Repeated Games

Game Theoretic Analysis

Kevin Leyton-Brown

University of British Columbia
Canada CIFAR AI Chair, Amii

James R. Wright

University of Alberta
Canada CIFAR AI Chair, Amii

Lecture Overview

Introduction

- Play the same normal-form game over and over
  - each round is called a **stage game**
- Questions we'll need to answer:
  - what will agents be able to **observe** about others' play?
  - how much will agents be able to **remember** about what has happened?
  - what is an agent's **utility** for the whole game?
- Some of these questions will have different answers for finitely- and infinitely-repeated games.

Finitely Repeated Games

- Everything is straightforward if we repeat a game a **finite number of times**
- we can write the whole thing as an extensive-form game with imperfect information
  - at each round players don't know what the others have done; afterwards they do
  - overall payoff function is additive: sum of payoffs in stage games

Example

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

$\Rightarrow$

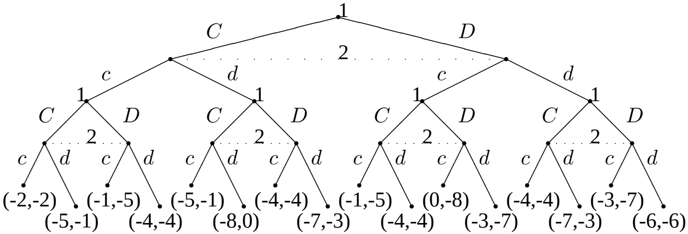
	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

Example

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

$\Rightarrow$

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

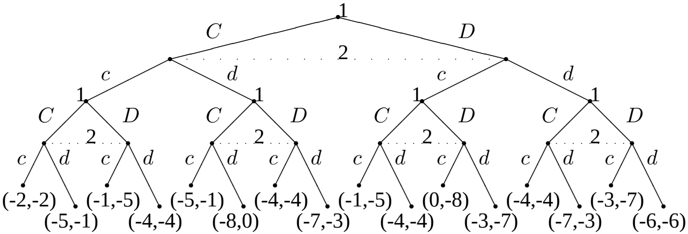


Example

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

$\Rightarrow$

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$



Play repeated prisoner's dilemma with one or more partners. Repeat the game five times.

Notes

- Observe that the strategy space is **much richer** than it was in the NF setting
- **Repeating a Nash strategy** in each stage game will be an equilibrium in behavioral strategies (called a stationary strategy)
- In general strategies adopted can depend on actions played so far
- We can apply **backward induction** in these games when the normal form game has a dominant strategy.

Lecture Overview

Repeated Games

Infinitely Repeated Games

Folk Theorem

Stochastic Games

Infinitely Repeated Games

- Consider an infinitely repeated game in extensive form:
  - an infinite tree!
- Thus, payoffs cannot be attached to terminal nodes, nor can they be defined as the sum of the payoffs in the stage games (which in general will be infinite).

Definition

Given an infinite sequence of payoffs $r_1, r_2, \ldots$ for player i, the **average reward** of i is

$$\lim_{k \to \infty} \sum_{j=1}^{k} \frac{r_j}{k}.$$

Discounted reward

Definition

Given an infinite sequence of payoffs $r_1, r_2, \ldots$ for player i and discount factor β with $0 < \beta < 1$, i's **future discounted reward** is

$$\sum_{j=1}^{\infty} \beta^j r_j.$$

- Interpreting the discount factor:
  1. the agent cares more about her well-being in the near term than in the long term
  2. the agent cares about the future just as much as the present, but with probability $1 - \beta$ the game will end in any given round.
- The analysis of the game is the same under both perspectives.

Strategy Space

- What is a pure strategy in an infinitely-repeated game?

Strategy Space

- What is a pure strategy in an infinitely-repeated game?
  - a choice of action at every decision point
  - here, that means an action at every stage game
  - ...which is an infinite number of actions!
- Some famous strategies (repeated PD):
  - **Tit-for-tat**: Start out cooperating. If the opponent defected, defect in the next round. Then go back to cooperation.
  - **Trigger**: Start out cooperating. If the opponent ever defects, defect forever.

Nash Equilibria

- With an infinite number of equilibria, what can we say about Nash equilibria?
  - we **won't** be able to construct an induced normal form and then appeal to Nash's theorem to say that an equilibrium exists
  - Nash's theorem only applies to finite games
- Furthermore, with an infinite number of strategies, there could be an infinite number of pure-strategy equilibria!
- It turns out we can characterize a set of **payoffs** that are achievable under equilibrium, without having to enumerate the equilibria.

Lecture Overview

Repeated Games

Infinitely Repeated Games

Folk Theorem

Stochastic Games

Definitions

- Consider any n-player game $G = (N, A, u)$ and any payoff vector $r = (r_1, r_2, \ldots, r_n)$.
- Let $v_i = \min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} u_i(s_{-i}, s_i)$.
  - i's **minmax value**: the amount of utility i can get when $-i$ play a minmax strategy against him

Definition

A payoff profile r is **enforceable** if $r_i \geq v_i$.

Definition

A payoff profile r is **feasible** if there exist rational, non-negative values α_a such that for all i, we can express r_i as $\sum_{a \in A} \alpha_a u_i(a)$, with $\sum_{a \in A} \alpha_a = 1$.

- a payoff profile is feasible if it is a convex, rational combination of the outcomes in G.

Folk Theorem

Theorem (Folk Theorem)

Consider any n-player game G and any payoff vector $(r_1, r_2, \ldots, r_n)$.

1. *If r is the payoff in any Nash equilibrium of the infinitely repeated G with average rewards, then for each player i, r_i is enforceable.*
2. *If r is both feasible and enforceable, then r is the payoff in some Nash equilibrium of the infinitely repeated G with average rewards.*

Folk Theorem (Part 1)

Payoff in Nash $\rightarrow$ enforceable

Part 1: Suppose r is not enforceable, i.e. $r_i < v_i$ for some i. Then consider a deviation of this player i to $b_i(s_{-i}(h))$ for any history h of the repeated game, where b_i is any best-response action in the stage game and $s_{-i}(h)$ is the equilibrium strategy of other players given the current history h. By definition of a minmax strategy, player i will receive a payoff of at least v_i in every stage game if he adopts this strategy, and so i's average reward is also at least v_i. Thus i cannot receive the payoff $r_i < v_i$ in any Nash equilibrium.

Folk Theorem (Part 2)

Feasible and enforceable $\rightarrow$ Nash

Part 2: Since r is a feasible payoff profile, we can write it as $r_i = \sum_{a \in A} \left(\frac{\beta_a}{\gamma}\right) u_i(a)$, where β_a and γ are non-negative integers.[1] Since the combination was convex, we have $\gamma = \sum_{a \in A} \beta_a$.

We're going to construct a strategy profile that will cycle through all outcomes $a \in A$ of G with cycles of length γ, each cycle repeating action a exactly β_a times. Let (a^t) be such a sequence of outcomes. Let's define a strategy s_i of player i to be a trigger version of playing (a^t): if nobody deviates, then s_i plays a_i^t in period t. However, if there was a period t' in which some player $j \neq i$ deviated, then s_i will play $(p_{-j})_i$, where (p_{-j}) is a solution to the minimization problem in the definition of v_j.

[1] Recall that α_a were required to be rational. So we can take γ to be their common denominator.

Folk Theorem (Part 2)

Feasible and enforceable $\rightarrow$ Nash

First observe that if everybody plays according to s_i, then, by construction, player i receives average payoff of r_i (look at averages over periods of length γ). Second, this strategy profile is a Nash equilibrium. Suppose everybody plays according to s_i, and player j deviates at some point. Then, forever after, player j will receive his $\min\max$ payoff $v_j \leq r_j$, rendering the deviation unprofitable.

Lecture Overview

Introduction

- What if we didn't always repeat back to the same stage game?
- A stochastic game is a **generalization** of repeated games
  - agents repeatedly play games from a set of normal-form games
  - the game played at any iteration depends on the previous game played and on the actions taken by all agents in that game
- A stochastic game is also a generalization of **Markov decision processes**
  - there are multiple players
  - one reward function for each agent
  - the state transition function and reward functions depend on the action choices of **both** players

Formal Definition

Definition

A **stochastic game** is a tuple (Q, N, A, P, R), where

- Q is a finite set of states,
- N is a finite set of n players,
- $A = A_1 \times \cdots \times A_n$, where A_i is a finite set of actions available to player i,
- $P : Q \times A \times Q \rightarrow [0, 1]$ is the transition probability function; $P(q, a, \hat{q})$ is the probability of transitioning from state q to state $\hat{q}$ after joint action a, and
- $R = r_1, \ldots, r_n$, where $r_i : Q \times A \rightarrow \mathbb{R}$ is a real-valued payoff function for player i.

Remarks

- This assumes strategy space is the same in all games
  - otherwise just more notation
- Again we can have average or discounted payoffs.
- Interesting special cases:
  - zero-sum stochastic game
  - single-controller stochastic game
    - transitions (but not payoffs) depend on only one agent

Strategies

- What is a pure strategy?

Strategies

- What is a pure strategy?
  - pick an action conditional on every possible history
  - of course, mixtures over these pure strategies are possible too!
- Some interesting restricted classes of strategies:
  - **behavioral strategy**: $s_i(h_t, a_{i_j})$ returns the probability of playing action a_{i_j} for history h_t.
    - the substantive assumption here is that mixing takes place at each history independently, not once at the beginning of the game
  - **Markov strategy**: s_i is a behavioral strategy in which $s_i(h_t, a_{i_j}) = s_i(h'_t, a_{i_j})$ if $q_t = q'_t$, where q_t and q'_t are the final states of h_t and h'_t, respectively.
    - for a given time t, the distribution over actions only depends on the current state
  - **stationary strategy**: s_i is a Markov strategy in which $s_i(h_{t_1}, a_{i_j}) = s_i(h'_{t_2}, a_{i_j})$ if $q_{t_1} = q'_{t_2}$, where q_{t_1} and q'_{t_2} are the final states of h_{t_1} and h'_{t_2}, respectively.
    - no dependence even on t

Equilibrium (discounted rewards)

- **Markov perfect equilibrium**:
  - a strategy profile consisting of only Markov strategies that is a Nash equilibrium regardless of the starting state
  - analogous to subgame-perfect equilibrium

Theorem

Every n-player, general sum, discounted reward stochastic game has a Markov perfect equilibrium.

Equilibrium (average rewards)

- **Irreducible stochastic game**:
  - every strategy profile gives rise to an irreducible Markov chain over the set of games
    - irreducible Markov chain: possible to get from every state to every other state
  - during the (infinite) execution of the stochastic game, each stage game is guaranteed to be played infinitely often—for any strategy profile
  - without this condition, limit of the mean payoffs may not be defined

Theorem

Every 2-player, general sum, average reward, irreducible stochastic game has a Nash equilibrium.

Folk Theorems for Stochastic Games

Theorem

For every 2-player, general sum, irreducible stochastic game, and every feasible outcome with a payoff vector r that provides to each player at least their minmax value, there exists a Nash equilibrium with a payoff vector r. This is true for games with average rewards, as well as games with large enough discount factors (i.e. with players that are sufficiently patient).