Available online at www.sciencedirect.com

ScienceDirect

Electronic Notes in DISCRETE MATHEMATICS

Electronic Notes in Discrete Mathematics 55 (2016) 155–159 www.elsevier.com/locate/endm

Determining the Optimal Strategies for Zero-Sum Average Stochastic Positional Games

Dmitrii Lozovanu

Applied Mathematics
Institute of Mathematics and Computer Science of Moldova Academy of Sciences
Chisinau, Moldova

Stefan Pickl

Institute for Theoretical Computer Science, Mathematics and Operations Research
Universität der Bundeswehr, München,
Neubiberg-München, Germany

Abstract

We consider a class of zero-sum stochastic positional games that generalizes the deterministic antagonistic positional games with average payoffs. We prove the existence of saddle points for the considered class of games and propose an approach for determining the optimal stationary strategies of the players.

Keywords: Average Markov decision process, Zero-sum stochastic positional games, Optimal strategies, Saddle points.

[1] Email: `dmitrii.lozovanu@math.md`
[2] Email: `stefan.pickl@unibw.de`

http://dx.doi.org/10.1016/j.endm.2016.10.039
1571-0653/

1 Introduction and Problem Formulation

The aim of this paper is to prove the existence of saddle points for a class of antagonistic stochastic games that extends deterministic positional games with average payoffs from [1,2,5]. The considered class of games we formulate by using the framework of a Markov decision process (X, A, c, p, x_0) with a finite set of states X, a finite set of actions $A = \cup_{x \in X} A(x)$ where $A(x)$ is the set of actions in the state $x \in X$, a probability transition function $p : X \times A \times X \to [0, 1]$ that satisfies the condition $\sum_{y \in X} p^a_{x,y} = 1,\ \forall x \in X, a \in A(x)$, and the cost function $c : X \times X \to R$. We assume that the Markov process describes the evolution of a dynamic system that is controlled by two players as follows: The set of states is divided into two subsets $X = X_1 \cup X_2$ with $X_1 \cap X_2 = \emptyset$, where X_1 represents the position set of the first player and X_2 represents the position set of the second player. At time moment $t = 0$ the dynamical system is in the state x_0. If this state belongs to the set of positions of the first player then the action $a_0 \in A(x_0)$ in this state is selected by the first player, otherwise the action $a_0 \in A(x_0)$ is chosen by the second player. After that the dynamical system passes randomly to an another state according to the probability distribution $\{p^{a_0}_{x_0,y}\}$. At time moment $t = 1$ the players observe the state $x_1 \in X$ of the system. If x_1 belongs to the set of positions of the first player then the action $a \in A(x_1)$ is chosen by the first player, otherwise the action is chosen by the second one and so on, indefinitely. In this process the first player intends to maximize $\lim\limits_{t\to\infty} \inf \frac{1}{t} \sum\limits_{\tau=1}^{t} \mu(x_{\tau-1})$, where $\mu(x_\tau) = \sum\limits_{y \in X} p^{a_\tau}_{x_\tau,y} c_{x_\tau,y}$, and the second player intends to minimize $\lim\limits_{t\to\infty} \sup \frac{1}{t} \sum\limits_{\tau=1}^{t} \mu(x_{\tau-1})$. We assume that the players use stationary strategies of a selection of the actions in their position sets. We define the stationary strategies of the players as maps: $s^1 : x \to a \in A(x)$ for $x \in X_1$; $\quad s^2 : x \to a \in A(x)$ for $x \in X_2$. Let s^1, s^2 be arbitrary strategies of the players.Then $s = (s^1, s^2)$ determines a Markov process induced by probability distributions $\{p^{s^i(x)}_{x,y}\}$ in the states $x \in X_i$, $i = 1, 2$ and a given starting state x_0. For this Markov process with transition costs $c_{x,y},\ x, y \in X$ we can determine the average cost per transition $F_{x_o}(s^1, s^2)$. The function $F_{x_o}(s^1, s^2)$ on $S = S^1 \times S^2$, where S^1 and S_2 represent the corresponding sets of the stationary strategies of player 1 and player 2, defines an antagonistic game. This game is determined by position sets X_1, X_2, the set of actions A, the probability function p, the cost function c and the starting state x_0. We denote this game by (X_1, X_2, A, c, p, x_0) and call it zero-sum average stochastic positional game. In this game, we are seeking for the saddle points.