Plan for Today

- 2-player zero-sum games
- 2-player general-sum games
- Many-player games with structure:
  - congestion games / exact potential games
  - Best-response dynamics
  - Price of anarchy, Price of stability

2-Player Zero-Sum games

- Two players R and C. Zero-sum means that what's good for one is bad for the other.
- Game defined by matrix with a row for each of R's options and a column for each of C's options.
- Matrix tells who wins how much.
  - an entry \((x,y)\) means \(x = \) payoff to row player, \(y = \) payoff to column player. "Zero sum" means that \(y = -x\).
- E.g., penalty shot:

\[
\begin{array}{c|cc}
 & \text{Left} & \text{Right} \\
\hline
\text{Left} & (0,0) & (1,-1) \\
\text{Right} & (1,-1) & (0,0) \\
\end{array}
\]

Game Theory terminology

- Rows and columns are called pure strategies.
- Randomized algs called mixed strategies.
- "Zero sum" means that game is purely competitive. \((x,y)\) satisfies \(x+y=0\). (Game doesn't have to be fair).

Minimax-optimal strategies

- Minimax optimal strategy is a (randomized) strategy that has the best guarantee on its expected gain, over choices of the opponent. [maximizes the minimum]
- I.e., the thing to play if your opponent knows you well.

Minimax-optimal strategies

- Can solve for minimax-optimal strategies using Linear programming
- No-regret strategies will do nearly as well or better.
- I.e., the thing to play if your opponent knows you well.
Minimax Theorem (von Neumann 1928)
- Every 2-player zero-sum game has a unique value $V$.
- Minimax optimal strategy for R guarantees R's expected gain at least $V$.
- Minimax optimal strategy for C guarantees C's expected loss at most $V$.

Existence of no-regret strategies gives one way of proving the theorem.

Can use notion of minimax optimality to explain bluffing in poker

Simplified Poker (Kuhn 1950)
- Two players A and B.
- Deck of 3 cards: 1,2,3.
- Players ante $1.
- Each player gets one card.
- A goes first. Can bet $1 or pass.
  - If A bets, B can call or fold.
  - If A passes, B can bet $1 or pass.
    - If B bets, A can call or fold.
- High card wins (if no folding). Max pot $2.

Writing as a Matrix Game
- For a given card, A can decide to
  - Pass but fold if B bets. [PassFold]
  - Pass but call if B bets. [PassCall]
  - Bet. [Bet]
- Similar set of choices for B.

Can look at all strategies as a big matrix...

<table>
<thead>
<tr>
<th>[FP,FP,CB]</th>
<th>[FP,CP,CB]</th>
<th>[FB,FP,CB]</th>
<th>[FB,CP,CB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$PF,PF,B$</td>
<td>0</td>
<td>0</td>
<td>-1/6</td>
</tr>
<tr>
<td>$PF,PF,B$</td>
<td>0</td>
<td>1/6</td>
<td>-1/3</td>
</tr>
<tr>
<td>$PF,PC,P$</td>
<td>-1/6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$PF,PC,B$</td>
<td>-1/6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$B,PF,P$</td>
<td>1/6</td>
<td>-1/3</td>
<td>0</td>
</tr>
<tr>
<td>$B,PF,B$</td>
<td>1/6</td>
<td>-1/6</td>
<td>-1/6</td>
</tr>
<tr>
<td>$B,PC,P$</td>
<td>0</td>
<td>-1/2</td>
<td>1/3</td>
</tr>
<tr>
<td>$B,PC,B$</td>
<td>0</td>
<td>-1/3</td>
<td>1/6</td>
</tr>
</tbody>
</table>

And the minimax optimal strategies are...
- A:
  - If hold 1, then 5/6 PassFold and 1/6 Bet.
  - If hold 2, then 1/2 PassFold and 1/2 PassCall.
  - If hold 3, then 1/2 PassCall and 1/2 Bet.
  - Has both bluffing and underbidding...
- B:
  - If hold 1, then 2/3 FoldPass and 1/3 FoldBet.
  - If hold 2, then 2/3 FoldPass and 1/3 CallPass.
  - If hold 3, then CallBet

Minimax value of game is $-1/18$ to A.
Now, to General-Sum games...

General-sum games

• In general-sum games, can get win-win and lose-lose situations.
• E.g., "what side of sidewalk to walk on?"

<table>
<thead>
<tr>
<th></th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>(1,1)</td>
<td>(-1,-1)</td>
</tr>
<tr>
<td>Right</td>
<td>(-1,-1)</td>
<td>(1,1)</td>
</tr>
</tbody>
</table>

E.g., "what side of sidewalk to walk on?"

<table>
<thead>
<tr>
<th></th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>(1,1)</td>
<td>(-1,-1)</td>
</tr>
<tr>
<td>Right</td>
<td>(-1,-1)</td>
<td>(1,1)</td>
</tr>
</tbody>
</table>

Person walking towards you

Economists use games and equilibria as models of interaction.

• E.g., pollution / prisoner’s dilemma:
  - (imagine pollution controls cost $4 but improve everyone’s environment by $3)

<table>
<thead>
<tr>
<th></th>
<th>don’t pollute</th>
<th>pollute</th>
</tr>
</thead>
<tbody>
<tr>
<td>don’t pollute</td>
<td>(2,2)</td>
<td>(-1,3)</td>
</tr>
<tr>
<td>pollute</td>
<td>(3,-1)</td>
<td>(0,0)</td>
</tr>
</tbody>
</table>

Need to add extra incentives to get good overall behavior.

Nash Equilibrium

• A Nash Equilibrium is a stable pair of strategies (could be randomized).
• Stable means that neither player has incentive to deviate on their own.
• E.g., "what side of sidewalk to walk on?"

<table>
<thead>
<tr>
<th></th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>(1,1)</td>
<td>(-1,-1)</td>
</tr>
<tr>
<td>Right</td>
<td>(-1,-1)</td>
<td>(1,1)</td>
</tr>
</tbody>
</table>

NE are: both left, both right, or both 50/50.

NE can do strange things

• Braess paradox:
  - Road network, traffic going from s to t.
  - travel time as function of fraction x of traffic on a given edge.

Fine. NE is 50/50. Travel time = 1.5
NE can do strange things

- Braess paradox:
  - Road network, traffic going from s to t.
  - Travel time as function of fraction x of traffic on a given edge.

Add new superhighway. NE: everyone uses zig-zag path. Travel time = 2.

Existence of NE

- Nash (1950) proved: any general-sum game must have at least one such equilibrium.
  - Might require mixed strategies.
- This also yields minimax thm as a corollary.
  - Pick some NE and let V = value to row player in that equilibrium.
  - Since it's a NE, neither player can do better even knowing the (randomized) strategy their opponent is playing.
  - So, they're each playing minimax optimal.

Existence of NE in 2-player games

- Proof will be non-constructive.
- Unlike case of zero-sum games, we do not know any polynomial-time algorithm for finding Nash Equilibria in \( n \times n \) general-sum games. [known to be "PPAD-hard"]

Notation:
- Assume an \( n \times n \) matrix.
- Use \( (p_1, \ldots, p_n) \) to denote mixed strategy for row player, and \( (q_1, \ldots, q_n) \) to denote mixed strategy for column player.

Proof

- We'll start with Brouwer's fixed point theorem.
  - Let \( S \) be a compact convex region in \( \mathbb{R}^n \) and let \( f:S \rightarrow S \) be a continuous function.
  - Then there must exist \( x \in S \) such that \( f(x)=x \).
  - \( x \) is called a "fixed point" of \( f \).
- Simple case: \( S \) is the interval \([0,1]\).
- We will care about:
  - \( S = \{(p,q): p,q \text{ are legal probability distributions on } 1,\ldots,n\} \). I.e., \( S = \text{simplex}_n \times \text{simplex}_n \)

Proof (cont)

- \( S = \{(p,q): p,q \text{ are mixed strategies}\} \).
- Want to define \( f(p,q) = (p',q') \) such that:
  - \( f \) is continuous. This means that changing \( p \) or \( q \) a little bit shouldn't cause \( p' \) or \( q' \) to change a lot.
  - Any fixed point of \( f \) is a Nash Equilibrium.
- Then Brouwer will imply existence of NE.

Try #1

- What about \( f(p,q) = (p',q') \) where \( p' \) is best response to \( q \), and \( q' \) is best response to \( p \)?
- Problem: not necessarily well-defined:
  - E.g., penalty shot: if \( p = (0.5,0.5) \) then \( q' \) could be anything.

<table>
<thead>
<tr>
<th></th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>(0,0)</td>
<td>(1,-1)</td>
</tr>
<tr>
<td>Right</td>
<td>(1,-1)</td>
<td>(0,0)</td>
</tr>
</tbody>
</table>
Try #1

- What about \( f(p,q) = (p',q') \) where \( p' \) is best response to \( q \), and \( q' \) is best response to \( p \)?
- Problem: also not continuous:
  - E.g., if \( p = (0.51, 0.49) \) then \( q' = (1,0) \). If \( p = (0.49, 0.51) \) then \( q' = (0,1) \).

<table>
<thead>
<tr>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0)</td>
<td>(1,1)</td>
</tr>
<tr>
<td>(1,1)</td>
<td>(0,0)</td>
</tr>
</tbody>
</table>

Instead we will use...

- \( f(p,q) = (p',q') \) such that:
  - \( q' \) maximizes \( (\text{expected gain wrt} \ p) - ||q-q'||^2 \)
  - \( p' \) maximizes \( (\text{expected gain wrt} \ q) - ||p-p'||^2 \)

\[ p \quad p' \]

Note: quadratic + linear = quadratic.

Instead we will use...

- \( f(p,q) = (p',q') \) such that:
  - \( q' \) maximizes \( (\text{expected gain wrt} \ p) - ||q-q'||^2 \)
  - \( p' \) maximizes \( (\text{expected gain wrt} \ q) - ||p-p'||^2 \)

\[ f \]

Note: quadratic + linear = quadratic.

Instead we will use...

- \( f(p,q) = (p',q') \) such that:
  - \( q' \) maximizes \( (\text{expected gain wrt} \ p) - ||q-q'||^2 \)
  - \( p' \) maximizes \( (\text{expected gain wrt} \ q) - ||p-p'||^2 \)

- \( f \) is well-defined and continuous since quadratic has unique maximum and small change to \( p,q \) only moves this a little.
- Also fixed point = NE. (even if tiny incentive to move, will move little bit).
- So, that’s it!

What if all players minimize regret?

- In zero-sum games, empirical frequencies quickly approaches minimax optimal.
- In general-sum games, does behavior quickly (or at all) approach a Nash equilibrium? (after all, a Nash Eq is exactly a set of distributions that are no-regret wrt each other).
- Well, unfortunately, no.
- (Even if it did, as we saw last time, you might not want to minimize regret in order to get other players to do what you want – e.g., ultimatum game)
A bad example for general-sum games

- Augmented Shapley game from [Z04]: “RPSF”
  - First 3 rows/cols are Shapley game (rock / paper / scissors but if both do same action then both lose).
  - 4th action “play foosball” has slight negative if other player is still doing r/p/s but positive if other player does 4th action too.
  - NR alg will cycle among first 3 and have no regret, but do worse than only Nash Equilibrium of both playing foosball.
- We didn’t really expect this to work given how hard NE can be to find...

More general forms of regret
1. “best expert” or “external” regret:
   - Given n strategies. Compete with best of them in hindsight.
2. “sleeping expert” or “regret with time-intervals”:
   - Given n strategies, k properties. Let \( S_i \) be set of days satisfying property i (might overlap). Want to simultaneously achieve low regret over each \( S_i \).
3. “internal” or “swap” regret: like (2), except that \( S_i \) = set of days in which we chose strategy i.

Internal/swap-regret
“Correlated equilibrium”
- Distribution over entries in matrix, such that if a trusted party chooses one at random and tells you your part, you have no incentive to deviate.
- E.g., Shapley game.

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>P</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>-1,1</td>
<td>-1,1</td>
<td>1,-1</td>
</tr>
<tr>
<td>P</td>
<td>1,-1</td>
<td>-1,1</td>
<td>-1,1</td>
</tr>
<tr>
<td>S</td>
<td>-1,1</td>
<td>1,-1</td>
<td>-1,1</td>
</tr>
</tbody>
</table>

What can we say?
- If algorithms minimize “internal” or “swap” regret, then empirical distribution of play approaches correlated equilibrium.
- Foster & Vohra, Hart & Mas-Colell, ...
- Though doesn’t imply play is stabilizing.

What are internal regret and correlated equilibria?

Internal/swap-regret
- E.g., each day we pick one stock to buy shares in.
  - Don’t want to have regret of the form “every time I bought IBM, I should have bought Microsoft instead”.
- Formally, regret is wrt optimal function \( f : \{1, \ldots, N\} \rightarrow \{1, \ldots, N\} \)
  - So, this says that correlated equilibria are a natural thing to see in multi-agent systems where individuals are optimizing for themselves
**Internal/swap-regret, contd**

Algorithms for achieving low regret of this form:
- Foster & Vohra, Hart & Mas-Colell, Fudenberg & Levine.
- Can also convert any "best expert" algorithm into one achieving low swap regret.
- Unfortunately, time to achieve low regret is linear in $n$ rather than $\log(n)$...

**Congestion games**

- Many multi-agent interactions have structure. One nice class: Congestion Games
- Always have a pure-strategy equilibrium.
- Have a potential function s.t. whenever a player switches, potential drops by exactly that player's improvement.
  - So, best-response dynamics always gives an equilibrium.
- Let's start with an example.

**Internal/swap-regret, contd**

Can convert any "best expert" algorithm $A$ into one achieving low swap regret. Idea:
- Instantiate one copy $A_i$ responsible for expected regret over times we play $i$.
- Each time step, if we play $p=(p_1,\ldots,p_n)$ and get cost vector $c=(c_1,\ldots,c_n)$, then $A_i$ gets cost-vector $p_i c$.
- If each $A_i$ proposed to play $q_i$, so all together we have matrix $Q$, then define $p = pQ$.
- Allows us to view $p_i$ as prob we chose action $i$ or prob we chose algorithm $A_i$.

**Fair cost-sharing**

Fair cost-sharing: $n$ players in weighted directed graph $G$.
Player $i$ wants to get from $s_i$ to $t_i$, and they share cost of edges they use with others.

**Good equilibria, Bad equilibria**

Fair cost-sharing: $n$ players in weighted directed graph $G$.
Player $i$ wants to get from $s_i$ to $t_i$, and they share cost of edges they use with others.

- **Good equilibrium**: all use edge of cost 1. (cost $1/n$ per player)
- **Bad equilibrium**: all use edge of cost $n$. (cost 1 per player)
- Cost(bad equilib) = $n \cdot$ Cost(good equilib)
Price of Anarchy and Price of Stability

- **Price of Anarchy**: ratio of worst equilibrium to social optimum. (worst-case over games in class)
  - We saw for cost-sharing PoA = Ω(n). Also O(n).
- **Price of Stability**: ratio of best equilibrium to social optimum. (worst-case over games in class)
  - For cost-sharing, PoS = O(log n).
- **Exact Potential function**: Function $\Phi$ s.t. if player moves, potential changes by exactly as much as cost of player who moved.
  - Guarantees that best-response dynamics will reach Nash equilibrium

Potential functions and PoS
For cost-sharing, PoS = O(log n):
- Given state $S$, let $n_e$ = # players on edge $e$. Cost($S$) =
- Define potential $\phi(S) = \ldots$
- So, cost($S$) \(\leq\) $\phi(S)$ \(\leq\) $\log(n) \times$ cost($S$).
- Now consider best-response dynamics starting from OPT. $\phi$ can only decrease.

Congestion games more generally
Game defined by $n$ players and $m$ resources.
- Each player $i$ chooses a set of resources (e.g., a path) from collection $S_i$ of allowable sets of resources (e.g., paths from $s_i$ to $t_i$).
- Cost of a resource $j$ is a function $f_j(n_j)$ of the number $n_j$ of players using it.
- Cost incurred by player $i$ is the sum, over all resources being used, of the cost of the resource.
- Generic potential function:

- Best-response dynamics may take a long time to reach equil, but if gap between $\phi$ and cost is small, can get to apx-equilib fast.