Learning mixed equilibria

Fudenberg and Kreps (1993)

Geelon So, agso@ucsd.edu
Game theory reading group — February 13, 2023

Learning in strategic settings

Question
If players play a game repeatedly, how do their beliefs and behaviors evolve over time?

Game setting

Alice and Bob repeatedly play a two-player game

Game setting

Alice and Bob repeatedly play a two-player game

- Let $S = S^{(A)} \times S^{(B)}$ be the set of pure-strategy profiles

Game setting

Alice and Bob repeatedly play a two-player game

- Let $S = S^{(A)} \times S^{(B)}$ be the set of pure-strategy profiles
- Let $\Sigma = \Sigma^{(A)} \times \Sigma^{(B)}$ be the set of mixed-strategy profiles

Game setting

Alice and Bob repeatedly play a two-player game

- Let $S = S^{(A)} \times S^{(B)}$ be the set of pure-strategy profiles
- Let $\Sigma = \Sigma^{(A)} \times \Sigma^{(B)}$ be the set of mixed-strategy profiles
- Let $u = (u^{(A)}, u^{(B)})$ be their utilities

A model of learning and behavior

Suppose that $\zeta_T = (\mathbf{s}_1, \ldots, \mathbf{s}_T)$ is a history of observed plays $\mathbf{s}_t \in S^{(A)} \times S^{(B)}$.

A model of learning and behavior

Suppose that $\zeta_T = (\mathbf{s}_1, \ldots, \mathbf{s}_T)$ is a history of observed plays $\mathbf{s}_t \in S^{(A)} \times S^{(B)}$.

- **assessment rule:** Alice predicts that Bob will play according to:

$$\mu_T^{(A)} \equiv \mu^{(A)}(\zeta_T) \in \Sigma^{(B)}.$$

A model of learning and behavior

Suppose that $\zeta_T = (\mathbf{s}_1, \ldots, \mathbf{s}_T)$ is a history of observed plays $\mathbf{s}_t \in S^{(A)} \times S^{(B)}$.

- **assessment rule:** Alice predicts that Bob will play according to:

$$\mu_T^{(A)} \equiv \mu^{(A)}(\zeta_T) \in \Sigma^{(B)}.$$

- **behavior rule:** Alice makes a play according to:

$$\phi_T^{(A)} \equiv \phi^{(A)}(\zeta_T) \in \Sigma^{(A)}.$$

A model of learning and behavior

Suppose that $\zeta_T = (\mathbf{s}_1, \ldots, \mathbf{s}_T)$ is a history of observed plays $\mathbf{s}_t \in S^{(A)} \times S^{(B)}$.

- **assessment rule:** Alice predicts that Bob will play according to:

$$\mu_T^{(A)} \equiv \mu^{(A)}(\zeta_T) \in \Sigma^{(B)}.$$

- **behavior rule:** Alice makes a play according to:

$$\phi_T^{(A)} \equiv \phi^{(A)}(\zeta_T) \in \Sigma^{(A)}.$$

Alice can also have a *subjective belief* about how Bob learns and behaves.

Fictitious play

Fictitious play: a simple model of learning and behavior

Suppose Bob has K pure strategies.

Fictitious play: a simple model of learning and behavior

Suppose Bob has K pure strategies.

- Alice initializes K counters: $N_1, \ldots, N_K \leftarrow 1$

Fictitious play: a simple model of learning and behavior

Suppose Bob has K pure strategies.

- Alice initializes K counters: $N_1, \ldots, N_K \leftarrow 1$
- For round $t = 1, 2, \ldots$

Fictitious play: a simple model of learning and behavior

Suppose Bob has K pure strategies.

- Alice initializes K counters: $N_1, \ldots, N_K \leftarrow 1$
- For round $t = 1, 2, \ldots$
  - Alice computes empirical mixed strategy μ_t:

$$(t + K) \cdot \mu_t = (N_1, \ldots, N_K).$$

Fictitious play: a simple model of learning and behavior

Suppose Bob has K pure strategies.

- Alice initializes K counters: $N_1, \ldots, N_K \leftarrow 1$
- For round $t = 1, 2, \ldots$
  - Alice computes empirical mixed strategy μ_t:

$$(t + K) \cdot \mu_t = (N_1, \ldots, N_K).$$

  - Alice plays the best response to assessment μ_t:

$$\phi_t = \arg\max_{\sigma \in \Sigma^{(A)}} u^{(A)}(\sigma, \mu_t).$$

Fictitious play: a simple model of learning and behavior

Suppose Bob has K pure strategies.

- Alice initializes K counters: $N_1, \ldots, N_K \leftarrow 1$
- For round $t = 1, 2, \ldots$
  - Alice computes empirical mixed strategy μ_t:
$$(t + K) \cdot \mu_t = (N_1, \ldots, N_K).$$
  - Alice plays the best response to assessment μ_t:
$$\phi_t = \underset{\sigma \in \Sigma^{(A)}}{\arg\max}\ u^{(A)}(\sigma, \mu_t).$$
  - Alice observes Bob's actual play s, and updates the counter for s:
$$N_s \leftarrow N_s + 1.$$

Connection to Bayesian players

The assessments μ_t are consistent with a Bayesian model:

Connection to Bayesian players

The assessments μ_t are consistent with a Bayesian model:

- Alice believes that Bob is playing by drawing i.i.d. from a fixed strategy $\sigma^* \in \Sigma^{(B)}$,

Connection to Bayesian players

The assessments μ_t are consistent with a Bayesian model:

- Alice believes that Bob is playing by drawing i.i.d. from a fixed strategy $\sigma^* \in \Sigma^{(B)}$,
- Alice's prior over assessments is a Dirichlet distribution.

Connection to Bayesian players

The assessments μ_t are consistent with a Bayesian model:

- Alice believes that Bob is playing by drawing i.i.d. from a fixed strategy $\sigma^* \in \Sigma^{(B)}$,
- Alice's prior over assessments is a Dirichlet distribution.

Remark. Alice believes that Bob is non-adaptive, and therefore plays *myopically*.

- This belief justifies maximizing immediate utility.

Strict Nash equilibria are absorbing states

Proposition

Let $\mathbf{s}_1, \mathbf{s}_2, \ldots$ *be a history generated by fictitious play.*

Strict Nash equilibria are absorbing states

Proposition

Let $\mathbf{s}_1, \mathbf{s}_2, \ldots$ *be a history generated by fictitious play.*

- *Strict Nash equilibria are absorbing states.*

Strict Nash equilibria are absorbing states

Proposition

Let $\mathbf{s}_1, \mathbf{s}_2, \ldots$ be a history generated by fictitious play.

- *Strict Nash equilibria are absorbing states.*
  - *If s_t is a strict Nash equilibrium, then:*

$$\mathbf{s}_{t'} = \mathbf{s}_t, \qquad \text{for all } t' \geq t.$$

Strict Nash equilibria are absorbing states

Proposition

Let $\mathbf{s}_1, \mathbf{s}_2, \ldots$ be a history generated by fictitious play.

- *Strict Nash equilibria are absorbing states.*
  - *If s_t is a strict Nash equilibrium, then:*

$$\mathbf{s}_{t'} = \mathbf{s}_t, \qquad \text{for all } t' \geq t.$$

- *Absorbing states are Nash equilibria.*

Strict Nash equilibria are absorbing states

Proposition

Let $\mathbf{s}_1, \mathbf{s}_2, \ldots$ be a history generated by fictitious play.

- *Strict Nash equilibria are absorbing states.*
  - *If s_t is a strict Nash equilibrium, then:*

$$\mathbf{s}_{t'} = \mathbf{s}_t, \qquad \text{for all } t' \geq t.$$

- *Absorbing states are Nash equilibria.*
  - *If the above equation holds, then $\mathbf{s}_t$ is a Nash equilibrium.*

Convergence toward mixed Nash equilibrium

Proposition

Suppose that in some history generated by fictitious play, the empirical frequencies converge to some mixed strategy profile. Then, that strategy profile is a Nash equilibrium.

Convergence toward mixed Nash equilibrium

Proposition

Suppose that in some history generated by fictitious play, the empirical frequencies converge to some mixed strategy profile. Then, that strategy profile is a Nash equilibrium.

Remark. The convergence of empirical frequencies is not guaranteed.

- Shapley (1964) gives the first nontrivial example; rock–paper–scissors.
- Convergence is guaranteed for zero-sum games (Robinson, 1951) and two-by-two games (Miyasawa, 1961).

Extensions of fictitious play

Myopic behavior

Definition

Let μ be the assessment rule for Alice. Her behavior rule ϕ is myopic relative to μ if:

$$\phi(\zeta_t) = \underset{\sigma \in \Sigma^{(A)}}{\arg\max}\ u^{(A)}\big(\sigma, \mu(\zeta_t)\big).$$

That is, she maximizes her immediate payoff with respect to her assessment $\mu(\zeta_t)$ at time t.

Myopic behavior

Definition

Let μ be the assessment rule for Alice. Her behavior rule ϕ is myopic relative to μ if:

$$\phi(\zeta_t) = \arg\max_{\sigma \in \Sigma^{(A)}} \, u^{(A)}\big(\sigma, \mu(\zeta_t)\big).$$

That is, she maximizes her immediate payoff with respect to her assessment $\mu(\zeta_t)$ at time t.

- The behavior rule of fictitious play is myopic.
- We can weaken myopia to an asymptotic notion.

Asymptotic myopia

Definition

At each time step, let the optimal expected utility with respect to the assessment $\mu(\zeta_t)$ be:

$$u_t^* = \max_{\sigma \in \Sigma^{(A)}} u^{(A)}\big(\sigma, \mu(\zeta_t)\big).$$

Asymptotic myopia

Definition

At each time step, let the optimal expected utility with respect to the assessment $\mu(\zeta_t)$ be:

$$u_t^* = \max_{\sigma \in \Sigma^{(A)}} u^{(A)}\big(\sigma, \mu(\zeta_t)\big).$$

- *ϕ is asymptotically myopic if:*

$$\lim_{t \to \infty} \Big(u_t^* - u^{(A)}\big(\phi(\zeta_t), \mu(\zeta_t)\big)\Big) = 0.$$

Asymptotic myopia

Definition

At each time step, let the optimal expected utility with respect to the assessment $\mu(\zeta_t)$ be:

$$u_t^* = \max_{\sigma \in \Sigma^{(A)}} u^{(A)}\big(\sigma, \mu(\zeta_t)\big).$$

- ϕ *is asymptotically myopic if:*

$$\lim_{t\to\infty} \Big(u_t^* - u^{(A)}\big(\phi(\zeta_t), \mu(\zeta_t)\big) \Big) = 0.$$

- ϕ *is strongly asymptotically myopic if:*

$$\lim_{t\to\infty} \left(u_t^* - \min_{s \in \mathrm{supp}(\phi(\zeta_t))} u^{(A)}\big(s, \mu(\zeta_t)\big) \right) = 0.$$

Settings where myopia may be justified

Suppose there are two large indistinguishable population of Alices and Bobs.

Settings where myopia may be justified

Suppose there are two large indistinguishable population of Alices and Bobs.

1. Two are randomly drawn to play the game, which is broadcasted.

Settings where myopia may be justified

Suppose there are two large indistinguishable population of Alices and Bobs.

1. Two are randomly drawn to play the game, which is broadcasted.
2. Alices and Bobs are randomly matched, the population game statistics are shared.

Settings where myopia may be justified

Suppose there are two large indistinguishable population of Alices and Bobs.

1. Two are randomly drawn to play the game, which is broadcasted.
2. Alices and Bobs are randomly matched, the population game statistics are shared.
3. Alices and Bobs are randomly matched, and only know their own histories.

Settings where myopia may be justified

Suppose there are two large indistinguishable population of Alices and Bobs.

1. Two are randomly drawn to play the game, which is broadcasted.
2. Alices and Bobs are randomly matched, the population game statistics are shared.
3. Alices and Bobs are randomly matched, and only know their own histories.

Justifications. In (1), players play very infrequently, and so the future is highly discounted.

Settings where myopia may be justified

Suppose there are two large indistinguishable population of Alices and Bobs.

1. Two are randomly drawn to play the game, which is broadcasted.
2. Alices and Bobs are randomly matched, the population game statistics are shared.
3. Alices and Bobs are randomly matched, and only know their own histories.

Justifications. In (1), players play very infrequently, and so the future is highly discounted. In (2) and (3), the players may believe that their actions will have little influence, so that the other population is essentially non-adaptive.

Adaptive assessments

Definition

Alice's assessment rule μ is ***adaptive*** *if for every $\varepsilon > 0$ and t, if Bob did not play a set of pure strategies S' during times t to t', then Alice's assessment of this strategy set is small:*

$$\phi_{t'}^{(A)}(S') < \varepsilon,$$

as long as $t' > T(\varepsilon, t)$ is sufficiently large.

Adaptive assessments

Definition

*Alice's assessment rule μ is **adaptive** if for every $\varepsilon > 0$ and t, if Bob did not play a set of pure strategies S' during times t to t', then Alice's assessment of this strategy set is small:*

$$\phi_{t'}^{(A)}(S') < \varepsilon,$$

as long as $t' > T(\varepsilon, t)$ is sufficiently large.

- That is, if Bob stops playing some strategy, then Alice's assessment that that strategy will be played converges to zero.

Adaptive assessments

Definition

*Alice's assessment rule μ is **adaptive** if for every $\varepsilon > 0$ and t, if Bob did not play a set of pure strategies S' during times t to t', then Alice's assessment of this strategy set is small:*

$$\phi_{t'}^{(A)}(S') < \varepsilon,$$

as long as $t' > T(\varepsilon, t)$ is sufficiently large.

- That is, if Bob stops playing some strategy, then Alice's assessment that that strategy will be played converges to zero.
- The empirical assessment from fictitious play is adaptive.

Adaptive assessments

Definition

Alice's assessment rule μ is ***adaptive*** *if for every $\varepsilon > 0$ and t, if Bob did not play a set of pure strategies S' during times t to t', then Alice's assessment of this strategy set is small:*

$$\phi_{t'}^{(A)}(S') < \varepsilon,$$

as long as $t' > T(\varepsilon, t)$ is sufficiently large.

- That is, if Bob stops playing some strategy, then Alice's assessment that that strategy will be played converges to zero.
- The empirical assessment from fictitious play is adaptive.
- This is a fairly weak condition.

Absorbing states are Nash equilibria

Proposition

Let Alice and Bob be strongly asymptotically myopic players with adaptive assessment rules. If $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots$ is their sequence of plays and $\boldsymbol{s}_t = \boldsymbol{s}_*$ for all $t \geq T$, then $\boldsymbol{s}_*$ is a Nash equilibrium of the game.

Proof: absorbing states are Nash equilibria

Proof by contradiction.

Proof: absorbing states are Nash equilibria

Proof by contradiction.

1. Suppose that $\mathbf{s}_*$ is not a Nash equilibrium. WLOG, Alice has a best response $\hat{s}$:

$$u^{(A)}\left(\hat{s}, s_*^{(B)}\right) > u^{(A)}\left(s_*\right).$$

Proof: absorbing states are Nash equilibria

Proof by contradiction.

1. Suppose that $\mathbf{s}_*$ is not a Nash equilibrium. WLOG, Alice has a best response $\hat{s}$:

$$u^{(A)}\left(\hat{s}, s_*^{(B)}\right) > u^{(A)}\left(s_*\right).$$

2. Because Alice's assessment is adaptive, ϕ_t becomes arbitrarily close to $\delta_{\mathbf{s}_*}$, the delta distribution on $\mathbf{s}_*$. Thus:

$$\lim_{t\to\infty} u^{(A)}(\phi_t) = u^{(A)}(\mathbf{s}_*).$$

Proof: absorbing states are Nash equilibria

Proof by contradiction.

1. Suppose that $\mathbf{s}_*$ is not a Nash equilibrium. WLOG, Alice has a best response $\hat{s}$:

$$u^{(A)}\left(\hat{s}, s_*^{(B)}\right) > u^{(A)}\left(s_*\right).$$

2. Because Alice's assessment is adaptive, ϕ_t becomes arbitrarily close to $\delta_{\mathbf{s}_*}$, the delta distribution on $\mathbf{s}_*$. Thus:

$$\lim_{t \to \infty} u^{(A)}(\phi_t) = u^{(A)}(\mathbf{s}_*).$$

3. Because Alice is strongly asymptotically myopic, given enough time, she plays arbitrarily near-optimal responses to $\phi_t^{(A)}$, which must be nearly as good as $\hat{s}$.

Proof: absorbing states are Nash equilibria

Proof by contradiction.

1. Suppose that $\mathbf{s}_*$ is not a Nash equilibrium. WLOG, Alice has a best response $\hat{s}$:

$$u^{(A)}\left(\hat{s}, s_*^{(B)}\right) > u^{(A)}\left(s_*\right).$$

2. Because Alice's assessment is adaptive, ϕ_t becomes arbitrarily close to $\delta_{\mathbf{s}_*}$, the delta distribution on $\mathbf{s}_*$. Thus:

$$\lim_{t \to \infty} u^{(A)}(\phi_t) = u^{(A)}(\mathbf{s}_*).$$

3. Because Alice is strongly asymptotically myopic, given enough time, she plays arbitrarily near-optimal responses to $\phi_t^{(A)}$, which must be nearly as good as $\hat{s}$.

But then, $\mathbf{s}_t$ does not remain on $\mathbf{s}_*$ for all time. □

Convergence to mixed strategies

Empirical frequencies

Definition

Let $\mathbf{s}_1, \mathbf{s}_2, \ldots$ *be the sequence of plays. Their sequence of* ***empirical frequencies*** $\overline{\sigma}_t \in \Sigma^{(A)} \times \Sigma^{(B)}$, *where* $\overline{\sigma}_t$ *is the empirical distribution over plays by time* t*.*

Empirical frequencies

Definition

Let $\mathbf{s}_1, \mathbf{s}_2, \ldots$ be the sequence of plays. Their sequence of ***empirical frequencies*** *$\overline{\sigma}_t \in \Sigma^{(A)} \times \Sigma^{(B)}$, where $\overline{\sigma}_t$ is the empirical distribution over plays by time t.*

- Note that $\overline{\sigma}_t$ is not necessarily a product distribution.

Empirical frequencies

Definition

Let $\mathbf{s}_1, \mathbf{s}_2, \ldots$ *be the sequence of plays. Their sequence of* ***empirical frequencies*** $\overline{\sigma}_t \in \Sigma^{(A)} \times \Sigma^{(B)}$, *where* $\overline{\sigma}_t$ *is the empirical distribution over plays by time* t.

- Note that $\overline{\sigma}_t$ is not necessarily a product distribution.
- If the marginal frequencies have limits, $\lim_{t\to\infty} \overline{\sigma}_t^{(A)} = \overline{\sigma}_*^{(A)}$ and $\lim_{t\to\infty} \overline{\sigma}_t^{(B)} = \overline{\sigma}_*^{(B)}$ exist,

Empirical frequencies

Definition

Let $\mathbf{s}_1, \mathbf{s}_2, \ldots$ *be the sequence of plays. Their sequence of* ***empirical frequencies*** $\overline{\sigma}_t \in \Sigma^{(A)} \times \Sigma^{(B)}$, *where* $\overline{\sigma}_t$ *is the empirical distribution over plays by time* t.

- Note that $\overline{\sigma}_t$ is not necessarily a product distribution.
- If the marginal frequencies have limits, $\lim_{t\to\infty} \overline{\sigma}_t^{(A)} = \overline{\sigma}_*^{(A)}$ and $\lim_{t\to\infty} \overline{\sigma}_t^{(B)} = \overline{\sigma}_*^{(B)}$ exist, define the product distribution:

$$\sigma_*^{\otimes} = \overline{\sigma}_*^{(A)} \otimes \overline{\sigma}_*^{(B)}.$$

Convergence implies Nash equilibrium?

Question. What sort of learners do Alice and Bob need to be so that we can obtain the guarantee that if $\sigma_*^{\otimes}$ exists, then it is a mixed Nash equilibrium?

Asymptotically empirical assessments

Definition

Alice's assessment rule μ is ***asymptotically empirical*** *if:*

$$\lim_{t\to\infty} \left\| \mu_t - \overline{\sigma}_t^{(B)} \right\| = 0.$$

That is, her assessment converges to Bob's empirical frequencies of play.

Convergence implies mixed Nash equilibrium

Proposition

Let Alice and Bob be strongly asymptotically myopic players with asymptotically empirical assessment rules. If $\overline{\sigma}_*^{\otimes}$ exists, then it is a Nash equilibrium.

Proof.

Analogous to proof for convergence to pure Nash equilibria. □

Non-convergence in behavior

Convergence of the empirical marginal frequencies does not imply convergence of behavior to a Nash equilibrium.

- Often, in fictitious play, players jump from pure strategies to another, typically in cycles of increasing lengths (e.g. rock–paper–scissors).

Non-convergence of rock–paper–scissors

R	|
P	
S	
R	
P	|
S	

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

<table>
<tr><td>R</td><td>|</td></tr>
<tr><td>P</td><td></td></tr>
<tr><td>S</td><td>|</td></tr>
<tr><td>R</td><td></td></tr>
<tr><td>P</td><td>||</td></tr>
<tr><td>S</td><td></td></tr>
</table>

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

<table>
<tr><td>R</td><td>|</td></tr>
<tr><td>P</td><td></td></tr>
<tr><td>S</td><td>||</td></tr>
<tr><td>R</td><td>|</td></tr>
<tr><td>P</td><td>||</td></tr>
<tr><td>S</td><td></td></tr>
</table>

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

<table>
<tr><td>R</td><td>|</td></tr>
<tr><td>P</td><td></td></tr>
<tr><td>S</td><td>|||</td></tr>
<tr><td>R</td><td>||</td></tr>
<tr><td>P</td><td>||</td></tr>
<tr><td>S</td><td></td></tr>
</table>

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

<table>
<tr><td>R</td><td>|</td></tr>
<tr><td>P</td><td>|</td></tr>
<tr><td>S</td><td>|||</td></tr>
<tr><td>R</td><td>|||</td></tr>
<tr><td>P</td><td>||</td></tr>
<tr><td>S</td><td></td></tr>
</table>

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

<table>
<tr><td>R</td><td>|</td></tr>
<tr><td>P</td><td>||</td></tr>
<tr><td>S</td><td>|||</td></tr>
<tr><td>R</td><td>||||</td></tr>
<tr><td>P</td><td>||</td></tr>
<tr><td>S</td><td></td></tr>
</table>

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

R	|
P	|||
S	|||
R	$\cancel{||||}$
P	||
S	

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

R	|
P	||||
S	|||
R	$\cancel{||||}$
P	||
S	|

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

R	𝍷
P	𝍸
S	𝍷𝍷𝍷
R	𝍸
P	𝍷𝍷
S	𝍷𝍷

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

R	$\vert$
P	$\cancel{\vert\vert\vert\vert}\ \vert$
S	$\vert\vert\vert$
R	$\cancel{\vert\vert\vert\vert}$
P	$\vert\vert$
S	$\vert\vert\vert$

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

R	$|$
P	$\cancel{||||}\ ||$
S	$|||$
R	$\cancel{||||}$
P	$||$
S	$||||$

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

R	$\vert$
P	$\cancel{\vert\vert\vert\vert}\ \vert\vert\vert$
S	$\vert\vert\vert$
R	$\cancel{\vert\vert\vert\vert}$
P	$\vert\vert$
S	$\cancel{\vert\vert\vert\vert}$

Table 1: Non-convergence of behavior

Non-convergence of rock–paper–scissors

R	𝍷𝍷
P	𝍸𝍷𝍷𝍷
S	𝍷𝍷𝍷
R	𝍸
P	𝍷𝍷
S	𝍸𝍷

Table 1: Non-convergence of behavior

References

Drew Fudenberg and David M Kreps. Learning mixed equilibria. *Games and economic behavior*, 5(3): 320–367, 1993.