CSC2556

Lecture 11

Noncooperative Games 2: Zero-Sum Games, Stackelberg Games

Zero-Sum Games

- Total reward is constant in all outcomes (w.l.o.g. 0)
- Focus on two-player zero-sum games (2p-zs)
  - "The more I win, the more you lose"
  - Chess, tic-tac-toe, rock-paper-scissor, …

P2 / P1	Rock	Paper	Scissor
Rock	(0 , 0)	(-1 , 1)	(1 , -1)
Paper	(1 , -1)	(0 , 0)	(-1 , 1)
Scissor	(-1 , 1)	(1 , -1)	(0 , 0)

Zero-Sum Games

- Reward for P2 = - Reward for P1
  - Only need a single matrix A : reward for P1
  - P1 wants to maximize, P2 wants to minimize

P1 \ P2	Rock	Paper	Scissor
Rock	0	-1	1
Paper	1	0	-1
Scissor	-1	1	0

Rewards in Matrix Form

- Reward for P1 when…
  - P1 uses mixed strategy x_1
  - P2 uses mixed strategy x_2
  - $x_1^T \, A \, x_2$ (where x_1 and x_2 are column vectors)

Maximin/Minimax Strategy

- Worst-case thinking by P1…
  - If I commit to x_1 first, P2 would choose x_2 to minimize my reward (i.e., maximize his reward)

- P1's best worst-case guarantee:

$$V_1^* = \max_{x_1} \min_{x_2} x_1^T * A * x_2$$

  - A maximizer x_1^* is a maximin strategy for P1

Maximin/Minimax Strategy

- P1's best worst-case guarantee:

$$V_1^* = \max_{x_1} \min_{x_2} x_1^T * A * x_2$$

- P2's best worst-case guarantee:

$$V_2^* = \min_{x_2} \max_{x_1} x_1^T * A * x_2$$

  - P2's minimax strategy x_2^* minimizes this

- $V_1^* \leq V_2^*$ (both play their "safe" strategies together)

The Minimax Theorem

- Jon von Neumann [1928]

- Theorem: For any 2p-zs game,
  - $V_1^* = V_2^* = V^*$ (called the minimax value of the game)
  - Set of Nash equilibria =

    $\{(\mathrm{x}_1^*, \mathrm{x}_2^*) : \mathrm{x}_1^* = \text{maximin for P1}, \mathrm{x}_2^* = \text{minimax for P2}\}$

- Corollary: x_1^* is best response to x_2^* and vice-versa.

The Minimax Theorem

- Jon von Neumann [1928]

  "As far as I can see, there could be no theory of games … without that theorem …

  I thought there was nothing worth publishing until the Minimax Theorem was proved"

- Indeed, much more compelling and predictive than Nash equilibria in general-sum games (which came much later).

Computing Nash Equilibria

- General-sum games: Computing a Nash equilibrium is PPAD-complete even with just two players.
  - Trivia: Another notable PPAD-complete problem is finding a three-colored point in Sperner's Lemma.

- 2p-zs games: Polynomial time using linear programming
  - Polynomial in #actions of the two players: m_1 and m_2

Computing Nash Equilibria

Maximize v

Subject to

$$(x_1^T A)_j \geq v, \; j \in \{1, \dots, m_2\}$$

$$x_1(1) + \cdots + x_1(m_1) = 1$$

$$x_1(i) \geq 0, i \in \{1, \dots, m_1\}$$

Minimax Theorem in Real Life?

- If you were to play a 2-player zero-sum game (say, as player 1), would you always play a maximin strategy?

- What if you were convinced your opponent is an idiot?

- What if you start playing the maximin strategy, but observe that your opponent is not best responding?

Minimax Theorem in Real Life?

Minimax Theorem in Real Life?

Kicker \ Goalie	L	R
L	0.58	0.95
R	0.93	0.70

Kicker

Maximize v

Subject to

$0.58p_L + 0.93p_R \geq v$

$0.95p_L + 0.70p_R \geq v$

$p_L + p_R = 1$

$p_L \geq 0, p_R \geq 0$

Goalie

Minimize v

Subject to

$0.58q_L + 0.95q_R \leq v$

$0.93q_L + 0.70q_R \leq v$

$q_L + q_R = 1$

$q_L \geq 0, q_R \geq 0$

Minimax Theorem in Real Life?

Kicker \ Goalie	L	R
L	0.58	0.95
R	0.93	0.70

Kicker
Maximin:
$p_L = 0.38, p_R = 0.62$

Reality:
$p_L = 0.40, p_R = 0.60$

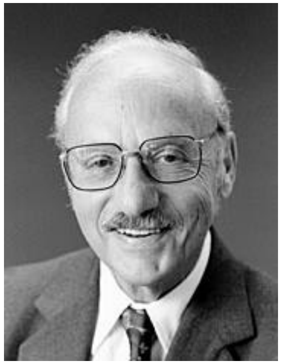
Goalie
Maximin:
$q_L = 0.42, q_R = 0.58$

Reality:
$p_L = 0.423, q_R = 0.577$

Minimax Theorem

- Implies Yao's minimax principle
- Equivalent to linear programming duality

John von Neumann



George Dantzig

von Neumann and Dantzig

George Dantzig loves to tell the story of his meeting with John von Neumann on October 3, 1947 at the Institute for Advanced Study at Princeton. Dantzig went to that meeting with the express purpose of describing the linear programming problem to von Neumann and asking him to suggest a computational procedure. He was actually looking for methods to benchmark the simplex method. Instead, he got a 90-minute lecture on Farkas Lemma and Duality (Dantzig's notes of this session formed the source of the modern perspective on linear programming duality). Not wanting Dantzig to be completely amazed, von Neumann admitted:

"I don't want you to think that I am pulling all this out of my sleeve like a magician. I have recently completed a book with Morgenstern on the theory of games. What I am doing is conjecturing that the two problems are equivalent. The theory that I am outlining is an analogue to the one we have developed for games.“

- (Chandru & Rao, 1999)

Sequential Move Games

- Focus on two players: “leader” and “follower”
- Leader first commits to playing x_1, follower chooses a best response x_2
  - We can assume x_2 to be a pure strategy w.l.o.g.
  - We don’t need x_1 to be a best response to x_2

A Curious Case

P1 \ P2	Left	Right
Up	(1 , 1)	(3 , 0)
Down	(0 , 0)	(2 , 1)

- Q: What are the Nash equilibria of this game?

- Q: You are P1. What is your reward in Nash equilibrium?

A Curious Case

P2 / P1	Left	Right
Up	(1 , 1)	(3 , 0)
Down	(0 , 0)	(2 , 1)

- Q: As P1, you want to commit to a pure strategy. Which strategy would you commit to?

- Q: What would your reward be now?

Commitment Advantage

P2 / P1	Left	Right
Up	(1 , 1)	(3 , 0)
Down	(0 , 0)	(2 , 1)

- With commitment to mixed strategies, the advantage could be even more.
  - If P1 commits to playing Up and Down with probabilities 0.49 and 0.51, respectively…
  - P2 is still better off playing Right than Left, in expectation
  - $\mathbb{E}$[Reward] for P1 increases to ~2.5

Stackelberg Equilibrium

- Leader chooses a minimax strategy, follower chooses a best response

- Commitment is always advantageous
  - The leader always has the option to commit to a Nash equilibrium strategy.

- What about the police trying to catch a thief?

Zero-Sum Stackelberg

- This can be computed using the same LP that we used for 2p-zs Nash equilibrium:

Maximize v

Subject to

$$(x_1^T A)_j \geq v, \; j \in \{1, \dots, m_2\}$$

$$x_1(1) + \cdots + x_1(m_1) = 1$$

$$x_1(i) \geq 0, i \in \{1, \dots, m_1\}$$

General-Sum Stackelberg

- Reward matrices A, B with $B \neq -A$

$$\max_{x_1} \ (x_1)^T A\, f(x_1)$$

$$\text{where } f(x_1) = \max_{x_2} \ (x_1)^T B\, x_2$$

- How do we compute this?

Stackelberg Games via LPs

- S_1, S_2 = sets of actions of leader and follower
- $|S_1| = m_1, |S_2| = m_2$
- $x_1(s_1)$ = probability of leader playing s_1
- π_1, π_2 = reward functions for leader and follower

$$\max \Sigma_{s_1 \in S_1} x_1(s_1) \cdot \pi_1(s_1, s_2^*)$$

subject to

$$\forall s_2 \in S_2, \;\; \Sigma_{s_1 \in S_1} x_1(s_1) \cdot \pi_2(s_1, s_2^*) \;\geq\; \Sigma_{s_1 \in S_1} x_1(s_1) \cdot \pi_2(s_1, s_2)$$

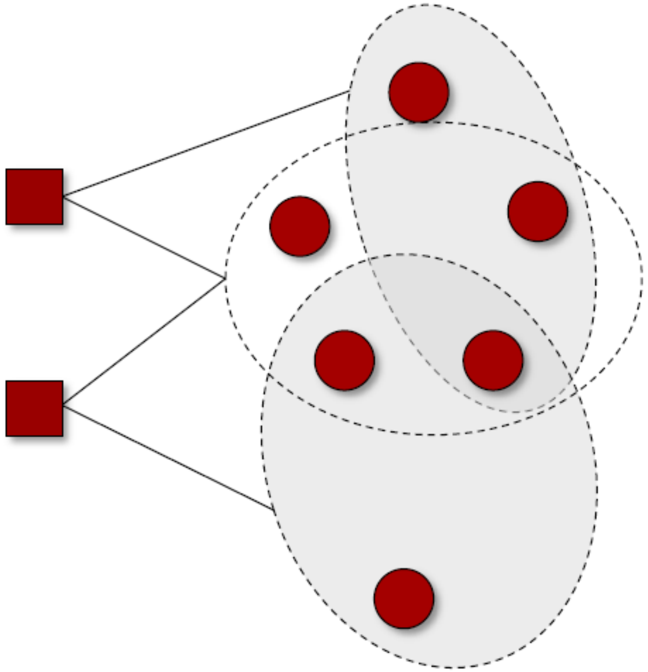
$$\Sigma_{s_1 \in S_1} x_1(s_1) = 1$$

$$\forall s_1 \in S_1, x_1(s_1) \geq 0$$

- One LP for each s_2^*, take the maximum over all m_2 LPs
- The LP corresponding to s_2^* optimizes over all x_1 for which s_2^* is the best response

Real-World Applications

- Security Games
  - Defender (leader) wants to deploy security resources to protect targets
  - A resource can protect one of several subsets of targets
  - Attacker (follower) observes the defender's strategy, and chooses a target to attack
  - Both players get a reward/penalty



- Number of actions is exponential

LAX

Newsweek National News

Subscribe Now | Make Newsweek Your Homepage | Newsletters | RSS

The Element of Surprise

To help combat the terrorism threat, officials at Los Angeles Inter Airport are introducing a bold new idea into their arsenal: randon of security checkpoints. Can game theory help keep us safe?

WEB EXCLUSIVE

By Andrew Murr
Newsweek
Updated: 1:00 p.m. PT Sept 28, 2007

Security forces work the sidewalk

Sept. 28, 2007 - Security officials at Los Angeles International Airport now have a new weapon in their fight against terrorism: complete, baffling randomness. Anxious to thwart future terror attacks in the early stages while plotters are casing the airport, LAX security patrols have begun using a new software program called ARMOR, NEWSWEEK has learned, to make the placement of security checkpoints completely unpredictable. Now all airport security officials have to do is press a button labeled "Randomize," and they can throw a sort of digital cloak of invisibility over where they place the cops' antiterror checkpoints on any given day.

Real-World Applications

- Protecting entry points to LAX
- Scheduling air marshals on flights
  - Must return home
- Protecting the Staten Island Ferry
  - Continuous-time strategies
- Fare evasion in LA metro
  - Bathroom breaks !!!
- Wildlife protection in Ugandan forests
  - Poachers are not fully rational
- Cyber security

…