

研究ノート

ランダムな局所的対戦構造をもつ適応プロセス

がゲームの均衡選択に与える影響

宮下 春樹・福住 多一

1 イントロダクション

本研究ノートは、進化ゲーム理論におけるプレイヤーどうしの対戦構造の違いが均衡選択におよぼす影響について考察する。通例、各プレイヤーは集団内のあらゆるプレイヤーと、ランダムに対戦すると想定される。本研究ノートは、各プレイヤーの対戦相手が、ランダムに選ばれた集団の一部である場合の分析を進める。先行研究として Hilbe (2011) はプレイヤーどうしの対戦が、無限母集団からランダムに選ばれた有限集団という局所的構造をもつ連続時間のレプリケーター動学を定式化している。本研究ノートは、次のような適応プロセスが均衡選択に与える影響を検討する。每期、有限母集団から有限人数のランダムな対戦相手が選ばれる局所対戦構造を想定する。その対戦相手の戦略分布に対して、プレイヤーは最適反応により適応する。我々は、この定式化した適応

動学を鹿狩りゲームと制裁が加わった囚人のジレンマゲームに適用し、その場合のプレイヤー集団の安定状態が示す特徴を数値例で検討する。

2 モデル

我々は Hilbe (2010) と異なり、有限の母集団から一部の対戦プレイヤーがピックアップされ、最適反応によって母集団の状態が変化していく適応プロセスを定式化する。 $M = \{1, 2, \dots, m\}$ をプレイヤー母集団とする。このプレイヤー集団から毎期 $t = 0, 1, 2, \dots$ に、その部分集合 $N \subseteq M$ がランダムにピックアップされる。ここで $N = \{1, 2, \dots, n\}$, $(n \leq m)$ である。 N 内でプレイヤーはランダムに 2 人の組を形成する。その各組は次のゲーム $G(S, u)$ をプレイする。 $S = \{s_1, s_2, \dots, s_k\}$ を各プレイヤーの純粋戦略の集合とする。 $u(s_i, s_j) \in \mathbb{R}$ を戦略 s_i をとるプレイヤーが s_j をとるプレイヤーとのプレイから得る利得とする。

2.1 適応プロセス

各 t 期のプレイヤー母集団の状態をベクトル $m(t) = (m_1(t), m_2(t), \dots, m_k(t))$ で表す。ただし、 $\sum_{i=1}^k m_i = m$ であり、各 m_i は戦略 s_i にコミットメントしているプレイヤーの人口である。また各 t 期において M から N がランダムにピックアップされるとし、 N における t 期の状態は次のベクトル $n(t) := (n_1(t), \dots, n_k(t))$

で表現される。ただし n_i は N の中で s_i にコミットメントしているプレイヤーの人口である。各期の N においてランダムにプレイヤーが1人ピックアップされ、実現している $n(t)$ に対して $t+1$ 期の期首に最適反応することにより、その戦略を変更すると仮定する。

2.2 状態の推移確率

推移確率 $\text{Pr}(m(t+1)|m(t))$ を求める。 $m(t)$ から $n(t) = (n_1(t), n_2(t), \dots, n_k(t))$ をピックアップする確率は超幾何分布 (hyper geometric distribution) に従う。つまり、この確率は

$$\frac{m_1 C n_1 \cdots m_k C n_k}{M C N} \quad (1)$$

である。戦略 s_i をとっているプレイヤーがピックアップされる確率は $\frac{n_i(t)}{n(t)}$ である。このプレイヤーが戦略 s_j をとる場合の期待利得は、

$$Eu_i(s_j|n(t)) = \frac{n_i(t)-1}{n(t)-1} u_i(s_i, s_i) + \sum_{j \neq i} \frac{n_j(t)}{n(t)-1} u_i(s_i, s_j).$$

$t+1$ 期に戦略 i のプレイヤーが最適反応に基づいてコミットメントする戦略は $s_j^* \in \arg \max_{s_j \in S} Eu_i(s_j|n(t))$ である。ここで $\text{BR}_i(n(t)) := \arg \max_{s_j \in S} Eu_i(s_j|n(t))$ と書く。各 s_j^* をとる確率は $\frac{1}{|\text{BR}(n(t))|}$ とする。 s_i にコミットメントしているプレイヤーがピックアップされ、 $\text{BR}(n(t)) = \{s_j^*\}$ であるとき、 $m_i(t)$ は一人減り、 $m_j(t)$ は一人増える。各戦略 s_i にコミットメントしているプレイヤーをそれぞれピックアップする

確率は、 $\frac{n_i(t)}{n(t)}$ であるから、 $m(t+1) = (m_1(t), \dots, m_i(t)-1, \dots, m_j(t)+1, \dots, m_k(t))$

である場合、

$$\text{pr}(m(t+1)|m(t)) = \sum_{i=1}^k \frac{m_1 C n_1 \cdots m_k C n_k}{M C N} \times \frac{n_i(t)}{n(t)} \times \frac{1}{|\text{BR}(n(t))|} \quad (2)$$

である。

3 具体的なゲームへの適用

本章では我々の定式化した適応プロセスでパラメーター値を特定化したものを、鹿狩りゲームと制裁戦略を加えた囚人のジレンマゲームに適用し、母集団の状態の安定性を分析する。

3.1 鹿狩りゲーム

プレイヤーの純粋戦略の集合を $S = \{S, H\}$ とし、式 (3) の利得行列 Π_1 をもつ鹿狩りゲームを考える：

$$\Pi_1 = \begin{array}{cc} & \begin{array}{cc} S & H \end{array} \\ \begin{array}{c} S \\ H \end{array} & \begin{pmatrix} 5 & 0 \\ 4 & 2 \end{pmatrix} \end{array} \quad (3)$$

$m = 10$ とする. このゲームをプレイするプレイヤーの対戦人数が $n = 3, 5$ の場合における状態推移図を図 1, 図 2 にそれぞれ示す. 図 1, 図 2 内の各状態から次の状態へ正の確率 $\text{pr}(m(t+1)|m(t))$ で推移する場合, そのパスを矢印で示し, 矢印がない状態への推移確率は 0 である. 図 1, 図 2 において, 全プレイヤーが S をとる状態 $(10,0)$ と全プレイヤーが H をとる状態 $(0,10)$ に注目しよう. $n = 3$ のとき, 状態 $(10,0)$ から $(0,10)$ へ到達するには, 集団状態の突然変異が 1 回だけ生じ, $(10,0)$ から $(9,1)$ へと置き換えられればよい. 一方, $(0,10)$ から $(10,0)$ へ到達するのに必要な突然変異数は 2 回である. $n = 5$ の場合には, $(10,0)$ から $(0,10)$ へ到達するための突然変異数は 2 回であり, $(0,10)$ から $(10,0)$ へ到達するには 3 回必要である. つまり戦略 S はピックアップされる $N \subseteq M$ の人数が少ないとき, 少数の突然変異によって H に置き換わるが, 人数が増えるにつれて置き換わりにくくなる. この結果は, 最適反応動学を用いて集団全体とのマッチングを想定したときの長期確率安定状態が H となるよく知られた結果 (Kandori, Mailath and Rob (1993), Young (1993)) と類似している.

3.2 制裁戦略を加えた囚人のジレンマゲーム

Weibull and and Salomonsson (2006) に従い, プレイヤーが非協力者を制裁する戦略をもつ囚人のジレンマゲームを導入する. 各プレイヤーの戦略は協力する (CN), 協力しない (D), そして制裁をする (CP) である. CP は CN との

全プレイヤーがSをとる状態

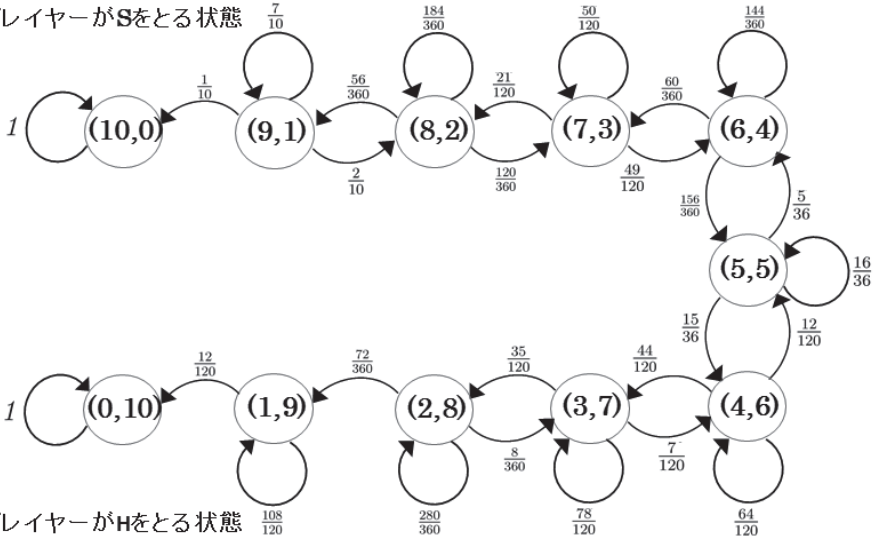


図 1: $n = 3$ のときの状態推移図

全プレイヤーがsをとる状態

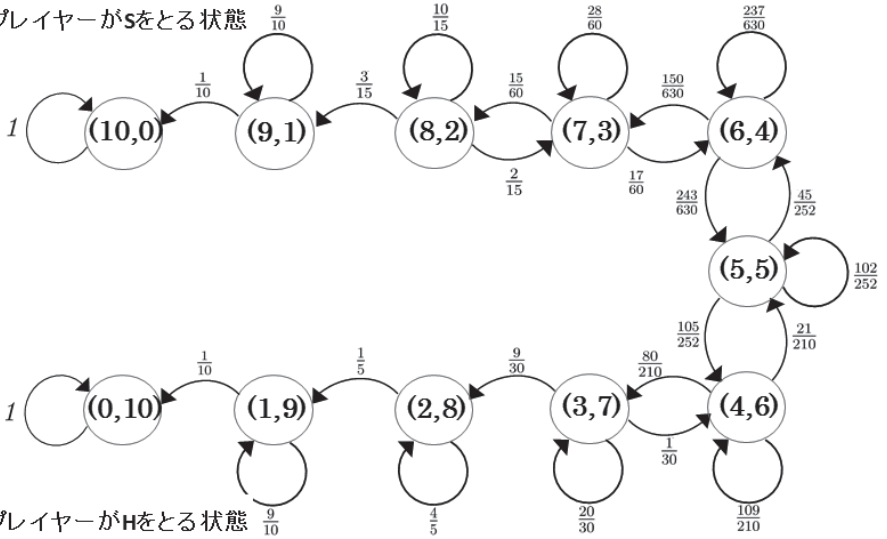


図 2: $n = 5$ のときの状態推移図

対戦ではCNをとり、Dとの対戦においては $c > 0$ のコストを負担することで相手の利得を $d > 0$ だけ減少させる。CNとDの対戦において、CNを選択するプレイヤーは $a < 1$ の利得を得る。この対戦において、Dを選択するプレイヤーは $b > 3$ の利得を得る。CP対Dの対戦においてCPをとるプレイヤーの利得は $a - c < 1$ であり、Dをとるプレイヤーは $b - d < 2$ の利得を得る。このゲームの利得行列は Π_2 で与えられる。

$$\Pi_2 = \begin{matrix} & \begin{matrix} CN & CP & D \end{matrix} \\ \begin{matrix} CN \\ CP \\ D \end{matrix} & \begin{pmatrix} 2 & 2 & a \\ 2 & 2 & a - c \\ b & b - d & 1 \end{pmatrix} \end{matrix} \quad (4)$$

以下の分析では $a = \frac{1}{2}, b = 3$ とし、制裁コストが低く($c = \frac{1}{4}$)、被制裁コストが高い場合($d = \frac{14}{5}$)に限定する。

$m = 8, n = 3$ とする。すると第 t 期における集団の状態は $(8,0,0), (7,1,0), \dots, (0,0,8)$ であり、合計45個存在する。式(2)を用いることにより、 45×45 推移確率行列 $P = [p_{ij}]_{i,j=1, \dots, 45}$ を得る。各成分 p_{ij} は式(2)によって求められる。この推移確率行列を表1に示してある。

表1より、すべてのプレイヤーがCPをとる状態 $(0,8,0)$ は、確率 $\frac{1}{2}$ で $(1,7,0)$ に推移する。また $(1,7,0)$ は確率 $\frac{3}{8}$ で $(2,6,0)$ へ推移する。このような推移確率行

列に従っていくことで、 $(0,8,0)$ は 10 ステップで $(0,0,8)$ へ到達する。すなわち、集団の状態は微小な確率で CN をとるプレイヤーが増加し、CN を選択するプレイヤーの人数の多い状態へと推移する。CN をとるプレイヤーが増加するのに伴い、各プレイヤーには D を選択するインセンティブが働く。すると D をとるプレイヤーの人数の多い状態へと到達する。つまり我々の定式化した適応プロセスにおいては制裁コストが低く、被制裁コストが高い場合であっても、長期的には全員が D を選択する状態で集団が落ち着くのである。

4 ディスカッション

Hilbe (2011) は、無限集団内で局所的にゲームがプレイされるとき連続時間のレプリケーター動学を導入している。この集団から $n \geq 2$ 人のプレイヤーをピックアップする。プレイヤーは、この集団内でランダムにゲームをプレイする。このときのレプリケーター動学を局所的レプリケーター動学 (local replicator dynamics) と呼ぶ。II₂ ゲームの混合戦略空間を $\Delta = \{(x_{CN}, x_{CP}, x_D) \in \mathbb{R}_+^3 \mid x_{CN} + x_{CP} + x_D = 1\}$ と書き、第 i 成分が 1 であり、残りが 0 であるベクトルを $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ と書く。

定理 1. (Hilbe, 2011) ゲームの利得行列を $A \in \mathbb{R}^{k \times k}$ とし、 A の転置行列を A^t

と書く。局所的レプリケーター動学は

$$\dot{x}_i = x_i \cdot [e_i \cdot \tilde{A}x - x \cdot \tilde{A}x] \tag{5}$$

によって与えられる。ただし x_i の上付き文字のドットは時間微分 $\frac{dx_i}{dt}$ を表し、

$$\tilde{A} = A - \frac{A+A^t}{n}$$

である。

$n = 3$ とする。 π_2 ゲームの Δ 内における局所的レプリケーター動学の軌道を位相図として図3に表す。*1 図中の黒丸 (●) はその点が安定的な定常状態であることを表し、白丸 (○) は不安定であることを表す。 $p \in [0, 1]$ を CN と CP をとるプレイヤーが混在する状態において CN をとるプレイヤーの割合とする。辺 CN-CP 上に CN と CP が混在する安定的な均衡の集合 $X = \{pe_{CN} + (1-p)e_{CP} \mid p \leq \frac{d-(b-2)+\frac{1}{n}(a+b-c-d-4)}{d-\frac{1}{n}(c+d)}\}$ が存在する。

先に定式化した式 (2) の適応プロセスにおいては、 Π_2 ゲームでは CP をとるプレイヤーのみからなる均衡は安定性をもたず、集団の全プレイヤーが D を選択する状態へと到達する。局所的レプリケーター動学は、模倣学習を通しての戦略分布の変化を表現し、CN をとるプレイヤーと CP をとるプレイヤーのみからなる状態においては、協力行動が模倣されやすく、集団の状態は D へ向かいにくい。これが局所的レプリケーター動学と我々の適応プロセスとの間に動学的

*1 オープンソフトウェア Dynamo : <http://www.ssc.wisc.edu/whs/dynamo/> を用いて描写している。図3の各軌道は $n \rightarrow \infty$ のとき、Weibull and Salomonsson (2006) および Sethi and Somanathan (1996) で示されている軌道に近づく。

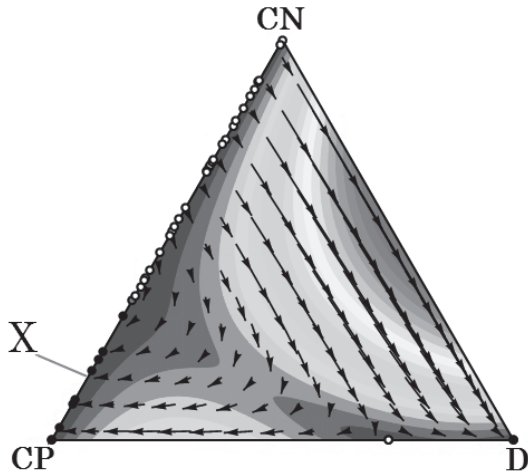


図 3: Π_2 ゲームの位相図

な差が生じる理由である．局所的な動学プロセスは，レプリケーター動学を用いるのか，あるいは最適反応動学を用いるのかによって異なる帰結が生じるのである．今後，我々は離散時間型の局所的レプリケーター動学の定式化を試みる．そこで CP と CN のみからなる集団状態が長期確率安定性をもつかどうかは興味深い問題となるであろう．

参考文献

- [1] Hilbe, C. “Local replicator dynamics: A simple link between deterministic and stochastic models of evolutionary game theory.” *Bulletin of Mathematical Biology*, 73: 2068-2087, 2011.

- [2] Kandori, M, Mailath, G J and Rob, R. “Learning mutation, and long-run equilibria in games.” *Econometrica*, 61: 29-56, 1993.

- [3] Sandholm, B, Dokumaci, E and Franchetti, F. Dynamo: Diagrams for evolutionary game dynamics : <http://www.ssc.wisc.edu/whs/dynamo/> 2016/10/29.

- [4] Sethi, R and Somanathan, E. “The evolution of social norms in common property resource use.” *American Economic Review*, 86(4): 766-788, 1996.

- [5] Weibull, J W and Salomonsson, M. “Natural selection and social preference.” *Journal of Theoretical Biology*, 239: 79-92, 2006.

- [6] Young , P. “The evolution of conventions.” *Econometrica*, 61(1): 57-84, 1993.

(2.5,1)	(2.4,3)	(2.3,2)	(2.2,4)	(2,1,5)	(2,0,8)	(1,7,0)	(1,6,1)	(1,5,2)	(1,4,3)	(1,3,4)	(1,2,3)	(1,1,6)	(1,0,7)	(0,9,0)	(0,7,1)	(0,6,2)	(0,5,3)	(0,4,4)	(0,3,5)	(0,2,6)	(0,1,7)	(0,0,8)
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1/66	3/66	0	0	0	0	1/6	5/66	1/64	0	0	0	0	0	0	0	0	0	0	0	0	0	
1/8	9/112	3/28	0	0	0	0	0	5/84	1/28	1/14	0	0	0	0	0	0	0	0	0	0	0	
0	0	5/112	5/28	0	0	0	0	0	0	1/66	5/42	0	0	0	0	0	0	0	0	0	0	
0	0	0	1/66	16/66	0	0	0	0	0	0	1/68	5/28	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	3/6	1/6	5/66	1/64	0	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	5/84	1/28	0	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	1/28	1/14	0	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	5/42	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	1/66	1/68	5/28	0	1/16	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	1/4	1/16	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5/112	0	0	0	0	0	0	
23/66	65/66	0	0	0	0	1/16	38/112	1/28	0	0	0	0	0	0	0	5/68	1/68	0	0	0	0	
2/68	13/56	0	0	0	0	0	5/56	5/66	5/64	1/7	0	0	0	0	0	1/68	1/68	1/68	0	0	0	
0	0	13/56	2/7	0	0	0	0	5/66	3/14	1/4	0	0	0	0	0	0	1/68	1/68	0	0	0	
0	0	0	17/112	0	0	0	0	0	3/14	5/112	5/28	0	0	0	0	0	0	0	0	0	0	
0	0	0	95/336	11/68	0	0	0	0	0	1/66	4/7	1/8	0	0	0	0	0	0	0	0	0	
0	0	0	0	11/66	0	0	0	0	0	0	0	7/8	0	0	0	0	0	0	0	5/68	1/8	
0	0	0	0	0	0	1/2	3/16	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	
0	0	0	0	0	0	1/16	3/16	15/68	0	0	0	0	0	1/16	5/16	5/68	0	0	0	0	0	
0	0	0	0	0	0	0	5/66	15/66	5/66	0	0	0	0	0	5/68	5/66	5/68	5/68	0	0	0	
0	0	0	0	0	0	0	0	15/66	5/64	9/28	0	0	0	0	0	5/66	3/28	3/28	0	0	0	
0	0	0	0	0	0	0	0	0	0	45/112	3/16	0	0	0	0	0	0	0	1/4	0	0	
0	0	0	0	0	0	0	0	0	0	3/8	1/4	0	0	0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0	0	0	0	0	1/4	0	0	0	0	0	0	1/66	5/28	1/8	
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5/74	5/8	1	