Game Theory 2024

Ulle Endriss
Institute for Logic, Language and Computation
University of Amsterdam

Plan for Today

Today we are going to focus on the special case of *zero-sum games* and discuss two positive results that do not hold for games in general.

- new solution concepts: *maximin* and *minimax solutions*
- *Minimax Theorem:* maximin = minimax = NE for zero-sum games
- *fictitious play:* basic model for learning in games
- *convergence result* for the case of zero-sum games

The first part of this is also covered in Chapter 3 of the *Essentials*.

K. Leyton-Brown and Y. Shoham. *Essentials of Game Theory: A Concise, Multidisciplinary Introduction*. Morgan & Claypool Publishers, 2008. Chapter 3.

Zero-Sum Games

Today we focus on *two-player games* $\langle N, \boldsymbol{A}, \boldsymbol{u} \rangle$ with $N = \{1, 2\}$.

Notation: Given player $i \in \{1, 2\}$, we refer to her opponent as $-i$.

Recall: A *zero-sum game* is a two-player normal-form game $\langle N, \boldsymbol{A}, \boldsymbol{u} \rangle$ for which $u_i(\boldsymbol{a}) + u_{-i}(\boldsymbol{a}) = 0$ for all action profiles $\boldsymbol{a} \in \boldsymbol{A}$.

Examples include (but are not restricted to) games in which you can *win* $(+1)$, *lose* (-1), or *draw* (0), such as *Matching Pennies* (left):

	H	T
H	1, -1	-1, 1
T	-1, 1	1, -1

	L	R
T	5, -5	-3, 3
B	0, 0	2, -2

Constant-Sum Games

A *constant-sum game* is a two-player normal-form game $\langle N, \boldsymbol{A}, \boldsymbol{u} \rangle$ for which there exists a $c \in \mathbb{R}$ such that $u_i(\boldsymbol{a}) + u_{-i}(\boldsymbol{a}) = c$ for all $\boldsymbol{a} \in \boldsymbol{A}$.

Thus: A zero-sum game is a constant-sum game with constant $c = 0$.

Everything about zero-sum games to be discussed today also applies to constant-sum games, but for simplicity we only talk about the former.

Fun Fact: Football is *not* a constant-sum game, as you get 3 points for a win, 0 for a loss, and 1 for a draw. But prior to 1994, when the "three-points-for-a-win" rule was introduced, World Cup games were constant-sum (with 2, 0, 1 points, for win, loss, draw, respectively).

Maximin Strategies

The definitions on this slide apply to arbitrary normal-form games . . .

Suppose player i wants to maximise her worst-case expected utility (e.g., if all others conspire against her). Then she should play:

$$s_i^\star \in \operatorname*{argmax}_{s_i \in S_i} \min_{\boldsymbol{s}_{-i} \in \boldsymbol{S}_{-i}} u_i(s_i, \boldsymbol{s}_{-i})$$

Any such $s_i^\star$ is called a *maximin strategy* (typically there is just one).

Solution concept: assume each player will play a maximin strategy.

Call $\max\limits_{s_i} \min\limits_{\boldsymbol{s}_{-i}} u_i(s_i, \boldsymbol{s}_{-i})$ player i's *maximin value* (or *security level*).

Exercise: Maximin and Nash

Consider the following two-player game:

	L	R
T	8, 2	0, 0
B	0, 0	8, 4

What is the maximin solution?

How does this relate to Nash equilibria?

Note: This is neither a zero-sum nor a constant-sum game.

Exercise: Maximin and Nash Again

Now consider this fairly similar game, which *is* zero-sum:

	L	R
T	8, -8	0, 0
B	0, 0	8, -8

What is the maximin solution?

How does this relate to Nash equilibria?

Minimax Strategies

Now focus on two-player games only, with players i and $-i$ ...

Suppose player i wants to minimise $-i$'s best-case expected utility (e.g., to *punish* her). Then i should play:

$$s_i^\star \in \underset{s_i \in S_i}{\operatorname{argmin}} \max_{s_{-i} \in S_{-i}} u_{-i}(s_i, s_{-i})$$

Any such $s_i^\star$ is called a *minimax strategy* (typically there is just one).

Call $\min\limits_{s_i} \max\limits_{s_{-i}} u_{-i}(s_i, s_{-i})$ player $-i$'s *minimax value*.

So, by analogy, player i's minimax value is $\min\limits_{s_{-i}} \max\limits_{s_i} u_i(s_i, s_{-i})$.

<u>Remark:</u> An alternative interpretation of player i's minimax value is what she gets when her opponent has to play *first* and i can *respond*.

Equivalence of Maximin and Minimax Values

Recall: For two-player games, we have seen the following definitions.

- Player i's *maximin value* is $\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i})$.
- Player i's *minimax value* is $\min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$.

Lemma 1 *In a two-player game, maximin and minimax value coincide:*

$$\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i}) = \min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$$

Exercise: *Can you see why?*

Interlude

To see that the lemma is not trivial, observe that it becomes *false* if we quantify over actions rather than strategies:

$$\max_{a_i} \min_{a_{-i}} u_i(a_i, a_{-i}) \stackrel{?}{=} \min_{a_{-i}} \max_{a_i} u_i(a_i, a_{-i})$$

Take *Matching Pennies* with players being confined to pure strategies. If you go first (LHS) you get -1 but if you go second (RHS) you get 1.

	H	T
H	1, -1	-1, 1
T	-1, 1	1, -1

Proof of Lemma

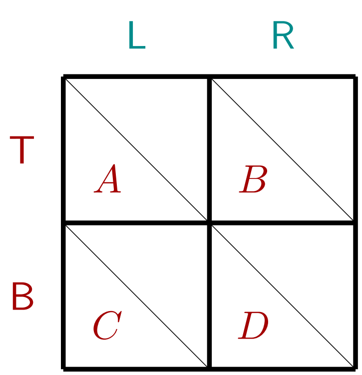
Let us now prove the lemma. The claim is, for any two-player game:

$$\max_{s_i} \min_{s_{-i}} u_i(s_i, s_{-i}) = \min_{s_{-i}} \max_{s_i} u_i(s_i, s_{-i})$$

One direction is straightforward:

($\leqslant$) LHS is what i can achieve when she *has to* move first, while RHS is what i can achieve when she *can* move second. ✓

For the full equation, we sketch the proof for 2x2-games only:



Rowena's maximin strategy is to play T with probability p so Colin cannot affect her EU: $Ap + C(1-p) = Bp + D(1-p)$
So *maximin value* is $Ap + C(1-p)$ for this $p \Rightarrow \frac{AD-BC}{A-B-C+D}$

Colin's minimax strategy is to play L with probability q so Rowena cannot affect her EU: $Aq + B(1-q) = Cq + D(1-q)$
So *minimax value* is $Aq + B(1-q)$ for this $q \Rightarrow \frac{AD-BC}{A-B-C+D}$

So we really get the same value! ✓ (Exercise: *Verify this!*)

The Minimax Theorem

Recall: A zero-sum game is a two-player game with $u_i(\boldsymbol{a}) + u_{-i}(\boldsymbol{a}) = 0$.

Theorem 2 (Von Neumann, 1928) *In a zero-sum game, a strategy profile is a NE iff each player's expected utility equals her minimax value.*

Proof: Let v_i be the minimax/maximin value of player i (and $v_{-i} = -v_i$ that of player $-i$).

(1) Suppose $u_i(s_i, s_{-i}) \neq v_i$. Then one player does worse than she could (v_i as *maximin* value). So she can deviate: (s_i, s_{-i}) *is not a NE*. ✓

(2) Suppose $u_i(s_i, s_{-i}) = v_i$. Then you cannot do better even if you were allowed to move second (v_i as *minimax* value). So (s_i, s_{-i}) *is a NE*. ✓

John von Neumann (1903–1957)

J. von Neumann. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, 1928.

Computing Nash Equilibria in Zero-Sum Games

The Minimax Theorem suggests a way of computing Nash equilibria for zero-sum games that is simpler than the general approach.

The reason why this simplifies matters is that, to compute the maximin (or minimax) value of a player, you only need to consider *her* payoffs.

Learning in Games

Suppose you keep playing the same game against the same opponents. You might try to *learn* their *strategies*.

A good hypothesis might be that the *frequency* with which player i plays action a_i is approximately her *probability* of playing a_i.

Now suppose you always best-respond to those hypothesised strategies. And suppose everyone else does the same. *What will happen?*

We are going to see that for *zero-sum games* this process *converges* to a NE. This yields a method for *computing a NE* for the (non-repeated) game: just *imagine* players engage in such "*fictitious play*".

Empirical Mixed Strategies

Given a *history* of actions $H_i^\ell = a_i^0, a_i^1, \ldots, a_i^{\ell-1}$ played by player i in ℓ prior plays of game $\langle N, \boldsymbol{A}, \boldsymbol{u}\rangle$, fix her *empirical mixed strategy* $s_i^\ell \in S_i$:

$$s_i^\ell(a_i) \quad = \quad \underbrace{\frac{1}{\ell} \cdot \#\{k < \ell \mid a_i^k = a_i\}}_{\text{relative frequency of } a_i \text{ in } H_i^\ell} \quad \text{for all } a_i \in A_i$$

Best Pure Responses

Recall: Strategy $s_i^\star \in S_i$ is a *best response* for player i to the (partial) strategy profile $\boldsymbol{s}_{-i}$ if $u_i(s_i^\star, \boldsymbol{s}_{-i}) \geqslant u_i(s_i', \boldsymbol{s}_{-i})$ for all $s_i' \in S_i$.

Due to expected utilities being convex combinations of plain utilities:

Observation 3 *For any given (partial) strategy profile $\boldsymbol{s}_{-i}$, the set of best responses for player i must include at least one pure strategy.*

So we can restrict attention to *best pure responses* for player i to $\boldsymbol{s}_{-i}$:

$$a_i^\star \in \underset{a_i \in A_i}{\operatorname{argmax}}\, u_i(a_i, \boldsymbol{s}_{-i})$$

Fictitious Play

Take any action profile $\boldsymbol{a}^0 \in A$ for the normal-form game $\langle N, \boldsymbol{A}, \boldsymbol{u} \rangle$.

Fictitious play of $\langle N, \boldsymbol{A}, \boldsymbol{u} \rangle$, starting in $\boldsymbol{a}^0$, is the following process:

- In round $\ell = 0$, each player $i \in N$ plays action a_i^0.
- In any round $\ell > 0$, each player $i \in N$ plays a *best pure response* to her opponents' *empirical mixed strategies:*

$$a_i^\ell \in \underset{a_i \in A_i}{\operatorname{argmax}}\, u_i(a_i, \boldsymbol{s}_{-i}^\ell), \text{ where}$$

$$s_{i'}^\ell(a_{i'}) = \frac{1}{\ell} \cdot \#\{k < \ell \mid a_{i'}^k = a_{i'}\} \text{ for all } i' \in N \text{ and } a_{i'} \in A_{i'}$$

Assume some deterministic way of *breaking ties* between maxima.

This yields a sequence $\boldsymbol{a}^0 \twoheadrightarrow \boldsymbol{a}^1 \twoheadrightarrow \boldsymbol{a}^2 \twoheadrightarrow \ldots$ with a corresponding sequence of empirical-mixed-strategy profiles $\boldsymbol{s}^0 \twoheadrightarrow \boldsymbol{s}^1 \twoheadrightarrow \boldsymbol{s}^2 \twoheadrightarrow \ldots$

Question: Does $\lim_{\ell \to \infty} \boldsymbol{s}^\ell$ exist and is it a meaningful strategy profile?

Example: Matching Pennies

Let's see what happens when we start in the upper lefthand corner HH (and break ties between equally good responses in favour of H):

	H	T
H	1, -1	-1, 1
T	-1, 1	1, -1

Any strategy can be represented by a single probability (of playing H).

$$\text{HH}\ (\tfrac{1}{1}, \tfrac{1}{1}) \twoheadrightarrow \text{HT}\ (\tfrac{2}{2}, \tfrac{1}{2}) \twoheadrightarrow \text{HT}\ (\tfrac{3}{3}, \tfrac{1}{3}) \twoheadrightarrow \text{TT}\ (\tfrac{3}{4}, \tfrac{1}{4}) \twoheadrightarrow \text{TT}\ (\tfrac{3}{5}, \tfrac{1}{5})$$
$$\twoheadrightarrow \text{TT}\ (\tfrac{3}{6}, \tfrac{1}{6}) \twoheadrightarrow \text{TH}\ (\tfrac{3}{7}, \tfrac{2}{7}) \twoheadrightarrow \text{TH}\ (\tfrac{3}{8}, \tfrac{3}{8}) \twoheadrightarrow \text{TH}\ (\tfrac{3}{9}, \tfrac{4}{9})$$
$$\twoheadrightarrow \text{TH}\ (\tfrac{3}{10}, \tfrac{5}{10}) \twoheadrightarrow \text{HH}\ (\tfrac{4}{11}, \tfrac{6}{11}) \twoheadrightarrow \text{HH}\ (\tfrac{5}{12}, \tfrac{7}{12}) \twoheadrightarrow \cdots$$

Exercise: *Can you guess what this will converge to?*

Convergence Profiles are Nash Equilibria

In general, $\lim_{\ell \to \infty} \boldsymbol{s}^\ell$ does not exist (no guaranteed convergence). But:

Lemma 4 *If fictitious play converges, then to a Nash equilibrium.*

Proof: Suppose $\boldsymbol{s}^\star = \lim_{\ell \to \infty} \boldsymbol{s}^\ell$ exists. To see that $\boldsymbol{s}^\star$ is a NE, note that $s_i^\star$ is the strategy that i *seems* to play when she best-responds to $\boldsymbol{s}^\star{}_{-i}$, which she *believes* to be the profile of strategies of her opponents. ✓

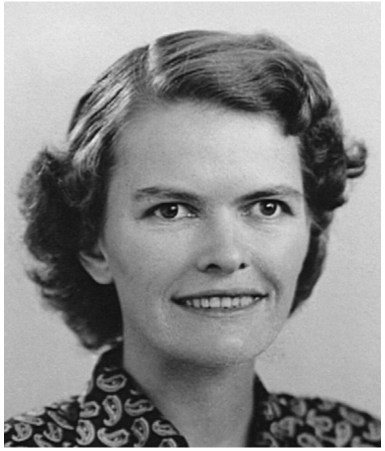
Remark: This lemma is true for arbitrary (not just zero-sum) games.

Convergence for Zero-Sum Games

Good news:

Theorem 5 (Robinson, 1951) *For any zero-sum game and initial action profile, fictitious play will converge to a Nash equilibrium.*

We know that if FP converges, then to a NE. Thus, we still have to show that it will converge. The proof of this fact is difficult and we are not going to discuss it here.



Julia Robinson
(1919–1985)

J. Robinson. An Iterative Method of Solving a Game. *Annals of Mathematics*, 54(2):296–301, 1951.

Summary

We have seen that *zero-sum games* are particularly well-behaved:

- *Minimax Theorem:* your expected utility in a Nash equilibrium will simply be your minimax/maximin value
- Convergence of *fictitious play:* if each player keeps responding to their opponent's estimated strategy based on observed frequencies, these estimates will converge to a Nash equilibrium

Both results give rise to alternative methods for computing a NE.

What next? Players who have incomplete information (are uncertain) about certain aspects of the game, such as their opponents' utilities.