2.1 Miscellanea

Things I forgot to mention at the beginning of class last time:

- The first homework is out today, and is due in two weeks.
- The recording / zoom for last class didn't come out great. I'll try to remember to have it more pointing towards me in the future. But if I forget or if it's not working well, let me know!

2.2 Nash Equilibria

Let's be a bit more formal than we were last time, and actually define things. I'm still going to abuse notation a bit, but hopefully it's all clear.

Definition 2.2.1 *A* one-shot simultaneous move game *consists of the following:*

- *n players $[n] = \{1, 2, \ldots, n\}$*
- *Set S_i of possible strategies/actions for each $i \in [n]$. We will let $S = S_1 \times S_2 \times \cdots \times S_n$ be the set of strategy profiles.*
- *Utility function $u_i : S \to \mathbb{R}$ for each $i \in [n]$.*

Since we'll talk about probability distributions over actions all the time, let's set up some notation:

$$\Delta_k = \left\{ x \in \mathbb{R}^k : x_i \geq 0 \ \forall i \in [k], \ \sum_{i=1}^{k} x_i = 1 \right\}$$

is the set of all probability distributions/vectors over k coordinates/things. We might instead use a set instead of k, in which case it's the set of all probability distributions over the set.

Now we can give an actual definition of Nash equilibrium (although we'll slightly simplify this later):

Definition 2.2.2 (Nash Equilibrium) *Let $\mu_i \in \Delta_{S_i}$ for all $i \in [n]$. Then $(\mu_1, \mu_2, \ldots, \mu_n)$ is a* Nash Equilibrium *if for every $i \in [n]$:*

$$\mathop{\mathbf{E}}_{s_j \sim \mu_j \ \forall j \in [n]} [u_i(s_1, \ldots, s_n)] \geq \mathop{\mathbf{E}}_{\substack{s_i \sim \mu_i' \\ s_j \sim \mu_j \ \forall j \in [n] \setminus \{i\}}} [u_i(s_1, \ldots, s_n)] \quad \forall \mu_i' \in \Delta_{S_i}$$

In other words, every player has a distribution over their actions/strategies, and for every player, if they deviated to some other distribution then they would not be any better off.

As some more notation, a distribution $\mu_j \in \Delta_{S_j}$ is called a *mixed strategy* since it is a "mix" over actions. Moreover, a single action is known as a *pure strategy*. A *pure Nash equilibrium* is a Nash equilibrium with the special property that each μ_i is actually a pure strategy (a single action, or equivalently the vector μ_i has a single nonzero entry). Sometimes, in order to contrast with pure Nash, we'll call a Nash equilibrium a *mixed Nash*.

While the definition above captures the intuitive notion that players don't want to deviate from their distribution to a different distribution, we can actually slightly simplify it. Suppose that player i *does* have incentive to deviate from μ_i to some other μ_i'. Then it is easy to see (good exercise to do at home!) that in fact player i will have incentive to deviate to some *pure* strategy. Intuitively, this is because all other players have fixed distributions, so any mixed strategy for i gives (expected) utility that is a convex combination of utilities of pure strategies, and any linear combination of pure strategies is at most the best pure strategy.

So in our definition of Nash, we can change the quantification to be over all actions $a \in S_i$ rather than over all $\mu_i' \in \Delta_{S_i}$. So often when we analyze Nash we'll just check whether any player has incentive to deviate to any pure strategy.

A few more notes about Nash equilibria:

- We need to assume that both the number of players and the strategy sets are finite, or else we don't necessarily have a guarantee that a Nash exists.
- The intuition is that there's no incentive to deviate, but we're playing a one-shot game. This is kind of weird. We'll talk a bit more about this in the future, but generally, it's always worth keeping in mind how our formalisms actually correspond to "real life".
- We also are implicitly assuming that every player knows the mixed strategies of all other players. Also kind of a weird assumption.

There's an enormous literature in game theory which generalizes Nash equilibria to other kinds of games. In games with turns/time there's something called a *subgame perfect Nash*, in games without full information there's something called a *Bayes Nash*, and in games where players can cooperate/collude there's something called a *strong Nash*. We might talk a little about some of these, but they're not the focus of the course – we're going to spend almost all of our time on one-shot games. These are all good topics to explore in your final project, though!

2.3 Two-Player Zero-Sum Games

Today we're going to focus on the very special case of two-player zero-sum games. Next class we'll generalize to other two-player games, and after that we'll talk about higher player counts. But we're going to start in this very special setting both for historical reasons (they were one of the first classes studied, even pre-Nash) and because there are some very interesting results which don't hold in general.

We're going to slightly change notation to specialize to the two-player case.

Definition 2.3.1 *In a 2-player game, also called a* bimatrix game*, there are two players and* $S_1 = [N]$ *and* $S_2 = [M]$*. There are two matrices* $A, B \in \mathbb{R}^{N \times M}$*. The utility functions are*

$$u_1(i,j) = A_{ij} \qquad u_2(i,j) = B_{ij}.$$

Player 1 is the row player *and player 2 is the* column player*.*

This matrix-focused way is particularly convenient because everything becomes simple linear algebra. For example, let $x \in \Delta_N$ and $y \in \Delta_M$ be mixed strategies for the two players. Then it's not hard to see that the expected utility of player 1 is $x^T Ay$, while the expected utility of player 2 is $x^T By$. Let's prove this for player 1 (the proof for player 2 is analogous):

$$\begin{aligned}
x^T Ay &= \sum_{i=1}^{N} \sum_{j=1}^{M} A_{ij} x_i y_j \\
&= \sum_{i=1}^{N} \sum_{j=1}^{M} A_{ij} \mathbf{Pr}[\text{player 1 plays } i] \mathbf{Pr}[\text{player 2 plays } j] \\
&= \sum_{i=1}^{N} \sum_{j=1}^{M} A_{ij} \mathbf{Pr}[\text{player 1 plays } i \wedge \text{player 2 plays } j] \\
&= \sum_{i=1}^{N} \sum_{j=1}^{M} u_1(i,j) \mathbf{Pr}[\text{player 1 plays } i \wedge \text{player 2 plays } j] \\
&= \mathop{\mathbf{E}}_{\substack{i \sim x \\ j \sim y}} [u_1(i,j)]
\end{aligned}$$

Definition 2.3.2 *A bimatrix game is* zero-sum *if* $B_{ij} = -A_{ij}$ *for all* $i \in [N], j \in [M]$*.*

This means that in a zero-sum bimatrix game, we can basically just forget B: we can think of player 1 as trying to maximize $x^T Ay$ and player 2 as trying to minimize $x^T Ay$.

We're going to spend basically the rest of class proving the following theorem.

Theorem 2.3.3 (von Neumann) *In a two-player zero-sum game, we can compute a Nash equilibrium in polynomial time.*

Note that this is not the form in which it was originally written given – this is a reinterpretation for this class. For example, von Neumann actually talked about "minimax strategies" rather than Nash, but in this context they're the same thing. He also didn't talk about "polynomial time", but that is a consequence of his actual proof.

To prove this, let's start reasoning about the structure of these games. Suppose that the row player uses a mixed strategy $x \in \Delta_N$. Then $x^T A$ is a vector in $\mathbb{R}^M$ where $(x^T A)_i$ is the (expected) utility of player 1 (cost to player 2) if the column player uses strategy $i \in [M]$. So if the column player knows that the row players is using x, what will they do? They'll just choose the action which minimizes this value! (If there are multiple such actions, it will choose some distribution over them). In other words, if the row player uses x, then the column player will play a strategy which results in it having cost $\min_{i \in [M]} (x^T A)_i$ (and row player gaining that much utility).

So now let's go back the row player: what will they do, now that they've done this analysis? They'll try to choose the x which maximizes $\min_{i \in [M]} (x^T A)_i$! In other words, the row player will try to solve the following problem:

$$\begin{aligned}
v_r = \max\ & v \\
\text{s.t. }\ & x_i \geq 0 && \forall i \in [N] \\
& \sum_{i=1}^{N} x_i = 1 \\
& (x^T A)_j \geq v && \forall j \in [M]
\end{aligned}$$

where recall that $(x^T A)_j = \sum_{i=1}^{N} x_i A_{ij}$.

This should look familiar to you: it's a linear program! And remember that linear programs can be solved in polynomial time using the Ellipsoid algorithm or interior point methods! This means that the row player can actually compute this strategy (which it will do since it is rational by assumption), ensuring that it gets utility v_r (no matter what the column player does). Let p be the optimal solution to this linear program.

Fix some arbitrary Nash equilibrium (p^*, q^*), and let v^* be the value obtained by the row player in such a Nash.

Lemma 2.3.4 $v_r = v^*$.

Proof: First, suppose for contradiction that $v_r > v^*$. Then if the row player switches instead to strategy p from p^*, it is guaranteed to get utility at least $v_r > v^*$, and thus the row player has incentive to switch, contradicting the definition of Nash. Thus $v_r \leq v^*$.

Now we just need to show that $v^* \leq v_r$. To see this, note that the column player is trying to minimize its cost, which is equal to the utility of player 1. Thus it must be the case that $v^* = \min_{j \in [M]} ((p^*)^T A)_j$. Since p^* is a valid solution to the LP with objective value v^*, this means that $v^* \leq v_r$ (since v_r is the value of the optimal LP solution). ∎

Note that we fixed an *arbitrary* Nash, and showed that v_r was equal to its value. This clearly implies that *every* Nash has the same value.

Now we can do the same thing for the column player! If the column player uses mixed strategy y, then the row player will choose the $i \in [N]$ that maximizes $(Aq)_i$. Knowing this, we get that the column player will play the mixed strategy corresponding to the optimal solution of the following LP:

$$\begin{aligned}
v_c = \min\ & v \\
\text{s.t. }\ & y_j \geq 0 && \forall j \in [M] \\
& \sum_{j=1}^{M} y_j = 1 \\
& (Ay)_i \leq v && \forall i \in [N]
\end{aligned}$$

where recall that $(Ay)_i = \sum_{j=1}^{M} A_{ij} y_j$.

Now by an analogous argument to our previous lemma, we can prove that $v_c = v^*$, which in turn implies that $v_r = v_c$. This is already a really cool fact: the maximum utility that the row player can guarantee for itself is equal to the minimum costs that the column player can guarantee it pays! Side note: if you're familiar with LPs, it's very easy to see this fact via strong duality.

Let p be the optimal solution to the row player LP, and let q be the optimal solution to the column player LP (note that we can find both in polynomial time). Then we can prove the main theorem of today:

Theorem 2.3.5 *(p, q) is a Nash equilibrium.*

Proof: Since p is an optimal solution to the row player's LP and we showed that $v_r = v^*$, we know that $p^T Aq \geq v_r = v^*$. Similarly, since q is an optimal solution to the column player's LP and we showed that $v_c = v^*$, we know that $p^T Aq \leq v_c = v^*$. Hence $p^T Aq = v^*$.

Suppose that the row player plays x instead of p. Then player 1 gets utility

$$x^T Aq = \sum_{i=1}^{N} x_i (Aq)_i \leq \max_{i \in [N]} (Aq)_i = v_c = v^*,$$

and hence player 1 has no incentive to deviate. Similarly, if the column player plays y instead of q, then it has cost

$$p^T Ay = \sum_{j=1}^{M} (p^T A)_j y_j \geq \min_{j \in [M]} (p^T A)_j = v_r = v^*,$$

and thus the column player also has no incentive to deviate. The (p, q) is a Nash. ∎