Operations Research Letters 44 (2016) 575–577

Contents lists available at ScienceDirect

Operations Research Letters

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

Uniform value for recursive games with compact action sets

Xiaoxi Li [a,*], Sylvain Sorin [b]

[a] Economics and Management School, Wuhan University, Luojia Hill, 430072 Wuhan, China

[b] CNRS, IMJ-PRG, UMR 7586, Sorbonne Universités, UPMC Univ Paris 06, Univ Paris Diderot, Sorbonne Paris Cité, Case 247, 4 Place Jussieu, 75252 Paris, France

ARTICLE INFO

Article history:
Received 12 April 2016
Received in revised form
16 June 2016
Accepted 16 June 2016
Available online 24 June 2016

Keywords:
Stochastic games
Recursive games
Value
Long-term analysis

ABSTRACT

Mertens, Neyman and Rosenberg (Mertens et al., 2009) used the Mertens and Neyman theorem (Mertens and Neyman, 1981) to prove the existence of a uniform value for absorbing games with a finite state space and compact action sets. We provide an analogous proof for another class of stochastic games, recursive games with a finite state space and compact action sets. Moreover, both players have stationary ε-optimal strategies.

1. Introduction

Zero-sum stochastic games were introduced by [10], and the model $\Gamma = \langle K, I, J, g, q\rangle$ is as follows: K is the finite state space, I (resp. J) is the action set for player 1 (resp. player 2), $g : K \times I \times J \to \mathbb{R}$ is the stage payoff function and $q : K \times I \times J \to \Delta(K)$ is a probability transition function ($\Delta(S)$ stands for the set of probabilities on a measurable set S). We assume throughout that I and J are compact metric sets and that both g and q are separately continuous on $I \times J$ (this implies their measurability, cf. I.1.Ex.7a in [8]).

The game is played as follows. Let $k_1 \in K$ be the initial state. At each stage $t \geq 1$, after observing a t-stage history $h_t = (k_1, i_1, j_1, \ldots, k_{t-1}, i_{t-1}, j_{t-1}, k_t)$, player 1 chooses an action $i_t \in I$ and player 2 chooses an action $j_t \in J$. This profile (k_t, i_t, j_t) induces a current stage payoff $g_t := g(k_t, i_t, j_t)$ and a probability $q(k_t, i_t, j_t)$ which is the law of k_{t+1}, the state at stage $t + 1$.

Let $H_t = K \times (I \times J \times K)^{t-1}$ be the set of t-stage histories for each $t \geq 1$, and $H_\infty = (K \times I \times J)^\infty$ be the set of infinite histories. We endow H_t with the product sigma-algebra $\mathcal{H}_t$ (discrete on K, Borel on I and J) and endow $H_\infty = (K \times I \times J)^\infty$ with the product sigma-algebra $\mathcal{H}_\infty$ spanned by $\cup_{t\geq 1} \mathcal{H}_t$. A *behavior strategy* σ for player 1 is a sequence $\sigma = (\sigma_t)_{t\geq 1}$ where for each $t \geq 1$, σ_t is a measurable map from $(H_t, \mathcal{H}_t)$ to $\Delta(I)$. A behavior strategy τ for player 2 is defined analogously. The set of player 1's (resp. player 2's) behavior strategies is denoted by Σ (resp. $\mathcal{T}$).

Given an initial state k_1, any strategy profile $(\sigma, \tau) \in \Sigma \times \mathcal{T}$ induces a unique probability distribution $\mathbb{P}^{k_1}_{\sigma,\tau}$ over the histories $(H_\infty, \mathcal{H}_\infty)$. The corresponding expectation is denoted by $\mathbb{E}^{k_1}_{\sigma,\tau}$.

In a λ-discounted game Γ_λ (for $\lambda \in (0, 1]$), the (global) payoff is defined as $\gamma_\lambda(k_1, \sigma, \tau) = \mathbb{E}^{k_1}_{\sigma,\tau}\left[\sum_{t\geq 1} \lambda(1-\lambda)^{t-1} g(i_t, j_t, k_t)\right]$ and the corresponding (minmax) value is v_λ. The n-stage game Γ_n (for $n \geq 1$) is defined analogously by taking the n-stage averaged payoff, and its value is denoted by v_n.

In the finite actions setup, [10] introduced the operator Φ where $\forall f \in \mathbb{R}^{|K|}$,

$$\Phi(\lambda, f)(k) = \mathtt{val}_{x\in\Delta(I), y\in\Delta(J)} \mathbb{E}^k_{x,y}\left[\lambda g(i, j, k) + (1-\lambda) f(k')\right], \quad (1)$$

with $\mathtt{val}_{x\in\Delta(I),y\in\Delta(J)} = \max_{x\in\Delta(I)} \min_{y\in\Delta(J)} = \min_{y\in\Delta(J)} \max_{x\in\Delta(I)}$. He proved that

$$v_\lambda = \Phi(\lambda, v_\lambda) \quad (2)$$

and moreover that *stationary* optimal strategies exist for each λ, i.e. depending at each stage t only on the current state k_t. These results extend to the current framework (cf. VII.1.a in [8]).

We are interested in long-run properties of Γ. A first notion corresponds to the existence of an *asymptotic value*: convergence of v_λ as λ tends to zero and convergence of v_n as n tends to infinity, to the same limit. Moreover, one can ask for the existence of ε-optimal strategies for both players that guarantee the asymptotic value in all sufficiently long games, explicitly:

* Corresponding author.
E-mail addresses: xxleewhu@gmail.com (X. Li), sylvain.sorin@imj-prg.fr (S. Sorin).

http://dx.doi.org/10.1016/j.orl.2016.06.005
0167-6377/

Definition 1.1. Let $w \in \mathbb{R}^{|K|}$. Play 1 can guarantee w in Γ if, for any $\varepsilon > 0$ and for every $k_1 \in K$, there exist $N(\varepsilon) \in \mathbb{N}$ and a behavior strategy $\sigma^* \in \Sigma$ for player 1 s.t.: $\forall \tau \in \mathcal{T}$,

$$(A)\ \frac{1}{n}\mathbb{E}^{k_1}_{\sigma^*,\tau}\Big[\sum_{t=1}^{n} g_t\Big] \geq w(k_1) - \varepsilon, \quad \forall n \geq N(\varepsilon),$$

$$(B)\ \mathbb{E}^{k_1}_{\sigma^*,\tau}\Big[\liminf_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} g_t\Big] \geq w(k_1) - \varepsilon.$$

A similar definition holds for player 2.

v is the uniform value of Γ if both players can guarantee it.

Remark. The existence of a uniform value v implies the existence of an *asymptotic value*, equal to v.

For stochastic games with a finite state space and finite action sets, [1] proved the convergence of v_λ as λ tends to zero (and later deduced the convergence of v_n as n tends to infinity to the same limit), relying on an algebraic argument. Using the property that the function $\lambda \mapsto v_\lambda$ has bounded variation, that follows from [1], [6] proved the existence of a uniform value. Actually, Mertens and Neyman's main result is even applicable to a stochastic game Γ with compact action sets under the following form:

Theorem 1.2 ([6]). *Let $\lambda \to w_\lambda \in \mathbb{R}^{|K|}$ be a function defined on $(0, 1]$. Player 1 can guarantee $\limsup_{\lambda\to 0^+} w_\lambda$ in the stochastic game Γ if w_λ satisfies:*

(i) *for some integrable function $\phi : (0, 1] \to \mathbb{R}_+$, $\|w_\lambda - w_{\lambda'}\|_\infty \leq \int_\lambda^{\lambda'} \phi(x)dx$, $\forall \lambda, \lambda' \in (0, 1)$, $\lambda < \lambda'$;*
(ii) *for every $\lambda \in (0, 1)$ sufficiently small, $\Phi(\lambda, w_\lambda) \geq w_\lambda$.*

Remark. In the construction of an ε-optimal strategy in [6], w_λ is taken to be v_λ, so condition (i) is implied by the bounded variation property of v_λ and condition (ii) is implied by Eq. (2).

Below we focus on two important classes of stochastic games: absorbing games and recursive games.

An *absorbing state* is such that the probability of leaving it is zero. Without loss of generality one assumes that at any absorbing state, the payoff is constant (equal to the value of the static game to be played after absorption), as long as one states that both players are informed of the current state.

A stochastic game Γ is an *absorbing game* if all states but one are absorbing. [7] used Theorem 1.2 to prove the existence of a uniform value for absorbing games with a finite state space and compact action sets, extending a result of [4] for the finite actions case.

Recursive games, introduced by [3], are stochastic games where the stage payoff is zero in all nonabsorbing states.

This note proves the existence of a uniform value for recursive games with a finite state space and compact action sets, using an approach analogous to [7] for absorbing games. Moreover, due to the specific payoff structure, we show that ε-optimal strategies in recursive games can be taken stationary. This is not the case for a general stochastic game, in which an ε-optimal strategy has to be usually a function of the whole past history, even in the finite actions case (cf. [2] for the "Big match" as an example).

[3] proved the existence of stationary ε-optimal strategies for the "limiting-average value" (property (B) in Definition 1.1). As our proof relies on his characterization of this value (and on its existence), we describe here the result.

Given $S \subseteq \mathbb{R}^d$, let $\overline{S}$ denote its closure.

Let $K^0 \subseteq K$ be the set of nonabsorbing states.

$\Phi(0, \cdot)$ refers to the operator $\Phi(\lambda, \cdot)$ with $\lambda = 0$ in Eq. (1).

When working with the operator $\Phi(0, \cdot)$ or $\Phi(\lambda, \cdot)$, it is sufficient to consider those vectors $u \in \mathbb{R}^K$ identical to the absorbing payoffs on the absorbing states $K \setminus K_0$. Whenever no confusion is caused, we identify u with its projection on $\mathbb{R}^{|K_0|}$.

Theorem 1.3 ([3]). *A recursive game Γ has a limiting-average value v and both players have stationary ε-optimal strategy, in the sense that $\forall \varepsilon > 0$, there are stationary strategies $(\sigma^*, \tau^*) \in \Sigma \times \mathcal{T}$ s.t.: for any $(\sigma, \tau) \in \Sigma \times \mathcal{T}$,*

$$\mathbb{E}^{k_1}_{\sigma^*,\tau}\Big[\liminf_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} g_t\Big] \geq v(k_1) - \varepsilon \quad \text{and}$$

$$\mathbb{E}^{k_1}_{\sigma,\tau^*}\Big[\limsup_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} g_t\Big] \leq v(k_1) + \varepsilon.$$

Moreover, the limiting-average value v is characterized by $\{v\} = \overline{\xi^+} \cap \overline{\xi^-}$, where

$$\xi^+ = \Big\{u \in \mathbb{R}^{|K^0|} : \Phi(0, u) \geq u,\ \text{and}\ \Phi(0, u)(k) > u(k)\ \text{whenever}\ u(k) > 0\Big\},$$

$$\xi^- = \Big\{u \in \mathbb{R}^{|K^0|} : \Phi(0, u) \leq u,\ \text{and}\ \Phi(0, u)(k) < u(k)\ \text{whenever}\ u(k) < 0\Big\}.$$

[3]'s proof of the above result consists of the following two arguments: first, any vector $u \in \xi^+$ (*resp.* $u \in \xi^-$) can be guaranteed by player 1 (*resp.* player 2); second, the intersection of $\overline{\xi^+}$ and $\overline{\xi^-}$ is nonempty.

2. Main results and the proof

We prove that v (as characterized in Theorem 1.3) is also the uniform value of Γ, and that players have stationary ε-optimal strategies.

Theorem 2.1. *A recursive game has a uniform value. Moreover, both players can guarantee the uniform value in stationary strategies.*

Remark. We emphasize that our definition of uniform value includes that of limiting-average value, thus our results extend [3] to a much stronger set-up.

Proof. We first prove that v is the uniform value of Γ using Theorem 1.2. Let u be any vector in ξ^+. An equivalent characterization for u is:

$$\Phi(0, u) \geq u \text{ and } u(k) \leq 0 \text{ whenever } \Phi(0, u)(k) = u(k), \quad \forall k \in K^0.$$

Define $w_\lambda = u$, $\forall \lambda \in (0, 1)$. We check that w_λ satisfies the two conditions (i) and (ii) in Theorem 1.2.

(i) It is trivial since u does not depend on λ;
(ii) The crucial point is that $\Phi(\lambda, u) = (1 - \lambda)\Phi(0, u)$ on K^0 for recursive games. Then the condition $\Phi(\lambda, w_\lambda) \geq w_\lambda$ for all λ close to 0 is satisfied in the following two cases:
   - for $k \in K^0$ with $\Phi(0, u)(k) > u(k)$, we have $(1 - \lambda)\Phi(0, u)(k) > u(k)$ for λ close to 0^+;
   - for $k \in K^0$ with $\Phi(0, u)(k) = u(k)$, we have $u(k) \leq 0$, thus $(1 - \lambda)\Phi(0, u)(k) = (1 - \lambda)u(k) \geq u(k)$, $\quad \forall \lambda \in (0, 1)$.

Now Theorem 1.2 states that player 1 can guarantee any $u \in \xi^+$. A symmetric argument proves that player 2 can guarantee any vector $u \in \xi^-$. As $\{v\} = \overline{\xi^+} \cap \overline{\xi^-}$ is nonempty by Theorem 1.3, this proves that v is the uniform value of Γ.

Next, we point out that any ε-optimal strategy appearing in the Mertens and Neyman theorem can be taken *stationary* when Γ is a recursive game. Indeed, let w_λ be the function satisfying conditions (i) and (ii) in Theorem 1.2. The general construction of σ for player 1 to guarantee $\limsup_{\lambda\to 0^+} w_\lambda$ is as follows: at each stage $t \geq 1$,

- Step 1. the t-stage history is used to compute in an inductive way a small enough discount factor λ_t;
- Step 2. play an optimal strategy $x_{\lambda_t}(w_{\lambda_t}, k_t) \in \Delta(I)$ in the zero-sum game corresponding to $\Phi(\lambda_t, w_{\lambda_t})(k_t)$, defined in Eq. (1) by letting $f = w_{\lambda_t}, \lambda = \lambda_t, k = k_t$.

Since (1) we have chosen in the first part of the proof $w_\lambda = u$ to be constant, (2) $\Phi(\lambda, u) = (1-\lambda)\Phi(0, u)$ for recursive games, this implies that $x_{\lambda_t}(w_{\lambda_t}, k_t)$ can be taken as $x(u, k_t)$, an optimal strategy of player 1 in the zero-sum game associated to $\Phi(0, u)(k_t) = \texttt{val}_{x\in\Delta(I), y\in\Delta(J)}\mathbb{E}^{k_t}_{x,y}[u(k')]$. Hence, to define σ, we do not need (as in [6]) the whole t-stage history to compute λ_t, and the only necessary information is the current state k_t. $\square$

3. Concluding remarks

1. To have a better understanding of the existence of stationary ε-optimal strategy in recursive games, one can compare our construction to the one in [7] for absorbing games. Indeed, they have also chosen w_λ to be some constant function u. However, there is no such equality $\Phi(\lambda, u) = (1-\lambda)\Phi(0, u)$ for absorbing games, so the optimal strategy $x_{\lambda_t}(w_{\lambda_t}, k_t)$ at each stage t for $\Phi(\lambda_t, u)$ depends on λ_t, hence on the whole history. On the other hand, choosing w_λ as v_λ (like in [13]) will induce strategies that will still depend upon λ.
2. Since the constructed ε-optimal strategy is stationary, our proof extends to *recursive games with signals on actions*: there is no need to assume perfect observation of the opponent's actions, as long as both players are informed of the current state.
3. For absorbing games with a finite state space and compact action sets, [9] characterized and proved the existence of an asymptotic value, via the so-called *operator approach*, following [4]. Using the same approach, [11] provided the corresponding result for recursive games with a finite state space and compact action sets and in particular showed that the asymptotic value is equal to the unique vector characterized by $\overline{\xi^+} \cap \overline{\xi^-}$ in [3], see also [12].
   However, these convergence results do not extend to general stochastic games with a finite state space and compact action sets: [14] provided an example with no convergence of v_λ. This implies a fortiori that the existence result of a uniform value for absorbing/recursive games with compact actions does not extend to general stochastic games.
4. Our proof does not extend to *recursive games with an infinite state space*. In fact, we used the nonemptiness of $\overline{\xi^+} \cap \overline{\xi^-}$, which is obtained in [3] by an inductive proof on the (finite) number of states.
   [5] provided a sufficient condition for recursive games with an infinite state space to have a uniform value, that is, the family of n-stage values $\{v_n\}$ being totally bounded for the uniform norm. They presented also an example for which this condition is not satisfied and no uniform value exists.

Acknowledgments

The authors are grateful to Guillaume Vigeral for helpful comments. Part of Xiaoxi Li's research is done when he was an ATER (teaching and research fellow) at THEMA, Université Cergy-Pontoise during the academic year 2015–2016.

References

[1] T. Bewley, E. Kohlberg, The asymptotic theory of stochastic games, Math. Oper. Res. 1 (1976) 197–208.
[2] D. Blackwell, T.S. Ferguson, The big match, Ann. Math. Statist. 39 (1968) 159–163.
[3] H. Everett, Recursive games, in: M. Dresher, A.W. Tucker, P. Wolfe (Eds.), Contributions to the Theory of Games III, in: Annals of Mathematical Studies, vol. 39, Princeton University Press, 1957, pp. 47–78.
[4] E. Kohlberg, Repeated games with absorbing states, Ann. Statist. 2 (1974) 724–738.
[5] X. Li, X. Venel, Recursive games: uniform value, Tauberian theorem and the Mertens conjecture "*Maxmin* $= \lim_{n\to\infty} v_n = \lim_{\lambda\to 0} v_\lambda$", Internat. J. Game Theory 45 (2016) 155–189.
[6] J.-F. Mertens, A. Neyman, Stochastic games, Internat. J. Game Theory 10 (1981) 53–66.
[7] J.-F. Mertens, A. Neyman, D. Rosenberg, Absorbing games with compact action spaces, Math. Oper. Res. 34 (2009) 257–262.
[8] J.-F. Mertens, S. Sorin, S. Zamir, Repeated Games, Cambridge University Press, 2015.
[9] D. Rosenberg, S. Sorin, An operator approach to zero-sum repeated games, Israel J. Math. 121 (2001) 221–246.
[10] L.S. Shapley, Stochastic games, Proc. Natl. Acad. Sci. 39 (1953) 1095–1100.
[11] S. Sorin, The operator approach to zero-sum stochastic games, in: A. Neyman, S. Sorin (Eds.), Stochastic Games and Applications, Kluwer Academic Publishers, 2003, pp. 417–426. (Chapter 27).
[12] S. Sorin, G. Vigeral, Existence of the limit value of two person zero-sum discounted repeated games via comparison theorems, J. Optim. Theory Appl. 157 (2013) 564–576.
[13] F. Thuijsman, O.J. Vrieze, Note on recursive games, in: B. Dutta, D. Mookherjee, T. Parthasarathy, T.E.S. Raghavan, S.H. Tijs (Eds.), Game Theory with Economic Applications, in: Lecture Notes in Economics and Mathematical Systems, vol. 389, Springer-Verlag, 1992, pp. 133–145.
[14] G. Vigeral, A zero-sum stochastic game with compact action sets and no asymptotic value, Dyn. Games Appl. 3 (2013) 172–186.