Learning in games

• Why learning?
  – For introspection, the rules of the game, rationality of the players, payoff functions – all common knowledge
  – Another problem: for multiple equilibria, how players come to expect the same equilibrium?

• Applicability
• Repeated games
• Teach opponent to play a best response to a particular action, by repeating it over and over again
Example of sophisticated learning

- How would you play this game, if you were player 1?

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>1,0</td>
<td>3,2</td>
</tr>
<tr>
<td>D</td>
<td>2,1</td>
<td>4,0</td>
</tr>
</tbody>
</table>
Sophisticated learning?

- Most learning theory → models for which the incentive is small to alter the future play of the opponents.

  - Examples:

    - large anonymous population: population size large compared to the discount factor
    - Players locked in their choices and discount factor small compared to maximum speed at which the system can possibly adjust
Common models for learning

- **Fictitious play**
  - Players observe only their own matches and play a best response to the historical frequency of play

- **Partial best-response**
  - A fixed portion of users switches each period from its current action to a BR to the aggregate statistics from the previous period

- **Replicator Dynamics**
  - The fraction of the population using a given strategy, grows proportionally to that strategy’s current payoff.
One type of learning: Cournot adjustment

- Unique Nash eq. is at the intersection of the reaction curves

\[ q_1 = r_1(q_2) \]
\[ q_2 = r_2(q_1) \]

The process converges to Nash equilibrium from any starting point → eq. globally stable
Fictitious play

- Repeated game
- Stationary assumption
- Each player: belief of opponents “strategy” by looking at what happened
- Player then plays best response (BR) according the his belief
- Belief: a prediction of the opponent action distribution, i.e. the degree to which player i believes player j will play a certain action.
- Players choose their actions in each period, s.t. to maximize their expected payoff, with respect to their belief for the current period.
Updating beliefs

- Player $i$: initial weight function
  \[ K^i_0 : S^{-i} \to \mathcal{R}^+ \]

- Game iteratively repeated $\Rightarrow K$ 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 $K$ $\Rightarrow$ updates beliefs
  - The belief player $i$ has at time $t$ about its opponent to play $s^{-i}$ at time $t$:
    \[ \gamma^i_t(s^{-i}) = \frac{K^i_t(s^{-i})}{\sum_{\hat{s} \in S^{-i}} K^i_t(\hat{s}^{-i})} \quad \text{Simple normalization} \]