Announcements

- Minbiao's office hour will be changed to Thursday 1-2 pm, starting from next week, at Rice Hall 442

CS6501: Topics in Learning and Game Theory
(Fall 2019)

Introduction to Game Theory (II)

Instructor: Haifeng Xu

Outline

- Correlated and Coarse Correlated Equilibrium
- Zero-Sum Games
- GANs and Equilibrium Analysis

Recap: Normal-Form Games

- n players, denoted by set $[n] = \{1, \cdots, n\}$
- Player i takes action $a_i \in A_i$
- An outcome is the action profile $a = (a_1, \cdots, a_n)$
  - As a convention, $a_{-i} = (a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_n)$ denotes all actions excluding a_i
- Player i receives payoff $u_i(a)$ for any outcome $a \in \Pi_{i=1}^n A_i$
  - $u_i(a) = u_i(a_i, a_{-i})$ depends on other players' actions
- $\{A_i, u_i\}_{i \in [n]}$ are public knowledge

A mixed strategy profile $x^* = (x_1^*, \cdots, x_n^*)$ is a **Nash equilibrium (NE)** if for any i, x_i^* is a best response to x_{-i}^*.

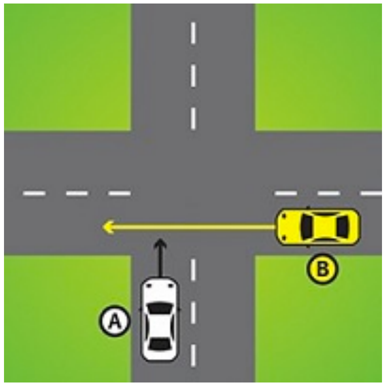
NE Is Not the Only Solution Concept

- NE rests on two key assumptions
    1. Players move simultaneously (so they cannot see others' strategies before the move)
    2. Players take actions independently
- Last lecture: sequential move results in different player behaviors
    - The corresponding game is called Stackelberg game and its equilibrium is called Strong Stackelberg equilibrium

Today: we study what happens if players do not take actions independently but instead are "coordinated" by a central mediator

- This results in the study of correlated equilibrium

An Illustrative Example



		B	
		STOP	GO
A	STOP	(-3, -2)	(-3, 0)
	GO	(0, -2)	(-100, -100)

The Traffic Light Game

Well, we did not see many crushes in reality… Why?

- There is a mediator – the traffic light – that coordinates cars' moves
- For example, recommend (GO, STOP) for (A,B) with probability 3/5 and (STOP, GO) for (A,B) with probability 2/5
  - GO = green light, STOP = red light
  - Following the recommendation is a best response for each player
  - It turns out that this recommendation policy results in equal player utility – $6/5$ and thus is "fair"

This is exactly how traffic lights are designed!

Correlated Equilibrium (CE)

- A (randomized) recommendation policy π assigns probability $\pi(a)$ for each action profile $a \in A = \Pi_{i\in[n]} A_i$
  - A mediator first samples $a \sim \pi$, then recommends a_i to i *privately*
- Upon receiving a recommendation a_i, player i's expected utility is

$$\frac{1}{c} \sum_{a_{-i}\in A_{-i}} u_i(a_i, a_{-i}) \cdot \pi(a_i, a_{-i})$$

  - c is a normalization term that equals the probability a_i is recommended

A recommendation policy π is a **correlated equilibrium** if

$$\sum_{a_{-i}} u_i(a_i, a_{-i}) \cdot \pi(a_i, a_{-i}) \geq \sum_{a_{-i}} u_i(a'_i, a_{-i}) \cdot \pi(a_i, a_{-i}), \forall\, a'_i \in A_i, \forall i \in [n].$$

- That is, any recommended action to any player is a best response
  - CE makes incentive compatible action recommendations
- Assumed π is public knowledge so every player can calculate her utility

Basic Facts about Correlated Equilibrium

Fact. Any Nash equilibrium is also a correlated equilibrium.

- True by definition. Nash equilibrium can be viewed as independent action recommendation
- As a corollary, correlated equilibrium always exists

Fact. The set of correlated equilibria forms a convex set.

- In fact, distributions π satisfies a set of linear constraints

$$\sum_{a_{-i}} u_i(a_i, a_{-i}) \cdot \pi(a_i, a_{-i}) \geq \sum_{a_{-i}} u_i({a'}_i, a_{-i}) \cdot \pi(a_i, a_{-i}), \forall\, {a'}_i \in A_i, \forall i \in [n].$$

Basic Facts about Correlated Equilibrium

Fact. Any Nash equilibrium is also a correlated equilibrium.

- True by definition. Nash equilibrium can be viewed as independent action recommendation
- As a corollary, correlated equilibrium always exists

Fact. The set of correlated equilibria forms a convex set.

- In fact, distributions π satisfies a set of linear constraints
- This is nice because that allows us to optimize over all CEs
- Not true for Nash equilibrium

Coarse Correlated Equilibrium (CCE)

- A weaker notion of correlated equilibrium
- Also a recommendation policy π, but only requires that any player does not have incentives to opting out of our recommendations

A recommendation policy π is a **coarse correlated equilibrium** if

$$\sum_{a \in A} u_i(a) \cdot \pi(a) \geq \sum_{a \in A} u_i(a'_i, a_{-i}) \cdot \pi(a), \forall\, a'_i \in A_i, \forall i \in [n].$$

That is, for any player i, following π's recommendations is better than opting out of the recommendation and "acting on his own".

Compare to correlated equilibrium condition:

$$\sum_{a_{-i}} u_i(a_i, a_{-i}) \cdot \pi(a_i, a_{-i}) \geq \sum_{a_{-i}} u_i(a'_i, a_{-i}) \cdot \pi(a_i, a_{-i}), \forall\, a'_i \in A_i, \forall i \in [n].$$

Coarse Correlated Equilibrium (CCE)

- A weaker notion of correlated equilibrium
- Also a recommendation policy π, but only requires that any player does not have incentives to opting out of our recommendations

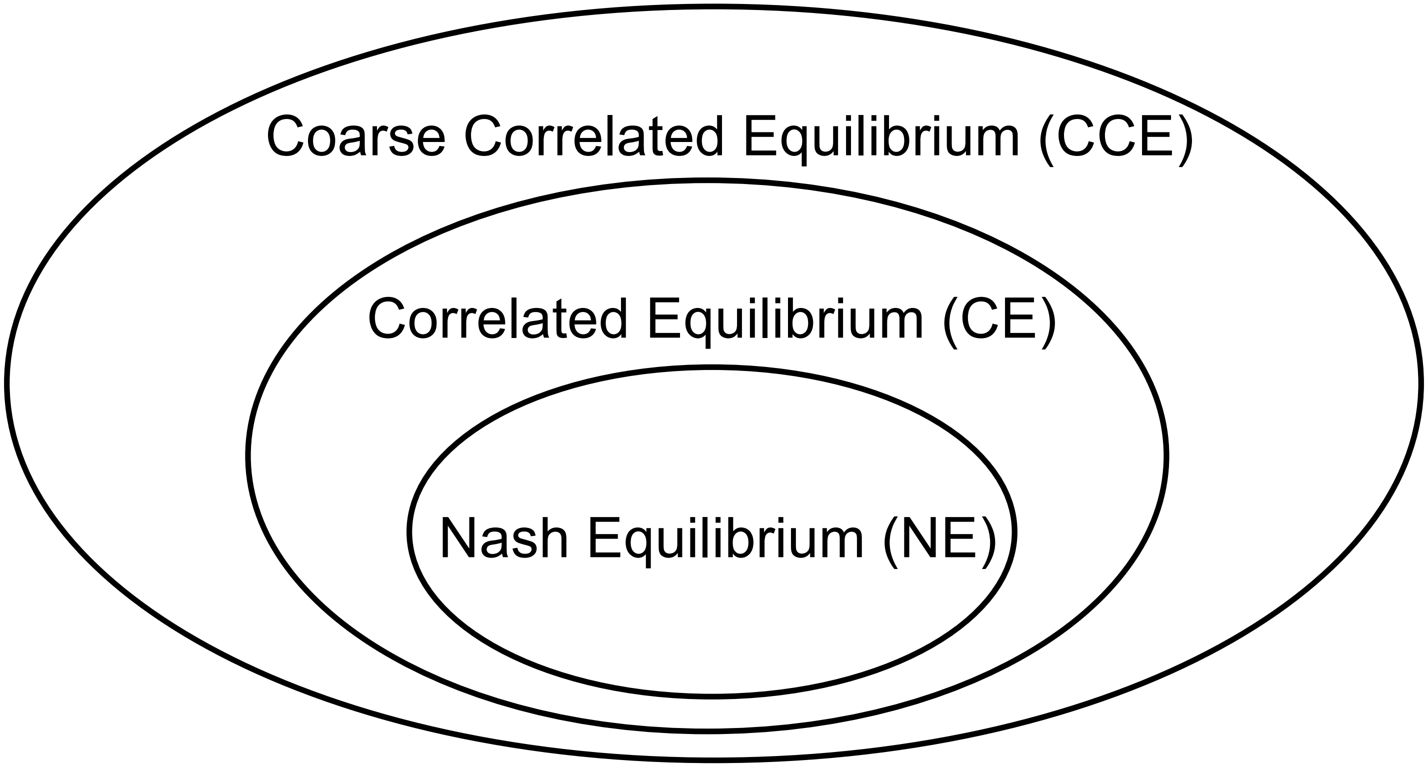
A recommendation policy π is a **coarse correlated equilibrium** if

$$\sum_{a \in A} u_i(a) \cdot \pi(a) \geq \sum_{a \in A} u_i(a'_i, a_{-i}) \cdot \pi(a), \forall a'_i \in A_i, \forall i \in [n].$$

That is, for any player i, following π's recommendations is better than opting out of the recommendation and "acting on his own".

Fact. Any correlated equilibrium is a coarse correlated equilibrium.

The Equilibrium Hierarchy



There are other equilibrium concepts, but NE and CE are most often used. CCE is not used that often.

Outline

- Correlated and Coarse Correlated Equilibrium
- Zero-Sum Games
- GANs and Equilibrium Analysis

Zero-Sum Games

- Two players: player 1 action $i \in [m] = \{1, \cdots, m\}$, player 2 action $j \in [n]$
- The game is **zero-sum** if $u_1(i,j) + u_2(i,j) = 0, \ \forall i \in [m], \ j \in [n]$
  - Models the strictly competitive scenarios
  - "Zero-sum" almost always mean "2-player zero-sum" games
  - n-player games can also be zero-sum, but not particularly interesting
- Let $u_1(x,y) = \sum_{i \in [m], j \in [n]} u_1(i,j) x_i y_j$ for any $x \in \Delta_m, y \in \Delta_n$
- (x^*, y^*) is a NE for the zero-sum game if: (1) $u_1(x^*, y^*) \geq u_1(i, y^*)$ for any $i \in [m]$; (2) $u_1(x^*, y^*) \leq u_1(x^*, j)$ for any j $\in [m]$
  - Condition $u_1(x^*, y^*) \leq u_1(x^*, j) \Leftrightarrow u_2(x^*, y^*) \geq u_2(x^*, j)$
  - We can "forget" u_2; Instead think of player 2 as minimizing player 1's utility

Maximin and Minimax Strategy

- Previous observations motivate the following definitions

Definition. $x^* \in \Delta_m$ is a maximin strategy of player 1 if it solves

$$\max_{x \in \Delta_m} \min_{j \in [n]} u_1(x, j).$$

The corresponding utility value is called maximin value of the game.

Remarks:

- x^* is player 1's best action if he was to move first

Maximin and Minimax Strategy

- Previous observations motivate the following definitions

Definition. $x^* \in \Delta_m$ is a maximin strategy of player 1 if it solves

$$\max_{x \in \Delta_m} \min_{j \in [n]} u_1(x, j).$$

The corresponding utility value is called maximin value of the game.

Definition. $y^* \in \Delta_n$ is a minimax strategy of player 2 if it solves

$$\min_{y \in \Delta_n} \max_{i \in [m]} u_1(i, y).$$

The corresponding utility value is called minimax value of the game.

Remark: y^* is player 2's best action if he was to move first

Duality of Maximin and Minimax

Fact. $\max_{x\in\Delta_m}\min_{j\in[n]} u_1(x,j) \le \min_{y\in\Delta_n}\max_{i\in[m]} u_1(i,y)$.

That is, moving first is no better.

- Let $y^* = \arg\min_{y\in\Delta_n}\max_{i\in[m]} u_1(i,y)$, so

$$\min_{y\in\Delta_n}\max_{i\in[m]} u_1(i,y) = \max_{i\in[m]} u_1(i,y^*)$$

- We have

$$\max_{x\in\Delta_m}\min_{j\in[n]} u_1(x,j) \le \max_{x\in\Delta_m} u_1(x,y^*) = \max_{i\in[m]} u_1(i,y^*)$$

Duality of Maximin and Minimax

Fact. $$\max_{x\in\Delta_m}\min_{j\in[n]} u_1(x,j) \le \min_{y\in\Delta_n}\max_{i\in[m]} u_1(i,y).$$

Theorem. $$\max_{x\in\Delta_m}\min_{j\in[n]} u_1(x,j) = \min_{y\in\Delta_n}\max_{i\in[m]} u_1(i,y).$$

- Maximin and minimax can both be formulated as linear program

Maximin

$$\begin{aligned} \text{max} \quad & u \\ \text{s.t.} \quad & u \le \textstyle\sum_{i=1}^m u_1(i,j)\, x_i, & \forall j \in [n] \\ & \textstyle\sum_{i=1}^m x_i = 1 \\ & x_i \ge 0, & \forall i \in [m] \end{aligned}$$

Minimax

$$\begin{aligned} \text{min} \quad & v \\ \text{s.t.} \quad & v \ge \textstyle\sum_{j=1}^n u_1(i,j)\, y_j, & \forall i \in [m] \\ & \textstyle\sum_{j=1}^n y_j = 1 \\ & y_j \ge 0, & \forall j \in [n] \end{aligned}$$

- This turns out to be primal and dual LP. Strong duality yields the equation

"Uniqueness" of Nash Equilibrium (NE)

Theorem. In 2-player zero-sum games, (x^*, y^*) is a NE if and only if x^* and y^* are the maximin and minimax strategy, respectively.

$\Leftarrow$: if x^* [y^*] is the maximin [minimax] strategy, then (x^*, y^*) is a NE

- Want to prove $u_1(x^*, y^*) \geq u_1(i, y^*), \forall i \in [m]$

$$
\begin{aligned}
u_1(x^*, y^*) &\geq \min_{j} u_1(x^*, j) \\
&= \max_{x \in \Delta_m} \min_{j} u_1(x, j) \\
&= \min_{y \in \Delta_n} \max_{i \in [m]} u_1(i, y) \\
&= \max_{i \in [m]} u_1(i, y^*) \\
&\geq u_1(i, y^*), \forall i
\end{aligned}
$$

- Similar argument shows $u_1(x^*, y^*) \leq u_1(x^*, j), \forall j \in [n]$
- So (x^*, y^*) is a NE

"Uniqueness" of Nash Equilibrium (NE)

Theorem. In 2-player zero-sum games, (x^*, y^*) is a NE if and only if x^* and y^* are the maximin and minimax strategy, respectively.

⇒: if (x^*, y^*) is a NE, then x^* [y^*] is the maximin [minimax] strategy

➢ Observe the following inequalities

$$
\begin{aligned}
u_1(x^*, y^*) &= \max_{i\in[m]} u_1(i, y^*) \\
&\geq \min_{y\in\Delta_n} \max_{i\in[m]} u_1(i, y) \\
&= \max_{x\in\Delta_m} \min_{j} u_1(x, j) \\
&\geq \min_{j} u_1(x^*, j) \\
&= u_1(x^*, y^*)
\end{aligned}
$$

➢ So the two "≥" must both achieve equality.

- The first equality implies y^* is the minimax strategy
- The second equality implies x^* is the maximin strategy

“Uniqueness” of Nash Equilibrium (NE)

Theorem. In 2-player zero-sum games, (x^*, y^*) is a NE if and only if x^* and y^* are the maximin and minimax strategy, respectively.

Corollary.

- NE of any 2-player zero-sum game can be computed by LPs
- Players achieve the same utility in any Nash equilibrium.
  - Player 1’s NE utility always equals maximin (or minimax) value
  - This utility is also called the game value

The Collapse of Equilibrium Concepts in Zero-Sum Games

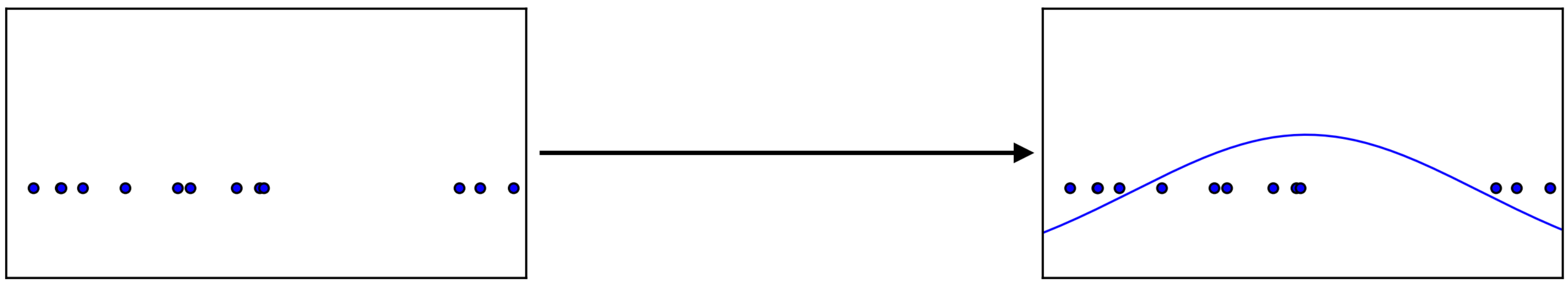
Theorem. In a 2-player zero-sum game, a player achieves the same utility in any Nash equilibrium, any correlated equilibrium, any coarse correlated equilibrium and any Strong Stackelberg equilibrium.

- Can be proved using similar proof techniques as for the previous theorem
- The problem of optimizing a player's utility over equilibrium can also be solved easily as the equilibrium utility is the same

Outline

- Correlated and Coarse Correlated Equilibrium
- Zero-Sum Games
- GANs and Equilibrium Analysis

Generative Modeling



Input data points drawn from distribution P_{true}

Output data points drawn from distribution P_{model}

Goal: use data points from P_{true} to generate a P_{model} that is close to P_{true}

Applications

Celeb training data

[Karras et al. 2017]

Input images from true distributions

Generated new images, i.e., samples from P_{model}

A few another Demos:

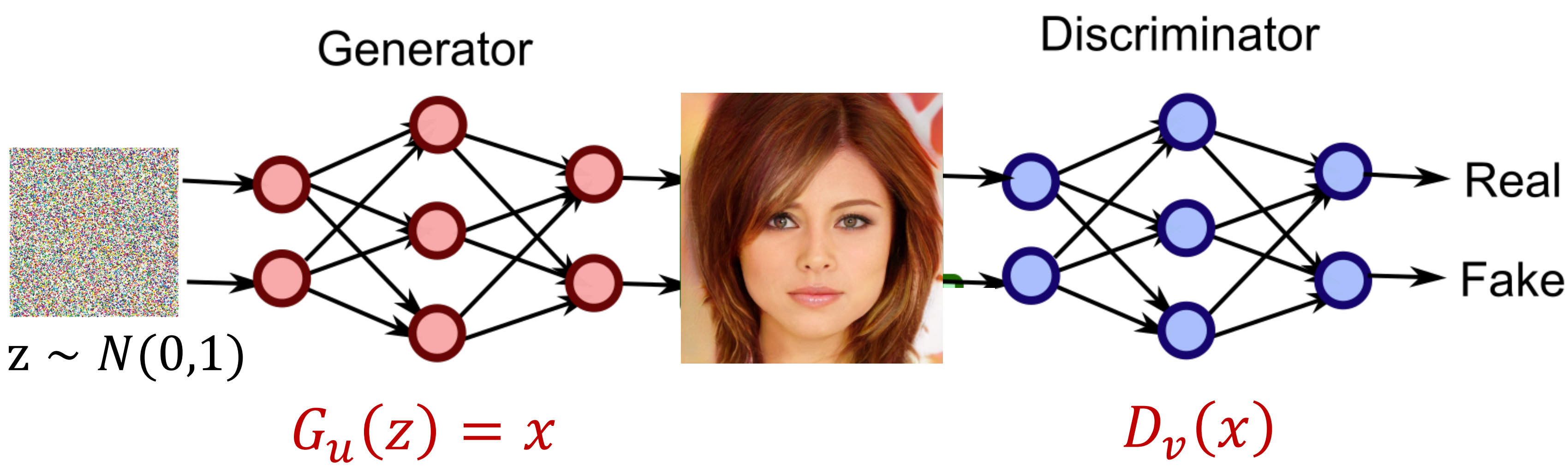
https://miro.medium.com/max/928/1*tUhgr3m54Qc80GU2BkaOiQ.gif

https://www.youtube.com/watch?v=PCBTZh41Ris&feature=youtu.be

http://ganpaint.io/demo/?project=church

GANs: Generative Adversarial Networks

- GAN is one particular generative model – a zero-sum game between the Generator and Discriminator



Objective: select model parameter u such that distribution of $G_u(z)$, denoted as P_{model}, is close to P_{real}

Objective: select model parameter v such that $D_v(x)$ is large if $x \sim P_{\text{real}}$ and $D_v(x)$ is small if $x \sim P_{\text{model}}$

GANs: Generative Adversarial Networks

- GAN is one particular generative model – a zero-sum game between the Generator and Discriminator
- The loss function originally formulated in [Goodfellow et al.'14]
  - $D_v(x) =$ probability of classifying x as "Real"
  - Log of the likelihood of being correct

$$L(u, v) = \mathbb{E}_{x \sim P_{\text{true}}} \log[D_v(x)] + \mathbb{E}_{z \sim N(0,1)} \log[1 - D_v(G_u(z))]$$

- The game: Discriminator maximizes this loss function whereas Generator minimizes this loss function
  - Results in the following zero-sum game

$$\min_u \max_v \; L(u, v)$$

  - The design of Discriminator is to improve training of Generator

GANs: Generative Adversarial Networks

- GAN is a large zero-sum game with intricate player payoffs
- Generator strategy G_u and Discriminator strategy D_v are typically deep neural networks, with parameters u, v
- Generator's utility function has the following general form where ϕ is an increasing concave function (e.g., $\phi(x) = \log x, x$ etc.)

$$\mathbb{E}_{x \sim P_{\text{true}}} \phi([D_v(x)]) + \mathbb{E}_{z \sim N(0,1)} \phi([1 - D_v(G_u(z))])$$

GAN research is mainly about modeling and solving this extremely large zero-sum game for various applications

WGAN – A Popular Variant of GAN

- Drawbacks of log-likelihood loss: unbounded at boundary, unstable
- Wasserstein GAN is a popular variant using a different loss function
  - I.e., substitute log-likelihood by the likelihood itself

$$\mathbb{E}_{x\sim P_{\text{true}}} D_v(x) - \mathbb{E}_{z\sim N(0,1)} D_v(G_u(z))$$

  - Training is typically more stable

Research Challenges in GANs

$$\min_{u} \max_{v} \ \mathbb{E}_{x\sim P_{\text{true}}}\phi([D_v(x)]) + \mathbb{E}_{z\sim N(0,1)}\phi([1 - D_v(G_u(z))])$$

- What are the correct choice of loss function ϕ?
- What neural network structure for G_u and D_v?
- Only pure strategies allowed – equilibrium may not exist or is not unique due to non-convexity of strategies and loss function
- Do not know P_{true} exactly but only have samples
- How to optimize parameters u, v?
- . . .

A Basic Question

Even if we computed the equilibrium w.r.t. some loss function, does that really mean we generated a distribution close to P_{true}?

Research Challenges in GANs

$$\min_u \max_v \; \mathbb{E}_{x\sim P_{\text{true}}}\phi([D_v(x)]) + \mathbb{E}_{z\sim N(0,1)}\phi([1 - D_v(G_u(z))])$$

A Basic Question

Even if we computed the equilibrium w.r.t. some loss function, does that really mean we generated a distribution close to P_{true}?

- Intuitively, if the discriminator network D_v is strong enough, we should be able to get close to P_{true}
- Next, we will analyze the equilibrium of a stylized example

(Stylized) WGANs for Learning Mean

- True data drawn from $P_{\text{true}} = N(\alpha, 1)$
- Generator $G_u(z) = z + u$ where $z \sim N(0,1)$
- Discriminator $D_v(x) = vx$

Remarks:

a) Both Generator and Discriminator can be deep neural networks in general

b) We picked particular format for illustrative purpose and also convenience of theoretical analysis

(Stylized) WGANs for Learning Mean

- True data drawn from $P_{\text{true}} = N(\alpha, 1)$
- Generator $G_u(z) = z + u$ where $z \sim N(0,1)$
- Discriminator $D_v(x) = vx$
- WGAN then has the following close-form format

$$\min_u \max_v \; \mathbb{E}_{x \sim P_{\text{true}}}[D_v(x)] + \mathbb{E}_{z \sim N(0,1)}[1 - D_v(G_u(z))]$$

$$\Rightarrow \; \min_u \max_v \; \mathbb{E}_{x \sim N(\alpha,1)}\,[vx] \; + \mathbb{E}_{z \sim N(0,1)}[1 - v(z + u)]$$

$$\Rightarrow \; \min_u \max_v \; [v\alpha] \; + \; [1 - vu]$$

- This minimax problem solves to $u^* = \alpha$
- I.e, WGAN does precisely learn P_{true} at equilibrium in this case

See paper “**Generalization and Equilibrium in GANs**” by Aaora et al. (2017) for more analysis regarding the equilibrium of GANs and whether they learn a good distribution at equilibrium

Thank You

Haifeng Xu

University of Virginia

hx4ad@virginia.edu