STOCHASTIC SYSTEMS

On the Existence of Optimal Strategies
in the Control Problem for a Stochastic Discrete Time
System with Respect to the Probability Criterion

A. I. Kibzun* and A. N. Ignatov*

*Moscow Aviation Institute, Moscow, Russia
e-mail: kibzun@mail.ru, alexei.ignatov1@gmail.com

Received December 16, 2016

Abstract—We study a control problem for a stochastic system with discrete time. The optimality criterion is the probability of the event that the terminal state function does not exceed a given limit. To solve the problem, we use dynamic programming. The loss function is assumed to be lower semicontinuous with respect to the terminal state vector, and the transition function from the current state to the next is assumed to be continuous with respect to all its arguments. We establish that the dynamic programming algorithm lets one in this case find optimal positional control strategies that turn out to be measurable. As an example we consider a two-step problem of security portfolio construction. We establish that in this special case the future loss function on the second step turns out to be continuous everywhere except one point.

Keywords: dynamic programming, stochastic system, discrete time, measurable positional strategy.

DOI: 10.1134/S0005117917100083

1. INTRODUCTION

Stochastic systems with discrete time occur in many applied problems, e.g., in the aerospace industry for correcting the position of a flying vehicle [1], and also in economics in the problem of security portfolio construction [2]. But usually such problems consider as their criterion the expectation of the loss function on the last step [3, 4]. It is known that in a number of problems this criterion leads to unsatisfactory results. For example, in the security portfolio construction problem the use of the average revenue criterion leads to a stock market paradox [5], when for the resulting optimal strategy average revenue tends to infinity under multiple uses, while the probability of ruin tends to one.

In many applied problems, it makes more sense to use the quantile and probabilistic criteria [5]. Here the probability criterion is defined as the probability of the loss function on the last step not exceeding a given limit. The quantile criterion is, in a way, inverse with respect to the probability criterion and equals to the limit that the loss function will not exceed with a given confidence probability.

Usually, control problems for stochastic control systems choose strategies on every step in the class of positional strategies that depend on the current system state; unlike program strategies, they result in a better value for the criterion. On the other hand, it is a rather hard problem to optimize the quality functional on the space of functions. To solve the problem of finding optimal positional strategies, we use the dynamic programming method [6]. The work [7] shows that dynamic programming is applicable to solving the optimal control problem for a stochastic system with discrete time with respect to the criterion of loss function expectation. This is related
to the fact that the expectation operator is additive and possesses the Markov property [7], namely, it can be represented as a superposition of operators each of which is defined for the next step of the discrete time system. The work [7] also formulates the measurable choice problem which is to show that the optimal strategies obtained by applying the dynamic programming algorithm turn out to be measurable with respect to current state vectors. In the opposite case, expectations in the dynamic programming algorithm turn out to be uncertain from the point of view of probability theory. To overcome this problem, the work [7] proposes to look for optimal control strategies in the class of the so-called universally measurable strategies. A mathematical description of this class is rather complex and hard to implement in practice. Note that similar problems have also been considered in the book [8], but [7] considers operators in a more general form, not only in the form of an expectation. The work [8] shows conditions under which there exists an optimal measurable strategy in the so-called semicontinuous model [8, pp. 71–74]. In this work we consider a special case of that model. However, conditions shown in [8] are hard to check; in essence, checking these conditions for the case considered in this paper will lead to a different proof of the statements formulated there. In this work, we choose a scheme of proving the statements shown in [7].

The quantile criterion is not additive and does not possess the Markov property [1], so one cannot apply dynamic programming to solve the control problem for a stochastic system with respect to the quantile criterion. The work [9] suggests to use the confidence approach, where the control problem with respect to the quantile criterion is replaced with a minimax problem for which dynamic programming turns out to be applicable since the minimax criterion is additive and possesses the Markov property [7]. But the question arises of how accurate this approximation actually is.

The probability criterion can be written in the form of expectation of the indicator function for the terminal state. Therefore, to solve the optimal control problem for a stochastic system with respect to the probability criterion dynamic programming turns out to be applicable. Studies of control problems with the probability criterion have been conducted in [10–12], but they did not set the conditions under which the optimal strategies obtained with dynamic programming are measurable.

In this work, we consider the optimal control problem for a stochastic system with discrete time where the transition function from the current state to the next one is continuous with respect to their arguments, and the terminal state function is lower semicontinuous. Under these conditions, and also given that the random vectors on every step are independent, it turns out that after applying the dynamic programming algorithm for solving the formulated control problem for a stochastic system with respect to the probability criterion one gets optimal measurable strategies, and the future loss function on every step turns out to be upper semicontinuous. We consider an example of sequentially constructing a security portfolio in two steps. It turns out that in this case all existence conditions for optimal measurable strategies hold. Moreover, we show that the future loss function on the second step is not only upper semicontinuous, but also continuous at all points except one.

2. PROBLEM SETTING

Consider a dynamic stochastic system in discrete time

\[ z_{i+1} = f_i(z_i, u_i, X_i), \quad i = 1, N, \]  

(1)

where \( z_i \in \mathbb{R}^s \) is the state vector, \( u_i \in U_i \subset \mathbb{R}^m \) is the control vector, \( X_i \) is the vector of random disturbances with values from \( \mathbb{R}^m \), \( N \) is the system horizon. Suppose that random vectors \( X_i \) are continuous and have distribution functions \( F_{X_i}(x_i) \).