Advanced Algorithms

Lecture 24: Game Theory and Nash Equilibria

Lecturer: Nathan Klein

1 Game Theory

Game theory studies strategic interactions between entities. In the simplest kinds of games, each player has a set of actions and knows a *payoff matrix*. In each entry of the matrix, there is a tuple (a, b) which indicates that in this outcome the row player receives a units of utility (let's call them utils as they often are) and the column player receives b utils. These numbers are not necessarily non-negative.

A classic example is the Prisoner's Dilemma. Here, there are two suspects for a crime deciding whether to confess or not. They are held in separate rooms. If one confesses and the other does not, the one that confesses gets no jail time and the other one has a long sentence. If they both confess, they both get medium sentences. If neither confesses, they both get short sentences.

	Confess	Don't Confess
Confess	(2, 2)	(5, 0)
Don't Confess	(0, 5)	(3, 3)

Table 1: Payoff matrix for the prisoner's dilemma.

1.1 Nash Equilibria

A **Nash equilibrium** for a game with a payoff matrix is a pair of strategies, one for each player, which neither one can unilaterally improve. In other words, even if each player *knows* the strategy of the other, there is no way to improve their expected payoff by changing their strategy. So, it's a "fixed point," a situation where neither player will change their mind.

Nash equilibria can either be **pure** or **mixed**. An equilibrium is pure if the strategy involves playing one move. It is mixed if the players randomize over their moves. In the prisoner's dilemma, there is exactly one Nash equilibrium, and it is pure: both players confess. Why is this the only Nash? It's because it's *always* better to confess, no matter what your opponent plays. If they confess, you get 2 utils if you confess and 0 otherwise, so you would prefer to confess. If they don't confess, you get 5 units if you confess and 3 otherwise.

The prisoner's dilemma is often used to demonstrate the importance of cooperation. A particularly clean example of this is the use of helmets in hockey. Robert Schelling, in his book "Micromotives and Macrobehavior" from the 70s, described the situation as follows. If hockey players are allowed to skate without helmets, they will: this was the case until 1979, when incoming players were required to wear helmets (existing players were allowed to not wear helmets and some continued to do so until the 90s). Before this law, it was well understood that hockey players who didn't wear helmets were much more prone to injury. And yet few players wore one, as without a helmet, you could see better and move more freely. So you increased your odds of winning. This gave players without helmets an edge over those who wear helmets. So players that valued winning enough would take off their helmets. And most did.

But at that point, no one has a helmet. So the risk of injury is high, and no one gets a competitive advantage. Let's model the payoff matrix here.

	No helmet	Helmet
No helmet	(2, 2)	(5, 1)
Helmet	(1, 5)	(4, 4)

Table 2: Payoff matrix for the helmet problem.

If no one wears a helmet, everyone's utility is low: injuries and no competitive advantage. Clearly the competitive advantage was worth a lot to the players, so we put this as $(1, 5)$, and if both players wear a helmet this is $(4, 4)$. This is exactly the prisoner's dilemma, I just slightly changed some of the numbers. Again the pure Nash equilibrium is for both players not to wear a helmet.

Now wouldn't it be nice if we could just make a rule and agree to wear helmets, locking in the highest utility of $(4, 4)$ instead of $(2, 2)$? Yes, and that's what was done. Some economists argue that situations like give us a window into the issues of a completely free-market economy, and why some regulation can increase everyone's utility in the long run.

1.2 Zero-Sum Games

The simplest types of games are zero-sum, where one player's utility results in an equal loss of the other player's utility. In these situations the payoff matrix contains just one number in each entry, which is the amount of utility gained or lost by the row player.

	Rock	Paper	Scissors
Rock	0	-1	1
Paper	1	0	-1
Scissors	-1	1	0

Table 3: Payoff matrix for rock, paper, scissors.

Let A be the payoff matrix for a zero-sum game. Then a pure Nash equilibrium would be one where:

$$\max_i \min_j A_{ij} = \min_j \max_i A_{ij}$$

For the LHS, the row player will announce their strategy first. They know that the column player will be strategic, so they want to choose the row which has the highest minimum. In this game that is -1. For the RHS, the column player will announce first, and then the row player will choose the max value in that column. In this game that is 1. So, there is no pure Nash here.

However, it does have a mixed Nash: play each of rock, paper, and scissors with probability $\frac{1}{3}$. Then the expected utility is 0 for *any* strategy your opponent plays, so it's impossible to exploit.

There is a theorem we will prove shortly:

Theorem 1.1 (von Neumann's Minimax Theorem)**.** *Any zero sum game has a mixed Nash equilibria. In other words, given a payoff matrix $A \in \mathbb{R}^{m \times n}$, if $X = \{x \in \mathbb{R}^m_{\geq 0} \mid \|x\|_1 = 1\}$ and $Y = \{y \in \mathbb{R}^n_{\geq 0} \mid \|y\| = 1\}$, we have:*

$$\max_{x \in X} \min_{y \in Y} x^T A y = \min_{y \in Y} \max_{x \in X} x^T A y$$

So, we can find a mixed Nash for any zero-sum game, even one as complex as poker. In fact the "game theory optimal" (GTO) strategy has been computed up to a small error for two-player Texas hold 'em with a limit on bet size. By GTO strategy I mean one side of a mixed Nash, as it is a strategy that cannot be beat, even though of course it will have expected payoff 0 for Poker and all fair games against itself. Near GTO strategies have been computed for a range of Poker settings.

1.3 The Minimax Theorem

We will use linear programming duality to prove the minimax theorem. We want to show that given a payoff matrix $A \in \mathbb{R}^{m\times n}$, if $X = \{x \in \mathbb{R}^m_{\geq 0} \mid \|x\|_1 = 1\}$ and $Y = \{y \in \mathbb{R}^n_{\geq 0} \mid \|y\| = 1\}$, we have:

$$\max_{x\in X} \min_{y\in Y} x^T Ay = \min_{y\in Y} \max_{x\in X} x^T Ay$$

First notice that $x^T Ay$ corresponds to what we intend: it is $\sum_{i,j} x_i y_i A_{ij}$. This is the probability we land in state (i, j) and this yields utility A_{ij} for the row player and $-A_{ij}$ for the column player.

Given a fixed x, y will then solve $\min_{y\in Y} x^T Ay$. But $x^T A$ is just a vector. To maximize the inner product with $x^T A$, y can WLOG choose a single index and put all of its mass there. So, if $a^{(1)}, \ldots, a^{(n)}$ are the rows of A:

$$\min_{y\in Y} x^T Ay = \min_{i\in[n]} (x^T A)_i = \min_{i\in[n]} x^T a^{(i)}$$

Therefore an equivalent problem is:

$$\max_{x\in X} \min_{i\in[n]} x^T a^{(i)}$$

This now can be written as an LP, where WLOG we assume the game has positive value. (This allows us to drop the equality on the sum of the x_i and assume $x_{n+1} \geq 0$.)

$$\begin{aligned} \max \quad & x_{n+1} \\ \text{s.t.} \quad & x^T A \geq x_{n+1} \\ & \sum_{i=1}^m x_i \leq 1 \\ & x \in \mathbb{R}^{m+1}_{\geq 0} \end{aligned}$$

Let's put things in standard form.

$$\begin{aligned} \max \quad & x_{n+1} \\ \text{s.t.} \quad & x_{n+1} - A^T x \leq 0 \\ & \sum_{i=1}^m x_i \leq 1 \\ & x \in \mathbb{R}^{m+1}_{\geq 0} \end{aligned}$$

This can be modeled using a big matrix $\tilde{A}$, which is made up of $-A^T$ plus a row of all 1s appended on the right and a row of all 1s appended on the bottom with a 0 as the last coordinate. Now we can use the form $\max c^T x$ subject to $\tilde{A}x \leq b$, where c is all 0s except a 1 in the last coordinate and b is all 0s except a 1 in the last coordinate.

But now, take the dual of this LP. Translating from $\min b^T y$ subject to $\tilde{A}^T y \geq c$ we get:

$$\begin{aligned}
\min \quad & y_{n+1} \\
\text{s.t.} \quad & y_{n+1} - Ay \geq 0 \\
& \sum_{i=1}^{m} y_i \geq 1 \\
& y \in \mathbb{R}_{\geq 0}^{n+1}
\end{aligned}$$

But by the same logic as above, this is exactly calculating the RHS of the minimax theorem. It just says to pick the vector y so that the smallest entry of Ay is as small as possible. By strong duality these values are equal, giving the theorem.

1.4 Nash's Theorem

While we proved this for zero-sum games, famously, John Nash proved that there are mixed equilibria for all games, whether or not the payoff matrix is zero-sum.

Theorem 1.2 (Nash Existence Theorem)**.** *Every finite game has at least one (possibly mixed) Nash equilibrium.*

It has a very nice proof using Brouwer's fixed point theorem. We will not show it in this class.

1.5 More on GTO in Poker

In rock paper scissors, it may not make sense to play the GTO strategy. If you think that your opponent is very likely to play rock, you should play paper more often. However, this then opens a window for you to be exploited, as you are playing paper too often.

The GTO strategy in rock paper scissors is a bit boring because it will *never* beat an expected utility of 0, even against an inexperienced opponent. In Poker, while exploiting is also important, just playing GTO is more interesting. Consider the following simplified game. The ante is \$2 and you can raise \$1 or do nothing. Suppose also that each player gets a pair independently with probability $\frac{1}{2}$, and you either get a pair or nothing. Let's suppose the row player goes first and is the only one who is allowed to raise.

There are only two strategies for the row player that are worth playing:

- S_1: always bet.
- S_2: only bet with a pair.

And two strategies for the column player:

- S_1': always call.
- S_2': only call with a pair.

	S'_1	S'_2
S_1	0	$\frac{1}{4}$
S_2	$\frac{1}{4}$	0

Table 4: Payoff matrix for a simplified (and unfair) poker game.

Let's compute the expected reward for one entry, when the row player uses S_1 and the column player S'_2. Here the row player will always bluff and the column player will always fold with nothing. So the expected value for the row player is:

$$\frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot (-3) + \frac{1}{4} \cdot 2 = \frac{1}{4}$$

Where the events are pair/pair, pair/nothing, nothing/pair, and nothing/nothing for the row and column player respectively. In other words, it is worth it to bluff if the other player will never call when they have nothing.

Here there is no pure Nash, since the column player always wants to swap to play the same indexed strategy as the row player. However, there is a mixed Nash. The row player can play S_1 half the time and S_2 half the time. This corresponds to bluffing with probability $\frac{1}{2}$ when the column player has nothing and betting nothing otherwise. The best thing for the column player to do is to call with nothing half the time, leading to an expected payoff of the row player of $+\frac{1}{8}$, as half the time they will get a $+\frac{1}{4}$ entry of the matrix.

But the column player can do much worse than this, even against the row player's GTO strategy. They could, for example, never call, a terrible strategy which would give us expected value bigger than $+\frac{1}{4}$. In actual poker, which is much more complicated, there are many reasonable sounding strategies which GTO play has positive expected value against (if you don't believe me, try playing against a GTO opponent, for example here).