CSC2541: Neural Net Training Dynamics

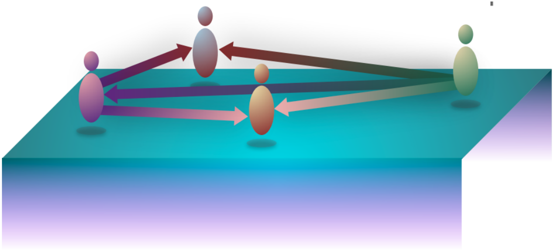
Tutorial 9 - Game Theory

Juhan Bae

University of Toronto, Winter 2022

March 16, 2022

Overview



- Future lectures: Differentiable games & Bilevel optimization
- Agenda for today's tutorial:
  - Introduction to Game Theory (Game Theory 101)
    - Types of games
    - Solution concepts
  - Generative Adversarial Network (GAN)

Game Theory

The study of mathematical models of strategic **interaction** among **rational** decision-makers

- **Game:** Situation in which multiple decision makers (**players**) make choices which influence each other's welfare (**utility**, **loss**)
- Real world examples: Competition between firms, voting strategy, soccer penalty kicks, etc
- ML examples: Generative Adversarial Networks (GANs), multi-agent RL, PCA, off-policy evaluation, robust optimization, etc

Components of a Game

- **Players:** Participants of the game, each may be an individual, organization, a machine, or an algorithm, etc.
- **Strategies:** Actions available to each player
- **Outcome:** The profile of player strategies
- **Payoffs:** A function mapping an outcome to a utility for each player

Types of Games

- Symmetric vs. Asymmetric
- Perfect vs. Imperfect information
- Cooperative vs. Non-Cooperative
- Simultaneous vs. Sequential
  - Simultaneous game: Players choose actions without knowing what other players are choosing.
  - Sequential game: Players choose actions before or after other players. Later players have information about previous choices.
- Zero-Sum vs. Non-Zero-Sum
  - Zero-Sum game: Each player's gain or loss balanced by loss or gain of the others. Total amount of utility remains unchanged.
  - Non-Zero-Sum game: Players can increase or decrease total amount of utility. One player's gain doesn't necessarily come at someone else's expense.

Example 1: Prisoner's Dilemma

- **Set-up:**
  - Two members A, B of a criminal gang are arrested
  - They are questioned in two separate rooms
  - No communications between them

	Coop	Defect
Coop	-1 -1	0 -3
Defect	-3 0	-2 -2

- **Question:** How should each prisoner act?

Example 1: Prisoner's Dilemma

	Coop	Defect
Coop	-1 -1	0 -3
Defect	-3 0	-2 -2

- What A thinks:
  - If B is going to cooperate...
    - Better for me to betray (my reward: 0)
    - Than for me to cooperate (my reward: -1)
  - If B is going to betray...
    - Better for me to betray (my reward: -2)
    - Than for me to cooperate (my reward -3)

Example 2: Battle of the Sexes

- **Set-up:**
  - A man and a woman are deciding on how to spend their evening
  - Two possibilities: going to a football game or going to a movie theatre
  - The man prefers football and the woman prefers movie

M \ W	Football	Movie
Football	1 2	0 0
Movie	0 0	2 1

Example 3: Rock, Paper, Scissor

	Rock	Paper	Scissors
Rock	0 0	+1 -1	-1 +1
Paper	-1 +1	0 0	+1 -1
Scissors	+1 -1	-1 +1	0 0

Simultaneous Games: Normal-Form Representation

- A set of players $[n] = \{1, ..., n\}$
- Player i takes action $a_i \in A_i$
- An outcome is the action profile $a = (a_1, ..., a_n)$
  - Note: as a convention, $a_{-i} = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n)$ represents all actions excluding a_i
- Player i receives payoff $u_i(a)$ for any outcome $a \in \Pi_{i=1}^{n} A_i$

Back to Example 1: Prisoner's Dilemma

	Coop	Defect
Coop	-1 -1	0 -3
Defect	-3 0	-2 -2

- 2 Players
- $A_i = \{\text{Coop, Defect}\}$ for $i = 1, 2$
- One outcome can be $a = (\text{Coop}, \text{Coop})$
- $u_1(a)$ and $u_2(a)$ are predefined.
- Players take actions simultaneously

Back to Example 3: Rock, Paper, Scissor

	Rock	Paper	Scissors
Rock	0 0	+1 -1	-1 +1
Paper	-1 +1	0 0	+1 -1
Scissors	+1 -1	-1 +1	0 0

Strategies in Normal Form Games

- Pure strategy
  - Choose an action to play $a_i \in A_i$ (e.g. defect)

- Mixed Strategy
  - Choose a probability distribution over actions $a_i \sim s_i$
  - Randomize over pure strategies (e.g. cooperate with probability 0.4 and defect with probability 0.6)

Dominant Strategy

- An action a_i is a **dominant strategy** (dominant action) for a player i if a_i is better than any other actions $a'_i \in \mathcal{A}_i$ regardless of what actions other player take
- More formally,

$$u_i(a_i, a_{-i}) \geq u_i(a'_i, a_{-i}), \forall a'_i \neq a_i \text{ and } \forall a_{-i}$$

- Dominant strategies do not always exist

Back to Example 1: Prisoner's Dilemma

	Coop	Defect
Coop	-1 -1	0 -3
Defect	-3 0	-2 -2

- What is a dominant strategy for each player?

Nash Equilibrium

- An outcome a^* is a **pure Nash equilibrium** if no player has incentive to deviate unilaterally
- More formally,

$$u_i(a_i^*, a_{-i}^*) \geq u_i(a_i, a_{-i}^*), \forall a_i \tag{1}$$

- If all players are playing the dominant strategy, simultaneously, they form an equilibrium

Back to Example 1: Prisoner's Dilemma

	Coop	Defect
Coop	-1 -1	0 -3
Defect	-3 0	-2 -2

- What is the pure Nash equilibrium?

Back to Example 2: Battle of the Sexes

M \ W	Football	Movie
Football	1 2	0 0
Movie	0 0	2 1

- What is the pure Nash equilibrium?

Back to Example 3: Rock, Paper, Scissor

	Rock	Paper	Scissors
Rock	0 0	+1 -1	-1 +1
Paper	-1 +1	0 0	+1 -1
Scissors	+1 -1	-1 +1	0 0

- What is the pure Nash equilibrium?

Best Responses

- A mixed strategy s_i^* is a **best-response** to a strategy profile s_{-i} of other players if $u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i}), \forall s_i \in S_i$
- Note: There always exist a pure best response

Nash Equilibrium

- A mixed strategy profile $s^* = (s_1^*, s_2^*, ..., s_n^*)$ is a **Nash equilibrium** if

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*), \forall s_i \in S_i, \forall i \in [n] \tag{2}$$

- For any player i, s_i^* is a best response to s_{-i}^*
- Every finite game (e.g. finite players and actions) admits at least one mixed Nash equilibrium

Back to Example 3: Rock, Paper, Scissor

	Rock	Paper	Scissors
Rock	0 / 0	+1 / -1	-1 / +1
Paper	-1 / +1	0 / 0	+1 / -1
Scissors	+1 / -1	-1 / +1	0 / 0

- What is the pure Nash equilibrium?
- What is the mixed Nash equilibrium?

Recall: Normal-Form Representation

- A set of players $[n] = \{1, ..., n\}$
- Player i takes action $a_i \in A_i$
- An outcome is the action profile $a = (a_1, ..., a_n)$
  - Note: as a convention, $a_{-i} = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n)$ represents all actions excluding a_i
- Player i receives payoff $u_i(a)$ for any outcome $a \in \Pi_{i=1}^{n} A_i$
- A mixed strategy profile s^* is a Nash equilibrium if for any i, s_i^* is a best response to s_{-i}^*

Nash Equilibrium is Not the Only Solution Concept

- Nash Equilibrium makes the following key assumptions:
  - Players move **simultaneously**
  - Players take actions **independently**

- A sequential move (sequential game) results in different player behaviors.
  - The corresponding game is called Stackelberg game and its equilibrium is called Stackelberg equilibrium
  - Bilevel optimization was first realized from Stackelberg game

Zero-Sum Games

- Zero-Sum game is a special case where total utility sum is constant in every outcome
- WLOG, the sum of total utility equals to 0 for all actions
- More formally, a two player zero-sum game is any two player game such that for every $a \in A_1 \times A_2, u_1(a) = -u_2(a)$

Back to Example 3: Rock, Paper, Scissor

	Rock	Paper	Scissors
Rock	0, 0	-1, +1	+1, -1
Paper	+1, -1	0, 0	-1, +1
Scissors	-1, +1	+1, -1	0, 0

- Is this a zero-sum game?

Back to Example 1: Prisoner's Dilemma

	Coop	Defect
Coop	-1 -1	0 -3
Defect	-3 0	-2 -2

- Is this a zero-sum game?

(Rough Idea) Maximin, Minimax Strategy

- We use the notation $[n] = \{1, ..., n\}$ and $\Delta[n]$ to denote the set of probability distribution over $[n]$
- For an $n \times m$ matrix U (this can be interpreted as the payoff matrix in a two player zero sum game):

$$\max\min(U) = \max_{p \in \Delta[n]} \min_{y \in [m]} \sum_{i=1}^{n} p_i U(i, y)$$

$$\min\max(U) = \min_{q \in \Delta[m]} \max_{x \in [n]} \sum_{i=1}^{m} q_j U(x, j)$$

- If U is a zero-sum game, then $\max\min(U)$ represents the payoff that first player can guarantee if it goes first and $\min\max(U)$ represents the payoff that it can guarantee if second player goes first

(Rough Idea) The Minimax Theorem

- For any game U, we get the inequality:

$$\min\max(U) \geq \max\min(U)$$

- It turns out that if U is a zero sum game:

$$\min\max(U) = \max\min(U)$$

- In 2-player zero-sum games, (s_1^*, s_2^*) is a Nash equilibrium if and only if s_1^* and s_2^* are the same maximin and minimax strategy, respectively.

Generative Adversarial Networks

Generative Modeling

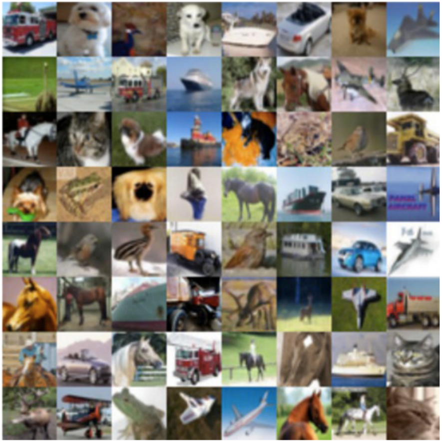
- In **generative modeling**, we would like to train a network that models a distribution, such as a distribution over images.
- One way to judge the quality of the model is to sample from it.
- This field has seen rapid progress:

2009

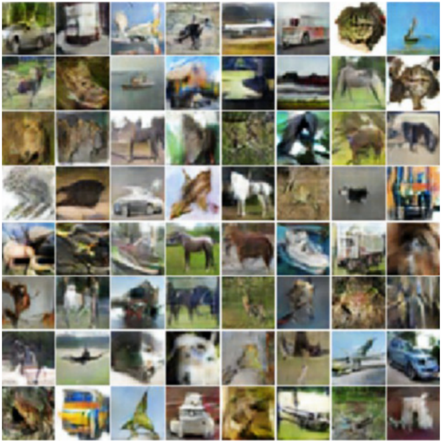
2015

2018

Applications



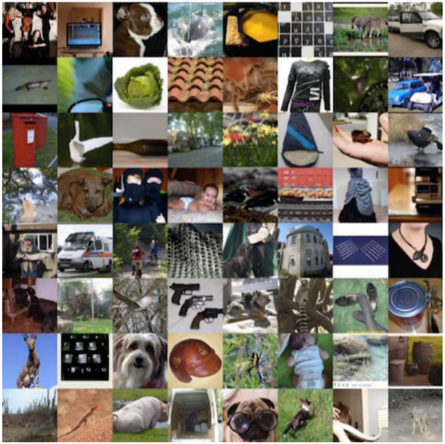
Training Data



Samples

Applications

ImageNet:

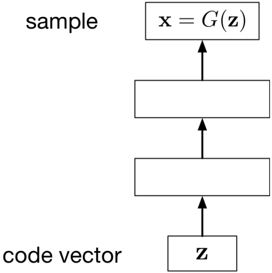


Applications

Applications

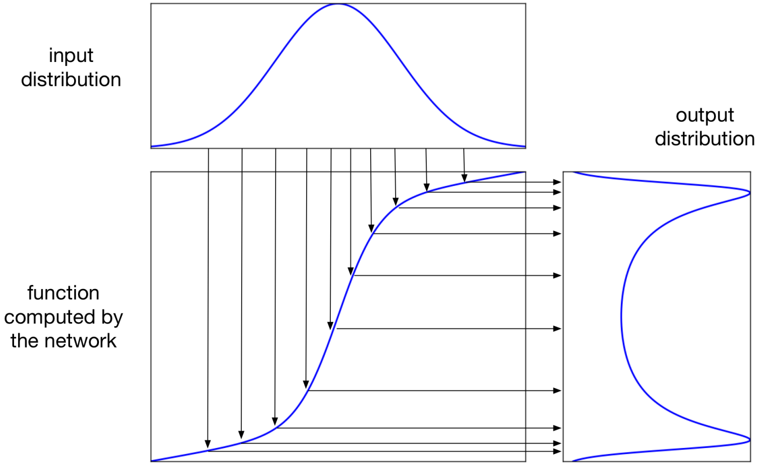
Implicit Generative Models

- **Implicit generative models** implicitly define a probability distribution
- Start by sampling the **code vector** $\mathbf{z}$ from a fixed, simple distribution (e.g. spherical Gaussian)
- The **generator network** computes a differentiable function G mapping $\mathbf{z}$ to an $\mathbf{x}$ in data space

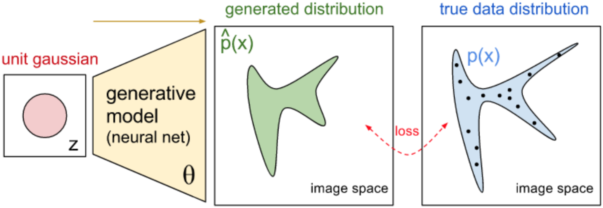


Implicit Generative Models

A 1-dimensional example:



Implicit Generative Models

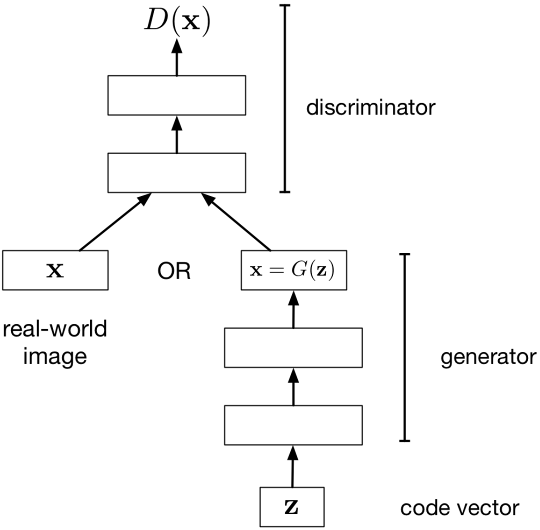


https://blog.openai.com/generative-models/

Generative Adversarial Networks

- The advantage of implicit generative models: if you have some criterion for evaluating the quality of samples, then you can compute its gradient with respect to the network parameters, and update the network's parameters to make the sample a little better
- The idea behind **Generative Adversarial Networks (GANs)**: train two different networks
  - The **generator network** tries to produce realistic-looking samples
  - The **discriminator network** tries to figure out whether an image came from the training set or the generator network
- The generator network tries to fool the discriminator network

Generative Adversarial Networks



Generative Adversarial Networks

- Let D denote the discriminator's predicted probability of being data
- Discriminator's cost function: cross-entropy loss for task of classifying real vs. fake images

$$\mathcal{J}_D = \mathbb{E}_{\mathbf{x}\sim\mathcal{D}}[-\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z}}[-\log(1 - D(G(\mathbf{z})))]$$

- One possible cost function for the generator: the opposite of the discriminator's

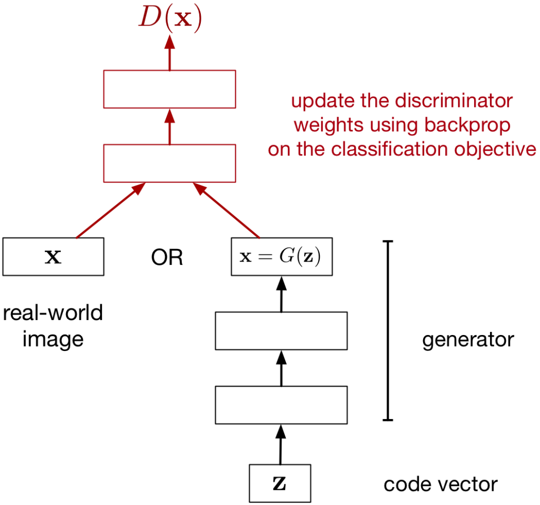
$$\begin{aligned}\mathcal{J}_G &= -\mathcal{J}_D \\ &= \text{const} + \mathbb{E}_{\mathbf{z}}[\log(1 - D(G(\mathbf{z})))]\end{aligned}$$

- This is called the minimax formulation, since the generator and discriminator are playing a zero-sum game against each other:

$$\max_G \min_D \mathcal{J}_D$$

Generative Adversarial Networks

Updating the discriminator:



Generative Adversarial Networks

Updating the generator: