Lecture 13: Social Learning with Finite Samples and a Continuum of Players

Harry PEI
Department of Economics, Northwestern University

Spring Quarter, 2021

Introduction

Last lecture: A sequence of myopic agents observe all their predecessors' actions and a private signal, in order to learn about a persistent state.

- With one rational type and a finite action space, asymptotic efficiency *if and only if* agents' private signals are unbounded.

  Bounded signals: agents may rationally ignore their private signals.

- With one rational type and bounded signals, agents will take the correct action asymptotically *if and only if* their action set is rich enough.
- Multiple rational types can lead to confounded learning.

  Agents' actions depend on their private signals, but the public history is uninformative about the state.

Common feature: Every agent observes all predecessors' actions.

- What if every agent only observes a finite sample?

Finite Sample Learning

Today: Finite sample learning in two scenarios.

- Banerjee and Fudenberg (2004): Learning from actions/payoffs.
- Wolitzky (2018): Learning from outcomes, but cannot observe actions.

Applications: Word-of-mouth communication.

- Conley and Udry (2001,2010): Pineapple farmers in Ghana only know about what a few other farmers are doing.
- Chen, Cai and Fang (2009): Restaurant choices.

Examine models with a continuum of players.

- The system is deterministic at the aggregate level.
- Complement the papers on social learning in networks where the number of agents is countable.

Model (Banerjee and Fudenberg 2004)

- Two payoff-relevant states $\theta \in \{\theta_a, \theta_b\}$.
- Two actions $\{a, b\}$.
- Payoffs $u(\theta_a, a) = u(\theta_b, b) = 1$ and $u(\theta_a, b) = u(\theta_b, a) = 0$.
- Prior belief $\Pr(\theta = \theta_a) = \pi > 1/2$.

- Time $t = 0, 1, 2, \ldots$
- Period 0: A continuum of individuals are born.
- In state θ_a, a fraction $x(\theta_a)$ take action a, others take action b.

  In state θ_b, a fraction $x(\theta_b)$ take action a, others take action b.
- In period t, a fraction $\gamma \in (0, 1)$ of old players are replaced.
- Every new player observes $N \in \mathbb{N}$ old players' actions, and a signal s whose distribution depends on θ and the sample, and takes an action.

Model (Banerjee and Fudenberg 2004)

- Every new player observes $N \in \mathbb{N}$ old players' actions, and a signal s whose distribution depends on θ and the sample, and takes an action.
- **Assumption:** Every new player samples uniformly.

  Suppose a fraction $x \in [0,1]$ of existing players play a, then

  $$\Pr(\text{there are } n \text{ players choosing } a \text{ in the sample}) = \binom{N}{n} x^n (1-x)^{N-n}.$$

- Let $\zeta \in \{0, 1, ..., N\}$ denote the number of action a in a sample.

  Signal distribution: $s \sim f(\cdot|\theta, \zeta) \in \Delta(S)$, with S finite.
- **Assumption:** Players sample independently and their signals are conditionally independent.

  By the LLN (Judd 1985), the fraction of population choosing a conditional on each state evolves deterministically.

Law of Motion

Recall that the initial conditions in period 0 are:

- a fraction $x(\theta_a)$ of players take action a if $\theta = \theta_a$,
- a fraction $x(\theta_b)$ of players take action a if $\theta = \theta_b$.

Let $\hat{x}_t(\zeta, s)$ be the prob with which a player chooses a in period t after observing sample ζ and signal s.

Let $\boldsymbol{x}_t \equiv (x_t(\theta_a), x_t(\theta_b))$, where $x_t(\theta)$ is the fraction of agents choosing a in period t conditional on the state being θ. By definition,

$$x_t(\theta) = (1-\gamma)x_{t-1}(\theta) + \gamma\Big(\sum_{\zeta,s} \Pr(\zeta, s | \theta, x_{t-1}(\theta)) \cdot \hat{x}_t(\zeta, s)\Big).$$

We say that $\boldsymbol{x}_t$ is a steady state if $\boldsymbol{x}_{t+1} = \boldsymbol{x}_t$ in some equilibrium.

The Improvement Principle

The average payoff of surviving players in period t:

$$U(\boldsymbol{x}_t) \equiv \pi x_t(\theta_a) + (1-\pi)(1 - x_t(\theta_b)).$$

- Conditional on $\theta = \theta_a$, a fraction $x_t(\theta_a)$ of them chose a.
- Conditional on $\theta = \theta_b$, a fraction $x_t(\theta_b)$ of them chose a.

Lemma: The Improvement Principle

Fixing π, the initial conditions, and any equilibrium,

- $U(\boldsymbol{x}_t)$ *is nondecreasing in* t.
- *For every* $t \in \mathbb{N}$, $U(\boldsymbol{x}_{t+1}) = U(\boldsymbol{x}_t)$ *if and only if no decision rule strictly improves on the rule "copy the action of the first person in the sample".*

The Improvement Principle: Intuition

Lemma: The Improvement Principle

Fixing π, the initial conditions, and any equilibrium,

1. $U(\boldsymbol{x}_t)$ *is nondecreasing in* t.
2. *For every* $t \in \mathbb{N}$, $U(\boldsymbol{x}_{t+1}) = U(\boldsymbol{x}_t)$ *if and only if no decision rule strictly improves on the rule "copy the action of the first person in the sample".*

Suppose a new player in period $t+1$ uses the following decision rule:

- *copy the action of the first person in their sample*.

He cannot do worse than the average player who survives in period t.

- This relies on uniform unbiased sampling.

His optimal decision rule (i.e., mapping from observed sample and signals to distribution over his actions) must yield a weakly higher expected payoff.

Convergence Theorem with Informative Signals

Convergence Theorem with Informative Signals

Assume that $N \geq 2$ *and* $f(s|\theta_a, \zeta) \neq f(s|\theta_b, \zeta)$ *for every* ζ.

1. *If at least one entry of* $\boldsymbol{x}_t$ *is neither* 0 *nor* 1, *then* $U(\boldsymbol{x}_{t+1}) > U(\boldsymbol{x}_t)$.
2. *If* $\boldsymbol{x}$ *is a steady state, then every entry of* $\boldsymbol{x}$ *must be either* 0 *or* 1.
3. *The system must converge to a steady state.*

Proof of Statement 1: If $U(\boldsymbol{x}_{t+1}) = U(\boldsymbol{x}_t)$, then

- Any new agent in $t+1$ cannot do better than imitating the first person in their sample.
- Therefore, for any sample $\zeta \in \{1, ..., N-1\}$ and any $s \in S$, the new agent is indifferent between a and b.
- This contradicts the presumption that s is informative about θ.

Convergence Theorem with Informative Signals

Convergence Theorem with Informative Signals

Assume that $N \geq 2$ *and* $f(s|\theta_a, \zeta) \neq f(s|\theta_b, \zeta)$ *for every* ζ.

1. *If at least one entry of* $\boldsymbol{x}_t$ *is neither* 0 *nor* 1*, then* $U(\boldsymbol{x}_{t+1}) > U(\boldsymbol{x}_t)$.
2. *If* $\boldsymbol{x}$ *is a steady state, then every entry of* $\boldsymbol{x}$ *must be either* 0 *or* 1.
3. *The system must converge to a steady state.*

Proof of Statement 2:

- In any steady state $\boldsymbol{x}_t$, we have $U(\boldsymbol{x}_{t+1}) = U(\boldsymbol{x}_t)$.
- The conclusion of Statement 2 follows from Statement 1.

Convergence Theorem with Informative Signals

Convergence Theorem with Informative Signals

Assume that $N \geq 2$ *and* $f(s|\theta_a, \zeta) \neq f(s|\theta_b, \zeta)$ *for every* ζ.

1. *If at least one entry of* $\boldsymbol{x}_t$ *is neither* 0 *nor* 1, *then* $U(\boldsymbol{x}_{t+1}) > U(\boldsymbol{x}_t)$.
2. *If* $\boldsymbol{x}$ *is a steady state, then every entry of* $\boldsymbol{x}$ *must be either* 0 *or* 1.
3. *The system must converge to a steady state.*

Proof of Statement 3: The key step is to show that

- If $\boldsymbol{x}_t$ is bounded away from the steady state, then $U(\boldsymbol{x}_{t+1}) - U(\boldsymbol{x}_t)$ is bounded away from 0.

Convergence Theorem without Informative Signals

Convergence Theorem without Informative Signals

Assume that $N \geq 3$.

1. *If at least one entry of $\boldsymbol{x}_t$ is neither 0 nor 1, then $U(\boldsymbol{x}_{t+1}) > U(\boldsymbol{x}_t)$.*
2. *If $\boldsymbol{x}$ is a steady state, then every entry of $\boldsymbol{x}$ must be either 0 or 1.*
3. *The system must converge to a steady state.*

If $U(\boldsymbol{x}_{t+1}) = U(\boldsymbol{x}_t)$, then any new agent in $t+1$ cannot do better than imitating the first person in their sample.

- The new agent is indifferent between a and b when there is one a in their sample, and when there are two as in their sample.
- This implies that $x_t(\theta_a) = x_t(\theta_b) \in (0,1)$, so

$$U(\boldsymbol{x}_t) \leq \max_{x\in[0,1]} \left\{\pi x + (1-\pi)(1-x)\right\} \leq \max\{\pi, 1-\pi\} = \pi.$$

- $U(\boldsymbol{x}_t) < \pi$ when $x_t(\theta_a) = x_t(\theta_b) \in (0,1)$.
- However, an agent's expected payoff is π under his prior.

Efficiency Theorem

Efficiency Theorem

Suppose $N \geq 2$ and for every sample $\zeta \in \{0, \ldots, N\}$, there is positive probability of a signal realization s such that

$$\frac{f(s|\theta_a, \zeta)}{f(s|\theta_b, \zeta)} \cdot \frac{\pi}{1-\pi} < 1.$$

then the system converges to the efficient point $\boldsymbol{x} = (1, 0)$.

Proof: We know that the system must converge to $\boldsymbol{x}$ consists only of 0 and 1.

- It cannot converge to $(0,0)$ or $(0,1)$ since the expected payoff is less than π (the achievable payoff under the prior).
- It cannot converge to $(1,1)$ since there exists $s \in S$ such that

$$\frac{f(s|\theta_a, \zeta)}{f(s|\theta_b, \zeta)} \cdot \frac{\pi}{1-\pi} < 1,$$

after which the new player should choose b after observing s.

Efficiency Theorem

Efficiency Theorem

Suppose $N \geq 2$ *and for every sample* $\zeta \in \{0, \ldots, N\}$, *there is positive probability of a signal realization s such that*

$$\frac{f(s|\theta_a, \zeta)}{f(s|\theta_b, \zeta)} \cdot \frac{\pi}{1-\pi} < 1.$$

then the system converges to the efficient point $\boldsymbol{x} = (1, 0)$.

This theorem requires s to be sufficiently informative.

- When the informativeness of s is low, there can be multiple steady states, some of them are inefficient.
- However, inefficient steady states are never stable,

  i.e., there exists small perturbations s.t. the distribution over actions drifts away from the inefficient steady state.

Inefficient steady states must be unstable

Let U^* be the payoff in an inefficient steady state $\boldsymbol{x}'$.

- Consider the hyperplane defined by the isoprofit curve

$$\pi x'(\theta_a) + (1-\pi)(1-x'(\theta_b)) - U^* = 0.$$

  By definition, this curve crosses $\boldsymbol{x}'$.

- The efficient point is $\boldsymbol{x}^* \equiv (1,0)$.

- Suppose in period 0, the action distribution $\boldsymbol{x}$ is at the side of the red hyperplane containing $\boldsymbol{x}^*$.

  The agent's expected payoff from $\boldsymbol{x}$ is strictly greater than U^*.

  Improvement principle implies that $\{\boldsymbol{x}_t\}_{t\in\mathbb{N}}$ can never converge to $\boldsymbol{x}'$.

What happens when $N = 1$?

When $N = 1$, the steady state $\boldsymbol{x}$ may contain sth other than 0 and 1.

- Suppose the initial value of $\boldsymbol{x}$ is $(1-\varepsilon_1, \varepsilon_2)$.
- For every bounded signal s, $\exists\, \varepsilon > 0$ s.t. when $\varepsilon_1, \varepsilon_2 < \varepsilon$, every player finds it optimal to play the action he observes.
- Therefore, $(1-\varepsilon_1, \varepsilon_2)$ is a steady state for small enough $\varepsilon_1, \varepsilon_2$.

Generic Ineffiency when $N = 1$

When $N = 1$ and $\boldsymbol{x} \neq (1,0)$ in period 0, the system converges to an inefficient steady state.

Intuition: The system gets stuck once it reaches $(1-\varepsilon_1, \varepsilon_2)$.

Model

- Time is continuous $t \in [0, +\infty)$. State $\theta \in \{0, 1\}$, with $\Pr(\theta = 1) = p$.
- Action $a_t \in \{0, 1\}$, outcome $y_t \in \{0, 1\}$.
  $\Pr(y_t = 1 | a_t = 0) = \chi$, $\Pr(y_t = 1 | a_t = 1, \theta = i) = \pi_i$ for $i \in \{0, 1\}$.

- At time 0, a continuum of players whose choices are exogenous.
- Old players die at rate γ and new players arrive at rate γ.
- When a new player arrives, he randomly samples K outcomes of surviving old players, and makes an irreversible choice $a_t \in \{0, 1\}$.
- Player t's payoff is $y_t - ca_t$ where $c \in \mathbb{R}$ is the relative cost of 1.
- **Assumptions:** $\pi_1 - c > \chi > \pi_0 - c$, (optimal action is state dependent)
  $p > p^*$, where $p^*(\pi_1 - c) + (1 - p^*)(\pi_0 - c) = \chi$, (1 is optimal ex ante)

$$\frac{1-p}{p} \cdot \left(\frac{1-\pi_0}{1-\pi_1}\right)^K > \frac{1-p^*}{p^*} \quad \text{(everyone chooses 1 is not an equilibrium)}$$

Aligned Points and Misaligned Points

The population at time t can be described by $\boldsymbol{x}_t \equiv (x_t(0), x_t(1))$, where $x_t(\theta)$ is the fraction of agents choosing action 1 in state θ.

- Prob of good outcome in state θ is $\sigma_\theta(\boldsymbol{x}_t) \equiv x_t(\theta)\pi_\theta + (1 - x_t(\theta))\chi$.
- Observing more $y = 1$ is good news iff $\sigma_1(\boldsymbol{x}_t) \geq \sigma_0(\boldsymbol{x}_t)$,

  or equivalently, $x_t(1)(\pi_1 - \chi) \geq x_t(0)(\pi_0 - \chi)$.

$\boldsymbol{x}_t \equiv (x_t(0), x_t(1)) \in [0,1]^2$ is *aligned* if $x_t(1)(\pi_1 - \chi) \geq x_t(0)(\pi_0 - \chi)$, and is *misaligned* otherwise.

Results

Theorem: Aligned Points are Absorbing

If the initial point $\boldsymbol{x}_0$ is aligned, then $\boldsymbol{x}_t$ is aligned for every $t \in \mathbb{R}_+$.

Suppose $x_0(1)(\pi_1 - \chi) \geq x_0(0)(\pi_0 - \chi)$ but $x_t(1)(\pi_1 - \chi) < x_t(0)(\pi_0 - \chi)$.

- $\exists\, s \in [0,t]$ s.t. $x_s(1)(\pi_1 - \chi) - x_s(0)(\pi_0 - \chi) = 0$ and the derivative of the LHS w.r.t. time is negative.
- Since $x_s(1)(\pi_1 - \chi) - x_s(0)(\pi_0 - \chi) = 0$, observing outcomes is uninformative, so all new agents at s choose ex ante optimal action 1.
- Therefore, $\dot{x}_s(\theta) = \gamma(1 - x_s(\theta))$. This yields

$$\dot{x}_s(0)(\pi_0 - \chi) = \gamma(1 - x_s(0))(\pi_0 - \chi) = \gamma\Big\{\pi_0 - \chi - x_s(1)(\pi_1 - \chi)\Big\}$$

$$< \gamma\Big\{\pi_1 - \chi - x_s(1)(\pi_1 - \chi)\Big\} = \dot{x}_s(1)(\pi_1 - \chi),$$

which leads to a contradiction.

Cost Saving Innovation & Outcome Improving Innovation

Theorem: Aligned Points are Absorbing

If the initial point $\boldsymbol{x}_0$ is aligned, then $\boldsymbol{x}_t$ is aligned for every $t \in \mathbb{R}_+$.

Intuition: Consider two cases.

1. Outcome improving: $\pi_1 > \chi$.

   If $x_s(1)(\pi_1 - \chi) = x_s(0)(\pi_0 - \chi)$, then adoption rate is higher in state 0.

   Increase in adoption rate is higher in state 1, making $\boldsymbol{x}$ more aligned.

2. Cost saving: $\pi_1 < \chi$.

   If $x_s(1)(\pi_1 - \chi) = x_s(0)(\pi_0 - \chi)$, then adoption rate is higher in state 1.

   Increase in adoption rate is higher in state 0, making $\boldsymbol{x}$ more aligned.

Another interesting observation:

- Efficient point $(0, 1)$ is aligned in the outcome improving case.
- Efficient point $(0, 1)$ is misaligned in the cost saving case.

Result: Long Term Inefficiency

Theorem: Long Term Inefficiency

If $\pi_1 < \chi$ and $\boldsymbol{x}_0$ is aligned, then the steady state is bounded away from efficiency no matter how large K is.

Why? If $\boldsymbol{x}_0$ is aligned, then $\boldsymbol{x}_t$ is aligned.

- The efficient point $(0, 1)$ is not aligned in the cost saving case.

Result: Long Term Efficiency

Theorem: Long Term Efficiency

If $\pi_1 > \chi$,

then for every $\boldsymbol{x}_0$ *and* $\varepsilon > 0$,

there exist $\overline{K} \in \mathbb{N}$ *and* $T \in \mathbb{R}_+$ *such that when* $K > \overline{K}$ *and* $t > T$,

every equilibrium path is within an ε *neighborhood of* $(0,1)$.