Game Theory: Repeated Games

Adaptive and Cooperative Algorithms (ECE 457A)

ECE, MME, and MSCI Departments,
University of Waterloo, ON, Canada

Course Instructor: Benyamin Ghojogh
Fall 2023

Repeated Games

Repeated Games

- In the **repeated games**, players make actions repeatedly in the same setting.
- Repeated games are different from one-shot unrepeated games.

Finitely Repeated Games

Finitely Repeated Games

- In the **finitely repeated games**, players make actions repeatedly in the same setting in the **finite** number of repetitions.
- There are two ways to find the equilibrium strategy in finitely repeated games:
  - solve it from beginning conditioning on previous history.
  - solve it **backwards** from the end of repetition.
    - ★ Soren **Kierkegaard**, the famous philosopher, has said: "Life can only be understood backwards, but it must be lived forwards."

Finitely Repeated Games: prisoner's dilemma

- Recall the prisoner's dilemma game:

prisoner 2

		cooperate	defect
prisoner 1	cooperate	-1, -1	-3, 0
	defect	0, -3	$\boxed{-2, -2}$

- Consider the finitely repeated prisoner's dilemma game. At every time slot, the two prisoners choose their actions simultaneously.
- The Nash equilibrium of the one-shot prisoner's dilemma game is (defect, defect).
- Backward analysis: in the last time slot, they both choose defect action. Or, in the one-to-last time slot, one or both choose defect action and then in the last time slot, the other one also chooses defect action.

Infinitely Repeated Games

Infinitely Repeated Games

- In the **infinitely repeated games**, players make actions repeatedly in the same setting in the **infinite** number of repetitions (forever).
- There are two ways to find the equilibrium strategy in infinitely repeated games:
  - **Grim strategy**
  - **Tit-for-Tat** (alternating approach)
- We will explain these with an example (prisoner's dilemma game) in the next slide.

Infinitely Repeated Games: prisoner's dilemma

- Recall the prisoner's dilemma game:

prisoner 2

prisoner 1	cooperate	defect
cooperate	$-1, -1$	$-3, 0$
defect	$0, -3$	$\boxed{-2, -2}$

- Consider the infinitely repeated prisoner's dilemma game. At every time slot, the two prisoners choose their actions simultaneously.
    - Grim strategy:
        - ★ Start with the cooperate action.
        - ★ Continue to choose the cooperate action unless another player chooses the defect action. Then, choose the defect action.
    - Tit-for-Tat:
        - ★ Start with the cooperate action.
        - ★ Thereafter, in the period n, choose the action that the other player chose in the period $(n-1)$.
        - ★ It is an alternating approach and its average payoff is less than the strategy (cooperate, cooperate).

The Folk Theorem

The Folk Theorem

- The formal statement of the **folk theorem** [1, 2]: In an **infinitely repeated** *n*-person game with **finite action sets** at each repetition, any profile of actions observed in any finite number of repetitions is the unique outcome of some subgame perfect equilibrium given:
  - **Condition 1**: The **rate** of time preference is zero, or positive and sufficiently small.
  - **Condition 2**: The probability that the **game ends** at any repetition is zero, or positive and sufficiently small.
  - **Condition 3**: The set of payoff profiles that strictly Pareto dominate the **minimax** payoff profiles in the mixed extension of the one-shot game is *n*-dimensional.
- What the folk theorem talks about is:
  - If an infinite time remains in a game, then there is always at least one player that will punish another player in order to guarantee a better future, even if the punishment hurts both parties.
  - Any finite time period is insignificant relative to eternity.
- We talk about the three conditions one by one in the next slides.

Condition 1: Discounting

- With **discounting**, the present gain from defecting is weighted more heavily than future gains from defecting (in the prisoner's dilemma game).
- If the **discount rate** is **very high**, the game almost becomes one-shot because the players will all choose defecting in the first time slot.
- If the **discount rate** is **zero or very low**, the game becomes infinitely repeated game.

prisoner 2

		cooperate	defect
prisoner 1	cooperate	-1, -1	-3, 0
	defect	0, -3	-2, -2

Condition 2: Probability of Game Ending

- Let the probability of ending the game in the repetitions be denoted by $\theta \geq 0$.
- If θ is **large**, the game becomes finitely repeated game.
- If θ is **very large (very close to one)**, the game almost becomes one-shot because it will most probably end after the first time slot.
- If θ is **zero or very low**, the game becomes infinitely repeated game.

Condition 3: Minimax

- **Minimax** strategy: the strategy in which:
  - All the other players pick strategies solely to **punish** player i. In other words, they **gang up** on the player i.
  - Player i **protects itself** the best it can.
- The set of strategies s^*_{-i} is a set of $(n-1)$ minimax strategies chosen by all the players except player i to keep the payoff of the player i as low as possible, no matter how it responds. In other words, s^*_{-i} solves:

$$\underset{s_{-i}}{\text{minimize}} \; \underset{s_i}{\text{maximize}} \; \pi_i(s_i, s_{-i}). \tag{1}$$

- The payoff the player i, obtained from the above equation, is called the **minimax payoff**, **minimax value**, or **security value**.

Condition 3: Minimax

- **Maximin** strategy: the strategy in which An offender trying to protect itself from punishment.
  - In the **minimax** strategy, the player i maximizes its payoff and the others minimize that maximum payoff of player i.
  - In the **maximin** strategy, the other players minimize that payoff of player i and then the player i maximizes its payoff which was minimized by others.
- The strategy s_i^* is a maximin strategy for player i if, given that the other players pick strategies to make the payoff of player i as low as possible, s_i^* gives the player i the highest possible payoff. In other words, s_i^* solves:

$$\star \quad \underset{s_i}{\text{maximize}} \; \boxed{\underset{s_{-i}}{\text{minimize}} \quad \pi_i(s_i, s_{-i}).} \tag{2}$$

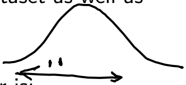
worst case scenario for player i

- In **maximin** strategy, each player protects itself from the worst harm possible made by others.
- The maximin and minimax strategies have opposite order of optimization.
- **Minimax Theorem** [2]: minimax equilibrium exists in pure or mixed strategies for every **two-person zero-sum** game and it is **identical** to the maximin equilibrium.

$$\min_x \max_y \pi$$
$$\max_y \min_x \pi$$

Example of Minimax in Machine Learning: GAN

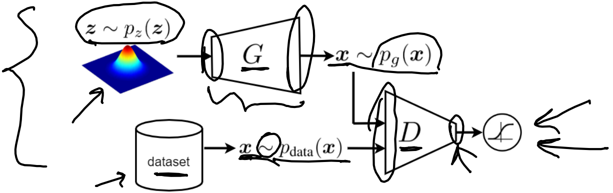
- An example of minimax game strategy in machine learning is Generative Adversarial ~~Learning~~ Network (GAN), proposed in 2014 [3].
- We denote the probability distributions of dataset and noise by $p_{\text{data}}(\boldsymbol{x})$ and $p_z(\boldsymbol{z})$, respectively.
- As the figure shows, the **discriminator** is trained by real points from dataset as well as generated points from the generator.
- The discriminator and generator are **trained simultaneously**.



- The optimization **loss function** for both the discriminator and generator is:

$$\min_{G} \max_{D} V(D, G) := \mathbb{E}_{\boldsymbol{x} \sim p_{\text{data}}(\boldsymbol{x})}\Big[\log\big(D(\boldsymbol{x})\big)\Big] + \mathbb{E}_{\boldsymbol{z} \sim p_z(\boldsymbol{z})}\Big[\log\Big(1 - D\big(G(\boldsymbol{z})\big)\Big)\Big], \qquad (3)$$

where $\mathbb{E}[.]$ denotes the expectation operator and the loss function $V(D, G)$ is also called the **value function** of the game.



- For more information on GAN, see our tutorial paper: [4]

Acknowledgment

- Some slides of this slide deck are inspired by teachings of Prof. Stanko Dimitrov at the University of Waterloo, Department of Management Science and Engineering.
- Some slides of this slide deck are based on the following book: Eric Rasmusen, "Games and Information: An Introduction to Game Theory", 4th Edition, 2007, [2] https://www.rasmusen.org/GI/download.htm
- For more information on GAN, see our tutorial paper: [4]

References

[1] J. W. Friedman, "A non-cooperative equilibrium for supergames," *The Review of Economic Studies*, vol. 38, no. 1, pp. 1–12, 1971.

[2] E. Rasmusen, *Games and information: An introduction to game theory*. Wiley-Blackwell, 4 ed., 2007.

[3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, vol. 27, 2014.

[4] B. Ghojogh, A. Ghodsi, F. Karray, and M. Crowley, "Generative adversarial networks and adversarial autoencoders: Tutorial and survey," *arXiv preprint arXiv:2111.13282*, 2021.