Stochastic Evolutionary Implementation on (or of?) a Network

Abstract

0.1 Introduction

I found the paper by Sandholm *Pigouvian Pricing and Stochastic Evolutionary Implementation (JET 2007)* interesting and I like that motivation. Also, the logic upon which the results are obtained is very close (in fact identical) to the one I used in *Learning Correlated Equilibria in a Population Game (MSS 2001).*

Please find below:

an executive summary of my setup (which includes Sandholm's as a special case);

an executive summary of Sandholm's motivation and results;

the couple of ideas I talked to you over lunch.

1 Ianni (2001)

Consider a population game $\Gamma(\Omega, G, \mu)$ where Ω is a set of N players, $G(A_i, \Pi_i)_{i=1,\ldots,n}$ is an underlying n-player normal form game and μ is a commonly known probability distribution over n-tuples of players that

a) partitions Ω in $\{\Omega_i\}_{i=1,..,i,.,n}$ and

b) such that $\mu(\omega) = \sum_{-\omega} \mu(\omega, -\omega) > 0$.

We can represent Ω in terms of a graph $\mathcal{G}(\Omega, E)$ where E is the set of edges connecting any n-tuple of players to which μ assigns positive probability.

(I guess μ defines a network in the standard sense)

a_ω is any strategy adopted by player ω, leading to expected payoff

$$\sum_{-\omega:\mu(\omega,-\omega)>0} \pi_\omega(a_\omega, a_{-\omega})\mu(-\omega \mid \dot{\omega}) \equiv E_{-\omega}\pi_\omega(a_\omega, a_{-\omega})$$

1.1 Equilibria

$E(\Gamma) \subseteq \times_\omega a_\omega$ is the set of equilibria of Γ. $a \in E(\Gamma)$ iff. for all ω

$$a_\omega \in Arg \max_{a_\omega} E_{-\omega}\pi_\omega(a_\omega, a_{-\omega})$$

$SE(\Gamma) \subseteq E(\Gamma)$ is the set of strict equilibria of Γ iff for all ω

$$a_\omega = Arg \max_{a_\omega} E_{-\omega}\pi_\omega(a_\omega, a_{-\omega})$$

$N(G) \subseteq \times_i A_i$ is the set of Nash equilibria of G.

$\Psi(G) \subseteq \{$set of all probability distributions over $\times_i A_i\}$ is the set of Correlated Equilibria of G.

1.2 Dynamics

Time runs discretely. At $t = 0$, $\times_\omega a_\omega = a^0$, which can only be changed if and whenever updating opportunity arise. In between any two time periods, one updating opportunity is allocated to one player, chosen with probability $\mu(\omega)$. At any $t > 0$ an n-tuple of players is drawn from $\mu(\omega, -\omega)$; they observe a_ω^{t-1} and they play one shot of G.

Action choices are made on the basis of a rule. We consider the following rules:

MBR: ω chooses a Myopic Best Reply at time t if $a_\omega^t \in Arg \max_{a_\omega} E_{-\omega}\pi_\omega(a_\omega, a_{-\omega}^{t-1})$

NBR: ω chooses a Noisy Myopic Best Reply at time t if

$$\Pr[a_\omega^t = a_j] = \frac{\exp[\sigma E_{-\omega}\pi_\omega(a_j, a_{-\omega}^{t-1})]}{\sum_j \exp[\sigma E_{-\omega}\pi_\omega(a_j, a_{-\omega}^{t-1})]}$$

with $j \in A_i$ and $\omega \in \Omega_i$

1.3 Results

Proposition 1 (Mailath, Samuleson, Shaked 1997) *Every equilibria of Γ induces, via μ, a probability distribution over the Cartesian product of the action spaces that constitutes a correlated equilibrium of G :*

$$\textit{If } a \in E(\Gamma), \textit{ then } \sum_{(\omega,-\omega)=a} \mu(\omega, -\omega) \in \Psi(G).$$

For any $\psi \in \Psi(G)$ there exists an a and a μ such that $\sum_{(\omega,-\omega)=a} \mu(\omega, -\omega) = \psi$.

Proposition 2 (Ianni 2001) *Suppose G is a potential game.*

Let $\Pi_\mu(a) = E_{\omega,-\omega}\pi_\omega(a_\omega, a_{-\omega})$. Then

1. $\Pi_\mu(a)$ is non-decreasing along any MBR path

2. $\Pi_\mu(a)$ is locally maximized at any $a \in \theta(\Gamma)$ and, if Γ has no ties, the only maximizers are to be found in $\theta(\Gamma)$

3. Under NBR the unique limit distribution is given by

$$P_\sigma(a) = \frac{\exp[\sigma\Pi_\mu(a)]}{\sum_a \exp[\sigma\Pi_\mu(a)]}$$

4. If G admits $e = a_1, a_2, ..., a_n$ as a Pareto dominant Nash equilibrium, then

$$\lim_{\sigma\to\infty} P_\sigma(a_e) = 1 \text{ where } a_e \text{ is such that, for any } i,\ a_\omega = a_i \text{ for all } \omega \in \Omega_i$$

Remark 3 *A special case of the above model is when matching is uniform as in the following Definition*

Definition 4 *Matching is uniform if*

$$\mu(\omega_1, \omega_2, ..., \omega_n) = \times_{i=1,...,n}\#\Omega_i \text{ for all } (\omega_1, \omega_2, ..., \omega_n)$$

In this case, for $\omega \in \Omega_i$, $\mu_U(\omega) = (\#\Omega_i)^{-1}$ and $\sum_{\omega\in\Omega_i:a_\omega=j} \mu_U(\omega) \equiv x_j^i$(i.e. it is the fraction of players choosing action j in population i) and $x^i = \{x_j^i\}_{j=1,...,\#A_i}$ is the vector of action adoption in population i and $x = \{x^i\}_{i=1...,n}$ is the distribution of actions in the game.

Remark 5 *Sandholm's model deals with a symmetric game. This could easily be derived in this framework by disposing of assumption a) on μ, but I prefer in general to keep the game potentially asymmetric.*

2 Sandholm (2007)

Matching is uniform and payoffs are specified as:

$$\pi_j^\omega = \pi_j(x) + \varepsilon_j^\omega$$

where ω is the player, j is the action and ε is the 'type realization '. Hence payoffs depend on the action chosen, on the distribution of actions in the population and a disturbance that depends on the player and on the action chosen (check this). The underlying game is defined as $G(\Pi, \varepsilon)$

Problem 6 *A social planner would like to ensure efficient behaviour on the part of players. Efficiency depends on the realization of ε, and this is only known to the players. Can the planner design a pricing mechanism (anonymous, in the sense that it can only be made conditional on action choices) that ensures the long run selection of the efficient outcome, regardless of the realization of types?*

Proposition 7 (Sandholm 2007) *In the above set up, with the game accounting for an outside option, the price mechanism that ensures the long run selection of the efficient outcome (regardless of the realizations of types) exists. It involves charging each player ω choosing an active action j, with a (negative) price, equal, in magnitude to the net benefit to ω's opponents when (s)he chooses j instead of the outside option (the "inactive" action).*

Logic of the proof. Let $\phi : \varepsilon \to \phi(\varepsilon)$ be a social choice correspondence mapping each realization of types into the socially optimal outcome given those types. We say that a price mechanism $p_j(x)$ stochastically implements ϕ if for each ε, the game with payoffs $G(\Pi - p, \varepsilon)$ is a potential game, with potential function $\Pi_{\mu_U}(x)$. The results then follows by using the result 3. and 4. in Proposition (2), once we notice that by construction, the payoff functions are additively separable in types. The logic of the proof is identical to the one used in Proposition (2) and it involves showing that for a potential game the limit distribution is unique and it is reversible, i.e. it is such that

$$\frac{P_\sigma(a^1)}{P_\sigma(a)} = \frac{P_\sigma(a^1 \mid a)}{P_\sigma(a \mid a^1)}$$

for any two configurations a and a^1 for which the transition under MBR occurs with positive probability. ■

3 Conjectures

Conjecture 8 *I believe that extending the result of Sandholm to a locally interactive setting can be done. This would lead to something like Stochastic Evolutionary Implementation on a Network. We do know that in a population game, equilibria (efficient and inefficient ones) under non uniform matching can be different from replica versions of the equilibria of the underlying game. It might be interesting to ask whether, given an exogenous network of interactions, the same result on the implementability of efficient outcome via price mechanisms still holds.*

Conjecture 9 *A more ambitious project instead is to model a game in which the network is chosen endogenously by players (maybe because players need to choose to form links and these come at a cost). Can a social planner affect incentives in such a way to induce players to form the type of network that allows society to get to an efficient outcome? This would lead to something like Stochastic Evolutionary Implementation of a Network.*

4 Action

The setup of Sandholm involves N players playing an N player game, where the payoff functions include a disturbance term.

Consider -instead- a population game where N players are matched in $n-$tuples to play a one shot of an n-player game.

For example, consider a bilateral game, take $n = 2$ and assume μ partitions Ω into Ω_1 and Ω_2 with action spaces $\{j \in A_1\}$ and $\{k \in A_2\}$.(this assumption is not necessary). The normal form game is denoted by $\Pi = \{\pi^1_{jk}, \pi^2_{jk}\}$. Matching is defined by μ_U (described above). Expected payoffs from one shot of interaction are:

$$E_{-\omega}\pi_\omega(a_\omega, a_{-\omega}) \equiv \begin{cases} \pi^1_j(\omega) = \sum\limits_k (\pi^1_{jk} + \varepsilon(\omega))x^2_k & \text{for } \omega \in \Omega_1, a_\omega = j \\ \pi^2_k(\omega) = \sum\limits_j (\pi^2_{jk} + \varepsilon(\omega))x^1_j & \text{for } \omega \in \Omega_2, a_\omega = k \end{cases}$$

Question: Can we design a price mechanism that guarantees the evolutionary implementation of the (a) socially efficient outcome in this population setting (same logic as Sandholm)?

Conjecture 10 *Answer: yes, if the game is symmetric, disturbances are* fixed effects*, i.e. they depend on the player, but not on the action chosen, and the price mechanism is such that the underlying game is transformed into a game with identical interests.*

Remarks:

1. this can always be done for an underlying 2by2 game

2. this can always be done for a 2-player partnership game (or balanced)

3. the price mechanism depends on the current configuration of play only in the sense that it charges a fixed amount conditional on the action played (these obviously may change)

4. the result works for any exogeneously given μ

5. if disturbances depended on the action, as well as on the player, things would not work in a population game, as the planner should know more. On the other hand disturbances that depend only on the player do affect social efficiency (as in Sandholm), but do not affect incentives in the underlying game (unlike in Sandholm).

All of the above is for a fixed exogeneous network. Now suppose agents cannot change actions, but can choose with whom to play. Suppose they do so in order to maximize expected payoff, but this time over μ, for a fixed a_ω:

$$Arg \max_{-\omega} E_{-\omega}\pi_\omega(a_\omega, a_{-\omega})$$

Furthermore, suppose that matching can only take place between say ω_1 and ω_2,currently playing $a_{\omega_1} = j$ and $a_{\omega_2} = k$ if $\pi^1_{jk} \geq \pi^1_{jl}$ for all l in A_1 and $\pi^2_{jk} \geq \pi^2_{jl}$ for all l in A_2 (this is sloppy but it means that links -and hence play- needs to be bilaterally agreed upon). Notice that 1. if the game is symmetric this condition implies $\pi_{kj} \geq \pi_{lj}$ and $\pi_{jk} \geq \pi_{jl}$ for all l; 2. if the game is a common interest game this implies $\pi_{jk} \geq \pi_{jl}$ for all l.

Example 11 (Exogeneous network) *Coordination game*

$$\begin{array}{cc} 5,5 & 0,3 \\ 3,0 & 3,3 \end{array}$$

Price mechanism that guarantees evolutionary implementation of (T, L) *involves charging* 3 *to whoever plays* B *or* R*, so that incentives, net of charges are:*

$$\begin{array}{cc} 5,5 & 0,0 \\ 0,0 & 0,0 \end{array}$$

NBR will then select (T, L)*, which is socially efficient given any payoff specific disturbance.*

Example 12 (Endogeneous network) *Given the above game in its original form (i.e. with no taxes or subsidies), consider the following story:*

Time runs discretely. At $t = 0$, $\times_\omega a_\omega = a^0$ *and* $\mu = \mu^0$.

Actions can only be changed if and whenever an action-updating opportunity arises. In between any two time periods, one updating opportunity is allocated to one player, chosen with probability $\mu(\omega)$*. This player updated her/his action to an MBR to the current configuration of play. Then a pair of players is drawn from* μ *and plays one shot of* G.

Links, i.e. pairs $(\omega_1, \omega_2) : \mu(\omega_1, \omega_2) > 0$*, can be changed if and whenever a link-updating opportunity arises. In between any two time periods, one updating opportunity is allocated to one player, chosen with probability* $\mu(\omega)$*. The configuration of actions is fixed at* a^{t-1}*, and player* ω_1 *can choose to update her or his current set of opponents,* $\{\omega_2 : \mu(\omega_1, \omega_2) > 0\}$ *to myopically maximize his or her payoff.*

What are the steady states of this process? Suppose there are only two players:

For example consider a bilateral game ($n = 2$) being played by randomly matched pairs of players, under uniform global matching.

1. What would efficiency imply in this setup? 2. What would the price mechanism look like?

Socially efficient is a profile of actions that maximizes the sum of the payoffs over all players.

$$\begin{aligned} \sum_\omega \pi(a_\omega, a_{-\omega}) &= \sum_l \sum_k \pi_{lk} x_l x_k \text{ symm} \\ &= \sum_j \sum_l \sum_k \pi^i_{lk} x^i_l x^j_k \text{ asymm} \end{aligned}$$

But there may be idiosyncratic types... to be added.

To guarantee convergence, the price mechanism should be such that for any player switching say from action j to action l, the average payoff in the population is also increasing.

A switch from j to l under MBR occurs iff:

$$\begin{aligned} E_{-\omega}\pi_\omega(a_\omega &= l, a_{-\omega}^{t-1}) - E_{-\omega}\pi_\omega(a_\omega = j, a_{-\omega}^{t-1}) > 0 \\ &\Leftrightarrow \sum_k (\pi_{lk}^i - \pi_{jk}^i) x_k^j > 0 \end{aligned}$$

The change in the population average payoff is:

$$\sum_k (\pi_{lk}^i - \pi_{jk}^i) x_k^j + \sum_k (\pi_{lk}^j - \pi_{jk}^j) x_k^j$$

This is trivially positive if:

1. The underlying game is (symmetric and) doubly symmetric ($\pi_{ij} = \pi_{ji}$)
2. The underlying game is a common interest game
3. If the game is balanced ($\pi_{ij} = \widetilde{\pi}_{ij} + \pi_j = \widetilde{\pi}_{ji} + \pi_i = \pi_{ji}$), where the tilda-ed payoffs are symmetric, then one price mechanism that would do the job involves charging any player playing action j a price of $-\pi_j$. This would work also for a generic network, not only global uniform matching.

Would this guarantee that NBR selects the efficient outcome for any value of the idosyncratic type? Think so (it should be the same as in Sandholm)

Since matching is uniform, the efficient outcome would be one of the Nash equilibria (pure or mixed). Hence the selection by NBR would take place over the convex hull of Nash equilibria of the underlying game. For some games the set of correlated equilbria expands this set. (one may look at an example)

If matching is not uniform, the mechanism works as well. Here everything depends on μ.

Find a game to start with.

I have started working on the following game (but not had the time to work it out):

$$\begin{array}{ccc} 6,6 & 2,7 & 0,0 \\ 7,2 & 0,0 & 0,0 \\ 0,0 & 0,0 & 0,0 \end{array}$$

The efficient correlated equilibrium, leading to a payoff of $(21/4, 21/4)$ is

$$\begin{array}{ccc} 1/2 & 1/4 & 0 \\ 1/4 & 0 & 0 \\ 0 & 0 & 0 \end{array}$$

Question: Is this the appropriate game to look at? Need to find a situation where by constructing a local pricing mechanism, one can get efficiency.

Remark:

Clearly if one could condition a pricing mechanism on action and identity, the answer is trivial.

Sandholm shows that for the games he looks at, one can design a pricing mechanism conditional on action and state that does the job.