BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 3(76), 2014, Pages 80–88
ISSN 1024–7696

Algorithms for solving stochastic discrete optimal control problems on networks

Dmitrii Lozovanu, Maria Capcelea

Abstract. In this paper we consider the stationary stochastic discrete optimal control problem with average cost criterion. We formulate this problem on networks and propose polynomial time algorithms for determining the optimal control by using a linear programming approach.

Mathematics subject classification: 93E20, 90C05.
Keywords and phrases: Time-discrete system, optimal control, stochastic network, Markov process, optimal strategy, linear programming.

1 Problem Formulation

Let a discrete dynamical system $\mathbb{L}$ with finite set of states X be given, where $|X| = n$. At every discrete moment of time $t = 0, 1, 2, \dots$ the state of $\mathbb{L}$ is $x(t) \in X$. The dynamics of the system is described by a directed graph of states' transitions $G = (X, E)$ where the set of vertices X corresponds to the set of states of the dynamical system and an arbitrary directed edge $e = (x, y) \in E$ expresses the possibility of the system $\mathbb{L}$ to pass from the state $x = x(t)$ to the state $y = x(t+1)$ at every discrete moment of time t. So, a directed edge $e = (x, y)$ in G corresponds to a stationary control of the system in the state $x \in X$ which provides a transition from $x = x(t)$ to $y = x(t+1)$ for every discrete moment of time t. We assume that graph G does not contain deadlock vertices, i.e., for each vertex x there exists at least one leaving directed edge $e = (x, y) \in E$. In addition, we assume that with each edge $e = (x, y) \in E$ a quantity $c_e \in \mathbb{R}$ is associated, which expresses the cost of the system $\mathbb{L}$ to pass from the state $x = x(t)$ to the state $y = x(t)$ for every $t = 0, 1, 2, \dots$.

A sequence of directed edges $E' = \{e_0, e_1, e_2, \dots, e_t, \dots\}$, where $e_t = (x(t), x(t+1))$, $t = 0, 1, 2, \dots$, determines in G a control of the dynamical system with a fixed starting state $x_0 = x(0)$. An arbitrary control in G generates a trajectory $x_0 = x(0), x(1), x(2), \dots$ for which the average cost per transition can be defined in the following way

$$f(E') = \lim_{t \to \infty} \frac{1}{t} \sum_{\tau=0}^{t-1} c_{e_\tau}.$$

In [1] it is shown that this value exists and $|f_{x_0}(E')| \le \max_{e \in E'} |c_e|$. Moreover, in [1] it is shown that if G is strongly connected, then for an arbitrary fixed starting state $x_0 = x(0)$ there exists the optimal control $E^* = \{e_0^*, e_1^*, e_2^* \dots\}$ for which

$$f(E^*) = \min_{E'} \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} c_{e_\tau}$$

and this optimal control does not depend either on the starting state or on time. Therefore, the optimal control for this problem can be found in the set of stationary strategies $\mathbb{S}$.

We assume that the set of states X of the dynamical system may admit states in which the system $\mathbb{L}$ makes transitions to the next state in a random way according to a given distribution function of probabilities on the set of possible transitions from these states [2]. So, the set of states X is divided into two subsets X_C and X_N $(X = X_C \cup X_N,\ X_C \cap X_N = \emptyset)$, where X_C represents the set of states $x \in X$ in which the transitions of the system to the next state y can be controlled by the decision maker at every discrete moment of time t and X_N represents the set of states $x \in X$ in which the decision maker is not able to control the transition because the system passes to the next state y randomly. Thus, for each $x \in X_N$ a probability distribution function $p_{x,y}$ on the set of possible transitions (x, y) from x to $y \in X(x)$ is given, i. e.,

$$\sum_{y\in X(x)} p_{x,y} = 1, \quad \forall x \in X_N; \quad p_{x,y} \geq 0, \quad \forall y \in X(x). \tag{1}$$

Here $p_{x,y}$ expresses the probability of the system's transition from the state x to the state y for every discrete moment of time t.

We call the graph G, with the properties mentioned above, decision network and denote it by (G, X_C, X_N, c, p, x_0). So, this network is determined by the directed graph G with a fixed starting state x_0, the subsets X_C, X_N, the cost function $c : E \to \mathbb{R}$ and the probability function $p : E_N \to [0, 1]$ on the subset of the edges $E_N = \{e = (x, y) \in E \,|\, x \in X_N,\ y \in X\}$, where p satisfies the condition (1). If the control problem is considered for an arbitrary starting state, then we denote the network by (G, X_C, X_N, c, p).

We define a stationary strategy for the control problem on networks as a map:

$$s : x \to y \in X(x) \quad \text{for} \quad x \in X_C,$$

where $X(x) = \{y \in X | e = (x, y) \in E\}$.

Let s be an arbitrary stationary strategy. Then we can determine the graph $G_s = (X, E_s \cup E_N)$, where $E_s = \{e = (x, y) \in E \,|\, x \in X_C,\ y = s(x)\}$. This graph corresponds to a Markov process with the probability matrix $P^s = (p^s_{x,y})$, where

$$p^s_{x,y} = \begin{cases} p_{x,y} & \text{if} \quad x \in X_N \quad \text{and} \quad y = X; \\ 1 & \text{if} \quad x \in X_C \quad \text{and} \quad y = s(x); \\ 0 & \text{if} \quad x \in X_C \quad \text{and} \quad y \neq s(x). \end{cases}$$

In the considered Markov process, for an arbitrary state $x \in X_C$, the transition $(x, s(x))$ from the states $x \in X_C$ to the states $y = s(x) \in X$ is made with the probability $p_{x,s(x)} = 1$ if the strategy s is applied. For this Markov process we can determine the average cost per transition for an arbitrary fixed starting state $x_i \in X$. Thus, we can determine the vector of average costs ω^s, which corresponds to the strategy s, according to the formula $\omega^s = Q^s \mu^s$, where Q^s is the limit matrix of the Markov process, generated by the stationary strategy s, and μ^s is the corresponding vector of the immediate costs, i.e., $\mu_x^s = \sum_{y \in X(x)} p_{x,y}^s c_{x,y}$ [3]. A component ω_x^s of the vector ω^s represents the average cost per transition in our problem with a given starting state x and a fixed strategy s, i.e., $f_x(s) = \omega_x^s$.

In such a way we can define the value of the objective function $f_{x_0}(s)$ for the control problem on a network with a given starting state x_0, when the stationary strategy s is applied.

The control problem on the network (G, X_C, X_N, c, p, x_0) consists of finding a stationary strategy s^* for which

$$f_{x_0}(s^*) = \min_s f_{x_0}(s).$$

2 A Linear Programming Approach for Determining Optimal Stationary Strategies on Perfect Networks

We consider the stochastic control problem on the network (G, X_C, X_N, c, p, x_0) with $X_C \neq \emptyset$, $X_N \neq \emptyset$ and assume that G is a strongly connected directed graph. Additionally, we assume that in G for an arbitrary stationary strategy $s \in \mathbb{S}$ the subgraph $G_s = (X, E_s \cup E_N)$ is strongly connected. This means that the Markov chain induced by the probability transition matrix P^s is irreducible for an arbitrary strategy s. We call the decision network with such a condition a *perfect network*. At first we describe an algorithm for determining the optimal stationary strategies for the control problem on perfect networks.

So, in this section we consider the control problem that the average cost per transition is the same for an arbitrary starting state, i. e., $f_x(s) = \omega^s, \quad \forall x \in X$.

Let $s \in \mathbb{S}$ be an arbitrary strategy. Taking into account that for every fixed $x \in X_C$ we have a unique $y = s(x) \in X(x)$, we can identify the map s with the set of boolean values $s_{x,y}$ for $x \in X_C$ and $y \in X(x)$, where

$$s_{x,y} = \begin{cases} 1 & \text{if} \quad y = s(x); \\ 0 & \text{if} \quad y \neq s(x). \end{cases}$$

For the optimal stationary strategy s^* we denote the corresponding boolean values by $s^*_{x,y}$.

Assume that the network (G, X_C, X_N, c, p, x_0) is perfect. Then the following lemma holds.

Lemma 1. *A stationary strategy s^* is optimal if and only if it corresponds to an optimal solution q^*, s^* of the following mixed integer bilinear programming problem:*

Minimize

$$\psi(s,q) = \sum_{x\in X_C}\sum_{y\in X(x)} c_{x,y}s_{x,y}q_x + \sum_{z\in X_N}\mu_z q_z \tag{2}$$

subject to

$$\begin{cases} \sum\limits_{x\in X_C} s_{x,y}q_x + \sum\limits_{z\in X_N} p_{z,y}q_z = q_y, \quad \forall y\in X; \\ \sum\limits_{x\in X_C} q_x + \sum\limits_{z\in X_N} q_z = 1; \\ \sum\limits_{y\in X(x)} s_{x,y} = 1, \quad \forall x\in X_C; \\ \qquad s_{x,y}\in\{0,1\}, \quad \forall x\in X_C,\ y\in X; \quad q_x\geq 0, \quad \forall x\in X, \end{cases} \tag{3}$$

where

$$\mu_z = \sum_{y\in X(z)} p_{z,y}c_{z,y}, \quad \forall z\in X_N.$$

Proof. Denote $\mu_x = \sum_{y\in X(x)} c_{x,y}s_{x,y}$ for $x\in X_C$. Then μ_x for $x\in X_C$ and μ_z for $z\in X_N$ represent, respectively, the immediate cost of the system in the states $x\in X_C$ and $z\in X_N$ when the strategy $s\in S$ is applied. Indeed, we can consider the values $s_{x,y}$ for $x\in X_C$ and $y\in X(x)$ as probability transitions from the state $x\in X_C$ to the state $y\in X(x)$.

Therefore, for fixed s the solution $q^s = (q^s_{x_{i_1}}, q^s_{x_{i_2}}, \dots, q^s_{x_{i_n}})$ of the system of linear equations

$$\begin{cases} \sum\limits_{x\in X_C} s_{x,y}q_x + \sum\limits_{z\in X_N} p_{z,y}q_z = q_y, \quad \forall y\in X; \\ \sum\limits_{x\in X_C} q_x + \sum\limits_{z\in X_N} q_z = 1; \end{cases} \tag{4}$$

corresponds to the vector of limit probabilities in the ergodic Markov chain determined by the graph $G_s = (X, E_s\cup E_N)$ with the probabilities $p_{x,y}$ for $(x,y)\in E_N$ and $p_{x,y} = s_{x,y}$ for $(x,y)\in E_C$ $(E_C = E\setminus E_N)$. Therefore, for given s the value

$$\psi(s,q^s) = \sum_{x\in X_C}\mu_x q_x + \sum_{z\in X_N}\mu_z q_z,$$

expresses the average cost per transition for the dynamical system if the strategy s is applied, i. e.,

$$f_x(s) = \psi(s,q^s), \quad \forall x\in X.$$

So, if we solve the optimization problem (2), (3) on a perfect network then we find the optimal strategy s^*. □

In the following for an arbitrary vertex $y\in X$ we will denote by $X_C^-(y)$ the set of vertices from X_C which contain directed leaving edges $e=(x,y)\in E$ that end in y, i. e., $X_C^-(y) = \{x\in X_C \mid (x,y)\in E\}$; in an analogues way we define the set $X^-(y) = \{x\in X \mid (x,y)\in E\}$.

Based on the lemma above we can prove the following result.

Theorem 1. *Let* $\alpha^*_{x,y}$ $(x \in X_C,\ y \in X)$, q^*_x $(x \in X)$ *be a basic optimal solution of the following linear programming problem:*
Minimize

$$\overline{\psi}(\alpha, q) = \sum_{x \in X_C} \sum_{y \in X(x)} c_{x,y}\alpha_{x,y} + \sum_{z \in X_N} \mu_z q_z \tag{5}$$

subject to

$$\begin{cases} \sum\limits_{x \in X_C^-(y)} \alpha_{x,y} + \sum\limits_{z \in X_N} p_{z,y} q_z = q_y, \quad \forall y \in X; \\ \sum\limits_{x \in X_C} q_x + \sum\limits_{z \in X_N} q_z = 1; \\ \sum\limits_{y \in X(x)} \alpha_{x,y} = q_x, \quad \forall x \in X_C; \\ \alpha_{x,y} \geq 0, \quad \forall x \in X_C,\ y \in X; \quad q_x \geq 0, \quad \forall x \in X. \end{cases} \tag{6}$$

Then the optimal stationary strategy s^* *on a perfect network can be found as follows:*

$$s^*_{x,y} = \begin{cases} 1 & if \quad \alpha^*_{x,y} > 0; \\ 0 & if \quad \alpha^*_{x,y} = 0, \end{cases}$$

where $x \in X_C$, $y \in X(x)$. *Moreover, for every starting state* $x \in X$ *the optimal average cost per transition is equal to* $\overline{\psi}(\alpha^*, q^*)$, *i. e.,*

$$f_x(s^*) = \sum_{x \in X_C} \sum_{y \in X(x)} c_{x,y}\alpha^*_{x,y} + \sum_{z \in X_N} \mu_z q^*_z$$

for every $x \in X$.

Proof. To prove the theorem it is sufficient to apply Lemma 1 and to show that the bilinear programming problem (2), (3) with boolean variables $s_{x,y}$ for $x \in X_C$, $y \in X$ can be reduced to the linear programming problem (5), (6). Indeed, we observe that the restrictions $s_{x,y} \in \{0, 1\}$ in the problem (2), (3) can be replaced by $s_{x,y} \geq 0$ because the optimal solutions after such a transformation of the problem are not changed. In addition, the restrictions

$$\sum_{y \in X(x)} s_{x,y} = 1, \quad \forall x \in X_C$$

can be changed by the restrictions

$$\sum_{y \in X(x)} s_{x,y} q_x = q_x, \quad \forall x \in X_C,$$

because for the perfect network it holds $q_x > 0, \ \forall x \in X_C$.

Based on the properties mentioned above in the problem (2), (3) we may replace the system (3) by the following system

$$\begin{cases} \sum\limits_{x \in X_C^-(y)} s_{x,y} q_x + \sum\limits_{z \in X_N} p_{z,y} q_z = q_y, \quad \forall y \in X; \\ \sum\limits_{x \in X_C} q_x + \sum\limits_{z \in X_N} q_z = 1; \\ \sum\limits_{y \in X(x)} s_{x,y} q_x = q_x, \quad \forall x \in X_C; \\ \qquad s_{x,y} \geq 0, \ \ \forall x \in X_C, \ y \in X; \quad q_x \geq 0, \ \ \forall x \in X. \end{cases} \tag{7}$$

Thus, we may conclude that problem (2), (3) and problem (2), (7) have the same optimal solutions. Taking into account that for the perfect network $q_x > 0, \ \ \forall x \in X$ we can introduce in problem (2), (7) the notations $\alpha_{x,y} = s_{x,y} q_x$ for $x \in X_C$, $y \in X(x)$. This leads to the problem (5), (6). It is evident that $\alpha_{x,y} \neq 0$ if and only if $s_{x,y} = 1$. Therefore, the optimal stationary strategy s^* can be found according to the rule given in the theorem. □

So, if the network (G, X_C, X_N, c, p, x_0) is perfect then we can find the optimal stationary strategy s^* by using the following algorithm.

Algorithm 1. Determining the Optimal Stationary Strategy on Perfect Networks

1) Formulate the linear programming problem (5), (6) and find a basic optimal solution $\alpha^*_{x,y}$ $(x \in X_C,\ y \in X)$, q^*_x $(x \in X)$.

2) Fix a stationary strategy s^* where $s^*_{x,y} = 1$ for $x \in X_C$, $y \in X(x)$ if $\alpha^*_{x,y} > 0$; otherwise put $s^*_{x,y} = 0$.

3 Extension of the Algorithm 1 for Solving the Unichain Control Problem

We show that the algorithm 1 can be extended for the problem in which an arbitrary strategy s generates a Markov unichain. For a unichain control problem the graph G_s induced by a stationary strategy may not be strongly connected, but it contains a unique deadlock strongly connected component that is reachable from every $x \in X$. A basic optimal solution α^*, q^* of the linear programming problem (5), (6) determines the strategy

$$s^*_{x,y} = \begin{cases} 1 & \text{if} \ \ \alpha^*_{x,y} > 0; \\ 0 & \text{if} \ \ \alpha^*_{x,y} = 0, \end{cases}$$

and a subset $X^* = \{x \in X \,|\, q^*_x > 0\}$, where s^* provides the optimal average cost per transition for the dynamical system $\mathbb{L}$ when it starts transitions in the states

$x_0 \in X^*$. This means that for an arbitrary network algorithm 1 determines the optimal stationary strategy of the problem only in the case if the system starts transitions in the states $x \in X^*$.

For a unichain control problem algorithm 1 determines the strategy s^* and the recurrent class X^*. The remaining states $x \in X \setminus X^*$ in X correspond to transient states and the optimal stationary strategies in these states can be chosen in order to reach X^*.

We show how to use the linear programming model (5), (6) for determining the optimal stationary strategies of the control problem on the nonperfect network in which for an arbitrary stationary strategy s the matrix P^s corresponds to a recurrent Markov chain.

An arbitrary strategy s in G generates a graph $G_s = (X, E_s \cup E_N)$ with unique deadlock strongly connected components $G'_s = (X'_s, E'_s)$ that can be reached from any vertex $x \in X$. The optimal stationary strategy s^* in G can be found from a basic optimal solution by fixing $s^*_{x,y} = 1$ for the basic variables. This means that in G we can find the optimal stationary strategy as follows:

We solve the linear programming problem (5), (6) and find a basic optimal solution α^*, q^*. Then we find the subset of vertices $X^* = \{x \in X \mid q^*_x > 0\}$ which in G corresponds to a strongly connected subgraph $G^* = (X^*, E^*)$. On this subgraph we determine the optimal solution of the problem using the algorithm 1. It is evident that if $x_0 \in X^*$ then we obtain the solution of the problem with fixed starting state x_0. To determine the solution of the problem for an arbitrary starting state we may select successively vertices $x \in X \setminus X^*$ which contain outgoing directed edges that end in X^* and will add them at each time to X^* using the following rule:

- if $x \in X_C \cap (X \setminus X^*)$ then we fix an directed edge $e = (x, y)$, put $s^*_{x,y} = 1$ and change X^* by $X^* \cup \{x\}$;
- if $x \in X_N \cap (X \setminus X^*)$ then change X^* by $X^* \cup \{x\}$.

4 An Approach for Solving the Multichain Control Problem Using a Reduction Procedure to a Unichain Problem

We consider the multichain control problem on the network (G, X_C, X_N, c, p, x_0), i. e., the case that for different starting states the average cost per transition may be different. We describe an approach for determining the optimal solution which is based on a reduction procedure of the multichain problem to the unichain case.

The graph G satisfies the condition that for an arbitrary vertex $x \in X_C$ each outgoing directed edge $e = (x, y)$ ends in X_N, i. e., we assume that

$$E_C = \{e = (x, y) \in E \mid x \in X_C,\ y \in X_N\}.$$

If the graph G does not satisfy this condition then the considered control problem can be reduced to a similar control problem on an auxiliary network $(G', X'_C, X'_N, c', p', x_0)$, where the graph G' satisfies the condition mentioned above. Graph $G' = (X', E')$ is obtained from $G = (X, E)$, where each directed edge $e = (x, y) \in E_C$ is changed by the following two directed edges $e^1 = (x, x_e)$ and $e^2 = (x_e, y)$.

We include each vertex x_e in X'_N and with each edge $e' = (x_e, y)$ we associate the cost $c'_{x_e,y} = 0$ and the transition probability $p'_{x_e,y} = 1$. With the edges $e' = (x, x_e)$ we associate the cost $c'_{x,x_e} = c_{(x,y)}$, where $e = (x, y)$. For the edges $e \in E_N$ in the new network we preserve the same costs and transition probabilities as in the initial network, i. e., the cost function c' on E_N and on the set of edges (x, x_e) for $x \in X_C,\ e \in E_C$ is induced by the cost function c. Thus, in the auxiliary network the graph G' is determined by the set of vertices $X' = X'_C \cup X'_N$ and the set of edges $E' = E'_C \cup E'_N$, where $X'_C = X_C$; $X'_N = X_N \cup \{x_e,\ e \in E_C\}$; $E'_C = \{e' = (x, x_e) \mid x \in X_C,\ e = (x, y) \in E_C\}$; $E'_N = E_N \cup \{e' = (x_e, y) \mid\ e = (x, y) \in E_C, y \in X\}$. There exists a bijective mapping between the set of strategies in the states $x \in X_C$ of the network (G, X_C, X_N, c, p, x_0) and the set of strategies in the states $x \in X_C$ of the network $(G', X'_C, X'_N, c', p', x_0)$ that preserves the average costs of the problems on the corresponding networks.

Thus, without loss of generality we may consider that G possesses the property that for an arbitrary vertex $x \in X_C$, each outgoing directed edge $e = (x, y)$ ends in X_N. Additionally, let us assume that the vertex x_0 in G is reachable from every vertex $x \in X_N$. Then an arbitrary strategy s in the considered problem induces a transition probability matrix $P^s = (p^s_{x,y})$ that corresponds to a Markov unichain with a positive recurrent class X^+ that contains the vertex x_0.

Therefore, if we solve the control problem on the network then we obtain the solution of the problem with fixed starting state x_0. So, we obtain such a solution if the network satisfies the condition that for an arbitrary strategy s the vertex x_0 in G_s is attainable for every $x \in X_N$. Now let us assume that this property does not take place. In this case we can reduce our problem to a similar problem on a new auxiliary network $(G'', X''_C, X''_N, p'', c'', x_0)$ for which the property mentioned above holds. This network is obtained from the initial one by the following way: we construct the graph $G'' = (X, E'')$ which is obtained from $G = (X, E)$ by adding new directed edges $e''_{x_0} = (x, x_0)$ from $x \in X_N \setminus \{x_0\}$ to x_0, if for some vertices $x \in X_N \setminus \{x_0\}$ in G there are no directed edges $e = (x, x_0)$ from x to x_0. We define the costs of directed edges $(x, y) \in E''$ in G'' as follows: if $e'' = (x, y) \in E$ then the cost $c''_{e''}$ of this edge in G'' is the same as in G, i. e., $c''_{e''} = c_{e''}$ for $e'' \in E$; if $e'' = (x, x_0) \in E'' \setminus E$ then we put $c''_{e''} = 0$. The probabilities $p''_{x,y}$ for $(x, y) \in E''$ where $x \in X_N$ we define by using the following rule: we fix a small positive value ε and put $p''_{x,y} = p_{x,y} - \varepsilon p_{x,y}$ if $(x, y) \in E'' \setminus E,\quad y \neq x_0$ and in G there is no directed edge $e = (x, x_0)$ from x to x_0; if in G for a vertex $x \in X \setminus \{x_0\}$ there exists a leaving directed edge $e = (x, x_0)$ then for an arbitrary outgoing directed edge $e = (x, y),\ y \in X(x)$ we put $p''_{x,y} = p_{x,y}$; for the directed edges $(x, x_0) \in E' \setminus E$ we put $p''_{x,x_0} = \varepsilon$.

Let us assume that the probabilities $p_{x,y}$ for $(x, y) \in E$ are given in the form of irreducible decimal fractions $p_{x,y} = a_{x,y}/b_{x,y}$.

Additionally, assume that the values ε satisfy the condition

$$\varepsilon \leq 2^{-2L-2},$$

where

$$L = \sum_{(x,y)\in E} \log(a_{x,y}+1) + \sum_{(x,y)\in E} \log(b_{x,y}+1) + \sum_{e\in E}(|c_e|+1) + 2\log(n) + 1.$$

Then, based on the results from [4] for our auxiliary optimization problem (with approximated data) we can conclude that the solution of this problem will correspond to the solution of our initial problem.

So, to find the optimal solution of the problem on the network (G, X_C, X_N, c, p, x_0) it is necessary to construct the auxiliary network $(G', X'_C, X'_N, c', p', x_0)$, where for each vertex $x \in X'_N$ an arbitrary directed edge $e' = (x, y)$ ends in X_N. Then we construct the network $(G'', X''_C, X''_N, c'', p'', x_0)$ and the auxiliary stochastic optimal control problem on this network. If the optimal stationary strategy s'^* in the auxiliary problem is found, then we fix $s^* = s'^*$ on X_C.

References

[1] Bellman R. *Functional equations in the theory of dynamic programming, XI-Limit theorems.* Rand. Circolo Math. Palermo, 1959, **8(3)**, 343–345.

[2] Lozovanu D., Pickl S. *Algorithmic solution of discrete control problems on stochastic networks.* Procceedings of CTW09 Workshop on Graphs and Combinatorial Optimization, Paris, 2009, 221–224.

[3] Puterman M. *Markov Decision Processes: Stochastic Dynamic Programming.* John Wiley, New Jersey, 2005.

[4] Khachian L. G. *Polynomial time algorithm in linear programming.* USSR Computational Mathematics and Mathematical Phisics, 1980, **20**, 51–58.

Dmitrii Lozovanu
Institute of Mathematics and Computer Science
Academy of Sciences of Moldova
5, Academiei str., Chisinau, MD-2028
Moldova
E-mail: *lozovanu@math.md*

Received September 10, 2014

Maria Capcelea
Moldova State University
60, Mateevici str., Chisinau, MD-2009
Moldova
E-mail: *mariacapcelea@yahoo.com*