Fictitious Play and Best-Response Dynamics in Identical Interest and Zero-Sum Stochastic Games

ICML22

Lucas Baudin and Rida Laraki
Université Paris-Dauphine

july 2022

Stochastic Games

Definition (Stochastic Game)

$G = (S, I, (A^i)_{i \in I}, (r^i_s)_{i \in I, s \in S}, (P_s)_{s \in S})$

- S is a finite state space
- I is the set of players
- A^i is the action set of player i
- $P_s : A \to \Delta(S)$ is the transition probability
- $r^i_s : A \to \mathbb{R}$ is the stage reward map.

Playing Stochastic Games

How to play stochastic games?

Playing Stochastic Games

How to play stochastic games?

- initial state s_0

Playing Stochastic Games

How to play stochastic games?

- initial state s_0
- for all steps $n \in \mathbb{N}$, the system is in s_n:
  - every player i plays an action a_n^i
  - every player i receives $r_{s_n}^i(a_n)$
  - new state $s_{n+1} \sim P_{s_n}(a_n)$

Two Widely Studied Learning Procedures

Fictitious Play for Repeated Games

- Brown [1] Robinson [3]
- play a best response to the empirical average of past actions of other players

Two Widely Studied Learning Procedures

Fictitious Play for Repeated Games

- Brown [1] Robinson [3]
- play a best response to the empirical average of past actions of other players

- stochastic games: repeated games with a state variable

Q-Learning for One-Player Stochastic Games

- Watkins [5]
- estimates a table of state-action continuation values

Two Widely Studied Learning Procedures

Fictitious Play for Repeated Games

- Brown [1] Robinson [3]
- play a best response to the empirical average of past actions of other players

- stochastic games: repeated games with a state variable

Q-Learning for One-Player Stochastic Games

- Watkins [5]
- estimates a table of state-action continuation values

- **how can we combine these procedures for multiplayer stochastic games?**

Our paper: based on ideas of Q-Learning and Fictitious Play, we propose a definition of Fictitious Play for multiplayer stochastic games.

Learning in Stochastic Games

Our paper: based on ideas of Q-Learning and Fictitious Play, we propose a definition of Fictitious Play for multiplayer stochastic games.

Convergence results: If all players follow the procedure, then empirical actions converge to:

- the set of stationary Nash equilibria for ergodic, identical-interest stochastic games.
- the set of approximate Nash equilibria for ergodic, zero-sum stochastic games

Fictitious Play for Stochastic Games

Inspired by Leslie et al. [2]; Sayin et al. [4].

- Two sets of variables at step n:
  - $u_{s,n}$: estimate of continuation payoffs
  - $x_{s,n}$: empirical action of all players

Proof via Best-Response Dynamics

1. Define an analogous system in continuous time, best-response dynamics

Proof via Best-Response Dynamics

1. Define an analogous system in continuous time, best-response dynamics
2. Characterize the internally chain transitive sets (*i.e.* the convergence)

Proof via Best-Response Dynamics

1. Define an analogous system in continuous time, best-response dynamics
2. Characterize the internally chain transitive sets (*i.e.* the convergence)
3. Use stochastic approximation theorems to link this with the limit sets of the discrete time system

Conclusion

All details and proofs in the paper!

References

[1] George W Brown. Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1):374–376, 1951.

[2] David S. Leslie, Steven Perkins, and Zibo Xu. Best-response dynamics in zero-sum stochastic games. *Journal of Economic Theory*, 189:105095, September 2020. ISSN 00220531. doi: 10.1016/j.jet.2020.105095.

[3] Julia Robinson. An Iterative Method of Solving a Game. *The Annals of Mathematics*, 54(2):296, September 1951. ISSN 0003486X. doi: 10.2307/1969530.

[4] Muhammed O. Sayin, Francesca Parise, and Asu Ozdaglar. Fictitious play in zero-sum stochastic games. *arXiv:2010.04223 [cs, math]*, October 2020.

[5] C. J. C. H. Watkins. *Learning from Delayed Rewards*. PhD thesis, King's College, Oxford, 1989.