INVENTORY MANAGEMENT WITH APPROXIMATE METHOD OF DYNAMIC PROGRAMMING

Saloni Srivastava*

Dr. R. K. Shrivastava**

ABSTRACT

In this paper, under multiple suppliers and stochastic lead time, we introduce an approximation method for inventory management. The concept of this method is that it is based on parametric dynamic programming using convex piecewise linear approximations of the differential cost function in approximate relative value iteration steps and works for a range of stochastic control problems with linear dynamics and convex piecewise linear immediate costs. It is to be noted that the differential cost approximation which is obtained containing the lower bounds through which we come to know about the quality of the corresponding policy or other heuristics and it also gives us a policy that is found to be very good when it is considered in general.

Keywords: *Multiple suppliers, parametric dynamic programming and different cost approximation.*

*Department of Applied Science and Humanities, Sachdeva Institute of Technology, Farah, Mathura

**Department of Mathematics, Agra College, Agra

1. PROBLEM PORTRAYAL

As we know that the uncertainty in the reaction of a system to a sequence of actions or applied action, comes under the problems of operational management. To make any of the operational decisions one should know about the above mentioned concept. Generally, historical data exists that can be used to quantify these uncertainties and build a stochastic model that can be used for decision-making under uncertainty. A decision in a multi stage setting can depend on information exposed up to the corresponding time step, and it is necessary to find adaptive policies that utilize this information as better as it can.

In this paper, we introduce an approximate dynamic programming approach for systems with linear dynamics and piecewise affine convex state and control cost, using piecewise affine convex approximations of the value function of bounded complexity. Here we considered the application in the inventory management. In general, the focus is on the management of a warehouse facing stochastic demand and having multiple suppliers available with stochastic lead tiles. Decisions at each time are how to order at each supplier. On considering an infinite time horizon, our aim is to compute a replenishment policy for the warehouse that minimizes the average cost per time step. The costs consist of order costs, holding costs for inventory on stock and penalty costs for unmet demand, which is backlogged. Negative inventory level signifies the backlogged demand. To satisfy the non-order-crossing (NOC) assumption taken by Ehrhardt in 1984 & Kaplan in 1970, we consider the lead times for each supplier and it is a standard assumption in inventory management and means that open orders of one supplier have to arrive in the sequence they were placed. The structure of the optimal policy is known for a single supplies setting under NOC, and more efficiently we can computed the policy. In general it is not true for the problem based on multi-supplier stochastic lead-time.

2. ALLIED ANALYSIS

In the literature it is found that the single sourcing inventory management problem with NOC has been investigated very thoroughly. Iglehard in 1963 and Kaplan in 1970 introduced the optimal stationary policy [which is of (S, S) type and Ehrhardt and Veinott gives the several ways for the computation of these parameters. The extensions to this problem are to use multiple suppliers, to drop NOC, or both. The approach presented here is applicable for all these cases, while we present computational results on instances with two suppliers assuming NOC. General order cost structure is allowed by the single sourcing theory, on the other hand, if order costs are piecewise linear and convex, excluding fixed order costs then only there is a existence of the work of our theory. It is found that the results based on the multi-supplier problems are found very less in amount and the optimal policy structure is known only for some special cases e.g, for two suppliers with deterministic lead times of difference at most one. Fukuda shows that in this case the known optimal policy is to of (s, S)-type, but Whittemore shows that this does not hold anymore for the general lead time case. Houtum, Veeraraghavan and Scheller-Wolf computed that as soon as the difference gets larger than one, the optimal policy becomes highly state dependent. Some more results are found about optimal policies with two suppliers, since none allows stochastic lead time due to the order crossing between different suppliers, all these assumptions are become quite restricted. Presence of complication in the structure of optimal policies for a multi-supplier stochastic lead time inventory management, many works propose to use parameterized heuristics and focus on ways to find the optimal set of parameters. On the other hand, most approaches cover the case of deterministic lead times. Schedler-Wolf, Veeraraghavan and Houtum introduce the heuristics to which we compare our approach are the so-called Single Index Policy and Dual Index Policy. In the former, the decision is only dependent on the inventory position; the latter also considers the so-called expedited inventory position, which only counts open orders that will arrive within the lead time of the faster supplier. To approximate optimal policies for problems that are too big to solve exactly is approximated by the dynamic programming with a variety of methods is found in recent area of research. Berteskas and Tsitsiklis analyzed that the examples are temporal difference learning or Q-learning, which are so-called model-free learning methods. Another approach is TD- learning with basis functions instead of a look-up table to represent the value function approximation. To find good basis functions to approximate the optimal value function is a problem. Farias and Van Roy analyzed that this also holds for the linear programming approach using constraint sampling, which approximates the LP-formulation of the corresponding MDP. As a disadvantage, approximate dynamic programming methods are sensitive to the trade-off between exploitation and exploration. To find a good exploration/exploitation-strategy is a complicated problem by itself and a major difficulty when being applied to a particular problem, often requiring a lot of experience of the user and trial-and-error. Bertsekas and Tsitsiklis gives the too strong assumption for the considered inventory management problem, which shows that convergence results about the learning methods usually require that all states of the system are visited many times.

3. APPROACH AND METHODOLOGY

In this paper, we present a different approach to develop an approximation algorithm based on parametric dynamic programming, which was inspired by Lincoln and Rantzer. We represent the state (in our case the inventory and open orders of the past time steps up to the maximum lead time) as a continuous vector and then apply value iteration using a parametric function over this space instead of tabulating it. A value iteration step can be formulated as a parametric linear program with the state vector as the parameter when taken in our case then the value function can be shown to be piecewise linear and convex for any fixed number of value iteration steps. As the number of planes representing the different linear pieces can grow exponentially, the exact value iteration is become only of theoretical interest. Yet we use this structural approach to develop suitable approximation architecture. The approximation strategy we propose is to use relative value iteration (RVI) and to perform approximate parametric RVI steps by finding a limited representative set of planes for the next approximation of the differential cost function and it is containing the policy which can be replicated to obtain a confidence interval for the corresponding average cost per stage. Also we show that to obtain a valid lower bound for the average cost per stage, any such differential cost approximation can be used. To implement this scheme we introduce two algorithms:

1. The randomized relative value iteration (RRVI) randomly samples a number of states and compute the corresponding planes (while skipping redundant ones) for the next differential cost function approximation.
2. The approximate relative value iteration (ARVI) identifies those states that are important to further increase the lower bound for the average cost and adds the planes corresponding to these states.

Both algorithms together work best. To improve the average costs or to achieve a higher service level by shifting the differential cost approximation based on information about the system behavior under the obtained policy, we also give a post-processing method.

4. RESULTS

To test the developed algorithms, several problem instances of different state space dimension are used. The resulting policies are evaluated by simulation and the estimated average costs are compared to the lower bounds provided by the algorithms themselves. This allows providing an upper bound for the optimality gap. Also estimated costs are compared to results achieved using single and dual index policies. When RRVI & ARVI are applied, two different objectives can be formulated either for minimization of the average cost (computing a good policy) or maximization of the lower bound. It is found that by starting the approximation with RRVI and continuing with ARVI, achieved the best lower bounds. ARVI was developed to increase the lower bound and the policies provided by this method are very expensive to estimate the corresponding average cost or simply it's very complicated. On the other hand, the maximal number of planes are used to approximate the value function are bounded by RRVI and hence to construct policies for evaluation, respectively application purposes, are used by RRVI. For smaller problems it is not necessary to tackle these two objectives separately, since the policies resulting from the combined algorithm are simple enough. It is found that the computed policy is optimal because the gap between average cost and lower bound is essentially is zero. For larger problem instances, the RRVI does not provide good lower bounds but leads to reasonable policies, whereas the mixed algorithm leads to quite good lower bounds, but policies that are too complicated to simulate from. Within a few percent of the computed lower bound, the upper bound for the relative error can be computed by the results of both algorithms. Compared to the heuristics, the developed algorithm achieved always comparable or better results.

5. CONCLUSION

For the inventory management problem with multiple suppliers and stochastic lead time, we have developed and tested an approximation method. The method is based on parametric dynamic programming using convex piecewise linear approximations of the differential cost function is approximate relative value iteration steps. Besides being applicable to a broad range of problems with linear dynamics and convex piecewise linear costs, there are no artificial parameters to set besides the limit of the number of planes (and hence the complexity of policy). To judge the quality of the obtained policy as well as any other policy is used by the method yields with valid lower bounds for the average cost. On the considered test cases, the method was able to find an optimal policy for small instances, while on larger cases the estimated optimality gap was in the range of a few percent. For the cases of the considered problem we only know about the result that are described in this paper and apart of this we don't know any other method to find the better result than this.

6. REFERENCES

[1] D.P. Bertsekas and J. Tsitsiklis. Neuro-Dynamic Programming. Athena Scientific, 1996.

[2] R. Ehrhardt, (S, S) policies for a dynamic inventory model with stochastic lead times. Oper. Res. 32(1): 121-132, 1984.

[3] D.P. Farias and B. Van Roy. On constraint sampling in the linear programming approach to approximate dynamic programming. Math. Oper. Res., 29(3): 462-478, 2004.

[4] Y. Fukuda, Optimal policies for the inventory problem with negotiable leadtime. Management Sci., 10(4): 690-708, 1964.

[5] R.S. Kaplan, A dynamic inventory model with stochastic lead times. Management Sci., 16(7): 491-507, 1970.

[6] D.L. Iglehart, Optimality of (s, S) policies in the infinite horizon dynamic inventory problem, Management Sci., 9(2): 259-267, 1963.

[7] B. Lincoln and A. Rantzer, Relaxing dynamic programming, IEEE Trans. Automat. Contr., 51(8): 1249-1260, 2006.

[8] A. Scheller-Wolf and S. Veeraraghavan, Now or later: a simple policy for effective dual sourcing in capacitated systems. Oper. Res., 56(4): 850-864, 2008.

[9] A. Scheller-Wolf, S. Veeraraghavan, and G. J. Van Houtum, Effective dual sourcing with a single index policy, 2006.

[10] A.F. Veinott Jr. and H. M. Wagner, Computing optimal (S, S) inventory policies. Management Sci., 11(5): 525-552, 1965.

[11] A.S. Whittemore and S. C. Saunders, Optimal inventory under stochastic demand with two supply options, SIAM J. Appl. Math., 32(2): 293-305, 1977