Game Theory

Department of Electronics

EL-766

Spring 2011

Lecture 18

May 11, 2011

Lecture Outline

- Learning in games
  - Fictitious play – existence of Nash eq.
  - No regret learning – an introduction

- Correlated Nash equilibrium
- Types of games
  - Dummy, coordination games
  - Super-modular games
  - Potential Games

Fictitious Play: Review

- Player i: initial weight function

$$K_0^i : S^{-i} \rightarrow \mathcal{R}^+$$

- Game iteratively repeated → K updated:

$$K_t(s^{-i}) = K_{t-1}(s^{-i}) + \begin{cases} 1, & \text{if } \; s_{t-1}^{-i} = s^{-i} \\ 0, & \text{ow.} \end{cases}$$

- Given the frequency vector K → updates beliefs
  - The belief player i has at time t about its opponent to play s^{-i} at time t:

$$\gamma_t^i\left(s^{-i}\right) = \frac{K_t^i\left(s^{-i}\right)}{\sum_{\hat{s} \in S^{-i}} K_t^i\left(\hat{s}^{-i}\right)}$$

← Simple normalization

Fictitious play

- Given the updated belief γ_t^i
- Fictitious play: any rule $p_t^i(\gamma_t^i) \in BR^i(\gamma_t^i)$
- Not a unique fictitious play rule → there may be more than one best response to a particular assessment

Convergence properties for fictitious play

- Proposition:
  - (1) If s is a strict Nash equilibrium, and s is played at date t in the process of fictitious play, then s is played at all subsequent dates (Nash equilibria are absorbing for the process of fictitious play)
  - (2) Any pure-strategy steady state of fictitious play must be Nash equilibrium
- Definition: marginal empirical distribution of play:

$$d_t^j(s^j) = \frac{k_t(s^j) - k_0(s^j)}{t}$$

- Proposition:
  - If the empirical distributions over each player's choices converge, the strategy profile corresponding to the product of these distributions is a Nash eq.
- Proposition: the fictitious play process converges for a two person zero-sum game

No Regret Learning

- **No regret learning strategies:** probabilistic learning alg. which specify that players ***explore*** the space of actions by playing all actions with some non-zero probability, and ***exploit*** successful actions, by increasing the probability of employing those actions that generate high profits.
  - Learn mixed strategy equilibria
- No external regret algorithms
  - **External regret:** difference between the payoffs achieved by the strategies prescribed by the given algorithm, and the payoffs achieved by any other fixed sequence of decisions, in the worst case.
- No internal regret algorithms
  - **Internal regret:** difference between the payoffs achieved by the strategies prescribed by the given algorithm, and the payoffs that could be achieved by a remapped sequence of strategies.
    - Sequence is remapped, if there is a mapping f of the strategy space into itself, s.t. for each occurrence of strategy x (e.g. x=169), the mapped strategy y (e.g. y = 170) appears in the re-mapped sequence
  - No internal regret → no external regret

No External Regret Algorithm – an example alg.

- Freud & Schapire:
  - "Game theory, on-line prediction and boosting", Proc. of the 9th Annual Conference on Computational Learning Theory, pp. 325-332, ACM press, May 1996.
- No external regret via multiplicative updating
- ρ_i^t = cumulative payoff obtained through time t via strategy i:

$$\rho_i^t = \sum_{x=0}^{t} r_i^x$$

- The weight associated with strategy i, for $\beta \in [0,1)$:

$$w_i^{t+1} = \frac{(1+\beta)^{\rho_i^t}}{\sum_{j=1}^{S} (1+\beta)^{\rho_j^t}}$$

No external regret –cont.

- The weight calculation represents a measure of regret
- Need to know payoff that would be obtained for all possible strategies
- Naive player: knows only his payoff for the currently played strategy
  - Previous algorithm may be modified to work for naive players: [Auer, Chesa-Bianchi, Freud and Schapire, 1995]
    - Gambling in a rigged casino: the adversarial multi-armed bandit problem, proc. of the 36th Annual Symposium on Foundations of Computer Science, pp. 322-331. ACM Press, Nov. 1995.

No internal regret – an example alg.

- The regret for a player at time t: difference between the payoffs achieved using its strategy of choice, e.g. i, and the payoffs that could have been achieved had strategy $j \neq i$ been played instead:

$$R_{i \to j}^{t} = 1_{i}^{t}\left(r_{j}^{t} - r_{i}^{t}\right)$$

Indicator function: 1 if strategy i employed at t, 0 ow.

- Cumulative regret:

$$R_{i \to j}^{T} = \sum_{t=0}^{T} R_{i \to j}^{t}$$

- Internal regret:

$$IR_{i \to j}^{T} = (R_{i \to j}^{T})^{+}$$

$$X^{+} = \max(X, 0)$$

No internal regret – cont.

- Cumulative regret for playing all strategies but j:

$$IR^T_{S\to j} = \sum_{i=1}^{S} IR^T_{i\to j}$$

- Updating weights
  - If strategy i is played at time t,

$$w_j^{t+1} = \frac{1}{\mu} IR^t_{i\to j}$$

$$w_i^{t+1} = 1 - \sum_{j\neq i} w_j^{t+1}$$

  - μ = normalizing term:

$$\mu > (|S|-1)\max_{j\in S} IR^t_{i\to j}$$

No internal regret alg. – cont.

Some observations:

- Achieves no internal regret in the limit as T→ ∞
- Learning converges to the correlated Nash equilibrium
- Naive version also proposed
- Reference:
  - S. Hart & A. Mass Colell, “A simple adaptive procedure leading to correlated equilibrium”.

Correlated Nash equilibrium

- If players can engage in preplay communication, then go in separate rooms and choose their strategy independently
  - Might gain if they can build a signaling device → send signals to the separate rooms.
- Example of game with correlated equilibria

	L	R
U	5,1	0,0
D	4,4	1,5

  - Three Nash equilibria: (U,L), (D,R), and a mixed strategy eq. with equal probability on each pure strategy: payoff 2.5 for each player

Correlated Nash equilibria: cont

- If they can jointly observe a random variable, e.g. a coin flip
  - Player 1: U if heads, D if tails
  - Player 2: L if heads, R if tails

    → Payoff: (3,3)
  - More general: players can obtain any payoff vector in the convex hull of the set of Nash equilibrium payoffs
  - Cannot obtain any payoff outside the convex hull of the set of Nash equilibrium payoffs
- Can gain further if players receive different signals, but correlated
  - Build a signaling device that sends different, but correlated, signals to each of them
  - Device has three equally likely states: A, B, C
  - If A – player 1 completely informed, but it cannot distinguish between B and C
  - Player 2 – informed when C, cannot distinguish from A and B

Correlated Nash eq. – cont.

- A Nash eq. for the transformed game:
  - Player 1 plays U when A and D ow
  - Player 2: R when C, and L ow
  - Eq: (U,L), (D,L) and (D,R), occur with probability 1/3, payoff 3.33
  - Payoff outside the convex hull of the set of Nash equilibrium payoffs
- Note: signaling device may be interpreted as a recommendation on how to play
- Definition: A correlated eq. is any probability distribution p(.) over the pure strategies S1 x S2 x… Si, s.t. for every player i and every si, with p(si)>0,

$$\sum_{s_{-i} \in S_{-i}} p(s_{-i} \mid s_i) u_i(s_i, s_{-i}) \geq \sum_{s_{-i} \in S_{-i}} p(s_{-i} \mid s_i) u_i(s_i', s_{-i}) \quad \forall \; s_{i'} \in S_i$$

- Player i should not be able to gain by disobeying the recommendation to play s_i, if every other player obeys its recommendation