

確率進化モデルと均衡選択、行動伝播 (尾山ゼミ：ゼミ論文)

山岸 敦*

2015/2/18[†]

概要

4つの確率進化モデル (Kandori, Mailath, and Rob, 1993 (以下 KMR と略記する); Ellison, 1993; Williams, 2014; Okada and Tercieux, 2012) をシミュレートするプログラムを python を用いて構築しその挙動や差異を検証する。はじめに、それぞれのモデルを均衡選択の理論と捉えた時のその性能の違いを考察する。特に、Young のゲーム (Young, 1993) と呼ばれるゲームにおいては KMR(1993) のみ以外の 3 つのモデルと異なりパレート効率的でない均衡を選択するという興味深い性質が確認できた。続いて、Ellison(1993) のモデルを「円形ネットワーク上でのゲーム」と捉え、KMR(1993) を「complete network 上でのゲーム」と捉えて両者の挙動の違いを考察する。この観点から見ると、両者はネットワーク上での行動の伝播を分析するモデル (Morris, 2000; Oyama and Takahashi, 2015) らにランダムショックを加えたモデル、と考えることができる。Oyama and Takahashi(2015) に倣い”bilingual game”を用いて分析を行った。その結果、円形ネットワークと完全ネットワークの行動伝播の性能についての追加的な知見が得られた。続いて wheel graph を用いて帝国の言語選択モデルを構築してシミュレートした結果、どのようなときに被支配者の言語が帝国の共通語となるかについての興味深い示唆が得られた。

*東京大学経済学部 3 年

[†]preliminary

1 はじめに

通常のゲーム理論のモデルでは、プレーヤーが常に最適反応を取ることが前提をされている。確率進化モデルはそこにシステムティックな形で「ランダム要素」を追加し、さらに何らかの動学的な過程を想定することで、これによりそれぞれのナッシュ均衡に関しより詳細な情報を得ることが可能にしている。しかし、どのような形でランダム要素を追加するか、どのようなゲームの構造を想定するかはモデルごとに異なり、それがモデルの挙動や導かれる結論に大きく影響する可能性もある。本論文では、4つの確率進化モデル (Kandori, Mailath, and Rob, 1993 (以下 KMR と略記する) ; Ellison, 1993; Williams, 2014; Okada and Tercieux, 2012) を比較する。そのため、分析に入る前に簡単に4つのモデルの概略を紹介する。続いて、均衡選択の理論として4つのモデルが選ぶ均衡の差異を検討する。均衡選択とは、複数のナッシュ均衡のうちでどれがもっともらしいかを議論することである (グレーヴァ, 2013)。つづいて、KMR(1993) と Ellison(1993) をネットワーク上の行動の伝播のモデルにランダムネスを追加したモデルだと解釈し、その挙動を”bilingual game”を用いて分析する。

この論文は以下のように構成されている。第2、3、4、5章はそれぞれ KMR モデル、Ellison モデル、Williams の Stochastic Fictitious Play モデル、Okada and Tercieux の Logit dynamics モデルの簡単な解説にあてられている。第6章においては、Young のゲーム (Young, 1993) を用いて均衡選択の理論として4つのモデルを比較する。第7章では、bilingual game を用いて KMR モデルと Ellison モデルの設定での行動の伝播について議論する。第8章で結論と今後の展望を示す。

2 KMR モデル

表 1: coordination game

| 1/2 | C | D |
|-----|------|------|
| C | 4, 4 | 0, 3 |
| D | 3, 0 | 2, 2 |

一般的な議論は複雑になるので、ここでは10人の原始集団内で、ハイリターンだが協力が得られないと食料を得られない戦略 C か、ローリターンだが安定的に食料を得ることができる戦略 D をとろうとしていると考える。C 戦略をとる人間の数を X_t と書けば、今自分が C を取っていて自分が D に行

動変更した際の利得は

$$3\frac{X_t - 1}{9} + 2(1 - \frac{(X_t - 1)}{9}) \quad (1)$$

であり、他の場合も同様に利得を計算できる。「自分が行動変更する場合、対戦相手は一人ランダムに選択され、それとの対戦により利得が定まる」と考えたときの期待利得になっていることに注意する。これが KMR で想定するゲームの構造である。C と D とで期待利得の高い方に行動を切り替えとする。同点の場合は、どちらかをランダムに選ぶとしておけばよい。さらに、ここが KMR のポイントであるが、 $\epsilon > 0$ の確率で行動変更する際に「突然変異」を起こし、期待利得計算など関係なくランダムに行動を変更すると想定する。

各期において各人に行動変更の機会をランダムに与えられると考え、「集団内で C 戦略をとる者の数 X_t 」を state とみなせば、11 個の state について遷移行列を書くことが可能になる。この定常分布はマルコフ連鎖の議論から一意に存在することが示され、それは「ある state にどれだけの割合とどまっているか」についての情報と解釈できる。さらに、 ϵ を 0 に収束されると、定常分布はある均衡に常にとどまることを示すようになる。これを KMR が選ぶ均衡と考える。表 1 の coordination game の場合、(D, D) 均衡が選択され、これはリスク支配の概念が選ぶ均衡と一致している。

KMR モデルには、収束が遅いという特徴がある。例として、 $\epsilon = 0.01$ で初期状態ではみな C を選んでいるとしよう。このとき、一人が行動を D に変更してもすぐに C に戻ってしまう可能性が高く、3～4 人がほぼ同時に行動を変更しないとイケない。そのため、これが起こる確率は低く、収束が遅くなるのである。

3 Ellison モデル

Ellison モデルは基本的に KMR モデルと同構造であるが、一点重要な違いがある。

Ellison モデルでは図 1 のように円形にプレーヤーが並んでいることが想定されていて、KMR では集団全体から対戦相手が選ばれたのに対し、Ellison モデルでは両隣のうちどちらかとランダムに対戦すると考えられているのである。表 1 の coordination game に関しては、KMR モデルと Ellison モデルの選択する均衡は一致することが知られている (Ellison, 1993)。

しかしながら、収束の早さに大きな違いがある。KMR モデルと同様に、10 人集団で表 1 のゲームをプレイし、 $\epsilon = 0.01$ であり、初期状態ではみな C を取る状況を考えよう。ここで、誰か一人が突然変異を起こして行動を D にしたとする。その D を取る人が選ばれる前に、両隣が選ばれれば彼らは行動

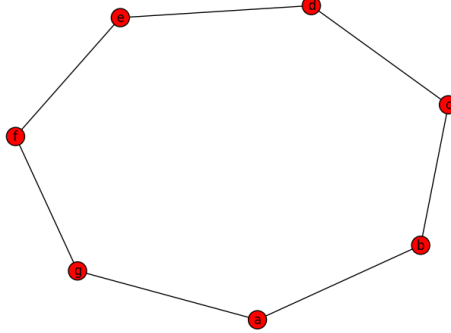


図 1: Ellison モデルの円形ネットワーク

を D に変更する（D の方が期待利得が高い）。このように、Ellison モデルには一度の突然変異が伝播しやすいメカニズムが備わっているのである。

4 Stochastic Fictitious Play モデル

この章では、Williams(2014) による Stochastic Fictitious Play モデルを簡単に解説する。まずは、"Stochastic" のつかない Fictitious Play を考えよう。これは、同じゲームを同じ対戦相手と繰り返し行なった時に相手の戦略をどう学習するかに関する理論である。特に、Fictitious Play においては「前回までに相手が各戦略をそれぞれ何回ずつ選んだか、その頻度を計算し、その頻度にしたがって相手が戦略を選択すると予想を立てたうえで最適反応を取る」（川越, 2011）という想定が置かれている。例として、表 1 のゲームを今までに 100 回行い、相手が C を 70 回、D を 30 回取ったとしよう。このとき、相手が次に C を取る確率 P_C は $\frac{70}{100}$ であると予想する。この下で各戦略の期待利得を計算すると

$$\begin{aligned} C : 4 \frac{70}{100} &= \frac{14}{5} \\ D : 3 \frac{70}{100} + 2 \frac{30}{100} &= \frac{27}{10} \end{aligned}$$

となり、C を選択する。

ただし、この定式化では現在の情報に対するウェイトは次第に低下していく。Williams(2014) は、現在の情報のウェイトが一定となる定式化をしており、これは「相手の真の戦略分布が途中で変わる可能性がある」と想定していることを示す。原文の記号を用いて、「 t 期における、プレーヤ 2 の、プレ

プレイヤー 1 が戦略 i をとってくる確率の予想」を θ_{1it} と表し、 ϵ は定数とすれば

$$\theta_{1it+1} = \theta_{1it} + \epsilon(1_{(Player\ 1\ plays\ i)} - \theta_{1it}) \quad (2)$$

ただし、 $1_{(Player\ 1\ plays\ i)}$ は、プレイヤー 1 が i をすると 1 で、そうでないと 0 である。(Indicator Function)

Stochastic Fictitious Play において、ランダム要素は期待利得の値に導入されている。すなわち、上記のルールに基づいて相手の戦略について予想を立て、各戦略について通常通り期待利得を求めた後で各戦略の期待利得に確率変数を加算する。こうしてランダム性を加味したあとの利得が最も大きい戦略が選ばれようとする。数学的に記述しよう。 $\theta_{2t} = (\theta_{21t}, \theta_{22t}, \dots)$ とし、さらに α_{i1} で相手が戦略 1 を取った時、戦略 i のもたらす利得を表記して $\alpha_i = (\alpha_{i1}, \alpha_{i2}, \dots)$ とする。 e_{it}^1 は、プレイヤー 1 の戦略 i に加えられる独立同一分布に従う確率変数である。 t 期に、プレイヤー 1 が戦略 i を取るとき、戦略 i の期待利得が最大になっている、すなわち

$$\theta_{2t} \cdot \alpha_i + e_{it}^1 \geq \max_{j \neq i} \{\theta_{2t} \cdot \alpha_j + e_{jt}^1\} \quad (3)$$

が成立する。なお、以下のシミュレーションでは簡単のため確率変数は期待値 0、分散 σ の正規分布に従うと仮定されている。こうした過程をシミュレートし、最終的に落ち着いた均衡を、選択された均衡と考える。例えば、表 1 のゲームでは (D, D) が選択された。

5 Logit Dynamics モデル

この章では、Okada and Tercieux(2012) の Logit Dynamics モデルの概略を紹介する。このモデルでは Logit choice rule と呼ばれる理論がその根幹をなすので、まずはじめにそれを解説する。Logit choice rule は常に最適反応を取らず、どんな戦略も正の確率で取るという限定合理性の理論であるが、「より利得の高い戦略をより高い確率で選択」することをシステマティックに表現し、またパラメーター β を導入し、その値によって合理性の度合いを表現していることに特徴がある。簡単のため、2 人 N 戦略ゲームを想定する。プレイヤー i の戦略、戦略空間はそれぞれ s_i, S_i とおき、利得関数は $u_i(s_1, s_2)$ とおく。 $\beta \geq 0$ を合理性のパラメーターとする。このとき、Logit choice rule プレイヤー 1 は戦略 s_1' を取る確率を

$$p_i(s_1', s_2) = \frac{e^{\beta u_i(s_1', s_2)}}{\sum_{s_1'' \in S_1} e^{\beta u_i(s_1'', s_2)}} \quad (4)$$

と指定する。ここから、 $\beta = 0$ のときにランダムに、 $\beta \rightarrow \infty$ のとき、最適な戦略集合からランダムに戦略を確定することがわかるだろう。これに加えて「各プレイヤーが各期に行動変更の機会を与えられる確率」を与えれば、

繰り返されるゲームを logit choice rule の基づいてプレーする際の戦略プロフィールの遷移行列を書くことができる。そして、その定常状態を求めた後に $\beta \rightarrow \infty$ を考え、そこで残る均衡を Logit Dynamics で選択された均衡と考える。例えば、表 1 のゲームでは、(D, D) 均衡が選択される。

6 Young ゲームの均衡選択

この章では、表 2 に示す Young ゲームと呼ばれるゲームの均衡選択について、上に紹介した 4 つのモデルを比較する。

表 2: coordination game

| 1/2 | 0 | 1 | 2 |
|-----|------|------|------|
| 0 | 6, 6 | 0, 5 | 0, 0 |
| 1 | 5, 0 | 7, 7 | 5, 5 |
| 2 | 0, 0 | 5, 5 | 8, 8 |

純粋戦略ナッシュ均衡は (0, 0)、(1, 1)、(2, 2) の 3 つである。このうち、(2, 2) 均衡はパレート効率的で、かつリスク支配している。表 1 のゲームでは (D, D) がリスク支配均衡であり、上に紹介した 4 つのモデルはみな (D, D) を選択する。以下で Young ゲームにおいても、(2, 2) 均衡が選択されるのかどうか確認していく。

KMR モデルと Ellison モデルのシミュレーション結果は、筆者の github ページに ipynb ファイルとしてアップされている (https://github.com/haru110jp/StochEvolution/blob/master/kmr_ellison_young.ipynb)¹。KMR モデルは収束が非常に遅く、膨大な計算量を必要とするため完全なシミュレーションは行えなかったものの、(0, 0) 均衡や (2, 2) 均衡から (1, 1) 均衡への移動は容易なのに比べ、(1, 1) 均衡からの脱出は難しいことが見てとれる。実際、KMR においては理論的に定常分布はみな戦略 1 を取る状態であることが示されている²。KMR モデルは (1, 1) 均衡を選択するのである。

その一方で、Ellison モデルでは少ない試行数でも収束していることが見て取れる。しかしより大切なことは、収束先、すなわち選択される均衡が (2, 2) となっていることである。これは、初期戦略を 0 や 2 のみにしても変わらない結果である。

続いて、Stochastic Fictitious Play モデルの均衡選択を考える。図 2 は、 $\sigma = 1$ 、 $\epsilon = 0.2$ とし、1000 回対戦を行った時に選ばれた戦略の頻度である (2 人ゲームなので、計 2000 の観察がある)。初期信念は、両者とも相手が

¹ 「nbviewer」というサイトにこの URL を貼り付けることで、閲覧できる。

² 定常分布を理論的に計算するプログラムも、<https://github.com/haru110jp/StochEvolution> で公開されている。

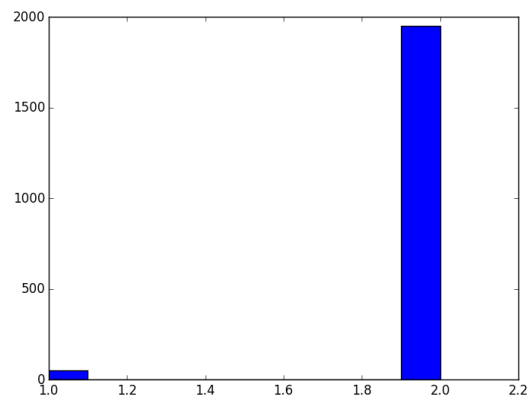


図 2: Stochastic Fictitious Play: Young ゲーム

必ず 0 を取る、というものである。明らかに、戦略 2 に収束している、すなわち均衡 $(2, 2)$ が選択されていることが読み取れる。

最後に、Logit Dynamics をシミュレートする。図 3 に示すのは、 $\beta = 8$ とし、2 人のプレイヤーには每期等確率で選択変更の機会が与えられる設定でのシミュレーション結果である³。初期行動は、2 人とも戦略 0 とした。

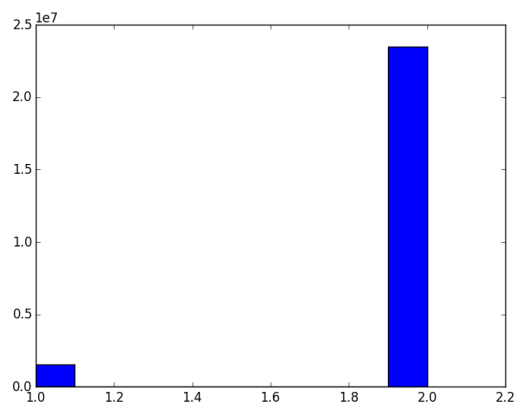


図 3: Logit Dynamics: Young ゲーム

³結果を見やすくするために 1000000 回の試行のうち最初の 500000 回のデータは除外した。

7 bilingual game を用いた、行動伝播の分析

この章では、Oyama and Takahashi(2015) に倣い bilingual game を用いてネットワーク上の行動伝播を分析する。前述したとおり、KMR や Ellison モデルは Morris(2000) や Oyama and Takahashi(2015) のネットワーク上での行動伝播モデルに突然変異を追加したものと解釈することができるため、両者の関連は深いのである。なお、この章の分析は始まったばかりであり、体系的なものでないことをお断りしておく⁴。

bilingual game の利得表を提示しよう。ただし、 $c > 0$ は定数である。

表 3: bilingual game

| 1/2 | A | AB | B |
|-----|----------|------------|------------|
| A | 11, 11 | 11, 11-c | 0, 3 |
| AB | 11-c, 11 | 11-c, 11-c | 10-c, 10-c |
| B | 3, 0 | 10, 10-c | 10, 10 |

なぜこのゲームが bilingual game と呼ばれるかについて簡単な説明を加える。いま、言語 A と言語 B が存在しており、プレーヤーはどちらの言語を使うか選択する。互いが同じ言語を使っていると、コミュニケーションが取れて利得が高くなる構造になっている。さらに、「AB」というバイリンガルになるという選択肢もある。しかし、この場合は2言語を使いこなさねばならないことで様々なコストが発生する。これが、パラメーター c で表現されている。

Oyama and Takahashi(2015) は”contagious”と”uninvadable”という2つの概念を定義した。ラフな形で説明すると、前者は、なんらかのネットワーク上において一部のプレーヤーがある戦略 a を取った時、プレーヤーが最適反応を繰り返していくと皆が a を取るようになることがありうる場合、 a は contagious であるという。後者は、どんなネットワーク上でも、一部の人間が他の戦略 a' を取り始めても最適反応を繰り返した結果それが集団全体には広がらないような戦略 a のことを指す。ただし、ここではネットワークの大きさが無限であると想定されているため、必ずしもそのまま有限のシミュレーションにあてはまるとは限らないことに注意する。

無限の大きさのネットワークでの理論という留保条件はつくものの、Oyama and Takahashi(2015) の Theorem 1 によると、 $c < \frac{40}{19}$ のとき、A は contagious であり $c < \frac{5}{3}$ なら uninvadable である。B については、 $c > \frac{5}{3}$ なら contagious であり、 $c > \frac{40}{19}$ なら uninvadable である。よって、これをベンチマークに c の値を操作してモデルの挙動を分析することにした。なお現時点でランダム

⁴分析の結果は、随時私の github に追加していく <https://github.com/haru110jp/StochEvolution/>

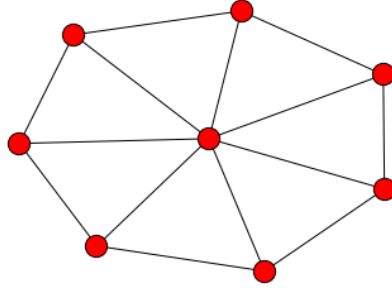


図 4: ノード数 8 の wheel graph

ネスを加えると分析の結果がどの要素に起因するかわかりにくくなるため、以下ではひとまず $\epsilon = 0$ とした。

2 つ興味深い現象を観察した。一つには、Ellison モデルのような円形ネットワークが必ずしもよく行動伝播をもたらすとは限らないということである。ノード数 8 の complete graph と cycle graph を作成し、 $c = 3$ と設定して行動 B のみが contagious になるようにした。初期時点では 3 名が行動 B、5 名が行動 A を取るとして一万回逐次最適反応を取る、という過程を 500 回繰り返した。すると、complete graph では 14 回目で行動 B が伝播したのに対し、cycle graph では一度も行動 B が伝播しなかった。

もう一つには、wheel graph の分析がある。wheel graph とは、円形グラフに「中心」を加えたものである。中心のノードは、円周上に並ぶ他のどのノードとも接続されている。例えば、中央のノードを中央政府と捉え、円周上のノードを各民族と考えれば、原始的な帝国における言語選択の問題を考察することができる。

まず、図 4 のようなノード数 8 (中央にひとつ、周辺に 7 つ) の wheel graph を作成した。パラメーター設定、分析法は先ほどと同様である。中央のノードだけが戦略 B、ほかは A を取るとき、9 回目で行動 B の伝播に成功した。行動改訂の順番に依存するものの、中央政府の戦略は言語選択に大きな影響をもつことがわかった。

次に、周辺民族のいくつが行動 B を選んでいれば、たとえ中央政府がはじめ行動 A を選んでいても行動 B が全体に伝播するか考える。つまり、侵略者が現地の言語を使い始める可能性のある critical mass を考えるのである。その結果、初期時点での行動 B 採用者が 2 名の時、500 回繰り返しても行動伝播が発生しなかったのに対し、3 名の時には 5 回目で行動伝播が発生した。このことは、周辺民族の言語 B が伝播しうる critical mass はこの場合 3 であったことを示している。この critical mass についてはさらなる網羅的な分析が

期待される。

8 結論と今後の展望

本論文では、4つの確率モデル、すなわち KMR モデル、Ellison モデル、Stochastic Fictitious Play モデル、Logit Dynamics モデルの挙動をシミュレーションによって比較検討した。Young ゲームを用いた均衡選択理論としてのモデル比較の結果、KMR モデルのみ他と違う結論に至った。さらに、KMR モデルと Ellison モデルはその構造上の類似にもかかわらず収束速度が大きく異なった。また、KMR モデルと Ellison モデルをネットワーク上でのゲームと考えると行動伝播の分析を行った結果、円形ネットワークが必ずしも行動を伝播させる性能が高いわけではないことが判明した。続いて wheel graph を用いた言語選択のシミュレーションでは、被支配者の言語が帝国の共通語となる条件についての考察の端緒となる結果が得られた。

今後の発展であるが、まずひとつには Young ゲーム以外の bimatrix game での均衡選択を検討することが挙げられる。次に、bilingual game のパラメーター c を変化させることで、どのような場合に行動伝播が発生し、どのような時に発生しないかをより網羅的に検討することが挙げられる。特に、Morris(2000) は coordination game を用いて行動伝播の特徴付けを行う一方で Oyama and Takahashi(2015) は bilingual game を用いて特徴付けを行っている。どのようなネットワークがどのような条件の下行動伝播を生じさせるかについて、両者の予測の違いを浮き彫りにするネットワークがどのようなものか特徴付けすることも重要であろう。

参考文献

- [1] 川越敏司 (2011) 「行動ゲーム理論入門」, 第二刷, NTT 出版
- [2] グレーヴァ香子 (2013) 「非協力ゲーム理論」, 第二刷, 知泉書院
- [3] Ellison, G(1993): "Learning, Local Interaction, and Communication," *Econometrica*, 61, 1047-1071
- [4] Kandori, M, Mailath, G, J, and Rob, R(1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 61, 29-56
- [5] Morris, S(2000): "Contagion," *Review of Economic Studies*; 67, 57-78
- [6] Okada, D, and Tercieux, O(2012): "Log-linear Dynamics and Local Potential," *Journal of Economic Theory*; 147, 1140-1164

- [7] Oyama, D and Takahashi, S(2015): "Contagion and Uninvadability in Local Interaction Games: The Bilingual Game and General Supermodular Games," Forthcoming in *Journal of Economic Theory*
- [8] Williams, N(2014): "Long Run Equilibrium in Discounted Stochastic Fictitious Play", *mimeo*