Learning in Games

Introduction and Basic Concepts

David K. Levine
October 5, 2010
Definition of Extensive Form Game

a finite game tree $X$ with nodes $x \in X$

nodes are partially ordered and have a single root (minimal element)

terminal nodes are $z \in Z$ (maximal elements)
Players and Information Sets

player 0 is nature

information sets \( h \in H \) are a partition of \( X \setminus Z \)

*each node in an information set must have exactly the same number of immediate followers*

each information set is associated with a unique player who “has the move” at that information set

\( H_i \subset H \) information sets where \( i \) has the move
More Extensive Form Notation

information sets belonging to nature $h \in H_0$ are singletons

$A(h)$ feasible actions at $h \in H$

each action and node $a \in A(h), x \in h$ is associated with a unique node that immediately follows $x$ on the tree

each terminal node has a payoff $r_i(z)$ for each player

by convention we designate terminal nodes in the diagram by their payoffs
Example: a simple simultaneous move game
Behavior Strategies

A *pure strategy* is a map from information sets to feasible actions
\[ s_i(h_i) \in A(h_i) \]

\( S_i \) are the set of pure strategies

\[ \sigma_i \in \Sigma_i \] are mixed strategies, probability distributions over pure strategies

A *behavior strategy* is a map from information sets to probability distributions over feasible actions
\[ \pi_i(h_i) \in P(A(h_i)) \]

*Nature’s move* is a behavior strategy for Nature and is a fixed part of the description of the game.

We may now define \( u_i(\pi) \)

*normal form* are the payoffs \( u_i(s) \) derived from the game tree
(1,1) (2,2) (3,3) (4,4)

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>1,1</td>
<td>2,2</td>
</tr>
<tr>
<td>D</td>
<td>3,3</td>
<td>4,4</td>
</tr>
</tbody>
</table>
Kuhn’s Theorem

every mixed strategy gives rise to a unique behavior strategy

\[ \hat{\pi}(h_i \mid \sigma_i) \] map from mixed to behavior strategies

The converse is NOT true

however: if two mixed strategies give rise to the same behavior strategy, they are equivalent, that is they yield the same payoff vector for each opponents profile

\[ u(\sigma_i, s_{-i}) = u(\sigma'_i, s_{-i}) \]
Additional Notation

$\overline{H}(\sigma)$ reached with positive probability under $\sigma$

$\hat{\rho}(\pi), \hat{\rho}(\sigma) \equiv \hat{\rho}(\hat{\pi}(\sigma))$ distribution over terminal nodes

$\mu_i$ a probability measure on $\Pi_{-i}$

$u_i(s_i | \mu_i)$ preferences

$\Pi_{-i}(\sigma_{-i} | J) \equiv \{\pi_{-i} | \pi_i(h_i) = \hat{\pi}(h_i | \sigma_i), \forall h_i \in H_{-i} \cap J \}$
Nash Equilibrium

a mixed profile $\sigma$ such that for each $s_i \in \text{supp}(\sigma_i)$ there exist beliefs $\mu_i$ such that

- $s_i$ maximizes $u_i(\cdot | \mu_i)$
- $\mu_i(\Pi_{-i}(\sigma_{-i} | H)) = 1$
Why Might We Be At Nash Equilibrium?

The rush hour traffic game
Potential games
Dynamics versus statics: two different questions
- What sort of outcomes can arise from asymptotic of learning? Nash? Self-confirming?
- What does the adjustment path look like?
Focus on statics first
Active versus passive learning
Unitary Self-Confirming Equilibrium

What does learning tell us in extensive form games?

\[ \mu_i (\Pi_{-i} (\sigma_{-i} \mid \bar{H}(\sigma))) = 1 \]

**Theorem:** *Path equivalent to Nash equilibrium when there are two players*

Why?
A₁, A₂ is self-confirming, but not Nash.

any strategy for 3 makes it optimal for either 1 or 2 to play down.

but in self-confirming, 1 can believe 3 plays R; 2 that he plays L.
Heterogeneous Self-Confirming equilibrium

- \( \mu_i(\Pi_{-i}(\sigma_{-i} \mid \bar{H}(s_i, \sigma))) = 1 \)
The “observation function”

\[ J(s_i, \sigma) = H, \bar{H}(\sigma), \bar{H}(s_i, \sigma) \]
Remark: In games with perfect information, the set of heterogeneous self-confirming equilibrium payoffs (and the probability distributions over outcomes) are convex.
Example Without Public Randomization

(6,8,6,8)
pass

4
drop→(2,3,4,3)
pass (50%)

3
drop→(0,4,5,4)
pass (50%)

2
drop→(7,5,3,5)
Knowing and Unknowing Losses

The relative importance of learning
Ultimatum Bargaining Results

A

($10.00 - x, x)

R

(0, 0)
Raw US Data for Ultimatum

<table>
<thead>
<tr>
<th>x</th>
<th>Offers</th>
<th>Rejection Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2.00</td>
<td>1</td>
<td>100%</td>
</tr>
<tr>
<td>$3.25</td>
<td>2</td>
<td>50%</td>
</tr>
<tr>
<td>$4.00</td>
<td>7</td>
<td>14%</td>
</tr>
<tr>
<td>$4.25</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>$4.50</td>
<td>2</td>
<td>100%</td>
</tr>
<tr>
<td>$4.75</td>
<td>1</td>
<td>0%</td>
</tr>
<tr>
<td>$5.00</td>
<td>13</td>
<td>0%</td>
</tr>
</tbody>
</table>

27

US $10.00 stake games, round 10
<table>
<thead>
<tr>
<th>Trials</th>
<th>Rnd</th>
<th>Cntry</th>
<th>Case</th>
<th>Expected Loss</th>
<th>Max Gain</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pl 1</td>
<td>Pl 2</td>
<td>Both</td>
</tr>
<tr>
<td>27</td>
<td>10</td>
<td>US</td>
<td>H</td>
<td>$0.00</td>
<td>$0.67</td>
<td>$0.34</td>
</tr>
<tr>
<td>27</td>
<td>10</td>
<td>US</td>
<td>U</td>
<td>$1.30</td>
<td>$0.67</td>
<td>$0.99</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>USx3</td>
<td>H</td>
<td>$0.00</td>
<td>$1.28</td>
<td>$0.64</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>USx3</td>
<td>U</td>
<td>$6.45</td>
<td>$1.28</td>
<td>$3.86</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
<td>Yugo</td>
<td>H</td>
<td>$0.00</td>
<td>$0.99</td>
<td>$0.50</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
<td>Yugo</td>
<td>U</td>
<td>$1.57</td>
<td>$0.99</td>
<td>$1.28</td>
</tr>
<tr>
<td>29</td>
<td>10</td>
<td>Jpn</td>
<td>H</td>
<td>$0.00</td>
<td>$0.53</td>
<td>$0.27</td>
</tr>
<tr>
<td>29</td>
<td>10</td>
<td>Jpn</td>
<td>U</td>
<td>$1.85</td>
<td>$0.53</td>
<td>$1.19</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
<td>Isrl</td>
<td>H</td>
<td>$0.00</td>
<td>$0.38</td>
<td>$0.19</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
<td>Isrl</td>
<td>U</td>
<td>$3.16</td>
<td>$0.38</td>
<td>$1.77</td>
</tr>
<tr>
<td>WC</td>
<td></td>
<td></td>
<td>H</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Rnds=Rounds, WC=Worst Case, H=Heterogeneous, U=Unitary
Comments on Ultimatum

- every offer by player 1 is a best response to beliefs that all other offers will be rejected so player 1’s heterogeneous losses are always zero.
- big player 1 losses in the unitary case
- player 2 losses all knowing losses from rejected offers; magnitudes indicate that “subgame perfection” does quite badly; but really a matter of social preference
- tripling the stakes increases the size of losses a bit less than proportionally (losses roughly double)
- key fact: unknowing losses considerably larger than knowing losses – relative importance of learning
This game has a unique self-confirming equilibrium; in it player 1 with probability 1 plays $T_1$
### Summary of Experimental Results

<table>
<thead>
<tr>
<th>Trials Rnd</th>
<th>Rnds</th>
<th>Stake</th>
<th>Expected Loss</th>
<th>Max Gain</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>PI 1</td>
<td>PI 2</td>
<td>Both</td>
</tr>
<tr>
<td>29*</td>
<td>6-10</td>
<td>1x</td>
<td>H</td>
<td>$0.00</td>
<td>$0.03</td>
</tr>
<tr>
<td>29*</td>
<td>6-10</td>
<td>1x</td>
<td>U</td>
<td>$0.26</td>
<td>$0.17</td>
</tr>
<tr>
<td>WC</td>
<td>1x</td>
<td>H</td>
<td></td>
<td>$0.80</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>1-10</td>
<td>1x</td>
<td>H</td>
<td>$0.00</td>
<td>$0.08</td>
</tr>
<tr>
<td>10</td>
<td>1-10</td>
<td>4x</td>
<td>H</td>
<td>$0.00</td>
<td>$0.28</td>
</tr>
</tbody>
</table>

Rnds=Rounds, WC=Worst Case, H=Heterogeneous, U=Unitary

*The data on which from which this case is computed is reported above.
Comments on Experimental Results

- heterogeneous loss per player is small; because payoffs are doubling in each stage, equilibrium is very sensitive to a small number of player 2’s giving money away at the end of the game.
- unknowing losses far greater than knowing losses
- quadrupling the stakes very nearly causes $\epsilon$ to quadruple
- theory has substantial predictive power: see WC
- losses conditional on reaching the final stage are quite large--inconsistent with “subgame perfection” indicative however of social preference. McKelvey and Palfrey estimated an incomplete information model where some “types” of player 2 liked to pass in the final stage. This cannot explain many players dropping out early so their estimated model fits poorly.