AN APPROACH FOR STUDYING AND SOLVING STOCHASTIC DISCRETE CONTROL PROBLEMS WITH FINITE TIME HORIZON

Maria CAPCELEA

Chair of Applied Mathematics

Este formulată o clasă de probleme stocastice de control optimal discret ce extinde clasa problemelor deterministe cu un orizont de timp finit. Se propune o metodă de soluționare a acestor probleme bazată pe metoda rețelei temporale extinse.

Introduction and Problems Formulations

We consider a time-discrete system L with a finite set of states $X \subset R^n$. At every time-step $t = 0,1,2,\ldots$, the state of the system L is $x(t) \in X$. Two states x_0 and x_f are given in X, where $x_0 = x(0)$ represents the starting state of system L and x_f is the state in which the system L must be brought, i.e. x_f is the final state of L. We assume that the system L should reach the final state x_f at the time-moment $T(x_f)$ such that $T_1 \le T(x_f) \le T_2$, where T_1 and T_2 are given. The dynamics of the system L is described as follows

$$x(t+1) = g_t(x(t),u(t)),\ t = 0,1,2,\ldots, \tag{1}$$

where

$$x(0) = x_0 \tag{2}$$

and $u(t) = (u_1(t),u_2(t),\ldots,u_m(t)) \in R^m$ represents the vector of control parameters. For any time-step t and an arbitrary state $x(t) \in X$ a feasible finite set $U_t(x(t)) = \{u^1_{x(t)}, u^2_{x(t)},\ldots,u^{k(x(t))}_{x(t)}\}$, for the vector of control parameters $u(t)$ is given, i.e.

$$u(t) \in U_t(x(t)),\ t = 0,1,2,\ldots. \tag{3}$$

We assume that in (1) the vector functions $g_t(x(t),u(t))$ uniquely are determined by $x(t)$ and $u(t)$, i.e. the state $x(t+1)$ is determined uniquely by $x(t)$ and $u(t)$ at every time-step $t = 0,1,2,\ldots$. In addition we assume that at each moment of time t the cost $c_t(x(t),x(t+1)) = c_t(x(t),g_t(x(t),u(t)))$ of system's transaction from the state $x(t)$ to the state $x(t+1)$ is known.

Let

$$x_0 = x(0), x(1), x(2),\ldots, x(t),\ldots$$

be a trajectory generated by given vectors of control parameters

$$u(0),u(1),\ldots,u(t-1),\ldots.$$

Then either this trajectory passes through the state x_f at the time-moment $T(x_f)$ or it does not pass through x_f.

We denote

$$F_{x_0x_f}(u(t)) = \sum_{t=0}^{T(x_f)-1} c_t(x(t),g_t(x(t),u(t))) \tag{4}$$

the integral-time cost of system's transactions from x_0 to x_f if $T_1 \le T(x_f) \le T_2$; otherwise we put $F_{x_0x_f}(u(t)) = \infty$.

In [1, 2, 4] have been formulated and studied the following problem: to determine the vectors of control parameters $u(0),u(1),\ldots,u(t),\ldots$ which satisfy conditions (1)-(3) and minimize functional (4).

This problem can be regarded as a control model with controllable states of dynamical system because for an arbitrary state $x(t)$ at every moment of time the choosing of vector of control parameter $u(t) \in U_t(x(t))$ is assumed to be at our disposition.

In the following we consider the stochastic versions of the control model formulated above. We assume that the dynamical system L may contains uncontrollable states, i.e. for the system L there exists dynamical states in which we are not able to control the dynamics of the system and the vector of control parameters $u(t) \in U_t(x(t))$ for such states is changing in the random way according to given distribution function

$$p : U_t(x(t)) \to [0,1], \quad \sum_{i=1}^{k(x(t))} p(u^i_{x(t)}) = 1 \tag{5}$$

on the corresponding dynamical feasible sets $U_t(x(t))$. If an arbitrary dynamic state $x(t)$ of system L at a given moment of time t we regard as position (x,t) then the set of positions

$$Z = \{(x,t) | x \in X, t = 0,1,2,\dots,T_2\}$$

of dynamical system can be divided into two disjoint subsets

$$Z = Z^C \bigcup Z^N \quad (Z^C \bigcap Z^N = \varnothing),$$

where Z^C represents the set of controllable positions of L and Z^N represents the set of positions $(x,t) = x(t)$ for which the distribution function (5) of the vectors of control parameters $u(t) \in U_t(x(t))$ are given. This mean that the dynamical system L works as follows. If the starting point belong to controllable positions then the decision maker fix a vector of control parameter and we obtain the state $x(1)$. If the starting state belong to the set of uncontrollable positions then the system passes to the next state in the random way. After that if at the time-moment $t = 1$ the state $x(1)$ belong to the set of controllable positions then the decision maker fix the vector of control parameter $u(t) \in U_t(x(t))$ and we obtain the state $x(2)$. If $x(1)$ belong to the set of uncontrollable positions then the system passes to the next state in the random way and so on. In this dynamic process the final state may be reached at given moment of time with a probability which depend on the control of the system in the deterministic states as well as the expectation of integral time cost by trajectory depends on control of the system in these states. Therefore our main concentration will be addressed on studying and solving the following classes of problems.

Problem 1. For given vectors of control parameters $u(t) \in U_t(x(t))$, $x(t) \in Z^C$, to determine the probability that the dynamical system L with given starting state $x_0 = x(0)$ will reach the final state x_f at the moment of time $T(x_f)$ such that $T_1 \le T(x_f) \le T_2$. This probability we denote $P_{x_0}(u(t), x_f, T_1 \le T(x_f) \le T_2)$; if $T_1 = T_2 = T$ then we use the notation $P_{x_0}(u(t), x_f, T)$.

Problem 2. To find the vectors of control parameters $u^*(t) \in U_t(x(t))$, $x(t) \in Z^C$ for which the probability in Problem 1 is maximal. This probability we denote $P_{x_0}(u^*(t), x_f, T_1 \le T(x_f) \le T_2)$; in the case $T_1 = T_2 = T$ we shall use the notation $P_{x_0}(u^*(t), x_f, T)$.

Problem 3. For given vectors of control parameters $u(t) \in U_t(x(t))$, $x(t) \in Z^C$ and given number of stages T to determine the expectation of integral-time cost of the system after T transactions when it starts transactions in the state $x_0 = x(0)$ at the moment of time $t = 0$. This expectation we denote $Exp_{x_0}(u(t),T)$.

Problem 4. To determine the vectors of control parameters $u^*(t) \in U_t(x(t))$, $x(t) \in Z^C$ for which the expectation of integral-time cost for dynamical system in Problem 3 is minimal. This expectation we denote $Exp_{x_0}(u^*(t), x_f, T)$.

Problem 5. For given vectors of control parameters $u(t) \in U_t(x(t))$, $x(t) \in Z^C$, to determine the expectation of integral-time cost of system's transactions from starting state x_0 to final state x_f when the final state is reached at the time-moment $T(x_f)$ such that $T_1 \le T(x_f) \le T_2$. This expectation we denote $Exp_{x_0}(u(t), x_f, T_1 \le T(x_f) \le T_2)$; in the case $T_1 = T_2 = T$ this expectation we denote $Exp_{x_0}(u(t), x_f, T)$.

Problem 6. To determine the vectors of control parameters $u^*(t) \in U_t(x(t))$, $x(t) \in Z^C$ for which the expectation of integral-time cost of system's transactions in Problem 5 is minimal. This expectation we denote $Exp_{x_0}(u^*(t), x_f, T_1 \le T(x_f) \le T_2)$; in the case $T_1 = T_2 = T$ this expectation we denote $Exp_{x_0}(u^*(t), x_f, T)$.

For an additional characterization of the finite stochastic processes we introduce also the notion of the variance of integral-time cost for the dynamical system and the corresponding problem of determining the variance in such processes will be considered. Note that in these problems the probability $P_{x_0}(u(t), x_f, T_1 \le T(x_f) \le T_2)$, the expectation $Exp_{x_0}(u(t), x_f, T_1 \le T(x_f) \le T_2)$ as well the variance need to be specified and more strictly defined. These notions we define in accordance with the basic notions of decision Markov processes and control theory. This will allow us to formulate more accurate our problems and to solve them in general form.

The considered problems comprises a large class of deterministic and stochastic dynamic problems from [1, 2, 3]. The problems from [3] related to finite Markov processes became Problems 1-3 in the case when $Z^C = \varnothing$, $T_1 = T_2 = T$ and when the probabilities $p(u^i_{x(t)})$ do not depend on time but depend only on the states. The discrete optimal control problems from [1, 2] became Problems 4-6 in the case $Z^N = \varnothing$. In the following we propose algorithms for solving the problem formulated above based on results from [1, 2, 3] and time-expended method from [4, 5].

The problems formulated above can be studied and solved separately, however the combined joint solution of some of them also may be asked and justified. For example, if we solve Problem 2 and find the control with the maximal probability of system transactions from the starting state to the final one then after that it has reason for the optimal control found to estimate the expected integral time cost of states transactions for the dynamical system , i.e. we have to solve additionally Problem 5. If we solve Problem 6 and find the control which provide the maximal expectation of integral time cost of states transactions of the system from x_0 to x_f then after that it has sense for such optimal control to estimate the probability of system passage from x_0 to x_f, i.e. we have to solve additionally Problem 1. So, such combined solution of the problems formulated above may be useful for practical point of view.

2. Definitions of the Basic Notions for Stochastic Discrete Control Problems

We have already noted that for studying and solving our problems it is necessary to define strictly the state probabilities and the expectation of integral-time cost. Below we specify and define these notions for Problems 1-6 in accordance with the basic notions from [3].

2.1. Definition of the State Probability and The Expectation of Integral-Time Cost

In this subsection we specify the notions of state probabilities $P_{x(0)}(u(t), x, T)$, $P_{x(0)}(u(t), x_f, T_1 \le T(x_f) \le T_2)$ and the expectations of integral-time cost $C_{x(0)}(u(t), T)$, $C_{x(0)}(u(t), x_f, T)$, $C_{x(0)}(u(t), x_f, T_1 \le T(x_f) \le T_2)$ for dynamical system mentioned in our problems using the definitions of state probabilities and the expectation of integral-time cost from previous chapters. First of all we stress our attention to the definition of the probability $P_{x_{i_0}}(u(t), x, T)$ for the dynamical system L. For given starting state x_{i_0}, given time-moment T and fixed control $u(t)$ we define this probability in the following way. We consider that a transaction of the system from an arbitrary controllable state $x = x(t)$ to the next state $y = x(t+1)$ generated by the control $u(t)$ is made with probability $p_{x,y} = 1$ and the rest of probabilities of system's transactions from x at the moment of time t to the next states are equal to zero. Thus we obtain a finite Markov process for which the probability of system passage from starting state x_{i_0} to final state x by using T unites of time can be defined. This probability we denote $P_{x_{i_0}}(u(t), x, T)$. The probability $P_{x_{i_0}}(u(t), x, T_1 \le T(x) \le T_2)$ for given T_1 and T_2 we define as probability of the dynamical system L to reach the state x at least at the one of the moment of times $T_1, T_1+1, \ldots, T_2$.

In order to define strictly the expectation of integral-time cost of dynamical system in Problems 3-6 we shall use the notion of expectation of integral-time cost for Markov processes with costs on transactions introduced in previous subsection. The expectation of integral-time cost $C_{x_{i_0}}(u(t), T)$ of system L in Problem 3 for fixed control $u(t)$ we define as the expectation of the integral-time cost during T transitions of dynamical system in the Markov process generated by the control $u(t)$ and the the corresponding costs of state's transactions of dynamical system.

In the following we shall use the random graph with given probability function $p: ER \to R$ on edge set ER and given distinguished vertices which correspond to starting and final states of the dynamical system. To the edges we will associate also the costs which correspond to the cost of system passage from one state to another. Such random graph we call stochastic network. Further the stochastic networks we will extend for the non-stationary Markov processes and will use for calculation of the probabilities $P_{x_{i_0}}(x, T_1 \le T(x) \le T_2)$ and the expectations $Exp_{x_{i_0}}(x, T)$, $Exp_{x_{i_0}}(x, T_1 \le T(x) \le T_2)$.

3. The Main Approach and Algorithms for Determining the State Probabilities of the System in the Control Problems on Stochastic Networks

In order to provide a better understanding of the main approach and to ground the algorithms for solving the problems formulated in Section 1 we shall use the network representation of the dynamics of the system and will formulate these problems on stochastic network. Note that in our control problems the probabilities and the costs of system's passage from one state to another depend on time. Therefore here we develop the time-expended network method from [4, 5] for the stochastic versions of control problems and reduce them to the static case of the problems. This will allow us to describe dynamic programming algorithms for solving the problems on static stochastic networks. At first we show how to construct the stochastic network and how to solve the problems with fixed number of stages, i.e. we consider the case $T_1 = T_2 = T$.

3.1. Construction of the Stochastic Time-Expended Network with Fixed Number of Transactions

If the dynamics of discrete system L and the information related to the feasible sets $U_t(x(t))$ and the cost functions $c_t(x(t), g_t(x(t), u(t)))$ in the problems with $T_1 = T_2 = T$ are known then our stochastic network can be obtained in the following way. Each position (x,t) which correspond to a dynamic state $x(t)$ we identify with a vertex $z = (x,t)$ of the network. So, the set of vertices Z of the network can be represented as follows

$$Z = Z_1 \cup Z_2 \cup \ldots \cup Z_T,$$

where

$$Z_t = \{(x,t) | x \in X\},\ t = 0,1,2,\ldots,T.$$

To each vector of control parameters $u(t) \in U_t(x(t)), t = 1,2,\ldots,T-1$ which provide a system passage from the state $x(t) = (x,t)$ to the state $x(t+1) = (y, t+1)$ we associate in our network a directed edge $e(z,w) = ((x,t),(y,t+1))$ from the vertex $z = (x,t)$ to the vertex $w = (y,t+1)$, i.e., the set of edges E of the network is determined by the feasible sets $U_t(x(t))$. After that to each directed edge $e = (z,w) = ((x,t),(y,t+1))$ originating in the uncontrollable positions (x,t) we put in correspondence the probability $p_e = p(u^i(t))$, where $u^i(t)$ is a vector of control parameter which provide the passage of the system from the state $x = x(t)$ to the state $x(t+1)=(y,t+1)$. Thus if we distinguish in E the subset of edges $E_N=\{e=(z,w)\in E \mid z\in Z^N\}$ originating in uncontrollable positions Z^N then on E_N we obtain the probability function $p:E\to R$ which satisfy the condition

$$\sum_{e\in E^+(z)} p_e=1,\ \ z\in Z^N\setminus Z_T$$

where $E^+(z)$ is the set of edges originating in z, i.e. $E^+(z)=\{e=(z,w)\,|\,e\in E, w\in Z\}$. In addition in the network we add to the edges $e=(z,w)=((x,t),(y,t+1))$ the costs $c_{(z,w)}=c((x,t),(y,t+1))=c_t(x(t),x(t+1))$ which correspond to the costs of system's passage from states $x(t)$ to the states $x(t+1)$. The subset of edges of the graph G originating in vertices $z\in Z^C$ we denote E_C, i.e. $E_C=E\setminus E_N$.

So, our network is determined by the tuple $(G,Z^C,Z^N,z_0,z_f,c,p,T)$, where $G=(Z,E)$ is the graph which describe the dynamics of the system; the vertices $z_0=(x_{i_0},0)$ and $z_f=(x_f,0)$ correspond to the starting and the final states of the dynamical system, respectively; c represents the cost function defined on the set of edges E and p is the probability function defined on the set of edges E_N which satisfy condition (5). Note that $Z=Z^C\bigcup Z^N$, where Z^C is a subset of vertices of G which correspond to the set of controllable positions of dynamical system and Z^N is a subset of vertices of G which correspond to the set of uncontrollable positions of system L. In addition we shall use the notation Z_t^C and Z_t^N, where $Z_t^C=\{(x,t)\in Z_t\,|\,(x,t)\in Z^C\}$ and $Z_t^N=\{(x,t)\in Z_t\,|\,\ (x,t)\in Z^C\}$.

The notation of stochastic network for different problems in the following may be specified. As example, for Problems 1 and 2 the information about the cost function c is not required, therefore this function in the notation of the network we will omit; for Problems 3, 4 the final state is not given and therefore in the notation of the stochastic network we will not use it.

It is easy to observe that after the construction described above the Problem 1 in the case $T_1=T_2=T$ can be formulated and solved on stochastic network (G,Z^C,Z^N,z_0,z_f,p,T). A control $u(t)$ of system L in this network means a fixing a passage from each controllable position $z=(x,t)$ to the next position $z=(x,t)$ through a leaving edge $e=(z,w)=((x,t),(y,t+1))$ generated by $u(t)$; this is equivalent with an association to these leaving edges the probability $p_e=1$ of the system's passage from the state (x,t) to the state $(y,t+1)$ considering $p_e=0$ for the rest of leaving edges. In other words a control on stochastic network means an extension of the probability function p from E_N to E by adding to the edges $e\in E\setminus E_N$ the probabilities p_e according to the mentioned above rule. We denote this probability function on E by p^u and will keep in mind that $p_e^u=p_e$ for $e\in E\setminus E_N$ and on E_C this function satisfy the following property

$$p^u:E_C\to\{0,1\},\quad \sum_{e\in E_C^+(z)} p_e^u=1 \text{ for } z\in Z^C,$$

induced by the feasible control $u(t)$ in the problems from Section 1. In general we can start with the definition of the control on stochastic network as a map

$$p^u:E_C\to\{0,1\}$$

which satisfy the condition $\sum_{e\in E_C^+(z)} p_e^u=1$ for $z\in Z\setminus\{z_f\}$ and then to show that this map uniquely determine a feasible control $u_p(t)$ for the problems from Section 1. So, each feasible control $u(t)$ uniquely define the function p_e on E and vice versa, i.e. each probability function p^u on E_C uniquely determine a feasible control $u_p(t)$ for the Problems 1-6. Therefore if the control $u(t)$ is given then the stochastic network we denote $(G, Z^C, Z^N, z_0, z_f, c, p^u, T)$. If the control $u(t)$ in the Problems 1-6 is not fixed then for the stochastic network we shall use the notation $(G, Z^C, Z^N, z_0, z_f, c, p, T)$. For the state probabilities of the system L on this stochastic network we shall use a similarly notations $P_{z_0}(u(t), z, T), P_{z_0}(u(t), z, T_1 \leq T(z) \leq T_2)$ and each time we will specify on which network they are calculated, i.e. will take into account that these probabilities are calculated by using the probability function on edges p^u which already do not depend on time.

References:

1. Bellman R., Kalaba R. Dynamic programming and modern control theory. - New York and London: Academic Press, 1965.
2. Boltjanski W.G. Optimale Steuerung diskreter Systeme. - Leipzig Akademische Verlagsgesell schaft Geest & Portig K.-G., Leipzig, 1976.
3. Howard R.A. Dynamic Programming and Markov Processes. - Wiley, 1960.
4. Lozovanu D., Pickl S. Optimization and Multiobjective Control of Time-Discrete Systems. - Springer, 2009.
5. Lozovanu D., Pickl S. Algorithms and the calculation of Nash equilibria for multi-objective control of time-discrete systems and polynomial-time algorithms for dynamic c-games on networks. - European Journal of Operational Research, 2007, 181, 1214-1232.

Prezentat la 07.12.2009