Evolution & Learning in Games

Econ 243B

Jean-Paul Carvalho

Lecture 6.

Properties of Deterministic Dynamics

Properties of Deterministic Dynamics

- Let us now examine specific properties of different revision protocols and mean dynamics.
- The properties studied in this lecture will help us later to characterize the global convergence and local stability properties of these dynamics.

Revision Protocols: Informational Burdens

(CI) **Complete information:** ρ_{ij} depends on $\pi_i, \ldots, \pi_n$ and on $x_1, \ldots x_n$.

(U) **Uncoupled:** ρ_{ij} depends on $x_1, \ldots x_n$ and π_i, but not on π_{-i}.

(U′) **Uncoupled′:** ρ_{ij} depends on $\pi_i, \ldots, \pi_n$, but not on x.

(CU) **Completely Uncoupled:** ρ_{ij} depends only on π_i.

(CU′) **Completely Uncoupled′:** ρ_{ij} depends only on π_j.

Imitative Protocols: Informational Burdens

Consider:

- *Imitation driven by dissatisfaction*: $\rho_{ij}(\pi, x) = (K - \pi_i)x_j$.
- *Imitation of success*: $\rho_{ij}(\pi, x) = x_j(\pi_j - K)$.
- *Pairwise proportional imitation*: $\rho_{ij}(\pi, x) = x_j[\pi_j - \pi_i]_+$.

These protocols are in classes CU, CU' and U' respectively.

Direct Protocols: Informational Burdens

Consider:

- *Logit Choice*: $\rho_{ij}(\pi, x) = \frac{exp(\eta^{-1}\pi_j)}{\sum_{k \in S} exp(\eta^{-1}\pi_k)}$.
- *Comparison to the Average Payoff*: $\rho_{ij}(\pi, x) = \left[\pi_j - \sum_{k \in S} x_k \pi_k\right]_+$.

These protocols are in classes U' and CI, respectively.

Properties of Aggregate Behavior

- Let us now introduce two desirable(?) properties of evolutionary dynamics. We shall then identify classes of revision protocols which generate mean dynamics that exhibit these properties.

- Firstly, consider:

Positive Correlation [PC]: $V_F(x) \neq \mathbf{0}$ implies that $\left(V_F\right)' F(x) > 0$.

- This can be conceived as follows: whenever a population is not at rest, the *covariance* between its strategies' growth rates and payoffs is positive.

Interpreting PC

- We shall show that:

$$Cov\big(V_F(x), F(x)\big) = \frac{1}{n}\big(V_F(x)\big)'F(x).$$

  Then the suggested interpretation follows.

- First, view the strategy set $S = \{1, \ldots, n\}$ as a probability space endowed with the uniform probability measure: $\mathbb{P}(\{i\}) = \frac{1}{n}$ for all $i \in S$.

- Then for each $x \in X$, we can think of the elements of the vectors $V_F(x)$ and $F(x)$ as different realizations of a random variable.

Interpreting PC

- Note that:

$$\mathbb{E}\big(V_F(x)\big) \equiv \sum_{k \in S} \mathbb{P}(\{k\}) V_{F,k}(x) = \sum_{k \in S} \tfrac{1}{n} V_{F,k}(x) = 0$$

because $V_{F,k}(x) \in TX$, i.e. changes in the population shares of each strategy must sum to zero.

Interpreting PC

Therefore:

$$\begin{aligned} Cov\big(V_F(x), F(x)\big) &= \mathbb{E}\big(V_F(x)F(x)\big) - \mathbb{E}\big(V_F(x)\big)\mathbb{E}\big(F(x)\big) \\ &= \sum_{k \in S} \mathbb{P}(\{k\}) V_{F,k}(x) F_k(x) - 0 \\ &= \frac{1}{n}\big(V_F(x)\big)' F(x). \end{aligned}$$

Hence if $\big(V_F(x)\big)' F(x) > 0$, then the covariance between strategies' growth rates and payoffs is positive.

Properties of Aggregate Behavior

- Secondly, consider:

Nash Stationarity [NS]: $V_F(x) = \mathbf{0}$ if and only if $x \in NE(F)$.

- This requires that the set of Nash equilibria equals the set of rest points of the dynamic.
- Dynamics with this property (partially) justify the concept of Nash equilibrium without strong equilibrium knowledge assumptions.

Properties of Aggregate Behavior

Nash Stationarity implies that:

(i) Every Nash equilibrium of F is a rest point of V_F: if there are no profitable unilateral deviations, then there is no change in the state.

(ii) Every rest point of V_F is a Nash equilibrium of F: profitable unilateral deviations are exploited.

Interpreting Rest Points

- Note: At a rest point of the (deterministic) mean dynamic, the underlying stochastic process is not necessarily at rest.
  - There may be some inflow from and outflow to each strategy, leaving the state x unchanged,
  - For the mean dynamic, it is only necessary that the *expected* inflow equals the *expected* outflow for each strategy.
- Hence we can think of a rest point of the mean dynamic as a *balance point* of the underlying stochastic process.

PC & Rest Points

Proposition 6.1. If V_F satisfies PC, then $x \in NE(F)$ implies that $V_F(x) = \mathbf{0}$.

Example

Consider the two-strategy coordination game:

$$F(x) = \begin{pmatrix} F_1(x) \\ F_2(x) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 2x_2 \end{pmatrix}.$$

The replicator dynamic for this game is:

$$V(x) = \begin{pmatrix} V_1(x) \\ V_2(x) \end{pmatrix} = \begin{pmatrix} x_1\big(F_1(x) - \overline{F}(x)\big) \\ x_2\big(F_2(x) - \overline{F}(x)\big) \end{pmatrix} =$$
$$\begin{pmatrix} x_1\big[x_1 - \big((x_1)^2 + 2(x_2)^2\big)\big] \\ x_2\big[2x_2 - \big((x_1)^2 + 2(x_2)^2\big)\big] \end{pmatrix}.$$

By inspection, $V(x) = \mathbf{0}$ if and only if $x \in \{(1,0), (0,1), (\frac{2}{3}, \frac{1}{3})\} = NE(F)$. Therefore, the replicator dynamic exhibits Nash stationarity in this game.

Example

The replicator dynamic in this game also exhibits Positive Correlation:

But what is the general behavior (i.e. in all games) of this and other dynamics?

Styles of Decision Making

- We shall now define classes of revision protocols that correspond to different styles of decision making.
- This allows us to systematically analyze how our specific choice of departure from "full rationality" affects the properties of evolutionary dynamics.

Families of Evolutionary Dynamics

Families of Evolutionary Dynamics and their Properties

Family	Leading Example(s)	C	$<$ CI	PC	NS
Imitation	Replicator	yes	yes	yes	no
Excess Payoff	BNN	yes	no	yes	yes
Pairwise Comparison	Smith	yes	yes	yes	yes
Best response	Best response	no	yes	yes*	yes*
Perturbed best response	Logit	yes	yes	no	no

C denotes Lipschitz continuity.

* Best response dynamics satisfy appropriately modified versions of PC and NS.

Imitative Dynamics

Imitative dynamics are based on revision protocols of the form:

$$\rho_{ij}(\pi, x) = x_j r_{ij}(\pi, x),$$

where r_{ij} is a conditional imitation rate.

These revision protocols generate a mean dynamic of the form:

$$\begin{aligned}\dot{x}_i &= \sum_{j \in S} x_j \rho_{ji}\big(F(x), x\big) - x_i \sum_{j \in S} \rho_{ij}\big(F(x), x\big) \\ &= \sum_{j \in S} x_j x_i r_{ji}\big(F(x), x\big) - x_i \sum_{j \in S} x_j r_{ij}\big(F(x), x\big) \\ &= x_i \sum_{j \in S} x_j \big[r_{ji}\big(F(x), x\big) - r_{ij}\big(F(x), x\big)\big]\end{aligned} \tag{1}$$

Imitative Dynamics

Definition. Suppose that the conditional imitation rates are Lipschitz continuous and that net conditional imitation rates are *monotone*, that is:

$$\pi_j \geq \pi_i \Leftrightarrow r_{kj}(\pi, x) - r_{jk}(\pi, x) \geq r_{ki}(\pi, x) - r_{ik}(\pi, x)$$

for all $i, j, k \in S$. Then the mean dynamic (2) is called an **imitative dynamic**.

Examples

Let us now consider two imitative revision protocols that do not generate the replicator dynamic as their mean dynamic.

Both are based on:

Imitation of Success with Repeated Sampling. When an agent's alarm clock rings he chooses an opponent at random. If the opponent is playing strategy j, he imitates him with *copying weight* $w(\pi_j)$. If he does not imitate the opponent, he draws a new opponent at random and repeats the procedure, stopping only when imitation occurs:

$$\rho_{ij}(\pi, x) = \frac{x_j w(\pi_j)}{\sum_{k \in S} x_k w(\pi_k)}.$$

Examples

This revision protocol yields the mean dynamic:

$$\dot{x}_i = \frac{x_i w\big(F_i(x)\big)}{\sum_{k \in S} x_k w\big(F_k(x)\big)} - x_i.$$

Examples

(a) *The Maynard-Smith Replicator Dynamic*.

Let the copying weights equal payoffs, i.e. $w(\pi_j) = \pi_j$. (This requires that payoffs are non-negative and average payoffs are positive.) Then the mean dynamic is:

$$\dot{x}_i = \frac{x_i F_i(x)}{\bar{F}(x)} - x_i,$$

which is known as the **Maynard-Smith Replicator Dynamic**.

Examples

(b) *The Imitative Logit Dynamic*.

Let the copying weights equal $w(\pi_j) = exp(\eta^{-1}\pi_j)$. Then the mean dynamic is:

$$\dot{x}_i = \frac{x_i exp\big(\eta^{-1}F_i(x)\big)}{\sum_{k\in S} x_k exp\big(\eta^{-1}F_k(x)\big)} - x_i,$$

which is known as the **Imitative Logit (or *i*-logit) Dynamic**.

Properties of Imitative Dynamics

- All imitative dynamics satisfy ***extinction***: if a strategy is unused, its growth rate is zero.

- Since imitative dynamics are Lipschitz continuous, they also exhibit ***uniqueness*** and ***forward and backward invariance***:

  Proposition 6.2. For every initial condition $\xi \in X$, an imitative dynamic admits a unique solution trajectory $\mathcal{T}_{(-\infty,\infty)} = \{x : (-\infty, \infty) \to X |\, x \text{ is continuous}\}$.

Properties of Imitative Dynamics

In addition, imitative dynamics exhibit ***support invariance***; the support of x_t is independent of t:

Theorem 6.1 If $\{x_t\}$ is a solution trajectory of an imitative dynamic, then the sign of component $(x_t)_i$ is independent of $t \in (-\infty, \infty)$. That is:

– if $(x_T)_i = 0$ for some T, then it equals zero for all t,

– if $(x_T)_i > 0$ for some T, then it is positive for all t.

- Extinction (i.e. if $x_i = 0$, then $V_i(x) = 0$) along with Lipschitz continuity of the dynamics implies that the speed of motion toward or away from the boundary of X must decline exponentially as the boundary is approached.

Properties of Imitative Dynamics

Monotonicity

As we have shown, imitative dynamics can be written in the following form:

$\dot{x}_i = V_i(x) = x_i G_i(x)$, where
$G_i(x) = \sum_{k \in S} x_k \big[r_{ki}\big(F(x), x\big) - r_{ik}\big(F(x), x\big) \big]$.

If strategy i is in use, then $G_i(x) = V_i(x)/x_i$ is the percentage growth rate of the number of agents using strategy i.

Properties of Imitative Dynamics

It follows from the imitation monotonicity condition in the definition of an imitative dynamic that all imitative dynamics exhibit ***monotone percentage growth rates***:

$G_i(x) \geq G_j(x)$ if and only if $F_i(x) \geq F_j(x)$.

Theorem 6.2. All imitative dynamics satisfy ***Positive Correlation***.

Properties of Imitative Dynamics

Rest Points and Restricted Equilibria

- Since all imitative dynamics satisfy PC, all Nash equilibria of F are rest points of an imitative dynamic (by Proposition 6.1).
- However, support invariance means that *non-Nash rest points can exist*: all pure states in X are rest points of an imitative dynamic, but they are not necessarily NE.

Properties of Imitative Dynamics

- The set of rest points can be characterized as follows. Recall that:

$$NE(F) = \{x \in X : x_i > 0 \implies F_i(x) = \max_{j \in S} F_j(x)\}.$$

- The set of rest points are the set of **restricted equilibria**:

$$RE(F) = \{x \in X : x_i > 0 \implies F_i(x) = \max_{j \in S : x_j > 0} F_j(x)\}.$$

These are the Nash equilibria of a restricted version of F in which only strategies in the support of x can be played.

Direct revision Protocols & Dynamics

- Let us turn to dynamics generated by direct revision protocols.
- Because strategies are directly selected, good strategies will be discovered and chosen even if they are unused.
- Hence there is some chance of the dynamic generated by a direct protocol satisfying Nash stationarity.
- We shall focus on the best response and logit dynamics (excess payoff and pairwise comparison dynamics are not studied here).

Best Response Dynamics

- The **best response dynamic** is generated by agents always switching to their current best response.
- This dynamic has some peculiar features because the best response correspondence is <u>discontinuous</u> (small changes in the state x can produce sharp changes in responses) and <u>multivalued</u> (there could be multiple best responses to a state).
- Differential inclusions—set-valued differential equations—can be used to analyze the best response dynamic.

The Best Response Protocol

- Suppose when an agent receives an opportunity to revise his strategy, he chooses a (myopic) best response to the current population state.
- The behavior is myopic because a best response to the current state may become superseded as the state changes.
- With enough inertia, however, it is likely to remain a best response for some time.
- Formally, the switching rate under best response protocol is independent of an agent's current strategy and does not depend directly on the state, $\rho_{ij}(\pi, x) = \sigma_j(\pi)$.
- The conditional switching rates also sum to one, $\sum_{j \in S} \sigma_j(\pi) = 1$, so that $\sigma(\pi) \in X$ is a mixed strategy.

The Best Response Dynamic

- Focussing on the single-population case again, this generates a mean dynamic of the form:

$$\begin{aligned} \dot{x}_i &= \sum_{j \in S} x_j \rho_{ji}\big(F(x), x\big) - x_i \sum_{j \in S} \rho_{ij}\big(F(x), x\big) \\ &= \sum_{j \in S} x_j \sigma_i\big(F(x)\big) - x_i \sum_{j \in S} \sigma_j\big(F(x)\big) \\ &= \sigma_i\big(F(x)\big) \sum_{j \in S} x_j - x_i \\ &= \sigma_i\big(F(x)\big) - x_i. \end{aligned} \tag{2}$$

- This can be expressed as:

$$\dot{x} = \sigma\big(F(x)\big) - x. \tag{3}$$

The Best Response Protocol

- What is $\sigma(\pi)$?

- The best response protocol is given by the multivalued map:

$$\sigma(\pi) = M(\pi) \equiv \arg\max_{y \in X} y'\pi,$$

where $M(\pi)$ is the set of mixed strategies that place mass only on pure strategies optimal under payoff vector π.

The Best Response Dynamic

- Substituting into (3), we have the following differential inclusion:

$$\dot{x} \in M\big(F(x)\big) - x. \tag{4}$$

Definition. A Carathéodory solution to the differential inclusion $\dot{x} \in V(x)$ is a Lipschitz continuous trajectory $\{x_t\}_{t \geq 0}$ that satisfies $\dot{x}_t \in V(x_t)$ at all but a measure zero set of times in $[0, \infty)$.

Theorem 6.3. Fix a continuous population game F. Then for each $\xi \in X$, there exists a trajectory $\{x_t\}_{t \geq 0}$ with $x_0 = \xi$ that is a Carathéodory solution to the differential inclusion (4).

Solution Trajectories

- As we shall see, while solutions to the best response dynamic exist, the best response protocol is discontinuous so the solutions *need not be unique*; multiple solution trajectories can emanate from a single initial condition.
- Yet they can be quite simple.
- Let $\{x_t\}$ be a solution to (4) and suppose that the best response to state x_t is the pure strategy $i \in S$ at all times $t \in [0, T]$.
- Then during this interval, evolution is described by the affine differential equation:

$$\dot{x} = e_i - x.$$

Solution Trajectories

- Hence the state x moves directly toward vertex e_i of the set X, proceeding more slowly as the vertex is approached.
- This means that the state x_t lies on the segment containing x_0 and e_i throughout the interval $[0, T]$.
- Solving $\dot{x} = e_i - x$ we get the following explicit formula for x_t:

$$x_t = (1 - \exp^{-t})e_i + \exp^{-t} x_0 \qquad \text{for all } t \in [0, T].$$

Examples

(a) Standard Rock-Paper Scissors

- One can construct a figure which appears to indicate that every solution trajectory converges to the unique Nash equilibrium $x^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.
- To formally prove this, we claim that along every solution trajectory $\{x_t\}$, whenever the best response is unique, we have:

$$\frac{d}{dt}\left(\max_{k \in S} F_k(x_t)\right) = -\max_{k \in S} F_k(x_t). \tag{5}$$

Examples

- To establish the claim, let x_t be a state in which there is a unique optimal strategy, say Paper.
- At this state $\dot{x}_t = e_P - x_t$. Since $F_P(x) = w(x_R - x_S)$:

$$\begin{aligned} \frac{d}{dt}F_P(x_t) &= \nabla F_P(x_t)'\dot{x}_t \\ &= (w \quad 0 \quad -w)(e_P - x_t) \\ &= -w(x_R - x_S) \\ &= -F_P(x_t). \end{aligned} \tag{6}$$

Examples

- Because any solution trajectory passes through states with multiple best responses at most a countable number of times (see Figure), (5) can be integrated with respect to time.

- This yields:

$$\max_{k \in S} F_k(x_t) = e^{-t} \max_{k \in S} F_k(x_0). \tag{7}$$

- In standard RPS, payoffs to each strategy are non-negative and equal zero only at the Nash equilibrium x^*.

- Then (7) implies that the maximal payoff across strategies $k \in S$ falls over time converging to zero as t approaches infinity; this occurs as x_t converges to the Nash equilibrium x^*.

Examples

(b) Two-Strategy Coordination

- Let the strategy set be $S = \{U, D\}$ and the payoff matrix be:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

- The game $F(x) = Ax$ has three Nash equilibria, two pure (e_U and e_D) and a mixed equilibrium $(x_U^*, x_D^*) = (\frac{2}{3}, \frac{1}{3})$.

Examples

- Denote the state by $\chi = x_D$, so that $\chi^* = \frac{1}{3}$.
- Then the best response-dynamic can be expressed as:

Examples

- From every initial condition except χ^*, there is a unique solution trajectory of the dynamic that converges to a pure Nash equilibrium:

$$\chi_0 < \chi^* \implies \chi_t = e^{-t}\chi_0. \tag{8}$$

$$\chi_0 > \chi^* \implies \chi_t = 1 - e^{-t}(1 - \chi_0). \tag{9}$$

Examples

- There are many solution trajectories from χ^*:
  - a stationary trajectory,
  - one that proceeds to $\chi = 0$ according to (8),
  - another that proceeds to $\chi = 1$ according to (9).
- Notice that solutions (8) and (9) quickly leave the vicinity of χ^*.
- In contrast, for Lipschitz continuous dynamics:
  1. solutions from all initial conditions are unique,
  2. solutions that start near a stationary point move very slowly near that point.

Properties of the Best Response Dynamic

- Let us now establish the analogue of PC and NS for the differential inclusion (4):

Theorem 6.4. The best response dynamic satisfies:

$$z'F(x) = max_{j \in S} \widehat{F}_j(x) \qquad \text{for all } z \in V_F(x). \tag{10}$$

$$\mathbf{0} \in V_F(x) \qquad \text{if and only if } x \in NE(F). \tag{11}$$

Properties of the Best Response Dynamic

- If condition (10) holds, then the correspondence $x \mapsto V_F(x)'F(x)$ is single-valued, always equaling the maximal excess payoff among strategies.
- This value is non-negative and equals zero if and only if all players are playing a best response, i.e. if x is a Nash equilibrium.
- If condition (11) holds, then the differential inclusion $\dot{x} \in V_F(x)$ has a stationary solution at every Nash equilibrium, but at no other states.
- As we have seen, this does not rule out the existence of additional solution trajectories that leave Nash equilibria.

Perturbed Best Response Dynamics

- Let us now introduce perturbations to the revision protocol:
  - random utility,
  - experimentation,
  - errors in perception or implementation (trembles).
- This leads to revision protocols that are a smooth function of payoffs.
- Such perturbed best response functions (or quantal response functions) are used in experimental economics to model experimental data.

Perturbed Best Response Protocols

- The leading example of a perturbed best response protocol is **logit choice**:

$$\tilde{M}_i(\pi) = \frac{\exp(\eta^{-1}\pi_i)}{\sum_{j \in S} \exp(\eta^{-1}\pi_j)}.$$

- Recall that as $\eta \to 0$ this converges to the (unperturbed) best response protocol.

- However, unlike the best response protocol, the logit protocol ρ is continuous, differentiable and single-valued.