A Simple Adaptive Procedure Leading To Correlated Equilibrium

Sergui Hart and Andreu Mas-Colell

Presented by: Arindam Banerjee
Nash Equilibrium (NE)

- \( \Gamma = (n, (S_i)_1^n, (u_i)_1^n) \) be a finite \( N \)-person game
- Let \( s_i \in S_i \) be the strategy of the \( i \)-th player
- Then, \( s = (s_i)_1^n \) is a (pure) strategy profile
- A strategy profile \( s^* \) is a NE if \( \forall i \)
  \[
  u_i(s^*) \geq u_i(s_i, s^*_{-i})
  \]
- In general, strategy profile may be pure or mixed
Theorem: Every finite $n$-player game admits at least one NE

The argument is based on a fixed point theorem

- Kakutani fixed point theorem
- Generalizes Brouwer’s fixed point theorem

Existence of fixed point implies existence of NE
Correlated Equilibrium (CE)

- A probability distribution \( \psi \) on \( S \), if \( \forall i, j, k, \)

\[
E_\psi [u_i(k, s_{-i}) - u_i(s)] \leq 0
\]

- NE implies a product of marginals
- CE is a general joint distribution
- All NE are CE
- The set of all CE is a nonempty, closed, convex set (polytope)
- The polytope may include points beyond convex hull of NE
A Simple Adaptive Algorithm

- A simple local strategy for repeated play
- The regret of $i$ for using $j$ instead of $k$

$$R^t_{i}(j, k) = \left[ \frac{1}{t} \sum_{\tau \leq t : s^\tau_i = j} \left[ u_i(k, s^\tau_{i-1}) - u^i(s^\tau) \right] \right] +$$

- For a large enough $\mu$,

$$p^{t+1}_i(k) = \begin{cases} \frac{1}{\mu} R^t_{i}(j, k), & \forall k \neq j \\ 1 - \sum_{k \in S_i : k \neq j} p^{t+1}_i(k) \end{cases} \quad (1)$$

- The empirical distribution of past plays

$$z^t(s) = \frac{1}{t} |\{\tau \leq t : s^\tau = s\}|$$

converges (a.s.) to the set of correlated equilibrium
Discussion

- The empirical distribution is \( \epsilon \)-close to the set of CE
- Hence, any projection (expectation) is \( \epsilon \)-close
- As a result, we get correlated \( \epsilon \)-equilibrium
- Randomization is over positive regret strategies
- Different from online (no-regret) algorithms we had seen earlier
- There is a inertia parameter \( \mu \)
No Regret and Stochastic Transition

- Strategy of player $i$ as an invariant probability distribution

$$q^t_i(j) = \sum_{k \neq j} q^t_i(k) \frac{1}{\mu} R^i_t(j, k) + q^t_i(j)[1 - \sum_{k \neq j} \frac{1}{\mu} R^i_t(j, k)]$$

$$\sum_k q^t_i(k) R^t_i(j, k) = q^t_i(j) \sum_k R^t_i(j, k)$$

- It is a solution of the system $\bar{R}q = q$ where $\bar{R}$ is a stochastic transition matrix

**Theorem:** If $p^{t+1} = q$ satisfies the above equation, then

$$\forall j, k, j \neq k, R^t_i(j, k) \to 0$$

- In this no-regret strategy, it does not matter what the other players are doing
Theorem: No Regret ⇔ empirical distribution “converges” to the set of CE

As a result, any no regret update algorithm leads to convergence to the set of CE

Theorem: In particular, the eigenvector update leads the empirical distribution to the set of CE

The proof is using Blackwell’s approachability theorem

It also gives rate of convergence:
- Expected regret converges at $O(1/\sqrt{t})$ (best possible)
- Also $z_t, t \geq T$, is a correlated $\epsilon$-equilibrium with probability at least $(1 - ce^{-cT})$
Discussion I

- For best response to calibrated forecasts, the empirical play converges to CE

- Calibrated smooth fictitious play leads to the set of CE

- Universal Consistency
  - There exists a procedure that ensures that \( \limsup \) of the max unconditional (external) regret becomes less/equal to 0
  - The eigenvector procedure is universally consistent
  - There are simpler procedures that are universally consistent, e.g., probabilities proportional to unconditional regrets

- Smooth Fictitious Play
  - Regularized fictitious play
  - It is universally \( \epsilon \)-consistent
  - Leads to correlated \( \epsilon \)-equilibrium
Discussion II

- Universal Calibration
  - All conditional (internal) regrets must go to 0
  - Eigenvector method and calibrated smooth fictitious play are universally calibrated
  - The simple procedure is not universally calibrated

- Better/Best response: The simple procedure considers better responses rather than only the best

- Eigenvector procedures: The simple procedure is much faster (per round) than the eigenvector procedures

- Inertia is due to $\mu$, converges for all $\mu$, although convergence rate depends on the value $\mu$

- “Friction” is necessary for the theorem to hold

- An extension to an unknown game (bandit setting) is possible