

Kandori, M., Mailath, G.J., Rob, R.,
Learning, Mutation, and Long Run
Equilibria in Games,
Econometrica, Vol.61, pp. 29-56,
1993,

生活研M2
齊藤いつみ

0. 目次

1. はじめに
2. モデルに関するラフな話
3. モデル
4. 極限分布の特徴
5. 調整ゲーム
6. 期待待ち時間
7. モデルの頑健性

1. はじめに

- 進化ゲーム, ゲームにおける学習とは
 - 各プレイヤーが自らの利得を上げる方向に行動を調整する動学的過程
 - 一定条件下でゲームにおける安定な状態 (ナッシュ均衡) に収束

しかし,

- 通常ゲームには**局所安定な均衡が複数存在**
- 伝統的な議論では, そのどれが実現しやすいかが明らかでなかった

1. はじめに

□ 確率進化ゲーム

- 確率的なゆらぎをモデルに明示的に組み込む
- 一つの均衡から別の均衡に移動する過程を**確率過程の定常分布**としてとらえる
- 長期的にみてどの均衡が起こりやすいか
- ゆらぎに対して長期的に頑強な均衡
- = **長期的確率安定的**

2. モデルに関するラフな話

- 今、ある寮で10人の学生が2つのコンピュータシステム (S1, S2) を選択するとする。
- 学生はランダムにマッチングし、同じシステムを持つ2人の学生が対面したときにはいつでも、「論文を書く」「ソフトウェアを交換」することができる。

II

	S1	S2	E1=(S1,S1)
I	S1	2,2	E2=(S2,S2)
	S2	0,0	

このゲームには、**純戦略均衡**として以下の2つの均衡が存在する (パレート効率はE1)

2. モデルに関するラフな話

		II	
		S1	S2
I	S1	2,2	0,0
	S2	0,0	1,1

また、**混合戦略**を考慮すると3つめの均衡状態が存在
→1/3がS1を選択し、2/3がS2を選択する

□ **純戦略**とは、

→プレイヤーが必ずどれかの戦略を選ぶゲームのこと

□ **混合戦略**とは

→プレイヤーが行動を確率的に選ぶような戦略をとること
でナッシュ均衡に達する非協力ゲームのこと

2. 純戦略と混合戦略

混合戦略の例

		B	
		q	1-q
A	p 1-p	S1	S2
		S1	2,2 0,0
		S2	0,0 1,1

プレイヤーAの期待効用

戦略S1: $2 \cdot q + 0 \cdot (1-q) = 2q$

戦略S2: $0 \cdot q + 1 \cdot (1-q) = 1-q$

プレイヤーBの期待効用

戦略S1: $2 \cdot p + 0 \cdot (1-p) = 2p$

戦略S2: $0 \cdot p + 1 \cdot (1-p) = 1-p$

プレイヤーは、期待効用の高くなる戦略を選択するので
 $p=q > 1/3$ のときS1, $p=q < 1/3$ のときS2を選択する

2. モデルに関するラフな話

- 最終的な均衡状態は**初期分布に大きく依存する**
- **少なくとも4人**の学生がS1を利用していれば、最終的に全ての学生がS1を利用するようになりパレート均衡**E1**が達成される。
- 上記以外の場合は、最終的な状態はE2となる。

しかし、

- 初期分布に完全に依存するのでは、長期的性質としてどちらがより安定かは議論しづらい
- できるだけ初期分布に依存しない性質を考えたい

2. モデルに関するラフな話

仮定：学生は、定期的にコンピュータを変更する機会を持つ

少しのランダム性を導入することで上記の問題を解消する

- 毎朝、微小な確率 ϵ で学生が出ていき、新しい学生が入ってくるとする



戦略分布の長期的過程は、一つにはとどまらず、初期分布にも依存しなくなる。

2. モデルに関するラフな話

□ システムとしては、よい方の均衡で長時間とどまる

このことを、2状態(E1,E2)のマルコフチェーンで考えてみる。

- 突然変異の結果、それぞれの状態に遷移する確率は0でない正の確率となる。
- E1→E2の遷移確率を p 、E2→E1の遷移確率を p' とする
- 前章の議論から、 p と p' のオーダーは ε^7 と ε^4 となる。

 $\varepsilon \rightarrow 0$ のとき、 $p/p' \rightarrow 0$ となる。 ($\varepsilon^4 / \varepsilon^7 \rightarrow 0$)

単純な計算から、E1とE2の定常分布はそれぞれ

$(p'/(p+p') \cdot p/(p+p'))$ となる。よって、

$\varepsilon \rightarrow 0$ のとき、 $(p'/(p+p') \cdot p/(p+p')) \rightarrow (1,0)$ となる。

3. モデル

状況設定

- N人の有限プレイヤーが繰り返しゲームを行い、時間がたつにつれ行動を適応させていくとする
- 以下のような 2×2 の対称ゲームを考える。
- 離散時間 $t=1, 2, \dots$ の各期間 t の初めに純粋戦略をとるとする
(将来効用は考慮しない)

II

		S1	S2
I	S1	a, a	b, c
	S2	c, b	d, d

戦略S1をとる人数が z 人のとき、それぞれの戦略の平均利得は以下ようになる

$$\pi_1(z) = \frac{(z-1)}{(N-1)}a + \frac{(N-z)}{(N-1)}b, \quad \pi_2(z) = \frac{z}{(N-1)}c + \frac{(N-z-1)}{(N-1)}d.$$

3. モデル

□ ダイナミクスの仮定 (確定的)

$$z_{t+1} = b(z_t)$$

$$\text{sign}(b(z) - z) = \text{sign}(\pi_1(z) - \pi_2(z)) \quad \text{for } z \neq 0, N$$

最適反応ダイナミクス

$$B(z) = \begin{cases} N, & \text{if } \pi_1(z) > \pi_2(z), \\ z, & \text{if } \pi_1(z) = \pi_2(z), \\ 0, & \text{if } \pi_1(z) < \pi_2(z). \end{cases}$$

調整ゲームに関しては, 上記の動的システムであれば
2つか3つの均衡点が存在 (0,N,1/3)

3. モデル

- ダイナミクスの仮定（確率的）
- ランダムなショックを考慮する。

$$z_{t+1} = b(z_t) + x_t - y_t$$

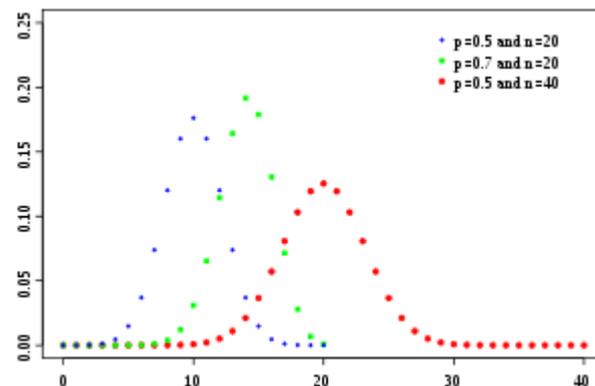
ここで, x_t, y_t は以下のような2項分布を仮定する。

$$x_t \sim \text{Bin}(N - b(z_t), \varepsilon) \quad \text{and} \quad y_t \sim \text{Bin}(b(z_t), \varepsilon)$$

二項分布

確率変数 X が二項分布 $X \sim \text{Bin}(n, p)$ に従うとき確率変数 X は以下の式を満たす

$$P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}$$



3. モデル

この動的システムは、有限状態空間におけるマルコフチェインとして定義され、遷移確率は以下のように表される。

$$p_{ij} = \text{Prob}(z_{t+1} = j | z_t = i)$$

マルコフチェインは一意的な定常状態を持つ。N次元の定常状態分布を μ とすると、以下の式が成り立つ

$$\mu = (\mu_0, \mu_1, \dots, \mu_N) \in \Delta_N$$

$$\mu P = \mu. \quad \rightarrow \quad \text{遷移確率をかけても状態が変化しない}$$

定義：極限分布 μ^* は次のように定義される

$$\mu^* = \lim_{\varepsilon \rightarrow 0} \mu(\varepsilon), \quad \text{if it exists.}$$

4. 極限分布の特徴

□ 遷移樹形図分析

- マルコフ過程の確率を書き下すよりも簡便な方法で定常分布（長期的確率安定状態）を求める方法（Freidlin and Wentzell(1984)）



□ 遷移樹形図定理：

長期的安定状態は、最小のコストを持つ樹形図の根となる。逆に、最小の遷移コストを持つ樹形図の根は、長期的確率安定状態である。

4. 極限分布の特徴

遷移コスト：

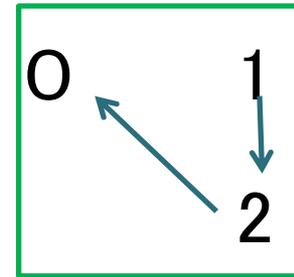
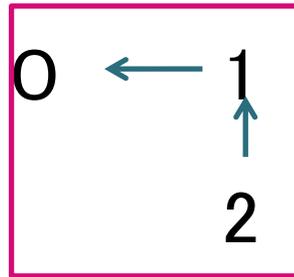
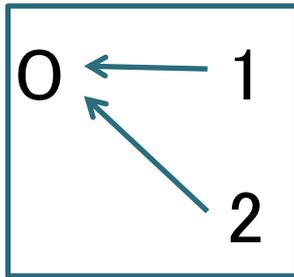
ある均衡から別の均衡に移る際に最低限必要なゆらぎ

最小の遷移コストを持つ樹形図の根は、ほかのさまざまな状態から最も少ない数の揺らぎで到達できるため、わずかな確率で揺らぎが起きるとき、長期的に見て最もおこりやすい均衡状態となる。

4. 極限分布の特徴

具体例

状態空間 Z が $Z=\{0,1,2\}$ の場合(均衡状態が3つある場合)
状態の遷移について, 全ての場合を書き下す



$$q_z = \sum_{h \in H_z} \prod_{(i \rightarrow j) \in h} p_{ij} \quad q_0 = p_{20} p_{10} + p_{21} p_{10} + p_{12} p_{20}$$

定理: ベクトル q は, 定常分布 μ に比例している

$$\mu_z^* = \lim_{\varepsilon \rightarrow 0} q_z(\varepsilon) / \sum_i q_i(\varepsilon)$$

5. 調整ゲーム

□再度調整ゲームについて考える

ある均衡状態に移行するための人口の限界レベルを z^* とすると,

$$\text{sign}(\pi_1(z) - \pi_2(z)) = \text{sign}(z - z^*).$$

$$z^* = [N(d - b) + a - d] / (a - c + d - b).$$

z^* は混合戦略均衡点と等しい

均衡状態は $\{0, N\}$ だが, それぞれの状態は誘導領域を持つ

誘導領域 (basin of attraction)

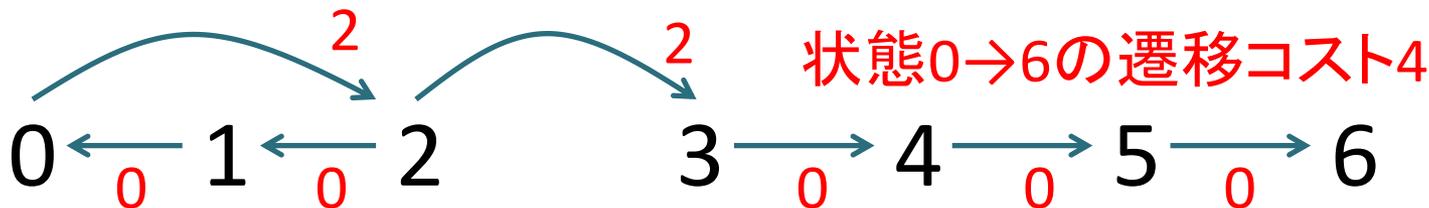
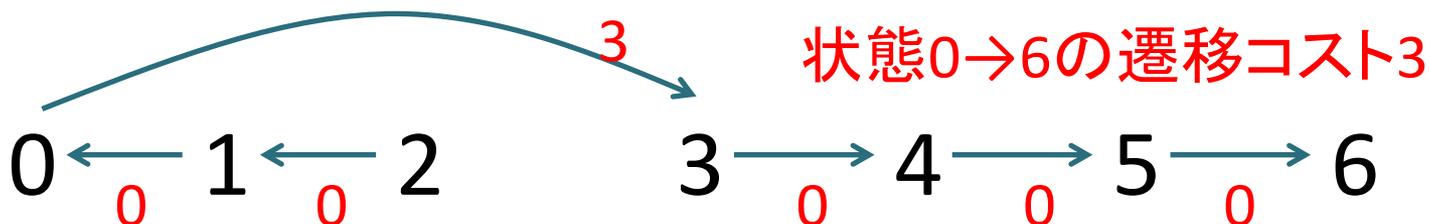
誘導領域とは, 均衡周辺の状態集合でそこに入ったら均衡へ向かうものを集めた領域

相対的に大きい誘導領域をもつ状態が長期的に安定な均衡状態となる.

5. 調整ゲーム

具体例

z^* は2, 3の間にある. 均衡状態は $\{0,6\}$ の2つ. それぞれの状態の誘導領域は $0\{0,1,2\}, 6\{3,4,5,6\}$



状態0→6のコスト3, 状態6→0のコスト4

➡ 状態6が長期的な確率安定状態となる

5. 調整ゲーム

調整ゲームで $z^* \neq N/2$ かつ $N \geq 2(a-d)(a-c-d+b)$ のとき, 長期的な均衡はリスク支配になる

□ 利得支配とリスク支配

利得支配

複数ナッシュ均衡の中で, もっとも二人にとって利得が高い組み合わせ

リスク支配

複数ナッシュ均衡の中で, どちらかが戦略を逸脱した際のリスクがもっとも低い組み合わせ

リスク支配の例

□ Stag Hunt games

	Stag	Here
Stag	3,3	0,2
Here	2,0	2,2

利得支配は, (stag,stag)

リスク支配は, (Here,Here)

ある戦略から別の戦略に移行することによる損

利得支配 $\rightarrow (3-2) \times (3-2)=1$

リスク支配 $\rightarrow (2-0) \times (2-0)=4$

$$\text{Risk(Stag,Stag)} > \text{Risk(Here,Here)}$$

6. 期待待ち時間

ある均衡状態が観測されるまでに、どのくらいの時間がかかるのか？ 長期的均衡とは、実際にどの程度の長さなのか？

□ 数値計算例

	S1	S2
S1	a, a	0, 0
S2	0, 0	1, 1

左図のような利得構造の調整ゲームを考える

$a > 1$

6. 期待待ち時間

a = THE PAYOFF FOR E_1 , ε = MUTATION RATE, N = POPULATION SIZE,
 t_1 = EXPECTED DURATION OF $E_1 = 1/p$, t_2 = EXPECTED DURATION OF $E_2 = 1/p'$,
 μ_2 = PROBABILITY OF E_2 IN STATIONARY DISTRIBUTION.

a	ε	N	t_1	t_2	μ_2
2	0.1	10	1.1×10^5	78	7.1×10^{-4}
2	0.1	20	4.6×10^9	420	9.0×10^{-8}
2	0.1	40	3.1×10^{17}	54,000	1.7×10^{-13}
3	0.1	10	2.7×10^6	14	5.3×10^{-6}
3	0.1	20	3.1×10^{12}	89	2.9×10^{-11}
3	0.1	40	9.1×10^{22}	680	7.4×10^{-21}
4	0.1	100	5.4×10^{61}	1,200	2.3×10^{-59}

$$a = 2, \varepsilon = 0.1, N = 10, (\mu_1(t), \mu_2(t)) = (0, 1)P^t$$

t	1	20	40	60	80	100
$\mu_2(t)$	0.93	0.23	0.054	0.013	0.0030	0.00070

7. モデルの頑健性とまとめ

- ゲームの利得構造が決定したとき、ゲームの長期的な結果は適応仮定によらない。
- 確率的ショックを与えることで、確率的に安定な戦略を示すことができた