate, we can assume that $L(s)$ entails negligible holding cost and find s such that $I_N((s - \mu)/\sigma = v(n, S)/(p\sigma)$. This value can be used to initiate the search for s, as it would yield a one sided bond.

Numerical Example

To illustrate the use of the formulas just presented, we apply them to a numerical example from Veinott and Wagner [9]. In it we have $K = 64$, $h = 1$, $p = 9$, and $\mu = \sigma^2 = 52$. We approximate the Poisson distribution with a normal distribution and assume that inventory level is a continuous variable.

We compute $\sqrt{2K/(h\mu)} \cong 1.6$ to estimate n^*, so we shall test $n = 1$ and $n = 2$. We shall seek S such that

$$\Phi(S - (n - 1)\mu) = (p - (n - 1)h)/(h + p).$$

For $n = 1$, we require $\Phi(S) = 0.9$ so $S \cong \mu + 1.28\sigma \cong 61.2$. Thus, $v(1, S_1) = L(S_1) + K = K + (h + p)\sigma I_N((S_1 - \mu)/\sigma) + h(S_1 - \mu) \cong 76.6$. For $n = 2$, we require $\Phi(S - \mu) = 0.8$ so $S \cong 2\mu + 0.84\sigma \cong 110.1$. Thus, $v(2, S_2) = (L(S_2) + L(S_2 - \mu) + K)/2 \cong [K + (h + p)\sigma I_N((S_2 - 2\mu)/\sigma)]/2 + hS_2 -$

$3h\mu/2 \cong 68.1$ so that $n^* = 2$ and $S^* = 110.1$ (a check did indicate that $n = 3$ was less preferable). To estimate s^* we use

$$I_N((s - \mu)/\sigma) = v(2{,}110.1)/(p\sigma) \cong 1.05$$

so $s \cong 45.1$. Our estimate of (s^*, S^*) is therefore (45.1, 110.1) whereas the actual discrete optimal solution is given in [9] to be (45, 112). It appears that our estimate yields better results than the estimate of (45, 61) found in [9] for this example after substantially more computation. It also appears to improve on the estimate of (43.3, 124.9) found by applying the normal approximation approach in [11] (see also Section 19.6 in [10]). It is unclear whether the proposed approach will consistently yield such attractive results.

Other Applications

The current approach can be readily applied to nonstationary problems. In the inventory area, this would lead to the study of deterministic inventory models wherein the holding cost function actually represents the expected holding and shortage cost function. Such a function does not vanish at the origin and may well be convex.

The approach can also be considered to be a dimension reducing technique for multidimensional state variable problems, such as multi-echelon, multi-product inventory control problems, multi-animal type livestock management problems, and multi-tree type forest management problems. Even a single policy improvement step may be prohibitive in some such problems. Finally, it is plausible that the approach may find use in the role of an analytic simulator.

ACKNOWLEDGMENTS

This research was supported in part by a grant from the Atlantic Richfield Foundation to the Graduate School of Business. It was carried out while the author was visiting the Department of Economics, University of Canterbury, New Zealand. The author is grateful for the administrative support provided during his visit.

REFERENCES

1. R. BELLMAN, *Dynamic Programming,* Princeton University Press, Princeton, N.J., 1957.
2. R. HOWARD, *Dynamic Programming and Markov Processes,* Technology Press of M.I.T., 1960.
3. T. MORTON, "A Critique of the Norman-White Dynamic Programming Approximation," *Opns. Res.* **17,** 751–753 (1969).
4. J. NORMAN AND D. WHITE, "A Method for Approximate Solutions to Stochastic Dynamic Programming Problems Using Expectations," *Opns. Res.* **16,** 296–306 (1968).

5. E. PORTEUS, "On the Optimality of Generalized (s, S) Policies," *Mgmt. Sci.* **17,** 411-426 (1971).
6. H. SCARF, "The Optimality of (S, s) Policies in the Dynamic Inventory Problem," in *Mathematical Methods in the Social Sciences,* H. Scarf, D. Gilford and M. Shelly (eds.), Stanford University Press, Stanford, Calif., 1963.
7. M. SCHÄL, "On the Optimality of (s, S)-Policies in Dynamic Inventory Models with Finite Horizon," *S.I.A.M. J. Appl. Math.* **30,** 528-537 (1976).
8. A. VEINOTT, JR., "On the Optimality of (s, S) Inventory Policies, New Conditions and a New Proof," *S.I.A.M. J. Appl. Math.* **14,** 1067-1083 (1966).
9. A. VEINOTT, JR., AND H. WAGNER, "Computing Optimal (s, S) Inventory Policies," *Mgmt. Sci.* **11,** 525-552 (1965).
10. H. WAGNER, *Principles of Operations Research,* Prentice-Hall, Englewood Cliffs, N.J., 1969.
11. H. WAGNER, M. O'HAGAN AND B. LUNDH, "An Empirical Study of Exactly and Approximately Optimal Inventory Policies," *Mgmt. Sci.* **11,** 690-723 (1965).

Approximations for Overflows from Queues with a Finite Waiting Room

JOHN H. RATH

Bell Laboratories, Holmdel, New Jersey

DIANE SHENG

Stanford University, Stanford, California

(Received September 1977; accepted July 1978)

We report in this note a study of the use of the interrupted Poisson process (IPP) as an approximation to the overflow process from a $M/M/C/K$ queue (a queue with a Poisson arrival process, C servers each with exponentially distributed service times, and a maximum of K customers in the system) The IPP has been used successfully by Kuczura to approximate the overflow from a $M/M/C/C$ queue Numerical comparisons are made for a number of examples and both the maximum absolute errors and percent errors are plotted.

WE REPORT in this note a study on the use of the interrupted Poisson process (IPP) as an approximation to the overflow process from a $M/M/C/K$ queue (a queue with a Poisson arrival process, C servers each with exponentially distributed service times, and a maximum of K ($K \geqq C$) customers in the system; customers arriving when K customers are already present are blocked and cleared from the system).

Operations Research 0030-364X/79/2706-1208 $01.25
Vol 27, No 6, November-December 1979