JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 106, No. 2, pp. 421–429, AUGUST 2000

TECHNICAL NOTE

Markov Perfect Equilibrium Existence for a Class of Undiscounted Infinite-Horizon Dynamic Games[1]

A. Garcia[2] and R. L. Smith[3]

Communicated by D. A. Carlson

Abstract. We prove the existence of Markov perfect equilibria (MPE) for nonstationary undiscounted infinite-horizon dynamic games with alternating moves. A suitable finite-horizon equilibrium relaxation, the ending state constrained MPE, captures the relevant features of an infinite-horizon MPE for a long enough horizon, under a uniformly bounded reachability assumption.

Key Words. Dynamic games, infinite horizon, average reward, alternating moves.

1. Introduction

The traditional approach to prove the existence of equilibria of infinite-horizon games relies heavily on the continuity of payoff functionals. The procedure begins by proving the existence of finite-horizon equilibria and follows by taking limits as the horizon diverges. Compactness of the infinite-horizon strategy space or the history space is usually required to ensure the existence of a limit point which will inherit by continuity the desired properties; see, for example, Fudenberg and Levine (Refs. 1–2), Harris (Ref. 3), and Borgers (Ref. 4).

[1]This research was supported in part by NSF Grant DMI-9713723 and Department of Defense Research and Engineering Multidisciplinary Initiative Grant ARO-DAAH-0496-1-0377.
[2]Associate, The Brattle Group, Washington, DC.
[3]Professor, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan.

In dynamic games with undiscounted or average reward payoffs, this approach fails, since the infinite-horizon payoff functionals are not continuous. Moreover, since future rewards are as valuable as present rewards, end-of-horizon effects are magnified; thus, there may exist infinite-horizon and finite-horizon equilibria of a substantially different nature. In other words, there are infinite-horizon equilibrium strategies that are not the limit of finite-horizon equilibrium strategies. These issues have been examined recently by Engwerda (Ref. 5) in the context of linear-quadratic games.

In this paper, we provide a new proof of the existence of Markov perfect equilibria (MPE) in the context of infinite-horizon nonstationary undiscounted dynamic games with alternating moves. The proof relies on a new method to overcome end-of-horizon effects as in Schochetman and Smith (Ref. 6). The idea is to restrict the deviation possibilities for players by forcing an ending target state for every finite horizon. This relaxation leads to the definition of a constrained MPE. A uniformly bounded reachability assumption, which essentially requires that every player in isolation can partially control the state dynamics, ensures that play in early periods (as opposed to play in late periods) is more relevant in identifying profitable deviations in the long run. Carlson and Haurie (Ref. 7) used a relatively similar technique for open-loop equilibrium with decoupled dynamics. We apply our results to an asynchronous dynamic duopoly [see Maskin and Tirole (Ref. 8)].

2. Preliminaries

2.1. Dynamic Games with Alternating Moves. For ease of exposition, we restrict our discussion to the case of two-player interaction. At every stage k, which is discretely indexed, each player $i \in \{1, 2\}$ takes an action a^i from the feasible action set A_k^i. The instantaneous reward $r_k^i: S \times A_k^1 \times A_k^2 \to \mathscr{R}$ is a function of the players actions and the current value of the state variable $s \in S$.

A transition function $f_k: S \times A_k^1 \times A_k^2 \to S$ determines the evolution of the state variable. A Markov strategy for player i in a T-horizon game, say π_i^T, is a T-tuple of maps $\pi_k^i: S \to A_k^i$, so that π_i^T is of the form

$$\pi_i^T = (\pi_0^i, \pi_1^i, \ldots, \pi_{T-1}^i).$$

We denote by $\Pi^i(T)$ the set of all such strategies for player $i \in \{1, 2\}$. We refer to the 2-tuple $\pi^T \in \Pi^1(T) \times \Pi^2(T)$ as a Markov strategy combination and denote by $\Pi(T)$ the set of all such strategy combinations.

The set of T-long feasible sequences of action profiles that players may exert is commonly referred to as the history space,

$$H(T)=\prod_{k=0}^{T-1} A_k^1 \times A_k^2.$$

We shall also denote by $h_T^{\pi^T}(s_0)\in H(T)$ the feasible history induced by strategy combination π^T from the initial state s_0. The total sum of rewards per stage for feasible history $h_T^{\pi^T}(s_0)\in H(T)$ is given by

$$P_T^i(h_T^{\pi^T}(s_0))=\sum_{k=0}^{T-1} r_k^i(s_k, \pi_1^T(s_k), \pi_2^T(s_k)).$$

Similarly, we shall denote by

$$h_T^{\pi^T}(s_k)\in \prod_{k}^{T-1} A_k^1 \times A_k^2$$

the feasible history of play induced by strategy combination π^T from intermediate state $s_k\in S$ at time period k. As above, the payoff obtained for this case will be denoted by

$$P_T^i(h_T^{\pi^T}(s_k))=\sum_{j=k}^{T-1} r_j^i(s_j, \pi_1^T(s_j), \pi_2^T(s_j)).$$

The extension of a dynamic game, when there is an infinite number of stages to play, follows straightforwardly by setting the history space to be the infinite Cartesian product

$$H=\prod_{k=0}^{\infty} A_k^1 \times A_k^2.$$

We shall denote by Π the set of all infinite-horizon feasible strategy combinations. The total aggregated reward received by player i under the infinite-horizon strategy combination π is defined as follows:

$$P^i(h^\pi(s_0))=\lim \inf_{T\to\infty} P_T^i(h_T^{\pi^T}(s_0))/T,$$

where π^T stands for the T-horizon truncation of the infinite-horizon strategy combination π.

Finally, in a game with alternating moves, players revise their actions in alternation: player 1 at odd periods $k=1, 3, 5, \ldots$, and player 2 at even periods $k=0, 2, 4, 6 \ldots$. Formally,

$$A_k^1 = A_{k+1}^2 = \{\emptyset\}, \qquad \text{for even } k.$$

2.2. Markov Perfect Equilibrium. For an excellent introduction to the concept of MPE, the reader is referred to Fudenberg and Tirole (Chapter 13, Ref. 9).

Definition 2.1. Markov Perfect Equilibrium. We say that π^T is a Markov perfect equilibrium (MPE) iff every player i who would like to deviate from π^T by playing $\gamma_i^T \in \Pi^i(T)$ from any intermediate state $s_k \in S$, $0 \leq k \leq T-1$, cannot find any incentive in doing so, i.e:

$$P_T^i(h_T^{(\gamma_i^T, \pi_{-i}^T)}(s_k)) \leq P_T^i(h_T^{\pi^T}(s_k)),$$

where $(\gamma_i^T, \pi_{-i}^T) \in \Pi(T)$ stands for the strategy combination in which player $j, j \neq i$, follow π_j^T and player i follows γ_i^T.

This definition carries over straightforwardly to the infinite-horizon setting with the above introduced framework.

We denote $\Pi^*(T)$ and Π^* the set of all Markov perfect equilibrium strategies for the T-horizon and infinite-horizon games, respectively.

Definition 2.2. Constrained Strategies. We denote by $\Pi(T, s)$ the set of constrained strategy combinations to state $s \in S$; $\pi^T \in \Pi(T, s)$ if and only if the history prescribed from every intermediate state reaches state s at time period T.

Definition 2.3. Constrained MPE. A strategy combination $\pi^T \in \Pi(T, s)$ is called a constrained MPE to state s iff, for every deviation $\gamma_i^T \in \Pi(T)$ such that $(\gamma_i^T, \pi_{-i}^T) \in \Pi(T, s)$ from every intermediate state $s_k \in S$ at time period k, we have

$$P_T^i(h_T^{(\gamma_i^T, \pi_{-i}^T)}(s_k)) \leq P_T^i(h_T^{\pi^T}(s_k)).$$

We denote by $\Pi^*(T, s)$ the set of all constrained MPE to state s.

3. Existence of Markov Perfect Equilibria

We make the following standing assumptions.

Assumption 2.1.

(i) Discreteness. Each set A_k^i is discrete and finite; hence, the history space H is compact in the product topology and Π is compact in the topology $\mathscr{L}$ (see the Appendix for a brief discussion).

(ii) Reward Boundedness. For every player i and for every time period k,

$$-\infty < -M \leq r_k^i(\cdot,\cdot) \leq M < \infty.$$

Assumption 2.2. Uniformly Bounded Reachability. There exists an infinite feasible sequence of states, say $s = (s_0, s_1, s_2, \ldots)$ such that, from any intermediate state of the sequence, say s, at time period k, for every player i, there exists a sequence of actions $\{a_j^i\}_{k<j\leq T}$, so that state s_T in the sequence is reached regardless of the other players actions. Moreover, the number of time periods required is bounded by a finite number L, i.e., $T-k \leq L < \infty$.

3.1. Application. Sequential Duopoly. In this section, we illustrate briefly the definitions introduced above for the case of a duopoly competition in prices, as in Maskin and Tirole (Ref. 10). Players move sequentially so that, in odd numbered periods k, firm 1 chooses its price which remains unchanged until period $k+2$; that is,

$$p_{k+1}^1 = p_k^1, \qquad \text{if } k \text{ is odd.}$$

Similarly, firm 2 chooses prices only in even numbered periods,

$$p_{k+1}^2 = p_k^2, \qquad \text{if } k \text{ is even.}$$

Hence, at time period k, the firm i instantaneous reward $r_k^i(\cdot)$ is a function of the state, i.e., the price that firm j sets on period $k-1$, say p_k^j, and the action, i.e. the price that firm i will establish p_k^i. The set of feasible pricing decisions (say P) is discrete and finite, goods are perfect substitutes that is, the firms share the market equally whenever they charge the same price. Firms have the same unit cost c. Let $D_k(\cdot)$ denote the market demand function at time period k. The total reward at time period k is given by

$$r_k(p) = (p-c)D_k(p), \qquad p \in P.$$

Then,

$$r_k^i(p^1, p^2) = \begin{cases} r_k(p^i), & \text{if } p^i < p^j, \\ r_k(p^i)/2, & \text{if } p^i = p^j, \\ 0, & \text{if } p^i > p^j. \end{cases}$$

Strategies are Markovian in that they depend on the current state, i.e., the rival action in the last period. Hence, the set of all histories is the same as the set of all feasible sequences of states.

Consider the infinite history $h=\{(p_k^1, p_k^2)\}_k$; then, the firm i undiscounted payoff is

$$P^i(h) = \lim_{T\to\infty} \inf (1/T) \sum_{k=0}^{T-1} r_k^i(p_k^1, p_k^2).$$

Now, let us assume that p_T^1 is a feasible price decision for firm 1 at the odd time period T. Then, $\Pi(T, p_T^1)$ stands for the set of all Markovian strategy combinations for horizon T in which player 1 is constrained to play p_T^1 at time period T. Similarly, $\Pi^*(T, p_T^1)$ is the set of T-long horizon constrained MPE strategy combinations to state p_T^1. Notice that, under the assumptions, by a backward induction argument one can see easily that $\Pi^*(T, p_T^1) \neq \varnothing$ and that the uniformly bounded reachability assumption holds (in fact, $L=2$).

3.2. Existence Results. The intuition for the next result lies in the fact that, under the uniformly bounded reachability assumption, a sequence of finite-horizon constrained MPE will encompass all possible deviations (and not just the constrained deviations) as the horizon diverges to infinity. Compactness of the strategy space ensures that every sequence of constrained MPE has a converging subsequence, and the limit strategy will be an MPE for the infinite-horizon game, by the above argument.

Lemma 3.1. Let s $=(s_0, s_1, s_2, \ldots)$ be the infinite feasible sequence of states defined in Assumption 2.2; let $\{\pi^T: \pi^T \in \Pi^*(T, s_T)\}_T$ be a sequence of finite-horizons constrained MPE to states s_T in the sequence s. If

$$\lim_{T\to\infty} \pi^T = \pi, \qquad \text{with respect to } \mathscr{L},$$

then $\pi \in \Pi^*$.

Proof. Let us first show that

$$P^i(h^{(\gamma_i,\pi_{-1})}(s_0)) \leq P^i(h^\pi(s_0)),$$

for any player i who would deviate by playing $\gamma_i \in \Pi$ from the initial state s_0. We recall that $h_T^{(\gamma_i,\pi_{-i})}(s_0)$ and $h_T^\pi(s_0)$ stand for the T-truncations of the histories induced by strategies (γ_i, π_{-i}) and π, respectively. By convergence in $\mathscr{L}$, there exists T^N such that, for any π^T with $T > \max\{T^N; N+L\}$, the play prescribed from the initial state by (γ_i^T, π_{-i}^T) and π^T coincide exactly with $h_T^{(\gamma_i,\pi_{-i})}(s_0)$ and $h_T^\pi(s_0)$, respectively, in the first $N<T$ periods.

Let us now consider the deviation for player i,

$$\bar{\gamma}_i^T = (\gamma_0^i, \gamma_1^i, \ldots, \gamma_N^i, a_{N+1}^i, \ldots, a_{N+L}^i, \pi_{N+L+1}^i, \ldots, \pi_T^i),$$

whereby we append, from period N, the actions $(a^i_{N+1}, \ldots, a^i_{N+L})$ as prescribed by the uniformly bounded reachability assumption to reach the sequence $\mathrm{s} = (s_0, s_1, s_2, \ldots)$ and, from period $N+L$, the actions prescribed by π^T_i. This deviation is such that $(\bar{\gamma}^T_i, \pi^T_{-i})$ reaches state s_T in $\mathrm{s} = (s_0, s_1, s_2, \ldots)$. Formally,

$$(\bar{\gamma}^T_i, \pi^T_{-i}) \in \Pi(T, s_T).$$

Hence, by hypothesis on π^T, we have

$$P^i_T(h_T^{(\bar{\gamma}^T_i, \pi^T_{-i})}(s_0)) \leq P^i_T(h_T^{\pi^T}(s_0)).$$

By cost boundedness and the choice of T^N and $\bar{\gamma}^t_i$, we have that total payoff accrued, up to period N, satisfies

$$\begin{aligned}
&P^i_N(h_T^{(\bar{\gamma}^T_i, \pi^T_{-i})}(s_0))/N \\
&\leq P^i_N(h_T^{\pi^T}(s_0))/N + (1/N) \sum_{k=N+1}^{N+L+1} [r^i_k(s_k, \pi^T(s_k)) - r^i_k(s_k; (a^i_k, \pi^T(s_k))] \\
&\leq P^i_N(h_T^{\pi^T}(s_0))/N + 2ML/N;
\end{aligned}$$

hence,

$$P^i_N(h_T^{(\gamma_i, \pi_{-i})}(s_0))/N \leq P^i_N(h_T^{\pi}(s_0))/N + 2ML/N.$$

Then, iterating on this construction, we have that

$$\begin{aligned}
P^i(h^{(\gamma_i, \pi_{-i})}(s_0)) &= \lim_{N\to\infty} \inf\, P^i_N(h_T^{(\gamma_i, \pi_{-i})}(s_0))/N \\
&\leq \lim_{N\to\infty} \inf\, P^i_N(h_T^{\pi}(s_0))/N \\
&= P^i(h^{\pi}(s_0)).
\end{aligned}$$

Thus, from the initial state, the proposed deviation is not profitable. For a deviation from any other state $s_k \in S$ with $0 < k$, we use the same argument. □

By a standard compactness argument the existence of MPE follows.

Theorem 3.1. Under Assumptions 2.1 and 2.2, there exists an MPE for the infinite-horizon undiscounted game.

Proof. By Assumption 2.1 and the alternating move structure, via backward induction one can always construct a sequence $\{\pi^T : \pi^T \in \Pi^*(T, s_T)\}_T$ of constrained MPE for the infinite feasible sequence of states $\mathrm{s} = (s_0, s_1, s_2, \ldots)$; see Aliprantis (Ref. 10). Then, by compactness of the strategy space, there exists a converging subsequence, say $\{\pi^{T_k}: \pi^{T_k} \in \Pi^*(T_k, s_{T_k})\}_k$ and

$$\lim_{k \to \infty} \pi^{T_k} = \pi, \qquad \text{with respect to } \mathscr{L}.$$

Finally, by Lemma 3.1, $\pi \in \Pi^*$. □

4. Conclusions

In this paper, we have presented a new approach to determine the existence of MPE in infinite-horizon nonstationary undiscounted dynamic games. The approach relies heavily on the structural properties of the game (the so called uniformly bounded reachability assumption). A new solution concept (the constrained MPE) for the finite-horizon game, captures the relevant features of an infinite-horizon MPE for a long enough horizon, under the aforementioned structural assumption.

An application to an asyncronous dynamic duopoly is presented. Of further research interest is the application of the techniques introduced here to linear-quadratic dynamic games [see Lau (Ref. 11)] where an analytical representation of the constrained MPE is possible.

5. Appendix: Topologies on the Set Π

Since our interest is to study the convergence of finite-horizon equilibrium strategies to infinite-horizon equilibrium strategies, it is very important to define carefully the relevant topologies on Π, and consequently the different notions of convergence which they induce. For a complete study, the interested reader is referred to Harris (Ref. 3).

We will adopt the convention that any finite-horizon strategy combination is trivially extended through any feasible choice of a continuation sequence of strategies, so that this extension is an element of Π.

First, we concentrate on a topology for H. Given $h = (a_0, a_1, a_2, \ldots)$ and $h' = (a'_0, a'_1, a'_2, \ldots)$, we define the metric $D: H \times H \to \mathscr{R}^+$ by

$$D(h, h') = \sup_k [\min\{d_k(a_t, a'_t), 1\}/k],$$

where d_k is any metric on $A_k^1 \times A_k^2 \times \cdots \times A_K^N$.

The metric $D(\cdot, \cdot)$ induces the product topology on H; see Munkres (Ref. 12, p. 123). As in Fudenberg and Levine (Ref. 2), we extend the notion of convergence to the strategy space Π [and implicitly to its subsets $\pi(T)$]

via the metric $\rho(\cdot, \cdot)$, defined as follows: For any $\pi, \pi' \in \Pi$,

$$\rho(\pi, \pi') = \sup_{k, s_k \in S} \left\{ D(h^{\pi}(s_k), h^{\pi'}(s_k)); \sup_{i, \gamma_i \in \Pi^i} [D(h^{(\gamma_i, \pi_{-i})}(s_k), h^{(\gamma_i, \pi'_{-i})}(s_k)] \right\}.$$

In words, strategies π and π' are close if, for every intermediate state $s_k \in S$, the play prescribed by them is close and the play prescribed after any deviation by any player is also close in the sense implied by the metric $D(\cdot, \cdot)$. We will denote by $\mathscr{S}$ the metric topology induced by $\rho(\cdot, \cdot)$. Notice that, when action sets are discrete, $\pi^T \to \pi'$ as $T \to \infty$ with respect to $\mathscr{S}$ if and only if, for all intermediate states, the early play prescribed by π^T and π' fully agree for large enough T.

References

1. Fudenberg, D., and Levine, D., *Subgame-Perfect Equilibria of Finite and Infinite-Horizon Games*, Journal of Economic Theory, Vol. 31, pp. 251–267, 1983.
2. Fudenberg, D., and Levine, D., *Limit Games and Limit Equilibria*, Journal of Economic Theory, Vol. 38, pp. 261–279, 1986.
3. Harris, C., *Existence and Characterization of Perfect Equilibrium in Games of Perfect Information*, Econometrica, Vol. 53, pp. 613–628, 1985.
4. Borgers, T., *Perfect Equilibrium Histories of Finite and Infinite-Horizon Games*, Journal of Economic Theory, Vol. 47, pp. 218–227, 1989.
5. Engwerda, J., *On the Open-Loop Nash Equilibrium in LQ Games*, Journal of Economic Dynamics and Control, Vol. 22, pp. 729–762, 1998.
6. Schochetman, I., and Smith, R. L., *Existence and Discovery of Average Optimal Solutions in Deterministic Infinite-Horizon Optimization*, Mathematics of Operations Research, Vol. 23, pp. 145–161, 1998.
7. Carlson, D., and Haurie, A., *A Turnpike Theory for Infinite-Horizon Open-Loop Differential Games with Decoupled Dynamics*, New Trends in Dynamic Games and Applications, Edited by G. Olsder, Annals of the International Society of Dynamic Games, Vol. 3, pp. 353–376, 1996.
8. Maskin, E., and Tirole, J., *A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles*, Econometrica, Vol. 56, pp. 571–599, 1988.
9. Fudenberg, D., and Tirole, J., *Game Theory*, MIT Press, Cambridge, Massachusetts, 1994.
10. Aliprantis, C., *On the Backward Induction Method*, Economics Letters, Vol. 64, pp. 125–131, 1999.
11. Lau, S. P., *A Note on the Markov Perfect Equilibrium of a Linear-Quadratic Alternating Move Game*, Economics Letters, Vol. 57, pp. 51–56, 1997.
12. Munkres, J., *Topology: A First Course*, Prentice Hall, Englewood Hills, New Jersey, 1975.