A Brief Tour of the Theory of Stochastic Games

Ivor Chen

University of Illinois at Urbana-Champaign
Department of Math and Department of Physics

November 16, 2022

Plan

Motivation

- In this course, we have studied the normal form of a 1-stage game.
- What if the game is played repeatedly? $\rightarrow$ Repeated games.
- What if the game at each stage is allowed to change? $\rightarrow$ Stochastic games.
- New questions in stochastic games:
  1. What is the pay-off? How do we calculate the pay-off?
  2. What are the strategies? How do we rigorously describe different classes of strategies?
  3. What are the relevant equilibria? When and how can we prove the existence of these equilibria?
  4. Stochastic games in extensive form?

Purposes and Applications

- Model interactions between the players and the environment
- Model interactions between the players
- As such, stochastic game theory has a wide range of applications:
  1. 1-player stochastic game = Markov decision process (reinforcement learning)
  2. Queueing theory
  3. Capital accumulation and resource extraction
  4. Profit maximization and risk minimization in financial market (Stochastic differential game)
- Stochastic game theory in an extensive form has a natural generalization to quantum information theory.

Mathematical Definition of Stochastic Game

Definition

A stochastic game G is given by a tuple $\Big(N, \mathcal{S}, (\mathcal{A}_i, q_i : \mathcal{S} \to \mathcal{A}_i, u_i : \mathcal{S} \times \prod_i \mathcal{A}_i \to \mathbb{R})_{i \in N}, p : \mathcal{S} \times \prod_{i \in N} \mathcal{A}_i \to \Delta(\mathcal{S})\Big)$ where

1. N is the set of players
2. $\mathcal{S}$ is the space of states of the game
3. $\mathcal{A}_i$ is the space of possible actions for the player i.
4. Given the state s of the game, $q_i(s)$ is the set of permissible actions for player i.
5. Given the state s of the game and the actions of all players $(a^{(i)})_{i \in N}$:
    1. The stage pay-off of player i is given by $u_i(s, (a^{(i)})_{i \in N})$.
    2. The subsequent state of the game is given by the distribution $p(s, (a^{(i)})_{i \in N})$.

How to Play a Stochastic Game

1. The game is initialized to a state s_1.
2. At each stage t $(t \geq 1)$:
    1. Each player i chooses an action from the permissible set of actions: $a_t^{(i)} \in q_i(s_t)$ where $q_i(s_t) \in \mathcal{A}_i$.
    2. Each player i receives stage pay-off $u_i(s_t, (a_t^{(i)})_{i \in N})$.
    3. The state of the game in the next stage s_{t+1} is sampled from the distribution $p(s_t, (a_t^{(i)})_{i \in N})$

Questions about the Definition

1. What information is available to each player at each stage?
   - Assume for now that **all past information (states and actions) and all players' current actions** are public information
   - This assumption is far too strong. Imperfect monitoring models weaken this assumption in various ways.
2. Why does the functions q_i, u_i, p only depend on the current state and action profile $(s_t, (a_t^{(i)})_i)$ but not the full history $(s_1, (a_1^{(i)})_i, s_2, \ldots, s_t, (a_t^{(i)})_i)$?
   - By a classical construction in probability theory (Kolmogorov and Rota), we can transform the state space so that the transformed game has functions q_i, u_i, p that only depend on the current state.

Plan

What is a Strategy for a Player?

- A strategy is a player's response to the available information. In other words, a strategy can be modeled by a function from the information set to the set of possible actions.
- What is the available information? This depends on the model assumptions.
- Assuming perfect monitoring, a strategy for player i at stage t can be described as a function $\sigma_i(s_t, (a_t^{(i)})_i, s_{t-1}, (a_{t-1}^{(i)})_i, \ldots, (a_1^{(i)})_i, s_1) \in \Delta(q_i(s_t))$. This called behavioral strategy.
- A pure strategy would map to the extremal points in $\Delta(q_i(s_t))$ at each stage t.
- **Kuhn's theorem**

Different Classes of Strategies

- Behavioral strategy is dependent on time.

Definition

A **Markov strategy** is a behavioral strategy where for each player i at each stage t, the function σ_i only depends on the final state of the game, i.e.: $\sigma_i^{Markov} : \mathcal{S} \times \mathbb{N} \to \Delta(\mathcal{A})$, where $supp(\sigma_i^{Markov}(s,t)) \subset q_i(s_t)$.

Definition

A **stationary strategy** is a Markov strategy where for each player i at each stage t, the function σ_i is independent of the time t, i.e.: $\sigma_i^{stationary} : \mathcal{S} \to \Delta(\mathcal{A})$, where $supp(\sigma_i^{stationary}(s)) \subset \cap_{t \geq 1} q_i(s_t)$.

Plan

Different Concepts of Pay-Off and Their Definitions

- In 1-stage games, pay-off functions are unambiguous. And their computations are straight-forward.
- In repeated games with potentially infinitely many stages, it is a priori unclear how to define a function to capture the infinite stream of pay-offs.
- Hard cut-off, discount, limit values.
- Fact: Given the initial state of the game s_1 and the behavioral strategy profile $(\sigma_i)_{i \in N}$, there exists an essentially unique probability distribution $\mathbb{P}_{\sigma, s_1}$ on $\mathcal{S} \times \prod_{i \in N} \mathcal{A}_i^{\mathbb{N}}$. (Path space construction)

Different Concepts of Pay-Off and Their Definitions

Definition

- For player i, a **T-stage pay-off** is given by:

$$u_i^T(s_1, \sigma) := \mathbb{E}_{s_1,\sigma}[\frac{1}{T} \sum_{1 \le i \le T} u_i(s_t, (a_t^{(j)})_{j \in N})]$$

- For player i, a η-**discounted pay-off** $(0 < \eta < 1)$ is given by:

$$u_i^\eta(s_1, \sigma) := \mathbb{E}_{s_1,\sigma}[(1-\eta) \sum_{t \ge 1} \eta^{t-1} u_i(s_t, (a_t^{(j)})_{j \in N})]$$

- For player i, a **limsup pay-off** is given by:

$$u_i^{limsup}(s_1, \sigma) := \mathbb{E}_{s_1,\sigma}[\limsup_{T \to \infty} \frac{1}{T} \sum_{1 \le i \le T} u_i(s_t, (a_t^{(j)})_{j \in N})]$$

Intuition behind the Definitions

- T-stage pay-off is suitable for a game that has only T stages.
- η-discounted pay-off allows the player to retain the memory of previous stage pay-offs. But the memory decays exponentially with time. The decay rate is given by $\log \eta$.
- Limsup pay-off allows the player to account for pay-offs at all stages. However, the pay-off from each single stage is insignificant compared with the overall averaged pay-off.

Plan

Different Concepts of Equilibria and Their Definitions

- Nash equilibrium depends on the pay-off function.

Definition

For $\epsilon \geq 0$, we have the following equilibria concepts:

- $\sigma_T^{opt} := (\sigma_i)_{i \in N}$ is a **T-stage ϵ-equilibrium** if for all i, initial state s_1 and strategy σ_i':

$$u_i^T(s_1, \sigma_T^{opt}) \geq u_i^T(s_1, \sigma_i', \sigma_{T,-i}^{opt}) - \epsilon$$

- $\sigma_\eta^{opt} := (\sigma_i)_{i \in N}$ is an η**-discounted ϵ-equilibrium** if for all i, initial state s_1 and strategy σ_i':

$$u_i^\eta(s_1, \sigma_\eta^{opt}) \geq u_i^\eta(s_1, \sigma_i', \sigma_{\eta,-i}^{opt}) - \epsilon$$

- $\sigma_{limsup}^{opt} := (\sigma_i)_{i \in N}$ is a **limsup ϵ-equilibrium** if for all i, initial state s_1 and strategy σ_i':

$$u_i^{limsup}(s_1, \sigma_{limsup}^{opt}) \geq u_i^{limsup}(s_1, \sigma_i', \sigma_{limsup,-i}^{opt}) - \epsilon$$

Uniform Equilibria

- Equilibria of stochastic games depends on the number of stages or the discount factor. What if we change these parameters? How will the corresponding equilibria change?

Definition

Given $\epsilon \geq 0$, a strategy profile $\sigma^{opt}_{uniform}$ is a **uniform ϵ-equilibrium** if there exists $T_0 \in \mathbb{N}$ and $\eta_0 \in (0, 1)$ such that

- For all $T \geq T_0$, $\sigma^{opt}_{uniform}$ is a T-stage ϵ-equilibrium
- For all $\eta \leq \eta_0$, $\sigma^{opt}_{uniform}$ is a η-discounted ϵ-equilibrium

A Taste of Known Results on the Existence of Equilibria

- For **finite state space and finite action spaces**, subgame perfect equilibrium in Markov strategies (**Markov perfect equilibrium (MPE)**) exists. (Reduction to 1-stage game).
- For **finite state space and finite action spaces**, *Takehashi* and *Fink* independently proved **the existence of discounted equilibrium in stationary strategies**. (Kakutani fixed point theorem). (The same idea dated back to *Shapely*.)

A Taste of Known Results on the Existence of Equilibria

- For **general state space and action spaces** with transition functions satisfying a **uniform continuity** condition, *Mertens and Parthasarathy* and later *Sloan* proved the existence of **subgame perfect equilibrium in general behavioral strategy**.
- For **finite state space and action spaces**, *Vieille* proved the existence of ϵ-**uniform equilibrium for any** $\epsilon > 0$ for two-player stochastic games.
- For **general state space and compact action spaces** with transition functions satisfying a L^1 **continuity** condition, *Nowak and Raghavan* and later *Nowak* proved the existence of **correlated equilibrium in stationary strategy** with public stochastic signal.

Plan

Complexity of Simple Stochastic Game

- *Condon* defined the **Simple Stochastic Game** based on directed graphs with two sinks where each vertex has out degree 2 and in degree 1.
- Apart from the two sinks, the graph has 3 classes of vertices: MIN, MAX, AVG.
- The game has 2 players: MIN and MAX. The player MIN can only make move when the game is at a vertex of the class MIN, and vice versa for MAX.
- If the game is at an AVG vertex, then it chooses one of its neighbors uniformly randomly.
- Each player pre-specifies the strategy. A strategy is a specification of how to go from a vertex to one of its neighbors.
- The game ends when it reaches either of the sink. The player MIN wins if it ends at sink 0. Otherwise, the player MAX wins.

Complexity of Simple Stochastic Game

Theorem (Condon '92)

The decision problem of deciding which player wins a simple stochastic game is both NP and coNP.

Algorithmic Problems of Stochastic Games

- The main problem in stochastic games is to compute the value of the game (finite stage, discounted, or uniform).
- 1-player stochastic game: value iteration, policy iteration.
- General results for multi-player stochastic games?

Plan

Quantum Games as a Natural Generalization of Stochastic Games in Extensive Form

- The framework of correlated stochastic game has a natural generalization to the quantum setting.
- More precisely, the players and the referee (the environment) uses *quantum correlation* instead of classical probability distributions to play and officiate the game. (**Quantum game**)
- Recently quantum games are used to prove (quantum) complexity theory results. (**MIP* = RE**)
- Quantum game theory is also used to prove results from pure mathematics (Connes' embedding problem in operator algebra theory).
-