Contributions to Game Theory and Management, XII, 246–260

Pure Stationary Nash Equilibria for Discounted Stochastic Positional Games

Dmitrii Lozovanu[1] **and Stefan Pickl**[2]

[1] *Institute of Mathematics and Computer Science of Moldova Academy of Sciences, Academiei 5, Chisinau, MD-2028, Moldova,*
E-mail: `lozovanu@math.md`

[2] *Institute for Theoretical Computer Science, Mathematics and Operations Research, Universität der Bundeswehr München, 85577 Neubiberg-München, Germany,*
E-mail: `stefan.pickl@unibw.de`

Abstract A discounted stochastic positional game is a stochastic game with discounted payoffs in which the set of states is divided into several disjoint subsets such that each subset represents the position set for one of the player and each player control the Markov decision process only in his position set. In such a game each player chooses actions in his position set in order to maximize the expected discounted sum of his stage rewards. We show that an arbitrary discounted stochastic positional game with finite state and action spaces possesses a Nash equilibrium in pure stationary strategies. Based on the proof of this result we present conditions for determining all optimal pure stationary strategies of the players.

Keywords: stochastic positional games, discounted payoffs, pure stationary strategies, mixed stationary strategies, Nash equilibria

1. Introduction

Stochastic games where introduced by Shapley, 1953. He considered two-person zero-sum stochastic games with finite state and action spaces for which he proved the existence of the value and the optimal stationary strategies of the players with respect to a discounted payoff criterion. Later this result has been extended to m-person stochastic games and the existence of Nash equilibria in stationary strategies have been obtained for a more general class of discounted stochastic games (see Fink, 1964; Takahashi, 1964; Sobol, 1971; Solan, 1998). Shapley defined a stationary strategy for a player as a map that provides in each state of the game a probability distribution over the set of feasible actions. Therefore a stationary strategy for a player in a stochastic game can be treated as a mixed stationary strategy. So, the existence of Nash equilibria results mentioned above are related to Nash equilibria in mixed stationary strategies for the considered games.

In this paper we study the problem of the existence of Nash equilibria in pure stationary strategies for a special class of m-player discounted stochastic games that we call *discounted stochastic positional games*. This class of games has been introduced by Lozovanu and Pickl, 2015. An m-player stochastic positional game with discounted payoffs is an m-player stochastic game where the set of states is divided into into m disjoint subsets such that each subset represents the position set for one of the players and each player controls the Markov decision process only in his position set. In such a game each player chooses actions in his position set in order to maximize the expected discounted sum of his stage rewards. We show that for an arbitrary discounted stochastic positional game with finite state and action spaces there exists a Nash equilibrium in pure stationary strategies. Based on the proof of this result we present conditions for determining all pure stationary Nash equilibria.

The paper is organized as follows. In Section 2 the general formulation of a discounted stochastic positional game is presented. Then in Sections 3 the formulation of a discounted stochastic positional game is specified when the players use pure and mixed stationary strategies of choosing the actions in position sets. In Section 4 some new basic properties of the solutions for a discounted Markov decision problem in terms of stationary strategies are presented. Additionally, it is shown that such a problem can be represented as a quasi-monotonic programming problem. Based on these results in Section 5 it is shown that a discounted stochastic game can be formulated in terms of stationary strategies, where the payoff of each player is quasi-monotonic with respect to his strategy. Using these properies a new proof of the existence of stationary Nash equilibrium for a discounted stochastic game is derived and new conditions for determining the optimal strategies of the players are obtained. In Section 6 it is shown that a stochastic positional game with discounted payoff represents a particulary case of a discounted stochastic game and the corresponding conditions for determining the stationary strategies of the players are specified. In Section 6 the proof of the existence of pure stationary Nash equilibria for an arbitrary discounted stochastic positional game is presented.

2. Formulation of the Discounted Stochastic Positional Game in the Term of Stationary Strategies

First we present the general model for a discounted stochastic positional game and then we specify the formulation of the game when the players use pure and mixed stationary strategies of choosing the actions in their state positions.

2.1. The General Model of a Discounted Stochastic Positional Game

A discounted stochastic positional game with m players consists of the following elements:

- a state space X (which we assume to be finite);
- a partition $X = X_1 \cup X_2 \cup \cdots \cup X_m$ where X_i represents the position set of player $i \in \{1, 2, \ldots, m\}$;
- a finite set $A(x)$ of actions in each state $x \in X$;
- a step reward $f^i(x, a)$ with respect to each player $i \in \{1, 2, \ldots, m\}$ in each state $x \in X$ and for an arbitrary action $a \in A(x)$;
- a transition probability function $p : X \times \prod_{x \in X} A(x) \times X \rightarrow [0, 1]$ that gives the probability transitions $p^a_{x,y}$ from an arbitrary $x \in X$ to an arbitrary $y \in X$ for a fixed action $a \in A(x)$, where $\sum_{y \in X} p^a_{x,y} = 1, \ \ \forall x \in X, \ a \in A(x)$;
- a discount factor $\gamma, \ 0 < \gamma < 1$;
- a starting state $x_0 \in X$.

The game starts at the moment of time $t = 0$ in the state x_0 where the player $i \in \{1, 2, \ldots, m\}$ who is the owner of the state position x_0 ($x_0 \in X_i$) chooses an action $a_0 \in A(x_0)$ and determines the rewards $f^1(x_0, a_0), f^2(x_0, a_0), \ldots, f^m(x_0, a_0)$ for the corresponding players $1,2,\ldots,m$. After that the game passes to a state $y=x_1\in X$ according to a probability distribution $\{p^{a_0}_{x_0,y}\}$. At the moment of time $t=1$ the player $k\in\{1,2,\ldots,m\}$ who is the owner of the state position x_1 ($x_1\in X_k$) chooses an action $a_1\in A(x_1)$ and players $1,2,\ldots,m$ receive the corresponding rewards $f^1(x_1,a_1), f^2(x_1,a_1),\ldots,f^m(x_1,a_1)$. Then the game passes to a state $y=x_2\in X$ according to a probability distribution $\{p^{a_1}_{x_1,y}\}$ and so on indefinitely. Such a play of the game produces a sequence of states and actions $x_0,a_0,x_1,a_1,\ldots,x_t,a_t,\ldots$ that defines a stream of stage rewards $f^1(x_t,a_t), f^2(x_t,a_t),\ldots,f^m(x_t,a_t)$, $t=0,1,2,\ldots$. *The discounted stochastic positional game* is the game with payoffs of the players

$$\sigma^i_{x_0}=\mathsf{E}\left(\sum_{\tau=0}^{\infty}\gamma^\tau f^i(x_\tau,a_\tau)\right),\quad i=1,2,\ldots,m$$

where E is the expectation operator with respect to the probability measure in the Markov process induced by actions chosen by players in their position sets and given starting state x_0. Each player in this game has the aim to maximize the expected discounted sum of his stage rewards. In the case $m=1$ this game becomes the discounted Markov decision problem with given action sets $A(x)$ for $x\in X$, a transition probability function $p: X\times\prod_{x\in X}A(x)\times X\rightarrow[0,1]$, step rewards $f(x,a)=f^1(x,a)$ for $x\in X$, $a\in A(x)$, given discount factor λ and starting state x_0.

In the paper we will study the discounted stochastic positional game when the players use pure and mixed stationary strategies of choosing the actions in the states.

2.2. A Discounted Stochastic Positional Games in Pure and Mixed Stationary Strategies

A strategy of player $i\in\{1,2,\ldots,m\}$ in a stochastic positional game is a mapping s^i that provides for every state $x_t\in X_i$ a probability distribution over the set of actions $A(x_t)$. If these probabilities take only values 0 and 1, then s^i is called *a pure strategy*, otherwise s^i is called *a mixed strategy.* If these probabilities depend only on the state $x_t=x\in X_i$ (i. e. s^i does not depend on t), then s^i is called *a stationary strategy*, otherwise s^i is called a *non-stationary strategy.*

Thus, we can identify the set of mixed stationary strategies $\mathbf{S}^i$ of player i with the set of solutions of the system

$$\begin{cases}\sum\limits_{a\in A(x)} s^i_{x,a}=1, & \forall x\in X_i;\\ s^i_{x,a}\ge 0, & \forall x\in X_i,\ \ \forall a\in A(x).\end{cases}\tag{1}$$

Each basic solution s^i of this system corresponds to a pure stationary strategy of player $i\in\{1,2,\ldots,m\}$. So, the set of pure stationary strategies S^i of player i corresponds to the set of basic solutions of system (1).

Let $\mathbf{s}=(s^1,s^2,\ldots,s^m)\in\mathbf{S}=\mathbf{S}^1\times\mathbf{S}^2\times\cdots\times\mathbf{S}^m$ be a profile of stationary strategies (pure or mixed strategies) of the players. Then the elements of probability transition matrix $P^{\mathbf{s}}=(p^{\mathbf{s}}_{x,y})$ in the Markov process induced by $\mathbf{s}$ can be calculated as follows:

$$p^{\mathbf{s}}_{x,y} = \sum_{a\in A(x)} s^i_{x,a} p^a_{x,y} \quad \text{for} \quad x \in X_i, \quad i = 1, 2, \ldots, m. \tag{2}$$

Let us consider the matrix $W^{\mathbf{s}} = (w^s_{x,y})$ where $W^{\mathbf{s}} = (I - \gamma P^{\mathbf{s}})^{-1}$. Then in a discounted stochastic positional game the payoff of player $i \in \{1, 2, \ldots, m\}$ for a given profile $\mathbf{s}$ and initial state $x_0 \in X$ is determined as follows

$$\sigma^i_{x_0}(\mathbf{s}) = \sum_{k=1}^{m} \sum_{y\in X_k} w^{\mathbf{s}}_{x_0,y} f^i(y, s^k), \quad i = 1, 2, \ldots, m, \tag{3}$$

where

$$f^i(y, s^k) = \sum_{a\in A(y)} s^k_{y,a} f^i(y, a), \quad \text{for} \;\; y \in X_k, \; k \in \{1, 2, \ldots, m\} \tag{4}$$

The functions $\sigma^1_{x_0}(\mathbf{s}),\ \sigma^2_{x_0}(\mathbf{s}),\ \ldots,\ \sigma^m_{x_0}(\mathbf{s})$ on $\mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \cdots \times \mathbf{S}^m$, defined according to (3),(4), determine a game in normal form that we denote by $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\sigma^i_{x_0}(\mathbf{s})\}_{i=\overline{1,m}} \rangle$. This game corresponds to a *discounted stochastic positional game in mixed stationary strategies* that in extended form is determined by the tuple $(\{X_i\}_{i=\overline{1,m}},\ \{A(x)\}_{x\in X},\ \{f^i(x,a)\}_{i=\overline{1,m}},\ p,\ \gamma,\ x_0)$. The functions $\sigma^1_{x_0}(\mathbf{s}),\ \sigma^2_{x_0}(\mathbf{s}), \ldots, \sigma^m_{x_0}(\mathbf{s})$ on $S = S^1 \times S^2 \times \cdots \times S^m$, determine the game $\langle \{S^i\}_{i=\overline{1,m}}, \{\sigma^i_{x_0}(\mathbf{s})\}_{i=\overline{1,m}} \rangle$ that corresponds to a *discounted stochastic positional game in pure strategies.*

A stochastic positional games can be considered also for the case when the starting state is chosen randomly according to a given distribution $\{\theta_x\}$ on X. So, for a given stochastic positional game we may assume that the play starts in the state $x \in X$ with probability $\theta_x > 0$ where $\sum_{x\in X} \theta_x = 1$. If the players use mixed stationary strategies then the payoff functions

$$\sigma^i_\theta(\mathbf{s}) = \sum_{x\in X} \theta_x \sigma^i_x(\mathbf{s}), \quad i = 1, 2, \ldots, m$$

on $\mathbf{S}$ define a game in normal form $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\sigma^i_\theta(\mathbf{s})\}_{i=\overline{1,m}} \rangle$ that in extended form is determined by $(\{X_i\}_{i=\overline{1,m}}, \{A(x)\}_{x\in X}, \{f^i(x,a)\}_{i=\overline{1,m}}, p, \gamma,\ \{\theta_x\}_{x\in X})$. In the case $\theta_x = 0, \forall x \in X \setminus \{x_0\},\ \theta_{x_0} = 1$ the considered game becomes a stochastic positional game with a fixed starting state x_0.

3. Some Auxiliary Results

To prove the main results we need some properties of reward optimality equations for a discounted Markov decision problem with finite state and action spaces. Based on these properties we show how to determine the solutions of a discounted Markov decision problem and how to formulate such a problem in terms of stationary strategies as a quasi-monotonic programming problem. We shall use these results in the sequel for the discounted stochastic positional games.

3.1. Optimality Equations for a Discounted Markov Decision Process

Here we present the optimality equations for a discounted Markov decision process determined by a tuple $(X, \{A(x)\}_{x\in X}, \{f(x,a)\}_{x\in X, a\in A(x)},\ p, \gamma)$ where X is a finite set of states; $A(x)$ is a finite set of actions in $x \in X$; $f(x,a)$ is a step reward in $x \in X$ for $a \in A(x)$, $p : X \times \prod_{x\in X} A(x) \times X \ \rightarrow\ [0,1]$ is a probability transition function that satisfies the condition $\sum_{y\in X} p^a_{y\in X} = 1,\ \forall x \in X, a \in A(x)$ and γ is a discount factor.

Theorem 1. *Let a Markov decision process* $(X,\{A(x)\}_{x\in X}, \{f(x,a)\}_{x\in X, a\in A(x)}, p, \gamma)$ *be given. Then the system of equations*

$$\sigma_x = \max_{a\in A(x)} \left\{ f(x,a) + \gamma \sum_{y\in X} p^a_{x,y}\sigma_y \right\}, \quad \forall x \in X; \tag{5}$$

has a unique solution with respect to $\sigma_x,\ x \in X$*. If* $\sigma^*_x,\ x \in X$*, is the solution of system (5) then*

$$\max_{a\in A(x)} \left\{ f(x,a) + \gamma \sum_{y\in X} p^a_{x,y}\sigma^*_y - \sigma^*_x \right\} = 0 \ \ \forall x \in X$$

and an arbitrary stationary strategy

$$s^* : x \ \rightarrow\ a \in A(x) \ for\ x \in X$$

such that

$$s^*(x) = a^* \in \operatorname*{argmax}_{a\in A(x)} \left\{ f(x,a) + \gamma \sum_{y\in X} p^a_{x,y}\sigma^*_y - \sigma^*_x \right\} \ \ for \ \ x \in X$$

represents an optimal stationary strategy for the discounted Markov decision problem with an arbitrary starting state $x \in X$*; the values* ω^*_x *for* $x \in X$ *represent the optimal expected discounted sum of the rewards that correspond to optimal strategy* s^* *when the process starts in* x*.*

The proof of this theorem can be found in Puterman, 2005. Based on this theorem the optimal values $\omega^*_x,\ x \in X$ for a discounted Markov decision problem can be determined by solving the following linear programming problem:
Minimize

$$\phi_\theta(\sigma) = \sum_{x\in X} \theta_x \sigma_x \tag{6}$$

subject to

$$\sigma_x \geq f(x,a) + \gamma \sum_{y\in X} p^a_{x,y}\sigma_y,\ \forall x \in X, \ \ \forall a \in A(x) \tag{7}$$

where $\theta_x,\ x \in X$ represent arbitrary positive values such that $\sum_{x\in X} \theta_x = 1$.

3.2. Dual Linear Programming Model for a Discounted Markov Decision Problem

The dual problem for the linear programming problem (6), (7) is the following:

Maximize

$$\varphi_\theta(\alpha) = \sum_{x\in X}\sum_{a\in A(x)} f(x,a)\alpha_{x,a} \tag{8}$$

subject to

$$\begin{cases} \sum\limits_{a\in A(y)} \alpha_{y,a} - \gamma \sum\limits_{x\in X}\sum\limits_{a\in A(x)} p_{x,y}^a\,\alpha_{x,a} = \theta_y, \quad \forall y \in X; \\ \qquad \alpha_{x,a} \geq 0, \qquad \forall x \in X,\ a \in A(x), \end{cases} \tag{9}$$

where θ_y for $y \in X$ represent arbitrary positive values that satisfy the condition $\sum\limits_{y\in X} \theta_y = 1$. Here θ_y for $y \in X$ can be treated as the probabilities of choosing the starting state $y \in X$ in the decision problem. In the case $\theta_y = 1$ for $y = x_0$ and $\theta_y = 0$ for $y \in X \setminus \{x_0\}$ we obtain the linear programming model for the discounted Markov decision problem with fixed starting state x_0.

In Puterman, 2005 the following relationship is shown between feasible solutions of problem (8), (9) and stationary strategies in the discounted Markov decision problem determined by the tuple $(X,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X, a\in A(x)},\ p)$: If α is an arbitrary feasible solution of the linear programming problem (8), (9) then $\sum\limits_{a\in A(x)} \alpha_{x,a} > 0,\ \forall x \in X$ and a stationary strategy $s : x \ \rightarrow\ a \in A(x)$ for $x \in X$ that corresponds to this feasible solution is determined as follows

$$s_{x,a} = \frac{\alpha_{x,a}}{\sum\limits_{a\in A(x)} \alpha_{x,a}} \quad \text{for} \ \ x \in X_\alpha,\ a \in A(x), \tag{10}$$

where $s_{x,a}$ expresses the probability of choosing the action $a \in A(x)$ in $x \in X$.

3.3. A Discounted Markov Decision Problem in Terms of Stationary Strategies

Using the relationship between feasible solutions of problem (8), (9) and stationary strategies (12) we can formulate the discounted Markov decision problem in terms of stationary strategies as follows:
Maximize

$$\psi_\theta(\mathbf{s},\mathbf{q}) = \sum_{x\in X}\sum_{a\in A(x)} f(x,a)s_{x,a}q_x \tag{11}$$

subject to

$$\begin{cases} q_y - \gamma \sum\limits_{x\in X}\sum\limits_{a\in A(x)} p_{x,y}^a\, s_{x,a}q_x = \theta_y, \qquad \forall y \in X; \\ \qquad\qquad \sum\limits_{a\in A(y)} s_{y,a} = 1, \qquad \forall y \in X; \\ s_{x,a} \geq 0, \quad \forall x \in X,\ \forall a \in A(x); \end{cases} \tag{12}$$

where θ_y are the same values as in problem (8), (9) and $s_{x,a}, q_x$ for $x \in X, a \in A(x)$ represent the variables that must be found. It is easy to observe that for fixed $s_{x,a},\ x \in X, a \in A(x)$ system (12) uniquely determines q_x for $x \in X$. This means that $\psi_\theta(\mathbf{s},\mathbf{q})$ depends only on $\mathbf{s}$ and for a given $\mathbf{s} \in \mathbf{S}$ we have $\psi_\theta(\mathbf{s},\mathbf{q}) = \sigma_\theta(\mathbf{s})$ i.e. in (11) we can set

$$\sigma_\theta(\mathbf{s}) = \sum_{x\in X}\sum_{a\in X} f(x,a)s_{x,a}q_x.$$

So the decision problem in stationary strategies (problem (11), (12)) can be derived from (8), (9) if we introduce the following notations

$$q_x = \sum_{a\in A(x)} \alpha_{x,a},\ \forall x \in X; \qquad s_{x,a} = \frac{\alpha_{x,a}}{\sum\limits_{a\in A(x)} \alpha_{x,a}}, \quad \forall x \in X,\ a \in A(x). \tag{13}$$

This means that if $\alpha_{x,a},\ x \in X, a \in A(x)$ is a feasible solution of problem (8), (9) then $s_{x,a}, x \in X, a \in A(x)$ and $q_x,\ x \in X$, determined according to (13), represent a feasible solution of problem (11),(12). Conversely, if $s_{x,a},\ x \in X, a \in A(x)$; $q_x,\ x \in X$ is a feasible solution of problem (11),(12) then $\alpha_{x,a} = s_{x,a}q_x$, $x \in X, a \in A(x)$ represent a feasible solution of problem (8), (9).

3.4. A Quasi-Monotonic Programming Model in Stationary Strategies for a Discounted Markov Decision Problem

Based on the results from the previous section we show that a discounted Markov decision problem in stationary strategies can be represented as a quasi-monotonic programming problem. We assume that an average Markov decision problem is determined by a tuple $(X,\ \{A(x)\}_{x\in X},\ \{f(x,a)\}_{x\in X},\ p,\ \{\theta_x\}_{x\in X}, \lambda)$.

Theorem 2. *Let an average Markov decision problem be given and consider the function*

$$\sigma_\theta(\mathbf{s}) = \sum_{x\in X}\sum_{a\in A(x)} f(x,a)s_{x,a}\, q_x,$$

where q_x for $x \in X$ satisfy the condition

$$q_y - \gamma \sum_{x\in X}\sum_{a\in A(x)} p^a_{x,y}\, s_{x,a}q_x = \theta_y, \qquad \forall y \in X. \tag{14}$$

Then on the set S of solutions of the system

$$\begin{cases} \sum\limits_{a\in A(x)} s_{x,a} = 1, & \forall x \in X; \\ s_{x,a} \geq 0, & \forall x \in X,\ a \in A(x) \end{cases}$$

the function $\sigma_\theta(\mathbf{s})$ depends only on $s_{x,a}$ for $x \in X,\ a \in A(x)$ and $\sigma_\theta(s)$ is quasi-monotonic on $\mathbf{S}$ (i.e. $\sigma_\theta(\mathbf{s})$ is quasi-convex and quasi-concave on $\mathbf{S}$).

Proof. For an arbitrary $\mathbf{s} \in \mathbf{S}$ system (14) uniquely determines q_x for $x \in X$ and therefore $\psi(\mathbf{s})$ is determined uniquely for an arbitrary $\mathbf{s} \in \mathbf{S}$, i.e. the first part of the theorem holds.

Now let us prove the second part of the theorem. We show that the function $\psi(\mathbf{s})$ is quasi-monotonic on $\mathbf{S}$. To prove this it is sufficient to show that for an arbitrary $c \in \mathbf{R}^1$ the *sublevel set*

$$L_c^-(\sigma_\theta) = \{\mathbf{s} \in \mathbf{S} | \ \sigma_\theta(\mathbf{s}) \le c\}$$

and the *superlevel set*

$$L_c^+(\sigma_\theta) = \{\mathbf{s} \in \mathbf{S} | \ \sigma_\theta(\mathbf{s}) \ge c\}$$

of function $\sigma(\mathbf{s})$ are convex. These sets can be obtained respectively from the *sublevel set* $L_c^-(\varphi_\theta) = \{\alpha | \ \varphi_\theta(\alpha) \le c\}$ and the *superlevel set* $L_c^+(\varphi_\theta) = \{\alpha | \ \varphi_\theta(\alpha) \ge c\}$ of function $\varphi_\theta(\alpha)$ for the linear programming problem (8), (9).

Denote by $\alpha^i, \ \ i = \overline{1,k}$ the basic solutions of system (9). All feasible strategies of problem (8), (9) can be obtained as convex combination of basic solutions $\alpha^i, \ \ i = \overline{1,k}$. Each $\alpha^i \in \{1, 2, \ldots, k\}$ determines a stationary strategy

$$s^{(i)}_{x,a} = \frac{\alpha^i_{x,a}}{q^i_x}, \quad x \in X, \ a \in A(x) \tag{15}$$

for which $\sigma(s^{(i)}) = \varphi(\alpha^i)$ where

$$q^i_x = \sum_{a \in A(x)} \alpha^i_{x,a}, \quad \forall x \in X. \tag{16}$$

An arbitrary feasible solution α of system (9) determines a stationary strategy

$$s_{x,a} = \frac{\alpha_{x,a}}{q_x} \quad \text{for} \quad x \in X, \ a \in A(x) \tag{17}$$

for which $\sigma_\theta(\mathbf{s}) = \varphi_\theta(\alpha)$ where $q_x = \sum_{a \in A(x)} \alpha_{x,a}, \ \forall x \in X$. Taking into account that α can be represented as $\alpha = \sum_{i=1}^k \lambda^i \alpha^i$, where $\sum_{i=1}^k \lambda^i = 1, \quad \lambda^i \ge 0, \ i = \overline{1,k}$ we have $\varphi_\theta(\alpha) = \sum_{i=1}^k \varphi_\theta(\alpha^i)\lambda^i$ and we can consider

$$\alpha = \sum_{i=1}^k \lambda^i \alpha^i; \quad q = \sum_{i=1}^k \lambda^i q^i; \tag{18}$$

Using (15)–(18) we obtain

$$s_{x,a} = \frac{\alpha_{x,a}}{q_x} = \frac{\sum_{i=1}^k \lambda^i \alpha^k_{x,a}}{q_x} = \frac{\sum_{i=1}^k \lambda^i s^i_{x,a} q^i_x}{q_x} = \sum_{i=1}^k \frac{\lambda^i q^i_x}{q_x} s^{(i)}_{x,a}, \ \forall x \in X_\alpha, \ a \in A(x)$$

and

$$q_x = \sum_{i=1}^k \lambda^i q^i_x, \ \text{for} \quad x \in X. \tag{19}$$

So,

$$s_{x,a} = \sum_{i=1}^k \frac{\lambda^i q^i_x}{q_x} s^{(i)}_{x,a} \ \ \text{for} \ \ x \in X, \ a \in A(x) \tag{20}$$

where q_x for $x \in X$ are determined according to (19). The strategy s defined by (20) is a feasible strategy because $s_{x,a} \geq 0, \forall x \in X, a \in A(x)$ and $\sum_{a\in A(x)} s_{x,a} = 1$, $\forall x \in X$. Moreover, we can observe that $q_x = \sum_{i=1}^{k} \lambda^i q_x^i$, for $x \in X$ represent a solution of system (14) for the strategy s defined by (20). This can be verified by introducing (19) and (20) in (14); after such a substitution all equations from (14) are transformed into identities. For $\sigma(\mathbf{s})$ we have

$$\sigma_\theta(\mathbf{s}) = \sum_{x\in X}\sum_{a\in A(x)} f(x,a)s_{x,a}q_x = \sum_{x\in X}\sum_{a\in A(x)} f(x,a)\sum_{i=1}^{k}\left(\frac{\lambda^i q_x^i}{q_x}s_{x,a}^{(i)}\right)q_x =$$

$$\sum_{i=1}^{k}\left(\sum_{x\in X}\sum_{a\in A(x)} f(x,a)s_{x,a}^{(i)}q_x^i\right)\lambda^i = \sum_{i=1}^{k}\sigma_\theta(s^{(i)})\lambda^i,$$

i.e.

$$\sigma_\theta(s) = \sum_{i=1}^{k}\sigma_\theta(s^{(i)})\lambda^i, \tag{21}$$

where s is the strategy that corresponds to α. This means that if strategies $s^{(1)}, s^{(2)}, \dots, s^{(k)}$ correspond to basic solutions $\alpha^1, \alpha^2, \dots, \alpha^k$, of problem (8), (9) and $s \in \mathbf{S}$ corresponds to an arbitrary solution α that can be expressed as convex combination of basic solutions of problem (8), (9) with the corresponding coefficients $\lambda^1, \lambda^2, \dots, \lambda^k$ then we can express the strategy s and the corresponding value $\sigma_\theta(s)$ by (19)–(21).

Thus, an arbitrary strategy $s \in \mathbf{S}$ is determined according to (19), (20) where $\lambda^1, \lambda^2, \dots, \lambda^k$ correspond to a solution of the following system

$$\sum_{i=1}^{k}\lambda^i = 1; \quad \lambda^i \geq 0, \quad i = \overline{1,k}.$$

Consequently, the sublevel set $L_c^-(\sigma_\theta)$ of function $\sigma(s)$ represents the set of strategies s determined by (19), (20), where $\lambda^1, \lambda^2, \dots, \lambda^k$ satisfy the condition

$$\begin{cases} \sum\limits_{i=1}^{k}\sigma_\theta(s^i)\lambda^i \leq c; \\ \sum\limits_{i=1}^{k}\lambda^i = 1; \quad \lambda^i \geq 0, \quad i = \overline{1,k} \end{cases} \tag{22}$$

and the superlevel set $L_c^+(\sigma_\theta)$ of $\sigma_\theta(s)$ represents the set of strategies s determined by (19),(20), where $\lambda^1, \lambda^2, \dots, \lambda^k$ satisfy the condition

$$\begin{cases} \sum\limits_{i=1}^{k}\sigma_\theta(s^{(i)})\lambda^i \geq c; \\ \sum\limits_{i=1}^{k}\lambda^i = 1; \quad \lambda^i \geq 0, \quad i = \overline{1,k}. \end{cases} \tag{23}$$

Let us show that $L_c^-(\sigma_\theta)$, $L_c^+(\sigma_\theta)$ are convex sets. We present the proof of convexity of sublevel set $L_c^-(\sigma_\theta)$. The proof of convexity of $L_c^+(\sigma_\theta)$ is similar to the proof of convexity of $L_c^-(\sigma_\theta)$.

Denote by Λ the set of solutions $(\lambda^1, \lambda^2, \ldots, \lambda^k)$ of system (22). Then from (19), (20), (22) we have

$$L_c^-(\sigma_\theta) = \prod_{x \in X} \hat{S}_x$$

where $\hat{S}_x$ represents the set of strategies

$$s_{x,a} = \frac{\sum_{i=1}^k \lambda^i q_x^i s_{x,a}^{(i)}}{\sum_{i=1}^k \lambda^i q_x^i}, \text{ for } a \in A(x)$$

in the state $x \in X$ determined by $(\lambda^1, \lambda^2, \ldots, \lambda^k) \in \Lambda$. Here $\sum_{i=1}^k \lambda^i q_x^i > 0$ and $s_{x,a}$ for a given $x \in X$ represents a linear-fractional function with respect to $\lambda^1, \lambda^2, \ldots, \lambda^k$ defined on a convex set Λ_x and $\hat{S}_x$ is the image of $s_{x,a}$ on Λ_x. Therefore $\hat{S}_x$ is a convex set (see Boyd and Vandenberghe, 2004). □

4. Stationary Nash Equilibria for Discounted Stochastic Games

As we have noted the problem of the existence of stationary Nash equilibria for discounted stochastic games have been studied by Fink, 1964; Takahashi, 1964; Sobol, 1971 and Solan, 1998. In this section we present a normal form game for a discounted stochastic game in mixed stationary strategy and show that the payoffs of the players in such a game are continuous and quasi-monotonic with respect to the corresponding strategies of the players. Based on these properties and the results of Dasgupta and Maskin, 1986 we obtain a new proof of the existence of stationary equilibria in a discounted stochastic game. Moreover using such a model we can derive the conditions for determining the optimal stationary strategies of the players..

4.1. A Normal Form of a Discounted Stochastic Game in Stationary Strategies

In general, an m-player discounted stochastic game is determined by the following elements:

- a state space X (which we assume to be finite);
- a finite set $A^i(x)$ of actions with respect to each player $i \in \{1, 2, \ldots, n\}$ for an arbitrary state $x \in X$;
- a payoff $f^i(x, a)$ with respect to each player $i \in \{1, 2, \ldots, n\}$ for each state $x \in X$ and for an arbitrary action vector $a \in \prod_i A^i(x)$;
- a transition probability function $p : X \times \prod_{x \in X} \prod_{i=1}^n A^i(x) \times X \rightarrow [0, 1]$ that gives the probability transitions $p_{x,y}^a$ from an arbitrary $x \in X$ to an arbitrary $y \in Y$ for a fixed action vector $a \in \prod_i A^i(x)$, where $\sum_{y \in X} p_{x,y}^a = 1, \ \forall x \in X, \ a \in \prod_i A^i(x)$;

a discount factor γ, $0 < \gamma < 1$;

- a starting state $x_0 \in X$.

The game starts in the state x_0 and the play proceeds in *a sequence of stages.* At *stage* t the players observe state x_t and simultaneously and independently choose actions $a_t^i \in A^i(x_t), \quad i = 1, 2, \dots, n$. Then nature selects a state $y = x_{t+1}$ according to probability transitions $p_{x_t,y}^{a_t}$ for the given action vector $a_t = (a_t^1, a_t^2, \dots, a_t^n)$. Such a play of the game produces a sequence of states and actions $\; x_0, a_0, x_1, a_1, \dots, x_t, a_t, \dots \;$ that defines a stream of stage payoffs $f_t^1 = f^1(x_t, a_t),\ f_t^2 = f^2(x_t, a_t), \dots,\ f_t^n = f^n(x_t, a_t), \quad t = 0, 1, 2, \dots$. *The discounted stochastic game* is the game with payoffs of the players

$$\sigma_{x_0}^i = \mathsf{E}\left(\sum_{\tau=0}^{\infty} \gamma^\tau f^i(x_\tau, a_\tau)\right), \quad i = 1, 2, \dots, m$$

We will assume that players use stationary strategies of choosing the actions in the states. A stationary strategy s^i of player $i \in \{1, 2, \dots, m\}$ we define as a mapping s^i that provides for for every state $x \in X$ a probability distribution over the set of actions $A(x)$. So we can identify the set of stationary (mixed stationary strategies) strategies $\mathbf{S}^i$ of player i with the set of solutions of the system

$$\begin{cases} \sum\limits_{a \in A(x)} s_{x,a}^i = 1, & \forall x \in X; \\ s_{x,a}^i \geq 0, & \forall x \in X, \quad \forall a \in A(x). \end{cases} \tag{24}$$

Each basic solution s^i of this system corresponds to a pure stationary strategy of player $i \in \{1, 2, \dots, m\}$. So, the set of pure stationary strategies S^i of player i corresponds to the set of basic solutions of system (24).

Let $\mathbf{s} = (s^1, s^2, \dots, s^m) \in \mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \dots \times \mathbf{S}^m$ be a profile of stationary strategies (pure or mixed strategies) of the players. Then the elements of the probability transition matrix $P^{\mathbf{s}} = (p_{x,y}^{\mathbf{s}})$ in the Markov process induced by $\mathbf{s}$ can be calculated as follows:

$$p_{x,y}^s = \sum_{(a^1, a^2, \dots, a^n) \in A(x)} \prod_{k=1}^{n} s_{x,a^k}^k p_{x,y}^{(a^1, a^2, \dots, a^n)}. \tag{25}$$

Let us consider the matrix $W^{\mathbf{s}} = (w_{x,y}^s)$ where $W^{\mathbf{s}} = (I - \gamma P^{\mathbf{s}})^{-1}$. Then in a discounted stochastic positional game the payoff of player $i \in \{1, 2, \dots, m\}$ for a given profile $\mathbf{s}$ and initial state $x_0 \in X$ is determined as follows

$$\sigma_{x_0}^i(s) = \sum_{y \in X} w_{x_0,y}^s f^i(y, s), \quad i = 1, 2, \dots, n, \tag{26}$$

where

$$f^i(y, s) = \sum_{(a^1, a^2, \dots, a^n) \in A(y)} \prod_{k=1}^{n} s_{y,a^k}^k f^i(y, a^1, a^2, \dots, a^n) \tag{27}$$

The functions $\sigma_{x_0}^1(\mathbf{s}),\ \sigma_{x_0}^2(\mathbf{s}),\ \dots,\ \sigma_{x_0}^m(\mathbf{s})$ on $\mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \dots \times \mathbf{S}^m$, determined according to (26),(27), define a game in normal form that we denote by $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\sigma_{x_0}^i(\mathbf{s})\}_{i=\overline{1,m}} \rangle$. This game corresponds to a *discounted stochastic game in stationary strategies.*

The discounted stochastic game can be considered also for the case when the starting state is chosen randomly according to a given distribution $\{\theta_x\}$ on X. This means that for a given stochastic game the play starts in the state $x \in X$

with probability $\theta_x > 0$ where $\sum_{x\in X} \theta_x = 1$. If the players use mixed stationary strategies then the payoff functions

$$\sigma^i_\theta(\mathbf{s}) = \sum_{x\in X} \theta_x \sigma^i_x(\mathbf{s}), \quad i = 1, 2, \ldots, m$$

on $\mathbf{S}$ define a game in normal form $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}}, \ \{\sigma^i_\theta(\mathbf{s})\}_{i=\overline{1,m}} \ \rangle$. In the case $\theta_x = 0, \forall x \in X \setminus \{x_0\}, \ \theta_{x_0} = 1$ the considered game becomes a discounted stochastic game with a fixed starting state x_0.

Bellow we show how to represent explicitly the payoff functions $\sigma^i_\theta(\mathbf{s}^1, \mathbf{s}^2, \ldots, \mathbf{s}^m)$ on $\mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \cdots \times \mathbf{S}^m$. Based on Theorem 2 and (26), (27) the payoffs in the game $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}}, \ \{\sigma^i_\theta(\mathbf{s})\}_{i=\overline{1,m}} \ \rangle$ can be defined as follows

$$\begin{cases} \sigma^i_\theta(s^1, s^2, \ \ldots, \ s^m) = \sum\limits_{x\in X} \ \sum\limits_{(a^1,a^2,\ldots,a^m)\in A(x)} \prod\limits_{k=1}^{m} s^k_{x,a^k} f^i(x, a^1, a^2 \ldots a^m) q_x, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i = 1, 2, \ldots, m, \end{cases} \tag{28}$$

where $q_x, \ x \in X$ are determined uniquely from the following system of equations

$$q_y - \gamma \sum_{x\in X} \ \sum_{(a^1,a^2,\ldots,a^m)\in A(x)} \ \prod_{k=1}^{m} s^k_{x,a^k} p^{(a^1,a^2,\ldots,a^m)}_{x,y} q_x = \theta_y, \qquad \forall y \in X; \tag{29}$$

for an arbitrary $\mathbf{s} = (s^1, s^2, \ldots, s^m) \in \mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \cdots \times \mathbf{S}^m$, where each $\mathbf{S}^i, \ i \in \{1, 2, \ldots, m\}$ represents the set of solutions of system (24).

Each payoff function $\sigma^i_\theta(s^1, s^2, \ \ldots, \ s^m)$ on $\mathbf{S}$ is continuous. Additionally, according to Theorem 2, each $\sigma^i_\theta(s^1, s^2, \ \ldots, \ s^m)$ is quasi-monotonic with with respect to strategy $\mathbf{s}^i$ on $\mathbf{S}^i$. Therefore, based on results of Dasgupta and Maskin, 1986 and Debreu, 1952 we obtain the following theorem.

Theorem 3. *The game $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}}, \ \{\sigma^i_\theta(\mathbf{s})\}_{i=\overline{1,m}} \ \rangle$ has a Nash equilibrium $\mathbf{s} = (s^1, s^2, \ldots, s^m) \in \mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \cdots \times \mathbf{S}^m$ that is a stationary Nash equilibrium of the discounted stochastic game with an arbitrary starting state $x \in X$.*

4.2. A Normal Form of a Discounted Stochastic Positional Game in Stationary Strategies

It is easy to see that a discounted stochastic positional game determined by a tuple $(\{X_i\}_{i=\overline{1,n}}, \ \{A(x)\}_{x\in X}, \ \{f^i(x, a\}_{i=\overline{1,m}}, \ p, \ \{\theta_y\}_{y\in X})$ represent a particular case of the discounted stochastic game from Section 4.1. Therefore if we specify the game model from the previous section for the positional game then we obtain the normal form of the positional game in stationary strategies $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}}, \ \{\sigma^i_{x_0}(\mathbf{s})\}_{i=\overline{1,m}} \ \rangle$, where $\mathbf{S}^i$ and $\sigma^i_{x_0}(\mathbf{s}), \ \ i \in \{1, 2, \ldots, m\}$ are defined as follows.

Let $\mathbf{S}^i, \ i \in \{1, 2, \ldots m\}$ be the set of solutions of the system (1) that determines the set of stationary strategies of player i. On the set $\mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \cdots \times \mathbf{S}^m$ we define m payoff functions

$$\sigma^i_\theta(s^1, s^2, \ldots, s^m) = \sum_{k=1}^{m} \sum_{x\in X_k} \sum_{a\in A(x)} s^k_{x,a} f^i(x, a) q_x, \qquad i = 1, 2, \ldots, m, \tag{30}$$

where q_x for $x \in X$ are determined uniquely from the following system of linear equations

$$q_y - \gamma \sum_{k=1}^{m} \sum_{x \in X_k} \sum_{a \in A(x)} s^k_{x,a}\ p^a_{x,y}\ q_x = \theta_y, \qquad \forall y \in X; \tag{31}$$

for an arbitrary $\mathbf{s} = (s^1, s^2, \dots, s^m) \in \mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \dots \times \mathbf{S}^m$.

Note that here the payoff functions $\sigma^i_\theta(s^1, s^2, \dots, s^m)$ defined according to (30), (31) for the positional game differ from the payoff functions $\sigma^i_\theta(s^1, s^2, \dots, s^m)$ defined according (28), (29) in the general case of the game. As a corollary from Theorem 3 we obtain that for the game $\langle \{\mathbf{S}^i\}_{i=\overline{1,m}},\ \{\sigma^i_\theta(\mathbf{s})\}_{i=\overline{1,m}} \rangle$, defined according to (1), (30), (31), there exists a Nash equilibrium $\mathbf{s} = (s^1, s^2, \dots, s^m) \in \mathbf{S} = \mathbf{S}^1 \times \mathbf{S}^2 \times \dots \times \mathbf{S}^m$ that is a stationary Nash equilibrium of the discounted stochastic positional game with an arbitrary starting state $x \in X$.

5. Existence of Pure Stationary Equilibria for a Discounted Stochastic Positional Game

The existence of Nash equilibria in pure stationary strategies for a discounted stochastic positional game can be derived on the basis of the following theorem.

Theorem 4. *Let a discounted stochastic positional game be given that is determined by the tuple* $(\{X_i\}_{i=\overline{1,m}}, \{A(x)\}_{x \in X}, \{f^i(x,a\}_{i=\overline{1,m}}, p)$. *Then there exist the values* σ^i_x *for* $x \in X,\ i = 1, 2, \dots, m$ *that satisfy the following conditions:*

1) $f^i(x,a) + \gamma \sum_{y \in X} p^a_{x,y} \sigma^i_y - \sigma^i_x \le 0, \quad \forall x \in X_i, \quad \forall a \in A(x),\ i = 1, 2, \dots, m,$

2) $\max_{a \in A(x)} \{f^i(x,a) + \gamma \sum_{y \in X} p^a_{x,y} \sigma^i_y - \sigma^i_x\} = 0, \quad \forall x \in X_i,\ i = 1, 2, \dots, m;$

3) *on each position set* $X_i, i \in \{1, 2, \dots, m\}$ *there exists a map* $s^{i^*}: X_i \to \cup_{x \in X_i} A(x)$ *such that*

$$s^{i^*}(x) = a^* \in Arg \max_{a \in A(x)} \left\{ f^i(x,a) + \gamma \sum_{y \in X} p^a_{x,y} \sigma^i_y - \sigma^i_x \right\}$$

and

$$f^j(x,a^*) + \gamma \sum_{y \in X} p^{a^*}_{x,y} \sigma^j_y - \sigma^j_x = 0, \quad \forall x \in X_i, \quad j = 1, 2, \dots, m.$$

The maps $s^{1^*}, s^{2^*}, \dots, s^{m*}$ *determine a Nash equilibrium* $s^* = (s^{1^*}, s^{2^*}, \dots, s^{m*})$ *for the discounted stochastic positional game determined by* $(\{X_i\}_{i=\overline{1,m}}, \{A(x)\}_{x \in X}, \{f^i(x,a\}_{i=\overline{1,m}}, p, \gamma)$ *and* $s^* = (s^{1^*}, s^{2^*}, \dots, s^{m*})$ *is a pure stationary Nash equilibrium for the game with an arbitrary starting position* $x \in X$.

Proof. According to Theorem 3 for the discounted stochastic positional game determined by $(\{X_i\}_{i=\overline{1,m}}, \{A(x)\}_{x \in X}, \{f^i(x,a\}_{i=\overline{1,m}}, p)$ there exists a stationary Nash equilibrium $\overline{s}^* = (\overline{s}^{1^*}, \overline{s}^{2^*}, \dots, \overline{s}^*)$. If $\overline{s}^{i^*}$ is a mixed stationary strategy of player $i \in \{1, 2, \dots, m\}$ then for a fixed $x \in X_i$ the strategy $\overline{s}^{i^*}(x)$ represents a convex combination of actions determined by the probability distribution $\{\overline{s}^{i^*}_{x,a}\}$ on $A^*(x) = \{a \in A(x) |\ \overline{s}^{i^*}_{x,a} > 0\}$.

Let us consider the Markov process induced by the profile of mixed stationary strategies $\overline{s}^* = (\overline{s}^{1^*}, \overline{s}^{2^*}, \dots, \overline{s}^*)$. Then according to (2) the elements of transition probability matrix $P^{\overline{\mathbf{s}}^*} = (p^{\overline{\mathbf{s}}^*}_{x,y})$ of this Markov process can be calculated as follows

$$p^{\overline{\mathbf{s}}^*}_{x,y} = \sum_{a \in A(x)} \overline{s}^{i^*}_{x,a} p^a_{x,y} \quad \text{for} \quad x \in X_i, \quad i = 1, 2, \dots, m. \tag{32}$$

and the step rewards in the states induced by $\overline{\mathbf{s}}$ can be determined according to (4), i. e.

$$f^i(x, \overline{s}^{k^*}) = \sum_{a \in A(x)} \overline{s}^{k^*}_{x,a} f^i(x, a), \ \text{ for } \ x \in X_k, \ k \in \{1, 2, \dots, m\}. \tag{33}$$

Based on Theorem 1 for this Markov process we can write the following equations

$$f^j(x, \overline{s}^{i^*}) + \gamma \sum_{y \in X} p^{\overline{s}^{i^*}}_{x,y} \sigma^j_y - \sigma^j_x = 0, \ \ \forall x \in X_i, \ \forall i, j \in \{1, 2, \dots, m\}. \tag{34}$$

From these equations we determine uniquely $\sigma^i, \ i = 1, 2, \dots, m$ (Puterman, 2005). These values satisfy the condition

$$f^j(x, a) + \gamma \sum_{y \in X} p^a_{x,y} \sigma^j_y - \sigma^j_x \le 0, \ \ \forall x \in X_i, \forall a \in A(x) \ \ \forall i, j \in \{1, 2, \dots, m\}. \tag{35}$$

By introducing (32) and (33) in (34) we obtain

$$\sum_{a \in A(x)} \overline{s}^{i^*}_{x,a} f^j(x, a)) + \gamma \sum_{y \in X} \sum_{a \in A(x)} \overline{s}^{i^*}_{x,a} p^a_{x,y} \sigma^j_y - \sigma^j = 0, \ \forall x \in X_i, \forall i, j \in \{1, 2, \dots, m\}.$$

In these equations we can set $\sigma^i_x = \sum_{a \in A(x)} \overline{s}^{i^*}_{x,a} \sigma^j_x$. After these substitutions and some elementary transformations of the equations we obtain

$$\sum_{a \in A(x)} \overline{s}^{i^*}_{x,a} (f^j(x, a) + \gamma \sum_{y \in X} p^a_{x,y} \sigma^j_y - \sigma^j) = 0, \ \ \forall x \in X_i, \ \forall i, j \in \{1, 2, \dots, m\}.$$

So, for the Markov process induced by the profile of mixed stationary strategies $\overline{s}^* = (\overline{s}^{1^*}, \overline{s}^{2^*}, \dots, \overline{s}^*)$ there exists the values $\sigma^i_x, \ x \in X, \ i = 1, 2, \dots, m$ that satisfy the following condition

$$f^j(x, a) + \gamma \sum_{y \in X} p^a_{x,y} \sigma^j_y - \omega^j_x = 0, \ \ \forall x \in X_i, \ \forall a \in A^*(x), \ j = 1, 2, \dots, m. \tag{36}$$

Now let us fix the strategies $\overline{s}^{1^*}, \ \overline{s}^{2^*}, \ \dots, \ \overline{s}^{i-1^*}, \ \overline{s}^{i+1^*}, \ \dots, \overline{s}^*$ of the players $1, \ 2, \dots, \ i-1, \ i+1, \dots, \ m$ and consider the problem of determining the maximal expected total discounted reward with respect to player $i \in \{1, 2, \dots, m\}$. Obviously, if we solve this decision problem then we obtain the strategy $\overline{s}^{i^*}$. However for this decision problem there exists also a pure optimal strategy s^{i^*}. If we write the optimality equations for the discounted Markov decision problem with respect to player i then we obtain that there exists the values ω^i_x for $x \in X$ such that

1) $f^i(x, a) + \sum_{y \in X} p^a_{x,y} \varepsilon^i_y - \varepsilon^i_x - \omega^i \le 0, \quad \forall x \in X_i, \ \forall a \in A(x);$

2) $\max_{a\in A(x)}\left\{f^i(x,a)+\sum_{y\in X} p^a_{x,y}\varepsilon^i_y-\varepsilon^i_x-\omega^i\right\}=0, \quad \forall x\in X_i.$

We can observe that $\sigma^i_x,\ x\in X$, determined from (34), satisfy conditions 1), 2) above and (36) holds. So, if for an arbitrary $i\in\{1,2,\dots,m\}$ we fix a map $s^{i^*}: X_i\to\cup_{x\in X_i}A(x)$ such that

$$s^{i^*}(x)=a^*\in Arg\max_{a\in A(x)}\left\{f^i(x,a)+\gamma\sum_{y\in X}p^a_{x,y}\sigma^i_y-\sigma^i\right\}, \quad \forall x\in X$$

and

$$f^j(x,a^*)+\gamma\sum_{y\in X}p^{a^*}_{x,y}\sigma^j_y-\sigma^j=0, \quad \forall x\in X_i, \quad j=1,2,\dots,m$$

then we obtain a Nash equilibrium in pure stationary strategies. □

6. Conclusion

Discounted stochastic positional games represents a special class of discounted stochastic games with finite state and action spaces for which pure stationary Nash equilibria exist. The considered class of games represents a generalization of discounted deterministic positional games on graphs considered by Gurvich et al., 1988. A pure and a mixed stationary Nash equilibria for a discounted stochastic positional game can be obtained by using the game models and conditions from Sections 6,7. Stationary Nash equilibria for a discounted stochastic game can be determined by using the game model in mixed stationary strategies from Section 5.

References

Boyd, S. and Vandenberghe, L. (2004). *Convex optimization.* Cambridge University Press.

Dasgupta, P. and Maskin, E. (1986). *The existence of equilibrium in discontinuous economic games.* Rev. Econ. Stud. **53**, 1–26.

Debreu, G. (1952). *A social equilibrium existence theorem.* Proceedings of the National Academy of Sciences, **38**, 886–893.

Fink, A. (1964). *Equilibria in a stochastic n-person game.* Journal of Science of Hiroshima University, Series A-I, **28**, 89–93.

Gurvich, V. Karzanov, A. and Khachiyan, L. (1988). *Cyclic games and an algorithm to find minimax cycle means in directed graphs.* USSR Comput. Math. Math. Phis., **28**, 85–91.

Lozovanu, D. and Pickl, S. (2015). *Optimization of Stochastic Discrete Systems and Control on Complex Networks.* Springer.

Puterman, M. (2005). *Markov Decision Processes: Discrete Dynamic Programming.* Wiley, Hoboken.

Shapley, L. (1953). *Stochastic games.* Proc. Natl. Acad. Sci. USA, **39**, 1095–1100.

Sobol, M. (1971). *Non-cooperative stochastic games.* Ann. Math. Statist., **42**, 1930–1935.

Solan, E. (1998). *Discounted stochastic games.* Mathematics of Operation Research, **23(4)**, 1010–1021.

Takahashi, M. (1964). *Equilibrium points of stochastic, noncooperative n-person games.* J. Sci. Hiroshima Univ., Series A-I, **28**, 95–99.