Lecture 3: Zero Sum Games & Online Learning

Lecture 3: 28 Aug 2024

CS6208 Fall 2024: Computational Game Theory

Recall from last week

Definition of a game

- Matrix games
- Examples of 2x2 games

Definition of Nash equilibrium for normal-form games

- Expressing things in matrix-vector representation
- Computational complexity
- Integer programming
- Support Enumeration
- Lemke Howson Algorithm

Today’s Agenda

First half: Limitations and practical considerations with Nash

Second half: **2-player zero-sum games** overcomes some of these

- The Minimax theorem, Nash in 2-player zero-sum games
- Finding NE (for 2-player zero-sum games) via linear programming
- Approximate NE, Exploitability
- ~~No-regret learning, multiplicative weights, obtaining NE via self-play~~
- ~~Regret matching, RM+ and Blackwell approachability [optional]~~

Limitations of Nash Equilibria

“I know the nuclear launch codes… but strangely, they don’t improve my dating life.”

Don’t romanticize Nash equilibrium

The **Grandmaster** (**En Dwi Gast**) is a fictional character appearing in American comic books published by Marvel Comics. The character first appeared in *The Avengers* #69.[2] The Grandmaster is one of the ageless Elders of the Universe and has mastered most civilizations' games of skill and chance. Different media appearances depict him as the Collector's brother.

He has a highly developed superhuman intellect, with vast knowledge and comprehension of games and game theory far beyond present-day Earth, as well as encyclopedic knowledge of thousands of exotic games played throughout the universe. He can calculate diverse low information probabilities within a

NE has **many** limitations!

… The theory can be conceived as both normative and as descriptive. It is normative in the sense that it wishes to indicate what the optimal course of action is in different situations: for example, **in the zero-sum situation**…

… If game theory is to be thought of in the latter sense, **then it is really a model**, and like all models must be looked upon as such….

…Now I think it is important for everybody to realize that there are many situations for which game theory is a good model; but the question arises whether it is a model for **every conflict situation** or every decision situation which might occur.

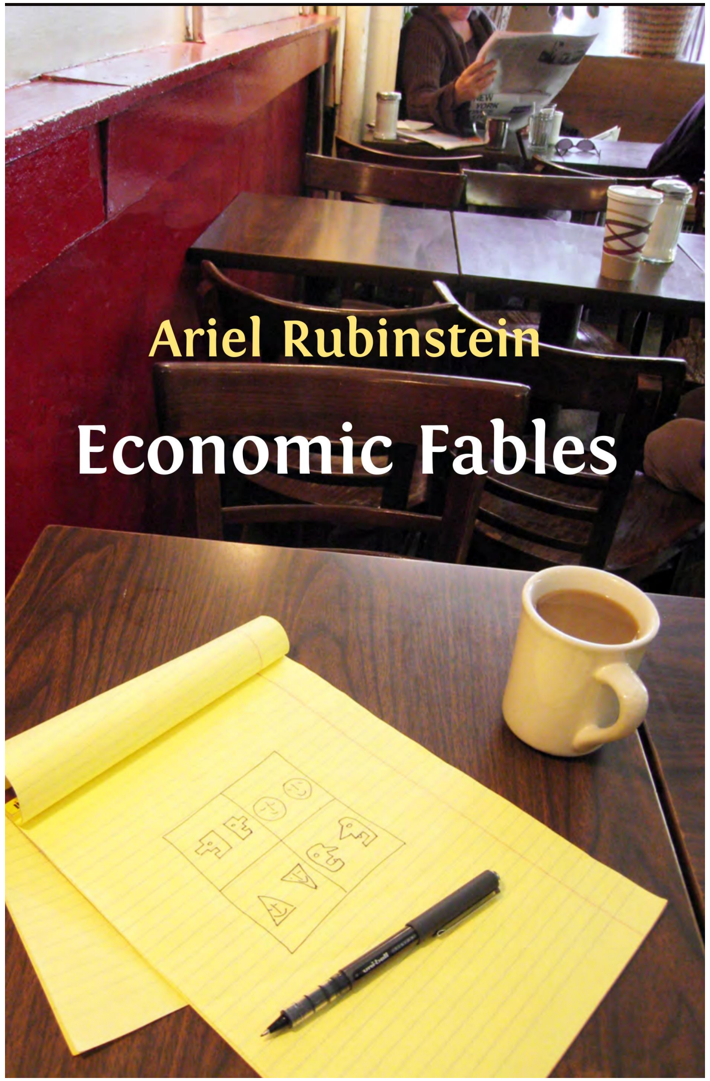
… Of course, it has never been asserted by anybody who seriously writes on game theory that **every single decision is a game situation**. Certainly a judge does not engage himself in playing a game. He makes a legal decision and that is a very different thing. He has to classify; he has to judge. That is totally different….

-Oskar Morgenstern, On some criticisms of Game Theory (1964)

John Von Neumann, one of the founders of game theory, was not only a genius in mathematics, he was also a genius in public relations. The choice of the name “theory of games” was brilliant as a marketing device. The word “game” has friendly, enjoyable associations. It gives a good feeling to people. It reminds us of our childhood, of chess and checkers, of children’s games.

The associations are very light, not heavy, even though you may be trying to deal with issues like nuclear deterrence. I think it’s a very tempting idea for people, **that they can take something simple and apply it to situations that are very complicated, like the economic crisis or nuclear deterrence.** But this is an illusion. Now my views, I have to say, are extreme compared to many of my colleagues. I believe that game theory is very interesting. I’ve spent a lot of my life thinking about it, **but I don’t respect the claims that it has direct applications**.

- Ariel Rubinstein, Economic Fables



My personal thoughts

There are many people who make a living off publishing or studying games or rational decision making.

I personally don’t think they are wiser than the average person

- At least, they do not apply their expertise in everyday decision making

Games are powerful & expressive, but just because a situation involves multiple ~~agents~~ actors does not mean they have to be mathematically modeled by games

Is Nash Practical?

Where do the payoffs come from?

In general, a very difficult problem

- Utilities are very subjective! Sometimes, only ordinal (rankings) are available, cardinal utilities are a “by-product” (see VNM utility theorem)

Does it make sense that I know my own utility for every outcome?

Does it make sense for me to know the opponent’s utility…?

- Can I assume opponent shares utility with me?
- Or should I assume opponent has opposing utilities as me?

Can we *learn* from data?

- Is this learning process itself strategyproof/vulnerable to attacks?
- See for example the tutorial on strategic ML (ICML 2024)

What constitutes a game?

- Do we include every person in the world?

Why do we even care about equilibria?

Two common goals

- We want to **understand** players' behavior better
- We want to **prescribe** good strategies to players
- Reminder: this class focuses on the second

Very common remark

- Game theory "doesn't work" in practice
  - Real life problems are about more data, foundation models etc.
  - I agree, sometimes!
- What are the alternatives?
  - Is condition of being incentive compatible strong or mild?
- Are modern algorithms secretly trying to achieve Nash without explicitly writing down equilibrium conditions?
  - What does it mean for a RL algorithm to "converge"?

Can Nash predict behavior?

Let's do an experiment!

Choose an integer from {1,…,100}

We are going to collate everyone's choice and take the mean $\bar{x}$

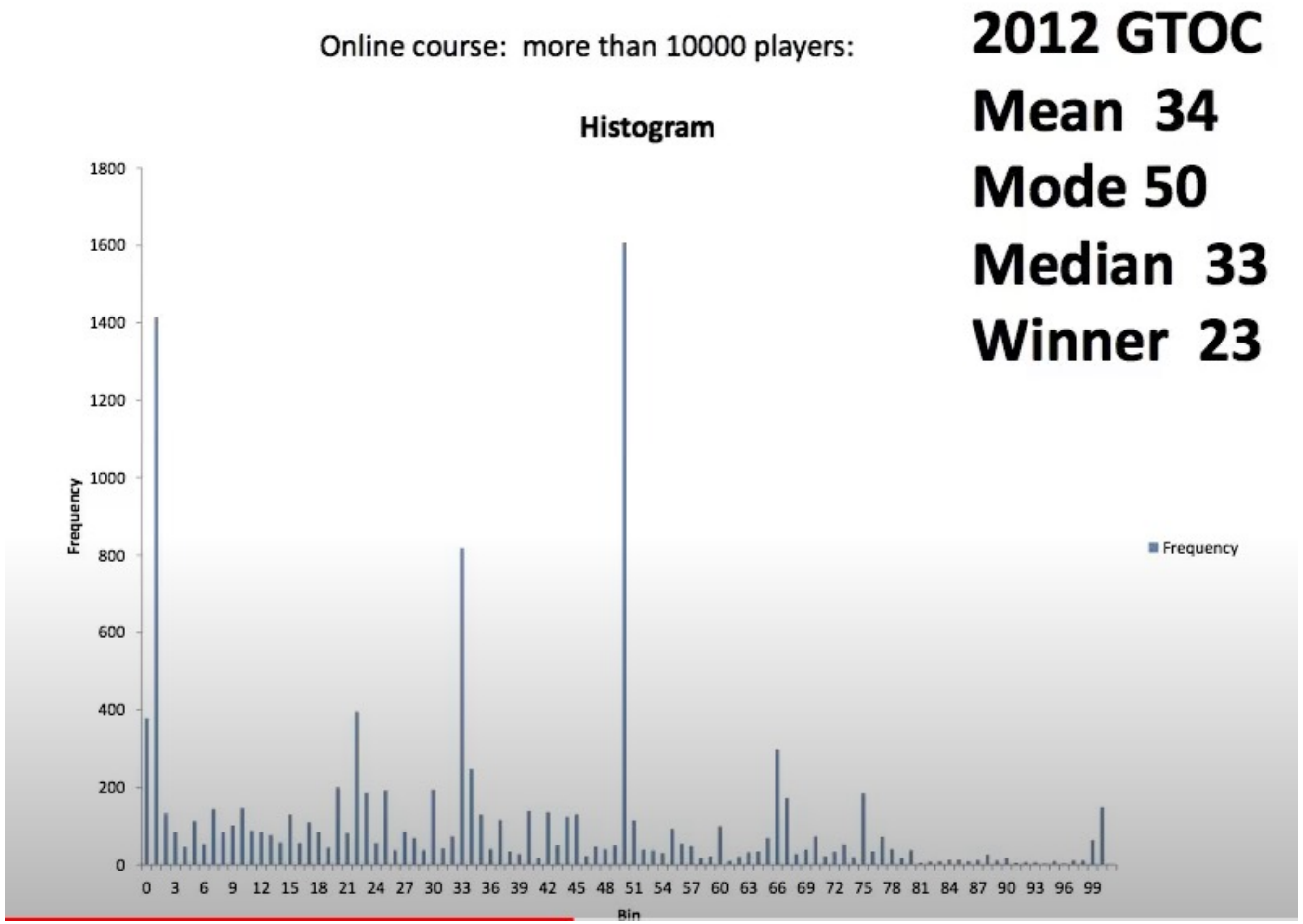
Consider the number $\mathrm{x}^* = \mathrm{int}(\frac{2}{3}\bar{x})$

The person who chose the number closest to x^* will win **a prize**

- Ties broken uniformly at random

*This is known as the
beauty-contest game

Some results…



Source: Lectures by Jackson, Leyton-Brown, and Shoham (http://www.game-theory-class.org), the videos can be found on youtube

More Practical Issues

Bounded Rationality

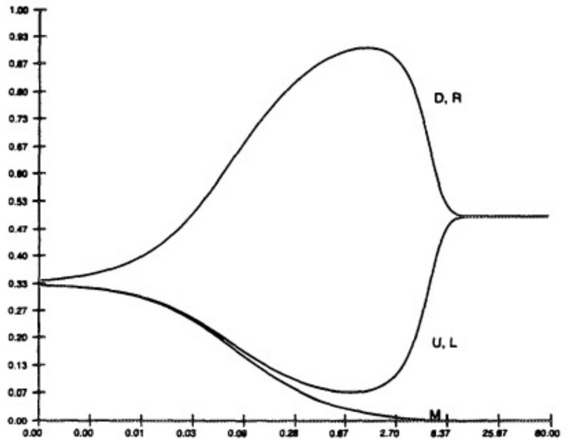


FIG. 2. QRE for game of Table 1, with $A = B = 100$.

Assumption from Nash: both players are behaving *rationally*

- Recall for NE: $x \in BR_1(y)$ AND $y \in BR_2(x)$
- Will play some best response to other player's strategy
  - Can be a bit brittle! x, y is Argmax --- a set function over expected utilities

Quantal Response Equilibrium

- Replace Argmax by **softmax** for both players
  - $x^* = softmax(\frac{1}{\lambda} A y^*)$ AND $y^* = softmax(\frac{1}{\lambda} B^T x^*)$
  - $x^* = argmax_x x^T A y^* + \lambda H(x), y = argmax_y (x^* B y) + \lambda H(y)$

*Be careful with whether you use λ or $1/\lambda$

- Equivalent to adding Gumbel noise to all utilities and playing according to the probability that an action is the best (random choice models)
- Always exists, converges to a NE as temperature → 0 if starting at principal branch
  - Can be used to "select" Nash (Quantal Response Equilibria for Normal Form Games, McKelvey and Palfrey, 1995)
- Lots of experiments "fitting" data to QRE. But introduces one more problem of estimating λ

Level-k

- NE is some kind of "infinite" nesting of beliefs
- E.g.: population contains different types of people, some think many "steps" ahead, others few. Common belief over distribution over levels
- http://www.columbia.edu/~md3405/Behave_Bounded_6_15.pdf

Equilibrium Selection

Can have multiple equilibrium, recall chicken game

6,6	1,7
7,1	0,0

Chicken Game

Very different payoffs depending on which equilibrium is chosen. Which is the "right" equilibrium? Who decides? How?

- You? Your opponent? Use the one that is worst for you? Or be optimistic?
- If we are being pessimistic, how about just assume opponent's is (irrationally) just trying to hurt us?
  - Becomes a zero-sum game between you and opponent

Payoff and risk dominance, Equilibrium refinements

Problem is accentuated if considering repeated games

Others

Multiplayer games

- Hard to solve (no longer LPs), equilibria can contain irrational numbers
- Even more equilibrium selection problems
- Large strategy profiles

Games with incomplete information

- Not clear what the payoffs of the game are
- Resolved (somewhat…) by *Bayesian games*
  - Each player is assigned at private type t_i at the beginning drawn from some joint distribution over types. Payoff depends on types and action. Bayes Nash equilibrium

Not all games have discrete actions/time

- Example: choosing a path that a robot takes, flight routes, logistics
- Distribution over continuous domain is somewhat harder

A note on testing game solving algorithms

Test our algorithms on “random games”

- E.g., payoffs are i.i.d. gaussian, in U[0, 1]
- I did (and sometimes still) do this

Not very realistic

- Degenerate games frequently
- We usually aren’t too concerned with “random” games

Gives overly optimistic view or running time

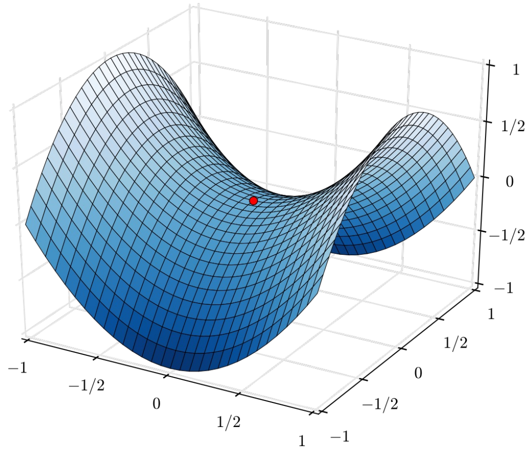
- Random games often have “easy” equilibria, e.g., games with pure NE
  - Solve easily in poly-time by enumeration
  - Provable: Goldberg et. al (1968), Krever et al (2025), Pradelski and Tarbush (2025) etc
  - Empirically we also see this
- Many games have some kind of structure underneath; payoffs depend on some latent variables (maybe modify those instead?)

Part 2: Zero-sum Games

“Oh that? Not a problem — more of a recurring plot twist.”

What are zero-sum Games?

Recall P1 and P2 payoffs = $x^T Ay, x^T By$



Zero-sum games: $A = -B$

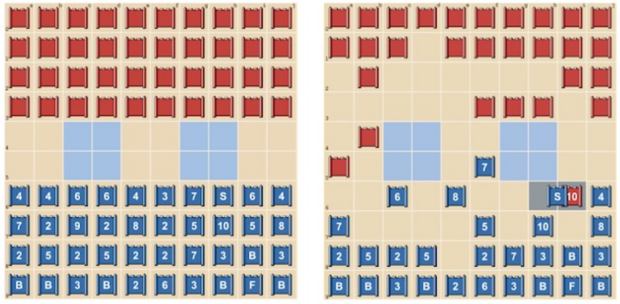
- Fully competitive
- Can just use a single matrix A
- P1's payoff is $x^T Ay$, P2's is $-x^T Ay$

By convention P1 (P2) is minimizing (maximizing) $x^T Ay$

- Can use other convention but you need to rotate your head/screen 180°

1v1 poker

Starcraft II



Stratego

Street Fighter

Review of Convex Optimization

"Premature optimization is the root of all evil"

-Donald Knuth

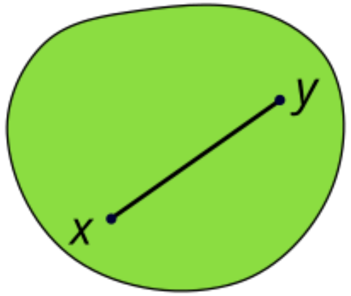
Convex Sets

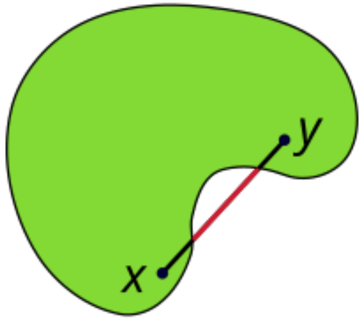
C is convex if it contains the line segment between any 2 of its points

- Assume $C \subseteq R^n$ and nonempty
- Can be open set, or unbounded
- Many tricky cases. Thankfully, those we encounter will be obvious
- Most of our C will also be compact (closed and bounded)

Examples:

- Simplex
- Convex hull of points (for this class, finite)
- Intersection of finite number of half-spaces (polyhedron/polytope)





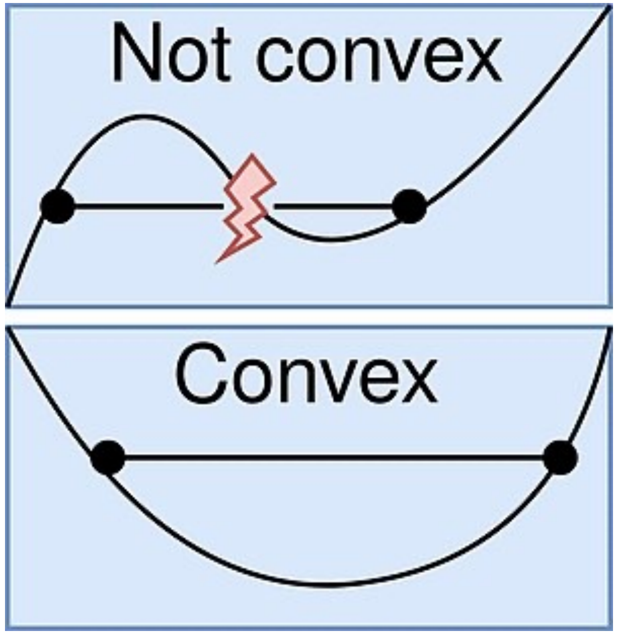
Source: wikipedia

Convex Functions

Intuition: bends upwards, $\underbrace{\text{Hessian}}_{\text{Second derivative}}$ PSD everywhere, $\underbrace{\text{epigraph}}_{\text{Set of points lying “above” function}}$ is convex

Properties

- Scaling a convex function by a nonnegative constant
- Sum of convex functions
- Maximum of convex functions



Source: wikipedia

Examples

- Linear functions $f(x) = c^T x$
- Quadratic: $f(x) = x^T Ax$ for A positive-semidefinite
- Maximum over linear functions $f(x) = \max_i c_i^T x$

Concave function = negative of convex function

This class: convex functions with convex, compact domains

The Minimax Theorem

“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**”

-John Von Neumann

Who has the advantage?

Suppose you and I are in a competition

- Let $f(x, y)$ be some real function (known to both) on X, Y
- You choose x, I choose y. Outcome is $f(x, y)$.
- You are minimizing the outcome, I am maximizing it

No assumptions on f or X, Y (except for assumption that min/max exists)

Choices are made sequentially and publicly

Which order of choosing is better for you? Me first, or you?

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} f(x, y) \quad \textbf{V.S.} \quad \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} f(x, y)$$

Option 1: You choose first **Option 2: I choose first**

The min-max inequality

General statement, **always** holds

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} f(x, y) \geq \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} f(x, y)$$

"Inner optimization wins"

- It can optimize for specific value the outer optimizer chose
- Other direction needs to account for all opponent's choices

Very much related to *weak duality* in optimization [optional]

- **Any** dual feasible solution yields a lower bound on the primal problem
- But solution to dual may not be primal optimal

Von Neumann's Minimax Theorem

If $f(x, y)$ continuous and
- $\mathcal{X}, \mathcal{Y}$ compact (closed and bounded)
- convex in x, i.e., $f(x, y_0)$ is convex for any fixed y_0
- concave in y, i.e., $f(x_0, y)$ is concave for any fixed x_0

Zur Theorie der gesellschaftsspiele (1920s) (on the theory of games and strategy).

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} f(x, y) = \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} f(x, y)$$

No disadvantage in choosing x (or y) first!

Generalizations and strict weakening of assumptions exist, e.g., Sion's minimax theorem

Can be proven by Nash's Theorem (but predates Nash)

Related to *strong duality* in optimization [optional]
- Zero duality gap, solving dual will give same answer as optimal
- Not generally true there, need constraint qualifications

History: https://web.math.ucsb.edu/~crandall/math201b/vnminimax.pdf

Saddle Points

Gradient is 0, minimizes x for fixed y, maximizes y for fixed x

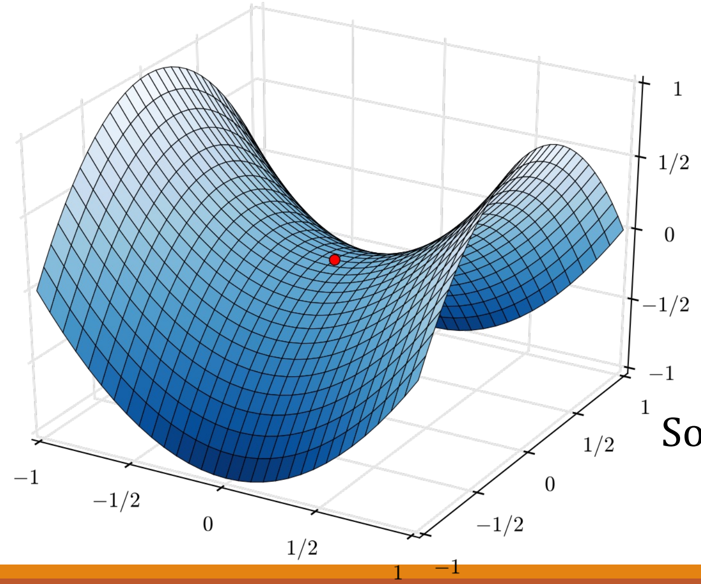
- Hessian (if twice differentiable) has at least one positive and negative eigenvalue

Usually bad for finding local minima (over non-convex functions)

- e.g., training of neural network parameters

For this class, we **want** to find the saddle point

Reminder: the domains $\mathcal{X}, \mathcal{Y}$ can be high dimensional!



Source: Wikipedia

Minimax Theorem in matrix games

Instantiating $f, \mathcal{X}, \mathcal{Y}$

- $f = x^T A y, \mathcal{X} = \Delta_n, \mathcal{Y} = \Delta_m$ (simplexes of dimension n and m)
- f is bilinear, i.e. *linear* in x and y. Linear →convex (and concave)
- Simplexes are convex, compact

$$\underbrace{\min_{x\in\Delta_n}}_{x^*} \max_{y\in\Delta_m} x^T A y = \underbrace{\max_{y\in\Delta_m}}_{y^*} \min_{x\in\Delta_n} x^T A y$$

Implication: no disadvantage in choosing x (or y) first!

$$\max_y \min_x x^T A y = \min_x x^T A y* \le \underbrace{x^{*T} A y^*}_{\text{Recall: expected payoff of the game}} \le \max_y x^{*T} A y = \min_x \max_y x^T A y$$

Caution: Choosing, in this sense means choosing of a **mixed strategy** x (or y) and obtaining rewards $x^T A y$.

- It is **not** the players playing their actions in turn!

Nice properties of 2P zero-sum games

Unique value→no need for equilibrium selection

Exchangeable: if (x, y) and (x', y') then (x, y') is also Nash

Set of NE form a polytope for each player

Adding additional actions cannot hurt player

- Not the case for general-sum games!

Can be quantitatively evaluated (next few slides)

Can be computed in polynomial time (recall NE in general is PPAD hard)

Caution: All of these apply only to **2-player** zero-sum games

- If n-player zero-sum games are "anything special" then we can take any general-sum game, add dummy player to make it zero-sum, and say interesting things about n-1 player general-sum games too
- Special cases exist (e.g., polymatrix games)

Examples

Hagrid: “Do you mean ter tell me that this boy—this boy!—knows nothin' abou'—about ANYTHING?”

Harry: “I know some things. **I can, you know, do math and stuff**.”

-Harry Potter and the Philosopher’s Stone, before Harry learns he’s a wizard

Recall: The crippled shooter

		Goalkeeper L	Goalkeeper R
Crippled Shooter	L	0	0.5
	R	1	0

50% chance of missing even if goalkeeper dives incorrectly

Shooter wins less if shooting left (e.g., 50% chance to outright miss)

How should the crippled shooter play? Shoot left more, equal, or less frequently?
How should the goalkeeper play? Dive left more, equal, or less frequently?
If both players play optimally, what is the expected utility of shooter?

Visualizing Nash

I added a negative sign here, by convention, the row player minimizes

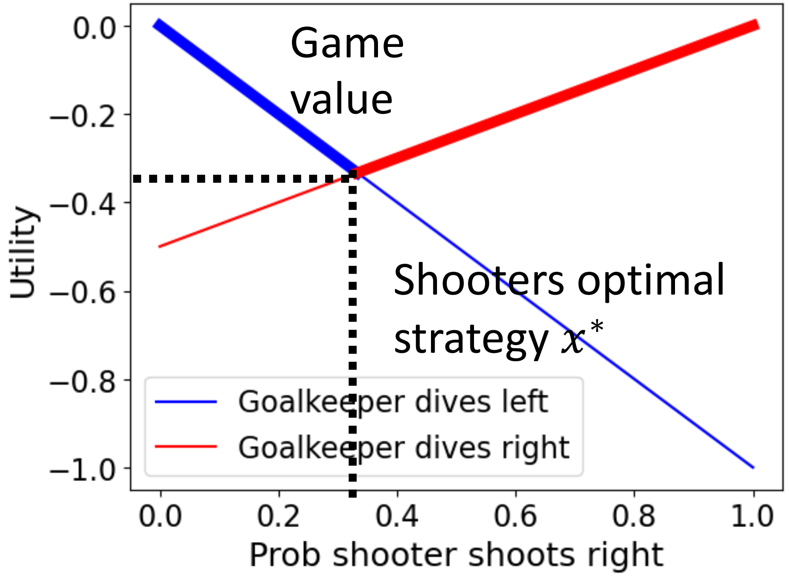
		Goalkeeper L	Goalkeeper R
Crippled Shooter	L	0	-0.5
	R	-1	0

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} f(x, y)$$

Treat this as a function of x

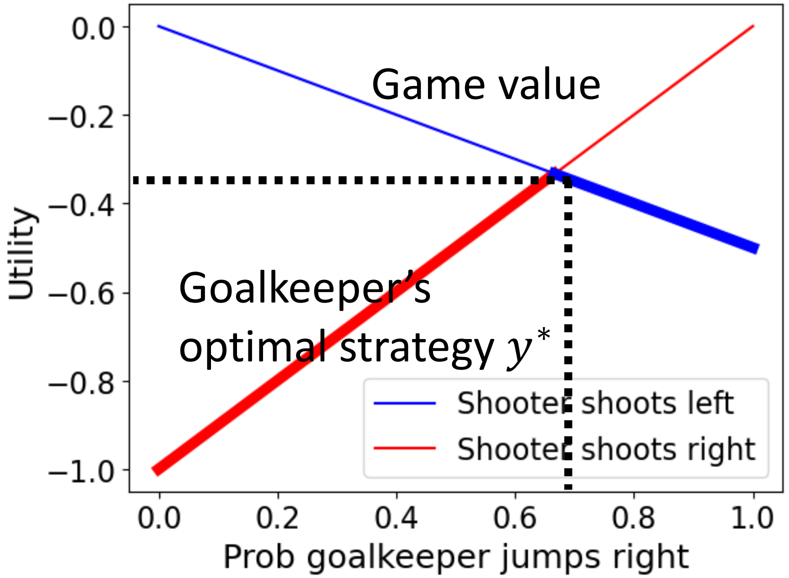
Minimizing a piecewise linear function

Each line is an opponent's action.



Max of linear functions is convex

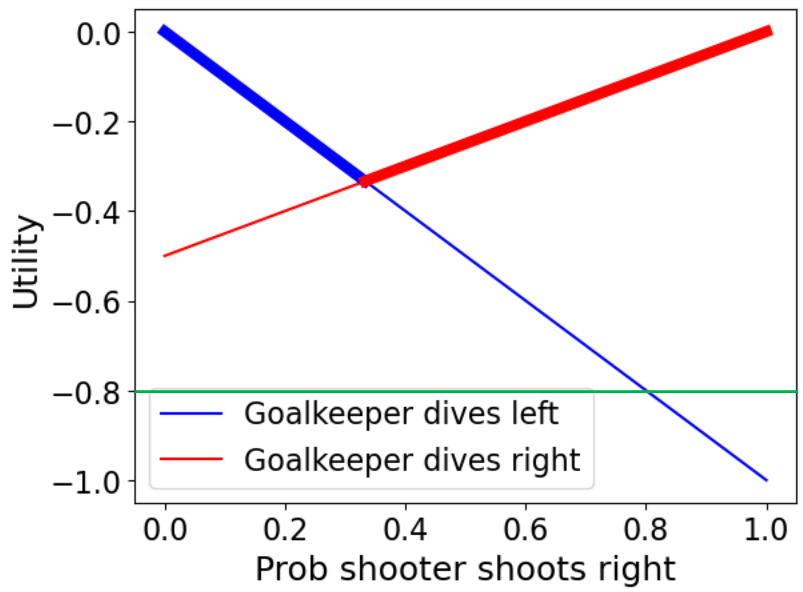
Q: why max for minimizing player?



Min of linear functions is concave

Nash as being conservative

		Goalkeeper L	Goalkeeper R	Walk around randomly
Crippled Shooter	L	0	-0.5	-0.8
	R	-1	0	-0.8



Why do we shoot in the "bad" direction more frequently?

- Hedging against all opponent strategies!
- No matter what goalkeeper does, we will get no worse than 0.33 (-0.33)

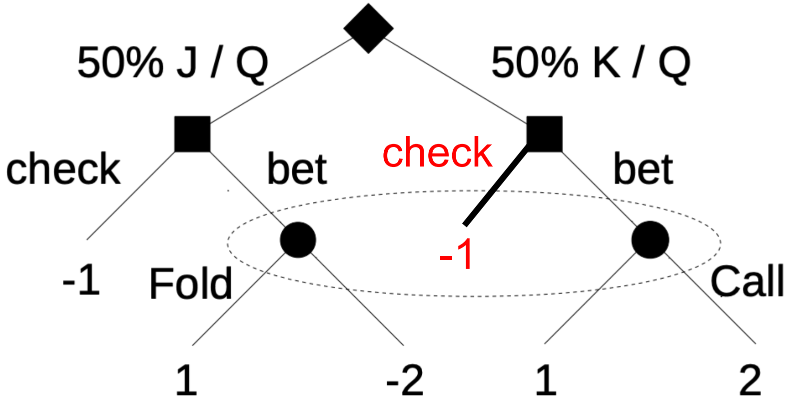
Caution:

- This does not mean that we will get exactly 0.33 for **all** opponent action. Opponent could have an action dominated strategies)
- Not just about finding any intersection

Exercise

	J→Check K→Check	J→Check K→Bet	J→Bet K→Check	J→Bet K→Bet
Fold	-1	0	0	1
Call	-1	0.5	-1.5	0

We will learn a more systemic and efficient way of doing this transformation in future



From the thesis of Neil Burch

In Lecture 1, we reasoned heuristically that the player moving first (the max player) can guarantee payoff of 1/3.

Solve the game from the minimizing player's perspective

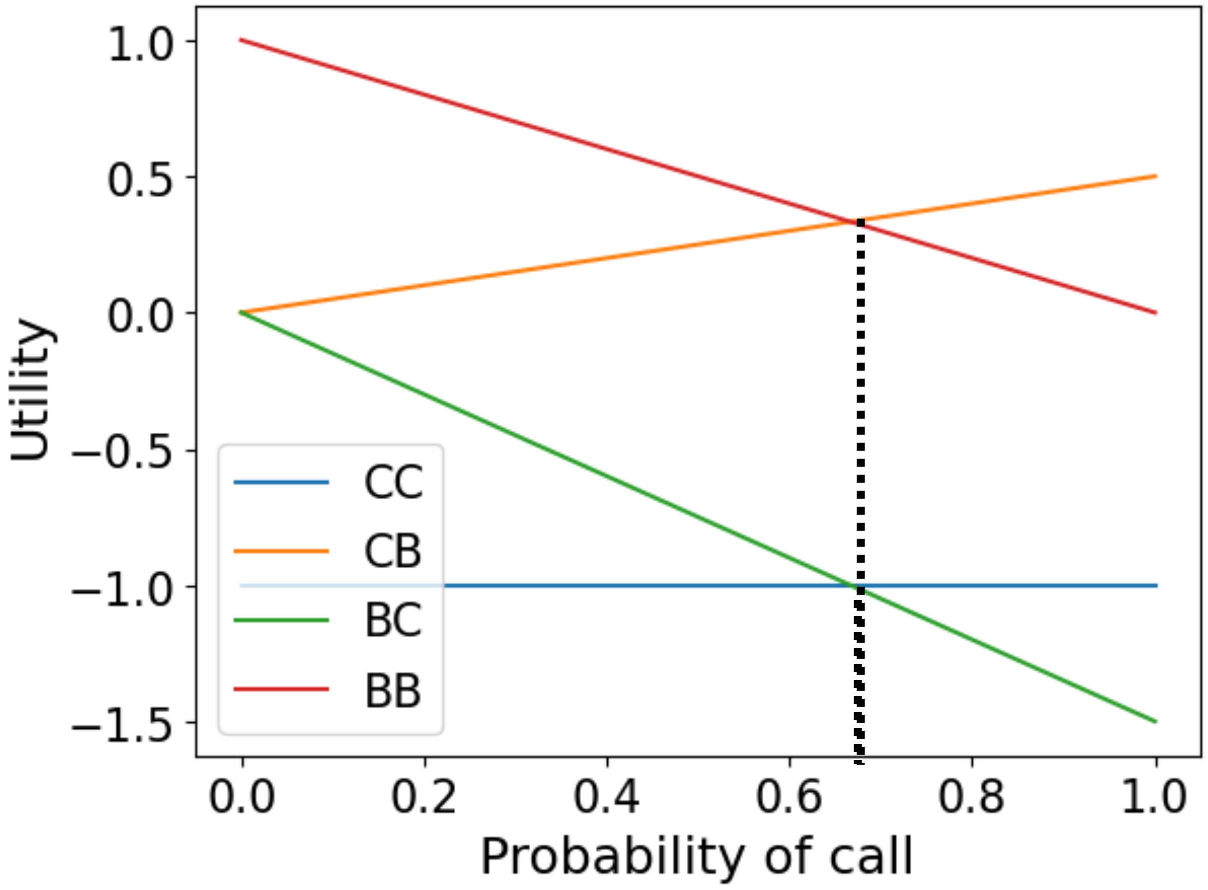
- What do you expect the value of the game to be?

Exercise: solution

	Strictly dominated J→Check K→Check	J→Check K→Bet	Strictly dominated by (B, B), weakly by (C, B) J→Bet K→Check	J→Bet K→Bet
Fold	-1	0	0	1
Call	-1	0.5	-1.5	0

Eliminate dominated actions

	J→Check K→Bet	J→Bet K→Bet
Fold	0	1
Call	0.5	0



Linear Programming

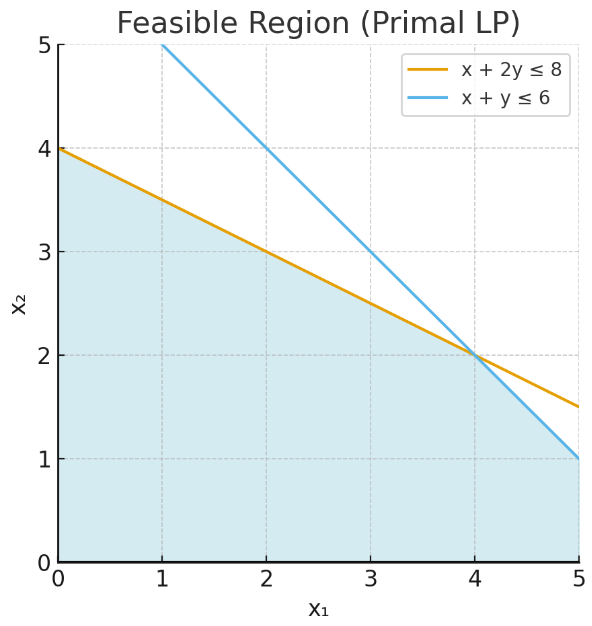
Linear Programming (Primal)

Generated this with GPT...

Optimize linear objective with linear constraints

Example:

- Maximize: $c^T x$
  Subject to: $Ax \leq b$
  $x \geq 0$



Many reformulations

- Maximize: $c^T x$
  Subject to: $Ax = b$
  $x \geq 0$

Linear Programming (Dual)

Every linear program has a "twin sibling" known as the *dual*

Example (primal)

Maximize $x^T c$

Subject to: $x^T 1 = 1$

$x \geq 0$

$\longleftrightarrow$

Example (dual)

Minimize t

Subject to: $t \geq c_i$ for all $i \in [n]$

Just pick highest value of c

Optimal values for primal and dual match

- Known as strong duality
- Feasible point in dual is a bound on primal value and vice versa

Systemic approaches exist to take duals

Super useful for proofs, optimization etc.

Game Solving via Linear Programming

"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**.“

- Von Neumann, 1947 to George Dantzig, on Farkas Lemma and Duality

Nash by Mathematical Programming

Bilinear Saddle-Point Problem (BSPP)

*Saddle-point problem

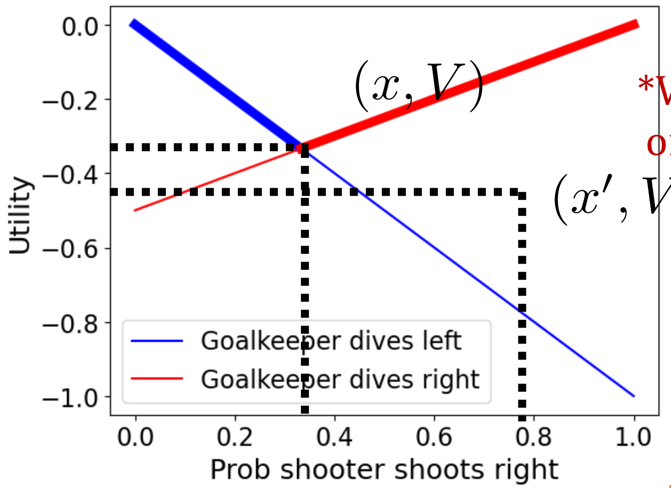
$$\min_{x \in \mathbb{R}^n} \max_{y \in \mathbb{R}^m} x^T A y$$

*Bilinear term

$$\text{such that } 1^T x = 1^T y = 1$$

$$x, y \geq 0$$

*Simplex constraints



*Violates the orange line!

Duals

Linear Programming

$$\min_{x \in \mathbb{R}^n, V \in \mathbb{R}} V$$

$$\text{such that } 1^T x = 1$$

$$A^T x \leq V$$

$$x \geq 0$$

*No matter what max-player does, min player cannot do worse than *V*

Obtained by taking dual of inner maximization problem

- V is the value of the game

To get y^*

- Resolve the LP by transposing A
- Alternatively, duals of constraint $P^T x \leq V$ give opponent distribution

Polyhedral strategy spaces [optional]

LP solver is generally quite slow

- For example, simplex methods require inversion of large matrix

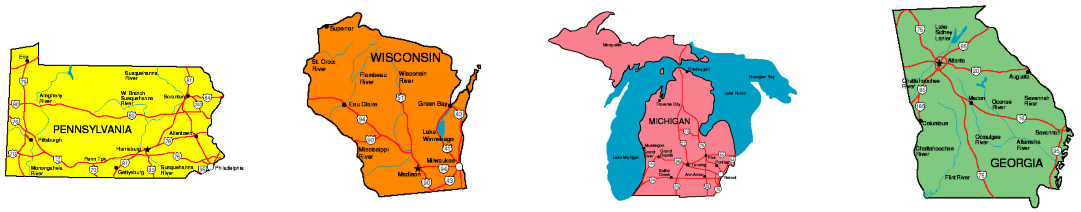
Not restricted to simplexes, allows for quite general constraints

- If $\mathcal{X}, \mathcal{Y}$ are polyhedral, replace Δ_n, Δ_m by them
- Examples:
  - Action a_5 must be played ≥17% of the time, but ≤62% →$0.17 \leq x(a_5) \leq 0.62$
  - a_1, a_4, a_7 played ≤34% of the time →$x(a_1) + x(a_4) + x(a_7) \leq 0.34$
  - a_1, a_2 must be played no less frequently than a_3, a_4→$x(a_1) + a(a_2) \geq x(a_3) + x(a_4)$
- Note: x, y *cannot have cross dependencies*, i.e., x cannot depend on y

Quite frequently, action spaces are much more structured, making for cleaner, more efficient formulations

- Paths, flows, circulations, matchings
- The teaser problem for this class!

Colonel Blotto Game



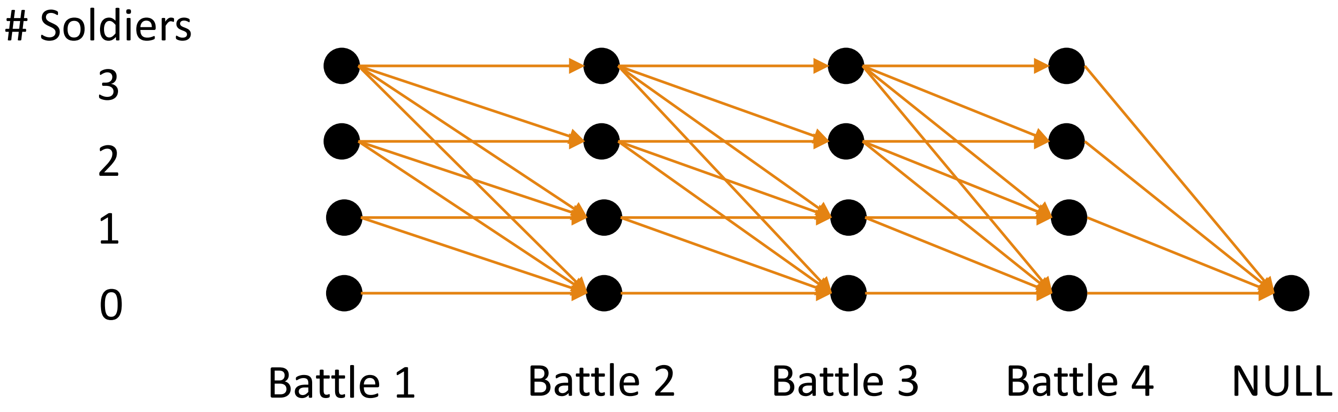
How should you allocate election funds?
Want to win, but no need to win by landslide

Allocating soldiers to battlefields

- b Battlefields, s soldiers per player
- Allocate soldiers between battlefields
- Battlefield i worth v_i, player with more soldiers allocated wins the battlefield

Number of pure strategies n, m are exponential in b

- Can solve in poly-time using LP



Arrows dictate expected number of "additional" soldiers are allocated given past number of soldiers used

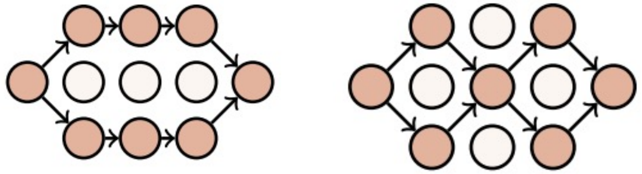
- Obeys flow-like structure

Can also solve by self-play via regret minimization on DAGs

Games played on DAGs

Professor pursuing a student over T steps

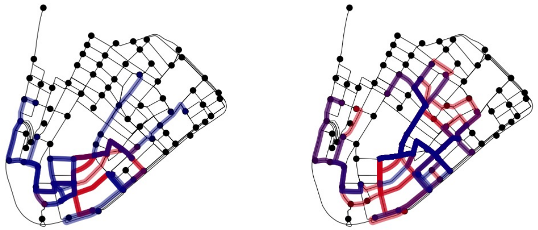
- **Each time** professor meets student, student will be assigned work
- Path taken is fixed once chosen
- Can be modeled by a DAG with diameter T



Payoffs are additively decomposable over edges

$$u_{\mathrm{LIN}}(p_d, p_a) = \sum_{e_d \in p_d} \sum_{e_a \in p_a} Q(e_d, e_a)$$

- Solvable in polytime, just like Blotto games



Example on a map of Manhattan with $T = 17$

If professor assigns work to a student at most once

$$u_{\mathrm{BIN}}(p_d, p_a) = r^{\odot}(p_a) \cdot \mathbb{1}\left[\sum_{e_d \in p_d} \sum_{e_a \in p_a} R(e_d, e_a) = 0\right]$$

- NP-hard to find Nash. Solution is not Markovian

*If both graphs are the same, then **empirically** we observe there always exists a Markovian strategy. Not sure why! Email me if you want to explore!