Convergence of No Regret Dynamics to Equilibrium in Separable Multi-player Zero Sum Games

Last class we saw that *two-player* zero sum games are special. Among other things:

1. They have a $\max\min = \min\max$ *value*. In particular, all equilibria have the same payoff (the value), and players can achieve the same payoff whether or not they need to first reveal their strategy to their opponent and commit to not changing it.

2. Equilibria can be computed "easily" – each player can compute their portion of an equilibrium strategy without needing to reason about exactly what their opponent is doing. There is no equilibrium selection problem.

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

Definition 1 *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. We may observe the following "Meta Theorem":

"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. The reason is easy to see:
"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.

Still, there are some cases in which we can generalize the interesting properties of two-player zero-sum games to multi-player games. Here is one of the most general:

Definition 2 *A **separable graphical game** is defined by a graph $G = (V, E)$. The set of players corresponds to the set of vertices V. Each player's utility function is decomposable as a sum of neighbor-specific utility functions, one for each of his neighbors in G:*

$$u_i(a) = \sum_{(i,j) \in E} u_i^{(i,j)}(a_i, a_j)$$

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.

Zero sum graphical games share many of the properties of two-player zero sum games: they have a value, and equilibria are easy to compute with efficient dynamics (in particular, the polynomial weights algorithm will converge to equilibrium when used to play the game by all players). Note that we do not require that each of the two-player games in the decomposition be zero sum – just that the game in aggregate is.

Definition 3 *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 $\Delta(T) = o_T(1)$.*

For example, one simple way to generate a no-regret sequence of actions in any game is to have each player play according to the polynomial weights algorithm. Recall that in this case, we get a sequence of actions with regret $\Delta(T) \leq 2\sqrt{\frac{\log k}{T}}$ in a k action game.

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 4 *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.

Remark   Note that if we have a way of generating a no-regret sequence of actions, then if we run the process for long enough, we will converge to arbitrarily fine approximations to Nash equilibrium in such games. For example, consider the polynomial weights algorithm, which has $\Delta(T) = 2\sqrt{\frac{\log k}{T}}$. Setting:

$$n \cdot 2\sqrt{\frac{\log k}{T}} \leq \epsilon$$

we find that we converge to an ϵ-approximate Nash equilibrium by having every player play according to the polynomial weights algorithm for:

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

rounds. This is fast – polynomial in all of the parameters of the problem. It is particularly fast in a two player game – it converges in $16 \log(k)/\epsilon^2$ steps.

Proof   We first note a useful fact: for every action $a_i^* \in A_i$ we have:

$$
\begin{aligned}
\frac{1}{T} \sum_{t=1}^{T} \sum_{(i,j) \in E} u_i^{i,j}(a_i^*, a_j^t) &= \sum_{(i,j) \in E} \sum_{t=1}^{T} \frac{1}{T} u_i^{i,j}(a_i^*, a_j^t) \\
&= \sum_{(i,j) \in E} u_i^{i,j}(a_i^*, \bar{a}_j)
\end{aligned}
$$

Now suppose every player i is playing according to $\bar{a}_i$. Let a_i^* be the best response of player i to the distribution of his opponents. By definition, we know:

$$\sum_{(i,j) \in E} u_i^{i,j}(a_i^*, \bar{a}_j) \geq \sum_{(i,j) \in E} u_i^{i,j}(\bar{a}_i, \bar{a}_j)$$

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

$$\underbrace{\frac{1}{T} \sum_{t=1}^{T} \sum_{(i,j) \in E} u_i^{(i,j)}(a_i^t, a_j^t)}_{LHS} \geq \underbrace{\sum_{(i,j) \in E} u_i^{(i,j)}(a_i^*, \bar{a}_j) - \Delta(T)}_{RHS}$$

Consider the left hand side of the above inequality, summed over all players i:

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

since the game is zero sum.

Now consider the right hand side summed over all n players:

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

Combining, since $LHS \geq RHS$ we get:

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

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

where the last equality again follows from the fact that the game is zero sum – the sum of the terms we added is 0.

Term by term, however, we have the following, by definition of a best response:

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

Therefore, it must be that for all players i:

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

which completes the proof. ∎