Finding Nash Equilibria of 2-Player Zero-Sum Games

We mentioned *Nash's Theorem* in the previous class: in every strategic game with a finite number of players and a finite number of actions per player, there is at least one mixed Nash equilibrium. In this class, we prove the theorem for the special case of 2-player zero-sum games.

(Note: Chapter 11 of Osborne's text is somewhat similar to these notes. However his approach is kind of like using a hammer (Nash's Theorem, whose proof relies on topological arguments and fixed point theorems) to open a walnut. I am taking a different approach in class and in these notes: we give a direct proof of Nash's Theorem in the special case of 2-player zero-sum games, using LP duality.)

Definition. We say that a game with von Neumann-Morganstern preferences is a *zero-sum game* if for all outcomes a,

$$\sum_{i=1}^{n} u_i(a) = 0.$$

In particular, for a two-player game, $u_1(a) = -u_2(a)$ for all outcomes a, so it is convenient to write down only the utility function for player 1.

Example of a 2-Player Zero-Sum Game. In the table below we show the payoffs for *player 1* for the game Rock-Paper-Scissors.

$p1 \backslash p2$	R	P	S
R	0	-1	1
P	1	0	-1
S	-1	1	0

(u_1 shown; $u_2 = -u_1$)

Consider the mixed strategy of player 1 which assigns probability $1/3$ to each of her actions. Question: what is special about this strategy? Answer: it guarantees no expected loss to player 1! That is to say, no matter which pure or mixed strategy player 2 responds with, the expected value to player 1 is zero.

Now let's consider a less familiar example. **Example: Matching Coins** (attributed to Vasek Chvátal 1981 in Martin Gardner's book "Mathematical Recreations.") Two players each privately pick a nickel, dime or quarter. Both are revealed simultaneously. If they are different, then player 1 wins both the coins, and if they match, player 2 wins both the coins. This can be written as a 2-player strategic game, with payoffs represented by cents won or lost. The game is zero-sum, since player 1's lost coins go directly to player 2, and vice-versa. Here is the payoff table for player 1:

$p1 \backslash p2$	N	D	Q
N	-5	10	25
D	5	-10	25
Q	5	10	-25

(u_1 shown; $u_2 = -u_1$)

Remark: one contrast between the 2 previously shown games is that Rock-Paper-Scissors is *symmetric* with respect to the two players ($u_1(A, B) = -u_1(B, A) = u_2(B, A)$ for all actions A and B) but Matching Coins is not symmetric. Intuitively, we expect neither player therefore to have an advantage in Rock-Paper-Scissors, while the same might not be true about Matching Coins.

Maxminimization

Question: Is it possible for player 1 to guarantee some positive expected profit in Matching Pennies? Answer: Yes. Perhaps the simplest way to see this is that, if player 1 uses a uniformly mixed distribution (the one that assigns probability $1/3$ to each action — which we will also write as $\frac{1}{3}N + \frac{1}{3}D + \frac{1}{3}Q$), then

- if player 2 uses the pure response N, then player 1 wins $5/3$ cents in expectation

- if player 2 uses the pure response D, then player 1 wins $10/3$ cents in expectation

- if player 2 uses the pure response Q, then player 1 wins $25/3$ cents in expectation

- by weighted averaging, when using $\frac{1}{3}N + \frac{1}{3}D + \frac{1}{3}Q$, player 1 wins at least $5/3$ cents in expectation against all mixed strategies of player 2

We can write this into our table,

$p1\backslash p2$	N	D	Q
N	-5	10	25
D	5	-10	25
Q	5	10	-25
$\frac{1}{3}N + \frac{1}{3}D + \frac{1}{3}Q$	5/3	10/3	25/3

(u_1 shown; $u_2 = -u_1$)

Question: what does this tell us about the *best* player 2 can do? Answer: there is no mixed strategy that can guarantee player 1 any expected utility better than $-5/3$ cents.

In other words, player 2 is going to lose money in this game no matter what he does. But, can he limit his losses? We could try the same trick as before: look at the payoffs for player 2 when he uses the uniformly mixed strategy, depending on the pure response of player 1. Visually, we add a new column to our table (caution, these are written as utilities for player 2, so the signs change).

	$\frac{1}{3}N + \frac{1}{3}D + \frac{1}{3}Q$
N	-10
D	-20/3
Q	10/3

(u_2 shown)

So by using the uniformly mixed strategy, player 2 can guarantee that he will lose no more than 10 cents per round, in expectation.

But: we still don't have the whole story for this game: we only know that player 2's "best" scenario leaves him losing between $5/3$ and 10 cents per round. Question: is there some way we can tighten our analysis so as to learn more? Can we get rid of this "gap?" Answer: there might be mixed strategies for player 1 and player 2 that "guarantee" them more than they are "guaranteed" by the uniformly mixed strategy.

What, exactly, do we mean by a better guarantee? Consider an arbitrary mixed strategy $\alpha_1 := (p_N, p_D, p_Q)$ for player 1, where p_N is the probability assigned to N, etc. Then as before, we can compute the values of α_1 to player 1 depending on the response of player 2, and we add a new row to the table:

$p1\backslash p2$	N	D	Q
α_1	$-5p_N + 5p_D + 5p_Q$	$10p_N - 10p_D + 10p_Q$	$25p_N + 25p_D - 25p_Q$

We see that α_1 guarantees a payoff to player 1 of $\min\{-5p_N + 5p_D + 5p_Q, 10p_N + -10p_D + 10p_Q, 25p_N + 25p_D - 25p_Q\}$. How can we maximize this guarantee? More generally, our approach suggests the following definition:

Definition. The *guarantee* of a mixed strategy α_i for player i is the minimum value of $u_i(\alpha_i, a_{-i})$ over all partial action profiles a_{-i} that the other players can choose. The *maxminimized value* of the game for player i, denoted $\mathrm{maxmin}(i)$, is the value of the largest guarantee,

$$\mathrm{maxmin}(i) := \max_{\alpha_i} \min_{a_{-i}} u_i(\alpha_i, a_{-i}),$$

where the max is taken over all mixed strategies player i has. A mixed strategy α_i that attains this maximum is called a *maxminimizer* (or *maxminimizing strategy*) for player i.

Now, let's go back to Matching Coins. Question: how can we find a maxminimizer for player 1? Answer: solve a linear program, with variables p_N, p_D, p_Q. A little trick is needed to compute the min.

We now give a more general presentation that works for any 2-player game. Without loss of generality write $A_1 = \{1, 2, \ldots, m\}$ and $A_2 = \{1, 2, \ldots, n\}$. Then the LP for player 1 is

$$\begin{aligned}
\max \; & v && \leftarrow \text{``value'' guaranteed} \\
s.t. \quad & p_i \geq 0 && \text{for } i = 1, \ldots, m \\
& \sum_{i=1}^{m} p_i = 1 \\
& \sum_{i=1}^{m} p_i \cdot u_1(i, j) \geq v && \text{for } j = 1, \ldots, n.
\end{aligned}$$

(Question: How does the trick work — why is v always $\min_j\{\sum_{i=1}^{m} p_i \cdot u_1(i,j)\}$ in an optimal solution?) Solving this LP for Matching Coins gives us an optimal solution (p^*, v^*) defined by $p_N^* = 7/34, p_D^* = 12/34, p_Q^* = 15/34, v^* = +50/17$. So by using the mixed strategy corresponding to p^*, player 1 wins at least $50/17$ expected cents per round.

Just as before, this also tells us that player 2 could at best guarantee himself an expected payoff of $-50/17$ cents per round.

We can also use the LP approach to compute a maxminimizer for player 2. In this case

- the unique maxminimizer for player 2 assigns probability $10/17$ to N, $5/17$ to D, and $2/17$ to Q

- the guarantee for this maxminimizer is $-50/17$

So amazingly, the "gap" that we had before has disappeared: we know that losing $50/17$ cents per game is the best possible scenario for player 2, and that winning $50/17$ cents per game is the best possible scenario for player 1. (In fact, as we will see, we have found a Nash equilibrium.) This is not a coincidence:

Theorem (von Neumann's minimax theorem, 1928)**.** In any 2-player zero-sum game,

$$\mathrm{maxmin}(1) + \mathrm{maxmin}(2) = 0.$$

We will prove this. (von Neumann looked at minimizing the maximum payoff of any opponent, which is why the name is different.) Unlike the original proof, we will give a short proof using the theory of LP duality. However, before getting in to these details, we show what this has to do with Nash equilibria.

Proposition. Suppose an n-player zero-sum game has the property that $\sum_{i=1}^{n} \text{maxmin}(i) = 0$. If α_i is a maxminimizing strategy for each player, then α is a mixed Nash equilibrium. Conversely, if **there are only 2 players**, and α' is a mixed Nash equilibrium, then α'_i is a maxminimizing strategy for each player i.

Proof. We will only prove the first of the two statements. Suppose otherwise for the sake of contradiction: α is not a NE, so some player j has a profitable deviation α'_j,

$$u_j(\alpha'_j, \alpha_{-j}) > u_j(\alpha).$$

Now for each player $i \neq j$, since α_i is a maxminimizer,

$$u_i(\alpha'_j, \alpha_{-j}) \geq \text{maxmin}(i)$$

and likewise, since α_j is a maxminimizer,

$$u_j(\alpha'_j, \alpha_{-j}) > u_j(\alpha) \geq \text{maxmin}(j).$$

Adding up the above inequalities for each player, we get

$$\sum_{i=1}^{n} u_i(\alpha'_j, \alpha_{-j}) > \sum_{i=1}^{n} \text{maxmin}(i) = 0$$

(where the equality uses the hypothesis of the proposition.) But by linearity of expectation and the fact that the game is zero-sum,

$$\sum_{i=1}^{n} u_i(\alpha'_j, \alpha_{-j}) = 0.$$

So the last two lines provide our contradiction.   □

Remark: it is not true that in all 3-player zero-sum games, maxmin(1)+maxmin(2)+maxmin(3)=0. However, it is still possible to compute maxmin values by an LP; e.g. in the LP for player 1, instead of $|A_2|$ constraints for each action of player 2, you would use $|A_2| \cdot |A_3|$ constraints for all partial action profiles a_{-1}.

Finally, what we really cared about:

Corollary. Every 2-player zero-sum has a mixed Nash equilibrium.

Proof. Combine the minimax theorem with the **Proposition**.   □

Proof of Minimax Theorem

Here is the LP to compute maxmin(1), modified into a standard form:

$$\text{maxmin}(1) = \max v \qquad \qquad \leftarrow \text{``value'' guaranteed}$$
$$s.t. \qquad p_i \geq 0 \qquad \qquad \text{for } i = 1, \ldots, m$$
$$\sum_{i=1}^{m} p_i = 1$$
$$v - \sum_{i=1}^{m} p_i \cdot u_1(i,j) \leq 0 \qquad \qquad \text{for } j = 1, \ldots, n.$$

We now compute its dual. Introduce an unbounded variable w for the equality constraint $\sum_{i=1}^{m} p_i = 1$ and a nonnegative variable q_j for each inequality constraint $\sum_{i=1}^{m} p_i \cdot u_1(i,j) - v \geq 0$. Then the new LP is

$$\text{maxmin}(1) = \min w$$
$$s.t. \qquad q_j \geq 0 \qquad \qquad \text{for } j = 1, \ldots, n$$
$$\sum_{j=1}^{n} q_j = 1 \qquad \qquad \text{since } v \text{ was unbounded}$$
$$w - \sum_{j=1}^{n} q_j \cdot u_1(i,j) \geq 0 \qquad \qquad \text{for } j = 1, \ldots, n.$$

(The initial equality follows from strong LP duality.)

Now we replace each $u_1(i,j)$ by $-u_2(i,j)$ (it is a zero-sum game) and we replace w by the variable $w' := -w$. Multiply the inequalities by -1. The minimum possible value of $-w'$ is the negative maximum of w', so we have

$$\text{maxmin}(1) = -\max w'$$
$$s.t. \qquad q_j \geq 0 \qquad \qquad \text{for } j = 1, \ldots, n$$
$$\sum_{j=1}^{n} q_j = 1$$
$$w' - \sum_{j=1}^{n} q_j \cdot u_2(i,j) \leq 0 \qquad \qquad \text{for } j = 1, \ldots, n.$$

But ignoring the minus sign, this is the LP to compute the maxmin value of player 2. Hence maxmin(1)=−maxmin(2). □

Other Ruminations

If you are reading this class as a student at Waterloo, it is pretty likely you have seen LPs and the simplex method before. Although the simplex algorithm takes exponential time in the worst case, there are other ways to solve LPs (interior point methods, the ellipsoid method) that run in polynomial time. Hence (by computing the maxmin strategies) we can find a mNE of any 2-player zero-sum game in polynomial time.

The simplex method always finds a rational optimum for any LP. So we can now deduce, assuming all $u_i(\cdot, \cdot)$'s are rational, that every 2-player zero-sum game has a *rational* mixed Nash equilibrium, i.e. one where all probabilities are rational. In fact, the zero-sumness is not important:

Proposition. Every 2-player game has a rational Nash equilibrium, if all $u_i(\cdot, \cdot)$ are rational.

Proof. Now we will use the hammer. By Nash's theorem, there is some Nash equilibrium α^*, but it might be irrational. However, let's consider the Support Characterization to derive another LP. Let $S_1 \subset A_1$ be the support of α_1^*, i.e. $S_1 = \{a_1 \in A_1 \mid \alpha_1^*(a_1) > 0\}$. Define S_2 similarly.

The Support Characterization says that every action in S_1 is a best response against α_2, and vice-versa. Let X be the best response value against α_2 and Y be the best response value against α_1. We can write these as linear constraints, where p_i stands for $\alpha_1^*(a_i)$ and q_j stands for $\alpha_2^*(a_j)$:

$$\Sigma_{i=1}^m p_i = 1$$
$$\Sigma_{j=1}^n q_j = 1$$
$$\text{all } p_i, q_j \text{ nonnegative}$$
$$p_i = 0, \qquad \forall i \notin S_1$$
$$q_j = 0, \qquad \forall j \notin S_2$$
$$\Sigma_{j=1}^n q_j u_1(i, j) = X, \qquad \forall i \in S_1$$
$$\Sigma_{j=1}^n q_j u_1(i, j) \leq X, \qquad \forall i \in A_1 \backslash S_1$$
$$\Sigma_{i=1}^m p_i u_2(i, j) = Y, \qquad \forall j \in S_2$$
$$\Sigma_{i=1}^m p_i u_2(i, j) \leq Y, \qquad \forall j \in A_2 \backslash S_2$$

Since it has a solution (the one corresponding to α^*), by the simplex method, it also has a rational solution. It is easy to check with the Support Characterization that this rational solution gives a rational Nash equilibrium. □

In fact, we can just "guess" the supports and use the above LP to compute a Nash equilibrium for any 2-player game. The Lemke-Howson algorithm (which we will see next class) is closely connected to the above approach; however, the Lemke-Howson algorithm will explicitly show that the equilibrium exists, without using Nash's Theorem in its powerful generality.

However, if we don't know the supports ahead of time, there are $2^{|A_1|+|A_2|}$ possibilities. So that naive algorithm for finding NE of 2-player games runs in exponential time. Is it possible to find NE of general 2-player games in polynomial time? What about 3-player and higher games? Recent developments in complexity theory show that this is probably not possible; much like the "P vs. NP" problem we don't know for sure, but if we could efficiently find (even approximate) NE of 2-player games in polynomial time, we could use it as a subroutine to do lots of other fancy stuff (compute fixed points, find NE of n-player general-sum games) in polynomial time.

As an exercise, by modifying a game discussed in class, find a 3-player finite game that has no rational Nash equilibria. What about a 3-player zero-sum game?