Operations Research Letters 43 (2015) 83–88

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

# Constrained stochastic games with the average payoff criteria

Qingda Wei [a], Xian Chen [b,*]

[a] School of Economics and Finance, Huaqiao University, Quanzhou, 362021, PR China

[b] School of Mathematical Sciences, Peking University, Beijing, 100871, PR China

ARTICLE INFO

*Article history:*
Received 27 July 2014
Received in revised form
27 November 2014
Accepted 5 December 2014
Available online 12 December 2014

*Keywords:*
Nonzero-sum games
Average payoff criteria
Vanishing discount approach
Constrained Nash equilibria

ABSTRACT

In this paper we study two-person nonzero-sum constrained stochastic games under the average payoff criteria. The state space is denumerable and the action spaces of the players are Borel spaces. Under the suitable conditions, we prove the existence of a constrained stationary Nash equilibrium via the vanishing discount approach. Furthermore, we use an example to illustrate our conditions.



## 1. Introduction

Since stochastic games were introduced by Shapley in [16], they have found wide applications to many areas such as economics, computer network, communication network, queueing system and evolutionary biology; see, for instance, [15,5,12,13,1,17] and the references therein. As we can see from the existing literature, the discounted payoff and average payoff criteria are commonly used and widely studied for the nonzero-sum stochastic games, and most of the literature concerns with the unconstrained cases. However, from the viewpoint of applications, nonzero-sum constrained stochastic games form an important class of stochastic control problems; see, for instance, [1,2,18]. More precisely, [1] deals with the case of finite states and finite actions under the discounted payoff and average payoff criteria, and [2,18] discuss the case of a denumerable state space and Borel action spaces under the discounted payoff criteria. As can be seen in [2,18], the generalization of the state space from a finite set to a denumerable set is not trivial for nonzero-sum constrained stochastic games. Moreover, to the best of our knowledge, there is no literature dealing with the case of a denumerable state space for nonzero-sum constrained stochastic games under the average payoff criteria.

In this paper we study the nonzero-sum constrained stochastic games with the average payoff criteria. The state space is denumerable and the action spaces of the players are Borel spaces. The reward and cost functions of the players are bounded. As we know, a fundamental problem for the nonzero-sum stochastic games with the average payoff criteria is to find optimality conditions for the existence of Nash equilibria. For the unconstrained case, for example, [17] deals with the case of finite states and finite actions and uses a technique of reducing the existence problem to a class of recursive games. [15] discusses the case of denumerable states, finite actions and nonnegative costs, and employs the vanishing discount approach to obtain the existence of a stationary Nash equilibrium under some conditions imposed on the relative difference of the discount optimal value functions. [6] establishes the existence result for the case of an uncountable state space, compact action spaces and bounded rewards under the ergodicity condition and certain separability conditions. For the constrained case, [1] uses the fixed point approach to show the existence of a constrained stationary Nash equilibrium under the unichain and Slater conditions. It should be noted that the assumption of a finite state space in [1] cannot be dropped. Therefore, the technique in [1] is not applicable to the case of a denumerable state space. We employ the vanishing discount approach, which is completely different from that in [1], to investigate the existence of constrained Nash equilibria for the average payoff criteria. First, we introduce the auxiliary nonzero-sum constrained stochastic game model with the discounted payoff criteria. Then under the usual continuity and compactness conditions, the uniform geometrical ergodicity condition and the Slater-like condition, we prove the existence of a constrained stationary Nash equilibrium for the average payoff criteria (see Theorem 4.1). This theorem generalizes the existence result in [1]. It should be mentioned that for ease of arguments,

* Corresponding author.
*E-mail addresses:* weiqd@hqu.edu.cn (Q. Wei), chenxian@amss.ac.cn (X. Chen).

http://dx.doi.org/10.1016/j.orl.2014.12.003
0167-6377/

we only consider the two-person nonzero-sum games, but the existence result for N-person nonzero-sum games can be obtained by the same approach with obvious notational changes. Finally, we present an example to illustrate our conditions (see Example 5.1).

The rest of this paper is organized as follows. In Section 2, we introduce the nonzero-sum constrained stochastic game model. In Section 3, we give optimality conditions for the existence of constrained Nash equilibria and some preliminary lemmas. In Section 4, we state and prove our main result. In Section 5, we illustrate our conditions with an example.

## 2. The game model

In this paper we consider the two-person nonzero-sum constrained stochastic game model

$$\mathcal{G} := \Big\{X, \big(A_m, \{A_m(i) \subseteq A_m, i \in X\}, \\ r_m(i, a^1, a^2), c_m(i, a^1, a^2), \theta_m\big)_{m=1,2}, Q(j|i, a^1, a^2)\Big\},$$

for which

(i) $X$, the so-called state space, is assumed to be a denumerable set.
(ii) For each $m = 1, 2$, $A_m$ is the action space of player $m$, and is assumed to be a Borel space with Borel $\sigma$-algebra $\mathcal{B}(A_m)$. For each $i \in X$, $A_m(i) \in \mathcal{B}(A_m)$ denotes the nonempty set of admissible actions for player $m$ in the state $i \in X$. Define $K_m := \{(i, a^m) : i \in X, a^m \in A_m(i)\}$ and $K := \{(i, a^1, a^2) : i \in X, a^1 \in A_1(i), a^2 \in A_2(i)\}$.
(iii) For each $m = 1, 2$, the functions $r_m$ and $c_m$ on $K$ denote the payoffs, while the real number $\theta_m$ denotes the constraint for player $m$. Moreover, for each fixed $i \in X$, $r_m(i, \cdot)$ and $c_m(i, \cdot)$ are real-valued Borel-measurable on $A_1(i) \times A_2(i)$.
(iv) The transition function $Q(j|i, a^1, a^2)$ denotes the probability of moving from state $i$ to $j$ when the actions $a^1 \in A_1(i)$ and $a^2 \in A_2(i)$ are chosen by the players 1 and 2, respectively. Moreover, for each fixed $i, j \in X$, $Q(j|i, a^1, a^2)$ is measurable in $(a^1, a^2) \in A_1(i) \times A_2(i)$.

Let $H_t$ be the set of admissible histories of the stochastic game up to time $t$, endowed with the standard Borel $\sigma$-algebra, i.e., $H_0 := X$ and $H_t := K^t \times X$ for all $t = 1, 2, \ldots$. For each $m = 1, 2$, $\Phi^m$ denotes the set of all stochastic kernels $\varphi^m$ on $A_m$ given $X$ satisfying $\varphi^m(A_m(i)|i) = 1$ for all $i \in X$.

Below we introduce the concept of a strategy.

**Definition 2.1.** A *randomized history-dependent strategy* for player $m$ is a sequence $\pi^m := \{\pi_t^m, t = 0, 1, \ldots\}$ of stochastic kernels $\pi_t^m$ on $A_m$ given $H_t$ satisfying $\pi_t^m(A_m(i_t)|h_t) = 1$ for all $h_t = (i_0, a_0^1, a_0^2, i_1, a_1^1, a_1^2, \ldots, i_t) \in H_t$ and $t = 0, 1, \ldots$. A strategy $\pi^m$ for player $m$ is said to be stationary if there exists a stochastic kernel $\varphi^m \in \Phi^m$ such that $\pi_t^m(\cdot|h_t) = \varphi^m(\cdot|i_t)$ for all $h_t \in H_t$ and $t = 0, 1, \ldots$.

For each $m = 1, 2$, we denote by $\Pi^m$ the class of all randomized history-dependent strategies, and by $\Phi^m$ the class of all stationary strategies for player $m$. $\Pi := \Pi^1 \times \Pi^2$ and $\Phi := \Phi^1 \times \Phi^2$ denote the set of all randomized history-dependent multi-strategies and the set of all stationary multi-strategies, respectively.

Let $\Omega := (X \times A_1 \times A_2)^\infty$ and $\mathcal{F}$ be the corresponding product $\sigma$-algebra. Then for an arbitrary multi-strategy $\pi \in \Pi$ and any initial distribution $\nu$ on $X$, employing the Tulcea theorem in [8, p.178], we obtain the existence of a unique probability measure $P_\nu^\pi$ on $(\Omega, \mathcal{F})$ and a stochastic process $\{(i_t, a_t^1, a_t^2), t = 0, 1, \ldots\}$, where $i_t$, $a_t^1$ and $a_t^2$ denote the state and the actions chosen by players 1 and 2 at the time $t$, respectively. The expectation operator with respect to $P_\nu^\pi$ is denoted by $E_\nu^\pi$. If $\nu$ is the Dirac measure concentrated at the initial state $i_0 = i$, we write $P_\nu^\pi$ and $E_\nu^\pi$ as $P_i^\pi$ and $E_i^\pi$, respectively.

Fix any initial distribution $\nu$ on $X$ throughout this paper. For each $\pi^1 \in \Pi^1$ and $\pi^2 \in \Pi^2$, we define the expected average reward and expected average cost for player $m$ ($m = 1, 2$) as

$$J_{r_m}(\nu, \pi^1, \pi^2) := \liminf_{n\to\infty} \frac{1}{n} E_\nu^{\pi^1,\pi^2}\left[\sum_{t=0}^{n-1} r_m(i_t, a_t^1, a_t^2)\right]$$

and

$$J_{c_m}(\nu, \pi^1, \pi^2) := \limsup_{n\to\infty} \frac{1}{n} E_\nu^{\pi^1,\pi^2}\left[\sum_{t=0}^{n-1} c_m(i_t, a_t^1, a_t^2)\right],$$

respectively. Fix any multi-strategy of the players $\pi = (\pi^1, \pi^2) \in \Pi$. For each $m = 1, 2$ and any strategy $\widetilde{\pi}^m \in \Pi^m$ of player $m$, we denote by $[\pi^{-m}, \widetilde{\pi}^m]$ the multi-strategy obtained from $\pi$ by replacing $\pi^m$ with $\widetilde{\pi}^m$. For each $m = 1, 2$, a strategy $\overline{\pi}^m \in \Pi^m$ is said to be *feasible* for player $m$ against $\pi$ if $J_{c_m}(\nu, [\pi^{-m}, \overline{\pi}^m]) \le \theta_m$. The set of all feasible strategies for player $m$ against $\pi$ is denoted by $\Delta^m(\pi)$. A multi-strategy $\overline{\pi} = (\overline{\pi}^1, \overline{\pi}^2) \in \Pi$ is said to be *feasible* for the constrained game model $\mathcal{G}$ if $\overline{\pi}^m \in \Delta^m(\overline{\pi})$ for $m = 1, 2$. The set of all feasible multi-strategies for the game model $\mathcal{G}$ is denoted by $\Delta$.

**Definition 2.2.** A multi-strategy $\pi^* = (\pi^{*1}, \pi^{*2}) \in \Pi$ is called a constrained Nash equilibrium for the game model $\mathcal{G}$ if

$$\pi^* \in \Delta \text{ and } J_{r_m}(\nu, \pi^*) = \sup_{\pi^m \in \Delta^m(\pi^*)} J_{r_m}(\nu, [\pi^{*-m}, \pi^m]) \quad \text{for } m = 1, 2.$$

Finally, for each $(f^1, f^2) \in \Phi$, $i, j \in X$ and $m = 1, 2$, we define

$$r_m(i, f^1, f^2) := \int_{A_1(i)} \int_{A_2(i)} r_m(i, a^1, a^2) f^1(da^1|i) f^2(da^2|i),$$
$$c_m(i, f^1, f^2) := \int_{A_1(i)} \int_{A_2(i)} c_m(i, a^1, a^2) f^1(da^1|i) f^2(da^2|i),$$
$$Q(j|i, f^1, f^2) := \int_{A_1(i)} \int_{A_2(i)} Q(j|i, a^1, a^2) f^1(da^1|i) f^2(da^2|i).$$

The main goal of this paper is to give conditions for the existence of constrained Nash equilibria under the average payoff criteria.

## 3. Preliminaries

In this section, we will give conditions for the existence of constrained Nash equilibria and some preliminary results.

In order to study the nonzero-sum constrained stochastic games with the average payoff criteria, we need to introduce the following auxiliary game model

$$\mathcal{G}_\alpha := \Big\{X, \big(A_m, \{A_m(i) \subseteq A_m, i \in X\}, \\ r_m(i, a^1, a^2), c_m(i, a^1, a^2), \theta_m^\alpha\big)_{m=1,2}, \alpha, Q(j|i, a^1, a^2)\Big\},$$

where the constants $\alpha \in (0, 1)$ and $\theta_m^\alpha$ denote the discount factor and the constraint imposed on the expected discounted cost of player $m$, respectively, and the other components are the same as in the game model $\mathcal{G}$.

For any discount factor $\alpha \in (0, 1)$, $\pi^1 \in \Pi^1$, and $\pi^2 \in \Pi^2$, the expected discounted reward and expected discounted cost for