Optimal Stationary Strategies for Stochastic Control Problems on Networks with Discounted Costs

Dmitrii Lozovanu[1], Stefan Pickl[2]

[1] Institute of Mathematics and Computer Science of Academy of Sciences of Moldova; e-mail:lozovanu@math.md

[2] Institute for Theoretical Computer Science, Mathematics and Operations Research Universität der Bundeswehr, Germany; e-mail:stefan.pickl@unibw.de

Abstract

We consider the infinite horizon discrete control problem on stochastic networks with discounted costs and propose an approach for determining the solutions of the problem based on linear programming. We formulate a linear programming model for the considered problem and propose polynomial time algorithms for determining the optimal stationary strategies of the problem on stochastic networks with discounted costs. Based on duality theory of linear programming we show also how to determine the solution of the problem by using efficient iterative procedures.

Keywords : *Control problem on networks, Optimal stationary strategies, Markov processes with discounted costs, Linear programming*

1 Introduction and problem formulation

In [4, 5] the following infinite horizon stochastic control problem on networks has been considered. Let a discrete time system $\mathbb{L}$ with a finite set of states X be given. 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 vertex set X corresponds to the set of states of $\mathbb{L}$ and a directed edge $e = (x, y) \in E$ expresses the possibility of the dynamical system to pass from the state $x = x(t) \in X$ to the state $y = x(t+1) \in X$ at every discrete moment of time $t = 0, 1, 2, \dots$. To each directed edge $e = (x, y) \in E$ a cost c_e that expresses the cost of the system $\mathbb{L}$ to pass from the state x to the state y is associated. Each such cost at every next discrete moment of time is rated with a given discount factor $\gamma, 0 < \gamma < 1$. So, if system $\mathbb{L}$ at the moment of time t passes from a state $x = x(t) \in X$ to a state $y = x(t+1) \in X$ through the directed edge $e = (x, y) \in E$ then the cost of states' transition of the system is $\gamma^t c_e$. We consider the control problem when the set of states X consists of two disjoint subsets X_C and X_N ($X = X_C \cup X_N,\ X_C \cap X_N = \emptyset$) where X_C represents the subset of controllable states for $\mathbb{L}$ and X_N is the subset of uncontrollable states for $\mathbb{L}$. This means that the decision maker in this problem may control the system only in the states $x \in X_C$ and can make a states' transition of $\mathbb{L}$ from a state $x = x(t)$ to the state $y = x(t+1)$ through the directed edge $e = (x, y) \in E$ at every discrete moment of time; if $x = x(t) \in X_N$ then the decision maker is unable to control the system $\mathbb{L}$ in $x = x(t)$ because the system passes to the next state $y = x(t+1) \in X$ randomly according to a given distribution $\{p_e\}$ on $E(x) = \{e = (x, y) : (x, y) \in E\}$, where $\sum_{e \in E(x)} p_e = 1, p_e \geq 0, \forall e \in E(x)$. So, if the starting state $x_0 = x(0)$ is given then the control process of $\mathbb{L}$ on G is the following: if $x(0) \in X_C$ then the decision maker chooses a transition of $\mathbb{L}$ from x_0 to a state x_1 such that $e_0 = (x_0, x_1) \in E$ where $x_1 = x(1)$. If $x_0 \in X_N$ then x_1 is chosen randomly according to the distribution $\{p_{x,y}\}$ on $E(x_0)$. At the moment of time $t = 1$ if $x_1 \in X_C$ then the decision maker chooses a transition for $\mathbb{L}$ from x_1 to a state x_2 such that $e_1 = (x_1, x_2)$ where $x_2 = x(2)$, otherwise (in the case $x_1 \in X_N$)

the state x_1 is chosen randomly according to the distribution $\{p_{x,y}\}$ on $E(x_0)$ and so on indefinitely. So, a stationary strategy (or a stationary control) for the system $\mathbb{L}$ on G can be defined as a map $s : x \mapsto y \in X(x)$ for $x \in X_C$, where $X(x) = \{y \in X|\ (x,y) \in E\}$. For a given starting state x_0 and a given stationary strategy s the total expected discounted cost $\sigma_{x_0}(s)$ for the dynamical system $\mathbb{L}$ is defined as follows. The strategy s induces the graph $G_s = (X, E_s \cup E_N)$, where $E_s = \{e = (x,y) \in E | x \in X_C, y = s(x)\}$, $E_N = \{e = (x,y) | x \in X_N, , y \in X\}$. This graph corresponds to a Markov process with a transition probability matrix $P^s = (p^s_{x,y})$, where

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

For this Markov process with associated costs $c_e, e \in E_s \cup E_N$ and given starting state x_0 we can define the total expected discounted cost $\sigma_{x_0}(s)$ (see [2]). So, we can consider the problem of determining the strategy s^* for which

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

The strategy s^* corresponds to an optimal stationary strategy for the stochastic control problem on G with discounted costs. In [4, 5] it is shown that the optimal stationary strategy s^* for the control problem with fixed starting state x_0 can be found by using a linear programming approach.

In this contribution we specify the linear programming model from [4, 5] for the case of the control problem with an arbitrary starting state. We show that the dual linear programming model for the control problem with an arbitrary starting state allows us to determine all optimal stationary strategies of the problem. Additionally, based on duality theory of linear programming and the results from [4, 5] we propose an efficient iterative algorithm for determining the optimal stationary strategies of the problem.

2 The main results

In this contribution we present the following result for the stochastic control problem with an arbitrary starting state $x \in X$.

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

$$\phi(\alpha, \beta) = \sum_{x \in X_C} \sum_{y \in X(x)} c_{x,y}\, \alpha_{x,y} + \sum_{x \in X_N} \mu_x \beta_x \tag{1}$$

subject to

$$\begin{cases} \sum_{x \in X(y)} \alpha_{y,x} - \gamma \sum_{x \in X_C^-(y)} \alpha_{x,y} - \gamma \sum_{x \in X_N^-(y)} p_{x,y}\, \beta_x = 1, \quad y \in X_C; \\ \beta_y - \gamma \sum_{x \in X_C^-(y)} \alpha_{x,y} - \gamma \sum_{x \in X_N^-(y)} p_{x,y}\, \beta_x = 1, \quad y \in X_N; \\ \beta_x \geq 0, \ \forall x \in X_N; \quad \alpha_{x,y} \geq 0, \ \forall x \in X_C, \, y \in X(x), \end{cases} \tag{2}$$

where

$$X_C^-(y) = \{x \in X_C |\ (x,y) \in E\}, \quad X_N^-(y) = \{x \in X_N |\ (x,y) \in E\} \quad \textit{for} \quad y \in X$$

and

$$\mu_x = \sum_{y \in X(x)} c_{x,y} p_{x,y} \ \textit{for} \ x \in X_N.$$

If in the graph $G = (X, E)$ each vertex $x \in X$ contains at least one leaving directed edge then $\sum_{y \in X(x)} \alpha^*_{x,y} > 0,\ \forall x \in X_C$ and

$$\frac{\alpha^*_{x,y}}{\sum_{y \in X} \alpha^*_{x,y}} \in \{0, 1\}, \quad \forall x \in X_C,\ y \in X(x).$$

The optimal stationary strategy s^* of the discounted stochastic control problem on the network with an arbitrary starting state $x \in X$ can be found by setting

$$s^*_{x,y} = \frac{\alpha^*_{x,y}}{\sum_{y \in X(x)} \alpha_{x,y}}, \quad \forall x \in X_C,\ y \in X(x).$$

The proof of this theorem is similar to the proof of Theorem 1 from [4] for the case of the problem with a fixed starting state x_0.

Using duality theory for the linear programming problem (1),(2) we obtain the following result.

Theorem 2 *Let σ^*_x $(x \in X)$ be the optimal solution of the linear programming problem: Maximize*

$$\varphi(\sigma, w) = \sum_{x \in X} \sigma_x \tag{3}$$

subject to

$$\begin{cases} \sigma_x - \gamma \sigma_y \leq c_{x,y}, & \forall x \in X_C,\ y \in X(x); \\ \sigma_x - \gamma \sum_{y \in X(x)} p_{x,y} \sigma_y \leq \mu_x, & \forall x \in X_N. \end{cases} \tag{4}$$

Then σ^*_x for $x \in X$ represents the discounted expected total costs for the problem with the corresponding starting states $x \in X$. An arbitrary optimal stationary strategy can be found by fixing

$$s^* : X_C \mapsto X$$

such that

$$(x, s^*(x)) \in E^*(x),\ \forall x \in X_C$$

where

$$E^*(x) = \{(x, y) \,|\, y \in X(x),\ \sigma^*_x - \gamma \sigma^*_y - c_{x,y} = 0\}.$$

According to this theorem the linear programming problem (3), (4) determines all optimal discounted costs σ^*_x for $x \in X$ and all optimal stationary strategies $s^* : X_C \mapsto X$ determined by the set $E^*(x)$. If $X_N = \emptyset$ then we obtain the deterministic infinite horizon control problem on a network with discounted costs. In the case when $X_N = \emptyset,\ \gamma = 1$ and G is a directed graph with nonnegative costs and has a sink vertex x_s then we obtain the problem of determining the minimum cost paths from every $x \in X \setminus \{x_0\}$ to sink vertex x_s.

Using the duality theory of linear programming we can also propose the following iterative algorithm for determining an optimal stationary strategy for the control problem.

Preliminary step (Step 0): Fix an arbitrary stationary strategy

$$s^0 : x_i \to x_j \in X(x_i) \text{ for } x_i \in X_C.$$

General step (Step $k,\ k > 0$): Determine the probability matrix $P^{s^{k-1}} = (p^{s^{k-1}}_{x_i,x_j})$, where

$$p^{s^{k-1}}_{x_i,x_j} = \begin{cases} p_{x_i,x_j}\,, & \text{if } x_i \in X_N \text{ and } (x_i, x_j) \in E_N; \\ 1, & \text{if } x_i \in X_C \text{ and } x_j = s^{k-1}(x_i); \\ 0, & \text{if } x_i \in X_C \text{ and } x_j \neq s^{k-1}(x_i). \end{cases}$$

Then calculate $\mu_{x_i,s^{k-1}(x_i)} = \sum_{y\in X(x_i)} p_{x_i,y}^{s^{k-1}(x_i)} c_{x_i,y}^{s^{k-1}(x_i)}$ for every $x_i \in X$ and solve the system of linear equations

$$\sigma_{x_i} = \mu_{x_i,s^{k-1}(x_i)} + \gamma \sum_{x_j\in X} p_{x_i,x_j}^{s^{k-1}(x_i)} \sigma_{x_j}, \quad i = 1, 2, \ldots, n$$

and determine the solution $\sigma_{x_1}^{k-1}, \sigma_{x_2}^{k-1}, \ldots, \sigma_{x_n}^{k-1}$. After that calculate a new strategy

$$s^k : x_i \to a \in A(x_i) \ \text{ for } \ x_i \in X_C,$$

where

$$s^k(x_i) = \arg \min_{a\in A(x_i)} \Big[\mu_{x_i,a} + \gamma \sum_{x_j\in X} p_{x_i,x_j}^{a} \sigma_{x_i}^{k-1}\Big], \quad \forall x_i \in X_C.$$

Check if the following condition $s^k(x_i) = s^{k-1}(x_i), \quad \forall x_i \in X_C$ is valid. If this condition holds then fix $s^* = s^k; \quad \sigma_{x_i}^* = \sigma_{x_i}^k, \quad \forall x_i \in X$ as the optimal solution of the problem; otherwise go to the next step $k + 1$.

The iterative algorithm described above can be regarded as a value iteration algorithm for a discounted Markov decision problem with the transition probability matrix P^s defined above. In [5] it is shown that the infinite stochastic control problem on stochastic networks with discounted costs is equivalent to a discounted Markov decision problem with a finite set of states and finite set of actions. The proposed linear programming approach and the iterative algorithms can be developed in a similar way for the discounted Markov decision problem.

3 Conclusion and perspectives

The considered infinite horizon discrete control problem on stochastic networks with discounted costs generalizes the deterministic control problems on networks from [1, 3]. The optimal stationary strategies of this problem can be found on the bases of Theorems 1 and 2 by solving the linear programming problem (1), (2) or the linear programming problem (3), (4). Additionally an optimal stationary strategy for the discounted stochastic control problem on networks can be found by using the proposed iterative algorithm. Therefore the linear programming approach and the proposed iterative algorithm for solving the considered problem can be extended for the infinite horizon discounted Markov decision processes with finite state and action spaces.

References

[1] R. Bellman and R. Kalaba. *Dynamic programming and modern control theory.* Academic Press, New York and London, 1965.

[2] L. Kallenberg. *Markov decision processes.* Leiden University, Netherlands, 2016.

[3] D. Lozovanu and S. Pickl. *Optimization and multiobjective control of time-discrete systems.* Springer, 2009.

[4] D. Lozovanu and S. Pickl. Determining the optimal strategies for discrete control problems on stochastic networks with discounted costs. *Discrete Applied Mathematics*, 182(1):169–180, 2015.

[5] D. Lozovanu and S. Pickl. *Optimization of stochastic discrete systems and control on complex networks.* Springer, 2015.