Adaptive and Sophisticated Learning

Adaptive and Sophisticated Learning

The idea of best reply dynamics goes back all the way to Cournot's study of duopoly and forms the foundation of Walrasian equilibrium in economy and is created by the classical *Tatonnement* learning process.

The underlying learning processes can be categorized into successively stronger versions:

- **Best-Reply Dynamics:** However, it's also known that this dynamics lead to non-convergent, cyclic behavior. In this model, an outsider with no information about the utilities (payoffs) of the agents could eventually predict the behavior of the agents more accurately than they themselves.

- **Fictitious-Play Dynamics:** The agents choose strategies that are best reply to predictions that the probability distributions of the competitors' play at the next round is based on the empirical distribution of the past plays. Even this dynamics lead to (if there is no zero-sum restriction) cycles of exponentially increasing lengths.

- **Stationary Bayesian Learning Dynamics:** The agents choose strategies as functions from the information set (empirical distribution of the past plays) without relying on any intermediate prediction. The distribution over the strategies changes as the empirical distribution changes. (Reactive Learning: involves no model building.)

  The dynamics may converge—but to a (mixed) strategy profile that is not necessarily the perfect (Nash) equilibrium.

Set-up

Player n plays a sequence of plays: $\{x_n(t)\}$. Each $x_n(t)$ is a pure strategy and is chosen by the rules of player n's learning algorithm. We are interested in two properties that may be satisfied by $\{x_n(t)\}$: it is *approximately* best-reply dynamics, then it is consistent with *adaptive learning*; it is *approximately* fictitious-play dynamics, then it is consistent with *sophisticated learning.*

Definition $\{x_n(t)\}$ **is consistent with adaptive learning.** *Player n eventually chooses only strategies that are* nearly *best replies to some probability distribution over his rivals joint strategy profiles, where near* <u>*zero*</u> *probabilities are assigned to strategies that have not been played for sufficiently long time.*

Definition $\{x_n(t)\}$ **is consistent with sophisticated learning.** *Player n eventually chooses only* nearly *best replies to his probabilistic forecast of rivals' joint strategy profiles, where the support of probability may include not only past plays but also strategies that the rivals may choose if they themselves were adaptive or sophisticated learners.*

We will look at the effect of these algorithms on *finite player games, with compact strategies and continuous pay-off functions*.

Note that these assumptions are consistent with the usual model of exchange economy with infinitely divisible goods. Note that in this model, serially undominated set is a singleton and thus the Walrasian equilibrium. One of the main results that we will see is that in any process, consistent with adaptive learning, play tends towards the serially undominated set and hence, in an exchange economy, adaptive learning would lead to equilibrium.

Formulation

Definition *Noncooperative game*

$$\Gamma = (N, (S_n; n \in N), \pi).$$

$$\begin{aligned} N &= \textit{Finite Player Set} \\ S_n &= \textit{Player } n\textit{'s strategy} \\ & \quad \textit{Compact Subset of some Normed Space} \\ \pi &= \textit{Pay-off Function} \\ & \quad \textit{Assumed Continuous.} \end{aligned}$$

$$S = \times_{n \in N} S_n \quad x \in S \Rightarrow x = (x_n, x_{-n}).$$

x_{-n} is the strategy choice of n's rivals.

$$\begin{aligned} \pi &: S \to \mathbb{R}^{|N|} = \textit{Pay-off Function, Continuous} \\ \pi_n &: S \to \mathbb{R} \\ &: (x_n, x_{-n}) \mapsto \pi_n(x). \end{aligned}$$

Let T be a set. Then $\Delta(T)$ = Set of all probability distributions over T.

$\Delta(S_n)$ = *Mixed strategies on S_n.* $\Delta_{-n} = \times_{j \neq n} \Delta(T_j)$ = *Mixed strategies of n's rivals.*