AN ALGORITHM FOR DETERMINING OPTIMAL STATIONARY STRATEGIES FOR UNICHAIN CONTROL PROBLEM

MARIA CAPCELEA, TITU CAPCELEA
Moldova State University,
Faculty of Mathematics and Computer Science,
Mateevici str., 60, Chisinau, MD-2009, Moldova

Abstract
A linear programming approach for determining optimal stationary strategies for the stochastic control problem on unichain network with average cost criterion and infinite time horizon is proposed.

Key words: Stochastic Discrete Control Problem, Optimal Stationary Strategies, Linear Programming Approach, Polynomial Time Algorithm

We consider the stationary stochastic discrete optimal control problem on decision network $R = (G, X_C, X_N, c, p)$ [1]. This network is determined by the directed graph $G = (X, \mathrm{E})$, the disjoint subsets $X_C, X_N \subset X$, the cost function $c : E \to E_N$ and the probability distribution function $p : E_N \to [0,1]$ on the set $E_N = \{(x, y) \in E \mid x \in X_N\}$. X_C represents the set of controllable states in which the transitions of the system to the next state can be controlled by the decision maker at every discrete moment of time and X_N represents the set of states in which the decision maker is not able to control the transitions because the system passes to the next state randomly.

Let $s : X_C \to X$ define a stationary strategy for the control problem on decision network R. We can determine the graph $G_s = (X, E_s \bigcup E_N)$, $E_s = \{(x, y) \in E \mid x \in X_C, y = s(x)\}$, that

corresponds to a Markov process with the probability matrix $P^S = \left(p_{x,y}^s \right)$, where

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

In the considered Markov process for an arbitrary state $x \in X_C$ the transition $\left(x, s(x) \right)$ from the states $x \in X_C$ to the states $y = s(x) \in X$ is made with 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 $x_i \in X$. For this, we can determine the vector of average costs w^s which corresponds to the strategy s, according to the formula $w^s = Q^s \mu^s$, 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. $\mu_x^s = \sum_{y \in X(x)} p_{x,y}^s c_{x,y}^s$. A component w_x^s of the vector w^s represents the average cost per transition in our problem with a given starting state x and a fixed strategy s, i.e. $f_x(s) = w_x^s$. 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 R with starting state x_0 consists of finding a stationary strategy s^* for which $f_{x_0}\left(s^* \right) = \min_s f_{x_0}(s)$.

We assume that R is a perfect network which means that in graph G for an arbitrary stationary strategy s the subgraph $G_s = \left(X, E_s \cup E_N \right)$ is strongly connected. So, the Markov chain induced by the probability transition matrix P^s is irreducible for an arbitrary strategy s.

At first we formulate an algorithm for determining the optimal stationary strategies for the control problem on perfect networks. Then we show that the proposed algorithm can be extended for the problem if 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 strongly connected component that is reachable from every state $x \in X$.

Let s be an arbitrary strategy. Taking into account that for every fixed $x \in X_C$ we have a unique $y = s(x) \in X(x)$, then 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 $s_{x,y} = \begin{cases} 1 & if \ y = s(x) \\ 0 & if \ y \neq s(x) \end{cases}$.

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:*

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

subject to

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

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

$$s_{x,y}^* = \begin{cases} 1 & if \ \alpha_{x,y}^* > 0 \\ 0 & if \ \alpha_{x,y}^* = 0 \end{cases},$$

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 $\bar{\psi}(\alpha^*, q^*)$ for every $x \in X$.*

For a unichain control problem the algorithm based on Theorem 1 determines the strategy s^* and the recurrent class X^*. In this case the remaining states $x \in X / X^*$ in X correspond to transient states and the optimal stationary strategies in the states $x \in X / X^*$ can be chosen in order to reach X^*. Therefore, the linear programming model from Theorem 1 can be used for determining the optimal stationary strategy for an arbitrary unichain control problem.

We solve the linear programming problem and find a basic optimal solution α^*, q^*. Then we find the subset of vertices $X^* = \left\{ x \in X \,\middle|\, q_x^* > 0 \right\}$ 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 based on Theorem 1. 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 $x \in X / X^*$ 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 $X^* \cup \{x\}$.

References

1. Lozovanu D., Pickl S. Optimization of stochastic discrete systems and control on complex networks. Computational networks. Springer, Advances in Computational Management Science, v.12, 2015.