Discounted stochastic game with unbounded costs on compact action spaces

Lu YU
email: yulumaths@gmail.com

September 28, 2024

1 Introduction

For a discounted, zero-sum stochastic game whose state space and action spaces are compact metric spaces, and whose payoff function is continuous (hence bounded), Parthasarathy [Par73, Thm. 3.1] shows that there is a value, and both players have optimal stationary strategies. This is used in the proof of [Par73, Thm. 4.1], which shows the following. For a discounted, positive, zero-sum stochastic game, if the sate spaces and action spaces are finite, then the stochastic game has a value and the two players have optimal stationary strategies.

For a stochastic game with countable state, compact metric action space and bounded payoff functions, Federgruen [Fed78, Thm. 1] establishes the existence of stationary discounted equilibrium point (DEP). He applies this to obtain the existence of average return equilibrium point in [Fed78, Theorems 2, 3] under a number of recurrency conditions.

The opposite of a payoff function is called a *cost function*. Sennott considers stochastic game with countable state space, finite actions spaces and unbounded cost functions. If the cost functions are determined by the discount optimality equation in the sense of [Sen94, Assumptions D1, D2], then [Sen94, Thm. 2.2] proves the existence of stationary λ-DEP for all $\lambda \in [0, 1)$. In addition, under [Sen94, Assumption A1] (among other assumptions) that there exists a λ-DEP for every $\lambda \in (0, 1)$, Sennott [Sen94, Thm. 4.2] demonstrates the existence of stationary average expected cost equilibrium point. Moreover, [Sen94, Thm. 2.2] is used in [Sen94, Prop. 5.1] to prove that any limit point of λ-DEP is an average expected cost equilibrium point.

Discounted, positive, zero-sum stochastic games with countable state, countable action spaces and unbounded payoff functions are considered in [FPS20]. By [FPS20, Thm. 11], if at every state, at least one player has a finite action space, then the game has a value.

In the case of [Sen94, Sec. 2], if the cost functions are bounded, then [Sen94, Assumptions D1, D2] hold. Sennott [Sen94, Sec. 3] gives a realistic example whose cost functions are unbounded and satisfy [Sen94, Assumptions D1, D2]. Thus, when the action spaces are finite, Sennott's theorem [Sen94, Thm. 2.2] is a strict generalization of Federgruen's theorem [Fed78, Thm. 1]. A natural question is to find a result covering both theorems. Theorem 3.7 provides a positive answer to this question.

It is likely that one can apply Theorem 3.7 to study the existence of stationary average expected cost equilibrium point for stochastic games, as well as the existence of value for positive zero-sum stochastic games, both with compact action spaces and unbounded payoff functions.

The text is organized as follows. In Section 2, we consider stochastic games with one player, i.e., Markov decision processes. We prove that the optimality equation characterizes stationary optimal strategies. In Section 3, we define the stochastic game model, then prove the existence of DEP. Finally, Section 4 explains how Theorem 3.7 specializes to Federgruen's theorem [Fed78, Thm. 1] and Sennott's theorem [Sen94, Thm. 2.2].

Notation

For a metric space X, let $P(X)$ be the set of Borel probability measures on X. We identify X with the subset of $P(X)$ of Dirac measures. Let $M(X)$ be the set of signed finite Borel measures on X.

2 Markov decision process

Definition 2.1. A Markov decision process is a vector $(S, (A(s))_{s\in S}, q, r)$, where

- S is a nonempty at most countable set of states, endowed with the discrete topology;
- A is a metric space and for every $s \in S$, $A(s) \subset A$ is a nonempty compact subset of actions available in state s. Let $\mathrm{SA} := \sqcup_{s\in S} A(s)$;
- $q : \mathrm{SA} \to P(S), \quad (s, a) \mapsto q_{s,\cdot}(a)$ is a transition rule;
- $r : \mathrm{SA} \to [0, +\infty)$ is a cost function.

A Markov decision process involves a decision maker. The process lasts for infinitely many stages. First, the initial state $s_0 \in S$ is given. At each stage $k \geq 0$, the current state s_k is announced to the decision maker. Then the decision maker chooses an action $a_k \in A(s_k)$ and pays the stage cost $r(s_k, a_k)$. A new state $s_{k+1} \in S$ is drawn according to $q_{s_k,\cdot}(a_k) \in P(A(s_k))$, and the game proceeds to stage $k + 1$.

Assumption 2.2. For any $s, t \in S$, the functions $r(s, \cdot) : A(s) \to [0, +\infty)$ and $q_{st} : A(s) \to [0, 1]$ are continuous.

For an integer $k \geq 0$, the set of histories till stage k is defined by $H_k := (\mathrm{SA})^k \times S$. The set of all histories is $H := \sqcup_{k\geq 0} H_k$, which is equipped with the natural sigma-algebra. Let $H_\infty = (\mathrm{SA})^{\mathbb{N}}$ be the set of plays. A strategy is a measurable map that assigns to each history $h = (s_1, a_1, \ldots, a_{k-1}, s_k)$ an element of $P(A(s_k))$. Let Π be the set of strategies. A map $\delta : S \to P(A)$ with $\delta(s)(A(s)) = 1$ for all $s \in S$ is called a *stationary strategy.* The set of all stationary strategies is $\Delta := \prod_{s\in S} P(A(s))$.

With the sigma-algebra generated by the cylinder sets, H_∞ is a measurable space. By Assumption 2.2, the functions $r : \mathrm{SA} \to [0, +\infty)$ and $q : \mathrm{SA} \to P(S)$ are measurable. Therefore, an initial state $s \in S$ and a strategy $\pi \in \Pi$ define a probability measure on H_∞ (see, e.g., [Sol22, p.9]). Let $\mathbb{E}_{s,\pi}$ be the corresponding expectation operator. Fix a discount factor $\lambda \in [0, 1)$. Define a function

$$V : S \times \Pi \to [0, +\infty], \quad (s, \pi) \mapsto (1 - \lambda)\mathbb{E}_{s,\pi}(\sum_{k\geq 0} \lambda^k r(s_k, a_k)).$$

It is the *expected total discounted cost* for the decision maker. For every $s \in S$, let $B(s) := \inf_{\pi\in\Pi} V(s, \pi)$. Assume Assumptions 2.2 and 2.3 in Section 2.

Assumption 2.3. [Sen89, Assumption 1] For every $s \in S$, the quantity $B(s)$ is finite.

Lemma 2.4. *For every $s \in S$, one has*

$$B(s) = \min_{a\in A(s)} ((1 - \lambda)r(s, a) + \lambda \sum_{t\in S} q_{st}(a)B(t)).$$

In particular, there is $a \in A(s)$ such that the series $\sum_{t\in S} q_{st}(a)B(t)$ converges.

Proof. Define a function

$$f : A(s) \to [0, +\infty], \quad a \mapsto \sum_{t\in S} q_{st}(a)B(t).$$

We prove that f is lower semicontinuous. Consider $c \in \mathbb{R}$ and a *sequence* $\{a_n\}_{n>0}$ in $[f \leq c] := \{a \in A(s) | f(a) \leq c\}$ with limit a in $A(s)$. For every $t \in S$, by continuity of $q_{st} : A(s) \to [0, 1]$ (Assumption 2.2) and Fatou's lemma [Rud87, 1.28], one has

$$f(a) \leq \liminf_{n\to+\infty} f(a_n) \leq c.$$

Since $A(s)$ is a metric space, its subset $[f \leq c]$ is closed.

By lower semicontinuity of f, continuity of $r(s, \cdot) : A(s) \to \mathbb{R}$ (Assumption 2.2) and compactness of $A(s)$, there is $a_0 \in A(s)$ at which point the function

$$(1 - \lambda)r(s, \cdot) + \lambda f : A(s) \to [0, +\infty]$$

attains its minimum (denoted $B'(s)$). The right hand side is thus well-defined. We shall prove $B(s) = B'(s)$.

Let $\pi \in \Pi$ be a strategy. Let $\mu := \pi(s) \in P(A(s))$ be the probability of the action at stage 0. For every $(t, \pi') \in S \times \Pi$, let $W(t, \pi')$ be the expected total discounted cost from stage 1 onward, with strategy π' and the state at stage 1 being t. When the state at stage 1 is t, the situation at this time is the same as if the process had started in state t, but all returns are multiplied by λ. Hence $W(t, \pi') \geq \lambda B(t)$. One has

$$\begin{aligned} V(s, \pi) &= \int_{A(s)} ((1-\lambda) r(s, a) + \sum_{t \in S} q_{st}(a) W(t, \pi)) d\mu(a) \\ &\geq \int_{A(s)} ((1-\lambda) r(s, a) + \sum_{t \in S} q_{st}(a) \lambda B(t)) d\mu(a) \\ &\geq \int_{A(s)} B'(s) d\mu(a) = B'(s). \end{aligned}$$

Hence $B(s) \geq B'(s)$. In particular, $B'(s)$ is a real number.

We prove $B(s) \leq B'(s)$. By Assumption 2.3, for every $\epsilon > 0$ and every $t \in S$, there is a strategy $\pi_t \in \Pi$ with $V(t, \pi_t) < B(t) + \epsilon$. Consider a strategy π that chooses a_0 at stage 0 and, if the next state is t, then follow the strategy π_t. One has

$$\begin{aligned} & B(s) \leq V(s, \pi) \\ =& (1-\lambda) r(s, a_0) + \lambda \sum_{t \in S} q_{st}(a_0) V(t, \pi_t) \\ \leq& (1-\lambda) r(s, a_0) + \lambda \sum_{t \in S} q_{st}(a_0)(B(t) + \epsilon) \\ =& B'(s) + \epsilon\lambda. \end{aligned}$$

Letting $\epsilon \to 0$, one has $B(s) \leq B'(s)$. □

For every $s \in S$ and every $\mu \in M(A(s))$, define $r(s, \mu) := \int_{A(s)} r(s, a) d\mu(a)$. For every $t \in S$, define $q_{st}(\mu) = \int_{A(s)} q_{st}(a) d\mu(a)$. Define a map

$$F_s : M(A(s)) \to \mathbb{R}, \quad \mu \mapsto (1-\lambda) r(s, \mu) + \lambda \sum_{t \in S} q_{st}(\mu) B(t).$$

A strategy $\pi \in \Pi$ is called λ-*discount optimal* if $V(s, \pi) = B(s)$. We prove that a stationary strategy satisfying the optimality equation is optimal.

Lemma 2.5. *Let $\delta \in \Delta$ be a stationary strategy. If for every $s \in S$, one has $\delta(s) \in \arg\min_{P(A(s))} F_s$, then δ is λ-discount optimal.*

Proof. For every $s \in S$ and every integer $k \geq 0$, let $V_k(s, \delta)$ be the expected discounted cost under δ before stage k. By convention, $V_0(s, \delta) = 0$. By induction on k, we prove $V_k(s, \delta) \leq B(s)$. Since $r \geq 0$ on SA, one has

$V_0(s,\delta) \le B(s)$. Assume the inequality for $k-1$. Let $\mu := \delta(s) \in P(A(s))$. One has

$$\begin{aligned}
V_k(s,\delta) &\overset{\text{(a)}}{=} \int_{A(s)} ((1-\lambda)r(s,a) + \lambda \sum_{t\in S} q_{st}(a) V_{k-1}(t,\delta)) d\mu(a) \\
&\overset{\text{(b)}}{\le} \int_{A(s)} ((1-\lambda)r(s,a) + \lambda \sum_{t\in S} q_{st}(a) B(t)) d\mu(a) \\
&= F_s(\mu) \overset{\text{(c)}}{=} \min_{P(A_s)} F_s \le \min_{A(s)} F_s \overset{\text{(d)}}{=} B(s),
\end{aligned}$$

where (a) is because that δ is stationary, (b) is from the inductive hypothesis, (c) is by $\mu \in \arg\min_{P(A(s))} F_s$, and (d) follows from Lemma 2.4.

The induction is completed. Then $V(s,\delta) \le B(s)$. Therefore, δ is λ-discount optimal. □

3 Discounted stochastic game

Model

Let S be a nonempty, at most countable set of states. Let I be a finite nonempty set of players. For every $i \in I$, let A^i be a metric space of actions for player i. For every $s \in S$, let $A^i(s) \subset A^i$ be a nonempty *compact* subset of actions feasible for player i in state s. Set $A(s) := \prod_{i\in I} A^i(s)$. Then $\mathrm{SA} = \sqcup_{s\in S} A(s)$ is the set of all feasible action profiles at all states. Endow it with the disjoint union topology. For every $i \in I$, let $r^i : \mathrm{SA} \to \mathbb{R}$ be the cost function for player i. The transition rule is a map $q : \mathrm{SA} \to P(S)$.

Definition 3.1. A stochastic game $G = (S, I, \{A^i(s)\}, r^i, q)$ lasts for countably infinitely many stages. An initial state $s_0 \in S$ is given. At each stage $k \ge 0$, the following takes place:

- The current state s_k is announced to all players.
- Each player $i \in I$ chooses an action $a^i_k \in A^i(s_k)$. The players' choices are made simultaneously and independently.
- The action profile $(a^i_k)_{i\in I}$ is publicly announced to all players.
- Each player $i \in I$ pays a stage cost $r^i(s_k, a_k)$.
- A new state $s_{k+1} \in S$ is drawn according to the probability $q_{s_k,\cdot}(a_k) \in P(S)$, and the game proceeds to stage $k+1$.

Remark 3.2. A stochastic game with I being a singleton is exactly a Markov decision process.

Fix a stochastic game G.

Assumption 3.3 ([Sen89, p.627], [Fed78, (1.8)]). For every $i \in I$ and any $s, t \in S$, the functions $r^i(s, \cdot) : A(s) \to [0, +\infty)$ and $q_{st} : A(s) \to [0, 1]$ are continuous.

The notation H_k, H, H_∞ can be defined as in Section 2. For a player $i \in I$, a strategy is a measurable map that assigns to each history $h_k = (s_0, a_0, \ldots, s_k) \in H$ an element of $P(A^i(s_k))$. The set of all strategies of player i is denoted by Π^i. Let $\Pi = \prod_{i \in I} \Pi^i$ and $\Pi^{-i} = \prod_{j \neq i} \Pi^j$. By Assumption 3.3, an initial state $s \in S$ and $\pi \in \Pi$ define a probability on the measurable space H_∞. Let $\mathbb{E}_{s,\pi}$ be the corresponding expectation operator. For each player $i \in I$, fix a discount factor $\lambda_i \in [0, 1)$. The function

$$V^i : S \times \Pi \to [0, +\infty], \quad (s, \pi) \mapsto (1 - \lambda_i)\mathbb{E}_{s,\pi}[\sum_{k \geq 0} \lambda_i^k r^i(s_k, a_k)]$$

is called the *expected total λ_i-discounted cost* to player i by strategy π at initial state s.

Definition 3.4. [Fed78, (1.12)] A strategy $\pi \in \Pi$ is called a λ-discounted equilibrium point if for every initial state $s \in S$, every $i \in I$ and every $\gamma^{-i} \in \Pi^{-i}$, one has $V^i(s, \pi) \leq V^i(s, \gamma^{-i}, \pi^i)$.

A stationary strategy of player i is a map $\delta^i : S \to P(A^i)$ with $\delta^i(s)(A^i(s)) = 1$. Let $\Delta^i := \prod_{s \in S} P(A^i(s))$ be the set of all stationary strategies of player i. Set $\Delta^{-i} := \prod_{j \neq i} \Delta^j$ and $\Delta := \prod_{i \in I} \Delta^i$.

Existence of stationary equilibrium

For every $i \in I$, every $\delta^{-i} \in \Delta^{-i}$ and every $s \in S$, let $B^i(s, \delta^{-i}) \in [0, +\infty]$ be the infimum of the function

$$\Pi^i \to [0, +\infty], \quad \pi^i \mapsto V^i(s, \delta^{-i}, \pi^i).$$

It is the best total cost of player i, given the initial state s and the strategy δ^{-i} of other players.

We adopt Assumption 3.5 introduced by Sennott.

Assumption 3.5. [Sen94, Assumptions D1, D2] For every convergent sequence $\{\delta_n\}_{n \geq 1}$ with limit δ in Δ, there is a function $R : I \times S \to [0, +\infty)$ such that

1. for every $s \in S$, every $i \in I$ and every integer $n \geq 1$, one has $B^i(s, \delta_n^{-i}) \leq R^i(s)$;

2. for every $s \in S$, every $i \in I$ and every $a \in A(s)$, one has $\sum_{t \in S} q_{st}(a) R^i(t) < \infty$, and the resulting function $\sum_{t \in S} q_{st}(\cdot) R^i(t) : A(s) \to \mathbb{R}$ is continuous;

3. for every function $U : I \times S \to [0, +\infty)$ with $U \leq R$ pointwise, the condition

$$\begin{aligned} U^i(s) =&(1-\lambda_i) r^i(s, \delta(s)) + \lambda_i \sum_{t \in S} q_{st}(\delta(s)) U^i(t) \\ \leq&(1-\lambda_i) r^i(s, \delta^{-i}(s), a) + \lambda_i \sum_{t \in S} q_{st}(\delta^{-i}(s), a) U^i(t) \end{aligned} \tag{1}$$

for all $i \in I$, $s \in S$ and $a \in A^i(s)$ implies $U^i = B^i(\cdot, \delta^{-i})$ on S for all $i \in I$.

Remark 3.6. For every $i \in I$ and every $\delta^{-i} \in \Delta^{-i}$, applying Assumption 3.5 1 to the constant sequence $\{\delta^{-i}\}_{n \geq 1}$, one has $0 \leq B^i(s, \delta^{-i}) < \infty$ for all $s \in S$. Because δ^{-i} is stationary, player i is faced with a Markov decision process (in the sense of Definition 2.1) satisfying Assumption 2.3. By Assumptions 3.3, the process also satisfies Assumption 2.2.

A function $f : C \to \mathbb{R}$ defined on a convex subset C of a real vector space is *quasiconvex* if for all $x, y \in C$ and $\lambda \in [0, 1]$, one has $f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}$. The proof of Theorem 3.7 relies on Kakutani's fixed point theorem.

Theorem 3.7. *Under Assumptions 3.3 and 3.5,*

1. *the stochastic game G admits a stationary λ-DEP.*
2. *Assume further that for every $i \in I$ and every $s \in S$, the space $A^i(s)$ is a convex and compact subset of some locally convex topological vector space, and the function $r^i(s, \cdot) : A^i(s) \to [0, +\infty)$ is quasiconvex. Then G admits a pure, stationary λ-DEP.*

Proof. By [Fol13, Thm. 7.8], for a compact metric space X, the set of signed Radon measures on X in the sense of [Fol13, 7.15–7.16] coincides with $M(X)$. Then by the Riesz representation theorem [Fol13, Cor. 7.18], $M(X)$ is dual to the Banach space $(C(X), \|\cdot\|_{\sup})$. Endow $M(X)$ with the weak star topology, which makes it a locally convex topological vector space.

For every $\delta^{-i} \in \Delta^{-i}$, define a function

$$\begin{gathered} F_{s,\delta^{-i}} : M(A^i(s)) \to \mathbb{R}, \\ \mu \mapsto (1-\lambda_i) r^i(s, \delta^{-i}(s), \mu) + \lambda_i \sum_{t \in S} q_{st}(\delta^{-i}(s), \mu) B^i(t, \delta^{-i}). \end{gathered}$$

Let $\Phi^i(\delta)_s := \arg\min_{P(A^i(s))} F_{s,\delta^{-i}}$. By construction, the function $F_{s,\delta^{-i}}$ is linear. By Assumption 3.3, the restriction $F_{s,\delta^{-i}} : P(A^i(s)) \to \mathbb{R}$ is continuous. From [Cho68, p.108], as $A^i(s)$ is compact, the subset $P(A^i(s)) \subset M(A^i(s))$ is convex compact, and the map taking the Dirac measures $A^i(S) \to P(A^i(S))$ is a topological closed embedding identifying $A^i(s)$ with the set of extreme points of $P(A^i(s))$. Therefore, $\Phi^i(\delta)_s$ is nonempty, convex and closed in $M(A^i(s))$. By Remark 3.6 and Lemma 2.4, one has $B^i(s, \delta^{-i}) = \min_{A^i(s)} F_{s,\delta^{-i}}$. Then by Bauer's maximum principle [NP06, Cor. A.3.3], one has $B^i(s, \delta^{-i}) = \min_{P(A^i(s))} F_{s,\delta^{-i}}$.

For every $i \in I$, define a correspondence

$$\Phi^i : \Delta \to 2^{\Delta^i}, \quad \delta \mapsto \prod_{s \in S} \Phi^i(\delta)_s.$$

Claim 3.8. For every $i \in I$, any sequences $\gamma_n \to \gamma$ in Δ and $\delta_n^i \to \delta^i$ in Δ^i with $\delta_n^i \in \Phi^i(\gamma_n)$ for all integers $n > 0$, one has $\delta^i \in \Phi^i(\gamma)$.

From [Dud66, Thm. 12], for every $i \in I$ and every $s \in S$, since $A^i(s)$ is a metric space, the weak star topology on $P(A^i(s))$ is metrizable. Because $I \times S$ is countable, both Δ^i and Δ are metrizable. So Claim 3.8 proves that $\Phi^i : \Delta \to 2^{\Delta_i}$ has closed graph.

1. Define a correspondence

   $$\Phi : \Delta \to 2^{\Delta}, \quad \delta \mapsto \prod_{i \in I} \Phi^i(\delta).$$

   By [Fan52, Lemma 3], Φ also has closed graph. Then every value of Φ is nonempty and closed in the locally convex vector space $\prod_{i \in I} \prod_{s \in S} M(A^i(s))$. By Tychonoff's theorem, Δ is compact in $\prod_{i \in I} \prod_{s \in S} M(A^i(s))$. From Kakutani's fixed point theorem [Fan52, Thm. 1], Φ admits a fixed point $\delta \in \Delta$.

   By Lemma 2.5, for every $i \in I$, the strategy δ^i is λ_i-discount optimal for the Markov decision process faced by player i induced by δ^{-i}. Thus, δ is a stationary λ-DEP.

2. For every $i \in I$, let $X^i = \prod_{s \in S} A^i(s)$. Set $X := \prod_{i \in I} X^i$. For every $s \in S$ and every $x \in X$, let $\Psi^i(x)_s := \arg\min_{A^i(s)} F_{s,x^{-i}}$. Then $\Psi^i(x)_s = \Phi^i(x)_s \cap A^i(s)$. By compactness of $A^i(s)$ and continuity of $r^i(s, \cdot)$ in Assumption 3.3, the subset $\Psi^i(x)_s$ of $A^i(s)$ is nonempty and closed. As $r^i(s, \cdot) : A^i(s) \to \mathbb{R}$ is quasiconvex, the subset is convex.

   Define a correspondence $\Psi^i : X \to 2^{X_i}, \quad x \mapsto \prod_{s \in S} \Psi^i(x)_s$. Then the graph of Ψ^i is the intersection of the graph of Φ^i with $X \times X_i$. Define $\Psi : X \to 2^X, \quad x \mapsto \prod_{i \in I} \Psi^i(x)$. Then Ψ has closed graph. By Tychonoff's theorem, the X^i and hence X are compact. As the $A^i(s)$ are convex compact subspaces of locally convex topological vector spaces, so is X. From Kakutani's fixed point theorem [Fan52, Thm. 1], Ψ admits a fixed point $x_0 \in X$. By Lemma 2.5, the strategy x_0 is a pure, stationary λ-DEP.

□

Proof of Claim 3.8. Indeed, for every $n > 0$ and every $a \in A^i(s)$, one has

$$\begin{aligned} B^i(s, \gamma_n^{-i}) &\overset{\text{(a)}}{=} (1 - \lambda_i) r^i(s, \gamma_n^{-i}(s), \delta_n^i(s)) + \lambda_i \sum_{t \in S} q_{st}(\gamma_n^{-i}(s), \delta_n^i(s)) B^i(t, \gamma_n^{-i}) \\ &\leq (1 - \lambda_i) r^i(s, \gamma_n^{-i}(s), a) + \lambda_i \sum_{t \in S} q_{st}(\gamma_n^{-i}(s), a) B^i(t, \gamma_n^{-i}), \end{aligned}$$

where (a) is from Lemma 2.4. Let $R : I \times S \to [0, +\infty)$ be the function given by Assumption 3.5 for the convergent sequence $\gamma_n \to \gamma$. By Tychonoff's theorem, for every $i \in I$, the product topology of $\prod_{s \in S}[0, R^i(s)]$ is compact. Because S is countable, the product space is metrizable. In this compact metrizable space, the sequence $\{(B^i(s, \gamma_n^{-i}))_{s\in S}\}_{n\geq 1}$ has a convergent subsequence (still denoted by the original sequence). For every $s \in S$, let $U^i(s) := \lim_{n\to\infty} B^i(s, \gamma_n^{-i})$. Then $U^i(s) \in [0, R^i(s)]$.

By finiteness of I and [Bil99, Thm. 2.8 (ii)], for every $s \in S$, the map taking product measure $\Delta(s) := \prod_{i\in I} P(A^i(s)) \to P(A(s))$ is continuous. Then by Assumption 3.3, one has

$$r^i(s, \gamma_n^{-i}(s), \delta_n^i(s)) \to r^i(s, \gamma^{-i}(s), \delta^i(s)),$$
$$r^i(s, \gamma_n^{-i}(s), a) \to r^i(s, \gamma^{-i}(s), a),$$

and for every $t \in S$, one has

$$q_{st}(\gamma_n^{-i}(s), \delta_n^i(s)) \to q_{st}(\gamma^{-i}(s), \delta^i(s)),$$
$$q_{st}(\gamma_n^{-i}(s), a) \to q_{st}(\gamma^{-i}(s), a).$$

From Assumption 3.5 2, the map $\sum_{t\in S} q_{st}(\cdot)R^i(t) : \Delta(s) \to \mathbb{R}$ is continuous. Thus, one has

$$\sum_{t\in S} q_{st}(\gamma_n^{-i}(s), \delta_n^i(s))R^i(t) \to \sum_{t\in S} q_{st}(\gamma^{-i}(s), \delta^i(s))R^i(t),$$
$$\sum_{t\in S} q_{st}(\gamma_n^{-i}(s), a)R^i(t) \to \sum_{t\in S} q_{st}(\gamma^{-i}(s), a)R^i(t).$$

Then by the generalized Lebesgue dominated convergence theorem [Roy68, Ch. 11, Prop. 18], one has

$$\sum_{t\in S} q_{st}(\gamma_n^{-i}(s), \delta_n^i(s))B^i(t, \gamma_n^{-i}) \to \sum_{t\in S} q_{st}(\gamma^{-i}(s), \delta^i(s))U^i(t),$$
$$\sum_{t\in S} q_{st}(\gamma_n^{-i}(s), a)B^i(t, \gamma_n^{-i}) \to \sum_{t\in S} q_{st}(\gamma^{-i}(s), a)U^i(t).$$

Thus, U satisfies condition (1). From Assumption 3.5 3, one has $U^i(s) = B^i(s, \gamma^{-i})$ and the result follows. □

4 Recover classical results

We explain how Theorem 3.7 contains both Federgruen's and Sennott's theorems, by showing that their hypotheses imply our assumptions respectively.

Remark 4.1. Sennott [Sen94, p.146] assumes that for every $i \in I$ and every $s \in S$, the set $A^i(s)$ is finite. With the discrete topology, it is compact, and Assumption 3.3 is satisfied.

Federgruen [Fed78, p.453] assumes that for every $i \in I$, the function $r^i : \mathrm{SA} \to \mathbb{R}$ is bounded. Since the results are not affected by the addition of a constant to all cost functions, one may assume that there is a constant $C_i > 0$ with $0 \leq r^i \leq C_i$ on SA. Define $R : I \times S \to [0, +\infty)$ by $R^i(s) = C_i$.

Lemma 4.2. *Then R satisfies Assumption 3.5.*

Proof.

1. For every $\pi \in \Pi^i$, one has $V^i(s, \delta_n^{-i}, \pi) \leq (1 - \lambda_i) \sum_{t \geq 0} \lambda_i^t C_i = R^i(s)$. Hence $B^i(s, \delta_n^{-i}) \leq R^i(s)$.

2. One has $\sum_{t \in S} q_{st}(a) R^i(t) = C_i < \infty$. The constant function $C_i : A(s) \to \mathbb{R}$ is continuous.

3. Let E be the vector space of bounded functions on S. Then $(E, \| \cdot \|_{\sup})$ is a Banach space. Fix $i \in I$. By [Bil99, Thm. 2.8 (ii)] and [Yu23, Lemma 4.5], for every $w \in E$, every $s \in S$ and every $\beta \in \Delta^{-i}$, the function $\sum_{t \in S} q_{st}(\beta, \cdot) w(t) : A^i(s) \to \mathbb{R}$ is continuous. By compactness of $A^i(s)$ and Assumption 3.3, the function

$$(1 - \lambda_i) r^i(s, \delta^{-i}(s), \cdot) + \lambda_i \sum_{t \in S} q_{st}(\delta^{-i}(s), \cdot) w(t) : A^i(s) \to \mathbb{R}$$

attains its minimum, denoted by $T(w)(s)$. Since $w : S \to \mathbb{R}$ is bounded, so is $T(w) : S \to \mathbb{R}$. Hence a map

$$T : E \to E, \quad T(w)(s) = \min_{a \in A^i(s)} ((1 - \lambda_i) r^i(s, \delta^{-i}(s), a) + \lambda_i \sum_{t \in S} q_{st}(\delta^{-i}(s), a) w(t))$$

for $w \in E$ and $s \in S$.

We prove that T is contracting. For any $w, u \in E$ and every $s \in S$, one has

$$\begin{aligned}
|Tw(s) - Tu(s)| &= | \min_{a \in A^i(s)} ((1 - \lambda_i) r^i(s, \delta^{-i}(s), a) + \lambda_i \sum_{t \in S} q_{st}(\delta^{-i}(s), a) w(t)) \\
&\qquad - \min_{a \in A^i(s)} ((1 - \lambda_i) r^i(s, \delta^{-i}(s), a) + \lambda_i \sum_{t \in S} q_{st}(\delta^{-i}(s), a) u(t))| \\
&\leq \max_{a \in A^i(s)} |((1 - \lambda_i) r^i(s, \delta^{-i}(s), a) + \lambda_i \sum_{t \in S} q_{st}(\delta^{-i}(s), a) w(t)) \\
&\qquad - ((1 - \lambda_i) r^i(s, \delta^{-i}(s), a) + \lambda_i \sum_{t \in S} q_{st}(\delta^{-i}(s), a) u(t))| \\
&= \lambda_i \max_{a \in A^i(s)} | \sum_{t \in S} q_{st}(\delta^{-i}(s), a)(w(t) - u(t))| \\
&\leq \lambda_i \max_{a \in A^i(s)} \sum_{t \in S} q_{st}(\delta^{-i}(s), a) \|w - u\|_{\sup} \\
&= \lambda_i \|w - u\|_{\sup}.
\end{aligned}$$

Hence $\|Tw - Tu\|_{\sup} \leq \lambda_i \|w - u\|_{\sup}$.

By Banach's fixed point theorem, T has a unique fixed point. By Part 1, one has $B^i(\cdot, \delta^{-i}) \in E$. From Remark 3.6 and Lemma 2.4, the function $B^i(\cdot, \delta^{-i})$ is a fixed point of T.

Since $U \leq R$ on $I \times S$, one has $U^i \in E$. By hypothesis on U, one has $U^i \leq T(U^i)$ on S. Since $\delta^i(s)$ is a probability on $A^i(s)$, one has $U^i \geq T(U^i)$. Then U^i is also a fixed point of T, so $U^i = B^i(\cdot, \delta^{-i})$ on S. $\square$

By Remark 4.1 (resp. Lemma 4.2), one recovers [Sen94, Thm. 2.2] (resp. [Fed78, Thm. 1]) from Theorem 3.7.

Acknowledgments

I thank Feng Shao and Sichong Zhang for answering my questions.

References

[Bil99] Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons, 1999.

[Cho68] Gustave Choquet. *Lectures on Analysis: Representation theory*, volume II. Benjamin/Cummings, 1968.

[Dud66] Rl Dudley. Convergence of Baire measures. *Studia Mathematica*, 3(27):251–268, 1966.

[Fan52] Ky Fan. Fixed-point and minimax theorems in locally convex topological linear spaces. *Proceedings of the National Academy of Sciences*, 38(2):121–126, 1952.

[Fed78] Awi Federgruen. On N-person stochastic games by denumerable state space. *Advances in Applied Probability*, 10(2):452–471, 1978.

[Fol13] Gerald B Folland. *Real analysis: modern techniques and their applications.* John Wiley & Sons, 2nd edition, 2013.

[FPS20] János Flesch, Arkadi Predtetchinski, and William Sudderth. Positive zero-sum stochastic games with countable state and action spaces. *Applied Mathematics & Optimization*, 82(2):499–516, 2020.

[NP06] Constantin Niculescu and Lars-Erik Persson. *Convex functions and their applications*, volume 23. Springer, 2006.

[Par73] T. Parthasarathy. Discounted, positive, and noncooperative stochastic games. *International Journal of Game Theory*, 2:25–37, 1973.

[Roy68] Halsey Lawrence Royden. *Real analysis*, volume 2. Macmillan company, 2nd edition, 1968.

[Rud87] Walter Rudin. *Real and complex analysis*. McGraw-Hill, 3rd edition, 1987.

[Sen89] Linn I. Sennott. Average cost optimal stationary policies in infinite state markov decision processes with unbounded costs. *Operations Research*, 37(4):626–633, 1989.

[Sen94] Linn I. Sennott. Nonzero-sum stochastic games with unbounded costs: discounted and average cost cases. *Zeitschrift für Operations Research*, 40:145–162, 1994.

[Sol22] Eilon Solan. *A course in stochastic game theory*, volume 103. Cambridge University Press, 2022.

[Yu23] Lu Yu. Existence of pure stationary equilibria for stochastic games, 2023.