Convergence of No-Regret Play to Nash Equilibrium

Aaron Roth

University of Pennsylvania

February 15 2024
Overview

- We’ve seen that two-player zero sum games are special.
  - They have a value, order of play doesn’t matter, equilibria can be computed “easily” i.e. it does not require counterspeculation — don’t need to reason about your opponent to compute a minmax strategy.
  - But you need to understand the game extremely well and make careful calculations.
  - Is there a natural dynamic that leads to Nash equilibrium if everyone uses it?
  - How many of these properties depend on the “two player” caveat?
Overview

▶ We’ve seen that two-player zero sum games are special.
▶ They have a value, order of play doesn’t matter, equilibria can be computed “easily”
We’ve seen that two-player zero sum games are special.

They have a value, order of play doesn’t matter, equilibria can be computed “easily”

i.e. it does not require counterspeculation — don’t need to reason about your opponent to compute a minmax strategy.
Overview

- We’ve seen that two-player zero sum games are special.
- They have a value, order of play doesn’t matter, equilibria can be computed “easily”
- i.e. it does not require counterspeculation — don’t need to reason about your opponent to compute a minmax strategy.
- But you need to understand the game extremely well and make careful calculations.
Overview

- We’ve seen that two-player zero sum games are special.
- They have a value, order of play doesn’t matter, equilibria can be computed “easily”
- i.e. it does not require counterspeculation — don’t need to reason about your opponent to compute a minmax strategy.
- But you need to understand the game extremely well and make careful calculations.
- Is there a natural dynamic that leads to Nash equilibrium if everyone uses it?
Overview

- We’ve seen that two-player zero sum games are special.
- They have a value, order of play doesn’t matter, equilibria can be computed “easily”
- i.e. it does not require counterspeculation — don’t need to reason about your opponent to compute a minmax strategy.
- But you need to understand the game extremely well and make careful calculations.
- Is there a natural dynamic that leads to Nash equilibrium if everyone uses it?
- How many of these properties depend on the “two player” caveat?
Two players?

Do these special properties carry over to general $n$ player zero sum games?
Two players?

Do these special properties carry over to general $n$ player zero sum games? We can certainly define such games:

Definition
An $n$ player game is zero-sum if for every action profile $a \in A$, $\sum_{i=1}^{n} u_i(a) = 0$. 

"Meta Theorem": $n$ player zero-sum games don't have any special properties that $n-1$ player general sum games don't have. In particular, we should not expect such games to have a value, nor that their equilibria should be easy to compute.

"Proof": Any $n-1$ player game can be made into an $n$ player zero sum game, by adding a new player $n$ (with a trivial action set), and $u_n(a) = -\sum_{i=1}^{n-1} u_i(a)$. Since player $n$ is payoff irrelevant to the $n-1$ other players, the equilibrium structure remains identical to the original game.
Two players?

Do these special properties carry over to general $n$ player zero sum games? We can certainly define such games:

**Definition**
An $n$ player game is zero-sum if for every action profile $a \in A$, $\sum_{i=1}^{n} u_i(a) = 0$.

The answer is no.

*"Meta Theorem":* $n$ player zero-sum games don’t have any special properties that $n-1$ player general sum games don’t have.

In particular, we should not expect such games to have a value, nor that their equilibria should be easy to compute.
Two players?

Do these special properties carry over to general $n$ player zero sum games? We can certainly define such games:

Definition
An $n$ player game is zero-sum if for every action profile $a \in A$, 
\[ \sum_{i=1}^{n} u_i(a) = 0. \]

The answer is no.

“Meta Theorem”: $n$ player zero-sum games don’t have any special properties that $n - 1$ player general sum games don’t have.

In particular, we should not expect such games to have a value, nor that their equilibria should be easy to compute.

“Proof”: Any $n - 1$ player game can be made into an $n$ player zero sum game, by adding a new player $n$ (with a trivial action set), and $u_n(a) = -\sum_{i=1}^{n-1} u_i(a)$. Since player $n$ is payoff irrelevant to the $n - 1$ other players, the equilibrium structure remains identical to the original game.
Definition
A separable graphical game is defined by a graph \( G = (V, E) \). The set of players corresponds to the set of vertices: \( P = V \). Each player’s utility function is decomposable as a sum of neighbor-specific utility functions, one for each of his neighbors in \( G \):

\[
u_i(a) = \sum_{(i, j) \in E} u_{i(j)}(a_i, a_j)
\]

i.e. it is as if each player is playing a 2-player game with each of his neighbors – except he must pick a single action \( a_i \) to play simultaneously against each of his neighbors.
Separable Graphical Games
Zero sum separable graphical games have many of the properties of two player zero sum games:
Zero sum separable graphical games have many of the properties of two player zero sum games:

1. They continue to have a value

Separable Graphical Games
Zero sum separable graphical games have many of the properties of two player zero sum games:

1. They continue to have a value
2. Equilibria are easy to compute with efficient dynamics.
Separable Graphical Games

Zero sum separable graphical games have many of the properties of two player zero sum games:

1. They continue to have a value
2. Equilibria are easy to compute with efficient dynamics.
3. We don’t require each of the constituent 2-player games are zero sum — just that the aggregate is.
Regret

Definition
A sequence of action profiles $a^1, \ldots, a^T$ has regret $\Delta(T)$ if for all players $i$ and actions $a^*_i$ we have:

$$\frac{1}{T} \sum_{t=1}^{T} u_i(a^t) \geq \frac{1}{T} \sum_{t=1}^{T} u_i(a^*_i, a^t_{-i}) - \Delta(T)$$

We say that such an action sequence is no-regret if $\Delta(T) = o_T(1)$. 

1. How to generate a sequence of no-regret play?
2. Have every player play polynomial weights. Then $\Delta(T) = O(2^q \log k T)$
3. But not the only way...
4. A permissive family of dynamics.
Definition
A sequence of action profiles \(a^1, \ldots, a^T\) has regret \(\Delta(T)\) if for all players \(i\) and actions \(a^*_i\) we have:

\[
\frac{1}{T} \sum_{t=1}^{T} u_i(a^t) \geq \frac{1}{T} \sum_{t=1}^{T} u_i(a^*_i, a^t_{-i}) - \Delta(T)
\]

We say that such an action sequence is no-regret if \(\Delta(T) = o_T(1)\).

1. How to generate a sequence of no-regret play?
Regret

Definition
A sequence of action profiles $a^1, \ldots, a^T$ has regret $\Delta(T)$ if for all players $i$ and actions $a_i^*$ we have:

$$\frac{1}{T} \sum_{t=1}^{T} u_i(a^t) \geq \frac{1}{T} \sum_{t=1}^{T} u_i(a_i^*, a^t_{-i}) - \Delta(T)$$

We say that such an action sequence is no-regret if $\Delta(T) = o_T(1)$.

1. How to generate a sequence of no-regret play?
2. Have every player play polynomial weights. Then

$$\Delta(T) = O(2\sqrt{\frac{\log k}{T}})$$
Regret

Definition
A sequence of action profiles $a^1, \ldots, a^T$ has regret $\Delta(T)$ if for all players $i$ and actions $a^*_i$ we have:

$$\frac{1}{T} \sum_{t=1}^{T} u_i(a^t) \geq \frac{1}{T} \sum_{t=1}^{T} u_i(a^*_i, a^t_{-i}) - \Delta(T)$$

We say that such an action sequence is no-regret if $\Delta(T) = o_T(1)$.

1. How to generate a sequence of no-regret play?
2. Have every player play polynomial weights. Then
   $$\Delta(T) = \mathcal{O}(2\sqrt{\log k})$$
3. But not the only way...
Regret

Definition
A sequence of action profiles $a^1, \ldots, a^T$ has regret $\Delta(T)$ if for all players $i$ and actions $a_i^*$ we have:

$$\frac{1}{T} \sum_{t=1}^{T} u_i(a_t) \geq \frac{1}{T} \sum_{t=1}^{T} u_i(a_i^*, a_{-i}^t) - \Delta(T)$$

We say that such an action sequence is no-regret if $\Delta(T) = o_T(1)$.

1. How to generate a sequence of no-regret play?
2. Have every player play polynomial weights. Then $
\Delta(T) = O(2\sqrt{\log k/T})$
3. But not the only way...
4. A permissive family of dynamics.
Dynamics

Given a sequence of action profiles $a^1, \ldots, a^T$, write

$$\bar{a}_i = \frac{1}{T} \sum_{t=1}^{T} a_i^t$$

to denote the mixed strategy for player $i$ that selects an action in \{${a_i^1, \ldots, a_i^T}$\} uniformly at random.
Dynamics

Given a sequence of action profiles $a^1, \ldots, a^T$, write
\[ \bar{a}_i = \frac{1}{T} \sum_{t=1}^{T} a_i^t \] to denote the mixed strategy for player $i$ that selects an action in \{a_i^1, \ldots, a_i^T\} uniformly at random.

**Theorem**

Consider any zero sum separable graphical game $G$. If a sequence of action profiles $a^1, \ldots, a^T$ has regret $\Delta(T)$, then the mixed strategies:

\[ (\bar{a}_1, \ldots, \bar{a}_n) \]

forms an $n\Delta(T)$-approximate Nash equilibrium.
Dynamics

Given a sequence of action profiles \( a^1, \ldots, a^T \), write \( \bar{a}_i = \frac{1}{T} \sum_{i=1}^{T} a_i^t \) to denote the mixed strategy for player \( i \) that selects an action in \( \{a_i^1, \ldots, a_i^T\} \) uniformly at random.

**Theorem**

Consider any zero sum separable graphical game \( G \). If a sequence of action profiles \( a^1, \ldots, a^T \) has regret \( \Delta(T) \), then the mixed strategies:

\[
(\bar{a}_1, \ldots, \bar{a}_n)
\]

forms an \( n\Delta(T) \)-approximate Nash equilibrium.

If every player plays using polynomial weights, they converge to an \( \epsilon \)-approximate Nash equilibrium by in:

\[
T = \frac{4n^2 \log k}{\epsilon^2}
\]

many rounds. In a two player game this is \( T = 16 \log(k)/\epsilon^2 \) steps.
Proof

1. A useful fact: for every action \( a_i^* \in A_i \) we have:

\[
\frac{1}{T} \sum_{t=1}^{T} \sum_{(i,j) \in E} u_{i,j}^t(a_i^*, a_j^t) = \sum_{(i,j) \in E} \sum_{t=1}^{T} \frac{1}{T} u_{i,j}^t(a_i^*, a_j^t) = \sum_{(i,j) \in E} u_{i,j}^t(a_i^*, \bar{a}_j)
\]
Proof

1. A useful fact: for every action $a^*_i \in A_i$ we have:

\[
\frac{1}{T} \sum_{t=1}^{T} \sum_{(i,j) \in E} u_{i,j}^{i,j}(a^*_i, a_j^t) = \sum_{(i,j) \in E} \sum_{t=1}^{T} \frac{1}{T} u_{i,j}^{i,j}(a^*_i, a_j^t) \\
= \sum_{(i,j) \in E} u_{i,j}^{i,j}(a^*_i, \overline{a}_j)
\]

2. Suppose every player $i$ is playing according to $\overline{a}_i$. Let $a^*_i$ be the best response of player $i$ to the distribution of his opponents. We know:

\[
\sum_{(i,j) \in E} u_{i,j}^{i,j}(a^*_i, \overline{a}_j) \geq \sum_{(i,j) \in E} u_{i,j}^{i,j}(\overline{a}_i, \overline{a}_j)
\]
Proof

1. We also know, since $a^1, \ldots, a^t$ have $\Delta(T)$ regret, that for all $i \in P$:

$$\frac{1}{T} \sum_{t=1}^{T} \sum_{(i,j) \in E} u_{i}^{(i,j)}(a^t_i, a^t_j) \geq \sum_{(i,j) \in E} u_{i}^{(i,j)}(a^*_i, \bar{a}_j) - \Delta(T)$$

This inequality holds because:

- **LHS** is the average of the utilities from interactions at time $t$ for all players.
- **RHS** is the sum of utilities from the optimal actions at time $t$ for all players minus the regret.
Proof

1. We also know, since $a^1, \ldots, a^t$ have $\Delta(T)$ regret, that for all $i \in P$:

$$\frac{1}{T} \sum_{t=1}^{T} \sum_{(i,j) \in E} u_i^{(i,j)}(a_i^t, a_j^t) \geq \sum_{(i,j) \in E} u_i^{(i,j)}(a_i^*, a_j^*) - \Delta(T)$$

2. Summing the LHS over all players:

$$LHS = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{(i,j) \in E} u_i^{(i,j)}(a_i^t, a_j^t) = \frac{1}{T} \sum_{t=1}^{T} 0 = 0$$

(why?)
Proof

1. For all $i \in P$:

$$\frac{1}{T} \sum_{t=1}^{T} \sum_{(i,j) \in E} u_{i,j}^{(i,j)}(a^t_i, a^t_j) \geq \sum_{(i,j) \in E} u_{i,j}^{(i,j)}(a^*_i, \bar{a}_j) - \Delta(T)$$

\[ \text{LHS} \quad \text{RHS} \]
1. For all \( i \in P \):

\[
\frac{1}{T} \sum_{t=1}^{T} \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_i^t, a_j^t) \geq \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_i^*, \bar{a}_j) - \Delta(T)
\]


2. Now summing the RHS:

\[
RHS = \sum_{i=1}^{n} \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_i^*, \bar{a}_j) - n \cdot \Delta(T)
\]
Proof

1. Combining the bounds (LHS $\geq$ RHS):
Proof

1. Combining the bounds (LHS $\geq$ RHS):

2. \[
n\Delta(T) \geq \sum_{i=1}^{n} \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_{i}^{*}, \bar{a}_{j})\]

3. (why?)

4. Let's think about each term...
Proof

1. Combining the bounds (LHS $\geq$ RHS):

2. 

$$n\Delta(T) \geq \sum_{i=1}^{n} \sum_{(i,j)\in E} u_{i}^{(i,j)}(a_i^*, \bar{a}_j)$$

$$= \sum_{i=1}^{n} \left( \sum_{(i,j)\in E} u_{i}^{(i,j)}(a_i^*, \bar{a}_j) - \sum_{(i,j)\in E} u_{i}^{i,j}(\bar{a}_i, \bar{a}_j) \right)$$
Proof

1. Combining the bounds (LHS $\geq$ RHS):

$$n \Delta(T) \geq \sum_{i=1}^{n} \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_{i}^{*}, \bar{a}_{j})$$

$$= \sum_{i=1}^{n} \left( \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_{i}^{*}, \bar{a}_{j}) - \sum_{(i,j) \in E} u_{i}^{(i,j)}(\bar{a}_{i}, \bar{a}_{j}) \right)$$

2. (why?)

3. (why?)
Proof

1. Combining the bounds (LHS $\geq$ RHS):

$$n\Delta(T) \geq \sum_{i=1}^{n} \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_{i}^{*}, \bar{a}_{j})$$

$$= \sum_{i=1}^{n} \left( \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_{i}^{*}, \bar{a}_{j}) - \sum_{(i,j) \in E} u_{i}^{(i,j)}(\bar{a}_{i}, \bar{a}_{j}) \right)$$

2. (why?)

3. (why?)

4. Lets think about each term...
Proof

\[ n\Delta(T) \geq \sum_{i=1}^{n} \left( \sum_{(i,j) \in E} u_i^{(i,j)}(a^*_i, \bar{a}_j) - \sum_{(i,j) \in E} u_i^{i,j}(\bar{a}_i, \bar{a}_j) \right) \]

1. For each term we have:

\[ \sum_{(i,j) \in E} u_i^{(i,j)}(a^*_i, \bar{a}_j) \geq 0 \]

(why?)

2. So for each player \( i \):

\[ \sum_{(i,j) \in E} u_i^{i,j}(\bar{a}_i, \bar{a}_j) \geq \sum_{(i,j) \in E} u_i^{i,j}(\bar{a}_i, \bar{a}_j) - n \Delta(T) \]

(why?)

3. Tada!
Proof

\[ n\Delta(T) \geq \sum_{i=1}^{n} \left( \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_{i}^{*}, \bar{a}_{j}) - \sum_{(i,j) \in E} u_{i}^{i,j}(\bar{a}_{i}, \bar{a}_{j}) \right) \]

1. For each term we have:

\[ \left( \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_{i}^{*}, \bar{a}_{j}) - \sum_{(i,j) \in E} u_{i}^{i,j}(\bar{a}_{i}, \bar{a}_{j}) \right) \geq 0 \]

(why?)
Proof

\[ n\Delta(T) \geq \sum_{i=1}^{n} \left( \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_i^*, \bar{a}_j) - \sum_{(i,j) \in E} u_{i}^{i,j}(\bar{a}_i, \bar{a}_j) \right) \]

1. For each term we have:

\[
\left( \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_i^*, \bar{a}_j) - \sum_{(i,j) \in E} u_{i}^{i,j}(\bar{a}_i, \bar{a}_j) \right) \geq 0
\]

(why?)

2. So for each player \(i\):

\[
\sum_{(i,j) \in E} u_{i}^{i,j}(\bar{a}_i, \bar{a}_j) \geq \sum_{(i,j) \in E} u_{i}^{(i,j)}(a_i^*, \bar{a}_j) - n\Delta(T)
\]

(why?)
Proof

\[ n\Delta(T) \geq \sum_{i=1}^{n} \left( \sum_{(i,j) \in E} u^{(i,j)}_i(a^*_i, \bar{a}_j) - \sum_{(i,j) \in E} u^{i,j}_i(\bar{a}_i, \bar{a}_j) \right) \]

1. For each term we have:

\[ \left( \sum_{(i,j) \in E} u^{i,j}_i(a^*_i, \bar{a}_j) - \sum_{(i,j) \in E} u^{i,j}_i(\bar{a}_i, \bar{a}_j) \right) \geq 0 \]

(why?)

2. So for each player \(i\):

\[ \sum_{(i,j) \in E} u^{i,j}_i(\bar{a}_i, \bar{a}_j) \geq \sum_{(i,j) \in E} u^{(i,j)}_i(a^*_i, \bar{a}_j) - n\Delta(T) \]

(why?)

3. Tada!
Thanks!

See you next class — stay healthy!