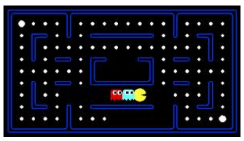
Games: simultaneous games



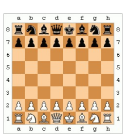
answer in chat

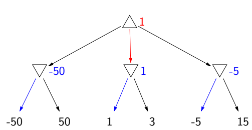
Question

For a simultaneous two-player zero-sum game (like rock-paper-scissors), can you still be optimal if you reveal your strategy?

- yes
- no

Turn-based games:





Simultaneous games:

?

- Game trees were our primary tool to model turn-based games. However, in simultaneous games, there is no ordering on the player's moves, so we need to develop new tools to model these games. Later, we will see that game trees will still be valuable in understanding simultaneous games.

Two-finger Morra

Example: two-finger Morra

Players A and B each show 1 or 2 fingers.
If both show 1, B gives A 2 dollars.
If both show 2, B gives A 4 dollars.
Otherwise, A gives B 3 dollars.

[play with a partner]

answer in chat

Question

What was the outcome?

- player A chose 1, player B chose 1
- player A chose 1, player B chose 2
- player A chose 2, player B chose 1
- player A chose 2, player B chose 2

Payoff matrix

Definition: single-move simultaneous game

Players $= \{\text{A}, \text{B}\}$
Actions: possible actions
$V(a, b)$: **A's utility** if A chooses action a, B chooses b
(let V be **payoff matrix**)

Example: two-finger Morra payoff matrix

A \ B	1 finger	2 fingers
1 finger	2	-3
2 fingers	-3	4

- In this lecture, we will consider only single move games. There are two players, A and B who both select from one of the available actions. The value or utility of the game is captured by a payoff matrix V whose dimensionality is $|\text{Actions}| \times |\text{Actions}|$. We will be analyzing everything from A's perspective, so entry $V(a, b)$ is the utility that A gets if he/she chooses action a and player B chooses b.

Strategies (policies)

Definition: pure strategy

A pure strategy is a single action:
$a \in$ Actions

Definition: mixed strategy

A mixed strategy is a probability distribution
$0 \leq \pi(a) \leq 1$ for $a \in$ Actions

Example: two-finger Morra strategies

Always 1: $\pi = [1, 0]$
Always 2: $\pi = [0, 1]$
Uniformly random: $\pi = [\frac{1}{2}, \frac{1}{2}]$

- Each player has a **strategy** (or a policy). A pure strategy (deterministic policy) is just a single action. Note that there's no notion of state since we are only considering single-move games.
- More generally, we will consider **mixed strategies** (randomized policy), which is a probability distribution over actions. We will represent a mixed strategy π by the vector of probabilities.

Game evaluation

Definition: game evaluation

The **value** of the game if player A follows π_{A} and player B follows π_{B} is

$$V(\pi_{\mathsf{A}}, \pi_{\mathsf{B}}) = \sum_{a,b} \pi_{\mathsf{A}}(a)\pi_B(b)V(a,b)$$

Example: two-finger Morra

Player A always chooses 1: $\pi_{\mathsf{A}} = [1, 0]$
Player B picks randomly: $\pi_{\mathsf{B}} = [\frac{1}{2}, \frac{1}{2}]$
Value: $\boxed{-\frac{1}{2}}$

[whiteboard: matrix]

- Given a game (payoff matrix) and the strategies for the two players, we can define the value of the game.
- For pure strategies, the value of the game by definition is just reading out the appropriate entry from the payoff matrix.
- For mixed strategies, the value of the game (that is, the expected utility for player A) is gotten by summing over the possible actions that the players choose: $V(\pi_{\mathsf{A}}, \pi_{\mathsf{B}}) = \sum_{a \in \mathsf{Actions}} \sum_{b \in \mathsf{Actions}} \pi_{\mathsf{A}}(a)\pi_{\mathsf{B}}(b)V(a,b)$. We can also write this expression concisely using matrix-vector multiplications: $\pi_A^\top V \pi_B$.

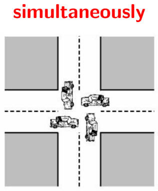
How to optimize?

Game value:

$$V(\pi_A, \pi_B)$$

Challenge: player A wants to maximize, player B wants to minimize...

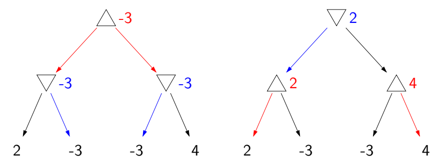
simultaneously



- Having established the values of fixed policies, let's try to optimize the policies themselves. Here, we run into a predicament: player A wants to maximize V but player B wants to minimize V **simultaneously**.
- Unlike turn-based games, we can't just consider one at a time. But let's consider the turn-based variant anyway to see where it leads us.

Pure strategies: who goes first?

Player A goes first: Player B goes first:



Proposition: going second is no worse

$$\max_a \min_b V(a,b) \le \min_b \max_a V(a,b)$$

- Let us first consider pure strategies, where each player just chooses one action. The game can be modeled by using the standard minimax game trees that we're used to.
- The main point is that if player A goes first, he gets -3, but if he goes second, he gets 2. In general, it's at least as good to go second, and often it is strictly better. This is intuitive, because seeing what the first player does gives more information.

Mixed strategies

Example: two-finger Morra

Player A reveals: $\pi_{\mathsf{A}} = [\frac{1}{2}, \frac{1}{2}]$

Value $V(\pi_{\mathsf{A}}, \pi_B) = \pi_B(1)(-\frac{1}{2}) + \pi_B(2)(+\frac{1}{2})$

Optimal strategy for player B is $\pi_B = [1, 0]$ **(pure!)**

Proposition: second player can play pure strategy

For any fixed mixed strategy π_{A}:

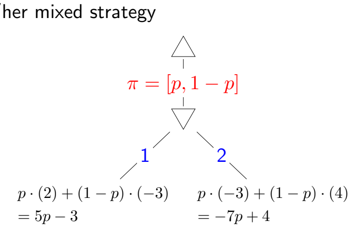
$$\min_{\pi_{\mathsf{B}}} V(\pi_{\mathsf{A}}, \pi_B)$$

can be attained by a pure strategy.

- Now let us consider mixed strategies. First, let's be clear on what playing a mixed strategy means. If player A chooses a mixed strategy, he reveals to player B the full probability distribution over actions, but importantly not a particular action (because that would be the same as choosing a pure strategy).
- As a warmup, suppose that player A reveals $\pi_{\mathsf{A}} = [\frac{1}{2}, \frac{1}{2}]$. If we plug this strategy into the definition for the value of the game, we will find that the value is a convex combination between $\frac{1}{2}(2) + \frac{1}{2}(-3) = -\frac{1}{2}$ and $\frac{1}{2}(-3) + \frac{1}{2}(4) = \frac{1}{2}$. The value of π_B that minimizes this value is $[1, 0]$. The important part is that this is a **pure strategy**.
- It turns out that no matter what the payoff matrix V is, as soon as π_{A} is fixed, then the optimal choice for π_{B} is a pure strategy. Ths is useful because it will allow us to analyze games with mixed strategies more easily.

Mixed strategies

Player A first reveals his/her mixed strategy



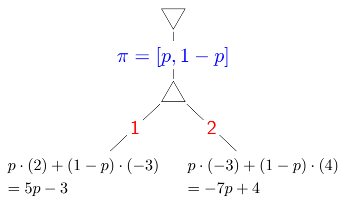
Minimax value of game:

$$\max_{0 \le p \le 1} \min\{5p - 3, -7p + 4\} = \boxed{-\frac{1}{12}} \text{ (with } p = \tfrac{7}{12}\text{)}$$

- Now let us try to draw the minimax game tree where the player A first chooses a mixed strategy, and then player B chooses a pure strategy.
- There are an uncountably infinite number of mixed strategies for player A, but we can summarize all of these actions by writing a single action template $\pi = [p, 1-p]$.
- Given player A's action, we can compute the value if player B either chooses 1 or 2. For example, if player B chooses 1, then the value of the game is $5p - 3$ (with probability p, player A chooses 1 and the value is 2; with probability $1-p$ the value is -3). If player B chooses action 2, then the value of the game is $-7p + 4$.
- The value of the min node is $F(p) = \min\{5p - 3, -7p + 4\}$. The value of the max node (and thus the minimax value of the game) is $\max_{0 \le 1 \le p} F(p)$.
- What is the best strategy for player A then? We just have to find the p that maximizes $F(p)$, which is the minimum over two linear functions of p. If we plot this function, we will see that the maximum of $F(p)$ is attained when $5p - 3 = -7p + 4$, which is when $p = \frac{7}{12}$. Plugging that value of p back in yields $F(p) = -\frac{1}{12}$, the minimax value of the game if player A goes first and is allowed to choose a mixed strategy.
- Note that if player A decides on $p = \frac{7}{12}$, it doesn't matter whether player B chooses 1 or 2; the payoff will be the same: $-\frac{1}{12}$. This also means that whatever mixed strategy (over 1 and 2) player B plays, the payoff would also be $-\frac{1}{12}$.

Mixed strategies

Player B first reveals his/her mixed strategy



Minimax value of game:

$$\min_{p \in [0,1]} \max\{5p-3, -7p+4\} = \boxed{-\frac{1}{12}} \text{ (with } p = \tfrac{7}{12}\text{)}$$

- Now let us consider the case where player B chooses a mixed strategy $\pi = [p, 1-p]$ first. If we perform the analogous calculations, we'll find that we get that the minimax value of the game is exactly the same $(-\frac{1}{12})$!
- Recall that for pure strategies, there was a gap between going first and going second, but here, we see that for mixed strategies, there is no such gap, at least in this example.
- Here, we have been computed minimax values in the conceptually same manner as we were doing it for turn-based games. The only difference is that our actions are mixed strategies (represented by a probability distribution) rather than discrete choices. We therefore introduce a variable (e.g., p) to represent the actual distribution, and any game value that we compute below that variable is a function of p rather than a specific number.

General theorem

Theorem: minimax theorem [von Neumann, 1928]

For every simultaneous two-player zero-sum game with a finite number of actions:

$$\max_{\pi_A} \min_{\pi_B} V(\pi_A, \pi_B) = \min_{\pi_B} \max_{\pi_A} V(\pi_A, \pi_B),$$

where π_A, π_B range over **mixed strategies**.

Upshot: revealing your optimal mixed strategy doesn't hurt you!

Proof: linear programming duality

Algorithm: compute policies using linear programming

- It turns out that having no gap is not a coincidence, and is actually one of the most celebrated mathematical results: the von Neumann minimax theorem. The theorem states that for any simultaneous two-player zero-sum game with a finite set of actions (like the ones we've been considering), we can just swap the min and the max: it doesn't matter which player reveals his/her strategy first, as long as their strategy is optimal. This is significant because we were stressing out about how to analyze the game when two players play simultaneously, but now we find that both orderings of the players yield the same answer. It is important to remember that this statement is true only for mixed strategies, not for pure strategies.
- This theorem can be proved using linear programming duality, and policies can be computed also using linear programming. The sketch of the idea is as follows: recall that the optimal strategy for the second player is always deterministic, which means that the $\max_{\pi_A} \min_{\pi_B} \cdots$ turns into $\max_{\pi_A} \min_b \cdots$. The min is now over n actions, and can be rewritten as n linear constraints, yielding a linear program.
- As an aside, recall that we also had a minimax result for turn-based games, where the max and the min were over agent and opponent policies, which map states to actions. In that case, optimal policies were always deterministic because at each state, there is only one player choosing.

Summary

- Challenge: deal with simultaneous min/max moves
- Pure strategies: going second is better
- Mixed strategies: doesn't matter (von Neumann's minimax theorem)