On Efficiency of Solutions of Stochastic Optimal Control Problem with Discrete Time

Igor I. Gasanov
Department of Computational and Information Systems

Computing Centre of Russian Academy of Sciences
40 Vavilov Street, 119991 Moscow GSP-1, Russia
E-mail: gasanov@ccas.ru

Iouldouz S. Raguimov
Department of Mathematics and Statistics

York University
4700 Keele Street, Toronto, Ontario, Canada, M3J 1P3
E-mail: raguimov@mathstat.yorku.ca

***Abstract*—For stochastic optimal control problem with discrete time, the efficiency of solutions corresponding to the parameters of a stochastic process determined by the method of optimization on time series is analyzed in comparison to the solutions related to the parameters obtained using a common statistical method of estimation. Parametric optimization problems for continuous and discrete stochastic optimization problems are introduced and the corresponding problems of optimization on time series are formulated. When a sample size is increasing, the asymptotic properties of solutions to the considered problems are investigated. Theorems on the convergence of the optimal objective value of discrete problem of parametric optimization on time series to the optimal objective value of the related discrete stochastic optimization problem have been formulated and proved.**

Keywords-*Markov decision process; stochastic optimization; parametric optimization; optimization on time series.*

I. Introduction

The challenges of dealing with uncertainty is a common problem in the management of economic and engineering systems. When uncertainty is modeled probabilistically with random variables, it is usually required to be described as a multidimensional stochastic process for which neither a structure nor parameters are known. Particular challenges related to the estimation of the characteristics of random variables of a process as well as to determining of their interrelationships appear to be very important for this type of problems. At the same time, as a rule, operations research analyst is experiencing a data insufficiency in determining the structure and/or calibrating parameters of a stochastic process. Even in the case, when the model of a stochastic process has been formulated, we obtain an optimization problem that is usually too complicated to solve analytically.

As examples, a decision-making problem with uncertainty related to the natural factors as well as problems of functioning and interaction of financial and economic institutions including a financial portfolio management problem can be referred. To the problems of this type also belongs the equipment replacement problem, which arises when a company has to determine how long a machine should be utilized before being traded in for a new one.

In situations when it is neither possible, nor affordable to obtain an optimal solution analytically, so-called method of optimization on time series [3] is often used to determine the best approximate solution to the problem. Relying on information about realizations of uncontrollable uncertain parameters, it is determined a control for a considered object that would be optimal once were used in the past. Here, it is assumed implicitly that since the uncertainty has a regular character, then a control, which would have been optimal during some sufficiently prolonged period of time in the past, will also be optimal in the future.

The abovementioned idea appears to be rational, especially since the necessity of making decisions in such systems arises frequently, and the authors do not know an efficient alternative approach to solving this type of decision-making problems. On the other hand, this technique raises certain questions and doubts. Particularly, since optimal control is determined and estimated on the same set of realizations of a stochastic parameter, while constructing an optimal control on time series, to what extend are we exploiting systematic properties of the stochastic process, and to what – just are making adjustments by utilizing only some insignificant for the future properties of the stochastic process?

The analysis of this problem seems to be interesting and represents an actual challenge. In Section II of the paper, a parametric optimal control problem with discrete time is introduced. In Section III, the corresponding problem of parametric optimization on time series is constructed and theorem on convergence of its optimal objective value to the optimal objective value of the parametric optimal control problem with discrete time is formulated. In Section IV, optimization of parameters of stochastic process on time series is analyzed and the corresponding control problem with modified stochastic process is introduced. Theorem on convergence of its optimal objective value to the optimal objective value of the corresponding optimal control problem

 ISBN: 978-1-61208-226-4

with discrete time is formulated.

II. Parametric Optimal Control Problem with Discrete Time

Consider one of the possible formalizations of a stochastic optimal control problem, namely, the mathematical model of a discrete-time Markov decision process [7]. Suppose that at any time t, $t = 1, 2, \cdots, \infty$, the state of a system is given by the characteristic vector $A_t \in \hat{A} \subset \mathbb{R}^n$. Once the control u_t has been chosen at the stage (time period) t and the value $\tilde{\xi}_t$ of the stochastic parameter ξ is realized, the system moves on from the state A_t to the state

$$A_{t+1} = \varphi(A_t, \tilde{\xi}_t, u_t),$$

where the parameter $\xi \in \Xi \subset \mathbb{R}^m$ is a stationary Markov process with the transition probability function $\Phi(\xi_t \mid \xi_{t-1})$. So, every ordered pair $S_t = (A_t, \tilde{\xi}_t)$ of arguments of a function φ determines a state of the system at stage $t+1$. It is assumed that the initial probability distribution, i.e., the probability distribution $F^1(S_1)$ on the set of initial states of the system is known, and at each state S the control $u \in U(S) \subset \hat{U} \subset \mathbb{R}^k$. Here $\hat{A}$ and $\hat{U}$ are bounded sets.

Suppose that every stage t of the process is associated with a certain payoff function (expected reward) $h_t = h_t(S_t, u_t)$ and assume a decision-maker is interested in maximization of the average reward earned per period, i.e., is solving the following maximization problem

$$Q = \lim_{n\to\infty} \frac{1}{n} E(\Sigma_{i=1}^n h_t(S_t, u_t)) \implies \max_u \qquad (1)$$

Generally speaking, the decision maker wants to maximize function (1) with respect to $u_t = u(S_t)$, where u is a mapping $\hat{A} \times \Xi \to \hat{U}$ (it is assumed that at stage t to the moment of choosing u_t the realization $\tilde{\xi}_t$ of ξ is known).

It is clear that for the existence of the expected value in (1), the functions involved in the model should satisfy certain conditions. Analysis of these conditions is out of the scope of this paper. Related existence problems have been solved in [1], [2].

In [5], discrete models of the stochastic optimization problems are studied and the corresponding discrete problems of optimization on time series are introduced. The convergence of optimal solutions of the discrete problems of optimization on time series to an optimal solution of the discrete stochastic optimization problem has been proved. Properties of optimal solutions of discrete problems of optimization on time series are analyzed and estimates for the optimal objective values are obtained.

It worth noting, that the considered formulation covers a wide range of stochastic control problems. Particularly, it includes the case when it is assumed that at different stage of a process, realizations of a stochastic parameter are independent. A decision-making problem for static models with infinite horizon has been solved in [6].

According to Gasanov and Raguimov [5], the method of optimization on time series is usually applied to parametric optimization problem where a certain parametric class of control functions is considered and a problem is formulated as a problem of choosing optimal values for parameters of a function from the considered class.

Let us consider the problem of maximization of (1) on the set of control functions $\hat{U_\alpha}$ with $\alpha \in \mathbb{A}$, where $\hat{U_\alpha}$ is a class of the functions $u(S;\alpha)$, such that there exists a one-to-one correspondence between $\hat{U}_\alpha$ and $\mathbb{A}$. Therefore, the original problem is reduced to a problem of finding a value of α, such that

$$Q = \lim_{n\to\infty} \frac{1}{n} E(\Sigma_{t=1}^n h_t(S_t, u(S_t;\alpha))) \implies \max_\alpha \qquad (2)$$

Denote the formulated problem as <u>Problem 1</u> and compare it with its discrete analogue – <u>Problem 1D</u>, of maximization of function (2) on $\hat{U_\alpha^D}$, the parametric class of discrete functions. Here, $\hat{U_\alpha^D}$ is the set of discrete analogues of $u(S;\alpha) \in \hat{U_\alpha}$. It is supposed that the state space, the set of values of the stochastic parameter and the decision set are finite sets, and consequently, the state vectors of the system, the stochastic parameter and controls take the values on a finite grid, i.e. $A_t \in \{A^i\}_{i=1}^I = \hat{A^D} \subset \hat{A},\ \xi \in \{\xi^j\}_{j=1}^J = \Xi^D \subset \Xi$ and $u \in \{u^s\}_{s=1}^S \subset \hat{U_\alpha^D}$. Functions φ and Φ are modified correspondingly.

Consider the case when there exists a solution to Problem 1 and assume that these two problems are such that when an appropriate (small) mesh for the grid is chosen, solutions to Problem 1D closely approximate solutions to Problem 1. Therefore, the determining of a solution of Problem 1D is assumed to be the same as the finding of an approximate solution to Problem 1. Certainly, this assumption can be investigated in order ot obtain important analytical results. However, the authors believe that the assumption is valid for a wide range of practical problems and consequently, plausible from the point of view of applications.

III. Problem of Parametric Optimization on Time Series

Suppose, the sequence of realizations $\{\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T}\} \subset \Xi^D$ of stochastic parameter ξ and the initial state $\tilde{A}_1 \in \hat{A^D}$ of the system are given. Consider the following discrete optimization problem.

<u>Problem 1R.</u> Maximize the function

$$\tilde{Q^T} = \frac{1}{T} \Sigma_{t=1}^T h_t(A_t, \tilde{\xi}_t, u(A_t, \tilde{\xi}; \alpha)) \qquad (3)$$

on the set of control functions $u_t = u(A_t, \tilde{\xi}_t)$, subject to the constraint $A_1 = \tilde{A}_1$.

The control function u is said to be *everywhere optimal*, if it is optimal for every initial distribution $F^1(S_1)$.

The following theorem is proved.

Theorem 1. If there exists an everywhere optimal control for Problem 1D, then the optimal objective value of Problem 1R converges almost everywhere to the optimal objective value of Problem 1D, provided that the size of the sample $\{\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T}\} \subset \Xi^D$ increases unboundedly, i.e., T approaches infinity.

Therefore, under the abovementioned assumptions an optimal solution of Problem 1R represents a well-grounded estimate for an optimal solution of Problem 1D. At the same time, it is clear that a substantial limitation of decision set will affect the optimal value. It is difficult to measure this effect, unless Problem 1 and the original maximization problem (1) both are solved.

Also, as it was mentioned above, an optimal control for Problem 1R is determined and estimated on the same sample of the realizations, which may result in a displacement (particularly, in an overstatement) of the estimates. The possible range of this displacement for considered series of realizations of ξ is not considered in this paper.

IV. Optimization of Parameters of Stochastic Process on Time Series

Often, when either a given data does not allow to construct a reliable model of a stochastic process or Problem 1 is overly complicated to be solved in the original form, the considered stochastic process is replaced with a simple one, which according to the opinion of an operations research analyst reflects essential characteristics of the original process. The parameters $\beta \in \mathbb{R}^n$ of this auxiliary process are calibrated on the given series of realizations $\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T}$ and the problem with the accordingly modified stochastic process is considered. Denote the obtained problem as Problem 1M. Let $Q(u)$ be the objective value of Problem 1 corresponding to the control u and consider the parametric class of optimization problems of the type 1M, where as parameters the calibrated coefficients of the modified stochastic process are considered. Suppose u^β is a solution to Problem 1M, which corresponds to fixed values of the coefficients β and $Q(u^\beta)$ is the corresponding objective value. Consider the problem of maximization of $Q(u^\beta)$ on the set of calibrated parameters β and denote it as Problem 1A.

Let us also estimate the parameters of the stochastic process using one of the commonly used statistical methods, namely, using Monte-Carlo method, and consider the problem corresponding to the determined parameters. Denote the obtained problem as Problem 1S. As before, discrete analogues of the problems 1M, 1A and 1S can be formulated. Denote these problems as Problem 1MD, Problem 1AD and Problem 1SD, respectively.

The efficiency of a control function obtained by solving Problem 1MD can be estimated on the sample $\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T}$ by calculating the value of objective function (3). Now, formulate Problem 1MR as a problem of finding the values of the parameters of the stochastic model that maximize the value of objective function (3).

Let $u^{1SD}(\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T})$ and $u^{1MR}(\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T})$ be optimal control functions for Problem 1SD and Problem 1MR, correspondingly.

Using Theorem 1, the following theorem has been proved.

Theorem 2. If there exists an everywhere optimal control function u^{1AD} for Problem 1AD, then the optimal objective value $Q(u^{1MR}(\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T}))$ of Problem 1MR converges almost everywhere to the optimal objective value $Q(u^{1AD})$, as T approaches infinity.
Moreover, for every sequence of realizations $\{\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T}\}$,

$$Q(u^{1SD}(\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T})) \leq Q(u^{1AD}).$$

Since the stochastic model of Problem 1S is, at most, one of the elements of an heuristic procedure, it would be unfounded to assume that the values $Q(u^{1SD}(\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T}))$ will converge to the value of $Q(u^{1AD})$, as T approaches infinity. Therefore, the following conclusion can be deduced. Provided that the size of a sample $\tilde{\xi_1}, \tilde{\xi_2}, \cdots, \tilde{\xi_T}$ increases, the solving Problem 1MR is, generally speaking, a more efficient method to solve Problem 1 than the solving Problem 1M with the parameters of stochastic model estimated on the same sample using any of the commonly used statistical methods. In other words, the solving Problem 1MR as a method of solving Problem 1 is asymptotically preferred to the solving Problem 1M with the parameters estimated by any other statistical method.

We understand that the asymptotic preference of one method to another does not provide formal grounds to consider the first method as more efficient in solving practical problems where data samples are always limited and mostly not large enough. For applied problems with a stochastic process of a non-established structure, the theoretical evaluation of the method based on solving of Problem 1MR appears to be difficult. Therefore, to estimate the presented method from the practical point of view it is necessary to carry out series of computational experiments.

In [4], mathematical models of a controlled system containing a model of a stochastic process in the form of Markov process are implemented where the Markov process is modeling a financial market. Using the Markov process, data imitating series of observations is generated. Then, from the point of view of an operations research analyst who does not know the structure of the stochastic process but knows only the series of observations, various problems of optimization on time series have been investigated. Computational experiments are implemented for different size of data imitating the series of realizations and different behavior of the operations research analyst. The results, that is, the obtained optimal controls and their estimates on the

 ISBN: 978-1-61208-226-4

given series of realizations can be compared with their “real” efficiency, i.e., with the efficiency on the original Markov process.

Certainly, such experiments cannot be considered as a formal proof of efficiency of the presented method. Nevertheless, from the point of view of their further utilization, the results of the experiments seem to be essential.

V. Conclusions and Future Work

For the stochastic optimal control problem with discrete time, the efficiency of solutions has been analyzed. Solutions related to the values of the parameters of a stochastic process determined by the method of optimization on time series has been compared with the solutions related to the parameters obtained using a common statistical method of estimation. Parametric optimization problems for the corresponding continuous and discrete stochastic optimization problems have been introduced and the related problems of optimization on time series have been formulated. When a sample size increases, the asymptotic properties of solutions to the considered problems have been analyzed. Theorems on the convergence of the optimal objective value of a discrete problem of parametric optimization on time series to the optimal objective value of the discrete stochastic optimization problem have been formulated and proved. The authors are intending to extend the obtained results to stochastic decision-making problems for hidden Markov processes.

References

[1] D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control. The Discrete Time Case*, New York: Academic Press, 1978.

[2] E. B. Dynkin and A. A. Yushkevich, *Controlled Markov Processes*, New York: Springer-Verlag, 1979.

[3] G. A. Agasandyan, I. I. Gasanov, I. S. Menshikov, A. N. Chaban, and Yu. M. Chebanyuk, Calculation methods for problems of control of reservoir modes, in *Cybernetics and Computational Techniques,* vol. 3, Moscow: Nauka, 1987, pp. 57-101 (Russian).

[4] I. I. Gasanov and I. S. Raguimov, “On algorithm of financial portfolio control by the method of optimization on time series”, In the proceedings of the *2004 International Conference on Computational Intelligence for Modeling Control and Automation-CIMCA”2004* /ISBN 1740881885, pp. 511-519.

[5] I. I. Gasanov and I. S. Raguimov, “On solution of stochastic control problem by the method of optimization on time series”, In the proceedings of the *2003 Hawaii International Conference on Statistics and Related Fields* /ISSN 1539-7211, pp. 1-7.

[6] I. S. Raguimov, “On decision making in operations with stochastic factors”, In the proceedings of the *2001 International Conference on Computational Intelligence for Modeling Control and Automation-CIMCA”2001* /ISBN 0858898470, pp. 580-587.

[7] W. L. Winston, *Operations Research: Applications and Algorithms*, Belmont, California: Duxbury Press, 1994.